



# Multivariate Verfahren

MSc Psychologie | MSc Klinische Psychologie und Psychotherapie

WiSe 2023/24

Prof. Dr. Dirk Ostwald

## (5) Einstichproben- $T^2$ -Tests

---

Wie im univariaten Fall unterscheidet man im multivariaten Fall

- Z-Tests
- Einstichproben- $T^2$ -Tests
- Zweistichproben- $T^2$ -Tests bei unabhängigen Stichproben
- Zweistichproben- $T^2$ -Tests bei abhängigen Stichproben

Wir betrachten hier exemplarisch Einstichproben- $T^2$ -Tests.

## Hypothesenszenarien

Einfache Nullhypothese  $H_0 : \mu = \mu_0$ , Einfache Alternativhypothese  $H_1 : \mu = \mu_1$

- Theoretisch wichtiges Szenario (Neymann-Pearson Lemma)
- Praktische Relevanz eher gering

Einfache Nullhypothese  $H_0 : \mu = \mu_0$ , Zusammengesetzte Alternativhypothese  $H_1 : \mu \neq \mu_0$

- Zweiseitiger Einstichproben- $T^2$ -Test mit ungerichteter Hypothese
- Ungerichtete Fragestellung nach einem Unterschied

Zusammengesetzte Nullhypothese  $H_0 : \mu \leq \mu_0$ , Zusammengesetzte Alternativhypothese  $H_1 : \mu > \mu_0$

- Einseitiger Einstichproben- $T^2$ -Test mit gerichteter Hypothese
- Gerichtete Fragestellung nach einem positiven Unterschied

Zusammengesetzte Nullhypothese  $H_0 : \mu \geq \mu_0$ , Zusammengesetzte Alternativhypothese  $H_1 : \mu < \mu_0$

- Gerichtete Fragestellung nach einem negativen Unterschied
- Qualitativ äquivalente Theorie zum umgekehrten Fall

Wir betrachten hier exemplarisch Einstichproben- $T^2$ -Tests

mit einfacher Nullhypothese und zusammengesetzter Alternativhypothese.

---

Anwendungsszenario

Modellformulierung und Modellschätzung

Modellevaluation

Anwendungsbeispiel

Selbstkontrollfragen

## Anwendungsszenario

- Eine Stichprobe/Gruppe  $n$  experimenteller Einheiten mit Datendimension  $m > 1$ .
- Annahme unabhängiger und identisch multivariat normalverteilter Daten.
- Erwartungswertparameter  $\mu$  und Kovarianzmatrixparameter  $\Sigma$  unbekannt.
- Quantifizieren der Unsicherheit beim Inferenzvergleich von  $\mu$  mit  $\mu_0$  beabsichtigt.

## Anwendungsbeispiel

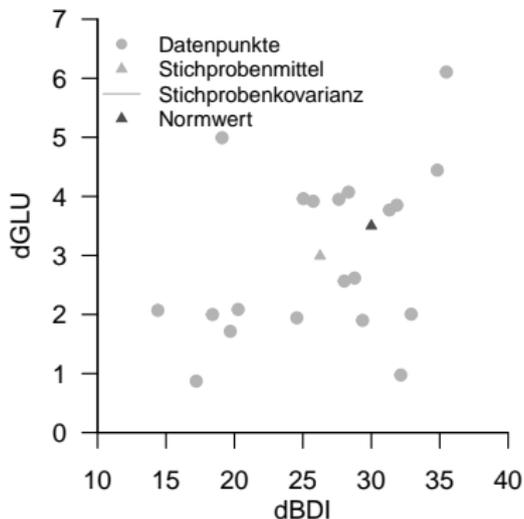
- Gruppenanalyse von BDI und Glukokortikoid Daten
  - $\mu \neq \mu_0$  als Evidenz für eine multivariate Abweichung von einem Normwert  $\mu_0$ .

Abweichung des wahren, aber unbekanntem, Erwartungswertparameters

vom Therapieerfolgsnormwert  $\mu_0 := (30, 3.5)^T$ ?

dBDI	dGLU
35	6.1
25	4.0
20	1.7
29	2.6
29	1.9
17	0.9
33	2.0
28	4.1
26	3.9
31	3.8
14	2.1
18	2.0
19	5.0
28	2.6
20	2.1
35	4.4
28	4.0
32	3.9
32	1.0
25	1.9

Abweichung des wahren, aber unbekanntem, Erwartungswertparameters  
vom Therapieerfolgswert  $\mu_0 := (30, 3.5)^T$ ?



---

Anwendungsszenario

## **Modellformulierung und Modellschätzung**

Modellevaluation

Anwendungsbeispiel

Selbstkontrollfragen

## Definition (Einstichproben- $T^2$ -Test-Modell)

Für  $i = 1, \dots, n$  seien  $v_i$   $m$ -dimensionale Zufallsvektoren, die die  $n$  Datenpunkte eines Einstichproben- $T^2$ -Test Szenarios modellieren. Dann hat das *Einstichproben- $T^2$ -Test-Modell* die strukturelle Form

$$v_i = \mu + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0_m, \Sigma) \text{ u.i.v. für } i = 1, \dots, n \text{ mit } \mu \in \mathbb{R}^m, \Sigma \in \mathbb{R}^{m \times m} \text{ pd} \quad (1)$$

und die Datenverteilungsform

$$v_i \sim N(\mu, \Sigma) \text{ u.i.v. für } i = 1, \dots, n \text{ mit } \mu \in \mathbb{R}^m, \Sigma \in \mathbb{R}^{m \times m} \text{ pd.} \quad (2)$$

### Bemerkungen

- Die Äquivalenz von struktureller Form und Datenverteilungsform des Einstichproben- $T^2$ -Test-Modells folgt direkt aus dem Theorem zu linear-affiner Transformation multivariat normalverteilter Zufallsvektoren.

## MODELLSCHÄTZUNG THEOREM OHNE BEWEIS

---

Anwendungsszenario

Modellformulierung und Modellschätzung

**Modellevaluation**

Anwendungsbeispiel

Selbstkontrollfragen

Zu Gliederung und Wiederholung des univariaten Falls, siehe [\(12\) Hypothesentests](#)

- (1) Teststatistik und Test
- (2) Analyse der Testgütefunktion
- (3) Testumfangkontrolle
- (4) p-Werte
- (5) Analyse der Powerfunktion
- (6) Bestimmung einer optimalen Stichprobengröße

## Definition (Einstichproben- $T^2$ -Teststatistik)

Gegeben seien das Einstichproben- $T^2$ -Test-Modell und ein Nullhypothesenparameter  $\mu_0 \in \mathbb{R}^m$ . Dann ist die Einstichproben- $T^2$ -Teststatistik definiert als

$$T^2 := n(\bar{v} - \mu_0)^T C^{-1}(\bar{v} - \mu_0), \quad (3)$$

wobei  $\bar{v}$  und  $C$  das Stichprobenmittel und die Stichprobenkovarianzmatrix der  $v_1, \dots, v_n$  bezeichnen.

### Bemerkungen

- $T^2$  ist die Stichprobenumfang-skalierte Mahalanobis Distanz von  $\bar{v}$  und  $\mu_0$  hinsichtlich  $C$ .
- $T^2 \uparrow$  für  $\|\bar{v} - \mu_0\| \uparrow$ ,  $C \downarrow$  und  $n \uparrow$ .

## Theorem (Verteilung der skalierten Einstichproben- $T^2$ -Teststatistik)

Es seien  $v_1, \dots, v_n \sim N(\mu, \Sigma)$  mit  $\mu \in \mathbb{R}^m$  und  $\Sigma \in \mathbb{R}^{m \times m}$  pd,

$$\nu := \frac{n - m}{(n - 1)m} \quad (4)$$

und für  $\mu \in \mathbb{R}^m$  sei die Einstichproben- $T^2$ -Teststatistik wie oben definiert. Dann gilt

$$\nu T^2 \sim f(\delta, m, n - m) \quad (5)$$

wobei  $f(\delta, m, n - m)$  die nichtzentrale  $f$ -Verteilung mit Nichtzentralitätsparameter

$$\delta := n(\mu - \mu_0)^T \Sigma^{-1} (\mu - \mu_0) \quad (6)$$

sowie mit Freiheitsgradparametern  $m$  und  $n - m$  bezeichnet.

### Bemerkungen

- Für einen Beweis von (5) verweisen wir auf Hotelling (1931) und Anderson (2003).
- Für  $\mu = \mu_0$  und damit  $\delta = 0$  entspricht  $f(m, n - m, \delta)$  der  $f$ -Verteilung  $f(m, n - m)$
- Für  $m := 1$  ist  $\nu = (n - 1)/(n - 1) \cdot 1 = 1$  und mit der Stichprobenvarianz  $S^2$  gilt

$$T^2 = n \frac{(\bar{v} - \mu_0)^2}{S^2} = \left( \sqrt{n} \frac{\bar{v} - \mu_0}{S} \right)^2 \quad (7)$$

- Das Quadrat der univariaten Einstichproben-T-Teststatistik  $T := \sqrt{n} \frac{\bar{v} - \mu_0}{S}$  ist also  $f(\delta, 1, n - 1)$  verteilt.
- Wir erinnern nachfolgend an die Begriffe der  $f$ -Verteilung und der nichtzentralen  $f$ -Verteilung.

## Definition ( $f$ -Zufallsvariable)

$\xi$  sei eine Zufallsvariable mit Ergebnisraum  $\mathbb{R}_{>0}$  und Wahrscheinlichkeitsdichtefunktion

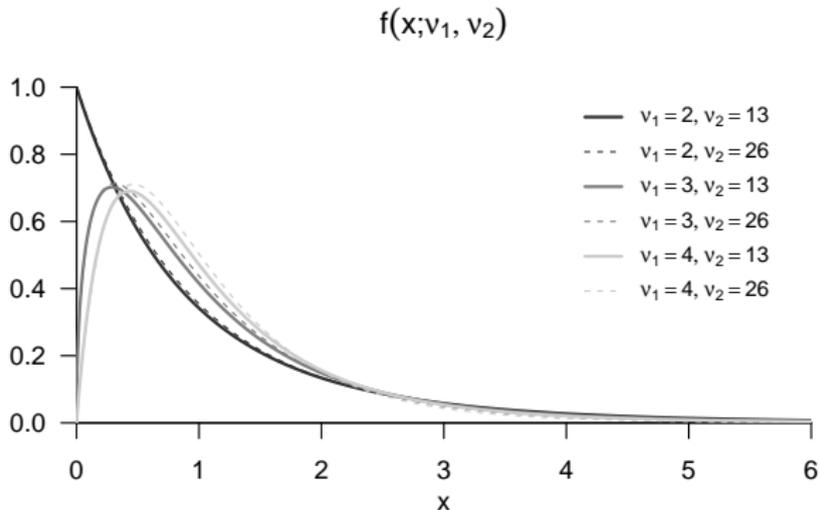
$$p_{\xi} : \mathbb{R} \rightarrow \mathbb{R}_{>0}, x \mapsto p_{\xi}(x) := \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right)} \frac{x^{\frac{\nu_1}{2} - 1}}{(\nu_1 x + \nu_2)^{\frac{\nu_1 + \nu_2}{2}}}, \quad (8)$$

wobei  $\Gamma$  die Gammafunktion bezeichne. Dann sagen wir, dass  $\xi$  einer  $f$ -Verteilung mit Freiheitsgradparametern  $\nu_1$  und  $\nu_2$  unterliegt und nennen  $\xi$  eine  $f$ -Zufallsvariable mit Freiheitsgradparametern  $\nu_1$  und  $\nu_2$ . Wir kürzen dies mit  $\xi \sim f(\nu_1, \nu_2)$  ab. Die Wahrscheinlichkeitsdichtefunktion (WDF) einer  $f$ -Zufallsvariable bezeichnen wir mit  $f(x; \nu_1, \nu_2)$ , die kumulative Verteilungsfunktion (KVF) einer  $f$ -Zufallsvariable bezeichnen wir mit  $F(x; \nu_1, \nu_2)$ , und die inverse kumulative Verteilungsfunktion einer  $f$ -Zufallsvariable bezeichnen wir mit  $F^{-1}(x; \nu_1, \nu_2)$ .

### Bemerkungen

- Im univariaten Fall ist die  $F$ -Statistik der Varianzanalyse bei Zutreffen der Nullhypothese  $f$ -verteilt
- Im multivariaten Fall ist z.B. die  $T^2$ -Statistik bei Zutreffen der Nullhypothese  $f$ -verteilt.

Wahrscheinlichkeitsdichtefunktionen von  $f$ -Verteilungen



## Definition (Nichtzentrale $f$ -Zufallsvariable)

$\xi$  sei eine Zufallsvariable mit Ergebnisraum  $\mathbb{R}_{>0}$  und Wahrscheinlichkeitsdichtefunktion

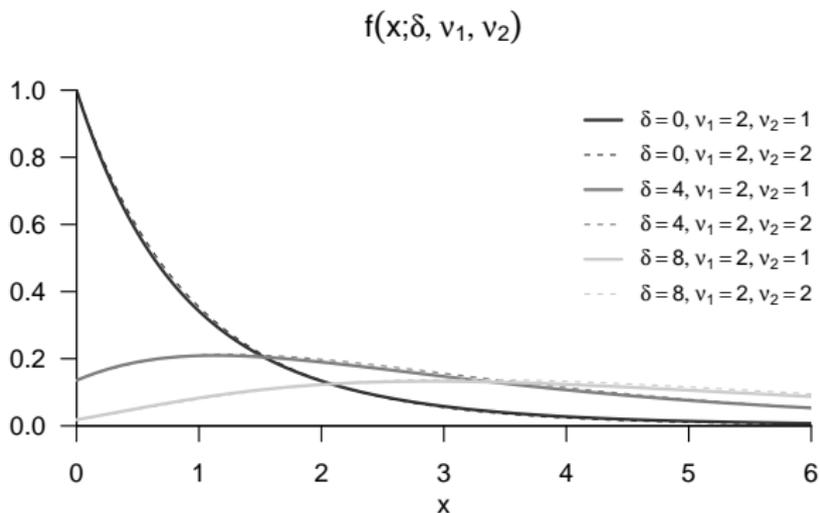
$$p_{\xi} : \mathbb{R} \rightarrow \mathbb{R}_{>0}, x \mapsto p_{\xi}(x) := \sum_{k=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^k}{\frac{\Gamma(\nu_2/2)\Gamma(\nu_1/2+k)}{\Gamma(\nu_2/2+\nu_1/2+k)} k!} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2+k} \left(\frac{\nu_2}{\nu_2 + \nu_1 x}\right)^{(\nu_1+\nu_2)/2+k} x^{\nu_1/2-1+k} \quad (9)$$

wobei  $\Gamma$  die Gammafunktion bezeichne. Dann sagen wir, dass  $\xi$  einer nichtzentralen  $f$ -Verteilung mit Nichtzentralitätsparameter  $\delta$  und Freiheitsgradparametern  $\nu_1$  und  $\nu_2$  unterliegt und nennen  $\xi$  eine nichtzentrale  $f$ -Zufallsvariable mit Nichtzentralitätsparameter  $\delta$  und Freiheitsgradparametern  $\nu_1$  und  $\nu_2$ . Wir kürzen dies mit  $\xi \sim f(\delta, \nu_1, \nu_2)$  ab. Die Wahrscheinlichkeitsdichtefunktion (WDF) einer  $f$ -Zufallsvariable bezeichnen wir mit  $f(x; \delta, \nu_1, \nu_2)$ , die kumulative Verteilungsfunktion (KVF) einer nichtzentralen  $f$ -Zufallsvariable bezeichnen wir mit  $F(x; \delta, \nu_1, \nu_2)$ , und die inverse kumulative Verteilungsfunktion einer nichtzentralen  $f$ -Zufallsvariable bezeichnen wir mit  $F^{-1}(x; \delta, \nu_1, \nu_2)$ .

### Bemerkungen

- Es gilt  $f(0, \nu_1, \nu_2) = f(\nu_1, \nu_2)$ .
- Im univariaten Fall ist die  $F$ -Statistik bei Nichtzutreffen der Nullhypothese nichtzentral  $f$ -verteilt
- Im multivariaten Fall ist z.B. die  $T^2$ -Statistik bei Nichtzutreffen der Nullhypothese nichtzentral  $f$ -verteilt.

## Wahrscheinlichkeitsdichtefunktionen von nichtzentralen $f$ -Verteilungen



## Theorem (WDF und KDF der Einstichproben- $T^2$ -Teststatistik)

Im Einstichproben- $T^2$ -Testszenario sei

$$\nu := \frac{n - m}{(n - 1)m} \quad (10)$$

Dann ist eine WDF der Einstichproben- $T^2$ -Teststatistik gegeben durch

$$p_{T^2} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}, t^2 \mapsto p_{T^2}(t^2) := \nu f(\nu t^2; \delta, m, n - m) \quad (11)$$

und eine KDF der Einstichproben- $T^2$ -Teststatistik ist gegeben durch

$$P_{T^2} : \mathbb{R}_{\geq 0} \rightarrow [0, 1], t^2 \mapsto P_{T^2}(t^2) := F(\nu t^2; \delta, m, n - m) \quad (12)$$

Bemerkungen

- $\nu T^2$  hat die WDF  $f(\delta, m, n - m)$ ,  $T^2$  dagegen hat die WDF  $\nu f(\nu t^2; \delta, m, n - m)$ .

## Beweis

Wir halten zunächst fest, dass das Theorem zur univariate WDF Transformation bei linear-affinen Abbildungen besagt, dass für eine Zufallsvariable  $\xi$  mit WDF  $p_\xi$  und der Definition  $v = f(\xi)$  mit  $f(\xi) := a\xi + b$  für  $a \neq 0$  eine WDF von  $v$  definiert ist durch  $p_v(y) := (1/|a|)p_\xi((y-b)/a)$ . Im vorliegenden Fall ist  $\xi = \nu T^2$  mit WDF  $f(\delta, m, n - m)$  und  $v := T^2 = \frac{1}{\nu}\nu T^2$ , also  $a = 1/\nu$  und  $b = 0$ . Mit  $\nu > 0$  ergibt sich (11) also aus

$$p_{T^2}(t^2) = \frac{1}{a} p_{\nu T^2} \left( \frac{t^2}{a} \right) = \nu f(\nu t^2; m, n - m) \quad (13)$$

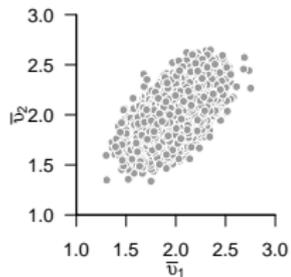
(12) folgt dann mit der Tatsache, dass WDFen bei kontinuierlichen Zufallsvariablen die Ableitungen der entsprechenden KVF sind, sowie der Kettenregel der Differentiation

$$\begin{aligned} \frac{d}{dt^2} P_{T^2}(t^2) &= \frac{d}{dt^2} (F(\nu t^2; m, n - m, \delta)) \\ &= \frac{d}{dt^2} F(\nu t^2; m, n - m, \delta) \frac{d}{dt^2} (\nu t^2) \\ &= \nu f(\nu t^2; m, n - m, \delta) \\ &= p_{T^2}(t^2). \end{aligned} \quad (14)$$

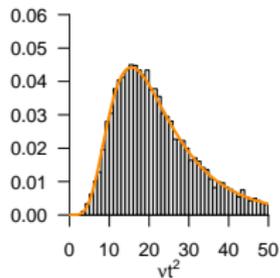
□

# Modellevaluation | (1) Teststatistik und Test

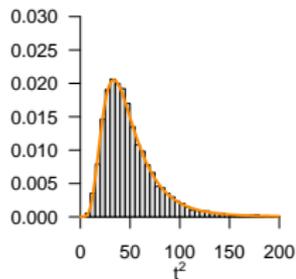
$\bar{v}$  Simulationen



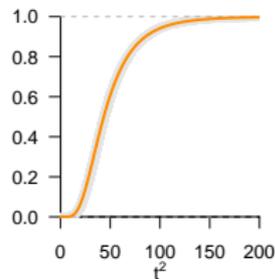
$vT^2 \sim f(\delta, m, n - m)$



$p(t^2) = vf(vt^2; \delta, m, n - m)$



$P(t^2) = F(vt^2; \delta, m, n - m)$



## Definition (Einstichproben- $T^2$ -Test)

Gegeben seien das Einstichproben- $T^2$ -Test-Modell und die Einstichproben- $T^2$ -Teststatistik. Dann ist für  $\Upsilon = (v_1 \quad \dots \quad v_n)$  und einen kritischen Wert  $k \geq 0$  der Einstichproben- $T^2$ -Test definiert als der kritische Wert-basierte Test

$$\phi(\Upsilon) := 1_{\{T^2 > k\}} := \begin{cases} 1 & T^2 > k \\ 0 & T^2 \leq k \end{cases}. \quad (15)$$

### Bemerkung

- Wie üblich repräsentiert  $\phi(\Upsilon) = 1$  das Ablehnen von  $H_0$ .
- Wie üblich repräsentiert  $\phi(\Upsilon) = 0$  das Nichtablehnen von  $H_0$ .

### Definition (Testgütefunktion)

Für einen Test  $\phi$  ist die *Testgütefunktion* definiert als

$$q_\phi : \Theta \rightarrow [0, 1], \theta \mapsto q_\phi(\theta) := \mathbb{P}_\theta(\phi = 1). \quad (16)$$

Für  $\theta \in \Theta_1$  heißt  $q_\phi$  auch *Powerfunktion* oder *Trennschärfefunktion*.

#### Bemerkung

- Der Wert der Testgütefunktion ist die Wahrscheinlichkeit dafür, dass der Test die Nullhypothese ablehnt.

### Theorem (Testgütefunktion)

$\phi$  sei der im obigen Testszenario definiert Test. Dann ist die Testgütefunktion von  $\phi$  gegeben durch

$$q_\phi : \mathbb{R}^m \rightarrow [0, 1], \mu \mapsto q_\phi(\mu) := 1 - F(\nu k; \delta_\mu, m, n - m) \quad (17)$$

wobei  $F(\cdot; \delta_\mu, m, n - m)$  die KVF der nichtzentralen  $f$ -Verteilung mit Freiheitsgradparametern  $m$  und  $n - m$  sowie mit Nichtzentralitätsparameter

$$\delta_\mu := n(\mu - \mu_0)^T \Sigma^{-1} (\mu - \mu_0) \quad (18)$$

bezeichnet.

#### Bemerkungen

- $q_\phi$  kann zur Bestimmung kritischer Werte für einen erwünschten Testumfang genutzt werden.
- $q_\phi$  kann zur Bestimmung der Testpower genutzt werden.

### Beweis

Die Testgütefunktion des betrachteten Tests im vorliegenden Testscenario ist definiert als

$$q_\phi : \mathbb{R}^m \rightarrow [0, 1], \mu \mapsto q_\phi(\mu) := \mathbb{P}_\mu(\phi = 1) \quad (19)$$

Da die Wahrscheinlichkeiten für  $\phi = 1$  und dafür, dass die zugehörige Teststatistik im Ablehnungsbereich des Tests liegt, gleich sind, benötigen wir also zunächst die Verteilung der Teststatistik. Wir haben oben aber bereits gesehen, dass

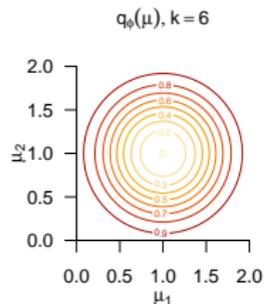
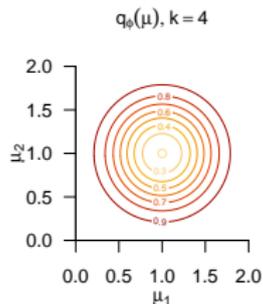
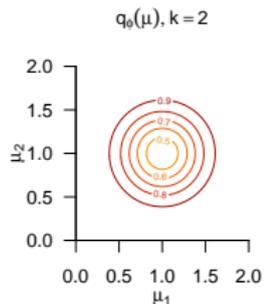
$$\frac{n-m}{m(n-1)} T^2 \sim f(\delta_\mu, m, n-m) \text{ mit } \delta_\mu := n(\mu - \mu_0)^T \Sigma^{-1} (\mu - \mu_0) \quad (20)$$

gilt. Der Ablehnungsbereich des betrachteten Tests ist  $A := ]k, \infty[$ . Also ergibt sich

$$\begin{aligned} q_\phi(\mu) &= \mathbb{P}_\mu(\phi = 1) \\ &= \mathbb{P}_\mu(T^2 \in ]k, \infty[) \\ &= \mathbb{P}_\mu(T^2 > k) \\ &= 1 - \mathbb{P}_\mu(T^2 \leq k) \\ &= 1 - F(\nu k; \delta_\mu, m, m-n). \end{aligned} \quad (21)$$

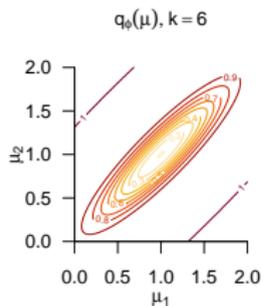
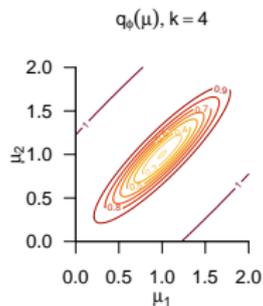
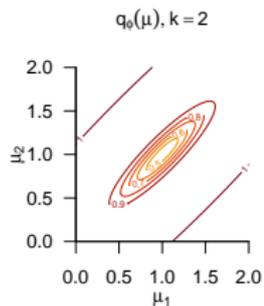
## Beispiele

$$m := 2, n := 15, \Sigma := \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}, \mu_0 := \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}$$



## Beispiele

$$m := 2, n := 15, \Sigma := \begin{pmatrix} 1.0 & 0.9 \\ 0.9 & 1.0 \end{pmatrix}, \mu_0 := \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}$$



### Definition (Level- $\alpha_0$ -Test, Signifikanzlevel $\alpha_0$ , Testumfang $\alpha$ )

$q_\phi$  sei die Testgütefunktion eines Tests  $\phi$  und es sei  $\alpha_0 \in [0, 1]$ . Dann heißt ein Test  $\phi$ , für den gilt, dass

$$q_\phi(\theta) \leq \alpha_0 \text{ für alle } \theta \in \Theta_0 \quad (22)$$

ein *Level- $\alpha_0$ -Test* und man sagt, dass der Test das *Signifikanzlevel*  $\alpha_0$  hat. Die Zahl

$$\alpha := \max_{\theta \in \Theta_0} q_\phi(\theta) \in [0, 1] \quad (23)$$

heißt der *Testumfang* von  $\phi$ .

#### Bemerkungen

- $\alpha$  ist die größtmögliche Wahrscheinlichkeit für einen Typ I Fehler.
- Ein Test ist dann, und nur dann, ein Level- $\alpha_0$ -Test, wenn  $\alpha \leq \alpha_0$  gilt.
- Bei einer einfachen Nullhypothese gilt für den Testumfang, dass  $\alpha = q_\phi(\theta_0) = \mathbb{P}_{\theta_0}(\phi = 1)$ .

### Theorem (Testumfangkontrolle)

$\phi$  sei der im obigen Testszenario definierte Test. Dann ist  $\phi$  ein Level- $\alpha_0$ -Test mit Testumfang  $\alpha_0$ , wenn der kritische Wert definiert ist durch

$$k_{\alpha_0} := \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m) \quad (24)$$

wobei  $\nu := (n - m)/((n - 1)m)$  und  $F^{-1}(\cdot; m, n - m)$  die inverse KVF der  $f$ -Verteilung mit Freiheitsgradparametern  $m$  und  $n - m$  ist.

### Beweis

Damit der betrachtete Test ein Level- $\alpha_0$ -Test ist, muss bekanntlich  $q_\phi(\mu) \leq \alpha_0$  für alle  $\mu \in \{\mu_0\}$ , also hier  $q_\phi(\mu_0) \leq \alpha_0$  gelten. Weiterhin ist der Testumfang des betrachteten Tests durch  $\alpha = \max_{\mu \in \{\mu_0\}} q_\phi(\mu)$ , also hier durch  $\alpha = q_\phi(\mu_0)$  gegeben. Wir müssen also zeigen, dass die Wahl von  $k_{\alpha_0}$  garantiert, dass  $\phi$  ein Level- $\alpha_0$ -Test mit Testumfang  $\alpha_0$  ist. Dazu merken wird zunächst an, dass für  $\mu = \mu_0$  gilt, dass

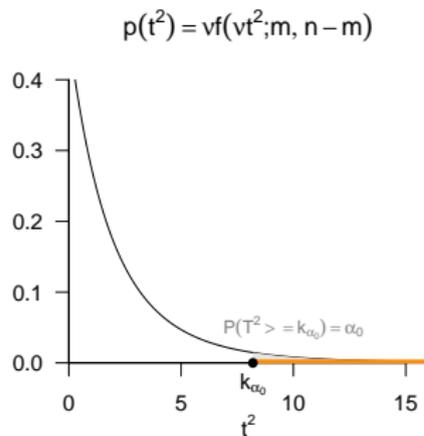
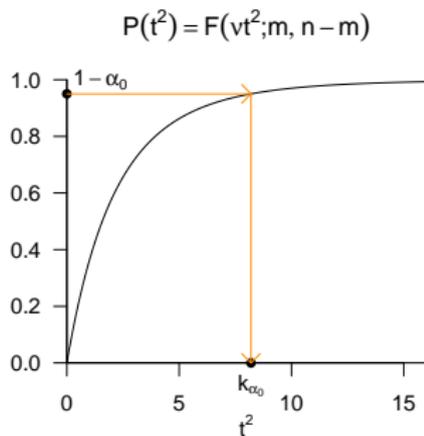
$$q_\phi(\mu_0) = 1 - F(\nu k; \delta_{\mu_0}, m, n - m) = 1 - F(\nu k; 0, m, n - m) = 1 - F(\nu k; m, n - m) \quad (25)$$

wobei  $F(\nu k; \delta, m, n - m)$  und  $F(\nu k; m, n - m)$  die KVF der nichtzentralen  $f$ -Verteilung mit Nichtzentralitätsparameter  $\delta$  und Freiheitsgradparametern  $m$  und  $n - m$  sowie der  $f$ -Verteilung mit Freiheitsgradparametern  $m$  und  $n - m$ , respektive, bezeichnen. Sei nun also  $k := k_{\alpha_0}$ . Dann gilt

$$\begin{aligned} q_\phi(\mu_0) &= 1 - F(\nu k_{\alpha_0}; m, n - m) \\ &= 1 - F(\nu \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m); m, n - m) \\ &= 1 - F(F^{-1}(1 - \alpha_0; m, n - m); m, n - m) \\ &= 1 - (1 - \alpha_0) = \alpha_0. \end{aligned} \quad (26)$$

Es folgt also direkt, dass bei der Wahl von  $k = k_{\alpha_0}$ ,  $q_\phi(\mu_0) \leq \alpha_0$  ist der betrachtete Test somit ein Level- $\alpha_0$ -Test ist. Weiterhin folgt direkt, dass der Testumfang des betrachteten Tests bei der Wahl von  $k = k_{\alpha_0}$  gleich  $\alpha_0$  ist.

Wahl von  $k_{\alpha_0} := \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m)$  mit  $m = 2, n = 15$  und  $\alpha_0 := 0.05$



### Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz  $\Upsilon = (v_1 \ \dots \ v_n)$  eine Realisation von  $v_1, \dots, v_n \sim N(\mu, \Sigma)$  mit unbekanntem Parametern  $\mu \in \mathbb{R}^m$  und  $\Sigma \in \mathbb{R}^{m \times m}$  pd ist.
- Man möchte entscheiden ob für ein  $\mu_0 \in \mathbb{R}^m$  eher  $H_0 : \mu = \mu_0$  oder  $H_1 : \mu \neq \mu_0$  zutrifft.
- Man wählt ein Signifikanzlevel  $\alpha_0$  und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert  $k_{\alpha_0}$ . Zum Beispiel gilt bei Wahl von  $\alpha_0 := 0.05, m = 2$  und  $n = 15$ , also Freiheitsgradparametern 2 und 13, dass  $k_{0.05} = \nu^{-1}F^{-1}(1 - 0.05; 2, 13) \approx 8.2$ .
- Anhand von  $m, n, \mu_0, \bar{v}$  und  $C$  berechnet man die Realisierung der Einstichproben- $T^2$ -Teststatistik

$$T^2 := n(\bar{v} - \mu_0)^T C^{-1}(\bar{v} - \mu_0) \quad (27)$$

- Wenn  $T^2$  größer als  $k_{\alpha_0}$  ist, lehnt man die Nullhypothese ab, andernfalls nicht.
- Die oben entwickelte Theorie garantiert dann, dass man in höchstens  $\alpha_0 \cdot 100$  von 100 Fällen die Nullhypothese fälschlicherweise ablehnt.

### Theorem (p-Wert)

Für den p-Wert den Einstichproben- $T^2$ -Test gilt

$$\text{p-Wert} = \mathbb{P}(T^2 \geq t^2) = 1 - F(\nu t^2; m, n - m). \quad (28)$$

#### Bemerkungen

Wir erinnern daran, dass per Definition der p-Wert das kleinste Signifikanzlevel  $\alpha_0$  ist, bei welchem man die Nullhypothese basierend auf einem vorliegenden Wert der Teststatistik ablehnen würde. Zum Beispiel ergeben sich

- Bei  $m = 2$  und  $n = 15$  der p-Wert für  $t^2 = 7.00$  zu 0.071
- Bei  $m = 2$  und  $n = 15$  der p-Wert für  $t^2 = 9.00$  zu 0.040
- Bei  $m = 2$  und  $n = 99$  der p-Wert für  $t^2 = 7.00$  zu 0.035
- Bei  $m = 4$  und  $n = 15$  der p-Wert für  $t^2 = 7.00$  zu 0.304

## Modellevaluation | (4) p-Wert

### Beweis

Bei einem beobachteten Wert  $t^2$  der Einstichproben- $T^2$ -Teststatistik  $T^2$  würde  $H_0$  für jedes  $\alpha_0$  mit  $t^2 \geq \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m)$  abgelehnt werden. Für diese  $\alpha_0$  gilt, wie unten gezeigt

$$\alpha_0 \geq \mathbb{P}(T^2 \geq t^2) \quad (29)$$

Das kleinste  $\alpha_0 \in [0, 1]$  mit  $\alpha_0 \geq \mathbb{P}(T^2 \geq t^2)$  ist dann  $\alpha_0 = \mathbb{P}(T^2 \geq t^2)$ , also folgt

$$\text{p-Wert} = \mathbb{P}(T^2 \geq t^2) = 1 - F(\nu t^2; m, n - m). \quad (30)$$

Es bleibt zu zeigen, dass gilt

$$\begin{aligned} t^2 &\geq \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m) \\ \Leftrightarrow \nu t^2 &\geq F^{-1}(1 - \alpha_0; m, n - m) \\ \Leftrightarrow \alpha_0 &\geq \mathbb{P}(T^2 \geq t^2). \end{aligned} \quad (31)$$

Dies aber folgt aus

$$\begin{aligned} t^2 &\geq \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m) \\ \nu t^2 &\geq F^{-1}(1 - \alpha_0; m, n - m) \\ F(\nu t^2; m, n - m) &\geq F(F^{-1}(1 - \alpha_0; m, n - m); m, n - m) \\ F(\nu t^2; m, n - m) &\geq 1 - \alpha_0 \\ \mathbb{P}(T^2 \leq t^2) &\geq 1 - \alpha_0 \\ \alpha_0 &\geq 1 - \mathbb{P}(T^2 \leq t^2). \end{aligned} \quad (32)$$

Wir betrachten die Testgütefunktion

$$q_\phi : \mathbb{R}^m \rightarrow [0, 1], \mu \mapsto q_\phi(\mu) := 1 - F(\nu k; \delta_\mu, m, n - m) \quad (33)$$

bei kontrolliertem Testumfang, also für

$$k_{\alpha_0} := \nu^{-1} F^{-1}(1 - \alpha_0; m, n - m) \quad (34)$$

mit festem  $\alpha_0$  als Funktion des Nichtzentralitätsparameters und des Stichprobenumfangs. Namentlich hängt hier  $k_{\alpha_0}$  auch von  $n$  ab.

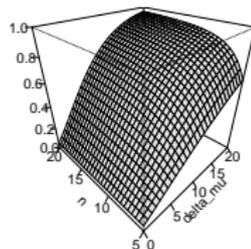
Es ergibt sich die bivariate reellwertige Funktion

$$\pi : \mathbb{R} \times \mathbb{N} \rightarrow [0, 1], (\delta_\mu, n) \mapsto \pi(\delta_\mu, n) := 1 - F(\nu k_{\alpha_0}; \delta_\mu, m, n - m) \quad (35)$$

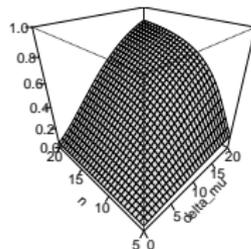
Bei festgelegtem  $\alpha_0$  hängt die Powerfunktion des Einstichproben- $T^2$ -Tests also vom unbekanntem Wert  $\delta_\mu$ , von der Datendimensionalität  $m$  und von der Stichprobengröße  $n$  ab. Wir evaluieren und visualisieren diese Abhängigkeiten untenstehend.

# Modellevaluation | (5) Analyse der Powerfunktion

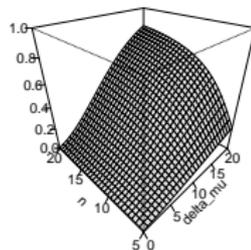
$\alpha_0 = 0.05, m = 2$



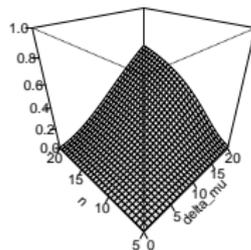
$\alpha_0 = 0.05, m = 4$



$\alpha_0 = 0.01, m = 2$



$\alpha_0 = 0.01, m = 4$



## Praktisches Vorgehen

Mit größerem  $n$  steigt die Powerfunktion des Tests an

- Ein großer Stichprobenumfang ist besser als ein kleiner Stichprobenumfang.
- Kosten für die Erhöhung des Stichprobenumfangs werden aber nicht berücksichtigt.

⇒ Die Theorie statistischer Hypothesentests ist nicht besonders lebensnah.

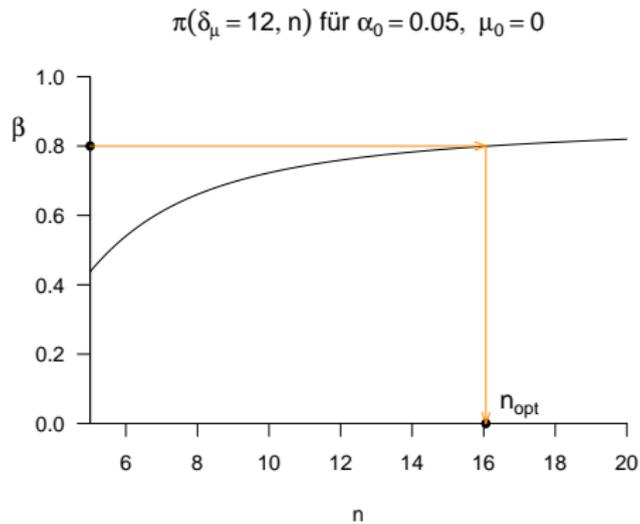
Die Powerfunktion hängt vom wahren, aber unbekanntem, Parameterwert  $\delta_\mu = n(\mu - \mu_0)^T \Sigma^{-1}(\mu - \mu_0)$  ab.

⇒ Wenn man  $\delta_\mu$  schon kennen würde, würde man den Test nicht durchführen.

Generell wird folgendes Vorgehen favorisiert

- Man legt das Signifikanzlevel  $\alpha_0$  fest und evaluiert die Powerfunktion.
- Man wählt einen Mindestparameterwert  $\delta_\mu^*$ , den man mit  $\pi(\delta_\mu, n) = \beta$  detektieren möchte.
- Ein konventioneller Wert ist  $\beta = 0.8$ .
- Man liest die für  $\pi(\delta_\mu = \delta_\mu^*, n) = \beta$  nötige Stichprobengröße  $n$  ab.

## Praktisches Vorgehen



---

Anwendungsszenario

Modellformulierung und Modellschätzung

Modellevaluation

**Anwendungsbeispiel**

Selbstkontrollfragen

# Anwendungsbeispiel

## Praktisches Vorgehen

```
D = read.csv("./5_Daten/5_Einstichproben_T2_Tests.csv") # Datensatzeinlesen
Y = rbind(D$y_1i, D$y_2i) # Datenmatrix
m = nrow(Y) # Dimensionalität der Zufallsvektoren/Daten
n = ncol(Y) # Anzahl der Datenpunkte
nu = (n-m)/(m*(n-1)) # Parameter
mu_0 = matrix(c(30,3.5), nrow = 2) # H0 Hypothesenparameter ("Normwert")
alpha_0 = 0.05 # Signifikanzlevel
k_alpha_0 = (1/nu)*qf(1-alpha_0,m,n-m) # kritischer Wert
j_n = matrix(rep(1,n), nrow = n) # 1_n
I_n = diag(n) # I_n
J_n = matrix(rep(1,n^2), nrow = n) # 1_{nn}
y_bar = (1/n)*(Y %>% j_n) # Stichprobenmittel
C = (1/(n-1))*(Y %>% (I_n-(1/n)*J_n) %>% t(Y)) # Stichprobenkovarianzmatrix
T2 = n*(y_bar - mu_0) %>% solve(C) %>% (y_bar - mu_0) # Einstichproben-T^2-Statistik
if(T2 > k_alpha_0){phi = 1} else {phi = 0} # Test 1_{T^2 >= k_alpha_0}
p = 1 - pf(nu*T2,m,n-m) # p-Wert
cat("Y_bar = ", y_bar,
    "\nC = ", C,
    "\nT^2 = ", T2,
    "\nalpha_0 = ", alpha_0,
    "\nk = ", k_alpha_0,
    "\nphi = ", phi,
    "\np = ", p)
```

```
Y_bar = 26.25615 2.991039
C = 38.8981 3.549813 3.549813 1.972143
T^2 = 7.546368
alpha_0 = 0.05
k = 7.504065
phi = 1
p = 0.04928746
```

## Black-Box-Verfahren

```
library(MVTests)
D      = read.csv("./5_Daten/5_Einstichproben_T2_Tests.csv")
Y      = rbind(D$y_1i, D$y_2i)
mu_0   = matrix(c(30,3.5) , nrow = 2)
alpha_0 = 0.05
phi    = OneSampleHT2(t(Y), mu_0, alpha_0)
cat("Y_bar = " , phi$Descriptive[2,],
    "\nT^2 = " , phi$HT2,
    "\nalpha_0 = " , phi$alpha,
    "\np = " , phi$p.value)

# R Paket
# Datensatzeinlesen
# Datenmatrix
# H0 Hypothesenparameter ("Normwert")
# Signifikanzlevel
# Einstichproben-T^2-Test
# Ausgabe

Y_bar = 26.25615 2.991039
T^2    = 7.546368
alpha_0 = 0.05
p      = 0.04928746
```

---

Anwendungsszenario

Modellformulierung und Modellschätzung

Modellevaluation

Anwendungsbeispiel

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Beschreiben Sie das Anwendungsszenario für einen Einstichproben- $T^2$ -Test.
2. Geben Sie die Definition des Einstichproben- $T^2$ -Test Modells wieder
3. Geben Sie die Definition der Einstichproben- $T^2$ -Teststatistik wieder.
4. Erläutern Sie, wann die Einstichproben- $T^2$ -Teststatistik hohe Werte annimmt.
5. Geben Sie das Theorem zu WDF und KDF der Einstichproben- $T^2$ -Teststatistik wieder.
6. Geben Sie das Theorem zur Testumfangkontrolle eines Einstichproben- $T^2$ -Tests wieder.
7. Erläutern Sie das praktische Vorgehen bei der Durchführung eines Einstichproben- $T^2$ -Tests.
8. Geben Sie das Theorem zum p-Wert eines Einstichproben- $T^2$ -Test an und erläutern Sie die Komponenten des entsprechenden Ausdrucks.

- Anderson, T. W. 2003. *An Introduction to Multivariate Statistical Analysis*. 3rd ed. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley-Interscience.
- Hotelling, Harold. 1931. "The Generalization of Student's Ratio." *The Annals of Mathematical Statistics* 2 (3): 360–78. <https://doi.org/10.1214/aoms/1177732979>.