

Data4: Naměřené hodnoty fiktivní veličiny, např. výška rostlin v cm.

Faktor1: První typ ošetření (treatment) zálivkou, 1 kóduje kontrolní plochy (normální zálivka), 2 kóduje plochy ošetřené dvojnásobnou zálivkou, faktor je s pevným efektem.

Data4 Faktor1

33	1
35	1
36	1
38	1
40	1
42	1
52	2
53	2
56	2
63	2
62	2
60	2

Příklad 7: Jednocestná analýza variance (1-way ANOVA), dvě hladiny faktoru s pevným efektem.

Použité proměnné: Data4, Faktor1

Zajímá nás otázka, jestli je výška rostliny závislá na množství zálivky. Založíme experiment, ve kterém budeme po určitou dobu jednu skupinu rostlin zalévat standardním množstvím vody a druhou skupinu rostlin dvojnásobným množstvím vody. Tento typ příkladů, kdy vysvětlujeme závislost spojitě proměnné proměnnou kategoriální (tzv. faktor), řešíme zde analýzou variance (ANOVA). Uvedený příklad odpovídá nejjednoduššímu modelu ANOVy – jednocestné (jednofaktoriální) analýze variance se dvěma hladinami jednoho faktoru (zálivka jedním či dvěma díly vody). Zadání dat do tabulky je následující: v jednom sloupečku zadáme studovanou proměnnou (výška rostliny) a ve druhém sloupečku kódujeme hladiny faktoru, 1 odpovídá zálivce jedním dílem vody, 2 pak zálivce dvěma díly vody.

Formulujeme nulovou hypotézu H_0 : Míra zálivky nemá vliv na růst rostlin. Smyslem analýzy variance je ověřit na jaké hladině pravděpodobnosti chyby 1. řádu můžeme tuto nulovou hypotézu zamítnout. Proceduru jednocestné analýzy variance voláme v oddíle **Analysis/ANOVA/One-way Anova** (ale lze použít i další možnosti, tedy **Analysis of Variance, GLM ANOVA**). Do řádku *Response Variable(s)* zadáme příslušný sloupeček s měřenou výškou rostlin, do řádku *Factor Variable* zadáme sloupeček, který kóduje hladiny faktoru. Jedná se o faktor s pevným efektem, proto v dalším řádku ponecháme typ faktoru *Fixed*. Zavoláním *Run/Run Procedure* získáme výpočet.

Analysis of Variance Report

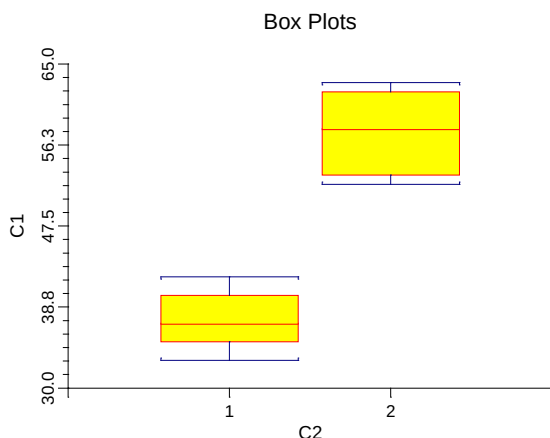
Proto, abychom mohli použít ANOVu, musí být splněny určité předpoklady. Jedním z nich je podmínka, aby naše data odpovídala normálnímu rozdělení. První oddíl výsledků nám nabízí test normality rozdělení reziduálů. Ve všech případech je výsledné doporučení *Accept*, tudíž můžeme předpokládat, že podmínky pro použití parametrického modelu ANOVy jsou splněny. Důležitá poznámka: Použití těchto testů je správnější, než použití testů o normálním rozdělení samotné závislé proměnné (viz Příklad6)!

Tests of Assumptions Section

Assumption	Test Value	Prob Level	Decision (0.05)
Skewness Normality of Residuals	-0.0558	0.955515	Accept
Kurtosis Normality of Residuals	-1.6177	0.105725	Accept
Omnibus Normality of Residuals	2.6201	0.269807	Accept
Modified-Levene Equal-Variance Test	1.9512	0.192680	Accept

Již pouhou vizuální kontrolou krabicových diagramů pro jednotlivé hladiny faktoru můžeme předvídat vlastní výsledek ANOVy. Když se dvojice krabic (tedy jejich mezikvartilová rozpětí) nepřekrývá, jako je to v tomto případě, ve většině případů se dá očekávat, že bude ANOVou nalezen průkazný rozdíl.

Box Plot Section



Tato část se týká obecného modelu použité analýzy variance. První řádek odpovídá testovanému faktoru, druhý řádek pak reziduální (zbytkové, nevysvětlené vlivem faktoru) variabilitě. Smyslem ANOVy je zjistit, jestli je signifikantní příspěvek té části variance, která je způsobena působením faktoru. Jinými slovy jestli variance mezi skupinami (among groups, mezi hladinami faktoru) je signifikantně vyšší než variance uvnitř skupin (within groups, v rámci jednotlivých hladin faktoru). Testové kritérium je tedy dáno podílem variance mezi skupinami ($S+sA$) a variance uvnitř skupin ($S(A)$).

Expected Mean Squares Section

Source	DF	Term	Denominator	Expected
Term		Fixed?	Term	Mean Square
A (C2)	1	Yes	S(A)	$S+sA$
S(A)	10	No		$S(A)$

Note: Expected Mean Squares are for the balanced cell-frequency case.

Tento oddíl je hlavním výsledkem analýzy variance. Řádek uvedený jako A(C2) odpovídá vlivu příslušného faktoru, řádek uvedený jako S(A) odpovídá reziduální variabilitě. Protože máme jen dvě hladiny faktoru, je odpovídající počet stupňů volnosti (DF) roven 1, počet stupňů volnosti reziduálů je roven počtu měření (12) sníženému o počet hladin faktoru (2), tedy 10. Dále jsou uvedeny sumy čtverců (*Sum of Squares*), průměrný čtverec (*Mean Square*, toto je vlastní variance – suma čtverců vydělená počtem stupňů volnosti). Podíl variance (průměrného čtverce) faktoru a reziduálů je testovací kritérium, *F-Ratio*. *Prob Level* udává hladinu pravděpodobnosti odpovídající získané hodnotě *F* statistiky (defaultová hodnota je 5%), hladina pravděpodobnosti je v tomto případě proklatě nízká. Závěrem tedy můžeme zamítnout nulovou hypotézu a na dosažené hladině pravděpodobnosti přijmout hypotézu alternativní, tedy že vyšší míra závlivky znamená průkazně vyšší růst rostlin.

Analysis of Variance Table

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level	Power (Alpha=(0.05))
A(C2)	1	1240.333	1240.333	75.32	0.000006*	1.000000
S(A)	10	164.6667	16.46667			
Total (Adjusted)	11	1405				
Total	12					

* Term significant at alpha = 0.05

Pokud by nebyl splněn některý z předpokladů pro použití parametrické analýzy variance, je možné použít metodu neparametrickou, tzv. Kruskal-Wallisovu ANOVu podle pořadí, založenou na testování polohy mediánů. Její výpočet je v NCSS proveden automaticky, pokud v zadávacím okně ANOVy tuto možnost před vlastním výpočtem nezměníme. Výsledek neparametrické ANOVy je zcela ve shodě s parametrickým testem, tedy můžeme zamítnout nulovou hypotézu ($p=0.004$).

Kruskal-Wallis One-Way ANOVA on Ranks

Hypotheses

H_0 : All medians are equal.

H_a : At least two medians are different.

Test Results

Method	DF	Chi-Square (H)	Prob Level	Decision(0.05)
Not Corrected for Ties	1	8.307693	0.003948	Reject Ho
Corrected for Ties	1	8.307693	0.003948	Reject Ho
Number Sets of Ties	0			
Multiplicity Factor	0			

Group Detail

Group	Count	Sum of Ranks	Mean Rank	Z-Value	Median
1	6	21.00	3.50	-2.8823	37
2	6	57.00	9.50	2.8823	58

Means and Effects Section

Term	Count	Mean	Standard Error	Effect
All	12	47.5		7.916667
A: C2				
1	6	37.33333	1.656636	29.41667
2	6	57.66667	1.656636	49.75

Plots of Means Section

