

דף נוסחאות מבוא לסטטיסטיקה א+ב

סטטיסטיקה תיאורית:

משתנה בדיד:

שכיח M_o הערך עם השכיחות הגובה ביותר

$$\bar{X} = \frac{\sum x}{n} \quad \text{ממוצע}$$

$$M_d = \frac{x_{n+1}}{2} \quad \text{חציון למקרה של מספר תצפיות אי-זוגי}$$

$$M_d = \frac{\frac{x_n}{2} + \frac{x_{n+1}}{2}}{2} \quad \text{חציון למקרה של מספר תצפיות זוגי}$$

$$S^2 = \frac{\sum x^2 \cdot f}{n} - \bar{X}^2 \quad \text{שונות} \quad S^2 = \frac{\sum (x - \bar{X})^2 \cdot f}{n} \quad \text{שונות נוסחת עבודה}$$

$$S = \sqrt{S^2} \quad \text{סטיית תקן}$$

$$X_{\max} - X_{\min} \quad \text{טווח}$$

$$Q_3 - Q_1 \quad \text{טווח בין רבעוני} \quad \text{רבעון עליון פחות רבעון תחתון}$$

משתנה רציף:

שכיח M_o קודם כל יש לחשב צפיפות שזה השכיחות חלקי רוחב המחלקה מחלקת השכיח תהיה המחלקה בעלת הצפיפות הגבוה ביותר ואת השכיח נחשב לפי הנוסחה:

$$M_o = L + \frac{f'_{(i)} - f'_{(i-1)}}{f'_{(i)} - f'_{(i-1)} + f'_{(i)} - f'_{(i+1)}} \times C$$

L – גבול תחתון של מחלקת השכיח

f'_i – הצפיפות של מחלקת השכיח C – רוחב מחלקת השכיח

$f'_{(i-1)}$ – הצפיפות של המחלקה לפני מחלקת השכיח

$f'_{(i+1)}$ – הצפיפות של המחלקה אחרי מחלקת השכיח

$$\bar{X} = \frac{\sum x}{n} \quad \text{ממוצע} \quad \text{אמצע מחלקה} - X$$

חציון

נחשב שכיחות מצטברת ואז נמצא את מחלק החציון לפי המחלקה הראשונה שבה המצטברת עוברת את סה"כ התצפיות חלקי 2, לאחר מציאת מחלקת החציון נחשב לפי הנוסחה הבאה:

$$Md = L + \left(\frac{C}{f}\right) \times \left(\frac{n}{2} - F_{(i-1)}\right)$$

L – הגבול העליון של מחלקת החציון

n – גודל המדגם

C – רוחב מחלקת החציון

f – השכיחות של מחלקת החציון

$F_{(i-1)}$ – השכיחות המצטברת של המחלקה שלפני מחלקת החציון

אחוזון P

אם במקום חציון רוצים אחוזון מסוים, נמצא תחילה את מחלקת האחוזון

לפי הפעם הראשונה שמצטברת עוברת את $\frac{P \cdot n}{100}$

נשים לב שרבעון עליון הוא אחוזון 75 ותחתן 25

$$Xp = L + \left(\frac{C}{f}\right) \times \left(\frac{P \cdot n}{100} - F_{(i-1)}\right)$$

L – הגבול העליון של מחלקת האחוזון

C – רוחב מחלקת האחוזון n – גודל המדגם

f – השכיחות של מחלקת האחוזון P – אחוזון בין 0 ל 100

$F_{(i-1)}$ – השכיחות המצטברת של המחלקה שלפני מחלקת האחוזון

$$S^2 = \frac{\sum x^2 \cdot f}{n} - \bar{x}^2 \quad \text{שונוות נוסחת עבודה} \quad S^2 = \frac{\sum (x - \bar{x})^2 \cdot f}{n} \quad \text{שונוות}$$

$$S = \sqrt{S^2} \quad \text{סטיית תקן} \quad X - \text{אמצע מחלקה}$$

$$X_{\max} - X_{\min} \quad \text{טווח}$$

$$Q_3 - Q_1 \quad \text{טווח בין רבעוני} \quad \text{רבעון עליון פחות רבעון תחתון}$$

צורת התפלגות

$$Q_2 - Q_1 = Q_3 - Q_2 \quad M_o = M_d = \bar{X} \quad \text{התפלגות סימטרית}$$

$$Q_2 - Q_1 < Q_3 - Q_2 \quad M_o < M_d < \bar{X} \quad \text{התפלגות אסימטרית ימנית \textbackslash חיובית}$$

$$Q_2 - Q_1 > Q_3 - Q_2 \quad M_o > M_d > \bar{X} \quad \text{התפלגות אסימטרית שמאלית \textbackslash שלילית}$$

ציון תקן:

$$Z_x = \frac{x - \mu}{\sigma} \quad \text{הציון המקורי} \quad x - \text{ממצוע הציונים} \quad \mu - \text{סטיית התקן} \quad \sigma$$

מקדם המתאם של פירסון:

$$r = \frac{\text{cov}(x, y)}{S_x \cdot S_y} \quad \text{cov}(x, y) = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{n}$$

$$\text{cov}(x, y) = \frac{\sum x \cdot y}{n} - \bar{x} \cdot \bar{y} \quad \text{cov}(x, y) = \frac{\sum x \cdot y - n \cdot \bar{x} \cdot \bar{y}}{n}$$

$$\hat{y} = a \cdot x + b \quad \text{רגרסיה לינארית פשוטה}$$

$$a = \frac{r \cdot S_y}{S_x} \quad a = \frac{\text{Cov}(x, y)}{S_x^2} \quad a = \frac{\sum x \cdot y - n \cdot \bar{x} \cdot \bar{y}}{\sum x^2 - n \cdot \bar{x}^2} \quad \text{שיפוע הקו:}$$

$$b = \bar{x} - a \cdot \bar{y} \quad \text{החתוך של הקו}$$

$$\frac{\sum (y - \bar{y})^2}{n} = \frac{\sum (\hat{y} - \bar{y})^2}{n} + \frac{\sum (y - \hat{y})^2}{n} \quad \text{פירוק השונות:}$$

$$R^2 = r^2 \quad R^2 = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \bar{y})^2} \quad \text{אחוז השונות המוסברת}$$

קומבניטוריקה והסתברות

סידור n עצמים בשורה n!

הוצאת k כדורים מתוך n ללא החזרה $\binom{n}{k}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \text{נוסחת ההסתברות השלמה}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{נוסחת בייס להסתברות מותנית}$$

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)} \quad \text{נוסחת בייס היפוך התנאי}$$

$$P(A \cap B) = P(A) \cdot P(B) \quad \text{מאורעות בלתי תלויים}$$

$$P(A) = 1 - P(\bar{A}) \quad \text{מאורע משלים} \quad P(A \cap B) = 0 \quad \text{מאורעות זרים}$$

התפלגות בינומית

סדרה של ניסויים בלתי תלויים בעלי הסתברות קבוצה להצלחה

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

הסתברות להצלחה – p מספר הצלחות – k מספר ניסויים – n

התפלגות נורמלית

$$Z \sim N(0,1)$$

$$P(Z < z) = \Phi(z) \quad P(Z < -z) = 1 - \Phi(z) \quad P(Z > z) = 1 - \Phi(z)$$

$$X \sim N(\mu, \sigma^2) \quad P(X < x) = P\left(Z < \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

מדדי קשר

מדד "פיי"

$$\phi = \sqrt{\frac{(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}}$$

Y2	Y1	Y / X
b	a	X1
d	c	X2

מדד "למדה"

$$L_y = n - [Y \text{ שכיח}]$$

$$L_{y/x} = n - [X \text{ של הערכים על פני השכיחים של } X]$$

$$\lambda_{y/x} = \frac{L_y - L_{y/x}}{L_y}$$

מדד "אתה"

סכום ריבועי הסטיות מהממוצע הכללי L_y

$$\eta_{y/x} = \frac{L_y - L_{y/x}}{L_y}$$

סכום ריבועי הסטיות בכל קטגוריה $L_{y/x} - X$

מתאם ספירמן

ההפרש בין הדרגה X לדרגה של $Y - d_i$

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

מספר הזוגות n

$$\chi^2 = \sum_i \frac{(E_i - O_i)^2}{E_i} \quad r_c = \sqrt{\frac{\chi^2}{n \cdot (L - 1)}}$$

מתאם קרמר

מספר התצפיות n מינימום מספר הקטגוריות L

סיכום תכונות מדדי קשר

מדד	סוגי משתנים וסולמות	טווח ערכים	מאפיינים
למדא	שמי+שמי, שמי+סדר	בין 0 ל 1	מבוסס על מרחק מהשכיח
קרמר	שמי+שמי, שמי+סדר	בין 0 ל 1	מבוסס על מרחק ריבועי ממצב של אי-תלות
פי	שמי+שמי, שמי+סדר	בין 0 ל 1	מקרה פרטי של קרמר עם 2 על 2
ספירמן	סדר+סדר	בין 1 ל -1	מבוסס על דרגות
פירסון	כמותי+כמותי	בין 1 ל -1	בודק קשר לינארי
אתה	סוג כלשהוא + כמותי	בין 0 ל 1	מבוסס על שונות כללית לעומת שונות לפי תת קבוצה, ומתאים גם לקשר של קו שאינו ישר

משפט הגבול המרכזי

אם למשתנה כלשהו יש התפלגות נורמלית או אם ההתפלגות היא לא נורמלית אבל גודל המדגם מספיק גדול (מעל 30) אז לממוצע ולסכום התצפיות תהיה התפלגות נורמלית
 μ - התוחלת של המשתנה, σ^2 - היא השונות

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{התפלגות הממוצע}$$

$$\sum_{i=1}^n x_i \sim N(n \cdot \mu, n \cdot \sigma^2) \quad \text{התפלגות הסכום}$$

רווח סמך

רווח סמך לתוחלת שונות ידועה $\bar{x} \pm l$

$$l = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \quad \text{אורך רווח הסמך -}$$

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{l} \right)^2 \quad \text{גודל המדגם נדרש לרמת סמך ואורך נתונים -}$$

רווח סמך לתוחלת שונות לא ידועה

$$l = t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{\hat{s}}{\sqrt{n}} \quad \text{אורך רווח הסמך -}$$

$$\hat{s} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \quad \text{אומד חסר הטיה לסטיית התקן}$$

$$\hat{p} = \frac{\text{\# הצלחות}}{n} \quad \hat{p} \pm l \quad \text{רווח סמך לפרופורציה}$$

$$l = Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} \quad \text{אורך רווח הסמך -}$$

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}}}{4 \cdot l} \right)^2 \quad \text{גודל המדגם נדרש לרמת סמך ואורך נתונים -}$$

$$(\bar{x}_1 - \bar{x}_2) \pm l \quad \text{רווח סמך להפרש תוחלות אוכלוסיות בלתי תלויות}$$

$$l = Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{אורך רווח הסמך שונות ידועה}$$

$$l = t_{n_1+n_2-2, 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}_1^2}{n_1} + \frac{\hat{s}_2^2}{n_2}} \quad \text{אורך רווח הסמך שהשוניות לא ידועות}$$

$$\hat{s}_1 = \sqrt{\frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1 - 1}} \quad \hat{s}_2 = \sqrt{\frac{\sum_{i=1}^{n_2} (x_i - \bar{x})^2}{n_2 - 1}}$$

$$\bar{d} \pm l \quad \text{רווח סמך להפרש תוחלות מדגם מזווג}$$

$$\bar{d} = \frac{\sum d_i}{n} \quad \hat{s}_d = \frac{\sum (d_i - \bar{d})^2}{n - 1} \quad l = t_{n-1, 1-\frac{\alpha}{2}} \cdot \frac{\hat{s}_d}{\sqrt{n}} \quad \text{אורך רווח הסמך -}$$

d_i – ההפרש מדדיה בין שני המצבים \(\backslash\)זמנים עבור תצפית

רווח סמך להפרש פרופורציות $(\hat{p}_1 - \hat{p}_2) \mp l$

$$l = Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1 \cdot (1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2 \cdot (1 - \hat{p}_2)}{n_2}} \quad \text{אורך רווח הסמך -}$$

רווח סמך לשונות

$$\hat{s}^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \quad \text{גבול עליון} \quad \frac{\hat{s}^2 \cdot (n - 1)}{\chi_{n-1, 1-\frac{\alpha}{2}}^2} \quad \text{גבול תחתון} \quad \frac{\hat{s}^2 \cdot (n - 1)}{\chi_{n-1, \frac{\alpha}{2}}^2}$$

בדיקת השערות

תמיד נכון: $p_{\text{value}} > \alpha \rightarrow H_0$ $p_{\text{value}} \leq \alpha \rightarrow H_1$

בדיקה השערות על תוחלת שונות ידועה

$$\begin{array}{ll} H_0: \mu = \mu_0 & C_{\text{ciriti}} = \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \\ H_1: \mu < \mu_0 & Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad p_{\text{value}} = \Phi(Z) \end{array}$$

$$\begin{array}{ll} H_0: \mu = \mu_0 & C_{\text{ciriti}} = \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \\ H_1: \mu > \mu_0 & Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad p_{\text{value}} = 1 - \Phi(Z) \end{array}$$

$$\begin{array}{ll} H_0: \mu = \mu_0 & C_1 = \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \\ H_1: \mu \neq \mu_0 & C_2 = \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \end{array} \quad Z = \frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \quad p_{\text{value}} = 2 \cdot (1 - \Phi(Z))$$

בדיקת השערות על תוחלת שונות לא ידועה

$$\begin{array}{lll} H_0: \mu = \mu_0 & C_{\text{criti}} = \mu_0 - t_{n-1, 1-\alpha} \cdot \frac{\hat{S}}{\sqrt{n}} & \bar{x} \leq C_{\text{criti}} \rightarrow H_1 \\ H_1: \mu < \mu_0 \end{array}$$

$$\begin{array}{lll} H_0: \mu = \mu_0 & C_{\text{criti}} = \mu_0 + t_{n-1, 1-\alpha} \cdot \frac{\hat{S}}{\sqrt{n}} & \bar{x} \geq C_{\text{criti}} \rightarrow H_1 \\ H_1: \mu > \mu_0 \end{array}$$

$$\begin{array}{lll} H_0: \mu = \mu_0 & C_1 = \mu_0 - t_{n-1, 1-\alpha/2} \cdot \frac{\hat{S}}{\sqrt{n}} & \bar{x} \leq C_1 \rightarrow H_1 \end{array}$$

$$\begin{array}{lll} H_1: \mu \neq \mu_0 & C_2 = \mu_0 + t_{n-1, 1-\alpha/2} \cdot \frac{\hat{S}}{\sqrt{n}} & \bar{x} \geq C_2 \rightarrow H_1 \end{array}$$

בדיקת השערות על פרופורציה

$$\begin{array}{lll} H_0: P = P_0 & H_1: P < P_0 & Z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} \quad p_{\text{value}} = \Phi(Z) \\ C_{\text{criti}} = P_0 - Z_{1-\alpha} \cdot \sqrt{\frac{P_0 \cdot (1 - P_0)}{n}} & & \hat{p} \leq C_{\text{criti}} \rightarrow H_1 \end{array}$$

$$\begin{array}{lll} H_0: P = P_0 & H_1: P > P_0 & Z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} \quad p_{\text{value}} = 1 - \Phi(Z) \\ C_{\text{criti}} = P_0 + Z_{1-\alpha} \cdot \sqrt{\frac{P_0 \cdot (1 - P_0)}{n}} & & \hat{p} \geq C_{\text{criti}} \rightarrow H_1 \end{array}$$

$$\begin{array}{lll} H_0: P = P_0 & H_1: P \neq P_0 & p_{\text{value}} = 2 \cdot (1 - \Phi(Z)) \\ C_{1,2} = P_0 \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{P_0 \cdot (1 - P_0)}{n}} & \begin{array}{l} \hat{p} \geq C_1 \rightarrow H_1 \\ \hat{p} \leq C_2 \rightarrow H_1 \end{array} & Z = \frac{|\hat{p} - P_0|}{\sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}} \end{array}$$

בדיקת השערות על הפרש תוחלות אוכלוסיות בלתי תלויות שונות ידועות

$$H_0: \mu_1 - \mu_2 = a \quad H_1: \mu_1 - \mu_2 < a \quad Z = \frac{a - (\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad p_{\text{value}} = \Phi(Z)$$

$$C_{\text{criti}} = a - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \bar{x}_1 - \bar{x}_2 \leq C_{\text{criti}} \rightarrow H_1$$

$$H_0: \mu_1 - \mu_2 = a \quad H_1: \mu_1 - \mu_2 > a \quad Z = \frac{a - (\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad p_{\text{value}} = 1 - \Phi(Z)$$

$$C_{\text{criti}} = a + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \bar{x}_1 - \bar{x}_2 \geq C_{\text{criti}} \rightarrow H_1$$

$$H_0: \mu_1 - \mu_2 = a \quad H_1: \mu_1 - \mu_2 \neq a \quad Z = \frac{a - (|\bar{x}_1 - \bar{x}_2|)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$C_{1,2} = a \mp Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \begin{array}{l} \bar{x}_1 - \bar{x}_2 \leq C_1 \rightarrow H_1 \\ \bar{x}_1 - \bar{x}_2 \geq C_2 \rightarrow H_1 \end{array} \quad p_{\text{value}} = 2 \cdot (1 - \Phi(Z))$$

במקרה שהשונות לא ידועות

נשתמש בשונות המצורפת של המדגם ובהתפלגות טי במקום נורמלית

למשל במקרה של קטן מ

$$H_0: \mu_1 - \mu_2 = a \quad H_1: \mu_1 - \mu_2 < a$$

$$\hat{S}_p^2 = \frac{(n_1 - 1) \cdot \hat{S}_1^2 + (n_2 - 1) \cdot \hat{S}_2^2}{n_1 + n_2 - 2} \quad \bar{x}_1 - \bar{x}_2 \leq C_{\text{criti}} \rightarrow H_1$$

$$C_{\text{criti}} = a - t_{n_1+n_2-2, 1-\alpha} \cdot \sqrt{\frac{\hat{S}_p^2}{n_1} + \frac{\hat{S}_p^2}{n_2}}$$

בדיקת השערות על הפרש תוחלות אוכלוסיות לא ב"ת מדגם מזווג

$$\bar{d} = \frac{\sum d_i}{n} \quad \hat{s}_d = \frac{\sum (d_i - \bar{d})^2}{n-1} \quad d_i - \text{הפרש בין שני המצבים עבור תצפית}$$

$$H_0: \mu_d = a \quad H_1: \mu_d < a \quad C_{\text{criti}} = a - t_{n-1, 1-\alpha} \cdot \frac{\hat{s}_d}{\sqrt{n}} \quad \bar{d} \leq C_{\text{criti}} \rightarrow H_1$$

$$H_0: \mu_d = a \quad H_1: \mu_d > a \quad C_{\text{criti}} = a + t_{n-1, 1-\alpha} \cdot \frac{\hat{s}_d}{\sqrt{n}} \quad \bar{d} \geq C_{\text{criti}} \rightarrow H_1$$

$$H_0: \mu_d = a \quad H_1: \mu_d \neq a \quad C_{1,2} = a \mp t_{n-1, 1-\alpha} \cdot \frac{\hat{s}_d}{\sqrt{n}} \quad \begin{aligned} \bar{d} \leq C_1 &\rightarrow H_1 \\ \bar{d} \geq C_2 &\rightarrow H_1 \end{aligned}$$

בדיקת השערות על הפרש פרופורציות

$$H_0: P_1 - P_2 = a \quad H_1: P_1 - P_2 < a \quad Z = \frac{a - (\hat{p}_1 - \bar{\bar{p}}_2)}{S_{\hat{p}}}$$

$$C_{\text{criti}} = a - Z_{1-\alpha} \cdot S_{\hat{p}} \quad \hat{p}_1 - \hat{p}_2 \leq C_{\text{criti}} \rightarrow H_1 \quad p_{\text{value}} = \Phi(Z)$$

$$H_0: P_1 - P_2 = a \quad H_1: P_1 - P_2 > a \quad Z = \frac{a - (\hat{p}_1 - \bar{\bar{p}}_2)}{S_{\hat{p}}}$$

$$C_{\text{criti}} = a + Z_{1-\alpha} \cdot S_{\hat{p}} \quad \hat{p}_1 - \hat{p}_2 \geq C_{\text{criti}} \rightarrow H_1 \quad p_{\text{value}} = 1 - \Phi(Z)$$

$$H_0: P_1 - P_2 = a \quad H_1: P_1 - P_2 \neq a \quad Z = \frac{a - (|\hat{p}_1 - \bar{\bar{p}}_2|)}{S_{\hat{p}}}$$

$$C_1 = a + Z_{1-\frac{\alpha}{2}} \cdot S_{\hat{p}} \quad \hat{p}_1 - \hat{p}_2 \geq C_1 \rightarrow H_1 \quad p_{\text{value}} = 2 \cdot (1 - \Phi(Z))$$

$$C_2 = a - Z_{1-\frac{\alpha}{2}} \cdot S_{\hat{p}} \quad \hat{p}_1 - \hat{p}_2 \leq C_2 \rightarrow H_1$$

חישוב טעות מסוג שני

בדיקת השערות על תוחלת

$$H_0: \mu = \mu_0 \quad H_1: \mu = \mu_1$$

$$\mu_1 > \mu_0 \quad \beta = P(\bar{X} < C_{\text{criti}}) = \Phi\left(\frac{C_{\text{criti}} - \mu_1}{\sigma/\sqrt{n}}\right)$$

$$\mu_1 < \mu_0 \quad \beta = P(\bar{X} > C_{\text{criti}}) = 1 - \Phi\left(\frac{C_{\text{criti}} - \mu_1}{\sigma/\sqrt{n}}\right)$$

בדיקת השערות על פרופורציה

$$H_0: P = P_0 \quad H_1: P = P_1$$

$$P_1 > P_0 \quad \beta = P(\hat{P} < C_{\text{criti}}) = \Phi\left(\frac{C_{\text{criti}} - P_1}{\sqrt{\frac{P_1 \cdot (1 - P_1)}{n}}}\right)$$

$$P_1 < P_0 \quad \beta = P(\hat{P} > C_{\text{criti}}) = 1 - \Phi\left(\frac{C_{\text{criti}} - P_1}{\sqrt{\frac{P_1 \cdot (1 - P_1)}{n}}}\right)$$

מבחן חי בריבוע לאי תלות

$$\chi^2_{\text{sample}} = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad \chi^2_{\text{criti}} = \chi^2_{(r-1) \cdot (c-1), 1-\alpha} \quad \chi^2_{\text{sample}} > \chi^2_{\text{criti}} \rightarrow H_1$$

מבחן חי בריבוע לטיב התאמה

$$\chi^2_{\text{sample}} = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad \chi^2_{\text{criti}} = \chi^2_{(c-1), 1-\alpha} \quad \chi^2_{\text{sample}} > \chi^2_{\text{criti}} \rightarrow H_1$$

מבחנים א-פרמטרים

מבחן ויילקוקסון למדגמים בלתי תלויים מאן ויטני

סכום הדרגות הקבוצה הראשונה W_1 סכום הדרגות הקבוצה השניה W_2

$$U_1 = W_1 - \frac{n_1 \cdot (n_1 + 1)}{2} \quad U_2 = W_2 - \frac{n_2 \cdot (n_2 + 1)}{2} \quad U = \min(U_1, U_2)$$

מטבלה מיוחדת U_c $U < U_c \rightarrow H1$

מבחן פישר

מאחדים את שתי הקבוצות ואז מחשבים לפי התפלגות היפר גאומטרית

$$X \sim \text{HG}(N, D, n) \quad P(X = k) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-k}}{\binom{N}{n}}$$

מספר תצפיות קבוצת הבדיקה n מספר הצלחות הכולל D מספר התצפיות הכולל N

מבחן ויילקוקסון למדגמים מזווגים

מחשבים את ההפרשים ואז מדרגים לפי ערך מוחלט וסוכמים לכל קבוצה

סכום הדרגות של ההפרשים החיוביים T^+

סכום הדרגות של ההפרשים השליליים T^-

$T = \min(T^+, T^-)$ מטבלה מיוחדת T_c $T > T_c \rightarrow H1$

מקנמר

$$\chi^2_{\text{sample}} = \frac{(B - C)^2}{B + C} \quad \chi^2_{\text{criti}} = \chi^2_{1, 1-\alpha} \quad \chi^2_{\text{sample}} > \chi^2_{\text{criti}} \rightarrow H1$$

B	
	C