Please print the following information:

Name: ___________________________  Instructor: ___________________________

Student ID #: _____________________  Section/Time: _______________________

THIS EXAM HAS TWO PARTS

PART I. Consists of 30 multiple choice questions worth a total of 60 points. Read all questions carefully. You may do calculations on the test paper. Mark the number of the opscan sheet corresponding to the test question number with a Number 2 pencil or a mechanical pencil with HB lead. Mark only one answer; otherwise the answer will be counted as incorrect. In case there is more than one answer, mark the best answer. Please make sure that your name appears on the opscan sheet in the spaces provided.

PART II. This part consists of 3 questions (40 points in total). You MUST show all work for each question in the space provided to receive full credit for that question. If you write your explanations in another part of the test, please indicate accordingly.

At the end of the examination, you MUST hand in this test booklet, your answer sheet and all scratch paper.

FOR DEPARTMENTAL USE ONLY:
PART II:

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>16</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Part I  Part II  Total
1. On a test, a sample of five students has a mean score of 78. The scores of four students of them are: 87, 81, 76, 53. Find the test score of the fifth student.

(a) 65  
(b) 98  
(c) 93  
(d) 71  
(e) 78

2. The standard deviation of a data set measures the _____ of the data set.

(a) most frequent value  
(b) variability  
(c) size  
(d) position  
(e) center

Use the following information to answer the questions 3 to 4.
Consider the sample data: 20, 18, 22, 26, 28, 21.

3. The **median** of the data is

(a) 22.5  
(b) 21.5  
(c) 21  
(d) 23  
(e) None of the above

4. The **sample standard deviation** of the data is about

(a) 14.30  
(b) 3.78  
(c) 11.92  
(d) 3.45  
(e) None of the above
A recent survey revealed that the distribution of the monthly utility bills of 3-bedroom houses has a mean of $125 and a standard deviation of $15. Assume that this distribution is bell-shaped and symmetric. Use the Empirical Rule to answer questions 5 to 6.

5. Find the interval which includes 95% of the monthly utility bills of 3-bedroom houses.

   (a) (95,155)
   (b) (110,140)
   (c) (80,155)
   (d) (95,140)
   (e) (80,170)

6. Approximately what percentage of the 3-bedroom homes’ monthly utility bills will be between $95 and $140?

   (a) 84%
   (b) 95%
   (c) 81.5%
   (d) 99%
   (e) 68%

7. Tom took a Biology test and an English test. The Biology test had a mean of 68 and standard deviation of 6. The English test had a mean of 32 and standard deviation of 3.

   Tom scored 80 in Biology, and 40 in English. Which of the following statements is correct?

   (a) Tom did better in Biology because 80 > 40.
   (b) Tom’s z-scores were -2.00 in Biology, and -2.67 in English. He did better in Biology.
   (c) Tom’s z-scores were -2.00 in Biology, and -2.67 in English. He did better in English.
   (d) Tom’s z-scores were 2.00 in Biology, and 2.67 in English. He did better in English.
   (e) Tom’s z-scores were 2.00 in Biology, and 2.67 in English. He did better in Biology.
Use the following information for questions 8 to 10

The following table gives a two-way classification on gender and salary of 1000 employees in a company.

<table>
<thead>
<tr>
<th></th>
<th>&lt;$40,000</th>
<th>&gt;$40,000</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>200</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>150</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. If one employee of this company is selected at random, what’s the probability that the salary of the selected person is greater than $40,000?

(a) 0.40  
(b) 0.60  
(c) 0.55  
(d) 0.60  
(e) 0.65

9. If one employee is randomly selected, what’s the probability that the person is male and the salary of this person is less than $40,000?

(a) 0.20  
(b) 0.60  
(c) 0.65  
(d) 0.75  
(e) 0.35

10. If one employee is randomly selected, what is the probability that the person is female or the salary of the person is greater than $40,000?

(a) 0.25  
(b) 0.75  
(c) 0.65  
(d) 0.80  
(e) 0.40
Consider the following probability distribution for a discrete random variable $X$. Find the missing probability and answer questions 11-13.

<table>
<thead>
<tr>
<th>$X$</th>
<th>-2</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.1</td>
<td>?</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

11. Find $P(X > 0)$

(a) 0.2  
(b) 0.9  
(c) 0.7  
(d) 0.6  
(e) 0.3

12. The mean (expected value) of $X$ is

(a) 2.0  
(b) 2.4  
(c) -2.0  
(d) 1.3  
(e) 1.5

13. The standard deviation of $X$ is closest to

(a) 7.65  
(b) 2.77  
(c) 17.50  
(d) 4.18  
(e) None of the above
Use the following information for questions 14 to 16.

The time spent (in minutes) in reading newspaper for an adult per day can be approximated by a normal distribution with $\mu = 15$ minutes and $\sigma = 3$ minutes.

14. Find the probability that the reading time per day for a randomly selected adult is more than 18 minutes.

(a) 0.1587
(b) 0.8413
(c) 0.6826
(d) 0.6587
(e) 1.0000

15. If 200 adults are randomly selected, approximately how many of them will spend in reading newspaper per day between 12 minutes and 19.5 minutes.

(a) 32
(b) 168
(c) 187
(d) 136
(e) 155

16. What is the shortest time spent in reading newspaper for an adult per day that would still place him in the top 10%?

(a) 15.30
(b) 15.90
(c) 18.84
(d) 11.16
(e) 15.10
Use the following information for questions 17 to 18.
Scores on a certain assessment test had a mean of 280 and a standard deviation of 31.5. A random sample of 49 students taking this test is selected.

17. Identifying the mean and standard deviation of the sample mean score
   (a) $\mu_{\bar{x}} = 49, \sigma_{\bar{x}} = 31.5$
   (b) $\mu_{\bar{x}} = 49, \sigma_{\bar{x}} = 4.5$
   (c) $\mu_{\bar{x}} = 280, \sigma_{\bar{x}} = 31.5$
   (d) $\mu_{\bar{x}} = 280, \sigma_{\bar{x}} = 4.5$
   (e) $\mu_{\bar{x}} = 31.5, \sigma_{\bar{x}} = 4.5$

18. Find the probability that the sample mean score is at least 289.
   (a) 0.6141
   (b) 0.3859
   (c) 0.0228
   (d) 0.9772
   (e) 0.2500
Use the following information for questions 19 to 20.

The manager of the dairy section of a large supermarket chose a random sample of 300 egg cartons and found that 48 cartons had at least one broken egg. Let $p$ denote the proportion of all cartons which have at least one broken egg.

19. Find a point estimate for $p$ and also construct a 90% confidence interval for $p$.

(a) 0.160, (0.119, 0.201)
(b) 0.840, (0.799, 0.881)
(c) 48, (0.125, 0.195)
(d) 48, (0.119, 0.201)
(e) 0.160, (0.125, 0.195)

20. Which of the following actions would result in a confidence interval wider than the 90% confidence interval computed above?

(a) Increasing the sample size
(b) Computing an 80% confidence interval rather than a 90% confidence interval
(c) Computing an 85% confidence interval rather than a 90% confidence interval
(d) Computing a 95% confidence interval rather than a 90% confidence interval
(e) None of the above

21. A company wants to estimate the mean time (in hours) per week for an adult who uses computers at home. Find the minimum sample size needed in order to construct a 97% confidence interval for the mean time of computer usage at home for an adult, to within 0.25 hours. The company assumes $\sigma$ is 1 hour.

(a) 76
(b) 57
(c) 19
(d) 15
(e) 97

22. In the test of hypothesis $H_0 : \mu = 100$ vs $H_a : \mu \neq 100$, a sample of size 250 yields the standardized test statistic $z = 1.47$. Find the p-value for the test and state your conclusion at $\alpha = .10$.

(a) The p-value is 0.93; Decision: Reject $H_0$.
(b) The p-value is 0.93; Decision: Fail to reject $H_0$.
(c) The p-value is 0.14; Decision: Fail to reject $H_0$.
(d) The p-value is 0.07; Decision: Reject $H_0$.
(e) The p-value is 1.86; Decision: Fail to reject $H_0$. 


Use the following information for questions 23 to 25

In the past, a company used to produce 40 tons of a certain food per day on average. Now, using a new technology, the mean daily yield of the food for a random sample of 25 days of production is 45 tons with a standard deviation of 5 tons. Assume that the daily yield of this food is normally distributed.

23. Set up the null and alternative hypotheses to test whether the new technology has **increased** the mean daily yield.
   (a) $H_0 : \mu = 40$ versus $H_a : \mu \neq 40$
   (b) $H_0 : \mu \leq 40$ versus $H_a : \mu > 40$
   (c) $H_0 : \mu \leq 45$ versus $H_a : \mu > 45$
   (d) $H_0 : \mu \geq 40$ versus $H_a : \mu < 40$
   (e) $H_0 : \mu \geq 45$ versus $H_a : \mu < 45$

24. The value of the standardized test statistic is
   (a) -1.00
   (b) 1.00
   (c) -5.00
   (d) 5.00
   (e) None of the above

25. Find the rejection region and state your conclusion at $\alpha = 0.05$.
   (a) Rejection region: $z < -1.96$ or $z > 1.96$; Decision: Fail to reject $H_0$.
   (b) Rejection region: $z > 1.645$; Decision: Reject $H_0$.
   (c) Rejection region: $t < -1.711$; Decision: Fail to reject $H_0$.
   (d) Rejection region: $t > 1.711$; Decision: Reject $H_0$.
   (e) Rejection region: $t < -2.064$ or $t > 2.064$; Decision: Reject $H_0$. 
A car dealer claims that on average, cars owned by professors are newer than cars owned by students. Random samples of 40 faculty cars and 36 student cars are taken and their ages are measured. The data was shown in the following table:

<table>
<thead>
<tr>
<th>ages for faculty cars</th>
<th>ages for student cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_1 = 40$</td>
<td>$n_2 = 36$</td>
</tr>
<tr>
<td>$\bar{x}_1 = 5.9$</td>
<td>$\bar{x}_2 = 7.1$</td>
</tr>
<tr>
<td>$s_1 = 1.6$</td>
<td>$s_2 = 1.8$</td>
</tr>
</tbody>
</table>

26. Choose the correct hypothesis to test the claim.

(a) $H_0 : \mu_1 \geq \mu_2$ vs. $H_a : \mu_1 < \mu_2$
(b) $H_0 : \bar{x}_1 \leq \bar{x}_2$ vs. $H_a : \bar{x}_1 > \bar{x}_2$
(c) $H_0 : \mu_1 \leq \mu_2$ vs. $H_a : \mu_1 > \mu_2$
(d) $H_0 : \bar{x}_1 \geq \bar{x}_2$ vs. $H_a : \bar{x}_1 < \bar{x}_2$
(e) $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 \neq \mu_2$

27. Find the value of the standardized test statistic.

(a) 1.20
(b) -1.20
(c) 2.575
(d) 3.52
(e) -3.06

28. Find the rejection region at $\alpha = 0.05$.

(a) Rejection region: $z > 1.96$
(b) Rejection region: $z < -1.96$
(c) Rejection region: $z > 1.645$
(d) Rejection region: $z < -1.645$
(e) Rejection region: $z > -1.645$

29. Find the p-value of this test.

(a) 0.0011
(b) 0.025
(c) 0.10
(d) 0.05
(e) 0.9989
30. In a simple linear regression study between two variables x (the independent variable) and y (the dependent variable), a random large sample is collected and the coefficient of correlation \( r = -0.98 \) is calculated. Which of the following conclusion may be made?

(a) \( x \) and \( y \) are very poorly correlated since \( r \) close to -1.
(b) \( x \) and \( y \) are almost perfectly correlated, and \( y \) increases as \( x \) is increased.
(c) \( x \) and \( y \) are almost perfectly correlated, and \( y \) decreases as \( x \) is increased.
(d) \( x \) and \( y \) are moderately correlated, and \( y \) increases as \( x \) is increased.
(e) None of the above

**End of Multiple Choice Section**
1. A comparison of the dexterity test scores of 5 assembly-line employees with their hourly productivity resulted the following data.

<table>
<thead>
<tr>
<th>x = score on dexterity test</th>
<th>y = units produced in one hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>55</td>
</tr>
<tr>
<td>14</td>
<td>63</td>
</tr>
<tr>
<td>17</td>
<td>67</td>
</tr>
<tr>
<td>16</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
</tr>
</tbody>
</table>

\[ n = 5, \sum x = 70, \sum y = 306, \sum x^2 = 1006, \sum y^2 = 18984, \sum xy = 4362. \]

(a) [3 pts.] Find the coefficient of correlation between \( x \) and \( y \) and interpret its meaning in the context of the problem.

(b) [5 pts.] At \( \alpha = .05 \), test the significance of the coefficient of correlation.

(c) [5 pts.] Find the equation of the regression line between \( y \) and \( x \).
(d) [2 pts.] A prospective employee has scored $x = 15$ on the dexterity test. Find the predicted value of $y$ for this employee.

(e) [3 pts.] Construct a 95% prediction interval for the above employee’s productivity. Given that the standard error of the estimate is $s_e = 2.757$.

2. The following table lists SAT scores of five students before and after they took a preparatory course.

<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT score before course</td>
<td>700</td>
<td>840</td>
<td>830</td>
<td>860</td>
<td>690</td>
</tr>
<tr>
<td>SAT score after course</td>
<td>720</td>
<td>840</td>
<td>820</td>
<td>900</td>
<td>700</td>
</tr>
</tbody>
</table>

Define the difference $d = \text{Second score} - \text{First score}$. Assume that the distribution of the difference of SAT scores is approximately normal.

(a) [2 pts.] Find the sample mean and standard deviation of the differences.

(b) [3 pts.] State the correct hypotheses to test whether the course is effective in improving the SAT score.
(c) [3 pts.] Find the value of the standardized test statistic.

(d) [3 pts.] Find the rejection region at $\alpha = .05$ and state your conclusion in the context of the problem.

3. In a survey of 400 students in a school district, it is found that 340 students use computers in school. Test the claim that the proportion of students in this district who use computers in school is at least 80%.

(a) [2 pts.] State the correct hypotheses.

(b) [3 pts.] Find the value of the standardized test statistic.
(c) [3 pts.] Find the rejection region at $\alpha = .05$ and state your conclusion in the context of the problem.

(d) [3 pts.] Find the p-value for the test.