

פרויקט תכנון ניסויים וניתוח שונות

סמסטר סתיו 2024

ד"ר בני יוסקוביץ'

קבוצה מספר 5

מגישים:

אופיר גוטליב

נדב בן מנחם

ניב פינטו

תרגיל מספר 1

	תוספת האלומיניום					
	0.12%	0.14%	0.16%	0.18%	0.20%	
K	2433	2474	1992	2108	2430	מתקן 1
11	2407	2415	1969	2042	2425	
	3654	2810	3024	2483	1785	מתקן 2
	3618	2745	3075	2425	1705	
	2128	2757	2700	3839	1800	מתקן 3
	2130	2786	2626	3818	1862	
	2305	2399	2636	3188	2710	מתקן 4
	2308	2390	2693	3205	2614	
	3157	3008	2081	3037	2235	מתקן 5
	3191	3006	2092	3053	2213	

בשאלה מתואר ניסוי ב-2 גורמים, 5 רמות לכל אחד, 2 חזרות, ומשתנה תגובה אחד.
נבחר להתייחס לגורם המתקן בתור בלוק ונבצע ניתוח באמצעות ניתוח שונות חד כיווני.

המודל:

$$Strength_{i,j,k} = \mu + \alpha_i + \beta_j + \varepsilon_{i,j,k}$$

כאשר: $i = 1, 2, 3, 4, 5$ $j = 1, 2, 3, 4, 5$ $k = 1, 2$

מרכיבים:

- $Strength_{i,j,k}$ המדידה ה- k של רמת האלומיניום ה- i במתקן ה- j .
- μ הממוצע הכללי.
- α_i השפעת רמת האלומיניום ה- i .
- β_j השפעת הבלוק של המתקן ה- j .
- $\varepsilon_{i,j}$ הטעות האקראית של המדידה ה- k של רמת האלומיניום ה- i במתקן ה- j .

הנחות:

- $\varepsilon_{i,j,k} \sim N(0, \sigma_\varepsilon^2) \quad \forall i, j, k$
- $\varepsilon_{i,j,k}$ – בלתי תלוי בין הקבוצות ובתוך הקבוצות
- β_j – בלתי תלוי
- $\sum_{i=1}^5 \alpha_i = 0$
- $\sum_{j=1}^5 \beta_j = 0$
- $\sum_{i=1}^5 n_i \alpha_i = 0$ לכל i ו- j

טבלת ניתוח שונות:

מקור	SS	df	$MS = \frac{SS}{df}$	$F_{st} = \frac{MS}{MSE}$	$F_{cr} = F_{4,41}^{0.05}$
אלומיניום	SSA	$a - 1 = 4$	$MSA = \frac{SSA}{4}$ $= 791,779.3$	$\frac{MSB}{MSE}$ $= 3.632$	2.6
מתקן	SSB	$b - 1 = 4$	$MSB = \frac{SSB}{4}$ $= 355,731.3$	$\frac{MSA}{MSE}$ $= 1.632$	2.6
טעויות	SSE	$nab - a - b + 1 = 41$	$MSE = \frac{SSE}{41}$ $= 217,973.11$		
סה"כ	SST	$nab - 1 = 49$			

$$\hat{\mu} = 2600$$

$\hat{B}_1 = \bar{y}_1 - \hat{\mu} = 2269 - 2600 = -330$	$\hat{A}_1 = \bar{y}_1 - \hat{\mu} = 2733 - 2600 = 133$
$\hat{B}_2 = \bar{y}_2 - \hat{\mu} = 263732 - 2600 = 133$	$\hat{A}_2 = \bar{y}_2 - \hat{\mu} = 2679 - 2600 = 79$
$\hat{B}_3 = \bar{y}_3 - \hat{\mu} = 2645 - 2600 = 45$	$\hat{A}_3 = \bar{y}_3 - \hat{\mu} = 2489 - 2600 = -111$
$\hat{B}_4 = \bar{y}_4 - \hat{\mu} = 2645 - 2600 = 45$	$\hat{A}_4 = \bar{y}_4 - \hat{\mu} = 2920 - 2600 = 320$
$\hat{B}_5 = \bar{y}_5 - \hat{\mu} = 2707 - 2600 = 108$	$\hat{A}_5 = \bar{y}_5 - \hat{\mu} = 2178 - 2600 = -422$

$$SSB = B \cdot n \cdot \sum_{i=1}^5 (\bar{y}_{i,j} - \hat{\mu})^2 = 5 \cdot 2 \cdot \sum_{i=1}^5 (\hat{B}_i)^2 = 1422925.326$$

$$SSA = A \cdot n \cdot \sum_{i=1}^5 (\bar{y}_{i,j} - \hat{\mu})^2 = 5 \cdot 2 \cdot \sum_{i=1}^5 (\hat{A}_i)^2 = 3167117$$

$$SST = \sum_{j=1}^5 \sum_{i=1}^5 (y_{i,j} - \bar{y})^2 = 13526939$$

$$SSE = SST - SSA - SSB = 8936897$$

```
> # Running a Randomized Complete Block ANOVA test
> model <- aov(strength ~ aluminum + facility, data)
> summary(model)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
aluminum	4	3167117	791779	3.632	0.0127 *
facility	4	1422925	355731	1.632	0.1846
Residuals	41	8936897	217973		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

טבלת ניתוח שונות
בשימוש R:

ניתן לראות כי השפעת הבלוקים *facility* אינה מובהקת ($p\text{-value} = 0.1846 > 0.05$)

לכן, בשלב הבא נעמיס את הבלוקים על השגיאה ונבחן בשנית את מובהקות המודל:

מקור	SS	df	$MS = \frac{SS}{df}$	$F_{st} = \frac{MS}{MSE}$	$F_{cr} = F_{4,41}^{0.05}$
אלומיניום	SSA	4	791,779	$\frac{MSB}{MSE}$ = 3.439	2.6
טעויות	$SSE + SSB$	$4 + 41$ = 45	$MSE = \frac{SSE + SSB}{45}$ = 230,218		
סה"כ	SST	49			

$$SSA = 3,167,117$$

$$SSE + SSB = 8,936,897 + 1,422,925 = 10,359,822$$

$$SST = 13,526,939$$

הטבלה המועמסת בשימוש R:

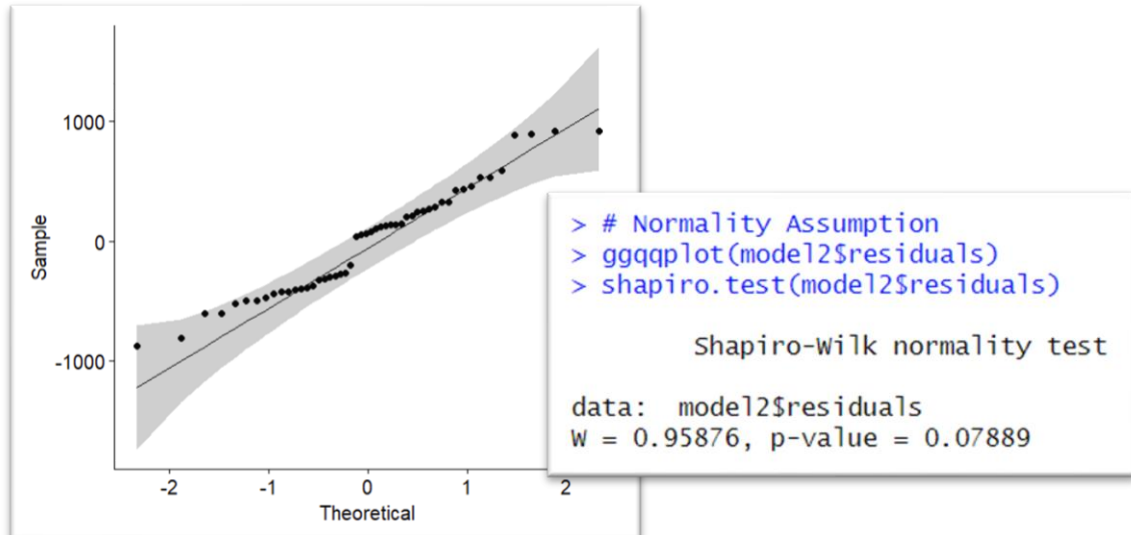
```
> # Facility wasn't significant
> # Running a standard one-way ANOVA
> model2 <- aov(strength ~ aluminum, data)
> summary(model2)
          Df Sum Sq Mean Sq F value Pr(>F)
aluminum    4  3167117   791779    3.439 0.0155 *
Residuals   45 10359823   230218
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ניתן לראות כי גורם האלומיניום נותר מובהק ($p\text{-value}=0.0155<0.05$).

נבדוק שהנחות המודל המועמס אכן מתקיימות, ולאחר מכן נתקדם לניתוחי המשך.

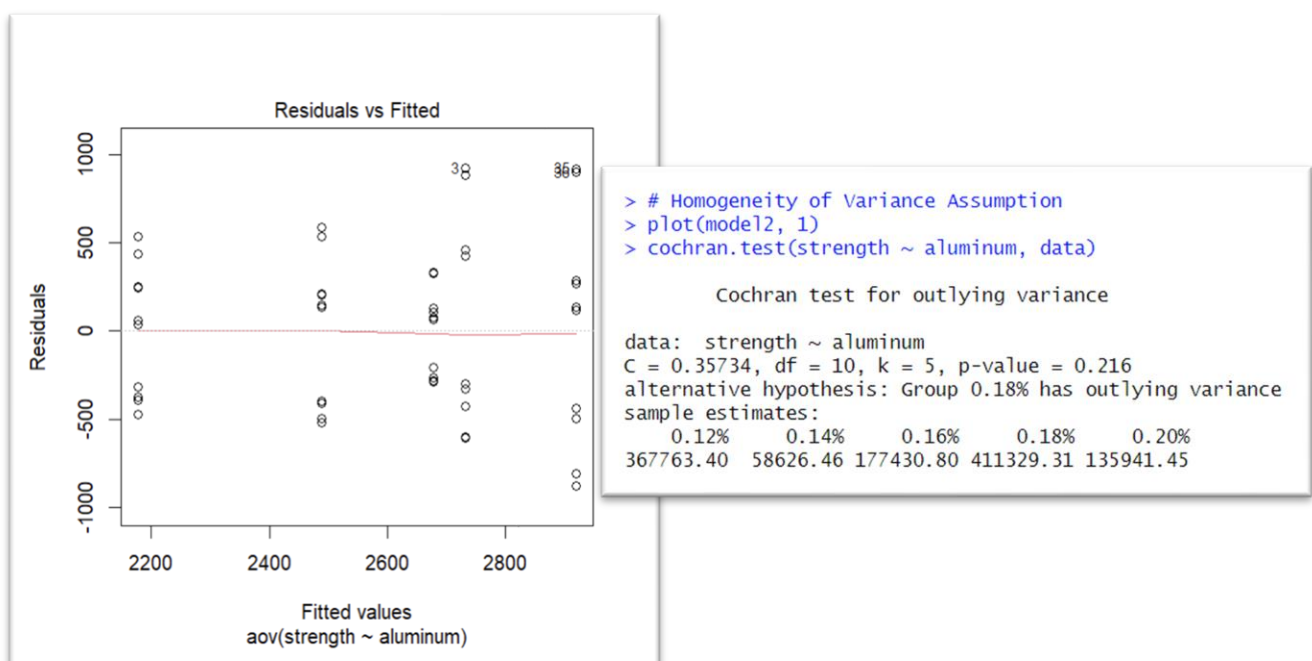
בדיקת הנחות המודל

1. הנחת הנורמליות – נבדוק את הנחת הנורמליות של השגיאה ע"י בחינה ויזואלית של תרשים QQ עבור השגיאות, וכן ננסה לגבות את המסקנה גם בצורה כמותית ע"י מובהקות מבחן סטטיסטי $Shapiro-Wilk$:



ניתן לראות כי השאריות נופלות בקירוב סמוך לקו המנחה, וכן ערך ה- p של המבחן לא מובהק מספיק על מנת לדחות את השערת האפס ($0.05 < 0.07889$). לכן, נקבע שהנחת הנורמליות מתקיימת.

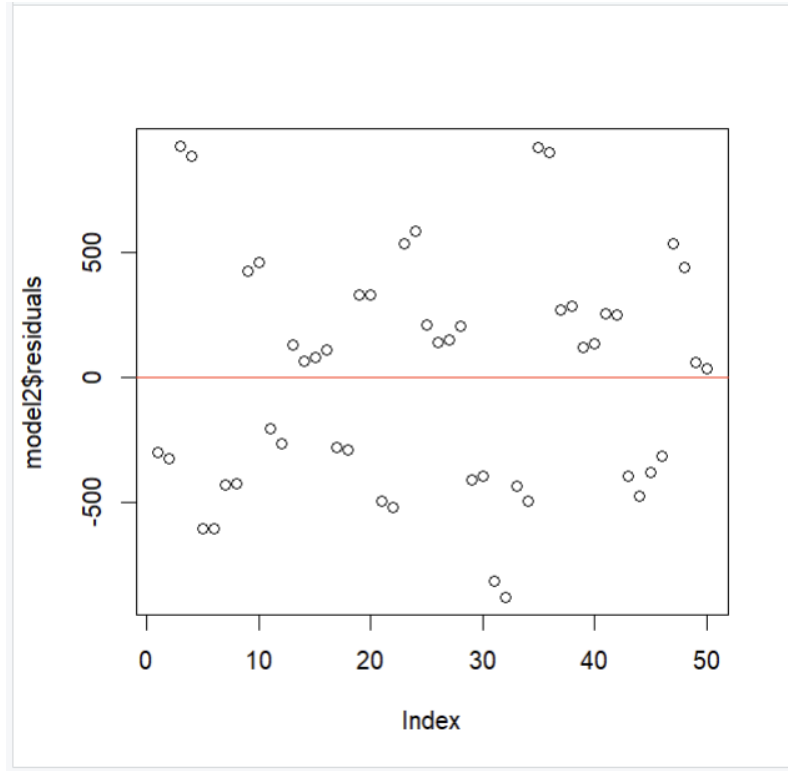
2. הנחת שוויון שוניות – נתבונן בתרשים השאריות כנגד הערך החזוי, וכן ניעזר כמותית במובהקות מבחן $Cochran C$, המתאים ביותר בהינתן ניסוי מאוזן.



ניתן לראות בכל התרשים רנדומליות יחסית של השאריות סביב קו ה-0, וכן ערך הק של המבחן רחוק מלהיות מובהק ולדחות את השערת האפס ($0.05 < 0.216$)

לכן, נקבע שהנחת שוויון השונויות מתקיימת.

3. הנחת אי תלות – נתבונן בתרשים השאריות לפי סדר הנתונים.



ניתן לראות אקראיות יחסית של ערך השגיאה.

לכן, נקבע שהנחת האי תלות מתקיימת.

רמת אלומיניום מועדפת

עתה כשגורם האלומיניום מובהק והנחות המודל מתקיימות, נבצע ניתוחי המשך. נעשה זאת ע"י השוואות מרובות באמצעות TukeyHSD, המתאים לניסוי מאוזן.

$$H_0: \mu_i - \mu_j = 0$$

$$H_1: \mu_i - \mu_j \neq 0$$

תחילה, נחשב את ערך HSD:

$$HSD = q_{k,N-k}^\alpha * \sqrt{\frac{MSE}{n}} = q_{5,45}^{0.05} * \sqrt{\frac{230,218}{10}} = 4.02 * 151.73 = 609.95$$

```
> mse <- summary(model2)[[1]]["Residuals", "Mean Sq"]
> df <- summary(model2)[[1]]["Residuals", "Df"]
> k <- 5
> n <- 10
> hsd <- qtkey(0.05, k, df, lower.tail = FALSE) * sqrt(mse/n)
> hsd
[1] 609.7122
```

וכעת נבצע $\binom{5}{2} = 10$ השוואות בין ממוצעי רמות שונות של אלומיניום, ונבדוק עבור אילו השוואות מתקיים $|\bar{y}_i - \bar{y}_j| < HSD$, כלומר אינטרקציה מובהקת.

$$\begin{aligned} |\bar{y}_{0.12} - \bar{y}_{0.14}| &= |2733 - 2679| = 54 < HSD \\ |\bar{y}_{0.12} - \bar{y}_{0.16}| &= |2733 - 2489| = 244 < HSD \\ |\bar{y}_{0.12} - \bar{y}_{0.18}| &= |2733 - 2920| = 187 < HSD \\ |\bar{y}_{0.12} - \bar{y}_{0.20}| &= |2733 - 2178| = 555 < HSD \\ |\bar{y}_{0.14} - \bar{y}_{0.16}| &= |2679 - 2489| = 190 < HSD \\ |\bar{y}_{0.14} - \bar{y}_{0.18}| &= |2679 - 2920| = 241 < HSD \\ |\bar{y}_{0.14} - \bar{y}_{0.20}| &= |2679 - 2178| = 501 < HSD \\ |\bar{y}_{0.16} - \bar{y}_{0.18}| &= |2489 - 2920| = 431 < HSD \\ |\bar{y}_{0.16} - \bar{y}_{0.20}| &= |2489 - 2178| = 311 < HSD \\ |\bar{y}_{0.18} - \bar{y}_{0.20}| &= |2920 - 2178| = 741.86 > HSD \end{aligned}$$

```
> tukey_comparisons <- TukeyHSD(model2)
> tukey_comparisons
Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = strength ~ aluminum, data = data)

$aluminum
      diff      lwr      upr      p adj
0.14%-0.12% -54.32366 -664.0359  555.38859 0.9990590
0.16%-0.12% -244.32756 -854.0398  365.38469 0.7852557
0.18%-0.12%  186.62504 -423.0872  796.33728 0.9064457
0.20%-0.12% -555.23189 -1164.9441  54.48036 0.0898555
0.16%-0.14% -190.00390 -799.7161  419.70835 0.9007790
0.18%-0.14%  240.94870 -368.7636  850.66094 0.7936465
0.20%-0.14% -500.90823 -1110.6205  108.80402 0.1530441
0.18%-0.16%  430.95260 -178.7597  1040.66484 0.2786684
0.20%-0.16% -310.90433 -920.6166  298.80792 0.6000912
0.20%-0.18% -741.85693 -1351.5692 -132.14468 0.0100662
```

קיימת אינטרקציה מובהקת ישירה או עקיפה בין כל שתי רמות, מה שאומר שאין קבוצות של רמות להן ערך חוזק ממוצע שונה באופן מובהק מקבוצות אחרות.

לכן, ניתן לקבוע כי ברמת מובהקות 5%, אין רמות אלומיניום עדיפות.

תרגיל מספר 2

א. מטריצת הניסוי:

ניסוי פקטוריאלי מלא עם 5 גורמים בשתי רמות עם שתי חזרות:

index	X1	X2	X3	X4	X5	Y1	Y2	Y1	Y2
1	1	1	1	1	1	97.07598	93.72402	104.7251	86.07495
2	1	1	1	1	-1	105.3034	88.49659	104.5417	89.25827
3	1	1	1	-1	1	109.2699	95.9301	112.4232	92.77679
4	1	1	1	-1	-1	110.3244	97.87556	117.892	90.30801
5	1	1	-1	1	1	108.1611	92.63894	115.6081	85.19195
6	1	1	-1	1	-1	111.2621	92.53786	113.1394	90.66058
7	1	1	-1	-1	1	122.049	93.15104	121.2283	93.97166
8	1	1	-1	-1	-1	118.463	99.73701	118.5997	99.60034
9	1	-1	1	1	1	93.60986	99.19014	96.24994	96.55006
10	1	-1	1	1	-1	94.8098	94.9902	96.42224	93.37776
11	1	-1	1	-1	1	102.2874	104.9126	107.0641	100.1359
12	1	-1	1	-1	-1	99.59758	104.6024	98.55411	105.6459
13	1	-1	-1	1	1	110.6039	92.19611	101.7738	101.0262
14	1	-1	-1	1	-1	101.6054	98.19458	106.0263	93.77371
15	1	-1	-1	-1	1	113.6384	103.5616	114.0994	103.1006
16	1	-1	-1	-1	-1	107.5809	106.6191	107.5831	106.6169
17	-1	1	1	1	1	99.46904	88.73096	99.69249	88.50751
18	-1	1	1	1	-1	103.2991	87.90092	100.4782	90.72184
19	-1	1	1	-1	1	89.23705	84.56295	86.81975	86.98025
20	-1	1	1	-1	-1	94.48823	82.31177	90.64371	86.15629
21	-1	1	-1	1	1	109.2084	96.99155	111.2273	94.97266
22	-1	1	-1	1	-1	113.0507	96.14929	110.4514	98.74858
23	-1	1	-1	-1	1	102.5621	89.23789	104.7123	87.08775
24	-1	1	-1	-1	-1	102.8573	91.9427	101.4277	93.37225
25	-1	-1	1	1	1	103.2074	96.99263	105.7748	94.42518
26	-1	-1	1	1	-1	95.49507	101.7049	99.5122	97.6878
27	-1	-1	1	-1	1	97.96454	87.83546	93.81637	91.98363
28	-1	-1	1	-1	-1	90.15759	92.64241	88.7946	94.0054
29	-1	-1	-1	1	1	108.9348	109.2652	114.6599	103.5401
30	-1	-1	-1	1	-1	106.6157	108.5843	107.5082	107.6918
31	-1	-1	-1	-1	1	109.6962	94.10383	109.1227	94.67734
32	-1	-1	-1	-1	-1	97.18208	103.6179	97.19035	103.6097

עבור ניסוי זה נקבל שמספר החזרות הוא $n=64$.

דרגות החופש של הניסוי הוא $df: 32 = 64 - (2^5 - 1) - 1$ (מספר החזרות פחות מס' הגורמים, האינטראקציות והממוצע).

רמת המובהקות $\alpha: 5\%$.

t קריטי עבור 32 דרגות חופש ורמת מובהקות של $5\%: 2.039$.

ב. מודל Y1

נרצה לבצע מבחן t על כל אחד מהגורמים והאינטראקציות. עבור זה נחשב את האפקטים של כל הגורמים והאינטראקציות, אך קודם נחשב את מספר פרמטרים של הניסוי כולו:

1. ממוצע של ערכי Y1:

$$104.482 = (103.6179 + \dots + 97.07598) / 64$$

2. סטיית התקן של הניסוי SE.exp:

$$8.25947 = [(103.6197 - 104.482)^2 + \dots + (97.05798 - 104.482)^2] / (64 - 1)$$

3. סטיית התקן של כל אפקט SE.Coeff:

$$1.03243 = 8.25947 / \sqrt{64} = SE.exp / \sqrt{n}$$

כעת נחשב את האפקט של כל גורם ואינטראקציה:

$$Effect(X_i) = \text{avg}(y(X_i=+1)) - \text{avg}(y(X_i=-1))$$

$$6.13485 = (100.901 + \dots + 107.582) / 16 - (99.5808 + \dots + 97.1862) / 16$$

לאחר מכן נחשב את ה-Coeff בהתאמה:

$$Coef(x_1) = 6.13485 / 2 = 3.06742 \quad \leftarrow \quad Coef(X_i) = Effect(X_i) / 2$$

כעת נגדיר את המבחן הסטטיסטי עבור כל גורם ואינטראקציה:

$$H_0: Coef(X_i) = 0$$

$$H_1: \text{else}$$

$$T_s = [Coef(X_i) - 0] / SE.Coeff(X_i)$$

$$T_s = 3.06742 / 1.03243 = 2.97106 \quad \text{לדוג' עבור } X_1$$

עבור כל גורם ואינטראקציה נבדוק מובהקות כך שאם הערך הקריטי בערך מוחלט גדול מערך הסטטיסטי הגורם/אינטראקציה מובהק. להלן סיכום של כל השלבים שעשינו עד כה עבור כל הגורמים והאינטראקציות:

Term	Effect	Coef	SE.Coeff	T-Value	Significance Test
X1	6.134856	3.067428	1.032434	2.971065	Yes
X2	4.142268	2.071134	1.032434	2.00607	No
X3	-9.65088	-4.82544	1.032434	-4.67385	Yes
X4	0.380493	0.190247	1.032434	0.18427	No
X5	2.034846	1.017423	1.032434	0.985461	No
X12	4.517737	2.258868	1.032434	2.187906	Yes
X13	0.821433	0.410717	1.032434	0.397814	No
X14	-7.86402	-3.93201	1.032434	-3.80848	Yes
X15	-0.89971	-0.44986	1.032434	-0.43572	No
X23	-0.24441	-0.12221	1.032434	-0.11837	No
X24	-0.1495	-0.07475	1.032434	-0.0724	No
X25	-3.45692	-1.72846	1.032434	-1.67416	No

X34	0.265238	0.132619	1.032434	0.128453	No
X35	-1.51154	-0.75577	1.032434	-0.73203	No
X45	-1.38104	-0.69052	1.032434	-0.66883	No
X123	0.704499	0.352249	1.032434	0.341184	No
X124	-1.17106	-0.58553	1.032434	-0.56714	No
X125	1.19863	0.599315	1.032434	0.580488	No
X134	-0.86605	-0.43303	1.032434	-0.41942	No
X135	-0.21608	-0.10804	1.032434	-0.10465	No
X145	-0.41694	-0.20847	1.032434	-0.20192	No
X234	-0.0604	-0.0302	1.032434	-0.02925	No
X235	-0.59868	-0.29934	1.032434	-0.28993	No
X245	0.758285	0.379143	1.032434	0.367232	No
X345	0.850595	0.425297	1.032434	0.411937	No
X1234	-0.10055	-0.05028	1.032434	-0.0487	No
X1235	-0.19241	-0.0962	1.032434	-0.09318	No
X1245	-0.00629	-0.00314	1.032434	-0.00305	No
X1345	-0.81421	-0.40711	1.032434	-0.39432	No
X2345	0.139494	0.069747	1.032434	0.067556	No
X12345	0.489912	0.244956	1.032434	0.237261	No

לאחר בדיקת המובהקות נקבל שרק הגורמים X1, X3 והאינטראקציות X12, X14 מובהקים ברמת מובהקות של 5%. לכן המודל שנקבל הוא:

$$\hat{Y}_1 = 104.482 - 3.067 * X_1 - 4.825 * X_3 + 2.258 * X_{12} - 3.932 * X_{14}$$

עבור רמות הגורמים הבאים (X2 = -1, X4 = 1) X1 = -1, X3 = -1, X12 = 1, X14 = -1 נקבל ערך מקסימלי שערכו 112.4306 במדד דרגת החוזק המרבית הנדרשת בייצור הבטון.

ג. מודל Y2

נרצה לבצע מבחן t על כל אחד מהגורמים והאינטראקציות. עבור זה נחשב את האפקטים של כל הגורמים והאינטראקציות, אך קודם נחשב את מספר פרמטרים של הניסוי כולו:

1. ממוצע של ערכי Y2:

$$95.529 = (93.724 + \dots + 103.6097) / 64$$

2. סטיית התקן של הניסוי SE.exp:

$$6.247964 = [(95.529 - 93.724)^2 + \dots + (103.6097 - 95.529)^2] / (64 - 1)$$

3. סטיית התקן של כל אפקט SE.Coef:

$$0.780956 = 6.247964 / \sqrt{64} = SE.exp / \sqrt{n}$$

נעת נחשב את האפקט של כל גורם ואינטראקציה:

$$Effect(X_i) = avg(y(X_i = +1)) - avg(y(X_i = -1))$$

$$1.8651 = (93.724 + \dots + 107.6169) / 16 - (88.7309 + \dots + 103.6097) / 16 \text{ – עבור } X_1 \text{ נקבל –}$$

לאחר מכן נחשב את ה-Coef בהתאמה:

$$\text{Coef}(x_1) = 1.8651/2 = 0.93257 \quad \leftarrow \text{Coef}(X_i) = \text{Effect}(X_i)/2$$

כעת נגדיר את המבחן הסטטיסטי עבור על גורם ואינטרקציה:

$$H_0: \text{Coef}(X_i) = 0$$

$$H_1: \text{else}$$

$$T_s = [\text{Coef}(X_i) - 0] / \text{SE.Coef}(X_i)$$

$$T_s = 0.93257 / 0.780956 = 2.97106 \quad \text{לדוג' עבור } X_1$$

עבור כל גורם ואינטראקציה נבדוק מובהקות כך שאם הערך הקריטי בערך מוחלט גדול מערך הסטטיסטי הגורם/אינטראקציה מובהק. להלן סיכום של כל השלבים שעשינו עד כה עבור על הגורמים והאינטראקציות:

Term	Effect	Coef	SE.Coef	T-Value	Significance Test
X1	1.865144	0.932572	0.780956	1.194141	No
X2	-8.14227	-4.07113	0.780956	-5.21301	Yes
X3	-4.34912	-2.17456	0.780956	-2.78448	Yes
X4	-0.38049	-0.19025	0.780956	-0.24361	No
X5	-2.03485	-1.01742	0.780956	-1.30279	No
X12	0.482263	0.241132	0.780956	0.308765	No
X13	3.178567	1.589283	0.780956	2.035049	No
X14	-6.53598	-3.26799	0.780956	-4.1846	Yes
X15	0.899712	0.449856	0.780956	0.576033	No
X23	0.244414	0.122207	0.780956	0.156484	No
X24	0.149502	0.074751	0.780956	0.095717	No
X25	0.456916	0.228458	0.780956	0.292536	No
X34	-0.26524	-0.13262	0.780956	-0.16982	No
X35	1.511541	0.755771	0.780956	0.967751	No
X45	1.381042	0.690521	0.780956	0.8842	No
X123	-0.7045	-0.35225	0.780956	-0.45105	No
X124	1.17106	0.58553	0.780956	0.74976	No
X125	-1.19863	-0.59932	0.780956	-0.76741	No
X134	0.866052	0.433026	0.780956	0.554482	No
X135	0.21608	0.10804	0.780956	0.138343	No
X145	0.416939	0.20847	0.780956	0.266942	No
X234	0.060399	0.0302	0.780956	0.03867	No
X235	0.598675	0.299338	0.780956	0.383296	No
X245	-0.75829	-0.37914	0.780956	-0.48549	No
X345	-0.85059	-0.4253	0.780956	-0.54459	No
X1234	0.10055	0.050275	0.780956	0.064376	No
X1235	0.192409	0.096205	0.780956	0.123188	No
X1245	0.006289	0.003144	0.780956	0.004026	No
X1345	0.814212	0.407106	0.780956	0.521292	No
X2345	-0.13949	-0.06975	0.780956	-0.08931	No
X12345	-0.48991	-0.24496	0.780956	-0.31366	No

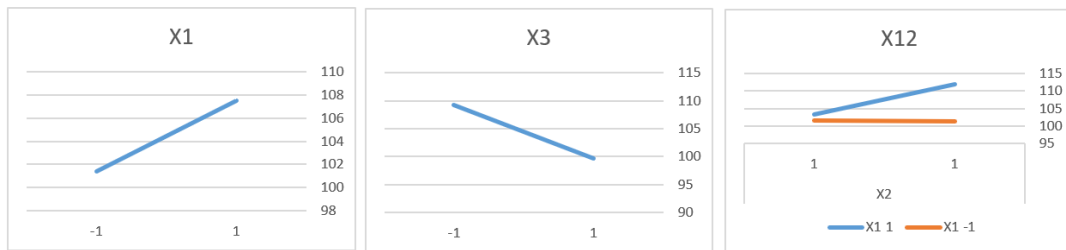
לאחר בדיקת המובהקות נקבל שרק הגורמים X_1 , X_3 והאינטראקציות X_{12} , X_{14} מובהקים ברמת מובהקות של 5%. לכן המודל שנקבל הוא:

$$\hat{Y}_2 = 95.5298 - 4.071 * X_2 - 2.174 * X_3 - 3.267 * X_{14}$$

עבור רמות הגורמים הבאים ($X_4 = -1$), $X_{14} = 1$, $X_3 = 1$, $X_2 = 1$ נקבל ערך מינימלי שערכו 86.0161 בממד הגמישות המזערית הנדרשת בייצור הבטון.

ד. שרטוט גורמים מודל Y_1 :

בחנו את האפקטים של הגורמים: X_1 , X_3 ושל האינטראקציה X_{12} . להלן הגרפים:



ניתן לראות ש- X_1 , X_3 הם גורמי כיוון (כפי שקיבלנו במבחן הסטטיסטי) ו- X_{12} הוא בעל אינטראקציה בינונית אך מובהקת כפי שהצגנו לעיל.

ה. אופטימיזר

בהמשך למודלים שמצאנו בסעיפים הקודמים ניקח את הגורמים המובהקים בלבד שמשותפים לשניהם – X_1 , X_2 , X_3 , X_{13} , X_{14} .

עבור כל קומבינציה של הגורמים נחשב את הפרדיקציה על פי המודלים ל- Y_1 ו- Y_2 . לדוג' עבור הקומבינציה $X_1=1, X_2=1, X_3=1, X_4=1, X_{13}=1, X_{14}$

$$\hat{Y}_1 = 104.482 - 3.067 * 1 - 4.825 * 1 + 2.258 * 1 - 3.932 * 1 = 101.0505$$

$$\hat{Y}_2 = 95.5298 - 4.071 * 1 - 2.174 * X_3 - 3.267 * 1 = 86.0161$$

לאחר מכן נחשב את מידת האופטימיזציה לכל תחזית כך שאת Y_1 נשאף למקסם ואת Y_2 נשאף להביא למינימום. בהמשך לקומבינציה הקודמת נעשה את החישובים הבאים:

$$d(Y_1) = (y.\text{est} - \min) / (\max - \min) = (101.0505 - 86.819) / (122.049 - 86.819) = 0.4039$$

$$d(Y_2) = (\max - y.\text{est}) / (\max - \min) = (109.2652 - 86.0161) / (109.2652 - 82.3117) = 0.8265$$

כעת נחשב את מידת האופטימיזציה המשותפת ל- Y_1 ו- Y_2 עם משקל שווה, נמשיך עם הדוגמה שלנו:

$$D(Y_1, Y_2) = \sqrt{d(Y_1) * d(Y_2)} = \sqrt{0.4039 * 0.8265} = 0.5902$$

את אותו הליך נבצע לכל הקומבינציות הקיימות, לאחר הרצה נקבל את התוצאות הבאות:

index	predict_Y1	predict_Y2	d(Y1)	d(Y2)	D(Y1*Y2)
1	101.050563	86.016126	0.403949	0.862565	0.590282
7	104.369601	91.056192	0.498162	0.675573	0.580125
2	106.703957	93.614861	0.564424	0.580644	0.572477
4	106.555725	94.678769	0.560216	0.541172	0.550612
13	104.597501	94.191559	0.504631	0.559248	0.531238

24	104.804910	94.809283	0.510518	0.536330	0.523265
9	105.065128	95.130106	0.517905	0.524427	0.521156
21	105.042942	95.257529	0.517275	0.519700	0.518486
28	104.554607	95.052958	0.503413	0.527289	0.515213
16	104.742457	95.244259	0.508746	0.520192	0.514437
26	104.773802	95.436751	0.509635	0.513050	0.511340
15	104.514964	95.266331	0.502288	0.519373	0.510759
5	103.237757	94.410117	0.466034	0.551139	0.506803
32	104.481715	95.529811	0.501344	0.509598	0.505454
23	104.428647	95.506219	0.499838	0.510473	0.505127
22	104.636083	95.801185	0.505726	0.499529	0.502618
31	104.236760	95.529811	0.494391	0.509598	0.501937
29	104.643866	95.844514	0.505947	0.497922	0.501918
30	104.411969	95.774767	0.499365	0.500509	0.499937
17	104.285277	95.742664	0.495768	0.501701	0.498726
25	104.029960	95.575740	0.488521	0.507894	0.498113
19	104.276854	95.896596	0.495529	0.495990	0.495759
14	103.598252	95.400009	0.476267	0.514413	0.494973
27	104.170814	95.933773	0.492519	0.494610	0.493564
20	104.038054	95.864685	0.488751	0.497173	0.492944
11	103.588019	95.768049	0.475976	0.500759	0.488210
6	110.730460	100.254467	0.678718	0.334308	0.476341
12	103.016042	95.985175	0.459740	0.492703	0.475937
18	103.648400	96.496174	0.477690	0.473745	0.475713
8	101.495055	95.166939	0.416566	0.523060	0.466786
10	102.736042	96.269048	0.451792	0.482171	0.466735
3	96.863897	94.199934	0.285109	0.558937	0.399196

ניתן לראות שעבור קומבינציה מספר 1 שהיא הקומבינציה $X_1=1, X_2=1, X_3=1, X_4=1, X_{13}=1, X_{14}=1$ נקבל את מידת האופטימיזציה הגבוהה ביותר. עבורה קומבינציה זו נקבל ערך חזוי של 101.0505 בדרגת החוזק המרבית וכן ערך של 86.0161 בדרגת הגמישות המזערית.

בנוסף ניתן להצביע על כך שמידת האופטימיזציה של קומבינציה זו בדרגת הגמישות היא שהביאה לאופטימיזר לבחור בה (הערך הנמוך ביותר שצפוי כתחזית בכל הניסוי).