UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2008/2009

April/May 2009

MAT 263 – Probability Theory
[Teori Kebarangkalian]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of FOURTEEN pages of printed materials before you begin the examination.

[Sila pastikan bahawa kertas peperiksaan ini mengandungi EMPAT BELAS muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]

Instructions: Answer **all four** [4] questions.

*[**Arahan**: Jawab **semua empat** [4] soalan.]*

1. (a) For each of the following, identify whether it can be distribution functions. Justify your answer.

(i) $F(x) = e^{-x}, x \geq 0$

(ii) $F(x) = 1 - \frac{1}{x}, x \geq 1$

(iii) $F(x) = x^2, x \geq 0$

(iv) $F(x) = \frac{1}{x^2}, x > 1$

[20 marks]

(b) The p.d.f. for a random variable X is given by

$$f(x) = \begin{cases} \frac{1}{4}x + \frac{1}{2} & \text{for } x = -2, -1, 0 \\ -\frac{1}{4}x + \frac{1}{2} & \text{for } x = 1, 2 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Obtain the distribution function of X.
(ii) Find the moment generating function of X.
(iii) Using (ii), find the p.d.f. of $Y = 3X - 2$.

[30 marks]

(c) The distribution function for a random variable X is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x^2}{c} & \text{if } 0 \leq x < 2 \\ x - \frac{x^2}{c} - 1 & \text{if } 2 \leq x < 4 \\ 1 & \text{if } x \geq 4 \end{cases}$$

(i) Find the value of c.
(ii) Using $F(x)$, determine $P(1 < X \leq 3)$.
(iii) Obtain the moment generating function of X.
(iv) Using (iii), find the mean and variance of X.

[35 marks]

(d) Let X and Y be two random variables with means, variances and correlation coefficients denoted by $\mu_X, \mu_Y, {\sigma_X}^2, {\sigma_Y}^2$ and ρ_{XY}, respectively.

Show that $E(Y \mid x) = \mu_Y + \rho_{XY} \frac{{\sigma_X}^2}{{\sigma_Y}^2}(x - \mu_X)$.

[15 marks]

1. (a) *Bagi setiap yang berikut, nyatakan sama ada ianya suatu fungsi taburan. Tentusahkan jawapan anda.*

(i) $F(x) = e^{-x}, x \geq 0$

(ii) $F(x) = 1 - \frac{1}{x}, x \geq 1$

(iii) $F(x) = x^2, x \geq 0$

(iv) $F(x) = \frac{1}{x^2}, x > 1$

[20 markah]

(b) *F.k.k bagi suatu pembolehubah rawak X diberi oleh*

$$f(x) = \begin{cases} \frac{1}{4}x + \frac{1}{2} & \text{bagi } x = -2, -1, 0 \\ -\frac{1}{4}x + \frac{1}{2} & \text{bagi } x = 1, 2 \\ 0 & \text{di sebaliknya.} \end{cases}$$

(i) *Dapatkan fungsi taburan bagi X.*

(ii) *Cari fungsi penjana momen bagi X.*

(iii) *Menggunakan (ii), cari f.k.k. bagi $Y = 3X - 2$.*

[30 markah]

(c) *Fungsi taburan bagi suatu pembolehubah rawak X diberi oleh*

$$F(x) = \begin{cases} 0 & \text{jika } x < 0 \\ \frac{x^2}{c} & \text{jika } 0 \text{ £ } x < 2 \\ x - \frac{x^2}{c} - 1 & \text{jika } 2 \text{ £ } x < 4 \\ 1 & \text{jika } x \text{ ³ } 4 \end{cases}$$

(i) *Cari nilai c.*

(ii) *Menggunakan $F(x)$, tentukan $P(1 < X \leq 3)$.*

(iii) *Dapatkan fungsi penjana momen bagi X.*

(iv) *Menggunakan (iii), cari min dan varians bagi X.*

[35 markah]

(d) *Katakan X dan Y adalah dua pembolehubah rawak dengan min, varians dan pekali korelasi yang ditandai $\mu_X, \mu_Y, \sigma_X{}^2, \sigma_Y{}^2$ dan ρ_{XY}, masing-masing. Tunjukkan bahawa $E(Y \mid x) = \mu_Y + \rho_{XY} \frac{\sigma_X{}^2}{\sigma_Y{}^2}(x - \mu_X)$.*

[15 markah]

2. (a) A university administers two tests to each new student to better predict student success in a certain course. Let L_1 be the event that a student passes the first test and let L_2 be the event that a student passes the second test. Let Q be the event that a student is qualified for that course. Given that

$$\alpha_1 = P(L_1 \mid Q), \qquad \alpha_2 = P(L_2 \mid Q),$$
$$\beta_1 = P(L_1^c \mid Q^c), \qquad \beta_2 = P(L_2^c \mid Q^c),$$
$$\omega = P(Q).$$

(i) Show that
$$P(Q^c \mid L_i) = \frac{(1-\beta_i)(1-\omega)}{\alpha_i \omega + (1-\beta_i)(1-\omega)}.$$

(ii) What must be true of $P(Q^c \mid L_1)$ and $P(Q^c \mid L_2)$ if the first test is considered better for screening unqualified applicants?

(iii) Using your answer in (ii), show that $(1-\beta_1)\alpha_2 < (1-\beta_2)\alpha_1$.

[30 marks]

(b) The folowing contingency table is the joint p.d.f. of the two independent random variables X and Y.

		Y: 1	2	3	4	
X	1	0.24			0.12	
	2					0.40
			0.30			

(i) Complete the table.

(ii) Obtain the distribution function of X and Y.

(iii) Find $E(Y \mid x)$.

[30 marks]

(c) The random variables X and Y are independent and identically distributed with common density

$$f_X(x) = e^{-x}, \quad x > 0.$$

(i) Determine the distribution function for the random variable $W = X + Y$.

(ii) Obtain the p.d.f. of $Z = W^{1/2}$.

[40 marks]

2. (a) *Pihak universiti menjalankan dua ujian ke atas setiap pelajar baru untuk meramal prestasi pelajar dalam suatu kursus. Katakan L_1 adalah peristiwa bahawa seseorang pelajar lulus ujian pertama dan L_2 adalah peristiwa bahawa seseorang pelajar lulus ujian kedua. Katakan Q adalah peristiwa bahawa seseorang pelajar layak bagi kursus tersebut. Diberi*

$$\alpha_1 = P(L_1 \mid Q), \qquad \alpha_2 = P(L_2 \mid Q),$$
$$\beta_1 = P(L_1^c \mid Q^c), \qquad \beta_2 = P(L_2^c \mid Q^c),$$
$$\omega = P(Q).$$

(i) *Tunjukkan bahawa*

$$P(Q^c \mid L_i) = \frac{(1-\beta_i)(1-\omega)}{\alpha_i\omega + (1-\beta_i)(1-\omega)}.$$

(ii) *Apakah yang mesti benar tentang $P(Q^c \mid L_1)$ dan $P(Q^c \mid L_2)$ jika ujian pertama adalah lebih baik bagi menyaring calun tidak layak?*

(iii) *Berdasarkan jawapan anda di (ii), tunjukkan bahawa* $(1-\beta_1)\alpha_2 < (1-\beta_2)\alpha_1$.

[30 markah]

(b) *Jadual kontinjensi berikut ialah f.k.k tercantum pembolehubah rawak X dan Y.*

		Y: 1	2	3	4	
X	1	0.24			0.12	
	2					0.40
			0.30			

(i) *Lengkapkan jadual.*

(ii) *Dapatkan fungsi taburan bagi X and Y.*

(iii) *Cari $E(Y \mid x)$.*

[30 markah]

(c) *Pembolehubah rawak X dan Y adalah tertabur secaman dan tak bersandar dengan ketumpatan sepunya*

$$f_X(x) = e^{-x}, \quad x > 0.$$

(i) *Tentukan fungsi taburan bagi pembolehubah rawak $W = X + Y$.*

(ii) *Dapatkan f.k.. bagi $Z = W^{1/2}$.*

[40 markah]

3. (a) For each of the following random variables X, write out the complete p.d.f. that best models X. Give the values of the parameters involved.

(i) There are 2000 students enrolled in a certain course. The exam papers in this course are graded by a team of teaching assistants; however, a sample of the papers is examined by the course professor for grading consistency. The professor selects 10 papers at random from the 2000 submitted and examines them for grading inconsistencies. Let X be the number of papers in the sample that are improperly graded.

(ii) Calls to an emergency center are placed independently and at random. During evening hours, the center receives an average of 100 calls an hour. Let X be the number of calls received between 7.00 p.m. and 7.30 p.m tonight.

(iii) Calls to an emergency center are placed independently and at random. It is currently 2 a.m. and the center's line are all open and ten calls are expected over the next half hour. Let X be the time in minutes until the next call is received.

(iv) A researcher is studying blood disorders exhibited by people with rare blood types. It is estimated that 10% of the population has the type of blood being investigated. Volunteers whose blood type is unknown are tested until 100 people with desired blood type are found. Let X be the number of people tested who do have the desired rare blood type.

[20 marks]

3. (a) *Bagi setiap pembolehubah rawak X berikut, tuliskan f.k.k. lengkap yang terbaik untuk memodel X. Beri nilai parameter terlibat.*

(i) *Terdapat 2000 pelajar mendaftar dalam suatu kursus. Kertas peperiksaan bagi kursus ini diperiksa oleh satu pasukan pembantu pengajar; walaubagaimana pun, suatu sampel rawak kertas di semak oleh profesor kursus bagi tujuan pemarkahan konsisten. Profesor ini memilih 10 kertas secara rawak dari 2000 yang dihantar dan memeriksanya bagi tujuan pemarkahan konsisten. Katakan X ialah bilangan kertas dalam sampel yang tidak diperiksa dengan betul.*

(ii) *Panggilan ke suatu pusat kecemasan tiba secara tak bersandar dan secara rawak. Pada waktu petang, pusat ini menerima 100 panggilan sejam. Katakan X ialah bilangan panggilan diterima antara 7.00 dan 7.30 malam ini.*

(iii) *Panggilan ke suatu pusat kecemasan tiba secara tak bersandar dan secara rawak. Waktu sekarang ialah 2 pagi dan semua talian ke pusat ini dibuka dan sepuluh panggilan dijangka masuk bagi setengah jam berikut. Katakan X ialah masa dalam minit sehingga panggilan berikutnya diterima.*

(iv) *Penyelidik membuat kajian tentang penyakit darah yang ditunjukkan oleh orang yang mempunyai jenis darah ganjil. Dianggarkan 10% dari populasi mempunyai jenis darah yang dikaji. Sukarelawan dengan jenis darah yang tidak diketahui diuji sehingga 100 orang dengan jenis darah yang dikehendaki diperolehi. Katakan X ialah bilangan orang diuji yang mempunyai jenis darah ganjil yang dikehendaki.*

[20 markah]

(b) A Ministry of Environment is concerned with the problem of setting criteria for the amount of certain toxic chemical to be allowed in rivers. A commom measure of toxicity for any pollutant is the concentration of the pollutant that will kill half of the test species in a given amount of time (usually 96 hours for fish species). This measure is called the LC50 (lethal concentration killing 50% of the test species). In many studies the $\ln(\text{LC50})$ measurements are normally distributed and hence the analysis is based on $\ln(\text{LC50})$ data. Studies on the effects of copper on a certain species of fish show the variance of $\ln(\text{LC50})$ measurements to be around 0.5 with concentration measurements in milligrams per liter.

(i) If it is desired that the sample mean differ from the population mean by no more than 0.5, with probability 0.90, how many tests should be run?

(ii) Suppose that the effects of copper on a second species of fish show the variance of $\ln(\text{LC50})$ measurements to be around 0.8. If the population means of $\ln(\text{LC50})$ for the two species are equal, find the probability that, with random samples of ten measurements from each species, the sample mean for the first species exceeds the sample mean for the second species by at least one unit.

(iii) If $n = 20$ observations are to be taken on $\ln(\text{LC50})$ measurements, with $\sigma^2 = 1.4$, find the two numbers a and b such that $P(a \le S^2 \le b) = 0.90$, where S^2 is the sample variance of the 20 measurements.

(iv) The study also look into the effects of lead on the fish. Let S_1^2 denote the sample variance for a random sample of ten $\ln(\text{LC50})$ values for copper and let S_2^2 denote the sample variance for a random sample of eight $\ln(\text{LC50})$ values for lead, both samples using the same species of fish. The population variance for measurements on copper is assumed to be twice the corresponding population variance for measurements on lead. Find two numbers c and d such that

$$P\left(c \le \frac{S_1^2}{S_2^2} \le d\right) = 0.90$$

[40 marks]

(c) Suppose that X and Y are random variables with joint density

$$f(x, y) = \begin{cases} 6(1-x-y) & \text{for } x+y \le 1, x \ge 0, y \ge 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Determine the marginal distribution function of X and Y.

(ii) Compute $P(X > \frac{1}{2} \mid Y = \frac{3}{4})$.

(iii) Calculate the expected value of $\dfrac{X^2+1}{X}$.

(iv) Determine the probability that $X > Y$.

[40 marks]

(b) *Kementerian Alam Sekitar sedang berfikir masalah penyediaan kriteria bagi amaun suatu bahan kimia beracun yang boleh dilepaskan ke dalam sungai. Suatu ukuran keracunan bagi sebarang pencemar ialah kepekatan pencemar yang akan membunuh setengah spesis ujian dalam amaun masa yang diberi (kebiasaannya 96 jam bagi setiap spesis ikan). Ukuran ini dipanggil LC50 (kepekatan lethal pembunuh 50% dari spesis ujian). Dalam banyak kajian ukuran* ln *(LC50) adalah tertabur normal dan dengan itu analisis adalah berdasarkan data* ln *(LC50). Kajian ke atas kesan tembaga terhadap suatu spesis tertentu ikan menunjukkan bahawa varians bagi ukuran* ln *(LC50) adalah sekitar 0.5 dengan ukuran kepekatan dalam milligrams per liter.*

(i) *Jika dikehendaki bahawa min sampel berbeza dari min populasi tidak lebih dari 0.5, dengan kebarangkalian 0.90, berapa banyakkah ujian yang perlu dijalankan?*

(ii) *Katakan kesan tembaga terhadap spesis kedua ikan menunjukkan bahawa varians bagi ukuran* ln *(LC50) adalah di sekitar 0.8. Jika min populasi* ln *(LC50) bagi kedua spesis adalah sama, cari kebarangkalian bahawa, dengan sampel rawak 10 ukuran bagi setiap spesis, min sampel bagi spesis pertama melebihi min sampel bagi spesis kedua sekurang-kurangnya satu unit.*

(iii) *Jika* $n = 20$ *ukuran* ln *(LC50) diambil, dengan* $\sigma^2 = 1.4$, *cari dua nombor* a *dan* b *supaya* $P(a \le S^2 \le b) = 0.90$, *yang mana* S^2 *adalah sampel varians bagi 20 ukuran.*

(iv) *Kajian juga melihat kesan plumbum ke atas ikan. Biar* S_1^2 *sebagai varians sampel suatu sampel rawak 10 nilai* ln *(LC50) bagi tembaga and* S_2^2 *sebagai varians sampel suatu sampel rawak lapan nilai* ln *(LC50) bagi plumbum, kedua-dua sampel menggunakan spesis ikan yang sama. Varians populasi bagi ukuran tembaga diandaikan dua kali ganda varians populasi bagi ukuran plumbum. Cari nombor* c *dan* d *supaya*

$$P(c \le \frac{S_1^2}{S_2^2} \le d) = 0.90$$

[40 markah]

(c) *Katakan* X *dan* Y *adalah pembolehubah rawak dengan ketumpatan tercantum*

$$f(x,y) = \begin{cases} 6(1-x-y) & bagi\ x+y \le 1, x \ge 0, y \ge 0 \\ 0 & di\ sebaliknya. \end{cases}$$

(i) *Tentukan fungsi taburan sut bagi* X *dan* Y.

(ii) *Hitung* $P(X > \frac{1}{2} \mid Y = \frac{3}{4})$.

(iii) *Kira nilai jangkaan bagi* $\dfrac{X^2+1}{X}$.

(iv) *Tentukan kebarangkalian bahawa* $X > Y$.

[40 markah]

4. (a) Show that if $(X \mid Y) \sim Gamma(r, \lambda)$ and $Y \sim Gamma(s, \beta)$, then $(Y \mid X = x) \sim Gamma(r + s, x + \beta)$.

[30 marks]

(b) (i) Show that if $Z \sim N(0,1)$, then $Z^2 \sim Gamma(1/2, 1/2)$.

(ii) Show that if Z_i's are independent $N(0,1)$, then $Y = \sum_{i=1}^{n} Z_i^2$ is $Gamma(n/2, 1/2)$.

[35 marks]

(c) If X and Y have joint density function

$$f(x, y) = \frac{1}{x^2 y^2} \quad x \geq 1, y \geq 1.$$

(i) Compute the joint density of $U = XY$, $V = X/Y$.
(ii) Obtain the marginal densities of U and V.
(iii) Are U and V independent? Justify your answer.

[35 marks]

4\. (a) *Tunjukkan bahawa, jika* $(X \mid Y) \sim Gamma(r, \lambda)$ *dan* $Y \sim Gamma(s, \beta)$, *maka* $(Y \mid X = x) \sim Gamma(r + s, x + \beta)$.

[30 markah]

(b) (i) *Tunjukkan bahawa, jika* $Z \sim N(0,1)$, *maka* $Z^2 \sim Gamma(1/2, 1/2)$.

(ii) *Tunjukkan bahawa, jika* Z_i *'s adalah tak bersandar* $N(0,1)$, *maka* $Y = \sum_{i=1}^{n} Z_i^2$ *adalah* $Gamma(n/2, 1/2)$.

[35 markah]

(c) *Jika* X *dan* Y *mempunyai fungsi ketumpatan tercantum*

$$f(x, y) = \frac{1}{x^2 y^2} \quad x \geq 1, y \geq 1.$$

(i) *Hitung ketumpatan tercantum bagi* $U = XY$, $V = X/Y$.

(ii) *Dapatkan ketumpatan sut bagi* U *dan* V.

(iii) *Adakah* U *dan* V *tak bersandar? Tentushakan jawapan anda.*

[35 markah]

APPENDIX/LAMPIRAN

DISCRETE DISTRIBUTIONS	
Bernoulli	$f(x)=p^x(1-p)^{1-x}, \quad x=0,1$ $M(t)=1-p+pe^t$ $\mu=p, \quad \sigma^2=p(1-p)$
Binomial	$f(x)=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}, \quad x=0,1,2,\ldots,n$ $M(t)=(1-p+pe^t)^n$ $\mu=p, \quad \sigma^2=np(1-p)$
Geometric	$f(x)=(1-p)^x p, \quad x=0,1,2,\ldots$ $M(t)=\frac{p}{1-(1-p)pe^t}, t<\text{In}(1-p)$ $\mu=\frac{1-p}{p}, \quad \sigma^2=\mu=\frac{(1-p)}{p^2}$
Negative Binomial	$f(x)=\frac{(x+r-p)!}{x!(r-1)!}p^r(1-p)^x, \quad x=0,1,2,\ldots$ $M(t)=\frac{p^r}{\left[1-(1-p)e^t\right]^r}, t<-\text{In}(1-p)$ $\mu=\frac{r(1-p)}{p}, \quad \sigma^2=\frac{r(1-p)}{p^2}$
Poisson	$f(x)=\frac{\lambda^x e^{-\lambda}}{x!}, \quad x=0,1,2,\ldots$ $M(t)=e^{\lambda(e^t-1)}$ $\mu=\lambda, \quad \sigma^2=\lambda$
Hipergeometric	$f(x)=\frac{\binom{n_1}{x}\binom{n_2}{r-x}}{\binom{n}{t}}, \quad x\le r, x\le n_1, r-x\le n_2,$ $\mu=\frac{rn_1}{n}, \quad \sigma^2=\frac{rn_1n_2(n-r)}{n^2(n-1)}$

CONTINUOUS DISTRIBUTION	
Uniform	$f(x)=\frac{1}{b-a},\quad a\le x\le b$ $M(t)=\frac{e^{tb}-e^{ta}}{t(b-a)},\quad t\ne 0,\quad M(0)=1$ $\mu=\frac{a+b}{2},\quad \sigma^2=\frac{(b-a)^2}{12}$
Exponential	$f(x)=\frac{1}{\theta}e^{-x/\theta},\ 0\le x\le\infty$ $M(t)=\frac{1}{1-\theta t},\quad t<1/\theta$ $\mu=\theta,\quad \sigma^2=\theta^2$
Gamma	$f(x)=\frac{1}{\Gamma(\alpha)\theta^{\alpha}}x^{\alpha-1}e^{-x/\theta},0\le x\le\infty$ $M(t)=\frac{1}{(1-\theta t)^{\alpha}},\quad t<1/0$ $\mu=\alpha\theta,\quad \sigma^2=\alpha\theta^2$
Chi Square	$f(x)=\frac{1}{\Gamma(r/2)2^{r/2}}r^{r/2-1}e^{-x/2},\quad 0\le x\le\infty$ $M(t)=\frac{1}{(1-2t)^{r/2}},\quad t<\frac{1}{2}$ $\mu=r,\quad \sigma^2=2r$
Normal	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-[(x-\mu)^2/2\sigma^2]},-\infty<x<\infty$ $M(t)=e^{\mu t+\sigma^2t^2/2}$ $E(X)=\mu,\quad Var(X)=\sigma^2$
Beta	$f(x)=\frac{1}{B(\alpha,\beta)}\quad x^{\alpha-1}(1-x)^{\beta-1},0<x<1$ $\mu=\frac{\alpha}{\alpha+\beta},\quad \sigma^2=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

FORMULA	
1.	$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$
2.	$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}, \quad \lvert r \rvert < 1$
3.	$\sum_{x=0}^{n} \binom{n}{x} b^x a^{n-x} = (a+b)^n$
4.	$\sum_{x=0}^{n} \binom{n}{x} \binom{r-n}{r-x} = \binom{n}{r}$
5.	$\sum_{x=0}^{n} \binom{n+k-1}{k} w^k = (1-w)^{-n}, \quad \lvert w \rvert < 1$
6.	$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad \Gamma(\alpha) = (\alpha-1)!$
7.	$B(\alpha,\beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1}\, dx$
8.	$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

- 000 O 000 -