Wahrscheinlichkeitstheorie und Statistik

Serie 10 - Lösungen

MC 10-1. Was bedeutet ein Signifikanzniveau von 5%? (Mehrere richtige Antworten sind möglich.)

(a) Bei einem positiven Testergebnis besteht eine 5%ige Wahrscheinlichkeit für die Alternativhypothese.

(b) Die Wahrscheinlichkeit eines Fehlers 2. Art beträgt 5%.

(c) Es besteht eine 5%ige Wahrscheinlichkeit, dass die Nullhypothese fälschlich verworfen wird.

(d) Das Testergebnis ist zu 95% richtig.

Lösung: Nur (c) ist richtig.

(a) Falsch. Für die Korrektheit der Alternativhypothese ist keine Wahrscheinlichkeit definiert.

(b) Falsch. Die Wahrscheinlichkeit eines Fehlers 2. Art hängt mit der Macht des Tests zusammen.

(c) Richtig.

(d) Falsch. Dies stimmt nur unter der Nullhypothese, aber nicht unter der Alternative.

MC 10-2. Was bedeutet ein P-Wert von 0.03 beim Testen eines Effekts auf Signifikanzniveau 0.05? (Mehrere richtige Antworten sind möglich.)

(a) Die Nullhypothese wird beibehalten.

(b) Der Effekt ist nicht signifikant.

(c) Die Nullhypothese wird verworfen.

(d) Ein Fehler 2. Art wurde gemacht.

Lösung: Nur (c) ist richtig.

(a) Falsch. Der P-Wert ist unterhalb des Signifikanzniveaus und daher wird die Nullhypothese verworfen.

(b) Falsch. Testergebnisse mit P-Werten unterhalb des Signifikanzniveaus werden als signifikant bezeichnet.

(c) Richtig.

(d) Falsch. Dies hängt vom tatsächlichen Bestehen eines Effekts ab, was nicht bekannt ist.

MC 10-3. Ordnen Sie den statistischen Fragestellungen (a) bis (d) geeignete Testmethoden (1) bis (4) zu. (Die Zuordnung ist bijektiv.)

(a) Unterscheidet sich die durchschnittliche Körpergröße von Männern und Frauen signifikant?

(b) Zeigt ein neuer Online-Werbetext eine höhere Klickrate als die bisherige Version?

(c) Gibt es einen Zusammenhang zwischen dem Impfstatus (geimpft /ungeimpft) und dem Auftreten von grippeähnlichen Symptomen (Ja / Nein) in einer Stichprobe von 500 Personen?

(d) Erreicht ein Logistik-Unternehmen die versprochene Quote von 95% pünktlichen Lieferungen?

(1) Einseitiger Z-Test

(2) Einseitiger t-Test

(3) Beidseitiger t-Test

(4) Chi-Quadrat Test

Lösung:

(a) — (3) Beidseitiger t-Test

(b) — (2) Einseitiger t-Test

(c) — (4) Chi-Quadrat Test

(d) — (1) Einseitiger Z-Test (die Varianz ist bei Normalapproximation der Binomialverteilung bekannt)

Aufgabe 10-4. Die Australier Mr. Smith und Dr. Thurston streiten sich über das Durchschnittsgewicht μ von Straußeneiern. Beide sind damit einverstanden, das Gewicht approximativ als normalverteilt aufzufassen. Mr. Smith behauptet, das mittlere Gewicht sei $\mu_0 = 1100$ Gramm, während Dr. Thurston darauf besteht, es liege im Schnitt bei $\mu_1 = 1200$ Gramm. Sie reisen also nach Afrika und suchen Straußeneier. Da diese aber gut versteckt sind, finden sie nur $n = 8$ Eier, deren Gewichte 1090, 1150, 1170, 1080, 1210, 1230, 1180 und 1140 Gramm betragen. Die Standardabweichung ist beiden bekannt und beträgt $\sigma = 55$ Gramm.

(a) Dr. Thurston schlägt vor, Mr. Smiths Behauptung als Hypothese $\mu = 1100$ gegen seine Alternative $\mu > 1100$ auf dem 5%-Niveau zu testen.

(b) Mr. Smith schlägt vor, Dr. Thurstons Behauptung als Hypothese $\mu = 1200$ gegen seine Alternative $\mu < 1200$ auf dem 5%-Niveau zu testen.

Wie lauten die Test-Ergebnisse?

Lösung: Der Mittelwert der 8 Eier ist

$$\bar{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{8}\sum_{k=1}^{8} x_k = 1156.25$$

(a) Unter Mr. Smiths Nullhypothese $\mu = \mu_0$ ist die folgende Test-Statistik ist approximativ standardnormalverteilt:

$$Z_0 = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{1156.25 - 1100}{55/\sqrt{8}} = 2.89$$

Das 95 % Quantil der Standardnormalverteilung ist $z_{0.95} = 1.645$. Da $Z_0 > z_{0.95}$, wird Mr. Smiths Nullhypothese $\mu = 1100$ auf dem 5%-Niveau verworfen.

(b) Unter Dr. Thurstons Nullhypothese $\mu = \mu_1$ ist die folgende Test-Statistik approximativ standardnormalverteilt:

$$Z_1 = \frac{\bar{x}_n - \mu_1}{\sigma/\sqrt{n}} = \frac{1156.25 - 1200}{55/\sqrt{8}} = -2.25$$

Das 5 % Quantil der Standardnormalverteilung ist $z_{0.05} = -1.645$. Da $Z_1 < z_{0.05}$, wird Dr. Thurstons Nullhypothese $\mu = 1200$ auf dem 5%-Niveau verworfen.

Dieses Beispiel illustriert sehr schön die Bedeutung der Wahl von Hypothese und Alternative und auch ihre asymmetrische Behandlung. Mit dem ersten Test würde man Dr. Thurston Recht geben, mit dem zweiten hingegen Mr. Smith. Und das passiert trotz völlig identischer Daten!

Aufgabe 10-5. Ein Pharmahersteller verkauft homöopathische Pillen gegen Stress und will deren Wirksamkeit prüfen. Dazu verwendet er ein standardisiertes Verfahren zum Messen des Stresslevels. Das Verfahren ist so geeicht, dass bei Einnahme von Placebo-Pillen eine Stressverminderung ($X = 1$) gleich häufig ausgewiesen wird wie eine Stresserhöhung ($X = 0$). 30 Testpersonen nehmen die homöopathischen Pillen ein und lassen ihr Stresslevel mit demselben Verfahren messen. Die Ergebnisse $X_1, \dots, X_{30}$ werden als unabhängig Bernoulli-verteilt mit unbekanntem Parameter $p \in [0, 1]$ angenommen.

(a) Der Hersteller würde gerne zeigen, dass seine Pillen tatsächlich Heilkraft haben ($p > 1/2$). Konstruieren Sie dafür mit Hilfe der Normalapproximation der Binomialverteilung einen Test mit Signifikanzniveau von 5%.

(b) Der Hersteller ist besorgt, dass das Testverfahren nicht mächtig genug ist, kleine Effekte nachzuweisen. Mit welcher Wahrscheinlichkeit würde der Test einen Effekt von $p = 0.6$ nachweisen? Würde es helfen, die Anzahl der Testpersonen von 30 auf 50 zu erhöhen?

(c) Da sich Pillen mit signifikantem Effekt gut verkaufen lassen und die Entwicklung wirksamer Medikamente teuer ist, verwendet ein unlauterer Konkurrent 100 verschiedene Placebos und lässt diese jeweils nach dem obigen Verfahren testen. Bei wie vielen dieser Placebos ist ein fälschlicher Nachweis eines Effekts zu erwarten?

Lösung:

(a) Sei $Y_{30} = \sum_{i=1}^{30} X_i$ die Anzahl der Erfolge. Der Test verwirft die Nullhypothese, wenn

$$\frac{Y_{30} - 30 * 0.5}{\sqrt{30 * 0.5 * 0.5}} \geq z_{0.95} \approx 1.64 \quad \iff \quad Y_{30} \geq 30 * 0.5 + \sqrt{30 * 0.5 * 0.5}\, z_{0.95} \approx 19.5,$$

wobei $z_{0.95}$ das 95%-Quantil der Standardnormalverteilung ist.

(b) Die Macht bei 30 Testpersonen und $p = 0.6$ ist

$$\begin{aligned} \mathbb{P}_{0.6}(Y_{30} \geq 19.5) &= \mathbb{P}_{0.6}\left(\frac{Y_{30} - 30 * 0.6}{\sqrt{30 * 0.6 * 0.4}} \geq \frac{19.5 - 30 * 0.6}{\sqrt{30 * 0.6 * 0.4}}\right) \\ &\approx 1 - \Phi\left(\frac{19.5 - 30 * 0.6}{\sqrt{30 * 0.6 * 0.4}}\right) \approx 28.8\%, \end{aligned}$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist. Bei 50 Testpersonen würde der Test die Nullhypothese verwerfen, wenn

$$\frac{Y_{50} - 50 * 0.5}{\sqrt{50 * 0.5 * 0.5}} \geq z_{0.95} \approx 1.64 \quad \Longleftrightarrow \quad Y_{50} \geq 50 * 0.5 + \sqrt{50 * 0.5 * 0.5}\, z_{0.95} \approx 30.8.$$

Die Macht bei 50 Testpersonen und $p = 0.6$ ist

$$\mathbb{P}_{0.6}(Y_{50} \geq 30.8) \approx 1 - \Phi\left(\frac{30.8 - 50 * 0.6}{\sqrt{50 * 0.6 * 0.4}}\right) \approx 44.3\%.$$

(c) Bei 5 Placebos.

Aufgabe 10-6. Beschreiben Sie den Algorithmus zur Entfernungsbestimmung mit Sonar [2] und beantworten Sie folgende Fragen:

(a) Wie lautet das statistische Modell?

(b) Mit welcher statistischen Methode wird die Verzögerung des Echos geschätzt?

Hinweis: Das Problem ist detailliert in [1, Illustation 7] beschrieben.

Lösung:

(a) Der unbekannte Parameter ist die Verzögerung t. Für diesen Parameter wird entweder gar keine oder die uniforme Verteilung angenommen, mit demselben Resultat für den entsprechenden Schätzer. Das beobachtete Signal ist ein Impuls mit rechts-Shift um t plus normalverteiltes Rauschen:

$$\texttt{signal} = S_t(\texttt{pulse}) + \epsilon.$$

(b) Der Schätzer ist ein Maximum-Likelihood Schätzer oder, äquivalent, ein Maximum-a-posteriori Schätzer mit uniform verteilter Verzögerung, aus folgendem Grund:

- Der Schätzer ist $\arg\max_t \texttt{matched\_filter\_output}_t$.
- $\texttt{matched\_filter\_output}_t$ ist das Skalarprodukt von `signal` und $S_t(\texttt{pulse})$, dem rechts-Shift des Impulses um t.
- Dieses Skalarprodukt ist bis auf additive Konstante, die nicht von t abhängt, gleich der quadratischen Abweichung $\|\texttt{signal} - S_t(\texttt{pulse})\|^2$.
- Diese quadratische Abweichung ist bis auf eine affine Transformation, die nicht von t abhängt, gleich der negativen log-likelihood $-\log f(\texttt{signal}|t)$.
- Wenn für die Verzögerung t eine uniform Verteilung angenommen wird, stimmt dies bis auf eine additive Konstante mit der negativen log-a-posteriori Dichte überein.

References

[1] Pierre Brémaud. *An Introduction to Probabilistic Modeling*. Springer, 1988.

[2] Philipp Harms. *Entfernungsbestimmung mit Sonar*. Verfügbar unter: `https://gist.github.com/philipp-harms/2aa880e579f318b5488ae1f237d216ab`. April 2025.