Mathematik 2, Statistik

Probeklausur

Zoltán Zomotor

Versionsstand: 13. Mai 2014, 16:44

Name	Vorname

Drei Punkte pro Aufgabe. Mindestpunktzahl zum Bestehen: 10 Punkte. Hilfsmittel: maximal zehn einseitig oder fünf beidseitig beschriftete DIN-A4-Spickzettel beliebigen Inhalts, möglichst selbst verfasst oder zusammengestellt; beliebiger Taschenrechner; kein Skript, keine andere Formelsammlung, kein Computer, kein Mobiltelefon.

Aufgabe 1:

Eine Firma erhält eine Lieferung von 12 Rechnern. Davon sind zwei Rechner defekt. Es werden zufällig fünf Geräte ausgewählt. Wie hoch ist die Wahrscheinlichkeit, dass unter diesen fünf Geräten *wenigstens* ein defektes Gerät ist?

Aufgabe 2:

Im Auftrag einer Winzergenossenschaft soll für die durchschnittliche Abfüllmenge einer Flaschenabfüllanlage, mit der 750 ml Weinflaschen gefüllt werden, ein 99 %-Schätzintervall bestimmt werden. Die Abfüllmenge X wird dabei als normalverteilt mit einer Standardabweichung von 10 ml angesehen. Es werden zehn auf dieser Anlage abgefüllte Flaschen zufällig ausgewählt und die Füllmenge kontrolliert. Die Stichprobe lieferte die folgenden Werte (Angaben in ml):

760 756 748 745 745 755 748 760 755 750.

Berechnen Sie das gesuchte Schätzintervall.

Aufgabe 3:

In einem Supermarkt werden grüne Gurken für 0.69 € das Stück angeboten. Angeblich wiegt jede Gurke 400 g. Ein Kunde wählt zufällig 20 Gurken aus und legt sie auf die Waage. Er berechnet aus den Messungen ein durchschnittliches Gewicht von 403.3 g und eine Standardabweichung von 6.73 g. Das Gewicht der Gurken kann als normalverteilt angesehen werden. Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob das durchschnittliche Gewicht der Gurken signifikant verschieden von 400 g ist.

Aufgabe 4:

Unter den Einwohnern einer Kleinstadt wird zufällig eine Frau über 40 ausgewählt. J und M bezeichnen die (zufällige) Anzahl von Jungen bzw. Mädchen, die die Frau geboren hat. Die Einzelwahrscheinlichkeiten stehen in der folgenden Tabelle:

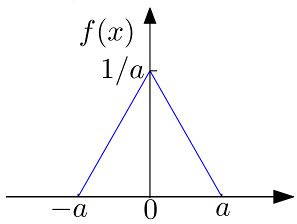
J \ M	0	1	2
0	0.10	0.15	0.15
1	0.15	0.05	0.10
2	0.15	0.10	0.05

Bestimmen Sie die Erwartungswerte, Varianzen, die Kovarianz und den Korrelationskoeffizienten von J und M.

Aufgabe 5:

Die Verteilung eines Merkmals X hängt von einem Parameter $a > 0$ ab. Die Verteilungsdichte ist

$$f(x) = \begin{cases} \frac{1}{a} + \frac{x}{a^2} & \text{falls } -a \le x \le 0, \\ \frac{1}{a} - \frac{x}{a^2} & \text{falls } 0 \le x \le a, \\ 0 & \text{sonst} \end{cases}$$



Bestimmen Sie die Varianz von X in Abhängigkeit von a.

Aufgabe 6:

Von einem Versicherungsunternehmen wurde für das vergangene Wirtschaftsjahr die folgende empirische Verteilung der Anzahl X der Schäden je Risiko in einer Haftpflichtversicherung registriert. Untersuchen Sie auf dem Signifikanzniveau von 0.05, ob angenommen werden kann, dass die Zufallsgröße X einer Poisson-Verteilung genügt.

Anzahl k der Schäden	Anzahl der Risiken mit genau k Schäden
0	51208
1	8105
2	642
3	45
4 oder mehr	0

Aufgabe 7:

Die folgende Tabelle enthält für Thüringen die Milchanlieferungen Y an Molkereien im Landkreis im Jahr 2006 und die Einwohnerzahl X am 31.12.2006:

Landkreis	Einwohnerzahl in 1000	Milchanlieferung in 1000 t
Eichsfeld	108.9	60.6
Nordhausen	92.6	34.3
Wartburgkreis	136.7	82.2
Unstrut-Hainich-Kreis	112.6	51.5
Kyffhäuserkreis	87.1	27.3
Schmalkalden-Meiningen	135.8	54.8
Gotha	142.5	41.7
Sömmerda	76.1	39.7
Hildburghausen	70.2	51.6
Ilm-Kreis	115.8	29.9
Weimarer Land	87.4	57.0
Sonneberg	63.1	16.7
Saalfeld-Rudolstadt	123.5	42.0
Saale-Holzland-Kreis	89.8	59.6
Saale-Orla-Kreis	92.1	110.3
Greiz	114.4	101.5
A1tenburger Land	104.7	41.0

(Quelle: Statistisches Jahrbuch Thüringen, 2007)

Daraus sind folgende Kennwerte berechnet worden:

$$\bar{x} = 103.135 \qquad s_x = 23.591$$
$$\bar{y} = 53.041 \qquad s_y = 25.055$$
$$s_{xy} = 144.97$$

Gibt es einen signifikanten Zusammenhang zwischen Einwohnerzahl und Milchanlieferung? Unterstellen Sie Normalverteilung der Merkmale und testen Sie mit der Irrtumswahrscheinlichkeit $\alpha = 0.05$.

Aufgabe 8:

Der wöchentliche Materialverbrauch (Angaben in Tonnen) zur Herstellung eines Produktes sei eine stetige Zufallsgröße X mit der folgenden Dichtefunktion:

$$f(x) = \begin{cases} 0.1 & \text{für } 0 \leq x \leq 5 \\ 0.04 \cdot (10 - x) & \text{für } 5 < x \leq 10 \\ 0 & \text{sonst} \end{cases}$$

Welche Materialmenge müsste gelagert werden, wenn die Wahrscheinlichkeit dafür, dass das gelagerte Material bereits vor Ablauf einer beliebigen Woche verbraucht ist, höchstens 0.05 betragen soll?

Lösung 1:

Die Anzahl X ist hypergeometrisch verteilt: $X \sim Hyp(N, M, n) = Hyp(12, 2, 5)$ ($n = 5$-faches Ziehen ohne Zurücklegen aus $N = 12$ Objekten, davon $M = 2$ markiert).

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} = 1 - \frac{\binom{2}{0}\binom{10}{5}}{\binom{12}{5}} = 1 - \frac{1 \cdot 252}{792}$$
$$= 0.682$$

Alternativ:

$$P(X \geq 1) = P(X = 1) + P(X = 2)$$

Bewertung:

- 1 Punkt: Hypergeometrische Verteilung angesetzt, x, N, M, n richtig zugeordnet
- 1 Punkt: Ansatz Wahrscheinlichkeit richtig
- 1 Punkt: Ergebnis korrekt.

Lösung 2:

Konfidenzintervall für μ bei Normalverteilung mit $\sigma = 10$ bekannt $\Rightarrow$ Kapitel 2.1

1. $1 - \alpha = 0.99$
2. $N(0, 1) : c = x_{1-\alpha/} = x_{0.995} = 2.5758$
3. $\bar{x} = 752.2$
4. $\dfrac{\sigma c}{\sqrt{n}} = \dfrac{10 \cdot 2.5758}{\sqrt{10}} = 8.1455$
5. $\left[744.055 \quad 760.346\right]$

Bewertung:

- 1 Punkt: Korrekter Ansatz μ bei Normalverteilung mit σ bekannt
- 1 Punkt: Alle 5 Schritte aufgeschrieben
- 1 Punkt: Ergebnis korrekt.

Lösung 3:

Einstichproben t-Test: Hypothesenpaar: $H_0 : \mu = 400$, $H_1 : \mu \neq 400$

1. Signifikanzniveau $\alpha = 0.05$
2. Testfunktionswert:

$$v = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{403.3 - 400}{6.73}\sqrt{20} = 2.193$$

3. Fraktil:

$$t(n-1) = t(19) : x_{1-\alpha/2} = x_{0.975} = 2.093$$

Verwerfungsbereich:

$$B = (-\infty, -2.093) \cup (2.093, \infty)$$

4. $v \in B \Rightarrow H_0$ verwerfen. Das mittlere Gewicht der Gurken ist signifikant von 400 g verschieden

Bewertung:

- 1 Punkt: Korrekten Test angesetzt
- 1 Punkt: Alle 4 Schritte aufgeschrieben
- 1 Punkt: Ergebnis korrekt

Lösung 4:

Randverteilungen bestimmen:

M / J	0	1	2	
0	0.10	0.15	0.15	0.40
1	0.15	0.05	0.10	0.30
2	0.15	0.10	0.05	0.30
	0.40	0.30	0.30	

J und M besitzen somit dieselbe Verteilung $\Rightarrow$

$$\begin{aligned}
&\mathrm{E}(J) = \mathrm{E}(M) = 0 \cdot 0.4 + 1 \cdot 0.3 + 2 \cdot 0.3 = 0.9 \\
&\mathrm{E}(J^2) = \mathrm{E}(M^2) = 0^2 \cdot 0.4 + 1^2 \cdot 0.3 + 2^2 \cdot 0.3 = 1.5 \\
&\mathrm{Var}(J) = \mathrm{Var}(M) = \mathrm{E}(J^2) - (\mathrm{E}(J))^2 = 1.5 - 0.9^2 = 0.69
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(J, M) &= \mathrm{E}(J \cdot M) - \mathrm{E}(J) \cdot \mathrm{E}(M) \\
&= 1 \cdot 1 \cdot 0.05 + 1 \cdot 2 \cdot 0.10 + 2 \cdot 1 \cdot 0.10 + 2 \cdot 2 \cdot 0.05 - 0.9 \cdot 0.9 \\
&= 0.65 - 0.81 = -0.16 \\
\varrho(J, M) &= \frac{\mathrm{Cov}(J, M)}{\sqrt{\mathrm{Var}(J)}\sqrt{\mathrm{Var}(M)}} \\
&= \frac{-0.16}{0.69} = -0.232
\end{aligned}$$

Lösung 5:

Erwartungswert ist offensichtlich null. Deswegen gilt:

$$\sigma^2 = \mathrm{Var}(X) = \int_{-a}^{a} (x-0)^2 f(x)\,\mathrm{d}x = 2 \cdot \int_{0}^{a} x^2 \cdot (\frac{1}{a} - \frac{x}{a^2})\,\mathrm{d}x$$
$$= 2 \cdot \left[\frac{x^3}{3a} - \frac{x^4}{4a^2}\right]_0^a = 2 \cdot \left(\frac{a^2}{3} - \frac{a^2}{4}\right) = \frac{a^2}{6}$$

Bewertung:

- 1 Punkt: Korrekter Ansatz für $\mathrm{Var}(X)$ mit $\mathrm{E}(x) = 0$
- 1 Punkt: Integriert
- 1 Punkt: Alle Rechenschritte und Ergebnis korrekt.

Lösung 6:

χ^2-Anpassungstest. Hypothesenpaar: $H_0 : F = P(\mu)$, $H_1 : F \neq P(\mu)$ Stichprobenumfang $n = 51208 + 8105 + 642 + 45 = 60\,000$ Bestimmung von

$$\lambda = \bar{x} = \frac{1}{60000}(0 \cdot 51208 + 1 \cdot 8105 + 2 \cdot 642 + 45 \cdot 3)$$
$$= 0.1587$$

1. Signifikanzniveau $\alpha = 0.05$
2. 2.1 Diskrete Verteilung, Intevalle gegeben

   2.2 h_j gegeben

   2.3

k	h_j	p_j	np_j
0	51208	0.8533	51195
1	8105	0.1354	8125
2	642	0.0107	645
3	45	$5.684 \cdot 10^-4$	34
4 oder mehr	0	$1 - P(x < 4) = 2.33 \cdot 10^{-5}$	1

   2.4 Testfunktionswert

$$v = \frac{1}{n} \sum_{j=1}^{5} \frac{h_j^2}{p_j} - n$$
$$= \frac{51208^2}{51195} + \frac{8105^2}{8124.7} + \frac{642^2}{644.7} + \frac{45^2}{34+1} - 60\,000$$
$$= 4.94$$

3. Die beiden letzten Intervalle zusammenfassen, 4 Intervalle mit $h_j \geq 5 \Rightarrow \chi^2(3) : x_{1-\alpha} = x_{0.95} = 7.82$. Verwerfungsbereich $B = (7.82, \infty)$

4. $v \notin B \Rightarrow H_0$ nicht ablehnen.

X genügt somit einer Poisson-Verteilung. Bewertung:

- 1 Punkt: χ^2-Anpassungstest angesetzt.
- 1 Punkt: Alle 4 Schritte notiert.
- 1 Punkt: Intervall zusammengelegt, korrekt gerechnet und Ergebnis korrekt.

Lösung 7:

Zweiseitiger Korrelationstest, Hypothesenpaar $H_0 : \varrho = 0$, $H_1 : \varrho \neq 0$

1. Signifikanzniveau $\alpha = 0.05$
2.

x_i	y_i	$x_i y_i$
108.9	60.6	6599.34
92.6	34.3	3176.18
136.7	82.2	11236.74
112.6	51.5	5798.90
87.1	27.3	2377.83
135.8	54.8	7441.84
142.5	41.7	5942.25
76.1	39.7	3021.17
70.2	51.6	3622.32
115.8	29.9	3462.42
87.4	57.0	4981.80
63.1	16.7	1053.77
123.5	42.0	5187.00
89.8	59.6	5352.08
92.1	110.3	10158.63
114.4	101.5	11611.60
104.7	41.0	4292.70

$$s_{xy} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right) = 144.97$$
$$r = \frac{s_{xy}}{s_x s_y} = \frac{144.97}{23.51 \cdot 25.055} = 0.245$$

Testfunktionswert:

$$\begin{aligned} v &= \sqrt{n-2}\frac{r}{\sqrt{1-r^2}} \\ &= \sqrt{15}\frac{0.245}{\sqrt{1-0.245^2}} \\ &= 0.98 \end{aligned}$$

3. Verwerfungsbereich:

$$t(15): x_{1-\alpha/2} = x_{0.975} = 2.13$$
$$\Rightarrow B = (-\infty, -2.13) \cup (2.13, \infty)$$

4. $v \notin B \Rightarrow H_0$ beibehalten,

Es gibt keinen signifikanten Zusammenhang zwischen Einwohnerzahl und Milchanlieferung.

Lösung 8:

Gesucht: x, so dass $P(X \geq x) \leq 0.05$ ist:

$$P(X \geq x) = 1 - P(X < x) = 1 - \int_0^x f(x)\,\mathrm{d}x = \int_x^{10} f(x)\,\mathrm{d}x$$

Für $x = 5 \Rightarrow P(X \leq 5) = \int_0^5 \frac{1}{10}\,\mathrm{d}x = 0.5 \Rightarrow x > 5$

$$\begin{aligned} P(X \geq x) &= \int_x^{10} f(x)\,\mathrm{d}x \\ &= \int_x^{10} 0.04(10-x)\,\mathrm{d}x = 0.04\left[10x - \frac{x^2}{2}\right]_x^{10} = 2 - 0.04x + 0.02x^2 = 0.05 \end{aligned}$$

Lösen der quadratischen Gleichung

$$0.02x^2 - 0.04x + 1.95 = 0 \Rightarrow x_1 = 11.58, x_2 = 8.42$$

Es müssen demnach 8.42 t Material gelagert werden.