

רווח-סמך ובדיקת השערות עבור הפרש פרופורציות באוכלוסייה

נוסחת רווח הסמך עבור הפרש פרופורציות ($\hat{q} = 1 - \hat{p}$) ברמת סמך $1 - \alpha$:

$$(\hat{p}_1 - \hat{p}_2) - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \leq p_1 - p_2 \leq (\hat{p}_1 - \hat{p}_2) + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

מקובל לדרוש שיתקיים: $5 \leq n_1 p_1, n_1 q_1, n_2 p_2, n_2 q_2$

בדיקת השערות:

עבור מבחן השערות חד-כיווני ימני

$$\begin{cases} H_0: p_1 - p_2 = 0 \\ H_1: p_1 - p_2 > 0 \end{cases}$$

אומדן עבור הפרופורציה p הוא: $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

$$K_\alpha = Z_{1-\alpha} \cdot \sqrt{\hat{p} \hat{q} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

נדחה H_0 אם $K_\alpha > \hat{p}$

האופן דומה -

למבחן החד-צדדי השמאלי נדחה H_0 אם: $K_\alpha < \hat{p}$.

למבחן הדו-כיווני נדחה H_0 אם: $K_{\alpha/2} > \hat{p}$ או $K_{\alpha/2} < \hat{p}$.

שאלה

במדגם מקרי של 400 מבוגרים ו-600 צעירים אשר צפו בתוכנית טלוויזיה, ציינו 100 מבוגרים ו-300 צעירים כי התוכנית מצאה חן בעיניהם.

- א. בנה רווחי סמך של 95% ושל 99%, עבור ההבדל בפרופורציות של כל המבוגרים ושל כל הצעירים אשר צפו בתוכנית והתוכנית מצאה חן ביניהם.
- ב. בדוק את ההשערה כי פרופורציית הצעירים שאהבו את התכנית גדולה מפרופורציית המבוגרים באוכלוסייה, בר"מ 1%.

פתרון:

(א) מהנתונים:

$$\hat{p}_1 = 300/600 = 0.5 \text{ (צעירים)}$$

$$\hat{p}_2 = 100/400 = 0.25 \text{ (מבוגרים)}$$

הן פרופורציות של צעירים ושל מבוגרים אשר צפו בתוכנית והתוכנית מצאה חן ביניהם בהתאמה.

נציב את הנתונים בנוסחת רו"ס ברמת הבטחון המתאימה. יש להשתמש ב- $Z_{1-\frac{\alpha}{2}}$.

הרווח ברמת סמך (רמת בטחון) של 95% הוא:

$$(Z_{0.975} = 1.96)$$

$$(0.5 - 0.25) \pm 1.96 \cdot \sqrt{\frac{0.5 \cdot 0.5}{600} + \frac{0.25 \cdot 0.75}{400}} = 0.25 \pm 0.06$$

\leq לכן אנו יכולים להיות בטוחים ב-95% שההפרש האמיתי בין הפרופורציות נמצא בין 0.19 ו-0.31.

הרווח ברמת סמך (רמת בטחון) של 99% הוא:

$$(Z_{0.995} = 2.58)$$

$$(0.5 - 0.25) \pm 2.58 \cdot \sqrt{\frac{0.5 \cdot 0.5}{600} + \frac{0.25 \cdot 0.75}{400}} = 0.25 \pm 0.08$$

\leq לכן אנו יכולים להיות בטוחים ב-99% שההפרש האמיתי בין הפרופורציות נמצא בין 0.17 ו-0.33. כצפוי, אורך רווח הסמך גדל עם גידול רמת הבטחון.

(ב). בדיקת ההשערה:

$$Z_{1-\alpha} = Z_{0.99} = 2.33 \quad ; \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{100 + 300}{400 + 600} = 0.4$$

הערך הקריטי של המבחן:

$$K = Z_{1-\alpha} \cdot \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 2.33 \cdot \sqrt{0.4 \cdot 0.6 \cdot \left(\frac{1}{400} + \frac{1}{600}\right)} = 0.074$$

השוואה של הערך הקריטי להפרש במדגם:

$$\hat{p} = \hat{p}_1 - \hat{p}_2 = 0.5 - 0.25 = 0.25 > K = 0.074$$

\leq מכאן שדוחים את H_0 , כלומר, פרופורציית הצעירים שאהבו את התוכנית גדולה מפרופורציית המבוגרים שאהבו את התוכנית באופן מובהק, בר"מ של 1%.