Biostats VHM 802 Course
Winter 2011, Atlantic Veterinary College, PEI
Henrik Stryhn

Exercises for Sessions 10–11: Random effects and repeated measures models

Exercise 10.1
Random effects models — modelling and construction of ANOVA tables
Exercises 10.2–4 below present three data sets and experimental designs that invite statistical models with random effect(s). Propose statistical models for each of these problems. Reflect on which effects should be fixed and which should be random; if the description of the context of the data seems insufficient, make appropriate assumptions. Sketch corresponding ANOVA tables by giving for all effects (sources of variation) in the model the degrees of freedom and the formula for an F-statistic for testing the hypothesis of no influence of the effect. The statistical analyses — using Minitab/Stata — are postponed to the corresponding exercises.

Exercise 10.2
Analysis of random effects models I
In an investigation of the moisture content in cheese, 3 lots (of cheese) were selected among a large number of lots, and in each of the lots the moisture content was measured for two randomly selected cheeses. (Data from Marcuse, M. (1949), Optimum allocation and variance components in nested sampling with an application to chemical analysis, *Biometrics* **5**, 189–206.)

Lot	Moisture content	
1	38.90	38.98
2	35.58	35.55
3	36.51	35.87

Analyse the data. Estimate all the parameters of the model, and give 95% confidence interval for mean parameter(s). Compare the results with a usual one-way ANOVA.

Exercise 10.3
Analysis of random effects models II
In a study of repeatability and reproducibility of laboratory tests the following measurements were made in 11 laboratories on 6 samples with 3 different materials (presumably levels of some chemical component). (Data from Youden, W. J. & Steiner, E. H. (1975), *Statistical manual of the association of official analytical chemists*, AOAC, Washington.)

	Material					
Laboratory	1		2		3	
1	21.2	21.4	12.7	12.9	16.0	16.0
2	21.4	21.6	13.2	13.0	16.1	15.8
3	20.8	20.7	13.1	12.8	16.3	16.0
4	21.9	21.6	13.5	13.1	17.1	16.8
5	21.0	20.9	12.9	13.0	16.5	16.4
6	20.9	20.4	12.8	12.7	16.5	16.2
7	21.2	20.9	12.8	12.7	16.7	16.7
8	22.0	21.1	13.0	12.9	16.6	16.9
9	20.7	21.0	12.6	12.9	16.3	16.5
10	20.9	21.3	12.1	12.8	16.5	16.7
11	21.1	20.6	13.0	12.8	16.5	16.2

Analyse the data. In particular, estimate the variance components as well as the repeatability and the reproducibility.

Exercise 10.4

Analysis of random effects models III

An experiment was conducted to clarify whether a breeding of two strains of mice to have low and high blood-pH values, respectively, had been successful. For each of the two strains, denoted here by pHL and pHH (in obvious notation), seven litters were selected, and within each litter the blood-pH value was recorded for four male mice. (Data from Sokal & Rohlf (1995): *Biometry*.)

Strain	Litter (code)	blood-pH values			
pHH	387	7.43	7.38	7.49	7.49
	388	7.39	7.46	7.50	7.55
	389	7.53	7.50	7.63	7.47
	401	7.39	7.39	7.44	7.55
	402	7.48	7.43	7.47	7.44
	404	7.43	7.55	7.44	7.50
	405	7.49	7.49	7.51	7.54
pHL	392	7.40	7.46	7.43	7.42
	408	7.35	7.40	7.46	7.38
	413	7.51	7.39	7.42	7.43
	414	7.46	7.53	7.49	7.45
	415	7.48	7.53	7.52	7.43
	434	7.43	7.40	7.48	7.47
	446	7.53	7.47	7.50	7.53

Analyse the data. Estimate the parameters of the model. Has the breeding succeeded in producing strains with markedly low and high blood-pH values, respectively?

Exercise 10.5

Split-plot type experimental design

In a study of certain conditions (I–IV) related to the watering of plants a number of hibiscus cuttings of two varieties (R=red rosa-sinensis and H=holiday) were planted in troughs and exposed to these conditions. A total of 16 troughs were used. Each trough contained cuttings of both varieties, and the whole trough had the same watering. Thus, there were four troughs with each watering condition. After 28 days the root weights of the plants were measured and summed to totals for each variety within troughs. The values are given in the table below. (Data from Skovgaard, I. (1994): *Statistisk Forsøgsplanlægning* (in Danish), DSR, KVL.)

Watering	Variety	Troughs			
I	H	100	153	80	145
	R	46	119	103	105
II	H	145	208	299	239
	R	87	220	219	168
III	H	291	236	234	215
	R	213	111	210	82
IV	H	190	193	282	207
	R	20	106	205	149

Analyse the data. If you have problems in identifying the experimental design, try making a draft of the experimental layout. Estimate the variance parameters of the model. For the significant (or close to significant) factors, think about how to present the results with appropriate estimated levels and corresponding standard errors. Calculate these as well.

Exercise 10.6

Unbalanced random effects model

The efficiency of three machines was examined in a study where six employees operated each machine in a certain period. For all working periods an overall rating was computed that took into account the number and quality of the components produced. Each employee operated each machine for either one, two, or three times in the study. Ratings are tabelled below. (Data from Milliken & Johnson (1984): *Analysis of Messy Data, Volume I: Designed Experiments.*)

Machine	Employee	Rating		
1	1	52.0		
	2	51.8	52.8	
	3	60.0		
	4	51.1	52.3	
	5	50.9	51.8	51.4
	6	46.4	44.8	49.2
2	1	64.0		
	2	59.7	60.0	59.0
	3	68.6	65.8	
	4	63.2	62.8	62.2
	5	64.8	65.0	
	6	43.7	44.2	43.0
3	1	67.5	67.2	66.9
	2	61.5	61.7	62.3
	3	70.8	70.6	71.0
	4	64.1	66.2	64.0
	5	72.1	72.0	71.1
	6	62.0	61.4	60.5

Analyse the data. In particular, estimate all variance components in the model (reflect on whether the employees should be taken as random or fixed effects). How can you determine which (if any) of the three machines is best, and worst?

Exercise 10.7

Analysis of longitudinal data I

In a study of the physiological effects of alcohol consumption two groups of moderately and strongly addicted persons had their salsolinol excretion measured during four consecutive days. Salsolinol is an alkaloid with similar structure as heroine. (Data from Hand & Taylor (1987): *Multivariate Analysis and Repeated Measures.*)

Salsolinol ($m$mol)		Day			
Group	Person	1	2	3	4
1	1	0.33	0.70	2.33	3.20
	2	5.30	0.90	1.80	0.70
	3	2.50	2.10	1.12	1.01
	4	0.98	0.32	3.91	0.66
	5	0.39	0.69	0.73	2.45
	6	0.31	6.34	0.63	3.86
2	7	0.64	0.70	1.00	1.40
	8	0.73	1.85	3.60	2.60
	9	0.70	4.20	7.30	5.40
	10	0.40	1.60	1.40	7.10
	11	2.60	1.30	0.70	0.70
	12	7.80	1.20	2.60	1.80
	13	1.90	1.30	4.40	2.80
	14	0.50	0.40	1.10	8.10

Explain why these data are longitudinal. Analyse the data along similar lines as the guinea pig growth data from the notes. Summarise the results of different analyses and conclude about on the influence of alcohol consumption on salsolinol excretion.

Exercise 10.8

Analysis of longitudinal data II

A study investigated the effect of the drugs morphine and trimethaphan on histamine release and hypotension in dogs. A total of 16 mangrel dogs were divided into four treatment groups of four dogs each. The dogs in the first two groups received intraveous morphine sulfate, and the remaining dogs received intravenous trimethaphan. In addition, dogs in the second and fourth groups had been treated so that their supplies of available histamine were depleted at the time of innoculation with the treatment drugs. The blood histamine levels were measured immediately prior to, and at intervals of one, three, and five minutes after administration of the treatment drugs, and the results are given in the table below. (Data from Morris, K. J. & Zeppa, R. (1963), Histamine-induced hypotension due to morphine and Arfonad in the dog, *J. Surg. Res.* **3**, 313–317.)

Treatment drug	Histamin levels at onset	Time of measurement			
		0	1	3	5
morphine	intact	0.04	0.20	0.10	0.08
morphine	intact	0.02	0.06	0.02	0.02
morphine	intact	0.07	1.40	0.48	0.24
morphine	intact	0.17	0.57	0.35	0.24
morphine	depleted	0.10	0.09	0.13	0.14
morphine	depleted	0.12	0.11	0.10	.
morphine	depleted	0.07	0.07	0.07	0.07
morphine	depleted	0.05	0.07	0.06	0.07
trimethaphan	intact	0.03	0.62	0.31	0.22
trimethaphan	intact	0.03	1.05	0.73	0.60
trimethaphan	intact	0.07	0.83	1.07	0.80
trimethaphan	intact	0.09	3.13	2.06	1.23
trimethaphan	depleted	0.10	0.09	0.09	0.08
trimethaphan	depleted	0.08	0.09	0.09	0.10
trimethaphan	depleted	0.13	0.10	0.12	0.12
trimethaphan	depleted	0.06	0.05	0.05	0.05

Describe the experimental design, in particular reflect on what type of analysis would be appropriate for single values for each dog. Then analyse the data using one or more of the methods for longitudinal data.

Exercise 10.9

Analysis of longitudinal data III

Visual acuity was measured as the response times of the eye to a stimulus, more precisely the time lag (in milliseconds) between a light flash and the electrical response at the back of the cortex. Seven volunteers saw the light, and recordings were made for left and right eyes. Of particular interest was the influence of lenses of powers 6/6, 6/18, 6/36 and 6/60 put before the eye. (The power 6/18 corresponds to a magnification such that the eye will perceive an object at 18 feet distance as being at 6 feet.) (Data from Crowder & Hand (1990): *Analysis of Repeated Measures.*)

	Left eye				Right eye			
Response time	Power				Power			
Person	6/6	6/18	6/36	6/60	6/6	6/18	6/36	6/60
1	116	119	116	124	120	117	114	122
2	110	110	114	115	106	112	110	110
3	117	118	120	120	120	120	120	124
4	112	116	115	113	115	116	116	119
5	113	114	114	118	114	117	116	112
6	119	115	94	116	100	99	94	97
7	110	110	105	118	105	105	115	115

Reflect also for these data on how the repeated measures aspect enters into the design. Analyse the data using a random effects model, and discuss the model assumptions made.

Exercise 10.10

Analysis of longitudinal data IV

A pilot trial was carried out to compare 2 drug treatments (tablets). Blood samples were taken from 5 (volunteer) subjects, and the outcome measured was the antibiotic serum level. All subjects were subjected to both drugs, in 2 periods with a wash-out period in between. In each of the periods, measurements were taken at 1, 2, 3 and 6 hours after medication. The observations are tabled below. (Data from Crowder & Hand (1990): *Analysis of Repeated Measures.*)

	Drug A				Drug B			
	Time (hours)				Time (hours)			
Subject	1	2	3	6	1	2	3	6
1	1.08	1.99	1.46	1.21	1.48	2.50	2.62	1.95
2	1.19	2.10	1.21	0.96	0.62	0.88	0.68	0.48
3	1.22	1.91	1.36	0.90	0.65	1.52	1.32	0.95
4	0.60	1.10	1.03	0.61	0.32	2.12	1.48	1.09
5	0.55	1.00	0.82	0.52	1.48	0.90	0.75	0.44

Describe the experimental design, and reflect on the effects of interest. What is the appropriate experimental unit for the drug treatment? Give also the design for a single measurement over time in both of the periods (e.g., 1 hour after medication). Analyse the data with the purpose of exploring the hypotheses of interest.

Exercise 10.11

Analysis of longitudinal data V

A diet (dietary regime treatment) was applied to 12 hospital patients, and the plasma ascorbic acid was measured at weeks 1, 2, 6, 10, 14, 15 and 16. The unequal spacing of the measurements corresponds to a division of the study period into 3 stages: pre-diet (weeks 1 and 2), during diet (weeks 6, 10 and 14) and post-diet (weeks 15 and 16). The observations are given in the table below. (Data from Crowder & Hand (1990): *Analysis of Repeated Measures.*)

	Week						
Patient	1	2	6	10	14	15	16
1	0.22	0.00	1.03	0.67	0.75	0.65	0.59
2	0.18	0.00	0.96	0.96	0.98	1.03	0.70
3	0.73	0.37	1.18	0.76	1.07	0.80	1.10
4	0.30	0.25	0.74	1.10	1.48	0.39	0.36
5	0.54	0.42	1.33	1.32	1.30	0.74	0.56
6	0.16	0.30	1.27	1.06	1.39	0.63	0.40
7	0.30	1.09	1.17	0.90	1.17	0.75	0.88
8	0.70	1.30	1.80	1.80	1.60	1.23	0.41
9	0.31	0.54	1.24	0.56	0.77	0.28	0.40
10	1.40	1.40	1.64	1.28	1.12	0.66	0.77
11	0.60	0.80	1.02	1.28	1.16	1.01	0.67
12	0.73	0.50	1.08	1.26	1.17	0.91	0.87

As in the previous exercise, describe the experimental design and reflect on the effects and comparisons of interest. Analyse the data with the purpose of quantifying the diet's impact on the levels of plasma ascorbic acid.