

דף נוסחאות סטטיסטיקה א

סטטיסטיקה תיאורית:

משתנה בדיד:

שכיח M_0 הערך עם השכיחות הגובה ביותר

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

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

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

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

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

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

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

משתנה רציף:

שכיח M_0 קודם כל יש לחשב צפיפות שזה השכיחות חלקי רוחב המחלקה

מחלקת השכיח תהיה המחלקה בעלת הצפיפות הגבוה ביותר ואת השכיח

נחשב לפי הנוסחה:

$$M_0 = L + \frac{f'_i - f'_{(i-1)}}{f'_i - f'_{(i-1)} + f'_i - f'_{(i+1)}} \times C$$

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

n – גודל המדגם C – רוחב מחלקת האחוזון

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

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

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

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

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

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

צורת התפלגות

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

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

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

ציון תקן:

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

מקדם המתאם של פירסון:

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

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

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

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

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

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

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

קומבניטוריקה והסתברות

סידור n עצמים בשורה n!

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

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

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

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

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

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

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"

$$L_y = n - [Y \text{ שכיח}]$$

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

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

מדד "אתה"

סכום ריבועי הסטיות מהמוצע 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

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