Lösung zu Aufgabe 10

a) a1) Ein zweiseitiges 95%-KI bei Normalverteilungsannahme mit bekanntem σ^2 hat die Gestalt

$$\left[\bar{X} - z_{0,975}\frac{\sigma}{\sqrt{n}}; \bar{X} + z_{0,975}\frac{\sigma}{\sqrt{n}}\right]$$

Dabei ist

* $\bar{X} \approx 3,734016$ das arithmetische Mittel der logarithmierten Genusszeiten $X_i = \log(Y_i)$. Berechnung mit R
* $\sigma^2 = \ln(1,1) \approx 0,3087235$
* $n = 20$ der Stichprobenumfang.
* $z_{0,975} = \Phi^{-1}(0,975) \approx 1,96$ das $97,5\%$-Quantil der Standardnormalverteilung[2]

Als konkretes Intervall ergibt sich damit
$I(X) = [3,734016 - 0,1353015; 3,734016 + 0,1353015] = [3,598715; 3,869318]$

a2) Ein zweiseitiges 95%-KI bei Normalverteilungsannahme mit unbekanntem σ^2 hat die Gestalt

$$\left[\bar{X} - t_{0,975}(n-1)\frac{S}{\sqrt{n}}; \bar{X} + t_{0,975}(n-1)\frac{S}{\sqrt{n}}\right]$$

Dabei ist

* $\bar{X} \approx 3,734016$ das arithmetische Mittel der logarithmierten Genusszeiten $X_i = \log(Y_i)$ wie in a)
* $S^2 \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2 \approx 0,1113083$ die Stichprobenvarianz bzw. $S \approx 0,3336291$ die Stichproben-Standardabweichung.
* $n = 20$ der Stichprobenumfang.
* $t_{0,975}(n-1) \approx 2,093024$ das $97,5\%$-Quantil der zentralen t-Verteilung mit n-Freiheitsgraden (Verteilung der ZV $\sqrt{n}\frac{X_0}{\sqrt{X_1^2 + \cdots X_2^2}}$, wenn $X_0, X_1, \ldots, X_n$ u.i.v. standardnormalverteilte ZV sind)

Als konkretes Intervall ergibt sich damit
$I(X) = [3,734016 - 0,1561432; 3,734016 + 0,1561432] = [3,577873; 3.890159]$
Man sieht, dass das KI breiter ist als bei bekannter geschätzer Varianz. Hier findet sich wieder, dass $\sqrt{\log(1,1)} < S$ ist und die Schätzung der unbekannten Varianz durch S ein anderes, größeres Quantil erfordert. Beide Effekte verbreitern das KI.

b) Laut VL hat unter der Normalverteilungsannahme $\mathcal{N}(\mu, \sigma^2)$ an die X_i die Statistik $T_n = \frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2$ eine Chi-Quadrat-Verteilung mit $n-1$ Freiheitsgraden, d.h. eine Verteilung mit der Wahrscheinlichkeitsdichte

$$f_{n-1}(z) = \begin{cases} \frac{z^{(n-1)/2-1}\exp(-z/2)}{2^{(n-1)/2}\Gamma((n-1)/2)} & z > 0 \\ 0 & z \leq 0 \end{cases},$$

[2] das $(1-\alpha)$-Quantil einer Verteilung wird auch als (oberes) α-Fraktil der Verteilung bezeichnet.

Bezeichnet man die Quantile mit q_α, so gilt $1-\alpha = P\left(q_{\alpha/2} \leq \frac{n-1}{\sigma^2}S^2 \leq q_{1-\alpha/2}\right)$

Stellt man beide Ungleichungen nach σ^2 um, so ergibt sich

$$1-\alpha = P\left(\frac{(n-1)S^2}{q_{1-\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_{\alpha/2}}\right)$$

D.h. ein $(1-\alpha)$-KI für σ^2 hat die Form $I_n = \left[\frac{(n-1)S^2}{q_{1-\alpha/2}} ; \frac{(n-1)S^2}{q_{\alpha/2}}\right]$

Für die logarithmierten Daten ergibt sich hier $S^2 = 0.1113083$, also $T = S^2 \times 19 \approx 2.114859$. Aus der Tabelle liest man für $\alpha = 0.05$ ab (in der Zeile für $n-1 = 19$ schauen):

- $q(0.975) = 32.852 \Rightarrow I_1 = 2.115/32.852 \approx 0.063$
- $q(0.025) = 8.907 \Rightarrow I_2 = 2.115/8.907 \approx 0.237$

Das KI hat also die Form $I_n = [0.063; 0.237]$.

Lösung zu Aufgabe 11

Aussage des ZGS: $X_1, .., X_n$ u.i.v. mit $\mathbb{E}X_i = \mu$, $Var(X_i) = \sigma^2$, dann

$P(\frac{\bar{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}} \leq y) \overset{n\to\infty}{\longrightarrow} \Phi(y)$, also näherungsweise standardnormalverteilt.

Es seien X_i Poisson-verteilt $(\lambda > 0)$. Dann

$\mathbb{E}X_i = \lambda, Var(X_i) = \lambda;$
$\mathbb{E}\bar{X}_n = \lambda, Var(\bar{X}_n) = \frac{\lambda}{n};$

Betrachten

$z := \frac{\bar{X}_n-\lambda}{\sqrt{\frac{\lambda}{n}}}$ näherungsweise $\mathcal{N}(0,1)$ (ZGS).

Forderung:

$P(\lambda \in I) = P(a \leq \lambda \leq b) \overset{!}{=} 1-\alpha.$

Das muss man sich jetzt aus:

$$1-\alpha = P(-c \leq \frac{\bar{X}_n-\lambda}{\sqrt{\frac{\lambda}{n}}} \leq c) = P(\frac{(\bar{X}_n-\lambda)^2}{\frac{\lambda}{n}} \leq c^2)$$

mit $c = z_{1-\frac{\alpha}{2}} \approx 1.96$ „zusammenbauen“.

Dazu gibt es zwei Ansätze:

a) Ersetze λ im Nenner durch einen Schätzer, d.h. durch $\bar{X}$, dann gilt die WS-Aussage immer noch approximativ

$$-1.96 \leq \frac{\bar{X}-\lambda}{\sqrt{\bar{X}/n}} \leq 1.96 \Leftrightarrow \bar{X} - 1.96\sqrt{\bar{X}/n} \leq \lambda \leq \bar{X} + 1.96\sqrt{\bar{X}/n}$$

Ergibt mit $n = 60$ und $\bar{X} = \frac{200}{60}$ das KI $[2.871; 3.795]$ mit der Breite 0.924

b) Löse die Ausgangsungleichung ohne Einsatz eines Schätzers für λ nach λ auf.

$$\frac{(\bar{X}_n-\lambda)^2}{\frac{\lambda}{n}} \leq c^2 \quad \Leftrightarrow \quad \bar{X}_n^2 - 2\lambda\bar{X}_n + \lambda^2 \leq \frac{c^2\lambda}{n}$$

$$\Leftrightarrow \quad \lambda^2 - 2\lambda(\bar{X}_n + \tfrac{c^2}{2n}) + (\bar{X}_n + \tfrac{c^2}{2n})^2 \leq -\bar{X}_n^2 + (\bar{X}_n + \tfrac{c^2}{2n})^2$$

$$\Leftrightarrow \quad [\lambda - (\bar{X}_n + \tfrac{c^2}{2n})]^2 \leq -\bar{X}_n^2 + \bar{X}_n^2 + \bar{X}_n\tfrac{c^2}{n} + \tfrac{c^4}{4n^2} = \frac{c^2(4\bar{X}_n n + c^2)}{4n^2}$$

$$\Rightarrow \quad -\frac{c\sqrt{n4\bar{X}_n+c^2}}{2n} \leq \lambda - \bar{X}_n - \frac{c^2}{2n} \leq \frac{c\sqrt{n4\bar{X}_n+c^2}}{2n}$$

Hinweis: Sie können diesen Schritt auch mit der pq-Formel lösen. Wähle also

$$I := [\bar{X}_n + \frac{c^2}{2n} - \frac{c\sqrt{n4\bar{X}_n+c^2}}{2n}; \bar{X}_n + \frac{c^2}{2n} + \frac{c\sqrt{n4\bar{X}_n+c^2}}{2n}] \ (*)$$

Es ist $c = u_{1-\frac{\alpha}{2}} = u_{97.5\%} = 1.96$

$\bar{X}_n = \frac{200}{60}, \quad n = 60$

Eingesetzt in (*):

$I \approx [2.902; 3.828]$. Das KI hat die Breite 0.931, ist also breiter als das erste.