Klausur Statistik (7.5 ECTS)

Name		**Prüfer**	Prof. Dr. I. Klein
Vorname		**Arbeitszeit**	Mittwoch, 03.08.2016 14:00 - 16:00 Uhr
Matrikelnummer			
Studienrichtung		**Sitzplatznummer**	
Semesterzahl		**Raum**	
Email (optional)			

Hinweis: Aufgabenblätter nicht auseinandertrennen!

Ergebnis:

Statistik	
Aufgabe	**Punkte**
1	
2	
3	
4	
Summe	
Note:	

Unterschrift des Kandidaten: ______________________

Unterschrift des Prüfers: ______________________

Hilfsmittel: Es gelten folgende Regelungen zu den erlaubten Hilfsmitteln:

- Nicht programmierbarer Taschenrechner
- Die vom Lehrstuhl seit dem WS 2014/15 offiziell herausgegebene Formelsammlung (DIN A5, gebunden, orangener Umschlag). Es sind prinzipiell keine weiteren Eintragungen oder Markierungen darin erlaubt. Ausgenommen sind farbliche Hinterlegungen von Textpassagen und/oder Formeln. Außerdem ist das Erratum zur 1. Auflage der Formelsammlung erlaubt.
- R Reference Card von Jonathan Baron, es sind keine weiteren Eintragungen oder Markierungen darin erlaubt.

Bewertung: Für jede Aufgabe werden maximal zehn Punkte vergeben. Bewertet werden grundsätzlich nur Lösungen, die im Lösungsteil stehen und für die Folgendes beachtet wird:

- Der Lösungsweg muss aus einer Darstellung der einzelnen Rechenschritte ersichtlich sein.
- Antworten sind stets zu begründen, es sei denn es wird ausdrücklich keine Begründung verlangt.
- Unleserliche Aufgabenteile werden mit 0 Punkten bewertet.

Viel Erfolg!

Aufgabe 1

An einer Universität finden Bauarbeiten statt.

1. Anstehende Klausuren werden in einem anderen Gebäude geschrieben, wenn der Baulärm B längerfristig 65 Dezibel (Abkürzung: dB) überschreitet.

Eine einfache Zufallsstichprobe, bei der an 14 Werktagen jeweils zur selben Uhrzeit Messungen durchgeführt wurden, hat folgendes ergeben:

Messung i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\mathbf{b}_i$	25	85	100	90	108	35	55	48	120	87	91	75	81	65

Nehmen Sie an, die gemessenen Daten stammen aus einer Normalverteilung mit unbekanntem Mittelwert μ und Varianz $\sigma^2 = 484$. Testen Sie die Vermutung
H_1: Mittlerer Baulärm in dB > 65 auf dem 10%-Signifikanzniveau.
Sie wissen, dass $\bar{b}_{14}$=76.07.

a) Berechnen Sie dazu zunächst die Teststatistik.

b) Wie lautet die kritische Schranke?

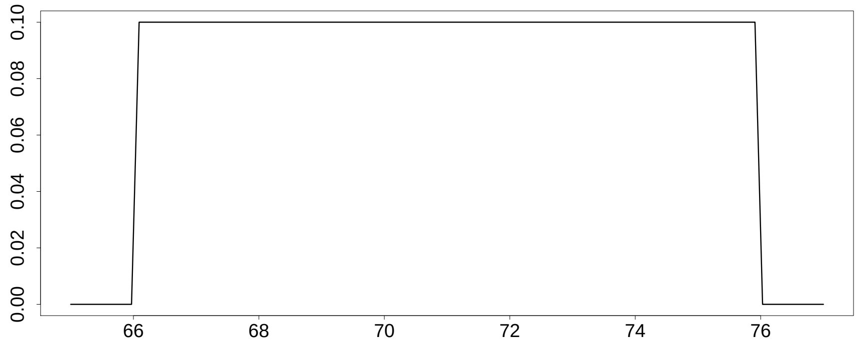
c) Berechnen Sie das rechtsseitige realisierte 90%-Konfidenzintervall.

d) Treffen Sie eine Testentscheidung und begründen Sie diese.
(Hinweis: Sollten Sie in den vorherigen Teilaufgaben keine Lösung erhalten haben, nutzen Sie für die Testentscheidung: p-Wert=0.0299).

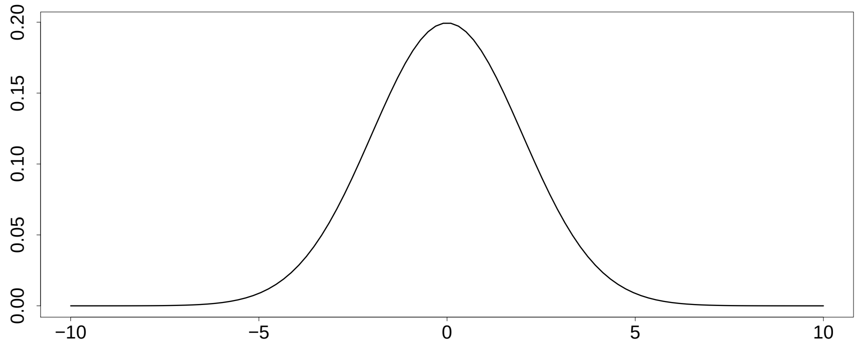
e) Berechnen Sie die Güte des Tests, wenn $\mu_1 = 71$.

2. Welche der folgenden Grafiken (Grafik 1, Grafik 2 oder Grafik 3) entspricht einer Normalverteilung mit Parametern $\mu = 71$ und $\sigma^2 = 484$?

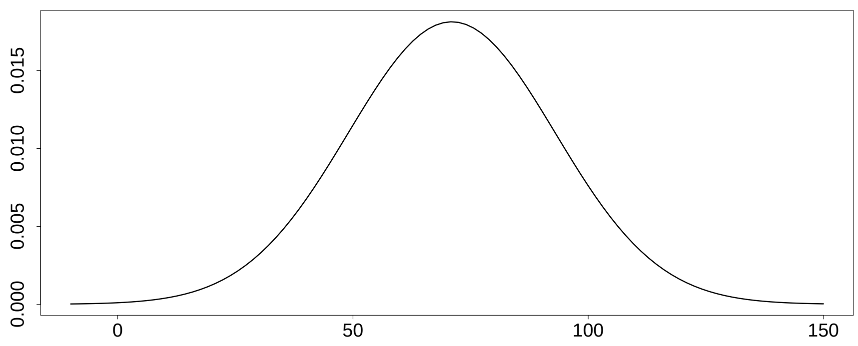
Grafik 1:



Grafik 2:



Grafik 3:



3. Nun wurde der Baulärm [in dB] während einer Woche jeweils vormittags, mittags und nachmittags gemessen.
   Folgende Datenpunkte ergaben sich:

Tageszeit/Baulärm in dB	**0-40**	**41-65**	**>65**	Σ
vormittags	4	2	8	14
mittags	5	8	1	14
nachmittags	7	4	3	14
Σ	16	14	12	42

a) Wie oft wurden mittags Werte über 40 dB gemessen?

Sie interessieren sich dafür, ob der 'Baulärm in dB' unabhängig von der 'Tageszeit' ist. Dazu führen Sie einen χ^2-Unabhängigkeitstest auf einem 5% Signifikanzniveau durch und erstellen folgende Indifferenztabelle:

Tageszeit/Baulärm in dB	**0-40**	**41-65**	**>65**	Σ
vormittags	w)	$4\frac{2}{3}$	4	14
mittags	$5\frac{1}{3}$	x)	4	14
nachmittags	$5\frac{1}{3}$	$4\frac{2}{3}$	y)	14
Σ	16	14	12	z)

Nehmen Sie im Weiteren an, dass die Bedingungen für eine gute Approximation an die χ^2 - Verteilung erfüllt sind.
Zusätzlich haben Sie folgende Information:

$$\Sigma_{i=1}^{k}\Sigma_{j=1}^{l}\frac{(n_{ij}-\frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n}} = 11.375.$$

b) Geben Sie die fehlenden Werte der Indifferenztabelle an.

w)

x)

y)

z)

c) Geben Sie den Wert der kritischen Schranke für den Unabhängigkeitstest an.

Sollten Sie keinen Wert für die kritische Schranke erhalten, verwenden Sie im Folgenden den Wert 11.

d) Können Sie anhand Ihrer Testentscheidung der Aussage 'Der Baulärm ist von der Tageszeit abhängig' zustimmen? Begründen Sie Ihre Antwort.

Cramérs V wurde berechnet und beträgt hier 0.3679.

e) Berechnen Sie die mittlere quadratische Kontingenz.

f) Welchen Wert nimmt die mittlere quadratische Kontingenz bei Unabhängigkeit zweier Merkmale an?

g) Welchen Wert würde die mittlere quadratische Kontingenz bei vollständiger Abhängigkeit der beiden Merkmale 'Baulärm in dB' und 'Tageszeit' annehmen?

Schmierpapier zu Aufgabe 1

Aufgabe 2

Die Zeitdauer T [in Jahren] bis zum Ausfall eines Kredits sei exponentialverteilt mit Parameter λ, d.h. $f(t) = \lambda \exp(-\lambda t)$, wobei $\lambda > 0$.

1. Zeigen Sie, dass die logarithmierte Likelihoodfunktion für n unabhängig und identisch verteilte Zeitdauern bis zum Ausfall gegeben ist durch

$$\ln(L(\lambda; t_1, \ldots, t_n)) = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i.$$

Die oben genannte logarithmierte Likelihoodfunktion ist global konkav.

2. Zeigen Sie, dass der Maximum-Likelihood-Schätzer für λ gegeben ist durch

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i}.$$

Sie haben von 10 Krediten folgende Zeitdauern bis zum Ausfall beobachtet.

i	1	2	3	4	5	6	7	8	9	10
t_i	1.8	2.2	0.5	0.8	3.5	0.9	3.7	0.2	5.1	2.3

3. Berechnen Sie den Schätzwert für $\widehat{\lambda}_{ML}$ konkret.

Für 19 ausgefallene Kredite liegen Beobachtungen über die Zufallsvariable "Höhe des jeweils eingetreten Verlusts" V [in Mio Euro] vor.

i	1	2	3	4	5	6	7	8	9	10
v_i	0.1	0.1	0.2	0.4	0.5	0.6	0.7	0.9	0.9	0.9

i	11	12	13	14	15	16	17	18	19
v_i	1.2	2.5	3.8	4.5	4.9	5.7	6.2	7.1	12.3

Hinweis: $\sum_{i=1}^{19} v_i = 53.5$ und $\sum_{i=1}^{19} v_i^2 = 342.77$

4. Bestimmen Sie das arithmetische Mittel der Beobachtungen aus V.

5. Bestimmen Sie den Schätzwert des Medians der Beobachtungen aus V.

6. Bestimmen Sie den Schätzwert des Modus der Beobachtungen aus V.

7. Bestimmen Sie den Wert des erwartungstreuen Schätzers

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(V_i - \bar{V}_n)^2$$

für die Varianz der Beobachtungen aus V.

Eine wichtige Maßzahl für die Quantifizierung finanzieller Risiken ist der so genannte Value-at-Risk VaR_α. Dieser ist definiert als das α-Quantil der Verteilung der Verluste.

8. Schätzen Sie den $VaR_{0.9}$ als empirisches Quantil aus den obigen Daten.

In einer Bank werden Kreditnehmer bezüglich ihrer Bonität in Ratingklassen A (gut), B (schlecht) und D (default - ausgefallen) eingeteilt, wobei sich die Bonitätseinschätzung im Lauf der Zeit ändern kann. Die folgende Tabelle zeigt **bedingte** Wahrscheinlichkeiten $P(R_{t+1}|R_t)$ für die Veränderungen der Zufallsvariable R ("Ratingklasse") eines einzelnen Kreditnehmers innerhalb eines Ein-Jahres-Zeithorizonts $[t, t+1)$. Die Veränderungen sollen dabei unabhängig vom konkreten Zeitpunkt t als konstant angenommen werden.

		$t+1$		
		A	B	D
	A	0.76	0.23	0.01
t	B	0.21	0.63	0.16
	D	0	0	1

9. Welches maximal zulässige Skalenniveau besitzt die Zufallsvariable "Ratingklasse"?

10. Wie groß ist die Wahrscheinlichkeit $P(R_1 = D|R_0 = B)$, d.h. dass ein Kreditnehmer innerhalb des Jahres $[0, 1)$ aus Ratingkategorie B ausfällt?

11. Wie groß ist die Wahrscheinlichkeit $P(R_2 = A|R_0 = B)$, d.h. dass ein Kreditnehmer innerhalb des Zwei-Jahres-Zeitraums $[0, 2)$ von Ratingkategorie B in Klasse A aufsteigt?

Schmierpapier zu Aufgabe 2

Aufgabe 3

Teil I:

1. Sie investieren 100 EUR in Aktie A. Der Aktienkurs besitzt im ersten Jahr eine Wachstumsrate von 3%. Nach drei Jahren hat die Aktie wieder ihren Anfangswert. Die Wachstumsraten im zweiten und dritten Jahr sind identisch. Berechnen Sie die Wachstumsrate im zweiten bzw. dritten Jahr.

2. Sie investieren 200 EUR in Aktie B. Der Aktienkurs hat eine durchschnittliche jährliche Wachstumsrate von 3 Prozent. Berechnen Sie den Wert nach 5 Jahren.

3. Sie entscheiden sich für eine Investition in eine Aktie. Sie investieren im ersten Jahr 100 EUR und im zweiten Jahr 200 EUR. Im ersten Jahr kostet eine Aktie 20 EUR. Der durchschnittliche Preis pro Aktie beträgt über alle 2 Jahre hinweg 12 EUR. Berechnen Sie den Preis für eine Aktie im zweiten Jahr.

Teil II:

Ein Anleger hätte gerne ein genaueres Verständnis bezüglich des Risikos seiner Anlage. Die Zufallsvariable W beschreibt die jährliche Rendite, wobei W normalverteilt ist mit den Parametern $\mu_W = 3$ und $\sigma_W^2 = 1.21$.

4. Bestimmen Sie die Wahrscheinlichkeit, dass die Rendite W mindestens 6 ist, unter der Bedingung, dass die Rendite mindestens 4 beträgt.

5. Mit welcher Wahrscheinlichkeit liegt die Rendite W zwischen 0.5 und 4?

6. Welche Rendite w wird mit einer Wahrscheinlichkeit von 50% maximal erreicht?

Teil III:

Im Folgenden beschreibt die Zufallsvariable Z die jährliche Rendite von einer Aktie, wobei Z den Erwartungswert $\mu_Z = 2$ und Varianz $\sigma_Z^2 = 0.5$ hat. Die Verteilung von Z ist unbekannt.

7. Berechnen Sie das zweite Potenzmoment von Z.

8. Wie hoch ist die Wahrscheinlichkeit mindestens, dass Z im 2-fachen Schwankungsintervall liegt?

Teil IV:

Aufgrund von Risikodiversifizierung wird im Folgenden ein Portfolio aus Aktien betrachtet. Die Zufallsvariable X beschreibt die Rendite von Aktie 1, wobei X den Erwartungswert μ_X und Varianz σ_X^2 besitzt. Die Rendite von Aktie 2 wird durch die Zufallsvariable Y mit Erwartungswert μ_Y und Varianz σ_Y^2 definiert. Die Korrelation zwischen den Zufallsvariablen X und Y beträgt ρ_{XY}. Die Rendite des Portfolios P wird durch folgende Zufallsvariable definiert:

$$P = \alpha X + (1 - \alpha)Y, \text{ wobei } \alpha \in [0, 1].$$

9. Bestimmen Sie den Erwartungswert von P in Abhängigkeit von α, μ_X, μ_Y.

10. Bestimmen Sie die Standardabweichung von P in Abhängigkeit von $\alpha, \sigma_X, \sigma_Y, \rho_{XY}$. Für welchen Wert von ρ_{XY} wird die Standardabweichung minimal?

Schmierpapier zu Aufgabe 3

Aufgabe 4

Mit einer Spritzgussmaschine werden weiße und schwarze Steine für ein Outdoor-Mühlespiel produziert.

1. Welchen Merkmalstyp hat das Merkmal „Farbe des Spielsteins“? Nennen Sie die maximal mögliche Skalierung.

Ihnen sei die normalverteilte Zufallsvariable G: „Gewicht eines gefertigten Steins [in g]“ gegeben. Als erste Qualitätskontrolle wird eine Gewichtsprüfung durchgeführt. Dafür ziehen Sie aus dem aktuellen Los eine Stichprobe von 10 Steinen und ermitteln ihr Gewicht.

Stein i	1	2	3	4	5	6	7	8	9	10
Gewicht in g	31.3	32.0	32.2	31.7	32.1	30.9	32.6	32.0	30.7	31.5

2. Bestimmen Sie Folgendes:

   a) Sprungstellen der empirischen Verteilungsfunktion und ihre zugehörigen Funktionswerte.

g_i									
$F(g_i)$									

   b) Unteres und oberes Quartil.

   c) Spannweite.

d) Stichproben-Standardabweichung.

3. Ergänzen Sie sinnvoll: Im Allgemeinen charakterisieren:
   - Median und Erwartungswert die ________________ einer Verteilung.
   - Varianz bzw. Standardabweichung die ________________ einer Verteilung.

4. Sie vermuten, dass das mittlere Gewicht der Steine bei $\alpha = 5\%$ signifikant von 32.0 g verschieden ist. Diese Vermutung überprüfen Sie mit einem geeigneten Test. Dafür betrachten Sie das Gewicht $G_1, G_2, \ldots, G_{25}$ von 25 Steinen als Zufallsstichprobe. Sie stellen folgende Hypothesen auf:

$$H_0 : \mu = \mu_0 = 32.0 \qquad \text{gegen} \qquad H_1 : \mu \neq \mu_0 = 32.0$$

   a) Bestimmen Sie die realisierte Prüfgröße. Nehmen Sie an, dass das realisierte Stichprobenmittel 31.8 g und die realisierte Stichproben-Standardabweichung 1.0 g betragen.

   b) Ermitteln Sie die kritische Schranke.

c) **Arbeiten Sie jetzt mit einem p-Wert von 0.5**. Treffen Sie die Testentscheidung und begründen Sie diese.

5. Gehen Sie jetzt davon aus, dass $G \sim \mathcal{N}(\mu = 32.0, \sigma^2 = 1.0)$. Ein Gewicht über 34.0 oder unter 30.0 Gramm bedeutet Ausschuss, dazwischen stimmt das Gewicht. Berechnen Sie die Wahrscheinlichkeit dafür, ein Produkt außerhalb der Gewichtstoleranz zu fertigen.

6. Nehmen Sie an, es seien folgende Fehlerwahrscheinlichkeiten beim Spritzgießen der Spielsteine gegeben:

   - Die Wahrscheinlichkeit p_1 für einen Farbfehler beträgt 1 Prozent.
   - Die Wahrscheinlichkeit p_2 für einen Oberflächenfehler beträgt 2 Prozent.
   - Die Wahrscheinlichkeit p_3 für einen Gewichtsfehler beträgt 1.5 Prozent.

   Die Fehlerwahrscheinlichkeiten seien voneinander unabhängig. Wie hoch ist die Wahrscheinlichkeit, ein fehlerfreies Produkt zu produzieren?

7. Nehmen Sie an, die Wahrscheinlichkeit, dass ein Fertigungsfehler auftritt liegt bei $p = 0.05$.

   a) Ein Mühlespiel besteht aus 18 Steinen. Wie hoch ist die Wahrscheinlichkeit 18 fehlerfreie Steine in Folge zu produzieren?

   b) Geben Sie die Standardabweichung der Anzahl fehlerfreier Steine bei einer Losgröße von 9000 Steinen an.

Schmierpapier zu Aufgabe 4

Lösung Klausur SS16 (7.5 ECTS)

Lösung 1

1. a) $t_n = \sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma} = 1.8829$ 0.5 P

   b) $\lambda_{0.90} = 1.2815$ 0.5 P

   c) $[76.07 - 1.2815\frac{22}{\sqrt{14}}; \infty) = [68.5362; \infty)$ 1.0 P

   d) 1.88>1.28; H_0 ablehnen 0.5 P

   e) 1- $\Phi(1.28 + \sqrt{14}\frac{65-71}{22}$)=0.3974 1.0 P

2. Grafik 3 0.5 P

3. a) 8+1=9 0.5 P

   b) w) $n_{11} = 5\frac{1}{3}$, 0.5 P

      x) $n_{22} = 4\frac{2}{3}$, 0.5 P

      y) $n_{33} = 4$, 0.5 P

      z) $N = 42$ 0.5 P

   c) $\chi^2_{0.95;2*2=4} = 9.49$ 0.5 P

   d) $\chi^2 = 11.375 > 9.49$ (oder >11); H_0 ablehnen bei $\alpha = 0.05$; 0.5 P
   Man kann der Aussage zustimmen. 0.5 P

   e) $(0.3679)^2 * 2 = 0.2707$ 1.0 P

   f) 0 0.5 P

   g) Maximalwert: $\min\{k-1; l-1\}$ =2 0.5 P

Lösung 2

1.)

$$L(\lambda; t_1, \ldots, t_n) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)$$

$$\Rightarrow \ln(L(\lambda; t_1, \ldots, t_n)) = \ln(\lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} t_i \qquad \square$$

1.5 P

2.)

$$\widehat{\lambda}_{ML} \Leftrightarrow \frac{\partial}{\partial \lambda} \ln(L(\lambda; t_1, \ldots, t_n)) = 0$$

$$\frac{\partial}{\partial \lambda} \ln(L(\lambda; t_1, \ldots, t_n)) = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i \stackrel{!}{=} 0$$

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i} \qquad \square$$

1.5 P

3.)

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{10}{21} \approx 0.4762$$

1 P

4.)

$$\bar{v} = \frac{1}{19} \sum_{i=1}^{19} v_i = \frac{53.5}{19} \approx 2.8158 \; [Mio.Euro]$$

0.5 P

5.) $v_{\text{med}} = 0.9$ 0.5 P

6.) $v_{\text{mod}} = 0.9$ 0.5 P

7.)

$$\widehat{s}^2 = \frac{1}{18} \left(\sum_{i=1}^{19} v_i^2 - \frac{1}{19} \left(\sum_{i=1}^{19} v_i \right)^2 \right) = \frac{1}{18} \left(342.77 - \frac{1}{19} 53.5^2 \right) \approx 10.6736$$

1 P

8.)

$$\widehat{VaR}_{0.9} = \min\{v | F_n(v) \geq 0.9)\} = 7.1$$

0.5 P

9.) Ordinales Skalenniveau 1 P

10.) $\mathbb{P}(R_1 = D | R_0 = B) = 0.16$ 1 P

11.)

$$\begin{aligned} \mathbb{P}(R_2 = A | R_0 = B) \quad &= \mathbb{P}(R_2 = A | R_1 = A) \mathbb{P}(R_1 = A | R_0 = B) \\ &+ \mathbb{P}(R_2 = A | R_1 = B) \mathbb{P}(R_1) = B | R_0 = B) \\ &= 0.76 \cdot 0.21 + 0.21 \cdot 0.63 = 0.2919 \end{aligned}$$

1 P

Lösung 3

1. -0.0147 — 1.0 P

2. 231.8548 — 1.0 P

3. 10 — 1.0 P

4. $P(W \geq 6 | W \geq 4) = 0.0176$ — 1.0 P

5. 0.8068 — 1.0 P

6. $w = 3$ — 1.0 P

7. $E(X^2) = Var(X) + E(X)^2 = 0.5 + 2^2 = 4.5$ — 1.0 P

8. $P(\mu - 2\sigma < X < \mu + 2\sigma) \geq 0.75$ — 1.0 P

9. $\alpha\mu_X + (1-\alpha)\mu_Y$ — 0.5 P

10. $\sqrt{\alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 + 2\alpha(1-\alpha)\sigma_X\sigma_Y\rho_{XY}}$
    $\rho_{XY} = -1$ — 1.5 P

Lösung 4

1. ungeordnet kategorial (qualitativ), Nominalskala 0.5 P+0.5 P

2. a) 0.5 P+0.5 P

g_i	30.7	30.9	31.3	31.5	31.7	32.0	32.1	32.2	32.6
$F(g_i)$	0.1	0.2	0.3	0.4	0.5	0.7	0.8	0.9	1.0

   b) 31.3, 32.1 0.5 P+0.5 P

   c) 1.9 0.5 P

   d) 0.6 0.5 P+0.5 P

3. Lage, Streuung 0.5 P+0.5 P

4. a) -1 0.5 P

   b) $t_{0.975;24} = 2.064$ 0.5 P

   c) p-Wert $= 0.5 > \alpha = 0.05$, H_0 nicht ablehnen bei $\alpha = 0.05$ 0.5 P+0.5 P

5. 0.0456 0.5 P+0.5 P

6. 0.9556 0.5 P

7. a) 0.3972 0.5 P

   b) 20.6760 0.5 P

Klausur Statistik (10 ECTS)

Name		**Prüfer**	Prof. Dr. I. Klein
Vorname		**Arbeitszeit**	Mittwoch, 03.08.2016
Matrikelnummer			14:00 - 16:00 Uhr
Studienrichtung		**Sitzplatznummer**	
Semesterzahl		**Raum**	
Email (optional)			

Hinweis: Aufgabenblätter nicht auseinandertrennen!

Ergebnis:

Statistik	
Aufgabe	**Punkte**
1	
2	
3	
4	
Summe	
Note:	

Unterschrift des Kandidaten: ________________________

Unterschrift des Prüfers: ________________________

Hilfsmittel: Es gelten folgende Regelungen zu den erlaubten Hilfsmitteln:

- Nicht programmierbarer Taschenrechner
- Die vom Lehrstuhl seit dem WS 2014/15 offiziell herausgegebene Formelsammlung (DIN A5, gebunden, orangener Umschlag). Es sind prinzipiell keine weiteren Eintragungen oder Markierungen darin erlaubt. Ausgenommen sind farbliche Hinterlegungen von Textpassagen und/oder Formeln. Außerdem ist das Erratum zur 1. Auflage der Formelsammlung erlaubt.
- R Reference Card von Jonathan Baron, es sind keine weiteren Eintragungen oder Markierungen darin erlaubt.

Bewertung: Für jede Aufgabe werden maximal zehn Punkte vergeben. Bewertet werden grundsätzlich nur Lösungen, die im Lösungsteil stehen und für die Folgendes beachtet wird:

- Der Lösungsweg muss aus einer Darstellung der einzelnen Rechenschritte ersichtlich sein.
- Antworten sind stets zu begründen, es sei denn es wird ausdrücklich keine Begründung verlangt.
- Unleserliche Aufgabenteile werden mit 0 Punkten bewertet.

Viel Erfolg!

Aufgabe 1

An einer Universität finden Bauarbeiten statt.

1. Anstehende Klausuren werden in einem anderen Gebäude geschrieben, wenn der Baulärm B längerfristig 65 Dezibel (Abkürzung: dB) überschreitet.

Eine einfache Zufallsstichprobe, bei der an 14 Werktagen jeweils zur selben Uhrzeit Messungen durchgeführt wurden, hat folgendes ergeben:

Messung i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\mathbf{b}_i$	25	85	100	90	108	35	55	48	120	87	91	75	81	65

Nehmen Sie an, die gemessenen Daten stammen aus einer Normalverteilung mit unbekanntem Mittelwert μ und Varianz $\sigma^2 = 484$. Testen Sie die Vermutung
H_1: Mittlerer Baulärm in dB > 65 auf dem 10%-Signifikanzniveau.
Sie wissen, dass $\bar{b}_{14}$=76.07.

a) Berechnen Sie dazu zunächst die Teststatistik.

b) Wie lautet die kritische Schranke?

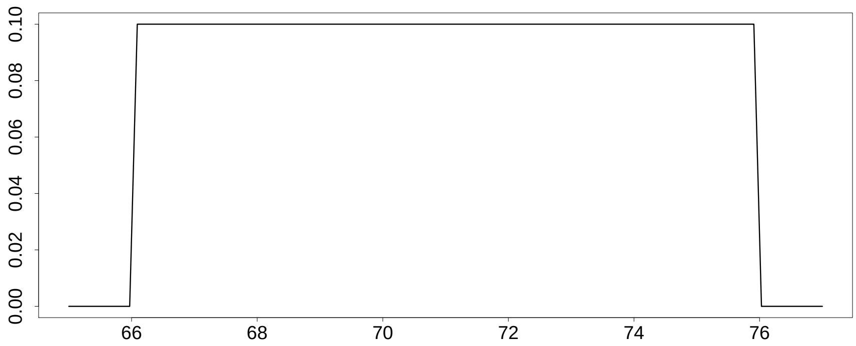
c) Berechnen Sie das rechtsseitige realisierte 90%-Konfidenzintervall.

d) Treffen Sie eine Testentscheidung und begründen Sie diese.
(Hinweis: Sollten Sie in den vorherigen Teilaufgaben keine Lösung erhalten haben, nutzen Sie für die Testentscheidung: p-Wert=0.0299).

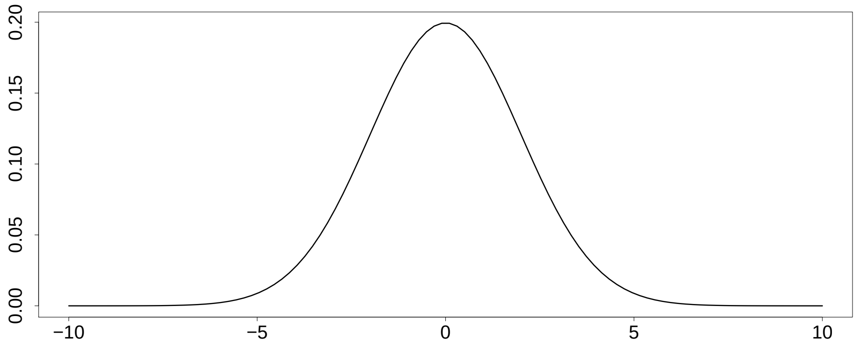
e) Berechnen Sie die Güte des Tests, wenn $\mu_1 = 71$.

2. Welche der folgenden Grafiken (Grafik 1, Grafik 2 oder Grafik 3) entspricht einer Normalverteilung mit Parametern $\mu = 71$ und $\sigma^2 = 484$?

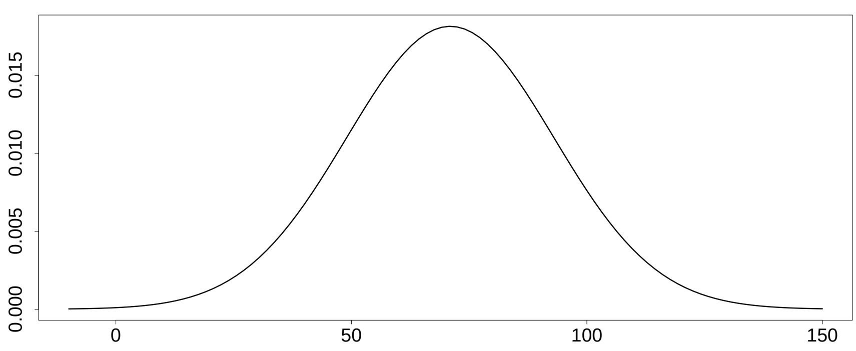
Grafik 1:



Grafik 2:



Grafik 3:



3. Nun wurde der Baulärm [in dB] während einer Woche jeweils vormittags, mittags und nachmittags gemessen.
   Folgende Datenpunkte ergaben sich:

Tageszeit/Baulärm in dB	**0-40**	**41-65**	**>65**	Σ
vormittags	4	2	8	14
mittags	5	8	1	14
nachmittags	7	4	3	14
Σ	16	14	12	42

a) Wie oft wurden mittags Werte über 40 dB gemessen?

Sie interessieren sich dafür, ob der 'Baulärm in dB' unabhängig von der 'Tageszeit' ist. Dazu führen Sie einen χ^2-Unabhängigkeitstest auf einem 5% Signifikanzniveau durch und erstellen folgende Indifferenztabelle:

Tageszeit/Baulärm in dB	**0-40**	**41-65**	**>65**	Σ
vormittags	w)	$4\frac{2}{3}$	4	14
mittags	$5\frac{1}{3}$	x)	4	14
nachmittags	$5\frac{1}{3}$	$4\frac{2}{3}$	y)	14
Σ	16	14	12	z)

Nehmen Sie im Weiteren an, dass die Bedingungen für eine gute Approximation an die χ^2 - Verteilung erfüllt sind.
Zusätzlich haben Sie folgende Information:

$$\Sigma_{i=1}^{k}\Sigma_{j=1}^{l}\frac{(n_{ij}-\frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n}} = 11.375.$$

b) Geben Sie die fehlenden Werte der Indifferenztabelle an.

w)

x)

y)

z)

c) Geben Sie den Wert der kritischen Schranke für den Unabhängigkeitstest an.

Sollten Sie keinen Wert für die kritische Schranke erhalten, verwenden Sie im Folgenden den Wert 11.

d) Können Sie anhand Ihrer Testentscheidung der Aussage 'Der Baulärm ist von der Tageszeit abhängig' zustimmen? Begründen Sie Ihre Antwort.

Cramérs V wurde berechnet und beträgt hier 0.3679.

e) Berechnen Sie die mittlere quadratische Kontingenz.

f) Welchen Wert nimmt die mittlere quadratische Kontingenz bei Unabhängigkeit zweier Merkmale an?

g) Welchen Wert würde die mittlere quadratische Kontingenz bei vollständiger Abhängigkeit der beiden Merkmale 'Baulärm in dB' und 'Tageszeit' annehmen?

Schmierpapier zu Aufgabe 1

Aufgabe 2

Die Zeitdauer T [in Jahren] bis zum Ausfall eines Kredits sei exponentialverteilt mit Parameter λ, d.h. $f(t) = \lambda \exp(-\lambda t)$, wobei $\lambda > 0$.

1. Zeigen Sie, dass die logarithmierte Likelihoodfunktion für n unabhängig und identisch verteilte Zeitdauern bis zum Ausfall gegeben ist durch

$$\ln(L(\lambda; t_1, \ldots, t_n)) = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i.$$

Die oben genannte logarithmierte Likelihoodfunktion ist global konkav.

2. Zeigen Sie, dass der Maximum-Likelihood-Schätzer für λ gegeben ist durch

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i}.$$

Sie haben von 10 Krediten folgende Zeitdauern bis zum Ausfall beobachtet.

i	1	2	3	4	5	6	7	8	9	10
t_i	1.8	2.2	0.5	0.8	3.5	0.9	3.7	0.2	5.1	2.3

3. Berechnen Sie den Schätzwert für $\widehat{\lambda}_{ML}$ konkret.

Für 19 ausgefallene Kredite liegen Beobachtungen über die Zufallsvariable "Höhe des jeweils eingetreten Verlusts" V [in Mio Euro] vor.

i	1	2	3	4	5	6	7	8	9	10
v_i	0.1	0.1	0.2	0.4	0.5	0.6	0.7	0.9	0.9	0.9

i	11	12	13	14	15	16	17	18	19
v_i	1.2	2.5	3.8	4.5	4.9	5.7	6.2	7.1	12.3

Hinweis: $\sum_{i=1}^{19} v_i = 53.5$ und $\sum_{i=1}^{19} v_i^2 = 342.77$

4. Bestimmen Sie das arithmetische Mittel der Beobachtungen aus V.

5. Bestimmen Sie den Schätzwert des Medians der Beobachtungen aus V.

6. Bestimmen Sie den Schätzwert des Modus der Beobachtungen aus V.

7. Bestimmen Sie den Wert des erwartungstreuen Schätzers

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(V_i - \bar{V}_n)^2$$

für die Varianz der Beobachtungen aus V.

Eine wichtige Maßzahl für die Quantifizierung finanzieller Risiken ist der so genannte Value-at-Risk VaR_α. Dieser ist definiert als das α-Quantil der Verteilung der Verluste.

8. Schätzen Sie den $VaR_{0.9}$ als empirisches Quantil aus den obigen Daten.

In einer Bank werden Kreditnehmer bezüglich ihrer Bonität in Ratingklassen A (gut), B (schlecht) und D (default - ausgefallen) eingeteilt, wobei sich die Bonitätseinschätzung im Lauf der Zeit ändern kann. Die folgende Tabelle zeigt **bedingte** Wahrscheinlichkeiten $P(R_{t+1}|R_t)$ für die Veränderungen der Zufallsvariable R ("Ratingklasse") eines einzelnen Kreditnehmers innerhalb eines Ein-Jahres-Zeithorizonts $[t, t+1)$. Die Veränderungen sollen dabei unabhängig vom konkreten Zeitpunkt t als konstant angenommen werden.

		$t+1$		
		A	B	D
	A	0.76	0.23	0.01
t	B	0.21	0.63	0.16
	D	0	0	1

9. Welches maximal zulässige Skalenniveau besitzt die Zufallsvariable "Ratingklasse"?

10. Wie groß ist die Wahrscheinlichkeit $P(R_1 = D|R_0 = B)$, d.h. dass ein Kreditnehmer innerhalb des Jahres $[0, 1)$ aus Ratingkategorie B ausfällt?

11. Wie groß ist die Wahrscheinlichkeit $P(R_2 = A|R_0 = B)$, d.h. dass ein Kreditnehmer innerhalb des Zwei-Jahres-Zeitraums $[0, 2)$ von Ratingkategorie B in Klasse A aufsteigt?

Schmierpapier zu Aufgabe 2

Aufgabe 3

Teil I:

1. Sie investieren 100 EUR in Aktie A. Der Aktienkurs besitzt im ersten Jahr eine Wachstumsrate von 3%. Nach drei Jahren hat die Aktie wieder ihren Anfangswert. Die Wachstumsraten im zweiten und dritten Jahr sind identisch. Berechnen Sie die Wachstumsrate im zweiten bzw. dritten Jahr.

2. Sie investieren 200 EUR in Aktie B. Der Aktienkurs hat eine durchschnittliche jährliche Wachstumsrate von 3 Prozent. Berechnen Sie den Wert nach 5 Jahren.

3. Sie entscheiden sich für eine Investition in eine Aktie. Sie investieren im ersten Jahr 100 EUR und im zweiten Jahr 200 EUR. Im ersten Jahr kostet eine Aktie 20 EUR. Der durchschnittliche Preis pro Aktie beträgt über alle 2 Jahre hinweg 12 EUR. Berechnen Sie den Preis für eine Aktie im zweiten Jahr.

Teil II:

Ein Anleger hätte gerne ein genaueres Verständnis bezüglich des Risikos seiner Anlage. Die Zufallsvariable W beschreibt die jährliche Rendite, wobei W normalverteilt ist mit den Parametern $\mu_W = 3$ und $\sigma_W^2 = 1.21$.

4. Bestimmen Sie die Wahrscheinlichkeit, dass die Rendite W mindestens 6 ist, unter der Bedingung, dass die Rendite mindestens 4 beträgt.

5. Mit welcher Wahrscheinlichkeit liegt die Rendite W zwischen 0.5 und 4?

6. Welche Rendite w wird mit einer Wahrscheinlichkeit von 50% maximal erreicht?

Teil III:

Im Folgenden beschreibt die Zufallsvariable Z die jährliche Rendite von einer Aktie, wobei Z den Erwartungswert $\mu_Z = 2$ und Varianz $\sigma_Z^2 = 0.5$ hat. Die Verteilung von Z ist unbekannt.

7. Berechnen Sie das zweite Potenzmoment von Z.

8. Wie hoch ist die Wahrscheinlichkeit mindestens, dass Z im 2-fachen Schwankungsintervall liegt?

Teil IV:

Aufgrund von Risikodiversifizierung wird im Folgenden ein Portfolio aus Aktien betrachtet. Die Zufallsvariable X beschreibt die Rendite von Aktie 1, wobei X den Erwartungswert μ_X und Varianz σ_X^2 besitzt. Die Rendite von Aktie 2 wird durch die Zufallsvariable Y mit Erwartungswert μ_Y und Varianz σ_Y^2 definiert. Die Korrelation zwischen den Zufallsvariablen X und Y beträgt ρ_{XY}. Die Rendite des Portfolios P wird durch folgende Zufallsvariable definiert:

$$P = \alpha X + (1 - \alpha)Y, \text{ wobei } \alpha \in [0, 1].$$

9. Bestimmen Sie den Erwartungswert von P in Abhängigkeit von α, μ_X, μ_Y.

10. Bestimmen Sie die Standardabweichung von P in Abhängigkeit von $\alpha, \sigma_X, \sigma_Y, \rho_{XY}$. Für welchen Wert von ρ_{XY} wird die Standardabweichung minimal?

Schmierpapier zu Aufgabe 3

Aufgabe 4

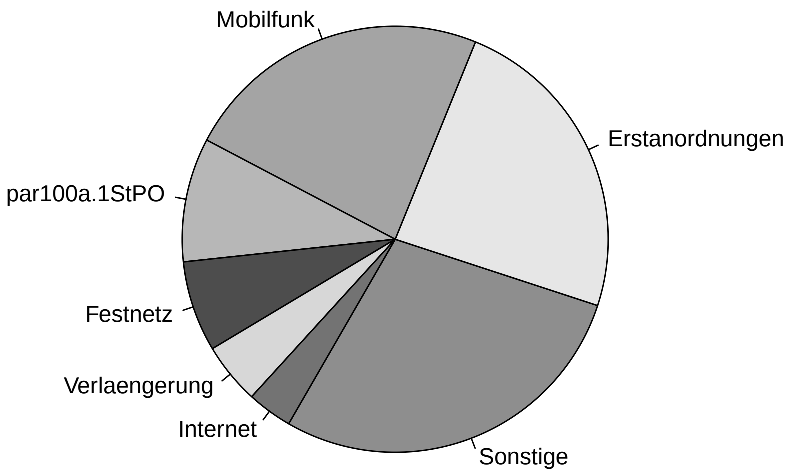
In Ihrem Workspace liegen zwei Dataframes `TKU.2013` und `TKU.2014`, welche die Daten aus den Telekommunikationsüberwachungsberichten des Bundesamtes für Justiz aus den Jahren 2013 und 2014 enthalten[1]. In beiden Dataframes sind die Anzahlen an Überwachungen nach Art und Bundesland aufgeschlüsselt, woraus sich folgende Merkmale ergeben:

Ausprägung Merkmal Überwachung	**Wert**	**Spaltenname**
Verfahren nach § 100a Abs. 1 StPO:	Anzahl	`par100a.1StPO`
Erstanordnungen:	Anzahl	`Erstanordnungen`
Verlängerungsanordnungen:	Anzahl	`Verlaengerung`
Festnetztelekommunikation:	Anzahl	`Festnetz`
Mobilfunktelekommunikation:	Anzahl	`Mobilfunk`
Internettelekommunikation:	Anzahl	`Internet`
Sonstige:	Anzahl	`Sonstige`

Weitere Merkmale	**Wert**	**Spaltenname**
Neues Bundesland:	1 wenn neues Bundesland	`neuebundeslaender`

Die folgende Grafik enthält ein Tortendiagramm der Überwachungen in Bayern aus dem Jahr 2013, das durch den Befehl erstellt wurde:

```
> pie(TKU.2013["BY", ])
```



[1] https://offenedaten.de/dataset/telekommunikationsuberwachung

1. Welche drei Merkmalsausprägungen kamen in Bayern im Jahr 2013 zusammen in knapp 75% der Fälle vor?

2. Gab es mehr oder weniger als 2000 sonstige Überwachungen in Bayern, wenn es in diesem Bundesland zu insgesamt 12213 Überwachungen im Jahr 2013 kam? **Mit Begründung!**

Sie finden zunächst an Ihrem Arbeitsplatz folgenden R-Output vor:

```
> str(TKU.2013)

'data.frame':         16 obs. of  8 variables:
 $ par100a.1StPO     : num  623 1149 465 104 33 ...
 $ Erstanordnungen   : num  1835 2917 1696 261 263 ...
 $ Verlaengerung     : num  358 568 350 28 48 137 659 79 246 452 ...
 $ Festnetz          : num  357 841 123 44 28 115 548 112 262 295 ...
 $ Mobilfunk         : num  2000 2863 1903 268 282 ...
 $ Internet          : num  84 419 20 2 10 ...
 $ Sonstige          : num  2292 3456 2036 289 317 ...
 $ neuebundeslaender: num  0 0 0 1 0 0 0 1 0 0 ...

> rownames(TKU.2013)

 [1] "BW" "BY" "BE" "BB" "HB" "HH" "HE" "MV" "NI" "NW" "RP" "SL" "SN" "ST" "SH"
[16] "TH"

> neubund = c("BB","MV","SN","ST","TH")
> altbund = rownames(TKU.2013[which(TKU.2013$neuebundeslaender == 0),])
```

Im Vektor `neubund` sind die Kürzel der neuen Bundesländer abgespeichert.

3. Welcher Datentyp wäre für die Spalte `neuebundeslaender` ebenfalls geeignet gewesen?

4. Wie viele Überwachungen im Bereich *Mobilfunktelekommunikation* gab es 2013 im Bundesland Baden-Württemberg (BW)?

5. Wie oft steht in der Spalte `neuebundeslaender` der Wert 0?

Betrachten Sie nun folgenden `R`-Output:

```
> c(sum(TKU.2013[neubund,-8]),sum(TKU.2014[neubund,-8]))

[1]  9982 10521

> c(sum(TKU.2013[altbund,-8]),sum(TKU.2014[altbund,-8]))

[1] 68314 70631
```

6. Berechnen Sie jeweils die prozentuale Zunahme der Anzahl von Überwachungen von 2013 auf 2014 in den alten sowie den neuen Bundesländern.

7. Liegt die prozentuale Zunahme an Überwachungen in ganz Deutschland von 2013 auf 2014 zwischen, über oder unter den Werten aus Teilaufgabe 6.?

8. Geben Sie den `R`-Befehl für das Erstellen eines logischen Vektors `mehr` an, der für jedes Bundesland den Wert `TRUE` enthält, falls die Anzahl der Überwachungen im Bereich Festnetztelekommunikation im Jahr 2014 höher waren als im Jahr 2013.

Betrachten Sie nun folgenden `R`-Output:

```
> tab = table(mehr,TKU.2014$neuebundeslaender)
> addmargins(tab)/sum(tab)

mehr          0      1    Sum
  FALSE 0.3750 0.1250 0.5000
  TRUE  0.3125 0.1875 0.5000
  Sum   0.6875 0.3125 1.0000
```

9. Wie viele Bundesländer hatten im Jahr 2014 mehr Überwachungen im Bereich der Festnetztelekommunikation als noch 2013?

10. Wie viel Prozent der relevanten Bundesländer aus Teilaufgabe 9. sind neue Bundesländer?

11. Geben Sie den `R`-Befehl an um den Teildatensatz des Jahres 2014 zu erhalten, welcher nur Bundesländer mit höchstens 100 Verlängerungsanordnungen, aber mehr als 20 Internettelekommunikations-Überwachungen enthält.

12. Vervollständigen Sie folgenden `R`-Code, um nachstehende log-Likelihoodfunktion der Exponentialverteilung in `R` zu berechnen:

$$\ln(L(\lambda; x_1, ..., x_n)) = \sum_{i=1}^{n} \ln f_{\exp}(x_i; \lambda)$$

```
LL_exp = function(     ,data){
    sum(       (       (        ,        )        )        )
}
```

Schmierpapier zu Aufgabe 4

Lösung Klausur SS16 (10 ECTS)

Lösung 1

1. a) $t_n = \sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma} = 1.8829$ 0.5 P
   b) $\lambda_{0.90} = 1.2815$ 0.5 P
   c) $[76.07 - 1.2815\frac{22}{\sqrt{14}}; \infty) = [68.5362; \infty)$ 1.0 P
   d) 1.88>1.28; H_0 ablehnen 0.5 P
   e) 1- $\Phi(1.28 + \sqrt{14}\frac{65-71}{22}$)=0.3974 1.0 P

2. Grafik 3 0.5 P

3. a) 8+1=9 0.5 P
   b) w) $n_{11} = 5\frac{1}{3}$, 0.5 P
      x) $n_{22} = 4\frac{2}{3}$, 0.5 P
      y) $n_{33} = 4$, 0.5 P
      z) $N = 42$ 0.5 P
   c) $\chi^2_{0.95;2*2=4} = 9.49$ 0.5 P
   d) $\chi^2 = 11.375 > 9.49$ (oder >11); H_0 ablehnen bei $\alpha = 0.05$; 0.5 P
      Man kann der Aussage zustimmen. 0.5 P
   e) $(0.3679)^2 * 2 = 0.2707$ 1.0 P
   f) 0 0.5 P
   g) Maximalwert: $\min\{k-1; l-1\}$ =2 0.5 P

Lösung 2

1.)

$$L(\lambda; t_1, \dots, t_n) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)$$

$$\Rightarrow \ln(L(\lambda; t_1, \dots, t_n)) = \ln(\lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i)) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} t_i \qquad \square$$

1.5 P

2.)

$$\widehat{\lambda}_{ML} \Leftrightarrow \frac{\partial}{\partial \lambda} \ln(L(\lambda; t_1, \dots, t_n)) = 0$$

$$\frac{\partial}{\partial \lambda} \ln(L(\lambda; t_1, \dots, t_n)) = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i \stackrel{!}{=} 0$$

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i} \qquad \square$$

1.5 P

3.)

$$\widehat{\lambda}_{ML} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{10}{21} \approx 0.4762$$

1 P

4.)

$$\bar{v} = \frac{1}{19} \sum_{i=1}^{19} v_i = \frac{53.5}{19} \approx 2.8158 \; [Mio.Euro]$$

0.5 P

5.) $v_{\text{med}} = 0.9$ 0.5 P

6.) $v_{\text{mod}} = 0.9$ 0.5 P

7.)

$$\widehat{s}^2 = \frac{1}{18} \left(\sum_{i=1}^{19} v_i^2 - \frac{1}{19} \left(\sum_{i=1}^{19} v_i \right)^2 \right) = \frac{1}{18} \left(342.77 - \frac{1}{19} 53.5^2 \right) \approx 10.6736$$

1 P

8.)

$$\widehat{VaR}_{0.9} = \min\{v | F_n(v) \geq 0.9)\} = 7.1$$

0.5 P

9.) Ordinales Skalenniveau 1 P

10.) $\mathbb{P}(R_1 = D | R_0 = B) = 0.16$ 1 P

11.)

$$\begin{aligned} \mathbb{P}(R_2 = A | R_0 = B) \quad &= \mathbb{P}(R_2 = A | R_1 = A)\mathbb{P}(R_1 = A | R_0 = B) \\ &+\mathbb{P}(R_2 = A | R_1 = B)\mathbb{P}(R_1) = B | R_0 = B) \\ &= 0.76 \cdot 0.21 + 0.21 \cdot 0.63 = 0.2919 \end{aligned}$$

1 P

Lösung 3

1. -0.0147 1.0 P

2. 231.8548 1.0 P

3. 10 1.0 P

4. $P(W \geq 6 | W \geq 4) = 0.0176$ 1.0 P

5. 0.8068 1.0 P

6. $w = 3$ 1.0 P

7. $E(X^2) = Var(X) + E(X)^2 = 0.5 + 2^2 = 4.5$ 1.0 P

8. $P(\mu - 2\sigma < X < \mu + 2\sigma) \geq 0.75$ 1.0 P

9. $\alpha\mu_X + (1-\alpha)\mu_Y$ 0.5 P

10. $\sqrt{\alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 + 2\alpha(1-\alpha)\sigma_X\sigma_Y\rho_{XY}}$
    $\rho_{XY} = -1$ 1.5 P

Lösung 4

1. Erstanordnungen, Mobilfunktelekommunikation und Sonstige. 0.5 P
2. Mehr, da Sonstige ca ein Viertel der Fläche belegt, also ca. 3000 0.5 P + 0.5 P
3. `factor` 0.5 P
4. 2000 0.5 P
5. 11 0.5 P
6. 5.3997% bzw. 3.3917% 0.5 Ansatz, 0.5 P + 0.5 P Ergebnis
7. Zwischen den beiden Werten aus TA 6. 0.5 P
8. `mehr = TKU.2014$Festnetz > TKU.2013$Festnetz` 1.5 P (-0.5 P pro Fehler)
9. 8 0.5 P
10. 37.5% 0.5 P
11. Z.B.
```
TKU.2014[TKU.2014$Verlaengerung <= 100 &
         TKU.2014$Internet > 20, ]
```
1.5 P (-0.5 P pro Fehler)
12.
```
LL_exp = function(lambda,data){
    sum(log(dexp(data,lambda)))
}
```
1 P (-0.5 P pro Fehler)