3.2. Aktívny experiment

⇒ **Aktívny experiment** – zisťujú sa **odozvy** na **zámerne zavádzané vstupné** testovacie signály.

⇒ **Cieľ experimentu** – **vytvorenie** takých **podmienok**, aby sa dosiahli **kvalitné odhady** parametrov systému **pomocou neveľkého objemu meraní a výpočtov**.

⇒ **Plánovanie (navrhovanie) experimentov** (*Design of Experiments – DOE, Designed Experiments, Experimental Design*)

- **Disciplína matematickej štatistiky**, zaoberajúca sa **štúdiom vplyvu** viacerých nezávislých premenných (faktorov) na sledovanú závislú veličinu.
- Umožňuje **roztriediť faktory na významné a menej významné**.
- **Vznikla v agronómii** pri systematickej analýze dielčích vplyvov na úrodu **v prvej polovici 20. storočia.** Základy modernej teórie plánovania experimentov položil sir Ronald Fischer: The Arrangement of Field Experiments (1926), The Design of Experiments (1935).

⇒ **Plánovaný experiment** – postupnosť **vopred naplánovaných pokusov** pri ktorých sa **cieľavedome vykonáva zmena vstupných faktorov** procesu tak, aby mohli byť **pozorované** a vhodným vyhodnotením experimentálnych dát **identifikované odpovedajúce zmeny odozvy**.

3.2.1. Ortogonálny plán pre lineárny model

⇒ Predpokladáme **lineárny statický n-faktorový model**

$$\boxed{\hat{y} = \hat{\theta}'_0 + \hat{\theta}'_1 u'_1 + \ldots + \hat{\theta}'_n u'_n}$$

kde premenné u'_i aj parametre $\hat{\theta}'_i$ majú **fyzikálny význam** a každá premenná u'_i má vymedzený **pracovný rozsah** svojimi hraničnými hodnotami

$$\boxed{u'_{i\,min} \le u'_i \le u'_{i\,max}}$$

Pracovná oblasť bude v priestore vstupov $\mathbf{U}' \subset R^n$ mať tvar hyperkvádra s dĺžkou hrán

$$d_i = \left(u'_{i\,max} - u'_{i\,min}\right).$$

Počet neznámych parametrov je $1 + k = 1 + n$

⇒ Ak pre každý z **n vstupov** uskutočníme merania na **h hladinách (hodnotách)**, potom **celkový počet bodov**, v ktorých sa budú experimenty realizovať, je $N = h^n$.

⇒ **Skutočné hodnoty** výstupu y sa môžu **od odhadovaných** hodnôt $\hat{y}$ **líšiť**, čo môže byť spôsobené tým, že

- zvolený **model** je pre aproximáciu skutočnej závislosti **nevhodný,**
- výsledky **meraní** sú zaťažené takými **chybami**, že dochádza ku skresľovaniu hodnôt odhadovaných parametrov.

Vplyv týchto faktorov nemožno od seba oddeliť, ale **ich štatistickú separáciu** možno urobiť tak, že **v každom z N bodov urobíme q paralelných meraní** s výsledkami y_{ij}, kde $i = 1,2,\ldots,N$ a $j = 1,2,\ldots,q$ a potom v každom bode určíme výberový priemer

$$\boxed{\overline{y}_i = \frac{1}{q}\sum_{j=1}^{q} y_{ij}}.$$

Rozptyl jednotlivých meraní k priemernej hodnote je mierou presnosti merania, zatiaľ čo **rozptyl výstupu regresnej funkcie voči priemernej hodnote merania** v každom z N bodov hovorí o **adekvátnosti modelu**.

$\Rightarrow$ **V dôsledku paralelných meraní vzrastie celkový počet meraní** na $P = qN = qh^n$.
Hodnotu N možno minimalizovať tým, že **zredukujeme počet hladín na** $h = 2$ (t.j. **experimenty** budeme robiť **iba v krajných hodnotách vstupných premenných** – rohových bodoch hyperkvádra), potom $N = 2^n$. V každom z týchto bodov budeme robiť **minimálne** $q = 2$ merania.

$\Rightarrow$ **Vstupy** u_i' **pretransformujeme na bezrozmerné premenné** u_i

$$\boxed{u_i = \frac{u_i' - \frac{u'_{i\max} + u'_{i\min}}{2}}{\frac{u'_{i\max} - u'_{i\min}}{2}} = \frac{2u_i' - u'_{i\max} - u'_{i\min}}{u'_{i\max} - u'_{i\min}}} \qquad u_i \in \left\langle u'_{i\min}, u'_{i\max} \right\rangle \Rightarrow u_i \in \left\langle -1, 1 \right\rangle$$

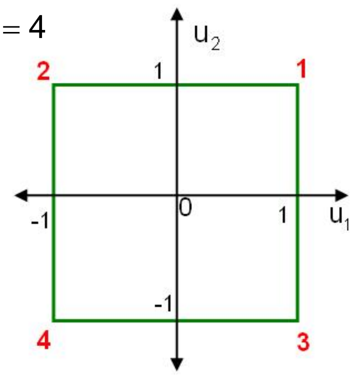
z pôvodného hyperkvádra sa stáva **hyperkocka s dĺžkou hrany rovnou 2**. Bezrozmerné hodnoty vstupov potom nadobúdajú iba dve hodnoty: 1 pre $u_i' = u'_{i\max}$ a -1 pre $u_i' = u'_{i\min}$.

Potom **modelom** je rovnica: $\boxed{\hat{y} = \hat{\theta}_0 + \hat{\theta}_1 u_1 + \ldots + \hat{\theta}_n u_n}$ a **počet meraní** je $\boxed{P = q2^n}$.

$\Rightarrow$ **Rozmiestnenie bodov merania v priestore**

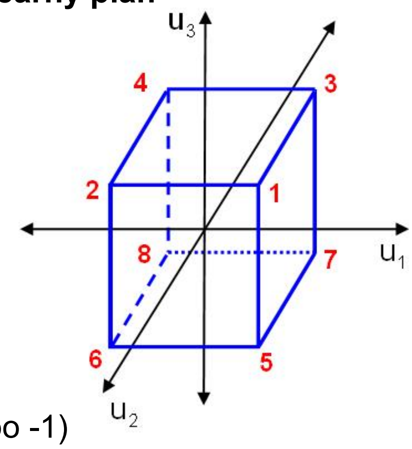
dvojfaktorový lineárny plán

$N = 2^2 = 4$



trojfaktorový lineárny plán

$N = 2^3 = 8$



Matica systému H bude mať koeficienty (iba hodnoty 1 alebo -1)

$$\boxed{h_{ij} = u_{ij} = (-1)^{\mathrm{int}\left(\frac{i-1}{2^{j-1}}\right)} \quad i = 1, 2, \ldots, 2^n \quad j = 0, 1, \ldots, n}$$

V prípade **trojfaktorového plánu** ide o problém riešenia **preurčeného systému 2^3=8 rovníc** pre **1+3=4 neznáme regresné koeficienty.** Tabuľka ortogonálneho plánu je nasledovná:

Číslo merania	Matica **H**				Výstupy					
	u_0	u_1	u_2	u_3	y_1	y_2	$\cdots$	y_q	$\bar{y}$	$\hat{y}$
1	1	1	1	1	x	x	…	x	x	
2	1	-1	1	1	x	x	…	x	x	
3	1	1	-1	1	x	x	…	x	x	
4	1	-1	-1	1	x	x	…	x	x	

5	1	1	1	-1	x	x	...	x	x	
6	1	-1	1	-1	x	x	...	x	x	
7	1	1	-1	-1	x	x	...	x	x	
8	1	-1	-1	-1	x	x	...	x	x	

Predposledný stĺpec obsahuje výberový priemer z q meraní.

⇒ Ak označíme jednotkový vektor dĺžky 2^n ako $\mathbf{L}_n = (1,1,\ldots,1)^T$, potom $\mathbf{H} = (\mathbf{L}_n, \mathbf{U}_n)$, kde $\mathbf{U}_n$ je matica rozmeru $2^n x\, n$.

ORTOGONALITA

Skalárny súčin ľubovoľných dvoch stĺpcov matice **H je rovný nule**, t.j. **stĺpce** tejto matice **sú navzájom ortogonálne**. Z toho vyplýva, že:

- **Súčet prvkov každého stĺpca** matice **H** je **rovný 0** (okrem prvého)**:**

$$\mathbf{L}_n^T \mathbf{u}_j = \sum_{i=1}^{2^n} u_{ij} = 0$$

- **Informačná matica R** aj **disperzná matica P** sú **diagonálne**:

$$\boxed{\mathbf{R} = \mathbf{H}^T\mathbf{H} = \begin{pmatrix} \mathbf{L}_n^T \\ \mathbf{U}_n^T \end{pmatrix} (\mathbf{L}_n, \mathbf{U}_n) = 2^n \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix} = 2^n \mathbf{I}_{n+1}}$$

$$\boxed{\mathbf{P} = \mathbf{R}^{-1} = \frac{1}{2^n} \mathbf{I}_{n+1}}$$

V dôsledku ortogonality plánu sa výrazne **zjednodušia výpočty** (**nemusíme počítať inverziu matice**) a tiež **jednotlivé koeficienty** θ_j sú navzájom **nezávislé**, t.j. ich kovariancia je nulová.

⇒ Zvýšením počtu faktorov o 1 dostaneme ortogonálnu maticu $\mathbf{H}_{n+1}$,ktorej prvým stĺpcom bude jednotkový vektor $\mathbf{L}_{n+1}$ a zvyšných n+1 stĺpcov predstavuje maticu $\mathbf{U}_{n+1}$, ktorú získame pomocou **rekurentného vzťahu**:

$$\mathbf{U}_{n+1} = \begin{pmatrix} \mathbf{U}_n & \mathbf{L}_n \\ \mathbf{U}_n & -\mathbf{L}_n \end{pmatrix}$$

⇒ Na **výpočet** jednotlivých **regresných koeficientov** $\hat{\theta}_j,\ j = 0,1,\ldots,n$ použijeme Gaussov vzorec, ktorý **sa zjednoduší na tvar**:

$$\boxed{\hat{\theta}_j^* = \frac{1}{2^n} \sum_{i=1}^{2^n} \bar{y}_i u_{ij} = \frac{1}{2^n} \sum_{i=1}^{2^n} \bar{y}_i (-1)^{\mathrm{int}\left(\frac{i-1}{2^{j-1}}\right)}} \qquad j = 0,1,\ldots,n$$

Tieto koeficienty sú **nevychýleným odhadom** parametrov θ_j.

⇒ Pre **disperzie** jednotlivých **regresných koeficientov** platí:

$$\sigma_{\theta_j}^2 = \sigma^2 p_{jj} = \frac{1}{2^n} \sigma^2$$

Disperziu šumu σ^2 **nepoznáme**, musíme ju odhadnúť na základe nameraných údajov

$$\boxed{S_{\bar{Y}}^2 = \frac{Q_{\bar{Y}}}{\nu_{\bar{Y}}} \qquad Q_{\bar{y}} = \sum_{i=1}^{2^n} \sum_{j=1}^{q} (y_{ij} - \bar{y}_i)^2 \qquad \nu_{\bar{y}} = 2^n (q-1)}$$

$S_{\bar{y}}^2$ je **výberový rozptyl jednotlivých vzoriek (meraní) voči priemeru skupiny** a **odhaduje disperziu šumov** σ^2, potom pre **odhad disperzie regresných koeficientov** platí:

$$\boxed{S_{\theta_j}^2 = \frac{1}{2^n} S_{\bar{y}}^2} \qquad \Rightarrow \qquad \boxed{S_{\theta_j} = 2^{\frac{-n}{2}} S_{\bar{y}}}$$

$\Rightarrow$ Ak má **niektorý z regresných koeficientov malú hodnotu**

a. jeho hodnota by **mala byť nulová**, t.j. príslušný faktor nemá vstupovať do modelu, jeho nenulová hodnota je **spôsobená chybami merania**

b. jeho hodnota **má byť nenulová** a nemožno ho zanedbať.

Pri **testovaní štatistickej významnosti vplyvu jednotlivých faktorov** na výslednú regresnú závislosť môžeme použiť štatistiku t, ktorá sa riadi Studentovým rozdelením s

$\nu_{\bar{y}} = 2^n(q-1)$ stupňami voľnosti $\qquad t = \dfrac{\hat{\theta}_j^* - \theta_j}{S_{\theta_j}} = \dfrac{\hat{\theta}_j^* - \theta_j}{S_{\bar{y}}} 2^{\frac{n}{2}}$

Dosadíme do $\quad P\left(-t_{1-\frac{\alpha}{2}}(\nu_{\bar{Y}}) \le t \le t_{1-\frac{\alpha}{2}}(\nu_{\bar{Y}})\right) = 1-\alpha$

po úprave $\quad P\left(\theta_j - t_{1-\frac{\alpha}{2}}(\nu_{\bar{Y}})\dfrac{S_{\bar{Y}}}{2^{\frac{n}{2}}} \le \hat{\theta}_j^* \le \theta_j + t_{1-\frac{\alpha}{2}}(\nu_{\bar{Y}})\dfrac{S_{\bar{Y}}}{2^{\frac{n}{2}}}\right) = 1-\alpha$

$$P\left(\theta_j - \delta \le \hat{\theta}_j^* \le \theta_j + \delta\right) = 1-\alpha$$

Ak θ_j=0, t.j. t.j. **člen $\theta_j u_j$ sa v štruktúre reálneho systému nenachádza**, potom platí

$$P\left(-\delta \le \hat{\theta}_j^* \le \delta\right) = 1-\alpha$$

Koeficient $\hat{\theta}_j = 0$ a teda **faktor** u_i **z modelu vypustíme**, ak

$$\boxed{-\delta \le \hat{\theta}_j^* \le \delta} \qquad \text{kde} \qquad \boxed{\delta = S_{\theta_j} t_{1-\frac{\alpha}{2}}(\nu_{\bar{y}}) = t_{1-\frac{\alpha}{2}}(\nu_{\bar{y}})\frac{S_{\bar{y}}}{2^{\frac{n}{2}}}}$$

$\Rightarrow$ **Adekvátnosť lineárneho modelu** (je lineárny model vhodný?)

overíme pomocou Fisher-Snedecorovho testu $\quad \boxed{w = \dfrac{S_{\tilde{y}}^2}{S_{\bar{y}}^2} = \dfrac{\dfrac{1}{2^n - n - 1} q \sum_{i=1}^{2^n} (\bar{y}_i - \hat{y}_i)^2}{\dfrac{1}{2^n(q-1)} \sum_{i=1}^{2^n} \sum_{j=1}^{q} (y_{ij} - \bar{y}_i)^2}}$

kde $\quad \boxed{S_{\tilde{Y}}^2 = \dfrac{Q_{\tilde{Y}}}{\nu_{\tilde{Y}}} \qquad Q_{\tilde{y}} = q\sum_{i=1}^{2^n} (\bar{y}_i - \hat{y}_i)^2 \qquad \nu_{\tilde{Y}} = 2^n - n - 1}$

je **výberový rozptyl výstupu regresného modelu voči priemeru skupiny údajov**.

Ak sú oba **rozptyly** približne **rovnaké** (t.j. **chyba modelu je približne na úrovni chyby merania**), platí $\quad P\left(w \le w_{1-\alpha}(\nu_{\tilde{y}}, \nu_{\bar{y}})\right) = 1-\alpha$

Lineárny model je adekvátny, ak $\quad \boxed{w \le w_{1-\alpha}(\nu_{\tilde{y}}, \nu_{\bar{y}})}$

v opačnom prípade je **potrebná iná štruktúra modelu**.

☑ ***Postup***:

1. Vstupy u_i' pretransformujeme na bezrozmerné premenné $u_i \in \langle -1,1 \rangle$

$$u_i = \frac{u_i' - \frac{u_{i\max}' + u_{i\min}'}{2}}{\frac{u_{i\max}' - u_{i\min}'}{2}} = \frac{2u_i' - u_{i\max}' - u_{i\min}'}{u_{i\max}' - u_{i\min}'} \qquad i = 1,\ldots,n$$

2. Vytvoríme maticu **H** – rozmiestnenie bodov merania. Zvolíme počet paralelných meraní v každom bode q. Odmeriame hodnoty výstupu y_{ij}, $i = 1,\ldots,2^n$, $j = 1,\ldots,q$. Vypočítame výberový priemer z q meraní $\bar{y}_i$ pre $i = 1,\ldots,2^n$.
3. Určíme hodnoty parametrov

$$\hat{\theta}_j = \frac{1}{2^n}\sum_{i=1}^{2^n} \bar{y}_i u_{ij} = \frac{1}{2^n}\sum_{i=1}^{2^n} \bar{y}_i (-1)^{\mathrm{int}\left(\frac{i-1}{2^{j-1}}\right)} \qquad j = 0,\ldots,n$$

4. Otestujeme štatistickú významnosť parametrov. Koeficient $\hat{\theta}_j = 0$ a teda faktor u_i z modelu vypustíme, ak

$$-\delta \le \hat{\theta}_j \le \delta \qquad \delta = S_{\theta_j} t_{1-\frac{\alpha}{2}}(\nu_{\bar{y}}) = \frac{S_{\bar{y}}}{2^{\frac{n}{2}}} t_{1-\frac{\alpha}{2}}(\nu_{\bar{y}})$$

kde $S_{\bar{Y}}^2 = \frac{Q_{\bar{Y}}}{\nu_{\bar{Y}}}$ $\qquad \nu_{\bar{Y}} = 2^n(q-1)$ $\qquad Q_{\bar{y}} = \sum_{i=1}^{2^n}\sum_{j=1}^{q}(y_{ij} - \bar{y}_i)^2$

5. Otestujeme adekvátnosť lineárneho plánu. Lineárny model je adekvátny, ak

$$\frac{S_{\tilde{Y}}^2}{S_{\bar{Y}}^2} = \frac{\frac{1}{2^n - n - 1} q \sum_{i=1}^{2^n}(\bar{y}_i - \hat{y}_i)^2}{\frac{1}{2^n(q-1)}\sum_{i=1}^{2^n}\sum_{j=1}^{q}(y_{ij} - \bar{y}_i)^2} \le w_{1-\alpha}(\nu_1, \nu_2) \qquad \begin{aligned}\nu_1 &= 2^n - n - 1 \\ \nu_2 &= 2^n(q-1)\end{aligned}$$

$$S_{\tilde{Y}}^2 = \frac{Q_{\tilde{Y}}}{\nu_{\tilde{Y}}} \qquad Q_{\tilde{y}} = q\sum_{i=1}^{2^n}(\bar{y}_i - \hat{y}_i)^2 \qquad \nu_{\tilde{Y}} = 2^n - n - 1$$

🗎 ***Príklad → priklad_ortog_plan.pdf + cvičenia***

3.2.2. <u>Ortogonálny plán pre kvadratický model</u>

⇒ Predpokladáme **n-faktorový kvadratický model**

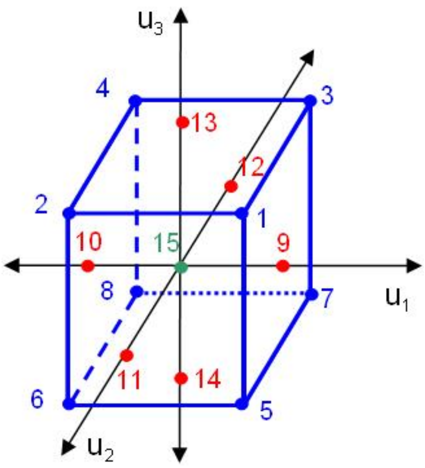
$$\boxed{\hat{y} = \hat{\theta}_0 + \hat{\theta}_1 u_1 + \ldots + \hat{\theta}_n u_n + \hat{\theta}_{n+1} u_1^2 + \ldots + \hat{\theta}_{2n} u_n^2 + \hat{\theta}_{2n+1} u_1 u_2 + \ldots + \hat{\theta}_k u_{n-1} u_n}$$

počet neznámych parametrov

$$\boxed{1 + k = \frac{1}{2}(n+1)(n+2) = 1 + \frac{1}{2}n(n+3)},$$

t.j. treba **rozšíriť počet stĺpcov** matice **H** a **zväčšiť aj počet riadkov** (počet bodov plánu experimentu)

⇒ **Kompozičný plán** – kompozícia troch skupín bodov plánu:



a) **rohové** – 2^n vrcholov hyperkocky s dĺžkou hrany rovnou dvom (pre 3-faktorový plán modré body 1-8)

b) **osové** – $2n$ bodov na osiach, symetricky položené voči počiatku, vzdialené od neho o hodnotu α (červené body 9-14)

c) **počiatok n-rozmerného priestoru** (zelený bod 15)

spolu: $N = 2^n + 2n + 1$ **bodov.**

$\Rightarrow$ **Ortogonalita plánu**

- nastavenie μ tak, aby súčet prvkov v stĺpcov prislúchajúcim kvadrátom faktorov bol nulový
- nastavenie α tak, aby stĺpce boli navzájom aj voči ostatným ortogonálne

$$\boxed{\mu = \sqrt{\frac{2^n}{N}}} \qquad \boxed{\alpha = \pm\sqrt{\frac{1}{2}\mu N(1-\mu)}}$$

Tabuľka ortogonálneho plánu pre kvadratický model:

Typy bodov	Počet bodov	Číslo polosi	Stĺpce matice H									
			u_0	Matica U			$u_1^2-\mu$	$u_2^2-\mu$	$u_3^2-\mu$	u_1u_2	u_1u_3	u_2u_3
				u_1	u_2	u_3						
			0	1	2	3	4	5	6	7	8	9
Rohové	2^n	1	1	1	1	1	1-μ	1-μ	1-μ	1	1	1
		2	1	-1	1	1	1-μ	1-μ	1-μ	-1	-1	1
		3	1	1	-1	1	1-μ	1-μ	1-μ	-1	1	-1
		4	1	-1	-1	1	1-μ	1-μ	1-μ	1	-1	-1
		5	1	1	1	-1	1-μ	1-μ	1-μ	1	-1	-1
		6	1	-1	1	-1	1-μ	1-μ	1-μ	-1	1	-1
		7	1	1	-1	-1	1-μ	1-μ	1-μ	-1	-1	1
		8	1	-1	-1	-1	1-μ	1-μ	1-μ	1	1	1
Osové	2n	9	1	α	0	0	α^2-μ	-μ	-μ	0	0	0
		10	1	-α	0	0	α^2-μ	-μ	-μ	0	0	0
		11	1	0	α	0	-μ	α^2-μ	-μ	0	0	0
		12	1	0	-α	0	-μ	α^2-μ	-μ	0	0	0
		13	1	0	0	α	-μ	-μ	α^2-μ	0	0	0
		14	1	0	0	-α	-μ	-μ	α^2-μ	0	0	0
Stred	1	15	1	0	0	0	-μ	-μ	-μ	0	0	0

$\Rightarrow$ **Informačná matica** má tvar

$$\boxed{\mathbf{R} = \mathbf{H}^T\mathbf{H} = \begin{pmatrix} r_0 & \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0} & r_1\mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & r_2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & r_3\mathbf{I}_{\binom{n}{2}} \end{pmatrix}}$$

$$\boxed{\begin{aligned} r_0 &= 2^2 + 2n + 1 = N \\ r_1 &= 2^n + 2\alpha^2 = \mu N \\ r_2 &= 2^2(1-\mu)^2 + 2(\alpha^2-\mu)^2 + (2n-1)\mu^2 = \frac{1}{2}\mu^2N^2(1-\mu)^2 \\ r_3 &= 2^n = \mu^2 N \end{aligned}}$$

kde $\mathbf{I}_n$ je jednotková matica rozmeru $n \times n$,

$\mathbf{I}_{\binom{n}{2}}$ je jednotková matica, ktorá má $\binom{n}{2} = \frac{1}{2}n(n-1)$ riadkov a stĺpcov

Disperzná matica $\quad \mathbf{P} = \mathbf{R}^{-1} \quad \rightarrow \quad \boxed{p_i = \frac{1}{r_i}}$

$$\mathbf{P} = \begin{pmatrix} p_0 & \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0} & p_1\mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & p_2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & p_3\mathbf{I}_{\binom{n}{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{r_0} & \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0} & \frac{1}{r_1}\mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{r_2}\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{r_3}\mathbf{I}_{\binom{n}{2}} \end{pmatrix}$$

⇒ **Odhad hľadaných parametrov**

$$\boxed{\begin{array}{lll} \hat{\theta}_j^* = p_1\sum_{i=1}^{N}\bar{y}_i u_{ij} & j = 1, 2, \cdots, n & \text{koef. pri lineárnych členoch} \\ \hat{\theta}_j^* = p_2\sum_{i=1}^{N}\bar{y}_i\left(u_{i,j-n}^2 - \mu\right) & j = n+1, \ldots, 2n & \text{koef. pri kvadratických členoch} \\ \hat{\theta}_j^* = p_3\sum_{i=1}^{N}\bar{y}_i u_{ix} u_{iz} & j = \frac{(x+1)(2n-z)}{2} + z = 2n+1, \cdots, k\,,\ 1 \le x < z \le n & \\ & & \text{koef. pri zmiešaných súčinoch} \\ \hat{\theta}_0^* = p_0\sum_{i=1}^{N}\bar{y}_i - \mu\sum_{i=1}^{n}\hat{\theta}_{i+n}^* & & \text{absolútny člen} \end{array}}$$

⇒ **Výberové rozptyly odhadovaných parametrov**

$$\mathrm{cov}\left(\hat{\boldsymbol{\theta}}\right) = \sigma^2\mathbf{P}$$

$$\boxed{\begin{array}{lll} S_{\theta_i}^2 = p_1 S_{\bar{y}}^2 & i = 1, 2, \ldots, n & \text{koef. pri lineárnych členoch} \\ S_{\theta_i}^2 = p_2 S_{\bar{y}}^2 & i = n+1, \ldots, 2n & \text{koef. pri kvadratických členoch} \\ S_{\theta_i}^2 = p_3 S_{\bar{y}}^2 & i = 2n+1, \ldots, k & \text{koef. pri zmiešaných súčinoch} \\ S_{\theta_0}^2 = \left(p_0 + \mu^2 n p_2\right) S_{\bar{y}}^2 & & \text{absolútny člen} \end{array}}$$

kde $S_{\bar{y}}^2 = \frac{1}{(q-1)N}\sum_{i=1}^{N}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_i\right)^2 \qquad \nu_{\bar{Y}} = (q-1)N$

je výberový rozptyl vzoriek voči priemeru skupiny.

⇒ **Testovanie štatistickej významnosti vplyvu jednotlivých faktorov**

Koeficient θ_j z modelu vypustíme, ak

$\boxed{-\delta_j \le \hat{\theta}_j^* \le \delta_j}$ kde $\boxed{\delta_j = S_{\theta_j} t_{1-\frac{\alpha}{2}}\left(\nu_{\bar{y}}\right)}$

⇒ **Adekvátnosť kvadratického modelu** overíme pomocou Fisher-Snedecorovho testu

$$\boxed{w = \frac{S_{\bar{y}}^2}{S_{\bar{y}}^2} = \frac{\frac{1}{N-k-1}\sum_{i=1}^{N} q\left(\bar{y}_i - \hat{y}_i\right)^2}{\frac{1}{N(q-1)}\sum_{i=1}^{N}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_i\right)^2}} \qquad \nu_1 = N-k-1,\ \ \nu_2 = N(q-1)$$

Kvadratický model je adekvátny, ak $\boxed{w \le w_{1-\alpha}\left(\nu_1, \nu_2\right)}$

☑ ***Postup***:

1. Vstupy u_i' pretransformujeme na bezrozmerné premenné $u_i \in \langle -1, 1 \rangle$

$$u_i = \frac{u_i' - \dfrac{u_{i\max}' + u_{i\min}'}{2}}{\dfrac{u_{i\max}' - u_{i\min}'}{2}} = \frac{2u_i' - u_{i\max}' - u_{i\min}'}{u_{i\max}' - u_{i\min}'} \qquad i = 1, \ldots, n$$

2. Počet bodov je $N = 2^n + 2n + 1$ a

$$\mu = \sqrt{\frac{2^n}{N}} \qquad \alpha = \pm\sqrt{\frac{1}{2}\mu N(1-\mu)}$$

3. Vytvoríme maticu **H** – rozmiestnenie bodov merania. Zvolíme počet paralelných meraní v každom bode q. Odmeriame hodnoty výstupu y_{ij}, $i = 1, \ldots, N$, $j = 1, \ldots, q$. Vypočítame výberový priemer z q meraní $\bar{y}_i$ pre $i = 1, \ldots, N$.
4. Určíme prvky disperznej matice

$$p_3 = \frac{1}{\mu^2 N} \qquad p_2 = \frac{2}{\mu^2(1-\mu)^2 N^2} \qquad p_1 = \frac{1}{\mu N} \qquad p_0 = \frac{1}{N}$$

5. Určíme hodnoty odhadu parametrov

$$\hat{\theta}_j^* = p_1 \sum_{i=1}^{N} \bar{y}_i u_{ij} \qquad j = 1, 2, \cdots, n \qquad \text{koef. pri lineárnych členoch}$$

$$\hat{\theta}_j^* = p_2 \sum_{i=1}^{N} \bar{y}_i \left(u_{i,j-n}^2 - \mu\right) \qquad j = n+1, \ldots, 2n \qquad \text{koef. pri kvadratických členoch}$$

$$\hat{\theta}_j^* = p_3 \sum_{i=1}^{N} \bar{y}_i u_{ix} u_{iz} \qquad j = \frac{(x+1)(2n-z)}{2} + z = 2n+1, \cdots, k\,,\ 1 \le x < z \le n$$

koef. pri zmiešaných súčinoch

$$\hat{\theta}_0^* = p_0 \sum_{i=1}^{N} \bar{y}_i - \mu \sum_{i=1}^{n} \hat{\theta}_{i+n}^* \qquad \text{absolútny člen}$$

6. Otestujeme štatistickú významnosť parametrov. Koeficient $\hat{\theta}_j$ z modelu vypustíme, ak

$$-\delta_j \le \hat{\theta}_j \le \delta_j \qquad \delta_j = S_{\theta_j} t_{1-\frac{\alpha}{2}}\left(\nu_{\bar{y}}\right)$$

$$S_{\theta_i}^2 = p_1 S_{\bar{y}}^2 \qquad i = 1, 2, \ldots, n \qquad \text{koef. pri lineárnych členoch}$$

$$S_{\theta_i}^2 = p_2 S_{\bar{y}}^2 \qquad i = n+1, \ldots, 2n \qquad \text{koef. pri kvadratických členoch}$$

$$S_{\theta_i}^2 = p_3 S_{\bar{y}}^2 \qquad i = 2n+1, \ldots, k \qquad \text{koef. pri zmiešaných súčinoch}$$

$$S_{\theta_0}^2 = \left(p_0 + \mu^2 n p_2\right) S_{\bar{y}}^2 \qquad \text{absolútny člen}$$

$$Q_{\bar{y}} = \sum_{i=1}^{N} \sum_{j=1}^{q} \left(y_{ij} - \bar{y}_i\right)^2 \qquad \nu_{\bar{Y}} = N(q-1) \qquad S_{\bar{Y}}^2 = \frac{Q_{\bar{Y}}}{\nu_{\bar{Y}}}$$

7. Otestujeme adekvátnosť kvadratického plánu. Kvadratický model je adekvátny, ak

$$\frac{S_{\tilde{Y}}^2}{S_{\bar{Y}}^2} = \frac{\dfrac{1}{N-k-1} \sum_{i=1}^{N} q\left(\bar{y}_i - \hat{y}_i\right)^2}{\dfrac{1}{N(q-1)} \sum_{i=1}^{N} \sum_{j=1}^{q} \left(y_{ij} - \bar{y}_i\right)^2} \le w_{1-\alpha}\left(\nu_1, \nu_2\right) \qquad \nu_1 = N-k-1,\ \nu_2 = N(q-1)$$

Príklad → priklad_ortog_plan.pdf + cvičenia