

Lösung der Nachklausur

1. Die Wahrscheinlichkeit, dass eine 5 gewürfelt wird ist

$$\frac{1}{3} \cdot 0.16 + \frac{1}{3} \cdot 0.24 + \frac{1}{3} \cdot \frac{1}{6} = \frac{17}{90} .$$

- a) Die Wahrscheinlichkeit, dass Würfel 1 gezogen und eine 5 geworfen wurde ist $\frac{1}{3} \cdot 0.16 = \frac{16}{300}$. Somit ist die bedingte Wahrscheinlichkeit, dass Würfel 1 gezogen wurde

$$\frac{\frac{16}{300}}{\frac{17}{90}} = \frac{24}{85} .$$

- b) Die Wahrscheinlichkeit, dass Würfel 2 gezogen und eine 5 geworfen wurde ist $\frac{1}{3} \cdot 0.24 = 0.08$. Somit ist die bedingte Wahrscheinlichkeit, dass Würfel 2 gezogen wurde

$$\frac{0.08}{\frac{17}{90}} = \frac{36}{85} .$$

- c) Die Wahrscheinlichkeit, dass der faire Würfel gezogen wurde ist somit

$$1 - \frac{24}{85} - \frac{36}{85} = \frac{25}{85} .$$

2. a) Die Wahrscheinlichkeit, dass man im ersten Spiel verliert ist $\frac{1}{2}$. In diesem Falle muss man im Mittel wieder $\mathbb{E}[T]$ warten, also insgesamt im Mittel $\mathbb{E}[T] + 1$. Das erste Spiel gewinnt man mit Wahrscheinlichkeit $\frac{1}{2}$. Verliert die Spielerin das zweite Spiel, so muss sie wieder im Schnitt $\mathbb{E}[T]$ warten, also insgesamt $\mathbb{E}[T] + 2$. Gewinnt die Spielerin die ersten beiden Spiele (Wahrscheinlichkeit $\frac{1}{4}$) und verliert sie das dritte Spiel, so muss sie im Mittel insgesamt $\mathbb{E}[T + 3]$ warten. Gewinnt sie das dritte Spiel auch noch, so hat sie dreimal gespielt. Damit erhält man

$$\mathbb{E}[T] = \frac{1}{2}(\mathbb{E}[T] + 1) + \frac{1}{2}\left[\frac{1}{2}(\mathbb{E}[T] + 2) + \frac{1}{2}\left\{\frac{1}{2}(\mathbb{E}[T] + 3) + \frac{1}{2} \cdot 3\right\}\right] ,$$

was die gesuchte Formel ist. Lösen wir nach $\mathbb{E}[T]$ auf, erhalten wir $\mathbb{E}[T] = 14$.

- b) Der Satz über die Unmöglichkeit gewinnträchtiger Spielsysteme ergibt, dass der mittlere Gewinn 0 ist.

Man denkt intuitiv, dass man im Mittel gewinnen muss, da ja der letzten drei Spiele zu einem Gewinn geführt haben. Spielt man aber relativ lange, so verliert man im Mittel, da ja drei Gewinne nacheinander lange nicht auftreten. Dieser Umstand wird besonders deutlich, wenn man nur auf den ersten Gewinn wartet.

- c) Aus Korollar 1.15 wissen wir, dass $\mathbb{E}[S_T^2] = \mathbb{E}[T] = 14$. Somit ist die Varianz des Gewinnes 14.

3. a) Die gemeinsame Dichte von (X, Y) ist

$$f_{X,Y}(x, y) = f_X(x)f_{Y|X=x}(y) = \frac{1}{2\pi\sigma\eta} \exp\left\{-\frac{x^2}{2\sigma^2} - \frac{(y-x)^2}{2\eta^2}\right\}.$$

- b) Wir brauchen die Randdichte von Y , und müssen daher über x integrieren. Den Exponenten können wir als

$$\begin{aligned} \frac{x^2}{2\sigma^2} + \frac{(y-x)^2}{2\eta^2} &= \frac{(\sigma^2 + \eta^2)x^2 - 2\sigma^2xy + \sigma^2y^2}{2\sigma^2\eta^2} \\ &= \frac{(x - \frac{\sigma^2}{\sigma^2 + \eta^2}y)^2}{2\frac{\sigma^2\eta^2}{\sigma^2 + \eta^2}} + \frac{y^2}{2(\sigma^2 + \eta^2)} \end{aligned}$$

schreiben. Integrieren wir über x , erhalten wir die Randdichte

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \frac{1}{\sqrt{2\pi(\sigma^2 + \eta^2)}} \exp\left\{-\frac{y^2}{2(\sigma^2 + \eta^2)}\right\}.$$

Somit Y normalverteilt mit Mittelwert 0 und Varianz $\sigma^2 + \eta^2$. Für die bedingte Dichte von X erhalten wir

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{\sqrt{2\pi\frac{\sigma^2\eta^2}{\sigma^2 + \eta^2}}} \exp\left\{-\frac{(x - \frac{\sigma^2}{\sigma^2 + \eta^2}y)^2}{2\frac{\sigma^2\eta^2}{\sigma^2 + \eta^2}}\right\}.$$

Somit ist X bedingt normalverteilt mit Mittelwert $\frac{\sigma^2}{\sigma^2 + \eta^2}y$ und Varianz $\frac{\sigma^2\eta^2}{\sigma^2 + \eta^2}$.

- c) Für festes $Y = y$ ist der bedingte Mittelwert von X also $\frac{\sigma^2}{\sigma^2 + \eta^2}y$. Somit erhalten wir

$$\mathbb{E}[X | Y] = \frac{\sigma^2}{\sigma^2 + \eta^2}Y.$$

4. a) Die momentenerzeugende Funktion der Exponentialverteilung ist

$$M_X(r) = \int_0^{\infty} e^{rx} \alpha e^{-\alpha x} dx = \frac{\alpha}{\alpha - r}$$

für $r < \alpha$. Sonst ist die Funktion unendlich.

b) Kennen wir $N = n$, so ist

$$\mathbb{E} \left[\exp \left\{ r \sum_{k=1}^n X_k \right\} \right] = \prod_{k=1}^n \mathbb{E} [e^{rX_k}] = \left(\frac{\alpha}{\alpha - r} \right)^n .$$

Dies gilt auch für $n = 0$. Somit ist

$$M_S(r) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \left(\frac{\alpha}{\alpha - r} \right)^n = \exp \left\{ \lambda \frac{\alpha}{\alpha - r} - \lambda \right\} = \exp \left\{ \frac{\lambda r}{\alpha - r} \right\} .$$

c) Nach Korollar 2.19 ist

$$\mathbb{P}[S \geq c] \leq \exp \left\{ - \left(cr - \frac{\lambda r}{\alpha - r} \right) \right\} .$$

Wir möchten dies möglichst klein machen, also $g(r) = cr - \frac{\lambda r}{\alpha - r}$ möglichst gross. Die ersten beiden Ableitungen sind

$$g'(r) = c - \frac{\lambda \alpha}{(\alpha - r)^2} , \quad g''(r) = -\frac{2\lambda \alpha}{(\alpha - r)^3} < 0 .$$

Das Maximum wird für $r = \alpha - \sqrt{\lambda \alpha / c}$ angenommen (die andere Lösung ist grösser als α). Da $r \geq 0$ gelten soll, erhalten wir die Bedingung $c \geq \lambda / \alpha$. Setzen wir die Lösung für r ein, erhalten wir

$$\mathbb{P}[S \geq c] \leq \exp \left\{ - \left(c \left(\alpha - \sqrt{\frac{\lambda \alpha}{c}} \right) - \frac{\lambda \left(\alpha - \sqrt{\frac{\lambda \alpha}{c}} \right)}{\sqrt{\frac{\lambda \alpha}{c}}} \right) \right\} = \exp \left\{ - (\sqrt{\alpha c} - \sqrt{\lambda})^2 \right\} .$$

5. Nehmen wir den Logarithmus, erhalten wir

$$\varepsilon(x) = \ln f(x) + a(x - 1994) .$$

Wir haben $n = 14$ Daten, wenn wir 1994 nicht dazuzählen. Der Mittelwert der Variablen ist

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \varepsilon(k + 1994) &= \frac{1}{n} \sum_{k=1}^n \left[\ln \frac{f(k + 1994)}{104890} + ak \right] \\ &= a \frac{n+1}{2} + \frac{1}{n} \sum_{k=1}^n \left[\ln \frac{f(k + 1994)}{104890} \right] . \end{aligned}$$

Die linke Seite ist ein Schätzer für $\mathbb{E}[\varepsilon(x)] = 0$. Setzen wir die rechte Seite 0, so erhalten wir den Schätzer

$$\hat{a} = \frac{2}{n(n+1)} \sum_{k=1}^n \left[\ln \frac{104890}{f(k+1994)} \right] = 0.140362 .$$

Die Vorhersage für 2012 ist

$$\hat{f}(2012) = 104890 \exp\{-18a\} = 8384.5659 .$$

Wir rechnen also mit 8385 gestohlenen Autos für 2012.

Es wäre besser, ein Modell der Form

$$f(x) = \exp\{b - a(x - 1994) + \varepsilon(x)\}$$

zu verwenden. Dann könnte man den Ausdruck $\sum_{x=0}^{14} (\varepsilon(1994+x))^2$ minimieren, und so Schätzer \hat{a} , \hat{b} für a und b erhalten.

6. a) Da der Durchschnitt $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ eine suffiziente Statistik ist, können wir den Test auch mit dem Mittelwert konstruieren. Der volle Likelihood führt zum gleichen Test. Der optimale Test ist der Likelihood-Quotienten Test. \bar{X} ist normalverteilt mit Mittelwert 25 (bzw. 26) und Varianz $\frac{4}{n}$. Somit ist der Likelihood-Quotient

$$\begin{aligned} c &\leq \frac{\frac{1}{\sqrt{8\pi/n}} \exp\left\{-\frac{(x-26)^2}{8/n}\right\}}{\frac{1}{\sqrt{8\pi/n}} \exp\left\{-\frac{(x-25)^2}{8/n}\right\}} = \exp\left\{\frac{n}{8}[(x-25)^2 - (x-26)^2]\right\} \\ &= \exp\left\{\frac{n}{8}[2x-51]\right\} . \end{aligned}$$

Wir können also den kritischen Bereich als

$$K = \{n(2\bar{X} - 51) \geq 8\tilde{c}\} = \left\{ \sqrt{n} \frac{\bar{X} - 25}{\sqrt{4}} \geq \frac{2\tilde{c}}{\sqrt{n}} + 0.25\sqrt{n} \right\}$$

wählen. Für $\alpha = 2.5\%$ erhalten wir $\frac{2\tilde{c}}{\sqrt{n}} + 0.25\sqrt{n} = 1.96$, also $\tilde{c} = 0.98\sqrt{n} - 0.125n$. Somit wird der kritische Bereich

$$K = \{n(2\bar{X} - 51) \geq 7.84\sqrt{n} - n\} = \{\sqrt{n}(\bar{X} - 25) \geq 3.92\} .$$

- b) Für den Fehler zweiter Art, soll die Wahrscheinlichkeit des Ereignisses

$$\{\sqrt{n}(\bar{X} - 25) < 3.92\} = \left\{ \sqrt{n} \frac{\bar{X} - 26}{\sqrt{4}} < 1.96 - \frac{\sqrt{n}}{2} \right\}$$

höchstens 2.5% betragen. Also muss $-1.6449 \geq 1.96 - \frac{\sqrt{n}}{2}$ gelten. Damit erhalten wir $n \geq (7.2098)^2 = 51.9812$. Wir brauchen damit mindestens 52 Beobachtungen.