Statistik II für Betriebswirte
Vorlesung 3

Dr. Andreas Wünsche

TU Bergakademie Freiberg
Institut für Stochastik

28. Oktober 2019

4.3.2 Tests für zwei verbundene Stichproben

- Liegen zwei Stichproben vor, deren Werte einander paarweise zugeordnet sind, spricht man von verbundenen Stichproben.
- Diese entstehen z.B. dann, wenn man jeweils zwei Merkmale an ein und demselben statistischen Objekt beobachtet.
- **Beispiele:**
  - Messwerte für die Wirkungen jeweils zweier Medikamente für jeweils ein und denselben Patienten;
  - Wert der Bestellungen einer Kundengruppe vor (1. Stichprobe) und nach (2. Stichprobe) einer Werbeaktion.
- Die mathematische Modellierung erfolgt über zwei endliche Folgen $X_1, X_2, \ldots, X_n$ und $Y_1, Y_2, \ldots, Y_n$ von jeweils n Zufallsgrößen.
- Dabei sind die Zufallsgrößen innerhalb einer Folge stochastisch unabhängig, aber für jedes $i = 1, \ldots, n$ können die Zufallsgrößen X_i und Y_i abhängig sein.
- Eine verbundene (mathematische) Stichprobe wird also durch unabhängige Zufallsvektoren $(X_1, Y_1), \ldots, (X_n, Y_n)$ modelliert.

Erwartungswertvergleich verbundener normalverteilter Stichproben

- **Geg.:** zwei abh. Merkmale $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$; entsprechend verbundene Stichprobe vom Umfang n, d.h. $(X_1, Y_1), \ldots, (X_n, Y_n)$.
- **Grundlage:** Ist der Zufallsvektor (X, Y) normalverteilt, dann ist die Differenz $D = X - Y$ normalverteilt mit dem Erwartungswert

$$\mu_D = \mathbf{E}D = \mathbf{E}X - \mathbf{E}Y = \mu_X - \mu_Y .$$

- Ein Test zum Beispiel mit

$$H_0 : \mu_X = \mu_Y , \qquad H_A : \mu_X \neq \mu_Y ,$$

kann dann als Test

$$H_0 : \mu_D = 0 , \qquad H_A : \mu_D \neq 0 ,$$

für die normalverteilte Stichprobe $D_1, \ldots, D_n$ durchgeführt werden (Einstichproben-t-Test).

Vorzeichentest für verbundene Stichprobe

- **Voraussetzung:** Verbundene Stichprobe $(X_1, Y_1), \ldots, (X_n, Y_n)$, so dass die Differenzzufallsgrößen $D_i = X_i - Y_i$ eine stetige Verteilungsfunktion F_D besitzen.
- Man bildet Differenzen: $D_i = X_i - Y_i, \;\; i = 1, \ldots, n$.
- **Hypothesen:** $H_0 \; : \; D_{0.5} = 0 \, , \;\; H_A \; : \; D_{0.5} \neq 0$ (zweiseitiger Test).
- Man verwendet den Vorzeichentest von oben zur Stichprobe der Differenzen $D_1, \ldots, D_n$ mit Median gleich 0.
- Treten Stichprobenwerte auf, die mit dem Median übereinstimmen, können diese (z.B.) weggelassen werden. Dann bleibt der Test konservativ, d.h. die Irrtumswahrscheinlichkeit 1. Art wird nicht vergrößert.
- Vorzeichentests können auch für nichtstetige Zufallsgrößen durchgeführt werden. Dann testet man z.B. die Hypothese $H_0 \; : \; P(D > 0) = P(D < 0)$.

Beispiel 4.5: Returns von 2 Fonds

- Returns in % von 2 Fonds (Fond A, Fond B) in 15 Monaten.
- **Daten:** (Quelle: Aczel, Sounderpandian: Complete Business Statistics, 2006, S.645)

A:	12	11	14	10	12	8	16	13	12	10	6	9	16	13	10
B:	14	15	16	9	10	8	18	12	17	13	10	12	15	19	14
Sign(A–B):	−	−	−	+	+		−	+	−	−	−	−	+	−	−

- **Aufgabe:** Durchführung des Vorzeichentests zur Untersuchung, ob beide Fonds „gleich“ sind.
- **Hypothesen:** $H_0 \; : \; D_{0.5} = 0 \, , \;\; H_A \; : \; D_{0.5} \neq 0$ mit $D = A - B \, .$
- 4 Monate $A > B$; 1 Monat $A = B$; 10 Monate $A < B$.
- Damit ist $t = 4$.

Beispiel 4.5: Returns von 2 Fonds

- **Kritischer Bereich** für $n' = 14 = 15 - 1$, $\alpha = 0.05$:
  $K = \{0, 1, 2\} \cup \{12, 13, 14\}$.

- $t = 4 \notin K \implies H_0$ wird angenommen $\implies$ Zwischen den beiden Fonds gibt es keine signifikanten Unterschiede.

- in Statgraphics :

Hypothesis Tests for Fond A - Fond B

<u>sign test</u>
Null hypothesis: median = 0
Alternative: not equal

Number of values below hypothesized median: 10
Number of values above hypothesized median: 4

Large sample test statistic = 1,33631 (continuity correction applied)
P-Value = 0,181449
Do not reject the null hypothesis for alpha = 0,05.

Vorzeichen-Rangtest nach Wilcoxon

- Mit dem Vorzeichen-Rangtest (Symmetrie-Test) nach Wilcoxon (oder Wilcoxon-Vorzeichen-Rangtest) kann man die Verteilung einer Zufallsgröße auf Symmetrie untersuchen.
- Die Verteilung von X ist symmetrisch zum Punkt M, falls für beliebige $x \in \mathbb{R}$ gilt: $P(X < M - x) = P(X > M + x)$.
  In diesem Fall ist M auch der Median der Zufallsgröße.
- Ist für zwei verbundene Stichproben die Differenzzufallsgröße symmetrisch zum Median verteilt, kann der Test für stetige Differenzzufallsgrößen wie der einfache Vorzeichentest für verbundene Stichproben durchgeführt werden, d.h. z.B.
  $H_0 : D_{0.5} = 0, \; H_A : D_{0.5} \neq 0$.
- Die Güte dieses Tests ist besser als die Güte eines entsprechenden einfachen Vorzeichentests, da hier die Größe der Differenzen mit berücksichtigt wird.

Wilcoxon-Vorzeichen-Rangtest: Rangzahlen

- **Vorgehen:** Man ordnet die **Beträge** der Differenzen und vergibt Rangzahlen R_i^+. Dabei erhält der kleinste Wert die Rangzahl 1 und der größte Wert die Rangzahl n. Tritt ein Wert mehrfach auf, erhalten (z.B.) alle diese Werte den arithmetischen Mittelwert der zugehörigen Ordnungsnummern als Rangzahl.
- Außerdem verwendet man die Indikatorgrößen

$$Z_i := \begin{cases} 1\,, & \text{falls } D_i > 0\,, \\ 0\,, & \text{falls } D_i < 0\,. \end{cases}$$

Beispiel 4.6: Bestellungen vor und nach einer Werbeaktion

- Wert von Bestellungen vor und nach einer Werbeaktion:

vor (X_i)	171.2	332.9	230.6	200.7	238.7
nach (Y_i)	238.7	260.1	203.1	133.2	171.2
Differenz ($D_i = X_i - Y_i$)	-67.5	72.8	27.5	67.5	67.5
Beträge Differenz	67.5	72.8	27.5	67.5	67.5
Ordnungsnummern	2-4	5	1	2-4	2-4
Rangzahlen (R_i^+)	3	5	1	3	3
Indikatorgrößen Z_i	0	1	1	1	1

Wilcoxon-Vorzeichen-Rangtest: Testgröße, krit. Bereich

- **Testgröße:** $T = W_n^+ := \sum_{i=1}^{n} R_i^+ \cdot Z_i$.
- Es gilt $\sum_{i=1}^{n} R_i^+ = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.
- **Kritischer Bereich:** $K = \{t \,:\, t \leq w_{n;\alpha/2}^+\} \cup \{t \,:\, t \geq w_{n,1-\alpha/2}^+\}$.
- Die Quantile $w_{n;\alpha/2}^+$ und $w_{n,1-\alpha/2}^+$ kann man aus Tabellen ablesen (z.B. im Anhang der Formelsammlung).
- Es gilt $w_{n,1-\alpha/2}^+ = \frac{n(n+1)}{2} - w_{n;\alpha/2}^+$.
- Für große n ($n \geq 20$) gilt: $T = \dfrac{W_n^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$

  ist näherungsweise standardnormalverteilt $\Rightarrow$ Quantile der Standardnormalverteilung können genutzt werden.

Beispiel 4.6: Bestellungen vor und nach einer Werbeaktion

- Wert von Bestellungen vor und nach einer Werbeaktion:

vor (X_i)	171.2	332.9	230.6	200.7	238.7
nach (Y_i)	238.7	260.1	203.1	133.2	171.2
Differenz ($D_i = X_i - Y_i$)	-67.5	72.8	27.5	67.5	67.5
Beträge Differenz	67.5	72.8	27.5	67.5	67.5
Rangzahlen (R_i^+)	3	5	1	3	3
Indikatorgrößen Z_i	0	1	1	1	1

- $H_0 \, : \, D_{0.5} = 0 \, , \quad H_A \, : \, D_{0.5} \neq 0 \, ; \quad \alpha = 0.1 \, .$
- Aus der Tabelle: $w^+_{5;0.95} = \frac{5 \cdot 6}{2} - w^+_{5;0.05} = 15 - 0 = 15$
  $\Rightarrow \; K = \{0\} \cup \{15\} \, .$
- $W_5^+ = 3 + 1 + 5 + 3 = 12 \notin K \;\; \Rightarrow \;\; H_0$ wird nicht abgelehnt, die Unterschiede sind nicht signifikant.

4.4 Verteilungstests

- Eine weitere Klasse von Tests beschäftigt sich mit der Prüfung, ob die Werte der Stichprobe aus einer Grundgesamtheit mit einer speziellen hypothetischen Verteilungsfunktion stammen.
- Ausführlicher wird hier der χ^2-Anpassungstest behandelt.
- Kurz vorgestellt werden auch der Kolmogorow-Smirnow-Test (auch Kolmogorow-Anpassungstest genannt) und der Shapiro-Wilk-Test.
- Weitere Tests, die zum Teil für ganz bestimmte Typen von Verteilungsfunktionen entwickelt wurden, kann man in der Literatur finden.

Der χ^2–Anpassungstest

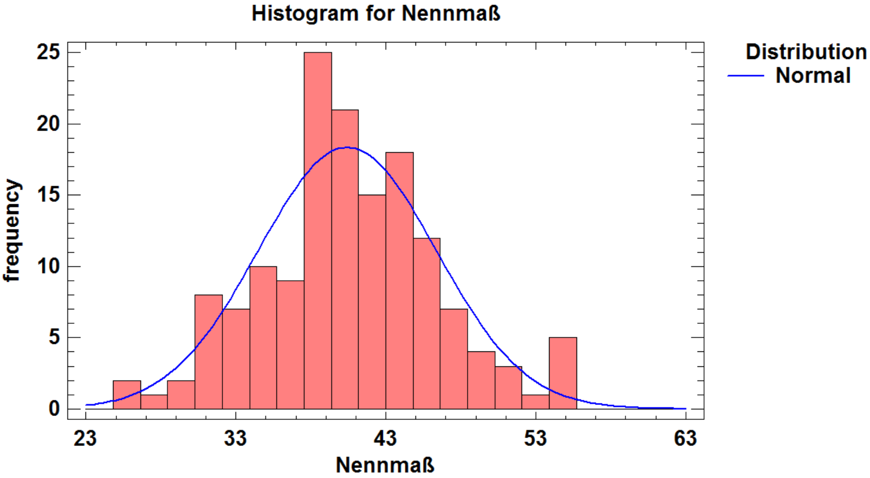
- Test, ob die vorliegende Stichprobe aus einer Grundgesamtheit mit einer hypothetischen Verteilungsfunktion F_0 entstammt.
- **Prinzipielles Vorgehen:**
  - Klasseneinteilung der Stichprobe;
  - Vergleich mit der hypothetischen Verteilung;
  - falls die Abweichungen zu groß sind, erfolgt eine Ablehnung der Nullhypothese.
- Dieser Test ist ein asymptotischer Test, d.h. man rechnet mit der asymptotischen Verteilung (für $n \to \infty$) der Testgröße unter H_0.
- **Hypothesen:**

$H_0 \; : \; F(x) = F_0(x)\,, \; x \in \mathbb{R}\,,$ F_0 ist eine Verteilungsfunktion,

z.B. $F_0(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ falls $X \sim \mathrm{N}(\mu, \sigma^2)\,;$

$H_A \; : \; F(x) \neq F_0(x)$ für mindestens ein $x \in \mathbb{R}\,.$

χ^2-Anpassungstest Grafik



χ^2-Anpassungstest – Testgröße T

- Einteilung der gesamten Merkmalsachse in k Klassen $A_1 = (-\infty, a_1), A_2 = [a_1, a_2), \ldots, A_k = [a_{k-1}, \infty)$.
- Bestimmung der absoluten Klassenhäufigkeiten $H_1, H_2, \ldots, H_k$ (Anzahl der Stichprobenwerte in der jeweiligen Klasse).
- Bestimmung der theoretischen Wahrscheinlichkeiten für die Klassenzugehörigkeiten unter der Annahme der Gültigkeit von H_0,

$$
\begin{aligned}
p_1 &= P_{H_0}(A_1) = P_{H_0}(X < a_1) = F_0(a_1)\,, \\
p_2 &= P_{H_0}(A_2) = P_{H_0}(a_1 \le X < a_2) = F_0(a_2) - F_0(a_1)\,, \\
&\ldots \\
p_k &= P_{H_0}(A_k) = P_{H_0}(a_{k-1} \le X) = 1 - F_0(a_{k-1}).
\end{aligned}
$$

- **Testgröße:** $T = \sum_{j=1}^{k} \frac{(H_j - np_j)^2}{np_j}$ („χ^2-Abstandsfunktion“),

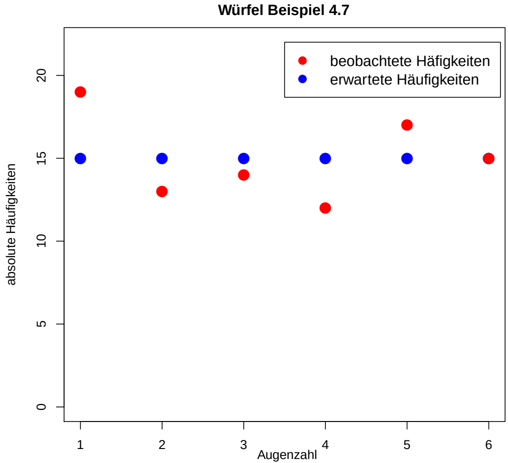
diese Größe ist unter H_0 asymptotisch $\chi^2_{k-1}-$verteilt.

χ^2-Anpassungstest – Kritischer Bereich

- **Kritischer Bereich:** $K = \{t \in \mathbb{R} : \ t > \chi^2_{k-1;1-\alpha}\}$.
- **Bemerkungen:**
  - Der Stichprobenumfang n sollte nicht zu klein sein.
  - Die Anzahl und die Größe der Klassen A_j sollten so sein, dass $np_j = nP_{H_0}(X \in A_j) > 1$ für alle $j = 1, \ldots, k$ gilt (und zusätzlich $np_j \geq 5$ für mindestens 80% der Klassen; ggf. Klassen zusammenfassen oder gesamte Klasseneinteilung ändern).
  - Bei diskreten Verteilungen und nicht zu kleinen Einzelwahrscheinlichkeiten sollte pro Merkmalswert jeweils eine Klasse gewählt werden.
- **Modifikation:**
  Unbekannte Parameter in der Verteilungsfunktion F_0 können durch (Maximum-Likelihood-)Schätzungen ersetzt werden.
  Sind m Parameter zu schätzen, so ist anstelle der $\chi^2_{k-1}-$Verteilung die $\chi^2_{k-m-1}-$Verteilung zu benutzen.

Beispiel 4.7: Test auf gerechten Würfel

Anhand einer Stichprobe von $n = 90$ Würfelergebnissen soll mit $\alpha = 0.05$ getestet werden, ob der Würfel gerecht ist, d.h. ob für die Augenzahl X gilt: $H_0 : p_i = P(X = i) = 1/6\,, \quad i = 1, \ldots, 6\,.$



Beispiel 4.7: Test auf gerechten Würfel

- **Daten:**

Augenzahl	1	2	3	4	5	6
H_j	19	13	14	12	17	15
np_j	15	15	15	15	15	15

- **Wert der Testgröße:**

$$t = \frac{(19-15)^2}{15} + \frac{(13-15)^2}{15} + \frac{(14-15)^2}{15} + \frac{(12-15)^2}{15} + \frac{(17-15)^2}{15}$$
$$= \frac{16}{15} + \frac{4}{15} + \frac{1}{15} + \frac{9}{15} + \frac{4}{15} + \frac{0}{15} = \frac{34}{15} = 2.2\overline{6}\,.$$

- **Kritischer Bereich:** $K = (\chi^2_{5;0.95}\,, \infty) = (11.07\,, \infty)\,.$
- **Testentscheidung:** $t \notin K$, H_0 wird nicht abgelehnt.
- **Testergebnis:** Die Abweichungen der beobachteten Häufigkeiten von einer diskreten Gleichverteilung sind nicht signifikant.

Beispiel 4.7: Statgraphics

Goodness-of-Fit Tests for Würfel

Chi-Square Test

	Lower	*Upper*	*Observed*	*Expected*	
	Limit	*Limit*	*Frequency*	*Frequency*	*Chi-Square*
at or below		1,0	19	15,00	1,07
	2,0	2,0	13	15,00	0,27
	3,0	3,0	14	15,00	0,07
	4,0	4,0	12	15,00	0,60
	5,0	5,0	17	15,00	0,27
	6,0		15	15,00	0,00

Chi-Square = 2,26667 with 3 d.f.   P-Value = 0,518933

- **Testentscheidung:** $p = 0.519 > 0.05 = \alpha$, H_0 wird nicht abgelehnt.
- **Testergebnis:** Die Abweichungen der beobachteten Häufigkeiten von einer diskreten Gleichverteilung sind nicht signifikant.

Beispiel 4.8: technisches Nennmaß

- $X \ldots$ technisches Nennmaß;
  $X \sim \mathrm{N}(\mu, \sigma^2)$ wird getestet.
- $H_0 \; : \; F(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right), \; x \in \mathbb{R};$
  $H_A \; : \; F(x) \neq \Phi\left(\dfrac{x-\mu}{\sigma}\right)$ für mindestens ein $x \in \mathbb{R}$.
- $\alpha = 0.05$.
- $n = 150$ Messungen,
  $\hat{\mu} = \overline{x} = 40.43\,, \quad \hat{\sigma} = s = 5.93 \quad \Rightarrow \quad m = 2$ Schätzparameter.
  Sei $k = 8$ (Anzahl der Klassen); $z_j = \dfrac{a_j - \overline{x}}{s}, \; j = 1, \ldots, k$.

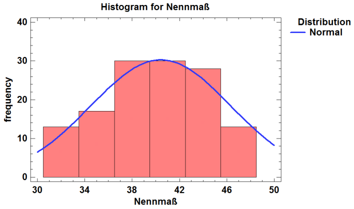
Beispiel 4.8: technisches Nennmaß – Daten und Test

a_j	H_j	z_j	$\Phi(z_j)$	p_j	np_j
…30.5	6	-1.67	0.0475	0.0475	7.12
30.5…33.5	13	-1.17	0.1210	0.0735	11.03
33.5…36.5	17	-0.66	0.2546	0.1336	20.04
36.5…39.5	30	-0.16	0.4364	0.1818	27.27
39.5…42.5	30	0.35	0.6368	0.2004	30.06
42.5…45.5	28	0.85	0.8023	0.1655	24.83
45.5…48.5	13	1.36	0.9131	0.1108	16.61
48.5…	13			0.0869	13.04
	150			1.0000	

- $$t = \sum_{j=1}^{k} \frac{(H_j - np_j)^2}{np_j} = 2.45\,, \quad K = (\chi^2_{8-2-1;0.95} = 11.1, \infty)\,,$$

  $t \notin K$, H_0 wird nicht abgelehnt, zum Niveau von 5% sind die Abweichungen zur Normalverteilung nicht signifikant.

Beispiel 4.8: Statgraphics



Goodness-of-Fit Tests for Nennmaß

Chi-Square Test

	Lower Limit	*Upper Limit*	*Observed Frequency*	*Expected Frequency*	*Chi-Square*
at or below		30,5	6	7,05	0,16
	30,5	33,5	13	11,14	0,31
	33,5	36,5	17	19,88	0,42
	36,5	39,5	30	27,61	0,21
	39,5	42,5	30	29,83	0,00
	42,5	45,5	28	25,08	0,34
	45,5	48,5	13	16,41	0,71
above	48,5		13	12,99	0,00

Chi-Square = 2,14148 with 5 d.f. P-Value = 0,829243

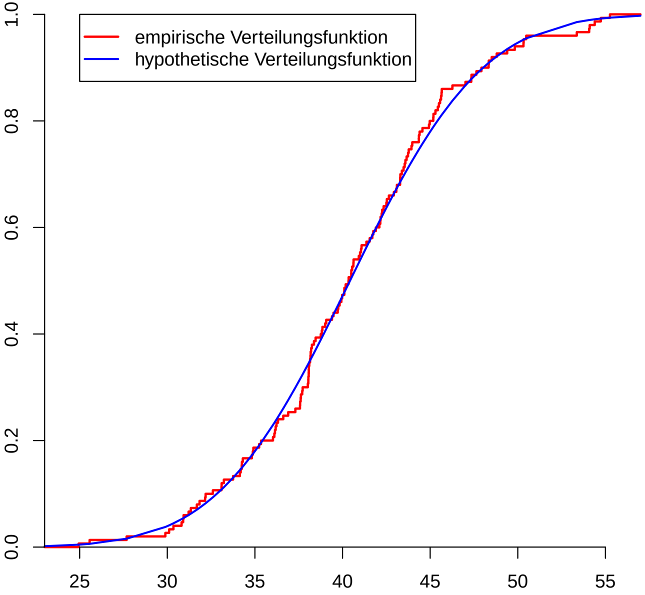
Der Kolmogorow-Smirnow-Test

- Der Kolmogorow-Smirnow-Test (Kolmogorow-Anpassungstest) basiert auf der empirischen Verteilungsfunktion $\hat{F}_n$ zur Stichprobe (vom Umfang n):

$$\hat{F}_n(x) := \frac{\text{Anzahl Stichprobenwerte } < x}{n}, \qquad x \in \mathbb{R}.$$

- **Voraussetzung:** Die hypothetische Verteilungsfunktion F_0 ist stetig und enthält keine unbekannten Parameter.
- **Hypothesen:**
  $H_0 : F(x) = F_0(x),\ x \in \mathbb{R};$
  $H_A : F(x) \neq F_0(x)$ für mindestens ein $x \in \mathbb{R}$.
- **Testgröße:** $T = \sup_{x \in \mathbb{R}} \left|\hat{F}_n(x) - F_0(x)\right|.$
- Der Test wird günstigerweise mit einem Computerprogramm durchgeführt.

Kolmogorow-Smirnow-Test im Beispiel 4.8:



Beispiel 4.8: Statgraphics

Goodness-of-Fit Tests for Nennmaß

Kolmogorov-Smirnov Test

	Normal
DPLUS	0,0485841
DMINUS	0,0550786
DN	0,0550786
P-Value	0,753039

$p = 0.753039 > 0.05 = \alpha \quad \Longrightarrow \quad H_0$ wird nicht abgelehnt, zum Niveau von 5% sind die Abweichungen zur Normalverteilung nicht signifikant.

Bemerkungen zum Kolmogorow-Smirnow-Test

- Der Kolmogorow-Smirnow-Test (kurz auch „K-S-Test“) ist im Gegensatz zum χ^2-Anpassungstest auch für kleine Stichproben anwendbar und das Testergebnis hängt nicht von einer Klasseneinteilung ab.
- Man kann einseitige Tests mit dem K-S-Test durchführen.
- Es gibt Verallgemeinerungen des K-S-Tests, bei denen statt festgelegter Parameterwerte der hypothetischen Verteilung F_0 geeignete Schätzwerte eingesetzt werden (z.B. der Lilliefors-Test im Fall von Normalverteilungen).
- Man kann mit einer Version des K-S-Testes auch prüfen, ob zwei Stichproben aus einer Grundgesamtheit stammen, also übereinstimmende Verteilungen zugrundeliegen.
- Der K-S-Test kann auch für diskrete Verteilungen genutzt werden, besitzt dann aber eine geringere Güte.

Der Shapiro-Wilk-Test zur Normalverteilungsprüfung

- Der Shapiro-Wilk-Test prüft ausschließlich, ob bei einer Stichprobe eine Normalverteilung vorliegt.
- Dieser Test besitzt eine hohe Güte, insbesondere auch im Fall von kleinen Stichprobenumfängen.
- Grundlage des Tests sind bestimmte Eigenschaften der Ordnungsstatistiken einer Stichprobe aus einer normalverteilten Grundgesamtheit.
- Der Test ist sehr rechenintensiv und sollte mit Statistik-Software durchgeführt werden.
- Statgraphics-Ergebnis in Beispiel 4.8:

Tests for Normality for Nennmaß

Test	*Statistic*	*P-Value*
Shapiro-Wilk W	0,979963	0,4061

4.5 χ^2–Unabhängigkeitstest; Homogenitätstest

- Mit dem χ^2–Unabhängigkeitstest überprüft man, ob zwei Merkmale X und Y stochastisch unabhängig sind, d.h. ob für beliebige (zulässige) Mengen A, B gilt:

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B).$$

- Konkreter prüft man, ob die relativen Häufigkeiten (berechnet aus einer verbundenen Stichprobe) näherungsweise diese Produktregel erfüllen.
- **Hypothesen:** H_0 : X und Y sind stochastisch unabhängig, H_A : X und Y sind abhängig.
- **Verbundene Stichproben:** $x_1, \quad x_2, \quad \ldots, \quad x_n;$
  $y_1, \quad y_2, \quad \ldots, \quad y_n.$
- **Einteilung der Merkmalsachsen in Klassen:**
  für X: $A_1, \ldots, A_k$; für Y: $B_1, \ldots, B_\ell$.

Kontingenztafel

- **Absolute Häufigkeiten:**
  H_{ij} : Anzahl der Beobachtungen, bei denen das Merkmal X in der Klasse A_i und gleichzeitig das dazugehörige Merkmal Y in B_j liegt.

- **Randhäufigkeiten:** $H_{i\bullet} = \sum_{j=1}^{\ell} H_{ij}\,, \quad H_{\bullet j} = \sum_{i=1}^{k} H_{ij}\,.$

- **Kontingenztafel:**

$X\backslash Y$	B_1	$\dots$	B_ℓ	
A_1	H_{11}		$H_{1\ell}$	$H_{1\bullet}$
$\vdots$				
A_k	H_{k1}		$H_{k\ell}$	$H_{k\bullet}$
	$H_{\bullet 1}$		$H_{\bullet\ell}$	n

- Im Fall $k = \ell = 2$ wird eine solche Tafel auch Vierfeldertafel oder $2 \times 2-$Felder-Tafel genannt. Sie wird häufig bei nominellen Merkmalen mit zwei Ausprägungen (dichotome Merkmale) benutzt.

χ^2–Unabhängigkeitstest – Testgröße, kritischer Bereich

- **Testgröße:** („empirische Kontingenz“)

$$T = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{\left(H_{ij} - \frac{H_{i\bullet} H_{\bullet j}}{n}\right)^2}{\frac{H_{i\bullet} H_{\bullet j}}{n}} .$$

- Für eine Vierfeldertafel kann die Testgröße einfacher berechnet werden durch

$$T = \frac{n\left(H_{11} H_{22} - H_{12} H_{21}\right)^2}{H_{1\bullet} H_{2\bullet} H_{\bullet 1} H_{\bullet 2}} .$$

- **Kritischer Bereich:** $K = \left\{t \in \mathbb{R} : \ t > \chi^2_{(k-1)(\ell-1);1-\alpha}\right\}$.

χ^2-Unabhängigkeitstest – Bemerkungen

- Von den Klassenhäufigkeiten sollten höchstens 20% kleiner als 5 sein, aber alle mindestens gleich 1.
- Der χ^2-Unabhängigkeitstest mit Hilfe einer Vierfeldertafel sollte für Stichprobenumfänge $n < 20$ nicht verwendet werden (sondern der „exakte Test von Fisher“).
- Für $20 \leq n \leq 60$ eignet sich die nach Yates korrigierte Teststatistik

$$T = \frac{n\left(|H_{11}H_{22} - H_{12}H_{21}| - \frac{n}{2}\right)^2}{H_{1\bullet}H_{2\bullet}H_{\bullet 1}H_{\bullet 2}} \,.$$

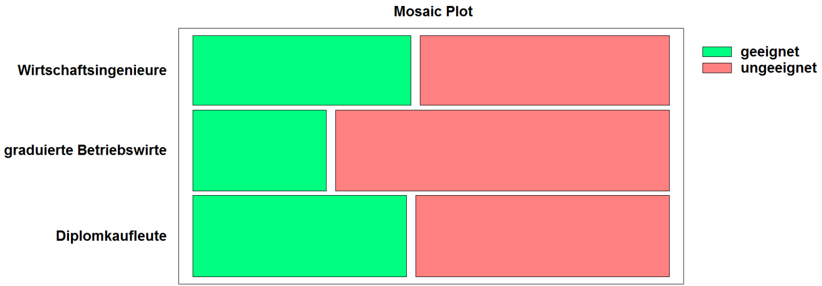
Beispiel 4.9: Eignung versus Studienabschluss

- 30 Wirtschaftsingenieure (b_1), 35 graduierte Betriebswirte (b_2) und 35 Diplomkaufleute (b_3), die sich bei einem Unternehmen beworben haben, werden nach einer Eignungsprüfung in die Kategorien „geeignet“ (a_1) und „ungeeignet“ (a_2) eingeordnet. Ist diese Eignung vom Studienabschluss abhängig oder nicht:
- Merkmal X (Eignung), Merkmal Y (Studienabschluss); $\alpha = 0.05$.
- **Hypothesen:**
  H_0 : Merkmal X und Merkmal Y sind unabhängig,
  H_A : Merkmal X und Merkmal Y sind abhängig.

Beispiel 4.9: Eignung versus Studienabschluss

- **Kontingenztafel:** (Quelle: Bleymüller, Gehlert, Gülicher: Statistik für Wirtschaftswissenschaftler, 2004, Abschn. 19.2.)

$Y \backslash X$	a_1	a_2	
b_1	14	16	30
b_2	10	25	35
b_3	16	19	35
	40	60	100



Beispiel 4.9: Testgröße, Testentscheidung

- **Kontingenztafel:**

$X \backslash Y$	b_1	b_2	b_3	
a_1	14	10	16	40
a_2	16	25	19	60
	30	35	35	100

- **Berechnung Testgröße:** $\dfrac{H_{i\bullet}H_{\bullet j}}{n}$: $\boxed{\begin{array}{ccc} 12 & 14 & 14 \\ 18 & 21 & 21 \end{array}}$

$$t = \frac{(14-12)^2}{12} + \frac{(10-14)^2}{14} + \frac{(16-14)^2}{14} + \frac{(16-18)^2}{18} + \frac{(25-21)^2}{21} + \frac{(19-21)^2}{21} = 2.937 < \chi^2_{2;0.95} = 5.99\,.$$

- Die Hypothese H_0 : „A und B sind unabhängig“ wird nicht abgelehnt, man kann nicht davon ausgehen, dass die Eignung signifikant vom Studienabschluss abhängt.

Beispiel 4.9: Statgraphics

Frequency Table

	geeignet	ungeeignet	Row Total
Wirtschaftsingenieure	14	16	30
	12,00	18,00	30,00%
graduierte Betriebswirte	10	25	35
	14,00	21,00	35,00%
Diplomkaufleute	16	19	35
	14,00	21,00	35,00%
Column Total	40	60	100
	40,00%	60,00%	100,00%

Cell contents:

Observed frequency

Expected frequency

Tests of Independence

Test	*Statistic*	*Df*	*P-Value*
Chi-Square	2,937	2	0,2303