AP Statistics Special Problems #2 Inference for Means

1. NORMAL TEMPERATURE A medical researcher measured the body temperatures of a sample of randomly selected adults. The data he collected are summarized below. We wish to estimate the average (or "normal") temperature among the adult population.

Summary	Temperature
Count (*n*)	52
Mean ($\bar{x}$)	98.285
Median	98.2
MidRange	98.6
StdDev (*s*)	0.6824
Range	2.8
IQR	1.05

a) Are the necessary conditions for a t-interval satisfied? Explain.
b) Find and interpret a 98% confidence interval for mean body temperature.
c) Explain what "98% confidence" means in this context.
d) 98.6°F is commonly assumed to be "normal." Do these data suggest otherwise? Explain.

2. SAT SCORES High school students who take the SAT Mathematics exam a second time generally score higher than on their first try. A random sample of 100 students gains an average of $\bar{x} = 22$ points on their second try with a standard deviation of 12 points.
(a) Construct and interpret a 95% confidence interval for the mean score gain μ in the population (Show your work). What is the margin of error?
(b) Calculate the 90% and 99% confidence intervals for μ. (Just use your calculator to get the interval).
(c) Make a sketch to compare these three intervals on the same axis (number the axis from 18 to 26). How does increasing the confidence level affect the length of the confidence interval?

3. PIZZA A researcher tests whether the mean cholesterol level among those who eat frozen pizza exceeds the value considered to indicate a health risk. She gets a p-value of 0.07. Explain in context what the p-value of 7% means.

4. STATISTICAL SIGNIFICANCE Explain in plain language why a significance test that is significant at the 1% level must always be significant at the 5% level. If a test is significant at the 5% level, what can you say about its significance at the 1% level?

5. LARGE FRIES PLEASE! A statistics class at the University of Wisconsin estimated the average mass of a large bag of French fries at McDonald's. They bought 30 bags during two different half-hour periods on two consecutive days at the same McDonald's and weighed the fries. The data are given below. McDonald's "target value" for the mass of a large order of fries was 171 g.
a) Is there significant evidence that this McDonald's was not meeting the "target value" of 171 grams? Provide statistical evidence to support your answer.
b) How strong is that evidence? Explain.
c) What if you went to the restaurant and complained to the manager and he fired the employees at that time shift. What type of error (Type I or Type II) would you make? How serious of an error is it?

125.7	139.9	132.9	131.6	152.1	151.3	145.2	138.6	142.7
132.8	127.5	141.1	152.4	145.8	143.8	155	162.7	140.2
145.7	156.6	154.3	143.7	115.9	162	175.4	139.6	153.8
133.6	142.2	137.9						

6. CHIPS AHOY In 1998, as an advertising campaign, the Nabisco Company announced a "1000 Chips Challenge" claiming that every 18-ounce bag of their Chips Ahoy cookies contained at least 1000 chocolate chips. Dedicated Statistics students at the Air Force Academy purchased some randomly selected bags of cookies, and counted the chocolate chips. Some of their data are given below. (*Chance*, 12, no. 1 [1999])

1219 1214 1087 1200 1419 1121 1325 1345
1244 1258 1356 1132 1191 1270 1295 1135

a) Check the conditions for inference. Comment on any concerns you have.
b) Construct and interpret a 95% confidence interval for the average number of chips in 18-ounce bags of Chips Ahoy cookies.
c) Is there enough evidence to claim that there at least 1000 chips in an 18-ounce Chips Ahoy cookies bag? **Use only your confidence interval** to test an appropriate hypothesis and state your conclusion.

7. WHAT TYPE OF SAMPLE DO WE HAVE? The following situations contain either a matched pairs t test **or** a two-sample t test. Which procedure would you choose in each of the following situations?
a) To test the wear characteristics of two tire brands, A and B. Brand A is mounted on 50 cars and Brand B on 50 other cars.
b) To test the wear characteristics of two tire brands, A and B, one Brand A tire is mounted on one side of each car in the rear, while Brand B tire is mounted on the other side. Which side gets which brand is determined by flipping a coin? The same procedure is used on the front.
c) To test the effect of background music on productivity, factory workers are observed. For one month they had no background music. For another month they had background music.
d) A random sample of 10 workers in Plant A are to be compared to a random sample of 10 workers in Plant B in terms of productivity.
e) A new weight-reducing diet was tried on 10 women. The weight of each woman was measured before the diet and again after 10 weeks on the diet.
f) An education researcher wants to learn whether it is more effective to put questions before or after introducing a new concept in an elementary school mathematics text. He prepares two test segments that teach the concept, one motivating questions before and the other with review questions after. He uses each test segment to teach a separate group of children.
g) Anther researcher approaches the same issue differently. She prepares test segments on two unrelated topics. Each segment comes in two versions, one with questions before and the other with questions after. The subjects are a single group of children. Each child studies both topics, one with questions before and the other with questions after.

8. STUDENTS' ATTITUDES TOWARD PARENTS Researchers at the University of South Alabama compared the attitudes of male college students toward their fathers with their attitudes toward their mothers (*Journal of Genetic Psychology*, March 1998). Each of a sample of 13 males was asked to complete the following statement about each of their parents: "My relationship with my father (mother) can be best described as (1) Awful, (2) Poor, (3) Average, (4) Good, and (5) Great." The following data were obtained:

Student	Attitude toward Father	Attitude toward Mother
1	2	3
2	5	5
3	4	3
4	4	5
5	3	4
6	5	4
7	4	5
8	2	4
9	4	5
10	5	4
11	4	5
12	5	4
13	3	3

Is there evidence that there is a difference in a student's attitude between their father and their mother, at the 1% significant level? Provide statistical evidence to support your answer.

9. MCDONALD'S VERSUS WENDY'S A student wanted to determine whether the wait time in the drive-through at McDonald's differed from that at Wendy's. She used a random sample of 30 cars at McDonald's and 27 cars at Wendy's and obtained these results:

Wait Time at McDonald's Drive Through (in seconds)

151.09	227.38	111.84	131.21	128.75	191.6	126.91	137.9	195.44
246.59	141.78	127.35	121.21	101.03	95.09	122.06	122.62	100.04
71.37	153.34	140.44	126.62	116.72	131.69	100.94	115.66	147.28
81.43	86.31	156.34						

Wait Time at Wendy's Drive Through (in seconds)

281.9	71.02	204.29	128.59	133.56	187.53	199.86	190.91	110.55
110.64	196.84	233.65	171.01	182.54	183.79	284.48	363.34	270.82
390.5	471.62	123.66	174.43	385.9	386.71	155.53	203.62	119.61

a) Is there significant evidence at the 5% significant level that there is a difference in wait time at each restaurant's drive-through? Provide statistical evidence to support your answer.
b) **Now** use a 95% confidence interval to support your answer to part (a).
c) Draw parallel boxplots on the same axis. Does this visual evidence support the results obtained in parts (a) and (b)? Explain.
d) Why do you think there is a huge difference in the wait times between these restaurants? (It is based on company policy)
e) Based on your answer in part (a), explain making a Type I and a Type II error.

10. TV SAFETY The manufacturer of a metal stand for home TV sets must be sure that its product will not fail under the weight of the TV. Since some larger sets weigh nearly 300 pounds, the company's safety inspectors have set a standard of ensuring that the stands can support an average of over 500 pounds. Their inspectors regularly subject a random sample of the stands to increasing weight until they fail. They test the hypothesis $H_o : \mu = 500$ against $H_a : \mu > 500$, using the level of significance $\alpha = 0.01$. If the sample of stands fails to pass this safety test, the inspectors will not certify the product for sale to the general public.
a) Is this an upper-tail or lower-tail test? In the context of the problem, why do you think this is important?
b) Explain what will happen if the inspectors commit a Type I error.
c) Explain what will happen if the inspectors commit a Type II error.
d) Which error is worse?

11. BEETLES IN THE WOODS An outbreak of the mountain pine beetle has affected several types of pine trees in British Columbia. The beetle leaves behind a fungus that produces blue-colored stains in the wood. Some consumers might worry the lumber obtained from the blue-stained trees is weaker as a result of the effects of the fungus. A Canadian company performed a test on the breaking strength of blue-stained wood. They measured the mean breaking strength of a sample of 100 pieces of blue-stained pine. The target breaking strength of lumber from healthy pine trees is 10,000 pounds per square inch (psi).
a) State appropriate null and alternative hypotheses for the company to test.
b) Describe a Type I and Type II error and give the consequences of each?
c) Which type of error is more serious? Why?