Übungen zur Vorlesung „Stochastik für die Informatik“

Abgabe der Lösungen zu den S-Aufgaben: Dienstag, 21. Dezember 2016, vor der Vorlesung (12:20-12:30 im Magnus HS)

29. S a) Z sei standard-normalverteilt. Bestimmen Sie $c > 0$ so, dass für das Intervall $I = [-c, c]$ gilt: $\mathbf{P}(Z \in I) = 0.80$. (Hinweis: Der R-Befehl `qnorm(p)` liefert $\Phi^{-1}(p)$. Mehr Informationen darüber bekommen Sie, wenn Sie `?qnorm` in die R-Konsole eingeben.)
b) Y sei $N(0, \sigma^2)$-verteilt. Bestimmen Sie $c > 0$ so, dass für das Intervall $I = [-c, c]$ gilt: $\mathbf{P}(Y \in I) = 0.80$.
c) Wie groß muss der Stichprobenumfang mindestens sein, damit das Intervall mit den Grenzen $M \pm 2$ den Populationsmittelwert mit Wahrscheinlichkeit
(i) 0.99
(ii) 0.9
(iii) 0.8
überdeckt? Dabei sei M der (zufällige) Stichprobenmittwelwert, und es sei bekannt, dass die Populationsvarianz den Wert 100 beträgt. Rechnen Sie mit der Normalapproximation.

30. Wir betrachten die Situation von Aufgabe 19. Es sei μ der in 19a) berechnete Populationsmittelwert, σ^2 die in 19a) berechnete Populationsvarianz, und $X_1, \dots, X_{50}$ die zufälligen Werte, die bei einem 50-maligen Ziehen ohne Zurücklegen entstehen (vgl. 19b)).
a) Warum sind die X_i nicht unabhängig?
b) Wir setzen $M := \frac{1}{50}(X_1 + \cdots + X_{50})$. Es sei Ihnen verraten, dass trotz der fehlenden Unabhängigkeit der X_i auch hier die asymptotische Normalität greift, vgl. dazu den Ausblick auf der vorletzten Folie von VL 7b). Deshalb dürfen Sie im Rest der Aufgabe mit der Normalapproximation rechnen, Sie sollten dabei aber die Standardabweichung von M verwenden, die wir in Aufgabe 19 berechnet haben (jetzt mit $n = 50$ anstelle des dortigen $n = 10$).
(i) Für welche Zahl δ ist die Wahrscheinlichkeit, dass M um mehr als δ von μ abweicht, ungefähr gleich 0.05?
(ii) Geben Sie ein um M zentriertes Intervall I an, sodass $\mathbf{P}(\mu \in I) \approx 0.95$.[1]
c) Überprüfen Sie, mit welcher Wahrscheinlichkeit das in b) ii) berechnete Intervall dem Populationsmittelwert μ überdeckt, indem Sie den Rechner viele Stichproben der Größe 50 ziehen lassen. Verwenden Sie dazu das unter dem Link A30.R von Benjamin Straub bereitgestellte R-Programm.

31 S a) Z_1, Z_2 und Z_3 seien unabhängig und $N(0, 1)$-verteilt. Es sei $X := 2Z_1 + 3Z_2$ und $Y := 6Z_2 + Z_3$. Berechnen Sie diejenigen Zahlen β_0 und β_1, für die der erwartete quadratische Abstand zwischen den beiden Zufallsvariablen $\beta_0 + \beta_1 X$ und Y minimal wird.

b) X und Y seien wie in a). Berechnen Sie die Regressionsgerade für $Y + 3$ auf der Basis von $X + 2$.

32. a) (X, Y) sei uniform verteilt auf der dreielementigen Menge $\{(0, 0), (1, 0), (3, 2)\} (\subset \mathbb{R}^2)$. Berechnen Sie die Regressionsgerade von Y auf der Basis von X.

b) Es sei $x_1 = 0$, $x_2 = 1$, $x_3 = 3$ sowie $y_1 = y_2 = 0$, $y_3 = 2$. Für welche Zahlen β_0 und β_1 wird $\sum_{i=1}^{3} (y_i - \beta_0 - \beta_1 x_i)^2$ minimal?

[1] Ein zufälliges Intervall I mit der Eigenschaft $\mathbf{P}(\mu \in I) \approx 0.95$ heißt *Konfidenzintervall für μ, approximativ zum Niveau* 0.95.