

מבוא לסטטיסטיקה והסתברות

פרק 22 - רוח סマー להפרש תוחלות (ממוצעים) במדגים בלתי תלויים

תוכן העניינים

1. כישONOיות האוכלוסייה לא ידועות ובהנחת שוויון שONOיות
2. כישONOיות האוכלוסייה לא ידועות והמדגים גדולים.
- 3.....

כשוניות האוכלוסייה לא ידועות ובהנחה שוויון שונוויות:

רקע:

המטרה היא לאמוד את פער התוחלות: $\mu_2 - \mu_1$, כולם ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רוח סמך:

$$\cdot \sigma^2_1 = \sigma^2_2 . 1$$

$$\cdot X_1, X_2 \sim N . 2$$

3. מוגדים בلتוי תלויים.

השונות המשוקלلت: כיון שאנו מניחים שבין שתי האוכלוסיות השונוויות שוות אנו אומדים את השונות הזו על ידי שקולול שתי השונוויות של שני המוגדים על ידי

$$\text{הנוסחה הבאה: } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

דרגות החופש: $d.f = n_1 + n_2 - 2$

$$\text{רוח סמך: } (\bar{x}_1 - \bar{x}_2) \pm t_{1-\frac{\alpha}{2}}^{n_1+n_2-2} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

אם הערך אפס נופל בגבולות רוח סמך נגיד שבביטחון של $\alpha - 1$, לא קיים הבדל בין התוחלות.

דוגמה (פתרון בהקלטה):

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לברק קנדל שבע מבחינות הכנסה הממוצעת של אקדמאים. להלן תוצאות המבחן שנעשה:

| | ברק קנדל | תל אביב | באר שבע |
|------------------------------|----------|---------|---------|
| מספר האקדמאים | | | |
| 9500 | 11,000 | 20 | 10 |
| סטיית התקן של הכנסות אקדמאים | 250 | 200 | |

בנו רוח סמך ברמת ביטחון של 90% להפרש תוחלות הכנסה בשני האזוריים. הניחו שהשכר מתפלג נורמלית עם אותה שוננות בכל אחד מהאזורים.

שאלות:

- 1) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו ל מבחן IQ.
להלן תוצאות המדגם:

| ארה"ב | ישראל | המדינה |
|---------|---------|---------------------|
| 15 | 15 | גודל המדגם |
| 1470 | 1560 | סכום הציונים |
| 147,560 | 165,390 | סכום ריבועי הציונים |

מצאו רוח סמך ברמת סמך של 95% לסתיטה בין ממוצע הציונים בישראל לממוצע הציונים בארה"ב. רשמו את כל הנקודות הדרושות לצורך פתרון התרגילים.

- 2) להלן 4 תוצאות על משתנה X שמתפלג: $N(\mu_x, \sigma^2)$, ומשתנה Y שמתפלג: $N(\mu_y, \sigma^2)$.

| | | | | |
|---|----|----|----|----|
| X | 22 | 20 | 21 | 25 |
| Y | 18 | 25 | 17 | 12 |

חשבו רוח סמך ל- $\mu_x - \mu_y$ ברמת הסמך 90%, בהנחה שני המדגמים בלתי תלויים.

תשובות סופיות:

- 1) הנקודות:
 1. השונות שווה.
 2. שהציונים מתפלגים נורמלית.
 3. המדגמים אינם תלויים זה בזה.
 $-5.52 < \mu_1 - \mu_2 < 17.52$
 $-9.6 < \mu_y - \mu_x < 1.6$ (2)

כשונותיות האוכלוסייה לא ידועות והמוגדים גדולים:

רקע:

המטרה היא לאמוד את פער התוחולות: $\mu_2 - \mu_1$, כלומר ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רוחח הסמך:

1. מוגדים גדולים.

2. מוגדים בלתי תלויים.

$$\text{רווח סמך} : Z_{\frac{1-\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} . (\bar{x}_1 - \bar{x}_2)$$

אם הערך אפס נופל בגבולות רוחח הסמך נגיד שבביטחון של $\alpha-1$, לא קיים הבדל בין התוחולות.

דוגמה (פתרון בהקלטה):

יבואן רכב מעוניין להשוות בין רכבים יפנים לעומת אמריקאים מבחינה תצורך הדלק שלהם. נדגמו 40 רכבים מכל סוג. לכל רכב נבדק כמות הק"מ שהמכונית נסעה עבור 1 ליטר דלק בmahירות של 100 קמ"ש. התוצאות שהתקבלו במדגם הם: ברכבים היפנים ממוצע 12.5 עם סטיית תקן 1.5 וברכבים האמריקאים המוצע 11.3 עם סטיית תקן של 1.4.

- א. בנו רוחח סמך ברמת סמך של 95% לפער הממוצעים של ציריכת הדלק של מכוניות יפניות לעומת אמריקאיות.
- ב. האם על סמך רוחח הסמך ניתן לקבוע שקיים הבדל בין שני סוגי הרכבים מבחינה ציריכת הדלק?

שאלות:

- 1)** במדגם של 45 תלמידים הלומדים משפטים התקבל ששבועם לומדים בממוצע 4 שעות עם סטטיסטית תקן 2 שעות מעבר לשעות הלימוד בכיתה. במדגם של 55 תלמידים הלומדים הנדסה התקבל ששבועם לומדים בממוצע 10 שעות עם סטטיסטית תקן של 3 שעות.
- א. אמדדו את הפרש ממוצעי שעות הלימידה של סטודנטים למשפטים לעומת סטודנטים להנדסה ברמת בטיחון של 90%.
- ב. האם קיים הבדל בין סטודנטים למשפטים ולסטודנטים להנדסה מבחינת ההשकעה שלהם מעבר לשעות הלימוד בכיתה?
- 2)** מחקר טוען שאנשים החיים במרכז הארץ צופים בממוצע בטלוויזיה יותר מאשר לא חיים במרכז. נדגמו 100 אנשים מהמרכז ו-107 אנשים לא מהמרכז. אנשים אלו נשאלו כמה שעות ביום הם נוהגים לצפות בטלוויזיה. במדגם של מרכז הארץ התקבל ממוצע 2.7 שעות וסטטיסטית תקן של 0.7 שעות. במדגם של מוחץ למרכז הארץ התקבל ממוצע 1.8 שעות וסטטיסטית תקן של 1.1 שעות. מצאו רוח סמך לפער בין ממוצע שעות הצפייה בטלוויזיה בין שני האזורים בביטחון של 95%.

תשובות סופיות:

- 1)** א. $\mu_1 - \mu_2 < -5.17$. ב. כן.
- 2)** כן, בביטחון של 90% סטודנטים להנדסה לומדים יותר מסטודנטים למשפטים.