UNIVERSITI SAINS MALAYSIA

First Semester Examination
2011/2012 Academic Session

January 2012

MST 561 – Statistical Inference
[Pentaabiran Statistik]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of EIGHT pages of printed material before you begin the examination.

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

Instructions: Answer **all five** [5] questions.

*[**Arahan:** Jawab **semua lima** [5] soalan.]*

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai].

1. (a) A box consists of 20 cards. The cards are written with different numbers. The numeral 2 is written on 8 cards, the numeral 4 is written on 10 cards and the numeral 6 is written on the remaining cards. The cards are mixed and we select one card randomly from the box. Let X be the number written on the selected card.

(i) Derive the probability mass function $f(x)$ of X

(ii) Derive the distribution function $F(x)$ of X

(iii) Plot the distribution function $F(x)$ and mention its points of discontinuities

(iv) Find $P(2 < X \leq 4)$

(b) Let X and Y be independent random variables, each with probability density function:

$$f(x) = \frac{x}{3} e^{-x/3} \quad 0 < x < \infty$$

$$f(y) = \frac{1}{3} e^{-y/3} \quad 0 < y < \infty$$

(i) Find the joint probability density function of X and Y, i.e. $h_{X,Y}(x, y)$

(ii) Find the joint probability density function of $z=x/y$ and $w=y$, i.e. $h_{Z,W}(z, w)$

[20 marks]

1. (a) Sebuah kotak mengandungi 20 kad. Kad ditulis dengan nombor yang berlainan. Angka 2 ditulis pada 8 kad, angka 4 ditulis pada 10 kad dan angka 6 ditulis pada kad selebihnya. Kad-kad tersebut dicampur dan kita memilih satu kad secara rawak daripada kotak. Katakan X adalah nombor yang ditulis pada kad yang dipilih.

(i) Terbitkan fungsi jisim kebarangkalian f (x) untuk X

(ii) Terbitkan fungsi taburan F(x) untuk X

(iii) Plotkan fungsi taburan F(x) dan nyatakan titik ketidakselanjaran

(iv) Cari $P(2 < X \leq 4)$

(b) Biarkan X dan Y sebagai pembolehubah rawak tak bersandar, setiap dengan fungsi kebarangkalian ketumpatan:

$$f(x) = \frac{x}{3} e^{-x/3} \quad 0 < x < \infty$$

$$f(y) = \frac{1}{3} e^{-y/3} \quad 0 < y < \infty$$

(i) Cari fungsi ketumpatan kebarangkalian tercantum X dan Y iaitu $h_{X,Y}(x, y)$

(ii) Cari fungsi ketumpatan kebarangkalian tercantum z=x/y dan w=y, iaitu $h_{Z,W}(z, w)$

[20 markah]

2. (a) We seek to measure the performance of the weather forecast system, i.e. how accurate the forecast of raining is. Let R denotes the event of rain, and $P(R) = 0.2$

There are two types of forecast inaccuracies. The first type of error is the weather forecast shows raining but it does not, $P(Y|R') = 0.05$. The second type is it is forecasted to be not raining but it does rain, $P(N|R) = 0.02$

Let Y denotes the forecast of raining and N the forecast with no raining.

(i) Calculate $P(Y)$

(ii) Calculate $P(R|Y)$

(iii) What can you say about the accuracy of the forecast?

(b) Let X indicates the score of students in a statistics course as a random variable with mean 72 and variance 24.

(i) What can be said about the probability that a student will score between 62 and 82?

(ii) Find the upper bound for the probability that a student's test score will exceed 80.

[20 marks]

2. (a) Kita ingin mengukur prestasi sistem ramalan cuaca, iaitu bagaimana tepat ramalan hujan.

Katakan R menandakan peristiwa hujan, dan $P(R) = 0.2$

Terdapat dua jenis ketidaktepatan ramalan. Jenis pertama adalah ramalan cuaca menunjukkan hujan tetapi ia tidak hujan, $P(Y|R') = 0.05$. *Jenis kedua adalah ia dijangka tidak akan hujan tetapi ia hujan,* $P(N|R) = 0.02$

Katakan Y menandakan ramalan hujan dan N ramalan dengan tiada hujan.

(i) Kirakan P(Y)

(ii) Kirakan $P(R|Y)$

(iii) Apakah yang boleh anda katakan tentang ketepatan ramalan?

(b) Andaikan X menunjukkan markah pelajar dalam kursus statistik, sebagai pembolehubah rawak dengan min 72 dan varians 24.

(i) Apakah yang boleh dikatakan dengan kebarangkalian bahawa seorang pelajar akan mendapat markah antara 62 dan 82?

(ii) Cari had atas untuk kebarangkalian bahawa markah seseorang pelajar akan melebihi 80.

[20 markah]

3. (a) Assume that the road accidents in a town follows a Poisson random variable with mean λ. The values of λ differs between car drivers and motorcyclist. λ

takes the value of 1.5 for car driver and 3 for motorcyclist. Assume that there are 60% of car driver and 40% of motorcyclist in that town. If a driver is chosen randomly,

(i) What is the probability that he will have no accident?

(ii) What is the conditional probability that he will have 2 accidents given that he had no accident in the preceding years?

(b) Assume that *X* is the weight of boxes in milligram in which it follows a normal distribution $N(\mu, 0.03)$. Let $X_1, X_2 ..., X_n$ be the random sample from this distribution. A hypothesis test is conducted to test $H_0 : \mu \geq 10.35$ versus $H_1 : \mu < 10.35$. What is the critical region of size $\alpha = 0.05$ specified by the likelihood ratio test criterion?

(c) Let *X* be a binomial random variable with parameters *n* and *p*. Show that $E\left[\frac{1}{X+1}\right] = \frac{1-(1-p)^{n+1}}{(n+1)p}$

[20 marks]

3. (a) *Andaikan bahawa kemalangan jalan raya di bandar tertentu merupakan pembolehubah rawak Poisson dengan min* λ. *Nilai-* λ *berbeza antara pemandu kereta dan penunggang motosikal. Nilai* λ *bagi pemandu kereta ialah 1.5 dan bagi penunggang motosikal ialah 3. Andaikan bahawa terdapat 60% pemandu kereta dan 40% penunggang motosikal di bandar itu. Jika pemandu dipilih secara rawak,*

(i) Apakah kebarangkalian bahawa dia tidak akan mempunyai kemalangan?

(ii) Apakah kebarangkalian bersyarat bahawa dia akan mempunyai 2 kemalangan, diberikan bahawa dia tidak mengalami kemalangan pada tahun sebelumnya?

(b) *Andaikan bahawa X merupakan berat kotak dalam milligram di mana ia mengikuti taburan normal* $N(\mu, 0.03)$. *Biarkan* $X_1, X_2 ..., X_n$ *merupakan sampel rawak daripada taburan ini. Satu ujian hipotesis dilakukan untuk menguji* $H_0 : \mu \geq 10.35$ *lawan* $H_1 : \mu < 10.35$. *Apakah rantau genting bagi saiz* $\alpha = 0.05$ *yang dispesifikasikan oleh kriteria ujian kebolehjadian?*

(c) *Biarkan X merupakan pembolehubah rawak binomial dengan parameter n dan p. Tunjukkan bahawa* $E\left[\frac{1}{X+1}\right] = \frac{1-(1-p)^{n+1}}{(n+1)p}$.

[20 markah]

4. (a) Assume that the probability mass function of X is given by $f(x|\theta)=\left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|}$ $,x=-1,0,1\ ,0\le\theta\le 1$

(i) Is X a complete sufficient statistic?

(ii) Is $|X|$ a complete sufficient statistic?

(b) The moment generating function of X is given by $M_x(t)=\exp\ 2e^{t-1}$ and that of Y by $M_y(t)=(\frac{3}{4}e^t+\frac{1}{4})^{10}$. If X and Y are independent, what is

(i) $P\ \ XY=0$

(ii) $E\ \ XY$

(c) Let X be a random sample of size n=6 from a Poisson distribution with mean λ. We reject $H_0:\lambda=0.5$ and accept $H_1:\lambda>0.5$ if the observed sum $\sum_{i=1}^{6}x_i\ge 6$. Get the significance level α of the test.

[20 marks]

4. *(a) Andaikan bahawa fungsi jisim kebarangkalian X diberikan oleh* $f(x|\theta)=\left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|}$
$,x=-1,0,1\ ,0\le\theta\le 1$

(i) Adakah X statistik yang cukup dan lengkap?

(ii) Adakah $|X|$ *statistik yang cukup dan lengkap?*

(b) Fungsi penjana momen untuk X diberi oleh $M_x(t)=\exp\ 2e^{t-1}$ *dan Y oleh* $M_y(t)=(\frac{3}{4}e^t+\frac{1}{4})^{10}$ *. Jika X dan Y tidak bersandar, apakah*

(i) $P\ \ XY=0$

(ii) $E\ \ XY$

(c) Biarkan X merupakan satu sampel rawak dengan saiz n=6 daripada taburan Poisson dengan min λ*. Kita menolak* $H_0:\lambda=0.5$ *dan menerima* $H_1:\lambda>0.5$ *jika jumlah yang diperhatikan ialah* $\sum_{i=1}^{6}x_i\ge 6$*. Dapatkan aras keertian* α *bagi ujian ini.*

[20 markah]

5. (a) Let $X_{1,\dots\dots,}X_n$ be a random sample from a $N(M,\sigma^2)$ population and σ^2 is unknown. Find the minimum value for n so that a 95% confidence interval for μ will have a length no more than $\frac{\sigma}{2}$.

(b) Let $X_{1,\dots\dots,}X_n$ be iid Bernouli (p). Show that the variance of $\bar{X}$ obtains the Cramér-Rao Lower Bound and hence $\bar{X}$ is the best unbiased estimator of p.

(c) One observation is taken from a normally distributed $N\ \ 0,\sigma^2$ population.

(i) Find an unbiased estimator of σ^2

(ii) Find the maximum likelihood estimator of σ

[20 marks]

5. *(a) Biarkan $X_{1,\dots\dots,}X_n$ sebagai sampel rawak dari populasi $N(M,\sigma^2)$ dan σ^2 tidak diketahui. Cari nilai minimum bagi n supaya selang keyakinan 95% bagi μ akan mempunyai panjang yang tidak lebih daripada $\frac{\sigma}{2}$.*

(b) Biarkan $X_{1,\dots\dots,}X_n$ bertaburan secaman dan tak bersandar Bernouli (p). Tunjukkan bahawa varians bagi $\bar{X}$ mencapai batas bawah Cramér-Rao dan oleh itu $\bar{X}$ adalah penganggar saksama terbaik bagi p.

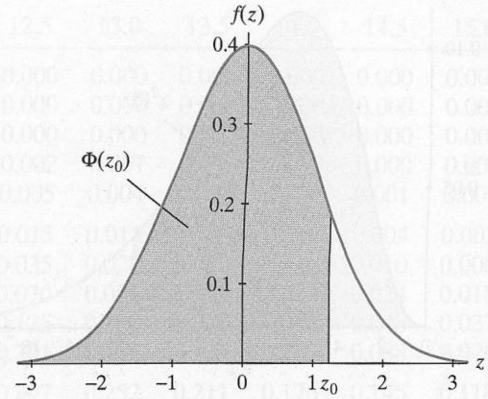
(c) Satu cerapan diambil dari populasi bertaburan normal $N\ \ 0,\sigma^2$.

(i) Cari penganggar saksama bagi σ^2

(ii) Cari penganggar kebolehjadian maksimum bagi σ

[20 markah]

The Normal Distribution



$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

α	0.400	0.300	0.200	0.100	0.050	0.025	0.020	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.054	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.326	2.576	2.807	3.291

Table of Distributions

Distribution	PDF or probability function	mean	variance	MGF
Point mass at a	$I(x=a)$	a	0	e^{at}
Bernoulli(p)	$p^x(1-p)^{1-x}$	p	$p(1-p)$	$pe^t+(1-p)$
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$	$(pe^t+(1-p))^n$
Geometric(p)	$p(1-p)^{x-1}I(x\geq 1)$	$1/p$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$ $(t<-\log(1-p))$
Poisson(λ)	$\frac{\lambda^x e^{-\lambda}}{x!}$	λ	λ	$e^{\lambda(e^t-1)}$
Uniform(a,b)	$I(a<x<b)/(b-a)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Normal(μ,σ^2)	$\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2	$\exp\left\{\mu t+\frac{\sigma^2t^2}{2}\right\}$
Exponential(β)	$\frac{e^{-x/\beta}}{\beta}$	β	β^2	$\frac{1}{1-\beta t}$ $(t<1/\beta)$
Gamma(α,β)	$\frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$	$\alpha\beta$	$\alpha\beta^2$	$\left(\frac{1}{1-\beta t}\right)^\alpha$ $(t<1/\beta)$
Beta(α,β)	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
t_ν	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\left(1+\frac{x^2}{\nu}\right)^{(\nu+1)/2}}$	0 (if $\nu>1$)	$\frac{\nu}{\nu-2}$ (if $\nu>2$)	does not exist
χ^2_p	$\frac{1}{\Gamma(p/2)2^{p/2}}x^{(p/2)-1}e^{-x/2}$	p	$2p$	$\left(\frac{1}{1-2t}\right)^{p/2}$ $(t<1/2)$