Roll No: --------------------------

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

End Semester Examination, May 2018
Program: B.Tech. GSE and GIE **Semester – VI**
Subject (Course): Statistical Methods in Geosciences **Max. Marks : 100**
Course Code : GSEG 331 **Duration : 3 Hrs**
No. of page/s: 3

Useful tabular values (single tail):

$P(Z \geq 1.96) = 0.025,\ P(Z \geq 2.58) = 0.005, P(Z \geq 1.645) = 0.05$
$t_{0.025,8} = 2.306, t_{0.025,9} = 2.262, t_{0.025,10} = 2.228, t_{0.025,14} = 2.145, t_{0.025,15} = 2.131,$
$t_{0.025,16} = 2.120,\ t_{0.025,20} = 2.086,\ t_{0.025,21} = 2.080, t_{0.025,22} = 2.074\ .$
$t_{0.05,8} = 1.860, t_{0.05,9} = 1.833,\ t_{0.05,10} = 1.812, t_{0.05,14} = 1.761, t_{0.05,15} = 1.753,$
$t_{0.05,16} = 1.746,\ t_{0.05,20} = 1.725,\ t_{0.05,21} = 1.721,\ t_{0.05,22} = 1.717\ .$
$F_{9,11,0.05} = 2.90, F_{11,9,0.05} = 3.10, F_{8,10,0.05} = 3.07, F_{10,8,0.05} = 3.35.$

Attempt all questions from **Section A** (each carrying 4 marks); attempt all questions from **Section B** (each carrying 8 marks)**;** attempt all questions from **Section C** (each carrying 20 marks)**.**

Section A
(Attempt all questions)

1.	Lots of 40 components each are called unacceptable if they contain as many as 3 defectives or more. The procedure for sampling the lot is to select 5 components at random and to reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?	**[4]**	**CO1**
2.	The probability that a certain kind of component will survive a shock test is 0.75. Find the probability that exactly 2 of the next 4 components tested survive.	**[4]**	**CO1**
3.	Let us consider the least-squares line be $\hat{y} = \widehat{\beta_0} + \widehat{\beta_1}x$, then construct a 95% confidence interval for β_1. Given that $n = 10, S_{xx} = 263.6, S_{xy} = 534.2, \bar{y} = 4.6,$ $\bar{x} = 3.8.$	**[4]**	**CO2**
4.	Find $SE(\overline{x_1} - \overline{x_2})$ under $H_0: \mu_1 = \mu_2$ where $\overline{x_1}$, $\overline{x_2}$ and μ_1, μ_2 are means of two samples and two populations, respectively.	**[4]**	**CO2**

<table>
<tr><td>5</td><td>If $\gamma(h)$ is a variogram function and $C(h)$ is a covariance function for a second order stationary random field following intrinsic hypothesis, prove that,
$$\gamma(h) = C(0) - C(h)$$</td><td>[4]</td><td>CO4</td></tr>
<tr><td colspan="4">SECTION B
(Attempt all questions and Q10 has internal choice)</td></tr>
<tr><td>6.</td><td>In a large city A, 25% of a random sample of 900 school children had defective eye-sight. In other large city B, 15.5% of random sample of 1600 schoolchildren had the same effect. Is this difference between the two proportions significant? Obtain 95% confidence limits for the difference in the population proportions?</td><td>[8]</td><td>CO2</td></tr>
<tr><td>7.</td><td>The heights of six randomly chosen sailors are in inches: 63, 65, 68, 69, 71 and 72. Those of 10 randomly chosen soldiers are 61, 62, 65, 66, 69, 69, 70, 71, 72 and 73. Discuss the light that these data throw on the suggestion that sailors are on the average taller than soldiers.</td><td>[8]</td><td>CO2</td></tr>
<tr><td>8.</td><td>Use the method of least squares to fit a straight line to the accompanying data points. Give the estimates of β_0 and β_1 and hence find the coefficient of determination.
<table>
<tr><td>x</td><td>-1</td><td>0</td><td>2</td><td>-2</td><td>5</td><td>6</td><td>8</td><td>11</td><td>12</td><td>-3</td></tr>
<tr><td>y</td><td>-5</td><td>-4</td><td>2</td><td>-7</td><td>6</td><td>9</td><td>13</td><td>21</td><td>20</td><td>-9</td></tr>
</table></td><td>[8]</td><td>CO3</td></tr>
<tr><td>9</td><td>Compute the variogram, $\gamma(h)$, for $h = 9$ for the data given below on a straight line.
4, 3, 3, 5, 5, 5, 4, 4, 5, 4, 5,17, 8, 2, 2, 3, 7, 7, 1, 6, 10, 9, 9, 10, 11, 12, 11, 3, 3, 4.
Each data is separated by 3 feet.</td><td>[8]</td><td>CO4</td></tr>
<tr><td>10.</td><td>Mathematically, define the simple kriging error variance, and express it as a function of variance-covariance function.
OR
Mathematically, define the ordinary kriging error variance, and express it as a function of variogram function.</td><td>[8]</td><td>CO4</td></tr>
<tr><td colspan="4">SECTION C
(Attempt all questions and Q12A, Q12B have internal choice)</td></tr>
<tr><td>11.A</td><td>Find the moment generating function of the binomial distribution and hence find the mean.</td><td>[10]</td><td>CO1</td></tr>
</table>

11.B

Two random samples gave the following results:

Sample	Size	Sample Mean	Sum of squares of deviations from the mean
1	10	15	90
2	12	14	108

Test whether the samples come from the same normal population at 10% level of significance.

[10] **CO2**

12.A

Let $Y = \beta_0 + \beta_1 x + \varepsilon$ be a simple linear regression model with $\varepsilon \sim N(0, \sigma^2)$ and let the errors ε_i associated with different observations y_i $(i = 1,2,\ldots,N)$ be independent. Then show that,

i. $\widehat{\beta_0}$ and $\widehat{\beta_1}$ have normal distributions.

ii. The mean and variance are given by

$$E[\,\widehat{\beta_0}\,] = \beta_0, Var(\,\widehat{\beta_0}\,) = \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)\sigma^2 \text{ and } E[\,\widehat{\beta_1}\,] = \beta_1,\ Var(\,\widehat{\beta_1}\,) = \frac{\sigma^2}{S_{xx}}$$

where $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$. In particular, the least-squares estimators $\widehat{\beta_0}$ and $\widehat{\beta_1}$ are unbiased estimators of β_0 and β_1 respectively.

OR

Show that $\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 + \sum_{i=1}^{n} (\widehat{y_i} - \bar{y})^2$

[10] **CO3**

12.B

Consider the given data from three location as mentioned below to estimate the value, $Z(x_0)$ at $x_0 = (180, 120)$ using simple kriging. Given $E[Z(x)] = 110$ and the variance-covariance function $2000 * \exp(\frac{-h}{250})$.

	x	y	$Z(x_i)$
x_1	387	72	50
x_2	392	81	56
x_3	388	56	53

OR

Consider the given data from two location as mentioned below to estimate the value, $Z(x_0)$ at $x_0 = (180, 120)$ using ordinary kriging. Given the variance-covariance function $2000 * \exp(\frac{-h}{250})$.

	x	y	$Z(x_i)$
x_1	387	72	50
x_2	392	81	56

[10] **CO4**