

2. Prüfung

Allgemeine Regeln: Die Prüfung dauert 100 Minuten. Die Benutzung eines (nicht graphischen) Taschenrechners ist erlaubt. Numerische Ergebnisse müssen auf 4 Dezimalstellen gerundet werden. Zum Erreichen der maximalen Punktzahl ist es notwendig, detaillierte Lösungen zu geben und die angewandten Behauptungen, Sätze und Eigenschaften aufzulisten. Es ist verboten, während der ersten 30 Minuten den Raum zu verlassen.

1. Geben wir die folgenden Definitionen an:

- (a) Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum.
 - i. Seien $A, B \in \mathcal{F}$ Ereignisse. Was bedeutet nach Definition, dass A und B *unabhängig* sind?
 - ii. Seien X und Y auf Ω definierte Zufallsvariablen. Was bedeutet nach Definition, dass X und Y *unabhängig* sind? (Zur Antwort dieser Frage müssen die im Teil i. definierten Begriffe nicht wieder definiert werden.)
- (b) Sei $X_1, X_2, \dots, X_n, \dots$ eine Folge von Zufallsvariablen und sei noch Z eine Zufallsvariable. Was bedeutet nach Definition, dass die obige Folge gegen Z *in Verteilung konvergiert*?

2. Die gemeinsame Verteilung der Zufallsvariablen X und Y ist in der folgenden Tabelle angegeben.

$\begin{array}{c} X \\ \backslash \\ Y \end{array}$	1	2
0	1/3	1/6
1	1/12	1/4
2	1/12	1/12

Berechnen wir die Kovarianz und den Korrelationskoeffizienten von X und Y , und bestimmen wir die Regressionsgerade von Y auf X .

3. Die Verteilungsfunktion einer absolut stetigen Zufallsvariablen ist durch

$$F_X: \mathbb{R} \rightarrow \mathbb{R}, \quad F_X(t) = \begin{cases} 0, & \text{falls } t \leq 1, \\ \ln t, & \text{falls } 1 < t < e, \\ 1, & \text{falls } t \geq e \end{cases}$$

angegeben, wo $e \approx 2,71 \dots$ die Basis des natürlichen Logarithmus ist.

- (a) Bestimmen wir den Erwartungswert und die Varianz von X .
- (b) 100 (gemeinsam) unabhängige und identisch verteilte Zufallsvariablen mit der obigen Verteilungsfunktion werden durch einen Computer simuliert. Schätzen wir durch eine geeignete Approximation die Wahrscheinlichkeit, dass die Summe dieser 100 Variablen höchstens 180 ist. (Die Bestimmung des genauen Wertes der Wahrscheinlichkeit wird in dieser Aufgabe nicht bewertet.)

Bemerkung: Statt symbolischer Werte können die (gerundete) numerische Werte der in Teil (a) bestimmten Ergebnisse in Teil (b) benutzt werden.

4. Die Dichtefunktion des Zufallsvektors (X, Y) ist durch

$$f_{X,Y}: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_{X,Y}(s, t) = \begin{cases} c, & \text{falls } 0 < t < 1 \text{ und } 0 < s < t, \\ 0 & \text{sonst,} \end{cases}$$

angegeben, wo $c > 0$ eine geeignete reelle Zahl ist.

- (a) Bestimmen wir den Wert von c .
- (b) Geben wir die Dichtefunktionen f_X bzw. f_Y der Zufallsvariablen X bzw. Y an.
- (c) Sind X und Y unabhängig? Sind X und Y identisch verteilt?

5. In einer tropischen Höhle ist die Temperatur das ganze Jahr über fast konstant. Die Temperatur wurde in 6 verschiedenen Zeitpunkten gemessen, und die folgende Werte (in Grad Celsius angegeben) wurden erhalten:

$$x_1 = 18,6 \quad x_2 = 18,0 \quad x_3 = 19,2 \quad x_4 = 17,4 \quad x_5 = 16,8 \quad x_6 = 18,0.$$

Der Datenvektor $\mathbf{x} = (x_1, \dots, x_6)$ wird als Realisierung einer einfachen Stichprobe $\mathbf{X} = (X_1, \dots, X_6)$ betrachtet, deren Stichprobenvariablen die Verteilung $N(\mu; 4)$ haben, wo der Erwartungswert μ unbekannt ist.

- (a) Geben wir ein Konfidenzintervall zum Konfidenzniveau 97% für den Parameter μ aufgrund des obigen Datenvektors an.
- (b) In einem alten Geographielehrbuch steht die Behauptung $\mu = 17,5$ °C. Es wurde vermutet, dass diese Behauptung wegen des Klimawandels nicht mehr haltbar sei. Konstruieren wir einen Hypothesentest zum Fehlniveau 5%, um zu entscheiden, ob die Behauptung gilt. Stellen wir eine der Situation entsprechende Nullhypothese H_0 und die zugehörige Alternativhypothese H_1 auf, berechnen wir den Wert der Teststatistik. Entscheiden wir, ob H_0 angenommen oder verworfen werden muss, und begründen wir die Entscheidung.
- *6. Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein zweidimensionaler nicht singular normalverteilter Zufallsvektor mit Erwartungswertvektor $\begin{pmatrix} 42 \\ -1 \end{pmatrix}$. Nehmen wir an, dass $\mathbb{E}(Y|X) = X + d$ gilt, wo d eine geeignete reelle Konstante ist.
- (a) Berechnen wir den Wert von d .
- (b) Bestimmen wir die Regressionsgerade von Y auf X .
- (c) Sind X und Y unabhängig?
- (d) Ist die Kovarianzmatrix von $\begin{pmatrix} X \\ Y \end{pmatrix}$ eine Diagonalmatrix?

Bemerkung: Wenn der Wert von d im Teil (a) nicht bestimmt wird, so können die anderen Teilaufgaben auch parametrisch gelöst werden.

Verteilung	Notation	$\text{ran} X$	$F_X(t)$	$f_X(k), f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
Indikatorvariable	$\mathbb{1}(p)$	$\{0, 1\}$		$1 - p, p$	p	$p(1 - p)$
Binomialverteilung	$\text{Bin}(n; p)$	$\{0, 1, \dots, n\}$		$\binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$
Poisson-Verteilung	$\text{Poi}(\lambda)$	$\mathbb{N}$		$\frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ
geometrische Verteilung	$\text{Geo}(p)$	$\mathbb{N}^+$		$(1 - p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
uniforme Verteilung	$U(a; b)$	$(a; b)$	$\frac{t-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentialverteilung	$\text{Exp}(\lambda)$	$[0; \infty)$	$1 - e^{-\lambda t} \quad (0 \leq t)$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normalverteilung	$N(\mu; \sigma^2)$	$\mathbb{R}$	$\Phi\left(\frac{t-\mu}{\sigma}\right)$	$\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	μ	σ^2
n -dim. Normalverteilung	$N(\underline{\mu}; \Sigma)$	$\mathbb{R}^n$	$f_X(\underline{t}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \Sigma}} e^{-\frac{1}{2}(\underline{t}-\underline{\mu})^T \Sigma^{-1}(\underline{t}-\underline{\mu})}$		$\underline{\mu}$	Σ

Standardnormalverteilungstabelle

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

Formelsammlung

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}_n^2, (s_n^*)^2 = \frac{n}{n-1} s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2, s_n^* = \sqrt{(s_n^*)^2}.$$

Tests und Konfidenzintervalle

- Zweiseitiger u -Test für eine Stichprobe: $u = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$, $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.
- Einseitiger u -Test für eine Stichprobe: $u = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$, $u_{\varepsilon} = \Phi^{-1}(1 - \varepsilon)$.
- Zweiseitiger t -Test für eine Stichprobe: $t = \frac{\bar{x}_n - \mu_0}{s_n^*} \sqrt{n}$,
 $t_{n-1, \varepsilon/2}$ ist das $(1 - \varepsilon/2)$ -Quantil der $t(n-1)$ -Verteilung.
- Einseitiger t -Test für eine Stichprobe: $t = \frac{\bar{x}_n - \mu_0}{s_n^*} \sqrt{n}$,
 $t_{n-1, \varepsilon}$ ist das $(1 - \varepsilon)$ -Quantil der $t(n-1)$ -Verteilung.
- Konfidenzintervall für μ bei bekannter σ : $\left[\bar{x}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \bar{x}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \right]$.
- Konfidenzintervall für μ bei unbekannter σ : $\left[\bar{x}_n - \frac{s_n^* t_{n-1, 1-\varepsilon/2}}{\sqrt{n}}, \bar{x}_n + \frac{s_n^* t_{n-1, \varepsilon/2}}{\sqrt{n}} \right]$.