

## Bachelorprüfung: Mathematik 4 - Statistik (2 Stunden)

### Bemerkungen:

- Es sind alle mitgebrachten schriftlichen Hilfsmittel und der Taschenrechner erlaubt.
- Natels sind auszuschalten!
- Lesen Sie zuerst alle Aufgaben durch! Verweilen Sie nicht zu lange bei einem Aufgabenteil, der Ihnen grosse Schwierigkeiten bereitet! Für die Note 6 brauchen nicht alle Aufgaben gelöst zu sein!
- Wenn nicht anders vermerkt, sind die Tests auf dem 5%-Niveau durchzuführen.
- Die nötigen Tabellen befinden sich auf den hintersten Seiten dieser Prüfung.
- Aufgaben 4 und 5 sind Multiple-Choice Aufgaben. Es ist jeweils genau eine Antwort korrekt. Eine korrekte Antwort gibt 1 **Plus**punkt und eine falsche Antwort  $\frac{1}{2}$  **Minus**punkt. Minimal erhält man für eine ganze Multiple-Choice Aufgabe 0 Punkte. Tragen Sie die korrekten Antworten der Multiple Choice Aufgaben mit Kreuzchen in das separate Antwortblatt ein.

**Viel Erfolg!**

### 1. (12 Punkte)

Corinne und Markus sind grosse Schlittel-Fans und möchten nun endlich klären, wer von ihnen schneller ist. Dazu gehen sie jeden Dienstag auf die Rigi schlitteln und messen die Zeit (in Sekunden), die sie für eine Abfahrt benötigen.

Daten:

Datum	10.1.06	17.1.06	24.1.06	31.1.06	7.2.06	14.2.06
Corinne	303	295	378	350	278	312
Markus	318	305	396	345	301	329
Differenz $d$	-15	-10	-18	5	-23	-17

Kennzahlen:

Corinne	$\bar{x}$	=	319.33	$\hat{\sigma}_x$	=	37.43
Markus	$\bar{y}$	=	332.33	$\hat{\sigma}_y$	=	35.10
Differenz	$\bar{d}$	=	-13.00	$\hat{\sigma}_d$	=	9.78

- Wieso ist es hier sinnvoll, mit gepaarten Stichproben zu arbeiten?
- Gib die Nullhypothese und die Alternativhypothese an. Handelt es sich hier um ein einseitiges oder zweiseitiges Testproblem?
- Berechne die Teststatistik des gepaarten t-Tests.
- Gib den Verwerfungsbereich für den t-Test an.
- Wie entscheidet der t-Test?

- f) Was für ein Testergebnis würdest Du bei einem ungepaarten t-Test erwarten? Begründe ohne den Test explizit durchzuführen.
- g) Welche Voraussetzungen macht der t-Test?
- h) Berechne auch den p-Wert des Vorzeichen-tests.

## 2. (6 Punkte)

Es soll untersucht werden, wieviele Felchen (eine Fischart) im Zürichsee leben. Es können nicht alle Felchen gefangen werden für diese Zählung. Daher kann man folgende Methode zur Schätzung der Anzahl Felchen benutzen:

Es werden 1000 Felchen auf zufällige Art gefangen. Diese Fische werden markiert und wieder in den See gesetzt. Nach einer gewissen Zeit werden immer wieder zufällig Felchen gefangen. Einige dieser gefangenen Felchen sind markiert, da sie bereits beim ersten Fang dabei waren. Jede gefangene Felche wird unmittelbar nach der Untersuchung wieder lebend in den See gesetzt. Der Anteil von markierten Felchen soll Aufschluss über die gesamte Anzahl Felchen geben. (Hinweis: Man nimmt immer vereinfachend an, dass keine Felchen im See verschwinden oder hinzukommen)

- a) Wenn die gesamte Anzahl Felchen im See 5000 beträgt, wie gross ist dann die Wahrscheinlichkeit  $p$ , dass eine zufällig gefangene Felche eine Markierung trägt?
- b) In Wahrheit weiss man nicht, wieviele Felchen sich im See befinden. Von 100 zufällig gefangenen Felchen sind 25 markiert. Gib eine sinnvolle (Punkt-) Schätzung für die Anzahl Felchen im Zürichsee an.
- c) Gib ein 95% Vertrauensintervall für die Anzahl Felchen an; basierend auf den Beobachtungen in Teilaufgabe b). Verwende hierfür eine geeignete Approximation. (Hinweis: Berechne zuerst das Vertrauensintervall für die Wahrscheinlichkeit, dass eine zufällig gefangene Felche eine Markierung trägt.)

## 3. (9 Punkte)

Die Fahrt mit dem Tram kostet 1 Franken. Wer kein Ticket löst und beim Schwarzfahren erwischt wird, zahlt eine gewisse Strafe (die den regulären Fahrpreis schon beinhaltet; es ist also nur die Strafe zu entrichten). Es wird im Folgenden angenommen, dass jede Person auf jeder Fahrt zufällig mit einer Wahrscheinlichkeit  $p = 0.02$  kontrolliert wird.

- a) Wie hoch müsste die Strafe mindestens gewählt werden, dass Schwarzfahren im Schnitt genauso teuer wird wie reguläres Lösen von Tickets?
- b) Wie gross ist die Wahrscheinlichkeit, auf 10 Fahrten mit dem Tram nie kontrolliert zu werden? Wie gross ist die Wahrscheinlichkeit, zweimal oder häufiger kontrolliert zu werden?
- c) Christoph denkt sich folgende Strategie aus für seine nächsten 3 Fahrten mit dem Tram: Er fährt schwarz, bis er das erste Mal erwischt wird. Falls er erwischt wird, so löst er für alle folgenden Fahrten wieder regulär ein Ticket. Wie hoch ist der Erwartungswert der Kosten für seine 3 Fahrten? Verwende als Strafe 100 Franken und als Ticketpreis 1 Franken. Hinweis: Überlege Dir z.B. alle möglichen Fälle, die bei 3 Fahrten eintreten können.
- d) Christoph fährt auf 100 Fahrten schwarz. Die Anzahl Kontrollen, in die Christoph auf allen 100 Fahrten gerät, wird mit  $X$  bezeichnet. Welche Verteilung hat  $X$ ? Gib auch die Parameter der Verteilung an. Mit welcher *diskreten* Verteilung kann diese Verteilung gut angenähert werden? Gib auch hier den Parameter der Verteilung an.

## 4. (9 Punkte)

Lukas ist auf der Suche nach einer neuen Wohnung. Dazu analysiert er das Wohnungsangebot in Zürich mit einer linearen Regression. Als Erklärende  $x$  wählt er die Wohnungsgrösse in  $m^2$  und die Zielvariable  $y$  ist der Mietpreis. Das Modell ist

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d. .}$$

Hier sind der R-Output und Plots:

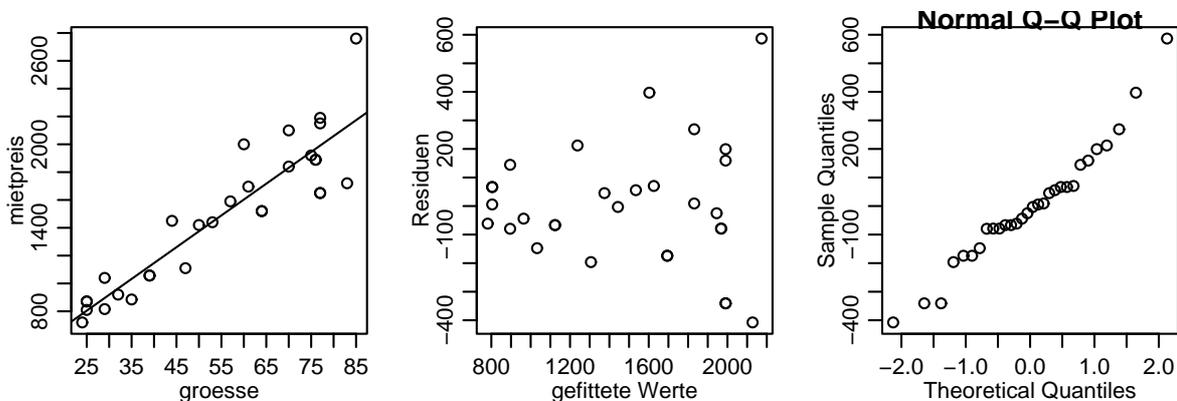
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	??	??	2.075	0.0473
x	22.81	1.93	??	2.13e-12

Residual standard error: 213.6 on 28 degrees of freedom

Multiple R-Squared: 0.833, Adjusted R-squared: 0.8271

F-statistic: 139.7 on 1 and 28 DF, p-value: 2.131e-12



- Mit wievielen Beobachtungen wurde die Regression berechnet?
  - 22
  - 26
  - 27
  - 28
  - 29
  - 30
- Wie gross ist der Intercept  $\hat{\beta}_0$ ?
  - 0
  - 234
  - 728
  - 1485
- Wie gross ist die t-Teststatistik für den Test der Nullhypothese  $H_0 : \beta_1 = 0$  (die Alternativhypothese ist  $H_A : \beta_1 \neq 0$ )?
  - 1.93
  - 11.819
  - 22.81
  - 139.7
  - 43.87
  - Keine Aussage möglich
- Wird  $H_0 : \beta_1 = 0$  auf dem 5% Niveau verworfen (die Alternative ist  $H_A : \beta_1 \neq 0$ )?
  - Ja
  - Nein
  - Keine Aussage möglich
- Welches der folgenden Intervalle ist ein exaktes zweiseitiges 95% Vertrauensintervall für  $\beta_1$ ?
  - $22.81 \pm 1.96 \cdot \frac{1.93}{\sqrt{28}}$
  - $22.81 \pm 2.048 \cdot \frac{1.93}{\sqrt{28}}$
  - $22.81 \pm 1.96 \cdot 1.93$
  - $22.81 \pm 2.048 \cdot 1.93$
- Lukas hätte gerne eine  $70m^2$  grosse Wohnung. Welchen Preis kann er im Mittel erwarten?
  - $70 \cdot 22.81$
  - $70 \cdot 22.81 + \hat{\beta}_0$
  - $70 \cdot 22.81 - \hat{\beta}_0$
  - $139.7 + 70 \cdot 22.81$
- Betrachte die gezeigten Plots. Welche der nachfolgenden Aussagen ist zutreffend?
  - Die Modellannahmen über die Fehler scheinen plausibel.

- b) Die Normalitätsannahme der Fehler  $\epsilon_i$  scheint grob verletzt zu sein.  
 c) Die Annahme der konstanten Varianz der Fehler  $\epsilon_i$  scheint verletzt zu sein.  
 d) Die Normalitätsannahme der Fehler  $\epsilon_i$  und die Annahme der konstanten Varianz der Fehler  $\epsilon_i$  scheinen grob verletzt zu sein.
- 8) Was passiert mit der Regression, wenn an der Stelle (80, 3000) noch eine Beobachtung hinzugefügt wird?
- a)  $\hat{\beta}_1$  wird kleiner,  $\hat{\sigma}$  wird kleiner.  
 b)  $\hat{\beta}_1$  wird kleiner,  $\hat{\sigma}$  wird grösser.  
 c)  $\hat{\beta}_1$  wird grösser,  $\hat{\sigma}$  wird kleiner.  
 d)  $\hat{\beta}_1$  wird grösser,  $\hat{\sigma}$  wird grösser.  
 e) Keine Aussage möglich.
- 9) Was passiert mit der t-Teststatistik für  $\beta_1$ , wenn an der Stelle (80, 1000) noch eine Beobachtung hinzugefügt wird?
- a) Sie wird kleiner  
 b) Sie bleibt gleich  
 c) Sie wird grösser  
 d) Keine Aussage möglich

### 5. (6 Punkte)

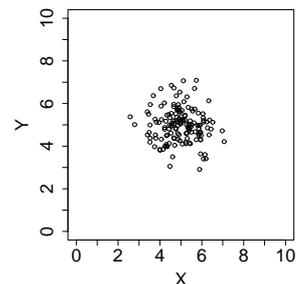
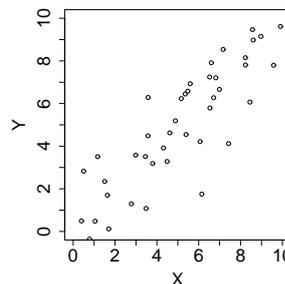
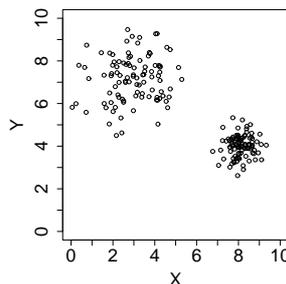
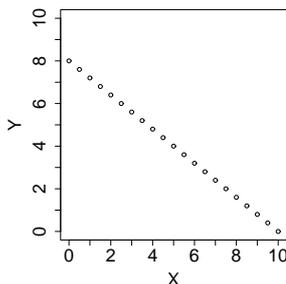
Verschiedene Aufgaben aus mehreren Gebieten:

- 1)  $X$  und  $Y$  seien unabhängige Zufallsvariablen mit

$$\mathbf{E}[X] = 10, \mathbf{E}[Y] = 15 \text{ und } \text{Var}(X) = 5, \text{Var}(Y) = 3.$$

Wie gross ist die Varianz von  $3X - 2Y$ ?

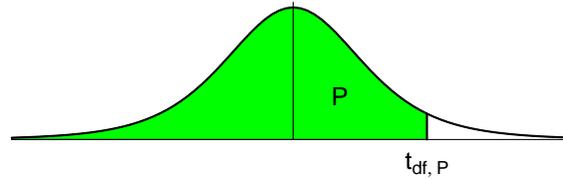
- a) 57  
 b) 33  
 c) 9  
 d) 21
- 2) In einer Schulklasse mit 23 Schülern wird für jeden Schüler die Note mit einem Würfel vergeben. Mit welcher Verteilung lässt sich die Anzahl Schüler mit der Note 6 beschreiben?
- a) Uniforme Verteilung  
 b) Poisson-Verteilung  
 c) Binomial-Verteilung  
 d) Normal-Verteilung
- 3) Die tägliche Überzeit (in Minuten) einer Angestellten einer Grossbank sei gut mit einer Exponential-Verteilung mit Erwartungswert 10 modellierbar. Welche Verteilung eignet sich (näherungsweise) für die Beschreibung der totalen jährlichen Überzeit, wenn die Überzeit an den einzelnen Tagen als unabhängig betrachtet werden kann?
- a) Uniforme Verteilung  
 b) Exponential-Verteilung  
 c) Binomial-Verteilung  
 d) Normal-Verteilung
- 4) Gegeben seien folgende 4 Streudiagramme von  $Y$  gegen  $X$ . Finde die dazugehörigen Korrelationen zwischen  $X$  und  $Y$  (gleiche Reihenfolge wie in den Plots!).
- a)  $-1, -0.02, 0.85, -0.01$   
 b)  $-0.8, -0.85, 0.85, -0.01$   
 c)  $1, -0.01, 0.3, 0.01$   
 d)  $-1, -0.85, 0.85, -0.01$



- 5) In einem Experiment soll der Eisengehalt [mg/kg] in Sojabohnen bestimmt werden. Aus Erfahrung weiss man, dass eine Normalverteilung mit bekannter Varianz gut als Modell für den Eisengehalt passt. Basierend auf 20 Beobachtungen wird ein zweiseitiges 95%-Vertrauensintervall berechnet. Dem Lebensmittelingenieur ist dies jedoch zu ungenau. Er möchte, dass das 95%-Vertrauensintervall halb so gross ist. Wie viele Beobachtungen sind dazu insgesamt nötig?
- a) 10
  - b) 40
  - c) 80
  - d) 60
- 6) Der Praktikant, der mit der Auswertung des Experimentes aus 5) (basierend auf 20 Beobachtungen) beauftragt wurde, liefert Dir das zweiseitige 95%-Vertrauensintervall  $[90 - 5, 90 + 5]$ . Dein Auftraggeber möchte nun aber auf einmal auf dem 5%-Niveau die Nullhypothese  $H_0 : \mu = 100$  gegen  $H_A : \mu > 100$  testen. Wie entscheidest Du?
- a) Die Nullhypothese wird verworfen
  - b) Die Nullhypothese wird beibehalten
  - c) Keine Aussage möglich



## Perzentile der t-Verteilung



Bsp.:  $t_{9; 0.975} = 2.262$

$df$	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
31	0.255	0.530	0.853	1.309	1.696	2.040	2.452	2.744
32	0.255	0.530	0.853	1.309	1.694	2.037	2.449	2.738
33	0.255	0.530	0.853	1.308	1.693	2.035	2.445	2.733
34	0.255	0.529	0.852	1.307	1.691	2.032	2.441	2.728
35	0.255	0.529	0.852	1.306	1.690	2.030	2.438	2.724
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
$\infty$	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576