Assignment 5

Due date : 6/10/2022

21102052 Lee Jeong-Yun

1. (Ex 12-16, p 488) The article "Characterization of Highway Runoff in Austin, Texas, Area" (J. of Envir. Engr., 1998: 131–137) gave a scatter plot, along with the least squares line, of x = rainfull volume (m^3) and y = runoff volume (m^3) for a particular location. The accompanying values were read from the plot.

x	5	12	14	17	23	30	40	47
y	4	10	13	15	15	25	27	46

x	55	67	72	81	96	112	127
y	38	46	53	70	82	99	100

a) Does a scatter plot of the data support the use of the simple linear regression model?

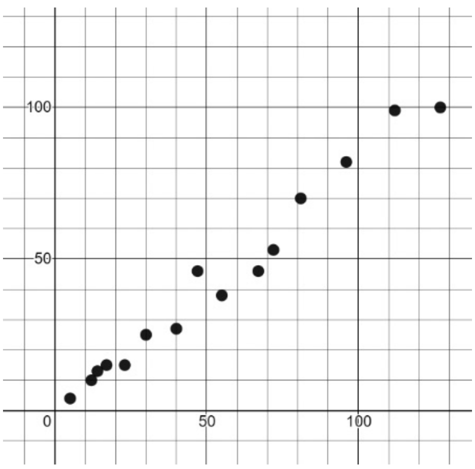
b) Calculate point estimates of the slope and intercept of the population regression line.

c) Calculate a point estimate of the true average runoff volume when rainfall volume is 50.

d) Calculate a point estimate of the standard deviation σ.

e) What proportion of the observed variation in runoff volume can be attributed to the simple linear regression relationship between runoff and rainfall?

a)



b) $n=15$ $\sum x_i = 798$ $\bar{x} = 53.2$ $\sum y_i = 643$ $\bar{y} = 42.87$

$\sum x_i y_i = 51232$ $\sum x_i^2 = 63040$ $\sum y_i^2 = 41999$

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} \qquad S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$= 51232 - \frac{798 \times 643}{15} = 17024.4$$

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 63040 - \frac{798^2}{15}$$

$$= 20586.4$$

$$\hat{\beta}_1 = \frac{17024.4}{20586.4} = 0.827$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 42.87 - 0.827 \times \frac{798}{15}$$

$$= -1.1264$$

$$y = -1.1264 + 0.827x$$

slope : 0.827 intercept : -1.1264

c) $y = -1.1264 + 0.827x$ $E(y) = -1.1264 + 0.827 \times 50$

$$= 40.2236$$

d) $SSE = S_{yy} - \hat{\beta}_1 S_{xy} = 41999 - \frac{643^2}{15} - 0.827 \times 17024.4$

$$= 357.01$$

$$\sigma = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{357.01}{15-2}} = 5.24$$

$$r^2 = 1 - \frac{SSE}{SST} = 1 - \frac{357.01}{41999 - \frac{643^2}{15}} = 1 - 0.025 = 0.975$$

2. (Slightly modified version of Ex 12-52, p 507) Plasma etching is essential to the fine-line pattern transfer in current semiconductor processes. The article "Ion Beam- Assisted Etching of Aluminum with Chlorine" (J. of the Electrochem. Soc., 1985: 2010– 2012) gives the accompanying data (read from a graph) on chlorine flow (x, in SCCM) through a nozzle used in the etching mechanism and etch rate (y, in 100 A/min).

x	1.5	1.5	2.0	2.5	2.5	3.0	3.5	3.5	4.0
y	23.0	24.5	25.0	30.0	33.5	40.0	40.5	47.0	49.0

a) Fit the simple linear regression model to this data.

b) Estimate the true average change in etch rate associated with a 1-SCCM increase in flow rate using a 95% confidence interval, and interpret the interval.

c) What proportion of observed variation in % removed can be attributed to the model relationship?

d) Does the simple linear regression model specify a useful relationship? Carry out an appropriate test of hypotheses using a significance level of 0.05. Yes

e) Calculate a 95% CI for $\mu_{Y\cdot 3.0}$, the true average etch rate when flow = 3.0. Has this average been precisely estimated?

f) Calculate a 95% PI for a single future observation on etch rate to be made when flow = 3.0. Is the prediction likely to be accurate?

g) Would the 95% CI and PI when flow = 2.5 be wider or narrower than the corresponding intervals of parts (c) and (d)? Answer without actually computing the intervals.

h) Would you recommend calculating a 95% PI for a flow of 6.0? Explain.

a) $s = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{45.3623}{9-2}} = 2.5456$

$S_{xx} = \sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2 = 70.5 - \frac{1}{9}\times 24^2 = 6.5$

$s_{\hat{\beta}_1} = \frac{s}{\sqrt{S_{xx}}} = \frac{2.5456}{\sqrt{6.5}} = 0.9985$

$y = \hat{\beta}_0 + \hat{\beta}_1 x = 6.4482 + 10.6026x$

$df = 7$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.05$

$t = \frac{10.6026}{0.9985} = 10.619$

$t_{0.025}(7) = 2.3646$

$t > 2.3646$. H_0 is rejected

b) $\hat{\beta}_1 \pm \underbrace{t_{\alpha/2,\, n-2}}_{=2.365} \cdot S_{\hat{\beta}_1}$ $\alpha=0.05$ $df=7$

$10.6026 \pm 2.365 \times 0.9985 = (8.2411,\ 12.9640)$

e) $(36.101,\ 40.412)$

f) $(31.863,\ \cancel{44}.650)$

g) $(30.91,\ 35)$

h) not recommend.

3. (Ex 10-6, p 401) The article "Origin of Precambrian Iron Formations" (Econ. Geology, 1964: 1025–1057) reports the following data on total Fe for four types of iron formation (1 = carbonate, 2 = silicate, 3 = magnetite, 4 = hematite).

1:	20.5	28.1	27.8	27.0	28.0
	25.2	25.3	27.1	20.5	31.3
2:	26.3	24.0	26.2	20.2	23.7
	34.0	17.1	26.8	23.7	24.9
3:	29.5	34.0	27.5	29.4	27.9
	26.2	29.9	29.5	30.0	35.6
4:	36.5	44.2	34.1	30.3	31.4
	33.1	34.1	32.9	36.3	25.5

Carry out an analysis of variance F test at significance level .01, and summarize the results in an ANOVA table.

	df	SS	MS	F
Tr	3	509.126	169.71	10.85
E	36	563.134	15.64	
T	39	1072.26		

p-value $< 0.01 = \alpha$, reject H_0

$SST = 33882.24 - 32809.98 = 1072.26$

$SSTr = \frac{1}{10}(333191 \cdot 1) - 32809.98 = 33319.11 - 32809.98 = 509.126$

$SSE = 1072.26 - 509.126 = 563.134$

$MSTr = 509.126/3 = 169.71$

$MSE = \frac{563.134}{36} = 15.64$

$F = \frac{MSTr}{MSE} = 10.85$

4. (A slightly changed form of Ex 10-18, p 408) Consider the accompanying data on plant growth after the application of five different types of growth hormone.

	$\bar{x}$	s^2
1: 12 17 7 14	51	12.75
2: 20 13 20 17	71	17.75
3: 18 15 20 17	70	17.5
4: 7 11 18 10	46	11.5
5: 6 11 15 8	40	10

a. Perform an F test at level $\alpha=0.05$.

b. What happens when Tukey's procedure is applied?

a) $4280 - \frac{1}{20}(278)^2 = 415.8 = SST$

$SSTr = \frac{1}{J}\sum x_{i.}^2 - \frac{1}{IJ}(x_{..})^2 = \frac{1}{4}(51^2+71^2+70^2+46^2+40^2) - \frac{1}{5\times4}278^2 = 4064.5 - 3864.2 = 200.3$

$SSE = 415.8 - 200.3 = 215.5$

$MSTr = \frac{200.3}{5-1} = 50.075$

$MSE = \frac{215.5}{5(4-1)} = 14.36$

$f = \frac{50.075}{14.36} = 3.487$

$\nu_1 = 4 \quad \nu_2 = 15$

$3.487 = f > f_{0.05,4,15} = 3.06$ H_0 is rejected.

b) $w = Q_{\alpha, I, I(J-1)} \sqrt{\frac{MSE}{J}} = 4.37 \times \sqrt{\frac{14.36}{4}} \approx 8.3$

$\bar{x}_5$	$\bar{x}_4$	$\bar{x}_1$	$\bar{x}_3$	$\bar{x}_2$
10	11.5	12.75	17.5	17.75

$17.75 - 10 = 7.75 < w = 8.3$

No same result. no differ

5. (Ex 14-30, p 620) Three different design configurations are being considered for a particular component. There are four possible failure modes for the component. An engineer obtained the follow- ing data on number of failures in each mode for each of the three configurations. Does the configuration appear to have an effect on type of failure?

		Failure Mode			
		1	2	3	4
Configuration	1	20	44	17	9
	2	4	17	7	12
	3	10	31	14	5

Expected

16.1053	43.5789	18	12.3158
7.1579	19.3684	8	5.47737
10.7368	29.0526	12	8.2105

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 0.9419 + 0.0041 + 0.0556 + 0.8927 + 1.3932 + 0.2896$$

$$+ 0.1250 + 7.7814 + 0.0506 + 0.1305 + 0.3333 + 1.2554$$

$$= 13.2532$$

$$\chi^2 = 13.2532 > \chi^2_{0.05,6} = 12.592$$

H_0 is rejected

6. (Ex 14-16, p 612) In a genetics experiment, investigators looked at 300 chromosomes of a particular type and counted the number of sister-chromatid exchanges on each ("On the Nature of Sister-Chromatid Exchanges in 5-Bromodeoxyuridine- Substituted Chromosomes," Genetics, 1979: 1251–1264). A Poisson model was hypothesized for the distribution of the number of exchanges. Test the fit of a Poisson distribution to the data by first estimating μ and then combining the counts for x = 8 and x = 9 into one cell.

x = Number of Exchanges	0	1	2	3	4	5	6	7	8	9	
frequency = f(λ) = **Observed Counts**	6	24	42	59	62	44	41	14	6	2	300
λf(λ)	0	24	84	177	248	220	246	98	48	18	1163

$$\lambda = \frac{\sum \lambda f(\lambda)}{f(\lambda)} = \frac{1163}{300} = 3.88 \qquad p(\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-3.88}(3.88)^{x}}{x!}$$

$\boxed{n = 300}$

λ	p(λ)	np(λ)	f(λ)
0	0.0207	6.21	6
1	0.0801	24.03	24
2	0.1554	46.62	42
3	0.201	60.3	59
4	0.195	58.5	62
5	0.1513	45.39	44
6	0.0979	29.37	41
7	0.0542	16.26	14
≥ 8	0.0283	13.3	8

$$\chi^2 = \sum\left(\frac{(f(\lambda) - np(\lambda))^2}{np(\lambda)}\right) = 7.77634$$

df = 8

$\chi^2 = 7.77634 < \chi^2_{0.05,\,8} = 15.5073$. Fit poisson.