Analýza rozptylu v analytické laboratoři

Milan Meloun a Jiří Militký,
Katedra analytické chemie, Univerzita Pardubice, 532 10 Pardubice
a
Katedra textilních materiálů, Technická univerzita Liberec,
461 17 Liberec

Souhrn: *Analýza rozptylu, ANOVA (z anglického* **Analysis of Variance***) se hojně užívá v technické praxi. Prvním krokem analýzy rozptylu je určit, zda jde o model analýzy rozptylu s pevnými, náhodnými nebo smíšenými efekty. Užitečné je i znát, zda jsou interakce faktorů principiálně možné. Postup analýzy rozptylu lze rozdělit do těchto kroků: 1. Odhad parametrů základního modelu ANOVA. 2. Testování jeho významnosti a konstrukce různých modelů s pevnými efekty. 3. Vyjádření složek rozptylů a testování jejich významnosti. 4. Ověření předpokladů normality, homogenity rozptylů a přítomnosti silně vybočujících pozorování. 5. Interpretace výsledků s ohledem na zadání dat a jejich případné úpravy. Jsou porovnány dva přístupy k analýze rozptylu, software ADSTAT a NCSS2000. Aplikace analýzy rozptylu má význam především k vyhodnocování reprodukovatelnosti a opakovatelnosti kvantitativních výsledků v analytické laboratoři.*

Analýza rozptylu, ANOVA (z anglického **An**alysis **o**f **V**ariance), se v technické praxi používá buď jako samostatná technika nebo jako postup, umožňující analýzu zdrojů variability u lineárních statistických modelů. V technické praxi se ANOVA uplatňuje jako samostatná technika v úlohách:

A. Určení významnosti způsobu přípravy vzorků na výsledek analýzy.

B. Určení vlivu přístroje, lidského faktoru a obsluhy na výsledek měření.

C. Zpracování různých mezilaboratorních experimentů.

D. Zpracování plánovaných experimentů, u kterých se systematicky sleduje vliv rozličných faktorů (teploty, času, koncentrace a dalších) na výsledek reakce či analýzy.

Podstatou analýzy rozptylu je rozklad celkového rozptylu dat na *složky objasněné* (známé zdroje variability) a *složku neobjasněnou*, o níž se předpokládá, že je náhodná. Následně se testují hypotézy o významnosti jednotlivých zdrojů variability.

A. Jednorozměrná analýza rozptylu

Budeme se zabývat případem, kdy daný faktor A má celkem K různých úrovní $A_1, \ldots, A_K$. Na každé úrovni A_i je provedeno n_i měření $\{y_{ij}\}$, $j = 1, \ldots, n_i$. *Celkový počet* měření je

$$N = \sum_{i=1}^{K} n_i$$

Sloupcový průměr $\hat{\mu}_i$ představuje součet prvků sloupce pro A_i dělený počtem opakování n_i,

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

Celkový průměr je součet všech hodnot dělený celkovým počtem dat. Platí také vztah

$$\hat{\mu} = \frac{1}{K} \sum_{i=1}^{K} \hat{\mu}_i$$

Rovnice umožňují stanovení odhadů parametrů μ_i a parametru μ. Pro výpočet odhadů efektů α_i lze pak použít jednoduchý vztah $\hat{\alpha}_i = \hat{\mu}_i - \hat{\mu}$. Při rozkladu μ_i vznikne přeurčený model, obsahující o jeden parametr více. Proto se při odhadu efektů α_i používá ještě jedna omezující podmínka $\sum_{i=1}^{K} n_i \alpha_i = 0$ a pro případ vyvážených experimentů lze použít podmínku $\sum_{i=1}^{K} \alpha_i = 0$.

Základním předpokladem statistické analýzy je, že náhodné chyby ε_{ij} jsou nezávislé náhodné veličiny s normálním rozdělením $N(0, \sigma^2)$. To znamená, že střední hodnota chyb je rovna nule a rozptyl σ^2 je konstantní. Součet čtverců odchylek od celkového průměru $\hat{\mu}$ definovaný vztahem

$$S_c = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu})^2$$

lze rozložit na dvě složky

$$S_c = \sum_{i=1}^{K} \sum_{j=1}^{n_i} [(y_{ij} - \hat{\mu}_i) + (\hat{\mu}_i - \hat{\mu})]^2 = S_A + S_R ,$$

kde S_A představuje rozptyl mezi jednotlivými úrovněmi daného faktoru

$$S_A = \sum_{i=1}^{K} n_i (\hat{\mu}_i - \hat{\mu})^2$$

a S_R je rozptyl reziduální, tj. uvnitř jednotlivých úrovní,

$$S_R = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu}_i)^2$$

Nevychýleným odhadem rozptylu chyb σ_e^2 je průměrný reziduální čtverec

$$\hat{\sigma}_e^2 = \frac{S_R}{N - K}$$

Cílem jednoduchého třídění je především testování, zda jsou efekty α_i nulové, tedy zda jednotlivé úrovně daného faktoru vedou ke statisticky nevýznamným rozdílům ve výsledcích. Nulová hypotéza se testuje dle schema

$$H_0\text{: } \alpha_i = 0,\ i = 1, ..., K, \text{ proti alternativní } H_A\text{: } \alpha_i \neq 0,\ i = 1, ..., K.$$

Při testování se využívá faktu, že veličina S_A / σ_e^2 má χ^2-rozdělení s $(K - 1)$ stupni volnosti a veličina S_R / σ_e^2 má nezávislé χ^2-rozdělení s $(N - K)$ stupni volnosti. Jejich podíl má pak F-rozdělení s $(K - 1)$ a $(N - K)$ stupni volnosti. Testovací statistika F_e má tvar

$$F_e = \frac{S_A (N - K)}{S_R (K - 1)}$$

Při platnosti nulové hypotézy H_0 má pak F_e statistika F-rozdělení s $(K - 1)$ a $(N - K)$ stupni volnosti. Vyjde-li F_e větší než kvantil $F_{1-\alpha}(K - 1, N - K)$, je nutné na hladině významnosti α nulovou hypotézu H_0 vyloučit a efekty považovat za nenulové, čili statisticky významné.

Postup výpočtu

1. Průměry a efekty úrovní: je proveden výpočet parametrů $\hat{\mu}_i$, $\hat{\mu}$, $\hat{\alpha}_i$, reziduí $\hat{e}_{ij}$, Jackknife reziduí e_{Ji} a diagonálních prvků H_{ii} projekční matice $\boldsymbol{H}$. Jsou identifikovány vlivné body, pro které je

$$\hat{e}_{Ji} > 2 \quad \text{a} \quad H_{ii} > 2N / \sum_{i=1}^{N} n_i$$

2. ANOVA tabulka: je sestavena tabulka ANOVA a proveden F-test významnosti efektů faktoru A.

3. Vícenásobné porovnávání Schéffeho procedurou: jsou testovány lineární kontrasty pro zadané kombinace úrovní, H_0: $\mu_i - \mu_j = 0$. Jsou tisknuty významné kontrasty. Je použita Schéffeho metoda mnohonásobného porovnávání.

4. Rankitový *Q-Q* graf: je konstruován rankitový graf Jackknife reziduí a transformační graf závislosti s_i vs. $\hat{\mu}_i$. Pokud jsou všechna data kladná a tato závislost je přibližně lineární, lze zvolit logaritmickou transformaci.

Na vzorové úloze **H5.11** *Vliv tavby na obsah mědi v bronzu* ukážeme postup analýzy: bylo zkoumáno, zda se obsah mědi v bronzu mění od tavby k tavbě. U každé tavby byly odebrány 4 vzorky a stanoven procentuální obsah mědi v bronzu. Na hladině významnosti $\alpha = 0.05$ vyšetřete, zda existuje vliv tavby (faktor A) na obsah mědi v bronzu. Ověřte rovněž homoskedasticitu vhodnou grafickou metodou.

Data: Obsah mědi v bronzu [%] pro různé tavby A, ..., E:

A	B	C	D	E
81	85	87	94	88
77	91	82	90	86
83	88	89	86	91
84	90	87	91	90

1. Průměry a efekty úrovní: je proveden výpočet parametrů sloupcových průměrů $\hat{\mu}_i$, celkového průměru $\hat{\mu}$, sloupcových efektů $\hat{\alpha}_i$, reziduí $\hat{e}_{ij}$.

Program ADSTAT: výstup modulu ANOVA

```
Celkový průměr        =  8.7000E+01
Reziduální rozptyl    =  8.2667E+00
Úroveň   Průměr         Efekt          Hii
  1  8.1250E+01       -5.7500E+00    2.5000E-01
  2  8.8500E+01       1.5000E+00     2.5000E-01
  3  8.6250E+01       -7.5000E-01    2.5000E-01
  4  9.0250E+01       3.2500E+00     2.5000E-01
  5  8.8750E+01       1.7500E+00     2.5000E-01
```

2. ANOVA tabulka: je sestavena tabulka ANOVA a proveden F-test významnosti efektů faktoru A. Jelikož nabývá Fisher-Snedecorovo testační kritérium 5.988 vyšší hodnoty než kvantil F(1-0.05, 7-1, 70-7) = 3.056, je nulová hypotéza H_0: *Efekty faktoru A jsou nulové* zamítnuta a faktor A je statisticky významný.

Program ADSTAT: výstup modulu ANOVA

H0: Efekty faktoru A jsou nulové, HA: ... nejsou nulové
Kvantil F(1-alfa,k-1,n-k) = 3.056

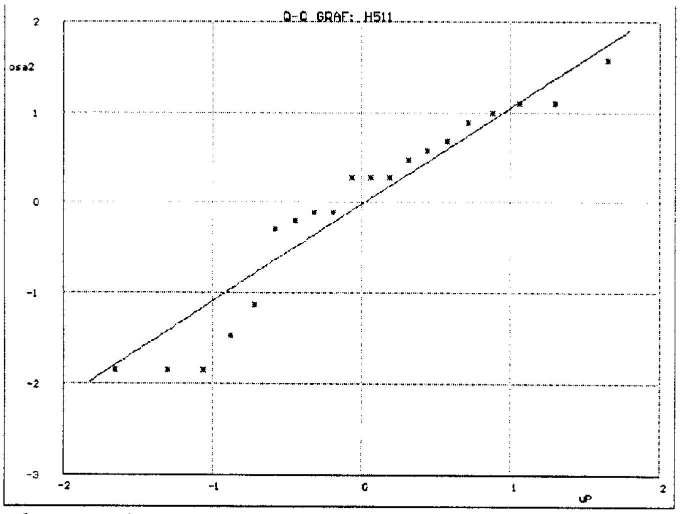
Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H0 je	Spočtená hlad.výz.
Mezi úrovněmi	k-1 = 4	1.9800E+02	4.9500E+01	5.988	Zamítnuta	0.004
Rezidua	n-k = 15	1.2400E+02	8.2667E+00			
Celkový	n-1 = 19	3.2200E+02	1.6947E+01			

3. Vícenásobné porovnávání Schéffeho procedurou: jsou testovány lineární kontrasty pro zadané kombinace úrovní, H_0: $\mu_i - \mu_j = 0$, Sheffého metodou mnohonásobného porovnávání. Tři dvojice P1=P2, P1=P4, P1=P5 vycházejí odlišně od ostatních, totiž nerovnají se sobě. Ostatní sloupce jsou co do hodnoty považovány za shodné.

Program ADSTAT: Vícenásobné porovnávání Schéffeho procedurou

Hypotéza H0	Průměrný párový rozdíl	Meze konfidenčního intervalu dolní	horní	Závěr
P1=P2	-7.250	-14.358	-0.142	Zamítnuta
P1=P3	-5.000	-12.108	2.108	Akceptována
P1=P4	-9.000	-16.108	-1.892	Zamítnuta
P1=P5	-7.500	-14.608	-0.392	Zamítnuta
P2=P3	2.250	-4.858	9.358	Akceptována
P2=P4	-1.750	-8.858	5.358	Akceptována
P2=P5	-0.250	-7.358	6.858	Akceptována
P3=P4	-4.000	-11.108	3.108	Akceptována
P3=P5	-2.500	-9.608	4.608	Akceptována
P4=P5	1.500	-5.608	8.608	Akceptována

4. Rankitový *Q-Q* graf: je konstruován rankitový graf Jackknife reziduí a transformační graf závislosti s_i vs. $\hat{\mu}_i$. Rankitový graf dokazuje, že většina hodnot odpovídá navrženému modelu, totiž jednofaktorové analýze rozptylu. V dolní části grafu je několik odlehlých hodnot, zbytek dobře splňuje lineární závislost. Transformační graf přináší body v náhodném mraku, což dokazuje, že transformace dat není nutná.



Obr. 1 Rankitový graf reziduí, **ADSTAT**

Program NCSS2000: Tests of Assumptions Section (*Oddíl testování předpokladů*).

Assumption	Test Value	Prob Level	Decision (0.05)
Skewness Normality of Residuals	-0.2942	0.768618	Accept
Kurtosis Normality of Residuals	-1.9323	0.053327	Accept
Omnibus Normality of Residuals	3.8202	0.148066	Accept
Modified-Levene Equal-Variance Test	0.1048	0.990183	Accept

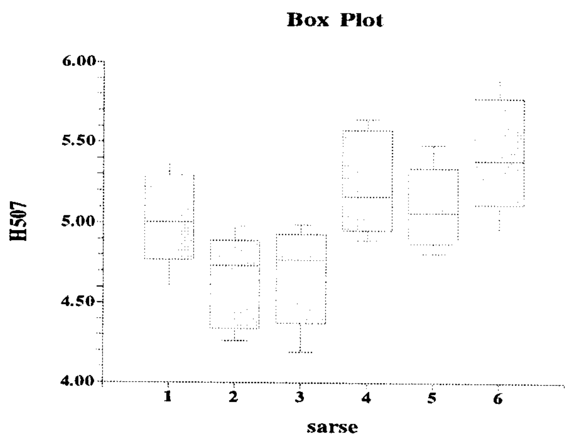
Vysvětlení: *přináší výsledky testů, ověřujících předpoklady normality a homoskedasticity (shodnosti rozptylů). ANOVA předpokládá, že kombinovaná rezidua (tj. odchylky od sloupcových průměrů) jsou normální. Ostatní předpoklady se týkají nezávislostí a náhodných chyb, jež však nejsou zde testovány. Musíte sledovat sloupec nadepsaný Decision(5%). Jestliže žádný z těchto testů nebyl zamítnut, můžeme si být jisti, že předpoklady jsou splněny. Dvě vlastnosti jsou zde testovány, šikmost a špičatost. Jestliže normalita reziduí selhává kvůli šikmosti, je možné k zajištění normality užít odmocninovou nebo logaritmickou transformaci původních dat. Bylo shledáno, že modifikovaný Levenův test je jedním z nejvýhodnějších a nejsilnějších testů homoskedasticity. Proto užijeme jako předběžný test vždy modifikovaný Levenův test.* **Normality (Skewness, Kurtosis, and Omnibus):** *tři testy umožňují testování šikmosti, špičatosti a celkové normality dat. Jestliže některý z testů zamítá normalitu, data nemohou být považována za Gaussovská, normální.*

Equal-Variance Test (Modified Levene): *modifikovaný test homoskedasticity byl shledán jedním z nejlepších na test homoskedasticity, tj. rovnosti rozptylů.*

Program NCSS2000: Kruskal-Wallis One-Way ANOVA on Ranks *(Kruskal-Wallisův test mediánů)*

Hypotheses					
Ho: All medians are equal.					
Ha: At least two medians are different.					
Method	**DF**	**Chi-Square (H)**	**Prob Level**	**Decision(0.05)**	
Not Corrected for Ties	5	15.88645	0.007176	Reject Ho	
Corrected for Ties	5	15.88999	0.007165	Reject Ho	
Number Sets of Ties	1				
Multiplicity Factor		6			
Group	**Count**	**Sum of Ranks**	**Mean Rank**	**Z-Value**	**Median**
1	5	80.50	16.10	0.1669	5.01
2	5	34.00	6.80	-2.4207	4.74
3	5	38.00	7.60	-2.1981	4.77
4	5	104.00	20.80	1.4747	5.17
5	5	87.50	17.50	0.5565	5.07
6	5	121.00	24.20	2.4207	5.39

***Vysvětlení:** V případě nenormality se užije Kruskal-Wallisův neparametrický test. Základní předpoklady o nezávislosti výběru, spojitosti náhodných prvků, a měřícím měřítku jsou pro tento test dodrženy. Kruskall-Wallisův test má však další předpoklad, že rozdělení ve sloupcích jsou stejná (i když nemusí být Gaussovská) a co do druhu ale i tvaru (tj. rozptylu) musí být stejná a mohou se lišit pouze v míře polohy (tj. mediánech). Konečně ověříme výsledky těchto testů na krabicových grafech sloupců. Jak je dále ukázáno, tento graf dobře zobrazí dodržení předpokladů normality a homoskedasticity.*



Obr. 2 Krabicový graf jednotlivých šarší, *NCSS2000*

Program NCSS2000: Analysis of Variance Table *(Tabulka analýzy rozptylu)*

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level (Alpha=0.05)
A: sarse	5	2.469747	0.4939493	5.21	0.002238*
S(A)	24	2.27624	9.484334E-02		
Total (Adj.)	29	4.745986			
Total	30				
* Term significant at alpha = 0.05					

***Vysvětlení:** tabulka přináší základní tabulku analýzy rozptylu.* **Source Term:** *zdroj proměnlivosti čili testovaný faktor.* **DF:** *stupně volnosti jsou zde počty pozorování (či naměřených hodnot), užívaných tímto faktorem.* **Sum of Squares:** *suma čtverců pro tento zdroj proměnlivosti. Uvádí se obvykle v ANOVA tabulce pro úplnost, nikoliv pro přímou interpretaci.* **Mean Square:** *Průměrný čtverec je vlastně odhad rozptylu způsobený tímto faktorem. Jde o sumu čtverců dělenou stupni volnosti.* **F-Ratio:** *Poměr průměrných čtverců tohoto faktoru v čitateli ku chybovému ve jmenovateli. Jde o hodnotu Fisher-Snedecorova F-testu.* **Prob Level:** *Hladina významnosti α pro F-test. Jde o pravděpodobnost, že*

F_{krit} poměr je větší než F_{exp}. Např. pro hladinu významnosti α = 0.05 bude pro statisticky významný faktor hodnota vypočteného α menší než 0.05. U výsledku F-testů se objeví hvězdička, když bude faktor statisticky významný.*

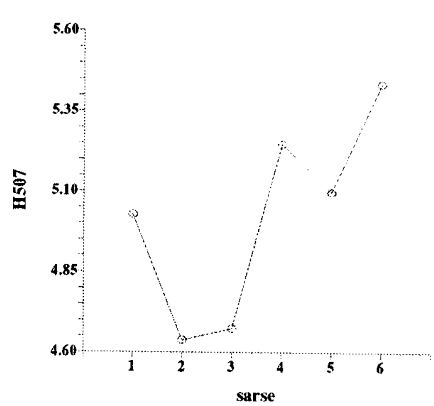
Program NCSS2000: Means and Effects Section *(Oddíl průměrů a efektů)*

Term	Count	Mean	Standard Error	Effect
All	30	5.019333		1.003867
A: sarse				
1	5	5.028	0.1377268	4.024133
2	5	4.638	0.1377268	3.634133
3	5	4.674	0.1377268	3.670133
4	5	5.246	0.1377268	4.242133
5	5	5.096	0.1377268	4.092134
6	5	5.434	0.1377268	4.430133

***Vysvětlení:** tabulka přináší průměry a efekty jednotlivých faktorů.* **Term:** *značí faktor.* **Count:** *počet pozorování či měření, zahrnuté v průměru.* **Mean:** *hodnota výběrového aritmetického průměru.* **Standard Error:** *standardní chyba průměru. Všimněte si, že standardní chyby jsou jednoduše odmocniny průměrného čtverce chybového výrazu pro daný faktor, podělené četností. Tyto standardní chyby nejsou totožné s jednoduchými standardními chybami, vypočtenými odděleně pro každý sloupec. Standardní chyby jsou vhodné k testování vzájemného porovnávání.* **Effect:** *Vypočtený příspěvek faktoru do průměru. Např. průměr prvního sloupce je roven sumě všech efektů (v řádku All) plus efekt prvního faktoru či sloupce.*

Plot of Means *(Grafy průměrů)*

Zobrazení průměru analyzovaných dat. Lze sledovat tvary závislosti v grafu.



Obr. 3 Diagram sloupcových průměrů

B. Dvoufaktorová analýza rozptylu bez opakování v cele

Při dvoufaktorové analýze rozptylu se provádí experimenty na různých úrovních dvou faktorů A a B. Kombinace úrovní faktorů tvoří typickou mřížkovou strukturu, jejímž elementem je tzv. ***cela.*** Platí, že (i, j)-tá cela odpovídá kombinaci úrovně A_i faktoru A a B_j faktoru B. V každé cele je obecně n_{ij} pozorování. Často se však setkáváme s případem ***bez opakování***, kdy v každé cele je pouze jediné pozorování, n_{ij} = 1. Kromě řádkových α_i a sloupcových β_j efektů se zde vyskytuje také interakční člen τ_{ij}. Tento člen je důsledkem různých kombinací sloupcových a řádkových efektů.

	B_1	B_2	...	B_M
A_1	.	.	...	.
A_2	.		...	.
.	.	.	...	.
.	.	.	...	.
.	.	.	...	.

cela A_2 B_2

A_N	.	.	...	.

Nejjednodušší je ***Tukeyův model interakce***, vyjádřený tvarem

$$\tau_{ij} = C\,\alpha_i\,\beta_j$$

kde C je konstanta. U těchto modelů obsahuje každá cela právě jednu hodnotu y_{ij}. O chybách ϵ_{ij} se předpokládá, že jsou to nezávislé a stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konstantním rozptylem. K testování se navíc předpokládá, že rozdělení chyb je normální. Definují se zde následující omezující podmínky

$$\sum_{i=1}^{N} \alpha_i = 0; \quad \sum_{j=1}^{M} \beta_j = 0; \quad \sum_{i=1}^{N} \tau_{ij} = 0; \quad \sum_{j=1}^{M} \tau_{ij} = 0$$

V případě pouze aditivního působení jednotlivých faktorů je $\tau_{ij} = 0$ pro všechna $i = 1, \ldots, N$ a $j = 1, \ldots, M$. Odhady parametrů μ, α_i, β_j lze pak určit ze vztahů

$$\hat{\mu} = \frac{1}{N\,M} \sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij}$$

$$\hat{\alpha}_i = \frac{1}{M} \sum_{j=1}^{M} y_{ij} - \hat{\mu}$$

$$\hat{\beta}_j = \frac{1}{N} \sum_{i=1}^{N} y_{ij} - \hat{\mu}$$

Pro rezidua $\hat{e}_{ij}$ platí výraz $\hat{e}_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$. K určení interakcí můžeme využít skutečnosti, že $\tau_{ij} = E(y_{ij}) - \mu - \alpha_i - \beta_j$, a pro odhad interakcí platí přibližně $\hat{\tau}_{ij} \approx \hat{e}_{ij}$.

Lze snadno identifikovat ***Tukeyův model*** interakce. Platí-li tento model, vyjde na grafu $\hat{e}_{ij}$ vs. $\hat{\alpha}_i \hat{\beta}_j$ lineární trend. Ze směrnice odpovídající regresní přímky se odhadne parametr C. Platí pro něj výraz

$$\hat{C} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \hat{e}_{ij}\,\hat{\alpha}_i\,\hat{\beta}_j}{\sum_{i=1}^{N} \sum_{j=1}^{M} \hat{\alpha}_i^2\,\hat{\beta}_j^2}\,.$$

Graf $\hat{e}_{ij}$ vs. $\hat{\alpha}_i\,\hat{\beta}_j / \hat{\mu}$ se označuje jako graf ***neaditivity***. Pokud vyjde v tomto grafu nenáhodný trend, znamená to, že je třeba uvažovat interakce. V tabulce představuje S_T součet čtverců odchylek odpovídající Tukeyově interakci

$$S_T = \frac{\left(\sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij}\,\hat{\alpha}_i\,\hat{\beta}_j\right)^2}{\sum_{i=1}^{N} \sum_{j=1}^{M} \hat{\alpha}_i^2\,\beta_j^2}$$

Analýza rozptylu pro dvojné třídění s interakcí Tukeyova typu

Součet čtverců pro	Stupně volnosti	Průměrný čtverec	Kritérium F

Faktor A $S_A = M \sum_{i=1}^{N} \alpha_i^2$	$N - 1$	$M_A = S_A/(N-1)$	$F_A = M_A/M_{AB}$
Faktor B $S_B = N \sum_{j=1}^{M} \beta_j^2$	$M - 1$	$M_B = S_B/(M-1)$	$F_B = M_B/M_{AB}$
Interakce (Tukey) S_T=rov.(5.53)	1	$M_T = S_T$	$F_T = M_T/M_E$
Reziduální $S_R = S_{AB} - S_T$	$NM - N - M$	$M_E = S_R/(NM-N-M)$	-
Celkový $S_C = \sum_{(i)} \sum_{(j)} (\hat{\mu} - y_{ij})^2$	$NM - 1$	-	-

Symbol S_{AB} označuje reziduální součet čtverců pro případ bez interakcí

$$S_{AB} = \sum_{i=1}^{N} \sum_{j=1}^{M} (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2$$

Odpovídající průměrný čtverec je $M_{AB} = \dfrac{S_{AB}}{(N-1)(M-1)}$. Hodnota M_{AB} je nevychýleným odhadem rozptylu σ^2. Pomocí F-kritéria lze opět provádět statistické testy. Začíná se testováním nulové hypotézy H_0: "Tukeyova interakce je nevýznamná", pro kterou lze použít testovací statistiku F_T. Za předpokladu platnosti nulové hypotézy H_0 má tato testovací statistika F-rozdělení s 1 a $(NM - N - M)$ stupni volnosti. Pokud nelze tuto hypotézu zamítnout, testuje se nulová hypotéza H_0: $\alpha_i = 0$, $i = 1, ..., N$ (efekty řádků, čili faktoru A, jsou nevýznamné) pomocí statistiky F_A nebo nulová hypotéza H_0: $\beta_j = 0, j = 1, ..., M$ (efekty sloupců, čili faktoru B, jsou nevýznamné) pomocí statistiky F_B. Obě tyto testovací statistiky jsou uvedeny v tabulce. Za předpokladu platnosti hypotézy H_0 má statistika F_A F-rozdělení s $(N - 1)$ a $(N - 1)(M - 1)$ stupni volnosti a statistika F_B také F-rozdělení s $(M - 1)$ a $(N - 1)(M - 1)$ stupni volnosti. Pokud však vyjde F_T vyšší než odpovídající kvantil F-rozdělení, je efekt Tukeyovy interakce významný.

Postup výpočtu

Je určen pro dvoufaktorovou analýzu rozptylu a modelu s pevnými efekty pro případ $n_{ij} = 1$, tedy bez opakování v cele. Předpokládá se interakce Tukeyova typu. Provádí se odhady parametrů, testy významnosti a ověření interakce, resp. transformace, vedoucí k aditivitě efektů. Je možné použít také robustní techniku *median polish*. Vstupem je obdélníková tabulka dat pro $A_1, ..., A_N$ úrovní faktoru A a $B_1, ..., B_M$ úrovní faktoru B, $\{y_{ij}\}$, $j = 1, ..., N$ a $j = 1, ..., M$. Pro všechny testy je standardně uvažována hladina významnosti $\alpha = 0.05$. Výstup výpočtu obsahuje

1. Průměry a efekty úrovní: jsou vypočteny parametry: celkový průměr $\hat{\mu}$, řádkové efekty $\hat{\alpha}_i$, sloupcové efekty $\hat{\beta}_j$, interakční člen $\hat{\tau}_{ij}$ a Tukeyho konstanta C.

2. ANOVA tabulka: je sestavena tabulka ANOVA a provedeny F-testy významnosti faktorů, resp. interakcí, včetně kombinovaných testů pro ověření celkové významnosti faktorů A, B.

3. Graf neaditivity: je kreslen graf neaditivity včetně určení optimální mocninné transformace λ pro zajištění aditivity. Lze zadat provedení analýzy pro transformovaná data, pokud jsou kladná.

4. Odhad parametrů metodou "median polish": lze zvolit odhad parametrů modelu ANOVA metodou "median polish" včetně určení robustních reziduí robustního grafu neaditivity a rankitového grafu.

Na úloze **H5.02** Vliv druhu svářecího kovu na pevnost sváru o zadání: vazebným článkem sváru zirkonové slitiny bývá nikl, železo a měď. Byly vytvořeny sváry o rozličném složení těchto svařecích komponent a cílem je vyšetřit pevnost sváru, t. zn. největší tlak v tisíci librách na čtvereční palec k přerušení sváru. Na sedmi svárech (faktor A) a hladině významnosti $\alpha = 0.05$ vyšetřete, zda záleží na druhu kovu (faktor B), užitého ve svářecím drátu, čili zda tlaky k přerušení sváru jsou u všech drátů stejné. Prozkoumejte, zda je třeba provést transformaci vedoucí ke stabilizaci rozptylu.

Data: Tlak k roztržení sváru zirkonové slitiny [10^3 liber/palec^2] pro sedm svárů a tři druhy svářecích drátů:

Svár	Nikl	Železo	Měď
1	67.0	71.9	72.2
2	67.5	68.8	66.4
3	76.0	82.6	74.5
4	72.7	78.1	67.3
5	73.1	74.2	73.2
6	65.8	70.8	68.7
7	75.6	84.9	69.0

1. Průměry a efekty úrovní: jsou vypočteny parametry celkový průměr $\hat{\mu}$, řádkové efekty $\hat{\alpha}_i$, sloupcové efekty $\hat{\beta}_j$, interakční člen $\hat{\tau}_{ij}$ a konstanta C.

Program ADSTAT: výstup modulu ANOVA

Celkový průměr = 7.2395E+01

Reziduální rozptyl = 1.1314E+01

FAKTOR A:			FAKTOR B:		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	7.0367E+01	-2.0286E+00	1	7.1100E+01	-1.2952E+00
2	6.7567E+01	-4.8286E+00	2	7.5900E+01	3.5048E+00
3	7.7700E+01	5.3048E+00	3	7.0186E+01	-2.2095E+00
4	7.2700E+01	3.0476E-01			
5	7.3500E+01	1.1048E+00			
6	6.8433E+01	-3.9619E+00			
7	7.6500E+01	4.1048E+00			

Tukeyho C = 1.5313E-01

2. ANOVA tabulka: je sestavena tabulka ANOVA a provedeny F-testy významnosti faktorů, resp. interakcí, včetně kombinovaných testů pro ověření celkové významnosti faktorů A, B.

H0: Efekty faktoru A jsou nulové, HA: ... nejsou nulové
Kvantil F(1-alfa,n-1,mn-m-n) = 3.095
H0: Efekty faktoru B jsou nulové, HA: ... nejsou nulové
Kvantil F(1-alfa,m-1,mn-m-n) = 3.982
H0: Interakce I je nulová, HA: ... není nulová
Kvantil F(1-alfa,1,mn-m-n) = 4.844 (Zde I znamená efekt Tukeyho interakce.)

Zdroj rozptylu	**Stupně volnosti**	**Součet čtverců**	**Průměrný čtverec**	**Testovací kriterium**	**Závěr H0 je**	**Spočtená hlad. význ.**
Mezi úrovněmi A, n-1 =	6	2.6829E+02	4.4715E+01	3.952	Zamítnuta	0.024
Mezi úrovněmi B, m-1 =	2	1.3190E+02	6.5951E+01	5.829	Zamítnuta	0.019
Interakce	1	3.9517E+01	3.9517E+01	3.493	Akceptována	0.088
Rezidua mn-m-n =	11	1.2446E+02	1.1314E+01			
Celkový mn-1 =	20	5.2465E+02	2.6232E+01			

TRANSFORMACE:

Odhad mocninné transformace : -1.0086E+01

Rozptyl odhadu transformace : 3.2258E+01

Akceptovatelný interval : (-1.5766E+01,-4.4066E+00)

Jelikož hodnota testačního kritéria 3.952 je vyšší než kvantil Fisher-Snedecorova rozdělení 3.095, je nulová hypotéza o nevýznamnosti faktoru A (druh sváru) zamítnuta a svár je statisticky významným faktorem. Jelikož je hodnota druhého testačního kritéria 5.829 vyšší než kvantil Fisher-Snedecorova rozdělení 3.982, je nulová hypotéza o nevýznamnosti faktoru B (druh kovu svářecího drátu) zamítnuta a tento kov je zde statisticky významným faktorem. Interakce má fyzikální významn, a proto ji budeme testovat: jelikož hodnota testačního kritéria 3.493 je nižší než kvantil Fisher-Snedecorova rozdělení 4.844, je nulová hypotéza o nevýznamnosti interakčního členu AB (interakce druhu sváru s kovem svářecího drátu) přijata a interakce je statisticky nevýznamná.

Jelikož odhad mocninné transformace -10.086 leží v akceptovatelném intervalu -15.766 až -4.4066, není třeba data transformovat mocninnou nebo Box-Coxovou transformací.

Následující výstup programu NCSS2000 přináší numericky poněkud jiné výsledky, protože tento program užívá při testování jiné stupně volnosti (totiž rovné počtu pozorování).

Program NCSS2000: Analysis of Variance Table

Source Term	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=0.05)
A: Svar	6	268.2895	44.71492	4.31	0.015087*	
B: Kov slitiny	2	131.901	65.95048	6.36	0.013094*	
AB	12	124.459	10.37159			
S	0	0				
Total (Adjusted)	20	524.6495				
Total	21					
* Term significant at alpha = 0.05						

Vysvětlení: přináší základní tabulku analýzy rozptylu. **Source Term:** *zdrojem proměnlivosti čili faktorem je zde svar, kov slitiny, nebo jejich interakce AB.* **DF:** *pro stupně volnosti jsou zde užity počty pozorování (či naměřených hodnot), užívaných dotyčným faktorem.* **Sum of Squares:** *suma čtverců pro tento faktor. Uvádí se zde spíše pro úplnost, než pro přímou interpretaci.* **Mean Square:** *průměrný čtverec představuje odhad rozptylu, způsobený tímto faktorem. Jde o sumu čtverců dělenou stupni volnosti.* **F-Ratio:** *Poměr průměrných čtverců testovaného faktoru v čitateli ku chybovému ve jmenovateli. Jde o hodnotu Fisher-Snedecorova testačního kritéria F.* **Prob Level:** *vypočtená hladina významnosti α pro uvažovaný F-test. Jde vlastně o vypočtenou pravděpodobnost, že F_{krit} je větší než F_{exp}. Např. pro uživatelem zvolenou hladinu významnosti $\alpha = 0.05$ bude pro statisticky významný faktor vypočtená hodnota hladiny významnosti α menší než 0.05. U hodnoty testačního kritéria F se pak objeví hvězdička*.* **Power (Alpha=0.05):** *síla testu obecně vyjadřuje pravděpodobnost zamítnutí hypotézy, že průměry jsou si rovny, když ony si ve skutečnosti rovny nejsou. Jde o hodnotu 1-α, tj. chybu typu II. Síla testu závisí na velikosti výběru, hodnotách rozptylu, zvolené hodnotě hladiny významnosti α a konečně na aktuálním rozdílu mezi průměry. Síla testu zde vyčíslená značí, že směrodatná odchylka souboru je rovna směrodatné odchylce pozorované a že rozdíl mezi průměry souboru jsou přesně rovny rozdílu mezi průměry výběru. Vysoká hodnota síly testu je totiž žádoucí. Vysoká hodnota značí, že je vysoká pravděpodobnost zamítnutí nulové hypotézy, když nulová hypotéza neplatí. To je kritická míra přesnosti v testování. Síla testu poskytuje informaci co máme udělat, aby naše výsledky byly významné. Jestliže přijmeme nulovou hypotézu s vysokou silou testu nezbyde nám už mnoho ke pozměnění. Konečně víme, že průměry* **nejsou** *různé. Když však přijmeme nulovou hypotézu s nízkou silou testu, můžeme si vybrat jednu z následujících možností:*

(1) Zvýšit hodnotu α. Lépe snad pak testovat $\alpha = 0.05$ než původní $\alpha = 0.01$. Zvýšením α se zvýší poněkud síla testu.

(2) Zvýšení velikosti výběru způsobí zvýšení síly testu, když byla dříve nízká. Když byla síla testu původně vysoká, zvýšení velikosti výběru bude mít malý efekt.

(3) Snížení velikosti rozptylu. Snad můžeme přeformulovat model a měření se stanou přesnějšími a extrémní zdroje proměnlivosti se odstraní).

Program NCSS2000: Means and Effects Section

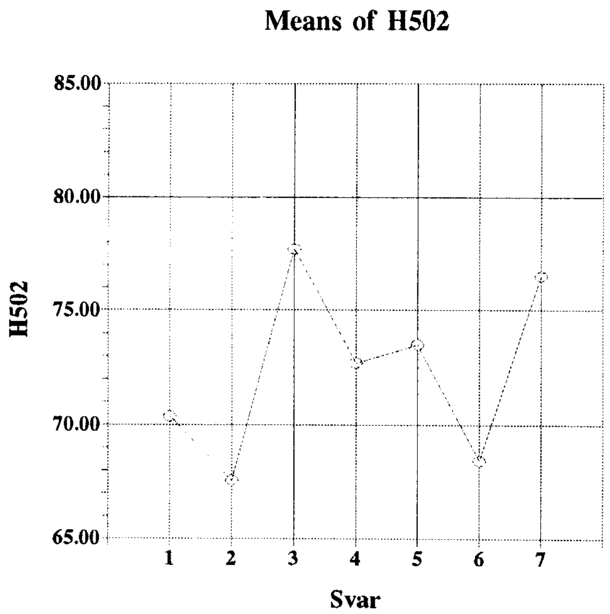
Term	Count	Mean	Standard Error	Effect
All	21	72.39524		72.39524
A: Svar				
1	3	70.36667	1.859354	-2.028571
2	3	67.56667	1.859354	-4.828571
3	3	77.7	1.859354	5.304762
4	3	72.7	1.859354	0.3047619

5	3	73.5	1.859354	1.104762
6	3	68.43333	1.859354	-3.961905
7	3	76.5	1.859354	4.104762
B: Slitina				
1	7	71.1	1.217233	-1.295238
2	7	75.9	1.217233	3.504762
3	7	70.18571	1.217233	-2.209524
AB: Svar,Slitina				
1,1	1	67	0	-2.071429
1,2	1	71.9	0	-1.971429
1,3	1	72.2	0	4.042857
2,1	1	67.5	0	1.228571
2,2	1	68.8	0	-2.271429
2,3	1	66.4	0	1.042857
3,1	1	76	0	-0.4047619
3,2	1	82.6	0	1.395238
3,3	1	74.5	0	-0.9904762
4,1	1	72.7	0	1.295238
4,2	1	78.1	0	1.895238
4,3	1	67.3	0	-3.190476
5,1	1	73.1	0	0.8952381
5,2	1	74.2	0	-2.804762
5,3	1	73.2	0	1.909524
6,1	1	65.8	0	-1.338095
6,2	1	70.8	0	-1.138095
6,3	1	68.7	0	2.476191
7,1	1	75.6	0	0.3952381
7,2	1	84.9	0	4.895238
7,3	1	69	0	-5.290476

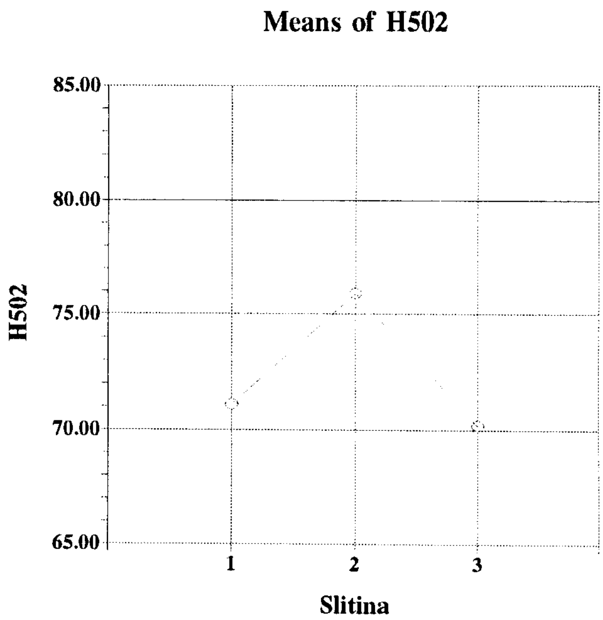
Vysvětlení: *přináší průměry a efekty jednotlivých faktorů.* **Term:** *testovaný faktor.* **Count:** *počet pozorování, měření zahrnuté do průměru.* **Mean:** *hodnota výběrového aritmetického průměru.* **Standard Error:** *standardní chyba průměru. Všimněte si, že standardní chyby jsou zde počítány jako odmocniny průměrného čtverce chybového výrazu pro daný faktor dělené četností. Tyto standardní chyby* ***nejsou totožné*** *s jednoduchými standardními chybami, vypočtenými odděleně pro každý sloupec. Standardní chyby jsou vhodné pouze k vzájemnému porovnávání.* **Effect:** *vypočtený příspěvek faktoru do průměru. Např. průměr prvního sloupce (skupiny) je roven sumě všech efektů (v řádku All) plus efekt prvního faktoru či sloupce.*

Plot of Means (*Grafy průměrů*)

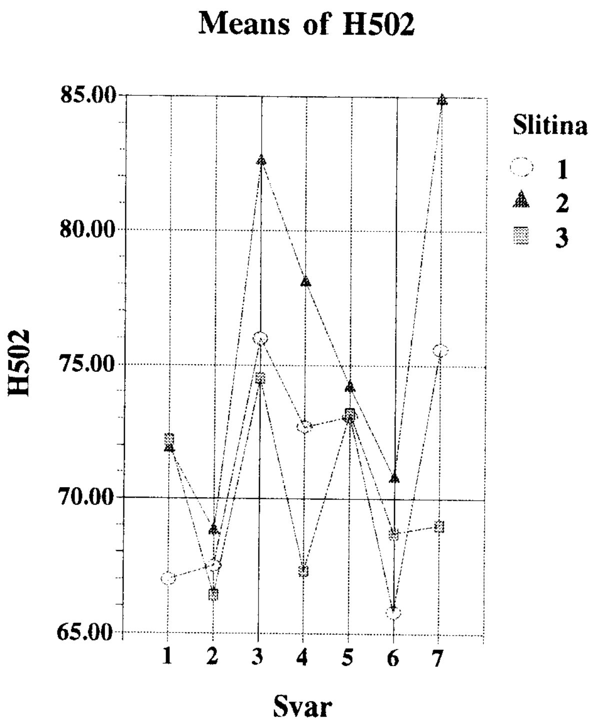
Zobrazení průměru analyzovaných dat. Lze sledovat tvary závislosti v grafu.



Obr. 4 Diagram průměrů pro různé sváry



Obr. 5 Diagram průměrů pro různé kovy slitiny



Obr. 6 Diagram průměrů při působení obou faktorů

C. Vyvážená dvoufaktorová analýza rozptylu

Slouží ke dvoufaktorové analýze rozptylu u vyvážených experimentů $n_{ij} = n$ a modelů s pevnými efekty. Je hledán optimální model ANOVA, odhadnuty jeho parametry a provedeny testy významnosti. Vstupem jsou pro úrovně A_1, ..., A_N faktoru A a úrovně B_1, ..., B_M faktoru B hodnoty $\{y_{ijk}\}$, $i = 1, ..., N$, $j = 1, ..., M$ a $k = 1, ..., n$. Pro všechny testy je standardně uvažována hladina významnosti $\alpha = 0.05$.

Postup výpočtu

Výstup výpočtu obsahuje

1. Průměry a efekty úrovní: jsou vypočteny parametry celkový průměr, řádkové efekty $\hat{\alpha}_i$, sloupcové efekty $\hat{\beta}_j$, interakční člen $\hat{\tau}_{ij}$ a rezidua $\hat{e}_{ijk}$.

2. ANOVA tabulka: je sestavena tabulka ANOVA a provedeny F-testy významnosti faktorů, resp. interakcí, včetně kombinovaných testů pro ověření celkové významnosti faktorů A a B.

3. Transformační graf: je kreslena závislost výběrových směrodatných odchylek s_{ij} v celách na průměrech $\hat{\mu}_{ij}$. Pokud je nalezena monotónní závislost, lze zadat vhodnou transformaci, ve které se provede opakovaná analýza.

4. Rankitový *Q-Q* graf: je konstruován rankitový graf pro rezidua $\hat{e}_{ijk}$.

Závěr

Jsou porovnány dva přístupy k analýze rozptylu, a to dvěma rozličnými programy, ADSTAT a NCSS2000. U programu NCSS2000 je užito chybného počtu stupňů volnosti, totiž rovného počtu pozorování a proto dochází k numerické neshodě. ADSTAT dosahuje správných hodnot.

Aplikace analýzy rozptylu má význam především k vyhodnocování reprodukovatelnosti a opakovatelnosti kvantitativních výsledků v analytické laboratoři.

Doporučená literatura:

1. M. Meloun, J. Militký: STATISTICKÉ ZPRACOVÁNÍ EXPERIMENTÁLNÍCH DAT - SBÍRKA ÚLOH S DISKETOU, Nakladatelství Univerzita Pardubice 1996.
2. M. Meloun, J. Militký: STATISTICKÉ ZPRACOVÁNÍ EXPERIMENTÁLNÍCH DAT, Nakladatelství PLUS Praha 1994, EAST PUBLISHING Praha 1998.
3. M. Meloun, J. Militký, M. Forina: CHEMOMETRICS FOR ANALYTICAL CHEMISTRY, Volume 1. PC-AIDED STATISTICAL DATA ANALYSIS, Ellis Horwood Chichester 1992.
4. M. Meloun, J. Militký, M. Forina: CHEMOMETRICS FOR ANALYTICAL CHEMISTRY, Volume 2. PC-AIDED REGRESSION AND RELATED METHODS, Ellis Horwood Chichester 1992.