Kurzlösungen
Übungsaufgaben Statistik II

Übung 9: Intervallschätzung

04.05.2010

Intervallschätzung

Aufgabe 1

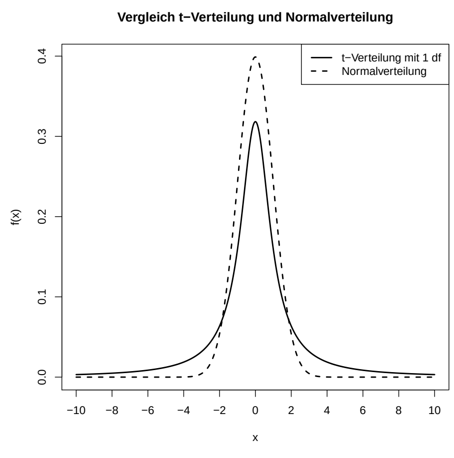
a) Konfidenzinterval für μ bei bekannten σ und Normalverteilung:
i) $[\overline{x} \mp z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}] = [5 \pm 1.96 \frac{3}{\sqrt{10}}] = [3.1406, 6.8594]$
UG = 3.1406
OG = 6.8594

ii) $[5 \pm 1.96 \frac{3}{\sqrt{100}}] = [4.412, 5.588]$
UG = 4.412
OG = 5.588

b) Stichprobenumfang gesucht:

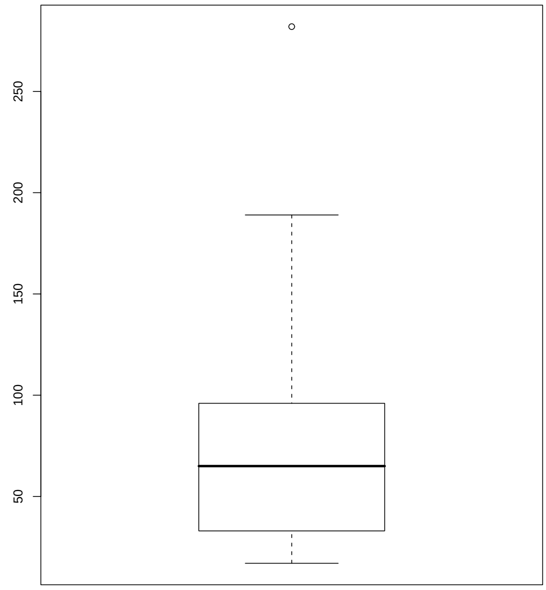
$$n \geq \left(\frac{2 \cdot z_{1-\frac{\alpha}{2}} \cdot \sigma}{L}\right)^2 = (\frac{2 \cdot 1.96 \cdot 3}{0.4})^2 = 864.36 \approx 865$$

c) Die Varianz wird geschätzt mit Stichprobenvarianz s^2. Durch die Schätzung der Varianz wird das Konfidenzintervall (KI) unpräziser. Es werden deshalb nicht die Quantile der Normalverteilung, sondern die der t-Verteilung für das KI genutzt. Das KI errechnet sich mit $[\overline{x} \mp t_{n-1,1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}]$. Die t-Verteilung hat mehr Wahrscheinlichkeitsmasse an den Rändern als die Normalverteilung.



Deshalb ist das Quantil der t-Verteilung stets grösser (wenn n nicht zu gross ist!) als das der Normalverteilung. Das KI wird somit breiter.

Aufgabe 2



a) Schiefe Daten und Ausreisser. Normalverteilungsannahme mit Vorbehalt.

b) Konfidenzintervall für μ bei unbekannten σ und Normalverteilung:
Aus den Daten müssen zuerst der Mittelwert und die Standardabweichung errechnet werden. Die Berechnung ergibt $\bar{x} = 88.7$ und $s = 83.79$.

$$[\overline{x} \mp t_{n-1,1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}] = [88.7 \pm 2.262 \frac{83.79}{\sqrt{10}}]$$
$$\mathbf{UG} = \mathbf{28.76}$$
$$\mathbf{OG} = \mathbf{148.64}$$

Der unbekannte Erwartungswert μ wird, unter Normalverteilung, mit einer Wahrscheinlichkeit von 95% von dem Konfidenzintervall mit den Grenzen 28.76 und 148.64 überdeckt.

Aufgabe 3

Die einzelnen Zufallsvariablen sind identisch unabhängig (wenn eine Person sich den Slogan merkt, hat das keinen Einfluss auf die nächste Person) Bernoulli verteilt:

$$
\begin{aligned}
X_i &= 1 \text{ Slogan wird nicht erkannt} \\
X_i &= 0 \text{ Slogan wird erkannt}
\end{aligned}
$$

Es gilt $E(X_i) = p$ und $Var(X_i) = p(1-p)$. Somit ist $Var(\hat{p}) = Var(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{p(1-p)}{n}$.

Die einzelnen Zufallsvariablen sind Bernoulli verteilt, da wir aber das Experiment 380 Mal wiederholen, ist der gesamte Prozess binomial verteilt. Y: Anzahl der Personen die sich den Slogan nicht merken konnten, $Y \sim B(n = 380, p)$

a) Punktschätzung für p: $\hat{p} = \bar{X} = \frac{123}{380}$

b) Konfidenzintervall für p. Approximation der Binomialverteilung durch die Normalverteilung (Bedingung $380\hat{p}(1-\hat{p}) \geq 9$ für $\hat{p}$ erfüllt).

i) 90% KI: $\alpha = 0.10$

$$
\begin{aligned}
\left[\hat{p} \mp z_{1-\frac{\alpha}{2}}\sqrt{Var(\hat{p})}\right] &= \left[\hat{p} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] \\
&= \left[\frac{123}{380} \pm 1.64\sqrt{\frac{\frac{123}{380}(1-\frac{123}{380})}{380}}\right] \\
&= [0.284, 0.363]
\end{aligned}
$$

UG = 0.284
OG = 0.363

ii) 95% KI: $\alpha = 0.05$

$$
\begin{aligned}
\left[\hat{p} \mp z_{1-\frac{\alpha}{2}}\sqrt{Var(\hat{p})}\right] &= \left[\frac{123}{380} \pm 1.96\sqrt{\frac{\frac{123}{380}(1-\frac{123}{380})}{380}}\right] \\
&= [0.277, 0.371]
\end{aligned}
$$

UG = 0.277
OG = 0.371

c) Intervalllänge nimmt zu (sichtbar auch in Teilaufgabe b).
Formel:

$$L = 2z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$z_{1-\frac{\alpha}{2}}$ nimmt zu, wenn das Konfidenzniveau zunimmt. Somit nimmt L zu, da $\hat{p}$ und n in diesem Beispiel gleich bleiben.

d) $L = 0.01$, wie gross muss dann n sein?

$$n \geq \left[\frac{2z_{1-\frac{\alpha}{2}}\sqrt{p(1-p)}}{L}\right]^2$$

Für $\hat{p} = \frac{123}{380}$ erhalten wir $n = 33'640$.