

Statistik I für Human- und Sozialwissenschaften

13. Übung

Lösungsvorschlag

Gruppenübung

G 42 Bei einem Fährunternehmen weiß man, dass im Mittel 16% derjenigen Personen, die einen Platz für ihren PkW auf einer Fähre buchen, zur Abfahrt nicht erscheinen. Um die Zahl der somit ungenutzten Plätze nicht zu groß werden zu lassen, werden daher für eine Überfahrt, bei der 120 Plätze zur Verfügung stehen, mehr als 120 Buchungen vorgenommen.

Man ermittle näherungsweise die Wahrscheinlichkeit dafür, daß alle zur Abfahrt erschienenen Personen, für die ein PkW-Platz reserviert wurde, diesen auch bekommen. Insgesamt wurden 130 Buchungen vorgenommen und wir können annehmen, dass das Erscheinen der einzelnen Personen zur Abfahrt unabhängig voneinander ist.

Wir benutzen hier den Zentralen Grenzwertsatz. Seien X_i , $i = 1, \dots, 130$ die Zufallsvariablen, die für das Erscheinen der 130 für die Fähre gebuchten PkWs stehen. Dann gilt

$$P(X_i = 1) = 1 - 0.16 = 0.84, \quad P(X = 0) = 0.16, \quad i = 1, \dots, 130.$$

Somit gilt

$$Y_{130} = \sum_{i=1}^{130} X_i \sim \text{Bin}(130, 0.84)$$

und

$$E(X_1) = 0.84, \quad \text{Var}(X_1) = 0.84 \cdot 0.16 = 0.1344.$$

Mit dem Zentralen Grenzwertsatz können wir nun schlußfolgern

$$\begin{aligned} P(Y_{130} \leq 120) &= P\left(\frac{Y_{130} - 130 \cdot E(X_1)}{\sqrt{130 \cdot \text{Var}(X_1)}} \leq \frac{120 - 130 \cdot E(X_1)}{\sqrt{130 \cdot \text{Var}(X_1)}}\right) \\ &\approx \Phi\left(\frac{120 - 109.2}{\sqrt{17.47}}\right) \\ &\approx \Phi(2.58) \approx 0.995. \end{aligned}$$

G 43 Zwölf Versuchsflächen wurden mit einer neuen Weizensorte bestellt. Diese Flächen erbrachten folgende Hektarerträge (in dz)

$$35.6, 33.7, 37.8, 31.2, 37.2, 34.1, 37.1, 34.9, 35.6, 34.0, 35.8, 36.6$$

Aus Erfahrung weiß man, dass die Hektarerträge als Realisierungen unabhängiger $N(\mu, 4.13)$ -verteilter Zufallsvariablen angesehen werden können.

Geben Sie für den Erwartungswert μ ein konkretes Schätzintervall zum Niveau 0.95 an.

$$\begin{aligned} &\left[\bar{X} - z_{0.975} \cdot \sqrt{\frac{\sigma^2}{n}}, \bar{X} + z_{0.975} \cdot \sqrt{\frac{\sigma^2}{n}}, \right] \\ &= \left[35.3 - 1.96 \cdot \sqrt{\frac{4.13}{12}}, 35.3 + 1.96 \cdot \sqrt{\frac{4.13}{12}} \right] \\ &= [32.96, 37.64] \end{aligned}$$

G 44 Vor der Einführung einer neuen Getreidesorte wurden vierzehn Flächen gleicher Größe mit dieser neuen Getreidesorte bestellt und die Erträge (in dz/ha) ermittelt

37.6, 35.1, 36.8, 38.2, 35.0, 32.2, 36.6,
38.8, 33.4, 35.9, 36.3, 34.7, 38.2, 36.4

Erfahrungsgemäß kann der Ertrag als normalverteilt angesehen werden.

- (a) Kann man annehmen, dass der mittlere Ertrag der neuen Sorte 37 dz/ha beträgt?
(b) Ist die Varianz σ^2 wesentlich größer als 3.3

Wählen Sie als Irrtumswahrscheinlichkeit in beiden Fällen $\alpha = 0.05$.

Hinweis: $\sum_{i=1}^{14} (x_i - \bar{x})^2 = 45.94$

(a) $H_0 : \mu_0 = 37, \quad H_1 \neq 37$

Da wir die Varianz σ^2 nicht kennen, wenden wir den zweiseitigen t -Test an.

- Teststatistik:

$$T = \frac{\bar{X} - \mu_0}{\bar{s}} \sqrt{n}$$

$$T_{20} = \frac{36.1 - 37}{\sqrt{3.5}} \cdot \sqrt{14} \approx -1.84$$

Kritischer Bereich: $|T| > t_{13;0.975}$

Wegen $|T| = 1.84 < 2.16 = t_{13;0.975}$ wird H_0 nicht abgelehnt.

(b) $H_0 : \sigma^2 \leq 3.3, \quad H_1 : \sigma^2 > 3.3$

Für das Testen dieser Hypothese verwenden wir den einseitigen χ^2 -Test.

- Teststatistik:

$$T = \frac{(n-1)\bar{S}^2}{\sigma^2}$$

$$T_{20} = \frac{13 \cdot 3.5}{3.3} = 13.79$$

- Kritischer Bereich: $T > \chi_{n-1;0.95}^2$

Wegen $13.79 < 22.36$ kann H_0 nicht abgelehnt werden.

G 45 Auf einer Versuchsfläche soll der Einfluß eines bestimmten Düngemittels auf die Weizenproduktion geprüft werden. Dazu wählte man 20 Äcker gleicher Flächengröße aus und düngte die Hälfte (Gruppe I). Die andere Hälfte (Gruppe II) wurde nicht gedüngt. Die Ernte lieferte folgende Ergebnisse (in dz):

Gruppe I:	3.8	4.2	6.1	2.6	4.9	3.7	0.8	2.1	5.3	4.7
Gruppe II:	2.7	3.7	3.6	5.3	3.1	3.8	3.0	2.1	4.1	3.6

Es soll untersucht werden, ob sich die mittleren Erträge in den beiden Gruppen wesentlich voneinander unterscheiden, d.h. ob die Düngung einen Einfluß auf den mittleren Ertrag hat. Setzen Sie dazu die Normalverteilung der Erträge voraus.

Prüfen Sie jedoch zuerst die Gleichheit der Varianzen (mit $\alpha = 0.01$). Können Sie aufgrund des Ergebnisses dieses Tests den zweiten Test (auf Gleichheit der mittleren Erträge) mit einem der Ihnen bekannten Tests durchführen? Begründen Sie Ihre Entscheidung.

benötigte Quantile: $F_{9,9;0.995} = 3.18$, $F_{9,9;0.005} = \frac{1}{3.18}$

Test auf Gleichheit der Varianzen: $H_0 : \sigma_I^2 = \sigma_{II}^2$, $H_1 : \sigma_I^2 \neq \sigma_{II}^2$
Wir führen demnach einen zweiseitigen F -Test durch.

- Teststatistik:

$$T = \frac{\bar{S}_I^2}{\bar{S}_{II}^2}$$

$$T_{10} = 3.41$$

- Kritischer Bereich: $T > F_{m-1,n-1;1-\alpha/2}$ oder $T < F_{m-1,n-1;\alpha/2}$

Wegen $3.41 > 3.18$ lehnen wir die Nullhypothese ab.

Demnach können wir auch nicht den t -Test anwenden, um auf die Gleichheit der mittleren Erträge zu testen, da dieser $\sigma_I^2 = \sigma_{II}^2$ voraussetzt. Es muß also ein neues Testverfahren gesucht werden (was an dieser Stelle jedoch ausbleibt.)