Hypothesentest für Binomialverteilungen

1. Definitionen

Ein Zufallsversuch, bei dem nur zwei Ereignisse unterschieden werden, heißt **Bernoulli-Versuch**. Ein wiederholter Bernoulli-Versuch, bei dem in jedem Durchgang die Wahrscheinlichkeiten beider Ereignisse gleich bleiben (mit Wiederholung bzw. Zurücklegen), und die Reihenfolge, in der die Ereignisse eintreten, keine Rolle spielt, heißt **Bernoulli-Kette**.

➔ Ein gewöhnlicher Spielwürfel hat zwar sechs verschiedene Augenzahlen (Ergebnisse), aber am Anfang eines *Mensch ärgere dich nicht*-Spiels unterscheidet man nur die Ereignisse „6“ und „nicht 6“. Mehrfaches Würfeln, um festzustellen, ob man diese höchste Augenzahl jeweils gewürfelt hat oder nicht, ist eine Bernoulli-Kette.

Die Wahrscheinlichkeitsverteilungen von Bernoulli-Ketten heißen **Binomialverteilungen**. Für die Wahrscheinlichkeit P, mit der ein bestimmtes Ereignis X bei einer Kette von n Bernoulli-Versuchen k-mal eintritt, gilt:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

mit n: Anzahl der Bernoulli-Versuche
k: Anzahl der Treffer des Ereignisses
p: Wahrscheinlichkeit des Ereignisses oder hypothetische Wahrscheinlichkeit beim Hypothesentest

➔ Die Wahrscheinlichkeit, beim zehnmaligen Würfeln genau dreimal eine 6 mit der Ereigniswahrscheinlichkeit 1/6 zu erzielen und dementsprechend siebenmal irgendeine andere Augenzahl mit der Wahrscheinlichkeit 5/6, berechnet sich mit

$$P(\text{"6" } gewürfelt = 3) = \binom{10}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^7 \approx 0{,}155 = 15{,}5\ \%$$

2. Rechnerischer Hypothesentest

2.1. Rechenregeln

Für Binomialverteilungen gilt:

Erwartungswert $\mu = n \cdot p$

Varianz $\sigma^2 = n \cdot p \cdot (1-p)$

Standardabweichung $\sigma = \sqrt{\sigma^2}$

2.2. Durchführung

a) Hypothese H_0 für die vermutete Wahrscheinlichkeit p_0 aufstellen
b) Signifikanzniveau α (auch Irrtumswahrscheinlichkeit oder Fehlertoleranz genannt) festlegen
c) Berechnung des Annahmebereichs:
$[0; \mu + c \cdot \sigma]$ für rechtsseitige Tests $p_0 \leq p$,
$[\mu - c \cdot \sigma; n]$ für linksseitige Tests $p_0 \geq p$ und
$[\mu - c \cdot \sigma; \mu + c \cdot \sigma]$ für zweiseitige Tests $p_0 = p$
d) Hypothese annehmen, wenn der Wert im Annahmebereich liegt, sonst verwerfen

Der Faktor c hängt vom Signifikanzniveau und von der Ein- oder Zweiseitigkeit des Tests ab und wird i.d.R. als Tabellenwert abgelesen:

α	einseitiger Test	zweiseitiger Test
10 %	1,28	1,64
5 %	1,64	1,96
1 %	2,33	2,58
1 ‰	3,09	3,27

3. Beispielaufgabe

a) Die erfahrene QM-Abteilung eines Radioherstellers vermutet, daß 10 % der hergestellten Radiogeräte defekt sind. Diese Vermutung soll als Hypothese mit dem Signifikanzniveau 5 % an einer Stichprobe von 100 Geräten geprüft werden.

b) Da eine Abweichung von der vermuteten Fehlerquote nach unten, also eine kleinere Anzahl defekter Geräte, für das Unternehmen natürlich kein Problem darstellt, wird die Hypothese aufgestellt, daß höchstens 10 % der Radios defekt sind.

a) Hypothese H_0: $\mathbf{p_0 = 0{,}1}$
$\mu = n \cdot p = 100 \cdot 0{,}1 = 10$
$\sigma^2 = n \cdot p \cdot (1 - p) = 100 \cdot 0{,}1 \cdot 0{,}9 = 9$ ➔ $\sigma = \sqrt{\sigma^2} = \sqrt{9} = 3$

α = 5 % für zweiseitigen Test ➔ c = 1,96 (abgelesen)
Annahmebereich: $[\mu - c \cdot \sigma; \mu + c \cdot \sigma] = [10 - 1{,}96 \cdot 3; 10 + 1{,}96 \cdot 3] = [4{,}12; 15{,}88]$

Antwort: Die Hypothese soll angenommen werden, wenn 5 bis 15 Geräte defekt sind.

☞ Gebrochene Werte nicht mathematisch runden, sondern so, daß sie innerhalb des Annahmebereichs liegen, weil sonst eine größere Irrtumswahrscheinlichkeit als beabsichtigt hingenommen wird.

b) Hypothese H_0: $\mathbf{p_0 \leq 0{,}1}$
α = 5 % für einseitigen Test ➔ c = 1,64 (abgelesen)
Annahmebereich, rechtsseitiger Test: $[0; \mu + c \cdot \sigma] = [0; 10 + 1{,}64 \cdot 3] = [0; 14{,}92]$

Antwort: Die Hypothese soll angenommen werden, wenn höchstens 14 Geräte defekt sind.

4. Kritische Beurteilung der Ergebnisse

Angenommen, beim tatsächlich durchgeführten Qualitätstest liegt die Anzahl defekter Radios außerhalb des Annahmebereichs, z. B. bei 18 Stück, dann wird die QM-Abteilung ihre Hypothese verwerfen. Was bedeutet das?

Hiermit ist **nicht** nachgewiesen, daß tatsächlich zweifelsfrei eine Fehlerquote von genau bzw. höchstens 10 % widerlegt worden ist, sondern nur, daß man diese Hypothese (begründete Vermutung) nicht annehmen wird und dabei mit einer Wahrscheinlichkeit von 95 % richtig liegt – aber auch mit 5 % Wahrscheinlichkeit einen Irrtum (Fehler 1. Art) in Kauf nimmt.

Stellen sich hingegen beispielsweise 12 Geräte als defekt heraus, wird die QM-Abteilung ihre Hypothese annehmen, da diese Anzahl (für beide Teilaufgaben) innerhalb des Annahmebereiches liegt. Dennoch ist nicht restlos auszuschließen, daß die tatsächliche Fehlerquote über 10 % liegt und die festgestellte Anzahl von 12 aus 100 durch Zufall oder aufgrund einer Überschneidung der Annahmebereiche auftrat. Die Wahrscheinlichkeit für diesen Irrtum (Fehler 2. Art) ist nicht einmal bestimmbar, weil die wahre Fehlerquote unbekannt ist!