Prof. H.R. Künsch Wahrscheinlichkeitsrechnung und Statistik Abt. IX, SS99

Musterlösung zur Übung 8

1. a) Es gilt

$$\begin{aligned}\mathbf{E}[(Y-\beta)^2] &= \mathbf{E}[((Y-\mathbf{E}[Y])+(\mathbf{E}[Y]-\beta))^2]\\ &= \mathbf{E}[(Y-\mathbf{E}[Y])^2]+(\mathbf{E}[Y]-\beta)^2\,,\end{aligned}$$

denn $\mathbf{E}[(Y-\mathbf{E}[Y])(\mathbf{E}[Y]-\beta)] = (\mathbf{E}[Y]-\beta)\,\mathbf{E}[Y-\mathbf{E}[Y]] = 0$. Also ist $\mathbf{E}[(Y-\beta)^2]$ minimal, wenn der zweite Term oben verschwindet, also wenn $\beta = \mathbf{E}[Y]$. Das Minimum ist $\mathbf{E}[(Y-\mathbf{E}[Y])^2] = \mathrm{Var}(Y)$.

b) Gesucht sind nun α und β, die den Ausdruck

$$\begin{aligned}\mathbf{E}[(Y-(\alpha X+\beta))^2] &= \mathrm{Var}(Y-(\alpha X+\beta))+\mathbf{E}[Y-(\alpha X+\beta)]^2\\ &= \mathrm{Var}(Y-\alpha X)+\mathbf{E}[Y-(\alpha X+\beta)]^2\end{aligned}$$

minimieren. Da der erste Term unabhängig von β ist, braucht β nur den zweiten Summanden zu minimieren. Offensichtlich ist

$$\beta = \mathbf{E}[Y]-\alpha\mathbf{E}[X]$$

die richtige Wahl, da dann $\mathbf{E}[Y-(\alpha X+\beta)]^2\,(\geq 0)$ verschwindet. Für α setzen wir die Ableitung des ersten Terms gleich Null:

$$\begin{aligned}\mathrm{Var}(Y-\alpha X) &= \mathbf{E}[(Y-\alpha X)^2]-\mathbf{E}[Y-\alpha X]^2\\ &= \mathbf{E}[Y^2]-\mathbf{E}[Y]^2+\alpha^2\,(\mathbf{E}[X^2]-\mathbf{E}[X]^2)\\ &\qquad -2\alpha\,(\mathbf{E}[XY]-\mathbf{E}[X]\mathbf{E}[Y])\\ &= \mathrm{Var}(Y)+\alpha^2\,\mathrm{Var}(X)-2\alpha\,\mathrm{Cov}(X,Y)\\ \frac{d}{d\alpha}\mathrm{Var}(Y-\alpha X) &= 2\alpha\,\mathrm{Var}(X)-2\,\mathrm{Cov}(X,Y)=0\\ \alpha &= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\end{aligned}$$

Der minimale mittlere quadratische Fehler ist also

$$\begin{aligned}&\mathrm{Var}(Y)+\frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)}\mathrm{Var}(X)-2\,\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\mathrm{Cov}(X,Y)\\ &= \mathrm{Var}(Y)-\frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)}\\ &= \mathrm{Var}(Y)\Big(1-\frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)\mathrm{Var}(Y)}\Big) = \mathrm{Var}(Y)(1-\rho^2)\,,\end{aligned}$$

wobei $\rho^2 = \mathrm{Corr}(X,Y)$ die Korrelation zwischen X und Y ist.

c) Das Verhältnis des mittleren quadratischen Prognosefehlers in b) zu demjenigen aus a) ist $1-\rho^2$. Bei unkorrelierten X und Y ($\rho = 0$) bringt der Einbezug von X in die Prognose also nichts, bei vollständig korrelierten Zufallsvariablen ist sogar eine fehlerfreie Prognose möglich, wenn auf X abgestützt wird.

2. a) Die Behauptung ergibt sich unmittelbar aus b) oder aus folgender Überlegung:

$$\begin{aligned}\mathbf{E}[\mathbf{E}[Y|X]] &= \sum_{x\in X(\Omega)}\mathbf{E}[Y|X=x]\,P[X=x]\\ &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} y\,P[Y=y|X=x]\,P[X=x]\\ &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} y\,P[Y=y,X=x]\end{aligned}$$

Wegen $|y\,P[Y=y,X=x]| = |y|P[Y=y,X=x]$ und $\mathbf{E}|Y|<\infty$ konvergiert die Doppelsumme absolut, und man darf daher die Summationsreihenfolge vertauschen.:

$$\begin{aligned}\mathbf{E}[\mathbf{E}[Y|X]] &= \sum_{y\in Y(\Omega)} y\sum_{x\in X(\Omega)} P[Y=y,X=x]\\ &= \sum_{y\in Y(\Omega)} y\,P[Y=y]\\ &= \mathbf{E}[Y]\end{aligned}$$

b) Genauso:

$$\begin{aligned}\mathbf{E}[\mathbf{E}[Y|X]g(X)] &= \sum_{x\in X(\Omega)}\mathbf{E}[Y|X=x]g(x)\,P[X=x]\\ &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} yg(x)\,P[Y=y|X=x]\,P[X=x]\\ &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} yg(x)\,P[Y=y,X=x]\\ &= \mathbf{E}[Y\,g(X)]\,,\end{aligned}$$

denn die Doppelsumme hängt wegen $|y\,g(x)|\leq M\,|y|$ (mit einer oberen Schranke M für $|g(x)|$) und $\mathbf{E}|Y|<\infty$ nicht von der Summationsreihenfolge ab.

c) Die erste Behauptung folgt aus der Ungleichung von Jensen und aus der Teilaufgabe a): $\mathbf{E}[\mathbf{E}[Y|X]^2]\leq\mathbf{E}[\mathbf{E}[Y^2|X]] = \mathbf{E}[Y^2]<\infty$. Für den zweiten Teil rechne

$$\begin{aligned}\mathbf{E}[(Y-g(X))^2] &= \mathbf{E}[(Y-\mathbf{E}[Y|X]+\mathbf{E}[Y|X]-g(X))^2]\\ &= \mathbf{E}[(Y-\mathbf{E}[Y|X])^2]+\mathbf{E}[(\mathbf{E}[Y|X]-g(X))^2]\\ &\qquad +2\,\mathbf{E}[(Y-\mathbf{E}[Y|X])(\mathbf{E}[Y|X]-g(X))]\,.\end{aligned}$$

Der letzte Term verschwindet, denn mit $h(x) = \mathbf{E}[Y|X=x]-g(x)$ folgt wegen Teilaufgabe b): $\mathbf{E}[(Y-\mathbf{E}[Y|X])(\mathbf{E}[Y|X]-g(X))] = \mathbf{E}[(Y-\mathbf{E}[Y|X])\,h(X)] = \mathbf{E}[Y\,h(X)]-\mathbf{E}[\mathbf{E}[Y|X]\,h(X)] = 0$.

Nun ist $\mathbf{E}[(\mathbf{E}[Y|X] - g(X))^2] \geq 0$ und unter den geforderten Voraussetzungen endlich. Also folgt die Behauptung aus der zweiten Zeile der obigen Gleichungskette.

3. a) Dank der Ungleichung von Čebyšev

$$P\Big[\Big|\frac{S_n}{n} - \mu\Big| \leq \varepsilon\Big] \leq \frac{1}{\varepsilon^2}\mathrm{Var}\Big(\frac{S_n}{n}\Big)$$

müssen wir zeigen, dass $\mathrm{Var}(\frac{S_n}{n}) \to 0\ (n \to \infty)$.

$$\begin{aligned}\mathrm{Var}\Big(\frac{S_n}{n}\Big) &= \mathrm{Var}\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) = \frac{1}{n^2}\Big(\sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\sum_{i<j}\mathrm{Cov}(X_i, X_j)\Big) \\ &= \frac{1}{n^2}\Big(n\sigma^2 + 2\sum_{k=1}^{n-1}(n-k)\,R(k)\Big) = \frac{1}{n}\Big(\sigma^2 + 2\sum_{k=1}^{n-1}\Big(1 - \frac{k}{n}\Big)R(k)\Big)\end{aligned}$$

Zu zeigen bleibt, dass

$$\lim_{n\to\infty}\frac{2}{n}\sum_{k=1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k) = 0\,.$$

Sei $\eta > 0$. Dann existiert ein $k_0 \in \mathbb{N}$, so dass $|R(k)| < \eta$ für alle $k > k_0$. Es gilt:

$$\begin{aligned}&\text{i)} & \Big|\frac{2}{n}\sum_{k=1}^{k_0}\Big(\frac{n-k}{n}\Big)R(k)\Big| &\leq \frac{2k_0}{n}\sigma^2 < \eta\,,\ \text{falls } n \text{ gross genug} \\ &\text{ii)} & \Big|\frac{2}{n}\sum_{k=k_0+1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k)\Big| &\leq \frac{2}{n}\sum_{k=k_0+1}^{n-1}\frac{n-k}{n}\,\eta \leq \frac{2}{n}\,\eta\sum_{k=0}^{n-1}\frac{n-k}{n} \\ &&&\leq \eta\Big(2 - \frac{2}{n}\sum_{k=0}^{n-1}\frac{k}{n}\Big) = \eta\Big(2 - \frac{n-1}{n}\Big) < 2\eta \\ &\text{Also} & \Big|\frac{2}{n}\sum_{k=1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k)\Big| &< 3\eta \qquad \forall\ n > n_o = \frac{2k_0\sigma^2}{\eta}\,.\end{aligned}$$

Da η beliebig war, folgt die Behauptung.

b) Zur Einstimmung gilt

$$\begin{aligned}\lim_{n\to\infty} n\,\mathrm{Var}\Big(\frac{S_n}{n}\Big) &= \lim_{n\to\infty}\Big(\sigma^2 + 2\sum_{k=1}^{n-1}\Big(1 - \frac{k}{n}\Big)R(k)\Big) \\ &= \sigma^2 + 2\sum_{k=1}^{\infty}R(k) - 2\lim_{n\to\infty}\sum_{k=1}^{n-1}\frac{k}{n}R(k)\,.\end{aligned}$$

Nun setzen wir

$$a_n(k) := \begin{cases}\frac{k}{n}R(k) & (k < n) \\ 0 & (k \geq n)\end{cases}$$

und berücksichtigen, dass $a_n(k) \to 0\ (n \to \infty)$ für festes k und dass die absolut summierbaren $R(k)$ die $|a_n(k)|$ dominieren. Grenzwert und Summe dürfen also im folgenden mit dem Satz von Lebesgue vertauscht werden:

$$\lim_{n\to\infty}\sum_{k=1}^{n-1}\frac{k}{n}R(k) = \lim_{n\to\infty}\sum_{k=1}^{n-1}a_n(k) = \lim_{n\to\infty}\sum_{k=1}^{\infty}a_n(k) = \sum_{k=1}^{\infty}\lim_{n\to\infty}a_n(k) = 0$$

Damit folgt

$$\lim_{n\to\infty} n\,\mathrm{Var}\Big(\frac{S_n}{n}\Big) = \sigma^2 + 2\sum_{k=1}^{\infty}R(k) = \sum_{k=-\infty}^{\infty}R(k)\,.$$