There are three versions of each problem. Some of you had an exam that was all one version, but most of you had exams for which two or three versions were represented. This process is fair if you agree with me that, for every question, the versions are of almost identical difficulty.

For each problem below, first read the information in the preamble and then find the solution that fits your personal exam.

1. **Problem 1: 8 points.** Each version presents the actual data from a balanced CRD on \( n = 6 \) subjects. The response of Subject 1 is 6 [18; 39] on Version 1 [2; 3].

Part (a) [(b); (c); (d)] is worth 1 [2.5; 3; 1.5] point(s).

**Version 1:**

(a) Reading from the table,

\[
\bar{x} = (12 + 27 + 24)/3 = 21.
\]

(b) Because the Skeptic is correct, the six given response values will not change. Thus,

\[
\bar{x} = (6 + 12 + 27)/3 = 15,
\]

\[
\bar{y} = (15 + 18 + 24)/3 = 19,
\]

and \( u = 15 - 19 = -4 \).

(c) I recommend that you take the table of actual data and use it to create the clone-enhanced study table for the CTE of \( c = +9 \):
Version 2:

(a) Reading from the table,

\[ \bar{x} = (18 + 33 + 30)/3 = 27. \]

(b) Because the Skeptic is correct, the six given response values will not change. Thus,

\[ \bar{x} = (18 + 18 + 21)/3 = 19, \]

\[ \bar{y} = (33 + 30 + 9)/3 = 24, \]

and \( u = 19 - 24 = -5. \)

(c) I recommend that you take the table of actual data and use it to create the clone-enhanced study table for the CTE of \( c = +6: \)

\[
\begin{array}{c|cccccc}
\text{Subject:} & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\text{Treat. 1:} & 18 & 33 & 24 & 27 & 30 & 15 \\
\text{Treat. 2:} & 12 & 27 & 18 & 21 & 24 & 9 \\
\end{array}
\]

Reading from this table we obtain:

\[ \bar{x} = (33 + 27 + 15)/3 = 25, \]

\[ \bar{y} = (12 + 18 + 24)/3 = 18, \]

and \( u = 25 - 18 = 7. \)

(d) For the males the Skeptic is correct; for the females there is a constant treatment effect of \(-6.\)

Note on grading: I deducted 1/2 point for not mentioning the sex of the subjects. Also, many people confused the sign of the CTE.

Version 3:

(a) Reading from the table,

\[ \bar{x} = (39 + 36 + 24)/3 = 33. \]

(b) Because the Skeptic is correct, the six given response values will not change. Thus,

\[ \bar{x} = (39 + 24 + 21)/3 = 28, \]

\[ \bar{y} = (36 + 12 + 24)/3 = 24, \]

and \( u = 28 - 24 = 4. \)

(c) I recommend that you take the table of actual data and use it to create the clone-enhanced study table for the CTE of \( c = +12: \)

\[
\begin{array}{c|cccccc}
\text{Subject:} & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\text{Treat. 1:} & 39 & 33 & 36 & 36 & 24 & 24 \\
\text{Treat. 2:} & 27 & 21 & 24 & 24 & 12 & 12 \\
\end{array}
\]

Reading from this table we obtain:

\[ \bar{x} = (39 + 33 + 24)/3 = 32, \]

\[ \bar{y} = (24 + 24 + 12)/3 = 20, \]

and \( u = 32 - 20 = 12. \)

(d) For the males there is a constant treatment effect of \(+3;\) for the females there is a constant treatment effect of \(-4.\)

Note on grading: I deducted 1/2 point for not mentioning the sex of the subjects. Also, many people confused the sign of the CTE.
2. **Problem 2: 7 points.** Each version presents the actual data from an unbalanced CRD. The response of Subject 2 is 15 [27; 9] on Version 1 [2; 3].

Part (a) [(b); (c) and (d) combined] is worth 3 [1; 3] point(s).

**Version 1:**

(a) I begin by creating the following table of ranks:

<table>
<thead>
<tr>
<th>Po:</th>
<th>1 2 3 4 5 6 7 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re:</td>
<td>15 15 21 24 27 27 27 36</td>
</tr>
<tr>
<td>Ra:</td>
<td>1.5 1.5 3 4 6 6 6 8</td>
</tr>
<tr>
<td>Tr:</td>
<td>1 2 2 2 1 2 2 1</td>
</tr>
</tbody>
</table>

Reading from this table,

\[ r_1 = 1.5 + 6 + 8 = 15.5 \text{ and } \]
\[ r_2 = 1.5 + 3 + 4 + 6 + 6 = 20.5. \]

(b) Treatment 1 gives larger responses because

\[ \bar{r}_1 = 15.5/3 = 5.17 \]

is larger than

\[ \bar{r}_2 = 20.5/5 = 4.1. \]

(c) Because the Skeptic is correct, the eight ranks will be the same for all 56 possible assignments. The P-value will be

\[ P(R_1 \geq r_1). \]

This probability is minimized by replacing \( r_1 \) by the largest possible value of \( R_1 \), which is 20 = 8 + 6 + 6.

By inspection, there are three assignments that yield \( r_1 = 20 \): combine the 8 with any two of the three 6’s.

Thus, the smallest possible P-value is 3/56 = 0.0536.

(d) The assignments are: 1,5,7; 1,5,8; and 5,7,8.

**Version 2:**

(a) I begin by creating the following table of ranks:

<table>
<thead>
<tr>
<th>Po:</th>
<th>1 2 3 4 5 6 7 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re:</td>
<td>21 27 27 27 30 30 33 42</td>
</tr>
<tr>
<td>Ra:</td>
<td>1 3 3 3 5.5 5.5 7 8</td>
</tr>
<tr>
<td>Tr:</td>
<td>2 1 1 2 2 2 2 1</td>
</tr>
</tbody>
</table>

Reading from this table,

\[ r_1 = 3 + 3 + 8 = 14 \text{ and } \]
\[ r_2 = 1 + 3 + 5.5 + 5.5 + 7 = 22. \]

(b) Treatment 1 gives larger responses because

\[ \bar{r}_1 = 14/3 = 4.67 \]

is larger than

\[ \bar{r}_2 = 22/5 = 4.40. \]

(c) Because the Skeptic is correct, the eight ranks will be the same for all 56 possible assignments. The P-value will be

\[ P(R_1 \geq r_1). \]

This probability is minimized by replacing \( r_1 \) by the largest possible value of \( R_1 \), which is 20.5 = 8 + 7 + 5.5.

By inspection, there are two assignments that yield \( r_1 = 20.5 \): combine the 8 and the 7 with either of the two 5.5’s.

Thus, the smallest possible P-value is 2/56 = 0.0357.

(d) The assignments are: 4,5,6; and 4,5,8.
Version 3:

(a) I begin by creating the following table of ranks:

<table>
<thead>
<tr>
<th>Po:</th>
<th>1 2 3 4 5 6 7 8 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re:</td>
<td>0 0 2 2 2 2 6 7 9</td>
</tr>
<tr>
<td>Ra:</td>
<td>1.5 1.5 4.5 4.5 4.5 4.5 7 8 9</td>
</tr>
<tr>
<td>Tr:</td>
<td>1 2 1 2 2 2 1 1 2</td>
</tr>
</tbody>
</table>

Reading from this table,

\[ r_1 = 1.5 + 4.5 + 7 + 8 = 21 \] and

\[ r_2 = 1.5 + 4.5 + 4.5 + 4.5 + 9 = 24. \]

(b) Treatment 1 gives larger responses because

\[ \bar{r}_1 = \frac{21}{4} = 5.25 \]

is larger than

\[ \bar{r}_2 = \frac{24}{5} = 4.80. \]

(c) Because the Skeptic is correct, the nine ranks will be the same for all 126 possible assignments. The P-value will be

\[ P(R_1 \geq r_1). \]

This probability is minimized by replacing \( r_1 \) by the largest possible value of \( R_1 \), which is \( 28.5 = 9 + 8 + 7 + 4.5 \).

By inspection, there are four assignments that yield \( r_1 = 28.5 \): combine the 9, 8 and 7 with any of the four 4.5’s. Thus, the smallest possible P-value is \( 4/126 = 0.0317 \).

(d) The assignments are: 2,3,4,6; 2,3,5,6; 2,3,6,7; and 2,3,6,9.
3. **Problem 3: 8 points.** Each version presents an exact sampling distribution for \( U \). The largest entry in the ‘\( u \)’ column is 4.0 [29.5; 20.0] in Version 1 [2; 3].

Part (a) [(b); (c)] is worth 2.5 [2.5; 3] points.

**Version 1:**

(a) The P-value is
\[
P(U \geq 2.2) = 0.0539 + 0.0090 + 0.0006 = 0.0635.
\]

(b) The P-value is
\[
P(U \leq -2.3) = 0.0446 + 0.0050 = 0.0496.
\]

(c) The P-value is
\[
P(U \geq 3.1) + P(U \leq -3.1).
\]

The first of these is:
\[
0.0090 + 0.0006 = 0.0096,
\]
and the second is 0.0050. Thus, the P-value is
\[
0.0096 + 0.0050 = 0.0146.
\]

**Version 2:**

(a) The P-value is
\[
P(U \geq 23.0) = 0.0117 + 0.0014 = 0.0131.
\]

(b) The P-value is
\[
P(U \leq -9.5) = 0.1750 + 0.0601 + 0.0086 = 0.2437.
\]

(c) The P-value is
\[
P(U \geq 16.5) + P(U \leq -16.5).
\]

The first of these is:
\[
0.0548 + 0.0117 + 0.0014 = 0.0679
\]
and the second is 0.0086. Thus, the P-value is
\[
0.0679 + 0.0086 = 0.0765.
\]

**Version 3:**

(a) The P-value is
\[
P(U \geq 11.0) = 0.0607 + 0.0120 + 0.0010 = 0.0737.
\]

(b) The P-value is
\[
P(U \leq -11.5) = 0.0516 + 0.0069 = 0.0585.
\]

(c) The P-value is
\[
P(U \geq 15.5) + P(U \leq -15.5).
\]

The first of these is:
\[
0.0120 + 0.0010 = 0.0130,
\]
and the second is 0.0069. Thus, the P-value is
\[
0.0130 + 0.0069 = 0.0199.
\]
4. **Problem 4: 3 points.** Each version asks you to draw a probability histogram. The first entry in the ‘Prob’ row is 0.21 [0.42; 0.30] in Version 1 [2; 3].

**Version 1:**
First, the distances between successive values are: 6, 3 and 9; thus, $\delta = 3$. The base of each rectangle is 3 and the height is probability divided by 3. Note that the second and third rectangles below must touch; the others must not.

**Version 2:**
First, the distances between successive values are: 6, 9 and 3; thus, $\delta = 3$. The base of each rectangle is 3 and the height is probability divided by 3. Note that the third and fourth rectangles below must touch; the others must not.

**Version 3:**
First, the distances between successive values are: 3, 6 and 9; thus, $\delta = 3$. The base of each rectangle is 3 and the height is probability divided by 3. Note that the first and second rectangles below must touch; the others must not.
5. **Problem 5: 6 points.** Each version asks questions about the three types of histograms for data. In part (a) $n$ equals 250 [400; 700] in Version 1 [2; 3]. Each part is worth 2 points. **A note on the grading:** As you will see below, for each version, three of the nine answers are *impossible*. (This problem is very similar to Practice Problem 5 in Chapter 2.) While grading, it seemed clear to me that many students never considered *impossible* to be an allowable answer (although many others did). As a result, I decided that the penalty for missing all three *impossibles* would be $-1$ instead the more natural $-2$.

**Version 1:**

(a) i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 50.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval; hence, the answer is $50/250 = 0.20$.

iii. For a density histogram, the height of a rectangle is the relative frequency of its class interval divided by the width of its class interval; hence, the answer is $0.20/10 = 0.02$.

(b) i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 5.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. A relative frequency may not exceed 1; hence, the described situation is impossible.

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case is $5(0.05) = 0.25$. Thus, the frequency of the class interval is $0.25(400) = 100$.

(c) i. For a frequency histogram, the height of a rectangle is the frequency of its class interval. But a frequency must be a nonnegative integer; hence, the described situation is impossible.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. Thus, the frequency of the class interval is $0.20(600) = 120$.

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case is $6(0.20) = 1.20$. A relative frequency, however, may not exceed one. Hence, the described situation is impossible.

**Version 2:**

(a) i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 100.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval; hence, the answer is $100/400 = 0.25$.

iii. For a density histogram, the height of a rectangle is the relative frequency of its class interval divided by the width of its class interval; hence, the answer is $0.25/5 = 0.05$.

(b) i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 3.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. A relative frequency may not exceed 1; hence, the described situation is impossible.

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case
is \(3(0.1) = 0.3\). Thus, the frequency of the class interval is \(0.3(200) = 60\).

(c)  

i. For a frequency histogram, the height of a rectangle is the frequency of its class interval. But a frequency must be a nonnegative integer; hence, the described situation is impossible.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. Thus, the frequency of the class interval is \(0.40(500) = 200\).

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case is \(3(0.40) = 1.20\). A relative frequency, however, may not exceed one. Hence, the described situation is impossible.

Version 3:

(a)  
i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 210.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval; hence, the answer is \(210/700 = 0.30\).

iii. For a density histogram, the height of a rectangle is the relative frequency of its class interval divided by the width of its class interval; hence, the answer is \(0.30/2 = 0.15\).

(b)  
i. For a frequency histogram, the height of a rectangle is the frequency of its class interval. But a frequency must be a nonnegative integer; hence, the described situation is impossible.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. Thus, the frequency of the class interval is \(0.20(800) = 160\).

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case is \(6(0.20) = 1.20\). A relative frequency, however, may not exceed one. Hence, the described situation is impossible.

(c)  
i. For a frequency histogram, the height of a rectangle is the frequency of its class interval; hence, the answer is 20.

ii. For a relative frequency histogram, the height of a rectangle is the relative frequency of its class interval. A relative frequency may not exceed 1; hence, the described situation is impossible.

iii. For a density histogram, the area of a rectangle is the relative frequency of its class interval. The area in this case is \(20(0.02) = 0.40\). Thus, the frequency of the class interval is \(0.40(300) = 120\).
6. **Problem 6: 8 points.** Each version presents you with the results of an unbalanced CRD with an ordered categorical response. The value of \( n \) is 40 [34; 30] in Version 1 [2; 3].

Parts (a) and (b) are worth 3 points each; (c) is worth 1.5 points and (d) is worth 1/2 point. A note on grading: In (b) the mean was worth one point and standard deviation two points. In (c) above or below was worth 1/2 point; using \( r_1 \) as the starting point for the entry was worth 1/2 point (other popular choices were the mean or the standard deviation); and correctly using the continuity correction was worth 1/2 point.

In (d), you received 1/2 point if your answer:

- While wrong, was consistent with earlier wrong answers; or
- Was correct, even if it was inconsistent with earlier wrong answers.

**Version 1:**

I begin by adding to the data table:

<table>
<thead>
<tr>
<th>Tr.</th>
<th>(Code)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>12</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Tot.</td>
<td></td>
<td>17</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Pos.</td>
<td>1–17</td>
<td>18–29</td>
<td>30–40</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>9</td>
<td>23.5</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

(a) From above,

\[
r_1 = 5(9) + 4(23.5) + 6(35) = 349 \]

and

\[
r_2 = 12(9) + 8(23.5) + 5(35) = 471.\]

(b) First,

\[
\mu = 15(41)/2 = 307.5.
\]

Next, note that the \( t_i \)'s are 17, 12 and 11. Thus,

\[
\sum (t_i^3 - t_i) = (17^3 - 17) + (12^3 - 12) + (11^3 - 11) = 4896 + 1716 + 1320 = 7932.
\]

The first part of the variance is

\[
15(25)(41)/12 = 1281.25,
\]

and the second part is

\[
\frac{15(25)(7932)}{12(40)(39)} = 158.89.
\]

Therefore,

\[
\sigma^2 = 1281.25 - 158.89 = 1122.36
\]

and

\[
\sigma = \sqrt{1122.36} = 33.50.
\]

(c) The exact P-value is

\[
P(R_1 \geq 349).
\]

We want the area to the right, so we choose Above and, using the continuity correction, enter 348.5 in the box.

(d) (iv). Because the number in the Above box is larger than the mean, 307.5, the area will be less than 0.5000. (The area under the normal curve above its mean is 0.5000.)

**Version 2:**

I begin by adding to the data table:

<table>
<thead>
<tr>
<th>Tr.</th>
<th>(Code)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>7</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Tot.</td>
<td></td>
<td>11</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Pos.</td>
<td>1–11</td>
<td>12–25</td>
<td>26–34</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>6</td>
<td>18.5</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

(a) From above,

\[
r_1 = 7(6) + 8(18.5) + 5(30) = 340 \]

and

\[
r_2 = 4(6) + 6(18.5) + 4(35) = 255.
\]
(b) First, 

$$\mu = 20(35)/2 = 350.$$ 

Next, note that the \( t_i \)'s are 11, 14 and 9. Thus, 

$$\sum(t_i^3 - t_i) = (11^3 - 11) + (14^3 - 14) + (9^3 - 9) = 1320 + 2730 + 720 = 4770.$$ 

The first part of the variance is 

$$20(14)(35)/12 = 816.67,$$

and the second part is 

$$\frac{20(14)(4770)}{12(34)(33)} = 99.20.$$ 

Therefore, 

$$\sigma^2 = 816.67 - 99.20 = 717.47$$

and 

$$\sigma = \sqrt{717.47} = 26.79.$$ 

(c) The exact P-value is 

$$P(R_1 \geq 340).$$

We want the area to the right, so we choose Above and, using the continuity correction, enter 339.5 in the box.

(d) (ii). Because the number in the Above box is smaller than the mean, 350, the area will be greater than 0.5000. (The area under the normal curve above its mean is 0.5000.)

Version 3:

I begin by adding to the data table:

<table>
<thead>
<tr>
<th>Tr.</th>
<th>(Code)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Tot.</td>
<td></td>
<td>7</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Pos.</td>
<td></td>
<td>1–7</td>
<td>8–19</td>
<td>20–30</td>
</tr>
<tr>
<td>Rank</td>
<td></td>
<td>4</td>
<td>13.5</td>
<td>25</td>
</tr>
</tbody>
</table>

(a) From above, 

$$r_1 = 4(4) + 6(13.5) + 2(25) = 147$$ and 

$$r_2 = 3(4) + 6(13.5) + 9(25) = 318.$$ 

(b) First, 

$$\mu = 12(31)/2 = 186.$$ 

Next, note that the \( t_i \)'s are 7, 12 and 11. Thus, 

$$\sum(t_i^3 - t_i) = (7^3 - 7) + (12^3 - 12) + (11^3 - 11) = 336 + 1716 + 1320 = 3372.$$ 

The first part of the variance is 

$$12(18)(31)/12 = 558,$$

and the second part is 

$$\frac{12(18)(3372)}{12(30)(29)} = 69.77.$$ 

Therefore, 

$$\sigma^2 = 558 - 69.77 = 488.23$$

and 

$$\sigma = \sqrt{488.23} = 22.10.$$ 

(c) The exact P-value is 

$$P(R_1 \leq 147).$$

We want the area to the left, so we choose Below and, using the continuity correction, enter 147.5 in the box.

(d) (iv). Because the number in the Below box is smaller than the mean, 186, the area will be less than 0.5000. (The area under the normal curve below its mean is 0.5000.)
7. **Problem 7: 4 points.** Each version asks you to identify all assignments that satisfy certain conditions. Patrick is subject 1 [5; 3] on Version 1 [2; 3].

Each part is worth 2 points.

**Version 1:**

(a) To make Leonard happy, we can place Teresa (2) on one treatment and Grace (6) on the other. There are eight such assignments:

- 1,2; 2,3; 2,4; and 2,5 put Teresa on treatment 1 and Grace on treatment 2.
- 1,6; 3,6; 4,6; and 5,6 put Grace on treatment 1 and Theresa on treatment 2.

(b) To make Sheldon happy, we can place either Craig or Grace on treatment 1 and either Patrick or Theresa on treatment 2. The assignments that do this are:

- 1,5; 1,6; 2,6; and 2,5.

**Version 2:**

(a) To make Leonard happy, we can place Teresa (6) on one treatment and Grace (4) on the other. There are eight such assignments:

- 1,6; 2,6; 3,6; and 5,6 put Teresa on treatment 1 and Grace on treatment 2.
- 1,4; 2,4; 3,4; and 4,5 put Grace on treatment 1 and Theresa on treatment 2.

(b) To make Sheldon happy, we can place either Craig or Grace on treatment 1 and either Patrick or Theresa on treatment 2. The assignments that do this are:

- 1,3; 1,4; 2,3; and 2,4.

**Version 3:**

(a) To make Leonard happy, we can place Teresa (4) on one treatment and Grace (2) on the other. There are eight such assignments:

- 1,4; 3,4; 4,5; and 4,6 put Teresa on treatment 1 and Grace on treatment 2.
- 1,2; 2,3; 2,5; and 2,6 put Grace on treatment 1 and Theresa on treatment 2.

(b) To make Sheldon happy, we can place either Craig or Grace on treatment 1 and either Patrick or Theresa on treatment 2. The assignments that do this are:

- 3,5; 3,6; 4,6; and 4,5.
8. **Problem 8: 6 points.** Each version asks you to use the results of a simulation experiment to approximate two P-values. The number of reps is 6000 [4000; 8000] on Version 1 [2; 3].

Parts (a) and (b) are worth 1.5 points each. Parts (c) and (d) combined are worth 3 points.

**Version 1:**

(a) The exact P-value is $P(U \geq 2.80)$. We approximate this by the relative frequency of $(U \geq 2.80)$, which is

$$(76 + 194)/6000 = 0.0450.$$  

(b) The nearly certain interval is

$$0.0450 \pm 3\sqrt{0.045(0.955)/6000} = 0.0450 \pm 0.0080.$$  

(c) Let $r$ denote the exact P-value in (a); i.e., $P(U \geq 2.80)$. By symmetry, the exact P-value in (c) is $2r$. Our single number approximation of $2r$ is $2(0.0450) = 0.0900$. Therefore, the NCI of $2r$ is twice the NCI of $r$; i.e.,

$$2(0.0450 \pm 0.0080) = 0.0900 \pm 0.0160.$$  

Literally, **nobody** got this correct; the popular wrong answer was

$$0.0900 \pm 3\sqrt{0.09(0.91)/6000} = 0.0900 \pm 0.0110.$$  

This answer is wrong because you **may not pretend** that you know that the frequency of $(U \leq -2.80)$ is the same as the frequency of $(U \geq 2.80)$.

Because everyone got this wrong, I conclude that this was a poor question to put on this exam. We will return to this topic in Chapter 13.

**Version 2:**

(a) The exact P-value is $P(U \geq 3.20)$. We approximate this by the relative frequency of $(U \geq 3.20)$, which is

$$(60 + 180)/4000 = 0.0600.$$  

(b) The nearly certain interval is

$$0.0600 \pm 3\sqrt{0.06(0.94)/4000} = 0.0600 \pm 0.0113.$$  

(c) Let $r$ denote the exact P-value in (a); i.e., $P(U \geq 3.20)$. By symmetry, the exact P-value in (c) is $2r$. Our single number approximation of $2r$ is $2(0.0600) = 0.1200$. Therefore, the NCI of $2r$ is twice the NCI of $r$; i.e.,

$$2(0.0600 \pm 0.0113) = 0.1200 \pm 0.0226.$$  

Literally, **nobody** got this correct; the popular wrong answer was

$$0.1200 \pm 3\sqrt{0.12(0.88)/4000} = 0.1200 \pm 0.0172.$$  

This answer is wrong because you **may not pretend** that you know that the frequency of $(U \leq -3.20)$ is the same as the frequency of $(U \geq 3.20)$.

Because everyone got this wrong, I conclude that this was a poor question to put on this exam. We will return to this topic in Chapter 13.
Version 3:

(a) The exact P-value is \( P(U \geq 3.80) \). We approximate this by the relative frequency of \( (U \geq 3.80) \), which is

\[
\frac{(63 + 185)}{8000} = 0.0310.
\]

(b) The nearly certain interval is

\[
0.0310 \pm 3\sqrt{0.031(0.969)/8000} =
0.0310 \pm 0.0058.
\]

(c) Let \( r \) denote the exact P-value in (a); i.e., \( P(U \geq 3.80) \). By symmetry, the exact P-value in (c) is \( 2r \). Our single number approximation of \( 2r \) is \( 2(0.0310) = 0.0620 \). Therefore, the NCI of \( 2r \) is twice the NCI of \( r \); i.e.,

\[
2(0.0310 \pm 0.0058) = 0.0620 \pm 0.0116.
\]

Literally, **nobody** got this correct; the popular wrong answer was

\[
0.0620 \pm 3\sqrt{0.062(0.938)/8000} =
0.0620 \pm 0.0081.
\]

This answer is wrong because you **may not pretend** that you know that the frequency of \( (U \leq -3.80) \) is the same as the frequency of \( (U \geq 3.80) \).

Because everyone got this wrong, I conclude that this was a poor question to put on this exam. We will return to this topic in Chapter 13.