

## 28. Nichtparametrische statistische Tests

*Man darf nicht das, was uns unwahrscheinlich und unnatürlich erscheint, mit dem verwechseln, was absolut unmöglich ist.*

(Carl Friedrich Gauß)

Die bisher betrachteten **parametrischen** Tests dienen zum Überprüfen von Behauptungen bezüglich Parameterwerten spezifischer Verteilungen. Die **nichtparametrischen Tests** dieses Kapitels überprüfen weitergehende Behauptungen, die sich auf die Verteilungsfunktion *als solche* beziehen. Die wichtigsten Vertreter nichtparametrischer Tests sind

- **Anpassungstests**: Überprüfen, ob eine Stichprobe mit einer gewissen Verteilung als Nullhypothese (z.B. Normalverteilung, Exponentialverteilung, ...) kompatibel ist. Die wichtigsten Anpassungstests sind der  **$\chi^2$ -Test** und der **Kolmogorow-Smirnov-Test**.
- **Unabhängigkeitstests**: Überprüfen, ob zwei Zufallsvariablen  $X$  und  $Y$  voneinander unabhängig sind, d.h., ob das aus der deskriptiven Statistik bekannte Kriterium  $f(x_i, y_j) = f(x_i)f(y_j)$  als Nullhypothese angenommen werden kann.
- **Homogenitätstests**: Überprüfen, ob zwei Stichproben  $X_i$  und  $Y_j$  derselben Verteilung gehorchen können.

## 29. Chi-Quadrat-Anpassungstest

Der  $\chi^2$ -Test ist anwendbar für diskrete Zufallsvariablen oder für klassierte Daten stetiger Zufallsvariablen. Dabei vergleicht man die tatsächlichen absoluten Häufigkeiten  $h_k$  der Stichprobe mit den gemäß der Nullhypothese erwarteten, *i.A. reellwertigen theoretischen Häufigkeiten*  $h_k^e$ . Die 4 Standardschritte der Tests werden durch einen Vorbereitungsschritt ergänzt:

0. **Vorbereitung:** Klassierung bzw. Umklassierung so, dass  $h_k^e \geq 5$  für alle Klassen  $k = 1, \dots, K$ . (Nach anderen Quellen  $h_k^e \geq 1$ .)
1. **Nullhypothese:** Die Zufallsgröße  $X$  gehorcht eine bestimmten Verteilungsfunktion  $F^{(0)}(x)$ . Dabei wird meist die *Form* der Verteilung (z.B. Gauß, Exponential, Poisson, ...) vorgegeben und aus der Stichprobe deren  $r$  Parameter abgeschätzt.
2. **Testvariable:**

$$Q = \sum_{k=1}^K \frac{(h_k - h_k^e)^2}{h_k^e} \sim \chi^2(K - 1 - r)$$

- $Q$  kann als mit  $h_k^e = n \left( F^{(0)}(x_k^o) - F^{(0)}(x_k^u) \right)$  gewichtete *Quadratsumme der relativen Abweichungen* zwischen den nach  $H_0$  erwarteten absoluten Häufigkeiten und denen der Stichprobe aufgefasst werden.
  - $Q \sim \chi^2(K - 1 - r)$ : Ein Freiheitsgrad geht wegen der Normierungsbedingung verloren; weitere  $r$ , wenn die Parameter aus der selben Stichprobe bestimmt werden.
3. **Realisierung**  $q$  von  $Q$  aus den klassierten Daten.
  4. **Entscheidung** mit Irrtumswahrscheinlichkeit  $\alpha$ :

Nullhypothese verworfen, falls  $q > q_{1-\alpha}^{(K-1-r)}$

## 29.2 Bemerkungen und Hinweise

- Zum “Merken” kann die Testvariable auch in Worten formuliert werden:

$$Q = \sum_{k=1}^K \left( \frac{(\text{empirische} - \text{theoretische abs. Häufigkeit})^2}{\text{theoretische abs. Häufigkeit}} \right)$$

*Vorteil:* Diese allgemeine Formulierung ist auch für den Unabhängigkeitstest (Kap. 31) und den Homogenitätstest (Kap. 32) gültig.

- Zum konkreten Ausrechnen kann man die Testvariable auch in der Form

$$Q = \left( \sum_{k=1}^K \frac{h_k^2}{h_k^e} \right) - n$$

schreiben mit dem Stichprobenumfang  $n = \sum_k h_k$ . Ausrechnen von  $Q$  in dieser Form benötigt i.A. weniger Rechenaufwand.

*Aufgabe:* Leiten Sie diese Formel her!

- Falls für eine Klasse die Bedingung  $h_k^e \geq 5$  nicht erfüllt ist, fasst man sie mit Nachbarklassen so lange zusammen, bis die zusammengelegte Klasse die Bedingung erfüllt.
- Falls der Stichprobenumfang so klein ist, dass nach dem Zusammenfassen  $K - 1 - r \leq 0$ , so ist der  $\chi^2$ -Test nicht möglich. Man kann aber in diesem Fall den später behandelten [Kolmogorow-Smirnow-Test](#) durchführen.

## 29.3 Beispiel: Klausurergebnisse gaußverteilt?

Bei der Teilprüfung zur Vorlesung “Operations Research und Logistik” im SS 2001 wurden folgende Punktzahlen erzielt (Umfang  $n = 45$ ):

36 38 50 33 18 46 35 45 22 30 17 42  
 46 31 41 06 31 46 18 47 12 39 39 23  
 26 27 42 34 31 27 29 33 35 50 48 40  
 41 36 43 37 41 45 34 32 36

Ist dies kompatibel mit einer Gaußverteilung? wenn ja, welchen Mittelwert und welche Varianz hat sie?

0. **Vorbereitung: Histogrammeinteilung** In der Praxis hat sich für die Intervallzahl des Histogramms bei nicht zu großen  $n$  die grobe Orientierung an  $K \approx \sqrt{n}$  bewährt, wobei Klassen solange zusammengefasst werden, bis für alle Klassen  $h_k^e \geq 5$ . Da dies allerdings nur eine “Faustformel” ist, wäre auch eine Klasse mit z.B.  $h_k^e = 4$  kein “Beinbruch”. Als “ersten Versuch” wählen wir z.B. Fünfer-Intervalle, die um  $x_1^* = 5, x_2^* = 10, \dots$  Punkte zentriert sind und erhalten damit

P	2.5- 7.5	7.5- 12.5	12.5- 17.5	17.5- 22.5	22.5- 27.5	27.5- 32.5	32.5- 37.5	37.5- 42.5	42.5- 47.5	> 47.5
$h_k$	1	1	1	3	4	6	10	9	7	3

## 29.3(b) Klausurergebnisse gaußverteilt? (Forts.)

Da die theoretischen Häufigkeiten zunächst noch nicht bekannt sind, orientieren wir uns bei der Zusammenfassung der Klassen an den empirischen Häufigkeiten und fassen alle Klassen so zusammen, dass die Klassenhäufigkeit der Stichprobe mindestens 3 beträgt (Der in der Literatur angegebene Grenzwert ist zwischen 1 und 5). Konkret fassen wir die untersten vier Klassen zusammen:

P	$\leq$ 22.5	22.5- 27.5	27.5- 32.5	32.5- 37.5	37.5- 42.5	42.5- 47.5	$>$ 47.5
$h_k$	6	4	6	10	9	7	3

- Nullhypothese:** Die Punkteverteilung gehorcht einer  $(\mu, \sigma^2)$ -Gaußverteilung mit aus den Daten bestimmten Parametern (Umfang  $n = 45$ )

$$\mu \stackrel{!}{=} \bar{X} = \frac{1}{n} \sum_{k=1}^{10} x_k^* h_k = 34.3,$$

$$\sigma^2 \stackrel{!}{=} S^2 = \frac{1}{n-1} \left( \sum_{k=1}^{10} (x_k^*)^2 h_k - n \bar{X}^2 \right) = 107.6$$

Hinweis: Auch wenn Klassen zusammengefasst wurden, werden die Parameter *aus den ursprünglichen Klassen* oder, noch besser, aus der Urliste (falls vorhanden), bestimmt. In dieser Aufgabe wurden die ursprünglichen Klassen verwendet (mit  $x_1^* = 5$  und  $x_{10}^* = 50$ ). Würde man die zusammengefassten Klassen nehmen, ergäbe sich bei  $x_1^* = 15$  die Schätzungen  $\bar{X} = 34.3$  und  $S^2 = 103.5$ , bei  $x_1^* = 20$  hingegen  $\bar{X} = 35$  und  $S^2 = 80.1$ . Der Varianzschätzer hängt also empfindlich vom etwas willkürlichen Wert für  $x_1^*$  ab.

## 29.3(c) Klausurergebnisse gaußverteilt? (Forts.)

### 2. Die Testvariable

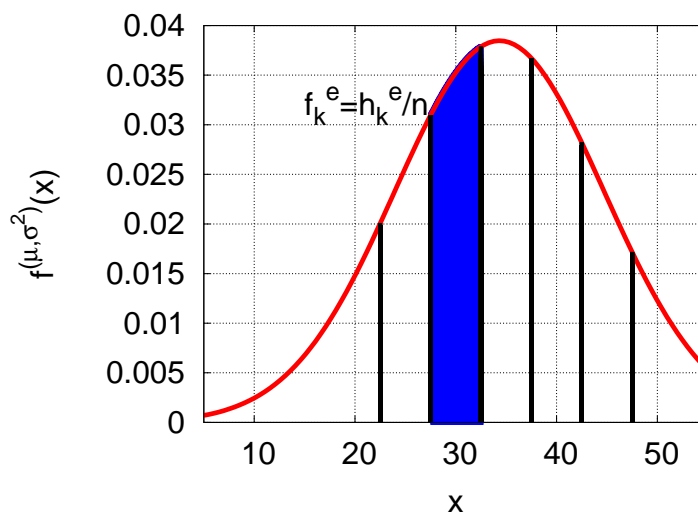
$$Q = \left( \sum_{k=1}^K \frac{h_k^2}{h_k^e} \right) - n \sim \chi^2(K - 1 - r) = \chi^2(4).$$

Berechnung der Zahl der Freiheitsgrade:

- $K = 7$  Klassen (nach der Zusammenfassung)
- $r = 2$  Parameter ( $\mu$  und  $\sigma^2$  wurden durch  $\bar{x}$  und  $s^2$  aus der Stichprobe abgeschätzt).
- 1 Freiheitsgrad geht wegen der Bedingung “Summe der Wahrscheinlichkeiten=1” verloren.

3. **Realisierung** der Testvariablen: Dazu muss man zunächst die erwarteten theoretischen Häufigkeiten  $h_k^e$  aus den relativen Häufigkeiten  $f_k^e$  der zur Nullhypothese gehörigen Verteilungsfunktion  $F(x)$  berechnen (die blaue Fläche gibt z.B.  $f_3^e$  an):

$$h_k^e = n f_k^e = [F(x_k^o) - F(x_k^u)]$$



### 29.3 (d) Klausurergebnisse gaußverteilt? (Forts.)

Hier ist  $F(x) = N^{(\mu, \sigma^2)}(x)$  und man muss, um Tabellen verwenden zu können, durch die bekannte Transformation  $Z = (X - \mu)/\sigma$  auf die Standardnormalverteilung übergehen <sup>1</sup>:

$$h_k^e = n \left[ \Phi \left( \frac{x_k^o - \mu}{\sigma} \right) - \Phi \left( \frac{x_k^u - \mu}{\sigma} \right) \right]$$

Sinnvollerweise erstellt man folgende Arbeitstabelle:

Klasse	$x_k^o$	$z_k^o = \frac{x_k^o - \mu}{\sigma}$	$\Phi(z_k^o)$	$h_k^e$	$h_k$	$\frac{h_k^2}{h_k^e}$
$\leq 22.5$	22.5	-1.14	0.127	5.71	6	6.30
22.5-27.5	27.5	-0.66	0.255	5.76	4	2.78
27.5-32.5	32.5	-0.18	0.430	7.87	6	4.57
32.5-37.5	37.5	0.31	0.620	8.56	10	11.69
37.5-42.5	42.5	0.79	0.785	7.40	9	10.94
42.5-47.5	47.5	1.27	0.898	5.10	7	9.60
$> 47.5$	$\infty$	$\infty$	1.000	4.59	3	1.96
$\Sigma$				45	45	47.84

Damit ist  $Q = \sum_{k=1}^K \frac{h_k^2}{h_k^e} - n = 47.84 - 45 = \underline{\underline{2.84}}$ .

(Hinweis: Die Tabelle wurde anhand der mit dem Computer ermittelten Werte von  $\Phi$  berechnet. Abweichungen in der letzten Stelle können durch Verwendung der Tabelle entstehen und sind natürlich in Klausuren ohne Bedeutung)

<sup>1</sup>bei kleinen  $n$  und unbekannter Varianz folgt aus der Nullhypothese eine student-verteilte Größe  $T = (X - \bar{X})/S$ , was hier nicht weiter betrachtet wird.

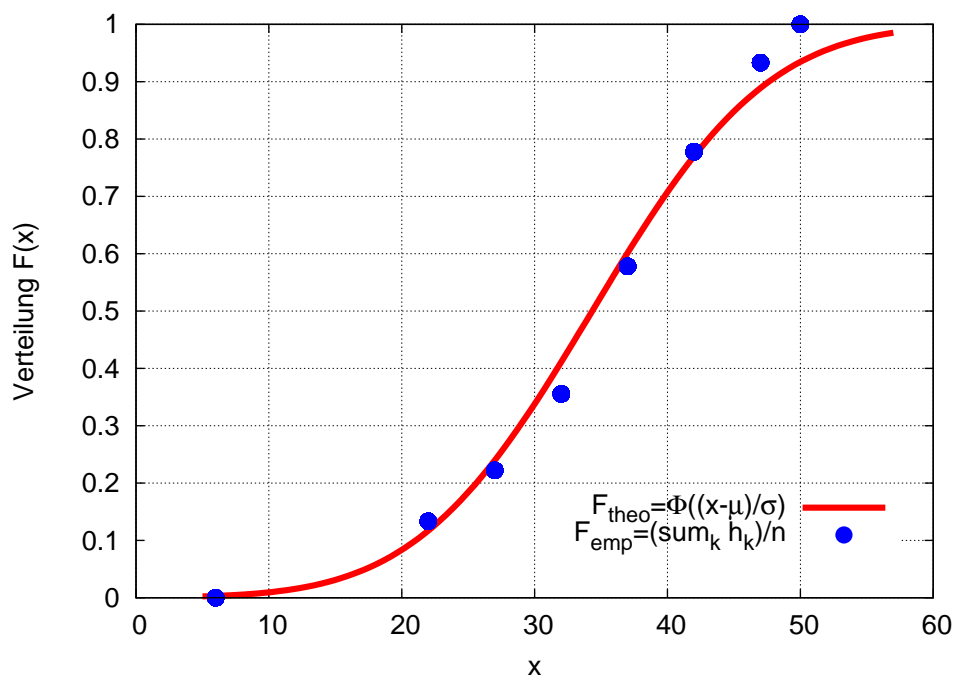
## 29.3 (e) Klausurergebnisse gaußverteilt? (Forts.)

### 4. Entscheidung:

Bei einer Fehlerwahrscheinlichkeit von  $\alpha = 5\%$  gilt "Falls

$$q > q_{1-\alpha}^{(K-1-r)} = q_{0.95}^{(4)} = 9.49,$$

verwirf die Nullhypothese". Dies ist nicht der Fall. Obwohl die Klausurergebnisse "etwas schief" aussehen, kann man die Annahme einer Gaußverteilung mit  $\mu = 34.3$  und  $\sigma^2 = 107.5$  *nicht* verwerfen! Das schließt jedoch nicht aus, dass die Ergebnisse tatsächlich einer anderen Verteilung gehorchen.



## 29.4 Aufgaben

1. Testen Sie die Klausurergebnisse nun auf eine Gleichverteilung, testen Sie also die Nullhypothese  $H_0 : X \sim G(a; b)$  mit den aus den Daten bestimmten Parametern  $a$  und  $b$ .
2. Wäre es ein Fehler bzw. Widerspruch, wenn auch dieser Test "bestanden" wird (die Verteilung kann doch nicht gleichzeitig eine Gauß- und eine Gleichverteilung sein!)?
3. Gegeben sind die relativen Häufigkeiten der ersten Ziffer der in Bytes angegebenen Dateigröße auf meiner Festplatte:

$i$	1	2	3	4	5	6	7	8	9
$f_i$ (%)	27.7	16.4	14.4	11.2	8.4	7.1	5.8	5.1	4.0

Testen Sie ( $n = 12757$ ), ob die relativen Häufigkeiten konsistent mit folgender Nullhypothese sind:

$$f_i^0 = f_i^{\text{Benford}} = \frac{\ln\left(\frac{i+1}{i}\right)}{\ln(10)} = \log_{10}\left(1 + \frac{1}{i}\right)$$

**Hintergrund:** Eine bemerkenswerte Theorie (vgl. Exkurse) besagt, dass unter sehr allgemeinen Bedingungen die Anfangsziffern von Zahlenreihen, welche viele Zehnerpotenzen umfassen (z.B. alle in einer Tageszeitung vorkommende Zahlen!) dem sogenannten **Gesetz von Benford** gehorchen. Dieses Gesetz besagt, dass die relative Häufigkeit der ersten Ziffer  $i$ , welche ungleich 0 ist (z.B. 2 bei der Zahl 257, 4 bei 0.48), gemäß obiger Formel gegeben ist. Sie könne sogar alle Zahlen der Stichprobe mit beliebigen Zahlen multiplizieren - Benford's Gesetz gilt für die multiplizierten Zahlen ebenso. Interessierte können näheres in den [Exkursen](#) zu diesem Kapitel nachschlagen.

## 30. Kolmogorow-Smirnov-Anpassungstest

*Student: "Herr Professor, mein ganzes Wissen habe ich Ihnen zu verdanken". Professor: "Ohh, bitte, erwähnen Sie doch solche Kleinigkeiten nicht"*

Der **Kolmogorow-Smirnov-Anpassungstest** (KS-Test) testet ebenfalls die Nullhypothese

$H_0$ : "Die Zufallsvariable  $X$  gehorcht einer bestimmten Verteilung  $F^{(0)}(x)$ "

Der KS-Test ist mathematisch exakt für stetige Verteilungen und der  $\chi^2$ -Test für diskrete Verteilungen. *In praxi* kann man jedoch beide Tests für beide Arten von Verteilungen anwenden.

- Vorteile:
  - Keinerlei Einschränkung bezüglich Stichprobenumfang
  - Keine Klassierung notwendig
  - Bestechend anschaulich und einfach in der Anwendung (nicht jedoch in der hier nicht behandelten Herleitung!)
- Nachteile:
  - Oft nicht sehr trennscharf: Es gibt Stichproben, die beim Kolmogorow-Test als konsistent mit der Nullhypothese "durchgehen", bei der aber der  $\chi^2$ -Test zur Ablehnung führen würde.

Es gibt allerdings verfeinerte (hier nicht behandelte) Erweiterungen des Kolmogorow-Tests, die trennschärfer sind.

## 30.1 Vorgehen beim Kolmogorow-Smirnov-Test

1. Nullhypothese wie beim  $\chi^2$ -Test: "Die Zufallsvariable  $X$  gehorcht einer bestimmten Verteilung  $F^{(0)}(x)$ " (Man kann auch einseitige Nullhypothesen formulieren, vgl. z.B. die [Statistik-Vorlesung der TU Wien](#)).

2. Die Test-Statistik ist die **D-Statistik**

$$D = \max_x \left| F(x) - F^{(0)}(x) \right| \sim D(n)$$

3. Berechnung einer Realisierung  $d$  von  $D$  aus der Stichprobe, am besten mit grafischer Hilfe.

*Hinweis:* Die Stichproben-Verteilungsfunktion  $F(x)$  ist eine Treppenfunktion bzw. stückweise linear, vgl. die Folien zur deskriptiven Statistik.

4. Test-Entscheidung: Verwerfe  $H_0$  bei einer Fehlerwahrscheinlichkeit  $\alpha$ , falls  $d$  größer als das entsprechende Quantil  $d_{n,1-\alpha}$  der D-Statistik ist. Es gilt näherungsweise:

$$d > d_{n,1-\alpha} \approx \frac{c(\alpha)}{\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}}$$

mit

$\alpha$	0.010	0.025	0.050	0.100
$c(\alpha)$	1.628	1.480	1.358	1.224

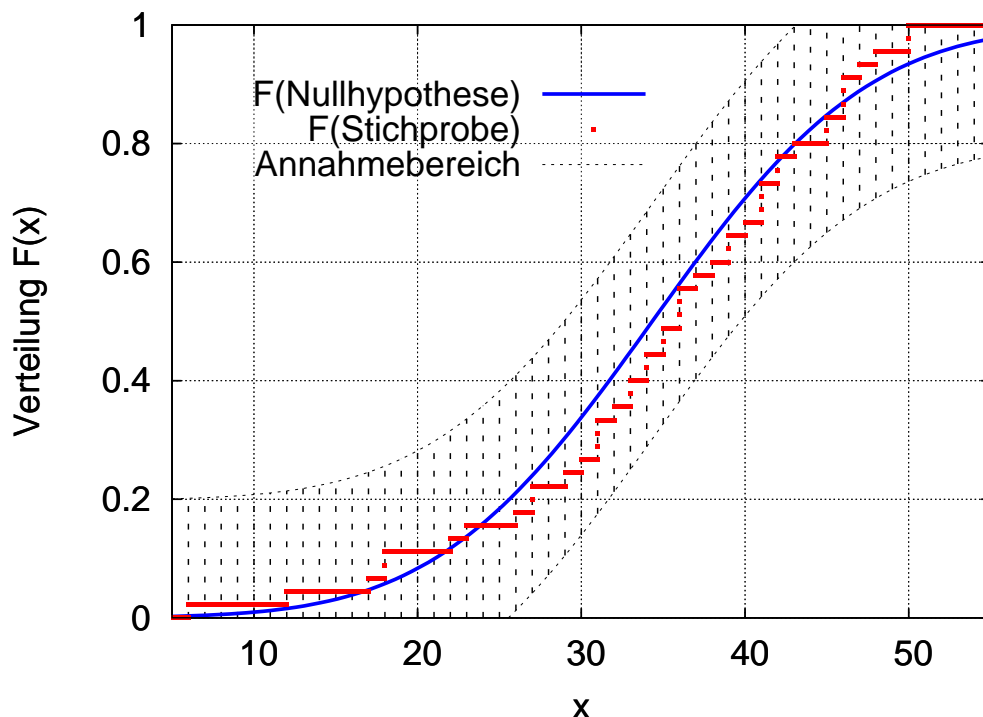
*Hinweis:* Im Gegensatz zum  $\chi^2$ -Test werden keine "Freiheitsgrade" abgezogen.

*Aufgabe:* Testen Sie die [Urliste der Klausurergebnisse](#) mit dem  $\chi^2$ -Test und dem Kolmogorow-Smirnov-Test auf eine (6,50)-Gleichverteilung

## 30.2 Beispiel zum Kolmogorow-Test

Wir testen nun die [Klausurergebnisse](#) vom vorigen Kapitel (mit Ausreißer) mit dem KS-Test auf eine Normalverteilung.

1. **Nullhypothese:**  $X \sim N(\mu, \sigma^2)$  mit  $\mu = 34.3$  und  $\sigma^2 = 107.5$
2. **Test-Statistik:**  $D = \max_{6 \leq x \leq 50} |F(x) - F^{(0)}(x)| \sim D(45)$
3. **Realisierung  $d$  von  $D$  aus der Stichprobe:**



Aus der Abbildung sieht man, dass die größte Abweichung wohl beim Übergang von 30 auf 31 Punkten stattfindet: Dabei springt die Verteilungsfunktion von  $12/45$  auf  $13/45$  und die größte Differenz ist

$$d = \left| F^{(0)}(31) - \frac{12}{45} \right| = \left| \Phi \left( \frac{31 - \mu}{\sigma} \right) - \frac{12}{45} \right| = 0.107$$

Hingegen ist  $d_{n,1-\alpha} = d_{45,0.95} = 0.198$ .  $\Rightarrow H_0$  kann nicht abgelehnt werden, was auch aus der Abbildung klar ist.