Aufgabe 1

Bei der Abfüllung von Mineralwasser in 1-Liter-Flaschen soll der Sollwert von 1 l eingehalten werden. Für die verwendete Abfüllanlage gilt nach Herstellerangaben, dass die Abfüllungen normalverteilt sind mit $\mu = 1\,000$ ml und $\sigma^2 = 100$ ml^2. Zur Überprüfung der Abfüllmenge wird von einem Getränkemarkt eine Stichprobe vom Umfang $n = 10$ aus einer Lieferung von $1\,000$ Flaschen gezogen. Dabei ergab sich eine durchschnittliche Abfüllmenge von $1\,020$ ml. Sie dürfen davon ausgehen, dass die Herstellerangabe $\sigma^2 = 100$ ml^2 korrekt ist.

Befindet sich der Abfüllprozess nicht mehr unter statistischer Kontrolle ($\alpha = 0.01$)?

Loesung 1

Sichprobenumfang $n = 10$ und $\mu_0 = \bar{x} = 1020$ ml

$$H_0: \ \mu = 1000 \quad vs. \quad H_1: \ \mu \neq 1000, \quad \alpha = 0.01.$$

Unter H_0 sind die Abfüllungen $X_i \sim N(\mu_0, \sigma^2)$, mit $\mu_0 = 1000$ ml, $\sigma^2 = 100(\text{ml})^2$

Das arithmetische Mittel $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ ist als Linearkombination von normalverteilten Zufallsgrößen wieder normalverteilt und zwar $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$.

Unter H_0: $\frac{\bar{X}-E(\bar{X})}{\sqrt{Var(\bar{X})}} = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \overset{H_0}{\sim} N(0,1)$

Der Abfüllprozess befindet sich nicht mehr unter statistischer Kontrolle, wenn die Wahrscheinlichkeit, dass H_0 verworfen wird, obwohl H_0 wahr ist, kleiner als α ist:

$$\begin{aligned}
P_{H_0}\left(\left|\frac{\bar{X}-\mu_0}{\sigma}\sqrt{n}\right| > \left|\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right|\right) &= 1 - P_{H_0}\left(-\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n} \leq \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \leq \frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right) \\
&\overset{\bar{X}\overset{H_0}{\sim}N(0,1)}{=} 1 - \left[\Phi\left(\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right) - \Phi\left(-\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right)\right] \\
&\overset{\Phi(-x)=1-\Phi(x)}{=} 1 - \left[\Phi\left(\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right) - \left(1 - \Phi\left(\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right)\right)\right] \\
&= 2 - 2\Phi\left(\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right) \\
&= 2 - 2\Phi\left(\frac{1020-1000}{10}\sqrt{10}\right) \\
&= 2 - 2\underbrace{\Phi\left(2\sqrt{10}\right)}_{\approx 1} \approx 0 < \alpha = 0.01
\end{aligned}$$

$\Rightarrow$ p-Wert $< \alpha \Rightarrow H_0$ wird verworfen, d.h. der Abfüllprozess befindet sich nicht mehr unter statistischer Kontrolle.

Aufgabe 2

Sei $X = (X_1, \dots, X_{n_X})$ eine unabhängig identisch verteilte Stichprobe, mit $X_i \sim N(\mu_X, \sigma_0^2)$ und $Y = (Y_1, \dots, Y_{n_Y})$ eine unabhängig identisch verteilte Stichprobe, mit $Y_i \sim N(\mu_Y, \sigma_0^2)$. X und Y seien unabhängig, die Parameter μ_X, μ_Y unbekannt und σ_0^2 bekannt.

a) Geben Sie einen Test zum Niveau α zur Überprüfung der Hypothese $H_0 : \mu_X = \mu_Y$ gegen die Alternative $H_1 : \mu_X \neq \mu_Y$ an.
Hinweis: Verwenden Sie zur Konstruktion dieses Tests die Differenz $\bar{X} - \bar{Y}$.

b) Bestimmen Sie ein Konfidenzintervall zum Vertrauensgrad $1 - \alpha$ für die Differenz der Mittelwerte $\mu_X - \mu_Y$.

c) Geben Sie das Konfidenzintervall zum Vertrauensgrad 0.95 und den Test zum Niveau $\alpha = 0.05$ an für den Fall, dass die Stichprobe von X die Werte

$$X = (0.4, -3.4, 1.3, 1.9, -0.2)$$

und die Stichprobe von Y die Werte

$$Y = (3.7, -0.6, 1.5, 1.9, -1.7, -1.1, 3.3)$$

ergeben hat.
Gehen Sie davon aus, dass $\sigma_0^2 = 2$ ist.

Loesung 2

Seien $X = (X_1, \dots, X_{n_X})$ und $Y = (Y_1, \dots, Y_{n_Y})$ zwei unabhängig identisch verteilte Stichproben mit $X_i \sim N(\mu_X, \sigma_0^2)$ und $Y_i \sim N(\mu_Y, \sigma_0^2)$. Dabei seien X und Y unabhängig, μ_X, μ_Y unbekannt und σ_0^2 bekannt.

a) Nun soll überprüft werden, ob beim Hypothesenpaar $H_0 : \ \mu_X = \mu_Y$ vs. $H_1 : \mu_X \neq \mu_Y$, H_0 zum Niveau α abgelehnt werden kann.
Analog dazu kann die Hypothese folgendermaßen formuliert werden, wenn man $\mu_D = \mu_X - \mu_Y$ setzt:

$$H_0 : \ \mu_D = 0 \quad \text{vs.} \quad H_1 : \ \mu_D \neq 0$$

μ_D kann durch $\bar{D} = \bar{X} - \bar{Y}$ geschätzt werden.
Bestimme zunächst Erwartungswert und Varianz von $\bar{X}$ bzw. $\bar{Y}$:

$$E(\bar{X}) = \frac{1}{n_X} \sum_{i=1}^{n_X} E(X_i) = \mu_X$$

$$Var(\bar{X}) = Var\left(\frac{1}{n_X} \sum_{i=1}^{n_X} X_i\right) = \frac{1}{n_X^2} \sum_{i=1}^{n_X} Var(X_i) = \frac{1}{n_X} \sigma_0^2$$

Damit gilt:

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma_0^2}{n_X}\right) \text{ und analog } \bar{Y} \sim N\left(\mu_Y, \frac{\sigma_0^2}{n_Y}\right)$$

Für die Differenz $\bar{D} = \bar{X} - \bar{Y}$ gilt:

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$$

und

$$Var(\bar{X}-\bar{Y}) = Var(\bar{X}) + Var(\bar{Y}) = \frac{\sigma_0^2}{n_X} + \frac{\sigma_0^2}{n_Y} = \frac{n_Y + n_X}{n_X n_Y}\sigma_0^2$$

Da die Differenz normalverteilter Zufallsvariablen wieder normalverteilt ist, gilt:

$$\bar{X}-\bar{Y} \sim N\left(\mu_X - \mu_Y, \frac{n_Y + n_X}{n_X n_Y}\sigma_0^2\right)$$

Die benötigte Teststatistik erhält man daraus durch Standardisierung:

$$|T| = \left|\frac{\bar{X}-\bar{Y}-(\mu_X-\mu_Y)}{\sigma_0}\sqrt{\frac{n_X n_Y}{n_Y + n_X}}\right|$$

Man sieht hieraus, dass der Ablehnbereich aus großen Werten besteht, da $|T|$ groß wird, wenn $\bar{X}-\bar{Y}$ stark von $\mu_X - \mu_Y$ abweicht.
Als Ablehnbereich für den Test erhält man:

$$K = \left\{(\bar{x},\bar{y}) : |\bar{x}-\bar{y}| > z_{1-\frac{\alpha}{2}}\sigma_0\sqrt{\frac{n_Y+n_X}{n_X n_Y}} + \underbrace{(\mu_X-\mu_Y)}_{\text{unter } H_0 = 0}\right\}$$

Also:

$$\varphi(x,y) = \begin{cases} 0 & \text{, falls } |\bar{x}-\bar{y}| \notin K \\ 1 & \text{, falls } |\bar{x}-\bar{y}| \in K \end{cases}$$

b) Gesucht ist nun das Konfidenzintervall zum Vertrauensgrad $1-\alpha$ für die Differenz $\mu_X - \mu_Y$. Aus dem in Teilaufgabe a) bestimmten kritischen Bereich ergibt sich das Konfidenzintervall

$$\mathrm{KI} = \left[(\bar{x}-\bar{y}) \pm z_{1-\frac{\alpha}{2}}\sigma_0\sqrt{\frac{n_Y+n_X}{n_X n_Y}}\right]$$

c) $\alpha = 0.05, n_X = 5, n_Y = 7, \sigma_0^2 = 2, z_{1-0.025} = z_{0.975} = 1.96$
$X = (0.4, -3.4, 1.3, 1.9, -0.2)$ $Y = (3.7, -0.6, 1.5, 1.9, -1.7, -1.1, 3.3)$
Es ergeben sich aus den Stichproben $\bar{x} = 0$ und $\bar{y} = 1$.
Das Konfidenzintervall berechnet sich wie folgt:

$$\begin{aligned} \mathrm{KI} &= \left[(0-1) \pm 1.96\sqrt{\sigma_0^2}\cdot\sqrt{\frac{7+5}{5\cdot 7}}\right] \\ &= [-1 \pm 1.623] \\ &= [-2.623, 0.623] \end{aligned}$$

Der Ablehnbereich für den Test lautet:

$$\begin{aligned} K &= \left\{(\bar{x},\bar{y}) : |\bar{x}-\bar{y}| > 1.96\cdot\sqrt{2}\cdot 0.59\right\} \\ &= \{(\bar{x},\bar{y}) : |\bar{x}-\bar{y}| > 1.623\} \end{aligned}$$

Im Beispiel ist $|\bar{x}-\bar{y}| = 1$ kleiner als 1.623
$\Rightarrow$ H_0 kann nicht verworfen werden.