Question 6

Intent of Question

The primary goals of this question are to assess a student's ability to (1) explain how to determine whether a sampling distribution is approximately normal; (2) calculate the standard error for the sampling distribution of a proportion and compare it to a value; (3) calculate the standard error for the average of two independent random variables; (4) conduct a test of hypotheses in a non-standard situation, using a rule called Chebyshev's inequality.

Solution

Part (a):

It is reasonable to assume that the distribution is approximately normal. The required condition is that there are at least 10 successes and 10 failures in the sample. In this case there are 44 defective lightbulbs and 356 non-defective lightbulbs, thus both exceed the minimum of 10 required.

Part (b):

Note that $\hat{p}_X = \frac{44}{400} = 0.11.$ So the standard error of $\hat{p}_X$ is $\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n}} = \sqrt{\frac{(0.11)(0.89)}{400}} = 0.0156.$

Part (c):

Since $\hat{p}_X = 0.11$, $0.11 - 0.10 = 0.01$, so that $\hat{p}_X$ is $\frac{0.01}{0.0156} = 0.64$ standard error away from 0.10.

Part (d):

(i) First compute $\hat{p}_Y = \frac{104}{400} = 0.26.$ So $\hat{D} = \frac{0.11 + 0.26}{2} = 0.185.$

(ii) The standard error of $\hat{p}_X = 0.0156$ is obtained from part (b). The standard error of $\hat{p}_Y$ is $\sqrt{\frac{\hat{p}_Y(1-\hat{p}_Y)}{n}} = \sqrt{\frac{(0.26)(0.74)}{400}} = 0.0219.$ So the standard error of $\hat{D}$ is $s_{\hat{D}} = \sqrt{\frac{1}{4}\left(0.0156^2 + 0.0219^2\right)} = 0.0134.$

Part (e)

$$W = \frac{0.185 - 0.10}{0.0134} = 6.34.$$

Question 6 (continued)

Part (f)

Suppose the true mean D is 0.10. Then the observed value of $\hat{D} = 0.185$ is 6.34 standard errors from the mean D. Using Chebyshev's inequality, the probability of observing a value of $\hat{D}$ within 6.34 standard errors of the mean of 0.10 is at least $1 - \frac{1}{6.34^2} = 0.975$. So the probability of observing a value as far from 0.10 as the one observed, or farther, is at most 0.025 if the true mean really is 0.10. Therefore, the p-value for this test is at most 0.025, which is less than 0.05, so the null hypothesis can be rejected. There is sufficient statistical evidence at the 0.05 level to conclude that the average proportion for all products that are defective is greater than 0.10.

<u>Scoring</u>

This question is scored in three sections. Section 1 consists of parts (a), (b) and (c), section 2 consists of part (d), and section 3 consists of parts (e) and (f). Each section is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. In part (a), states that it is reasonable to assume the condition is met *AND* provides appropriate justification by comparing the number of defectives and non-defectives to a reasonable number.
2. In part (b), calculates $\hat{p}_X$ correctly, either separately or in the process of showing the computation for its standard error.
3. In part (b), gives the correct formula for the standard error of $\hat{p}_X$.
4. In part (c), states the correct number of standard errors $\hat{p}_X$ is from 0.10.

Partially correct (P) if the response includes only two or three of the four components.

Incorrect (I) if the response includes at most one of the four components.

Notes

- In component 1, it is acceptable to check the sample size condition using $400 \times 0.1 = 40$ instead of the observed value of 44, because the sampling distribution refers to all possible samples.
- In component 4, transcription errors should not penalize a response if there is no ambiguity in how the error occurred. For instance, $\frac{0.11 - 0.10}{0.0156} = \frac{0.1}{0.0156} = 6.41$ is an acceptable transcription error.

Question 6 (continued)

Section 2 is scored as follows:

Essentially correct (E) if the response includes the following four components in part (d):

1. Correctly calculates $\hat{p}_Y$ either separately or in the process of showing the computation of $\hat{D}$.
2. Correctly calculates $\hat{D}$.
3. Gives the correct formula for the standard error of an average of two independent random variables.
4. Correctly computes the standard error of $\hat{D}$ *OR* if an incorrect but reasonable formula is given for the standard error, plugs the correct values into that formula. For instance, an incorrect but reasonable formula might use the standard errors rather than the squared standard errors of the estimated proportions.

Partially correct (P) if the response includes only two or three of the four components.

Incorrect if the response includes at most one of the four components.

Section 3 is scored as follows:

Essentially correct (E) if the response includes the following four components:

1. In part (e), correctly calculates W using the values from part (d).
2. In part (f), recognizes that Chebyshev's inequality should be used by substituting W for k.
3. In part (f), applies reasonable logic to make a conclusion based on using W and Chebyshev's inequality.
4. In part (f), makes a conclusion including linkage and context, consistent with the logic given in component 3.

Partially correct if the response includes only two or three of the four components.

Incorrect if the response includes at most one of the four components.

Question 6 (continued)

4 Complete Response

Three sections essentially correct

3 Substantial Response

Two sections essentially correct and one section partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and one or two sections partially correct

OR

Three sections partially correct

1 Minimal Response

One section essentially correct

OR

No sections essentially correct and two sections partially correct