Einführung in die Stochastik — Sommersemester 07
Dr. Walter Oevel — 2. 7. 2007

Übungsblatt 15

Hier ist zusätzliches Übungsmaterial zur Klausurvorbereitung quer durch die Inhalte der Vorlesung. Eine Korrektur ist nicht vorgesehen. Es handelt sich nicht um eine Probeklausur in dem Sinne, dass der Schwierigkeitsgrad/Bearbeitungsaufwand aller Aufgaben „typisch" für Klausuraufgaben sind.

Aufgabe 80: (Kombinatorik)

Mit welcher Wahrscheinlichkeit erhält man beim Skat (10 aus 32 Karten) genau 3 Asse und genau einen Buben?

Musterlösung:
Es gibt

$\binom{32}{10}$ Möglichkeiten, 10 aus 32 Karten zu bekommen,

$\binom{4}{3}$ Möglichkeiten, 3 der 4 Asse zu bekommen,

$\binom{4}{1}$ Möglichkeiten, einen der 4 Buben zu bekommen,

$\binom{24}{6}$ Möglichkeiten, die restlichen 6 Karten aus den verbleibenden 24 Karten (32 minus 4 Asse minus 4 Buben) zu bekommen.

Dies ergibt eine W'keit von $\dfrac{4 \cdot 4 \cdot \binom{24}{6}}{\binom{32}{10}} = \dfrac{5852}{175305} \approx 0.0334$.

Aufgabe 81: (Wahrscheinlichkeit von Ereignissen)

Die Koeffizienten $a, b \in \{1, 2, \ldots, 6\}$ der quadratischen Gleichung $x^2 + a\,x + b = 0$ werden (fair) gewürfelt. Mit welcher Wahrscheinlichkeit sind die Lösungen i) reell, ii) reell und rational?

Musterlösung:
Die beiden Lösungen der quadratischen Gleichung sind $\frac{1}{2}\,(a \pm \sqrt{a^2 - 4\,b})$.
i) Die Lösungen sind reell, wenn $a^2 \geq 4\,b$ gilt. Dies sind die folgenden Fälle:

$$\begin{array}{ll} b = 1, & a = 2,3,4,5,6 \\ b = 2, & a = 3,4,5,6 \\ b = 3, & a = 4,5,6 \\ b = 4, & a = 4,5,6 \\ b = 5, & a = 5,6 \\ b = 6, & a = 5,6 \end{array}$$

In 19 von 36 Fällen ergeben sich also reelle Lösungen.

ii) Die Lösungen sind genau dann rational, wenn $a^2 - 4\,b$ eine Quadratzahl ist. Spielt man die Möglichkeiten in i) durch, so findet man die 7 Fälle

$$(a, b) = (2, 1), (3, 2), (4, 3), (4, 4), (5, 4), (5, 6), (6, 5) \ .$$

Die gefragt Wahrscheinlichkeit ist also 7/36.

Aufgabe 82: (Unabhängigkeit von Ereignissen)

Es werden 2 faire Würfel geworfen. Sind die Ereignisse „die Augensumme ist durch 4 teilbar" und „die Differenz der Augen ist gerade" unabhängig?

Musterlösung:
Der sich anbietende Stichprobenraum ist $\Omega = \{(1,1), (1,2), \ldots, (6,6)\}$. Die durch 4 teilbaren Augensummen 4, 8, 12 entsprechen dem Ereignis

$$A = \{(1,3), (2,2), (3,1),\ (2,6), (3,5), (4,4), (5,3), (6,2),\ (6,6)\} \ .$$

Die durch 2 teilbaren Differenzen entsprechen dem halben Stichprobenraum:

$$B = \Big\{(1,1), (1,3), (1,5),\ (2,2), (2,4), (2,6),\ (3,1), (3,3), (3,5), \\ \ldots,\ (6,2), (6,4), (6,6)\Big\}\ , \quad |B| = \tfrac{1}{2} \cdot |\Omega| = 18.$$

Es gilt $A \subset B$, also $A \cap B = A$. Dies ist auch ohne explizites Aufschreiben klar: gilt für eine Summe ganzer Zahlen $x + y = 4\,k$ mit $k \in \mathbb{Z}$, so ist die Differenz $x - y = x - (4\,k - x) = 2\,x - 4\,k = 2\,(x - 2\,k)$ eine gerade Zahl. Mit

$$P(A \cap B) = P(A) \ \neq\ P(A) \cdot P(B) = P(A) \cdot \frac{1}{2}$$

sind die Ereignisse abhängig.

Aufgabe 83: (Rechenregeln für Erwartungswerte/Streuungen)

Es wird 2 Mal fair gewürfelt. Sie X die Differenz zwischen dem ersten und dem zweiten Wurf. Bestimme (auf möglichst einfache Weise) Erwartungswert und Streuung von X!

Musterlösung:
Sei Y das Ergebnis eines Wurfes. Es gilt

$$\mathrm{E}(Y) = \frac{1 + 2 + \cdots + 6}{6} = \frac{7}{2}\ ,\ \mathrm{E}(Y^2) = \frac{1 + 4 + \cdots + 36}{6} = \frac{91}{6}$$

und damit $\sigma^2(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 = 35/12$. Sei $X = X_1 - X_2$, wo X_i das Ergebnis des i-ten Wurfs ist. $X_{1,2}$ sind unabhängig und haben beide den Erwartungswert $\mathrm{E}(X_i) = 7/2$ und die Varianz $\sigma^2(X_i) = 35/12$. Es folgt

$$\mathrm{E}(X) = \mathrm{E}(X_1 - X_2) = \mathrm{E}(X_1) - \mathrm{E}(X_2) = 0$$

und wegen der Unabhängigkeit

$$\sigma^2(X) = \sigma^2(X_1 - X_2) = \sigma^2(X_1) + \sigma^2(-X_2) = \sigma^2(X_1) + \sigma^2(X_2) = \frac{35}{6}\ .$$

Aufgabe 84: (Linearität von Erwartungswerten)

Gegeben ist eine willkürlich ausgewählte Gruppe von 500 Personen. Sei

a) X die Anzahl aller Tage des Jahres, an denen mindestens eine dieser Personen Geburtstag hat,

b) Y die Anzahl aller Tage des Jahres, an denen mindestens zwei dieser Personen Geburtstag haben.

Bestimme den Erwartungswert von X und Y!

Musterlösung:
Fixiere $k \in \{1, 2, \ldots, 365\}$ (ein vorgegebener Tag des Jahres).

a) Sei

$$X_k = \begin{cases} 1, & \text{falls mindestens eine der Personen an diesem Tag Geburtstag hat,} \\ 0, & \text{falls keine der Personen an diesem Tag Geburtstag hat.} \end{cases}$$

Offensichtlich gilt

$$P(X_k = 0) = \left(\frac{364}{365}\right)^{500} \approx 0.2537, \quad P(X_k = 1) = 1 - P(X_k = 0) \approx 0.7463,$$

damit ist der Erwartungswert

$$\mathrm{E}(X_k) \;=\; 0 \cdot P(X_k = 0) + 1 \cdot P(X_k = 1) \;=\; P(X_k = 1)\ .$$

Mit $X = \sum_{k=1}^{365} X_k$ folgt

$$\mathrm{E}(X) = \sum_{k=1}^{365} \mathrm{E}(X_k) = 365 \cdot \left(1 - \left(\frac{364}{365}\right)^{500}\right) \approx 272.4\ .$$

b) Sei nun

$$Y_k = \begin{cases} 1, & \text{falls mindestens zwei der Personen an diesem Tag Geburtstag haben,} \\ 0, & \text{falls höchstens eine der Personen an diesem Tag Geburtstag hat.} \end{cases}$$

Das Ereignis „$Y_k = 0$“ entspricht „keine oder genau eine Person hat am Tag k Geburtstag“:

$$P(\text{„keine Person hat am Tag } k \text{ Geburtstag“}) = \left(\frac{364}{365}\right)^{500},$$

$$P(\text{„genau eine Person hat am Tag } k \text{ Geburtstag“}) = \binom{500}{1} \frac{1}{365} \left(\frac{364}{365}\right)^{499}$$

$$\Rightarrow \quad P(Y_k = 0) \;=\; \left(\frac{364}{365}\right)^{500} \;+\; \frac{500}{365}\left(\frac{364}{365}\right)^{499}\ .$$

Man beachte hierzu, dass das Ereignis $Y_k = j$ als j Erfolge bei 500-facher Wiederholungedes Bernoulli-Experiments „wähle eine Person, ist ihr Geburtstag k?“ interpretiert werden kann.
Wie in a) gilt

$$E(Y_k) \;=\; P(Y_k = 1) \;=\; 1 - P(Y_k = 0) \;=\; 1 \;-\; \left(\frac{364}{365}\right)^{500} \;-\; \frac{500}{365}\left(\frac{364}{365}\right)^{499}\ .$$

Mit $Y = \sum_{k=1}^{365} Y_k$ folgt

$$\mathrm{E}(Y) = \sum_{k=1}^{365} \mathrm{E}(Y_k) = 365 \cdot \Bigg(1 \;-\; \Big(\frac{364}{365}\Big)^{500} \;-\; \frac{500}{365}\Big(\frac{364}{365}\Big)^{499}\Bigg) \approx 145.2 \; .$$

Aufgabe 85: (Bedingte Erwartungswerte)

Sei X eine Zufallsvariable über einem diskreten Stichprobenraum Ω, sei $A \subset \Omega$ ein beliebiges Ereignis. Zeige formal mit der Definition 2.58 der Vorlesung:

$$\mathrm{E}(f(X) \,|\, A) = \frac{1}{P(A)} \sum_{\omega \in A} f(X(\omega))\; P(\omega) \; .$$

Musterlösung:
Für diskretes X mit $\Omega(X) = \{r_1, r_2, \dots\}$ gilt nach Definition 2.58 der Vorlesung:

$$\mathrm{E}(f(X) \,|\, A) = \sum_{r \in X(\Omega)} f(r)\; P(X = r \,|\, A) \; .$$

Für einen diskreten Stichprobenraum $\Omega = \{\omega_1, \omega_2, \dots\}$ gilt

$$P(X = r \,|\, A) = \frac{P(X^{-1}(r) \cap A)}{P(A)} = \frac{1}{P(A)} \sum_{\omega \in X^{-1}(r) \cap A} P(\{\omega\}) = \frac{1}{P(A)} \sum_{\substack{\omega \in A \\ X(\omega) = r}} P(\{\omega\}) \; .$$

Es folgt

$$\begin{aligned} \mathrm{E}(f(X) \,|\, A) &= \sum_{r \in X(\Omega)} f(r)\; \frac{1}{P(A)} \sum_{\substack{\omega \in A \\ X(\omega) = r}} P(\{\omega\}) = \frac{1}{P(A)} \sum_{r \in X(\Omega)} \sum_{\substack{\omega \in A \\ X(\omega) = r}} f(r)\, P(\{\omega\}) \\ &= \frac{1}{P(A)} \sum_{r \in X(\Omega)} \sum_{\substack{\omega \in A \\ X(\omega) = r}} f(X(\omega))\, P(\{\omega\}) = \frac{1}{P(A)} \sum_{\omega \in A} f(X(\omega))\, P(\{\omega\}) \; . \end{aligned}$$

Aufgabe 86: (Moivre–Laplace)

Es ist bekannt, dass Knabengeburten häufiger sind als Mädchengeburten. In der Tat hat man durch langjährige statistische Untersuchungen festgestellt, dass durchschnittlich bei 1 000 Geburten 514 Jungen geboren werden (also etwa 5.76% mehr Knaben als Mädchen). Mit welcher Sicherheit ist die Aussage „es werden mindestens 5% mehr Knaben als Mädchen geboren“ richtig, wenn beim Erstellen dieser Statistik 1 000 000 Geburten betrachtet wurden?

Musterlösung:
Dies ist dieselbe Überlegung wie in Aufgabe 55, lediglich etwas anders formuliert:
Betrachte „Knabe“ als Erfolg des Bernoulli–Experiments „Geburt“, das $n = 10^6$-fach wiederholt wurde. Sei $p = P(\text{„Knabe“})$. Die Aussage „mindestens 5% mehr Knaben als Mädchen“ entspricht $p \geq 1.05\,(1-p)$, also $p \geq 0.5122$. Gesucht ist die Sicherheit

$$P(0.5122 \leq p) \;=\; P(0.5122 \leq p < \infty) \; .$$

Für die mittlere Häufigkeit $\bar{X}_n$ der Knabengeburten gilt nach Moivre–Laplace:

$$P(A \le \bar{X}_n - p \le B) \;\approx\; \frac{1}{\sqrt{2\,\pi}} \int_{A\,\sqrt{\frac{n}{p\,q}}}^{B\,\sqrt{\frac{n}{p\,q}}} e^{-x^2/2}\,dx\;.$$

Wir rechnen zurück, welche Werte für A und B wir wählen müssen, damit wir diese W'keit als Sicherheit für die Aussage $0.5122 \le q$ interpretieren können. Setzen wir dazu gemäß Interpretation 2, 4.3 des Skripts den Messwert $\bar{X}_n = 514/1000$ ein, gilt

$$P(A \le \bar{X}_n - p \le B) = P\Big(A \le \frac{514}{1\,000} - p \le B\Big) = P\Big(\underbrace{\frac{514}{1000} - B}_{=0.5122} \le p \le \underbrace{\frac{514}{1\,000} - A}_{=\infty}\Big)\;.$$

Wählen wir also $B = 514/1000 - 0.5122 = 0.0018$, $A = -\infty$, so folgt

$$P(0.5122 \le p \le \infty) = P(A \le \bar{X}_n - p \le B)$$

$$\approx \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{B\,\sqrt{\frac{n}{p\,q}}} e^{-x^2/2}\,dx = \frac{1}{2}\left(1 + \Phi\Big(B\,\sqrt{\frac{n}{p\,q}}\,\Big)\right)\;.$$

Mit $p\,q \approx 1/4$ ergibt sich die Sicherheit

$$P(0.5122 \le p) \approx \frac{1}{2}\Big(1 + \Phi\Big(0.0018\,\sqrt{4 \cdot 10^6}\,\Big)\Big) = \frac{1}{2}\Big(1 + \Phi\Big(3.6\Big)\Big)$$

$$\approx \frac{1 + 0.99968}{2} = 0.99984\;.$$

Aufgabe 87: (Moivre–Laplace)

In einer Fabrik will man den Ausschussanteil der Produktion statistisch mit der Genauigkeit von $\pm 1\%$ und der Sicherheit von 95.4% schätzen. Wieviele Stichproben müssen dafür vorgenommen werden?

Musterlösung:

Die Wahrscheinlichkeit für das Bernoulli–Experiment

$$p = P(\text{„ein untersuchtes Teil ist Ausschuss“})$$

soll bis auf die absolute Genauigkeit $\epsilon = 0.01$ festgelegt werden. Für die relative Häufigkeit $\bar{X}_n$ der als Ausschuss festgestellten untersuchten Objekte bei n-facher Wiederholung des Bernoulli–Experiments gilt nach Moivre–Laplace (Bemerkung 4.11 der Vorlesung):

$$P(-\epsilon \le \bar{X}_n - p \le \epsilon) \;\approx\; \frac{1}{\sqrt{2\,\pi}} \int_{-\epsilon\,\sqrt{\frac{n}{p\,q}}}^{\epsilon\,\sqrt{\frac{n}{p\,q}}} e^{-x^2/2}\,dx \;=\; \Phi\Big(\epsilon\,\sqrt{\frac{n}{p\,q}}\,\Big)\;.$$

Mit $S = P(-\epsilon \le \bar{X}_n - p \le \epsilon) = 0.954$, $\Phi^{-1}(0.954) \approx 2.0$ muss für n gelten:

$$\epsilon\,\sqrt{\frac{n}{p\,q}} \approx \Phi^{-1}(S) \quad \Rightarrow \quad n \approx \frac{p\,q}{\epsilon^2}\,\Phi^{-1}(S) \approx \frac{p\,q \cdot 4.0}{0.01^2}\;.$$

Da keinerlei Information über die Größenordnung von p und $q = 1 - p$ vorliegen, verwenden wir die allgemein gültige Abschätzung $p\,q \le 1/4$. Damit sind für

$$n \ge \frac{1}{0.01^2} = 10\,000$$

Stichproben die Anforderungen auf jeden Fall erfüllt.

Aufgabe 88: (Moivre–Laplace, zentraler Grenzwertsatz)

In Aufgabe 43.c) war graphisch beobachtet worden, dass die Poisson–Verteilung Pois(λ) einer Variable X wie eine Normalverteilung aussieht, wenn λ sehr gross ist. Begründe dieses Phänomen über Moivre–Laplace bzw. den zentralen Grenzwertsatz! Genauer: zeige, dass für beliebiges α, β gilt:

$$P\Big(\lambda + \alpha \cdot \sqrt{\lambda} \le X \le \lambda + \beta \cdot \sqrt{\lambda}\,\Big) = \frac{1}{\sqrt{2\,\pi}} \int_\alpha^\beta e^{-x^2/2}\,dx + O\Big(\frac{1}{\sqrt{\lambda}}\Big).$$

Musterlösung:

Wir wissen, dass die Poisson–Verteilung Pois(λ) die Grenzverteilung der Binomial–Verteilung Bi(n, p) ist, wenn n groß und gleichzeitig p klein ist: im Grenzübergang $n \to \infty$ wird $\lambda = n \cdot p$ konstant gehalten, also $p = p(n) = \lambda/n$ (siehe Abschnitt 2.3.2 des Skripts).

Halte nun $p \ll 1$ fest, setze $n = n(\lambda) = \lambda/p$ und lasse λ gegen ∞ laufen. Damit läuft auch n gegen ∞. Nach Moivre–Laplace konvergiert $Bi(n, p) = Bi(\lambda/p, p)$ im folgenden Sinne gegen eine Normalverteilung: Für $Bi(\lambda/p, p)$ verteiltes X gilt nach den Folgerungen 4.10 des Skripts

$$P(a' \le X \le b') \;=\; \frac{1}{\sqrt{2\,\pi}} \int_{\frac{a'-n\,p}{\sqrt{n\,p\,q}}}^{\frac{b'-n\,p}{\sqrt{n\,p\,q}}} e^{-x^2/2}\,dx \;+\; O\Big(\frac{1}{\sqrt{n}}\Big)$$

$$= \;\frac{1}{\sqrt{2\,\pi}} \int_{\frac{a'-\lambda}{\sqrt{\lambda\,(1-p)}}}^{\frac{b'-\lambda}{\sqrt{\lambda\,(1-p)}}} e^{-x^2/2}\,dx \;+\; O\Big(\frac{1}{\sqrt{\lambda/p}}\Big)\;.$$

Setzen wir nun noch $\alpha = \frac{a'-\lambda}{\sqrt{\lambda\,(1-p)}}$, $\beta = \frac{b'-\lambda}{\sqrt{\lambda\,(1-p)}}$, so folgt mit

$$P(a' \le X \le b') = P(\underbrace{a' - \lambda}_{=\alpha\,\sqrt{\lambda\cdot(1-p)}} \le X - \lambda \le \underbrace{b' - \lambda}_{=\beta\,\sqrt{\lambda\cdot(1-p)}})$$

das Ergebnis

$$P\Big(\alpha \cdot \sqrt{\lambda\,(1-p)} \le X - \lambda \le \beta \cdot \sqrt{\lambda\,(1-p)}\,\Big)$$
$$= \;\frac{1}{\sqrt{2\,\pi}} \int_\alpha^\beta e^{-x^2/2}\,dx \;+\; O\Big(\frac{1}{\sqrt{\lambda/p}}\Big)\;.$$

Für $\lambda \gg p$ können wir $O(1/\sqrt{\lambda/p})$ vernachlässigen. Mit $p \ll 1$ gilt $1 - p \approx 1$ und es folgt

$$P\Big(\alpha \cdot \sqrt{\lambda\,(1-p)} \le X - \lambda \le \beta \cdot \sqrt{\lambda\,(1-p)}\Big) \;\approx$$

$$\boxed{P\Big(\lambda + \alpha \cdot \sqrt{\lambda} \le X \le \lambda + \beta \cdot \sqrt{\lambda}\,\Big) \;\approx\; \frac{1}{\sqrt{2\,\pi}} \int_\alpha^\beta e^{-x^2/2}\,dx.}$$

Alternative Begründung: Die Idee ist: nach Aufgabe 72 ist die Summe unabhängiger Pois(λ_i)–verteilter Variabler wieder Poisson–verteilt zum Parameter $\lambda = \sum_i \lambda_i$. Damit kann jede Poisson–verteilte Variable aufgefasst werden als Summe von Poisson–Variablen. Hiermit kommt der zentrale Grenzwertsatz ins Spiel:

Gegeben ein Pois(λ_0)–verteiltes Einzelexperiment X_0, das n-fach unabhängig wiederholt wird. Die Wiederholungen seien $X_1, \ldots, X_n$. Der gemeinsame Erwartungswert ist $\mu = \lambda_0$, die Streuung ist $\sigma = \sqrt{\lambda_0}$. Fixiere λ_0 im Folgenden.
Betrachte die Summenvariable $X = X_1 + \cdots + X_n$ zum Poisson–Parameter $\lambda = n\,\lambda_0$. Nach den Folgerungen 4.15 des Skripts gilt für die Summenvariable

$$P(a' \leq X \leq b') \;=\; \frac{1}{\sqrt{2\,\pi}} \int_{\frac{a'-n\,\mu}{\sqrt{n}\,\sigma}}^{\frac{b'-n\,\mu}{\sqrt{n}\,\sigma}} e^{-x^2/2}\;dx \;+\; O\Big(\frac{1}{\sqrt{n}}\Big)$$

$$= \frac{1}{\sqrt{2\,\pi}} \int_{\frac{a'-\lambda}{\sqrt{\lambda}}}^{\frac{b'-\lambda}{\sqrt{\lambda}}} e^{-x^2/2}\;dx \;+\; O\Big(\frac{1}{\sqrt{n}}\Big).$$

Es geht nun analog zur ersten Begründung über Moivre–Laplace weiter. Setzen wir $\alpha = \frac{a'-\lambda}{\sqrt{\lambda}}$, $\beta = \frac{b'-\lambda}{\sqrt{\lambda}}$, so folgt mit

$$P(a' \leq X \leq b') = P(\underbrace{a'-\lambda}_{=\alpha\,\sqrt{\lambda}} \leq X - \lambda \leq \underbrace{b'-\lambda}_{=\beta\,\sqrt{\lambda}}\,)$$

das Ergebnis

$$\boxed{P\Big(\lambda + \alpha \cdot \sqrt{\lambda} \leq X \leq \lambda + \beta \cdot \sqrt{\lambda}\,\Big) \;=\; \frac{1}{\sqrt{2\,\pi}} \int_{\alpha}^{\beta} e^{-x^2/2}\;dx \;+\; O\Big(\frac{1}{\sqrt{\lambda/\lambda_0}}\Big)\;.}$$

Nun betrachte $\lambda \gg \lambda_0$ für ein beliebig vorgegebenes λ_0.
