475.101/102/107/108
STATISTICS

Assignment 4, Semester 2, 2000
Due: 4pm Thursday, 28th September

Instructions for handing in: PLEASE

1. Use (standard) **A4** sized paper.
2. Number each page in the top centre and **print** your name legibly at the **top right hand corner** of each page with the surname or family name underlined.
3. Attach a white *Department of Statistics Assignment Cover Sheet* to the front of the assignment. Staple or clip all pages together in the extreme top left hand corner.
4. Fold the paper length-wise so that the printed side of the cover sheet faces out.
5. **Print** your name, paper number and assignment number on the outside of the cover sheet. **Print** your ID number on the inside of the cover sheet.

Notes:

- Statistics is about summarising, analysing and communicating information. Communication is an important part of statistics. For this reason you will be expected to write answers which clearly communicate your thoughts. **The mark you receive will be based on your written English** as well as your statistical/technical work.
- Assignment 4 will be marked out of 60 marks, 54 marks for questions 1 - 7 as shown below and 6 marks for communication, style and presentation. Please refer to Section B, Questions 14 to 16, in your Course Resource Book for examples of how to set out your assignment answers. Your final mark will be converted to a mark out of 10 which will be recorded towards your course work.

Question 1. [7 marks]

It is known that the average stay of tourists in a Queenstown hotel has been 3.4 nights. A tourism industry analyst wanted to know whether this is still the case due to a recent fall in the New Zealand dollar. The analyst obtained a random sample of 32 tourists and recorded the number of nights each spent at this hotel. The results are given below:

2, 5, 5, 5, 2, 3, 2, 4, 4, 2, 8, 4, 4, 3, 1, 3, 3, 5, 3, 1, 5, 6, 5, 3, 6, 3, 2, 5, 8, 5, 1, 2

Summary statistics: $\bar{x} = 3.75$, $s \approx 1.8316$, $n = 32$

(a) **By hand** calculate a 95% confidence interval for the number of nights spent at this hotel. Interpret your result. Use 1-2 sentences of plain English.

(b) Does your confidence interval from part (a) contain the true mean? Briefly explain your answer. Use 2-3 sentences of plain English.

Question 2. [17 marks]

(a) For each of the following *P-values* describe the strength of evidence against the null hypothesis:

(i) 0.1189 **(ii)** 0.0780 **(iii)** 0.5912 **(iv)** 0.0078

(b) Using your tables for the Student's *t*-distribution obtain upper and lower values that a *P-value* can take in the following cases:

(i) $H_0 : p_1 = p_2$ vs $H_1 : p_1 \neq p_2$ where $t_0 = 2.484$ and $df = \infty$

(ii) $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 < \mu_2$ where $t_0 = -4.532$ and $df = 44$

(c) In each of the following cases indicate whether a hypothesis test or a confidence interval is the more appropriate procedure to initially carry out. Briefly explain.

(i) A broker on the Auckland Stock Exchange wishes to investigate the amount of time between placement and execution of a market order. A sample of orders was taken over a period of weeks and the execution time was recorded. You are asked to estimate the true mean execution time.

(ii) A production line operates with a filling weight standard of 16 grams per container. Overfilling or underfilling is a serious problem, and the production line should be shut down for adjustment if this occurs. The quality inspector sampled 35 containers in order to assess whether the containers are being filled correctly, overfilled or underfilled. Is there any indication that the production line should be shut down for adjustment?

(d) In each of the following explain why the hypotheses are not valid. Rewrite each hypothesis correctly.

(i) $H_0 : p_1 \neq p_2$ vs $H_1 : p_1 = p_2$

(ii) $H_0 : \bar{x}_1 - \bar{x}_2 = 0$ vs $H_1 : \bar{x}_1 - \bar{x}_2 \neq 0$

(e) For each of the following statements, state whether they are true or false. If a statement is false, rewrite correctly.

(i) A *test statistic* is a measure of discrepancy between the true value and the hypothesised value.

(ii) We cannot rule in a hypothesised value for a parameter, we can only determine whether there is evidence to rule out a hypothesised value.

(iii) The margin of error is the quantity added to and subtracted from the standard error to construct a confidence interval.

(iv) A two-tailed *t*-test will reject (at the 5% level) values for the null hypothesis that are inside the corresponding 95% confidence interval.

(v) It makes no sense to try to interpret a *P-value* without explicitly stating the relevant hypotheses.

(vi) For 95% confidence intervals, our confidence in a particular interval comes from the fact that the method works 19 times out of 20.

NOTES:

- If you are asked to carry out a *t*-test, you will need to:

1. state the parameter in words.
2. state the null and alternative hypotheses.
3. when instructed use either *Excel* or MINITAB.
   - make sure you are prepared for both questions 4 and 5 before you begin to use the computer.
   - **include** computer output.
   - when using the computer report *P-values* to 3 decimal places.
4. when instructed to carry out a *t*-test using *Excel* **assume unequal variances**.

Question 3. [5 marks]

This question refers to the details in Question 1.

(a) Is there a significant difference in the mean number of nights spent at this hotel since the recent fall in the New Zealand dollar? **By hand** carry out a two-tailed *t*-test to investigate this. Interpret your result using plain English in 1-2 sentences.

(b) Can we justify carrying out a one- tailed *t*-test instead of a two- tailed *t*-test here? Briefly explain.

(c) Without doing any calculations, would a 90% confidence interval for μ contain 3.4? Briefly explain.

Question 4. [9 marks]

Top companies are often judged by productivity, as measured by revenue and/or profit per employee. The following table shows (in dollars) the revenue per employee for 22 companies with a head office in Sydney and 12 companies with a head office in Melbourne.

Sydney		*Melbourne*
172,880	144,797	206,347
86,734	46,614	131,076
165,025	391,224	158,892
172,614	220,765	312,297
335,198	180,733	240,692
103,284	123,439	224,420
347,609	421,784	91,030
355,656	355,275	130,247
275,355	331,528	195,909
139,678	281,302	398,902
342,659	35,817	351,615
		707,161

(a) Draw appropriate plots for this data set, **preferably using a computer**. Using your plots, in plain English, **briefly** comment on the data in terms of the original story.

(b) Is there a significant difference in the mean revenue per employee for companies with head offices in Sydney and the mean revenue per employee for companies with head offices in Melbourne? Use ***Excel*** to investigate this question. Note: A 95% confidence interval for difference in the mean revenue per employee for companies with head offices in Sydney and that of Melbourne is [–$149,217, $81,723]. Interpret your results (including that of the confidence interval).

Question 5. [6 marks]

In May 2000 the Consumer magazine published the results from a survey on international airlines sent to 12,000 Consumer members in January and February this year. Consumer obtained information on 4,483 trips on international airlines in the previous 12 months. One of the questions asked in the survey was "*Would you recommend the airline to a friend*?" The airline travelled and the number of travellers who answered "Yes, definitely" to this question is given in the table below:

Airline	*Numbers of Travellers*	*Travellers who answered "Yes, definitely"*
Singapore	316	194
Malaysia	62	30
Thai	66	29
Cathay Pacific	40	16
Freedom Air	184	16
Air New Zealand	1841	732
British Airways	100	24
Qantas	1072	256
Air Pacific	77	18
United Airlines	239	41

(a) Treating this as 2 independent samples, is there a significant difference in the proportion of travellers using Singapore Airlines who would definitely recommend the airline to a friend and the corresponding proportion of travellers using Thai Airlines? Use **MINITAB** to investigate this question. Interpret your results.

(b) In 1 or 2 sentences, briefly explain why this data set should not be treated as 2 independent samples.

Question 6. [5 marks]

The National Centre on Addiction and Substance Abuse (CASA) at Columbia University regularly surveys attitudes of teens and those who most influence them. As part of the 1999 CASA "National Survey of American Attitudes on Substance Abuse V: Teens and Their Parents", 2000 teenagers aged between 12 and 17 (1000 boys and 1000 girls), were interviewed by telephone on many lifestyle issues. The data in the table below is given for two questions answered by the teenagers in the survey.

Would you describe your relationship with your mother as excellent, very good, good, fair, or poor?			
	Total	**Male**	**Female**
Excellent	770	380	390
Very Good	680	350	330
Good	350	190	160
Fair	160	70	90
Poor	40	10	30
Would you describe your relationship with your father as excellent, very good, good, fair, or poor?			
	Total	**Male**	**Female**
Excellent	590	330	260
Very Good	580	310	270
Good	460	210	250
Fair	210	90	120
Poor	130	50	80
Don't Know	30	10	20

(a) Identify the sampling situation as (a) *Two independent samples*, (b) *Single sample, several response categories* or (c) *Single sample, two or more Yes/No items* in the following cases:

(i) We want to compare the proportion of male teenagers who describe their relationship with their mother as very good and the proportion of male teenagers who describe their relationship with their father as very good.

(ii) We are interested in comparing the proportion of male teenagers who describe their relationship with their father as poor and the proportion of female teenagers who describe their relationship with their father as poor.

(iii) We are interested in comparing the proportion of teenagers who describe their relationship with their mother as fair and the proportion of teenagers who describe their relationship with their mother as poor.

(b) Calculate the standard errors for the cases described in parts **(i)** and **(ii)** in **(a)** above.

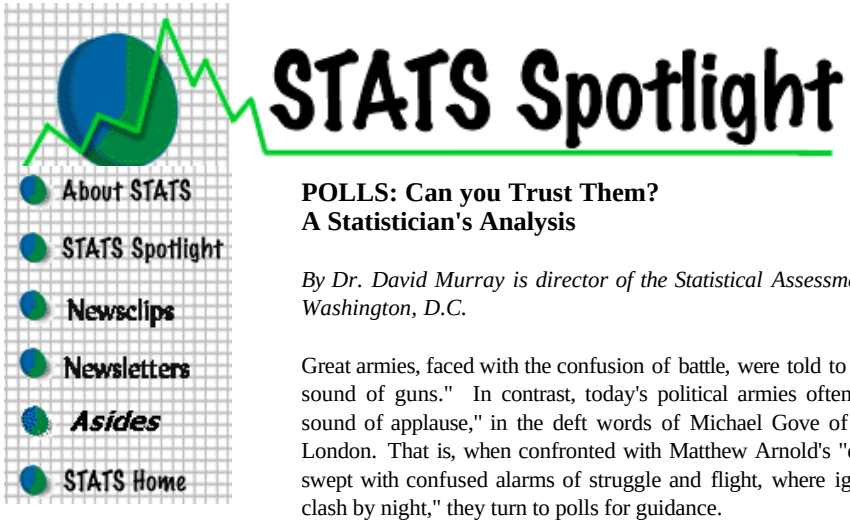
Question 7. [5 marks]

Read the article *POLLS: Can you Trust Them*? (written by Dr David Murray, obtained from TomPaine.com, July 24) on the next page and answer the following questions. Refer to §8.5.3 "Margins of Error" in Media Reports, page 349, in **Chance Encounters**.

(a) Why does Dr David Murray think that the CNN headline "Poll: Gore Closes Gap Against Bush," abuses the results from a recent poll? Briefly explain.

(b) What point is made about sub-samples in polls?

(c) In media reports of polls, what statements should be included so that it is difficult to abuse poll results as was done by the CNN report?



POLLS: Can you Trust Them?
A Statistician's Analysis

By Dr. David Murray is director of the Statistical Assessment Service in Washington, D.C.

Great armies, faced with the confusion of battle, were told to "march to the sound of guns." In contrast, today's political armies often "steer to the sound of applause," in the deft words of Michael Gove of the Times of London. That is, when confronted with Matthew Arnold's "darkling plain, swept with confused alarms of struggle and flight, where ignorant armies clash by night," they turn to polls for guidance.

So long as polls are kept in their proper place, used as occasional check points on the road towards one's principals, they are useful servants. But when polls are used to determine destinations in the first place, we've entered a Humpty Dumpty world, only now upside down-soothing words and pithy phrasings become the masters of political principal. This state of affairs creates two problems. Not only do voters detect this pusillanimous procedure and, seeking authenticity, turn away, there is an additional danger to political strategists themselves -- they might believe the polls, and change their tactics accordingly. Though polls have their scientific methods and their authoritative numerical window-dressing, they have a capacity to mislead. When that happens more than one egg-head is destined for a "great fall."

A case in point is the interpretation of a recent (July 17) CNN/USA Today/Gallup presidential poll. Under the headline, "Poll: Gore Closes Gap Against Bush," CNN announced a sudden turn around in performance in a very short time. Scarcely a week ago in a comparable poll, candidate George W. Bush led Vice President Gore by a comfortable margin -- 50 percent to 41 percent. But a poll of 635 "likely voters" taken July 14th through 16th showed a considerable swing. Bush now stood at only 48 percent, to Gore's 46 percent, a scant 2 percentage point difference. These numbers no doubt are an encouraging sign for a formerly dispirited Democratic party. But then things get a little chancy, as the electioneers interpret what we're seeing in the numbers. Gore is said to be "capitalising on gains among women and Independents," and the numbers do appear to make this case. In the prior poll of July 6th through 9th, Gore got only 43 percent of the female voters, while now he receives 51 percent. Only 32 percent of Independents went for Gore last week, but he now stands as 44 percent; that is, he picked up a whopping 12 percent of additional support.

The real problem emerges when the CNN analysis identifies the political factor that must have turned the trick. What happened during the week? The endorsement of Bill Bradley, whose "natural constituency during the Democratic primary" are those now said to be breaking for Gore. Leaving aside for the moment the common problems of all polls (it's only valid "if the election were to be held today," we don't know whether those polled were actually a representative sample of those who will vote on election day, we don't know the non-response rate or the effect of a telephone methodology, etc.), take a gander at those margins of error. For the "constituencies" in question, they range between 4 and 6 percent. Because the original sample is only 635 individuals, the overall margin of error is 4 percentage points, plus or minus. Hence, a result that shows a lead of 48 percent to 46 percent is really only one outcome taking place between two "error bars" that are actually 8 percentage points apart -- 4 percentage points higher or lower than the actual result. Thus, the actual finding of the poll is a range of outcomes spread between a Bush lead of 52 percent over Gore's 42 percent (to take one possible outcome of the 8 percentage point swing) or a Gore lead over Bush (50 percent to 42 percent), a result also contained within an 8 percentage point margin of error swing. It gets even worse when it comes to interpreting numbers for males, females and Independents. Since these are sub-samples of the overall sample, and have roughly 320 individuals in each category, the margin of error swells to 6 points plus or minus -- a 12 percentage point overall swing. Hence, this week's result show that Gore may lead Bush for female likely voters by 51 percent to 44 percent, but the actual possible spread for the finding runs between a huge Gore lead over Bush (57 percent to 37 percent), or a Bush advantage over Gore (50 percent to 45 percent). The same argument pertains to the results for Independents.

The critical point for the political prophets is that the apparent swing from the week of July 6-9 to the week of July 14-16 is within the margin of error of either poll. That is, it could well be an exercise in self-delusion if Democratic (or Republican) strategists read the results of this poll and concluded anything at all about the impact of Bill Bradley's endorsement.

It could have been meaningful. It could have been an anomaly (an interpretation strengthened by the odd finding that among eighteen to twenty-nine-year-olds, Bush walloped Gore 64 percent to 30 percent -- a 15 percentage point decline from Gore's standing a week before. Be endorsed by Bradley and lose 15 points among the young?) or it could just have been the fluctuation of statistical noise represented by the sloshing around within the margin of error. The take home message? Findings such as these make great headlines, but as a guide for political generals, they are not a map to steer by.

Note: The above article was obtained from the Statistical Assessment Service on http://www.stats.org/. The Statistical Assessment Service examines the way that scientific, quantitative, and social research are presented by the media, and works with journalists to help them convey this material more accurately and effectively.