MATH 243 (PHILLIPS, WINTER 2006): REAL FINAL SOLUTIONS

1. (Multiple choice: 4 points/part; 44 points total.) Circle the letter of the best answer.

1A. We choose a sample of 100 entering freshmen at the University of New Pinchland, ask for their math SAT scores, and find that the mean of these 100 numbers is 489. Meanwhile, the registrar tells us that the mean of the math SAT scores of all entering freshmen is 507. The number 507 is what:

a. A population.
[✓ b.] A parameter.
c. A sample.
d. A statistic.
e. Both (a) and (b).
f. Both (c) and (d).
g. None of the above.

1B. The margin of error for a 99% confidence interval is _______ the margin of error for a 95% confidence interval (calculated from the same sample). (Circle the letter of the answer that best fills in the blank.)

a. smaller than
[✓ b.] larger than
c. equal to
d. any of the above are possible
e. none of the above

If you want to be more certain that the true value is in your interval, your interval needs to be larger.

1C. An analysis of a particular two variable data set yielded a correlation $r = 0.5$. If the roles of the explanatory and response variables are switched, the new correlation will be:

a. 2
[✓ b.] 0.5
c. -0.5
d. -2
e. Impossible to determine from the information provided.
f. None of the above.

1D. Let p be the true proportion of successes in a large population. (This is our standard notation.) For large samples (but no more than a tenth of the population), and assuming the other conditions for the one proportion z procedures are met, what is the sampling distribution of the sample proportion $\widehat{p}$?

a. Exactly normal with mean p and standard deviation $\sqrt{p(1-p)/n}$.
[✓ b.] Approximately normal with mean p and standard deviation $\sqrt{p(1-p)/n}$.
c. Approximately normal with mean $\sqrt{p(1-p)/n}$ and standard deviation p.
d. Highly nonnormal with mean $\sqrt{p(1-p)/n}$ and standard deviation p.
e. A t distribution with mean p and $\sqrt{p(1-p)/n}$ degrees of freedom.
f. A t distribution with mean $\sqrt{p(1-p)/n}$ and p degrees of freedom.

1E. A researcher finds the mass (in grams) and tail length (in centimeters) of each member of a sample of white mice, and statistically analyzes the data he gets. With mass being the explanatory variable, the units of the correlation of mass and tail length are:

a. Tails per mouse.
b. Centimeters per gram.
c. Grams per centimeter.
d. Grams.
e. Centimeters.
[✓ f.] None: the correlation has no units.
g. Cannot be determined from the information given.

1F. A laboratory tests 200 chemicals to determine whether they cause cancer in mice. For 11 of them, the laboratory obtains evidence that they cause cancer ($P < 0.05$ for each of these 11 chemicals). Which of the following is it safe to conclude?

[✓ a.] The reported P-values greatly overstate the significance of the results.
b. It is very unlikely that, if these 11 chemicals did not cause cancer, the observed effects of these 11 chemicals would be as extreme or more extreme than they actually were.
c. At least one of these 11 chemicals is likely to cause cancer, although most probably only have a small effect on cancer rates.
d. If these 11 chemicals do not cause cancer, the probability is at most 0.05 that some 11 of these chemicals would give results as extreme or more extreme than the results actually observed with these 11 chemicals.
e. These 11 chemicals all show strong evidence causing cancer, although their effects on cancer rates may be small and of little practical significance.

Under the situation described, even if none of the chemicals causes cancer, you *expect* roughly ten of them to show evidence of causing cancer at $P < 0.05$. That is, significance 0.05 means: if the null hypothesis is true, the outcome is sufficiently unlikely that in the long run it will occur on average one time in twenty. Here it has been given 200 chances to occur, so having it occur 10, 11, or even 12 times is not surprising.

1G. A statistical analysis was done in Salt Lake City which showed a strong positive correlation between the sale of sweaters and freeway car accidents. Which of the following is the most plausible explanation of this correlation?

a. A mistake, since there should be no correlation.
b. Wearing sweaters causes people to crash their cars.
c. Car accidents make people feel cold.
[✓ d.] The effect of a lurking variable, such as the presence of snow.
e. Scientific dishonesty: the statistician doing the study didn't have enough publications to get tenure at his university, and had to produce some more in a great hurry.
f. There is no good explanation, since these two things are unrelated.

1H. The diameters of the balls in balls bearings from Wang's Bearings Inc. vary according to a normal distribution with mean 2.5 mm and standard deviation 0.1 mm. What is the sampling distribution for the sample mean of a simple random sample of 25 of these balls?

a. $N(0.5, 0.1)$.
b. $N(0.5, 0.02)$.
c. $N(2.5, 0.5)$.
d. $N(2.5, 0.1)$.
[✓ e.] $N(2.5, 0.02)$.
f. $N(2.5, 0.004)$.
g. Impossible to determine from the information given.

1I. Gary owns 15 pieces of land. The mean value of the lots is \$30,000, the median value is \$20,000, the standard deviation of the values is \$12,000, and the range of the values is \$55,000. The total value of Gary's land is:

$\boxed{\text{a.}}$ \$450,000
b. \$300,000
c. \$180,000
d. \$825,000
e. Impossible to tell from the information given.
f. None of the above.

The mean is the total value divided by the number of lots, so the total value is (\$30,000)(15) = \$450,000.

1J. The P-value for a hypothesis test:

a. is the probability, assuming H_0 is true, of an outcome at most as extreme as the outcome we actually observed.
$\boxed{\text{b.}}$ is the probability, assuming H_0 is true, of an outcome at least as extreme as the outcome we actually observed.
c. is computed assuming that H_0 is false.
d. is always smaller than α.
e. All of the above.

1K. While talking about the cars that fellow students drive, Jill made the claim that 15% of the students drive white cars. Jack found this hard to believe and decided to check the validity of Jill's claim with a confidence interval. In his simple random sample of size 200 there were 17 white cars. The standard error of $\widehat{p}$ is approximately:

a. 0.0004
b. 0.0252
c. 0.0006
$\boxed{\text{d.}}$ 0.0197
e. None of the above.

The standard error is

$$\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}.$$

Here $n = 200$ and $\widehat{p} = \frac{17}{200}$. This gives a standard error of

$$\sqrt{\frac{\frac{17}{200}\left(1-\frac{17}{200}\right)}{200}} \approx 0.0197199.$$

2. (5 points/part; 20 points total) In the following situations, determine which one of the following statistical procedures is appropriate. Fill in each blank with the letter of the correct procedure. (Choices may be used more than once.)

Procedures:

A. one sample z test
B. one sample t test
C. matched pairs t test
D. two sample t test
E. one proportion z test
F. two proportion z test

Situations:

(1) __________ A researcher wishes to measure the effect of alcohol on reaction time. He measures the reaction times of 100 subjects with no alcohol, and the reaction times of the same subjects after two drinks.

Solution: C: matched pairs t test.

(2) __________ In order to estimate the mean high school GPA of Oregonians, you take a random sample of 500 high school graduates in Oregon and obtain their GPAs.

Solution: B: one sample t test.

(3) __________ Senator Snort's pollster conducts a telephone survey to gauge Senator Snort's support before he decides whether to seek reelection.

Solution: E: one proportion z test.

(4) __________ Senator Snort's opponent wants to know the effectiveness of a series of negative television ads featuring Senator Snort's recent conviction for drunk driving. His pollster asks in a telephone survey before the ads are run whether the respondents would vote for Senator Snort, and does the same thing, with an independent sample, after the ads have run.

Solution: F: two proportion z test.

3. a. (33 points. The steps do not all have the same point value.) A simple random sample of 714 adults (18 years or older) in a particular state showed that 535 of them voted in a particular election. A simple random sample of 634 adults in the same state showed that 437 of them voted in the next election.

Because of changes to the state voter registration laws, a researcher believes that the rates at which adults voted in the two elections are different. Test the researcher's belief at significance level $\alpha = 0.01$, using the following steps.

(1) State which test you will use (for example, one sample t or z procedure, or some other appropriate test from this course).
(2) Check that you can safely use the test in this case.
(3) State the hypotheses you will test.
(4) Calculate the test statistic.
(5) Give a P-value (or give two values between which P lies), and illustrate your answer with a graph.
(6) Draw and state the appropriate conclusion, expressing it in words appropriate for the context of the problem.

Solution:

(1) The appropriate procedure is the two proportion z procedure for hypothesis testing.

(2) We check whether this test it safe to use. First, we need the samples to be simple random samples. This is given. Second, we need the populations to be at least 10 times the sizes of the samples for both elections. Since even the smallest state well over 100,000 adult residents, this is clear. Third, we need there to be at least 5 successes and at least 5 failures in each sample. This is clearly met.

Showing that you checked these is part of showing your work.

(3) Let p_1 be the true proportion of adults who voted in the first election, and let p_2 be the true proportion of adults who voted in the second election. The hypotheses are

$$\begin{aligned} \mathrm{H}_0 &: p_1 = p_2. \\ \mathrm{H}_\mathrm{a} &: p_1 \neq p_2. \end{aligned}$$

(4) Since we want to know the probability of getting a result this extreme under the null hypothesis, we assume for the purpose of computing probabilities that the two samples are from the same population, and we combine them. (You don't have to include the above sentence in your answer; just do what it says.) The pooled sample proportion is

$$\widehat{p} = \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1 + n_2} = \frac{535 + 437}{714 + 634} \approx 0.72107$$

and the pooled sample size $n = n_1 + n_2 = 1348$. The standard error is

$$\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \approx \sqrt{(0.72107)(1-0.72107)\left(\frac{1}{714} + \frac{1}{634}\right)} \approx 0.024473,$$

and

$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx 2.4527.$$

(5) This gives a P-value of about 0.014179 (calculator) or $2 \cdot (1 - 0.9929) = 0.0142$ (from Table A).

No graph yet.

(6) We are requiring significance $\alpha = 0.01 < 0.014179$, so we cannot reject H_0. We cannot conclude that there is a difference between the proportions of adults who voted in the two elections.

Note: A correct solution must state the conclusion in terms of the proportions of adults who voted in the two elections. Just saying that we fail to reject the null hypothesis, or do not conclude that $p_1 \neq p_2$, will not get full credit.

b. (7 points.) Give at least one plausible reason why your results may not correctly represent the effect of the change in the laws on the rate at which adults vote in elections in this state.

> *Solution:* One election might have generated more interest among the voters than the other. (Possible reasons, which are not needed for a correct solution: controversial ballot measures, a close race, one was a primary and the other a general election, one involved candidates to more important offices than the other, one was a special election. Any one of these individually is also acceptable as a reason.)
>
> Most of the credit will be given for merely pointing out that there is no control group.
>
> Note: Reasons must be plausible. Little credit will be given for suggesting, for example, that there was a large increase in the senility rate among adults in the state.

4. (32 points. The steps do not all have the same point value.) An automobile company is testing a an engine modification to determine if it results in better gas mileage. It tests with 20 new cars, identical except that 10 of them incorporate the engine modification and 10 don't. Each automobile gets 10 gallons of gasoline and is then driven until it runs out of gasoline.

The cars without the engine modification went for a mean distance of 278 miles on the ten gallons of gas, with standard deviation 2.09 miles. The cars with the engine modification went for a mean distance of 282 miles on the tank of gas, with standard deviation 3.10 miles.

Is there significant evidence at the 0.01 level that the mean miles driven with the engine modification is larger than the mean miles driven without it? Assuming that the appropriate test is safe to use, test using the following steps.

(1) State which test you will use (for example, one sample t or z procedure, or some other appropriate test from this course).
(2) State the hypotheses you will test.
(3) Calculate the test statistic.
(4) Give a P-value (or give two values between which P lies), and illustrate your answer with a graph.
(5) Draw and state the appropriate conclusion, expressing it in words appropriate for the context of the problem.

Solution:

(1) Since we want to compare the two population means using two independent samples, we use the two sample t procedure.

(2) Let μ_1 be the mean miles that can be driven on 10 gallons of gasoline for cars of this type without the engine modification, and let μ_2 be the mean miles that can be driven on 10 gallons of gasoline for cars of this type with the engine modification.

$\text{H}_0\colon\ \mu_1 = \mu_2.$
$\text{H}_\text{a}\colon\ \mu_1 < \mu_2.$

(3) The test statistic is

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{278 - 282}{\sqrt{\frac{(2.09)^2}{10} + \frac{(3.10)^2}{10}}} \approx -3.3833.$$

(4) If we plan to use Table C with the conservative choice of the number of degrees of freedom, then we take the number of degrees of freedom to be $\min(n_1 - 1, n_2 - 1) = 9$. The number 3.3833 lies between 3.250, which corresponds to upper tail probability 0.005, and 3.690, which corresponds to upper tail probability 0.0025. Thus $0.0025 < P < 0.005$. (Since we are being conservative, we can't actually be sure that $P > 0.0025$.)

I also entered the data in my calculator. It gave 15.781 degrees of freedom, and $P \approx 0.0019271$.

No graph yet.

(5) Since $P < 0.01$, we reject the null hypothesis. We conclude at the significance level 0.01 that cars of this type with the engine modification on average go farther on 10 gallons of gasoline than those without the engine modification.

Note: A correct solution must state the conclusion in terms of the distance the cars can travel on 10 gallons of gasoline, or something equivalent (such as gas mileage). Just saying that we reject the null hypothesis, or conclude that $\mu_1 < \mu_2$, will not get full credit.

5. (8 points.) A researcher is studying the effect of a new headache relief medication. He administers it to a group of 50 subjects with headaches (which we may treat as a simple random sample), and reports that 37 of them experienced substantial relief. Is it appropriate and safe to use the one proportion z procedures to estimate the effectiveness of this medication? Why or why not? (Do **not** carry out the test.)

Solution: No. There is no control group.

6. (15 points.) Based on a simple random sample of 250 credit card holders, a large department store found that 12 card holders missed at least one payment during the last 6 months. Find a 90% confidence interval for the proportion of credit card holders who missed at least one payment during the last 6 months. Be sure to check that the procedure is safe to use.

Solution: We use the one sample z procedure for confidence intervals. We use the "plus four" method, since the large sample version is not safe in this case. (The requirement that there be at least 15 successes is violated.)

We check that it is safe to use. First, we need the sample to be a simple random sample. This is given. Second, we need the population to be at least 10 times the size of the sample. This isn't stated, but surely any department store with its own credit card has at least 2500 card holders, and probably a lot more than that. Third, we need confidence at least 90% (given) and sample size at least 10 (given).

Let p be the true proportion of card holders who missed at least one payment during the last 6 months. The z critical value for 90% confidence has upper tail probability $\frac{1}{2}(1 - 0.9) = 0.05$. The bottom line of Table C gives $z^* \approx 1.645$. The "plus four" method calls for the use of

$$\widetilde{p} = \frac{12 + 2}{250 + 4} \approx 0.055118.$$

Our confidence interval is then

$$\widetilde{p} \pm z^* \sqrt{\frac{\widetilde{p}(1 - \widetilde{p})}{n + 4}},$$

which is approximately

$$0.055118 \pm (1.645)\sqrt{\frac{0.055118(1 - 0.055118)}{254}} \approx 0.055118 \pm 0.023555,$$

or 5.51% ± 2.36%.

b. (9 points.) How large should the sample size be if we want a 98% confidence interval with a margin of error of 0.03 (or less)?

Solution: We will base this on the version of the one sample z confidence interval which uses $\widehat{p}$. We need a guessed value p^* for the true proportion, and we take the sample proportion $12/250 = 0.048$ from the data we already have. (Note that it is not between 0.3 and 0.7.) Futher let z^* be the critical value for 98% confidence, which by Table C is 2.326. Then we want

$$0.03 \geq z^* \sqrt{\frac{p^*(1 - p^*)}{n}}.$$

Solving for n (or using the formula

$$n = \left(\frac{z^*}{m}\right)^2 p^*(1 - p^*)$$

from the book), we find that we want

$$n \geq \left(\frac{z^*}{0.03}\right)^2 p^*(1 - p^*) \approx \left(\frac{2.326}{0.03}\right)^2 (0.048)(1 - 0.048) \approx 253.85.$$

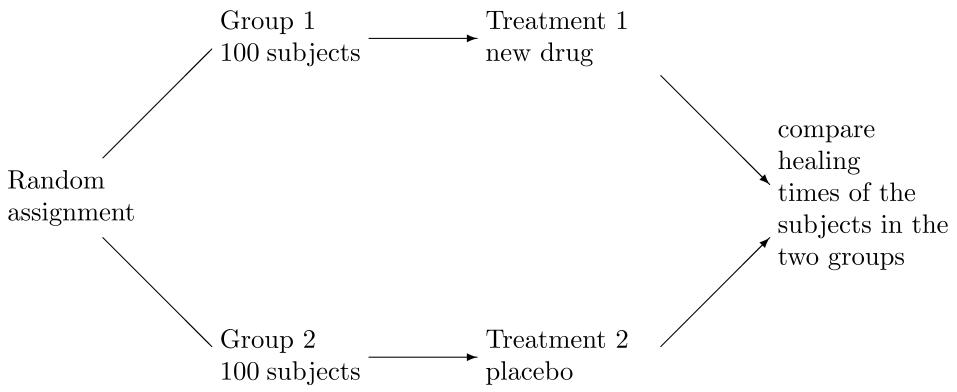
Since the sample size must be an integer, we take $n = 254$.

7. (7 points.) A researcher has been told that 8% of students at the University of Oregon are left handed. In a simple random sample of 100 University of Oregon students, 11 are found to be left handed. Is it appropriate and safe to use the one proportion z procedures on this outcome to carry out a hypothesis test on this claim? Why or why not? (Do **not** carry out the test.)

Solution: No. The sample size is $n = 100$. The null hypothesis is H_0: $p = 0.08$, so the proportion in the null hypothesis is $p_0 = 0.08$. One must have $np_0 \geq 10$, but here $np_0 = 8$.

8. (24 points) You want to determine whether a new drug (administered as a pill once per day) can speed the healing of broken bones in people. You have 200 test subjects available, all with similar bone fractures. Use a diagram to outline in detail the design of a randomized comparative double blind experiment. Include information about the treatment groups and the response variable. Be sure that one can tell from your description that your experiment has all the characteristics expected of such experiments.

Solution:



The response variable is the length of time it takes the subjects' bone fractures to heal. The assignment to the treatment groups is made randomly. The second treatment is a placebo, something which superficially is indistinguishable from the new drug but which does nothing. Until after the experiment is over and the healing times of the fractures are evaluated, neither the experimenters nor the subjects are to know who got the new drug and who got the placebo.

Extra credit. (20 extra credit points. Do not attempt this problem until you have checked all your answers to the other problems on this exam. It will be graded only if your score on the main part of the exam is 80% or better, and if your course grade without it is better than C^-. It will be graded much more severely; in particular, your answer must be clear, well written, and concise.)

Alice and Bob each have a bag containing a mixture of ordinary fair coins and coins with tails on both sides. Alice's bag contains 1024 ordinary coins and one coin with tails on both sides, and Bob's bag contains 512 coins of each kind. Each person chooses a coin at random from the corresponding bag. Alice flips her coin 10 times, and it comes up tails every time. Bob flips his coin 5 times, and it comes up tails every time. We want to understand the strength of the evidence that each of the coins has tails on both sides.

(1) (2 extra credit points.) State appropriate hypotheses.

Solution: For both:

H_0: The coin is fair.

H_a: The coin has tails on both sides.

(2) (6 extra credit points.) Give, with reasons, P-values for both outcomes. (The credit is all for the justification. In particular, no credit for just a number.)

Solution: The P-value is the probability of an outcome as extreme as what actually occurred, if the null hypothesis is true. Thus, for Alice it is the probability that one gets tails 10 times in a row if the coin is fair, which is

$$\frac{1}{2^{10}} = \frac{1}{1024} \approx 0.0009765.$$

Similarly, for Bob it is

$$\frac{1}{2^5} = \frac{1}{32} = 0.03125.$$

(3) (12 extra credit points.) Given the contents of the bags, for which of Alice and Bob is the evidence stronger that the coin has tails on both sides? Explain precisely the reasons for your conclusion.

Solution: Bob has stronger evidence. The explanation which follows is more detailed than is needed for the solution, but you must show that you understand the reasoning in it.

Imagine that the process is repeated a very large number of times. We ask, out of all the times that Alice's coin comes up tails on all 10 tosses, what fraction of the time does she have a coin with tails on both sides? We ask the same question for Bob, when his coin comes up tails all 5 tosses.

In both cases, we consider each possible outcome exactly once: which coin was chosen and what sequence of sides came up. For Alice, there are (1025)(1024) possibilities: the number of coins multiplied by the number of possible sequences of tosses. Exactly 2048 show tails on all 10 tosses: one for each of the 1024 for each fair coin, and all 1024 sequences for the single coin with tails on both sides. Thus, if we consider all the times that Alice gets tails on all 10 tosses, in $1024/2048 = \frac{1}{2}$ of them she has a fair coin and in the other half of them she has a coin with tails on both sides.

For Bob, there are (1024)(32) possibilities. Exactly (512)(1) come from fair coins and show tails on all 5 tosses (once for each fair coin), and exactly (512)(32) come from two tailed coins and show tails on all 5 tosses (every possible sequence of tosses for each such coin). Thus, if we consider all the times Bob gets tails on all 5 tosses,

$$\frac{(512)(32)}{(512)(1) + (512)(32)} = \frac{32}{33}$$

of the time, he has a coin with tails on both sides. Only $\frac{1}{33}$ of the time does he have a fair coin instead.