Prof. Vincent Tassion
Laurin Köhler-Schindler

Wahrscheinlichkeit & Statistik

Serie 11

Abgabe bis Mittwoch (18.05.2022) um 10:15 Uhr

Diese Serie beschäftigt sich mit Konfidenzintervallen und der t-Verteilung.

Weitere Informationen und Instruktionen zur Abgabe unter
https://metaphor.ethz.ch/x/2022/fs/401-0614-00L/

Wichtiger Hinweis: Ein Konfidenzintervall für θ mit Niveau $1-\alpha$ ist ein Zufallsintervall $I = [A, B]$, sodass für alle $\theta \in \Theta$ gilt

$$\mathbb{P}_\theta[A \leq \theta \leq B] \geq 1 - \alpha.$$

In der Definition, die in der Vorlesung gegeben wurde, wurde α und $1-\alpha$ vertauscht. Im Skript wurde dies nun korrigiert.

Aufgabe 11.1 [Konfidenzintervall für $\mathcal{N}(\mu, \sigma^2)$ mit bekannter Standardabweichung]
In einem Sägewerk wird das Schnittholz einer bestimmten Sortierklasse einer Qualitätskontrolle unterzogen. An jedem Produktionstag wird ein Brett auf seine Steifigkeit getestet. $X_n, n \geq 1$, bezeichne das Ergebnis am n-ten Tag. Basierend auf Erfahrung kann angenommen werden, dass die Steifigkeit eines Brettes normalverteilt ist mit bekannter Standardabweichung $\sigma = 1430$ MPa[1]. Wir wählen also den Parameterraum $\Theta = \mathbb{R}$ und die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$, wobei die $(X_n)_{n\geq 1}$ unter $\mathbb{P}_\theta$ unabhängig und $\mathcal{N}(\theta, \sigma^2)$-verteilt sind.

(a) Leite die Formel eines Konfidenzintervalls für θ mit Niveau 95% nach 150 Produktionstagen her.

(b) Berechne aus (a) das realisierte Konfidenzintervall für einen beobachteten Stichprobenmittelwert von $11'000$ MPa (nach 150 Produktionstagen).

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

Aufgabe 11.2 [t-Verteilung mit m Freiheitsgraden]
Seien X und Y unabhängige Zufallsvariablen mit $X \sim \mathcal{N}(0, 1)$ und $Y \sim \chi^2_m$. Ziel dieser Aufgabe ist es zu zeigen, dass der Quotient

$$Z := \frac{X}{\sqrt{\frac{1}{m}Y}}$$

t-verteilt ist mit m Freiheitsgraden.

(a) Bestimme die Dichte von $Y' := \sqrt{\frac{1}{m}Y}$.

(b) Seien $\hat{X}$ und $\hat{Y}$ zwei unabhängige Zufallsvariablen mit Dichte $f_{\hat{X}}$ resp. $f_{\hat{Y}}$. Definiere $\hat{Z} := \frac{\hat{X}}{\hat{Y}}$ und bestimme die Dichte von $\hat{Z}$ (in Abhängigkeit von $f_{\hat{X}}$ und $f_{\hat{Y}}$).

Hinweis: Benutze in (a) und (b) die Charakterisierung der Dichte aus Proposition 4.16.

(c) Zeige, dass Z t-verteilt ist mit m Freiheitsgraden.

[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.

Aufgabe 11.3 [Konfidenzintervall für $\mathcal{N}(\mu, \sigma^2)$ mit unbekannter Standardabweichung]
Wir betrachten erneut die Situation aus Aufgabe 11.1 und nehmen weiterhin an, dass die Steifigkeit eines Brettes normalverteilt ist. In diesem Fall gehen wir aber davon aus, dass die Standardabweichung σ unbekannt ist. Der unbekannte Parameter $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$ ist in diesem Fall also 2-dimensional und wir wählen die Modellfamilie $(\mathbb{P}_\theta)_{\theta \in \Theta}$, wobei die $(X_n)_{n \geq 1}$ unter $\mathbb{P}_\theta$ unabhängig und $\mathcal{N}(\mu, \sigma^2)$-verteilt sind.
Zur Erinnerung: Aus der Vorlesung kennen wir bereits zwei Schätzer

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{und} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

für μ und σ^2.

(a) Zeige für jedes $\theta = (\mu, \sigma^2) \in \Theta$, dass die Zufallsvariable

$$\frac{\bar{X}_n - \mu}{\sqrt{S^2/n}}$$

t_{n-1}-verteilt ist unter $\mathbb{P}_\theta$.
Hinweis: Verwende $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$, $\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$, die Unabhängigkeit von $\bar{X}_n$ und S^2 sowie das Resultat aus der vorherigen Aufgabe.

(b) Leite die Formel eines Konfidenzintervalls für μ mit Niveau 95% nach 10 Produktionstagen her.