

Analýza rozptylu, 1 faktor

Menu:	QCExpert	Anova	Jednofaktorová
-------	----------	-------	----------------

ANOVA je všeobecně používaná zkratka anglického Analysis Of Variance, česky analýza rozptylu. Modul ANOVA slouží k porovnání různých zdrojů nebo vlastností různých tříd materiálu na základě změřených hodnot nebo charakteristik. Tyto zdroje se nazývají *faktory* a mohou to být například čtyři různé laboratoře (což odpovídá čtyřem úrovním faktoru „laboratoř“), několik úrovní znečištění, kvalita výrobku ze tří stejných linek. Cílem je rozhodnout, zda se střední hodnota měřené veličiny pro různé faktory liší, či nikoliv. To se prokazuje testováním hypotézy o vlivu faktoru na střední hodnotu. Předpokladem pro výpočet je normalita dat pro jednotlivé úrovně faktorů a nepřítomnost vybočujících měření. Výpočet je doplněn tabulkou a grafem z-skóre, které porovnají střední hodnoty ve sloupcích s předepsanou nebo vypočítanou hodnotou a intervalem $\pm 3\sigma$ s ohledem na jejich rozptyl. Z-skóre se často používá při výběru dodavatele a při mezilaboratorních zkouškách.

Data a parametry

Podle zaškrtnutých položek *Druhu analýzy* se vypočítá základní jednofaktorová ANOVA a rozšířené porovnání všech dvojic úrovní mezi sebou (*Párové porovnání*) podle Scheffého. Pokud vyjde některá z úrovní významně odlišná v Scheffého porovnání i v závěru celkové analýzy, doporučuje se odlišné sloupce z výpočtu vyloučit a přepočítat analýzu znovu. Může se stát, že v důsledku snížení rozptylu po odstranění odlišující se úrovně (vyloučením příslušného sloupce z výpočtu) se odliší další úroveň. Tak lze postupovat až do potvrzení nevýznamnosti vlivu faktoru. Není-li zadána (zaškrtnuty, viz Obrázek 1) předepsaná referenční hodnota (*Střed*) a její směrodatná odchylka, použijí se při konstrukci grafu a tabulky z-skóre celkový průměr a reziduální směrodatná odchylka. Hladina významnosti určuje významnost všech testů, obvyklá hodnota je 0.05. Data je možné zadávat ve dvou formátech, které specifikujeme knoflíky *Ve sloupcích* a *Podle faktoru*.

Uspořádání dat ve sloupcích:

Data jsou ve sloupcích, každý sloupec odpovídá jedné úrovni faktoru. Počet dat v jednotlivých sloupcích nemusí být stejný. Minimální počet sloupců (tedy úrovní faktoru) je 2. Minimální počet řádků je 2. Názvy sloupců by měly odpovídat úrovním analyzovaného faktoru, např. Linka A, Linka B, Linka C. Výběr sloupců z aktuálního listu se může provést v poli *Sloupce* dialogového panelu ANOVA, Obrázek 1. Tlačítko *Vyber vše* označí všechny sloupce v daném listu. Implicitně jsou vybrány všechny sloupce obsažené v aktuálním listu. *Příklad dat:*

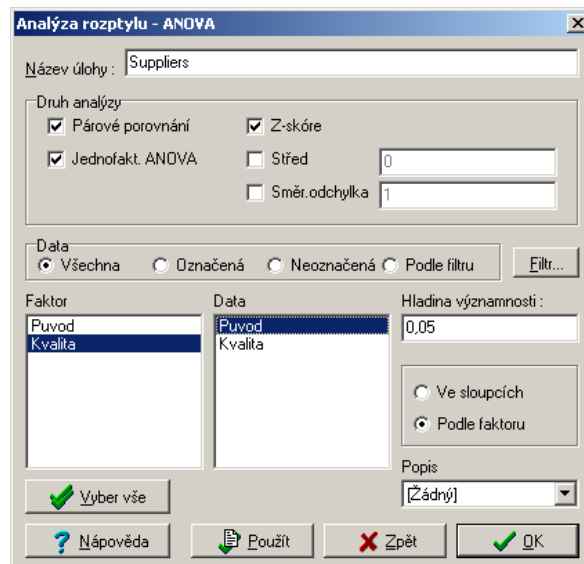
Externí lab.	Lab 1	Lab 2	Lab 3
1.47	1.24	2.32	3.6
1.75	0.94	2.4	7.3
1.09	1.84	1.45	2.65
3.09	0.3	1.86	8.2

Uspořádání dat podle faktoru:

Data jsou ve dvou sloupcích. V jednom sloupci je úroveň faktoru (sloupec *Faktor*), ve druhém jsou příslušné hodnoty (sloupec *Data*).

Příklad dat:

Původ	Kvalita	Původ	Kvalita
UKR	17.17	SWE	23.19
GER	23.73	BRA	26.87
GER	23.7	ARG	24.59
ARG	24.78	UKR	19.5
BRA	27.91	BRA	26.94



Obrázek 1 ANOVA - Vstupní dialogový panel

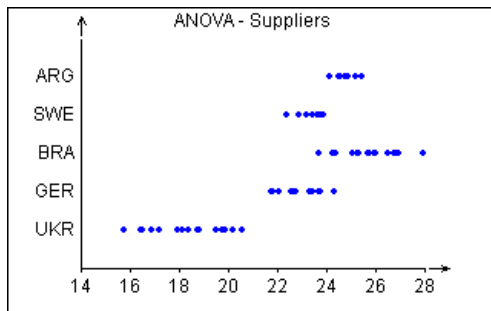
Protokol

Analýza rozptylu - ANOVA	
Počet úrovní faktoru	Počet sloupců vzatých do výpočtu.
Sloupec	
Počet hodnot	Počet platných dat ve sloupci.
Efekty faktorů	Střední vliv dané úrovně faktoru na odchylku hodnoty proti celkovému průměru, tedy rozdíl průměru úrovně a celkového průměru.
Průměr úrovně	Průměr dat pro jednotlivé úrovně (sloupce).
Celkový průměr	Průměr všech dat.
Celkový rozptyl	Rozptyl všech dat.
Průměrný čtverec	Průměrný čtverec rozdílu dat od celkového průměru.
Reziduální rozptyl	Střední rozptyl dat uvnitř jednotlivých úrovní.
Reziduální součet čtverců	Součet čtverců odchylek uvnitř jednotlivých úrovní, tedy rozdílů dat a průměru příslušné úrovně. Tato hodnota odpovídá předpokladu, že faktory mají vliv a vysvětlují část variability.
Celkový součet čtverců	Součet čtverců rozdílů dat od celkového průměru. Tato hodnota odpovídá předpokladu, že faktory nemají vliv a variabilita je způsobena jen náhodnou chybou.
Vysvětlený součet čtverců	Rozdíl celkového a reziduálního součtu čtverců. Odpovídá variabilitě, kterou lze přisoudit rozdílům mezi úrovněmi faktoru.
Závěr	Závěr celkové analýzy, tedy výsledek testu, zda má různá úroveň faktoru nějaký vliv na měřenou veličinu.
Faktor	Významnost nebo nevýznamnost vlivu faktoru.
Teoret.	Kritický kvantil pro danou hladinu významnosti.
Vypočt.	Testovací hodnota.
Pravděpodobnost	Pravděpodobnost nevýznamnosti vlivu faktoru; je-li menší než zvolená hladina významnosti, je vliv faktoru statisticky významný.
Porovnání dvojic úrovní, Scheffého metoda	Podrobné porovnání všech jednotlivých dvojic úrovní faktoru navzájem.

Srovnávané dvojice	Rozdíl	Rozdíl mezi středními hodnotami v úrovních.
	Závěr	Významnost nebo nevýznamnost rozdílu.
Pravděpodobnost		Pravděpodobnost nevýznamnosti, je-li menší než zvolená hladina významnosti, je rozdíl mezi úrovněmi statisticky významný.
Z-skóre		Srovnání střední hodnoty v jednotlivých úrovních s referenční hodnotou a intervalem 3 sigma.
Střední hodnota		Normovaná střední hodnota. Za nevyhovující lze považovat hodnoty větší než 3 a menší než -3.
95% interval		Normovaná polovina intervalu spolehlivosti stř. hodnoty.
	Rozdíl	Normovaná střední hodnota rozšířená o svůj 95% interval spolehlivosti, maximální očekávaný rozdíl od nastavené požadované střední hodnoty. Za nevyhovující lze považovat hodnoty větší než 3 a menší než -3.

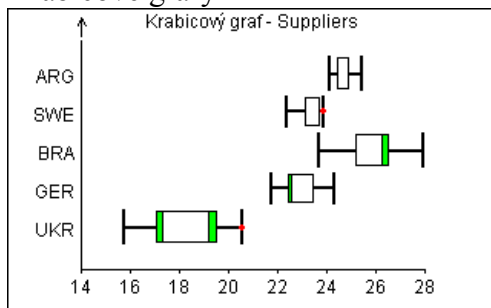
Graf

Graf ANOVA



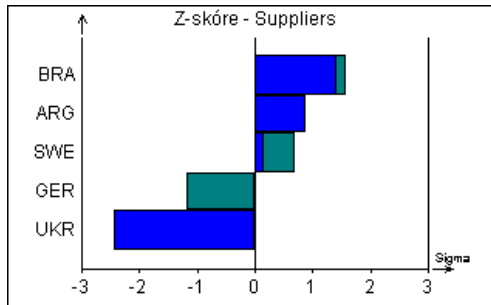
Graf ANOVA zobrazuje polohu měřených dat, takže je možné vizuálně posoudit rozdíly a rozptyl v jednotlivých úrovních. Na ose y je měřená hodnota, na ose x jsou zadány úrovně faktoru. Graf může sloužit i pro identifikaci odlehlých měření ve spolupráci s modulem *Základní statistika*.

Krabicové grafy



Krabicový graf - s výhodou jej lze využít k porovnání hodnot pro jednotlivé úrovně faktoru. Graf může sloužit i pro identifikaci odlehlých měření ve spolupráci s modulem *Základní statistika*.

Graf Z-skóre



Srovnání střední hodnoty v jednotlivých úrovních s referenční hodnotou umístěnou v nule a intervalem ± 3 sigma označeným svislými přímkami. Vnitřní část sloupce je úroveň normované střední hodnoty. Za nevyhovující lze považovat hodnoty větší než 3 a menší než -3. Vnější část sloupce je normovaná střední hodnota rozšířená o svůj 95% interval spolehlivosti. Chceme-li uvažovat průměry včetně intervalu spolehlivosti, je za nevyhovující nutno považovat hodnoty větší než 3 a menší než -3.