

Chi-Quadrat-Tests

statistik-online.ch

Mit **Chi-Quadrat-Tests** werden Häufigkeiten untersucht. Man vergleicht beobachtete Häufigkeiten mit erwarteten Häufigkeiten.

1 Grundidee

Ein Chi-Quadrat-Test prüft, ob beobachtete Häufigkeiten gut zu erwarteten Häufigkeiten passen.

Typische Fragestellungen sind:

- Treten mehrere Kategorien gleich häufig auf?
- Weicht eine beobachtete Verteilung von einer erwarteten Verteilung ab?
- Hängen zwei kategoriale Variablen miteinander zusammen?
- Sind Verteilungen in mehreren Gruppen gleich?

2 Teststatistik

Die Teststatistik lautet:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Symbol	Bedeutung
O	beobachtete Häufigkeit
E	erwartete Häufigkeit
χ^2	Chi-Quadrat-Teststatistik

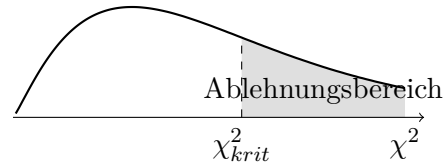
Interpretation: Je grösser χ^2 , desto stärker unterscheiden sich beobachtete und erwartete Häufigkeiten.

3 Entscheidungsregel

Chi-Quadrat-Tests verwenden den rechten Bereich der Chi-Quadrat-Verteilung. Die Entscheidung erfolgt ausschliesslich mit dem kritischen Wert:

$$\chi^2 > \chi_{krit}^2 \Rightarrow H_0 \text{ verwerfen}$$

$$\chi^2 \leq \chi_{krit}^2 \Rightarrow H_0 \text{ nicht verwerfen}$$



Der kritische Wert hängt vom Signifikanzniveau α und von den Freiheitsgraden df ab.

4 Wichtige Chi-Quadrat-Tests

Test	Fragestellung	Typische Daten
Anpassungstest	Passt eine Verteilung zu einer Erwartung?	eine Variable
Unabhängigkeitstest	Hängen zwei Variablen zusammen?	Kreuztabelle

Der Homogenitätstest wird rechnerisch gleich durchgeführt wie der Unabhängigkeitstest mit einer Kreuztabelle. Deshalb wird er hier nicht als eigener Test getrennt behandelt.

5 Chi-Quadrat-Anpassungstest

Der **Anpassungstest** prüft, ob eine beobachtete Verteilung zu einer erwarteten Verteilung passt.

Hypothesen

H_0 : Die beobachtete Verteilung passt zur erwarteten Verteilung.

H_1 : Die beobachtete Verteilung passt nicht zur erwarteten Verteilung.

Freiheitsgrade

Wenn es k Kategorien gibt, gilt:

$$df = k - 1$$

Beispiel

Ein Würfel wird 60-mal geworfen. Bei einem fairen Würfel würde jede Augenzahl etwa 10-mal erwartet. Geprüft wird mit $\alpha = 0.05$.

Augenzahl	1	2	3	4	5	6
beobachtet O	8	12	9	11	10	10
erwartet E	10	10	10	10	10	10

Kategorie	O	E	$\frac{(O-E)^2}{E}$
1	8	10	0.40
2	12	10	0.40
3	9	10	0.10
4	11	10	0.10
5	10	10	0.00
6	10	10	0.00
	χ^2		1.00

$$df = 6 - 1 = 5$$

Aus der Chi-Quadrat-Tabelle:

$$\chi_{krit}^2 = 11.070 \quad (\alpha = 0.05, df = 5)$$

$$\chi^2 = 1.00 \leq 11.070$$

Entscheidung: H_0 wird nicht verworfen. Die beobachteten Häufigkeiten sprechen nicht gegen einen fairen Würfel.

6 Chi-Quadrat-Unabhängigkeitstest

Der **Unabhängigkeitstest** prüft, ob zwei kategoriale Variablen miteinander zusammenhängen.

Hypothesen

H_0 : Die beiden Variablen sind unabhängig.

H_1 : Die beiden Variablen sind abhängig.

Erwartete Häufigkeiten

Bei einer Kreuztabelle wird jede erwartete Häufigkeit so berechnet:

$$E = \frac{\text{Zeilensumme} \cdot \text{Spaltensumme}}{\text{Gesamtsumme}}$$

Freiheitsgrade

$$df = (\text{Anzahl Zeilen} - 1) \cdot (\text{Anzahl Spalten} - 1)$$

Beispiel

Es wird untersucht, ob die Lernmethode mit dem Bestehen einer Prüfung zusammenhängt. Geprüft wird mit $\alpha = 0.05$.

	bestanden	nicht bestanden	Summe
Methode A	30	10	40
Methode B	20	20	40
Summe	50	30	80

Die erwarteten Häufigkeiten sind:

$$E_{A, \text{bestanden}} = \frac{40 \cdot 50}{80} = 25$$

$$E_{A, \text{nicht}} = \frac{40 \cdot 30}{80} = 15$$

Da beide Zeilensummen gleich gross sind, ergeben sich für Methode B ebenfalls 25 und 15.

Zelle	O	E	$\frac{(O-E)^2}{E}$
A, bestanden	30	25	1.00
A, nicht bestanden	10	15	1.67
B, bestanden	20	25	1.00
B, nicht bestanden	20	15	1.67
	χ^2		5.34

$$df = (2 - 1)(2 - 1) = 1$$

Aus der Chi-Quadrat-Tabelle:

$$\chi_{krit}^2 = 3.841 \quad (\alpha = 0.05, df = 1)$$

$$\chi^2 = 5.34 > 3.841$$

Entscheidung: H_0 wird verworfen. Die Daten sprechen für einen Zusammenhang zwischen Lernmethode und Prüfungsergebnis.

7 Voraussetzungen

- Die Daten sind Häufigkeiten, keine Mittelwerte.
- Die Kategorien schliessen sich gegenseitig aus.
- Die Beobachtungen sind unabhängig voneinander.
- Die erwarteten Häufigkeiten sollten nicht zu klein sein.

Als Faustregel sollten erwartete Häufigkeiten meistens mindestens 5 betragen.

8 Zusammenfassung

Schritt	Anpassungstest	Unabhängigkeitstest
1	Hypothesen formulieren	Hypothesen formulieren
2	erwartete Häufigkeiten bestimmen	erwartete Häufigkeiten berechnen
3	$\chi^2 = \sum \frac{(O-E)^2}{E}$	$\chi^2 = \sum \frac{(O-E)^2}{E}$
4	$df = k - 1$	$df = (r - 1)(c - 1)$
5	χ^2 mit χ_{krit}^2 vergleichen	χ^2 mit χ_{krit}^2 vergleichen

Merksatz: Grosse Abweichungen zwischen beobachteten und erwarteten Häufigkeiten führen zu einem grossen χ^2 -Wert.

9 Aufgaben

Aufgabe 1: Anpassungstest

In einer Stichprobe mit 60 Personen wird die Lieblingsfarbe erhoben. Wenn alle drei Farben gleich beliebt sind, würde man pro Farbe 20 Personen erwarten.

Farbe	Rot	Blau	Grün
beobachtet O	18	22	20
erwartet E	20	20	20

Prüfen Sie mit $\alpha = 0.05$, ob die Farben gleich beliebt sind.

Aufgabe 2: Unabhängigkeitstest

Es wird untersucht, ob die Teilnahme an einem Kurs mit dem Bestehen einer Prüfung zusammenhängt.

	bestanden	nicht bestanden	Summe
Kurs besucht	24	16	40
Kurs nicht besucht	12	28	40
Summe	36	44	80

Prüfen Sie mit $\alpha = 0.05$, ob ein Zusammenhang besteht.

Aufgabe 3: Multiple Choice

Kreuzen Sie alle richtigen Aussagen an.

- Beim Chi-Quadrat-Test werden Häufigkeiten verglichen.
- Die Teststatistik kann negativ sein.
- H_0 wird verworfen, wenn $\chi^2 > \chi_{krit}^2$.
- Beim Unabhängigkeitstest werden erwartete Häufigkeiten mit Zeilen- und Spaltensummen berechnet.

10 Lösungen

Lösung 1

Farbe	O	E	$\frac{(O-E)^2}{E}$
Rot	18	20	0.20
Blau	22	20	0.20
Grün	20	20	0.00
	χ^2		0.40

$$df = 3 - 1 = 2$$

$$\chi_{krit}^2 = 5.991 \quad (\alpha = 0.05, df = 2)$$

$$\chi^2 = 0.40 \leq 5.991$$

Entscheidung: H_0 wird nicht verworfen. Die Daten sprechen nicht gegen gleich beliebte Farben.

Lösung 2

Die erwarteten Häufigkeiten:

$$E_{Kurs, bestanden} = \frac{40 \cdot 36}{80} = 18$$

$$E_{Kurs, nicht} = \frac{40 \cdot 44}{80} = 22$$

Für die zweite Zeile ergeben sich ebenfalls 18 und 22.

Zelle	O	E	$\frac{(O-E)^2}{E}$
Kurs, bestanden	24	18	2.00
Kurs, nicht bestanden	16	22	1.64
kein Kurs, bestanden	12	18	2.00
kein Kurs, nicht bestanden	28	22	1.64
	χ^2		7.28

$$df = (2 - 1)(2 - 1) = 1$$

$$\chi_{krit}^2 = 3.841 \quad (\alpha = 0.05, df = 1)$$

$$\chi^2 = 7.28 > 3.841$$

Entscheidung: H_0 wird verworfen. Die Daten sprechen für einen Zusammenhang zwischen Kursteilnahme und Prüfungserfolg.

Lösung 3

- ✓ Richtig – Chi-Quadrat-Tests arbeiten mit Häufigkeiten.
- ✗ Falsch – χ^2 kann nicht negativ sein, weil quadrierte Abweichungen verwendet werden.
- ✓ Richtig – Der Ablehnungsbereich liegt rechts.
- ✓ Richtig – Die erwarteten Häufigkeiten entstehen aus Zeilensumme, Spaltensumme und Gesamtsumme.