620.370 Statistics for Mechanical Engineers — Semester 2, 2009

Not Homework set 12

Tutorial problems: Quiz 12; 212, 214, 219.

---

1. (a) The table below gives values of a variable x and a dependent variable y.

x	0	1	2	3	4	5
y	54.61	70.58	77.69	80.23	87.70	88.32
	65.44	68.12	72.16	75.75	86.67	
	54.83					

   i. Assuming that $\mathrm{E}(Y \mid x) = \alpha + \beta e^{-0.4x}$ and $\mathrm{var}(Y \mid x) = \sigma^2$, obtain estimates of α, β and σ^2 using the method of least squares.

   ii. Plot the observations and your fitted curve.

   iii. Find a prediction interval for an observation at x=5.

   (b) Consider the following bivariate data set:

x	31	46	41	52	36	47	53	36	27	58	39	50
y	82	60	87	63	73	50	55	76	80	42	70	49

   Assume that these data were obtained from a bivariate normal population with correlation ρ.

   i. Find the sample correlation coefficient, r.

   ii. Find a 95% confidence interval for ρ.

   iii. Test the hypothesis that X and Y are independent.

2. An investigation of a die-casting process resulted in the data below for x_1 = furnace temperature, x_2 = die close time, and y = temperature difference on the die surface ("A multiple-objective decision-making approach for assessing simultaneous improvement in die life and casting quality in a die casting process", *Quality Engineering*, 1994: 371–383).

x_1	1250	1300	1350	1250	1300	1250	1300	1350	1350
x_2	6	7	6	7	6	8	8	7	8
y	80	95	101	85	92	87	96	106	108

Output from fitting the model $\eta = \mathrm{E}(y \mid x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ is given below:

```
Predictor      Coef   SE Coef       T      P
Constant    -199.56     11.64  -17.14  0.000
x1           0.2100   0.00864   24.30  0.000
x2           3.0000    0.4321    6.94  0.000
S = 1.05848   R-Sq = 99.1%   R-Sq(adj) = 98.8%

Analysis of Variance
Source          DF      SS      MS       F      P
Regression       2  715.50  357.75  319.31  0.000
Residual Error   6    6.72    1.12
Total            8  722.22
```

   (a) The model utility test is a test of β_1=β_2=0, i.e. whether the regression is significantly different from zero. Use the regression MS to test this hypothesis.

   (b) Calculate and interpret the 95% confidence interval for β_2.

   (c) When x_1=1300 and x_2=7, the standard error of $\hat{\eta}$ is 0.353. Calculate a 95% confidence interval for η when furnace temperature is 1300 and die close time is 7.

   (d) Calculate a 95% prediction interval for the temperature difference resulting from an experimental run with a furnace temperature of 1300 and a die close time of 7.

Quiz 12

Q12.1 If the correlation coefficient r_{XY} is not significantly different from zero, then this indicates that

A. X and Y are independent

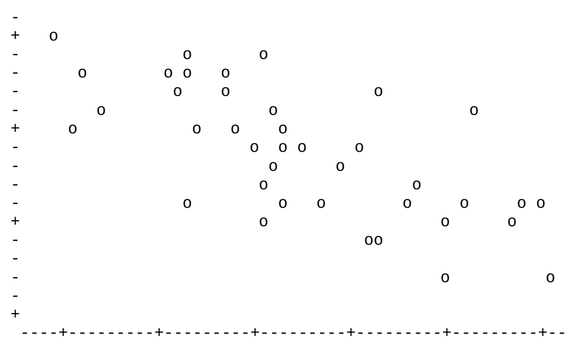
B. X and Y are not functionally related

C. X and Y are identically distributed

D. X and Y are uncorrelated

E. X and Y have equal means

Q12.2 For the data represented in the scatter diagram below, which one of the following statements about the correlation coefficient is true?



A. $0.6 < r < 1.0$

B. $0.2 < r < 0.6$

C. $-0.2 < r < 0.2$

D. $-0.6 < r < -0.2$

E. $-1.0 < r < -0.6$

Q12.3–12.4 refer to the information below:

The following Minitab output was obtained when a multiple regression model of the form $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + e_i$,
was fitted to a sample of 20 observations.

```
The regression equation is
y = 16.6 + 0.76 x1 - 0.694 x2 + 5.64 x3

Predictor      Coef   SE Coef       T       P
Constant     16.578     4.292    3.86   0.001
x1            0.764     1.480    0.52   0.613
x2          -0.6941    0.4697   -1.48   0.159
x3            5.640     2.540    2.22   0.041

Analysis of Variance
SOURCE        DF        SS        MS       F      P
Regression     3   24393.2    8131.1   22.14  0.000
Error         16    5874.9     367.2
Total         19   30268.0
```

Q12.3 The value of R^2 is:

A. 0.649

B. 0.806

C. 0.194

D. 0.898

E. -0.898

Q12.4 A 95% confidence interval for β_2 is:

A. $(-1.690, 0.302)$

B. $(-1.695, 0.307)$

C. $(-1.514, 0.126)$

D. $(-1.164, -0.224)$

E. $(-0.799, -0.589)$

Answers to Not Homework 12

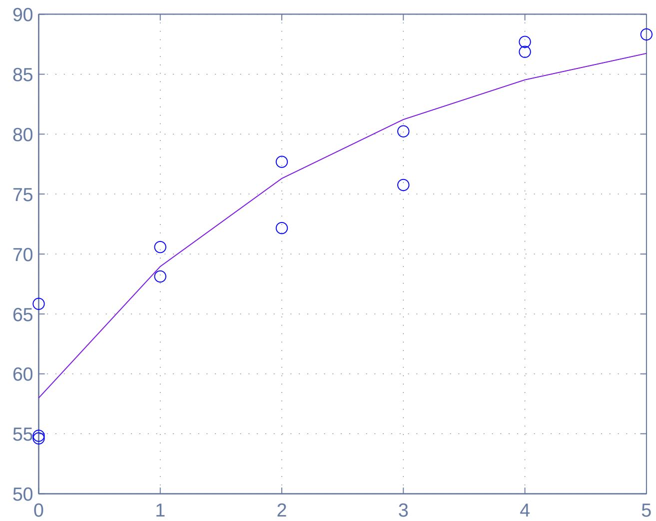
1. (a) We put $z = e^{-0.4x}$ and fit a linear regression of y on z, since $\mathrm{E}(Y \mid z) = \alpha + \beta z$.

   i. Using MATLAB:

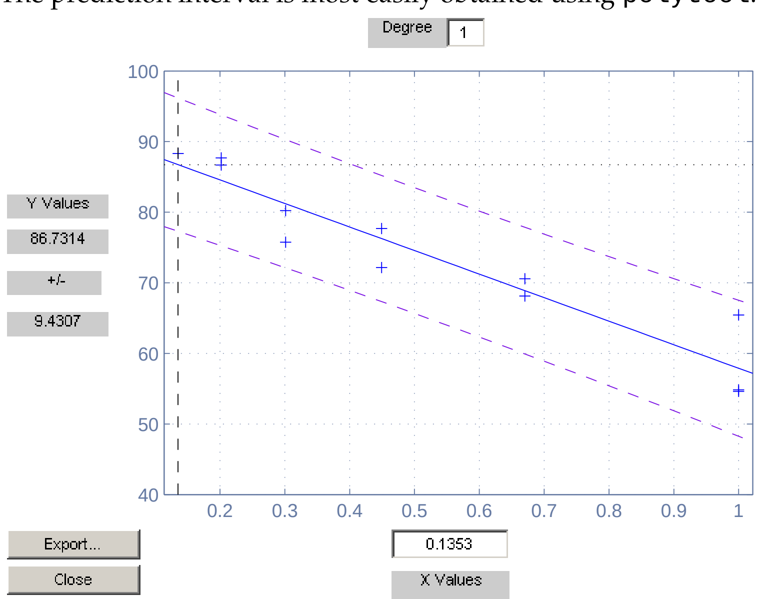
```
>> z=exp(-0.4*x)
>> [b,bi,r,ri,stat]=regress(y,z);
>> b
b =  91.2443  -33.3549
>> stat
stat =  0.8999   89.8770    0.0000   14.7408
```

   Thus $\hat{\alpha} = 91.24$, $\hat{\beta} = -33.35$ and $s^2 = 14.74$

   ii.



   iii. The prediction interval is most easily obtained using `polytool`:



   Thus, prediction interval = $86.73 \pm 9.43 = (77.30, 96.16)$

For interest, the computation details are as follows:

$$n = 12 \qquad \Sigma(z-\bar{z})(y-\bar{y}) = -39.7199$$
$$\Sigma x = 21 \qquad \Sigma(z-\bar{z})^2 = 1.190827$$
$$\Sigma y = 3348 \qquad \Sigma(y-\bar{y})^2 = 1472.2630$$

Therefore $\hat{\beta} = \frac{-39.7199}{1.1908} = -33.3549$, $\hat{\alpha} = 73.51 - (-33.3549)\times 0.5317 = 91.2443$; $s^2 = \frac{1}{12-2}(1472.2630 - \frac{39.7199^2}{1.1908}) = 14.7408$.

$\hat{\eta} = 91.2443 + 0.1353\times(-33.3549) = 86.73$;

95% PI: $86.73 \pm 2.228\times\sqrt{14.7408(1 + \frac{1}{12} + \frac{(0.1353-0.5317)^2}{1.190827})} = (77.30, 96.16)$

The MINITAB output is given below:

```
The regression equation is y = 91.2 - 33.4 z

Predictor     est      se      t     P
Constant    91.244    2.174  41.96  0.000
z          -33.355    3.518  -9.48  0.000
S = 3.83936   R-Sq = 90.0%   R-Sq(adj) = 89.0%

Analysis of Variance
Source          df      SS      MS      F      P
Regression       1  1324.9  1324.9  89.88  0.000
Residual Error  10   147.4    14.7
Total           11  1472.3

Predicted Values for New Observations
x        fit     se.fit      95% CI          95% PI
0.1353   86.73    1.78   (82.76, 90.70)  (77.30, 96.16)
```

(b) MATLAB gives:

```
>> [r,p,rl,ru]=corrcoef(x,y)
r = -0.8443
p = 0.0006
rl = -0.9553
ru = -0.5245
```

Thus $r = -0.8443$; 95% CI for ρ: $(-0.955, 0.525)$; and since the CI does not include zero, we reject the hypothesis of independence.

$n = 12$, $\bar{x} = 43$, $\bar{y} = 65.5833$;
$\Sigma(x-\bar{x})^2 = 998$, $\Sigma(x-\bar{x})(y-\bar{y}) = -1291$, $\Sigma(y-\bar{y})^2 = 2342.92$.
$r = \frac{-1291}{\sqrt{998\times 2342.92}} = -0.8443$;
$r = -0.8443 \overset{\text{atanh}}{\longrightarrow} z = -1.2359 \pm \frac{1.96}{\sqrt{9}} = (-1.8892, -0.5825) \overset{\text{tanh}}{\longrightarrow} (-0.9553, -0.5245)$;
or use the diagram in the Tables: $(-0.96, -0.52)$.
Since the CI does not include zero, we reject the hypothesis of independence;
or use, $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = -4.982 < c_{0.025}(t_{10}) = -2.228$, so reject ρ=0.

2. (a) $\frac{\text{regSS}}{\text{resMS}} = 319.3 \gg c_{0.95}(F_{2,6})$, so we reject $\beta_1 = \beta_2 = 0$, indicating that the model is useful in predicting y.

(b) 95% CI for β_2: $3.00 \pm 2.447\times 0.4321 = (1.94, 4.06)$.

(c) $\hat{\eta} = -199.56 + 0.21\times 1300 + 3.00\times 7 = 94.44$;
95% CI for η: $94.44 \pm 2.447\times 0.353 = (93.58, 95.30)$.

(d) $\text{se} = \sqrt{cs^2}$, $\text{pe} = \sqrt{(1+c)s^2}$, so $\text{pe}^2 = s^2 + \text{se}^2$; thus $\text{pe} = 1.116$;
95% PI for Y: $94.44 \pm 2.447\times 1.116 = (91.71, 97.17)$.

Quiz 12 answers: DDBA