Institut für Angewandte Mathematik
SS 2017
Patrik Ferrari, Eva Kopfer & Angelo Profeta

5. Übungsblatt „Einführung in die Statistik"

Abgabe: Mittwoch 31.5.2017 in der Vorlesung

---

1. (Sensitivität von Stichprobenmittel und -median) *[5 Pkt]*

Sei $T_n : \mathbb{R}^n \to \mathbb{R}$ eine Statistik, welche n reellen Beobachtungswerten einen "Mittelwert" zuordnet. Wie stark T_n bei gegebenen Werten $x_1, \ldots, x_n \in \mathbb{R}$ von einer Einzelbeobachtung $x \in \mathbb{R}$ abhängt, wird beschrieben durch die Sensitivitätsfunktion

$$S_n(x) \ = \ n\Big(T_n(x_1, \ldots, x_{n-1}, x) - T_{n-1}(x_1, \ldots, x_{n-1})\Big).$$

Bestimmen und zeichnen Sie S_n in den Fällen, wenn T_n a) der Stichprobenmittelwert, b) der Stichprobenmedian und c) das α-getrimmte Mittel

$$\big(x_{\lfloor n\alpha\rfloor+1:n} + \cdots + x_{n-\lfloor n\alpha\rfloor:n}\big)\big/\big(n - 2\lfloor n\alpha\rfloor\big)$$

zu einem Trimm-Niveau $0 \le \alpha < 1/2$ ist. Hier bezeichnet $x_{k:n}$ die k-te Ordnungsstatistik von $x_1, \ldots, x_n$.

2. (Verteilungsdichte von Ordnungsstatistiken) *[5 Pkt]*

Seien $X_1, \ldots, X_n$ unabhängige, identisch verteilte relle Zufallsvariablen mit stetiger Verteilungsdichte ϱ. Weiterhin sei die Menge $\mathcal{X} = \{x \in \mathbb{R} \,|\, \varrho(x) > 0\}$ ein Intervall. Bestimmen Sie die Verteilungsdichte der k-ten Ordnungsstatistik.

Hinweis: Beweisen Sie zunächst, dass für eine Zufallsvariable X mit Verteilungsfunktion F_X gilt, dass $F_X(X) \sim \mathrm{Unif}_{[0,1]}$.

3. (Normalapproximation der Chiquadrat-Quantile) *[5 Pkt]*

Für $0 < \alpha < 1$ und $n \in \mathbb{N}$ sei $\chi^2_{n;\alpha}$ das α-Quantil der $\mathrm{Chi}^2(n)$-Verteilung and $\Phi^{-1}_{0,1}(\alpha)$ das α-Quantil der Standardnormalverteilung. Zeigen Sie, dass $(\chi^2_{n;\alpha} - n)/\sqrt{2n} \to \Phi^{-1}_{0,1}(\alpha)$ für $n \to \infty$.

P. (Statistik von Lottozahlen) [*Extra 5 Pkt*]

Unter dem nachfolgend aufgeführten Link finden Sie die Häufigkeiten der Zahlen 1 bis 49 für sämtliche der 4854 Ziehungen der Lottozahlen des Spiels 6 aus 49 im Zeitraum von 1955 bis 2010 (File auf der Homepage)

Zahl	**Häufigkeit**	**Zahl**	**Häufigkeit**	**Zahl**	**Häufigkeit**
1	576	18	569	35	566
2	574	19	562	36	569
3	591	20	537	37	565
4	569	21	555	38	616
5	566	22	566	39	570
6	593	23	564	40	571
7	570	24	557	41	585
8	543	25	605	42	574
9	589	26	613	43	602
10	553	27	606	44	570
11	560	28	530	45	531
12	552	29	541	46	549
13	506	30	563	47	567
14	559	31	601	48	579
15	534	32	566	49	601
16	563	33	589		
17	567	34	560		

Abbildung 1: Lottozahlen in 4644 Ziehungen des Lotto 6 aus 49

a) Bestimmen Sie für jede einzelne der Zahlen ein Konfidenzintervall zum Irrtumsniveau $\alpha = 0.05$ hinsichtlich der Wahrscheinlichkeit, diese Zahl in einer Ziehung zu sehen. Stellen Sie die Konfidenzintervalle zusammen mit den Mittelwerten und dem theoretischen Wert graphisch dar.

b) Schreiben Sie eine Funktion mit dem Input-Argument n, die die Häufigkeitsverteilung der Zahlen 1 bis 49 für n Lottoziehungen simuliert.

c) Simulieren Sie nun $n = 10^2, 10^3, 10^4$ Lottozahlen und stellen Sie, wie im *Aufgabenteil a)*, die Konfidenzintervalle, die Mittelwerte und den theoretischen Wert wieder für jede einzelne Zahl graphisch dar.