Question 5

Intent of Question

The primary goals of this question were to assess students' ability to (1) state the appropriate hypotheses, (2) identify and compute the appropriate test statistic, (3) make a conclusion in the context of the problem for a one-sample t test, and (4) use simulation results to find a simulated p-value to make an inference about the standard deviation.

Solution

Part (a):

Step 1: State a correct pair of hypotheses.

H_0: $\mu = 12.1$
H_a: $\mu \neq 12.1$, where μ is the mean number of fluid ounces dispensed into all juice bottles filled in the past hour

Step 2: Identify a correct test (by name or by formula) and check appropriate conditions. (Stem of the question said to assume that conditions for inference are met.)

One-sample t test for a mean *OR* $t = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Step 3: Correct mechanics, including the value of the test statistic, df, and p-value (or rejection region).

test statistic: $t = \dfrac{12.05 - 12.1}{\frac{0.085}{\sqrt{4}}} = \dfrac{-0.05}{0.0425} = -1.176$

p-value: $2 \cdot P(T_{3df} < -1.176) = 0.324$

Step 4: State a correct conclusion in the context of the problem, using the result of the computations.

Because the p-value of 0.324 is larger than any reasonable significance level, such as $\alpha = .05$, do not reject the null hypothesis that the mean number of fluid ounces being dispensed is 12.1 fluid ounces. There is not sufficient evidence to conclude that the machine is filling the juice bottles with an average amount different from 12.1 fluid ounces.

Part (b):

In 300 simulated sample standard deviations, the value of the computed standard deviation (0.085) from our sample in part (a) or a value larger than 0.085 occurred only 12 times. This is a simulated p–value of $\frac{12}{300}$ or 0.04. If the actual population standard deviation is 0.05, then we estimate that the chance of observing a sample standard deviation of 0.085 or larger is 4 percent.

Because this simulated p-value is less than a significance level of 5 percent, the sample from part (a) provides strong evidence that the standard deviation of the juice being dispensed exceeds 0.05 ounces.

Scoring

This problem is scored in four sections. Section 1 consists of part (a), step 1. Section 2 consists of part (a), steps 2 and 3. Section 3 consists of part (a), step 4. Section 4 consists of part (b). Sections 1, 2, and 3 are each scored as essentially correct (E) or incorrect (I). Both step 2 and step 3 in section 2 must be completely correct to earn an E for section 2. Section 4 is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 [part (a), step 1] is scored as follows:

Essentially correct (E) if the student states a correct pair of hypotheses.

Incorrect (I) otherwise.

Section 2 [part (a), steps 2 and 3] is scored as follows:

Essentially correct (E) if the student identifies a correct test (by name or formula) and includes correct mechanics.

Incorrect (I) otherwise.

Section 3 [part (a), step 4] is scored as follows:

Essentially correct (E) if the student states a correct conclusion in the context of the problem.

Incorrect (I) otherwise.

Section 4 [part (b)] is scored as follows:

Essentially correct (E) if the student finds the correct simulated p-value from the dotplot or states that the actual standard deviation of 0.085 would be unusual if $\sigma = 0.05$ because 0.085 lies in the tail of the distribution *AND* an appropriate conclusion is made.

Partially correct (P) if only the simulated p-value or statement that 0.085 lies in the tail is correct, but the conclusion is weak, wrong, or missing.

Incorrect if the value 0.085 is not linked to the distribution.

Notes:

- If the p-value in section 2 is incorrect but the conclusion is consistent with the computed p-value, section 4 can be considered correct.
- In section 4, if both an α and a p-value are given together, the linkage between the p-value and the conclusion is implied. If no α is given, the solution must be explicit about the linkage by giving a correct interpretation of the p-value or explaining how the conclusion follows from the p-value.

Question 5 (continued)

Each essentially correct (E) response counts as 1 point, and a partially correct (P) response in section 4 counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 1½ points), use a holistic approach to determine whether to score up or down, depending on the strength of the response and communication.

5. A bottle-filling machine is set to dispense 12.1 fluid ounces into juice bottles. To ensure that the machine is filling accurately, every hour a worker randomly selects four bottles filled by the machine during the past hour and measures the contents. If there is convincing evidence that the mean amount of juice dispensed is different from 12.1 ounces or if there is convincing evidence that the standard deviation is greater than 0.05 ounce, the machine is shut down for recalibration. It can be assumed that the amount of juice that is dispensed into bottles is normally distributed.

During one hour, the mean number of fluid ounces of four randomly selected bottles was 12.05 and the standard deviation was 0.085 ounce.

(a) Perform a test of significance to determine whether the mean amount of juice dispensed is different from 12.1 fluid ounces. Assume the conditions for inference are met.

As the conditions are met, procede with a one sample t-test for popul. mean as ~~both~~ sample size $(n=4)$ is small and popul. variance is uncertain.

$H_0: \mu_0 = 12.1$ $\quad \bar{X} = 12.05$

$H_a: \mu \neq 12.1$ $\quad s = 0.085$

$n = 4$

$$\cancel{t = 12.05} \quad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = -1.176$$

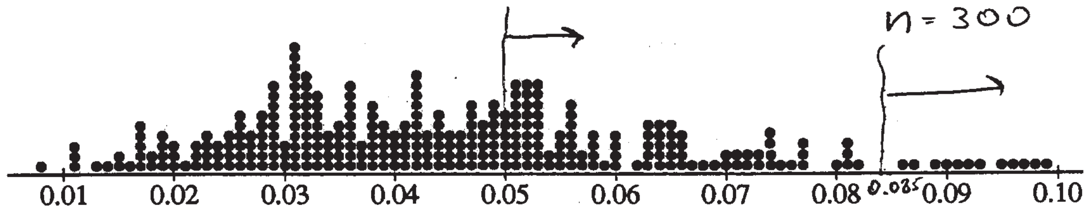
$p\text{-val} = 0.324$

With such a high p-value (greater than $\alpha = 0.1$), we ~~do~~ not have evidence to ~~reject~~ H_0. Thus, we cannot claim that the mean amount of juice has changed.

GO ON TO THE NEXT PAGE.

SA2

(b) To determine whether this sample of four bottles provides convincing evidence that the standard deviation of the amount of juice dispensed is greater than 0.05 ounce, a simulation study was performed. In the simulation study, 300 samples, each of size 4, were randomly generated from a normal population with a mean of 12.1 and a standard deviation of 0.05. The sample standard deviation was computed for each of the 300 samples. The dotplot below displays the values of the sample standard deviations.



Use the results of this simulation study to explain why you think the sample provides or does not provide evidence that the standard deviation of the juice dispensed exceeds 0.05 fluid ounce.

The probability that if true standard deviation is correct (is $\sigma = 0.05$), then the relative frequency of samples with the same size that will result in a stand. deviation of 0.085 or even more extreme is $\frac{12}{300} = 0.04$. This is approximately the corresponding theoretical probability (the p-value of the test with $H_0: \sigma = 0.05$ versus $H_a: \sigma > 0.05$). Thus, on $0.05 = \alpha$ signif. level we have enough evidence to reject H_0 and to say that stand. deviation of the dispenser has changed and has increased,

as p-val $\approx 0.04 < \alpha = 0.05$

GO ON TO THE NEXT PAGE.

5. A bottle-filling machine is set to dispense 12.1 fluid ounces into juice bottles. To ensure that the machine is filling accurately, every hour a worker randomly selects four bottles filled by the machine during the past hour and measures the contents. If there is convincing evidence that the mean amount of juice dispensed is different from 12.1 ounces or if there is convincing evidence that the standard deviation is greater than 0.05 ounce, the machine is shut down for recalibration. It can be assumed that the amount of juice that is dispensed into bottles is normally distributed.

During one hour, the mean number of fluid ounces of four randomly selected bottles was 12.05 and the standard deviation was 0.085 ounce.

(a) Perform a test of significance to determine whether the mean amount of juice dispensed is different from 12.1 fluid ounces. Assume the conditions for inference are met.

It is assumed that the conditions for inference are met.

$\mu_0 = 12.1$, $\bar{X} = 12.05$, $\sigma = 0.05$, $n = 4$

Z-Test

$H_0: \mu = \mu_0$

$H_a: \mu \neq \mu_0$ $\qquad z = \dfrac{\bar{X} - \mu_0}{\sigma}$

$$z = \frac{12.05 - 12.1}{0.05} = -2$$

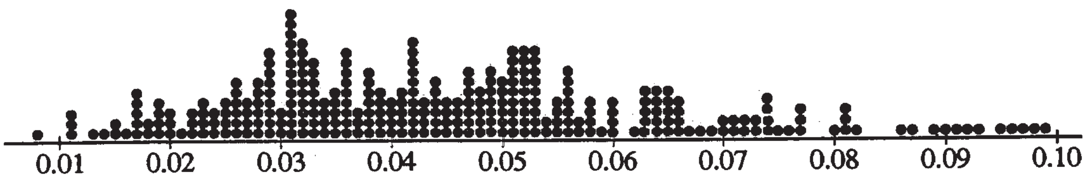
$2 \times P(t < -2) = 2 \times 0.02275 =$ PValue $= 0.0455$.

Thus, the null hypothosis can be rejected at $\alpha = 0.05$ and the is significant evidence suggesting that the mean amount of juice is different from 12.1 fluid ounce.

GO ON TO THE NEXT PAGE.

5B2

(b) To determine whether this sample of four bottles provides convincing evidence that the standard deviation of the amount of juice dispensed is greater than 0.05 ounce, a simulation study was performed. In the simulation study, 300 samples, each of size 4, were randomly generated from a normal population with a mean of 12.1 and a standard deviation of 0.05. The sample standard deviation was computed for each of the 300 samples. The dotplot below displays the values of the sample standard deviations.



Use the results of this simulation study to explain why you think the sample provides or does not provide evidence that the standard deviation of the juice dispensed exceeds 0.05 fluid ounce.

~~The standard~~

The distribution of the stimulation suggests that the center of the distribution is between 0.035 and 0.045. Moreover, this ~~graph~~ distribution is nearly symatrical and bell shaped.

We can also see that only 12 of the distribution out of the 300 simulation exceeds the value 0.085. Thus, considering ~~that~~ the center, shape ~~on~~ of the distribution, and the fact that only 8 of 300 stimulation's standard deviation exceeds 0.085, the distribution suggest evidence that the standard deviation of the juice dispensed exceeds 0. of fluid ounce.

GO ON TO THE NEXT PAGE.

5C1

5. A bottle-filling machine is set to dispense 12.1 fluid ounces into juice bottles. To ensure that the machine is filling accurately, every hour a worker randomly selects four bottles filled by the machine during the past hour and measures the contents. If there is convincing evidence that the mean amount of juice dispensed is different from 12.1 ounces or if there is convincing evidence that the standard deviation is greater than 0.05 ounce, the machine is shut down for recalibration. It can be assumed that the amount of juice that is dispensed into bottles is normally distributed.

During one hour, the mean number of fluid ounces of four randomly selected bottles was 12.05 and the standard deviation was 0.085 ounce.

(a) Perform a test of significance to determine whether the mean amount of juice dispensed is different from 12.1 fluid ounces. Assume the conditions for inference are met.

a). $H_0: \mu_{\bar{x}} = 12.1$ where $\mu_{\bar{x}}$ is the mean amount of 4 sample juice

$H_a: \mu_{\bar{x}} \neq 12.1$ ~~and μ_{juice} is the mean amount of juice dispensed., which is 12.1~~

use ~~two~~ one sample t-test.

check the condition

we note the ~~sample~~ is a simple random sample

and we assume that the conditions for inference are met

$$t = \frac{\bar{X} - \mu_{juice}}{\frac{S_x}{\sqrt{n}}} = \frac{12.05 - 12.1}{\frac{0.085}{\sqrt{4}}} = -1.176$$

$df = 4 - 1 = 3$

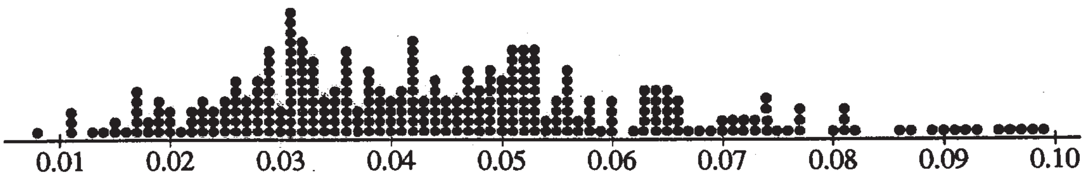
p-value = .3242

Since the p-value is large, (p-value > 0.1) we have not enough evidence to indicate that the mean of 4 sample-juice dispensed is different from 12.1 fluid ounces.

GO ON TO THE NEXT PAGE.

5C2

(b) To determine whether this sample of four bottles provides convincing evidence that the standard deviation of the amount of juice dispensed is greater than 0.05 ounce, a simulation study was performed. In the simulation study, 300 samples, each of size 4, were randomly generated from a normal population with a mean of 12.1 and a standard deviation of 0.05. The sample standard deviation was computed for each of the 300 samples. The dotplot below displays the values of the sample standard deviations.



Use the results of this simulation study to explain why you think the sample provides or does not provide evidence that the standard deviation of the juice dispensed exceeds 0.05 fluid ounce.

b. the dotplot shows a roughly skewed to right plots, the center of the data is approximately between 0.03–0.04 and the mean is just a little ~~so big~~ larger than median. maybe around 0.04. +① as a result, we can say that the samples does not provide the evidence that the standard deviation of the juice dispensed exceeds 0.05 fluid ounce.

①. What's more, the ~~varibility~~ variability of the plots is relatively small although there are some outliers at the tail.

Question 5

Sample: 5A
Score: 4

In this response, all of the steps in the test of significance in part (a) (which comprises sections 1, 2, and 3 for scoring purposes) were each scored as essentially correct. However, the student gives two reasons for using a one-sample *t* test—small sample size and uncertain population variance—but only the second is a valid reason. A *t* test is called for because the population standard deviation is unknown (it may have changed from 0.05 ounce), not because the sample size is small. The test of significance also would have been stronger if the parameter μ were defined as the mean amount of juice dispensed into all bottles by the machine in the past hour. These were considered minor errors. The conclusion is linked to the computations and written in the context of the situation. Section 4 [part (b)] is beautifully done and also was scored as essentially correct. The location of 0.085 is clearly marked on the dotplot. The estimated *p*-value is correctly computed and described as the proportion of values that are as large or larger than 0.085 (12/300) if $\sigma = 0.05,$ which is correct. The response then links this *p*-value to the conclusion that the null hypothesis, $\sigma = 0.05,$ must be rejected in favor of the conclusion that the standard deviation of the machine has increased. Because all the steps in sections 1–4 were essentially correct, this complete response received a score of 4.

Sample: 5B
Score: 3

This response includes the required steps in the test of significance in part (a) (which comprises sections 1, 2, and 3 for scoring purposes). Again, the test of significance would have been stronger if the parameter μ were defined as the mean amount of juice dispensed into bottles by the machine in the past hour. The student then chooses to calculate a *z* test statistic to test the hypothesis about the population mean, which would be appropriate if the population standard deviation, σ, were known. However, σ is not known, and the student incorrectly uses a σ value of 0.05. A more serious error, in mechanics, is that the student does not include the divisor $\sqrt{n}$ in the denominator of the test statistic. Thus section 2 [part (a), steps 2 and 3] was scored as incorrect. The other portions of this part—section 1 [part (a), step 1] and section 3 [part (a), step 4]—were scored as essentially correct. In section 4 [part (b)] the value 0.085 is located in the extreme right tail of the distribution, because only 12 out of the 300 simulated values exceed the value of 0.085, and the correct conclusion is drawn that there is evidence that the standard deviation of the juice dispensed exceeds 0.05 fluid ounce. (The fact that 12 out of 300 is later stated as "8 of 300" was considered an inconsequential slip.) Although the description of the shape of the distribution (symmetric) is not accurate, this part was scored as essentially correct because the student correctly states the meaning of the distribution and the meaning of the location of 0.085 on it. Because sections 1, 3, and 4 were all essentially correct, this substantial response received a score of 3.

Sample: 5C
Score: 2

In part (a) of this response the hypotheses and the conclusion are unclear. They appear to be about the mean of the sampling distribution of the means of the amount of juice in samples of four bottles, but they may be about the specific sample of four bottles. The parameter μ should represent the mean of the population of amounts of juice dispensed in the bottles from which the sample of four bottles was taken. Thus section 1, the hypotheses portion of part (a), was scored as incorrect. The correct test is identified by name, and correct mechanics are used, so section 2 was scored as essentially correct. Section 3, the

Question 5 (continued)

conclusion step, which is in the context of the situation and follows from the *p*-value of ".3242," was also scored as essentially correct. Section 4 [part (b)] was scored as incorrect because the discussion is entirely about the shape, center, and variability of the simulated distribution. The description of this distribution, rather than the location of 0.085, is used to decide whether there is evidence that the standard deviation is larger than 0.05. The sample standard deviation of 0.085 is not mentioned nor linked to this distribution. Because sections 2 and 3 were both essentially correct and sections 1 and 4 were incorrect, this developing response received a score of 2.