C Lösungen zu Kontrollfragen

1. **a)** Falsch, weil $A \cap B = \{\text{"genügend"}\} \neq \emptyset$.
b) Richtig.
c) Richtig.
d) Richtig.
e) Falsch. Es ist $(A \cap B^c) \cup C = \{\text{"sehr gut"}, \text{"gut"}, \text{"unbrauchbar"}\}$.

2. **a)** Richtig.
b) Richtig, weil $B = [10, \infty) \subset [7, \infty) = A$ (Skizze machen!).
c) Richtig.

3. Generell: Alle Lösungen kann man schnell mit einem Venn-Diagramm herleiten.

a) Falsch. Richtig ist $A^c \cap B^c$.
b) Richtig, weil immer gilt $A \cap B \subseteq A \cup B$.
c) Falsch, weil wegen $A \subseteq B$ gilt, dass $A \cap B = A$.
d) Richtig, weil wegen $A \subseteq B$ gilt, dass $A \cup B = B$.
e) Richtig, weil $A \cap B^c$, $A \cap B$ und A^c paarweise disjunkt sind und zusammen Ω "ergeben", d.h. $\Omega = (A \cap B^c) \cup (A \cap B) \cup A^c$. Daher gilt $1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c)$.
f) Richtig, falls $A \cap B = \emptyset$ wäre, so hätte man $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 1.1 > 1$!

4. **a)** Richtig, weil $\sum_{k=1}^{4} \mathbb{P}(\omega_k) = 1$ sein muss.
b) Falsch, $\mathbb{P}(B) = \mathbb{P}(\omega_2) + \mathbb{P}(\omega_3) = 0.5$.
c) Richtig, weil $A \cup B = \{\omega_1,\ \omega_2,\ \omega_3\}$.
d) Richtig, weil $A \cap B = \{\omega_2\}$.
e) Falsch, beim Laplace-Modell wäre $\mathbb{P}(\omega_k) = 0.25$ für $k = 1, \ldots, 4$.

5. **a)** Richtig, weil beide Ereignisse 3 Elemente enthalten: $\{2,\ 4,\ 6\}$ bzw. $\{1,\ 3,\ 5\}$.
b) Richtig.
c) Richtig. Es handelt sich immer noch um die gleichen 6 Elementarereignisse.

6. **a)** Richtig.
b) Falsch, weil die Ereignisse disjunkt sind.
c) Falsch.

7. **a)** Falsch. Es ist $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\, \mathbb{P}(B) = 0.2 \cdot 0.5$.
b) Richtig gemäss dem Satz der totalen Wahrscheinlichkeit.
c) Falsch, es ist $\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B) = 0.8$.
d) Richtig. Dies ist $\mathbb{P}(B^c \mid A)$ berechnet mit dem Satz von Bayes.

8. **a)** Richtig.

b) Falsch. Mit dem Satz der totalen Wahrscheinlichkeit erhalten wir:

$$\begin{aligned}\mathbb{P}(N) &= \mathbb{P}(N \mid M_1)\,\mathbb{P}(M_1) + \mathbb{P}(N \mid M_2)\,\mathbb{P}(M_2) + \mathbb{P}(N \mid M_3)\,\mathbb{P}(M_3) \\ &= 0.1 \cdot 0.5 + 0.3 \cdot 0.4 + 0.5 \cdot 0.1 = 0.22.\end{aligned}$$

c) Falsch. Mit dem Satz von Bayes haben wir

$$\mathbb{P}(M_1 \mid N^c) = \frac{\mathbb{P}(N^c \mid M_1)\,\mathbb{P}(M_1)}{\mathbb{P}(N^c)},$$

wobei gemäss Aufgabenstellung $\mathbb{P}(N^c \mid M_1) = 1 - \mathbb{P}(N \mid M_1) = 0.9$ und $\mathbb{P}(M_1) = 0.5$ gilt. Ferner ist $\mathbb{P}(N^c) = 1 - \mathbb{P}(N) = 0.78$ (siehe oben). Alles eingesetzt erhalten wir also

$$\mathbb{P}(M_1 \mid N^c) = \frac{0.9 \cdot 0.5}{0.78} = 0.58.$$

9. **a)** Richtig.

b) Richtig, weil $\mathbb{P}(T_+ \cap D) = \mathbb{P}(T_+ \mid D)\,\mathbb{P}(D) = 0.9 \cdot 0.1 = 0.09$.

c) Richtig. Es gilt weges des Satzes von Bayes

$$\mathbb{P}(D \mid T_-) = \frac{\mathbb{P}(T_- \mid D)\,\mathbb{P}(D)}{\mathbb{P}(T_- \mid D)\,\mathbb{P}(D) + \mathbb{P}(T_- \mid D^c)\,\mathbb{P}(D^c)}.$$

Ferner gilt gemäss Aufgabestellung $\mathbb{P}(T_- \mid D) = 1 - \mathbb{P}(T_+ \mid D) = 0.1$ und $\mathbb{P}(T_- \mid D^c) = 1 - \mathbb{P}(T_+ \mid D^c) = 0.95$. Wir haben also

$$\mathbb{P}(D \mid T_-) = \frac{0.1 \cdot 0.1}{0.1 \cdot 0.1 + 0.95 \cdot 0.9} = 0.0116.$$

10. **a)** Richtig. Falsch wäre $\mathbb{P}(x \geq 1)$.

b) Richtig.

11. **a)** Falsch. Richtig ist $\mathbb{P}(X = 6) = F(6) - F(5) = 0.61 - 0.45 = 0.16$.

b) Richtig. Dies ist $\mathbb{P}(X \leq 7) = F(7) = 0.74$.

c) Richtig. Es ist $\mathbb{P}(X \geq 10) = 1 - \mathbb{P}(X \leq 9) = 1 - F(9) = 1 - 0.92 = 0.08$.

12. **a)** Richtig. Es ist

$$\mathbb{E}[X] = \sum_{k=1}^{4} x_k p(x_k) = 30 \cdot 0.6 + 60 \cdot 0.2 + 180 \cdot 0.15 + 300 \cdot 0.05 = 72.$$

b) Falsch. Die Varianz ist 5436. Dies kann man entweder durch

$$\operatorname{Var}(X) = \sum_{k=1}^{4} (x_k - 72)^2 p(x_k)$$

berechnen oder via

$$\operatorname{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \sum_{k=1}^{4} x_k^2 p(x_k) - 72^2 = 10\,620 - 72^2 = 5436.$$

13. **a)** Richtig. Es ist $\mathbb{E}[50X] = 50 \cdot \mathbb{E}[X] = 50 \cdot 200 = 10\,000$.

b) Falsch. Die Standardabweichung beträgt $50 \cdot \sqrt{230} = 758.3$.

14. **a)** Richtig. Es ist $\mathbb{E}[X-1] = \mathbb{E}[X] - 1 = 1 - 1 = 0$.

b) Richtig. Es ist $\mathbb{E}[4X+3Y] = \mathbb{E}[4X] + \mathbb{E}[3Y] = 4 \cdot \mathbb{E}[X] + 3 \cdot \mathbb{E}[Y]$.

c) Falsch. Es ist $\operatorname{Var}(-Y) = \operatorname{Var}((-1) \cdot Y) = (-1)^2 \operatorname{Var}(Y) = \operatorname{Var}(Y) = 4$.

d) Falsch. Es ist $\operatorname{Var}\left(5 + \frac{1}{2}Y\right) = \left(\frac{1}{2}\right)^2 \operatorname{Var}(Y) = 1$.

e) Richtig. Es ist $\operatorname{Var}(-2Y) = (-2)^2 \cdot \operatorname{Var}(Y) = 16$. Die Standardabweichung ist daher $\sqrt{16} = 4$.

f) Richtig. Es ist $\mathbb{E}\left[X^2\right] = \operatorname{Var}(X) + (\mathbb{E}[X])^2$.

15. **a)** Richtig. Wenn wir mit X die Anzahl Chips mit Defekt bezeichnen, dann gilt $X \sim \operatorname{Bin}(n,p)$ mit $n = 1000$ und $p = 0.001$. Also ist

$$\mathbb{P}(X=2) = \binom{1000}{2} 0.001^2 0.999^{998} = \frac{1000 \cdot 999}{2} 0.001^2 0.999^{998} = 0.184.$$

b) Falsch. Die Wahrscheinlichkeit ist wegen der Unabhängigkeit gegeben durch $0.001^2(1-0.001)^{998}$.

c) Richtig. Wir können die Anzahl fehlerhafter Chips approximieren mit einer Poissonverteilung mit Parameter $\lambda = np = 1000 \cdot 0.001 = 1$. Dann gilt

$$\mathbb{P}(X=1) = e^{-\lambda} \frac{\lambda^1}{1!} = e^{-1} = 0.368.$$

d) Falsch.

e) Richtig. Der Wertebereich ist $W = \{0, 1, 2, \ldots, 1000\}$ und umfasst daher 1001 verschiedene Werte. Bei jedem dieser Werte "springt" die kumulative Verteilungsfunktion mit der entsprechenden Wahrscheinlichkeit nach oben.

16. **a)** Richtig. Es ist $\mathbb{P}(X=3) = e^{-7} \frac{7^3}{3!} = 0.052$.

b) Richtig. Es ist $\mathbb{E}\left[\frac{1}{7}X\right] = \frac{1}{7} \cdot \mathbb{E}[X] = 1$.

c) Falsch. Es ist $\operatorname{Var}\left(\frac{1}{7}X\right) = \left(\frac{1}{7}\right)^2 \cdot 7 = \frac{1}{7}$.

d) Falsch. $X + Y$ ist zwar immer noch eine diskrete Zufallsvariable, aber es ist $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 7 + 10 \cdot 0.5 = 12$.

e) Richtig. Es ist $X + Y \sim \operatorname{Pois}(15)$ und $e^{-15} \frac{15^{13}}{13!} = 0.096$.

f) Falsch. Der Parameter muss eine positive *reelle* Zahl sein.

17. **a)** Falsch. Es ist $\mathbb{P}(X = 0.3) = 0$.

b) Falsch. Eine Dichte kann auch grösser als 1 sein, solange die Fläche unter der Dichte 1 ist.

c) Richtig.

d) Richtig, weil das Ereignis $\{|X| \leq 1\}$ geschrieben werden kann als $\{-1 \leq X \leq 1\}$.

e) Falsch. Dies gilt nur umgekehrt. Falls die Dichte symmetrisch um 3 ist, dann liegt auch der Erwartungswert bei 3.

f) Richtig.

g) Falsch.

h) Richtig.

18. **a)** Richtig, weil der Wert 0 näher am Symmetriepunkt $\mu = 1$ liegt als der Wert 3 (Skizze machen!).

b) Richtig. Im Bereich $[\mu - \sigma, \mu + \sigma]$ liegt ca. Wahrscheinlichkeit $2/3$. Wir betrachten nur die eine Hälfte davon. Wegen der Symmetrie sind es dann $1/3$.

c) Richtig.

d) Richtig, weil das 80%-Quantil sicher grösser als der Median ist, welcher hier 1 beträgt.

19. **a)** Richtig, weil $\mathbb{E}\left[e^X\right] = \int_0^1 e^x \,\mathrm{d}x = e^1 - e^0 = e - 1$ gilt.

b) Richtig, weil $\lambda = \frac{1}{100}$.

c) Falsch. Weil $\mathbb{E}[X] = \frac{1}{\lambda}$ und $\mathrm{Var}(X) = \frac{1}{\lambda^2}$ gilt für alle $\lambda > 0$, muss $\sigma_X = \frac{1}{\lambda}$ sein.

d) Richtig, weil die Dichte monoton fallend ist, und daher die Fläche über einem Intervall mit Länge 1 immer kleiner wird.

20. **a)** Richtig.

b) Falsch. $Y = e^X$ wäre lognormal-verteilt.

c) Richtig. Es ist $\mathbb{P}(X \leq 0) = \Phi(-2.5) = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$.

21. **a)** Richtig. Die Anzahl Kaffees in einer 6-Minuten-Periode ist poisson-verteilt mit Parameter $\lambda = 20 \cdot 6/60 = 2$.

b) Richtig. Die Anzahl Kaffees in einer 3-Stunden-Periode ist poisson-verteilt mit Parameter $\lambda = 20 \cdot 3 = 60$.

c) Richtig. Die Wartezeit T zwischen zwei Kaffees (in Stunden) ist exponential-verteilt mit Parameter $\lambda = 20$. Gefragt ist $\mathbb{P}\left(T > \frac{12}{60}\right) = 1 - \mathbb{P}\left(T \leq \frac{12}{60}\right) = e^{-20 \cdot \frac{12}{60}} = e^{-4}$.

d) Richtig, weil die Intensität über 24 Stunden nicht konstant ist.

22. **a)** Falsch. Weil $20 \cdot 0.25 = 5$ gilt, ist das empirische 25%-Quantil gegeben durch $\frac{1}{2}\left(x_{(5)} + x_{(6)}\right) = 0.635$.

b) Falsch. Der empirische Median ist $\frac{1}{2}(1.22 + 1.26) = 1.24$.

c) Richtig.

d) Richtig.

23. **a)** Richtig.

b) Falsch.

c) Richtig.

d) Richtig.

24. **a)** Richtig.

b) Falsch. Erst für $x < 0.12$ gilt $F_n(x) = 0$.

25. **a)** Falsch. Wir wissen nur anhand der Histogramme nicht, wie die Punkte zusammenhängen.

b) Richtig.

c) Richtig.

d) Falsch. Die empirische Korrelation wäre dann 1.

26. **a)** Falsch.

b) Richtig.

27. **a)** Richtig.

b) Falsch, dies geht nur bei Unabhängigkeit.

c) Falsch.

d) Richtig.

28. **a)** Richtig, weil der linke Gipfel höher ist als der rechte Gipfel.
b) Richtig.
c) Falsch. Zum Beispiel hängt die Form der bedingten Dichte von Y gegeben $X = x$ vom Wert x ab.
d) Falsch.

29. **a)** Falsch.
b) Falsch.
c) Falsch. In diesem Fall wäre die Korrelation $3/2 > 1$.

30. **a)** Richtig.
b) Falsch. Es ist $\text{Var}(X - Y) = \text{Var}(X + (-Y)) = \text{Var}(X) + \text{Var}(-Y) = \text{Var}(X) + \text{Var}(Y)$.
c) Richtig, weil ein perfekter linearer (positiver) Zusammenhang vorhanden ist.

31. **a)** Richtig.
b) Richtig.
c) Falsch. Es gibt Abhängigkeiten und saisonale Schwankungen.

32. **a)** Falsch. Die Würfe sind ja unabhängig voneinander.
b) Richtig.
c) Falsch. Erst wenn wir summieren oder mitteln kommt der Zentrale Grenzwertsatz ins Spiel.

33. **a)** Falsch. Es werden immer die Quantile der Standardnormalverteilung verwendet.
b) Falsch. Es gibt immer eine natürliche Streuung.
c) Richtig.
d) Falsch.

34. **a)** Richtig. Dies ist gerade die erste Gleichung bei der Momentenmethode.
b) Falsch. Der Schätzer entspricht dem arithmetischen Mittel der Daten, welches natürlich in der Regel nicht exakt 2 ist.
c) Richtig.
d) Richtig.
e) Richtig. Das geschätzte Modell entspricht in der Regel *nicht* dem wahren Modell, sondern ist ja "nur" eine Schätzung dafür.

35. **a)** Falsch. Der Index i muss jeweils bei der Beobachtung x_i stehen und nicht bei den Parametern.
b) Richtig, weil die Likelihood im diskreten Fall ein Produkt von Wahrscheinlichkeiten ist.
c) Richtig, weil ein Parameterschätzer von den zufälligen Daten abhängt.
d) Falsch. Richtig wäre es gerade anders herum.

36. **a)** Falsch. Das Signifikanzniveau kontrolliert nur die Wahrscheinlichkeit für einen Fehler 1. Art.
b) Falsch. Darüber können wir keine Aussagen machen, weil die Parameter fix und nicht zufällig sind. Aber wir wissen: Wenn die Nullhypothese stimmt, dann machen wir nur mit Wahrscheinlichkeit 5% einen Fehler 1. Art.
c) Falsch. Wir müssen schauen, wie "weit aussen" das beobachtete Ereignis in

der Verteilung unter der Nullhypothese liegt. Bei diskreten Verteilungen mit vielen möglichen Ausgängen ist es ganz natürlich, dass die Einzelwahrscheinlichkeiten klein sind (weil sie sich zu 1 addieren müssen).

d) Falsch.

37. **a)** Richtig.

b) Richtig. Unter H_0 gilt für die Summe, dass diese Pois (50)-verteilt ist.

c) Richtig. Der Verwerfungsbereich ist gegeben durch $K = \{63, 64, \ldots\}$. Die Beobachtung 66 liegt daher im Verwerfungsbereich.

38. **a)** Richtig.

b) Falsch. Es ist $K = (-\infty, -t_{14,0.975}] \cup [t_{14,0.975}, \infty) = (-\infty, -2.145] \cup [2.145, \infty)$.

c) Richtig.

d) Richtig.

e) Falsch. Die Wahrscheinlichkeit für einen Fehler 1. Art entspricht bei einem t-Test immer exakt dem verwendeten Signifikanzniveau und nimmt daher insbesondere mit der Stichprobengrösse *nicht* ab.

39. **a)** Richtig. Unter der Nullhypothese gilt $X_1 + \ldots + X_{10} \sim \text{Pois}(10 \cdot 3)$. Diese Verteilung approximieren mit einer Normalverteilung mit dem passenden Erwartungswert bzw. Varianz.

b) Richtig.

c) Falsch. Es ist $z = \frac{36-30}{\sqrt{30}}$.

40. **a)** Falsch. Wir haben eine Wahrscheinlichkeit von 75%, dass wir ein signifikantes Resultat erhalten, wenn wir *annehmen*, dass tatsächlich diese Alternativhypothese gilt.

b) Falsch. Zur Berechnung der Macht eines statistischen Tests müssen wir "nur" eine entsprechende Annahme der Modellparameter im Bereich der Alternativhypothese treffen.

c) Richtig. Es wird generell "schwieriger", die Nullhypothese zu verwerfen, weil der Verwerfungsbereich *kleiner* wird (man schneidet an den Rändern der Verteilung unter der Nullhypothese weniger ab). Oder kurz: Man verwendet beim Test ein "strengeres" Kriterium.

41. **a)** Richtig, weil $0.01 <$ p-Wert < 0.05.

b) Richtig.

c) Richtig. Grund dafür ist die Symmetrie der t-Verteilung (Skizze machen!).

d) Richtig. Gemäss Bonferroni-Korrektur muss p-Wert $\leq \frac{\alpha}{20}$ gelten, damit man die Nullhypothese verwirft. Dies ist äquivalent zu p-Wert $\cdot\, 20 \leq \alpha$.

42. **a)** Richtig.

b) Falsch. Das Vertrauensintervall macht "nur" eine Aussage über einen *Parameter* einer Verteilung, nicht über eine Einzelbeobachtung einer Verteilung.

c) Richtig, weil der Wert $p = p_0 = 0.5$ *nicht* im Vertrauensintervall enthalten ist.

d) Richtig.

43. **a)** Richtig. Der p-Wert ist hier gegeben durch $\mathbb{P}(X = 14) + \mathbb{P}(X = 15)$ für $X \sim \text{Bin}(15, 0.5)$.

b) Richtig.

c) Falsch. Die Beobachtung x_2 ist und bleibt die betragsmässig grösste Abweichung von μ_0, daher bleibt $\text{Rang}(|x_2 - \mu_0|) = 10$.

44. **a)** Richtig.

b) Falsch. Der beobachtete Wert der Teststatistik ist $t = 5.64/(7.4/\sqrt{8}) = 2.16$. Es $t_{7,0.975} = 2.365$. Also können wir die Nullhypothese nicht verwerfen.

c) Richtig.

d) Falsch. Es wird angenommen, dass die Differenzen d_i i.i.d. Realisierungen einer Normalverteilung sind.

45. **a)** Richtig.

b) Richtig. Es ist $\widehat{\sigma}^2_{pool} = \frac{1}{18}\left(9 \cdot 0.2^2 + 9 \cdot 0.15^2\right) = \frac{1}{2}\left(0.2^2 + 0.15^2\right)$.

c) Richtig.

d) Falsch. Bei $\widehat{\sigma}_{pool}$ darf kein Quadrat stehen.