THE UNIVERSITY OF SUSSEX

G1042

BSc/MMath EXAMINATIONS 2009

STATISTICS FOR ENGINEERS

You may attempt as many questions as you wish, but marks will be given for the best FOUR answers only.

Time allowed: ONE hour.

Each question carries TWENTY marks. The numbers beside the questions indicate the approximate marks that can be gained from the corresponding parts of the questions.

Examination handout: Maths Dept Statistical Tables, Statistical Formulae (Engineering Statistics)

Turn over/

1. (a) Write down the formula for the conditional probability of A given B, $P(A|B)$, assuming that $P(B) > 0$. Hence show that, if A and B are independent, then

$$P(A \text{ and } B) = P(A) \times P(B).$$ **[4 marks]**

(b) A quality control system measures the masses of items coming off a production line. If the mass lies outside a certain range, the production line is stopped. Unfortunately, for each item there is a 6% chance of the line stopping even though the item is satisfactory, and there is an 8% chance of it not stopping even when the item is flawed. The probability of an item being flawed is 0.15. Find the probability that

(i) the line stops for a particular item, **[8 marks]**

(ii) if the line has stopped, it is because the item is flawed. **[8 marks]**

2. In recent weeks, 40% of my e-mails have been spam.

(a) If, in a given period, I receive 10 e-mails, write down the distribution of x, the number of spam e-mails I receive. Find Prob($x = 7$). **[8 marks]**

(b) In a week, I get 150 e-mails. Use the Normal approximation to find Prob ($y < 50$), where y is the number of spam e-mails I get. **[12 marks]**

3. A sample of widgets have the following diameters, in cms: 6.21, 6.30, 6.29, 6.27, 6.24.

(a) Find the mean and standard deviation of the diameters of the sample. **[6 marks]**

(b) Assuming that widget diameters are Normally distributed, find the 95% Confidence Interval for the mean diameter of the population. **[8 marks]**

(c) Find the number that should be sampled in order that this Confidence Interval should have a width of 0.02 cm. **[6 marks]**

4. The lifetime t in hours of a certain electronic component is a random variable having the probability density function:

$$\begin{aligned} f(t) &= 0 \text{ for } t < 100 \\ f(t) &= k/t^3 \text{ for } t \geq 100, \text{ where } k \text{ is a constant.} \end{aligned}$$

(a) Show that $k = 20,000$. **[6 marks]**

(b) Find the mean lifetime of the components. **[6 marks]**

(c) Find the probability that a component will need replacing within the first 150 hours of operation. **[4 marks]**

(d) Find the "half-life" of a component, that is, the time after which there is a probability of 1/2 that it will need replacing. **[4 marks]**

5. An investigation of the acoustics of different concert halls gave the following results for d, the front-back distance in metres, and t, the reverberation time in seconds:

d	20	35	41	42	50	56	60
t	1.2	1.4	2.7	2.6	2.9	3.1	4.5

   (a) Find the least squares estimate of the regression line of t on d. Give a physical interpretation of the constants in the equation of the line. **[8,2,2 marks]**

   (b) Predict the reverberation time when (i) $d = 38$ and (ii) when $d = 10$. Comment on the reliability of each of these estimates. **[8 marks]**

6. In a certain production process, the proportion of p of defective items is unknown. A two-stage plan for the quality control procedure is as follows: Select 40 items from a batch and accept the batch if none are defective. Reject the batch if two or more are faulty. If one is faulty take another 10 items and reject the batch if any of these are faulty; otherwise accept the batch.

   (a) Find, in terms of p, the probability $L(p)$ that a batch is accepted. Evaluate $L(p)$ at $p = 0$, $p = 0.01$ and $p = 0.1$; hence sketch the graph of the operating characteristic. Comment on the shape of this graph. **[14 marks]**

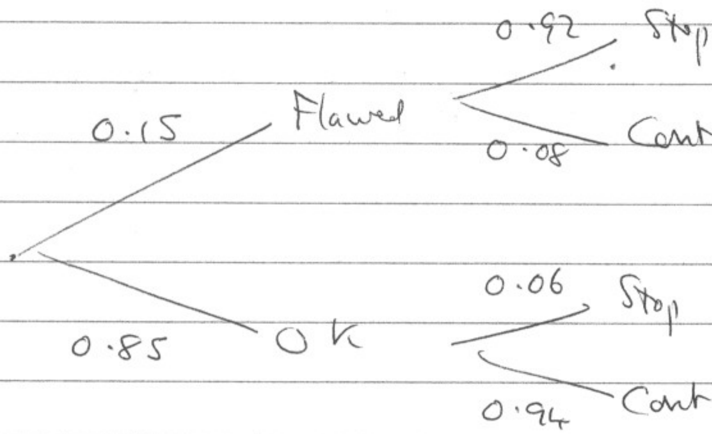
   (b) Show that the expected number of items sampled under this method is

$$40(1 + 10p(1 - p)^{39}).$$ **[6 marks]**

Stats for Engineers 2009

1 $P(A|B) = \dfrac{P(A \text{ and } B)}{P(B)} = P(A)$ if independent

$\therefore P(A \text{ and } B) = P(A) \times P(B)$



$P(\text{stops}) = 0.15 \times 0.92 + 0.85 \times 0.06 = \cancel{0.0966}\ 0.138 + 0.051$

$= \cancel{0.1476}\ 0.189$

$P(\text{Flaw}|\text{stop}) = \dfrac{\cancel{0.0966}\ 0.138}{\cancel{0.1476}\ 0.189} = \cancel{0.654}\ 0.730$ ✓

2 a) $x \sim Bin(10, 0.4)$

$P(x=7) = C^{10}_{7}\, 0.4^7 \times 0.6^3 = 120 \times 0.00164 \times 0.216 = 0.0425$

b) $y \sim Bin(150, 0.4) \simeq N(60, 36)$

$\therefore P(y < 50) = P(y < 49.5)$

$= P\left(z < \dfrac{49.5 - 60}{6}\right)$

$= P(z < -1.75) = 1 - 0.9599$

$= 0.0401$

3a) $m = \frac{31.31}{5} = 6.262$ $\quad s^2 = \frac{5}{4}\left(\frac{\cancel{196.07}\ 196.0687}{5} - 6.262^2\right) = \cancel{0.001695}\ 0.00137$

$\therefore\ s = \cancel{0.0412}\ 0.0370$

b) $95\%\ CI = 6.262 \pm t_{n-1}\frac{s}{\sqrt{n}} = 6.262 \pm 2.776 \times \frac{\cancel{0.0412}\ 0.0370}{\sqrt{5}}$

$= 6.262 \pm \cancel{0.0511}\ 0.0459 = \cancel{6.211}\ 6.216 \text{ to } \cancel{6.313}\ 6.308$

c) Need $t_{n-1}\frac{s}{\sqrt{n}} = \cancel{0.1}\ 0.01$

Assume n is large, so $t_{n-1} \approx 1.96$

$n = \frac{s^2 \times 1.96^2}{0.01^2} = 65\quad 53$

4 a) $\int_{100}^{\infty} \frac{k}{t^3}\,dt = 1 \qquad \left[\frac{-k}{2t^2}\right]_{100}^{\infty} = 1 \qquad 0 + \frac{k}{2\times 100^2} = 1 \qquad k = 20{,}000$

b) $\mu = \int_{100}^{\infty} t\frac{k}{t^3}\,dt = \left[\frac{-k}{t}\right]_{100}^{\infty} = 0 + \frac{k}{100} = 200$

c) $\int_{100}^{150} \frac{k}{t^3}\,dt = \left[\frac{-k}{2t^2}\right]_{100}^{150} = \frac{-k}{2}\left(\frac{1}{150^2} - \frac{1}{100^2}\right) = 0.5556$

d) Need $\int_{100}^{\tau} \frac{k}{t^3}\,dt = \frac{1}{2} \qquad \frac{-k}{2}\left(\frac{1}{\tau^2} - \frac{1}{100^2}\right) = \frac{1}{2}$

$\frac{1}{\tau^2} - \frac{1}{100^2} = -\frac{1}{k} \qquad \frac{1}{\tau^2} = 0.00005 \qquad \tau = 141.4$ hrs

5a) $\Sigma d = 304$ $\Sigma t = 18.4$ $\Sigma d^2 = 14306$ $\Sigma dt = 881.5$

$\therefore \ln t = a + bd$ $b = \dfrac{881.5 - \frac{304 \times 18.4}{7}}{14306 - 304^2/7} = 0.0747$

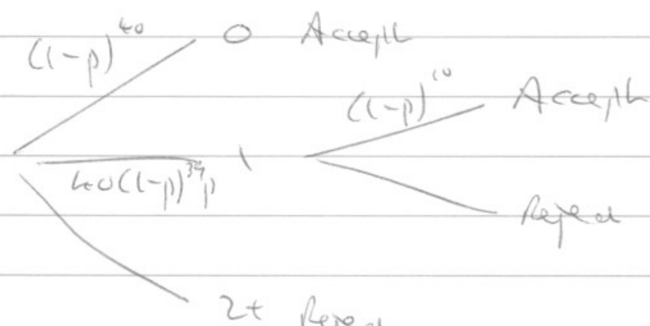
Then $a = \bar{t} - b\bar{d} = -0.6142$

$\therefore$ $t = 0.0747d - 0.6142$

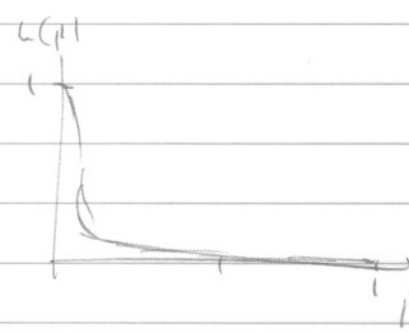
Intercept = value of t when $d = 0$ —
t is −ve for $d = 0$, which is clearly unrealistic
The gradient is the increase in time for each extra 1m increase in d.

b) $d = 38$ $t = 2.22s$, $d = 10$ $t = 0.133s$
— the $d = 10$ result is unreliable, since it's outside the given data range — the regression line may change.

6 a)



$L(p) = (1-p)^{40} + 40p(1-p)^{49}$



$L(0) = 1$ $L(0.01) = 0.913$ $L(0.1) = 0.038$

$L(p)$ decreases rapidly as p increases from 0, as required.

b) Expect $40 + 40p(1-p)^{39} \times 10 = 40(1 + 10p(1-p)^{39})$