Angewandte Stochastik Blatt 6

1. (Erwartungswert)

a) Sei X eine Zufallsvariable mit Werten in $\mathbb{Z}_+$. Zeige:

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} P[X \geq k].$$

(Eventuell sind beide Seiten $+\infty$.)
Sei X eine beliebige Zufallsvariable mit Werten in $[0, \infty)$. Zeige:

$$\mathbb{E}(X) = \int_0^{\infty} P[X \geq s] ds.$$

(Wieder können beide Seiten $+\infty$ sein.)

b) In einer Lostrommel befinden sich N Lose mit den Nummern $1, 2, \ldots, N$; N ist unbekannt. Der kleine Fritz will wissen, wieviele Lose sich in der Trommel befinden und entnimmt in einem unbeobachteten Augenblick ein Los, merkt sich die aufgedruckte Nummer und legt es wieder in die Trommel zurück. Das macht er n Mal. Berechne aus den gemerkten Nummern $X_1, \ldots, X_n$ einen Maximum-Likelihood-Schätzer T für N. Ist dieser erwartungstreu ?

c) Berechne approximativ für großes N den relativen Erwartungswert $\mathbb{E}_N(T)/N$. Betrachte dazu einen geeigneten Ausdruck als Riemann-Summe.

2. (Randomized Response) Um bei Umfragen zu heiklen Themen ("Nehmen Sie harte Drogen ?") die Privatsphäre der befragten Personen zu schützen und zuverlässige Antworten zu bekommen, wurde folgendes "Unrelated Question"-Befragungsmodell vorgeschlagen: Ein Stapel Fragekarten ist zut Hälfte mit der heiklen Frage A und zur anderen Hälfte mit einer harmlosen Frage B beschriftet, welche nichts mit der Frage A zu tun hat ("Waren Sie letzte Woche im Kino ?"). Der Interviewer läßt den Befragten die Karten mischen, eine Karte verdeckt ziehen und die darauf gestellte Frage beantworten. Die untersuchte Personengruppe enthalte einen bekannten Anteil p_B der Personen, welche Frage B bejahen (Kinogänger). Sei $\theta = p_A$ die Wahrscheinlichkeit, mit der die heikle Frage bejaht wird. Es werden n Personen unabhängig befragt. Präzisiere das Modell, gib einen erwartungstreuen Schätzer für θ an und bestimme dessen Varianz.

3. (Normalapproximation der Binomialverteilung) Wir nennen zwei Folgen (a_n) und (b_n) asymptotisch äquivalent und schreiben dafür $a_n \sim b_n$, wenn

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

Die *Stirlingsche Formel* besagt, dass

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n}.$$

a) Sei $p \in (0,1)$ und sei (k_n) eine Folge mit $k_n \sim np$. Zeige mit Hilfe der Stirlingschen Formel, dass für die Binomialverteilung gilt:

$$b_{n,p}(k_n) \sim \sqrt{\frac{n}{2\pi k_n(n-k_n)}} \left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1-p)}{n-k_n}\right)^{n-k_n}.$$

b) Betrachte

$$-\ln\left[\left(\frac{np}{k_n}\right)^{k_n} \left(\frac{n(1-p)}{n-k_n}\right)^{n-k_n}\right] = ng(t_n),$$

wobei

$$t_n = \frac{k_n}{n},\ g(t) = t\ln\left(\frac{t}{p}\right) + (1-t)\ln\left(\frac{1-t}{1-p}\right).$$

Zeige durch Untersuchung der Taylorentwicklung von g an der Stelle p, dass für Folgen (k_n) mit $n(k_n/n - p)^3 \to 0$ gilt:

$$e^{-ng(t_n)} \sim e^{-\frac{(k_n-np)^2}{2np(1-p)}},$$

und daher

$$b_{n,p}(k_n) \sim \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k_n-np)^2}{2np(1-p)}}.$$