STATISTICS 151 SECTION 1 FINAL EXAM
MAY 2 2009

This is an open book exam. Course text, personal notes and calculator are permitted. You have 3 hours to complete the test. Personal computers and cellphones are not allowed. If you have any queries about the meaning of a question, ask the instructor for advice.

Answers are to be written in a blue book.

Before you begin, copy out the following pledge and sign it. If your blue book has a preprinted pledge, you may sign that instead.

Honor pledge: On my honor, I have neither given nor received unauthorized aid in this exam.

SHOW ALL WORKING — even correct answers will not get full credit if it's not clear how they were obtained. Incorrect answers will gain substantial credit if the method of working is substantially correct.

Answer six of the eight questions. If you attempt more than six, all the answers will be graded but only the best six (complete questions) will count. Each question is worth a total of 20 points and the whole exam is worth 120 points (which will be rescaled to a maximum of 35 for grading purposes). Points for each individual part of a question are also given in square brackets.

1. One of the concerns about global warming is that higher temperatures may lead to higher concentrations of air pollutants, with consequent adverse effects for human health. This is of particular concern with ground-level ozone, which is more prevalent at high temperatures, and which is regulated by the EPA because of its adverse health effects.

   The following table shows the mean summer temperatures and the mean summer ozone levels for 12 US cities, for 1987–2000:

City	Temperature x (°F)	Ozone y (ppb)
Atlanta	72	54
Chicago	63	38
Denver	62	44
Dallas/Fort Worth	77	51
Houston	78	46
Los Angeles	68	56
New York	66	41
Phoenix	85	53
Salt Lake City	65	53
San Diego	67	52
Seattle	59	34
Washington	66	44

   [$\bar{x} = 69.0$, $s_x = 7.6$, $\bar{y} = 47.2$, $s_y = 7.1$, and the correlation coefficient r is 0.56.]

   (a) Draw a scatterplot of y (vertical axis) against x (horizontal axis). Comment briefly on its features (e.g. is there a close relationship between x and y? If so, is it linear?). [**7 points**]

   (b) Suppose we want a formal regression line $\hat{y} = a + bx$. Calculate a and b and sketch the line (roughly) on the plot. [**7 points**]

(c) Now let's suppose that this relationship (which is based on past data comparing different cities) is also valid for calculating future ozone levels as a function of temperature. If mean temperature rises by 5°F as a result of global warming, how much would you expect ozone to rise? [**6 points**]

2. A large number of patients at a medical facility are examined to determine whether they have (a) heart disease, or (b) lung disease, or (c) neither. In this study, we don't consider the possibility that a patient may have both heart and lung disease; assume that if the doctor sees signs of both, she decides which is more severe and reports that one.

   Over the whole population under study, 20% have heart disease, 10% have lung disease, and 70% have neither. Among those with heart disease, 30% are smokers; among those with lung disease, 50% are smokers; among those with neither, 10% are smokers.

   (a) Represent this information in the form of a tree diagram. [**8 points**]

   (b) What proportion of the population are (i) smokers who also have lung disease? (ii) non-smokers? [**4 points**]

   (c) Given that someone is a smoker, what is the probability that he or she (i) has heart disease, (ii) has lung disease? [**8 points**]

3. Alex is a professional baseball player with a hitting rate of .280; that is, of all the times he is "at bat", the proportion of times he successfully hits the ball and gets to first base is .280. Assume that individual attempts are independent and obey the assumptions of a binomial distribution.

   In a single game, Alex is at bat 5 times.

   (a) What is the probability Alex gets no hits (i.e. all 5 attempts are failures)? [**3 points**]

   (b) What is the probability Alex gets at least 2 hits out of the 5 attempts? [**5 points**]

   Now suppose that over the next 15 games, Alex will be at bat a total of 70 times. Assume that his overall performance does not change, i.e. the probability he gets a hit on any single at-bat remains at .280.

   (c) What are the mean and standard deviation of X, the total number of hits he gets in these 70 at-bats? [**3 points**]

   (d) What is the probability that X is at least 25? [**4 points**]

   (e) Find a number x so that the probability that Alex scores at least x hits in the next 70 at-bats is approximately 0.75. [**5 points**]

4. A researcher believes she has a new drug that will lower someone's resting heart rate without adverse side effects. To see whether the drug in fact reduces patients' heart rate, she performs a trial run with just five patients, intending to carry out a more systematic trial if this trial run is successful. The five patients to whom she administers the drug all show a reduction in heart rate: the amounts of the reductions (in beats per minute) are 2, 5, 10, 15, 17 for the five patients. Assume that these form a random sample of size 5 from a larger population of potential users of the drug.

   (a) Formulate a suitable null hypothesis and alternative hypothesis for this problem. [**3 points**]

   (b) Calculate $\bar{x}$ and the sample standard deviation s. [**5 points**]

(c) Carry out the test in (a). Do you reject the null hypothesis at a significance level .05? [**9 points**]

(d) Briefly discuss the test and its implications. Would you recommend the researcher go on to conduct a full study? [**3 points**]

5. An Extra-Sensory Perception (ESP) game consists of a subject being shown five cards with different symbols on them (a square, a triangle, a circle, a hexagon and a star). The cards are turned over and shuffled, and the subject draws out one of them face down. She is then asked to guess which of the five symbols she has selected. Since the expected proportion of successes by random guessing is one-fifth, an observed proportion that is significantly greater than this may be taken as evidence that the subject indeed has ESP. The experiment is repeated a total of 180 times.

   (a) Formulate this as a hypothesis testing problem and state H_0 and H_a. [**3 points**]

   (b) Show that a one-sided test of significance level .05 will reject the null hypothesis whenever $\hat{p} > .249$, where $\hat{p}$ is the observed proportion of successes in 180 trials. [**8 points**]

   (c) Suppose the true probability of success is 0.3. What is the probability of a type II error? [**9 points**]

6. In an election for student class president, 200 students (80 men, 120 women) are eligible to vote. In the election, 25 men (31.25%) and 61 women (50.83%) actually do vote. Assume that these numbers are representative of voting tendencies among some larger population of men and women.

   (a) Calculate a 90% confidence interval for the difference between the proportion of women who vote and the proportion of men who vote. [**9 points**]

   (b) Carry out a formal test of the hypothesis that the proportions of men and women who vote are the same. Be careful to state the null and alternative hypotheses, and any additional assumptions you make in the course of calculating the test. [**9 points**]

   (c) Briefly summarize your conclusions from parts (a) and (b). [**2 points**]

7. 15 students took exams in History and English, with the following results:

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	SD
History	58	81	61	47	43	62	58	59	61	12	46	41	32	22	71	50.3	18.3
English	64	80	54	73	54	61	57	70	54	55	61	65	52	40	72	60.8	10.2
Difference	–6	1	7	–26	–11	1	1	–11	7	–43	–15	–24	–20	–18	–1	–10.5	14.1

   (a) Analyze the test results as an "independent samples" test for the comparison of two means. Based on this test, do you believe there is a statistically significant difference between students' scores in History and English? [**8 points**]

   (b) Analyze the test results as an "paired comparisons" test for the comparison of two means. Based on this test, do you believe there is a statistically significant difference between students' scores in History and English? [**7 points**]

   (c) Which of the two tests is more appropriate in this situation? State the assumptions of the tests, and comment whether those assumptions appear to be satisfied for this example. [**5 points**]

8. Three professors, Smith, Jones and Higgins, give a statistics exam. The breakdown of grades for each professor is as follows:

Grade	Prof. Smith	Prof. Jones	Prof. Higgins	Total
A	15	30	20	65
B	35	40	45	120
C	30	25	35	90
D or F	10	5	10	25
Total	90	100	110	300

We are interesting in determining whether there are systematic difference of grading policy among the three professors.

(a) Formulate a suitable null and alternative hypothesis. Is this a test of independence, a test of homogeneity, or something else? Explain briefly. [**5 points**]

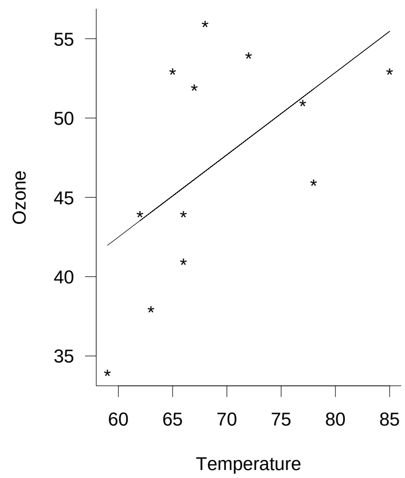
(b) Calculate a chi-squared test. Do you reject the null hypothesis in (a) at the .05 level of significance? [**9 points**]

(c) Of all the entries in the above table, which one contains the strongest evidence that there might, indeed, be a difference among the professors' grading policies? [**3 points**]

(d) Briefly discuss and interpret your answers to (b) and (c). [**3 points**]

SOLUTIONS

1. (a) See the following:

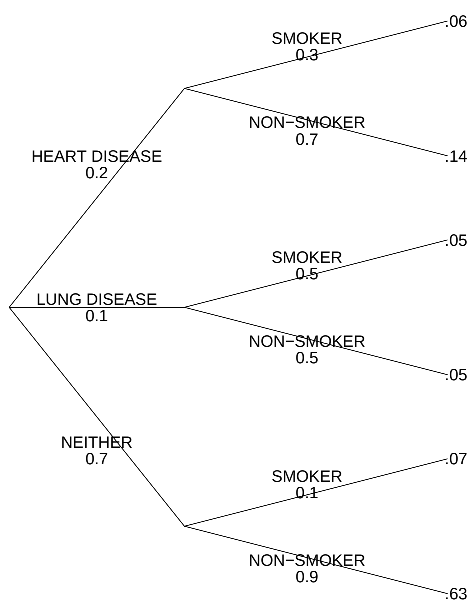


The plot shows a clear positive relationship between temperature and ozone, but there is also a lot of scatter. The relationship is linear as closely as can be judged, and there are no apparent outliers.

(b) $b = r\frac{s_y}{s_x} = \frac{0.56 \times 7.1}{7.6} = 0.5232, a = \bar{y} - b\bar{x} = 47.2 - 0.5232 \times 69.0 = 11.102$. The line is plotted on the above figure.

(c) A temperature rise of 5 °F would correspond to a mean increase in ozone of $5 \times 0.52 = 2.6$ ppb. (Alternatively: the new mean $\bar{y}$ would be 49.8 ppb, which is a rise of 2.6 ppb compared with the current $\bar{y}$.)

2. (a) See the following:



(b) (i) .05 (ii) .14+.05+.63=.82

(c) $\frac{6}{18}$ (or $\frac{1}{3}$) and $\frac{5}{18}$.

3. (a) $0.72^5 = 0.193$.

(b) The probability of no hits is 0.193 and the probability of 1 hit is $5\times.28\times.72^4 = .376$; hence the probability of at least 2 hits is $1 - .193 - .376 = .431$. [Some students incorrectly interpreted the question as the probability of exactly 2 hits, for which the answer is $\frac{5\times4}{2}\times(.28)^2\times(.72)^3 = .2926$. I gave partial credit for that.]

(c) The mean is $70\times.28 = 19.6$ and the standard deviation $\sqrt{70\times.28\times.72} = 3.757$.

(d) Corresponding to $x = 25$, we have $z = \frac{25-19.6}{3.757} = 1.44$ to two decimal places. The left tail probability is .9251; therefore, the right-tail probability is 1–.9251=.0749 (say, about 7%).

(e) The z statistic corresponding to a right-tail probability of .75 (left-tail probability .25) is –0.67 to two decimal places; that translates to $x = 19.6 - 0.67\times3.757 = 17.1$; say $x = 17$ to the nearest whole number.

4. (a) Let μ be the mean reduction in heart rate in the whole population. Then $H_0:\ \mu = 0$ against $H_a:\ \mu > 0$. (You could also consider a two-sided test, $H_a:\ \mu \neq 0$, though in this case the wording of the question, focussing specifically on the claim that the new drug reduces heart rate, suggests a one-sided test.)

(b) $\bar{x} = 9.8,\ s = \sqrt{\frac{7.8^2+4.8^2+0.2^2+5.2^2+7.2^2}{4}} = 6.38.$

(c) The standard error is $SE = \frac{s}{\sqrt{n}} = \frac{6.38}{\sqrt{5}} = 2.853$ and hence the t statistic is $t = \frac{9.8}{2.853} = 3.43$.

With $df = 4$, the right-tail probability is between .025 and .01. Since the test is one-sided, the P-value is between the same two values. Since the P-value is clearly less than .05, the result is statistically significant.

(d) Despite the small sample size, the reduction in mean heart rate is statistically significant. This suggests that the researcher should conduct a full study.

5. (a) Let p be the long-run proportion of successes (correct guesses). Since the probability of success under random guessing is .2, the natural null and alternative hypotheses are $H_0:\ p = 0.2$ and $H_a:\ p > 0.2$. We formulate a one-sided alternative because, in this setting, a result $p < 0.2$ would not make sense.

(b) The standard error under H_0 is $\sqrt{\frac{.2\times.8}{180}} = .0298$. A one-sided test with significance level .05 corresponds to a z value of 1.645; therefore, we reject H_0 if $\hat{p} > 0.2 + 1.645\times.0298 = .249$. [Some students did this by testing the value .249 as if this were actually the observed $\hat{p}$, finding a P-value of almost exactly .05. Some students then went a step beyond that, arguing that if $\hat{p}$ was a bit bigger than .249, the P-value would be less than .05, while if $\hat{p}$ was a bit smaller, the P-value would be greater than .05, so that .249 is actually the cut-off value at which the test goes from insignificant to significant. I was willing to give credit for that if the argument was spelled out, but the real point of the question was to explain where the number .249 came from, not to conduct a test under that $\hat{p}$.]

(c) If $p = 0.3$, then the standard error of $\hat{p}$ is $\sqrt{\frac{.3\times.7}{180}} = .0342$. The test will wrongly accept H_0 if $\hat{p} < .249$, so the task becomes to calculate the probability of this event under $p = 0.3$. The z statistic is $\frac{.249-.3}{.0342} = -1.49$. By Table A, the left-tail probability associated with this event is .0681. Allowing for some rounding error, the probability of type II error is about 7%.

6. (a) Let p_1 be the proportion of women who vote and p_2 be the proportion of men who vote. The estimate of $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2 = .5083 - .3125 = .1958$. The standard error is

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{120} + \frac{\hat{p}_2(1-\hat{p}_2)}{80}} \quad = \quad .0691.$$

Therefore, a 90% confidence interval is

$$.1958 \pm 1.645 \times .0691 = .1958 \pm .1137 = (.0821, .3095).$$

(b) The natural hypotheses are $H_0 : \ p_1 = p_2$ and $H_a : \ p_1 \neq p_2$. Assuming H_0 is true, the pooled value of $p_1 = p_2 = p$ is $\hat{p} = \frac{25+61}{200} = .43$, so the standard error (SE) for the hypothesis test is

$$\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{120} + \frac{1}{80}\right)} \quad = \quad .0715.$$

The test statistic is $z = \frac{\hat{p}_1 - \hat{p}_2}{SE} = \frac{.1958}{.0715} = 2.74$ to two decimal places. From Table A, the *left-hand* probability associated with this z value is .9969; therefore, the one-sided P-value is $1 - .9969 = .0031$, and the two-sided P-value is twice that, .0062. The P-value is clearly less than .05. Therefore, *reject* the null hypothesis.

(c) The statistical test shows that the differences in voting proportions among men and women is genuine. This could have been anticipated from part (a), since 0 is well outside the 90% confidence interval for $p_1 - p_2$.

7. (a) We have $\bar{x}_1 - \bar{x}_2 = -10.5$; the standard error is $\sqrt{\frac{18.3^2}{15} + \frac{10.2^2}{15}} = 5.409$; the t statistic is $-\frac{10.5}{5.409} = -1.941$. By looking up $t = 1.941$ in Table B with $df = 14$, we find that the one-sided tail probability is between 0.05 and 0.025; therefore, the two-sided P-value is between 0.1 and 0.05. Not statistically significant at the 5% level. Note that this is one example where it actually makes a difference whether you consider a one-sided or two-sided test (if a one-sided test, it would be considered statistically significant) but the context of the question here clearly indicates a two-sided test since there was no a priori reason to think that students would score better in English than History.

(b) A paired comparison test is based on the individual differences $x_1 - x_2 = -6, 1, 7, -26, ...$ for which the mean is –10.5 and the standard deviation is 14.1. Testing the null hypothesis that the mean difference is 0, we have a standard error $\frac{14.1}{\sqrt{15}} = 3.641$; the t statistic is $-\frac{10.5}{3.641} = -2.884$. By Table B for $df = 14$, the one-sided tail probability is between .01 and .005; the two-sided P-value is therefore between .02 and .01. This is statistically significant at the 5% level.

(c) They are not independent samples because it is the same 15 students taking both tests; therefore, a paired comparison test seems more appropriate. The other assumptions are that the data are quantitative and from a normal dstribution; these appear to be satisfied though there are some possible outliers (e.g. student 10). However with this caveat, it does appear that there is a statistically significant difference between the History and English scores, with students performing better in English.

8. (a) If we use symbols like $P_S(A)$ as an abbreviation for "the probability that Smith gives an A", then the natural null hypothesis is

$$H_0 : \quad P_S(A) = P_J(A) = P_H(A), \ P_S(B) = P_J(B) = P_H(B),$$
$$P_S(C) = P_J(C) = P_H(C), \ P_S(D/F) = P_J(D/F) = P_H(D/F),$$

and H_a is that at least one of the equalities in H_0 is false. This is a test of homogeneity: the column totals (number of students in each class) is fixed, but the hypothesis amounts to the claim that the conditional probability of getting a certain grade, given which professor, is the same for all three professors.

(b) Rounded to two decimal places, the tables of expected values and of $\frac{(\text{observed}-\text{expected})^2}{\text{expected}}$ values are

$$\begin{pmatrix} 19.5 & 21.67 & 23.83 \\ 36.0 & 40.00 & 44.00 \\ 27.0 & 30.00 & 33.00 \\ 7.5 & 8.33 & 9.17 \end{pmatrix} \text{ and } \begin{pmatrix} 1.04 & 3.21 & 0.62 \\ 0.03 & 0.00 & 0.02 \\ 0.33 & 0.83 & 0.12 \\ 0.83 & 1.33 & 0.08 \end{pmatrix}$$

The chi-squared statistic (total of all the numbers in the last table) is 8.44, with $3 \times 2 = 6$ degrees of freedom. According to Table C of the text, the right-tail probability is between 0.25 and 0.1. Since the P-value is greater than .05, the result is not statistically significant.

(c) The largest single entry in the table of $\frac{(\text{observed}-\text{expected})^2}{\text{expected}}$ is 3.21, corresponding to the number of As given by Professor Jones.

(d) The table suggests some differences among the three professors, in particular, Jones seems to give more As than the other two professors. However, the fact that the chi-squared test led did not reject the null hypothesis implies that the differences could be random — there is no systematic evidence of a discrepancy in grading policy.