STAT 1220
Common Final Exam

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

1. A real estate agent wishes to estimate the mean number of days on the market of all houses sold in the last two years in Mecklenburg County. To do so, he examines the sales data of fifty randomly selected houses sold in that period. The population of interest to the agent is:

(a) All houses sold in Mecklenburg County in the last two years.
(b) All houses sold that were not in the sample.
(c) The fifty houses in the sample.
(d) All houses sold except the fifty in the sample.
(e) None of the above.

Problems 2 through 4 pertain to the following sample data:

\[ 3, -2, 5, 4, 6, -8, 0, 5, 1 \]

2. The mean of this data set is about

(a) 4.4  (b) 14  (c) 4.2  (d) 1.6  (e) 1.1

3. The median of this data set is

(a) 6  (b) 4.5  (c) 2  (d) 0  (e) 3

4. The sample standard deviation of this data set is about

(a) 17.6  (b) 4.4  (c) 4.2  (d) 4.5  (e) 19.8

5. The standard deviation of a numerical data set measures the ________ of the data.

(a) average
(b) most frequent value
(c) variability
(d) size
(e) range

6. The distribution of the ages in years of missing children aged one to nine is roughly bell-shaped, with a mean of about 5.5 and a standard deviation of about 2. The proportion of such children who are three and a half or younger is about

(a) .84  (b) .68  (c) .32  (d) .95  (e) .16
7. A mother is told that her six year old son’s height is the 56th percentile. This means that

(a) The z-score of his height is $z = .56$.
(b) About 56% of the heights of all six year old boys are greater than his.
(c) He has attained 56% of his adult height.
(d) About 56% of the heights of all six year old boys are less than his.
(e) None of the above.

Problems 8 and 9 pertain to the following situation.
A card is selected at random from a well-shuffled ordinary deck consisting of 13 spades and 13 clubs (printed with black ink) and 13 hearts and 13 diamonds (printed in red ink). Three spades, three clubs, three hearts, and three diamonds are “face cards”.

8. The probability that the card is a heart or is a face card is about

(a) 0.23 (b) 0.42 (c) 0.19 (d) 0.48 (e) 0.25

9. The two events $R$: “the card is red” and $F$: “the card is a face card” are

(a) Independent because $P(R) + P(F) 
eq P(R \text{ or } F)$.
(b) Independent because $P(R) \cdot P(F) = P(R \text{ and } F)$.
(c) Independent because $P(R|F) = P(R)$.
(d) Dependent because $P(R) \cdot P(F) = P(R \text{ and } F)$.
(e) Dependent because $P(R) + P(F) 
eq P(R \text{ and } F)$.

Problems 10 and 11 are based on the following display of the results of a random survey done to investigate the relationship between species of household pets (restricted to dogs and cats) and the gender of their owners.

<table>
<thead>
<tr>
<th></th>
<th>Cat</th>
<th>Dog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>152</td>
<td>250</td>
</tr>
<tr>
<td>Female</td>
<td>172</td>
<td>112</td>
</tr>
</tbody>
</table>

10. The probability that a randomly selected owner of such a pet is male is about

(a) .41 (b) .47 (c) .53 (d) .72 (e) .59

11. The probability that a randomly selected owner of such a pet is male, given that the pet owned is a dog, is about

(a) .70 (b) .37 (c) .63 (d) .19 (e) .59
Problems 12 and 13 pertain to the following situation.
The number of minutes $x$ (rounded to the nearest whole minute) that a city bus takes to drive a particular route from one end to the other has the probability distribution shown.

<table>
<thead>
<tr>
<th>$x$</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>46</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$</td>
<td>.20</td>
<td>.25</td>
<td>.30</td>
<td>.15</td>
<td>.05</td>
<td>.05</td>
</tr>
</tbody>
</table>

12. The probability that the time (rounded to the nearest whole minute) of a randomly selected trip from one end of the route to the other will be at least 45 minutes is about

(a) 0.75   (b) 0.15   (c) 0.90   (d) 0.25   (e) .10

13. The average number of minutes (i.e., the mean) it takes the bus to drive the length of the route is about

(a) 44.0   (b) 42.6   (c) 46.2   (d) 43.8   (e) 44.5

14. It is estimated that 6% of the adult population has blood type O-negative (the “universal donor” type). If 25 blood donors will come to a blood donation center this afternoon, the probability that at least one will have type O-negative blood is about:

(a) .24   (b) .49   (c) .21   (d) .79   (e) .94

15. A convenience store sells 250 $1 lottery tickets a day. If 20% of all $1 lottery tickets win some sort of prize, then the average number of winning tickets sold by this store each day is about

(a) 23   (b) 44   (c) 102   (d) 11   (e) 50

In Problems 16 and 17, $z$ denotes the standard normal random variable.

16. $P(z \geq 0.44)$ is approximately

(a) 0.33   (b) -0.17   (c) 0.67   (d) 0.17   (e) 1.00

17. $P(-0.16 \leq z \leq 2.18)$ is approximately

(a) 0.49   (b) 0.55   (c) 0.48   (d) 0.42   (e) 0.04
Problems 18 and 19 pertain to the following situation.
Annual starting salaries in a certain region of the U. S. for college graduates with an engineering major are normally distributed with mean $39725 and standard deviation $7320.

18. The probability that a randomly selected graduate with an engineering major has a starting salary of at least $39000 is about

(a) .10  (b) .54  (c) .90  (d) .73  (e) .46

19. Suppose a school takes a sample of 125 such graduates and records the annual starting salary of each. The probability that the sample mean would be at least $39000 is about

(a) .87  (b) .46  (c) 1.00  (d) 0.00  (e) .54

20. Scores on a common final exam are normally distributed with mean 71 and standard deviation 9. Department policy is that the top 10% of students receive an A. The minimum exam score to be awarded an A is about:

(a) 95  (b) 83  (c) 98  (d) 87  (e) 90

21. If a trade association wishes to construct a 95% confidence interval for the proportion of all adults who would subscribe to an online shopping service, were it offered, to within .025, then (assuming no prior knowledge of the true proportion) it must choose a sample of size at least:

(a) 1083  (b) 2707  (c) 3604  (d) 1537  (e) 1230

22. A corporation's management wishes to estimate the difference in mean weekly sales between salesmen who are at least 20 lbs. overweight and those who are not. The data are: for 20 salesmen who are overweight, the mean weekly sales were $2355, sample standard deviation $183; for 35 other salesmen, the mean weekly sales were $2572, sample standard deviation $175. Assuming that the populations of both groups are normally distributed, with the same standard deviation, a 90% confidence interval for the difference in the population means is

(a) $152 to $282
(b) $134 to $300
(c) $186 to $248
(d) $50 to $50
(e) $167 to $267
23. In a random sample of eight beams constructed of engineered wood the load \( x \) at which each beam fails under successively heavier loading is recorded. The results are \( \bar{x} = 4085 \) and \( s = 221 \). Assuming that the population of maximum loads for all such beams is normally distributed, a 99% confidence interval for the mean maximum load for all such beams is:

(a) \( 4085 \pm (2.576) \frac{221}{\sqrt{8}} \)
(b) \( 4085 \pm (2.681) \frac{221}{\sqrt{8}} \)
(c) \( 4085 \pm (3.499) \frac{221}{\sqrt{8}} \)
(d) \( 4085 \pm (3.055) \frac{221}{\sqrt{8}} \)
(e) \( 4085 \pm (2.896) \frac{221}{\sqrt{8}} \)

24. An investigator wishes to estimate the mean number of flight hours of pilots who fly for regional airlines. Assuming that the standard deviation of the number of hours is 120 hours, the minimum sample size needed in order to construct a 99% confidence interval for the mean number of hours of flight time of all such pilots, to within 50 hours, is about

(a) 39     (b) 22     (c) 7     (d) 99     (e) 238

25. In a test of hypotheses of the form \( H_0 : \mu = 78.3 \) versus \( H_1 : \mu \neq 78.3 \) using \( \alpha = .01 \), when the sample size is 28 and the population is normal but with unknown standard deviation the rejection region will be the interval or union of intervals:

(a) \([2.771, \infty)\)
(b) \([2.473, \infty)\)
(c) \([2.576, \infty)\)
(d) \((-\infty, -2.771] \cup [2.771, \infty)\)
(e) \((-\infty, -2.473] \cup [2.473, \infty)\)

26. In a test of hypotheses \( H_0 : \mu = 2380 \) versus \( H_1 : \mu > 2380 \), the rejection region is the interval \([2.718, \infty)\), the value of the sample mean computed from a sample of size 12 is \( \bar{x} = 2413 \), and the value of the test statistic is \( t = 2.902 \). The correct decision and justification are:

(a) Do not reject \( H_0 \) because the sample is small.
(b) Do not reject \( H_0 \) because \( 2.718 < 2.902 \).
(c) Reject \( H_0 \) because 2413 is larger than 2380.
(d) Reject \( H_0 \) because 2.902 is positive.
(e) Reject \( H_0 \) because 2.902 lies in the rejection region.
27. In a test of hypotheses of the form $H_0 : p = .62$ vs. $H_1 : p < .62$ a sample of size 1200 produced the test statistic $z = -1.915$. The $p$-value (observed significance) of the test is about:

(a) 0.06  (b) 0.03  (c) 0.97  (d) -0.03  (e) 0.57

28. A large sample is collected in order to study the relationship between two variables $x$ and $y$ to investigate whether $x$ can be used to predict $y$. It is found that $r = -.68$. The true statement is:

(a) $x$ and $y$ are almost perfectly linearly correlated, and $y$ increases as $x$ is increased
(b) $x$ and $y$ are moderately linearly correlated, and $y$ increases as $x$ is increased
(c) $x$ and $y$ are almost perfectly linearly correlated, and $y$ decreases as $x$ is increased
(d) $x$ and $y$ are moderately linearly correlated, and $y$ decreases as $x$ is increased
(e) none of the statements above is correct

Problems 29 and 30 refer to a survey of 1450 people, in which it was found that 781 favor creation of a national high school graduation exam.

29. The setup of the null and alternative hypotheses to test whether there is sufficient evidence that a majority of people (i.e., more than 50%) favor creation of such an exam is:

(a) $H_0 : p = .54$ vs. $H_1 : p > .54$
(b) $H_0 : \hat{p} = .50$ vs. $H_1 : \hat{p} > .50$
(c) $H_0 : p = .50$ vs. $H_1 : p \neq .50$
(d) $H_0 : p = .50$ vs. $H_1 : p > .50$
(e) $H_0 : p = .50$ vs. $H_1 : p > .54$

30. The value of the test statistic is:

(a) 0.039  (b) 3.102  (c) 2.941  (d) 2.506  (e) 2.950
Part II

1. The administration of Whatsamatta University wishes to determine if the mean credit card debt of its students exceeds the national average of $2200. A survey of 50 students at Whatsamatta University gave a mean credit card debt of $2418 with sample standard deviation $683. The test is performed at the 1% level of significance.

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

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

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

   (d) State a conclusion about the mean credit card debt of all students at Whatsamatta University, based on the test you performed. [2 points]
2. Automotive engineers investigated the effect of a particular additive on fuel economy of automobiles by computing the mean miles per gallon of five vehicles operated normally on fuel with the additive and then operated normally on fuel without the additive, with results shown in the table.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>With</td>
<td>22.5</td>
<td>26.5</td>
<td>39.8</td>
<td>19.6</td>
<td>40.2</td>
</tr>
<tr>
<td>Without</td>
<td>22.7</td>
<td>25.2</td>
<td>36.7</td>
<td>19.6</td>
<td>39.9</td>
</tr>
</tbody>
</table>

Test whether the mean change in fuel economy is positive, at the 5% level of significance, in the following series of steps. The default is that the additive has no effect.

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

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

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

(d) State a conclusion about the mean change in fuel economy, based on the test you performed. [2 points]
3. A study of the relation between the waistline and percent body fat in men yielded the data shown (waistline measurement in inches).

<table>
<thead>
<tr>
<th>waistline (x)</th>
<th>34.7</th>
<th>46.5</th>
<th>32.8</th>
<th>41.3</th>
<th>29.8</th>
<th>39.3</th>
<th>35.7</th>
<th>29.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>percent fat (y)</td>
<td>21.8</td>
<td>33.6</td>
<td>13.4</td>
<td>27.1</td>
<td>13.7</td>
<td>30.2</td>
<td>18.5</td>
<td>11.2</td>
</tr>
</tbody>
</table>

\[ \bar{x} = 36.1375 \quad \bar{y} = 21.1875 \quad SS_{xx} = 248.53875 \quad SS_{xy} = 332.67375 \quad SS_{yy} = 494.30875 \]

(a) Find the proportion of the variability in percent body fat that is accounted for by the size of the waistline. Explain fully. [2 points]

(b) Find the regression line for predicting \( y \) from \( x \). [4 points]

(c) Construct a 95% confidence interval for the average percent body fat of all men whose waistline size is 34.0 inches. [4 points]