PLEASE PRINT THE FOLLOWING INFORMATION:

Name: ____________________________  Instructor: ____________________________

Student ID #: ____________________  Section/Time: ____________________________

THIS EXAM HAS TWO PARTS.

PART I.
Part I consists of 30 multiple choice questions. Each correct answer is scored 2 points; each incorrect (or blank) answer is scored 0, so there is no penalty for guessing. You may do calculations on the test paper, but your answers must be marked on the OPSCAN sheet with a soft lead pencil (HB or No. 2 lead). Any question with more than one choice marked will be counted as incorrect. If more than one choice seems correct, choose the one that is most complete or most accurate. Make sure that your name and ID number are written and correctly bubbled on the OPSCAN sheet.

PART II.
Part II consists of 3 free response questions, with values as indicated. You must show all work in the space provided or elsewhere on the exam paper in a place that you clearly indicate. Work on loose sheets will not be graded.

FOR DEPARTMENT USE ONLY:

Part II.

<table>
<thead>
<tr>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part I</th>
<th>Part II</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Part I

Problems 1 through 3 pertain to the following sample data:

9, −5, 6, 0, 4, 6, 3, −1, 1

1. The sample mean computed from this data set is about:

(a) 2.1  (b) 3.0  (c) 2.6  (d) 4.0  (e) 9.0

2. The sample median computed from this data set is about:

(a) 3.0  (b) 2.5  (c) 4.5  (d) 2.1  (e) 4.0

3. The sample standard deviation computed from this data set is about:

(a) 2.9  (b) 4.3  (c) 2.8  (d) 4.5  (e) 4.1

4. If the number 10 is added to each value of a data set, then:

(a) The mean goes up by 10 and the standard deviation stays the same.
(b) The mean goes up by 10 and the standard deviation goes up by $\sqrt{10}$.
(c) The mean goes up by 10 and the standard deviation goes down by $\sqrt{10}$.
(d) The mean and the standard deviation both stay the same.
(e) It is impossible to tell what happens to both the mean and the standard deviation.

5. Weights of one year old boys have a bell-shaped distribution with mean 27 pounds and standard deviation 3 pounds. The proportion of all one year old boys who weigh at least 30 pounds is about:

(a) 84%  (b) 95%  (c) 68%  (d) 5%  (e) 16%

6. A mother is told that her one year old son’s weight is the 75th percentile. This implies that:

(a) Her son weighs 75% of the average weight of all one year old boys.
(b) Her son weighs more than 75% of all one year old boys.
(c) Her son weighs less than 75% of all one year old boys.
(d) The z-score of her son’s weight is .75.
(e) None of the above.
Problems 7 through 10 pertain to the following situation:

Individuals in a sample were classified by age and Body Mass Index (BMI), with the following results.

<table>
<thead>
<tr>
<th>BMI</th>
<th>B: below healthy</th>
<th>H: healthy</th>
<th>A: above healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y: 18–44</td>
<td>21</td>
<td>295</td>
<td>471</td>
</tr>
<tr>
<td>M: 45-64</td>
<td>12</td>
<td>283</td>
<td>807</td>
</tr>
<tr>
<td>O: 65 and over</td>
<td>8</td>
<td>179</td>
<td>413</td>
</tr>
</tbody>
</table>

Suppose an individual is selected at random.

7. $P(A)$, the probability that his BMI is above a healthy value is about:

(a) .169 (b) .279 (c) .181 (d) .679 (e) .562

8. $P(Y \text{ or } H)$, the probability that either he is young or his BMI is at a healthy level is about:

(a) .620 (b) .304 (c) .502 (d) .732 (e) .316

9. The probability that he is young given that his BMI is at a healthy level is about:

(a) .375 (b) .316 (c) .304 (d) .620 (e) .390

10. The events $Y$ and $H$ are

(a) independent because age and weight are unrelated
(b) independent because $P(Y \text{ and } H) \neq P(Y)P(H)$
(c) independent because $P(Y) + P(H) \neq 1$
(d) dependent because $P(Y \text{ and } H) \neq P(Y)P(H)$
(e) dependent because $P(Y) + P(H) \neq 1$

11. The Chamber of Commerce of a certain city wishes to estimate the average incomes of its residents at 90% confidence and to within $500. If it estimates that the standard deviation of the incomes to be about $6,500, what is the minimum size of the sample that it must take?

(a) 458 (b) 278 (c) 49 (d) 22 (e) 650
12. To estimate the current average retirement age of U. S. citizens a researcher sampled 45 recently retired individuals. The sample data yielded mean 62.3 years and standard deviation .9 year. A 95% confidence interval for the true mean retirement age of all recently retired individuals is about:

(a) 62.3 ± .04  (b) 62.3 ± .32  (c) 62.3 ± .27  (d) 62.3 ± .22  (e) 62.3 ± .16

Problems 13 and 14 pertain to the following situation:
Every day for six months a commuter records the number \( x \) of minutes his train ride to work takes, and obtains the probability distribution:

\[
\begin{array}{c|cccccccc}
   x & 23 & 24 & 25 & 26 & 27 & 28 & 29 \\
P(x) & .09 & .14 & .18 & .23 & .17 & .11 & .08 \\
\end{array}
\]

13. The probability that a randomly selected train ride will last at least 25 minutes is about:

(a) .59  (b) .23  (c) .77  (d) .18  (e) .41

14. The average length of the daily train ride in minutes is about:

(a) 26.5  (b) 25.0  (c) 27.0  (d) 25.9  (e) 26.1

15. One in five $1 lottery tickets wins the buyer some kind of payoff. If a person buys five tickets, the probability that none will be a winning ticket is about:

(a) .20  (b) .33  (c) .80  (d) .48  (e) .52

16. If \( z \) denotes the standard normal random variable, then \( P(1.38 \leq z \leq 3.91) \) is about:

(a) .9162  (b) .0327  (c) .1805  (d) .2127  (e) .0838

17. Personal best finishing times for a particular race in high school track meets are normally distributed with mean 5.27 minutes and standard deviation .42 minutes. If the qualifying time for this event for the regional championship is to be set so that only four percent of all runners qualify, then that qualifying time in minutes is about:

(a) 4.54  (b) 3.52  (c) 4.69  (d) 6.01  (e) 5.46
Problems 18 and 19 pertain to the following information:

The number of hours sleep typically gotten by a randomly selected college student is normally distributed with mean 7 and standard deviation 1.5.

18. The probability that a randomly selected college student typically sleeps at least eight hours per night is about:

(a) .2486  (b) .2514  (c) .3548  (d) .7486  (e) .1769

19. Ten college students are selected at random. The probability that the average number of hours of sleep that they typically get each night is at least eight is about:

(a) .0321  (b) .1769  (c) .2514  (d) .0174  (e) .0000

20. In order to estimate the proportion of personal bankruptcies that are a result of medical bills a research group took a sample of 500 bankruptcy cases and counted 258 that resulted from medical bills. A 99% confidence interval for the proportion of all personal bankruptcies that result from medical bills is:

(a) .516 ± .029  (b) .516 ± .037  (c) .516 ± .052  (d) .516 ± .044  (e) .516 ± .058

21. A study to compare the means of the age \( x \) at first marriage of men and women obtained the information shown, where the men and women were sampled independently of one another. We may assume that the populations of ages are normally distributed with approximately equal standard deviations.

<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( \bar{x} )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>male</td>
<td>50</td>
<td>26.8</td>
<td>.21</td>
</tr>
<tr>
<td>female</td>
<td>50</td>
<td>25.1</td>
<td>.19</td>
</tr>
</tbody>
</table>

The correct formula to employ to use these data to construct a 90% confidence interval for the difference in mean age at first marriage between men and women is:

(a) \((\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s^2_p \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}\)

(b) \(b \pm t_{\alpha/2} \frac{s_e}{SS_{xx}}\)

(c) \(d \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}\)

(d) \(\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}\)

(e) \(\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}}\)
22. In a test of hypotheses of the form $H_0 : p = .37$ vs. $H_1 : p \neq .37$ a sample of size 2500 produced the test statistic $z = -1.272$. The $p$-value (observed significance) of the test is about:

(a) −.1020  (b) 0.0102  (c) .2040  (d) .1020  (e) −.2040

23. In a test of hypotheses $H_0 : \mu = 98.6$ versus $H_1 : \mu > 98.6$, the rejection region is the interval $[2.306, \infty)$, the value of the sample mean computed from a sample of size 9 is $\bar{x} = 99.7$, and the value of the test statistic is $t = 2.118$. The correct decision and justification are:

(a) Do not reject $H_0$ because the sample is small.
(b) Do not reject $H_0$ because $2.118 < 2.306$.
(c) Reject $H_0$ because 99.7 is larger than 98.6.
(d) Reject $H_0$ because 99.7 is larger than 2.118.
(e) Reject $H_0$ because 2.118 lies in the rejection region.

24. In the test of hypotheses $H_0 : \mu = 120$ versus $H_1 : \mu \neq 120$ at the 1% level of significance, when the sample size is 38 and $\sigma$ is unknown the rejection region will be the interval or union of intervals:

(a) $[2.431, \infty)$
(b) $(-\infty, -2.576] \cup [2.576, \infty)$
(c) $[2.429, \infty)$
(d) $(-\infty, -2.429] \cup [2.429, \infty)$
(e) $(-\infty, -2.715] \cup [2.715, \infty)$

25. A salesman knows that 3% of all shoppers will buy a product from his point of purchase display. If he can have it set up in a store that has 900 customers per day, then the average number of sales per day from his display will be about:

(a) 270  (b) 300  (c) 54  (d) 27  (e) 30

26. How large a sample is required to obtain a 95% confidence interval for the proportion of all small businesses that saw an increase in business last month (in comparison with the previous month) to within 2 percentage points?

(a) 1201  (b) 241  (c) 1225  (d) 2401  (e) 48
Problems 27 and 28 pertain to the following situation.

In a survey of 1250 people it was found that 733 favor increased federal regulation of the financial sector.

27. The setup of the null and alternative hypotheses to test whether there is sufficient evidence to conclude that a majority of people (i.e., more than 50%) favor increased regulation is:

(a) $H_0 : p = .50$ vs. $H_1 : p > .50$
(b) $H_0 : p = .50$ vs. $H_1 : p < .50$
(c) $H_0 : p = .50$ vs. $H_1 : p \neq .50$
(d) $H_0 : p = .59$ vs. $H_1 : p \neq .59$
(e) $H_0 : p = .50$ vs. $H_1 : p > .50$

28. The value of the test statistic is:

(a) 6.109 (b) 3.054 (c) 1.932 (d) 6.203 (e) .586

Problems 29 and 30 pertain to the following information.

A study investigating the relationship between size $x$ (in hundreds of square feet) and the assessed value $y$ (in thousands of dollars) of 1025 randomly selected dwelling houses in a certain area yielded $r = .77$, $s_e = 53.75$, and the regression equation $\hat{y} = -3.117 + 9.445x$.

29. The predicted average assessed value of dwelling houses of size 2550 square feet is about:

(a) $240,816$ (b) $70,853$ (c) $237,731$ (d) $240,536$ (e) $243,965$

30. For each additional 100 square feet in size the average assessed value of dwelling houses in this area

(a) increases by about $944.50$
(b) increases by about $9,445$
(c) increases by about $94.50$
(d) increases by about $5375$
(e) changes by an amount that cannot be determined from the information given
Part II

1. To investigate whether online automobile insurance policies might be cheaper than the policies he can offer, a local insurance agent compared the automobile insurance premiums for five of his customers with price quotations for the same coverage from online insurance companies. Results are shown in the table. Test the null hypothesis that the mean of the difference in premiums for all automobile policies is zero versus the relevant alternative, as indicated below. Use \( \alpha = .01 \). Assume a normal distribution of differences.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Local</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>568</td>
<td>391</td>
</tr>
<tr>
<td>2</td>
<td>451</td>
<td>488</td>
</tr>
<tr>
<td>3</td>
<td>605</td>
<td>677</td>
</tr>
<tr>
<td>4</td>
<td>783</td>
<td>703</td>
</tr>
<tr>
<td>5</td>
<td>907</td>
<td>1008</td>
</tr>
</tbody>
</table>

(a) State the null and alternative hypotheses for the test. [2 points]

(b) State the correct formula for the test statistic. Justify your answer. [2 points]

(c) Construct the rejection region. [2 points]

(d) Compute the value of the test statistic, and make a decision. [4 points]

(e) State a conclusion, in the context of the problem, about the mean of the difference in premiums for all policies, based on the test you performed. [2 points]
2. Price $x$ (in 2005 dollars per pound) and consumption $y$ (in pounds per capita) of beef were sampled for ten randomly selected years. Summary information is:

\[ n = 10, \Sigma x = 36.19, \Sigma x^2 = 134.171502, \Sigma y = 774.74, \Sigma y^2 = 60739.2316, \Sigma xy = 2832.21065 \]

(a) Compute $SS_{xx}$, $SS_{xy}$, and $SS_{yy}$. [3 points]

(b) Compute the least squares regression line. [4 points]

(c) Find the coefficient of determination $r^2$ and explain what it means in the context of this problem. [3 points]

(d) Assuming that $s_e = 7.619827$, compute a 90% confidence interval for consumption when price is $3.10 per pound. [2 points]
3. The value used for the mean weight of elevator passengers by the state Division of Labor is 190 pounds. The Division wishes to investigate whether this value needs to be changed (upward or downward). A sample of fifty individuals in buildings having an elevator yielded sample mean 196 pounds and sample standard deviation 35 pounds. Assume the population is normally distributed.

(a) State the null and alternative hypotheses for the test. [2 points]

(b) State the correct formula for the test statistic. Justify your answer. [2 points]

(c) Construct the rejection region for $\alpha = .05$. [4 points]

(d) Compute the value of the test statistic, and make a decision. [4 points]

(e) State a conclusion about the mean weight of elevator passengers, based on the test you performed. [2 points]

(f) Compute a 95% confidence interval for the mean weight of all elevator passengers. [2 points]