# حل تمرین روش‌های آماری

## جلسه‌ی چهارم

① فرض کنید که یک داروی استاندارد ۲۵٪ در درمان یک بیماری موثر است و شخصی ادعا می‌کند که دارویی که او ساخته است ۵۰٪ در درمان آن بیماری موثر است. فرض کنید که ۲۰ بیمار را انتخاب کنیم و متغیر تصادفی $X$ را برابر تعداد بیماران بهبود یافته توسط داروی جدید در بین ۲۰ بیمار درنظر بگیریم. با درنظر گرفتن $C = \{X \mid X \geqslant 9\}$، احتمال خطای نوع اول و دوم را محاسبه کنید.

$$\begin{cases} H_0: \ p = \frac{1}{4} \\ H_1: \ p = \frac{1}{2} \end{cases} \qquad X \sim B(20, p)$$

$$\alpha = P(X \in C \mid \text{درست } H_0) = P(X \geqslant 9 \mid p = \tfrac{1}{4})$$

$$= 1 - P(X \leqslant 8 \mid p = \tfrac{1}{4}) \simeq 1 - 0{,}9591 = 0{,}0409$$

$$\beta = P(X \notin C \mid \text{درست } H_1) = P(X \leqslant 8 \mid p = \tfrac{1}{2}) \simeq 0{,}2517$$

---

② با فرض $X \sim N(\mu, 4)$ و
$$\begin{cases} H_0: \ \mu = 0 \\ H_1: \ \mu = 1 \end{cases}$$

یک نمونه‌ی تصادفی $X_{25}, \dots, X_2, X_1$ از $X$ جمع آوری شده است. اگر ناحیه‌ی بحرانی

به صورت $C=\{X_1,\dots,X_{25} \mid \bar{X} > 0.4\}$ باشد، $\alpha$ و $\beta$ را بدست آوریم.

$$\alpha = P(\bar{X} > 0.4 \mid \mu = 0) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{0.4-0}{2/5}\right) = P(Z > 1)$$

$$= 1 - P(Z \leq 1) \simeq 1 - 0.8413 = 0.1587$$

$$\beta = P(\bar{X} \leq 0.4 \mid \mu = 1) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{0.4-1}{2/5}\right) = P(Z \leq -1.5)$$

$$\simeq 0.0668$$

---

۳) در مثال قبل، اگر ناحیه‌ی بحرانی به صورت $C=\{X_1,\dots,X_{25} \mid \bar{X} > c\}$ باشد، مقدار $c$ را به گونه‌ای تعیین کنید که $\alpha = 0.1$ باشد و احتمال خطای نوع دوم را بدست آوریم.

$$0.1 = \alpha = P(\bar{X} > c \mid \mu = 0) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{c-0}{2/5}\right) = P(Z > 2.5c)$$

$$\Longrightarrow P(Z \leq 2.5c) = 0.9 \Longrightarrow 2.5c = z_{0.9} = 1.28$$

$$\Longrightarrow c = \frac{1.28}{2.5} \simeq 0.512$$

$$\beta = P(\bar{X} \le c \mid \mu = 1) = P\left( \frac{\bar{X} - \mu}{6/\sqrt{n}} \le \frac{0.512 - 1}{2/5} \right) = P(Z \le -1.22)$$

$$\simeq 0.1112$$

---

④ فرض کنید $X$ یک متغیر تصادفی گسسته با یکی از تابع‌های احتمال زیر باشد:

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ |
|---|---|---|---|---|---|---|
| $f_0(X)$ | $0.05$ | $0.05$ | $0.1$ | $0.1$ | $0.2$ | $0.5$ |
| $f_1(X)$ | $0.1$ | $0.15$ | $0.25$ | $0.15$ | $0.25$ | $0.1$ |

در آزمون فرض‌های $\begin{cases} H_0 : X \sim f_0 \\ H_1 : X \sim f_1 \end{cases}$ کدام یک از ناحیه‌های بحرانی زیر را ترجیح می‌دهید؟

$C_1 = \{ x_1 , x_3 \}$ $\qquad\qquad$ $C_2 = \{ x_2 , x_3 \}$

برای $C_1$ : $\alpha = P(X = x_1$ یا $X = x_3 \mid H_0) = 0.05 + 0.1 = 0.15$

$$\beta = P(X \notin C_1 \mid H_1) = 0.65$$

برای $C_2$ : $\alpha = P(X = x_2$ یا $X = x_3 \mid H_0) = 0.05 + 0.1 = 0.15$

$$\beta = P(X \notin C_2 \mid H_1) = 0.6$$

چون $\alpha$ برای هر دو یکسان است، با توجه به اینکه $\beta$ برای $C_2$ کمتر است، ناحیه‌ی $C_2$ را ترجیح می‌دهیم.