Aufgabenstellung und Ergebnisse zur

Bachelor-Prüfung
Deskriptive Statistik und Wahrscheinlichkeitsrechnung
Sommersemester 2017

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 (= 16 + 12 + 15 + 18 + 6 + 10 + 17 + 18 + 8) 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						
Σ						

Aufgabe 1 (16 Punkte)

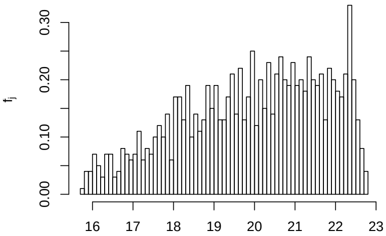
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

		wahr	falsch
1.	In einem Histogramm sind die Höhen der Rechtecke stets proportional zu den relativen Häufigkeiten der zugehörigen Klassen.	☐	■
2.	Ist x_p ein p-Quantil eines kardinalskalierten Merkmals X (mit $0 < p < 1$), dann ist mindestens ein Anteil von $1 - p$ der Urlisteneinträge zum Merkmal X größer als oder gleich x_p.	■	☐
3.	In einer bestimmten Prüfung waren 65% aller durchgefallenen Prüflinge männlich, 35% weiblich. Daraus kann man schließen, dass der Anteil der durchgefallenen unter den männlichen Prüflingen größer ist als unter den weiblichen Prüflingen.	☐	■
4.	Es seien A, B und C drei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(A) = 0.3$, $P(B) = 0.2$ und $P(C) = 0.4$. Damit gilt $P(A \cup B \cup C) \geq 0.9$.	☐	■
5.	Würfelt man fünf Mal mit einem fairen Würfel, so ist die Anzahl der Einsen geometrisch verteilt mit Parameter $p = 5 \cdot \frac{1}{6} = \frac{5}{6}$.	☐	■
6.	Für die stetige Zufallsvariable X mit Verteilungsfunktion F_X gelte $F_X(x) = 1 - F_X(-x)$ für alle $x \in \mathbb{R}$. Damit ist X symmetrisch um Null verteilt.	■	☐
7.	Die beiden Zufallsvariablen X und Y seien unabhängig und jeweils Poisson-verteilt mit Parameter $\lambda = 10$. Dann ist auch $X+Y$ Poisson-verteilt mit Parameter 10.	☐	■
8.	Sind X und Y zwei Zufallsvariablen mit positiven Varianzen, so ist der Korrelationskoeffizient von $2 \cdot X$ und $2 \cdot Y$ doppelt so groß wie der Korrelationskoeffizient von X und Y.	☐	■

Aufgabe 2 (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 das folgende Histogramm *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 Ränge $\mathrm{rg}(X)_1, \ldots, \mathrm{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$2, \quad 2, \quad 5, \quad 7, \quad 8, \quad 8, \quad 8, \quad 11$$

des ordinalskalierten Merkmals X lauten:

(a) 1.5, 1.5, 5, 7, 8.5, 8.5, 8.5, 11 ☐

(b) 1.5, 1.5, 2, 3, 4.5, 4.5, 4.5, 6 ☐

(c) 1.5, 1.5, 3, 4, 6, 6, 6, 8 ■

(d) 1.5, 1.5, 3, 4, 6.5, 6.5, 6.5, 8 ☐

3. Es seien X und Y zwei Zufallsvariablen über dem Wahrscheinlichkeitraum $(\Omega, \mathcal{F}, P)$ mit existierenden positiven Varianzen $\mathrm{Var}(X)$ und $\mathrm{Var}(Y)$ sowie existierender Kovarianz $\mathrm{Cov}(X, Y)$. Dann gilt für die Unkorreliertheit von X und Y sowie die stochastische Unabhängigkeit von X und Y:

(a) Es gibt keinen Zusammenhang zwischen Unkorreliertheit und Unabhängigkeit. ☐

(b) Aus der stochastischen Unabhängigkeit folgt die Unkorreliertheit. ■

(c) Aus der Unkorreliertheit folgt die stochastische Unabhängigkeit. ☐

(d) Unkorreliertheit und stochastische Unabhängigkeit sind äquivalent. ☐

4. Sind X_1, X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(30, 5^2)$, $X_2 \sim N(50, 10^2)$ und $X_3 \sim N(70, 10^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

(a) $N(50, 25^2)$-Verteilung. ☐

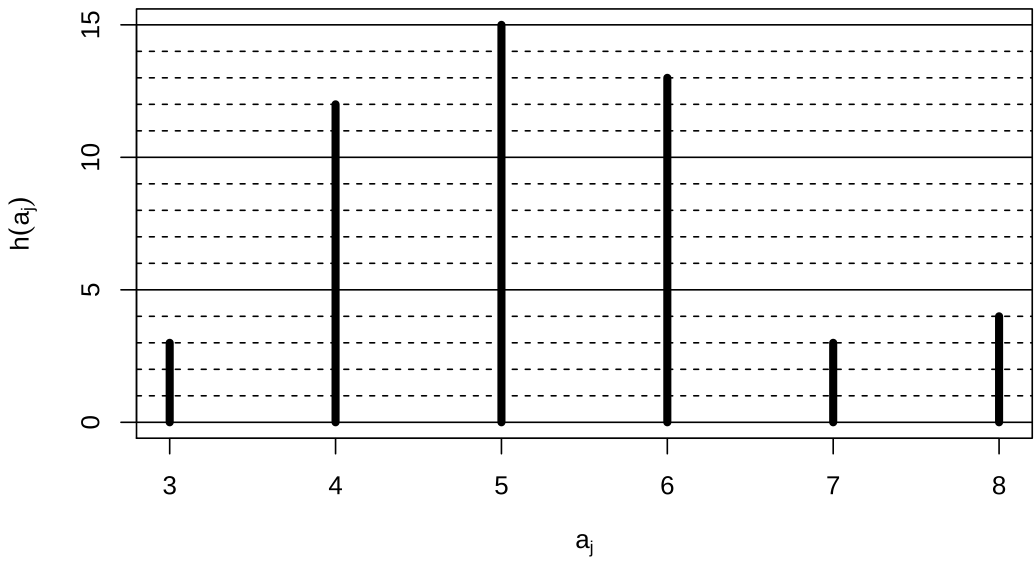
(b) $N(150, 25^2)$-Verteilung. ☐

(c) $N(50, 15^2)$-Verteilung. ☐

(d) $N(150, 15^2)$-Verteilung. ■

Aufgabe 3 (4 + 5 + 3 + 1 + 2 = 15 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



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

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

(c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(d) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mehr als 5 annehmen?

(e) Berechnen Sie ein unteres Quartil und ein oberes Quartil des Merkmals X.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Häufigkeitstabelle (mit absoluten und relativen Häufigkeiten):

a_j	3	4	5	6	7	8	Σ
$h(a_j)$	3	12	15	13	3	4	50
$r(a_j)$	0.06	0.24	0.30	0.26	0.06	0.08	1.00

(b) $\overline{x} = 5.26$, $s^2 = 1.6324$

(c) Empirische Verteilungsfunktion:

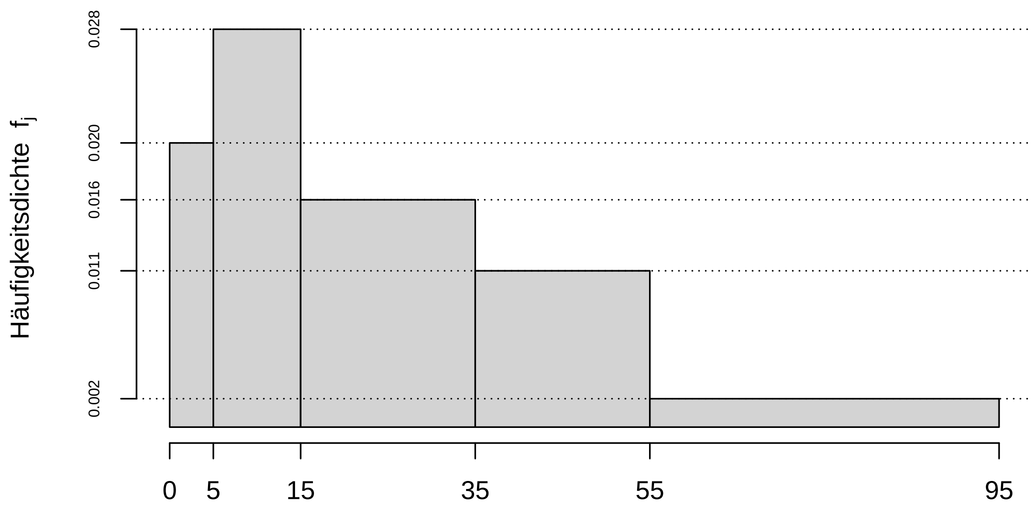
$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 3 \\ 0.06 & \text{für} \quad 3 \leq x < 4 \\ 0.30 & \text{für} \quad 4 \leq x < 5 \\ 0.60 & \text{für} \quad 5 \leq x < 6 \\ 0.86 & \text{für} \quad 6 \leq x < 7 \\ 0.92 & \text{für} \quad 7 \leq x < 8 \\ 1.00 & \text{für} \quad x \geq 8 \end{cases}$$

(d) Gesuchter Anteil: $0.4 = 40\%$

(e) $x_{0.25} = 4, x_{0.75} = 6$

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

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 50$:



(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 27.638?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 20 und 70 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 den Median.

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	$(0, 5]$	5	2.5	5	0.10	0.020	0.10
2	$(5, 15]$	10	10.0	14	0.28	0.028	0.38
3	$(15, 35]$	20	25.0	16	0.32	0.016	0.70
4	$(35, 55]$	20	45.0	11	0.22	0.011	0.92
5	$(55, 95]$	40	75.0	4	0.08	0.002	1.00

(b) (Approximative) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für} \quad x \leq 0 \\ 0.02 \cdot (x - 0) & \text{für} \quad 0 < x \leq 5 \\ 0.1 + 0.028 \cdot (x - 5) & \text{für} \quad 5 < x \leq 15 \\ 0.38 + 0.016 \cdot (x - 15) & \text{für} \quad 15 < x \leq 35 \\ 0.7 + 0.011 \cdot (x - 35) & \text{für} \quad 35 < x \leq 55 \\ 0.92 + 0.002 \cdot (x - 55) & \text{für} \quad 55 < x \leq 95 \\ 1 & \text{für} \quad x > 95 \end{cases}$$

(c) Mittelwert (näherungsweise): 26.95, relative Abweichung vom exakten Wert: -0.02489 bzw. -2.489%

(d) Mit emp. Verteilungsfunktion genäherte Anzahl: 24.5

(e) Median: 22.5

Aufgabe 5 (4 + 2 = 6 Punkte)

Der 23-köpfige Kader einer Fußballmannschaft bestehe aus drei Torwarten, acht Abwehrspielern, sieben Mittelfeldspielern und fünf Stürmern.

(a) Wie viele Möglichkeiten hat der Trainer, die Startelf – bestehend aus einem Torwart, vier Abwehrspielern, vier Mittelfeldspielern und zwei Stürmern – zusammenzustellen?

(b) Nehmen Sie nun an, dass der Trainer die Zusammenstellung gemäß Teil (a) rein zufällig bestimmt, und dass zwei der fünf Stürmer bereits mit einer gelben Karte vorbelastet sind. Wie groß ist die Wahrscheinlichkeit, dass genau diese beiden Stürmer bei der zufälligen Zusammenstellung der Startelf ausgewählt werden?

Hinweis: Sie können diesen Aufgabenteil auch ohne die Bearbeitung von Aufgabenteil (a) lösen!

Ergebnisse (ohne Begründung/Rechenweg):

(a) 73500

(b) 0.1

Aufgabe 6 (6 + 2 + 2 = 10 Punkte)

Ein Hersteller von Tiefkühlfertiggerichten bezieht seine Frischfischlieferungen von den vier Großhändlern A, B, C und D. Dabei werden einzelne Lieferungen mit einer Wahrscheinlichkeit von 20% von Lieferant A, 50% von Lieferant B, 15% von Lieferant C und 15% von Lieferant D geliefert. Bei den anschließenden Qualitätskontrollen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 2.5% bei Lieferant A, 2% bei Lieferant B, 7% bei Lieferant C und 3% bei Lieferant D Anlass zu Beanstandungen.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Lieferung in der Qualitätskontrolle beanstandet wird?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine von der Qualitätskontrolle nicht beanstandete Lieferung von Großhändler D geliefert wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Lieferung stammt von Großhändler D“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Ergebnisse (ohne Begründung/Rechenweg):

(a) 0.03

(b) 0.15

(c) Ja.

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

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

$$f_X(x) = \begin{cases} \frac{1}{8}x + \frac{1}{8} & \text{für } -1 \leq x < 1 \\ -\frac{1}{8}x + \frac{5}{8} & \text{für } 1 \leq x \leq 3 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X > 0\})$ und $P(\{1 \leq X \leq 2\})$.

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

(d) Bestimmen Sie den Median von X.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Verteilungsfunktion von X:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -1 \\ \frac{1}{16}x^2 + \frac{1}{8}x + \frac{1}{16} & \text{für } -1 < x \leq 1 \\ -\frac{1}{16}x^2 + \frac{5}{8}x - \frac{5}{16} & \text{für } 1 < x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$$

(b) $P(\{X > 0\}) = \frac{15}{16}, P(\{1 \leq X \leq 2\}) = \frac{7}{16}$

(c) $\mathrm{E}(X) = \frac{3}{2}$

(d) $x_{0.50} = 1.536$

Aufgabe 8 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

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

$X \setminus Y$	1	3	5	$p_{i\cdot}$
-1	0.05	0.08	0.12	
1	0.1	0.22	0.13	
3	0.2	0.05	0.05	
$p_{\cdot j}$				

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

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

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

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(3X - 2Y)$ sowie $\mathrm{Var}(3X - 2Y)$.

Ergebnisse (ohne Begründung/Rechenweg):

(a) Ergänzte Tabelle:

$X \setminus Y$	1	3	5	$p_{i\cdot}$
-1	0.05	0.08	0.12	0.25
1	0.1	0.22	0.13	0.45
3	0.2	0.05	0.05	0.3
$p_{\cdot j}$	0.35	0.35	0.3	1

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

x_i	-1	1	3
$p_{X\|Y=1}(x_i)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$
$p_{X\|Y=3}(x_i)$	$\frac{8}{35}$	$\frac{22}{35}$	$\frac{1}{7}$
$p_{X\|Y=5}(x_i)$	$\frac{2}{5}$	$\frac{13}{30}$	$\frac{1}{6}$

(c) Es gilt: $\mathrm{E}(X) = 1.1$, $\mathrm{E}(Y) = 2.9$, $\mathrm{Var}(X) = 2.19$, $\mathrm{Var}(Y) = 2.59$, $\mathrm{Cov}(X, Y) = -0.87$, $\mathrm{Korr}(X, Y) = -0.3653$

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

(e) $\mathrm{E}\left(3 \cdot X - 2 \cdot Y\right) = -\frac{5}{2}$, $\mathrm{Var}\left(3 \cdot X - 2 \cdot Y\right) = 40.51$

<u>Aufgabe 9</u> (2 + 2 + 4 = 8 Punkte)

Ein freiberuflicher Netzwerktechniker benötigt für das Auflegen eines Netzwerkanschlusses im Mittel 12 Minuten bei einer Standardabweichung von 3 Minuten. Man kann davon ausgehen, dass die benötigten Zeitdauern für die einzelnen Anschlüsse nicht gegenseitig voneinander abhängen. In einem bestimmten Schaltschrank sind insgesamt 36 Anschlüsse aufzulegen.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Arbeitszeit (für alle 36 Netzwerkanschlüsse) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Netzwerktechniker nicht länger als 7.5 Stunden bzw. 450 Minuten zum Auflegen aller Anschlüsse benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die gesamte Arbeitszeit für 36 Netzwerkanschlüsse mit einer Wahrscheinlichkeit von 0.95 realisiert.

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

Ergebnisse (ohne Begründung/Rechenweg):

(a) Erwartungswert: 432, Standardabweichung: 18

(b) Gesuchte (genäherte) Wahrscheinlichkeit: 84.13%

(c) $[396.72, 467.28]$

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