KAPITEL 10

Tests statistischer Hypothesen

In der Statistik muss man oft Hypothesen testen, z.B. muss man anhand einer Stichprobe entscheiden, ob ein unbekannter Parameter einen vorgegebenen Wert annimmt. Zuerst betrachten wir ein Beispiel.

10.1. Ist eine Münze fair?

Es sei eine Münze gegeben. Wir wollen testen, ob diese Münze fair ist, d.h. ob die Wahrscheinlichkeit von "Kopf", die wir mit θ bezeichnen, gleich $1/2$ ist. Dazu werfen wir die Münze z.B. $n = 200$ Mal. Sei S die Anzahl der Würfe, bei denen die Münze Kopf zeigt. Nun betrachten wir zwei Hypothesen:

(1) Nullhypothese H_0: Die Münze ist fair, d.h., $\theta = 1/2$.
(2) Alternativhypothese H_1: Die Münze ist nicht fair, d.h., $\theta \neq 1/2$.

Wir müssen uns entscheiden, ob wir die Nullhypothese H_0 verwerfen oder beibehalten. Die Entscheidung muss anhand des Wertes von S getroffen werden. Unter der Nullhypothese gilt, dass $\mathbb{E}_{H_0} S = 200 \cdot \frac{1}{2} = 100$. Die Idee besteht nun darin, die Nullhypothese zu verwerfen, wenn S stark von 100 abweicht. Dazu wählen wir eine Konstante $c \in \{0, 1, \ldots\}$ und verwerfen H_0, falls $|S - 100| > c$. Andernfalls behalten wir die Hypothese H_0 bei. Bei diesem Vorgehen können wir zwei Arten von Fehlern machen:

(1) Fehler 1. Art: H_0 wird verworfen, obwohl H_0 richtig ist.
(2) Fehler 2. Art: H_0 wird nicht verworfen, obwohl H_0 falsch ist.

Wie sollte nun die Konstante c gewählt werden? Man möchte natürlich die Wahrscheinlichkeiten der beiden Arten von Fehlern klein halten. In diesem Beispiel ist es allerdings nicht möglich, die Wahrscheinlichkeit eines Fehlers 2. Art zu bestimmen. Der Grund dafür ist, dass man für die Berechnung dieser Wahrscheinlichkeit den Wert von θ kennen muss, bei einem Fehler 2. Art ist allerdings nur bekannt, dass $\theta \neq 1/2$ ist. Die Wahrscheinlichkeit eines Fehlers 1. Art kann aber sehr wohl bestimmt werden und ist

$$\mathbb{P}_{H_0}[|S - 100| > c] = 2\mathbb{P}_{H_0}[S > 100 + c] = 2 \sum_{k=100+c+1}^{200} \binom{n}{k} \frac{1}{2^n},$$

da $S \sim \text{Bin}(200, 1/2)$ unter H_0. Wir wollen nun c so wählen, dass die Wahrscheinlichkeit eines Fehlers 1. Art nicht größer als ein kleines vorgegebenes Niveau $\alpha \in (0, 1)$ ist. Normalerweise wählt man $\alpha = 0.01$ oder 0.05. Hier wählen wir das Niveau $\alpha = 0.05$. Nun rechnet man nach, dass

$$\mathbb{P}_{H_0}[|S - 100| > c] = \begin{cases} 0.05596, & \text{für } c = 13, \\ 0.04003, & \text{für } c = 14. \end{cases}$$

Damit die Wahrscheinlichkeit eines Fehlers 1. Art kleiner als $\alpha = 0.05$ ist, müssen wir also $c \geq 14$ wählen. Dabei ist es sinnvoll, c möglichst klein zu wählen, denn sonst vergrößert man die Wahrscheinlichkeit eines Fehlers 2. Art. Somit wählen wir $c = 14$. Unsere Entscheidungsregel lautet nun wie folgt:

(1) Wir verwerfen H_0, falls $|S - 100| > 14$.
(2) Sonst behalten wir die Hypothese H_0 bei.

In diesem Beispiel kann man für die Berechnung der Wahrscheinlichkeiten die Approximation durch die Normalverteilung benutzen. Es soll ein c mit

$$\mathbb{P}_{H_0}[S - 100 < -c] \leq \frac{\alpha}{2}$$

bestimmt werden. Um die Güte der Approximation zu verbessern, benutzen wir den $\frac{1}{2}$–Trick. Da c ganz ist, ist die obige Ungleichung äquivalent zu

$$\mathbb{P}_{H_0}[S - 100 \leq -c - 0.5] \leq \frac{\alpha}{2}.$$

Unter H_0 gilt $S \sim \text{Bin}(200, 1/2)$ und somit $\mathbb{E}_{H_0} S = 100$, $\text{Var}_{H_0} S = 200 \cdot \frac{1}{2} \cdot \frac{1}{2} = 50$. Die obige Ungleichung ist äquivalent zu

$$\mathbb{P}_{H_0}\left[\frac{S - 100}{\sqrt{50}} \leq -\frac{c + 0.5}{\sqrt{50}}\right] \leq \frac{\alpha}{2}.$$

Nun können wir die Normalverteilungsapproximation benutzen und die obige Ungleichung durch die folgende ersetzen:

$$\Phi\left(-\frac{c + 0.5}{\sqrt{50}}\right) \leq \frac{\alpha}{2}$$

Somit muss für c die folgende Ungleichung gelten:

$$\frac{c + 0.5}{\sqrt{50}} \geq -z_{\frac{\alpha}{2}}.$$

Wegen der Symmetrie der Standardnormalverteilung gilt $-z_{\frac{\alpha}{2}} = z_{1-\frac{\alpha}{2}}$. Für $\alpha = 0.05$ ist $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$ und somit ist die obige Ungleichung äquivalent zu $c \geq 13.36$. Somit müssen wir $c = 14$ wählen. Die Entscheidungsregel bleibt genauso wie oben.

10.2. Allgemeine Modellbeschreibung

Wir beschreiben nun allgemein das statistische Testproblem. Sei $X = (X_1, \ldots, X_n)$ eine Stichprobe von unabhängigen und identisch verteilten Zufallsvariablen mit Dichte bzw. Zähldichte h_θ, wobei $\theta \in \Theta$ unbekannt sei. Es sei außerdem eine Zerlegung des Parameterraums Θ in zwei disjunkte Teilmengen Θ_0 und Θ_1 gegeben, d.h.

$$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset.$$

Wir betrachten nun zwei Hypothesen:

(1) Die Nullhypothese H_0: $\theta \in \Theta_0$.
(2) Die Alternativhypothese H_1: $\theta \in \Theta_1$.

Wir sollen anhand der Stichprobe $X_1, \ldots, X_n$ entscheiden, ob wir H_0 verwerfen oder beibehalten. Dazu wählen wir eine Borel–Menge $K \subset \mathbb{R}^n$, die Ablehnungsbereich genannt wird. Die Entscheidung wird nun wie folgt getroffen:

(1) Wir verwerfen H_0, falls $(X_1, \dots, X_n) \in K$.
(2) Wir behalten H_0 bei, falls $(X_1, \dots, X_n) \notin K$.

Diese Entscheidungsregel kann auch mit Hilfe einer Funktion $\varphi : \mathbb{R}^n \to \{0, 1\}$ formuliert werden, wobei

$$\varphi(x) = \begin{cases} 1, & \text{falls } x \in K, \\ 0, & \text{falls } x \notin K. \end{cases}$$

Die Nullhypothese wird verworfen, falls $\varphi(X_1, \dots, X_n) = 1$ und wird beibehalten, falls $\varphi(X_1, \dots, X_n) = 0$. Nun können zwei Arten von Fehlern machen:

(1) Fehler 1. Art: H_0 wird verworfen, obwohl H_0 richtig ist.
(2) Fehler 2. Art: H_0 wird nicht verworfen, obwohl H_0 falsch ist.

Normalerweise versucht man K bzw. φ so zu wählen, dass die Wahrscheinlichkeit eines Fehlers 1. Art durch ein vorgegebenes Niveau $\alpha \in (0, 1)$ beschränkt ist, typischerweise $\alpha = 0.01$ oder 0.05.

Definition 10.2.1. Eine Borel–Funktion $\varphi : \mathbb{R}^n \to \{0, 1\}$ heißt *Test* zum Niveau $\alpha \in (0, 1)$, falls

$$\mathbb{P}_\theta[\varphi(X_1, \dots, X_n) = 1] \leq \alpha \ \text{ für alle } \theta \in \Theta_0.$$

Im Folgenden werden wir zahlreiche Beispiele von Tests konstruieren.

10.3. Tests für die Parameter der Normalverteilung

Seien $X_1, \dots, X_n \sim \mathrm{N}(\mu, \sigma^2)$ unabhängige und mit Parametern (μ, σ^2) normalverteilte Zufallsvariablen. Wir wollen Hypothesen über die Parameter μ und σ^2 testen. Wir werden folgende vier Fälle betrachten:

(1) Tests für μ bei bekanntem σ^2.
(2) Tests für μ bei unbekanntem σ^2.
(3) Tests für σ^2 bei bekanntem μ.
(4) Tests für σ^2 bei unbekanntem μ.

Fall 1: Tests für μ bei bekanntem σ^2 (Gauß–z–Test).

Seien $X_1, \dots, X_n \sim \mathrm{N}(\mu, \sigma_0^2)$ unabhängig, wobei die Varianz σ_0^2 bekannt sei. Wir wollen nun verschiedene Hypothesen für μ testen, z. B. $\mu = \mu_0$, $\mu \geq \mu_0$ oder $\mu \leq \mu_0$, wobei μ_0 vorgegeben ist. Wir betrachten die Teststatistik

$$T := \sqrt{n}\, \frac{\bar{X}_n - \mu_0}{\sigma_0}.$$

Unter $\mu = \mu_0$ gilt $T \sim \mathrm{N}(0,1)$. Wir betrachten drei Fälle in Abhängigkeit davon, wie die zu testende Hypothese formuliert wird.

Fall 1A. $H_0 : \mu = \mu_0$; $H_1 : \mu \neq \mu_0$. Die Nullhypothese H_0 sollte verworfen werden, wenn $|T|$ groß ist. Dabei sollte die Wahrscheinlichkeit eines Fehlers 1. Art höchstens α sein. Dies führt zu der Entscheidungsregel, dass die Nullhypothese H_0 verworfen wird, falls $|T| > z_{1-\frac{\alpha}{2}}$.

Fall 1B. $H_0 : \mu \geq \mu_0$; $H_1 : \mu < \mu_0$. Die Nullhypothese H_0 sollte verworfen werden, wenn T klein ist. Dies führt zu der Entscheidungsregel, dass H_0 verworfen wird, falls $T < z_\alpha$.

Fall 1C. $H_0 : \mu \leq \mu_0$; $H_1 : \mu > \mu_0$. Hier sollte H_0 verworfen werden, wenn T groß ist. In diesem Fall wird H_0 verworfen, wenn $T > z_{1-\alpha}$.

Fall 2: Tests für μ bei unbekanntem σ^2 (Student–t–Test).

Seien $X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$, wobei μ und σ^2 unbekannt seien. Wir möchten Hypothesen über μ testen, z. B. $\mu = \mu_0$, $\mu \geq \mu_0$ oder $\mu \leq \mu_0$, wobei μ_0 vorgegeben ist. Die Teststatistik aus Fall 1 können wir dafür nicht verwenden, denn sie enthält den unbekannten Parameter σ^2. Deshalb schätzen wir zuerst σ^2 durch

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Wir betrachten die Teststatistik

$$T := \sqrt{n}\, \frac{\bar{X}_n - \mu_0}{S_n}.$$

Dann gilt unter $\mu = \mu_0$, dass $T \sim t_{n-1}$.

Fall 2A. $H_0 : \mu = \mu_0$; $H_1 : \mu \neq \mu_0$. Die Nullhypothese H_0 sollte verworfen werden, wenn $|T|$ groß ist. Dabei sollte die Wahrscheinlichkeit eines Fehlers 1. Art höchstens α sein. Wegen der Symmetrie der t–Verteilung erhalten wir die folgende Entscheidungsregel: H_0 wird verworfen, falls $|T| > t_{n-1,1-\frac{\alpha}{2}}$.

Fall 2B. $H_0 : \mu \geq \mu_0$; $H_1 : \mu < \mu_0$. Die Nullhypothese H_0 wird verworfen, wenn $T < t_{n-1,\alpha}$.

Fall 2C. $H_0 : \mu \leq \mu_0$; $H_1 : \mu > \mu_0$. Die Nullhypothese H_0 wird verworfen, wenn $T > t_{n-1,1-\alpha}$.

Fall 3: Tests für σ^2 bei bekanntem μ (χ^2–Streuungstest).

Seien $X_1, \ldots, X_n \sim \mathrm{N}(\mu_0, \sigma^2)$ unabhängig, wobei der Erwartungswert μ_0 bekannt sei. Wir wollen verschiedene Hypothesen über die quadratische Streuung σ^2 der Stichprobe testen, wie z. B. $\sigma^2 = \sigma_0^2$, $\sigma^2 \geq \sigma_0^2$ oder $\sigma^2 \leq \sigma_0^2$, wobei σ_0^2 vorgegeben ist. Ein natürlicher Schätzer für σ^2 ist

$$\widetilde{S}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_0)^2.$$

Unter $\sigma^2 = \sigma_0^2$ gilt

$$T := \frac{n\widetilde{S}_n^2}{\sigma_0^2} = \sum_{i=1}^{n} \left(\frac{X_i - \mu_0}{\sigma_0} \right)^2 \sim \chi_n^2.$$

Fall 3A. $H_0 : \sigma^2 = \sigma_0^2$; $H_1 : \sigma^2 \neq \sigma_0^2$. Die Nullhypothese H_0 sollte abgelehnt werden, wenn T zu groß oder zu klein ist. Die χ^2–Verteilung ist nicht symmetrisch. Dies führt zu folgender Entscheidungsregel: H_0 wird verworfen, wenn $T < \chi^2_{n,\frac{\alpha}{2}}$ oder $T > \chi^2_{n,1-\frac{\alpha}{2}}$.

Fall 3B. $H_0 : \sigma^2 \geq \sigma_0^2$; $H_1 : \sigma^2 < \sigma_0^2$. Die Nullhypothese H_0 sollte verworfen werden, wenn T zu klein ist. Dies führt zu folgender Entscheidungsregel: H_0 wird verworfen, wenn $T < \chi^2_{n,\alpha}$ ist.

Fall 3C. $H_0 : \sigma^2 \leq \sigma_0^2$; $H_1 : \sigma^2 > \sigma_0^2$. Die Nullhypothese H_0 sollte verworfen werden, wenn T zu groß ist. Dies führt zu folgender Entscheidungsregel: H_0 wird verworfen, wenn $T > \chi^2_{n,1-\alpha}$ ist.

Fall 4: Tests für σ^2 bei unbekanntem μ (χ^2–Streuungstest).

Seien $X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$, wobei μ und σ^2 unbekannt seien. Wir wollen Hypothesen über σ^2 testen, z. B. $\sigma^2 = \sigma_0^2$, $\sigma^2 \geq \sigma_0^2$ oder $\sigma^2 \leq \sigma_0^2$, wobei σ_0^2 vorgegeben ist. Ein natürlicher Schätzer für σ^2 ist in diesem Fall

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

Unter $\sigma^2 = \sigma_0^2$ gilt

$$T := \frac{(n-1)S_n^2}{\sigma_0^2} \sim \chi^2_{n-1}.$$

Die Entscheidungsregeln sind also die gleichen wie in Fall 3, lediglich muss man die Anzahl der Freiheitsgrade der χ^2–Verteilung durch $n-1$ ersetzen.

10.4. Zweistichprobentests für die Parameter der Normalverteilung

Nun betrachten wir zwei Stichproben $(X_1, \ldots, X_n)$ und $(Y_1, \ldots, Y_m)$. Wir wollen verschiedene Hypothesen über die Lage und die Streuung dieser Stichproben testen. Z. B. kann man sich für die Hypothese interessieren, dass die Erwartungswerte (bzw. Streuungen) der beiden Stichproben gleich sind. Wir machen folgende Annahmen:

(1) $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ sind unabhängige Zufallsvariablen.
(2) $X_1, \ldots, X_n \sim \mathrm{N}(\mu_1, \sigma_1^2)$.
(3) $Y_1, \ldots, Y_m \sim \mathrm{N}(\mu_2, \sigma_2^2)$.

Wir wollen nun Hypothesen über $\mu_1 - \mu_2$ und σ_1^2/σ_2^2 testen. Dabei werden wir uns auf die Nullhypothesen der Form $\mu_1 = \mu_2$ bzw. $\sigma_1^2 = \sigma_2^2$ beschränken. Nullhypothesen der Form $\mu_1 \geq \mu_2$, $\mu_1 \leq \mu_2$, $\sigma_1^2 \geq \sigma_2^2$, $\sigma_1^2 \leq \sigma_2^2$ können analog betrachtet werden.

Fall 1: Test für $\mu_1 = \mu_2$ bei bekannten σ_1^2 und σ_2^2 (Zweistichproben–z–Test).

Es seien also σ_1^2 und σ_2^2 bekannt. Wir können $\mu_1 - \mu_2$ durch $\bar{X}_n - \bar{Y}_m$ schätzen. Unter der Nullhypothese $H_0 : \mu_1 = \mu_2$ gilt, dass

$$T := \frac{\bar{X}_n - \bar{Y}_m}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim \mathrm{N}(0,1).$$

Die Nullhypothese H_0 wird verworfen, wenn $|T|$ groß ist, also wenn $|T| > z_{1-\frac{\alpha}{2}}$.

Fall 2: Test für $\mu_1 = \mu_2$ bei unbekannten aber gleichen σ_1^2 und σ_2^2 (Zweistichproben–t–Test).

Es seien nun σ_1^2 und σ_2^2 unbekannt. Um das Problem zu vereinfachen, werden wir annehmen, dass die Varianzen gleich sind, d.h. $\sigma^2 := \sigma_1^2 = \sigma_2^2$. Wir schätzen σ^2 durch

$$S = \frac{1}{n+m-2}\left(\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + \sum_{j=1}^{m}(Y_j - \bar{Y}_m)^2\right).$$

Wir betrachten die folgende Teststatistik:

$$T := \frac{\bar{X}_n - \bar{Y}_m}{S\sqrt{\frac{1}{n} + \frac{1}{m}}}.$$

Wir haben bei der Konstruktion der Konfidenzintervalle gezeigt, dass $T \sim t_{n+m-2}$ unter $\mu_1 = \mu_2$. Somit wird die Nullhypothese H_0 verworfen, wenn $|T| > t_{n+m-2,1-\frac{\alpha}{2}}$.

Fall 3: Test für $\sigma_1^2 = \sigma_2^2$ bei unbekannten μ_1 und μ_2 (F–Test).

Seien also μ_1 und μ_2 unbekannt. Wir wollen die Nullhypothese $H_0 : \sigma_1^2 = \sigma_2^2$ testen. Natürliche Schätzer für σ_1^2 und σ_2^2 sind gegeben durch

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2, \quad S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \bar{Y}_m)^2.$$

Bei der Konstruktion der Konfidenzintervalle haben wir gezeigt, dass für $\sigma_1^2 = \sigma_2^2$

$$T := \frac{S_X^2}{S_Y^2} \sim F_{n-1,m-1}.$$

Die Hypothese H_0 sollte verworfen werden, wenn T zu klein oder zu groß ist. Dabei ist die F–Verteilung nicht symmetrisch. Die Nullhypothese wird also verworfen, wenn $T < F_{n-1,m-1,\frac{\alpha}{2}}$ oder $T > F_{n-1,m-1,1-\frac{\alpha}{2}}$.

Fall 4: Test für $\sigma_1^2 = \sigma_2^2$ bei bekannten μ_1 und μ_2 (F–Test).

Analog zu Fall 3 (Übung).

10.5. Asymptotische Tests für die Erfolgswahrscheinlichkeit bei Bernoulli–Experimenten

Manchmal ist es nicht möglich oder schwierig, einen exakten Test zum Niveau α zu konstruieren. In diesem Fall kann man versuchen, einen Test zu konstruieren, der zumindest approximativ (bei großem Stichprobenumfang n) das Niveau α erreicht. Wir werden nun die entsprechende Definition einführen. Seien $X_1, X_2, \ldots$ unabhängige und identisch verteilte Zufallsvariablen mit Dichte bzw. Zähldichte h_θ, wobei $\theta \in \Theta$. Es sei außerdem eine Zerlegung des Parameterraumes Θ in zwei disjunkte Teilmengen Θ_0 und Θ_1 gegeben:

$$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset.$$

Wir wollen die Nullhypothese $H_0 : \theta \in \Theta_0$ gegen die Alternativhypothese $H_1 : \theta \in \Theta_1$ testen.

Definition 10.5.1. Eine Folge von Borel–Funktionen $\varphi_1, \varphi_2, \ldots$ mit $\varphi_n : \mathbb{R}^n \to \{0,1\}$ heißt *asymptotischer Test* zum Niveau $\alpha \in (0,1)$, falls

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta_0} \mathbb{P}_\theta[\varphi_n(X_1,\ldots,X_n) = 1] \leq \alpha.$$

Dabei ist φ_n die zum Stichprobenumfang n gehörende Entscheidungsregel.

Wir werden nun asymptotische Tests für die Erfolgswahrscheinlichkeit θ bei Bernoulli–Experimenten konstruieren. Seien $\mathrm{X}_1, \ldots, X_n$ unabhängige und mit Parameter $\theta \in (0,1)$ Bernoulli–verteilte Zufallsvariablen. Wir wollen verschiedene Hypothesen über den Parameter θ testen, z. B. $\theta = \theta_0$, $\theta \geq \theta_0$ oder $\theta \leq \theta_0$. Ein natürlicher Schätzer für θ ist $\bar{X}_n$. Wir betrachten die Teststatistik

$$T_n := \sqrt{n}\,\frac{\bar{X}_n - \theta_0}{\sqrt{\theta_0(1-\theta_0)}}.$$

Unter der Hypothese $\theta = \theta_0$ gilt nach dem Zentralen Grenzwertsatz

$$T_n \xrightarrow[n\to\infty]{d} \mathrm{N}(0,1).$$

Wir betrachten nun drei verschiedene Fälle.

Fall A. $H_0 : \theta = \theta_0$; $H_1 : \theta \neq \theta_0$. In diesem Fall sollte H_0 verworfen werden, wenn $|T_n|$ groß ist. Entscheidungsregel: H_0 wird verworfen, wenn $|T_n| \geq z_{1-\frac{\alpha}{2}}$.

Fall B. $H_0 : \theta \geq \theta_0$; $H_1 : \theta < \theta_0$. Die Nullhypothese H_0 sollte verworfen werden, wenn T_n klein ist. Entscheidungsregel: H_0 wird verworfen, wenn $T_n \leq z_\alpha$.

Fall C. $H_0 : \theta \leq \theta_0$; $H_1 : \theta > \theta_0$. Die Nullhypothese H_0 sollte verworfen werden, wenn T_n groß ist. Entscheidungsregel: H_0 wird verworfen, wenn $T_n \geq z_{1-\alpha}$.

Nun betrachten wir ein Zweistichprobenproblem, bei dem zwei Parameter θ_1 und θ_2 von zwei Bernoulli–verteilten Stichproben verglichen werden sollen. Wir machen folgende Annahmen:

(1) $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ sind unabhängige Zufallsvariablen.
(2) $X_1, \ldots, X_n \sim \mathrm{Bern}(\theta_1)$.
(3) $Y_1, \ldots, Y_m \sim \mathrm{Bern}(\theta_2)$.

Es sollen nun Hypothesen über die Erfolgswahrscheinlichkeiten θ_1 und θ_2 getestet werden, z. B. $\theta_1 = \theta_2$, $\theta_1 \geq \theta_2$ oder $\theta_1 \leq \theta_2$. Ein natürlicher Schätzer für $\theta_1 - \theta_2$ ist $\bar{X}_n - \bar{Y}_m$. Wir definieren uns folgende Größe

$$\tilde{T}_{n,m} = \frac{\bar{X}_n - \bar{Y}_m}{\sqrt{\frac{\theta_1(1-\theta_1)}{n} + \frac{\theta_2(1-\theta_2)}{m}}}.$$

Satz 10.5.2. *Unter $\theta := \theta_1 = \theta_2$ gilt*

$$\tilde{T}_{n,m} \xrightarrow{d} N(0,1) \text{ für } n, m \to \infty.$$

Beweis. Wir haben die Darstellung $\tilde{T}_{n,m} = Z_{1;n,m} + \ldots + Z_{n+m;n,m}$, wobei

$$Z_{k;n,m} = \begin{cases} \frac{X_k - \theta}{n\sqrt{\theta(1-\theta)}\sqrt{\frac{1}{n}+\frac{1}{m}}}, & \text{falls } k = 1, \ldots, n, \\ -\frac{Y_{k-n} - \theta}{m\sqrt{\theta(1-\theta)}\sqrt{\frac{1}{n}+\frac{1}{m}}}, & \text{falls } k = n+1, \ldots, n+m. \end{cases}$$

Wir wollen den Zentralen Grenzwertsatz von Ljapunow verwenden. Es gilt:

(1) Die Zufallsvariablen $Z_{1;n,m}, \ldots, Z_{n+m;n,m}$ sind unabhängig.
(2) $\mathbb{E}Z_{k;n,m} = 0$.
(3) $\sum_{k=1}^{n+m} \mathbb{E}Z_{k;n,m}^2 = 1$.

Die letzte Eigenschaft kann man folgendermaßen beweisen:

$$\sum_{k=1}^{n+m} \mathbb{E}Z_{k;n,m}^2 = \frac{1}{\theta(1-\theta)\left(\frac{1}{n}+\frac{1}{m}\right)}\left(n \cdot \frac{\theta(1-\theta)}{n^2} + m \cdot \frac{\theta(1-\theta)}{m^2}\right) = 1.$$

Wir müssen also nur noch die Ljapunow–Bedingung überprüfen. Sei $\delta > 0$ beliebig. Es gilt

$$\begin{aligned}\sum_{k=1}^{n+m} \mathbb{E}|Z_{k;n,m}|^{2+\delta} &= \frac{1}{\sqrt{\theta(1-\theta)}^{2+\delta}\left(\frac{1}{n}+\frac{1}{m}\right)^{\frac{2+\delta}{2}}}\left\{\frac{n}{n^{2+\delta}}\mathbb{E}|X_1-\theta|^{2+\delta} + \frac{m}{m^{2+\delta}}\mathbb{E}|Y_1-\theta|^{2+\delta}\right\} \\ &\le \frac{C(\theta)}{\left(\frac{1}{n}+\frac{1}{m}\right)^{\frac{2+\delta}{2}}}\left\{\frac{1}{n^{1+\delta}} + \frac{1}{m^{1+\delta}}\right\} \\ &= \frac{C(\theta)}{n^{\frac{\delta}{2}}\left(1+\frac{n}{m}\right)^{\frac{2+\delta}{2}}} + \frac{C(\theta)}{m^{\frac{\delta}{2}}\left(\frac{m}{n}+1\right)^{\frac{2+\delta}{2}}},\end{aligned}$$

was für $n, m \to \infty$ gegen 0 konvergiert. Dabei ist $C(\theta)$ eine von n, m unabhängige Größe. Nach dem Zentralen Grenzwertsatz von Ljapunow folgt die Behauptung des Satzes. □

Die Größe $\tilde{T}_{n,m}$ konvergiert zwar gegen die Standardnormalverteilung, wir können diese Größe allerdings nicht direkt zur Konstruktion von asymptotischen Tests verwenden, denn $\tilde{T}_{n,m}$ beinhaltet die unbekannten Parameter θ_1 und θ_2. Deshalb betrachten wir eine Modifizierung von $\tilde{T}_{n,m}$, in der θ_1 und θ_2 durch die entsprechenden Schätzer $\bar{X}_n$ und $\bar{Y}_m$ ersetzt wurden:

$$T_{n,m} = \frac{\bar{X}_n - \bar{Y}_m}{\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n} + \frac{\bar{Y}_m(1-\bar{Y}_m)}{m}}}.$$

Nach dem Gesetz der großen Zahlen gilt $\bar{X}_n \to \theta_1$ und $\bar{Y}_m \to \theta_2$ fast sicher für $n, m \to \infty$. Aus dem Satz von Slutsky kann man dann herleiten (Übungsaufgabe), dass

$$T_{n,m} \xrightarrow{d} \mathrm{N}(0,1) \text{ für } n, m \to \infty.$$

Wir betrachten nun drei verschiedene Nullhypothesen.

Fall A. $H_0 : \theta_1 = \theta_2$; $H_1 : \theta_1 \neq \theta_2$. In diesem Fall sollte H_0 verworfen werden, wenn $|T_{n,m}|$ groß ist. Entscheidungsregel: H_0 wird verworfen, wenn $|T_{n,m}| \ge z_{1-\frac{\alpha}{2}}$.

Fall B. $H_0 : \theta_1 \ge \theta_2$; $H_1 : \theta_1 < \theta_2$. Die Nullhypothese H_0 sollte verworfen werden, wenn $T_{n,m}$ klein ist. Entscheidungsregel: H_0 wird verworfen, wenn $T_{n,m} \le z_{\alpha}$.

Fall C. $H_0 : \theta_1 \le \theta_2$; $H_1 : \theta_1 > \theta_2$. Die Nullhypothese H_0 sollte verworfen werden, wenn $T_{n,m}$ groß ist. Entscheidungsregel: H_0 wird verworfen, wenn $T_{n,m} \ge z_{1-\alpha}$.