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 **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 zugehörige Antwortblatt ein.
- Die nötigen Tabellen befinden sich auf der hintersten Seiten dieser Prüfung.

Viel Erfolg!

1. **(9 Punkte)** F. Lauer möchte das Gerücht überprüfen, dass Blumen schneller wachsen, wenn man mit ihnen redet. Daher kauft sie acht identische Blumenzwiebeln, schickt jeweils zwei davon zu jedem ihrer vier Kinder und bittet sie, die beiden Blumenzwiebeln genau gleich zu behandeln. Mit dem einzigen Unterschied, dass sie nur mit der einen Blume reden sollen. Nach sechs Wochen erkundigt sie sich, wie hoch die Blumen gewachsen sind und erhält folgende Antworten (in cm):

Kind	1	2	3	4
Blume (beredet)	30.3	32.2	29.9	30.1
Blume (nicht beredet)	30.1	31.9	29.9	30.0

Nun möchte sie mit Hilfe dieser Daten herausfinden, ob an dem Blumengerede etwas dran ist und bittet Sie, einen geeigneten statistischen Test durchzuführen. Nehmen Sie an, dass die Differenzen Höhe Blume (beredet) minus Höhe Blume (nicht beredet) normalverteilt mit Erwartungswert μ und Varianz σ^2 sind.

a) Es handelt sich um einen gepaarten Test. Warum?

b) Geben Sie die Null- und die Alternativhypothese an und begründen Sie kurz Ihre Wahl.

c) Geben Sie eine Schätzung $\hat{\sigma}^2$ für die Varianz σ^2 der Differenz an (mit Lösungsweg).

d) Führen Sie den geeigneten t-Test zum Niveau 0.05 durch: Bestimmen Sie den Wert der Teststatistik T und dessen Verteilung unter der Nullhypothese, den Verwerfungsbereich für T und den Testentscheid. (Wenn Sie obige Aufgabe c) nicht lösen konnten, benutzen Sie im Folgenden $\hat{\sigma}^2 = \frac{5}{300}$.)

e) Bestimmen Sie ein einseitiges 95%-Vertrauensintervall für μ.

f) Sie sind der statistische Berater von F. Lauer und haben die Berechnungen durchgeführt. Was sagen Sie ihr nun?

2. (8 Punkte) Die Basler Werbefirma Luegane will die Wirksamkeit von Plakatwerbung untersuchen. Dazu befragt die Firma 10 Passanten kurz nachdem diese an einem Plakat vorbei gegangen sind. 7 Passanten geben an, das Plakat nicht betrachtet zu haben, 3 geben an, es betrachtet zu haben. Sei nun $\pi \in [0, 1]$ die Wahrscheinlichkeit, dass ein Passant am Plakat vorbei geht, dieses aber nicht betrachtet. Sei weiterhin $X \in \{0, 1, \ldots, 10\}$ die Anzahl der Passanten die das Plakat nicht betrachten. Die Passanten werden im Folgenden als unabhängig betrachtet.

Wir nehmen nun erst einmal an, dass $\pi = 0.6$.

a) Wie ist X verteilt?

b) Wie gross ist die Wahrscheinlichkeit, dass alle Passanten das Plakat betrachtet haben?

c) Wie gross ist die Wahrscheinlichkeit, dass sowohl der erste als auch der fünfte Passant das Plakat nicht betrachtet haben?

In der Werbebranche wird seit Jahren angenommen, dass die Wahrscheinlichkeit, dass ein Passant am Plakat vorbei geht, dieses aber nicht betrachtet höchstens 0.6 ist. Die Firma Luegane vermutet jedoch, dass die Wahrscheinlichkeit, dass ein Passant am Plakat vorbei geht, dieses aber nicht betrachtet grösser als 0.6 ist. Mit obigen Daten soll das nun getestet werden (auf dem Signifikanzniveau 0.05).

d) Was sind die Null- und die Alternativhypothese?

e) Führen Sie einen geeigneten, exakten Test durch. Geben Sie den Verwerfungsbereich für X und den Testentscheid an.

f) Die Genfer Konkurrenzfirma Tasvue will nun einen ähnlichen Versuch durchführen. Im Speziellen will Tasvue den gleichen Versuchsaufbau (mit n Passanten) und den gleichen Test (ebenfalls zum Signifikanzniveau 0.05) anwenden. Beschreiben Sie ganz kurz, wie der Fehler 2. Art im Vergleich zu oben reduziert werden könnte.

3. (9 Punkte) Marco und Fabio sind Spieler des FC ETH. Sie verfügen über eine unterschiedliche Passgenauigkeit (Wahrscheinlichkeit, dass ein gespielter Pass auch beim Mitspieler ankommt.) Der Trainer des FC ETH schätzt, dass Marco eine Passgenauigkeit von 90% und Fabio eine Passgenauigkeit von 50% hat.

a) Wie gross ist die Wahrscheinlichkeit, dass Fabio vier Pässe hintereinander spielt, die alle beim Mitspieler ankommen?

b) Wie gross ist die Wahrscheinlichkeit, dass Marco drei Pässe hintereinander spielt wovon nur einer beim Mitspieler ankommen?

c) Wie gross ist die Wahrscheinlichkeit, dass ein Doppelpass zwischen Fabio und Marco gelingt?

d) Fabio spielt in einem Fussballspiel 20 Pässe. Welche Verteilung besitzt die Anzahl der erfolgreichen Pässe? Berechnen Sie den Erwartungswert und die Varianz dieser Verteilung.

e) Marco spielt in einer ganzen Meisterschaft 952 Pässe. Mit welcher stetigen Approximation kann man die erfolgreichen Pässe nun gut modellieren? (Verteilungsfamilie und Parameter angeben.)

f) Von den 952 gespielten Pässen von Marco in der Meisterschaft sind 831 beim Mitspieler angekommen. Ist die Einschätzung des Trainers zur Passgenauigkeit von Marco plausibel? Führen Sie einen zweiseitigen Test auf dem 5% Niveau durch. Geben Sie die Null- und die Alternativhypothese, die Teststatistik sowie den Testentscheid an. (Benutzen Sie die in e) bestimmte Approximation).

4. (7 Punkte) Es wird untersucht, inwiefern die Regenmenge (rain) durch die Luftfeuchtigkeit (humidity) und den Taupunkt (dewpoint) beinflusst wird. Es wurden 100 Messungen gemacht und folgendes Modell angepasst:

$$\text{rain}_i = \beta_0 + \beta_1 \cdot \text{humidity}_i + \beta_2 \cdot \text{dewpoint}_i + \varepsilon_i, \qquad \varepsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad i = 1, \ldots, 100.$$

Der (unvollständige) Regressionsoutput sieht wie folgt aus:

```
Residuals:
   Min     1Q Median     3Q    Max
-2.652 -0.566  0.075  0.660  2.359

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.0158     0.0990    0.16     0.87
humidity    -12.5499        ???   -1.24     0.22
dewpoint    -15.5496    10.1130     ???      ???
Residual standard error: 0.974 on ?? degrees of freedom
Multiple R-squared: 0.896,       Adjusted R-squared: 0.894
F-statistic:  420 on 2 and ?  DF,  p-value: <2e-16
```

1) Was ist der Standardfehler von $\hat{\beta}_1$? (auf eine signifikante Stelle genau)
   a) 10
   b) 1
   c) 100
   d) 0.1
2) Was ist der p-Wert von $\hat{\beta}_2$? (auf eine signifikante Stelle genau)
   a) 0.001
   b) 0.01
   c) 0.1
   d) 1
3) Mit wievielen Freiheitsgraden wurde der "residual standard error" berechnet?
   a) 97
   b) 98
   c) 2
   d) 3
4) Berechnen Sie das exakte 99%-Konfidenzintervall für β_2 (verwenden Sie 90 Freiheitsgrade).
   a) $[-18.2, -12.9]$
   b) $[-39.5, 8.4]$
   c) $[-16.9, -14.1]$
   d) $[-42.2, 11.1]$
5) Gibt es einen signifikanten Zusammenhang (5% Niveau) zwischen rain und mindestens einer der erklärenden Variablen?

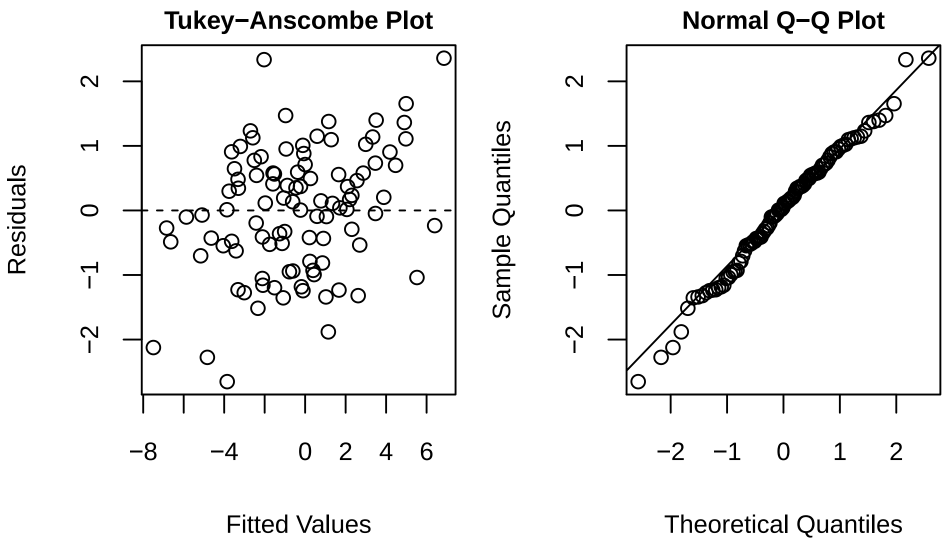
a) Ja.
b) Nein.
c) Keine Angabe möglich.

6) Wie ändert sich das multiple R-squared, wenn man eine dritte Variable aufnehmen würde, die für rain relevant und mit den beiden bisherigen Variablen unkorreliert ist?

a) Es wird grösser.
b) Es bleibt ungefähr gleich.
c) Es wird kleiner.
d) Keine Aussage möglich.

7) Betrachten Sie die nachfolgenden Plots. Welche der folgenden Aussagen trifft zu?

a) Alle Modellannahmen sind erfüllt.
b) Die Fehlervarianz ist nicht konstant, aber die Normalverteilungsannahme ist plausibel.
c) Die Fehlervarianz ist konstant, aber die Normalverteilungsannahme trifft nicht zu.
d) Sowohl konstante Fehlervarianz als auch Normalverteilungsannahme treffen nicht zu.



5. Die folgenden Aufgaben sind zufällig angeordnet und insbesondere nicht nach Schwierigkeitsgrad sortiert.

1) Seien A und B stochastisch unabhänginge Ereignisse. Dann gilt für das Komplement A^C von A:
   a) A^C und B sind stochastisch unabhänging.
   b) A^C und B sind nicht stochastisch unabhänging.
   c) Es kann keine Aussage über die stochastische Abhängigkeit von A^C und B gemacht werden.
2) Zwei Ereignisse A und B schließen sich aus. Welche Aussage trifft immer zu?
   a) $P(A \cup B) = P(A) + P(B)$.
   b) $P(A) = P(B)$.
   c) $P(A) + P(B) < 1$.
   d) $P(A) + P(B) = 1$.
3) In einer Kiste seien drei Spielkarten. Eine davon ist auf beiden Seiten schwarz, eine auf beiden Seiten weiss und eine auf einer Seite schwarz und auf der anderen Seite weiss. Es wird zufällig eine Karte gezogen und auf den Tisch gelegt. Angenommen, die Karte ist auf der sichtbaren Seite schwarz. Wie gross ist die Wahrscheinlichkeit, dass sie auf der anderen Seite weiss ist?
   a) 1/3.
   b) 1/2.
   c) 1/4.
   d) 2/3.
4) Für eine Zufallsvariable Z gelte $\mathrm{E}[Z] = 1$ und $\mathrm{Var}[Z] = 2$. Welchen Wert hat $\mathrm{E}[3Z + 2]$?
   a) 5.
   b) 6.
   c) 7.
   d) 8.
5) Für eine Zufallsvariable Z gelte $\mathrm{E}[Z] = 1$ und $\mathrm{Var}[Z] = 2$. Welchen Wert hat $\mathrm{Var}[3Z + 2]$?
   a) 6.
   b) 8.
   c) 18
   d) 20.
6) Eine Zufallsvariable $X \in [0, \infty)$ habe die Dichtefunktion $f(x) = \lambda^2 x e^{-\lambda x}$, $\lambda > 0$. Was ist die dazugehörige kummulative Verteilungsfunktion?
   a) $F(x) = \lambda x$, $x \in [0, \infty]$.
   b) $F(x) = 1 - \exp(-\lambda x)$, $x \in [0, \infty]$.
   c) $F(x) = 1 - (1 + \lambda x)e^{-\lambda x}$, $x \in [0, \infty]$.
   d) $F(x) = \Phi(x)$, $x \in \mathbb{R}$.
7) Eine Zufallsvariable $X \in \mathbb{R}$ habe die kummulative Verteilungsfunktion $F(x) = \alpha \frac{\exp(\lambda x)}{1+\exp(\lambda x)}$, $\lambda > 0$. Welchen Wert hat α?

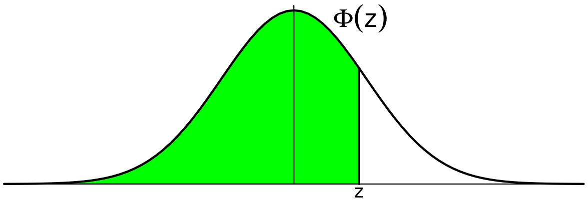
a) $\alpha = \frac{1}{4}$.
b) $\alpha = \frac{1}{2}$.
c) $\alpha = 1$.
d) Keine Aussage möglich.

8) Wir testen mit einem Binomialtest auf dem Signifikanzniveau 0.05 ob eine Münze gefälscht wurde so dass sie häufiger Kopf zeigt ($H_0 : \pi = 0.5$, $H_A : \pi > 0.5$). Wie gross ist die Wahrscheinlichkeit, dass wir die Münze als gefälscht bezeichnen (H_0 wird verworfen), wenn sie in Wahrheit fair (H_0 ist richtig) ist?
   a) Mindestens 0.05.
   b) Mindestens 0.95.
   c) Höchstens 0.05.
   d) Höchstens 0.95.

9) Es gilt für ein Ereignis E, dass $\text{odds}(E) = 4$. Was ist die Wahrscheinlichkeit für E?
   a) 0.2.
   b) 0.4
   c) 0.6.
   d) 0.8.

10) Die Wahrscheinlichkeit, dass eine zufällig gewählte Person die Krankheit K hat sei $P(K) = 0.01$. Für diese Krankheit wurde ein Test T entwickelt. Eine kranke Person wird mit Wahrscheinlichkeit $P(T|K) = 0.9$ positiv getestet, eine gesunde Person mit Wahrscheinlichkeit $P(T|K^C) = 0.1$. Nun wird eine zufällig gewählte Person positiv getestet. Wie gross ist ungefähr die Wahrscheinlichkeit, dass die Person krank ist?
   a) 0.0041.
   b) 0.041.
   c) 0.0083.
   d) 0.083.

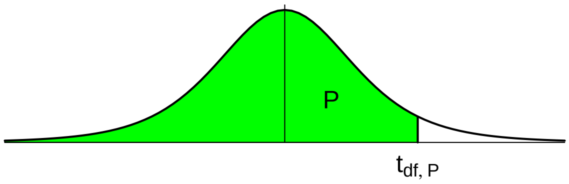
Tabelle der Kumulativen Normalverteilung $\Phi(z) = \mathrm{P}\left[Z \leq z\right], \; Z \sim \mathcal{N}(0,1)$



Bsp.: $\mathrm{P}\left[Z \leq 1.96\right] = 0.975$

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
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576