Konstruktion eines einfachen Tests

Wir konstruieren einen Test für den Parameter p einer Bernoulli-verteilten Zufallsvariablen X. Wir setzen

$$H_0 : p \geq p_0, \qquad H_1 : p < p_0.$$

Als Testgröße verwenden wir

$$T := X_1 + \ldots + X_n.$$

Für größere Wahrscheinlichkeiten p erwarten wir auch größere Werte für T. Deshalb ist es sinnvoll, einen Ablehnungsbereich der Art $K := [0, k]$ für T zu wählen, wobei $k \in \mathbb{R}$ geeignet festzulegen ist. Wir konstruieren hier also einen einseitigen Test, während für eine Nullhypothese $H_0 : p = p_0$ sowohl zu kleine als auch zu große Werte von T zur Ablehnung von H_0 führen sollten und somit ein zweiseitiger Test vorzuziehen wäre.

T ist binomialverteilt. Da wir von einem großen Stichprobenumfang n ausgehen, bietet es sich an, die Verteilung von T nach dem Grenzwertsatz von de Moivre (siehe Korollar 117) durch die Normalverteilung zu approximieren.

Sei

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

$\tilde{T}$ ist annähernd standardnormalverteilt.

Wir berechnen für jeden Wert von k das zugehörige Signifikanzniveau α des Tests.

$$
\begin{aligned}
\text{Fehlerwahrscheinlichkeit 1. Art} &= \max_{p \in H_0} \Pr_p[T \in K] \\
&= \max_{p \in H_0} \Pr_p[T \le k] \\
\text{Fehlerwahrscheinlichkeit 2. Art} &= \sup_{p \in H_1} \Pr_p[T \notin K] \\
&= \sup_{p \in H_1} \Pr_p[T > k]
\end{aligned}
$$

Für den Fehler 1. Art α erhalten wir

$$
\begin{aligned}
\alpha &= \max_{p \geq p_0} \Pr_p[T \leq k] = \Pr_{p=p_0}[T \leq k] \\
&= \Pr_{p=p_0}\left[\tilde{T} \leq \frac{k - np}{\sqrt{np(1-p)}}\right] \\
&= \Pr\left[\tilde{T} \leq \frac{k - np_0}{\sqrt{np_0(1-p_0)}}\right] \approx \Phi\left(\frac{k - np_0}{\sqrt{np_0(1-p_0)}}\right).
\end{aligned}
$$

Unter Verwendung der Quantile der Standardnormalverteilung ergibt sich damit:

- Ist k so gewählt, dass $(k - np_0)/\sqrt{np_0(1 - p_0)} = z_\alpha$, so ist das Signifikanzniveau gleich α.
- Ist das gewünschte Signifikanzniveau α des Tests vorgegeben, so erhält man den Wert $k = k(n)$ in Abhängigkeit vom Umfang n der Stichprobe durch

$$k = z_\alpha \cdot \sqrt{np_0(1 - p_0)} + np_0. \tag{8}$$

Kleinere Werte für k verkleinern zwar den Fehler 1. Art, vergrößern jedoch den Annahmebereich und damit die Wahrscheinlichkeit für einen Fehler 2. Art.

Verhalten der Testfehler

Wie verhalten sich die möglichen Testfehler des konstruierten Verfahrens? Was geschieht beispielsweise, wenn p nur geringfügig kleiner als p_0 ist?

In diesem Fall betrachten wir beim Fehler 2. Art die Wahrscheinlichkeit

$$\Pr_{p=p_0-\varepsilon}[T \geq k] \approx \Pr_{p=p_0}[T \geq k] \approx 1 - \alpha \; .$$

Wenn sich also die „wahren“ Verhältnisse nur minimal von unserer Nullhypothese unterscheiden, so werden wir diese „im Zweifelsfall“ annehmen.

Bei echten Alternativtests werden für hinreichend große Stichproben und einen geeignet eingestellten Ablehnungsbereich beide Testfehler klein.

Beispiel 129

Die Abbruchrate p der Transaktionen in einem Online-Datenbanksystem wurde bereits früher einmal ermittelt. Allerdings sind die entsprechenden Daten verloren gegangen und die Entwickler erinnern sich nur noch, dass das Ergebnis entweder $p = 1/3$ oder $p = 1/6$ lautete. Unter dieser Annahme würde man den Test wie folgt ansetzen:

$$H_0 : p \geq 1/3, \qquad H_1' : p \leq 1/6.$$

Beispiel (Forts.)

Für den Fehler 2. Art erhält man nun:

$$\textit{Fehlerwahrsch. 2. Art} = \max_{p \le 1/6} \Pr_p[T > k]$$
$$\approx 1 - \Phi\left(\frac{k - (1/6)\cdot n}{\sqrt{(1/6)\cdot(5/6)n}}\right).$$

Mit den obigen Werten $k = 25$ *und* $n = 100$ *ergibt sich mit*

$$\Phi\left(\frac{150-100}{\sqrt{5}\cdot 10}\right) = \Phi(\sqrt{5}) \approx 0{,}9871$$

ein Fehler 2. Art der Größe $0{,}0129$*, während sich für die triviale Alternative* $H_1 : p < 1/3$ *ein Wert von etwa* $0{,}95$ *ergibt.*