

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

### **Bemerkungen:**

- Es sind alle 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!
- 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. (8.5 Punkte)

Martin liest leidenschaftlich gerne Bücher. Seine Bücher kauft er mehrheitlich in den Buchhandlungen A und B. Er hat den Verdacht, dass die Buchhandlung A im Allgemeinen Taschenbücher teurer anbietet. Um dies zu prüfen geht er zuerst in die Buchhandlung A, wählt dort zufällig 15 Taschenbücher aus und notiert sich deren Preise. Danach geht er in die Buchhandlung B, wählt dort wiederum zufällig 15 Taschenbücher aus und notiert sich die Preise.

Die Preise (in CHF) der Taschenbücher aus Buchhandlung A sind  $X_1, \dots, X_{15}$ , mit Durchschnitt  $\bar{X}$  und Varianz  $\hat{\sigma}_X^2$ . Entsprechend werden die Preise der Taschenbücher aus Buchhandlung B mit  $Y_1, \dots, Y_{15}$  bezeichnet, mit Durchschnitt  $\bar{Y}$  und Varianz  $\hat{\sigma}_Y^2$ . Ferner nehmen wir an, die Messdaten seien jeweils unabhängig voneinander und normalverteilt: die  $X_i$  als  $\mathcal{N}(\mu_X, \sigma^2)$  und die  $Y_i$  als  $\mathcal{N}(\mu_Y, \sigma^2)$ . Martin berechnet folgende Werte:

$\bar{X}$	18
$\hat{\sigma}_X^2$	12.3
$\bar{Y}$	15.2
$\hat{\sigma}_Y^2$	11.5

- Muss Martin einen gepaarten oder einen ungepaarten Test durchführen? (Begründen Sie!)
- Geben Sie die Null- und Alternativhypothese an.
- Führen Sie einen t-Test auf dem 5% Niveau durch und formulieren Sie den Testentscheid.
- Berechnen Sie das einseitige 95%-Vertrauensintervall für die Differenz  $\mu_X - \mu_Y$ .

Martin ist nicht ganz überzeugt von seinem Experiment. Deshalb geht er nochmals in beide Buchhandlungen. Diesesmal notiert er sich die Preise von 15 Taschenbücher aus der Buchhandlung A und die Preise derselben Taschenbücher aus Buchhandlung B.

- Bilden die Messungen in diesem Experiment eine gepaarte oder eine ungepaarte Stichprobe? (Begründen Sie!)
- Geben Sie die Teststatistik des geeigneten Tests an.

## 2. (9 Punkte)

Das Unternehmen *Fine Waters* stellt das Mineralwasser *Fresh* her. Bei der maschinellen Abfüllung und Etikettierung der Flaschen kommt es immer wieder vor, dass einige Flaschen kleine Mängel aufweisen (z.B. falsches Abfüllvolumen oder nicht lesbare Etiketten). Solche Flaschen können nicht verkauft werden und müssen weggeworfen werden.

Nehmen Sie an, dass eine Flasche mit einer Wahrscheinlichkeit von 10% einen Mangel aufweist.

- a) Wie gross ist die Wahrscheinlichkeit, dass von 5 produzierten Flaschen genau drei einen Mangel aufweisen?
- b) Wie gross ist die Wahrscheinlichkeit, dass von 5 produzierten Flaschen die ersten drei einen Mangel aufweisen?

Ein Ingenieur des Unternehmens möchte wissen, wie effizient die maschinelle Abfüllung und Etikettierung funktioniert. Dazu nimmt er eine Stichprobe von 150 Flaschen.

- c) Welche Verteilung besitzt die Anzahl der Flaschen mit Mangel? Berechnen Sie den Erwartungswert und die Varianz dieser Verteilung.
- d) Verwenden Sie die Normal-Approximation zur Berechnung der Wahrscheinlichkeit, dass 15 bis 20 Flaschen weggeworfen werden müssen.

Der Ingenieur untersucht die 150 Flaschen. Er stellt fest, dass 20 Flaschen einen Mangel aufweisen. Ist diese Häufigkeit vereinbar mit der bisher angenommenen Hypothese, dass durchschnittlich 10% der Flaschen einen Mangel aufweisen?

- e) Führen Sie dazu einen zweiseitigen Test auf dem 5% Niveau durch. Geben Sie die Null- und Alternativhypothese, die Teststatistik sowie den Testentscheid an. Benutzen Sie die Normal-Approximation.

### 3. (8 Punkte)

Ein Sportartikelgeschäft beschäftigt drei MitarbeiterInnen, welche die Kunden beim Kauf beraten sollen.

- Frau Alleskönnerin kann 95% aller Kunden zufriedenstellend beraten.
- Herr Besserwisser gibt in 90% der Fälle eine nützliche Beratung ab.
- Frau Chancenlos kann nur 70% aller Kunden hilfreich beraten.

Es ist gleichwahrscheinlich, von einem der drei Mitarbeiter bedient zu werden. Beantworten Sie folgende Fragen:

- a) Wie gross ist die Wahrscheinlichkeit, dass ein Kunde mit der Beratung, die er erhält, nicht zufrieden ist?
- b) Wie gross ist die Wahrscheinlichkeit, dass der Kunde zufrieden ist, wenn wir wissen, dass er nicht von Frau Alleskönnerin bedient wurde?

Michael ist ein grosser Sportfan. Beim Einkauf in einem Sportartikelgeschäft geht er nach einer speziellen Strategie vor. Ist er mit der Beratung eines Mitarbeiters unzufrieden, wechselt er zum nächsten Mitarbeiter. Beantworten Sie folgende Fragen:

- c) Wie gross ist die Wahrscheinlichkeit, dass Michael unzufrieden das Geschäft verlässt? (unzufriedene Beratung bei allen drei MitarbeiterInnen)
- d) Wie gross ist die Wahrscheinlichkeit, dass Michael nach zwei Beratungen das Geschäft verlässt?

#### 4. (9 Punkte)

Fabio will ein Reisebüro eröffnen. Dazu analysiert er das Angebot an Büroräumlichkeiten in Zürich mit einer linearen Regression. Als erklärende Variable  $x$  wählt er die Bürogrösse in  $m^2$  und die Zielvariable  $y$  ist der Mietpreis in  $SFr.$ . 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:

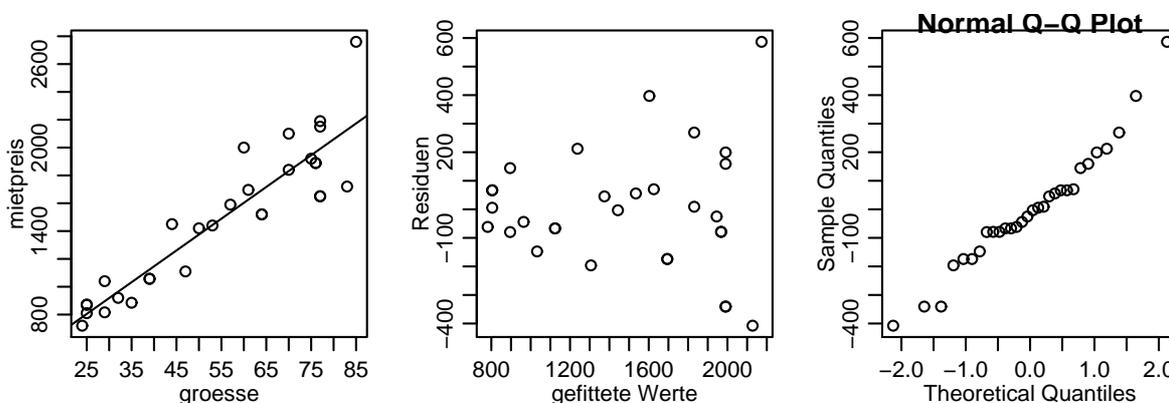
Coefficients:

	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



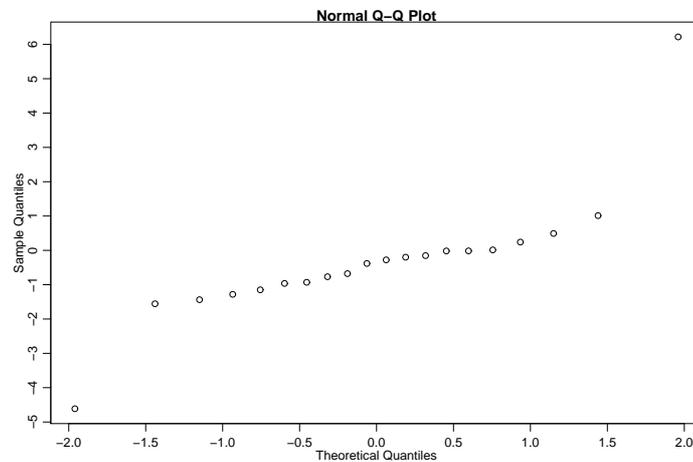
- Mit wievielen Beobachtungen wurde die Regression berechnet?
  - 22
  - 23
  - 28
  - 30
  - 29
- Wie gross ist der Achsenabschnitt (Intercept)  $\hat{\beta}_0$ ?
  - 234
  - 728
  - 0
  - 1485
- Wie gross ist die t-Teststatistik für den Test der Nullhypothese  $H_0 : \beta_1 = 0$  (die Alternativhypothese ist  $H_A : \beta_1 \neq 0$ )?
  - 22.81
  - 1.93
  - 11.819
  - 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  $\beta_1$ ?
  - $22.81 \pm 1.96 \cdot \frac{1.93}{\sqrt{28}}$
  - $22.81 \pm 2.048 \cdot \frac{1.93}{\sqrt{28}}$
  - $22.81 \pm 1.96 \cdot 1.93$
  - $22.81 \pm 2.048 \cdot 1.93$
- Fabio hätte gerne ein  $70m^2$  grosses Büro. 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?

- a) Nur die Annahme der konstanten Varianz der Fehler  $\epsilon_i$  scheint verletzt zu sein.
  - b) Nur die Normalitätsannahme der Fehler  $\epsilon_i$  scheint grob verletzt zu sein.
  - c) Die Modellannahmen über die Fehler scheinen plausibel.
  - 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 bleibt gleich
  - b) Sie wird kleiner
  - c) Sie wird grösser
  - d) Keine Aussage möglich

5. (10 Punkte)

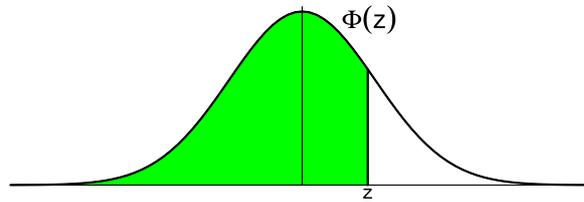
- 1) Es sei  $X \sim N(\mu, \sigma^2)$  mit  $\mu = 2$  und  $\sigma^2 = 9$ . Ferner sei  $Y \sim \text{Poisson}(\lambda)$  mit  $\lambda = 4$ . Nehmen Sie zudem an, dass  $X$  und  $Y$  unabhängig sind.  $Z$  sei gegeben durch  $Z = X + 2Y - 6$ . Wie gross ist  $\mathbf{E}[Z]$ ?
- a) 4                      b) 10                      c) 16  
d) 25                      e) 38                      f) Keine der Antworten trifft zu
- 2) Nehmen Sie  $Z$  aus Aufgabe 1). Wie gross ist die Varianz von  $Z$ ?
- a) 17                      b) 11                      c) 25  
d) 19                      e) 70                      f) Keine Aussage möglich
- 3) Zwei Statistiker A und B berechnen mittels zweier verschiedener Datensätze (unterschiedliche Anzahl Beobachtungen) das 95% Konfidenzintervall für den Erwartungswert einer normalverteilten Zufallsvariablen  $X$ . Beide verwenden die Annahme, dass  $X$  normalverteilt ist, und dass  $\text{Var}(X) = 25$  bekannt ist. Der Statistiker A erhält als Konfidenzintervall (9.36, 11.63), der Statistiker B erhält (9.02, 10.98). Welcher der beiden Statistiker benützt den Datensatz mit grösserer Anzahl Beobachtungen?
- a) Statistiker A  
b) Statistiker B  
c) Statistiker A und B benützen dieselbe Anzahl Beobachtungen  
d) Keine Aussage ist möglich
- 4) Welche der folgenden Verteilungen beschreibt die Anzahl des Eintretens seltener Ereignisse innerhalb eines bestimmten Zeitintervalls?
- a) Normalverteilung    b) Binomialverteilung  
c) Exponentialverteilung    d) Poisson-Verteilung
- 5) Welche der folgenden Aussagen verkleinert die Macht eines Tests?
- a) Vergrösserung des Fehlers 2. Art  
b) Vergrösserung des Signifikanzniveaus  
c) Vergrösserung der Anzahl Beobachtungen
- 6) Es sei  $X \sim \text{Exp}(1/500)$ . Was ist die zugehörige kumulative Verteilungsfunktion  $F(x)$ ?
- a)
- $$F(x) = \begin{cases} 0 & x < 0 \\ \exp(\frac{-x}{500}) & x \geq 0 \end{cases}$$
- b)
- $$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \exp(\frac{-x}{500}) & x \geq 0 \end{cases}$$
- c)
- $$F(x) = \begin{cases} 0 & x < 0 \\ 500 \cdot \exp(\frac{-x}{500}) & x \geq 0 \end{cases}$$
- d)
- $$F(x) = \begin{cases} 0 & x < 0 \\ 1/500 \cdot \exp(-x \cdot 500) & x \geq 0 \end{cases}$$
- 7) Es sei  $X \sim \text{Exp}(1/500)$ . Was ist der Median von  $X$ ?
- a) 500                      b)  $500^2$                       c)  $1/500$   
d)  $500 \cdot \ln(2)$                       e)  $\ln(2)/500$                       f) Aussagen a) bis e) sind falsch

- 8) Es sei  $X \sim N(\mu, \sigma^2)$  mit  $\mu = 1$  und  $\sigma^2 = 2$ . Was ist  $\mathbf{E}[X^2]$ ?
- $(\mathbf{E}[X])^2$
  - $(\text{Var}(X))^2$
  - $\sqrt{\text{Var}(X)}$
  - $(\mathbf{E}[X])^2 + \text{Var}(X)$
- 9) Es sei  $Y = \sum_{i=1}^{100} X_i$ , wobei die  $X_i, i = 1, \dots, 100$ , unabhängig und identisch Bernoulli(0.5)-verteilt sind. Was ist die exakte Verteilung von  $Y$ ?
- Bernoulli(0.5)
  - Bin(100, 0.5)
  - $\mathcal{N}(\mu = 50, \sigma^2 = 25)$
  - Pois(50)
- 10) Von einer Zufallsvariablen  $X$  mit Erwartungswert  $\mu$  wurden Daten gesammelt. Ein Normal-Plot der Daten ist unten aufgezeichnet. Wir möchten die folgende Hypothese auf dem 5% Niveau testen:  $H_0 : \mu = 0, H_A : \mu \neq 0$ .



- Der P-Wert des t-Tests mit diesen Daten ist 0.4614. Mit dem Wilcoxon-Test erhält man 0.0400 als P-Wert. Welche der folgenden Aussagen ist zutreffend?
- Der Wilcoxon-Test ist vorzuziehen und die Nullhypothese wird verworfen.
  - Der Wilcoxon-Test ist vorzuziehen und die Nullhypothese wird nicht verworfen.
  - Der t-Test ist vorzuziehen und die Nullhypothese wird verworfen.
  - Der t-Test ist vorzuziehen und die Nullhypothese wird nicht verworfen.

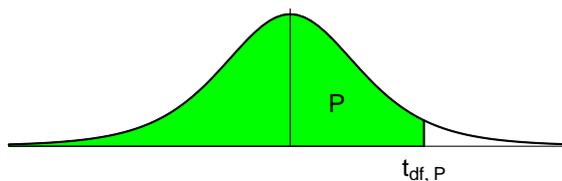
Tabelle der Kumulativen Normalverteilung  $\Phi(z) = P[Z \leq z]$ ,  $Z \sim \mathcal{N}(0,1)$



Bsp.:  $P[Z \leq 1.96] = 0.975$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

## 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