Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Deskriptive Statistik und Wahrscheinlichkeitsrechnung
Wintersemester 2017/18

Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Die Klausur besteht aus insgesamt 10 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 16 + 12 + 15 + 18 + 5 + 8 + 17 + 5 + 15 + 9) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbsterstellte DIN-A4 Blätter bzw. 4 selbsterstellte (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.

Bewertungsteil — Bitte nicht beschreiben						
Aufgabe	(a)	(b)	(c)	(d)	(e)	Σ
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
Σ						

Aufgabe 1 (16 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!).

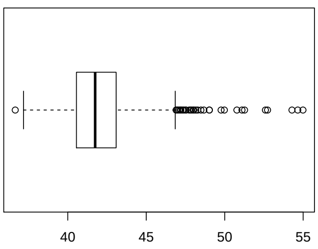
	Aussage	wahr	falsch
1.	Stetige Merkmale sind stets kardinalskaliert.	■	☐
2.	Empirische Verteilungsfunktionen F sind stets streng monoton wachsende Funktionen.	☐	■
3.	Die jährlichen Inflationsraten in Deutschland betrugen in den Jahren 2014–2017 im Einzelnen 0.9%, 0.3%, 0.5% und 1.8%. Damit beträgt die durchschnittliche jährliche Inflationsrate in diesem Zeitraum (gerundet) 0.875%.	☐	■
4.	Die Wahrscheinlichkeit, mit einem fairen (6-seitigen) Würfel dreimal nacheinander **keine** 6 zu würfeln, beträgt weniger als 60%.	■	☐
5.	Wenn Sie bei der Bearbeitung von Aufgabe 2 dieser Klausur (Vier MC-Teilaufgaben mit jeweils vier Antwortmöglichkeiten) jeweils genau eine der vier möglichen Antworten ankreuzen, haben Sie insgesamt 256 Bearbeitungsmöglichkeiten.	■	☐
6.	Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B \in \mathcal{F}$ zwei Ereignisse mit $P(A) > 0$, $P(B) > 0$ sowie $P(A) = P(A\|B)$. Dann gilt auch $P(B) = P(B\|A)$.	■	☐
7.	Es sei f_X eine Dichtefunktion zu einer stetigen Zufallsvariablen X. Dann gilt stets: $$\int_{-\infty}^{\infty} f_X(x)dx = 1$$	■	☐
8.	Die Zufallsvariablen X, Y und Z seien stochastisch unabhängig, ferner gelte $X \sim B(10, 0.2)$, $Y \sim B(15, 0.2)$ und $Z \sim B(25, 0.2)$. Dann ist die Summe $S := X + Y + Z$ ebenfalls binomialverteilt, es gilt genauer $S \sim B(50, 0.2)$.	■	☐

<u>Aufgabe 2</u> (12 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 der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ■

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

2. Die Wahrscheinlichkeit, durch eine rein zufällige Aneinanderreihung der Buchstaben E,K,K,O,R,R,R,T und U das Wort „KORREKTUR" zu erhalten, beträgt:

(a) $\frac{1! + 2! + 1! + 3! + 1! + 1!}{9!}$ ☐

(b) $\frac{1}{1 \cdot 2 \cdot 1 \cdot 3 \cdot 1 \cdot 1}$ ☐

(c) $\frac{1! \cdot 2! \cdot 1! \cdot 3! \cdot 1! \cdot 1!}{9!}$ ■

(d) $\frac{1}{9!}$ ☐

3. Das Merkmal X des zweidimensionalen Merkmals (X, Y) sei ordinalskaliert, das Merkmal Y kardinalskaliert. Damit ist die Berechnung folgender Abhängigkeitsmaße zwischen X und Y immer möglich:

(a) Pearsonscher Korrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐

(b) Spearmanscher Rangkorrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ■

(c) Nur Spearmanscher Rangkorrelationskoeffizient ☐

(d) Nur korrigierter Pearsonscher Kontingenzkoeffizient ☐

4. Der zweidimensionale diskrete Zufallsvektor (X, Y) besitze 8 Trägerpunkte, die alle auf einer Geraden mit Steigung 0.1 liegen. Dann gilt:

(a) $\mathrm{Korr}(X, Y) = -0.1$ ☐

(b) $\mathrm{Korr}(X, Y) = 0.1$ ☐

(c) $\mathrm{Korr}(X, Y) = 0.8$ ☐

(d) $\mathrm{Korr}(X, Y) = 1$ ■

Aufgabe 3 (3 + 3 + 1 + 5 + 3 = 15 Punkte)

Bei einer Umfrage wurden 25 Personen befragt, wie viele Smartphones sie in den vergangenen fünf Jahren gekauft haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

(c) Wie groß ist der Anteil der Personen in der Umfrage, die mindestens 3 Smartphones in den vergangenen fünf Jahren gekauft haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein unteres Quartil, ein oberes Quartil und den zugehörigen Interquartilsabstand des Merkmals X.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Häufigkeitstabelle:

a_j	0	1	2	3	4	Σ
$h(a_j)$	3	6	8	6	2	25
$r(a_j)$	0.12	0.24	0.32	0.24	0.08	1.00

(b) Empirische Verteilungsfunktion:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 0 \\ 0.12 & \text{für} \quad 0 \leq x < 1 \\ 0.36 & \text{für} \quad 1 \leq x < 2 \\ 0.68 & \text{für} \quad 2 \leq x < 3 \\ 0.92 & \text{für} \quad 3 \leq x < 4 \\ 1.00 & \text{für} \quad x \geq 4 \end{cases}$$

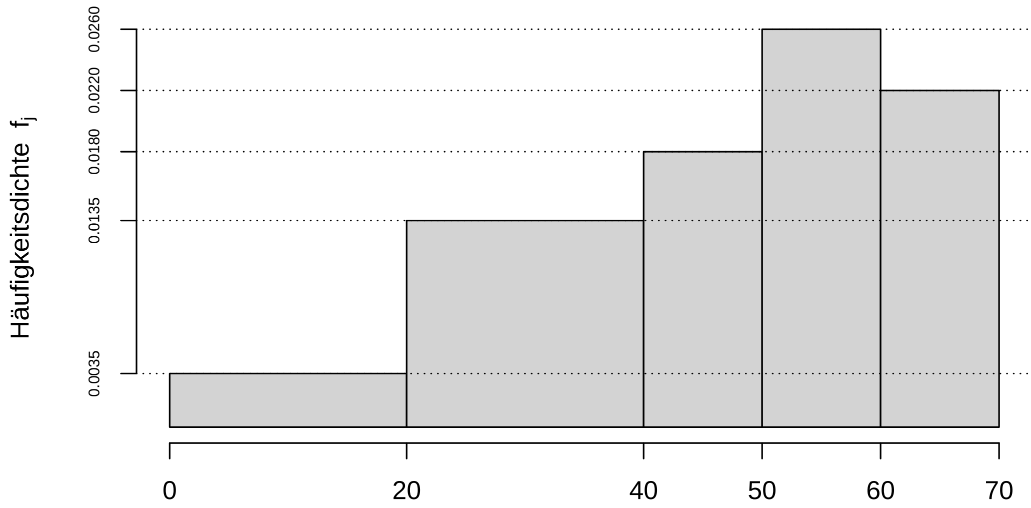
(c) Anteil der Personen, die mindestens 3 Smartphones in den vergangenen fünf Jahren gekauft haben: $0.32 = 32\%$

(d) $\overline{x} = 1.92$, $s^2 = 1.2736$

(e) $x_{0.25} = 1, x_{0.75} = 3, \text{IQA: } 2$

Aufgabe 4 (7 + 4 + 3 + 2 + 2 = 18 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 100$:



(a) 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.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) 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 45.687?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 30 und 50 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

(e) Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten das untere Quartil.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Klassierung:

Nr. j	Klasse $K_j = (k_{j-1}, k_j]$	Klassen-breite b_j	Klassen-mitte m_j	absolute Häufigkeit h_j	relative Häufigkeit $r_j = \frac{h_j}{n}$	Häufigkeits-dichte $f_j = \frac{r_j}{b_j}$	Verteilungs-funktion $F(k_j)$
1	$(0, 20]$	20	10	7	0.07	0.0035	0.07
2	$(20, 40]$	20	30	27	0.27	0.0135	0.34
3	$(40, 50]$	10	45	18	0.18	0.0180	0.52
4	$(50, 60]$	10	55	26	0.26	0.0260	0.78
5	$(60, 70]$	10	65	22	0.22	0.0220	1.00

(b) (Approximative) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ 0.0035 \cdot (x - 0) & \text{für } 0 < x \leq 20 \\ 0.07 + 0.0135 \cdot (x - 20) & \text{für } 20 < x \leq 40 \\ 0.34 + 0.018 \cdot (x - 40) & \text{für } 40 < x \leq 50 \\ 0.52 + 0.026 \cdot (x - 50) & \text{für } 50 < x \leq 60 \\ 0.78 + 0.022 \cdot (x - 60) & \text{für } 60 < x \leq 70 \\ 1 & \text{für } x > 70 \end{cases}$$

(c) Mittelwert (näherungsweise): 45.5, relative Abweichung vom exakten Wert: -0.004093 bzw. -0.4093%

(d) Mit emp. Verteilungsfunktion genäherte Anzahl: 31.5

(e) Unteres Quartil: $33.\overline{3}$

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

Eine Urne enthält 30 gleichartige Kugeln, von denen 3 lila und gepunktet, 6 orange und gepunktet, 12 lila und ungemustert sowie 9 orange und ungemustert sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

(a) eine rein zufällig aus der Urne entnommene Kugel orange und gepunktet ist?

(b) eine rein zufällig aus der Urne entnommene Kugel lila ist?

(c) eine rein zufällig aus der Urne entnommene Kugel ungemustert ist, wenn man weiß, dass sie lila ist?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $\frac{1}{5}$

(b) $\frac{1}{2}$

(c) $\frac{4}{5}$

Aufgabe 6 (5 + 2 + 1 = 8 Punkte)

Bei einem Chip-Hersteller verteilt sich die Produktion eines bestimmten Mikrocontrollers auf insgesamt drei verschiedene Produktionslinien A, B und C. Dabei werden im Mittel 15% der Chips auf Linie A, 40% der Chips auf Linie B und 45% der Chips auf Linie C hergestellt. Aus den Ergebnissen der Qualitätssicherung ist bekannt, dass 98% der Chips aus Linie A, 98.5% der Chips aus Linie B und 99% der Chips aus Linie C nicht fehlerhaft sind.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Chip nicht fehlerhaft ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein fehlerhafter Chip auf der Linie C produziert wurde?

(c) Sind die Ereignisse „Chip ist fehlerhaft“ und „Chip wurde auf Linie C produziert“ stochastisch unabhängig?

Ergebnisse (ohne Begründung/Rechenweg):

(a) 0.9865

(b) 0.3333

(c) Nein.

Aufgabe 7 (5 + 2 + 6 + 4 = 17 Punkte)

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

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

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < 0\})$ und $P\left(\left\{0 \leq X \leq \frac{3}{4}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie das obere Quartil von X.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Verteilungsfunktion von X:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -0.5 \\ \frac{1}{2}x^2 + \frac{1}{2}x + \frac{1}{8} & \text{für } -\frac{1}{2} < x \leq \frac{1}{2} \\ -2x^2 + 4x - 1 & \text{für } \frac{1}{2} < x \leq 1 \\ 1 & \text{für } x > 1 \end{cases}$$

(b) $P(\{X < 0\}) = \frac{1}{8}, P\left(\left\{0 \leq X \leq \frac{3}{4}\right\}\right) = \frac{3}{4}$

(c) $\mathrm{E}(X) = \frac{5}{12}$

(d) $x_{0.75} = 0.6464$

Aufgabe 8 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Chemie waren die Atommassen von 12 Elementen des Periodensystems auswendig zu lernen. Der Schüler Carl Clever hat 9 dieser Atommassen auswendig gelernt (die Chance, bei den anderen 3 Atommassen durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Carl die Hausaufgabe ordentlich erledigt hat, indem er 3 Mal rein zufällig und unabhängig voneinander eines der Elemente auswählt und die zugehörigen Atommassen abfragt. Kann Carl mindestens zu 2 dieser 3 Elemente die Atommassen korrekt angeben, so ist die Überprüfung bestanden.

(a) Welche Verteilung besitzt die Anzahl der von Carl abgegebenen richtigen Antworten?

(b) Mit welcher Wahrscheinlichkeit besteht Carl die Überprüfung?

Ergebnisse (ohne Begründung/Rechenweg):

(a) $B(3, 0.75)$

(b) 0.84375

Aufgabe 9 (2 + 9 + 1 + 3 = 15 Punkte)

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

$X \setminus Y$	-1	0	2	$p_{i\cdot}$
2	0.07	0.17	0.06	
4	0.13	0.28	0.09	
6	0.1	0.05	0.05	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}(4X - 5Y)$ sowie $\mathrm{Var}(4X - 5Y)$.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Ergänzte Tabelle:

$X \setminus Y$	-1	0	2	$p_{i\cdot}$
2	0.07	0.17	0.06	0.3
4	0.13	0.28	0.09	0.5
6	0.1	0.05	0.05	0.2
$p_{\cdot j}$	0.3	0.5	0.2	1

(b) Es gilt: $\mathrm{E}(X) = 3.8$, $\mathrm{E}(Y) = 0.1$, $\mathrm{Var}(X) = 1.96$, $\mathrm{Var}(Y) = 1.09$, $\mathrm{Cov}(X, Y) = -0.08$, $\mathrm{Korr}(X, Y) = -0.05473$

(c) X und Y sind **nicht** stochastisch unabhängig.

(d) $\mathrm{E}(4 \cdot X - 5 \cdot Y) = 14.7$, $\mathrm{Var}(4 \cdot X - 5 \cdot Y) = 61.81$

Aufgabe 10 (2 + 4 + 3 = 9 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{336}$ seien unabhängig identisch $B(1, 0.7)$-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{336} X_i = X_1 + \ldots + X_{336}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 220 und 240 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.9-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 13!

Ergebnisse (ohne Begründung/Rechenweg):

(a) $Y \sim B(336, 0.7)$, $\mathrm{E}(Y) = 235.2$, $\mathrm{Var}(Y) = 70.56$.

(b) $P\{220 \leq Y \leq 240\} \approx 0.6806$

(c) $y_{0.9} \approx 245.952$

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