Fachhochschule Nordwestschweiz

Übungsblatt LE2

Lösungen

Lösung 1.

(a) 60.65% ($\lambda = 1/6000$ und $\exp(-\lambda * 3000) = 0.6065$)

(b) 47.86% ($\exp(-1/6000 * 1000) - \exp(-1/6000 * 6000) = 0.4786$)

Lösung 2.

(a) 28.65% ($\lambda = 1/8$ und $\exp(-\lambda \cdot 10) = 0.2865$)

(b) 63.21% ($1 - \exp(-1/8 * 8) = 0.6321$)

Lösung 3.

(a) $\lambda = 0.1$ ($\lambda = -\ln(1 - .2592)/3.$)

(b) 10 Jahre (=$1/\lambda$)

(c) Gleichung: $0.5 = 1 - \exp(-\lambda x)$
$-0.5 = -\exp(-\lambda x)$
$0.5 = \exp(-\lambda x)$
$\ln(0.5) = -\lambda x$
$\frac{\ln(0.5)}{-\lambda} = x$
$-\ln(0.5) \cdot 10 = 6.93$ Jahre

Lösung 4.

(a) 0

(b) In python: `import scipy.stats`
`scipy.stats.norm(125, 5).cdf(130)-scipy.stats.norm(125, 5).cdf(120)` = 0.68

(c) `scipy.stats.norm(125, 5).cdf(110)` = 0.0013

(d) `scipy.stats.norm.ppf(0.05, loc=125, scale=5)` = 116.8 g

Lösung 5.

`scipy.stats.norm.ppf(0.9, loc=80, scale=10)` = 92.8 kg

Lösung 6.

(a) $\mu = 32$ (=(24+34+32+36+38+32+28)/7) und $\sigma = 4.761$.

(b) Konfidenzintervall $[\mu - t\frac{\sigma}{\sqrt{n}}, \mu + t\frac{\sigma}{\sqrt{n}}]$. Man hat 7 Datenpunkte, d.h. 6 Freiheitsgrade. t schaut man in einer Tabelle nach, z.B. ttable.org für one-tail 0.025 (=$1 - \alpha/2$, wobei $\alpha =$ 0.95) bzw. two-tails 0.05: $t_{0.975,6} = 2.447$. $\rightarrow [32 - 2.447 \cdot 4.761/\sqrt{7}, 32 + 2.447 \cdot 4.761/\sqrt{7}] =$ $[27.597; 36.403]$ (t-Verteilung!)

(c) Analog zu vorheriger Lösung mit $t_{0.95,6} = 1.943$. $[28.504; 35.496]$ (t-Verteilung!)

Lösung 7.

Es ist $\bar{x} = 50.14$ und σ=0.3962

(a) $[50.14 - 1.895 * 0.396232/\sqrt{8}, 50.14 - 1.895 * 0.396232/\sqrt{8}] = [49.875; 50.405]$ (t-Verteilung!)

(b) $[50.14 - 5.408 * 0.396232/\sqrt{8}, 50.14 - 5.408 * 0.396232/\sqrt{8}] = [49.382; 50.898]$ (t-Verteilung!)

Lösung 8.
Eine Irrtumswahrscheinlichkeit von 4.55% heisst, dass man das z sucht, welches 1-.0455/2 = 0.97725 (zweiseitiger Test), also den Flächeninhalt unter der Normalverteilung von 97.7% ergibt. In einer Tabelle der Standardnormalverteilung findet man $\Phi_{0;1}(2) = 0.97725$ oder via Python `from scipy.stats import norm` und `norm.ppf(.97725)`, was 2 gibt.
Das Konfidenzintervall ist $[\bar{x} - z\frac{\sigma}{\sqrt{n}}, \bar{x} + z\frac{\sigma}{\sqrt{n}}] = [600000 - 2 \cdot 90000/\sqrt{225}, 600000 + 2 \cdot 90000/\sqrt{225}] =$ $[588000; 612000]$ (Standardnormalverteilung!)

Lösung 9.

(a) lohn $= 2.83 * m_i + 12.67$ ($m_i = 1$ für Mann und 0 sonst)

(b) $R^2 = 0.41$, also nicht ideal, aber nicht unkorreliert.

(c) Man benützt den Wert der t-Verteilung für $\alpha/2$=0.025 (für ein 95% Konfidenzintervall). Die Anzahl der Freiheitsgrade ist $n-2=12-2=10$. $t_{(1-\alpha/2)}(n-2)=2.228$ (aus der t Tabelle). Die Standardabweichung für die Steigung ist 1.078 (s. Fahrmeir, S.444 oder z.B. via OLS summary() in Python). Konfidenzintervall: $2.83 \pm 2.228 \cdot 1.078$, d.h. $[0.43, 5.23]$.

(d) Der Zweistichproben-t-Test verlangt identische Varianzen. Wir berechnen die Standardabweichungen des Lohns von Schülerinnen und Schülern und erhalten 1.37 und 1.98. Da sie nicht identisch sind, ist der Zweistichproben-t-Test nicht verwendbar.

Lösung 10.

(a) Da die Stichprobe klein ist, verwenden wir den t-Wert. Wir berechnen

$$t = \frac{3.25-3.5}{0.31}\sqrt{10} = -2.55$$

In der t-Tabelle verwenden wir den Wert bei 10-1=9 Freiheitsgraden und einem einseitigen Signifikanzniveau von 5%: $t_{9;0.95}=1.833$. Wir testen, ob $|t|=2.55$ grösser oder kleiner ist als der Tabellenwert. Hier ist $2.55 > 1.833$, deswegen ist die Aussage falsch ("Stichprobe der 10 Akkus nicht aus einer Grundgesamtheit von Akkus mit mindestens 3.5 Stunden Laufzeit"). Die Akkulaufzeiten erreichen also die angegebene Zahl nicht.

(b) Man rechnet analog zu a), aber mit $t_{9;0.99}=2.821$. Nun ist 2.55 kleiner als der Tabellenwert. In diesem Fall kann man nicht sagen, ob der Hersteller lügt. Es kann nämlich sein, dass die Stichprobe aus der Grundgesamtheit von Akkus mit mindestens 3.5 Stunden Laufzeit stammt. Dies ist kein Widerspruch zu vorher. In a) hat man eine Irrtumswahrscheinlichkeit von 5% zugelassen, d.h. von 100 Stichproben können 5 zum falschen Resultat führen. Deswegen konnte man sagen, dass der Hersteller wahrscheinlich gelogen hat (aber in bis zu 5% der Fälle hätte man ihm Unrecht getan). Hier lässt man nur in 1% der Fälle eine Falschaussage zu, deswegen kann man nicht ausschliessen, dass die Akkus doch länger halten könnten.

Lösung 11.

(a) Wir berechnen die Mittelwerte und Standardabweichungen der Stichprobe von roten und nicht-roten Autos:

$$\mu_{rot} = \frac{42+58+...}{5} = 51.6$$

$$\mu_{nicht-rot} = \frac{40+50+...}{5} = 49.2$$

$$\sigma^2_{rot} = \frac{1}{n-1}\sum_{i=1}^{5}(x_i-\mu_{rot})^2 = \frac{1}{4}((42-51.6)^2+...+(52-51.6)^2) \rightarrow \sigma_{rot} = 9.74$$

$$\sigma^2_{nicht-rot} = \frac{374.8}{4} = 93.7 \rightarrow \sigma_{nicht-rot} = 9.68$$

Wir verwenden einen Zweistichproben-t-Test, da die Varianzen fast identisch sind. Das gewogene Mittel der Varianz ist:

$$\frac{(4 \cdot 9.68^2 + 4 \cdot 9.74^2)}{10 - 2} = 9.71$$

Der t-Wert ist:

$$t = \frac{51.6 - 49.2}{9.71} \cdot \sqrt{25/10} = 0.4$$

Dies ist ein zweiseitiger Test, da die Geschwindigkeiten grösser oder kleiner sein können (es gibt keine Vorgabe, bei in welchen Fällen die Bussen verteilt werden). In der Tabelle lesen wir ab: $t_{8,0.95} = 2.3$

Das berechnete t ist kleiner als das t aus der Tabelle. Dies heisst, dass die roten Autos aus der Grundgesamtheit aller Autos stammen können und man keinen Unterschied feststellen kann. Die Polizistin kann also nicht sagen, dass rote Autos schneller fahren würden.