ST104

UNIVERSITY OF WARWICK

December 1999

STATISTICAL LABORATORY I: Practical Examination

Open book exam: all course notes and books are allowed.

ALL questions should be attempted.

Your answer to each question should contain a final section headed: "Constructive Criticism".

Time allowed: 2 hours

Read carefully the instructions on the answer book and make sure that the particulars required are entered on each answer book.

---

1. The following table can be used to derive death rates (number of deaths divided by person-years) from coronary heart disease (CHD) amongst British male doctors, classified by age and according to whether or not they smoked.

	Person-years		**Coronary deaths**	
Age	**Non-smokers**	**Smokers**	**Non-smokers**	**Smokers**
35–44	18790	52407	2	32
45–54	10673	43248	12	104
55–64	5710	28612	28	206
65–74	2585	12663	28	186
75–84	1462	5317	31	102

[data from Breslow, N. (1985), 'cohort analysis in epidemiology'. Pp. 109–143 in *Celebration of Statistics*, eds. A. C. Atkinson & S. E. Fienberg, Springer-Verlag, New York]

Using simple numerical and graphical summaries, discuss the main features of the data.

What advice on smoking would you give to a 75 year old male doctor, and why?

Continued...

2. The following data were given in a newspaper article (*New York Times*, 31–May–1998), which claimed that Companies whose Chief Executive Officers had low golf handicaps, also had high performing stocks.

CEO	Company	Golf Handicap	Stock Rating
Melvin R. Goodes	Warner-Lambert	11.0	85
Jerry D. Choate	Allstate	10.1	83
Charles K. Gifford	Bank Boston	20.0	82
Harvey Golub	American Express	21.1	79
John F. Welch Jr.	General Electric	3.8	77
Louis V. Gerstner Jr.	IBM	13.1	75
Thomas H. O'Brien	PNC Bank	7.1	74
Walter V. Shipley	Chase Manhattan	17.2	73
John S. Reed	Citicorp	13.0	72
Terrence Murray	Fleet Financial	10.1	67
William T. Esrey	Sprint	10.1	66
Hugh L. McColl Jr.	Nationsbank	11.0	64
James E. Cayne	Bear Stearns	12.6	64
John R. Stafford	Amer. Home Products	10.9	58
John B. McCoy	Banc One	7.6	58
Frank C. Herringer	Transamerica	10.6	55
Ralph S. Larsen	Johnson & Johnson	16.1	54
Paul Hazen	Wells Fargo	10.9	54
Lawrence A. Bossidy	Allied Signal	12.6	51
Charles R. Shoemate	Bestfoods	17.6	49
James E. Perrella	Ingersoll-Rand	12.8	49
William P. Stiritz	Ralston Purina	13.0	48
Duane L. Burnham	Abbott Laboratories	15.6	46
Richard C. Notebaert	Ameritech	19.2	45
Raymond W. Smith	Bell Atlantic	13.7	44
Warren E. Buffett	Berkshire Hathaway	22.0	43
Donald V. Fites	Caterpillar	18.6	41
Vernon R. Loucks Jr.	Baxter International	11.9	40
Michael R. Bonsignore	Honeywell	22.0	38
Edward E. Whitacre Jr.	SBC Communications	10.0	37
Peter I. Bijur	Texaco	27.1	35
Mike R. Bowlin	Atlantic Richfield	16.6	35
H. Lawrence Fuller	Amoco	8.0	33
Ray R. Irani	Occidental Petroleum	15.5	31
Charles R. Lee	GTE	14.8	29
John W. Snow	CSX	12.8	29
Philip M. Condit	Boeing	24.2	25
Joseph T. Gorman	TRW	18.1	24
H. John Riley Jr.	Cooper Industries	18.0	22
Richard B. Priory	Duke Energy	10.0	22
Leland E. Tollett	Tyson Foods	16.0	20
Brude E. Ranck	Browning-Ferris	23.0	15
William H. Joyce	Union Carbide	19.0	13
Thomas E. Capps	Dominion Resources	18.0	12

The following further seven data points were omitted from the analysis, and were described as being 'scientifically sifted out'.

CEO	Company	Golf Handicap	Stock Rating
Scott G. McNealy	Sun Microsystems	3.2	97
William H. Gates	Microsoft	23.9	95
Sanford I. Weill	Travelers Group	18.0	95
Frank V. Cahouet	Mellon Bank	22.0	92
William C. Steere Jr.	Pfizer	34.0	89
Donald B. Marron	Paine Webber	25.0	89
Christopher B. Galvin	Motorola	11.7	3

[data also available from `http://www.dartmouth.edu/~chance/chance_news`].

The analysis was carried out by Graef Crystal, an investment compensation expert, who obtained data on golf scores from the magazine *Golf Digest* and who created a Stock Rating summarising stock market performance, with 100 being highest and 0 lowest.

The newspaper also reported that the study showed that 'executive wannabees... spend more time on the links', and quoted Graef Crystal as saying 'For all the different factors I've tested... this is certainly one of... the strongest.'

Plot the data, including the seven omitted points, in an appropriate way, and comment critically on the newspaper's analysis.

Continued...

3. Forty rats were divided randomly into four separate treatment groups of size 10, and each treatment group was given a different diet, distinguished by amount of protein (low/high) and by source of protein (beef/cereal). The reported weight gains were as follows:

Protein source	**Beef**	**Beef**	**Cereal**	**Cereal**
Protein amount	**Low**	**High**	**Low**	**High**
	90	73	107	98
	76	102	95	74
	90	118	97	56
	64	104	80	111
	86	81	98	95
	51	107	74	88
	72	100	74	82
	90	87	67	77
	95	117	89	86
	78	111	58	92

   (a) Using appropriate simple graphical displays and numerical summaries, discuss the effects of the different diets.

   (b) Carry out a formal statistical test on the combined beef and cereal data to investigate the diffence between a low- and a high-protein diet.

   (c) In your 'constructive criticism', discuss amongst other things whether or not the formal test you have carried out is a sensible procedure.

END.