UNIVERSITI SAINS MALAYSIA

First Semester Examination
2011/2012 Academic Session

January 2012

MST 564 – Statistical Reliability
[Kebolehpercayaan Statistik]

Duration : 3 hours
[Masa : 3 jam]

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

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

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

*[**Arahan:** Jawab **semua empat** [4] 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) The ordinary language usage for Reliability is absence of failures. But the definitions for Reliability are far more detailed and far more complex than this. What are these definitions?

(b) The graph of Reliability (Risk of Failure) vs Time or the BATHTUB CURVE is often used to represent what happens in a typical process. Do you think this is a realistic representation? If yes, why is it so? If no, why not?

(c) Four **W**s - **W**hat, **W**hy, **W**hen, **W**here – can be used to elucidate reliability data. Write short notes on

(i) What is reliability data?
(ii) Why use reliability data?
(iii) When to use reliability data?
(iv) Where to use reliability data?

[100 marks]

1. (a) Penggunaan bahasa biasa untuk Kebolehpercayaan adalah ketiadaan kegagalan. Tetapi takrif untuk Kebolehpercayaan adalah jauh lebih terperinci dan jauh lebih kompleks daripada ini. Apakah definisi ini?

(b) Graf Kebolehpercayaan (Risiko Kegagalan) berlawan Masa atau LENGKUNGAN TAB MANDI sering digunakan untuk mewakili apa yang berlaku dalam proses yang tipikal. Adakah anda fikir ini adalah perwakilan yang realistik? Jika ya, mengapa begitu? Jika tidak, mengapa tidak?

(c) Empat W - Apa, Mengapa, Bila, Di mana - boleh digunakan untuk menjelaskan data kebolehpercayaan. Tuliskan nota ringkas tentang

(i) Apakah data kebolehpercayaan?
(ii) Mengapakah menggunakan data kebolehpercayaan?
(iii) Bilakah menggunakan data kebolehpercayaan?
(iv) Di manakah menggunakan data kebolehpercayaan?

[100 markah]

2. (a) The three-parameter Weibull distribution has $\beta = 4\,\text{hr}$, $\alpha = 780$, and $t_0 = 100$ hr. Compute its

(i) R(500 hr)
(ii) MTTF
(iii) t_{median}
(iv) design life for a desired 95 percent reliability.

(b) A company recorded the failure or censored times (in hours) for 20 machine parts from a single manufacturer (see **Table 1**). Censored times (indicated by '+') resulted from parts that were no longer used because of design or manufacturing changes.

Table 1. Failure times (in hours) for 20 machine parts

16+	117	261+	323+	518
531+	643+	758	824	881
1323	1582	1795	1854+	2556
2829	5002	6298	13,991+	18,789

These parts were advertised by the manufacturer as having a 500-hr design life (with a 90 percent reliability). The company wants to determine whether the advertised design life is being met before it decides to order additional parts from this manufacturer.

(i) Conduct a least squares analysis using the exponential, Weibull, normal, and lognormal distributions.
(ii) What is the best fit from the results obtained in (i)?
(iii) Find $R(t)$ and a 90 percent design life if Maximum Likelihood Estimation (MLE) is used for the best fit reliability model.
(iv) What is the company's decision based on the best fit reliability model?

[100 marks]

2. (a) Taburan Weibull tiga parameter mempunyai $\beta = 4\,\text{jam}$, $\alpha = 780$, and $t_0 = 100$ jam. Kira

(i) R(500 jam)
(ii) MTTF
(iii) t_{median}
(iv) reka bentuk hayat bagi kebolehpercayaan 95 peratus yang dikehendaki.

*(b) Sebuah syarikat mencatatkan kegagalan atau masa tertapis (dalam jam) bagi 20 bahagian mesin daripada pengilang tunggal (lihat **Jadual 1**). Masa tertapis (ditunjukkan oleh '+') disebabkan oleh bahagian yang tidak lagi digunakan kerana perubahan reka bentuk atau pembuatan.*

***Jadual 1**. Masa kegagalan (dalam jam) bagi 20 bahagian mesin*

16+	*117*	*261+*	*323+*	*518*
531+	*643+*	*758*	*824*	*881*
1323	*1582*	*1795*	*1854+*	*2556*
2829	*5002*	*6298*	*13,991+*	*18,789*

Bahagian ini telah diiklankan oleh pengeluar sebagai mempunyai hayat reka bentuk 500-jam (dengan kebolehpercayaan 90 peratus). Syarikat itu ingin menentukan sama ada hayat reka bentuk yang diiklankan sedang dipenuhi sebelum ia memutuskan untuk menempah bahagian tambahan daripada pengeluar ini

(i) Laksanakan analisis kuasa dua terkecil dengan menggunakan taburan eksponen, Weibull, normal, dan lognormal.

(ii) Apakah penyuaian yang terbaik daripada keputusan yang diperoleh dalam (i)?

(iii) Cari $R(t)$ dan reka bentuk hayat 90 peratus jika Anggaran Kebolehjadian Maksimum (MLE) digunakan untuk model kebolehpercayaan penyuaian terbaik.

(iv) Apakah keputusan syarikat itu berdasarkan model kebolehpercayaan penyuaian terbaik?

[100 markah]

3. (a) Much cancer research has recently been directed at early detection through mass screening and an assessment of other variables that are predictive of remission times. In a study of breast cancer recurrence in female patients, only 17 observations were uncensored observed times until cancer recurrence and 56 observations were censored. **Table 2** displays the grouped data obtained from this study.

Table 2. Breast cancer recurrence data

Remission time (in months)	Number of patients	Number of recurrences	Number censored
0 - 50	73	2	5
50 - 70	66	5	15
70 - 80	46	1	6
80 - 90	39	4	3
90 - 100	32	1	8
100 - 120	23	3	11
120 -140	9	1	8

(i) What is the time to event in this study?
(ii) Why are some observations censored in this study?
(iii) What are the possible censoring mechanisms in this study?
(iv) Perform a suitable analysis on the grouped data in this study.
(v) Does the high proportion of censoring affect the estimates of survival probabilities? If it does, then explain.
(vi) What is the probability of survival of breast cancer patients beyond 120 months based on the results obtained in part (iv)?

(b) The data from the clinical trial examining steroid-induced remission times (in weeks) for two groups of leukemia patients are shown in **Table 3**.

One group of 21 patients were given 6-mercaptopurine (6-MP); a second group of 21 patients were given a placebo. Since the trial lasted 1 year and patients were admitted to the trial during the year, some of the data could not be gathered by the cut-off date when some patients were still in remission.

Table 3. Remission times (in weeks) for leukemia patients

6-MP	Placebo
6 6 6 6+ 7 9+ 10 10+ 11+ 13 16 17+ 19+ 20+ 22 23 25+ 32+ 32+ 34+ 35+	1 1 2 2 3 4 4 5 5 8 8 8 8 11 11 12 12 15 17 22 23

The notation '+' denotes 'censoring'.

(i) Give a general comparison of the survival prospects of the two groups by interpreting the graphs of their survival functions.
(ii) Read off the graph estimates of the probability of survival beyond 3 months.
(iii) What can you conclude from parts (i) and (ii)?

[100 marks]

3. (a) *Banyak penyelidikan kanser baru-baru ini telah diarahkan pada pengesanan awal melalui pemeriksaan besar-besaran dan penilaian pembolehubah lain yang meramalkan masa remitan. Dalam satu kajian kanser payudara berulang di kalangan pesakit wanita, hanya 17 pemerhatian adalah masa yang tidak tertapis sehingga berulangnya kanser dan 56 pemerhatian adalah tertapis.* ***Jadual 2*** *memaparkan data berkumpulan yang diperoleh daripada kajian ini.*

Jadual 2*. Data kanser payudara berulang*

Masa Peremitan (dalam bulan)	***Bilangan pesakit***	***Bilangan berulang***	***Bilangan tertapis***
0 - 50	*73*	*2*	*5*
50 - 70	*66*	*5*	*15*
70 - 80	*46*	*1*	*6*
80 - 90	*39*	*4*	*3*
90 - 100	*32*	*1*	*8*
100 - 120	*23*	*3*	*11*
120 -140	*9*	*1*	*8*

(i) Apakah masa kepada peristiwa dalam kajian ini?
(ii) Kenapakah sesetengah pemerhatian tertapis dalam kajian ini?
(iii) Apakah mekanisme penapisan yang mungkin dalam kajian ini?
(iv) Laksanakan analisis yang sesuai atas data berkumpulan dalam kajian ini.
(v) Adakah kadaran penapisan yang tinggi mempengaruhi anggaran kebarangkalian survival? Jika ya, jelaskan.
(vi) Apakah kebarangkalian survival pesakit kanser payudara yang melebihi 120 bulan berdasarkan keputusan yang diperoleh dalam bahagian (iv)?

(b) *Data daripada percubaan klinikal yang memeriksa masa remitan yang disebabkan steroid (dalam minggu) bagi dua kumpulan pesakit leukemia ditunjukkan dalam* ***Jadual 3****.*

Satu kumpulan dengan 21 pesakit telah diberikan 6-mercaptopurine (6-MP); kumpulan kedua dengan 21 pesakit telah diberikan plasebo. Oleh sebab percubaan itu berlangsung selama 1 tahun dan pesakit telah dimasukkan dalam percubaan itu pada tahun tersebut, beberapa data tidak dapat dikumpulkan pada tarikh 'cut-off' apabila sesetengah pesakit masih dalam remitan.

***Jadual 3**. Masa remitan (dalam minggu) bagi pesakit leukemia*

6-MP	***Plasebo***
6 6 6 6+ 7 9+ 10 10+ 11+ 13 16 17+ 19+ 20+ 22 23 25+ 32+ 32+ 34+ 35+	*1 1 2 2 3 4 4 5 5 8 8 8 8 11 11 12 12 15 17 22 23*

Notasi '+' menandakan 'penapisan'.

(i) Berikan perbandingan umum prospek survival kedua-dua kumpulan dengan mentafsirkan graf fungsi survival mereka.

(ii) Baca anggaran graf kebarangkalian untuk survival yang melebihi 3 bulan.

(iii) Apakah yang anda boleh simpulkan daripada bahagian-bahagian (i) dan (ii)?

[100 markah]

4. (a) In the treatment of certain disorders of the kidney, dialysis may be used to remove waste materials from the blood. One problem that can occur in patients on dialysis is the occurrence of an infection at the site at which the catheter is inserted. In a study to investigate the incidence of infection, the time from insertion of the catheter until infection was recorded for a group of kidney patients. The data in **Table 4** give the number of days from the insertion of the catheter until its removal following the first occurrence of an infection. The data set includes the values of a variable that indicates the infection status of an individual, which takes the value zero if the catheter was removed for a reason other than the occurrence of an infection, and unity otherwise. Also given are the age of each patient in years and a variable that denotes the sex of each patient (1 = male, 2 = female).

Table 4. Times to removal of a catheter following a kidney infection for 13 patients suffering from diseases of the kidney

Patient	Time (days)	Status	Age (years)	Sex
1	8	1	28	1
2	15	1	44	2
3	22	1	32	1
4	24	1	16	2
5	30	1	10	1
6	54	0	42	2
7	119	1	22	2
8	141	1	34	2
9	185	1	60	2
10	292	1	43	2
11	402	1	30	2
12	447	1	31	2
13	536	1	17	2

(i) Summarize the data set in **Table 4** before a more detailed analysis is performed.
(ii) What is the objective of this study?
(iii) Having determined the objective in part (ii), what is an appropriate method to achieve this objective?
(iv) Using the proposed methodology in part (iii), perform the detailed analysis on the data set in **Table 4**.
(v) From the results obtained in part (iv), what can you conclude?

(b) In a study undertaken to compare the treatments given to hypernephroma patients and to relate response and survival to surgery, metastasis, and treatment time, data from 58 patients were collected (see **Table 5**; only 5 cases are displayed). How would you analyze the data (use given data file) to answer these questions?

(i) Do patients who had nephrectomy have a higher response rate?
(ii) Is the time of nephrectomy related to response and survival?
(iii) Are there significant differences between the treatments?
(iv) What are the most important variables related to response and survival?

Table 5. Data of 58 patients with hypernephroma
(first 5 cases displayed)

Patient	Gender	Age (years)	Neph-rectomy	Time of Neph-rectomy	Treat-ment	Res-ponse	Sur-vival Time (weeks)	Status	Lung Meta-stasis	Bone Meta-stasis
1	2	53	1	0.0	1	1	77	0	1	0
2	1	69	1	4.0	1	2	18	1	0	1
3	1	61	0	-9.0	1	0	8	1	1	0
4	2	52	1	2.0	1	2	68	1	1	0
5	1	46	1	2.0	1	2	35	1	0	1

Note that:

Gender (1 = male, 2 = female)
Nephrectomy (1 = Yes, 0 = No)
Time of Nephrectomy (Number of years prior to treatment; negative value = no nephrectomy)
Treatment (1 = combined chemotherapy and immunotherapy, 2 = others)
Response (0 = no response, 1 = complete response, 2 = partial response, 3 = stable, 4 = increasing disease, 9 = unknown)
Status (1 = dead, 0 = alive)
Lung Metastasis (1 = Yes, 0 = No)
Bone Metastasis (1 = Yes, 0 = No)

[100 marks]

4. (a) *Dalam rawatan penyakit tertentu buah pinggang, dialisis boleh digunakan untuk membuang bahan-bahan buangan daripada darah. Satu masalah yang boleh berlaku pada pesakit yang menjalani rawatan dialisis ialah berlakunya jangkitan di tapak di mana kateter dimasukkan. Dalam satu kajian untuk menyiasat insiden jangkitan, masa bermula daripada pemasukan kateter sehingga jangkitan dicatatkan bagi kumpulan pesakit buah pinggang. Data dalam* ***Jadual 4*** *memberi bilangan hari dari penyisipan kateter sehingga penyingkiran berikutan kejadian pertama jangkitan. Set data termasuk nilai-nilai pembolehubah yang menunjukkan status jangkitan daripada seorang individu, yang mengambil nilai sifar jika kateter itu dikeluarkan atas sebab-sebab selain daripada berlakunya jangkitan, dan nilai satu sebaliknya. Juga diberikan adalah umur pesakit dalam tahun dan pembolehubah yang mewakili jantina setiap pesakit (1 = lelaki, 2 = perempuan).*

Jadual 4*. Masa untuk penyingkiran keteter berikutan jangkitan buah pinggang bagi 13 pesakit yang mengidap penyakit buah pinggang*

Pesakit	***Masa (hari)***	***Status***	***Umur (tahun)***	***Jantina***
1	*8*	*1*	*28*	*1*
2	*15*	*1*	*44*	*2*
3	*22*	*1*	*32*	*1*
4	*24*	*1*	*16*	*2*
5	*30*	*1*	*10*	*1*
6	*54*	*0*	*42*	*2*
7	*119*	*1*	*22*	*2*
8	*141*	*1*	*34*	*2*
9	*185*	*1*	*60*	*2*
10	*292*	*1*	*43*	*2*
11	*402*	*1*	*30*	*2*
12	*447*	*1*	*31*	*2*
13	*536*	*1*	*17*	*2*

(i) Ringkaskan set data dalam ***Jadual 4*** *sebelum analisis yang lebih terperinci dilakukan.*

(ii) Apakah objektif kajian ini?

(iii) Setelah menentukan objektif dalam bahagian (ii), apakah kaedah yang sesuai untuk mencapai objektif ini?

(iv) Dengan menggunakan kaedah yang dicadangkan dalam bahagian (iii), laksanakan analisis terperinci atas set data dalam ***Jadual 4****.*

(v) Daripada keputusan yang diperoleh dalam bahagian (iv), apakah kesimpulan anda?

(b) *Dalam satu kajian yang dijalankan untuk membandingkan rawatan yang diberikan kepada pesakit hipernefroma dan untuk mengaitkan respon dan survival pembedahan, metastasis, dan masa rawatan, data daripada 58 pesakit telah dikumpulkan (lihat **Jadual 5**; hanya 5 kes dipaparkan). Bagaimana anda akan menganalisis data (menggunakan fail data yang diberikan) untuk menjawab soalan-soalan berikut?*

(i) Adakah pesakit yang ada nefrektomi mempunyai kadar respon yang lebih tinggi?
(ii) Adakah masa nefrektomi berkaitan dengan respon dan survival?
(iii) Adakah terdapat perbezaan yang bererti antara rawatan?
(iv) Apakah pembolehubah yang paling penting yang berkaitan dengan respon dan survival?

***Jadual 5**. Data bagi 58 pesakit hipernefroma (5 kes pertama dipaparkan)*

Pesakit	*Jantina*	*Umur (tahun)*	*Nef-rektomi*	*Masa Nef-rektomi*	*Ra-watan*	*Res-pon*	*Masa Survival (minggu)*	*Status*	*Meta-stasis Paru-Paru*	*Meta-stasis Tulang*
1	*2*	*53*	*1*	*0.0*	*1*	*1*	*77*	*0*	*1*	*0*
2	*1*	*69*	*1*	*4.0*	*1*	*2*	*18*	*1*	*0*	*1*
3	*1*	*61*	*0*	*-9.0*	*1*	*0*	*8*	*1*	*1*	*0*
4	*2*	*52*	*1*	*2.0*	*1*	*2*	*68*	*1*	*1*	*0*
5	*1*	*46*	*1*	*2.0*	*1*	*2*	*35*	*1*	*0*	*1*

Nota:

Jantina (1 = lelaki, 2 = perempuan)
Nephrectomy (1 = Ya, 0 = Tidak)
Masa Nefrektomi (Bilangan tahun sebelum rawatan; nilai negatif = Tiada nefrektomi)
Rawatan (1 = gabungan kimoterapi dan immunoterapi, 2 = lain-lain)
Respon (0 = tiada respon, 1 = respon lengkap, 2 = respon sebahagian, 3 = stabil, 4 = peningkatan penyakit, 9 = tidak diketahui)
Status (1 = mati, 0 = hidup)
Metastasis Paru-Paru (1 = Ya, 0 = Tidak)
Metastasis Tulang (1 = Ya, 0 = Tidak)

[100 markah]

APPENDIX

Summary of Reliability Formulae

$$F(t) = \int_0^t f(t)dt$$

$$R(t) = 1 - F(t)$$

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}$$

$$h(t) = \frac{f(t)}{R(t)}$$

$$H(t) = \int_0^t h(t)dt$$

$$R(t) = e^{-H(t)}$$

$$H(t) = -\ln R(t)$$

$$MTTF = \int_0^\infty tf(t)dt = \int_0^\infty R(t)dt$$

APPENDIKS

Ringkasan Rumus-Rumus Kebolehpercayaan

$$F(t) = \int_0^t f(t)dt$$

$$R(t) = 1 - F(t)$$

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}$$

$$h(t) = \frac{f(t)}{R(t)}$$

$$H(t) = \int_0^t h(t)dt$$

$$R(t) = e^{-H(t)}$$

$$H(t) = -\ln R(t)$$

$$MTTF = \int_0^\infty tf(t)dt = \int_0^\infty R(t)dt$$

APPENDIX (contd)

Summary of Reliability Formulae (contd)

Lifetime following an **Exponential Distribution:**

$$f(t) = \lambda e^{-\lambda t}$$
$$F(t) = 1 - e^{-\lambda t}$$
$$R(t) = e^{-\lambda t}$$
$$h(t) = \lambda$$
$$H(t) = \lambda t$$
$$MTTF = \frac{1}{\lambda}$$

Lifetime following a **Weibull Distribution:**

$$f(t) = \beta \alpha^{-\beta} t^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$R(t) = e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$h(t) = \beta \alpha^{-\beta} t^{\beta-1}$$

$$H(t) = \left(\frac{t}{\alpha}\right)^{\beta}$$

$$MTTF = \alpha \Gamma\left(1 + \frac{1}{\beta}\right)$$

$$\text{Design Life} = t_R = \alpha(-\ln R)^{1/\beta}$$

APPENDIKS (sambung)

Ringkasan Rumus-Rumus Kebolehpercayaan (sambung)

Masahayat mengikut ***Taburan Eksponen:***

$$f(t) = \lambda e^{-\lambda t}$$
$$F(t) = 1 - e^{-\lambda t}$$
$$R(t) = e^{-\lambda t}$$
$$h(t) = \lambda$$
$$H(t) = \lambda t$$
$$MTTF = \frac{1}{\lambda}$$

Masahayat mengikut ***Taburan Weibull:***

$$f(t) = \beta \alpha^{-\beta} t^{\beta-1} \exp\left[-\left(\frac{t}{\alpha} \right)^{\beta} \right]$$

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$R(t) = e^{-\left(\frac{t}{\alpha}\right)^{\beta}}$$

$$h(t) = \beta \alpha^{-\beta} t^{\beta-1}$$

$$H(t) = \left(\frac{t}{\alpha} \right)^{\beta}$$

$$MTTF = \alpha \Gamma\left(1 + \frac{1}{\beta} \right)$$

$$\text{Design Life} = t_R = \alpha(-\ln R)^{1/\beta}$$

APPENDIX (contd)

Summary of Reliability Formulae (contd)

Lifetime following a **Normal Distribution:**

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \frac{(t-\mu)^2}{\sigma^2} \right]$$

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma} \right)$$

$$R(t) = 1 - \Phi\left(\frac{t-\mu}{\sigma} \right)$$

$$h(t) = \frac{f(t)}{1 - \Phi\left(\frac{t-\mu}{\sigma} \right)}$$

Lifetime following a **Lognormal Distribution:**

$$f(t) = \frac{1}{\sqrt{2\pi}st} \exp\left[-\frac{1}{2s^2} \left(\ln \frac{t}{t_{median}} \right)^2 \right]$$

$$F(t) = \Phi\left(\frac{1}{s} \ln \frac{t}{t_{median}} \right)$$

$$R(t) = 1 - \Phi\left(\frac{1}{s} \ln \frac{t}{t_{median}} \right)$$

$$MTTF = t_{median} \exp\left(\frac{s^2}{2} \right)$$

$$t_R = t_{median} \exp\left(s z_{1-R} \right)$$

APPENDIKS (sambung)

Ringkasan Rumus-Rumus Kebolehpercayaan (sambung)

Masahayat mengikut ***Taburan Normal:***

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \frac{(t-\mu)^2}{\sigma^2} \right]$$

$$F(t) = \Phi\left(\frac{t-\mu}{\sigma} \right)$$

$$R(t) = 1 - \Phi\left(\frac{t-\mu}{\sigma} \right)$$

$$h(t) = \frac{f(t)}{1 - \Phi\left(\frac{t-\mu}{\sigma} \right)}$$

Masahayat mengikut ***Taburan Lognormal:***

$$f(t) = \frac{1}{\sqrt{2\pi}st} \exp\left[-\frac{1}{2s^2} \left(\ln \frac{t}{t_{median}} \right)^2 \right]$$

$$F(t) = \Phi\left(\frac{1}{s} \ln \frac{t}{t_{median}} \right)$$

$$R(t) = 1 - \Phi\left(\frac{1}{s} \ln \frac{t}{t_{median}} \right)$$

$$MTTF = t_{median} \exp\left(\frac{s^2}{2} \right)$$

$$t_R = t_{median} \exp\left(s z_{1-R} \right)$$

- ooo O ooo -