תרגול 10 - השוואת חלופות

$A_1 - B_1 = C_1$

$A_2 - B_2 = C_2$

$A_3 - B_3 = C_3$

$\bar{C}$

$\bar{A}$ $\bar{B}$

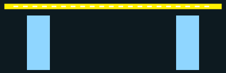
$\bar{A} - \bar{B}$

השוואת כמה חלופות דוגמא

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.

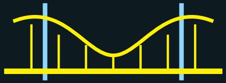
השוואת כמה חלופות דוגמא

חלופה A



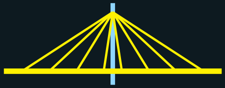
גשר קורה פשוט

חלופה B



גשר תלוי

חלופה C

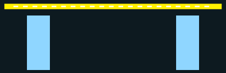


גשר כבלים

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.

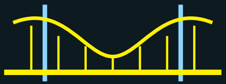
השוואת כמה חלופות דוגמא

חלופה A



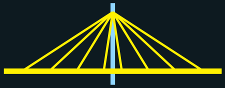
גשר קורה פשוט

חלופה B



גשר תלוי

חלופה C



גשר כבלים

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.

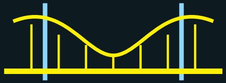
השוואת כמה חלופות דוגמא

חלופה A

גשר קורה פשוט

$I_A = [0.08, 0.16]$

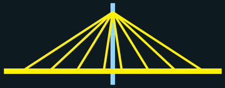
חלופה B



גשר תלוי

$I_B = [0.12, 0.20]$

חלופה C



גשר כבלים

$I_C = [0.28, 0.32]$

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.
- מכל חלופה נקבל אומד לשיעור הכשל ורווח סמך ברמת ביטחון 95%.

השוואת כמה חלופות דוגמא

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.
- מכל חלופה נקבל אומד לשיעור הכשל ורווח סמך ברמת ביטחון 95%.
- לפי האומד בלבד, חלופה A נראית הטובה ביותר.

השוואת כמה חלופות דוגמא

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.
- מכל חלופה נקבל אומד לשיעור הכשל ורווח סמך ברמת ביטחון 95%.
- לפי האומד בלבד, חלופה A נראית הטובה ביותר.

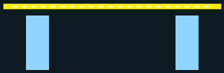
הבעיה
רווח סמך עבור כל חלופה בנפרד מבטיח:

$$\Pr(\mu_A \in \widehat{I}_A,\ \mu_B \in \widehat{I}_B,\ \mu_C \in \widehat{I}_C,\ ) \geq 0.95$$

אוניברסיטת בן-גוריון בנגב
جامعة بن غوريون في النقب
Ben-Gurion University of the Negev

השוואה כמה חלופות דוגמא

חלופה A



גשר קורה פשוט

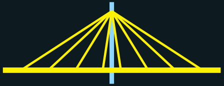
$I_A = [0.08, 0.16]$

חלופה B

גשר תלוי

$I_B = [0.12, 0.20]$

חלופה C



גשר כבלים

$I_C = [0.28, 0.32]$

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.
- מכל חלופה נקבל אומד לשיעור הכשל ורווח סמך ברמת ביטחון 95%.
- לפי האומד בלבד, חלופה A נראית הטובה ביותר.

הבעיה
רווח סמך עבור כל חלופה בנפרד מבטיח:

$$\Pr(\mu_A \in \widehat{I}_A,\ \mu_B \in \widehat{I}_B,\ \mu_C \in \widehat{I}_C,\ ) \geq 0.95$$

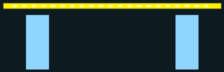
אבל השוואת חלופות שואלת: מה **ההפרש** בין הגשרים?

$$\Pr(\mu_A < \mu_B) \geq 0.95$$

אוניברסיטת בן-גוריון בנגב
جامعة بن غوريون في النقب
Ben-Gurion University of the Negev

השוואת כמה חלופות דוגמא

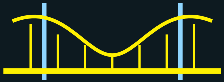
חלופה A



גשר קורה פשוט

$I_A = [0.08, 0.16]$

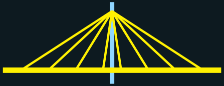
חלופה B



גשר תלוי

$I_B = [0.12, 0.20]$

חלופה C



גשר כבלים

$I_C = [0.28, 0.32]$

- ברצוננו להשוות בין שלושה גשרים לפי **שיעור הכשל** שלהם.
- עבור כל חלופה מריצים סימולציה, וחוזרים עליה מספר פעמים.
- מכל חלופה נקבל אומד לשיעור הכשל ורווח סמך ברמת ביטחון 95%.
- לפי האומד בלבד, חלופה A נראית הטובה ביותר.

הבעיה

רווח סמך עבור כל חלופה בנפרד מבטיח:

$$\Pr(\mu_A \in \widehat{I}_A,\ \mu_B \in \widehat{I}_B,\ \mu_C \in \widehat{I}_C,\ ) \geq 0.95$$

אבל השוואת חלופות שואלת: מה ההפרש בין הגשרים?

$$\Pr(\mu_A < \mu_B) \geq 0.95$$

אוניברסיטת בן-גוריון בנגב
جامعة بن غوريون في النقب
Ben-Gurion University of the Negev

השוואת כמה חלופות דוגמא

חלופה A

גשר קורה פשוט

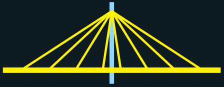
$I_A = [0.08, 0.16]$

חלופה B

גשר תלוי

$I_B = [0.12, 0.20]$

חלופה C



גשר כבלים

$I_C = [0.28, 0.32]$

השוואה	רווח סמך להפרש	מסקנה
$A - B$	$[-0.12, 0.04]$	אין מנצח מובהק
$B - C$	$[-0.20, -0.08]$	$B < C$
$A - C$	$[-0.24, -0.12]$	$A < C$

הבעיה

רווח סמך עבור כל חלופה בנפרד מבטיח:

$$\Pr(\mu_A \in \widehat{I}_A,\ \mu_B \in \widehat{I}_B,\ \mu_C \in \widehat{I}_C,\ ) \geq 0.95$$

אבל השוואת חלופות שואלת: מה ההפרש בין הגשרים?

$$\Pr(\mu_A < \mu_B) \geq 0.95$$

על סדר היום

- השוואה בין חלופות
- מבחן t מזווג עבור השוואת חלופות תלויות
- CRN - Common Random Numbers
- קירוב Welch להשוואה בין חלופות בלתי תלויות

השוואת חלופות - מוטיבציה

בתרגול הקודם שאלנו:

איך מדווחים מדדים של סימולציה עם רווחי סמך?

השוואת חלופות - מוטיבציה

בתרגול הקודם שאלנו:

איך מדווחים מדדים של סימולציה עם רווחי סמך?

עכשיו השאלה משתנה:

איך מחליטים איזו חלופה טובה יותר?

השוואת חלופות - מוטיבציה

בתרגול הקודם שאלנו:

איך מדווחים מדדים של סימולציה עם רווחי סמך?

עכשיו השאלה משתנה:

איך מחליטים איזו חלופה טובה יותר?

לדוגמה:

- מערכת קיימת מול מערכת חדשה
- שרת יחיד מול מספר שרתים
- מדיניות תור אחת מול מדיניות אחרת

השוואת חלופות - מוטיבציה

בתרגול הקודם שאלנו:

איך מדווחים מדדים של סימולציה עם רווחי סמך?

עכשיו השאלה משתנה:

איך מחליטים איזו חלופה טובה יותר?

לדוגמה:

מערכת קיימת מול מערכת חדשה

שרת יחיד מול מספר שרתים

מדיניות תור אחת מול מדיניות אחרת

הבעיה

לא מספיק לדעת כל ממוצע בנפרד.

צריך לבדוק האם ההפרש בין החלופות מובהק סטטיסטית.

דוגמה - מפעל הפלדה

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

- אחוז הפגומים במערכת
- רמות הזיהום הנפלטות כתוצאה מן הייצור

דוגמה - מפעל הפלדה

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

- אחוז הפגומים במערכת
- רמות הזיהום הנפלטות כתוצאה מן הייצור

בתרגול הקודם בדקנו כל מערכת בנפרד.

דוגמה - מפעל הפלדה

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

- אחוז הפגומים במערכת
- רמות הזיהום הנפלטות כתוצאה מן הייצור

בתרגול הקודם בדקנו כל מערכת בנפרד.

כעת נרצה להשוות ישירות בין המערכת החדשה למצב הקיים.

דוגמה - מפעל הפלדה

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

- אחוז הפגומים במערכת
- רמות הזיהום הנפלטות כתוצאה מן הייצור

בתרגול הקודם בדקנו כל מערכת בנפרד.

כעת נרצה להשוות ישירות בין המערכת החדשה למצב הקיים.

מידע חדש

בוצע CRN בעת הרצת החלופות -

כלומר קיימת תלות בין ריצות מקבילות של החלופות.

אוניברסיטת בן-גוריון בנגב
جامعة بن غوريون في النقب
Ben-Gurion University of the Negev

CRN - Common Random Numbers

נשתמש ב-CRN כדי להפחית שונות בעת השוואה בין חלופות.

CRN - Common Random Numbers

נשתמש ב-CRN כדי להפחית שונות בעת השוואה בין חלופות.

במקום להריץ את שתי החלופות עם אקראיות בלתי תלויה, נשתמש באותם זרמי מספרים אקראיים בריצות מקבילות (באמצעות אותו SEED).

CRN - Common Random Numbers

נשתמש ב-CRN כדי להפחית שונות בעת השוואה בין חלופות.

במקום להריץ את שתי החלופות עם אקראיות בלתי תלויה, נשתמש באותם זרמי מספרים אקראיים בריצות מקבילות (באמצעות אותו SEED).

$$\mathrm{Cov}(X_j, Y_j) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) > 0$$

CRN - Common Random Numbers

נשתמש ב-CRN כדי להפחית שונות בעת השוואה בין חלופות.

במקום להריץ את שתי החלופות עם אקראיות בלתי תלויה, נשתמש באותם זרמי מספרים אקראיים בריצות מקבילות (באמצעות אותו SEED).

$$\mathrm{Cov}(X_j, Y_j) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) > 0$$

ואז שונות ההפרש היא:

$$\mathrm{Var}(X_j - Y_j) = \mathrm{Var}(X_j) + \mathrm{Var}(Y_j) - 2\,\mathrm{Cov}(X_j, Y_j)$$

CRN - Common Random Numbers

נשתמש ב-CRN כדי להפחית שונות בעת השוואה בין חלופות.

במקום להריץ את שתי החלופות עם אקראיות בלתי תלויה, נשתמש באותם זרמי מספרים אקראיים בריצות מקבילות (באמצעות אותו SEED).

$$\mathrm{Cov}(X_j, Y_j) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) > 0$$

ואז שונות ההפרש היא:

$$\mathrm{Var}(X_j - Y_j) = \mathrm{Var}(X_j) + \mathrm{Var}(Y_j) - 2\,\mathrm{Cov}(X_j, Y_j)$$

המטרה
מתאם חיובי מקטין את שונות ההפרש, ולכן מקל לזהות הבדל בין החלופות.

דוגמה - נתוני הסימולציה

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	0.060	0.042	2	20.26	19.45
3	0.053	0.053	3	21.73	20.95
4	0.060	0.040	4	15.37	19.57
5	0.053	0.050	5	19.61	20.00
6	0.057	0.046	6	18.35	18.66
7	0.060	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	15.10	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:

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

$$\frac{\gamma}{1+\gamma} = \frac{0.05}{1.05} = 0.0476, \quad \alpha_{\text{total}} = 0.2, \quad \alpha_i = \frac{0.2}{4} = 0.05, \quad t_{15-1,1-\frac{0.05}{2}} = 2.145$$

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

הדרישה היא $\frac{\delta}{\bar{X}} \leq 0.0476$. לכן שתי החלופות המסומנות באדום אינן עומדות בדרישת הדיוק היחסי.

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:

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

$$\frac{\gamma}{1+\gamma} = \frac{0.05}{1.05} = 0.0476, \quad \alpha_{\text{total}} = 0.2, \quad \alpha_i = \frac{0.2}{4} = 0.05, \quad t_{15-1,1-\frac{0.05}{2}} = 2.145$$

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

הדרישה היא $\frac{\delta}{\bar{X}} \leq 0.0476$. לכן שתי החלופות המסומנות באדום אינן עומדות בדרישת הדיוק היחסי.

סעיף א - המשך

ניתן לראות ששתי החלופות לא עומדות בדרישת הדיוק היחסי.

נשערך את מספר הריצות הנדרש באמצעות דיוק גס:

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

מצב קיים:

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

מערכת חדשה:

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

השוואת חלופות ורמת מובהקות

כאשר יש כמה חלופות וכמה מדדים, נחלק את רמת המובהקות הכוללת בין כל ההשוואות.

השוואת חלופות ורמת מובהקות

כאשר יש כמה חלופות וכמה מדדים, נחלק את רמת המובהקות הכוללת בין כל ההשוואות.

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

השוואת חלופות ורמת מובהקות

כאשר יש כמה חלופות וכמה מדדים, נחלק את רמת המובהקות הכוללת בין כל ההשוואות.

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

עבור r חלופות ו k מדדים נבצע $r \cdot k$ השוואות.

השוואת חלופות ורמת מובהקות

כאשר יש כמה חלופות וכמה מדדים, נחלק את רמת המובהקות הכוללת בין כל ההשוואות.

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

עבור r חלופות ו k מדדים נבצע $r \cdot k$ השוואות.

אז בהשוואה זוגית מלאה מספר ההשוואות הוא:

השוואת חלופות ורמת מובהקות

כאשר יש כמה חלופות וכמה מדדים, נחלק את רמת המובהקות הכוללת בין כל ההשוואות.

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

עבור r חלופות ו k מדדים נבצע $r \cdot k$ השוואות.

אז בהשוואה זוגית מלאה מספר ההשוואות הוא:

$$\binom{r}{2} \cdot k$$

השוואה בין חלופות: מבחן t-מזווג

נתונות שתי סדרות (באורך זהה) של תוצאות מריצות סימולציה:

$$X_{1,j}, \qquad j = 1, \ldots, n \qquad\qquad X_{2,j}, \qquad j = 1, \ldots, n$$

השוואה בין חלופות: מבחן t-מזווג

נתונות שתי סדרות (באורך זהה) של תוצאות מריצות סימולציה:

$$X_{1,j}, \qquad j = 1, \dots, n \qquad\qquad X_{2,j}, \qquad j = 1, \dots, n$$

ניצור משתנה מקרי עבור ההפרש בין אותו מדד בריצות מקבילות:

$$Z_j = X_{1,j} - X_{2,j}$$

השוואה בין חלופות: מבחן t-מזווג

נתונות שתי סדרות (באורך זהה) של תוצאות מריצות סימולציה:

$$X_{1,j}, \qquad j = 1, \ldots, n \qquad\qquad X_{2,j}, \qquad j = 1, \ldots, n$$

ניצור משתנה מקרי עבור ההפרש בין אותו מדד בריצות מקבילות:

$$Z_j = X_{1,j} - X_{2,j}$$

נחשב רווח סמך עבור המשתנה המקרי החדש:

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

השוואה בין חלופות: מבחן t-מזווג

נתונות שתי סדרות (באורך זהה) של תוצאות מריצות סימולציה:

$$X_{1,j}, \qquad j = 1, \dots, n \qquad\qquad X_{2,j}, \qquad j = 1, \dots, n$$

ניצור משתנה מקרי עבור ההפרש בין אותו מדד בריצות מקבילות:

$$Z_j = X_{1,j} - X_{2,j}$$

נחשב רווח סמך עבור המשתנה המקרי החדש:

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

הנחות:

- יש תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בתוך כל חלופה בפני עצמה.
- ניתן להניח נורמליות עבור סדרת ההפרשים.

השוואה בין חלופות: מבחן t-מזווג

$$Z_j = X_{1,j} - X_{2,j}, \qquad \delta = t_{n-1,1-\alpha/2}\frac{S_Z(n)}{\sqrt{n}}$$

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad S_Z^2(n) = \frac{\sum_{j=1}^{n}(Z_j - \bar{Z}(n))^2}{n-1}$$

מבחן השערות:

$$\mu_0 = 0$$

השערת האפס: ממוצע ההפרשים הוא 0

$$H_0 : \mu_Z = \mu_0$$

$$H_1 : \mu_Z \neq \mu_0$$

$$t_{\text{stat}} = \frac{\bar{Z} - \mu_0}{S_Z/\sqrt{n}}$$

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

$$|t_{\text{stat}}| > t_{\text{crit}} \Rightarrow \text{Reject } H_0$$

השוואה בין חלופות: מבחן t-מזווג

$$Z_j = X_{1,j} - X_{2,j}, \qquad \delta = t_{n-1,1-\alpha/2}\frac{S_Z(n)}{\sqrt{n}}$$

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad S_Z^2(n) = \frac{\sum_{j=1}^{n}(Z_j - \bar{Z}(n))^2}{n-1}$$

מבחן השערות:

$$H_0 : \mu_Z = \mu_0$$

$$\mu_0 = 0$$

$$H_1 : \mu_Z \neq \mu_0$$

השערת האפס: ממוצע ההפרשים הוא 0

בעצם בדקנו אם

$$t_{\text{stat}} = \frac{\bar{Z} - \mu_0}{S_Z/\sqrt{n}}$$

$$|\bar{Z}| > \delta$$

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

$$|t_{\text{stat}}| > t_{\text{crit}} \Rightarrow \text{Reject } H_0$$

השוואה בין חלופות: מבחן t-מזווג

$$Z_j = X_{1,j} - X_{2,j}, \qquad \delta = t_{n-1,1-\alpha/2}\frac{S_Z(n)}{\sqrt{n}}$$

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad S_Z^2(n) = \frac{\sum_{j=1}^{n}(Z_j - \bar{Z}(n))^2}{n-1}$$

מבחן השערות:

$$\mu_0 = 0$$

השערת האפס: ממוצע ההפרשים הוא 0

בעצם בדקנו אם

$$|\bar{Z}| > \delta$$

$$\frac{|\bar{Z}|}{S_Z/\sqrt{n}} > \delta \cdot \frac{\sqrt{n}}{S_Z}$$

$$H_0 : \mu_Z = \mu_0$$

$$H_1 : \mu_Z \neq \mu_0$$

$$t_{\text{stat}} = \frac{\bar{Z} - \mu_0}{S_Z/\sqrt{n}}$$

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

$$|t_{\text{stat}}| > t_{\text{crit}} \Rightarrow \text{Reject } H_0$$

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 2 טובה יותר מחלופה 1.

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 2 טובה יותר מחלופה 1.
- אם רווח הסמך שלילי:

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 2 טובה יותר מחלופה 1.
- אם רווח הסמך שלילי:
  אז חלופה 1 טובה יותר מחלופה 2.

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 2 טובה יותר מחלופה 1.
- אם רווח הסמך שלילי:
  אז חלופה 1 טובה יותר מחלופה 2.
- אם רווח הסמך מכיל את 0 :

פירוש רווח הסמך

נניח שמטרתנו היא **למזער** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 2 טובה יותר מחלופה 1.
- אם רווח הסמך שלילי:
  אז חלופה 1 טובה יותר מחלופה 2.
- אם רווח הסמך מכיל את 0 :
  לא ניתן לקבוע הבדל מובהק.

פירוש רווח הסמך

נניח שמטרתנו היא **למקסם** את המדד.

$$Z_j = X_{1,j} - X_{2,j}$$

כאשר חלופה 1 היא המצב הקיים וחלופה 2 היא המערכת החדשה.

- אם רווח הסמך חיובי:
  אז חלופה 1 טובה יותר מחלופה 2.
- אם רווח הסמך שלילי:
  אז חלופה 2 טובה יותר מחלופה 1.
- אם רווח הסמך מכיל את 0 :
  לא ניתן לקבוע הבדל מובהק.

השוואה בין חלופות - קירוב Welch

אם אין תלות בין הריצות של החלופות או אם מספר הריצות שונה עבור כל חלופה, לא ניתן להשתמש במבחן מזווג, נשתמש בקירוב Welch.

השוואה בין חלופות - קירוב Welch

אם אין תלות בין הריצות של החלופות או אם מספר הריצות שונה עבור כל חלופה, לא ניתן להשתמש במבחן מזווג, נשתמש בקירוב Welch.

רווח הסמך עבור ההפרש הוא:

$$\bar{X}_1(n_1) - \bar{X}_2(n_2) \pm t_{f,1-\alpha/2}\sqrt{\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}}$$

השוואה בין חלופות - קירוב Welch

אם אין תלות בין הריצות של החלופות או אם מספר הריצות שונה עבור כל חלופה, לא ניתן להשתמש במבחן מזווג, נשתמש בקירוב Welch.

רווח הסמך עבור ההפרש הוא:

$$\bar{X}_1(n_1) - \bar{X}_2(n_2) \pm t_{f,1-\alpha/2}\sqrt{\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}}$$

נקרב את דרגות החופש ע"י:

$$\hat{f} = \frac{\left(\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}\right)^2}{\frac{\left(\frac{S_1^2(n_1)}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2(n_2)}{n_2}\right)^2}{n_2-1}}$$

הנחות:

- אין תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בתוך כל חלופה בפני עצמה.
- ניתן להניח נורמליות עבור סדרת ההפרשים.

סעיף ב'

ברמת מובהקות של 0.10, על פי שני המדדים, מצאו איזו מן החלופות טובה יותר.

סעיף ב'

ברמת מובהקות של 0.10, על פי שני המדדים, מצאו איזו מן החלופות טובה יותר.

כיוון שיש מתאם בין החלופות, נרצה לבצע מבחן t מזווג.

סעיף ב'

ברמת מובהקות של 0.10, על פי שני המדדים, מצאו איזו מן החלופות טובה יותר.

כיוון שיש מתאם בין החלופות, נרצה לבצע מבחן t מזווג.

נבדוק שההנחות מתקיימות:

- יש תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בתוך כל חלופה בפני עצמה.
- ניתן להניח נורמליות של ההפרשים.

סעיף ב'

ברמת מובהקות של 0.10, על פי שני המדדים, מצאו איזו מן החלופות טובה יותר.

כיוון שיש מתאם בין החלופות, נרצה לבצע מבחן t מזווג.

נבדוק שההנחות מתקיימות:

- יש תלות בין ריצות מקבילות בשתי הסדרות.
- אין תלות בין ריצות בתוך כל חלופה בפני עצמה.
- ניתן להניח נורמליות של ההפרשים.

השלב הבא
נחשב את סדרת ההפרשים עבור כל מדד:

$$Z_j = X_{\text{קיים},j} - X_{\text{חדש},j}$$

ואז נבנה רווח סמך עבור μ_Z.

אוניברסיטת בן-גוריון בנגב
جامعة بن غوريون في النقب
Ben-Gurion University of the Negev

דוגמה - נתוני הסימולציה

נחשב את ההפרש:

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	0.060	0.042	2	20.26	19.45
3	0.053	0.053	3	21.73	20.95
4	0.060	0.040	4	15.37	19.57
5	0.053	0.050	5	19.61	20.00
6	0.057	0.046	6	18.35	18.66
7	0.060	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	15.10	20.07
ממוצע	0.056	0.047	**ממוצע**	18.21	19.411
סטיית תקן	0.003	0.005	**סטיית תקן**	2.266	0.867

דוגמה - נתוני הסימולציה

נחשב את ההפרש:

אחוז הפגומים				רמת הזיהום			
מספר ריצה	מצב קיים	מערכת חדשה	הפרש	מספר ריצה	מצב קיים	מערכת חדשה	הפרש
1	0.053	0.053	0.000	1	21.54	19.11	2.43
2	0.060	0.042	0.018	2	20.26	19.45	0.81
3	0.053	0.053	0.000	3	21.73	20.95	0.78
4	0.060	0.040	0.020	4	15.37	19.57	−4.20
5	0.053	0.050	0.003	5	19.61	20.00	−0.39
6	0.057	0.046	0.011	6	18.35	18.66	−0.31
7	0.060	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	−1.10
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	−2.30
15	0.053	0.047	0.006	15	15.10	20.07	−4.97
ממוצע	0.056	0.047	0.009	**ממוצע**	18.21	19.411	−1.201
סטיית תקן	0.003	0.005	0.007	**סטיית תקן**	2.266	0.867	2.215

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- נחשב את רווח הסמך עבור כל אחד מהמדדים:

אחוז הפגומים:

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

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- 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.

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- 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.

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- 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:

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

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- 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

סעיף ב-המשך:

- את α נחלק במספר ההשוואות שאנחנו עושים:

$$\alpha_{total} = 0.1 \qquad \alpha_i = 0.05$$

- 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.

מימוש ב-Python

https://colab.research.google.com/drive/1FoxALcrsFrJ_Oj9UTG9dAGVfPbYcAsq3

סיכום

קירוב לפי Welch	מזווג *t*
ניתן להשתמש ב-*n* שונים עבור כל חלופה	אותו *n* לשתי 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 ה-seed ונעבור להרצה הבאה עם seed אחר הפעם - אך אותו ה-seed למצב הקיים והחלופות גם בעת, וכך הלאה..

```
For i in range(number_of_runs):
        np.random.seed(i*2+40) - any function of I
        measue1, measue2, measue3 = run_simulation()
```