EXAM 2

Practice A

- There are 110 possible points on this exam. The test is out of 100.

- You have one class period to complete this exam, but you should be able to complete it in less than that

- Please turn off all cell phones and other electronic equipment.

- Be sure to read all instructions and questions carefully.

- Remember to show all your work. Writing down what you put into Excel is sufficient to show your work.

- Try all questions! You get zero points for questions that are not attempted.

- Note the last sheet lists all the equations you will need for this exam.

- *Please print clearly and neatly.*

Part I: Matching. *Write the letter from the column on the right which best matches each word or phrase in the column on the left. You will not use all the options on the right and you cannot use the same option more than once.*
    2 points each.

1. **E** Central Limit Theorem
2. **A** Confidence interval
3. **F** Degrees of freedom
4. **C** Practical significance
5. **H** Statistical significance
6. **B** Type I Error
7. **D** Type II Error

A. Attempts to capture the true population mean but not always.
B. Example: Buying a video game that turned out to be boring.
C. Example: Determining that a machine is putting in twice as many chips in a bag as it was designed to.
D. Example: Refusing to major in a subject you find interesting and would lead to a great job.
E. Justifies why a difference between the population and sample means does not mean something interesting is going on.
F. Needed for critical t scores but not for critical z scores.
G. Needed for critical z scores but not for critical t scores.
H. When the absolute value of the calculated score is greater than the critical score.
I. When you know the population standard deviation.

1. *The CLT reminds us that sample means can be unusually large or small by pure chance. It's not enough to measure the sample average; you need to determine how likely its difference is due to randomness.*
2. *Confidence intervals try to capture the true population mean with a margin of error centered around the sample average. It's an application of the CLT. Yes, the sample average is the best guess but it could be a little lower or a little higher. How much difference is reasonable? The interval tells us.*
3. *When you don't know σ, you need degrees of freedom which inform you how closely the t distribution represents the z distribution; recall the t distribution is really a family of distributions; the degrees of freedom tell you which "family member" you care about.*
4. *"Determining" implies the test was statistically significant. And it was practically significant, too! While it technically depends on details, it's hard to imagine a scenario where it's not a big deal that the machine is so very off.*
5. *By definition; the absolute value of the difference between the population mean and the sample mean, adjusted for sample size and standard deviation,*

is compared to a critical value based on the normal distribution. This tells us the chance that the difference was a coincidence (alpha) or not.

6. *Your null hypothesis should be "not buy a game." That's your default state. If you reject it when you should have failed to reject it, you bought the game when you should have stuck with the null. Type I error.*

7. *Your null hypothesis is to not major in something. By failing to reject that null, you risk missing out on a great major. (Like economics!) Type II error.*

Part II: Multiple Choice. *Choose the best answer to the following.*
    4 points each.

8. Suppose you're running a poll to determine if Senator Andy Guzman will win his re-election. He has one opponent. Out of 100 likely voters surveyed, Sen. Guzman had 60% in favor of his re-election. What is the null hypothesis of this test?
    a. $\mu = 0.00$
    b. $\mu = 0.40$
    c. $\mu = 0.50$
    d. $\mu = 0.60$
    e. $\mu = 1.00$

The null hypothesis is whatever the default assumption is; here it's that he's split the vote, 50/50.

9. Which of the following is an example of a two-tailed test?
    a. If your new internet is faster than the average.
    b. Whether a clock keeps better time than the standard clock.
    c. Testing if a different car frame design changes its strength.
    d. Determining if a dispenser puts the right amount of catnip in a cat toy.
    e. None of the above

All these examples except D have extreme values which are not interesting—an internet that's extremely slow, a clock that keeps much worse time, a weaker body—because all of these deal with improvements. But it is noteworthy if there's too little catnip or too much catnip placed in a toy. Either could represent a problem with the machine (and thus suggesting maintenance).

10. If the null hypothesis is rejected at the 99% level, what must be true?
    a. The null hypothesis is rejected at the 95% level.
    b. The null hypothesis is rejected at the 99.9% level.
    c. The p value is less than 0.01
    d. A & C
    e. All of the above

If something statistically significant at one particular significance level, it's always going to be statistically significant at all significance levels <u>below</u> that one. Also, a p-value represents the lowest α one can claim and still get statistical significance. The α for 99% is 0.01, so the p-value would have to be below that value.

11. Suppose you sample various factories in Maryland and various factories in Virginia. You find that when you take into account the costs of labor, regulatory compliance, quality control, and other considerations, moving your factories to Virginia would save your firm $15 a year. The calculated z-value associated with this difference is 3.52. At 99% confidence, what should your conclusion be?
    a. It's statistically significant and practically significant.
    b. It's statistically significant but not practically significant.
    c. It's not statistically significant but it is practically significant.
    d. It's not statistically significant nor practically significant.
    e. None of the above / It is impossible to tell with the information provided.

At 99% confidence, a two-tailed test is significant at the 2.576 level. 3.52 is well above this. But at just a $15/year difference, it's not worth moving your whole operation across state lines.

12. Suppose you gambled on some farmland which was priced very cheaply because it was once occupied by a feed lot. This could make the land very good, very poor, or have no effect. To test if your investment was a good one, you weigh a sample of 13 randomly select watermelons you just grew. You're new at watermelon planting so you can only go on the standard deviation of your sample which was four pounds with an average of 18.25 pounds. The average watermelon weighs 20 pounds. At 95% confidence, describe the results of your test.
    a. Reject the null because the result is less than the critical value.
    b. Fail to reject the null because the result is less than the critical value.
    c. Reject the null because the result is more than the critical value.
    d. Fail to reject the null because the result is more than the critical value.
    e. None of the above

Since we don't know the population standard deviation, we have to use the t-table provided. Ignore the z-score chart.

$$t = \left|\frac{18.25 - 20}{4/\sqrt{13}}\right| = \left|\frac{-1.75}{1.109}\right| = 1.577$$

With 12 (13 – 1) degrees of freedom, 1.577 < 2.179.

If you thought the critical value was 1.782, that was for the one-tailed test, not two-tailed. Recall that our watermelons could be really good or really bad. Turns out, they might just be average. We fail to reject the null.

13. Huyen runs a shoe company needs to figure out how much leather to order (which changes due to variations of assembly, differences in shoe sizes, etc.). Based on a sample of 40 pairs, each pair uses, on average, 0.56 square meters of leather with a sample standard deviation of 0.17. At 99.9% confidence, what's the confidence interval of leather, in square meters, used per pair of shoes?
    a. **Between 0.464 & 0.656**
    b. Between 0.472 & 0.648
    c. Between 0.487 & 0.633
    d. Between 0.491 & 0.629
    e. None of the above

Note you are using the t distribution because you don't know the population standard deviation. You only know the sample standard deviation. Thus you type into Excel "=CONFIDENCE.T(0.001,0.17,40)" resulting in about 0.0956. Add and subtract that result from 0.56 (the sample mean, your point estimate) and get 0.656 on the high end and 0.464 on the low end.

14. Suppose you wanted to if people spent an unusually large amount of time looking at your website. Based on a sample of 300 users, you found they spent an average of 14.8 minutes looking at your website with 99% margin of error of 3 minutes. Suppose the average amount of time spent on a website is 12.2 minutes. Is your sample average statistically significant compared to the average?
    a. No, because the sample size is too small.
    b. **No, because the population average is in the confidence interval.**
    c. Yes, because the population average is larger than the margin of error.
    d. Yes, because the difference between the sample average and the population average is less than the confidence interval.
    e. Because you do not have the standard deviation, it is impossible to tell the answer to this question.

This example creates a 99% confidence interval between 11.8 minutes and 17.8 minutes. In other words, there is a 99% chance the true population mean is in that interval. Note this is indeed the case—the population mean is 12.2—so the fact that your sample average happens to be higher than your population average is probably a coincidence. The definition of not statistically significant.

Note as well that you must have the standard deviation; you could not have determined a margin of error without it.

15. Which of the following is a true statement?
    a. "As α increases, the margin of error decreases."
    b. "All practically significant results are statistically significant."
    c. "There is no maximum value for the standard deviation of a sample proportion."
    d. A & B
    e. All of these are true

A is true. As α increases, you establish a lower confidence level. Therefore, the margin of error shrinks (because at a lower confidence level the associated z or t score is smaller). Because the interval is defined as the margin of error plus and minus the sample mean, a smaller error means a narrower interval.

B is also true. If the result isn't statistically significant, then it's functionally zero which is never practically significant.

But C is not true. Because proportions are confined to between zero and one, there's a natural maximum for standard deviation. You can see that in the hypothesis testing equation for proportions, where you multiply

16. Consider a normal distribution with a mean of sixty and a standard deviation of eight. What percent of observations have a value more than 58 but no bigger than 61?
    a. 0.1%
    b. 14.8%
    c. 19.7%
    d. 24.5%
    e. None of the above

First begin by finding the portion of observations 61 or less. That's "=NORM.DIST(61,60,8,1)", or 0.5497. Then find the portion of observations that's 58 or less. That's "=NORM.DIST(58,60,8,1)", or 0.4013. The difference between the two is 14.8%.

17. Alice is an analyst for an airline and she wants to know if removing the charge for checked bags will reduce how often people carry on their luggage. Suppose 60% of airline passengers have no checked bags. The airline decides to offer 122 different customers across five flights free checked bags and 52% of those customers have no checked bags. Using standard thresholds for statistical significance, is this result statistically significant?
    a. It's not statistically significant.
    b. It is statistically significant at a confidence level of 95%, but no higher.
    c. It is statistically significant at a confidence level of 99%, but no higher.
    d. It is statistically significant at a confidence level of 99.9%.
    e. There is not enough information to determine an answer.

This is a test of proportion, since we're wondering what portion of passengers will check luggage. Before we run the numbers, note that (a) this is a one-tailed test and (b) while this is across five flights, the relevant sample size is 122.

$$z_p = \left| \frac{0.52 - 0.60}{\sqrt{\frac{0.6(1-0.6)}{122}}} \right| = \left| \frac{-0.08}{0.04435} \right| = 1.804$$

Since this is a one-tail z test and the critical value at 95% confidence is 1.645, we do have statistical significance. Note that if it was a two-tail test, the critical value would be 1.960 and wouldn't be statistically significant.

18. Where is the sample average in relation to a confidence interval?
    a. **In the exact middle.**
    b. Somewhat near the middle; it depends on how the sample was gathered.
    c. It depends on the standard deviation.
    d. It's either at the low end or the high end of the interval.
    e. It's impossible to know.

The confidence interval is calculated by adding and subtracting the margin of error from the sample average; thus the sample average is in the exact middle of the confidence interval.

19. Suppose you find that 10% of men are "tall," with "tall" being defined as more than 72 inches tall. Assuming a symmetric normal distribution of height, if the standard deviation of men's height increased, what would happen?
    a. Fewer than 10% of men would be "tall."
    b. Fewer than 10% of men would be "tall" for small increases, but it would be the opposite for large increases.
    c. 10% of men would still be "tall."
    d. More than 10% of men would be "tall" for small increases, but it would be the opposite for large increases.
    e. **More than 10% of men would be "tall."**

Increasing the standard deviation "flattens" the normal distribution; with greater dispersion, more observations are shoved to the extreme values and you get more "tall" men.

Part III: Short Answer. *Answer the following.*
    16 points each.

20. The self-checkout was invented by David R. Humble in 1984. It uses a scale to confirm the item scanned was placed in the bag. But all item weights vary. And some weights—like of the bread baked in the store—vary a lot. So the self-checkout uses a confidence

interval: as long as the weight of the item is in the interval, the computer reads it as the item in question. Suppose you are programming the computer with bread weights. You test the weight of 12 loaves of bread, getting an average of 1.4 pounds with a population standard deviation of 0.2 pounds. To 99% confidence, what's the range of the confidence interval? Remember to show all your work and/or relevant Excel commands.

I will do this both by hand and with Excel. First, by hand.

Note that we have the population standard deviation so we'll be using the z scores. Since this is a confidence interval, we'll use the two-tailed version. At 99% confidence, z=2.576.

Thus: $1.4 \mp 2.576 \left(0.2 / \sqrt{12} \right) = 1.4 \mp 2.576 \left(0.2 / 3.464102 \right)$
$$= 1.4 \mp 2.576(0.057735) = 1.4 \mp 0.148725$$

The lower boundary is about 1.2513
The upper boundary is about 1.5487

Using Excel, note, again, we know the population standard deviation. Thus we will use the CONFIDENCE.NORM function (instead of CONFIDENCE.T).

Type: "=CONFIDENCE.NORM(0.01,0.2,12)" and press ENTER. You should get 0.148716. This is the margin of error. Note that it's slightly difference than what I calculated above (0.148725). This is due to rounding.

Also note that Excel asked for probability—our alpha—not confidence. That's why we type 0.01 and not 0.99.

Adding and subtracting, you get:

The lower boundary is about 1.2513
The upper boundary is about 1.5487

Same as before. If rounding differences would have resulted in a different number, it's not a big deal. This is why I ask you to show your work. If you got a slightly different number than I have, I can see if it's due to rounding. If it is, and the rounding isn't misguided (e.g. rounding 1.41 to 1.4), then I won't penalize you.

21. Suppose 34% of people drink coffee in the morning and you want to know the effect of a new advertising campaign your coffee company is about to launch. After it's launched, a

survey (based on 2,395 people) reveals that 36% drink coffee in the morning. At 95%, 99%, and 99.9%, what should you conclude?

In answering this question, be sure to:
- Show your work on the calculated value;
- Indicate what your critical value or p-value is;
- Indicate what your critical values are;
- Determine if this is statistically significant

This is a two-tailed test—you're checking for an effect, not if it "improved" it—so you do some math:

$$z_p = \left|\frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}\right| = \left|\frac{0.36 - 0.34}{\sqrt{\frac{0.34(1 - 0.34)}{2,395}}}\right| = \left|\frac{0.02}{0.00968}\right| = 2.066$$

At 95% confidence (z=1.96), it is statistically significant. But at 99% (z=2.576) and 99.9% (z=3.291), it is not statistically significant.

22. Gustav Graves is slicing meat at his deli. Like all meat slicers, his slices meat as thinly as physically possible; that way he gets the largest number of slices, and the most flavor, per pound of meat. Thinner slices are always better.

    His meat slicer normally slices meat 0.8 mm thick but he's wondering if it needs to be cleaned and serviced. He uses it to make 14 slices. The average thickness of this sample is 1 mm with a standard deviation of 0.4.

    At 95% confidence, does Gustav need to have his meat slicer serviced? If this difference is statistically significant, how would you know if it's practically significant (name at least two factors)?

    In answering this question, be sure to:
    - Show your work on the calculated value;
    - Indicate what your calculated value is;
    - Indicate what your critical value or p-value is;
    - Determine if this is statistically significant;
    - Determine how you would know if this is practically significant

*This is a hard question. First and foremost, recognize that this is **not** a two-tailed test. It is a one-tailed test. Remember: thinner slices are always better. In other machine examples, too much or too little indicated a problem. But if the machine for whatever reason is slicing meat even thinner than 0.8 mm, that's cause for celebration, not a reason to service it.*

Also note this is a t-test, not a z-test. We have a sample standard deviation, not a population standard deviation

So we do the math: (1-0.8) / (0.4/sqrt(14)) = 1.87

We have 13 degrees of freedom (14-1). With a one-tailed test of an alpha of 0.05, (=T.INV(0.1,13)), you have a critical value of 1.771 (note that the =TINV command defaults to two-tailed so it must be converted; at an alpha of 0.1, 0.05 is in either half of the tail).

So yes, it is statistically significant. The slicer needs to be serviced.

But should Gustav service it? 0.2 mm difference doesn't seem like that much. It might be. It might not. It depends on:
- *How much would it cost, in time and money, to get it fixed? If it cost $600, it's probably best to let it slide until the problem is bad enough.*
- *Does Gustav have another slicer? If this is his only one, not having it for even a day would be a big problem.*
- *How much meat does Gustav use? At the current sample average, for every 4 slices Gustav cuts, he loses a slice if the machine was functioning properly (at four slices, he goes through 4 mm. If each slice was 0.8 mm thick, he'd get five slices out of 4 mm of meat). Turning five slices into four slices isn't a huge deal, but if he goes through a lot of meat every day, that can multiply into scores of sandwiches he's effectively giving away for free.*
- *How much does he pay for meat? If this is cheap meat he's losing, not a big deal. But if this is very fancy meat that's he's effectively giving away, service that slicer ASAP. This idea leads to an interesting prediction: high-end delis will have better slicers than low-end delis because their meat is probably more expensive and thus it pays to get the absolute most out of it.*
- *To answer this part of the question correctly, you wouldn't have to note all of these factors; as mentioned, two would suffice. But the basic goal is to demonstrate an understanding of the practical application behind statistical significance.*

Exam 2 Equation and Information Reference

Function	Output
ABS	The absolute value of an input
AVERAGE	Arithmetic mean of a dataset
CONFIDENCE.NORM	Determines the margin of error to make a confidence interval (known σ)

CONFIDENCE.T	Determines the margin of error to make a confidence interval (unknown σ)
CORREL	Correlation coefficient of two variables
CTRL + `	Show formulas
CTRL + F	Find
CTRL + P	Print
CTRL + X	Cut highlighted area
CTRL + C	Copy highlighted area
CTRL + V	Paste highlighted area
CTRL + Z	Undo
F4	Makes cell reference absolute
GEOMEAN	Geometric mean of a dataset (adjustments must be added manually)
LARGE	Larger values of a dataset (k=1 is largest, k=2 is second largest, k=3 is third largest…)
MAX	Maximum value of a dataset
MEDIAN	Median of a dataset
MIN	Minimum value of a dataset
MODE	Mode of a dataset
NORM.DIST	Returns the normal distribution for a specified mean and standard deviation.
NORM.INV	Returns the inverse of the normal cumulative distribution for a specified mean and standard deviation.
NORM.S.DIST	Returns the standard normal distribution.
NORM.S.INV	Returns the inverse of the standard normal cumulative distribution. Useful for finding critical z scores.
QUARTILE	The 0th to 4th quartile of a dataset
SQRT	Finds the square root of the value in question.
SMALL	Smaller values of a dataset (k=1 is smallest, k=2 is second smallest, k=3 is third smallest…)
STDEV.S	Standard deviation of a sample
T.INV	Finds area under a t distribution; useful for finding one-tailed critical t scores.
T.INV.2T	Finds area under a t distribution; useful for finding two-tailed critical t scores.

Geometric Mean

$$Geometric\ Mean = \sqrt[n]{\prod_{i=1}^{n}(1 + x_i)} - 1$$

Weighted Average

$$Weighted\ Average = \frac{\sum_i^n (w_i x_i)}{\sum_i^n w_i}$$

Coefficient of Variation

$$CV = \frac{s}{\bar{x}}$$

Confidence interval for proportion

$$\widehat{CI}_{\bar{p}} = \bar{p} \mp z_{\alpha/2}\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

z-test

$$z_{\bar{x}} = \left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right|$$

t-test

$$t_{\bar{x}} = \left|\frac{\bar{x} - \mu}{s/\sqrt{n}}\right|$$

z-test (proportion)

$$z_p = \left|\frac{\bar{p} - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}\right|$$

Critical z scores

Use =NORM.S.INV command

Confidence	α	$z_{\alpha/2}$	z_α
95%	0.05	1.960	1.645
99%	0.01	2.576	2.326
99.9%	0.001	3.291	3.090

Critical t scores

Use T.INV or T.INV.2T commands or see the table on the last page

p-values

Make your calculated value negative and then use one of the following (make sure cumulative is turned <u>on</u>):

	1 tail	*2 tails*
z	NORM.S.DIST	Multiply 1 tail
t	T.DIST	result by 2

Table B

t distribution critical values

df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
						Tail probability p						
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%

Confidence level C