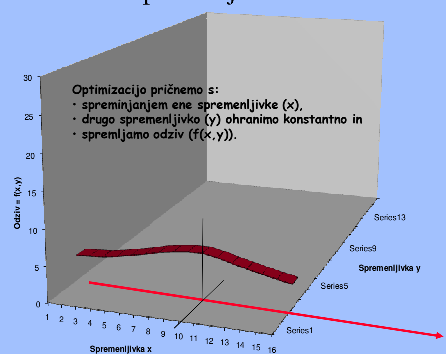
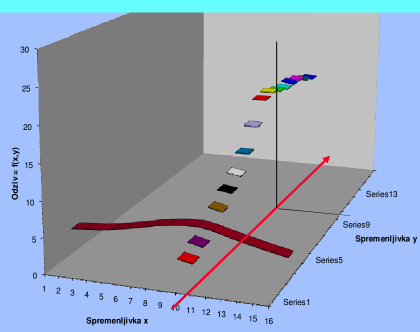
Eksperimentalni načrti

Načrtovanje eksperimentov je pomembno iz več razlogov. S pravilno izbiro eksperimentov prihranimo čas, kemikalije, izrabo opremo in število operaterjev. Pomembno je, da s primernim izborom eksperimentov pridemo do najboljšega možnega rezultata po najbolj ekonomični poti. Posebej nevarno je prepričanje, da lahko do optimalnih pogojev pridemo tako, da eksperimente izvajamo “zapovrstno” s spreminjanjem ene same spremenljivke, ostale pa držimo konstantne. Več o tem, t.i. “one-at-the-time” načinu, najdete še v poglavju o optimizacijah.

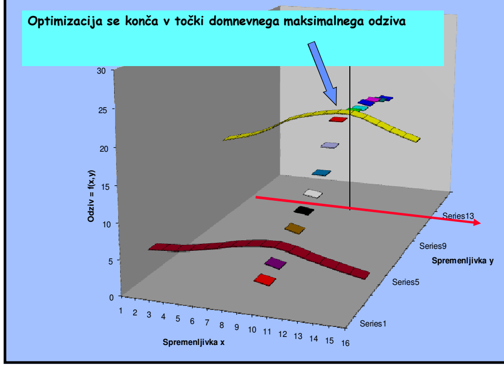
Optimizacija sistema z dvema spremenljivkama



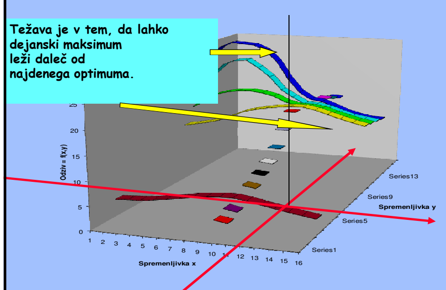
Optimizacijo nadaljujemo v točki x, kjer je odziv maksimalen:
- **spremenljivko x tokrat ohranimo konsantno,**
- **spreminjamo y in**
- **beležimo odziv.**



Optimizacija sistema z dvema spremenljivkama



Optimizacija sistema z dvema spremenljivkama



- Optimizacija posameznih spremenljivk je ena izmed najslabših optimizacijskih tehnik.
- Vodi le do lokalnih izboljšav in je močno odvisna od izbire začetnih pogojev.

Eksperimentalni načrti

Najprej moramo določiti ***spremenljivke*** (faktorje) in ***odgovore*** (responses), ki jih bomo pri eksperimentalnih načrtih obravnavali. Spremenljivke so lahko kvalitativne (npr. uporaba katalizatorja da ali ne) ali kvantitativne (numerične vrednosti koncentracije, temperature, pH itd).

Naslednja stvar je izbira ***nivojev*** (levels) vrednosti vseh izbranih spremenljivk, ki bodo vključene v eksperimentalni načrt. Pri vsaki spremenljivki moramo določiti vsaj dva nivoja. Poleg nivojev (najpomembnejši sta najvišja in najnižja vrednost) moramo določiti tudi celotno ***eksperimentalno področje*** (experimental domain).

V grobem delimo eksperimentalne načrte glede na namen uporabe v dve skupini oziroma vrsti:

a) Izbor eksperimentov, s katerim določamo ***vpliv spremenljivk*** na pregledovano (obravnavano) meritev oziroma na določen odgovor (response) ali lastnost in

b) izbor najprimernejših eksperimentov pri že izbranih spremenljivkah za izdelavo ***modelne funkcije***, ki naj napoveduje izbrano lastnost oziroma odgovor (glej poglavje o več faktorski linearni regresiji – MLR).

Poseben problem predstavlja izbor eksperimentov, iz večjega števila že opravljenih eksperimentov, ko nimamo možnosti, da bi naredili tiste, ki jih določa načrt. V takih primerih je treba narediti natanko tak eksperimentalni načrt, kot bi ga opravili za načrtovanje novih eksperimentov, potem pa poleg nivojev določiti še intervale nivojev za vsako spremenljivko posebej, označiti vse obstoječe eksperimente z oznakami nivojev oz. intervalov (++--0+…) in končno izbrati tiste eksperimente, ki se teoretičnemu načrtu najbolj približajo. Ker so nivoji spremenljivk podani z intervali, lahko v primerih, ko imamo več eksperimentov z isto nivojsko (intervalsko) oznako, izberemo tistega, ki ima vrednosti spremenljivk bliže pravim nivojskim vrednostim (bliže koncema ali sredini intervala).

Obe vrsti eksperimentalnih načrtov izhajata iz popolnih načrtov, ki zajemajo vse možne eksperimente z vsemi spremenljivkami na vseh izbranih nivojih.

Eksperimentalni ***nivo*** neke spremenljivke je vrednost, pri kateri se eksperiment izvaja. Vsaka spremenljivka mora imeti določena *najmanj dva nivoja*. Popolni eksperimentalni načrt *m* spremenljivk, od katerih ima vsaka *n* nivojev, vsebuje:

$N = n^m$ *eksperimentov.*

Če imajo posamezne spremenljivke x_i različno število nivojev n_i, potem ima popolni eksperimetalni načrt teh spremenljivk:

$N = n_1 n_2 \ \ldots \ n_m = \Pi_i n_i$ eksperimentov Oznaka Π pomeni produkt)

Poseben problem predstavlja izbor odgovorov (responses), glede na katere delamo eksperimentalni načrt. Najlaže je, če natanko vemo, za kateri kriterij moramo narediti eksperimentalni načrt, npr.: za določitev optimalni pogojev, ki bodo dali minimalni standardni odmik pri novi analitični metodi; določitev, kaj vpliva na specificirano lastnost nekega izdelka končnega izdelka; kaj vpliva na količinski izplen določene kemijske reakcije, določiti vplive na ločljivost kromatografske metode itd.).

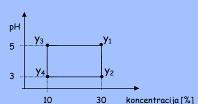
Velikokrat bi bilo treba obdelati več odgovorov hkrati. Na problem naletimo že, če moramo narediti eksperimentalni načrt za dva odgovora hkrati (npr. dve lastnosti izdelka, ki si med seboj nasprotujete, npr. odpornost na udarce in elastičnost prekrivnega laka, občutljivost in grobost fotografske emulzije, teža in mehanska odpornost embalaže, maksimalna ločljivost vrhov in minimalni retencijski čas kromatograma itd.). V takih primerih, če je le mogoče, poizkusimo najprej z ekserimentalnim načrtom za en sam odgovor (navadno je to najpomembnejša lastnost) in šele nato pričnemo iskati kompleksni kriterij, ki bi povezoval vse željene odgovore hkrati (glej poglavje o optimizaciji). Včasih zadostuje že, da kot odločujočo lastnost y_i definiramo kvocient (ali produkt) dveh ali več normaliziranih odgovorov.

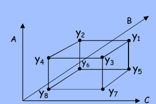
V tem poglavju se bomo držali eksperimentalnih načrtov ***za en odgovor*** (response), ki ga bomo označevali kot y_i (odgovor i-tega eksperimenta).

Dvo-nivojski eksperimentalni načrti za določanje vpliva spremenljivk

Vse spremenljivke nastopajo v eksperimentih samo z dvema vrednostima (nivojema), ki ju označimo s "+" (običajno je to najvišja ali maksimalna vrednost, ki jo spremenljivka lahko ima) in z "-" (najnižja ali minimalna vrednost spremenljivke). V anglosaksonski literaturi se v tem poglavju za *spremenljivke* vedno uporablja izraz ***faktor*** *(faktorski načrti).*

Pri *m* spremenljivkah ali faktorjih predvideva *popolni dvo-nivojski eksperimentalni načrt* (two-level full-factorial designs) 2^m eksperimentov.





A	B	C	y
+	+	+	y_1
+	+	-	y_2
+	-	+	y_3
+	-	-	y_4
-	+	+	y_5
-	+	-	y_6
-	-	+	y_7
-	-	-	y_8

Ker postanejo popolni eksperimentalni načrti prezahtevni takoj, ko imamo več kot štiri spremenljivke, večinoma uporabljamo delne eksperimentalne načrte (fractional factorial designs).

A	B	C	D	E	F	G	y
+	+	+	+	+	+	+	y_1
+	+	-	+	-	-	-	y_2
+	-	+	-	+	-	-	y_3
+	-	-	-	-	+	+	y_4
-	+	+	-	-	+	-	y_5
-	+	-	-	+	-	+	y_6
-	-	+	+	-	-	+	y_7
-	-	-	+	+	+	-	y_8

$$Vpliv\ \ faktorja\ \ i = \bar{y}_i^{+} - \bar{y}_i^{-}$$

$$Vpliv\ \ faktorja\ \ i = \frac{2}{vsi\ \ eksper.}\left(\sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ +}} y_l - \sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ -}} y_l \right)$$

$$Vpliv\ \ faktorja\ \ C = \frac{2}{8}[(y_1 + y_3 + y_5 + y_7) - (y_2 + y_4 + y_6 + y_8)]$$

Delni eksperimentalni načrt za sedem spremenljivk (faktorjev A, B, C D, E, F in G) smo dobili iz popolnega načrta za tri spremenljivke A, B in C, s tem, da smo v stolpcih D, E, F in G naredili produkte znakov: AB, AC, BC in ABC. S tem smo omogočili določitev vpliva sedmih faktorjev s samo osmimi eksperimenti, ali pa določitev interakcij (vpliv enega faktorja na drugega) med tremi prvotnimi spremenljivkami. Podobno lahko dobimo tudi večje dvo-nivojske načrte.

$$Vpliv\ \ faktorja\ \ i\ \ na\ \ j = \frac{2}{vsi\ \ eksper.}\left(\sum_{l}^{\substack{vsi\ odgovori,\ ko\ sta\ oba \\ faktorja\ na\ istem\ nivoju}} y_l - \sum_{l}^{\substack{vsi\ odgovori,\ ko\ sta\ oba \\ faktorja\ na\ razlicnih\ \ nivojih}} y_l \right)$$

$$Vpliv\ \ faktorja\ \ B\ \ na\ \ C = \frac{2}{8}[(y_1 + y_4 + y_5 + y_8) - (y_2 + y_3 + y_6 + y_7)]$$

Vpliv vsake spremenljivke je signifikanten (pri stopnji zaupanja α), če je njegova absolutna vrednost večja ali enaka intervalu zaupanja (confidence limit):

$$|Vpliv| \geq t(\alpha, N-2)s_{skupna}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \Rightarrow vpliv\ \ je\ \ signifikanten$$

Pri izračunu standardnega odmika s_{vpliva} uporabljamo tako imenovano skupno varianco (pooled variance), ki jo lahko določimo na več načinov.

$$s^2_{skupna} = \frac{\sum_{i=1}^{\frac{N}{2}} (x_{i1} - x_{i2})^2}{N} = \frac{\sum_{i=1}^{\frac{N}{2}} d_i^2}{N}$$

$$s^2_{skupna} = \frac{(n_1 - 1)s_+^2 + (n_2 - 1)s_-^2}{n_1 + n_2 - 2}$$

Na levi strani je prikazana formula izračuna skupne variance dveh skupin, ki zahteva po ***dve*** meritvi y_{l1} in y_{l2} pri vsakem od *N* eksperientov y_l. Če delamo s Plackett-Burmanovim načrtom za 8 eksperi-mentov, pomeni da je $N = 16$, $N = n_1 + n_2$ in $n_1 = n_2 = 8$, od koder sledi, da je $N = 2n$. Za vse eksperimente načrta moramo kvadrate razlik $d_l = y_{l1} - y_{l2}$ med obema meritvama sešteti.

Ciklični Plackett-Burmanovi eksperimentalni načrti.

Popolni 3-faktorski načrt ($8 = 2^3$ eksperimentov) smo z upoštevanjem štirih kombinacij povečali v 7-faktorski delni načrt, ki še vedno vsebuje le 8 eksperimentov. Podobno lahko naredimo s katerimkoli popolnim dvo-nivojskim faktorskim načrtom. Npr.: <u>popolni</u> 4-faktorski ($16 = 2^4$ ekserimentov) ali 5-faktoski dvo-nivojski načrt ($32 = 2^5$ eksperimentov) povečamo na <u>delni</u> 15- oziroma 31-faktorski načrt s tem, da dodamo vse (lahko pa tudi samo nekaj) možne kombinacije štirih oz. petih faktorjev. V teh primerih moramo seveda narediti vseh 16 ali 32 eksperimentov.

Če si želimo pri izogniti 16 eksperimentom in bi radi testirali do 11 faktorjev, lahko uporabimo delni 11-faktorski Plackett-Burmanov <u>ciklični</u> načrt, katerega prvi eksperiment je podan z vrednostmi v spodnji shemi, vse ostale eksperimente pa dobimo tako, da oznake prvega eksperimenta premikamo po eno mesto v desno:

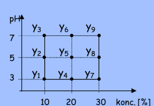
1.	+ + - + + + - - - + -	Osnovni eksperiment
2.	- + + - + + + - - - +	Vak naslednji eksperiment dobimo tako, da
3.	+ - + + - + + + - - -	prejšnjo vrsto ciklično premaknemo za eno mesto
	… … … … … …	v desno.
	… … … … … …	
10.	- + + + - - - + - + +	
11.	+ - + + + - - - + - +	Tako nadaljujemo, dokler ne dobimo 11 eksperimentov.
12.	- - - - - - - - - - -	Dvanajsti eksperiment vsebuje vse spremneljivke na nivoju minus.

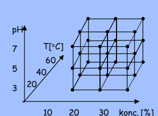
Ko je ciklični Plackett-Burmanov eksperimentalni načrt za 11 spremenljivk v celoti napisan, je za vsako spremenljivko v njem 6 plusov in 6 minusov. To pomeni, da lahko, tako kot v prejšnjih primerih, izračunamo vplive vseh spremenljivk na zelo podoben način, kot je to opisano na prejšnji strani:

$$\boldsymbol{Vpliv_i = (1/6)(\Sigma_i\, y_l^{i+} - \Sigma_i\, y_l^{i-})}.$$

Več-nivojski delni eksperimentalni načrti za modeliranje in optimizacijo odgovorov

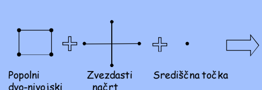
Tro-nivojski dvo-faktorski eksperimentalni načrt je verjetno edini popolni načrt, ki ga v praksi še lahko uporabljamo. Vsebuje le devet (= 3^2) eksperimentov. Že tro-nivojski tro-faktroski vsebuje kar 27 (= 3^3) eksperimentov. Ostali več-nivojski in več-faktorski pa seveda še veliko več.

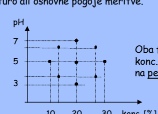




Ker je število eksperimentov preveliko, uporabljamo samo del popolnih eksperimentalnih načrtov. Najbolj znani delni (fractional) eksperimentalni načrt je ***<u>središčni sestavljeni načrt</u>*** (central composite design). Vedno je sestavljen iz treh delov:

a) <u>dvo-nivojskega načrta</u>, ki vsebuje toliko faktorjev kot menimo, da jih potrebujemo,
b) <u>zvezdastega delnega načrta</u>, ki predstavlja samo eksperimente z ekstremnimi vrednostmi in
c) <u>središčne točke</u>, ki največkrat predstavlja osnovno recepturo ali osnovne pogoje meritve.





Oba faktorja, pH in konc., sta upoštevana na <u>petih</u> nivojih

Središčni sestavljeni eksperimentalni načrti za različno število fakotrjev.

Dvo-faktorski		Tri-faktorski			m-faktorski							
x_1	x_2	x_1	x_2	x_3	x_1	x_2	x_3…	…		x_{n-1}	x_n	
0	0	0	0	0	0	0	0	…	…	0	0	središčni eksperiment (vedno en sam)
1	0	1	0	0	1	0	0	…	…	0	0	<u>zvezdasti</u> tro-nivojski *m*-faktorski načrt, ki ima vedno *2m* eksperimentov
−1	0	−1	0	0	−1	0	0	…	…	0	0	
0	1	0	1	0	0	1	0	…	…	0	0	
0	−1	0	−1	0	0	−1	0	…	…	0	0	
		0	0	1	0	0	1	…	…	0	0	
		0	0	−1	0	0	−1	…	…	0	0	
					…	…	…	…	…	…	…	
					…	…	…	…	…	0	1	
					…	…	…	…	…	0	−1	
α	α	α	α	α	α	α	α	…	…	α	α	<u>popolni</u> dvo-nivojski *m*-faktorski načrt, ki ima vedno *2^m^* eksperimentov
α	$-\alpha$	α	α	$-\alpha$	α	α	α	…	…	α	$-\alpha$	
$-\alpha$	α	α	$-\alpha$	α	α	α	α	…	…	$-\alpha$	α	
$-\alpha$	$-\alpha$	α	$-\alpha$	$-\alpha$	α	α	α	…	…	$-\alpha$	$-\alpha$	
		$-\alpha$	α	α	…	…	…	…	…	…	…	
		$-\alpha$	α	$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$	…	…	α	$-\alpha$	
		$-\alpha$	$-\alpha$	α	$-\alpha$	$-\alpha$	$-\alpha$	…	…	$-\alpha$	α	
		$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$	$-\alpha$	…	…	$-\alpha$	$-\alpha$	
$\boldsymbol{\alpha = \frac{1}{\sqrt{2}} = 0.707}$		$\alpha = \frac{1}{\sqrt[4]{8}} = 0.577$			$\alpha = \frac{1}{\sqrt[4]{2^m}}$							

V splošnem zahteva vsak središčni sestavljeni eksperimentalni načrt, ki zajame *m* spremenljivk:

$2^m + 2m + 1$ eksperimentov

Eksperiment v središčni točki navadno ponavljamo večkrat, ker iz ponovitev določimo varianco in standardni odmik

Računski primer za dvonivojski načrt - določitev vplivov posameznih faktorjev z eno meritvijo

Podatki:

A = čas ekstrakcije (ure)	1	3
B = koncentracija kisline (N)	0.05	0.01
C = material posode	steklo	celuloza
D = vrsta kisline	HCl	HCLO4
E = vrsta topila	heksan	toluen
F = temperatura (oC)	50	75
G = svetlobni pogoji	svetlo	temno

A	B	C	D	E	F	F		%ekstrakcije
1	1	1	1	1	1	1		81.0
1	1	-1	1	-1	-1	-1		75.0
1	-1	1	-1	1	-1	-1		85.0
1	-1	-1	-1	-1	1	1		60.0
-1	1	1	-1	-1	1	-1		95.0
-1	1	-1	-1	1	-1	1		96.0
-1	-1	1	1	-1	-1	1		91.0
-1	-1	-1	1	1	1	-1		89.0
-17.50	5.50	8.00	0.00	7.50	-5.50	-4.00	Vpliv	
10.97	10.40	6.22	7.39	6.40	15.28	15.94	s(+)	
3.30	14.38	15.94	16.75	16.03	9.03	8.41	s(-)	
7.13	12.39	11.08	12.07	11.21	12.16	12.17	s(+)+s(-)]/2	
1.89	2.49	2.35	2.46	2.37	2.47	2.47	s*sqrt(2/n)	
4.62	6.09	5.76	6.01	5.79	6.03	6.04	*2.447	

Kaj moramo in kaj moremo izračunati:

$$Vpliv\ faktorja\ i = \frac{2}{vsi\ eksper.}\left(\sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ +}} y_l - \sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ =}} y_l\right)$$

$$s^2_{pooled} = \frac{(s_1^2 + s_2^2)}{2}$$

$$s_a = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2}{n-1}}$$

t(0.05,6) = 2.447

$$t_{meritev} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_{pooled}\sqrt{\frac{2}{n}}} \Rightarrow |Vpliv_A| \geq t(\alpha, 2n-2) s_{pooled}\sqrt{\frac{2}{n}} \Rightarrow vpliv\ signifikanten$$

Računski primer za dvonivojski načrt - določitev vplivov posameznih faktorjev in medsebojnih vplivov

Podatki:

A	B	C	D	E	F	G	y_1	y_2	$(y_1-y_2)/2$	y_1-y_2
+	+	+	+	+	+	+	1.4	2.4	1.9	1.00
+	+	-	+	-	-	-	2.1	2.7	2.4	0.36
+	-	+	-	+	-	-	3.0	2.0	2.5	1.00
+	-	-	-	-	+	+	2.3	3.1	2.7	0.64
-	+	+	-	-	+	-	2.1	3.1	2.6	1.00
-	+	-	-	+	-	+	4.1	3.3	3.7	0.64
-	-	+	+	-	-	+	1.4	2.6	2.0	1.44
-	-	-	+	+	+	-	1.3	1.7	1.5	0.16

Kaj moramo in kaj moremo izračunati:

$|vpliv_A| = ?$ $|vpliv_{A\ na\ B}| = ?$

$|vpliv_B| = ?$ $|vpliv_{A\ na\ C}| = ?$

$|vpliv_C| = ?$ $|vpliv_{B\ na\ C}| = ?$

$|vpliv_{A,B,C}| = ?$

Skupno varianco s_{pooled} bomo računali na tri načine:
1. Iz povprečja štirih meritev v vsaki skupini,
2. Iz povprečja osmih meritev v vsaki skupini in
3. Iz dveh ponovitev vsake od osmih meritev

$s_{pooled,1} = ?$ $s_{vpliva,1} = ?$ t(0.05,14) = ?

$s_{pooled,2} = ?$ $s_{vpliva,2} = ?$ t(0.05,8) = ?

$s_{pooled,3} = ?$ $s_{vpliva,3} = ?$ t(0.05,6) = ?

Formule:

$$Vpliv\ faktorja\ i = \frac{2}{vsi\ eksper.}\left(\sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ +}} y_l - \sum_{l}^{\substack{vsi\ odgovori,\ ko\ je \\ faktor\ i\ na\ nivoju\ -}} y_l\right)$$

$$Vpliv\ faktorja\ i\ na\ j = \frac{2}{vsi\ eksper.}\left(\sum_{l}^{\substack{vsi\ odgovori,\ ko\ sta\ oba \\ faktorja\ na\ istem\ nivoju}} y_l - \sum_{l}^{\substack{vsi\ odgovori,\ ko\ sta\ oba \\ faktorja\ na\ razlicnih\ nivojih}} y_l\right)$$

$$t_{meritev} = \frac{|\bar{y}_{zg} - \bar{y}_{sp}|}{s_{pooled}\sqrt{\frac{2}{n}}} \Rightarrow |Vpliv_A| \geq t(\tfrac{\alpha}{2}, 2n-2) s_{pooled}\sqrt{\frac{2}{n}} \Rightarrow vpliv\ signifikanten$$

$$s^2_{pooled} = \frac{(s_1^2 + s_2^2)}{2} \qquad s_{pooled} = \sqrt{\frac{1}{N}\sum_{i}^{N} d_i^2} \qquad s_a = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} y_i)^2}{n-1}}$$

Računski primer za 3-faktorski petnivojski načrt - določitev eksperimentov za izdelavo modela

Tri-faktorski

x_1	x_2	x_3	
0	0	0	središčni eksperiment (vedno en sam)
1	0	0	
−1	0	0	
0	1	0	zvezdasti tro-nivojski
0	−1	0	*m*-faktorski načrt,
0	0	1	ki ima vedno
0	0	−1	*2m* eksperimentov
α	α	α	
α	α	−α	popolni dvo-nivojski
α	−α	α	*m*-faktorski načrt,
α	−α	−α	ki ima vedno 2^m eksperimentov
−α	α	α	
−α	α	−α	
−α	−α	α	
−α	−α	−α	

$$\alpha = \frac{1}{\sqrt[4]{2^m}} \quad \alpha = \frac{1}{\sqrt[4]{8}} = 0.595$$

Naloga: Pripraviti moramo petnivojski trifaktorski eksperimentalni načrt za izdelavo modela, ki bi napovedoval ločljivost HPLC metode. Faktorji, ki vplivajo na kromatogram, so podani v tabeli

Nivoji	Temperatura	pH	Koncentracija
	st. C		%
minimalni	20	4.5	5
nivo -a			
delovni	25	5	20
nivo a			
maksimalni	35	7	30

Formula za izračun vmesnih nivojev:

nivo"α" = nivo"0" + α (nivo"1" – nivo"0")
nivo"-α" = nivo"0" + α (nivo"-1" – nivo"0")