

## 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!
- Die nötigen Tabellen befinden sich nach den Aufgaben auf den hintersten Seiten dieser Prüfung.
- Bei den Aufgaben 1, 2 und 3 muss der Lösungsweg ersichtlich sein, sonst gibt es keine Punkte.
- Für die Antworten zu den Aufgaben 1, 2 und 3 sind die zugehörigen Antwortblätter zu benutzen. Bitte schreiben Sie alle Antworten, welche Sie abgeben möchten, nur auf die vorbereiteten Antwortblätter. Antworten auf eigens mitgebrachten Blättern werden nicht korrigiert. Es sind zwei Reserve-Antwortblätter vorhanden.
- Aufgaben 4 und 5 sind Multiple-Choice-Aufgaben. Es ist jeweils genau eine Antwort korrekt. Eine korrekte Antwort gibt 1 **Pluspunkt**; eine falsche Antwort (oder mehrere angekreuzte Antworten)  $\frac{1}{2}$  **Minuspunkt**. Minimal erhält man für eine ganze Multiple-Choice-Aufgabe 0 Punkte. Tragen Sie die korrekten Antworten der Multiple-Choice-Aufgaben mit Kreuzchen in das zugehörige Antwortblatt ein.
- Rechnen Sie immer mit den exakten Zwischenresultaten weiter. Runden sie alle auf dem Papier angegebenen Endergebnisse sinnvoll.

**Viel Erfolg!**

### 1. (8 Punkte)

Homer S. kauft seine Lebensmittel entweder im Kwik-E-Mart oder im Wull-Mart. Bis jetzt ist er davon ausgegangen, dass beide Läden im Mittel gleich teuer sind. Er möchte nun überprüfen, ob dies wirklich so ist. Deshalb geht er an einem Tag in beide Geschäfte, und kauft von seinen Lieblingsprodukten je eines. Die Produkte sind identisch in beiden Geschäften. Er stellt folgende Preise (in USD) fest:

Produkt	Donuts	Burgers	Hot-Dogs	Speck	Duff Bier	Pizza
Kwik-E-Mart	4	6	3	1	3	3
Wull-Mart	5	5	4	2	5	2

- a) Handelt es sich hier um einen gepaarten oder einen ungepaarten Test? Begründen Sie.
- b) Welche Annahme über die Verteilung der Differenzen der Preise macht man typischerweise, um einen  $t$ -Test durchzuführen? Wie kann man diese Annahme überprüfen?
- c) Geben Sie die Null- und die Alternativhypothese an.
- d) Wir betrachten die Differenzen (Preis Kwik-E-Mart)  $-$  (Preis Wull-Mart). Schätzen Sie die Varianz  $\sigma^2$  dieser Differenzen.
- e) Führen Sie einen  $t$ -Test auf dem 10%-Niveau durch. Was ist der Testentscheid?
- f) Geben Sie ein zweiseitiges 90%-Vertrauensintervall für den Mittelwert der Differenzen  $\mu$  an. Geben Sie die Antwort auf drei Nachkommastellen genau an.

## 2. (8 Punkte)

Eine der gefährlichsten Komplikationen nach einer Bluttransfusion ist die sogenannte *transfusionsassoziierte akute Lungeninsuffizienz* (“TRALI”), die sich durch akute Atemnot einige Stunden nach einer Bluttransfusion bemerkbar macht. TRALI tritt im Mittel nach einer von 55'000 Bluttransfusionen auf.

- a) In einem grösseren Spital werden in einem Jahr 30'000 Bluttransfusionen durchgeführt. Welcher (exakten) Verteilung entspricht die Anzahl der Patienten, die in diesem Jahr nach einer Bluttransfusion an TRALI erkranken? (Verteilungsfamilie und Parameter angeben!)
- b) Mit welcher Verteilung kann man die Zahl der Patienten, die an TRALI erkranken, gut *approximativ* beschreiben? Begründen Sie die Wahl Ihrer Verteilung in wenigen Sätzen, und berechnen Sie die dazugehörigen Parameter.
- c) Wie gross ist die Wahrscheinlichkeit, dass in dem Spital in einem Jahr 2 oder mehr Fälle von TRALI auftreten? Berechnen Sie die Wahrscheinlichkeit einmal mit der exakten Verteilung aus Aufgabe a) und einmal mit der genäherten Verteilung aus Aufgabe b) und geben Sie die Resultate jeweils auf 4 Stellen genau an.
- d) Sie möchten nun die Wahrscheinlichkeit, nach einer Bluttransfusion an TRALI zu erkranken, selber schätzen. Dazu machen Sie eine Umfrage bei 6 grösseren Spitälern ( $i = 1, \dots, 6$ ) und fragen jeweils nach der Anzahl Bluttransfusionen ( $n_i$ ) sowie den Anzahl Fällen von TRALI ( $k_i$ ) im letzten Jahr. Sie erhalten folgende Angaben:

$i$	$n_i$	$k_i$
1	34'855	2
2	25'177	0
3	29'816	0
4	31'164	2
5	30'021	3
6	27'469	0

Schätzen Sie allein auf Grund dieser Angaben die Wahrscheinlichkeit, nach einer Bluttransfusion an TRALI zu erkranken.

### 3. (10 Punkte)

Marco ist ein Spargelliebhaber. Deshalb konsumiert er während der Spargelsaison jeden Mittag einen Spargelsalat in der Kantine, der jeweils 10 Spargelstücke enthält. Diese werden in der Küche durchs Schneiden ganzer Spargeln in 6 Stücke hergestellt.

- a) Marco mag besonders die Spargelspitzen, deshalb zählt er jeweils deren Anzahl (Variable  $X$ ) auf seinem Teller. Wie ist  $X$  verteilt unter der Annahme, dass bei der Herstellung des Salats zufällig 10 Stücke aus einem grossen, gut durchmischten Topf genommen werden? (Verteilungsfamilie sowie Parameter angeben!)
- b) Wie gross ist die Wahrscheinlichkeit  $p = P(X \geq 3)$ , mindestens 3 Spargelspitzen auf seinem Teller zu finden? (Numerisches Resultat auf 4 Stellen.)
- c) Marco freut sich vom Beginn der Spargelsaison an auf den ersten Tag, an dem er einen Teller mit mindestens 3 Spargelspitzen erwischt; er zählt daher die Tage  $T$  bis zu diesem Ereignis. Als Beispiel:  $T = 5$  bedeutet, dass er die ersten 4 Tage der Spargelsaison jeweils 2 oder weniger Spitzen auf dem Teller hat und am 5. Tag endlich 3 oder mehr.

Berechnen Sie die Verteilung von  $T$ , d.h. die Wahrscheinlichkeiten  $P(T = k)$  für  $k = 1, 2, 3, \dots$  (als Funktion von bzw. Formel in  $k$  und  $p$  aus Aufgabe b); Sie brauchen keine konkreten Zahlenwerte für  $k$  oder  $p$  einzusetzen).

- d) Wie lange muss Marco im Durchschnitt auf den ersten Tag mit mindestens 3 Spitzen im Teller warten? Berechnen Sie den Erwartungswert von  $T$ . Sie können dafür die folgende Formel verwenden: für eine reelle Zahl  $q$  mit  $0 < q < 1$  gilt

$$\sum_{k=1}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}.$$

- e) Marco verdächtigt das Kantinenpersonal, beim Zubereiten der Salate gezielt Spargelspitzen herauszusuchen und zu naschen. Deshalb zählt er die Zahl der Spitzen, die er über die ganze Spargelsaison verspeist; er nennt diese Zahl  $Y$ . Während einer Spargelsaison isst er 48 Spargelsalate, verspeist also insgesamt  $48 \cdot 10 = 480$  Spargelstücke. Wie ist  $Y$  *approximativ* verteilt unter der Annahme, dass das Personal *keine* Spitzen nascht? (Verteilungsfamilie und Parameter angeben.)
- f) Während einer Saison zählt Marco  $Y = 72$  Spargelspitzen. Stimmt seine Vermutung, dass das Personal Spitzen nascht? Führen Sie mit der Verteilung aus e) einen Test auf dem 5%-Niveau durch. Geben Sie explizit die Null- und Alternativhypothese an; muss ein- oder zweiseitig getestet werden?

#### 4. (8 Punkte)

Es liegen uns Daten über die Entwicklung der Geburtenrate von 20 lateinamerikanischen Ländern vor. Die Daten beinhalten die prozentuale Abnahme der Geburtenrate ( $y$ ) zwischen den Jahren 1965 und 1975 in den Ländern und einen Index ( $x$ ) der den sozialen Hintergrund der Länder beschreibt. Folgendes lineare Modell wurde an diese Daten angepasst:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, 20.$$

Der Regressionsoutput sieht wie folgt aus:

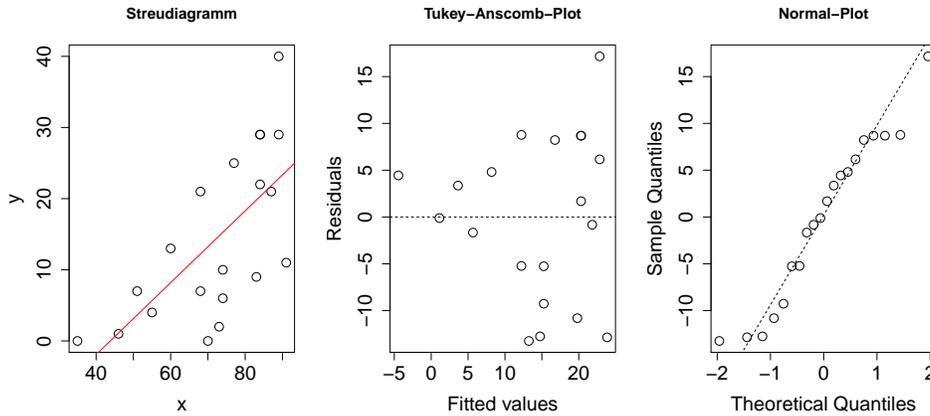
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	???	9.6416	-2.295	0.03398
x	0.5052	0.1308	???	???

Residual standard error: 8.973 on ??? degrees of freedom  
Multiple R-squared: 0.4532, Adjusted R-squared: 0.4228  
F-statistic: 14.92 on 1 and 18 DF, p-value: 0.001141

- 1) Wie gross ist der Intercept  $\hat{\beta}_0$ ?
  - a) 9.6416
  - b) 30.863
  - c) -3.863
  - d) 24.679
  - e) -23.245
  - f) -22.125
- 2) Wieviele Freiheitsgrade hat der "Residual standard error"?
  - a) 20
  - b) 19
  - c) 16
  - d) 17
  - e) 18
  - f) 21
- 3) Wie gross ist der  $t$ -Wert der Schätzung von  $\beta_1$ ?
  - a) 1.863
  - b) 0.52
  - c) 3.863
  - d) 2.458
  - e) 10.253
  - f) -3.121
- 4) Welches der folgenden Intervalle ist ein exaktes zweiseitiges 95% Vertrauensintervall für  $\beta_1$ ?
  - a)  $0.5052 \pm 2.101 \cdot \frac{0.1308}{\sqrt{18}}$
  - b)  $0.5052 \pm 1.960 \cdot \frac{0.1308}{\sqrt{18}}$
  - c)  $0.5052 \pm 2.101 \cdot 0.1308$
  - d)  $0.5052 \pm 1.960 \cdot 0.1308$
- 5) Wird die Nullhypothese  $H_0 : \beta_1 = 0$  auf dem 1%-Niveau verworfen?
  - a) Ja
  - b) Nein
  - c) Keine Aussage möglich
- 6) Wie gross ist die Schätzung von  $\sigma^2$ ?
  - a) 9.23
  - b) 50.34
  - c) 80.51
  - d) 8.973
  - e) 40.86
  - f) Keine Aussage möglich

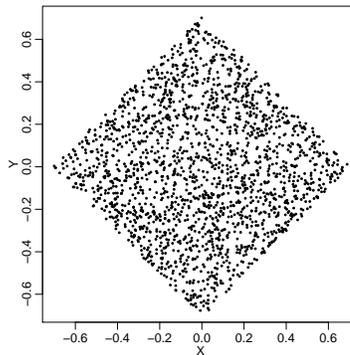
7) Betrachten Sie die Plots. Welche der folgenden Aussagen ist zutreffend?



- Die Modellannahmen über die Fehler sind erfüllt.
  - Die Modellannahmen über die Fehler scheinen plausibel. Es gibt aber Ausreisser.
  - Die Annahme der konstanten Varianz scheint grob verletzt zu sein, aber die Normalitätsannahme der Fehler ist plausibel.
  - Es gibt systematische Fehler im Modell.
- 8) Was passiert mit der Regression, wenn an der Stelle (40, 30) noch eine Beobachtung hinzugefügt wird?
- $\hat{\beta}_1$  wird kleiner,  $\hat{\sigma}$  wird kleiner
  - $\hat{\beta}_1$  wird kleiner,  $\hat{\sigma}$  wird grösser
  - $\hat{\beta}_1$  wird grösser,  $\hat{\sigma}$  wird kleiner
  - $\hat{\beta}_1$  wird grösser,  $\hat{\sigma}$  wird grösser

5. (8 Punkte)

- 1)  $Z$  sei eine Zufallsvariable mit Verteilung  $Z \sim \mathcal{N}(\mu, \sigma^2)$  mit  $\mu = 2$  und  $\sigma^2 = 4$ . Was ist  $P(Z \geq 5)$ ?
- a)  $\Phi(5)$       b)  $\Phi(-\frac{3}{2})$       c)  $\Phi(2)$   
d)  $\Phi(\frac{3}{2})$       e)  $1 - \Phi(\frac{1}{2})$       f)  $1 - \Phi(5)$
- 2) Das zu einem  $t$ -Test gehörende *einseitige* 95%-Vertrauensintervall für  $\mu$  sei  $[2.45, \infty)$ ; das *zweiseitige* 95%-Vertrauensintervall des  $t$ -Tests auf denselben Daten sei  $[a, b]$ . Dann gilt für die Grenzen  $a$  und  $b$ :
- a)  $a, b < 2.45$       b)  $a < 2.45, b = 2.45$   
c)  $a < 2.45, b > 2.45$       d)  $a = 2.45, b > 2.45$   
e)  $a, b > 2.45$       f) keine Aussage möglich
- 3)  $X$  und  $Y$  seien zwei unabhängige Zufallsvariablen mit  $\mathcal{E}(X) = 1$ ,  $\text{Var}(X) = 1$ ,  $\mathcal{E}(Y) = 3$  und  $\text{Var}(Y) = 2$ . Was ist die Varianz von  $Z = X + 2Y$ ?
- a) 3      b) 5      c) 7  
d) 9      e) 11      f) keine Aussage möglich
- 4) Das folgende Streudiagramm zeigt eine grosse Zahl von Realisierungen von zwei Zufallsvariablen  $X$  und  $Y$ :



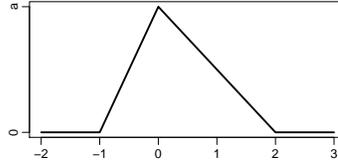
Welche der folgenden Aussagen ist zutreffend?

- a)  $X$  und  $Y$  sind unkorreliert.  
b)  $X$  und  $Y$  haben Korrelation 1.  
c)  $X$  und  $Y$  haben Korrelation im Bereich  $[0.1, 0.7]$ .  
d) Es ist keine Aussage über die Korrelation von  $X$  und  $Y$  möglich.

Claudia spielt mit einem präparierten Würfel, der mit Wahrscheinlichkeit  $\frac{2}{5}$  eine 6 zeigt und jeweils mit Wahrscheinlichkeit  $\frac{3}{25}$  eine der Zahlen 1 bis 5.

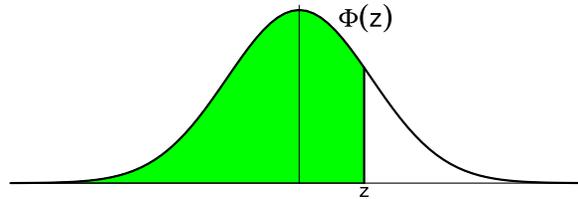
- 5) Wie gross ist die Wahrscheinlichkeit, dass der Würfel in 4 Würfeln genau 2 mal eine 6 zeigt (gerundet auf 4 Nachkommastellen)?
- a) 0.0138      b) 0.0576      c) 0.3456  
d) 0.4250      e) 0.9600      f) Antworten a) bis e) sind falsch
- 6) Wie gross ist die Wahrscheinlichkeit, dass der Würfel in drei Würfeln einmal eine 1, einmal eine 2, und einmal eine der Zahlen  $\{3, 4, 5, 6\}$  zeigt (gerundet auf 4 Nachkommastellen)?
- a) 0.0109      b) 0.0172      c) 0.0328  
d) 0.0657      e) 0.1250      f) Antworten a) bis e) sind falsch

Die folgende Abbildung zeigt die Wahrscheinlichkeitsdichte  $f_X$  einer Zufallsvariablen  $X$ :



- 7) Welchen Wert hat das Maximum  $a$  der Dichte  $f_X$ ?
- a)  $\frac{1}{2}$                       b)  $\frac{4}{5}$                       c) 1  
d)  $\frac{3}{2}$                       e) 4                          f) keinen der genannten Werte
- 8) Ist der Median  $\mu$  der Zufallsvariablen  $X$  grösser oder kleiner als ihr Erwartungswert  $\mathcal{E}(X)$ ?
- a)  $\mu < \mathcal{E}(X)$   
b)  $\mu = \mathcal{E}(X)$   
c)  $\mu > \mathcal{E}(X)$   
d) das hängt von der Anzahl Beobachtungen von  $X$  ab

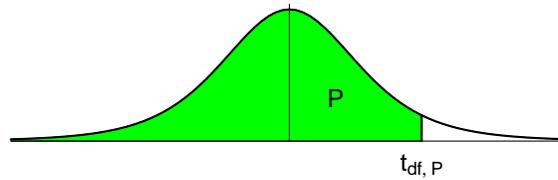
Tabelle der Kumulativen Normalverteilung  $\Phi(z) = P[Z \leq z]$ ,  $Z \sim \mathcal{N}(0, 1)$



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
90	0.254	0.526	0.846	1.291	1.662	1.987	2.368	2.632
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