Blatt 5

Data Science 2

Sommersemester 2023

Aufgabe 1:

Sei $X_1, X_2, \ldots, X_n$ eine Stichprobe. Die Verteilung der X_i hänge von den unbekannten Paramtern $a, b \in \mathbb{R}, c > 0$ ab.

a) Sei $T(X_1, \ldots, X_n) := 2a + \frac{1}{n^2}(X_1^2 + X_2^2 + \ldots + X_n^2) \sim U(0, 10)$
b) Sei $T(X_1, \ldots, X_n) := c \cdot \min X_1, \ldots, X_n \sim Exp(1)$
c) Sei $T(X_1, \ldots, X_n) := (\prod_{i=1}^n X_i)^{\frac{1}{n}} \sim U(b, b+1)$
d) Sei $T(X_1, \ldots, X_n) := \frac{\bar{X}}{S} \sim Exp(c^2)$.

Geben Sie je Teilaufgabe ein zweiseitiges $(1-\alpha)$ %-Konfidenzintervall für den Parameter a, b oder c an.

Aufgabe 2:

Seien $X_1, X_2, \ldots, X_n$ unabhängige Zufallsvariablen mit Werten in $\mathbb{R}_{>0}$, Erwartungswert $\mu > 0$ und Varianz $\sigma^2 > 0$. Sei

$$f(X_1, \ldots, X_n) := \left(\prod_{i=1}^n X_i\right)^{\frac{1}{n}}$$

a) Stellen Sie einen Zusammenhang zwischen $f(X_1, \ldots, X_n)$ und dem ZGWS her. b) Bestimmen Sie analytisch ein 95%-Konfidenzintervall für μ auf Basis des ZGWS. Hierzu dürfen Sie $\sigma > 0$ als bekannt voraussetzen. c) Realisieren Sie $X_1, \ldots, X_n$ nun entsprechend einer Paretoverteilung mit Parametern $x_{min} = 5$ und $k = 3$. Geben Sie ein entsprechend b) bestimmtes Konfidenzintervall auf Basis dieser Realisierungen an. d) Realisieren Sie nun **erneut** paretoverteilte Zufallsvariablen entsprechend c). Teilen Sie die Realisierungen in 10 Gruppen ein und schätzen Sie den Erwartungswert der Realisierungen je Gruppe. Geben Sie anschließend den Anteil der Erwarungswerte innerhalb Ihres Konfidenzintervalls an. Passen Ihre Ergebnisse zur Konstruktion Ihres Konfidenzintervalls?

Hinweis: -

Lösung

Aufgabe 3: Ist Aufgabe b) einfach genug und so einfach wie ich sie mir vorstelle?

Seien $X_1, X_2, \ldots, X_n \sim N(0, \sigma^2)$ mit $\sigma > 0$ unbekannt.

a) Schätzen Sie σ^2 mit dem hierfür üblichen, erwartungstreuen Schätzer. Simulieren Sie dazu $X_1, \ldots, X_n$ für $\sigma^2 = 9$
b) Ermitteln Sie die approximative Verteilung von $S(X_1, \ldots, X_n)$ mit Hilfe des ZGWS und auf Basis von σ. Sie dürfen $E(X_i^4) = 3\sigma^4$ benutzen.
c) Bestimmen Sie ein zweiseitiges 95%-Konfidenzintervall für σ^2

d) Testen Sie Ihr Konfidenzintervall, indem Sie analog zu Aufgabe 2c und 2d verfahren.

Hinweis: -

Lösung

Aufgabe 4:

Seien $X_1, X_2, \ldots, X_n \sim U[a, b]$ unabhängige, identisch verteilte Zufallsvariablen mit $a < b$, wobei $a := 0$ und b unbekannt sei. Wir schätzen b mittels

$$\hat{b}(X_1, \ldots, X_n) := max(X_1, \ldots, X_n)$$

a) Berechnen Sie die Verteilungsfunktion

$$\hat{b}$$

. Hinweis: Für alle $t \in \mathbb{R}$ gilt:

$$max(X_1, \ldots, X_n) \leq t \Leftrightarrow X_1 \leq t, \ldots X_n \leq t$$

.

b) Ist der Schätzer

$$\hat{b}$$

erwartungstreu für b? Hinweis:

$$E(\hat{b}) = \int_0^\infty P(\hat{b} \geq t)dt$$

c) Ergänzen Sie

$$\hat{b}$$

, sodass ein erwartungstreuer Schätzer entsteht. Bedenken Sie, dass ein Schätzer nur auf $X_1, \ldots, X_n$ und bekannten Werten beruhen darf.

d) Simulieren Sie $X_1, \ldots, X_n$ für $b := 10$. Führen Sie anschließend jeweils 100 Schätzungen mittels des erwartungstreuen und des verzerrten Schätzers durch. Berechnen Sie den mittleren quadratischen Fehler bezüglich des wahren Werts von $b = 10$ zuerst unter Nutzung des erwartungstreuen Schätzers, dann unter Nutzung des verzerrten Schätzers. Welcher Fehler ist größer?

Hinweis: -

Lösung

Aufgabe 5: Sind hier Dopplungen enthalten? se() nachschlagen.

Seien $X_1, X_2, \ldots, X_n \sim Exp(\lambda)$ mit $\lambda > 0$.

a) Erzeugen Sie Realisationen für $\lambda := 10$.
b) Schätzen Sie $\frac{1}{\lambda}$ und $\frac{1}{\lambda^2}$ jeweils 10 mal auf Basis von 1 000 Realisationen mittels $\bar{X}$
c) Schätzen Sie $\frac{1}{\lambda}$ und $\frac{1}{\lambda^2}$ jeweils 10 mal auf Basis von 1 000 Realisationen mittels S <— Ist klar welches S gemeint ist?
d) Bestimmen Sie $se(\bar{X})$, $se(S)$ sowie den mittleren quadratischen Fehler all dieser Schätzer zum wahren Wert von λ.
e) Ist $\bar{X}$ erwartungstreu für λ? Ist S erwartungstreu für $\frac{1}{\lambda}$?

Hinweis: -

Lösung
