

# סטטיסטיקה והסתברות

פרק 36 - מבחני חי בריבוע

תוכן העניינים

1. מבחן לאי תלות ..... 1

## מבחן חי בריבוע לאי תלות בין משתנים – רקע

מבחן לאי תלות מטרתו לבדוק האם קיים קשר בין שני משתנים. שני המשתנים שנבדקים צריכים להיות מחולקים למספר קטגוריות.

### מבנה המבחן:

#### השערות:

אין תלות בין המשתנים  $H_0$ .

יש תלות בין המשתנים  $H_1$ .

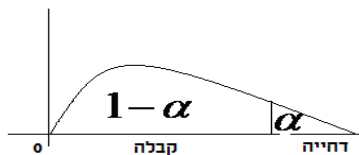
#### כלל הכרעה:

הערך הקריטי נקבע על סמך התפלגות חי בריבוע. התפלגות זו היא אסימטרית חיובית ותלויה בדרגות החופש  $d.f = (r-1)(c-1)$ . כאשר:  $r$  - מספר הקטגוריות של המשתנה שבשורות.  $c$  - מספר הקטגוריות של המשתנה שבעמודות.

הערך הקריטי הוא:  $\chi^2_{1-\alpha, (r-1)(c-1)}$ , כלומר האחוזון ה- $1-\alpha$  בהתפלגות חי בריבוע שדרגות החופש הן  $(r-1)(c-1)$ . אם  $\chi^2 > \chi^2_{1-\alpha, (r-1)(c-1)}$  אז דוחים את השערת האפס.

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

כאשר:



$O_i$  - השכיחות נצפית במדגם בתא  $i$ .

$E_i$  - שכיחות צפויה במדגם בתא  $i$  בהנחת השערת האפס.

$$E_i = \frac{f(x) \cdot f(y)}{n}$$

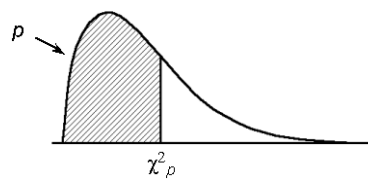
**הערה:**

תנאי כדי לבצע את המבחן הוא  $E_i \geq 5$  לכל  $i$ . במידה ותנאי זה לא מתקיים יש אפשרות לאחד קטגוריות סמוכות עד שהתנאי יתקיים.  
 תנאי חלופי: אין  $E$  קטן מ-1 וגם אין ביותר מ 20% מהתאים  $E$  קטן מ-5.

**דוגמה (הפתרון בהקלטה):**

האם יש תלות בין המגדר לבין דעה מסוימת?  
 יש לבדוק ברמת מובהקות של 5% על סמך תוצאות הסקר:

| המגדר / דעה | בעד | נגד | נמנע | סה"כ |
|-------------|-----|-----|------|------|
| גברים       | 50  | 40  | 10   |      |
| נשים        | 20  | 60  | 20   |      |
| סה"כ        |     |     |      |      |

טבלת התפלגות חי-בריבוע – ערכי החלוקה  $\chi^2_p$ 

| df | $p$      |         |         |         |        |       |       |      |      |      |      |      |      |
|----|----------|---------|---------|---------|--------|-------|-------|------|------|------|------|------|------|
|    | .005     | .01     | .025    | .05     | .10    | .25   | .50   | .75  | .90  | .95  | .975 | .99  | .995 |
| 1  | 0.004393 | 0.00157 | 0.00982 | 0.02393 | 0.0158 | 0.102 | 0.455 | 1.32 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 |
| 2  | 0.0100   | 0.0201  | 0.0506  | 0.103   | 0.211  | 0.575 | 1.39  | 2.77 | 4.61 | 5.99 | 7.38 | 9.21 | 10.6 |
| 3  | 0.0717   | 0.115   | 0.216   | 0.352   | 0.584  | 1.21  | 2.37  | 4.11 | 6.25 | 7.81 | 9.35 | 11.3 | 12.8 |
| 4  | 0.207    | 0.297   | 0.484   | 0.711   | 1.06   | 1.92  | 3.36  | 5.39 | 7.78 | 9.49 | 11.1 | 13.3 | 14.9 |
| 5  | 0.412    | 0.554   | 0.831   | 1.15    | 1.61   | 2.67  | 4.35  | 6.63 | 9.24 | 11.1 | 12.8 | 15.1 | 16.7 |
| 6  | 0.676    | 0.872   | 1.24    | 1.64    | 2.20   | 3.45  | 5.35  | 7.84 | 10.6 | 12.6 | 14.4 | 16.8 | 18.5 |
| 7  | 0.989    | 1.24    | 1.69    | 2.17    | 2.83   | 4.25  | 6.35  | 9.04 | 12.0 | 14.1 | 16.0 | 18.5 | 20.3 |
| 8  | 1.34     | 1.65    | 2.18    | 2.73    | 3.49   | 5.07  | 7.34  | 10.2 | 13.4 | 15.5 | 17.5 | 20.1 | 22.0 |
| 9  | 1.73     | 2.09    | 2.70    | 3.33    | 4.17   | 5.90  | 8.34  | 11.4 | 14.7 | 16.9 | 19.0 | 21.7 | 23.6 |
| 10 | 2.16     | 2.56    | 3.25    | 3.94    | 4.87   | 6.74  | 9.34  | 12.5 | 16.0 | 18.3 | 20.5 | 23.2 | 25.2 |
| 11 | 2.60     | 3.05    | 3.82    | 4.57    | 5.58   | 7.58  | 10.3  | 13.7 | 17.3 | 19.7 | 21.9 | 24.7 | 26.8 |
| 12 | 3.07     | 3.57    | 4.40    | 5.23    | 6.30   | 8.44  | 11.3  | 14.8 | 18.5 | 21.0 | 23.3 | 26.2 | 28.3 |
| 13 | 3.57     | 4.11    | 5.01    | 5.89    | 7.04   | 9.30  | 12.3  | 16.0 | 19.8 | 22.4 | 24.7 | 27.7 | 29.8 |
| 14 | 4.07     | 4.66    | 5.63    | 6.57    | 7.79   | 10.2  | 13.3  | 17.1 | 21.1 | 23.7 | 26.1 | 29.1 | 31.3 |
| 15 | 4.60     | 5.23    | 6.26    | 7.26    | 8.55   | 11.0  | 14.3  | 18.2 | 22.3 | 25.0 | 27.5 | 30.6 | 32.8 |
| 16 | 5.14     | 5.81    | 6.91    | 7.96    | 9.31   | 11.9  | 15.3  | 19.4 | 23.5 | 26.3 | 28.8 | 32.0 | 34.3 |
| 17 | 5.70     | 6.41    | 7.56    | 8.67    | 10.1   | 12.8  | 16.3  | 20.5 | 24.8 | 27.6 | 30.2 | 33.4 | 35.7 |
| 18 | 6.26     | 7.01    | 8.23    | 9.39    | 10.9   | 13.7  | 17.3  | 21.6 | 26.0 | 28.9 | 31.5 | 34.8 | 37.2 |
| 19 | 6.84     | 7.63    | 8.91    | 10.1    | 11.7   | 14.6  | 18.3  | 22.7 | 27.2 | 30.1 | 32.9 | 36.2 | 38.6 |
| 20 | 7.43     | 8.26    | 9.59    | 10.9    | 12.4   | 15.5  | 19.3  | 23.8 | 28.4 | 31.4 | 34.2 | 37.6 | 40.0 |
| 21 | 8.03     | 8.90    | 10.3    | 11.6    | 13.2   | 16.3  | 20.3  | 24.9 | 29.6 | 32.7 | 35.5 | 38.9 | 41.4 |
| 22 | 8.64     | 9.54    | 11.0    | 12.3    | 14.0   | 17.2  | 21.3  | 26.0 | 30.8 | 33.9 | 36.8 | 40.3 | 42.8 |
| 23 | 9.26     | 10.2    | 11.7    | 13.1    | 14.8   | 18.1  | 22.3  | 27.1 | 32.0 | 35.2 | 38.1 | 41.6 | 44.2 |
| 24 | 9.89     | 10.9    | 12.4    | 13.8    | 15.7   | 19.0  | 23.3  | 28.2 | 33.2 | 36.4 | 39.4 | 43.0 | 45.6 |
| 25 | 10.5     | 11.5    | 13.1    | 14.6    | 16.5   | 19.9  | 24.3  | 29.3 | 34.4 | 37.7 | 40.6 | 44.3 | 46.9 |
| 26 | 11.2     | 12.2    | 13.8    | 15.4    | 17.3   | 20.8  | 25.3  | 30.4 | 35.6 | 38.9 | 41.9 | 45.6 | 48.3 |
| 27 | 11.8     | 12.9    | 14.6    | 16.2    | 18.1   | 21.7  | 26.3  | 31.5 | 36.7 | 40.1 | 43.2 | 47.0 | 49.6 |
| 28 | 12.5     | 13.6    | 15.3    | 16.9    | 18.9   | 22.7  | 27.3  | 32.6 | 37.9 | 41.3 | 44.5 | 48.3 | 51.0 |
| 29 | 13.1     | 14.3    | 16.0    | 17.7    | 19.8   | 23.6  | 28.3  | 33.7 | 39.1 | 42.6 | 45.7 | 49.6 | 52.3 |
| 30 | 13.8     | 15.0    | 16.8    | 18.5    | 20.6   | 24.5  | 29.3  | 34.8 | 40.3 | 43.8 | 47.0 | 50.9 | 53.7 |

## שאלות

- (1) נבדקה התלות בין גודל הארגון לבין שביעות הרצון של העובדים. להלן התוצאות:

| שביעות רצון<br>גודל המפעל | נמוכה | בינונית | גבוהה | סה"כ |
|---------------------------|-------|---------|-------|------|
| גדול                      | 182   | 203     | 215   | 600  |
| קטן                       | 154   | 110     | 136   | 400  |
| סה"כ                      | 336   | 313     | 351   | 1000 |

מה המסקנה ברמת מובהקות של 2.5%?

- (2) מפעל עובד בשלוש משמרות. להלן מספר המוצרים הפגומים והתקינים בכל אחת מן המשמרות לפי מדגם שנעשה:

| לילה | ערב | יום |        |
|------|-----|-----|--------|
| 70   | 60  | 50  | פגומים |
| 800  | 700 | 600 | תקינים |

האם יש הבדל בין שיעורי הפגומים במשמרות השונות? הסיקו עבור רמת מובהקות  $\alpha = 0.05$ .

- (3) נדגמו 50 מוצרים ממפעל מסוים מתוך 30 מוצרים שיוצרו ביום 17 נבחרו לייצוא מתוך המוצרים שיוצרו בלילה 10 נבחרו לייצוא. האם יש קשר בין היות מוצר לייצוא למועד שבו הוא יוצר? בדקו ברמת בטחון של 95%.

- (4) במטרה לבדוק האם השתנו דפוסי ההצבעה למפלגות השונות בין שבוע שעבר לשבוע נלקחו שני סקרים אחד מהשבוע שעבר והאחר מהשבוע. להלן דפוסי ההצבעה שהתקבלו בסקרים אלה.

א. מהי רמת המובהקות המינמלית עבורה ניתן להחליט שהשתנו דפוסי ההצבעה משבוע שעבר לשבוע באופן מובהק?

ב. כיצד הייתה התשובה לסעיף א משתנה אם כל השכיחויות בטבלה של תוצאות המדגם היו מוכפלות פי 2?

ג. בנו רווח סמך לשיעור המצביעים למפלגה א השבוע ברמת סמך של 95%.

|           | מפלגה א | מפלגה ב | מפלגות אחרות | סה"כ |
|-----------|---------|---------|--------------|------|
| שבוע שעבר |         |         |              |      |
| השבוע     | 143     |         | 253          | 550  |
| סה"כ      | 243     | 314     |              | 1050 |

- (5) סטודנט קיבל בבדיקת השערות ערך  $\chi^2$  (chi-square) השוו לאפס. הסטודנט הסיק כי לא קיימת תלות בין שני המשתנים שבדק, בכל רמת מובהקות. נכון / לא נכון? נמקו.

- (6) להלן טבלת O של שני משתנים שהתקבל במדגם כלשהו :

| $f(x)$ | $Y_4$ | $Y_3$ | $Y_2$ | $Y_1$ |        |
|--------|-------|-------|-------|-------|--------|
| 200    |       |       |       |       | $X_1$  |
| 200    |       |       |       |       | $X_2$  |
|        | 160   | 120   | 60    | 60    | $f(y)$ |

מה צריכות להיות השכיחויות בתוך הטבלה כדי שמובהקות התוצאה (PV) תהיה 100%?

### תשובות סופיות

- (1) נסיק שיש קשר בין גודל הארגון לשביעות הרצון של העובדים.
- (2) נסיק שאין הבדל מובהק בין שיעור הפגומים במשמרות השונות.
- (3) נסיק שאין קשר בין היות מוצא לייצוא למועד שבו הוא יוצר.
- (4) א. 10% ב. קטן ג. (0.223, 0.297)
- (5) נכון
- (6) להלן טבלה :

| $f(x)$ | $Y_4$ | $Y_3$ | $Y_2$ | $Y_1$ |        |
|--------|-------|-------|-------|-------|--------|
| 200    | 80    | 60    | 30    | 30    | $X_1$  |
| 200    | -8    | 60    | 30    | 30    | $X_2$  |
| 400    | 160   | 120   | 60    | 60    | $f(y)$ |