Statistik für Informatiker **SS 2017**

Prof. Dr. Matthias Birkner
M.Sc. Frederik Klement

Blatt 6

Aufgabe 1 (2 + 2 Punkte)

a): Es sei X geometrischverteilt mit Parameter $p \in (0,1)$. Es seien $n, k \in \mathbb{N}$, bestimmen Sie $P(X = n + k | X \geq n)$.

b): Es sei $N \sim \mathcal{N}_{0,1}$ normalverteilt. Sei $a > 0$, bestimmen Sie $P(N \in [0, a] | N^2 \in [0, a^2])$.

Aufgabe 2 (2 + 2 + 2 Punkte)
Von allen in Deutschland lebenden Personen tragen 10% den Virus $\mathcal{V}_1$, 15% den Virus $\mathcal{V}_2$ und 5% beide Virusarten in sich. Ist X nun eine zufällig ausgewählte Person, so nehmen wir an, dass

$$P(X \in V_1) = 0.1, \qquad P(X \in V_2) = 0.15, \qquad P(X \in V_1 \cap V_2) = 0.05,$$

hierbei steht $\{X \in V_i\}$ für das Ereignis, dass die ausgewählte Person das Virus $\mathcal{V}_i$ in sich trägt. Ein Pharmaziekonzern hat einen medizinischen Test T entwickelt, der je nach Virus mit folgenden Wahrscheinlichkeiten positiv ausfällt, d.h. $\{T = 1\}$,:

$$P(T = 1 | X \in V_1 \cap V_2^c) = 0.9, \qquad P(T = 1 | X \in V_1 \cap V_2) = 0.8,$$
$$P(T = 1 | X \in V_1^c \cap V_2) = 0.75, \qquad P(T = 1 | X \notin V_1 \cup V_2) = 0.1.$$

a): Bestimmen Sie die Wahrscheinlichkeit, dass der Test T positiv ausfällt.

b): Bestimmen Sie die Wahrscheinlichkeit, dass die Person X den Virus $\mathcal{V}_1$ in sich trägt, falls der Test positiv ausfällt.

c): Bestimmen Sie die Wahrscheinlichkeit, dass die Person X mindestens einen Virus in sich trägt, falls der Test positiv ausfällt.

Aufgabe 3 (2 + 2 + 2 Punkte)
Sie haben in der Vorlesung gelernt, dass die Verteilungsfunktion F_X einer Zufallsvariable X gegeben ist durch

$$F_X : \mathbb{R} \to [0,1], F_X(x) = P((-\infty, x]).$$

Die Quantilfunktion F_X^{-1} ist gegeben durch

$$F_X^{-1} : [0,1] \to \mathbb{R} \cup \{\infty\}, F_X^{-1}(t) := \inf\{x : F_X(x) \geq t\}.$$

Diese erfüllt die Eigenschaft, dass für alle $x \in \mathbb{R}$ und $t \in [0,1]$ gilt

$$F_X^{-1}(t) \leq x \Leftrightarrow t \leq F_X(x). \tag{1}$$

Sei nun U eine auf $[0,1]$ Zufallsvariable, dann besitzt $\tilde{X} = F_X^{-1}(U)$ die Verteilung von X, denn aufgrund von (1) gilt

$$P(\tilde{X} \leq x) = P(U \leq F_X^{-1}(x)) = F_X(x).$$

Somit erlaubt uns die Relation (1) Zufallsvariablen mit einer gewünschten Verteilung aus uniformverteilten Zufallsvariablen zu generieren, dies wird auch Inversionsmethode genannt. Verwenden Sie die Inversionsmethode, um exponentialverteilte Zufallsvariablen zu simulieren.

a): Schreiben Sie dazu zuerst eine Funktion `quantil_exp(t,`λ`)`, welche für $t \in [0,1)$ and $\lambda > 0$ den Wert $F_X^{-1}(t)$ mit $X \sim \mathrm{Exp}_\lambda$ berechnet (ohne Verwendung von *qexp* natürlich).
Hinweis: Für F_X^{-1} siehe Bsp.1.32 aus dem Vorlesungsskript.

Für eine Stichprobe $x_1, x_2, ..., x_n$ ist die empirische Verteilungsfunktion $\hat{F}_x$ gegeben durch

$$\hat{F}_x(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{[x_i,\infty)}(t).$$

Für einen festen Wert $t \in \mathbb{R}$ entspricht $\hat{F}_x(t)$ also der Zahl der Stichproben x_i, welche kleiner gleich t sind, geteilt durch die Gesamtzahl n aller Stichproben.

b): Schreiben Sie die Funktion `empVF(x,t)`, welche den Wert $\hat{F}_x(t)$ berechnet für die Stichprobe `x` $= x$ und den Wert `t` $= t$.

c): Simulieren Sie per Inversionmethode 1000 Zufallsvariablen für $\lambda = \frac{1}{2}$ und vergleichen Sie die daraus gewonnene empirische Verteilungsfunktion mit der theoretischen Verteilungsfunktion F_X, indem Sie beide zusammen in einen Plot mit unterschiedlichen Farben zeichnen lassen. Verwenden Sie hierbei als x-Achse die Werte von 0 bis 10 mit Schrittlänge 10^{-1}. Ergänzen Sie den Plot um eine passende Beschriftung der Achsen, eine Legende und einen Titel.

Achtung: Zusätzlich zur elektronischen Abgabe per Mail, drucken Sie bitte Ihren Code aus und reichen Sie diesen zusammen mit Ihrer restlichen Abgabe ein.

Abgabe: 10 Uhr 09.06.2017