Stochastik Formelsammlung

Prof. Dr. Barbara Staehle

WS 2019/2020

Teil I
Beschreibende Statistik

1 Charakterisierung einer Stichprobe

1.1 Häufigkeitsverteilung einer Stichprobe

1. Seien n **Messwerte** in einer **Urliste** (unsortierte Stichprobe) gegeben: $x_1, x_2, \ldots, x_2$.
2. Ermittle die **verschiedenen auftretenden** Werte als $a_1, a_2, \ldots, a_k$ (eliminiere mehrfach vorkommende).
3. Ermittle für jeden Wert a_i dessen **absolute Häufigkeit** h_i durch Zählen dessen Vorkommens in der Urliste.
4. Ermittle die **relative Häufigkeit** als $f_i = \frac{h_i}{n}$.
5. Plausibilitätscheck:
$$\sum_{i=1}^{k} h_i = n \quad \text{und} \quad \sum_{i=1}^{k} f_i = 1.$$

Darstellung: Balken- oder Stabdiagramm

- auch: **Histogramm**
- x-Achse: a_i (Werte der ZV)
- y-Achte: h_i/f_i (absolute oder relative Häufigkeiten) als Säule / Stab / Balken

Darstellung: empirische Verteilungsfunktion

- berechne
$$\bar{F}(x) = \sum_{i:a_i \le x} f_i$$
- $\bar{F}$ ist Stufenfunktion mit min>0, max=1, Sprünge bei a_i

1.2 Kennwerte einer Stichprobe

Definition 1. Sei $x_1, \ldots, x_n$ eine Stichprobe mit den verschiedenen Werte $a_1, \ldots, a_k$ und den absoluten Häufigkeiten $h_1, \ldots, h_k$ bzw. relativen Häufigkeiten $f_1, \ldots, f_k$. Das **(arithmetische) Mittel** (oder **Mittelwert**) der Stichprobe ist
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{i=1}^{k} h_i a_i = \sum_{i=1}^{k} f_i a_i.$$

Definition 2. Der **Median** (auch **Zentralwert**) einer geordneten Stichprobe $x_1, \ldots, x_n$ ist
$$\tilde{x} = \begin{cases} x_{m+1} & \text{falls } n = 2m+1, \\ \frac{1}{2}(x_m + x_{m+1}) & \text{falls } n = 2m. \end{cases}$$

Der Modalwert

Der **Modalwert** gibt den am häufigsten auftretenden Stichprobenwert an. Vorteil: auch für nicht-numerische Stichproben verwendbar.

Bemerkung: Kommen mehrere Werte am häufigsten vor, heißt die Stichprobe **multimodal** (**bimodal** , falls es zwei Modi gibt)

Definition 3. Für die **geordneten** Stichprobenwerte $x_1, \ldots, x_n$ und $0 < p < 1$ heißt

$$\tilde{x}_p = \begin{cases} x_{\lceil np \rceil} & \text{falls } np \notin \mathbb{N}, \\ \frac{1}{2}(x_{np} + x_{np+1}) & \text{falls } np \in \mathbb{N} \end{cases}$$

das **p-Quantil** .

- Das 0.5-Quantil ist genau der Median: $\tilde{x}_{0.5} = \tilde{x}$.
- $\tilde{x}_{0.25}, \tilde{x}, \tilde{x}_{0.75}$ werden als **Quartile** bezeichnet.
- $\tilde{x}_{0.1}, \tilde{x}_{0.2}, \ldots, \tilde{x}_{0.9}$ heißen **Dezile** .
- $\tilde{x}_{0.01}, \tilde{x}_{0.02}, \ldots, \tilde{x}_{0.99}$ heißen **Perzentile** .

Definition 4. Für die Stichprobe $x_1, \ldots, x_n$ gibt die **(Stichproben-)Varianz** oder **empirische Varianz** an, wie sehr die Stichprobenwerte x_i um ihren Mittelwert $\bar{x}$ streuen:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Als Maß für die Streuung wird auch die Wurzel der Varianz verwendet

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2},$$

die so genannte **(Stichproben-)Standardabweichung** oder **empirische Standardabweichung** .

Alternative Varianz-Berechnung: Für die k verschiedenen Werte der Stichprobe a_i mit Häufigkeiten h_i gilt

$$s^2 = \frac{(\sum_{i=1}^{n} x_i^2) - n \cdot \bar{x}^2}{n-1} = \frac{(\sum_{i=1}^{k} h_i a_i^2) - n \cdot \bar{x}^2}{n-1}.$$

Weitere Streuungsmaße

Spannweite $R = x_{max} - x_{min}$ (Differenz von größtem und kleinstem Stichprobenwert); Vorteil: einfach zu berechnen, Nachteil: starke Beeinflussung durch Ausreißer.

Interquartilabstand $I = \tilde{x}_{0.75} - \tilde{x}_{0.25}$ (Differenz zwischen 75% und 25% Quantil); Vorteil: resistent gegen Ausreißer, Nachteil: aufwändiger zu berechnen. Abkürzung: IQR.

2 Multivariate Statistik (v3 only)

2.1 Lineare Korrelation (v3 only)

Definition 5. Gegeben seien die Wertepaare $(x_1, y_1), \ldots, (x_n, y_n)$ wobei nicht alle x_i gleich sind bzw. nicht alle y_i gleich sind. Die Zahl

$$r_{x,y} = \frac{s_{x,y}}{s_x s_y}$$

heißt **(empirischer) Korrelationskoeffizient** . Dabei ist

$$s_{x,y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

die **(empirische) Kovarianz** , $\bar{x}, \bar{y}$ sind die arithmetischen Mittelwerte und

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

sind die **(empirischen) Standardabweichungen** der x_i bzw. der y_i-Werte.

Eigenschaften des Korrelationskoeffizienten

- $-1 \leq r_{x,y} \leq 1$
- Falls $r_{x,y} > 0$, sind die Stichproben **(linear) positiv korreliert**
- Falls $r_{x,y} < 0$, sind die Stichproben **(linear) negativ korreliert**
- Falls $r_{x,y} = 0$, sind die Stichproben **(linear) unkorreliert**
- Falls $|r_{x,y}| = 1$, besteht perfekte lineare Abhängigkeit.

2.2 Lineare Regression (v3 only)

Korrelationsanalyse Untersuchung des **Ausmaßes des (linearen) Zusammenhangs** zwischen zwei Merkmalen x und y; beide Merkmale sind **gleichrangig** , d.h., sowohl gemessene x-Werte als auch gemessene y-Werte können streuen.

Regressionsanalyse Untersuchung der **Art des (linearen) Zusammenhangs** zwischen zwei Merkmalen x und y; beide Merkmale sind **nicht gleichrangig** , man betrachtet y als abhängig von x. Man geht davon aus, dass x festgehalten und exakt messbar ist und nur die y-Werte streuen.

Lineare Regression

Lineare Regression ist das einfachstes Regressionsmodell: Ermittlung einer Regressionsgerade, die den mittleren quadratischen Fehler (die mittlere quadratische Abweichung der Messwerte von den Funktionswerten) minimiert (Gauß'sche Methode der kleinsten Quadrate).

Berechnung einer Regressionsgerade

Gegeben Wertepaare $(x_1, y_1), \ldots, (x_n, y_n)$

Gesucht Gerade $f(x) = kx + d$

Bedingung minimiere den quadratischen Fehler $\sum_{i=1}^{n}(y_i - f(x_i))^2$

Lösung Berechne k und d als

$$k = r_{x,y}\frac{s_y}{s_x} = \frac{s_{x,y}}{s_x^2} \quad \text{und} \quad d = \bar{y} - k\bar{x}$$

Legende $r_{x,y}$: empirischer Korrelationskoeffizient, s_x, s_y: Standardabweichungen, $\bar{x}, \bar{y}$: arithmetische Mittelwerte, $s_{x,y}$: empirische Kovarianz

Die Qualität der Regressionsgeraden

Bestimmtheitsmaß $R^2 = r_{x,y}^2$ (Quadrat des Korrelationskoeffizienten) sagt aus, welcher Anteil der Variation in der abhängigen Variablen durch die Regressionsgerade erklärt werden kann. Je größer R^2, desto besser beschreibt die Gerade den Zusammenhang zwischen x und y.

Winkel zwischen den Regressionsgeraden für x und y $f(x) = kx + d$ bzw. $g(y) = k'y + d'$; kleiner Winkel visualisiert hohe Korrelation zwischen x und y, großer Winkel eine kleine Korrelation; Schnittpunkt im Daten-Schwerpunkt.

Teil II
Diskrete Wahrscheinlichkeitstheorie

3 Wahrscheinlichkeitsrechnung

3.1 Zufallsexperimente, Ereignisse und Wahrscheinlichkeit

Definition 6. Ein **Zufallsexperiment** ist ein Vorgang, der

- beliebig oft unter gleichartigen Bedingungen wiederholt werden kann
- und dessen Ergebnis nicht mit Sicherheit vorhergesagt werden kann.

Die Menge aller möglichen (sich gegenseitig ausschließenden) Ergebnisse des Zufallsexperiments wird **Ergebnismenge, Ereignismenge** oder **Ergebnisraum** genannt und mit Ω bezeichnet.

Definition 7.
- Ein **Ereignis** A ist eine Teilmenge von Ω und heißt **eingetreten**, wenn das Ergebnis des Experiments ein Element von A ist.
- Die einelementigen Teilmengen von Ω enthalten genau die möglichen Ergebnisse des Experiments und heißen **Elementarereignisse**.
- Ω selbst heißt **das sichere Ereignis**, da es auf jeden Fall eintritt.
- Die leere Menge $\emptyset$ steht für ein Ereignis, das nie eintritt und heißt **das unmögliche Ereignis**.

Definition 8. Gegeben sind die Ereignisse $A, B \subseteq \Omega$

- Das Ereignis A **und** B entspricht dem Durchschnitt $A \cap B$.
- Das Ereignis A **oder** B entspricht der Vereinigung $A \cup B$.
- Das **Gegenereignis** von A ist jenes Ereignis, das eintritt, wenn A nicht eintritt. Schreibweise: $\bar{A}$. Es entspricht dem Komplement $\bar{A} = \Omega \setminus A$.
- A, B heißen **unvereinbar**, wenn $A \cap B = \emptyset$, sie nicht gleichzeitig eintreten können, andernfalls heißen A, B **vereinbar**.

Definition 9. Ein **Laplace-Experiment** ist ein Zufallsexperiment mit folgenden Eigenschaften:

- Das Zufallsexperiment hat nur **endlich viele** mögliche Ergebnisse.
- Jedes dieser Ergebnisse ist **gleich wahrscheinlich**.

Satz 10. *Bei einem Laplace-Experiment mit n möglichen Ergebnissen hat jedes dieser Ergebnisse die Wahrscheinlichkeit $\frac{1}{n}$. Wenn ein Ereignis A in k dieser Fälle eintritt, dann ist die Wahrscheinlichkeit von A gleich*

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} = \frac{k}{n}.$$

Definition 11 (Axiome von Kolmogorov)**.** Ein **Wahrscheinlichkeitsmaß** $P : \Omega \to [0, 1]$ muss folgendes erfüllen:

1. Die Wahrscheinlichkeit jedes Ereignisses ist eine reelle Zahl zwischen 0 und 1:
$$\forall_{A \subseteq \Omega}\ 0 \leq P(A) \leq 1$$

2. Das sichere Ereignis hat die Wahrscheinlichkeit 1:

$$P(\Omega) \ = 1$$

3. **Additionsregel** : Für abzählbar viele unvereinbare Ereignisse $A_i \subseteq \Omega$, $i \in I \subseteq \mathbb{N}$ ist die Wahrscheinlichkeit der Vereinigung dieser Ereignisse gleich der Summe der Einzel-Wahrscheinlichkeiten:

$$P(A_1 \cup A_2 \cup \ldots) = \sum_{i \in I} P(A_i) \text{ falls } \forall_{i,j \in I, i \neq j} \ A_i \cap A_j = \emptyset$$

Satz 12 (Eigenschaften von Wahrscheinlichkeiten)**.** *1. Die Wahrscheinlichkeit des Gegenereignisses von A ist*

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

2. Additionsregel für beliebige Ereignisse A und B:

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

3. Die Wahrscheinlichkeit ist monoton:

$$P(A) \leq P(B) \text{ für } A \subseteq B,$$

wenn A eine Teilmenge von B ist, ist $P(A)$ kleiner oder gleich $P(B)$.

3.2 Kombinatorik

Satz 13 (Summenregel)**.** *Für zwei endliche,* ***disjunkte*** *Mengen A und B (deren Elemente jeweils unterschiedliche Eigenschaften haben) ist die Anzahl der Elemente ihrer Vereinigungsmenge gleich*

$$|A \cup B| = |A| + |B|.$$

Verallgemeinerung für disjunkte Mengen $A_1, A_2, \ldots, A_k$:

$$|A_1 \cup A_2 \cup \ldots \cup A_k| = |A_1| + |A_2| + \ldots + |A_k|.$$

Satz 14 (Produktregel)**.** *Wenn A_1 und A_2 beliebige endliche Mengen sind, welche die Möglichkeiten für den ersten und zweiten Schritt eines Prozesses beschreiben, dann ist die Anzahl der Möglichkeiten den gesamten Prozess durchzuführen, beschrieben durch:*

$$|A_1 \times A_2| = |A_1| \cdot |A_2|.$$

Verallgemeinerung für k Schritte $A_1, A_2, \ldots, A_k$:

$$|A_1 \times A_2 \times \ldots \times A_k| = |A_1| \cdot |A_2| \cdot \ldots \cdot |A_k|.$$

3.3 Bedingte Wahrscheinlichkeit

Definition 15. Seien A, B Ereignisse über dem selben Ereignisraum Ω. Die Wahrscheinlichkeit des Ereignisses B **unter der Bedingung** , dass Ereignis A eingetreten ist, heißt **bedingte Wahrscheinlichkeit** $P(B|A)$ und ist definiert als

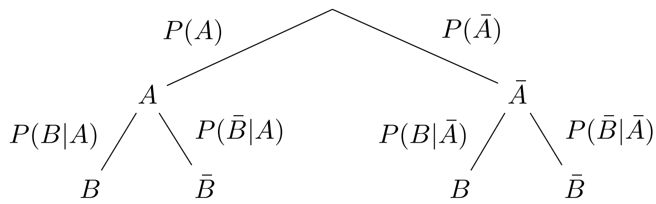
$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Satz 16 (Multiplikationssatz)**.** *Gegeben sind Ereignisse A und B über dem selben Ereignisraum Ω mit Wahrscheinlichkeiten ungleich null. Dann gilt:*

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

Satz 17 (von der totalen Wahrscheinlichkeit)**.** *Für eine Partition $E_1, \ldots, E_n$ von Ω und ein beliebiges Ereignis $A \subseteq \Omega$ gilt*

$$P(A) = \sum_{k=1}^{n} P(A \cap E_k) = \sum_{k=1}^{n} P(E_k)P(A|E_k)$$



Satz 18 (Formel von Bayes)**.** *Gegeben ist eine beliebige Partition $E_1, \ldots, E_n$ von Ω und ein beliebiges Ereignis $A \subseteq \Omega$. Die Wahrscheinlichkeit, dass eines der Ereignisse E_j unter der Bedingung von A eintritt, ist*

$$P(E_j|A) = \frac{P(E_j)P(A|Ej)}{P(A)} = \frac{P(E_j)P(A|Ej)}{\sum_{k=1}^{n} P(E_k)P(A|E_k)}$$

Bemerkung: Meistens ist $n = 2$ und $E_1 = E, E_2 = \bar{E}$, und damit

$$P(E|A) = \frac{P(E)P(A|E)}{P(E)P(A|E) + P(\bar{E})P(A|\bar{E})}.$$

Definition 19. Zwei Ereignisse A und B über dem selben Ereignisraum Ω heißen **unabhängig** , wenn eine (und damit alle drei) der folgenden Eigenschaften erfüllt ist:

- $P(B|A) = P(B)$ (falls $P(A) > 0$)
- $P(A|B) = P(A)$ (falls $P(B) > 0$)
- $P(A \cap B) = P(A) \cdot P(B)$ **Multiplikationssatz für unabhängige Ereignisse**

4 Zufallsvariablen

4.1 Diskrete Zufallsvariablen

Definition 20. Eine **Zufallsvariable** X ist eine Funktion, die zu einem Zufallsexperiment mit Ereignisraum Ω gehört und die jedem Elementarereignis dieses Zufallsexperimentes eine reelle Zahl zuordnet: $X : \Omega \to \mathbb{R}, \omega \mapsto X(\omega)$.

Eine **diskrete** Zufallsvariable X

- kann nur endlich oder abzählbar unendlich viele Werte $T = \{x_1, x_2, x_3, \ldots\}$ annehmen; T heißt **Träger** von X (oder Wertemenge von X).
- hat Realisierungen x_i; das Ereignis $X = x_i$ tritt mit der Wahrscheinlichkeit $p_i = P(X = x_i)$ eintritt. Die Realisierungen x_i gemeinsam mit den zugehörigen Wahrscheinlichkeiten p_i heißen **(Wahrscheinlichkeits-)Verteilung** der Zufallsvariablen.

Definition 21. Für eine (diskrete oder stetige) Zufallsvariable X heißt die Funktion

$$F(x) = P(X \leq x), \quad \text{für } x \in \mathbb{R},$$

Verteilungsfunktion (CDF) (E: cumulative distribution function) von X.

Eigenschaften der Verteilungsfunktion

Ist $F(x)$ die Verteilungsfunktion einer (diskreten oder stetigen) Zufallsvariablen X, so gilt:

- $F(x)$ **wächst monoton von 0 bis 1** :

$$\lim_{x\to-\infty} F(x) = 0, \quad F(x) \leq F(y) \text{ für } x < y, \quad \lim_{x\to\infty} F(x) = 1$$

- Wenn $F(x)$ an einer Stelle x_0 springt, so ist die Sprunghöhe genau $P(X = x_0)$.

4.2 Erwartungswert und Varianz

Definition 22. Für eine **diskrete** Zufallsvariable X ist der **Erwartungswert** , $\mathrm{E}(X)$ oder μ_X die mit den Wahrscheinlichkeiten gewichtete Summe über alle Realisierungen x_i (aus dem Träger T) von X:

$$\mathrm{E}(X) = \sum_{x_i \in T} x_i p_i$$

Eigenschaften des Erwartungswertes

Für zwei beliebige Zufallsvariablen X, Y, $a, b \in \mathbb{R}$, sowie eine Funktion $g : \mathbb{R} \to \mathbb{R}$ gilt:

- Der Erwartungswert ist linear, daher
  - $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$
  - $\mathrm{E}(aX) = a\mathrm{E}(X)$
  - $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$
- Auch $g(X)$ ist eine Zufallsvariable, daher

$$\mathrm{E}(g(X)) = \sum_{x_i \in T} g(x_i) p_i$$

  Achtung: Meist gilt $\mathrm{E}(g(X)) \neq g(\mathrm{E}(X))$.

- Sind X und Y **unabhängig** , dann gilt

$$\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$$

Definition 23. Für eine diskrete Zufallsvariable X mit Erwartungswert μ_X ist die **Varianz** , $\mathrm{Var}(X)$ oder σ_X^2 die mit den Wahrscheinlichkeiten gewichtete Summe über die quadratische Abweichung aller Realisierungen x_i (aus dem Träger T) von X vom Erwartungswert:

$$\mathrm{Var}(X) = \sum_{x_i \in T} (x_i - \mu)^2 p_i$$

Die **Standardabweichung** von X ist $\sigma_\mathbf{X} = \sqrt{\mathrm{Var}(X)}$.

Satz 24 (Vereinfachte Varianz-Berechnung)**.** *Für eine Zufallsvariable X mit Erwartungswert μ_X und Varianz σ_X^2 gilt*

$$\sigma_X^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \mathrm{E}(X^2) - \mu_X^2$$

Eigenschaften der Varianz

Für zwei beliebige Zufallsvariablen X, Y, $a, b \in \mathbb{R}$ gilt:

- $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$
- $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$
- Für $Y = aX + b$ gilt: $\sigma_Y = |a|\sigma_X$
- Nur wenn X und Y **unabhängig** sind, gilt $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

5 Wichtige diskrete Verteilungen

5.1 Verteilungen in Bernoulli-Ketten

Definition 25.
- Ein Zufallsexperiment bei dem ein Ereignis A mit Wahrscheinlichkeit p eintritt oder nicht heißt **Bernoulli-Experiment** .
- Falls Ereignis A eintritt heißt dies **Erfolg** , falls A nicht eintritt, ist dies ein **Misserfolg** .
- $p = P(A)$ heißt auch **Erfolgswahrscheinlichkeit** , $q = 1 - p$ die Wahrscheinlichkeit eines Misserfolgs.

Definition 26. Eine Zufallsvariable X, welche mit Wahrscheinlichkeit p den Wert 1 (Erfolg), mit Wahrscheinlichkeit $q = 1 - p$ den Wert 0 (Misserfolg) annimmt, heißt **Bernoulli-verteilt** : $X \sim Ber(p)$.

Definition 27. Wird ein Bernoulli-Experiment n-mal hintereinander unter denselben Bedingungen ausgeführt, und sind die Experimente unabhängig voneinander, so spricht man von einer **Bernoulli-Kette** der Länge n.

Definition 28. Gegeben sei eine Bernoulli-Kette mit Erfolgswahrscheinlichkeit p und Misserfolgswahrscheinlichkeit $q = 1 - p$. Die Anzahl X der Versuche, die bis zum 1. Erfolg unternommen werden müssen, ist eine **geometrisch verteilte** Zufallsvariable mit Parameter p: $X \sim geom(p)$. Es gilt

$$P(X = x) = p \cdot q^{x-1}$$

Definition 29. Gegeben sei eine Bernoulli-Kette der Länge n mit Erfolgs- wahrscheinlichkeit p und Misserfolgswahrscheinlichkeit $q = 1 - p$. Die Anzahl X der erfolgreichen Versuche, ist eine **binomialverteilte** Zufallsvariable mit Parametern n und p: $X \sim Bin(n, p)$. Es gilt

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$$

5.2 Die hypergeometrische Verteilung (v3 only)

Definition 30. Gegeben ist eine Grundgesamtheit aus N Elementen, von denen M eine bestimmte Eigenschaft haben Man entnimmt eine Stichprobe vom Umfang n (ohne Zurücklegen). Die Anzahl X der Elemente in der Stichprobe mit der bestimmten Eigenschaft ist eine **hypergeometrisch verteilte** Zufallsvariable mit den Parametern n, M, N: $X \sim H(n, M, N)$. Es gilt

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

5.3 Die Poisson-Verteilung

Definition 31. Eine Zufallsvariable X, die jede Zahl $x \in \mathbb{N}_0$ mit der Wahrscheinlichkeit

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad (\lambda > 0)$$

annehmen kann, heißt **poissonverteilt mit dem Parameter** λ . Kurzschreibweise: $X \sim Po(\lambda)$.

Spezialfall Poisson-Verteilung
Für unabhängige Zufallsvariablen $A \sim Po(\lambda_A)$ und $B \sim Po(\lambda_B)$ gilt

$$A + B \sim Po(\lambda_A + \lambda_B)$$

5.4 Näherungsweise (v3 only)

Faustregel
Wenn $n \gtrapprox 50$ und $p \lessapprox 0.1$ ist, dann kann eine Binomialverteilung mit den Parametern n und p durch die Poisson-Verteilung mit dem Parameter $\lambda = n \cdot p$ angenähert werden.

Faustregel
Wenn aus sehr vielen Elementen nur wenige ausgewählt werden (wenn der **Auswahlsatz** $\frac{n}{N} \lessapprox 0.005$ ist), dann kann eine hypergeometrische Verteilung mit den Parametern n, M, N durch die Binomialverteilung mit den Parametern n und $p = \frac{M}{N}$ angenähert werden.

Teil III
Kontinuierliche Wahrscheinlichkeitstheorie

6 Zufallsvariablen

6.1 Kontinuierliche Zufallsvariablen

Für eine **stetige** Zufallsvariable gilt:

- die Verteilungsfunktion (CDF) $F(x) = P(X \leq x)$ ist eine **stetige** und (mindestens stückweise) **differenzierbare** Funktion mit
  - $\lim\limits_{x \to -\infty} F(x) = 0$
  - $\lim\limits_{x \to \infty} F(x) = 1$
- $f(x) = F'(x)$ heißt **(Wahrscheinlichkeits-) Dichtefunktion (PDF),** (E: probability density function)
- die Verteilungsfunktion erhält man als Integral der Dichtefunktion:
$$F(x) = \int_{-\infty}^{x} f(t)\ \mathbf{d}t.$$

Für eine stetige Zufallsvariable X mit Dichte $f(x)$ und Verteilungsfunktion $F(x)$ gilt:

- f ist auf ganz $\mathbb{R}$ definiert und immer ≥ 0,
- f ist so normiert, dass die Gesamtfläche unter dem Graphen von f gleich 1 ist:
$$\int_{-\infty}^{\infty} f(t)\ \mathbf{d}t = 1,$$
- die Wahrscheinlichkeit, dass X einen Wert im Intervall $[a; b]$ (oder $(a; b]$ oder $[a; b)$ oder $(a; b)$) annimmt, ist

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = F(b) - F(a) = \int_{a}^{b} f(t)\ \mathbf{d}t.$$

6.2 Quantile, Erwartungswert und Varianz

Definition 32. Gegeben sei eine (diskrete oder stetige) Zufallsvariable X und $p \in (0; 1)$. $x_p \in \mathbb{R}$, für das
$$\mathbf{F(x_p) = p}$$
gilt, heißt **p-Quantil** von X. Ein Quantil zu $p = 0.5$ heißt **Median** .

Bemerkungen:

- $p\%$ aller Werte die X annimmt sind kleiner als das $p\%$-Quantil.
- Für stetige ZV existiert für jedes $p \in (0; 1)$ ein eindeutiges p-Quantil $x_p = F^{-1}(p)$.
- Für diskrete ZV muss ein Quantil nicht für jedes $p \in (0; 1)$ existieren (Sprungstellen)!

Definition 33. Für eine **stetige** Zufallsvariable X definiert man

- den **Erwartungswert**, $\mathrm{E}(X)$ oder μ_X als
$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, \mathrm{d}x$$
- die **Varianz**, $\mathrm{Var}(X)$ oder σ_X^2 als
$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu_X)^2 f(x)\, \mathrm{d}x$$
- die **Standardabweichung** von X als $\sigma_{\mathbf{X}} = \sqrt{\mathrm{Var}(X)}$

	diskrete ZV	**stetige ZV**
X	$X: \Omega \to \mathbb{R}, \quad \omega \mapsto X(\omega)$	
Ω	endlich oder abzählbar unendlich	überabzählbar unendlich
Verteilung	$p_i = P(X = x_i)$	-
Verteilungsdichte	-	$f(x) = F'(x)$ $P(X = x) = 0$
Verteilungsfunktion	$F(x) = P(X \leq x)$ $F(x) = \sum\limits_{i: x_i \leq x} p_i$	$F(x) = \int\limits_{-\infty}^{x} f(t)\mathrm{d}t$

Tabelle 1: Diskrete und stetige Zufallsvariablen: Allgemeines

	diskrete ZV	**stetige ZV**
X	$X: \Omega \to \mathbb{R}, \quad \omega \mapsto X(\omega)$	
p-Quantil $p \in (0;1)$	$x_p = F^{-1}(p)$, existiert für manche p-Werte	alle p-Werte
Erwartungswert	$\mathrm{E}(X) = \mu_X = \sum\limits_{x_i \in T} x_i p_i$	$\mathrm{E}(X) = \mu_X = \int\limits_{-\infty}^{\infty} x f(x)\, \mathrm{d}x$
Varianz	$\mathrm{Var}(X) =$ $\sum\limits_{x_i \in T} (x_i - \mu_X)^2 p_i$	$\mathrm{Var}(X) =$ $\int\limits_{-\infty}^{\infty} (x-\mu_X)^2 f(x)\, \mathrm{d}x$
Standardabweichung	$\sigma_X = \sqrt{\mathrm{Var}(X)}$	

Tabelle 2: Diskrete und stetige Zufallsvariablen: Kennwerte

	allgemein	**diskrete Zufallsvariable**	**stetige Zufallsvariable**
$P(a < X \leq b)$	$F(b) - F(a)$	$\sum\limits_{i: a < x_i \leq b} p_i$	$\int\limits_{a}^{b} f(t)\mathrm{d}t$
$P(X \leq b)$	$F(b)$	$\sum\limits_{i: x_i \leq b} p_i$	$\int\limits_{-\infty}^{b} f(t)\mathrm{d}t$
$P(X > a)$	$1 - F(a)$	$\sum\limits_{i: x_i > a} p_i$	$\int\limits_{a}^{\infty} f(t)\mathrm{d}t$

Tabelle 3: Diskrete und stetige Zufallsvariablen: Intervallwahrscheinlichkeiten

7 Wichtige stetige Verteilungen

7.1 Die Gleichverteilung

Definition 34. Eine Zufallsvariable X, die alle Werte des rellen Intervalls $[a;b]$ mit der gleichen Wahrscheinlichkeit annehmen kann, heißt **gleichverteilt** (manchmal auch **rechteckverteilt** :) $X \sim U(a,b)$. Es gilt:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \le x \le b \\ 0 & \text{sonst} \end{cases} \qquad F(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \le x \le b \\ 1 & \text{für } x > b \end{cases}$$

7.2 Die Exponentialverteilung

Definition 35. Eine Zufallsvariable X, welche nur Werte > 0 mit Erwartungswert $\frac{1}{\lambda}$ annehmen kann und die Zeit zwischen zwei Ereignissen beschreibt, heißt **exponentialverteilt:** $X \sim exp(\lambda)$, falls gilt:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \ge 0 \\ 0 & \text{sonst} \end{cases} \qquad F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{für } x \ge 0 \\ 0 & \text{sonst} \end{cases}$$

7.3 Die Normalverteilung

Definition 36. Eine Zufallsvariable X heißt **normalverteilt** mit den Parametern μ und σ, $\mathbf{X \sim N}(\mu, \sigma)$, wenn sie die Dichtefunktion

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

besitzt. Die zugehörige Wahrscheinlichkeitsverteilung heißt **Normalverteilung** oder auch **Gauß-Verteilung.** Der Graph der Dichtefunktion wird **Gauß'sche Glockenkurve** genannt. Die Parameter μ und σ sind Erwartungswert bzw. Standardabweichung von X.

Weitere Bemerkungen:

- μ beschreibt die Verschiebung der Glocke auf der x-Achse relativ zum Ursprung.
- Die Fläche unter der Glockenkurve ist immer gleich 1 (unabhängig von μ und σ). Daher:
  - Je größer σ, desto breiter und niedriger ist die Glocke,
  - Je kleiner σ, desto schmaler und höher ist die Glocke.

Achtung: Die Verteilungsfunktion der Normalverteilung

$$F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

kann nur numerisch berechnet werden, da f keine Stammfunktion besitzt.

Definition 37. Eine Zufallsvariable Z heißt **standardnormalverteilt** , wenn sie normalverteilt mit den Parametern $\mu = 0$ und $\sigma = 1$, $\mathbf{Z \sim N(0,1)}$, ist. Ihr Dichte- und Verteilungsfunktion sind gegeben durch

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

und

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt.$$

68–95–99.7-Regel / 3σ-Grenzen
Bei **jeder** normalverteilten Zufallsvariable X (mit beliebigem Mittelwert μ und Standardabweichung σ) ist die Wahrscheinlichkeit, dass X

- einen Wert zwischen $\mu - \sigma$ und $\mu + \sigma$ annimmt, etwa 68.3%,

- einen Wert zwischen $\mu - 2\sigma$ und $\mu + 2\sigma$ annimmt, etwa 95.5%,
- einen Wert zwischen $\mu - 3\sigma$ und $\mu + 3\sigma$ annimmt, etwa 99.7%.

Satz 38 (Additionssatz der Normalverteilung)**.** *Für X und Y unabhängig und normalverteilte ZV mit $X \sim N(\mu_X, \sigma_X)$ bzw. $Y \sim N(\mu_Y, \sigma_Y)$ gilt: $X + Y$ ist ebenfalls normalverteilt mit Erwartungswert $\mu_X + \mu_Y$ und Standardabweichung $\sqrt{\sigma_X^2 + \sigma_Y^2}$:*

$$X + Y \sim N(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}) :$$

7.4 Die Normalverteilung als Näherung (v3 only)

Original X_B sei eine binomialverteilte ZV mit $X_B \sim Bin(n,p)$, mit $\mu = np$ und $\sigma^2 = np(1-p)$ und Verteilung $F_B(x)$.

Näherung durch eine normalverteilte ZV X_N mit Parametern $\mu = np$ und $\sigma = \sqrt{np(1-p)}$ und Verteilung $F_N(x)$.

Es gilt

$$F_B(x) \approx F_N(x + 0.5)$$

Bedingung np und $n(1-p)$ müssen „groß genug“ sein, $n \cdot p \cdot (1-p) \gtrapprox 9$.

Original X_P sei eine poissonverteilte ZV mit $X_P \sim Po(\lambda)$, mit $\mu = \lambda$ und $\sigma^2 = \lambda$ und Verteilung $F_P(x)$.

Näherung durch eine normalverteilte ZV X_N mit Parametern $\mu = \lambda$ und $\sigma = \sqrt{\lambda}$ und Verteilung $F_N(x)$.

Es gilt

$$F_P(x) \approx F_N(x + 0.5)$$

Bedingung λ muss „groß genug“ sein, $\lambda \gtrapprox 9$.

Teil IV
Schließende Statistik

8 Grundbegriffe

Definition 39. Eine **(Zufalls)stichprobe** vom Umfang n ist eine Folge $X_1, \ldots, X_n$ von unabhängigen, **identisch verteilten** Zufallsvariablen. X_i ist die **Merkmalsausprägung** des i-ten Elements der Stichprobe. Die X_i heißen **Stichprobenvariablen** . Wird eine Stichprobe gezogen, dann nehmen die ZV $X_1, \ldots, X_n$ die **konkreten Werte** oder **Realisierung** $x_1, \ldots, x_n$ an.

9 Große und zentrale Gesetze

9.1 Das Gesetz der großen Zahlen

Satz 40 (Das (starke) Gesetz der großen Zahlen)**.** *Seien $X_1, X_2, \ldots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert μ und sei $\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$ ihr arithmetisches Mittel. Dann gilt für jede noch so kleine Zahl $\varepsilon > 0$:*

$$\lim_{n\to\infty} P(|\bar{X} - \mu| < \varepsilon) = 1$$

Satz 41 (Theorem von Bernoulli)**.** *Ein Zufallsexperiment, bei dem das Ereignis A mit der Wahrscheinlichkeit p eintritt, werde n-mal unabhängig wiederholt, f_n sei die relative Häufigkeit des Eintretens von A. Dann gilt für jede noch so kleine Zahl $\varepsilon > 0$*

$$\lim_{n\to\infty} P(|f_n - p| \le \varepsilon) = 1.$$

Satz 42 (Hauptsatz der Statistik)**.** *Seien $X_1, X_2, \ldots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Verteilungsfunktion $F(x)$ und empirischer Verteilungsfunktion $\bar{F}(x)$. Dann gilt für jede noch so kleine Zahl $\varepsilon > 0$ und jedes $x \in \mathbb{R}$:*

$$\lim_{n\to\infty} P(|\bar{F}(x) - F(x)| < \varepsilon) = 1$$

9.2 Der zentraler Grenzwertsatz

Satz 43. *Seien $X_1, \ldots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert jeweils μ und Standardabweichung σ. Dann hat das arithmetische Mittel $\bar{X} = \frac{X_1+X_2+\ldots+X_n}{n}$ den Erwartungswert μ und die Varianz $\frac{\sigma^2}{n}$ und ist (ungefähr)* ***approximativ normalverteilt*** *:*

$$\bar{X} \overset{a}{\sim} N(\mu, \sigma/\sqrt{n}), \qquad \lim_{n\to\infty} P(\bar{X} \le x) = \Phi\left(\frac{x-\mu}{\sigma/\sqrt{n}}\right).$$

Die zugehörige standardisierte Zufallsvariable Z ist ***approximativ standardnormalverteilt:***

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1), \qquad \lim_{n\to\infty} P(Z \le z) = \Phi(z).$$

Vokabular: $\sigma/\sqrt{n}$ wird auch als **Standardfehler** des arithmetischen Mittels bezeichnet. Verwendung: Konfidenzintervalle.

10 Schätzer (v3 only)

10.1 Punktschätzungen (v3 only)

Definition 44. Eine Funktion $T(X_1, \ldots, X_n)$ der Stichprobenvariablen, die zur Schätzung eines Parameters θ der Grundgesamtheit verwendet wird, heißt **Schätzfunktion** (oder **Schätzstatistik** oder **(Punkt)Schätzer**) für θ. $T(X_1, \ldots, X_n)$ ist ebenfalls eine Zufallsvariable. Eine konkrete Realisierung $T(x_1, \ldots x_n)$ heißt **Schätzwert** und wird mit $\hat{\theta}$ bezeichnet.

θ	Schätzfunktion	Schätzwert (für konkrete Realisierung)
μ	$\bar{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$	$\hat{\mu} = \bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$
σ^2	$S^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2$	$\hat{\sigma^2} = s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \bar{x})^2$
π_m	$\bar{P_m} = \frac{1}{n}\lvert\{X_i \mid X_i \text{ hat Merkmal } m\}\rvert$	$\hat{\pi_m} = \bar{p_m} = \frac{1}{n}\lvert\{x_i \mid x_i \text{ hat Merkmal } m\}\rvert$

Tabelle 4: Schätzung verschiedener Parameters θ der Grundgesamtheit

Definition 45. Eine Schätzfunktion T für eine Stichprobe der Größe n heißt

- **erwartungstreu** (oder **unverzerrt**, engl. **unbiased**), wenn ihr Erwartungswert gleich dem zu schätzenden Parameter ist:
$$E(T) = \theta$$
- **konsistent**, wenn sie stochastisch gegen θ konvergiert:
$$\forall_{\varepsilon>0} \quad \lim_{n\to\infty} P(|T-\theta| < \varepsilon) = 1$$
- **konsistent im quadratischen Mittel**, oder **effizient** wenn die erwartete quadratische Abweichung im Grenzwert verschwindet:
$$\lim_{n\to\infty} \mathrm{E}((T-\theta)^2) = 0$$

Eigenschaften bekannter Schätzer

erwartungstreu und konsistent im quadratischen Mittel (effizient) sind

$\bar{X}$ das arithmetische Mittel als Schätzer für den wahren Mittelwert μ

$\bar{P}_m$ die empirische Häufigkeit von Merkmal m als Schätzer für die wahre Wahrscheinlichkeit des Merkmals π_m

$\bar{F}$ die empirische Verteilungsfunktion als Schätzer für die wahre Verteilungsfunktion F

erwartungstreu und konsistent ist

S^2 die empirische Varianz als Schätzer für die tatsächliche Varianz σ^2

10.2 Intervallschätzungen (v3 only)

Vokabular: α und $1-\alpha$

Irrtumswahrscheinlichkeit Mit einer Wahrscheinlichkeit von α überdeckt das Konfidenzintervall den wahren Wert von θ **nicht**; typische Werte: 5%, 1%, ...

Vertrauenswahrscheinlichkeit Konfidenzniveau, Sicherheit: Mit einer einer Wahrscheinlichkeit von $1-\alpha$ überdeckt das Konfidenzintervall den wahren Wert von θ; typische Werte: 95%, 99%, ...

Definition 46. Ein Intervall
$$[g_u(X_1,\dots,X_n), g_o(X_1,\dots,X_n)]$$
dessen Grenzen g_u und g_o aus den Stichprobenwerten berechnet werden, und das mit einer vorgegebenen Wahrscheinlichkeit $1-\alpha$ den gesuchten Parameter θ der Grundgesamtheit überdeckt, d.h.
$$P(\theta \in [g_u(X_1,\dots,X_n), g_o(X_1,\dots,X_n)]) = 1-\alpha$$
heißt **Konfidenzintervall** (oder **Vertrauensintervall** oder **Vertrauensbereich**) zum **Niveau** $1-\alpha$. Man nennt $1-\alpha$ **Konfidenzniveau** (oder **Vertrauenswahrscheinlichkeit**, auch **Sicherheit**).

Konfidenzintervall für den Erwartungswert μ einer normalverteilten ZV X bei bekannter Standardabweichung σ

1. Wähle ein Konfidenzniveau $1-\alpha$ (z.B. 0.90, 0.95, 0.99).

2. Ziehe eine Stichprobe vom Umfang n und berechne $\bar{x}$.

3. Bestimme das Quantil $z_{1-\frac{\alpha}{2}}$ der Standardnormalverteilung.

4. Das Konfidenzintervall
$$\left[\bar{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$
überdeckt den gesuchten Erwartungswert μ mit der Wahrscheinlichkeit $1-\alpha$.

Konfidenzintervall für den Erwartungswert μ einer normalverteilten ZV X bei unbekannter Standardabweichung σ

1. Wähle ein Konfidenzniveau $1-\alpha$ (z.B. 0.90, 0.95, 0.99).

2. Ziehe eine Stichprobe vom Umfang n und berechne $\bar{x}$ sowie s^2.

3. Bestimme das Quantil $t_{n-1;1-\frac{\alpha}{2}}$ der t-Verteilung.

4. Das Konfidenzintervall
$$\left[\bar{x} - t_{n-1;1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \bar{x} + t_{n-1;1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$$
überdeckt den gesuchten Erwartungswert μ mit der Wahrscheinlichkeit $1-\alpha$.

Faustregel: Ab 30 Freiheitsgraden ($m \geq 30$) kann die t-Verteilung durch die Standardnormalverteilung approximiert werden.

Konfidenzintervall für den Erwartungswert μ einer beliebig verteilten ZV X bei großem Stichprobenumfang ($n \geq 30$)

1. Wähle ein Konfidenzniveau $1-\alpha$ (z.B. 0.90, 0.95, 0.99).

2. Ziehe eine Stichprobe vom Umfang $n \geq 30$ und berechne $\bar{x}$ sowie die empirische Standardabweichung $s2$ (falls σ unbekannt).

3. Bestimme das Quantil $z_{1-\frac{\alpha}{2}}$ der Standardnormalverteilung.

4. Das Konfidenzintervall
$$\left[\bar{x} - z_{1-\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}}\right] \quad \text{(falls } \sigma \text{ bekannt)}$$
$$\left[\bar{x} - z_{1-\frac{\alpha}{2}}\tfrac{s}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}}\tfrac{s}{\sqrt{n}}\right] \quad \text{(falls } \sigma \text{ unbekannt)}$$
überdeckt den gesuchten Erwartungswert μ **annähernd** mit der Wahrscheinlichkeit $1-\alpha$.

Teil V

Anhang

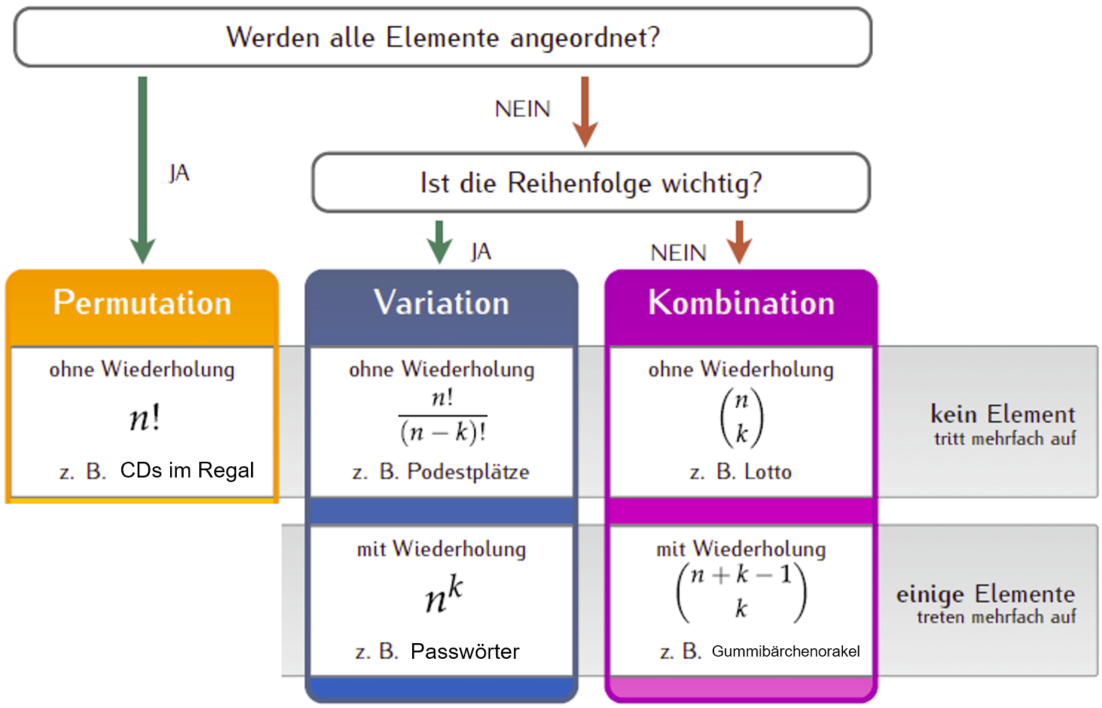


Bild 1: Übersicht über die verschiedenen Zählverfahren

Fakten zur Bernoulli-Verteilung

Für eine Bernoulli-verteilte Zufallsvariable X, $X \sim Ber(p)$ gilt

Wertebereich $X \in \{0, 1\}$

Parameter p: Erfolgswahrscheinlichkeit, $q = 1 - p$

Verteilung $P(X = x) = \begin{cases} q & \text{falls } x = 0, \\ p & \text{falls } x = 1, \\ 0 & \text{sonst.} \end{cases}$

Verteilungsfunktion $P(X \leq x) = F(x) = \begin{cases} 0 & \text{falls } x < 0, \\ q & \text{falls } 0 \leq x < 1, \\ 1 & \text{falls } x \geq 1. \end{cases}$

Erwartungswert $\mathrm{E}[X] = p$

Varianz $\mathrm{Var}[X] = pq$

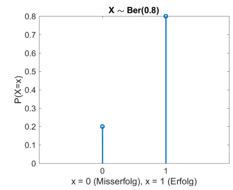


Bild: Verteilung

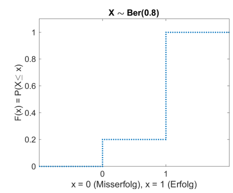


Bild: Verteilungsfunktion

Wahrscheinlichkeitsrechnung
Zufallsvariablen
Wichtige diskrete Verteilungen

Verteilungen in Bernoulli-Ketten
Die hypergeometrische Verteilung (v3 only)
Die Poisson-Verteilung
Näherungsweise (v3 only)

Fakten zur geometrischen Verteilung

Für eine geometrisch verteilte Zufallsvariable $X \sim geom(p)$ gilt

Wertebereich $X \in \mathbb{N}$

Parameter p: Erfolgswahrscheinlichkeit, $q = 1 - p$

Verteilung $P(X = x) = \begin{cases} pq^{x-1} & \text{falls } x \in \mathbb{N}, \\ 0 & \text{sonst.} \end{cases}$

Verteilungsfunktion $P(X \leq x) = F(x) = \begin{cases} 1 - q^{\lfloor x \rfloor} & \text{falls } x \geq 1, \\ 0 & \text{sonst.} \end{cases}$

Erwartungswert $\mathrm{E}[X] = \frac{1}{p}$

Varianz $\mathrm{Var}[X] = \frac{q}{p^2}$

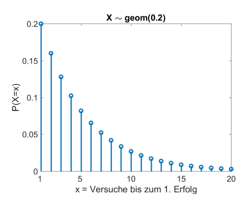


Bild: Verteilung

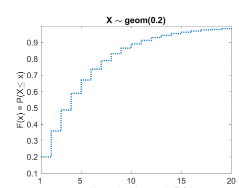


Bild: Verteilungsfunktion

Fakten zur Binomialverteilung

Für eine binomialverteilte Zufallsvariable $X \sim Bin(n, p)$ gilt

Wertebereich $X \in \mathbb{N}_0, X \leq n$

Parameter p: Erfolgswahrscheinlichkeit, n: Anzahl der durchgeführten Versuch

Verteilung $P(X = x) =$
$$\begin{cases} \binom{n}{x} p^x q^{n-x} & \text{falls } x \in \mathbb{N}_0, \\ 0 & \text{sonst.} \end{cases}$$

Verteilungsfunktion $P(X \leq x) = F(x) =$
$$\begin{cases} 0 & \text{falls } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k q^{n-k} & \text{falls } 0 \leq x \leq n, \\ 1 & \text{falls } x > n. \end{cases}$$

Erwartungswert $\mathrm{E}[X] = np$

Varianz $\mathrm{Var}[X] = npq$

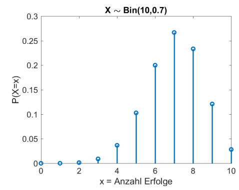


Bild: Verteilung

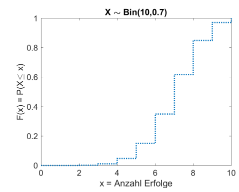


Bild: Verteilungsfunktion

Fakten zur Poisson-Verteilung

Für eine poissonverteilte Zufallsvariable $X \sim Po(\lambda)$ gilt

Wertebereich $X \in \mathbb{N}_0$

Parameter $\lambda > 0$: Auftrittsrate

Verteilung $P(X = x) =$
$$\begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{falls } x \in \mathbb{N}_0, \\ 0 & \text{sonst.} \end{cases}$$

Verteilungsfunktion $P(X \leq x) = F(x) =$
$$\begin{cases} 0 & \text{falls } x < 0 \\ e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!} & \text{falls } 0 \leq x. \end{cases}$$

Erwartungswert $\mathrm{E}[X] = \lambda$

Varianz $\mathrm{Var}[X] = \lambda$

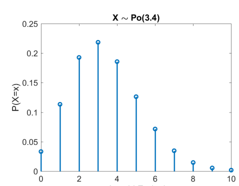


Bild: Verteilung

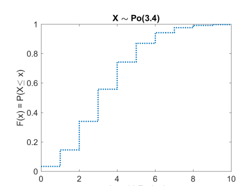


Bild: Verteilungsfunktion

Fakten zur hypergeometrischen Verteilung

Für eine hypergeometrisch verteilte Zufallsvariable $X \sim H(n, M, N)$ gilt

Wertebereich $X \in \mathbb{N}_0, X \leq M$

Parameter n: Größe der Stichprobe
M: Anzahl der Elemente mit der gewünschten Eigenschaft
N: Anzahl aller Elemente

Verteilung $P(X = x) =$

$$\begin{cases} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & \text{falls } x \in \mathbb{N}_0, \\ 0 & \text{sonst.} \end{cases}$$

Verteilungsfunktion $P(X \leq x) = F(x) =$

$$\begin{cases} 0 & \text{falls } x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} & \text{falls } 0 \leq x \leq M, \\ 1 & \text{falls } x > M. \end{cases}$$

Erwartungswert $\mathrm{E}[X] = n\frac{M}{N}$

Varianz $\mathrm{Var}[X] = n\frac{M}{N}(1 - \frac{M}{N})\frac{N-n}{N-1}$

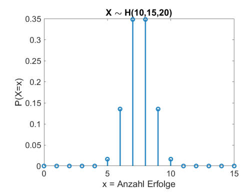


Bild: Verteilung

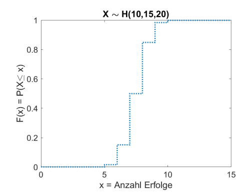


Bild: Verteilungsfunktion

Fakten zur Gleichverteilung

Für eine gleichverteilte Zufallsvariable X, $X \sim U(a, b)$ gilt

Wertebereich $X \in [a, b]$

Parameter $a, b \in \mathbb{R}$: minimaler und maximaler Wert, den X annehmen kann

Verteilungsdichte $f(x) =$

$$\begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion $F(x) = P(X \leq x) =$

$$\begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x \leq b \\ 1 & \text{für } x > b \end{cases}$$

Erwartungswert $\mathrm{E}[X] = \frac{a+b}{2}$

Varianz $\mathrm{Var}[X] = \frac{(b-a)^2}{12}$

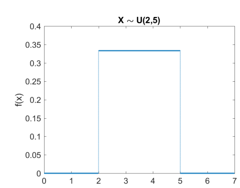


Bild: Verteilungsdichtefunktion

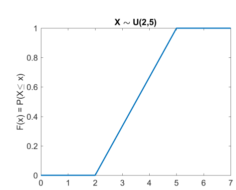


Bild: Verteilungsfunktion

Fakten zur Exponentialverteilung

Für eine exponentialverteilte Zufallsvariable X, $X \sim exp(\lambda)$ gilt

Wertebereich $X \in \mathbb{R}_0^+$

Parameter $\lambda \in \mathbb{R}^+$: Ankunftsrate der Ereignisse

Verteilungsdichte $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$

Verteilungsfunktion $P(X \leq x) = F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$

Erwartungswert $\mathrm{E}[X] = \frac{1}{\lambda}$

Varianz $\mathrm{Var}[X] = \frac{1}{\lambda^2}$

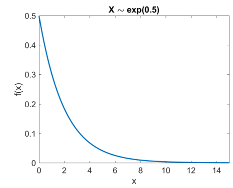


Bild: Verteilungsdichtefunktion

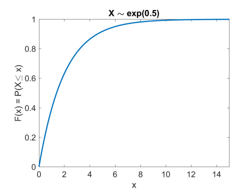


Bild: Verteilungsfunktion

Fakten zur Normalverteilung

Für eine normalverteilte Zufallsvariable X, $X \sim N(\mu, \sigma)$ gilt

Wertebereich $X \in \mathbb{R}$

Parameter $\mu \in \mathbb{R}$: Erwartungswert (Ortsparameter), $\sigma \in \mathbb{R}_0^+$: Standardabweichung (Skalierungsparameter)

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

Verteilungsfunktion $F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$

Erwartungswert $\mathrm{E}[X] = \mu$

Varianz $\mathrm{Var}[X] = \sigma^2$

$X \sim N(0,1)$ heißt standardnormalverteilt mit den Parametern $\mu = 0$ und $\sigma = 1$, (siehe Bilder)

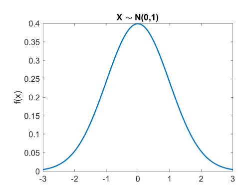


Bild: Verteilungsdichtefunktion

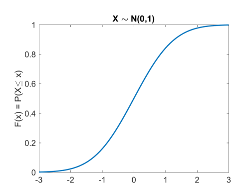


Bild: Verteilungsfunktion

Bedingung np und $n(1-p)$ müssen „groß genug" sein, $n \cdot p \cdot (1-p) \gtrapprox 9$.

xx