Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine-Universität Düsseldorf

20. Januar 2011

1. Der F-Test zum Vergleich zweier Varianzen

2. ANOVA
   - Beispielhafte Fragestellung
   - Bonferroni-Korrektur

Warum vergleicht man zwei Varianzen?

- Um Unterschiedlichkeit zweier Verteilungen nachzuweisen
- Um Voraussetzungen eines anderen Tests zu prüfen
- Um eine ANOVA zu rechnen
- Was ist eine ANOVA?
- “Analysis of Variance”: Ein Test, mit welchem man den Einfluss der Gruppenzugehörigkeit auf einen Parameter prüfen kann, indem man Varianzen vergleicht

F-Test zum Vergleich zweier Varianzen

- $X_1, \ldots, X_{n_1}$ und $Y_1, \ldots, Y_{n_2}$ bezeichnen zwei Gruppen von Messwerten
- Verteilungsvoraussetzungen:
  - Die X_j sind verteilt gemäß $N(\mu_1, \sigma_1^2)$, wobei μ_1 und σ_1 unbekannt sind
  - Die Y_j sind verteilt gemäß $N(\mu_2, \sigma_2^2)$, wobei μ_2 und σ_2 unbekannt sind
- Ziel: σ_1 und σ_2 sollen verglichen werden

F-Test, Fortsetzung

- x_j und y_j seien Realisierungen.
- Bestimme arithmetische Mittelwerte und Stichprobenstreuungen

$$\overline{x} = \frac{1}{n_1} \sum_{j=1}^{n_1} x_j \qquad s_x = \sqrt{\frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (x_j - \overline{x})^2}$$

$$\overline{y} = \frac{1}{n_2} \sum_{j=1}^{n_2} y_j \qquad s_y = \sqrt{\frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (y_j - \overline{y})^2}$$

- Die Teststatistik ist

$$t = \frac{s_x^2}{s_y^2}$$

F-Test, Fortsetzung

- Das Signifikanzniveau sei α
- Wir bestimmen die Quantile der F-Verteilung (genaueres s. u.)

$F_{n_1-1,\, n_2-1,\, 1-\alpha/2}$	beim zweiseitigen Test
$F_{n_1-1,\, n_2-1,\, 1-\alpha}$	bei einem einseitigen Test

- Entscheidung

$H_0 = \{\sigma_1 = \sigma_2\}$: Die Nullhypothese H_0 wird abgelehnt, wenn

$$t > F_{n_1-1,\, n_2-1,\, 1-\alpha/2} \text{ oder } t < \frac{1}{F_{n_2-1,\, n_1-1,\, 1-\alpha/2}}$$

$H_0 = \{\sigma_1 \leq \sigma_2\}$: Die Nullhypothese H_0 wird abgelehnt, wenn

$$t > F_{n_1-1,\, n_2-1,\, 1-\alpha}$$

$H_0 = \{\sigma_1 \geq \sigma_2\}$: Die Nullhypothese H_0 wird abgelehnt, wenn

$$t < \frac{1}{F_{n_2-1,\, n_1-1,\, 1-\alpha}}$$

Die F-Verteilung

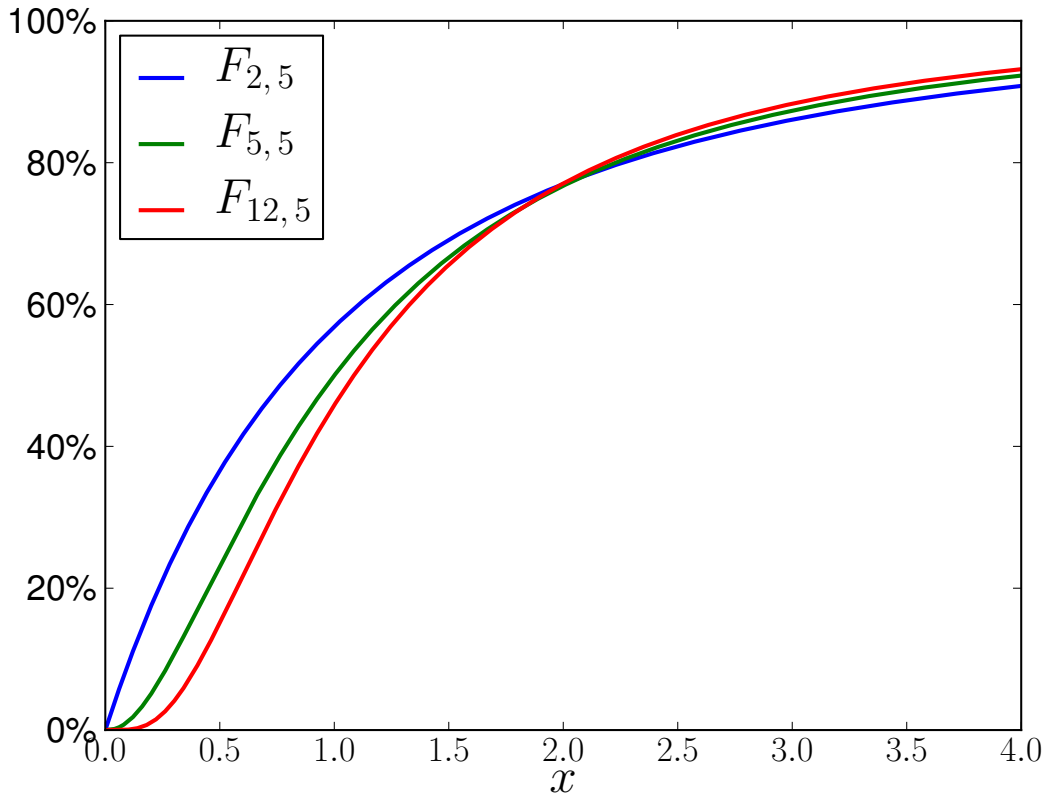
- Die F-Verteilung mit $n_1 - 1$ und $n_2 - 1$ Freiheitsgraden ist definiert als die Verteilung derjenigen Zufallsvariablen Z, die erklärt ist durch

$$Z = \frac{(n_2 - 2) \cdot \sum_{j=1}^{n_1} (X_j - \overline{X})^2}{(n_1 - 1) \cdot \sum_{j=1}^{n_2} (Y_j - \overline{Y})^2}$$

  wenn $X_1, \ldots, X_{n_1}$ und $Y_1, \ldots, Y_{n_2}$ unabhängige, standardnormalverteilte Zufallsvariablen sind.
- Für f_1 bzw. f_2 Freiheitsgrade sind die Quantile der F-Verteilung für α nahe 1 tabelliert.
- Für α nahe 0 benutzt man die Formel

$$F_{f_1, f_2, \alpha} = \frac{1}{F_{f_2, f_1, 1-\alpha}}$$

Verteilungsfunktionen von F-Verteilungen



Quantile $f_{f_1, f_2, 0.95}$ der F-Verteilungen

	f_1					
f_2	1	2	3	4	5	6
1	161.45	199.50	215.71	224.58	230.16	233.99
2	18.51	19.00	19.16	19.25	19.30	19.33
3	10.13	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	**6.26**	6.16
5	6.61	5.79	5.41	**5.19**	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87
8	5.32	4.46	4.07	3.84	3.69	3.58
9	5.12	4.26	3.86	3.63	3.48	3.37
10	4.96	4.10	3.71	3.48	3.33	3.22

Beispiele:

- $f_{4, 5, 0.95} = 5.19$
- $f_{5, 4, 0.95} = 6.26$

F-Test, Beispiel

- Schweine der Rasse B nehmen gleichmäßiger an Gewicht zu als solcher der Rasse A. Zum Signifikanzniveau $\alpha = 0.05$ soll nachgewiesen werden, dass bei der Rasse B das Schlachtgewicht weniger streut als bei solchen der Rasse A.
- Dazu wird bei einer Stichprobe von 7 Schweinen der Rasse A die Stichprobenstreuung $s_X = 15\,kg$, bei 11 Schweinen der Rasse B die Stichprobenstreuung $s_Y = 8\,kg$ festgestellt.
- Ist der Unterschied signifikant?

F-Test, Fortsetzung

- Die Nullhypothese ist $\sigma_1 \leq \sigma_2$
- Die Teststatistik ist

$$t = \frac{s_x^2}{s_y^2} = \frac{15^2}{8^2} = 3.516$$

- Das muss mit dem Quantil $f_{6,10,0.95} = 3.22$ verglichen werden.
- Wegen $t > f_{6,10,0.95}$ kann die Nullhypothese abgelehnt werden.

Beispielhafte Fragestellung: Unterrichtsmethoden

- Jeweils 4 bis 5 Schüler wurden nach einer von 4 Methoden in Statistik unterrichtet. Hat die Wahl der Unterrichtsmethode überhaupt einen Einfluss auf den Lernerfolg?
- Daten: (Der Erfolg wurde auf einer Skala von 0 bis 8 gemessen)

	Unterrichtsmethode			
	1	2	3	4
	2	3	6	5
	1	4	8	5
	3	3	7	5
	1	5	6	3
		0	8	2

Unterrichtsmethoden, Fortsetzung

- Die Nullhypothese ist, dass alle Methoden bis auf zufällige Abweichungen dasselbe Ergbnis liefern
- Wir könnten je zwei Unterrichtsmethoden mit einem t-Test für unverbundene Stichproben testen
- Das sind $\binom{4}{2} = 6$ Paarvergleiche

Unterrichtsmethoden, Fortsetzung

- Für jeden einzelnen Paarvergleich sei α die Fehlerwahrscheinlichkeit erster Art
- Im schlimmsten Falle sind die Paarvergleiche unabhängig
- Mit Wahrscheinlichkeit $p = 1 - \alpha$ vermeidet ein einzelner Paarvergleich den Fehler erster Art
- Mit Wahrscheinlichkeit $(1 - \alpha)^6$ vermeiden ihn alle
- Das Signifikanzniveau ist

$$1 - (1 - \alpha)^6$$

- Für $\alpha = 0.05$ erhält man

$$1 - 0.95^6 = 0.2649$$

Bonferroni-Korrektur

- Die Bonferroni-Korrektur löst dieses Problem auf Kosten des Fehlers zweiter Art
- Die Probanden sind in k-Gruppen aufgeteilt
- Die Gruppen unterscheiden sich in klar festgelegten Parametern
- In allen Gruppen wird ein Messwert erhoben

Bonferroni-Korrektur, Fortsetzung

- Die Nullhypothese ist, dass der Messwert nicht von der Gruppe abhängt
- Zu diesem Zweck werden

$$\binom{k}{2} = \ell$$

  Paarvergleiche vorgenommen
- Wenn ein Paarvergleich signifikante Unterschiede zeigt, wird die Nullhypothese abgelehnt
- Damit der Gesamttest das Signifikanzniveau α aufweist, muss jeder einzelne Paarvergleich zum Signifikanzniveau

$$\beta = \frac{\alpha}{\ell}$$

  durchgeführt werden

Bonferroni-Korrektur: Beispiel

- Um den Einfluss des Tierkreiszeichens auf den Blutdruck zu prüfen, werden die mittleren Blutdrücke von 12 Probandengruppen untersucht
- Jede Gruppe umfasst 20 Mitglieder
- Durch Paarvergleiche soll zum Signifikanzniveau $\alpha = 0.05$ nachgewiesen werden, dass das Tierkreiszeichen einen Einfluss auf den Blutdruck hat
- Die Nullhypothese ist, dass das Tierkreiszeichen keinen Einfluss auf den Blutdruck hat
- Zu welchem Signifikanzniveau müssen die einzelnen Paarvergleiche gerechnet werden?

Beispiel Bonferroni: Fortsetzung

- Anzahl der Paarvergleiche

$$\ell = \binom{12}{2} = \frac{12 \cdot 11}{2} = 66$$

- Signifikanzniveau des einzelnen Vergleichs

$$\beta = \frac{\alpha}{\ell} = \frac{0.05}{66} = 0.00076$$

- Quantil

$$t_{n_1+n_2-2,\,1-\beta/2} = t_{38,\,0.9996} = 3.66$$

- Zum Vergleich

$$t_{n_1+n_2-2,\,1-\alpha/2} = t_{38,\,0.975} = 2.02$$