תרגול 10- חלופות

על סדר היום

-חזרה על תרגול קודם

-מבחן טי מזווג ומבחן וולש

-השוואה בזוגות

דוגמה

מפעל לייצור פלדה רוצה לבדוק את יכולתה של מערכת ייצור חדשה לשפר שני מדדים חשובים למפעל :

- אחוז הפגומים במערכת .
- רמות הזיהום הנפלטות כתוצאה מן הייצור.

במפעל בצעו מחקר סימולציה בו מודלו המצב הקיים והמערכת החדשה.

כל ריצות הסימולציה הן בלתי תלויות .

בוצע CRN בעת הרצת החלופות -כלומר קיימת תלות בין החלופות.

דוגמה -המשך

כחלק מהמחקר נאספו הנתונים הבאים:

אחוז הפגומים			רמת הזיהום		
מספר ריצה	מצב קיים	מערכת חדשה	מספר ריצה	מצב קיים	מערכת חדשה
1	0.053	0.053	1	21.54	19.11
2	0.06	0.042	2	20.26	19.45
3	0.053	0.053	3	21.73	20.95
4	0.06	0.04	4	15.37	19.57
5	0.053	0.05	5	19.61	20
6	0.057	0.046	6	18.35	18.66
7	0.06	0.048	7	16.13	19.36
8	0.057	0.041	8	16.52	18.29
9	0.057	0.042	9	17.76	18.43
10	0.057	0.054	10	17.36	18.46
11	0.051	0.042	11	21.13	19.16
12	0.055	0.054	12	16.66	20.48
13	0.057	0.045	13	19.41	20.66
14	0.054	0.043	14	16.21	18.51
15	0.053	0.047	15	15.1	20.07
ממוצע	0.056	0.047	ממוצע	18.21	19.411
סטיית תקן	0.003	0.005	סטיית תקן	2.266	0.867

סעיף א':

המפעל מעוניין בדיווח תוצאות של כל מערכת בודדת עם דיוק יחסי של 0.05.

מה מספר הריצות הנדרש ברמת מובהקות של 0.20?

חישוב דיוק יחסי:

$$\alpha_i = \frac{\alpha_{total}}{Num\ of\ Comparisons}$$

$$\frac{\gamma}{1+\gamma} = 0.0476 \qquad \alpha_{total} = 0.2 \qquad \alpha_i = \frac{0.2}{4} \qquad t_{15-1,1-\frac{0.05}{2}}=2.145$$

	אחוז הפגומים	מדד הזיהום
מצב קיים	$\frac{\delta}{\bar{X}} = \frac{t_{15-1,1-\frac{0.05}{2}} * \frac{0.003}{\sqrt{15}}}{0.056}$=0.0296	✖ $\frac{\delta}{\bar{X}} = \frac{t_{15-1,1-\frac{0.05}{2}} * \frac{2.265}{\sqrt{15}}}{18.210}$=0.0688
מערכת חדשה	✖ $\frac{\delta}{\bar{X}} = \frac{t_{15-1,1-\frac{0.05}{2}} * \frac{0.005}{\sqrt{15}}}{0.047}$=0.0589	$\frac{\delta}{\bar{X}} = \frac{t_{15-1,1-\frac{0.05}{2}} * \frac{0.869}{\sqrt{15}}}{19.411}$=0.0247

סעיף א -המשך:

ניתן לראות ששתי החלופות לא עומדות בדרישת הדיוק היחסי.

נשערך את מספר הריצות הנדרש באמצעות דיוק גס:

$$n = n_0 * \left(\frac{\delta_0}{\delta_t}\right)^2 = n_0 * \left(\frac{\delta_0}{\bar{X} * \frac{\gamma}{1+\gamma}}\right)^2$$

מצב קיים:

$$n = 15 * \left(\frac{t_{15-1,1-\frac{0.05}{2}} * \frac{2.265}{\sqrt{15}}}{18.210 * 0.0476}\right)^2 = \lceil 31.416 \rceil = 32$$

מערכת חדשה:

$$n = 15 * \left(\frac{t_{15-1,1-\frac{0.05}{2}} * \frac{0.005}{\sqrt{15}}}{0.047 * 0.0476}\right)^2 = \lceil 22.981 \rceil = 23$$

השוואת חלופות: רמת מובהקות

כאשר יש מספר חלופות ומספר מדדים ,רמת המובהקות הסטטיסטית אלפא, תתחלק **במספר רווחי סמך (השוואות)** שנבצע.

- השוואה בזוגות

$$\alpha_i = \frac{\alpha_{total}}{Num\ of\ Comparisons}$$

מספר ההשוואות הכולל =מספר ההשוואות *מספר המדדים

כל השוואה בודדת תתבצע ברמת מובהקות של האלפא המחולקת.

השוואה בין חלופות: מבחן t-מזווג

נתונות שתי סדרות של תוצאות של ריצות סימולציה: $\mathbf{x_{1,j}}$, j=1...n_1 $\mathbf{x_{2,j}}$, j=1,...,n_2 כאשר $n_1 = n_2$

ניצור משתנה מקרי חדש שמייצג את ההפרש בין אותו מדד בריצות מקבילות:

$$z = x_{1,j} - x_{2,j}$$

ונחשב רווח סמך כפי שלמדנו, עבור המשתנה המקרי החדש:

$$\overline{Z}(n) \pm t_{n-1,1-\alpha/2} \frac{S_Z(n)}{\sqrt{n}}$$

הנחות-

- יש תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בכל חלופה בפני עצמה.
- ניתן להניח שוויון שונויות בין הסדרות.
- התוצאות מתפלגות נורמאלית.

נכון יותר לבצע מבחן השערות ורק לאחריו רווח סמך.

השוואה בין חלופות: מבחן t - מזווג

$$z = x_{1,j} - x_{2,j}$$

$$\overline{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad V\hat{a}r[\overline{Z}(n)] = S(n)^2 = \frac{\sum_{j=1}^{n} (Z_j - \overline{Z}(n))^2}{(n-1)}$$

מבחן השערות:

השערת האפס: ממוצע ההפרשים $\mu_0 = 0$

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

$$t_{st} = \frac{z - \mu_0}{\frac{s_z}{\sqrt{n}}}$$

$$t_{cr} = t_{1-\frac{\alpha_i}{2}, n-1}$$

$|t_{st}| > t_{cr}$ → Reject H0, there is a difference.

השוואה בין חלופות :מבחן t - מזווג

רווח הסמך:

$$\bar{Z}(n) \pm t_{n-1,1-\alpha/2} \frac{S_Z(n)}{\sqrt{n}}$$

- רווח הסמך חיובי – חלופה 1 גדולה מחלופה 2
- רווח הסמך שלילי – חלופה 2 גדולה מחלופה 1
- רווח הסמך כולל את 0 – לא ניתן להבדיל ביניהן
- במבחן יש לשים לב מה החסרתם ממה והאם נרצה מדד מקסימלי או מינימלי, ולנמק למה זה יותר/פחות טוב לבחור בחלופה זו.

השוואה בין חלופות: קירוב ע"פ Welch

נגדיר את רווח הסמך:

$$\bar{x}_1(n_1) - \bar{x}_2(n_2) \pm t_{f,1-\frac{\alpha}{2}} * \sqrt{\frac{s_{1(n_1)}^2}{n_1} + \frac{s_{2(n_2)}^2}{n_2}}$$

הנחות-

- אין תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בכל חלופה בפני עצמה.
- לא ניתן להניח שוויון שונויות בין הסדרות.
- התוצאות מתפלגות נורמאלית
- את דרגות החופש נעריך באמצעות-

$$\hat{f} = \left(\frac{\left[\frac{s_{1(n_1)}^2}{n_1} + \frac{s_{2(n_2)}^2}{n_2}\right]^2}{\frac{\left[\frac{s_{1(n_1)}^2}{n_1}\right]^2}{n_1 - 1} + \frac{\left[\frac{s_{2(n_2)}^2}{n_2}\right]^2}{n_2 - 1}} \right)$$

אם לא יוצא שלם ,נבצע קירוב לינארי

(due to Welch, 1938)

השוואה בין חלופות: CRN

נשתמש בשיטה זו בכדי להפחית שונות בעת השוואה בין חלופות (x, y).
אנו אומדים את השונות של ההפרש בין החלופות

$$\mathbb{V}(x-y) = \mathbb{V}(x) + \mathbb{V}(y) - 2 \cdot Cov(x,y)$$

המטרה היא ליצור מתאם חיובי בין החלופות:

$$Cov(x_i, y_i) > 0$$

נייצר מתאם בין ריצות מקבילות של חלופות שונות ע"י מתן זרמי מספרים אקראיים זהים לחלופות השונות בהרצות מקבילות.

הדרך שבה נוכל לשלוט זרם המספרים האקראיים בסימולציה הוא על ידי הגדרת ה-seed.

סעיף ב המשך

ברמת מובהקות של 0.10 על פי שני המדדים מצא איזו מן החלופות טובה יותר ?פרט מה ההנחות שהנחת בעת ביצוע הבדיקה.

כיוון שיש מתאם בין החלופות נרצה לבצע מבחן טי מזווג.

נבדוק שההנחות מתקיימות-

- יש תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בכל חלופה בפני עצמה.
- ניתן להניח שוויון שונויות בין הסדרות.
- התוצאות מתפלגות נורמאלית.

סעיף ב -המשך:

1. נחשב את ההפרש בין תוצאות של כל שתי הרצות מקבילות.
2. נחשב ממוצע וסטיית תקן להפרש.

אחוז הפגומים				רמת הזיהום			
מספר ריצה	מצב קיים	מערכת חדשה	הפרש	מספר ריצה	מצב קיים	מערכת חדשה	הפרש
1	0.053	0.053	0	1	21.54	19.11	2.43
2	0.06	0.042	0.018	2	20.26	19.45	0.81
3	0.053	0.053	0	3	21.73	20.95	0.78
4	0.06	0.04	0.02	4	15.37	19.57	-4.2
5	0.053	0.05	0.003	5	19.61	20	-0.39
6	0.057	0.046	0.011	6	18.35	18.66	-0.31
7	0.06	0.048	0.012	7	16.13	19.36	-3.23
8	0.057	0.041	0.016	8	16.52	18.29	-1.77
9	0.057	0.042	0.015	9	17.76	18.43	-0.67
10	0.057	0.054	0.003	10	17.36	18.46	-1.1
11	0.051	0.042	0.009	11	21.13	19.16	1.97
12	0.055	0.054	0.001	12	16.66	20.48	-3.82
13	0.057	0.045	0.012	13	19.41	20.66	-1.25
14	0.054	0.043	0.011	14	16.21	18.51	-2.3
15	0.053	0.047	0.006	15	15.1	20.07	-4.97
ממוצע			0.009	ממוצע			-1.201
סטיית תקן			0.007	סטיית תקן			2.215

סעיף ב -המשך:

3. את אלפא נחלק במספר ההשוואות שאנחנו עושים: $\alpha_i = 0.05$ $\alpha_{total} = 0.1$

4. נחשב את רווח הסמך עבור כל אחד מהמדדים:

- אחוז הפגומים: $0.009 \pm t_{1-\frac{0.05}{2},15-1} * \frac{0.007}{\sqrt{15}} = [0.00512, 0.0128]$

ניתן לראות כי 0 לא נמצא בגבולות רווח הסמך.

כיוון שרווח הסמך חיובי ואנו שואפים למזער את אחוז הפגומים ,נאמר כי עבור מדד זה החלופות שונות ברמת מובהקות של 5%, חלופה 2 עדיפה על חלופה 1.

- רמת הזיהום: $-1.201 \pm t_{1-\frac{0.05}{2},15-1} * \frac{2.215}{\sqrt{15}} = [-2.427, 0.0257]$

ניתן לראות כי 0 נמצא בגבולות רווח הסמך ולכן נאמר כי ברמת מובהקות של 5%, עבור מדד הזיהום אין הבדל בין המערכות.

דוגמא Welch

https://colab.research.google.com/drive/1FoxALcrsFrJ_Oj9UTG9dAGVfPbYcAsq3?usp=sharing

סיכום

קירוב לפי welch	t מזווג
ניתן להשתמש ב חשונים עבור כל חלופה	אותו חלשתי החלופות
אין מתאם בין ההרצות	קיים מתאם בין ההרצות
לא ניתן להשתמש אם בוצע CRN	ניתן להשתמש אם בוצע CRN

שימו לב! כי בעבודה נבצע CRN (כלומר set seed)-- נריץ את המצב הקיים ואת החלופות עם אותו הseed ונעבור להרצה הבאה עם seed אחר הפעם- אך אותו הseed למצב הקיים והחלופות גם כעת, וכך הלאה..

```
For i in range(number_of runs):

    np.random.seed(i*2+40) – any function of I
    measue1,measue2,measue3 = run_simulation()
```