

UNIVERZITA PARDUBICE
Fakulta chemicko-technologická
Katedra analytické chemie

*Licenční studium Pythagoras – Statistické zpracování
experimentálních dat*

Semestrální práce
ANOVA

vypracoval:
Ing. David Dušek
*Výzkumný ústav lesního hospodářství a myslivosti, v.v.i.
Výzkumná stanice Opočno*

duben 2008

OBSAH

ÚLOHA 1 – JEDNOFAKTOROVÁ ANOVA	2
1) ZADÁNÍ	2
2) ŘEŠENÍ	2
3) ZÁVĚR.....	3
ÚLOHA 2 – DVOUFAKTOROVÁ ANOVA BEZ OPAKOVÁNÍ.....	4
1) ZADÁNÍ	4
2) ŘEŠENÍ	0
3) ZÁVĚR.....	5
ÚLOHA 3 – DVOUFAKTOROVÁ ANOVA S OPAKOVÁNÍM.....	6
1) ZADÁNÍ	6
2) ŘEŠENÍ	6
3) ZÁVĚR.....	8

ÚLOHA 1 – JEDNOFAKTOROVÁ ANOVA

1) ZADÁNÍ

Bylo stanoveno pH půdy v porostu smrku pichlavého (SMP), na holině s buření (bur) a v porostu smrku ztepilého (SMZ). V každém porostu bylo odebráno sedm vzorků půdy. Jednofaktorovou analýzou rozptylu bylo ověřováno, lze-li střední hodnoty pH půdních vzorků jednotlivých porostů považovat za shodné či nikoliv. Byl použit software ADSTAT a NCSS.

Data: pH

SMP	4.2	4.0	4.1	3.8	3.9	3.8	4.0
bur	4.1	3.9	3.8	3.9	4.0	4.1	3.9
SMZ	4.2	4.3	4.2	4.2	4.3	4.2	4.2

2) ŘEŠENÍ

Testy základních předpokladů:

Testovaný předpoklad	Hodnota testu	Pravděpodobnost	H_0 akceptována? ($\alpha=0.05$)
Normální šikmost reziduí	0.559	0.576	Ano
Normální špičatost reziduí	0.031	0.975	Ano
Omnibus test normality reziduí	0.314	0.855	Ano
Mod. Levenův test homogenity rozptylů	2.154	0.145	Ano

Průměry a efekty úrovní:

Úroveň	Průměr	Efekt	H_{ii}
1	3.9714	-0.0810	0.1429
2	3.9571	-0.0952	0.1429
3	4.2286	0.17619	0.1429

Detekce vlivných bodů:

Závěr	Odlehlé	Vybočující
	Ano: 1	Žádné

Tabulka ANOVA:

Zdroj variability	df	Součet čtverců	Průměrný čtverec	F-Ratio	Pravděpodobnost	Síla testu ($\alpha=0.05$)
Úrovně	2	0.3267	0.1633	13.03	0.0003	0.991
Reziduální	18	0.2257	0.0125			
Celkem (adj.)	20	0.5524				
Celkem	21					

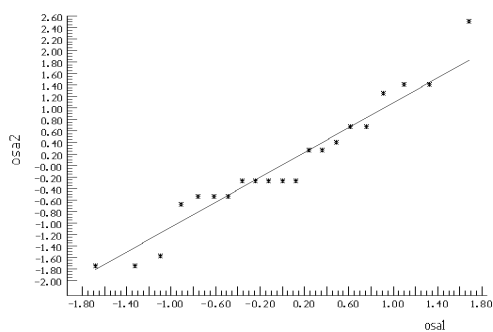
Scheffeho metoda vícenásobného porovnání:

Hypotéza	Průměrný párový rozdíl	Meze konfidenčního intervalu		Závěr
H0		dolní	horní	
P1 = P2	0.014	-0.145	0.174	Akceptována
P1 = P3	-0.257	-0.417	-0.098	Zamítnuta
P2 = P3	-0.271	-0.431	-0.112	Zamítnuta

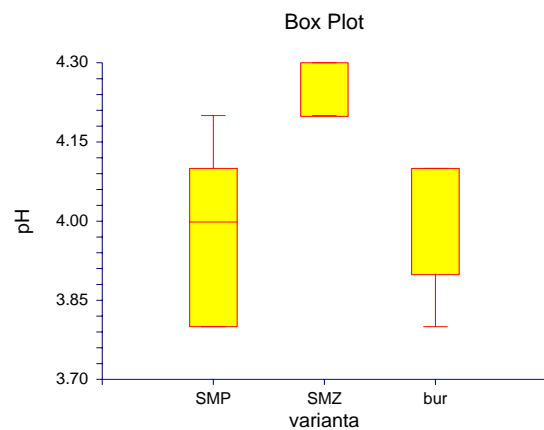
P1=SPM, P2=bur, P3=SMZ

Zkouška transformace:

Korelační koeficient, R: -0.941



Obr.1: Q-Q graf reziduí



Obr.2: Krabicové grafy

3) ZÁVĚR

Hypotézu o normalitě reziduí a hypotézu o shodě rozptylů nebylo možno na hladině alfa 0.05 zamítnout. Byla identifikována jedna odlehlá hodnota. To je patrné i z Q-Q grafu reziduí (Obr.1). Vypočítaný korelační koeficient nelze považovat za blízký nule, ale po provedené logaritmické transformaci se jeho hodnota vůbec nezměnila, proto je uvedena analýza netransformovaných dat. Z tabulky ANOVA je zřejmé zamítnutí nulové hypotézy o shodě středních hodnot výběrů, a proto byl proveden Scheffeho test. Na základě tohoto testu lze konstatovat, že pH půdy pod porostem smrku ztepilého (SMZ) je statisticky signifikantně odlišné (vyšší) od zbývajících dvou ploch (SMP a bur). Rozdíly v pH jsou do jisté míry patrné i z krabicových grafů (Obr.2).

ÚLOHA 2 – DVOUFAKTOROVÁ ANOVA BEZ OPAKOVÁNÍ

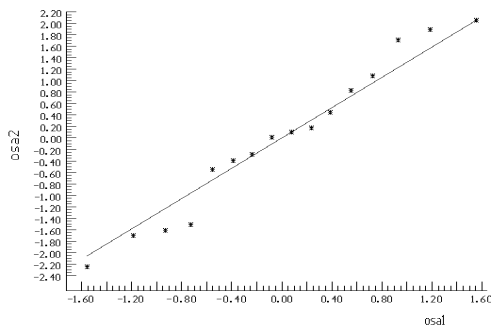
1) ZADÁNÍ

Byl zjišťován vliv tloušťky stromu, vyjádřené jeho příslušností do daného tloušťkového stupně (faktor A) a vliv olistění (faktor B) na velikost tloušťkového přírůstu stromů. Výpočet byl proveden pomocí programu ADSTAT 2.0.

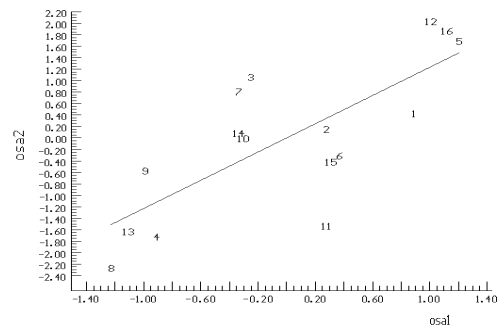
Data: tloušťkový přírůst za dvouleté období (mm)

		Olistění			
		1 (30 – 40 %)	2 (50 – 60 %)	3 (70 – 80 %)	4 (90 – 100 %)
Tloušťkový stupeň	1 (6 cm)	0.25	1.67	4.00	3.00
	2 (8 cm)	1.06	0.75	3.29	2.00
	3 (10 cm)	2.00	4.25	4.15	9.50
	4 (12 cm)	1.10	4.50	5.43	9.50

2) ŘEŠENÍ



Obr.3: Q-Q graf reziduí



Obr.4: Graf neaditivity

Průměry a efekty úrovní:

Celkový průměr = 3.528

Reziduální rozptyl = 3.273

FAKTOR A:			FAKTOR B:		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	2.230	-1.298	1	1.103	-2.426
2	1.775	-1.753	2	2.793	-0.736
3	4.975	1.447	3	4.218	0.689
4	5.133	1.604	4	6.000	2.472

Tukeyho C = 3.4794E-01

Tabulka ANOVA pro model s Tukeyho interakcí:

Faktor A: Kvantil $F(1-\alpha, n-1, mn-m-n) = 4.066$

Faktor B: Kvantil $F(1-\alpha, m-1, mn-m-n) = 4.066$

Tukeyho interakce: Kvantil $F(1-\alpha, 1, mn-m-n) = 5.318$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H_0 je	Spočtená hlad.výz.
Mezi úrovněmi A	$n-1 = 3$	37.704	12.568	3.840	Akceptována	0.057
Mezi levels B	$m-1 = 3$	52.041	17.347	5.300	Zamítnuta	0.026
Interakce	1	14.847	14.847	4.536	Akceptována	0.066
Rezidua	$mn-m-n = 8$	26.184	3.273			
Celkový	$mn-1 = 15$	115.93	7.729			

Transformace:

Odhad mocninné transformace: -0.228

Rozptyl odhadu transformace: 0.294

Akceptovatelný interval: (-0.771, 0.316)

3) ZÁVĚR

Kvantilový graf reziduí (Obr.3) ukazuje na jejich přibližně normální rozdělení. V grafu neaditivity (Obr.4) je patrný nenáhodný trend, aditivitu by bylo možno zlepšit mocninnou transformací, i když ne statisticky významně. Byla akceptována nulová hypotéza o nevýznamnosti Tukeyho interakce mezi faktory. Faktor tloušťka stromu (A) vyšel jako statisticky nesignifikantní, ale nulová hypotéza nebyla zamítnuta na hladině významnosti 0.05 jen velmi těsně. Nulová hypotéza pro faktor olistění (B) byla zamítnuta, hodnota spočtené statistiky je větší než kritický kvantil, stupeň olistění tedy má signifikantní vliv na tloušťkový přírůst.

ÚLOHA 3 – DVOUFAKTOROVÁ ANOVA S OPAKOVÁNÍM

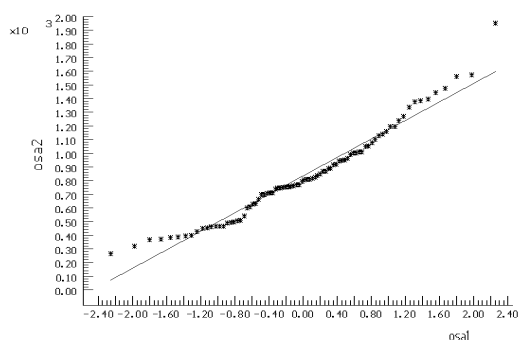
1) ZADÁNÍ

Na třech lokalitách (faktor B) – v porostu smrku pichlavého (SMP), na holině (bur) a v porostu smrku ztepilého (SMZ) – byly odebrány vzorky půdy pro stanovení obsahu Ca ve čtyřech půdních horizontech (faktor A: L, F, H, Ah). Na každé lokalitě bylo odebráno vždy sedm vzorků pro každý půdní horizont. K analýze byl použit software ADSTAT a UNISTAT.

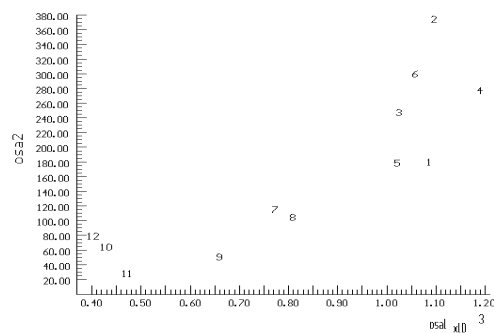
Data: obsah Ca ($\text{mg}\cdot\text{kg}^{-1}$)

	SMP	bur	SMZ		SMP	bur	SMZ
L	1254	860	1082	H	858	798	674
L	904	806	744	H	736	746	724
L	1224	848	1568	H	938	712	570
L	984	956	1054	H	760	674	594
L	874	1206	994	H	814	806	670
L	992	1964	812	H	760	1000	674
L	1382	1022	920	H	530	914	708
F	1640	918	820	Ah	559	489	382
F	766	972	772	Ah	465	515	399
F	1224	804	1462	Ah	414	448	282
F	1168	880	1352	Ah	448	416	337
F	884	1112	1392	Ah	427	473	389
F	1194	1394	770	Ah	358	489	512
F	1450	1056	824	Ah	348	484	524

2) ŘEŠENÍ



Obr. 5: Q-Q graf reziduí původních dat

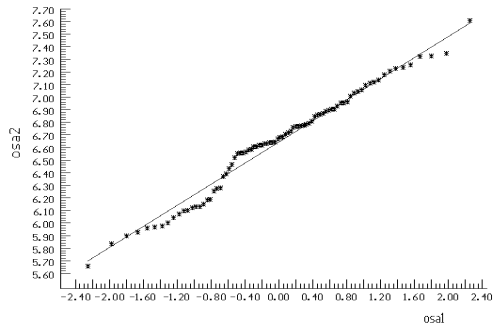


Obr.6: Graf neaditivity původních dat

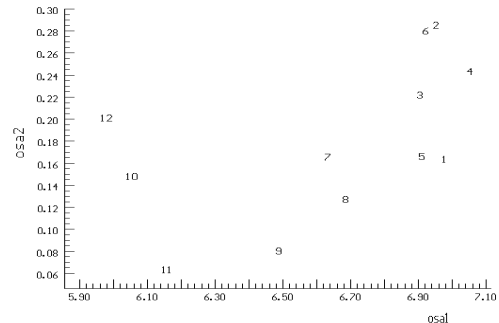
Zkouška transformace:

Korelační koeficient, $R: 0.852$

Q-Q graf (Obr.5) pro původní netransformovaná data indikuje rozdělení reziduí odlišné od normálního. V grafu neaditivnosti (Obr.6) je patrný lineární trend a korelační koeficient je příliš vysoký. Proto byla provedena logaritmická transformace.



Obr.7: Q-Q graf pro transformovaná data



Obr.8: Graf neaditivnosti pro transformovaná data

Zkouška transformace:

Korelační koeficient, R: 0.541

Průměry a úrovně efektů:

Celkový průměr = 6.644

Reziduální rozptyl = 0.043

FAKTOR A:			FAKTOR B:		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	6.945	0.302	1	6.680	0.036
2	6.962	0.319	2	6.677	0.033
3	6.602	-0.041	3	6.574	-0.070
4	6.065	-0.579			

Tabulka ANOVA pro model bez interakcí faktorů A, B:

Faktor A: Kvantil $F(1-\alpha, m(n-1), mn(o-1)) = 2.013$

Faktor B: Kvantil $F(1-\alpha, n(m-1), nm(o-1)) = 2.070$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr H_0 je	Spočtená hlad.výz.
Mezi úrovněmi A	n-1= 3	11.122	3.707	67.902	Zamítnuta	0.000
Mezi úrovněmi B	m-1= 2	0.204	0.102	1.868	Akceptována	0.095
Rezidua	mno-m-n+1= 60	3.276	0.055			
Celkový	mno-1= 83	15.602	0.176			

Scheffeho metoda vícenásobného porovnání pro faktor A:

Srovnání	Rozdíl	Směrodat. chyba	q Stat	Tabulka q	P- value	Dolní 95%	Horní 95%	Výsledek H ₀
F - Ah	0.898	0.063	14.195	2.858	0.000	0.717	1.079	zamítnuta
L - Ah	0.881	0.063	13.927	2.858	0.000	0.700	1.062	zamítnuta
H - Ah	0.538	0.063	8.506	2.858	0.000	0.357	0.719	zamítnuta
F - H	0.340	0.063	5.690	2.858	0.000	0.179	0.541	zamítnuta
L - H	0.342	0.063	5.421	2.858	0.000	0.162	0.524	zamítnuta
F - L	0.017	0.063	0.268	2.858	0.995	-0.163	0.198	akceptována

3) ZÁVĚR

Po logaritmické transformaci klesla hodnota korelačního koeficientu z 0.852 na 0.541. V Q-Q grafu (Obr.7) po transformaci rezidua lépe sledují přímku, tedy lépe odpovídají normálnímu rozdělení. V grafu neaditivní (Obr.8) došlo po transformaci k menšímu zeslabení trendu. Byla tedy provedena ANOVA na logaritmicky transformovaných datech. Byl použit model bez interakcí. Z tabulky ANOVA je zřejmé, že nulovou hypotézu pro faktor A je na hladině významnosti 0.05 možno zamítnout, mezi úrovněmi faktoru B nebyl nalezen statisticky signifikantní rozdíl, tedy vliv stanoviště na obsah vápníku není průkazný. Pro faktor A (půdní horizonty) byl v programu UNISTAT 5.6 spočítán Scheffeho test vícenásobného porovnání. Všechny půdní horizonty vykazují statisticky signifikantní rozdíly v obsahu vápníku s výjimkou horizontů L a F.