**Math 247: Test 4** Name: KEY

1. (2 pts) What is a P-value? A P-value is the probability of getting a value as extreme (or more extreme) as the observed value <u>if</u> the null hypothesis is true.

2. (2 pts) Can a P-value ever be zero? YES [✓ NO]

   If Minitab reports a P-value as 0.000, what does this mean about the P-value?
   The P-value is so small (like .0001) that Minitab doesn't show enough decimial places to see any non-zero digits.

3. (2 pts) Why do we use the t-distribution instead of the z-distribution (the Normal distribution) for making inferences about population means?
   Since we rarely if ever know $\sigma$ we have to use S which introduces extra variability into the distribution

4. (2 pts) The two types of statistical <u>inference</u> that you've learned about are

   a. Confidence intervals b. Hypothesis testing

**Multiple Choice: Circle the best answer to each question.**
(Each question is worth 3 points)

---

For problems 5, 6, and 7, use this information: Suppose we have a population, take a sample of size n = 10, and find the sample mean. We then repeat this process 1000 times.

5. If we were to graph the result of all those sample means in a histogram or dotplot we will have created which of the following:

(a) the population distribution [✓ (b) the sampling distribution of the mean] (c) the distribution of the sample

6. The shape of this distribution would be which of the following: due to the small sample size (n< 25) the sampling dist. isnt guaranteed to be normal

(a) Normal (b) skewed left (c) skewed right [✓ (d) can't tell from this information]

7. If we increased the sample size to n = 50 then the shape of this distribution would be which of the following:

[✓ (a) Normal] (b) skewed left (c) skewed right (d) can't tell from this information

---

8. For a Before-and-After study, you would use which test:

   a. One sample t-test [✓ b. Paired t-test] c. Two sample t-test d. ANOVA

9. In a Before-and-After study, the samples are

   a. Independent [✓ b. Dependent] c. Can't tell

10. If a study finds a statistically significant result, that automatically means that result is clinically significant.

True [✓ False]

11. In an ANOVA test, you will reject the null hypothesis if which of the following are true:
   - a. All the means are equal
   - [✓ b.] There is a large amount of variation between the sample means and a small amount of variation within the samples.
   - c. There is a small amount of variation between the sample means and a large amount of variation within the samples.
   - d. You know all of the population standard deviations.

12. In an ANOVA test, the Test Statistic is
   - a. A z-value
   - b. A t-value
   - c. A Confidence Interval
   - [✓ d.] An F-value

**Short answer:**

13. (6 pts) In finding a confidence interval for a random sample of 30 Cuesta students' GPA's, one interval was (2.60, 3.20) and the other was (2.65, 3.15). One of these intervals is a 95% CI and the other is a 99% CI.

(a) Which interval is the 95% CI? (2.65, 3.15)

(b) Interpret the interval from part (a) (whichever one you chose) in the context of the problem.

We are 95% confident that the population mean GPA for Cuesta students is between 2.65 and 3.15.

14. (6 pts) A company has developed a new type of lightbulb, and wants to estimate its mean lifetime. A simple random sample of 26 bulbs had a mean lifetime of 560 hours with a standard deviation of 30 hours.

Construct a 90% confidence interval for the population mean lifetime of all bulbs manufactured by this new process. Do this by hand, using the formula. The t-chart is at the end of the test. <u>Assume the conditions for the CLT are satisfied.</u>

CI: $\bar{x} \pm t^* SE$

$560 \pm 1.708(5.883)$

$560 \pm 10.0$

$SE = \frac{30}{\sqrt{26}}$

$= 5.883$

Parking Lot

$n = 26 \quad df = 25$

$\bar{x} = 560$

$s = 30$

$t^* = 1.708$

CI: (550, 570) (you do not have to interpret the interval)

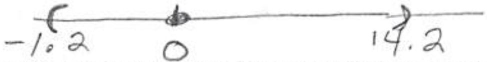
15. (8 pts) A researcher is investigating whether non-vegetarian women weigh a different amount on average than vegetarian women. A sample of 30 non-vegetarian women had a mean of 142 pounds and a sample of 40 vegetarian women had a mean of 135 pounds.

2 (a) What would the hypotheses be for the test? Write them in symbols and in words.

$H_0: \mu_{NON} = \mu_{VEG} \Rightarrow \mu_{NON} - \mu_{VEG} = 0$ — The mean weight of non-vegetarian women is the same as vegetarian women

$H_a: \mu_{NON} \neq \mu_{VEG} \Rightarrow \mu_{NON} - \mu_{VEG} \neq 0$ — The mean weights are different

(b) The 95% confidence interval for the difference between the mean weight of non-vegetarian and vegetarian women is ( -1.2, 14.2).



3 Does this interval show that there is a significant difference between weights of vegetarians and non-vegetarians? Explain your answer.

No there is not a significant difference. Since the interval captures zero, and this implies $\mu_{NON} - \mu_{VEG}$ could be zero, we have to conclude there is not a significant difference.

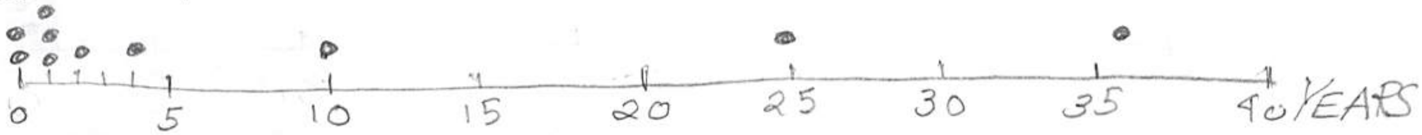
3 Based on this interval, would you reject or not reject the null hypothesis (assume the significance level for the test was .05)? (Circle your answer.)

Reject the null [✓ Do not reject the null]

16. (4 pts) To study the ages of pennies, a random sample of 10 pennies is drawn with the ages (in years) shown below.

0, 0, 1, 1, 1, 2, 4, 10, 25, 36

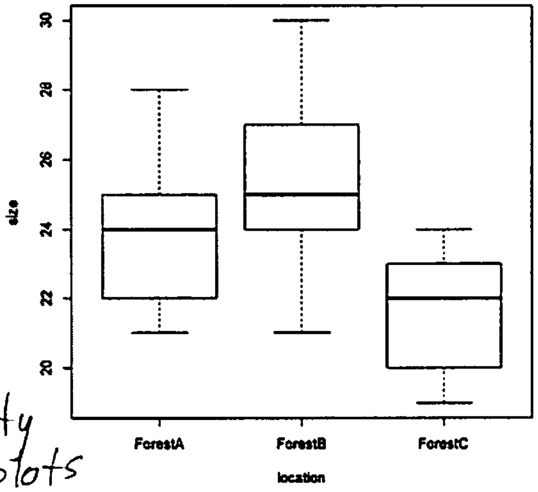
1 (a) Make a dotplot of this data



3 (b) Why would it NOT be appropriate to construct a confidence interval for the mean age of pennies using this data? Include your observations about the dotplot in your answer.

The dotplot shows the distribution of ages is strongly skewed to the right and has outliers so this is strong evidence that the underlying population is not normal. If the underlying pop is not normal and the sample size is small (n=10<25) then the CLT conditions are not satisfied.

17. (6 pts) A biologist did a study on the diameter of trees in 3 different forests. She gathered random, independent samples of 15 trees in each of the 3 forests. Given the boxplots of the data shown, and considering just the conditions of sample size and equal variance, would it be appropriate for her to use ANOVA to see whether there is a significant difference in mean tree size in the different forests?



It [✓ IS] or IS NOT appropriate to use ANOVA (circle one)

Explain. Since the sample size is small ($n = 15 < 25$), we have to check normality of the underlying populations. The boxplots don't show strong skewing or outliers so we can assume normality. We also have to assume equal variances in the populations The IQR's of the boxplots are roughly the same so this condition is also satisfied.

18. (8 pts) A random sample of students was studied to see whether seating position in the class is associated with GPA. The seating position of the students was observed (front, middle, back) and their GPA's were noted.

What is the factor (treatment) in this study? Seating position

What is the response? GPA

What are the hypotheses for the test?

$H_0: \mu_{front} = \mu_{mid} = \mu_{back}$ (Seating position is not associated with GPA)

$H_a$: At least one mean is different. (There is an association.)

An ANOVA test was performed with the following Minitab result:

```
Analysis of Variance
Source   DF  Adj SS  Adj MS  F-Value  P-Value
Factor    2   2.350  1.1752     7.50    0.003
Error    27   4.229  0.1566
Total    29   6.580
```

← $.003 < .05$ (default) & Reject $H_0$

Which number measures the total variance BETWEEN the groups? $MS_{factor} = 1.1752$

Which number measures the total variance WITHIN the groups? $MS_{error} = .1566$

Show how the F-value is determined from these values: $F = \frac{1.1752}{.1566} = 7.50$

What is the conclusion of the test? (Assume all of the conditions were met).

Reject $H_0$. At least one mean GPA is significantly different from the others.
There is an association between seating position and GPA.

**Take home (due Wednesday 5/3/17)** **Name:** KEY

(10 pts) In trying to determine whether air pollution causes reduction in children's lung health a researcher proposes to see if the lung volume of 10-year-old boys who live in high ozone pollution is significantly less than the lung volume of all 10-year-old boys. The mean volume for all 10-year-old boys is 2.05 liters A random sample of 100 10-year-old boys who live in a community with high levels of ozone pollution is found to have a mean volume of 1.98 liters, with a standard deviation of 0.3 liters.

Perform all steps of the hypothesis test. For full credit, include a sketch of the sampling distribution with two axes. Shade in the area that represents the P-value. Find the P-value using Minitab (but do all other parts of the test by hand). Include a printout of your Minitab work.

(1) $H_0: \mu_{Polluted} = 2.05$ Liters (Same lung volume)

$H_a: \mu_{polluted} < 2.05$ (Lung volume is reduced.)

Parking Lot

$\mu_0 = 2.05$ l (All boys)

$n = 100$

$\bar{x} = 1.98$ l (Boys in polluted areas)

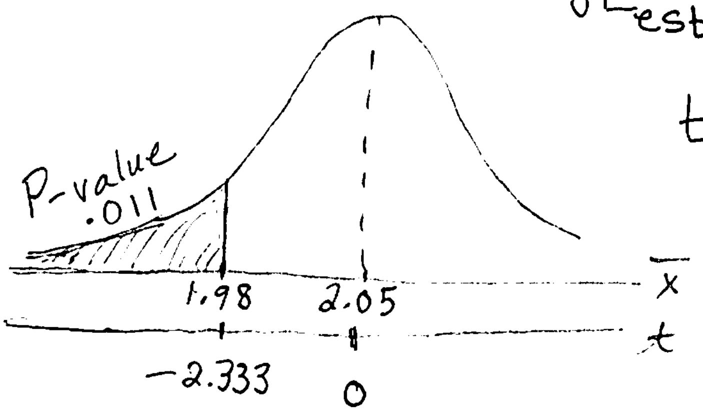
$s = .3$ L

(2) 1) Sample is random (stated) and observation are independent (assume no relatives)

2) Large sample. $n = 100 > 25$

3) Large pop $\geq 10(100) = 1000$ (There are more than 1000 boys living in high ozone areas.)

(3)



$$SE_{est} = \frac{s}{\sqrt{n}} = \frac{.3}{\sqrt{100}} = .03$$

$$t = \frac{\bar{x} - \mu_0}{SE_{est}} = \frac{1.98 - 2.05}{.03} = -2.333$$

P-value = .011 (Minitab)

(4) Using a .05 level of significance, .011 < .05 so reject Ho. We have evidence to conclude there is a statistically significant reduction in lung volume in 10-year-old boys living in high ozone polluted areas.

## One-Sample T

```
Test of μ = 2.05 vs < 2.05

  N    Mean   StDev  SE Mean  95% Upper Bound      T      P
100  1.9800  0.3000   0.0300           2.0298  -2.33  0.011
```

# *t* Table

| cum. prob | $t_{.50}$ | $t_{.75}$ | $t_{.80}$ | $t_{.85}$ | $t_{.90}$ | $t_{.95}$ | $t_{.975}$ | $t_{.99}$ | $t_{.995}$ | $t_{.999}$ | $t_{.9995}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| one-tail | 0.50 | 0.25 | 0.20 | 0.15 | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 | 0.0005 |
| two-tails | 1.00 | 0.50 | 0.40 | 0.30 | 0.20 | 0.10 | 0.05 | 0.02 | 0.01 | 0.002 | 0.001 |
| df | | | | | | | | | | | |
| 1 | 0.000 | 1.000 | 1.376 | 1.963 | 3.078 | 6.314 | 12.71 | 31.82 | 63.66 | 318.31 | 636.62 |
| 2 | 0.000 | 0.816 | 1.061 | 1.386 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.327 | 31.599 |
| 3 | 0.000 | 0.765 | 0.978 | 1.250 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.215 | 12.924 |
| 4 | 0.000 | 0.741 | 0.941 | 1.190 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 | 8.610 |
| 5 | 0.000 | 0.727 | 0.920 | 1.156 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 | 6.869 |
| 6 | 0.000 | 0.718 | 0.906 | 1.134 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 | 5.959 |
| 7 | 0.000 | 0.711 | 0.896 | 1.119 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 | 5.408 |
| 8 | 0.000 | 0.706 | 0.889 | 1.108 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 | 5.041 |
| 9 | 0.000 | 0.703 | 0.883 | 1.100 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 | 4.781 |
| 10 | 0.000 | 0.700 | 0.879 | 1.093 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 | 4.587 |
| 11 | 0.000 | 0.697 | 0.876 | 1.088 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 | 4.437 |
| 12 | 0.000 | 0.695 | 0.873 | 1.083 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 | 4.318 |
| 13 | 0.000 | 0.694 | 0.870 | 1.079 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 | 4.221 |
| 14 | 0.000 | 0.692 | 0.868 | 1.076 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 | 4.140 |
| 15 | 0.000 | 0.691 | 0.866 | 1.074 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 | 4.073 |
| 16 | 0.000 | 0.690 | 0.865 | 1.071 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 | 4.015 |
| 17 | 0.000 | 0.689 | 0.863 | 1.069 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 | 3.965 |
| 18 | 0.000 | 0.688 | 0.862 | 1.067 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 3.610 | 3.922 |
| 19 | 0.000 | 0.688 | 0.861 | 1.066 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 | 3.883 |
| 20 | 0.000 | 0.687 | 0.860 | 1.064 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 | 3.850 |
| 21 | 0.000 | 0.686 | 0.859 | 1.063 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 3.527 | 3.819 |
| 22 | 0.000 | 0.686 | 0.858 | 1.061 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 3.505 | 3.792 |
| 23 | 0.000 | 0.685 | 0.858 | 1.060 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 3.485 | 3.768 |
| 24 | 0.000 | 0.685 | 0.857 | 1.059 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 3.467 | 3.745 |
| 25 | 0.000 | 0.684 | 0.856 | 1.058 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 3.450 | 3.725 |
| 26 | 0.000 | 0.684 | 0.856 | 1.058 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 3.435 | 3.707 |
| 27 | 0.000 | 0.684 | 0.855 | 1.057 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 3.421 | 3.690 |
| 28 | 0.000 | 0.683 | 0.855 | 1.056 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 3.408 | 3.674 |
| 29 | 0.000 | 0.683 | 0.854 | 1.055 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 3.396 | 3.659 |
| 30 | 0.000 | 0.683 | 0.854 | 1.055 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.385 | 3.646 |
| 40 | 0.000 | 0.681 | 0.851 | 1.050 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 | 3.307 | 3.551 |
| 60 | 0.000 | 0.679 | 0.848 | 1.045 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 | 3.232 | 3.460 |
| 80 | 0.000 | 0.678 | 0.846 | 1.043 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 | 3.195 | 3.416 |
| 100 | 0.000 | 0.677 | 0.845 | 1.042 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 | 3.174 | 3.390 |
| 1000 | 0.000 | 0.675 | 0.842 | 1.037 | 1.282 | 1.646 | 1.962 | 2.330 | 2.581 | 3.098 | 3.300 |
| *z* | 0.000 | 0.674 | 0.842 | 1.036 | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 3.090 | 3.291 |
| | 0% | 50% | 60% | 70% | 80% | 90% | 95% | 98% | 99% | 99.8% | 99.9% |
| | **Confidence Level** | | | | | | | | | | |