8 Signifikanztests Teil 2

Zoltán Zomotor

Versionsstand: 12. April 2016, 11:27

Die nummerierten Felder bitte während der Vorlesung ausfüllen.

Bitte hier notieren, was beim Bearbeiten unklar geblieben ist:

Inhaltsverzeichnis

1 Differenztests [Pap11, Teil III, Kapitel 4.5.3.2]

- Gegeben
  - Zwei verbundene einfache Stichproben

    1

  - mit

    2

- Hypothesenpaare:

3

- „Trick“:
  - Übergang zu 4
  - mit 5

⇒ Einstichprobentests mit $\mu_0 = 0$:

6

- Änderung gegenüber Einstichproben-t- und approximativem Gauß-Test:
  Berechnung der Testfunktion v in Schritt 2:

Voraussetzung	anzuwendender Test	Testfunktionswert
Z_i normalverteilt	Einstichproben-t-Test	7
X_i, Y_i dichotom mindestens 5 der $z_i > 0$ mindestens 5 der $z_i < 0$	approximativer Gaußtest	
Z_i beliebig verteilt $n > 30$	approximativer Gaußtest	

- Alternative Berechnugnsmöglichkeit für v, falls X_i, Y_i dichotom:

8

Bedingungen an z_i dann: 9

- Beispiel: Aufgabe 14.12 in [BBK12]:
  500 zufällig ausgewählten Bundesbürgern wurden die beiden Fragen vorgelegt, ob sie
  a) den Bau weiterer Kernkraftwerke befürworten oder ablehnen,
  b) ein Energiesparprogramm für notwendig erachten oder nicht.

  Dabei ergaben sich folgende Daten: 236 Personen befürworten den KKW-Bau, von denen 71 das Sparen befürworten. Insgesamt befürworteten 217 Personen das Sparen.

  Testen Sie zum Signifikanzniveau $\alpha = 0.05$ die Hypothese H_0, dass die Anteile p_1 bzw. p_2 der Personen, die den Bau weiterer Kernkraftwerke ablehnen bzw. ein Energisparprogramm für notwendig ansehen, gleich groß sind gegen H_1: $p_1 > p_2$

  Lösung: Definition der Zufallsvariablen X_i, Y_i mit:

10

11

	$y_i = 0$	$y_i = 1$	$\sum$
$x_i = 0$	12		
$x_i = 1$			
$\sum$			

Hypothesenpaar: 13

Fall: 14

Voraussetzungen 15

1. 16

2. 17

3. 18

4. 19

Interpretation: Zum Signifikanzniveau 5 % kann statistisch bestätigt werden, dass der Anteil der KKW-Gegner größer als der der Sparbefürworter ist.

2 χ^2 -Test für die Varianz [Pap11, Teil III, Kapitel 4.5.4]

- Gegeben: einfache Stichprobe 20
- Hypothesenpaare:

21

- Vorgehensweise:

Schritt 1: Ein Signifikanzniveau 22 festlegen.

Schritt 2: Testfunktionswert v in Abhängigkeit von μ bekannt / unbekannt berechnen:

23

Schritt 3: Verwerfungsbereich festlegen:

24

Fraktilswerte entnehmen aus

25
- -Verteilung bei μ bekannt

26
- -Verteilung bei μ unbekannt

27

Schritt 4: H_0 genau dann ablehnen, wenn gilt

- Beispiel ([Pap11, Seite 592]): Bei der Serienherstellung von Schrauben mit einer bestimmten Länge kann die Zufallsvariable

  X = *Länge einer Schraube*

  als eine normalverteilte Größe betrachtet werden. Aufgrund langjähriger Erfahrungen weiß man, dass die Standardabweichung einen Wert von $\sigma_0 = 1.2$ mm besitzt. Eine zu Kontrollzwecken entnommene Zufallsstichprobe vom Umfang $n = 25$ ergab jedoch eine empirische Standardabweichung von $s = 1.5$ mm. Kann diese Abweichung noch durch *zufällige* Schwankungen erklärt werden, oder ist sie bei einem Signifikanzniveau von 1 % *signifikant*?

  28

  Hypothesenpaar

  29

  Fall:

30

1.

31

2.

32

3.

33

4.

3 Zweistichprobentests [Pap11, Teil III, Kapitel 4.5.3.3]

- Gegeben:
  - Zwei unabhängige einfache Stichproben

    34

    mit
  - Stichprobenumfängen 35
  - Erwartungswerten 36
  - Varianzen 37
  - Stichprobenmittel 38

39

– Stichprobenvarianzen

- Gesucht: Aussagen über den Vergleich der Erwartungswerte
- Hypothesenpaare:

40

- Vorgehensweise:

41

Schritt 1: Ein Signifikanzniveau festlegen.

Schritt 2: Testfunktionswert v gemäß Tabelle 1 auf Seite 8 berechnen

Schritt 3: Verwerfungsbereich festlegen:

42

Fraktilswerte aus der Spalte „Verteilung von V unter $\mu_1 = \mu_2$“ in Tabelle 1 entnehmen.

43

Schritt 4: H_0 genau dann ablehnen, wenn gilt

Tabelle 1: Testfunktionen zum Vergleich zweier Erwartungwerte

	Voraussetzung	Testfunktion V	Verteilung von V unter $\mu_1 = \mu_2$	
1.	$X_i \sim N(\mu_1, \sigma_1)$ $Yi \sim N(\mu_2, \sigma_2)$ σ_1, σ_2 bekannt	44	$N(0, 1)$	Zweistichproben-Gaußtest
2.	$X_i \sim N(\mu_1, \sigma_1)$ $Y_i \sim N(\mu_2, \sigma_2)$ σ_1, σ_2 unbekannt, aber $\sigma_1 = \sigma_2$	$\dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}}$	$t(n_1 + n_2 - 2)$ (für $n_1 + n_2 > 32$ Fraktile aus $N(0, 1)$)	Zweistichproben-t-Test
3.	$X_i \sim B(1, p_1)$ $Y_i \sim B(1, p_2)$ $5 \leq \sum x_i \leq n_1 - 5$ $5 \leq \sum y_i \leq n_2 - 5$	$\dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{(\sum X_i + \sum Y_i) \cdot (n_1 + n_2 - \sum X_i - \sum Y_i)}{(n_1+n_2) \cdot n_1 \cdot n_2}}}$	approximativ $N(0, 1)$	approximativer Zweistichproben-Gaußtest
4.	X_i, Y_i beliebig verteilt $n_1 > 30, n_2 > 30$	45	approximativ $N(0, 1)$	approximativer Zweistichproben-Gaußtest

- Beispiel: Zwei Werkzeugmaschinen sollen anhand Ihrer Ausschussrate verglichen werden. Es werden zwei unabhängige Stichproben genommen: 80 Teile von Maschine 1, von denen 20 Ausschuss sind und 100 Teile von Maschine 2, von denen 50 Ausschuss sind. Lässt sich zum Signifikanzniveau 10 % anhand dieser Stichproben nachweisen, dass Maschine 1 weniger Ausschuss als Maschine 2 liefert?
  - Zufallsvariablen, Verteilung:

    46

    47

    mit
  - Hypothesenpaar: 48
  - Fall: 49
  - Voraussetzung gemäß Tabelle 1:

    50
  - 51

    1.

    52

    2.

    53

    3.

    54

    4.

4 χ^2-Anpassungstest [Pap11, Teil III, Kapitel 5.3]

- Gegeben: Einfache Stichprobe 55
- Hypothesenpaar:

  56

- Grundgedanke:

  Unterteile die 57 in möglichst viele 58 und vergleiche für jedes 59 die tatsächliche mit der theoretischen (aus F_0 errechneten) 60

- Vorgehensweise

Schritt 1: Ein Signifikanzniveau 61 festlegen.

Schritt 2: Den Testfunktionswert v wie folgt ermitteln:

2.1: Die x-Achse in $k \geq 2$ disjunkte, aneinander angrenzende Intervalle

62

unterteilen.

2.2: Für jedes 63 die Anzahl h_j der in A_j liegenden Stichprobenwerte notieren.

2.3: Für jedes 64 die Wahrscheinlichkeit

65 berechnen, das heißt dass ein Beobachtungswert zum Merkmal X in das Intervall A_j fällt, wenn $G \sim F_0$ bezüglich X ist.

2.4: Testfunktionswert berechnen:

66

Schritt 3: Mit dem Fraktilswert 67 ______ der 68 ______ -Verteilung den Verwerfungsbereich 69 ______ festlegen.

Schritt 4: H_0 genau dann ablehnen, wenn 70 ______ gilt.

- Bemerkungen:
  (i) Test nur anwendbar, wenn
  71 ______
  Falls nicht erfüllt ⇒ Intervalle zusammenlegen!
  (ii) Im Allgemeinen sinkt die Wahrscheinlichkeit für 72 ______ , wenn k steigt.
  (iii) Falls diskrete Verteilung:
  - pro Ausprägung ein Intervall, falls (i) erfüllt
  - Schritt 2.1 entfällt

- Beispiel: Gegeben ist folgendes Stichprobenergebnis:

a_j	0	1	2	3	4	5
h_j	1	2	5	6	8	8

Prüfe zum Signifikanzniveau 5 %, ob $G \sim B(20, 0.15)$ gilt. Voraussetzung 73 ______ prüfen wir in Schritt 2.3.

1. 74 ______

2. Testfunktionswert v berechnen:

2.1 75 ______ (sofern Voraussetzung Lücke 71 erfüllt)

2.2 76 ______ (sofern Voraussetzung Lücke 71 erfüllt)

77

2.3

a_j	0	1	2	3	4	5
h_j	1	2	5	6	8	8
p_j	0.0388	0.1368	0.2293	0.2428	0.1821	0.1702
$30p_j$	1.164	4.104	6.879	7.285	5.463	5.106

78

Die Intervalle erfüllen $30p_j \geq 5$ nicht!

$\Rightarrow$ Zusammenlegen!

A_j	$(-\infty, 1]$	$(1, 2]$	$(2, 3]$	$(3, 4]$	$(4, \infty]$
h_j	3	5	6	8	8
p_j	0.1756	0.2293	0.2428	0.1821	0.1702

79

2.4

80

3.

81

4.

Literatur

[BBK12] Günter Bamberg, Franz Baur und Michael Krapp. *Statistik*. 17. Auflage. Oldenbourg Verlag, 2012.

[Pap11] Lothar Papula. *Mathematik für Ingenieure und Naturwissenschaftler*. 6. Auflage. Bd. 3. Vieweg + Teubner, 2011.