1. **(9 points.)** Below are 100 sorted integer responses. Note that I am not telling you the values of \( x_{(14)} \) and \( x_{(94)} \); I do this to discourage you from spending your time summing the 100 numbers. In order to answer the questions below, all you need to know about these numbers are the facts obvious from this listing; namely, \( 449 \leq x_{(14)} \leq 481 \) and \( 677 \leq x_{(94)} \leq 686 \). Call this data set A.

\[
\begin{array}{rrrrrrrrrrrrrr}
400 & 401 & 403 & 404 & 404 & 405 & 431 & 437 & 443 & 446 & 446 & 447 & 449 & x_{(14)} & 481 \\
493 & 494 & 513 & 515 & 517 & 519 & 521 & 522 & 524 & 525 & 525 & 526 & 528 & 531 & 536 \\
539 & 542 & 546 & 550 & 552 & 559 & 560 & 564 & 564 & 571 & 571 & 573 & 582 & 587 \\
588 & 588 & 588 & 597 & 601 & 603 & 607 & 608 & 608 & 609 & 609 & 609 & 610 & 612 \\
612 & 612 & 619 & 626 & 630 & 631 & 635 & 637 & 639 & 641 & 642 & 642 & 646 & 648 & 649 \\
676 & 677 & 677 & x_{(94)} & 686 & 690 & 694 & 695 & 696 & 697 \\
\end{array}
\]

The mean of data set A equals 580.00 and the standard deviation equals 82.84.

(a) According to the Empirical Rule, approximately 68\% of the data fall between \( c \) and \( d \). Determine the numerical values of \( c \) and \( d \).

Answer:

(b) Refer to part (a) and the values of \( c \) and \( d \) that you determined. What percentage of the observations actually fall between the values of \( c \) and \( d \)?

Answer:

(c) Refer to the observation equal to 539. Determine its deviation (from the mean).

Answer:

(d) Calculate the median of data set A.

Answer:

(e) Delete observations 400, 401, 403 and 697 from data set A; call the resultant set of 96 numbers data set B. Calculate the mean of data set B. **Hint:** Do not try to obtain your answer by summing the 96 numbers!

Answer:

(f) Refer to the part(e). Calculate the median of data set B.

Answer:
2. (5 points.) The sampling distribution of the random variable $X$ is given in the table below.

<table>
<thead>
<tr>
<th>$x$ :</th>
<th>2</th>
<th>6</th>
<th>8</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$ :</td>
<td>0.20</td>
<td>0.30</td>
<td>0.40</td>
<td>0.10</td>
</tr>
</tbody>
</table>

(a) What is the value of $\delta$?

Answer:

(b) Draw the probability histogram of this distribution on the numberline below. Remember to label the heights of each rectangle.

3. (3 points.) I have one rectangle from a histogram for data, but I don’t remember which type of histogram I have. The boundaries of the rectangle are 2 and 6; its height is 0.12. Given that $n = 300$, how many observations are in the class interval 2 to 6 if my histogram is:

(a) A frequency histogram.

Answer:

(b) A relative frequency histogram.

Answer:

(c) A density histogram.

Answer:

4. (3 points.) I have the following sorted data from a balanced CRD:

1 2 3 6 7 8 12
2 3 4 5 6 7 C

I obtained a simulation study with $m = 1,000$ runs for the test statistic $U$ using the vassarstats website.

For each of the cases below, I give you the value of $C$ and the one-tailed P-value given by vassarstats. Your job is to determine the alternative ($>$, $<$ or $\neq$) that corresponds to the P-value.

(a) $C = 16$ and one-tailed = 0.431.

Answer:

(b) $C = 8$ and one-tailed = 0.408.

Answer:

(c) $C = 12$ and one-tailed = 0.059.

Answer:
5. (6 points.) An unbalanced CRD yields the following data, with numerical codes in parentheses.

<table>
<thead>
<tr>
<th>Tr.</th>
<th>Low (1)</th>
<th>Med. (2)</th>
<th>High (3)</th>
<th>Tot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>21</td>
<td>16</td>
<td>46</td>
</tr>
</tbody>
</table>

(a) Calculate the values of $r_1$ and $r_2$. In order to receive credit, you must show your work.

**Answer:**

(b) Given that the data are analyzed by ranks, which treatment gives larger responses? In order to receive credit, you must show explain your answer.

**Answer:**

6. (6 points.) Refer to the data in the previous problem. I am interested in the sampling distribution of the test statistic $R_1$ for the sum of ranks test. In order to receive credit for your answers below, you must show your work.

(a) Calculate the mean of the sampling distribution.

**Answer:**

(b) Calculate the variance of the sampling distribution.

**Answer:**

(c) Calculate the standard deviation of the sampling distribution.

**Answer:**
7. **(10 points.)** I performed a power analysis to compare the test statistics $U$ and $R_1$. In order to make your computations a bit easier, my simulation experiment consisted of only 1,000 runs instead of the usual 10,000 runs. I won’t tell you explicitly which alternative I used to create the table below. My results are summarized in the table below.

<table>
<thead>
<tr>
<th></th>
<th>$R_1$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fail to Reject $H_0$</td>
<td>Reject $H_0$</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
<td>572</td>
<td>31</td>
<td>603</td>
<td></td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>51</td>
<td>346</td>
<td>397</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>623</td>
<td>377</td>
<td>1,000</td>
<td></td>
</tr>
</tbody>
</table>

(a) Based on this simulation study for the alternative of interest to me, what is the approximate power using the test statistic $R_1$?
   **Answer:**

(b) Let $r_U$ denote the exact power for the alternative of interest to me and the test statistic $U$. Let $r_{R_1}$ denote the exact power for the alternative of interest to me and the test statistic $R_1$. Calculate the nearly certain interval for $r_U$.
   **Answer:**

(c) Refer to the notation in part (b). Calculate the nearly certain interval for $r_U - r_{R_1}$.
   **Answer:**

(d) Briefly interpret your answer in (c); in particular, what can you conclude about the two powers?
   **Answer:**
8. **(3 points.)** I performed a balanced CRD and obtained the number $u$ for the observed value of the test statistic $U$. You are given the following two probabilities.

- $P(U > u) = 0.1420$.
- $P(U = u) = 0.0689$.

Obtain the P-value for each of the alternatives given below.

(a) $>$.  
   **Answer:**

(b) $<$.  
   **Answer:**

(c) $\neq$.  
   **Answer:**

9. **(5 points.)** As discussed in Chapter 3 of the *Course Notes*, for a balanced CRD every assignment has a mirror image. (If you don’t remember this, see the notes that were distributed with this exam.)

The following fact was given in the *Course Notes*: If the assignment $a_1$ gives the number $u$ as the observed value of the test statistic $U$ and the Skeptic is correct, then its mirror image assignment, denoted by $-a_1$, gives the number $-u$ as the observed value of the test statistic $U$.

For each of the questions below, select your answer from the following ten options:

-20, -15, -10, -5, 0, 5, 10, 15, 20

and ‘unable to determine’.

Note that I do **not** want you to explain your answer. Also, the balanced CRD has $n = 8$ and the actual assignment 1,2,3,4 yields $u = -5$.

(a) On the assumption that the Skeptic is correct, what would be the value of $u$ for the assignment 1,3,5,7?  
   **Answer:**

(b) On the assumption that the Skeptic is correct, what would be the value of $u$ for the assignment 5,6,7,8?  
   **Answer:**

(c) On the assumption that there is a constant treatment effect of size 5, what would be the value of $u$ for the assignment 5,6,7,8?  
   **Answer:**