Prof. Dr. P. Embrechts — ETH Zürich — Winter 2015

Wahrscheinlichkeit und Statistik
BSc D-INFK

1. **a)** 3. **b)** 1. **c)** 2. **d)** 3. **e)** 2. **f)** 1. **g)** 3. **h)** 1. **i)** 3. **j)** 2.

2. Seien die folgenden Ereignisse mit den unten angegebenen Symbolen notiert:

A={Die Server sind bereits ausgelastet.}
A^c={Die Server sind nicht ausgelastet.}
E={Die Einstellung wird aktiviert.}
E^c={Die Einstellung wird nicht aktiviert.}

Folgende Wahrscheinlichkeiten werden in der Aufgabenstellung gegeben:
$P[E^c|A] = 0.1$ und $P[E|A^c] = 0.2$

a) Folgende Angabe wird in der Aufgabenstellung gemacht: $P[A] = p = \frac{1}{3}$.
Gefragt wird nach den Wahrscheinlichkeiten $P[A \cap E^c]$ und $P[A^c \cap E^c]$.
Es gilt

$$P[A \cap E^c] = P[E^c|A]\ P[A] = 0.1 \cdot \frac{1}{3} = \frac{1}{30},$$

und

$$P[A^c \cap E^c] = (1 - P[E|A^c])\ P[A^c] = 0.8 \cdot \frac{2}{3} = \frac{8}{15}.$$

b) Folgende Angabe wird in der Aufgabenstellung gemacht: $P[A] = p$, wobei der Wert von p nun unbekannt ist.
Gefragt wird nach einer Formel für $P[E]$ in Abhängigkeit von dem Wert p.
Dies kann wie folgt berechnet werden:

$$\begin{aligned} P[E] = 1 - P[E^c] &= 1 - (P[A \cap E^c] + P[A^c \cap E^c]) \\ &= 1 - 0.1 \cdot p - 0.8 \cdot (1 - p) \\ &= 0.2 + 0.7 \cdot p, \end{aligned}$$

Bitte wenden!

oder

$$\begin{aligned} P[E] &= P[E|A] \cdot P[A] + P[E|A^c] \cdot P[A^c] \\ &= (1 - P[E^c|A]) \cdot P[A] + P[E|A^c] \cdot (1 - P[A]) \\ &= 0.9 \cdot p - 0.2 \cdot (1-p) \\ &= 0.2 + 0.7 \cdot p. \end{aligned}$$

Es gilt:

$$0.2 + 0.7 \cdot p \geq \frac{1}{2} \quad \text{genau dann wenn} \quad p \geq \frac{3}{7}.$$

c) Folgende Angabe wird in der Aufgabenstellung gemacht: $P[E] = 0.3$ und die Anzahl der Versuche beträgt $n = 3$.
Es gilt

$$T \sim \text{Bin}(3, 0.3) - \text{verteilt.}$$

Gefragt wird nach der Wahrscheinlichkeit $P[T = k]$ in Abhängigkeit von p.
Diese ist

$$\binom{3}{k} \left(\frac{7}{10}\right)^{3-k} \left(\frac{3}{10}\right)^{k}.$$

Für $k = 1$ hat sie den Wert $\left(\frac{49}{100}\right)\left(\frac{9}{10}\right) < \frac{50}{100}$ und liegt somit nicht über 50%.

d) Folgende Angabe wird in der Aufgabenstellung gemacht: $P[E] = 0.3$.
Gefragt wird nach dem Wert $p = P[A]$.
Dies lässt sich berechnen aus

$$\begin{aligned} P[E] &= P[E|A] \cdot P[A] + P[E|A^c] \cdot P[A^c] \\ &= (1 - P[E^c|A]) \cdot P[A] + P[E|A^c] \cdot (1 - P[A]), \end{aligned}$$

woraus durch Auflösen nach $P[A]$

$$P[A] = \frac{P[E] - P[E|A^c]}{1 - P[E^c|A] - P[E|A^c]}$$

folgt. Alternativ kann man direkt die Formel $P[E] = 0.2 + 0.7 \cdot p$ verwenden, woraus

$$p = \frac{P[E] - 0.2}{0.7} = \frac{1}{7}$$

folgt.

Siehe nächstes Blatt!

3. 2 Punkte: **Antwort 1**

Die Dichte der Summe zweier unabhängiger Zufallsvariablen ist gegeben durch die Faltung ihrer Dichten. Sei $z \in \mathbb{R}$. Dann gilt

$$\begin{aligned} f_W(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_0^{\infty} \lambda \exp(-\lambda x) I_{\{0 \le z-x \le 1\}}\,\mathrm{d}x \\ &= \int_0^{\infty} \lambda \exp(-\lambda x) I_{\{z-1 \le x \le z\}}\,\mathrm{d}x \end{aligned}$$

Falls $z < 0$, dann ist $f_W(z) = 0$. Falls $z \in [0,1]$, so ist

$$f_W(z) \underset{(0.5)}{=} \int_0^z \lambda \exp(-\lambda x)\,\mathrm{d}x = -\exp(-\lambda x)\big|_{x=0}^{z} \underset{(0.5)}{=} 1 - \exp(-\lambda z)$$

und für $z \ge 1$ gilt

$$\begin{aligned} f_W(z) &\underset{(0.5)}{=} \int_{z-1}^{z} \lambda \exp(-\lambda x)\,\mathrm{d}x = \exp(-\lambda(z-1)) - \exp(-\lambda z) \\ &\underset{(0.5)}{=} \exp(-\lambda z)(\exp(\lambda) - 1). \end{aligned}$$

Zusammengefasst haben wir also

$$f_W(z) = \begin{cases} 0, & \text{falls } z < 0, \\ 1 - \exp(-\lambda z), & \text{falls } 0 \le z < 1, \\ \exp(-\lambda z)(\exp(\lambda) - 1), & \text{falls } 1 \le z. \end{cases}$$

2 Punkte: **Antwort 2**

$$\begin{aligned} f_W(z) &= \int_{-\infty}^{\infty} f_X(z-x) f_Y(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \lambda \exp(-\lambda(z-x)) I_{\{z-x \ge 0\}} I_{\{0 \le x \le 1\}}\,\mathrm{d}x \\ &= \int_0^1 \lambda \exp(-\lambda(z-x)) I_{\{z-x \ge 0\}}\,\mathrm{d}x \end{aligned}$$

Falls $z < 0$, dann ist $f_W(z) = 0$. Falls $z \in [0,1]$, so ist

$$f_W(z) \underset{(0.5)}{=} \int_0^z \lambda \exp(-\lambda(z-x))\,\mathrm{d}x = \exp(-\lambda z)\left(\exp(\lambda x)\right)\big|_{x=0}^{z} \underset{(0.5)}{=} 1 - \exp(-\lambda z)$$

und für $z \ge 1$ gilt

$$f_W(z) \underset{(0.5)}{=} \int_0^1 \lambda \exp(-\lambda(z-x))\,\mathrm{d}x \underset{(0.5)}{=} \exp(-\lambda(z))\left(\exp(-\lambda) - 1\right).$$

Zusammengefasst haben wir also

$$f_W(z) = \begin{cases} 0, & \text{falls } z < 0, \\ 1 - \exp(-\lambda z), & \text{falls } 0 \le z < 1, \\ \exp(-\lambda z)(\exp(\lambda) - 1), & \text{falls } 1 \le z. \end{cases}$$

Bitte wenden!

b) 2 Punkte: Die Linearität des Erwartungswertes gibt $E[W] = E[X] + E[Y]$. Wir wissen, dass $E[X] = \frac{1}{\lambda}$ und $\underbrace{E[Y] = \int_0^1 y \, \mathrm{d}y = \frac{1}{2}}_{(0.5)}$. Also ist $\underbrace{E[W] = \frac{1}{\lambda} + \frac{1}{2}}_{(0.5)}$.

Für die Varianz gilt wegen Unabhängigkeit $\mathrm{Var}[W] = \mathrm{Var}[X] + \mathrm{Var}[Y]$. Wir wissen, dass $E[X] = \frac{1}{\lambda^2}$ und

$$\underbrace{\mathrm{Var}[Y] = E[Y^2] - E[Y]^2 = \int_0^1 y^2 \, \mathrm{d}y - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.}_{(0.5)}$$

Also ist $\underbrace{\mathrm{Var}[W] = \frac{1}{\lambda^2} + \frac{1}{12}}_{(0.5)}$.

Bemerkung: Wenn für den Erwartungswert und für die Varianz nur die Formel hingeschrieben wurde aber nicht vollständig ausgerechnet wurde, gab es einen halben Punkt Abzug.

c) 2 Punkte: Für $t \leq 0$ gilt $f_Z(t) = 0$. Sei $t > 0$. Dann gilt mit partieller Integration

$$\begin{aligned}
F_Z(t) &= \int_{-\infty}^{t} f_Z(z) \, \mathrm{d}z \\
&= \int_{-\infty}^{t} \alpha^2 z \exp(-\alpha z) \, \mathrm{d}z \\
&\underset{(1\text{Punkt})}{=} -\alpha z \exp(-\alpha z)\Big|_{z=0}^{t} + \int_0^t \alpha \exp(-\alpha z) \, \mathrm{d}z \\
&= -\alpha t \exp(-\alpha t) + 1 - \exp(-\alpha t) \\
&\underset{(1\text{Punkt})}{=} 1 - \exp(-\alpha t)(1 + \alpha t).
\end{aligned}$$

d) 2 Punkte: Die Log-Likelihood-Funktion lautet

$$\begin{aligned}
l(z_1, \ldots, z_n; \alpha) &\underset{(0.5)}{=} \log \prod_{i=1}^{n} \alpha^2 z_i \exp(-\alpha z_i) = \sum_{i=1}^{n} (2 \log \alpha + \log z_i - \alpha z_i) \\
&= 2n \log \alpha + \sum_{i=1}^{n} \log z_i - \alpha \sum_{i=1}^{n} z_i.
\end{aligned}$$

Die Ableitung der Log-Likelihood-Funktion nach α ist

$$\frac{\partial}{\partial \alpha} l(z_1, \ldots, z_n; \alpha) \underset{(0.5)}{=} \frac{2n}{\alpha} - \sum_{i=1}^{n} z_i$$

Siehe nächstes Blatt!

und diese ist 0 für

$$\alpha \underset{(0.5)}{=} \frac{2}{\frac{1}{n}\sum_{i=1}^{n} z_i} = \frac{2}{\overline{z}_n}.$$

Der Maximum-Likelihood-Schätzer für α lautet also $\underbrace{T = 2/\overline{Z}_n}_{(0.5)}$.

e) 2 Punkte: Sei S die Anzahl der CPUs die innerhalb der nächsten 10 Jahre kaputt gehen. Dann ist S $\mathrm{Bin}(n,p)$-verteilt mit $n = 100$ und $p = 0.5$. Daher ist $\underbrace{E[S] = np = 50 \text{ und } \mathrm{Var}[S] = np(1-p) = 25}_{(0.5)}$. Nun approximieren wir mit dem zentralen Grenzwertsatz

$$\begin{aligned}
P[S = 40] = P[39.5 < S \le 40.5] &\underset{(0.5)}{=} P\left[-\frac{10.5}{5} < \frac{S - E[S]}{\sqrt{\mathrm{Var}[S]}} \le -\frac{9.5}{5}\right] \\
&\underset{(0.5)}{=} P\left[\frac{S - E[S]}{\sqrt{\mathrm{Var}[S]}} \le -\frac{9.5}{5}\right] - P\left[\frac{S - E[S]}{\sqrt{\mathrm{Var}[S]}} - \frac{10.5}{5}\right] \\
&\approx \Phi(-1.9) - \Phi(-2.1) = \Phi(2.1) - \Phi(1.9) \\
&\underset{(0.5)}{\approx} 0.9821 - 0.9713 = 0.0108.
\end{aligned}$$

4. **a)** Da die Daten nach Annahme normalverteilt sind mit unbekannter Varianz und unbekanntem Mittelwert, bietet sich ein t-Test an. Da wir Zweifel an der Aussage des Anbieters haben, dass seine mittlere Übertragungsrate *mindestens* 10 Mbit/s beträgt, interessieren uns nur Abweichungen nach unten. Somit ist der Test einseitig.

b) Seien $X_1, \dots, X_9$ unter $P_{(\mu,\sigma)}$ i.i.d. mit Verteilung $\mathcal{N}(\mu, \sigma^2)$.

(i) Die Nullhypothese und die Alternative lauten

$$H_0 : \mu = \mu_0 := 10 \quad \text{und} \quad H_A : \mu < \mu_0.$$

(ii) Die Teststatistik des t-Tests lautet $T = \frac{\overline{X}_n - \mu_0}{S/\sqrt{n}} = \frac{\overline{X}_9 - 10}{S/3}$. Unter H_0 ist T t-verteilt mit 8 Freiheitsgraden.

(iii) Der Verwerfungsbereich hat die Form $K = (-\infty, c]$ für ein $c \in \mathbb{R}$. Damit der Test das Niveau $\alpha = 5\%$ einhält, muss man für c das α-Quantil der t_8-Verteilung wählen. Wegen Symmetrie der Dichte der t-Verteilung gilt $t_{8,\alpha} = -t_{8,1-\alpha}$. Aus der Tabelle der Quantile der t-Verteilung erhalten wir $t_{8,0.95} = 1.860$. Somit ist der gesuchte Verwerfungsbereich $K = (-\infty, -1.860]$.

(iv) Der beobachtete Wert der Teststatistik ist $T(\omega) = \frac{\overline{x}_9 - 10}{s/3} = \frac{9.8-10}{0.3/3} = -2$.

(v) Da $T(\omega) = -2 \in K$ lautet der Testentscheid, die Nullhypothese auf dem 5%-Niveau zu verwerfen.

Bitte wenden!

c) Die Teststatistik T des z-Tests ist unter der Nullhypothese standardnormalverteilt. Der Test verwirft, falls $T(\omega) = 2 \geq c$. Je grösser c, desto kleiner das Niveau des Tests. Das Grösstmögliche Niveau, für das der Test noch verwirft erhalten wir also, wenn wir $c = 2$ wählen. Das zu diesem Test gehörige Niveau ist dann gerade

$$P_{\mu_0}[T \geq 2] = 1 - \Phi(2) \approx 1 - 0.97725 = 0.02275.$$

Somit ist der P-Wert gleich 2.275%.

d) Die Teststatistik des z-Tests ist $T = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}}$ mit $\mu_0 = 100$ und $\sigma = 10$. Wir wollen

$$\begin{aligned}
0.9 \overset{!}{\leq} \beta(\mu_A) = P_{\mu_A}[T \in K] = P_{\mu_A}[T \geq c] &= P_{\mu_A}\left[\frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} \geq c\right] \\
&= P_{\mu_A}\left[\frac{\overline{X}_n - \mu_A}{\sigma/\sqrt{n}} - \frac{\mu_0 - \mu_A}{\sigma/\sqrt{n}} \geq c\right] = P_{\mu_A}\left[\frac{\overline{X}_n - \mu_A}{\sigma/\sqrt{n}} \geq c + \frac{\mu_0 - \mu_A}{\sigma/\sqrt{n}}\right] \\
&= 1 - \Phi\left(c - \sqrt{n}\right) = \Phi(-c + \sqrt{n}).
\end{aligned}$$

Aus der Tabelle erhalten wir somit $-c + \sqrt{n} \geq 1.28$. Einsetzen von $c = 2.72$ und Auflösen nach n liefert dann $n \geq (2.72 + 1.28)^2 = 16$. Wir brauchen also mindestens 16 Beobachtungen damit $\beta(\mu_A) \geq 0.9$.