



Karlsruher Institut für Technologie

Institut für Stochastik

Prof. Dr. Ursula Müller-Harknett

Celeste Mayer, M.Sc.

**Lösungsvorschlag Klausur zur Vorlesung  
Wahrscheinlichkeitstheorie und Statistik  
für die Fachrichtungen Maschinenbau/ Informatik**

Verteilungsfunktion  $\Phi(x)$  der Standard – Normalverteilung  $\mathcal{N}(0, 1)$

| $x$ | 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 | 0.7703 | 0.7734 | 0.7764 | 0.7793 | 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 | 0.9996 | 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 |
| 3.5 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 |
| 3.6 | 0.9998 | 0.9998 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.7 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.8 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.9 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

**Aufgabe 1** (10 Punkte)

Ein Ornithologe möchte herausfinden, ob ein linearer Zusammenhang zwischen Schnabellänge  $x$  und Flügelspannweite  $y$  bei Rabenvögeln besteht. Dafür misst er bei 9 unterschiedlichen Vögeln die beiden Merkmale (in cm) und notiert sie handschriftlich in einer Tabelle. Dummerweise, ist seine Schrift an 2 Stellen unleserlich, sodass nicht alle Ziffern gelesen werden können. Sie erhalten folgende Tabelle:

| $j$   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8    | 9  |
|-------|-----|-----|-----|-----|-----|-----|-----|------|----|
| $x_j$ | 5.6 | 5.5 | 6.2 | 6.5 | 6.4 | 6.5 | 5.8 | 1x.x | 10 |
| $y_j$ | 85  | 88  | 99  | 111 | 108 | 101 | 98  | 140  | x  |

- a) Berechnen Sie den Median und das 0.2-getrimmte Stichproben-Mittel von  $(x_1, \dots, x_9)$ . (2P)

Median =  getrimmtes Mittel =

- b) Geben Sie die kleinste untere Schranke für das arithmetische Mittel der Schnabellänge an, d.h. die kleinste reelle Zahl  $x_u$ , die  $x_u \leq \bar{x}$  erfüllt. (1P)

$x_u =$

- c) Nehmen Sie nun an, dass die mittlere Schnabellänge  $\bar{x} = 7.4$  ist und das Stichproben-0.1-Quantil der Flügelspannweite bei  $\tilde{y}_{0.1} = 83$  liegt. Wie sehen in diesem Fall die fehlenden Werte aus? (2P)

$x_8 =$    $y_9 =$

Es wird nun bekannt, dass es sich bei Vogel 8 und 9 nicht um Rabenvögel handelte, weshalb die zugehörigen Daten aus dem Datensatz entfernt werden. Gehen Sie ab jetzt also von diesem reduzierten Datensatz aus.

Im Folgenden können Sie die Ergebnisse  $\bar{x} = 6.0714, \bar{y} = 98.5714, s_x^2 = 0.1857, s_y^2 = 90.9524$  und  $\frac{1}{6} \sum_{j=1}^7 (x_j - \bar{x})(y_j - \bar{y}) = 3.6857$  verwenden.

- d) Berechnen Sie den Korrelationskoeffizienten der Merkmale  $x$  und  $y$ . Können Sie eine lineare Abhängigkeit vermuten? Begründen Sie Ihre Antwort.

$r_{xy} =$

Da  $r_{xy}$  nahe bei 1 ist kann lineare Abhängigkeit vermutet werden.

- e) Bestimmen Sie die Koeffizienten der Regressionsgerade  $f(x) = a^* + b^*x$  für den reduzierten Datensatz. (2P)

$$a^* =$$

-21.9283

$$b^* =$$

19.8471

- f) Bei einem weiteren Rabenvogel wird nun eine Schnabellänge von  $x_{10} = 6$  gemessen. Welche Flügelspannweite  $y_{10}$  erwarten Sie? (1P)

$$y_{10} =$$

97.1543

## Lösungsvorschlag:

- a) Offenbar ist  $x_8$  der größte Wert in der Beobachtungsreihe. Damit sind Median und oberes Quartil unabhängig von  $x_8$  und entsprechen  $x_{(5)} = 6.4$  und  $x_{(7)} = 6.5$ .
- b) Das kleinste arithmetische Mittel wird für  $x_8 = 10.0$  realisiert und ist dann  $\frac{1}{9} \sum_{j=1}^9 x_j = 6.9444 = \frac{47}{7}$ .
- c) Umformen von  $\frac{1}{9} \sum_{j=1}^9 x_j = 7.4$  liefert  $x_8 = 14.1$ . Da  $\tilde{y}_{0.1} = y_{(1)}$  folgt  $y_9 = 83$ .
- d) Folgt aus direktem Einsetzen der Daten in

$$r_{xy} = \frac{\frac{1}{6} \sum_{j=1}^7 (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{s_x^2} \sqrt{s_y^2}}.$$

Diesen Wert erhält man ebenfalls, wenn man direkt die Werte aus der Tabelle nutzt und die Werte in der Aufgabenstellung nicht verwendet.

- e) Einsetzen der gegebenen Werte in  $b^* = r_{xy} \frac{s_y}{s_x}$  und  $a^* = \bar{y} - b^* \bar{x}$  liefert obige Ergebnisse, wenn man mit dem gerundeten Wert für  $r_{xy}$  rechnet.

Rechnet man mit dem exakten Wert, so erhält man:

$$a^* = -21.9313 \quad b^* = 19.8476.$$

Rechnet man direkt mit den Werten in der Tabelle ohne die angegebenen Werte der Aufgabenstellung zu nutzen, so erhält man

$$a^* = -21.9234 \quad b^* = 19.8462.$$

- f) Einsetzen von  $x = 6$  in die in e) bestimmte Regressionsgerade liefert das gewünschte Ergebnis. In den drei Fällen aus vorheriger Teilaufgabe erhalten wir:
  - 1. 97.1543,
  - 2. 97.1543,
  - 3. 97.1538.

**Aufgabe 2** (9 Punkte)

Es sei  $X$  eine normalverteilte Zufallsvariable mit Erwartungswert 2 und Varianz 1. Weiter sei  $Y$  eine von  $X$  unabhängige Zufallsvariable mit  $Y \sim \mathcal{N}(1, 1)$ .

- a) Wie ist die Zufallsvariable  $Z := 2 - X + Y$  verteilt? (2P)

$$2 - X + Y \sim \boxed{\mathcal{N}(1, 2)}$$

- b) Berechnen Sie die Kovarianz von  $X$  und  $Z + X$ . (1P)

$$C(X, Z + X) = \boxed{0}$$

- c) Berechnen Sie den Erwartungswert  $\mathbb{E}[X(Z + X)]$ . (1P)

$$\mathbb{E}[X(Z + X)] = \boxed{6}$$

- d) Bestimmen Sie ein  $t \in \mathbb{R}$  so, dass  $\mathbb{P}(X - 2 \leq t) = 0.8812$  gilt. Es genügt, wenn Sie das Ergebnis auf 2 Nachkommastellen genau angeben. (2P)

$$t = \boxed{1.22}$$

- e) Geben Sie die Wahrscheinlichkeit dafür an, dass das Minimum von  $X$  und  $Y$  größer als 1 ist. (2P)

$$\mathbb{P}(\min\{X, Y\} \geq 1) = \boxed{0.4207}$$

- f) Schätzen Sie die folgende Wahrscheinlichkeit mit der Tschebyschev-Ungleichung bestmöglich ab: (2P)

$$\mathbb{P}(Y \in (-1, 3)) \geq \boxed{\frac{3}{4}}$$

### Lösungsvorschlag:

a) Wegen der Unabhängigkeit von  $X$  und  $Y$  gilt mit den Rechenregeln für die Normalverteilung

$$2 - X + Y \sim \mathcal{N}(2 - 2 + 1, (-1)^2 + 1^2) = \mathcal{N}(1, 2).$$

b) Mit der Unabhängigkeit von  $X$  und  $Y$  gilt mit den Rechenregeln für die Kovarianz

$$C(X, Z + X) = C(X, 2 - X + Y - X) = C(X, Y) = 0.$$

c) Da nach b)  $X$  und  $(Z + X)$  unabhängig sind gilt

$$\mathbb{E}[X(X + Z)] = \mathbb{E}[X][Z] = 2 \cdot 3 = 6.$$

d) Es ist

$$\begin{aligned}\mathbb{P}(X - 2 \leq t) &= \mathbb{P}(X \leq t + 2) = \mathbb{P}\left(\frac{X - \mathbb{E}[X]}{\sqrt{V(X)}} \leq t\right) \\ &= \Phi(t) \stackrel{!}{=} 0.8888.\end{aligned}$$

Wir erhalten  $0.8888 = \Phi(t)$ . Damit ist mit der Normalverteilungstabelle  $t = \Phi^{-1}(0.8888) = 1.22$ .

e) Es ist

$$\begin{aligned}\mathbb{P}(\min\{X, Y\} \geq 1) &= \mathbb{P}(X \geq 1 \text{ und } Y \geq 1) \\ &= \mathbb{P}(X \geq 1)\mathbb{P}(Y \geq 1) = \mathbb{P}(\mathcal{N}(0, 1) \geq -1)\mathbb{P}(\mathcal{N}(0, 1) \geq 0) \\ &= (1 - \Phi(-1))\Phi(0) = (0.8413)\frac{1}{2} = 0.4207.\end{aligned}$$

f) Es ist

$$\begin{aligned}\mathbb{P}(Y \in (-1, 3)) &= 1 - \mathbb{P}(Y \notin (-1, 3)) = 1 - \mathbb{P}(|Y - 1| \geq 2) \\ &\geq 1 - \frac{\mathbb{V}(Y)}{4} = \frac{3}{4}.\end{aligned}$$

### Aufgabe 3 (11 Punkte)

Alex und Judith spielen seit Jahren gemeinsam Tischtennis. Nun treten sie bei einem Wettkampf gegeneinander an. Als begeisterter Tischtennis-Zuschauer wollen Sie sich gründlich auf den Wettkampf vorbereiten. Sie finden eine Liste mit allen Satzständen vergangener Spiele von Judith und Alex. Dieser entnehmen Sie, dass Judith außerhalb von Wettkämpfen 182 von 250 Sätzen gewonnen und 68 Sätze verloren hat. Da all die Sätze unter den gleichen Bedingungen gespielt wurden, kann man davon ausgehen, dass Judith bei jedem Satz die gleiche Gewinnwahrscheinlichkeit hatte. Die Ergebnisse aller Sätze können zusätzlich als voneinander unabhängig betrachtet werden.

- a) Geben Sie ein approximatives Konfidenzintervall zum 90%-Niveau für die Gewinnwahrscheinlichkeit  $p$  eines Satzes von Judith an. Nutzen Sie dafür die Normalverteilungstabelle auf S.2. (2P)

$$\mathcal{C}(x) = \boxed{[0.6816, 0.7744]/[0.6817, 0.7743]/[0.6818, 0.7742]}$$

Bei diesem Wettkampf wird nun auf zwei Gewinnsätze gespielt, d.h. Judith und Alex spielen so viele Sätze bis entweder Judith zwei Sätze gewonnen hat oder Alex zwei Sätze gewonnen hat. Anhand neuer Erkenntnisse gehen Sie nun davon aus, dass Judith einen Satz mit einer Wahrscheinlichkeit von  $p = 0.7$  gegen Alex gewinnt. Allerdings ist Judith nicht in der Lage, unter Wettkampfbedingungen bei einem Satzverlust einen kühlen Kopf zu bewahren. Verliert Judith also einen Satz und damit ihre Nerven, so fällt ihre Gewinnwahrscheinlichkeit beim nächsten Satz auf  $p = 0.4$ . Gewinnt Sie jedoch im Anschluss an einen verlorenen Satz den nächsten Satz, so erholen sich ihre Nerven und ihre Gewinnwahrscheinlichkeit steigt wieder auf den ursprünglichen Wert von  $p = 0.7$ .

- b) Wie hoch ist die Wahrscheinlichkeit, dass Judith das Spiel gegen Alex gewinnt? (2P)

$$\mathbb{P}(\text{Judith gewinnt}) = \boxed{0.658}$$

Sie haben nun leider das komplette Match verpasst, erfahren aber durch Zufall, dass Judith siegte.

- c) Wie hoch ist die Wahrscheinlichkeit, dass Judith auch den ersten Satz gewonnen hat? (2P)

$$\mathbb{P}(\text{Judith gewinnt Satz 1} | \text{Judith gewinnt}) = \boxed{0.8723}$$

Um künftig nicht weitere Tischtennis-Matches komplett zu verpassen, machen Sie sich nun Gedanken über die Zeit, die das Spiel dauerte. Sie gehen davon aus, dass die Spielzeit  $T$  eine verschobene Exponentialverteilung hat, d.h.  $T \sim \text{Exp}(\lambda) + c$  (in Minuten), wobei Ihnen sowohl  $c$  als auch  $\lambda$  unbekannt sind. Sie wissen jedoch, dass jedes Spiel mindestens 5 Minuten gehen muss und es auch Spiele geben kann, die genau 5 Minuten dauern. Des Weiteren wissen Sie, dass ein Spiel im Durchschnitt 8 Minuten dauert.

d) Geben Sie  $\lambda$  und  $c$  an. (2P)

$$\lambda = \boxed{\frac{1}{3}} \quad c = \boxed{5}$$

e) Nehmen Sie an, dass  $\lambda = 1/2$  und  $c = 4$  gelten. Berechnen Sie die Wahrscheinlichkeit, dass ein Match länger als 5 Minuten dauert. (2P)

$$\mathbb{P}(T \geq 5) = \boxed{e^{-1/2} = 0.6065}$$

f) Nehmen Sie an, dass  $\lambda = 1/2$  und  $c = 4$  gelten. Berechnen Sie die Wahrscheinlichkeit, dass ein Match genau 5 Minuten dauert. (1P)

$$\mathbb{P}(T = 5) = \boxed{0}$$

### Lösungsvorschlag:

a) Wir modellieren die Anzahl an Gewinnen mithilfe der Indikator-Zufallsvariablen

$$X_j = \mathbf{1}\{\text{Gewinn von Judith im } j\text{-ten Testdurchlauf}\}, \quad j = 1, \dots, n.$$

Nach Voraussetzung sind die  $X_1, \dots, X_n$  uiv. mit einer  $\text{Bin}(1, p)$ -Verteilung für die unbekannte Gewinnwahrscheinlichkeit  $p$ . Daher gilt

$$S_n = \sum_{j=1}^n X_j \sim \text{Bin}(n, p).$$

Es liegt somit die Situation aus den Beispielen 18.6 und 18.7 vor. Daher verwenden wir die Größe  $T_n = \frac{S_n}{n}$ . Für die zugrundeliegende Stichprobe ist bekannt, dass genau  $x = 182$  Gewinne aufgetreten sind, d.h. es gilt

$$T_n(x) = \frac{x}{n} = \frac{182}{250}.$$

Mithilfe der Vorgehensweise aus den oben genannten Beispielen erhalten wir ein Konfidenzintervall der Form

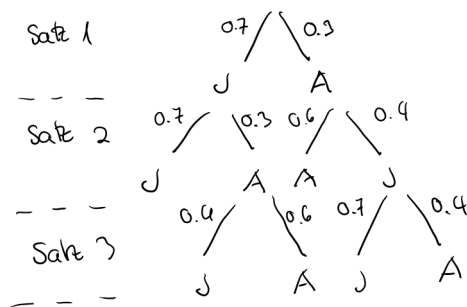
$$\mathcal{C}(x) = [l_n^*(x), L_n^*(x)]$$

für

$$l_n^*(x) = T_n(x) - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \sqrt{T_n(x)(1 - T_n(x))}, \quad L_n^*(x) = T_n(x) + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \sqrt{T_n(x)(1 - T_n(x))}.$$

Mit  $u_{1-\frac{\alpha}{2}}$  wird dabei das  $(1 - \frac{\alpha}{2})$ -Quantil der Standardnormalverteilung bezeichnet. In unserem Fall erhalten wir aus der Tabelle  $u_{0.95} = 1.65/1.645/1.64$ . Insgesamt erhalten wir daher das gesuchte Konfidenzintervall zum Konfidenzniveau 0.9 durch

$$\mathcal{C}(x) = [0.6816, 0.7744][0.6817, 0.7743]/[0.6818, 0.7742].$$



- b) Jeder Pfad in der Zeichnung, der mit Judith endet, impliziert, dass Judith gewinnt. Wir berechnen die Wahrscheinlichkeit für das Eintreten dieser Pfade und addieren diese Wahrscheinlichkeiten und erhalten

$$\mathbb{P}(\text{Judith gewinnt}) = 0.7^2 + 0.7 \cdot 0.3 \cdot 0.4 + 0.3 \cdot 0.4 \cdot 0.7 = 0.658.$$

- c) Gesucht ist  $\mathbb{P}(\text{J gewinnt S1} | \text{J gewinnt Spiel})$ . Wir benutzen die Bayes-Formel und erhalten

$$\begin{aligned} \mathbb{P}(\text{J gewinnt S1} | \text{J gewinnt Spiel}) &= \frac{\mathbb{P}(\text{J gewinnt Spiel} | \text{J gewinnt S1}) \mathbb{P}(\text{J gewinnt S1})}{\mathbb{P}(\text{J gewinnt Spiel})} \\ &= \frac{(0.7 + 0.3 \cdot 0.4) \cdot 0.7}{0.658} = 0.8723. \end{aligned}$$

- d) Da jedes Spiel mindestens 5 Minuten dauert und die Exponentialverteilung nur Wahrscheinlichkeitsmasse auf den positiven reellen Zahlen hat, muss  $c = 5$  sein. Weiter ist

$$\mathbb{E}[T] = \frac{1}{\lambda} + c \stackrel{!}{=} 8,$$

damit ist  $\lambda = \frac{1}{3}$ .

- e) Es gilt

$$\mathbb{P}(T \geq 5) = \mathbb{P}(\text{Exp}(1/2) + 4 \geq 5) = 1 - \mathbb{P}(\text{Exp}(1/2) < 1) = 1 - 1 + e^{-1/2} = e^{-1/2} = 0.6065.$$

- f) Da es sich bei  $T$  um eine stetig verteilte Zufallsvariable handelt, ist die Wahrscheinlichkeit von jedem Einpunkt-Ereignis 0.

**Aufgabe 4** (9 Punkte)

Die Zufallsvariablen  $X_1, \dots, X_n$  seien stochastisch unabhängig und identisch verteilt mit der Verteilung

$$\mathbb{P}_\vartheta(X = k) = \frac{2}{\vartheta} \cdot \left( \frac{\vartheta - 2}{\vartheta} \right)^k, \quad k = 0, 1, 2, \dots,$$

wobei  $\vartheta \in \Theta := (2, \infty)$  ein unbekannter Parameter ist. Der unbekannte Parameter  $\vartheta$  soll aufgrund einer unabhängigen Stichprobe  $x = (x_1, \dots, x_n)$  geschätzt werden, wobei  $x_1, \dots, x_n \geq 0$  sind.

**Hinweis:** Der Erwartungswert von  $X_1$  ist  $\mathbb{E}_\vartheta[X_1] = \frac{\vartheta}{2} - 1$ .

- a) Bestimmen Sie die Verteilungsfunktion  $F_\vartheta(t)$  für  $t \in \mathbb{N}_0$ . (1P)

$$F_\vartheta(t) = \boxed{1 - \left(1 - \frac{2}{\vartheta}\right)^{t+1}}$$

- b) Geben Sie die Likelihood-Funktion  $L_x(\vartheta)$  zur Stichprobe  $x$  an. (1P)

$$L_x(\vartheta) = \boxed{\frac{2^n}{\vartheta^n} \left( \frac{\vartheta-2}{\vartheta} \right)^{\sum_{l=1}^n x_l}}$$

- c) Geben Sie die Log-Likelihood-Funktion  $M_x(\vartheta)$  zur Stichprobe  $x$  an. (1P)

$$M_x(\vartheta) = \boxed{-n \log(\vartheta) + n \log(2) + \sum_{j=1}^n x_j (\log(\vartheta - 2) - \log(\vartheta))}$$

- d) Bestimmen Sie den Maximum-Likelihood-Schätzer  $\hat{\vartheta}$  für  $\vartheta$  zur Stichprobe  $x$ . (2P)

$$\hat{\vartheta}(x) = \boxed{\frac{2}{n} \sum_{j=1}^n x_j + 2}$$

- e) Geben Sie den Momentenschätzer  $\hat{\vartheta}_M$  für  $\vartheta$  zur Stichprobe  $x$  an. (2P)

$$\hat{\vartheta}_M(x) = \boxed{\frac{2}{n} \sum_{j=1}^n x_j + 2}$$

- f) Ist der Momenten-Schätzer erwartungstreu? Begründen Sie Ihre Antwort. (2P)

s. Lösungsvorschlag

**Lösungsvorschlag:**

a) Es liegt eine geometrische Verteilung mit Parameter  $p = \frac{\vartheta}{2}$  vor. Damit ist

$$F_{\vartheta}(t) = 1 - \left(1 - \frac{\vartheta}{2}\right)^{t+1}.$$

Dasselbe Ergebnis erhält man auch durch

$$F_{\vartheta}(t) = \sum_{j=0}^t \frac{2}{\vartheta} \left(\frac{\vartheta-2}{\vartheta}\right)^j.$$

b) Die Likelihood-Funktion lautet für die Stichprobe  $x = (x_1, \dots, x_n)$

$$\begin{aligned} L_x(\vartheta) &= \prod_{j=1}^n \mathbb{P}(X_j = x_j) = \frac{2^n}{\vartheta^n} \prod_{j=1}^n \left(\frac{\vartheta-2}{\vartheta}\right)^{x_j} \\ &= \frac{2^n}{\vartheta^n} \left(\frac{\vartheta-2}{\vartheta}\right)^{\sum_{l=1}^n x_l}. \end{aligned}$$

c) Die Log-Likelihood-Funktion ist

$$M_x(\vartheta) = \log(L_x(\vartheta)) = -n \log(\vartheta) + n \log(2) + \sum_{j=1}^n x_j (\log(\vartheta-2) - \log(\vartheta)).$$

d) Zur Bestimmung des Maximum-Likelihood Schätzers betrachten wir die Ableitung nach  $\vartheta$

$$\begin{aligned} M_x(\vartheta)' &= \frac{-n}{\vartheta} + \sum_{j=1}^n x_j \left(\frac{1}{\vartheta-2} - \frac{1}{\vartheta}\right) \\ &= \frac{2}{\vartheta(\vartheta-2)} \left(-\frac{n}{2}(\vartheta-2) + \sum_{j=1}^n x_j\right) \\ &= \frac{2}{\vartheta(-2+\vartheta)} \left(-\frac{n}{2}\vartheta + \sum_{j=1}^n x_j + n\right) \stackrel{!}{=} 0. \end{aligned}$$

Also ist

$$\hat{\vartheta} = \frac{2}{n} \sum_{j=1}^n x_j + 2$$

ein stationärer Punkt von  $M_x(\vartheta)$ . Da  $1/(\vartheta(\vartheta-2))$  positiv ist, gilt

$$M_x(\vartheta)' \begin{matrix} > \\ = 0 \\ < \end{matrix} \iff \sum_{i=1}^n x_i \begin{matrix} > \\ = \frac{n}{2} \cdot (\vartheta-2) \\ < \end{matrix} \iff \vartheta \begin{matrix} < \\ = \frac{2}{n} \sum_{i=1}^n x_i + 2 = 2\bar{x} + 2. \\ > \end{matrix}$$

$M_x(\vartheta)$  ist also unterhalb von  $2\bar{x} + 2$  steigend und danach fallend, also  $\hat{\vartheta}_n(x) = 2(\bar{x} + 1)$  die (einzige) Maximumstelle und damit der gesuchte Maximum-Likelihood-Schätzer.

e) Für eine geometrische Verteilung mit Parameter  $p$  ist der Erwartungswert gerade  $\frac{1-p}{p}$ . Nun liegt eine geometrische Verteilung mit Parameter  $p = \frac{\vartheta}{2}$  vor. Umformen und Einsetzen ergibt

$$\vartheta = 2(\mathbb{E}_{\vartheta}[X_1] + 1).$$

Damit entspricht der Momentenschätzer dem Maximum-Likelihood-Schätzer. Der Hinweis führt zu denselben Erkenntnissen.

f) Nach dem Hinweis gilt  $\mathbb{E}_\vartheta[X_1] = \frac{\vartheta}{2} - 1$ . Damit ist

$$\mathbb{E}[\hat{\vartheta}_M] = \mathbb{E}\left[\frac{2}{n}\sum_{j=1}^n X_j + 2\right] = \vartheta.$$

**Aufgabe 5** (9 Punkte)

Sie sind mit 4 Freunden in einem Restaurant. Dort werden 7 verschiedene Gerichte (nummeriert von 1 bis 7) angeboten. Im Folgenden möchte jeder von Ihnen genau ein Gericht bestellen. Des Weiteren betrachten wir nur die Bestellungen der gesamten Gruppe und sagen, dass zwei Bestellungen identisch sind, wenn sie dieselben Gerichte mit derselben Häufigkeit enthalten.

- a) Wie viele unterschiedliche Bestellungen sind möglich? (1.5P)

$$\text{Anzahl} = \boxed{\binom{11}{5} = 462}$$

- b) Wie viele unterschiedliche Bestellungen gibt es, wenn niemand das Gleiche wie ein anderer der Gruppe bestellen möchte? (1.5P)

$$\text{Anzahl} = \boxed{\binom{7}{5} = 21}$$

Bei der Bestellung erfahren Sie, dass jedes Gericht nur noch 2 Mal verfügbar ist. Sie wählen Ihre Gerichte nacheinander, wobei keine Präferenzen bestehen und deshalb jeder sein Gericht zufällig gleichverteilt aus den noch zur Verfügung stehenden Gerichten wählt.

- c) Wie groß ist die Wahrscheinlichkeit, dass kein Gericht doppelt bestellt wird? (2P)

$$\text{Wahrscheinlichkeit} = \boxed{\frac{360}{2401} = 0.1499}$$

- d) Wie groß ist die Wahrscheinlichkeit, dass Sie ein beliebiges Gericht bereits nicht mehr bestellen können, wenn Sie als 4. bestellen? (2P)

$$\text{Wahrscheinlichkeit} = \boxed{\frac{19}{49} = 0.3878}$$

Nun möchten Sie zusätzlich zu Ihren Gerichten Getränke bestellen. Nehmen Sie an, dass Sie und Ihre Freunde zufällig aus den Optionen “Mineralwasser”, “Cola” und “Apfelschorle” wählen können und die Zufallsvariable  $X$  die zufällige Anzahl an bestellten “Mineralwasser” angibt.

- e) Welche Verteilung hat  $X$ , wenn jeder von Ihnen genau ein Getränk bestellt und dabei jedes Getränk mit gleicher Wahrscheinlichkeit wählt? (1P)

$$X \sim \boxed{\text{Bin}(5, \frac{1}{3})}$$

- f) Geben Sie in der Situation von e) die Wahrscheinlichkeit an, dass mehr als 3 “Mineralwasser” bestellt werden. (1P)

$$\mathbb{P}(X > 3) = \boxed{0.0453}$$

## Lösungsvorschlag:

- a) Wir sind in einem Urnenmodell mit 7 durchnummerierten Kugeln (7 Gerichten) und 5 (Ziehungen). Wir ziehen mit Zurücklegen aber ohne Betrachtung der Reihenfolge. Damit liegt eine “5-Kombination” mit Wiederholung vor, also haben wir  $\binom{5+7-1}{5} = 462$  Möglichkeiten.
- b) Man möchte nun aus genau 7 Optionen 5 Optionen ohne Zurücklegen und ohne Betrachtung der Reihenfolge ziehen. Damit liegt eine “5-Kombination” ohne Wiederholung vor, also haben wir  $\binom{7}{5} = 21$  Möglichkeiten.
- c) Kein Gericht wird doppelt bestellt, wenn jeder etwas Anderes bestellt. Der erste Bestellende bestellt mit Wahrscheinlichkeit 1 etwas Neues. Der Zweite mit Wahrscheinlichkeit  $6/7$ ... Insgesamt erhalten wir

$$\frac{6}{7} \frac{5}{7} \frac{4}{7} \frac{3}{7} = \frac{360}{2401} = 0.1499.$$

- d) Ein Gericht kann nicht mehr bestellt werden, wenn bereits zwei der drei Bestellenden vor Ihnen das Gericht bestellt haben. Der erste Bestellende wählt beliebig. Nun ist es möglich, dass entweder der zweite Bestellende dasselbe wählt (Wahrscheinlichkeit dafür ist  $1/7$ ) oder etwas Anderes (Wahrscheinlichkeit dafür ist  $6/7$ ). Wählt er dasselbe, so ist irrelevant, was der dritte Bestellende wählt. Andernfalls muss dieser entweder dasselbe wie der Erste oder der Zweite bestellen (Wahrscheinlichkeit ist  $2/7$ ). Insgesamt haben wir

$$\frac{1}{7} + \frac{6}{7} \cdot \frac{2}{7} = \frac{19}{49} = 0.3878.$$

- e) Es handelt sich um ein Treffer-Niete Modell, wobei ausschließlich die Wahl “Mineralwasser” als Treffer gewertet wird. Da jede Wahl gleichwahrscheinlich ist, ist die Trefferwahrscheinlichkeit  $p = 1/3$  und damit, da 5 Mal unabhängig voneinander gewählt wird:

$$X \sim \text{Bin}(5, \frac{1}{3}).$$

- f) Da  $X \sim \text{Bin}(5, \frac{1}{3})$  gilt

$$\mathbb{P}(X > 3) = \mathbb{P}(X = 4) + \mathbb{P}(X = 5) = 5 \left(\frac{1}{3}\right)^4 \frac{2}{3} + \left(\frac{1}{3}\right)^5 = 0.0453.$$