

Vaje za drugi delni izpit

1. V prašičereji spremljamo število živorojenih pujskov na gnezdo (output 1). Ustvariti želimo ustrezen mešani model, ki ga bomo lahko uporabljali pri napovedovanju plemenske vrednosti za žival. Naša prva naloga je poiskati najprimernejši sistematski del modela. Oznake lastnosti in vplivov so naslednje: liveborn - število živorojenih pujskov, breed- pasma in parity - dolžina laktacije.

- Koliko pasem je bilo vključenih analizo?
- Koliko meritev je bilo vključenih v analizo?
- Stopinje prostosti za model in ostanek (razvidno naj bo, koliko stopinj prostosti porabimo za posamezni vpliv)?
- Delež pojasnjene variance (izračunajt sami!)?
- Varianca za ostanek in standardni odklon za ostanek?
- Napišite model v skalarni obliki (ne pozabite obrazložiti uporabljenih črk)!
- Izpišite neznane in ocenljive parametre v modelu?

Output 1

```
Class          Levels  Values
  breed          5    11 12 21 22 33
Number of observations    4410
NOTE: Due to missing values, only 3279 observations can be used in this analysis.
      The SAS System                               10:45 Wednesday, November 19, 2008 20

                                The GLM Procedure

Dependent Variable: liveborn

Source          DF          Sum of Squares    Mean Square    F Value    Pr > F
Model           5          549.67817        109.93563      15.54    <.0001
Error          3273          23152.75183         7.07386
Corrected Total 3278          23702.43001

R-Square          0.023191
Coeff Var         27.44546
Root MSE          2.659673
liveborn Mean     9.690759

Source          DF          Type I SS    Mean Square    F Value    Pr > F
breed           4          366.4598789     91.6149697     12.95    <.0001
parity          1          183.2182960     183.2182960     25.90    <.0001
Source          DF          Type III SS   Mean Square    F Value    Pr > F
breed           4          355.4465228     88.8616307     12.56    <.0001
parity          1          183.2182960     183.2182960     25.90    <.0001
```

2. V prašičereji spremljamo število odstavljenih pujskov na gnezdo (output 2). Ustvariti želimo ustrezen mešani model, ki ga bomo lahko uporabljali pri napovedovanju plemenske vrednosti za žival. Naša prva naloga je poiskati najprimernejši sistematski del modela. Oznake lastnosti in vplivov so naslednje: p_weaned1 - število odstavljenih pujskov, genotip- pasma, zap_pras - zaporedna prasitev in plact - dolžina predhodne laktacije.

- Koliko pasem je bilo vključenih analizo?
- Koliko meritev je bilo vključenih v analizo?
- Stopinje prostosti za model in ostanek (razvidno naj bo, koliko stopinj prostosti porabimo za posamezni vpliv)?
- Delež pojasnjene variance (izračunajt sami!)?
- Varianca za ostanek?
- Napišite model v skalarni obliki (ne pozabite obrazložiti uporabljenih črk)!
- Izpišite neznane in ocenljive parametre v modelu?

Output 2

```

Class Level Information
Class      Levels  Values
genotip    6      11 12 15 22 44 55
Number of observations    11964
NOTE: Due to missing values, only 7082 observations can be used in this analysis.

                The SAS System                               19:01 Thursday, November 20, 2008    6
                The GLM Procedure
Dependent Variable: p_weaned1

Source          DF          Sum of Squares    Mean Square    F Value    Pr > F
Model           7          2988.11725      426.87389     49.18     <.0001
Error          7074        61405.42582       8.68044
Corrected Total 7081        64393.54307

R-Square      Coeff Var    Root MSE    p_weaned1 Mean
0.046404      36.14809     2.946258     8.150522

Source          DF          Type I SS    Mean Square    F Value    Pr > F
genotip         5          162.648701    32.529740     3.75     0.0022
zap_prasitev    1          2277.641542   2277.641542   262.39     <.0001
plact           1          547.827004    547.827004    63.11     <.0001
Source          DF          Type III SS   Mean Square    F Value    Pr > F
genotip         5          99.063446     19.812689     2.28     0.0439
zap_prasitev    1          2118.990003   2118.990003   244.11     <.0001
plact           1          547.827004    547.827004    63.11     <.0001

```

3. V prašičereji spremljamo debelino hrbtne slanine (output 3). Ustvariti želimo ustrezen mešani model, ki ga bomo lahko uporabljali pri napovedovanju plemenske vrednosti za žival. Naša prva naloga je poiskati najprimernejši sistematski del modela. Oznake lastnosti in vplivov so naslednje: dhs_100 - debelina hrbtne slanine pri 100 kg, farma- farma, pasma - pasma in masa_100 - dejanska telesna masa masa ob meritvi debeline hrbtne slanine.

- Koliko pasem je bilo vključenih analizo?
- Koliko farm je bilo vključenih v nalizo?
- Koliko meritev je bilo vključenih v analizo?
- Stopinje prostosti za model in ostanek (razvidno naj bo, koliko stopinj prostosti porabimo za posamezni vpliv)?
- Delež pojasnjene variance (izračunajte sami!)?
- Varianca za ostanek?
- Napišite model v skalarni obliki (ne pozabite obrazložiti uporabljenih črk)!
- Ali je model primeren? Utemeljite!
- Izpišite neznanne in ocenljive parametre v modelu?
- Ocena regresijskega koeficienta s standardno napako za dejansko telesno maso?
- Izpišite povprečje za debelino hrbtne slanine (s standardno napako) za pasmo 55?
- Izpišite povprečje za debelino hrbtne slanine (s standardno napako) za farmo 1?

Output 3

The SAS System 08:08 Thursday, February 5, 2009 14
The GLM Procedure

Class Level Information

Class	Levels	Values
farma	2	1 4
pasma	3	11 22 55

Number of observations 5529

NOTE: Due to missing values, only 2996 observations can be used in this analysis.

The SAS System 08:08 Thursday, February 5, 2009 15
The GLM Procedure

Dependent Variable: dhs_100

	DF	Sum of Squares	Source Mean Square	F Value	Pr > F
Model	4	1528.43974	382.10993	109.97	<.0001
Error	2991	10392.38369	3.47455		
Corrected Total	2995	11920.82343			

R-Square	Coeff Var	Root MSE	dhs_100 Mean
0.128216	16.93377	1.864015	11.00768

Source	DF	Type I SS	Mean Square	F Value	Pr > F
farma	1	438.1831348	438.1831348	126.11	<.0001
pasma	2	930.7026429	465.3513215	133.93	<.0001
masa_100	1	159.5539596	159.5539596	45.92	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
farma	1	0.8620436	0.8620436	0.25	0.6185
pasma	2	970.0381060	485.0190530	139.59	<.0001
masa_100	1	159.5539596	159.5539596	45.92	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	3.114532153 B	1.37498745	2.27	0.0236
farma 1	-0.039774403 B	0.07985249	-0.50	0.6185
farma 4	0.000000000 B	.	.	.
pasma 11	-1.743691942 B	0.10557495	-16.52	<.0001
pasma 22	-1.496154617 B	0.11075642	-13.51	<.0001
pasma 55	0.000000000 B	.	.	.
masa_100	0.093199026	0.01375330	6.78	<.0001

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The SAS System 08:08 Thursday, February 5, 2009 16
The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey-Kramer

	dhs_100 LSMEAN	Standard Error	H0:LSMEAN=0 Pr > t	H0:LSMean1=LSMean2 Pr > t
farma				
1	11.2710392	0.0624716	<.0001	0.6185
4	11.3108136	0.0485445	<.0001	

The SAS System 08:08 Thursday, February 5, 2009 17
The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey-Kramer

	dhs_100 LSMEAN	Standard Error	Pr > t	LSMEAN Number
pasma				
11	10.6271833	0.0471482	<.0001	1
22	10.8747207	0.0672618	<.0001	2
55	12.3708753	0.0890685	<.0001	3

Least Squares Means for effect pasma

Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: dhs_100

i/j	1	2	3
1		0.0077	<.0001
2	0.0077		<.0001
3	<.0001	<.0001	

4. V prašičereji spremljamo debelino hrbtnne slanine (output 4). Ustvariti želimo ustrezen mešani model, ki ga bomo lahko uporabljali pri napovedovanju plemenske vrednosti za žival. Naša prva naloga je poiskati najprimernejši sistematski del modela.

- Koliko pasem je bilo vključenih analizo?
- Koliko farm je bilo vključenih v nalizo?
- Koliko sezon je bilo vključenih v nalizo?
- Koliko meritev je bilo vključenih v analizo?
- Stopinje prostosti za model in ostanek (razvidno naj bo, koliko stopinj prostosti porabimo za posamezni vpliv)?
- Delež pojasnjene variance (izračunajte sami!)?
- Varianca in standardni odklon za ostanek?
- Napišite model v skalarni obliki (ne pozabite obrazložiti uporabljenih črk)!
- Izpišite neznane in ocenljive parametre v modelu?
- Ocena regresijskega koeficienta s standardno napako za maso_100?
- Izpišite povprečje za debelino hrbtnne slanine (s standardno napako) za pasmo 22?
- Izpišite povprečje za debelino hrbtnne slanine (s standardno napako) za farmo 4?

Output 4

The SAS System 08:08 Thursday, February 5, 2009 5
The GLM Procedure

Class Level Information

Class	Levels	Values
sezona	34	200101 200102 200103 200104 200105 200106 200107 200108 200109 200110
		200111 200112 200201 200202 200203 200204 200205 200206 200207 200208
		200209 200210 200211 200212 200301 200302 200303 200304 200305 200306
		200307 200308 200309 200310
		200311 200312 200313 200314 200315 200316 200317 200318 200319 200320
		200321 200322 200323 200324 200325 200326 200327 200328 200329 200330
		200331 200332 200333 200334 200335 200336 200337 200338 200339 200340
		200341 200342 200343 200344 200345 200346 200347 200348 200349 200350
		200351 200352 200353 200354 200355 200356 200357 200358 200359 200360
		200361 200362 200363 200364 200365 200366 200367 200368 200369 200370
farma	2	1 4
pasma	3	11 22 55

Number of observations 5529
NOTE: Due to missing values, only 2996 observations can be used in this analysis.

The SAS System 08:08 Thursday, February 5, 2009 6
The GLM Procedure

Dependent Variable: dhs_100

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	37	3082.86294	83.32062	27.89	<.0001
Error	2958	8837.96049	2.98782		
Corrected Total	2995	11920.82343			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sezona	33	1353.639115	41.019367	13.73	<.0001
farma	1	684.613154	684.613154	229.13	<.0001
pasma	2	949.108039	474.554019	158.83	<.0001
masa_100	1	95.502629	95.502629	31.96	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sezona	33	1554.423199	47.103733	15.77	<.0001
farma	1	40.264372	40.264372	13.48	0.0002
pasma	2	977.225089	488.612545	163.54	<.0001
masa_100	1	95.502629	95.502629	31.96	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	5.010314786 B	1.36916216	3.66	0.0003
sezona 200101	1.158234113 B	0.54783957	2.11	0.0346
sezona 200102	1.147362050 B	0.55410027	2.07	0.0385
sezona 200103	1.235006053 B	0.54973852	2.25	0.0247
sezona 200104	1.597650196 B	0.55093374	2.90	0.0038
sezona 200105	1.265372096 B	0.55644705	2.27	0.0230
sezona 200106	0.240099111 B	0.55859928	0.43	0.6674
sezona 200107	0.041854946 B	0.55566094	0.08	0.9400
sezona 200108	0.391111416 B	0.55114865	0.71	0.4780
sezona 200109	0.639232839 B	0.55792337	1.15	0.2520
sezona 200110	0.999485751 B	0.54547038	1.83	0.0670
sezona 200111	1.159096501 B	0.55695902	2.08	0.0375
sezona 200112	1.084313105 B	0.54843720	1.98	0.0481
sezona 200201	1.155492243 B	0.54992873	2.10	0.0357
sezona 200202	0.727069970 B	0.56119475	1.30	0.1952
sezona 200203	0.564140172 B	0.55400991	1.02	0.3086
sezona 200204	0.663365179 B	0.55222166	1.20	0.2297
sezona 200205	0.148720166 B	0.55623421	0.27	0.7892

The SAS System 08:08 Thursday, February 5, 2009 7
The GLM Procedure

Dependent Variable: dhs_100

Parameter	Estimate	Standard Error	t Value	Pr > t
sezona 200206	0.078367107 B	0.55106750	0.14	0.8869
sezona 200207	-0.378457637 B	0.55330308	-0.68	0.4940
sezona 200208	-0.262050816 B	0.55430109	-0.47	0.6364
sezona 200209	-0.010093005 B	0.55034007	-0.02	0.9854
sezona 200210	0.270108817 B	0.56648164	0.48	0.6335
sezona 200211	-0.261032756 B	0.55365538	-0.47	0.6373
sezona 200212	-0.390647538 B	0.55456937	-0.70	0.4812
sezona 200301	-0.232873564 B	0.54715762	-0.43	0.6704
sezona 200302	-0.453810214 B	0.55486023	-0.82	0.4135
sezona 200303	-0.505352865 B	0.55600129	-0.91	0.3635
sezona 200304	-0.925885037 B	0.56195818	-1.65	0.0995
sezona 200305	-0.536410873 B	0.54928389	-0.98	0.3289
sezona 200306	-0.609164843 B	0.56310592	-1.08	0.2794
sezona 200307	-1.159250624 B	0.56466589	-2.05	0.0402
sezona 200308	-0.842546852 B	0.55708390	-1.51	0.1305
sezona 200309	-0.250004952 B	0.55368640	-0.45	0.6516
sezona 200310	0.000000000 B	.	.	.
farma 1	-0.277823759 B	0.07568083	-3.67	0.0002
farma 4	0.000000000 B	.	.	.

asma	11	-1.742650604 B	0.09871720	-17.65	<.0001
asma	22	-1.582319509 B	0.10358986	-15.27	<.0001
asma	55	0.000000000 B	.	.	.
masa_100		0.072787745	0.01287442	5.65	<.0001

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The SAS System 08:08 Thursday, February 5, 2009 8
The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey-Kramer

asma	dhs_100	Standard	HO:LSMEAN=0		HO:LSMean1=LSMean2	
			Pr > t	Pr > t	Pr > t	Pr > t
1	11.0967316	0.0600410	<.0001			0.0002
4	11.3745553	0.0480125	<.0001			

The SAS System 08:08 Thursday, February 5, 2009 9
The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey-Kramer

asma	dhs_100	Standard	Pr > t	LSMEAN	Number
11	10.6013162	0.0464906	<.0001		1
22	10.7616473	0.0640792	<.0001		2
55	12.3439668	0.0844111	<.0001		3

Least Squares Means for effect asma

Pr > |t| for HO: LSMean(i)=LSMean(j)

Dependent Variable: dhs_100

i/j	1	2	3
1		0.0961	<.0001
2	0.0961		<.0001
3	<.0001	<.0001	