DEPARTMENT OF STATISTICS AND DEMOGRAPHY

MAIN EXAMINATION, 2017/18

COURSE TITLE: INTRODUCTION TO STATISTICS

COURSE CODE: STA 141

TIME ALLOWED: TWO (2) HOURS

INSTRUCTION: ANSWER <u>ALL QUESTIONS IN SECTION A</u> AND <u>ANY TWO</u>
                        QUESTIONS IN <u>SECTION B</u>

SPECIAL REQUIREMENTS: SCIENTIFIC CALCULATORS, GRAPH PAPER AND
                        STATISTICAL TABLES

DO NOT OPEN THIS PAGE UNTIL PERMISSION HAS BEEN GRANTED BY THE
INVIGILATOR

SECTION A

Question 1

The number of raisins in each of 14 mini boxes (1/2-ounce size) was counted for a generic brand and for Sunmaid brand raisins. The two data sets are shown here:

Generic Brand				Sunmaid			
25	26	25	28	25	29	24	24
26	28	28	27	28	24	28	22
26	27	24	25	25	28	39	27
26	26			28	24		

    a.   What are the mean and standard deviation for the generic brand?
    b.   What are the mean and standard deviation for the Sunmaid brand?
    c.   Compare the centres and variabilities of the two brands using the results of parts a and b.

(5+5+2 marks)

Question 2

In a state where cars have to be tested for the emission of pollutants, 25% of all cars emit excessive amounts of pollutants. When tested, 99% of all cars that emit excessive amounts of pollutants will fail, but 17% of the cars that do not emit excessive amounts of pollutants will also fail. What is the probability that a car that fails the test actually emits excessive amounts of pollutants? **(10 marks)**

Question 3

Determine the following probabilities using the standard normal distribution:
    (i) $P(z < -2.14)$
    (ii) $P(0.67 < z < 2.49)$
    (iii) $P(-2.07 < z < -0.93)$
    (iv) $P(z < 1.78)$
    (v) $P(-1.2 < z < 2.5)$ **(15 marks)**

Question 4

    (a)   An experiment can result in one or both of events A and B with the probabilities shown in this probability table:

	A	A^c
B	.34	.46
B^c	.15	.05

Find the following probabilities:
    (i) $P(A)$ (ii) $P(B)$ (iii) $P(A \cap B)$
    (iv) $P(A \cup B)$ (v) $P(A|B)$ (vi) $P(B|A)$ **(2+2+2+3+3+3 marks)**

    (b)   Refer to question (a)
        (i)   Are events A and B mutually exclusive? Explain.
        (ii) Are events A and B independent? Explain.   **(4+4 marks)**

SECTION B

Question 5

To estimate the amount of lumber in a tract of timber, an owner decided to count the number of trees with diameters exceeding 12 inches in randomly selected 50-by-50-foot squares. Seventy 50-by-50-foot squares were chosen, and the selected trees were counted in each track. The data are listed here.

7	8	7	10	4	8	6	8	9	10
9	6	4	9	10	9	8	8	7	9
3	9	5	9	9	8	7	5	8	8
10	2	7	4	8	5	10	7	7	7
9	6	8	8	8	7	8	9	6	8
6	11	9	11	7	7	11	7	9	13
10	8	8	5	9	9	8	5	9	8

    a. Construct a relative frequency histogram to describe the data.
    b. Calculate the sample mean $\bar{x}$ as an estimate of μ, the mean number of timber trees for all 50-by-50-foot squares in the track.
    c. Calculate s for the data.
    d. Construct the intervals $\bar{x} \pm s$, $\bar{x} \pm 2s$, and $\bar{x} \pm 3s$. Calculate the percentage of squares falling into each of the three intervals, and compare with the corresponding percentages given by the Empirical Rule.

(8+3+4+5 marks)

Question 6

    (a)   Over a long period of time it has been observed that a given marksman can hit a target on a single trial with probability equal to .8. Suppose he fires four shots at the target.

        (i)   What is the probability that he will hit the target exactly two times?
        (ii)   What is the probability that he will hit the target at least once?   **(2+3 marks)**

    (b)   Use the cumulative binominal table for $n = 5$ and $p = .6$ to find the probabilities of these events:
        (i) Exactly three successes
        (ii)Three or more successes   **(5 marks)**

    (c)   In southern California, a growing number of persons pursuing a teaching credential are choosing paid internship over traditional student teaching programs. A group of eight candidates for three local teaching positions consisted of five candidates who had enrolled in paid internships and three candidates who had enrolled in traditional student teaching programs. Let us assume that all eight candidates are equally qualified for the positions. Let x represent the number of internship-trained candidates who are hired for these three positions.

(i) Does x have a binomial distribution or a hyper geometric distribution? Support your answer.
(ii) Find the probability that three internship-trained candidates are hired for these positions.
(iii) What is the probability that none of the three hired was internship-trained?
(iv) Find $P(x \leq 1)$.   **(2+3+2+3 marks)**

Question 7

(a)   Let x be a normally distributed random variable with a mean of 10 and a standard deviation of 2. Find the probability that x lies between 11 and 13.6.   **(3 marks)**

(b)   Studies show that gasoline use for compact cars sold in the United States is normally distributed, with a mean of 25.5 miles per gallon (mpg) and a standard deviation of 4.5 mpg. What percentage of compacts gets 30 mpg or more?   **(3 marks)**

(c)   The scores on a national achievement test were approximately normally distributed, with a mean of 540 and a standard deviation of 110.

(i) If you achieve a score of 680, how far, in standard deviations, did your score depart from the mean?
(ii) What percentage of those who took the examination scored higher than you?   **(2+3 marks)**

(d) Given a normal distribution with $\mu=40$ and $\sigma=6$, find the value of 'x' that has (i) 45% of the area to the left and (ii) 14% of the area to the right.   **(5 marks)**

(e) The probability that a patient recovers from a rare disease is 0.4. If 100 people are known to have contracted this disease, what is the probability that less than 30 survive? Use the Normal approximation to the Binomial to answer this question.   **(4 marks)**

Question 8

A random sample of eight drivers insured with a company and having similar auto insurance policies was selected. The following table lists their driving experiences (in years) and monthly auto insurance premiums:

Driving Experience (years)	Monthly Auto Insurance Premium (E)
5	6, 400
2	8, 700
12	5, 000
9	7, 100
15	4, 400
6	5, 600
25	4, 200
16	6, 000

(a) Which is a dependent variable and which is the independent variable?
(b) Find the least squares regression line by choosing appropriate dependent and independent variables based on your answer in part (a).
(c) Interpret the meaning of the values of 'a' and 'b' calculated in (b).
(d) Plot the scatter diagram and the regression line.   **(2+10+5+3 marks)**

END OF EXAM!!