

AUFGABENSTELLUNG UND ERGEBNISSE ZUR
 BACHELOR-PRÜFUNG
 DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHKEITSRECHNUNG
 WINTERSEMESTER 2023/24

apl. Prof. Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Die Klausur besteht aus insgesamt 9 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 18 + 15 + 13 + 17 + 7 + 8 + 16 + 17 + 9) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbstgestaltete DIN-A4 Blätter bzw. 4 selbstgestaltete (einseitige) DIN-A4 Seiten. Benötigte Tabellen finden Sie am Ende dieses Klausurheftes.
- Bei der Korrektur werden **nur** die Lösungen auf diesen Blättern berücksichtigt.
- Bei mehreren Lösungsvorschlägen muss die gültige Lösung eindeutig gekennzeichnet sein.
- Mit Ausnahme der Multiple-Choice-Aufgaben muss der Lösungsweg klar ersichtlich sein.
- Alle Punkte innerhalb von Zahlen sind stets Dezimalpunkte (und keine Tausenderpunkte).

Bewertungsteil — Bitte nicht beschreiben						
Aufgabe	(a)	(b)	(c)	(d)	(e)	Σ
1		■	■	■	■	
2		■	■	■	■	
3						
4						
5					■	
6				■	■	
7					■	
8					■	
9				■	■	
Σ						

Aufgabe 1 (18 Punkte)

Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.

Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

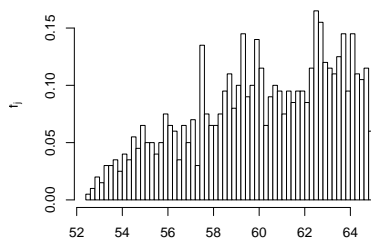
- | | wahr | falsch |
|--|-------------------------------------|-------------------------------------|
| 1. Beim Modus (oder Modalwert) eines kardinalskalierten Merkmals X handelt es sich um die größte angenommene Merkmalsausprägung von X . | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 2. Am Vorzeichen des Bravais-Pearsonschen Korrelationskoeffizienten zweier Merkmale X und Y kann man erkennen, ob X auf Y wirkt oder umgekehrt (Y auf X). | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 3. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) mit $P(B \setminus A) = 0.3$ und $P(A \setminus B) = 0.4$. Dann gilt stets $P(A \cup B) \geq 0.7$. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 4. Die Wahrscheinlichkeit, dass sich bei rein zufälliger Anordnung der Ziffern 1, 1, 1, 3, 3 und 7 die Zahl 317131 ergibt, beträgt (ggf. gerundet) 1.667%. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 5. Es seien A , B und C drei Ereignisse in einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) mit $P(C) > 0$. Gilt sowohl $P(A C) < P(A)$ als auch $P(B C) < P(B)$, dann gilt stets auch $P(A \cap B C) < P(A \cap B)$. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 6. Verteilungsfunktionen stetiger Zufallsvariablen sind stets streng monoton wachsend auf \mathbb{R} . | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 7. Binomialverteilte Zufallsvariablen sind nie symmetrisch (um ihren Erwartungswert) verteilt. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 8. Die Funktion $p_X : \mathbb{R} \rightarrow \mathbb{R}$ mit | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| $p_X(x) = \begin{cases} \frac{1}{10} & \text{für } x \in \{11, 12, \dots, 20\} \\ 0 & \text{sonst} \end{cases}$ | | |
| erfüllt alle Voraussetzungen für Wahrscheinlichkeitsfunktionen diskreter Zufallsvariablen. | | |
| 9. Existieren alle beteiligten Momente, so gilt: Ist die Kovarianz zweier Zufallsvariablen X und Y negativ, dann ist die Varianz der Differenz von X und Y größer als die Summe der einzelnen Varianzen (von X und Y). | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

Aufgabe 2 (15 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.

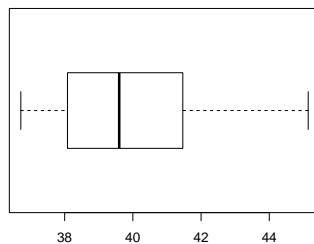
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



- (a) leptokurtisch und rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

2. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



- (a) leptokurtisch und rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

3. Die Ränge $rg(X)_1, \dots, rg(X)_8$ zur (der Einfachheit halber sortierten) Urliste

sehr gut, sehr gut, sehr gut, gut, gut, befriedigend, befriedigend, ausreichend

des ordinalskalierten Merkmals X lauten:

- (a) 1, 1, 1, 4, 4, 6, 6, 8
- (b) 1, 1, 1, 2, 2, 4, 3, 4
- (c) 2, 2, 2, 4.5, 4.5, 6.5, 6.5, 8
- (d) 1.5, 1.5, 1.5, 4.5, 4.5, 6.5, 6.5, 8

4. In einer Eisdiele werden insgesamt 13 verschiedene Eissorten als Kugeln für Eiswaffeln angeboten. Wie viele Möglichkeiten haben Sie zur Zusammenstellung einer Eiswaffel mit 3 Kugeln, wenn Sie nicht mehrere Kugeln derselben Sorte in der Waffel haben möchten und die Reihenfolge, in der die Kugeln in der Waffel angeordnet sind, egal ist?

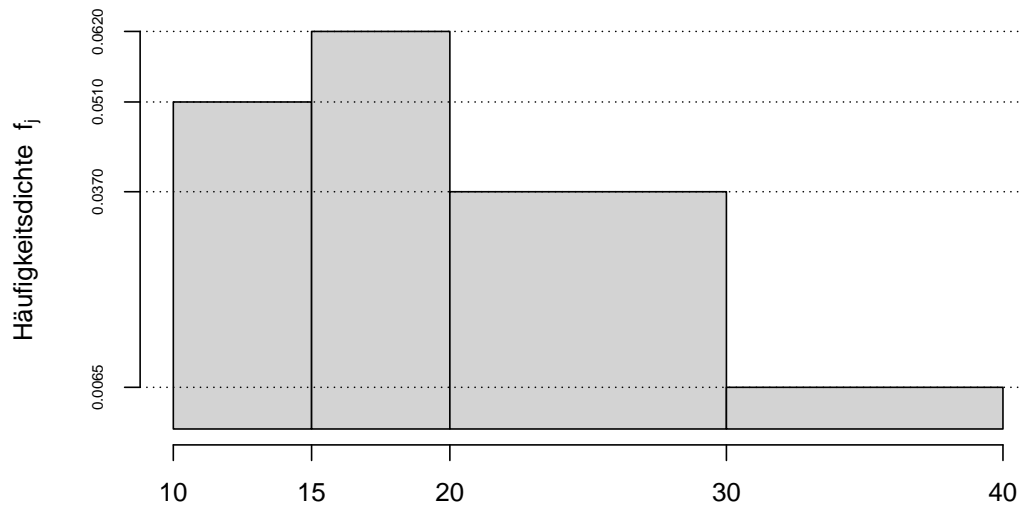
- (a) 3^{13} Möglichkeiten.
- (b) 13^3 Möglichkeiten.
- (c) $(13)_3 = \frac{13!}{10!}$ Möglichkeiten.
- (d) $\binom{13}{3} = \frac{13!}{3! \cdot 10!}$ Möglichkeiten.

5. Sind X_1 , X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(30, 6^2)$, $X_2 \sim N(70, 12^2)$ und $X_3 \sim N(50, 4^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

- (a) $N(50, 22^2)$ -Verteilung.
- (b) $N(50, 14^2)$ -Verteilung.
- (c) $N(150, 22^2)$ -Verteilung.
- (d) $N(150, 14^2)$ -Verteilung.

Aufgabe 4 (6 + 4 + 3 + 2 + 2 = 17 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 200$:



- Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.
- Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.
- Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 19.895?
- Welche Näherung für die Anzahl von Merkmalswerten zwischen 20 und 35 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?
- Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten das obere Quartil.

Ergebnisse (ohne Begründung/Rechenweg):

- (a) Klassierung:

Nr. j	Klasse $K_j = (k_{j-1}, k_j]$	Klassenbreite b_j	Klassenmitte m_j	absolute Häufigkeit h_j	relative Häufigkeit $r_j = \frac{h_j}{n}$	Häufigkeitsdichte $f_j = \frac{r_j}{b_j}$	Verteilungsfunktion $F(k_j)$
1	(10, 15]	5	12.5	51	0.255	0.0510	0.255
2	(15, 20]	5	17.5	62	0.310	0.0620	0.565
3	(20, 30]	10	25.0	74	0.370	0.0370	0.935
4	(30, 40]	10	35.0	13	0.065	0.0065	1.000

(b) (Approximative) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x \leq 10 \\ 0.051 \cdot (x - 10) & \text{für } 10 < x \leq 15 \\ 0.255 + 0.062 \cdot (x - 15) & \text{für } 15 < x \leq 20 \\ 0.565 + 0.037 \cdot (x - 20) & \text{für } 20 < x \leq 30 \\ 0.935 + 0.0065 \cdot (x - 30) & \text{für } 30 < x \leq 40 \\ 1 & \text{für } x > 40 \end{cases}$$

(c) Mittelwert (näherungsweise): 20.137, relative Abweichung vom exakten Wert: 0.01216 bzw. 1.216%

(d) Mit emp. Verteilungsfunktion genäherte Anzahl: 80.5

(e) Oberes Quartil: 25

Aufgabe 5 (1 + 2 + 2 + 2 = 7 Punkte)

Eine Urne enthält 128 gleichartige Kugeln, von denen 8 blau und gestreift, 32 rot und gestreift, 64 blau und gepunktet sowie 24 rot und gepunktet sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

- (a) eine rein zufällig aus der Urne entnommene Kugel rot und gestreift ist?
- (b) eine rein zufällig aus der Urne entnommene Kugel blau ist?
- (c) eine rein zufällig aus der Urne entnommene Kugel gepunktet ist, wenn man weiß, dass sie rot ist?
- (d) bei dreimaligem rein zufälligen Ziehen *mit Zurücklegen* die erste Kugel blau, die zweite Kugel rot und die letzte Kugel gestreift ist?

Ergebnisse (ohne Begründung/Rechenweg):

- (a) $\frac{1}{4}$
- (b) $\frac{9}{16}$
- (c) $\frac{3}{7}$
- (d) $\frac{315}{4096}$

Aufgabe 6 (4 + 2 + 2 = 8 Punkte)

An einer seltenen Krankheit seien 1.5% der Bevölkerung einer bestimmten Altersgruppe erkrankt. Zum Einsatz in flächendeckenden Früherkennungsuntersuchungen existiere ein medizinisches Diagnoseverfahren, welches erkrankte Personen mit einer Wahrscheinlichkeit von 96% (korrekterweise) auch als krank einstuft, bei gesunden (bzw. nicht an dieser Krankheit erkrankten) Personen allerdings mit einer Wahrscheinlichkeit von 3% (fälschlicherweise) ebenfalls eine entsprechende Erkrankung diagnostiziert.

- (a) Mit welcher Wahrscheinlichkeit wird ein Patient in der betrachteten Altersgruppe im Rahmen einer Früherkennungsuntersuchung als krank eingestuft?
- (b) Mit welcher Wahrscheinlichkeit wird sich eine positive Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe als falsch herausstellen?
- (c) Wie groß ist die Wahrscheinlichkeit für eine richtige Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe?

Ergebnisse (ohne Begründung/Rechenweg):

- (a) 0.04395
- (b) 0.6724
- (c) 0.9699

Aufgabe 7 (5 + 1 + 6 + 4 = 16 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} -\frac{1}{4}x + \frac{1}{2} & \text{für } 0 \leq x < 2 \\ \frac{1}{6}x - \frac{1}{4} & \text{für } 2 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

- (a) Berechnen Sie die Verteilungsfunktion F_X von X .
- (b) Berechnen Sie $P(\{X < 1\})$.
- (c) Bestimmen Sie den Erwartungswert $E(X)$.
- (d) Bestimmen Sie das untere Quartil von X .

Ergebnisse (ohne Begründung/Rechenweg):

- (a) Verteilungsfunktion von X :

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ -\frac{1}{8}x^2 + \frac{1}{2}x & \text{für } 0 < x \leq 2 \\ \frac{1}{12}x^2 - \frac{1}{4}x + \frac{2}{3} & \text{für } 2 < x \leq 4 \\ 1 & \text{für } x > 4 \end{cases}$$

- (b) $P(\{X < 1\}) = \frac{3}{8}$
- (c) $E(X) = 1.9\bar{4}$
- (d) $x_{0.25} = 0.5858$

Aufgabe 8 (2 + 3 + 9 + 3 = 17 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y) :

$X \setminus Y$	-2	0	2	$p_{i\cdot}$
1	$\frac{1}{40}$	$\frac{3}{20}$	$\frac{3}{40}$	
2	$\frac{3}{40}$	$\frac{1}{5}$	$\frac{1}{10}$	
4	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{40}$	
$p_{\cdot j}$				

- (a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.
- (b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.
- (c) Berechnen Sie $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Cov}(X, Y)$ sowie $\text{Korr}(X, Y)$.
- (d) Berechnen Sie $E(4X - 2Y)$ sowie $\text{Var}(4X - 2Y)$.

Ergebnisse (ohne Begründung/Rechenweg):

- (a) Ergänzte Tabelle:

$X \setminus Y$	-2	0	2	$p_{i\cdot}$
1	$\frac{1}{40}$	$\frac{3}{20}$	$\frac{3}{40}$	$\frac{1}{4}$
2	$\frac{3}{40}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{3}{8}$
4	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{40}$	$\frac{3}{8}$
$p_{\cdot j}$	$\frac{1}{5}$	$\frac{11}{20}$	$\frac{1}{4}$	1

- (b) Tabelle der bedingten Wahrscheinlichkeitsfunktionswerte von $Y|X = x_i, i \in \{1, 2, 3\}$:

y_j	-2	0	2
$p_{Y X=1}(y_j)$	$\frac{1}{10}$	$\frac{3}{5}$	$\frac{3}{10}$
$p_{Y X=2}(y_j)$	$\frac{1}{5}$	$\frac{8}{15}$	$\frac{4}{15}$
$p_{Y X=4}(y_j)$	$\frac{4}{15}$	$\frac{8}{15}$	$\frac{1}{5}$

- (c) Es gilt: $E(X) = \frac{5}{2}$, $E(Y) = 0.1$, $\text{Var}(X) = \frac{3}{2}$, $\text{Var}(Y) = 1.79$, $\text{Cov}(X, Y) = -\frac{1}{4}$, $\text{Korr}(X, Y) = -0.1526$
- (d) $E(4 \cdot X - 2 \cdot Y) = 9.8$, $\text{Var}(4 \cdot X - 2 \cdot Y) = 35.16$

Aufgabe 9 (2 + 4 + 3 = 9 Punkte)

Die Zufallsvariablen X_1, \dots, X_{625} seien unabhängig identisch $B(1, 0.8)$ -verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{625} X_i = X_1 + \dots + X_{625}$$

bezeichnet.

- (a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $E(Y)$ und Varianz $\text{Var}(Y)$ an.
- (b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 485 und 508 annimmt.
- (c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.90-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 12!

Ergebnisse (ohne Begründung/Rechenweg):

- (a) $Y \sim B(625, 0.8)$, $E(Y) = 500$, $\text{Var}(Y) = 100$.
- (b) $P\{485 \leq Y \leq 508\} \approx 0.7213$
- (c) $y_{0.90} \approx 512.8$

Tabelle zur Standardnormalverteilung

$$F_{N(0,1)}(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

$$\Phi(x) = 1 - \Phi(-x)$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	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