Národní informační středisko pro podporu jakosti

ANALÝZA ROZPTYLU

a její využití při vyhodnocování experimentálních dat

Eva Jarošová, VŠE Praha

Obsah

- Podstata metody, jednofaktorová ANOVA
- F-test
- Mnohonásobná porovnávání
- Dvoufaktorová ANOVA, Excel
- Aplikace při faktoriálním návrhu s jednou replikací
- ANOVA ve Statistice a Minitabu
- Předpoklady a jejich ověření
- Odhad složek rozptylu

Použití

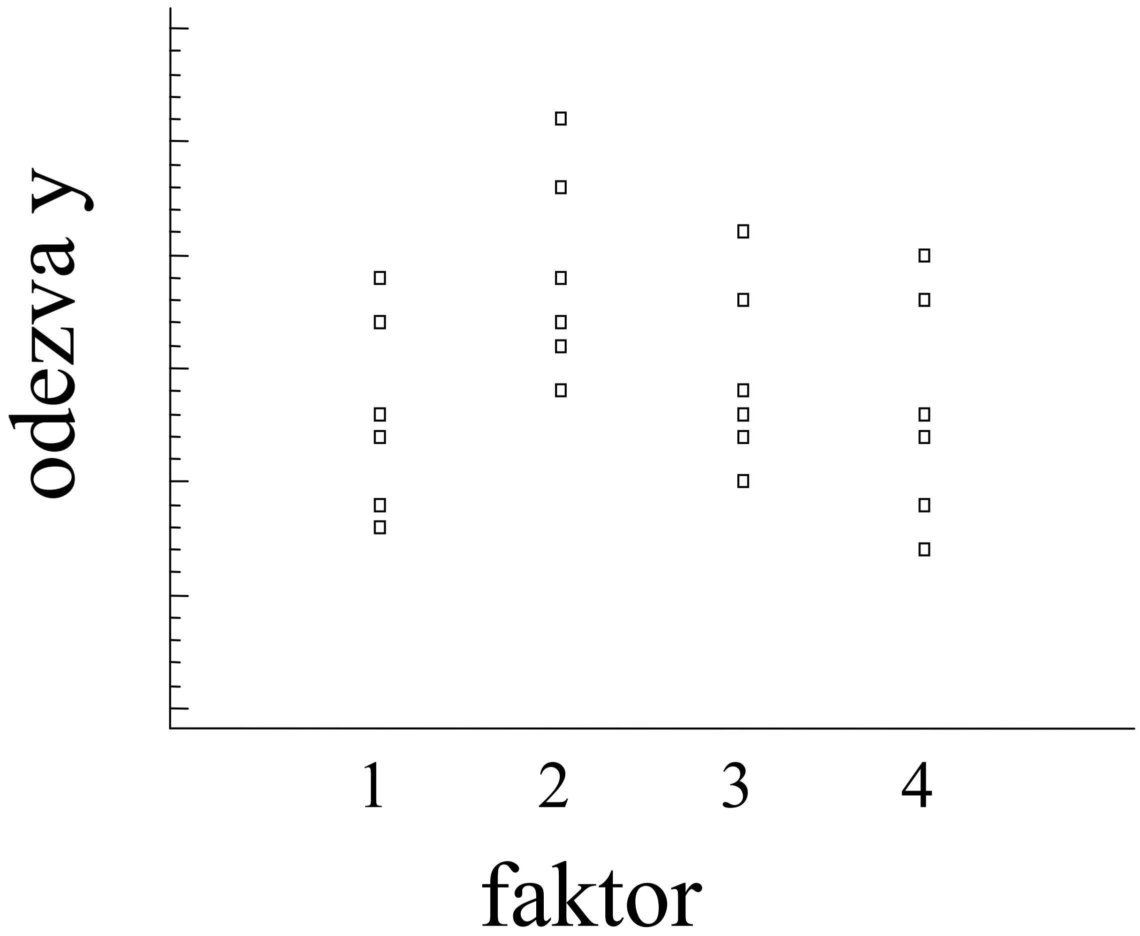
Při analýze experimentů pro zkoumání vlivu jednoho či více faktorů na sledovanou veličinu (odezvu)

- Nutnost u faktorů s více než dvěma úrovněmi
- Využití v případě více faktorů s dvěma úrovněmi (jinak t-test), především v případě jedné replikace experimentu
- Odhad složek rozptylu při identifikaci zdrojů variability

Jednofaktorová ANOVA

- Roztřídění naměřených hodnot do skupin podle úrovní jediného faktoru
- Výpočet průměru v každé skupině
- Porovnávání průměrů pomocí statistického testu

Znázornění výsledků měření bodovým diagramem



F-test

Předpoklad: $y_{ij} \sim N(\mu_i, \sigma^2)$

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

faktor nemá vliv, změny jeho úrovní nevedou ke změně střední hodnoty

H_1: non H_0

faktor má vliv

F - test

Testová statistika

pro vyvážený návrh

$$F = \frac{\frac{SS_A}{a-1}}{\frac{SS_E}{a(r-1)}}$$

a … počet úrovní faktoru A

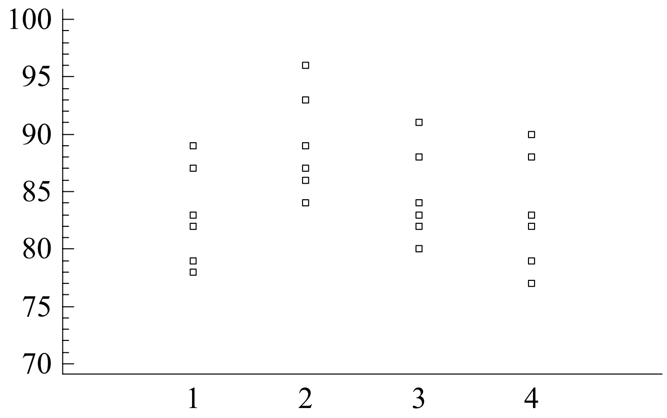
r … počet replikací

SS_A měří variabilitu mezi skupinami výsledků při různých úrovních zkoumaného faktoru

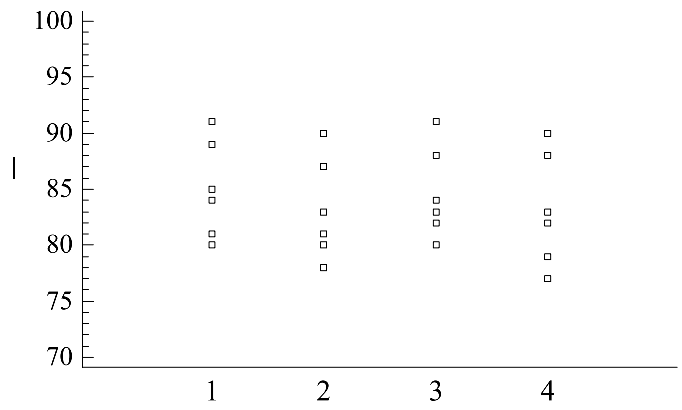
SS_E měří variabilitu výsledků uvnitř skupin odpovídající experimentální chybě

Součet čtverců SS_A

$$SS_A = r\sum_{i=1}^{a}(\overline{y}_i - \overline{y})^2$$



větší SS_A

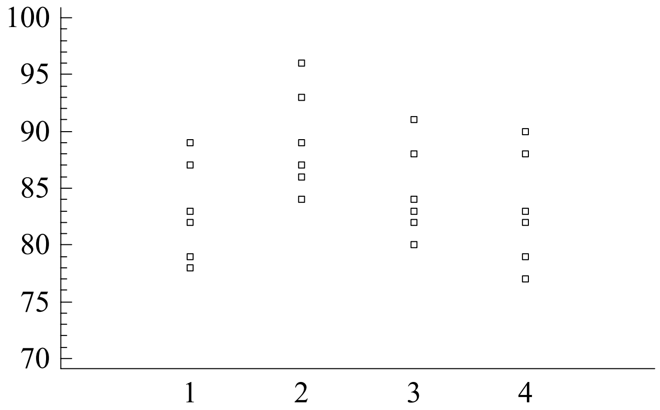


menší SS_A

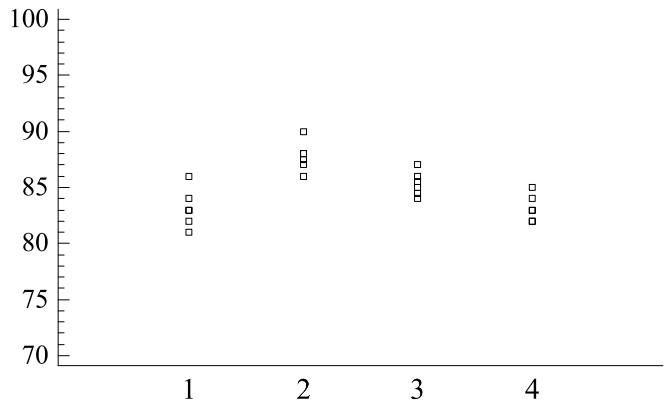
Větší hodnoty SS_A svědčí proti hypotéze H_0

Součet čtverců SS_E

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{r}(y_{ij} - \overline{y}_i)^2$$



větší SS_E



menší SS_E

Větší experimentální chybě odpovídají větší hodnoty SS_E

ANOVA

Rozklad celkové variability měřené celkovým součtem čtverců na část způsobenou rozdíly mezi skupinami a na vnitroskupinovou variabilitu

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{r} (y_{ij} - \overline{y})^2$$

$$SS_T = SS_A + SS_E$$

Tabulka ANOVA

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	F
faktor A	SS_A	a - 1	$\frac{SS_A}{a-1}$	$F = \frac{\frac{SS_A}{a-1}}{\frac{SS_E}{a(r-1)}}$
reziduální	SS_E	$a(r - 1)$	$\frac{SS_E}{a(r-1)}$	
celkový	SS_T	ar - 1		

Rozhodování

- Porovnání vypočtené hodnoty F s kritickou hodnotou

  $$F_{1-\alpha}(\nu_1,\nu_2), \quad \nu_1 = a-1, \quad \nu_2 = a(r-1)$$

  Platí-li F > $F_{1-\alpha}(\nu_1,\nu_2)$
  zamítáme H_0 na hladině významnosti α.

- Porovnání p-hodnoty s hladinou významnosti α
  Je-li p-hodnota < α,
  zamítáme H_0 na hladině významnosti α

Příklad 1

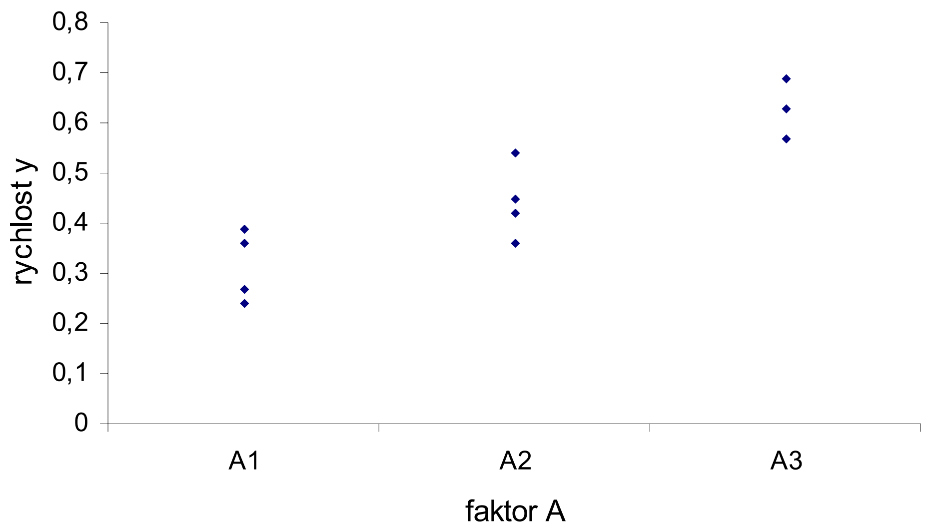
Cílem experimentu je určit vhodnou rychlost proudu kapaliny z ostřikovače předního skla automobilu. Zkouší se tři různé velikosti otvoru trysky. V experimentu jsou použity vždy čtyři trysky se stejnou velikostí otvoru.

Velikost (mm^2)		rychlost (m/s)			
A_1	5,2	0,24	0,36	0,27	0,39
A_2	3,9	0,45	0,42	0,36	0,54
A_3	2,6	0,69	0,57	0,57	0,63

Tabulka ANOVA

ANOVA

Zdroj variability	*SS*	*st. vol.*	*MS*	*F*	*Hodnota P*	*F krit*
faktorA	0,18135	2	0,090675	19,39572	0,000546	4,256492
reziduální	0,042075	9	0,004675			
celkový	0,223425	11				



Mnohonásobná porovnávání

Následují poté, kdy F-test vyšel významně. Nejčastěji zjišťujeme, které průměry se od sebe významně liší. Přitom porovnáváme různé dvojice středních hodnot. V podstatě jde o konstrukci intervalu spolehlivosti pro rozdíl středních hodnot. Rozlišujeme dva případy:

- předem zvolená dvojice průměrů
- všechny možné dvojice průměrů (mnohonásobná porovnávání)

Porovnání dvou předem zvolených středních hodnot

Např. interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$

$$\bar{y}_{1.} - \bar{y}_{2.} \pm t_{1-\alpha/2} \sqrt{\frac{2 \cdot MS_E}{r}}$$

MS_E je průměrný čtverec z tabulky ANOVA, *r* počet replikací, t značí kvantil t rozdělení s a(r-1) stupni volnosti.

Testujeme hypotézu H_0: $\mu_1 - \mu_2 = 0$

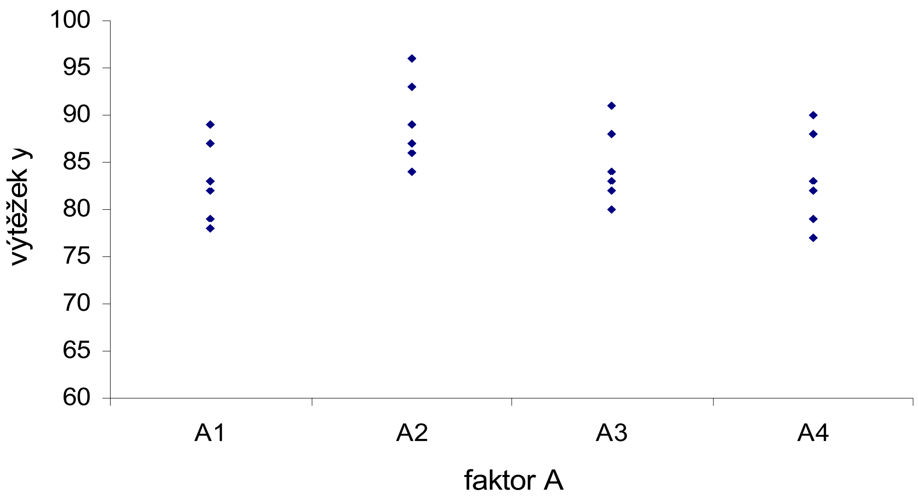
Neobsahuje-li interval 0, zamítneme H_0 na hladině významnosti α.

Příklad 2

Pomocí experimentu zkoumáme vliv použitého katalyzátoru na výtěžek chemického procesu a chceme vybrat nejvhodnější ze čtyř typů. Odezvou je výtěžek chemického procesu, tj. množství vyráběné látky, zkoumaný faktor, katalyzátor, má čtyři úrovně. Předpokládejme, že můžeme provést dvacet čtyři zkoušek, při vyváženém návrhu to znamená vždy šest zkoušek se stejným katalyzátorem. Jedna várka vstupní suroviny ovšem stačí jen na provedení čtyř zkoušek. Při plánování experimentu je třeba počítat s tím, že kvalita várek vstupní suroviny bude kolísat a přispívat k větší experimentální chybě. Proto je vhodné volit uspořádání do bloků tak, že bloky budou určeny jednotlivými várkami vstupní suroviny. Při jedné várce vstupní suroviny vystřídáme všechny úrovně zkoumaného faktoru (katalyzátoru), jejich pořadí se pro každý blok volí náhodně.

Příklad 2

	várka					
katalyzátor	1	2	3	4	5	6
A_1	87	79	82	89	83	78
A_2	93	84	89	96	86	87
A_3	88	80	84	91	83	82
A_4	88	77	83	90	82	79



ANOVA – model s dvěma faktory, bez interakce

Rozklad celkového součtu čtverců

$$SS_T = SS_A + SS_b + SS_E$$

faktor A   bloky

p-hodnotu porovnáváme s α

ANOVA

Zdroj variability	*SS*	*St.vol*	*MS*	*F*	*Hodnota P*	*F krit*
Faktor A	149	3	49,66667	49,66667	5,03E-08	3,287383
Bloky	392	5	78,4	78,4	3,28E-10	2,901295
Reziduální	15	15	1			
Celkový	556	23				

klasický postup testu:
F porovnáváme s *F krit*

ANOVA - závěr

Klasický postup $F > F_{krit}$

Faktor *katalyzátor* má vliv na výtěžek procesu (rozdíl mezi katalyzátory existuje). F *krit* odpovídá zvolené hladině významnosti α = 0,05.

Postup pomocí p-hodnoty

Používají se hodnoty α 0,05 0,01 0,001

$$p - hodnota < 0,001$$

Vliv faktoru prokázán na hladině významnosti 0,001.

Příklad 3

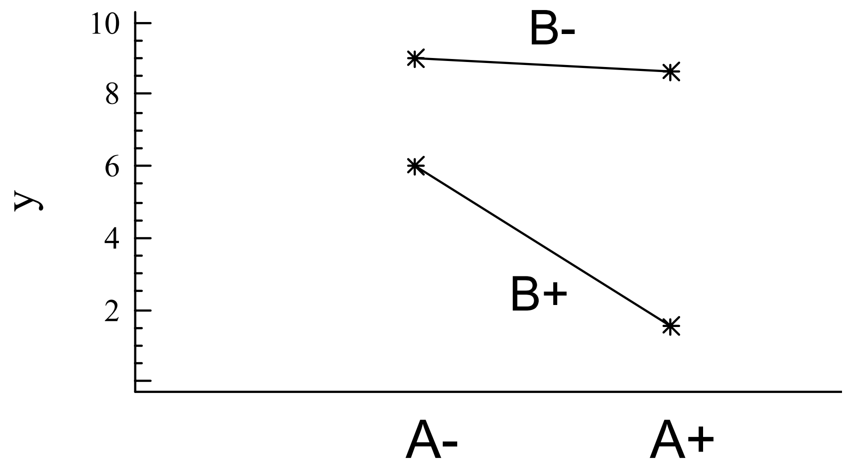
Produkt se v hnětači zpracovává v dávkách a k tomu, aby se dosáhlo požadované viskozity, je třeba, aby reakce trvala 7 až 9 hodin. Při kratší době reakce má vyrobená dávka nežádoucí vlastnosti. Doba reakce závisí na procentu složky X ve zpracovávané směsi (faktor A) a na teplotě (faktor B). Je třeba najít takovou kombinaci podílu složky X a teploty, která povede k reakci požadované délky. Cílem experimentu je zjistit, zda faktory A a B ovlivňují dobu reakce a zda nedochází k jejich interakci.

Příklad 3

A	B	
	-	+
-	(4) 9,0 (6) 9,0	(2) 5,5 (3) 6,5
+	(1) 9,3 (8) 8,0	(7) 1,8 (5) 1,3

(pořadí zkoušky) doba reakce

2 replikace - 2 zkoušky při každé kombinaci procenta složky a teploty



ANOVA – model s dvěma faktory, s interakcí

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

faktor A faktor B interakce AB

ANOVA

Zdroj variability	*SS*	*St. vol.*	*MS*	*F*	*Hodnota P*	*F krit*
Faktor A	11,52	1	11,52	31,34694	0,004996	7,70865
Faktor B	51,005	1	51,005	138,7891	0,000297	7,70865
Interakce	8,405	1	8,405	22,87075	0,008761	7,70865
Reziduální	1,47	4	0,3675			
Celkový	72,4	7				

Příklad 4

Automobilová součást se skládá z několika snýtovaných prvků. Cílem experimentu je určit, jak různé konfigurace prvků ovlivňují pevnost v tahu.

Faktory: výška nýtu (A), průměr otvoru v plechu (B), tloušťka plechu (C), průměr dříku (D), délka dříku (E), průměr podložky (F), tloušťka podložky (G)

7 faktorů, dílčí faktoriální experiment 2^{7-4}, 1 replikace

A	B	C	D	E	F	G	y
-	-	-	+	+	+	-	513
+	-	-	-	-	+	+	461
-	+	-	-	+	-	+	488
+	+	-	+	-	-	-	481
-	-	+	+	-	-	+	523
+	-	+	-	+	-	-	558
-	+	+	-	-	+	-	532
+	+	+	+	+	+	+	546

Příklad 4

Model ANOVA

$$A + B + C + D + E + F + G$$

V modelu uvažujeme jen hlavní efekty faktorů A až G. Protože jde o dílčí faktoriální návrh 2^{7-4}, hlavní efekty jsou smíšeny s dvou a vícefaktorovými interakcemi.

Ukázka alias struktury - Statistica

Aliasing of Effects (Computed from Generators) (Spreadsheet1 in Workb
2**(7-4) design
(Factors are denoted by numbers)

Factor	124	135	236	1237	2345	1346	1256	347	257
A	24	35	1236	237	12345	346	256	1347	1257
B	14	1235	36	137	345	12346	156	2347	57
C	1234	15	26	127	245	146	12356	47	2357
D	12	1345	2346	12347	235	136	12456	37	2457
E	1245	13	2356	12357	234	13456	126	3457	27
F	1246	1356	23	12367	23456	134	125	3467	2567
G	1247	1357	2367	123	23457	13467	12567	34	25
12	4	235	136	37	1345	2346	56	12347	157
13	234	5	126	27	1245	46	2356	147	12357
23	134	125	6	17	45	1246	1356	247	357
14	2	345	12346	2347	1235	36	2456	137	12457
24	1	12345	346	1347	35	1236	1456	237	457
34	123	145	246	1247	25	16	123456	7	23457
15	245	3	12356	2357	1234	3456	26	13457	127
25	145	123	356	1357	34	123456	16	23457	7
35	12345	1	256	1257	24	1456	1236	457	237

Tabulka ANOVA

Výstup Statistica, Experimental Design (DOE)

Factor	ANOVA; Var.:y; R-sqr=1, (Spreadsheet1 in W 2**(7-4) design DV: y SS	df	MS	F	p
(1)A	12,5	1	12,5		
(2)B	8,0	1	8,0		
(3)C	5832,0	1	5832,0		
(4)D	72,0	1	72,0		
(5)E	1458,0	1	1458,0		
(6)F	0,5	1	0,5		
(7)G	544,5	1	544,5		
Error	0,0	0			
Total SS	7927,5	7			

Nezbývají žádné stupně volnosti pro test významnosti efektů – musíme upravit model.

Redukce modelu

$$C + D + E + G$$

Factor	ANOVA; Var.:y; R-sqr=,99735; Adj:,99382 (NPJ_4_nyty. 4 factors at two levels; MS Residual=7, DV: y				
	SS	df	MS	F	p
(1)C	5832,000	1	5832,000	833,1429	0,000091
(2)D	72,000	1	72,000	10,2857	0,049063
(3)E	1458,000	1	1458,000	208,2857	0,000721
(4)G	544,500	1	544,500	77,7857	0,003072
Error	21,000	3	7,000		
Total SS	7927,500	7			

Příklad 5

Cílem experimentu je zjistit, které faktory mají vliv na pevnost svaru.
Faktory: teplota (A), tlak (B), doba svařování (C), doba držení (D), doba stisknutí (E).

Dílčí faktoriální návrh 2^{5-1}

A	B	C	D	E	y
-	-	-	-	+	1194
+	-	-	-	-	871
-	+	-	-	-	764
+	+	-	-	+	1463
-	-	+	-	-	1205
+	-	+	-	+	1256
-	+	+	-	+	616
+	+	+	-	-	1384
-	-	-	+	-	1152
+	-	-	+	+	1398
-	+	-	+	+	533
+	+	-	+	-	1382
-	-	+	+	+	1170
+	-	+	+	-	920
-	+	+	+	-	776
+	+	+	+	+	1410

Řešení MINITAB

Alias Structure

I + A*B*C*D*E

A + B*C*D*E

B + A*C*D*E

C + A*B*D*E

D + A*B*C*E

E + A*B*C*D

A*B + C*D*E

A*C + B*D*E

A*D + B*C*E

A*E + B*C*D

B*C + A*D*E

B*D + A*C*E

B*E + A*C*D

C*D + A*B*E

C*E + A*B*D

D*E + A*B*C

Interpretace

A + B*C*D*E

Hlavní efekt faktoru A je smíšen se čtyřfaktorovou interakcí BCDE

A*B + C*D*E

Efekt dvoufaktorové interakce AB je smíšen s třífaktorovou interakcí CDE

Při dalším postupu předpokládáme, že jsou uvedené interakce vyššího řádu nulové. Pokud se na základě uvedeného experimentu ukáže např. vliv interakce AB (smíšený s CDE) významný, přisuzujeme zatím efekt interakci AB. V dalším experimentu, např. úplném faktoriálním, se tato hypotéza buď potvrdí nebo ne.

Tabulka ANOVA

Model

$$A + B + C + D + E + AB + AC + ... + DE$$

V tabulce jsou sečteny součty čtverců příslušejících hlavním efektům a součty čtverců dvoufaktorových interakcí.

5 hlavních efektů A B C D E   (5 stupňů volnosti)

10 interakcí AB AC … DE   (10 stupňů volnosti)

Protože nezbyly žádné stupně volnosti pro odhad velikosti reziduální variability, nemůže se určit hodnota testové statistiky F.

```
Analysis of Variance for y (coded units)

Source                 DF   Seq SS   Adj SS   Adj MS    F    P
Main Effects            5   512279   512279   102456    *    *
2-Way Interactions     10   874131   874131    87413    *    *
Residual Error          0        0        0        0
Total                  15  1386410
```

Redukce modelu

Využijeme tabulku s vypočtenými hodnotami efektů a najdeme několik nejmenších: C D BC DE. V Minitabu musíme odstranit i všechny interakce obsahující C nebo D.

Term	Effect	Coef	Term	Effect	Coef
Constant		1093,38	Constant		1093,38
A	334,25	167,12			
B	-104,75	-52,37	A*E	169,25	84,63
C	-2,50	-1,25	B*C	13,50	6,75
D	-1,50	-0,75	B*D	-30,00	-15,00
E	73,25	36,62	B*E	-144,25	-72,13
A*B	403,25	201,63	C*D	-44,75	-22,38
A*C	-33,50	-16,75	C*E	-31,50	-15,75
A*D	35,50	17,75	D*E	-3,00	-1,50

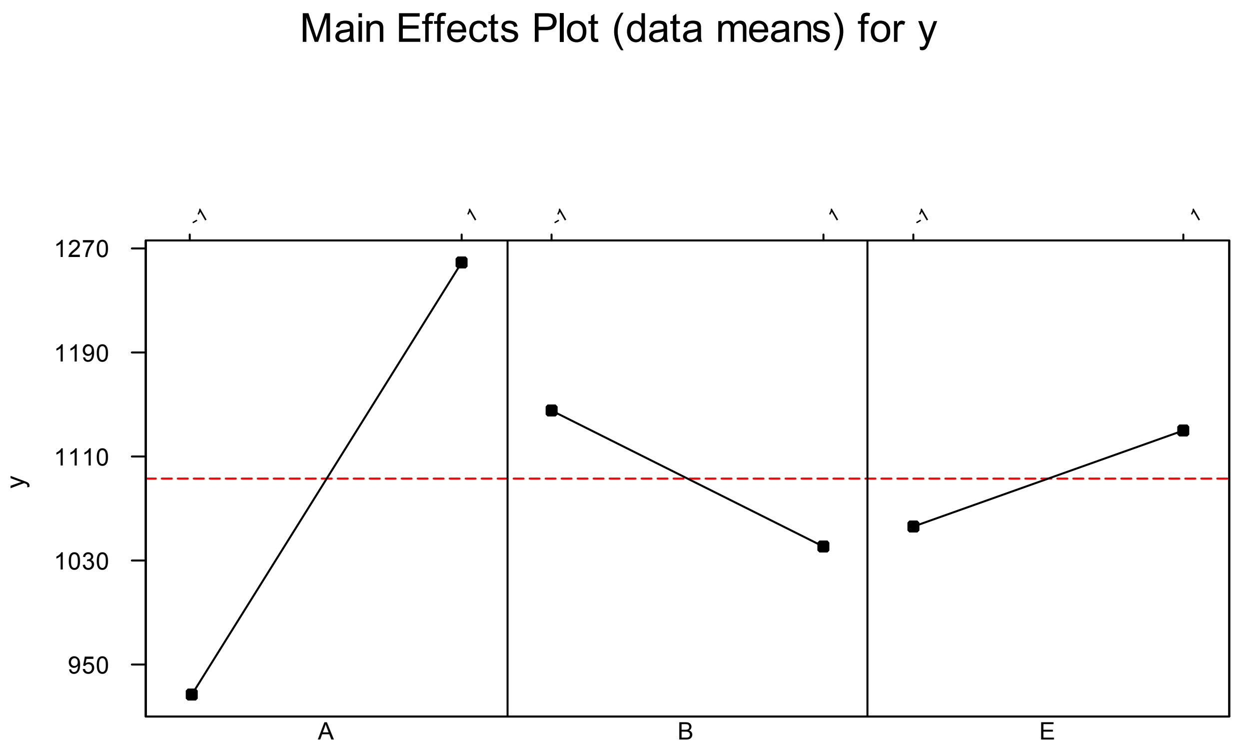
Redukovaný model

Fractional Factorial Fit: y versus A; B; E

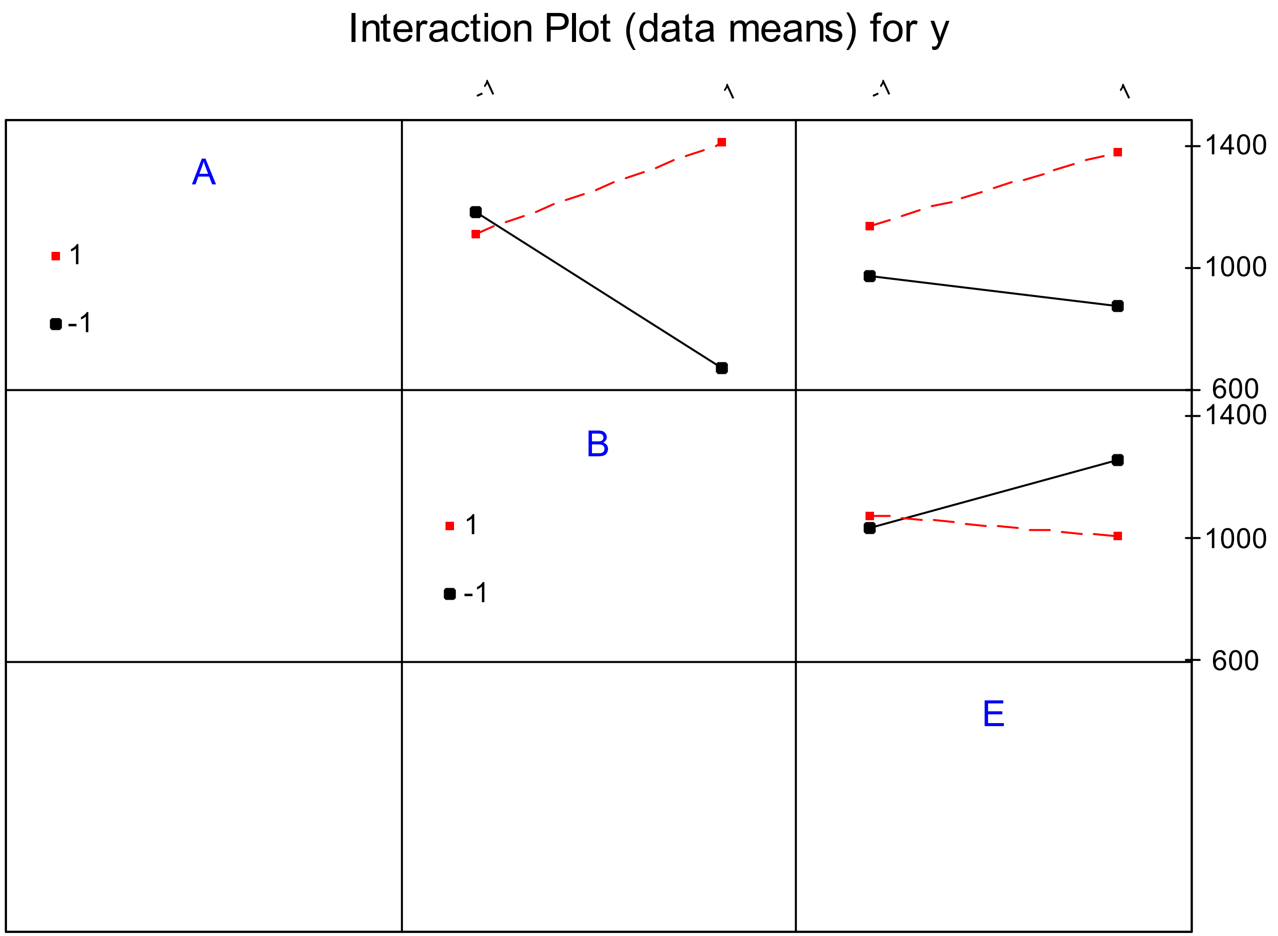
Estimated Effects and Coefficients for y (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		1093,38	13,41	81,51	0,000
A	334,25	167,12	13,41	12,46	0,000
B	-104,75	-52,37	13,41	-3,90	0,004
E	73,25	36,62	13,41	2,73	0,023
A*B	403,25	201,63	13,41	15,03	0,000
A*E	169,25	84,63	13,41	6,31	0,000
B*E	-144,25	-72,12	13,41	-5,38	0,000

Graf hlavních efektů

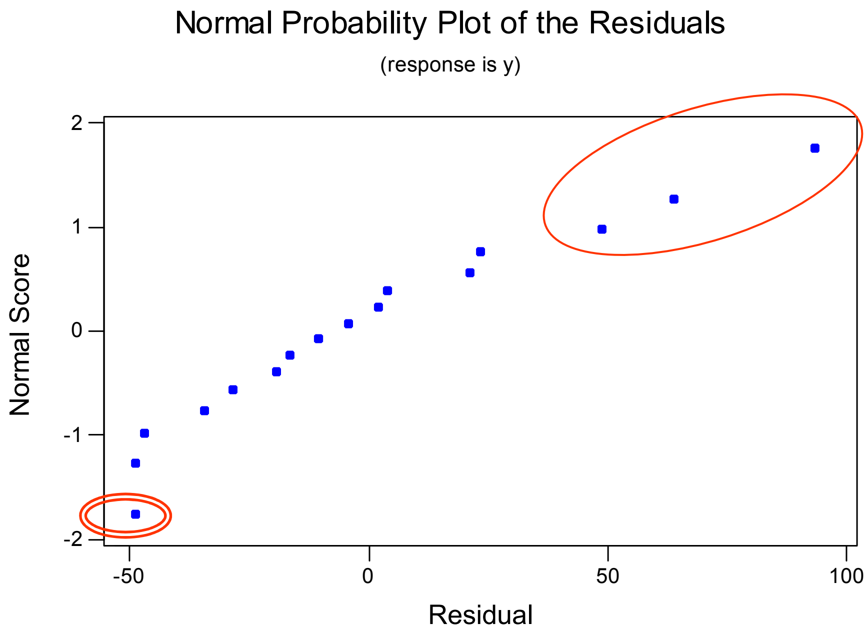


Graf interakcí



Ověření předpokladů

Body vynesené v normálním pravděpodobnostním grafu by měly ležet přibližně v přímce. Vyznačené body se této představě poněkud vymykají. Zvlášť bod vlevo dole by bylo vhodné identifikovat a posoudit, zda příslušné pozorování nemohlo ovlivnit předchozí závěry.



Příklad 6

Experiment pro určení opakovatelnosti a reprodukovatelnosti

základní návrh

n měřených jednotek

g operátorů

r opakovaných měření stejného kusu stejným operátorem

Model ANOVA

efekt interakce jednotka-operátor

$$y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + e_{ijk}$$

efekt jednotky efekt operátora

Zdroj variability	Součet čtverců	Stupně volnosti	Průměrný čtverec	Střední hodnota průměrného čtverce
jednotka	SS_1	$f_1 = n - 1$	MS_1	$gr\sigma_P^2 + r\sigma_{PO}^2 + \sigma_e^2$
operátor	SS_2	$f_2 = g - 1$	MS_2	$nr\sigma_O^2 + r\sigma_{PO}^2 + \sigma_e^2$
jednotka-operátor	SS_3	$f_3 = (n - 1)(g - 1)$	MS_3	$r\sigma_{PO}^2 + \sigma_e^2$
reziduální	SS_4	$f_4 = ng(r - 1)$	MS_4	σ_e^2

F-testy v Excelu

ANOVA

Zdroj variability	SS	St. vol.	MS	F	Hodnota P	F krit
Jednotka	2,058708	9	0,228745	[redacted]		
Operátor	0,048	2	0,024			
Interakce	0,103667	18	0,005759	4,458781	0,000156	1,960117
Reziduální	0,03875	30	0,001292			
Celkem	2,249125	59				

Výpočet statistiky F v prvních dvou řádcích :

$$MS_1 / MS_3 \qquad\qquad MS_2 / MS_3$$

ANOVA

Zdroj variability	SS	St. vol.	MS	F	Hodnota P	F krit
Jednotka	2,058708	9	0,228745	39,71785	4,65E-10	2,456282
Operátor	0,048	2	0,024	4,167203	0,032564	3,554561
Interakce	0,103667	18	0,005759	4,458781	0,000156	1,960117
Reziduální	0,03875	30	0,001292			
Celkem	2,249125	59				

Odhad složek rozptylu

Složka rozptylu	Odhad
Jednotka	$\hat{\sigma}_P^2 = \dfrac{MS_1 - MS_3}{gr}$
Operátor	$\hat{\sigma}_O^2 = \dfrac{MS_2 - MS_3}{nr}$
Interakce	$\hat{\sigma}_{PO}^2 = \dfrac{MS_3 - MS_4}{r}$

Zdroj variability	*Rozptyl*	*Směr. odch.*
Jednotka	0,037164	0,192781
Operátor	0,000912	0,0302
Interakce	0,002234	0,047263