University of Toronto Scarborough
STAB22 Final Examination

April 2009

For this examination, you are allowed two handwritten letter-sized sheets of notes (both sides) prepared by you, a non-programmable, non-communicating calculator, and writing implements.

This question paper has 24 numbered pages. Before you start, check to see that you have all the pages. You should also have a Scantron sheet on which to enter your answers, and a set of statistical tables. If any of this is missing, speak to an invigilator.

This examination is multiple choice. Each question has equal weight, and there is no penalty for guessing. To ensure that you receive credit for your work on the exam, fill in the bubbles on the Scantron sheet for your correct student number (under "Identification"), your last name, and as much of your first name as fits.

Mark in each case the best answer out of the alternatives given (which means the numerically closest answer if the answer is a number and the answer you obtained is not given.)

If you need paper for rough work, use the back of the sheets of this question paper. The question paper will be collected at the end of the examination, but any writing on it will not be read or marked.

Before you begin, two more things:

- Check that the colour printed on your Scantron sheet matches the colour of your question paper. If it does not, get a new Scantron from an invigilator.
- Complete the signature sheet, but *sign it only when the invigilator collects it.* The signature sheet shows that you were present at the exam.

The correct alternative has an asterisk next to it.

1. A simple random sample of 70 coffee drinkers was taken. Each sampled coffee drinker was asked to taste two unmarked cups of coffee, one of which is actually Brand A and the other is Brand B, and was asked which one they preferred. 44 coffee drinkers preferred Brand A, and the other 26 preferred Brand B. Use this information to answer this question and the next one.

   The people who commissioned the survey are trying to find out whether a majority of all coffee-drinkers (that is, more than 50% of them) prefer brand A. Carry out a suitable test of significance to assess the evidence. What is the P-value of your test?

   (a) between 0.05 and 0.10
   (b) between 0.025 and 0.05
   (c) greater than 0.10
   (d) * between 0.01 and 0.025
   (e) less than 0.01

   Null $p = 0.50$, alternative $p > 0.50$; $\hat{p} = 44/70 = 0.629$; standard error, using the null value of p, 0.50, is $\sqrt{(0.50)(1-0.50)/70} = 0.060$, so $z = (0.629 - 0.5)/0.060 = 2.15$. One-sided P-value is 0.0158.

2. In the situation of Question 1, it turned out that the coffee drinkers had always been given an unmarked cup containing Brand A coffee first, and Brand B coffee second. How do you react to this knowledge?

   (a) The P-value was small, so there will still be good evidence that Brand A is preferred.
   (b) The coffee drinkers received their coffee in unmarked cups, so it doesn't matter which brand is actually tasted first.
   (c) Brand A must have an advantage by being tasted first, so there cannot be a significant difference between brands A and B.
   (d) The P-value was large, so there will still be no evidence that Brand A is preferred.
   (e) * A better approach would have been to toss a coin to decide whether each drinker gets Brand A first or Brand B first.

   If the test gives a small P-value (as it did), this could be because brand A really is preferred, or because the brand given first, whichever it was, is preferred, and we have no way of telling which is the reason for the small P-value. So (a) and (c) could be true, or could be false. (d) was actually false anyway, but if you thought the P-value from the previous question was large, (d) here could also be either true or false. (b) is a red herring: the fact that the cup is unmarked doesn't stop Brand A from being tasted first, so if being first is the reason for the difference, whether the cup is unmarked or not will not matter. That leaves (e), and we already know that randomization is a good thing.

3. A university financial aid office took a simple random sample of students to see how many of them were employed the previous summer. Of the 750 men sampled, 703 had been employed the previous summer; of the 650 women sampled, 592 had been employed the previous summer. Use this information for this question and the next one.

   Test whether the proportion of all male students employed last summer is different from the proportion of all female students employed last summer. What is the P-value of your test?

   (a) * between 0.05 and 0.10
   (b) between 0.01 and 0.025
   (c) between 0.025 and 0.05
   (d) less than 0.01

(e) greater than 0.05

Two-proportion test from §8.2, so go through the procedure, using a two-sided alternative. $\hat{p}_1 = 0.9373, \hat{p}_2 = 0.9108$; $\hat{p} = (703 + 592)/(750 + 650) = 0.9250$;

$$SE_{Dp} = \sqrt{(0.9250)(0.0750)(1/750 + 1/650)} = 0.0141;$$

$z = (0.9373 - 0.9108)/0.0141 = 1.88$; P-value $2 \times (1 - 0.9699) = 0.0602$. (e) was supposed to say "greater than 0.10"; with the question as written, (e) is also correct.

4. From the information given in Question 3, what is the upper limit of a 95% confidence interval for the difference between the proportion of men employed the previous summer and the proportion of women employed the previous summer? (Take the difference as men minus women.)

(a) * 0.05

(b) -0.10

(c) 0.10

(d) 0.00

(e) -0.05

Section 8.2 again, only this time using SE_D because it's a CI:

$$SE_D = \sqrt{(0.9373)(0.0627)/750 + (0.9108)(0.0892)/650} = 0.0143,$$

so the upper limit is $0.9373 - 0.9108 + 1.96(0.0143) = 0.0545$. (The lower limit is close to 0.)

5. A report states that 3% of all births are "multiple births" (that is, twins, triplets, etc.). A random sample of 5000 births is taken. Assuming the report to be correct, what is the probability that 2.5% or less of the births in the sample are multiple births?

(a) * less than 0.02

(b) between 0.05 and 0.10

(c) greater than 0.30

(d) between 0.02 and 0.05

(e) between 0.10 and 0.30

Normal approximation to binomial: for $\hat{p}$, mean is $p = 0.03$ and SD is $\sqrt{(0.03)(0.97)/5000} = 0.0024$, so $z = (0.025 - 0.03)/0.0024 = -2.07$, and the probability of less is 0.0191.

6. A population has a mildly non-normal shape with mean 35 and standard deviation 7.5. A sample of size 225 is taken from this population. What is the probability that the sample mean is less than 33.5?

(a) * about 0.001

(b) 0.05

(c) 0.42

(d) more than 0.90

(e) 0.12

Asking about sample mean, so use method of §5.2: $z = (33.5 - 35)/(7.5/\sqrt{225}) = -3$, so the probability of less is 0.0013. Don't forget to divide by $\sqrt{225}$!

The mild nonnormality doesn't affect the calculation. The sample size is so large that the approximation from the central limit theorem will be very good.

7. A simple random sample of 30 observations is taken from some population. The sample mean is 120. You wish to test the null hypothesis $\mu = 116$ against the alternative $\mu \neq 116$, but you have only the output below.

```
 N     Mean  SE Mean          95% CI
30  120.000    1.826  (116.422, 123.578)

 N     Mean  SE Mean          99% CI
30  120.000    1.826  (115.297, 124.703)
```

What can you say about the P-value of your test?

(a) It is less than 0.01.
(b) It is greater than 0.95.
(c) It is less than 0.05.
(d) It is greater than 0.05.
(e) * It is less than 0.05 but greater than 0.01.

116 is outside the 95% confidence interval, so the two-sided P-value would be less than 0.05 (an implausible value); 116 is inside the 99% interval, so the P-value is greater than 0.01 (plausible).

8. It is desired to estimate the mean of a population using a 95% confidence interval. The population is known to have a highly skewed shape. A sample of size 45 is proposed. What do you think of this choice of sample size?

(a) * the sample size may not be large enough to allow the Central Limit Theorem to apply
(b) the sample size is bigger than 30, so the calculation should be accurate
(c) the sampling distribution of the sample mean has exactly a normal shape regardless of the shape of the population
(d) the normal distribution will be an accurate approximation to the t distribution in this case
(e) The Law of Large Numbers says that the sample mean and population mean will be close, so there is no need to calculate a confidence interval

Don't get hung up on a "magic" sample size of 30: larger is better, and if the population is highly skewed, larger is *necessary* for a normal approximation to be good. So (b) and (d) are wrong. The Central Limit Theorem is only an approximation, never the exact truth, so (c) is no good (replace "exactly" by "approximately" and you would be all right). In (e), the Law of Large Numbers does indeed say they will be close, but not *how* close; the normal approx, if it applies, does say how close. That leaves (a), which is right on the money.

9. A student's mark on a Psychology exam has a normal distribution with mean 65 and SD 10. The same student's mark on a Chemistry exam has a normal distribution with mean 60 and SD 15. The two marks are independent of each other. Use this information for this question and the next two.

What is the probability that the student scores over 80 on both exams?

(a) between 0.01 and 0.10
(b) between 0.20 and 0.50
(c) between 0.10 and 0.20
(d) * less than 0.01

(e) greater than 0.50

Two parts to this: use the normal distribution to find the chance of the student getting 80 on each exam, and then use the laws of probability to figure out "both".

On the psychology exam, a score of 80 goes with a z of $(80-65)/10 = 1.5$, so, from Table A, the probability of a score over 80 is 0.0668. Likewise, on the chemistry exam, $z = (80-60)/15 = 1.33$, and a prob of 0.0912. For independent events (as here), the probability of both is gotten by multiplying each probability together: $(0.0668)(0.0912) = 0.0061$.

You can guess that the answer would be pretty small because getting over 80 on even one of the exams would be hard enough without having to get over 80 on both.

10. What is the probability that the student scores over 80 on exactly one of the two exams?

(a) between 0.01 and 0.10

(b) between 0.20 and 0.50

(c) * between 0.10 and 0.20

(d) less than 0.01

(e) greater than 0.50

Use 0.0668 and 0.0912 from the previous question. The student has to get (over 80 on psych and not on chem) or (less than 80 on psych and over 80 on chem). Figure out the probability of each one by multiplying, then add the results: $0.0668(1-0.0912) + (1-0.0668)(0.0912) = 0.1458$.
You can't use the binomial here because the chances of getting over 80 on the two exams are different (if they were the same, you could).

11. What is the probability that the student's mean mark in the two courses is over 80?

(a) between 0.10 and 0.20

(b) greater than 0.50

(c) * between 0.01 and 0.10

(d) between 0.20 and 0.50

(e) less than 0.01

The total score has mean $65 + 60 = 125$ and SD $\sqrt{10^2 + 15^2} = 18.03$ (remember rules for variances and SDs). For the student's mean mark to be above 80, the total on the two exams has to be above 160, so $z = (160-125)/18.03 = 1.94$, and the probability of being above this is 0.0261. (It's easier to get a mean score above 80 than it is to get over 80 on both exams because a high enough score on one exam will allow you to have a score below 80 on the other and still get an average above 80, such as 84 and 78.)

12. The time it takes (for any student) to complete a STAB22 final exam is a random variable having a normal distribution with mean 160 minutes and standard deviation 15 minutes. Use this information for this question and the next one.

Anne and Bob are two friends writing this exam. What is the probability that Anne will complete the exam at least ten minutes before Bob completes the exam? (Assume that all students start the exam at the same time and that their completion times are independent.)

(a) this probability is greater than 0.45 but less than 0.55

(b) this probability is less than 0.01

(c) this probability is greater than 0.35 but less than 0.45

(d) * this probability is greater than 0.01 but less than 0.35

(e) this probability is greater than 0.55

This requires careful thought. Let A be the time Anne takes, and B be the time Bob takes. We want the probability that $A - B \leq -10$ (a lower time is better). A and B are both normally distributed, so their difference is as well, with a mean of $160 - 160 = 0$ minutes (difference in means) and an SD of $\sqrt{15^2 + 15^2} = 21.2$ minutes (variance of difference is sum of variances).

To get the probability that the difference is less than -10, find $z = (-10 - 0)/21.2 = -0.47$, and a probability of 0.3192.

(You can guess that the answer will be less than 0.5 because Anne and Bob take the same time on average and half the time Anne will be quicker than Bob, and only some of that time will Anne be as much as 10 minutes quicker. Because the SD is quite large, Anne will be at least 10 minutes quicker more of the time than you might expect.)

13. Using the information given in Question 12, if 300 students are writing the exam, what is the approximate probability that more than 275 students will complete the exam in less than 180 minutes? Choose the closest value from the options below.

(a) 0.75
(b) 0.00
(c) 1.00
(d) 0.50
(e) * 0.25

This one is binomial, with $n = 300$ students. We want more than 275 students to "succeed" (complete the exam in less than 180 minutes), but we don't know the probability of success, yet. There is enough information to figure it out, though: completion time has a normal distribution with mean 160 and SD 15 minutes. So for completing in less than 180 minutes, $z = (180 - 160)/15 = 1.33$, with a probability of 0.9082, which we use as p below.

This binomial problem has a large n, so we'll need to use the normal approx to the binomial. You can check that the rule of thumb is all right, since $300(1 - 0.9082)$ is about 30. The mean number of successes is $300(0.9082) = 272.46$, and the sd of the number of successes is $\sqrt{300(0.9082)(1 - 0.9082)} = 5.001$. So z for 275 successes is $(275 - 272.46)/5.001 = 0.51$. The probability of more students than this completing the exam in time is $1 - 0.6950 = 0.3050$, so 0.25 is the answer to choose. (You might guess this, since 275 is a bit above the mean, so the answer should be a bit less than 0.50, but not as small as 0, for which we would need to be asking about 290 students completing the exam in time.)

14. In a certain city, the probability that a call to a randomly-chosen telephone number will reach a live person is 0.3. Suppose 10 such calls are made. What is the probability that at least one of them reaches a live person?

(a) less than 0.10
(b) between 0.50 and 0.90
(c) * greater than 0.90
(d) between 0.30 and 0.50
(e) between 0.10 and 0.30

This is a typical "at least one" problem: find the probability of none, and then subtract that from 1. The probability that a single call fails to reach a live person is $1 - 0.3 = 0.7$, so the probability that none of the 10 calls reach a live person is $(1 - 0.3)^{10} = 0.0282$ (using the multiplication rule because of the independence implied by the random choice of phone numbers). Thus the probability that at least one does is $1 - 0.0282 = 0.9718$. (Even though the chance of reaching a live person is pretty small when you only call once, making 10 calls gives you a decent chance of catching *somebody*; the mean number of live people you encounter is 3.)

15. An experiment was conducted to compare a new drug with a standard drug with respect to mean recovery time. It is known that the mean recovery time for the standard drug is 26 days. Following the experiment, which involved 65 patients, a 95% confidence interval estimate was constructed for the mean recovery time (in days) for patients on the new drug. The 95% confidence interval was found to be from 24.6 to 27.8. Which of the following statements is a valid conclusion based on this information?

    (a) The experimenter should reject the claim that the new drug is the same as the standard drug with respect to mean recovery time.

    (b) There is evidence, at the 5% level of significance, that the new drug is better than the standard drug.

    (c) * There is insufficient evidence, at the 1% level of significance, to reject the claim that there is no difference between the new drug and the standard drug with respect to mean recovery time.

    (d) There is sufficient evidence, at the 5% level of significance, to reject the claim that there is no difference between the new drug and the standard drug with respect to mean recovery time.

    (e) None of the other statements is a valid conclusion based on the information from this study.

    The confidence interval contains 26, the mean recovery time for the standard drug, so there is not evidence (at the 0.05 level, at least) that the new drug is different from the older one. This rules out (a), (b) and (d). If the results are not significant at the 0.05 level, the P-value is bigger than 0.05, which means it must be bigger than 0.01 as well, and so the results are not significant at the 0.01 level either. This means that (c) is true. (Assessment of significance from a confidence interval is appropriate for a *two-sided* test, so that, if 26 had been outside the confidence interval, we would have had to think more carefully about whether (b) is true.)

16. Following the analysis of a well-designed completely randomized experiment it was reported that the observed effect was "statistically significant". Which of the following statements best explains the meaning of the phrase "statistically significant"?

    (a) The observed result made sense to the experimenter since it was what was hoped would happen.

    (b) The observed effect happened because the experiment was properly designed and carried out without bias.

    (c) The experimenter carefully employed the basic principles of experimental design in conducting the study.

    (d) * The observed effect was sufficiently large so that it would rarely occur simply by chance.

    (e) The laws of probability say that this observed result would be expected to happen by chance.

    "Statistically significant" means that the result was unlikely to happen by chance if there was actually no effect (and therefore we conclude that there *is* an effect). This is what (d) says. (b) and (c) are true but irrelevant: a properly designed experiment is a given, before you can start talking about statistical significance. (a) is not how science works (or should work), and (e) is missing a "not".

17. During a recent year, the average cost of making a movie was \$58 million. A simple random sample of 15 action movies was taken during that year; the sampled action movies cost a mean of \$62.5 million to make, with a sample standard deviation of \$9.5 million. Is this evidence that action movies cost more to make, on average, than movies in general? Use $\alpha = 0.05$.

(a) * The P-value is small, so we can conclude that action movies cost more to make on average.

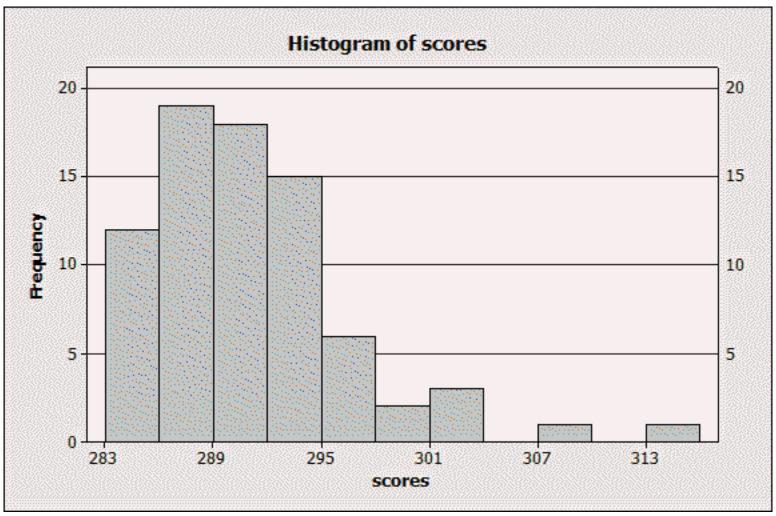
(b) The P-value is small, so we have evidence that action movies cost the same to make as movies generally.

(c) The P-value is not small, so we cannot conclude that action movies cost more to make on average.

(d) The P-value is not small, so we have evidence that action movies cost the same to make as movies generally.

We only have a sample SD, so we have to use a t-test, one-sided since we are looking for evidence of "more". The test statistic is $t = (62.5 - 58)/(9.5/\sqrt{14}) = 1.83$, with $14 - 1 = 13$ df. In Table D, the P-value is between 0.025 and 0.05, so we *do* have evidence at $\alpha = 0.05$ that the action movies cost more to make on average, which is (a). (b) and (d) are wrong because we don't have evidence in favour of "the same", only against, and in (c) the P-value *is* smaller than 0.05.

18. The histogram below shows the scores of 77 players in a golf tournament. Use the histogram for this question and the one following.



What is the median score for these golfers?

(a) * between 289 and 292

(b) between 286 and 289

(c) cannot tell from a histogram

(d) between 289 and 292

(e) greater than 292

The median score is the $(77 + 1)/2 = 39$th smallest (or largest) one. 12 of the golfers scored between 283 and 286, 19 scored between 286 and 289, and 18 scored between 289 and 292. (You might be off by one reading the frequencies from the scale, but that doesn't matter.) So $12 + 19 = 31$ of the golfers scored less than 289; this is less than 39 of them, so we need to take the next group: $12 + 19 + 18 = 49$ which is more than 39, so the median is in the last group we added, between 289 and 292.

19. From the histogram in Question 18, how does the mean score compare with the median?

(a) cannot estimate mean from a histogram

(b) the mean score is about the same as the median

(c) * the mean score is bigger than the median

(d) the mean score is smaller than the median

The shape is skewed to the right, so the mean should be bigger than the median (the mean is pulled into the long tail). (a) is true, but you can do better: even though you can't say what the mean is, you do know that it's bigger than the median.

20. The distribution of the heights of students in a large class is approximately normal. Moreover, the average height is 68 inches, and approximately 16% of the students were taller than 71 inches. Based on the 68-95-99.7 rule, approximately what percentage of students in this class are taller than 65 inches?

(a) 2.5%
(b) * 84%
(c) 5%
(d) 16%
(e) 95%

If 16% are taller than 71 inches, and the mean is 68 inches, there should also be 16% shorter than 65 inches, because the normal distribution is symmetric about the mean. If 16% are shorter, $100 - 16 = 84\%$ are taller.

21. A company producing a certain item has two machines: Machine A and Machine B. Machine A is an old machine and 10% of the items produced by this machine are defectives. Machine B is a new machine and 5% of the items produced by this machine are defectives. A random sample of two items was selected from each machine for quality control purposes. What is the probability that the sample from machine B will have MORE defective items than the sample from machine A? Choose the closest answer from the options below.

(a) 0.15
(b) 0.18
(c) 0.25
(d) 0.01
(e) * 0.08

Think of all the ways in which machine B can have more defective items than machine A, work out the probability of each, and add them up. You can take advantage of the random sampling to see that the number of defective items produced by the two machines are independent.

Since only two items are sampled from each machine, B will give more defectives than A if B gives 2 and A gives 1 or 0, or if B gives 1 and A gives 0. The number of defective items produced by each machine has a binomial distribution with $n = 2$ and $p = 0.10$ or 0.05 (or you can figure out the probabilities directly), giving, from Table C, these results:

Machine	0 defectives	1 defective	2 defectives
A	0.8100	0.1800	0.0100
B	0.9025	0.0950	0.0025

This means that the answer we want is

$$(0.0025)(0.8100 + 0.1800) + (0.0950)(0.8100) = 0.002475 + 0.0769,$$

which is about 0.08. Since defectives are rare, much the most likely way to have B giving more defectives than A is for B to give 1 and A to give none.

22. In an experiment, the soles of boys' shoes are made of two different synthetic materials, A and B. To see which material is better (lasts longer), measurements were made on the amount of wear of the soles of shoes worn by 10 boys. Each boy wore a special pair of shoes - the left sole was made with material A and the right sole with material B.

    Based on this information, which one of the following statements is true?

    (a) The decision as to whether the left or the right sole is to be made with material A or B should be determined by using a table of the standard normal distribution.
    (b) The type of material is the response variable, with two levels, A and B.
    (c) * None of the above other statements is true.
    (d) This is an example of a completely randomized design.
    (e) Increasing the sample size (number of replications) is one way to eliminate the need for any randomization in assigning the materials.

    All of the statements apart from (c) are false: in (a), a table of random digits ought to be used (or you could toss a coin). In (b), the response is how long the sole lasts; the material type is the factor (or is an explanatory variable). This is a matched pairs design (the two materials are compared on the same boy), so (d) is false, and in (e), randomization is always required, regardless of the sample size.

23. The amount of money spent by a customer at a discount store is a random variable with mean \$100 and standard deviation \$30. What is the approximate probability that a randomly selected group of 50 shoppers will spend a total of more than \$5300? Choose the number closest to the answer from the options given below.

    (a) 0.20
    (b) *0.10
    (c) 0.50
    (d) 0.30
    (e) 0.40

    The easiest way to tackle this is via the sample mean (and using the ideas of §5.2). If the $n = 50$ shoppers spend a total of more than \$5300, the sample mean is bigger than $5300/50 = 106$. According to the Central Limit Theorem, the sampling distribution of the sample mean will be approximately normal with mean 100 and SD $30/\sqrt{50} = 4.243$ (remember, we are talking about a sample mean, so divide by $\sqrt{n}$). To find the prob of the sample mean being bigger than 106, find $z = (106 - 100)/(30/\sqrt{50}) = 1.41$, and the prob of being more than this is $1 - 0.9207 = 0.0793$, which is closer to 0.10 than any of the other options.

    Or you can work directly with the total, using the rules for means and SDs from Chapter 4. The sample total has mean $50(100) = 5000$, and variance $50(30^2) = 45000$, so the SD of the sample total is $\sqrt{45000} = 212.13$. Again we have to use the Central Limit Theorem to claim that the sampling distribution of the sample total is approximately normal (the CLT works for means *and* totals). For the sample total being bigger than 5300, $z = (5300 - 5000)/212.13 = 1.41$, which is the same z as above, and therefore the probability is the same, as you would guess.

24. According to a report, the mean salary for mayors in all Canadian cities is \$120,000. A newspaper believes that this figure is not correct, and wishes to assess the evidence against it. A reporter at the newspaper takes a simple random sample of 10 Canadian cities, and finds that the mean salary for mayors in those cities is \$132,200. Use this information for this question and the following one.

    What would a suitable *null hypothesis* for a test of significance say?

(a) The sample mean salary is equal to $132,200.
(b) * The mean salary of all mayors is equal to $120,000.
(c) The mean salary of all mayors is not equal to $120,000.
(d) The sample mean salary is not equal to $132,200.
(e) The mean salary of all mayors is greater than $120,000.

A null hypothesis makes a claim that the population mean is equal to some value, which rules out (c), (d) and (e). The value in (a) is the *sample* mean, so the right answer must be (b).

25. Using the information in Question 24, what would a suitable *alternative hypothesis* for a test of significance say?

(a) * The mean salary of all mayors is not equal to $120,000.
(b) The sample mean salary is not equal to $132,000.
(c) The mean salary of all mayors is equal to $120,000.
(d) The sample mean salary is equal to $132,200.
(e) The mean salary of all mayors is greater than $120,000.

An alternative hypothesis makes the claim that the population mean is in some way different from some value. That rules out (c) and (d). (b) uses the sample mean. The newspaper believes only that the null-hypothesis $120,000 is "not correct", so a two-sided test is called for, as in (a). (e) would be correct if we were doing a particular one-sided test.

26. A fair 10-sided die has faces numbered from 1 to 10. The probability of rolling a 1 on any roll is 0.1. If the die is rolled 12 times, what is the value of r such that the probability of rolling r 1's or less is 0.8891?

(a) there is no value of r that works
(b) 2.5
(c) * 2
(d) 5
(e) 1

The number of 1's rolled has a binomial distribution with $n = 12$ and $p = 0.1$ (12 independent rolls, and the chance of rolling a 1 is always the same). So look at Table C and cast your eye down the $n = 12, p = 0.1$ section until you have gone far enough to make the probabilities add up to 0.8891 (since you want r *or less*). The first two add up to 0.6590, and the first three add up to 0.8891, so you need to go as far as 2 (k in the table).

In the days before computers were used for role-playing games, people used to have real 10-sided dice. They were 10-sided prisms: if you looked at them end-on, you saw a regular 10-sided decagon, but if you looked at them side-on, they looked just like a rectangle. You rolled them sideways, and got the score of the face that was (horizontal and) uppermost.

27. In a population, 45% of people have type O blood, 40% have type A, 11% have type B, and 4% have type AB. Use this information for this question and the next one.

Consider an accident victim with type B blood. She can only receive a transfusion from a person with type B or type O blood. There are three people from this population who are willing to donate blood, but they don't know what their blood type is. Assuming that the blood types of the three individuals are independent, what is the probability that at least one of them will be a suitable donor? Choose the closest answer from the following options.

(a) * 0.9
(b) 0.6
(c) 0.7
(d) 0.8
(e) 0.5

If there were just one person donating blood, that person could be type B or type O, which are disjoint outcomes, so the probability of that person being a suitable donor is $0.45+0.11 = 0.56$. But there are three willing donors, so the chance of at least one of them being a match will be bigger than that.

This is your standard "at least one": work out the chance that *none* of the donors will be suitable, and then take one minus that. Here, the prob of a single donor not being suitable is $1 - 0.56 = 0.44$, so the prob of all three failing to be suitable is $0.44^3 = 0.0852$, and therefore the chance of at least one (one or more) of them being suitable is $1 - 0.0852 = 0.9148$. (The more opportunities you have for success, the greater the chance of at least one success.)

28. Using the distribution of blood types given in Question 27, what is the probability that both people in a couple will have the SAME blood type if we assume that one partner's blood type does not influence the blood type of the other partner? Choose the closest answer from the following options.

(a) 0.1
(b) 0.3
(c) 0.2
(d) 0.5
(e) * 0.4

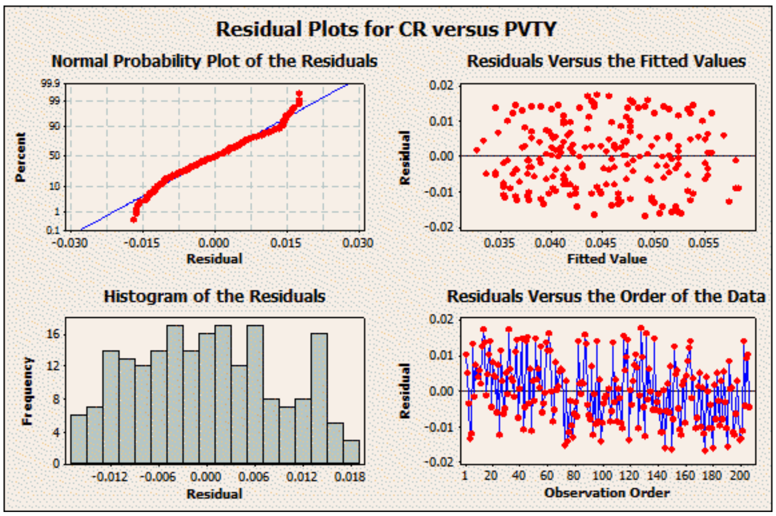
The two people could be: both O's, both A's, both B's, both AB's. Nothing else will work. Work out the probabilities of each of these four events by multiplying (independence) and then add the results up (disjoint). This gives $0.45^2 + 0.40^2 + 0.11^2 + 0.04^2 = 0.3762$.

29. A researcher examined the relationship between crimes and other demographic variables for 206 of the most populous counties in the US. The MINITAB output below is a part of the statistical analysis of this data. The variables in the output are:

CR	Crime rate (Number of crimes divided by total population)
PVTY	percent below poverty level
HSC	percent high school graduates

```
Correlations: CR, PVTY, HSC

           CR    PVTY
PVTY   0.573
       0.000

HSC   -0.293  -0.572
       0.000   0.000


Cell Contents: Pearson correlation
               P-Value

Regression Analysis: CR versus PVTY
```

```
The regression equation is
CR = 0.0294 + 0.00235 PVTY

Predictor        Coef    SE Coef      T      P
Constant     0.029387   0.001670  17.59  0.000
PVTY        0.0023498  0.0002356   9.97  0.000


S = 0.00896965   R-Sq = 32.8%   R-Sq(adj) = 32.5%
```




```
Regression Analysis: CR versus HSC

The regression equation is
CR = 0.0952 - 0.000630 HSC


Predictor         Coef    SE Coef      T      P
Constant       0.09521    0.01155   8.24  0.000
HSC         -0.0006303  0.0001442  -4.37  0.000


S = 0.0104614   R-Sq = omitted  R-Sq(adj) = omitted
```

Use this information for this question and the next one.

Which of the following statements is true?

(a) The linear regression of CR on PVTY explains more than 50% of the variability in CR.

(b) * Using the regression of CR on PVTY, the residual for an observation with PVTY = 3.7 and CR = 0.05 is greater than 0.005.

(c) The linear regression of CR on HSC explains more than 25% of the variability in CR.

(d) None of the other four statements is true.

(e) The MINITAB output for the regression analysis of CR versus PVTY above indicates the need for a higher order (curved) model rather than a linear model.

Judge each of the four statements individually. You can stop as soon as you reach a true one, though it's probably sensible to check the others as well in case you made a mistake.

In (a), the percent of variability explained by the regression is R-squared, the square of the correlation. Here, the correlation between CR and PVTY is 0.573, and the square of this is less than 0.50, so (a) is false.

In (b), the regression equation for predicting CR from PVTY is $0.0294 + 0.00235 PVTY$, so when PVTY is 3.7, the predicted CR is $0.0294 + (0.00235)(3.7) = 0.0381$. Residual is observed minus predicted, $0.05 - 0.0381 = 0.0119$, which is less than 0.05, so this statement is true. If you're pressed for time, go ahead and mark B and move on to the next question, otherwise...

(c) is the same idea as (a). Find the correlation between CR and HSC, which is -0.293, and square it to get 0.086. This is safely less than 0.25, so the statement is false.

For (e), look for any evidence that a curve would be better. The best place to look is the plot of residuals vs. fitted values (which you interpret the same way as a plot of residuals vs. x): there is no discernible pattern at all, let alone a curved pattern, so there is no evidence that a curve would be better.

In short, (b) is true and the other statements are false.

30. Using the information in Question 29 above, calculate the slope of the regression line for predicting PVTY from HSC). Choose the closest value from the options below.

    (a) -0.5
    (b) * -0.3
    (c) -0.2
    (d) -0.4
    (e) -0.1

    The slope of the line for predicting PVTY from HSC is rSD_{PVTY}/SD_{HSC}, where r is the correlation between PVTY and HSC (-0.572). We don't know the standard deviations, but we do know two other slopes: predicting CR from PVTY and predicting CR from HSC. So we know that:

$$\begin{aligned} 0.00235 &= 0.573(SD_{CR}/SD_{PVTY}) \\ -0.000630 &= -0.293(SD_{CR}/SD_{HSC}) \end{aligned}$$

    If we knew SD_{PVTY}/SD_{HSC}, we'd be able to find the slope we want. A little thought reveals that dividing the second equation by the first will get that (because SD_{CR} will cancel, and everything else is numbers). So we get

$$(-0.000630/0.00235) = (-0.293/0.573)(SD_{PVTY}/SD_{HSC})$$

    (careful with the zeroes!) and rearranging,

$$SD_{PVTY}/SD_{HSC} = (-0.000630/0.00235)/(-0.293/0.573) = 0.5243.$$

    This is correctly positive since it is two positive things divided by each other. The last step is to mutiply by the correlation: the slope is $(-0.572)(0.5243) = -0.2999$.

    Yes, this was a difficult question.

31. A car magazine commissions a study to determine whether single drivers do more driving for pleasure than married drivers. The magazine collects samples of 35 single drivers and 35 married drivers, and records how many kilometres per week of pleasure driving each driver does. Some Minitab output from an analysis is shown below:

```
Two-Sample T-Test and CI: single, married

Two-sample T for single vs married

          N   Mean  StDev  SE Mean
single   35  193.3   27.0      4.6
married  35  190.4   24.7      4.2


Difference = mu (single) - mu (married)
Estimate for difference:  2.82857
95% CI for difference:  (-9.50924, 15.16639)
T-Test of difference = 0 (vs not =): T-Value = 0.46  P-Value = 0.649  DF = 67
```

Using the output as given, what conclusion should the car magazine draw from its study, using $\alpha = 0.05$ if necessary?

(a) There is evidence that the amount of pleasure driving done by single and married people is different on average.

(b) * There is no evidence that single people do more pleasure driving than married people, because the P-value is about 0.325.

(c) There is evidence that single people do more pleasure driving than married people, on average.

(d) Single people do about 15 kilometres per week of pleasure driving more than married people do.

(e) There is no evidence that single people do more pleasure driving than married people, because the P-value is about 0.649.

As with any test, go hunting for the P-value, here 0.649, which is not small by any standards. So there is no evidence that the mean amount of pleasure driving done by single and married people is different. The test done was two-sided, but even if it had been one-sided, the P-value would have been $0.649/2 = 0.325$ which is far from small either. That rules out (a) and (c). (d) is looking for the *middle* of the confidence interval (difference in sample means), but 15 is the upper end, so that's no good.

That leaves a choice between (b) and (e). Looking for evidence of "more" implies a one-sided test, but the P-value in (e) comes from a two-sided test. The discussion in the previous paragraph says that the right P-value for a one-sided test would be about 0.325, so (b) is the answer.

32. A carnival game offers a \$100 cash prize for anyone who can break a balloon by throwing a dart at it. It costs \$15 for each throw. John is willing to spend up to \$45 trying to win. That means he is going to stop if he breaks the balloon on or before the 3rd throw and stop at the 3rd throw even if he failed to break it on the 3rd throw.

The chance that John will break the balloon in any throw is 0.1. Assume that breaking the balloon on any attempt is independent of what happens on the other attempts.

John's net gain is the amount of money he won minus the total cost. Denote this net gain by X. Which of the following numbers is the closest to the mean of X?

(a) \$20

(b) \$10

(c) \$0

(d) $-$\$20

(e) * $-$\$10

If John breaks the balloon on the first throw, he will win $100 - 15 = 85$ (the prize minus the cost of making one throw. The probability of doing this is 0.1.

If John breaks the balloon on the second throw (which means that he missed the first time), he will win $100 - 2(15) = 70$, with probability $(0.9)(0.1) = 0.09$.

If he breaks the balloon on the third throw, having missed the first two times, he wins $100 - 3(15) = 55$ with probability $(0.9)^2(0.1) = 0.081$.

The rest of the time, he misses with all three throws and then quits. In this case, he loses $3(15) = 45$, that is, wins $-\$45$. The probability of this is $1 - 0.1 - 0.09 - 0.081 = 0.729$.

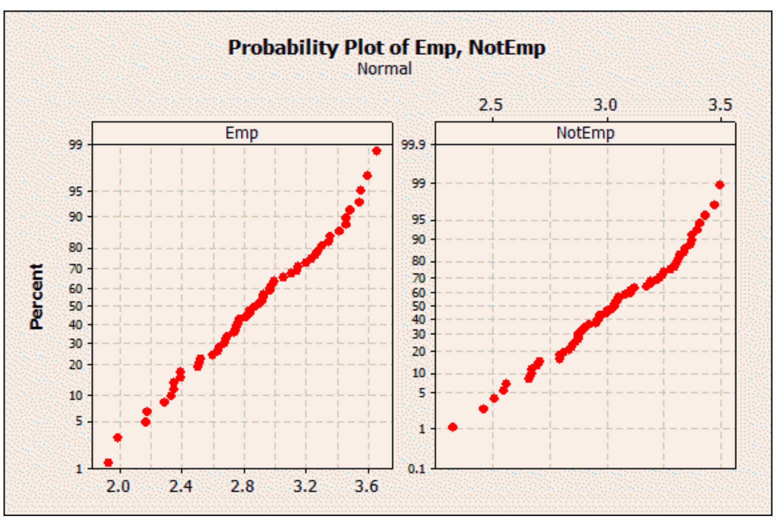
Then go through and multiply winnings by probabilities to get a mean for X of

$$85(0.1) + 70(0.09) + 55(0.081) - 45(0.729) = -13.55.$$

Even though there are 3 different ways to win, the most likely outcome is that he will lose, and therefore the mean comes out negative.

33. In a study of the effects of college student employment on academic performance, the researchers analyzed the GPAs of a random sample of students who were employed (denoted Emp) and a random sample of students who were not employed (denoted NotEmp). Some MINITAB output obtained from this study is given below. In the questions below, μ_{Emp} denotes the population mean GPA of all students employed and μ_{NotEmp} denotes the population mean GPA of all students not employed.

```
Descriptive Statistics: Emp, NotEmp

Variable   N  N*    Mean   StDev  SE Mean  Minimum      Q1   Median      Q3  Maximum
Emp       55   0  2.8734  0.4294   0.0579   1.9257  2.5987   2.8761  3.2340   3.6547
NotEmp    65   0  3.0224  0.2717   0.0337   2.3223  2.8569   3.0286  3.2633   3.4923
```



Use this information for this question and the following two questions.

Which of the following statements is true?

(a) * Look at the sample of employed students. At least 25% of these students have a GPA equal to or below 2.6000.

(b) None of the other four statements is true.

(c) Look at the distribution of the GPAs in the sample of students who were not employed. This distribution is right skewed.

(d) Look at the sample of students who were employed. In this sample, more than 15 students have a GPA of 3.25 or higher.

(e) Look at the maximum observed GPA of the sample of employed students. According to the $1.5 \times IQR$ criterion, this value is an outlier.

In (a), for 25% less than a certain value, we need to look at Q1. Here Q1 is 2.5987, so 25% of the GPAs are less than 2.5987, and because 2.6 is slightly more than Q1, slightly more than 25% of the GPAs are less than 2.6. (Or exactly 25% if there are no data values between 2.5987 and 2.6). So (a) is true. You could mark A and move on to the next question, or check the others:

Skipping (b), you can use the right-hand normal quantile plot to assess normality (and thereby symmetry/skewness). I'd say the plot is pretty straight, which would imply the data are normally distributed and thus symmetric. If there is any curve, it is lower in the middle than you would expect for a straight line: this means that the lower values are more spaced out than you would expect from a normal distribution and the higher ones are closer together. This would mean that the values are skewed to the *left*. Either thought process is good, but right-skewness is definitely wrong.

In (d), Q3 says that 25% of employed students have a GPA above 3.234. 25% of 55 is 13.75. So there must be 13 or fewer students with a GPA above 3.234, and fewer than this with a GPA above 3.25. So (d) is false.

In (e), use the criterion to figure out which upper-end values would be outliers. $1.5 \times IQR$ is $1.5(3.2340 - 2.5987) = 0.953$, and Q3 plus this is $3.2340 + 0.953 = 4.187$. The highest observed value is just above 3.6, nowhere near as big as this, so the highest observed value is not an outlier.

34. Using the information in Question 33 above, which of the following statements is true?

(a) Consider the 90% confidence interval for μ_{Emp}, the population mean GPA of students who were employed. The margin of error of this confidence interval is greater than 0.11.

(b) Based on the on the information given in Question 33, we see that t-procedures cannot be applied for this data set.

(c) Consider the t-test for testing the null hypothesis $\mu_{Emp} = 3.0$ against the alternative hypothesis $\mu_{Emp} < 3.0$. The P-value of this test is less than 0.01.

(d) * None of the other four statements is true.

(e) Consider the t-statistic for testing the null hypothesis $H_0 : \mu_{Emp} = 3.0$ against the alternative hypothesis $H_a : \mu_{Emp} < 3.0$. This test statistic is between -2.00 and 2.00.

Same idea again: evaluate each statement and stop as soon as you find a true one (or go on and check the others). Another approach is to check the easiest one to check first (so that, with luck, you don't even have to check the hardest one).

In (a), go ahead and figure out the margin of error of the CI, using t procedures (since the population SD is not known). For the employed students: $df = 55 - 1 = 54$, so use 50 in Table D, and $t^* = 1.676$, which is a smidgen bigger than z^* would have been. The margin is $1.676(0.4294)/\sqrt{55} = 0.097$. This is not greater than 0.11, so (a) is false.

In (b), t procedures are good if the sample size is big enough to overcome any non-normality in the population. Here, (i) 55 is a decently big sample, and (ii) in the normal quantile plot, there is no serious evidence of non-normality, so the t procedure should be (doubly) OK. (b) is therefore false.

In (c), work out the test statistic and then get the P-value from it. The null hypothesis says $\mu = 3$, so $t = (2.8734 - 3)/(0.4294/\sqrt{55}) = -2.186$. We are on the correct side, so look up 2.186 without the minus sign in the 50 df line of Table D (closest). The one-sided P-value is between 0.01 and 0.02. The P-value is not less than 0.01, so this one is also false. (Or do it backwards: if the P-value is going to be less than 0.01, t would have to be bigger than 2.403 when you remove the minus sign. It isn't, so the P-value is not less than 0.01).

In (e), this is the same test as (c), so use the work from there. The test statistic was not between -2 and 2, so this one is false too. The answer to mark is (d).

35. Consider the information given in Question 33 above. If we were interested in testing the null hypothesis that the population mean GPAs of the two groups were equal, against the alternative that the means were different, what could we say about the P-value for this test?

It must be

(a) * between 0.02 and 0.05
(b) between 0.05 and 0.10
(c) between 0.01 and 0.02
(d) greater than 0.10
(e) less than .01

To figure out the P-value, you have to go all the way through the testing procedure. This is a two-sample t-test, since you are comparing two sample means. So work out the test statistic, which is

$$t = (2.8734 - 3.0224)/\sqrt{0.4294^2/55 + 0.2717^2/55} = -2.1746$$

(or the positive version of this if you did the subtraction the other way around). With 54 df (the smaller of $55 - 1$ and $55 - 1$!), the P-value is between twice 0.01 and twice 0.02, that is, between 0.02 and 0.04, it being a two-sided test. If this is true, (a) is certainly true as well.

36. A random variable X might have one of two probability distributions. The null hypothesis is that X has this distribution:

Value	0	1	2	3
Probability	0.50	0.40	0.08	0.02

The alternative hypothesis is that X has this distribution:

Value	0	1	2	3
Probability	0.10	0.30	0.35	0.25

A test of significance is carried out by observing one value of X, and rejecting the null hypothesis in favour of the alternative if the observed value is 2 or larger. Use this information for this question and the following one.

What is the probability of a Type I error for this test?

(a) 0.60
(b) cannot determine it, because we need to know the exact value that was observed.
(c) 0.05
(d) 0.08
(e) * 0.10

A type 1 error happens if the null hypothesis is true, but it is rejected. In this case, that means the first distribution is the correct one, but we got a value of X that would make us reject it (2 or bigger). The probability of this is $0.08 + 0.02 = 0.10$.

37. In Question 36, a test of significance was described for the distribution of a random variable X. What is the probability of a Type II error for this test?

(a) 0.35

(b) * 0.40

(c) 0.60

(d) 0.05

(e) 0.10

A type II error means that the alternative hypothesis is true, but we do not decide to reject the null. Here, that means that the second distribution is the true one, but we got a value of X that was 0 or 1 (not 2 or bigger). The probability of this is $0.10 + 0.30 = 0.40$.

38. When would you prefer to use t procedures (confidence interval, test of significance) instead of z procedures?

(a) When you want to obtain a smaller confidence interval for the same sample size.

(b) When you have a small sample.

(c) * When the population standard deviation is not known.

(d) When the population distribution is approximately normal.

(e) When you have a large sample.

The only distinction between t and z is that you do a t when you don't know the population SD. Some texts assert that it has to do with the sample size as well, but when you have a large sample it doesn't matter much which test you do, and in this course we've said it only matters whether you know the population SD or not. So (b) is defensible, but not the best answer. The other alternatives are wrong; for example, if the population distribution is approx normal, you could use either t or z depending on whether you know the population SD or not.

39. The costs of major surgery can vary substantially from one place to another. A study of the costs involved in a particular surgery was done in California and Montana. The 95% confidence intervals for the mean costs in the two states were reported to be from \$5826.76 to \$6173.24 in Montana and from \$6061.41 to \$6338.59 in California. No other information was given in the report. Assume that these intervals were calculated based on the normal distribution (i. e. assume that both the sample sizes were very large).

Calculate the upper limit of the 95% confidence interval for the difference between the two population means (i.e. mean cost in California minus mean cost in Montana). Choose the closest answer from the following options.

(a) * \$420

(b) \$375

(c) \$340

(d) \$200

(e) \$360

"Assume that both sample sizes were very large" means to use the bottom line of the t table, Table D. That means taking $t^* = 1.96$. We don't have the sample SDs, but we can work out what we need from the individual CI's given, because the margin of error is half the length of the confidence interval (we go up and down from the sample mean).
Label California as 1 and Montana as 2. Then:

$$\begin{aligned} t^* s_1/\sqrt{n_1} &= (6338.59 - 6061.41)/2 = 138.59 \\ t^* s_2/\sqrt{n_2} &= (6173.24 - 5826.76)/2 = 173.24 \end{aligned}$$

Since we are taking $t^* = 1.96$, divide both of these by 1.96 to get $s_1/\sqrt{n_1} = 70.71$ and $s_2/\sqrt{n_2} = 88.39$.

So that's what we have. What we need is the difference in sample means, and the margin of error is calculated as $t^*\sqrt{s_1^2/n_1 + s_2^2/n_2}$. The sample means are the midpoints of the given intervals, 6200 and 6000, while the things inside the square root are the squares of what we found above: for example, $s_1^2/n_1 = (s_1/\sqrt{n_1})^2$. So those are 4999.8 and 7812.4.

Finally, the margin of error of your two-sample interval is $1.96\sqrt{4999.8 + 7812.4} = 222$, and so the upper limit is $6200 - 6000 + 222 = 422$. This one was difficult, but it has in common with all the difficult questions that you have to figure out what you want, and then try to get it from what you have. (Was it worth struggling through all the above for one mark? Your call.)

40. A 90% confidence interval for a population mean, based on a sample of size 30, is from 40 to 65. This confidence interval turned out to be too big. Which of the following is a way of making the confidence interval shorter?

    (a) * Make the sample size bigger.
    (b) Use the median instead of the mean.
    (c) Make the sample size smaller.
    (d) Choose a higher level of confidence.

    Choosing a higher level of confidence like 99%, or taking a smaller sample size, will make the interval *longer*. The confidence intervals we have seen do not involve medians at all. So that leaves (a).

41. A magazine is considering the launch of an online edition. In a small survey, the magazine contacted a random sample of 500 current subscribers and asked whether they would be interested in an online edition. Some of these subscribers showed an interest and some did not. Based on the information from this sample, the investigators tested the null hypothesis $H_0 : p = 0.3$ against $H_a : p > 0.3$, where p is the population proportion of subscribers who will be interested in an online edition. The value of the Z-statistic for testing $H_0 : p = 0.3$ against $H_a : p > 0.3$ was 5.86.

    If we were interested in testing $H_0 : p = 0.4$ against $H_a : p > 0.4$ based on the same sample used above, what could we say about the value of the Z-statistic for this test? This value is:

    (a) between 1.5 and 3.0.
    (b) between 3.0 and 4.5.
    (c) greater than 4.5.
    (d) * between 0.5 and 1.5
    (e) less than 0.5.

    Start with what we have, and figure out how to get what we want.

    What we have is the z-statistic for testing $p = 0.3$. This is

    $$z = \frac{\hat{p} - 0.3}{\sqrt{(0.3)(0.7)/500}} = 5.86.$$

    We don't know $\hat{p}$, but we know everything else, so we can work it out:

    $$\hat{p} = 5.86\sqrt{(0.3)(0.7)/500} + 0.3 = 0.4201.$$

    This makes sense: the test statistic was (very) positive, so $\hat{p}$ should be quite a bit bigger than 0.3. Then go back and figure out z when $H_0 : p = 0.4$:

    $$z = (0.4201 - 0.4)/\sqrt{(0.4)(0.6)/500} = 0.92.$$

42. A simple random sample of 20 observations is taken from a population of successes and failures where the proportion p of successes is believed to be about 0.8. In the sample, 15 successes are observed. Which of the following statements best applies to this situation? ("Standard procedures" means what we learned in lectures.)

   (a) The probability of observing exactly 15 successes is 0.1746.
   (b) * The sample size is too small to use standard procedures to obtain a confidence interval for p.
   (c) There should have been exactly 16 successes observed in the sample.
   (d) A 95% confidence interval for p is from 0.56 to 0.94.

   The number of successes that would be observed in this situation has a binomial distribution with $n = 20$ and $p = 0.8$ approximately. Standard procedures use the normal approximation to the binomial (to get confidence intervals and do tests), but here $n(1 - p) = 20(0.2) = 4$ which is less than 10, so the normal approximation, and hence standard procedures, are no good. So (b) is true, and (d) should not be calculated at all (so we don't need to check it). (c) is false, because there is always sampling variability: we expect to see *about* 16 successes, but a particular experiment may give more or less. To check (a) using Table C, $p > 0.5$ so look up $20 - 15 = 5$ successes with $p = 1 - 0.8 = 0.2$: the value is indeed 0.1746, but our value of p, 0.8, was only approximately correct, so we ought to take this four-digit accuracy with a grain of salt. This means (b) is a better answer than (a).

43. A government agency is interested in whether air quality in Ontario cities is changing. The agency made a study: 10 Ontario cities are randomly selected, and the air quality in those cities is recorded on a date in 2007, and also the same date in 2008. A lower figure indicates better air quality. Summaries of the data collected are shown below ("difference" is year 2 minus year 1):

```
Variable     N    Mean   StDev      Q1  Median     Q3
year1       10   28.00    9.98   20.00   29.50  36.00
year2       10   38.90    9.07   29.50   43.00  46.50
difference  10   10.90    9.75    3.50   14.00  17.50
```

Carry out a suitable test of significance to determine whether mean air quality has changed. You may assume normality as necessary to carry out your test. What P-value do you get?

   (a) * between 0.005 and 0.01
   (b) between 0.02 and 0.05
   (c) greater than 0.05
   (d) between 0.01 and 0.02
   (e) less than 0.005

   This one is a matched pairs t-test, because the same 10 cities were used in both years. So do the calculation based on the "difference" line, and test the null hypothesis that the mean difference is 0.
   Here $t = (10.90 - 0)/(9.75/\sqrt{10}) = 3.535$ with 9 df. The one-sided P-value is between 0.0025 and 0.005, which needs to be doubled to get "between 0.005 and 0.01". (If you do a two-sample test or do it one-sided, you will get one of the other alternatives.)

44. A random variable X has this distribution:

Value	1	3	5
Probability	0.1	0.6	0.3

X has mean 3.4 and SD 1.2. A sample of 36 values is taken from this distribution. What is the approximate probability that the sample mean is bigger than 3.9?

(a) 0.38

(b) 0.30

(c) 0.25

(d) * 0.006

(e) 0.34

You don't actually need the individual values and probabilities, just the mean and SD. This is because we have a large sample from an only-very-slightly skewed distribution, so the Central Limit Theorem applies. (If it didn't, this would be a very difficult question!)

Using §5.2 procedures, $z = (3.9 - 3.4)/(1.2/\sqrt{36}) = 2.5$, and the prob of being greater than this is 0.0062 (approximately).

If you don't divide by $\sqrt{36}$, you are finding the approximate probability that an *individual value* is bigger than 3.9. But you don't need an approximation to find this: you can get it from the original distribution as 0.30. But this is not what was wanted.

45. An experimenter has decided to use a small value of α like $\alpha = 0.01$ in her experiment. Which of the following is a consequence of her decision?

(a) the probability of a Type II error is small.

(b) the probability of a Type I error is large

(c) * there is a small chance of rejecting the null hypothesis when it is true

(d) if the null hypothesis is false, the experimenter will easily be able to reject it

(e) even if the null hypothesis is far from being true, it will be hard to reject it.

The value of α only relates to type I errors, not type II errors, and likewise relates only to rejecting the null hypothesis when it is true (it says nothing about what might happen if H_0 is false). So that rules out (a), (d) and (e). α *is* the probability of a type I error, so (b) is also false. That leaves (c).

46. A nutrition laboratory tested a random sample of 50 "reduced sodium" hot dogs. The mean sodium content of the sample was 309mg. Let μ denote the population mean sodium content. A test was carried out to test the null hypothesis $H_0 : \mu = 300$ against $H_a : \mu > 300$. The P-value of this test was 0.038. A test was also carried out of the null hypothesis $H_0 : \mu = 298$ against $H_a : \mu > 298$ (using the same sample); the P-value was 0.015. Assume that the data satisfy all assumptions required for the tests.

If we calculate a 95% confidence interval for μ using this sample, what can we say about its margin of error? Choose the correct range for this margin of error from the following alternatives.

(a) between 12.00mg and 16.00mg

(b) * between 8.00mg and 12.00mg

(c) greater than 16.00mg

(d) between 4.00mg and 8.00mg

(e) less than 4.00mg

This requires care, because the tests are one-sided. We can make them two-sided by doubling the P-values, getting 0.076 for $\mu = 300$ vs. $\mu \neq 300$, and 0.030 for $\mu = 298$ vs. $\mu \neq 298$ (we had to change the alternatives to make them two-sided). So 300 is *inside* the 95% confidence interval (the P-value is greater than 0.05 after we doubled it) and 298 is outside (the P-value, even after we doubled it, is less than 0.05). The sample mean is 309 (given), so the margin of error has to be such that 300 is inside the interval (must be at least 9) and 298 is outside (must be smaller than 11). If the margin of error is between 9 and 11, it must also be true that it's between 8 and 12, so mark (b).

47. A random variable X has this probability distribution:

Value	1	3	5
Probability	0.1	0.3	0.6

What is the variance of X?

(a) 2.8
(b) 3.3
(c) * 1.8
(d) 1.2
(e) 4.0

Find the mean first: $1(0.1) + 3(0.3) + 5(0.6) = 4$. Then find the variance by summing (value-mean) squared times probability: $3^2(0.1) + 1^2(0.3) + 1^2(0.6) = 0.9 + 0.3 + 0.6 = 1.8$.

48. Let Z be a standard normal random variable and let T be a random variable with a t distribution with 5 degrees of freedom. Which of the following events is most likely (i.e. has the highest probability) to occur?

(a) Z and T are both greater than 2.0
(b) Z is less than -2.0
(c) * T is greater than 2.0
(d) Z is greater than 2.0
(e) Z is greater than or equal to 2.0

T is more likely than Z to be far away from 0 (in either direction) because the t distribution has longer tails than the normal. "Or equal" makes no difference for continuous random variables, so the probabilities in (d) and (e) are the same. Also, the chance of two things both happening is less than the chance of either one happening.

49. A 99% confidence interval for a population mean goes from 27 to 33. The interval was based on a sample size of 50. The interval was calculated using a known population standard deviation, but unfortunately the value has been lost. What is the population standard deviation?

(a) 10.8
(b) 24.6
(c) * 8.2
(d) 1.2
(e) 16.4

Again, work from what you know so that you can work out what you don't. The margin of error of the confidence interval is half its length, $(33 - 27)/2 = 3$, which is equal to $z^*\sigma/\sqrt{n}$. Here, $z^* = 2.576$ and $n = 50$, so

$$3 = 2.576\sigma/\sqrt{50}$$

and rearrange to find that $\sigma = 8.23$. Mark the closest answer to that.

50. An insurance company is trying to estimate the mean number of sick days that food service workers take per year. The company wishes to use a 90% confidence interval. The standard deviation of the number of sick days is known to be 2.5. How big a sample does the insurance company need to take if they wish to estimate the mean to within 0.5 days?

(a) 42

(b) 30
(c) 8
(d) * 68

Use the formula for sample size to get

$$n = (z^*\sigma/m)^2 = ((1.645)(2.5)/0.5)^2 = 67.65$$

and round up to get 68. Or write $m = z^*\sigma/\sqrt{n}$, put in the values for m, z^*, σ and solve for n.

51. An association of Christmas tree growers in Indiana commissioned a survey, in which a simple random sample of all Indiana residents was taken. One of the questions in the survey was "Did you have a Christmas tree during the last holiday season"? 400 people were surveyed, and 331 answered "yes". What is the upper limit of a 99% confidence interval for the proportion of all Indiana residents who had a Christmas tree during the last holiday season?

(a) 0.83
(b) 0.86
(c) 0.80
(d) * 0.88
(e) 0.78

Work it out: the interval is $\hat{p} \pm z^* \sqrt{\hat{p}(1-\hat{p})/n}$, and plug in 2.576 for z^*, $331/400 = 0.8275$ for $\hat{p}$ and 400 for n. The upper limit is 0.876.

52. A regression was carried out to predict a variable y from a variable x. The residuals from the regression were plotted against x, as shown below:

What do you learn from the plot?

(a) A curve should be fitted rather than a straight line.
(b) The residuals have a straight-line relationship with x.
(c) A straight-line relationship describes the relationship well.
(d) * The predictions become less accurate as x increases.
(e) The correlation between x and y is low.

The residuals "fan out": they tend to get further from 0 as x gets larger. That is, the predictions get less accurate as x gets bigger. There is no curved pattern, so no evidence that a curve should be fitted instead, but there *is* a pattern, so the regression is not satisfactory.