

# Statistik

## *Vorlesung 6 (Tests II)*

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

# Das Entscheidungsverfahren

Das Test (also das Entscheidungsverfahren)

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

besteht normalerweise aus:

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

# Die Berechnung des p-Wertes

Das Entscheidungsverfahren besteht normalerweise aus:

- ... und der resultierenden Entscheidungsregel

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

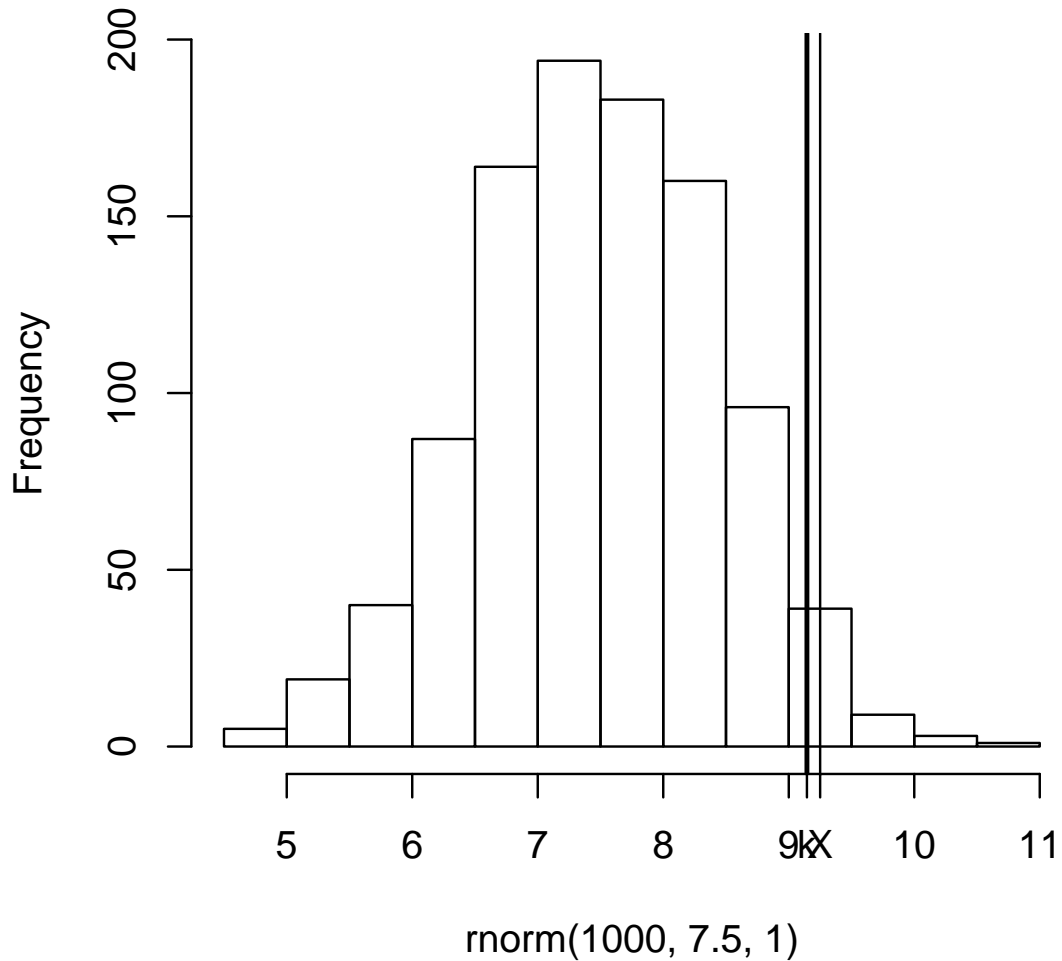
der p-Wert (kleinstes  $\alpha$ -Niveau zu dem abgelehrt wird) wird dann wie folgt berechnet:

$$p = 1 - F^{-1}(S(X))$$

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

# Kritischer Wert

Histogram of `rnorm(1000, 7.5, 1)`



$p = \text{Fläche oberhalb } S(X)$

# 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
- 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
- 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.

# 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.

# 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 betrachtet.
- multivariat: Es werden mehrere Merkmale betrachtet.

# Ein-Stichproben-Tests

## Verteilung

irgendeine Normalverteilung  $\Rightarrow$  Shapiro-Wilk-Test

eine bestimmte stetige Verteilung  $\Rightarrow$

(Ein-Stichproben)-KS-Test

# 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

```
> x <- rexp(20)
> shapiro.test(x)
```

Shapiro-Wilk normality test

data: x

W = 0.9069, p-value = 0.05563

# 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, F0)`

# Ein-Stichproben-Tests

## 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  $\chi^2$ -Test (später)

# Gausstest

## ● Gausstest

Situation:

Test auf Mittelwert bei bekannter Varianz

$H_0$  : Erwartungswert  $\mu$  von  $X$  ist  $\mu_0$ :  $\mu = \mu_0$

$H_1$  : Erwartungswert  $\mu$  von  $X$  ist ungleich  $\mu_0$ :  $\mu \neq \mu_0$

Voraussetzungen:

repräsentative Stichprobe

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

$\sigma_0^2$  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  $\mu$  von  $X$  ist  $\mu_0$ :  $\mu = \mu_0$

$H_1$  : Erwartungswert  $\mu$  von  $X$  ist ungleich  $\mu_0$ :  $\mu \neq \mu_0$

Voraussetzungen:

repräsentative Stichprobe

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

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

$\sigma^2$  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<m0))`



# Ein-Stichproben-Tests

## Streuung

normalverteilt  $\Rightarrow \chi^2$ -Test (Chi-Quadrat-Test)

nicht normal  $\Rightarrow \dots$  problematisch

# $\chi^2$ -Test auf Varianz

## • $\chi^2$ -Test auf Varianz

**Situation:**

Test auf gegebenen Varianz bei normalverteilten Daten.

$$H_0 : \sigma^2 = \sigma_0^2$$

$$H_1 : \sigma^2 \neq \sigma_0^2 \text{ oder } \sigma^2 > \sigma_0^2 \text{ oder } \sigma^2 < \sigma_0^2$$

**Voraussetzungen:**

repräsentative Stichprobe

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

**Bemerkung:**

**Befehl:** --

# Zwei-Stichproben-Tests

## Verteilung

stetig  $\Rightarrow$  zwei Stichproben KS-Test

diskret  $\Rightarrow$  ... Tafeltests (später)

# Zwei Stichproben Kolmogorov-Smirnov-Test

## • Zwei Stichproben Kolmogorov-Smirnov-Test

Situation:

Testet die Gleichheit der stetigen Verteilungen  
der Stichproben

$$H_0 : F_X \equiv F_Y$$

$$H_1 : F_X \neq F_Y$$

Voraussetzungen:

repräsentative Stichproben

Für X und Y wurde die gleiche Rundungsregel  
verwendet

Bemerkung:

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

# Zwei-Stichproben-Tests

## Lage

normalverteilt  $\Rightarrow$  Zwei Stichproben t-test

nicht normal aber stetig  $\Rightarrow$  Wilcoxon-Vorzeichen-Rang Test

# Zwei-Stichproben-t-Test

## • Zwei-Stichproben-t-Test

Situation:

Vergleich von Mittelwerten zweier Stichproben bei Normalverteilung und gleicher Varianz

$$H_0 : \mu_X = \mu_Y$$

$$H_1 : \mu_X \neq \mu_Y \text{ oder } \mu_X > \mu_Y \text{ oder } \mu_X < \mu_Y$$

Voraussetzungen:

repräsentative Stichproben

$$X_i \sim N(\mu_X, \sigma^2) \text{ und } Y_i \sim N(\mu_Y, \sigma^2)$$

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

$$H_0 : \mu_X = \mu_Y$$

$$H_1 : \mu_X \neq \mu_Y \text{ oder } \mu_X > \mu_Y \text{ oder } \mu_X < \mu_Y$$

Voraussetzungen:

repräsentative Stichproben

$$X_i \sim N(\mu_X, \sigma_X^2) \text{ und } Y_i \sim N(\mu_Y, \sigma_Y^2), \text{ i.i.d.}$$

Bemerkung:

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

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

# Wilcoxon–Rang–Summen–Test

## • Wilcoxon–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

## Streuung

normalverteilt  $\Rightarrow$  F-test

nicht normal aber stetig  $\Rightarrow$  Fligner Test

# F-Test

## ● F-Test

**Situation:**

Test auf Gleichheit der Varianz bei  
Normalverteilung

$$H_0 : \sigma_X^2 = \sigma_Y^2$$

$$H_1 : \sigma_X^2 \neq \sigma_Y^2 \text{ oder } \sigma_X^2 > \sigma_Y^2 \text{ oder } \sigma_X^2 < \sigma_Y^2$$

**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

## Lage

normalverteilt  $\Rightarrow$  gepaarter t-test

nicht normal aber stetig  $\Rightarrow$  Wilcoxon-Vorzeichen-Rang Test

# gepaarter t-Test

## ● gepaarter t-Test

Situation:

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

$$H_0 : E[X - Y] = 0$$

$$H_1 : E[X - Y] \neq 0 \text{ oder}$$

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)`

# Wilcoxon-Vorzeichen-Rang-Test

## ● Wilcoxon-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)`

# Tests auf Abhängigkeit

stetige Größen

⇒ Korrelationstests

# Pearson–Korrelationstest

## • Pearson–Korrelationstest

Situation:

Test auf Pearson-Korrelation gleich 0

$$H_0 : \text{cor}(X, Y) = 0$$

$$H_1 : \text{cor}(X, Y) \neq 0$$

Voraussetzungen:

repräsentative Stichprobe

X, Y gemeinsam normalverteilt

Bemerkung:

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



# Spearman–Korrelationstest

## • Spearman–Korrelationstest

Situation:

Test auf Spearman-Rang-Korrelation

$$H_0 : \text{cor}(X, Y) = 0$$

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

Voraussetzungen:

repräsentative Stichproben

stetige Verteilung / wenige Bindungen

Bemerkung:

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

# Tests auf Abhängigkeit

diskrete Größen

dichotom  $\Rightarrow$  Fishers exakter Test

viele  $\Rightarrow$  (spezieller)  $\chi^2$ -Test

sonst  $\Rightarrow$  ... problematisch

# $\chi^2$ -Test auf Unabhängigkeit in Kontingenztafeln

## • $\chi^2$ -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  
kategoriale 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 dichter Merkmale

$H_0$  : Die Merkmale sind stochastisch Unabhängig

$H_1$  : Die Merkmale sind stochastisch abhängig

Voraussetzungen:

repräsentative Stichprobe

dichotome Merkmale

Bemerkung:

Im Gegensatz zum  $\chi^2$ -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

## 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 Verfahren und ist damit potentiell anfällig gegeben zu viele gleiche Messwerte.

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



# Mehrstichproben Tests

## 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

		Voraussetzung	Stichprobensituation			
		Ein-	Zwei-	Mehr-	gepaart/bivariat	
interessierende Parameter	Lage	homo-skedastisch	Ein-Stichproben t-test	Zwei-Stichproben t-test	Varianzanalyse (ANOVA)	gepaarter t-test
		hetero-skedastisch		Welchs t-test	???	
	stetig		Vorzeichentest	Wilcoxon Rang-Summen Test	Kruskal-Wallis-Test	Wilcoxon Vorzeichen-Rang Test
		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
identisch?			Z.-S.-K.-S.-Test	???		
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
	speziell/stetig		Generalisierte Lineare Modelle			Spearman-Korrelations-Test

# Zwischenschritte

