11. Übungsblatt

Eigenschaften von Summen von normalverteilten Zufallsvariablen, Tschebyscheff-Ungleichung.
Grundbegriffe der Statistik, empirische Verteilungsfunktion

1. Seien $X \sim \mathrm{N}(1;9)$ und $Y \sim \mathrm{N}(2;16)$ unabhängige, normalverteilte Zufallsvariablen.
   (a) Wie groß ist $\mathbb{P}(X > 0{,}5)$? *(Wiederholungsaufgabe, cf. Übungsblatt 7.)*
   (b) Bestimmen wir die Kovarianz und den Korrelationskoeffizient von $X + 2Y$ und $X - 4Y$.
   (c) Was für eine Verteilung hat die Zufallsvariable $2X+Y$ und mit welchen Parametern (und warum)?

   (Prüfung, 23.01.2025., Aufgabe 4)

2. Seien X_1, X_2, X_3 gemeinsam unabhängige, normalverteilte Zufallsvariablen, jeweils mit Erwartungswert μ und Varianz σ^2. Wie ist die Zufallsvariable $Y = \frac{X_1+X_2+X_3-3\mu}{\sqrt{3}\sigma}$ verteilt? Wie groß ist $\mathrm{corr}(X_1, Y)$?

3. Seien X und Y unabhängige standardnormalverteilte Zufallsvariablen, und seien noch $U = 3X-2Y+1$ und $V = 7X - 5Y - 3$.
   a) Bestimmen wir die Kovarianzmatrix des Zufallsvektors $(X, Y)^T$.
   b) Bestimmen wir den Erwartungswertvektor und die Kovarianzmatrix des Zufallsvektors $(U, V)^T$.
   c) Berechnen wir $\mathrm{cov}(-2U + V, 12U - 5V)$.
   d) Sind $-2U + V$ és $12U - 5V$ unabhängig?

4. Seien $X \sim \mathrm{N}(1;4)$ und $Y \sim \mathrm{N}(-2;1)$ unabhängige, normalverteilte Zufallsvariablen und sei $U := X+Y$ und $V := X - 2Y$.
   (a) Geben wir den Erwartungswertvektor (also $(\mathbb{E}(U), \mathbb{E}(V))^T$) und die Kovarianzmatrix des Zufallsvektors $(U, V)^T$ an.
   (b) Sind U und V unabhängig?
   (c) Was ist die Regressionsgerade von V auf U?
   (d*) Dementsprechend bestimmen wir die Regression $\mathbb{E}(U^2(2V + 1)|U)$.

   (Prüfung, 26.06.2025., Aufgaben 3c)+6. in modifizierter Form, für insgesamt 29 Punkte.)*

5. In der Vorlesung: Seien $X \sim \mathrm{N}(5;2)$ und $Y \sim \mathrm{N}(4;3)$ unabhängige Zufallsvariablen. Geben wir die Wahrscheinlichkeit $\mathbb{P}(X < Y)$ an.

Auch nur in der Vorlesung (aber die Tschebyscheff-Ungleichung ist Teil des Prüfungsstoffes):

6. Ein gegebener Datenbankserver erhält im Durchschnitt 50 Abrufe innerhalb einer Zeiteinheit. Die Standardabweichung der Anzahl der Abrufe ist nach Erfahrung gleich 5. Geben wir eine untere Schranke für die Wahrscheinlichkeit an, dass die Anzahl der Abrufe innerhalb einer Zeiteinheit mehr als 40, aber weniger als 60 sein wird.

7. Wir werfen 10 reguläre Würfel, sei X die Summe der Augenzahlen. Wir haben früher mittels des zentralen Grenzwertsatzes die Wahrscheinlichkeiten $\mathbb{P}(24 < X < 46)$ und $\mathbb{P}(31 \leq X \leq 39)$ abgeschätzt. Welche Abschätzungen können wir für dieselben Wahrscheinlichkeiten mit Hilfe der Tschebyscheff-Ungleichung bekommen?

---

8. In einer mündlichen Prüfung prüfte ein Prüfer 13 Studierende an einem Tag und vergab die folgende Noten (in dieser Reihenfolge): $3, 2, 4, 5, 2, 1, 3, 3, 4, 5, 4, 1, 2$. Betrachten wir die obigen Daten als die Realisierung einer einfachen Stichprobe $\mathbf{X} = (X_1, \ldots, X_{13})$. Berechnen wir das empirische Mittel, die korrigierte empirische Standardabweichung, den Stichprobenmedian, den Modus und den der Variablen X_8^* entsprechenden Wert für diese Realisierung.

9. Eine Mine hat 8 Mitarbeiter(innen): Schneewittchen (die Direktorin der Mine), Chef (der Schichtleiter), ferner Happy, Brummbär, Hatschi, Pimpel, Schlafmütz und Seppel (die Bergmänner). `Bitte wenden!`

a) Das Monatsgehalt der Direktorin ist 2000 Taler, der Schichtleiter verdient 300 Taler pro Monat und jeder Bergmann verdient 200 Taler monatlich. Bestimmen wir das Mittel, den Median und den Modus der Gehälter der Arbeitnehmer der Mine.

b) In seiner Freizeit beschäftigt sich Happy auch mit offenen Problemen der Informatik. Nehmen wir an, dass er das P-NP-Problem löst, für das er 1 Million Dollar erhält, und nach Umwechslung erhöht sich sein Gehalt für diesen Monat auf genau 5 Millionen Taler. Wie ändern sich das Mittel, der Median und der Modus der Gehälter in diesem Monat?

Was können wir anhand der Aufgabe über das Verhalten des Mittels und des Medians sagen?

---

10. In der Vorlesung: Ein regulärer Würfel wird n Mal geworfen.

    (a) Im Fall $n = 7$ und $x = 3$, welche Verteilung hat $nF_n^*(x)$ (also der Wert der empirischen Verteilungsfunktion an der Stelle $x = 3$ multipliziert mit der Stichprobengröße) und mit welchen Parametern? $F_6^*(1)$=?

    (b) Wir erhalten die Ergebnisse $1, 3, 4, 5, 6, 1, 4$. Skizzieren wir den Graphen der dieser Realisierung gehörigen empirischen Verteilungsfunktion F_7^* dar. $F_7^*(1) =?$ $F_7^*(6) =?$

    (c) Sei nun $(X_1, \ldots, X_n)$ der Stichprobenvektor, die das Ergebniss der n-maligen Würfelns angibt. Bestimmen wir die Werte $\lim_{n\to\infty} F_n^*(1)$, $\lim_{n\to\infty} F_n^*(3{,}5)$, und $\lim_{n\to\infty} F_n^*(6)$ (genauer gesagt, die Werte, mit denen diese Grenzwerte mit Wahrscheinlichkeit 1 übereinstimmen).

11. Für $n \geq 1$ seien $X_1, \ldots, X_n$ unabhängige und identisch verteilte stetige Zufallsvariablen, deren Dichte wie folgt gegeben ist:

$$f_{X_1}(x) = \begin{cases} 3x^2, & \text{ha } 0 < x < 1, \\ 0, & \text{sonst.} \end{cases}$$

    (a) Im Fall $n = 8$, wie ist die Zufallsvariable $8F_8^*(0{,}5)$ verteilt und mit welchen Parametern? Wie groß ist $\mathbb{D}(F_8^*(0{,}5))$?

    (b) Nach der Auswahl der Stichprobe (für $n = 8$) haben wir die folgende Realisation erhalten:

$$\Big(0{,}11, \quad 0{,}42, \quad 0{,}55, \quad 0{,}76, \quad 0{,}29, \quad 0{,}38, \quad 0{,}24, \quad 0{,}95\Big).$$

    Skizzieren wir den Graphen der empirischen Verteilungsfunktion F_8^*. Wie groß sind $F_8^*(0{,}11)$ és $F_8^*(0{,}95)$? Bestimmen wir den Stichprobenmedian und die korrigierte empirische Varianz.

    (c) Wie groß ist $\lim_{n\to\infty} F_n^*(0{,}5)$ (mit Wahrscheinlichkeit 1)?

12. Wir kaufen n billige alte Tastaturen auf dem Flohmarkt. Die Anzahlen der klebrigen Tasten auf den einzelnen Tastaturen können mit guter Annäherung durch n unabhängigen Poisson-verteilten Zufallsvariablen beschrieben werden mit Parameter $\lambda = 0{,}4$. Für $1 \leq i \leq n$ bezeichne X_i die Anzahl der klebrigen Tasten auf der i-ten Tastatur, und bezeichne F_n^* die zur Stichprobe $(X_1, \ldots, X_n)$ gehörige empirische Verteilungsfunktion. Wie groß ist $\lim_{n\to\infty} F_n^*(3)$ (mit Wahrscheinlichkeit 1)?

13. * Ein Konzertort im Freien hat die Kapazität von 12 000 Personen. Laut Beobachtungen wird an jedem Konzerttag mit Wahrscheinlichkeit 30% unabhängig von den anderen Konzerttagen mindestens 90% dieser Kapazität ausgenutzt. In einem Jahr gibt es 85 Konzerttage; an jedem solchen Tag schreiben wir auf, wie viel Prozent der Kapazität ausgenutzt wurde, bezeichne F_{85}^* die dazugehörige empirische Vereilungsfunktion. Wie ist $85F_{85}^*(0{,}9)$ verteilt? Schätzen wir die Wahrscheinlichkeit $\mathbb{P}(F_{85}^*(0{,}9) > 0{,}75)$ mittels geeigneter Approximation ab.

14. * Für $n \geq 1$ sei $X_1, \ldots, X_n$ eine einfache Stichprobe, wobei für jedes $1 \leq i \leq n$, $X_i \sim \mathrm{U}(0; 1)$. Für großes n, schätzen wir mittels geeigneter Approximation die Wahrscheinlichkeit $\mathbb{P}(F_n^*(0{,}3/n) > 1/n)$ ab, wobei F_n^* die entsprechende empirische Verteilungsfunktion bezeichnet.

15. Zeigen wir, dass die emprische Verteilungsfunktion F_n^*, die zu einer Realisierung einer beliebigen Stichproben vom Umfang n gehört, für alle $n \in \mathbb{N}^+$ wirklich eine Verteilungsfunktion ist. Ist es möglich, dass eine Zufallsvariable mit Verteilungsfunktion F_n^* auch eine Dichtefunktion hat?

Formeln

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

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

Name der Verteilung	Notation	Ran(X)	$F_X(t)$	$p_X(k)$, $f_X(t)$	$\mathbb{E}(X)$	$\mathbb{D}^2(X)$
Indikatorvariable Bernoulli-Verteilung	$\mathbb{1}_A$ $\mathrm{B}(p)$	$\{0,1\}$		$p_X(0)=1-p$, $p_X(1)=p$ $(p=\mathbb{P}(A))$	p	$p(1-p)$
Binomialverteilung	$\mathrm{Bin}(n;p)$	$\{0,1,\ldots,n\}$		$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Poisson-Verteilung	$\mathrm{Pois}(\lambda)$	$\{0,1,2,\ldots\}$		$\frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ
geometrische Verteilung	$\mathrm{Geo}(p)$	$\{1,2,\ldots\}$		$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
uniforme Verteilung	$\mathrm{U}(a;b)$	$(a;b)$	$\frac{t-a}{b-a}$ (falls $t\in(a;b)$)	$\frac{1}{b-a}$ (falls $t\in(a;b)$)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentialverteilung	$\mathrm{Exp}(\lambda)$	$[0;\infty)$	$1-e^{-\lambda t}$ (falls $t\in(0;\infty)$)	$\lambda e^{-\lambda t}$ (falls $t\in(0;\infty)$)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normalverteilung	$\mathrm{N}(\mu;\sigma^2)$	$\mathbb{R}$	$\Phi\left(\frac{t-\mu}{\sigma}\right)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	μ	σ^2