BC 203 Introduction to Biostatistics

Problem Set One

Due: Thursday September 20, 2007 at the beginning of class

---

1. Pagano #7, 9, 15 (Chapter2, p.30)

2. Pagano #8 (Chapter 2, p. 30), calculate also the mean, median, range, variance, and standard deviation.

3. Pagano #6, 8 (Chapter 3, p. 59). For part (b) of #6, simply simplify.

4. (a) Do you think the mean daily rainfall for providence in 1999 was greater than, less than, or about the same as in Tucson, Arizona (which is located in a desert)?

   (b) How about the mode daily rainfall in Providence vs. Tucson?

   (c) How about the median daily rainfall in Providence vs. Tucson?

   (Lame Hint: What was the median daily rainfall in Providence last year? You do know the answer. Think about it.)

5. Pagano #9 (Chapter 3, p. 61).

6. Attached are tables derived from the UNICEF publication <u>The state of the World's Children 1997</u>. Listed are 150 countries, divided into 23 "industrialized" countries and 127 "developing" countries in order of their 1995 under 5 years mortality rate (per 1000). The latter category is further divided into six separate geographic regions.

   (a) Make side-by-side Boxplots showing how the rates for the seven groups of countries compare. (Don't worry if the group where the rates are lowest does not reveal much about the distribution of rates within that group. The purpose of this graph is to show comparisons <u>between</u> the seven groups.) What are your conclusions?

7. Attached is data showing the fluoride content (ppm) of the drinking water and the number of dental caries per 100 children in 21 communities (ref. Pagano 1993). Define two groups where Group 1 consists of the communities whose fluoride concentration is less than 0.5 ppm and Group 2 consists of those whose fluoride concentration is greater than or equal to 0.5 ppm.

   (a) Make side-by-side Boxplots comparing the number of dental caries in the two groups.

   (b) Make a graph showing the relationship between the caries rate and the fluoride content in these 21 communities. Does this graph reveal more about the relationship than the Boxplots in (a)? Explain.

   (c) Give a rough estimate of the caries rate that you would expect to see in a community where the fluoride concentration is 3.0 ppm.

   (d) What do these data suggest about the effect of increasing fluoride concentration (by 1 ppm) from 0.25 to 1.25 ppm?

   (e) What do these data suggest about the effect of increasing the fluoride concentrations (by 1 ppm) from 2.0 to 3.0 ppm?

8. Pagano #8 (Chapter 4, p. 90).

9. In the P.A.H.O. publication “Urban patterns of Mortality”, deaths and death rates from all causes and from specific causes of death are presented for males aged 15-74 in the years 1962-64:

   (i) The percentage of all deaths in Guatemala City that were reported to be due to malignant neoplasms was 13.5, and the same percentage in Lima, Peru was 16.9.

   (ii) The death rate in Guatemala City for malignant neoplasms was 173.6 per 100,000 population, and in Lima it was 91.9 per 100,000.

   On the basis of the information in (i) and (ii), what can you say about the all-cause death rates in Lima and Guatemala city?

   (Hint: What is generic formula for the percentage of all deaths due to cancer? And the death rate due to cancer? And the all-cause death rate?)

10. In an article entitled “Premature Baby Statistics” (British Medical Journal, 1997:1,1313-1315) neonatal mortality rates were presented for infants weighing 250 grams or less, born in the maternity Department of the Northern General Hospital, Sheffield, England, during a 27-year period. The data for two intervals of calendar time are shown below.

	1953-57	1968-72
Number of babies weighing 2500 grams or less	1080	1432
Number of neonatal deaths	231	183
Neonatal mortality rate/100 babies	21.4	12.8

(a) What factors might account for the decline in mortality rates?

The data on these premature babies was also presented by birth weight in 500 gram classes, as follows:

	1953-57		1968-72	
Birth Weight (grams)	**Number of Babies**	**Number of Deaths**	**Number of Babies**	**Number of Deaths**
501-1000	126	116	49	45
1001-1500	112	48	106	59
1501-2000	241	31	92	40
2001-2500	601	36	985	39
Total	1080	231	1232	183

(b) For each time period, calculate the relative frequencies of the four weight classes and the weight specific mortality rates. Using the 1953-57 population as the reference population, calculate mortality rate adjusted for birth weight (Just Direct adjustment will do).

(d) Using the rates of 1968-72 population as the reference rates, calculate the SMR (standardized mortality ratio) for the 1953-57 population.

(e) Following (d), calculate the indirectly adjusted infant mortality rate for the 1953-57 population.

(c) Can you now explain the decline in (crude) mortality rates? Which methods were most helpful?

11. Pagano #15 (Chapter 4, p. 93).
For (g) plot two graphs: (1) rates versus age and (2) log(rates) versus age (use the midpoint of the age interval). Because of a property of logarithms, parallel lines on a log scale imply that the actual rates are proportional.

Data for this homework is posted on the website in separate files. A text file called **hwk1_data.txt** contains all the data for this homework in tab-delimited form. If you want to use this file, you will have to edit it so that Stata can read it.

Useful Stata Commands (Stata's help is useful as well):

Command	Function
list	lists data
graph x y	graphs a scatterplot of x versus y
graph x, box	boxplot of x
graph x, hist	histogram of x
stem	stem and leaf display
summarize	summary statistics of data

infile var1 var2 using *path/filename*

The infile command tells Stata to read in data from a tab delimited text file (such as notepad). Each column in the file would be a variable (var1 and var2 in this example). **Note: check the help to be sure there is not a comma before `using' .**

If you don't feel like using the 'infile' command, you can type the data directly into the Stata editor (basically a spreadsheet) by clicking on the 'editor' button.

DATA: 1995 Under-5 mortality rate by region/country

Sub-Saharan Africa

	Country	Mort.
1.	Mauritius	23
2.	Botswana	52
3.	South Africa	67
4.	Zimbabwe	74
5.	Namibia	78
6.	Kenya	90
7.	Cameroon	106
8.	Congo	108
9.	Senegal	110
10.	Gambia	110
11.	Togo	128
12.	Ghana	130
13.	Rwanda	139
14.	Benin	142
15.	Gabon	148
16.	Cote d'Ivoire	150
17.	Chad	152
18.	Lesotho	154
19.	Tanzania	160
20.	Burkina Faso	164
21.	Madagascar	164
22.	Central African Rep.	165
23.	Burundi	176
24.	Zaire	185
25.	Uganda	185
26.	Nigeria	191
27.	Ethiopia	195
28.	Eritrea	195
29.	Mauritania	195
30.	Zambia	203
31.	Mali	210
32.	Somalia	211
33.	Liberia	216
34.	Malawi	219
35.	Guinea	219
36.	Guinea-Bissau	227
37.	Mozambique	275
38.	Sierra Leone	284
39.	Angola	292
40.	Niger	320

South Asia

	Country	Mort.
1.	Sri Lanka	19
2.	Nepal	114
3.	India	115
4.	Bangladesh	115
5.	Pakistan	137
6.	Bhutan	189
7.	Afghanistan	257

Middle East & North Africa

	Country	Mort.
1.	Kuwait	14

2.	United Arab Emirates	19
3.	Oman	25
4.	Jordan	25
5.	Saudi Arabia	34
6.	Syrian Arab Rep.	36
7.	Tunisia	37
8.	Lebanon	40
9.	Iran	40
10.	Egypt	51
11.	Algeria	61
12.	Libyan Arab J.	63
13.	Iraq	71
14.	Morocco	75
15.	Yemen	110
16.	Sudan	115

East Asia and Pacific

	Country	Mort.
1.	Hong Kong	6
2.	Singapore	6
3.	Korea, Rep. of	9
4.	Malaysia	13
5.	Korea, Dem. Rep. of	30
6.	Thailand	32
7.	Viet Nam	45
8.	China	47
9.	Philippines	53
10.	Mongolia	74
11.	Indonesia	75
12.	Papua New Guinea	95
13	Lao P. Dem. Rep.	134
14.	Myanmar	150
15.	Cambodia	174

Latin America and Caribbean

	Country	Mort.
1.	Cuba	10
2.	Jamaica	13
3.	Chile	15
4.	Costa Rica	16
5.	Trinidad and Tobago	18
6.	Panama	20
7.	Uruguay	21
8.	Venezuela	24
9.	Argentina	27
10.	Mexico	32
11.	Paraguay	34
12.	Colombia	36
13.	Honduras	38
14.	Ecuador	40
15.	El Salvador	40
16.	Dominican Rep.	44
17.	Peru	55
18.	Brazil	60
19.	Nicaragua	60
20.	Guatemala	67
21.	Bolivia	105
22.	Haiti	124

Central and Eastern Europe, Commonwealth of Independent States, and Baltic States

	Country	Mort.

1.	Czech Rep.	10
2.	Hungary	14
3.	Croatia	14
4.	Slovakia	15
5.	Poland	16
6.	Bosnia & Herzegovina	17
7.	Lithuania	19
8.	Bulgaria	19
9.	Belarus	20
10.	Estonia	22
11.	Yugoslavia	23
12.	Ukraine	24
13.	Latvia	26
14.	Georgia	26
15.	Romania	29
16.	Russian Red.	30
17.	Armenia	31
18.	TFRY Macedonia	31
19.	Moldova	34
20.	Albania	40
21.	Kazakhstan	47
22.	Turkey	50
23.	Azerbaijan	50
24.	Kyrgyzstan	54
25.	Uzbekistan	62
26.	Tajikistan	79
27.	Turkmenistan	85

Industrialized Countries

	Country	Mort.
1.	Sweden	5
2.	Finland	5
3.	Japan	6
4.	Denmark	7
5.	United Kingdom	7
6.	Switzerland	7
7.	Ireland	7
8.	Austria	7
9.	Germany	7
10.	Slovenia	8
11.	Canada	8
12.	Norway	8
13.	Austria	8
14.	Italy	8
15.	Netherlands	8
16.	France	9
17.	Israel	9
18.	New Zealand	9
19.	Spain	9
20.	Greece	10
21.	United States	10
22.	Belgium	10
23.	Portugal	11

Data: Number of Carries and Fluoride Concentrations for 21 Communities.

Community	Fluoride	Caries	Group
1	0.0	673	1
2	0.0	722	1
3	0.0	810	1
4	0.1	823	1
5	0.1	1037	1
6	0.1	772	1
7	0.1	706	1
8	0.2	703	1
9	0.2	733	1
10	0.3	652	1
11	0.4	556	1
12	0.5	444	2
13	0.6	412	2
14	0.9	343	2
15	1.2	258	2
16	1.2	303	2
17	1.2	281	2
18	1.3	323	2
19	1.8	252	2
20	1.9	236	2
21	2.6	246	2