Practice Test-Statistics

Section 2-Free Response
Part A: Questions 1-5

75% of grade for Section 2

Directions: Show all of your work.   You will be graded on the correctness of the methods you use as well as the accuracy of your answers and your explanation. Spend about 65 minutes on Part A.

1. A random sample of size 30 from a normal distribution has sample first and third quartiles of 6.94 and 11.45. The sample mean is 9.29.

(a) Estimate the true standard deviation of the distribution.

(b) Suppose that this sampling experiment is repeated. Estimate the probability that the sample mean of the new sample exceeds 12.

2. Ten graduating students were given a comprehensive test over four core subjects. Their grades are supposed to reflect their aggregate performance. Their scores along with their high school grade point averages are shown in the table below.

| Subject	1	2	3	4	5	6	7	8	9	10
Test	77	68	92	84	68	67	82	63	88	72
GPA	2.92	2.21	2.56	3.09	2.92	2.18	2.92	3.30	3.49	2.21

(a) Find the sample correlation between test scores and grade point averages.

(b) An administrator did a student-t test for the null hypothesis that the true population correlation is 0. The p-value for the test statistic was 0.45. Can we conclude that the true correlation is different from 0?

(c) Another administrator remarked that the conclusion in part (b) could not be applied to students in a neighboring state because, although they use the same test, it is graded on a 200 point scale rather than a 100 point scale. Is this remark valid? Why, or why not?

3. Genetic theory predicts that if a certain population is in equilibrium, genotypes aa, Aa, and AA occur with frequencies 0.16, 0.48, and 0.36, respectively. A sample of 100 members of the population was obtained and classified as follows.

aa	Aa	AA
10	67	23

(a) If the population is in equilibrium, what are the expected numbers of counts in the three genotype categories?

(b) At a significance level of 5% can we reject the hypothesis that the population is in equilibrium? Show your work.

4. The figure below is a scatterplot of observed values of two variables and the fitted least-squares line. Below that are the details of the linear regression from Excel.
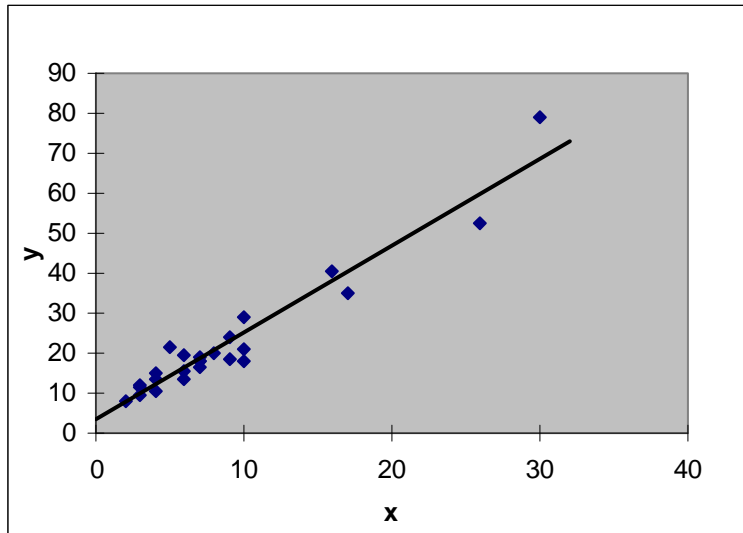
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.96461459
R Square	0.93048131
Adjusted R Square	0.92745876
Standard Error	4.1813967
Observations	25

ANOVA

	df	SS	MS	F	Significance F
Regression	1	5382.408797	5382.409	307.8463	8.22019E-15
Residual	23	402.133803	17.48408		
Total	24	5784.5426			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	3.3207799	1.371073673	2.422029	0.023721	0.484497925	6.157061869
x	2.17616668	0.124029556	17.54555	8.22E-15	1.919591995	2.432741362

(A) What is the predicted value of y when x = 25?

(B) What is the estimated error standard deviation σ?

(C) Find a 95% confidence interval for the slope.

(D) At a significance level of 5%, can we conclude that the true slope is not equal to 2?

5. The figure below is a stemplot of 40 test grades in a course. The stems are the tens digits and the leaves are the units digits.

The decimal point is 1 digit(s) to the right of the |

```
5 | 34
5 | 6688
6 | 112334
6 | 567899
7 | 011112333
7 | 5778889
8 | 13
8 | 577
9 | 0
```

 (a) Write down the five number summary of this data.

(b) In the empty box below, sketch a boxplot (box and whisker diagram). Use the convention that outliers are observations more than 1.5 times the interquartile range from the nearest quartile. Whiskers extend to either the most extreme observation in either direction or to a distance of 1.5 times the interquartile range, whichever is shorter.

60   65   70   75   80   85   90
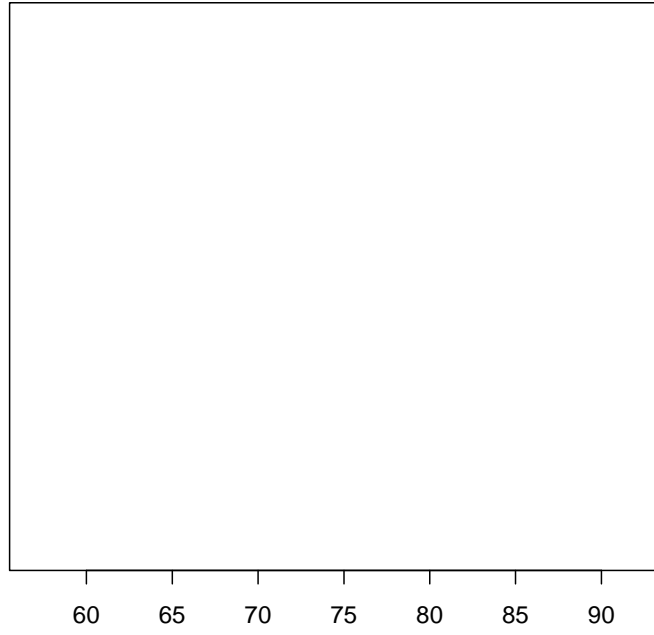
Part B – Question 6

25% of grade on Section 2

Directions: Part B consists of a single question. Answer as fully as you can. Spend about 25 minutes on this question.

6. Twenty adult subjects are available for a clinical study of the effectiveness of a medication for chronic sleeplessness.   The experiment will be conducted over a period of three weeks under controlled living conditions.

Researcher A proposes to divide the subjects into two groups of size 10, a control group and a treatment group. The control group would be given a placebo and the treatment group would be given the medication. The total amount of nighttime sleep would be recorded for each subject.

Researcher B proposes to administer both the placebo and the medication to each subject, separated by a one-week resting period, and to compare the sleep totals for the week under each treatment regimen. Both A's and B's plans would be conducted as double-blind experiments.

(a) Assuming that A's plan is followed, describe how the control and treatment groups should be selected, how the effectiveness of the medication would be defined in terms of population parameters, and describe in detail a statistical procedure that could be used to assess the medication's effectiveness.

(b) Assuming that B's plan is followed, describe how the effectiveness of the medication would be defined in terms of population parameters and describe in detail a statistical procedure for assessing the effectiveness. Has B's plan adequately compensated for the possibility of a "carry over" effect, where the order of the treatments affects the outcome? What else might be done to compensate?

(c) Discuss the relative merits of the two plans.