UNIVERSITY OF GHANA

BACHELOR OF SCIENCE IN ENGINEERING

SECOND SEMESTER EXAMINATIONS: 2015/2016

DEPARTMENT OF COMPUTER ENGINEERING

FAEN 302: STATISTICS FOR ENGINEERS (3 CREDITS)

INSTRUCTION: ***ANSWER ALL QUESTIONS IN SECTION A AND TWO QUESTIONS FROM SECTION B.***

Attached to the questions are all needed Statistical Tables.

TIME ALLOWED: ***THREE (3) HOURS***

Section A (50 marks)
Answer all questions

1. Identify an appropriate statistical test and the tail of the test for the following problems. Assume all samples used for study come from a normal population.

   (a) A packaging engineer of a brewery wishes to determine whether the average monthly production volume differ from 5000hl. **[2 marks]**

   (b) A civil engineer is tasked to determine whether the mean compressive strength of Ghacem cement is greater than the mean compressive strength of Diamond cement. **[2 marks]**

   (c) Management of a manufacturing company wishes to know whether the preference for a pension plan is independent of job classification. **[2 marks]**

   (d) An engineer wishes to determine whether the average diameter of bolts produced by a machine prior to overhaul is larger than it was before overhaul. **[2 marks]**

   (e) A lecturer wishes to determine whether the average score of students offering Statistics for Engineers differ with respect to the group (G1, G2, G3) they belong to. **[2 marks]**

2. Let X be a binomial random variable following $\sim b(n,p)$, $0 < p < 1$ such that, $16P\{X=1\} = 4Var(X) = E(X)$. Find $E(X)$.

   [4 marks]

3. A machine in a factory must be repaired if it produces more than 10% defectives among the large lot of items that it produces in a day. A random sample of 100 items from the day's production contains 15 defectives, and the supervisor says that the machine must be repaired. Does the sample evidence support his decision? Use a test with 5% significance level. [4 marks]

4. (a) State the Central Limit Theorem. [3 marks]

   (b) A bottling machine can be regulated so that it discharges an average of μ ounces per bottle. It has been observed that the amount of fill dispensed by the machine is normally distributed with $\sigma = 1.0$ ounce. A sample of $n = 9$ filled bottles is randomly selected from the output of the machine on a given day (all bottled with the same machine setting), and the ounces of fill are measured for each. Find the probability that the sample mean will be within 0.3 ounce of the true mean μ for the chosen machine setting. [4 marks]

5. (a) If X_1, X_2, X_3 and X_4 are pairwise uncorrelated random variables, each having mean 0 and variance 1. Compute the correlations of;

      i. $X_2 + X_3$ and $X_3 - X_4$ [3 marks]

      ii. $X_1 + X_2$ and $X_3 + X_4$ [3 marks]

   (b) If Z is a standard normal random variable, $Z \sim N(0,1)$, find $Cov(Z, Z^2)$. [4 marks]

6. An oil prospector will drill a succession of holes in a given area to find a productive well. The probability that he is successful on a given trial is 0.2.

   (a) What is the probability that the third hole drilled is the first to yield a productive well? [3 marks]

   (b) If the prospector can afford to drill at most ten wells, what is the probability that he will fail to find a productive well? [3 marks]

7. The random variables X has the probability density function given by

$$f(x) = \begin{cases} 4x^2e^{-2x}, & 0 < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

   Obtain $E[X]$ and $V[X]$ by inspection. [5 marks]

8. The Rockwell hardness index for steel is determined by pressing a diamond point into the steel and measuring the depth of penetration. For 50 specimens of an alloy of steel, the Rockwell hardness index averaged 62 with standard deviation 8. The manufacturer claims that this alloy has an average hardness index of at least 64. Is there sufficient evidence to refute the manufacturer's claim at the 5% significance level? [4 marks]

Section B (50 marks)
Answer any two questions from this section.
All questions carry equal marks.

1. (a) State the Total Probability Rule (TPR). [2 marks]

   (b) Suppose that the number of people who enter a coffee shop on a given day is a Poisson random variable with parameter λ. Show that if each person who enters the coffee shop is a male with probability p and a female with probability $1-p$, then the number of males and females entering the shop are independent Poisson random variable with respective parameters λp and $\lambda(1-p)$. [6 marks]

   (c) Suppose that X and Y are jointly continuous with joint density function $f(x,y)$ and marginal densities $f(x)$ and $f(y)$ respectively. Show that the;

      i. expectation of Y, $E[Y] = E\{E[Y|X]\}$. [3 marks]

      ii. variance of Y, $V(Y) = E[V(Y|X)] + V[E(Y|X)]$. [5 marks]

      iii. A quality control plan for an assembly line involves sampling $n = 10$ finished items per day and counting Y, the number of defectives. If p denotes the probability of observing a defective, then Y has a binomial distribution, assuming that a large number of items are produced by the line. But p varies from day to day and is assumed to have a uniform distribution on the interval form 0 to 1/4. Find the expected value of Y, variance of Y and identify the distribution of Y. [5 marks]

   (d) The weekly amount of downtime Y (in hours) for an industrial machine has approximately a gamma distribution with $\alpha = 3$ and $\beta = 2$. The loss L (in cedis) to the industrial operation as a result of this downtime is given by $L = 30Y + 2Y^2$. Find the expected value and variance of L. [4 marks]

2. (a) Define the following:

      i. Statistical hypothesis. [2 marks]

      ii. P-value of a test. [2 marks]

   (b) Aptitude tests should produce scores with a large amount of variation so that an administrator can distinguish between persons with low aptitude and persons with high aptitude. The standard test used by a certain industry has been producing scores with a standard deviation of 10 points for 21 employees. A new test is given to 21 prospective employees and produces a sample standard deviation of 12 points.

i. Compute a point estimate for the ratio of variance $\frac{\sigma_1^2}{\sigma_2^2}$. [2 marks]

ii. Construct a 95% confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$. [3 marks]

iii. Are scores from the new test significantly more variable than scores from the standard test? Use $\alpha = 0.05$. [3 marks]

(c) The random variable (X, Y) has a joint probability density function given by

$$f(x,y) = \begin{cases} \dfrac{e^{-x/y}e^{-y}}{y}, & 0 < x < \infty\ ,\ 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Compute $E[X^2|Y = y]$ [6 marks]

(d) If X and Y are independent binomial random variable with identical parameters n and p, $0 < p < 1$.

i. Identify the distribution of $X|(X + Y)$. [5 marks]

ii. Calculate $E[X|X + Y = m]$ [2 marks]

3. (a) Shear strength measurements derived from unconned compression tests for two types of soils gave the results shown in Table 1.0 below (measurements in tons per square foot).

Table 1.0

Soil Type I	$n_1 = 30$	$\bar{x}_1 = 1.65$	$s_1 = 0.26$
Soil Type II	$n_2 = 35$	$\bar{x}_2 = 1.43$	$s_2 = 0.22$

Do the soils appear to differ with respect to average shear strength, at the 5% significance level? [5 marks]

(b) Operators of gasoline-fueled vehicles complain about the price of gasoline in gas stations. Given that, the national gas tax per gallon is constant (GHC 18.4 *as of January 13, 2005*), but regional and local taxes vary from GHC 7.5 to GHC 32.10 for $n = 18$ key metropolitan areas around the country. The total tax per gallon for gasoline at each of these 18 locations are collected. The sample descriptives are shown in Table 2.0 below.

Table 2.0: Descriptives

			Statistic	Std. Error
Total tax per gallon for gasoline	Mean		39.5561	1.68236
	95 % CI for Mean	Lower Bound	36.0066	
		Upper Bound	43.1056	
	5% Trimmed Mean		39.4312	
	Median		38.3000	
	Std. Deviation		7.13764	
	Skewness		0.30300	0.536
	Kurtosis		-0.45600	1.038

i. Discuss the descriptives of the gasoline price sample shown in Table 2.0 and indicate whether normality can be suspected in the sample distribution. **[3 marks]**

(c) The Tables below represent the normality test and one sample t-test on the sample data.

Table 3.0: Tests of Normality

	Kolmogorov Smirnov			Shapiro Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total tax / gallon	0.117	18	0.200	0.962	18	0.647

Table 4.0: One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean	df
Total tax / gallon	18	39.5561	7.13764	d	

Table 5.0: One-Sample Test

	Test Value= 45					
	t	df	Sig. (2-tailed)	Mean Difference	95% CI of the Diff. Lower	Upper
Total tax / gallon	c	17	0.005	-5.44389	a	b

i. From Table 3.0, what conclusion can you make about the sample distribution? **[3 marks]**

ii. Find a, b, c and d shown in Table 5.0 and Table 4.0 respectively. **[6 marks]**

iii. From Table 5.0, is there sufficient evidence to claim that the average per gallon gas tax is less than GHC 45? **[2 marks]**

(d) The pH of water samples from a specific lake is a random variable Y with probability density function given by

$$f(y) = \begin{cases} \frac{3}{8}(7-y)^2, & 5 \leq y \leq 7 \\ 0, & \text{otherwise.} \end{cases}$$

i. Find an interval shorter than $[5, 7]$ in which at least three-fourth $(3/4)$ of the pH measurements must lie. **[3 marks]**

ii. How often would you expect to see a pH measurement below 5.5? **[3 marks]**