Department of Statistics

COURSE STATS 330/762

Assignment 1, 2010

Instructions: Hand in your completed assignment to the Student Resource Centre by **4pm August 5th**

The data sets for this assignment are in the files **kent.txt** (for Question 1) and **Sitka.csv** (for Question 2) which are available on the course web page.

Question 1.

The data in the file **kent.txt** relate to a series of 93 crash tests. In each test, a cadaver is subjected to an impact meant to mimic the effect of an automobile crash. A major factor determining the degree of injury in a car crash is the "chest displacement" (the maximum displacement of the sternum relative to the spine). A major aim of these tests is to relate this, (and other factors such as gender and age) to the actual degree of injury, which is measured by the number of rib fractures.

The tests were carried out in four different ways, mimicking different crash conditions such as whether a seat belt was worn, whether an air bag is fitted and so on. The four conditions considered are

Blunt hub loading: This mimics the impact of the steering wheel in an unrestrained crash.

Seatbelt loading: This mimics the effect of a crash in which a seatbelt is worn, but no airbag is fitted.

Distributed loading: This mimics the effect of a crash with an airbag but no seat belt.

Combined loading: This mimics the effect of a crash with an airbag and a seatbelt.

The data set is in row-column form with 93 rows corresponding to the 93 tests. The variables are

load.cond :The test condition, one of **blunt** (Blunt hub loading), **comb** (Combined loading), **seatbelt** (Seatbelt loading) or **dist** (Distributed loading).
cmax:The maximum chest displacement.
fractures:The number of rib fractures.
age: Age.
gender:gender (m=male, f=female) **.**
testid:The test id number (you can ignore this).

Instructions

1. Load the data into R, and make a data frame **kent.df** to contain the data. Check for any typographical errors (the data below may be taken to be the correct data, but the data on the web may have been corrupted). Print out the last 10 lines of the data file. [5 marks]

2. What is the relationship between the number of fractures and cmax? Does the relationship depend on the test condition, age and gender? If so, how? Draw suitable plots to answer this question. Don't try and fit any models. [8 marks, 4 for the plots and 4 for the discussion]

3. The researchers were interested in whether the injury was "severe", defined as more than six rib fractures. Does the proportion of severe injuries depend on cmax and gender? How? Again, do not fit any models to answer this, but draw a suitable plot or plots. Hint: divide the range of cmax into suitable intervals, and plot the proportion of severe injuries falling in each cell of the resulting cmax-gender two-way table. The function **barplot** will be useful. [5 marks]

4. Are there any features of the data that might make fitting a regression model difficult? [2 marks]

More information about this question may be found in the article by Kent and Patrie, a copy of which is available on the web.

Question 2

The csv file Sitka.csv contains growth data on 40 Sitka spruce trees, 20 of which were grown in an ozone-rich environment, and 20 in a control environment. Each tree corresponds to five lines in the data file, with log-size measurements taken at days 152, 174, 201 , 227, 258. (The numbers represent days elapsed since 1 Jan 1988). The data are organized in row-column format, with 200 rows in all , corresponding to 5 measurements on each of 40 trees. The variables are

size: The log of the tree size (log of height times trunk diameter squared),
time: The days at which the measurements were taken, measured in days after Jan 1 1988,
tree: The tree ID
treat: The treatment given, either "ozone" or "control"

Instructions

1. Load the data into R, and make a data frame **sitka.df** to contain the data. Check for any typographical errors (the data below may be taken to be the correct data, but the data on the web may have been corrupted). Print out the last 5 lines of the data file, for each treatment. [5 marks]

2. Draw a Trellis plot showing the growth of each tree in the two groups. Does ozone have an effect on growth? Are there any trees whose growth is different from the others in the group? [7 marks]

3. Draw a plot similar to the one shown in class for the rat data, that shows the two groups on the same plot. Add lines to the plot representing the average size for the group at each time point. What do you conclude? [8 marks]

Total for assignment: 40 marks

R hints

1. Use **read.table** to create the data frames.

2. Use regular R graphics and trellis graphics to help answer Question 1.

3. Don't forget to label the axes. Use **xlab** and **ylab**. Skim the paper "A Tour of Trellis Graphics" on the web site for helpful hints.

Crash test data

load.cond	cmax	fractures	age	gender	testif
blunt	32.9	9	59	m	172/43fm
blunt	32.1	0	61	m	171/42fm
blunt	31.5	10	64	m	177/45fm
blunt	26.9	9	66	m	200/60fm
blunt	25.7	3	75	m	189/53fm
blunt	37.3	4	53	m	203/63fm
blunt	37.1	6	72	m	204/64fm
blunt	39.3	9	80	m	69/15fm
blunt	44.4	12	81	m	65/13fm
blunt	42	14	67	f	61/12ff
blunt	43.5	6	76	f	66/14ff
blunt	32.8	6	48	m	104/37fm

blunt	45.9	11	51	m	93/31fm
blunt	42.5	16	65	m	86/24fm
blunt	45.8	13	75	m	94/32fm
blunt	28.9	17	60	m	47/5fm
blunt	32.5	11	83	m	50/6fm
blunt	44.7	11	64	m	96/34fm
blunt	40.7	7	49	f	190/54ff
blunt	41.8	11	58	f	85/23ff
blunt	39.5	10	65	m	87/25fm
blunt	35	6	66	m	219/120fm
blunt	37	10	69	m	218/119fm
blunt	18.5	0	75	m	88/26fm
blunt	19.4	0	54	m	92/28fm
blunt	31	3	52	f	92/30ff
blunt	37.7	0	60	m	187/51fm
blunt	39.4	3	65	m	192/56fm
blunt	48.6	9	65	m	188/52fm
blunt	43.1	10	66	m	186/50fm
blunt	39	4	68	m	196/58fm
blunt	39.7	0	69	m	182/48fm
blunt	37.5	0	19	m	77/19fm
blunt	35	0	29	m	79/20fm
blunt	41.7	17	72	m	83/22fm
blunt	41.8	14	78	m	76/18fm
blunt	31	0	46	m	178/46fm
blunt	34.6	7	52	m	99/36fm
blunt	40.7	8	46	f	191/55ff
blunt	37	9	58	m	220/123fm
blunt	41.4	6	54	m	145
comb	23	0	57	m	577
comb	25	4	69	f	578
comb	34	11	72	f	579
comb	28	0	57	m	580
comb	28	3	55	m	665
comb	32	3	69	m	666
comb	36	13	59	f	667
comb	14.5	1	67	f	533
comb	18.3	4	47	m	534
comb	31.5	16	57	f	535
comb	15.7	3	67	m	545
comb	26.7	1	63	f	C11
comb	23.8	2	58	m	C12
comb	28.2	0	50	m	C13

seatbelt	27.1	6	60	m	THC75
seatbelt	26.1	6	64	f	THC77
seatbelt	28.7	3	43	m	THC79
seatbelt	27.4	10	63	m	THC93
seatbelt	25.2	2	63	m	THC91
seatbelt	11.8	0	64	f	THC76
seatbelt	11.8	0	43	m	THC78
seatbelt	10.2	0	63	m	THC90
seatbelt	12.1	0	63	m	THC92
seatbelt	34.1	8	47	f	THC11
seatbelt	35.4	0	17	f	THC12
seatbelt	25.1	2	86	f	THC13
seatbelt	29.8	17	69	m	THC14
seatbelt	29.1	3	60	m	THC15
seatbelt	35.4	4	59	m	THC16
seatbelt	30	7	71	m	THC17
seatbelt	9.5	0	72	m	THC61
seatbelt	14.4	0	71	m	THC64
seatbelt	11.3	0	40	m	THC68
seatbelt	36.3	6	67	m	THC18
seatbelt	28.4	4	83	f	THC19
seatbelt	30.1	18	70	m	THC20
seatbelt	22.8	4	72	m	THC62
seatbelt	36.1	10	71	m	THC65
seatbelt	30.6	1	40	m	THC69
seatbelt	40.8	14	63	f	147
dist	29.6	9	61	f	386
dist	71	29	45	f	387
dist	36.6	4	34	f	388
dist	43.7	25	68	f	421
dist	40.6	17	67	f	422
dist	42.2	13	51	f	423
dist	55.2	20	55	f	424
dist	0.35	4	69	m	116
dist	0.35	0	29	f	143
dist	0	4	40	m	650
dist	11	0	70	m	651
dist	12	0	46	m	652

Sitka Spruce data

size	time	tree	treat
4.51	152	1	ozone
4.98	174	1	ozone
5.41	201	1	ozone
5.9	227	1	ozone
6.15	258	1	ozone
4.24	152	2	ozone
4.2	174	2	ozone
4.68	201	2	ozone
4.92	227	2	ozone
4.96	258	2	ozone
3.98	152	3	ozone
4.36	174	3	ozone
4.79	201	3	ozone
4.99	227	3	ozone
5.03	258	3	ozone
4.36	152	4	ozone
4.77	174	4	ozone
5.1	201	4	ozone
5.3	227	4	ozone
5.36	258	4	ozone
4.34	152	5	ozone
4.95	174	5	ozone
5.42	201	5	ozone
5.97	227	5	ozone
6.28	258	5	ozone
4.59	152	6	ozone
5.08	174	6	ozone
5.36	201	6	ozone
5.76	227	6	ozone
6	258	6	ozone
4.41	152	7	ozone
4.56	174	7	ozone
4.95	201	7	ozone
5.23	227	7	ozone
5.33	258	7	ozone
4.24	152	8	ozone
4.64	174	8	ozone
4.95	201	8	ozone
5.38	227	8	ozone
5.48	258	8	ozone

4.82	152	9	ozone
5.17	174	9	ozone
5.76	201	9	ozone
6.12	227	9	ozone
6.24	258	9	ozone
3.84	152	10	ozone
4.17	174	10	ozone
4.67	201	10	ozone
4.67	227	10	ozone
4.8	258	10	ozone
4.07	152	11	ozone
4.31	174	11	ozone
4.9	201	11	ozone
5.1	227	11	ozone
5.1	258	11	ozone
4.28	152	12	ozone
4.8	174	12	ozone
5.27	201	12	ozone
5.55	227	12	ozone
5.65	258	12	ozone
4.47	152	13	ozone
4.89	174	13	ozone
5.23	201	13	ozone
5.55	227	13	ozone
5.74	258	13	ozone
4.46	152	14	ozone
4.84	174	14	ozone
5.11	201	14	ozone
5.34	227	14	ozone
5.46	258	14	ozone
4.6	152	15	ozone
4.08	174	15	ozone
4.17	201	15	ozone
4.35	227	15	ozone
4.59	258	15	ozone
3.73	152	16	ozone
4.15	174	16	ozone
4.61	201	16	ozone
4.87	227	16	ozone
4.93	258	16	ozone
4.67	152	17	ozone
4.88	174	17	ozone
5.18	201	17	ozone

5.34	227	17	ozone
5.49	258	17	ozone
2.96	152	18	ozone
3.47	174	18	ozone
3.76	201	18	ozone
3.89	227	18	ozone
4.3	258	18	ozone
3.24	152	19	ozone
3.93	174	19	ozone
4.76	201	19	ozone
4.62	227	19	ozone
4.64	258	19	ozone
4.36	152	20	ozone
4.77	174	20	ozone
5.02	201	20	ozone
5.26	227	20	ozone
5.45	258	20	ozone
4.53	152	1	control
5.05	174	1	control
5.18	201	1	control
5.41	227	1	control
5.42	258	1	control
4.97	152	2	control
5.32	174	2	control
5.83	201	2	control
6.29	227	2	control
6.45	258	2	control
4.37	152	3	control
4.81	174	3	control
5.03	201	3	control
5.19	227	3	control
5.4	258	3	control
4.58	152	4	control
4.99	174	4	control
5.37	201	4	control
5.68	227	4	control
5.93	258	4	control
4	152	5	control
4.5	174	5	control
4.92	201	5	control
5.44	227	5	control
5.87	258	5	control
4.73	152	6	control

5.05	174	6	control
5.33	201	6	control
5.92	227	6	control
6.01	258	6	control
5.15	152	7	control
5.63	174	7	control
6.11	201	7	control
6.39	227	7	control
6.61	258	7	control
4.1	152	8	control
4.46	174	8	control
4.84	201	8	control
5.29	227	8	control
5.48	258	8	control
3.22	152	9	control
3.85	174	9	control
4.47	201	9	control
4.85	227	9	control
5.11	258	9	control
2.23	152	10	control
2.89	174	10	control
3.16	201	10	control
3.4	227	10	control
3.52	258	10	control
3.65	152	11	control
4.36	174	11	control
4.76	201	11	control
5.18	227	11	control
5.44	258	11	control
3.4	152	12	control
3.92	174	12	control
4.5	201	12	control
4.97	227	12	control
5.14	258	12	control
5.16	152	13	control
5.49	174	13	control
5.74	201	13	control
6.05	227	13	control
6.21	258	13	control
4.04	152	14	control
4.52	174	14	control
5.15	201	14	control
5.59	227	14	control

5.87	258	14	control
4.52	152	15	control
4.91	174	15	control
5.04	201	15	control
5.71	227	15	control
5.97	258	15	control
4.56	152	16	control
5.12	174	16	control
5.4	201	16	control
5.69	227	16	control
5.89	258	16	control
4.9	152	17	control
5.35	174	17	control
5.71	201	17	control
6.12	227	17	control
6.25	258	17	control
4.83	152	18	control
5.1	174	18	control
5.43	201	18	control
5.59	227	18	control
6.04	258	18	control
5.46	152	19	control
5.79	174	19	control
6.12	201	19	control
6.41	227	19	control
6.63	258	19	control
4.17	152	20	control
4.67	174	20	control
5.16	201	20	control
5.56	227	20	control
5.75	258	20	control