Übungsblatt: Hypothesentests II

Mit Lösungen

BSc Wahrscheinlichkeit und Statistik (Module B09 bis B11) · Gruppenarbeit

Auftrag. Bearbeitet die Aufgaben in Gruppen. Rechnet zuerst von Hand (Hypothesen, Teststatistik, Entscheidung gegen $\alpha = 0{,}05$, Interpretation), prüft danach in R. Die Lösungen werden gemeinsam besprochen.

Teil A: Tests auswählen und durchführen

Aufgabe 1: Welcher Test passt?

Ordnet jeder Situation den passenden Test zu und formuliert H_0 und H_1.

(a) Ein Manager schlägt den Index in 14 von 20 Monaten. Trefferquote $> 50\%$?
(b) Mittlere Tagesrendite einer Strategie ($n = 250$): weicht sie von 0 ab?
(c) Zwei *unabhängige* Fonds: unterscheiden sich die mittleren Renditen?
(d) Dieselben 10 Tage, zwei Strategien: hat A einen Vorteil gegenüber B?

(a) Binomialtest. H_0: $p = 0{,}5$ $\quad H_1$: $p > 0{,}5$.
(b) Einstichproben-t-Test. H_0: $\mu = 0$ $\quad H_1$: $\mu \neq 0$.
(c) Zweistichproben-t-Test. H_0: $\mu_A = \mu_B$ $\quad H_1$: $\mu_A \neq \mu_B$.
(d) Paardifferenzentest (gepaart). H_0: $\mu_d = 0$ $\quad H_1$: $\mu_d \neq 0$.

Aufgabe 2: Binomialtest

Ein Manager schlägt die Benchmark in **14 von 20** Monaten. Ist er besser als ein Münzwurf? H_0: $p = 0{,}5$ $\quad H_1$: $p > 0{,}5$ (einseitig). $X = 14$, $n = 20$; p-Wert $= P(X \geq 14 \mid p = 0{,}5) = 0{,}0577 > 0{,}05 \Rightarrow$ H_0 **nicht** verwerfen. Kein signifikanter Skill. *Typischer Fehler:* $p = 0{,}058$ ist nicht "fast signifikant"; α wird vorher festgelegt.

```
binom.test(14, 20, 0.5, alternative = "greater")   # p = 0.0577
```

Aufgabe 3a: Einstichproben-t-Test (grosse Stichprobe)

Tagesrenditen, $n = 250$, $\bar{x} = 0{,}12\%$, $s = 0{,}90\%$. H_0: $\mu = 0$ $\quad H_1$: $\mu \neq 0$. SE $= 0{,}90/\sqrt{250} = 0{,}0569\%$; $t = 0{,}12/0{,}0569 = 2{,}11$ (gross $n \Rightarrow t \approx z$); $|2{,}11| > 1{,}96 \Rightarrow$ verwerfen, $p \approx 0{,}035$. Mittlere Tagesrendite signifikant von 0 verschieden.

```
t.test(renditen, mu = 0)   # t = 2.11, p = 0.035
```

Aufgabe 3b: Einstichproben-t-Test (kleine Stichprobe)

12 Monatsrenditen (%): 0,5 −0,2 1,1 0,8 −0,5 1,4 0,2 0,9 −0,1 1,2 0,6 0,7. Liegt der Mittelwert beim Ziel $\mu_0 = 0{,}3\%$? H_0: $\mu = 0{,}3$ $\quad H_1$: $\mu \neq 0{,}3$. $\bar{x} = 0{,}55\%$, $s = 0{,}593\%$, $n = 12$; SE $= 0{,}593/\sqrt{12} = 0{,}171\%$; $t = (0{,}55 - 0{,}30)/0{,}171 = 1{,}46$, df $= 11$; $t_{0{,}975;11} = 2{,}201$; $|1{,}46| < 2{,}201 \Rightarrow$ **nicht** verwerfen ($p \approx 0{,}17$). Mittel liegt über dem Ziel, aber $n = 12$ reicht nicht (geringe Power).

```
x <- c(0.5,-0.2,1.1,0.8,-0.5,1.4,0.2,0.9,-0.1,1.2,0.6,0.7)
t.test(x, mu = 0.3)   # t = 1.46, df = 11, p = 0.172
```

Aufgabe 4: Zweistichproben-t-Test

Zwei *unabhängige* Fonds: A ($n = 30$, Mittel 1,20%, $s = 0{,}80\%$), B ($n = 30$, Mittel 0,60%, $s = 0{,}90\%$). Unterscheiden sich die mittleren Renditen? H_0: $\mu_A = \mu_B$ $\quad H_1$: $\mu_A \neq \mu_B$. $s_p^2 = \frac{29 \cdot 0{,}80^2 + 29 \cdot 0{,}90^2}{58} = 0{,}725$, $s_p = 0{,}852$; SE $= 0{,}852\sqrt{1/30 + 1/30} = 0{,}220$; $t = (1{,}20 - 0{,}60)/0{,}220 = 2{,}73$, df $= 58$; $t_{0{,}975;58} = 2{,}00 \Rightarrow$ verwerfen ($p \approx 0{,}008$). Fonds A signifikant besser. Welch (ungleiche Varianzen, Default in R): hier gleiche $n \Rightarrow$ fast identisch.

```
sp <- sqrt((29*0.80^2 + 29*0.90^2)/58)          # 0.852
t  <- (1.20 - 0.60)/(sp*sqrt(1/30 + 1/30))      # 2.73, df = 58
2*pt(-abs(t), df = 58)                            # p = 0.008
```

Aufgabe 5: Paardifferenzentest (gepaart)

Zwei Strategien an *denselben* 10 Tagen (%):

A	1,2	-0,5	0,8	1,5	-0,2	0,9	1,1	0,4	-0,3	1,0
B	0,9	-0,7	0,5	1,0	-0,4	0,6	0,8	0,2	-0,5	0,7

Hat A einen Vorteil gegenüber B (Differenz $d = A - B$)?
H_0: $\mu_d = 0$ H_1: $\mu_d \neq 0$ $d = 0{,}3\,;0{,}2\,;0{,}3\,;0{,}5\,;0{,}2\,;0{,}3\,;0{,}3\,;0{,}2\,;0{,}2\,;0{,}3$; $\bar{d} = 0{,}28\%$, $s_d = 0{,}092\%$; $t = 0{,}28/(0{,}092/\sqrt{10}) = 9{,}64$, df $= 9$; $t_{0{,}975;9} = 2{,}262 \Rightarrow$ klar verwerfen ($p < 0{,}001$). Der kleine, aber *konsistente* Vorsprung ist hoch signifikant (Power der Paarung).

```
A <- c(1.2,-0.5,0.8,1.5,-0.2,0.9,1.1,0.4,-0.3,1.0)
B <- c(0.9,-0.7,0.5,1.0,-0.4,0.6,0.8,0.2,-0.5,0.7)
t.test(A, B, paired = TRUE)   # t = 9.64, df = 9, p < 0.001
```

Teil B: χ^2-Tests

Aufgabe 6: χ^2-Anpassungstest

Tage mit positiver Rendite je Wochentag (Total 100):

	Mo	Di	Mi	Do	Fr
beobachtet O	29	15	14	27	15

Sind die positiven Tage gleich über die Woche verteilt? H_0: gleichverteilt H_1: nicht gleichverteilt. $E = 100/5 = 20$ je Tag; $\chi^2 = \frac{9^2+5^2+6^2+7^2+5^2}{20} = \frac{216}{20} = 10{,}8$; df $= 4$; $\chi^2_{0{,}95;4} = 9{,}488 \Rightarrow$ verwerfen ($p \approx 0{,}029$). Es gibt einen Wochentagseffekt.

```
chisq.test(c(29,15,14,27,15))    # X-squared = 10.8, df = 4, p = 0.029
```

Aufgabe 7: χ^2-Unabhängigkeitstest

Rating gegen Ausfall (200 Anleihen):

	Ausfall ja	Ausfall nein	Summe
Investment Grade	5	95	100
High Yield	20	80	100
Summe	25	175	200

Sind Rating und Ausfall unabhängig? H_0: unabhängig H_1: abhängig. Erwartet E_{ij} = Zeile·Spalte/n: alle 12,5 bzw. 87,5; $\chi^2 = \sum(O - E)^2/E = 10{,}29$ (ohne Korrektur), df $= 1$; $\chi^2_{0{,}95;1} = 3{,}841 \Rightarrow$ verwerfen. Ausfall hängt vom Rating ab (HY fällt häufiger aus).

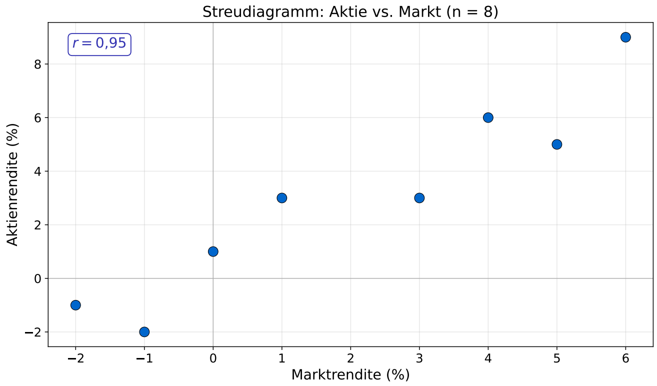
```
M <- matrix(c(5,95,20,80), nrow = 2, byrow = TRUE)
chisq.test(M)    # mit Yates (Default 2x2): X-squared = 8.96, p = 0.003
```

Teil C: Korrelation und Regression

Aufgabe 8: Kovarianz, Korrelation, Streudiagramm

8 Monatsrenditen (%): Markt x und Aktie y.

x (Markt)	-2	-1	0	1	3	4	5	6
y (Aktie)	-1	-2	1	3	3	6	5	9

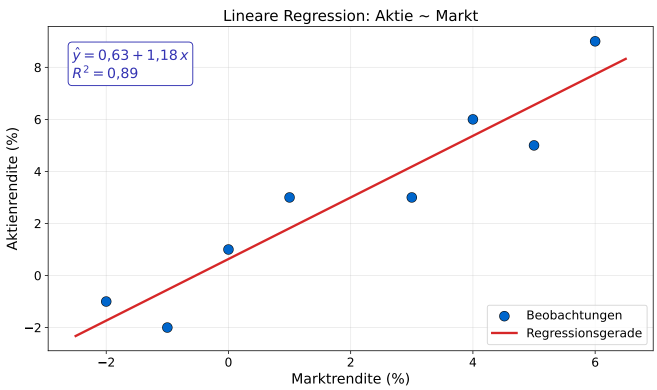


Berechnet Kovarianz, Pearson-r und Spearman-ρ und deutet das Streudiagramm. Mit $\bar{x} = 2$, $\bar{y} = 3$: $S_{xx} = 60$, $S_{yy} = 94$, $S_{xy} = 71$. $\text{cov} = 71/7 = 10{,}14$; $r = 71/\sqrt{60 \cdot 94} = 0{,}945$; $\rho \approx 0{,}95$. Starker positiver linearer Zusammenhang. Pearson misst *linearen*, Spearman *monotonen* Zusammenhang; ein Ausreisser drückt r stark, ρ kaum.

```
cov(x, y)                        # 10.14
cor(x, y)                        # 0.945
cor(x, y, method = "spearman")   # 0.946
```

Aufgabe 9: Lineare Regression (CAPM)

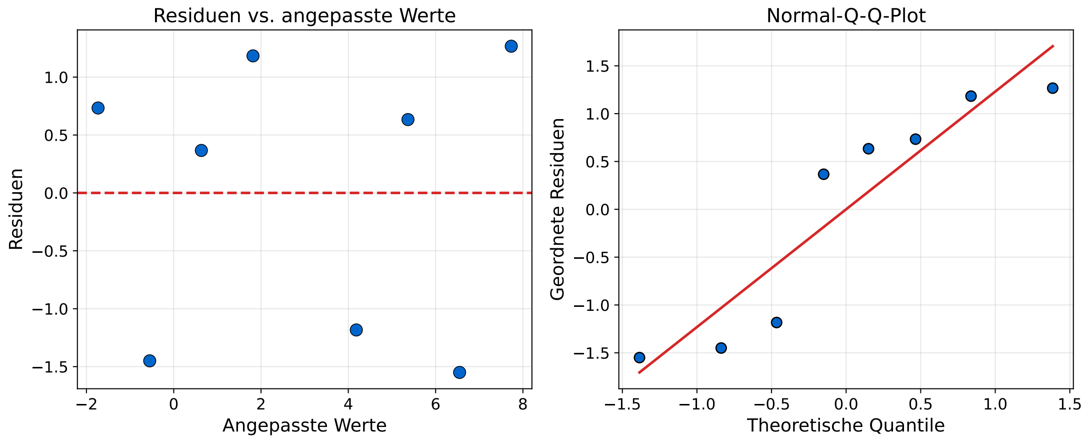
Gleiche Daten wie Aufgabe 8. Regressiert die Aktie auf den Markt, $y = \alpha + \beta x + \varepsilon$. Schätzt β (Marktsensitivität), α (Überrendite) und R^2 und interpretiert. $\hat{\beta} = S_{xy}/S_{xx} = 71/60 = 1{,}18$; $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 3 - 1{,}18 \cdot 2 = 0{,}63$; $R^2 = r^2 = 0{,}89$. $\beta = 1{,}18 > 1$: die Aktie verstärkt Marktbewegungen; $\alpha = 0{,}63\% > 0$: kleine Überrendite; 89% der Varianz erklärt.




```
m <- lm(y ~ x); summary(m)    # Intercept 0.633, x 1.183, R^2 0.894
abline(m)
```

Aufgabe 10: Residuen- und Q-Q-Plot

Prüft die Modellannahmen der Regression aus Aufgabe 9. Was sagen die beiden Plots?



Links (Residuen vs. angepasste Werte): kein Muster, gleichmässige Streuung um 0 $\Rightarrow$ Linearität und konstante Varianz plausibel. **Rechts** (Q-Q-Plot): Punkte nahe der Geraden $\Rightarrow$ Residuen annähernd normal. Trichterform wäre Heteroskedastizität, starke Krümmung wäre Nicht-Normalität.

```
plot(fitted(m), resid(m)); abline(h = 0)
qqnorm(resid(m)); qqline(resid(m))
```

Teil D: Gruppenaufgaben

G1: Welcher Test? (6 Szenarien)

Ordnet jedem Szenario den Test zu und begründet.

1. Anteil defekter Trades $> 2\%$?
2. Mittlere Sharpe-Ratio zweier Strategien (gleiche Tage) verschieden?
3. Renditen vor und nach einem Regimewechsel (zwei Gruppen) verschieden?
4. Sind Branche und Ausfall (ja/nein) unabhängig?
5. Hängt die Aktienrendite linear von der Marktrendite ab?
6. Ist die Verteilung der Ratings wie vom Index vorgegeben?

1. Binomialtest 2. Paardifferenzentest 3. Zweistichproben-t 4. χ^2-Unabhängigkeit 5. Korrelation/Regression 6. χ^2-Anpassung.

G2: R-Ausgabe interpretieren

Lest die Ausgabe und nennt Teststatistik, df, p-Wert, Entscheidung und einen Schlusssatz.

```
        Welch Two Sample t-test
t = 2.73, df = 57.2, p-value = 0.0084
95 percent confidence interval:  0.16  1.04
mean of x  mean of y :  1.20  0.60
```

$t = 2{,}73$, df ≈ 57, $p = 0{,}008 < 0{,}05 \Rightarrow H_0$ verwerfen. Die mittlere Rendite von A (1,20%) ist signifikant höher als die von B (0,60%); das KI [0,16; 1,04] enthält die 0 nicht.

G3: Findet den Fehler

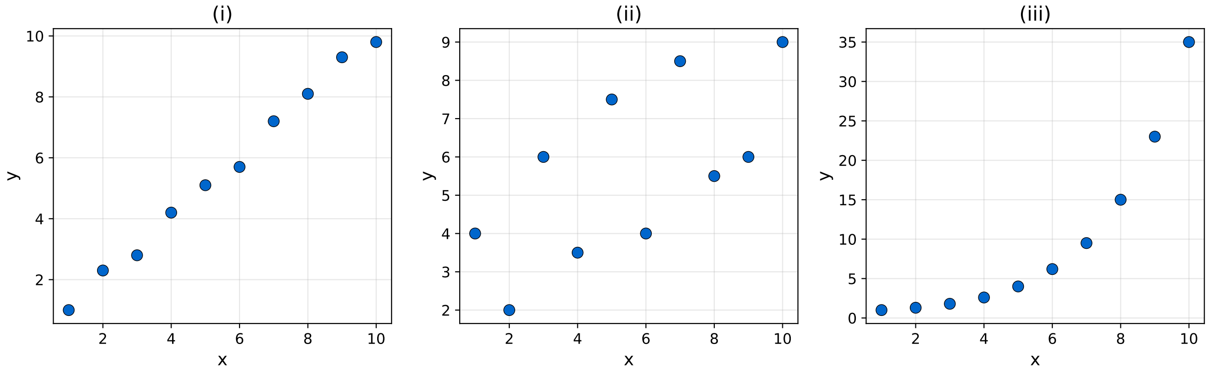
Was ist jeweils falsch, und wie korrigiert ihr es?

1. “Wir testen zwei Strategien an denselben 10 Tagen mit `t.test(A, B)` ohne `paired`.”
2. “$p = 0{,}12 > 0{,}05$, also ist H_0 bewiesen.”
3. “$r = 0{,}9$ zwischen Werbung und Umsatz, also verursacht Werbung den Umsatz.”

1. Gepaarte Daten brauchen `paired = TRUE` (sonst Power verschenkt). 2. Nicht ablehnen ist kein Beweis von H_0 (fehlende Evidenz $\neq$ Evidenz für H_0). 3. Korrelation ist nicht Kausalität (Drittvariablen, Scheinkorrelation).

G4: Streudiagramm mit Augenmass

Unten seht ihr drei Punktwolken (i), (ii), (iii).



(a) Ordnet sie nach Stärke des *linearen* Zusammenhangs.
(b) Skizziert je die ungefähre Regressionsgerade.
(c) Bei welcher Wolke weichen Pearson-r und Spearman-ρ am stärksten ab, und warum?

(a) Linear (Pearson): (i) $r \approx 1{,}0 \ >$ (iii) $r \approx 0{,}9 \ >$ (ii) $r \approx 0{,}65$. (b) Die Gerade folgt der Wolke; bei (iii) passt eine Gerade nur grob. (c) Bei **(iii)**: $\rho \approx 1{,}0$ (perfekt monoton), aber $r \approx 0{,}89$ (gekrümmt, nicht linear). Spearman erkennt die Monotonie, Pearson misst nur den linearen Anteil. Bei (i) und (ii) sind r und ρ fast gleich.