Math. II, Statistik, Uebung3, Termin 25.04.2013 (AI), 29.04.2013 (BI)

1-Kontingenztest 2-Stetige Zufallsvariable 3-, 4-Diskrete Zufallsvariable

ALLGEMEINES

1c) "Da die Altersgruppenzugehörigkeit und die Intensität der sportlichen Beschäftigung voneinander unabhängig sind, muss der spezielle Multiplikationssatz für relative Häufigkeiten verwender werden."

Ursache und Folge sind hier vertauscht:
…weil hier der Spez. Multiplikationssatz gilt, kann man schließen, dass…

Der Multiplikationssatz gilt nur für Wahrscheinlichkeiten
…man benutzt ihn hier näherungsweise für die relativen Häufigkeiten, weil die Wahrscheinlichkeiten nicht bekannt sind

Wenn man den Multiplikationssatz auf relative Häufigkeiten anwendet, kann man prüfen, ab die Zufallsvariablen voneinander unabhängig sind.

3a) Die Approximationsbedingungen ausdrücklich benutzen:
Stichprobe ohne Zurücklegen. Für den Auswahlsatz gilt:
$n/N = 6/50 = 0,12 > 0,05$. $n = 6 < 100$ => Binomialverteilung nicht verwendbar
$n/p = 6/0,08 = 75 < 1500$ => POISSON-Verteilung nicht verwendbar

4a) <u>mindestens ein fehlerhafter Schalter</u>
Z.B. aus Skript sta4wahr.doc, Seite 4.6
d) Wahrscheinlichkeit für eine Inkubationszeit von mindestens 5 Tagen: $1 - F_4 = 0,27$
Z.B. aus Skript sta5verteil.doc, Seite 5.3
c) mindestens 5 c) $W(X \geq 5) = 1 - F(4) = 1 - 0,6331 = 0,3669$

AUFGABE 1 10 PUNKTE

Es soll untersucht werden, ob ein Zusammenhang besteht zwischen der Altersgruppe von Personen und ihrer sportlichen Betätigung. In der folgenden Tabelle sind die Altersgruppen den Intensitäten "selten", "gelegentlich", "regelmäßig" gegenübergestellt.
Bei 200 zufällig ausgewählten Personen fand man folgende Ergebnisse:

a) Testen Sie mit Hilfe des Chi^2-Unabhängigkeitstests, ob bei einem Sicherheitsgrad von 95% der Zusammenhang zwischen der Altersgruppen-Zugehörigkeit und der Intensität der sportlichen Betätigung statistisch gesichert ist.
Ergänzen Sie die Arbeitstabelle und bestimmen Sie die Freiheitsgrade. [1 P.]
Formulieren Sie Ihre Folgerung unter expliziten Nennung der beiden Zufallsvariablen [7 P.]

b) Prüfen Sie, ob man bei einem Signifikanzsniveau von $\alpha = 2\%$ die Unabhängigkeit der beiden Zufallsvariablen bestätigen kann oder ablehnen muss.
Formulieren Sie unter der Verwendung des Begriffes "Nullhypothese". [1 P.]

c) Erläutern Sie, wieso hier der spezielle Multiplikationssatz für relative Häufigkeiten bzw. Wahrscheinlichkeiten angewendet wird. [1 P.]

	Gruppe 1	< 30 Jahre	Gruppe 2	30 - 50 Jahre	Gruppe 3	> 50 Jahre	Summen
	n_{ij}	u_{ij}	n_{ij}	u_{ij}	n_{ij}	u_{ij}	
selten <1x	30	29,2	27	22,265	16	21,535	73
pro Woche	0,021917808		1,007		1,423		
gelegentlich	30	29,6	25	22,57	19	21,83	74
1x pro Woche	0,005		0,262		0,367		
regelmäßig	20	21,2	9	16,165	24	15,635	53
≥ 2x p.Woche	0,068		3,176		4,475		
Summen	80		61		59		200
						200	
	n = (3-1)(3-1) =		4				
	$\chi 2_{empir}$ =		10,805				
	$\chi 2_{crit}$ (ν=4, α=0,05) =		9,488				
	$\chi 2_{crit}$ (ν=4, α=0,02) =		11,668				

$$chi^2_{empirisch} = \chi^2_{empirisch} = \sum q_{ij} = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\left(n_{ij}-u_{ij}\right)^2}{u_{ij}} = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\left(n_{ij}-\frac{n_{i\bullet}\,n_{\bullet j}}{n}\right)^2}{\frac{n_{i\bullet}\,n_{\bullet j}}{n}}$$

a) Beim Sicherheitsgrad von 95% ist der Zusammenhang zwischen der Altersgruppen-Zugehörigkeit und der Intensität der sportlichen Betätigung statistisch gesichert, weil $\chi^2_{empirisch} > \chi^2_{crit.}$ *10,805 > 9,488*

b) Bei einem Sicherheitsniveau von $\alpha_2 = 2\%$ *kann man die Nullhypothese (Hypothese der Unabhängigkeit) beibehalten, weil* $\chi^2_{empirisch} < \chi^2_{crit.}$ *10,805 < 11,668*

c) Die Merkmale A und B sind voneinander unabhängig, wenn $h(A\cap B) = h(A)\cdot h(B)$
Diese Multiplikationen werden hier benutzt und führen zu den theoretischen Häufigkeiten u_{ij}. *(Die relativen Häufigkeiten werden als beste Schätzungen für die Wahrscheinlichkeiten verwendet.)*

Aufgabe 2 10 Punkte

An einer bestimmten Abfertigungsstelle warten die LKWs 0 bis 10 Minuten lang.

Die Wartezeit sei die stetige Zufallsvariable X [Minuten], für die die folgende

Dichtefunktion gilt: $f(x) = \begin{cases} 0,2 - 0,02\,x & \text{für } 0 \le x \le 10 \\ 0 & \text{für alle übrigen } x \end{cases}$

a) Zeichnen Sie den Funktionsgraphen der Dichtefunktion in die gegebene Koordinatenebene. [1 P.]

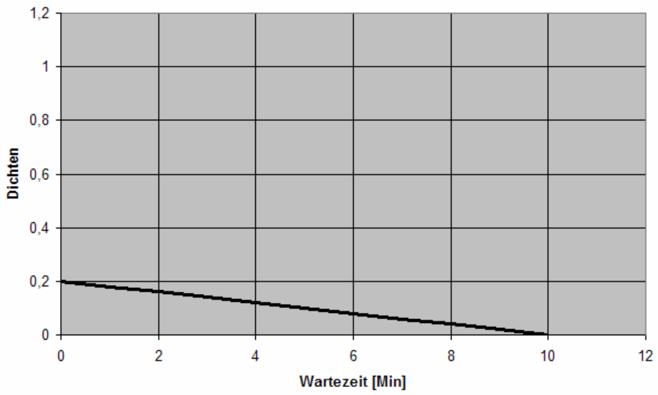
b) Bestimmen Sie die Funktionsgleichung der Verteilungsfunktion zu der gegebenen Dichtefunktion. [1,5 P.]

c) Bestimmen Sie die Wahrscheinlichkeit dafür, dass ein LKW höchstens 3 Minuten warten muss. [1 P.]

d) Bestimmen Sie die Wahrscheinlichkeit dafür, dass ein LKW zwischen 4 und 6 Minuten warten muss. [1,5 P.]

e) Bestimmen Sie den Erwartungswert μ. [2 P.]

f) Bestimmen Sie die Standardabweichung σ. [3 P.]



$$b)\ \int (0,2 - 0,02x)\,dx = 0,2x - 0,01x^2 \Rightarrow F(x) = \begin{cases} 0 & \text{für } x < 0 \\ 0,2x - 0,01x^2 & \text{für } 0 \le x \le 10 \\ 1 & \text{für } x > 10 \end{cases}$$

$$\text{c) } W(X \le 3) = \int_0^3 (0,2 - 0,02x)\,dx = \left[0,2x - 0,01x^2\right]_0^3 = 0,6 - 0,09 = 0,51 = 51\%$$

$$d) W(4 \le X \le 6) = \int_4^6 (0,2 - 0,02x)\,dx = \left[0,2x - 0,01x^2\right]_4^6 = 1,2 - 0,36 - (0,8 - 0,16) = 0,2$$

$$e) \mu = \int_0^{10} x \cdot (0,2 - 0,02x)\,dx = \int_0^{10} (0,2x - 0,02x^2)\,dx = \left[\frac{1}{10}x^2 - \frac{2}{300}x^3\right]_0^{10} = 10 - 6,667 = 3,333\,[\text{Minuten}]$$

$$f)\ \sigma^2 = \int_0^{10} x^2 \cdot (0,2 - 0,02x)\,dx - \left(\frac{10}{3}\right)^2 = \int_0^{10} (0,2x^2 - 0,02x^3)\,dx - \frac{100}{9}$$

$$= \left[\frac{0,2}{3}x^3 - \frac{0,02}{4}x^4\right]_0^{10} - \frac{100}{9} = 66,667 - 50 - 11,111 = 5,556\left[\text{Min}^2\right]$$

$$\sigma = \sqrt{5,556} = 2,36\,[\text{Min}]$$

AUFGABE 3 10 PUNKTE

Ein Betrieb fertigt 50 Fenster, bei denen im Durchschnitt 8% nachgearbeitet werden müssen. Aus der Produktion werden 6 Stück als Stichprobe ohne Zurücklegen entnommen.

a) Erläutern Sie genau, wieso man in diesem Fall weder die Binomialverteilung noch die POISSON-Verteilung benutzen kann. [2 P.]

b) Bestimmen Sie mit Hilfe der hypergeometrischen Verteilung die Wahrscheinlichkeit dafür, dass sich in der Stichprobe höchstens zwei Fenster befinden, die Nacharbeit erforderlich machen. Geben Sie Zwischenergebnisse an. [5 P.]

c) Bestimmen Sie diese Wahrscheinlichkeit (zum Vergleich) mit Hilfe der POISSON-Verteilung. Geben Sie Zwischenergebnisse an. [3 P.]

Lösung

a) Stichprobe ohne Zurücklegen. Für den Auswahlsatz gilt:
n/N = 6/50 = 0,12 > 0,05. n = 6 < 100 => Binomialverteilung nicht verwendbar
n/p = 6/0,08 = 75 < 1500 => POISSON-Verteilung nicht verwendbar

*b) p = 0,08, M = 0,08*50 = 4 x = 2 N=50 n=6*

$$W(X \le 2) = F_{Hyp|6;50;4}(2) = \frac{\binom{4}{0}\cdot\binom{50-4}{6-0}}{\binom{50}{6}} + \frac{\binom{4}{1}\cdot\binom{50-4}{6-1}}{\binom{50}{6}} + \frac{\binom{4}{2}\cdot\binom{50-4}{6-2}}{\binom{50}{6}} =$$

$$\frac{\binom{4}{0}\cdot\binom{46}{6}}{\binom{50}{6}} + \frac{\binom{4}{1}\cdot\binom{46}{5}}{\binom{50}{6}} + \frac{\binom{4}{2}\cdot\binom{46}{4}}{\binom{50}{6}} = 0,589 + 0,345 + 0,062 = 0,996$$

c) $\mu = np = 6 \cdot 0,08 = 0,48$ dafür steht keine Tabelle zur Verfügung

$$F_{Poi|0,48}(2) = \frac{0,48^0}{0!}e^{-0,48} + \frac{0,48^1}{1!}e^{-0,48} + \frac{0,48^2}{2!}e^{-0,48} = 0,619 + 0,297 + 0,07 = 0,986$$

AUFGABE 4 10 PUNKTE

Ein Betrieb produziert Schalter. Der Warenausgang wird durch Stichproben kontrolliert:
Ein Kunde entnimmt aus einem Karton à 60 Stück 12 Schalter.
Der Betrieb rechnet mit einem Ausschuss-Anteil von 5 %.
Der Karton geht an den Produzenten zurück, wenn dabei mindestens
ein fehlerhafter Schalter gefunden wird.

a) Mit welcher Wahrscheinlichkeit wird ein Karton zurückgewiesen, wenn
   (1) jeder geprüfte Schalter vor dem Ziehen des nächsten Schalters zurückgelegt wird? [2,5 P.]
   (2) wenn die geprüften Schalter nicht zurückgelegt werden? [2,5 P.]
   (3) man mit Hilfe der POISSON-Verteilung rechnet? [1,5 P.]

b) Bestimmen Sie den Erwartungswert, die erwartete Varianz und Standardabweichung für den Fall des Nicht-Zurücklegens (Aufgabenteil a (2)) [2,5 P.]

c) Wie groß muss die Gesamtproduktion an Schalter dieses Typs sein, damit man im Falle des Nicht-Zurücklegens hier die Binomial-Verteilung als Approximation benutzen kann? [1 P.]

Zufallsvariable X ist die Anzahl fehlerhafter Schalter.
Die Treffer-Wahrscheinlichkeit für fehlerhafter Schalter ist p = 0,05. n = 12 Stück.

a1) mit Zurücklegen, also BERNOULLI-Kette, p konstant, X ist binomialverteilt.
p=0,05, n=12 für diese Parameter ist keine Tabelle vorhanden.
$W(X \geq 1) = 1 - W(X \leq 0) = 1 - W(X=0) =$

$$= 1 - \binom{12}{0} 0,05^0 \cdot 0,95^{12} = 1 - 0,95^{12} = 1 - 0,540 = 0,460$$

a2) ohne Zurücklegen n/N = 12/60 = 0,2 > 0,05
also ist die Zufallsvariable X hypergeometrisch verteilt mit N=60, M=60·0,05=3
wieder gilt: $W(X \geq 1) = 1 - W(X \leq 0) = 1 - W(X=0) =$

$$= 1 - \frac{\binom{3}{0} \cdot \binom{60-3}{12-0}}{\binom{60}{12}} = 1 - \frac{\binom{3}{0} \cdot \binom{57}{12}}{\binom{60}{12}} = 1 - \frac{7,073 \cdot 10^{11}}{1,3994 \cdot 10^{12}} = 1 - 0,505 = 0,495$$

a3) $\mu = 0,05 \cdot 12 = 0,6$. *Keine Tabelle mit* $\mu = 0,6$
wieder gilt: $W(X \geq 1) = 1 - W(X \leq 0) = 1 - W(X=0) =$

$$= 1 - \frac{0,6^0}{0!} e^{-0,6} = 1 - 0,549 = 0,451$$

b) ohne Zurücklegen, X ist hypergeometrisch verteilt
$\mu = n\,p = 12 \cdot 0,05 = 0,6$ *Stück*

$$\sigma^2 = n\,p\,q \frac{N-n}{N-1} = 12 \cdot 0,05 \cdot 0,95 \cdot \frac{60-12}{60-1} = 0,57 \cdot 0,8136 = 0,464 \cdot \quad \sigma = 0,68 \text{ Stück.}$$

c) Die Grenze wäre bei n/N =0,05, hier: 12/N = 0,05 => 240 Stück