WWZ
Wirtschaftswissenschaftliche Fakultät der
Universität Basel

Dr. Thomas Zehrt

Bitte in Druckbuchstaben ausfüllen:

Name	
Vorname	

Statistik
Probeprüfung 1

- Zeit: 90 Minuten, Maximale Punktzahl: 72
- Zur Orientierung: mit 36 Punkten haben Sie **sicher** bestanden.
- Die Prüfung umfasst 12 Aufgaben (1 bis 12) und die Bewertung der Teilaufgaben erfolgt gemäss den eingerahmten Punktzahlen.
- Provisorische Berechnungen sind auf separaten Blättern auszuführen. Diese Blätter sind -als Entwurf gekennzeichnet- ebenfalls abzugeben.
- Manche Aufgaben können auf verschiedene Arten gelöst werden. Eventuell gibt es einen sehr einfachen Weg. Überlegen Sie (kurz), bevor Sie drauf los rechnen!
- Die definitive Lösung darf von jeder Aufgabe nur eine Version enthalten und hat direkt im Anschluss an diese Aufgabe (bzw. auf der Rückseite des entsprechenden Aufgabenblattes) zu erfolgen. Dabei sollten alle Rechenschritte klar ersichtlich sein, denn die ,,Qualität" Ihrer Fehler wird bewertet. Ist die Antwort auf die Frage richtig und vollständig, erhalten Sie stets die volle Punktzahl.
- Bei den folgenden Fehlern erhalten Sie keine Punkte für die Aufgabe bzw. den Aufgabenteil:
  - eine Wahrscheinlichkeit < 0 oder > 1 oder eine negative Varianz (natürlich falsch) berechnet und nicht kommentiert;
  - Unabhängigkeit von Ereignissen bzw. Zufallsvariablen grundlos angenommen;
  - ein Summenzeichen bzw. Betragstriche (grundlos) weggelassen;
  - eine Gleichung durch eine Variable teilen, die Null sein könnte.
- Die ausgeteilten Formelsammlungen dürfen nicht beschriftet werden und sind ebenfalls mit der Prüfung abzugeben.

1. Bei einem Abfüllprozess wird eine Stichprobe vom Umfang $n = 15$ genommen. Folgende Abfüllmengen (in Gramm) werden dabei notiert (Urliste):

$$400, 399, 398, 400, 398, 399, 397, 400, 402, 399, 401, 399, 400, 402, 398.$$

   (a) Bestimmen Sie die empirische Verteilungsfunktion F dieser Daten. $\boxed{4}$

   (b) Bestimmen Sie (falls möglich) die Werte $F(-17)$, $F(400)$ und $F(500)$. $\boxed{2}$

   Lösung:

2. (a) Eine Aktie zeigt an 5 aufeinanderfolgenden Tagen die folgende Wertentwicklung:

Tag	1	2	3	4	5
Kurs	60.00	63.50	61.50	65.00	62

.

i. Welchen durchschnittlichen Wachstumsfaktor hat der Prozess? $\boxed{2}$

ii. Welchen Wert wird diese Aktie, basierend auf diesem durchschnittlichen Wachstumsfaktor, am 10. Tag haben? $\boxed{2}$

(b) Beweisen oder widerlegen Sie den folgenden Zusammenhang (zwischen arithmetischen und geometrischen Mittel):

Für alle reellen Zahlen $a, b > 0$ gilt $\dfrac{a+b}{2} \geq \sqrt{ab}$. $\boxed{3}$

Lösung:

3. Es wurden $n = 20$ Studenten nach ihren gerundeten Noten in Mathematik (Merkmal X) und in Statistik (Merkmal Y) befragt. Die Ergebnisse sind in folgender Urliste festgehalten:

$$(6,6), (4,4), (4,4), (4,5), (5,4), (3,4), (4,4), (4,4), (1,1), (4,5),$$
$$(3,4), (1,2), (3,1), (4,5), (5,4), (3,4), (5,5), (4,4), (4,4), (4,5).$$

(a) Erstellen Sie die zweidimensionale Häufigkeitstabelle (Kontingenztafel). $\boxed{2}$

(b) Berechnen Sie $\bar{x}$ und $var(X)$. $\boxed{4}$

Lösung:

		Y						Vert.
		1	2	3	4	5	6	von X
	1							
	2							
	3							
X	4							
	5							
	6							
Vert. von Y								

4. (a) Bestimmen Sie mit Hilfe der Methode der kleinsten Quadrate für die Wertepaare

x_i	1	2	3
y_i	1	5	10

eines Messvorgangs die Regressionskurve der Gestalt $y = f(x) = Ce^{Dx}$. $\boxed{4}$

(b) Erläutern Sie kurz, wie man ein Regressionsproblem mit einer Ansatzfunktion des Typs $y = f(x) = a \cdot x^b$ in ein lineares Regressionsproblem umwandeln kann. $\boxed{2}$

Lösung:

5. Erläutern Sie (kurz) die Bedeutung des Summanden G im additiven Zeitreihenmodell $y_i = y(t_i) = G(t_i) + S(t_i) + R(t_i)$. Wie kann G bestimmt werden? Wie ist die trendbereinigte Zeitreihe $y_i^\star$ definiert? $\boxed{6}$

   Lösung:

6. (a) Aus den Zahlen 1 bis 30 werden beim Zahlenlotto ,,5 aus 30" fünf verschiedene Zahlen ausgewählt. Wieviele **Möglichkeiten** gibt es

   i. genau fünf Richtige zu tippen oder $\boxed{1}$

   ii. mindestens zwei Richtige zu tippen? $\boxed{2}$

   (b) Die acht Ecken eines Würfels sind gleichmässig abgeschliffen, so dass der Würfel auch auf jeder Ecke liegen bleiben kann. Allerdings ist die Wahrscheinlichkeit jeder Ecke nur 1/4 so gross wie jede Seite. Wie gross ist die Wahrscheinlichkeit dafür, eine Sechs zu würfeln? $\boxed{3}$

Lösung:

7. Aus zwei Urnen U_1 und U_2 wird zufällig eine Urne ausgewählt, wobei jede Urne dieselbe Wahrscheinlichkeit besitzt, in die Auswahl zu gelangen. Die zwei Urnen enthalten weisse und schwarze Kugeln, wobei sich in Urne

   - U_1 : zwei weisse und fünf schwarze
   - U_2 : vier weisse und vier schwarze

   Kugeln befinden. Aus der zufällig gewählten Urne wird nun eine Kugel gezogen.

   (a) Wie gross ist die Wahrscheinlichkeit, dass die gezogenen Kugel weiss ist? $\boxed{3}$

   (b) Die gezogene Kugel ist schwarz. Wie gross ist die Wahrscheinlichkeit, dass sie aus Urne U_2 stammt? $\boxed{3}$

   <u>Lösung:</u>

8. (a) Sei X eine Zufallsvariable mit $E(X) = 3$ und $Var(X) = 2$. Schätzen Sie die Wahrscheinlichkeit $P(-1 < X < 7)$ (nach unten) ab. Können Sie auch (irgend) eine obere Schranke für diese Wahrscheinlichkeit angeben? Begründen Sie Ihre Aussage. $\boxed{4}$

(b) In einer Urne seien 6 Kugeln, durchnummeriert von 1 bis 6. Es werden zwei Kugeln **mit** Zurücklegen gezogen. X sei die Nummer der ersten Kugel und Y die der zweiten Kugel. Sind X und Y abhängig oder unabhängig? Begründen Sie Ihre Aussage. $\boxed{2}$

Lösung:

9. (a) In einer Telefonzentrale gehen im Durchschnitt 120 Anrufe pro Stunde ein. Wie gross ist die Wahrscheinlickeit, dass innerhalb einer **Minute** $\boxed{4}$

   i. kein Anruf eingeht,

   ii. genau ein Anruf eingeht oder

   iii. mehr als drei Anrufe eingehen?

   (b) Gegeben seien die x-Werte $x_1 = -2$, $x_2 = -1$, $x_3 = 0$ und $x_4 = 1$. Bestimmen Sie **zwei** Reihen von y-Werten $y_1, \ldots, y_4$, so dass der Korrelationskoeffizient der Wertepaare $(x_1, y_1), \ldots, (x_4, y_4)$ gleich -1 ist. $\boxed{2}$

Lösung:

10. Es sei folgende Familie von Funktionen gegeben:

$$\phi(x; \mu, \sigma) \quad := \quad \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

(a) Skizzieren Sie $\phi(x; -1, 1)$. $\boxed{2}$

(b) Berechnen Sie $\displaystyle\int_{\mu-2\sigma}^{\mu+2\sigma} \phi(x; \mu, \sigma)\, dx$. $\boxed{4}$

Lösung:

11. Es gelte $X \sim B(10; p)$. Bestimmen Sie einen ML-Schätzer p_{ML} für p bezüglich der konkreten Stichprobe $(0, 3, 2)$. Der Weg zur Herleitung **muss** angegeben werden. $\boxed{6}$

Lösung:

12. (a) In einer Prüfung werden 15 Fragen gestellt, die nur mit ,,Ja" oder ,,Nein" zu beantworten sind. Der Dozent legt fest, dass ein Student, der 13 oder mehr Fragen richtig beantwortet, die Prüfung bestanden hat. Die Nullhypothese H_0 sei: der Student hat (ausschliesslich) geraten. Bestimmen Sie die Wahrscheinlichkeit α eines Fehlers 1. Art. $\boxed{4}$

(b) Welche zwei Fehler können bei einem Hypothesentest auftreten, und wie sind sie zu interpretieren? $\boxed{2}$

Lösung:

Lösungen

1. (a) $F(x) = \begin{cases} 0 & \text{für } x < 397 \\ 1/15 & \text{für } x \in [397, 398) \\ 4/15 & \text{für } x \in [398, 399) \\ 8/15 & \text{für } x \in [399, 400) \\ 12/15 & \text{für } x \in [400, 401) \\ 13/15 & \text{für } x \in [401, 402) \\ 1 & \text{für } x \geq 402 \end{cases}$

   (b) $F(-17) = 0$ da $-17 < 397$, $F(400) = 12/15$ da $400 \in [400, 401)$ und $F(500) = 1$ da $500 \geq 402$ (**alle** Werte können stets bestimmt werden)

2. 1.00823 und 64.593

   Hinweis: Sicher gilt $(a-b)^2 \geq 0$ für alle reellen Zahlen.

3. $\bar{x} = 15/4$ und $var(X) = 111/76 = 1.4605$

4. (a) $C = 0.3684$ und $D = 1.1513$

   (b) Auf die Gleichung den ln anwenden ...

5. siehe Skript

6. (a) 1 und $26'126$

   (b) 1/8

7. 0.3928 und 0.4118

8. (a) $P(-1 < X < 7) = P(|X-3| < 4) \geq 1 - \frac{Var(X)}{4^2} = 7/8$ ist eine untere Schranke. Eine obere Schranke ist trivialerweise 1 (für jede Wahrscheinlichkeit) und eine bessere obere Schranke kann aus den vorhandenen Informationen wohl nicht bestimmt werden!?

   (b) -

9. (a) 0.1353, 0.2707 und 0.1428

   (b) -

10. Glockenkurve (siehe Skript), 0.955

11. Der Lösungsweg ist entscheidend! Hier die wichtigsten Schritte:

    - Likelihood-Funktion:
      $L(0,3,2;p) = \binom{10}{0} p^0 (1-p)^{10} \cdot \binom{10}{3} p^3 (1-p)^7 \cdot \binom{10}{2} p^2 (1-p)^8$
    - Log-Likelihood-Funktion:
      $\ln L(0,3,2;p) = \ln \binom{10}{0} + \ln p^0 + \ln(1-p)^{10} + \ln \binom{10}{3} + \ln p^3 + \ln(1-p)^7 + \ln \binom{10}{2} + \ln p^2 + \ln(1-p)^8$
    - $\frac{\partial}{\partial p} \ln L(0,3,2;p) = \frac{3}{p} + \frac{2}{p} - \frac{10}{1-p} - \frac{7}{1-p} - \frac{8}{1-p} = 0$
    - $p = \frac{5}{30} = \frac{1}{6}$

12. (a) $\alpha = 0.0037$

    (b) -