Wahrscheinlichkeit und Statistik (deutsch) FS 2020 Sessions-Prüfung

13.08.2020 **Zeit-Limit: 180 Minutes**

Nachname Vorname	**Legi-Nummer**	**Prüfungsnr.**
jeweils die ersten zwei Buchstaben	*letzte sechs Ziffern*	*Nicht ausfüllen*

Trag **jetzt die ersten zwei Buchstaben deines Namens und Vornamens ein**, sowie die **letzten sechs Ziffern deiner Legi-Nummer**. Wenn du separate Blätter abgibst, schreib obige, und **nur obige Informationen deutlich oben auf jedes Blatt**.

Die Prüfung enthält 10 Aufgaben und zusätzlich statistische Tabellen.

Punkte-Tabelle (nur für die Benotung, bitte leer lassen)

Aufgabe:	1	2	3	4	5	6	7	8	9	10	Total
Maximal:	10	10	10	10	10	10	10	10	10	10	100
Punkte:											
Kontrolle:											

Instruktionen

Hilfsmittel bei der Prüfung sind keine erlaubt (auch keine Taschenrechner oder Handys).

Vor der Prüfung

- Leg deine Legi (Studentenausweis) auf den Tisch.
- Fülle das Deckblatt wie oben angegeben aus. Öffne die Prüfung nicht.

Während der Prüfung:

- Lies die Fragen genau durch, und versuche so viele wie möglich zu beantworten. Starte jede Aufgabe auf einem neuen Blatt, und schreib die obigen Informationen auf jedes Blatt. Verwende keine rote oder grüne Tinte und keinen Bleistift.
- Du kannst Resultate aus der Vorlesung oder Formelsammlung ohne Beweis verwenden, wenn nicht explizit ein Beweis verlangt wird. Um volle Punkte zu bekommen, muss aber aus deiner Lösung klar hervorgehen, wie du auf deine Antwort gekommen bist. Ein korrektes Resultat allein gibt nicht immer alle Punkte.
- Vereinfache die Ergebnisse so weit wie möglich. Wenn ein numerischer Ausdruck einen Taschenrechner benötigt, kannst du einen numerischen Ausdruck als Antwort geben — z.B. wäre $\frac{\sqrt{3}}{7^2}$ eine gültige Lösung (aber beispielsweise $\int_0^1 \frac{\sqrt{3}}{7^2} dx$ oder $\frac{\sqrt{300}}{490}$ gäbe nicht die volle Punktzahl).

Nach der Prüfung

- Ordne deine Antwortblätter sinnvoll und leg alles in das bereitgestellte Couvert.
- Nachdem deine Prüfung von einem Supervisor eingesammelt wurde, bleib sitzen und folge den Anweisungen.

1. (10 Punkte) Eine Urne enthält 5 rote, 5 schwarze und 5 weisse Kugeln. Wir ziehen zufällig 3 Kugeln und legen sie jeweils direkt zurück, nachdem wir ihre Farbe notiert haben.

   (a) (3 Punkte) Beschreibe diese Situation durch einen geeigneten Wahrscheinlichkeitsraum.

   (b) (3 Punkte) Was ist die Wahrscheinlichkeit, dass die 3 gezogenen Kugeln genau zwei unterschiedliche Farben aufweisen?

   (c) (4 Punkte) Was ist die Wahrscheinlichkeit des Ereignisses in (b), wenn wir die Kugeln jeweils nicht zurücklegen?

2. (10 Punkte) Ein Oktaeder ist ein geometrischer Körper mit 8 gleichen Seiten. Es kann als ein Würfel mit 8 möglichen Ausgängen benutzt werden. Wir haben zwei faire Oktaeder-Würfel. Das eine hat auf seinen Seitenflächen die ungeraden Zahlen $1, 3, 5, \ldots, 15$, das andere die geraden Zahlen $2, 4, 6, \ldots, 16$. Wir werfen eine Münze, um zu entscheiden, mit welchem Würfel wir würfeln. Falls die Münze Kopf zeigt, so nehmen wir den Würfel mit den ungeraden Zahlen, und andernfalls den mit den geraden Zahlen.

   (a) (3 Punkte)

      (i) Beschreibe dieses Experiment mit einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Benutze ein Laplace-Modell.

      (ii) Was ist die Kardinalität $|\mathcal{F}|$ von $\mathcal{F}$? Gib Beispiele von Ereignissen E_1, E_2, E_3, E_4 in $\mathcal{F}$ mit $\mathbb{P}[E_i] \neq \mathbb{P}[E_j]$ für alle $i \neq j$.

   (b) (2 Punkte) Was ist die Wahrscheinlichkeit, dass beim Würfeln eine durch 3 teilbare Zahl herauskommt?

   (c) (5 Punkte)

      (i) Definiere Zufallsvariablen $X : \Omega \to \mathbb{R}$ und $Y : \Omega \to \mathbb{R}$, so dass X und Y die Ergebnisse des Münzwurfs bzw. des Würfelwurfs beschreiben.

      (ii) Sind die Zufallsvariablen X und Y unabhängig? Gib einen Beweis für deine Antwort.

3. (10 Punkte) Eine Auto-Abschleppfirma bedient eine Autobahnstrecke von 200 Meilen. Die Firma befindet sich innerhalb der Strecke in einem Abstand von b Meilen von einem Streckenende. Autopannen treten gleichverteilt entlang der Strecke auf, und die Abschleppfahrzeuge fahren mit einer konstanten Geschwindigkeit von 40 Meilen pro Stunde. Wir nehmen an, dass nach einem Anruf bei der Abschleppfirma sofort ein Abschleppfahrzeug vom Firmenstandort zum Pannenort losfährt.

   (a) (3 Punkte) Berechne (als Funktion von b) die Wahrscheinlichkeit, dass die Wartezeit auf ein Abschleppfahrzeug mehr als 30 Minuten beträgt. Welchen numerischen Wert erhält man für $b = 50$?

   (b) (2 Punkte) Berechne den Erwartungswert der Wartezeit auf ein Abschleppfahrzeug, wieder als Funktion von b. Welchen numerischen Wert erhält man für $b = 50$?

   (c) (5 Punkte) Die Abschleppfirma möchte damit werben, dass die Wartezeit auf das Abschleppen eine kleine Standardabweichung hat. Was ist dann die beste Wahl für den Standort, und wie gross ist der numerische Wert der zugehörigen Standardabweichung?

4. (10 Punkte) Um die Zuverlässigkeit eines Krankheitstests zu untersuchen, betrachten wir eine Gruppe von Leuten. Wir wissen, dass eine zufällig ausgewählte Person die Krankheit mit einer Wahrscheinlichkeit von $\frac{1}{3}$ hat. Das Testverfahren sucht bestimmte Marker im Blut, und man stellt fest, dass die Wahrscheinlichkeit für ein korrektes positives Ergebnis (d.h. bei einer kranken Person werden Marker im Blut gefunden) $\frac{4}{5}$ beträgt. Man stellt auch fest, dass die Wahrscheinlichkeit für ein falsches positives Ergebnis (d.h. bei einer gesunden Person werden Marker im Blut gefunden) $\frac{3}{10}$ beträgt.

   (a) (1 Punkt) Gib für zwei beliebige Ereignisse A, B die allgemeine Definition der bedingten Wahrscheinlichkeit $\mathbb{P}[A|B]$.

   (b) (4 Punkte) Berechne die Wahrscheinlichkeit des Ereignisses C, dass eine Person krank ist und der Test positiv ausfällt.

   (c) (2 Punkte) Berechne die Wahrscheinlichkeit, dass das Testergebnis positiv ist.

   (d) (3 Punkte) Was ist die Wahrscheinlichkeit, dass eine Person krank ist, wenn Marker im Blut gefunden wurden (so dass der Test positiv ausfällt)?

5. (10 Punkte) Frank und Johann verbringen jede Woche die zufälligen Zeiten X (für Frank) und Y (für Johann) für ihre Übungsblätter in Wahrscheinlichkeit und Statistik. Die gemeinsame Dichte von X und Y ist gegeben durch

$$f_{X,Y}(x,y) = \frac{1}{y} I_{(0,1]}(y) I_{(0,y]}(x).$$

   (a) (3 Punkte) Bestimme die Randdichten von X und Y.

   (b) (1 Punkt) Sind X und Y unabhängig? Begründe deine Antwort.

   (c) (4 Punkte) Berechne den erwarteten Zeitaufwand $\mathbb{E}[X]$ und $\mathbb{E}[Y]$ von Frank bzw. Johann. Welcher der beiden ist im Schnitt schneller fertig?

   (d) (2 Punkte) Berechne die Kovarianz von X und Y.

6. (10 Punkte) In einer grossen Warensendung gibt es einen Anteil von 10% von defekten Einheiten. Wir nehmen eine Stichprobe von n Einheiten und untersuchen diese auf Defekte. Dabei nennen wir D_n die Anzahl der defekten Einheiten in dieser Stichprobe.

   (a) (4 Punkte) Gib ein probabilistisches Modell, um diese Situation zu beschreiben. Was ist die Verteilung von D_n, und was sind deine Annahmen?

   (b) (6 Punkte) Ein Kunde sagt uns, dass er bei einer Bestellung von n Einheiten höchstens einen Anteil von 13% von defekten Einheiten in seiner Lieferung akzeptieren wird. Wie gross müssen wir n mindestens wählen, um diesen Kunden unter den Annahmen in (a) mit einer approximativen Wahrscheinlichkeit von 99% zufriedenstellen zu können?
   Hinweis: $\sqrt{5.41} \approx 2.3259$

7. (10 Punkte) Wir betrachten i.i.d. Zufallsvariablen $X_1, X_2, \ldots$, wobei alle X_i die uniforme Verteilung auf $[a, b]$ mit $a < 0 < b$ besitzen.

   (a) (4 Punkte) Bestimme den Maximum-Likelihood-Schätzer T_n für a aufgrund einer Stichprobe $(X_1, \ldots, X_n)$ vom Umfang n.

   Hinweis: Die Likelihood-Funktion ist nicht differenzierbar.

   (b) (6 Punkte) Aus der gleichen Stichprobe $(X_1, \ldots, X_n)$ schätzen wir b durch den Schätzer $U_n := c_n \max_{i=1,\ldots,n} X_i$ für eine Konstante c_n. Wie muss c_n gewählt werden, damit U_n für $a = -b$ erwartungstreu wird?

8. (10 Punkte) In einem Schlaflabor werden zwei verschiedene Typen von Schlaftabletten untersucht. Bei der Studie nahmen 11 Personen teil; jede nahm in einer Nacht eine Tablette vom Typ I und in einer anderen Nacht eine Tablette vom Typ II vor dem Einschlafen. Die resultierenden Schlafzeiten in Stunden waren

Person	1	2	3	4	5	6	7	8	9	10	11
Schlaftablette Typ I	3.8	5.5	5.6	6.3	6.9	7.1	7.4	7.8	7.9	8.0	8.1
Schlaftablette Typ II	3.0	4.2	6.4	4.5	4.6	4.6	7.0	6.7	7.3	6.7	6.9

Vor der Studie war unklar, welche der Tabletten die stärkere Wirkung hat.

(a) (1 Punkt) Haben wir hier eine gepaarte oder eine ungepaarte Stichprobe? Begründe deine Antwort.

(b) (3 Punkte) Welcher Test wäre angebracht, wenn man keine spezifische Verteilung annehmen will? Was sind seine Voraussetzungen?

(c) (6 Punkte) Formuliere die Nullhypothese H_0 für diese Situation. Kann H_0 auf dem Signifikanzniveau $\alpha = 0.05$ mit dem Test aus (b) verworfen werden?

9. (10 Punkte) Unterhalb einer Kläranlage wurden 16 Wasserproben aus einem grossen Fluss entnommen und jeweils deren Ammoniumkonzentration X_i (angegeben in $\mu\mathrm{gNH}_4^+/\ell$) mit einem Messgerät bestimmt. Wir wählen die Messzeitpunkte zufällig und dürfen annehmen, dass die Messergebnisse einer Normalverteilung folgen. Der Mittelwert der Proben ergab $\overline{x}_{16} = 204.2$, und die empirische Varianz in der Stichprobe beträgt $s_{16}^2 = (9.8)^2 = 96.04$.

   Wir wollen nun wissen, ob mit diesen Daten eine Überschreitung des gesetzlichen Grenzwerts von 200 $\mu\mathrm{gNH}_4^+/\ell$ nachgewiesen werden kann (auf dem 5%-Niveau).

   (a) (2 Punkte) Die Standardabweichung der Messungen sei im Voraus aufgrund früherer Studien bekannt und betrage 10 $\mu\mathrm{gNH}_4^+/\ell$. Finde einen geeigneten statistischen Test, um die obige Hypothese zu überprüfen. Wie lauten deine Modellannahmen?

   (b) (4 Punkte) Führe den Test aus (a) durch. Gib dazu folgendes explizit an: die Null- und Alternativhypothesen H_0 und H_A, die Teststatistik T_n, den realisierten Wert von T_n, den Verwerfungsbereich und das Testergebnis.

   (c) (3 Punkte) Wie wahrscheinlich ist es, dass man mit dem obigen Test eine Grenzwertüberschreitung nachweisen kann, wenn die wahre Ammoniumkonzentration bei 205 $\mu\mathrm{gNH}_4^+/\ell$ liegt?

   (d) (1 Punkt) Wie wahrscheinlich ist es, dass man mit dem obigen Test eine Grenzwertüberschreitung nachweist, wenn die wahre Ammoniumkonzentration bei 200 $\mu\mathrm{gNH}_4^+/\ell$ liegt?

10. (10 Punkte) Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von Poisson-verteilten Zufallsvariablen. Jedes X_n hat einen Poisson-Parameter λ_n, und es gelte $0 < \lambda_n \leq C$ für eine Konstante $C \in (0, \infty)$.

    (a) (3 Punkte) Zeige, dass gilt

    $$\mathbb{P}\left[X_n \geq n \text{ für unendlich viele } n\right] = 0.$$

    (b) (3 Punkte) Diskutiere ausführlich, ob oder wann die Summe $S_n := \sum_{i=1}^{n} X_i$ eine Poisson-Verteilung hat. Beweise deine Aussagen detailliert, ohne Resultate über Poisson-Verteilungen aus der Vorlesung zu benutzen

    (c) (4 Punkte) Beweise, dass für alle n gilt

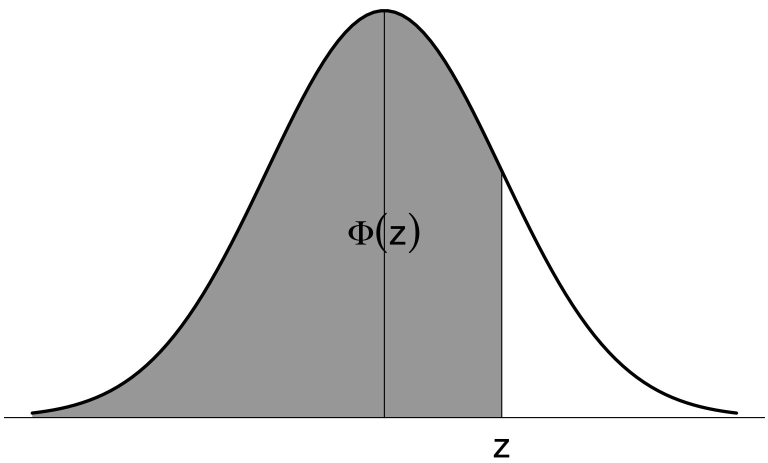
    $$\mathbb{P}\left[X_n \geq b\right] \leq e^{b - b\log\frac{b}{C} - C},$$

    sofern $b > C$.

⋆ ⋆ ⋆ ⋆ ⋆ ⋆ ⋆

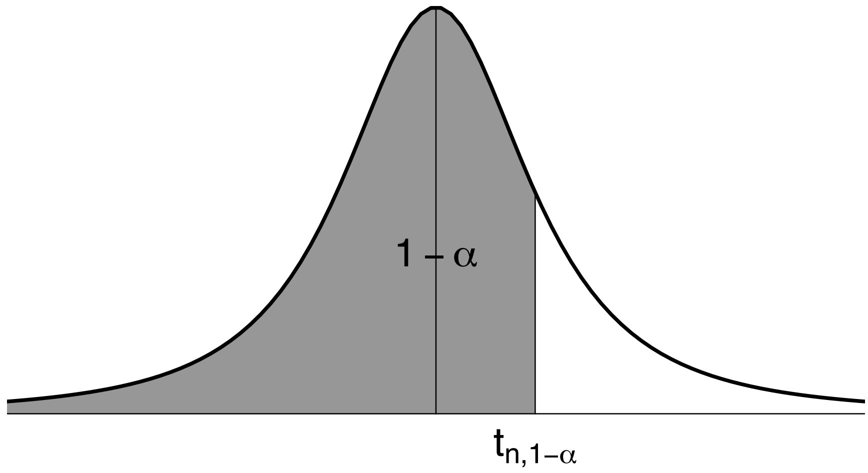
⋆ ⋆ ⋆ Viel Glück! ⋆ ⋆ ⋆

Tabellen



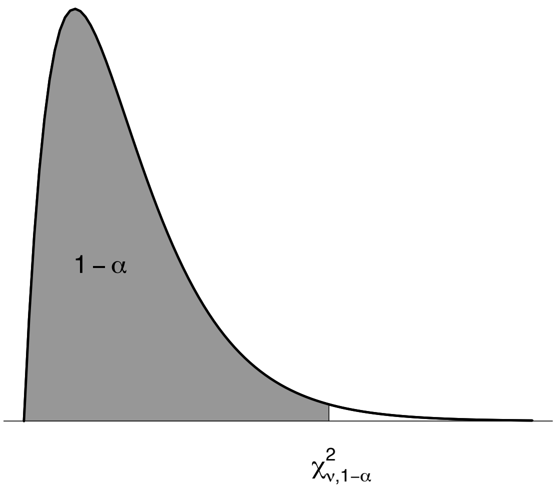
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	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

Tabelle der Standard-Normalverteilungsfunktion $\Phi(z) = P[Z \leq z]$ mit $Z \sim \mathcal{N}(0,1)$


1 − α
t_{n,1−α}

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
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576

Ausgewählte Quantile $t_{n,1-\alpha}$ der t-Verteilung; in der Tabelle ist $n = df$.
Für $df = \infty$ erhält man die Quantile $z_{1-\alpha}$ der Standard-Normalverteilung.



	$p = 0.90$	$p = 0.95$	$p = 0.975$	$p = 0.999$	$p = 0.9995$
$\nu = 1$	2.7055	3.8415	5.0239	10.8276	12.1157
$\nu = 2$	4.6052	5.9915	7.3778	13.8155	15.2018
$\nu = 3$	6.2514	7.8147	9.3484	16.2662	17.7300
$\nu = 4$	7.7794	9.4877	11.1433	18.4668	19.9974
$\nu = 5$	9.2364	11.0705	12.8325	20.5150	22.1053
$\nu = 6$	10.6446	12.5916	14.4494	22.4577	24.1028
$\nu = 7$	12.0170	14.0671	16.0128	24.3219	26.0178
$\nu = 8$	13.3616	15.5073	17.5345	26.1245	27.8680
$\nu = 9$	14.6837	16.9190	19.0228	27.8772	29.6658
$\nu = 10$	15.9872	18.3070	20.4832	29.5883	31.4198
$\nu = 11$	17.2750	19.6751	21.9200	31.2641	33.1366
$\nu = 12$	18.5493	21.0261	23.3367	32.9095	34.8213

Ausgewählte Quantile $\chi^2_{\nu,1-\alpha}$ der Chiquadrat-Verteilung; in der Tabelle ist $p = 1 - \alpha$.