

Multiple-Choice-Klausur

Statistik II

Prof. Dr. Dr. H. Toutenburg

Bitte ankreuzen
Richtig/Falsch

1. Ist X eine exponentialverteilte Variable mit $\lambda = 3$, dann ist die Varianz von $3 + 3X$ genau 9.
($Var(3 + 3X) = Var(3X) = 9Var(X) = 9 \frac{1}{\lambda^2} = 9 \frac{1}{3^2} = 9 \frac{1}{9} = 1 \neq 9$)
2. Der Fehler 2. Art bezeichnet die Wahrscheinlichkeit, dass H_0 beibehalten wird, obwohl H_1 richtig ist.
3. Ist $E(X) = 0$ und $E(Y) = 0$, so gilt: $Cov(X, Y) = E(X \cdot Y)$.
($Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$)
4. Zwei zufällige Ereignisse A und B heißen disjunkt, falls ihr gleichzeitiges Eintreten unmöglich ist.
5. Eine Modellannahme im multiplen linearen Regressionsmodell $Y = X\beta + \epsilon$ ist, dass die Fehlerterme normalverteilt sind.

1. Sei $X \sim N(5; 2^2)$ und $Y \sim N(2; 2^2)$. Dann folgt $E\left(\frac{X}{Y}\right) = \frac{5}{2}$.
2. Die Varianz einer auf dem Intervall $[0;2]$ gleichverteilten Zufallsvariablen X ist kleiner als die Varianz einer auf dem Intervall $[0;3]$ gleichverteilten Zufallsvariable Y .
3. 3 grüne, 5 gelbe und 2 blaue Zettel können auf 2520 unterscheidbare Arten auf 10 nebeneinanderliegende Kästchen verteilt werden.
4. Eine diskrete Zufallsvariable X besitzt folgende Wahrscheinlichkeitsfunktion: $P\{X = 0\} = 0.25, P\{X = 1\} = 0.5, P\{X = 2\} = 0.25$. Es gilt für die Varianz von X : $\text{Var}X = 1$.
5. Für die $N(2; 2^2)$ verteilte Zufallsgröße Y gilt: $P\{Y \leq 3\} = 0.5987$
6. Für die Schätzung des unbekanntes Parameters μ einer normalverteilten Grundgesamtheit ergibt sich aufgrund der Stichprobenwerte $\bar{x} = 60$ und $s^2 = 49$ mit $n = 16$ das Konfidenzintervall $[56.57; 63.43]$ bei einem Konfidenzniveau von 0.95.
7. Gegeben seien zwei stochastisch unabhängige Ereignisse A und B . Es gilt $P(B) = 0.4$ und $P(A \cap B) = 0.2$. Dann gilt für $P(A \cup B) = 0.9$.
8. Die Hypothese „Zwei Stichproben entstammen derselben Grundgesamtheit“ kann sowohl mit dem Iterationstest als auch mit dem Anpassungstest von Kolmogoroff Smirnow überprüft werden.

Bitte ankreuzen
Richtig / Falsch

1. Die Funktion $f(x) = x$, $x \in [0, 1]$, ist eine Dichte.
2. Sind zwei beliebige Zufallsvariablen unkorreliert, so sind sie auch unabhängig.
3. Das Stichprobenmittel ist ein ML-Schätzer für den Erwartungswert einer Verteilung.
4. Wirft man einen Würfel zweimal hintereinander, so ist die Wahrscheinlichkeit für jedes Elementarereignis $1/36$.
5. Hat die Zufallsgröße X den Erwartungswert 2, so hat $\frac{1}{X}$ den Erwartungswert $\frac{1}{2}$.
6. X sei eine beliebig verteilte Zufallsgröße mit der Streuung σ . Die Wahrscheinlichkeit, daß eine Realisation im 2σ -Bereich liegt, beträgt mindestens 0.75.
7. Der McNemar-Test setzt zwei unabhängigen Stichproben voraus.
8. Im folgenden Vierfeldermodell

Werbung	Umsatz	
	hoch	niedrig
wenig	40	60
viel	20	80

 besteht ein positiver Zusammenhang zwischen Werbung und Umsatz (Werbung wirkt umsatzsteigernd).
9. Die beiden Zufallsvariablen X und Y seien unkorreliert. Dann folgt $E(XY) = E(X)E(Y)$.
10. Aufgrund des nachfolgenden SPSS-Listing für den 2-Stichproben K-S-Test wird die Hypothese der Gleichheit der Verteilungen nicht abgelehnt.

```

- - - - Kolmogorov - Smirnov 2-Sample Test
      XY
      by GRUPPE

```

```

Cases
  15  GRUPPE = 1.00  X
  12  GRUPPE = 2.00  Y
  --
  27  Total

```

```

      Most extreme differences
Absolute   Positive   Negative   K-S Z   2-Tailed P
.31667     .31667     -.15000   .8176   .5157

```

- Bitte ankreuzen
Richtig / Falsch
1. Will man nachweisen, daß eine Zufallsgröße normalverteilt ist, so verwendet man den Kolmogorov-Smirnov-Anpassungstest.
 2. Die hypergeometrische Verteilung entspricht dem Urnenmodell ohne Zurücklegen und die Binomialverteilung dem Urnenmodell mit Zurücklegen.
 3. Sind zwei normalverteilte Zufallsgrößen X und Y unkorreliert, so sind sie auch unabhängig.
 4. Die Varianz eines Merkmals mit lauter gleichen Merkmalsausprägungen ist gleich Null.
 5. Es gilt stets $E(X + Y) = E(X) + E(Y)$.
 6. Da im nachfolgenden Listing die Gleichheit der Varianzen abgelehnt wird, ist als t -Wert -7.24 zu wählen

t-tests for Independent Samples of TYP

Variable	Number of Cases	Mean	SD	SE of Mean

BRENND Brenndauer				
X	25	5996.4863	65.304	13.061
Y	22	6125.5776	56.961	12.144

Mean Difference = -129.0914

Levene's Test for Equality of Variances: F= .157 P= .694

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff

Equal	-7.17	45	.000	17.993	(-165.331, -92.851)
Unequal	-7.24	45.00	.000	17.834	(-165.012, -93.171)

Bitte ankreuzen
Richtig / Falsch

1. \bar{X} ist ein erwartungstreuer Schätzer für den Parameter μ einer $N(\mu, \sigma^2)$ -verteilten Zufallsgröße X .
2. Hat X die Varianz $\frac{1}{2}$, dann hat $\frac{1}{2}X$ die Varianz $\frac{1}{8}$.
3. Sind $X \sim B(n_1, p_1)$, $Y \sim B(n_2, p_2)$, und sind X und Y unabhängig, dann gilt: $X + Y \sim B(n_1 + n_2, p_1 + p_2)$.
4. Das Modell $y = a + b \cos(3x)$ ist ein lineares Regressionsmodell.
5. Der F -Test ist ein Zweistichprobentest und setzt Normalverteilung und Gleichheit der Varianzen voraus.
6. Die Varianz eines Merkmals mit lauter gleichen Merkmalsausprägungen ist gleich Eins.
7. Zur Anwendung der Tschebyscheffschen Ungleichung für die Zufallsvariable X mit $E(X) = \mu$ und $\text{Var}(X) = \sigma^2$ muß X normalverteilt sein.
8. Sind Z_1, \dots, Z_n unabhängig und identisch $N(0, 1)$ -verteilt, dann gilt: $\sum_{i=1}^n Z_i^2$ ist χ^2 -verteilt mit n Freiheitsgraden.
9. Die hypergeometrische Verteilung entspricht dem Urnenmodell ohne Zurücklegen und die Binomialverteilung dem Urnenmodell mit Zurücklegen.
10. Sind zwei Zufallsgrößen X und Y unkorreliert, so sind sie auch unabhängig.
11. Die Wirksamkeit einer Therapie soll anhand des Blutparameters $Lp(a)$ gemessen werden. Aufgrund des Ergebnisses des paired t -Test kann die Nullhypothese $H_0 : \mu_{\text{nachher}} \geq \mu_{\text{vorher}}$ zum Signifikanzniveau $\alpha = 0.05$ abgelehnt werden.

Paired Samples Test

	Paired Differences			t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
Lp(a) nachher - Lp(a) vorher	-.9000	1.3776	.4356	-2.066	9	.069

Ein Autohändler möchte wissen, welche Einflußgrößen den monatlichen Umsatz in den Jahren 1984 bis 1991 beeinflusst haben könnten. Neben den Umsatzzahlen wurde die Mitarbeiterzahl, die Werbungsausgaben und das Jahr berücksichtigt.

```

* * * * MULTIPLE REGRESSION * * * *

Listwise Deletion of Missing Data

Equation Number 1   Dependent Variable..  UMSATZ

Block Number  1.  Method:  Enter

Multiple R   .64098 R Square .41085
Adjusted R Square .39164

Standard Error   16.93213

Analysis of Variance
      DF      Sum of Squares      Mean Square
Regression      3      18393.59180      6131.19727
Residual      92      26376.11745      286.69693

Signif F = .0000

----- Variables in the Equation -----
Variable      B      SE B      Beta      T      Sig T
MITARB      3.579099      .919831      .944229      3.891      .0002
WERBUNG      .043395      .663900      .005235      .065      .9480
JAHR      -3.119514      2.286515      -.330986     -1.364      .1758
(Constant)  449.432830  156.102248      2.879      .0050

```

Bitte ankreuzen
Richtig / Falsch

1. In die Stichprobe gehen die Zahlen von 95 Monaten ein.
2. Der Wert der F -Statistik beträgt 21.38564.
3. Die Variablen 'Werbung' und 'Jahr' sind relevant für das Modell, weil die p -values der zugehörigen Tests größer sind als 0.05.
4. 'Beta' steht für die unstandardisierten Parameterschätzungen.
5. Die Konstante übt den größten Einfluß auf den Response aus, weil ihre Schätzung betraglich am größten ist.

Um ihr Übergewicht etwas zu reduzieren, unterziehen sich 30 Personen einem mehrwöchigen Fitnessprogramm, 15 Personen bei Trainer A, 15 Personen bei Trainer B. Im Anschluß an die Einheiten wird ihr Gewicht gemessen, um Unterschiede zwischen den Trainern feststellen zu können. Der t -Test für zwei unabhängige Stichproben liefert im SPSS folgende Ergebnisse.

	<i>N</i>	Mean	Std. Dev.	Std. Error Mean
Trainer A	15	69.164	4.3649	1.1270
Trainer B	15	63.808	3.2463	0.8382

Tabelle 1: Group Statistics.

	Levene's Test for Equality of Variances	t -test for Equality of Means					
	<i>F</i>	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Diff.
Equal variances assumed	3.350	0.078	???	28	0.001	5.356	.14045

Tabelle 2: Independent Samples Test.

- Bitte ankreuzen
Richtig / Falsch
- Der Test auf Varianzhomogenität lehnt $H_0 : \sigma_{\text{Trainer A}}^2 = \sigma_{\text{Trainer B}}^2$ nicht ab, denn der p -value ist größer als 0.05.
 - Der Wert der fehlenden Teststatistik beträgt 3.813, wenn man bekannte Varianzen voraussetzt (Gauß-Test).
 - Der p -value des Tests auf gleiche Mittelwerte ist kleiner als 0.05, folglich wird die Hypothese $H_0 : \mu_{\text{Trainer A}} = \mu_{\text{Trainer B}}$ abgelehnt.
 - Die Stichprobenumfänge beim Zweistichproben- t -Test für unabhängige Stichproben müssen gleich groß sein.
 - Das Gesamtgewicht der Teilnehmer bei Trainer A beträgt 1037.46kg

Ein skeptischer Bundesliga-Kapitän notiert 19mal ‘Zahl’ und 15mal ‘Kopf’ bei der Seitenwahl nach der abgelaufenen Bundesligasaison. Der Hobby-Statistiker erhält den folgenden SPSS-Output:

χ^2 -Test

	Category	<i>N</i>	Obs. Prop.	Test Prop.	Asymp. Sig. (2-tailed)
Group 1	Kopf	15	.44	0.5	0.607
Group 2	Zahl	19	.56		
Total		34	1.00		

Bitte ankreuzen
Richtig / Falsch

1. Die zugehörige Hypothese $H_0 : p = p_0$ mit $p_0 = 0.5$ kann mittels Binomialtest überprüft werden.
2. Der Wert der zugehörigen Teststatistik beträgt 0.6997.
3. $np(1 - p)$ ist hier kleiner als 9, also kann die Standardnormalverteilung als Näherung verwendet werden.
4. Laut SPSS ist der p -value größer als 0.05 und somit muß H_0 verworfen werden.
5. Generell kann $p(1 - p)$ maximal den Wert 0.25 annehmen.

Betrachten Sie die folgende (2×2) -Kontingenztafel:

	$X = 1$	$X = 0$	
$Y = 1$	$f_{11} = 0.2$	$f_{12} = 0.1$	
$Y = 0$	$f_{21} = 0.3$	$f_{22} = 0.5$	

In dieser Tafel ist ein Fehler enthalten. Es sei aber von der ‘wahren’ Tabelle bekannt, dass

$$P(X = 1 | Y = 1) = 2/3, \quad P(X = 0 | Y = 1) = 1/3$$

$$P(X = 1 | Y = 0) = 2/7, \quad P(X = 0 | Y = 0) = 5/7.$$

Bitte kreuzen Sie an, ob die folgenden Aussagen richtig oder falsch sind. Ein richtiges Statement ergibt 2 Punkte, bei einem falschen Statement werden 2 Punkte abgezogen. Eine negative Gesamtpunktzahl wird gleich Null gesetzt.

- | | Bitte ankreuzen
Richtig / Falsch | |
|---|-------------------------------------|-----------------------|
| 1. Bei einem Stichprobenumfang von $N = 10$ hat der Odds-Ratio der ‘wahren’ Tafel den Wert 5. | <input type="radio"/> | <input type="radio"/> |
| 2. Die richtigen Zellhäufigkeiten lauten:
$f_{11} = 0.2, f_{12} = 0.1, f_{21} = 0.3, f_{22} = 0.4$. | <input type="radio"/> | <input type="radio"/> |
| 3. Für die bedingte Wahrscheinlichkeit der ‘wahren’ Tabelle gilt
$P(Y = 0 X = 1) = 1/2$ | <input type="radio"/> | <input type="radio"/> |
| 4. Bei einem Stichprobenumfang von $N = 100$ gilt in der ‘wahren’ Tafel
$n_{22} = 40$. | <input type="radio"/> | <input type="radio"/> |
| 5. In (2×2) -Kontingenztafeln muss der Stichprobenumfang größer als 10 sein. | <input type="radio"/> | <input type="radio"/> |