

## Schriftliche Prüfung (2 Stunden)

### Bemerkungen:

- Alle schriftlichen Hilfsmittel und ein Taschenrechner sind erlaubt.
- Mobiltelefone 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.
- Der Lösungsweg muss immer ersichtlich sein.
- Bei den Multiple-Choice-Aufgaben ist jeweils genau eine Antwort korrekt. Eine korrekte Antwort gibt 1 **Pluspunkt** und eine falsche Antwort  $\frac{1}{2}$  **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.
- Die nötigen Tabellen befinden sich auf der hintersten Seiten dieser Prüfung.

**Viel Erfolg!**

1. (9 Punkte) Lukas und Markus haben bisher immer "Feinste Mini-Brezeln 100g" des Herstellers Gammelbrot und Söhne zum Znüni gegessen. Vom ständigen Hungerklagen von Markus genervt schlägt Lukas nun vor auf "Rustikale Brez'n Bayrischer Art 100g" des Herstellers Wolpertinger Backwaren zu wechseln. Lukas vermutet, dass - obwohl beide Hersteller das selbe Gewicht angeben - die Brezeln der Firma Gammelbrot und Söhne leichter als das Konkurrenzprodukt sind. Um dies zu untersuchen bringt Markus seine Küchenwaage mit und wiegt einige Brezeln beider Marken. Die Ergebnisse (alles in g) sind:

Brezeln von Gammelbrot	97	97	97	98	93
Brezeln von Wolpertinger	100	101	99	99	98

Nun wollen die beiden testen ob die Vermutung von Lukas richtig ist. Dazu nehmen sie an, dass die Gewichts-Differenzen normalverteilt mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$  sind (mit Gewichts-differenz ist hier und im Folgenden gemeint: Gewicht einer Brezel von Gammelbrot - Gewicht einer Brezel von Wolpertinger).

*Anmerkung: Die Aufgabe war unklar gestellt. Folgender Zusatz hat gefehlt: "Es werden von jedem Brezelhersteller 5 Brezeln unterschiedlicher Sorten gekauft (mit Mohn, Sonnenblumenkernen, Kürbiskernen, Lauge, und gemischt). In der Tabelle entspricht jede Spalte einer Sorte."*

- Handelt es sich um einen gepaarten oder einen ungepaarten Test?
- Geben Sie die Null- und die Alternativhypothese an.
- Geben Sie eine Schätzung für die Varianz  $\sigma^2$  der Differenz an.
- Führen Sie den geeigneten t-Test durch: Bestimmen Sie den Wert der Teststatistik  $T$ , den Verwerfungsbereich für  $T$  und den Testentscheid. (Wenn sie obige Aufgabe nicht lösen konnten, benutzen Sie im Folgenden als Ersatzwert  $\sigma^2 = 2.5$ .)
- Bestimmen Sie ein einseitiges 95%-Vertrauensintervall für  $\mu$ .
- Angenommen, der Wert  $\sigma^2$  wäre nicht aus den Daten geschätzt, sondern bekannt: Wie lautet dann das einseitige 95%-Vertrauensintervall? Geben Sie eine kurze Erklärung für den Unterschied zu e)!

2. (9 Punkte) Das Pharmaunternehmen Life Co. hat ein neues Medikament zur Bekämpfung von ADHS entwickelt. Um die Wirksamkeit festzustellen wurde das Medikament mit  $n = 10$  Patienten getestet. Die derzeitige Standardmethode zeigt bei 30% der behandelten Patienten eine Wirkung.
- a) Angenommen das neue Medikament ist genauso wirksam wie die Standardmethode, wie gross ist die Wahrscheinlichkeit, dass die Behandlung bei genau 2 Patienten eine Wirkung zeigt? Wie gross ist die Wahrscheinlichkeit, dass sie bei höchstens 2 Patienten eine Wirkung zeigt?
  - b) Die Behandlung mit dem neuen Medikament war bei 4 Patienten erfolgreich. Führen Sie einen einseitigen Hypothesentest durch um festzustellen ob das neue Medikament wirksamer ist als die Standardmethode (bei einem Signifikanzniveau von 5%). Geben Sie explizit alle Schritte an.
  - c) Wie ist die *Macht* eines Hypothesentests definiert? Geben Sie die Macht an für den Test  $H_0 : \pi = 0.3$  vs.  $H_A : \pi = 0.6$  ( $\pi$  ist die Wirksamkeit).
  - d) An einem zweiten Test nehmen 200 Patienten teil. Diesmal zeigt das Medikament bei 80 Patienten eine Wirkung. Führen Sie erneut einen einseitigen Hypothesentest durch und berechnen Sie anschliessend das (einseitige!) 95%-Intervall für  $\pi$ .

3. (7 Punkte) Ein Mass für die technische Entwicklung (bzw. Rückständigkeit) eines Landes ist der Anteil der arbeitenden Bevölkerung im Agrarsektor. Es soll nun der Einfluss auf das Pro-Kopf-Einkommen untersucht werden. Dafür liegen Daten aus 20 Ländern aus dem Jahr 1960 vor. Folgendes Modell wurde angepasst:

$$\text{pcinc}_i = \beta_0 + \beta_1 \cdot \text{agr}_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2),$$

wobei  $\text{pcinc}_i$  das durchschnittliche Pro-Kopf-Einkommen (in USD) und  $\text{agr}_i$  der Anteil der arbeitenden Bevölkerung im Agrarsektor (in % · 100) sind. D.h. wenn z.B. 30% der arbeitenden Bevölkerung im Land  $i$  im Agrarsektor tätig sind, dann ist  $\text{agr}_i = 30$ . Der (unvollständige) Regressionsoutput sieht wie folgt aus:

Residuals:

Min	1Q	Median	3Q	Max
-351.60	-211.36	-3.75	105.93	590.10

Coefficients:

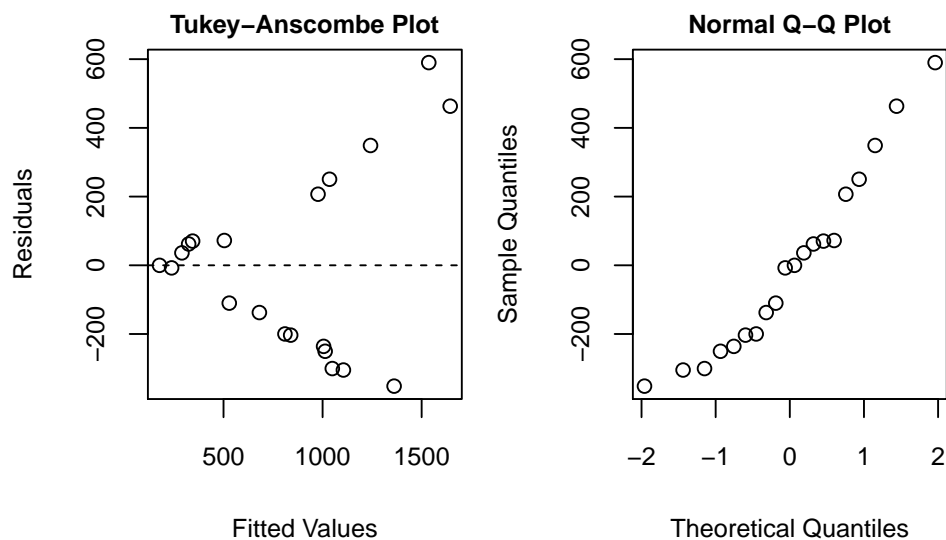
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1317.905	???	12.406	2.95e-10
agr	-18.858	???	-5.554	2.84e-05

---

Residual standard error: 273.1 on ? degrees of freedom  
 Multiple R-squared: 0.6315, Adjusted R-squared: 0.611  
 F-statistic: 30.85 on 1 and ? DF, p-value: 2.845e-05

- 1) Was ist  $\hat{\beta}_1$ ?
  - a) 1317.905
  - b) 12.406
  - c) -18.858
  - d) -5.554
- 2) Was ist der Standardfehler von  $\hat{\beta}_0$ ?
  - a) 0.009
  - b) 0.295
  - c) 3.395
  - d) 106.230
- 3) Mit wievielen Freiheitsgraden wurde der "residual standard error" berechnet?
  - a) 18
  - b) 22
  - c) 2
  - d) 20
- 4) Berechnen Sie das 99%-Konfidenzintervall für  $\beta_1$  (ohne Normalapproximation).
  - a) [-22.25, -15.46]
  - b) [-25.99, -11.73]
  - c) [-27.52, -10.19]

- d)  $[-28.63, -9.09]$
- 5) Kann die Nullhypothese  $H_0 : \beta_0 = 0$  auf dem 5% Signifikanzniveau verworfen werden?
- Ja.
  - Nein.
  - Keine Angabe möglich.
- 6) In der Schweiz betrug 1960 das durchschnittliche Pro-Kopf-Einkommen 1361 USD, der Agrarsektor hatte einen Anteil von 11%. Wie hoch ist das Residuum für diesen Datenpunkt in unserem Modell?
- 45.17
  - 250.53
  - 1110.47
  - 1315.83
- 7) Betrachten Sie die nachfolgenden Plots. Welche der folgenden Aussagen trifft zu?
- Alle Modellannahmen sind erfüllt.
  - Die Fehlervarianz scheint nicht konstant zu sein, aber die Normalverteilungsannahme ist plausibel.
  - Die Fehlervarianz scheint konstant zu sein, aber die Normalverteilungsannahme scheint nicht zuzutreffen.
  - Sowohl konstante Fehlervarianz als auch Normalverteilungsannahme treffen nicht zu.



#### 4. (8 Punkte)

- 1) Was ist der Median und das arithmetische Mittel (Mean) der folgenden Zahlen: 2, 4, 5, 6, 7, 10, 270?
  - a) Median = 136, Mean = 43.4
  - b) Median = 6.5, Mean = 35.7
  - c) Median = 6, Mean = 43.4
  - d) Median = 43.4, Mean = 6
  - e) Median = 35.7, Mean = 6.5
  - f) Median = 43.4, Mean = 136
- 2) Was ist die Varianz von  $Z = 3X - Y$ , wenn  $X \sim \mathcal{N}(2, 2)$ ,  $Y \sim \mathcal{N}(-2, 2)$  und  $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$ ?
  - a)  $\text{var}(Z) = 4$
  - b)  $\text{var}(Z) = 16$
  - c)  $\text{var}(Z) = 20$
  - d)  $\text{var}(Z) = 22$
  - e) Kann ohne weitere Informationen nicht berechnet werden.
- 3) Angenommen, ein zweiseitiges 95%-Vertrauensintervall für den wahren Erwartungswert  $\mu$  einer Population ist  $[490, 497]$ . Nehmen wir weiter an, wir machen nun mit den selben Daten, mit denen das Vertrauensintervall berechnet wurde, einen zweiseitigen Ein-Stichproben t-Test mit der Nullhypothese  $H_0 : \mu = 500$  und der Alternative  $H_A : \mu \neq 500$  auf dem 5% Signifikanzniveau. Wird die Nullhypothese verworfen?
  - a) Ja.
  - b) Nein.
  - c) Kann man ohne weitere Informationen nicht beantworten.
- 4) An einer Losbude werden  $n = 120$  Lose gezogen. Die geschätzte Gewinnwahrscheinlichkeit beträgt 0.35. Das 95%-Vertrauensintervall ist  $[0.26, 0.44]$ . Angenommen wir wollen die Unsicherheit des Vertrauensintervalls von  $\pm 0.09$  auf  $\pm 0.03$  reduzieren. Circa wieviele Lose sind dafür insgesamt notwendig?
  - a)  $n = 40$
  - b)  $n = 120$
  - c)  $n = 360$
  - d)  $n = 1080$
  - e)  $n = 3240$
- 5) Der p-Wert eines einseitigen Binomialtests mit  $H_0 : \pi = 0.5$  und  $H_A : \pi > 0.5$  ist 0.07. Welche der Regeln für das Verwerfen der Nullhypothese ist korrekt?
  - a) Auf dem 10%-Niveau verwerfen und auf dem 5%-Niveau nicht verwerfen.
  - b) Auf dem 10%-Niveau verwerfen und auf dem 5%-Niveau verwerfen.
  - c) Auf dem 10%-Niveau nicht verwerfen und auf dem 5%-Niveau nicht verwerfen.
  - d) Auf dem 10%-Niveau nicht verwerfen und auf dem 5%-Niveau verwerfen.

- 6) In einem Callcenter ist die Zeit zwischen zwei eingehenden Anrufen Exponentialverteilt. Welche Verteilung modelliert exakt die Anzahl der Anrufe innerhalb eines Jahres?
- a) Normal-Verteilung
  - b) Poisson-Verteilung
  - c) Exponential-Verteilung
  - d) Binomial-Verteilung
  - e) Uniforme Verteilung
- 7) Welche der folgenden Aussagen ist korrekt?
- a)  $\text{Corr}(X, Y) = 0 \Leftrightarrow X, Y$  sind unabhängig.
  - b)  $\text{Corr}(X, Y) = 0 \Rightarrow X, Y$  sind unabhängig.
  - c)  $X, Y$  sind unabhängig  $\Rightarrow \text{Corr}(X, Y) = 0$ .
  - d) Keine.
- 8) Eine Person arbeitet 256 Tage im Jahr. Am  $i$ -ten Tag macht sie unabhängig von den anderen Tagen  $X_i$  Überstunden. Die Verteilung der Überstunden ist unbekannt, allerdings ist sie für jeden Tag gleich und es gilt  $\mathbf{E}[X_i] = 0.5$  und  $\text{Var}(X_i) = 1$ . Wie wahrscheinlich ist es (approximativ), dass die Person am Ende vom Jahr mehr als 160 Überstunden gemacht hat?
- a) 97.7%
  - b) 84.1%
  - c) 69.2%
  - d) 30.8%
  - e) 15.9%
  - f) 2.3%

5. (7 Punkte) Durch die hohe Nachfrage nach Rohstoffen könnte bald ein profitabler Abbau großer Zinnvorkommen in Ostdeutschland möglich werden. Die Schweizer Firma Schmelzli untersuchte deshalb potentielle Lagerstätten in Dresden, Gera und Chemnitz und führte an jedem dieser drei Standorte jeweils eine Schallmessung  $X \in \{0, 1\}$  durch. Es ist aus langjähriger Erfahrung bekannt, dass bei einer Messung  $X = 1$  die Wahrscheinlichkeit für Zinnvorkommen bei 60% und bei einer Messung  $X = 0$  die Wahrscheinlichkeit für Zinnvorkommen bei 40% liegt.

An allen drei potentiellen Lagerstätten wurde  $X = 0$  gemessen.

- a) Wie groß ist die Wahrscheinlichkeit, dass in allen drei potentiellen Lagerstätten Zinn vorhanden ist?
- b) Wie gross ist die Wahrscheinlichkeit, dass in mindestens einer potentiellen Lagerstätte Zinn vorhanden ist?

Eine weitere Messung in Görlitz hat zum Ergebnis  $X = 1$ . Die Besitzer der Firma Schmelzli, Ottfried Zübel und Jonathan Ackert, sind nun uneins über das weitere Vorgehen. Während Zübel ohne die Hilfe einer anderen Firma mit den Bohrungen in Görlitz beginnen will, plädiert Ackert für gemeinsame Bohrungen in Görlitz mit der deutschen Firma Borifex. Die Verluste für Schmelzli (in Mio. Dollar) beider Strategien bei tatsächlichem Zinnvorkommen und ohne tatsächlichem Zinnvorkommen sind in folgender Tabelle notiert:

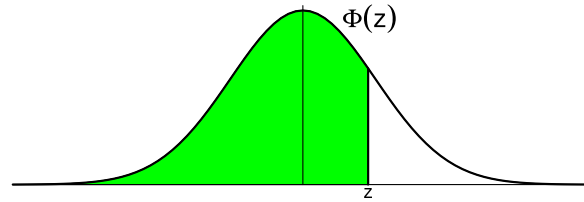
	Alleiniges Bohren	Gemeinsames Bohren mit Borifex
Zinnvorkommen	0	4
kein Zinnvorkommen	12	6

Hier ist so normiert worden, dass im optimalen Fall (alleiniges Bohren bei Zinnvorkommen) der Wert 0 steht.

- c) Berechnen Sie die Erwartungswerte der Verluste beider Strategien.
- d) Wie groß ist die Varianz der jeweiligen Strategien? Welche Strategie würde eine auf wenig Risiko bedachte Firma wählen?



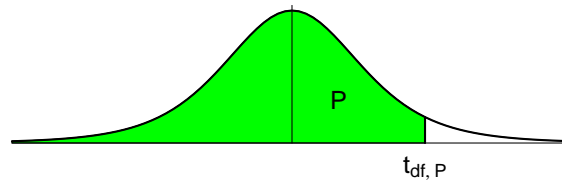
Tabelle der Kumulativen Normalverteilung  $\Phi(z) = P[Z \leq z]$ ,  $Z \sim \mathcal{N}(0, 1)$



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