Variante **A**

Statistik

Technische Universität Dortmund | Fakultät Wirtschaftswissenschaften | 20. März 2024

Bitte tragen sie ihre Daten sorgfältig und leserlich ein:

1. Prüfungsversuch ☐ Ja Nein ☐

Matrikelnummer ☐ ☐ ☐ ☐ ☐ ☐ Nachname ____________________

Studiengang ____________________ Vorname ____________________

Bearbeitungshinweise:

Diese Klausur besteht aus 10 Aufgaben. Alle Aufgaben sind unabhängig voneinander lösbar.

Jede Aufgabe hat vier Antwortmöglichkeiten, von denen jeweils genau eine zutreffend ist.

Markieren sie die jeweils zutreffende Antwortmöglichkeit jeder Aufgabe auf diesem Deckblatt. Es werden ausschließlich ihre Markierungen der jeweiligen Antwortmöglichkeiten a) bis d) auf diesem Deckblatt gewertet. Skizzen und Anmerkungen werden nicht bewertet.

Sie dürfen entweder eine oder zwei Antwortmöglichkeiten für jede Aufgabe markieren.

Markieren sie genau eine Antwortmöglichkeit, so erhalten sie bei Markierung der zutreffenden Antwortmöglichkeit drei Punkte für die entsprechende Aufgabe.

Markieren sie genau zwei Antwortmöglichkeiten, so erhalten sie bei Markierung der zutreffenden Antwortmöglichkeit einen Punkt für die entsprechende Aufgabe.

In allen anderen Fällen erhalten sie null Punkte für diese Aufgabe.

Bitte verwenden sie einen Kugelschreiber oder nicht zu starken Filzstift.

Taschenrechner sind gemäß Liste als Hilfsmittel zugelassen.

Bei 18 von maximal 30 erreichbaren Punkten ist die Klausur in jedem Fall bestanden.

Die Bearbeitungszeit beträgt 60 Minuten.

Viel Erfolg!

Markierung: ☒ Korrektur: ■

Korrekturhinweis: Wenn sie irrtümlich ein falsches Kästchen angekreuzt haben, malen sie dieses bitte vollständig aus und kreuzen sie eindeutig erkennbar die zutreffende Antwort an.

	a)	b)	c)	d)		a)	b)	c)	d)		a)	b)	c)	d)
Aufgabe 1	☐	☐	☐	☐	Aufgabe 5	☐	☐	☐	☐	Aufgabe 9	☐	☐	☐	☐
Aufgabe 2	☐	☐	☐	☐	Aufgabe 6	☐	☐	☐	☐	Aufgabe 10	☐	☐	☐	☐
Aufgabe 3	☐	☐	☐	☐	Aufgabe 7	☐	☐	☐	☐					
Aufgabe 4	☐	☐	☐	☐	Aufgabe 8	☐	☐	☐	☐					

Aufgabe 1 (zu Teil 1 deskriptive Statistik)

Laut Bundesnetzagentur hat im Jahr 2023 die Windkraft an Land („Wind Onshore") einen Anteil von etwa 26,4% an der Gesamterzeugung von Energie in Deutschland.

Welcher Winkel entfällt auf „Wind Onshore" bei einer Darstellung mittels eines Kreissektorendiagramms?

Die möglichen Antworten wurden auf ganze Zahlen gerundet.

a) 95°

b) 75°

c) 100°

d) 26°

Aufgabe 2 (zu Teil 1 deskriptive Statistik)

In Deutschland wurden im Jahr 2023 insgesamt 745 Windenergieanlagen (WEA) aufgestellt. Diese verteilen sich auf die Bundesländer laut folgender Tabelle (Quelle: Deutsche WindGuard):

Bundesland	#WEA	Bundesland	# WEA
Baden-Württemberg	15	Niedersachsen	131
Bayern	7	Nordrhein-Westfalen	114
Berlin	0	Rheinland-Pfalz	33
Brandenburg	77	Saarland	6
Bremen	1	Sachsen	10
Hamburg	1	Sachsen-Anhalt	17
Hessen	37	Schleswig-Holstein	249
Mecklenburg-Vorpommern	41	Thüringen	6

Die zugrundeliegende Stichprobe lautet $(x_1, x_2, \ldots, x_{16})$, wobei i das Bundesland bezeichnet und x_i die Anzahl der WEA dieses Bundeslandes.

Wie lauten der Modalwert oder die Modalwerte dieser Stichprobe?

a) 1 und 6

b) 249

c) Schleswig-Holstein

d) 16

Aufgabe 3 (zu Teil 1 deskriptive Statistik)

Betrachtet seien wieder die Daten aus Aufgabe 2:

Bundesland	#WEA	Bundesland	#WEA
Baden-Württemberg	15	Niedersachsen	131
Bayern	7	Nordrhein-Westfalen	114
Berlin	0	Rheinland-Pfalz	33
Brandenburg	77	Saarland	6
Bremen	1	Sachsen	10
Hamburg	1	Sachsen-Anhalt	17
Hessen	37	Schleswig-Holstein	249
Mecklenburg-Vorpommern	41	Thüringen	6

Wie groß ist der Anteil der Bundesländer, die im Jahr 2023 weniger als 50 Windenergieanlagen (WEA) aufgestellt haben?

a) 75%

b) 50%

c) 25%

d) 87,5%

Aufgabe 4 (zu Teil 1 deskriptive Statistik)

Laut Bundesverband WindEnergie betrugen in Deutschland die installierten Windenergieleistungen im Jahr 2022 58.106 Megawatt und im Jahr 2023 61.010 Megawatt. Um wieviel Prozent ist die installierte Windenergieleistung von 2022 auf 2023 gestiegen?

Die Prozentzahlen sind auf ganze Zahlen gerundet.

a) 5%

b) 6%

c) 7%

d) 8%

Aufgabe 5 (zu Teil 2 Wahrscheinlichkeitsrechnung)

Die Ereignisse A und B seien in Bezug auf einen bestimmten Tag folgendermaßen beschrieben:

- A: Es scheint mindestens vier Stunden die Sonne.
- B: Es herrscht den gesamten Tag Windstille.

Die Wahrscheinlichkeit, dass A oder B auftreten lautet $\mathbb{P}(A \cup B) = 55\%$ und die Wahrscheinlichkeit, dass A und B auftreten lautet $\mathbb{P}(A \cap B) = 5\%$.

Wie lautet die Summe der Wahrscheinlichkeiten der einzelnen Ereignisse $\mathbb{P}(A) + \mathbb{P}(B)$?

a) $\mathbb{P}(A) + \mathbb{P}(B) = 60\%$

b) $\mathbb{P}(A) + \mathbb{P}(B) = 50\%$

c) $\mathbb{P}(A) + \mathbb{P}(B) = 5\%$

d) $\mathbb{P}(A) + \mathbb{P}(B) = 11\%$

Aufgabe 6 (zu Teil 2 Wahrscheinlichkeitsrechnung)

In der näheren Umgebung von Dortmund stehen 10 Windenergieanlagen (WEA), welche mit der identischen Wahrscheinlichkeit $p = 1,554\%$ unabhängig von einander an einem Tag einen Betriebsschaden aufweisen.

Wie groß ist die Wahrscheinlichkeit, dass an einem Tag nicht mehr als eine WEA einen Schaden aufweist?

Die Prozentzahlen sind auf ganze Zahlen gerundet.

a) 99%

b) 97%

c) 95%

d) 90%

Aufgabe 7 (zu Teil 2 Wahrscheinlichkeitsrechnung)

Familie Tovolta erwägt, Solarzellen auf dem Dach ihres Hauses zu installieren. Sie glaubt, dass sich dies wirtschaftlich rentieren würde, wenn mindestens 1992 Stunden im Jahr die Sonne auf ihr Dach scheint.

Die Anzahl der Sonnenstunden S (pro Jahr) sei normalverteilt mit Erwartungswert $\mu = 2012$ und Varianz $\sigma^2 = 256$.

Wie hoch ist die Wahrscheinlichkeit, dass mindestens 1992 Stunden im Jahr die Sonne scheint?

Die Prozentzahlen sind auf ganze Zahlen gerundet.

a) $\mathbb{P}(S \geq 1992) = 89\%$

b) $\mathbb{P}(S \geq 1992) = 84\%$

c) $\mathbb{P}(S \geq 1992) = 95\%$

d) $\mathbb{P}(S \geq 1992) = 99\%$

Aufgabe 8 (zu Teil 2 Wahrscheinlichkeitsrechnung)

Betrachtet seien wieder die Ereignisse:

- A: Es scheint mindestens vier Stunden die Sonne.
- B: Es herrscht den gesamten Tag Windstille.

Es gelte wieder $\mathbb{P}(A \cup B) = 55\%$ und $\mathbb{P}(A \cap B) = 5\%$.

Zusätzlich sei bekannt, dass die Wahrscheinlichkeit, dass den gesamten Tag Windstille herrscht, $\mathbb{P}(B) = 10\%$ beträgt.

Wie groß ist die Wahrscheinlichkeit, dass an einem Tag mindestens vier Stunden die Sonne scheint, gegeben dass den gesamten Tag Windstille herrscht?

a) $\mathbb{P}(A|B) = 50\%$

b) $\mathbb{P}(A|B) = 60\%$

c) $\mathbb{P}(A|B) = 11\%$

d) $\mathbb{P}(A|B) = 5\%$

Aufgabe 9 (zu Teil 3 Inferenzstatistik)

Familie Tovolta aus Aufgabe 7 kennt den Parameter μ der Verteilung der Sonnenstunden S pro Jahr nicht, weiß aber, dass S normalverteilt ist mit Varianz $\sigma^2 = 256$.

Für neun zufällig ausgewählte Jahre liest die Familie die Sonnenstunden aus einer Datenbank ab und berechnet das arithmetische Mittel $\bar{x}$ aus dieser Stichprobe. Familie Tovolta stellt die Nullhypothese $H_0 : \mu \leq 1992$ auf und möchte einen Gaußtest zum Signifikanzniveau $\alpha = 5\%$ durchführen. Der benötigte kritische Wert der Normalverteilung lautet $z_{95\%} = 1,645$. Nehmen Sie an, dass die zufälligen Werte identisch und unabhängig verteilt sind.

Für welche Werte $\bar{x}$ sollte die Familie Nullhypothese ablehnen?

Die möglichen Antworten wurden auf ganze Zahlen gerundet.

a) Lehne H_0 genau dann ab, falls $\bar{x} > 2001$.

b) Lehne H_0 genau dann ab, falls $\bar{x} > 1992$.

c) Lehne H_0 genau dann ab, falls $\bar{x} > 2024$.

d) Lehne H_0 genau dann ab, falls $\bar{x} > 1983$.

Aufgabe 10 (zu Teil 3 Inferenzstatistik)

Familie Tovolta möchte nun die in Aufgabe 9 erhobene Stichprobe $x_1, x_2, \ldots, x_9$ benutzen, um den unbekannten Parameter μ zu schätzen.

Wieder sei bekannt, dass die Zufallsvariablen $X_1, X_2, \ldots, X_9$ unabhängig und identisch verteilt sind mit Varianz $\sigma^2 = 256$.

Der Familie stehen nun vier Schätzer zur Verfügung:

- $\hat{\mu}_1 = X_1 - 2X_2 + 3X_3 - 4X_4 + 5X_5 - 6X_6 + 7X_7 - 8X_8 + 9X_9$
- $\hat{\mu}_2 = \frac{1}{9}\left(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9\right)$
- $\hat{\mu}_3 = X_1 - 2X_2$
- $\hat{\mu}_4 = \frac{1}{3}\left(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9\right)$

Welcher dieser Schätzer ist erwartungstreu?

a) $\hat{\mu}_2$ ist erwartungstreu, es gilt $\mathbb{E}[\hat{\mu}_2] = \mu$ für alle μ.

b) $\hat{\mu}_1$ ist erwartungstreu, es gilt $\mathbb{E}[\hat{\mu}_1] = \mu$ für alle μ.

c) $\hat{\mu}_3$ ist erwartungstreu, es gilt $\mathbb{E}[\hat{\mu}_3] = \mu$ für alle μ.

d) $\hat{\mu}_4$ ist erwartungstreu, es gilt $\mathbb{E}[\hat{\mu}_4] = \mu$ für alle μ.

Formelsammlung zu Teil 1 deskriptive Statistik

Eindimensionale Daten (Kapitel 3)

n: Stichprobenumfang, $x_1, \ldots, x_n$ Stichprobenwerte, $x_{(1)} \leq \ldots \leq x_{(n)}$ geordnete Stichprobenwerte
$a_j,\ j = 1, \ldots, k$: Ausprägungen, $h(a_j)$ bzw. $f(a_j) = \frac{h(a_j)}{n}$: absolute bzw. relative Häufigkeit von a_j
$H(x) = \sum_{a_j \leq x} h(a_j)$ bzw. $F(x) = \sum_{a_j \leq x} f(a_j)$: absolute bzw. relative kumulierte Häufigkeitsverteilung

Lageparameter

Modalwert x_{Mod} (beliebige Skalierung): für Ausprägung $a = x_{Mod}$ gilt $h(x_{Mod}) \geq h(a_j)$ für alle $j = 1, \ldots, k$
Median x_{Med} (ordinale Skalierung): Für $a = x_{Med}$ gilt $\sum_{a_j \leq x_{Med}} f(a_j) \geq \frac{1}{2}$ und $\sum_{a_j \geq x_{Med}} f(a_j) \geq \frac{1}{2}$.
Arithmetisches Mittel (kardinale Skalierung): $\bar{x} = \sum_{j=1}^{k} a_j f(a_j) = \frac{1}{n} \sum_{i=1}^{n} x_i$
Geometrisches Mittel x_{Geom} (kardinale Skalierung): $x_{Geom} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots x_n}$ (nur für $x_i \geq 0,\ i = 1, \ldots, n$)

Streuungsmaße (alle: kardinale Skalierung)

Spannweite: $SP = \max_i x_i - \min_i x_i$
Durchschnittliche Abweichung von $\lambda \in \mathbb{R}$: $\bar{s} = \frac{1}{n} \sum_{i=1}^{n} |x_i - \lambda|$
Mittlere quadratische Abweichung $s^2 = \sum_{j=1}^{k} (a_j - \bar{x})^2 \cdot f(a_j) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$
Standardabweichung $s = \sqrt{s^2}$

Verschiebungssatz:

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

Konzentrationsmaße (alle: kardinale Skalierung)

Variationskoeffizient: $V = s/\bar{x}$
Lorenzkurve: Polygonzug durch (u_k, v_k), $k = 0, \ldots, n$ mit $u_0 = v_0 = 0$, $u_k = \frac{k}{n}$, $v_k = \frac{\sum_{i=1}^{k} x_{(i)}}{\sum_{i=1}^{n} x_i}$ für $k = 1, \ldots, n$
Gini-Koeffizient: $G = 2 \frac{\sum_{i=1}^{n} i \cdot x_i}{n \sum_{i=1}^{n} x_i} - \frac{n+1}{n} = \sum_{k=2}^{n} u_{k-1} v_k - u_k v_{k-1}$
Herfindahl-Index: $H = \sum_{i=1}^{n} \left(\frac{x_i}{\sum_{k=1}^{n} x_k} \right)^2$

Mehrdimensionale Daten (Kapitel 4)

Merkmale X und Y mit Ausprägungen $a_1, \ldots, a_k$ und $b_1, \ldots, b_l$
Abs. Häufigkeit von (a_i, b_j): $h_{ij} = h(a_i, b_j)$, *Randhäufigkeit von* a_i *bzw.* b_j: $h_{i\bullet} = \sum_{j=1}^{l} h_{ij}$, $h_{\bullet j} = \sum_{i=1}^{k} h_{ij}$
Relative Häufigkeiten: $f_{ij} = \frac{1}{n} h_{ij}$, $f_{i\bullet} = \frac{1}{n} h_{i\bullet}$, $f_{\bullet j} = \frac{1}{n} h_{\bullet j}$
Bedingte relative Häufigkeiten: $f_1(a_i|b_j) = \frac{f_{ij}}{f_{\bullet j}} = \frac{h_{ij}}{h_{\bullet j}}$, $f_2(b_j|a_i) = \frac{f_{ij}}{f_{i\bullet}} = \frac{h_{ij}}{h_{i\bullet}}$
X *und* Y *unabhängig*: $f_1(a_i|b_j) = f_{i\bullet} \Leftrightarrow h_{ij} = \frac{h_{i\bullet} h_{\bullet j}}{n} \Leftrightarrow f_2(b_j|a_i) = f_{\bullet j}$ für alle i, j
Kovarianz von X *und* Y (kardinale Skalierung): $s_{xy} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$ $(= \overline{xy} - \bar{x}\bar{y})$
(Bravais-Pearson-)Korrelationskoeffizient von X *und* Y (kardinale Skalierung): $r_{xy} = \frac{s_{xy}}{s_x \cdot s_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$
Rangkorrelation von Spearman (ordinale Skalierung): $r_{xy}^{SP} = \frac{\sum_{i=1}^{n} (R(x_i) - \frac{n+1}{2})(R(y_i) - \frac{n+1}{2})}{\sqrt{\sum_{i=1}^{n} (R(x_i) - \frac{n+1}{2})^2} \cdot \sqrt{\sum_{i=1}^{n} (R(y_i) - \frac{n+1}{2})^2}}$,
wobei $R(x_i)$: Rang von x_i, $R(y_i)$: Rang von y_i
Kontingenzkoeffizient (beliebige Skalierung): $K = \sqrt{\frac{\chi^2}{n + \chi^2}}$ mit $\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(h_{ij} - \tilde{h}_{ij})^2}{\tilde{h}_{ij}}$, wobei $\tilde{h}_{ij} = \frac{h_{i\bullet} h_{\bullet j}}{n}$

Lineare Transformation: $y_i = a + b \cdot x_i,\ a, b \in \mathbb{R}$:

$$y_{Mod} = a + b \cdot x_{Mod},\ y_{Med} = a + b \cdot x_{Med},\ \bar{y} = a + b \cdot \bar{x},\ s_y^2 = b^2 \cdot s_x^2,\ s_y = |b| \cdot s_x,\ s_{zy} = b \cdot s_{zx},\ s_{xy} = \begin{cases} s_x \cdot s_y & , b \geq 0 \\ -s_x \cdot s_y & , b < 0 \end{cases}$$

Indexzahlen (Kapitel 5)

\# Güter: n, Preis p und Menge q von Gut i in Basis- bzw. Berichtsperiode: $p_0(i)$ und $q_0(i)$ bzw. $p_t(i)$ und $q_t(i)$
Preisindizes nach Laspeyres, Paasche & Fisher: $P_{0t}^L = \frac{\sum_{i=1}^{n} p_t(i) \cdot q_0(i)}{\sum_{i=1}^{n} p_0(i) \cdot q_0(i)}$, $P_{0t}^P = \frac{\sum_{i=1}^{n} p_t(i) \cdot q_t(i)}{\sum_{i=1}^{n} p_0(i) \cdot q_t(i)}$, $P_{0t}^F = \sqrt{P_{0t}^L \cdot P_{0t}^P}$
Mengenindizes nach Laspeyres, Paasche & Fisher: $Q_{0t}^L = \frac{\sum_{i=1}^{n} p_0(i) \cdot q_t(i)}{\sum_{i=1}^{n} p_0(i) \cdot q_0(i)}$, $Q_{0t}^P = \frac{\sum_{i=1}^{n} p_t(i) \cdot q_t(i)}{\sum_{i=1}^{n} p_t(i) \cdot q_0(i)}$, $Q_{0t}^F = \sqrt{P_{0t}^L \cdot P_{0t}^P}$

Formelsammlung zu Teil 2 Wahrscheinlichkeitsrechnung

ω: Elem.-ereignis, Ω: Ergebnismenge, $A \subset \Omega$: Ereignis, $\mathcal{A}$: Ereignissystem mit $A, B \in \mathcal{A} \Rightarrow \bar{A}, A \cap B, A \cup B \in \mathcal{A}$

Zufallsvorgänge, Ereignisse und Wahrscheinlichkeiten (Kapitel 7)

Wahrscheinlichkeit $\mathbb{P}: \mathcal{A} \to \mathbb{R}_{\geq}$ mit $\mathbb{P}(\Omega) = 1, \mathbb{P}(\emptyset) = 0, \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
Bedingte Wahrscheinlichkeit: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ (falls $\mathbb{P}(B) > 0$).
Satz der vollständigen Wahrscheinlichkeit: $\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$, wobei $A_1, \ldots, A_k$ Partition von Ω
Formel von Bayes: $\mathbb{P}(A_j|B) = \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j) / \sum_{i=1}^{k} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$
Ereignisse A und B sind stochastisch unabhängig, falls: $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Zufallsvariablen und Verteilungen (Kapitel 8)

Zufallsvariable: Funktion $X : \Omega \to \mathbb{R}$, *W'keit von Zufallsv.*: $\mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\})$ für $B \subseteq \mathbb{R}$
Verteilungsfunktion: $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) = \mathbb{P}(X \leq x)$ und $F(b) - F(a) = \mathbb{P}(a < X \leq b)$ für $a < b$.
Diskrete Zufallsvariable: Falls Wertebereich $\{x_1, x_2, \ldots\}$ abzählbar. $p_i = \mathbb{P}(X(\omega) = x_i)$
- *Wahrscheinlichkeitsfunktion* $f(x) = \mathbb{P}(X = x)$, *Verteilungsfunktion* $F(x) = \mathbb{P}(X \leq x)$
- *Gemeinsame W'keitsfunktion* $f(x, y) = \mathbb{P}(X = x, Y = y)$, *gem. V-funktion* $F(x, y) = \mathbb{P}(X \leq x, Y \leq y)$
- *Randwahrscheinlichkeiten*: $f_1(x) = \sum_{j=1}^{m} f(x, y_j)$, $f_2(y) = \sum_{i=1}^{k} f(x_i, y)$
- *Randverteilungen*: $F_1(x) = \sum_{x_i \leq x} f_1(x_i)$, $F_2(y) = \sum_{y_j \leq y} f_2(y_j)$

Stetige Zufallsvariable: Falls Wertebereich überabzählbar, z.B. $\mathbb{R}, \mathbb{R}_{\geq}$ oder $[a, b] \subset \mathbb{R}$ und $\mathbb{P}(X = x) = 0$.
- *Dichtefunktion* $f : \mathbb{R} \to \mathbb{R}$ mit $\int_{-\infty}^{\infty} f(x)dx = 1$ und $f(x) \geq 0$ für alle $x \in \mathbb{R}$.
- *Verteilungsfunktion* $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) = \int_{-\infty}^{x} f(t)dt = \mathbb{P}(X \leq x)$
- Gemeinsame Dichtefunktion $f(x, y)$, gemeinsame Verteilungsfunktion $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t, s)dsdt$
- *Randdichten*: $f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy$, $f_2(y) = \int_{-\infty}^{\infty} f(x, y)dx$
- *Randverteilungen*: $F_1(x) = \int_{-\infty}^{x} f_1(t)dt$, $F_2(y) = \int_{-\infty}^{y} f_2(s)ds$
- *Bedingte W'keits- bzw. Dichtefunktion*: $f_1(x|y) = \frac{f(x,y)}{f_2(y)}$, $f_2(y|x) = \frac{f(x,y)}{f_1(x)}$ (falls $f_1(x), f_2(y) > 0$)

Zufallsvariablen X und Y sind stochastisch unabhängig, falls: $F(x, y) = F_1(x) \cdot F_2(y)$ bzw. $f(x, y) = f_1(x) \cdot f_2(y)$

Verteilungsparameter (Kapitel 9)

α-Quantil x_α: $\mathbb{P}(X \geq x_\alpha) \geq 1 - \alpha$ und $\mathbb{P}(X \leq x_\alpha) \geq \alpha$. Falls X stetig und $f(x) > 0$ für alle x: $F(x_\alpha) = \alpha$
Erwartungswert: $\mathbb{E}[X] = \sum_i x_i f(x_i)$ (falls X diskret) bzw. $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)dx$ (falls X stetig)
Rechenregeln: $\mathbb{E}[a \cdot X + b \cdot Y + c] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y] + c$. Falls X, Y unabhängig: $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$
Bedingte Erwartungswerte: $\mathbb{E}[X|Y = y] = \sum_i x_i f_1(x_i|y)$ (diskret) bzw. $\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} x f_1(t|y)dt$ (stetig)
Varianz σ^2 und Standardabweichung σ: $\sigma^2 = Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, $\sigma = \sqrt{Var(X)}$
Kovarianz von X und Y: $Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$
Korrelation von X und Y: $\rho(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$
Rechenregeln: $Var(a + b \cdot X) = b^2 Var(X)$, $Var(a + b \cdot X + c \cdot Y) = b^2 Var(X) + 2bcCov(X, Y) + c^2 Var(Y)$
Ungleichung von Tschebyscheff: $\mathbb{P}(|X - \mathbb{E}[X]| \geq c) \leq \frac{Var(X)}{c^2}$

Einige wichtige Verteilungen

Bernoulliverteilung, $X \sim B(1, p)$: $\mathbb{P}(X = 1) = p$ und $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{E}[X] = p$, $Var(X) = p(1 - p)$
Binomialverteilung, $X \sim B(n, p)$: $X = \sum_{i=1}^{n} X_i$, wobei $X_i \overset{iid}{\sim} B(1, p)$, $\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$, $\mathbb{E}[X] = np$, $Var(X) = np(1 - p)$
Poissonverteilung, $X \sim P(\lambda)$: $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \cdot \exp(-\lambda)$ für $k = 0, 1, 2, \ldots$, $\mathbb{E}[X] = Var(X) = \lambda$
Gleichverteilung $X \sim U(a, b)$: $f(x) = \frac{1}{b-a}$, $F(x) = \frac{x-a}{b-a}$ für $a \leq x \leq b$, $\mathbb{E}[X] = \frac{a+b}{2}$, $Var(x) = \frac{(b-a)^2}{12}$
Normalverteilung $X \sim \mathcal{N}(\mu, \sigma^2)$: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $\mathbb{E}[X] = \mu$, $Var(X) = \sigma^2$
Standardnormalverteilung $Z \sim \mathcal{N}(0, 1)$ mit Dichte ϕ und Verteilungsfunktion Φ (siehe Anhang).
Für $X \sim \mathcal{N}(\mu, \sigma^2)$ gilt $F(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ (*Standardisierung*: $Z = \frac{X-\mu}{\sigma}$).

Gesetz der großen Zahlen und zentraler Grenzwertsatz (Kapitel 10)

X_i unabhängig und identisch verteilt (iid) mit $\mathbb{E}[X_i] = \mu$ und $Var(X_i) = \sigma^2 < \infty$, $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.
Schwaches Gesetz der großen Zahlen: $\lim_{n\to\infty} \mathbb{P}(|\bar{X}_n - \mu| \leq \epsilon) = 1$ für alle $\epsilon > 0$.
Zentraler Grenzwertsatz: $\lim_{n\to\infty} \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$

Formelsammlung zu Teil 3 Inferenzstatistik

Grundlagen der Induktiven Statistik (Kapitel 11) (mit $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$)

Wichtige Testverteilungen

χ^2-Verteilung mit n Freiheitsgraden: $C \sim \chi^2(n)$, $C = \sum_{i=1}^n X_i^2$ mit $X_i \overset{iid}{\sim} \mathcal{N}(0,1)$. $\mathbb{E}[C] = n$, $Var(C) = 2n$
t-Vert. mit n FG: $T = \frac{X}{\sqrt{C/n}} \sim t(n)$ mit $X \sim \mathcal{N}(0,1)$, $C \sim \chi^2(n)$, X, C unabh. $\mathbb{E}[T] = 0$, $Var(T) = \frac{n}{n-2}$.

Wichtige Stichprobenfunktionen

Stichprobenmittel: $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$, *Gauß-Statistik*: $G = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$
Mittlere quadratische Abweichung bzgl. μ: $M^2(\mu) = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$, $\mathbb{E}[M^2(\mu)] = \sigma^2$, $\frac{n}{\sigma^2}M^2(\mu) \sim \chi^2(n)$
Mittlere quadratische Abweichung: $M^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$, $\mathbb{E}[M^2] = \frac{n-1}{n}\sigma^2$
Stichprobenvarianz: $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, $\mathbb{E}[S^2] = \sigma^2$, $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$
t-Statistik: $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

Punktschätzung (Kapitel 12)

Schätzer für unbekannten Parameter θ: Funktion $\hat{\theta} : \mathbb{R}^n \to \Theta$ (mit Θ: Menge aller möglichen Parameter)
Erwartungstreue: $\mathbb{E}[\hat{\theta}] = \theta$ für alle $\theta \in \Theta$, *Effizienz*: $Var(\hat{\theta}) \leq Var(\tilde{\theta})$ für alle erwartungstreuen Schätzer $\tilde{\theta}$
Konsistenz: $\lim_{n\to\infty} \mathbb{P}\left(|\hat{\theta} - \theta| > \epsilon\right) = 0$ für alle $\epsilon > 0$.
Kleinste Quadrate Schätzer $\hat{\theta}^{LS}$ für $\theta = \mu$: $\hat{\theta}^{LS} = \bar{x}$ (minimiert $\sum_{i=1}^n (x_i - \hat{\theta})^2$. $\hat{\theta}^{LS} = \bar{x}$ ist e-treu und effizient.)

Intervall-Schätzung (Kapitel 13)

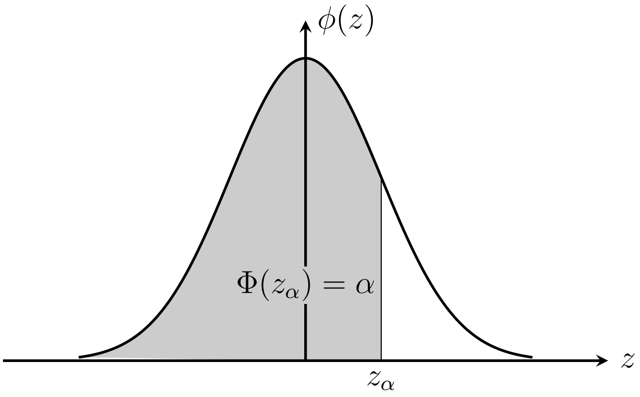
α-Quantil z_α für $Z \sim \mathcal{N}(0,1)$: $\Phi(z_\alpha) = \alpha$. Symmetrie $\Rightarrow \Phi(-z_\alpha) = 1 - \Phi(z_\alpha) = 1 - \alpha = \Phi(z_{1-\alpha})$
Einseitige Konfidenzintervalle für μ bei σ^2 bekannt: $\left(-\infty, \bar{X} + z_{1-\alpha} \cdot \sigma/\sqrt{n}\right]$ bzw. $\left(\bar{X} - z_{1-\alpha} \cdot \sigma/\sqrt{n}, \infty\right]$
Zweiseitiges Konfidenzintervall für μ bei σ^2 bekannt: $\left[\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}, \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sigma/\sqrt{n}\right]$
α-Quantil $t_{(n-1,\alpha)}$ für $T \sim t(n-1)$: $\mathbb{P}(T \leq t_{(n-1,\alpha)}) = \alpha$
Eins. Konfidenzintervalle für μ bei σ^2 unbekannt: $\left(-\infty, \bar{X} + t_{(n-1,1-\alpha)} \cdot S/\sqrt{n}\right]$ bzw. $\left(\bar{X} - t_{(n-1,1-\alpha)} \cdot S/\sqrt{n}, \infty\right]$
Zweiseitiges Konfidenzintervall für μ bei σ^2 unbekannt: $\left[\bar{X} - t_{(n-1,1-\frac{\alpha}{2})} \cdot S/\sqrt{n}, \bar{X} + t_{(n-1,1-\frac{\alpha}{2})} \cdot S/\sqrt{n}\right]$
α-Quantil $c_{(n-1,\alpha)}$ für $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$: $\mathbb{P}\left(\frac{n-1}{\sigma^2}S^2 \leq c_{(n-1,\alpha)}\right) = \alpha$
Zweiseitiges Konfidenzintervall für σ^2: $\left[\frac{n-1}{c_{(n-1,\frac{\alpha}{2})}}S^2, \frac{n-1}{c_{(n-1,1-\frac{\alpha}{2})}}S^2\right]$

Signifikanztests (Kapitel 14)

Gaußtest: $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ unbekannt, σ^2 bekannt. Teststatistik: $G = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$. Lehne $H_0 : \mu = \mu_0$ ab, falls $|g| > z_{1-\frac{\alpha}{2}}$. Lehne $H_0 : \mu \leq \mu_0$ ab, falls $g > z_{1-\alpha}$. Lehne $H_0 : \mu \geq \mu_0$ ab, falls $g < -z_{1-\alpha}$
Binomialtest: $X_i \overset{iid}{\sim} B(n,p)$, p unbek. Teststat.: $V = \sum_{i=1}^n X_i$. Größtes c_1 mit $\mathbb{P}(X \leq c_1) \leq \frac{\alpha}{2}$, $\mathbb{P}(X \leq c_1 + 1) > \frac{\alpha}{2}$. Kleinstes c_2 mit $\mathbb{P}(X \geq c_2) \leq \frac{\alpha}{2}$, $\mathbb{P}(X \geq c_2 - 1) > \frac{\alpha}{2}$. Lehne $H_0 : p = p_0$ ab, falls $v < c_1 \vee v > c_2$.
t-Test: $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ, σ^2 unbekannt. Teststatistik: $T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$. Lehne $H_0 : \mu = \mu_0$ ab, falls $t > t_{(n-1,1-\frac{\alpha}{2})}$. Lehne $H_0 : \mu \leq \mu_0$ ab, falls $t > t_{(n-1,1-\alpha)}$. Lehne $H_0 : \mu \geq \mu_0$ ab, falls $t < -t_{(n-1,1-\alpha)}$.
Approximativer Gaußtest: X_i iid bel. vert., μ, σ^2 unbek., $n \geq 30$. Teststatistik: T. Testentsch.: wie Gauß-Test.
χ^2-Test: $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ, σ^2 unbekannt. Teststatistik: $V = \frac{(n-1)}{\sigma_0^2}S^2$. Lehne $H_0 : \sigma^2 = \sigma_0^2$ ab, falls $v < c_{(n-1,\frac{\alpha}{2})}$ oder $v > c_{(n-1,1-\frac{\alpha}{2})}$. Lehne $H_0 : \sigma^2 \leq \sigma_0^2$ ab, falls $v > c_{(n-1,1-\alpha)}$. Lehne $H_0 : \sigma^2 \geq \sigma_0^2$ ab, falls $v < c_{(n-1,\alpha)}$.
Kontingenztest: $H_0 : X, Y$ unabhängig. Teststatistik: χ^2 aus Kapitel 4. Lehne H_0 ab, falls $\chi^2 > c_{(m,1-\alpha)}$, wobei $m = (\#$ Auspr. von $X - 1)(\#$ Auspr. von $Y - 1)$

Je nach Fortschritt der Vorlesung können hier noch weitere Formeln platziert werden.

Wahrscheinlichkeiten der Standardnormalverteilung $\mathcal{N}(0,1)$



z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,504	0,508	0,512	0,516	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Hinweis zur Benutzung dieser Tabelle:
Die Wahrscheinlichkeit, dass eine standardnormalverteilte Zufallsvariable Z einen Wert annimmt, welcher kleiner ist als $x+y$, ist in der Zeile x und Spalte y abzulesen. Zum Beispiel gilt: $\mathbb{P}(Z \leq 1,5+0,06) = \Phi(1,56) = 0,9406$, also 94,06%.

Kritische Werte der t-Verteilung

		Signifikanzniveau				
	einseitig:	10%	5%	2,5%	1%	0,5%
	zweiseitig:	20%	10%	5%	2%	1%
Freiheits-	1	3,078	6,314	12,706	31,821	63,657
grade	2	1,886	2,920	4,303	6,965	9,925
	3	1,638	2,353	3,182	4,541	5,841
	4	1,533	2,132	2,776	3,747	4,604
	5	1,476	2,015	2,571	3,365	4,032
	6	1,440	1,943	2,447	3,143	3,707
	7	1,415	1,895	2,365	2,998	3,499
	8	1,397	1,860	2,306	2,896	3,355
	9	1,383	1,833	2,262	2,821	3,250
	10	1,372	1,812	2,228	2,764	3,169
	11	1,363	1,796	2,201	2,718	3,106
	12	1,356	1,782	2,179	2,681	3,055
	13	1,350	1,771	2,160	2,650	3,012
	14	1,345	1,761	2,145	2,624	2,977
	15	1,341	1,753	2,131	2,602	2,947
	16	1,337	1,746	2,120	2,583	2,921
	17	1,333	1,740	2,110	2,567	2,898
	18	1,330	1,734	2,101	2,552	2,878
	19	1,328	1,729	2,093	2,539	2,861
	20	1,325	1,725	2,086	2,528	2,845
	21	1,323	1,721	2,080	2,518	2,831
	22	1,321	1,717	2,074	2,508	2,819
	23	1,319	1,714	2,069	2,500	2,807
	24	1,318	1,711	2,064	2,492	2,797
	25	1,316	1,708	2,060	2,485	2,787
	26	1,315	1,706	2,056	2,479	2,779
	27	1,314	1,703	2,052	2,473	2,771
	28	1,313	1,701	2,048	2,467	2,763
	29	1,311	1,699	2,045	2,462	2,756
	30	1,310	1,697	2,042	2,457	2,750
	40	1,303	1,684	2,021	2,423	2,704
	60	1,296	1,671	2,000	2,390	2,660
	90	1,291	1,662	1,987	2,368	2,632
	120	1,289	1,658	1,980	2,358	2,617
	∞	1,282	1,645	1,960	2,326	2,576

Hinweise zur Nutzung dieser Tabelle:

Es sei T eine Zufallsvariable, welche t-verteilt ist mit n Freiheitsgraden.

Einseitig: Das Quantil $t_{(n,1-\alpha)}$, für welches gilt $\mathbb{P}(T \leq t_{(n,1-\alpha)}) = 1-\alpha = \mathbb{P}(-t_{(n,1-\alpha)} \leq T)$ ist in der Tabelle abzulesen in der Zeile n und in der Spalte, welche unter „einseitig“ das Signifikanzniveau α angibt. Zum Beispiel gilt für $\alpha = 1\%$, $n = 22$ und $t_{(22,1-1\%)} = 2,508$, dass $\mathbb{P}(T \leq 2,508) = 1 - 1\% = 99\%$.

Zweiseitig: Das Quantil $t_{(n,1-\alpha/2)}$, für welches gilt $\mathbb{P}(-t_{(n,1-\alpha/2)} \leq T \leq t_{(n,1-\alpha/2)}) = 1-\alpha$ ist in der Tabelle abzulesen in der Zeile n und in der Spalte, welche unter „zweiseitig“ das Signifikanzniveau α angibt. Zum Beispiel gilt für $\alpha = 5\%$, $n = 17$ und $t_{(17,1-2.5\%)} = 2,110$, dass $\mathbb{P}(-2,110 \leq T \leq 2,110) = 1 - 5\% = 95\%$.

Kritische Werte der χ^2-Verteilung

		Signifikanzniveau		
		10%	5%	1%
Freiheits-Grade	1	2,71	3,84	6,63
	2	4,61	5,99	9,21
	3	6,25	7,81	11,34
	4	7,78	9,49	13,28
	5	9,24	11,07	15,09
	6	10,64	12,59	16,81
	7	12,02	14,07	18,48
	8	13,36	15,51	20,09
	9	14,68	16,92	21,67
	10	15,99	18,31	23,21
	11	17,28	19,68	24,72
	12	18,55	21,03	26,22
	13	19,81	22,36	27,69
	14	21,06	23,68	29,14
	15	22,31	25,00	30,58
	16	23,54	26,30	32,00
	17	24,77	27,59	33,41
	18	25,99	28,87	34,81
	19	27,20	30,14	36,19
	20	28,41	31,41	37,57
	21	29,62	32,67	38,93
	22	30,81	33,92	40,29
	23	32,01	35,17	41,64
	24	33,20	36,42	42,98
	25	34,38	37,65	44,31
	26	35,56	38,89	45,64
	27	36,74	40,11	46,96
	28	37,92	41,34	48,28
	29	39,09	42,56	49,59
	30	40,26	43,77	50,89

Hinweis zur Nutzung dieser Tabelle:

Es sei C eine Zufallsvariable, welche χ^2-verteilt ist mit n Freiheitsgraden. Das Quantil $c_{(n,1-\alpha)}$, für welches gilt $\mathbb{P}(C \leq c_{(n,1-\alpha)}) = 1 - \alpha$, ist in der Tabelle abzulesen in der Zeile n und in der Spalte, welche das Siginifikanzniveau α angibt. Zum Beispiel gilt für $\alpha = 1\%$, $n = 26$ und $c_{(26,1-1\%)} = 45,64$, dass $\mathbb{P}(C \leq 45,64) = 1 - 1\% = 99\%$.

Lösung zu Aufgabe 1

Der Anteil des Winkels von Windkraft an 360° muss 26,4% betragen. Demnach lautet der Winkel $360 \cdot 0,264 = 95,04$, was gerundet auf ganze Zahlen 95° ist.

Lösung zu Aufgabe 2

Die Werte 1 und 6 werden von je zwei Bundesländern „genannt", alle anderen Werte jeweils nur von einem Bundesland. Deswegen sind 1 und 6 die Modalwerte.

Lösung zu Aufgabe 3

Die Länder Baden-Württemberg (15), Bayern (7), Berlin (0), Rheinland-Pfalz (33), Saarland (6), Bremen (1), Sachsen (10), Hamburg (1), Sachsen-Anhalt (17), Hessen (37), Mecklenburg-Vorpommern (41) und Thüringen (6) haben weniger als 50 WEA aufgestellt, also 12 Länder. Dies entspricht einem Anteil von $\frac{12}{16} = 75\%$ aller 16 Bundesländer.

Lösung zu Aufgabe 4

Der relative Anstieg beträgt $(61010 - 58106)/58106 = 2904/58106 = 0,049977... \hat{=} 4,9977\%$, also 5%.

Lösung zu Aufgabe 5

Es gilt:

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

Daraus folgt:

$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B) + \mathbb{P}(A \cap B)$$

Also lautet die richtige Antwort $\mathbb{P}(A) + \mathbb{P}(B) = 55\% + 5\% = 60\%$.

Lösung zu Aufgabe 6

Die Wahrscheinlichkeit, dass keine WEA einen Schaden aufweist ist durch

$$(1-p)^{10} = (1-0,01554)^{10} = 0,98446^{10} \approx 0,8550 = 85,5\%$$

gegeben. Die Wahrscheinlichkeit, dass genau eine WEA einen Schaden aufweist ist durch

$$10 \cdot p \cdot (1-p)^9 = 10 \cdot 0,01554 \cdot (1-0,01544)^9 = 0,1554 \cdot 0,98446^9 \approx 0,1554 \cdot 0,8685 = 13,50\%$$

gegeben. In Summe ergibt sich die Wahrscheinlichkeit $85,5\% + 13,5\% = 99\%$.

Lösung zu Aufgabe 7

Wir müssen die normalverteilte Zufallsvariable S zunächst standardisieren:

$$S \geq 1992 \Leftrightarrow S - \mu \geq 1992 - \mu \Leftrightarrow \frac{S-\mu}{\sqrt{\sigma^2}} \geq \frac{1992-\mu}{\sqrt{\sigma^2}}$$

Mit $\mu = 2012$ und $\sqrt{\sigma^2} = \sqrt{256} = 16$ gilt $\frac{1992-\mu}{\sqrt{\sigma^2}} = \frac{1992-2012}{16} = \frac{-20}{16} = -\frac{5}{4} = -1,25$.

$$\mathbb{P}(S \geq 1992) = \mathbb{P}\left(\frac{S-\mu}{\sqrt{\sigma^2}} \geq -1,25\right) = \mathbb{P}\left(\frac{S-\mu}{\sqrt{\sigma^2}} \leq 1,25\right) = \Phi(1,25)$$

Diesen Wert können wir in der Tabelle auf Seite 18 ablesen: $\Phi(1,25) = 0,8944$. Gerundet auf ganze Prozentzahlen ergibt dies 89%.

Lösung zu Aufgabe 8

Die Formel für bedingte Wahrscheinlichkeiten ist gegeben durch:

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

Also gilt $\mathbb{P}(A|B) = \frac{5\%}{10\%} = \frac{1}{2} = 50\%$.

Lösung zu Aufgabe 9

Für den Gaußtest lautet die Testentscheidung: Lehne $H_0 : \mu \leq \mu_0$ genau dann ab, falls $g = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$. Mit $\mu_0 = 1992$, $\sigma = \sqrt{\sigma^2} = 16$, $\sqrt{n} = \sqrt{9} = 3$ und $z_{95\%} = 1,645$ ist die Ungleichung gegeben durch:

$$\frac{\bar{x} - 1992}{16/3} > 1,645 \Leftrightarrow \bar{x} > 1,645 \cdot \frac{1}{3} \cdot 16 + 1992 = 2000,77\bar{3} \approx 2001$$

Lösung zu Aufgabe 10

$$\begin{aligned}
\mathbb{E}[\hat{\mu}_1] &= \mathbb{E}[X_1 - 2X_2 + 3X_3 - 4X_4 + 5X_5 - 6X_6 + 7X_7 - 8X_8 + 9X_9] \\
&= \mathbb{E}[X_1] - 2\mathbb{E}[X_2] + 3\mathbb{E}[X_3] - 4\mathbb{E}[X_4] + 5\mathbb{E}[X_5] - 6\mathbb{E}[X_6] + 7\mathbb{E}[X_7] - 8\mathbb{E}[X_8] + 9\mathbb{E}[X_9] \\
&= \mu - 2\mu + 3\mu - 4\mu + 5\mu - 6\mu + 7\mu - 8\mu + 9\mu \\
&= \mu\,(1 - 2 + 3 - 4 + 5 - 6 + 7 - 8 + 9) = 5\mu \neq \mu \\
\mathbb{E}[\hat{\mu}_2] &= \mathbb{E}[\frac{1}{9}\,(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9)] \\
&= \frac{1}{9}\,(\mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] + \mathbb{E}[X_4] + \mathbb{E}[X_5] + \mathbb{E}[X_6] + \mathbb{E}[X_7] + \mathbb{E}[X_8] + \mathbb{E}[X_9]) \\
&= \frac{1}{9}\,(\mu + \mu + \mu + \mu + \mu + \mu + \mu + \mu + \mu) = \frac{1}{9} 9\mu = \mu \\
\mathbb{E}[\hat{\mu}_3] &= \mathbb{E}[X_1 - 2X_2] \\
&= \mathbb{E}[X_1] - 2\mathbb{E}[X_2] = \mu - 2\mu = -\mu \neq \mu \\
\mathbb{E}[\hat{\mu}_4] &= \mathbb{E}[\frac{1}{3}\,(X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9)] \\
&= \frac{1}{3}\,(\mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] + \mathbb{E}[X_4] + \mathbb{E}[X_5] + \mathbb{E}[X_6] + \mathbb{E}[X_7] + \mathbb{E}[X_8] + \mathbb{E}[X_9]) \\
&= \frac{1}{3}\,(\mu + \mu + \mu + \mu + \mu + \mu + \mu + \mu + \mu) = \frac{1}{3} 9\mu = 3\mu \neq \mu
\end{aligned}$$