PART I. Consists of 25 multiple choice questions worth a total of 60 points. Read all questions carefully. You may do calculations on the test paper. Mark only one answer; otherwise the answer will be counted as incorrect. In case there is more than one answer, mark the best answer.

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.

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>13</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The following is used for questions 1, 2 and 3.

Given below are the ages of motorcyclists at the time they were fatally injured in traffic accidents.

17 38 27 14 18 16 21

1  Find the mean of the data.
   (a) 18.20
   (b) 20.58
   (c) 21.57
   (d) 25.65
   (e) 29.32

2  Find the standard deviation of the sample data.
   (a) 1.25
   (b) 3.35
   (c) 6.78
   (d) 8.38
   (e) 10.25

3  Find the median of the data.
   (a) 17
   (b) 18
   (c) 20
   (d) 22
   (e) 24
Use the following to answer questions 4, 5 and 6.
The IQ scores have a mean of 100 and a standard deviation of 15.

4 According to Chebycheb’s Theorem, what percentage of people will have IQ scores between 70 and 130?
   (a) At least 75%
   (b) At most 60%
   (c) At least 50%
   (d) At most 40%
   (e) At most 20%

5 If the distribution of the IQ scores can be approximated by normal distribution, then approximately what percentage of people will have IQ scores greater than 115?
   (a) 6.5%
   (b) 16%
   (c) 25%
   (d) 54%
   (e) 90%

6 The z-score of a person with IQ score of 140 is.
   (a) $z = 0.00$
   (b) $z = 1.25$
   (c) $z = 2.00$
   (d) $z = 2.67$
   (e) $z = 3.15$
The following is used for questions 7 and 8.

Based on the data from the National Health Survey, men’s heights are normally distributed with mean 69 in and the standard deviation 3.2 in.

7. The standard doorway height is 80 in. What proportion of men are too tall to fit through a standard doorway without bending?
   (a) .0003
   (b) .0044
   (c) .0351
   (d) .2090
   (e) .4207

8. If a statistician designs a house so that all of the doorway have heights that are sufficient for all men except the tallest 5%, what doorway height would be used?
   (a) 65.325 in.
   (b) 70.854 in.
   (c) 74.264 in.
   (d) 82.214 in.
   (e) 85.542 in.

The following is used for questions 9 and 10. A study by the University of Texas Southwestern Medical Center examined 626 people to see if there was an increased risk of contracting hepatitis C associated with having a tattoo.

<table>
<thead>
<tr>
<th></th>
<th>Tattoo</th>
<th>No tattoo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has hepatitis C</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td>No hepatitis C</td>
<td>88</td>
<td>495</td>
</tr>
</tbody>
</table>

9. If one person is randomly selected, what’s the probability that this person has hepatitis C and no tattoo?
   (a) 0.029
   (b) 0.040
   (c) 0.141
   (d) 0.181
   (e) 0.790
10. If one person is randomly selected, what’s the probability that this person does not have hepatitis C?
(a) 0.0111  
(b) 0.0288  
(c) 0.1405  
(d) 0.7907  
(e) 0.9313

The following is used for questions 11 and 12.
For women aged 18-24, the systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. 100 women between the ages of 18 and 24 are randomly selected, let $\bar{X}$ represent the mean systolic pressure of these 100 women.

11. Find the mean and standard deviation of $\bar{X}$, i.e., $\mu_{\bar{X}}, \sigma_{\bar{X}}$.
(a) 114.8, 131  
(b) 114.8, 13.1  
(c) 114.8, 1.31  
(d) 114.8, 13.1  
(e) 11.48, 1.31

12. What is the probability that the mean systolic blood pressure $\bar{X}$ is between 112.2 and 116.4?
(a) .8649  
(b) .3879  
(c) .1578  
(d) .0571  
(e) .0324
13. A random sample of 21 desktop PCs is selected. The mean life span is 6.8 years with a standard deviation of 2.4 years. Construct a 95% confidence interval for the mean life span of all desktop PCs. Assume that the life spans of all desktop PCs are approximately normally distributed.

(a) (5.85, 7.75 )
(b) (1.68, 3.12)
(c) (5.60, 8.00)
(d) (5.71, 7.89)
(e) (5.77, 7.83)

14. Which of the following statements are true about the sampling distribution of $\bar{x}$?

I. The mean of the sampling distribution is equal to the mean of the population.
II. The standard deviation of the sampling distribution is equal to the population standard deviation.
III. The shape of the sampling distribution is always approximately normal, when $n$ is sufficiently large ($n \geq 30$).

(a) II only
(b) III only
(c) I and II only
(d) I and III only.
(e) None of above.
15. A study of 4200 cell phone users found that 135 of them developed cancer of the brain or nervous system. Construct a 95% confidence interval of the proportion of cell phone users who develop cancer of the brain or nervous system.

(a) (0.0175, 0.0225)  
(b) (0.0267, 0.0375)  
(c) (0.190, 0.214)  
(d) (195, 209)  
(e) (0.75, 0.85)

16. The following table gives the probability distribution of the number of girls in the families of four children.

<table>
<thead>
<tr>
<th>Number of Girls $x$</th>
<th>$P(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.506</td>
</tr>
<tr>
<td>1</td>
<td>0.369</td>
</tr>
<tr>
<td>2</td>
<td>0.102</td>
</tr>
<tr>
<td>3</td>
<td>0.015</td>
</tr>
<tr>
<td>4</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Find the mean and the standard deviation of number of the girls.

(a) $\mu = 0.23, \sigma = 1.23$  
(b) $\mu = 0.65, \sigma = 0.786$  
(c) $\mu = 2.00, \sigma = 1.937$  
(d) $\mu = 2.35, \sigma = 0.521$  
(e) $\mu = 3.45, \sigma = 0.281$. 
The following is used for questions 17, 18 and 19. 

The following data gives the brain volumes for a group of patients with obsessive-compulsive disorders and a control group of healthy persons. Assume that the distributions of the brain volumes for these two groups of people are approximately normal with equal standard deviation.

<table>
<thead>
<tr>
<th>Control group</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 = 10 )</td>
<td>( n_1 = 10 )</td>
</tr>
<tr>
<td>( \bar{x}_1 = 0.45 )</td>
<td>( \bar{x}_2 = 0.34 )</td>
</tr>
<tr>
<td>( s_1 = 0.08 )</td>
<td>( s_2 = 0.07 )</td>
</tr>
</tbody>
</table>

At 0.01 significant level, test the claim that the means of the brain volumes for these two groups are different.

17. Choose the correct hypotheses to test the claim.
   (a) \( H_0 : \mu_1 \geq \mu_2 \) versus \( H_a : \mu_1 < \mu_2 \)
   (b) \( H_0 : \bar{x}_1 \geq \bar{x}_2 \) versus \( H_a : \bar{x}_1 < \bar{x}_2 \)
   (c) \( H_0 : \mu_1 \leq \mu_2 \) versus \( H_a : \mu_1 > \mu_2 \)
   (d) \( H_0 : \mu_1 = \mu_2 \) versus \( H_a : \mu_1 \neq \mu_2 \)
   (e) \( H_0 : \mu_1 < \mu_2 \) versus \( H_a : \mu_1 \geq \mu_2 \)

18. Find the value of the standardized test statistic.
   (a) \( 0.62 \)
   (b) \( 1.85 \)
   (c) \( 2.27 \)
   (d) \( 3.27 \)
   (e) \( 3.53 \)

19. Find the critical value for the test and state your conclusion at \( \alpha = .01 \).
   (a) The critical value is \( 2.861 \). The decision: Accept \( H_0 \)
   (b) The critical value is \( 2.552 \). The decision: Reject \( H_0 \)
   (c) The critical value is \( 2.878 \). The decision: Reject \( H_0 \)
   (d) The critical value is \( 2.539 \). The decision: Accept \( H_0 \)
   (e) The critical value is \( 4.604 \). The decision: Reject \( H_0 \)
The following is used for questions 20 and 21.
Listed below are the self-reported heights and measured heights for males aged 12-16.

<table>
<thead>
<tr>
<th>Reported heights</th>
<th>68</th>
<th>71</th>
<th>63</th>
<th>70</th>
<th>71</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured heights</td>
<td>67.9</td>
<td>69.9</td>
<td>64.9</td>
<td>68.3</td>
<td>70.3</td>
</tr>
</tbody>
</table>

The difference of the heights (self − measured) for this sample results in $\bar{d} = 0.34$ and $s_d = 1.381$. Assume that the heights are approximately normally distributed.

20. Is the mean of self-reported heights greater than that of the measured heights? Choose the appropriate hypotheses to test the claim.
   (a) $H_0 : \mu_d = 0$ versus $H_a : \mu_d \neq 0$
   (b) $H_0 : \bar{d} \leq 0$ versus $H_a : \bar{d} > 0$
   (c) $H_0 : \mu_d \leq 0$ versus $H_a : \mu_d > 0$
   (d) $H_0 : \mu_d \geq 0$ versus $H_a : \mu_d < 0$.
   (e) $H_0 : \mu_d < 0$ versus $H_a : \mu_d \geq 0$.

21. Find the rejection region and state your decision at $\alpha = .05$.
   (a) Rejection Region: $t < 2.353$; Decision: Reject $H_0$
   (b) Rejection Region: $t < -3.182$; Decision: Reject $H_0$
   (c) Rejection Region: $t > 2.132$; Decision: Fail to reject $H_0$
   (d) Rejection Region: $t > 3.182$; Decision: Fail to reject $H_0$
   (e) Rejection Region: $t < -2.132$; Decision: Fail to reject $H_0$
The following is used for questions 22, 23 and 24.
Among 734 randomly selected Internet users, it was found that 360 of them use the Internet for making travel plans. Use a 0.01 significance level to test the claim that among all Internet users, less than 50% use it for making travel plans.

22. Set up the null and alternative hypotheses to test the claim.
   (a) $H_0 : p \leq 0.50$ versus $H_a : p > 0.50$
   (b) $H_0 : p = 0.76$ versus $H_a : p \neq 0.76$
   (c) $H_0 : p \geq 0.50$ versus $H_a : p < 0.50$
   (d) $H_0 : p \geq 367$ versus $H_a : p < 367$
   (e) $H_0 : \bar{x} \leq 0.49$ versus $H_a : \bar{x} > 0.49$

23. Find the P-value for the above mentioned test.
   (a) 0.0735
   (b) 0.3015
   (c) 0.4207
   (d) 0.4840
   (e) 0.7054

24. At the significance level of 0.01, your conclusion is.
   (a) Accept $H_0$ since P-value is less than 0.01
   (b) Fail to accept $H_0$ correct since P-value is greater than 0.01
   (c) Accept $H_0$ since P-value is greater than 0.01
   (d) Fail to accept $H_0$ since P-value is less than 0.01
   (e) None of above
The following is used for questions 25.
The following table gives the car weight, x (in lb) and the fuel consumption, y (in miles/gal) for 7 randomly selected cars.

<table>
<thead>
<tr>
<th>Weight</th>
<th>MPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>3175</td>
<td>27</td>
</tr>
<tr>
<td>3450</td>
<td>29</td>
</tr>
<tr>
<td>3225</td>
<td>27</td>
</tr>
<tr>
<td>3985</td>
<td>24</td>
</tr>
<tr>
<td>2440</td>
<td>37</td>
</tr>
<tr>
<td>2500</td>
<td>34</td>
</tr>
<tr>
<td>2290</td>
<td>37</td>
</tr>
</tbody>
</table>

The regression equation is $\hat{y} = -0.00797x + 54.7$.

25. Predict the fuel consumption in MPG if the car weight is $x = 3000$ lb.
   (a) The MPG is 25.16
   (b) The MPG is 28.51
   (c) The MPG is 30.79
   (d) The MPG is 35.22
   (e) The MPG cannot be predicted

End of Multiple Choice Section
1. 10 people were randomly selected, and their systolic (x) and diastolic (y) blood pressures were measured.

<table>
<thead>
<tr>
<th></th>
<th>Systolic</th>
<th>138</th>
<th>130</th>
<th>135</th>
<th>140</th>
<th>120</th>
<th>125</th>
<th>120</th>
<th>130</th>
<th>130</th>
<th>144</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diastolic</td>
<td>82</td>
<td>91</td>
<td>100</td>
<td>100</td>
<td>80</td>
<td>90</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>98</td>
</tr>
</tbody>
</table>

From the data, we have:

\[ \sum x = 1312, \quad \sum x^2 = 172730, \quad \sum y = 881, \quad \sum y^2 = 78309, \quad \sum xy = 116008. \]

(a). (6 pts.) Calculate the correlation coefficient, \( r \), between the systolic and diastolic blood pressures.

(b). (5 pts.) Find the equation of the regression line.

(c). (5 pts.) Use the equation in part (c) to predict \( y \) when \( x = 128 \).
2. The following is the list of the ages of 12 Oscar-winning Actresses.

```
22  24  27  39  33  29  29  33  42  74  27  28
```

(a) (4 pts.) Find the five-number summaries, i.e., (Minimum, First Quartile, Second Quartile, Third Quartile, Maximum).

(b) (4 pts.) Find the inter-quartile range.

(c) (4 pts.) Draw a stem-and-leaf plot.
3. A scientist claimed that the average weight of the bumblebee bats is 1.8g. 15 bumblebee bats were randomly selected, and their weights are measured:

\[ 1.7 \quad 1.6 \quad 1.5 \quad 2.0 \quad 2.3 \quad 1.6 \quad 1.6 \quad 1.8 \quad 1.5 \quad 1.7 \quad 2.2 \quad 1.4 \quad 1.6 \quad 1.6 \quad 1.6 \]

Assume that the weights of the bumblebees are approximately normally distributed. Test if the scientist’s claim is correct or not.

(a) (5 pts.) In the context of the problem, explain Type I error.

(b) (6 pts.) State the null and the alternative hypotheses.

\[ H_0 : \]
\[ H_a : \]

(c) (5 pts.) Find the value of the standardized test statistic.

(d) (6 pts.) Find the rejection region for the test and draw your decision at the significance level of \( \alpha = 0.05 \).
The Keys to the multiple choice problems:
1C 2D 3B 4A 5B 6D 7A 8C 9A 10E 11C 12A 13D 14D 15B 16B 17D 18D 19C 20C 21C
22C 23B 24C 25C