Zusammenfassung
Wahrscheinlichkeit und Statistik

Autor:
André Gasser, gassera@student.ethz.ch

Datum:
Januar 2013

1 Kombinatorik

1.1 Permutation

n unterschiedlichen Kugeln:

$$P(n) = n!$$

n unterschiedliche Kugeln mit $n_1, n_2, ..., n_k$ gleichen Kugeln:

$$P(n, n_1, n_2, ..., n_k) = \frac{n!}{n_1!n_2!...n_k!}$$

Bemerkungen:

- Ermitteln der Anzahl Anordnungen.

1.2 Kombination

Mit Zurücklegen:

$$C_w(n,k) = \left(\begin{array}{c} n+k-1 \\ k \end{array} \right) = \frac{(n+k-1)!}{k!(n-1)!}$$

Ohne Zurücklegen:

$$C(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}$$

Bemerkungen:

- Reihenfolge spielt keine Rolle.

1.3 Variation

Mit Zurücklegen:

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

Ohne Zurücklegen:

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

Bemerkungen:

- Reihenfolge ist wesentlich.

2 Wahrscheinlichkeit

2.1 Allgemeine Rechenregeln

$$\begin{array}{rcl} P\left[A^C\right] & = & 1 - P[A] \\ P\left[\Omega\right] & = & P[A] + P\left[A^C\right] = 1 \\ P[\emptyset] & = & 0 \\ P\left[B^C|A\right] & = & 1 - P[B|A] \\ A \subseteq B & \Rightarrow & P[A] \leq P[B] \end{array}$$

2.2 DeMorgan'sche Gesetze

$$\begin{array}{rcl} \overline{A \cap B} & = & \overline{A} \cup \overline{B} \\ \overline{A \cup B} & = & \overline{A} \cap \overline{B} \end{array}$$

2.3 Additionssatz

Zur Berechnung der Wahrscheinlichkeit, dass Ereignis A oder B eintritt.

Allgemein:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

A,B disjunkt:

$$P[A \cup B] = P[A] + P[B]$$

Bemerkungen:

- Wenn die Ereignisse nicht offensichtlich disjunkt sind, die **erste Formel verwenden**!

2.4 Multiplikationssatz

Zur Berechnung der Wahrscheinlichkeit, dass die Ereignisse A und B eintreten.

$$P[A \cap B] = P[A] \cdot P[B|A]$$

Bemerkungen

- $P[A, B] = P[A \cap B]$.

2.5 Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit von B gegeben A entspricht der Wahrscheinlichkeit, dass B eintritt, wenn man schon weiss, dass A eingetreten ist. Es gilt:

$$P[B|A] = \frac{P[A \cap B]}{P[A]}$$

$$P[A \cap B] = P[A] \cdot P[B|A] = P[B] \cdot P[A|B]$$

Bemerkungen:

- Die **Pfadregel**, nach der Wahrscheinlichkeiten in einem **Wahrscheinlichkeitsbaum** multipliziert werden, um die Wahrscheinlichkeit eines Blattes zu erhalten, entspricht einer Verkettung bedingter Wahrscheinlichkeiten.

2.5.1 Satz der totalen Wahrscheinlichkeit

$$P[B] = \sum_{i=1}^{n} P[A_i] \cdot P[B|A_i]$$

2.5.2 Satz von Bayes

Zur Berechnung einer bestimmten Zwischenstation A_k in einem Ereignisbaum, wobei mehrere Ereignisse A_i zu Ereignis B führen.

$$P[A_k|B] = \frac{P[A_k] \cdot P[B|A_k]}{\sum\limits_{i=1}^{n} P[A_i] \cdot P[B|A_i]}$$

2.6 Eigenschaften von Ereignissen

2.6.1 Unabhängigkeit

Wenn zwischen zwei Ereignissen A und B kein kausaler Zusammenhang besteht (d.h. es gibt keine gemeinsamen Ursachen oder Ausschliessungen), dann sind sie *unabhängig* voneinander. In diesem Fall gilt:

$$P[A \cap B] = P[A] \cdot P[B]$$

3 Zufallsvariablen

3.1 Diskrete Zufallsvariablen

3.1.1 Gewichtsfunktion

Die Summe aller Gewichte ist immer 1 und die Werte immer im Intervall [0,1]:

$$\sum_{i=1}^{\infty} f(x) = 1$$

3.1.2 Verteilungsfunktion

3.2 Stetige Zufallsvariablen

3.2.1 Dichtefunktion

Die Fläche unter der Dichtefunktion ist immer 1 und die Werte immer im Intervall [0,1]:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Bemerkungen

- Durch Integrieren der Dichtefunktion erhält man die Verteilungsfunktion.

3.2.2 Verteilungsfunktion

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(u)du$$

Bemerkungen

- Durch Ableiten der Verteilungsfunktion erhält man die Dichtefunktion.

3.3 Gemeinsame Verteilungen

Gemeinsame Verteilungsfunktion: Die gemeinsame Verteilungsfunktion von n Zufallsvariablen $X_1, ..., X_n$ ist die Abbildung $F : \mathbb{R}^n \to [0, 1]$,

$$F(t_1, ..., t_n) := P[X_1 \leq t_1, ..., X_n \leq t_n]$$

Gemeinsame Gewichtsfunktion: Falls $X_1, ..., X_n$ *diskrete* Zufallsvariablen sind, ist ihre gemeinsame Gewichtsfunktion $p : \mathbb{R}^n \to [0, 1]$ definiert durch

$$p(x_1, ..., x_n) := P[X_1 = x_1, ..., X_n = x_n]$$

Gemeinsame Dichte: Seien $X_1, ..., X_n$ *stetige* Zufallsvariablen mit gemeinsamer Verteilungsfunktion $F(t_1, ..., t_n)$. Die Funktion $f : \mathbb{R}^n \to [0, \infty)$ heisst *gemeinsame Dichte* von $X_1, ..., X_n$, falls für alle $t_i \in \mathbb{R}$ gilt

$$F(t_1, ..., t_n) = \int_{-\infty}^{t_1} ... \int_{-\infty}^{t_n} f(x_1, ..., x_n)\ dx_n...dx_1$$

3.3.1 Randverteilungen

Seien X und Y Zufallsvariablen mit gemeinsamer Verteilungsfunktion F_{XY}, dann ist die Randverteilung $F_X : \mathbb{R} \to [0, 1]$ von X definiert durch

$$F_X = P[X \leq x] = P[X \leq x, Y \leq \infty] = \lim_{y\to\infty} F_{XY}(x, y)$$

Für zwei *diskrete* Zufallsvariablen X und Y mit gemeinsamer Gewichtsfunktion $f_{XY}(x, y)$ ist die Gewichtsfunktion der Randverteilung von X gegeben durch

$$f_X = P[X = x] = \sum_j P[X = x, Y = y_j] = \sum_j f_{XY}(x, y_j)$$

Für zwei *stetige* Zufallsvariablen X und Y mit gemeinsamer Dichte $f_{XY}(x, y)$ ist die Dichtefunktion der Randverteilung (Randdichte) von X gegeben durch

$$f_X(x) = \int_{\infty}^{\infty} f_{XY}(x, y)\ dy$$

3.3.2 Bedingte Verteilung

Bedingte Gewichtsfunktion: Seien X und Y *diskrete* Zufallsvariablen mit gemeinsamer Gewichtsfunktion $f_{XY}(x, y)$, dann ist die bedingte Gewichtsfunktion $f_{X|Y}(x|y)$ von X gegeben Y definiert durch

$$f_{X|Y}(x|y) = P[X = x|Y = y] = \frac{f_{XY}(x, y)}{f_Y(y)}$$

Bedingte Dichte: Für zwei *stetige* Zufallsvariablen X und Y mit gemeinsamer Dichte $f_{XY}(x, y)$ ist die bedingte Dichte $f_{X|Y}$ von X gegeben Y definiert durch

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

3.4 Funktionen diskreter Zufallsvariablen

3.4.1 Summe von Zufallsvariablen

...

3.4.2 Produkte von Zufallsvariablen

...

3.5 Funktionen stetiger Zufallsvariablen

Fall $Y = aX + b$:
Verteilungsfunktion:

$$\begin{aligned} F_Y(t) &= P[Y \leq t] = P[aX + b \leq t] = \\ &= P[X \leq \tfrac{t-b}{a}] = F_X(\tfrac{t-b}{a}) \end{aligned}$$

Dichtefunktion:

$$f_Y(t) = \frac{1}{a} f_X\left(\frac{t-b}{a}\right)$$

Fall $Y = X^2$:
Verteilungsfunktion:

$$\begin{aligned} F_Y(t) &= P[Y \leq t] = P[X^2 \leq t] = P[-\sqrt{t} \leq t \leq \sqrt{t}] \\ &= F_X(\sqrt{t}) - F_X(-\sqrt{t}) \end{aligned}$$

Dichtefunktion:

$$f_Y(t) = \frac{1}{\sqrt{2\pi}} t^{-\frac{1}{2}} e^{-\frac{1}{2}t}$$

Fall $Y = \frac{1}{X}$:
Verteilungsfunktion:

$$\begin{aligned} F_Y(t) &= P[Y \leq t] = P[\tfrac{1}{X} \leq t] \\ &= P[X \geq \tfrac{1}{t}] = 1 - P[X \leq \tfrac{1}{t}] = 1 - F_X(\tfrac{1}{t}) \end{aligned}$$

Dichtefunktion:

$$f_Y(t) = \frac{1}{t^2} f_X\left(\frac{1}{t}\right)$$

3.6 Chebyshev-Ungleichung

Die Chebyshev-Ungleichung liefert eine Abschätzung von Wahrscheinlichkeiten, auch wenn die genaue Verteilungsfunktion nicht bekannt ist. Es muss nur der Erwartungswert $E[X]$ und die Varianz $Var[X] < \infty$ einer Zufallsvariablen X bekannt sein, dann gilt für jedes $k > 0$:

$$P\left[|X - E[X]| \geq k\right] \leq \frac{Var[X]}{k^2}$$

3.7 Eigenschaften von Zufallsvariablen

3.7.1 Unabhängigkeit

Zwei Zufallsvariablen X und Y heissen unabhängig, wenn stets gilt

$$\begin{aligned} F(x,y) &= F_X(x) \cdot F_Y(y) \\ f(x,y) &= f_X(x) \cdot F_Y(y) \end{aligned}$$

wobei $F_X(x), F_X(y)$ die Verteilungsfunktionen und $f_X(x), f_Y(y)$ die Gewichts- bzw. Dichtefunktionen von X und Y sind.

3.7.2 Unkorreliert

Zwei Zufallsvariablen X und Y heissen *unkorreliert*, falls gilt $Cov(X,Y) = 0$.
Eine Menge von Zufallsvariablen $X_1, X_2, ..., X_n$ heissen *paarweise unkorreliert*, wenn alle Paare X_i, X_j mit $i \neq j$ unkorreliert sind.

3.8 Erwartungswert

Der Erwartungswert ist das langfristige Durchschnittsergebnis bei einem Zufallsexperiment mit vielen Wiederholungen.

Diskrete Verteilung:

$$E[X] = \sum_i x_i \cdot f(x_i)$$

Stetige Verteilung:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

3.8.1 Additionssatz für Erwartungswerte

Der Erwartungswert einer aus n (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Summe

$$Z = a_1X_1 + a_2X_2 + ... + a_nX_n$$

ist gleich der Summe der Erwartungswerte der einzelnen Zufallsvariablen:

$$E[Z] = a_1E[X_1] + a_2E[X_2] + ... + a_nE[X_n]$$

3.8.2 Multiplikationssatz für Erwartungswerte

Der Erwartungswert eines aus n *stochastisch unabhängigen* (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Produkts

$$Z = X_1 \cdot X_2 \cdot ... \cdot X_n$$

ist gleich dem Produkt der Erwartungswerte der einzelnen Zufallsvariablen:

$$E[Z] = E[X_1 \cdot X_2 \cdot ... \cdot X_n] = E[X_1] \cdot E[X_2] \cdot ... \cdot E[X_n]$$

3.8.3 Formeln für 2. Moment

Dieses wird z.B. für die Momenten-Methode und zur Berechnung der Varianz benötigt.

Diskrete Verteilung:

$$E[X^2] = \sum_i x_i^2 \cdot f(x_i)$$

Stetige Verteilung:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx$$

3.8.4 Weitere Rechenregeln

$$\begin{aligned} E[a] &= a (\text{für a = const.}) \\ E[aX] &= a \cdot E[X] (\text{für a = const.}) \\ E[aX + b] &= a \cdot E[X] + b \\ E[E[X]] &= E[X] \end{aligned}$$

3.9 Varianz

Die Varianz ist ein Streuungsmass, also ein Mass für die Abweichung einer Zufallsvariable von ihrem Erwartungswert.

Diskrete Verteilung:

$$Var[X] = \sum_i (x_i - E[X])^2 \cdot f(x_i)$$

Stetige Verteilung:

$$Var[X] = \int_{-\infty}^{\infty} (x - E[X])^2 \cdot f(x)dx$$

Vereinfachte Berechnung:

$$Var[X] = E[X^2] - E[X]^2$$

3.9.1 Additionssatz für Varianzen

Die Varianz einer aus n *stochastisch unabhängigen* (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Summe

$$Z = a_1X_1 + a_2X_2 + ... + a_nX_n$$

ist gleich der Summe der Varianzen der einzelnen Zufallsvariablen:

$$Var[Z] = a_1^2Var[X_1] + a_2^2Var[X_2] + ... + a_n^2Var[X_n]$$

3.9.2 Weitere Rechenregeln

$$\begin{aligned} Var[aX + b] &= a^2 \cdot Var[X] \\ Var[aX + bY] &= a^2 \cdot Var[X] + b^2 \cdot Var[Y] + \\ & \quad 2ab \cdot Cov[X, Y] \end{aligned}$$

3.10 Kovarianz

Die Kovarianz ist ein Mass für den Zusammenhang zwischen zwei Zufallsvariablen X und Y bzw. der Streuung zwischen ihnen.

$$\begin{aligned} Cov[X, Y] &= E[(X - E[X])(Y - E[Y])] \\ Cov[X, Y] &= E[XY] - E[X] \cdot E[Y] \\ Cov[X, Y] &= \tfrac{1}{2} \cdot (Var[X + Y] - Var[X] - Var[Y]) \\ Cov[X, Y] &= Cov[Y, X] \\ Cov[X, X] &= Var[X] \\ Cov[aX + b, Y] &= a \cdot Cov[X, Y] \\ Cov[X + Y, Z] &= Cov[X, Z] + Cov[Y, Z] \\ \forall a \in \mathbb{R} : Cov[X, a] &= 0 \\ \forall b \in \mathbb{R} : Cov[X, bY] &= b \cdot Cov[X, Y] \end{aligned}$$

3.11 Standardabweichung

$$sd[X] = \sqrt{Var[X]}$$

4 Indikatorfunktion $\mathbb{1}$

Die Indikatorfunktion ist eine spezielle Zufallsvariable, welche angibt ob ein Ereignis A eingetreten ist.

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases} \tag{1}$$

5 Diskrete Verteilungen

5.1 Bernoulli-Verteilung

Verteilung eines Experiments mit zwei Ausgängen (Erfolg, Misserfolg) mit Erfolgsparameter p.

Notation:

$$X \sim Be(p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = p^x \cdot (1 - p)^{1-x}$$

Erwartungswert:

$$E[X] = p$$

Varianz:

$$Var[X] = p(1 - p)$$

Beispiel (Münzwurf): Ein fairer Münzwurf ist bernoulliverteilt mit Parameter $p = \frac{1}{2}$. Für einen Parameter $p \neq \frac{1}{2}$ wäre der Münzwurf unfair.

5.2 Binomial-Verteilung

Beschreibt die Anzahl der Erfolge in einer Serie von *gleichartigen* und *unabhängigen* Versuchen, die jeweils genau zwei mögliche Ergebnisse haben ("Erfolg" oder "Misserfolg"). n ist die Anzahl der Versuche bzw. Wiederholungen, p ist die Wahrscheinlichkeit für einen Ërfolg".

Notation:

$$X \sim Bin(n, p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \binom{n}{x} p^x(1 - p)^{n-x}$$

Verteilungsfunktion:

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

Erwartungswert:

$$E[X] = np$$

Varianz:

$$Var[X] = np(1 - p)$$

Bemerkungen:

- Die Binomialverteilung $Bin(n, p)$ darf für grosse n und kleine p näherungsweise durch die Poisson-Verteilung mit Parameter $\lambda = np$ ersetzt werden (Faustregel: $np < 10$ und $n > 1500p$).

Beispiel: Geburtstagsproblem, Anzahl der Köpfe beim 10-maligen Münzwurf.

5.3 Geometrische Verteilung

Wartezeit auf ersten Erfolg bei einer Folge von 0-1-Experimenten mit Erfolgsparameter p. X ist die Nummer des ersten erfolgreichen Experiments.

Notation:

$$X \sim Geom(p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = p(1-p)^{x-1}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = 1 - (1-p)^x$$

Erwartungswert:

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

Varianz:

$$Var[X] = \frac{1-p}{p^2}$$

Beispiel: Wartezeit auf Kopf bei wiederholtem Münzwurf

5.4 Poisson-Verteilung

Ereignet sich in einem Intervall (z.B. in einer gewissen Zeit, auf einer gewissen Fläche, in einem gewissen Volumen, usw.) ein völlig zufällig auftretendes Ereignis im Schnitt μ mal (Erwartungswert), dann ist die Zufallsgrösse, welche die Häufigkeit des Ereignisses in diesem Intervall angibt, poissonverteilt mit Parameter $\lambda = \mu$.

Notation:

$$X \sim P(\lambda), X \sim Pois(\lambda)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = e^{-\lambda}\frac{\lambda^x}{x!}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = e^{-\lambda}\sum_{k=0}^{x}\frac{\lambda^k}{k!}$$

Erwartungswert:

$$E[X] = \lambda$$

Varianz:

$$Var[X] = \lambda$$

Bemerkungen

- Falls $Z = X + Y$ und $X \sim Pois(\lambda)$, $Y \sim Pois(\mu)$ dann gilt $Z \sim Pois(\lambda + \mu)$.

Beispiel: Anzahl eingehender Druckaufträge in $n = 3600$ Sekunden mit einer Jobwahrscheinlichkeit von $p = 1\%$ pro Sekunde, wenn pro Sekunde maximal ein Job eintreffen kann.

5.5 Negativbinomiale Verteilung

Wartezeit auf den r-ten Erfolg bei einer Folge von unabhängigen 0-1-Experimenten mit Erfolgsparameter p.
Notation:

$$X \sim NB(r, p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \sum_{i=r}^{x} \binom{i-1}{r-1} p^r (1-p)^{i-r}$$

Erwartungswert:

$$E[X] = \frac{r}{p}$$

Varianz:

$$Var[X] = \frac{r(1-p)}{p^2}$$

Beispiel: Wartezeit auf dritten Kopf beim wiederholten Münzwurf

5.6 Hypergeometrische Verteilung

Urnenmodell für Ziehen ohne Zurücklegen. In einer Urne befinden sich n Gegenstände. Davon sind r Gegenstände vom Typ A und $n-r$ Gegenstände vom Typ B. Es werden m Gegenstände ohne Zurücklegen gezogen. X beschreibt die Wahrscheinlichkeitsverteilung für die Anzahl k der Gegenstände vom Typ A in der Stichprobe.

Notation:

$$X \sim HypGeom(n, m, r)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \frac{\binom{r}{x}\binom{n-r}{m-x}}{\binom{n}{m}}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \sum_{k=max(0,m-n)}^{x} \frac{\binom{r}{k}\binom{n}{m-k}}{\binom{r+n}{m}}$$

Erwartungswert:

$$E[X] = \frac{mr}{n}$$

Varianz:

$$Var[X] = \frac{mr}{n^2(n-1)}(n-r)(n-m)$$

Beispiel (Lotto): Anzahl Zahlen $n = 45$, richtige Zahlen $r = 6$, meine Zahlen $m = 6$. Die Wahrscheinlichkeit für 4 Richtige ist

$$f(4) = \frac{\binom{6}{4}\binom{39}{2}}{\binom{35}{6}} \approx 0.00136$$

6 Stetige Verteilungen

6.1 Gleichverteilung

Alle Ereignisse zwischen a und b sind gleich wahrscheinlich.

Notation:

$$X \sim U(a,b)$$

Dichtefunktion:

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

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:

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

Varianz:

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

6.2 Exponentialverteilung

Die Exponentialverteilung ist ein Modell für Wartezeiten und Lebensdauern. Beispielanwendungen sind die Berechnung der Lebensdauer von Bauteilen und Zugverspätungen.

Notation:

$$X \sim Exp(\lambda)$$

Dichtefunktion:

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

Verteilungsfunktion:

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

Erwartungswert:

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

Varianz:

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

Bemerkungen:

- Sind $X_1 \sim Exp(\lambda_1), ..., X_n \sim Exp(\lambda_n)$ *stochastisch unabhängig*, so ist $\min(X_1, ..., X_n) \sim Exp(\lambda_1 + ... + \lambda_n)$.
- Die Summe von n exponentialverteilten Zufallsvariablen mit *gleichem Parameter* λ ist gammaverteilt mit $\Gamma(n, \lambda)$.
- Erwartungswert und Standardabweichung sind gleich ($E[X] = sd[X]$).

Beispiel (Lebensdauer:) Die Exponentialverteilung ist eine typische Lebensdauerverteilung. So ist z.B. die Lebensdauer von elektronischen Bauteilen häufig annähernd exponentialverteilt.

6.3 Normalverteilung

Die besondere Bedeutung der Normalverteilung beruht unter anderem auf dem zentralen Grenzwertsatz, der besagt, dass eine Summe von n unabhängigen, identisch verteilten Zufallsvariablen mit endlicher Varianz im Grenzwert $n \to \infty$ normalverteilt ist.

Notation:

$$X \sim N(\mu, \sigma^2)$$

Dichtefunktion:

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

Verteilungsfunktion:

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

Erwartungswert:

$$E[X] = \mu$$

Varianz:

$$Var[X] = \sigma^2$$

Transformation einer normalverteilten Zufallsvariable X in eine standardnormalverteilte Zufallsvariable Z:

$$Z = \frac{X - \mu}{\sigma}$$

Beispiel: Gewicht von Kürbissen, Streuung von Messungen um den Mittelwert, Gewichte oder Grössen von Individuen in einer großen Bevölkerung.

6.4 Standardnormalverteilung

Notation:

$$X \sim N(0,1)$$

Dichtefunktion:

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

Verteilungsfunktion:

$$\Phi(u) = P[X \leq u] = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{u} e^{-\frac{t^2}{2}} dt$$

Erwartungswert:

$$E[X] = 0$$

Varianz:

$$Var[X] = 1$$

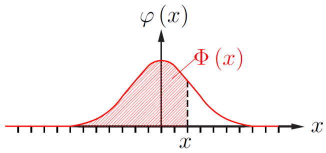
Bemerkungen:

- Die Standardnormalverteilung entspricht einer Normalverteilung mit den Parametern $\mu = 0$ und $\sigma = 1$.

6.4.1 Tabelle ablesen

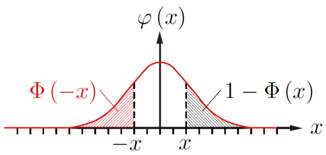
$\Phi(x)$ bedeutet, dass man zu einem Wert x die Wahrscheinlichkeit in der Tabelle auslesen will. $\Phi^{-1}(x)$ bedeutet, dass für eine Wahrscheinlichkeit der Tabellenwert ermittelt werden soll.

Positive Werte: $\Phi(x)$



Bsp: $\Phi(1.25) = 0.8944$

Negative Werte: $\Phi(-x) = 1 - \Phi(x)$



Bsp: $\Phi(-1.25) = 1 - \Phi(1.25) = 1 - 0.8944$

Für nicht tabellierte Werte:

$$\Phi^{-1}(0.03) = -\Phi^{-1}(0.97) = -1.88$$

Regeln zum Auslesen der Werte:

- Ist ein gesuchter Wert c nicht tabelliert, so wird derjenige Wert a oder b genommen, der näher bei c liegt.
- Sind zwei Werte a und b gleich weit vom gesuchten Wert c entfernt, so wird der Mittelwert von a und b verwendet.

6.5 Pareto-Verteilung

Die Pareto-Verteilung, ist eine stetige Wahrscheinlichkeitsverteilung auf einem rechtsseitigen unendlichen Intervall $[x_0, \infty)$. Eine stetige Zufallsvariable X heisst pareto-verteilt $Par(\alpha, x_0)$ mit den Parametern $\alpha > 0$ und $x_0 > 0$, mit folgenden Eigenschaften:

Notation:

$$X \sim Par(\alpha, x_0)$$

Dichtefunktion:

$$f(x) = \begin{cases} \frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1} & x \geq x_0 \\ 0 & x < x_0 \end{cases}$$

Verteilungsfunktion:

$$F(x; x_0, \alpha) = \begin{cases} 1 - \left(\frac{x}{x_0}\right)^{-\alpha} & x \geq x_0, \alpha > 0, \\ 0 & \text{sonst} \end{cases}$$

Erwartungswert:

$$E[X] = \begin{cases} x_0 \frac{\alpha}{\alpha - 1} & \alpha > 1 \\ \infty & \alpha \leq 1 \end{cases}$$

Varianz:

$$Var[X] = \begin{cases} x_0^2 \left(\frac{\alpha}{\alpha-2} - \frac{\alpha^2}{(\alpha-1)^2}\right) & \\ = x_0^2 \frac{\alpha}{(\alpha-2)(\alpha-1)^2} & \alpha > 2 \\ \infty & \alpha \leq 2 \end{cases}$$

Standardabweichung:

$$sd[X] = \frac{x_0}{\alpha - 1}\sqrt{\frac{\alpha}{\alpha - 2}} \text{ für } \alpha > 2$$

Beispiel: Die Verteilung wurde zunächst zur Beschreibung der Einkommensverteilung Italiens verwendet. Paretoverteilungen finden sich charakteristischerweise dann, wenn sich zufällige, positive Werte über mehrere Größenordnungen erstrecken und durch das Einwirken vieler unabhängiger Faktoren zustande kommen.

7 Grenzwertsätze

In vielen Situationen taucht die Summe von *vielen gleichartigen Zufallsvariablen* auf. Wir möchten wissen, wie sich diese Summe etwa verhält, und untersuchen deshalb ihre Asymptotik, wenn die Anzahl der Summanden gegen unendlich geht.
Für die folgenden Sätze definieren wir einige Grössen: $S_n = \sum_{i=1}^{n} X_i$ und $\overline{X}_n = \frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

7.1 Schwaches Gesetz der grossen Zahlen

Seien $X_1, X_2, ..., X_n$ *unabhängige* oder *paarweise unkorrelierte* Zufallsvariablen mit gleichem Erwartungswert $E[X_i] = \mu$ und gleicher Varianz $Var[X_i] = \sigma^2$. Sei $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Dann gilt:

$$\forall \epsilon > 0 : \lim_{n\to\infty} P\left[|\overline{X}_n - \mu| > \epsilon\right] = 0$$

Bemerkungen:

- Für hinreichend grosse n konvergiert der Mittelwert $\overline{X}_n$ gegen den Erwartungswert μ.
- Der Satz funktioniert nicht, falls der Erwartungswert oder die Varianz nicht definiert sind.

7.2 Starkes Gesetz der grossen Zahlen

Seien $X_1, ..., X_n$ *unabhängige* Zufallsvariablen mit *gleicher Verteilung* (i.i.d) und Erwartungswert $E[X_i] = \mu$. Sei $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Dann gilt:

$$P\left[\lim_{n\to\infty} \overline{X}_n = \mu\right] = 1$$

Bemerkungen:

- Für hinreichend grosse n konvergiert der Mittelwert $\overline{X}_n$ gegen den Erwartungswert μ.

7.3 Zentraler Grenzwertsatz

$X_1, X_2, X_3, ..., X_n, ...$ seien *stochastisch unabhängige* Zufallsvariablen, die alle der *gleichen* Verteilungsfunktion mit Erwartungswert μ und Varianz σ^2 genügen. Dann konvergiert die Verteilungsfunktion $F_Z(u)$ der standardisierten Zufallsvariablen

$$Z_n = \frac{(X_1 + X_2 + ... + X_n) - n\mu}{\sqrt{n}\sigma}$$

im Grenzfall $n \to \infty$ gegen die Verteilungsfunktion $\Phi(t)$ der Standardnormalverteilung:

$$\lim_{n\to\infty} F_Z(u) = \Phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^u e^{-\frac{t^2}{2}} dt$$

Bemerkungen:

- Für ein hinreichend grosses n ist $Z_n = X_1 + X_2 + ... + X_n$ annähernd *normalverteilt* mit Erwartungswert $E[Z_n] = n\mu$ und Varianz $Var[Z_n] = n\sigma^2$.

8 Statistik

8.1 Schätzer

Schätzer sind Funktionen von Zufallsvariablen und somit selbst wieder Zufallsvariablen. Sie verfügen deshalb über einen Erwartungswert und eine Varianz.

8.1.1 Erwartungstreuer Schätzer

Ein Schätzer T ist *erwartungstreu* wenn der Erwartungswert des Schätzers gleich dem zu schätzenden Parameter ϑ ist:

$$E_\vartheta[T] = \vartheta$$

8.1.2 Konsistenter Schätzer

Ein Folge von Schätzern $T^{(n)}$ ist *konsistent* wenn diese mit zunehmendem Stichprobenumfang n gegen den gesuchten Parameter ϑ konvergiert:

$$\lim_{n\to\infty} P_\vartheta \left[|T^{(n)} - \vartheta| > \epsilon \right] = 0 \text{ (für jedes } \epsilon > 0)$$

8.1.3 Momenten-Methode

Sei $X_1, X_2, ..., X_n$ eine Stichprobe vom Umfang n mit gegebener Verteilung t. Die Parameter von t seien unbekannt. Mit der Momenten-Methode können diese geschätzt werden. Der Momentenschätzer ist i.d.R. nicht erwartungstreu.

Vorgehen:

1. 1. Moment berechnen:
$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$
2. 2. Moment berechnen:
$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right) - \overline{x}^2$$
3. Nimm die Formel für den Erwartungswert der geg. Verteilung t und setze sie gleich $\overline{x}$
4. Nimm die Formel für die Varianz der geg. Verteilung t und setze sie gleich s^2
5. Löse das Gleichungssystem auf. Du erhälst die gesuchten Parameter.

8.1.4 Maximum-Likelihood-Methode

Methode zur systematischen Gewinnung von Schätzfunktionen.

Vorgehen:

1. Likelihood-Funktion aufstellen:
$$L(x_1, ..., x_n, \vartheta) = \prod_{i=1}^{n} f(x_i, \vartheta)$$
2. Log-Likelihood-Funktion aufstellen (logarithmieren von L):
$$\log L(x_1, ..., x_n, \vartheta)$$
3. $\log L$ mit Hilfe elementarer Logarithmenregeln möglichst vereinfachen
4. $\log L$ nach jedem unbekannten Parameter partiell ableiten
5. Partielle Ableitungen $= 0$ setzen
6. Gleichungssystem nach Parametern auflösen

8.2 Tests

8.2.1 Wichtige Begriffe

- **Einseitiger Test:** Einen einseitigen Test führt man durch, wenn man wissen will, ob sich ein Wert vergrössert oder verkleinert hat.
- **Zweiseitiger Test:** Einen zweiseitigen Test führt man durch, wenn man lediglich wissen will, ob sich ein Wert verändert hat.
- **Nullhypothese H_0:** Die zu prüfende Annahme.
- **Alternativhypothese** H_A: Alternative Annahme, falls H_0 verworfen werden muss.
- **Signifikanzniveau** α: Ist die Wahrscheinlichkeit, dass die Nullhypothese verworfen werden muss (auch Wahrscheinlichkeit für Fehler 1. Art).
- **Teststatistik T:** Test- oder Prüfwert.
- **P-Wert:** Das kleinste Niveau, auf dem der Test die Nullhypothese noch verwirft.
- **Fehler 1. Art:** H_0 wird verworfen, obwohl sie richtig wäre.
- **Fehler 2. Art:** H_0 wird beibehalten, obwohl H_A stimmt.

8.2.2 Allgemeines Vorgehen

1. Wahl des Modells
2. Formulieren der Nullhypothese H_0
3. Formulieren der Alternativhypothese H_A
4. Teststatistik T aufstellen
5. Verteilung der Teststatistik unter der Nullhypothese
6. Bestimmen des Verwerfungsbereichs
7. Konkreter Wert für Teststatistik T berechnen
8. Testentscheidung: H_0 beibehalten oder verwerfen?

8.2.3 z-Test

Der z-Test ist ein Test für den Erwartungswert bei *bekannter* Varianz σ^2. Es seien also $X_1, X_2, ..., X_n \sim N(\mu, \sigma^2)$ *i.i.d.*. Wir wollen die Nullhypothese $\mu = \mu_0$ testen.

Nullhypothese und Alternativhypothese:

a) $H_0 : \mu = \mu_0$ gegen $H_A : \mu \neq \mu_0$ (zweiseitig)

b) $H_0 : \mu = \mu_0$ gegen $H_A : \mu > \mu_0$ (einseitig)

c) $H_0 : \mu = \mu_0$ gegen $H_A : \mu < \mu_0$ (einseitig)

Teststatistik:

Mittelwert $\overline{X}$: $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$

Teststatistik T: $T = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

Verwerfungsbereich:

a) $|T| > t_{1-\frac{\alpha}{2}} = \Phi^{-1}(1-\frac{\alpha}{2}) = c$
$(P_{\vartheta_0}[|T|] > \Phi^{-1}(1-\frac{\alpha}{2}] = \alpha)$
Verwerfungsbereich $K = (-\infty, -c) \cup (c, \infty)$

b) $T > \Phi^{-1}(1-\alpha) = c$
Verwerfungsbereich $K = (c, \infty)$

c) $T < t_\alpha = -t_{1-\alpha} = -\Phi^{-1}(1-\alpha) = c$
Verwerfungsbereich $K = (-\infty, -c)$

8.2.4 t-Test

Der t-Test ist ein Test für den Erwartungswert bei *unbekannter Varianz* σ^2. Es seien also $X_1, X_2, ..., X_n \sim N(\mu, \sigma^2)$*i.i.d.*. Wir wollen die Nullhypothese $\mu = \mu_0$ testen.

Nullhypothese und Alternativhypothese:

a) $H_0 : \mu = \mu_0$ gegen $H_A : \mu \neq \mu_0$ (zweiseitig)

b) $H_0 : \mu = \mu_0$ gegen $H_A : \mu > \mu_0$ (einseitig)

c) $H_0 : \mu = \mu_0$ gegen $H_A : \mu < \mu_0$ (einseitig)

Teststatistik:

Mittelwert $\overline{X}$: $\overline{X} = \frac{X_1+X_2+...+X_n}{n}$

Schätzfunktion für S^2: $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2$

Teststatistik T: $T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$

Die Teststatistik T ist **t-verteilt** mit $n-1$ **Freiheitsgraden**.

Verwerfungsbereich:

a) $|T| > t_{n-1,1-\frac{\alpha}{2}} = c$
Verwerfungsbereich $K = (-\infty, -c) \cup (c, \infty)$

b) $T > t_{n-1,1-\alpha} = c$
Verwerfungsbereich $K = (c, \infty)$

c) $T < -t_{n-1,1-\alpha} = -c$
Verwerfungsbereich $K = (-\infty, -c)$

8.2.5 Likelihood-Quotienten-Test

Der Likelihood-Quotienten-Test kann man verwenden, um eine geeignete Teststatistik T zu erhalten.

Vorgehen:

1. Verallgemeinerter Likelihood-Quotient aufstellen:

$$R(x_1, x_2, ..., x_n; \vartheta_0, \vartheta_A) := \frac{L(x_1, x_2, ..., x_n; \vartheta_0)}{L(x_1, x_2, ..., x_n; \vartheta_A)}$$

2. Formel vereinfachen. Im Exponent muss eine Summenformel stehen.
3. Überlegen, welcher Wert grösser ist, ϑ_0 oder ϑ_A?

Ist dieser Quotient klein, sind die Beobachtungen für die Alternativhypothese deutlich wahrscheinlicher als für die Nullhypothese. Der Verwerfungsbereich $K := [0, c)$ wird so gewählt, dass der Test das gewünschte Signifikanzniveau einhält.

8.2.6 Ungepaarter Zweistichproben-z-Test

Seien $X_1, X_2, ..., X_n \sim N(\mu_X, \sigma^2)$ und $Y_1, Y_2, ..., Y_m \sim N(\mu_Y, \sigma^2)$ zwei Stichproben mit $m = n$ oder $m \neq n$. Die Erwartungswerte μ_X und μ_Y seien unbekannt, die Varianz σ^2 sei *bekannt*. Der Test kann wie ein normaler z-Test durchgeführt werden, als Teststatistik T wird jedoch folgende Formel verwendet:

Falls $m \neq n$:

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1)$$

Falls $m = n$:

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{2}{n}}} \sim N(0,1)$$

8.2.7 Ungepaarter Zweichstichproben-t-Test

Seien $X_1, X_2, ..., X_n \sim N(\mu_X, \sigma^2)$ und $Y_1, Y_2, ..., Y_m \sim N(\mu_Y, \sigma^2)$ zwei Stichproben mit $m = n$ oder $m \neq n$. Die Erwartungswerte μ_X und μ_Y seien unbekannt, die Varianz σ^2 sei ebenfalls *unbekannt*. Der Test kann wie ein normaler t-Test durchgeführt werden, als Teststatistik T wird jedoch folgende Formel verwendet:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

$$S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(Y_j - \overline{Y}_m\right)^2$$

Falls $m \neq n$:

$$S^2 = \frac{1}{m+n-2}\left((n-1)S_X^2 + (m-1)S_Y^2\right)$$

$$T = \frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{S\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$$

Falls $m = n$:

$$S^2 = \frac{1}{2(n-1)}(n-1)\left(S_X^2 + S_Y^2\right)$$

$$T = \frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{S\sqrt{\frac{2}{n}}} \sim t_{2n-2}$$

8.2.8 Gepaarter Zweistichproben-Test

In diesem Fall können die beiden Stichproben $X_1, X_2, ..., X_n\ i.i.d \sim N(\mu_X, \sigma_X^2)$ und $Y_1, Y_2, ..., Y_n\ i.i.d \sim N(\mu_Y, \sigma_Y^2)$ durch Definition einer neuen Zufallsvariable $Z = X_i - Y_i$ auf eine Stichprobe vereinfacht werden. Der Test kann dann wie ein normaler z-Test oder t-Test durchgeführt werden, wobei $Z_1, Z_2, ..., Z_n\ i.i.d \sim N(\mu_X - \mu_Y, 2\sigma^2)$ gilt.

8.3 Konfidenzbereiche

Ein Konfidenzbereich gibt ein Intervall an, in dem sich ein gesuchter Parameter mit sehr hoher Wahrscheinlichkeit befindet.

Standardabweichung σ *bekannt*:

$$C(X_1, ..., X_n) =$$

$$\left[\overline{X}_n - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}, \overline{X}_n + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right]$$

Standardabweichung σ *nicht bekannt*:

$$C(X_1, ..., X_n) =$$

$$\left[\overline{X}_n - t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \overline{X}_n + t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$$

9 Differentialrechnung

9.1 Faktorregel

Ein konstanter Faktor bleibt beim Differenzieren erhalten:

$$y = C \cdot f(x) \Rightarrow y' = C \cdot f'(x)$$

9.2 Summenregel

Bei einer endlichen Summe von Funktionen darf gliedweise differenziert werden:

$$y = f_1(x) + f_2(x) + ... + f_n(x) \Rightarrow y' = f_1'(x) + f_2'(x) + ... + f_n'(x)$$

9.3 Produktregel

Die Ableitung einer in Produktform $y = u(x) \cdot v(x)$ darstellbaren Funktion erhält man nach folgender Produktregel:

$$y' = u'(x) \cdot v(x) + u(x) \cdot v'(x) = u'v + uv'$$

9.4 Quotientenregel

Die Ableitung einer Funktion, die als Quotient zweier Funktionen $u(x)$ und $v(x)$in der Form $y = \frac{u(x)}{v(x)}$ darstellbar ist, erhält man nach der Quotientenregel:

$$y' = \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{v^2(x)}$$

9.5 Kettenregel

Die Ableitung einer zusammengesetzten (verketteten) Funktion $y = F(u(x)) = f(x)$ erhält man als Produkt aus äusserer und innerer Ableitung:

$$y' = \frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

9.6 Partielle Differentiation

Summanden, die keine Variable beinhalten nach der abgeleitet wird, fallen WEG!

9.7 Wichtige elementare Ableitungen

$$\begin{array}{lll}
f(x) = c & \rightarrow & f'(x) = 0 \\
f(x) = x^n & \rightarrow & f'(x) = n \cdot x^{n-1} \\
f(x) = \sqrt{x} & \rightarrow & f'(x) = \frac{1}{2\sqrt{x}} \\
f(x) = e^x & \rightarrow & f'(x) = e^x \\
f(x) = a^x & \rightarrow & f'(x) = (\ln a) \cdot a^x \\
f(x) = \ln x & \rightarrow & f'(x) = \frac{1}{x} \\
f(x) = \log_a x & \rightarrow & f'(x) = \frac{1}{(\ln a) \cdot x}
\end{array}$$

10 Integralrechnung

10.1 Faktorregel

Ein konstanter Faktor darf vor das Integral gezogen werden:

$$\int_a^b C \cdot f(x)dx = C \cdot \int_a^b f(x)dx$$

10.2 Summenregel

Eine endliche Summe von Funktionen darf giedweise integriert werden:

$$\int_a^b (f_1(x) + f_2(x) + ... + f_n(x))\, dx =$$

$$\int_a^b f_1(x)dx + \int_a^b f_1(x)dx + ... + \int_a^b f_1(x)dx$$

10.3 Vertauschungsregel

Vertauschen der beiden Integrationsgrenzen bewirk einen Vorzeichenwechsel des Integrals:

$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$

10.4 Gleiche Integrationsgrenzen

Fallen die Integrationsgrenzen zusammen ($a = b$), so ist der Integralwert gleich Null:

$$\int_a^b f(x)dx = 0$$

10.5 Partielle Integration

$$\int f(x)dx = \int u(x)\cdot v'(x)dx = u(x)\cdot v(x) - \int u'(x)\cdot v(x)dx$$

Vorgehen:

1. Integrand in $u(x)$ und $v'(x)$ zerlegen.
2. $u(x)$ ableiten, $v'(x)$ integrieren.
3. Formel aufschreiben und lösen.

10.6 Wichtige Stammintegrale

$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad \int a^x dx = \frac{a^x}{\ln a} + C$

$\int \frac{1}{x} dx = \ln|x| + C \quad \int e^x dx = e^x + C$

10.7 Weitere Integrale

$$\begin{array}{rcll}
\int a\ dx & = & ax & \\
\int x^a\ dx & = & \frac{1}{a+1}x^{a+1}, & a \neq -1 \\
\int (ax+b)^c\ dx & = & \frac{1}{a(c+1)}(ax+b)^{c+1}, & c \neq -1 \\
\int \frac{1}{x}\ dx & = & \log|x|, & x \neq 0 \\
\int \frac{1}{ax+b}\ dx & = & \frac{1}{a}\log|ax+b| & \\
\int \frac{1}{x^2+a^2}\ dx & = & \frac{1}{a}\arctan\frac{x}{a} & \\
\int e^{ax}\ dx & = & \frac{1}{a}e^{ax} & \\
\int xe^{ax}\ dx & = & \frac{e^{ax}}{a^2}(ax-1) & \\
\int x^2e^{ax}\ dx & = & e^{ax}\left(\frac{x^2}{a} - \frac{2x}{a^2} + \frac{2}{a^3}\right) & \\
\int \log|x|\ dx & = & x(\log|x| - 1) & \\
\int \log_a|x|\ dx & = & x(\log_a|x| - \log_a e) & \\
\int x^a \log x\ dx & = & \frac{x^{a+1}}{a+1}\left(\log x - \frac{1}{a+1}\right), & a \neq -1, x > 0 \\
\int \frac{1}{x}\log x\ dx & = & \frac{1}{2}\log^2 x, & x > 0
\end{array}$$

11 Verschiedenes

Binomialkoeffizient
Der Binomialkoeffizient

$$\binom{n}{k}$$

gibt für $n, k \in \mathbb{N}$ an, **wie viele Möglichkeiten es gibt,** k **Objekte aus** n **Objekten auszuwählen**. Damit gibt der Binomialkoeffizient an, wie viele k-elementige Teilmengen aus einer n-elementigen Menge gebildet werden können (gesprochen: "k aus n" oder "k tief n").

Für $k = 0$ ist der Binomialkoeffizient 1:

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

Für $k = 1$ ist der Binomialkoeffizient n:

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

Für $k > n$ ist der Binomialkoeffizient stets 0.

Gammafunktion

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt = (x-1)!$$

Diverse Summenformeln

$$\sum_{k=0}^{\infty} \frac{1}{k!} = e$$

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^\lambda$$

$$\sum_{k=1}^{\infty} \frac{\lambda^k}{k!} = e^\lambda - 1$$

$$\sum_{i=0}^{n} i = \sum_{i=1}^{n} i = \frac{n(n-1)}{2}$$

$$\sum_{i=1}^{n} i^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6}$$

$$\sum_{i=0}^{\infty} a \cdot p^i = \sum_{i=1}^{\infty} a \cdot p^{i-1} = \frac{a}{1-p}$$

Mitternachtsformel
Formel zum Auflösen von allgemeinen quadratischen Gleichungen der Form $ax^2 + bx + c = 0$:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Produkteformeln

$$\prod_{k=m}^{n} a \cdot x_i = a^{n-m+1} \cdot \prod_{k=m}^{n} x_i$$

Logarithmenregeln

$$\begin{aligned}
\log(uv) &= \log(u) + \log(v) \\
\log\left(\tfrac{u}{v}\right) &= \log(u) - \log(v) \\
\log_b(r) &= \tfrac{\log_a(r)}{\log_a(b)} \\
\log_a(u^k) &= k \cdot \log_a(u) \\
\log_a(\sqrt[n]{u}) &= \log_a(u^{\frac{1}{n}}) = \tfrac{1}{n} \cdot \log_a(u) \\
\log_a(a^b) &= b \\
\log 1 &= 0 \\
\log_a(a) &= 1 \\
\ln e &= 1
\end{aligned}$$

12 Beispiele

12.1 t-Test

Ein Waschmittelhersteller bringt 5kg-Packungen in den Umlauf. Die Konsumentenschutzorganisation kauft 25 Packungen. Es ergibt sich ein Mittel von $\overline{X} = 4.9kg$ und eine empirische Stichprobenvarianz $S^2 = 0.1kg^2$. Die einzelnen Gewichte seine durch unabhängige $N(\mu, \sigma^2)$ Zufallsvariablen beschrieben.

1. Wie lauten die Hypothesen H_0 und H_A?
2. Wie ist $(\overline{X} - 5)/(S/5))$ verteilt unter H_0?
3. Führen Sie den t-Test auf dem 5
4. Berechnen Sie das 95%-Vertrauensintervall für den in 3. beschriebenen Test.

Lösung:

1. $H_0 : \mu = 5, H_1 : \mu < 5$
2. Verteilung: $\sim t_{5^2-1} = t_{24}$
3. $T_1 = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} = -1.581$. Wir ermitteln das t-Quantil für $n-1 = 24$ Freiheitsgrade und $t^{-1}(\alpha) = t^{-1}(0.05) = t^{-1}(0.95) = 1.711$. Somit erhalten wir eine kritische Grenze von $c = -1.711$. Weil $T_1 = -1.581 > c$ ist, wird die Nullhypothese nicht verworfen.
4. ???

12.2 P-Wert berechnen

Für einen t-Test oder einen z-Test kann der P-Wert wie folgt berechnet werden:

P-Wert bei zweiseitigem Test:
Beispiel mit $c = 3.43$ als Grenze des Verwerfungsbereichs:

$$\begin{aligned}
P_{H_0}[|T| > 3.43] &= 2 \cdot P_{H_0}[T > 3.43] \\
&= 2 \cdot (1 - P_{H_0}[T \leq 3.43] \\
&= 2 \cdot (1 - \Phi(3.43)) \\
&= 2 \cdot (1 - 0.999698) \\
&\approx 0.0006
\end{aligned}$$

P-Wert bei einseitigem Test (rechts):
Beispiel mit $c = 3.43$ als Grenze des Verwerfungsbereichs:

$$\begin{aligned}
P_{H_0}[T > 3.43] &= (1 - P_{H_0}[T \leq 3.43] \\
&= (1 - \Phi(3.43)) \\
&= (1 - 0.999698) \\
&\approx 0.0003
\end{aligned}$$

P-Wert bei einseitigem Test (links):
Beispiel mit $c = 3.43$ als Grenze des Verwerfungsbereichs:

$$\begin{aligned}
P_{H_0}[T < 3.43] &= \Phi(3.43) \\
&\approx 0.999698)
\end{aligned}$$

12.3 Likelihood-Quotienten-Test (Neyman-Pearson-Lemma

Du erhältst den Auftrag, die Anzahl Ausfälle eines Systems zu überprüfen. Nach Angaben des Herstellers sollen erwartungsgemäss 0.5 Ausfälle/h eintreten. Gehe davon aus, dass die Anzahl Ausfälle poisson-verteilt mit unbekanntem Parameter λ ist und dass die einzelnen Ausfälle unabhängig voneinander sind. Die Analyse nach 6 Betriebsstunden hat nun folgendes Ergebnis gebracht:

Betriebsstunde i	1	2	3	4	5	6
Anz. Ausfälle X_i	1	0	1	1	2	1

Aufgrund der hohen Anzahl Ausfälle haben wir den Verdacht, dass λ grösser als die vom Hersteller angegebenen Anzahl Ausfälle ist. Prüfe anhand eines einseitigen Tests auf dem Niveau 2.5%, ob tatsächlich $\lambda = 0.5$ Ausfälle/h angenommen werden kann. Gib

1. das Modell,
2. die Nullhypothese,
3. die Alternativhypothese,
4. die Teststatistik,
5. die Verteilung der Teststatistik unter der Nullhypothese,
6. den Verwerfungsbereich,
7. den beobachteten Wert der Teststatistik, sowie
8. den Testentscheid

an.

Lösung:

1. **Modell:** Unter P_λ sind die $X_i,\ i.i.d. \sim Pois(\lambda),\ i = 1, ..., 6,\ \lambda$ unbekannt.
2. **Nullhypothese:** $H_0: \ \lambda = \lambda_0 = 0.5$
3. **Alternativhypothese:** $H_A: \ \lambda = \lambda_A > \lambda_0$
4. **Teststatistik:** $T = \sum_{i=1}^{6} X_i$, denn

$$R(x_1, ..., x_6; \lambda_0, \lambda_A) = \frac{L(x_1, x_2, ..., x_6; \lambda_0)}{L(x_1, x_2, ..., x_6; \lambda_A)} =$$

$$\frac{e^{-6\lambda_0} \prod_{i=1}^{6} \frac{\lambda_0^{x_i}}{x_i!}}{e^{-6\lambda_A} \prod_{i=1}^{6} \frac{\lambda_A^{x_i}}{x_i!}} = e^{-6(\lambda_0 - \lambda_A)} \left(\frac{\lambda_0}{\lambda_A}\right)^{\sum_{i=1}^{6} x_i}$$

   Da $\lambda_0 < \lambda_A$ wird $R(x_1, ..., x_6; \lambda_0, \lambda_A)$ klein, genau dann, wenn $\sum_{i=1}^{6} x_i$ gross ist. Statt des komplizierten Quotienten wählen wir als Teststatistik also

$$T = \sum_{i=1}^{6} X_i$$

5. **Verteilung der Teststatistik unter H_0:** $T \sim Pois(6\lambda_0) = Pois(3)$
6. **Verwerfungsbereich:** Der kritische Bereich "Quotient klein" hat die äquivalente Form "Summe gross", also ist der Verwerfungsbereich von der Form $K = [k, \infty)$. Um das Signifikanzniveau einzuhalten, muss gelten $P_{\lambda_0}[T \geq k] \leq 2.5\% \Leftrightarrow P_{\lambda_0}[T < k] \geq 97.5\%$:

k	$P_{\lambda_0}[T = k]$	$P_{\lambda_0}[T \leq k]$
0	0.050	0.050
1	0.149	0.199
2	0.224	0.423
3	0.224	0.647
4	0.168	0.815
5	0.101	0.916
6	0.050	0.966
7	0.022	0.988

   Deshalb haben wir als Verwerfungsbereich $= [8, \infty)$.

7. **Beobachteter Wert der Teststatistik:** $t = 6$
8. **Testentscheid:** Da 6 nicht im Verwerfungsbereich liegt, wird die Nullhypothese **nicht** verworfen.

12.4 Erwartungstreuer Schätzer

Wir prüfen, ob der Schätzer $T = \frac{1}{n} \sum_{i=1}^{n} X_i$ erwartungstreu ist.

$$E[T] = E\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=0}^{n} E[X_i] = \frac{1}{n} \cdot n \cdot \mu = \mu$$

Der Schätzer T ist erwartungstreu.

12.5 Verteilung der Summe zweier normalerverteilter ZV

Seien $X \sim N(\mu_1, \sigma_1^2)$ und $Y \sim N(\mu_2, \sigma_2^2)$. Wie ist $Z = 1 + aX + bY$ verteilt?

Lösung:

$$\begin{aligned} Z &\sim N(1 + aE[X] + bE[Y], a^2 Var[X] + b^2 Var[Y]) \\ &\sim N(1 + a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2) \end{aligned}$$

Wahrscheinlichkeit und Statistik (BSc D-INFK)

1. (9 Punkte) Bei den folgenden 9 Fragen wird nur die Lösung verlangt. Der Lösungsweg wird nicht bewertet. Es gibt pro richtig beantwortete Frage 1 Punkt.

a) Seien A und B Ereignisse in einem Wahrscheinlichkeitsraum. Es gelte $P[A \cup B^c] = \frac{2}{3}$, $P[B] = \frac{2}{3}$ und $P[A \cap B^c] = \frac{1}{6}$. Berechnen Sie $P[A]$.

b) Sei die Zufallsvariable $X \sim \text{Binom}(n, \frac{1}{2})$ für ein $n \in \mathbb{N}$ und es gelte $P[X \leq 1] = \frac{5}{16}$. Bestimmen Sie $P[X \geq 3]$.

c) Seien X_1 und X_2 Zufallsvariablen mit

$$E[X_1] = 2,\ E[X_2] = 4,\ \text{Var}[X_1] = 12,\ \text{Var}[X_2] = 9,\ \text{Var}[1 + \frac{1}{2}X_1 + \frac{1}{3}X_2] = 3.$$

Bestimmen Sie $E[X_1 X_2]$.

d) Seien X_1 bzw. X_2 unabhängige Zufallsvariablen mit Verteilung $\mathcal{N}(\mu_1, \sigma_1^2)$ bzw. $\mathcal{N}(\mu_2, \sigma_2^2)$. Bestimmen Sie die Verteilung von $X_3 := 1 + aX_1 + bX_2$ für $a, b \in \mathbb{R}$.

e) Seien X_1, X_2 unabhängige, exponentialverteilte Zufallsvariablen mit Parameter $\lambda > 0$, d.h. die Dichte von X_i, $i \in \{1, 2\}$ ist gegeben durch $f_{X_i}(x) = \lambda e^{-\lambda x},\ x \geq 0$. Sei die Zufallsvariable Y gegeben als $Y := X_1 + X_2$. Bestimmen Sie die Verteilung von Y.

f) Seien nun $X_1, \ldots, X_n,\ n \in \mathbb{N}$ wie in Teilaufgabe **e)**, und seien mit $\bar{X}_n,\ n \in \mathbb{N}$ die Zufallsvariablen $\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ bezeichnet. Bestimmen Sie diejenige Verteilung, mit der sich die Verteilung von $\bar{X}_n$ für grosse n approximieren lässt.

g) Sei $X > 0$ eine Zufallsvariable mit Dichtefunktion f_X und sei $Y := \frac{1}{X}$. Bestimmen Sie die Verteilungsfunktion $F_Y(t)$ und die Dichtefunktion $f_Y(t)$, $t > 0$ von Y in Abhängigkeit von F_X und f_X.

h) Welcher Typ von Fehler kann bei einem statistischen Test vorliegen, wenn die Nullhypothese akzeptiert wurde?

i) Die Zufallsvariable X hat die Dichte

$$f_X(x) = \begin{cases} \theta e^{-\theta x}, & x \geq 0 \\ 0 & , \text{ sonst.} \end{cases}$$

Es liegen folgende Beobachtungen von X vor: $x_1 = 0.27, x_2 = 0.22, x_3 = 0.23, x_4 = 0.25, x_5 = 0.28$. Bestimmen Sie den Maximum-Likelihood Schätzer $\hat{\theta}(x_1, \ldots, x_5)$ für θ.

Bitte wenden!

2. (5 Punkte) Wir betrachten eine Eishockey-Play-Off-Serie der Kloten Flyers gegen die ZSC Lions. Die Serie besteht aus maximal drei Runden. Jede Runde hat einen eindeutigen Sieger, es gibt also keine Unentschieden. Die Mannschaft, die zuerst zwei Runden für sich entscheiden kann, gewinnt die Serie. Falls eine Mannschaft die Serie bereits nach zwei Runden gewonnen hat, wird keine dritte Runde mehr gespielt.

Im Folgenden bezeichne F_i für $i = 1, 2, 3$ das Ereignis, dass die Kloten Flyers die i-te Runde gewinnen. Ferner sei R die Gesamtzahl der gespielten Runden und X die Anzahl der Runden, die die Kloten Flyers gewinnen. Folgende Wahrscheinlichkeiten seien gegeben:

$$P[F_1] = \frac{1}{3},\ P[R = 2, X = 0] = \frac{1}{4},\ P[R = 2, X = 2] = \frac{1}{6},\ P[R = 3, X = 1] = \frac{1}{3}.$$

Beachten Sie: Setzen Sie erst am Ende Ihrer Berechungen konkrete Zahlen ein. Es muss klar werden, woher die Zahlen stammen.

a) Mit welcher Wahrscheinlichkeit gewinnen die Flyers die Serie in insgesamt drei Runden?

b) Berechnen Sie $P[F_2]$.

Wir nehmen nun an, die zwei Mannschaften treffen jede Spielsaison in einer Serie dieser Form aufeinander. Ausserdem nehmen wir an, dass die Ergebnisse der Serien unabhängig voneinander sind. Es bezeichne N die Saison (mit $N = 1$ falls die ZSC Lions in der ersten Saison gewinnen), in der die ZSC Lions zum ersten Mal eine Serie für sich entscheiden.

c) Benennen Sie die Verteilung von N und bestimmen Sie den/die zugehörigen Parameter. Berechnen Sie ausserdem den Erwartungswert $E[N]$.

3. (8 Punkte)

In einer Versicherung wird die Schadenshöhe eines Schadensfalls als eine Zufallsvariable S mit Dichte

$$f_S(x) = \begin{cases} 0 & \text{für } x < 0, \\ \frac{1}{2000} & \text{für } x \in [0, 1000], \\ \frac{c}{x^4} & \text{für } x > 1000, \end{cases}$$

modelliert, wobei c eine Konstante ist.

a) Wie muss man c wählen, damit f_S eine Dichte ist?

b) Berechnen Sie die Verteilungsfunktion der Schadenshöhe in diesem Modell.

c) Berechnen Sie die erwartete Schadenshöhe in diesem Modell. Berechnen Sie die Varianz.

d) Die Versicherung hat ein System, wobei Schadensfälle nach Schadenshöhe kategorisiert werden, und entsprechend unterschiedlich behandelt werden. Die erste Kategorie besteht aus Schadensfällen mit Schadenshöhen von weniger als 500 Franken, und die zweite Kategorie besteht aus Schadensfällen mit Schadenshöhen zwischen 500 und 2000 Franken. Wie viel Prozent der Schadensfällen gehören in diesem Modell im Durchschnitt zur *zweiten* Kategorie?

e) Die höchste Kategorie besteht aus den Schadensfällen, deren Schadenshöhen zu den 0.4 Prozent höchsten aller Schadenshöhen gehören. Wie hoch muss die Schadenshöhe (in diesem Modell) sein, damit ein Schadensfall zur höchsten Kategorie gehört?

Bitte wenden!

4. **(7 Punkte)** Die Post hat festgestellt, dass normalerweise 5% aller Sendungen auf dem Postweg verloren gehen. Der Online-Shop *Azamon.com* möchte diese Information benutzen, um betrügerische Kunden zu erkennen.

a) Welche Verteilung können wir benutzen, um die Anzahl X der verlorenen Pakete für einen Kunde, der n Bestellungen gemacht hat, zu modellieren?

b) *Shopper99* hat 15 Bestellungen gemacht und zwei von ihnen als "auf dem Postweg verloren gegangen" angezeigt. *Azamon.com* möchte testen, ob dieser Kunde betrügerisch ist.

i) Formulieren Sie eine geeignete Null- und Alternativ-Hypothese.

ii) Bestimmen Sie den Verwerfungsbereich K für X zum Niveau $\alpha = 5\%$. Ist *Shopper99* nach Auswertung dieses Tests als Betrüger anzusehen?

iii) Wie gross ist die Wahrscheinlichkeit, unter H_0, dass zwei oder mehr Pakete von 15 verloren gehen?

Hinweis: Verwenden Sie die folgende Binomial-Tabelle für $n = 15$ und $p = 0.05$:

k	0	1	2	3	4
$P(X \leq k)$	0.463	0.829	0.964	0.995	0.999

c) Der Einfachheit halber nehmen wir an, dass ein Test mit Signifikanzniveau von genau $\alpha = 5\%$ und Macht gleich $1 - \beta = 20\%$ zur Verfügung steht. Das Geschäft hat $N = 100$ Kunden getestet und festgestellt, dass nach diesem Test Y von ihnen als betrügerisch gelten. Wir möchten den Anteil f aller Kunden, die wirklich betrügerisch sind, abschätzen.

i) Was ist die Wahrscheinlichkeit f^* für einen Kunden als betrügerisch betrachtet zu werden? *(Schreiben Sie diese als Ausdruck von α, β und f.)*

ii) Geben Sie den Momenten-Schätzer $\hat{f}(Y)$ für f an.

iii) Berechnen Sie $\hat{f}(8)$.

Wahrscheinlichkeit und Statistik - Musterlösung (BSc D-INFK)

1. (9 Punkte)

a) $P[A] = 1/2$ **b)** $5/16$ **c)** $E[X_1 X_2] = 5$
d) $\mathcal{N}(1 + a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$ **e)** $\Gamma(2, \lambda)$ **f)** $\mathcal{N}(1/\lambda, 1/n\lambda^2)$
g) $F_Y(t) = 1 - F_X(\frac{1}{t})$, $f_Y(t) = \frac{1}{t^2} f_X(\frac{1}{t})$ **h)** Fehler 2. Art **i)** $\hat{\theta} = \frac{5}{\sum_{i=1}^5 x_i} = 4$

2. **a)** (1 Punkt) R nimmt nur die Werte $2, 3$ an, X nur die Werte $0, 1, 2$. Ausserdem ist auf Grund des Spielmodus $P[R = 2, X = 1] = 0$ und $P[R = 3, X = 0] = 0$. Daher gilt

$$\begin{aligned} P[R = 3, X = 2] &= 1 - P[R = 2, X = 0] - P[R = 2, X = 2] - P[R = 3, X = 1] \\ &= 1 - \frac{1}{4} - \frac{1}{6} - \frac{1}{3} = \frac{1}{4}. \end{aligned}$$

b) (2 Punkte) Mit $F_1 \cap F_2 = \{R = 2, X = 2\}$ und $F_1^c \cap F_2^c = \{R = 2, X = 0\}$ gilt

$$\begin{aligned} P[F_2|F_1] &= \frac{P[F_2 \cap F_1]}{P[F_1]} = \frac{1/6}{1/3} = \frac{1}{2}, \\ P[F_2^c|F_1^c] &= \frac{P[F_2^c \cap F_1^c]}{P[F_1^c]} = \frac{1/4}{2/3} = \frac{3}{8} \end{aligned}$$

und somit $P[F_2|F_1^c] = 1 - P[F_2^c|F_1^c] = \frac{5}{8}$. Mit dem Satz der totalen Wahrscheinlichkeit bekommt man schliesslich

$$P[F_2] = P[F_2|F_1]P[F_1] + P[F_2|F_1^c]P[F_1^c] = \frac{1}{2} \cdot \frac{1}{3} + \frac{5}{8} \cdot \frac{2}{3} = \frac{7}{12}.$$

c) (2 Punkte) N ist geometrisch verteilt mit Erfolgsparameter

$$p' = P[X \neq 2] = P[X = 0] + P[X = 1] = P[R = 2, X = 0] + P[R = 3, X = 1] = \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$$

und hat Erwartungswert

$$E[N] = \frac{1}{p'} = \frac{12}{7}.$$

Bitte wenden!

3. **a)** (1 Punkt) Wegen der Normierungsbedingung muss es gelten, dass

$$\int_{-\infty}^{\infty} f_S(x)dx = 1.$$

Wir haben

$$\begin{aligned}
\int_{-\infty}^{\infty} f_S(x)dx &= \int_0^{1000} \frac{1}{2000}dx + \int_{1000}^{\infty} cx^{-4}dx \\
&= \frac{1}{2} + c\int_{1000}^{\infty} x^{-4}dx \\
&= \frac{1}{2} + c(-\frac{x^{-3}}{3})|_{1000}^{\infty} \\
&= \frac{1}{2} + \frac{c}{3 \cdot 1000^3}.
\end{aligned}$$

Somit ist

$$c = (1 - \frac{1}{2}) \cdot 3 \cdot 1000^3 = \frac{3}{2}1000^3 (= 15 \cdot 10^8).$$

b) (1 Punkt) Die Verteilungsfunktion $F_S(x)$ ist definiert durch

$$F_S(x) = \int_{-\infty}^{x} f_S(y)dy.$$

Also haben wir $F_S(x) = 0$ für $x < 0$. Für $x \in [0, 1000]$ haben wir

$$F_S(x) = \int_0^x \frac{1}{2000}dy = \frac{x}{2000}.$$

Schliesslich für $x > 1000$

$$\begin{aligned}
F_S(x) &= \int_0^{1000} \frac{1}{2000}dx + c\int_{1000}^{x} y^{-4}dy \\
&= \frac{1}{2} + c\left(-\frac{y^{-3}}{3}\right)|_{1000}^{x} \\
&= \frac{1}{2} + \frac{c}{3}\left(\frac{1}{1000^3} - \frac{1}{x^3}\right) \\
&= \frac{1}{2} + \frac{\frac{3}{2}1000^3}{3}\left(\frac{1}{1000^3} - \frac{1}{x^3}\right) \\
&= 1 - \frac{1}{2}\left(\frac{1000}{x}\right)^3.
\end{aligned}$$

Somit ist

$$F_S(x) = \begin{cases} 0 & \text{für } x < 0, \\ \frac{x}{2000} & \text{für } x \in [0, 1000], \\ 1 - \frac{1}{2}\left(\frac{1000}{x}\right)^3 & \text{für } x > 1000. \end{cases}$$

Siehe nächstes Blatt!

c) (2 Punkte) Für den Erwartungswert gilt

$$\begin{aligned}
E[S] &= \int_{-\infty}^{\infty} x f_S(x) dx \\
&= \int_0^{1000} \frac{x}{2000} dx + \int_{1000}^{\infty} x \cdot c x^{-4} dx \\
&= \frac{1}{2000} \frac{x^2}{2} |_0^{1000} + c \int_{1000}^{\infty} x^{-3} dx \\
&= \frac{1000^2}{4000} + c\left(-\frac{x^{-2}}{2}\right) |_{1000}^{\infty} \\
&= 250 + \frac{c}{2 \cdot 1000^2} \\
&= 250 + \frac{\frac{3}{2} \cdot 1000^3}{2 \cdot 1000^2} \\
&= 1000.
\end{aligned}$$

Für die Varianz gilt

$$\begin{aligned}
\operatorname{Var}[S] &= E[S^2] - (E[S])^2 \\
&= E[S^2] - 10^6.
\end{aligned}$$

Wir haben

$$\begin{aligned}
E[S^2] &= \int_{-\infty}^{\infty} x^2 f_S(x) dx \\
&= \int_0^{1000} \frac{x^2}{2000} + \int_{1000}^{\infty} x^2 \cdot c x^{-4} dx \\
&= \frac{1}{2000} \frac{x^3}{3} |_0^{1000} + c \int_{1000}^{\infty} x^{-2} dx \\
&= \frac{1000^3}{3 \cdot 2000} + c(-x^{-1}) |_{1000}^{\infty} \\
&= \frac{1}{6} 10^6 + \frac{3}{2} \frac{1000^3}{1000} \\
&= \frac{10}{6} 10^6.
\end{aligned}$$

Also ist

$$\begin{aligned}
\operatorname{Var}[S] &= \frac{10}{6} 10^6 - 10^6 \\
&= \frac{2}{3} 10^6.
\end{aligned}$$

d) (2 Punkte) Der Anteil von Schadensfällen, die zur zweiten Kategorie gehören, beträgt

$$P[500 \leq S \leq 2000].$$

Bitte wenden!

Wir haben

$$
\begin{aligned}
P[500 \leq S \leq 2000] &= F_S(2000) - F_S(500) \\
&= 1 - \frac{1}{2}\left(\frac{1000}{2000}\right)^3 - \frac{500}{2000} \\
&= 1 - \frac{1}{16} - \frac{1}{4} \\
&= \frac{11}{16}.
\end{aligned}
$$

In Prozent sind es also im Durchschnitt

$$\frac{11}{16} \cdot 100\% = \frac{11}{4} \cdot 25\% = 68.75\%$$

von den Schadensfällen, die zur zweiten Kategorie gehören.

e) (2 Punkte) Der Anteil von Schadensfällen, deren Schadenshöhe x ubersteigt, ist $P[S \geq x]$. Also suchen wir ein x, so dass

$$P[S \geq x] = 0.004 = 4 \cdot 10^{-3}.$$

Es ist klar, dass $x \geq \frac{1}{2}$ (sonst ist $P[S \geq x] \geq \frac{1}{2} > 4 \cdot 10^{-3}$), und für $x \geq \frac{1}{2}$ haben wir

$$P[S \geq x] = 1 - F_S(x) = \frac{1}{2}\left(\frac{1000}{x}\right)^3.$$

Also brauchen wir

$$\frac{1}{2}\left(\frac{1000}{x}\right)^3 = 4 \cdot 10^{-3} \implies x = 5000.$$

Ein Schadensfall gehört zur höchsten Kategorie, wenn die Schadenshöhe mindestens 5000 ist.

Siehe nächstes Blatt!

4. **a)** (1 Punkt) Unter der Annahme, dass die Sendungen unabhängig sind, ist die Anzahl X der verlorenen Pakete Binom$(n, 0.05)$-verteilt.

b) (3 Punkte)

i) H_0: $p = 0.05$ (nicht betrügerisch)
H_A: $p > 0.05$ (betrügerisch)

ii) Aus der Tabelle

$$\mathbb{P}_{H_0}(X \geq 2) = 1 - \mathbb{P}_{H_0}(X \leq 1) = 0.171 > \alpha$$
$$\mathbb{P}_{H_0}(X \geq 3) = 0.036 \leq \alpha.$$

Also ist der Verwerfungsbereich zum Niveau $\alpha = 5\%$ gleich $K = [3, 15]$. Da $2 \notin K$, wird die Nullhypothese nicht verworfen und somit ist *Shopper99* nicht als Betrüger anzusehen.

iii) P-Wert $\mathbb{P}_{H_0}(X \geq 2) = 0.171$.

c) (3 Punkte) $Y \sim$ Binom$(100, f^*)$

i) Der Satz von der totalen Wahrscheinlichkeit ergibt

$$\begin{aligned} f^* =& \mathbb{P}(H_0 \text{ verworfen}) \\ =& \mathbb{P}(H_0 \text{ verworfen} | \text{ nicht betrügerisch}) \cdot \mathbb{P}(\text{nicht betrügerisch}) \\ &+ \mathbb{P}(H_0 \text{ verworfen} | \text{ betrügerisch}) \cdot \mathbb{P}(\text{betrügerisch}) \\ =& \alpha(1-f) + (1-\beta)f \end{aligned}$$

ii) $\mathbb{E}[Y] = Nf^* = N(\alpha + f(1 - \beta - \alpha))$. Also ist der Momenten-Schätzer für f

$$\hat{f}(Y) = \frac{Y/N - \alpha}{1 - \beta - \alpha}.$$

iii)

$$\hat{f}(8) = \frac{8/100 - 0.05}{0.20 - 0.05} = \frac{0.03}{0.15} = \frac{1}{5} = 20\%$$

Wahrscheinlichkeit und Statistik (BSc D-INFK)

1. (7 Punkte)

Bei den folgenden 7 Fragen wird nur die Lösung verlangt. Der Lösungsweg wird nicht bewertet. Es gibt pro richtig beantwortete Frage 1 Punkt.

a) Es seien $P[B] = 3P[A]$, $P[A \cup B] = \frac{5}{6}$ und $P[A \cap B] = \frac{1}{6}$. Wie gross ist $P[A^c]$?

b) Sei

$$f(x) = \begin{cases} \frac{c}{x} & \text{falls } 1 \leq x \leq 2, \\ 0 & \text{sonst.} \end{cases}$$

Wie ist c zu wählen, damit f eine Wahrscheinlichkeitsdichte darstellt?

c) Wir nehmen an, es seien drei (faire) Münzen geworfen worden. Wir wissen, dass mindestens zwei Würfe Kopf ergeben haben. Was ist dann die Wahrscheinlichkeit, dass alle drei Würfe Kopf ergeben haben?

d) Es seien X, Y Zufallsvariablen mit $E[X] = 5$, $E[Y] = 1$ und $E[(X+1)(Y+2)] = 20$. Bestimmen Sie $\operatorname{Cov}(X, Y)$.

e) Für gewisses $n \in \mathbb{N}$ seien $X_1, \ldots, X_n$ unabhängige, identisch verteilte Zufallsvariablen. X_1 habe Verteilungsfunktion F. Bestimmen Sie die Verteilungsfunktion der Zufallsvariablen $Y = \max\{X_1, \ldots, X_n\}$ in Abhängigkeit von F.

f) Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit $E[X_i] = \mu$ und $\operatorname{Var}(X_i) = \sigma^2$ für alle i. Wir definieren

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \text{für } n \in \mathbb{N}.$$

Durch welche Verteilung kann man die Verteilung von Y_n für "grosse n" approximieren?

g) Für gewisses $\vartheta > 1$ habe die Zufallsvariable X die Dichte

$$f_\vartheta(x) = \begin{cases} (\vartheta - 1)x^{-\vartheta} & \text{falls } x \geq 1, \\ 0 & \text{sonst.} \end{cases}$$

Es seien n Beobachtungen $x_1, \ldots, x_n$ von X gegeben. Bestimmen Sie den Maximum-Likelihood Schätzer für ϑ.

Bitte wenden!

2. **(9 Punkte)** Student Christoph fährt jeden Tag mit dem Velo an die ETH. Dabei kommt er an genau zwei Kreuzungen mit Ampeln vorbei. Aus Erfahrung weiss Christoph, dass er an der ersten Ampel mit Wahrscheinlichkeit 25% warten muss. Auf Grund der Ampelsteuerung und Christophs Fahrgeschwindigkeit hängt die Wahrscheinlichkeit, dass Christoph an der zweiten Ampel warten muss davon ab, ob er an der ersten Ampel warten musste oder nicht. Falls er an der ersten Ampel nicht warten musste, ist die Wahrscheinlichkeit an der zweiten auch nicht warten zu müssen $\frac{2}{3}$. Hatte er an der ersten Ampel hingegen "rot", muss er mit Wahrscheinlichkeit $\frac{1}{7}$ auch an der zweiten Ampel warten. Christoph kommt genau dann zu spät zur Vorlesung, wenn er an wenigstens einer Ampel warten muss.

a) Mit welcher Wahrscheinlichkeit muss Christoph an genau einer Ampel warten?

b) Mit welcher Wahrscheinlichkeit kommt Christoph zu spät zur Vorlesung?

c) Mit welcher Wahrscheinlichkeit hatte Christoph an der ersten Ampel "grün", gegeben, dass er an der zweiten Ampel nicht warten musste?

Hinweis: Falls Sie Teil **b)** nicht gelöst haben, benutzen Sie im Folgenden für die Wahrscheinlichkeit, dass Christoph zu spät zur Vorlesung kommt den Wert $\frac{1}{2}$.

d) Mit welcher Wahrscheinlichkeit musste Christoph an der ersten Ampel warten, gegeben, dass er zu spät zur Vorlesung kam? Sind die Ereignisse "Christoph kommt zu spät zur Vorlesung" und "Christoph muss an der ersten Ampel warten" unabhängig?

e) Christoph fährt im Februar 10-mal an die ETH. Nehmen Sie an, die Ereignisse, dass er an einem der 10 Tage zu spät zur Vorlesung kommt, seien unabhängig. Wie gross ist die Wahrscheinlichkeit, dass er im Februar mindestens zweimal zu spät zur Vorlesung kommt?

Hinweis: Potenzen wie beispielsweise 2^{10} müssen nicht weiter vereinfacht werden.

3. (12 Punkte)

Wir betrachten einen Eisstand am Zürichsee im späten Frühling.

a) Der Stand öffnet um 10 Uhr. An einem Tag, an dem die Temperatur t Grad beträgt, ist die Zeit Z_1 (in Stunden), bis der erste Kunde kommt, exponentialverteilt mit Parameter t. Was ist die Dichte von Z_1? Was ist die Verteilungsfunktion von Z_1?

b) Z_1 ist die Zeit *in Stunden*, bis der erste Kunde kommt. Sei M dieselbe Zeit, aber in Minuten. Welche Verteilung hat M?

c) Der Stand ist nur in Betrieb, falls die Temperatur zumindest 20 Grad ist. An einem Tag, an dem dies eintrifft, bezeichen wir die Temperatur mit T. Wir nehmen an, dass T eine Zufallsvariable mit folgender Dichte ist.

$$f_T(t) = \begin{cases} \frac{c}{t} & 20 \leq t \leq 30 \\ 0 & \text{sonst} \end{cases}, \tag{1}$$

wobei c eine Konstante ist. Berechne c. (Falls die Antwort so weit wie möglich vereinfacht ist, muss sie nicht explizit ausgerechnet werden.)

d) Was ist die erwartete Temperatur? Was ist die Varianz? (Falls die Antworten so weit wie möglich vereinfacht sind, müssen sie nicht explizit ausgerechnet werden.)

e) In dieser Teilaufgabe und der nächsten, nehmen wir an, dass die Temperatur T die Dichte von *d*) hat, und, dass die Wartezeit Z_1 *gegeben die Temperatur* $t = T$, die Dichte in (a) hat. Was ist in diesem Fall die gemeinsame Dichte von T und Z_1?

f) Wenn T und Z_1 diese gemeinsame Dichte haben, was ist die Randdichte von Z_1? (Die Randdichte ist nicht gleich der Dichte von (a), weil wir hier eine zufällige Temperatur betrachten.)

g) Jetzt betrachten wir wieder den Fall, dass die Temperatur *nicht* zufällig ist, sondern gleich t ist für t konstant. Also hat Z_1 wieder die Dichte von (a). Wir bezeichnen die Zeit, nachdem der erste Kunde gegangen ist, bis der zweite Kunde kommt, mit Z_2. Wir nehmen an, dass auch Z_2 exponentialverteilt ist mit Paramter t. Ferner nehmen wir an, dass Z_1 und Z_2 unabhängig sind. Bezeichne mit X die Zeit, in Stunden nach 10 Uhr, bis der zweite Kunde ankommt. Was ist die Dichte von X?

Bitte wenden!

4. **(9 Punkte)** Ihre Webseite wird auf einem entfernten Server gehostet. Der Host behauptet, dass die durchschnittliche Antwortzeit des Servers gleich $400ms$ ist. Da das Signal von Ihrem Haus zum Server über k Router gesendet wird, modellieren wir die Antwortzeit X mit der Gamma(k, θ) Verteilung. Die entsprechende Dichte ist:

$$f(x) = \frac{x^{k-1}}{(k-1)!} \cdot \frac{e^{-x/\theta}}{\theta^k}$$

a) Sei $X_i \overset{iid}{=} X$.

i) Bestimmen Sie die Verteilung von $Y = \sum_{i=1}^n X_i$.

ii) Berechnen Sie $\mathbb{E}[X]$ und $\text{Var}(X)$

Hinweis: Falls $X = \sum_{j=1}^k E_j$, wobei $E_j \overset{iid}{\sim} \text{EXP}(1/\theta)$, $j = 1, \ldots, k$, dann gilt $X \sim \text{Gamma}(k, \theta)$.

b) Sie haben zehn Experimente durchgeführt, und die folgenden Ergebnisse (in ms) erhalten:

i	1	2	3	4	5	6	7	8	9	10
x_i	700	244	696	495	1164	864	113	427	510	1211

Kennzahl: $\bar{x} = 642.4$

Schätzen Sie θ mittels der Maximum-Likelihood-Methode wenn $k = 4$ gegeben ist. Geben sie dazu sowohl den Schätzer $\hat{\theta}(\underline{X})$ wie auch den Schätzwert $\hat{\theta}(\underline{x})$ an.

c) Warum ist es vernünftig anzunehmen, dass $\hat{\theta}(\underline{X})$ näherungsweise normalverteilt ist? Begründen Sie kurz.

d) Sie vermuten, dass die durchschnittliche Antwortzeit des Servers von Ihrem Computer eigentlich länger als $400ms$ ist.

i) Formulieren Sie eine geeignete Null- und Alternativ-Hypothese.

ii) Berechnen Sie der Verwerfungsbereich für $\hat{\theta}$ zum Signifikanzniveau $\alpha = 5\%$. Wird die Null-Hypothese auf diesem Niveau abgelehnt?

iii) Berechnen Sie die ungefähre Macht des Tests unter der Annahme, dass die wahre durchschnittliche Antwortzeit $600ms$ beträgt.

Hinweis: $50/\sqrt{10} = 15.8$ und $2\sqrt{10}/25 = 1.01$. Benutzen Sie $\hat{\theta} \approx$ normalverteilt.

Wahrscheinlichkeit und Statistik - Musterlösung (BSc D-INFK)

1. (7 Punkte) **a)** $\frac{3}{4}$ **b)** $1/\log(2)$ **c)** $\frac{1}{4}$ **d)** 2 **e)** F^n **f)** $\mathcal{N}\Big(\mu, \frac{\sigma^2}{n}\Big)$ **g)** $1 + \frac{n}{\sum_{i=1}^n \log(x_i)}$

2. Wir definieren $W_i :=$ "Christoph muss an der i-ten Ampel warten" für $i = 1, 2$. Aus der Aufgabenstellung erhält man $P[W_1] = \frac{1}{4}$, $P[W_2^c \mid W_1^c] = \frac{2}{3}$, $P[W_2 \mid W_1] = \frac{1}{7}$.

a) (2 Punkte) Die gesuchte Wahrscheinlichkeit ist $P[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)]$. Wir berechnen

$$\begin{aligned} P[W_1 \cap W_2^c] &= P[W_2^c \mid W_1]P[W_1] = (1 - P[W_2 \mid W_1])P[W_1] \\ &= (1 - \frac{1}{7}) \cdot \frac{1}{4} = \frac{6}{7} \cdot \frac{1}{4} = \frac{6}{28} \end{aligned}$$

und

$$\begin{aligned} P[W_1^c \cap W_2] &= P[W_2 \mid W_1^c]P[W_1^c] = (1 - P[W_2 \mid W_1^c])(1 - P[W_1]) \\ &= (1 - \frac{2}{3}) \cdot (1 - \frac{1}{4}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}. \end{aligned}$$

Da $W_1 \cap W_2^c$ und $W_1^c \cap W_2$ disjunkt sind haben wir also

$$P[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)] = P[W_1 \cap W_2^c] + P[W_1^c \cap W_2] = \frac{6}{28} + \frac{1}{4} = \frac{13}{28}.$$

b) (1 Punkt) $S := W_1 \cup W_2$ beschreibt das Ereignis, dass Christoph zu spät zur Vorlesung kommt.

$$\begin{aligned} P[S] &= 1 - P[S^c] = 1 - P[W_1^c \cap W_2^c] \\ &= 1 - P[W_2^c \mid W_1^c]P[W_1^c] = 1 - \frac{2}{3} \cdot (1 - \frac{1}{4}) = \frac{1}{2}. \end{aligned}$$

c) (2 Punkte) Mit dem Satz von Bayes erhält man

$$\begin{aligned} P[W_1^c \mid W_2^c] &= \frac{P[W_2^c \mid W_1^c]P[W_1^c]}{P[W_2^c \mid W_1^c]P[W_1^c] + P[W_2^c \mid W_1]P[W_1]} \\ &= \frac{\frac{2}{3} \cdot (1 - \frac{1}{4})}{\frac{2}{3} \cdot (1 - \frac{1}{4}) + (1 - \frac{1}{7}) \cdot \frac{1}{4}} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{6}{28}} = \frac{\frac{1}{2}}{\frac{20}{28}} = \frac{7}{10}. \end{aligned}$$

Bitte wenden!

d) (2 Punkte) Mit der Definition der bedingten Wahrscheinlichkeit folgt

$$\begin{aligned} P[W_1 \mid S] &= \frac{P[W_1 \cap S]}{P[S]} = \frac{P[W_1 \cap (W_1 \cup W_2)]}{P[S]} = \frac{P[W_1]}{P[S]} \\ &= \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}. \end{aligned}$$

Da $P[W_1 \mid S] = \frac{1}{2} \neq \frac{1}{4} = P[W_1]$, sind die Ereignisse W_1 und S abhängig.

e) (2 Punkte) Sei E_i, $i = 1, \dots, 10$, das Ereignis, dass Christoph am i-ten Tag zu spät zur Vorlesung kommt. Wir wissen $P[E_i] = P[S] = \frac{1}{2}$. Da die E_i unabhängig sind, ist die Anzahl der Verspätungen X binomialverteilt mit Parametern $n = 10$ und $p = \frac{1}{2}$.

$$\begin{aligned} P[X \geq 2] &= 1 - P[X = 1] - P[X = 0] = 1 - \binom{10}{1}\left(\frac{1}{2}\right)^{10} - \left(\frac{1}{2}\right)^{10} \\ &= 1 - \frac{11}{2^{10}} \left(= \frac{2^{10} - 11}{2^{10}}\right). \end{aligned}$$

3. **a)** (1 Punkt)

$$f_{Z_1}(z) = te^{-tz}$$
$$P[Z_1 \leq 1 - z] = 1 - e^{-tz}.$$

b) (2 Punkte)

$$P[M \leq z] = P[60Z_1 \leq z] = P[Z_1 \leq \frac{z}{60}] = 1 - e^{-\frac{tz}{60}} \implies Exp(\frac{t}{60}).$$

c) (1 Punkt)

$$\int_{20}^{30} \frac{1}{t} dt = \ln t|_{20}^{30} = \ln \frac{30}{20} = \ln \frac{3}{2}$$
$$c = \frac{1}{\ln \frac{3}{2}}.$$

d) (2 Punkte)

$$E[T] = \frac{1}{\ln \frac{3}{2}} \int_{20}^{30} \frac{1}{t} t dt = \frac{10}{\ln \frac{3}{2}}.$$
$$E[T^2] = \frac{1}{\ln \frac{3}{2}} \int_{20}^{30} t dt = \frac{1}{\ln \frac{3}{2}} \frac{t^2}{2}|_{20}^{30} = \frac{30^2 - 20^2}{2 \ln \frac{3}{2}} = \frac{250}{\ln \frac{3}{2}}.$$
$$Var[T] = \frac{250}{\ln \frac{3}{2}} - \frac{100}{(\ln \frac{3}{2})^2}.$$

Siehe nächstes Blatt!

e) (2 Punkte)

$$f_{T,Z_1}(t,z) = \frac{1}{\ln\frac{3}{2}} 1_{[20,30]}(t) \cdot e^{-tz} 1_{[0,\infty)}(z).$$

f) (2 Punkte)

$$f_{Z_1}(z) = \frac{1}{\ln\frac{3}{2}} \int_{20}^{30} e^{-tz} dz = \frac{1}{\ln\frac{3}{2}} \frac{e^{-tz}}{z}\Big|_{20}^{30} = \frac{e^{-20z} - e^{-30z}}{z\ln\frac{3}{2}} \qquad \text{für } z \geq 0$$

g) (2 Punkte)

$$\begin{aligned} f_X(x) &= \\ (f_{Z_1} * f_{Z_1})(x) &= \\ t^2 \int_0^x e^{-t(x-u)} e^{-tu} du &= \\ t^2 \int_0^x e^{-tx} du &= \\ t^2 x e^{-tx} & \qquad \text{für } x \geq 0 \end{aligned}$$

4. a) (3 Punkte)

i) $Y = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \sum_{j=1}^{k} E_{ij} \sim \text{Gamma}(nk, \theta)$, wobei $E_{ij} \overset{\text{iid}}{\sim} \text{EXP}(1/\theta)$.

ii) Falls $E_j \sim \text{EXP}(1/\theta)$, dann $\mathbb{E}[E_j] = \theta$ und $\text{Var}(E_j) = \theta^2$.
Also $\mathbb{E}[X] \overset{\text{lin.}}{=} \sum_{j=1}^{k} \mathbb{E}[E_j] = k\theta$ und $\text{Var}(X) \overset{\text{indep.}}{=} \sum_{j=1}^{k} \text{Var}(E_j) = k\theta^2$

b) (2 Punkte)

$$\begin{aligned} L(x_1, \ldots, x_n; \theta) &= \prod_{i=1}^{n} f(x_i; \theta) = \frac{(\prod x_i)^{k-1}}{((k-1)!)^n} \cdot \frac{e^{-\sum x_i/\theta}}{\theta^{nk}} \\ l(x_1, \ldots, x_n; \theta) &= (k-1) \sum_{i=1}^{n} \log x_i - n \sum_{j=1}^{k-1} \log j - \sum_{i=1}^{n} x_i/\theta - nk \log\theta \\ \frac{dl}{d\theta} &= \sum_{i=1}^{n} x_i/\theta^2 - nk/\theta \overset{!}{=} 0 \\ \hat{\theta} &= \sum_{i=1}^{n} x_i/nk = \bar{x}/k \end{aligned}$$

Wenn $k = 4$, dann $\hat{\theta}(\underline{X}) = \bar{X}/4$ und $\hat{\theta}(\underline{x}) = \bar{x}/4 = 160.6$.

c) (1 Punkt) Da $\hat{\theta} = Y/nk$ der Mittelwert von 40 iid exponentialverteilten Zufallsvariablen ist, können wir mit Hilfe des zentralen Grenzwertsatzes annehmen, dass $\hat{\theta}$ näherungsweise normalverteilt ist.

Bitte wenden!

d) (3 Punkte)
i)

$$H_0 : \theta k \leq 400, \text{ d.h. } \theta \leq 100.$$
$$H_A : \theta k > 400, \text{ d.h. } \theta > 100.$$

ii) ZGS $\Rightarrow \hat{\theta}(\underline{\mathrm{X}}) \approx \mathrm{N}(\theta, \frac{n \cdot k\theta^2}{n^2k^2}) = \mathrm{N}(\theta, \frac{\theta^2}{nk})$. Unter H_0, $\theta = 100$.
Aus der Tabelle der Standardnormalverteilung, $\Phi(0.95)^{-1} = 1.64$.
Der Verwerfungsbereich ist $K = (\mu + 1.64\sigma, \infty) = (100 + 1.64 \cdot 15.8, \infty) = (126, \infty)$.
Da $\hat{\theta}(\underline{\mathrm{x}}) = 160.6 \in K$, wird die Null-Hypothese abgelehnt.
iii) Falls $\theta = 600/4 = 150$,

$$\mathbb{P}(\hat{\theta} \in K) \approx \bar{\Phi}\left(\frac{126 - 150}{150/\sqrt{40}}\right) = \bar{\Phi}(-1.01) = \Phi(1.01) = 0.84$$

Also, wenn die wahre durchschnittliche Antwortzeit $600ms$ ist, wird die Null-Hypothese mit Wahrscheinlichkeit 84% (=Macht des Tests) abgelenht.

Übungsserie 13

1. Untenstehend findest du mehrere Beispiele für Vergleiche von 2 Stichproben. Beantworte für jedes Beispiel kurz die folgenden Fragen:

- Handelt es sich um gepaarte oder um ungepaarte Stichproben? Begründe!
- Ist der Test einseitig oder zweiseitig durchzuführen? Begründe!
- Wie lautet die Nullhypothese in Worten?
- Wie lautet die Alternativhypothese in Worten?

a) In einem Experiment sollte der Effekt von Zigarettenrauchen auf Blutplättchenanhäufungen untersucht werden. Dazu wurden 11 Probanden vor und nach dem Rauchen einer Zigarette Blutproben entnommen, und es wurde gemessen, wie stark sich die Blutplättchen anhäuften. Es interessiert, ob sich Blutplättchen durch das Rauchen vermehrt anhäufen.

b) Die nächsten Daten sind aus einer Studie von Charles Darwin über die Fremd- und Selbstbefruchtung. 15 Paare von Setzlingen mit demselben Alter wurden gezüchtet. Dabei wurde bei jedem Paar je einer durch Selbst- und einer durch Fremdbefruchtung produziert. Beide Teile je eines Paares hatten nahezu gleiche Bedingungen. Das Ziel bestand darin zu sehen, ob die fremdbefruchteten Pflanzen mehr Lebenskraft besitzen als die selbstbefruchteten (d.h., ob sie grösser werden). Es wurden die Höhen jeder Pflanze nach einer fixen Zeitspanne gemessen.

c) Beeinflusst der Kalziumgehalt in der Nahrung den systolischen Blutdruck? Zur Überprüfung dieser Frage wurde einer Versuchsgruppe von 10 Männern während 12 Wochen ein Kalziumzusatz verabreicht. Einer Kontrollgruppe von 11 Männern gab man ein Placebopräparat.

d) In einem Experiment wurde untersucht, ob Mäuse zwei Formen von Eisen (Fe^{2+} und Fe^{3+}) unterschiedlich gut aufnehmen. Dazu wurden 36 Mäuse in zwei Gruppen zu je 18 unterteilt und die eine Gruppe mit Fe^{2+} und die andere mit Fe^{3+} "gefütter". Da das Eisen radioaktiv markiert war, konnte sowohl die Anfangskonzentration wie auch die Konzentration einige Zeit später gemessen werden. Daraus wurde für jede Maus der Anteil des aufgenommenen Eisens berechnet.

Bitte wenden!

2. Die Geschwindigkeit von einem alten und einem neu gekauften PC soll miteinander verglichen werden. 10 zufällig ausgewählte Arbeiten (computer jobs) werden mit beiden Maschinen einmal durchgeführt und die nötigen Ausführungszeiten in Sekunden gemessen.

Man findet (mit A werden die Ausführungszeiten des alten PC bezeichnet, mit N die Ausführungszeiten des neuen PC):

i	1	2	3	4	5	6	7	8	9	10
A_i	20	30	25	50	10	42	27	34	14	45
N_i	14	19	16	32	6	28	78	23	9	32
$D_i = N_i - A_i$	-6	-11	-9	-18	-4	-14	51	-11	-5	-13

Die aus den Daten berechneten Mittelwerte und Standardabweichungen sind:

$\overline{a} = 29.7$; $\quad \overline{n} = 25.7$; $\quad \overline{d} = -4$
$s_A = 13.22$; $\quad s_N = 20.47$; $\quad s_D = 19.80$.

Man möchte nun untersuchen, ob der neue PC generell schneller als der alte ist.

a) Es handelt sich hier um 2 gepaarte Stichproben. Wieso? Begründe *kurz*.

b) Führe einen t-Test auf dem 5%-Niveau durch.

3. Die durchschnittliche Fahrzeit von Zürich nach Bellinzona mit einem Intercity Zug beträgt 146 Minuten. Mit dem Cisalpino werden die folgenden Zeiten gemessen:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
152	145	141	137	145	146	139	147	138

Wir nehmen an, dass diese Werte unabhängige Realisierungen einer normalverteilten Zufallsvariable mit unbekanntem Erwartungswert μ und unbekannter Varianz σ^2 sind.

a) Teste die Hypothese $H_0 : \mu = 146$ gegen die Alternative $H_A : \mu \neq 146$ auf dem 5%-Niveau.

b) Berechne das 2-seitige 95%-Vertrauensintervall für den wahren Erwartungswert μ.

Bei einem neuen Zug misst man folgende Zeiten:

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
150	140	138	137	145	146	139	147	136

Man möchte diese Werte mit denen des Cisalpinos vergleichen.

c) Handelt es sich um einen gepaarten oder um einen ungepaarten Vergleich (Begründung)?

d) Führe den entsprechenden t-Test für die Nullhypothese “die erwartete Zeit ist gleich für die beiden Züge” auf dem 5%-Niveau durch.

Siehe nächstes Blatt!

Kennzahlen: $\overline{x}_9 = 143.33$, $\overline{y}_9 = 142$, $s_x^2 = 24.25$, $s_y^2 = 25.5$.

4. Ein Kunde hat den Eindruck, dass die Brote einer Bäckerei im Mittel zu leicht sind (Sollgewicht 1000 g). Er kauft sich 7 Brote und wiegt sie: 993, 974, 1008, 1014, 969, 986 und 979 g. Die Gewichte dürfen als unabhängig und normalverteilt angenommen werden.

a) Üblicherweise haben die Gewichte von 1 kg - Broten eine Standardabweichung von 15 g. Bestimme ein 90 %-Vertrauensintervall für den wahren Mittelwert der Brote dieser Bäckerei.

b) Der Inspektor zweifelt daran, dass die Gewichte der Brote dieser Bäckerei die übliche Streuung haben. Bestimme deshalb noch ein 90%-Vertrauensintervall für den wahren Mittelwert bei unbekannter Streuung.

Lösungsskizze Serie 13

1. a) Gepaarte Stichprobe: Zu jeder Blutplättchenmenge vor dem Rauchen gehört die Blutplättchenmenge der selben Person nach dem Rauchen.
Einseitiger Test: Wir wollen nicht wissen, ob sich die Bluplättchenmenge verändert hat, sondern ob sie sich erhöht hat.

H_0 : Rauchen hat keinen Einfluss auf die Anhäufung der Blutplättchen. ($\mu_R = \mu_{NR}$)
H_A: Durch Rauchen erhöht sich die Anhäufung der Blutplättchen. ($\mu_R > \mu_{NR}$)

b) Gepaarte Stichprobe: Zu jeder Höhe eines selbstbefruchteten Setzlings gehört die Höhe des fremdbefruchteten "Partners".
Einseitiger Test: Wir wollen nicht wissen, ob sich die Höhen unterscheiden, sondern ob die fremdbefruchteten Setzlinge grösser werden als die selbstbefruchteten.

H_0 : Die Höhen unterscheiden sich nicht. ($\mu_f = \mu_s$)
H_A: Fremdbefruchtete Setzlinge werden grösser als selbstbefruchtete. ($\mu_f > \mu_s$)

c) Ungepaarte Stichprobe: Ungleiche Anzahl in den Gruppen. Zu einem Blutdruck aus der Versuchsgruppe gehört nicht ein spezifischer aus der Kontrollgruppe.
Zweiseitiger Test: Wir wollen nur wissen, ob das Kalzium einen Einfluss hat auf den Blutdruck, egal ob nach oben oder unten.

H_0: Kalzium hat keinen Einfluss auf den Blutdruck. ($\mu_{Kalz} = \mu_{Kontr}$)
H_A: Kalzium hat einen Einfluss auf den Blutdruck. ($\mu_{Kalz} \neq \mu_{Kontr}$)

d) Ungepaarte Stichprobe: Die Anzahlen in den beiden Gruppen brauchen nicht gleich zu sein. Zur Eisenmessung einer "Fe^{2+}-Maus" gehört nicht eine bestimmte Messung einer "Fe^{3+}-Maus".
Zweiseitiger Test: Wir wollen nur wissen, ob die Mäuse die verschiedenen Eisenformen unterschiedlich gut aufnehmen.

H_0: Die Eisenaufnahme ist von der Form unabhängig. ($\mu_2 = \mu_3$)
H_A: Die Eisenaufnahme ist von der Form abhängig. ($\mu_2 \neq \mu_3$)

2. a) Gepaart, denn man soll Paare von Typ $(A_i, N_i), i = 1, \ldots, n = 10$ betrachten.

Bitte wenden!

b) Der t-Test wird folgendermassen formal durchgeführt:

Modellannahmen	:	D_i: $i-$te Differenz, $i = 1, \ldots, 10$. $D_i \overset{iid}{\sim} N(\mu, \sigma^2)$ mit σ unbekannt.
Nullhypothese H_0	:	$\mu = \mu_0 = 0$
Alternative H_A	:	$\mu < 0$ (einseitig)
Teststatistik	:	$T = \frac{\overline{D} - \mu_0}{S_D/\sqrt{n}}$
Verwerfungsbereich	:	V.B. $= \{T < t_{n-1,\alpha} = t_{9,0.05} = -1.833\}$.
Beobachtung	:	$T(w) = \sqrt{10} \cdot (-4)/19.80 = -0.6388$.
Testentscheid	:	$T(w) = -0.6388 \notin$ V.B.: die Nullhypothese wird beibehalten.

3. **a)** Seien $X_1, \ldots, X_n$ iid $\sim \mathcal{N}(\mu, \sigma^2)$. Null- und Alternativhypothese lauten:

$$H_0: \quad \mu = 146,$$
$$H_A: \quad \mu \neq 146.$$

Die Teststatistik

$$T = \frac{\overline{X}_n - 146}{S_X/\sqrt{n}}$$

ist unter H_0 t-verteilt mit $n-1$ Freiheitsgraden. Auf dem Signifikanzniveau $\alpha = 5\%$ wird die Nullhypothese verworfen, falls $|T| > t_{n-1,1-\alpha/2}$, wobei $t_{n-1,1-\alpha/2}$ das $1 - \alpha/2$-Quantil der t-Verteilung mit $(n-1)$ Freiheitsgraden ist. Für die obigen Daten haben wir

$$|T(w)| = \left| \frac{143.33 - 146}{\sqrt{24.25}/3} \right| = 1.63 < t_{8,0.975} = 2.306.$$

Die Nullhypothese kann somit nicht verworfen werden.

b) Das Vertrauensintervall zum Niveau $1 - \alpha$ besteht aus allen Parameterwerten μ, bei denen der Test zum Niveau α nicht verwirft. Gesucht sind also alle Werte von μ so dass:

$$-t_{n-1,1-\alpha/2} \leq \frac{\overline{X}_n - \mu}{S_X/\sqrt{n}} \leq t_{n-1,1-\alpha/2}.$$

Also

$$C(X_1, \ldots, X_n) = \left[\overline{X}_n - t_{n-1,1-\alpha/2} \frac{S_X}{\sqrt{n}}, \overline{X}_n + t_{n-1,1-\alpha/2} \frac{S_X}{\sqrt{n}} \right],$$

und aus den Daten erhalten wir $C(x_1, \ldots, x_n) = [139.54, 147.12]$.

c) Ungepaarte Stichproben. Zur Messung eines neuen Zugs gehört nicht die Messung eines bestimmten alten Zugs.

Siehe nächstes Blatt!

d) Da die Stichproben ungepaart sind, muss man einen 2-Stichproben-t-Test durchführen. Die zwei Stichproben seien durch $X_1, \ldots, X_n \sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_1, \ldots, Y_n \sim \mathcal{N}(\mu_Y, \sigma^2)$ gegeben. Null- und Alternativhypothese lauten:

$$\begin{aligned} H_0: &\quad \mu_X = \mu_Y, \\ H_A: &\quad \mu_X \neq \mu_Y. \end{aligned}$$

Die Teststatistik

$$T = \frac{\overline{X}_n - \overline{Y}_n}{S_{pool}\sqrt{\frac{2}{n}}},$$

mit $S^2_{Pool} = \frac{1}{2n-2}\left((n-1)S^2_X + (n-1)S^2_Y\right) = \frac{1}{2}(S^2_X + S^2_Y)$, ist unter H_0 t-verteilt mit $2n-2$ Freiheitsgraden. Auf dem Signifikanzniveau $\alpha = 5\%$ wird die Nullhypothese verworfen, falls $|T| > t_{2n-2,1-\alpha/2}$, wobei $t_{2n-2,1-\alpha/2}$ das $1-\alpha/2$-Quantil der t-Verteilung mit $2n-2$ Freiheitsgraden ist. In diesem Fall haben wir

$$|T(w)| = 0.57 < t_{16,0.975} = 2.12.$$

Die Nullhypothese kann somit nicht verworfen werden.

4. Modell: X_i= Gewicht von Brot i, $i = 1, \ldots, 7 = n$, unabhängig und normalverteilt $\mathcal{N}(\mu, \sigma^2)$.

$$\Longrightarrow \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1),$$

mit $\overline{x}_n = 989$.

a) $\sigma = 15$.
Für ein 90 %-Vertrauensintervall für μ muss gelten:

$$\begin{aligned} &P\Big[\Big|\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma}\Big| \le q_{0.95}\Big] = 0.90 \text{ mit } q_{0.95} = 1.645 \\ &\Longrightarrow \sqrt{n}\Big|\frac{\overline{X}_n - \mu}{\sigma}\Big| \le 1.645 \\ &\Longrightarrow \overline{X}_n - 1.645 \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}_n + 1.645 \cdot \frac{\sigma}{\sqrt{n}}. \end{aligned}$$

Also erhalten wir das Konfidenzintervall: $[979.7; 998.3]$.

b) Ersetze σ durch $s_X = 17.01$. Dann

$$\sqrt{n}\frac{\overline{X}_n - \mu}{S_X} \sim t_{n-1,0.95}, \text{ mit } n = 7.$$

Aus der Tabelle für die t-Verteilung lesen wir ab, dass $t_{n-1,0.95} = 1.943$. Daher finden wir das Konfidenzintervall

$$[\overline{X}_n - 1.943 \cdot \frac{S_X}{\sqrt{n}}, \overline{X}_n + 1.943 \cdot \frac{S_X}{\sqrt{n}}].$$

Mit den Daten erhalten wir $[976.5; 1001.5]$.

Übungsserie 11

1. In einer Gruppe verheirateter Männer (keiner von ihnen war jemals verwitwet) ist ein gewisser Anteil $0 < p < 1$ mindestens einmal geschieden. Mit Hilfe der folgenden Prozedur soll p geschätzt werden: Sei $x \in (0, 1)$ bekannt. Eine Urne enthält hundert Umschläge. $100x$ von ihnen enthalten die Frage: "Wurden Sie jemals geschieden?" Die anderen $100(1 - x)$ enthalten die Frage: "Ist Ihre momentane Ehe die erste?"
Beachten Sie, dass eine Person die zweite Frage notwendig mit "nein" beantworten muss, wenn sie die erste mit "ja" beantwortet. Wir nehmen an, dass jeder nur wahrheitsgemässe Antworten gibt.
N verheiratete Männer werden nun zufällig aus der Gruppe ausgewählt. Nacheinander wird jeder von ihnen gebeten, einen Umschlag aus der Urne zu nehmen, seinen Inhalt zu lesen und ihn dann wieder in die Urne zurückzulegen. Er muss dann die Frage, die er gelesen hat, laut mit "ja" oder "nein" beantworten. Also kennt nur er allein die Frage, die in dem Umschlag enthalten war. Es bezeichne Y die Anzahl derjenigen Männer, die mit "ja" geantwortet haben.

a) Bestimme die Verteilung von Y.

b) Bestimme den Momentenschätzer von p aus der Zufallsvariablen Y für den Fall, dass $x \neq 1/2$.

c) Berechne den Erwartungswert und die Varianz des Schätzers aus **b)**. Gibt es einen Wert von x, so dass die Varianz des Schätzers minimal ist?

Bemerkung: Diese Aufgabe beschreibt eine Stichprobenprozedur namens *randomized response*. Es handelt sich dabei um eine Technik, die nützlich ist, wenn es darum geht, Informationen über heikle Themen zu erhalten.

2. Unter P_ϑ seien $X_1, \ldots, X_n$ iid $\sim \mathcal{U}(0, \vartheta)$. Wir betrachten den Schätzer

$$T^{(n)} = \max(X_1, \ldots, X_n)$$

für ϑ.

a) Bestimme die Verteilungsfunktion von $T^{(n)}$ unter P_ϑ.

b) Ist die Folge von Schätzern $T^{(n)}$, $n \in \mathbb{N}$, konsistent für ϑ?

Bitte wenden!

Lösungsskizze Serie 11

1. a) Sei X_i die Zufallsvariable, welche angibt, ob der i-te Mann mit "ja" ($X_i = 1$) oder "nein" ($X_i = 0$) antwortet. Mit dem Gesetz der totalen Wahrscheinlichkeit folgt

$$P[X_i = 1] = px + (1-p)(1-x) = p(2x-1) + 1 - x =: p^*.$$

Dieses Experiment wird N-mal unabhängig, identisch durchgeführt, also ist $Y = \sum_{i=1}^N X_i$ binomialverteilt mit Erfolgswahrscheinlichkeit p^*:
$Y \sim \mathrm{Bin}(N, p^*)$, $E[Y] = Np^*$ und $\mathrm{Var}[Y] = Np^*(1-p^*)$.

b) Das erste theoretische Moment ist $E[Y] = Np^* = N(p(2x-1) + 1 - x)$. Also ist für $x \neq 1/2$ der Momentenschätzer von p

$$\hat{p} = \frac{Y/N - 1 + x}{2x - 1}.$$

c) Aus den Eigenschaften von Erwartungswert und Varianz erhalten wir

$$E[\hat{p}] = \frac{E[Y]/N - 1 + x}{2x - 1} = p$$

(der Schätzer ist erwartungstreu). Analog:

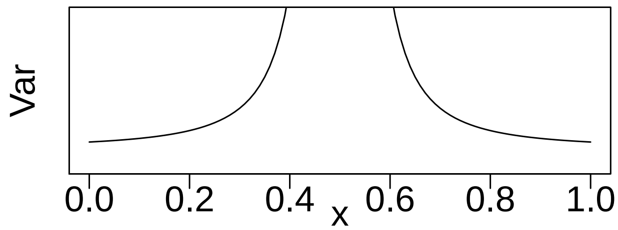
$$\begin{aligned}\mathrm{Var}[\hat{p}] &= \frac{\mathrm{Var}[Y]/N^2}{(2x-1)^2} = \frac{p^*(1-p^*)/N}{(2x-1)^2} = \frac{(p(2x-1)+1-x)(x-p(2x-1)}{N(2x-1)^2} \\ &= \frac{(2x-1)^2(p-p^2) + x(1-x)}{N(2x-1)^2} = \frac{p(1-p)}{N} + \frac{x(1-x)}{N(2x-1)^2}.\end{aligned}$$

Um die Extrema zu finden, leitet man nach x ab:

$$\begin{aligned}\frac{d}{dx}\mathrm{Var}[\hat{p}] &= \frac{(1-2x)(2x-1)^2 N - 2N(2x-1)2(x-x^2)}{N^2(2x-1)^4} \\ &= \frac{-(2x-1)^2 - 4(x-x^2)}{N(2x-1)^3} = -\frac{1}{N(2x-1)^3}.\end{aligned}$$

D.h. die Varianz steigt monoton auf $(0, 0.5)$ und fällt monoton auf $(0.5, 1)$, nimmt also Extrema nur in den Randpunkten $0, 1$ (die nicht zugelassen sind) an. Je näher an 0.5, desto grösser die Varianz. Man muss also zwischen Genauigkeit (gewöhnliche Umfrage mit $x = 0$ oder $x = 1$) und Anonymität ($x = 0.5$) abwägen.

Bitte wenden!



2. **a)** Unter P_ϑ erhalten wir für die Verteilungsfunktion von $T^{(n)}$

$$\begin{aligned} F_{T^{(n)}}(x) &= P_\vartheta[T^{(n)} \leq x] = P_\vartheta[\cap_{i=1}^n X_i \leq x] = \prod_{i=1}^n P_\vartheta[X_i \leq x] \\ &= \left(\frac{x}{\vartheta}\right)^n \mathbb{1}\{x \in [0, \vartheta]\}, \quad x \leq \vartheta, \end{aligned}$$

und $F_{T^{(n)}}(x) = 1$ für $x \geq \vartheta$.

b) Mit der Verteilungsfunktion aus **a)** folgt für jedes $\varepsilon > 0$

$$\begin{aligned} P_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] &= 1 - P_\vartheta[|T^{(n)} - \vartheta| \leq \varepsilon] \\ &= 1 - P_\vartheta[-\varepsilon \leq T^{(n)} - \vartheta \leq \varepsilon] \\ &= 1 - \left(P_\vartheta[T^{(n)} - \vartheta \leq \varepsilon] - P_\vartheta[T^{(n)} - \vartheta \leq -\varepsilon]\right) \\ &= 1 - \left(1 - \left(\frac{\vartheta - \varepsilon}{\vartheta}\right)^n \mathbb{1}_{\{(\vartheta - \varepsilon) \in [0, \vartheta]\}}\right) \\ &= \left(\frac{\vartheta - \varepsilon}{\vartheta}\right)^n \mathbb{1}_{\{(\vartheta - \varepsilon) \in [0, \vartheta]\}} \xrightarrow{n \to \infty} 0. \end{aligned}$$

Somit ist die Konsistenz bewiesen.

3. Im Modell P_p seien $X_1, \ldots, X_n$ iid Bernoulli(p)-verteilte Zufallsvariablen mit der Interpretation, dass $X_i = 1$, falls der i-te Wurf Kopf ist und 0 sonst, $i = 1, \ldots, n = 100$. Sei $S_n := \sum_{i=1}^n X_i$ die Anzahl Köpfe in den n Würfen. Nach Voraussetzung ist S_n binomialverteilt mit Parametern n und p. Einseitiger Test:

$$\begin{aligned} H_0 &: p = p_0 = 1/2, \\ H_A &: p > p_0. \end{aligned}$$

Siehe nächstes Blatt!

Die Nullhypothese möchten wir verwerfen, falls zu viele Köpfe geworfen werden. Wir wählen daher einen Verwerfungsbereich der Form $K = (c, 100]$. Der Verwerfungsbereich ist durch die Ungleichung

$$0.01 \geq P_{p_0}[S_n > c] \approx 1 - \Phi\left(\frac{c - 50}{5}\right)$$

bestimmt. Also $c \geq 61.6$ und wir erhalten den Verwerfungsbereich $K_{1\%} = [62, 100]$. Damit der Test zum 1% Niveau die Annnahme einer fairen Münze nicht verwirft, darf höchstens 61 mal Kopf fallen. In unserem Fall wird die Nullhypothese zum Signifikanzniveau 1% nicht verworfen.