

LICENČNÍ STUDIUM  
STATISTICKÉ ZPRACOVÁNÍ DAT

**SEMESTRÁLNÍ PRÁCE**  
1.5 ANOVA

Ing. Věra Fialová

BIOPHARM  
VÝZKUMNÝ ÚSTAV BIOFARMACIE A VETERINÁRNÍCH LÉČIV a.s.  
Jílové u Prahy

2005

SEMESTRÁLNÍ PRÁCE  
1.5 ANOVA

## Obsah

<b>Úloha 1. Jednorozměrná analýza</b> .....	<b>3</b>
<b>1. ANOVA bez transformace</b> .....	<b>3</b>
1.1. Podmínky.....	3
1.2. Průměry a efekty.....	3
1.3. Detekce vlivných bodů (Vybočující a odlehlé body) .....	3
1.4. ANOVA.....	4
1.5. Vícenásobné porovnávání (Sheffeho procedura) .....	4
1.6. Zkouška transformace .....	4
<b>2. ANOVA s transformací</b> .....	<b>5</b>
2.1. Podmínky.....	5
2.2. Průměry a efekty.....	5
2.3. Detekce vlivných bodů (Vybočující a odlehlé body) .....	5
2.4. ANOVA.....	5
2.5. Vícenásobné porovnávání (Sheffeho procedura) .....	6
2.6. Zkouška transformace .....	6
<b>3. Závěr</b> .....	<b>6</b>
<b>Úloha 2. Dvourozměrná ANOVA bez opakování</b> .....	<b>7</b>
<b>1. ANOVA bez transformace</b> .....	<b>7</b>
1.1. Podmínky.....	7
1.2. Průměry a efekty.....	7
1.3. ANOVA pro model s Tukeyho iteracemi .....	8
1.4. Zkouška transformace .....	8
<b>3. Závěr</b> .....	<b>8</b>
<b>Úloha 3. Dvourozměrná ANOVA s opakováním</b> .....	<b>9</b>
<b>1. ANOVA bez transformace</b> .....	<b>9</b>
1.1. Podmínky.....	9
1.2. Průměry a efekty.....	9
1.3. ANOVA pro model s iteracemi faktorů A, B.....	10
1.4. Zkouška transformace .....	10
<b>2. ANOVA s transformací</b> .....	<b>11</b>
2.1. Podmínky.....	11
2.2. Průměry a efekty.....	11
2.3. ANOVA pro model s iteracemi faktorů A, B.....	11
2.4. Zkouška transformace .....	12
<b>3. Závěr</b> .....	<b>12</b>

SEMESTRÁLNÍ PRÁCE  
1.5 ANOVA

**Úloha 1. Jednorozměrná analýza**

**Zadání:** Při validaci oxytetracyclinu v prasečí svalovině bylo ve třech následujících dnech měřeno 6 vzorků na stanovení opakovatelnosti. Určete jestli jsou mezi daty v jednotlivých dnech statisticky významné rozdíly. (ADSTAT)(OTC4.txt)

**Data:**

c (ng.mL<sup>-1</sup>)

Den 1	Den 2	Den 3
103,0	96,8	112,8
96,1	98,4	111,1
93,1	102,8	87,9
94,8	107,2	86,6
105,8	93,8	85,1
104,6	105,6	85,4

**1. ANOVA bez transformace**

*1.1. Podmínky*

Hladina významnosti  $\alpha$ : 0.050  
Transformace: Ne  
Počet úrovní faktoru A, k: 3  
Celkový počet n: 18

*1.2. Průměry a efekty*

Byl proveden výpočet celkového průměru (aritmetický) a rozptylu. Dále byly spočteny sloupcové průměry, jejich efekty (velikosti komponenty, kterými tato komponenta přispívá do průměru) a rezidua  $H_{ii}$ .

Celkový průměr 98,4  
Reziduální rozptyl 78,4

Úroveň	Průměr	Efekt	$H_{ii}$
1	99,5	1.15	0,167
2	100,8	2.40	0,167
3	94,8	-3.55	0,167

*1.3. Detekce vlivných bodů (Vybočující a odlehlé body)*

Závěr            Odlehlé            Vybočující  
                  Ano: 2            Žádné

**SEMESTRÁLNÍ PRÁCE**  
**1.5 ANOVA**

*1.4. ANOVA*

Byl proveden F-test významnost efektů.

$H_0$ : Efekty faktoru jsou nulové

$H_A$ : Efekty faktoru nejsou nulové

Kvantil  $F_{1-\alpha}(k-1, n-k)$                       3,682

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr $H_0$ je	Spočtená hlad.výz.
Mezi úrovněmi	$k-1 = 2$	118,1	59,06	0,754	Akceptována	0,488
Rezidua	$n-k = 15$	1175,3	78,36			
Celkově	$n-1 = 17$	1293,4	76,09			

Protože Fischer-Snedecorovo testační kritérium  $F_{exp}$  je 0,754 vyšší než tabulkový kvantil  $F_{1-\alpha}(2,15)$  je 3,682 je nulová hypotéza  $H_0$  přijata a  $H_A$  zamítnuta. Efekty faktoru A nejsou statisticky významné. Mezi jednotlivými dny není statisticky významný rozdíl.

*1.5. Vícenásobné porovnávání (Sheffeho procedura)*

Byly testovány lineární konstanty pro zadané kombinace úrovní  $H_0: \mu_i - \mu_j = 0$ .

Hypotéza $H_0$	Průměrný párový rozdíl	Meze konfidenčního intervalu		Závěr
		dolní	horní	
$P1 = P2$	-1,250	-15,119	12,619	Akceptována
$P1 = P3$	4,700	-9,169	18,569	Akceptována
$P2 = P3$	5,950	-7,919	19,819	Akceptována

Nulová hypotéza byla přijata pro všechny kombinace. Nejsou mezi jednotlivými dny statisticky významné rozdíly.

*1.6. Zkouška transformace*

Korelační koeficient, R:                      -0.985

Korelační koeficient není blízky 0 – transformace je nutná.

**SEMESTRÁLNÍ PRÁCE**  
**1.5 ANOVA**

## 2. ANOVA s transformací

### 2.1. Podmínky

Hladina významnosti $\alpha$ :	0.050
Transformace:	Logaritmická
Počet úrovní faktoru A, k:	3
Celkový počet n:	18

### 2.2. Průměry a efekty

Byl proveden výpočet celkového průměru (aritmetický) a rozptylu hodnot po logaritmické transformaci. Dále byly spočteny sloupcové průměry, jejich efekty (velikosti komponenty, kterými tato komponenta přispívá do průměru) a rezidua  $H_{ii}$ .

Celkový průměr	4,58
Reziduální rozptyl	0,008

Úroveň	Průměr	Efekt	$H_{ii}$
1	4,60	1,41	0,167
2	4,61	2,67	0,167
3	4,54	-4,08	0,167

### 2.3. Detekce vlivných bodů (Vybočující a odlehlé body)

Odlehlé	Vybočující	
Závěr	Ano: 2	Žádné

### 2.4. ANOVA

Byl proveden F-test významnost efektů.

$H_0$ : Efekty faktoru jsou nulové

$H_A$ : Efekty faktoru nejsou nulové

Kvantil  $F(1-\alpha, k-1, n-k)$ : 3,682

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr $H_0$ je	Spočtená hlad.výz.
Mezi úrovněmi	$k-1 = 2$	0,0155	0,0077	0,965	Akceptována	0,404
Rezidua	$n-k = 15$	0,1204	0,0080			
Celkově	$n-1 = 17$	0,1359	0,0080			

Protože Fischer-Snedecorovo testační kritérium  $F_{\text{exp}}$  je 0,965 vyšší než tabulkový kvantil  $F_{1-\alpha}(2,15)$  je 3,682 je nulová hypotéza  $H_0$  přijata a  $H_A$  zamítnuta. Efekty faktoru A nejsou statisticky významné. Mezi jednotlivými dny není statisticky významný rozdíl.

**SEMESTRÁLNÍ PRÁCE**  
**1.5 ANOVA**

*2.5. Vícenásobné porovnávání (Sheffeho procedura)*

Byly testovány lineární konstanty pro zadané kombinace úrovní  $H_0: \mu_i - \mu_j = 0$ .

Hypotéza $H_0$	Průměrný párový rozdíl	Meze konfidenčního intervalu		Závěr
		dolní	horní	
$P1 = P2$	-0,013	-0,153	0,128	Akceptována
$P1 = P3$	0,055	-0,085	0,195	Akceptována
$P2 = P3$	0,068	-0,073	0,208	Akceptována

Nulová hypotéza byla přijata pro všechny kombinace. Nejsou mezi jednotlivými dny statisticky významné rozdíly.

*2.6. Zkouška transformace*

Korelační koeficient, R: -0.990

Korelační koeficient po a před transformací nebyl významně rozdílný – transformace nebyla nutná.

**3. Závěr**

Pomocí analýzy rozptylu bylo zjištěno, že mezi jednotlivými skupinami nejsou významné rozdíly a celkově není rozdíl mezi přípravou vzorků a měřením v jednotlivých dnech.

Při porovnání transformovaných a netransformovaných výsledků, nebyla transformace nutná a v obou případech vyšly srovnatelné výsledky.

SEMESTRÁLNÍ PRÁCE  
1.5 ANOVA

**Úloha 2. Dvourozměrná ANOVA bez opakování**

**Zadání:** Při vývoji metody na stanovení trimethoprimu (TMP) byla měřena koncentrace této látky při třech vlnových délkách a ve třech mobilních fázích s jiným obsahem MeCN. Ověřit zda má na obsah TMP vliv vlnová délka nebo složení mobilní fáze. (ADSTAT)(TMP.txt)

**Data:**

$c$  ( $\mu\text{g}\cdot\text{mL}^{-1}$ )

		Mobilní fáze 1	Mobilní fáze 2	Mobilní fáze 3
Vlnová délka (nm)	261	0,062	0,050	0,048
	271	0,049	0,046	0,042
	281	0,038	0,039	0,038

**1. ANOVA bez transformace**

*1.1. Podmínky*

Hladina významnosti  $\alpha$ : 0.050

Transformace: Ne

Metoda analýzy: Nejmenší čtverce

*1.2. Průměry a efekty*

Byl proveden výpočet celkového průměru (aritmetický) a rozptylu. Dále byly spočteny sloupcové a řádkové průměry, jejich efekty (velikosti komponenty, kterými tato komponenta přispívá do průměru).

Celkový průměr 0,046

Reziduální rozptyl  $2,12 \cdot 10^{-5}$

Faktor A			Faktor B		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	0,053	$7,55 \cdot 10^{-3}$	1	0,050	$3,89 \cdot 10^{-3}$
2	0,046	$-1,11 \cdot 10^{-4}$	2	0,045	$-7,78 \cdot 10^{-4}$
3	0,038	$-7,44 \cdot 10^{-3}$	3	0,043	$-3,11 \cdot 10^{-3}$

Tukeyho C = 140,87

## SEMESTRÁLNÍ PRÁCE 1.5 ANOVA

### 1.3. ANOVA pro model s Tukeyho iteracemi

$H_0$ : Efekty faktoru A jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, n-1, mn-m-n) = 9,552$

$H_0$ : Efekty faktoru B jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, m-1, mn-m-n) = 9,552$

$H_0$ : Iterace I je nulová      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, mn-m-n) = 10,128$

(I znamená efekt Tukeyho interakce.)

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr $H_0$ je	Spočtená hlad.výz.
Mezi úrovněmi A	$n-1 = 2$	$3,38 \cdot 10^{-4}$	$1,69 \cdot 10^{-4}$	7,939	Akceptována	0,063
Mezi úrovněmi B	$m-1 = 2$	$7,62 \cdot 10^{-5}$	$3,81 \cdot 10^{-5}$	1,793	Akceptována	0,307
Iterace	1	$5,67 \cdot 10^{-5}$	$5,67 \cdot 10^{-5}$	2,668	Akceptována	0,201
Rezidua	$mn-m-n = 3$	$6,38 \cdot 10^{-5}$	$2,13 \cdot 10^{-5}$			
Celkově	$mn-1 = 8$	$4,78 \cdot 10^{-4}$	$5,97 \cdot 10^{-5}$			

Protože testační kritérium  $F_{\text{exp}}$  je 7,939 menší než tabulkový kvantil  $F_{1-\alpha}(1,3)$  je 9,552 je nulová hypotéza  $H_0$  přijata a faktor A je statisticky nevýznamný.

Protože testační kritérium  $F_{\text{exp}}$  je 1,793 menší než tabulkový kvantil  $F_{1-\alpha}(1,3)$  je 9,552 je nulová hypotéza  $H_0$  přijata a faktor B je statisticky nevýznamný.

Protože testační kritérium  $F_{\text{exp}}$  je 2,668 menší než tabulkový kvantil  $F_{1-\alpha}(1,3)$  je 10,128 je nulová hypotéza  $H_0$  přijata a iterace faktorů A a B je statisticky nevýznamný.

Na koncentraci látky nemá statisticky významný vliv ani vlnová délka ani složení mobilní fáze.

### 1.4. Zkouška transformace

Odhad mocninné transformace: -5,449

Rozptyl odhadu transformace: 11,688

Akceptovatelný interval: (-8,867; -2,030)

Odhad mocninné transformace leží v akceptovatelném intervalu, proto není třeba data transformovat.

## 3. Závěr

Pomocí analýzy rozptylu bylo zjištěno, že na koncentraci TMP nemá statisticky významný vliv ani vlnová délka ani složení mobilní fáze.



SEMESTRÁLNÍ PRÁCE  
1.5 ANOVA

**Úloha 3. Dvourozměrná ANOVA s opakováním**

**Zadání:** V rámci bioekvivalenční studie na hovězím dobytku byl stanovován 4-methylaminoantipyrine (4-MAA) ve dvou přípravcích se stejným obsahem 4-MAA, které byly podávány ve dvou periodách (I, II) dvěma skupinám zvířat. Otestujte parametr plochy pod křivkou zda je závislý na periodě nebo na podávaném přípravku (ADSTAT)(4\_MAA.txt)

**Data:**

AUC<sub>n</sub> (μg.hr.mL<sup>-1</sup>)

	Perioda I	Perioda II
Testovaný přípravek TI	430,4	244,6
	315,1	177,9
	275,8	201,9
	234,1	149,8
	206,1	106,5
	211,5	174,3
Referenční přípravek RI	397,2	161,4
	344,2	129,9
	312,3	105,7
	415,5	184,8
	306,6	123,4
	268,4	151,1

**1. ANOVA bez transformace**

*1.1. Podmínky*

Hladina významnosti  $\alpha$ : 0.050  
 Transformace: Ne  
 Počet úrovní parametru A, n: 2  
 Počet úrovní parametru B, m: 2  
 Počet opakování v jedné buňce, o: 6

*1.2. Průměry a efekty*

Byl proveden výpočet celkového průměru (aritmetický) a rozptylu. Dále byly spočteny sloupcové a řádkové průměry, jejich efekty (velikosti komponenty, kterými tato komponenta přispívá do průměru).

Celkový průměr 234,5  
 Reziduální rozptyl 3355,8

Faktor A			Faktor B		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	227,3	-7,19	1	309,8	7,52
2	241,7	7,19	2	159,3	-7,52

**SEMESTRÁLNÍ PRÁCE**  
**1.5 ANOVA**

*1.3. ANOVA pro model s iteracemi faktorů A, B*

Byl použit model s iteracemi, protože by mohl být vztah mezi podávaným přípravkem a periodě ve které byl podán.

$H_0$ : Efekty faktoru A jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, n-1, mn(o-1)) = 4,351$

$H_0$ : Efekty faktoru B jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, m-1, mn(o-1)) = 4,351$

$H_0$ : Iterace I je nulová      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, (n-1)(m-1), mn(o-1)) = 4,351$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr $H_0$ je	Spočtená hlad.výz.
Mezi úrovněmi A	$n-1 = 1$	1239,8	1239,8	0,369	Akceptována	0,550
Mezi úrovněmi B	$m-1 = 1$	135890	135890	40,493	Zamítnuta	0,000
Iterace	$(n-1)(m-1) = 1$	13533	13533	4,033	Akceptována	0,058
Rezidua	$mn(o-1) = 20$	67116	3355,8			
Celkově	$mno-1 = 23$	217770	9468,5			

Protože testační kritérium  $F_{exp}$  je 0,369 menší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  přijata a faktor A je statisticky nevýznamný.

Protože testační kritérium  $F_{exp}$  je 40,493 větší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  zamítnuta a faktor B je statisticky významný.

Protože testační kritérium  $F_{exp}$  je 4,033 menší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  přijata a iterace faktorů A a B je statisticky nevýznamný.

Mezi přípravky není statisticky významný rozdíl, mezi periodami je rozdíl a mezi skupinami není žádný vztah.

*1.4. Zkouška transformace*

Korelační koeficient, R:     0.6815

Korelační koeficient není blízký 0 – transformace je nutná.

**SEMESTRÁLNÍ PRÁCE**  
1.5 ANOVA

**2. ANOVA s transformací**

2.1. Podmínky

Hladina významnosti $\alpha$ :	0.050
Transformace:	Logaritmická
Počet úrovní parametru A, n:	2
Počet úrovní parametru B, m:	2
Počet opakování v jedné buňce, o:	6

2.2. Průměry a efekty

Byl proveden výpočet celkového průměru (aritmetický) a rozptylu. Dále byly spočteny sloupcové a řádkové průměry, jejich efekty (velikosti komponenty, kterými tato komponenta přispívá do průměru).

Celkový průměr	5,374
Reziduální rozptyl	0,057

Faktor A			Faktor B		
Úroveň	Průměr	Efekt	Úroveň	Průměr	Efekt
1	5,367	-0,0074	1	5,708	0,333
2	5,382	0,0074	2	5,041	-0,333

2.3. ANOVA pro model s iteracemi faktorů A, B

Byl použit model s iteracemi, protože by mohl být vztah mezi podávaným přípravkem a periodě ve které byl podán.

$H_0$ : Efekty faktoru A jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, n-1, mn(o-1)) = 4,351$

$H_0$ : Efekty faktoru B jsou nulové      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, m-1, mn(o-1)) = 4,351$

$H_0$ : Iterace I je nulová      $H_A$ : ... nejsou nulové

Kvantil  $F(1-\alpha, (n-1)(m-1), mn(o-1)) = 4,351$

Zdroj rozptylu	Stupně volnosti	Součet čtverců	Průměrný čtverec	Testovací kritérium	Závěr $H_0$ je	Spočtená hlad.výz.
Mezi úrovněmi A	$n-1 = 1$	0,0013	0,0013	0,023	Akceptována	0,880
Mezi úrovněmi B	$m-1 = 1$	2,668	2,668	46,997	Zamítnuta	0,000
Iterace	$(n-1)(m-1) = 1$	0,261	0,261	4,604	Zamítnuta	0,044
Rezidua	$mn(o-1) = 20$	67116	3355,8			
Celkově	$mno-1 = 23$	217770	9468,5			

Protože testační kritérium  $F_{exp}$  je 0,023 menší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  přijata a faktor A je statisticky nevýznamný.

## SEMESTRÁLNÍ PRÁCE

### 1.5 ANOVA

Protože testační kritérium  $F_{\text{exp}}$  je 46,997 větší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  zamítnuta a faktor B je statisticky významný.

Protože testační kritérium  $F_{\text{exp}}$  je 4,604 menší než tabulkový kvantil  $F_{1-\alpha}(1,20)$  je 4,351 je nulová hypotéza  $H_0$  přijata a iterace faktorů A a B je statisticky významný.

Mezi přípravky není statisticky významný rozdíl, mezi periodami je rozdíl a mezi skupinami je pravděpodobně vztah.

#### 2.4. Zkouška transformace

Korelační koeficient, R: -0.248

Korelační koeficient po a před transformací byl významně rozdílný – transformace byla nutná.

### 3. Závěr

Pomocí analýzy rozptylu bylo zjištěno, že na koncentraci 4-MAA v hovězí plazmě nemá vliv podávaný přípravek, ale statisticky významný vliv má perioda ve které byly přípravky podávány. Také byl prokázán vztah mezi podávaným přípravkem a periodou.