Question 5: Multi-Focus — 4 points

General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

	Model Solution	Scoring
(a)	The sample median reduction in blood pressure for those eating dark chocolate, 7 mmHg, is greater than the sample median reduction in blood pressure for those eating white chocolate, about 0 mmHg.	**Essentially correct (E)** if the response satisfies the following three components: 1. The medians are correctly determined with the median for dark chocolate equal to 7 mmHg and the median for white chocolate between −2 mmHg and 2 mmHg, inclusive 2. The computed medians are compared correctly 3. The context of the response variable (reduction in blood pressure) and the treatment groups (dark chocolate and white chocolate) is included **Partially correct (P)** if the response satisfies two of the three components and, if stated, the reported values of the medians are reasonable *OR* if the response satisfies one of components 2 or 3 and the reported values of the medians are reasonable. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- To be scored as reasonable, the reported value of the median for the dark chocolate group must be between 5 mmHg and 8 mmHg, inclusive, and the reported value for the median of the white chocolate group must be between −2 mmHg and 3 mmHg, inclusive.

	Model Solution	Scoring
(b)	The researcher's conclusion may not necessarily be true because looking at the difference in sample means alone does not consider the variability in the sampling distribution of the differences in sample means. Another random assignment of subjects to dark and white chocolate may result in a sample mean reduction in blood pressure for subjects assigned to dark chocolate that is smaller than the mean reduction for subjects assigned to white chocolate. The variability in the sampling distribution of potential differences in sample means must be considered in making a conclusion about convincing statistical evidence. Further, an inference procedure can assess the likelihood that the observed difference in sample means occurred by random chance if the population means are equal.	**Essentially correct (E)** if the response satisfies at least two of the following three components: 1. Indicates that the researchers failed to consider the variability in the sampling distribution of differences in sample means (or that this should be considered) OR states that a different random assignment could produce different sample means 2. Includes an explanation that a difference of 5.66 mmHg could have occurred by random chance, even if the population means are equal 3. Indicates that an inference procedure is needed **Partially correct (P)** if the response satisfies only one of the three components. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- A response may be scored E if it states that no conclusion can be made about the relationship between the population mean blood pressure reduction for those who consume dark chocolate and those who consume white chocolate because the sample was drawn only from the subpopulation of *healthy* adults.

Model Solution	Scoring
(c) The observed value of the sample statistic $\bar{x}_{dark} - \bar{x}_{white}$ is 5.66 mmHg. The graph of simulation results reveals that a difference of 5.66 mmHg or larger occurred in only 3 of the 120 trials. Thus, the p-value is approximately equal to $\frac{3}{120} = 0.025$. Thus, assuming that there is no difference in mean blood pressure reduction, there is an approximate probability of 0.025 of getting a difference of 5.66 mmHg or larger by chance alone. Because this approximate p-value is less than 5 percent, there is convincing evidence that adding dark chocolate to the diet will result in a greater mean reduction in blood pressure than adding white chocolate to the diet for people similar to those who participated in the study.	**Essentially correct (E)** if the response satisfies the following four components: 1. Calculates the correct p-value of 0.025 2. Provides supporting work from the simulation for the calculation of the p-value 3. Provides a correct conclusion in the context of whether there is convincing evidence that adding dark chocolate to the diet will result in greater average reduction in blood pressure than adding white chocolate to the diet 4. Justifies the conclusion by directly comparing the calculated p-value to 0.05 **Partially correct (P)** if the response satisfies only two or three of the four components *OR* if the conclusion (component 3) is correct but is not in context and at least one additional component is satisfied *OR* if the conclusion (component 3) is in context and correct, but the justification is based on an analysis of the data shown in the dotplots (e.g., a two-sample t-confidence interval, a two-sample t-test, an appropriate nonparametric test, or an "exact" randomization test). **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- Component 2 is satisfied by simply noting that only three of the simulated trials resulted in a difference of at least 5.66 mmHg, the observed difference.
- A response that incorrectly evaluates the p-value may still satisfy components 3 and 4 if the response is consistent with the use of the incorrect p-value.
- The hypotheses are not required because they are given in the stem.

Scoring for Question 5	Score
Complete Response Three parts essentially correct	**4**
Substantial Response Two parts essentially correct and one part partially correct	**3**
Developing Response Two parts essentially correct and no part partially correct *OR* One part essentially correct and one or two parts partially correct *OR* Three parts partially correct	**2**
Minimal Response One part essentially correct and no part partially correct *OR* No part essentially correct and two parts partially correct	**1**