Kísérlettervezés 2. (Design of experiments (DOE))

Bevezető

Mihalovits Máté

mihalovits@mail.bme.hu

Kísérlettervezés célja és haszna

- Folyamat megismerése, viselkedésének matematikai leírása

→ matematikai modell: a függő változó (válasz) viselkedésének leírása, hogyan függ a független változóktól (faktorok)

- Például: Kitermelés vizsgálata

  → 2 becsülendő paraméter: tengelymetszet és meredekség (faktor hatása):

  $\hat{Y} = b_0 + b_1 * hőmérséklet$

  $Y = \beta_0 + \beta_1 * hőmérséklet$

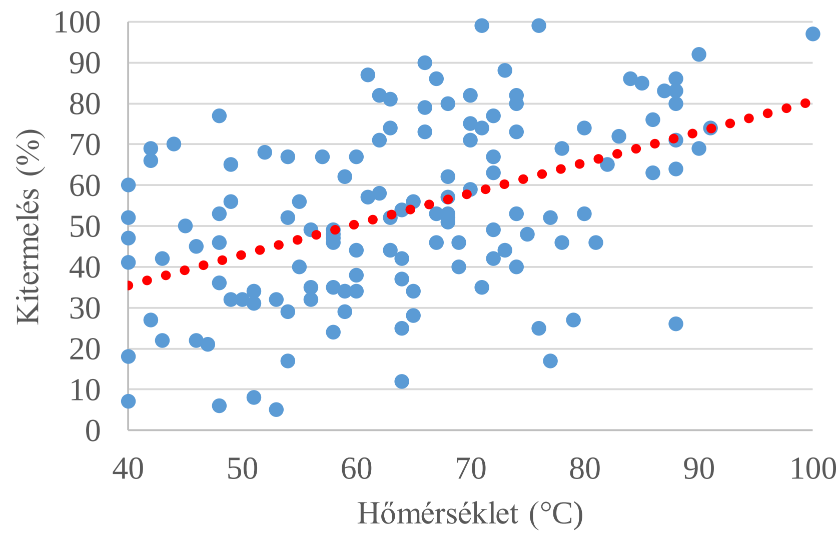
  $y_{mért} = \beta_0 + \beta_1 * hőmérséklet + \varepsilon$

  További faktorok?

- Ehhez mérések szükségesek

  1) Faktorok egyenkénti változtatása

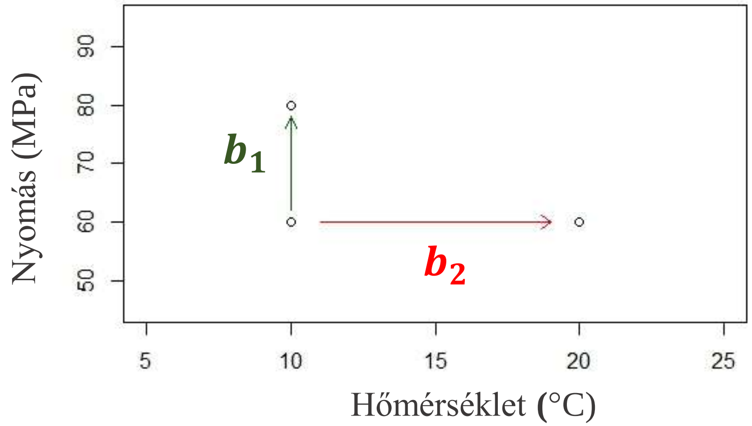
  2) Kísérlettervezés



Kísérlettervezés célja és haszna

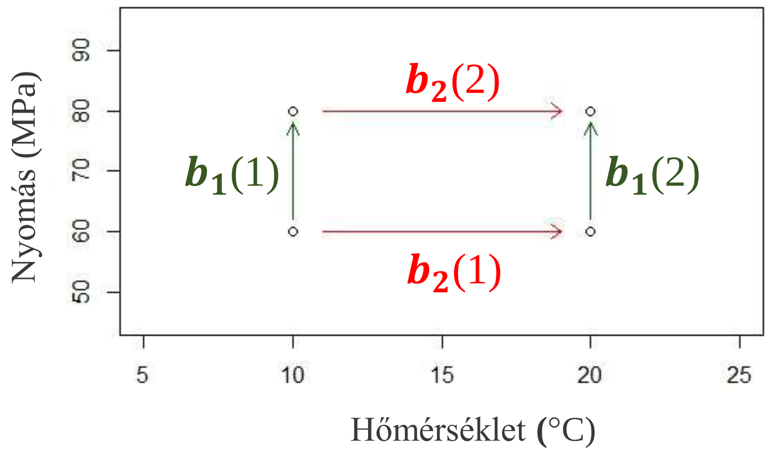
- szituáció: reakció kitermelés (y) a nyomás (z_1) és hőmérséklet (z_2) függvényében
- becsült paraméterek: b_1, b_2
- **faktorok egyenkénti változtatása:**

#	Hőmérséklet (°C)	Nyomás (MPa)
1	10	60
2	20	60
3	10	80

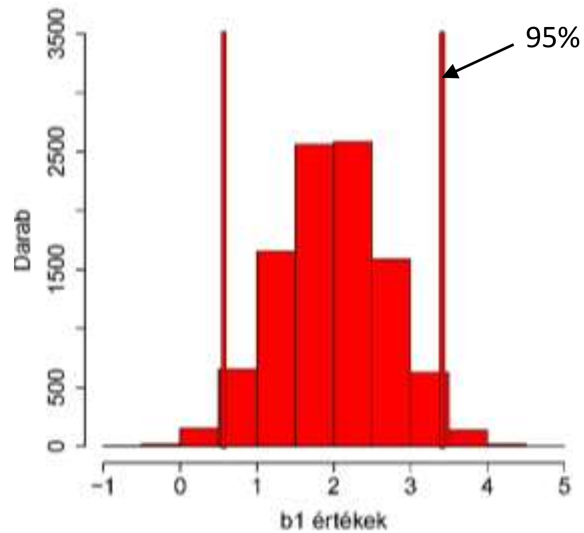


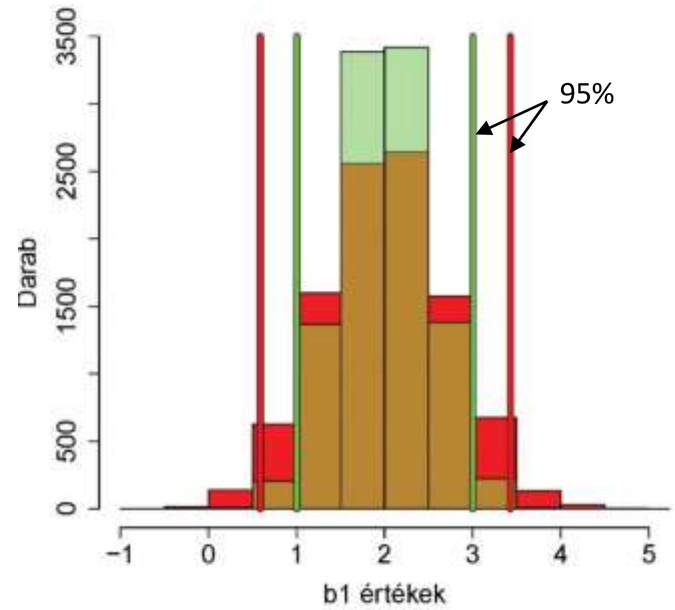
- **kísérleti terv (2^2 terv):**

#	Hőmérséklet (°C)	Nyomás (MPa)
1	10	60
2	20	60
3	10	80
4	20	80

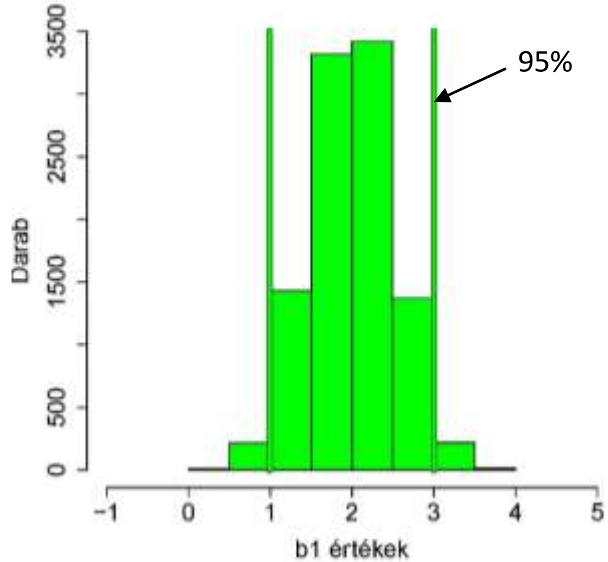


faktorok egyenkénti változtatása

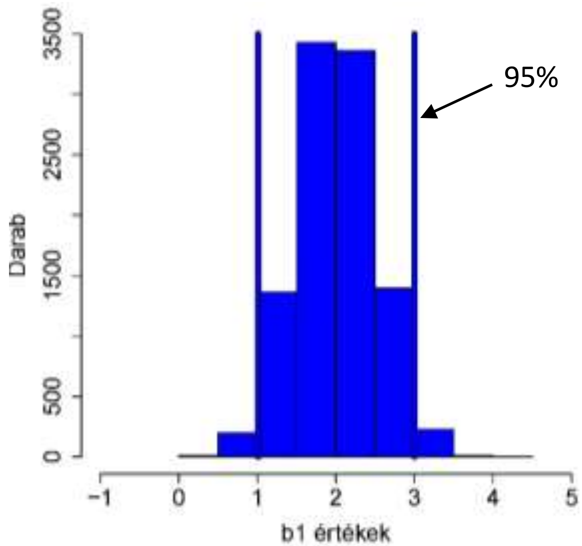




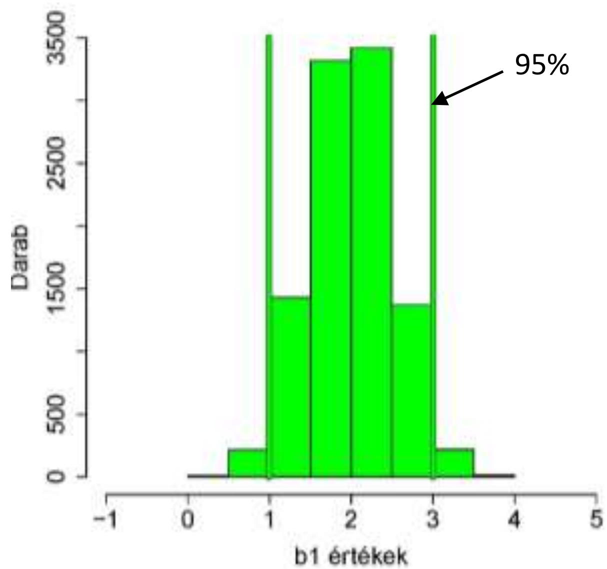
kísérleti terv

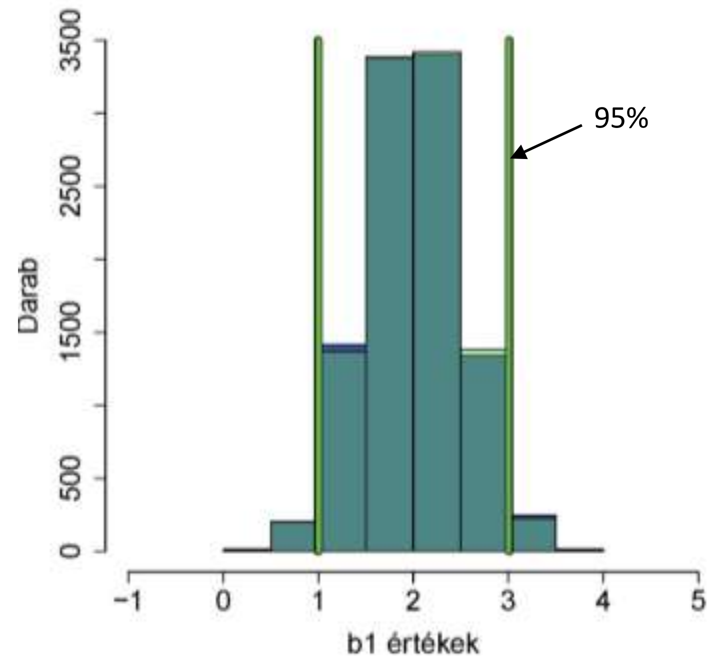


Kétszer elvégezve a faktorok egyenkénti változtatásával készült tervet (6 mérés):



kísérleti terv:





Kísérlettervezés célja és haszna

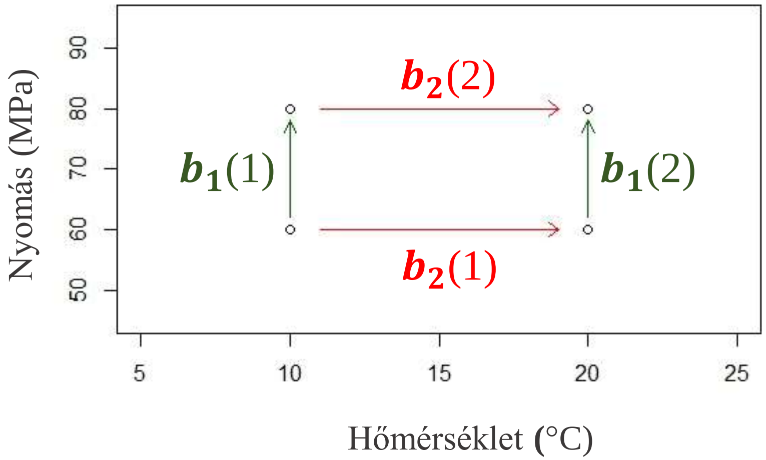
Építhető modell:

faktorok egyenkénti változtatásával:

$$\hat{Y} = \boldsymbol{b_0} + \boldsymbol{b_1 z_1} + \boldsymbol{b_2 z_2}$$

kísérleti terv:

$$\hat{Y} = \boldsymbol{b_0} + \boldsymbol{b_1 z_1} + \boldsymbol{b_2 z_2} + \boldsymbol{b_{12} z_1 z_2}$$



Összefoglalva:

1. Adott mérészám mellett a legnagyobb nyert információ (legkisebb a becsült paraméter bizonytalansága)
2. Kölcsönhatás (kcsh) kiértékelhető
3. Ortogonalitás: paraméterek (koefficiensek) egymástól függetlenül kiértékelhetőek

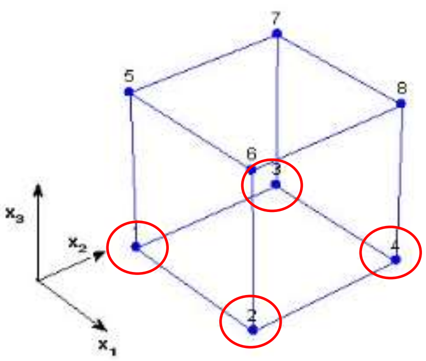
Kísérleti terv felírása

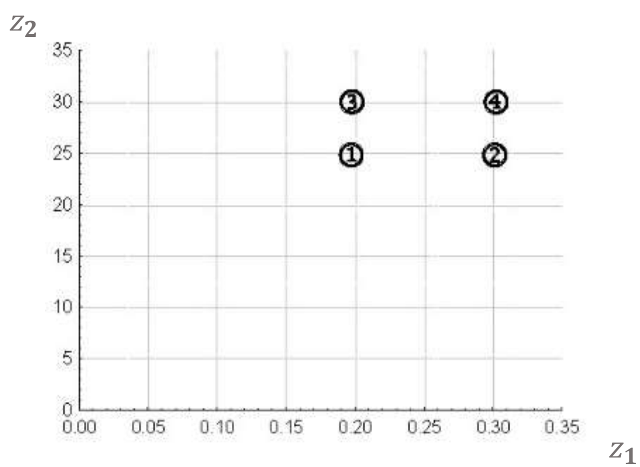
Vizsgáljuk a baracklekvár-főzés technológiai paramétereinek hatását a baracklekvár minőségére (2^3 terv).

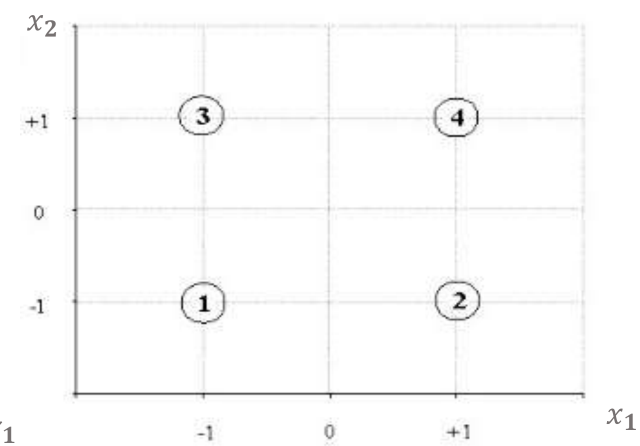
z_1 cukor mennyisége :	0.2	és	0.3 kg/kg lekvár
z_2 forralási idő:	25	és	30 perc
z_3 barack beszerzési hely:	Tesco	és	piac

Kísérleti terv felírása

- Értékek transzformálása: $x_i = \frac{z_i - z_i^0}{\Delta z_i}$







Faktor	z_1	z_2	z_3	x_1	x_2	x_3
Közép beállítás z_i^0	0.25	27.5	-	0	0	-
Variációs intervallum Δz_i	0.05	2.5	-	0.5	0.5	-
Fölső szint z_i^{max}	0.3	30	piac	+1	+1	+1
Alsó szint z_i^{min}	0.2	25	Tesco	-1	-1	-1

Kísérleti terv felírása

#	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	+	+	-	-	-	-	+	+	64
2	+	-	+	-	-	+	-	+	71
3	+	-	-	+	+	-	-	+	17
4	+	+	+	+	+	+	+	+	34,5
5	+	+	+	-	+	-	-	-	37
6	+	+	-	+	-	+	-	-	62
7	+	-	+	+	-	-	+	-	72
8	+	-	-	-	+	+	+	-	16,5

Kísérleti terv statisztikai kiértékelése

Szakmai kérdés: Van-e hatása a faktornak (a), és ha van, mekkora (b)?

→ statisztikai kérdés:

a) H_0: $\beta_i = 0$ (a faktor hatásának várható értéke 0)

b) Konfidenciaintervallum a hatás várható értékére

Kiértékelés (http://kkft.bme.hu/attachments/article/46/Kisterv%20gyakorlo_kisterv_20190905.pdf):

1) Hatások kiszámítása
2) Hatások vizsgálata:
   a) vizuális: hatás-ábra, Pareto-ábra, hatások normal plotja
   b) statisztikai: t-próba
3) Reziduum vizsgálat
4) Modell (redukált) felírása

1.) Hatások kiszámítása

#	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	+	+	-	-	-	-	+	+	64
2	+	-	+	-	-	+	-	+	71
3	+	-	-	+	+	-	-	+	17
4	+	+	+	+	+	+	+	+	34,5
5	+	+	+	-	+	-	-	-	37
6	+	+	-	+	-	+	-	-	62
7	+	-	+	+	-	-	+	-	72
8	+	-	-	-	+	+	+	-	16,5

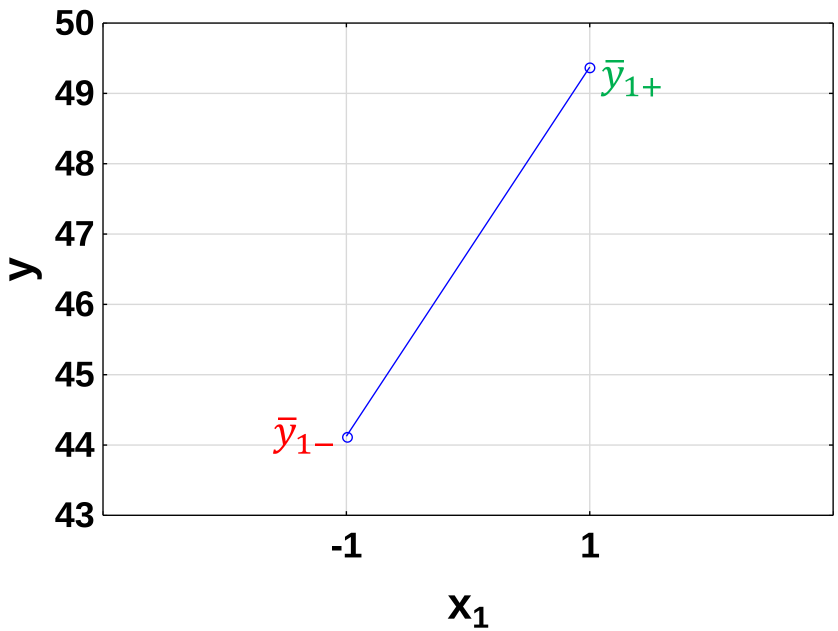
Hatások számítása:

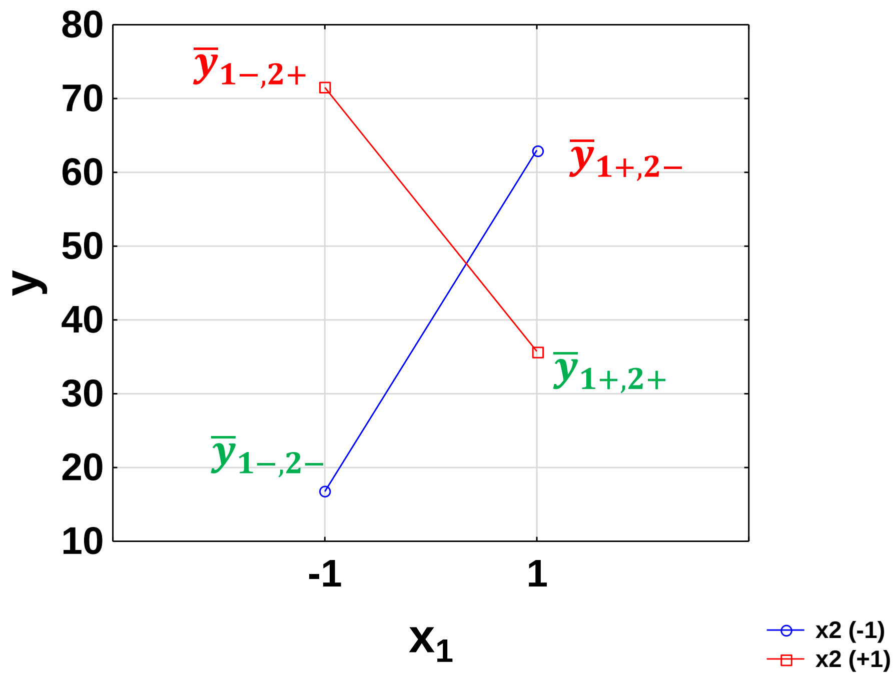
$h_i = \bar{y}_{i+} - \bar{y}_{i-}$ pl: $h_1 = \bar{y}_{1+} - \bar{y}_{1-} = \frac{64+34,5+37+62}{4} - \frac{71+17+72+16,5}{4} = 5,25$

Paraméterek becslése: $b_i = \frac{h_i}{2}$ és $b_0 = \bar{y}$ pl: $b_1 = \frac{5.25}{2} = 2,625$

2.) Hatások vizuális vizsgálata

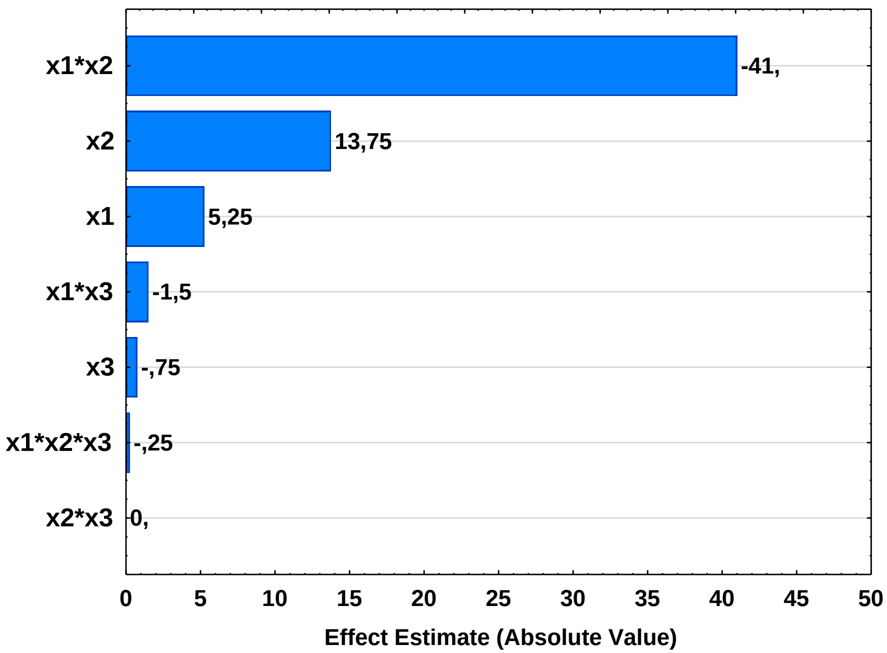
a) Hatás-ábra és kölcsönhatás-ábra



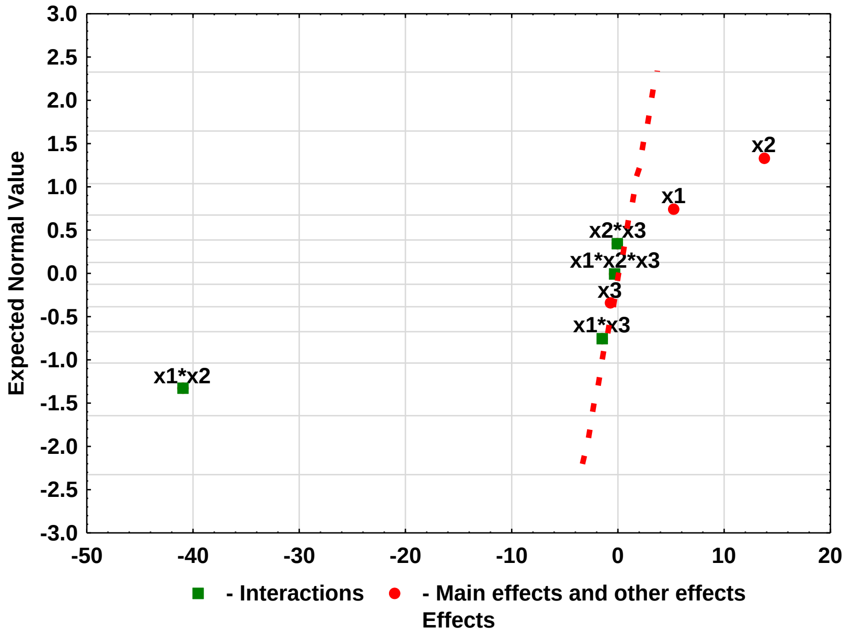


2.) Hatások vizuális vizsgálata

b) Pareto-ábra



c) normal probability plot



2.) Hatások statisztikai vizsgálata

- H_0 : E(b_i) = β_i = 0
- t-próba: $t_0 = \frac{b_i}{s_{b_i}}$, ahol $s_{b_i} = \sqrt{\frac{s_r^2}{N}}$, ahol N a pontok száma a tervben

→ t-kritikus értékhez hasonlítás: H_0 elfogadása: $-t_{(1-\frac{\alpha}{2},\nu)} < t_0 < t_{(1-\frac{\alpha}{2},\nu)}$

→ szabadsági fok (ν) = $N - l$, ahol l a paraméterek száma

- 2^2 terv: 4 mérés, 4 vizsgálandó paraméter: b_0, b_1, b_2, b_{12}

→ 0 a szabadsági fok (nem lehet szórást számolni), az illesztett felület átmegy minden ponton, túlhatározott (túldeterminált) a modell

Megoldás: a) megismételni az egész tervet → 2 * N

b) centrumponti mérés → ν = $N - l + k_c$

c) a modellből kivenni a vizuális vizsgálat alapján nem szignifikáns tago(ka)t (*modell redukálás*) → l csökken

- konfidenciaintervallum a várható értékre: $b_i \pm t_{(1-\frac{\alpha}{2},\nu)} s_{b_i}$

2.) Hatások statisztikai vizsgálata

$\boldsymbol{x_1 x_2 x_3}$ tagot kivéve a modellből, nyerünk 1 szabadsági fokot, és számolható reziduális szórásnégyzet

→ illesztésből kapjuk: $s_r^2 = 0.125$

→ paraméter szórása (Std. Err. Coeff.): $s_b = \sqrt{\frac{0.125}{8}} = 0.125$

Factor	Coeff.	Std.Err. Coeff.	t(1)	p	-95.% Cnf.Limt	+95.% Cnf.Limt
Mean/Interc.	46.7500	0.125000	374.000	0.001702	45.1617	48.3383
(1)x1	2.6250	0.125000	21.000	0.030292	1.0367	4.2133
(2)x2	6.8750	0.125000	55.000	0.011574	5.2867	8.4633
(3)x3	-0.3750	0.125000	-3.000	0.204833	-1.9633	1.2133
1 by 2	-20.5000	0.125000	-164.000	0.003882	-22.0883	-18.9117
1 by 3	-0.7500	0.125000	-6.000	0.105137	-2.3383	0.8383
2 by 3	0.0000	0.125000	0.000	1.000000	-1.5883	1.5883

3.) Reziduum vizsgálat

Regresszióanalízis feltételei:

1) A lineáris modell adekvát,

tehát $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$

hiba: $y_{mért} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$

reziduum: $y_{mért} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + \hat{\varepsilon}$

2) A hibák várható értéke 0

3) A hibák normális eloszlásúak (normal probability plot)

4) A hibák konstans varianciájúak (becsült érték vs reziduum ábra)

5) A hibák függetlenek (modellbe bele nem vett faktor vs reziduum ábra)

4.) Modell (redukált) felírása

Teljes modell:

$$\hat{Y} = 46{,}75 + 2{,}625x_1 + 6{,}875x_2 - 0{,}375x_3 - 20{,}5x_1x_2 - 0{,}75x_1x_3 + 0x_2x_3 - 0{,}125x_1x_2x_3$$

Redukált modell (nem-szignifikáns tagok elhagyása):

→hierarchia szabály: a szignifikáns kcsh-ban szereplő tagok akkor is bekerülnek főhatásként a modellbe, ha azok nem szignifikánsak

→heredity szabály: csak akkor fogadunk el kcsh-t szignifikánsnak, ha legalább az egyik benne szereplő tag főhatása szignifikáns

$$\hat{Y} = 46{,}75 + 2{,}625x_1 + 6{,}875x_2 - 20{,}5x_1x_2$$

A becsült paraméterek változatlanok redukálás után az x_i értékeket tartalmazó modellben (ortogonalitás):

Factor	Coeff.	Std.Err. Coeff.	t(1)	p	-95.% Cnf.Limt	+95.% Cnf.Limt
Mean/Interc.	46.7500	0.125000	374.000	0.001702	45.1617	48.3383
(1)x1	2.6250	0.125000	21.000	0.030292	1.0367	4.2133
(2)x2	6.8750	0.125000	55.000	0.011574	5.2867	8.4633
(3)x3	-0.3750	0.125000	-3.000	0.204833	-1.9633	1.2133
1 by 2	-20.5000	0.125000	-164.000	0.003882	-22.0883	-18.9117
1 by 3	-0.7500	0.125000	-6.000	0.105137	-2.3383	0.8383
2 by 3	0.0000	0.125000	0.000	1.000000	-1.5883	1.5883

Factor	Coeff.	Std.Err. Coeff.	t(4)	p	-95.% Cnf.Limt	+95.% Cnf.Limt
Mean/Interc.	46.7500	0.423896	110.2866	0.000000	45.5731	47.9269
(1)x1	2.6250	0.423896	6.1926	0.003457	1.4481	3.8019
(2)x2	6.8750	0.423896	16.2186	0.000085	5.6981	8.0519
1 by 2	-20.5000	0.423896	-48.3610	0.000001	-21.6769	-19.3231

Ugyanez nem igaz z_i értékeket tartalmazó modellre:

Factor	Regressn Coeff.
Mean/Interc.	-445.000
(1)x1	1975.000
(2)x2	-8.000
(3)x3	-7.000
1 by 2	43.000
1 by 3	25.000
2 by 3	0.500
1*2*3	-2.000

$\longrightarrow$

Factor	Regressn Coeff.
Mean/Interc.	-1169.5
(1)x1	4562.5
(2)x2	43.75
1 by 2	-164

+1: Validálás / prediktálás

Validálás:

Véletlenszerűen kiválasztott helyen történő méréssel ellenőrizhetjük, hogy a modell megfelelő-e $\quad H_0 : E(\hat{Y}) = E(y_{mért})$

Prediktálás:

Adott beállításokhoz (z_i) prediktálhatunk várható válasz értéket

a) $\hat{Y} = 46{,}75 + 2{,}625x_1 + 6{,}875x_2 - 20{,}5x_1x_2$ átírható az $x_i \rightarrow z_i$ transzformációval

$\hat{Y} = -1169 + 4562{,}5z_1 + 43{,}75z_1z_2 - 164z_1z_2$

VAGY

b) $z_i \rightarrow x_i$ transzformáció és számolás a $\hat{Y} = 46{,}75 + 2{,}625x_1 + 6{,}875x_2 - 20{,}5x_1x_2$ egyenlettel