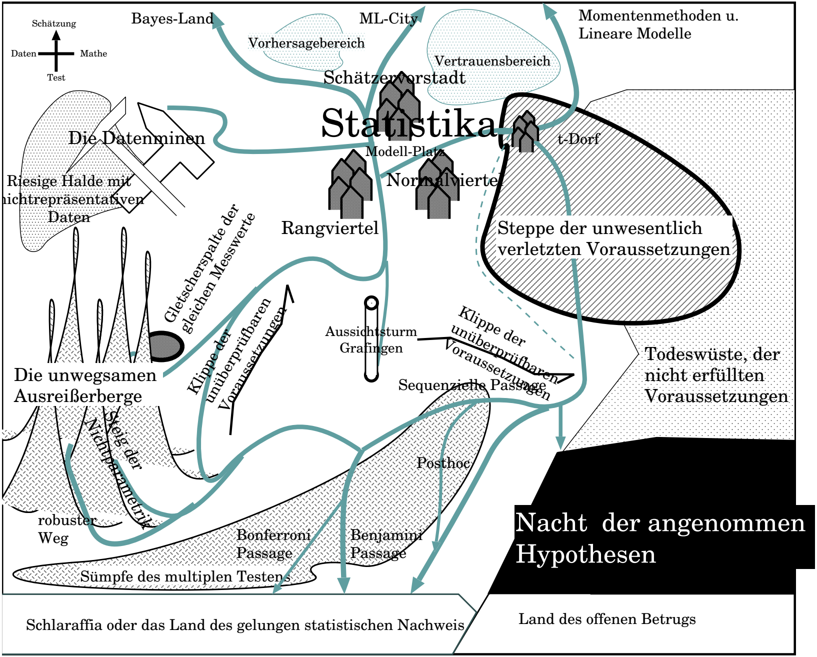
Datenanalyse und Statistik

Vorlesung 6 (Tests II)

K.Gerald van den Boogaart
http://www.stat.boogaart.de

18. November 2019



Schätzung
Daten
Mathe
Test
Bayes-Land
ML-City
Momentenmethoden u.
Lineare Modelle
Vorhersagebereich
Vertrauensbereich
Schätzervorstadt
Statistika
Die Datenminen
Modell-Platz
t-Dorf
Riesige Halde mit
nichtrepräsentativen
Daten
Normalviertel
Rangviertel
Steppe der unwesentlich
verletzten Voraussetzungen
Gletscherspalte der
gleichen Messwerte
Klippe der
unüberprüfbaren
Voraussetzungen
Aussichtsturm
Grafingen
Klippe der
unüberprüfbaren
Voraussetzungen
Todeswüste, der
nicht erfüllten
Voraussetzungen
Die unwegsamen
Ausreißerberge
Sequenzielle Passage
Steig der
Nichtparametrik
Posthoc
robuster
Weg
Bonferroni
Passage
Benjamini
Passage
Nacht der angenommen
Hypothesen
Sümpfe des multiplen Testens
Schlaraffia oder das Land des gelungen statistischen Nachweis
Land des offenen Betrugs

Konstruktion des einfachen einseitigen Gauss-Tests

Je größer der Messwert, deso eher wurde der Grenzwert überschritten:

$$T(X) = \begin{cases} 0, & \text{falls } X < k \\ 1, & \text{falls } X \geq k \end{cases}$$

k heißt kritischer Wert.

Konstruktion des einfachen einseitigen Gauss-Tests

Je größer der Messwert, deso eher wurde der Grenzwert überschritten:

$$T(X) = \begin{cases} 0, & \text{falls } X < k \\ 1, & \text{falls } X \geq k \end{cases}$$

k heißt kritischer Wert.

Wähle k so, dass (wenn die Hypothese stimmt)

$$P(X > k) = \alpha = 0.05$$

Konstruktion des einfachen einseitigen Gauss-Tests

Je größer der Messwert, deso eher wurde der Grenzwert überschritten:

$$T(X) = \begin{cases} 0, & \text{falls } X < k \\ 1, & \text{falls } X \geq k \end{cases}$$

k heißt kritischer Wert.

Wähle k so, dass (wenn die Hypothese stimmt)

$$P(X > k) = \alpha = 0.05$$

Falls X normalverteilt mit Mittelwert $\mu = 7,5$ und Standardabweichung $\sigma = 1$ so ist

$$k = F^{-1}_{N(\mu,\sigma^2)}(1 - \alpha) = 9.14485362695147$$

Konstruktion des einfachen einseitigen Gauss-Tests

Je größer der Messwert, deso eher wurde der Grenzwert überschritten:

$$T(X) = \begin{cases} 0, & \text{falls } X < k \\ 1, & \text{falls } X \geq k \end{cases}$$

k heißt kritischer Wert.

Wähle k so, dass (wenn die Hypothese stimmt)

$$P(X > k) = \alpha = 0.05$$

Falls X normalverteilt mit Mittelwert $\mu = 7,5$ und Standardabweichung $\sigma = 1$ so ist

$$k = F^{-1}_{N(\mu,\sigma^2)}(1-\alpha) = 9.14485362695147$$

F_P ist die Verteilungsfunktion der Verteilung P.
F_P^{-1} ist ihre Umkehrfunktion, die Quantilfunktion.

Ein Test hat

- Namen
  (hier: einfacher einseitiger Gauss-Test)

Ein Test hat

- Namen
- Anwendungssituation
  Überprüfen ob der wahre Erwartungswert einen festen Wert (hier $\mu = 7,5$) übersteigt, wenn eine normalverteilte Messung mit bekannter Varianz σ_0^2 (hier $\sigma_0 = 1$ also $\sigma_0^2 = 1$) vorliegt.

Ein Test hat

- Namen
- Anwendungssituation
- Hypothese und Alternative
  $H_0 : \mu = 7,5$ vs. $H_1 : \mu > 7,5$

Ein Test hat

- Namen
- Anwendungssituation
- Hypothese und Alternative
- Voraussetzungen
  $X \sim N(\mu, \sigma_0^2)$

Ein Test hat

- Namen
- Anwendungssituation
- Hypothese und Alternative
- Voraussetzungen
- Ein Entscheidungsverfahren
  Meist im Computer implementiert. Hier:

$$T(X) = \begin{cases} 0, & \text{falls } X < F^{-1}_{N(\mu_0,\sigma_0^2)}(1-\alpha) \\ 1, & \text{falls } X \geq F^{-1}_{N(\mu_0,\sigma_0^2)}(1-\alpha) \end{cases}$$

bzw.

$$p(X) = 1 - F_{N(\mu_0,\sigma_0^2)}(X)$$

Das Entscheidungsverfahren

Das Test (also das Entscheidungsverfahren)

$$T(X) = \begin{cases} 0, & \text{falls } S(X) < F^{-1}_{P_0^S}(1-\alpha) \\ 1, & \text{falls } S(X) \geq F^{-1}_{P_0^S}(1-\alpha) \end{cases}$$

besteht normalerweise aus:

- Einer Funktion S(X), genannt die Teststatistik (hier $S(X) = X$)

Das Entscheidungsverfahren

Das Test (also das Entscheidungsverfahren)

$$T(X) = \begin{cases} 0, & \text{falls } S(X) < F^{-1}_{P_0^S}(1-\alpha) \\ 1, & \text{falls } S(X) \geq F^{-1}_{P_0^S}(1-\alpha) \end{cases}$$

besteht normalerweise aus:

- Einer Funktion S(X), genannt die Teststatistik
  (hier $S(X) = X$)
- und der Verteilung P_0^S der Teststatistik wenn Nullhypothese und Voraussetzung gelten:
  (hier $P_0^S = N(\mu_0, \sigma_0^2)$)

Das Entscheidungsverfahren

Das Test (also das Entscheidungsverfahren)

$$T(X) = \begin{cases} 0, & \text{falls } S(X) < F^{-1}_{P^S_0}(1-\alpha) \\ 1, & \text{falls } S(X) \geq F^{-1}_{P^S_0}(1-\alpha) \end{cases}$$

besteht normalerweise aus:

- Einer Funktion S(X), genannt die Teststatistik
  (hier $S(X) = X$)
- und der Verteilung P^S_0 der Teststatistik wenn Nullhypothese und Voraussetzung gelten:
  (hier $P^S_0 = N(\mu_0, \sigma^2_0)$)
- Dann ist $p = 1 - F_{P^S_0}(S(X))$

Die Berechnung des p-Wertes

Das Entscheidungsverfahren besteht normalerweise aus:

- ... und der resultierenden Entscheidungsregel

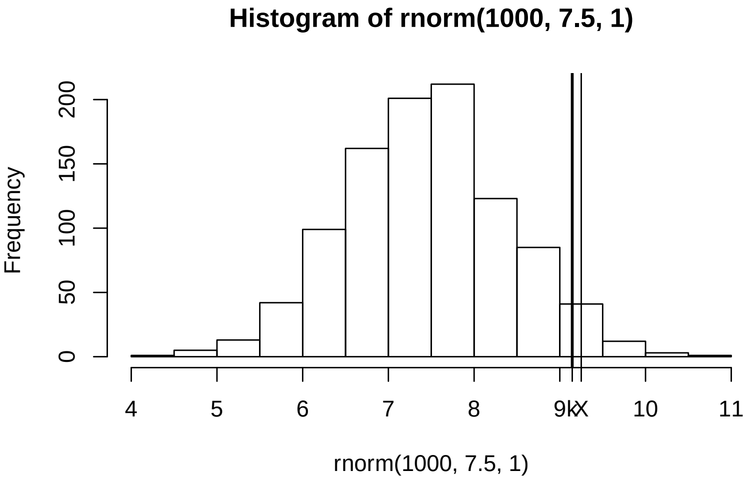
$$T(X) = \begin{cases} 0, & \text{falls } S(X) < F^{-1}_{P_0^S}(1-\alpha) \\ 1, & \text{falls } S(X) \geq F^{-1}_{P_0^S}(1-\alpha) \end{cases}$$

der p-Wert (kleinstes α-Niveau zu dem abgeleht wird) wird dann wie folgt berechnet:

$$p = 1 - F(S(X))$$

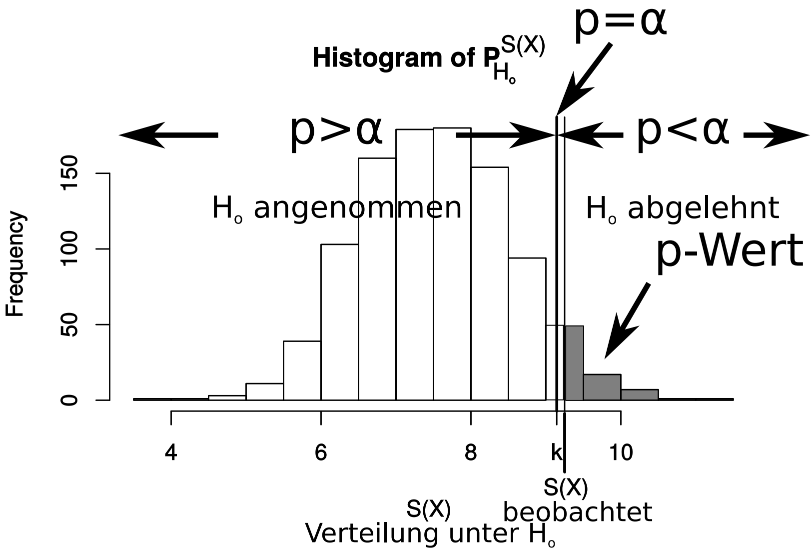
da dann $S(X) = F^{-1}(1-p)$ genau auf der Grenze liegt.

Kritischer Wert und p-Wert



$p =$Fläche oberhalb $S(X)$

Kritischer Wert und p-Wert



$p =$Fläche oberhalb $S(X)$

Der Test ist im Computer implementiert

```
EinfacherGauss.test <- function(x,mean=0,var=1) {
  parameter <-c(mean=mean,sd=sqrt(var))
  statistic <- c(T=x)
  structure(list(
                data.name=deparse(substitute(x)),
                method="Ein Stichproben Gauss-Test",
                alternative="greater",
                parameter=parameter,
                statistic=statistic,
                p.value=1-pnorm(statistic,
                  mean=parameter["mean"],
                  sd=parameter["sd"])
                ),
            class="htest")
}
```

Der Test wird im Computer durchgeführt

```
> EinfacherGauss.test(9.46,mean=7.5,var=1)

        Ein Stichproben Gauss-Test

data:  9.46
T = 9.46, mean = 7.5, sd = 1.0, p-value = 0.025
alternative hypothesis: greater
```

Jeder im Computer implementierte Test gibt irgendwo einen p-Wert aus. Diesen gilt es zu finden.

Es gibt viele Tests.

Wie finde ich den richtigen?

Die Testsituationen

Die Testsituationen werden nach mehrere Kriterien unterteilt:

- **Anzahl der beteiligten Stichproben**
- **Zu testende Größe**
- **Art der Alternative**
- **Art der Voraussetzungen**
- **Anzahl der beteiligten Merkmale**

Beteiligte Stichproben

- **Einzelbeobachtung**
  z.B. ein einzelner Messwert, wie im Beispiel

Beteiligte Stichproben

- **Einzelbeobachtung**
- **Ein-Stichproben-Tests**
  Es werden Eigenschaften einer Grundgesamtheit untersucht, z.B. wenn das Labor mehrere Messungen gemacht hat.

Beteiligte Stichproben

- **Einzelbeobachtung**
- **Ein-Stichproben-Tests**
- **Zwei-Stichproben-Tests**
  wenn zwei Grundgesamtheiten verglichen werden sollen (z.B. behandelte und unbehandelte Flächen).

Beteiligte Stichproben

- **Einzelbeobachtung**
- **Ein-Stichproben-Tests**
- **Zwei-Stichproben-Tests**
- **Gepaarte Tests**
  wenn zwei Messungen am gleichen statistischen Individuum verglichen werden sollen (z.B. vorher – nachher Vergleiche). Eigentlich: Ein-Stichprobentests mit zwei Merkmalen.

Beteiligte Stichproben

- **Einzelbeobachtung**
- **Ein-Stichproben-Tests**
- **Zwei-Stichproben-Tests**
- **Gepaarte Tests**
- **Mehr-Stichproben-Tests**
  Überprüfen, ob mehrere Grundgesamtheiten gleich sind (z.B. nachweisen, dass sich verschiedene ethnische Gruppen in ihrem Sozialverhalten unterscheiden.)

Zu testende Größe

- **Mittelwert/Lage**
  z.B. verdienen Männer und Frauen im Schnitt gleich viel, wenn sie als Geologen/Maschinenbauer arbeiten?

Zu testende Größe

- **Mittelwert/Lage**
- **Varianz/Streuung**
  z.B. Welches von zwei Verfahren mißt genauer?

Zu testende Größe

- **Mittelwert/Lage**
- **Varianz/Streuung**
- **Verteilung**
  z.B. sind die Daten wirklich normalverteilt?

Zu testende Größe

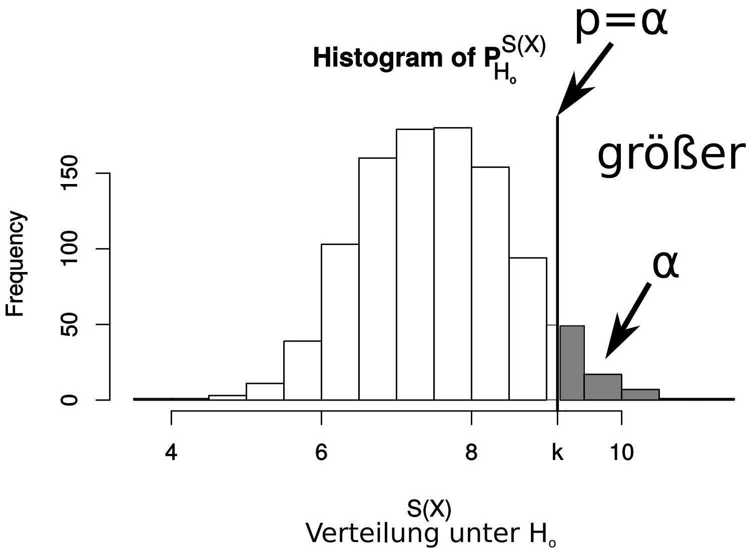
- **Mittelwert/Lage**
- **Varianz/Streuung**
- **Verteilung**
- **Unabhängigkeit**
  z.B. bekommen Raucher öfter Krebs? d.h. ist Krebs vom Rauchen abhängig.

Art der Alternative

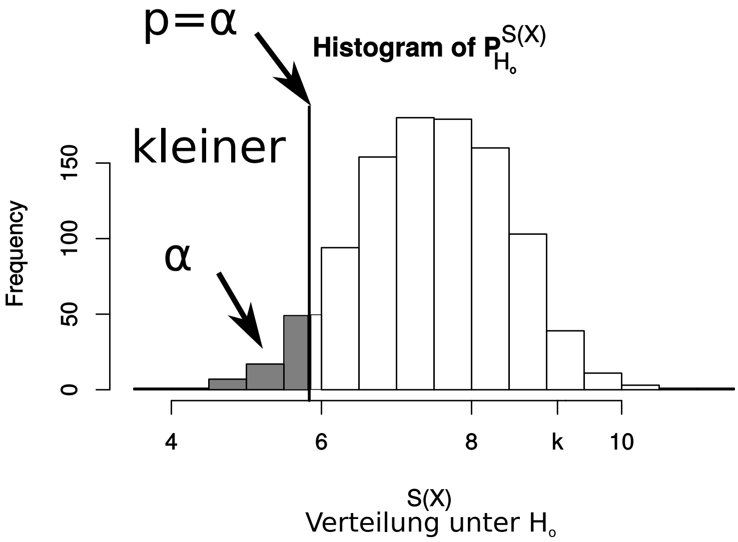
- **“Größer”**: $H_0 : \mu \leq 7,5$ vs. $H_1 : \mu > 7,5$
- **“Kleiner”**: $H_0 : \mu \geq 7,5$ vs. $H_1 : \mu < 7,5$
- **“Ungleich”**: $H_0 : \mu = 7,5$ vs. $H_1 : \mu \neq 7,5$

Die Tests auf größer und kleiner heißen auch **einseitige Tests**, da von der Hypothese aus die Alternative nur in einer Richtung liegt. Tests bei denen die Alternative in beiden Richtungen von der Hypothese liegt heißen auch **zweiseitige Tests**.

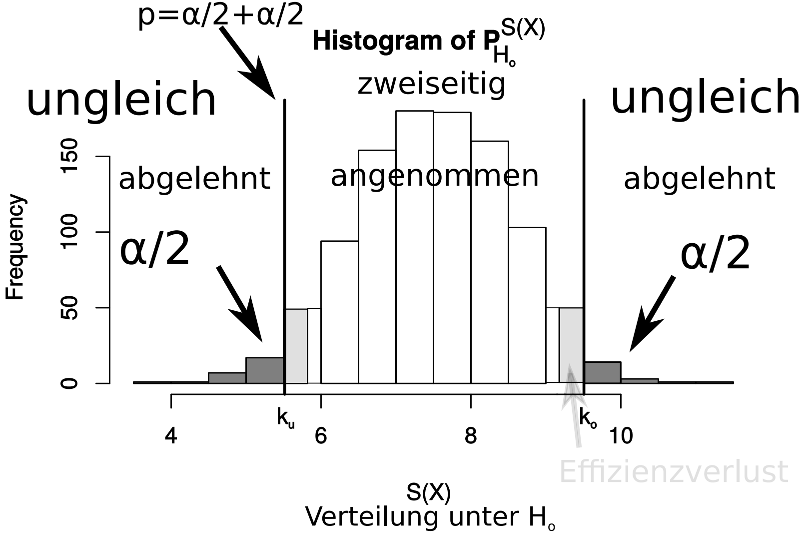
Test auf größer



Test auf kleiner



Test auf ungleich



Art der Voraussetzung

- Generalvoraussetzung: Unabhängigkeit / repräsentative Stichprobe(n)

Art der Voraussetzung

- Generalvoraussetzung: Unabhängigkeit / repräsentative Stichprobe(n)
- **Normalverteilungsbasierte Tests**
  Voraussetzung: Die Zufallsvariablen sind normalverteilt.
  Problem: Falsche Ergebnisse, wenn Ausreißer oder bimodale Verteilungen vorliegen.
  Betrachtet: Mittelwerte, Varianzen

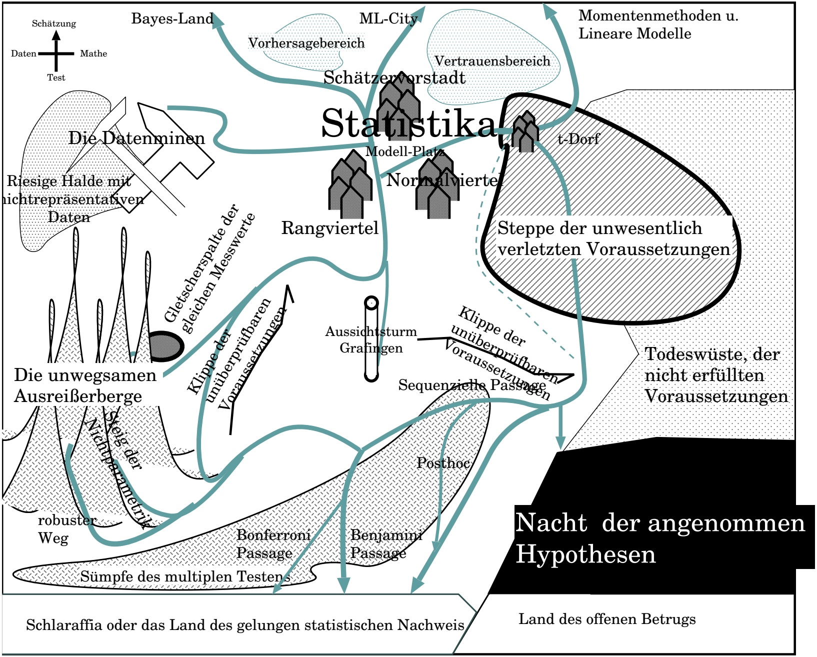
Art der Voraussetzung

- Generalvoraussetzung: Unabhängigkeit / repräsentative Stichprobe(n)
- **Normalverteilungsbasierte Tests**
- **Nichtparametrische Test/Rangtests**
  Voraussetzung: Die Zufallsvariablen sind stetig verteilt.
  Problem: Falsche Ergebnisse, wenn zu viele Messwerte gleich sind.
  Betrachtet: Ränge, relative Verschiebung, $<$.
  Info: Weniger effizient als normalerverteilungsbasierte Tests

Art der Voraussetzung

- Generalvoraussetzung: Unabhängigkeit / repräsentative Stichprobe(n)
- **Normalverteilungsbasierte Tests**
- **Nichtparametrische Test/Rangtests**
- **Robuste Tests**

  Voraussetzung: Die Zufallsvariablen sind Normalverteilung, aber es dürfen falsche Werte/Ausreißer vorhanden sein
  Problem: Verfügbarkeit, Maximaler Anteil der falschen Werte muß angegeben werden.
  Betrachtet: Mittelwert, Varianz des "Hauptanteils der Daten"
  Info: Liegen von der Effizienz her zwischen den beiden anderen.


Schätzung
Daten
Mathe
Test
Bayes-Land
ML-City
Momentenmethoden u.
Lineare Modelle
Vorhersagebereich
Vertrauensbereich
Schätzervorstadt
Statistika
Die Datenminen
t-Dorf
Modell-Platz
Normalviertel
Riesige Halde mit
ichtrepräsentativen
Daten
Rangviertel
Steppe der unwesentlich
verletzten Voraussetzungen
Gletscherspalte der
gleichen Messwerte
Klippe der
unüberprüfbaren
Voraussetzungen
Aussichtsturm
Grafingen
Klippe der
unüberprüfbaren
Voraussetzungen
Todeswüste, der
nicht erfüllten
Voraussetzungen
Sequenzielle Passage
Die unwegsamen
Ausreißerberge
Steig der
Nichtparametrik
Posthoc
robuster
Weg
Bonferroni
Passage
Benjamini
Passage
Nacht der angenommen
Hypothesen
Sümpfe des multiplen Testens
Schlaraffia oder das Land des gelungen statistischen Nachweis
Land des offenen Betrugs

Voraussetzungen an die Varianz

Bei Zwei– und Mehrstichproben-Problemen mit Normalverteilungsvoraussetzung ist oft noch die Unterscheidung nach der Gleichheit der Varianz wichtig.

- **homoskedastisch**: Streuung in allen Teilgrundgesamtheiten gleich.
- **heteroskedastisch**: Streuungen nicht unbedingt gleich.

Anzahl der beteiligten Merkmale

- **univariat**: Es wird nur ein Merkmal betrachtet.
- **bivariat**: Es werden zwei Merkmale betrachet.
- **multivariat**: Es werden mehrere Merkmale betrachtet.

Die Testsituationen

Die Testsituationen werden nach mehrere Kriterien unterteilt:

- **Anzahl der beteiligten Stichproben**
  Ein-, Zwei-, Mehrstichprobensituation (+ gepaarte Tests)
- **Zu testende Größe**
  Lage, Streuung, Verteilung, Abhängigkeit
- **Art der Alternative**
  Größer, Kleiner, Ungleich / Einseitig, Zweiseitig
- **Art der Voraussetzungen**
  Generalvoraussetzung: representativ
  Normalverteilung, Stetige Verteilung, oder Diskrete Verteilung
  Homo-, oder Heteroskedastisch
- **Anzahl der beteiligten Merkmale**
  Univariat, Bivariat, Multivariat

Ein-Stichproben-Tests

... auf Verteilung

irgendeine Normalverteilung $\Rightarrow$ Shapiro-Wilk-Test
eine bestimmte stetige Verteilung $\Rightarrow$ (Ein-Stichproben)-KS-Test
diskrete Größe $\Rightarrow$ χ^2-test auf Verteilung

Shapiro-Wilk-Test

- **Shapiro-Wilk-Test**

  Situation:

  Test auf Normalverteilung

  H_0 : X normalverteilt ($X \sim N(\mu, \sigma^2)$)
  H_1 : X nicht normalverteilt

  Voraussetzungen:

  repräsentative Stichprobe
  nicht zu stark gerundet

  Bemerkung:

  d.h. es darf keine "falschen Bindungen" geben

  Befehl: `shapiro.test(X)`

Beispiel

```
        Shapiro-Wilk normality test

data:  x
W = 0.79384, p-value = 0.0007002
```

Kolmogorov-Smirnov-Test

- **Kolmogorov-Smirnov-Test**

  Situation:

  Test auf spezielle Verteilung (stetig)

  H_0 : X hat die Verteilungsfunktion F_0: $F_X \equiv F_0$
  H_1 : Die Verteilungsfunktion F_X ist ungleich F_0

  Voraussetzungen:

  repräsentative Stichprobe
  nicht zu stark gerundet

  Bemerkung:

  d.h. es darf keine "falschen Bindungen" geben
  Der Test nimmt zu oft an, wenn die Parameter aus den Daten geschätzt wurden.
  Der Test lehnt zu oft ab, wenn die Parameter aus anderen Daten geschätzt wurden.
  Eine kleinere Verteilungsfunktion gehört zu größeren Werten.

  Befehl: `ks.test(X,`F_0`)`

Kolmogorov-Smirnov-Test (einseitiger)

- **Kolmogorov-Smirnov-Test (einseitiger)**

Situation:

Test auf spezielle Verteilung (stetig)

H_0 : X eine kleinere Verteilungsfunktion als F_0
H_1 : Die Verteilungsfunktion F_X ist ungleich F_0

Voraussetzungen:

repräsentative Stichprobe
nicht zu stark gerundet

Bemerkung:

d.h. es darf keine "falschen Bindungen" geben
Der Test nimmt zu oft an, wenn die Parameter aus den Daten geschätzt wurden.
Der Test lehnt zu oft ab, wenn die Parameter aus anderen Daten geschätzt wurden.
Eine kleinere Verteilungsfunktion gehört zu größeren Werten.

Befehl:

χ^2-test auf Verteilung einer diskreten Größe

- **χ^2-test auf Verteilung einer diskreten Größe**

  Situation:

  Test der Verteilung einer diskreten Größe

  H_0 : X ist diskret Verteilt
  H_1 : ...

  Voraussetzungen:

  repräsentative Daten

  Bemerkung:

  Das ist ein asymtotischer Test: Es müssen mindestens 5 Punkte pro Kategorie erwartet werden

  Befehl:

- **Binomial-test auf Verteilung einer dichtomen Größe**

Situation:

H_0 : X hat eine dichtome Scale
H_1 : ...
Voraussetzungen:

repräsentative Daten

Bemerkung:

Keine Probleme

Befehl:

Ein-Stichproben-Tests

... auf Lage

normalverteilt, Varianz bekannt $\Rightarrow$ Gauss-Test
normalverteilt, Varianz unbekannt $\Rightarrow$ spezieller t-Test
nicht normal $\Rightarrow$ Vorzeichentest
dichotom $\Rightarrow$ Binomial Test
diskret $\Rightarrow$ ... spezieller χ^2-Test (später)

Gausstest

- **Gausstest**

Situation:

Test auf Mittelwert bei bekannter Varianz

H_0 : Erwartungswert μ von X ist μ_0: $\mu = \mu_0$
H_1 : Erwartungswert μ von X ist ungleich μ_0: $\mu \neq \mu_0$

Voraussetzungen:

repräsentative Stichprobe
X ist normalverteilt: $X_i \sim N(\mu, \sigma_0^2)$
σ_0^2 bekannt

Bemerkung:

Der Gauss-Test wird sehr selten auf reale Datensätze angewendet, da die Varianz fast nie bekannt ist. Er ist jedoch der wohl am leichtesten theoretisch zu verstehende Test und daher immer noch überall zu finden.

Befehl: ––

Einstichproben t-Test

- **Einstichproben t-Test**

  Situation:

  Test auf Mittelwert bei unbekannter Varianz

  H_0 : Erwartungswert μ von X ist μ_0: $\mu = \mu_0$
  H_1 : Erwartungswert μ von X ist ungleich μ_0: $\mu \neq \mu_0$
  Voraussetzungen:

  repräsentative Stichprobe
  X ist normalverteilt: $X_i \sim N(\mu, \sigma^2)$
  σ^2 ist unbekannt

  Bemerkung:

  Befehl: `t.test(X)`

Einstichproben t-Test (einseitig)

- **Einstichproben t-Test (einseitig)**

  Situation:

  Test auf Mittelwert bei unbekannter Varianz

  $H_0 : \mu \leq \mu_0$
  $H_1 : \mu > \mu_0$

  Voraussetzungen:

  repräsentative Stichprobe
  X ist normalverteilt: $X_i \sim N(\mu, \sigma^2)$
  σ^2 ist unbekannt

  Bemerkung:

  Entsprechende einseitige Test gibt es auch für kleiner und auch für viele andere Tests, wo wir das nicht im Einzelnen aufführen werden.

  Befehl: `t.test(X,alternative="greater")`

Binomial Test

- **Binomial Test**

  Situation:

  Test auf Erfolgswahrscheinlichkeit

  H_0 : Die Wahrscheinlichkeit p für einen Erfolg ist p_0: $p = p_0$
  H_1 : Die Wahrscheinlichkeit p für einen Erfolg ist nicht p_0: $p \neq p_0$

  Voraussetzungen:

  repräsentative Stichprobe
  X ist Binomialverteilt mit Erfolgswahrscheinlichkeit p: $X \sim Bi(p, n_0)$
  Die Anzahl der Versuche n_0 ist bekannt

  Bemerkung:

  Befehl: `binom.test(sum(X),`n_0`*length(X))`

Vorzeichentest

- **Vorzeichentest**

  Situation:

  Test auf bestimmten Median

  H_0 : Der Median der Verteilung von X ist m_0: $F_X(0.5) = m_0$
  H_1 : Der Median der Verteilung von X ist nicht m_0: $F_X(0.5) \neq m_0$

  Voraussetzungen:

  repräsentative Stichprobe
  Verteilungsfunktion F_X im Median stetig

  Bemerkung:

  Befehl: `binom.test(table(X<`m_0`))`

Ein-Stichproben-Tests

... auf Streuung

normalverteilt $\Rightarrow$ χ^2-Test (Chi-Quadrat-Test)
nicht normal $\Rightarrow$ ... problematisch

χ^2-Test auf Varianz

- **χ^2-Test auf Varianz**

  Situation:

  Test auf gegebenen Varianz bei normalverteilten Daten.

  $H_0 : \sigma^2 = \sigma_0^2$
  $H_1 : \sigma^2 \neq \sigma_0^2$ oder $\sigma^2 > \sigma_0^2$ oder $\sigma^2 < \sigma_0^2$

  Voraussetzungen:

  repräsentative Stichprobe
  $X_i \sim N(\mu, \sigma^2)$

  Bemerkung:

  Befehl: --

Zwei-Stichproben-Tests

... auf Verteilung

stetig $\Rightarrow$ zwei Stichproben KS-Test
diskret $\Rightarrow$ ... Tafeltests (später)

- **Zwei Stichproben Kolmogorov-Smirnov-Test**

  Situation:

  Testet die Gleichheit der stetigen Verteilungen der Stichproben

  $H_0 : F_X \equiv F_Y$
  $H_1 : F_X \not\equiv F_Y$

  Voraussetzungen:

  repräsentative Stichproben
  Für X und Y wurde die gleiche Rundungsregel verwendet

  Bemerkung:

  Befehl: `ks.test(X,Y)`

Zwei-Stichproben-Tests

... auf Lage

normalverteilt $\Rightarrow$ Zwei Stichproben t-test
nicht normal aber stetig $\Rightarrow$ Wilcoxen-Vorzeichen-Rang Test

Zwei-Stichproben-t-Test

- **Zwei-Stichproben-t-Test**

  Situation:

  Vergleich von Mittelwerten zweier Stichproben bei Normalverteilung und gleicher Varianz

  $H_0 : \mu_X = \mu_Y$
  $H_1 : \mu_X \neq \mu_Y$ oder $\mu_X > \mu_Y$ oder $\mu_X < \mu_Y$

  Voraussetzungen:

  repräsentative Stichproben
  $X_i \sim N(\mu_X, \sigma^2)$ *und* $Y_i \sim N(\mu_Y, \sigma^2)$

  Bemerkung:

  Die Normalverteilungsvoraussetzung ist relativ unkritisch, solange keine Ausreißer vorliegen und Verteilung ungefähr normal ist. Die Normalverteilungsvoraussetzung kann mit dem Shapiro-Wilk Test und die Varianzgleichheit mit dem F-Test überprüft werden.

  Befehl: `t.test(X,Y,var.equal=TRUE)`

Welchs t–Test

- **Welchs t–Test**

  Situation:

  Vergleich von Mittelwerten zweier Stichproben bei Normalverteilung und verschiedener Varianz

  $H_0 : \mu_X = \mu_Y$
  $H_1 : \mu_X \neq \mu_Y$ oder $\mu_X > \mu_Y$ oder $\mu_X < \mu_Y$

  Voraussetzungen:

  repräsentative Stichproben
  $X_i \sim N(\mu_X, \sigma_X^2)$ *und* $Y_i \sim N(\mu_Y, \sigma_Y^2)$, *i.i.d.*

  Bemerkung:

  Die Normalverteilungsvoraussetzung ist relativ unkritisch, solange keine Ausreißer vorliegen und Verteilung ungefähr normal ist.

  Befehl: `t.test(X,Y)`

Wilcoxen–Rang–Summen–Test

- **Wilcoxen–Rang–Summen–Test**

  Situation:

  Vergleich der Lage zweier Stichproben mit stetiger Verteilung

  $H_0 : \forall x : F_X(x) = F_Y(x)$
  $H_1 : \exists c \neq 0 : \forall x : F_X(x) = F_Y(x - c)$

  Voraussetzungen:

  repräsentative Stichproben
  die Verteilungen F_X und F_Y sind stetig.

  Bemerkung:

  Dieser Test arbeitet auch für andere Alternativen gut, bei der eine Gruppe größere Werte hat.

  Befehl: `wilcox.test(X,Y)`

Zwei-Stichproben-Tests

normalverteilt $\Rightarrow$ F-test
nicht normal aber stetig $\Rightarrow$ Fligner Test

F-Test

- **F-Test**

  Situation:

  Test auf Gleichheit der Varianz bei Normalverteilung

  $H_0 : \sigma_X^2 = \sigma_Y^2$
  $H_1 : \sigma_X^2 \neq \sigma_Y^2$ oder $\sigma_X^2 > \sigma_Y^2$ oder $\sigma_X^2 < \sigma_Y^2$

  Voraussetzungen:

  repräsentative Stichproben
  $X_i \sim N(\mu_X, \sigma_X^2)$, $Y_i \sim N(\mu_Y, \sigma_Y^2)$

  Bemerkung:

  Befehl: `var.test(X,Y)`

Fligner-Test

Für den nichtparametrischen Streuungsvergleich eignet sich auch der Fligner-Test, der als Mehrstichprobentest besprochen wird.

```
fligner.test(list(X,Y))
```

Gepaarte-Tests
... auf Lage

normalverteilt $\Rightarrow$ gepaarter t-test
nicht normal aber stetig $\Rightarrow$ Wilcoxen-Vorzeichen-Rang Test

gepaarter t-Test

- **gepaarter t-Test**

  Situation:

  Die Werte jeweils zweier aufeinanderfolgender Messungen am gleichen statistischen Individuum sollen verglichen werden.

  $H_0 : E[X - Y] = 0$
  $H_1 : E[X - Y] \neq 0$ oder

  Voraussetzungen:

  repräsentative Stichprobe
  $X_i - Y_i \sim N(\mu, \sigma^2)$

  Bemerkung:

  Dieses normalverteilungsbasierte Verfahren hat Probleme mit Ausreißern in der Differenz.

  Befehl: `t.test(X,Y,paired=TRUE)`

- **Wilcoxen-Vorzeichen-Rang-Test**

Situation:

Testet auf eine mittlere Änderung von 0 zwischen beiden Beobachtungen am gleichen Individuum.

H_0 : Die Verteilung von $X_i - Y_i$ ist symmetrisch um 0.
H_1 : Die Verteilung von $X_i - Y_i$ ist symmetrisch um ein $c \neq 0$

Voraussetzungen:

Die Verteilung ist für alle Paare gleich.

Bemerkung:

Dieses rangbasierte Verfahren hat Probleme mit Bindungen in den Differenzen.

Befehl: `wilcox.test(X,Y,paired=TRUE)`

Gepaarte Tests

... auf Abhängigkeit

Lineare Abhängigkeit $\Rightarrow$ (Pearson) Korrelationstests
Monotone Abhängigkeit $\Rightarrow$ Spearman Korrelationstests

- **Pearson–Korrelationstest**

  Situation:

  Test auf Pearson-Korrelation gleich 0

  $H_0 : \mathrm{cor}(X, Y) = 0$
  $H_1 : \mathrm{cor}(X, Y) \neq 0$

  Voraussetzungen:

  repräsentative Stichprobe
  X,Y gemeinsam normalverteilt

  Bemerkung:

  Befehl: `cov.test(X,Y)`

Spearman–Korrelationstest

- **Spearman–Korrelationstest**

  Situation:

  Test auf Spearman-Rang-Korrelation

  $H_0 : \mathrm{cor}(X, Y) = 0$
  $H_1 : \mathrm{cor}(X, Y) \neq 0$ oder $\mathrm{cor}(X, Y) > 0$ oder $\mathrm{cor}(X, Y) < 0$

  Voraussetzungen:

  repräsentative Stichproben
  stetige Verteilung / wenige Bindungen

  Bemerkung:

  Befehl: `cov.test(X,Y,method="spearman")`

Gepaarte Tests

... auf Abhängigkeit diskreter Größen

dichotom $\Rightarrow$ Fishers exakter Test
viele $\Rightarrow$ (spezieller) χ^2-Test
sonst $\Rightarrow$ ... problematisch

- **χ^2-Test auf Unabhängigkeit in Kontingenztafeln**

  Situation:

  Test auf Unabhängigkeit von kategoriellen Merkmalen

  H_0 : X und Y sind stochastisch unabhängig
  H_1 : X und Y sind stochastisch abhängig

  Voraussetzungen:

  repräsentative Stichprobe
  kategorielle Merkmale

  Bemerkung:

  Der p-Wert des Tests wird nur approximativ berechnet. Die Approximation ist schlecht, wenn in einzelnen Zellen der Datentafel unter der Unabhängigkeitsannahme weniger als 3-5 Werte zu erwarten sind.

  Befehl: `chisq.test(table(X,Y))`

Fishers exakter Test

- **Fishers exakter Test**

  Situation:

  Test auf Unabhängigkeit von 2x2 Kontingenztafeln und dichtomer Merkmale

  H_0 : Die Merkmale sind stochastisch Unabhängig
  H_1 : Die Merkmale sind stochastisch abhängig

  Voraussetzungen:

  repräsentative Stichprobe
  dichotome Merkmale

  Bemerkung:

  Im Gegensatz zum χ^2-Test auf Unabhängigkeit wird hier keine Approximation verwendet. Der Test ist also immer dann vorzuziehen, wenn er in der Situation anwendbar ist.

  Befehl: `fisher.test(table(X,Y))`

Tests auf Abhängigkeit

Stetige Größe Abhängig von diskreter Größe

dichotom $\Rightarrow$ Zwei-Stichproben-Test
mehrere Kategorien $\Rightarrow$ Mehr-Stichproben-Test

Mehrstichproben Tests

... auf Lage

Normalverteilt, homoskedastisch $\Rightarrow$ Varianzanalyse
immerhin stetig $\Rightarrow$ ... Kruskal-Wallis-Test

Einfache Varianzanalyse

- **Einfache Varianzanalyse**

Situation:

Test auf Gleichheit der Erwartungswerte mehrerer normalverteilter Stichproben.

$H_0 : \forall g, g' : \mu_g = \mu_{g'}$
$H_1 : \exists g, g' : \mu_{g_i} \neq \mu_j$

Voraussetzungen:

repräsentative Stichproben
$X_i \sim N(\mu_{g_i}, \sigma^2)$ *wobei g_i die Gruppenzugehörigkeit des Individuums i beschreibt.*

Bemerkung:

Die Varianzanalyse setzt die Gleichheit der Varianz und Normalverteilung voraus.

Befehl: `anova(lm(X~G))`

Kruskal–Wallis–Test

- **Kruskal–Wallis–Test**

  Situation:

  Test auf Gleichheit der Lage mehrerer stetig verteilter Stichproben.

  H_0 : Alle Gruppen haben die gleiche Verteilung
  H_1 : Die Verteilungen der Gruppen sind gegeneinander verschoben.

  Voraussetzungen:

  repräsentative Stichproben
  eigentlich: gleiche nur verschobene Verteilung

  Bemerkung:

  Der Kruskal Wallis Test ist ein Rangbasiertes Verfahrung und ist damit potentiell anfällig gegegen zu viele gleiche Messwerte.

  Befehl: `kruskal.test(X,G)`

Mehrstichproben Tests

... auf Streuung

normalverteilt $\Rightarrow$ Bartlett–Test
immerhin stetig $\Rightarrow$ Fligner–Test

Bartlett–Test

- **Bartlett–Test**

  Situation:

  Testet auf gleiche Varianz in mehrere Stichproben

  H_0 : Die Varianzen der Stichproben sind gleich
  H_1 : Die Varianzen der Stichproben sind nicht alle gleich.

  Voraussetzungen:

  repräsentative Stichproben
  $X_i \sim N(\mu_{G_i}, \sigma_{G_i})$

  Bemerkung:

  Dieser Test wird oft eingesetzt, um eine Voraussetzung der Varianzanalyse zu überprüfen.

  Befehl: `bartlett.test(X,G)`

Fligner–Test

- **Fligner–Test**

  Situation:

  Testet auf gleiche Streuung in mehreren Stichproben

  H_0 : Die Streuungen der Stichproben sind gleich.
  H_1 : Die Streuungen der Stichproben sind unterschiedlich.

  Voraussetzungen:

  repräsentative Stichproben
  stetige Verteilung

  Bemerkung:

  Befehl: `fligner.test(X,G)`

Übersicht

Stetige Skala

interessierende Parameter		Voraussetzung	Stichprobensituation: Ein-	Zwei-	Mehr-	gepaart/bivariat
Lage	normal	homoskedastisch	Ein-Stichproben t-test	Zwei-Stichproben t-test	Varianzanalyse (ANOVA)	gepaarter t-test
Lage	normal	heteroskedastisch		Welchs t-test	???	
Lage	stetig		Vorzeichentest	Wilcoxen Rang-Summen Test	Kruskal-Wallis-Test	Wilcoxen Vorzeichen-Rang Test
Streuung	normal		$\chi 2$-Test auf Streuung	F-Test	Bartlett-Test	$\chi 2$-Test auf Streuung auf Differenzen
Streuung	stetig		???	???	Fligner-Test	???
Verteilung	normal?		Shapiro-Wilk-Test	Shapiro-Wilk-Test auf beiden Stichp	Shapiro-Wilk-Test auf jeder Stichp	Shapiro-Wilk-Test auf Differenzen
Verteilung	identisch?			Z.-S.-K.-S.-Test	???	
Verteilung	speziell?		E.-S.-K.-S.-Test $\chi 2$-Test auf Vert		E.-S.-K.-S.-Test auf jeder Stichp	E.-S.-K.-S.-Test auf Differenzen
Abhängigkeit	normal		Lineare Modelle			Pearson-Korrelations-Test
Abhängigkeit	speziell/stetig		Generalisierte Lineare Modelle			Spearman-Korrelations-Test

Zwischenschritte