

Mathematik 2

Dr. Thomas Zehrt

Testen

Benötigtes Vorwissen: Der Stoff der Vorlesung „Statistik“ wird als bekannt vorausgesetzt, insbesondere **Kapitel 11** „(Punkt)schätzen“ und **Kapitel 12** „Testen“

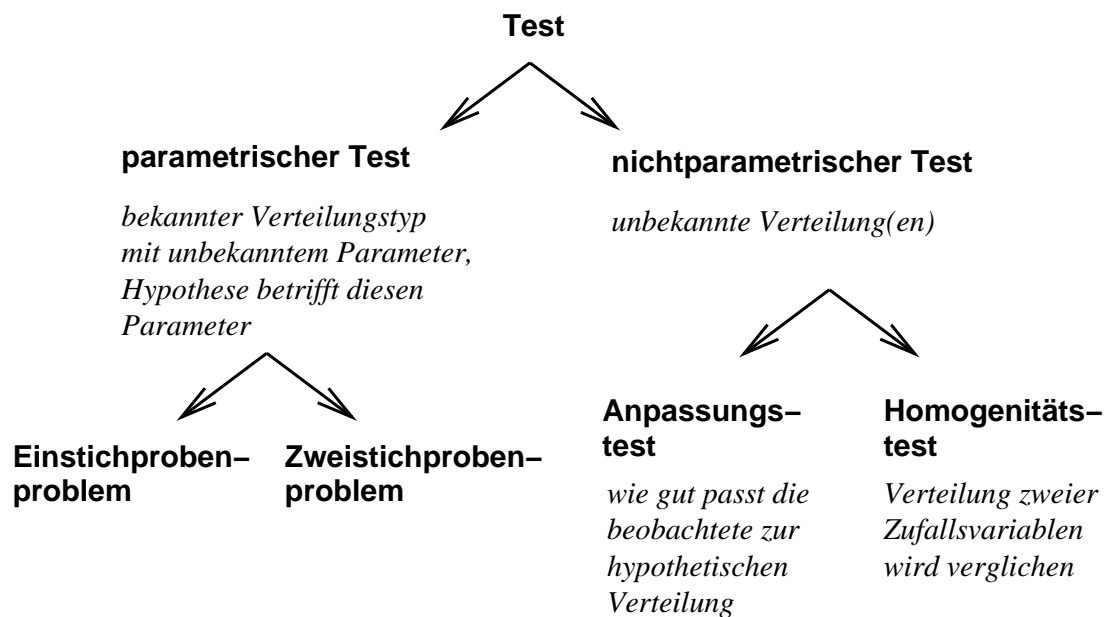
Inhaltsverzeichnis

1	Einleitung und Beispiel	2
1.1	Überblick	2
1.2	Beispiel	2
2	Parametrische Testprobleme	4
2.1	Allgemeines	4
3	Wichtige parametrische Tests	8
3.1	Der einfache Gaußtest	8
3.2	Weitere Tests	11
4	*Nichtparametrische Tests*	17
4.1	Anpassungstests	17
4.1.1	Der χ^2 -Anpassungstest	17
4.1.2	Der Kolmogorov-Smirnov Anpassungstest	19
5	Übungsaufgaben	20

1 Einleitung und Beispiel

1.1 Überblick

Hier sehen Sie eine Übersicht über typische allgemeine Testprobleme. Wir werden uns auf so genannte parametrische Tests einschränken. Wir werden also immer davon ausgehen, dass das zu untersuchende Merkmal (Zufallsvariable) einer bekannten Verteilung folgt, der Normalverteilung, dass wir aber (mindestens) einen der Parameter μ und σ nicht kennen. Nichtparametrische Testprobleme untersuchen, ob ein bestimmtes Merkmal einer vermuteten Verteilung folgt.



1.2 Beispiel

Beispiel 1.1 Eine Firma stellt Schokoladentafeln mit dem Gewicht 100 g her. Durch zufällige Einflüsse wiegt aber nicht jede Schokoladentafel exakt 100 g. Was können wir über das (echte) Durchschnittsgewicht **aller** hergestellten Tafeln aussagen? Eine Totalerhebung ist sicher zu aufwändig, könnte aber unsere Frage sicher und richtig beantworten.

Verteilungsannahme: Die Zufallsvariable $X =$ „Gewicht einer Schokoladentafel“ ist normalverteilt mit unbekanntem Mittelwert μ und der Einfachheit halber bekannter Varianz σ^2 . Zu X gehört also der Hypothesenraum aller Normalverteilungen $\{N(\mu, \sigma^2) \mid \mu \in \mathbb{R}\}$.

Irgendeine dieser Normalverteilungen beschreibt also X perfekt. Nur welche?

Wir könnten eine Stichprobe von n Schokoladentafeln nehmen und deren durchschnittliches Gewicht bestimmen.

Testgröße (Zufallsvariable bzw. Punktschätzer für μ):

$\bar{X} = \bar{X}_n =$ „arithmetisches Mittel der Gewichte von n zufällig ausgewählten Schokoladentafeln“. \bar{X} ist unbedingt als Zufallsvariable anzusehen, jede

mögliche Auswahl von n der sehr, sehr vielen Schokoladentafeln könnte in der Stichprobe landen und würden zu einem der vielen möglichen Werten von \bar{X} gehören. Da X normalverteilt ist, ist auch \bar{X} normalverteilt mit Mittelwert μ und Varianz σ^2/n .

Mit dieser Testgrösse lässt sich schon recht gut arbeiten, in der Praxis wird aber meist noch standardisiert.

Die typische Gedankenkette könnte jetzt wie folgt aussehen:

- Angenommen (Nullhypothese), dass $\mu = \mu_0 = 100$ wäre.
- Die Verteilung von \bar{X} ist dann vollständig bekannt: $\bar{X} \sim N(100, \sigma^2)$
- \bar{X} würde mit grosser Wahrscheinlichkeit einen Wert nahe bei 100 realisieren. Wir könnten diese Wahrscheinlichkeit sogar für jedes Intervall konkret berechnen!
- Realisiert \bar{X} einen Wert nahe bei 100, glauben wir weiterhin an unsere Anfangsannahme (Nullhypothese).
- Realisiert \bar{X} dagegen einen (unwahrscheinlichen, aber nicht unmöglichen) Wert weit entfernt von 100 (im Verwerfungsbereich), verwerfen wir die Nullhypothese und glauben, dass $\mu \neq 100$ ist.

Die Zuordnung der Hypothesen hängt davon ab, was man beweisen will. H_1 sollte aber die eigentliche Forschungshypothese sein.

Die folgenden drei Testprobleme könnten dann von Interesse sein:

1. $H_0 : \mu = 100 \text{ g}$

$H_1 : \mu \neq 100 \text{ g}$

Wir sind hier nur daran interessiert, ob der Hersteller das angegebene Gewicht **exakt** einhält.

2. $H_0 : \mu \leq 100 \text{ g}$

$H_1 : \mu > 100 \text{ g}$

Dieser Test wäre sinnvoll, wenn in der Firma der Verdacht besteht, dass die Schokoladentafeln mehr als 100 g wiegen, man also Schokolade verschenkt.

3. $H_0 : \mu \geq 100 \text{ g}$

$H_1 : \mu < 100 \text{ g}$

Hier wird untersucht, ob die Schokoladentafeln im Schnitt weniger als 100 g wiegen. Dieses Testproblem ist wohl aus Sicht des Kunden am wichtigsten.

2 Parametrische Testprobleme

2.1 Allgemeines

Sei X eine Zufallsvariable mit den möglichen Verteilungen

$$\text{Hypothesenraum: } \Omega = \left\{ P_\theta : \theta \in \Theta \right\}$$

Bei einem (parametrischen) Testproblem wird der Hypothesenraum in zwei Teilmengen aufgeteilt:

- die zu testende Nullhypothese $H_0 : \theta \in \Theta_0$ und
- die Alternative $H_1 : \theta \in \Theta_1$.

Die Nullhypothese H_0 ist diejenige Hypothese, welche auf ihren Wahrheitsgehalt hin überprüft werden soll. Sie beinhaltet den Zustand des Parameters der Grundgesamtheit, der bis zum jetzigen Zeitpunkt bekannt ist oder als allgemein akzeptiert gilt.

Die Alternativhypothese H_1 ist ein Teil der zur Nullhypothese entgegengesetzten Aussage. Sie ist die eigentliche Forschungshypothese und drückt aus, was mittels der statistischen Untersuchung gezeigt werden soll.

Es gilt stets: $\Theta_0 \cap \Theta_1 = \emptyset$ und bei Signifikanztests (die wir im weiteren ausschliesslich behandeln werden) auch $\Theta_0 \cup \Theta_1 = \Theta$.

Die typischen Situationen dabei sind:

	H_0	H_1	
1.	$H_0 : \theta = \theta_0$	$H_1 : \theta = \theta_1, \theta_0 \neq \theta_1$	<u>einfache Hypothese</u>
2.	$H_0 : \theta = \theta_0$	$H_1 : \theta \neq \theta_0$	<u>zweiseitige Fragestellung</u>
3.	$H_0 : \theta \leq \theta_0$	$H_1 : \theta > \theta_0$	<u>einseitige Fragestellung</u>
4.	$H_0 : \theta \geq \theta_0$	$H_1 : \theta < \theta_0$	<u>einseitige Fragestellung</u>

Eine Funktion $T(\mathbf{X}) = T(X_1, \dots, X_n)$ der Stichprobenvariablen $\mathbf{X} = (X_1, \dots, X_n)$ heisst Testgrösse. Für die konkrete Stichprobe (x_1, \dots, x_n) ergibt sich $t = T(x_1, \dots, x_n)$ als Realisierung der Testgrösse.

Der Wertebereich der Zufallsgrösse $T(\mathbf{X})$ wird in folgende zwei Teile zerlegt:

- Verwerfungsbereich (oder auch kritischer Bereich bzw. Ablehnbereich) R
- Annahmehereich $\bar{R} = R^c$.

Aufgrund der Realisierung (x_1, \dots, x_n) wird dann folgende Testentscheidung getroffen:

- H_0 ablehnen, falls $t = T(x_1, \dots, x_n) \in R$
- H_0 beibehalten, falls $t = T(x_1, \dots, x_n) \in \bar{R} = R^c$

Innerhalb des gewählten Modells gibt es nun vier Möglichkeiten, wie die Realität und die Testentscheidung zusammentreffen können.

		Realität	
		H_0 ist richtig	H_0 ist falsch
Testent- scheidung	H_0 beibehalten	ok	Fehler 2. Art, β -Fehler
	H_0 verwerfen	Fehler 1. Art, α -Fehler	ok

Bei der Testkonstruktion gibt man die Wahrscheinlichkeit α eines Fehlers 1. Art vor. Diese Schranke bezeichnet man als Signifikanzniveau und der Test heisst Signifikanztest zum Niveau α .

Der kritische Bereich R wird dann so konstruiert, dass die Wahrscheinlichkeit eines Fehlers 1. Art nicht grösser als α wird:

$$P(\underbrace{H_0 \text{ verwerfen}}_{T(\mathbf{X}) \in R} \mid H_0 \text{ richtig}) \leq \alpha$$

Dabei sollte aber die Wahrscheinlichkeit eines Fehlers 2. Art $P(H_0 \text{ beibehalten} \mid H_0 \text{ falsch})$ nicht zu gross werden.

Beim Testen wird nur die Wahrscheinlichkeit eines Fehlers 1. Art durch α kontrolliert. Wenn also H_0 wirklich gilt, werden wir sicher nur in $\alpha \cdot 100\%$ der durchgeführten Tests für H_1 entscheiden.

Die Entscheidung **für** H_1 ist somit statistisch gesichert und wir sprechen dann von einem signifikanten Ergebnis.

Die Wahrscheinlichkeit für einen Fehler 2. Art wird dagegen nicht kontrolliert und die Entscheidung H_0 beizubehalten ist statistisch **nicht** abgesichert. Es spricht nur nichts gegen H_0 .

Allgemeines Vorgehen:

1. Verteilungsannahme über die Zufallsvariable X (bzw. über deren Verteilungsfunktion F) machen.
2. Formulieren der Nullhypothese und der Alternative.
3. Vorgabe der Irrtumswahrscheinlichkeit α .
4. Konstruktion einer geeigneten Testgrösse $T(\mathbf{X}) = T(X_1, \dots, X_n)$ als Funktion der Stichprobenvariablen \mathbf{X} , deren Verteilung unter der Nullhypothese vollständig bekannt sein muss.
5. Wahl eines kritischen Bereichs R aus dem möglichen Wertebereich von $T(\mathbf{X})$, so dass $P(T(\mathbf{X}) \in R | H_0 \text{ richtig}) \leq \alpha$ gilt.
6. Bestimmung der Realisierung $t = T(x_1, \dots, x_n)$ der Testgrösse $T(\mathbf{X})$ anhand einer konkreten Stichprobe.
7. Entscheidungsregel: Liegt t in R , so wird die Nullhypothese abgelehnt, sonst beibehalten.

Der p -Wert Beim Einsatz von Statistiksoftware zum Hypothesentest wird der kritische Bereich meist nicht angezeigt. Stattdessen wird der konkrete Wert t der Stichprobe zusammen mit dem zugehörigen p -Wert angezeigt.

Definition 2.1 *Der p -Wert ist definiert als die Wahrscheinlichkeit unter Gültigkeit von H_0 , dass die Testgrösse T den konkreten Wert t oder einen im Sinne der Alternativhypothese noch extremeren Wert annimmt.*

Konkret bedeutet das:

$$\text{rechtsseitiger Test: } p\text{-Wert} = P(T \geq t \mid H_0 \text{ stimmt})$$

$$\text{linksseitiger Test: } p\text{-Wert} = P(T \leq t \mid H_0 \text{ stimmt})$$

$$\text{zweiseitiger Test: } p\text{-Wert} = P(|T| \geq t \mid H_0 \text{ stimmt})$$

Der p -Wert kann ebenfalls als Entscheidungskriterium für das Verwerfen der Nullhypothese genutzt werden.

Je kleiner der p -Wert ist, desto stärker sprechen die Daten gegen die Nullhypothese und damit für die Alternativhypothese.

Gebäuchliche Grenzen sind:

- $p\text{-Wert} > 0.100$: schwache Beweislast gegen H_0
- $0.050 < p\text{-Wert} \leq 0.100$: mässige Beweislast gegen H_0
- $0.010 < p\text{-Wert} \leq 0.050$: moderate Beweislast gegen H_0
- $0.001 < p\text{-Wert} \leq 0.010$: starke Beweislast gegen H_0

3 Wichtige parametrische Tests

3.1 Der einfache Gaußtest

Untersuchen des Mittelwertes bei bekannter Varianz (einfacher Gaußtest)

Gegeben: $X \sim N(\mu, \sigma^2)$ mit $\sigma^2 = \sigma_0^2$ bekannt

Zu prüfen: Besitzt μ (unbekannt) einen bestimmten Wert $\mu = \mu_0$?

1. Verteilungsannahme: $X \sim N(\mu, \sigma_0^2)$
2. Festlegen von H_0 und H_1
 - I: $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
 - II: $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$
 - III: $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$
3. Vorgabe von α : $\alpha = 0.05$

4. Konstruktion der Testgrösse:

$$\text{Stichprobenmittelwert } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \stackrel{H_0}{\sim} N(\mu_0, \sigma_0^2/n)$$

$$\text{Standardisierung } T(\mathbf{X}) = \frac{\bar{X} - \mu_0}{\sigma_0} \sqrt{n} \stackrel{H_0}{\sim} N(0, 1)$$

5. Kritischer Bereich:

	H_0	H_1	R
I:	$\mu = \mu_0$	$\mu \neq \mu_0$	$(-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty)$
II:	$\mu \leq \mu_0$	$\mu > \mu_0$	$(z_{1-\alpha}, \infty)$
III:	$\mu \geq \mu_0$	$\mu < \mu_0$	$(-\infty, z_\alpha)$

6. Realisierung der Testgrösse: Aus einer konkreten Stichprobe x_1, \dots, x_n wird der Stichprobenmittelwert $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ und daraus die Realisierung der Testgrösse ermittelt $t = \frac{\bar{x} - \mu_0}{\sigma_0} \sqrt{n}$

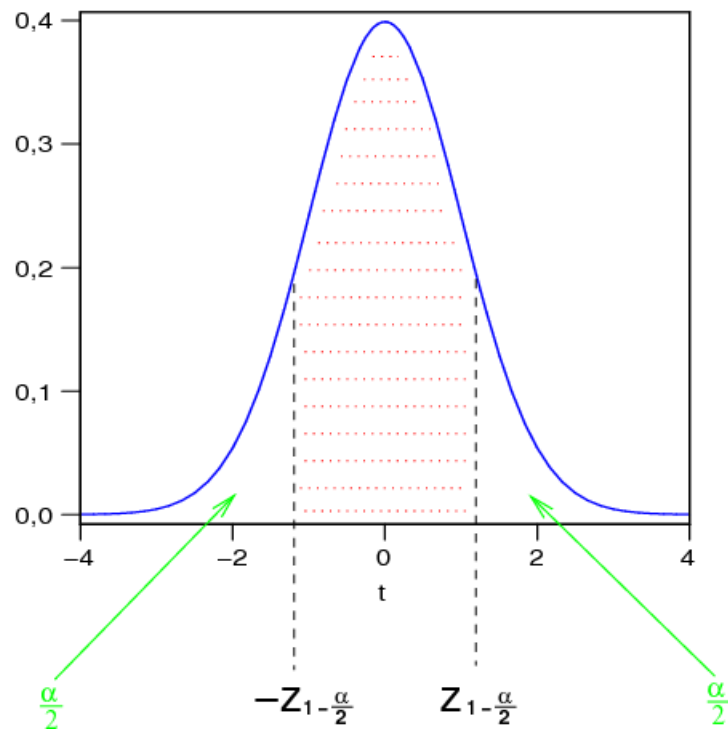
7. Testentscheidung: Liegt t in R , so muss die Nullhypothese verworfen werden. Daraus folgt insgesamt: Lehne H_0 ab, wenn

- I. $|t| > z_{1-\frac{\alpha}{2}}$
- II. $t > z_{1-\alpha}$
- III. $t < -z_{1-\alpha} = z_\alpha$

Beispiel 3.1 Für den zweiseitigen Test $H_0 : \mu = \mu_0$ und $H_1 : \mu \neq \mu_0$ wissen wir also, dass die Testgrösse T unter der Nullhypothese standardnormalverteilt ist. Wir kennen also die Wahrscheinlichkeiten, mit denen T bestimmte Werte realisiert.

Natürlich werden wir die Nullhypothese $\mu = \mu_0$ verwerfen, wenn der Stichprobenmittelwert \bar{X} (der den Erwartungswert μ_0 hat) viel kleiner oder viel grösser als μ_0 ist. Das ist gleichbedeutend damit, dass die Testgrösse T einen zu kleinen oder einen zu grossen Wert realisiert.

Zu klein bedeutet dabei, dass $t < -z_{1-\frac{\alpha}{2}}$, zu gross bedeutet $t > z_{1-\frac{\alpha}{2}}$



Wegen der Symmetrie der Standardglockenkurve gilt auch noch:

$$-z_{1-\frac{\alpha}{2}} = z_{\frac{\alpha}{2}}$$

Aufgabe 3.1 Bei der Produktion von Bolzen beträgt der Sollwert für den Durchmesser 10 mm. Der Bolzendurchmesser X sei dabei normalverteilt mit dem unbekanntem Erwartungswert μ und der durch die Konstruktion der Maschine festgelegten Standardabweichung $\sigma = 0.5$ mm. Zur Überprüfung der Einstellung werden 100 Teile entnommen und daraus der mittlere Durchmesser $\bar{x} = 10.15$ mm berechnet. Prüfen Sie die Hypothese $H_0 : \mu = 10$ mm mit der Irrtumswahrscheinlichkeit $\alpha = 0.05$.

Lösung:

1. $H_0 : \mu = 10 = \mu_0$

2. $H_1 : \mu \neq 10$

3. Test(Name): Gauss-Test

4. T (allgemein): $T = \frac{\bar{X} - \mu_0}{\sigma_0} \sqrt{n}$

5. t (speziell): $t = \frac{10.15 - 10}{0.5} \sqrt{100} = 3$

6. vollständiger Verwerfungsbereich:

$$\begin{aligned} R &= (-\infty, -z_{1-\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}}, \infty) \\ &= (-\infty, -z_{0.975}) \cup (z_{0.975}, \infty) \\ &= (-\infty, -1.96) \cup (1.96, \infty) \end{aligned}$$

7. Entscheidung: H_0 ablehnen, da $3 \in R$

Alternativ zu diesem Vorgehen könnten wir auch den p -Wert bestimmen. Da es sich um einen zweiseitigen Test handelt, unsere Testgröße $T = \frac{X-10}{0.5} \sqrt{100}$ unter der Annahme der Korrektheit der Nullhypothese $\mu = 10$ eine standardnormalverteilte Zufallsvariable ist und sich mit dem Wert $T = 3$ realisiert hat, gilt:

$$p\text{-Wert} = P(|T| \geq 3 \mid \mu = 10) = P(T \geq 3) + P(T \leq -3) \approx 0.001.$$

Wir haben hier also eine starke Beweislast gegen H_0 .

3.2 Weitere Tests

Signifikanztests für μ und σ einer normalverteilten Grundgesamtheit (Signifikanzniveau α)

Hypothesen H_0 H_1	Voraussetzungen	Testgrösse T	Verteilung von T	Kritischer Bereich
<u>Gauß-Test</u>				
$\mu = \mu_0$ $\mu \neq \mu_0$ $\mu \leq \mu_0$ $\mu > \mu_0$ $\mu \geq \mu_0$ $\mu < \mu_0$	$\sigma^2 = \sigma_0^2$ bekannt	$\frac{\bar{X} - \mu_0}{\sigma_0} \sqrt{n}$	$N(0, 1)$ -Verteilung	$ t > z_{1-\frac{\alpha}{2}}$ $t > z_{1-\alpha}$ $t < -z_{1-\alpha}$
<u>Einfacher t-Test</u>				
$\mu = \mu_0$ $\mu \neq \mu_0$ $\mu \leq \mu_0$ $\mu > \mu_0$ $\mu \geq \mu_0$ $\mu < \mu_0$	σ^2 unbekannt	$\frac{\bar{X} - \mu_0}{S_X} \sqrt{n}$	t_{n-1} -Verteilung	$ t > t_{n-1; 1-\frac{\alpha}{2}}$ $t > t_{n-1; 1-\alpha}$ $t < -t_{n-1; 1-\alpha}$
<u>χ^2-Streuungstest</u>				
$\sigma^2 = \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$ $\sigma^2 \leq \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$ $\sigma^2 < \sigma_0^2$	μ unbekannt	$\frac{(n-1) \cdot S_X^2}{\sigma_0^2}$	χ_{n-1}^2 -Verteilung	$t > \chi_{n-1; 1-\frac{\alpha}{2}}^2$ oder $t < \chi_{n-1; \frac{\alpha}{2}}^2$ $t > \chi_{n-1; 1-\alpha}^2$ $t < \chi_{n-1; \alpha}^2$
Ergänzung (nicht prüfungsrelevant)				
$\sigma^2 = \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$ $\sigma^2 \leq \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$ $\sigma^2 < \sigma_0^2$	μ bekannt	$\frac{n \cdot S_X^{*2}}{\sigma_0^2}$	χ_n^2 -Verteilung	$t > \chi_{n; 1-\frac{\alpha}{2}}^2$ oder $t < \chi_{n; \frac{\alpha}{2}}^2$ $t > \chi_{n; 1-\alpha}^2$ $t < \chi_{n; \alpha}^2$

Signifikanztests zum Vergleich zweier Erwartungswerte μ_X und μ_Y und Varianzen σ_X^2 und σ_Y^2 von normalverteilten Grundgesamtheiten (Signifikanzniveau α)

Hypothesen H_0 H_1	Voraussetzungen	Testgrösse T	Verteilung von T	Kritischer Bereich
<u>Doppelter t-Test</u>				
$\mu_X = \mu_Y$ $\mu_X \neq \mu_Y$	$(X_1, \dots, X_{n_1}), (Y_1, \dots, Y_{n_2})$ unabh. Stichproben	$\frac{\bar{X} - \bar{Y}}{S} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$ mit	$t_{n_1+n_2-2}$ -Verteilung	$ t > t_{n_1+n_2-2; 1-\frac{\alpha}{2}}$
$\mu_X \leq \mu_Y$ $\mu_X > \mu_Y$	$\sigma_X^2 = \sigma_Y^2$	$S =$		$t > t_{n_1+n_2-2; 1-\alpha}$
$\mu_X \geq \mu_Y$ $\mu_X < \mu_Y$	$X \sim N(\mu_X, \sigma_X^2)$ $Y \sim N(\mu_Y, \sigma_Y^2)$	$\sqrt{\frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2}}$		$t < -t_{n_1+n_2-2; 1-\alpha}$
<u>Welch-Test</u>				
$\mu_X = \mu_Y$ $\mu_X \neq \mu_Y$	$(X_1, \dots, X_{n_1}), (Y_1, \dots, Y_{n_2})$ unabh. Stichproben	$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}}$	t_m -Verteilung $m =$	$ t > t_{m; 1-\frac{\alpha}{2}}$
$\mu_X \leq \mu_Y$ $\mu_X > \mu_Y$	$\sigma_X^2 \neq \sigma_Y^2$		$\frac{(S_X^2/n_1 + S_Y^2/n_2)^2}{\left(\frac{S_X^2/n_1}{n_1-1} + \frac{S_Y^2/n_2}{n_2-1}\right)}$	$t > t_{m; 1-\alpha}$
$\mu_X \geq \mu_Y$ $\mu_X < \mu_Y$	$X \sim N(\mu_X, \sigma_X^2)$ $Y \sim N(\mu_Y, \sigma_Y^2)$		ganzzahlig gerundet	$t < -t_{m; 1-\alpha}$
<u>F-Test</u>				
$\sigma_X^2 = \sigma_Y^2$ $\sigma_X^2 \neq \sigma_Y^2$	$(X_1, \dots, X_{n_1}), (Y_1, \dots, Y_{n_2})$ unabh. Stichproben	S_X^2/S_Y^2	F_{n_1-1, n_2-1} -Verteilung	$t > F_{n_1-1, n_2-1; 1-\frac{\alpha}{2}}$ oder $t < F_{n_1-1, n_2-1; \frac{\alpha}{2}}$
$\sigma_X^2 \leq \sigma_Y^2$ $\sigma_X^2 > \sigma_Y^2$	μ_X, μ_Y unbekannt	S_X^2/S_Y^2	F_{n_1-1, n_2-1} -Verteilung	$t > F_{n_1-1, n_2-1; 1-\alpha}$
$\sigma_X^2 \geq \sigma_Y^2$ $\sigma_X^2 < \sigma_Y^2$	$X \sim N(\mu_X, \sigma_X^2)$ $Y \sim N(\mu_Y, \sigma_Y^2)$	S_Y^2/S_X^2	F_{n_2-1, n_1-1} -Verteilung	$t > F_{n_2-1, n_1-1; 1-\alpha}$

Beachte: Gilt $X \sim F_{m,n}$, so ist $1/X \sim F_{n,m}$ und für die Quantile folgt $F_{m,n;\alpha} = \frac{1}{F_{n,m;1-\alpha}}$.

Hypothesen H_0 H_1	Voraussetzungen	Testgrösse T	Verteilung von T	Kritischer Bereich
<u>paired t-Test</u>				
$\mu_X = \mu_Y$ $\mu_X \neq \mu_Y$	$(X_1, \dots, X_n), (Y_1, \dots, Y_n)$ (verbundene) Stichproben	$\frac{\bar{D}}{S_D} \sqrt{n}$ mit		$ t > t_{n-1; 1-\frac{\alpha}{2}}$
$\mu_X \leq \mu_Y$ $\mu_X > \mu_Y$	$D = X - Y \sim N(0, \sigma_D^2)$	$S_D^2 = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2$	t_{n-1} -Verteilung	$t > t_{n-1; 1-\alpha}$
$\mu_X \geq \mu_Y$ $\mu_X < \mu_Y$				$t < -t_{n-1; 1-\alpha}$

Aufgabe 3.2 Zur Überprüfung einer Produktionsanlage, die Kilotüten mit Mehl abfüllt, werden 20 Tüten entnommen, gewogen und daraus das mittlere Gewicht $\bar{x} = 998.92$ g und die Standardabweichung $s = 0.738$ g berechnet. Es kann davon ausgegangen werden, dass das Gewicht normalverteilt ist. Prüfen Sie, ob die Anlage im Mittel wesentlich weniger ($\alpha = 0.05$) als 1000 g abfüllt.

Lösung:

1. H_0 :
2. H_1 :
3. Test(Name):
4. T (allgemein):
5. t (speziell):
6. vollständiger Verwerfungsbereich:
7. Entscheidung:

Resultat:

- Lösung direkt: H_0 wird abgelehnt, denn $t = -6.58 < -1.73 = t_{19;0.95}$
- Lösung mit p -Wert:

$$p - \text{Wert} = P(T \leq -6.54 \mid \mu = 1000) \approx 0$$

also starke Beweislast gegen H_0 .

Aufgabe 3.3 Zur Überprüfung einer Produktionsanlage, die Kilotüten mit Mehl abfüllt, werden 20 Tüten entnommen, gewogen und daraus das mittlere Gewicht $\bar{x} = 998.92$ g und die Standardabweichung $s = 0.738$ g berechnet. Es kann davon ausgegangen werden, dass das Gewicht normalverteilt ist. Unterscheidet sich die empirische Streuung s^2 wesentlich ($\alpha = 0.05$) von $\sigma_0^2 = 0.3$ (g^2)

Lösung:

1. H_0 :
2. H_1 :
3. Test(Name):
4. T (allgemein):
5. t (speziell):
6. vollständiger Verwerfungsbereich:
7. Entscheidung:

Resultat:

- Lösung direkt: H_0 wird abgelehnt, denn $t = 34.49 > 32.85 = \chi_{19;0.975}^2$
- Lösung mit p -Wert:

p - Wert =

Aufgabe 3.4 In einer landwirtschaftlichen Versuchsanstalt erhielten 9 von 22 Masttieren (Gruppe I) Grünfütterzumischungen, während die übrigen 13 Tiere (Gruppe II) ausschliesslich mit dem proteinhaltigen Mastfutter gefüttert wurden. Nach einer gewissen Zeit wurden folgende Gewichtszunahmen in kg bei den Tieren festgestellt.

Gruppe I: 7.0, 11.8, 10.1, 8.5, 10.7, 13.2, 9.4, 7.9, 11.1 und Gruppe II: 13.4, 14.6, 10.4, 11.9, 12.7, 16.1, 10.7, 8.3, 13.2, 10.3, 11.3, 12.9, 9.7

Testen Sie die Nullhypothese $\sigma_I^2 = \sigma_{II}^2$ mit Hilfe des F-Tests (für $\alpha = 0.1$).

Lösung:

1. H_0 :
2. H_1 :
3. Test(Name):
4. T (allgemein):
5. t (speziell):
6. vollständiger Verwerfungsbereich:
7. Entscheidung:

Resultat:

- Lösung direkt: H_0 wird nicht abgelehnt
- Lösung mit p -Wert:

$$p - \text{Wert} =$$

4 *Nichtparametrische Tests*

4.1 Anpassungstests

Bei einem Anpassungstest wird untersucht, wie gut sich eine beobachtete Verteilung der hypothetischen Verteilung „anpasst“. Die Nullhypothese ist hier stets die Vermutung, dass eine ganz bestimmte (hypothetische) Verteilung vorliegt. Es kann also (statistisch) nur bewiesen werden, dass dieser bestimmte Verteilungstyp **nicht** vorliegt.

4.1.1 Der χ^2 -Anpassungstest

Gegeben: Zufallsvariable X (ordinal oder stetig)

Zu prüfen: Besitzt X die Verteilungsfunktion F_0 ?

1. Verteilungsannahme: X hat die Verteilungsfunktion F_0
2. Festlegen von H_0 und H_1

H_0 : „Die Verteilungsfunktion F einer Stichprobe stimmt mit F_0 überein: $F(x) = F_0(x)$ “

H_1 : $F(x) \neq F_0(x)$
3. Vorgabe von α : $\alpha = 0.05$
4. Konstruktion der Testgrösse: Eine Stichprobe $\mathbf{X} = (X_1, \dots, X_n)$ wird in k Klassen eingeteilt:

Klasse	1	2	...	k	Total
Anz. Beobacht.	n_1	n_2	...	n_k	n

$$T(\mathbf{X}) = \sum_{i=1}^k \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i} \stackrel{H_0}{\sim} \chi_{k-1-r}^2$$

- N_i absolute Häufigkeit der Stichprobe für die Klasse i (das sind Zufallsvariablen mit den Realisierungen n_i)
 - p_i die mit Hilfe der (hypothetischen) Verteilungsfunktion F_0 berechneten Wahrscheinlichkeiten dafür, dass X einen Wert in der Klasse i annimmt
 - $n \cdot p_i$ die unter H_0 erwartete Häufigkeit in der Klasse i
 - r Anzahl Parameter von F_0
5. Kritischer Bereich: $(\chi_{k-1-r;1-\alpha}, \infty)$
 6. -
 7. Testentscheidung: Lehne H_0 ab, wenn $t > \chi_{k-1-r;1-\alpha}$

Aufgabe 4.1 Eine Versuchsreihe mit $n = 250$ Würfeln einer belgischen 1-Euromünze zeigt 140-mal den (massigen) Kopf des belgischen Königs Albert und nur 110-mal die Zahl. Entscheiden Sie mittels χ^2 -Anpassungstest (und mit $\alpha = 0.05$), ob die Münze fair ist.

Lösung:

1. Verteilungsannahme: Ausgang 1 sei „Kopf“ und Ausgang 2 sei „Zahl“, dann ist X (Anzahl „Kopf“) gleichverteilt mit $p(X = 1) = p(X = 2) = 1/2$
2. H_0 : Verteilung der Stichprobe entspricht der Gleichverteilung
3. Vorgabe von α : $\alpha = 0.05$
- 4./6. $k = 2$

Verteilung von T : $\chi_{2-1-0}^2 = \chi_1^2$

Klasse	1	2
Anz. Beobacht.	$n_1 = 140$	$n_2 = 110$

$$t = \sum_{i=1}^2 \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i} = \frac{(140 - 250 \cdot 1/2)^2}{250 \cdot 1/2} + \frac{(110 - 250 \cdot 1/2)^2}{250 \cdot 1/2} = 3.6$$

5. Kritischer Bereich: Es gilt $c_{1;0.95} = 3.84$ also $K = (3.84, \infty)$
7. Testentscheidung:

$t = 3.6 < \chi_{1;0.95} = 3.84$ also kann die Nullhypothese nicht verworfen werden.

4.1.2 Der Kolmogorov-Smirnov Anpassungstest

Ein Nachteil beim χ^2 -Anpassungstest ist die (willkürliche) Klasseneinteilung, die (insbesondere bei kleinen Stichproben) das Testergebnis beeinflusst. Hier ist der Kolmogorov-Smirnov Anpassungstest besser geeignet!

Gegeben: Zufallsvariable X (ordinal oder stetig)

Zu prüfen: Besitzt X die Verteilungsfunktion F_0 ?

1. Verteilungsannahme: X hat die Verteilungsfunktion F_0
2. Festlegen von H_0 und H_1

H_0 : „Die Verteilungsfunktion F einer Stichprobe stimmt mit F_0 überein: $F(x) = F_0(x)$ “

H_1 : $F(x) \neq F_0(x)$
3. Vorgabe von α : $\alpha = 0.05$
4. Konstruktion der Testgröße: Eine Stichprobe (x_1, \dots, x_n) wird geordnet $x_{(1)} \leq \dots \leq x_{(n)}$ und die Verteilungsfunktion konstruiert.

$$\hat{F}(x) := \begin{cases} 0 & \text{für } -\infty < x < x_{(1)} \\ i/n & \text{für } x_{(i)} \leq x < x_{(i+1)} \\ 1 & \text{für } x_{(n)} \leq x < \infty \end{cases}$$

Testgröße:

$$D = \sup_{x \in \mathbb{R}} |F_0(x) - \hat{F}(x)|$$

5. Kritischer Bereich: $(d_{n;1-\alpha}, \infty)$ wobei

n	3	4	5	6	7	8	9
$d_{n;1-\alpha}$	0.708	0.624	0.563	0.519	0.483	0.454	0.430
n	10	15	20	25	30	40	> 40
$d_{n;1-\alpha}$	0.409	0.338	0.294	0.264	0.242	0.210	$\approx 1.36/\sqrt{n}$

6. -

7. Testentscheidung: Lehne H_0 ab, wenn $t > d_{n;1-\alpha}$ gilt.

5 Übungsaufgaben

1. In einem verregneten Land beträgt die Regenwahrscheinlichkeit in den Herbstmonaten 50%.
Jeden Morgen im Herbst fragt sich Susi, ob sie einen Regenschirm mitnehmen soll oder nicht. Um zu einer Entscheidung zu kommen, wirft sie eine faire Münze. Wirft sie Kopf, nimmt sie einen Regenschirm mit, ansonsten lässt sie den Schirm zu Hause.
 - (a) Betrachten Sie die Situation wie einen statistischen Test. Wie müssen die Hypothesen gewählt werden, damit der Fehler 1. Art die schlimmere Auswirkung darstellt?
 - (b) Bestimmen Sie die Wahrscheinlichkeit für den Fehler 1. Art.
2. Bei der Produktion von Bolzen beträgt der Sollwert für den Durchmesser 10 mm. Der Bolzendurchmesser X sei dabei normalverteilt mit dem unbekanntem Erwartungswert μ und der durch die Konstruktion der Maschine festgelegten Standardabweichung $\sigma = 0.5$ mm. Zur Überprüfung der Einstellung werden 100 Teile entnommen und daraus der mittlere Durchmesser $\bar{x} = 10.15$ mm berechnet. Prüfen Sie die Hypothese $H_0 : \mu = 10$ mm mit der Irrtumswahrscheinlichkeit $\alpha = 0.05$.
3. Zur Überprüfung einer Produktionsanlage, die Kilotüten mit Mehl abfüllt, werden 20 Tüten entnommen, gewogen und daraus das mittlere Gewicht $\bar{x} = 998.92$ g und die Standardabweichung $s = 0.738$ g berechnet. Es kann davon ausgegangen werden, dass das Gewicht normalverteilt ist. Prüfen Sie, ob die Anlage im Mittel wesentlich weniger ($\alpha = 0.05$) als 1000 g abfüllt.
4. Zur Überprüfung einer Produktionsanlage, die Kilotüten mit Mehl abfüllt, werden 20 Tüten entnommen, gewogen und daraus das mittlere Gewicht $\bar{x} = 998.92$ g und die Standardabweichung $s = 0.738$ g berechnet. Es kann davon ausgegangen werden, dass das Gewicht normalverteilt ist. Unterscheidet sich die empirische Streuung s^2 wesentlich ($\alpha = 0.05$) von $\sigma_0^2 = 0.3$ (g^2) ?
5. In einer landwirtschaftlichen Versuchsanstalt erhielten 9 von 22 Masttieren (Gruppe I) Grünfütterzumischungen, während die übrigen 13 Tiere (Gruppe II) ausschließlich mit dem proteinhaltigen Mastfutter gefüttert wurden. Nach einer gewissen Zeit wurden folgende Gewichtszunahmen in kg bei den Tieren festgestellt.
Gruppe I: 7.0, 11.8, 10.1, 8.5, 10.7, 13.2, 9.4, 7.9, 11.1 und Gruppe II: 13.4, 14.6, 10.4, 11.9, 12.7, 16.1, 10.7, 8.3, 13.2, 10.3, 11.3, 12.9, 9.7
Testen Sie die Nullhypothese $\sigma_I^2 = \sigma_{II}^2$ mit Hilfe des F-Tests (für $\alpha = 0.1$).

Lösungen einiger Übungsaufgaben

1. (a) Testproblem: Regen oder kein Regen

Entscheidungsregel: Ich glaube, dass es regnen wird, wenn Münzwurf „Kopf“ zeigt.

Entscheidung / Realität	Regen	kein Regen
„Kopf“ \rightarrow Regen \rightarrow Schirm	trocken	Schirm umsonst
„Zahl“ \rightarrow kein Regen \rightarrow kein Schirm	nass	trocken

Mögliche Fehler: Schirm umsonst mitgenommen bzw. nass werden (ist wohl die schlimmere Konsequenz)

Damit der Fehler 1. Art die schlimmere Konsequenz darstellt, müssen die Hypothesen wie folgt gewählt werden:

H_0 : Regen

H_1 : kein Regen

- (b) $P(H_0 \text{ verworfen} \mid H_0 \text{ stimmt}) = 0.5$

2. $3 = |t| > z_{0.975} = 1.96$, H_0 wird abgelehnt
3. $-6.54 = t < -t_{19;0.95} = -1.73$, H_0 wird abgelehnt
4. $34.49 = t > \chi_{19;0.975} = 32.85$, H_0 wird abgelehnt
5. $t = 0.85 \not< 0.3$ und $\not> 2.85$, H_0 wird nicht abgelehnt