

Opisna statistika in kvaliteta procesov in meritev

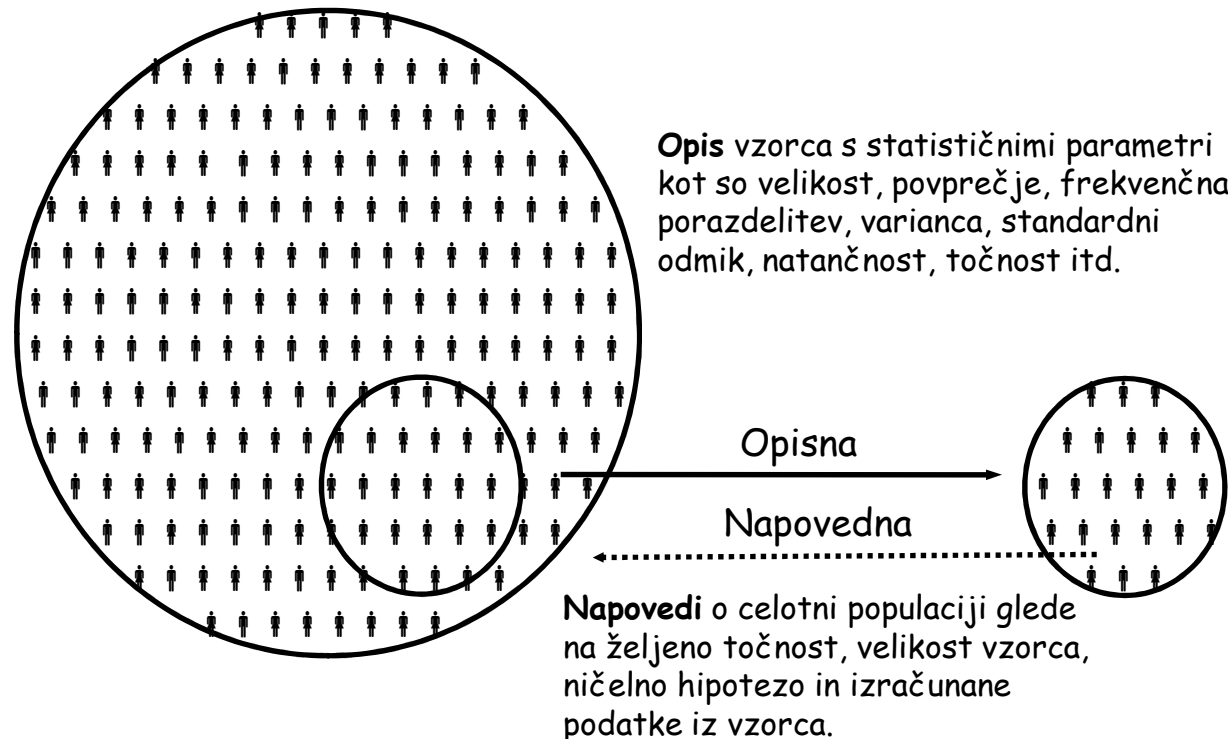
Razlika med opisno in napovedno statistiko

Populacija in vzorec

Statistika je del matematičnih ved, ki preučuje množice objektov, njihove opise in značilnosti ter ugotavlja zveze, ki veljajo med različno velikimi deli skupin istovrstnih objektov. Skupino istovrstnih objektov (kemijske analize izdelkov, živil, zdravil, procesnih postopkov, kemijskih struktur, receptur izdelkov itd.) imenujemo populacijo. Vsem objektom v populaciji lahko izmerimo eno ali več lastnosti. Numerične vrednosti posameznih lastnosti objektov v populaciji so različno porazdeljene. **Zelo malo** lastnosti je porazdeljenih v normalni ali Gaussovi porazdelitvi.

1. primer. Poznamo opis in parametre velike množice. Z merjenjem lastnosti objektov v manjšem vzorcu, želimo ugotoviti ali mali vzorec pripada veliki množici ali ne. (Ugotavljanje kompatibilnosti s standardi).

2. primer. Želimo opisati zelo veliko množico, ki nas zanima, a ne poznamo njenih statističnih parametrov. Zato izberemo manjši vzorec in na njem izmerimo in izračunamo ustrezne vrednosti. (Napoved lastnosti velike množice).



Povprečje, varianca, standardni odmik in napaka povprečja

Osnovni parametri, ki opisujejo vzorce in populacije so:

- povprečje -
- modus - vrednost, ki jo ima največ objektov (meritev) v vzorcu,
- mediana - vrednost, ki razdeli vzorec z N objekti (meritvami) na dva številčno enaka dela, ki imata po delitvi ali $N/2$ (sodi vzorci) ali $(N-1)/2$ objektov (lihi vzorci).
- varianca - povprečni kvadrat odmikov posameznih vrednosti od središča vzorca.
- standardni odmik - kvadratni koren variance.
- razpon - razlika med najmanjšo in največjo vrednostjo iste lastnosti objektov v populaciji
- napaka povprečja - je odvisna od velikosti vzorca (N), medtem ko standardni odmik ni!

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\text{modus}\{x_1, x_2, \dots, x_i, \dots, x_m\} = x_i \leftarrow \max\{f(x_i)\}$$

$$x_{\text{mediana}} = \frac{1}{2} \left[x_{\frac{N}{2}} + x_{\frac{(N+1)}{2}} \right], \text{ ali}$$

$$x_{\text{mediana}} = x_{\frac{N+1}{2}}$$

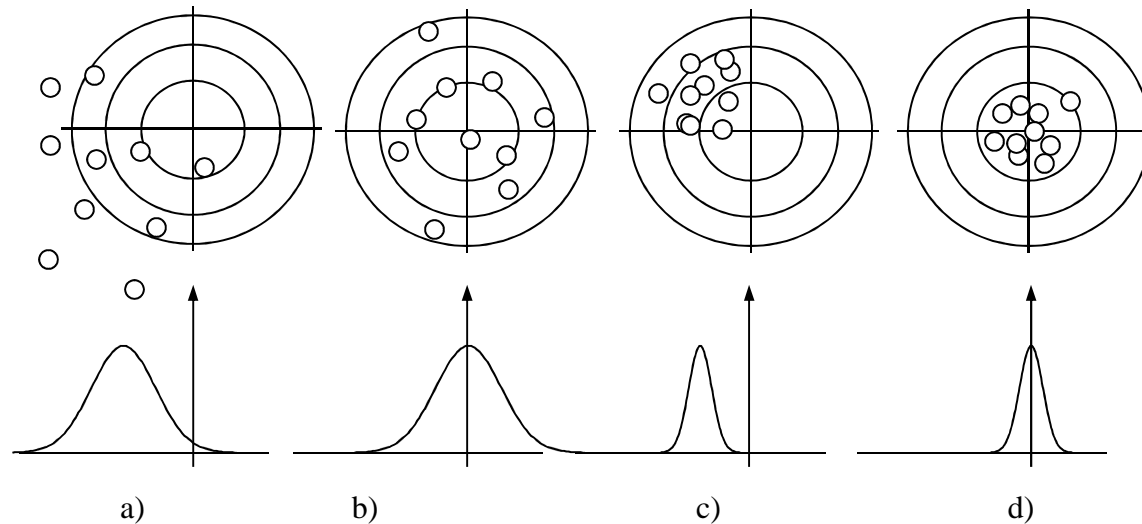
$$v = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1} = \frac{\sum_{i=1}^N x_i^2 - \frac{1}{N} \left(\sum_{i=1}^N x_i \right)^2}{N-1}$$

$$s = \sqrt[2]{v} = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} = \sqrt{\frac{\sum_{i=1}^N x_i^2 - \frac{1}{N} \left(\sum_{i=1}^N x_i \right)^2}{N-1}}$$

$$\text{Razpon} = x_{\max} - x_{\min}$$

$$s_{\bar{x}} = \frac{s}{\sqrt{N}}$$

Točnost in natančnost (accuracy and precision)



Točnost (accuracy) in natančnost (precision) meritev. Meritev na skrajni levi a) je nenatančna in netočna, naslednja b) je točna ni pa natančna, meritev c) je natančna, a ni točna in zadnja meritev d) je hkrati točna in natančna.

Točnost (accuracy) meritve je odvisna od razlike med izmerjenim povprečjem vzorca in dejansko povprečno vrednostjo populacije (tarče). Čim večja je ta razlika, tem manjša je točnost. Točnost je pogosto povezana z neznanimi napakami pri meritvah (bias), s slabimi standardi in s pojavi, ki brez naše vednosti vplivajo na meritve.

Natančnost (precision) meritve je neposredno povezana z velikostjo standardnega odmika. Čim večji je standardni odmik, tem slabša je natančnost. Natančnost je povezana z naravo meritve in jo navadno zelo dobro poznamo in tudi določimo. Natančnost povprečja vzorcev lahko popravimo z večanjem števila meritev. Natančnosti same metode pa ne moremo izboljšati.

Normalna porazdelitev ter uvod v testiranje hipotez

Ničelna hipoteza

Osnova napovedne statistike je preverjanje hipotez: njihova potrditev ali zavrnitev. Najbolj običajna hipoteza v statistiki je ničelna hipoteza H_0 , ki trdi, da med povprečno vrednostjo meritve μ_1 in poznano (standardno) vrednostjo povprečja μ_0 ni razlike.

Nasprotje ničelni hipotezi je hipoteza H_1 , imenovana tudi alternativna hipoteza in jo vedno lahko izoblikujemo v eni od treh možnih različic:

1. $\mu_1 \neq \mu_0$

2. $\mu_1 > \mu_0$

3. $\mu_1 < \mu_0$

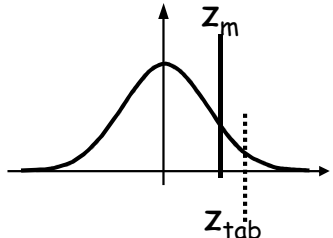
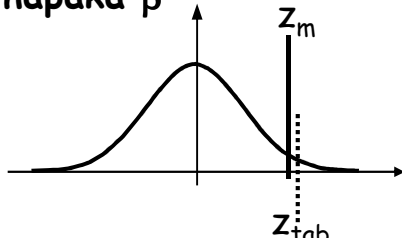
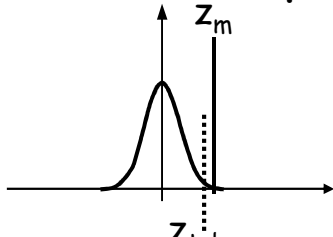
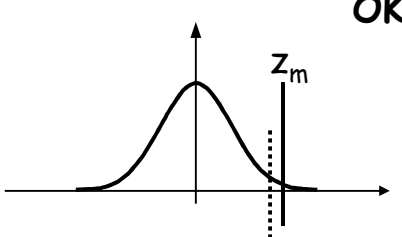
Zato da ugotovimo ali potrebujemo enostranski ali obojestranski (dvostranski) test moramo vedno oblikovati alternativno hipotezo H_1 . Prva oblika zahteva obojestranski, drugi dve pa enostranski test ničelne hipoteze.

Pri testu s katerokoli hipotezo je potrebno vnajprej predpisati kritično mejo ali interval zaupanja, se pravi tveganje α , s katerim bo odločitev o hipotezi sprejeta ali zavržena.

V največjem številu primerov privzemamo (predpišemo) 5 % tveganje.

Napaki I in II reda (napaki α in β)

Pri testiranju ničelne hipoteze H_0 imamo vedno štiri možne izide testa. Dve možnosti glede na dejansko stanje, ko H_0 drži ali ne, in dve glede na ugotovitev testa, ki H_0 bodisi potrdi bodisi ovrže.

	Dejansko stanje	
	H_0 drži	H_0 ne drži
Izid testa: H_0 drži $ z_m < z_{tab} $	OK 	napaka β 
Izid testa: H_0 ne drži $ z_m \geq z_{tab} $	napaka α 	OK 

Primerjava povprečja s kritično vrednostjo

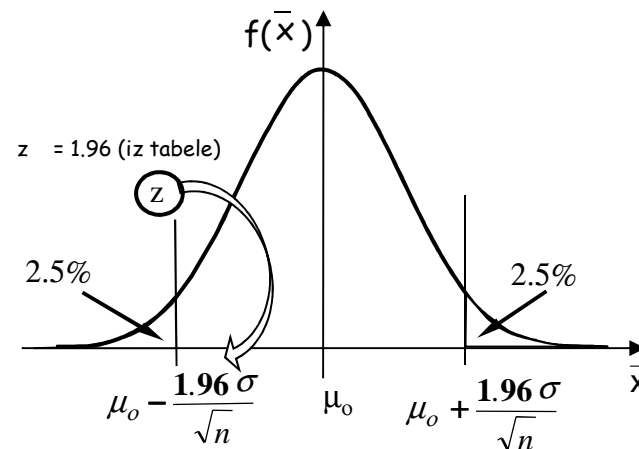
Ničelno hipotezo H_0 , da se izmerjeno povprečje \bar{x} ne razlikuje od standardne vrednosti μ_0 , lahko potrdimo s tveganjem α , če je kritična vrednost z , manjša od vrednosti v tabelah:

$$|z|_{meritev} = \frac{|\bar{x} - \mu_0|}{\sigma_{\bar{x}}} = \frac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}}$$

$|z|_{meritev} < |z|_{tabele, \alpha} \Rightarrow$ hipoteza sprejeta

$|z|_{meritv} = |t|_{meritev} < |t|_{tabele, \alpha, ps} \Rightarrow$ hipoteza sprejeta

Dvostranski test



V tabelah gledamo vrednosti pri $\alpha/2$

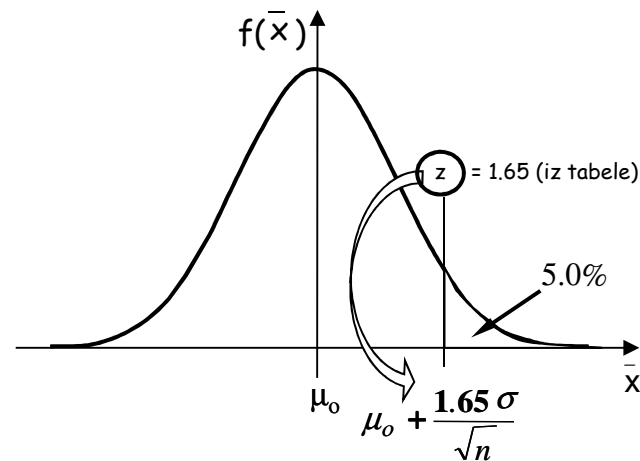
Primerjava povprečja s kritično vrednostjo

$$|z| = \frac{|\bar{x} - \mu_0|}{\sigma_{\bar{x}}} = \frac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}}$$

$|z|_{meritev} < |z|_{tabele, \alpha} \Rightarrow$ hipoteza sprejeta

$|z|_{meritv} = |t|_{meritev} < |t|_{tabele, \alpha, ps} \Rightarrow$ hipoteza sprejeta

Enostranski test



V tabelah gledamo vrednosti pri α

Primerjava dveh povprečnih vrednosti, Z-test, skupna varianca

Veliki vzorci: n_1 in $n_2 \geq 30$

$$|z| = \frac{|(\bar{x}_1 - \bar{x}_2) - \mu_0|}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$\sigma_{\bar{x}_1 \pm \bar{x}_2 \pm \bar{x}_3 \pm \dots \bar{x}_m}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \frac{\sigma_3^2}{n_3} + \dots + \frac{\sigma_m^2}{n_m}$$

$|z|_{\text{meritev}} < |z|_{\text{tabele}, \alpha} \Rightarrow$ hipoteza sprejeta

Primerjava dveh povprečnih vrednosti, t -test, skupna varianca majhnih vzorcev

Primer: majhni vzorci: n_1 in $n_2 < 30$

Prvi korak k primerjavi povprečij dveh vzorcev je ta, da preverimo ali varianci obeh vzorcev pripadata isti populaciji ali ne. Za to opravimo F -test, ki je obdelan v naslednjem poglavju.

Ko ugotovimo (z F -testom), da sta obe varianci statistično enaki (pripadata isti populaciji), lahko izračunamo skupno varianco obeh vzorcev (pooled variance za majhne vzorce) na naslednji način:

$$s_{pooled}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad \text{splošno} \Rightarrow s_{pooled}^2 = \frac{\sum_{i=1}^k (n_i - 1)s_i^2}{\sum_{i=1}^k (n_i - 1)}$$

Poseben, a zelo pogost primer je izračun skupne (pooled) variance k vzorcev, ki imajo vsak le po dve meritvi x_{i1} in x_{i2} :

$$s_{pooled}^2 = \frac{\sum_{i=1}^k (x_{i,1} - x_{i,2})^2}{2k}$$

Primerjava dveh povprečnih vrednosti, *t*-test, skupna varianca majhnih vzorcev

Primer: majhni vzorci: n_1 in $n_2 < 30$

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2 - \mu_0|}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_{pooled}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad s_{pooled}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$|t|_{meritev} < |t|_{tabele}(\alpha, n_1 + n_2 - 2) \Rightarrow$ hipoteza H_0 sprejeta

1	1.24	1.32	1.280	-0.080	0.0064	
2	1.56	1.33	1.445	0.230	0.0529	
		x1p-x2p	0.165		0.0593	vsota
					0.0148	vsota/4 = s^2
					0.1218	s(pooled)
					0.1218	s(razlike)
		t = x1p-x2p /s(razlike)	1.355			
		t tabele (0.025,2)	4.303			

Primer: Preveri ali sta povprečji dveh danih vzorcev enaki ali ne. Prvi vzorec ima pet, drugi pa štiri meritve, podane v spodnji tabeli.

1	1.79	1.32	3.2041	1.7424	
2	1.56	1.33	2.4336	1.7689	
3	1.47	1.24	2.1609	1.5376	
4	1.71	1.52	2.9241	2.3104	
5	1.52		2.3104		
Vsota	8.05	5.41	13.0331	7.3593	
Povprečje	1.61	1.35			
Vsota ² /(n-1)	12.9605	7.3170			
Varianca	0.0181	0.0141		s(pooled) ² =	0.0164
t(meritev)	2.996				
t(0.025,7)	2.365				

$$t_{meritev} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_{pooled}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$s_i^2 = \frac{\sum_{j=1}^{n_i} x_{ij}^2 - \frac{1}{n_i} \left(\sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i - 1}$$

$$s_{pooled}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Število prostostnih stopenj je $n_1 + n_2 - 2$

$$t_{meritev} < t_{\alpha, n_1 + n_2 - 2} \Rightarrow \text{hipoteza sprejeta}$$

Primerjava razpršenosti ali primerjava dveh varianc, F -test

Da lahko statistične primerjave naredimo, moramo najprej preveriti hipotezo ali sta oba vzorca vzeta iz iste populacije ali ne.

Test za primerjavo varianc se imenuje primerjava razpršenosti ali F -test (črka F je uporabljena v čast statistiku Fisherju) .

$$F = \frac{s_1^2}{s_2^2} < F_{\text{tabela}}(\alpha, n_1 - 1, n_2 - 1) \Rightarrow \text{vzorca pripadata isti populaciji}$$

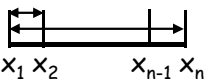
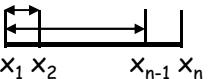
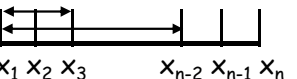
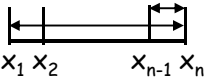
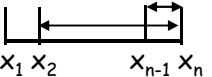
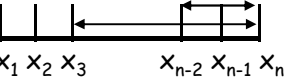
Ker damo v števec vedno večji standardni odmik, je vrednost F vedno večja ali kvečejmu enaka ena. O F -testu bomo govorili še pri poglavju o analizi variance (ANOVA) in poglavju o kalibracijski premici.

Test za odkrivanje ubežnike (outlier test), Dixonov in Grubbsov test

Pri večini analitičnega dela računamo povprečja iz meritev pri katerih se pogosto pojavi vprašanje ali je katera od meritev (največja ali najmanjša) "ubežnik" oziroma outlier in bi jo bilo treba ponovno preveriti, ali morda pri računu povprečja in standardnega odmika celo izpustiti. Poleg zelo priljubljenega Dixonovega testa, je bil z ISO 5725-2:1994(E) normo predlagan Grubbsov test ubežnikov, ki je ravno tako enostaven, a nima nekaterih pomanjkljivosti Dixonovega testa (n.pr. občutljivosti na premike vrednosti znotraj enakega razpona).

Dixonov test, ISO 5725-1986(E), se sedaj večinoma uporablja le pri medlaboratorijskih testih.

Oba testa predpostavljata, da je vseh n vrednosti x_i v vzorcu urejenih po velikosti: od najmanjšega x_1 , do največjega x_n .

Dixon			Grubbs	
Q_{10} , $n=3..7$	Q_{11} , $n=8..12$	Q_{22} , $n=13..30$	G_1 (single)	G_2 (pair)
 $x_1 \ x_2 \ \dots \ x_{n-1} \ x_n$	 $x_1 \ x_2 \ \dots \ x_{n-1} \ x_n$	 $x_1 \ x_2 \ x_3 \ \dots \ x_{n-2} \ x_{n-1} \ x_n$	$G_1 = \frac{x_1 - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=3}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
 $x_1 \ x_2 \ \dots \ x_{n-1} \ x_n$	 $x_1 \ x_2 \ \dots \ x_{n-1} \ x_n$	 $x_1 \ x_2 \ x_3 \ \dots \ x_{n-2} \ x_{n-1} \ x_n$	$G_1 = \frac{x_n - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
$Q_{10} = \frac{x_2 - x_1}{x_n - x_1}$	$Q_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1}$	$Q_{22} = \frac{x_3 - x_1}{x_{n-2} - x_1}$		
$Q_{10} = \frac{x_n - x_{n-1}}{x_n - x_1}$	$Q_{11} = \frac{x_n - x_{n-1}}{x_n - x_2}$	$Q_{22} = \frac{x_n - x_{n-2}}{x_n - x_3}$		

Q_{ij} in D_i tabele so na voljo za vrednosti $\alpha = 0.05$ in 0.01 ter za velikosti vzorcev od 3 do 30 najdete v knjigi: Massart,... (Part A), strani 111 in 113

Določanje ubežnikov

Ugotovite, ali so v nizu vrednosti: 22.1, 22.4, 22.9, 23.0, 23.5, 23.7, 23.9, 26.5, ubežniki ali ne.

Grubbs	
G_1 (en ubežnik)	G_2 (dva ubežnika)
$G_1 = \frac{x_1 - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=3}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
$G_1 = \frac{x_n - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
Dixon	
Ker imamo 8 meritev uporabimo:	
$Q_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1}$	$Q_{11} = \frac{x_n - x_{n-1}}{x_n - x_2}$

Meritve	x_i	
1	22.1	
2	22.4	
3	22.9	
4	23.0	
5	23.5	
6	23.7	
7	23.9	
8	26.5	
povprečje μ_0		23.5
povprečje brez prvih dveh		23.9
povprečje brez zadnjih dveh		22.9
Dixonov test za 8 meritev		
Tabela: Q11(0.05,8)		0.608
$(x_2 - x_1)/(x_{n-1} - x_1)$		0.167
$(x_n - x_{n-1})/(x_n - x_2)$		0.634
Grubsov test za enega ubežnika		
Tabela: G1(0.05,8)		2.126
$(x_1 - \mu_0)/s_0$		1.028
$(x_n - \mu_0)/s_0$		2.203

Meritve	x_i	$(x_i - \mu_0)^2$	$[x_i - \mu(-p_2)]^2$	$(x_i - \mu(-z))^2$
1	22.1	1.960		0.694
2	22.4	1.210		0.284
3	22.9	0.360	1.034	0.001
4	23.0	0.250	0.840	0.004
5	23.5	0.000	0.174	0.321
6	23.7	0.040	0.047	0.588
7	23.9	0.160	0.000	
8	26.5	9.000	6.674	
		SS0	SS(1,2)	SS(n,n-1)
		12.980	8.768	1.893
Tabela: G2(0.05,8)		0.110		
G2 = SS(1,2)/SS0		0.676	nista ubežnika	
G2 = SS(n,n-1)/SS0		0.146	nista ubežnika	

Za test ubežnikov moramo urediti vrednosti po velikosti od najmanjšega x_1 , do največjega x_n .

Če je $Q_{ij}^{\text{iz meritev}} > Q_{ij}^{\text{tabele}}$ je vzorec ubežnik
 Če je $G_1^{\text{iz meritev}} > G_1^{\text{tabele}}$ je vzorec ubežnik
 Če je $G_2^{\text{iz meritev}} < G_2^{\text{tabele}}$ sta vzorca ubežnika

Dixonov test
 Grubbsov test
 Grubbsov test

Rezultat Dixonovega testa: najmanjša vrednost ni, največja pa je ubežnik.

Rezultat Grubsovega testa: prva vrednost ni, zadnja pa je ubežnik; niti zadnji dve meritvi niti prvi dve nista ubežnika!

Ugotovite, ali so v nizu vrednosti: 21.1, 22.4, 22.9, 23.0, 23.5, 23.7, 26.0, 26.5, ubežniki ali ne.

Grubbs	
G_1 (en ubežnik)	G_2 (dva ubežnika)
$G_1 = \frac{x_1 - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=3}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
$G_1 = \frac{x_n - \bar{x}}{s}$	$G_2 = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$
Dixon	
Ker imamo 8 meritev uporabimo:	
$Q_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1}$	$Q_{11} = \frac{x_n - x_{n-1}}{x_n - x_2}$

Dixon-ov test

$$Q_{11} = (26.5 - 26.0)/(26.5 - 22.4) = 0.122$$

Grubbs-ov test

$$G = SS_{7,8}/SS_0 = 1.89/18.52 = 0.1021$$

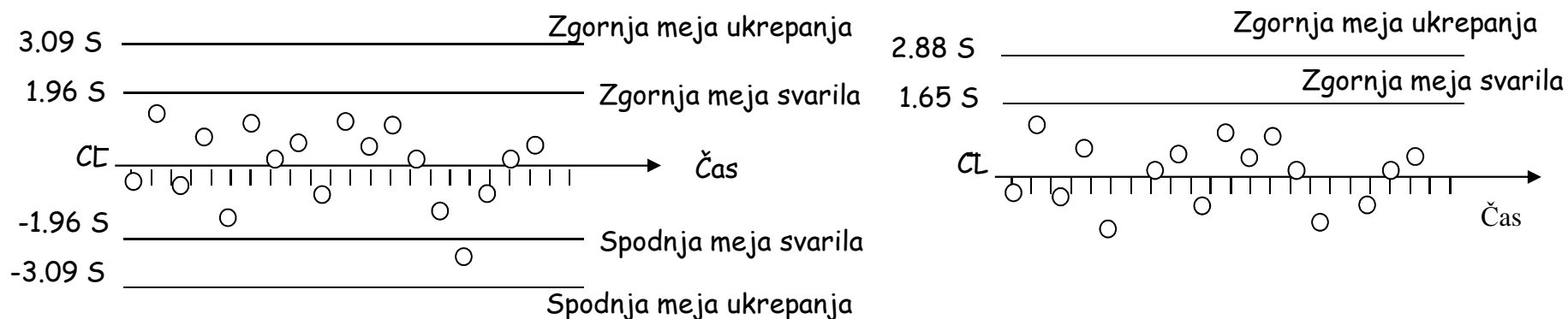
Rezultat Dixonovega testa: niti najmanjša vrednost, niti največja nista ubežnika.

Rezultat Grubsovega testa: zadnji dve meritvi sta ubežnika!

Kontrolne karte (control charts)

Kontrolne karte je že leta 1931 razvil Shewhart in so osnova kontrole kvalitete in statistične procesne kontrole. Glavni cilj take kontrole je oceniti ali se proces nahaja v okviru statistično določenih meja, z drugimi besedami ali se centralna točka sistema in njegova disperzija bisveno spreminjata tekom časovne skale. S tako kontrolo dejansko spremljamo spremembe sistematičnih in naključnih napak.

Za doseg omenjenega cilja si izberemo n objektov ali vzorcev iz produkcijske linije in jim izmerimo kvaliteto oz. količino, ki bo indikator stabilnosti procesa proizvodnje. Izračunamo centralno točko in disperzijo in ju narišemo vzdolž časovne skale. Najpogosteje uporabljamo povprečne kontrolne karte 'mean charts'.



Oddaljenost meje ukrepanja in meje svarila od povprečne oziroma standardne vrednost \bar{CL} , sta pri enostranski in obojestranski kontroli različni. Obe sta prirejeni za 95 oziroma 99.8 % verjetnost. Centralno linijo (CL) navadno določimo kot povprečje meritev vsaj 10 vzorcev, vendar je priporočljivo uporabiti 20 vzorcev. S^2 je lahko skupna (pooled) varianca varianc vseh vzorcev, lahko je to povprečna varianca posameznih varianc vzorcev ali pa se izračuna iz območja vseh meritev z uporabo Hartleyjeve konstante (Massart, Part A, str. 153).

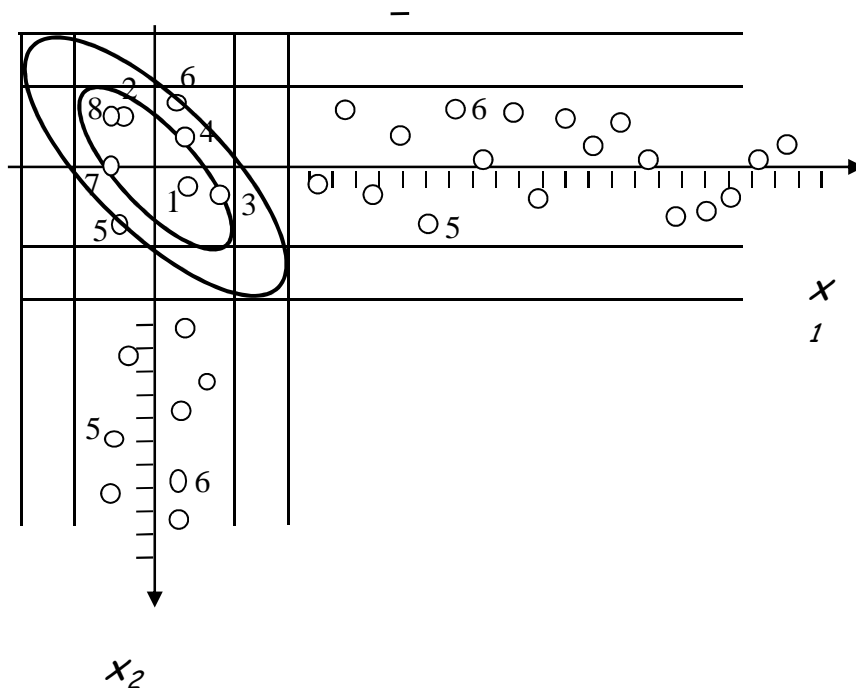
Povprečne kontrolne karte nam pomagajo pri odkrivanju naslednjih pojavov:

- pojav pristaranih (biased) vrednosti,
- cikličnih in periodičnih sprememb,
- pojav trendov (drift).

Pravila za ukrepanje:

- točka pade izven meja ukrepanja,
- dve zaporedni točki padeta izven meja svarila,
- sedem zaporednih točk je na isti strani CL ali 10 od 11 točk je na isti strani CL,
- sedem zaporednih točk kaže naraščanje.

Dobro znan nabor pravil za ukrepanje je poznan pod imenom 'Western Electric rules'.

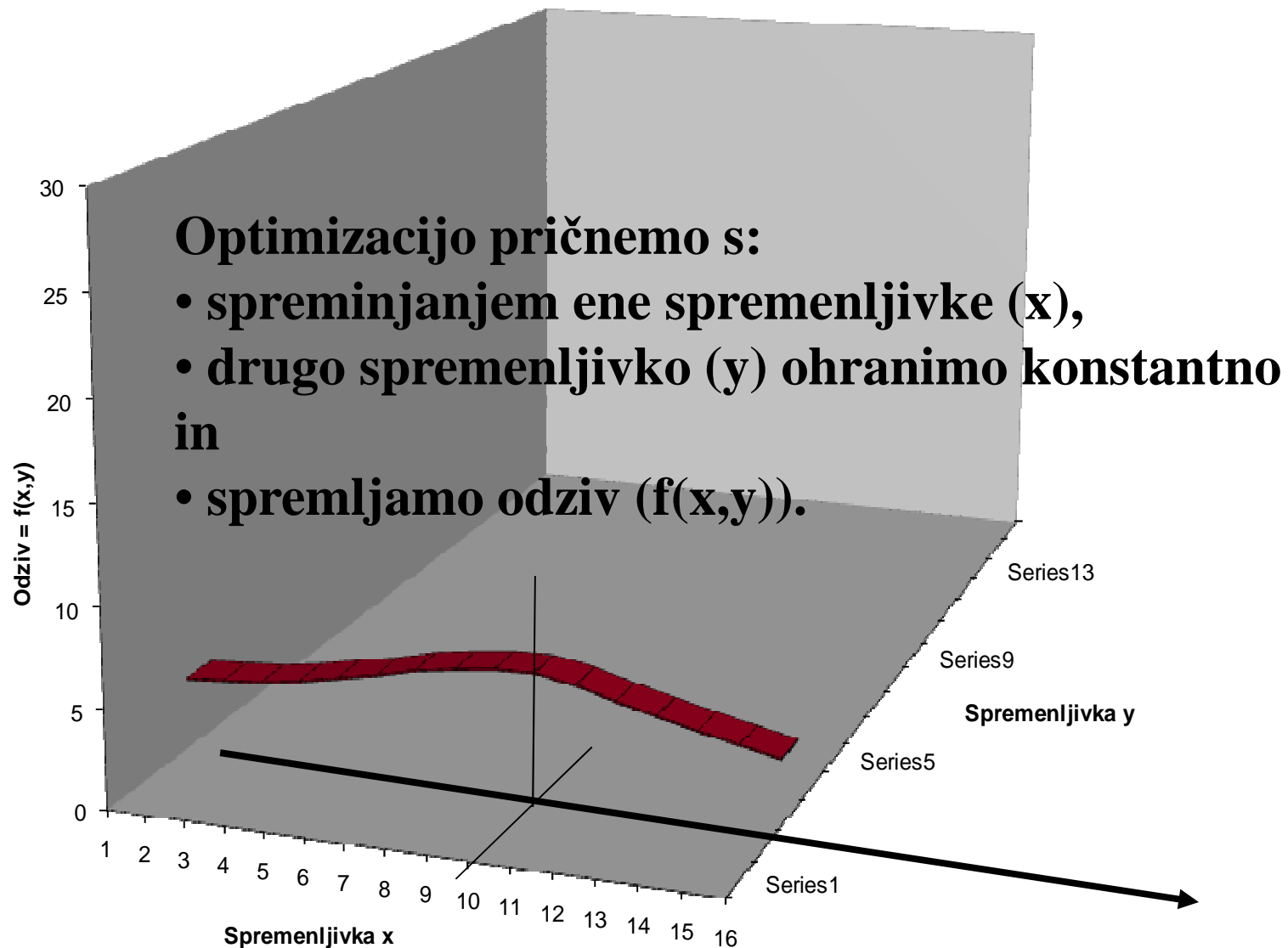


Pomembna je tudi hkratna kontrolna karte za spremljanje dveh ali celo več spremenljivk istega procesa ali postopka. Izmerjene vrednosti dveh spremenljivk x_1 in x_2 vnašamo vsako v svojo kontrolno karto, ki sta postavljeni pravokotno druga na drugo. Hkrati delamo s pomočjo obojnih vrednosti x še 2d-projekcijo (zgoraj levo) iz katere lahko vidimo, kdaj padejo nekatera stanja procesa ali postopka preko svarilne ali celo preko akcijske meje, čeprav je vsaka posamezna spremenljivka še znotraj lastnega intervala zaupanja. S spremljanjem vsake kontrolne karte posebej, takih anomalij ne bi mogli odkriti.

Eksperimentalni načrti

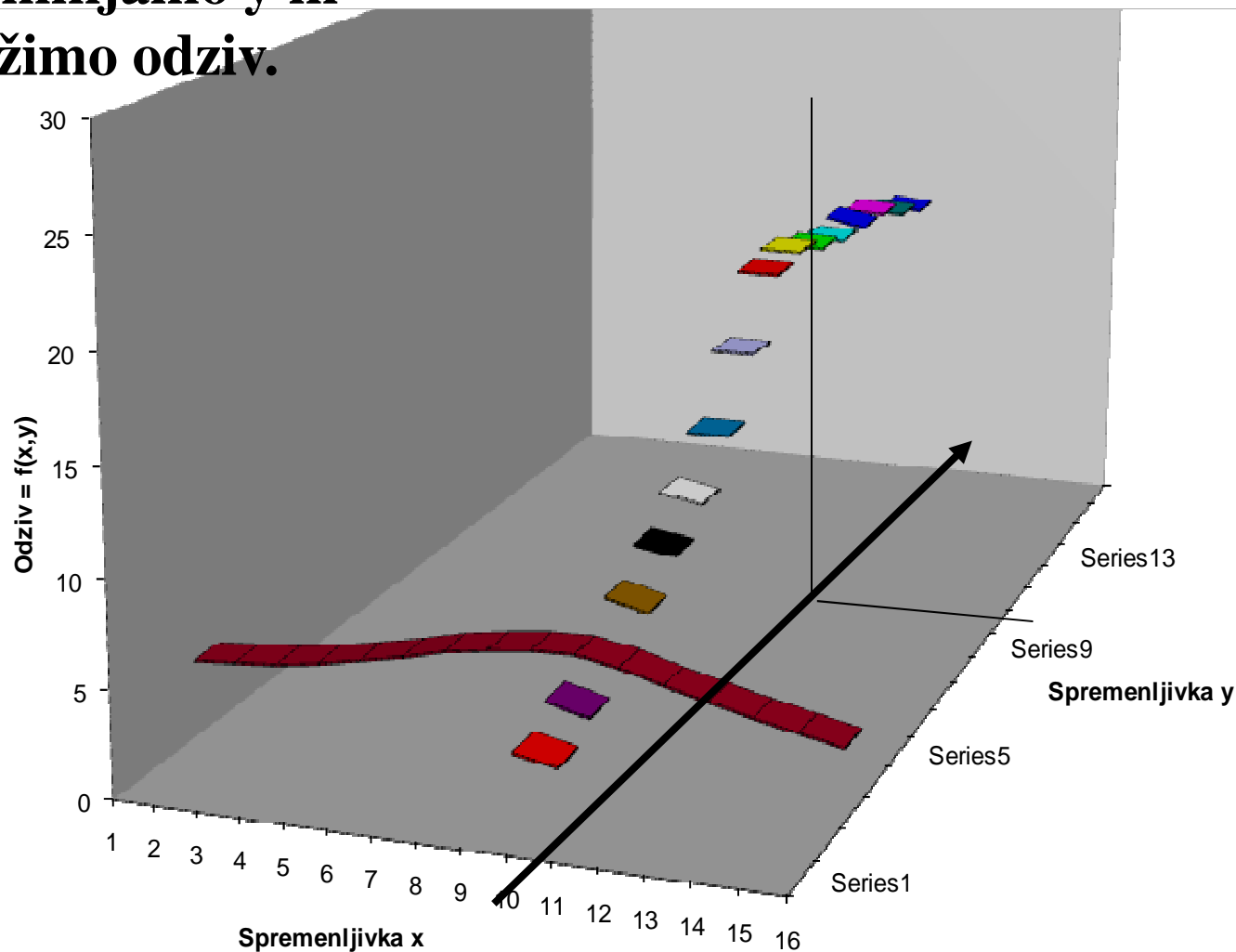
Načrtovanje eksperimentov je pomembno iz več razlogov. S pravilno izbiro eksperimentov prihranimo čas, kemikalije, izrabo opreme in število operaterjev. Pomembno je, da s primernim izborom eksperimentov pridemo do najboljšega možnega rezultata po najbolj ekonomični poti. Posebej nevarno je prepričanje, 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



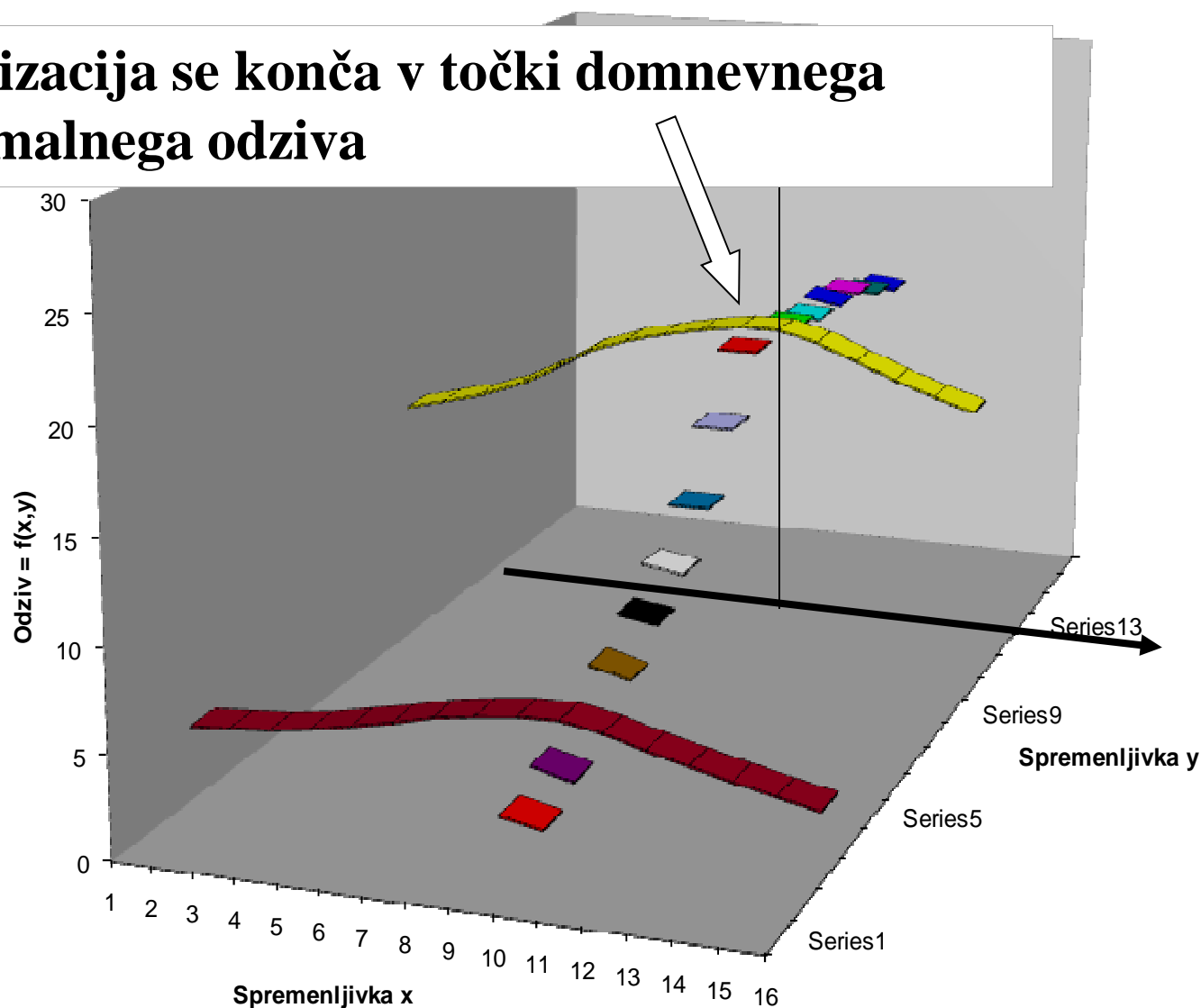
Optimizacijo nadaljujemo v točki x, kjer je odziv maksimalen:

- **spremenljivko x tokrat ohranimo konsantno,**
- **spreminjamo y in**
- **beležimo odziv.**



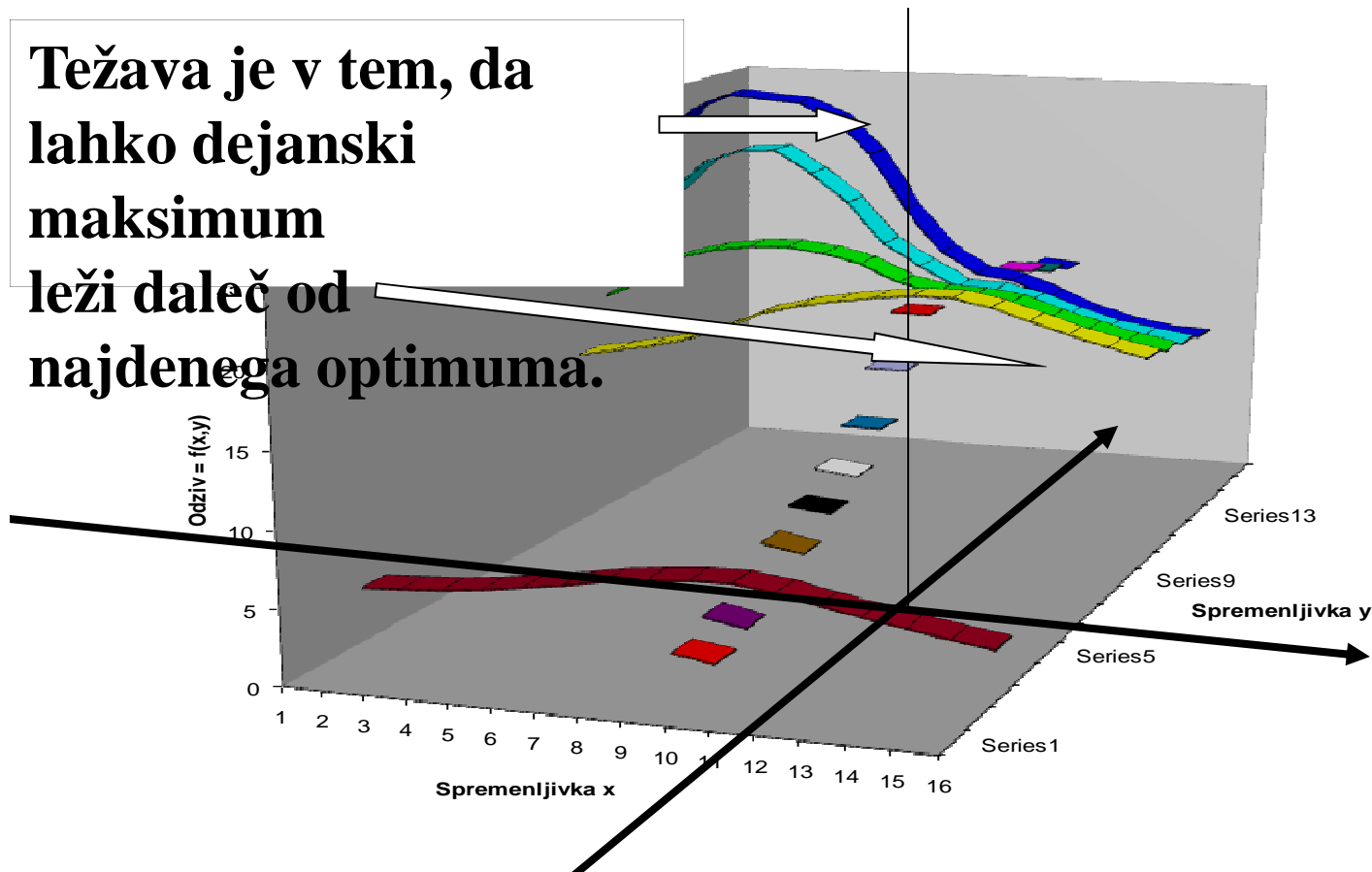
Optimizacija sistema z dvema spremenljivkama

Optimizacija se konča v točki domnevnega maksimalnega odziva



Optimizacija sistema z dvema spremenljivkama

Težava je v tem, da lahko dejanski maksimum leži daleč od najdenega optimuma.



- 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 (intervalno) 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 eksperimentalni načrt teh spremenljivk:

$N = n_1 n_2 \dots n_m = \prod_i n_i$ eksperimentov (Oznaka \prod 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.).

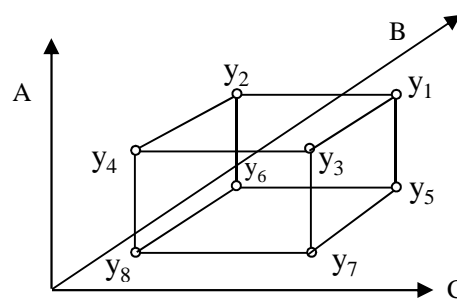
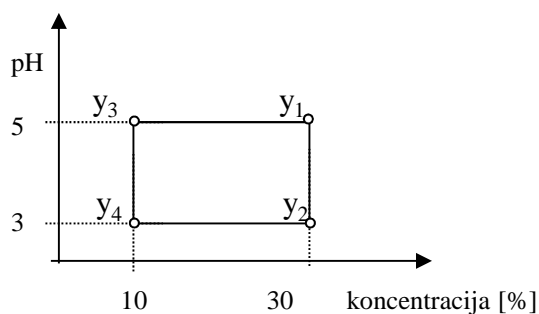
Velikokrat bi bilo treba obdelati več odgovorov hkrati. Na problem naletimo že, če moramo narediti eksperimentalni načrt za dva odgovora hkrati (npr. dve lastnosti izdelka, ki si med seboj nasprotujeta, npr. odpornost na udarce in elastičnost prekrivnega laka, občutljivost in grobost fotografske emulzije, teža in mehanska odpornost embalaže, maksimalna ločljivost vrhov in minimalni retencijski čas kromatograma itd.). V takih primerih, če je le mogoče, poizkusimo najprej z eksperimentalnim 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

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

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

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

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.

$$\text{Vpliv faktorja } i \text{ na } j = \frac{2}{\text{vsi eksper.}} \left(\overset{\text{vsi odgovori, ko sta oba faktorja na istem nivoju}}{\sum_l y_l} - \overset{\text{vsi odgovori, ko sta oba faktorja na različnih nivojih}}{\sum_l y_l} \right)$$

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

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

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

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

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

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

Na levi strani je prikazana formula izračuna skupne variance dveh skupin, ki zahteva po **dve** meritvi y_{i1} in y_{i2} pri vsakem od N eksperimentov y_i . Če delamo s Plackett-Burmanovim načrtom za 8 eksperimentov, 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 $d_i = y_{i1} - y_{i2}$ 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 katerikoli popolnim dvo-nivojskim faktorskim načrtom. Npr.: popolni 4-faktorski ($16 = 2^4$ eksperimentov) ali 5-faktorski dvo-nivojski načrt ($32 = 2^5$ eksperimentov) povečamo na delni 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 izoginiti 16 eksperimentom in bi radi testirali do 11 faktorjev, lahko uporabimo delni 11-faktorski Plackett-Burmanov ciklični 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:

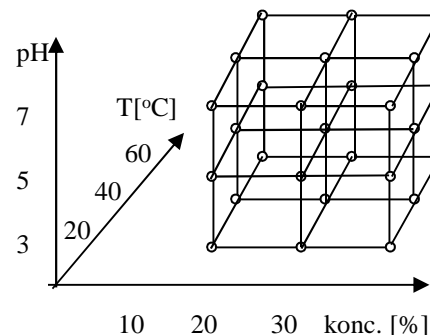
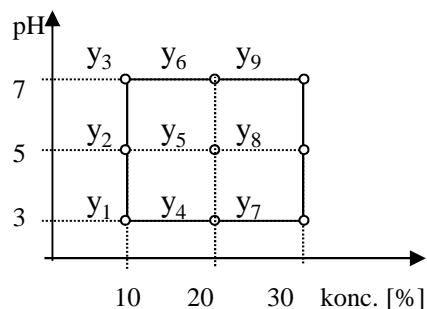
1. + + - + + + - - - + - Osnovni eksperiment
2. - + + - + + + - - - + Vak naslednji eksperiment dobimo tako, da
3. + - + + - + + + - - - prejšnjo vrsto ciklično premaknemo za eno mesto
... .. v desno.
... ..
10. - + + + - - - + - + +
11. + - + + + - - - + - + Tako nadaljujemo, dokler ne dobimo 11 eksperimentov.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

$$Vpliv_i = (1/6)(\sum_i y_i^{i+} - \sum_i y_i^{i-}).$$

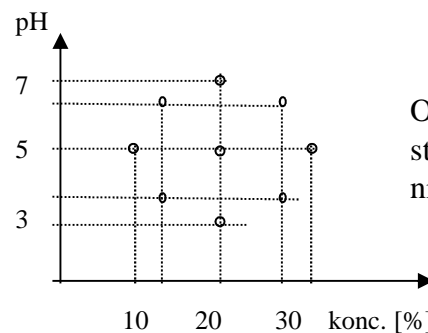
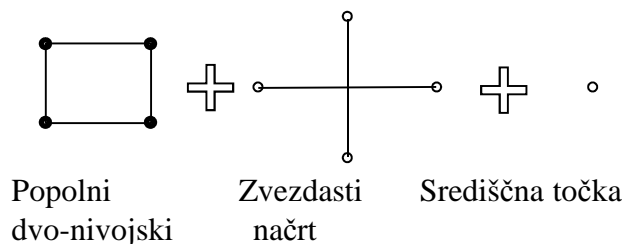
Več-nivojski delni eksperimentalni načrti za modeliranje in optimizacijo odgovorov

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-faktorski 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 *središčni sestavljeni načrt* (central composite design). Vedno je sestavljen iz treh delov:

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



Oba faktorja, pH in konc., sta upoštevana na petih nivojih

Središčni sestavljeni eksperimentalni načrti za različno število faktorjev.

| Dvo-faktorski | | Tri-faktorski | | | m-faktorski | | | | | | |
|---------------|-----------|---------------|-----------|-----------|-------------|-----------|-----------------------|-----------|-----------|-----------|--|
| x_1 | x_2 | x_1 | x_2 | x_3 | x_1 | x_2 | $x_3 \dots$ | \dots | x_{n-1} | x_n | |
| 0 | 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 | 0 | } zvezdasti tro-nivojski
<i>m</i> -faktorski načrt,
ki ima vedno

<i>2m</i> eksperimentov |
| -1 | 0 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | |
| 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | |
| 0 | -1 | 0 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | |
| | | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | |
| | | 0 | 0 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | |
| | | | | | ... | ... | | ... | ... | ... | |
| | | | | | ... | ... | | 0 | 1 | 1 | |
| | | | | | ... | ... | | 0 | 0 | -1 | |
| α | α | α | α | α | α | α | $\alpha \dots \dots$ | α | α | α | } popolni dvo-nivojski
<i>m</i> -faktorski načrt,
ki ima vedno

<i>2^m</i> eksperimentov |
| α | $-\alpha$ | α | α | $-\alpha$ | α | α | $\alpha \dots \dots$ | α | $-\alpha$ | $-\alpha$ | |
| $-\alpha$ | α | α | $-\alpha$ | α | α | α | $\alpha \dots \dots$ | $-\alpha$ | α | α | |
| $-\alpha$ | $-\alpha$ | α | $-\alpha$ | $-\alpha$ | α | α | $\alpha \dots \dots$ | $-\alpha$ | $-\alpha$ | $-\alpha$ | |
| | | $-\alpha$ | α | α | ... | ... | | ... | ... | ... | |
| | | $-\alpha$ | α | $-\alpha$ | $-\alpha$ | $-\alpha$ | $-\alpha \dots \dots$ | α | $-\alpha$ | $-\alpha$ | |
| | | $-\alpha$ | $-\alpha$ | $-\alpha$ | $-\alpha$ | $-\alpha$ | $-\alpha \dots \dots$ | $-\alpha$ | $-\alpha$ | $-\alpha$ | |

$$\alpha = \frac{1}{\sqrt{2}} = 0.707$$

$$\alpha = \frac{1}{\sqrt[4]{8}} = 0.577$$

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

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 odklik

V analizo dobimo jekleno pločevino iz dveh šarž. Z analizo vsebnosti mangana v jeklih želimo določiti ali gre za enakovredno pločevina. V prvem vzorcu smo dobili naslednje rezultate: 3.46, 3.56, 3.48, 3.52, 3.51 ($\mu\text{g/g}$). Rezultati drugega vzorca pa so bili: 3.35, 3.42, 3.33, 3.45, 3.38, 3.4 ($\mu\text{g/g}$). Ali sta vzorca enaka pri 5% tveganju?

Pri tehnološkem postopku izdelave lepila testiramo vpliv 3 aditivov na kvaliteto. Kvaliteto smo definirali med 1 in 2. Višja vrednost predstavlja boljše lastnosti. Naredili smo 8 eksperimentov za dvonivojski eksperimentalni načrt. Podatki v μg aditiva na kg so zbrani v tabeli. V zadnji vrstici so podane meje, ki razdelijo vsebnost posameznih aditivov na zgornji in spodnji nivo. Izračunajte vpliv 1 aditiva na kvaliteto laka!

| | 1 | 2 | 3 | kvaliteta |
|------|----------|----------|----------|------------------|
| 1 | 8.75 | 18 | 90.4 | 1.5 |
| 2 | 6.25 | 27 | 271.2 | 1.7 |
| 3 | 5.69 | 9 | 113 | 1.85 |
| 4 | 9.375 | 23 | 339 | 1.9 |
| 5 | 8.44 | 8 | 271.2 | 1.7 |
| 6 | 7.875 | 5 | 135.6 | 1.95 |
| 7 | 5.315 | 11 | 271.2 | 1.4 |
| 8 | 4.69 | 18 | 135.6 | 1.1 |
| Meje | 6.565 | 16 | 180.8 | |