Math7501 Examination 2009: Sample Solutions and Marking Schemes

1. (a) (i) A sample space is a collection of all the possible outcomes of an experiment.

   (ii) An event is a subset of the sample space.

   (iii) A random variable is a real valued function defined on the sample space.

   (b) The three axioms of probability for a probability function $P(\cdot)$ are:

   1. For any $E \in \mathcal{F}$, $P(E) \geq 0$.
   2. $P(\Omega) = 1$.
   3. If $E_1 \in \mathcal{F}, E_2 \in \mathcal{F}, \ldots$ are mutually disjoint, then $P\left(\cup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} P(E_j)$.

   (c) (i) Events A, B and C are mutually disjoint if $A \cap B = \emptyset$, $A \cap C = \emptyset$ and $B \cap C = \emptyset$.

   (ii) Events A, B and C are mutually independent if $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$, and $P(A \cap B \cap C) = P(A)P(B)P(C)$.

   (d) If $P(A \cap B) > 0$, then both $P(B|A)$ and $P(C|A \cap B)$ are well defined and

   $$P(A)P(B|A)P(C|A\cap B) = P(A) \times \frac{P(A \cap B)}{P(A)} \times \frac{P(A \cap B \cap C)}{P(A \cap B)} = P(A \cap B \cap C).$$

   (e) $P(A|B) > P(A) \Rightarrow P(A \cap B)/P(B) > P(A) \Rightarrow P(A \cap B)/P(A) > P(B)$, i.e. $P(B|A) > P(B)$.

2. Let H_j denote the event that a head on the jth toss. Let F, BH and BT denote respectively the events that the chosen coin is one of the fair ones, one of those that always land showing a head and one of those that always land showing a tail.

(a)

$$\begin{aligned} P(H_1) &= P(H_1|F)P(F) + P(H_1|BH)P(BH) + P(H_1|BT)P(BT) \\ &= \frac{1}{2} \times \frac{2}{7} + 1 \times \frac{2}{7} + 0 \times \frac{3}{7} = \frac{3}{7} \end{aligned}$$

and

$$E(X) = 1 \cdot P(H_1) + 0 \cdot (1 - P(H_1)) = \frac{3}{7}$$

(b)

$$P(F|H_1) = \frac{P(H_1|F)P(F)}{P(H_1)} = \frac{1/7}{3/7} = \frac{1}{3}$$

(c)

$$\begin{aligned} P(H_2|H_1) &= P(H_2|H_1 \cap F)P(F|H_1) + P(H_2|H_1 \cap BH)P(BH|H_1) \\ &= \frac{1}{2} \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{5}{6} \end{aligned}$$

or

$$\begin{aligned} P(H_2|H_1) &= \frac{P(H_1 \cap H_2)}{P(H_1)} \\ &= \frac{P(H_1 \cap H_2|F)P(F) + P(H_1 \cap H_2|BH)P(BH)}{P(H_1)} \\ &= \frac{(1/4) \cdot (2/7) + 1 \cdot (2/7)}{(3/7)} = \frac{5}{6} \end{aligned}$$

(d)

$$\begin{aligned} P(F|H_1 \cap H_2) &= \frac{P(H_1 \cap H_2|F)P(F)}{P(H_1 \cap H_2|BH)P(BH) + P(H_1 \cap H_2|F)P(F)} \\ &= \frac{(1/4) \times (2/7)}{1 \times (2/7) + (1/4) \times (2/7)} = \frac{1}{5} \end{aligned}$$

3. (a) Let X be the bulb lifetime; $X \sim Exp(1/1000)$. Then

$$P(X > 100) = e^{-100/1000} = e^{-0.1} = 0.905 \quad \text{(3 d.p.)}$$

(b) Let Y be the number of bulbs in a box with lifetimes in excess of 100 hours; then $Y \sim Bin(2, 0.905)$. We require

$$P(Y = 2) = (0.905)^2 = 0.819.$$

So, 81.9% of boxes meet the guarantee.

(c) (i) For each t, the number of bulbs burnt out by time t is a Poisson random variable with parameter $t/1000$.

(ii) Let N be number of bulbs burnt out within t hours; $N \sim Pois(t/1000)$. We require

$$\begin{aligned} P(N \geq 4) &= 1 - P(N \leq 3) \\ &= 1 - \sum_{k=0}^{3} \frac{e^{-t/1000}(t/1000)^k}{k!} \end{aligned}$$

(d) $Z_2 \sim \Gamma(4, 0.001)$. So, the required probability density function is

$$f(z) = \begin{cases} (0.001)^4 z^3 e^{-0.004z}/\Gamma(4) & z \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and the mean is $E(Z_2) = 4/0.001 = 4000$ hours.

4. (a) Possible values of S_n are $0, 1, \ldots, n$. For $s = 0, 1, \ldots, n$

$$P(S_n = s) = \binom{n}{s} p^s (1-p)^{n-s}$$

i.e. $S_n \sim Bin(n, p)$. Hence, $E(S_n) = np$ and $Var(S_n) = np(1-p)$.

(b)

$$\begin{aligned} P(S_{100} \leq 2) &= P(S_{100} = 0) + P(S_{100} = 1) + P(S_{100} = 2) \\ &= (1-p)^{100} + 100p(1-p)^{99} + \frac{100 \cdot 99}{2} p^2 (1-p)^{98} \\ &= (1-p)^{98} \left\{ (1-p)^2 + 100p(1-p) + \frac{100 \cdot 99}{2} p^2 \right\} \\ &= (0.96)^{98} \left\{ (0.96)^2 + 100(0.04)(0.96) + 50 \cdot 99 (0.04)^2 \right\} \\ &= 0.01830547 \{0.9216 + 3.84 + 7.92\} \\ &= 0.2321426 \end{aligned}$$

(c) (i) Since $n = 100$ is large and $p = 0.04$ is small, $Bin(100, 0.04) \approx Pois(4)$ and

$$\begin{aligned} P(S_{100} \leq 2) &\approx e^{-4} + e^{-4} 4 + e^{-4} 4^2/2 \\ &= e^{-4}(1 + 4 + 8) = 0.2381033 \end{aligned}$$

(ii) By CTL, $S_{100}/100 \approx N(0.04, (0.04)(0.96)/100)$

$$\begin{aligned} P(S_{100} \leq 2) &= P\left(\frac{S_{100}}{100} \leq 0.02\right) \\ &\approx \Phi\left(\frac{0.025 - 0.04}{\sqrt{0.04(0.96)/100}}\right) \quad \text{(continuity correction)} \\ &= \Phi(-0.7654655) = 0.2219972. \end{aligned}$$

(d) The Poisson approximation is better since $n = 100$ is large and 0.04 is small. The Normal approximation is not as good since 0.04 is small.

5. (a) $X_1, \ldots, X_n \sim Exp(1/\alpha)$ are independent, so

$$E(Z) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n}(1/\alpha)^{-1} = n\alpha$$

$$Var(Z) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n}(1/\alpha)^{-2} = n\alpha^2$$

(b) For $y > 0$,

$$\begin{aligned} P(Y > y) &= P(X_i > y, i = 1, \ldots, n) \\ &= \prod_{i=1}^{n} P(X_i > y) \quad \text{(by independence)} \\ &= \left[\int_y^\infty \frac{1}{\alpha} e^{-x/\alpha} dx\right]^n = \left[e^{-y/\alpha}\right]^n \\ &= e^{-ny/\alpha} \end{aligned}$$

Hence, $Y \sim Exp(n/\alpha)$ and $E(Y) = \alpha/n$, $Var(Y) = (\alpha/n)^2$.

(c) Since $E(AZ) = An\alpha$, AZ is unbiased for α if $An = 1$ i.e. $A = 1/n$.
Since $E(BY) = B\alpha/n$, BY is unbiased for α if $B/n = 1$ i.e. $B = n$.

(d)

$$Var(Z/n) = Var(Z)/n^2 = (n\alpha^2)/n^2 = \alpha^2/n$$

$$Var(nY) = n^2 Var(Y) = n^2(\alpha/n)^2 = \alpha^2$$

We prefer Z/n as an estimator of α; it has a smaller variance than nY for all $n > 1$, while both estimators are unbiased. (In fact, Z/n is a consistent estimator of α and nY is not.)

6. Let Player A's scores be $X_1, \ldots, X_{10} \sim N(\mu_1, \sigma_1^2)$ independent. Let Player B's scores be $Y_1, \ldots, Y_{10} \sim N(\mu_2, \sigma_2^2)$ independent.

(a) For testing $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 \neq \sigma_2^2$, use the test statistic

$$F = \frac{S_1^2/(n_1 - 1)}{S_2^2/(n_2 - 1)} = \frac{17312/9}{13208/9} = 1.310721$$

H_0 would not be rejected, at the 5% level, for all values of F between the lower and upper 2.5% of a $F_{9,9}$ distribution, i.e. for all values in the range (0.2483859, 4.025994). Since the observed value of F is 1.310721, do not reject H_0 at the 5% level.

(b) $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$ with $\sigma_1^2 = \sigma_2^2$. The pooled sample variance is

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{9 \cdot 17312 + 9 \cdot 13208}{10 + 10 - 2} = 15260$$

the test statistic is

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{1/n_1 + 1/n_2}} = \frac{375 - 360}{\sqrt{15260(1/10 + 1/10)}} = 0.2715182$$

H_0 would not be rejected, at the 5% level, for all values of T between the lower and upper 2.5% of a t distribution with 18 degrees of freedom, i.e. for all values in the range (-2.101,2.101). Since the observed value of T is 0.2715182 and falls in the the range (-2.101,2.101), do not reject H_0 at the 5% level.

(c) A 95% confidence interval for $\mu_1 - \mu_2$ is

$$(375 - 360) \pm 2.101\sqrt{15260(1/10 + 1/10)} = (-101.0696, 131.0696)$$