Stat 112: Homework

1 Descriptive Statistics

1. A day care class consists of 16 children of ages three and four years old. We will refer to this class as the population, and we will identify each child by a number. The weight in pounds of each child is:

Child	Weight	Child	Weight	Child	Weight	Child	Weight
1	35	5	47	9	39	13	35
2	50	6	33	10	36	14	38
3	43	7	35	11	41	15	43
4	38	8	27	12	34	16	32

   (a) Draw a leaf and stem plot.

   (b) What is the range of the data?

   (c) Draw a histogram using an interval of 5.

   (d) Compute the population mean.

   (e) Compute the population variance.

   (f) Find the median.

   Rather than weigh all her children, the teacher decided to weigh just a few of them to save time. The night before, the teacher selected at random four names from the class roster. The names of these four children now constitute a sample of the population. Coincidentally, on the next day, all children were present for class, and no new children were added to the class. The roster, luckily, is in one-to-one correspondence with the population.

   The following four children were weighed:

Child	Weight
3	43
8	27
14	38
2	50

   (g) Compute the sample mean.

   (h) Compute the sample variance.

   (i) Find the sample median.

   (j) Compare these sample statistics with the population statistics to see how close they are to each other.

2. Prior to the 1970's, Iowa, Minnesota, and Wisconsin played a small role in world trade due to their economic policies but now they play an important role in the world economy. A table taken from *The Fleming Report, December 2000* of data in billions of dollars shows the relationship between mid-western America and industrial countries in terms of annual imports:

	1950	1960	1970	1980	1990	2000
Industrial Countries, x	39.8	85.4	226.9	1,370.2	1896	2,237.9
Middle America, y	21.2	21.7	25.5	248.4	534.8	819.4

   (a) Compute the following: $\sum_{i=1}^{6} x_i, \quad \sum_{i=1}^{6} x_i^2, \quad \sum_{i=1}^{6} x_i y_i, \quad \sum_{i=1}^{6} y_i, \quad \sum_{i=1}^{6} y_i^2$

   (b) Depict, that is, make a picture of the set of data which was taken from *The Fleming Report*.

3. Suppose two samples of the same size were drawn from the same population.

Account I	Account II
30	1
32	1
28	1
30	1
30	146

   (a) Compute the sample mean, sample variances, ranges, and medians for both accounts.

   (b) These accounts correspond to two different teams of five members who sell millions of dollars of contracts to the government. Which is the stronger of the two teams and explain your opinion? Or to put the question any way, which team would you hire based only on performance?

4. Use the statistics mode of your calculator to compute the sample mean and sample variance of the following data taken from STAT51 students:

   188 212 1432 318 209 243 378 233 250 247 290 201 294 238 250 306 232 267 300 240 400 402 212 308 296 280 200 234 284 261 405 241 257 290 265 379 311 379 300 380 215 375 318 229 248 427 1512 134 333 242 378 300 382 255 327 240 312 416

5. Virginia Department of Transportation traffic engineers were compelled to measure the speed of traffic on a residential street in a Northern Virginia suburb because the residents complained enough to their elected officials about speeding vehicles on their street. Reluctantly, the engineers measured the speeds of the motor vehicles and from 1,729 vehicles in 24 hours, they reported that the average speed was 28 MPH. Since the posted speed limit is 25 MPH, the engineers concluded that "the traffic conforms to the posted speed limit"; therefore, no remedial action was taken. However, some residents obtained the raw traffic data and discovered that the speeds of three vehicles exceeded 65 MPH, that 15% of the vehicles traveled faster than 40 MPH, and that 80% exceeded 25 MPH. If you are a state legislator representing the community, would you conclude that both factions or one of them or neither were distorting the truth with descriptive statistics? Explain. (All of the numbers are reliable.)

2 Confidence Intervals

1. A random sample of 10 observations is drawn from a normal population. The sample mean, $\bar{x} = 4$ and the sample variance, $s^2 = .5$. Find the 95% confidence interval for the mean, μ.

2. The demand for meat at a grocery store for a weekend is normally distributed with mean equal to 4900 pounds and standard deviation equal to 500 pounds. Find the 90% confidence interval for the expected demand.

3. Given the sample, $\mathcal{S} = \{3\ 5\ 1\ 5\ 5\ 1\ 5\ 4\ 0\ 4\ 5\ 4\ 5\ 6\ 2\ 6\ 1\ 4\ 5\ 6\}$

   (a) Find the 95% confidence interval of $E[\bar{x}]$. (Hint: use Student's t-distribution).

   (b) Draw a histogram of $\mathcal{S}$.

   (c) Draw 95% confidence interval of $E[\bar{x}]$ on the histogram.

4. A day care class consists of 12 children of ages three and four years old. We will refer to this class as a sample. The weight in pounds of each child is:

   $\mathcal{S} = \{30\ 50\ 40\ 40\ 45\ 35\ 35\ 25\ 40\ 35\ 40\ 30\}$

   Find the 95% confidence interval about the population mean.

5. From a list of 90,000 farmers who operate a farm in Ohio, a sample of 2,000 is drawn, but only 64% of them cooperate in giving an interview. Let X be the number of useful interviews. Find two numbers, a and b, which are symmetric about X such that $P(a \leq X \leq b) = 95\%$. (Hint: There is only one state , so n=1.) Note: $z_{.025} = 1.96$

6. The grade point averages (GPA) of a sample of 100 students were obtained. Denote the GPA of a student by X_i. From the data, it was found that $\bar{x} = 3.5$ and $s = .5$. Find the 90% confidence interval about the population mean.

7. Find $z_{\frac{\alpha}{2}}$ for $\alpha = .08$

8. The following sample of 16 measurements was selected from a population that is approximately normally distributed:
   $\mathcal{S} = \{91\ 80\ 99\ 110\ 95\ 106\ 78\ 121\ 106\ 100\ 97\ 82\ 100\ 83\ 115\ 104\}$

   (a) Construct a 80% confidence interval for the population mean.

   (b) Interpret the meaning of this confidence interval for your STAT51 professor.

   (c) The 95% confidence interval is: (91.19876,104.6762). Explain why the 80% confidence interval is narrower than the 95% confidence interval.

9. A random sample of 49 observations is drawn from a normal population with mean equal to 50 and $\sigma = 15$. Find c such that P($\bar{x} \leq c$)=.89.

10. Suppose T_{10} is a Student's t distribution with 10 degrees of freedom. Find t_0 such that $P(T_{10} \leq t_0) = .05$.

3 Testing the Hypothesis between a Mean and a Constant

1. A manufacturer of sports equipment has developed a new synthetic fishing line that he claims based on numerous experiments has a mean braking strength of 15 pounds with a standard deviation of .5 pounds. Test the hypothesis that $\mu = 15$ against the alternative that $\mu \neq 15$ if a random sample of 50 lines is tested and found to have a mean braking strength of 14.8 pounds. That is, should the null hypothesis be rejected or accepted? Use a .01 level of significance.

2. A random sample of n=25 observations is drawn from a normal population with standard deviation equal to 9. The sample mean was found to be, $\bar{x} = 16.5$ Test the hypothesis: $H_0 : \mu = 15$ vs $H_1 : \mu > 15$ at the .05 level of significance.

3. According to the American Medical Association, approximately 50% of all U. S. physicians whose ages are under 35 are known to be women. Four samples of 15 physicians were interviewed each. Let X_i be the number of women that were found in the i^{th} random sample of 15 physicians whose ages are less than 35. The number of women in each of the four samples were counted and it was found that $\bar{x} = 5$ and $s^2 = \frac{2}{3}$. What is the 95% confidence interval about μ? (Hint: Let X_i be approximately distributed as a normal, $N(\mu, \sigma^2)$. Use n=4 in the formula for the lower and upper limits and the t distribution).

4. Continuation of problem 3. Given that $\bar{x} = 5$ and $s^2 = \frac{2}{3}$ from the 4 samples of 15 physicians each, test the hypothesis that $\mu = 7.5$ against the alternative that $\mu \neq 7.5$ at the .05 level of significance. (Hint: Use the test where σ^2 is unknown and use n=4 in the formula of the test statistic).

5. Given the sample, $\mathcal{S} = \{3\ 5\ 1\ 5\ 5\ 1\ 5\ 4\ 0\ 4\ 5\ 4\ 5\ 6\ 2\ 6\ 1\ 4\ 5\ 6\}$. Test the hypothesis: $H_0 : \mu = 3$ vs $H_1 : \mu \neq 3$ at the .05 level of significance.

6. Let t_0 be a specific value of t. Find t_0 such that $P(t \geq t_0)$=.01 where df=14.

7. The temperature in Fahrenheit of six ovens which are used in a manufacturing process of cintering ferrite yokes for television sets were recorded $\mathcal{S} = \{1325\ 1323\ 1340\ 1331\ 1328\ 1323\}$. Test the hypothesis : $H_0 : \mu = 1325$ vs $H_1 : \mu \geq 1325$ at the .05 level of significance when the p-value is 0.1279207.

8. A random sample of 20 observations selected from a normal population produced $\bar{x} = 72.6$ and $s^2 = 19.4$. Test $H_0 : \mu = 80$ vs $H_1 : \mu < 80$. Use $\alpha = .05$

9. A professor of social work observed that the attitude of students with regard to social issues seems to change in the span of time from senior year in college to receiving a masters degree. To test his observation, the professor asked 10 students who were in their senior year the question, "Is there too little space exploration?". The answers were reported using a five point scale, 1 strongly disagree to 5 strong agree, they appear in Table 9. All 10 students were women. Test the null hypothesis that going to graduate school changes one's attitude on the issue of funding the space program against the alternative that going to graduate school does not cause a change in attitude at a level of significance of .05.

Student	Before Bachelor's Degree x	After Masters Degree y	Difference d=x-y
1	2	5	
2	3	2	
3	2	4	
4	2	3	
5	2	3	
6	3	5	
7	5	3	
8	3	5	
9	2	2	
10	1	2	
mean	2.5	3.4	
std	1.166667	1.6	

10. STAT51 students counted the number of e's for an experiment in counting on the first day of class. The average number of e's from the 47 students is 218.957 with a sample standard deviation of 71.143. The hypothesis is that the class average equals the true value of 315 versus the alternative hypothesis that the class average is less than the true value. Use a .05 level of significance and only the p-value=2.25×10^{-12} to decide whether to reject or not reject the null hypothesis.

11. Construct a 90% confidence interval about the mean for the number of e's which the STAT51 students in the previous problem counted. Why is the confidence interval for the STAT51 students so far away from the true number of e's?

4 Testing the Hypothesis between Two Means

1. An amateur gardener wanted to discover whether one fertilizer is better than another when applied to his tomato plants. Having taken a statistics course in the design of experiments, this gardener knew that conditions for all 11 of his tomato plants must be the same so that the effects of applying the two different fertilizers will not be confounded by something else. Therefore, he planted the 11 tomato plants of the same hybrid in the same plot with the same exposure to sunlight, etc. He applied fertilizer A to five tomato plants which were chosen at random, and he applied fertilizer B to six other tomato plants. At harvest, the average weight of a tomato from each plant was obtained and is tabulated below:

   Is there a difference in the yield due to the different fertilizers at the .05 level of significance? (Hint:Test the hypothesis:$H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$).

2. A farmer claims that the average yield of corn of variety A exceeds the average yield of variety B by at least 12 bushels per acre. To test this claim, 20 one acre plots of each variety are

Average Weight of a Tomato	
Fertilizer A	Fertilizer B
29.9 oz	26.6 oz
11.4	23.7
25.3	28.5
16.5	14.2
21.1	17.9
.	24.3

planted and grown under similar conditions. Variety A yielded, on the average, 86.7 bushels per acre with a sample standard deviation of 6.28 bushels per acre, while variety B yielded, on the average, 77.8 bushels per acre with a sample standard deviation of 5.61 bushels per acre. Test the farmer's claim using a .05 level of significance. (Hint: Test the hypothesis: $H_0 : \mu_1 - \mu_2 = 12$ vs $H_1 : \mu_1 - \mu_2 > 12$).

3. Let $\bar{x}_1 = 61,600$, $n_1 = 12$, and $s_1 = 3300$. Let $\bar{x}_2 = 64,000$, $n_2 = 12$, and $s_2 = 3500$. Assume that $\sigma_1^2 = \sigma_2^2$. Test the hypothesis:

   $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$ at $\alpha = .05$

4. A Cuckoo will lay eggs in the nests of other birds. It is claimed by some biologists that the size of an egg which a Cuckoo will lay conforms to the size of the egg of a warbler or of a wren depending in which nest the Cuckoo lays her egg. The following descriptive statistics given in Table 1 were obtained for Cuckoo eggs which were found in randomly sampled nests of warblers and wrens. Test the hypothesis: $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$ at a level of significance of .01 where #1 designates the population of warbler nests and #2 designates the population of wren nests.

Table 1:

	#1	#2
n	29	35
$\bar{x}$	22.20 mm	21.12 mm
s	.65 mm	.75 mm

5. Residents usually welcome the installation of speed humps on their streets in order to reduce speeding traffic whereas traffic engineers despise speed humps because by impeding cut-through traffic they divert cut-through traffic onto the arterial routes. The engineers recorded the speeds before and after the installation of speed humps to see if they have an effect. A tabulation of speeds on some selected streets are given below. At a level of $\alpha = .10$, can the engineers conclude that speed humps on the average reduce the speed of traffic by 10 mph?

Street	Before	After
Northborne Drive	45 mph	33 mph
Kings Park Drive	50	27
Hollinger Avenue	43	28
Springhaven Drive	40	30

6. A company bakes computer chips in two ovens, A and B. Chips are randomly assigned to an oven and hundreds of chips are baked each hour. The percentage of defective chips are tabulated below. For example, oven A produced chips which were 45% defective during hour 1. Does there appear a difference between oven A and oven B with respect to the rate of producing defective chips at a 95% level of significance?

	1	2	3	4	5	6	7	8	9
A	45%	32	34	31	35	37	31	30	27
B	36%	37	33	34	33	32	33	30	24