Übungen zur Vorlesung „Stochastik für die Informatik“

Abgabe der Lösungen zu den S-Aufgaben: Freitag, 31. Januar 2020, vor der Vorlesung (12:10-12:15 im Magnus HS)

45. S a) Es geht darum, in einer großen Population den Anteil p der Individuen mit einem bestimmten Merkmal mit einem approximativen 99%-Konfidenzniveau auf $\pm 2\%$ genau zu schätzen. Wie groß muss dafür der Stichprobenumfang n mindestens sein? Finden Sie jeweils eine möglichst gute untere Schranke für n,
(i) die vom Anteilschätzer $\hat{p}$ abhängt,
(ii) die nicht von $\hat{p}$ abhängt.
b) Bei 28% der Individuen einer großen Population wird ein bestimmtes Merkmal festgestellt. Wie groß muss die Stichprobengröße n mindestens sein, damit die Hypothese "der Anteil des Merkmals in der Population ist 30 %" mit einem p-Wert 0.01 abgelehnt werden kann?

Verwenden Sie in beiden Aufgabenteilen die asymptotische Normalität des Anteilschätzers $\hat{p}$.

46. S Aus zwei reellwertigen Stichproben $x_1, x_2, \dots, x_{100}$ und $y_1, y_2, \dots, y_{200}$, die aus zwei großen Populationen $\mathcal{P}_1$ und $\mathcal{P}_2$ rein zufällig gezogen wurden, ergaben sich die Stichprobenmittelwerte $\bar{x} = 10$ und $\bar{y} = 11$ sowie die Stichprobenstandardabweichungen $s_x = 2.5$ und $s_y = 3.0$.

a) Was ist die geschätzte Standardabweichung
(i) des zufälligen Stichprobenmittelwertes M_X
(ii) des zufälligen Stichprobenmittelwertes M_Y
(iii) der Differenz $M_X - M_Y$?

b) Geben Sie ein approximatives 99%-Konfidenzintervall für die Differenz der beiden Populationsmittelwerte an.

c) Zu welchem p-Wert können Sie die Hypothese ablehnen, dass die beiden Populationsmittelwerte gleich sind?

47. Wir betrachten die in Aufgabe 13 beschriebene Situation. Zu welchem p-Wert lässt sich die Hypothese, das das Ziehen der Stichprobe rein zufällig erfolgte, ablehnen unter Verwendung

(i) von Fishers exaktem Test

(ii) der asymptotischen Normalität des Anteilschätzers einschließlich der Varianzkorrektur für endliche Populationen?

48. $\mathcal{P}$ sei eine Menge bestehend aus 40 reellen Werten, und T sei eine zufällige vierelementige Teilmenge von $\mathcal{P}$. Sie bekommen gesagt, dass T aus dem kleinsten, zweitkleinsten, viert- und achtkleinsten Element von $\mathcal{P}$ besteht. Zu welchem p-Wert können Sie unter Verwendung des Wilcoxon-Rangsummentests die Hypothese der reinen Zufälligkeit der Wahl von T (zugunsten einer Tendenz "hin zum Rand von $\mathcal{P}$") verwerfen?

Zusatzaufgabe.[1] $X_1, \dots, X_{10}$ seien reellwertige unabhängig und identisch verteilte Zufallsvariable. Ihre Verteilung ρ besitze eine Dichte. Berechnen Sie dier Wahrscheinlichkeit, mit der das Intervall $[X_{(2)}, X_{(9)}]$ den Median von ρ überdeckt.

[1] Ein schriftliches Bearbeiten dieser Aufgabe ist optional. Es wird wie auch die anderen S-Aufgaben auf die Bonuspunkte angerechnet.