STATISTICS II
MARCH 2007 – SOLUTIONS

(1)

(a) i) Probability of "success" is $\frac{3}{4}$. $X \sim \text{Geometric}(0.75)$

ii) $P(X \geq 4) = P(X > 3)$

$= P(\text{first 3 draws are Diamonds})$

$= (1/4)^3$ [by independence]

$= \frac{1}{64} = 0.0156$

(b) i) $P(Y > 2) = 1 - P(Y \leq 2)$

$= 1 - \frac{15}{16} = \frac{1}{16}$

$= 0.0625$

$$P(Y>3 / Y>1) = \frac{P(Y>3 \cap Y>1)}{P(Y>1)} = \frac{P(Y>3)}{P(Y>1)}$$

$$= \frac{1 - P(Y \leq 3)}{1 - P(Y \leq 1)} = \frac{1 - 63/64}{1 - 3/4}$$

$$= \frac{1}{16} = 0.0625$$

ii) Answers in i) suggest that the distribution of Y has the lack-of-memory property.

2

(a) $P(150 < T < 200) = P(T \leq 200) - P(T \leq 150)$

$= F_T(200) - F_T(150) = e^{-0.005\times 150} - e^{-0.005\times 200}$

$= e^{-0.75} - e^{-1} = 0.4724 - 0.3679$

$= \underline{0.1045}$

(b) Clearly $Y = \frac{1}{7} T$.

Then, using the rescaling property of the exponential distribution, we have

$$Y \sim \text{Exp}\left(\frac{0.005}{1/7}\right) \quad \text{i.e.} \quad \underline{Y \sim \text{Exp}(0.035)}.$$

(c) We want

$P(Y > 32) = 1 - (1 - e^{-0.035\times 32})$

$= e^{-1.12} = \underline{0.3263}$

[OR, $P(T > 32\times 7) = P(T > 224)$

$= e^{-0.005\times 224}$

$= e^{-1.12} = 0.3263$].

3

(a) $$E(X) = \int_0^{\beta} x f(x)\, dx$$

$$= \int_0^{\beta} x \frac{2(\beta - x)}{\beta^2}\, dx = \frac{2}{\beta^2}\left[\frac{x^2\beta}{2} - \frac{x^3}{3}\right]_0^{\beta}$$

$$= \frac{2}{\beta^2}\left(\frac{\beta^3}{2} - \frac{\beta^3}{3}\right) = \frac{2}{\beta^2}\,\frac{\beta^3}{6}$$

$$= \frac{\beta}{3}$$

(b) For the method of moments we want

$$E(X) = \bar{x} \Rightarrow \frac{\hat{\beta}}{3} = \bar{x}$$

$$\Rightarrow \frac{\hat{\beta}}{3} = 0.65$$

$$\Rightarrow \hat{\beta} = 1.95$$

(4)

(a) Clearly $f_X(x) \geq 0$ for $1 < x < \infty$, so we need $\int_1^\infty f_X(x)\,dx = 1$. We have:

$$\int_1^\infty f_X(x)\,dx = \int_1^\infty x^{-2}\,dx = \left[-x^{-1}\right]_1^\infty$$

$$= \left(0 - (-1)\right) = \underline{1}$$

(b) i) For $1 < x < \infty$ and $Y = X^{-1}$ we have

$$\underline{Y \in (0,1)}.$$

ii) $g(Y) = X^{-1} \Rightarrow g^{-1}(Y) = Y^{-1}$ (or $X = Y^{-1}$)

with $g^{-1}()$ being a decreasing function.

$$f_Y(y) = f_Y\{g^{-1}(y)\} \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$= f_X(x) \left| \frac{dx}{dy} \right| = \frac{1}{(y^{-1})^2} \left| -y^{-2} \right|$$

$$= \underline{1}$$

iii) From i) and ii), Y has the pdf of a uniform$(0,1)$ r.v.

$$\therefore \quad \underline{Y \sim U(0,1)}.$$

5

(a) If X is the r.v. giving the test score, we have

$$X \sim N(60, 15^2).$$

$$P(85 < X < 95) = P\left(\frac{85-60}{15} < Z < \frac{95-60}{15} \right)$$

[where $Z \sim N(0,1)$].

$$= P(Z < 2.33) - P(Z < 1.67)$$

$$= \Phi(2.33) - \Phi(1.67) = 0.9901 - 0.9525$$

$$= \underline{\underline{0.0376}}$$

(b) From (a), proportion of students between 85 and 95 is $\underline{3.76\%}$.

(c) We want a value $x_{0.1}$ which is such that

$$P(X > x_{0.1}) = 0.1 \Rightarrow P\left(Z > \frac{x_{0.1} - 60}{15}\right) = 0.1$$

where $Z \sim N(0,1)$.

From tables we find

$$P(Z > 1.2816) = 0.1$$

Thus,

$$\frac{x_{0.1} - 60}{15} = 1.2816$$

$$\Rightarrow x_{0.1} = 15 \times 1.2816 + 60 \Rightarrow \underline{\underline{x_{0.1} = 79.224}}$$

6

(a) $E(\bar{X}) = E\left\{\frac{1}{n}(X_1 + \cdots + X_n)\right\}$

$$= \frac{1}{n}\left(E(X_1) + \cdots + E(X_n)\right) = \frac{1}{n} n\mu = \underline{\mu}$$

$$\text{var}(\bar{X}) = \text{var}\left\{\frac{1}{n}(X_1 + \cdots + X_n)\right\}$$

$$= \frac{1}{n^2}\left\{\text{var}(X_1) + \cdots + \text{var}(X_n)\right\} \quad \text{[by \underline{independence}]}$$

$$= \frac{1}{n^2} n\sigma^2 = \underline{\frac{\sigma^2}{n}}$$

(b) CLT gives: $\bar{X} \sim N\left(\mu, \sigma^2/n\right)$ approximately.

(c) If Y is the number of "heads", then $Y \sim \text{Bin}(100, \frac{1}{2})$

and we can use $X \sim N(50, 25)$ to approximate it.

Then,

$$P(Y > 40) \simeq P(X > 40.5)$$ [where $X \sim N(50,25)$ and using continuity correction]

$$= P\left(Z > \frac{40.5 - 50}{5}\right)$$ [where $Z \sim N(0,1)$]

$$= P(Z > -1.9)$$

$$= P(Z < 1.9)$$ [by symmetry of normal distⁿ.]

$$= \Phi(1.9) = \underline{0.971}$$

(7)

(a) $Q_1 = x_{\left(\frac{30+1}{4}\right)} = x_{(7.75)} = 29.1 + \frac{3}{4} \times 0.6 = \underline{29.55}$

Median $= x_{\left(\frac{30+1}{2}\right)} = x_{(15.5)} = \underline{31.5}$

$Q_3 = x_{\left(\frac{3(30+1)}{4}\right)} = x_{(23.25)} = 33.3 + \frac{1}{4} 0.2 = \underline{33.35}$

So, 5-point summary is given by

min	Q1	Median	Q3	max
25.1	29.55	31.5	33.35	36.3

(c)

25	1 5
26	
27	2 9
28	8 9
29	1 7 9
30	1 3 6 6
31	1 3 7 7 9
32	2 4 6 9
33	3 5 8
34	2 5 9
35	3
36	3

The distribution looks fairly symmetric and "bell" shaped.

Plot suggests data are Normal.

(b) $IQR = Q_3 - Q_1 = 33.35 - 29.55 = \underline{3.8}$. It gives the length of the interval in which the central 50% of the data lies.

(d) $\bar{x} = \frac{\sum x_i}{n} = \frac{937.3}{30} = 31.243$

Then, for known variance, a 90% for μ is given by

$$\bar{x} \pm z_{0.05} \frac{\sigma}{\sqrt{n}} = 31.243 \pm 1.645 \sqrt{\frac{8.1}{30}}$$

$$= 31.243 \pm 0.855, \quad \text{i.e.} \quad (30.388, 32.098)$$

(8)

(a) $\bar{x} = \frac{\sum x_i}{7} = \frac{523}{7} = 74.714$

$$s^2 = \frac{1}{6}\left\{\sum x_i^2 - \frac{1}{7}\left(\sum x_i\right)^2\right\} = \frac{1}{6}\left(39321 - \frac{523^2}{7}\right)$$

$$= 40.905$$

Now, assuming that the data come from a <u>normal population</u>, a 95% CI for μ is given by

$$\bar{x} \pm t_{6,0.025}\frac{s}{\sqrt{n}} = 74.714 \pm 2.447\sqrt{\frac{40.905}{7}}$$

$$= 74.714 \pm 5.915$$

i.e. $\underline{(68.799\,,\ 80.629)}$

(b) If the value 86 mph is replaced by 76 mph, then obviously the <u>sample variance will be smaller</u>.

With the sample size remaining the same, this will give smaller standard error for the sample mean $(s/\sqrt{n})$, and therefore a <u>narrower</u> CI for μ.