HONG KONG STATISTICAL SOCIETY

2015 EXAMINATIONS − SOLUTIONS

HIGHER CERTIFICATE – MODULE 2

The Society is providing these solutions to assist candidates preparing for the examinations in 2017.

The solutions are intended as learning aids and should not be seen as "model answers".

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.

HC2 Exam Solutions 2015

1(a)(i) $P(A|B) = \frac{P(A \cap B)}{P(B)}$ ① for numerator ① for denominator

②

(ii) (A) If A and B are independent $P(A \cap B) = P(A) \times P(B)$

$\Rightarrow P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$ ①

①

(B) A and B are mutually exclusive if $P(A \cap B) = 0$

$\Rightarrow P(A|B) = 0$ ①

①

(C) $B \subset A \Rightarrow P(A \cap B) = P(B)$

$\Rightarrow P(A|B) = \frac{P(B)}{P(B)} = 1$ ①

①

(b) let D be person has disease
+ be positive test result
− be negative test result

$\bar{D}$ be person does not have disease

$P(+|D) = 0.98$ $P(-|\bar{D}) = 0.99$

$P(D) = 0.03$

(i) Want $P(D|+)$

$= \frac{P(+|D)P(D)}{P(+)}$

$P(+) = P(+|D)P(D) + P(+|\bar{D})P(\bar{D})$ ①

$= 0.98 \times 0.03 + 0.01 \times 0.97$

$= 0.0385$ ①

③ $P(D|+) = \frac{0.98 \times 0.03}{0.0385} = 0.7634$ ①

(ii) $P(D|++) = \frac{P(++|D)P(D)}{P(++|D)P(D) + P(++|\bar{D})P(\bar{D})}$ ①

$= \frac{0.98^2 \times 0.03}{0.98^2 \times 0.03 + 0.01^2 \times 0.97}$ ① $= 0.9966$ ①

③

(c) $P(M \text{ wins}) = \frac{1}{2}$ $P(M \text{ draws}) = \frac{1}{6}$ $P(M \text{ loses}) = \frac{1}{3}$

$P(C \text{ wins}) = \frac{2}{3}$ $P(C \text{ draws}) = \frac{1}{6}$ $P(C \text{ loses}) = \frac{1}{6}$

(i) M wins if M wins and C does not win
or M draws and C loses

$P(M \text{ wins championship}) = P(M \text{ wins} \times C \text{ does not win}) + P(M \text{ draws} \times C \text{ loses})$ ①

$= P(M \text{ wins}) \times P(C \text{ does not win}) + P(M \text{ draws}) \times P(C \text{ loses})$ ①
as indep. games

$= \frac{1}{2} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{6}$ ①

$= \frac{7}{36}$ ①

④

(ii) $P(C \text{ draws} \mid M \text{ wins championship})$

$= \frac{P(C \text{ draws} \times M \text{ wins championship})}{P(M \text{ wins championship})}$ ① (ie M wins game)

$= \frac{P(C \text{ draws}) \times P(M \text{ wins game})}{P(M \text{ wins championship})}$ ① by indep. ①

$= \frac{\frac{1}{6} \times \frac{1}{2}}{\frac{7}{36}}$ ①

$= \frac{3}{7}$ ①

⑤

20

2 (a) A continuity correction is needed when a discrete random variable is approximated by a continuous random variable

$$P(X \leq x) = P(X < x+1)$$

① ① ①

and so you need to allow for this in your approximation

③ Accept other equivalent explanations.

(b) Poisson approximation to Binomial is used when n is large and p is small (or p large ie interchange success & failure),

① possible values $p < 0.1$ or $(p > 0.9)$ and ① $np > 5$.

Normal approximation to Binomial is used when n is large and p is not small, ie $0.1 < p < 0.9$, better

① approximation if ① p closer to 0.5.

④ Marks given for sensible answers.

(c) Fair die $n = 300$ $P(X < 45 \text{ sixes})$

Let X = no. of sixes $X \sim \text{Binomial}\left(300, \frac{1}{6}\right)$

Use a normal approximation to Binomial with $\mu = 50, \sigma^2 = \frac{125}{3}$

$$P(X < 45) = P(X \leq 44)$$

$$= P\left(Z < \frac{44\frac{1}{2} - 50}{\sqrt{\frac{125}{3}}}\right) \quad ①$$

$$= P(Z < -0.852) \quad ①$$

$$= P(Z > 0.852)$$

$$= 1 - \Phi(0.852) \quad ①$$

$$= 1 - 0.80298 \quad ① \text{ using interpolation}$$

⑤ $$= 0.19702 \quad ①$$

No marks if use a Poisson approximation.

(d) X = no. of accidents in one working week

$$X \sim Po(0.2)$$

Let Y = no. of accidents in 150 working weeks

$$Y \sim Po(30.0)$$

② ① ①

2. (d) continued.

$P(Y \geq 35)$ $\mu = 30.0$ ① $\sigma^2 = 30.0$ ①

Use a Normal approximation to Poisson.

$$P(Y \geq 35) = P\left(Z > \frac{34.5 - 30.0}{\sqrt{30.0}}\right) \text{①}$$

$$= P(Z > 0.8216)$$

$$= 1 - P(Z < 0.8216) \quad \text{①}$$

$$= 1 - 0.7943 \quad \text{①} \quad \text{using interpolation}$$

⑥ $$= 0.2057 \quad \text{①}$$

[20]

3. $f(y) = \begin{cases} k(y+y^3) & 0 < y < 2 \\ 0 & \text{ow} \end{cases}$

(i) $\int_0^2 f(y)\,dy = 1$ ①

$k\int_0^2 y + y^3\,dy = 1$ ①

$k\left[\frac{y^2}{2} + \frac{y^4}{4}\right]_0^2 = 1$ ①

④ $k[2+4] = 1 \Rightarrow k = \frac{1}{6}$ ①

(ii) $F(y) = P(Y \le y)$ ①

$= \int_0^y \frac{1}{6}(u+u^3)\,du$ ①

$= \left[\frac{1}{6}\left(\frac{u^2}{2} + \frac{u^4}{4}\right)\right]_0^y$ ①

$= \frac{1}{6}\left(\frac{y^2}{2} + \frac{y^4}{4}\right)$ ①

$= \frac{y^2}{12}\left(\frac{y^2+2}{2}\right)$ $0 < y < 2$ ①

$\therefore F(y) = \begin{cases} 0 & y \le 0 \\ \frac{y^2}{12}\left(\frac{y^2+2}{2}\right) & 0 < y < 2 \\ 1 & y \ge 2 \end{cases}$

⑤

$P\left(\frac{1}{2} < y < \frac{3}{2}\right) = F\left(\frac{3}{2}\right) - F\left(\frac{1}{2}\right)$ ①

$= \frac{\left(\frac{3}{2}\right)^2}{12}\left(\frac{\left(\frac{3}{2}\right)^2+2}{2}\right) - \frac{\left(\frac{1}{2}\right)^2}{12}\left(\frac{\left(\frac{1}{2}\right)^2+2}{2}\right)$ ①

$= \frac{9}{48}\left(\frac{9+8}{8}\right) - \frac{1}{48}\left(\frac{1+8}{8}\right) = \frac{9}{24} = \frac{3}{8}$ ①

③

(iii) $\text{Var}[Y] = E[Y^2] - (E[Y])^2$ ①

$$E[Y] = \frac{1}{6}\int_0^2 y(y+y^3)\,dy \quad ①$$

$$= \frac{1}{6}\left[\frac{y^3}{3} + \frac{y^5}{5}\right]_0^2 \quad ①$$

$$= \frac{1}{6}\left[\frac{8}{3} + \frac{32}{5}\right]$$

$$= \frac{1}{6}\left[\frac{40+96}{15}\right] = \frac{136}{\frac{6}{3}\times 15} = 1.511\left(\frac{68}{45}\right) \quad ①$$

$$E[Y^2] = \frac{1}{6}\int_0^2 y^2(y+y^3)\,dy \quad ①$$

$$= \frac{1}{6}\left[\frac{y^4}{4} + \frac{y^6}{6}\right]_0^2 \quad ①$$

$$= \frac{1}{6}\left[4 + \frac{32}{3}\right] = \frac{44}{18} = \frac{22}{9} \quad ①$$

$$\therefore \text{Var}[Y] = E[Y^2] - (E[Y])^2$$

$$= \frac{22}{9} - \left(\frac{68}{45}\right)^2 = 0.161 \quad ①$$

(g)

20

4. $X_1, X_2, \ldots, X_n$ are independent $X_i \sim N(\mu_i, \sigma^2)$ $i = 1, 2, \ldots, n$

$$T = \sum_{i=1}^{n} X_i$$

(i)(a) $E[T] = \sum_{i=1}^{n} \mu_i$ ①

(b) $Var[T] = \sum_{i=1}^{n} \sigma_i^2$ ① by independence

(c) Normal ①

(3)

(ii) Let X_i be weight of each hardback book $i = 1, 2, 3$

$$X_i \sim N(1.5, 0.5^2)$$

Let Y_j be weight of each paperback book $j = 1, 2, \ldots, 6$

$$Y_j \sim N(0.8, 0.3^2)$$

(a) Let W_1 be total weight of parcel ⇒ $W = X_1 + X_2 + X_3 + Y_1 + \ldots + Y_6$ ①

$$E[W_1] = \sum_{i=1}^{3} E[X_i] + \sum_{j=1}^{6} E[Y_j]$$

$$= 3 \times 1.5 + 6 \times 0.8 = 9.3 \quad ①$$

$$Var[W_1] = \sum_{i=1}^{3} Var[X_i] + \sum_{j=1}^{6} Var[Y_j] \quad ① \text{ by indep}$$

$$= 3 \times 0.5^2 + 6 \times 0.3^2 = 1.29 \quad ①$$

$$P(W_1 > 11) = P\left(Z > \frac{11 - 9.3}{\sqrt{1.29}}\right) \quad ①$$

$$= P(Z > 1.4968)$$

$$= 1 - \Phi(1.4968) \quad ①$$

$$= 1 - 0.9328$$

(7) $$= 0.0672 \quad ①$$

cont. 4 (b) $P(X_1 + X_2 + X_3 > Y_1 + Y_2 + \ldots + Y_6)$

$= P(X_1 + X_2 + X_3 - Y_1 - Y_2 - \ldots - Y_6 > 0)$

Let $W_2 = X_1 + X_2 + X_3 - Y_1 - Y_2 - Y_3 - Y_4 - Y_5 - Y_6$

$E[W_2] = \sum_{i=1}^{3} E[X_i] - \sum_{j=1}^{6} E[Y_j]$

$= 3 \times 1.5 - 6 \times 0.8 = -0.3$ ①

$Var[W_2] = \sum_{i=1}^{3} Var[X_i] + \sum_{j=1}^{6} Var[Y_j]$ by indep

$= 1.29$ ①

$P(W_2 > 0) = P\left(Z > \frac{0 - (-0.3)}{\sqrt{1.29}}\right)$ ①

$= P(Z > 0.2641)$

$= 1 - \Phi(0.2641)$ ①

$= 1 - 0.6042$

$= 0.3958$ ① (5)

(c) Let n_H be the maximum number of hardback books on pallet

Let $W_3 = \sum_{i=1}^{n_H} X_i$

$E[W_3] = 1.5 n_H$ ①

$Var[W_3] = 0.5^2 n_H = 0.25 n_H$ ①

$P(W_3 > 100) < 0.01 \Rightarrow P\left(Z > \frac{100 - 1.5 n_H}{\sqrt{0.25 n_H}}\right) < 0.01$ ①

$\Rightarrow \frac{100 - 1.5 n_H}{\sqrt{0.25 n_H}} > 2.3263$ ① $\Rightarrow 100 - 1.5 n_H > 1.1632\sqrt{n}$

$\Rightarrow 1.5n + 1.1632\sqrt{n} - 100 < 0$ ①

20 (5)