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

1. In order to estimate the proportion of all U. S. college students who are registered to vote, a political party took a random sample of 1600 students. In the sample 688 students were registered to vote. The population of interest to the party is:

(a) The 1600 students surveyed.
(b) The 688 registered voters in the sample.
(c) All registered voters in the U. S.
(d) All college students who are registered to vote.
(e) All college students.

Problems 2 through 4 pertain to the following sample data:

\[ 2, -7, 3, 3, 0, -1, -4, 2, 1 \]

2. The mean of this data set is about

(a) 2.6  (b) -0.1  (c) 4.5  (d) 1.1  (e) 2

3. The median of this data set is

(a) -2  (b) 0  (c) 1  (d) -1  (e) 2

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

(a) 3.41  (b) 1.42  (c) 3.21  (d) 2.56  (e) 11.6

5. The distribution of the times a city bus takes to complete a particular route at off-peak hours is roughly bell-shaped with mean 73 minutes and standard deviation 1.5 minutes. The proportion of times that the route is completed within 76 minutes is about

(a) 0.975  (b) 0.95  (c) 0.680  (d) 0.84  (e) 0.50

6. Twelve percent of all males aged 18–30 fail to meet physical standards for training as a fireman. Eight percent have an educational deficiency. Four percent fail to meet both the physical and educational standards. The probability that a randomly selected man will fail to qualify for training (i.e., will fail to meet the physical or educational standard) is

(a) 0.48  (b) 0.20  (c) 0.16  (d) 0.05  (e) 0.24
Problems 7 through 9 pertain to the data in the following table, summarizing a random sample of 3600 college students who were classified according to GPA and the number of hours worked at non-academic jobs. Assume the proportions in the sample accurately reflect those in the population of all college students.

<table>
<thead>
<tr>
<th>GPA</th>
<th>few</th>
<th>many</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>684</td>
<td>432</td>
<td>1116</td>
</tr>
<tr>
<td>average</td>
<td>1728</td>
<td>360</td>
<td>2088</td>
</tr>
<tr>
<td>high</td>
<td>360</td>
<td>36</td>
<td>396</td>
</tr>
<tr>
<td>Total</td>
<td>2772</td>
<td>828</td>
<td>3600</td>
</tr>
</tbody>
</table>

7. The probability that a randomly selected student has a low GPA is

(a) 0.43  (b) 0.77  (c) 0.25  (d) 0.18  (e) 0.31

8. The probability that a randomly selected student has a low GPA given that the number of hours he works is many is

(a) 0.52  (b) 0.39  (c) 0.31  (d) 0.12  (e) 0.23

9. The events \( L \): "has a low GPA" and \( M \): "number of hours worked is many" are

(a) Independent because \( P(L) = P(L|M) \).
(b) Independent because \( P(L) \neq P(L|M) \).
(c) Dependent because \( P(L) \neq P(L|M) \).
(d) Dependent because \( P(L) \neq P(L|M) \).
(e) Dependent because \( P(L \text{ and } M) \neq 0 \).

Problems 10 and 11 are based on the following probability distribution of the number \( x \) of vehicles lined up at the drive-through of a fast food restaurant at 4:00 pm.

<table>
<thead>
<tr>
<th>( x )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(x) )</td>
<td>0.18</td>
<td>0.26</td>
<td>0.23</td>
<td>0.17</td>
<td>0.12</td>
<td>0.04</td>
</tr>
</tbody>
</table>

10. On a randomly selected day the probability that at least one car will be at the drive-through at 4:00 pm is

(a) 0.82  (b) 0.26  (c) 0.44  (d) 0.18  (e) 0.38

11. The average number of cars lined up at the drive-through at 4:00 pm is about

(a) 1.1  (b) 1.5  (c) 2.5  (d) 3.0  (e) 1.9
12. Two evenly matched teams (meaning that the probability that either one will win a randomly selected game between them is 0.5) have a best two out of three play-off for a championship. Assuming that the result of one game is independent of the others, the probability that one or the other team will win the first two games (so there is no third game) is

(a) 0.50  (b) 0.25  (c) 0.75  (d) 0.67  (e) 0.33

13. Twenty-seven percent of airline passengers prefer an aisle seat. A particular flight has ten empty seats, of which three are aisle seats. The probability that exactly three of the next ten people who book a seat on this flight will prefer an aisle seat is about

(a) 0.81  (b) 0.26  (c) 0.09  (d) 0.54  (e) 0.32

14. For the standard normal random variable $z$, $P(-2.01 < z < 0.36)$ is

(a) 0.6184  (b) 0.6628  (c) 0.6406  (d) 0.8214  (e) 0.2128

15. If $x$ is a normally distributed random variable with mean 762 and standard deviation 34 then $P(x \geq 704)$ is

(a) 0.8212  (b) 0.8311  (c) 0.9564  (d) 0.0436  (e) 0.5436

16. Scores on a standardized exam given to school children are normally distributed with mean 203 and standard deviation 12. The exam score that is the 33rd percentile is about

(a) 208  (b) 67  (c) 198  (d) 236  (e) 190

Problems 17 and 18 are based on the following information: weights of carry-on luggage on international flights are normally distributed with mean 14.3 lb. and standard deviation 3.8 lb.

17. The probability that a randomly selected carry-on bag weighs at most 15 lb. is about

(a) 0.43  (b) 0.76  (c) 0.83  (d) 0.57  (e) 0.04

18. The probability that in a random sample of 36 bags the mean weight will be at most 15 lb. is about

(a) 0.93  (b) 0.54  (c) 0.36  (d) 0.87  (e) 0.73
19. In a random sample of five children, the age at which they first began to combine words in speaking had mean 16.5 months and standard deviation 9.6 months. Assuming normality of the population, a 90% confidence interval for the age at which all children begin to combine words is

(a) (12.4, 20.6)  (b) (7.8, 25.2)  (c) (9.9, 23.1)  (d) (10.2, 22.8)  (e) (7.3, 25.7)

20. In a random sample of 425 elderly people, 102 suffered from sleep apnea. A 95% confidence interval for the proportion of all elderly people who suffer from sleep apnea is

(a) (0.21, 0.27)  (b) (0.20, 0.28)  (c) (0.19, 0.29)  (d) (0.15, 0.31)  (e) (0.23, 0.25)

21. A travel agency wishes to estimate, at 99% confidence and to within $100, the mean airfare for flights between the east coast of the U. S. and continental Europe. If the standard deviation is assumed to be $275, the minimum estimated sample size is

(a) 41  (b) 8  (c) 51  (d) 73  (e) 102

22. A university administrator wishes to estimate to within three percentage points the proportion of all students who graduate within four years of their first enrollment, at 90% confidence. The estimated minimum sample size needed is

(a) 457  (b) 1068  (c) 752  (d) 330  (e) 1267

23. An advertising agency wishes to estimate, to within 95% confidence and to within two percentage points, the proportion of all adults who recognize its client’s logo. Knowing that two years ago the proportion was 19% and working from that figure the estimated minimum sample size needed is

(a) 2401  (b) 1042  (c) 1479  (d) 966  (e) 1566
Problems 24 and 25 pertain to the following information: a multinational beverage corporation claims that its U.S. market share is 23%. An independent research firm wishes to test this claim at the 5% level of significance. Their research, based on a sample of size 1750, yields a sample proportion of 21% yielding the test statistic $z = -1.99$.

24. In the test $H_0 : p = 0.23$ versus $H_1 : p < 0.23$ the decision is

(a) Do not reject $H_0$ because the test statistic does not fall in the rejection region.
(b) Do not reject $H_0$ because the test statistic is negative.
(c) Reject $H_0$ because $0.21 < 0.23$.
(d) Reject $H_0$ because the test statistic falls in the rejection region.
(e) Reject $H_0$ because the test statistic is negative.

25. The p-value (or observed significance) of the test is

(a) 0.04    (b) 0.02    (c) 0.01    (d) 0.10    (e) 0.05

Problems 26–28 pertain to the following information: a sample of 200 people in the year 2000 that measured the time spent per week on the internet yielded sample mean 6.8 hours and sample standard deviation 7.7 hours. A similar sample of 500 people in 2005 yielded sample mean 7.3 hours and sample standard deviation 10.9 hours. Test whether the mean time spent on the internet increased between 2000 and 2005 against the default that it remained the same, at the 10% level of significance. Assume that the populations of times are normally distributed with the same standard deviation.

26. Choosing the times in 2000 as Population 1 and the times in 2005 as Population 2 the value of the test statistic is

(a) -0.592    (b) -1.335    (c) -2.053    (d) -0.701    (e) -2.673

27. The rejection region is

(a) $(-\infty, -2.326]$    (b) $(-\infty, -1.645]$    (c) $(-\infty, -1.960]$    (d) $(-\infty, -1.282]$    (e) $(-\infty, -2.576]$

28. The decision is:

(a) Reject $H_0$ because the test statistic falls in the rejection region.
(b) Reject $H_0$ because $7.3 > 6.8$.
(c) Reject $H_0$ because the test statistic is negative.
(d) Do not reject $H_0$ because the test statistic does not fall in the rejection region.
(e) Do not reject $H_0$ because the test statistic is negative.
Problems 29 and 30 refer to a study relating the speed $x$ (in miles per hour) of passenger cars and level $y$ of noise (in decibels) they produced. From their data the researchers computed $\hat{y} = 71.4 + 0.39x$, $r = 0.957$, and $r^2 = 0.916$.

29. Consider the three statements

I. About 96% of the noise produced by a car is explained by its speed.
II. For every 1 mph increase in speed, noise produced increases by 0.39 dB.
III. The average sound level produced by all cars is 71.4 dB.

The correct statements are

(a) Only I.  (b) Only II.  (c) Only III.  (d) I and II.  (e) II and III.

30. The predicted mean level of noise produced by cars travelling 35 mph is

(a) 75 dB  (b) 58 dB  (c) 85 dB  (d) 106 dB  (e) 33 dB
Part II

1. A company compared sales performance produced by two different compensation programs by switching nine randomly selected salesmen from the usual program (Program 1) to a new program (Program 2). Sales in thousands of dollars generated by each salesman over time periods of the same length are shown. Test, at the 1% level of significance, whether means sales volume differs between the two programs, in the following series of steps.

<table>
<thead>
<tr>
<th>Salesman</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>55</td>
</tr>
<tr>
<td>B</td>
<td>22</td>
</tr>
<tr>
<td>C</td>
<td>34</td>
</tr>
<tr>
<td>D</td>
<td>22</td>
</tr>
<tr>
<td>E</td>
<td>25</td>
</tr>
<tr>
<td>F</td>
<td>61</td>
</tr>
<tr>
<td>G</td>
<td>55</td>
</tr>
<tr>
<td>H</td>
<td>36</td>
</tr>
<tr>
<td>I</td>
<td>68</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 about the difference in mean sales volumes under the two programs, based on the test you performed. [2 points]
2. Data on inflation and unemployment for 25 years were examined in order to discern a possible relationship between a given year's inflation rate $x$ and the following year's unemployment rate $y$. For example, for one year the rate of inflation was 1.6% ($x = 1.6$) and the unemployment the following year was 5.8% ($y = 5.8$). Summary information is:

\[
\begin{align*}
   n & = 25, & 1.1 \leq x \leq 8.9, & \quad 4.0 \leq y \leq 9.7, & \quad x = 3.36, & \quad y = 6.048 \\
   SS_{xx} & = 62.22, & SS_{xy} & = 31.808, & SS_{yy} & = 51.9024, & s_x & = 6.354
\end{align*}
\]

(a) Find the proportion of the variability in one year's unemployment that is explained by the previous year's rate of inflation. [2 points]

(b) Construct the regression line based on these data. [4 points]

(c) Use your answer to (b) to estimate next year's unemployment if this year's inflation is 3.5%. [2 points]

(d) Fill in the blanks and justify in the space provided: “For every one percentage point increase in the rate of inflation the unemployment rate [i]___________ (increases/decreases) by [ii]_________ percentage points. [4 points]

(e) Construct a 90% confidence interval for the unemployment rate, on average, in years for which inflation the previous year was 3.5%. [6 points]
3. Based on his long experience, a wildlife ranger knows that lengths of mature fish of a particular species of fish are normally distributed with mean 20.25 inches and standard deviation 1.25 inches.

*Note: if you cannot do part (a) then use the substitute answer \( p = 0.1384 \) for (a) (which is not correct) in the remaining two parts, for full credit on those parts if worked correctly.*

(a) Find the probability that a randomly selected such fish is at least 22 inches long. [4 points]

(b) A fisherman who caught and released four such fish one afternoon bragged to the ranger that three of them were 22 inches long. Compute the probability that at least three of four randomly selected fish will be 22 inches long or longer, using your answer to (a). [4 points]

(c) If the fisherman caught and released four fish on every outing, find the mean number of fish at least 22 inches long caught per outing, assuming no skill on the part of the fisherman (that is, as if the fish were merely a random selection). [2 points]