Franzettis Mathematik
Weg zur Maturprüfung

Teil VI
Stochastik

7. Juni 2023

Inhaltsverzeichnis

Teil VI

Stochastik

Normalprogramm / Erweitertes Niveau

– auf einfache Situationen die Begriffe Population, Bestand und relative Häufigkeit anwenden
– eine Verteilung anhand eines Kreis- oder Stabdiagrammes oder eines Histogrammes darstellen
– Masszahlen einer Verteilung (arithmetisches Mittel, Median, Modus, Varianz und Standardabweichung) definieren und interpretieren

Kapitel 30

Beschreibende Statistik

30.1 Grundlagen

Beschreibenden Statistik betreibt man heutzutage mit Software-Paketen. Das händische Rechnen, auch mit dem Taschenrechner, ist somit auf extrem kleine Datenmengen beschränkt. Somit legen wir hier viel mehr Wert auf das Verständnis als auf Rechenfertigkeiten.

30.1.1 Sinn und Zweck

Das Wort Statistik leitet sich vom Lateinischen ab, vom Verb *sto*[1] (stare) mit der Bedeutung stehen, aufgestellt sein, oder von verwandten *status* (pl. statu(u)s), Zustand, Lage. Daraus leitet sich auch der Begriff Staat und Stadt ab. Im Englischen ist *state* sowohl Staat als auch Zustand.
Die Statistik dient der Beschreibung eines Zustands, einer Lage. Die Mittel der Beschreibung sind

- die Tabelle,
- die Graphik oder
- Kennzahlen.

Die Beschreibung wiederum dient als Grundlage der Vorausschau, der Prognose und der Entscheidungsfindung. Die vertiefte Analyse der Daten zielt auch auf die Erkundung von Zusammenhängen. Es mag dümmlich klingen, aber wir kennen nur die Vergangenheit, wenn überhaupt, wir sind aber hauptsächlich an der Zukunft interessiert. Unser gegenwärtiges Handeln ist auf die Zukunft gerichtet, wird dereinst Folgen haben.
Die Statistik ist eng verwandt mit dem Zählen von grösseren Mengen wie Vieherden, Getreidevorräte, die wiederum mit der Sesshaftig-Werdung der Menschheit in Mesopotamien, dem Ackerbau, der Lagerhaltung von Lebensmitteln, der Arbeitsteilung und dem Dienst an der Allgemeinheit zusammenhängt. Die Organisation einer solchen Gesellschaft bedarf der Schrift für Wort und Zahl, der Schriftträger wie Tontafeln und der Buchhaltung. Statistik ist auch Buchhaltung des Staates. Die ersten Beispiele sind die Volkszählung (Weihnachtsgeschichte) und der Zensus, die Feststellung des Vermögens für die Steuern. Bis zur Aufklärung war die Statistik des Staates geheim.

[1]Für Puristen wird im Latein nicht der Infinitiv angegeben, sondern der Präsens der ersten Person singular, wie der alte Stowasser zeigt.

30.1.2 Sammeln und Ordnen

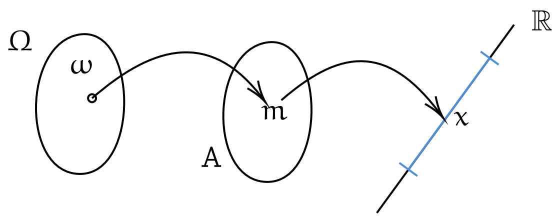
Die Datenmengen, die man untersuchen will, sind nicht einfach schon da. Man muss sie organisieren. Meist interessiert eine spezifische Fragestellung woraus dann die zu suchenden Daten bestimmt werden. Dies trifft u.a. zu für

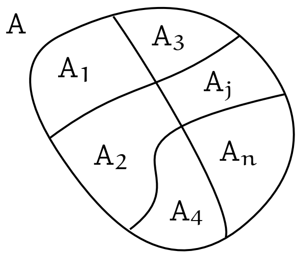
- wissenschaftliche Experimente,
- Beobachtungen in Natur (Niederschlagsmenge), in der Wirtschaft (Börsenkurse) beim Staat (Zoll),
- Befragung in den Sozialwissenschaften.

Die Daten sind Fakten, die man konstruieren muss.

Begriffe

Bei einer statistischen Auswertung muss man die Menge der möglichen Beobachtungsobjekte oder Experimentereignisse festlegen. Diese Menge nennt man *Grundgesamtheit* oder *Population* Ω. Ein Element dieser Menge nennt man *statistische Einheit* ω.





Was will man messen? *Merkmale* oder Eigenschaften der statistischen Einheiten. Die möglichen Ausprägungen dieser Merkmale legen den *Ausprägungsbereich* fest. Will ich also die Farbe vorbeifahrender Autos messen, so muss ich den Bereich akzeptabler Farben festlegen. Dieser könnten sein {schwarz, weiss, silbern-grau, bunt}. Wichtig ist, dass die Ausprägungen paarweise disjunkt sind, sie sich gegenseitig ausschliessen. Würden wir die Menge wie folgt festlegen {schwarz, weiss, silbern-grau, bunt, allrad}, so würde dies nicht zutreffen (siehe Abbildung).

30.1.3 Messen und Zählen

Die Ausprägungen können Merkmalswerte mit verschiedenen Eigenschaften aufweisen:

- nominal,
- ordinal,
- kardinal (intervallmässig und verhältnismässig)

sein. Nominale und ordinale werden auch als *qualitative* Skalen bezeichnet, wogegen Intervall- un Verhältnisskala *quantitative*, *kardinale* oder *metrische* Skalen sind.

Nominalskala

Nominal bedeutet einfach, dass sie einen Namen haben. Beispielsweise sind Farben wie rot, grün, gelb etc. oder Geschlecht weiblich, männlich und sonst, Religionszugehörigkeit aber auch Postleitzahl nominal. Die wichtigste Relation ist einfach, dass sie eindeutig sind. Diese Eigenschaften werden dann gezählt.

Es kann aus technischen Gründen wünschenswert sein, solche Eigenschaften zu kodieren, d.h. eine Zahl zuzuordnen. Also männlich=1, weiblich=2, sonst=3. Mit diesen Zahlen kann man allerdings nicht rechnen.

Ordinalskala

Schon im Wort steckt die Eigenschaft, dass man die Ausprägungen ordnen kann. Damit gilt die Ordnungsrelation "$>$". Beispielsweise ist der Intelligenzquotient ordinal skaliert. Eine Wert von 120 ist grösser als 100, und bedeutet, dass die statistische Einheit mit diesem Wert am Test besser abgeschlossen hat als die Einheit mit 100. Ob er "intelligenter" ist, kann man debattieren.

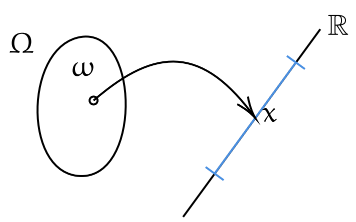
Intervallskala

Bei der Intervallskala ist zusätzlich zur Ordnungsrelatio auch der Abstand eine sinnvolle Grösse. Wenn wir die Temparatur in Grad Celsius betrachten, dann kann man sagen, dass 40° höher als 20° sind, und somit wärmer, und, dass der Temperaturunterschied zu 30° gleichgross ist. Was man aber nicht sagen kann, ist, dass 40° doppelt so warm sind als 20°. Das Verhältnis ist sinnlos.

Verhältnisskala

Zusätzlich zu Eindeutigkeit, Ordnung und Abstand wird bei dieser Skala auch das Verhältnis sinnvoll. Ein Einkommen von 150'000 Fr. ist drei Mal mehr als eines von 50'000 Fr.
Verhältnisgrössen können ganze oder reelle Zahlen sein. Die Anzahl Einwohner eines Dorfes ist eine ganze, sogar natürliche Zahl, das Körpergewicht der Gemeindepräsidenten sind reelle (positive) Zahlen wie 120.3456 kg.

30.1 Übung Wie steht es mit den Schulnoten? Zu welcher Skala gehören sie? Die möglichen Ausprägungen in der Schweiz sind $\{1,2,3,4,5,6\}$, wobei 4 gerade noch genügend ist. In Deutschland gilt die Skala "sehr gut"=1, "gut"=2, "befriedigend"=3, "ausreichend"=4, "mangelhaft"=5 und "ungenügend"=6. Das ist klar eine Ordinalskala mit einer Kodierung, Zahlzuordnung. Wie würde man ein deutsches Zeugnis auf Schweizer System abbilden? Obwohl es eindeutig ist, dass Schulnoten ordinal sind, wird auch per Gesetz davon ausgegangen, dass sie wie Verhältniszahlen zu verwenden sind. Denn eine Addition von ordinalskalierten Grössen ist unsinnig, so dass auch ein Durchschnitt unsinnig ist. Z.B. (männlich + weiblich)/2=? oder A hat eine 6, B eine 2, ergo ist A dreimal besser.
Zumindest im Kanton Zürich wird seit 1994 das System für Mittelschulen noch kurioser. Denn es werden negative Differenzen von 4 doppelt gewichtet zu positiven Differenzen über 4. Also $(4-3) \sim (6-4)$. Das ist die Krönung zur Förderung des Mittelmasses. ◁



Sind die Ausprägungen schon selber kardinale Grössen, also Verhältnis- oder Intervallgrössen, dann sieht das Messen einfacher aus gemäss der Abbildung. Dies trifft z.B. auf die monatliche Niderschlagsmenge oder den Temperaturdurchschnitt zu. Wenn allerdings sehr viele Datenpunkte erfasst werden, dann kann es nützlich sein, diese in eine übersichtliche Anzahl Intervalle einzuteilen. Man bilden z.B. Klassen für das Einkommen.

30.2 Häufigkeit

Dia Häufigkeit ist die natürlichste Art der Quantifizierung, denn jede Menge, hat eine Zahleigenschaft, ihre Mächtigkeit. Bei einer Untersuchung hat man die statistischen Einheiten definiert und ihre Ausprägungen festgestellt. Bei nominal und ordinal skalierten Merkmalen muss man nur noch zählen. Bei kardinalen Skalen muss man eventuell zuerst Klassen bilden, auf englisch *bins*, also Behälter oder Eimer, in die man die Merkmalsausprägungen rein tut. Eine Strichliste für beobachtete Autos sieht wie folgt aus

Farbe	Anzahl	
schwarz	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	35
weiss	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	30
silbern-grau	𝍸 𝍸 𝍸 𝍸 |||	23
bunt	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 |	46
Total		134

Mit der folgenden Setzung

Definition 1. Die *absolute Häufigkeit* $H_n(A)$ eines Ereignisses A ist die Anzahl dessen Eintretens bei n Versuchen. Die *relative Häufigkeit* $h_n(A)$ ist die absolute Häufigkeit im Verhältnis zur Gesamtanzahl Experimente n. Also

$$H_n(A) = n_A \qquad h_n(A) = \frac{n_A}{n}$$

können wir die Beobachtung beschreiben als

	Ausprägung				
	schwarz	weiss	silbern-grau	bunt	Total
Absolute Häufigkeit	35	30	23	46	134
Relative Häufigkeit	0.261	0.224	0.172	0.343	1.000

Die relative Häufigkeit kann man natürlich auch als Prozentzahlen angeben, also 26.1%, 22.4% etc.

30.2.1 Häufigkeitsfunktion

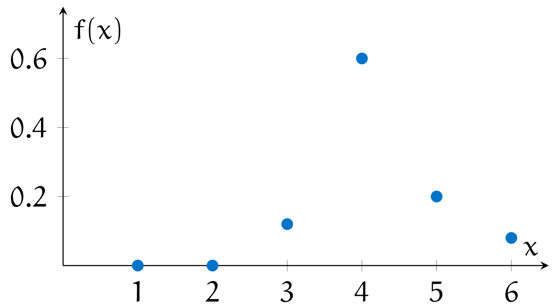
Als Grundlage für die folgenden Theorie gehen wir von den Noten einer Schulklasse mit 25 Schülern aus und nehmen an, Noten seien kardinale Grössen.

Note x_j	Anzahl $n(x_j)$	Häufigkeit $h_n(x_j)$
1	0	0.00
2	0	0.00
3	3	0.12
4	15	0.60
5	5	0.20
6	2	0.08

Definition 2. Die Häufigkeitsfunktion der Stichprobe für kardinale Daten ist

$$f(x) = \begin{cases} h_n(x_j) = \dfrac{n(x_j)}{n} & \text{falls} \quad x = x_j, \quad j = 1, 2, \ldots, k \\ 0 & \text{sonst} \end{cases}$$

Die Graphik zeigt die diskreten Punkte der Häufigkeitsfunktion $h(x)$. Hier werden nur ganze Noten verteilt. Diese Darstellung ist eine Vorbereitung für die sogenannte Verteilung. Dort fragen wir nach der kumulierten Häufigkeit, dass die Noten schlechter sind als ein bestimmter Wert.



30.2.2 Verteilungsfunktion, Summenhäufigkeit

Die relative Häufigkeit beschreibt die Anteile. Nun fragen wir mit der Summenhäufigkeit nach dem Anteil der Messwerte, die *kleiner oder gleich* einem Ausprägungswert sind.
Die Summenhäufigkeit wird auch *Verteilungsfunktion* genannt. Man kann sie wie folgt festlegen.

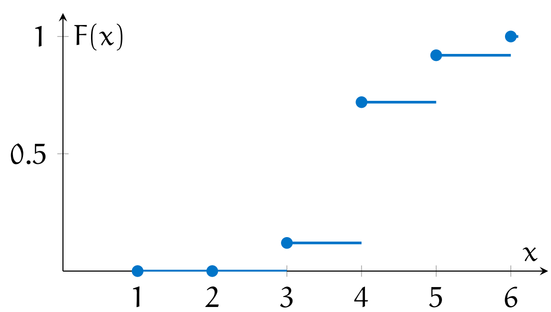
Definition 3. Die *Summenhäufigkeit* oder *Verteilungsfunktion* $F(x)$ ist

$$F(x) = \sum_{x_i \leqslant x} h(x_i)$$

Anmerkung 30.1. Die Funktion $F(x)$ ist jetzt für alle x festgelegt, nicht nur für die diskreten Punkte x_i.

Die Summenhäufigkeiten ergeben sich aus der Summe der relativen Häufigkeiten.

Note x	1	2	3	4	5	6
Summenhäufigkeit $F(x)$	0.00	0.00	0.12	0.72	0.92	1.00



Die Verteilungsfunktion links zeigt die Summenhäufigkeiten. Für $x = 4.25$ folgt der Wert $F(4.5) = 0.72$. Der Punkt in der Abbildung besagt, dass dieser Wert dazugehört, das Ende der Striche rechts aber nicht. Die Note 6 ist das Maximum. Der Wertebereich einer Verteilungsfunktion ist immer $[0, 1]$.

Folgende Eigenschaften der Verteilungsfunktion herrschen:

- Die Funktion ist reellwertig,
- sie ist eine Treppenfunktion, rechst der Ausprägung verläuft sie auf konstantem Niveau,
- die Verteilungsfunktion ist monoton steigend, d.h. $F(x + d) \geqslant F(x)$,
- ihr Wertebereiche ist das Intervall $[0, 1]$.

30.3 Graphische Darstellung

Der optische Eindruck einer Graphik kann und soll schnell einen ersten Überblick über potentiell sehr viele Datenpunkte gewähren. Eine geschickte Darstellung dient auch als wertvolle Ausgangsbasis für Analysen.

30.3.1 Durchgehendes Beispiel

Wir haben ein Beispiel gewählt, dass einen aktuellen Bezug hat, nämlich die Klimaerwärmung. Was erkennt man aus historischen metereologischen Daten?
Im folgenden verwenden wir schon erhobene Daten, genauer Messungen der Metereologischen Anstalt zu Temperatur und Niederschlagsmenge im Monatsdurchschnitt aus den Jahren 1864-2021 in Zürich-Fluntern[2]. Die Tabelle 30.1 auf Seite 30-27 zeigt einen Auszug, nämlich nur den Monat Februar. Die Nidergschlagsmenge kann man in Hhöhe in mm angeben oder in Liter pro Quadratmeter. Die Koeffizienten sind gleich, denn $1[mm]$ entspricht $1[l/m^2]$. Wir haben den Februar gewählt, weil er die grösste Streuung aufweist, also variabler ist als andere Monate. Der Datensatz bietet die Möglichkeit, Temperatrur und Niderschlag einzeln zu betrachten, aber auch zusammen. Damit ist etwa die Fragestellung möglich: "Gibt es einen Zusammenhang zwischen Temperatur und Niederschlag?"

Year	Month	Temperature	Precipitation
1864	1	-6.6	25.7
1864	2	-1.5	32.9
1864	3	4.5	51.0
1864	4	6.8	46.9
...	...	...	...
1864	12	-4.2	4.0
1865	1	0.0	63.7
1865	2	$\boxed{-2.9}$	52.8
1865	3	-1.2	51.9
...	...	...	...

Die Beobachtungen sind kardinal oder metrisch, d.h. ihre Skala ist für die Temperatur eine Intervall- und für die Niederschläge eine Verhältnisskala. Für Häufigkeiten muss man also die Zahlen zuerst in Klassen einteilen und dann die Mächtigkeit der Klasse bestimmen. Wie man Klassen bildet, sehen wir später.
Der Datensatz enthält 153 Messpunkte pro Monat. Dies ist zu viel, um von Hand zu rechnen. Die gezeigten Analysen und Bilder wurden mit einem Programm erstellt. Deshalb wird hier der Fokus mehr auf das Verständnis als auf die Produktion gelegt.
Es ist ein Merkmal der beschreibenden Statistik, dass man eine grosse Zahl von Daten sinnvoll kondensieren kann, also von Daten zu Informationen zu verarbeiten.

30.3.2 Stamm-Blatt-Diagramm

Die einfachste graphische Darstellung ist das Stamm-Blatt-Diagramm, bei dem die vorhandenen Zahlen eventuell gerundet, aber sonst nur geordnet werden. Diese Graphik konnte man schon mit einer Schreibmaschine darstellen.

[2] https://www.meteoschweiz.admin.ch/product/output/climate-data/homogenous-monthly-data-processing/data/homog_mo_SMA.txt

Mittlere Temperaturen Monat Februar °C

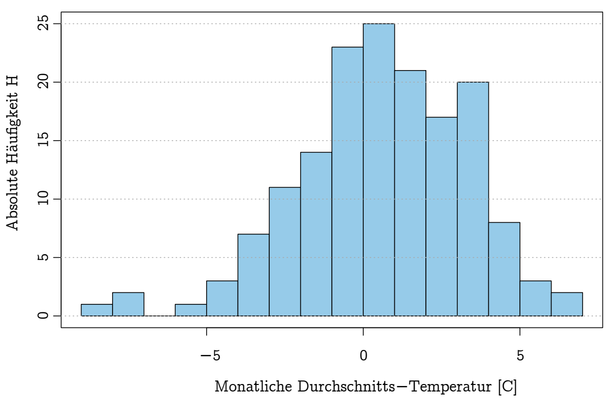
```
-8 | 9
-7 | 74
-6 |
-5 | 2
-4 | 831
-3 | 8765310
-2 | 98875442110          (-2.9 ist hier)
-1 | 96665553221110
-0 | 9976666555555544332111
 0 | 001222333334445566678899999
 1 | 1222344445666777788999
 2 | 1112233466777
 3 | 0000122223456667778999
 4 | 0001367789
 5 | 0156
 6 | 25
```

Niederschlagsmenge Monat Februar [×100 mm]

```
 0 | 34678800002224455557889999
 2 | 002334445556779991222234556777788 9
 4 | 0011223345778899922333456667778899
 6 | 00233345556778900000001125677 7
 8 | 44568899902389
10 | 3312799
12 | 58836
14 | 383
16 | 196
18 | 2
20 |
22 |
24 |
26 | 6
```

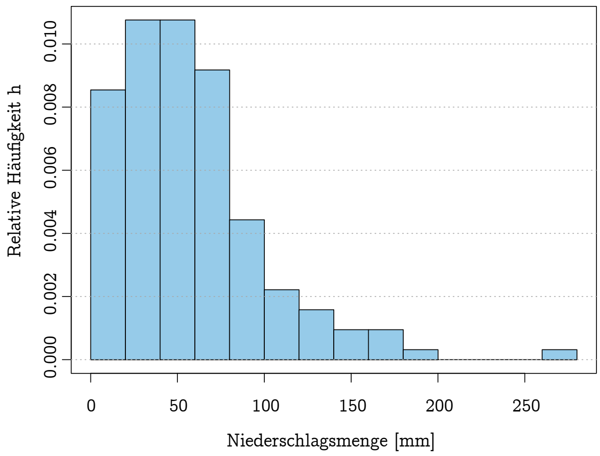
Als erstes muss man die Zahlen der Gröss nach ordnen, von klein zu gross. Die erst oder zwei ersten Ziffern legen die Klassen fest. Die Klassenbreite ist also $x.9\overline{9} - x.0 = 1$. Nach dem vertikalen Strich folgen jeweils die erste Ziffer nach dem Dezimalpunkt. Diese Zahlenaufstellung vermittelt einen visuellen Eindruck der Zahlen.
Nun könnte man wie bei den Autos die Häufigkeit bestimmen, das ist, die Anzahl Ziffern pro Reihe zählen und auch durch 153, der Anzahl Messungen, teilen.
Der Vergleich von Temperatur und Niederschlag zeigt, dass die Temperatur etwas symmetrischer um 0 herum verteilt ist, wogegen der Niederschlag viel einseitiger verteil ist. Das hat auch damit zu tun, dass es keine negativen Niederschlagsmengen gibt.

30.3.3 Histogramm



Wir zählen nun die Ausprägungen des obigen Stamm-Blatt-Diagramms und erhalten das nebenstehende Histogramm. Wie man ziemlich gut sieht, entspricht es dem Stamm-Blatt-Diagramm, wenn man diese um 90 Grad dreht. Beim Histogramm sind die Abschnitte auf der x-Achse bedeutungsvoll, denn sie stellen Intervalle für die kardinalskalierten Ergebnisse dar. In der Abbildung sind die Klassenbreiten gleichgross oder aequidistant. Dies muss aber nicht so sein.
Die Summer aller Häufigkeiten der Klassen muss 153, die Anzahl Messungen ergeben.

Für die Niederschläge wählen wir die relative Häufigkeit als Höhe der Säulen, ebenfalls unter Verwendung aequidistanter Klassen. Nun muss die Summe der Säulenhöhen 100% oder 1 ergeben. Denn sonst hat man nicht alle Werte berücksichtigt. Wiederum entspricht das Bild dem Stamm-Blatt-Diagramm, das um 90 Grad gedreht wird.
Sind die Klassenbreiten nicht konstant, dann muss man die Rechtecksfläche als die entsprechende Häufigkeit betrachten.



Klasseneinteilung

Wir besprechen nur die identischen Klassenbreiten. Hier gibt es einige Fausregeln. Denn, was möchte man erreichen? Je weniger Klasssen desto mehr Information geht verloren, aber, je mehr Klassen desto mehr treten Ausreisser und Zufallsschwankungen hervor. Mit der Klassenbildung gehen die Werte verloren, denn es wird angenommen, die Klassenmitte sei der Wert für alle Mitglieder der Klasse. Zwischen der untersten und der obersten Klassengrenze müssen alle Messungen liegen. Deshalb sollte die Klassen zusammen etwas mehr überdecken als den Messbereich. Zudem muss festgelegt werden, zu welcher Klasse Messungen auf den Grenzen gehören. Z.B. ist das Intervall $(e_{i-1}, e_i]$ oder $[e_{i-1}, e_i)$ mit e_k Klassengrenze. Die einfachste Fausregel ist bei n Werten

$$5 \leqslant k = \sqrt{n} \leqslant 25$$
$$5 \leqslant k \approx 5 \cdot \lg(n) \leqslant 25$$
$$5 \leqslant k \approx 1 + 3.3 \cdot \lg(n) \leqslant 25$$

Das ergibt für unser $n = 153$ ca. 12. Die Programme und Tabellenkalkulatoren (Spread Sheets) haben solche Graphikfunktionen eingebaut, die selber Klassen bilden. Wenn man nicht zufrieden ist, kann man die Vorgaben übersteuern.

30.1 Übung Ein Betrieb beschäftige Arbeiter zu unterschiedlichen Stundenlöhnen. Die Übersicht ist wie folgt:

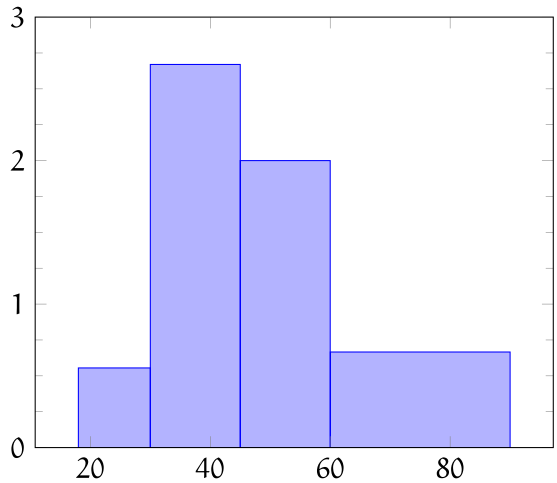
Stundenlohn x	Anzahl der Beschäftigten
$18 \leqslant x < 30$	20
$30 \leqslant x < 45$	80
$45 \leqslant x < 60$	60
$60 \leqslant x \leqslant 90$	40

Wir zeichnen eine Histogramm der Lohnkosten. Wir bemerken, dass die Klassenbreiten unterschiedlich sind.

Die relativen Häufigkeiten sind 10, 40, 30 und 20 %. Die Klassenbreiten sind 12, 15, 15 und 30. Die Balkenhöhen ergeben sich aus

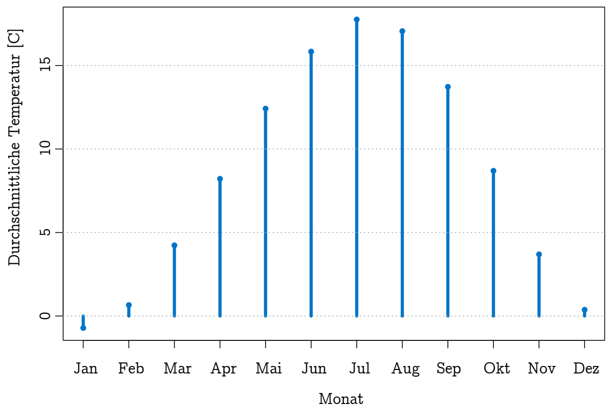
$$\text{Höhe} = \frac{\text{Fläche}}{\text{Breite}} = \frac{\text{Anteil}}{\text{Klassenbreite}}.$$

Damit folgen diese zu 10/12, 40/15, 30/15 und 20/30 in Prozent. Beim nicht aequidistanten Histogramm muss man auf die Flächen achten. ◁



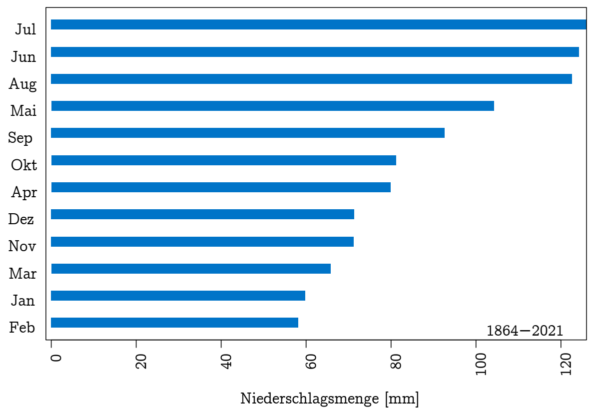
30.3.4 Säulen- oder Stabdiagramm

Wir betrachten die Säulendiagramme, die eine grosse Ähnlichkeit mit den Histogrammen aufweisen. Ein wichtiger Unterschied ist die Bedeutung der Werte der x-Achse.



In der Abbildung sieht man die durchschnittlichen Monatstemperaturen (immer aus denselben Ur-Daten). Wie bei den Histogrammen gibt es senkrecht stehende Säulen, deren Höhe Werten der y-Achse entsprechen. Die x-Achse besteht aus nominalen Ausprägungen. Hier sind sie im Jahresverlauf dargestellt. Inhaltlich ist klar, dass der Januar der kälteste und Juli der wärmste Monat sind.

Man kann die Säulen auch zu Balken machen, die dann horizontal zu liegen kommen. Hier zeigen wir die Niederschlagsmenge nach Monaten der Menge nach sortiert.

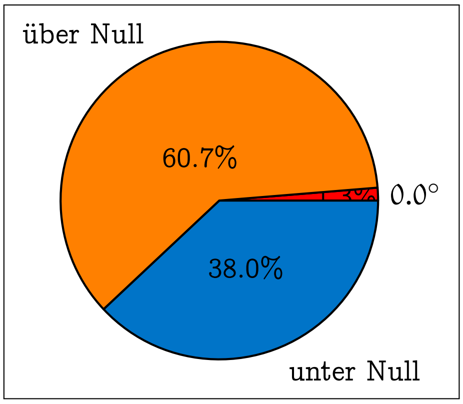


Was man hier schon vermuten kann, ist, dass die Niederschlagsmenge mit der Höhe der Temperatur zusammenhängt. Denn die wärmsten Monate sind auch die regenreichsten. Hier wird allerdings die Menge gezeigt und nicht die Anzahl Tage, an denen es regnet. Die Dicke der Balken hat keine Bedeutung.

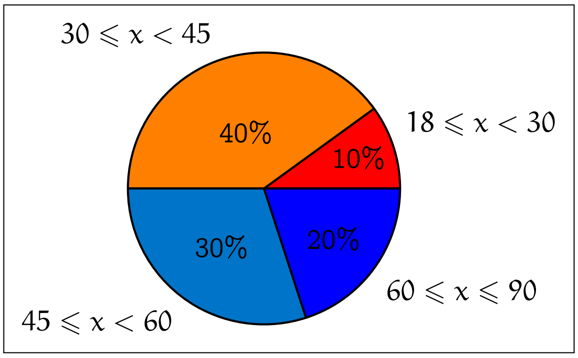
Und nun zu einem weiteren Klassiker.

30.3.5 Kreisdiagramm

Das Kreisdiagramm, aus dem Englischen "Torten-Diagramm", zeigt die Anteile der verschiedenen Ausprägungen. In unseren Temperatur-Daten können wir die Anteile der Jahre mit mittleren Temperaturen über Null, unter Null und Null unterscheiden und anteilsmässig als 60.7%, 38.0% und 1.3% bestimmen. Die Anteile werden als Winkel und damit Flächen übersetzt. In unserem Beispiel muss der Anteil von 0.607 sich im Winkel $\varphi = 0.607 \cdot 360 = 218.5$ zeigen. Aber auch die Fläche, $F = \pi 1^2 \frac{218.5}{360} = 0.607 \cdot \pi$.



30.2 Übung Aus dem Beispiel der Übung 30.1 machen wir nun aus dem Histogramm ein Kreisdiagramm. Das Kreisdiagramm ist eine Alternative zum Histogramm. Die Anteile kennen wir schon, es sind dies nach Lohnklassen 10, 40, 30, 20%. Die unterschiedlichen Klassenbreiten sind hier, ohne Verlust der Übersicht, nicht mehr sichtbar. ◁



30.3.6 Streudiagramm

Nun gehen wir von einen Merkmal zu zwei verbundenen Messungen über, also dem Paar Temperatur und Niederschlag. Gibt es mögliche Zusammenhänge zwischen den beiden? Das Streudiagramm, engl. Scatter-Plot, zeichnet alle Wertepaare als Koordinaten ein.

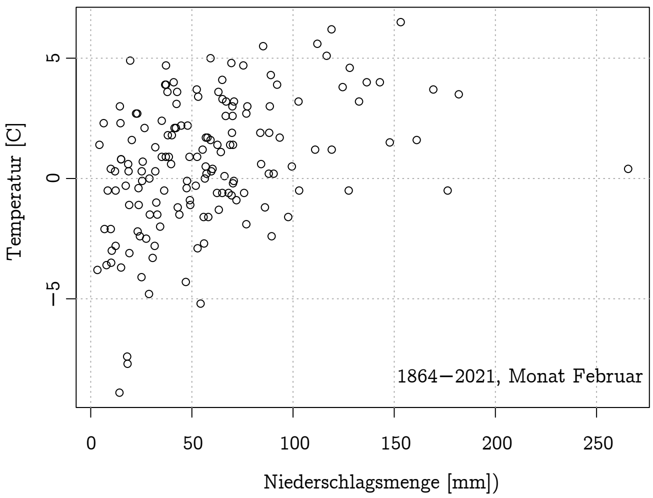


Abbildung 30.1

Wir haben die zwei Dimensionen Temperatur T und Niederschlag N mit n Messpunkten. Es werden alle n Paare (T_j, N_j) oder (N_j, T_j) aufgetragen. Welche Zahlenreihe die Ordinate und die Abszisse bilden, ist mehr eine ästhetische Frage. Es entsteht eine Wolke mit einer charakteristischen Form. Diese gilt es zu interpretieren. Wir gehen später darauf ein.

Wichtig 1. ☠ Man sollte es unbedingt vermeiden, Diagramme unnötig mit leeren Merkmalen wie dreidimensionalen Stäben und Scheiben, Schattenwürfe u.d.g zu schmücken. ⊣

30.4 Daten und Information

Daten sind die Basis, denn sie werden, von Maschinen erzeugt oder von Menschen erfragt, gesammelt. Mit der Zusammenfassung, mit Kontext versehen, mit dem Erkennen von Mustern und anderen Resultaten der Analyse werden Informationen erzeugt. Informationen dienen der Wissenserweiterung.

30.4.1 Stichprobenparameter

30.4.2 Lage

Für nominal und ordinal skalierte Messwerte ist der Modus eine sehr eingängiger Parameter.

Definition 4. Der *Modus* einer Stichprobe ist der am häufigsten vorkommende Wert.

Anmerkung 30.1. Der Modus ist nicht eindeutig. Es kann zwei oder mehr Modi geben. Bei zwei Modi hat man eine Häufigkeitsfunktion mit zwei Höckern.

Definition 5. Der *Median* oder *Zentralwert* einer geordneten Stichprobe ist der mittlere Wert, der Wert in der Mitte.

$$\tilde{x} = \begin{cases} x_{m+1} & \text{für ungerades } n = 2m{+}1 \\ \frac{1}{2}(x_m + x_{m+1}) & \text{für gerades } n = 2m \end{cases}$$

30.2 Beispiel In unserem laufenden Beispiel sind die Werte:

$$\tilde{x} = 0.65\,^\circ\mathrm{C} \qquad \text{und} \qquad \tilde{y} = 52.5\,\mathrm{mm}$$

◁

Definition 6. Das *arithmetische Mittel* (oder *Durchschnitt* oder *Populationsmittelwert*) ist die Summe der gegebenen Werte geteilt durch die Anzahl der Werte.

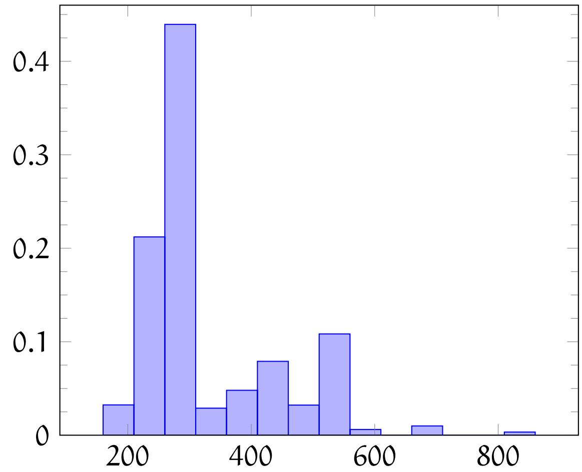
$$m_X = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

30.3 Beispiel Wir bestimmen die Mittelwerte der Temperaturen und Niederschläge. Es ist

$$m_X = 0.666\,^\circ\mathrm{C} \qquad \text{und} \qquad m_Y = 58.1\mathrm{mm}$$

Diese sind recht nahe bei den Medianen. ◁

Der Median gilt als sehr robustes Mass, weil geringe Veränderung in der Stichprobe sich nicht auswirken. Anderseits kann man aus der Angabe des Medians nicht viel schliessen, insbesondere nicht auf auf den Gesamtwert der Stichprobe. Der Mittelwert anderseits änderst sich auch bei geringen Änderungen.



30.4 Übung Für frei praktizierende Schweizer Ärzte sind für das Jahr 2014 folgende Angaben zum Einkommen bekannt: Median: Fr. 256'706, Mittelwert Fr. 320'209. Wie kann man das interpretieren? Angenommen jeder der unteren Hälfte bezöge den Median, dann würden die oberen 50% Fr. 383'712 verdienen, damit es mit dem Mittelwert wieder aufgeht. Nun ist es aber viel extremer, denn ein Jugendpsychiater verdient im Mittel Fr. 187'002 und eine Neurochirurg Fr. 817'897. Mit dem Mittelwert kann ich einen Teil der Gesundheitskosten berechnen: Anzahl Ärzte mal Durchschnittseinkommen. Mir dem Median ist dies nicht möglich. Bis 2018 wurde immer nur der Median veröffentlicht. Wieso? ◁

30.5 Übung ** Die Anzahl praktizierender Ärzte hat von 2014 von rund 17'800 zu 2020 auf 20'300 zugenommen. Angenommen diese Zuflüsse würden alle den Median verdienen und der alte Bestand erführe eine Einkommenserhöhung von 2.5%. Wie sieht der neue Mittelwert aus? Die Rechnung ist

$$m_{neu} = \frac{1}{20300}[17800 \cdot 320209 \cdot 1.025 + 2500 \cdot 256706] = 319407$$

Ist der Lohn der Ärzte gesunken? ◁

Mit dem Median haben wir die geordnete Stichprobe halbiert, in eine obere und eine untere Hälfte. Nun kann man die Liste aber noch weiter unterteilen.

Anzahl Teile	Benennung	Anzahl Werte
2	Median	1
4	Quartile	3
10	Dezile	9
100	Perzentile	99

Das erste und dritte Quartil wird häufig verwendet, das zweite ist der Median. Der sogenannte *Interquartilabstand* ist die Differenz von drittem und erstem. Zwischen diesen sind ebenfalls 50% der Messpunkte eingeschlossen. Der Abstand beantwortet die Frage, innerhalb welcher Spanne variieren die mittleren Messwerte.

30.6 Übung Für die Ärzteeinkommen kann man die Quartile als 230'000, 257'000 (Median) und 375'000 abschätzen. Dass das zweite Quartil, Median genannt, so nahe am ersten liegt, zeigt, wie asymmetrisch die Verteilung ist. Zudem liegen 25% aller Einkommen in diesem kleinen Bereich von $230 < X < 257$ in Tausend. Der Interquartilabstand ist 145'000. ◁

30.4.3 Streuung

Das einfachste, aber nicht sehr wichtige, Mass für die Streuung ist die Spannweite, also die Differenz von maximalen und minimalem Wert

$$R = max(x_i) - min(x_i)$$

Das Mass aller Dinge für die Streuung ist die Varianz.

Definition 7. Die *Stichprobenvarianz* einer statistischen Variablen X der Stichprobe $(x_1, x_2, ..., x_n)$ ist ein Mass für die Streuung der einzelnen x_i-Werte, um den Mittelwert m_X:

$$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - m_X)^2$$

Anmerkung 30.7. Dass man bei der Stichprobe die um eins verringerte Anzahl $n-1$ verwendet, ist eine schwierige Sache. Die mittlere quadratische Abweichung verwendet den Nenner n. Verwendet man die Varianz als *Schätzer* der Grundmenge der Stichprobe, so soll man $n-1$ verwenden. Für die Darstellung der Stichprobe ist n besser geeignet.

30.8 Übung Eine Stichprobe zum IQ-Test ergibt bei 5 Probanden die Werte 92,95,84,117,103. Der Mittelwert ist 98.2. Die Summe der quadratischen Abweichungen vom Mittelwert ist $38.44 + 10.24 + 201.64 + 353.44 + 23.04 = 626.8 = S$. Damit folgt $\frac{S}{5} = 125.36$ und $\frac{S}{4} = 156.7$. Nun wissen wir aber, dass der Mittelwert der Grundgesamtheit beim IQ-Test 100 ist und nicht wie in der Stichprobe 98.2. Die "wahre" quadratische Abweichung von 100 ist 128.6 und damit grösser als 125.36. Die Verwendung von $n-1$ führt zu einer grösseren Varianz, weil eben die Abweichung vom Mittelwert der Grundgesamtheit grösser ist als die vom Stichprobenmittelwert. ◁

Wichtig 2. Die korrigierte Stichprobenvarianz ist insbesondere dann vorzuziehen, wenn die Stichprobe klein ist, da sie einen besseren Schätzwert für die Varianz der Grundgesamtheit liefert. ⊣

30.9 Beispiel Die Varianzen und ihre Quadratwurzeln sind

$$s_X^2 = 7.41 \qquad \text{und} \qquad s_Y^2 = 1718$$
$$s_X = 2.72 \qquad\qquad\qquad s_Y = 41.45$$

hier mit n gerechnet, weil wir die Angaben später für Graphiken verwenden werden. ◁

Anmerkung 30.10. Die Varianz ist auch die *mittlere quadratische Abweichung* vom Mittelwert (wenn man n anstatt $n-1$ verwendet).

Die Varianz kann man auch umformen (Verschiebungssatz der Varianz)

$$s_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 = \frac{1}{n}[\sum_{i=1}^{n}(x_i^2 - 2mx_i] + m^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2m^2 + m^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - m^2$$

Definition 8. Die Quadratwurzel der Varianz nennt man *Standardabweichung* und schreibt

$$s_X = \sqrt{s_X^2}.$$

Anmerkung 30.11. Die Standardabweichung hat wieder die Einheit der gemessenen Variablen.

30.4.4 Kovarianz, Korrelation

Da die Stichprobe oder Erhebung in unserem Fall zwei Variablen betrifft, nämlich Temperatur X und Niederschlagsmenge Y, fragen wir nach Gemeinsamkeiten, nach möglichen Zusammenhängen. Die naheliegendste Grösse ist die gemeinsame Variation, die Kovariation.

Definition 9. Ist $(x_1, y_1), \ldots, (x_n, y_n)$ eine Stichprobe zweier statistischer Variablen X und Y, dann ist die *Stichprobenkovarianz* definiert als durchschnittliches Abweichungsprodukt:

$$c_{XY} = \frac{1}{n} \sum_{i=1}^{n} (x_i - m_X)(y_i - m_Y)$$

mit m_X und m_Y den Mittelwerten von X und Y.

30.12 Beispiel Die Kovarianz von Temperatur und Niederschlag ist

$$c_{XY} = 43.93\,[C \cdot mm]$$

◁

Anmerkung 30.13. Auch für die Kovarianz kann man eine Verschiebung darstellen:

$$c_{XY} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i - m_X m_Y$$

Definition 10. Die *Korrelation* r_{XY} zweier statistischer Variablen X und Y der Stichprobe $(x_1, y_1), \ldots, (x_n, y_n)$ ist

$$r_{XY} = \frac{c_{XY}}{\sqrt{s_X^2}\sqrt{s_Y^2}}$$

mit der Kovarianz c_{XY} und den Varianzen s_X^2 und s_Y^2.

30.14 Beispiel Die Korrelation zwischen Temperatur und Niederschlag ist

$$r_{XY} = 0.389$$

Da die Korrelation positiv ist, besteht ein Relation, die mehr von einem mit mehr vom anderen beinhaltet. Also höhere Temperaturen gehen mit höheren Niederschlägen einher. ◁

Anmerkung 30.15. Die Korrelation ist eine dimensionslose Zahl. In unserem Beispiel haben die zwei Variablen die Einheiten °C und mm. Durch die Division mit den Varianzen fallen die Einheiten weg.

Anmerkung 30.16. Schreibt man die Korrelation mit der Setzung $a_i = (x_i - m_X)$ und $b_i = (y_i - m_Y)$ wie folgt

$$r_{ab} = \frac{a_i b_1 + a_2 b_2 + a_3 b_3 \ldots}{\sqrt{a_1^2 + a_2^2 + a_3^2 + \ldots}\sqrt{b_1^2 + b_2^2 + b_3^2 + \ldots}}$$

dann fällt einem, vielleicht, die Formel ein

$$\cos(\beta) = \frac{\vec{a} \bullet \vec{b}}{a \cdot b} = \frac{a_i b_1 + a_2 b_2 + a_3 b_3}{\sqrt{(a_1^2 + a_2^2 + a_3^2)}\sqrt{(b_1^2 + b_2^2 + b_3^2)}}$$

Die Korrelation kann man als Kosinus des Zwischenwinkels der n-dimensionalen Vektoren $\vec{a}$ und $\vec{b}$ auffassen. Der Kosinus kann nur Werte zwischen -1 und 1 annehmen. Da es nur einen Zwischenwinkel gibt, folgt $r_{XY} = r_{YX}$.

Interpretation Korrelation

Erste Untersuchungen der Korrelation betrafen die Höhe von Vätern und Söhnen. Dabei wurde die Korrelation eingeführt. Sie zeigt das gemeinsame Schwanken zweier Grössen, eben der Höhe von Vätern und Söhnen. Nun wird oft fälschlicherweise geschlossen, dass es einen ursächlichen Zusammenhang zeige, also eine *Kausalität*. Es ist klar, dass die Höher der Väter über die Vererbung einen Einfluss auf die Söhne haben. Da aber die Korrelation zwischen Vätern uns Söhnen dieselbe ist wie zwischen Söhnen und Vätern, kann man nicht schliessen, dass die Höhe der Söhne die Höhe der Väter bewirkt.
Was man vermuten kann, ist, dass es einen gemeinsamen Faktor im Hintergrund gibt, der für beide Höhen verantwortlich ist, die Gene. Diese haben partiell einen Einfluss und eine Ursächlichkeit. Nur werden sie in der Korrelation nicht gemessen. Es werden zweimal die Wirkungen derselben Ursache gemessen.
Nun muss es aber nicht sein, dass es immer eine versteckte Ursache gibt.

30.17 Übung Es wird eine Korrelation zwischen dem Verkauf von Sonnenbrillen und Hitzeschlägen dokumentiert. Was könnte der Zusammenhang sein? Das sonnige Wetter. Der Kauf der Sonnenbrillen verursacht keine Hitzeschläge. Man kann eine positive Korrelation vermuten. Anders sähe der Verkauf von Regenschirmen und der Verkauf von Sonnencrème aus. $\triangleleft$

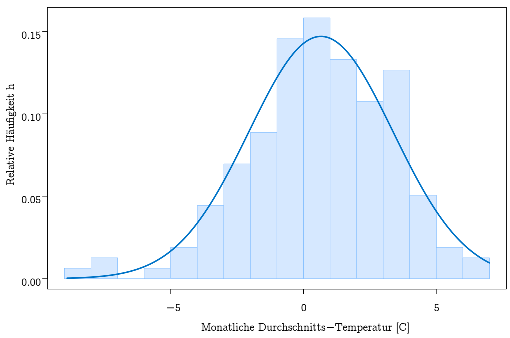
Die Korrelation kann als Informationsquelle für Vorhersagen aufgefasst werden. Bei einer hohen positiven Korrelation kann ich aus den Hitzeschlägen auf den Verkauf der Sonnenbrillen schliessen, oder aus der Höhe des Sohnes die Höhe des Vaters abschätzen oder aus den Niederschlagsmengen eine Abschätzung der Temperatur vornehmen.

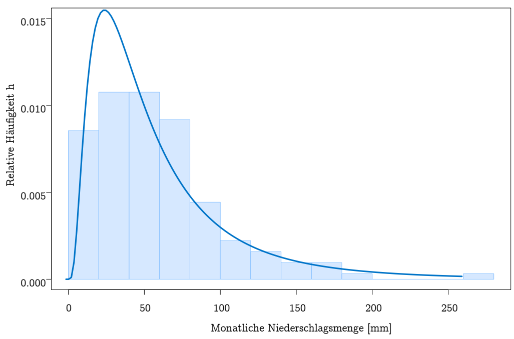
30.4.5 Form, Schiefe**

Wenn wir die Häufigkeitsfunktionen von Temperatur und Niederschlagsmenge anschauen, dann erkennen wir, dass Lage und Streuung nicht alles, und vielleicht nicht genug aussagen. Das Mass für die Schiefe ist definiert als:

$$g_X = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - m_X}{s_X}\right)^3 .$$

Sie zeigt an, ob und wie stark die Verteilung nach rechts (rechtssteil, linksschief) oder nach links (linkssteil, rechtsschief) geneigt ist. Jede nicht symmetrische Verteilung heisst schief.





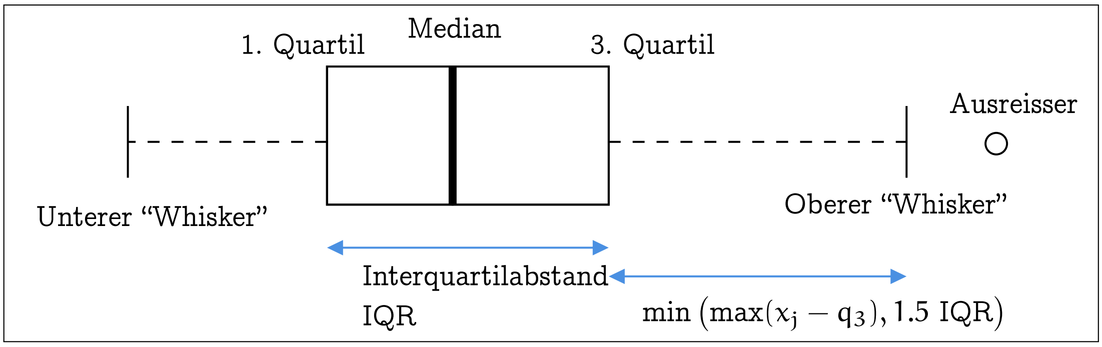
30.18 Beispiel Für unsere Daten aus dem laufenden Beispiel folgt:

$$g_X = -0.53 \qquad \text{und} \qquad g_Y = 1.54$$

Die Temperatur X ist leicht rechststeil oder linksschief, die Niederschlagsmenge ist deutlich rechtsschief und asymmetrisch. Dies erkennt man auch gut aus den Abbildungen, wo idealisierte Kurven zur Verdeutlichung überlagert sind. ◁

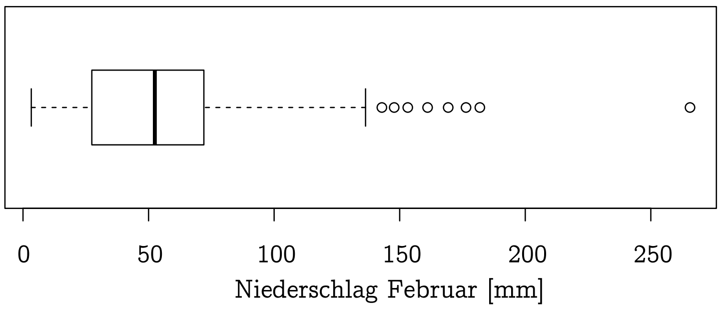
30.4.6 Box-Plot

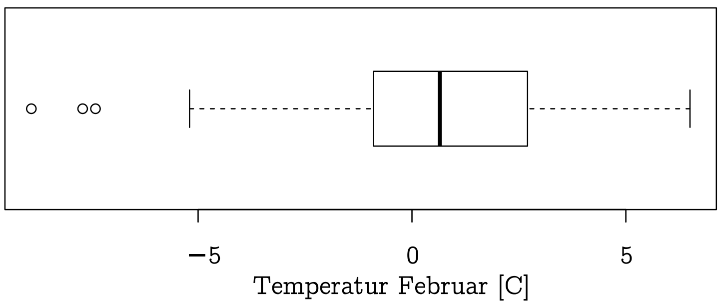
Der Box-Plot oder Kastendiagramm ist die beste zusammenfassende Darstellung für eine Verteilung kardinaler Messungen.



Den Median kann man als Symmetrieachse anschauen, denn er trennt genau hälftig die unteren von den oberen Messpunkten. Liegt er ziemlich mittig zwischen den zwei Whiskers ("Schnurrhaaren" oder auch Antennen), dann ist die Verteilung eher symmetrisch, liegt er mehr bei einem Whisker, dann ist die Verteilung asymmetrisch.
Die Whisker geben das Maximum und Minimum an, sofern diese nicht mehr als 1.5 Interquartilabstände von den Quartilen entfernt liegen. Messwerte ausserhalb dieses Spanne werden als *Ausreisser* klassifiziert. Somit: gibt es keine Ausreisser, dann ist der Whisker der extremste Wert ansonsten nicht.

30.19 Übung Wir wenden die Box-Plots auf unser laufendes Beispiel an. Die Situation stellt sich wie folgt dar.





Die Interpretationen kreisen um die dargestellten Kennzahlen, also 2 Quartile, Median, Maximum und Minimum und ihrer gegenseitigen Lage.

Bei den Niderschlägen sieht man, dass die Box mit 50% der Masse weit links im Verhältnis zu den Extrema liegt, die Verteilung also asymmetrisch und linkssteil ist. Der Median ist nur leicht rechts im Quartilintervall. Es hat viele Ausreisser nach oben, unten ist das Minimum viel näher zum Median.
Bei der Temperatur liegt die Box mittig zwischen den Whiskern. Sie ist damit recht symmetrisch. Es hat wenige Aussreisser nach unten. ◁

30.4.7 Konzentrationsmass**

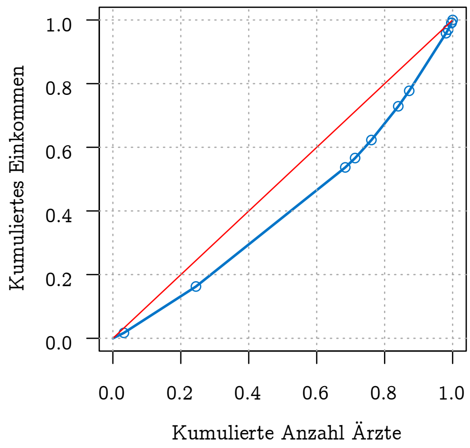
Dieses Mass ergänzt das Streuungsmass. Das Konzentrationmass will eine Aussage darüber machen, wie stark einzelne Ausprägungen oder Klassen zur Summe aller Ausprägungen beitragen. Das gebräuchlichste stammt aus er Volkswirtschaft, wo gefragt wird, wie das Einkommen bei den Einwohnern verteilt ist. Das bekannteste Mass ist der sogenannte Gini-Index. Um es plastisch zu verstehen, bietet sich die *Lorenzkurve* an.

30.20 Beispiel Wir nehmen die Daten von den Ärzteeinkommen, für die wir im Histogramm schon die Klassen mit den Mitten m_j gebildet haben und und zugehörigen relative Häufigkeiten h_j. Uns interessiert der Anteil der Ärzte am Anteil des Einkommens, also "x% der Ärzte verdienen y% des gesamten Einkommens". Dazu werden die Häufigkeiten der Ärzte kumuliert und der Anteil des Gesamteinkommens ebenso. Dieses ergibt sich aus der Häufigkeit mal Klassenmitten. Dann werden beide Variablen kumuliert und normiert, so dass beide Variabeln von 0 bis 1 laufen.

Klassenmitten m_j	Häufigkeiten h_j	Kumulierte H_j	$q_j = m_j h_j$	Kumuliert	Anteil G_j
160	0.03241	0.03241	5.1856	5.19	0.01698
210	0.21210	0.24454	44.541	49.73	0.16280
260	0.43950	0.68402	114.27	164.00	0.53692
310	0.02893	0.71295	8.9683	172.96	0.56628
360	0.04804	$\boxed{0.76099}$	17.2944	190.26	$\boxed{0.62290}$
410	0.07902	0.84001	32.3982	222.66	0.72897
460	0.03222	0.87224	14.8212	237.48	0.77749
510	0.10835	0.98059	55.2585	292.74	0.95841
560	0.00612	0.98671	3.4272	296.16	0.96963
610	0.00000	0.98670	0	296.16	0.96963
660	0.00992	0.99663	6.5472	302.71	0.99106
710	0.00000	0.99663	0	302.71	0.99106
760	0.00000	0.99663	0	302.71	0.99106
810	0.00337	1.00000	2.7297	305.44	1.00000

Aus der Tabelle lesen wir heraus, dass die unteren 76% der Ärzte 62% des Gesamteinkommens verdienen oder die untersten 3.2% verdienen nur 1.7% des Gesamten.
Die Lorenzkurve beschreibt die Punkte (G_j, H_j)

In der Abbildung sind neben den Punkten (G_j, H_j) auch die Diagonale eingetragen. Diese Kurve besagt, dass genau alle gleichviel verdienen, denn der relative Anteil ist konstant. Die Ungleichheit wird als Differenz zu diesem Ideal gemessen und zwar als der Flächenanteil, der zwischen den beiden Kurven liegt. Je grösser dieser Anteil desto "ungerechter" die Verteilung.



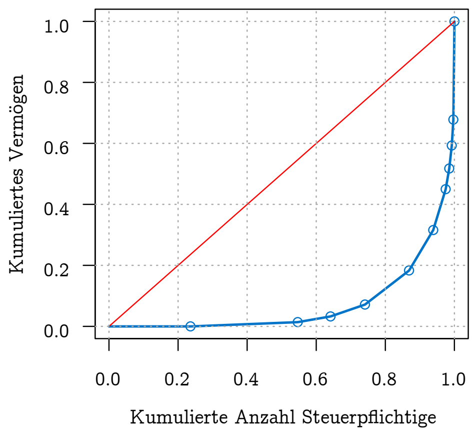
Mit (G_j, H_j) lässt sich die Fläche berechnen. Wir bestimmen die Fläche unter der blauen Kurve, multiplizieren mit 2 und ziehen diesen Wert von 1, der die Gesamtfläche darstellt. Diese Zahl ist der Gini-Index. Wie man aus der Abbildung ersehen kann, liegt der Gini-Index zwischen 0 und 1. Ein Wert von 1 würde bedeuten, dass nur eine Person alles verdient.
Die Formel lautet (es gibt verschiedene Formulierungen):

$$G = \frac{\sum_{i=1}^{n}(H_i + H_{i-1})q_i}{\sum_{i=1}^{n} q_i} - 1$$

wobei $H_0 = 0$ gilt. In unserem Beispiel folgt $G = 0.183$. Die Ungleichheit ist sehr gering. Der Gini-Index für das gesamtschweizerische Einkommen ist rund 0.35 - 04. ◁

30.21 Beispiel Von der Steuerverwaltung sind für das Jahr 2018 folgende Daten bekannt, die für alle 5.42 Millionen Steuerpflichtige erhoben wurden:

Klassenmittelwert [Mio CHF] m_j	Prozentualer Anteil h_j
0.0000	23.582
0.0167	31.049
0.0727	9.500
0.1447	9.972
0.3238	12.760
0.7012	6.993
1.3763	3.604
2.4230	1.035
3.8092	0.732
6.8612	0.458
37.7037	0.316



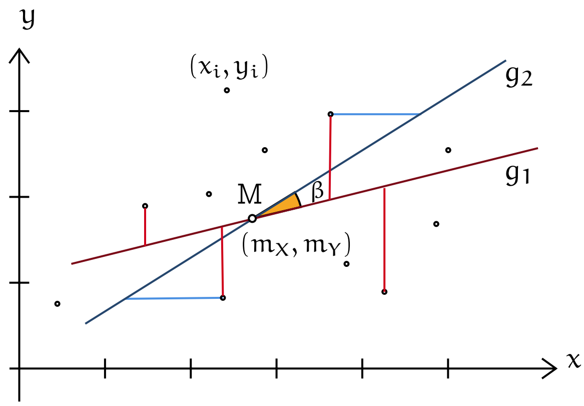
Der Gini-Index ist 0.846, was in der Lorenzkurve plausibel wird. Die reichsten 3.2‰besitzen 32% des Vermögens. Diese Verteilung ist in der Zeit recht stabil. Derselbe Wert für die Welt ist 2018 0.904.

Man erkennt, woher die 80-20-Paretoregel kommt. Die Kurve läuft nahe am Punkt $(0.8, 0.2)$ vorbei, wo 80% der Steuerpflichtigen 20% des Vermögens besässen.

◁

30.5 Modellieren**

30.5.1 Lineare Regression



Wir kommen auf die Abb. 30.3.6 zurück, dem Streudiagramm von Temperatur X und Niederschlag Y. Eine einfache Idee ist es, den Punkthaufen (x_i, y_i) durch eine lineare Funktion zu ersetzen, also das gemessene y_i durch $\hat{y}_i = b_1 x_i + c_1$ zu ersetzen. Oder umgekehrt, x_i durch $\hat{x}_i = a_2 y_i + c_2$ zu modellieren. Es gibt also zwei mögliche lineare Gleichung, je nach dem, welche Variable auf die Gerade gezwungen wird. Es gibt $n > 2$ Punkte der Stichprobe, die Gerade hat zwei Parameter. Das Gleichungssystem ist überbestimmt. Wir fordern, dass die Mittelwerte m_X und m_Y auf den Geraden liegen soll. Die Punkt-Richtungsform der Geraden (Differenz $y - y_0$ gleich Steigung mal Differenz $x - x_0$) ist

$$\hat{y} - m_Y = b_1(x - m_X) \quad \text{und} \quad \hat{x} - m_X = b_2(y - m_Y) \qquad \Leftrightarrow \qquad y - m_Y = \frac{1}{b_2}(\hat{x} - m_X)$$

Die Steigungen bestimmt man aus der Minimierung der Quadrate der Abweichungen. Diese sogenannte *Methode der kleinsten Quadrate* stammt von Carl F. Gauss. Die Zielfunktionen sind

$$z_1 = \frac{1}{n}\sum_{k=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - [b_1(x_i - m_X) + m_Y])^2 = \min!$$

Nun leiten wir z_1 nach b_1 ab, wobei wir die Substitution $\delta_i = x_i - m_X$ vornehmen. Dabei brauchen wird die Kettenregel $[(f(x))^2]' = 2f(x)f'(x)$:

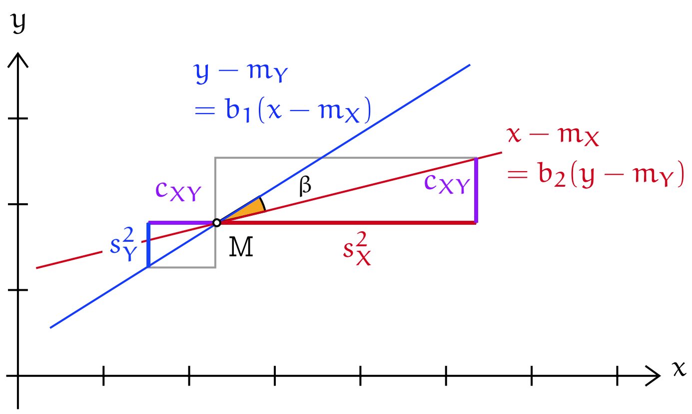
$$\begin{aligned}
\frac{dz_1}{db_1} = 0 &= \frac{1}{n}\sum_{k=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n} 2(y_i - b_1\delta_i - m_Y)(-\delta_i) \\
&= \frac{2}{n}\sum_{i=1}^{n}\left(-\delta_i y_i + b_1\delta_i^2 + \delta_i m_Y\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}[-\delta_i(y_i - n_Y)] + b_1\frac{1}{n}\sum_{i=1}^{n}\delta_i^2 \\
&= -c_{XY} + b_1 s_X^2
\end{aligned}$$

Damit folgt

$$b_1 = \frac{c_{XY}}{s_X^2} \quad \text{und ganz analog} \quad \frac{1}{b_2} = \frac{c_{XY}}{s_Y^2}$$

Aus den Formel sieht man sofort, dass gelten muss

$$c_{XY} = \sqrt{\frac{b_1}{b_2}}.$$



In der Abbildung sieht man die Steigungen der zwei Geraden als Verhältnis von Kovarianz c_{XY} und den Varianzen s_X^2 und s_Y^2. Die zwei Regressionsgeraden kann man also einfach bestimmen, wenn man die Stichprobenparameter schon berechnet hat. Die Zeichnung hängt von den Einheiten der zwei Variablen ab. Die Niederschlagsmenge ist eine andere Grössenordnung als die Temperatur.

30.1 Übung Wir betrachten unser laufendes Beispiel. Wir haben schon bestimmt $c_{XY} = 43.93$, $s_X^2 = 7.41$ und $s_Y^2 = 1718$ sowie die Mittelwerte $M = (0.666, 58.1)$. Damit folgen die Steigungen der Regressionsgeraden und die Geraden:

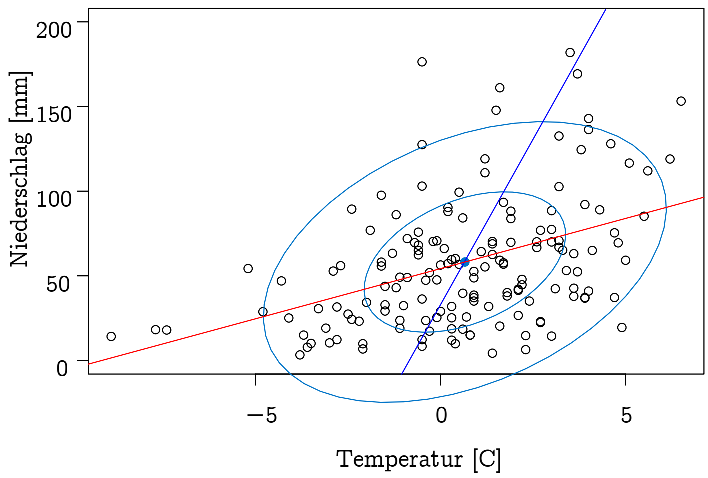
$$\begin{aligned}
g_1 : \quad & \hat{y} - 58.1 = \frac{43.93}{7.41}(x - 0.666) = 5.93 \cdot (x - 0.666) \\
g_2 : \quad & \hat{x} - 0.66 = \frac{1718}{43.93}(y - 58.1) = 39.1 \cdot (y - 58.1)
\end{aligned}$$

Die Gerade g_2 kann man auch umschreiben, so dass die y-Koordinate von $\hat{x}$ abhängt:

$$g_2 : \quad y_2 = \frac{1}{39.1}(x - 0.666) + 58.1 = 0.0256 \cdot (x - 0.666) + 58.1$$

und analog

$$g_1 : \quad y_1 = 5.93 \cdot (x - 0.666) + 58.1$$



Wie man gut erkennt, unterscheiden sich die zwei Geraden erheblich. Die Korrelation kennen wir, sie ist $r_{XY} = 0.389$. Die Ellipsen schliessen Punkte gleichen Abstands vom Mittelwert dar, also Punkte (x_i, y_i), die in den Intervallen $(m_X - s_X < x_i < m_X + s_X)$ und $(m_Y - m_Y < y_i < m_Y + s_Y)$ liegen. $\triangleleft$

Jetzt kann man immer eine Gerade durch einen Punkthaufen zeichnen. Mit der Optimierung wird der beste Wert gefunden, die Gerade, welche die kleinste quadratische Abweichung aufweist. Es stellt sich aber zusätzlich die Frage, wie gut die Gerade die Punkte darstellt. Hier gibt es das *Bestimmtheitsmass* R^2, das von den Statistik-Programmen mitgerechnet wird. Die Summe der quadratischen Abweichung wird wie folgt aufgespalten:

$$\underbrace{\sum_{i=1}^{n}(y_i - m_Y)^2}_{\text{SST}} = \underbrace{\sum_{i=1}^{n}(\hat{y}_i - m_Y)^2}_{\text{SSR}} + \underbrace{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{\text{SSE}}$$

Dabei steht SS für *sums of squares* und T für total, R für Regression und E für Error. Damit legt man das Bestimmtheitsmass fest als

$$R^2 = \frac{\text{SSR}}{\text{SST}} = 1 - \frac{\text{SSE}}{\text{SST}}$$

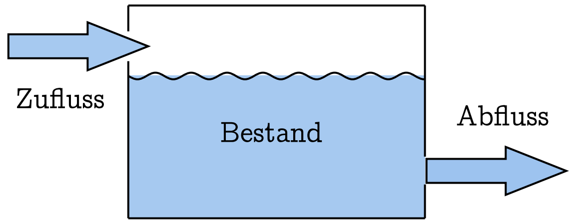
Der Wert liegt zwischen 0 und 1, und je näher bei 1 desto besser. Unsere Regression weist einen Wert von 0.15 auf, ist also relativ schlecht. Das heisst, der Zusammenhang im Punkthaufen wird durch eine lineare Beziehung schwach erklärt.

30.6 Bestand und Bewegung

30.6.1 Reservoirmodell

Das Reservoir oder Brunnen ist ein sehr weit verbreitetes Modell. Es fusst auf der Massenerhaltung in einem Kontrollvolumen. Es besagt, dass die Änderung des Bestandes im Kontrollvolumen gleich der Differenz von Zu- und Abfluss sein muss. Denn es geht nichts verloren.

Der Bestand im Volumen ist eine Funktion des *Zeitpunkts*. Zu jedem Zeitpunkt t gibt es einen Wert $B(t)$. Die Bewegungsgrössen von Zu- und Abfluss Z, A sind die Mengen, die sich über eine *Zeitperiode* ansammeln, $Z(t_2 - t1)$ sowie $Z(t_2 - t_1)$. Es gilt also allgemein



$$B(t_2) - B(t_1) = Z(t_2 - t_1) - A(t_2 - t_1)$$

oder

$$B(t_2) = B(t_1) + Z(t_2 - t_1) - A(t_2 - t_1)$$

Der Endbestand ist der Anfangsbestand plus Zuflüsse minus Abflüsse in der Zwischenzeit.

30.1 Übung Ein Lager an Kupfer enthält anfangs Jahr 12 Tonnen. 3.5 Tonnen werden im Laufe des Jahres verbraucht, 2.5 t werden verkauft und 7t werden gekauft. Wie gross ist der Bestand am Ende des Jahres? $B(t_2) = B(t_1) + 7 - 3.5 - 2.5 = 12 + 1 = 13$. ◁

Eine der wichtigsten staatlichen Statistiken ist die Bevölkerungsstatistik. Sie gibt Auskunft über die Anzahl Einwohner eines bestimmten Gebietes, Staat, Kanton, Städte, Gemeinden etc. Die Zu- und Abflüsse stammen aus zwei Quellen: der Wanderung und der natürlichen Änderung aus Geburt und Tod. Zufluss ist Zuwanderung + Geburt, Abfluss ist Abwanderung + Todesfälle.

30.2 Übung Die Bevölkerungsstatistik ist ein wichtiges planerisches Werkzeug der Behörden. Wir betrachten die Bevölkerungsstatistik der Stadt Zürich. Der Bestand ist die Bevölkerungszahl in einem Zeitpunkt, hier Jahresanfang und Jahresende. Die Bewegungen stammen aus der Zu- und Abwanderung (Migration) und den Geburten und Todesfällen (natürlicher Saldo). Man prüfe, ob die allgemeine Gleichung eingehalten ist. Ja, denn Bestand am Jahresanfang $B(t_1)$ plus Bewegungen $Z(t_2 - t_1) - A(t_2 - t_1)$ stimmen.

Jahr	2013	2014	2015	2016	2017	2018	Total
Bestand am 1.1.	394012	398575	404783	410404	415682	423310	2446766
Geburten	4920	5145	5191	5176	5240	5212	30884
Todesfälle	-3465	-3334	-3400	-3178	-3279	-3235	-19891
Zuzüge	43599	42152	42473	42318	43354	42998	256894
Wegzüge	-40491	-37755	-38643	-39038	-37687	-39548	-233162
Bestand am 31.12	398575	404783	410404	415682	423310	428737	
Saldo	4563	6208	5621	5278	7628	5427	34725

Die Bevölkerung hat um $p\%$ zugenommen, numerisch $B(2019) = (1+p)^6 B(2013)$ und daraus $\left(\frac{B(2019)}{B(2013)}\right)^{1/6} - 1 = \left(\frac{428737}{394012}\right)^{1/6} - 1 = 0.0142 = 1.42\%$. Die mittlere Verweildauer (siehe unten) ist rund 10 Jahre. ◁

Verweildauer, Umschlag

Man kann sich bei diesem Modell fragen, wie lange ein bestimmter Tropfen im Reservoir bleibt. Dazu muss man ein paar Annahmen treffen. Erstens soll der Abfluss in der Reihenfolge des Zuflusses erfolgen. Wenn A vor B ins Reservoir gekommen ist, so soll es auch vor B den Behälter verlassen. Das ist die fifo-Regel für “first in, first out”.

30.3 Übung Ein Spital hat folgenden Bestand und Bewegung in einer Woche. Insgesamt 74 Zugänge, 67 Abgänge und 692 Patiententage.

Wochentag t	$Z(t)$	$A(t)$	$B(t)$
1	21	9	91
2	10	23	93
3	8	0	103
4	5	10	105
5	17	6	96
6	2	6	101
7	11	13	103
Total	74	67	692

Die mittlere *Verweildauer* V ergibt sich zu

$$V = \frac{692}{\frac{1}{2}74 + \frac{1}{2}67} = \frac{692}{70.5} = 9.82$$

Daraus ergibt sich die mittlere *Umschlagshäufigkeit* U zu:

$$U = \frac{7}{9.82} = 0.71$$

Wenn nun Massnahmen getroffen werden, um die Verweildauer um 20% zu verringern, wieviele Betten könnten dann eingespart werden? Die Bewegungen bleiben gleich, also müssen die Patiententage um 20% zurückgehen, also kann man 20% der Betten einsparen. ◁

30.6.2 Abgangsmodell

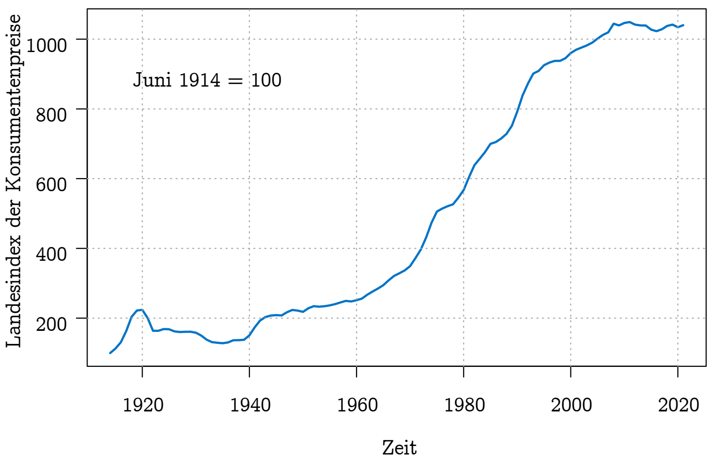
Ein Spezialfall des Reservoirmodells ist das Abgangsmodell, bei dem es keine Zuflüsse gibt sondern nur Abflüsse. Wir kennen bereits einen Vertreter dieses Modells: die Sterbetafel. Sie geht von 10000 lebend Geborenen aus und verzeichnet die abnehmenden Bestände bis zum Tafelende im Alter $\omega = 100$.

30.7 Zeitreihen

Sehr viele Daten werden im Zeitverlauf erhoben. Beispiele sind Bevölkerung, Staatsfinanzen, Wetterdaten, Börsendaten, Befragungen zu gesellschafts-politischen Einstellungen und Befindlichkeit u.v.m. Man möchte nicht nur die Vergangenheit besser verstehen, sondern Prognosen für die Zukunft machen.
Bei Zeitreihen ist also die x-Koordinate die Zeit.

30.7.1 Liniendiagramm

Wir beginnen mit einem Beispiel, dem Landesindex der Konsumentenpreise. Das ist im wesentlichen der normierte Preis eines für die Bevölkerung massgebenden Warenkorb. Am Verlauf über die Zeit kann man den Wertverzehr des Geldes abschätzen.

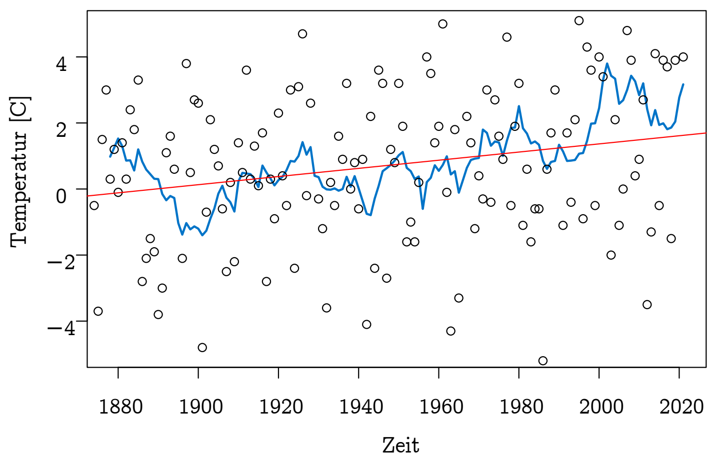


Nun könnten wir ja auch die Temperatur und die Niederschläge als Liniendiagramm darstellen, denn die Daten sind ja zeitlich aufgenommen worden. Nun sind aber gerade diese Werte von einer starken zufälligen Komponente überlagert. Deshalb ist die Darstellung ein wildes Auf und Ab. In solchen Fällen drängt sich eine Glättung der Daten auf.

Glättung

Ein von vielen Möglichkeiten ist die Methode des gleitenden Durchschnitts. Ein Durchschnitt mindert die Zufälligkeit, da diese meist symmetrisch ist, also Ausschläge nach oben und unten aufweist. Der Durchschnitt darf nicht allzu viele Punkte enthalten, denn dann ist die Aussage auch wieder stark relativiert. Wir wählen 15 Jahre. Somit ist die neue Variable

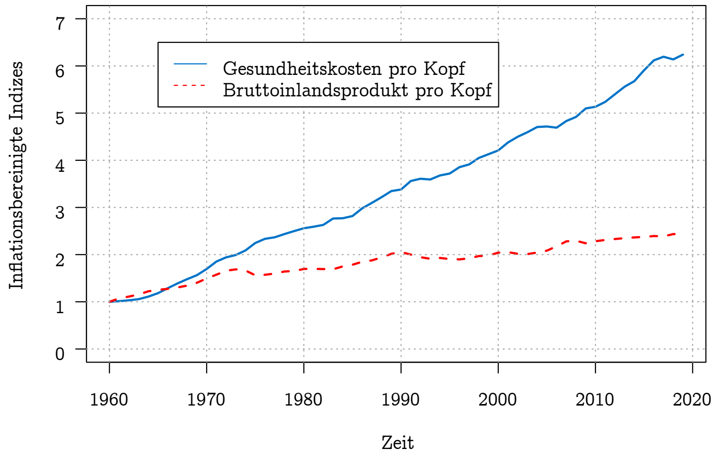
$$H(t) = \frac{1}{10} \sum_{i=t-14}^{t} X_i$$



Nun nähert die geglättete blaue Kurve die wilden Punkte der tatsächlichen Messungen an. Nur jetzt ist eine Interpretation möglich. Zudem machen wir eine lineare Regression, welche die rote Gerade hervorbringt. Jetzt ist zu erkennen, dass die Temperatur am steigen ist. Die blaue Kurve würde eine Zunahme der Steigung für die 40er-Jahre nahelegen.

30.7.2 Indexierung

Will man verschiedene Zeitreihen zusammen darstellen, weil sie thematisch ähnlich sind, so hat man zwei Möglichkeiten, die Zeitreihen auf dasselbe Bild zu bringen: entweder man verwendet zwei verschiedene Massstäbe, so dass links und rechts der Graphik je eine Beschriftung mit Skalierung dargestellt werden muss oder man normiert die Werte auf einen Bezugspunkt. In der Abbildung des Landesindex ist diese Normierung gemacht worden, indem der Wert des Warenkorbes jeweils durch den Wert für Juni 1914 geteilt wurde. Meist wählt man einen Wert weit links, so dass die Kurve bei 100% beginnt.



In der Abbildung sieht man die Gesundheitskosten und Bruttoinlandsprodukt BIP pro Kopf, wobei mit dem Landesindex der Konsumentenpreise die Frankenbeträge bereinigt wurden. Man sieht, dass die Gesundheitskosten ungebremst über dem Niveau des BIP wachsen. Das BIP ist natürlich kein perfekter Ersatz für das Einkommen, aber tendenziell kann man schon davon ausgehen, dass das verfügbare Einkommen proportional zum BIP verläuft.

Aufgaben

2.1 Die sechs Mitglieder einer Weitsprungleistungsgruppe springen (in Metern): 6.5, 6.3, 6.8, 7.4, 7.3, 7.7 Wie gross sind Median, arithmetischer Mittelwert und Standardabweichung? [71‰]

1 Wir odnen die Punkte: 6.3, 6.5, 6.8, 7.3, 7.4, 7.7.
Der Median bei gerader Anzahl hier 6, ist das Mittel von 3. und 4. Messwert, also $(6.8 + 7.3)/2 = 14.1/2 = 7.05$. Die Summe beträgt 42, die Anzahl Messungen 6, damit ist der Mittelwert 7. Für die Standardabweichung bestimmen wir zuerst die Varianz, wobei wir direkt die Abweichungen eingeben: $((0.7)^2+0.5^2+0.2^2+0.3^2+0.4^2+0.7^2)/6 = (0.49+0.25+0.04+0.09+0.16+0.49)/6 = 1.52/6 = 0.26$ und daraus die Wurzel 0.510 (Wir haben hier durch n und nicht $n-1$ dividiert).

2.2 Die monatlichen Bruttosaläre aller Buchhaltungsangestellten einer Firma wurden wie folgt erhoben: 9'500.-, 8'100.-, 6'750.-, 4'850.-, 5'100.-, 5'300.-
Berechnen Sie den Mittelwert $\bar{x}$, den Median $\tilde{x}$ sowie die Standardabweichung s der Bruttosaläre. [54‰]

2 Die Summe der Saläre ist 39600, dividiert durch 6 ergibt den Durchschnitt Fr 6'000.–. Der median ist der Zentralwert, hier Durchschnitt von 3. und 4. Wert der geordneten Liste. Also $\tilde{x} = (5300 + 6750)/2) = 12050/2 = 6025$. Die Varianz ergibt sich aus (Verschiebungssatz) $s^2 = (9500^2 + 8100^2 + 6750^2+4850^2+5100^2+5300^2)/6-6600^2 = (90250000+65610000+45562500+23522500+26010000+28090000)/-43560000 = 279045000/6-43560000 = 46507500-43560000 = 2947500$. Damit $s = \sqrt{s^2} = 1717$. (Hätte man $n-1 = 5$ anstatt $n = 6$ verwendet, dann wäre $s = 1881$.) Wenn der Taschenrechner über statistische Funktionen verfügt, nutze man diese.

2.3 Die Zeugnisnoten einer Maturitätsprüfung im Fach Mathematik sehen wie folgt aus:

5	4	4	5	3.5	5	3.5	3	4	4.5
4.5	5.5	4.5	4.5	4.5	3	5	4.5	4.5	3

Berechnen Sie den Median, den Mittelwert und die Standardabweichung der Zeugnisnoten. [69‰]

3 Wir machen eine Häufigkeitstabelle:

3	3.5	4	4.5	5	5.5	Total
3	2	3	7	4	1	22
3	5	8	15	19	20	

Der Median ist das Mittel von 10. und 11. Wert. Beide sind 4.5, also $\tilde{x} = 4.5$

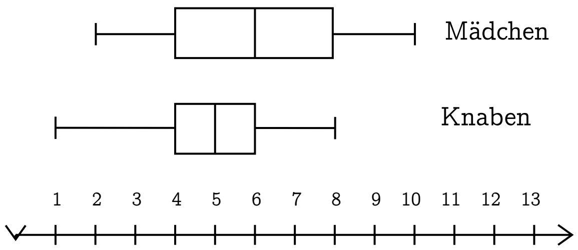
2.4 [3]Gegeben ist eine geordnete Liste mit neun Werten $a_1, a_2, \ldots, a_9$. Der Wert a_1 wird um 5 vergrössert, der Wert a_9 wird um 5 verkleinert, die restlichen Werte der Liste bleiben unverändert. Durch die Abänderung der beiden Werte a_1 und a_9 kann sich eine neue, nicht geordnete Liste ergeben.
Aufgabenstellung: Welche statistischen Kennzahlen der Liste werden durch die genannten Änderungen in keinem Fall verändert? Kreuzen Sie die entsprechende(n) statistische(n) Kennzahl(en) an!
Aussage 1: arithmetisches Mittel
Aussage 2: Median
Aussage 3: Modus
Aussage 4: Spannweite
Aussage 5: Standardabweichung

4 Beachte: $a_1 + 5$ muss nicht mehr an erster, a_9 nicht mehr an letzter Stelle stehen.
X Aussage 1: Diese Aussage ist richtig, weil sich der Mittelwert nicht ändert, da die gesamte Summe mit $5-5 = 0$ nicht ändert.
Aussage 2: Diese Aussage ist falsch, weil sich eine neue, anders sortierte Liste ergeben kann, bei der ein anderer Wert in der Mitte stehen kann. Darauf wird in der Angabe auch hingewiesen.

[3]`https://www.maths2mind.com/` aus Oesterreich.

Aussage 3: Diese Aussage ist falsch, weil 2 neue Werte ($a_1 + 5$ bzw. $a_9 - 5$) hinzukommen (und 2 alte Werte entfallen). Dadurch kann ein anderer Wert am häufigsten in der Stichprobe vorkommen.
Aussage 4: Diese Aussage ist falsch, weil sich das bisherige Minimum a_1 auf $a_1 + 5$ erhöht, während sich das bisherige Maximum a_9 auf $a_9 - 5$ reduziert. Dadurch muss sich der Abstand zwischen Minimum und Maximum verkleinern.
Aussage 5: Diese Aussage ist falsch, weil sich das bisherige Minimum a_1 auf $a_1 + 5$ erhöht, während sich das bisherige Maximum a_9 auf $a_9 - 5$ reduziert. Dadurch ändert sich die Streuung der Werte um den Mittelwert und entsprechend ändert sich die Standardabweichung.

2.5 Für Mädchen und Knaben einer Schulklasse wurde die Länge ihres Schulweges ermittelt und als Boxplots dargestellt.



Kreuzen Sie die zutreffenden Aussagen an.
Aussage 1: Mehr als 60 % der befragten Mädchen haben einen Schulweg von mindestens 4 km.
Aussage 2: Der Median der erhobenen Daten ist bei Knaben und Mädchen gleich.
Aussage 3: Mindestens 50 % der Mädchen und mindestens 75 % der Knaben haben einen Schulweg, der kleiner oder gleich 6 km ist.
Aussage 4: Höchstens 40 % der befragten Knaben haben einen Schulweg zwischen 4 km und 8 km.
Aussage 5: Die Spannweite ist bei den Umfragedaten der Knaben genauso groß wie bei den Umfragedaten der Mädchen.

5 X Aussage 1: Diese Aussage ist richtig, weil zwischen dem "linken Rand der Box" und dem "Whisker" 75% der Daten liegen und somit 75% Mädchen einen Schulweg haben, der länger als 4km ist.
Aussage 2: Diese Aussage ist falsch, weil der Median dem "Strich in der Box" entspricht und dieser für Mädchen bei 6km und für Knaben bei 5km liegt.
X Aussage 3: Diese Aussage ist richtig, weil 6km bei den Mädchen dem Median (und somit 50% entspricht) bzw. weil 6km bei den Knaben dem "rechten Rand der Box" und somit 75% entspricht.
Aussage 4: Diese Aussage ist falsch, weil 75 und nicht 40% der Knaben einen Schulweg zwischen 4km (linker Rand der Box) und 8km (rechter Whisker) haben.
Aussage 5: Diese Aussage ist falsch, weil die Spannweite bei den Mädchen beträgt 10-2=8km bei den Knaben hingegen beträgt sie 8-1=7km

Tabelle 30.1: Mittlere Temperatur [°C] und Niderschlagshöhe [mm] im Februar in Zürich-Fluntern.

Jahr	T	N	Jahr	T	N	Jahr	T	N	Jahr	T	N
1864	-1.5	32.9	1903	2.1	26.6	1943	2.2	47.9	1983	-1.6	55.8
1865	-2.9	52.8	1904	1.2	110.9	1944	-2.4	89.4	1984	-0.6	62.4
1866	3.7	169.3	1905	0.7	25.7	1945	3.6	63.1	1985	-0.6	65.0
1867	4.7	75.4	1906	-0.6	68.1	1946	3.2	102.7	1986	-5.2	54.3
1868	2.3	6.4	1907	-2.5	27.4	1947	-2.7	56.0	1987	0.6	84.2
1869	4.9	19.5	1908	0.2	88.0	1948	1.2	55.3	1988	1.7	57.7
1870	-3.1	19.1	1909	-2.2	23.2	1949	0.8	15 .0	1989	3.0	70.0
1871	0.9	37.1	1910	1.4	70.3	1950	3.2	66.9	1990	$\boxed{6.5}$	153.2
1872	0.3	59.6	1911	0.5	56.8	1951	1.9	83.8	1991	-1.1	19.0
1873	-0.1	25.4	1912	3.6	42.7	1952	-1.6	97.6	1992	1.7	56.9
1874	-0.5	12.3	1913	0.3	18.7	1953	-1.0	32.4	1993	-0.4	23.6
1875	-3.7	15.0	1914	1.3	31.9	1954	-1.6	58.1	1994	2.1	42.1
1876	1.5	147.8	1915	0.1	66.1	1955	0.2	90.4	1995	5.1	116.6
1877	3.0	88.5	1916	1.7	93.4	1956	$\boxed{-8.9}$	14.2	1996	-0.9	49.0
1878	0.3	25.2	1917	-2.8	12.3	1957	4.0	142.9	1997	4.3	89.0
1879	1.2	119.1	1918	0.3	31.9	1958	3.5	181.9	1998	3.6	37.9
1880	-0.1	47.6	1919	-0.9	72.0	1959	1.4	4.3	1999	-0.5	176.4
1881	1.4	62.6	1920	2.3	14.7	1960	1.9	69.8	2000	4.0	136.4
1882	0.3	12.0	1921	0.4	9.9	1961	5.0	59.2	2001	3.4	53.1
1883	2.4	35.1	1922	-0.5	103	1962	-0.1	70.7	2002	5.5	85.2
1884	1.8	40.1	1923	3	77.4	1963	-4.3	47.0	2003	-2.0	34.3
1885	3.3	65.1	1924	-2.4	24.3	1964	1.8	38.1	2004	2.1	41.4
1886	-2.8	31.6	1925	3.1	42.4	1965	-3.3	30.6	2005	-1.1	49.2
1887	-2.1	6.8	1926	4.7	37.2	1966	5.6	112	2006	0.0	56.4
1888	-1.5	29.2	1927	-0.2	70.3	1967	2.2	44.7	2007	4.8	69.5
1889	-1.9	76.9	1928	2.6	70.1	1968	1.4	68.9	2008	3.9	37.2
1890	-3.8	$\boxed{3.3}$	1929	-7.4	18.0	1969	-1.2	43.0	2009	0.4	60.2
1891	-3.0	10.4	1930	-0.3	17.3	1970	0.4	$\boxed{265.6}$	2010	0.9	35.1
1892	1.1	64.3	1931	-1.2	86.1	1971	-0.3	51.9	2011	2.7	22.4
1893	1.6	59.2	1932	-3.6	7.8	1972	3.0	14.4	2012	-3.5	10.1
1894	0.6	39.7	1933	0.2	57.2	1973	-0.4	47.4	2013	-1.3	63.3
1895	-7.7	18.2	1934	-0.5	8.4	1974	2.7	76.9	2014	4.1	65.0
1896	-2.1	9.9	1935	1.6	161.1	1975	1.6	20.3	2015	-0.5	36.3
1897	3.8	124.5	1936	0.9	48.8	1976	0.9	38.6	2016	3.9	92.1
1898	0.5	99.4	1937	3.2	132.6	1977	4.6	128.0	2017	3.7	52.4
1899	2.7	23.0	1938	0.0	29.0	1978	-0.5	127.5	2018	-1.5	43.8
1900	2.6	66.7	1939	0.8	15.0	1979	1.9	88.2	2019	3.9	36.8
1901	-4.8	28.8	1940	-0.6	75.8	1980	3.2	70.8	2020	6.2	119.0
1902	-0.7	69.6	1941	0.9	52.6	1981	-1.1	23.7	2021	4.0	41.0
			1942	-4.1	25.1	1982	0.6	18.5			

Normalprogramm / Erweitertes Niveau

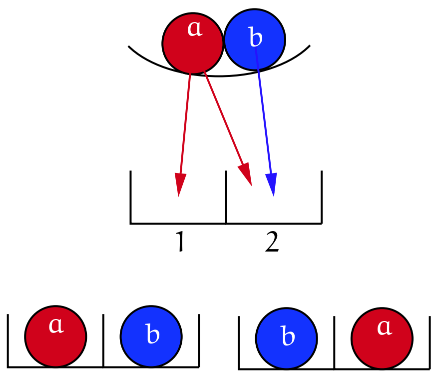
– einfache Anordnungen (Variationen, Kombinationen) mit oder ohne Wiederholungen, Permutationen mit oder ohne Wiederholungen erkennen und unterscheiden, diese abzählen und zur Lösung einfacher kombinatorischer Probleme anwenden
– die Koeffizienten des Pascal'schen Dreieckes berechnen und im Zusammenhang mit dem binomischen Lehrsatz anwenden.

Kapitel 31

Kombinatorik

31.1 Permutation

Definition 11. Dinge *permutieren* heisst, sie in allen denkbaren Anordnungen zusammenstellen.

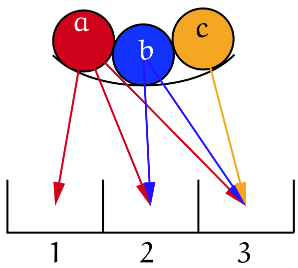


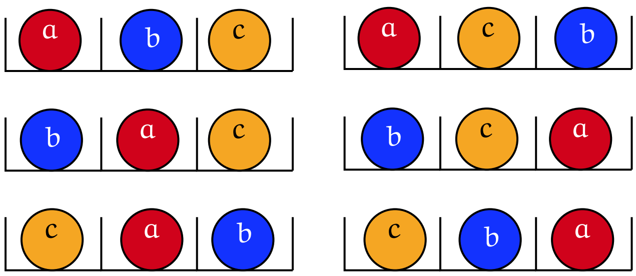
Das Resultat jeder einzelnen Zusammenstellung heisst Komplex. Zunächst nehmen wir an, dass die zu permutierenden Dinge von einander unterschieden werden können. Wir wähle der Einfachheit halber kleine Buchstaben, wie in der Abbildung ersichtlich.

Die Anzahl der durch Permutation von n Dingen entstehenden Komplexe heisse $P(n)$. Dann ist $P(2) = 2$, weil a und b keine anderen Komplexe liefern, als (a, b) und (b, a). Hieraus folgt aber, dass $P(3)$ dreimal so gross ist, als $P(2)$, weil erstens auf a die zwei Komplexe der Buchstaben b und c folgen können, zweitens auf b die zwei Komplexe der Buchstaben a und c, drittens auf c die zwei Komplexe der Buchstaben a und b, so dass

$$P(3) = P(2) + P(2) + P(2) = 3 \cdot P(2) = 3 \cdot 2 = 6$$

Man kann sich auch vorstellen, wie man n Sachen auf n Plätze verteilt. Für die erste Sache hat man n Möglichkeiten, für die zweite Sache nur noch $n - 1$ Plätze und so fort.





Ebenso ergibt sich:

$$
\begin{aligned}
P(4) &= 4 \cdot P(3) = 4 \cdot 3 \cdot 2 \cdot 1, \\
P(5) &= 5 \cdot P(4) = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1, \\
&\vdots \\
P(n) &= n \cdot P(n-1) = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1.
\end{aligned}
$$

Satz 31.1. Die Anzahl aller durch Permutation von n Dingen entstehenden Komplexe ist gleich dem Produkte aller natürlichen Zahlen von 1 bis n.

Definition 12. Das Produkt aller natürlichen Zahlen von 1 bis n schreibt man

$$n!$$

gelesen: "n Fakultät".

Wichtig 3. ☠ Für $n = 0$ ist festgesetzt: $0! = 1$, gleich wie $1! = 1$. ⊣

Damit kann man für $P(n)$ schreiben:

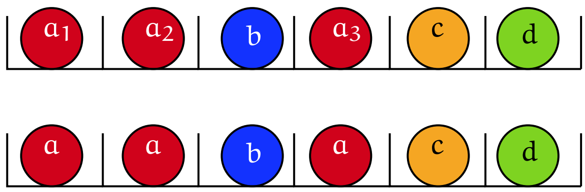
Formel 31.2. Permutationen

$$P(n) = n!$$

31.3 Übung Wir bestimmen die ersten Fakultäten von 1 bis 9. Durch rekursives Multiplizieren ergibt sich: $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$, $7! = 5040$, $8! = 40320$ und $9! = 362880$. Wie man sieht, nehmen die Zahlen fast explosionsartig zu. Das Rechnen mit grossen Fakultäten ist schnell mühsam. Ein Umweg ist der Gebrauch von Logarithmen, denn die Multiplikation verwandelt sich in eine Addition. Es gilt $\log(a \cdot b) = \log(a) + \log(b)$ etc. Somit ist $\log(n!) = \log(1) + \log(2) + \ldots + \log(n)$. ◁

Wir nehmen nun an, dass mehrere Kugeln dieselbe Farbe haben. Sie sind mit a_1, a_2 und a_3 angeschrieben. Im abgebildeten Beispiel sind es 6 Sachen, die zu $P(6) = 720$ Anordnungen führen. Nun nehme wir die roten Kugeln von ihren Plätzen und machen den Index weg, so dass die roten Kugeln alle mit a beschriftet sind. Wie viele Möglichkeiten gibt es, diese 3 Sachen auf 3 Plätze zu verteilen? Antwort. $P(3)$. Die zweite Reihe de Abbildung hat also insgesamt $P(6) \div P(3)$ verschieden Anordnungen. Das sind 120.

Alternativ hätten wir uns vorstellen können, die 3 nicht-roten Kugeln zuerst zu verteilen. Für die erste hätte man 6 Plätze zur Verfügung, für die zweit dann 5 und für die dritte noch 4. Die restlichen freien Plätze sind für die roten Kugeln; man hat keine Wahl mehr. Damit hat man $6 \cdot 5 \cdot 4$ Möglichkeiten, also 120.



Wir können verallgemeinern und sagen, dass n Dinge da sind, die zu permutieren sind, a von ihnen ununterscheidbar sind, so entstellen durch Permutation nicht $n!$, sondern $\frac{n!}{a!}$. Wenn nun unter den n Dingen noch eine zweite Gruppe von b Dingen sich befindet, die auch unter sich ununterscheidbar sind, so werden unter den entstehenden Anordnungen je $b!$ genau übereinstimmend, so dass man die soeben gefundene Zahl noch durch $b!$ zu dividieren hat, um die wirklich verschiedenen Anordnungen allein zu erhalten.

Satz 31.4. Unter den n Dingen befinden sich jeweils n_1, $n_2 \dots n_k$, einander gleiche. Es gibt dann

$$\frac{P(n)}{P(n_1) \cdot P(n_2) \cdot \ldots \cdot P(n_k)}$$

verschiedene Anordnungen für die $n_1 + n_2 + \ldots + n_k \leqslant n$.

31.5 Übung Wie viel mögliche Anordnungen gibt es für die Buchstaben aaabbc? Antwort: $\frac{P(6)}{P(3)P(2)} = \frac{6 \cdot 5 \cdot 4}{2} = 60$. Und wie viele Möglichkeiten gibt es für die angegebene Platzierung? Antwort: $3!2!1! = 12$. ◁

31.6 Übung Die Anordnung aaaabbb hat wie viele Möglichkeiten? Es sind genau 35. Denn $\frac{7!}{4!3!} = \frac{7 \cdot 6 \cdot 5}{6} = 7 \cdot 5$. ◁

31.7 Übung 3) Die Buchstaben aabbccdd ergeben 2520 Anordnungen. Wir rechnen $8! \div (2!)^4 = 40320 \div 16 = 2520$. ◁

31.8 Übung In einem Fussballspiel erzielt eine Mannschaft 5 Tore. Wie viele Halbzeitresultate sind möglich? Man veranschaue sich das Problem als 11/111 mit einem Trennstrich, der an jeder Position stehen kann, also auch als erster oder letzter. Damit folgt $M = \frac{6!}{1!5!} = 6$. Bei einem Eishockey-Spiel mit zwei Unterbrüchen und 5 Toren: 1//1111 sind es $\frac{7!}{5!2!} = 21$. ◁

31.2 Variation

Die Variation unterscheidet sich von der Permutation dadurch, dass man n Dinge nun auf nur k Plätze verteilen kann. Man kann etwa an das Gesellschaftsspiel Reise nach Jerusalem (auf engl. musical chairs) denken, wo ein Stuhl weniger als Tanzende vorhanden ist.

31.2.1 Ohne Wiederholung

Diese Aufgabe haben wir eigentlich schon als Permutation gelöst. Die Grundaufgabe ist, aus n verschiedenen Dingen deren k auf k Plätze zu verteilen. Dabei ist k kleiner (oder gleich) n. Das erste Ding (oder Element) hat n Möglichkeiten, das zweite $n-1$ usf. bis man die k Plätze belegt hat. Das sind $n! \div (n-k)!$.
Als Beispiel nehmen wir die Buchstaben a, b, c, d und e und platzieren sie nach allen Möglichkeiten auf 3 Plätze. Die Auswahl des ersten Buchstabens hat 5 Möglichkeiten, der zweite Buchstabe noch 4 und der dritte dann 3. Also insgesamt ergibt dies $5 \cdot 4 \cdot 3$, also 60 Anordnungen. Hier gehen wir von der Anzahl Dinge aus.

Definition 13. Variation n Dinge zur kten Ordnung ohne Wiederholung *variieren* heisst, je k unter den n Dingen in allen möglichen Anordnungen zusammenstellen, aber so, dass kein Ding öfter als einmal vorkommt.

Satz 31.1. Die Variation von n Dingen zur k-ten Ordnung ohne Wiederholung hat

$$V(n,k) = \frac{n!}{(n-k)!}$$

Anordnungen.

Anmerkung 31.2. Falls $k = n$, also gleichviel Dinge wie Plätze vorhanden sind, dann geht die Variation als Spezialfall in die Permutation über, denn $V(n,0) = P(n)$.

31.3 Übung Beim Pferdetoto muss in der sog. Dreierwette der Zieleinlauf der ersten drei Pferde in der richtigen Reihenfolge vorausgesagt werden. Wie viel verschiedene Dreier-Wetten sind möglich, wenn 10 Pferde starten?
Es handelt sich hier um Variation ohne Wiederholung von 10 Pferden zur 3-ten Ordnung. Also ergeben sich:

$$V(10,3) = \frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$$

Möglichkeiten. $\triangleleft$

31.4 Übung Aus den Ziffern 1 bis 9 sollen dreistellige Zahlen gebildet werden. Wie viele Möglichkeiten gibt es, wenn jede Ziffer nur einmal verwendet werden darf? Antwort: 9 Ziffern werden auf 3 Plätze verteilt, wobei jede Ziffer nur einmal verwendet werden darf und die Anordnung der Ziffern von Bedeutung ist. Es handelt sich also um Variationen 3. Ordnung ohne Wiederholung. Ihre Anzahl beträgt $V(9,3) = 9 \cdot 8 \cdot 7 = 504$. $\triangleleft$

31.2.2 Mit Wiederholung

Eine Wiederholung macht auch dann Sinn, wenn $k > n$ ist, also mehr Plätze als Dinge vorhanden sind. Man denke etwa an grosse Zahlen, die man mit 10 Ziffern bilden kann.

Definition 14. Variation mit Wiederholung n Dinge zur k-ten Ordnung mit Wiederholung variieren heisst, je k unter den Dingen in allen möglichen Anordnungen zusammenstellen, und zwar so, dass in einer Anordnung ein und dasselbe Ding beliebig oft, also 1 bis k mal, vorkommen darf.

Satz 31.5. Es gibt genau

$$V_w(n, k) = n^k$$

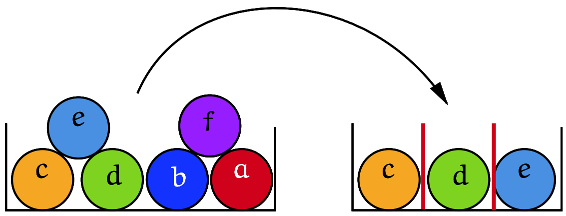
verschiedene Variationen k-ter Ordnung mit Wiederholung, wobei auch $k > n$ sein darf.

31.6 Übung Ein Affe hat eine Schreibmaschine mit den Buchstaben r, o, k, t und e. Wie viele Variationen gibt es für Wörter der Länge 13? Antwort: 5^{13}. Das sind über 1.2 Milliarden. Angenommen der Affe schreibt "Rokokokokotte". Auf wie viele Arten kann man das Wort schreiben? Es enthält 5 "o", 4 "k" und 2 "t". Es gibt $5!4!2! = 5760$ nicht unterscheidbare Arten. Die Wahrscheinlichkeit, dass er dieses Wort schreibt ist daher rund 0.000005. ◁

31.7 Übung Wir suchen siebenstellige Telefonnummern, die mindestens zwei gleichen Ziffern aufweisen. Wir erinnern die Definitionen und können die Anzahl Variationen mit Wiederholung und die ohne Wiederholung in Abzug bringen:

$$M = V_w(10, 7) - V(10, 7) = 10^7 - \frac{10!}{7!} = 10000000 - 720 = 9999280$$

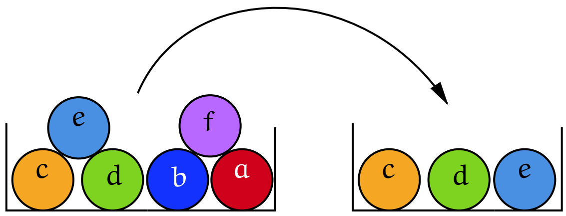
◁



Bei Permutation und Variation kommt es auf die Position an. Man zieht aus einer Urne und *platziert* die Kugel an einer bestimmten Stelle. Beim folgenden Modell der Kombination hingegen zieht man aus einer Urne und legt wieder in eine Urne ab, sodass es keinen bestimmten Platz gibt. In der Abbildung wird dieser Sachverhalt durch die roten Trennlinien symbolisiert. Die Anzahl Anordnungen mit Platzordnung sind viel mehr als solche ohne.

31.3 Kombination

Das Modell hier ist das Ziehen aus einem Gefäss und ablegen in einem anderen. Es gibt keine Platzordnung. Bis jetzt hatten wir es mit Permutationen und Variationen zu tun, bei denen die Reihenfolge wichtig war. Hier nicht mehr.



Da wir aber schon das Instrumentarium für geordnete Anordnungen haben, gehen wir von diesen aus und Vernichten dann die Ordnung.

31.3.1 Kombination ohne Wiederholung

In der Abbildung sehen wir sechs Dinge, von denen 3 ausgewählt werden. Wenn man zuerst annimmt, dass die Auswahl geordnet ist, dann ist es ein Variationsproblem mit der Lösung $V(6, 3)$. Die Anordnung wird aufgelöst, d.h. (c, d, e) ist gleich c, e, d usw. Wie viele Anordnungen dieser 3 Dinge gibt es? Es sind $P(3) = 3!$. Somit ist die Anzahl ungeordneter Anordnungen also $V(6, 3) \div P(3)$. Verallgemeinert folgt:

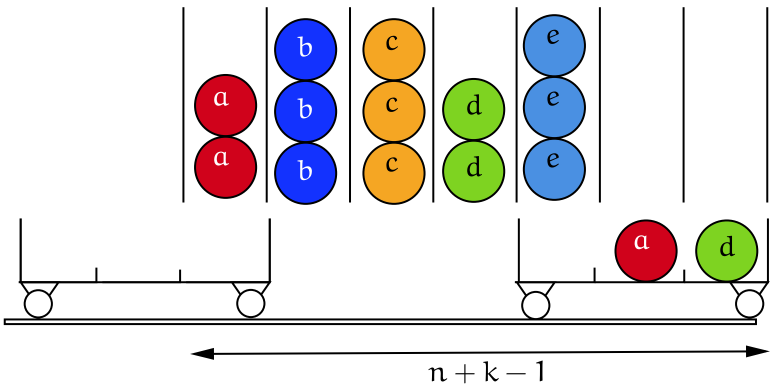
Satz 31.1. Es gibt genau
$$C(n,k)=\frac{V(n,k)}{P(k)}=\frac{n!}{(n-k)!k!}$$
Möglichkeiten, aus n Dingen k zu wählen ohne Wiederholung und ohne Berücksichtigung der Reihenfolge.

31.2 Übung Es dürfen aus einer Schulklasse von 24 Schülern deren 4 an einem Wettbewerb teilnehmen. Wie viele Auswahlmöglichkeiten hat die Lehrerin? Sie kann $(24\cdot 23\cdot 22\cdot 21)/4!$ (oder $\binom{24}{4}$) auswählen. Das sind 10626 Möglichkeiten. $\triangleleft$

31.3.2 Kombination mit Wiederholung

Wider Erwarten ist dies nun das schwierigste Problem. Zwei Aspekte sind zu berücksichtigen: Innerhalb einer Anordnung dürfen gleiche Elemente wiederholt vorkommen und verschiedene Anordnungen, die die gleichen Dinge enthalten, werden als gleich angesehen.

Wenn man mit einem Problem Schwierigkeiten hat, kann man versuchen, einen indirekte Weg zu suchen. Bekannt ist das Prinzip der Schäfer, wonach die Schäfer nicht direkt die Schafe zählen sondern die Beine und dann durch vier teilen. Im nächsten Abschnitt werden wir darauf zurückkommen.



Wir betrachten einen Mechanismus, der aus einem Wägelchen mit k Abteilungen besteht. Es fährt unter den n Silos mit den Elementen durch. Damit auch k-malige Wiederholungen möglich sind, hat jedes Silo k gleiche Elemente. Im konkreten Fall ist $n=5$ und $k=3$. Wieviele Positionen muss das Wägelchen abfahren können, damit es in jeder Abteilung ein beliebiges Element aufnehmen kann? Es muss soweit fahren, dass die letzte Abteilung unter dem letzten Silo zu stehen kommt. Das heisst, es muss $n+(k-1)$ Positionen abdecken.

Satz 31.3. Es gibt genau
$$C_w(n,k)=C(n+k-1,k)=\frac{(n+k-1)!}{(n-1)!k!}$$
Möglichkeiten, aus n Dingen k zu wählen *mit* Wiederholung und ohne Berücksichtigung der Reihenfolge.

Berechnen wir nun $C_w(5,3)$. Es folgt $\frac{7!}{4!3!}=\frac{7\cdot 6\cdot 5}{6}=35$.

31.4 Urnenmodell und Zusammenfassung

Anhand des Urnenmodells mit wollen wir die vier Situationen nochmals veranschaulichen. Das Beispiel ist möglichst einfach: Wir ziehen 2 Kugeln von 3 unterscheidbaren aus der Urne:

Mit Reihenfolge, mit Wiederholung		
$(1,1)$	$(1,2)$	$(1,3)$
$(2,1)$	$(2,2)$	$(2,3)$
$(3,1)$	$(3,2)$	$(3,3)$
$V_w(3,2) = 3^2 = 9$		

Mit Reihenfolge, ohne Wiederholung		
	$(1,2)$	$(1,3)$
$(2,1)$		$(2,3)$
$(3,1)$	$(3,2)$	
$V(3,2) = 6$		

Ohne Reihenfolge, mit Wiederholung		
$(1,1)$	$(1,2)$	$(1,3)$
	$(2,2)$	$(2,3)$
		$(3,3)$
$C_w(n,k) = 6$		

Ohne Reihenfolge, ohne Wiederholung		
–	$(1,2)$	$(1,3)$
	–	$(2,3)$
		–
$C(n,k) = 3$		

Wie man erkennt, bedeutet ohne Wiederholung, dass die Diagonalemente wegfallen. Ist die Reihenfolge unerheblich, fallen alle unteren Ausserdiagonalelemente weg.
Die Permutation erhalten wir, wenn $n = k$ und sowohl mit Reihenfolge als auch mit Wiederholung.

Wichtig 4. Bei schwierigen Zählproblemen ist es meist möglich, Teile der Menge mit dem Urnenmodell zu berechnen und dann das Resultat zusammenzustellen. ⊣

31.5 Binomischer Lehrsatz

31.5.1 Binomialkoeffizient

Bei der Kombination ohne Wiederholung sind wir auf den Term

$$\frac{n!}{(n-k)!k!}$$

gestossen. Dafür gilt folgende Abkürzung:

Definition 15. Wir schreiben

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

und sagen "n tief k" oder "n über k" oder "k aus n".

Anmerkung 31.1. Auf dem Taschenrechner findet man oft eine Taste "nCr" für " n choose r".

Auch Binomialkoeffizienten befolgen Rechenregeln. Aus der Definition ist ersichtlich, dass $\binom{n}{k} = \binom{n}{n-k}$. Zusammenfassend:

Eigenschaften 31.2. Es gelten folgende Regeln:

$$\binom{n}{k} = \binom{n}{n-k}$$

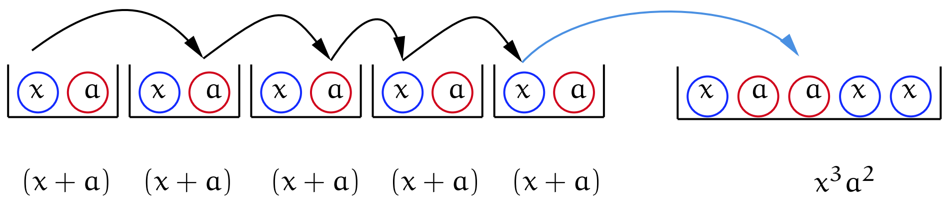
$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Die zweite Formel folgt aus der Erweiterung der Brüche

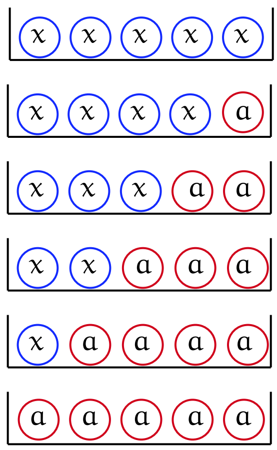
$$\begin{aligned}\frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} &= \frac{n!(k+1)}{(k+1)!(n-k)!} + \frac{n!(n-k)}{(k+1)!(n-k)!} \\ &= \frac{n![(k+1)+(n-k)]}{(k+1)!(n-k)!} = \frac{(n+1)!}{(k+1)!(n-k)!} \\ &= \binom{n+1}{k+1}\end{aligned}$$

31.5.2 Lehrsatz

Wir kennen aus der Algebra Terme wie $(x+a)^2 = x^2 + 2ax + a^2$. Die Koeffizienten, hier 1,2,1 ergeben sich aus dem Ausmultiplizieren. Sie sind durch das Pascal'sche Dreieck gegeben. Hier wollen wir eine Begründung für diese Zahlen angeben. Dazu verwenden wir ein Modell wie in der Abbildung



Wie man sieht, kann man den Ausdruck x^3a^2 mit x und a aus verschiedenen Urnen ziehen. Wir haben das Grundmodell der Kombination vor uns: Ziehe k aus n ohne Berücksichtigung der Reihenfolge. Das Resultat ist bekanntlich $C(n,k) = \binom{n}{k}$.
Am Beispiel von $(x+a)^5$ spielen wir die Anordnungen durch. Es ist einsichtig, dass es am Rand (alles x oder alles a) nur eine Möglichkeit gibt, gegen die gemischten Terme die möglichen Anordnungen anwachsen und wieder zurückgehen bis auf 1. Die Koeffizienten, d.h. Anzahl Kombinationen, sind symmetrisch.
Der Satz lautet also



$\binom{5}{0} = 1$

$\binom{5}{1} = 5$

$\binom{5}{2} = 10$

$\binom{5}{3} = 10$

$\binom{5}{4} = 5$

$\binom{5}{5} = 1$

Satz 31.3. Binomischer Lehrsatz

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

31.4 Übung Man bestimme $(3x+2y)^4$. Wir substituieren $a = 3x$ und $b = 2y$. Damit erhalten wir $(a+b)^4$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Mit zurücksubstituieren

$$(3x+2y)^4 = (3x)^4 + 4(3x)^3(2y) + 6(3x)^2(2y)^2 + 4(3x)(2y)^3 + (2y)^4$$

und

$$(3x+2y)^4 = 81x^4 + 216x^3y + 216x^2y^2 + 94xy^3 + 16y^4$$

$\triangleleft$

31.5 Übung Wir berechnen das 4. Glied von $(2a+3b)^{10}$.
Achtung, die Nummerierung beginnt mit $k=0$. Das vierte Glied ist $k=3$. Angenommen wir hätten $(c+d)^{10}$, dann wäre das 4. Glied

$$\binom{10}{3}c^7d^3 \quad \text{rückeingesetzt} \quad \binom{10}{3}(2a)^7(3b)^3$$

und somit

$$120 \cdot 2^7 \cdot 3^3 a^7 b^3 = 120 \cdot 128 \cdot 9 a^7 b^3 = 138240 a^7 b^3$$

$\triangleleft$

31.6 Indirektes Zählen, Ein- und Ausschalten**

Wir haben uns in diesem Kapitel vor allem um Zähltechniken gekümmert. Die Frage war durchgehend “wie viele?” Eine weitere Zählmethode stellt das Prinzip von Ein- und Ausschalten dar. Sie zählt die Anzahl Dinge, die mindestens eine von mehreren Eigenschaft aufweisen und zählen die Dinge wieder ab, die mehr als eine Eigenschaft aufweisen, damit sie nicht mehrfach berücksichtigt werden.

31.1 Übung An einer Party mit 6 Personen soll angestossen werden, Wie viel Mal erklingen die Gläser? Jeder stösst mit jedem ausser sich selber an. Insgesamt ergibt dies $6 \cdot 5$ Mal. Nun sind aber alle Stösse doppelt gezählt. Wir müssen die Zahl durch 2 dividieren und erhalten 15. $\triangleleft$

31.2 Übung Es soll ein Tischtennisturnier ausgetragen werden mit 6 Teilnehmern. Wie viele Partien sind anzusetzen? Das ist das gleiche Problem wie oben, also 15 Spiele. Man kann aber auch folgende Überlegung anstellten. Wenn zwei spielen, gibt es ein Spiel. Wenn einer dazu kommt, muss er gegen 2 spielen, also $1+2=3$. Es kommt ein vierter, der gegen 3 spielt, also $3+3$. Es kommt noch einer, der gegen vier spielt, $6+4$. Nun kommt der sechste, der gegen fünf spielt, also $10+5=15$. Dies ist die rekursive Variante. Und mit einem siebten folgt dann $15+61=21$. $\triangleleft$

31.3 Übung Gesucht ist die Anzahl Personen, die *mindestens* eine Katze oder einen Hund besitzen. Nimmt man die Anzahl Personen, die einen Hund halten und addiert die Anzahl Personen, die eine Katze besitzen, dann subtrahiert man die Anzahl Personen, die beides haben.

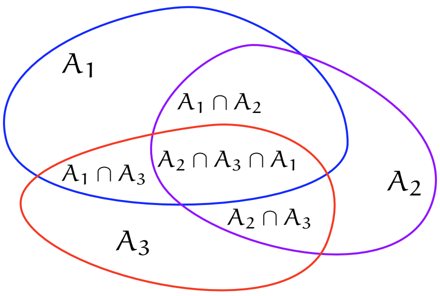
Wir erinnern uns an die Mächtigkeit von Mengen. Die zwei Mengen Hundebesitzer H und Katzenhalter K sind nicht elementfremd, haben also eine Schnittmenge $H \cap K \neq \varnothing$. Die Mächtigkeit der Tierbesitzer $|H \cup K|$ ist deshalb $|H| + |K| - |H \cap K|$. $\triangleleft$

Wir verallgemeinern für drei Mengen, etwas Hunde-, Katzen- und Hamsterbesitzer. Wir suchen $|A_1 \cup A_2 \cup A_3|$ und formen um, ausgehend von $|A_i \cup A_j| = |A_i| + |A_j| - |A_i \cap A_j|$

$$\begin{aligned}
&|A_1 \cup A_2 \cup A_3| \\
&= |(A_1 \cup A_2) \cup A_3| \\
&= |A_1 \cup A_2| + |A_3| - |(A_1 \cup A_2) \cap A_3| \\
&= |A_1 \cup A_2| + |A_3| - |(A_1 \cap A_3) \cup (A_2 \cap A_3)| \\
&= |A_1| + |A_2| - |A_1 \cap A_2| + |A_3| - |(A_1 \cap A_3) \cup (A_2 \cap A_3)| \\
&= |A_1| + |A_2| - |A_1 \cap A_2| + |A_3| - |A_1 \cap A_3| - |A_2 \cap A_3| + |(A_1 \cap A_3) \cap (A_2 \cap A_3)| \\
&= |A_1| + |A_2| - |A_1 \cap A_2| + |A_3| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_3 \cap A_2| \\
&= |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_3 \cap A_2|
\end{aligned}$$

oder kürzer

$$|A_1 \cup A_2 \cup A_3| = \sum_{j=1}^{3} |A_j| - \sum_{j_1<j_2} |A_{j_1} \cap A_{j_2}| + |A_1 \cap A_2 \cap A_3|$$



Man summiert alle Besitzer eines Tieres und subtrahiert alle Besitzer von zwei Tieren, worunter die mit dreien doppelt abgezählt wurden. Diese fügt man wieder hinzu. In der Abbildung erkennt man gut die Situation. Die Schnittmenge von zwei Mengen ist doppelt gezählt, die von drei Mengen dreimal.

Die Erweiterung auf $n = 4$ ist sinngemäss:

$$|A_1 \cup A_2 \cup A_3 \cup A_4| = \sum_{j=1}^{4} |A_j| - \sum_{j_1<j_2} |A_{j_1} \cap A_{j_2}| + \sum_{j_1<j_2<j_3} |A_{j_1} \cap A_{j_2} \cap A_{j_3}| - |A_1 \cap A_2 \cap A_3 \cap A_4|$$

Jetzt wird auch klar, wieso man von Aus- und Einschalten spricht. Es sind die Vorzeichen der Summanden, die von $-$ zu $+$ alternieren.

Dérangements

Ein Klasse von Fragestellungen heisst *Dérangement-Problem*. Die obigen Mengenschreibweise kann man hier vereinfachen, weil alle Mengen A_i gleichmächtig sind. Generell gehören je zwei Objekte zueinander, z.B. Tanzpartner, Hutbesitzer, Briefe und Umschläge und ähnliches. Die Anzahl Dérangements ist die Gesamtzahl von Anordnungen ohne Anordnungen, die korrekte Elemente enthalten, also mit der Grundmenge G interessiert $G \setminus (A_1 \cup A_2 \cup \ldots)$ oder $(\overline{A_1 \cup A_2 \cup \ldots})$.
Die Anzahl Dérangements in einer Menge der Mächtigkeit n, wo also kein Ding am richtigen Platz ist, beträgt somit wie folgt.

Satz 31.4. Die Anzahl Dérangements in einer Menge mit Mächtigkeit n ist:

$$D_n = n! \left\{ 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} \right\}$$

31.5 Übung Es liegen vier Briefe a, b, c, d vor und vier Umschläge A, B, C, D. Wie viele Möglichkeiten gibt es, dass alle Briefe falsch verpackt sind? Die Anzahl möglicher Verpackungen ist $P(4) = 4! = 24$. Das ist die Mächtigkeit der Grundmenge G. Wir betrachten die 4 Teilmengen, bei denen ein Brief i falsch ist. Davon gibt es pro i $P(3)$.
Für die Formel $|A_1 \cup A_2 \cup A_3 \cup A_4|$ bestimmen wir:

- $|A_i| = 3!$ für Anordnungen, die mindestens eine richtige Zuordnung habe, für alle i,
- $|A_{j_1} \cap A_{j_2}| = 2!$ für $j_1 < j_2$,
- $|A_{j_1} \cap A_{j_2} \cap A_{j_3}| = 1!$ für $j_1 < j_2 < j_3$ und
- $|A_1 \cap A_2 \cap A_3 \cap A_4| = 0! = 1$.

Die Anzahl Teilmengen sind der Reihe nach $\binom{4}{1}$, $\binom{4}{2}$, $\binom{4}{3}$ und $\binom{4}{4}$. Damit und dem Vorzeichengebrauch folgt:

$$\begin{aligned} D &= 24 - \left\{ \binom{4}{1} 3! + \binom{4}{2} 2! - \binom{4}{3} 1! + \binom{4}{4} 1 \right\} \\ &= 24 - 24 + 12 - 4 + 1 \\ &= 9 \end{aligned}$$

Wir überprüfen mit der Formel:

$$D_n = 4! \cdot \left\{ 1 - \frac{1}{1} + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right\}$$

ergibt $24 - 24 + 12 - 4 + 1 = 9$.
Anhand der Zusammenstellung aller Permutationen kann man an der Anzahl Striche die einzelnen Teilmengen festmachen. Es hat 9 Zusammenstellungen ohne Strich.

$$\begin{array}{llll}
\cancel{A}\cancel{B}\cancel{C}\cancel{D} & BA\cancel{C}D & CAB\cancel{D} & DABC \\
\cancel{A}\cancel{B}DC & BADC & CADB & DA\cancel{C}B \\
\cancel{A}CBD & BCA\cancel{D} & C\cancel{B}A\cancel{D} & D\cancel{B}AC \\
\cancel{A}CDB & BCDA & C\cancel{B}DA & D\cancel{B}\cancel{C}A \\
\cancel{A}DBC & BDAC & CDAB & DCAB \\
\cancel{A}DCB & BD\cancel{C}A & CDBA & DCBA
\end{array}$$

◁

31.6 Übung Sechs Ehepaare nehmen an einem Tanzkurs teil. Alle sind gleichzeitig auf der Tanzfläche. Wie viele Paare sind möglich, bei denen kein einziges ein Ehepaar ist? Es ist laut Formel

$$D_6 = 6!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!}) = 720(\frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720}) = 265$$

Insgesamt sind $M = 6! = 720$ Paare möglich. ◁

Rencontre

Rencontre oder Treize ist ein Kartenspiel mit folgender Regel: Zwei Spieler A und B besitzen jeweils ein vollständiges Kartenspiel mit M Karten. Sie mischen ihre Karten und legen diese als Stapel vor sich ab. Nun ziehen beide Spieler gleichzeitig immer wieder die oberste

Karte von ihrem Stapel. Erscheint zu irgendeinem Zeitpunkt zweimal die gleiche Karte, so gewinnt der eine Spieler, andernfalls der andere. Es stellt sich die Frage, wie gross ist die Gewinnwahrscheinlichkeit von den Spielern.
Man bemerke, dass man sich die Tanzpaare von obigem Beispiel auch als zwei Kartenspiele von Männern und Frauen vorstellen kann mit $M = 6$.
Leonhard Euler hat dieses Problem als Rekursion gelöst. Wir betrachten den Fall $M = 3$. Wir stellen uns vor, dass die Karten fortlaufend durchnumeriert sind und das erste Kartenspiel aufsteigend geordnet ist. Damit haben wir nur die Anordnung vereinfacht aber keine Einschränkung der Allgemeinheit erwirkt. Wir tabellieren die Möglichkeiten für 1,2 und 3 Karten.

	B
A	1
1	$\boxed{1}$

	B	
A	1	**2**
1	$\boxed{1}$	2
2	$\boxed{\cancel{2}}$	1

	B					
A	1	2	3	**4**	**5**	6
1	$\boxed{1}$	$\boxed{1}$	2	2	3	3
2	$\boxed{\cancel{2}}$	$\cancel{3}$	1	3	1	$\boxed{2}$
3	$\boxed{\cancel{3}}$	$\cancel{2}$	$\boxed{3}$	1	2	$\cancel{1}$

Pro Zug (Zeilen) gibt es $M!$ Permutationen, hier also 1, 2 und 6. Wenn A gewinnt, wir die Position eingerahmt. Die dem Gewinn nachfolgenden Karten interessieren nicht mehr, weil das Spiel schon aus ist. Deshalb sind diese Positionen durchgestrichen. Die zwei ersten Kolonnen der dritten Tabelle beschreiben den Gewinn von A im ersten Zug, Kolonne 6 Gewinn im zweiten und Kolonne 3 im dritten Zug. In den Kolonnen ohne Einrahmung gewinnt B, als zwei Mal. Wir tabellieren die Verhältnisse für $M = 4$ mit $4! = 24$ Permutationen:

Tabelle 31.1

A	B																							
1	**1**	**1**	**1**	**1**	**1**	**1**	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4
2	$\cancel{2}$	$\cancel{2}$	$\cancel{3}$	$\cancel{3}$	$\cancel{4}$	$\cancel{4}$	1	1	3	3	4	4	1	1	**2**	**2**	4	4	1	1	**2**	**2**	3	3
3	$\cancel{3}$	$\cancel{4}$	$\cancel{2}$	$\cancel{4}$	$\cancel{2}$	$\cancel{3}$	**3**	4	1	4	1	**3**	2	4	$\cancel{1}$	$\cancel{4}$	1	2	2	**3**	$\cancel{1}$	$\cancel{3}$	1	2
4	$\cancel{4}$	$\cancel{3}$	$\cancel{4}$	$\cancel{2}$	$\cancel{3}$	$\cancel{2}$	$\cancel{4}$	3	**4**	1	3	$\cancel{1}$	**4**	2	$\cancel{4}$	$\cancel{1}$	2	1	3	$\cancel{2}$	$\cancel{3}$	$\cancel{1}$	2	1

Aus Platzgründen haben wir anstatt Rahmen fette Schrift genommen. Wir sehen hier wiederum M Gewinne im ersten Zug bei M Karten mit 6! Permutationen. Im zweiten Zug dann 4, im dritten 3 und im vierten Zug dann 2. Wir machen eine Tabelle mit den Gewinnen

Tabelle 31.2

		Anzahl Karten M			
	k	1	2	3	4
	1	1	1	2	6
Gewinn	2		0	1	4
in k	3			1	3
	4				2

Euler macht folgenden Schritt: Aus der Tabelle 31.1 betrachtet er den Gewinn im k-ten Zug, indem er den Ausschnitt bestimmt, für den B ein k aufweist. Mit $M = 4$ gibt es dann $(4-1)! = 6$ Kolonnen. Spieler A gewinnt also im k-ten Zug jeweils 6 minus die Anzahl der Fälle, bei denen er schon vor k gewonnen hat. Für $k = 6$ gibt es 6 Gewinne, bei $k = 2$ dann 6-2=4, bei $k = 3$ 6-3 und für $k = 4$ 6-4.

$k=1$	B					
A	**1**	**2**	**3**	**4**	**5**	**6**
2	$\cancel{2}$	$\cancel{2}$	$\cancel{3}$	$\cancel{3}$	$\cancel{4}$	$\cancel{4}$
3	$\cancel{3}$	$\cancel{4}$	$\cancel{2}$	$\cancel{4}$	$\cancel{2}$	$\cancel{3}$
4	$\cancel{4}$	$\cancel{3}$	$\cancel{4}$	$\cancel{2}$	$\cancel{3}$	$\cancel{2}$

$k=2$	B					
A	1	2	**3**	**4**	**5**	**6**
1	$\boxed{1}$	$\boxed{1}$	3	3	4	4
$\cancel{3}$	$\cancel{3}$	$\cancel{4}$	$\cancel{1}$	$\cancel{4}$	$\cancel{1}$	$\cancel{3}$
$\cancel{4}$	$\cancel{4}$	$\cancel{3}$	$\cancel{4}$	$\cancel{1}$	$\cancel{3}$	$\cancel{1}$

$k=3$	B					
A	1	2	**3**	**4**	**5**	6
1	$\boxed{1}$	$\boxed{1}$	2	2	4	4
2	$\cancel{2}$	$\cancel{4}$	1	4	1	$\boxed{2}$
4	$\cancel{4}$	$\cancel{2}$	$\cancel{4}$	$\cancel{1}$	$\cancel{2}$	$\cancel{1}$

$k=4$	B					
A	1	2	3	**4**	**5**	6
1	$\boxed{1}$	$\boxed{1}$	2	2	3	3
2	$\cancel{2}$	$\cancel{3}$	1	3	1	$\boxed{2}$
3	$\cancel{3}$	$\cancel{2}$	$\boxed{3}$	1	2	$\cancel{1}$

Man bemerkt, dass die Tabelle (ohne Markierungen) identisch ist mit derjenigen von $M = 3$, sofern man sich die 4 durch die 3 ersetzt denkt. Hier will man zeigen, dass man für das Spiel mit 4 Karten auf (mehrere) Tabellen für drei Karten zurückgreifen kann.
Tabelle 31.2 fasst auch die vier obigen Tableaus zusammen. Der Zusammenhang ist wie folgt: Im ersten Zug kann man auf 6 Arten gewinnen. Im zweiten Zug ebenfalls 6 Mal unter Abzug derjenigen von diesen 6, die schon im ersten Zug gewonnen hätten. Im dritten Zug zieht man von 6 wiederum die Gewinne im ersten Zug ab und die vom zweiten Zug, die nicht auch schon im ersten gewonnen wurden. Bei $M = 4$ ergeben sich $(M-1)! = 6$ Gewinnchancen.
Wir bezeichnen mit $g_{m,k}$ die Anzahl Gewinne im k-ten Zug bei $m = n+1$ Karten. Es folgt

$$\begin{aligned}
g_{n+1,1} &= n! \\
g_{n+1,2} &= g_{n+1,1} - g_{n,1} \\
g_{n+1,3} &= g_{n+1,2} - g_{n,2} \\
&\dots \\
g_{n+1,k} &= g_{n+1,k-1} - g_{n,k-1}
\end{aligned}$$

Nun gehen wir von der Anzahl Gewinnsituationen zur Wahrscheinlichkeit über. Dazu setzen wir $P_{m,k}$ für die Wahrscheinlichkeit, dass der erste Spieler im k-ten Spiel mit m Karten gewinnt. Es gilt:

$$P_{m,1} = \frac{g_{m,1}}{m!} = \frac{(m-1)!}{m!} = \frac{1}{m}$$

Nun weiter

$$\begin{aligned}
P_{m,2} &= \frac{g_{m,1} - g_{m-1,1}}{m!} = \frac{(m-1)! - (m-2)!}{m!} \\
&= \frac{(m-2)!(m-1-1)}{m!} = \frac{(m-1)-1}{m(m-1)} \\
&= \frac{1}{m} - \frac{1}{m(m-1)}
\end{aligned}$$

Weiter

$$\begin{aligned}
P_{m,3} &= \frac{g_{m,2} - g_{m-1,2}}{m!} = \frac{(m-1)! - (m-2)! - [(m-2)! - (m-3)!]}{m!} \\
&= \frac{(m-3)![(m-1)(m-2) - 2(m-2) + 1]}{m!} \\
&= \frac{[(m-1)(m-2) - 2(m-2) + 1]}{m(m-1)(m-2)} \\
&= \frac{1}{m} - \frac{2}{m(m-1)} + \frac{1}{m(m-1)(m-2)}
\end{aligned}$$

und, ohne Rechnung

$$P_{m,4} = \frac{1}{m} - \frac{3}{m(m-1)} + \frac{3}{m(m-2)(m-2)} - \frac{1}{m(m-1)(m-2)(m-3)}$$

Und schön untereinander

$$\begin{aligned}
P_{m,1} &= \frac{1}{m} \\
P_{m,2} &= \frac{1}{m} - \frac{1}{m(m-1)} \\
P_{m,3} &= \frac{1}{m} - \frac{2}{m(m-1)} + \frac{1}{m(m-1)(m-2)} \\
P_{m,4} &= \frac{1}{m} - \frac{3}{m(m-1)} + \frac{3}{m(m-2)(m-2)} - \frac{1}{m(m-1)(m-2)(m-3)}
\end{aligned}$$

Die Koeffizienten der Terme sind die Binome aus dem Pascal'schen Dreieck, das wir in dieser häufigsten Darstellung kennen.

$$\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & \swarrow & & \searrow & & \\
 & & 1 & & 1 & & \\
 & \swarrow & \searrow & & \swarrow & \searrow & \\
 & 1 & & 2 & & 1 & \\
\swarrow & \searrow & \swarrow & & \searrow & \swarrow & \searrow \\
1 & & 3 & & 3 & & 1
\end{array}$$

Nun kann man das Dreieck auch scheren und wie folgt wiedergeben:

m	$\binom{m-1}{0}$	$\binom{m-1}{1}$	$\binom{m-1}{2}$	$\binom{m-1}{3}$	$\binom{m-1}{4}$
1	1				
2	1	1			
3	1	2	1		
4	1	3	3	1	
5	1	4	6	4	1
6	1	5	10	10	5
7	1	6	15	20	15
8	1	7	21	35	35

Man beachte, dass die Summe einer Kolonne, hier z.B. die dritte in blau, der vertikal nächsten, horizontal um eine Position verschobenen Zahl entspricht! Das gilt für jede Kolonne. Diese Tatsache ist sehr hilfreich in den nächsten Überlegungen.

Die gesamte Wahrscheinlichkeit für den Spielgewinn ist die Summe. Wir summieren als erstes die erste Kolonne, die sich einfach zu $S_1 = m\frac{1}{m} = 1$ ergibt. Die zweite Kolonne ist

$$S_2 = -\frac{\sum_{k=2}^{m} \binom{k-1}{1}}{m(m-1)} = -\frac{\binom{m}{2}}{m(m-1)} = -\frac{1}{1 \cdot 2} = -\frac{1}{2!}$$

Weiters gilt nach derselben Logik

$$S_3 = -\frac{\sum_{k=3}^{m} \binom{k-1}{2}}{m(m-1)(m-2)} = \frac{\binom{m}{3}}{m(m-1)(m-2)} = \frac{1}{1 \cdot 2 \cdot 3} = \frac{1}{3!}$$

und

$$S_4 = -\frac{1}{4!}$$

Alles zusammen ergibt

$$P_m = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + (-1)^m \frac{1}{m!}$$

Ein Vergleich mit der Formel für die Anzahl Dérangements D_m gemäss Satz 31.4 zeigt, dass

$$P_m = 1 - \frac{D_m}{m!}$$

Der zweite Term ist die Wahrscheinlichkeit, dass B gewinnt, da er siegt, wenn keine Karten übereinstimmen. Es gilt folgender Grenzwert

$$\lim_{m \to \infty} \frac{D_m}{m!} = \frac{1}{e}$$

und deshalb

$$\lim_{m \to \infty} P_m = 1 - \frac{1}{e} \approx 0.632$$

31.7 Übung Nochmals ein Beispiel: Zehn Männer gingen, als man noch Hüte trug, ins Restaurant. Beim verlassen gingen die Lichter aus und jeder nahm einen Hut. Wie gross ist die Wahrscheinlichkeit, dass mindestens einer seinen Hut genommen hatte? Einfach:

$$P_{10} = 1 - \frac{D_m}{m!} = 0.632$$

◁

Aufgaben

0.8 Wie viele verschiedene Zusammenstellungen von genau 5 Buchstaben können aus den 26 Buchstaben des Alphabets gebildet werden, wenn Wiederholungen zulässig bzw. nicht zulässig sind?

8 Mit Wiederholung: Für erste Wahl 26 Möglichkeiten, dann wieder 26 etc. Also total $M = 26^5 = 11881376$.
Ohne Wiederholung: Für erste 26, dann 25, dann 24 etc. Somit $M = 26 \cdot 25 \cdot 24 \ldots = \frac{26!}{21!} = 789360$.

0.9 Wieviel geordnete Paare je zweier verschiedener Buchstaben kann man aus den 26 Buchstaben des Alphabets bilden?

9 Es ist eine Variationsaufgabe ohne Wiederholung, denn "geordnete Paare" und "verschiedene". Es gilt deshalb $V(n,k) = V(26,2) = \frac{n!}{(n-k)} = \frac{26!}{24!} = 650$.

0.10 Gegeben seien die neuen Ziffern 1, 2, 3, 4, 5, 6, 7, 8, 9.

(a) Wie viele dreistellige Zahlen können daraus gebildet werden, wenn jede Ziffer höchstens einmal vorkommen darf?

(b) Wie viele der so gebildeten Zahlen sind gerade, wie viele ungerade?

(c) Wie viele dieser Zahlen sind durch 5 teilbar?

(d) Wie viele dieser Zahlen sind kleiner als 200 bzw. größer als 500?

10 **(a)** Man wählt 3 aus 9 aus: $M = 9 \cdot 8 \cdot 7 = 504$

(b) Gerade, wenn Endzahl 2, 4, 6 oder 8, also 4 Ziffern von 9. Man wählt zwei Zahlen aus 8 aus und eine aus 4, den geraden. Also $M = 8 \cot 7 \cdot 4 = 224$.

(c) Nur Zahl mit Endziffer 5, als eine Möglichkeit. Somit $M = 8 \cdot 7 \cdot 1 = 56$.

(d) Kleiner als 200 heisst, erste Zahl ist 1. Somit $M = 1 \cdot \cdot 8 \cdot 7 = 56$. Grösser als 500 heisst, erste Zahl aus $\{5, 6, 7, 8, 9\}$. Also $M = 5 \cdot 8 \cdot 7 = 280$.

0.11 a) Wieviel verschiedene Würfe mit drei Würfeln sind möglich, bei denen alle Würfel verschiedene Augenzahlen zeigen? b) Wie viel verschiedene Würfe mit drei Würfeln sind möglich?

11 a) (Ob mit drei Würfeln oder 3 Mal mit einem spielt hier keine Rolle.) Es sind $6 \cdot 5 \cdot 4$ Möglichkeiten oder formell $C(6,3)$
b) Das Modell ist Kombination mit Wiederholung $C_w(6,3) = \frac{(6+2)!}{3!(6-1)!} = \frac{8!}{3!5!} = \frac{8 \cdot 7 \cdot 6}{6} = 56$.

0.12 a) Wie viel Telefonanschlüsse lassen sich einrichten, wenn nur siebenstellige Rufnummern verwendet werden? b) Wie viel Telefonanschlüsse sind bei siebenstelligen Rufnummern möglich, wenn solche, die mit 0 beginnen, nicht erlaubt sind?

12 a) Die Reihenfolge zählt, Wiederholungen der Ziffern erlaubt, somit Variation mit Wiederholung $V_w(10,7) = 10^7$.
b) Jede zehnte Nummer beginnt mit einer 0, deshalb ist die Anzahl $\frac{9}{10} V_w(10,7) = 9000000$.

0.13 Ein Eishockeyspiel mit drei Spielabschnitten von je 20 Minuten endet mit dem Resultat 5:3. Wie viele Möglichkeiten gibt es für die Drittelsresultate?

13 Für die erste Mannschaft folgt mit der Veranschaulichung $11/1/11$ die Permutationen $\frac{7!}{2!5!} = 21$ und die zweite $\frac{5!}{2!3!} = 10$. Deshalb total $21 \cdot 10 = 210$.

0.14 Ein Kartenspiel mit 36 verschiedenen Karten soll so unter 4 Spieler aufgeteilt werden, dass jeder genau 9 Karten erhält.

(a) Wie viele unterschiedliche Möglichkeiten gibt es?

(b) Wie viele Möglichkeiten gibt es, dass ein Spieler alle vier Asse erhält?

(c) Wie gross ist die Wahrscheinlichkeit?

14 **(a)** Mit jeder verteilten Karte nimmt die Möglichkeit um 1 ab. Der ereste Spieler hat die Möglichkeiten $36 \cdot 35 \cdot 34 \ldots$ oder $\binom{36}{9}$, der zweite dann $27 \cdot 26 \ldots$, oder $\binom{27}{9}$. Der dritte $\binom{18}{9}$ und der vierte $\binom{9}{9} = 1$. Zusammen $M = \binom{36}{9}\binom{27}{9}\binom{18}{9}\binom{9}{9}$.

(b) Wir beginnen mit der ersten Spieler, der die Asse bekommen soll. Er hat neben den Assen noch 5 Karten aus $36-4$ zu ziehen als $M_1 = \binom{32}{5}$. Der zweite Spieler dann $\binom{27}{9}$, der dritte $\binom{18}{9}$ und der vierte den Rest, also 1 oder $\binom{9}{9}$. Nun heisst es ein Spieler, von denen es 4 gibt; also mal 4. Daraus folgt $M = 4 \cdot \binom{32}{9}\binom{27}{9}\binom{18}{9} \cdot 1$.

(c) Die Wahrscheinlichkeit ist das Verhältnis von Möglichkeiten von 4 Assen zu allen Möglichkeiten: $WS = \frac{4 \cdot \binom{32}{5}\binom{27}{9}\binom{18}{9} \cdot 1}{\binom{36}{9}\binom{27}{9}\binom{18}{9}\binom{9}{9}} = \frac{4 \cdot \binom{32}{5}}{\binom{36}{9}} \approx 0.00856$

0.15 Wie viele Möglichkeiten gibt es, im Zahlenlotto “6 aus 49” genau 3,4,5, beziehungsweise 6 richtige Zahlen anzukreuzen?

15 Es gibt 6 richtige Zahlen und 43 falsche Zahlen. Man zieht die Anzahl richtiger Zahlen aus 6 und die Anzahl falscher aus 43 und multipliziert diese Möglichkeiten. Also 6 richtige: Ein Möglichkeit aus 6 richtigen 6 richtige zu ziehen, und 0 aus 43. Zusammen $M_6 = 1$. Mit drei richtigen: $M_3 = \binom{6}{3}\binom{43}{3} = 246820$, $M_4 = \binom{6}{4}\binom{43}{2} = 13545$ und $M_5 = \binom{6}{5}\binom{43}{1} = 258$. Alle Möglichkeiten sind $M = \binom{49}{6} = 13983816$ und nur falsche $\binom{43}{6} = 6096454$.

0.16 Es sollen 10 verschiedene Weine verglichen werden. Dabei werden immer 2 Proben gereicht. Um der Objektivität willen werden auch Proben mit zwei mal demselben Wein ausgeschenkt. Wieviele Tests ergeben sich aus dieser Anordnung?

16 Man wählt 2 aus 10, wobei die Reihenfolge unerheblich ist. Es entspricht dem Modell “mit Wiederholung”, da identische Proben zulässig sind. Somit Wählt man 2 aus 10 mit Wiederholung oder $C_w(10,2) = C(10+2-1,2) = \binom{11}{2} = 55$.

Normalprogramm / Erweitertes Niveau

die Begriffe Zufallsexperiment, Ergebnis, Ereignis, Wahrscheinlichkeit eines Ereignisses erklären
die Ereignisse nicht-A, A oder B, A und B, unabhängige und unvereinbare (disjunkte) Ereignisse definieren
das sichere und unmögliche Ereignis definieren
bedingte Wahrscheinlichkeiten berechnen
unabhängige Ereignisse und Folgen von Zufallsversuchen erkennen
einen Ergebnisbaum aufstellen und anwenden
die Begriffe Zufallsvariable, Erwartungswert und Standardabweichung einer Zufallsvariablen definieren, insbesondere bei einer Binomialverteilung
die Binomialverteilung anwenden
die Normalverteilung als Näherung der Binominalverteilung kennen

Kapitel 32

Wahrscheinlichkeitstheorie

Die Wahrscheinlichkeitstheorie hat ihren Ursprung in praktischen Fragestellung, d.h. dem Glücksspiel. Ein Leiter einer Spielbank, der Chevalier De Méré, beobachtete die Resultate, die er sich vom Mathematiker und Philosophen Blaise Pascal erklären liess. Einfache Vorgänge wie ein Wurf einer Münze, der Zug eines Loses, das Drehen am Roulette haben ungewisse Ausgänge. Aber lange Folgen von Versuchen zeigen Regelmässigkeiten. Dieser Zusammenhang von Ungewissheit im Einzelnen und der Regelmässigkeit im Kollektiv ist das Wesen der Wahrscheinlichkeitstheorie.

32.1 Zufallsexperiment

32.1.1 Begriffe

Definition 16. Ein *Zufallsexperiment* $\mathcal{E}$ bezeichnet einen Versuch, der unter genau festgelegten Versuchsbedingungen durchgeführt wird und einen zufälligen Ausgang (Ergebnis) hat.

Wichtig 5. Wir verwenden die Wörter "Ereignis", "Ergebnis" und "Ausgang" synonym. ⊣

Damit ein Experiment ein Zufallsexperiment ist, muss es folgende Eigenschaften aufweisen:

(1) Es gibt einen genau festgelegten Plan zur Durchführung,

(2) alle möglichen Ergebnisse des Experiments sind vorab bekannt,

(3) das Ergebnis jedes einzelnen Experiments kann nicht vorhergesagt werden (Zufälligkeit).

Anmerkung 32.1. Das Wort Experiment klingt sehr wissenschaftlich. Die Experimente sind aber auch, oder vor allem, Glücksspiele wie Würfeln oder Kartenspiele. Der Wurf eines Würfels ist ein Zufallsexperiment mit den drei obigen Eigenschaften. Die Wartezeit an einer Bushaltestelle ist auch ein Zufallsexperiment.

Anmerkung 32.2. Der Zufall kann von einem künstlichen Ding erzeugt werden, z.B. Würfel, oder aus der Natur (Geschlecht eines Neugeborenen) oder der menschlichen Umwelt (Buswartezeit) stammen.

Definition 17. Die Menge Ω aller möglichen Elementarereignisse ω_i eines Zufallsexperiments nennt man *Ereignisraum*.

Es ist das Zufallsexperiment das zweimalige Werfen einer Münze mit Kopf oder Zahl. Dann ist der Ereignisraum $\Omega = \{(K,K),(K,Z),(Z,K),(Z,Z)\}$. Beim Experiment "Augenzahl eines Würfels" ist $\Omega = \{1,2,3,4,5,6\}$.

Anmerkung 32.3. Der Ereignisraum enthält alle wesentlichen Informationen eines Zufallsexperiments.

Definition 18. Ein *zusammengesetztes Ereignis* A ist eine Teilmenge des Ereignisraums Ω, d.h. $A \subset \Omega$.

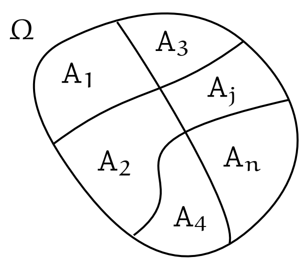
32.4 Beispiel Das Experiment besteht im Werfen eines blauen und eines roten Würfels. Es interessieren die zusammengesetzte Ergebnisse A: "Summe gleicher Zahlen". Der Ereignisraum ist $\Omega = \{(1,1),(1,2),\ldots(6,5),(6,6)\}$ und $A = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$, also Teilmenge des Ereignisraums.

Formel 32.5. Ein zusammengesetztes Ereignis A tritt genau dann ein, wenn das eingetretene Elementarereignis ω Element von A ist, also $\omega \in A$.

Das Beispiel fortführend: Es wird $\omega = (1,6)$ geworfen, es ist nicht Elemente von A, damit ist A nicht eingetreten. Es wird $(4,4)$ geworfen, dieses Elementarereignis ist Element von A, A ist eingetroffen.

Definition 19. Eine *Zerlegung* oder *Partition* vom Ereignisraum Ω sind n paarweise disjunkte Ereignisse A_j, die Ω aufteilen. Es ist

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n \quad \text{mit} \quad A_i \cap A_j = \varnothing$$



Bei einem Zufallsexperiment tritt genau ein Ereignis A_j ein.

32.1.2 Verknüpfung von Ereignissen

Ereignisse können auf *drei* Arten verknüpft werden:

- Komplement,
- (logische) Summe und
- (logisches) Produkt.

In Analogie zur Mengenlehre und zur mathematischen Logik kann man Ereignisse wie Mengen behandeln. Es gibt das komplementäre Ereignis zu A.

Definition 20. Das zum Ereignis A *komplementäre Ereignis* $\overline{A}$ besteht darin, dass A nicht eintritt, also $\overline{A} = \{\omega | \omega \notin A\}$

Man kann auch schreiben: $\overline{A} = \Omega \setminus A$.

Definition 21. Die *Summe zweier Ereignisse* A und B, $A \cup B$, ist das Ereignis, dass A oder B eintritt.

32.6 Beispiel Wurf eines Würfels: Es ist A die Augenzahl kleiner als 4 $\{1,2,3\}$ und B die ungerade Augenzahl $\{1,3,5\}$. Daraus $A \cup B = \{1,2,3,5\}$.

Definition 22. Das *Produkt zweier Ereignisse* A und B, $A \cap B$, ist das Ereignis, dass sowohl A als auch B eintritt.

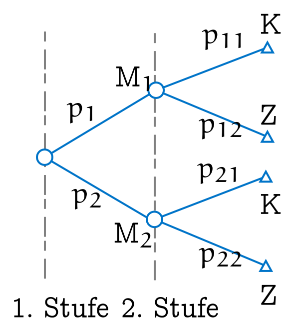
32.7 Beispiel Wurf eines Würfels: Es ist A die Augenzahl kleiner als 4 $\{1,2,3\}$ und B die ungerade Augenzahl $\{1,3,5\}$. Daraus $A \cap B = \{1,3\}$.

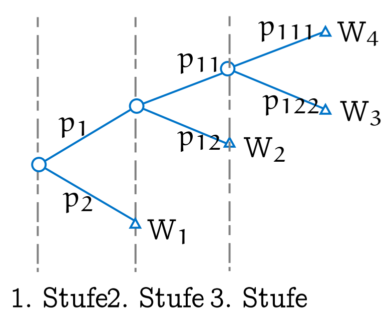
Definition 23. Ω heisst *sicheres Ereignis*, $\varnothing = \overline{\Omega}$ heisst *unmögliches Ereignis*.

Definition 24. Zwei Ereignisse A und B heissen *unvereinbar* oder *disjunkt*, wenn sie nicht gleichzeitig eintreten können, also $A \cap B = \varnothing$.

32.1.3 Ereignisbaum

Zufallsexperimente können einstufig oder mehrstufig sein. Der einmalig Wurf einer Münze ist einstufig (und wird Bernoulli-Experiment genannt). Das Experiment, ziehe ein von zwei Münzen aus dem Sack und Werfe einmal, ist ein zweistufiges Experiment. Mehrstufige Zufallsexperimente, und damit auch Glücksspiele, lassen sich übersichtlich als Graphen darstellen. Ausgehend von einem sogenannten Zufallsknoten (◯) (1. Stufe) gelangt man zu den nächsten Zufallsknoten durch eine Kante (Linie), deren Anzahl von der Anzahl Ereignisse abhängt. Am Schluss erscheinen die Endknoten als Elementarereignisse. Sie werden Blätter (△) genannt. Man beachte, dass häufig, aber nicht zwangsmässig, die Bäume von links nach rechts wachsen.





32.8 Übung Wir zeichnen einen Baum für das obige zweistufige Zufallsexperiment. Vom ersten Zufallsknoten gehen zwei Kanten ab, eine für jede Münze. Nun gehen von jedem Zufallsknoten wieder je zwei Kanten ab, die in den Blättern enden. Die Abbildung oben links gibt dies wieder. ◁

Die Bäume müssen nicht symmetrisch sein und können auch mehr als zwei Äste aufweisen. In der obigen Abbildung rechts könnte ein Würfelspiel dargestellt sein, das entweder abbricht oder weitergeht, und nach 3 Zügen fertig ist.
Bei den Blättern stehen die Ereignisse oder Ergebnisse oder Ausgänge. Als Vorgriff sein erwähnt, dass an den Kanten (Ästen) die Wahrscheinlichkeiten für das Eintreten dargestellt wird.

32.1.4 Zufallsvariable

Der Ereignisraum enthält alle Ereignisse als Elemente. Diese sind nicht nur Zahlen sondern können beliebige Elemente sein. Zum Beispiel ist $\Omega_1 = \{(K, K), (K, Z), (Z, K), (Z, Z)\}$ oder $\Omega_2 = \{⚀, ⚂, ⚄, ⚅\}$ und $\Omega_3 = \{(\clubsuit, \diamondsuit), (\heartsuit, \spadesuit), (\clubsuit, \heartsuit)\}$. Häufig ist man nicht so sehr am konkreten Ergebnis interessiert als an einer zahlenmässigen Darstellung. Beim Doppelwurf ist es die Summe der Augenzahl, oder beim Münzwurf die Anzahl von "Kopf" usw. Diese Zuordnung von Zahlen X_i zu den Ereignisse ω_i nennt man Zufallsvariable.

Definition 25. Eine *Zufallsvariable* ist die einem Ereignis $\omega \in \Omega$ zugeordnete reelle Zahl $X(\omega)$.

Anmerkung 32.9. Aus der Definition geht hervor, dass die Zufallsvariable, oder auch Zufallsgrösse, eine *Funktion* ist (und der Name nicht glücklich gewählt ist).

32.10 Beispiel Es ist $\omega = KKZKZZKZKZZZZKKKZKKZ$ die Folge von zwanzigmaligem Werfen einer Münze. Zu diesem Ergebnis wird die Anzahl "Kopf" zugeordnet, also $X(\omega) = 10$.

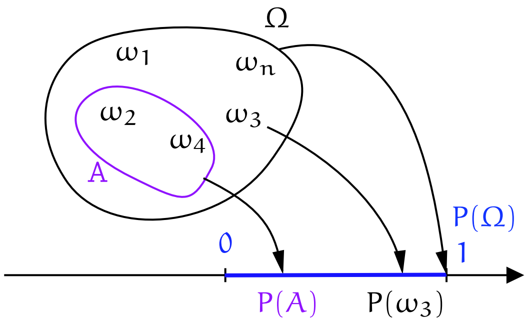
Definition 26. Die Werte $W = \{x_1, x_2, \ldots, x_n\}$ der Zufallsvariablen X heissen *Realisationen* von X.

32.2 Wahrscheinlichkeit

Die Wahrscheinlichkeit als Begriff besitzt mehrere Interpretationen. Man übergeht diese Schwierigkeit, indem man grundlegende Forderungen an das Rechnen mit Wahrscheinlichkeiten stellt. Man spricht dann von den Axiomen der Wahrscheinlichkeitsrechnung. Wie man zu den Wahrscheinlichkeiten kommt, mit Rechnung oder Schätzung, ist der erste Schritt, der mehrere Möglichkeiten offenlässt. Sind sie aber einmal festgelegt, gelten die Axiome.
Die Vielzahl von Interpretationen hat auch mit der Vielseitigkeit der Verwendung des Begriffs zu tun. Die Ungewissheit ist mit der unvorhersehbaren Zukunft verknüpft, die in vielen Lebensbereichen vorkommt. Wie wahrscheinlich ist ein Sieg von Kandidat Q? Wird es morgen regnen? Wie sicher ist ein Atomkraftwerk? Wie entwickeln sich die Aktienkurse? Ist die Prämie meiner Versicherung gerechtfertigt? Ist das ein faires Spiel? Ist das neue Medikament wirksam?
Aber auch rückblickend gibt es Fragestellungen mit Ungewissheit. War Homer eine Frau? Zu 99% nicht. Hat A seine Frau umgebracht, gegeben die verschiedenen Indizien?
Gewisse Ungewissheiten werden sich klären, wenn ein Ereignis eintritt, z.B. die Wahl von Kandidat P, oder viele Daten mit der Zeit anfallen, z.B. die Erfahrung mit den neuen Medikament oder die Gewinne der Versicherungsgesellschaft. Andere wird man nie klären können, d.h. das Skelett von Homer finden. Gewissen Wahrscheinlichkeiten werden immer Schätzungen bleiben.

Wir stellen die drei wichtigsten Interpretationen und Herleitungen dar. Wir sprechen vor allem von Wahrscheinlichkeiten für bestimme Ereignisse aber auch für Aussagen.

Die Wahrscheinlichkeit ist eine Funktion der Elementarereignisse oder zusammengesetzten Ereignisse eines Zufallsexperiments in das reelle Intervall $(0, 1)$.



32.2.1 Klassische Wahrscheinlichkeit

Die klassische Auffassung stammt vom Glückspiel, vom Wurf von Münzen, der Augenzahl von Würfeln, dem Roulette oder von Kartenspielen. Aus den Regeln des Spiels und der Beschaffenheit des Spielmaterials ergibt sich die Wahrscheinlichkeit als Verhältnis von günstigen Ergebnissen zu möglichen.

Definition 27. Ein Zufallsexperiment besitzt n gleichwahrscheinliche Ausgänge. Es ist n_A die Anzahl Ausgänge, bei denen A eintritt. Dann ist die *Wahrscheinlichkeit* für das Ereignis A

$$P(A) = \frac{n_A}{n}.$$

Anmerkung 32.1. Schon in der Definition sieht man, dass das Argument zirkulär ist, auf sich selber bezieht. Die Wahrscheinlichkeit ist Voraussetzung des zu definierenden. Meist wird dies kaschiert, indem von *fairen* Würfel o.ä. gesprochen wird.

Diese Definition, spezialisiert für Glücksspiele, lässt viele Fragestellungen offen, z.B. wie wahrscheinlich regnet es morgen?

32.2 Übung Wir betrachten den zweimaligen Wurf einer Münze mit Kopf oder Zahl. Gesucht ist die Wahrscheinlichkeit eines Doppelkopfes. Es gibt drei Ausgänge ohne Berücksichtigung der Reihenfolge, nämlich $(KK), (ZK)$ und (ZZ). Diese Ausgänge sind nicht gleichwahrscheinlich, denn (ZK) kann auf zwei Arten zustande kommen. Deshalb gibt es 4 Möglichkeiten für 3 Ausgänge. Die günstigen Fälle für (KZ) sind 2, für (KK) und (ZZ) jeweils 1. Damit sind die Wahrscheinlichkeiten $P(KZ) = 2/4$, $P(KK) = P(ZZ) = 1/4$. ◁

32.3 Übung Wir betrachten das Würfelexperiment "Doppelwurf zweier Würfel".

Der Ausgang oder das Ereignis ist die Summe der Augenzahlen. Damit ist der Ereignisraum $\Omega = \{2, 3, 4, 5, \dots, 11, 12\}$. Die Tabelle zeigt alle möglichen Kombinationen von Summen. Die Diagonalen ergeben die gleichen Summen. Deren Anzahl kann man als die für das entsprechende Ergebnis günstigen Fälle anschauen. Wir zählen und erstellen die folgende Tabelle:

Σ	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Ω	2	3	4	5	6	7	8	9	10	11	12
günstige Fälle ω_i	1	2	3	4	5	6	5	4	3	2	1
$P(\omega_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

◁

32.4 Übung Der Spielbankleiter De Méré hatte die Erfahrung gemacht, dass beim Dreifachwurf eines Würfels die Summe 11 häufiger auftrat als die Summe 12, obwohl es je sechs Anordnungen gibt, die zu den Summen führen. Es sind dies

11	146	155	236	245	335	344
12	156	246	255	336	345	444

Bei Dreifachwurf gibt es total $6^3 = 216$ Möglichkeiten. Eine Anordnung 146 mit lauter unterschiedlicher Zahlen hat $3! = 3 \cdot 2 \cdot 1 = 6$ Möglichkeiten. Eine Anordnung mit zwei unterschiedlichen Zahlen hat $3!/2! = 3 \cdot 2 \cdot 1/2 = 3$ Möglichkeiten und schliesslich eine Anordnung mit drei gleichen Zahlen hat 1 Möglichkeit. Wenn man diese Einsicht auf die Anordnungen anwendet, dann erhält man für 11: $6+3+6+6+3+3 = 27$ Möglichkeiten und für 12: $6+6+3+3+6+1 = 25$ Möglichkeiten. Die Wahrscheinlichkeiten sind also:

$$P(X = 11) = \frac{27}{216} \qquad P(X = 12) = \frac{25}{216}$$

Die Beobachtung von De Méré waren also richtig. ◁

32.5 Übung Auf dem Pausenplatz hat es genau 140 Schüler aus 7 Klassen zu 20 Schülern. Es werden 5 herausgegriffen. Wie hoch ist die Wahrscheinlichkeit, dass alle aus verschiedenen Klassen stammen? Es ist hier die Wahrscheinlichkeit das Verhältnis von "günstigen" zu "möglichen" Fällen. Zuerst zu den Möglichen: die Anordnungen ergeben sich aus "ziehe 5 aus 140" als Kombination $M = \binom{140}{5} = 416,965,528$. Die Günstigen erfordern zwei Schritte: bestimme die möglichen Klassen mit "ziehe 5 aus 7" oder $\binom{7}{5} = 21$. Nun pro Klasse "ziehe 1 aus 20" $\binom{20}{1} = 20$ und das fünfmal. Somit folgt $G = \binom{7}{5}\binom{20}{1}^5 = 67,200,000$. Damit $P \approx 0.161$. ◁

32.2.2 Häufigkeitsinterpretation

Die Häufigkeit hat zwei Ausprägungen, nämlich die absolute und die relative. Wir unterstellen ein wiederholtes Zufallsexperiment.

Definition 28. Die *absolute Häufigkeit* $H_n(A)$ eines Ereignisses A ist die Anzahl dessen Eintretens bei n Versuchen. Die *relative Häufigkeit* $h_n(A)$ ist die absolute Häufigkeit im Verhältnis zur Gesamtanzahl Experimente n. Also

$$H_n(A) = n_A \qquad h_n(A) = \frac{n_A}{n}$$

32.6 Beispiel Es wurde eine Münze 120 Mal geworfen und dabei ist Kopf 63 Mal erschienen. Die absolute Häufigkeit ist 63, die relative 63/120=0.525.

Der Zusammenhang mit der Wahrscheinlichkeit wird hergestellt, indem gefordert wird, dass bei sehr grossen n die relative Häufigkeit $h_n(A)$ gegen den Wert $P(A)$ strebt.

Definition 29. Die *Wahrscheinlichkeit* für das Ereignis A ist

$$P(A) = \lim_{n\to\infty} h_n(A)$$

Anmerkung 32.7. Aus dieser Definition ersieht man sofort, dass die Wahrscheinlichkeit für Ereignisse, die sich nicht wiederholen lassen, nicht existiert. Die Wahrscheinlichkeit, dass es morgen regnet, kann so nicht festgelegt werden, weil es nur ein "morgen" gibt. Ebenso kann die Wahrscheinlichkeit, eine schwere Krankheit für Frau Q zu überleben, nicht festgelegt werden. Für den Meteorologen sieht die Situation anders aus als für das Brautpaar, das morgen heiraten will. Er sagt sich: morgen ist ein Tag mit bestimmten Eigenschaften. Tage mit solchen Eigenschaften zeitigen in 25% der Fälle Regen. Oder für den Arzt: Patienten mit solchen Symptomen wie Frau Q sterben in 45% aller Fälle.

32.2.3 Subjektive Wahrscheinlichkeit

Die subjektive Wahrscheinlichkeit, drückt den subjektiven Überzeugungsgrad über das Eintreffen oder Nicht-Eintreffen eines Ereignisses aus und entsteht durch Überlegungen und Vorwissen. Zur Erfassung der subjektiven Wahrscheinlichkeiten benutzt man üblicherweise Wetten oder aleatorische Verträge.
Mit subjektiv ist nicht gemeint, dass jeder seine Wahrscheinlichkeit selber festlegen kann. Diese Wahrscheinlichkeit muss alle bekannten Fakten und Analogien miteinbeziehen.
Buchmacher legen die Preise für Wetten fest. Da es mehrere Anbieter gibt, gleichen sich die Preise an. Die Anbieter machen die Preise implizit miteinander. Die einfache Preismechanik ist: Gewinn $=$ Einsatz $\times$ Quote. Die Quote ist der Kehrwert der Wahrscheinlichkeit. Umgeformt resultiert: implizite Wahrscheinlichkeit $=$ Einsatz $/$Gewinn.
Ein Anbieter stellt die Quoten, ohne Profit für ihn, für ein Fussballspiel zwischen Bayern München und Leverkusen mit

Ereignis	Quote	Implizite WS
Sieg München	1.82	0.54
Unentschieden	3.7	0.27
Sieg Leverkusen	5.25	0.19

Bei Hunde- und Pferdewetten gibt es auch zusammengesetzte Ereignisse, z.B. Hund wird erster oder zweiter oder dritter.

32.2.4 Axiome der Wahrscheinlichkeitsrechnung

Axiome sind grundlegende Forderungen, die nicht bewiesen werden können. Sie sind das Fundament, auf dem eine Theorie gebaut wird.
Es besteht Einigkeit darüber, was eine Wahrscheinlichkeit ist, nämlich das Mass für den Grad der Gewissheit eines Sachverhalts. Wie man diesen bestimmt, ist kontrovers. Sobald man annimmt, man kennt die Wahrscheinlichkeiten, kann man mit den Axiomen fortfahren. Wir betrachten Axiome wie Sätze.

Satz 32.8. Axiome der Wahrscheinlichkeitsrechnung Es ist $P(A)$ die Wahrscheinlichkeit des Ereignisses $A \in \Omega$ und Ω der Ereignisraum.

- **A1**: $P(A) \geqslant 0$ für alle $A \subset \Omega$
- **A2**: $P(A \cup B) = P(A) + P(B)$ für alle A und B mit $A \cap B = \varnothing$
- **A3**: $P(\Omega) = 1$

32.2.5 Elementare Sätze

Die folgenden Sätze folgen aus den Axiomen der Wahrscheinlichkeitsrechnung.

Unmögliches Ereignis

Satz 32.9. Ist $\varnothing$ das unmögliche Ereignis, dann gilt

$$P(\varnothing) = 0$$

Denn mit Axiom 2 gilt: $P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing)$ und damit $P(\varnothing) = 0$.

Komplement

Satz 32.10. Ist $\overline{A}$ das komplementäre Ereignis (*Gegenereignis*) von A, dann gilt

$$P(\overline{A}) = 1 - P(A)$$

Denn $\Omega = A \cup \overline{A}$ und (Axiom 2 und 3) $P(\Omega) = 1 = P(A) + P(\overline{A})$ und somit $P(\overline{A}) = 1 - P(A)$.

Anmerkung 32.11. Dieses Gesetz braucht man häufig, dann nämlich, wenn $P(\overline{A})$ viel einfacher zu berechnen ist als $P(A)$. Dies ist auch dann der Falle, wenn die Ereignismenge A viel mächtiger ist als $\overline{A}$.

Additionssatz

Aus den Axiomen, als Wiederholung:

Formel 32.12. Axiom 2 Für zwei unvereinbare, disjunkte, Ereignisse A und B $(A \cap B = \varnothing)$ gilt

$$P(A \cup B) = P(A) + P(B)$$

Satz 32.13. Für n paarweise disjunkte Ereignisse $A_i \cap A_j = \varnothing$ und $i \neq j$

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots P(A_n).$$

Nun verallgemeinert für nicht-disjunkte Ereignisse

Satz 32.14. Für zwei beliebige Ereignisse A und B gilt

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Die Schnittmenge ist doppelt gezählt, weshalb sie in Abzug gebracht werden muss. Man erinnere sich an die Hundehalter und die Katzenhalter, von deren Vereinigung man die Hunde- und Katzenhalter abziehen muss.

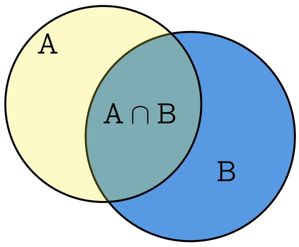
32.15 Übung Es interessiert die Wahrscheinlichkeit, dass zwei Würfel bei gleichzeitigem Wurf *mindestens* eine Sechs aufweisen. Eine Sechs hat die Wahrscheinlichkeit 1/6. Ein Student gibt die Antwort: 1/6+1/6=2/6=1/3. Bei näherem Betrachten muss man sich die Frage stellen, sind $A = \{(6)\}$ und $B = \{(6)\}$ unvereinbar? Oder anders $A \cap B = \varnothing$? Nein, denn es ist ja $(6, 6)$ möglich. Diese wurde doppelt "gezählt". Diese Wahrscheinlichkeit ist 1/36. Also ist die richtige Antwort $P(X) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$. Hätte man das Gegenereignis "keine Sechs" verwendet, hätte man argumentiert: Wahrscheinlichkeit erster Würfel keine Sechs $p = 5/6$, ebenso für den zweiten. Die Ereignisse sind unabhängig, also total $p = 25/36$. Damit $P(x) = 1 - p^2 = 1 - 25/36 = 11/36$. $\triangleleft$

Bedingte Wahrscheinlichkeit

Definition 30. Die *bedingte Wahrscheinlichkeit* $P(A|B)$ ist die Wahrscheinlichkeit von A unter der Voraussetzung, dass das Ereignis B mit $P(B) \neq 0$ eingetreten ist, formell

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Anmerkung 32.16. Man kann die Formel auch als Normierung auf die Basis B verstehen, als Verhältnis des Teils von B, der auch zu A gehört, zu B.



Es zeigt das Verhältnis der Hunde- und Katzenhalter $A \cap B$ im Verhältnis zu den Katzenhalter B.

Wichtig 6. ☠ Im allgemeinen gilt $(P(A|B) \neq P(B|A)$. ⊣

32.17 Übung Wir betrachten Familien mit zwei Kindern. Die Wahrscheinlichkeit für beide Geschlechter sei 50%. Es ist A die Menge von Familien mit 2 Knaben und B die Menge von Familien mit *mindestens* einem Knaben. Wir wählen eine Familie aus B und fragen uns, wie gross die Wahrscheinlichkeit ist, dass sie 2 Knaben aufweist. Die Wahrscheinlichkeit $P(B)$ ist das Komplement zu "keine 2 Mädchen". Zwei Mädchen haben die Wahrscheinlichkeit $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Damit $P(B) = 1 - 1/4 = 3/4$. Da A Teilmenge von B ist, folgt $P(A \cap B) = P(A)$ mit $P(A) = 1/4$. Damit folgt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

Das ist nicht für alle besonders intuitiv. ◁

Multiplikationssatz

Die Definition der bedingten Wahrscheinlichkeit kann man umschreiben als

Formel 32.18.

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Für das gleichzeitige eintreten von n Ereignissen A_i gilt die Verallgemeinerung

Satz 32.19.

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

Für drei Ereignisse folgt:

Formel 32.20.

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$$

32.21 Übung Das Lottospiel sieht vor, 6 aus 49 Zahlen ohne Wiederholung zu ziehen. Alle 6 Zahlen richtig angekreuzt zu haben bedeutet, die Ereignisse $A_1, A_2 \ldots, A_6$ stimmen mit den gezogenen Nummern überein. Wir suchen also $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6)$. Es ist $P(A_1) = \frac{6}{49}$, $P(A_2|A_1) = \frac{5}{48}$, $P(A_3|A_1 \cap A_2) = \frac{4}{47}$ usf. Es gibt immer eine Zahl weniger, die man auf einen Platz weniger setzen kann.

$$P(A_1, A_2 \ldots, A_6) = \frac{6}{49} \cdot \frac{5}{48} \cdot \frac{4}{47} \cdot \frac{3}{46} \cdot \frac{2}{45} \cdot \frac{1}{44} = \frac{1}{\binom{49}{6}}$$

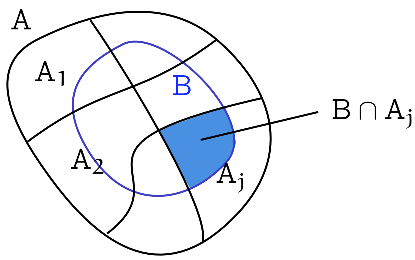
Wir erinnern uns an die Kombination ohne Zurücklegen. Wie viele Möglichkeiten gibt es, 6 aus 49 zu wählen? Antwort $C(49, 6) = \binom{49}{6}$. Es gibt nur eine günstige Wahl. Damit ist die Wahrscheinlichkeit eben

$$\frac{1}{\binom{49}{6}}.$$

Die Überlegung mit der Kombinatorik im Hinterkopf ist bei gleichwahrscheinlichen Ereignissen einfacher als das Handhaben von Wahrscheinlichkeiten. ◁

Totale Wahrscheinlichkeit

Wir betrachten eine Menge A, die in n disjunkte Teilmengen A_j zerlegt ist. Die Menge B ist eine echte Teilmenge von A. Damit ist die Menge B auch darstellbar als $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$. Damit folgt



$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_n \cap B) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots + P(B|A_n)P(A_n) \end{aligned}$$

Deshalb:

Satz 32.22. Satz von der totalen WS

$$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

Formel 32.23. Für $n = 2$ gilt

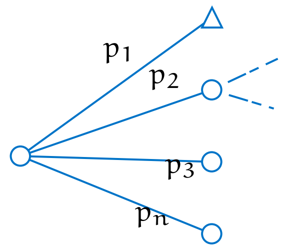
$$P(B) = P(B|A) \cdot P(A) + P(B|\overline{A}) \cdot P(\overline{A})$$

32.24 Übung Ein Fabrikant hat drei Lieferanten, die fehlerhafte Produkte mit Wahrscheinlichkeiten 0.06, 0.1 und 0.05 liefern. Sein Lager besteht aus 40%, 25% und 35% Produkte je Lieferant. Wie gross ist die Wahrscheinlichkeit, dass ein zufällig entnommenes Element fehlerhaft ist? Es sind B die fehlerhaften und A_i die Produkte von Lieferant i. Wir nehmen die Formel der totalen Wahrscheinlichkeit für $n = 3$, die lautet:

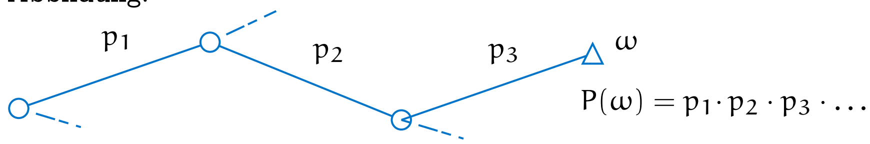
$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) \\ &= 0.06 \cdot 0.4 + 0.1 \cdot 0.25 + 0.05 \cdot 0.35 \\ &= 0.0665 \end{aligned}$$

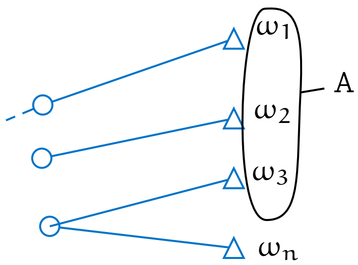
Intuitiv ist es einfacher mit ganzen Zahlen zu operieren. Wir nehmen an, das Lager bestehe aus 40, 25 und 35 Produkten. Von diesen sind 2.4, 2.5 und 1.75 defekt, total 6.65 von Hundert, oder 0.0665. ◁

32.2.6 Regeln für Baumdiagramme



Die Baumdiagramme, die ein meist mehrstufiges Zufallsexperiment beschreiben, sind so konstruiert, dass gewisse Regelmässigkeiten herrschen. Die Wahrscheinlichkeiten der Kanten (Äste) aus einem Knoten summieren sich zu 1, also $p_1 + p_2 + \ldots + p_n = 1$. Siehe Abbildung.





Die Wahrscheinlichkeit eines Ereignisses (Blatt) ist das Produkt der Wahrscheinlichkeiten längs des Pfades.
Die Wahrscheinlichkeit $P(A)$ eines zusammengesetzten Ereignisses $A = \{\omega_1, \omega_2, \ldots\}$ ist die Summe der Wahrscheinlichkeiten der Elementarereignisse $\omega \in A$, also

$$P(A) = \sum_{\omega \in A} p(\omega)$$

Wir fassen zusammen

Eigenschaften 32.25. Ereignisbaum

- die Wahrscheinlichkeit des Ereignisses am Ende eines bestimmten Pfades ist das Produkt der Wahrscheinlichkeiten entlang eben dieses Pfades
- Die Wahrscheinlichkeit für ein Ereignis bestehend aus mehreren Pfaden ist die Summe der Wahrscheinlichkeiten der einzelnen Pfade

32.2.7 Bayes'sche Formel**

Aus der Definition der bedingten Wahrscheinlichkeit und dem Satz von der totalen Wahrscheinlichkeit folgt folgender Zusammenhang.

Satz 32.26. Die A_j bilden eine Zerlegung von Ω und B ist irgendein Ereignis, so dass

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum\limits_{i=1}^{n} P(B|A_i)P(A_i)}$$

Formel 32.27.

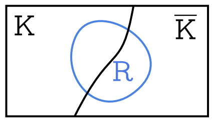
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})}$$

32.28 Rezept Die Bayes'sche Formel kann man als Modell verstehen für Ursache und Wirkung. Z.B. ist bekannt, dass eine Ereignis B eingetreten ist, die Wirkung. Es steht eine Menge von Ursachen A_j zur Verfügung, die Zerlegung von Ω. Welche der Ursachen A_j die Wirkung erzielt hat ist nicht bekannt aber die Wahrscheinlichkeit schon $P(B|A_j)$. Damit lässt sich berechnen, mit welcher Wahrscheinlichkeit A_j als Ursache, als Bedingung, in Frage kommt.

32.29 Übung Mit Röntgenuntersuchungen kann man Tuberkulosekranke finden. Es ist K das Ereignis krank mit der Wahrscheinlichkeit $P(K) = 0.001$ und folglich nicht-krank $P(\overline{K}) = 0.999$. Wir suchen das Ereignis $P(K|R)$, also Kranke, die röntgenpositiv sind. Die Röntgenuntersuchung liefert positive Resultate für Kranke mit Wahrscheinlichkeit $P(R|K) = 0.9$ und für Nicht-Kranke von $P(R|\overline{K}) = 0.01$. Mit der Formel folgt:

$$P(K|R) = \frac{P(R|K)P(K)}{P(R|K)P(K) + P(B|\overline{K})P(\overline{K})}$$

$$P(K|R) = \frac{0.9 \cdot 0.001}{0.9 \cdot 0.001 + 0.01 \cdot 0.999} = 0.083$$

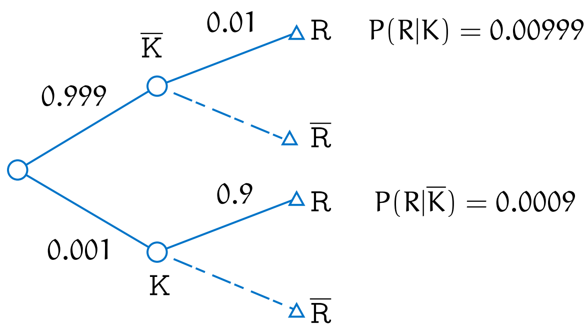


Psychologen haben gezeigt, dass man die Formel mit ganzen Zahlen besser beherrschen kann. Gehen wir von einer Bevölkerung von 100000 Personen aus. Davon sind 100 krank und 99900 gesund. Von 100 kranken werden 90 positiv getestet, von den 99900 Gesunden sind 1% positiv, also 999. Insgesamt sind 1089 positiv. Die Wahrscheinlichkeit einen Kranken aus den Positiven zu ziehen ist

$$\frac{90}{1089} = 0.083$$

Die Quote von 8.3% ist sehr gering, weshalb das Röntgenverfahren nicht effizient ist. Dies hätten wir aus den gegebenen Zahlen, vor allem wegen der 90%, intuitiv nicht erahnt. Es ist das 1%, das zu hoch ist. ◁

32.30 Übung Wir stellen denselben Zusammenhang mit einem Ereignisbaum dar. Dieser sieht z.B. so aus



Die Ergebnisse mit $\overline{R}$ interessieren nicht. Man sieht sofort, dass $P(R) = P(R|\overline{K}) + P(R|K) = 0.01089$ ist. $P(K|R) = 0.0009/0.01809 = 0.083$. Der Baum ist sehr nützlich für den Überblick. ◁

32.2.8 Unabhängige Ereignisse

Wir besprechen ein wichtiges Konzept ausgehend von den bedingten Wahrscheinlichkeiten und von disjunkten (unvereinbaren) Ereignissen. Sind A und B unvereinbar, d.h. $A \cap B = \varnothing$, dann folgt $P(A|B) = 0$. Ist B eine Teilmenge von A, also $B \subset A$, dann folgt $P(A|B) = 1$. Ist eine Frau gewählt worden, dann ist auch ein Mensch gewählt worden, als Illustration. Das Ereignis B bestimmt die Wahrscheinlichkeit von A mit.

Nun wird gezeigt, dass es Fälle gibt, bei denen der Eintritt von B keine Wirkung auf A zeigt. Zur Illustration der Doppelwurf von zwei Würfeln. Es ist natürlich nicht so, dass das Resultat des ersten Würfels den Ausgang des zweiten irgendwie beeinflusst.

32.31 Übung Es ist das Ereignis A: Augenzahl auf Würfel 1 ungerade und B: Augenzahl Würfel 2 ist grösser als 2. Es ist dann $P(A) = 1/2$ und $P(B) = 1/3$. Die Wahrscheinlichkeit, dass beide eintreten ist $P(A \cap B) = 1/6$. Wenden wir die Definition der bedingten Wahrscheinlichkeit an, so folgt

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/3} = 1/2 = P(A)$$

und

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/6}{1/2} = 1/3 = P(B)$$

Sind $P(A) > 0$ und $P(B) > 0$, dann folgt

$$P(A \cap B) = P(A) \cdot P(B)$$

Darauf bauen wir die Definition. ◁

Definition 31. Zwei Ereignisse A und B sind genau dann *unabhängig*, wenn gilt

$$P(A \cap B) = P(A) \cdot P(B)$$

Aus der Definition entspringen folgende Sätze.

Satz 32.32. Die Ereignisse A und B mit $P(A) > 0$ und $P(B) > 0$ sind genau dann unabhängig, wenn eine der Bedingungen

$$P(A|B) = P(A) \qquad \text{oder} \qquad P(B|A) = P(B)$$

erfüllt ist

Anmerkung 32.33. Es gilt auch für A und B unabhängig für die Komplemente:

$$P(A) = P(A) \cdot P(B) = P(A) \cdot P(\overline{B})$$

32.2.9 Beispiel Endlosspiel mit geometrischer Reihe

32.34 Übung Zwei Basketballer messen sich in ihrer Treffsicherheit: Sie werfen nacheinander abwechselnd auf einen Korb. Wer zu erst den Korb trifft, gewinnt. Spieler A trifft mit einer Wahrscheinlichkeit von $p_A = 0.6$. Spieler B trifft den Korb mit einer Wahrscheinlichkeit von $p_B = 0.8$. Die Versuche aller Spieler sind als unabhängige Ereignisse anzunehmen. Spieler A beginnt das Spiel.

Wir bestimmen die Wahrscheinlichkeit, dass Spieler A gewinnt (und damit auch die Wahrscheinlichkeit, dass B gewinnt). Intuitiv ist klar, dass der Beginnende bei gleicher Wahrscheinlichkeit die bessere Gewinnchance hat. Anderseits ist das Spiel nicht abbrechend, wenn nie getroffen wird. Im ersten Versuch hat A die Wahrscheinlichkeit $p_{A,1} = 0.6$. Im zweiten Versuch, der voraussetzt, dass B nicht getroffen hat und man selber im ersten Wurf ebenfalls nicht, ist $P_{A,2} = (1 - p_A)(1 - p_B)p_A = 0.4 \cdot 0.2 \cdot 0.6$. Weiter ist $P_{A,3} = (1 - p_A)(1 - p_B)(1 - p_A)(1 - p_B)p_A$. Damit die Rekursion

$$P_{A,k} = (1 - p_A)^{k-1}(1 - p_B)^{k-1}p_A = [(1 - p_A)(1 - p_B)]^{k-1}p_A$$

Das ist die Wahrscheinlichkeit, im k-ten Wurf zu treffen. Die Wahrscheinlichkeit zu treffen summiert sich über alle Versuche. Somit

$$S_{A,\infty} = \sum_{k=1}^{\infty} [(1-p_A)(1-p_B)]^{k-1} p_A$$

Das ist eine geometrische Reihe nach dem Muster $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$, sodass gilt

$$S_{A,\infty} = \frac{p_A}{1-(1-p_A)(1-p_B)} = \frac{p_A}{1-0.4\cdot 0.2} = \frac{0.6}{0.92} \approx 0.652$$

Obwohl er der schlechtere Schütze ist, hat er wegen der Reihenfolge der Würfe die höhere Gewinnwahrscheinlichkeit. Wäre p_B auch 0.6, dann wäre $S_{A,\infty} = \frac{0.6}{0.84} \approx 0.714$. ◁

32.3 Zufallsvariablen*

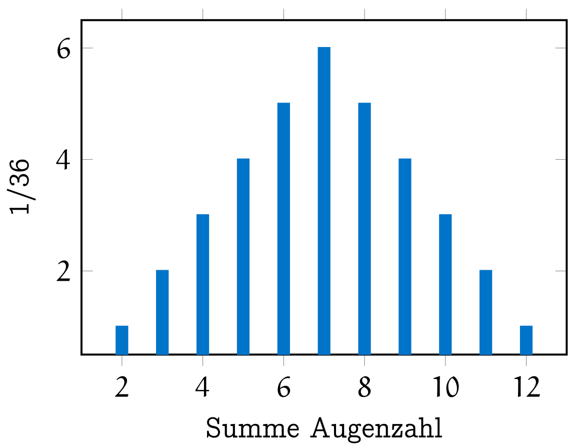
Nun da wir den Begriff der Wahrscheinlichkeit beschrieben haben, kommen wir auf das Zufallsexperiment zurück. Es wird durch zwei oder drei Elemente definiert: die Ereignisse $\Omega = \{\omega\}$, evtl. die zusammengesetzten Ereignisse $S = A \cap B \cup C \ldots$ mit $A \subset \Omega$ etc. und die Wahrscheinlichkeiten P. Das Duo (Ω, P) oder das Trio (Ω, S, P) nennt man *Wahrscheinlichkeitsraum*.

Wir haben bereits die Zufallszahl definiert (Def. 25 auf Seite 32-3) als Abbildung $X(\omega_i) = x_i$ vom Ereignis ω_i in die reellen Zahlen, und die Wertemenge $W = \{x_1, x_2, \ldots, x_n\}$ als Realisationen (Def. 26) der Zufallszahl.

32.3.1 Wahrscheinlichkeitsfunktion, Verteilung

Es ist X eine (diskrete) Zufallsvariable "über" Ω. Es interessiert die Wahrscheinlichkeit, mit denen X die Realisationen $x_1, x_2, \ldots, x_n$ annimmt. Das Ereignis $A_j \subset \Omega$ umfasst genau jene Elemente von Ω, für die gilt $X(\omega) = x_i$. A_j ist also die Menge $\{\omega | X(\omega) = x_i\}$. Damit ist $P(A_j) = P(X = x_j)$.

32.1 Übung Wir betrachten den Doppelwurf mit den Ereignissen $\Omega = \{(i,j)\}$. Wir haben die Abbildung $X((i,j)) = i + j$, d.h. die Summe der Augenzahlen. $A_k = \{(i,j) | i + j = k\}$ sind die Ereignisse, dass die Summe k erscheint. In der Abbildung sind die Wahrscheinlichkeiten $P(A_k)$ in 1/36 angegeben.



◁

Definition 32. Für $A_i = \{\omega \,|\, X(\omega) = x_i\}$ ist die *Wahrscheinlichkeitsfunktion* $f(x)$

$$f(x) = \begin{cases} P(X = x_i) = P(A_i) & \text{falls} \quad x = x_i \\ 0 & \text{sonst} \end{cases}$$

Für das obige Beispiel des Doppelwurfs ist die Funktion

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

32.2 Übung Für den einfachen Wurf mit x der Augenzahl ist die Wahrscheinlichkeitsfunktion $f(x)$ einfach

x	1	2	3	4	5	6
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

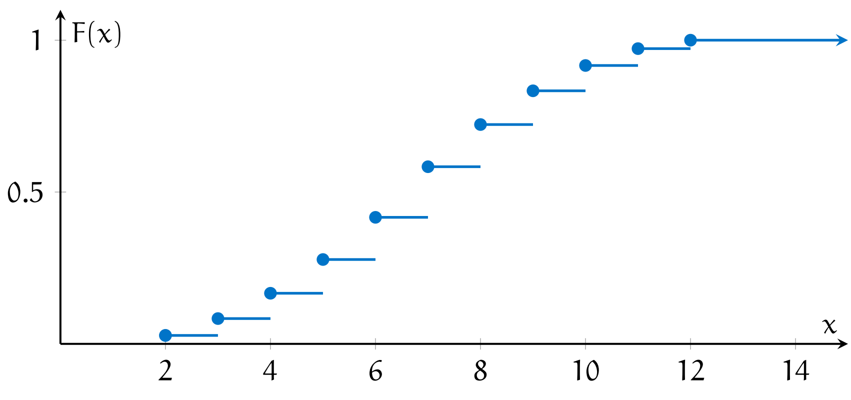
◁

Verteilungsfunktion

Die Verteilungsfunktion geht aus der Wahrscheinlichkeitsfunktion hervor. Es wird gefragt nach der Wahrscheinlichkeit, dass eine Zufallsvariable X einen bestimmten Wert x nicht übersteigt. Dazu müssen die Wahrscheinlichkeiten $P(X = x_i)$, für die $x_i \leqslant x$ ist, aufsummiert werden.

Definition 33. Die *Verteilungsfunktion* $F(x)$ zu $f(x)$ ist

$$F(x) = P(X \leqslant x) = \sum_{x_i \leqslant x} f(x_i)$$



Satz 32.3.

$$P(a < X \leqslant b) = F_X(b) - F_X(a)$$

Eigenschaften 32.4. Verteilungsfunktion

$$0 \leqslant F(x) \leqslant 1$$
$$F(x_{i+1}) = F(x_i) + f(x_{i+1})$$

32.3.2 Erwartungswert, Streuung

Unser Zufallsmodell kann man tabellarisch wie folgt darstellen, als Abbildung der Ereignisse in die Zufallszahlen und die Zufallszahlen in die Wahrscheinlichkeiten

Ω	ω_1	ω_2	ω_3	$\ldots$	ω_n
$X(\omega)$	x_1	x_1	x_3	$\ldots$	x_n
$P(X)$	p_1	p_2	p_3	$\ldots$	p_n

oder konkret für den fairen Würfel mit der Augenzahl als Zufallsgrösse

Ω	⚀	⚁	⚂	⚃	⚄	⚅
$X(\omega)$	1	2	3	4	5	6
$P(X)$	1/6	1/6	1/6	1/6	1/6	1/6

Eine übersichtliche Charakterisierung von Verteilungs- oder Wahrscheinlichkeitsfunktionen sind Mittelwert oder Erwartungswert und Varianz. Der Erwartungswert gibt Auskunft über die Lage (wo ungefähr?) und die Varianz über die Streuung (wie weit auseinander?).

Definition 34. Der *Erwartungswert* $E(X)$ einer diskreten Zufallsvariablen X ist

$$E(X) = \sum_i x_i \cdot P(x = x_i)$$

Beim Würfel ergibt sich der Erwartungswert (auch Mittelwert genannt) als

$$E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3.5$$

Anmerkung 32.5. Der Erwartungswert ist vor dem Gebrauch von Wahrscheinlichkeiten verwendet worden, nämlich als "gerechter" Preis von Verträgen, wo eine feste Zahlung für ein ungewisses Resultat entschädigt wird, z.B. Preis des nächsten Fischfangs, oder Prämie für eine Versicherung. Der Erwartungswert ist der Einsatz für ein faires Spiel.

Definition 35. Die *Varianz* $V(X) = \sigma_X^2$ einer diskreten Zufallsvariablen X mit Erwartungswert μ_X ist

$$V(X) = \sum_i (x_i - \mu_X)^2 \cdot P(X = x_i)$$

$\sigma_X = \sqrt{V(X)}$ heisst *Standardabweichung*.

Für den fairen Würfel ist die Varianz:

$$\begin{aligned} V(X) &= \frac{1}{6}\left[(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2\right] \\ &= \frac{1}{6}\left[6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.252\right] \\ &= 2.92 \end{aligned}$$

Die Standardabweichung ist $\sigma = 1.71$.

Anmerkung 32.6. Die Varianz hat nicht dieselbe Masseinheit wie X. Wenn X in Meter gemessen würde, dann ist die Varianz in Quadratmeter. Die Standardabweichung hingegen hat dasselbe Mass wie die Zufallsgrösse.

Eigenschaften 32.7. Erwartungswert E(X) Mit Konstanten c gilt:

$$E(c) = c$$
$$E(cX) = cE(X)$$
$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

Die Varianz ist auch ein Erwartungswert und zwar von $(X - \mu)^2$, also $V(X) = E((X - \mu)^2)$. Man kann also mit $E(X) = \mu$ schreiben:

$$\begin{aligned} V(X) = E((X-\mu)^2) &= E(X^2 - 2X\mu + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

Man kann alternativ auch folgende Darstellung zeigen:

$$V(X) = E(X^2 - 2X\mu + \mu^2) = E(X^2 - X + X - 2X\mu + \mu^2) = E(X^2 - X) + E(X) - \mu^2$$

und mit $\mu = E(X)$ folgt

$$V(X) = E(X(X-1)) - \mu(\mu - 1) \tag{32.1}$$

Eigenschaften 32.8. Varianz V(X) Mit Konstanten c gilt:

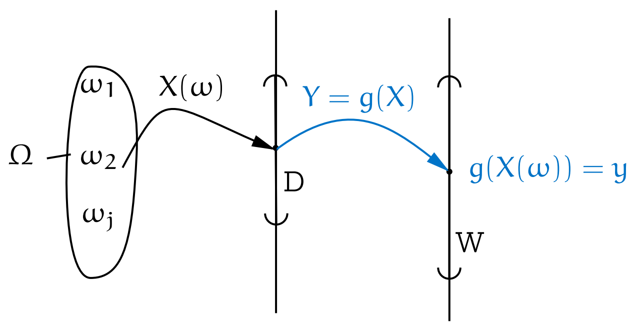
$$V(c) = 0$$
$$V(cX) = cV(X) \quad \text{und} \quad \sigma_{cX} = \sqrt{c}\sigma_x$$
$$V(X) = E(X^2) - \mu^2 = E(X(X-1)) - \mu(\mu-1)$$

32.3.3 Funktion einer Zufallsvariablen**

Wir haben mit der Zufallsvariablen eine Abbildung, genauer eine Funktion, der Ereignisse auf die Zahlenebene $X(\omega)$. Diese Zahlen können wiederum die unabhängige Variable einer beliebigen Funktion $Y = g(X)$ sein. Wenn X eine bestimmte Wahrscheinlickeitsfunktion und damit Verteilung besitzt, so wird die Zufallszahl Y eine andere, von der Funktion $g(X)$ bestimmte Verteilung aufweisen.

Satz 32.9. Ist $Y = g(X)$ eine Funktion der Zufallsvariablen X, dann ist auch Y eine Zufallsvariable.

Die Verteilung von Y ergibt sich aus $P(W) = P(D|g(B) = W)$. Die Berechnung ist meist sehr aufwendig und auch häufig nicht einfach. Wir gehen nicht weiter darauf ein. Als Beispiel nehmen wir die Funktion $Y = X^2$. Ein weiteres Beispiel könnte eine Auszahlungsfunktion in Abhängigkeit von einem Glücksspiel sein.



Kurioserweise braucht man für $E(Y)$ nicht die Verteilung von Y. Es gilt nämlich $E(Y) = E(g(X)) = \sum_i g(x_i)P(X = x_i)$ respektive $E(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx$.

32.4 Binomialverteilung*

Es gibt unendlich viele Verteilungen, denn gemäss Definition sind nur gewisse Eigenschaften einzuhalten. Hier wird vor allem auf die Binomialverteilung eingegangen.

32.4.1 Grundmodell

Die Grundvorstellung kann eine Serie von Münzwürfen sein. Es könnte dies z.B. folgende Ereignis eintreten bei 20 Würfen:

$$\mathsf{KKZZKZZZKKKZKZKKKZZK}$$

Uns interessiert nicht die Reihenfolge, somit ist diese Serie der folgenden äquivalent

$$\mathsf{KKKKKKKKKKKZZZZZZZZZ}$$

Es sind 11 mal "Kopf" und 9 mal "Zahl". Aus der Kombinatorik wissen wir, dass es $C(20, 11) = \binom{20}{11}$ Möglichkeiten gibt für diese Muster und total 2^{20} Anordnungen. Das Verhältnis ist $\frac{167960}{1048576} = 0.1602$, die Quote von günstigen zu mögliche Anordnungen. Dies ist eine Wahrscheinlichkeit unter der Voraussetzung, dass die Münze fair ist, also gleichwahrscheinlich die Seiten zeigt.
Die Wahrscheinlichkeit für die Köpfe ist, da alle Würfe unabhängig sind

$$P(KKKKKKKKKKK) = P(K) \cdot P(K) \ldots = P(K)^{11}$$

und analog für Zahl

$$P(ZZZZZZZZZ) = P(Z)^9$$

und damit

$$P(KKKKKKKKKKKZZZZZZZZZ) = P(K)^{11} \cdot P(Z)^9$$

Wir interessieren uns nicht für die Reihenfolge. Deshalb gibt es $\binom{20}{11}$ Mögliche Anordnung. Die Frage nach der Wahrscheinlichkeit in einer Serie von 20 Würfen 11 mal Kopf zu bekommen ist somit

$$P(n_K = 11, n = 20) = \binom{20}{11} P(K)^{11} \cdot P(Z)^9$$

Weil $Z = \overline{K}$ ist, d.h. entweder Kopf oder Zahl tritt mit Wahrscheinlichkeit 1 ein, muss $P(Z) = 1 - P(K)$ gelten. Wenn die Münze fair ist, dann gilt auch $P(Z)P(K) = \frac{1}{2}$. Damit wird

$$P(n_K = 11, n = 20) = \binom{20}{11} (\frac{1}{2})^{20} = 0.1602$$

Das ist die Zahl, die wir schon mit kombinatorischer Anschauung oben bestimmt haben. Wenn nun der Würfel nicht fair ist, sondern $P(K) = 0.53$ und damit $P(Z) = 0.47$, dann ist

$$P(n_K = 11, n = 20) = \binom{20}{11} (0.53)^{11} (0.47)^9 = 0.1742$$

Wir fassen zusammen

Formel 32.1. Die Anzahl von Sequenzen der Länge n von K und Z, in der K m mal erscheint, ist

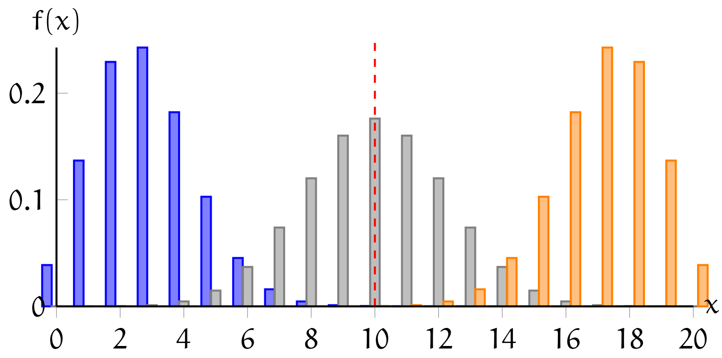
$$C(n, k) = \binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Die *Wahrscheinlichkeitsfunktion* für dieses Zufallsexperiment, Anzahl Erfolge x mit Wahrscheinlichkeit p in n Versuchen, ist:

Satz 32.2. WS-Funktion x aus n mit p

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{für} \quad x = 0, 1, 2 \ldots, n \\ 0 & \text{sonst} \end{cases}$$

Rechts sehen wir ein paar Ausprägungen der Binomialverteilung. Falls $p = 0.5$ ist, ist sie symmetrisch, andernfalls nicht. Es besteht eine Symmetrie zwischen Funktionen mit Parameter p und $1-p$. Mit der Definition 33 folgt sie sogenannte Binomialverteilung, häufig mit $B(n, p)$ bezeichnet:



Satz 32.3. Binomialverteilung

$$F(x) = P(X \leqslant x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

Man schreibt für eine Zufallsvariable X, die binomialverteilt ist, $X \sim B(n, p)$.

Satz 32.4. Eine binomialverteile Zufallvariabel $X \sim B(n, p)$ besitzt

- den Erwartungswert $E(X) = np$
- die Varianz $V(X) = np(1-p)$

32.5 Übung ** Die Herleitung des Erwartungswerts setzt die Kenntnis des binomischen Lehrsatzes voraus. Repetition:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Der Erwartungswert folgt aus der Definition:

$$\begin{aligned}
E(X) &= \sum_{k=0}^{n} k \cdot P(X=k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \frac{k \cdot n!}{(n-k)!k!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n \cdot (n-1)!}{(n-k)!(k-1)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \binom{n-1}{k-1} p \cdot p^{k-1} (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \\
&= np \sum_{l=0}^{m} \binom{m}{l} p^{l} (1-p)^{m-l}
\end{aligned}$$

mit $m = n-1$ und $l = k-1$ und dem Lehrsatz folgt

$$E(X) = np \cdot (p + (1-p))^m = np \cdot 1 = np.$$

Die Varianz bestimmt sich mit denselben Argumenten, wobei noch ein bisschen mehr umgeformt werden muss. Wir verwenden Gleichung 32.1. Der zweite Term ist einfach

$$\mu(\mu - 1) = np(np-1)$$

Der erste ist

$$E(X(X-1)) = \sum_{k=0}^{n} (k-1) \cdot k \cdot P(X=k)$$

Die rechte Seite unterscheidet sich von der obigen Herleitung nur durch den blauen Term $(k-1) = l$. Somit können wir ihn bis zur letzten Form einfach stehen lassen und bekommen

$$E(X(X-1)) = np \sum_{l=0}^{m} l \binom{m}{l} p^{l} (1-p)^{m-l}$$

Die Summe ist der Erwartungswert $mp = (n-1)p$. Also ist $E(X(X-1)) = np(n-1)p$. Die zwei Teile zusammen ergeben

$$\begin{aligned}
V(X) &= np(n-1)p - np(np-1) = n^2p^2 - np^2 - n^2p^2 + np \\
&= np(1-p)
\end{aligned}$$

◁

Die Wahrscheinlichkeiten der Binomialverteilung können auch rekursiv, d.h. mit dem Vorgänger bestimmt werden.

Formel 32.6. Rekursionseigenschaft von $B(n,k)$

$$\begin{aligned}
f(0) &= (1-p)^n \\
f(k+1) &= f(k) \frac{(n-k)}{(k+1)} \frac{p}{1-p}
\end{aligned}$$

für $k = 0, 1, \ldots, n-1$.

32.4.2 Typische Fragestellungen

Die Binomialverteilung $B(n, q)$ hat vier potentiell zu bestimmende Grössen:

(1) die Wahrscheinlichkeit $P(X)$,

(2) die elementare Wahrscheinlichkeit p,

(3) die Anzahl Versuche n und

(4) die Anzahl Treffer/Erfolge k.

Daraus ergeben sich vier typische Fragestellungen.

Wahrscheinlichkeit

32.7 Übung Ein Schüler hat für eine Multiple-Choice-Prüfung nichts gelernt. Die Prüfung stellt 10 Fragen mit je 4 Antworten, von denen nur eine richtig ist. Wie gross ist die Wahrscheinlichkeit, zufällig 6 richtige anzukreuzen?
Aus der Frage bestimmen wir $n = 10$, $k = 6$ und $p = 1/4$, denn für jede Frage stehen 4 möglich Antworten zur Verfügung. Die Lösung ist

$$P(X = 6) = f(6) = \binom{10}{6} 0.25^6 \cdot 0.75^4 = 210 \cdot 0.0002 \cdot 0.32 = 0.0162$$

Eine genügende Note bedeutet 6, 7, 8, 9 oder 10 richtige. Wie gross ist die Wahrscheinlichkeit, zufällig genügend zu sein? Man muss die einzelnen Wahrscheinlichkeiten summieren, also

$$\begin{aligned} P(X \geqslant 6) &= P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10) \\ &= 0.0162 + 0.00309 + 0.000386 + 0.000028 + 0.00000095 \\ &= 0.0197 \end{aligned}$$

Diese Zahlen wurden mit einem programmierbaren Rechner berechnet. Von “Hand” ist die Rechnung mühsam. Fazit: sich nicht auf die Prüfung vorzubereiten ist nicht besonders schlau, denn die Wahrscheinlichkeit ist keine 2% genügend zu sein. ◁

Anzahl Versuche

32.8 Übung Eine Fabrik stellt Glühbirnen her, von denen 0.5% defekt sind. Wie viele muss sie herstellen, so dass mindestens eine defekt ist? Das Wort “mindesten” löst den Reflex aus, an das Gegenereignis “kein” zu denken. Also $P(X \geqslant 1) = 1 - P(0)$. Wir kennen $p = 0.005$, $k = 0$ (keine defekt) und $P(0) = 0.99$. Gesucht ist n. Daraus gibt die Formel

$$P(0) = 0.99 = \binom{n}{0} 0.005^0 \cdot 0.995^n$$

und damit $P(X \geqslant 1) = 1 - P(0) = 0.01$

$$0.01 = 1 \cdot 1 \cdot 0.995^n$$

und damit

$$\log(0.01) = n \cdot \log(0.995)$$

und

$$n = \frac{\log(0.01)}{\log(0.995)} = \frac{\ln(0.01)}{\ln 0.995)} = \frac{\lg(0.01)}{\lg(0.995)} = 918.7 \approx 919$$

◁

Einzelwahrscheinlichkeit

32.9 Übung In einem Gugelhopf sind 50 Rosinen verteilt. Wie viele Stücke x darf man schneiden, wenn mit 99%-iger Wahrscheinlichkeit mindestens eine Rosine pro Stück sein sollen? Die Wahrscheinlichkeit für eine Rosine in x Stücken ist $p = 1/x$ und damit $1 - p = (x-1)/x$. Mit dem Gegenereignis $P(k \geqslant 1) = 1 - P(0) \geqslant 0.99$ folgt $P(0) \leqslant 0.01$. Es ist

$$P(0) = \binom{50}{0}\left(\frac{1}{x}\right)^0\left(\frac{x-1}{x}\right)^{50} = \left(\frac{x-1}{x}\right)^{50} \leqslant 0.01$$

und daraus

$$\frac{x-1}{x} \leqslant 0.01^{\frac{1}{50}} = 0.912$$

damit

$$x - 1 \leqslant 0.912x \quad \Leftrightarrow \quad x(1 - 0.912) \leqslant 1 \quad \Leftrightarrow \quad x \leqslant 11.37$$

Es dürfen höchstens 11 Stücke geschnitten werden. $\triangleleft$

Anzahl Treffer

32.10 Übung Ein fairer Würfel wird 20 Mal geworfen. Wir möchten wissen, welche Anzahl k von Sechsen dazu führt, dass die Wahrscheinlichkeit für dieses Ereignis eine Wahrscheinlichkeit von weniger als 5% aufweist. Die Binomialverteilung gibt $P(X \leqslant k)$ wieder. Hier wird aber $P(X > k) < 0.05$ verlangt. Nun ist das Gegenereignis $1 - P(X \leqslant k-1) < 0.05$ und somit $P(X \leqslant k-1) > 0.95$. Die Lösung lässt sich nicht analytisch, d.h. durch elementares Rechnen finden. Für diese Fragestellung muss man in der Wertetabelle der Binomialverteilung die Werte abgleichen.

0	1	2	3	4
0.026084	0.130420	0.328659	0.566545	0.768749
5	6	7	8	9
0.898159	0.962864	0.988746	0.997158	0.999401

An der Stelle 6 ($P(X \leqslant 6)$) ist die (kumulierte) Wahrscheinlichkeit zum ersten Mal grösser als 0.95. Das heisst, dass $k - 1 = 6$ ist und somit das gesuchte k 7 ist. $\triangleleft$

32.4.3 Multinomiale Verteilung**

Die Binomialverteilung beruht auf dem Modell einer Urne mit zwei Sorten von Kugeln, oder mit mehr Sorten aber der Fragestellung von "diese Sorte" (rot) gegen "einer der anderen Sorten" (nicht rot).

Eine Urne enthalte 6 Kugeln, 1 rote, 2 blaue und 3 gelbe. Die Wahrscheinlichkeiten (Ziehen mit Zurücklegen) sind $p_1 = 1/6$, $p_2 = 1/3$ und $p_3 = 1/2$. Es werden 5 Ziehungen gemacht. Wir möchten die Wahrscheinlichkeit wissen, dass 2 Mal rot, einmal blau und zweimal gelb gezogen wird, also $P(x_1 = 2, x_2 = 1, x_3 = 2)$. Für "rot/nicht rot" folgt $P(x_1 = 2) = \binom{5}{2}p_1^2(1-p_1)^3 = \binom{5}{2}p_1^2 \cdot Q_1$. Jetzt gilt für die übrigen 3 Ziehungen mit einer blauen $P(x_2 = 1) = Q_1 = \binom{3}{1}p_2^1(1-p_2)^2 = \binom{3}{1}p_2^1 \cdot Q_2$ und weiter $P(x_3 = 2) = Q_2 = \binom{2}{2}p_3^2(1-p_3)^0$. Zusammengefasst und eingesetzt ergibt sich

$$P(x_1 = 2, x_2 = 1, x_3 = 2) = \binom{5}{2}p_1^2 \cdot \binom{3}{1}p_2^1 \cdot \binom{2}{2}p_3^2(1-p_3)^0$$

Vereinfacht und mit aufgelösten Fakultäten

$$P(x_1 = 2, x_2 = 1, x_3 = 2) = \frac{5!}{2!1!2!} p_1^2 \cdot p_2^1 \cdot p_3^2$$

Daraus folgt die richtige Vermutung

Formel 32.11.

$$P(n_1, ..., n_s) = \frac{n!}{n_1! \cdots n_s!} p_1^{n_1} \cdots p_s^{n_s}$$

mit $n = n_1 + n_2 + \ldots + n_s$ und $p_1 + p_2 + \ldots + p_s = 1$.

Um das Beispiel zu Ende zu führen rechnen wir noch aus

$$P(x_1 = 2, x_2 = 1, x_3 = 2) = \frac{120}{4} \left(\frac{1}{6}\right)^2 \left(\frac{1}{3}\right)^1 \left(\frac{1}{2}\right)^2 = 0.0694$$

32.5 Stetige Zufallsvariablen*

Bis hierhin haben wir diskrete Zufallsvariablen angeschaut, also solche, die nur ganzzahlige Werte annehmen können. Dies ergibt sich in vielen Grundmodelle wie z.B. dem Münzwurf, dem Würfeln, dem Roulette, dem Lotto etc. Eine Zufallsgrösse (Zufallsvariable) kann aber auch reelle Werte annehmen. Man denke hier an die Wartezeit am Postschalter, der Zeigerstellung bei der Draufsicht auf die Uhr u.v.m. Zudem werden Zufallsvariablen reell, wenn man Anteile betrachtet, z.B. Anteil der Schüler, die genügende Noten schreiben (rationale Zahlen). Weil das Rechnen mit reellen Zahlen erheblich einfacher sein kann als der Kalkül mit ganzen Zahlen wählt man Näherungen im Reellen. Wir werden eben die Binomialverteilung durch eine stetige Verteilung annähern.

Die Begriffe, die wir für reelle Zufallszahlen erweitern wollen sind etwa die Wahrscheinlichkeitsfunktion, die Verteilung und die Masszahlen Erwartungswert und Varianz. Es zeigt sich, dass man im Wesentlichen einfach die Summen durch Integrale ersetzt.

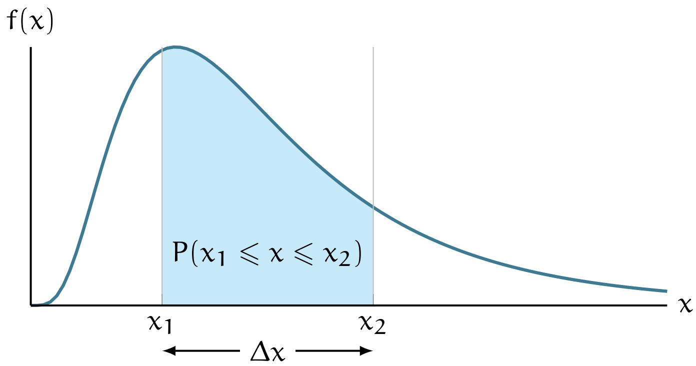
Definition 36. Eine stetige Wahrscheinlichkeitsfunktion $f(x)$, *Wahrscheinlichkeitsdichtefunktion* genannt, besitzt die Eigenschaften:

(1) $f(x) \geqslant 0$,

(2) $P(x_1 \leqslant x \leqslant x_2) = \int\limits_{x_1}^{x_2} f(x)dx$,

(3) $\int\limits_{-\infty}^{\infty} f(x)dx = 1$.

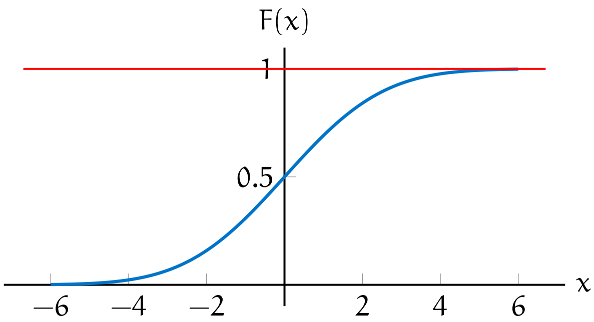
Die Dichte $f(x)$ ist keine Wahrscheinlichkeit; sie muss mit Δx multipliziert werden, um es zu werden. Die Wahrscheinlichkeit ist ein Flächenstück unter der Dichtekurve. In der Abbildung entspricht $P(x_1 \leqslant x \leqslant x_2)$ der Fläche unter der Kurve $f(x)$ zwischen x_2 und x_1.



Definition 37. Die *Verteilungsfunktion* der stetigen Zufallsvariablen X ist

$$F(X) = P(X \leqslant x) = \int_{-\infty}^{x} f(z)dz$$

Die Verteilungsfunktion $F(x)$ ist *Stammfunktion* von der Wahrscheinlichkeitsfunktion $f(x)$.



Die Abbildung zeigt die Verteilungsfunktion einer stetigen Zufallsvariable. Die Funktion geht für wachsendes x asymptotisch nach 1. An einer Stelle a bedeutet $F(a)$ die Wahrscheinlichkeit, dass $x \leqslant a$ ist. Für $P(x_1 \leqslant x \leqslant x_2)$ muss man die Differenz $F(x_2) - F(x_1)$ bilden. Die Inverse von $F(x)$, also die zu $F(x) = p$ gehörigen x-Werte nennt man *Quantile* x_p. Das Quantil $x_{0.5}$ nennt man Median. In der Abbildung ist der Median null.

Eigenschaften 32.1. Stetige Zufallsvariable

- $F'(X) = \frac{dF(x)}{dx} = f(x)$
- $P(X = a) = 0$
- $P(a < X < b) = P(a \leqslant X \leqslant b) = F(b) - F(a)$
- $E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)dx$
- $V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)dx$

Anmerkung 32.2. Bei einer stetigen Zufallsvariable ist $P(X = a)$ null, denn $P(X = a) = P(a \leqslant x \leqslant a) = F(a) - F(a) = 0$. Für eine diskrete Zufallsvariable existiert $P(X = k)$, für eine stetige aber nicht.

Auch stetige Zufallsvariablen haben Erwartungswert und Varianz. In Analogie zur diskreten Zufallsgrösse gelten folgende Definitionen

Definition 38. Momente Hat eine reelle Zufallsvariable X eine Wahrscheinlichkeitsdichtefunktion f, so ist der *Erwartungswert* von X:

$$E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)\,dx.$$

Die *Varianz* von X ist, sofern sie existiert,

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx.$$

Anmerkung 32.3. Erwartungswert und Varianz werden auch als erstes und zweites Moment bezeichnet.

Anmerkung 32.4. Die Varianz ist der Erwartungswert der quadratischen Abweichung vom Mittelwert.

Anmerkung 32.5. Auch für stetige Variablen gibt es den Verschiebungssatz, hier für die Varianz, der Berechnungen einfacher machen kann:

$$V(X) = E((X - E(X))^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - E(X)^2 \ .$$

32.6 Normalverteilung*

Die Normalverteilung beschreibt die berühmte Gauss'sche Glockenkurve. Sie dient auch als Näherungsfunktion für die Binomialverteilung, je nach Parameter besser oder schlechter. Die Binomialverteilung hat den grossen Nachteil, für grosse Zahlen n die Fakultäten bestimmen zu müssen. Deshalb hat man im 18. Jahrhundert nach Wegen gesucht, diese Rechnung zu vereinfachen. Zur Vorbereitung betrachten wir grosse Fakultäten N. Dafür existiert eine relativ einfache Näherung, die aus folgender Überlegung hervorgeht:

$$\begin{aligned}\ln(N!) &= \ln(1 \cdot 2 \cdot 3 \cdot \ldots \cdot N)\\ &= \ln(1) + \ln(2) + \ln(3) + \ldots + \ln(N) = \sum_{k=1}^{N} \ln(k)\end{aligned}$$

Nun machen wir einen Übergang vom diskreten $k \in \mathbb{N}$, d.h ganzzahligen, zum reellen $k \in \mathbb{R}$ mit der Näherung (Integral anstatt Summe):

$$\begin{aligned}\sum_{k=1}^{N} \ln(k) \approx \int_1^N \ln(k)dk\\ &= [x\ln(x) - x]_1^N = N\ln(N) - N + \cancel{1}\\ &\approx N\ln(N) - N\end{aligned}$$

und daraus dann einfach

$$N! = \exp(N\ln(N) - N) = \exp(\ln(N^N) - N) = \frac{N^N}{e^N} = \left(\frac{N}{e}\right)^N$$

Formel 32.1. Näherung Fakultät Mit der Euler'sche Zahl $e = 2.71828\ldots$

$$n! \approx \left(\frac{n}{e}\right)^n$$

N	10	20	50	100	150
Exakt	15.1	42.3	148.5	363.7	605.0
Genähert	13.0	39.9	145.6	360.5	601.6

32.6.1 Näherung von Moivre-Laplace

Die Näherung für grosse n der Binomialverteilung $B(n,p)$ besitzt die Wahrscheinlichkeitsfunktion $f(x)$ gemäss folgender Formel

Satz 32.2.

$$f(x) = P\left(X = k\right) \;=\; B(k \mid p, n) \;\approx\; \frac{1}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{n}{2\sigma^2}\left(\frac{k}{n} - p\right)^2\right)$$

Die (kumulierte) Verteilungsfunktion ist:

Satz 32.3.

$$F(x) = P\left(x_1 \leqslant \frac{X - np}{\sqrt{n}} \leqslant x_2\right) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}} \int_{x_1}^{x_2} \exp\left(\frac{-x^2}{2\sigma^2}\right) dx$$

Definition 39. Die Wahrscheinlichkeitsfunktion $f(x) = \phi(x)$ zur *Standard-Normalverteilung* $F(X) = \Phi(X)$ ist

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

Man nennt sie auch *Gauss'sche Glockenkurve*.

Die zugehörige Verteilung ist die Summe aller Werte der Wahrscheinlichkeitsfunktion. Weil die Zufallsgrösse reell ist, wird aus der Summe das Integral über den gesamte Wertebereich von X. Der Wert $1/\sqrt{2\pi}$ stellt sicher, dass die totale Wahrscheinlichkeit gleich 1 ist[1]

Definition 40. Die *Standard-Normalverteilung* $F(X) = \Phi(X)$ ist

$$F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{(x-\mu)/\sigma} e^{-\frac{1}{2}z^2} dz = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

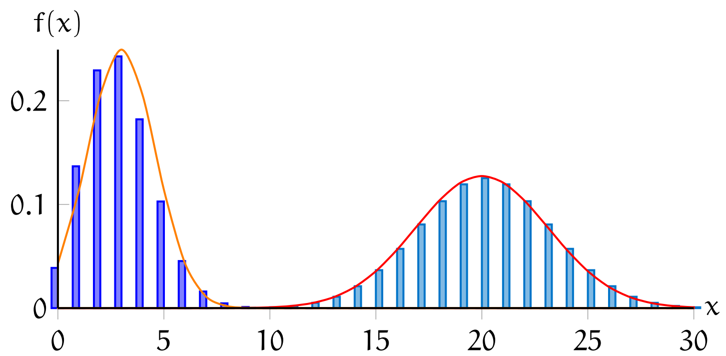
Anmerkung 32.4. Die Standard-Normalverteilung $\Phi(Z)$ für eine Zufallsvariable $Z = \frac{X-\mu}{\sigma}$ ist tabelliert (siehe Seite 32-45), weil nicht elementar berechenbar, und besitzt den Erwartungswert $\mu_Z = E(Z) = 0$ und die Varianz $V(Z) = 1$. Mit $X = Z\sigma + \mu$ kann man Z zurücktransformieren.

[1] $\int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$.

Formel 32.5. Approximation der Binomialverteilung Mit $\mu = np$ und $\sigma = np(1-p)$:

$$P(x_1 \leqslant X \leqslant x_2) = \underbrace{\sum_{k=x_1}^{x_2} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}}_{\text{Binomialverteilung}}$$
$$\approx \underbrace{\Phi\left(\frac{x_2 + 0.5 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - 0.5 - \mu}{\sigma}\right)}_{\text{Normalverteilung}}.$$

Anmerkung 32.6. Man übernimmt also die Parameter μ und σ von der Binomialverteilung für die Normalverteilung. Damit stimmen sie in Lage und Streuung ungefähr überein.



In der Abbildung sieht man zwei Approximationen der binomialen Wahrscheinlichkeitsfunktion durch die Dichte der Normalverteilung. Im linken Fall ist sie nicht besonders gut, weil sowohl n nicht besonders gross ist und p nicht nahe bei 0.5. Im rechten Fall ist die Näherung sehr gut, die rote Kurve schmiegt sich mit geringer Abweichung an.

Wichtig 7. Wann ist die Näherung sinnvoll, weil qualitativ genug gut? Es hat sich die Bedingung nach Laplace eingebürgert, wonach gelten sollte

$$V(X) = n \cdot p \cdot (1-p) > 9.$$

Die Güte der Näherung ist hoch, wenn

- n möglichst gross ist und
- p möglichst nahe bei 0.5.

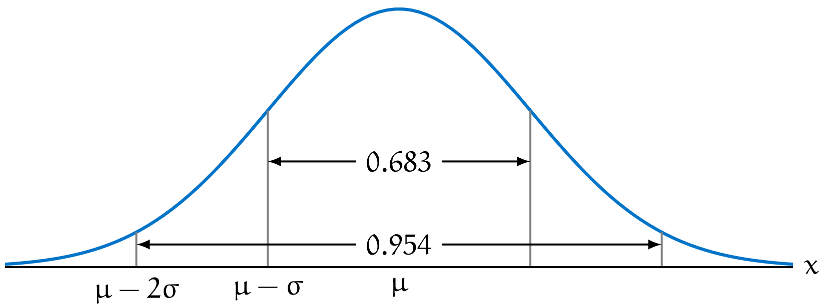
⊣

32.6.2 Eigenschaften der Standard-Normalverteilung

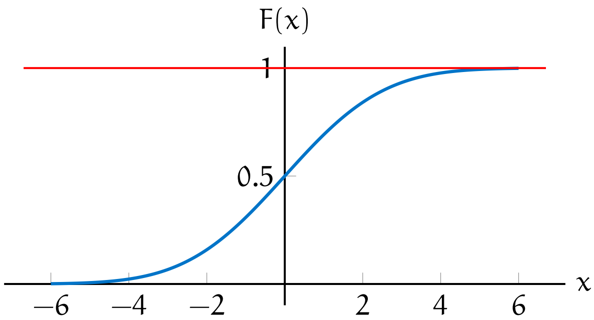
Die Wahrscheinlichkeitsfunktion $f(x)$, siehe Abbildung, ist symmetrisch zur Achse $x = \mu$. Die Fläche unter dieser Kurve ist 1. Die x-Werte liegen im Intervall $(-\infty, \infty)$, auch wenn die $f(x)$ relativ schnell nach 0 streben. Die Kurve hat zwei Parameter, nämlich Mittelwert und Varianz. Der Mittelwert bestimmt die *Lage* der Kurve, die Varianz die *Streuung*.

Eigenschaften 32.7. Normalverteilung

- $P(-a \leqslant X \leqslant -b) = P(b \leqslant X \leqslant a)$ oder $\Phi(-b) - \Phi(-a) = \Phi(a) - \Phi(b)$,
- $P(-a) = 1 - P(a)$ oder $\Phi(-z) = 1 - \Phi(z)$



Die Wahrscheinlichkeitsfunktion oder Dichte der Normalverteilung $\phi(x)$ hat die typische Glockenform. Die Verteilung ist das Integral, d.h. Fläche, dieser Funktion und hat den Wertebereich von $(0, 1)$, wie abgebildet:



32.6.3 Stetigkeitskorrektur

In der Formel 32.5 sieht man sehr gut, dass mit der Näherung eine Umstellung von k, einer natürlichen Zahl, auf x, eine reelle Zahl, stattfindet. Zwei Dinge müssen beim Übergang beachtet werden:
Erstens muss man die Ordnungsrelationen “grösser (gleich)” und “kleiner (gleich)” genau verstehen, bevor man die Näherung verwendet

$$P(X_{BV} < x) = P(X_{BV} \leqslant x - 1) \quad \text{respektive} \quad P(X_{BV} > x) = P(X_{BV} \geqslant x + 1)$$

$$P(X_{BV} \leqslant x) = P(0 \leqslant X_{BV} \leqslant x) \quad \text{respektive} \quad P(X_{BV} \geqslant x) = P(x \leqslant X_{BV} \leqslant n)$$

und besonders

Formel 32.8.

$$P(X_{BV} = x) = P(x - 0.5 \leqslant X_{BV} \leqslant x + 0.5)$$

32.6.4 Beispiel mit Tabelle

32.9 Übung Wir vergleichen eine Binomialverteilung $B(50, 0.3)$ mit der Approximation. Als erstes nehmen wir $P(X = 10)$, also genau 10 Treffer aus 50 Versuchen. Es ist $\mu = np = 15$ und $V(X) = np(1-p) = 10.5$ (nach unserem Kriterium ist ein Approximation zulässig, weil $V(X) > 9$).

$$P(X = 10) = \binom{50}{10} \cdot \left(0.3\right)^{10} \cdot \left(0.7\right)^{40} = 0.0386$$

Die Näherung ist

$$P\left(\frac{9.5 - 15}{\sqrt{10.5}} \leqslant X \leqslant \frac{10.5 - 15}{\sqrt{10.5}}\right)$$

und daher

$$P(-1.7 \leqslant N \leqslant -1.39) = P(1.39 \leqslant N \leqslant 1.7) = \Phi(1.7) - \Phi(1.39) = 0.9554 - 0.9177 = 0.0377$$

Wir zeigen hier eine Ausschnitt der Tabelle 32.7.3 von Seite 32-45.

z	0.00	0.01	0.02	...	0.07	0.08	0.09
1.2	0.8849	0.8869	0.8888	...	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	...	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	...	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	...	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	...	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	...	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	...	0.9693	0.9699	0.9706

Für $\Phi(1.7\,0)$ ist die erste Spalte mit der Zeilenbeschriftung 1.7 die gesuchte Zahl, für $\Phi(1.3\,9)$ sucht man die Zeile mit 1.3 und geht zur Spalte mit 0.09. $\triangleleft$

32.10 Übung Eine Industriebäckerei stellt sehr viele Brote her mit einem Gewicht von x mit einem Mittelwert von $\mu = 1$kg und einer Streuung von $\sigma = 0.05$kg. Wir gross ist die Wahrscheinlichkeit, dass $x \leqslant 0.95$ ist? Gesucht wird $P(x \leqslant 0.95)$, für die Tabelle brauchen wir $z = \frac{x-1}{0.05}$ und damit $x = 0.05z + 1$, eingesetzt $P(0.05z + 1 \leqslant .95)$, daraus $P(0.05z \leqslant -0.05) = P(z \leqslant -1)$. Das ist $P(z \leqslant -1) = 1 - \Phi(1) = 1 - 0.413 = 0.159$. $\triangleleft$

32.6.5 Standardfehler**

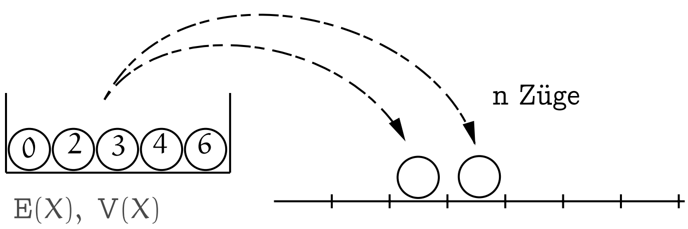
Der wohl wichtigste Satz der Wahrscheinlichkeitstheorie ist der *Zentrale Grenzwertsatz*, von dem es viele Varianten gibt. Ganz einfach ausgedrückt lautet der einfachste:

Satz 32.11. Zentraler Grenzwertsatz Eine Summe von sehr vielen unabhängigen identisch verteilten Zufallsvariablen X_i mit endlicher Varianz ist approximativ normalverteilt.

Anmerkung 32.12. Man beachte, die X_i müssen nicht normalverteilt sein; sie müssen derselben Verteilung angehören, wie es z.B. die gleichverteilten Würfelaugen sind.

Anmerkung 32.13. Messungen sind regelmässig fehlerbehaftet. Führt man viele unabhängige Messungen (mit denselben Instrumenten) durch, so sind die Fehler als Abweichung vom wahren Wert als normalverteilt anzusehen.

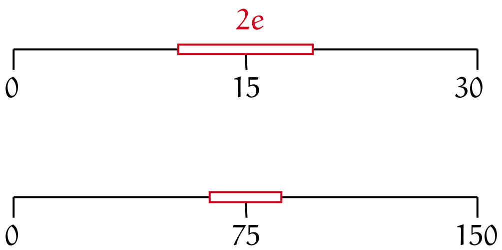
32.14 Übung Wir betrachten eine Urne mit fünf Kugeln. Sie tragen Zahlen. Alle Kugeln haben dieselbe Wahrscheinlichkeit gezogen zu werden.



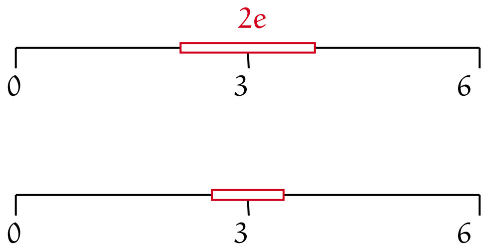
Das Urnenmodell hat den Erwartungswert $E(X) = \frac{1}{5}(0+2+3+4+6) = 3$. Die Varianz ist $V(X) = \frac{1}{5}((-3)^2+1+0+1+9) = 20/5 = 4$. Die Standardabweichung ist $\sigma = \sqrt{V(X)} = \sqrt{4} = 2$.
Unser Experiment soll das n-fache Ziehen von Kugeln mit Zurücklegen sein, wobei die Zufallsvariable Y_n die Summe der gezogenen Zahlen sein soll, also $Y_n = X_1 + X_2 + \ldots + X_n$. Wir wählen $n = 25$. Die Summe Y_n kann theoretisch von minimal 0 bis maximal 150 reichen. Es ist aber eher anzunehmen, dass Y_n in der Nähe von $E(Y) = 25 \cdot E(X) = 25 \cdot 3 = 75$ zu liegen kommt. Die Streuung, also Standardabweichung, von Y ist $\sigma_Y = \sqrt{n} \cdot \sigma_X$ (siehe Eigenschaften 32.8 auf Seite 32-16). ◁

Definition 41. Der *Standardfehler* einer Summe Y mit n Summanden von Zufallsgrössen X_i einer Verteilung mit Standardabweichung σ ist die Grösse

$$e = \sigma_Y = \sqrt{n} \cdot \sigma.$$



Wir vergleichen das Intervall um den Erwartungswert für unser $n = 25$ und $e = 10$ mit dem Fall $n = 5$. Relativ gesehen verringert sich das Intervall mit zunehmendem n. Nun verwenden wir die Normal-Näherung. Aus der Tabelle 32.7.3 von Seite 32-45 lesen wir bei 1.00, das ist σ in standardisierter Form, 0.8413 ab. Für $P(X \leqslant -\sigma)$ würden wir $1 - 0.8413$ erhalten. Für $P(\mu - \sigma < Y < \mu + \sigma)$ also $2 \cdot (0.8413 - 0.5) = 0.6826$.



Wir betrachten anstelle der Summe den Mittelwert, also Summe geteilt durch Anzahl Züge. Wir übernehmen die Abbildung und beschriften sie neu. Man erkennt, dass der Standardfehler des Mittelwertes mit wachsendem n geringer wird. Aufgrund der Eigenschaften 32.8 ist klar, dass gilt $E(Y/n) = E(Y)/n$ und $V(Y/n) = V(Y)/n$ und damit

$$\sigma_{\overline{Y}} = \frac{\sigma_Y \sqrt{n}}{n} = \frac{\sigma}{\sqrt{n}}$$

Satz 32.15. Der Standardfehler $e_{\overline{Y}}$ des Mittelwertes $\overline{Y}$ ist

$$\sigma_{\overline{Y}} = \frac{\sigma}{\sqrt{n}}$$

Anmerkung 32.16. Dies ist eine äusserst wichtige Erkenntnis. Bei gegebener Wahrscheinlichkeit $P(\mu - e_{\overline{Y}}\sigma < \overline{Y} < \mu - e_{\overline{Y}})$ wird das Intervall $\mu \pm e_{\overline{Y}})$ mit n zwar immer kleiner, aber nur nach dem Wurzelgesetz. Wenn man von 10 Würfen auf 100 Würfe geht, den Aufwand also verzehnfacht, dann wird der Standardfehler nur um $\sqrt{10} \approx 3$ mal kleiner.

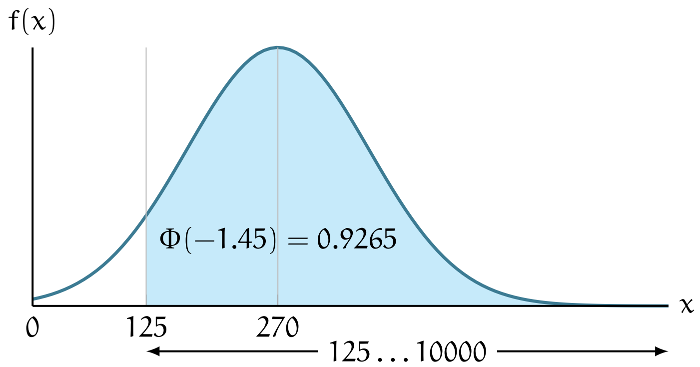
Für den Mittelwert von n-fach wiederholten Zufallsexperimenten besagt der Zentrale Grenzwertsatz, dass die Wahrscheinlichkeit, dass $\overline{y}$ im Intervall $\mu \pm e_{\overline{Y}}$ zu liegen kommt, 0.683 ist. Das ist der Wert aus der Tabelle 32.7.3 für $\Phi(1) - \Phi(-1)$ und $\Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1 = 2 \cdot 0.8413 = 0.683$.

32.17 Übung Beim (europäischen) Roulette gibt es insgesamt 37 Elementarereignisse, d.h. die Zahlen 0,1,2, ...,36. (Beim amerikanischen gibt es noch die Doppelnull, also 38.) Neben den Zahlen kann man auch auf gerade/ungerade, hoch/tief und rot/schwarz setzen. Wir analysieren nur das Setzen auf rot oder schwarz. Es werde 10,000 mal gedreht, der Einsatz sei immer 1 Geldeinheit, die Auszahlung jeweils 2 Geldeinheiten. Wir betrachten die Zufallsgrösse Nettogewinn der Spielbank.

Das analoge Urnenmodell ist eine Urne mit 19 Kugel "+1" und 18 Kugel "-1". In 19 von 37 Fällen behält die Bank den Einsatz, in 18 von 37 bezahlt sie eine Geldeinheit. Der Erwartungswert des Gewinns eines Spieles für die Bank ist $E(X) = \frac{19}{37} \cdot 1 + \frac{18}{37}(-1) = \frac{1}{37}$. Die Varianz ist $V(X) = \frac{19}{37} \cdot (1 - \frac{1}{37})^2 + \frac{18}{37} \cdot (-1 - \frac{1}{37})^2 \approx 1$.

Eine mögliche Frage ist: Wie gross ist die Wahrscheinlichkeit, dass die Spielbank einen Gewinn von mindestens 150 Geldeinheiten macht?

Der Standardfehler der Gewinnsumme ist $e_Y = \sqrt{10000} \cdot 1 = 100$, der Erwartungswert $E(Y) = 10000 \cdot \frac{1}{37} = 270$. Mit 68.3% Wahrscheinlichkeit ist der Gewinn Y zwischen $y = 270 \pm 100$. Für unsere Frage kommt wieder der Zentrale Grenzwertsatz zum tragen mit der Standard-Normalverteilung als Tabelle. Gesucht ist $P(Y \geqslant 125) = 1 - P(Y \leqslant 125)$. Deshalb transformieren wir Y mit $Z = \frac{Y-270}{100}$, speziell $z = \frac{125-270}{100} = -1.45$. $P(z \geqslant -1.45) = 1 - \Phi(-1.45) = \Phi(1.45)$, aus der Tabelle $\Phi(1.45) = 0.9265$. ◁

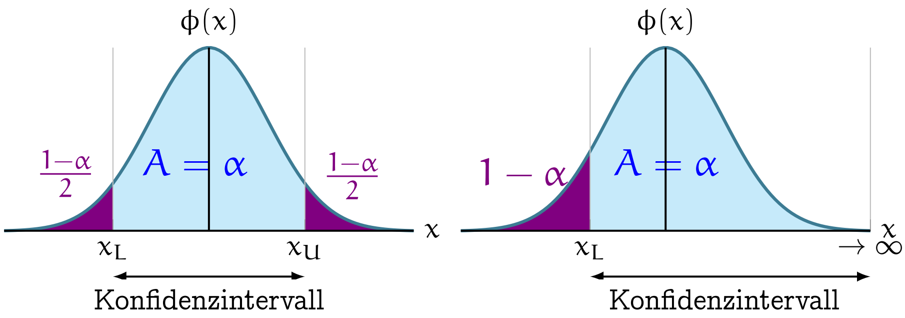


Das Beispiel stellt sich graphisch wie in der Abbildung dar: Uns interessiert die Strecke von 125 bis zum maximal möglichen, also 10000 mal gewonnen. Weil uns ein Resultat "grösser als" oder "mindestens" interessiert, ist die Wahrscheinlichkeit einseitig. In der Abbildung sieht man auch, wie eine Zufallsgrösse, die nur positiv, grösser als null, sein kann, durch die Normalverteilung, die von $-\infty$ bis ∞ nur angenähert wird. Die Fläche der Funktion im negativen Bereich wird vernachlässigt. Man könnte dies korrigieren, indem man die Fläche neu normiert. Man nennt das Stutzen.

Konfidenzintervall

Konfidenzintervalle sind eine Art Verallgemeinerung des Standardfehlers, der ja das Intervall $\mu \pm \sigma$ umfasst und mittels Normalverteilung damit 68.3% abdeckt.

Nun kann man das Intervall in Abhängigkeit der Wahrscheinlichkeit festlegen und anstatt 68.3% etwa 90, 95 oder 99% wählen. Wie aus den obigen Beispielen ersichtlich, kann man das Intervall symmetrisch um den Mittelwert wählen; dann ist es zweiseitig. Oder man wählt ein Intervall einseitig, wie beim Roulette-Beispiel.



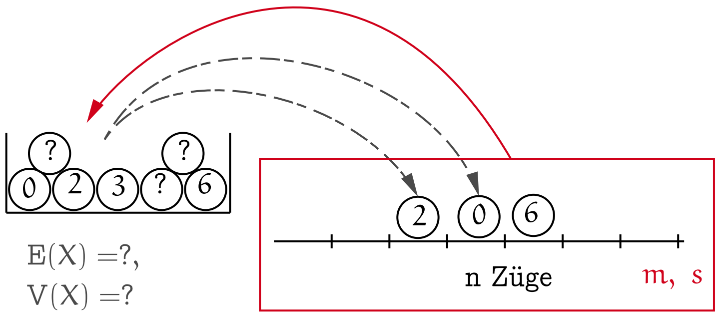
Definition 42. Ein Intervall (x_L, x_U) mit der Eigenschaft

$$P(x_L \leqslant Y \leqslant x_U) = 1 - \alpha$$

heisst *Konfidenzintervall* für den Zufallsvariable Y zum Konfidenzniveau $1 - \alpha$.

Anmerkung 32.18. Übliche Konfidenzniveaus sind 90%, 95% und 99%.

In unseren Beispielen war das Urnenmodell bekannt, damit auch der Mittelwert und Varianz von X und der Summe Y. In vielen Fällen ist aber nur die Realisierung, also das Resultat eines Zufallsversuchs bekannt. Damit ist der Mittelwert m und die Standardabweichung s eine Zufallsvariable. Man verwendet Konfidenzintervalle, um solche Parameter zu schätzen. Solche Fragestellungen werden in der beurteilenden Statistik behandelt.



Von hier kann man ins nächste Kapitel wechseln, wenn man sich für die beurteilende Statistik interessiert.

32.7 Ein paar weitere Verteilungen**

Der Zufalls kann sich in viele Möglichkeiten zeigen, vom Spiel zur Realität mit Unfällen und Vorkommnisse von Höherer Gewalt, auf englisch *acts of God*. Wir haben gesehen, dass Zufallsgrössen Merkmale verschiedener Skalen aufweisen, also von ordinal zu kardinal, oder diskret zu stetig, von ganzzahlig zu reellen Zahlen. Wir haben im Wesentlichen vier Verteilungen betrachtet:

- die Gleichverteilung beim Würfel,
- die Binomialverteilung bei der Urne und

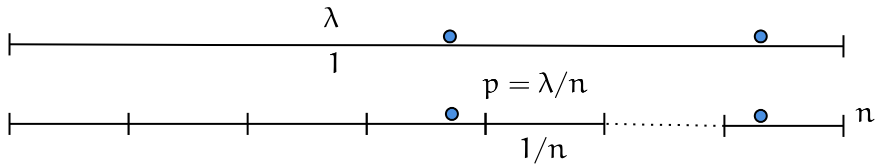
- die Normalverteilung als Vereinfachung und Verallgemeinerung der Binomialverteilung.

Es gibt also eine Korrespondent zwischen Zufallsmodell und Verteilung. Andere Verteilungen folgen also aus Zufallsmodellen. Im folgenden betrachten wir ein paar Modell und deren Verteilung. Im nächsten Kapitel zeigen wir dann, wie man aus Zufallsgrössen mit Verteilung dann die Verteilung von Funktionen von Zufallsgrössen bestimmt.

32.7.1 Poissonverteilung, "Gesetz der kleinen Zahl"

Die Poissonverteilung modelliert ganze positive Zahlen, die zeitlich (oder örtlich) selten eintreten. Das berühmteste Beispiel sind preussische Soldaten, die durch Hufschlag eines Pferdes zu Tode kommen. Somit eignet sich das Modell auch für die Anzahl Unfälle an einer Kreuzung, Anzahl Tippfehler pro Seite, Anzahl Blitzeinschläge pro Fläche usw.

Angenommen, innerhalb des Intervalls der Länge 1 gibt es im Mittel λ Ereignisse. Die Wahrscheinlichkeit, ein Ereignis im Intervall der Länge $1/n$ zu finden, ist proportional zur Länge des Intervalls mit Proportionalitätsfaktor λ, also λ/n. Mann nennt λ *Intensität* oder *Rate*.



Das Eintreten eines Ereignisses im Intervall w wird nicht beeinflusst von Ereignissen, die in der Vorgeschichte stattgefunden haben. Die Ereignisse sind unabhängig.

Somit ist die Wahrscheinlichkeit, ein Ereignis im Intervall $1/n$ zu finden, gegeben als

$$p_1 = \lambda \cdot \frac{1}{n}$$

sowie die Wahrscheinlichkeit eines leeren Intervalls durch $p_0 = 1 - p_1 = 1 - \lambda \cdot \frac{1}{n}$.

Die Wahrscheinlichkeit, k Ereignisse (Kugeln) auf n Plätze zu verteilen, ist binomial mit

$$P(X = k) = \binom{k}{n} (p)^k (1-p)^{n-k}$$

mit $p = \frac{\lambda}{n}$

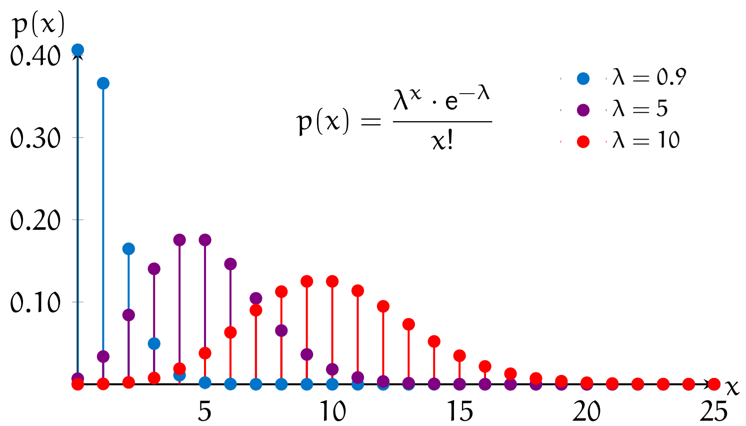
$$P(X = k) = \binom{k}{n} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

Nun wollen wir den Übergang machen, so dass die Ereignisse irgendwo stattfinden können.

D.h. wir machen die Intervalle verschwinden klein mit dem Grenzübergang $n \to \infty$:

$$\begin{aligned}
\lim_{n\to\infty} P(X=k) &= \lim_{n\to\infty} \binom{k}{n} (p)^k (1-p)^{n-k} \\
&= \lim_{n\to\infty} \binom{k}{n} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \left(\frac{\lambda^k}{k!}\right) \left(\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\right) \left(1-\frac{\lambda}{n}\right)^{n} \left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \underbrace{\left(\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}\right)}_{\to 1} \underbrace{\left(1-\frac{\lambda}{n}\right)^{n}}_{\to e^{-\lambda}} \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{\to 1} \\
&= \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

Da die Poissonverteilung ein Grenzwert der Binomialverteilung ist, und diese wiederum mit der Normalverteilung als Grenzübergang verbunden ist, gibt es auch einen Zusammenhang zwischen Poisson- und Normalverteilung. Man siehe dazu die Abbildung mit den Verläufen.



Wir bestimmen die Momente der Poissonverteilung. Mit der Definition des Erwartungswerts, wir summieren ab $k = 1$, denn für $k = 0$ ist der Summand null, folgt:

$$E(K) = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda}{k} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \underbrace{\sum_{j=0}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}}_{=1} = \lambda$$

Der Term in der letzten Summe ist gerade die Wahrscheinlichkeitsfunktion der Poisson-

Verteilung, die zwingend 1 ergeben muss. Mit dem Verschiebungssatz folgt für die Varianz

$$\begin{aligned}
V(K) + \lambda^2 &= \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{(k)!} e^{-\lambda} \\
&= \sum_{j=0}^{\infty} (j+1) \cdot \frac{\lambda^{j+1}}{j!} e^{-\lambda} \\
&= \lambda \underbrace{\sum_{j=0}^{\infty} j \cdot \frac{\lambda^j}{j!} e^{-\lambda}}_{E(K)} + \lambda \underbrace{\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda}}_{=1} \\
&= \lambda^2 + \lambda
\end{aligned}$$

Und damit

$$V(K) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Formel 32.1. Poissonverteilung Die Wahrscheinlichkeiten der Poissonverteilung mit der diskreten Zufallsvariablen $K \geqslant 0$ ist

$$f(k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Es ist $E(k) = \lambda$ und $V(k) = \lambda$.

32.2 Übung In einer Anlage sind Teile verbaut, die ausfallen können. Aus Erfahrung weiss man, dass pro Woche 1.5 Teile ausgewechselt werden müssen. Wie gross ist die Wahrscheinlichkeit, dass in einer Woche 4 Teile zu ersetzen sind? Die Intensität ist $\lambda = 1.5$. Die Wahrscheinlichkeit ist $P(x = 4) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{1.5^4 e^{-1.5}}{4!} = \frac{5.063 \cdot 0.223}{24} = 0.0471 \approx 5\%$. $\triangleleft$

32.3 Übung Nun mit der gleichen Ausgangslage die Frage nach der Wahrscheinlichkeit, dass innerhalb von zwei Wochen kein Ausfall eintritt und allgemein innerhalb von t Wochen. Der Prozess ist gedächtnislos, d.h. die Ausfälle der Vorwoche haben keinen Einfluss auf die Ausfälle der Folgewoche. Somit ist

$$P_0(2) = P_0(1) \cdot P_0(1) = \frac{\lambda^0 e^{-\lambda}}{0!} \cdot \frac{\lambda^0 e^{-\lambda}}{0!} = \frac{\lambda^{2 \cdot 0} e^{-\lambda \cdot 2}}{0!}$$

Allgemein mit t folgt

$$P_0(t) = e^{-\lambda \cdot t}$$

$\triangleleft$

Die Wahrscheinlichkeiten der Poissonverteilung kann man als Rekursion berechnen, d.h. mit dem Vorgänger auf den Nachfolger schliessen. Es ist

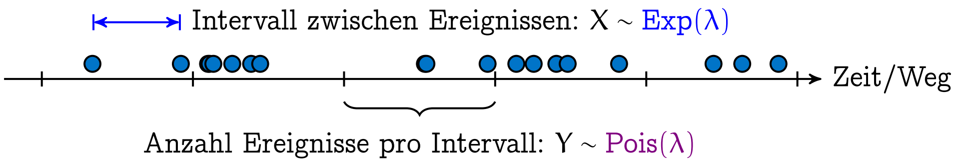
$$f(k+1) = \frac{\lambda}{k+1} f(k)$$

denn

$$f(k+1) = \frac{\lambda}{k+1} f(k) = \frac{\lambda \cdot \lambda^k e^{-\lambda}}{(k+1)k!} = \frac{\lambda^{k+1} e^{-\lambda}}{(k+1)!}$$

32.7.2 Exponentialverteilung

Die Exponentialverteilung leitet sich aus dem obigen Poisson-Prozess ab. Die Frage ist nach dem Modell, das die Länge oder Dauer X beschreibt, das zwischen zwei poissonverteilten Ereignissen liegt. Wir erkennen sofort, dass diese Zufallsgrösse X eine stetige, positive Grösse in $\mathbb{R}$ ist.



Die Wahrscheinlichkeit, dass X höchstens den Wert x annimmt, berechnet sich als

$$P(X \leqslant x) = 1 - P_0$$

mit P_0 der Wahrscheinlichkeit, dass kein Ereignis im Intervall der Länge x auftritt. Diese Grösse ist die poissonverteilte Wahrscheinlichkeit von $Y = 0$, nicht mit der Ereignisrate λ, die sich auf das Einheitsintervall bezieht, sondern mit der skalierten Rate $x \cdot \lambda$. Somit ist

$$P_0 = \frac{(\lambda x)^0}{0!} e^{-\lambda x} = e^{-\lambda x}$$

Somit ist

$$P(X \leqslant x) = F(x) = 1 - e^{-\lambda x}$$

Die Dichte ist die Ableitung der Verteilung, also

$$f_X(x) = \lambda e^{-\lambda x}$$

Zur Verteilung bestimmen wir Mittelwert und Varianz. Mir den Definitionen 38 folgt:

$$E(X) = \lambda \int_{-\infty}^{\infty} x \cdot \exp(-\lambda x) dx$$

Mit partieller Integration $(u'v = uv - uv')$ folgt

$$E(X) = \lambda x \frac{-1}{\lambda} \exp(-\lambda x)\Big|_0^{\infty} - \lambda \int_0^{\infty} \exp(-\lambda x) dx = 0 - \lambda \frac{1}{\lambda^2} \exp(-\lambda x)\Big|_0^{\infty} = \frac{-1}{\lambda}(0-1) = \frac{1}{\lambda}$$

Die Varianz bestimmt sich am einfachsten mit dem Verschiebungssatz und partieller Integration sowie Einsetzen des Erwartungswerts ergibt für den ersten Term:

$$\begin{aligned}
V(X) + \mu^2 &= \lambda \int_0^{\infty} x^2 \exp(-\lambda x) dx \\
&= \lambda x^2 \frac{-1}{\lambda} \exp(-\lambda x)\Big|_0^{\infty} - \lambda \int_0^{\infty} 2x \frac{-1}{\lambda} \exp(-\lambda x) dx \\
&= 0 + \frac{1}{\lambda} \int_0^{\infty} 2x \exp(-\lambda x) dx \\
&= \frac{2}{\lambda} E(X) = \frac{2}{\lambda}\mu = \frac{2}{\lambda^2}
\end{aligned}$$

Damit folgt

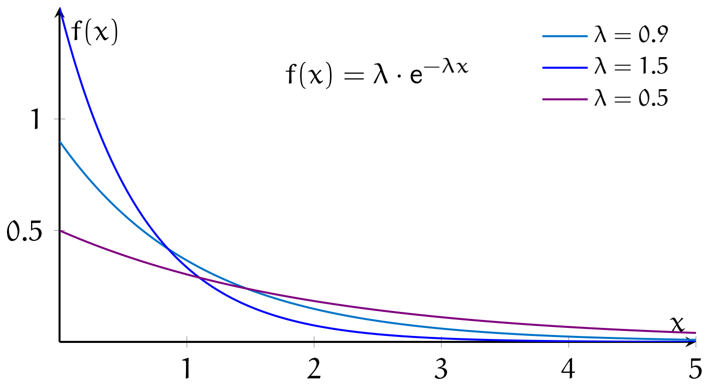
$$V(X) = \frac{2}{\lambda^2} - \mu^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Formel 32.4. Exponentialverteilung Die Wahrscheinlichkeitsdichte der Exponentialverteilung mit der stetigen Zufallsvariablen $X \geqslant 0$ ist

$$f(k) = \lambda e^{-\lambda x}.$$

Es ist $E(X) = \frac{1}{\lambda}$ und $V(k) = \frac{1}{\lambda^2}$.

Graphisch sieht die Verteilung wie folgt aus



32.5 Übung Ein Bauteil hat eine konstante Ausfallrate. D. h. die Wahrscheinlichkeit, dass es innerhalb eines Jahres defekt wird, ist unabhängig vom Alter des Bauteils. Es ist bekannt, dass am Ende des ersten Jahres 10% der Geräte ausgefallen sind. Wann ist die Hälfte der Bauteile defekt? Das Modell ist die Exponentialverteilung, weil hier die Ausfallsrate konstant ist. Es gilt die Wahrscheinlichkeit $F(x) = 1 - \exp(-\lambda x)$ und $F(x = 1) = 0.1 = 1 - \exp(-\lambda x)$. Somit ist $\lambda = \ln(1 - 0.1) = \ln(0.9) = -0.105$. Jetzt suchen wir x. Es folgt $F(x) = 1 - \exp(-0.105x) = 0.5$ und $\ln(0.5) = -0.105x$ und damit $x = \frac{-0.693}{-0.105} \approx 6.6$. Nach 6.6 Jahren ist die Hälfte defekt. $\triangleleft$

32.6 Übung Eine Fussgängerin benötigt 10s, um die Strasse zu überqueren. Autos kommen im Schnitt alle 8s vorbei. Wie gross ist die Wahrscheinlichkeit, dass ein ausreichende zeitliche Lücke entsteht? Mit der Exponentialverteilung gilt $\lambda = 1/8$. Gefragt wird nach $F(X \geqslant 10) = 1 - F(X < 10) = 1 - [1 - \exp(-1/8 \cdot 10)] = \exp(-1/8 \cdot 10) \approx 0.29$. Die Wahrscheinlichkeit ist 29%, also recht hoch. Wenn alle 5s ein Auto kommt, dann wird die Wahrscheinlichkeit$\exp(-1/5 \cdot 10) \approx 0.14$, also halb so viele. $\triangleleft$

32.7.3 Paretoverteilung

Die Paretoverteilung beschreibt das Einkommen der Bürger eines Landes, das über einem bestimmten Schwellenwert liegt. Diese Verteilung ist auch in der Versicherungswirtschaft sehr beliebt zur Modellierung von Schadenhöhen. Mit gewissen Parameterwerten wird die Verteilung stark endlastig, wie es für Naturkatastrophen üblich ist.

Formel 32.7. Eine stetige Zufallsvariable X ist pareto-verteilt $\mathrm{Par}(k, x_{\min})$ mit den Parametern $k > 0$ und $xmin > 0$ mit der Wahrscheinlichkeitsdichte

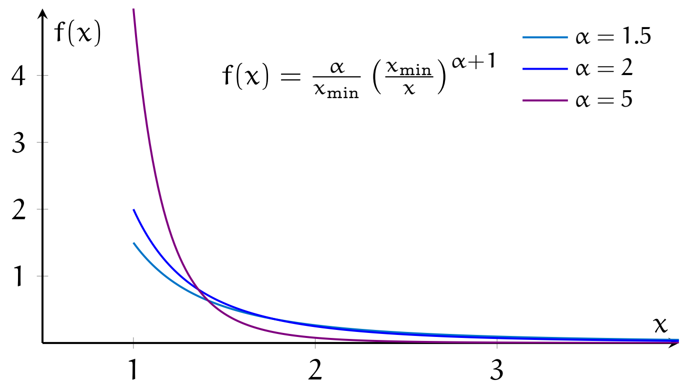
$$f(x) = \begin{cases} \frac{\alpha}{x_{\min}} \left(\frac{x_{\min}}{x}\right)^{\alpha+1} & x \geqslant x_{\min} \\ 0 & x < x_{\min} \end{cases}$$

Der Erwartungswert ergibt sich zu:

$$E(X) = \begin{cases} x_{\min} \frac{\alpha}{\alpha - 1} & \alpha > 1 \\ \infty & \alpha \leqslant 1 \end{cases}$$

Die Varianz ergibt sich zu

$$V(X) = \begin{cases} x_{\min}^2 \left(\frac{k}{k-2} - \frac{k^2}{(k-1)^2}\right) = x_{\min}^2 \frac{k}{(k-2)(k-1)^2} & k > 2 \\ \infty & k \leqslant 2 \end{cases}$$



Aus der Versicherungswirtschaft sind folgende Werte für den Paretoparameter α nach Versicherungstyp für eine Schadenhöhe bekannt. Je kleiner der Parameter desto endlastiger die Kurve und damit die Wahrscheinlichkeit grosser Schäden.

Typ	α
Krankenversicherung	3.0 ...
Motorfahrzeughaftpflicht	2.8 ...
Haftpflicht	2.0 ...
Unfallversicherung	2.0 ...
Motorfahrzeugkasko-Hagel	1.85 ...
Transport	1.5 ...
Feuer	1.5 ... 2.5
Sturm	0.8 ... 1.3
Erdbeben	0.5 ... 0.8

Wir beenden hier das Kapitel. Im nächsten werden noch ein paar weitere Verteilungen besprochen, nachdem Methoden zu deren Bestimmung eingeführt werden.

Aufgaben

3.8 Anlässlich einer Umfrage unter 750 Personen geben 225 der Befragten an, im Jahr 2019 eine Flugreise unternommen zu haben.

(a) Wie gross ist die Wahrscheinlichkeit, dass von 3 Personen mindestens 1 Person geflogen ist? [54 ‰]

Von den übrigen 525 Personen sind 40% Vegetarier. Insgesamt sind 68% der Befragten keine Vegetarier. Eine der befragten Personen werde nun zufällig ausgewählt. Wie gross ist die Wahrscheinlichkeit, dass

(b) sie 2019 eine Flugreise unternahm und zu den Vegetariern gehört? [54 ‰]

(c) sie 2019 eine Flugreise unternahm, falls sie zu den Vegetariern gehört? [27 ‰]

(d) sie 2019 eine Flugreise unternahm oder zu den Vegetariern gehört? [27 ‰]

(e) sie 2019 eine Flugreise unternahm und zu den Vegetariern gehört oder beides nicht der Fall ist? [27 ‰]

8 **(a)** "Mindestens", d.h. Gegenereignis $P(\text{mindestens } 1) = 1 - P(\text{keine})$. $p_g = \frac{225}{750} = 0.3$ und $p_{ng} = 0.7$. $P(\text{keine}) = 0.7^3 = 0.343$. Deshalb $P(\text{mindestens } 1) = 1 - 0.343 = 0.657$.

(b) Von den nicht-geflogenen sind $0.4 \cdot 525 = 210$ Vegetarier. Insgesamt sind $1-0.68 = 0.32$ Vegetarier, also $0.32 \cdot 750 = 240$. Somit sind $30 = 240 - 210$ geflogene Vegetarier. Tabelle

	geflogen	nicht geflogen	total
Anzahl	225	525	750
Vegetarier	30	210	240
Nicht Veg.	195	315	510

Antwort: $\frac{30}{750} = 0.04$.

(c) Antwort: $\frac{225+210}{750} = 0.58$.

(d) Antwort: $\frac{30+315}{750} = 0.46$

3.9 Erkan ist der Elfmeterschütze seiner Fussballmannschaft. Zu Spielbeginn verwertet er einen Elfmeter mit einer Wahrscheinlichkeit von 90% zu einem Tor. Nachher sinkt seine Treffsicherheit bis zum Spielende nach 90 Minuten linear auf 65%. Erkan schiesst nach der 18., nach der 54. und nach der 81. Minute je einen Elfmeter. Mit welcher Wahrscheinlichkeit

(a) verwertet er den dritten Elfmeter? [24 ‰]

(b) verwertet er keinen Elfmeter? [48 ‰]

(c) verwertet er genau einen Elfmeter? [48 ‰]

(d) verwertet er im Training genau drei von fünf Versuchen, wenn seine Treffsicherheit dort konstant 90% beträgt? [48 ‰]

(e) Für welche Elfmeter hätte die Mannschaft im Spiel auf Roland als Schützen setzen sollen, dessen Treffsicherheit $g(t)$ folgender Zeit-Wahrscheinlichkeits-Funktion gehorcht? [24 ‰]

$$g(t) = 0.2 \cdot \ln(-t + 100)$$

9 Die Formel für die Wahrscheinlichkeit ist gemäss $y = m \cdot t + b$, hier $y = -0.25 \cdot t/90 + 0.9$. (Prüfung: bei $t = 0$ folgt 0.9, bei $t = 90$ folgt 0.65) $y_1 = -0.25 \cdot 18/90 + 0.9 = 0.85$, $y_2 = -0.25 \cdot 54/90 + 0.9 = 0.75$ und $y_3 = -0.25 \cdot 81/90 + 0.9 = 0.675$.

(a) $p = 0.675$

(b) $p = (1 - 0.85)(1 - 0.75)(1 - 0.675) = 0.0122$

(c) $P = 0.85(1 - 0.75)(1 - 0.675) + (1 - 0.85)(0.75)(1 - 0.675) + (1 - 0.85)(1 - 0.75)0.675 = 0.131$

(d) Binomialverteilung $B(5, 0.9)$, also $P = \binom{5}{3} 0.9^3 0.1^2 = 9 \cdot 0.729 \cdot 0.01 = 0.0656$

(e) Wir rechnen die Funktion für $t = (18, 54, 81)$ aus, das gibt $g(\{18, 54, 81\}) = \{0.88, 0.77, 0.59\}$. Auf die ersten zwei, denn 0.88>0.85 und 0.77>0.75.

3.10 Eine Urne enthält 10 Kugeln, nämlich 1 weisse, 2 blaue, 3 rote und 4 schwarze. Es wird viermal eine Kugel herausgezogen, ihre Farbe notiert und wieder zurückgelegt.

(a) Mit welcher Wahrscheinlichkeit haben alle vier Kugeln die gleiche Farbe ?

(b) Mit welcher Wahrscheinlichkeit sind alle vier Kugeln von verschiedener Farbe?

Nun werden vier Kugeln gleichzeitig gezogen.

(c) Mit welcher Wahrscheinlichkeit sind alle gezogenen Kugeln blau oder rot?

(d) Mit welcher Wahrscheinlichkeit halten wir mehr blaue als weisse Kugeln in der Hand?

10 **(a)** Die "gleiche Farbe" bedeutet die disjunkten Ereignisse (w,w,w,w), (b,b,b,b), (r,r,r,r) und (s,s,s,s). Die Wahrscheinlichkeiten sind nach Anteil an der Urne $p_w = 1/10 = 0.1$, $p_b = 0.2$, $p_r = 0.3$ und $p_s = 0.4$. Es ist wegen der Unabhängigkeit $P((w,w,w,w)) = 0.1^4$, $P((b,b,b,b)) = 0.2^4$, $P((r,r,r,r)) = 0.3^4$ und $P((s,s,s,s)) = 0.4^4$. Total $P(\text{gleichfarbig}) = 0.1^4 + 0.2^4 + 0.3^4 + 0.4^4 = 0.0001 + 0.0016 + 0.0081 + 0.256 = 0.0354$.

(b) Gesucht $P((w,b,r,s)) = 0.1 \cdot 0.2 \cdot 0.3 \cdot 0.4 = 0.0024$.

(c) Blau oder rot bedeutet, fünf Kugel von 10, also $p = 0.5$. Damit $P(b oder r) = 0.5^4 = 0.0625$.

(d) "Mehr blaue als weisse" sind die Ereignisse (w,b,b,x) mit $x \in \{b,r,s\}$, also (w,b,b,b), (w,b,b,r) und (w,b,b,s). Mit $P((w,b,b,b)) = 0.1 \cdot 0.2^3$, $P((w,b,b,r)) = 0.004 \cdot 0.3$, $P((w,b,b,s))) = 0.004 \cdot 0.4$ und zusammen $P = 0.004(0.2 + 0.3 + 0.4) = 0.004 \cdot 0.9 = 0.0036$.

3.11 An einer Cocktail-Bar kontrollieren zwei Angestellte die Gäste auf Volljährigkeit. Der Angestellte A begutachtet 40% aller Ausweise und die Angestellte B die restlichen 60%. Die Wahrscheinlichkeit, dass A den gefälschten Ausweis eines inderjährigen Gastes als solchen erkennt, ist 50%. Die Erfolgsquote von B beträgt 30%.

(a) Wie gross ist die Wahrscheinlichkeit, dass ein minderjähriger Gast nicht erwischt wird?

(b) In einer Partygruppe von 10 Personen hat es genau einen minderjährigen Gast. B wählt aus dieser Gruppe zufällig drei Personen aus und kontrolliert sie. Wie gross ist die Wahrscheinlichkeit, dass sie ihn entdeckt?

(c) Einem minderjährigen Gast ist es gelungen, an der Bar harten Alkohol zu rhalten. Wie gross ist die Wahrscheinlichkeit, dass er von A kontrolliert wurde?

(d) Wie viele Minderjährige müssen es versuchen, damit die Wahrscheinlichkeit, dass mindestens einer Alkohol erhält, 99.99% übersteigt?

11 **(a)** Zur Übersicht kann man eine Tabelle, einen Ereignisbaum oder ein Mengendiagramm machen. Z.B. Tabelle

	A	B	Total
Minderj. Gäste	40	60	100
Erkannt	20	18	38
Nicht erkannt	20	42	62

Die Nicht-Erkannten sind 62 von Hundert, also 0.62=62%.

(b) WS ausgewählt zu werden (Binomialverteilung): $p = \binom{3}{1}0.1 \cdot 0.9^2 = 3 \cdot 0.1 \cdot 0.81 = 0.243$. WS erkannt zu werden von B. $\frac{18}{60} = 0.3$. Zusammen $P = 0.429 \cdot 0.3 = 0.129$.

(c) Aus der Tabelle: $p = \frac{20}{62} = 0.323$.

(d) "Mindestens", deshalb Gegenereignis "keiner": $P(x \geqslant 1) = 1 - P(x = 0) > 0.999$. Daraus $0.001 > P(x = 0)$, $P(x = 0) = p^n$ mit $p = \frac{38}{100} = 0.38$, also $P(x = 0) = 0.38^n$. Einsetzen $0.001 > 0.38^n$. Logarithmieren $\ln(0.001) > n \cdot \ln(0.38)$, mit Werten $-6.91 > n \cdot (-0.968)$. Mit (-1) multiplizieren und Zeichen wechseln: $6.91 < 0.968n$, $7.14 < n$. Damit $n = 8$.

3.12 Bei einem idealen Würfel sind 3 Seiten rot, 2 Seiten gelb und 1 Seite blau gefärbt.

(a) Der Würfel werde einmal geworfen. Mit welcher Wahrscheinlichkeit zeigt der Würfel rot oder gelb? [34‰]

(b) Der Würfel werde zweimal geworfen. Mit welcher Wahrscheinlichkeit zeigt der Würfel nie blau? [17‰]

(c) Mit welcher Wahrscheinlichkeit zeigt der Würfel zweimal dieselbe Farbe? [34‰]

(d) Der Würfel werde dreimal geworfen. Mit welcher Wahrscheinlichkeit wird genau einmal rot geworfen? [52‰]

(e) Mit welcher Wahrscheinlichkeit wird mindestens einmal rot geworfen?[34‰]

(f) Mit welcher Wahrscheinlichkeit wird jede Farbe geworfen? [52‰]

12 **(a)** Es ist $p_r = 1/2$, $p_g = 1/3$ und $p_b = 1/6$. Es hat 5 rote oder gelbe Seiten, also ist $p = \frac{5}{9}$.

(b) "Nie blau" heisst "nur rot oder geld". Also $p = \frac{5}{6}\frac{5}{6} = \frac{25}{36} = 0.694$.

(c) Disjunkte Ereignisse (r,r), (b,b) und (g,g). Damit $P = P((r,r)) + P((g,g)) + P((b,b)) = \frac{1}{2}\frac{1}{2} + \frac{1}{3}\frac{1}{3} + \frac{1}{6}\frac{1}{6} = \frac{18+9+1}{36} = \frac{28}{36} = 0.778$.

(d) Typ Binomialverteilung: $P = \binom{3}{1}\frac{1}{2}(\frac{1}{2})^2 = 3 \cdot \frac{1}{8} = \frac{3}{8} = 0.375$.

(e) "Mindestens", also Gegenereignis $P(r \geqslant 1) = 1 - P(r = 0)$ $P(r = 0) = (\frac{1}{3})^3 = 1/8$ also $P(r \geqslant 1) = 7/8$.

(f) $P(r, g, b) = \frac{1}{2}\frac{1}{3}\frac{1}{6} = \frac{1}{216}$.

3.13 In einem Frühjahr trägt ein Apfelbaum viele Blüten. Die Wahrscheinlichkeit, dass eine Blüte bestäubt wird, ist 0.5. Die Wahrscheinlichkeit, dass aus einer bestäubten Blüte ein reifer Apfel wird, ist 0.4.

(a) An einem Ast hat es 6 Blüten. Wie gross ist die Wahrscheinlichkeit, dass höchstens eine Blüte bestäubt wird? [36‰]

(b) Wie gross ist die Wahrscheinlichkeit, dass aus zwei Blüten zwei reife Äpfel werden? [36‰]

(c) An einem Ast hat es 12 Blüten. Wie gross ist die Wahrscheinlichkeit, dass dieser Ast genau 3 reife Äpfel tragen wird? [36‰]

(d) Wie viele Blüten muss ein Ast haben, damit die Wahrscheinlichkeit, dass an diesem Ast mindestens ein reifer Apfel wächst, grösser als 99,9% ist? [71‰]

13 **(a)** "Höchstens eine" bedeutet $k \in \{0, 1\}$ also $P = P(k = 0) + P(k = 1)$. $P(k = 0) = 0.5^0 \cdot 0.5^6 = 0.5^6$, $P(k = 1) = 0.5 \cdot 0.5^5 = 0.5^6$, damit $P = 2 \cdot 0.5^6 = 0.0313$.

(b) $P(\text{reifer Apfel}) = 0.5 \cdot 0.4 = 0.2$, zwei reife Äpfel: $p = 0.2 \cdot 0.2 = 0.04$

(c) Binomialmodell, $P = \binom{12}{3}0.2^3 0.8^9 = 0.236$.

(d) Tipp: "mindestens", damit Gegenereignis, $P(x \geqslant 1) = 1 - P(x < 1) = 1 - P(x = 0) > 0.999$. Damit $0.001 > P(x = 0)$. Mit $P(x = 0) = 0.8^n$ folgt $0.001 > 0.8^n$. Logarithmieren: $\ln(0.001) > n \cdot \ln(0.8)$ oder $-6.91 > n(-0.223)$ und $6.91 < 0.223n$. Daraus $n > 30.99$. Damit $n = 31$.

3.14 * Ein gewöhnlicher Spielwürfel wird sechs Mal hintereinander geworfen. [200‰]

(a) Wie gross ist die Wahrscheinlichkeit, dass dabei genau drei Mal eine '5' gewürfelt wird?

(b) Wie gross ist die Wahrscheinlichkeit, dass dabei jede der sechs Augenzahlen genau einmal vorkommt?

(c) Wie gross ist die Wahrscheinlichkeit, dass dabei alle Augenzahlen genau in der Reihenfolge '1', '2', '3', '4', '5', '6' nacheinander auftreten?

(d) Ein anderer, gefälschter Würfel hat eine veränderte Wahrscheinlichkeit $p \neq \frac{1}{6}$ für die Augenzahl '6'. Die Wahrscheinlichkeit, in zwei Würfen mit diesem Würfel genau einmal eine '6' zu würfeln, ist $\frac{3}{8}$. Wie gross ist diese Wahrscheinlichkeit p ?

(e) Zwei gewöhnliche Würfel sollen 180 Mal gleichzeitig miteinander geworfen werden. Wie oft kann dabei eine Augensumme beider Würfel von 7 erwartet werden? Wie gross ist die zu erwartende Standardabweichung von diesem Mittelwert?

14 **(a)** Das ist die Grundmodell zur Binomialverteilung. Antwort: $P(X = 3) = \binom{6}{3}(\frac{1}{6})^3(\frac{5}{6})^3 = 20 \cdot \frac{5^3}{6^6} = 0.0536$.

(b) Angenommen wir haben alles unterschiedliche Zahlen gewürfelt. Da es nicht auf die Reihenfolge ankommt, gibt es 6! Kombinationen. Insgesamt gibt es 6^6 Anordnungen. Damit also $p = \frac{6!}{7^6} = \frac{720}{46656} = 0.0154$.

(c) Es gibt nur eine Möglichkeit für dieses Resultat. Alle Möglichkeiten sind 6^6. Damit ist $p = 1/6^6 = 1/46656 = 0.000021$.

(d) Die Wahrscheinlichkeit für eine Sechs sei p, für nicht eine Sechs deshalb $(1-p)$ und somit für genau eine Sechs $\binom{2}{1}p(1-p)$ als $P = 2p(1-p) = 3/8$. Damit folgt $p - p^2 = -3/16$. Quadratische Ergänzen gibt $(p-0.5)^2 - 0.25 = -3/16$ und $(p-0.5)^2 = 1/4 - 3/16 = 1/16$. Daraus $p - 0.5 = \pm 1/4$ und $p = 0.5 \pm 0.25$ und $p_{1,2} = \{0.25, 0.75\}$. Es gibt zwei Lösungen.

(e) Es gibt 6 Möglichkeiten für eine 7 (das ist (1,6),(2,5),(3,4),(4,3),(5,2),(6,1)). Insgesamt gibt es 36 Anordnungen. Also ist $P(7) = 1/6$. Der Erwartungswert bei 180 Würfen ist 180/6=30, nämlich 180 mal den Erwartungswert eines Doppelwurfes. Wir haben die Ereignisse $(7, \text{nicht}7)$ mit den Zufallsgrössen $(1, 0)$. Der Erwartungswert ist $1/6$. Die Varianz des Doppelwurfs ist $V(x) = 1/6(1-1/6)^2 + 5/6(0-1/6)^2$ und somit $\frac{5^2}{6^3} + \frac{5}{6^3} = \frac{30}{216}$. Die Standardabweichung für einen Zug ist $\sigma = \sqrt{30/216}$. Für die Summe von 180 Zügen ist dann $\sigma_{180} = \sqrt{180}\sqrt{30/216} = \sqrt{25} = 5$.

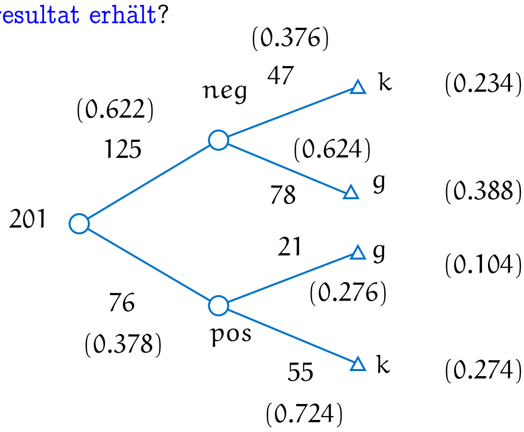
3.15 * Eine Gruppe von insgesamt 201 Personen wird mit einem noch nicht erprobten Test auf eine Erbkrankheit getestet. Bei 125 Personen fällt dieser Test negativ aus; 47 Personen haben diese Krankheit, obwohl der Test negativ ausgefallen ist, und 21 Personen haben diese Krankheit nicht, obwohl der Test positiv ausgefallen ist. [200‰]

(a) Erstellen Sie ein Baumdiagramm, das für diese Gruppe die absoluten Häufigkeiten sowie die jeweiligen Wahrscheinlichkeiten jedes Teilpfades enthält.

(b) Eine Person wird nun zufällig herausgegriffen. Wie gross ist die Wahrscheinlichkeit, dass diese Person diese Krankheit aufweist?

(c) Wie gross ist die Wahrscheinlichkeit, dass eine Person mit negativem Testergebnis diese Krankheit aufweist?

(d) Wie gross ist die Wahrscheinlichkeit, dass eine Person mit dieser Krankheit ein negatives Testresultat erhält?



15 **(a)**

(b) $P(k) = 0.234 + 0.274 = 0.508$.

(c) $P(k|neg) = 0.234$

(d) $P(neg|k) = \frac{47}{47+55} = 0.461$.

3.16 * Wir platzieren 16 rote und 24 blaue Kugeln in einer Urne und mischen kräftig. Das Ziehen der Kugeln erfolgt mit Zurücklegen darf als vollkommen zufällig betrachtet werden. [186‰]

(a) Wie gross ist die Wahrscheinlichkeit, dass bei dreimaligem Ziehen nur gleiche Farben gezogen werden? [21‰]

(b) Wie gross ist die Wahrscheinlichkeit, bei fünfmaligem Ziehen mindestens einmal eine rote Kugel zu ziehen? [21‰]

(c) Wie oft muss man ziehen, damit die Wahrscheinlichkeit, mindestens einmal eine blaue Kugel zu ziehen, grösser als 0.95 ist? [29‰]

(d) Wir betrachten folgendes Spiel:

- Man zieht nacheinander Kugeln aus der Urne.
- Zieht man eine blaue Kugel beim ersten Zug, so erhält man Fr. 5.–, zieht man eine blaue beim zweiten Zug Fr. 10.–, zieht man eine blaue beim dritten Zug Fr. 15.–.
- Zieht man eine rote Kugel, so muss man Fr. 5.– bezahlen.

- Das Spiel ist beendet, wenn entweder eine blaue Kugel gezogen wird, oder wenn die dritte Kugel gezogen wurde.

Berechnen Sie den Erwartungswert für den Gewinn bei diesem Spiel. [43‰]

(e) Sie gewinnen beim Spiel aus d) Fr. 5.–. Mit welcher Wahrscheinlichkeit haben Sie beim zweiten Zug das Spiel beendet? [36‰]

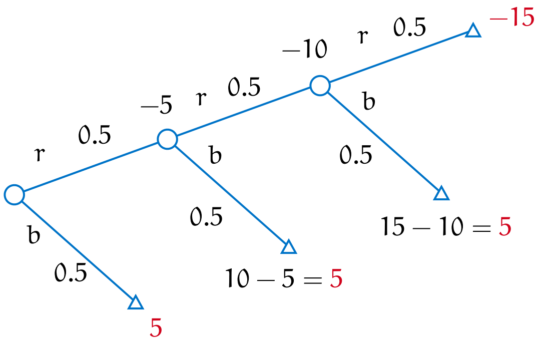
(f) Wir ziehen 20 Mal. Wie gross ist die Wahrscheinlichkeit, dass von den 20 Kugeln 8, 9 oder 10 Kugeln rot sind? [36‰]

16 **(a)** Gleiche Farbe, 2 Ereignisse (r, r, r) und (b, b, b). $P((r, r, r)) = (\frac{2}{5})^3 = 0.0604$ und $P((b, b, b)) = (\frac{3}{5})^3 = 0.216$, total $P() = 0.276$

(b) Tipp: "mindestens", deshalb Gegenereignis. $P((b, b, b, b)) = (\frac{24}{40})^4 = (\frac{3}{10})^4 = 0.000006$, damit $P() = 1 - 0.000006 = 0.999994$

(c) Mit der binomischen Verteilung und Gegenereignis $P(b \geqslant 1) > 0.95$ wird dann $1 - P(b = 0) > 0.95$ und $-P(b = 0) > -0.05$ weiter $P(b = 0) < 0.05$. Mit Formel $P(b = 0) = \binom{n}{0}(\frac{3}{5})^0(\frac{2}{5})^n = (\frac{2}{5})^n$ also $(\frac{2}{5})^n < 0.05$. Damit $n \log(0.4) < \log(0.05)$ und $n(-0.92) < (-3)$ und $n(0.92) > 3$, $n > \frac{3}{0.92} = 3.26$ also $n = 4$.

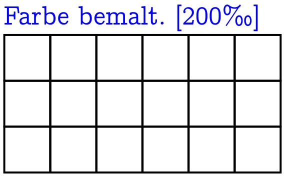
(d) Der Erwartungswert berechnet sich als $E(X) = 0.5 \cdot 5 + 0.25 \cdot 5 + 0.125 \cdot 5 - 0.125 \cdot 15 = 1.25$.



(e) Es gibt drei Möglichkeiten für eine 5. Die Wahrscheinlichkeiten sind 1/2, 1/4 und 1/8. Eine 5 im zweiten Zug hat die WS 1/4. Daraus folgt $P(5|(b, b)) = \frac{1/4}{1/4+1/2+1/8} = \frac{2/8}{2/8+4/8+1/8} = \frac{2/8}{7/8} = \frac{2}{7}$.

(f) Frage nach binomialverteilten Ereignissen, denn Modell "n mal Kopf in 20 Zügen". $P(A = \{8, 9, 10\}) = P(k = 8) + P(k = 9) + P(k = 10)$. $P(k = 8) = \binom{20}{8}(0.4)^8(0.6)^{12} = 125970 \cdot 0.000655 \cdot 0.002176782 = 0.180$, $P(k = 9) = \binom{20}{9}(0.4)^9(0.6)^{11} = 0.160$ und $P(k = 10) = \binom{20}{10}(0.4)^10(0.6)^{10} = 0.117$, zusammen $P = 0.457$.

3.17 Jedes der 18 Felder des rechts skizzierten Rechtecks wird entweder mit schwarzer oder mit weisser Farbe bemalt. [200‰]



(a) Wie viele verschieden bemalte Rechtecke sind möglich?

(b) Wie viele verschiedene Rechtecke mit 10 weissen und 8 schwarzen Feldern gibt es

(c) Jedes Feld werde zufällig – entsprechend dem Ausfall eines Münzenwurfs – schwarz oder weiss bemalt. Berechnen Sie die Wahrscheinlichkeit, dass es auf diese Weise mindestens eine Spalte (3 übereinander liegende Felder) gibt, die schwarz ist.

(d) Sie haben ein Rechteck mit 10 weissen und 8 schwarzen Feldern vor sich. Darauf wer- fen Sie zufällig 12 Sandkörner. Wie gross ist die Wahrscheinlichkeit, dass genau 5 dieser Sandkörner in einem schwarzen Feld zu liegen kommen?

(e) Auf ein anders bemaltes Rechteck werden nun wieder zufällig 2 Sandkörner geworfen. Die Wahrscheinlichkeit, dass eines davon auf einem weissen und das andere auf einem schwarzen Feld zu liegen kommt, ist gleich $\frac{5}{18}$. Wie viele Felder dieses Rechtecks sind schwarz bemalt?

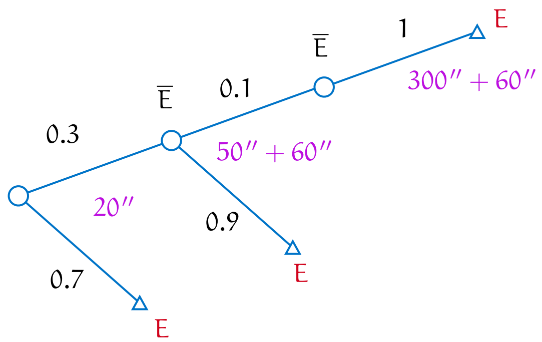
17 **(a)** Mit zwei Farben: $M = 2^{18} = 262'144$.

(b) $\binom{18}{10} = \binom{18}{8} = 43758$.
(c) $P((s,s,s)) = 0.5^3 = 1/8 = 0.125$, Modell: schwarze Spalten auf 6 Plätze verteilen. "mindestens", deshalb Gegenereignis "keine schwarze Spalte". $P = 1 - P(k=0) = 1 - \binom{6}{0}(7/8)^6 = 1 - 0.449 = 0.551$
(d) $p_s = 8/18 = 4/9$, Binomialmodell $P = \binom{12}{5}(4/9)^5(5/9)^7 = 0.224$.
(e) Zwei Ereignisse (s,w) und (w,s). Jedes hat WS $p(1-p)$, zusammen $2p(1-p)$, damit $2p(1-p) = 5/18$ oder äquivalent $p(1-p) = 5/36$ und weiter $p^2 - p = -5/36$, quadratisch ergänzt $(p-1/2)^2 - 1/4 = -5/36$, $(p-1/2)^2 = 4/36$, radiziert (Wurzel gezogen) $p - 1/2 = \pm 1/3$ und schliesslich $p = 1/2 \pm 1/3$ und $\{1/6, 5/6\}$. Anzahl Felder entweder 3 oder 15.

3.18 (Teilaufgabe) Das Handgepäck auf diesem Flughafen wird mit maximal drei Kontrollen geprüft: Eine erste Kontrolle mit einem Schnellscanner führt mit 70%iger Wahrscheinlichkeit zu einem eindeutigen Resultat; andernfalls wird eine zweite Kontrolle mit einem hochauflösenden 3D–Scanner durchgeführt, die mit 90%iger Wahrscheinlichkeit zu einem eindeutigen Resultat führt. Bringt auch diese kein eindeutiges Resultat, wird in einer dritten Kontrolle das Gepäckstück eöffnet und genau durchsucht, was auf jeden Fall zu einem eindeutigen Resultat führt.

(a) Zeichnen Sie ein Baumdiagramm, welches diesem Ablauf entspricht. Wie gross ist die Wahrscheinlichkeit, dass spätestens nach der zweiten Kontrolle ein eindeutiges Resultat vorliegt? [40‰]

(b) Die erste Kontrolle dauert 20 Sekunden, die zweite Kontrolle dauert 50 Sekunden, und die dritte Kontrolle dauert 5 Minuten. Zwischen zwei Kontrollvorgängen vergeht jeweils eine Minute. Wie lange dauert eine solche Gepäckkontrolle im Mittel? [40‰]



18 (a) Nach zwei Stufen haben die Ereignisse E die Wahrscheinlichkeit $p = 0.7 + 0.3 \cdot 0.9 = 0.97$.

(b) Wahrscheinlichkeitsgewichtete Zeiten als Mittel. $M = 0.7 \cdot 20 + 0.3 \cdot 0.9 \cdot (20 + 110) + 0.3 \cdot 0.1 \cdot 1 \cdot (20 + 110 + 360) = 63.80\text{s}$

3.19 In einem Geldbeutel befinden sich 15 Münzen, von denen 5 gezinkt sind. [200‰]

(a) Es werden drei Münzen ohne Zurücklegen aus dem Geldbeutel genommen. Bestimmen Sie die Wahrscheinlichkeit, dass mindestens eine davon gezinkt ist.

(b) Wie viele Male muss eine Münze zufällig gezogen und sofort wieder zurückgelegt werden, damit mit einer Wahrscheinlichkeit von mindestens 95% mindestens eine gezinkte gezogen wird?

(c) Eine zufällig gezogene Münze aus dem Geldbeutel wird zwei Mal geworfen. Zeigt sie zuerst Kopf und dann Zahl, so gewinnt ein erster Spieler. Bei der umgekehrten Reihenfolge gewinnt ein zweiter Spieler. In den anderen Fällen wird das Spiel wiederholt. Begründen Sie, warum dadurch tatsächlich ein faires Spiel garantiert wird.

(d) Ein Casino bietet das folgende Spiel an: Ein Spieler bezahlt einen Einsatz von 18 CHF. Dafür wird eine gezinkte Münze, die mit einer Wahrscheinlichkeit von 60% Kopf zeigt, zweimal geworfen. Zeigt die Münze zweimal Zahl, erhält der Spieler 30 CHF, zeigt sie zwei Mal Kopf, erhält er 10 CHF; in allen anderen Fällen erhält er x CHF. Wie gross muss x sein, damit dieses Spiel fair ist?

(e) Eine erste Münze zeigt mit einer Wahrscheinlichkeit von 70% Kopf, eine zweite zeigt mit einer Wahrscheinlichkeit von 40% Kopf. Die beiden Münzen werden miteinander geworfen: Eine der beiden Münzen zeigt nun Kopf, die andere Zahl. Wie gross ist die Wahrscheinlichkeit, dass es die erste Münze ist, die Kopf zeigt?

19 **(a)** "Mindestens", Gegenereignis betrachte: $P(k \geqslant 1) = 1 - P(k = 0)$ mit $p_z = 5/15$ und $p_{nz} = 10/15$ folgt $P(k = 0) = (10/15)^3 = 0.3$, damit $P(k \geqslant 1) = 1 - 0.3 = 0.7$

(b) $P(k \geqslant 1) \geqslant 0.95$ oder $1 - (10/15)^x \geqslant 0.95$ und $0.05 \geqslant (10/15)^x$. Logarithmieren $\ln(0.05) \geqslant x \ln(10/15)$ und $-3 \geqslant x(-0.406)$ ist gleich (Ordnungsrelation umkehren) $3 \leqslant 0.406x$ und $3/0.406 = 7.39 \leqslant x$, so dass $x = 8$.

(c) Bei jeder Münze, ob gezinkt oder nicht, gilt $p_K + p_Z = 1$. Die Ereignisse (K, Z) und (Z, K) sind gleichwahrscheinlich, denn die Ausgänge sind unabhängig. Es ist fair, weil jeder Spieler die gleiche Gewinnwahrscheinlichkeit besitzt.

(d) Faires Spiel heisst, Einsatz gleich erwartetem Gewinn. Hier also $E(\text{Gewinn}) = 18$. Es ist $P(Z, Z)) = 0.4^2 = 0.16$, $P((K, K)) = 0.6^2 = 0.36$ und $P(\text{sonst}) = 1 - P(Z, Z)) - P((K, K)) = 1 - 0.52 = 0.48$. Der Erwartungswert $E(G) = 30P(Z, Z)) + 10P(K, K)) + x0.48 = 18$ oder $30 \cdot 0.16 + 10 \cdot 0.36 + x0.48 = 18$ und $8.4 + 0.48x = 18$, $0.48x = 9.6$ und $x = 20$.

(e) Die Wahrscheinlichkeit ist $p = \frac{70}{110} = 0.636$, denn der eine zeigt Kopf in 70 von Hundert und der andere in 40 von Hundert. Damit erscheint in 110 von 200 Kopf, oder in 55 von 100 Doppelwürfen. Von den 110 stammen 70 vom ersten Würfel.

Standard-Normalverteilung $\Phi(z = \frac{x-\mu}{\sigma^2})$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Kapitel 33

Beurteilende Statistik**

Anstatt beurteildende Statistik kennt man auch die Namen schliessende oder mathematische Statistik, Inferenzstatistik usf. Der Buchtitel von Jakob Bernoulli von 1713 sagt es am besten: *ars conjectandi*, die Kunst des Vermutens. Beide Wörter sind wichtig, also Kunst und Vermutung. Da Kunst von "können" kommt, könnte man sagen: *gekonnte Vermutung*. Eine Vermutung ist immer unsicher.
Lesenden muss klar sein, dass hier nur ein ganz kleiner Einblick in die beurteilende Statistik gegeben werden kann. Wir wollen nur die grossen Linien zeichnen.
Häufig wird das Testen wie Kuchenbacken verwendet. Man schaut im Kochbuch nach, welchen Test man wofür verwendet. Hier wollen wir eher darlegen, was unter der Oberfläche liegt, und deshalb auf viele Rezepte verzichten.

33.1 Grundlagen

Im vorangegangenen Kapitel haben wir Urnen betrachtet, deren Eigenschaften bekannt sind und aus denen Stichproben gezogen werde. Dann haben wir die Eigenschaften der Stichprobe untersucht. Die Auswertung der Stichprobe erfolgt mit den Instrumenten der beschreibenden Statistik, also z.B. Mittelwert, Standardabweichung etc. In diesem Kapitel liegt die Stichprobe vor, aber die Eigenschaften der Urne sind unbekannt.

Anmerkung 33.1. Im Folgenden sind Zufallsvariablen oder -grössen durch Grossbuchstaben angezeigt. Es ist z.B. X eine Zufallsgrösse und x eine Realisierung dieser Grösse. Dies ist in der Statistik üblich.

Zufallsstichprobe

Stichproben erhebt man, weil eine sogenannte Vollerhebung nicht praktisch, zu teuer oder nicht durchführbar ist. Bis noch vor wenigen Jahrzehnten wurden Volkszählungen durchgeführt, wo jeder Haushalt eine statistische Einheit war. Man behilft sich mit Teilerhebungen.
Die ideale Stichprobe ist eine zufällige, bei der jedes Element die gleiche Chance hat ausgewählt zu werden, jede Wahl also unabhängig von anderen ist, und die Ausprägungen gleich verteilt sind. Das ist für eine Urne mit jeweils gleich vielen Kugeln pro Farbe (Merkmal) und Ziehen mit Zurücklegen gegeben.
Diese Stichprobe ist ideal, weil sie am besten die Voraussetzungen des zentralen Grenzwertsatzes für grosse Stichproben erfüllt, der lautet: Eine Summe von sehr vielen unabhängigen

identisch verteilten Zufallsvariablen X_i mit endlicher Varianz ist approximativ normalverteilt. Aus dieser Kenntnis kann man dann neben der Summe oder Mittelwert nach anderen Kenngrössen der Probe fragen.

Ein Ziehen ohne Zurücklegen verstösst gegen die Unabhängigkeit. Angenommen man wählt eine zufällige Stichprobe von Probanden für einen Test der Wirksamkeit eines Medikaments. Dann wählt man wieder zufällig die Hälfte der Probanden mit dem Wirkstoff. Diese müssen ohne Zurücklegen gezogen werden. Die Stichprobe ist also nicht unabhängig.

Die Population kann andere Merkmale besitzen, die für die Untersuchung oder Repräsentanz wichtig sind. Z.B. Fragen zur Politik betreffen alle Wahlberechtigten, die eine gewisse Altersstruktur aufweisen und Land- und Stadtbevölkerung darstellen. Gewisse Stichproben werden proportional geschichtet und unterteilt, um solche demographischen Effekte zu berücksichtigen.

33.1.1 Fragestellungen

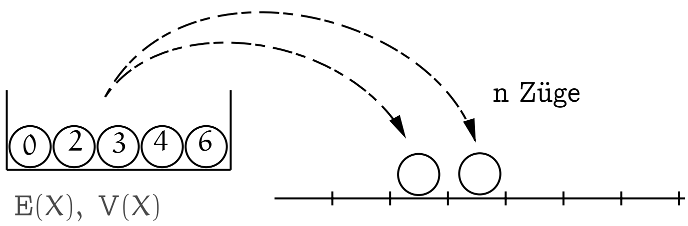
Es gibt zwei grundsätzliche Fragestellungen bei der beurteilenden Statistik, und zwar in Abhängigkeit von unserem Wissen über die Grundgesamtheit oder Population:

(1) Wir wissen nichts von den Merkmalen der Grundgesamtheit und erarbeiten deshalb aus der Stichprobe *Schätzungen*,

(2) wir haben *Vermutungen* zur Gesamtheit und suchen Bestätigung.

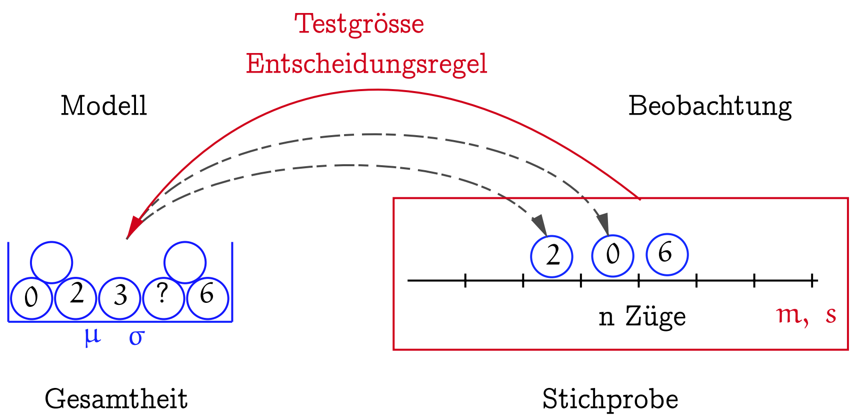
33.1.2 Modell

Wir betrachten eine Urne mit fünf Kugeln. Sie tragen Zahlen. Alle Kugeln haben dieselbe Wahrscheinlichkeit gezogen zu werden.



Das Urnenmodell hat den Erwartungswert $E(X) = \frac{1}{5}(0+2+3+4+6) = 3$. Die Varianz ist $V(X) = \frac{1}{5}((-3)^2 + 1 + 0 + 1 + 9) = 20/5 = 4$. Die Standardabweichung ist $\sigma = \sqrt{V(X)} = \sqrt{4} = 2$.

Unser Experiment soll das n-fache Ziehen von Kugeln mit Zurücklegen sein, wobei die Zufallsvariable Y_n die Summe der gezogenen Zahlen sein soll, also $Y_n = X_1 + X_2 + \ldots + X_n$. Wir wählen $n = 25$. Die Summe Y_n kann theoretisch von minimal 0 bis maximal 150 reichen. Es ist aber eher anzunehmen, dass Y_n in der Nähe von $E(Y) = 25 \cdot E(X) = 25 \cdot 3 = 75$ zu liegen kommt. Die Streuung, also Standardabweichung, von Y ist $\sigma_Y = \sqrt{n} \cdot \sigma_X$ (siehe Eigenschaften 32.8 auf Seite 32-16).



33.2 Testen von Hypothesen

33.2.1 Einführungsbeispiel

Wir ziehen 25 mal mit Zurücklegen aus einer Urne mit der Zusammensetzung von oben und erhalten folgende Stichprobe:

$$x = (4, 2, 2, 3, 6, 4, 2, 3, 0, 4, 0, 0, 3, 4, 0, 6, 2, 0, 2, 2, 0, 0, 2, 3, 6)$$

Die Summe y ist 60 mit einem Durchschnitt von 2.4 und einer Stichprobenvarianz von $s_X^2 = 3.293$. Nun argumentieren Fritz und Frida, die die Zusammensetzung der Urne nicht kennen. Fritz meint, die Summe ist eigentlich 75, der Zufall hat halt 68 ergeben. Frida meint, der Durchschnitt ist sicher nicht 75 sondern kleiner, denn die Abweichung von $75 - 60 = 15$ ist doch eindeutig. Nun schauen sie auf die Zahlen. Sie erinnern den Standardfehler, der $\sqrt{n}\sigma$ ist. Sie einigen sich, dass σ_X nicht bekannt ist, aber von $s_X = \sqrt{3.83}$ ersetzt werden kann. Somit ist $SE_Y = \sqrt{25 \cdot 3.83} = 9.79$. Nun setzen sie die Abweichung zwischen Vermutung und Beobachtung ins Verhältnis zum Standardfehler und bilden die *Testgrösse*

$$Z = \frac{\text{beobachtet} - \text{vermutet}}{\text{SE}} = \frac{60 - 75}{9.79} = -1.53$$

Die Diskrepanz zwischen Vermutung und Beobachtung ist $-1.53SE_Y$. Nun stellt sich die Frage, ist das viel oder wenig? Was bedeutet dieser Wert überhaupt? Das Argument von Fritz ist, die Abweichung vom wahren Wert 75 ist nur zufällig, es gibt keine andere Begründung. Die Zufälligkeit würde mit mehr Ziehungen immer mehr verschwinden. Frieda ist überzeugt, dass 15 ein zu grosser Unterschied ist, um nur zufällig zu sein. Wie kann man dies entscheiden? Aufgrund der relativ grossen Zahl von $n = 25$ kann man die Normalverteilung heranziehen. Wie gross ist der Wert für $z = -1.53$? Aus der Tabelle lesen wir für $z = 1.53$ den Wert $\Phi(z) = 0.9370$ heraus, womit $\Phi(-z) = 1 - 0.9370 = 0.063$ oder 6.3%. Die Chance, dass der beobachtet Wert zur Vermutung passt, ist 6.3%. Können wir das besser einordnen?
Zu Recht oder Unrecht sei dahingestellt, es hat sich eingebürgert, eine Schwelle von 5% zu nehmen. Damit hätte Frieda recht, wenn der Wert noch ein bisschen tiefer wäre. Bei einem Unterschied von 17 würde man Frieda recht geben. Aber hier ist $6.3 > 5$, so dass man Fritz recht gibt. Nun haben wir die Stichprobenvarianz mit 3.29 genommen, weil wir so getan haben, als kennten wir den wahren Wert 4 nicht. Mit 4 allerdings wäre z kleiner geworden, und hätte die Meinung von Fritz noch zusätzlich gestützt. Die *Entscheidungsregel* ist hier, die Vermutung zu verwerfen, wenn die Chance geringer als 5% ist.

Nun gehen wir einen Schritt weiter, nämlich jeder Schüler hat eine Stichprobe aus der Urne gezogen. Während der Lehrer 60 gezogen hat, haben die Schüler gemäss Abbildung 33.1 gezogen.

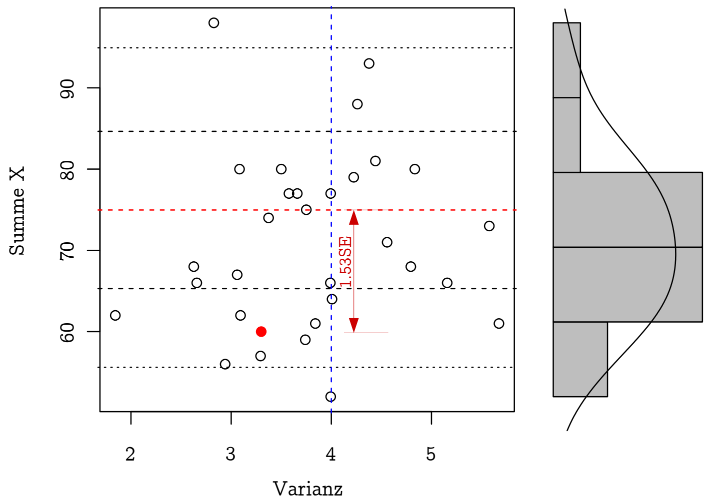


Abbildung 33.1: 25 Stichproben

In der Abbildung erkennen wir einige Proben, die ausserhalb des Annahmebereichs liegen. Zwei Proben sind jenseits von 2 Standardfehlern. Andere liegen viel näher am Erwartungswert der Vermutung. Aber genau dies ist zu erwarten: nicht alle Stichproben einer gegebenen Gesamtheit führen zum selben Resultat, zur selben Entscheidung. Im Histogramm rechts sind die Proben zusammengefasst. Man kann schon ein paar Eigenschaften erkennen, die es mit einer Normalverteilung verbindet, z.B. die Symmetrie, die meiste Fläche um den Mittelwert etc.

Aus dem Beispiel lernen wir einiges. Wir fassen zusammen.

Definition 43. Als *Testgrösse* (auch Teststatistik oder Prüffunktion) bezeichnet man eine bestimmte Stichprobenfunktion, die bei einem Hypothesentest dazu verwendet wird, die Testentscheidung zu treffen.

Die Testfunktion muss Grössen aus Grundgesamtheit und Stichprobe verbinden. Bei z ist die die Stichprobe anhand der beobachteten Summe und die Gesamtheit anhand des vermuteten Wertes und Varianz. Anstatt Hypothese könnte man auch Vermutung setzen. Hier war die Vermutung, dass der Erwartungswert der Urne 3 ist und somit beim 25-maligen Ziehen die Summe 75 folgt.

Anmerkung 33.1. Die Testgrösse ist ein Funktion auch von Zufallsvariablen und damit ebenfalls eine Zufallsgrösse. Zufallsgrössen werden durch Wahrscheinlichkeitsfunktionen beschrieben.

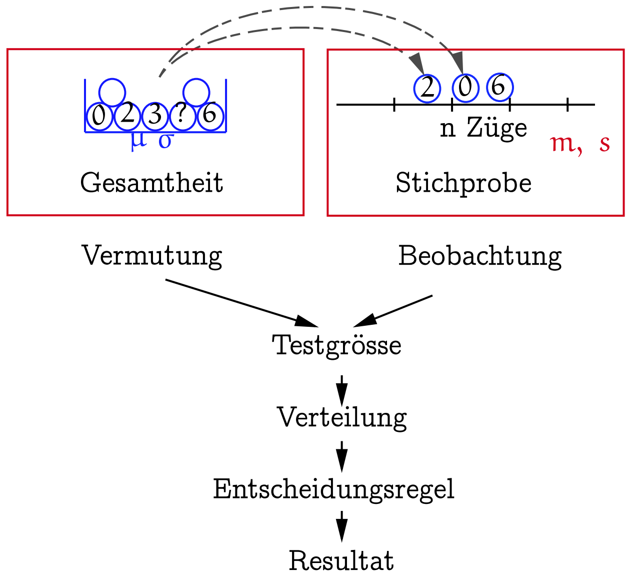


Abbildung 33.2: Übersicht Hypothesentest

33.2.2 Testkomponenten

Die für Tests benötigten Zutaten sind in der Abbildung 33.2 ersichtlich. Sie hängen alle voneinander ab.

Stichprobe

Wenn man die fundamentale Aufgabe hat, aus der Stichprobe auf die Grundgesamtheit zu schliessen, so ist dies der Ein-Stichprobenfall. Damit will man die Situation abgrenzen, wonach mehrere Stichproben miteinander verglichen werden, um zwei Grundgesamtheiten zu vergleichen oder z.B. zwei Gruppen zu bilden, die unterschiedliche behandelt werden, also Wirkstoff- und Placebogabe. Es liegt der Zwei-Stichprobenfall vor. Wir beginne mit einer Stichprobe.

Vermutung, Hypothese

Wir verwenden die Ausdrücke Vermutung und Hypothese synonym. Es gibt zwei Arten von Vermutungssituationen: Es liegt eine Vermutung H vor und damit auch ihr Komplement $\overline{H} = G \setminus H$, oder zweitens eine Hypothese H und mehrere Alternativen H_1, H_2 usw. Die grundlegende Hypothese ist die sogenannte *Null-Hypothese* H_0, die deshalb speziell ist, weil sie die vermuteten Grössen für die Testgrösse beisteuert. Testen bedeutet deshalb vor allem, die Null-Hypothese verwerfen wollen. In unserem Einführungsbeispiel war H_0: Die Summe ist 75, und das Komplement, die Summe ist kleiner als 75.

Testgrösse

Gemäss Definition 43 ist die Testgrösse eine Funktion. Sie soll Grundgesamtheit mit der Stichprobe verbinden und mit ihrer zugehörigen Verteilung die Entscheidung ermöglichen. Häufig wir eine Differenz zwischen beobachteten Grössen und vermuteten hergestellt, denn das Ausmass des Abstands sagt etwas über die Ähnlichkeit aus. In unserem Beispiel haben wir die Abweichung der Summenzahl bezogen auf den Standardfehler herangezogen, also eine relative Differenz angenommen:

$$Z = \frac{Y - \mu_0 \cdot n}{\sigma\sqrt{n}}$$

Wie haben etwas vereinfacht, weil wir den Standardfehler und somit die Varianz als fest angenommen habe. Eigentlich ist es auch ein Zufallsvariable S und somit

$$T = \frac{Y - \mu_0 \cdot n}{S\sqrt{n}}.$$

Unter Berücksichtigung dieser Tatsache ist aber wiederum eine andere Verteilung massgebend, die berücksichtigt, dass auch der Nenner zufällig schwankt. Für grössere Stichproben spielt diese Vereinfachung ein untergeordnete Rolle.

Definition 44. Wenn eine Zufallsvariable X einer gewissen Wahrscheinlichkeitsverteilung F gehorcht, dann schreibt man $X \sim F$.

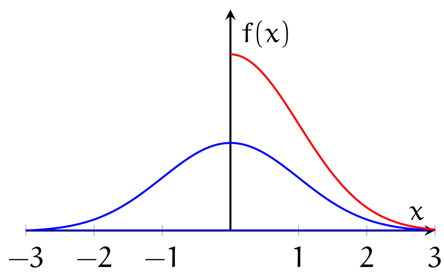
33.2 Beispiel Wir könnten schreiben $Y \sim \Phi(0,1)$, wobei Φ für die Normalverteilung steht oder $T \sim t_n$, wo t_n eine entsprechende Verteilung darstellt. ◁

Testverteilung

Die zu benutzende Verteilung bestimmt sich aus der Testgrösse. Wir haben Z als normalverteilt betrachtet, denn Y in $Z = \frac{Y - \mu_0 \cdot n}{\sigma}$ ist normalverteilt. Nun betrachten wir eine Testgrösse $Y = X^2$, wobei $X \sim \Phi(0,1)$ ist. Wir suchen $P(Y < y)$ und setzen deshalb ein

$$\begin{aligned} P(Y < y) &= P(X^2 < y) \\ &= P(X < \pm\sqrt{y}) \\ &= P(-\sqrt{y} < X < \sqrt{y}) \\ &= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) \\ &= 2\Phi(\sqrt{y}) \end{aligned}$$

Das war ein besonders einfaches Beispiel. Diese Verteilung ist nun nur für positive Werte $\sqrt{y}$ definiert. Da die Verteilung gegen 1 gehen muss, ist der Faktor 2 gerade richtig.



33.3 Übung Eine Stadt verzeichnet 15941 Geburten, wovon 8173 Knaben sind und deshalb 7768 Mädchen. Nun möchte man wissen, ob diese Zahlen mit einer Wahrscheinlichkeit von $p = 0.5$ für die Aufteilung verträglich sind und in welchem Masse. Die Nullhypothese H_0 ist: Mädchen- und Knabengeburten sind gleich wahrscheinlich. Wir setzen X als die Anzahl Knabengeburten. Diese ist binomialverteilt und wegen der grossen Anzahl ungefähr normalverteilt mit $\mu = n \cdot p$ und $\sigma^2 = n \cdot p(1-p)$. Somit

$$Z = \frac{X - n \cdot p}{\sqrt{n \cdot p(1-p)}} \quad \sim \quad \Phi(0,1)$$

und damit

$$Y = Z^2 = \frac{(X - n \cdot p)^2}{n \cdot p(1-p)} \quad \sim \quad 2\Phi(\sqrt{y})$$

Wir setzen ein (mit $x = 8173$) und erhalten einen Wert

$$y = \frac{(8173 - 15941 \cdot 0.5)^2}{15941 \cdot 0.25} = 10.29$$

Den dazugehörigen Wahrscheinlichkeitswert p finden wir, indem wir $\sqrt{y} = 3.21$ bestimmen und rechnen $p = 2 \cdot [\Phi(3.21) - 0.5) = 2[0.9993 - 0.5] = 0.9986 = 1 - 0.0014$. Die Nullhypothese ist äusserst unwahrscheinlich. Wieso haben wir 0.5 abgezogen? Da $\sqrt{y} > 0$ ist, müssen wir die ganze Wahrscheinlichkeit von $\Phi(x)$ von $(-\infty, 0)$ weglassen. Dafür multiplizieren wir aber mit 2, das aus der Transformation hervorgegangen ist. Sieh Abbildung mit den Dichten, rote Kurve. ◁

Als Vorarbeit wollen wir noch die Testgrösse umformen. Der Grund wird später ersichtlich. Es ist die Wahrscheinlichkeit einer Mädchengeburt p_2 das Komplement der Knabengeburt $p_2 = 1 - p_1$ und somit $p_1 + p_2 = 1$. Somit schreiben wir mit X_1 Anzahl Knabengeburten und X_2 Anzahl Mädchengeburten

$$\begin{aligned}
Y &= \frac{(X_1 - n \cdot p_1)^2}{n \cdot p_1(1 - p_1)} \\
&= \frac{(X_1 - n \cdot p_1)^2(p_1 + p_2)}{n \cdot p_1 \cdot p_2} \\
&= \frac{(X_1 - n \cdot p_1)^2}{n \cdot p_1} + \frac{(X_1 - n \cdot p_1)^2 p_1}{n \cdot p_2} \\
&= \frac{(X_1 - n \cdot p_1)^2}{n \cdot p_1} + \frac{((n - X_2) - n \cdot (1 - p_2))^2}{n \cdot p_2} \\
&= \frac{(X_1 - n \cdot p_1)^2}{n \cdot p_1} + \frac{(-X_2 + n \cdot p_2)^2}{n \cdot p_2} \\
&= \sum_{k=1}^{2} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k}
\end{aligned}$$

Die Testgrösse ist auch die Summe der quadratischen Abweichungen von zwei (und später mehr) Anteilen. Wir werden diese Transformation noch brauchen.

33.4 Übung Wir kommen auf die Übung 33.3 zurück. Wir testen den Sachverhalt mit einem anderen Test. Wir haben die Testgrösse Z angegeben. Der entsprechende Testwert ist

$$z = \frac{8173 - 15491 \cdot 0.5}{\sqrt{15491 \cdot 0.25}} = 6.87$$

Wir können nun die entsprechende Wahrscheinlichkeit dazu mit $\Phi(0, 1)$ bestimmen. Nun sehen wir aber, dass der Wert z so gross ist, dass er nicht mehr auf der Tabelle zu finden ist. Damit kann die Hypothese von 50% Knabengeburten nicht aufrecht gehalten werden. ◁

Wichtig 8. Zu einem bestimmten zu testenden Sachverhalt kann es mehrere zulässige Testgrössen und damit Tests geben. ⊣

Entscheidungsregeln

Es gibt viele Möglichkeiten zu entscheiden in einer ungewissen Lage. Deshalb gibt es die perfekte und einzige Lösung nicht. Man kann aber versuchen, durch zusätzliche Überlegungen, Kriterien und Präferenzen die Anzahl einzuschränken.

Die Entscheidung über die Annahme oder Ablehnung einer Vermutung, kann richtig oder falsch sein. Denn es gibt keine absolute Gewissheit.
Angenommen ein Bettler steht vor einem und bittet um Geld. Nun kann er in einer echten Notlage sein oder einen betrügen wollen. Der Entscheider kann abwägen, ob es besser ist, nichts zu geben, und damit nicht betrogen zu werden oder etwas zu geben, um die Not zu lindern auch wenn vielleicht der Bettler nicht ehrlich ist. Zwischen Linderung und Betrug kann man abwägen.

Definition 45. Beim Test einer Hypothese liegt ein *Fehler 1. Art* (auch α-Fehler) vor, wenn die Nullhypothese zurückgewiesen wird, obwohl sie in Wirklichkeit wahr ist. Die Fehlentscheidung, die Nullhypothese zurückzuweisen, obwohl sie wahr ist, heisst *Fehler 2. Art*.

Um beim Beispiel zu bleiben sei die Vermutung (Nullhypothese), der Bettler ist in Not. Die unbekannte Tatsache ist, dass er echt bedürftig ist. Ich halte ihn für einen Betrüger und gebe nichts. Das ist ein Fehler erster Art. Ich gebe ein Almosen, obwohl er betrügt. Das ist der Fehler zweiter Art. Richtig wäre, etwas zu geben, wenn er bedürftig ist und nichts zu geben, wenn er schwindelt.

Wahrheitswert

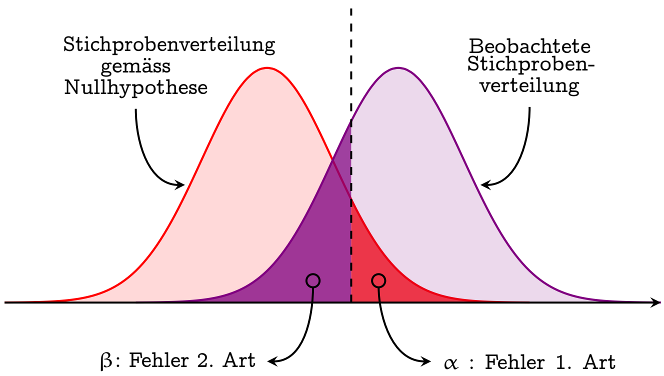
Entscheidung	Vermutung wahr	falsch	
Vermutung annehmen	richtiger Entscheid $p = 1 - \alpha$	Fehler 2. Art $p = \beta$	A
ablehnen	Fehler 1. Art $p = \alpha$	richtiger Entscheid $p = 1 - \beta$	$\overline{A}$
	W	$\overline{W}$	

In der Abbildung sind die vier Möglichkeiten dargestellt. Die Wahrscheinlichkeiten sind strenggenommen wie folgt:

$$P(A|W) = 1 - \alpha, \quad P(\overline{A}|W) = \alpha, \quad P(A|\overline{W}) = \beta \quad \text{und} \quad P(\overline{A}|\overline{W}) = 1 - \beta.$$

oder mit den zwei Verteilungen

$$P_0(A) = 1 - \alpha, \quad P_0(\overline{A}) = \alpha, \quad P_1(A) = \beta \quad \text{und} \quad P_1(\overline{A}) = 1 - \beta.$$



33.5 Übung Wir betrachten die als wahr angenommene Null-Hypothese $\mu_0 = 5$, normalverteilt mit $\Phi(5, 0.5)$ mit der Alternativen $\mu = 6$, einseitig mit $\alpha = 0.05$. Wie gross ist β? Der Annahmebereich der Nullhypothese reicht bis $5 + z_\alpha \cdot 0.5$, also $5 + 1.645 \cdot 0.5 = 5.822$. Für die alternative Verteilung mit $\Phi(6, 02.)$ entspricht dies dem Wert $z = \frac{5.822-6}{0.5} = -0.356$. Aus der Standardnormal-Tafel folgt $P(Z < z) = 0.331 = \beta$. Formell:

$$P(Z < z) = P(Z < \frac{\mu_0 + z_\alpha \cdot \sigma - \mu_1}{\sigma}) = P(Z < \frac{\mu_0 - \mu_1}{\sigma} + z_\alpha).$$

Man erkennt, dass β mit kleinerem σ sinkt. In der Praxis erreicht man dies mit der Erhöhung des Stichprobenumfangs. $\triangleleft$

Wir wissen, dass eine vernünftige Testgrösse Grundgesamtheit und Stichprobe verbindet und dass die Testgrösse einer bestimmten Wahrscheinlichkeitsverteilung folgt. Darauf muss die Entscheidungsregel aufbauen. Die Verteilung ist ja einfach eine Funktion $F(y) = P(Y < y)$, die von einer unabhängigen Variablen y und vielleicht von Parametern abhängt.
Im Einführungbeispiel haben wir den Standardfehler herangezogen, um zu entscheiden, ob die Testgrösse zur Vermutung passt. Im vorangehenden Kapitel haben wir gesagt, Konfidenzintervalle seien eine Art Verallgemeinerung des Standardfehlers.

Definition 46. Die vor einem Test festgelegte Wahrscheinlichkeit α, bei einer Entscheidung einen Fehler 1. Art zu begehen, falls die Nullhypothese wahr ist, nennt man *Signifikanzniveau* oder *Irrtumswahrscheinlichkeit*.

Nun kann man diese Wahrscheinlichkeit mit dem Test verbinden. Weil die Testgrösse im Kern eine Abweichung von "beobachtet" minus "vermutet" enthält, gilt, etwas salopp geschrieben

$$P(T = \text{"beobachtet"} - \text{"vermutet"} > k) = \alpha$$

Nun teilt die Grösse k den möglichen Bereich in zwei Teilbereiche. Falls $T \leqslant k$ bei fester Irrtumswahrscheinlichkeit $\alpha = P(\overline{A}|W)$, dann die Vermutung annehmen, falls $T > k$, die Vermutung ablehnen. Die Grösse k heisst *kritischer Wert*.

Anmerkung 33.6. Die Vermutung annehmen heisst nicht, dass sie wahr ist, sondern dass das Resultat mit der Vermutung verträglich, die Hypothese nicht widerlegt werden kann. Ob wahr oder falsch wissen wir nie mit Gewissheit, nur nach Wahrscheinlichkeiten.

Jetzt haben wir eine Verbindung zwischen Irrtumswahrscheinlichkeit α, Testgrösse T und Verteilung $F(y) = P(Y < y)$ geschaffen, aber eine neue Variable eingeführt, nämlich k. Wenn aber die Verteilung $F(y)$ ist, dann ist die Inverse $F^{-1}(y) = q(y)$. Dieser Wert nennt sich *Quantil* oder *Perzentil*. In unserem Beispiel ist

$$P(T > k) = 1 - P(T \leqslant k) = \alpha, \qquad P(T \leqslant k) = 1 - \alpha, \qquad k = F^{-1}(1 - \alpha)$$

33.7 Beispiel Es ist $F = \Phi(0, 1)$, eine Normalverteilung. Wir wollen die Quantile $q(0.95)$ und $q(0)$ bestimmen. In der Tabelle suchen wir den Wert $\Phi = 0.95$. Wir finden 0.9495 bei $z = 1.64$ und 0.9505 bei $z = 1.65$. Somit ist $q(0.95) \approx 1.645$. Sodann $q(0.5)$ ist einfach, nämlich $z = 0.5$. Das Quantil teilt den Wertebereich von $[0, 1]$ in zwei Stücke $[0, F(q)] \cup (F(q), 1]$, z.B. $[0, 0.95] \cup (0.95, 1]$. $\triangleleft$

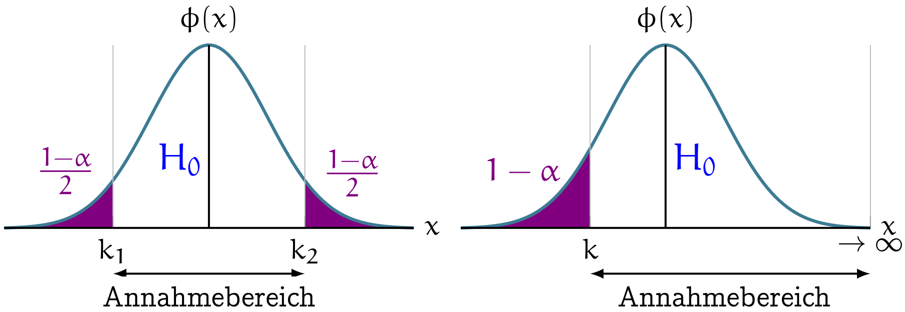
Nun kann man auch das Quantil bestimmen, das zur Testgrösse t passt. Es ist

$$P(T > t) = p \qquad \Leftrightarrow \qquad q(p) = F^{-1}(p)$$

Diese Grösse nennt man *p-Wert*. Anstatt die Grösse t und k zu vergleichen, kann man auch Irrtumswahrscheinlichkeit und P-Wert vergleichen:

$$(t > k) \qquad \Leftrightarrow \qquad (p < \alpha).$$

Wir haben bis hierhin immer $T =$ "beobachtet" $-$ "vermutet" betrachtet und dann $P(T > k)$ beurteilt. Nun kann die Richtung der Abweichung unerheblich sein. Somit sollte man $P(|T| > k) = 1 - P(|T| \leqslant k)$ betrachten und somit $1 - P(-k \leqslant T \leqslant k)$.



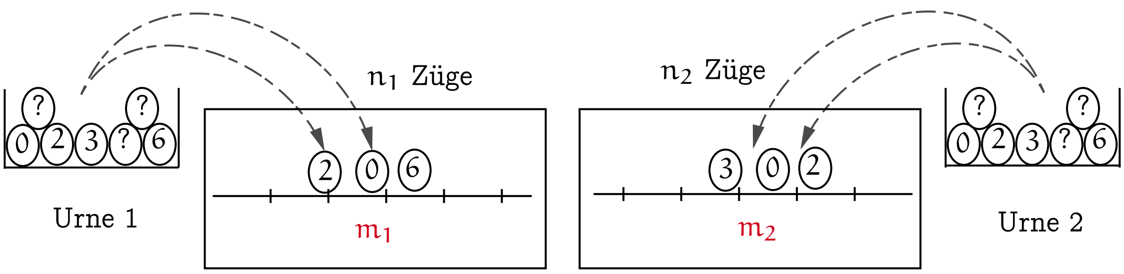
Formel 33.8. Entscheidungsregeln für den Ein-Stichprobenfall Man verwendet zwei Methoden:

- Signifikanzniveau (Irrtumswahrscheinlichkeit) α wählen, Annahmebereich mit Verteilung bestimmen (einseitig, zweiseitig), Testwert t prüfen;
- Testwert t bestimmen, daraus p-Wert aus Verteilung errechnen und beurteilen

Wichtig 9. Eine Annahme der Vermutung ist nur eine notwendige, aber nicht hinreichende Bedingung für die Relevanz der Vermutung. ⊣

33.2.3 Zwei-Stichprobenfall

Wir wollen zwei Stichproben aus zwei Urnen vergleichen. Der einfachste Vergleich betrifft die Differenz der Mittelwerte. Denn dieser gilt als erster und wichtigster Parameter eines Modells. Wie in der Abbildung ersichtlich, sind die zwei Stichproben von unterschiedlichem Umfang, d.h. zum einen n_1 Züge und zum anderen n_2.



Die Testgrösse wird ganz analog sein:

$$Z = \frac{\text{beobachtete Differenz} - \text{vermutete Differenz}}{\text{SE der Differenz}}$$

Der Standardfehler SE benötigt die entsprechende Standardabweichung, hier der Differenz. Das ist nun eine neue Fragestellung: wie bestimmt sich diese? Wir erinnern, dass der Erwartungswert einer Summe oder Differenz $Y_1 - Y_2$ der Summe oder Differenz der Erwartungswerte ist

$$E(Y_1 - Y_2) = E(Y_1) - E(Y_2).$$

Anderseits ist die Varianz gleich dem Erwartungswert der quadratischen Abweichung, also $Y_1 = (X_1 - \mu_1)^2$ etc. Somit

$$\begin{aligned} V(Y_1 - Y_2) &= E[\big((X_1 - \mu_1) - (X_2 - \mu_2)\big)^2] \\ &= E[(X_1 - \mu_1)^2 + (X_2 - \mu_2)^2 - 2Y_1Y_2] \\ &= E[(X_1 - \mu_1)^2] + E[(X_2 - \mu_2)^2] - 2E[Y_1Y_2] \\ &= V(Y_1) + V(Y_2) - 2E(Y_1Y_2) & \text{mit} \quad E(Y_1Y_2) = 0 \\ &= V(Y_1) + V(Y_2) \end{aligned}$$

Weil die Stichproben unabhängig sind, ist der Term $E(Y_1Y_2)$, die Kovarianz null. Obwohl sich die Zufallszahlen subtrahieren, ist die Varianz der Differenz grösser. Anders geschrieben

$$\sigma^2_{Y_1-Y_2} = \sigma^2_{Y_1} + \sigma^2_{Y_2} \quad \Rightarrow \quad \sigma_{Y_1+Y_2} = \sqrt{\sigma^2_{Y_1} + \sigma^2_{Y_2}}$$

Aus dieser allgemeinen Betrachtung leiten wir den Standardfehler ab. Von Satz 32.15 auf Seite 32-29 wissen wir, dass für den Mittelwert $\overline{Y}$ gilt: $\sigma_{\overline{Y}} = \sigma X/\sqrt{n}$. Damit für die Differenz

$$\sigma^2_{X_1-X_2} = \frac{\sigma^2_{X_1}}{n_1} + \frac{\sigma^2_{X_2}}{n_2}$$

oder eben mit den Standardfehlern

$$SE_{X_1-X_2} = \sqrt{SE^2_{X_1} + SE^2_{X_2}}$$

33.9 Übung Wir betrachten zwei ungleiche Stichproben aus zwei unabhängigen Urnen und hegen die Vermutung, die Mittelwerte sind gleich. Das ist die Nullhypothese. Die alternative Hypothese ist: sie sind ungleich. Das Experiment hat folgende Kennzahlen geliefert

	Stichprobe 1	Stichprobe 2
Umfang n	50	100
Mittelwert	2.006	2.058
Standardabweichung σ_X	0.101	0.099
Standardfehler SE_X	0.014	0.010

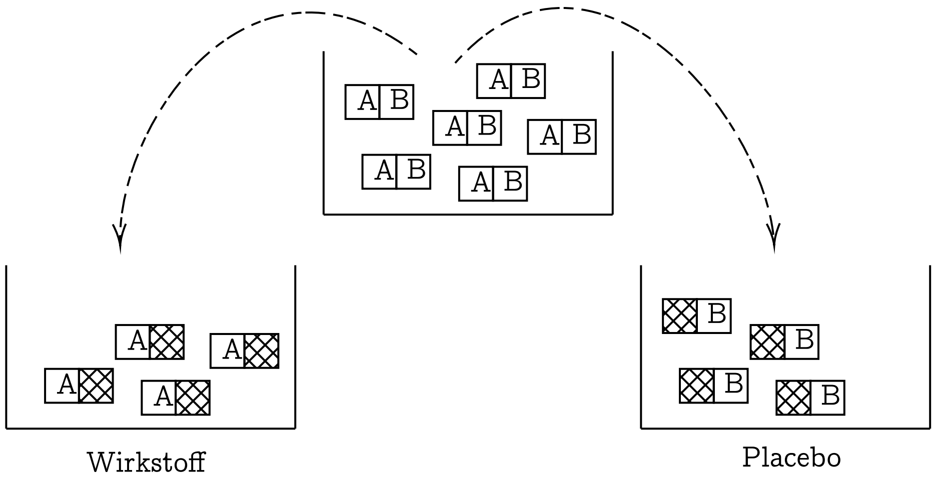
Der gemeinsame Standardfehler ist $\sqrt{0.014^2 + 0.010^2} \approx 0.018$ und die Testgrösse

$$z = \frac{(2.006 - 2.058) - 0}{0.018} \approx -3.0$$

Die Null im Zähler ist die erwartete Differenz unter der Nullhypothese. Obwohl wir die echten Standardabweichungen nicht kennen, verwenden wir die Normalverteilung als Verteilung der Testgrösse, denn die recht hohe Anzahl von Zügen, also 50 und 100, rechtfertigen diese einfachste Wahl. Da der Wert -3 extrem hoch ist, aus der Tabelle $\Phi(3) = 0.9987$ herausgelesen wird, und somit $\Phi(-3) = 1 - 0.9987 = 0.0013$, der p-Wert, sehr gering ist, wird man die Vermutung verwerfen. Die Mittelwerte sind höchstwahrscheinlich nicht gleich. ◁

Randomisierte Studie

Random heisst Zufall auf englisch. Eine randomisierter Test bedeutet, dass ein Losverfahren entscheidet, wer von den Probanden ein Wirkstoff bekommt und wer ein Placebo. Mit dem Losverfahren will man andere Faktoren ausschliessen, z.B. einen Forscher, der besonders kränklichen Probanden den vermeintlichen Wirkstoff zuhält.



Schon aus der Abbildung sieht man, dass die Situation nicht identisch ist mit der echten Zwei-Stichprobensituation. Das Urnenmodell sieht vor, aus einer Grundmenge von Probanden die Hälfte mit Wirkstoff A und die anderen mit Placebo B zu versorgen. Wir ziehen aus der Urne zufällig ohne Zurücklegen. Das bedeutet beispielsweise, dass eine hochsensible Person nur einer Gruppe zugehören kann und es keinen Ausgleich gibt. Sie beeinflusst beide Gruppen, denn bei einer wird sich ein Effekt zeigen, bei der anderen aber fehlt eine hochsensible Person. Deshalb sind die Stichproben nicht unabhängig voneinander sind.
Es wird ein Vitaminhaltiges Präparat gegen Erkältung getestet mit einer Grundmenge von Probanden mit 2000 Teilnehmern. Beobachtet wird die mittlere Anzahl Erkältungen in einem gewissen Zeitraum. Die Probanden mit Wirkstoff hatten einen Durchschnitt von $m_1 = 2.2$ Erkältungen bei $s = 2.9$ und die Placebogruppe $m_2 = 2.4$ mit $s = 3.1$. Der gemischte oder "gepoolte" Standardfehler ist

$$SE = \frac{1}{\sqrt{1000}}\sqrt{2.9^2 + 3.1^2} \approx 4.24/31.6 = 0.13$$

Die Testgrösse ist

$$z = \frac{(2.2 - 2.4) - 0}{0.13} \approx -1.54$$

Die Normalverteilung ergibt $\Phi(1.54) = 0.9382$ und damit $\Phi(-1.54) = 0.0618$ oder rund 6%. Damit würde man die Nullhypothese, wonach die Mittelwerte gleich sind, nicht verwerfen.
Wir kommen nochmals auf die Voraussetzungen zurück. Wir haben zwei Fehler gemacht:

- Abhängigkeit infolge fehlendem Zurücklegen,
- Kovarianz weggelassen beim Standardfehler.

Den ersten Fehler haben wir schon besprochen,. Der zweite ergibt sich aus der Gleichung auf Seite 33-10, wo wir wegen Unabhängigkeit den Term $-E(Y_1 Y_2)$ null gesetzt haben. Dieser

Term würde tendenziell das SE verringern, und somit die Testgrösse vergrössern. Der erste Fehler verringert aber die Differenz im Zähler. Somit wirken sich die Fehler entgegen und machen den Test zulässig.

33.2.4 Testverteilung

Zweidimensionale Wahrscheinlichkeit

Wir haben bereits Zufallsfunktionen eingeführt. Nun kann eine Funktion z von zwei unabhängigen Zufallsgrössen abhängen. Wenn z die Summe der Augenzahl von zwei Würfeln ist, die jeweils einmal geworfen werden, dann ist $Z = X + Y$. Wir kennen die Verteilung des ersten Würfels f_X und des zweiten Würfels f_Y. Wenn sie nicht fair sind, dann unterscheiden sich die Dichten. Hier sind die beiden Würfel unabhängig, d.h., die Wahrscheinlichkeiten multiplizieren sich. Eine Wurf mit $x = 5$ und $y = 2$ hat die Wahrscheinlichkeit $P(X = 5, Y = 2) = P(X = 5) \cdot P(Y = 2)$. Man beachte, dass die gemeinsame Wahrscheinlichkeit nicht davon abhängt, was wir mit den zugehörigen x und y machen, ob wir sie addieren, multiplizieren oder irgend eine Funktion $w(x, y)$ anwenden.

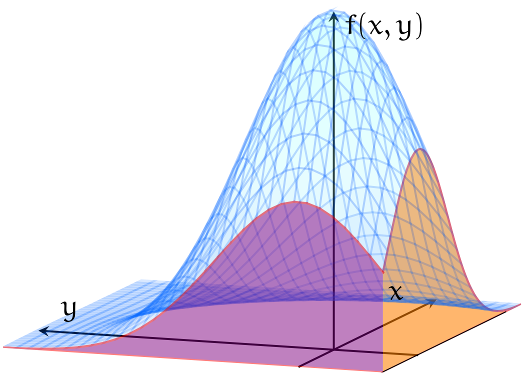


Abbildung 33.3: Dichte einer zweidimensionalen Verteilung $f_{X,Y}(x, y)$ mit zwei Schnitten.

In der Abbildung 33.3 sieht man die gemeinsame Dichte zweier Zufallsgrössen, die normalverteilt sind. Die Darstellung zeigt eine Fläche, die wir bei einem bestimmten x-Wert und y-Wert aufgeschnitten haben. Man erkennt, dass die Schnittflächen zu den typischen Glockenkurven gehören. Diese Dichte ist $f_Z(x, y) = \phi(x)\phi(y)$.
Die gemeinsame Dichte $f(x, y)$ zweier stetiger unabhängiger Zufallsgrössen X und Y ist

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

Man bedenke, dass dies auch für Funktionen $u(X)$ und $w(Y)$ gilt, d.h. $f(u(X), w(Y)) = f_u \cdot f_w$. Zum Beispiel würfelt man mit dem ersten Würfel und verwendet diese Augenzahl, um soviel mal mit dem zweiten Würfel zu werfen und die Augenzahl zu addieren. Die Fragestellung ist dann z.B., wie gross ist die Wahrscheinlichkeit, eine Summe von 12 zu werfen.

Summen und Faltung

Wir erinnern den Doppelwurf mit dem Würfel und die kumulierte Augenzahl. Wenn X die Augenzahl des ersten Würfels ist und Y die des zweiten, so kann man für die Wahrscheinlichkeit schreiben

$$\begin{aligned} f_{X+Y}(z) &= P(X+Y=z) \\ &= \sum_{\{x,y)|x+y=z\}} P(X=x, Y=y) \\ &= \sum_x P(X=x, Y=z-x) \\ &= \sum_x f_X(x) f_Y(z-x) \end{aligned}$$

Wir haben die Wahrscheinlichkeiten schon berechnet in Abschnitt 32.2.1. Es ist

z	2	3	4	5	6	7	8	9	10	11	12
$f_{X+Y}(z)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Σ	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Dass wir hier einfach die Wahrscheinlichkeiten multiplizieren dürfen, ist dem Umstand geschuldet, dass X und Y unabhängig sind. Mit der Formel kann man natürlich noch weitere Summanden hinzunehmen. Diese Summe wird *Faltung* genannt.

Es ist intuitiv einfach, dieses Summe als Integral für kontinuierliche Verteilungen zu erweitern. Wir nehmen wieder zwei unabhängige Variablen X und Y, welche die Wahrscheinlichkeiten f_X und f_Y aufweisen. Wir schreiben analog zu oben

$$\begin{aligned} P(X+Y \leqslant z) = F &= \iint_{\{(x,y)|x+y\leqslant z\}} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x) f_Y(y) dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_X(x) f_Y(u-x) du dx \qquad \text{mit} \quad u = x+y \\ \frac{dF(z)}{du} = f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \end{aligned}$$

Satz 33.10. Faltung Es sind X_1 und X_2 zwei stetige Zufallsgrössen mit Wahrscheinlichkeitsdichten f_1 und f_2. Dann gilt für die Dichte der Summe $X_1 + X_2$:

$$g(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z-x) dx.$$

Für zwei diskrete Zufallsgrössen X_1 und X_2 ergibt sich die Wahrscheinlichkeitsfunktion der Summe $f_{X_1+X_2}$ zu:

$$f_{X_1+X_2}(z) = \sum_x f_{X_1}(x) \cdot f_{X_2}(z-x).$$

33.11 Übung Wir wollen zwei standardnormalverteilte Variablen X und $Y \sim \Phi(0,1)$ addieren und somit deren Wahrscheinlichkeiten falten. Dazu nehmen wir die obige Formel sowie die Wahrscheinlichkeitsdichten der Normalverteilung

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$$

womit folgt

$$\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} \phi(x)\phi(z-x)\,dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(\frac{-x^2}{2}\right)\exp\left(\frac{-(z-x)^2}{2}\right)dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(\frac{-x^2-(x-z)^2}{2}\right)dx \\
&= \frac{1}{2\pi}\exp\left(\frac{-z^2}{2}\right)\int_{-\infty}^{\infty} \exp\left(\frac{-2x^2-2xz}{2}\right)dx \\
&= \frac{1}{2\pi}\exp\left(\frac{-z^2}{2}\right)\int_{-\infty}^{\infty} \exp\left(-[x^2+xz]\right)dx \\
&= \frac{1}{2\pi}\exp\left(\frac{-z^2}{2}\right)\int_{-\infty}^{\infty} \exp\left(-[(x+\frac{z}{2})^2-\frac{z^2}{4}]\right)dx \\
&= \frac{1}{2\pi}\exp\left(\frac{-z^2}{2}+\frac{z^2}{4}\right)\underbrace{\int_{-\infty}^{\infty} \exp\left(-[(x+\frac{z}{2})^2]\right)dx}_{\sqrt{\pi}} \\
&= \frac{1}{\sqrt{4\pi}}\exp\left(\frac{-z^2}{4}\right)
\end{aligned}$$

Wir haben die Formel $\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}dt = \sqrt{2\pi}$ von Pierre Simon Laplace angewendet, die auch Teil der Normalverteilung ist. Zudem haben wir quadratisch ergänzt, d.h. $x^2+xz = (x+z/2)^2 - z^2/4$ verwendet. Nun vergleichen wir das Resultat mit der Normalverteilung mit den Parametern μ und σ^2:

$$\phi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Wir sehen, dass die Summe zweier Zufallsvariablen, welche die Dichte $\phi(0,1)$ haben, zu $f_{X+Y} = \phi(0,2)$ sich summieren. Allgemein gilt

$$\left.\begin{array}{l} X \sim \Phi(\mu_X, \sigma_X^2) \\ Y \sim \Phi(\mu_Y, \sigma_Y^2) \end{array}\right\} \Rightarrow (X+Y) \sim \Phi(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)$$

Man nennt das Phänomen, dass Summen die Verteilung der Summanden beibehalten, die *Reproduktionseigenschaft*. Diese gilt für normalverteilte Zufallsgrössen und für ein paar andere, aber nicht allgemein. ◁

X^2-Verteilung

Wir kommen nun zurück zum Beispiel mit den Knaben- und Mädchengeburten von Übung 33.3 auf Seite 33-5. Wir hatten gefunden, dass das Quadrat Y einer normalverteilten Zufallsgrösse $X \sim \Phi(x; 0, 1)$ gemäss $G_1 = 2\Phi(\sqrt{y}; 0, 1)$ verteilt ist. Für $y < 0$ ist $G_1(y) = 0$. Der Zusammenhang zwischen Dichte g und Verteilung G ist $g(x) = \frac{dG(x)}{dx}$. Hier also

$$g_1(y) = \frac{dG_1(y)}{dy} = \frac{2\cdot d\Phi(\sqrt{y})}{dy} = 2\frac{\phi(\sqrt{y})}{\sqrt{y}} = 2\cdot y^{-\frac{1}{2}}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-y}{2}\right)$$

Uns interessiert die Summe von Quadraten von Zufallsgrössen, die normalverteilt sind. Dazu verwendet man die vollständige Induktion und zwar mit einer Variante, bei der man von n

auf $n+2$ schliesst. Dies ist mit viel weniger Aufwand verbunden. Allerdings muss man die Verankerung für $n = 1$ und $n = 2$ zuerst zeigen. Dann verwendet man den behaupteten Satz wie folgt.

Satz 33.12. Die Dichte der X_n^2-Verteilung ist gegeben durch

$$g_n(y) = \begin{cases} \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} y^{\frac{n}{2}-1} \exp\left(-\frac{y}{2}\right) & \text{für} \quad y > 0 \\ 0 & \text{für} \quad y \leqslant 0 \end{cases}$$

Es ist $E(y) = \frac{(\frac{n-1}{2})!}{(\frac{n}{2}-1)!} = \mu$ und $V(y) = \mu - \mu^2$.

Anmerkung 33.13. Die Grösse $(\frac{n}{2})!$, eine Fakultät von halben Zahlen, ist folgendermassen zu verstehen. Es ist $(\frac{n}{2}-1)!\frac{n}{2} = (\frac{n}{2})!$ und speziell $(\frac{-1}{2})! = \sqrt{\pi}$. Also $(\frac{5}{2})! = \frac{5}{2}\frac{3}{2}\frac{1}{2}(\frac{-1}{2})! = \frac{5}{2}\frac{3}{2}\frac{1}{2}\sqrt{\pi} = \frac{15}{8}\sqrt{\pi}$.

Verankerung für $n = 1$, ist gemäss Dichte für $n = 1$. Nun für $n = 2$: Wir müssen die Faltung anwenden gemäss

$$g_2(y) = \int_{-\infty}^{\infty} g_1(x) g_1(y-x) dx$$

Nun kann man die Grenzen noch reduzieren, denn es muss ja $x > 0$ und $y - x > 0$ sein oder kurz $0 < x < y$. Somit

$$\begin{aligned} g_2(y) &= \int_0^y 2 \cdot x^{-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x}{2}\right) 2 \cdot (y-x)^{-\frac{1}{2}} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(y-x)}{2}\right) dx \\ &= \frac{2}{\pi} \int_0^y x^{-\frac{1}{2}} (y-x)^{-\frac{1}{2}} \cdot \exp\left(\frac{-x-(y-x)}{2}\right) dx \\ &= \frac{2}{\pi} \exp\left(\frac{-y}{2}\right) \int_0^y x^{-\frac{1}{2}} (y-x)^{-\frac{1}{2}} dx \\ &= \frac{2}{\pi} \exp\left(\frac{-y}{2}\right) \int_0^y \frac{1}{\sqrt{x(y-x)}} dx \end{aligned}$$

Nun müssen wir das Integral lösen mit einer geschickten Substitution. Wir versuchen

$$u = \sqrt{1 - \frac{x}{y}} \qquad \Leftrightarrow \qquad x = y(1-u^2)$$

sowie $dx/du = -2uy$ und $dx = -2uy\,du$. Die Grenzen sind $0 = 1 - u^2$ und damit $u_0 = 1$ sowie $y = y(1-u^2)$ woraus $u_1 = 0$. In das Integral eingesetzt folgt

$$\int_1^0 \frac{-2uy\,du}{\sqrt{y(1-u^2)yu^2}} = \int_0^1 \frac{2\,du}{\sqrt{(1-u^2)}} = 2\arcsin(u)\Big|_0^1 = 2\frac{\pi}{2} = \pi$$

Daraus folgt endlich

$$g_2(y) = \frac{1}{2} \exp\left(\frac{-y}{2}\right)$$

Das entspricht ebenfalls dem Satz 33.12 für $n = 2$. Nun schliessen wir von n auf $n+2$, wiederum mit dem Faltungsintegral

$$g_{n+2}(y) = \int_{-\infty}^{\infty} g_n(x) g_2(y-x)\, dx$$

Eingesetzt

$$\begin{aligned}
g_{n+2}(y) &= \int_0^y \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) \cdot \frac{1}{2} \exp\left(\frac{-(y-x)}{2}\right) dx \\
&= \frac{1}{2} \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} \int_0^1 x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) \cdot \exp\left(\frac{-(y-x)}{2}\right) dx \\
&= \frac{1}{2} \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} \exp\left(\frac{-y}{2}\right) \int_0^1 x^{\frac{n}{2}-1} \, dx \\
&= \frac{1}{2^{\frac{n}{2}+1}(\frac{n}{2}-1)!} \exp\left(\frac{-y}{2}\right) \frac{2}{n} x^{\frac{n}{2}} \Big|_0^y \\
&= \frac{1}{2^{\frac{n}{2}}(\frac{n}{2})!} \exp\left(\frac{-y}{2}\right) y^{\frac{n}{2}}
\end{aligned}$$

Wir haben $(\frac{n}{2}-1)!n = (\frac{n}{2})!$ verwendet und Übereinstimmung mit der vermuteten allgemeinen Formel für die Dichte erzielt. Somit ist der Satz 33.12 bewiesen. Damit können wir einiges anfangen.

Zur besseren Vorstellung zeigen wir die Verläufe dieser Verteilungsdichte mit einem Parameter.

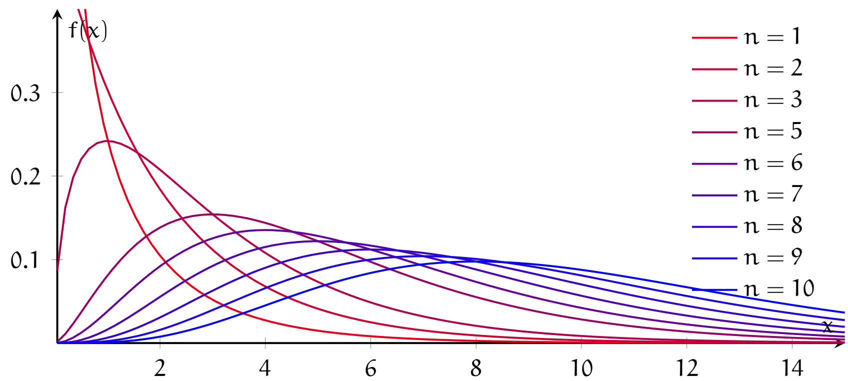


Abbildung 33.4: X_n^2-Verteilungsdichten

In der Abbildung sieht man mit ein bisschen Vorstellungsvermögen, dass mit wachsendem Freiheitsgrad die Dichte sich der Normalverteilung annähert, also glockenförmig wird.

Anpassungstest

Wir erweitern das Beispiel mit den Knaben- und Mädchengeburten mit einem Würfel. Hier gibt es nicht zwei sondern sechs Möglichkeiten. Wie bei den Geburten gibt es eine Schliessbedingung, die besagt, dass wenn man fünf Ergebnisse kennt das sechste gegeben ist, analog zu den Knabengeburten, die aus den Mädchengeburten gegeben sind.

Das Experiment sieht 60 Würfe voraus, die wie folgt ausgefallen sind

Augenzahl	1	2	3	4	5	6
Experiment H_i	4	6	17	16	8	9
Erwartet	10	10	10	10	10	10

Wenn man eine Augenzahl wie 3 betrachtet und das Gegenereignis Nicht-Drei, denn entspricht dies dem Binomialmodell mit den Wahrscheinlichkeit $b(x) = \binom{x}{n} p^n (1-p)^{n-x}$ mit dem Erwartungswert $\mu = p$ und der Varianz $\sigma^2 = np(1-p)$. Mit $n = 60$ ist $\sigma^2 = 60 \cdot \frac{1}{6}\frac{5}{6} = \frac{50}{60}$

$$X^2 = \sum_{i=1}^{6} \frac{\Big(\text{beobachtete Häufigkeit} - \text{erwartete Häufigkeit}\Big)^2}{\text{erwartete Häufigkeit}}$$

und konkret

$$X^2 = \sum_{i=1}^{6} \frac{(H_i - 10)^2}{10} = \frac{1}{10}(36 + 16 + 49 + 36 + 4 + 1) = \frac{142}{10} = 14.2$$

Wenn die tatsächlichen Häufigkeiten nahe bei den erwarteten sind, ist die Testgrösse klein. Hier sind die Abweichungen vom erwarteten Wert zum Teil sehr hoch, was sich in der Testgrösse wiederfindet. Nun kennen wir die Verteilung der Testgrösse, welche die Summe von sechs Zufallsgrössen ist, nämlich die mit der Faltung hergeleitete Verteilung G_k, die aus g_k durch Summieren entsteht. Hier ist $k = 6 - 1 = 5$. Man bezeichnet k als Freiheitsgrad, es ist aber die Anzahl Summanden weniger eins wegen der Schliessbedingung. Statistikpakete können solche einfachen Test in Windeseile ausführen und sogleich den p-Wert von 0.01439 ausgeben. Von Alters her nimmt man die Tabelle zur Hand und findet für den Freiheitsgrad $n = 5$:

X^2-Quadrat

n	0.001	0.005	0.01	0.025	0.05	0.10	0.125	0.20	0.25	0.333	0.50
1	0.000	0.000	0.000	0.001	0.004	0.016	0.025	0.064	0.102	0.186	0.455
2	0.002	0.010	0.020	0.051	0.103	0.211	0.267	0.446	0.575	0.811	1.386
3	0.024	0.072	0.115	0.216	0.352	0.584	0.692	1.005	1.213	1.568	2.366
4	0.091	0.207	0.297	0.484	0.711	1.064	1.219	1.649	1.923	2.378	3.357
5	0.210	0.412	0.554	0.831	1.145	1.610	1.808	2.343	2.675	3.216	4.351
n	0.60	0.667	0.75	0.80	0.875	0.90	0.95	0.975	0.990	0.995	0.999
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	$\boxed{12.833}$	$\boxed{15.086}$	16.750	20.515

Nun interpoliert man den Wert mit

$$q = \frac{14.2 - 12.833}{15.086 - 12.833}(0.99 - 0.975) + 0.975 = 0.9841$$

damit wird $(1 - q) = 0.0159$ und der interpolierte p-Wert ist 0.0159. Mit diesem Wert würde man die Nullhypothese, alle Augenzahlen sind gleichhäufig, verwerfen, der Würfel ist nicht fair. Zur Abrundung schauen wir nochmals den Geburtenfall an, es war $X^2 = 10.29$ bei $n = 1$ und wir hatten den p-Wert von 0.0014 errechnet. Aus der Tabelle wäre interpoliert:

$$q = \frac{10.29 - 7.879}{10.828 - 7.879}(0.999 - 0.995) + 0.995 = 0.9983$$

und damit $(1 - 0.9983) = 0.0017$. Das lineare interpolieren ist hier nicht besonders genau.

Wichtig 10. Mit dem X^2-Anpassungstest kann man testen, ob die mehreren Ausprägungen in der Stichprobe die Vermutung einer *bestimmten Verteilung* (Urnenmodell) der Zufallsvariablen in der Grundgesamtheit zulassen. ⊣

Hier war die Verteilung eine Gleichverteilung.

Test der Unabhängigkeit

Der X^2-Test eignet sich auch zur Klärung der Unabhängigkeit von Merkmalen. Die Daten sind meist in Tabellen verfügbar. Beispielsweise wird untersucht, ob Linkshänder unter Männern häufiger ist als unter Frauen. Die Merkmale sind Rechtshänder, Linkshänder und Beidhänder. Eine Tabelle oder Tafel sieht wie folgt aus

Händigkeit	Männer	Frauen	Total
Rechts	934	1070	2004
Links	113	92	205
Beid	20	8	28
Total	1067	1170	2237

Wir haben zwei Verteilungen, eine für Männer und eine für Frauen, die wir vergleichen wollen. Wir wollen die Stichprobe mit der Erwartung koppeln. Was ist die Erwartung? Die Nullhypothese ist, es gibt keinen Unterschied. Somit wird erwartet, dass die totalen Häufigkeiten die vermuteten sind. Für Rechtshänder also $\frac{2004}{2237} = 89.6\%$, für Linkshänder $\frac{205}{2237} = 9.2\%$ und für Beidhänder $\frac{28}{2237} = 1.2\%$. Diese Werte multiplizieren wir mit der Anzahl Probanden und erhalten die erwarteten Werte, z.B. für rechtshändige Männer $1067 \cdot 0.896 = 956$. Tabellarisch

Händigkeit	Männer	Frauen
Rechts	959	1048
Links	98	107
Beid	13	15

Die Anzahl Freiheitsgrade ist bei einer $m \times n$-Tafel $k = (m-1)(n-1)$. In beide Richtungen gibt es zeilen- und kolonnenweise eine Schliessbedingung. In diesen Beispiel sind $m = 2$ und $m = 3$ und somit $k = 1 \cdot 2 = 2$.
Wir bestimmen die Testgrösse

$$X^2 = \frac{(934-959)^2}{959} + \frac{(113-98)^2}{98} + \frac{(20-13)^2}{13} + \frac{(1070-1048)^2}{1048} + \frac{(92-107)^2}{107} + \frac{(8-15)^2}{15}$$

Ausgerechnet ist

$$X^2 \approx 12$$

Aus der obigen Tabelle sehen wir für $n = 2$, dass das Quantil grösser als 0.995 und damit der p-Wert bei weniger als 0.005 liegt. Dies ist klar im Ablehnungebereich. Wir folgern, dass die Händigkeit geschlechterspezifisch ist.

Transformation

Wie kann man generell Verteilungen von Funktionen von Zufallsvariablen bestimmen? Wir haben in Abschnitt 33.2.2 auf Seite 33-5 bereits ein Beispiel gemacht für $Y = X^2$. Ohne es weiter zu erwähnen, kommt bereits in dieser Transformation die Inverse der Verteilung vor. Das impliziert, dass für Transformationen die Verteilung invertierbar und damit auch monoton sein muss. Das sei vorausgesetzt.
Als erstes betrachten wir den Fall einer Zufallsvariablen. Es ist $y = u(x)$ und $x = w(y)$, u und w sind ihre Inversen. Die Dichte von X ist $f(x)$. Nun ist

$$P(a < Y < b) = P\big(w(a) < X < w(b)\big) = \int\limits_{w(a)}^{w(b)} f(x)dx = \int\limits_{a}^{b} f(w(y))\frac{dw}{dy}dy = G(y)$$

Weil $\frac{dw}{dy}$ entweder immer positiv oder immer negativ ist (weil $\frac{dw}{dy} \neq 0$ sein muss), je nach Verlauf von $w(y)$, kann man die Betragszeichen um die Ableitung legen und so keine Fallunterscheidung machen. Die Dichte von Y folgt dann aus

$$g(y) = \frac{dG(y)}{dy} = f(w(y))|\frac{dw}{dy}| \tag{33.1}$$

Nun betrachten wir den Fall von Funktionen zweier Variablen X und Y, speziell von Quotienten $Z = \dfrac{Y}{X}$. X und Y seien unabhängig und ihre Verteilungen seien f_X und f_Y. Die gemeinsame Wahrscheinlichkeit ist

$$P(Z \leqslant z) = \iint\limits_{\{(x,y)|y/x \leqslant z\}} f_X(x)f_Y(y)dxdy$$

Nun muss man zwei Fälle unterscheiden:

- $x > 0$, dann $y \leqslant xz$ und
- $x < 0$, dann $y \geqslant xz$.

Deshalb muss man diese zwei Bedingungen im Integral berücksichtigen und wir schreiben

$$P(Z \leqslant z) = \underbrace{\int\limits_{-\infty}^{0} \Big[\int\limits_{xz}^{\infty} f_X f_Y \, dy \Big] \, dx}_{x<0} + \underbrace{\int\limits_{0}^{\infty} \Big[\int\limits_{-\infty}^{xz} f_X f_Y \, dy \Big] \, dx}_{x>0}$$

Nun machen wir die Substitution $y = xv$ und $dy/dv = x$ oder $dy = x\,dv$ mit den Grenzen $y_0 = xz$, daraus $v_0 = z$.

$$\begin{aligned} P(Z \leqslant z) &= \int\limits_{-\infty}^{0} \Big[\int\limits_{z}^{-\infty} x \cdot f_X(x)f_Y(xv)\,dv \Big] \, dx + \int\limits_{0}^{\infty} \Big[\int\limits_{-\infty}^{z} x \cdot f_X(x)f_Y(xv)\,dv \Big] \, dx \\ &= \int\limits_{-\infty}^{0} \Big[\int\limits_{-\infty}^{z} (-x) \cdot f_X(x)f_Y(xv)\,dv \Big] \, dx + \int\limits_{0}^{\infty} \Big[\int\limits_{-\infty}^{z} x \cdot f_X(x)f_Y(xv)\,dv \Big] \, dx \end{aligned}$$

Weil im ersten Summanden $x < 0$ und $(-x)$ steht, im zweiten gilt $x > 0$ und x, kann man zusammenfassen zu

$$P(Z \leqslant z) = \int_{-\infty}^{z} \int_{-\infty}^{\infty} |x| \cdot f_X(x) f_Y(xv)\, dx\, dv$$

Aus der Wahrscheinlichkeitsfunktion folgt die Dichte durch Ableiten nach v zu

$$f_Z(z) = \int_{-\infty}^{\infty} |x| \cdot f_X(x) f_Y(xz)\, dx \tag{33.2}$$

Das ist die Formel für die Bestimmung von Dichten von Quotienten unabhängiger Zufallsgrössen. Wir werden diese Formel zweimal brauchen. konkret muss man die Dichten einsetzen und x "ausintegrieren". Wir gehen die Berechnung nochmals durch für $Z = \frac{Y}{aX}$. Überall würde man anstatt z dann az schreiben. Nur bei der Substitution wäre dann $dy = a \cdot x\, dv$. Und somit

$$f_Z(z) = a \int_{-\infty}^{\infty} |x| \cdot f_X(x) f_Y(x \cdot az)\, dx$$

Die F-Verteilung

Die Testgrösse T ist ein Quotient aus zwei unabhängigen Zufallsgrössen U und V, die beide X^2-verteilt sind mit unterschiedlichen Freiheitsgraden m und n. Also

$$Z = \frac{Y/m}{X/n}$$

Die Dichten sind gemäss Satz 33.12 auf Seite 33-15:

$$g_n(y) = \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} y^{\frac{n}{2}-1} \exp\left(-\frac{y}{2}\right)$$

Das setzen wir in die Gleichung 33.2 ein, wobei wir zur Übersichtlichkeit den Bruch durch A_n und A_m abkürzen und setzen die untere Integrationsgrenze auf null und lassen die Betragszeichen weg, denn die Variablen sind immer positiv. Um die Gleichung 33.2 für Quotienten zu verwenden, setzen wir anstattt z aber $az = z\frac{m}{n}$ mit $a = \frac{m}{n}$ ein, um die Analogie zu vervollkommnen:

$$f_Z(z) = \tfrac{m}{n} A_m A_n \int_0^{\infty} x \cdot x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) \cdot (xz\tfrac{m}{n})^{\frac{m}{2}-1} \exp\left(-\frac{xz\frac{m}{n}}{2}\right) dx$$

Alles, was nicht x enthält vorziehen und Terme zusammenfassen ergibt

$$f_Z(z) = \tfrac{m}{n} A_m A_n (z\tfrac{m}{n})^{\frac{m}{2}-1} \int_0^{\infty} x^{\frac{n+m}{2}-1} \exp\left(-\frac{x}{2}(1 + z\tfrac{m}{n})\right) dx$$

Nun müssen wir das Integral lösen und machen als erstes eine Substitution, um den Ausdruck einfacher darzustellen. Es bietet sich der Ausdruck des Exponenten an:

$$u = \frac{x}{2}(1 + z\tfrac{m}{n}) \qquad \Leftrightarrow \qquad x = \frac{2}{1 + z\frac{m}{n}} u$$

$$\frac{du}{dx} = \frac{1}{2}(1 + z\tfrac{m}{n}) \qquad \Leftrightarrow \qquad dx = \frac{2}{1 + z\frac{m}{n}}\, du$$

Damit folgt dann mit Ersetzen von x und dx:

$$f_Z(z) = \tfrac{m}{n} A_m A_n (z\frac{m}{n})^{\frac{m}{2}-1} (\frac{2}{1 + z\frac{m}{n}})^{\frac{n+m}{2}} \int\limits_0^\infty u^{\frac{n+m}{2}} \exp\left(-u\right)\, du$$

Das Integral ist im Wesentlichen

$$I_k = \int\limits_0^\infty y^k \exp(-y) dy$$

Durch partielle Integration, wobei der erste Term f' entspricht und der zweite g folgt

$$I_k = -y^k \exp(-y)\Big|_0^\infty + k \int_0^\infty y^{k-1} \exp(-y) dy$$

Der erste Summand ist 0, denn die Exponentialfunktion geht schneller nach 0 als die Potenzfunktion nach Unendlich geht. Somit kann man schreiben

$$I_k = k \cdot I_{k-1}$$

Wendet man die partielle Integration auf I_{k-1} am, so folgt $I_{k-1} = (k-1)I_{k-2}$ etc. Hier also

$$I_k = k!$$

ein wunderbares Resultat! Der letzte Term ist

$$I_0 = \exp(-y)\Big|_0^\infty = 1$$

Wir ersetzten also das Integral durch die Fakultät und erhalten dann

$$f_Z(z) = \tfrac{m}{n} A_m A_n (z\frac{m}{n})^{\frac{m}{2}-1} (\frac{2}{1 + z\frac{m}{n}})^{\frac{n+m}{2}} (\frac{n+m}{2})!$$

Nun setzen wir noch die zwei Term A_m und A_n ein

$$f_Z(z) = \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!} \cdot \frac{1}{2^{\frac{m}{2}}(\frac{m}{2}-1)!} (z\frac{m}{n})^{\frac{m}{2}-1} (\frac{2}{1 + z\frac{m}{n}})^{\frac{n+m}{2}} (\frac{n+m}{2})!$$

Mit Kürzen der 2 und schöner Gruppieren folgt dann das Endresultat für unsere gesuchte Dichte des Quotienten zweier X^2-verteilter Zufallsgrössen zu:

$$f_Z(z) = \frac{(\frac{n+m}{2})!}{(\frac{n}{2}-1)!(\frac{m}{2}-1)!} \left(\tfrac{m}{n}\right)^{\frac{m}{2}} (z)^{\frac{m}{2}-1} \left(1 + z\tfrac{m}{n}\right)^{-\frac{n+m}{2}}$$

Üblicherweise sind die Werte tabelliert, wobei aufgrund der zwei Parameter diese Tafeln recht gross sind. Statistikprogramme enthalten entsprechende Prozeduren.

Formel 33.14. Dichte F-Verteilung Die Dichte der F-Verteilung ist

$$f_{n,m}(f) = \begin{cases} \frac{(\frac{n+m}{2})!}{(\frac{n}{2}-1)!(\frac{m}{2}-1)!} \left(\frac{m}{n}\right)^{\frac{m}{2}} (f)^{\frac{m}{2}-1} \left(1 + f\frac{m}{n}\right)^{-\frac{n+m}{2}} & \text{für} \quad f > 0 \\ 0 & \text{für} \quad f \leqslant 0 \end{cases}$$

33.15 Übung Das Standardbeispiel für den F-Test sind Vergleiche von Varianzen aus zwei unterschiedlichen Urnen, wobei der Mittelwert als gleich angenommen wird. Konkret ist die Preisschwankung von Tickets on-line oder im Laden gleich? Es liegen zwei Stichproben A und B vor mit 4 und 5 Elementen. Die Nullhypothese ist $H_0 : \sigma_A^2 = \sigma_B^2$, die Alternative $H_A : \sigma_A^2 \neq \sigma_B^2$. Man bestimme die Testgrösse, wobei die grössere Zahl in den Zähler, die kleinere in den Nenner kommt! Man wählt das Signifikanzniveau (oder Irrtumswahrscheinlichkeit) α und entscheidet sich für einen zweiseitigen Test. Im Versuch ergeben sich die Varianzen $s_1 = 4.2$ und $s_2 = 4.7$ mit $m = 4$ und $n = 5$ Elementen. Die Freiheitsgrade sind jeweils um 1 reduziert, also $\nu_1 = 3$ und $\nu_2 = 4$. Daraus folgt die Prüfgrösse

$$t = \frac{4.2/3}{4.7/4} \approx 1.19$$

Die Entscheidung sei $\alpha = 0.05$ zweiseitig, also der kritische Wert $q_{1-\alpha/2}(\nu_1, \nu_2)$. Aus der Tabelle 33.3.4 auf Seite 33-36 folgt für k der Wert $k = 9.98$. Nun ist $t < k$, wir verwerfen die Nullhypothese nicht. ◁

In der folgenden Abbildung33.5 sieht man ein paar Kurven, die mit wachsender Anzahl Freiheitsgrade, hier $m = n = 30$ sich der Normalverteilung annähern, obwohl die Verteilung nur für $x > 0$ gilt.

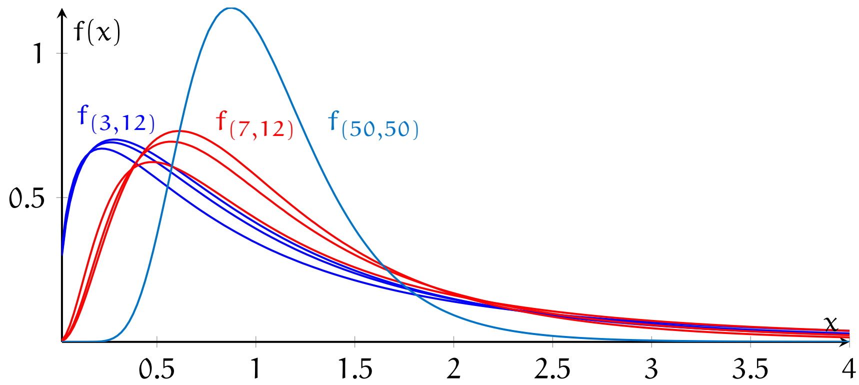


Abbildung 33.5: F-Verteilungsdichten mit zwei Parametern

Die t-Verteilung

Eine normalverteilte Zufallsvariable X führt zur standardnormalverteilten Zufallsgrösse

$$Z = \frac{X - \mu}{\sigma/\sqrt{n}} \quad \sim \quad \Phi(0, 1)$$

Nun ist es etwas unrealistisch, den Mittelwert einer Grundgesamtheit nicht zu kennen und Vermutungen darüber anzustellen, aber die Varianz σ^2 als fest anzunehmen. Mit der Stichprobenvarianz wird die Testgrösse

$$T = \frac{X - \mu}{S\sqrt{n}} \quad \sim \quad F?$$

X und S sind unabhängig. Wir studieren zuerst aber

$$Z = \frac{X}{\sqrt{Y/n}} = \frac{X}{\frac{1}{\sqrt{n}}Q}$$

Wir kennen die Verteilung von Y und suchen nun die von Q. Wie in Gleichung 33.1 dargelegt, ist $q = \sqrt{y}$ und $y = w(q) = q^2$. Damit ist $\frac{dw(q)}{dq} = 2q$. Somit folgt die Verteilungsdichte von Q:

$$f_Q(q) = \frac{1}{2^{\frac{n}{2}}(\frac{n}{2}-1)!}(q^2)^{\frac{n}{2}-1}\exp\left(-\frac{q^2}{2}\right)\cdot 2q$$

Die Verteilungsdichte des Zählers ist standardnormal, d.h.

$$f_X(x) = \phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)$$

Mit den Gleichungen 33.2.4 und 33.1 folgt:

$$\begin{aligned}
f_Z(z) &= a\int\limits_{-\infty}^{\infty}|q|\cdot f_Q(q)f_X(q\cdot az)\,dq \\
&= \frac{1}{\sqrt{n}}\int_0^\infty q\cdot\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}(q^2)^{\frac{m}{2}-1}\exp\left(-\frac{q^2}{2}\right)\cdot 2q\cdot\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{q^2z^2}{2n}\right)\,dq \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}\int_0^\infty(q^2)^{\frac{m}{2}}\exp\left(-\frac{q^2(1+\frac{z^2}{n})}{2}\right)\,dq
\end{aligned}$$

Wir substituieren $u = \frac{q^2}{2}(1+\frac{z^2}{n})$ und berechnen $q^2 = \frac{2u}{1+\frac{z^2}{n}}$, $q = [\frac{2u}{1+\frac{z^2}{n}}]^{1/2}$und $\frac{dq}{du} = [\frac{2u}{1+\frac{z^2}{n}}]^{-1/2}$. Damit wird die Dichte

$$\begin{aligned}
f_Z(z) &= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}\int_0^\infty(q^2)^{\frac{m}{2}}\exp\left(-\frac{q^2(1+\frac{z^2}{n})}{2}\right)\,dq \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}\int_0^\infty\left(\frac{2u}{1+\frac{z^2}{n}}\right)^{\frac{m}{2}}\exp(-u)\left[\frac{2u}{1+\frac{z^2}{n}}\right]^{-1/2}\,du \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}(1+\frac{z^2}{n})^{-\frac{m-1}{2}}\int\limits_0^\infty(2u)^{\frac{m-1}{2}}\exp(-u)\,du \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}(1+\frac{z^2}{n})^{-\frac{m+1}{2}}2^{\frac{m-1}{2}}\int_0^\infty u^{\frac{m-1}{2}}\exp(-u)\,du \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{\sqrt{n}}\frac{1}{2^{\frac{m}{2}}(\frac{n}{2}-1)!}(1+\frac{z^2}{n})^{-\frac{m+1}{2}}2^{\frac{m-1}{2}}(\frac{m-1}{2})! \\
&= \frac{1}{\sqrt{\pi n}}\frac{(\frac{m-1}{2})!}{(\frac{n}{2}-1)!}(1+\frac{z^2}{n})^{-\frac{m+1}{2}} \\
&= \frac{1}{\sqrt{\pi n}}\frac{(\frac{m+1}{2}-1)!}{(\frac{n}{2}-1)!}(1+\frac{z^2}{n})^{-\frac{m+1}{2}}
\end{aligned}$$

Anstatt $(n-1)!$ schreibt man auch $\Gamma(n)$, die sogenannte *Gammafunktion*.

Formel 33.16. Dichte t-Verteilung Die Dichte der t-Verteilung ist

$$f_n(t) = \begin{cases} \frac{1}{\sqrt{\pi n}}\frac{(\frac{n+1}{2}-1)!}{(\frac{n}{2}-1)!}(1+\frac{t^2}{n})^{-\frac{n+1}{2}} & \text{für} \quad t > 0 \\ 0 & \text{für} \quad t \leqslant 0 \end{cases}$$

Es ist $E(t) = 0$ für $n \geqslant 1$ und $V(t) = \frac{n}{n-2}$ für $n \geqslant 3$.

Satz 33.17. Eine Zufallsstichprobe der Länge n mit Mittelwert $\overline{X}$ und Stichprobenvarianz S^2 aus einer normverteilten Grundmenge mit μ und σ^2 hat die Testgrösse

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

Diese ist t-verteilt mit $n-1$ Freiheitsgraden.

Die Herleitung von Testverteilungen wird in den meisten Anfängerdarstellungen unterlassen, weil, wie gesehen, der Aufwand erheblich sein kann. Dann fallen aber die Verteilungen wie vom Himmel und man gerät in den Rezeptmodus. Deshalb kann es nicht schaden, sich ein bisschen durchzubeissen.

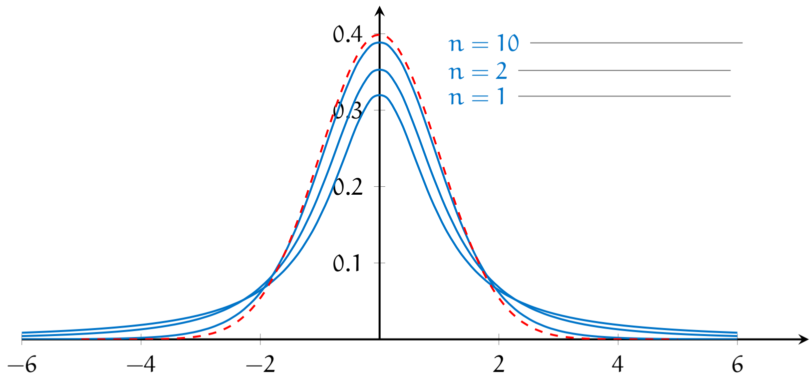


Abbildung 33.6: t-Verteilungsdichten, rot Normalverteilung

In der Abbildung 33.6 sieht man gut, dass mit wachsendem Freiheitsgrad die t-Verteilung zur Normalverteilung konvergiert. Bedeutsam ist die Tatsache, dass an den beiden Enden der t-Verteilung viel mehr Fläche ("Masse") liegt, die für die Irrtumswahrscheinlichkeit wesentlich ist. Sie ist *endlastig*. Für kleine n sind die Unterschiede sehr gross, weshalb man für kleine Stichproben nicht die Normalverteilung als Testverteilung nehmen sollte.

33.2.5 Warnhinweise

Ein paar Merkpunkte:

- Lege die Irrtumswahrscheinlichkeit α vor dem Experiment fest.
- Der p-Wert hängt von der Stichprobenlänge ab. Deshalb kann bei grossem n auch ein kleiner Effekt signifikant erscheinen. Umgekehrt kann ein wesentlicher Effekt bei kleinem n als nicht signifikant erscheinen. n sollte man angeben.
- Gib den p-Wert und den verwendeten Test an anstatt nur "signifikant" oder "hoch signifikant" zu melden.
- Es sollte immer ein Urnenmodell für die Stichprobenziehung darstellbar sein.
  - Wenn ein Signifikanztest mit der Grundgesamtheit ausgeführt wird, pass auf.
  - Wenn die Stichprobe unsystematisch erzeugt wurde, sei vorsichtig.
- Ein Signifikanztest prüft nicht das Testverfahren.

33.3 Schätztheorie

Wir wenden uns dem zweiten Problemkreis zu. Wir haben eine Zufallsstichprobe aus der Grundgesamtheit genommen. Wir haben keine Vermutung, Nullhypothese zur Grundmenge, sondern benutzen die Stichprobe, um gewisse Parameter zu schätzen. Beispielsweise ist es nicht möglich oder wirtschaftlich, alle produzierten Glühbirnen zu testen. Deshalb nimmt man ein möglichst aussagekräftige Stichprobe, um die fehlerhaften Birnen abzuschätzen.
Nun kann man einen festen Wert suchen, einen *Punktschätzer* oder ein Intervall, in dem sich der "wahre" Wert befinden sollte. Dann spricht man von einem *Intervallschätzer*.

33.3.1 Schätzfunktion

Wir stellen uns die Frage, die Intensität einer Poissonverteilung aus einer Stichprobe zu schätzen. Diese ist $x = (0, 0, 0, 2, 0, 2, 0, 0, 1, 1, 0, 0, 2, 1, 1, 0, 0, 2, 0, 0)$. Das sind die toten Soldaten des Corps II in den Jahren 1875-1894, die durch Hufschlag gestorben sind. Für die Poissonverteilung wissen wir, dass gilt

$$\lambda = E(X) \qquad \text{und} \qquad \lambda = V(X)$$

Wenn man nun den Erwartungswert mit dem Stichprobenmittelwert oder die Varianz mit der Stichprobenvarianz identifiziert, dann hat man zwei mögliche Schätzer für denselben Parameter. Bei unserer Stichprobe folgt $m = 0.6$ und $s^2 = 0.674$ mit $n - 1$ als Nenner.

Definition 47. Ist $(X_1, X_2, \ldots, X_n)$ eine Stichprobe aus der Zufallsgrösses $X \sim F_\vartheta$, so heisst jede reelwertige Funktion

$$T_n = g(X_1, X_2, \ldots, X_n)$$

Schätzfunktion (oder Schätzer) für den reellen Parameter ϑ der Zufallsgrösse X.
Der *Schätzwert* $t_n = g(x_1, x_2, \ldots, x_n)$ ist eine Realisierung von T_n in der Stichprobe.

Diese Definition verdeutlicht das Beispiel. Denn *jede* Funktion ist sehr allgemein. Man kann sogar soweit gehen, die beobachteten Daten ganz zu ignorieren. Aber nochmals: es gibt viele mögliche Schätzer für Parameter aus einer Stichprobe.
Um diese Beliebigkeit einzuschränken, macht man sich zusätzliche Überlegungen zu Eigenschaften der Schätzfunktion und zu Optimierungskriterien. Daraus folgen die Beurteilungskriterien

(1) erwartungstreu,

(2) konsistent (asymptotisch zutreffend),

(3) effizient,

(4) suffizient, erschöpfend.

Das letzte Kriterium werden wir nicht behandeln.

33.3.2 Beurteilungskriterien

Erwartungstreue

Definition 48. Eine Schätzfunktion T_n für den Parameter ϑ einer Wahrscheinlichkeitsverteilung von X heisst *erwartungstreu*, wenn

$$E(T_n) = \vartheta$$

Eine Schätzfunktion ist also erwartungstreu, wenn sie im Mittel den wahren Wert ϑ liefert.

33.1 Übung Ist der Stichprobenmittelwert eine erwartungstreue Schätzfunktion? Gemäss Voraussetzung hat jede Grösse der Srichprobe X_j dieselbe Verteilung, also gleichen Erwartungswert $E(X_j) = \mu$ und gleiche Varianz $V(X_j) = \sigma^2$. Der Erwartungswert der Summe ist

$$E(\frac{1}{n}[X_1 + X_2 + \ldots + X_n]) = \frac{1}{n} E(X_1) + E(X_2) + \ldots + E(X_n) = \frac{1}{n} n \cdot \mu = \mu$$

Der Schätzer ist erwartungstreu. $\triangleleft$

33.2 Übung Nun die Frage, ob die Stichprobenvarianz s^2 erwartungstreu ist. Wir benutzen die Variable k, um nachher entweder n oder $n-1$ einzusetzen. (Nach unserer Definition ist $n-1$ zu verwenden.) Es ist m der Stichprobenmittelwert $m = \sum_{i=1}^{n} X_i$. Wir verwenden die Erweiterung $(X - m) = (X - \mu) - (m - \mu)$ im folgenden mit $V(m) = \frac{\sigma^2}{n}$

$$\begin{aligned}
E(s^2) &= E\left[\frac{1}{k}\sum_{i=1}^{n}(X_j - m)^2\right] \\
&= \frac{1}{k} E\left[\sum_{i=1}^{n}((X_j - \mu) - (m - \mu))^2\right] \\
&= \frac{1}{k} E\left[\sum_{i=1}^{n}(X_j - \mu)^2 - \sum_{i=1}^{n} 2(X_j - \mu)(m - \mu)) + \sum_{i=1}^{n}(m - \mu))^2\right] \\
&= \frac{1}{k} E\left[\sum_{i=1}^{n}(X_j - \mu)^2 - \sum_{i=1}^{n} 2(X_j - \mu)(m - \mu)) + \sum_{i=1}^{n}(m - \mu))^2\right] \\
&= \frac{1}{k}\left[\sum_{i=1}^{n} E(X_j - \mu)^2 - 2\,E[(m - \mu)\underbrace{\sum_{i=1}^{n}[X_j - \mu]}_{n(m-\mu)} + n\,E(m - \mu)^2\right] \\
&= \frac{1}{k}\left[\sum_{i=1}^{n} E(X_j - \mu)^2 - n\,E(m - \mu)^2\right] \\
&= \frac{1}{k}[n\,V(X) - n\,V(m))] = \frac{1}{k}\left[n\,V(X) - n\frac{V(X)}{n}\right] \\
&= \frac{n-1}{k} V(X) = \frac{n-1}{k}\sigma^2
\end{aligned}$$

Wir sehen, dass man für die Stichprobenvarianz den Nenner $k = n - 1$ wählen muss, damit der Schätzer erwartungstreu ist, d.h. $E(s^2) = \sigma^2$. $\triangleleft$

Nun machen wir eine etwas absurde Annahme. Für die Stichprobe $(X_1, X_2, \ldots, X_n)$ mit $E(X_j) = \mu$ nehmen wir den den Wert x_5, d.h. den Wert der fünften Ziehung ($n > 5$ angenommen). Nun ist ja auch $E(X_5) = \mu$ uns somit dieser Schätzer erwartungstreu. Da dieser Schätzer intuitiv viel schlechter als der Mittelwert ist, muss es zusätzliche Kriterien geben, um die Güte eines Schätzers abzuschätzen.

Effizienz

Die Grundidee ist bei mehreren erwartungstreuen Schätzfunktionen für einen Parameter ϑ, diejenige zu bevorzugen, welche bei gegebenen Stichprobenumfang n die kleinste Streuung, und somit Varianz, hat.

Definition 49. Ein erwartungstreue Schätzfunktion $T_n = g(X_1, X_2, \ldots, X_m)$ für den Parameter ϑ heisst *effizient*, falls für jede andere Schätzfunktion $H_n = h(X_1, X_2, \ldots, X_m)$ gilt

$$V(T_n) \leqslant V(H_n).$$

33.3 Übung Wir vergleichen die zwei Schätzfunktionen $T_1 = \sum_{j=1}^{1} 0X_j$ mit $T_2 = X_5$. Die Varianzen sind $V(T_1) = \sigma^2/10$ und $V(T_2) = \sigma$. Somit ist der Mittelwert viel effizienter als ein einzelner Wert. $\triangleleft$

Konsistenz

Definition 50. Eine Schätzfunktion $T_n = g(X_1, X_2, \ldots, X_n)$ für den Parameter ϑ heisst *konsistent* oder *asymptotisch zutreffend*, wenn die Folge (T_n) mit wachsendem n gegen ϑ konvergiert als

$$P(|T_n - \vartheta| < \epsilon) \to 1 \quad \text{für} \quad n \to \infty.$$

Anmerkung 33.4. Beim Grenzwert in der Definition spricht man von Konvergenz "in Wahrscheinlichkeit", denn einzelne Werte von der Zufallsgrösse T_n können durchaus zufällig mit n anwachsen. Was damit gemeint ist, sieht man in der folgenden Abbildung 33.7

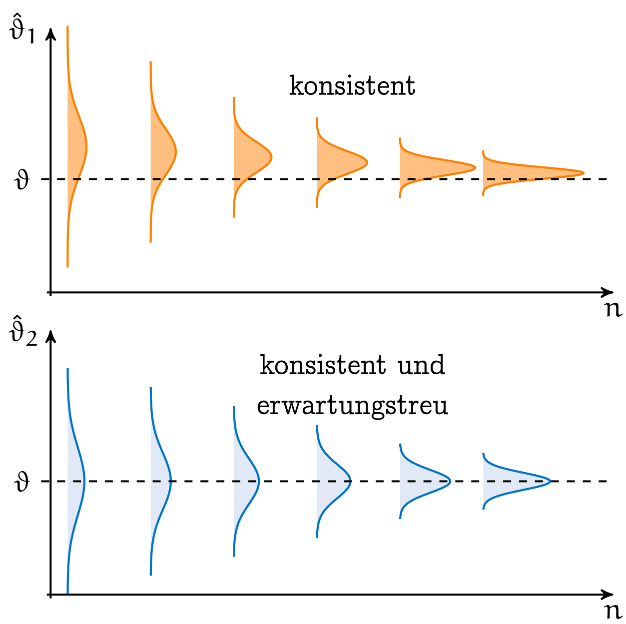


Abbildung 33.7: Konsistenz

Satz 33.5. Ungleichung von Tschebyscheff Für eine Zufallsgrösse mit Mittelwert μ und Varianz σ^2 gilt für jedes $t > 0$:

$$P(|X - \mu| > t) \leqslant \frac{\sigma^2}{t^2}$$

Im stetigen Fall kann betrachten wir den Bereich $D = \{x||x - \mu| > t\}$ und schreiben

$$P(|X - \mu| > t) = \int\limits_{\{x||x-\mu|>t} f(x)dx$$

Falls $x \in D$, ist

$$\frac{|x - \mu|^2}{t} \geqslant 1$$

Somit folgt

$$P(|X - \mu| > t) = \int\limits_D f(x)dx \leqslant \int\limits_D \frac{(x-\mu)^2}{t^2} f(x)dx \leqslant \int\limits_{-\infty}^{\infty} \frac{(x-\mu)^2}{t^2} f(x)dx = \frac{\sigma^2}{t^2}$$

Somit ist es gezeigt. Der Satz besagt, dass wenn die Varianz sehr klein ist, es sehr wahrscheinlich ist, dass X stark von μ abweicht. Dies erkannt man besonders gut, wenn man schreibt $t = k\sigma$

$$P(|X - \mu| > k\sigma) \leqslant \frac{1}{k^2}$$

Für diskrete Zufallsgrössen gilt bei Summation nur über D, d.h. x-Werte, welche die Ungleichung erfüllen, so ist diese Summe kleiner als die Summe über beliebige x, also

$$\sigma^2 = \sum_{j=-\infty}^{\infty} (x_j - \mu)^2 P(X = x_j) > \sum_{\{x_j||x_j-\mu|>t\}} (x_j - \mu)^2 P(X = x_j)$$

Für die rechten Terme gilt

$$(x_j - \mu)^2 > t^2$$

und somit

$$\sigma^2 > t^2 \sum_{\{x_j||x_j-\mu|>t\}} (x_j - \mu)^2 P(X = x_j) = t^2 P(|x - \mu| > t)$$

Damit ist ein Schätzer konsistent, wenn die Stichprobenvarianz mit n kleiner wird.

Satz 33.6. Gegeben die Stichprobe $(X_1, X - 2, \ldots, X_n)$ aus der Zufallsgrösse $X \sim F_X$, dann ist der Stichprobenmittelwert $m = \sum_{j=1}^{n} X_j$ eine erwartungstreu, konsistente und effiziente Schätzfunktion für den Erwartungswert μ von F_X.

Bevor wir weiterschreiten, fassen wir noch zusammen.

Satz 33.7. Gegeben die Stichprobe $(X_1, X - 2, \ldots, X_n)$ aus der Zufallsgrösse $X \sim F_X$, dann ist der Stichprobenvarianz

$$s^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - m)^2$$

eine erwartungstreue Schätzfunktion für die Varianz σ^2 von F_X.

Wir haben bis hierhin Stichprobenmittelwert und -varianz als Schätzer verwendet. Mann nennt dies *Momentenschätzer*, denn Mittelwert ist das erste und die Varianz das zweite Moment.

33.3.3 Schätzmethoden

Maximale Mutmasslichkeit, Maximum-Likelihood

Für die Verwendung dieser Methode ist die Kenntnis der Verteilung unerlässlich.

Satz 33.8. Prinzip der maximalen Mutmasslichkeit Eine Zufallsstichprobe $(X_1, X_2, \ldots, X_n)$ aus der Zufallsgrösse $X \sim F_\vartheta$ hat eine Realisierung $(a_1, a_2, \ldots, a_n)$. Als Schätzwert für den Parameter ϑ dient dann jeder Wert $\hat{\vartheta}$, für den die Wahrscheinlichkeit

$$P_\vartheta(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n)$$

maximal wird.

Wir erkennen im obigen Prinzip eine Extremalaufgabe, das Finden eines Maximums. Dazu brauchen wir die entsprechende Funktion und die Ableitbarkeit. Die Funktion ergibt sich aus dem Prinzip und der Tatsache, dass die einzelnen Züge in der Stichprobe voneinander unabhängig sind und damit deren Wahrscheinlichkeiten multipliziert werden. Die definieren die Likelihood-Funktion $L(\vartheta)$ wie folgt

$$L(\vartheta) = P_\vartheta(X_1 = a_1) \cdot P_\vartheta(X_2 = a_2) \cdot \ldots \cdot P_\vartheta(X_m = a_n)$$

oder für stetige Zufallsgrössen

$$L(\vartheta) = f_\vartheta(X_1) \cdot f_\vartheta(X_2) \cdot \ldots \cdot f_\vartheta(X_n)$$

Mit der Logarithmieren der Funktion gehen die Produkte in Summe über. Der Logarithmus ist eine streng monotone Funktion, so dass die Lage der Extrema nicht verändert werden. Somit

$$\log(L(\vartheta)) = \sum_{j=1}^{n} \log(P_\vartheta(X_j))$$

Vorgehensweise

(1) Setze die Stichprobendaten in die Dichten $f_\vartheta(a_j)$ oder $P_\vartheta(a_j)$ ein

$$L(\vartheta) = f_\vartheta(a_1) \cdot f_\vartheta(a_2) \cdot \ldots \cdot f_\vartheta(a_n)$$

(2) Finde $\hat{\vartheta}$, das die Funktion maximiert

$$\frac{d}{d\vartheta} L(\vartheta) = f_\vartheta(a_1) \cdot f_\vartheta(a_2) \cdot \ldots \cdot f_\vartheta(a_n) = 0$$

oder

$$\frac{d}{d\vartheta} \log(L(\vartheta)) = \sum_{j=1}^{n} \log(P_\vartheta(X_j)) = 0$$

(3) $\hat{\vartheta}$ ist der Maximum-Likelihood-Schätzer.

33.9 Übung Wir bestimmen nun den Maximum-Likelihood-Schätzer für eine poissonverteilte Stichprobe $a = (0, 0, 0, 2, 0, 2, 0, 0, 1, 1, 0, 0, 2, 1, 1, 0, 0, 2, 0, 0)$. Die Verteilung folgt aus $f_\lambda =$

$e^{-\lambda}\frac{\lambda^x}{x!}$. Die logarithmierte Likelihoodfunktion ist

$$\begin{aligned}
\log(L(\lambda) &= \log\left(e^{-\lambda}\frac{\lambda^{X_1}}{X_1!}\cdot e^{-\lambda}\frac{\lambda^{X_2}}{X_2!}\cdot\ldots\cdot e^{-\lambda}\frac{\lambda^{X_n}}{X_n!}\right)\\
&= \log\left(e^{-n\lambda}\frac{\lambda^{X_1}}{X_1!}\cdot\frac{\lambda^{X_2}}{X_2!}\cdot\ldots\cdot\frac{\lambda^{X_n}}{X_n!}\right)\\
&= \log\left(e^{-n\lambda}\frac{\lambda^{X_1+X-2+\ldots+X_n}}{X_1!}\cdot\frac{1}{X_2!}\cdot\ldots\cdot\frac{1}{X_n!}\right)\\
&= -n\lambda+\log(\lambda)\sum_{j=1}^{n}X_j+\sum_{j=1}^{n}\frac{1}{X_j!}
\end{aligned}$$

Ableiten nach λ und nullsetzen

$$0=-n+\frac{1}{\hat{\lambda}}\sum_{j=1}^{n}X_j$$

und somit

$$\hat{\lambda}=\frac{1}{n}\sum_{j=1}^{n}X_j$$

Und mit den Zahlen $\hat{\lambda}=\frac{12}{20}=0.6$.
Der Maximum-Likelihood-Schätzer ist wiederum der Stichprobenmittelwert. ◁

33.10 Übung Wir wollen die Zusammensetzung einer Urne bestimmen. Sie enthält rote und andere Kugeln. Das Experiment ist Ziehen mit Zurücklegen. Das Modell ist die Binomialverteilung gemäss $P(X=k)=\binom{n}{k}p^k\cdot(1-p)^{n-k}$. Wir wollen aus der Stichprobe $a=(0,0,0,1,0,1,0,0,1,1,0,0,1,1,1,0,0,1,0,0)$, wobei a_j entweder 1 oder 0 ist, bestimmen. Es hat $k=8$ rote Kugeln in der Stichprobe.

(1) Die Likelihood-Funktion ist

$$L(\vartheta)=P_\vartheta(X_1=a_1)\cdot P_\vartheta(X_2=a_2)\cdot\ldots\cdot P_\vartheta(X_m=a_n)=\binom{n}{k}p^n(1-p)^{n-k}$$

(2) Ableiten mach p (Produktregel), nullsetzen

$$\begin{aligned}
0 &= \binom{k}{n}\left[kp^{k-1}(1-p)^{n-k}+p^k(1-p)^{n-k-1}(n-k)\right]\\
&= -\binom{n}{k}\left[p^{k-1}(1-p)^{n-k-1}(np-k)\right]
\end{aligned}$$

Mit dem Nullproduktsatz ist jeder Faktor null, insbesondere

$$0=np-k \qquad\Leftrightarrow\qquad p=\frac{k}{n}$$

Auch hier liefert der ML-Schätzer, wie man häufig schreibt, das intuitive Resultat. ◁

Nun gibt es ja bekanntlich Verteilungen mit mehr als einem Parameter. Dann ist wie folgt zu verfahren: Man stellt die Likelihoodfunktion auf und bestimmt die Maxima der Funktionen, die durch jeweiliges Ableiten nach den Parametern entstehen.

Für die Normalverteilung $\Phi(\mu, \sigma^2)$ ist die log-Likelihoodfunktion

$$\begin{aligned}
\log(L(\mu,\sigma) &= \sum_{j=1}^{n} \log\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(X_1-\mu)^2}{2\sigma^2}\right)\right) \\
&= -\frac{1}{2\sigma^2}\sum_{j=1}^{n}(X_j-\mu)^2 - \sum_{j=1}^{n}\log\left(\sigma\sqrt{2\pi}\right) \\
&= -\frac{1}{2\sigma^2}\sum_{j=1}^{n}(X_j-\mu)^2 - n\log\left(\sigma\sqrt{2\pi}\right) \\
&= -\frac{1}{2\sigma^2}\sum_{j=1}^{n}(X_j-\mu)^2 - \frac{1}{2}n\log\left(\sigma^2 4\pi\right) \\
&= -\frac{1}{2\sigma^2}\sum_{j=1}^{n}(X_j-\mu)^2 - \frac{1}{2}n\log\left(\sigma^2\right) - \frac{1}{2}n\log\left(4\pi\right)
\end{aligned}$$

Zuerst ableiten nach μ und nullsetzen, dabei Faktoren ohne μ weglassen ergibt

$$0 = +\sum_{j=1}^{n}(X_j-\mu) \qquad \Leftrightarrow \qquad \sum_{j=1}^{n} X_j - n\mu = 0 \qquad \Leftrightarrow \qquad \mu = \frac{1}{n}\sum_{j=1}^{n} X_j = m$$

Nun die Ableitung nach σ^2, man denke sich σ^2 als eine Zahl u

$$0 = \frac{1}{u^2}\sum_{j=1}^{n}(X_j-\mu)^2 - \frac{n}{2}\frac{1}{u}$$

Mal $u^2/2$ und umgruppieren

$$\sum_{j=1}^{n}(X_j-\mu)^2 = nu \qquad \Leftrightarrow \qquad u = \frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)^2 = \hat{\sigma^2}$$

Wir wissen, dass dieser Schätzer nicht erwartungstreu ist, weil n anstelle von $n-1$ steht.

33.3.4 Intervallschätzer

Bis hierhin haben wir Punktschätzer betrachtet, die zur weiteren Berechnung gebraucht werden aber wenig Auskunft geben über ihre Genauigkeit. Wir betrachten nur ein Beispiel, das genügen soll, den Grundgedanken zu vermittel. Wir nehmen die für die Normalverteilung ermittelten Maximum-Likelihood-Schätzer für Erwartungswert und Varianz, d.h.

$$\hat{\mu} = m \qquad \text{und} \qquad \hat{\sigma^2} = \frac{n-1}{n}s^2$$

(Wir haben ja s^2 mit $n-1$ definiert und berücksichtigen, dass der ML-Schätzer eben nicht s^2 liefert.) Von der Normalverteilten Zufallsgrössen wissen wir (Satz 33.17, dass die Zufallsfunktion

$$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$$

t-verteilt ist mit $n-1$ Freiheitsgraden. Mit eine Quantil, das unten und oben der Verteilung jeweils $\alpha/2$ abgrenzt, folgt für die Wahrscheinlichkeit unter Berücksichtigung, dass die Verteilung symmetrisch ist,

$$P(-t_{n-1}(\alpha/2) \leqslant T \leqslant t_{n-1}(\alpha/2) = 1-\alpha$$

eingesetzt

$$P(-t_{n-1}(\alpha/2) \leqslant \frac{\overline{X}-\mu}{S/\sqrt{n}} \leqslant t_{n-1}(\alpha/2) = 1-\alpha$$

und damit und $\overline{X} = m$

$$P\left(m - t_{n-1}(\alpha/2)S/\sqrt{n} \leqslant \mu \leqslant m + t_{n-1}(\alpha/2)S/\sqrt{n}) = 1-\alpha\right)$$

Dies ist das geschätzte Intervall mit der Irrtumswahrscheinlichkeit α.
Nun hat die Varianz eine andere Testfunktion, nämlich

$$Z = \frac{(n-1)}{\sigma^2}s^2 \quad \sim \quad X^2_{n-1}$$

Analog zu oben setzen wir die Wahrscheinlichkeit für Z

$$P\left(X_{n-1}(1-\alpha/2) \leqslant \frac{(n-1)}{\sigma^2}s^2 \leqslant X_{n-1}(\alpha/2)\right) = 1-\alpha$$

und umgeformt, invertiert

$$P\left(\frac{s^2(n-1)}{X_{n-1}(1-\alpha/2)} \leqslant \sigma^2 \leqslant \frac{s^2(n-1)}{X_{n-1}(\alpha/2)}\right) = 1-\alpha$$

Weil die X^2-Funktion nicht symmetrisch ist, sind die Grenzen nicht gleichweit weg von σ^2. Für Intervallschätzer sind die Verteilungen der Schätzfunktion notwendig.

Damit lassen wir es für dieses Einführung bewenden.

χ^2_ν-Quantile in %

ν	0.1%	0.5%	1.0%	2.5%	5.0%	10.0%	12.5%	20.0%	25.0%	33.3%	50.0%
1	0.000	0.000	0.000	0.001	0.004	0.016	0.025	0.064	0.102	0.186	0.455
2	0.002	0.010	0.020	0.051	0.103	0.211	0.267	0.446	0.575	0.811	1.386
3	0.024	0.072	0.115	0.216	0.352	0.584	0.692	1.005	1.213	1.568	2.366
4	0.091	0.207	0.297	0.484	0.711	1.064	1.219	1.649	1.923	2.378	3.357
5	0.210	0.412	0.554	0.831	1.145	1.610	1.808	2.343	2.675	3.216	4.351
6	0.381	0.676	0.872	1.237	1.635	2.204	2.441	3.070	3.455	4.074	5.348
7	0.598	0.989	1.239	1.690	2.167	2.833	3.106	3.822	4.255	4.945	6.346
8	0.857	1.344	1.646	2.180	2.733	3.490	3.797	4.594	5.071	5.826	7.344
9	1.152	1.735	2.088	2.700	3.325	4.168	4.507	5.380	5.899	6.716	8.343
10	1.479	2.156	2.558	3.247	3.940	4.865	5.234	6.179	6.737	7.612	9.342
11	1.834	2.603	3.053	3.816	4.575	5.578	5.975	6.989	7.584	8.514	10.341
12	2.214	3.074	3.571	4.404	5.226	6.304	6.729	7.807	8.438	9.420	11.340
13	2.617	3.565	4.107	5.009	5.892	7.042	7.493	8.634	9.299	10.331	12.340
14	3.041	4.075	4.660	5.629	6.571	7.790	8.266	9.467	10.165	11.245	13.339
15	3.483	4.601	5.229	6.262	7.261	8.547	9.048	10.307	11.037	12.163	14.339
16	3.942	5.142	5.812	6.908	7.962	9.312	9.837	11.152	11.912	13.083	15.338
17	4.416	5.697	6.408	7.564	8.672	10.085	10.633	12.002	12.792	14.006	16.338
18	4.905	6.265	7.015	8.231	9.390	10.865	11.435	12.857	13.675	14.931	17.338
19	5.407	6.844	7.633	8.907	10.117	11.651	12.242	13.716	14.562	15.859	18.338
20	5.921	7.434	8.260	9.591	10.851	12.443	13.055	14.578	15.452	16.788	19.337
21	6.447	8.034	8.897	10.283	11.591	13.240	13.873	15.445	16.344	17.720	20.337
22	6.983	8.643	9.542	10.982	12.338	14.041	14.695	16.314	17.240	18.653	21.337
23	7.529	9.260	10.196	11.689	13.091	14.848	15.521	17.187	18.137	19.587	22.337
24	8.085	9.886	10.856	12.401	13.848	15.659	16.351	18.062	19.037	20.523	23.337
25	8.649	10.520	11.524	13.120	14.611	16.473	17.184	18.940	19.939	21.461	24.337
26	9.222	11.160	12.198	13.844	15.379	17.292	18.021	19.820	20.843	22.399	25.336
27	9.803	11.808	12.879	14.573	16.151	18.114	18.861	20.703	21.749	23.339	26.336
28	10.391	12.461	13.565	15.308	16.928	18.939	19.704	21.588	22.657	24.280	27.336
29	10.986	13.121	14.256	16.047	17.708	19.768	20.550	22.475	23.567	25.222	28.336
30	11.588	13.787	14.953	16.791	18.493	20.599	21.399	23.364	24.478	26.165	29.336
35	14.688	17.192	18.509	20.569	22.465	24.797	25.678	27.836	29.054	30.894	34.336
40	17.916	20.707	22.164	24.433	26.509	29.051	30.008	32.345	33.660	35.643	39.335
45	21.251	24.311	25.901	28.366	30.612	33.350	34.379	36.884	38.291	40.407	44.335
50	24.674	27.991	29.707	32.357	34.764	37.689	38.785	41.449	42.942	45.184	49.335
55	28.173	31.735	33.570	36.398	38.958	42.060	43.220	46.036	47.610	49.972	54.335
60	31.738	35.534	37.485	40.482	43.188	46.459	47.680	50.641	52.294	54.770	59.335

χ^2_ν-Quantile in %

ν	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	12.833	15.086	16.750	20.515
6	6.211	6.867	7.841	8.558	9.992	10.645	12.592	14.449	16.812	18.548	22.458
7	7.283	7.992	9.037	9.803	11.326	12.017	14.067	16.013	18.475	20.278	24.322
8	8.351	9.107	10.219	11.030	12.636	13.362	15.507	17.535	20.090	21.955	26.125
9	9.414	10.215	11.389	12.242	13.926	14.684	16.919	19.023	21.666	23.589	27.877
10	10.473	11.317	12.549	13.442	15.198	15.987	18.307	20.483	23.209	25.188	29.588
11	11.530	12.414	13.701	14.631	16.457	17.275	19.675	21.920	24.725	26.757	31.264
12	12.584	13.506	14.845	15.812	17.703	18.549	21.026	23.337	26.217	28.300	32.910
13	13.636	14.595	15.984	16.985	18.939	19.812	22.362	24.736	27.688	29.819	34.528
14	14.685	15.680	17.117	18.151	20.166	21.064	23.685	26.119	29.141	31.319	36.123
15	15.733	16.761	18.245	19.311	21.384	22.307	24.996	27.488	30.578	32.801	37.697
16	16.780	17.840	19.369	20.465	22.595	23.542	26.296	28.845	32.000	34.267	39.252
17	17.824	18.917	20.489	21.615	23.799	24.769	27.587	30.191	33.409	35.718	40.790
18	18.868	19.991	21.605	22.760	24.997	25.989	28.869	31.526	34.805	37.156	42.312
19	19.910	21.063	22.718	23.900	26.189	27.204	30.144	32.852	36.191	38.582	43.820
20	20.951	22.133	23.828	25.038	27.376	28.412	31.410	34.170	37.566	39.997	45.315
21	21.991	23.201	24.935	26.171	28.559	29.615	32.671	35.479	38.932	41.401	46.797
22	23.031	24.268	26.039	27.301	29.737	30.813	33.924	36.781	40.289	42.796	48.268
23	24.069	25.333	27.141	28.429	30.911	32.007	35.172	38.076	41.638	44.181	49.728
24	25.106	26.397	28.241	29.553	32.081	33.196	36.415	39.364	42.980	45.559	51.179
25	26.143	27.459	29.339	30.675	33.247	34.382	37.652	40.646	44.314	46.928	52.620
26	27.179	28.520	30.435	31.795	34.410	35.563	38.885	41.923	45.642	48.290	54.052
27	28.214	29.580	31.528	32.912	35.570	36.741	40.113	43.195	46.963	49.645	55.476
28	29.249	30.639	32.620	34.027	36.727	37.916	41.337	44.461	48.278	50.993	56.892
29	30.283	31.697	33.711	35.139	37.881	39.087	42.557	45.722	49.588	52.336	58.301
30	31.316	32.754	34.800	36.250	39.033	40.256	43.773	46.979	50.892	53.672	59.703
35	36.475	38.024	40.223	41.778	44.753	46.059	49.802	53.203	57.342	60.275	66.619
40	41.622	43.275	45.616	47.269	50.424	51.805	55.758	59.342	63.691	66.766	73.402
45	46.761	48.510	50.985	52.729	56.052	57.505	61.656	65.410	69.957	73.166	80.077
50	51.892	53.733	56.334	58.164	61.647	63.167	67.505	71.420	76.154	79.490	86.661
55	57.016	58.945	61.665	63.577	67.211	68.796	73.311	77.380	82.292	85.749	93.168
60	62.135	64.147	66.981	68.972	72.751	74.397	79.082	83.298	88.379	91.952	99.607

t_ν-Quantile in %

ν	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
35	0.255	0.434	0.682	0.852	1.170	1.306	1.690	2.030	2.438	2.724	3.340
40	0.255	0.434	0.681	0.851	1.167	1.303	1.684	2.021	2.423	2.704	3.307
45	0.255	0.434	0.680	0.850	1.165	1.301	1.679	2.014	2.412	2.690	3.281
50	0.255	0.433	0.679	0.849	1.164	1.299	1.676	2.009	2.403	2.678	3.261
55	0.255	0.433	0.679	0.848	1.163	1.297	1.673	2.004	2.396	2.668	3.245
60	0.254	0.433	0.679	0.848	1.162	1.296	1.671	2.000	2.390	2.660	3.232
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Quantile der F-Verteilung

$\nu_2 \backslash \nu_1$	q	2	3	4	5	6	7	8	10	12	15	20	30	50	∞
1	0.900	49.5	53.6	55.8	57.2	58.2	59.1	59.7	60.5	61.0	61.5	62.0	62.6	63.0	63.3
	0.950	199.	216.	225.	230.	234.	237.	239.	242.	244.	246.	248.	250.	252.	254.
	0.975	800.	864.	900.	922.	937.	948.	957.	969.	977.	985.	993.			
	0.990														
	0.999														
2	0.900	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.39	9.41	9.43	9.44	9.46	9.47	9.49
	0.950	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
	0.975	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5
	0.990	99.0	99.2	99.2	99.3	99.3	99.4	100.	100.	100.	100.	100.	100.	100.	99.5
	0.999	999.	999.												
3	0.900	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.23	5.22	5.20	5.18	5.17	5.15	5.13
	0.950	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.74	8.70	8.66	8.62	8.58	8.53
	0.975	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.4	14.3	14.3	14.2	14.1	14.0	13.9
	0.990	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.2	27.1	26.9	26.7	26.5	26.4	26.1
	0.999	149.	141.	137.	135.	133.	132.	131.	129.	128.	127.	126.	125.	125.	123.
4	0.900	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.92	3.90	3.87	3.84	3.82	3.79	3.76
	0.950	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.86	5.80	5.75	5.70	5.63
	0.975	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.84	8.75	8.66	8.56	8.46	8.38	8.26
	0.990	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.5	14.4	14.2	14.0	13.8	13.7	13.5
	0.999	61.2	56.2	53.4	51.7	50.5	49.7	49.0	48.0	47.4	46.8	46.1	45.4	44.9	44.1
5	0.900	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.30	3.27	3.24	3.21	3.17	3.15	3.10
	0.950	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.62	4.56	4.50	4.44	4.36
	0.975	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.62	6.52	6.43	6.33	6.23	6.14	6.02
	0.990	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.1	9.89	9.72	9.55	9.38	9.24	9.02
	0.999	37.1	33.2	31.1	29.8	28.8	28.2	27.6	26.9	26.4	25.9	25.4	24.9	24.4	23.8
6	0.900	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.94	2.90	2.87	2.84	2.80	2.77	2.72
	0.950	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.94	3.87	3.81	3.75	3.67
	0.975	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.46	5.37	5.27	5.17	5.07	4.98	4.85
	0.990	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.87	7.72	7.56	7.40	7.23	7.09	6.88
	0.999	27.0	23.7	21.9	20.8	20.0	19.5	19.0	18.4	18.0	17.6	17.1	16.7	16.3	15.7
7	0.900	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.70	2.67	2.63	2.59	2.56	2.52	2.47
	0.950	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.51	3.44	3.38	3.32	3.23
	0.975	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.76	4.67	4.57	4.47	4.36	4.28	4.14
	0.990	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.62	6.47	6.31	6.16	5.99	5.86	5.65
	0.999	21.7	18.8	17.2	16.2	15.5	15.0	14.6	14.1	13.7	13.3	12.9	12.5	12.2	11.7
8	0.900	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.54	2.50	2.46	2.42	2.38	2.35	2.29
	0.950	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.22	3.15	3.08	3.02	2.93
	0.975	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.29	4.20	4.10	4.00	3.89	3.81	3.67
	0.990	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.81	5.67	5.52	5.36	5.20	5.07	4.86
	0.999	18.5	15.8	14.4	13.5	12.9	12.4	12.0	11.5	11.2	10.8	10.5	10.1	9.80	9.33

Index