



Lösungsskizzen zu Übungsblatt 3

Aufgabe 1

Es sei p_1 die Zähldichte der Binomial-Verteilung $\mathcal{B}(n, p)$ für $n \in \mathbb{N}$ und $0 < p < 1$.

(i) Zeigen Sie die Rekursionsformel

$$p_1(k+1) = \frac{n-k}{k+1} \cdot \frac{p}{1-p} \cdot p_1(k)$$

für $k = 0, 1, \dots, n-1$.

Sei p_2 die Zähldichte der hypergeometrischen Verteilung $\mathcal{H}(N, M, n)$ für $N, M, n \in \mathbb{N}_0$. Wir wollen zeigen, dass sich diese für $N, M \rightarrow \infty$ und $\frac{M}{N} \rightarrow p$ durch die Zähldichte der Binomialverteilung mit Parametern (n, p) approximieren lässt.

(iii) Zeigen Sie, dass für $k = 0, \dots, n$

$$p_2(k) \rightarrow \binom{n}{k} p^k (1-p)^{n-k}$$

für $N, M \rightarrow \infty$.

(iii) Sie haben eine Münze, welche bei einem Wurf mit Wahrscheinlichkeit $p \in (0, 1)$ 'Kopf' anzeigt. Sie werfen diese Münze nun n -mal hintereinander und zählen die Anzahl k_1 der Köpfe. Danach wirft jemand anderes die Münze noch weitere m -mal und zählt ebenfalls die Anzahl k_2 der Köpfe. Welcher Verteilung folgt die Summe der Anzahlen $k_1 + k_2$ der Köpfe?

Sie dürfen für den Beweis folgende Rechenregel verwenden:

$$\sum_{k_1=0}^k \binom{n}{k_1} \cdot \binom{m}{k-k_1} = \binom{m+n}{k}$$

für $k \leq m+n$, $m, n \in \mathbb{N}_0$.

(1+2+2 Punkte)

Beweis. (i) Sei $k \in \{0, 1, \dots, n-1\}$. Es gilt:

$$\begin{aligned} p_1(k+1) &= \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1} \\ &= \frac{n!}{(n-k-1)!(k+1)!} \cdot \frac{p}{1-p} p^k (1-p)^{n-k} \\ &= \frac{p}{1-p} \cdot \frac{n-k}{k+1} \cdot \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\ &= \frac{p}{1-p} \cdot \frac{n-k}{k+1} p_1(k). \end{aligned}$$

(ii) Die Zähldichte der hypergeometrischen Verteilung $H(N, M, n)$ lautet:

$$\begin{aligned}
 p_2(k) &= \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\frac{M!}{k!(M-k)!} \cdot \frac{(N-M)!}{(n-k)!(N-M-n+k)!}}{\frac{N!}{n!(N-n)!}} \\
 &= \frac{n!}{k! \cdot (n-k)!} \cdot \frac{\frac{M!}{(M-k)!} \cdot \frac{(N-M)!}{(N-M-n+k)!}}{\frac{N!}{(N-n)!}} \\
 &= \binom{n}{k} \cdot \frac{\prod_{i=0}^{k-1} (M-i) \cdot \prod_{i=0}^{n-k-1} (N-M-i)}{\prod_{i=0}^{n-1} (N-i)} \cdot \frac{\frac{1}{N^n}}{\frac{1}{N^n}} \\
 &= \binom{n}{k} \cdot \frac{\prod_{i=0}^{k-1} \left(\frac{M}{N} - \frac{i}{N}\right) \cdot \prod_{i=0}^{n-k-1} \left(1 - \frac{M}{N} - \frac{i}{N}\right)}{\prod_{i=0}^{n-1} \left(1 - \frac{i}{N}\right)}
 \end{aligned}$$

Alle auftretenden Produkte sind endlich, also können wir diese ohne weiteres Wissen berechnen und insbesondere den Limes $M, N \rightarrow \infty$ in die Produkte hereinziehen. Wir nutzen noch $M/N \rightarrow p$ und $i/N \rightarrow 0$ (da i fest).

$$p_2(k) \xrightarrow{M, N \rightarrow \infty} \binom{n}{k} \cdot \frac{\prod_{i=0}^{k-1} (p-0) \cdot \prod_{i=0}^{n-k-1} (1-p-0)}{\prod_{i=0}^{n-1} (1-0)} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

(iii) Wir definieren für $k_1 \in \{0, \dots, n\}, k_2 \in \{0, \dots, m\}, k \in \{0, \dots, m+n\}$ die folgenden Ereignisse.

- $A_{k_1} :=$ Im ersten Durchgang k_1 -mal Kopf,
- $B_{k_2} :=$ Im zweiten Durchgang k_2 -mal Kopf,
- $C_k :=$ Bei beiden Runden zusammen wurde k -mal Kopf geworfen.

Laut Voraussetzung sind A_k, B_l unabhängig für alle k, l . Außerdem gilt

$$\mathbb{P}(A_k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{und} \quad \mathbb{P}(B_k) = \binom{m}{k} p^k (1-p)^{m-k}.$$

Nun gilt nach Definition der Ereignisse A_{k_1}, C_k , dass $A_{k_1} \cap C_k = \emptyset$ für $k_1 > k$. Wir berechnen

$$\begin{aligned}
 \mathbb{P}(C_k) &= \sum_{k_1=0}^{\infty} \mathbb{P}(C_k \cap A_{k_1}) = \sum_{k_1=0}^k \mathbb{P}(C_k \cap A_{k_1}) = \sum_{k_1=0}^k \mathbb{P}(B_{k-k_1} \cap A_{k_1}) \\
 &= \sum_{k_1=0}^k \mathbb{P}(B_{k-k_1}) \cdot \mathbb{P}(A_{k_1}) \\
 &= \sum_{k_1=0}^k \left[\binom{n}{k_1} p^{k_1} (1-p)^{n-k_1} \right] \cdot \left[\binom{m}{k-k_1} p^{k-k_1} (1-p)^{m-(k-k_1)} \right] \\
 &= p^k \cdot (1-p)^{(m+n)-k} \sum_{k_1=0}^k \binom{n}{k_1} \cdot \binom{m}{k-k_1} \\
 &= \binom{m+n}{k} \cdot p^k \cdot (1-p)^{(m+n)-k}.
 \end{aligned}$$

Das ist die Zähldichte der Binomialverteilung mit Parametern $(m+n, p)$.

□

Aufgabe 2

Entscheiden Sie jeweils, welche diskrete Wahrscheinlichkeitsverteilung zur Modellierung der folgenden Probleme geeignet ist und geben Sie zur Berechnung der Wahrscheinlichkeiten deren Parameter an.

- (i) Sie werfen einen fairen Würfel, bis das erste Mal eine '6' auftritt. Mit welcher Wahrscheinlichkeit muss man mehr als 9-mal würfeln?
- (ii) Eine Firma produziert insgesamt 350 elektronische Bauteile des gleichen Typs für einen Kunden. Die Firma weiß, dass genau 5 Teile defekt sind, aber nicht welche. Um die Qualität zu prüfen untersucht der Kunde die Bauteile mittels einer Stichprobe. Dazu werden zufällig 20 Bauteile herausgenommen und untersucht. Wenn mehr als ein defektes Bauteil gefunden wird, wird die Sendung zurückgeschickt. Wie groß ist die Wahrscheinlichkeit, dass die Warensendung wieder zurück geschickt wird?
- (iii) Ein Insekt legt 100 Eier, die sich unabhängig voneinander entwickeln. Aus jedem Ei schlüpft mit Wahrscheinlichkeit 0.01 ein Nachkomme. Wie groß ist die Wahrscheinlichkeit, dass es mindestens 2 Nachkommen gibt?

(1+1+1 Punkte)

Beweis. (i) Wir modellieren die Anzahl der Würfe durch die geometrische Verteilung mit $p = \frac{1}{6}$ für Erfolg (d.h. eine '6'), d.h.

$$p(k) = p \cdot (1 - p)^{k-1}$$

mit k als Anzahl der Würfe. Sei A das Ereignis, dass man mehr als 9-mal würfeln muss, dass eine '6' gewürfelt wurde. Damit

$$\mathbb{P}(A) = \sum_{k=10}^{\infty} p(k) = p \sum_{k=9}^{\infty} (1 - p)^k = \frac{1}{6} \cdot \sum_{k=9}^{\infty} \left(\frac{5}{6}\right)^k \approx 0,194,$$

wobei wir benutzt haben, dass $\sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^k = 6$.

- (ii) Anzahl der defekten Bauteile, die vom Kunden gezogen werden, modellieren wir durch eine hypergeometrische Verteilung mit Parametern $(N, M, n) = (350, 5, 20)$, d.h.

$$p(k) = \frac{\binom{N-M}{n-k} \cdot \binom{M}{k}}{\binom{N}{n}}, \quad k = 0, \dots, n$$

Gesucht ist die Wahrscheinlichkeit für das Ereignis $A :=$ 'Es wird mehr als ein defektes Bauteil gefunden'.

$$\mathbb{P}(A) = 1 - p(0) - p(1) \approx 1 - 0.744 - 0.228 = 0.028.$$

Special für die Übungsgruppe:

Approximation durch Binomialverteilung mit Parametern (n, p) wie in Aufgabe 1, $p := \frac{M}{N} = \frac{5}{350} = \frac{1}{70}$ und

$$\tilde{p}(k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}, \quad k = 0, \dots, n$$

Gesucht ist

$$\mathbb{P}(A) \approx 1 - \tilde{p}(0) - \tilde{p}(1) \approx 1 - 0.750 - 0.217 = 0.033.$$

- (iii) Die Eier entwickeln sich unabhängig voneinander und eine feste Anzahl von $n = 100$ Versuchen mit Erfolgswahrscheinlichkeit $p = 0.01$ wird durchgeführt. Also folgt die Anzahl der Nachkommen einer Binomialverteilung mit Parametern (n, p) und hat also eine Zähldichte $p(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$.

Die exakte Wahrscheinlichkeit für mindestens zwei Nachkommen (wir nennen dieses Ereignis B) ist demzufolge

$$\begin{aligned} \mathbb{P}(B) &= 1 - p(0) - p(1) = 1 - (1-p)^n - n \cdot p \cdot (1-p)^{n-1} \\ &\approx 1 - 0.366 - 0.370 = 0.264 \end{aligned}$$

Special für die Übungsgruppe:

Nach der Vorlesung (Proposition 3.8) kann die Zähldichte der Binomialverteilung mit Parametern (n, p) für großes n und kleines p durch die Zähldichte einer Poisson-Verteilung mit Parameter $\lambda = n \cdot p$ approximiert werden.

Hier approximieren wir also $p(k)$ durch $\hat{p}(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, die Zähldichte einer Poisson-Verteilung mit Parameter $\lambda = n \cdot p = 1$. In diesem Fall erhalten wir

$$\mathbb{P}(B) \approx 1 - \hat{p}(0) - \hat{p}(1) = 1 - \frac{e^{-\lambda} \lambda^0}{0!} - \frac{e^{-\lambda} \lambda^1}{1!} = 1 - \frac{2}{e} \approx 0.264.$$

□

Aufgabe 3

Wir haben einen Würfel, bei welchem die Augenzahl '3' bei einem Würfelwurf mit Wahrscheinlichkeit $p \in (0, 1)$ auftritt. Wir haben die Vermutung, dass die '3' häufiger kommt als es bei einem fairen Würfel (alle Zahlen gleichwahrscheinlich) der Fall wäre. Ganz konkret vermuten wir, dass die '3' tatsächlich mit einer Wahrscheinlichkeit von 25% auftritt. Um dies zu überprüfen, werfen wir den Würfel solange, bis das erste Mal '3' erscheint. Die Anzahl der benötigten Würfe (einschließlich des Wurfs mit der '3') bezeichnen wir mit k . Auf Basis dieses Versuchs wollen wir nun einen Test durchführen, ob unsere Vermutung richtig ist.

- Formulieren Sie das Testproblem mathematisch. Wählen Sie Null- und Alternativhypothese so, dass die folgende Forderung den Fehler 1. Art beschränkt: "Der peinliche Irrtum zu behaupten, dass ein unfairer Würfel (bzgl. der '3') vorliegt, obwohl der Würfel eigentlich fair ist, soll maximal mit einer Wahrscheinlichkeit von 20% auftreten".
- Bestimmen Sie den Neyman-Pearson-Test für das Testproblem aus (a) und berechnen Sie die Wahrscheinlichkeit für den Fehler 2. Art.
- Sei nun wieder $\alpha \in [0, 1]$ beliebig. Unter welchen Bedingungen ist der Test aus (b) ein gleichmäßig bester Test für die Hypothesen $H_0 : p = \frac{1}{6}$ gegen $H'_A : p \in \{p' : p' > \frac{1}{6}\}$?

(2+3+1 Punkte)

Beweis. (a) Die Anzahl der benötigten Würfe bis zur ersten '3' (einschließlich des Wurfs mit der 3) ist geometrisch verteilt.

Wir haben also einen Stichprobenraum $\Omega = \mathbb{N}$ versehen mit der Potenzmenge $\mathcal{P}(\Omega)$ vorliegen, und es gibt nun zwei Hypothesen für die Verteilung der Stichprobe.

$$\begin{aligned} H_0 : \quad \mathbb{P} &= \mathcal{G}(p), \quad p = \frac{1}{6}, \quad p(k) = p \cdot (1-p)^{k-1} \\ H_A : \quad \mathbb{Q} &= \mathcal{G}(q), \quad q = \frac{1}{4}, \quad q(k) = q \cdot (1-q)^{k-1}. \end{aligned}$$

Zu finden ist nun ein bester Test $\phi^*(k)$ zu einem vorgegebenen Niveau $\alpha = 0.2$, der anhand der Anzahl benötigten Würfe k bis zur ersten '3' entweder die Nullhypothese H_0 annimmt oder verwirft.

Wir wollen eine Schranke für den Fehler 1. Art vorgeben:

$$\mathbb{P}(\phi = \mathbb{Q}) = \text{''Wahrscheinlichkeit, sich für unfairen Würfel zu entscheiden, obwohl eigentlich fairer Würfel vorliegt.''} \leq \alpha := 0.2.$$

Damit die anschauliche Interpretation des Fehler 1. Art mit der linken Seite $\mathbb{P}(\phi = \mathbb{Q})$ übereinstimmt, muss \mathbb{P} (bzw. H_0) als die Wahrscheinlichkeitsverteilung des fairen Würfels und \mathbb{Q} (bzw. H_A) als die Wahrscheinlichkeitsverteilung des unfairen Würfels definiert werden. Dies wurde oben genauso getan.

(b) Ein (approximativer) Neyman-Pearson-Test für das in (a) formulierte Testproblem ist gegeben durch:

$$\phi(k) = \begin{cases} \mathbb{P}, & L(k) := \frac{q(k)}{p(k)} < c^* \\ \mathbb{Q}, & L(k) \geq c^* \end{cases},$$

für $k = 0, \dots, n$ und einem c^* mit $\mathbb{P}(\phi = \mathbb{Q}) = \mathbb{P}(L(k) \geq c^*) \leq \alpha$, so dass die linke Seite der Ungleichung maximal ist.

Wir wollen den Test mit Hilfe der Monotonie von $L(k)$ vereinfachen. Es gilt wegen $p < q$, dass

$$L(k) = \frac{q(k)}{p(k)} = \frac{q}{p} \cdot \underbrace{\left(\frac{1-q}{1-p}\right)^{k-1}}_{<1},$$

d.h. $L(\cdot)$ ist monoton fallend. Der Test kann daher äquivalent umgeformt werden zu

$$\phi(k) = \begin{cases} \mathbb{P}, & k > k^* \\ \mathbb{Q}, & k \leq k^* \end{cases}$$

für $k \in \Omega$ und einem k^* maximal, so dass $\mathbb{P}(\phi = \mathbb{Q}) = \mathbb{P}(k \leq k^*) \leq \alpha$ gilt.

(Achtung: An der Stelle sollte man sich noch einmal klar machen, dass der erhaltene Test wirklich sinnvoll ist, weil kleine Werte von p , d.h. kleine Erfolgswahrscheinlichkeiten, eine große Anzahl von nötigen Würfeln produzieren. Daher wird \mathbb{P} , der kleinere Wert p durch den Test ausgewählt, wenn k groß ist).

Hier ist das erfüllt, wenn

$$\alpha \geq \mathbb{P}(k \leq k^*) = \sum_{k=1}^{k^*} p(k) = p \cdot \sum_{k=0}^{k^*-1} (1-p)^k = p \cdot \frac{1 - (1-p)^{k^*}}{1 - (1-p)} = 1 - (1-p)^{k^*},$$

d.h. wenn $(1-p)^{k^*} \geq 1 - \alpha$ bzw. $k^* \leq \frac{\log(1-\alpha)}{\log(1-p)}$. Das bedeutet, wir haben k^* als die größte natürliche Zahl zu wählen, die noch $k^* \leq \frac{\log(1-\alpha)}{\log(1-p)} \approx 1.2$ (beachte $\alpha = 0.2, p = 1/6$) erfüllt, d.h. wir müssen $k^* = 1$ wählen.

In diesem Fall beträgt der Fehler 2. Art

$$\mathbb{Q}(\phi = \mathbb{P}) = \mathbb{Q}(k > k^*) = \sum_{k=2}^{\infty} q(k) = 1 - q(1) = 1 - q = 0.75.$$

(c) Das Neyman-Pearson-Lemma trifft nur eine Optimalitätsaussage, wenn die Gleichung $\mathbb{P}(\phi = \mathbb{Q}) = \alpha$ exakt erfüllt ist. $\alpha \in [0, 1]$ muss also so gewählt sein, dass

$$\alpha = \mathbb{P}(\phi = \mathbb{Q}) = \mathbb{P}(k \leq k^*) = 1 - (1-p)^{k^*}$$

für ein k^* exakt erfüllt werden kann. Die Wahl von k^* hängt hierbei nicht vom konkreten Wert von q ab, sondern nur von der Eigenschaft $q > p = \frac{1}{6}$ von q .

Sei \mathbb{Q} (wie oben) die geometrische Wahrscheinlichkeitsverteilung mit Parameter q . Der resultierende Test ϕ^* mit diesem k^* hat laut dem Neyman-Pearson-Lemma folgende Eigenschaft.

Für alle $q > p = \frac{1}{6}$ gilt: ϕ^* minimiert $\phi \mapsto \mathbb{Q}(q)(\phi = \mathbb{P})$ unter allen Tests mit $\mathbb{P}(\phi = \mathbb{Q}(q)) \leq \alpha$.

Das ist gerade die Definition, dass ϕ^* gleichmäßig bester Test für die Hypothesen $H_0 : p = \frac{1}{6}$ gegen $H'_A : p \in \{p' : p' > \frac{1}{6}\}$ ist. □

Aufgabe 4

Auf $\Omega = \{1, 2, 3, 4\}$, versehen mit der Potenzmenge $\mathcal{P}(\Omega)$, seien zwei Wahrscheinlichkeitsverteilungen \mathbb{P} und \mathbb{Q} durch die folgenden Zähldichten p und q gegeben:

ω	1	2	3	4
$p(\omega)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$q(\omega)$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$

Es soll die Hypothese $H_0 = \{\mathbb{P}\}$ gegen die Alternative $H_A = \{\mathbb{Q}\}$ getestet werden.

- Formulieren Sie für dieses Testproblem den Test ϕ aus dem Neyman-Pearson-Lemma. Für welche $\alpha \in [0, 1]$ existiert ϕ und wann liefert er einen optimalen Test zum Niveau α ?
- Berechnen Sie für die Werte von α aus (a) auch jeweils die Wahrscheinlichkeit für den Fehler 2. Art des optimalen Tests.

(4 Punkte)

Beweis. (a) Der Neyman-Pearson-Test lautet

$$\phi^*(\omega) := \begin{cases} \mathbb{P}, & L(\omega) := \frac{q(\omega)}{p(\omega)} < c^*, \\ \mathbb{Q}, & L(\omega) \geq c^* \end{cases}$$

mit $\mathbb{P}(\phi^* = \mathbb{Q}) = \mathbb{P}(L(\omega) \geq c^*) = \alpha$. Hier lautet der Likelihoodquotient:

$$L(1) = \frac{2}{5}, \quad L(2) = \frac{4}{5}, \quad L(3) = \frac{6}{5}, \quad L(4) = \frac{8}{5}.$$

Offensichtlich ist L sogar monoton wachsend in ω , wodurch wir den Test äquivalent umformen können zu

$$\phi^*(\omega) := \begin{cases} \mathbb{P}, & \omega < c^{**}, \\ \mathbb{Q}, & \omega \geq c^{**} \end{cases}$$

mit einem neuen c^{**} . Dieses wird aus der Bedingung

$$\alpha = \mathbb{P}(\phi^* = \mathbb{Q}) = \mathbb{P}(\{\omega : \omega \geq c^{**}\})$$

bestimmt. Laut dem Neyman-Pearson-Lemma muss diese Bedingung (exakt) erfüllt sein, damit der Test ϕ^* überhaupt existiert und die Optimalitätseigenschaft besitzt (Test mit kleinstem Fehler 2. Art unter der Nebenbedingung Fehler 1. Art $\leq \alpha$). Wegen $\omega \in \Omega = \{1, 2, 3, 4\}$ genügt es, $c^{**} \in \{1, 2, 3, 4, 5\}$ zu betrachten, um alle möglichen Werte der rechten Seite der obigen Gleichung zu ermitteln. Wir erhalten

$$\mathbb{P}(\{1, 2, 3, 4\}) = 1, \quad \mathbb{P}(\{2, 3, 4\}) = \frac{3}{4}, \quad \mathbb{P}(\{3, 4\}) = \frac{1}{2}, \quad \mathbb{P}(\{4\}) = \frac{1}{4}, \quad \mathbb{P}(\emptyset) = 0.$$

Das heißt, nur für $\alpha \in \{0, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, 1\}$ existiert der Neyman-Pearson-Test und es gilt die Optimalitätsaussage aus dem Neyman-Pearson-Lemma.

(b) Der Fehler 2. Art lautet:

$$\mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\{\omega : \omega < c^{**}\}).$$

Wir erhalten nacheinander

$$\alpha = 0 : \quad c^{**} = 5 \quad \Rightarrow \quad \mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\{1, 2, 3, 4\}) = 1,$$

$$\alpha = \frac{1}{4} : \quad c^{**} = 4 \quad \Rightarrow \quad \mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\{1, 2, 3\}) = \frac{6}{10},$$

$$\alpha = \frac{2}{4} : \quad c^{**} = 3 \quad \Rightarrow \quad \mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\{1, 2\}) = \frac{3}{10},$$

$$\alpha = \frac{3}{4} : \quad c^{**} = 2 \quad \Rightarrow \quad \mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\{1\}) = \frac{1}{10},$$

$$\alpha = 1 : \quad c^{**} = 1 \quad \Rightarrow \quad \mathbb{Q}(\phi^* = \mathbb{P}) = \mathbb{Q}(\emptyset) = 0.$$

Dies zeigt noch einmal explizit, dass der Fehler 2. Art ansteigt, wenn der Fehler 1. Art kleiner wird. □