סטטיסטיקה – תרגול 10

כיצד בוחרים סטטיסטי מבחן?

סטטיסטי המבחן הרגיש ביותר הוא יחס הנראות. אם נצליח למצוא התאמה 1-1 לסקאלה של יחס הנראות למבחן אחר, מובטח גם לו שתהיה עוצמה מקסימלית (ע"פ למת פירסון).

לדחות	$TP - V$	טעות $FP - I$
לא לדחות	טעות FN - II	$TN - V$

הגדרות

α - סיכוי שאנו מוכנים לחות איתו בשלום לטעות מסוג 1 – P_{H_0}(דחייה)

π - עוצמה של מבחן, כלומר הסיכוי שנדחה אם האלטרנטיבה נכונה - P_{H_1}(דחייה)

β - איננו דוחים כאשר האלטרנטיבה נכונה - P_{H_0}(אי דחייה)

$p - value$ – הסיכוי לראות את התוצאה שקיבלנו או הקיצונית ממנה, תחת השערת ה-0.

פתרון תרגיל 9

שאלה 1

$$H_0: X_i \sim N(2,0.5), H_1: X_i \sim N(3,0.5)$$

נשתמש בממוצע כי יש לו עוצמה גבוהה יותר (כפונקציה מונוטונית למת פירסון), נדחה לערכים גבוהים של הממוצע.

אם זמן השליפה הממוצע הוא 2.5 ומעלה נדחה. מהו π:

$$\pi = P_{H_1}(\text{דחייה}) = P_{H_1}(\overline{X_3} > 2.5) = 1 - \Phi\left(\frac{2.5-3}{\frac{0.5}{\sqrt{3}}}\right) = 0.958$$

הסיכוי לטעות מסוג I:

$$P(I) = P_{H_0}(\text{דחייה}) = P_{H_0}(\overline{X_3} > 2.5) = 1 - \Phi\left(\frac{2.5-2}{\frac{0.5}{\sqrt{3}}}\right) = 0.041$$

הסיכוי לטעות מסוג שני הוא $1 - \pi$.

אם ידוע כי $\bar{X}_3 = 2.45$ נחליט שזהו שרת מסוג א' כי אנחנו לא נמצאים באזור הדחייה.

$$p_{value} = P_{H_0}(\overline{X_3} > 2.45) = 1 - \Phi\left(\frac{2.45-2}{\frac{0.5}{\sqrt{3}}}\right) = 0.059$$

כלומר, הסיכוי שזמן השליפה יהיה ארוך מזמן זה הוא 5.9%. מקובל שהסף הוא 5% ולכן התוצאה הזו לא מספיק בלתי סבירה כדי לדחות את השערת ה-0.

נוכל להגדיל את כמות השליפות וע"י כך לזהות בצורה טובה יותר על איזה שרת מדובר. תיקון ל-4 יגדיל את העוצמה ל-0.9772 והסיכוי לטעות מסוג 1 ירד ל-0.022.

תיקון ל-4 יגדיל את העוצמה ל-0.9772 והסיכוי לטעות מסוג 1 ירד ל-0.022.

שאלה 5

$$H_1: X_i \sim Pois(1), \qquad H_0: X \sim Pois(2)$$

יחס הנראות:

$$L(\lambda) = P(X_1 = x_1, \dots, X_n = x_n) = \underbrace{\prod_{i=1}^{n} P(X_i = x_i)}_{P(X=x)=\frac{\lambda^x e^{-\lambda}}{x!}} = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = e^{-n\lambda} \lambda^{\sum x_i} \prod \frac{1}{x_i!}$$

$$\Longrightarrow \Lambda = \frac{e^{-n\lambda_1} {\lambda_1}^{\sum x_i} \prod \frac{1}{x_i!}}{e^{-n\lambda_0} {\lambda_0}^{\sum x_i} \prod \frac{1}{x_i!}} = \underbrace{e^{n(\lambda_0 - \lambda_1)}}_{>0} \left(\frac{\lambda_1}{\lambda_0}\right)^{\sum x_i}$$

יחס הנראות יורד.
ביחס הנראות נדחה תמיד לערכים גבוהים של יחס הנראות (בהגדרה של סהרון זה הפוך). לכן נדחה לערכים נמוכים של חס השאילתות.

$$\pi = P_{H_1}\left(\sum x_i \le 3\right) = 0.43$$