Einführung in die Stochastik
14. Übung
Gruppenübung: 30.06.2008
Lösungsvorschlag

Gruppenübung

G 24 Eine Apparatur füllt bestimmte Mengen (in g) eines pulverförmigen Medikaments ab. Es wird angenommen, daß diese Menge durch eine normalverteilte Zufallsvariable X mit unbekanntem Erwartungswert μ und unbekannter Varianz σ^2 beschrieben werden kann. Zehn Messungen ergaben ein empirisches Mittel von $\bar{x}_{10} = 19.6$ und eine empirische Varianz von $v_{10}(x) = 1.41$.

(i) Geben Sie eine Realisierung des Konfidenzintervalls für μ zum Niveau $1 - \alpha = 0.9$ an.
Hinweis: $\sum_{i=1}^{10} x_i^2 = 3854.3$

(ii) Wie verändern sich die Intervallgrenzen des Konfidenzintervalls für μ zum Niveau $1 - \alpha = 0.9$, falls zusätzlich $\sigma^2 = 1$ bekannt ist?

(i) Das Konfidenzintervall lautet (vgl. Vorlesung, VIII.2. Satz 32):

$$\left[\bar{X}_n - t_{n-1;1-\alpha/2} \cdot \sqrt{\frac{v_n(X)}{n}}\, , \bar{X}_n + t_{n-1;1-\alpha/2} \cdot \sqrt{\frac{v_n(X)}{n}}\right].$$

Durch Einsetzen des empirischen Mittelwertes $\bar{x}_{10} = 19.6$, der empirischen Varianz $v_{10}(x) = 1.41$ und des 95%-Quantils der t_9-Verteilung erhalten wir als konkretes Schätzintervall für μ

$$\begin{aligned}
&\left[\bar{x}_{10} - t_{9;1-0.1/2} \cdot \sqrt{\frac{v_{10}(x)}{10}}\, , \bar{x}_{10} + t_{9;1-0.1/2} \cdot \sqrt{\frac{v_{10}(x)}{10}}\right] \\
&= \left[19.6 - 1.83 \cdot \sqrt{\frac{1.41}{10}}\, , 19.6 + 1.83 \cdot \sqrt{\frac{1.41}{10}}\right] \\
&\approx [18.91, 20.29]
\end{aligned}$$

(ii) Das Konfidenzintervall ergibt sich in diesem Fall, da σ^2 als bekannt vorausgesetzt wird, durch (vgl. Vorlesung VIII.2. Satz 25)

$$\left[\bar{X}_{10} - u_{1-\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n}}\, , \bar{X}_{10} + u_{1-\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n}}\right].$$

Mit $\sigma^2 = 1$ und $u_{0.95} = 1.645$ ergibt sich als konkretes Schätzintervall für μ

$$\begin{aligned}
&\left[\bar{x}_{10} - u_{1-0.1/2} \cdot \sqrt{\frac{\sigma^2}{10}}\, , \bar{x}_{10} + u_{1-0.1/2} \cdot \sqrt{\frac{\sigma^2}{10}}\right] \\
&= \left[19.6 - 1.65 \cdot \sqrt{\frac{1}{10}}\, , 19.6 + 1.65 \cdot \sqrt{\frac{1}{10}}\right] \\
&\approx [19.08, 20.12]
\end{aligned}$$

G 25 Sei X eine normalverteilte Zufallsvariable mit unbekanntem Erwartungswert $\mu \in \mathbb{R} =: \Theta$ und bekannter Varianz σ^2. Wir betrachten ein Testproblem zur *einseitigen* Hypothese

$$\Theta_0 = \{\mu \in \Theta : \mu \geq \mu_0\}.$$

(i) Geben Sie einen Verwerfungsbereich R_n an, der einen Signifikanztest zu einem gegebenen Niveau α definiert.

(ii) Für eine gegebene Stichprobe von 20 unabhängigen $\mathbf{N}(\mu, 4)$-verteilten Zufallsvariablen sei $\bar{x}_{20} = 11$. Testen Sie unter Verwendung von (i) die Hypothese

$$\Theta_0 = \{\mu \in \Theta : \mu \geq 12\}$$

für $\alpha = 0.01$.

(i) Da die Varianz σ^2 bekannt ist, bietet sich als Teststatistik

$$g_n(X) = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$

an. Unter $\mu = \mu_0$ gilt $g_n(X) \sim \mathbf{N}(0,1)$. Da wir Θ_0 genau dann ablehnen, wenn die Realisierung der Teststatistik zu klein ist, wählen wir als Verwerfungsbereich

$$R_n = \{g_n \leq u_\alpha\}.$$

Damit ergibt sich nämlich für die Wahrscheinlichkeit eines Fehlers 1. Art, also für $\mu \in \Theta_0$:

$$\begin{aligned} P^\mu(\{X \in R_n\}) &= P^\mu(\{g_n(X) \leq u_\alpha\}) \\ &= P^\mu\left(\left\{\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \leq u_\alpha\right\}\right) \\ &\leq P^\mu\left(\left\{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq u_\alpha\right\}\right) \leq \alpha, \end{aligned}$$

so dass tatsächlich ein Signifikanztest zum Niveau α vorliegt. Die vorletzte Ungleichung folgt aus der Tatsache, dass für $\mu \in \Theta_0$ gerade $\mu \geq \mu_0$ gilt und damit

$$\bar{X}_n - \mu \leq \bar{X}_n - \mu_0,$$

so dass

$$\{\bar{X}_n - \mu_0 \leq u_\alpha \cdot \sigma/\sqrt{n}\} \subset \{\bar{X}_n - \mu \leq u_\alpha \cdot \sigma\sqrt{n}\}.$$

(ii) Wegen $u_{0.01} = -u_{0.99} = -2.326$ und $g_{20}(x) = \frac{11-12}{2}\sqrt{20} = -2.236$ lehnen wir die Nullhypothese nicht ab.

G 26 Der Personalchef einer Bank möchte untersuchen, ob Männer und Frauen die gleiche Chance besitzen, die Aufnahmeprüfung zu bestehen. Für 36 zufällig ausgewählte Stellenbewerber, von denen 23 Männer waren, wurde das Ergebnis der Prüfung ermittelt. Es zeigte sich, dass genau 17 Bewerber die Prüfung bestanden, davon 7 Frauen. Wir möchten die Unabhängigkeit des Prüfungsergebnisses von der Geschlechtszugehörigkeit überprüfen.

(i) Sei X die Anzahl der Männer, welche den Test bestehen. Es kann gezeigt werden, dass X unter der Annahme der Unabhängigkeit zwischen Prüfungsergebnis und Geschlechtszugehörigkeit hypergeometrisch verteilt ist. Bestimmen Sie für obige Realisierung die zugehörigen Parameter k, n_0, n der hypergeometrischen Verteilung.

(ii) Ausgehend von Ihrem Ergebnis in (i) bestimmen Sie einen Verwerfungsbereich für einen zweiseitigen Test der Unabhängigkeitsannahme (Prüfungsergebnis bzgl. Geschlechtszugehörigkeit) zum Niveau $\alpha = 0.05$. Zu welcher Entscheidung gelangen Sie anhand der oben gegebenen Stichprobe?

Hinweis:

- Die Verteilung von X ist durch folgende Tabelle gegeben (Aufgabenteil (i) muss trotzdem gelöst werden!):

x	**0**	**1**	**2**	**3**	**4**	**5**
$P(\{X=x\})$	0	0	0	0	$1.03 \cdot 10^{-6}$	$5.088 \cdot 10^{-5}$
x	**6**	**7**	**8**	**9**	**10**	**11**
$P(\{X=x\})$	$9.16 \cdot 10^{-4}$	$8.16 \cdot 10^{-3}$	0.041	0.122	0.228	0.270
x	**12**	**13**	**14**	**15**	**16**	**17**
$P(\{X=x\})$	0.232	0.095	0.027	$4.44 \cdot 10^{-3}$	$3.71 \cdot 10^{-4}$	$1.17 \cdot 10^{-5}$

- Bei diesem Test handelt es sich um einen exakten Test (die Wahrscheinlichkeit dafür, dass die Testgröße Werte im Ablehnungsbereich annimmt, lässt sich exakt bestimmen). In der Literatur wird er auch als *Fisher-Test* bezeichnet.

(i) Der Übersicht halber fassen wir die gegebenen Daten in folgender Tabelle zusammen:

	Geschlechtszugehörigkeit	
Prüfungsergebnis	*männlich*	*weiblich*
bestanden	*10*	*7*
nicht bestanden	*13*	*6*

Unter der Annahme der Unabhängigkeit des Prüfungsergebnisses von der Geschlechterzugehörigkeit folgt $X \sim \mathbf{H}(36, 23, 17)$, also

$$P(\{X=k\}) = \frac{\binom{23}{k}\binom{13}{17-k}}{\binom{36}{17}}, \qquad k = 4, \ldots, 17.$$

(ii) Unter der Unabhängigkeitsannahme ist X hypergeometrisch (s. (i)). Unser Verwerfungsbereich enthält nun alle die Werte, welche zu klein bzw. zu groß sind (zweiseitiger Test!) für diese Annahme. Seien dazu $h_{\alpha/2}$ und $h_{1-\alpha/2}$ die Quantile der $H(36, 23, 17)$-Verteilung, dann gilt

$$P(\{h_{\alpha/2} \leq X \leq h_{1-\alpha/2}\}) = F_X(h_{1-\alpha/2}) - F_X(h_{\alpha/2}-) \leq 1 - \alpha,$$

wobei F_X die Verteilungsfunktion von X, also der hypergeometrischen Verteilung, bezeichnet und für das h_α-Quantil gilt $F_X(h_\alpha-) \leq \alpha \leq F_X(h_\alpha)$. Wir müssen demnach das $h_{0.025}$- und das $h_{0.975}$-Quantil bestimmen (mittels der im Hinweis gegeben Werte der Wahrscheinlichkeitsfunktion von X):

$$P(\{X \leq 8\}) = 0.0499 > 0.025 \Longrightarrow h_{0.025} = 8$$
$$P(\{X \geq 14\}) = 0.032 > 0.025 \Longrightarrow h_{0.975} = 14$$

Da in unserem Falle $x = 10$, kann die Annahme beim Niveau $\alpha = 0.05$ nicht abgelehnt werden.