KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICS AND STATISTICS
Term 133

STAT 319 Statistics for Engineers and Scientists

Final Exam Wednesday August 13, 2014

Please circle your instructor name

Mr. Malik Mr. Al- Sawi Mr. Saleh

Name: ______________________________________ ID #:_______________________Section #____

Important Note:

- Show all your work including formulas, intermediate steps and final answer
- In hypothesis testing problems, write the null and the alternative hypotheses, test statistic, decision rule, critical values, and your conclusion, unless otherwise is specified.
- You may assume α = 0.05 for testing and 95% for confidence interval estimation if not otherwise stated.

Question No	Full Marks	Marks Obtained
1	3	
2	5	
3	8	
4	9	
5	10	
6	30	
Total	65	

Q1: Ahmad will take two books with him during the summer vacation. Suppose that the probability that he will like book 1 is 0.6, the probability that he will like book 2 is 0.5, and the probability that he will like both books is 0.4. What is the conditional probability that he will like book 2 given that he did not like book 1 *(3pts.)*

Q2: A communication cable company has determined that the number of push – button terminal switches demanded daily has a normal distribution with mean 200 and standard deviation 50

a. On what percentage of days will the demand be less than 90 switches? *(2pts.)*

b. Based on cost considerations, the company has determined that its best strategy is to produce a sufficient number of switches so that it will fully supply demand on 94% of all days. How many terminal switches should the company produce per day? *(3pts.)*

Q3: A radioactive mass emits particles from time to time. The time between two emissions is random. Let T represent the time in seconds between two emissions. Assume the probability density function of T given by

$$f(t) = \begin{cases} c\, e^{-0.2t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

a. Find the value of c so that $f(t)$ is a probability density function. *(1pt.)*

b. Find the median time between emissions. *(2pts.)*

c. If a sample of size 36 selected from this distribution. What is the approximate probability distribution of the sample mean? *(3pts.)*

d. Find the probability that the sample mean is more than 10. *(2pts.)*

Q4: An article in the Transactions of the Institution of Chemical Engineers (1956, 34, 280-293) reported data from an experiment investigating the effect of several process variable on the vapor phase oxidation of naphthalene. A sample of percentage mole conversion of naphthalene to maleic anhydride follows:

2	2.8	2.8	2.8	3	3.1	3.3	3.6	3.8	4
4.2	4.3	4.7	4.8	4.8	5	5	5.1	5.2	

You may use the following information's

$$\sum x = 74.3, \ \sum x^2 = 307.77$$

a. Construct a 95% confidence interval to estimate the mean of the vapor phase oxidation of naphthalene. *(5pts.)*

b. Do you need any other assumptions? If yes, what? If no, why? *(1pt.)*

c. How many more items you need to estimate the population mean to be within ±0.2 with 99% confidence level? *(3pts.)*

Q5: Two different types of injection – molding machines are used to form plastic parts. A part is considered defective if it has excessive shrinkage or is discolored. Two random samples, each of size 300, are selected, and 21 defective parts are found in the sample from first machine while 9 defective parts are found in the sample from the second machine

a. Is it reasonable to conclude that both machines produce the same fraction of defective parts? *(8pts.)*

b. Find the p – value of the test. *(2pts.)*

Q6: Suppose the mean peak power load (eV) can be modeled as a function of the maximum temperature (Fahrenheit) for the day. Following is the data for Peak Power Load y and Maximum Temperature for the days:

Temperature (Fahrenheit)	95	82	90	81	99	100	93	95	93	87
peak power load (eV)	214	152	156	129	254	266	210	204	213	150

You may use the following information's:

$$n = 10, \ \sum x = 915, \ \sum y = 1948, \quad Sxx = 380.5, \quad Syy = 19263.6, \quad Sxy = 2556$$

a. Estimate the linear regression model between Peak Power Load and Maximum Temperature. *(3pts.)*

b. Interpret the coefficients of the estimate. *(2pts.)*

c. Calculate the value of the corresponding residual for the maximum temperature. *(3pts.)*

d. Find the mean square error. *(3pts.)*

e. Construct a 90% confidence interval for the regression coefficient. *(5pts.)*

f. Depend on your answer in part (e), do you think that there is no relation between Peak Power Load and Maximum Temperature? Explain. *(2pts.)*

g. Calculate and interpret the co-efficient of determination and discuss what its value tells you about the two variables. *(3pts.)*

h. Test, at 1% level of significance, the hypotheses that the higher the temperature, the higher is the peak power load. Use p-value to do this test. *(5pts.)*

i. Estimate the average Peak power load if the maximum temperature is 97 with a 95% confidence coefficient. *(4pts.)*

- $\bar{X} = \dfrac{\sum X}{n}$
- $S^2 = \dfrac{\sum (X - \bar{X})^2}{n-1} = \dfrac{\sum X^2 - \frac{1}{n}(\sum X^2)}{n-1}$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A') = 1 - P(A)$
- $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$
- $\dfrac{P(A \mid B_r)P(B_r)}{\sum_{i=1}^{k} P(A \mid B_i)P(B_i)}$
- $E(X) = \sum xf(x) or \int_{-\infty}^{\infty} xf(x)dx$
- $\sigma^2 = \sum [x-\mu]^2 f(x) or \int_{-\infty}^{\infty} [x-\mu]^2 f(x)dx$

  $\sigma^2 = E(x^2) - [E(x)]^2$

- $b(x;n,p) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,\ldots.n$

$\binom{n}{x} = \dfrac{n!}{x!(n-x)!},$

$\mu = np, q = 1-p, \sigma^2 = npq$

- $p(x;\lambda) = \Pr[X = x] = \dfrac{e^{-\lambda}\lambda^x}{x!}, X = 0,1,2\ldots$

  $E(X) = Var(X) = \lambda$

- $g(x;p) = pq^{x-1}, \mu = \dfrac{1}{p}, \sigma^2 = \dfrac{q}{p^2}$
- $h(x,n,k,N) = P(X = x) = \dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$

$\mu = E(X) = \dfrac{nK}{N}, \sigma^2 = \dfrac{nK}{N}(1-\dfrac{K}{N})(\dfrac{N-n}{N-1})$

- $\mathrm{f}(\boldsymbol{x}) = \boldsymbol{\lambda e^{-\lambda x}},\ \boldsymbol{x > 0}$

  $\boldsymbol{\mu = \frac{1}{\lambda},\ \sigma^2 = \frac{1}{\lambda^2}}$

- $\mu_{\bar{X}} = \mu,\ \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$
- $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2,$
- $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$
- $t = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}, df = n-1$
- $\bar{x} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$ or $\bar{x} \pm z_{\alpha/2} \dfrac{S}{\sqrt{n}}$

  or $\bar{x} \pm t_{\alpha/2} \dfrac{S}{\sqrt{n}}$

- $n \geq \left(\dfrac{Z_{\alpha/2}\sigma}{e}\right)^2$
- $\hat{P} \pm z_{\alpha/2} \sqrt{\dfrac{\hat{P}(1-\hat{P})}{n}}$
- $n \geq \dfrac{Z^2_{\alpha/2} \hat{P}(1-\hat{P})}{e^2}$
- $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$
- $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$
- $(\bar{x}_1 - \bar{x}_2) + \pm t_{\alpha/2} s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$, where

  $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2},$

  $df = n_1 + n_2 - 2$

- $(\hat{P}_1 - \hat{P}_2) \pm z_{\alpha/2} \sqrt{\dfrac{\hat{P}_1\hat{Q}_1}{n_1} + \dfrac{\hat{P}_2\hat{Q}_2}{n_2}},$

  where $\hat{P}_i = \dfrac{X_i}{n_i}, \hat{q}_i = 1 - \hat{P}_i$

- $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ **or** $Z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ **or**
- $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}, df = n-1$
- $z = \dfrac{X - np_0}{\sqrt{np_0q_0}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0q_0}{n}}}$

- $$z = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
- $$z = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
- $$t = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
- $z = \frac{\hat{P}_1 - \hat{P}_2 - 0}{\sqrt{\hat{p}\hat{q}(\frac{1}{n_1} + \frac{1}{n2})}}$, where

  $\hat{p} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$, and

  $\hat{q} = 1 - \hat{p}$
- $\hat{y} = a + bx$, where:

  $$b = \frac{\sum xy - \sum x \sum y / n}{\sum x^2 - (\sum x)^2 / n} = \frac{S_{xy}}{S_{xx}},$$

  and $a = \bar{y} - b\bar{x}$
- $$S_{xx} = \sum (x - x)^2 = \sum x^2 - (\sum x)^2 / n$$
- $$S_{yy} = \sum (y - y)^2 = \sum y^2 - (\sum y)^2 / n$$

  =SST
- $$S_{xy} = \sum xy - (\sum x \sum y) / n$$
- $$SSE = \sum (y_i - \hat{y}_i)^2 = S_{yy} - bS_{xy}$$
- $$S^2 = MSE = \frac{SSE}{n-2}, S = \sqrt{MSE}$$
- $s_{b_1} = \frac{S}{\sqrt{Sxx}}$
- $b_1 \pm t_{\frac{\alpha}{2}, n-2} s_{b_1}$
- $t = \frac{b_1 - \beta_0}{s_{b_1}}$
- $$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$
- $$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$
- $$R^2 = 1 - \frac{SSE}{SST}$$
- $$r = b\sqrt{\frac{S_{XX}}{S_{YY}}} = \frac{S_{XY}}{\sqrt{(S_{XX})(S_{YY})}}, r^2 = R^2$$
- $$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, df = n-2$$