1. (2 pts) Given a normal probability distribution, which of the following is true?
   (a) The median is to the right of the mean.
   (b) The mean is to the right of the median.
   (c) The mean is equal to the median.
   (d) Any of the above can occur depending on whether the mean is positive or negative.

2. (2 pts) For events A and B in a sample space S, we are told \( P(A) = .6 \) and \( P(B) = .3 \) and \( P(A \text{ and } B) = .2 \). Which of the following is true?
   (a) A and B independent events.
   (b) A and B mutually exclusive events.
   (c) \( P(A \text{ or } B) = .9 \)
   (d) \( P(A \text{ or } B) = .7 \)

3. (2 pts) Which of the following is true about a binomial random variable for 10 trials with probability of success on each trial given as \(.1\)?
   (a) The probability of no successes is \(.1^{10}\)
   (b) Its standard deviation is equal to \(10(.9)(.1)\)
   (c) The probability of 10 successes is \(.1^{10}\)
   (d) None of the above.

4. An hypothesis test on the mean reports a \( P \)-value of .07. Which of the following is true?
   (a) The null hypothesis should be rejected if the level of significance is \( \alpha = .05 \).
   (b) The null hypothesis should be accepted if \( \alpha = .08 \).
   (c) The probability of a Type I error is .93.
   (d) The null hypothesis should be rejected if the level of significance is \( \alpha = .10 \).

5. Which of the following is true for confidence intervals coming from the same population?
   (a) A 95% confidence interval is always longer than a 99% confidence interval if they are both derived from the same sample.
   (b) Increasing the sample size increases the length of a 95% confidence interval.
   (c) Long confidence intervals provide more information than short confidence intervals.
   (d) The sample mean is always the midpoint of the confidence interval.
6. (2 pts) A two-tailed hypothesis test on the mean of an approximately normal population is conducted with a sample size of \( n = 15 \). For what values of \( t \) should the null hypothesis be rejected given that \( \alpha = .1 \)?

(a) \( t \leq -1.645 \) or \( t \geq 1.645 \)  
(b) \( t \leq -1.761 \) or \( t \geq 1.761 \)  
(c) \( t \leq -1.345 \) or \( t \geq 1.345 \)  
(d) \( t \geq -1.753 \) or \( t \geq 1.753 \)

7. (2 pts) The number of ways of choosing 3 objects from 90 objects without replacement is:

(a) \( \frac{90!}{3!} \)  
(b) \( \frac{90 \cdot 89 \cdot 88}{87!} \)  
(c) \( \frac{90 \cdot 89 \cdot 88}{3 \cdot 2 \cdot 1} \)  
(d) \( 3^{90} \)

8. (2 pts) Unaware that 33% of the 8,000 voters in his district still support him, a politician decides to estimate his political strength. A sample of 900 voters shows that 36% support him. Which of the following is true?

(a) The parameter of interest is 36%.  
(b) The statistic of interest is 33%.  
(c) The sample size is 8000.  
(d) None of the above.

9. (9 pts) Consider the sample of 30 numbers

\[
\begin{array}{ccccccccccccccc}
31 & 37 & 44 & 46 & 47 & 51 & 52 & 65 & 68 & 68 & 69 & 70 & 71 & 72 & 73 \\
75 & 75 & 76 & 76 & 77 & 78 & 79 & 80 & 81 & 85 & 89 & 92 & 97 & 101 & 104
\end{array}
\]

(a) (3 pts) Construct a stem and leaf plot for this data with stems 3,4,5,6,7,8,9,10

(b) (6 pts) Given that \( Q_1 = 65 \) and \( Q_2 = 74 \). Find \( Q_3 \) and the interquartile range, and construct a modified boxplot for the data.
10. (9 pts) Given the data 9,12,15,17,18,19,23,45,52,61,63,63. One has $\sum x = 397$ and $\sum x^2 = 18341$. Find:
(a) (1 pts) the mean
(b) (2 pts) the sample standard deviation
(c) (2 pts) the median
(d) (2 pts) the 83rd percentile of this data
(e) (2 pts) the percentile rank of the data value 45

11. The following data is for investigating the relation between salary in thousands (x) and average number of absences per year (y).

<table>
<thead>
<tr>
<th>Salary (x)</th>
<th>20</th>
<th>23</th>
<th>28</th>
<th>30</th>
<th>33</th>
<th>35</th>
<th>37</th>
<th>40</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absences (y)</td>
<td>2.4</td>
<td>2.2</td>
<td>1.9</td>
<td>2.1</td>
<td>1.5</td>
<td>1.4</td>
<td>1.3</td>
<td>0.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

For this data: $\sum x = 288 , \sum x^2 = 9660, \sum y = 13.7, \sum y^2 = 24.93, \sum xy = 398.2$.

(a) (4 pts) Find the equation of the least squares regression line.

(b) (2 pts) Given that the correlation coefficient is $r = -.945$. Does it indicate a good linear relationship? Explain.

(c) (3 pts) Assuming that the properties of the linear regression model are satisfied, is there sufficient evidence to support that $\beta_1 < 0$ where $\beta_1$ is the slope of the population regression line. Test at level of significance $\alpha = .01$. 
12. (9 pts) A business employs 300 men and 400 women. Seven percent of men and nine percent of women have been working there for more than 10 years. Let $A$ be the event the employee is a woman, and $B$ be the event the employee has been employed for more than 10 years.

(a) (1 pt) Find $P(A)$.

(b) (2 pts) Find $P(B)$.

(c) (2 pts) Find $P(A \text{ and } B)$.

(d) (2 pts) Find $P(A \text{ or } B)$.

(e) (2 pts) Find the probability an employee is female, given that the length of employment is more than 10 years.

13. (9 pts) The scores of the 1977 Mathematical Scholastic aptitude test were normally distributed with mean 620 and standard deviation 80.

(a) What score indicates a percentile rank of 43?

(b) What proportion of scores are between 460 and 700?

(c) What proportion of scores are greater than 700?

(d) What proportion of scores are less than 700?
14. In a recent survey of 250 randomly selected engineering graduates starting work in the United States, it was found that average annual starting salary for the sample was $43,200 with a standard deviation of $8,000.

(a) (6 pts) Conduct an hypothesis test to determine whether the average starting salary for engineers is more than $42,000 per year using a significance level of $\alpha = 0.01$. Make sure to state $H_0$, $H_a$, the critical region, and your conclusion.

(b) (3 pts) If the true mean starting salary for engineers in the United States is actually $43,000 with standard deviation $8,000, what is the probability that a random sample of size 250 engineers will have a sample mean starting salary greater than $42,000?

15. A “Gallop” poll of 1100 randomly selected Americans found that 49 percent think the Elian Gonzales story was the top news event of 2000.

(a) (5 pts) Find a 99 percent confidence interval for the true proportion of Americans that think the Elian Gonzales story was the top news event of 2000.

(b) (2 pts) Explain in words what this interval means.

(c) (2 pts) Based on this information, would you be comfortable saying that less than 50 percent of Americans think the Elian Gonzalez story was the top news event in 2000. Explain.
16. In a 1993 survey of 150 Education graduates and 200 Social Science graduates, the following data were obtained for their average starting salaries.

<table>
<thead>
<tr>
<th>Major</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>24,500</td>
<td>2600</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>21,500</td>
<td>2400</td>
</tr>
</tbody>
</table>

(a) (5 pts) Conduct an hypothesis test to determine whether $\mu_1 - \mu_2 < 3500$ using a 5% level of significance where $\mu_1$ is the population mean salary for the Education graduates and $\mu_2$ is the population mean salary for the Social Science graduates. Make sure to state $H_0$, $H_a$, the rejection region, and your conclusion.

(c) (4 pts) Report the $P$-value of your test, and explain what it means.

17. (9 pts) A person decides to test whether a die is fair. On 120 tosses, the following results were observed.

<table>
<thead>
<tr>
<th>$x$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>17</td>
<td>25</td>
<td>22</td>
<td>15</td>
<td>24</td>
<td>17</td>
</tr>
</tbody>
</table>

Conduct an hypothesis test at a level of significance of $\alpha = .05$ to determine whether the die is fair. Make sure to state: (i) Hypotheses (ii) Critical region (iii) Conclusion (iv) Justification that you could use the test you used.
18. (a) (3 pts) A shoe company claimed runners would record quicker times on average when using their brand of sneaker. To test this, 10 runners were chosen and were timed in a 100 meter dash with their own shoes, and timed in another 100 meter dash with the company’s shoes. To test the differences of the means, would you use a dependent or independent sample method? Explain. What distribution would you use? What additional assumptions must be made to use this distribution?

(b) A company wishes to check whether the mean weekly production levels at their five factories are equal using the method of analysis of variance.
(i) (2 pts) State what assumptions should be made.

(ii) (1 pts) State the null hypothesis.

(iii) (3 pts) The company took samples of size 11 from each factory and found the means and standard deviations from those samples:

<table>
<thead>
<tr>
<th>Factory</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}$</td>
<td>20.75</td>
<td>15.50</td>
<td>16.50</td>
<td>15.50</td>
<td>9.00</td>
</tr>
<tr>
<td>$s^2$</td>
<td>30.25</td>
<td>25.67</td>
<td>27.67</td>
<td>29.67</td>
<td>20.67</td>
</tr>
</tbody>
</table>

Compute the $F$ statistic and state the degrees of freedom for the numerator and denominator. Because you don’t have the table, you do not need to find the critical region or report the conclusion of the test.