Statistik für Informatiker **SS 2017**

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

Blatt 8

Aufgabe 1 (1 + 2 + 2 + 2 Punkte)
Seien $X_1, X_2, ..., X_n$ unabhängige Zufallsvariablen mit $X_i \sim \mathrm{Ber}_p$, wobei $p \in (0,1)$. Zeigen Sie für alle $a \in (p, 1)$, dass

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right) \leq e^{-na\log\frac{a}{p} - n(1-a)\log\frac{1-a}{1-p}}. \tag{1}$$

a): Zeigen Sie zuerst, dass für alle $s > 0$

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right) = P\left(e^{\sum\limits_{i=1}^{n} sX_i - san} \geq 1\right). \tag{2}$$

b): Schätzen Sie (2) mit einer geeigneten Ungleichung nach oben ab, und folgern Sie, dass für alle $s > 0$

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right) \leq e^{-nas}\mathbb{E}[e^{sX_1}]^n = f(s)^n,$$

wobei $f(s) = pe^{(1-a)s} + (1-p)e^{-as}$. **Hinweis:** Benutzen Sie ohne Beweis, dass $g(U_1), ..., g(U_n)$ unabhängig sind, wenn $U_1, ..., U_n$ unabhängig sind.

c): Bestimmen Sie das Minimum von f auf $s > 0$.

d): Kombinieren Sie Ihre bisherigen Ergebnisse und zeigen Sie (1).

Aufgabe 2 (1 + 3 + 1 Punkte)
Ihr Freund, ein begeisterter Roulette-Spieler, verschleppt Sie ins Kasino, genauer in die Spielbank in Wiesbaden. Umgeben von Herren in feinen Smokings, stehen Sie nun mit Ihrem Freund am Roulettetisch. In jeder Runde wird durch das Drehen der Roulettescheibe und dem Einsatz einer Kugel eine Zahl zwischen 0 und 36 gezogen, wobei jede Zahl mit gleicher Wahrscheinlichkeit gezogen wird. Die Zahlen **1, 3, 5, 7, 9, 12, 14, 16, 18, 19, 21, 23, 25, 27, 30, 32, 34, 36** sind rot und die Zahlen **2, 4, 6, 8, 10, 11, 13, 15, 17, 20, 22, 24, 26, 28, 29, 31, 33, 35** schwarz gefärbt, die **0** ist grün. Es gibt folgende Setzmöglichkeiten:

	Gewinnzahlen	Gewinnauszahlung
Rouge	Die roten Zahlen	2×Einsatz
Impair	Die ungeraden Zahlen	2×Einsatz
Zéro	Die Null	36×Einsatz

Ihr Freund spielt mit Jetons im Wert von jeweils einen Euro.

a): Da er nicht weiß, wie er setzen soll, wirft Ihr Freund eine faire Münze. Bei Kopf setzt er 3 Jetons auf Rouge und bei Zahl einen auf Zéro. Berechnen Sie den zu erwartenden Reingewinn Ihres Freundes. (Reingewinn = gewonne Jetons − gesetzte Jetons).

b): Ab jetzt spielt Ihr Freund nur noch Rouge, wobei er 3 Jetons setzt. Monsieur Dostojewski, ein gemeinsamer Bekannter von Ihnen, spielt am selben Tisch und setzt auf Impair. Sein Einsatz beträgt 5 Jetons. Berechnen Sie die Kovarianz und Korrelation zwischen dem Reingewinn Ihres Freundes und dem von Monsieur Dostojewski. **Hinweis:** Berechnen Sie die Kovarianz und die Korrelation erst für den Fall, dasss beide Spieler nur einen Jeton setzen. Nutzen Sie dann die Bilinearität der Kovarianz aus.

c): Ihr Freund spielt weiterhin nur Rouge. Er überlegt, ob er in einer Runde 10 Jetons auf einmal setzen soll, oder in zehn aufeinander folgenden Runden jeweils einen. Berechnen Sie den Erwartungswert und die Varianz des Reingewinns ihres Freundes für beide Varianten. Nehmen Sie hierbei an, dass die Ergebnisse der unterschiedlichen Runden stochastisch unabhängig sind.

Aufgabe 3 (1 + 1 + 1 + 1 Punkte)
Sie verwalten einen Webserver, der in jeder Sekunde eine zufällige Anzahl von neuen Anfragen erhält und der in jeder Sekunde eine zufällige Anzahl der bisherigen Anfragen bearbeiten kann. Wir modellieren dies durch folgendes, vereinfachtes Modell: Sei S_n die Zahl der noch zu bearbeitenden Anfragen in der n-ten Sekunde, dann gilt

$$S_0 = s_0; \quad S_{n+1} := \max\{S_n + X_n - Y_n, 0\},$$

hierbei sind $X_1, X_2, ...,$ und $Y_1, Y_2, ...$ zwei Folgen von unabhängigen Zufallsvariablen, wobei $X_i \sim \mathrm{Bin}_{p_1,m_1}$ und $Y_i \sim \mathrm{Bin}_{p_2,m_2}$. **Hinweis:** $S_n \neq S_0 + \sum_{i=1}^{n} X_i - Y_i$.

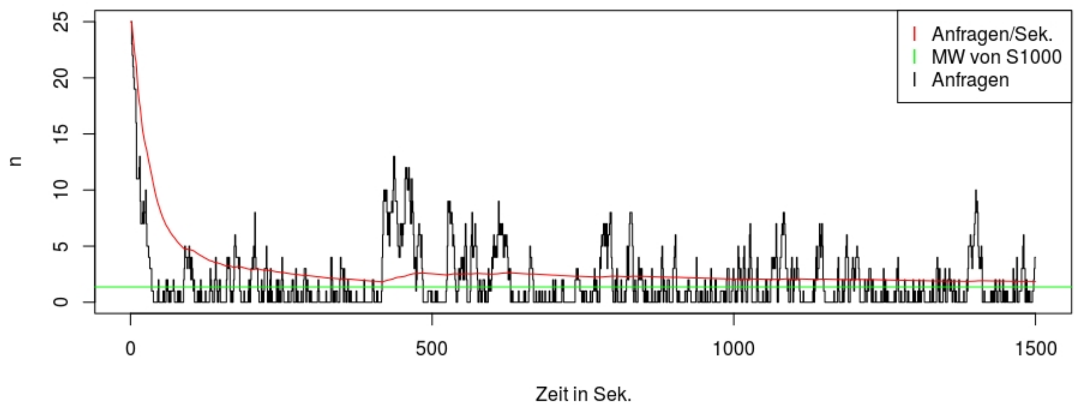
a): Schreiben Sie eine Funktion `mk(s0,m1,m2,p1,p2,n)`, welche den Server bis zur `n`-ten Sekunde simuliert mit $S_0 =$ `s0`, $X_i \sim \mathrm{Bin}_{\mathtt{p1},\mathtt{m1}}$ und $Y_i \sim \mathrm{Bin}_{\mathtt{p2},\mathtt{m2}}$. Geben Sie den Pfad $(S_1, S_2, ..., S_n)$ zurück.

b): Simulieren Sie den Pfad 1000 mal für `s0`=25,`m1=5`,`m2=6`,`p1=0.6`, `p2=0.6` und `n=1000`. Erstellen Sie ein Histogramm jeweils von S_5, S_{10}, S_{100} und S_{1000} auf einer Seite. Berechnen Sie für alle vier Zeitpunkte den Mittelwert und geben Sie diesen in der Überschrift an.

c): Schreiben Sie eine Funktion `mw(pfad)`, die für einen übergebenen Pfad `pfad` $= (x_1, x_2, ..., x_n)$ den Vektor `V` $= (v_1, v_2, ..., v_n)$ mit $v_k := \sum_{i=1}^{k} x_i/k$ berechnet.

d): Simulieren Sie einen Pfad `pfad` mit `n` = 1500 und den restlichen Parametern aus **b)**. Berechnen Sie `V`. Zeichnen Sie `pfad` und `V` in unterschiedlichen Farben in einem gemeinsamen Plot. Zeichnen Sie zusätzlich noch den in **b)** berechneten Mittelwert von S_{1000} als horizontale Linie ein (`abline`).

Hinweis: Ihr Bild in **d)** sollte etwa so aussehen:



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 23.06.2017