Statistik
Dr. Thomas Zehrt Testen von Hypothesen

Motivation

Bei einem Testverfahren wird die aus einer Stichprobe gewonnene Information dazu verwendet, eine Entscheidung über eine **Hypothese** zu treffen. Hypothesen sind hier meist Annahmen über die Verteilung oder über einzelne Parameter der Verteilung eines Merkmals in einer Grundgesamtheit.

Hier ein klassisches Beispiel:

Eine Lady behauptet, dass sie am Geschmack erkennen kann, ob zuerst die Milch oder zuerst der Tee in die Tasse gegossen wurde. Wir wollen versuchen, diese Behauptung zu überprüfen, wobei uns wohl kein direkter Weg einfällt. Wir könnten zwar den Tee mit Milch und die Milch mit Tee physikalisch-chemisch untersuchen, um eventuell existierende Unterschiede aufzuspüren, aber jede hier gewonnene Erkenntnis wäre bestenfalls ein Indiz, aber kein Beweis. Wir werden deshalb die Methoden der Statistik heranziehen, um das Problem (höchstwahrscheinlich) zu lösen. Es gibt offensichtlich zwei (sich gegenseitig ausschliessende) Hypothesen, an die man hier glauben kann:

- die Lady hat die behauptete Fähigkeit nicht und
- die Lady hat die Fähigkeit

In der Testtheorie werden diese beiden Hypothesen nicht symmetrisch behandelt, wir müssen eine von beiden (die dann so genannte Nullhypothese) bevorzugen. Das wird im Allgemeinen die Hypothese sein, die der (im Moment) gültigen Weltanschauung am nächsten kommt.

Ein nicht ganz typischer statistischer Test für die behauptetet Fähigkeit der Lady könnte wie folgt aussehen:

In vier Tassen wird zuerst Tee und dann ein Zusatz Milch gegeben (Tassen vom **Typ 1**). In vier weitere Tassen wird zuerst Milch und dann der Tee eingegossen (Tassen vom **Typ 2**). Der Lady werden die acht Tassen in zufälliger Reihenfolge (d.h. mit der Gleichverteilung auf der Menge der 8! Permutationen) gereicht. Sie wird aufgefordert, **genau** vier Tassen zu benennen, die sie für die vom Typ 1 hält. Benennt sie alle vier Tassen korrekt, so soll der Beweis ihrer Behauptung erbracht sein, denn die Wahrscheinlichkeit, dass dieses Resultat zufällig zu Stande kommt ist mit der hypergeometrischen Verteilung ($N = 8, S = 4, n = 4$ und $s = 4$):

$$\frac{\binom{4}{4}\binom{4}{0}}{\binom{8}{4}} \;=\; \frac{1}{70} \;=\; 0.014.$$

Schwieriger wird das Problem, wenn die Lady behauptet, zwar nicht unfehlbar zu sein, aber doch öfter die richtige Klassifikation herauszuschmecken als das dem Zufall entspricht. Sollte man ihre Behauptung schon glauben, wenn sie (mindestens) drei Tassen vom Typ 1 herausfindet? Die Wahrscheinlichkeit, dass dieses Resultat zufällig zu Stande kommt ist:

$$\frac{\binom{4}{4}\binom{4}{0} + \binom{4}{3}\binom{4}{1}}{\binom{8}{4}} \quad = \quad \frac{17}{70} \;\approx\; 0.24.$$

Daher wären drei richtig klassifizierte Tassen vom Typ 1 kein wirklich überzeugender Beweis ihrer Fähigkeiten.

Natürlich kann man bei solchen Tests Irrtümer (die Lady klassifiziert zufällig genügend viele Tassen richtig) nie ausschliessen, aber man möchte doch meist eine **Grenze für diese Fehlerwahrscheinlichkeit** α setzen. Oft verlangt man, dass der Test so aufgebaut wird, dass die Wahrscheinlichkeit eines solchen Fehlers höchstens $\alpha = 0.05$ ist. Diese Grenze kann natürlich stets eingehalten werden, indem man die Zahl der Tassen erhöht. z.B. kann man bei mehr als 13 Tassen einen Fehler der Lady zulassen, ohne die Schranke $\alpha = 0.05$ zu überschreiten.

Mathematisch ist die Zahl α die obere Grenze für die bedingte Wahrscheinlichkeit:

$$P(\text{Lady klassifiziert oft genug richtig} \mid \text{(obwohl) sie ausschlisslich rät}) \quad \leq \quad \alpha.$$

Benötigtes Schulwissen

- Rechnen mit Binomialkoeffizienten
- Ungleichungen

1 Typisches Testverfahren für die ,,tea tasting Lady“

Eine Lady behauptet, dass sie am Geschmack erkennen kann, ob zuerst die Milch oder zuerst der Tee in die Tasse gegossen wurde. Wir wollen versuchen, diese Behauptung zu überprüfen.

Die folgende Testidee ist aus verschiedenen Gründen besser geeignet, um solche Probleme (statistisch) zu untersuchen als der Test aus der Motivation.

Der Lady wird 10-mal die Aufgabe gestellt, zwei Tassen, von denen eine vom Typ 1 und die andere vom Typ 2 ist, korrekt zu klassifizieren. Damit die Lady unabhängig von früheren Entscheidungen urteilen kann, wird jedes Teilexperiment an einem anderen Tag ausgeführt. Sei X die Zahl der Tage, an denen sie die Tassen richtig klassifiziert.

Als Modell bietet es sich an, X als binomialverteilt mit den Parametern $n = 10$ und p anzunehmen: $X \sim B(10; p)$. X kann also nur die Werte

$$\{\, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \,\} \quad \text{mit} \quad P(X = k) = \binom{10}{k} p^k (1-p)^{n-k}$$

annehmen.

Die sogenannte **Nullhypothese** H_0 ist der Fall $p = 1/2$ und in diesem Fall werden wir $X = 5$ erwarten, d.h. dass die Lady 5 Tassen richtig klassifiziert. Die **Alternative** H_1 kann man durch $p > 1/2$ beschreiben. Wir nehmen also an, dass die Lady, wenn ihre Behauptung stimmt, an jedem Tag unabhängig von den anderen Tagen mit der Wahrscheinlichkeit $p > 1/2$ einen Erfolg erziehlt. Die Auswertung des Testes läuft ähnlich wie im ersten Fall:

Wir werden die Nullhypothese nur dann verwerfen und somit der Lady glauben schenken, wenn sie **genügend** oft Erfolg hat, also sagen wir mindestens x-mal, oder anders ausgedrückt, wenn

$$X \in R := \{\, x, x+1, \ldots, 10 \,\} \subset \{\, 0, 1, 2, \ldots, 5, \ldots, x, x+1, \ldots, 10 \,\}$$

ist. Falls sie allerdings weniger als x Tassen richtig klassifiziert, glauben wir weiterhin an die Nullhypothese. Bei der Bestimmung der Zahl x orientieren wir uns an einer (von uns) festgelegten Irrtumswahrscheinlichkeit α. Wir verlangen also:

$$\begin{aligned}
& P(\, H_0 \text{ verwerfen} \mid H_0 \text{ richtig} \,) \\
= \;& P(X \in R \mid p = 1/2 \,) \\
= \;& \underbrace{P(\, X \in \{x, x+1, \ldots, 10\} \mid p = 1/2 \,)}_{\substack{\text{Wahrscheinlichkeit, dass die Lady mindestens,} \\ \text{x Tassen richtig klassifiziert, obwohl sie die Fähigkeit nicht hat}}} \\
= \;& \sum_{k=x}^{10} \binom{10}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{10-k}
\end{aligned}$$

$$
\begin{aligned}
&= \binom{10}{x} \underbrace{\left(\frac{1}{2}\right)^x \left(1-\frac{1}{2}\right)^{10-x}}_{=\left(\frac{1}{2}\right)^{10}} + \ldots + \binom{10}{9} \underbrace{\left(\frac{1}{2}\right)^9 \left(1-\frac{1}{2}\right)^{10-9}}_{=\left(\frac{1}{2}\right)^{10}} + \binom{10}{10} \underbrace{\left(\frac{1}{2}\right)^{10} \left(1-\frac{1}{2}\right)^{10-10}}_{=\left(\frac{1}{2}\right)^{10}} \\
&= \left(\frac{1}{2}\right)^{10} \sum_{k=x}^{10} \binom{10}{k} \\
&\leq \alpha
\end{aligned}
$$

Oder äquivalent dazu: Finde die kleinste natürliche Zahl x, so dass

$$\sum_{k=x}^{10} \binom{10}{k} \leq 2^{10} \cdot \alpha$$

oder (mit dem Gegenereignis)

$$\sum_{k=0}^{x-1} \binom{10}{k} \geq 2^{10} \cdot (1-\alpha)$$

gilt.

Beispiel 1.1 *Sei $\alpha = 0.05$ und wir nehmen an, dass die Nullhypothese gilt, also $p = 1/2$. Wir suchen eine Zahl x, so dass*

$$
\begin{aligned}
P(\, X \geq x \mid H_0 \text{ richtig }) &= \sum_{k=x}^{10} \binom{10}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10-k} \\
&= \left(\frac{1}{2}\right)^{10} \sum_{k=x}^{10} \binom{10}{k} \\
&= \left(\frac{1}{2}\right)^{10} \left(\binom{10}{x} + \binom{10}{x+1} + \ldots + \binom{10}{10} \right) \\
&\leq 0.05
\end{aligned}
$$

gilt, oder

$$\binom{10}{x} + \binom{10}{x+1} + \ldots + \binom{10}{10} \leq 51.2.$$

Wir untersuchen also die Summe

$$\ldots + \binom{10}{5} + \binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \underbrace{\binom{10}{10}}_{=1}$$

$$\underbrace{\qquad\qquad + \binom{10}{9} + \binom{10}{10}}_{=11}$$

$$\underbrace{\qquad\qquad + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}}_{=56}$$

$$\underbrace{\qquad\qquad + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}}_{\ldots}$$

$$\underbrace{\qquad\qquad + \binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}}_{\ldots}$$

$$\underbrace{\qquad\qquad + \binom{10}{5} + \binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}}_{\ldots}$$

und erkennen, dass wir $x = 9$ wählen müssen, um die Schranke $\alpha = 0.05$ einzuhalten.

2 Grundbegriffe der Testtheorie

Sei X eine Zufallsvariable mit bekanntem Verteilungstyp, d.h. wir kennen den Typ der Verteilung P von X. Unbekannt sei aber ein Parameter θ der Verteilung P von X und wir schreiben deshalb besser P_θ. Bei unserem Startbeispiel ist die genutzte Zufallsvariable binomialverteilt $X \sim B(10, p)$ und der unbekannte Parameter ist $\theta = p$. Die Menge aller möglichen Verteilungen von X bilden den

$$\underline{\text{Hypothesenraum}}: \quad \Omega = \left\{ P_\theta \; = \; P_\theta(X = x) \; : \; \theta \in \Theta \right\}$$

und Θ ist die Menge der überhaupt erlaubten Werte des Parameters. Bei einem (parametrischen) Testproblem wird die Menge Θ in zwei Teilmengen aufgeteilt:

- die zu testende Nullhypothese $H_0 : \theta \in \Theta_0$ und
- die Alternative $H_1 : \theta \in \Theta_1$.

Dabei gilt stets: $\Theta_0 \cap \Theta_1 = \emptyset$ und bei Signifikanztests (die wir im weiteren ausschliesslich behandeln werden) auch $\Theta_0 \cup \Theta_1 = \Theta$.

Die typischen Situationen dabei sind:

	H_0	H_1	
1.	$H_0 : \theta = \theta_0$	$H_1 : \theta = \theta_1$, $\theta_0 \neq \theta_1$	einfache Hypothese
2.	$H_0 : \theta = \theta_0$	$H_1 : \theta \neq \theta_0$	zweiseitigen Fragestellung
3.	$H_0 : \theta \leq \theta_0$	$H_1 : \theta > \theta_0$	einseitige Fragestellung
4.	$H_0 : \theta \geq \theta_0$	$H_1 : \theta < \theta_0$	einseitige Fragestellung

Eine Funktion $T(\mathbf{X}) = T(X_1, \ldots, X_n)$ der Stichprobenvariablen $\mathbf{X} = (X_1, \ldots, X_n)$ heisst Testgrösse. Für die konkrete Stichprobe $(x_1, \ldots, x_n)$ ergibt sich $t = T(x_1, \ldots, x_n)$ als Realisierung der Testgösse.

Der Wertebereich der Zufallsgrösse $T(\mathbf{X})$ wird in folgende zwei Teile zerlegt:

- Verwerfungsbereich, kritischer Bereich oder Ablehnbereich R
- Annahmebereich R^c.

Beispiel 2.1 *Für das Beispiel der Lady gilt somit:*

$$\Omega = \left\{ B(10, p) \; : \; p \in [0, 1] = \Theta \right\}$$

- $H_0 : p \in \Theta_0 = \{1/2\}$ *(nur ein Wert) und*
- $H_1 : p \in \Theta_1 = (1/2, 1]$ *(ganzes Intervall)*

Hier gilt **nicht** $\Theta = [0, 1] = \Theta_0 \cup \Theta_1 = \{1/2\} \cup (1/2, 1]$ *(kein typischer Test).*

$$T \; = \; X \quad = \quad X_1 + X_2 + \ldots + X_{10}$$

wobei X_i entweder 1 *oder* 0 *ist, je nach dem ob die Lady am i-ten Tag richtig oder falsch klassifiziert.*

Aufgrund der Realisierung $(x_1, \ldots, x_n)$ wird dann folgende Testentscheidung getroffen:

- H_0 ablehnen, falls $t = T(x_1, \ldots, x_n) \in R$
- H_0 beibehalten, falls $t = T(x_1, \ldots, x_n) \in R^c$

Innerhalb des gewählten Modells gibt es nun vier Möglichkeiten, wie die Realität und die Testentscheidung zusammentreffen können.

		Realität	
		H_0 ist richtig	H_0 ist falsch
Testent-scheidung	H_0 beibehalten	ok	Fehler 2. Art, β-Fehler
	H_0 verwerfen	Fehler 1. Art, α-Fehler	ok

Bei der Testkonstruktion gibt man die Wahrscheinlichkeit α eines Fehlers 1. Art vor. Diese Schranke bezeichnet man als Signifikanzniveau und der Test heisst dann natürlich Signifikanztest zum Niveau α.

Der kritische Bereich R wird dann so konstruiert, dass die Wahrscheinlichkeit eines Fehlers 1. Art nicht grösser als α wird:

$$P(\underbrace{H_0 \text{ verwerfen}}_{T(\mathbf{X}) \in R} | H_0 \text{ richtig }) \leq \alpha$$

Dabei sollte aber die Wahrscheinlichkeit eines Fehlers 2. Art $P(H_0$ beibehalten $|H_0$ falsch $)$ nicht zu gross werden. Hierin liegt eine **unsymmetrische Behandlung** der beiden Risiken. Das Risiko einen Fehler 1. Art zu machen wird stärker gescheut.

Ist man daran interessiert, ob irgendwelche Daten innerhalb einer bestehenden Theorie erklärbar sind oder auf eine neue Theorie hindeuten, so sollte man im Zweifelsfall bei der bestehenden Theorie bleiben.

Wird ein neues mit einem alten Medikament verglichen, so wird man bei unklaren Werten beim alten und erprobten Medikament bleiben. Im Testproblem trägt man dieser Überlegung Rechnung, in dem man als Hypothese die Verteilung(en) wählt, die der etablierten Theorie oder der reinen Zufälligkeit entspricht. Deshalb verwendet man auch oft das Wort Nullhypothese.

Allgemeines Vorgehen:

1. Verteilungsannahme über die Zufallsvariable X (bzw. über deren Verteilungsfunktion F) machen
2. Formulieren der Nullhypothese und der Alternative.
3. Vorgabe der Irrtumswahrscheinlichkeit α
4. Konstruktion einer geeigneten Testgrösse $T(\mathbf{X}) = T(X_1, \ldots, X_n)$ als Funktion der Stichprobenvariablen $\mathbf{X}$, deren Verteilung unter der Nullhypothese vollständig bekannt sein muss.
5. Wahl eines kritischen Bereichs R aus dem möglichen Wertebereich von $T(\mathbf{X})$, so dass $P(T(\mathbf{X}) \in R | H_0 \text{ richtig }) \leq \alpha$ gilt.
6. Bestimmung der Realisierung $t = T(x_1, \ldots, x_n)$ der Testgrösse $T(\mathbf{X})$ anhand einer konkreten Stichprobe.
7. Entscheidungsregel: Liegt t in R, so wird die Nullhypothese abgelehnt, sonst beibehalten.

3 Signifikanztests für die Wahrscheinlichkeit p

Ein Zufallsexperiment, bei dem ein Ereignis E mit einer unbekannten Wahrscheinlichkeit $\theta = p$ eintritt, wird n-mal durchgeführt. Die Zufallsvariable X zähle die Anzahl der Durchführungen, bei denen das Ereignis E eingetreten ist.

Natürlich gilt wieder $X \sim B(n;p)$ für irgend ein unbekanntes $p \in [0,1]$.

Nun führen wir das Experiment tatsächlich n-mal durch und erhalten $X = t$, d.h. t-mal den Ausgang E (und $(n-t)$-mal den Ausgang E^c).

3.1 Zweiseitiger Test

gegeben:	$H_0 : p = p_0$ $H_1 : p \neq p_0$ α	
Verwerfungsbereich:	$R = \underbrace{\{0,1,\ldots,x_l\}}_{R_l} \cup \underbrace{\{x_r,\ldots,n\}}_{R_r}$	
x_l grösste Zahl s.d.	$\sum_{k=0}^{x_l} \binom{n}{k} p_0^k \, (1-p_0)^{n-k}$	$\leq \frac{\alpha}{2}$
x_r kleinste Zahl s.d	$\sum_{k=0}^{x_r-1} \binom{n}{k} p_0^k \, (1-p_0)^{n-k}$	$\geq 1 - \frac{\alpha}{2}$

Konstruktion des Verwerfungsbereiches

$$\begin{aligned} P(H_0 \text{ verwerfen } \mid H_0 \text{ richtig }) &= P(X \in R \mid p = p_0) \\ &= P(X \in R_l \text{ oder } X \in R_r \mid p = p_0) \\ &= P(X \in R_l \mid p = p_0) + P(X \in R_r \mid p = p_0) \\ &\leq \alpha \end{aligned}$$

Wir bestimmen nun die Grenzen x_l und x_r der beiden Komponenten von R durch die Forderungen:

$$P(X \in R_l \mid p = p_0) = \sum_{k=0}^{x_l} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \leq \frac{\alpha}{2}$$

$$P(X \in R_r \mid p = p_0) = \sum_{k=x_r}^{n} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} = 1 - \sum_{k=0}^{x_r-1} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \leq \frac{\alpha}{2}$$

3.2 Linksseitiger Test

gegeben:	$H_0 : p \geq p_0$ $H_1 : p < p_0$ α
Verwerfungsbereich:	$R = \{0, 1, \ldots, x\}$
x grösste Zahl s.d.	$\sum_{k=0}^{x} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \leq \alpha$

Konstruktion des Verwerfungsbereiches

$$\begin{aligned} P(H_0 \text{ verwerfen } \mid H_0 \text{ richtig }) &= P(X \in R \mid p \geq p_0) \\ &\leq P(X \in R \mid p = p_0) \\ &\leq \alpha \end{aligned}$$

Wir bestimmen nun die Grenze x durch die Forderung:

$$P(X \in R \mid p = p_0) = \sum_{k=0}^{x} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \leq \alpha$$

3.3 Rechtsseitiger Test

gegeben:	$H_0 : p \leq p_0$ $H_1 : p > p_0$ α
Verwerfungsbereich:	$R = \{x, x+1, \ldots, n\}$
x kleinste Zahl s.d.	$\sum_{k=0}^{x-1} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \geq 1 - \alpha$

Konstruktion des Verwerfungsbereiches

$$\begin{aligned} P(H_0 \text{ verwerfen } \mid H_0 \text{ richtig }) &= P(X \in R \mid p \leq p_0) \\ &\leq P(X \in R \mid p = p_0) \\ &\leq \alpha \end{aligned}$$

Wir bestimmen nun die Grenze x durch die Forderung:

$$P(X \in R \mid p = p_0) = \sum_{k=x}^{n} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} = 1 - \sum_{k=0}^{x-1} \binom{n}{k} p_0^k \, (1-p_0)^{n-k} \leq \alpha$$

Aufgabe 3.1 *Ein Hersteller von Überraschungseiern behauptet, dass in mindestens* 14% *der Eier Figuren stecken. Sie wollen das testen und nehmen eine Stichprobe von* $n = 1000$ *Eier in denen Sie* 130 *Figuren finden. Genügt dieses Ergebnis, um die Behauptung des Herstellers, mit einer genügend hohen Wahrscheinlichkeit, zu widerlegen?*

Lösung:

- Stichprobengrösse $n = 1000$
- $T = X :=$ Anzahl Figuren in der Stichprobe also $X \sim B(1000, p)$
- Linksseitiger Test $H_0 : p \geq 0.14$ (Behauptung des Herstellers) und $H_1 : p < 0.14$
- Verwerfungsbereich: $R = \{0, 1, 2, \ldots, x\}$, s.d. x die grösste Zahl mit

$$P(0 \leq X \leq x \mid p = 0.14) \;=\; \sum_{k=0}^{x} \binom{1000}{k} \cdot 0.14^k \cdot 0.86^{1000-k} \;\leq\; 0.05$$

- Lokaler Grenzwertsatz $\mu = 1000 \cdot 0.14 = 140$ und $\sigma = \sqrt{1000 \cdot 0.14 \cdot 0.86} = 10.97$ und somit

$$\sum_{k=0}^{x} \binom{1000}{k} \cdot 0.14^k \cdot 0.86^{1000-k} \;\approx\; \Phi\left(\frac{x - 140 + 0.5}{10.97}\right) - \underbrace{\Phi\left(\frac{0 - 140 - 0.5}{10.97}\right)}_{\approx 0}$$

- Bestimme x so, dass folgendes gilt:

$$\Phi\left(\frac{x - 139.5}{10.97}\right) \;=\; 0.05$$

- Gleichwertig dazu (denn stets gilt $\Phi(z) + \Phi(-z) = 1$):
  Bestimme x so, dass folgendes gilt:

$$\Phi\left(-\frac{x - 139.5}{10.97}\right) \;=\; 1 - 0.05 \;=\; 0.95$$

- Mit Hilfe der Formelsammlung erhält man, dass Φ an der Stelle 1.645 den Wert 0.95 hat. Dann müssen wir die Gleichung

$$-\frac{x - 139.5}{10.97} \;=\; 1.645$$

  nach x auflösen und wir erhalten $x = 121.45$. Sicherheitshalber abrunden auf $x = 121$.

- Verwerfungsbereich: $R = \{0, 1, 2, \ldots, 120, 121\}$ und da wir 130 Figuren gefunden haben, können wir die Behauptung des Herstellers **nicht** (mit der gewünschten Sicherheit) verwerfen.

4 Übungsaufgaben

1. In einem verregneten Land beträgt die Regenwahrscheinlichkeit in den Herbstmonaten 50%.
   Jeden Morgen im Herbst fragt sich Susi, ob sie einen Regenschirm mitnehmen soll oder nicht. Um zu einer Entscheidung zu kommen, wirft sie eine faire Münze. Wirft sie Kopf, nimmt sie einen Regenschirm mit, ansonsten lässt sie den Schirm zu Hause.

   (a) Betrachten Sie die Situation wie einen statistischen Test. Wie müssen die Hypothesen gewählt werden, damit der Fehler 1.Art die schlimmere Auswirkung darstellt?

   (b) Bestimmen Sie die Wahrscheinlichkeit für den Fehler 1. Art.

2. In einer Prüfung werden 10 Fragen gestellt, die nur mit ,,Ja" oder ,,Nein" zu beantworten sind. Der Dozent legt fest, dass ein Student der 8 oder mehr Fragen richtig beantwortet, die Prüfung besteht. Die Nullhypothese H_0 sei: der Student hat (ausschliesslich) geraten. Bestimmen Sie die Wahrscheinlichkeit eines Fehlers 1. Art.

3. Max hat die Vermutung, dass Moritz einen Würfel so gezinkt hat, dass die Zahl 6 seltener als das für einen fairen Würfel zu erwarten wäre, fällt. Er überlegt sich folgendes: ,,Ich werde den Würfel 30-mal werfen. Dann sollte ich etwa 5-mal eine 6 erhalten. Erhalte ich keine oder nur eine 6, so werde ich annehmen, dass der Würfel gezinkt ist".

   Modellieren Sie das Experiment. Was ist H_0? Was ist H_1? Verwerfungsbereich? Wahrscheinlichkeit eines Fehlers 1. Art?

4. Bei der Massenproduktion eines Produktes treten immer wieder unbrauchbare Stücke auf. Der Produzent versichert, dass ihr Anteil p nicht mehr als 2% beträgt. Eine Stichprobe vom Umfang $n = 5000$ ergab 120 unbrauchbare Produkte. Formulieren Sie H_0 und H_1. Bestimmen Sie den Verwerfungsbereich des Testes zum Signifikanzniveau $\alpha = 0.05$. Sollte man H_0 ablehnen?

Lösungen einiger Übungsaufgaben

1. (a) Testproblem: Regen oder kein Regen

   Entscheidungsregel: Ich glaube, dass es regnen wird, wenn Münzwurf ‚Kopf, zeigt.

Entscheidung / Realität	Regen	kein Regen
„Kopf" → Regen → Schirm	trocken	Schirm umsonst
„Zahl" → kein Regen → kein Schirm	nass	trocken

   Mögliche Fehler: Schirm umsonst mitgenommen bzw. nass werden(ist wohl die schlimmere Konsequenz)

   Damit der Fehler 1. Art die schlimmere Konsequenz darstellt, müssen die Hypothesen wie folgt gewählt werden:

   H_0 : Regen

   H_1 : kein Regen

   (b) $P(\ H_0 \text{ verworfen } \mid H_0 \text{ stimmt }) \ = \ 0.5$

2. $\alpha = 5.5\%$

3. $R = \{0, 1\}$ (kann man der Aufgabenstellung entnehmen) und $\alpha = 2.9\%$

4. $H_0 : p \leq 0.02$ gegen $H_1 : p > 0.02$,

   X zähle die Anzahl unbrauchbarer Produkte in der Stichprobe, $X \sim B(5000, p)$

   Hinweis: Nutzen Sie den Grenzwertsatz von De Moivre und Laplace.