

Chi²-Anpassungstest (Bernsteinwürfel.sav)

Voraussetzung: 1 kategoriale Variable (nominal, ordinal), erwartete Häufigkeit in jeder Kategorie ≥ 1 und < 5 nur in 20% der Kategorien.

Der Chi²-Anpassungstest (**Goodness of fit test**) prüft, ob sich eine beobachtete Verteilung einer kategorialen Variablen von einer theoretisch erwarteten Verteilung unterscheidet. Eine häufige Anwendung ist der Test auf Normalverteilung oder auf Gleichverteilung.

Beispiel: Entspricht die Verteilung der Größe von Kindergärten einer Stichprobe jener in der Grundgesamtheit oder wurden z.B. bevorzugt kleine Kitas ausgewählt?

Demo 1:

Die Variable „Augenzahl“ beim Würfel hat 6 Ausprägungen, die bei einem idealen Würfel und häufigen Würfeln etwa gleich oft erscheinen. Bei 90 Würfeln mit einem Bernsteinwürfel ergab sich folgende Verteilung der Variablen „zahl“ in der Datei **Bernsteinwürfel.sav**. Laden Sie die Datei **Bernsteinwürfel.sav** in SPSS (oder PSPP).

Analysieren / Nichtparametrische Tests / Chi-Quadrat

Als „Testvariable“ übernehmen Sie „zahl“ und belassen die Markierung bei „Erwartete Werte“ auf „Alle Kategorien gleich“.

Ergebnis: Bis auf die „1“ gibt es wenig Abweichungen von der erwarteten Anzahl 15

NPAR TESTS

/CHISQUARE=zahl.

zahl			
	Beobachtetes N	Erwartete Anzahl	Residuum
1	7	15,0	-8,0
2	16	15,0	1,0
3	19	15,0	4,0
4	17	15,0	2,0
5	16	15,0	1,0
6	15	15,0	,0
Gesamt	90		

Statistik für Test	
	zahl
Chi-Quadrat	5,733 ^a
df	5
Asymptotische Signifikanz	,333

a. Bei 0 Zellen (,0%) werden weniger als 5 Häufigkeiten erwartet. Die kleinste erwartete Zellenhäufigkeit ist 15,0.

Trotz der geringen Anzahl gewürfelter „1“-en entscheidet der Test für einen idealen Würfel, da alle Kategorien berücksichtigt werden. Die insgesamt Abweichung gegenüber einer Gleichverteilung ist nicht signifikant ($p = 0,333$)

Ein anderes Ergebnis erhält man bei alleiniger Betrachtung der Kategorie „1“ in der Variablen „zahl1“ mit dem dichotomen Merkmal {1, 0}. Nach absteigender Sortierung liefert der Binomialtest folgendes Ergebnis: Der Würfel ist bei dieser Betrachtung nicht ideal, da die „1“ zu wenig erscheint ($p = 0,012$, einseitig). Siehe auch Beitrag "Binomialtest".

Test auf Binomialverteilung

	Kategorie	N	Beobachteter Anteil	Testanteil	Exakte Signifikanz (1-seitig)
zahl1	Gruppe 1	1	7	,077778	,012 ^a
	Gruppe 2	0	83	,922222	
	Gesamt		90	1,000000	

a. Nach der alternativen Hypothese ist der Anteil der Fälle in der ersten Gruppe $<$, 166666.

Demo 2: Größenverteilung von Kindergärten

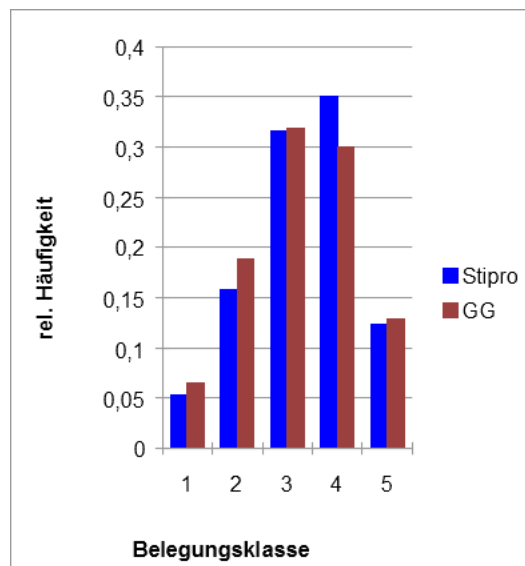
Es wurde eine Zufallsstichprobe von 57 aus 170 Kindergärten (Kigä) gezogen, d.h., eine Clusterstichprobe, die hinsichtlich der Größenverteilung der Kigä repräsentativ sein sollte in Bezug auf die Grundgesamtheit aller 170 Kigä im Landkreis. Es könnte ja sein, dass der mittlere Kariesbefall in kleinen Kigä anders ist, als in großen Kigä. Hierzu erfolgt zunächst eine gleiche Klassierung der Größe (Belegung) in 5 Klassen von Stichprobe und Grundgesamtheit. Dies ist möglich, da man alle Kigä des Landkreises mit ihrer Größe kennt. Nun vergleicht man die beobachteten und erwarteten Häufigkeiten in den Klassen z.B. mit dem **Chi² - Anpassungstest**.

Belegungsklassen Stichprobe

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig <= 20	3	5,3	5,3	5,3
21 - 35	9	15,8	15,8	21,1
36 - 50	18	31,6	31,6	52,6
51 - 64	20	35,1	35,1	87,7
65+	7	12,3	12,3	100,0
Gesamt	57	100,0	100,0	

Belegungsklassen der Grundgesamtheit

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig <= 20	11	6,5	6,5	6,5
21 - 35	32	18,8	18,8	25,3
36 - 50	54	31,8	31,8	57,1
51 - 64	51	30,0	30,0	87,1
65+	22	12,9	12,9	100,0
Gesamt	170	100,0	100,0	



Berechnung mit WinPepi/Describe/C2: Handberechnung von „Expected“

Frequencies:	Observed	Expected
	3	3.7
	9	10.7
	18	18.1
	20	17.1
	7	7.4
Total	57	57

Chi-square goodness-of-fit tests:

Pearson:

chi-sq. (4 d.f.) = 0.917 P = 0.922

Ergebnis:

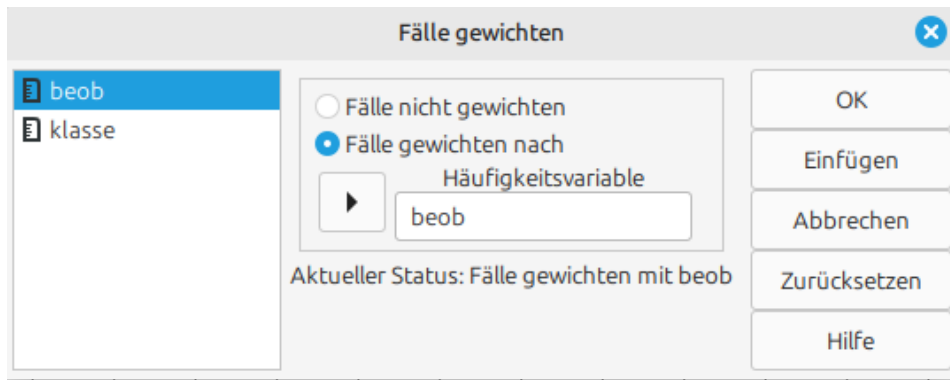
Die Belegungszahlen in den Kindergärten der Stichprobe unterscheiden sich nicht von denen der Grundgesamtheit.

Berechnung mit PSPP (unter LINUX):

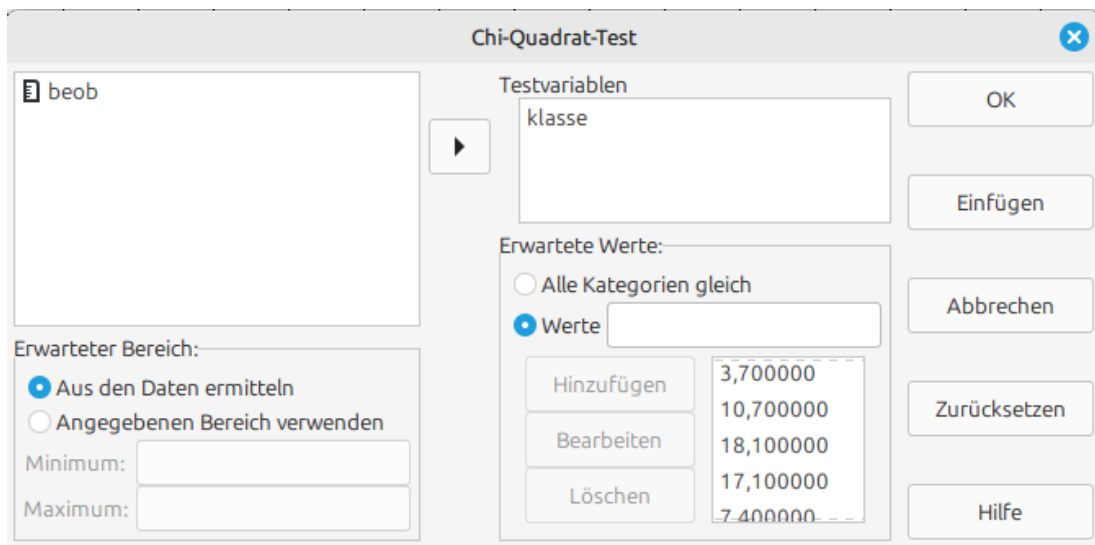
Es werden 2 Variable angelegt, die Variable „klasse“, entspricht den obigen 5 Größenklassen und die Variable „beob“ mit der Häufigkeit der Kindergärten in der jeweiligen Klasse.

Die Testvariable „klasse“ wird nun gewichtet mit der Häufigkeit der Beobachtungen (Variable "beob"):

beob	klasse
3	1
9	2
18	3
20	4
7	5



Mit „Analysieren / Nichtparametrische Tests / Chi-Quadrat“ öffnet sich folgendes Fenster, in das man unter „Erwartete Werte“ die handberechneten Werte eingibt.



Im Ergebnisfenster von PSPP erhält man in Übereinstimmung mit WinPepi:

WEIGHT BY beob.

```
NPAR TEST
/CHISQUARE= klasse
/EXPECTED = 3.7 10.7 18.1 17.1 7.4.
```

klasse

Wert	Beobachtete N	Erwartete N	Residual
1	3	3,70	-,70
2	9	10,70	-1,70
3	18	18,10	-,10
4	20	17,10	2,90
5	7	7,40	-,40
Gesamt	57		

Teststatistiken

	Chi-Quadrat	df	Asymp. Sig.
klasse	,92	4	,922

Die Form der Dialogfenster von PSPP unter Windows variieren.