NM317/318 Exam Paper 10/11
Three Questions for the Statistical Part

Three 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) Four parts of a machine are manufactured independently. They have probabilities p_1, p_2, p_3 and p_4 respectively of failing. Find the probability that (i) at least one part fails, (ii) exactly three parts fail.

(8 marks)

(b) In an effort to control cost, a quality-control inspector is interested in whether the mean weight of sauce dispensed by bottle-filling machines differs from 1kg. From the bottling process, the inspector collects the following measurements (in kg):

0.98 0.99 1.01 1.03 0.99 1.04 1.00 1.01 1.01 1.03

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

(9 marks)

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

$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ cx & \text{if } 0 \le x \le 4; \\ 0 & \text{if } x > 4, \end{cases}$$

where c is a constant. Find the value of c, and then find the mean and variance of X.

(8 marks)

2.(a) A marketing resarch survey shows that 80% of the car owners indicated that their next car purchase would be either a compact or an economy car. Assume the 80% figure is correct, and five prospective buyers are interviewed. (i) Find the probability that all five indicate that their next car purchase would be either a compact or an economy car. (ii) Find the probability that at most one indicates that his/her next car purchase would be either a compact or an economy car.

(7 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	7	8
y	1	3	4	6	6	7	9	9

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

(8 marks)

(c) 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 1.98 cm or longer than 2.02 cm are rejected as 'too-short' or 'too-long' respectively. A sample of $N = 600$ bolts is checked. If for this sample we find 15 'too-short' bolts and 15 'too-long' bolts, find estimates for the mean and standard deviation of the length of these bolts.

(10 marks)

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

Source of variation	Sum of squares	d.f.	Mean square	F-ratio
Treatments		3		
Error			1.12	
Total variation	36.6	19		

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 $\alpha = 0.05$.

(9 marks)

(b) Assume that the random variable X obeys the Rayleigh distribution with the probability density function

$$f(x) = \begin{cases} xe^{-x^2/2} & \text{if } x \geq 0; \\ 0 & \text{if } x < 0. \end{cases}$$

Find (i) $P(X \leq 1)$, (ii) $P(1 < X \leq 4)$, (iii) $P(X > 4)$.

(7 marks)

(c) Test, using $\alpha = 0.05$, the hypothesis that the following set of 200 numbers are 'random digits', that is, each number is equally likely to be $0, 1, 2, \cdots, 9$.

Number	0	1	2	3	4	5	6	7	8	9
Frequency	22	16	15	18	16	25	23	17	24	24

(9 marks)

Solutions to NM317/318 Exam Paper 10/11

1(a) Let $E_i = \{\text{part } i \text{ fails}\}$, $i = 1, 2, 3, 4$.

(i) P(at least one part fails)

$= 1 - P(\text{all 4 parts function})$

$= 1 - P(\bar{E}_1 \bar{E}_2 \bar{E}_3 \bar{E}_4)$

$= 1 - P(\bar{E}_1) P(\bar{E}_2) P(\bar{E}_3) P(\bar{E}_4)$

$= 1 - (1-p_1)(1-p_2)(1-p_3)(1-p_4)$

(ii) P(exactly 3 parts fail)

$= P(E_1 E_2 E_3 \bar{E}_4) + P(E_1 E_2 \bar{E}_3 E_4) + P(E_1 \bar{E}_2 E_3 E_4)$

$+ P(\bar{E}_1 E_2 E_3 E_4)$

$= p_1 p_2 p_3 (1-p_4) + p_1 p_2 (1-p_3) p_4 + p_1 (1-p_2) p_3 p_4$

$+ (1-p_1) p_2 p_3 p_4$

1(b) Let μ = the mean weight.

$H_0 : \mu = 1$ (kg)

$H_1 : \mu \neq 1$ (kg)

Compute

$\bar{x} = \frac{1}{10}[0.98 + 0.99 + \cdots + 1.03] = 1.009$

$s^2 = \frac{1}{10-1}[0.98^2 + 0.99^2 + \cdots + 1.03^2 - 10 \times 1.009^2]$

$= 0.000388$

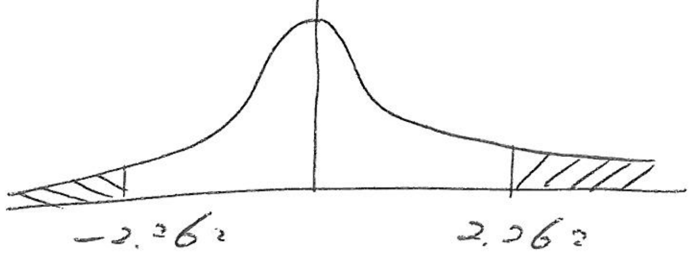
$s = 0.0197$

$t = \frac{\bar{x} - 1}{s/\sqrt{n}} = \frac{1.009 - 1}{0.0197/\sqrt{10}} = 1.445$

Given $\alpha = 0.05$. $d.f. = n-1 = 9$.

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

The reject region is



$\because$ $t = 1.445$ is not in the rejection region

$\therefore$ we don't reject H_0. That is, the machines do not need adjusting.

1(c) For

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^4 cx\,dx = \frac{1}{2}cx^2\Big|_{x=0}^{4}$$

$$= \frac{1}{2}c \cdot 4^2 = 8c,$$

we then have $c = \frac{1}{8}$. Compute then

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^4 \frac{1}{8}x^2\,dx$$

$$= \frac{1}{24}x^3\Big|_{x=0}^{4} = \frac{1}{24} \times 4^3 = \frac{8}{3}$$

$$Var(X) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

$$= \int_0^4 \frac{1}{8}x^3\,dx - \mu^2 = \frac{1}{32} \times 4^4 - \left(\frac{8}{3}\right)^2$$

$$= \frac{8}{9}$$

2(a) Let X = the number of car owners who indicate that their next purchase will be a compact or an economy car. Then $X \sim B(5, 0.8)$

(i) $P(X=5) = \binom{5}{5} 0.8^5 \, 0.2^0 = 0.8^5 = 0.32768$

(ii) $P(X \le 1) = P(X=0) + P(X=1)$

$= \binom{5}{0} 0.8^0 \, 0.2^5 + \binom{5}{1} 0.8^1 \, 0.2^4$

$= 0.00064 + 0.00032$

$= 0.00672$

2(b) $\bar{x} = \frac{1}{8}(1+2+\cdots+8) = 4.5$

$\bar{y} = \frac{1}{8}(1+3+\cdots+9) = 5.625$

$S_{xx} = 1^2+2^2+\cdots+8^2 - 8\times 4.5^2 = 42$

$S_{xy} = 1\times 1 + 2\times 3 + \cdots + 8\times 9 - 8\times 4.5\times 5.625 = 47.5$

$S_{yy} = 1^2+3^2+\cdots+9^2 - 8\times 5.625^2 = 55.875$

$\hat{a}_1 = \frac{S_{xy}}{S_{xx}} = \frac{47.5}{42} = 1.130952$

$\hat{a}_0 = \bar{y} - \hat{a}_1 \bar{x} = 5.625 - 1.13095 \times 4.5 = 0.53576$

$\therefore \hat{y} = 0.53576 - 1.13095x$

The correlation coefficient

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{47.5}{\sqrt{42 \times 55.875}} = 0.9805$$

Compute

$$|r|\sqrt{\frac{n-2}{1-r^2}} = 0.9805 \times \sqrt{\frac{8-2}{1-0.9805^2}} = 12.021$$

Given $\alpha = 0.05$, find

$$t_{\alpha/2,\, n-2} = t_{0.025,\, 6} = 2.447$$

$$\therefore \quad |r|\sqrt{\frac{n-2}{1-r^2}} > 2.447$$

We conclude that the correlation coefficient is significantly different from 0 at $\alpha = 0.05$.

2c) Let X be the length of a bolt. Then $X \sim N(\mu, \sigma^2)$.

Note

$$P(X < 1.98) = \frac{15}{600} = 0.025$$

So

$$P\left(Z < \frac{1.98-\mu}{\sigma}\right) = 0.025$$

where $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. This implies

$$P\left(Z > \frac{\mu-1.98}{\sigma}\right) = 0.025$$

or

$$P\left(Z \le \frac{\mu-1.98}{\sigma}\right) = 0.975$$

But, by the normal distribution table,

$$P(Z \le 1.96) = 0.975$$

Hence

$$\frac{\mu-1.98}{\sigma} = 1.96 \qquad (1)$$

On the other hand,

$$P(X > 2.02) = \frac{15}{600} = 0.025$$

$$P\left(Z > \frac{2.02-\mu}{\sigma}\right) = 0.025$$

$$P\left(Z \le \frac{2.02-\mu}{\sigma}\right) = 0.975$$

$$\therefore \quad \frac{2.02-\mu}{\sigma} = 1.96 \qquad (2)$$

(1) and (2) $\Rightarrow$ $\mu - 1.98 = 2.02 - \mu$, $\mu = 2$

$$\sigma = \frac{2-1.98}{1.96} = 0.0102$$

3(a) $d.f_{error} = d.f_{total} - d.f_{treat} = 19 - 3 = 16$

$SS_{error} = df_{error} \times MS_{error} = 16 \times 1.12 = 17.92$

$SS_{treat} = SS_{total} - SS_{error} = 36.6 - 17.92 = 18.68$

$$MS_{treat} = \frac{SS_{treat}}{d.f_{treat}} = \frac{18.68}{3} = 6.2267$$

$$F = \frac{MS_{treat}}{MS_{error}} = \frac{6.2267}{1.12} = 5.56$$

So the ANOVA is

Source	S.S	d.f.	M.S	F
treat	18.68	3	6.2267	5.56
Error	17.92	16	1.12	
Total	36.6	19		

There are 4 treatments used.

H_0: treat means are the same

H_1: treat means differ

Given $\alpha = 0.05$, $v_1 = 3$, $v_2 = 16$, find

$F_{0.05, 3, 16} = 3.24$

$\because F = 5.56 > 3.24$

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

3(b) (i) $P(X \le 1) = \int_{-\infty}^{1} f(x)\,dx = \int_0^1 x e^{-x^2/2}\,dx$

$= \left[-e^{-x^2/2}\right]_{x=0}^{1} = 1 - e^{-\frac{1}{2}} = 0.3935$

(ii) $P(1 < X \le 4) = \int_1^4 x e^{-x^2/2}\,dx$

$= \left[-e^{-x^2/2}\right]_{x=1}^{10} = e^{-\frac{1}{2}} - e^{-8} = 0.6062$

(iii) $P(X > 4) = 1 - P(X \le 1) - P(1 < X \le 4)$

$= 1 - (1 - e^{-\frac{1}{2}}) - (e^{-\frac{1}{2}} - e^{-8}) = e^{-8} = 0.00034$

3(c) H_0 : the set of numbers are random digits

H_1 : - - - - are Not . . .

Compute

$$X^2 = \sum_{i=0}^{9} \frac{(n_i - np_i)^2}{np_i}$$

when $n = 200$, $p_i = \frac{1}{10}$, p_i's are the freq.

Hence

$$X^2 = \frac{(22-20)^2}{20} + \frac{(16-20)^2}{20} + \cdots \frac{(24-20)^2}{20} + \frac{(24-20)^2}{20}$$

$$= 7.$$

Given $\alpha = 0.05$, d.f. $= 10 - 1 = 9$. find

$$\chi_9^2(0.05) = 16.92$$

As $X^2 < \chi_9^2(0.05)$, we accept H_0 and conclude that the # are random digits