1) From 9 names on a ballot, a committee of 3 will be elected to attend a political national convention. How many different committees are possible?
   A) 252     B) 729     C) 504     D) 60,480     E) 84

2) Forty-seven math majors, 22 music majors and 31 history majors are randomly selected from 585 math majors, 279 music majors and 393 history majors at the state university. What sampling technique is used?
   A) convenience
   B) stratified
   C) cluster
   D) systematic
   E) simple random

3) A doctor at a local hospital is interested in estimating the birth weight of infants. How large a sample must she select if she desires to be 90% confident that her estimate is within 2 ounces of the true mean? Assume that $\sigma = 4.9$ ounces and that birth weights are normally distributed.
   A) 13     B) 15     C) 16     D) 17     E) 19
4) According to the law of large numbers, as more observations are added to the sample, the difference between the sample mean and the population mean
   A) Remains about the same
   B) Tends to become smaller
   C) Is inversely affected by the data added
   D) Tends to become larger

5) The length of time it takes college students to find a parking spot in the library parking lot follows a normal distribution with a mean of 6.5 minutes and a standard deviation of 1 minute. Find the probability that a randomly selected college student will take between 5.0 and 7.5 minutes to find a parking spot in the library lot.
   A) 0.4938
   B) 0.0919
   C) 0.7745
   D) 0.2255

6) A researcher wishes to construct a confidence interval for a population mean μ. If the sample size is 19, what conditions must be satisfied to compute the confidence interval?
   A) The population standard deviation σ must be known.
   B) It must be true that np(1-p) ≥ 10 and n ≤ 0.05N.
   C) The data must come from a population that is approximately normal with no outliers.
   D) The confidence level cannot be greater than 90%.

7) Investing is a game of chance. Suppose there is a 36% chance that a risky stock investment will end up in a total loss of your investment. Because the rewards are so high, you decide to invest in five independent risky stocks. Find the probability that at least one of your five investments becomes a total loss.
   A) 0.8926
   B) 0.0604
   C) 0.006
   D) 0.302

8) If we do not reject the null hypothesis when the null hypothesis is in error, then we have made a
   A) Type I error
   B) Correct decision
   C) Type II error
   D) Type β error
9) What effect would increasing the sample size have on a confidence interval?
   A) No change  B) Change the confidence level
   C) Increase the width of the interval  D) Decrease the width of the interval

10) A seed company has a test plot in which it is testing the germination of a hybrid seed. They plant 50 rows of 40 seeds per row. After a two-week period, the researchers count how many seeds per row have sprouted. They noted that the least number of seeds to germinate was 33 and some rows had all 40 germinate. The germination data is given below in the table. The random variable $X$ represents the number of seeds in a row that germinated and $P(x)$ represents the probability of selecting a row with that number of seeds germinating. Determine the expected number of seeds per row that germinated.

<table>
<thead>
<tr>
<th>$x$</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$</td>
<td>0.02</td>
<td>0.06</td>
<td>0.10</td>
<td>0.20</td>
<td>0.24</td>
<td>0.26</td>
<td>0.10</td>
<td>0.02</td>
</tr>
</tbody>
</table>

   A) 1.51  B) 4.61  C) 36.50  D) 36.86  E) 37.00

---

**PART II Instructions:** Questions 11 – 20 are open response. Answer all TEN questions carefully and completely, for full credit you must show all appropriate work and clearly indicate your answers.

11) The owner of a computer repair shop has determined that their daily revenue has mean $7200 and standard deviation $1200. The daily revenue totals for the next 30 days will be monitored. What is the probability that the mean daily revenue for the next 30 days will exceed $7500? Round your answer to 4 decimal places.
12) The costs in dollars of a random sample of 20 college textbooks are given in the stem-and-leaf plot below.

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7 8</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>0 2</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>1 6 9 9</td>
</tr>
<tr>
<td>8</td>
<td>2 4 4</td>
</tr>
<tr>
<td>9</td>
<td>0 3 5 7</td>
</tr>
<tr>
<td>10</td>
<td>5 7</td>
</tr>
</tbody>
</table>

Legend: 2|7 represents $27

i) Find the five number summary for this data set. Include the name or correct symbol for each of the numbers as well as its value.

ii) Draw a boxplot of this data set.

iii) Use complete sentences to briefly describe the shape of the distribution for this data.
13) Find the mean, median, and mode of the following statistic students' test scores. Round to the nearest tenth if necessary.

68 73 82 82 86 87 90 91 97

mean = __________________

median = __________________

mode = __________________

14) A physical fitness association is including the mile run in its secondary-school fitness test. The time for this event for boys in secondary school is known to have a normal distribution with a mean of 470 seconds and a standard deviation of 40 seconds. The fitness association wants to recognize the fastest 10% of the boys with certificates of recognition. What time would the boys need to beat in order to earn a certificate of recognition from the fitness association? Note that the fastest runners have the shortest times. Round to the nearest second.
i) Based on the scatterplot, is it reasonable to suggest that there is a linear relationship between hours of sleep and exam scores? Yes or No (circle one)

ii) Find the correlation coefficient for the given data. Round to 4 decimal places.

iii) Determine if there is a significant linear correlation. Report the critical value and state your conclusion.

iv) Find the equation of the least-squares regression line for this data. Round values to 2 decimal places.

v) Use the regression equation to predict the exam score of a student who slept for 7 hours the night before the exam. Is the predicted exam score a good estimate? Briefly explain your answer.
A random sample of 20 college students is selected. Each student is asked how much time he or she spent on the Internet during the previous week. The following times (in hours) are recorded:

<table>
<thead>
<tr>
<th>Time (in hours)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
</tr>
</tbody>
</table>

i) Create a frequency and relative frequency table for this data. Use 3 as the lower class limit of the first class, and use a class width of 4.

<table>
<thead>
<tr>
<th>Class (optional)</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
</table>

ii) Create a relative frequency histogram for the data. Be sure to label your axes.
17) When 440 junior college students were surveyed, 200 said they have a passport. Construct a 95% confidence interval for the proportion of junior college students that have a passport. Round to the nearest thousandth.

18) The National Association of Realtors estimates that 23% of all homes purchased in 2004 were considered investment properties. If a sample of 800 homes sold in 2004 is obtained what is the probability that at most 200 homes are going to be used as investment property? Round your answer to 4 decimal places.
19) In 2010, 36% of adults in a certain country were morbidly obese. A health practitioner suspects that the percent has changed since then. She obtains a random sample of 1042 adults and finds that 393 are morbidly obese. Is this sufficient evidence to support the practitioner's suspicion that the percent of morbidly obese adults has changed at the $\alpha = 0.10$ level of significance?

Are you using the Classical or P-Value approach? (circle one)

Null Hypothesis:

Alternative Hypothesis:

Test Statistic:

Critical Value(s) or P-Value (circle which of these you are using):

Conclusion about the Null Hypothesis:

Do the data support the health practitioner's suspicion? Answer with complete sentences.
20) A shipping firm suspects that the mean life of a certain brand of tire used by its trucks is less than 40,000 miles. To check the hypothesis, the firm randomly selects and tests 18 of these tires and finds that they have a mean lifetime of 39,300 miles with a standard deviation of 1200 miles. At $\alpha = 0.05$, test the shipping firm’s hypothesis. Assume that the life of the tires is normally distributed with no outliers. Show a complete solution including all your steps.
SOLUTIONS

Any calculator is okay. Necessary tables and formulas are attached to the back of the exam. All problems are equally weighted.

Computers, cell phones or other devices that connect to the Internet or communicate with others are not allowed. Students may not bring notes, formulas, or tables into the exam.

This exam has two parts:
PART I: 10 multiple choice questions
PART II: 10 open ended questions

PART I Instructions: Questions 1–10 are multiple choice. Answer all TEN questions and circle the correct answer. It is not necessary to show work. No partial credit will be awarded on this portion of the exam.

1) From 9 names on a ballot, a committee of 3 will be elected to attend a political national convention. How many different committees are possible?
   A) 252  B) 729  C) 504  D) 60,480  E) 84
   \( \binom{9}{3} = \frac{9!}{(9-3)!} = 84 \)  
   \( \text{order is unimportant - combinations} \)

2) Forty-seven math majors, 22 music majors and 31 history majors are randomly selected from 585 math majors, 279 music majors and 393 history majors at the state university. What sampling technique is used?
   A) convenience  B) stratified  C) cluster  D) systematic  E) simple random

3) A doctor at a local hospital is interested in estimating the birth weight of infants. How large a sample must she select if she desires to be 90% confident that her estimate is within 2 ounces of the true mean? Assume that \( \sigma = 4.9 \) ounces and that birth weights are normally distributed.
   A) 13  B) 15  C) 16  D) 17  E) 19
   \[ n = \left( \frac{z_{0.05} \cdot \sigma}{E} \right)^2 = \left( \frac{1.645 \cdot 4.9}{2} \right)^2 = 16.24 \]
   \[ \text{always round up to a whole person} \]
4) According to the law of large numbers, as more observations are added to the sample, the difference between the sample mean and the population mean
   A) Remains about the same
   B) Tends to become smaller
   C) Is inversely affected by the data added
   D) Tends to become larger

5) The length of time it takes college students to find a parking spot in the library parking lot follows a normal distribution with a mean of 6.5 minutes and a standard deviation of 1 minute. Find the probability that a randomly selected college student will take between 5.0 and 7.5 minutes to find a parking spot in the library lot.
   A) 0.4938
   B) 0.0919
   C) 0.7745
   D) 0.2255

\[
\begin{align*}
   & P(5.0 < x < 7.5) \\
   & = P(\frac{5.0 - 6.5}{1} < z < \frac{7.5 - 6.5}{1}) \\
   & = P(-1.5 < z < 1) \\
   & = 0.8413 - 0.0668 \\
   & = 0.7745
\end{align*}
\]

6) A researcher wishes to construct a confidence interval for a population mean \( \mu \). If the sample size is 19, what conditions must be satisfied to compute the confidence interval?
   A) The population standard deviation \( \sigma \) must be known.
   B) It must be true that \( np(1-p) \geq 10 \) and \( n \leq 0.05N \).
   C) The data must come from a population that is approximately normal with no outliers.
   D) The confidence level cannot be greater than 90%.

7) Investing is a game of chance. Suppose there is a 36% chance that a risky stock investment will end up in a total loss of your investment. Because the rewards are so high, you decide to invest in five independent risky stocks. Find the probability that at least one of your five investments becomes a total loss. Round to the nearest ten-thousandth when necessary.
   A) 0.8926
   B) 0.0604
   C) 0.006
   D) 0.302

\[
\begin{align*}
   P(\text{at least one loss}) & = 1 - P(\text{none are a total loss}) \\
   & = 1 - (1 - 0.36)^5 \\
   & = 0.8926
\end{align*}
\]

8) If we do not reject the null hypothesis when the null hypothesis is in error, then we have made a
   A) Type I error
   B) Correct decision
   C) Type II error
   D) Type \( \beta \) error
9) What effect would increasing the sample size have on a confidence interval?
   A) No change  C) Increase the width of the interval  B) Change the confidence level  D) Decrease the width of the interval

10) A seed company has a test plot in which it is testing the germination of a hybrid seed. They plant 50 rows of 40 seeds per row. After a two-week period, the researchers count how many seeds per row have sprouted. They noted that the least number of seeds to germinate was 33 and some rows had all 40 germinate. The germination data is given below in the table. The random variable X represents the number of seeds in a row that germinated and P(x) represents the probability of selecting a row with that number of seeds germinating. Determine the expected number of seeds per row that germinated.

<table>
<thead>
<tr>
<th>x</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x)</td>
<td>0.02</td>
<td>0.06</td>
<td>0.10</td>
<td>0.20</td>
<td>0.24</td>
<td>0.26</td>
<td>0.10</td>
<td>0.02</td>
</tr>
</tbody>
</table>

A) 36.1  B) 41.6  C) 36.5  D) 36.8  E) 37.0

\[ E(V) = \mu = \sum x \cdot P(x) = 36.8 \]

---

**PART II Instructions:** Questions 11 - 20 are open response. Answer all TEN questions carefully and completely, for full credit you must show all appropriate work and clearly indicate your answers.

11) The owner of a computer repair shop has determined that their daily revenue has mean $7200 and standard deviation $1200. The daily revenue totals for the next 30 days will be monitored. What is the probability that the mean daily revenue for the next 30 days will exceed $7500? Round your answer to 4 decimal places.

\[ Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{7500 - 7200}{1200/\sqrt{30}} \approx 1.37 \]

\[ P(\bar{X} > 7500) = P(Z > 1.37) \]

\[ = 1 - 0.9147 \]

\[ = 0.0853 \]
12) The costs in dollars of a random sample of 20 college textbooks are given in the stem-and-leaf plot below.

<table>
<thead>
<tr>
<th>Stem</th>
<th>Leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7 8</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>0 2/3</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>1 6 9/9</td>
</tr>
<tr>
<td>8</td>
<td>2 4 4</td>
</tr>
<tr>
<td>9</td>
<td>0/3 5 7</td>
</tr>
<tr>
<td>10</td>
<td>5 7</td>
</tr>
</tbody>
</table>

Legend: 2|7 represents $27

\[ Q_1 = \frac{42 + 53}{2} \]
\[ Q_3 = \frac{90 + 93}{2} \]

i) Find the five number summary for this data set. Include the name or correct symbol for each of the numbers as well as its value.

\[ \text{minimum} = 27 \quad Q_1 = 47.5 \quad \text{median} = 79 \]
\[ Q_3 = 91.5 \quad \text{maximum} = 107 \]

ii) Draw a boxplot for this data set.

![Boxplot]

iii) Use complete sentences to briefly describe the shape of the distribution for this data.

The data is skewed left.
13) Find the mean, median, and mode of the following statistic students' test scores. Round to the nearest tenth if necessary.

68 73 82 82 86 87 90 91 97

mean = \text{ } 83.8

median = \text{ } 84

mode = \text{ } 82

14) A physical fitness association is including the mile run in its secondary-school fitness test. The time for this event for boys in secondary school is known to have a normal distribution with a mean of 470 seconds and a standard deviation of 40 seconds. The fitness association wants to recognize the fastest 10% of the boys with certificates of recognition. What time would the boys need to beat in order to earn a certificate of recognition from the fitness association? Round to the nearest second.

\[ Z \sigma + \mu = x \]

\[ -1.28 (40) + 470 \]

414 seconds
15) The data below are the final exam scores of 10 randomly selected history students and the number of hours they slept the night before the exam.

<table>
<thead>
<tr>
<th>Hours, x</th>
<th>3</th>
<th>5</th>
<th>2</th>
<th>8</th>
<th>2</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scores, y</td>
<td>65</td>
<td>80</td>
<td>60</td>
<td>88</td>
<td>66</td>
<td>78</td>
<td>85</td>
<td>90</td>
<td>90</td>
<td>71</td>
</tr>
</tbody>
</table>

The scatterplot for this data:

```
\begin{center}
\begin{tikzpicture}
\begin{axis}[
    xlabel=Hours of Sleep,
    ylabel=Exam Scores,
    xmin=1, xmax=10,
    ymin=60, ymax=90,
]
\addplot[only marks] table [x=Hours, y=Scores] {data.csv};
\end{axis}
\end{tikzpicture}
\end{center}
```

i) Based on the scatterplot, is it reasonable to suggest that there is a linear relationship between hours of sleep and exam scores? Yes or No (circle one)

ii) Find the correlation coefficient for the given data. Round to 4 decimal places. 

$r = 0.8465$

iii) Determine if there is a significant linear correlation. Report the critical value and state your conclusion.

$|r| > 0.632$

$\text{There is a significant linear correlation}$

iv) Find the equation of the least-squares regression line for this data. Round values to 2 decimal places.

$\hat{y} = 5.04 x + 56.11$

v) Use the regression equation to predict the exam score of a student who slept for 7 hours the night before the exam. Is the predicted exam score a good estimate? Briefly explain your answer.

When $x = 7$, the predicted exam score is 91. The estimate is good because a significant correlation exists and because $x = 7$ is within the range of the collected data.
16) A random sample of 20 college students is selected. Each student is asked how much time he or she spent on the Internet during the previous week. The following times (in hours) are recorded:

<table>
<thead>
<tr>
<th>8</th>
<th>12</th>
<th>3</th>
<th>15</th>
<th>16</th>
<th>5</th>
<th>16</th>
<th>5</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>12</td>
<td>13</td>
<td>4</td>
<td>4</td>
<td>11</td>
<td>9</td>
<td>17</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

i) Create a frequency and relative frequency table for this data. Use 3 as the lower class limit of the first class, and use a class width of 4.

<table>
<thead>
<tr>
<th>Class (optional)</th>
<th>Tally (optional)</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 - 6</td>
<td>HHH 11</td>
<td>7</td>
<td>0.35</td>
</tr>
<tr>
<td>7 - 10</td>
<td>III</td>
<td>3</td>
<td>0.15</td>
</tr>
<tr>
<td>11 - 14</td>
<td>HHH 1</td>
<td>6</td>
<td>0.3</td>
</tr>
<tr>
<td>15 - 18</td>
<td>III</td>
<td>4</td>
<td>0.2</td>
</tr>
</tbody>
</table>

ii) Create a relative frequency histogram for the data. Be sure to label your axes.

(hours spent on the Internet)
17) When 440 junior college students were surveyed, 200 said they have a passport. Construct a 95% confidence interval for the proportion of junior college students that have a passport. Round to the nearest thousandth.

\[ \hat{p} = \frac{200}{440} \]

The conditions are met:

\[ n \hat{p} (1 - \hat{p}) = 109 \geq 10, \quad n \leq 0.05 N \]

\[ E = 1.96 \sqrt{\frac{\hat{p} (1 - \hat{p})}{440}} \approx 0.0465 \]

\[ 0.408 < p < 0.501 \]

18) The National Association of Realtors estimates that 23% of all homes purchased in 2004 were considered investment properties. If a sample of 800 homes sold in 2004 is obtained what is the probability that at most 200 homes are going to be used as investment property? Round your answer to 4 decimal places.

\[ \rho = 0.23 \quad \hat{p} = \frac{200}{800} = 0.25 \]

\[ Z = \frac{\hat{p} - \mu_p}{\sigma_p} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.25 - 0.23}{\sqrt{(0.23)(0.77)/800}} \approx 1.34 \]

\[ P(\hat{p} \leq 0.25) = P(Z \leq 1.34) = 0.9099 \]
19) In 2010, 36% of adults in a certain country were morbidly obese. A health practitioner suspects that the percent has changed since then. She obtains a random sample of 1042 adults and finds that 393 are morbidly obese. Is this sufficient evidence to support the practitioner’s suspicion that the percent of morbidly obese adults has changed at the α = 0.1 level of significance? Round p to five decimal places when calculating the test statistic.

Are you using the Classical or P-Value approach? (circle one)

Null Hypothesis:

\[ H_0: \hat{p} = 0.36 \]

Alternative Hypothesis:

\[ H_1: \hat{p} \neq 0.36 \]

Test Statistic:

\[ Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx 1.15 \]

\[ n = 1042 \]

\[ \hat{p} = \frac{393}{1042} \]

\[ \alpha = 0.10 \]

Critical Value(s) or P-Value (circle which of these you are using):

\[ \text{critical values} = \pm 1.645 \]

\[ \text{p-value} = 0.2502 \]

\[ \text{Using technology} \quad \text{p-value} = 0.2485 \]

Conclusion about the Null Hypothesis:

Fail to reject \( H_0 \)

\( Z_0 \) is not in the critical region \( \quad \text{p-value} > \alpha \)

Do the data support the health practitioner’s suspicion? Answer with complete sentences.

There is not sufficient evidence to support the health practitioner's suspicion that the proportion of morbidly obese adults has changed since 2010.
20) A shipping firm suspects that the mean life of a certain brand of tire used by its trucks is less than 40,000 miles. To check the hypothesis, the firm randomly selects and tests 18 of these tires and finds that they have a mean lifetime of 39,300 miles with a standard deviation of 1200 miles. At $\alpha = 0.05$, test the shipping firm’s hypothesis. Assume that the life of the tires is normally distributed with no outliers. Show a complete solution including all your steps.

$H_0: \mu = 40,000 \quad H_1: \mu < 40,000$

Test statistic:

$$t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t_0 = \frac{39,300 - 40,000}{\frac{1200}{\sqrt{18}}} \approx -2.47$$

Critical value $t = -1.740$ or $p$-value $< 0.02$

$t_0$ is in the critical region, so reject $H_0$.

There is sufficient evidence to support the shipping firm’s suspicion that the mean life of the tires is less than 40,000 miles.