Kapitel 1

WAHRSCHEINLICHKEITSTHEORIE

by RÉMY

Bemerkung:
Der Oberbegriff **Stochastik** wird hier nicht verwendet, da wir uns SMP nicht mit der überflüssigen **Statistik** beschäftigen.

1.1 Wiederholungen: Unter- & Mittelstufe

1.1.1 Zufallsexperimente

Definition

Als Zufallsexperiment bezeichnet man Versuche, deren Ergebnisse sich nicht vorhersagen lassen, also vom Zufall abhängig sind.
Vor der Durchführung eines Zufallsexperiments muss eine **Ergebnismenge** S festgelegt werden. Sie beinhaltet alle möglichen Ergebnisse: $S = \{e_1, e_2, ..., e_n\}$
Ein Versuch heißt Zufallsexperiment, falls:

- er unter gleichen Bedingungen beliebig oft wiederholbar ist
- alle möglichen Ergebnisse vor Durchführung bekannt sind
- sein Ergebnis sich nicht mit Sicherheit vorhersagen lässt
- bei jeder Durchführung genau ein Ergebnis aus S auftritt

Beispiel:
Bekannte Zufallsexperimente sind:

- das Werfen einer Münze
- das Werfen eines Würfels
- das Ziehen einer Kugel aus einer Urne
- ...

Definition - Laplace-Experiment

Laplace-Experimente sind Experimente, deren Ergebnisse jeweils gleichwahrscheinlich sind.

Beispiel:
Ein Beispiel hierfür wäre der Wurf eines **perfekten** Würfels.

Theorem

In einem Laplace-Experiment gilt für die Wahrscheinlichkeit P, dass ein Ergebnis A von n möglichen Ergebnis eintritt:

$$P(A) = \frac{1}{n}$$

Mehrstufige Zufallsexperimente

Definition

Werden mehrere (n) Zufallsexperimente nacheinander ausgerführt, so kann man sie als ein einziges Zufallsexperiment zusammenfassen. Man nennt dies ein **mehrstufiges Zufallsexperiment**. Die Ergebnisse eines solchen Experiments kann man als geordnete n-Tupel auffassen.

Beispiel:
Zweifaches Werfen einer Münze: $S = \{(Z/Z), (Z/K), (K/Z), (K/K)\}$.

Bemerkung:
Alternativ kann man S als einfache Menge definieren. Bei zweifachem Münzwurf wäre eine mögliche Darstellung $S = \{0, 1, 2\}$ mit der Anzahl an Kopf-Würfen als Ergebnis möglich.

Ereignisse

Definition - Ereignisse

Jede Teilmenge A von der Ergebnismenge S nennt man ein Ereignis. Endet das Zufallsexperiment mit einem Ergebnis aus A, sagt man: A ist eingetreten.

Beispiel:
Werfen eines Würfels: $S = \{1, 2, 3, 4, 5, 6\}$

A: Augenzahl ist gerade:	$A = \{2, 4, 6\}$
B: Augenzahl ist ungerade:	$B = \{1, 3, 5\}$
C: Augenzahl ist Primzahl:	$C = \{2, 3, 5\}$
D: Augenzahl < 7:	$D = S$
E: Augenzahl $= 6$:	$E = \{6\}$
F: Augenzahl > 6:	$F = \{\}$

Bemerkung:

Definition

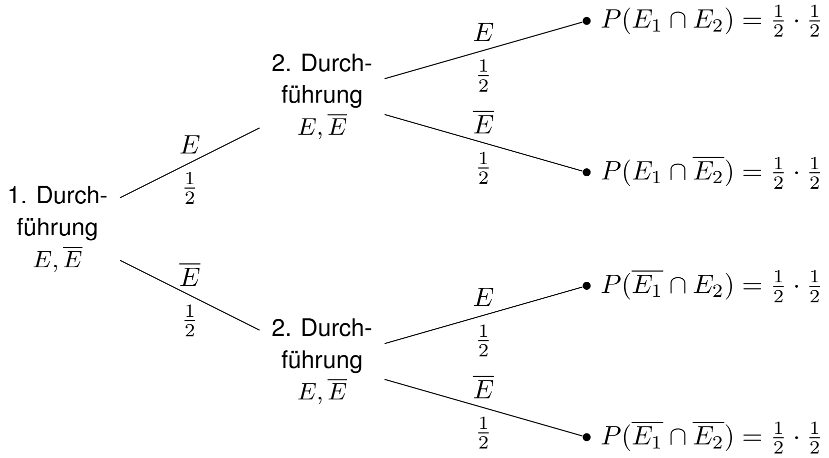
- Ein Ereignis, das nur aus einem Ergebnis besteht, heißt **Elementarereignis**.
- $B = \bar{A}$ (A quer) ist das **Gegenereignis** von A. Es gilt: $B = S \backslash A$
- Ein Ereignis, das immer eintritt, heißt **sicheres Ereignis**
- Ein Ereignis, das niemals eintritt, heißt **unmögliches Ereignis**

Baumdiagramme

Definition - Bernoulli-Experiment

Bei Bernoulli-Experimenten gibt es nur 2 mögliche Ausgänge: Erfolg / Miserfolg ($=$ $\overline{\text{Erfolg}}$). Mehrfaches Ausführen (l Mal) von Bernoulli-Experimenten ergibt eine **Bernoulli-Kette** der Länge l.

Eine Bernoulli-Kette kann als **Baumdiagramm** dargestellt werden:



E bezeichnet das Erfolgs-Ereignis, $\overline{E}$ bezeichnet somit den Miserfolg. Jeder Pfad trägt 2 Informationen: das jeweilige Ereignis und seine Wahrscheinlichkeit.

Definition - Pfadregeln

- Die Summe der Wahrscheinlichkeiten auf den Ästen, die von einem Knoten (Ort der Verzweigung), ist $= 1$.
- Die Wahrscheinlichkeit eines Pfades (also eines Elementarereignisses) ist gleich dem **Produkt** der Wahrscheinlichkeiten aller Äste des Pfades.
- Die Wahrscheinlichkeit eines Ereignissesb ist gleich der **Summe** der Wahrscheinlichkeit der Pfade, die zu diesem Ereignis führen.

Bemerkung:
Besonders beim Urnenmodell eines Zufallsexperiments muss beachtet werden, ob nach dem Ziehen zurückgelegt wird, oder nicht, weil sich die Wahrscheinlichkeiten der Äste sonst entsprechend verändern.

Empirisches Gesetz der großen Zahlen

Definition - Häufigkeiten

Nach der n-fachen Durchführung eines Zufallsexperiments betrachtet man, wie oft Ereignisse eingetreten sind.
Ist das Ereignis A H-mal eingetreten, so nennt man H die **absolute Häufigkeit** und $\frac{H}{n}$ die **relative Häufigkeit** von A.

Theorem

Wird ein Zufallsexperiment sehr häufig durchgeführt, so stabilisieren sich die relativen Häufigkeiten der Ereignisse. Es gilt:

$$\lim_{n\to\infty} \left(\frac{H_n(E)}{n} \right) = P(E)$$

mit E einem Ereignis, und H_n seiner Häufigkeit nach n Wiederholungen.

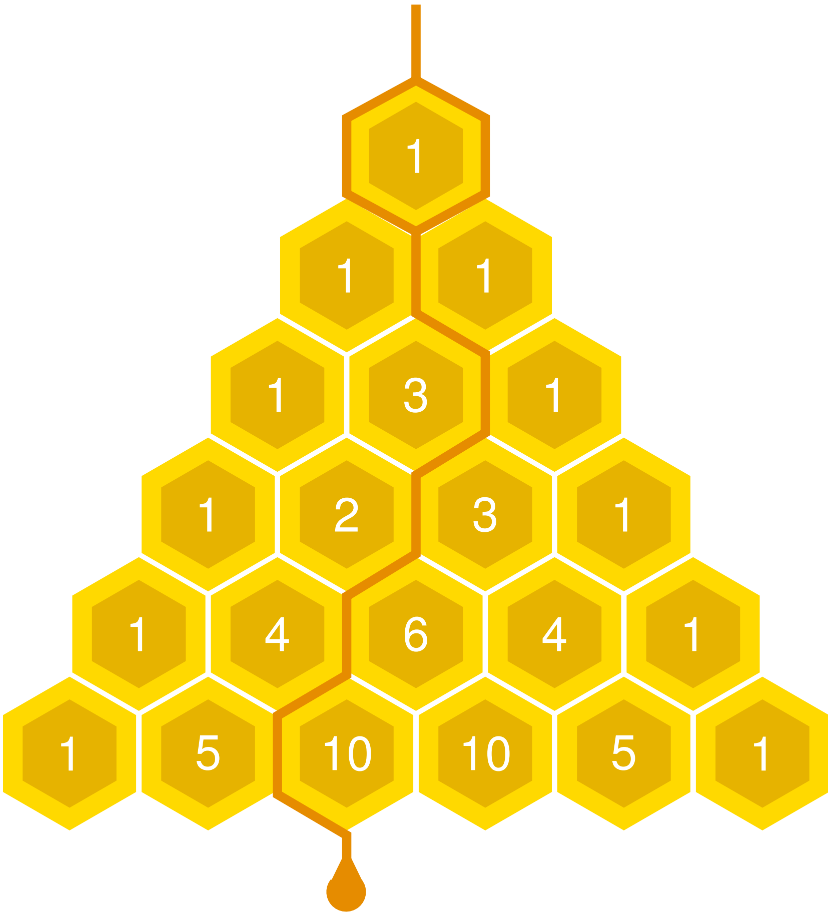
1.1.2 Zufallsvariable

Definition

Sind die Ergebnisse eines Zufallsexperiments Zahlen, oder kann man den Ergebnissen Zahlen zuornen, so nennt man die Variable für diese Zahlen **Zufallsvariable** X.
Mit Hilfe von Zufallsvariablen kann man Zufallsexperimente einfacher bechreiben.

Beispiel:
Zählen von Erfolgen (1) und Miserfolgen (0) bei Bernoulli-Ketten: statt $P(($Erfolg, Erfolg, ..., Erfolg, $))$ (n Mal) schreibt man einfach: $P(x = k * 1)$ mit k der gewünschten Anzahl an Erfolgen.

1.1.3 Das Pascal'sche Dreieck



Bekannt aus der Mittelstufe. Es ergibt sich, wenn man die Summe von zwei Werten eine Stufe tiefer, zwischen die beiden Werte schreibt. Es wird vor Allem für Binomialkoeffizienten verwendet, findet aber auch in der Wahrscheinlichkeitsrechnung Verwendung.

1.2 Kombinatorik

1.2.1 Binomialkoeffizienten

Definition

Der Binomialkoeffizient ist die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge.

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \qquad \text{(gesprochen "n über k")} \qquad \forall 0 \leq k \leq n$$

Bemerkung:
Anschaulich entspricht das den Möglichkeiten, genau k bestimmte Kugeln von n Kugeln zu ziehen, wobei die gezogenen Kugeln nicht zurückgelegt werden, und die Reihenfolge, in der sie gezogen wurden, nicht beachtet wird.

Bemerkung:
Die gefundenen Werte entsprechen den Vorfaktoren, die man für das k-te Element aus der $n$ten Reihe aus dem Pascalschen Dreieck ablesen kann.

Das bedeutet, dass Potenzen von Binomen auch über Binomialkoeffizienten darstellbar sind:

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

Bemerkung:
Die obige Definition gilt nur für $0 \leq k \leq n$ da die Fakultät (!) nicht für negative Zahlen definiert ist.

GTR-Tipp:
Im englischen wird $\binom{n}{r}$ als "n choose r"gesprochen. Auf dem GTR findet sich die Option unter `MATH` > `PRB`. Es handelt sich um `nCr`.
Benutzung: <ZAHL$_1$> `nCr` <ZAHL$_2$>. (Entspricht $\binom{Z_1}{Z_2}$)

Theorem

Für Binomialkoeffizienten gelten mehrere Eigenschaften, unter ihnen wollen wir folgende zwei hervorheben:

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

2.
$$\binom{n}{k} = \binom{n}{n-k}$$

Beweis

1.
$$\begin{aligned}
\binom{n+1}{k+1} &= \frac{(n+1)!}{(n-k)!k!} \\
&= \frac{n!}{k!} \frac{(n+1)}{(n-k)!(k+1)} \\
&= \frac{n!}{k!} \frac{n-k+k+1}{(n-k)!(k+1)} \\
&= \frac{n!}{k!} \left(\frac{n-k}{(n-k)!(k+1)} + \frac{k+1}{(n-k)!(k+1)} \right) \\
&= \frac{n!}{k!} \left(\frac{1}{(n-k-1)!(k+1)} + \frac{1}{(n-k)!} \right) \\
&= \frac{n!}{k!} \left(\frac{1}{(n-(k+1))!(k+1)} + \frac{1}{(n-k)!} \right) \\
&= \frac{n!}{(n-k)!k!} + \frac{n!}{(n-k-1)!(k+1)!} \\
&= \binom{n}{k} + \binom{n}{k+1}
\end{aligned}$$

2.
$$\begin{aligned}
\binom{n}{k} &= \frac{n!}{(n-k)!k!} \\
&= \frac{n!}{k!(n-k)!} \\
&= \frac{n!}{(n-(k-n))!(n-k)!} \\
&= \binom{n}{n-k}
\end{aligned}$$

□

Bemerkung:

1. Entspricht der Aussage, dass ein Glied im Pascal'schen Dreieck sich aus der Summe der zwei überliegenden Glieder ergibt.
2. Entspricht der Aussage, dass das Pascal'sche Dreieck symmetrisch ist.

1.2.2 Kombinatorik

Theorem

Kombinatorik bezeichnet die Anzahl der Möglichkeiten der Anordnung von k Elementen auf n Stellen.
Es werden 2 Fälle unterschieden:

- Die Reihenfolge der Elemente wird berücksichtigt:
$$n(n-1)(n-2)...(n-(k-1))$$
- Die Reihenfolge der Elemente wird **nicht** berücksichtigt:
$$\binom{n}{k}$$

Bemerkung:
Auch hier gilt die Einschränkung $0 \leq k \leq n$. Diese ist sinnvoll, denn es ist nicht möglich, eine k-elementige Teilmenge einer n-elementigen Menge zu nehmen, wenn $k > n$. Deshalb gilt:

$$\text{Anzahl Möglichkeiten } = 0 \qquad \text{für } n \leq k$$

Beweis

Es handelt sich hier eher um eine logische Begründung:

- Reihenfolge berücksichtigt:
1. Auswahl: n Möglichkeiten
2. Auswahl: $n-1$ Möglichkeiten
..
k-te Auswahl: $n-(k-1)$ Möglichkeiten
$$\Rightarrow \quad \text{Möglichkeiten insgesamt: } n(n-1)(n-2)...(n-(k-1))$$
- Reihenfolge **nicht** berücksichtigt:
1. Auswahl: n Möglichkeiten, 1 mögliche Permutation
1. Auswahl: $n-1$ Möglichkeiten, 2 mögliche Permutationen
..
k-te Auswahl: $n-(k-1)$ Möglichkeiten, k mögliche Permutationen

$$\begin{aligned}\Rightarrow \quad \text{Möglichkeiten insgesamt: } & \frac{n(n-1)(n-2)...(n-(k-1)) \quad \} \text{ 1 neue Möglichkeit pro unterschiedlichem Fall}}{1*2*...*(k-1)(k-2)k \quad \} \text{ Anzahl der Wege, die zur selben Teilmenge führen}} \\ &= \frac{n(n-1)(n-2)...(n-(k-1))}{k!} \\ &= \frac{n!}{(n-k)!k!} \\ &= \binom{n}{k}\end{aligned}$$

□

Beispiel:

Nehmen wir das Ereignis E: bei einer Lotto-Ziehung “6 aus 49” sind genau 4 Zahlen richtig.

$$P(E) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}}$$

Zunächst nimmt man die Anzahl an Möglichkeiten, 4 richtige Kugeln von 6 zu ziehen, man multipliziert diese durch die Anzahl an Möglichkeiten, 2 Kugeln aus den 43 “unerwünschten” zu ziehen. Um die Wahrscheinlichkeit zu erhalten teilt man durch die Gesamtanzahl an Möglichkeite, 6 Kugeln aus 49 zu ziehen.

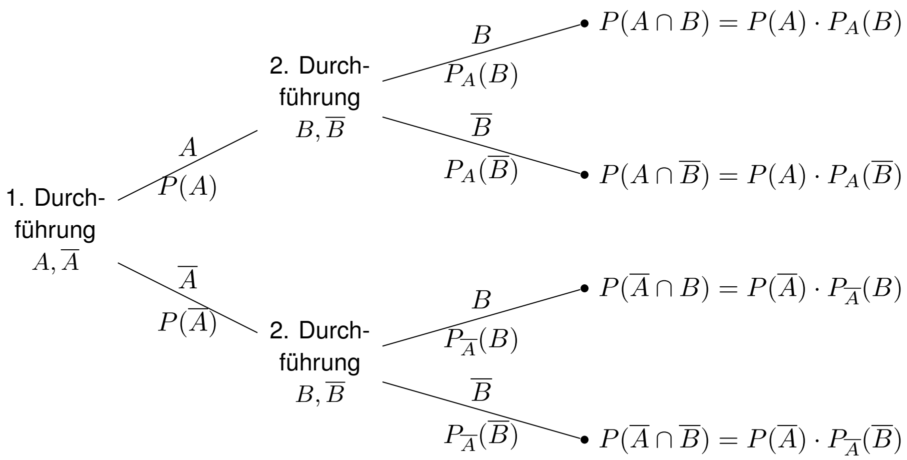
1.3 Bedingte Wahrscheinlichkeiten

Definition

Sind A und B beliebige Ereignisse mit $P(A) \neq 0$, so bezeichnet man $P_A(B)$ oder $P(B|A)$ die **durch** A **bedingte Wahrscheinlichkeit von** B. Es gilt:

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

Daraus ergibt sich die korrekte Darstellung als Baumdiagramm:



Theorem - Satz von Bayes

$$P_B(A) = \frac{P_A(B) \cdot P(A)}{P(B)}$$

Bemerkung:

Folgende Aussagen sind zu dieser Formulierung äquivalent:

- $P_A(B) \cdot P(B) = P_B(A) \cdot P(A)$
- $P(A) = P_A(B) \cdot P(B) + P_A(\overline{B}) \cdot P(\overline{B})$ (totale Wahrscheinlichkeit von A)

Beweis

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{P(A \cap B)}{P(A)} \cdot P(A)}{P(B)} = \frac{P_A(B) \cdot P(A)}{P(B)}$$

□

1.3.1 Stochastische Unabhängigkeit

Definition

Die Ergebnisse A und B werden stochastisch unabhängig genannt, wenn das Eintreten As die Wahrscheinlichkeit Bs nicht verändert. Es gilt dann:

$$P_A(B) = P(B)$$

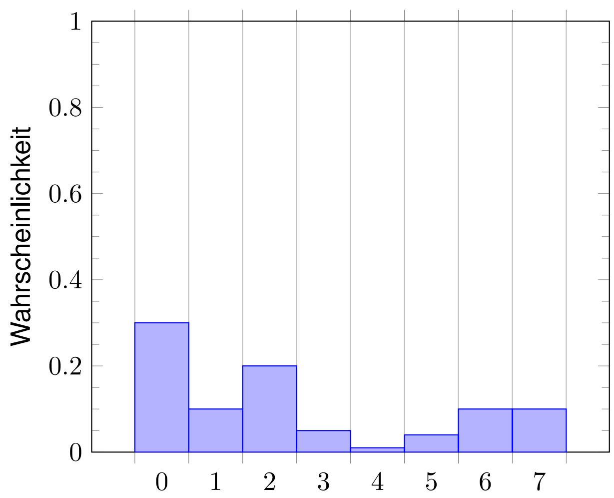
1.4 Wahrscheinlichkeitsverteilung

Definition

Man nennt $P(e_i)$ **Wahrscheinlichkeit** und die Zuordnung $e_i \longmapsto P(e_i)$ **Wahrscheinlichkeitsverteilung**, wenn gilt:

- $e_i \in S = \{e_1, e_2, ..., e_n\} \wedge P(e_i) \in \mathbb{R}$
- $!\exists P(e_i) \forall e_i \in S$
- $0 \leq P(e_i) \leq 1 \forall i \leq n$
- $P(e_1) + P(e_2) + ... + P(e_n) = 1$

Darstellung:



Bemerkung:
Dieses Balkendiagramm wird als Histogramm bezeichnet und ist in vielen Fällen eine gute Darstellungsmöglichkeit.

GTR-Tipp:

1. Mit `seq()` (in `LIST > OPS`) wird die Liste mit allen X-Werten generiert:
   `seq(X,X,0,<Anzahl an Versuchen>,1)→<Variable>`.
   (→ wird durch die Taste `STO>` aufgerufen)
2. Mit `<gewünschte Verteilung>()` (in `DISTR > DISTR`) wird die Liste mit allen Y-Werten generiert:
   `<Verteilung>(<Anzahl an Versuchen>,<Erfolgswahrscheinlichkeit>)→<Variable>`.
3. Im Menü `STAT PLOT > Plot 1` kann der gewünschte Anzeigemodus gewählt werden, für die `Xlist` wird die erste Liste gewählt, für `Freq` die zweite. Beim Anzeige achte man auf die richtigen Maßstäbe: Die Anzahl an Versuchen für `Xmax`, 1 für `Ymax`.

Bemerkung:

Der GTR ist ab einer Anzahl von über 48 überfordert, dann können die Balken wegen der niedrigen Auflösung nicht alle angezeigt werden. In einem solchen Fall kann nur ein Ausschnitt des Histogramms angezeigt werden, sonst erscheint eine Fehlermeldung.

Bemerkung:

Alternativ kann eine Funktion definiert werden über
`<Verteilung>(<Anzahl an Versuchen>,<Erfolgswahrscheinlichkeit>,round(X,0))` und anschließend angezeigt werden.
Dies funktioniert nur für ganze x, weshalb X durch `round(X,0)` (zu finden in... (KP, catalog durchsuchen)) auf den nächsten ganzen Wert gerundet wird.

Definition - Erwartungswert

Es sei eine Zuffalsvariable X mit Werten $x_1, x_2, ..., x_n$ und die zugehörige Wahrscheinlichkeitsverteilung $P(X)$. Als Erwartungswert μ wird das arithmetische Mittel der Werte von $P(X)$ bezeichnet:

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(X = x_i)$$

Definition - Median einer Wahrscheinlichkeitsverteilung

Als Median einer Wahrscheinlichkeitsverteilung bezeichnenn wir den den Wert X_n, ab dem die kumulierte Wahrscheinlichkeit von $P(X = x_i)$ den Wert von 50% überschreitet.

Bemerkung:

Anschaulich bezeichnet der Erwartungswert den durchschnittlichen Wert, den eine Zufallsgröße annimt, wenn ein Zufallsexperiment unendlich oft durchgeführt wird.

1.4.1 Binomialverteilung

Theorem

Für eine Bernoulli-Kette der Länge $l \in \mathbb{N}$ und der Trefferwahrscheinlichkeit P (Wahrscheinlichkeit für einen Pfad) gilt:

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

mit k der gerwünschten Anzahl an Erfolgen.
Außerdem gilt:

$$P(X \leq k) = \sum_{i=0}^{k} \binom{l}{i} p^i (1-p)^{l-i}$$

Bemerkung:

Definition

X heißt in diesem Fall **binomialverteilte Zufallsvariable**. Die entsprechende Wahrscheinlichkeitsverteilung heißt Binomialverteilung. Es gilt $P(X = k) = B_{l,p}(k)$.
Die kumulierte Binomialverteilung wird durch ist gegeben durch: $P(X \leq k) = F_{l,p}(k)$.
l ist die Länge der Bernoulli-Kette, p die Erfolgswahrscheinlichkeit und k ist die gewünschte Anzahl an Erfolgen.

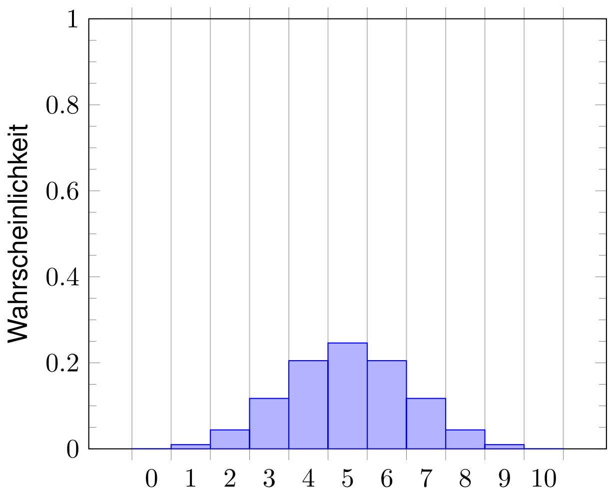
GTR-Tipp:

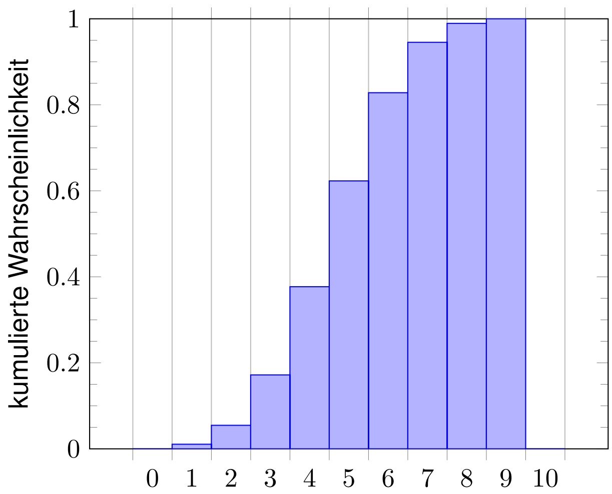
Beide Berechnungen werden durch einen GTR-Befehl automatisiert:

- $\binom{l}{k} p^k (1-p)^{l-k}$ wird duch den Befehl `binomPDF(l,p,k)` berechnet.

> - $\sum_{i=0}^{k} \binom{l}{i} p^i (1-p)^{l-i}$ wird duch den Befehl `binomCDF(l,p,k)` berechnet.
>
> Beide Befehle befinden sich im `DISTR`-Menü (`2ND+VARS`)

Darstellung:





Erwartungswert

Theorem - Erwartungswert einer Binomialverteilung

Für eine binomialverteilte Zufallsvariable X gilt:

$$E(x) = n \cdot p$$

mit n der Anzahl an Versuchen, und p der Erfolgswahrscheinlichkeit eines Versuchs.

Beweis

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

> Bemerkung:
>
> Der Startindex der Summe kann verändert werden, da die kumulierte Wahrscheinlichkeit sich nicht verändert. Für $k = 0$ ist $\binom{n-1}{k-1} = 0$.

$\square$

1.5 Kontinuierliche Wahrscheinlichkeiten

Manche Ereignisse lassen sich nicht als ganze Zahl modellieren oder können nicht einzeln abgezählt werden, weshalb man ihre Wahrscheinlichkeitsverteilung nur durch kontinuierliche Funktionen darstellen kann. Man kann dies als Erweiterung der diskreten Wahrscheinlichkeit auf die reellen Zahlen sehen, analog zum Übergang von den Folgen zu den Funktionen.

Definition - Dichtefunktion

Eine auf $I = [a; b]$ stetige Funktion f heißt Dichtefunktion einer Wahrscheinlichkeit, wenn gilt:

$$\int_a^b f(x)dx = 1$$

Definition

Die Wahrscheinlichkeitsverteilung P, die jedem Intervall $[c; d] \subset I$ den Wert $P([c; d]) = \int_c^d f(x)dx$ zuordnet, heißt Wahrscheinlichkeitsverteilung über I.

Bemerkung:
Durch die Definition des Integrals lassen sich folgende Eigenschaften direkt ableiten:

- $P(I) = 1$
- $[c; d] \subset I \Rightarrow P([c; d]) \in [0; 1]$
- mit $I = [a; b]$ gilt: $P([a; c]) = 1 - P([c; b])$
- Die Wahrscheinlichkeit eines "Singleton", einer ein-elementigen Menge, also eines einzelnen Ergebnisses ist: $P(c) = P([c; c]) = \int_c^c f(x)dx = 0$.
- Das bedeutet, dass $P(A \cap B)$ null sein kann, ohnen dass A und B disjunkt sind. (Zum Beispiel, wenn der Schnitt genau ein Element enthält.)
- Es folgt außerdem, dass $P([c; d]) = P(]c; d]) = P([c; d[) = P(]c; d[) = \int_c^d f(t)dt$

Definition - Stetige Zufallsvariable

Sei P eine Wahrscheinlichkeitsverteilung über I mit Dichtefunktion f.

- Eine Zufallsvariable X mit Werten Auswahl I folgt der Wahrscheinlichkeitsverteilung P, wenn gilt:

$$P(c \leq X \leq d) = \int_c^d f(t)dt$$

- Die Verteilungsfunktion F ist definiert durch $F(X) = P(X \leq x) = P(a \leq x) = \int_a^x f(t)dt$

Bemerkung:
F besitzt per Definition folgende Eigenschaften:

- F ist differenzierbar und $F' = f$
- F ist monoton steigend auf I
- Für $h > 0$ und $x + h \in I$ gilt: $P(x + h) = F(x + h) - F(x)$.
  (Gleichbedeutend mit der Aussage $P(c \leq X \leq d) = F(d) - F(c)$)

Definition - Erwartungswert

Sei X eine Zufallsvariable mit Werten aus $I = [a; b]$. Der Erwartungswert ist gegeben durch:

$$E(X) = \mu = \int_a^b t \cdot f(t)dt$$

Definition - Varianz

Sei X eine Zufallsvariable mit Werten aus $I = [a; b]$. Die Varianz ist gegeben durch:

$$\sigma^2 = \int_a^b (t - \mu)^2 \cdot f(t)dt$$

1.5.1 Exponentialverteilung: $X \sim \mathcal{E}(\lambda)$

Definition

Eine stetige Zufallsvariable X hat eine Exponentialverteilung wenn X die folgende Wahrscheinlichkeitsdichte hat:

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

$f_\lambda(x)$ heißt Dichtefunktion der Exponentialverteilung.

Bemerkung:
Diese Funktion erfüllt die Bedingungen der Wahrscheinlichkeitsverteilung, denn sie ist auf $\mathbb{R}_+$ definiert und stetig. Außerdem gilt: $\displaystyle\int_0^\infty f_\lambda dt = 1$.

Definition

Die Verteilungsfunktion der Exponentialverteilung ist gegeben durch:

$$P(X \leq x) = F_\lambda(x) = \int_0^x f_\lambda(t)dt = 1 - e^{-\lambda x}$$

Bemerkung:
Solche Funktionen werden verwendet, um Halbwärtszeiten oder Lebensdauern zu modellieren.

Theorem - Gedächtnislosigkeit

Für eine exponentialverteilte Zufallsvariable X gilt folgende Eigenschaft:

$$P_{(X \geq x)}(X \geq x + t) = P(X \geq t) \qquad \forall t, x \geq 0$$

Bemerkung:
Das bedeutet beispielsweise für die Zersetzung eines radioaktiven Atoms, dass die Wahrscheinlichkeit sich bei $x + t$ zu zersetzen unter der Voraussetzung, dass diese noch nicht bei x stattgefunden hat, nicht von x abhängt.
Solche Lebensdauern werden als “alterungslos” bezeichnet.

Bemerkung:
Die Gedächtnislosigkeit ist definierte Eigenschaft der Exponentialverteilung. (Es handelt sich somit um eine notwendige und hinreichende Bedingung)

Beweis

$$\begin{aligned}
P_{(X\geq x)}(X \geq x+t) &= \frac{P((X \geq x) \cap (X \geq x+t))}{P(X \geq x)} \\
&= \frac{P(X \geq x+t)}{P(X \geq x)} \\
&= \frac{1-(1-e^{-\lambda(x+t)})}{1-(1-e^{-\lambda x})} \\
&= \frac{e^{-\lambda(x+t)}}{e^{-\lambda x}} \\
&= e^{-\lambda t} \\
&= 1-(1-e^{-\lambda(t)}) \\
&= P(X \geq t)
\end{aligned}$$

□

Theorem - Erwartungswert

Für die Exponentialverteilung gilt:

$$\mu = \frac{1}{\lambda}$$

Beweis

$$\begin{aligned}
\mu &= \int_0^\infty t \cdot f_\lambda(t) dt \\
&= \int_0^\infty t \cdot \lambda e^{-\lambda t} dt \\
&= \left[-t \cdot e^{-\lambda t}\right]_0^\infty - \int_0^\infty -1 \cdot e^{-\lambda t} dt \\
&= \left[-t \cdot \lambda e^{-\lambda t}\right]_0^\infty - \left[\frac{1}{\lambda} e^{-\lambda t}\right]_0^\infty \\
&= 0 - 0 - \Big(0 - \frac{1}{\lambda}) \\
&= \frac{1}{\lambda}
\end{aligned}$$

□

Theorem - Varianz

Für die Exponentialverteilung gilt:

$$\sigma^2 = \frac{1}{\lambda^2}$$

Beweis

$$\begin{aligned}
\sigma^2 &= \int_0^\infty (t-\mu)^2 \cdot f_\lambda(t)dt \\
&= \int_0^\infty (t-\mu)^2 \cdot \lambda e^{-\lambda t}dt \\
&= \int_0^\infty t^2 \cdot \lambda e^{-\lambda t}dt - \int_0^\infty 2\mu t \cdot \lambda e^{-\lambda t}dt + \int_0^\infty \mu^2 \cdot \lambda e^{-\lambda t}dt \\
&= \left[-t^2 \cdot e^{-\lambda t}\right]_0^\infty - \int_0^\infty -2t \cdot e^{-\lambda t}dt - \left(\left[-2\mu t \cdot e^{-\lambda t}\right]_0^\infty - \int_0^\infty -2\mu \cdot e^{-\lambda t}dt\right) + \left[-\mu^2 \cdot e^{-\lambda t}\right]_0^\infty \\
&= \left[-t^2 \cdot e^{-\lambda t}\right]_0^\infty - \left(\left[\frac{2t}{\lambda}e^{-\lambda t}\right]_0^\infty - \int_0^\infty \frac{1}{\lambda} \cdot e^{-\lambda t}dt\right) - \left(\left[-2\mu t \cdot e^{-\lambda t}\right]_0^\infty - \left[\frac{2\mu}{\lambda} \cdot e^{-\lambda t}\right]_0^\infty\right) + \left[-\mu^2 \cdot e^{-\lambda t}\right]_0^\infty \\
&= \left[-t^2 \cdot e^{-\lambda t}\right]_0^\infty - \left[\frac{2t}{\lambda}e^{-\lambda t}\right]_0^\infty + \left[-\frac{2}{\lambda^2}e^{-\lambda t}\right]_0^\infty - \left[-2\mu t \cdot e^{-\lambda t}\right]_0^\infty + \left[\frac{2\mu}{\lambda} \cdot e^{-\lambda t}\right]_0^\infty + \left[-\mu^2 \cdot e^{-\lambda t}\right]_0^\infty \\
&= 0 - 0 - (0-0) + 0 + \frac{2}{\lambda^2} - (0-0) + 0 - \frac{2\mu}{\lambda} + 0 + \mu^2 \qquad \Big| \mu = \frac{1}{\lambda} \\
&= \frac{2}{\lambda^2} - \frac{2}{\lambda^2} + \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2}
\end{aligned}$$

□

Bemerkung:

$$\Rightarrow \sigma = \sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda} = \mu$$

1.5.2 Normalverteilung: $X \sim \mathcal{N}(\mu, \sigma^2)$

Definition

Eine stetige Zufallsvariable X hat eine (Gauß- oder) Normalverteilung mit Erwartungswert μ und Varianz σ^2 wenn X die folgende Wahrscheinlichkeitsdichte hat:

$$\varphi_{\mu,\sigma^2}(x) = f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\varphi(x)$ ist die Dichtefunktion der Normalverteilung.

Definition

Die Verteilungsfunktion der Normalverteilung ist gegeben durch:

$$P(X \leq x) = \Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^x \varphi_{\mu,\sigma^2}(t)dt$$

Durch Substitution erhält man:

$$\Phi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2}dt$$

Bemerkung:

Diese Definition ist somit analog zur Definition der Verteilungsfunktion der Binomialverteilung (Integral als Summe der Flächen).
Außerdem bedeutet das, dass $\Phi(x) \xrightarrow{x\to\infty} 1$.

GTR-Tipp:

Zur Berechnung geht man, wie üblich, ins Menü `DISTR` und wählt für φ `normalpdf`, und für Φ `normalcdf` aus.

Theorem

Für eine normalverteilte Zufallsvariable X mit den Parametern n, p und der Standardabweichung $\sigma = \sqrt{\text{Var}(X)} = \sqrt{\sigma^2}$ erhält man folgende Näherungen:

- $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 68,3\%$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 95,4\%$
- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 99,7\%$

Beweis

Die Wahrscheinlichkeit für ein bestimmtes Streuintervall $I = [\mu - z\sigma; \mu + z\sigma]$ kann berechnet werden als:

$$\begin{aligned} P(I) &= 2\Phi(z) - 1 \\ &= 2\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt - 1 \end{aligned}$$

□

Bemerkung:
Andersherum kann die Größe von I berechnet werden über:

$$z = \Phi^{-1}\left(\frac{P+1}{2}\right)$$

GTR-Tipp:
Diese Berechnung kann auch vom GTR getätigt werden: Im Menü DISTR wählt man `invNorm()` aus, und gibt in der Eingabe die Fläche (also die Wahrscheinlichkeit), μ und σ an. Man erhält somit die Breite des Intervalls, und kann dies wieder in Abhängigkeit von σ angeben.

Rückführung: Binomialverteilung

Warum interessiert uns das? Weil die Binomialverteilung diese symmetrische Kurve für $n \to \infty$ Versuche annähert. Noch nicht überzeugt? Über die Varianz können wir interessante Aussagen über die Verteilung der Wahrscheinlichkeiten machen.

Theorem - Laplace-Bedingung

Für eine Binomialverteilung mit $\sigma = \sqrt{np(1-p)}$ und $\mu = np$ gilt:
Wenn $\sigma > 3$, dann kann diese Binomialverteilung durch die Normalverteilung angenähert werden.

$$B_{l,p}(k) = P(X = k) = \binom{l}{k} p^k (1-p)^{l-k} \approx \varphi_{\mu,\sigma^2}(k) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2}\left(\frac{k-\mu}{\sigma}\right)^2} ; k \in \mathbb{Z}$$

Daraus folgt, dass die Verteilungsfunktion ebenfalls angenähert werden kann:

$$P(a \leq X \leq b) \approx \int_{a-0,5}^{b+a,5} \varphi_{\mu,\sigma^2}(x)dx = \Phi_{\mu,\sigma^2}(b + 0,5) - \Phi_{\mu,\sigma^2}(a - 0,5)$$

Beweis

Die Abweichung wird so gering, dass sie vernachlässigbar wird. □

Theorem

Für eine binomialverteilte Zufallsvariable X mit den Parametern n, p und der Varianz $\sigma = \sqrt{np(1-p)}$ erhält man folgende Näherungen:

- $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 68,3\%$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 95,4\%$

- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 99,7\%$

1.6 Hypothesen

1.6.1 Signifikanztests

Einseitiger Test

Zweiseitiger Test

1.6.2 Hypothesentests

Fehler 1. Art

Fehler 2. Art