**NM303/NM309 Exam Paper 08/09**
**Two Questions for the Statistical Part**

*Two questions in this part are compulsory and each carries 25 marks. Battery-powered pocket calculators may be used. However, answers without any working are not acceptable, so full marks will not be awarded, even if the answer is correct, unless an outline of the calculations carried out is clearly shown.*

**1.(a)** Three parts of a machine are manufactured independently. These three parts have probabilities $p_1, p_2$ and $p_3$ respectively of failing. Find the probability that (i) at least one part fails, (ii) exactly two parts fail.

**(7 marks)**

**(b)** The lengths of bolts produced in a certain factory may be taken to be normally distributed. The bolts are checked on two 'go-no go' gauges so that those shorter than 2.983 in. or longer than 3.021 in. are rejected as 'too-short' or 'too-long' respectively. A sample of $N = 600$ bolts is checked. If for this sample we find 20 'too-short' bolts and 15 'too-long' bolts, find estimates for the mean and standard deviation of the length of these bolts.

**(11 marks)**

**(c)** Assume that a random variable $X$ has the probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ 2x & \text{if } 0 \le x \le 1; \\ 0 & \text{if } x > 1. \end{cases}$$

Find the mean and variance of $X$.

**(7 marks)**

**2.(a)** In an effort to control cost, a quality-control inspector is interested in whether the mean number of ounces of sauce dispensed by bottle-filling machines differs from 16 ounces. From the bottling process, the inspector collects the following measurements:

16.3 16.2 15.8 15.4 16.0 15.6 15.5 16.1 15.9 16.1

Assume that the number of ounces of sauce is normally distributed. Test at $\alpha = 0.05$ whether the bottle-filling machines need adjusting.

**(8 marks)**

**(b)** The following data represent the measurements on a controlled variable $x$ and its dependent variable $y$:

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $y$ | 7 | 5 | 5 | 3 | 2 | 0 |

Find the linear regression equation of $y$ on $x$. Calculate the correlation coefficient and show that it is significantly different from zero at $\alpha = 0.05$ level of significance.

**(9 marks)**

(c) A portion of the ANOVA table for a completely randomised design is shown below:

| Source of variation | SS | d.f. | MS | $F$-ratio |
|---|---|---|---|---|
| Treatments | | 2 | 6.71 | |
| Error | | | | |
| Total variation | 26.86 | 14 | | |

Fill in the blanks in the ANOVA table (you need to show clearly how you work them out). How many treatments are used in this design? Do the data present sufficient evidence to indicate differences among the treatment means? Test using $\alpha = 0.05$.

**(8 marks)**

D

NM 303/309 Exam 08/09
Solutions to Statistics Part

1. (a) Define $E_i = \{\text{part } i \text{ fails}\}$, $i = 1, 2, 3.$

(i) $P(\text{at least one part fails})$

$= 1 - P(\text{no ~~parts~~ fails})$

$= 1 - P(\bar{E}_1 \bar{E}_2 \bar{E}_3) = 1 - p(\bar{E}_1)p(\bar{E}_2)p(\bar{E}_3)$

$= 1 - (1-p_1)(1-p_2)(1-p_3)$

(ii) $P(\text{exactly two parts fail})$

$= P(E_1E_2\bar{E}_3 + E_1\bar{E}_2E_3 + \bar{E}_1E_2E_3)$

$= P(E_1E_2\bar{E}_3) + p(E_1\bar{E}_2E_3) + P(\bar{E}_1E_2E_3)$

$= P(E_1)P(E_2)P(\bar{E}_3) + p(E_1)P(\bar{E}_2)P(E_3) + p(\bar{E}_1E_2E_3)$

$= p_1p_2(1-p_3) + p_1(1-p_2)p_3 + (1-p_1)p_2p_3$

(b) Let $X$ be the lengths of the bolts.

$X \sim N(\mu, \sigma^2)$

By the conditions,

$P(X < 2.983) = \frac{20}{600} = 0.0333$

$P(X > 3.021) = \frac{15}{600} = 0.025$

On the other hand, compute

$P(X < 2.983) = P(Z < \frac{2.983 - \mu}{\sigma}) = 0.0333$

$P(Z > \frac{\mu - 2.983}{\sigma}) = 0.0333$

$P(Z \leq \frac{\mu - 2.983}{\sigma}) = 1 - 0.0333 = 0.9667$

By the normal distribution table,

$P(Z \le 1.84) = 0.9667$

$\therefore \quad \frac{\mu - 2.983}{\sigma} = 1.84$

i.e. $\quad \mu - 1.84\sigma = 2.983 \qquad (1)$

Also, $P(X > 3.021) = P(Z > \frac{3.021-\mu}{\sigma}) = 0.025$

$P(Z \le \frac{3.021-\mu}{\sigma}) = 1 - 0.025 = 0.975$

But, $\quad P(Z \le 1.96) = 0.975$

$\therefore \quad \frac{3.021 - \mu}{\sigma} = 1.96$

$\mu + 1.96\sigma = 3.021 \qquad (2)$

Eq(2) − Eq(1) $\Rightarrow \quad 3.8\sigma = 0.038$

$\sigma = 0.01$ (in.)

$\therefore \quad \mu = 3.021 - 1.96 \times 0.01$

$= 3.001$ (in.)

(c) $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^1 2x^2\, dx$

$= \frac{2}{3} x^3 \Big|_{x=0}^{1} = \frac{2}{3}$

$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_0^1 2x^3\, dx$

$= \frac{2}{4} x^4 \Big|_{x=0}^{1} = \frac{1}{2}$

$Var(X) = E(X^2) - \mu^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2$

$= \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$

2(a) Let $\mu$ = the mean number of ounces of sauce.

$H_0: \mu = 16$ (machines do not need adjusting)

$H_1: \mu \neq 16$ (machines do need adjusting)

Compute

$$\bar{x} = \frac{1}{10}[16.3 + 16.2 + \cdots + 16.1] = 15.89$$

$$s^2 = \frac{1}{10-1}[16.3^2 + 16.2^2 + \cdots + 16.1^2 - 10 \times 15.89^2]$$

$$= 0.09433$$

$$s = 0.307083$$

$$t = \frac{\bar{x} - 16}{s/\sqrt{n}} = \frac{15.89 - 16}{0.307083/\sqrt{10}} = -1.1327$$

Given $\alpha = 0.05$, $n = 10$ find

$$t_{\alpha/2,(n-1)} = t_{0.025;\,9} = 2.262$$

Since $|t| = 1.1327 < t_{0.025,9}$

we do not reject $H_0$. That is, the machines do not need adjusting.

2(c) Compute

$$\bar{x} = \frac{1}{6}(1 + 2 + \cdots + 6) = 3.5$$

$$\bar{y} = \frac{1}{6}(7 + 5 + \cdots + 0) = 3.67$$

$$S_{xx} = 1^2 + 2^2 + \cdots + 6^2 - 6 \times 3.5^2 = 17.5$$

$$S_{xy} = 1 \times 7 + 2 \times 5 + \cdots + 6 \times 0 - 6 \times 3.5 \times 3.67 = -23$$

$$S_{yy} = 7^2 + 5^2 + \cdots + 0^2 - 6 \times 3.67^2 = 31.33$$

$$\hat{a}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-23}{17.5} = -1.314$$

$\hat{a}_0 = \bar{y} - \hat{a}_1 \bar{x} = 3.67 - (-1.314) \times 3.5 = 8.269$

$\therefore \hat{y} = 8.269 - 1.314x$

The correlation coefficient

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{-23}{\sqrt{17.5 \times 31.33}} = -0.982$$

Compute

$$|r| \sqrt{\frac{n-2}{1-r^2}} = 0.982 \times \sqrt{\frac{6-2}{1-0.982^2}} = 10.4$$

Given $\alpha = 0.05$, find

$$t_{\alpha/2, n-2} = t_{0.025, 4} = 2.776$$

Since $|r| \sqrt{\frac{n-2}{1-r^2}} > t_{\alpha/2, n-2}$

we conclude that the correlation coefficient is significantly diffent from 0 at $\alpha = 0.05$ level of sign.

2(c) Note

$SS_{treat} = df_{treat} \times MS_{treat} = 2 \times 6.71 = 13.42$

$SS_{error} = SS_{total} - SS_{treat} = 26.86 - 13.42 = 13.44$

$df_{error} = df_{total} - df_{treat} = 14 - 2 = 12$

$MS_{error} = \frac{SS_{error}}{df_{error}} = \frac{13.44}{12} = 1.12$

$F = \frac{MS_{treat}}{MS_{error}} = \frac{6.71}{1.12} = 5.99$

So the ANOVA is

| Source | SS | df | MS | F |
|---|---|---|---|---|
| treat | 13.42 | 2 | 6.71 | 5.99 |
| Error | 13.44 | 12 | 1.12 | |
| Total | 26.86 | 14 | | |

There are 3 treatments used

$H_0$ : treat. means are the same

$H_1$ : - — — — differ

Given $\alpha = 0.05$, $v_1 = 2$, $v_2 = 12$

find $F_{0.05, 2, 12} = 3.89$

$\because$ $F = 5.99 > 3.89$

$\therefore$ reject $H_0$ and conclude that ~~that~~ treat. means differ.