Prof. Josef Teichmann
David Martins

Wahrscheinlichkeit & Statistik

Serie 13

Die Übungen mit (*) markiert sind fakultativ.

Aufgabe 13.1 [Zweiseitiger z-Test]
Die durchschnittliche Fahrzeit von Zürich nach Bellinzona mit einem Intercity-Zug beträgt 146 Minuten. Mit dem Cisalpino werden die folgenden Zeiten gemessen:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
152	145	141	137	145	146	139	147	138

Wir nehmen an, dass diese Werte Realisierungen einer u.i.v. Stichprobe $X_1, \ldots, X_n$ sind mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$, wobei μ ein unbekannter Parameter und $\sigma^2 = 9$ ist. Führen Sie einen geeigneten Test durch, um auf dem 5%-Niveau festzustellen, ob die mittlere Fahrzeit des Cisalpino von jener des Intercity abweicht.

Aufgabe 13.2 [Konfidenzintervall für $\mathcal{N}(\mu, \sigma^2)$ mit bekannter Standardabweichung]
In einem Sägewerk wird das Schnittholz einer bestimmten Sortierklasse einer Qualitätskontrolle unterzogen. Pro Produktionstag wird eine Stichprobe mit zehn Brettern entnommen und jedes Brett auf seine Steifigkeit getestet. Basierend auf Erfahrung kann angenommen werden, dass die Steifigkeit eines Brettes normalverteilt ist mit bekannter Standardabweichung $\sigma = 1430$ MPa [1].

(a) Leiten Sie die Formel des 95%-Konfidenzintervalls für μ nach 15 Produktionstagen her.

(b) Berechnen Sie aus a) das realisierte Konfidenzintervall für einen beobachteten Stichprobenmittelwert von $\overline{x} = 11'000$ MPa (nach 15 Produktionstagen).

(c) Wie viele Stichproben wären nötig, damit die Breite des Konfidenzintervalls kleiner als 200 MPa ist?

Aufgabe 13.3 [Approximatives Konfidenzintervall I: Geometrische Verteilung]
Sei $\Theta = [1/2, 1]$. Wir betrachten die Modellfamilie $(\mathbb{P}_\theta)_{\theta \in \Theta}$, wobei $X_1, \ldots, X_n$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_1 \sim \text{Geom}(\theta)$. Der Maximum-Likelihood Schätzer für θ ist gegeben durch

$$T_{ML} = \frac{n}{\sum_{i=1}^n X_i}.$$

Bestimme ein approximatives Konfidenzintervall für θ mit Niveau 95%.
Hinweis: Für alle $\theta \in [1/2, 1]$ gilt $\frac{\sqrt{1-\theta}}{\theta} \leq \sqrt{2}$. Weiterhin wissen wir, dass $\mathrm{E}_\theta[X_1] = 1/\theta$ und $Var_\theta[X_1] = (1-\theta)/\theta^2$.

Aufgabe 13.4 [Approximatives Konfidenzintervall II: Binomialverteilung]
Um die Anzahl N der Forellen in einem See zu bestimmen, wird folgendermassen vorgegangen (Capture-Recapture-Methode): In einem ersten Schritt werden 500 Forellen gefangen, markiert und wieder ausgesetzt. In einem zweiten Schritt werden nochmals 200 Forellen gefangen und die Anzahl X der markierten Forellen bestimmt.

[1] Das Pascal ist eine abgeleitete SI-Einheit des Drucks sowie der mechanischen Spannung. Sie ist nach Blaise Pascal benannt und folgendermassen definiert: 1 Pa= $1\text{kg}\times\text{m}^{-1}\times\text{s}^{-2} = 1\text{N}\times\text{m}^{-2}$. Ein Pascal ist also der Druck, den eine Kraft von einem Newton auf eine Fläche von einem Quadratmeter ausübt. 1 MPa $= 10^6$ Pa $= 10$ bar.

(a) Für X wird oft eine Binomialverteilung angenommen, $X \sim \mathrm{Bin}(n, \theta)$ (θ bezeichne die Wahrscheinlichkeit, dass ein im zweiten Schritt gefangener Fisch markiert ist).
Wie gross ist n? Wie gross ist der Parameter θ, wenn die Gesamtzahl der Forellen im See $N = 2000$ bzw. $N = 5000$ ist?

(b) Die tatsächliche Beobachtung für X ergibt den Wert 40. Gebe eine vernünftige Schätzung für den Parameter θ an, und leite daraus eine Schätzung für die Gesamtzahl N der Forellen im See ab.

(c) Bestimme ein approximatives Konfidenzintervall für θ mit Level 95% und daraus ein approximatives Konfidenzintervall für N mit Level 95%.

Hinweis: Benutze für den Schätzer von θ den zentralen Grenzwertsatz.