# MATH364 Summer 2003 Exam Solutions

1. (a) *Prevalence* is the proportion of the population who have the disease of interest;
*Sensitivity* is the probability that a diseased person will produce a positive test result;
*Specificity* is the probability that a non-diseased person will produce a negative test result;
*Positive predictive value* is the probability that a person who tests positive is in fact diseased.

$$\begin{aligned} \text{PPV} &= \Pr(\text{Diseased} \mid \text{Test positive}) \\ &= \frac{\Pr(\text{D and } +\text{ve})}{\Pr(+\text{ve})} \\ &= \frac{\Pr(\text{D})\Pr(+\text{ve} \mid \text{D})}{\Pr(\text{D})\Pr(+\text{ve} \mid \text{D}) + \Pr(\text{ND})\Pr(+\text{ve} \mid \text{ND})} \\ &= \frac{\text{Prevalence} \times \text{Sensitivity}}{\text{Prev} \times \text{Sens} + (1 - \text{Prev})(1 - \text{Specificity})} \end{aligned}$$

[4 marks for definitions;
4 marks for calculation.]

(b) (i) Sensitivity $= 7/15 = 47\%$, Specificity $= 28/34 = 82\%$

(ii) PPV $= 7/13 = 54\%$

(iii) With prevalence 10%, PPV $= \dfrac{0.1 \times 0.47}{0.1 \times 0.47 + 0.9 \times 0.18} = 23\%$

(iv) Comment: Sensitivity seems very low; specificity is fine. For a population with prevalence 10%, PPV is only 23%, so that most of those with high viral load will not in fact develop severe disease.

[1 mark for sensitivity;
1 mark for specificity;
1 mark for PPV for sample;
2 marks for PPV for population;
2 marks for comment.]

(c) For a diagnostic test based on a continuous measurement, an ROC curve is a plot of Sensitivity versus Specificity for various different cutoff values. For the screening test of (b), a cutoff value other than 4.5 might be better. To find out, plot an ROC curve, using various different cutoff values. Points towards the top right corner of the graph (high sensitivity and high specificity) correspond to the best choices of cutoff value. Would need to collect more detailed data on the viral load of individuals with/without severe disease, not just whether the load is above or below 4.5, but also whether it's above or below various other cutoff values.

[5 marks.]

2. (a) In a cohort study, identify subjects according to risk factors then follow up to see whether disease develops. In a case-control study, identify subjects according to whether they have disease or not, then obtain information on risk factors.

Advantages of cohort study include: can distinguish antecedent causes from associated factors; can estimate incidence, relative and attributable risk; bias in risk factor information not a problem.

Disadvantages of cohort study include: cost of large long-term studies; bias from loss to follow-up.

Advantages of case-control study include: quick and cheap; good for rare diseases; minimise number of non-diseased studied; suitable for multiple hypotheses.

Disadvantages of case-control study include: temporal relationship in doubt; no estimate of incidence; recall bias.

With population proportions

| | Case | Healthy |
|---|---|---|
| Exposed | $P_1$ | $P_2$ |
| Not Exposed | $P_3$ | $P_4$ |

$$\text{Relative Risk} = \frac{P_1/(P_1+P_2)}{P_3/(P_3+P_4)} \qquad \text{Odds Ratio} = \frac{P_1/P_2}{P_3/P_4}$$

If $P_1$ and $P_3$ are small, ie for a 'rare' disease, the odds ratio is approximately equal to the relative risk.

[1 mark each for defining cohort study, case-control study, relative risk, odds ratio;
4 marks for advantages/disadvantages;
1 mark for circumstances for equality.]

(b) Odds ratio estimate $\hat{\psi} = \dfrac{36/78}{11/95} = 3.99$

95% CI for $\ln(\psi)$ is $\ln(3.99) \pm 1.96\sqrt{(1/36)+(1/78)+(1/11)+(1/95)} = 1.383 \pm 1.96 \times 0.377 = [0.644, 2.121]$

95% CI for $\psi$ is $[1.90, 8.34]$

Interpretation: The odds ratio estimate suggests that women who first had sexual intercourse before the age of 15 are at 4 times the risk of cervical cancer compared to those who did not. The CI says we're fairly confident that early sexual intercourse is associated with an increase in risk of between 2 times and 8 times.

There is evidence of an association between age of first intercourse and development of cervical cancer.

[1 mark for OR estimate;
4 marks for CI;
2 marks for interpretation;
1 mark for saying there is an association.]

(c) Would prefer a case-control study, because the condition is quite rare, so that in order to pick up a sufficient number of cases a cohort study would need to involve a large number of women over a long period of time, and would be very expensive.

[3 marks.]

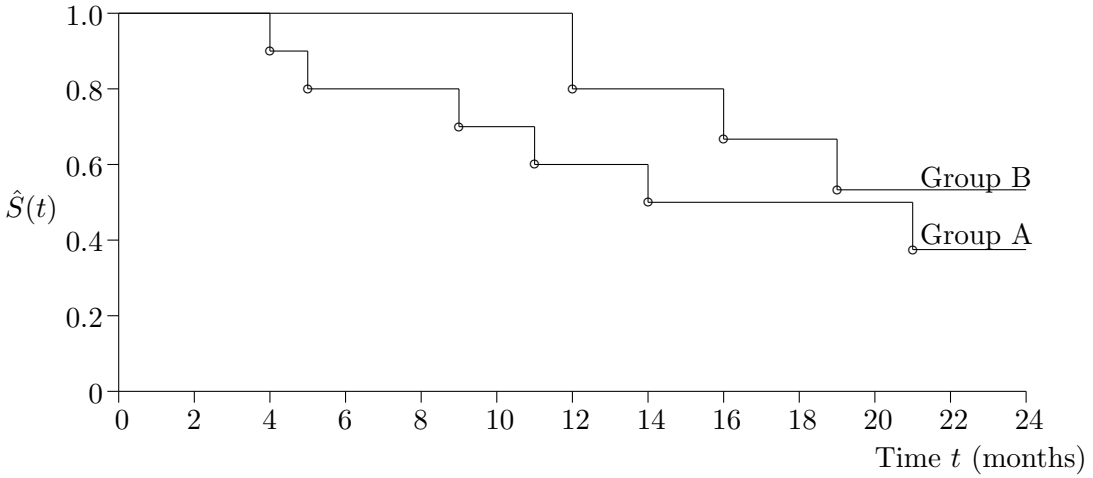
3. (a) Group A:

| $t_{(i)}$ | $n_i$ | $d_i$ | $1-(d_i/n_i)$ | $S(t_i+)$ |
|---|---|---|---|---|
| 4 | 10 | 1 | 9/10 | 0.900 |
| 5 | 9 | 1 | 8/9 | 0.800 |
| 9 | 8 | 1 | 7/8 | 0.700 |
| 11 | 7 | 1 | 6/7 | 0.600 |
| 14 | 6 | 1 | 5/6 | 0.500 |
| 21 | 4 | 1 | 3/4 | 0.375 |

Group B:

| $t_{(i)}$ | $n_i$ | $d_i$ | $1-(d_i/n_i)$ | $S(t_i+)$ |
|---|---|---|---|---|
| 12 | 10 | 2 | 8/10 | 0.800 |
| 16 | 6 | 1 | 5/6 | 0.667 |
| 19 | 5 | 1 | 4/5 | 0.533 |

Kaplan-Meier estimates:




Comment: Group A graph is consistently below Group B graph, suggesting that survival is better in Group B.

[2 marks for $t$, $n$, $d$ values; 3 marks for $S(t)$ values;
3 marks for K-M graphs; 2 marks for comment.]

(b) Log-rank test:

| $t_{(i)}$ | $d_i$ | $n_{Ai}$ | $n_{Bi}$ | $e_{Ai}$ |
|---|---|---|---|---|
| 4 | 1 | 10 | 10 | 0.500 |
| 5 | 1 | 9 | 10 | 0.474 |
| 9 | 1 | 8 | 10 | 0.444 |
| 11 | 1 | 7 | 10 | 0.412 |
| 12 | 2 | 6 | 10 | 0.750 |
| 14 | 1 | 6 | 7 | 0.462 |
| 16 | 1 | 5 | 6 | 0.455 |
| 19 | 1 | 4 | 5 | 0.444 |
| 21 | 1 | 4 | 4 | 0.500 |

Expected deaths in each group are $E_A = 4.44$, $E_B = 5.56$.

Test statistic $X^2 = \frac{(6-4.44)^2}{4.44} + \frac{(4-5.56)^2}{5.56} = 0.985$, to be compared with $\chi^2_1$. From tables, $\chi^2_1(0.05) = 3.84$, so at the 5% level no evidence against $H_0$.

Interpretation: there's no evidence of a difference in survival between the two groups.

Comment: Although the graphs in (a) suggested a consistent difference, the difference wasn't very large, and the sample sizes are small, so evidence of a difference is not statistically significant. Larger samples are needed.

[2 marks for $t$, $n$ values; 2 marks for $e$ values;
1 mark for $E_A$, $E_B$; 1 mark for $X^2$ value;
1 mark for critical value and test;
1 mark for interpretation; 2 marks for comment.]

4. (a) Have an unknown parameter $\theta$ and observed data $\mathbf{X}$. Before observing the data, express our beliefs about $\theta$ in a prior distribution $f(\theta)$. Using the data and a statistical model, can work out the likelihood $f(\mathbf{X} \mid \theta)$. Our beliefs about $\theta$ are then modified in the light of the data, and expressed as a posterior distribution $f(\theta|\mathbf{X})$, which is found using Bayes' theorem, so that $f(\theta \mid \mathbf{X}) \propto f(\theta) f(\mathbf{X} \mid \theta)$.

   For a clinical trial, possible sources of prior information include data from previous trials; expert clinical opinion; ignorance priors.

   A conjugate prior has the property that the posterior distribution is of the same form as the prior, but with updated parameter values.

   [5 marks for features of a Bayesian analysis;
   2 marks for sources of prior information;
   1 mark for conjugate prior definition.]

   (b) Prior beliefs:

$$\begin{aligned} \mu &= \log_e \left( \frac{(\alpha_T - 0.5)\,(\beta_N - 0.5)}{(\alpha_N - 0.5)\,(\beta_T - 0.5)} \right) \\ &= \log_e \left( \frac{(50-0.5)(320-0.5)}{(80-0.5)(280-0.5)} \right) = \log_e(0.7117) = -0.3400 \\ \sigma^2 &= \frac{1}{\alpha_N} + \frac{1}{\beta_N} + \frac{1}{\alpha_T} + \frac{1}{\beta_T} = 0.0392 \end{aligned}$$

   Prior for log odds ratio is $N(-0.34, 0.0392)$.

   Hence 95% CI for log odds ratio is $-0.34 \pm 1.96\sqrt{0.0392} = (-0.728, 0.048)$

   95% CI for odds ratio is $(0.483, 1.049)$.

   Interpretation: The CI includes 1, so there's no real evidence of a difference between the treatment and no treatment.

   [2 marks for distribution of log(OR);
   2 marks for CI for log odds ratio;
   1 mark for CI for odds ratio;
   1 mark for interpretation.]

   (c) Sample estimate of odds ratio is $(21/159)/(24/136) = 0.748$.

   Posterior CI for OR excludes 1, so there is evidence that the treatment has an effect. Appears that risk of death within one year is lower on the treatment than without treatment.

   The prior belief was that there was no real evidence of the treatment being better than no treatment; in the sample, the treatment did perform better than no treatment; this extra evidence has altered the CI to exclude 1, so that posterior belief is that the treatment is better than no treatment. Note that prior CI only just included 1, whereas posterior CI only just excludes 1, so that the change in conclusion could be regarded as an artefact caused by our choice of 5% significance level.

   [1 mark for sample OR;
   2 marks for interpretation of posterior CI;
   3 marks for comment.]

5. (a) Calculating log odds ratio and weight for each trial (and precision values needed for part (c)),

| Trial | $\hat{y}_i$ | $w_i$ | $w_i\hat{y}_i$ | Precision |
|---|---|---|---|---|
| 1 | -0.365 | 16.178 | -5.911 | 4.022 |
| 2 | -0.080 | 15.175 | -1.209 | 3.895 |
| 3 | 0.396 | 11.915 | 4.720 | 3.452 |
| 4 | 0.666 | 5.964 | 3.975 | 2.442 |
| 5 | 0.000 | 2.763 | 0.000 | 1.662 |
| 6 | -1.639 | 0.813 | -1.333 | 0.902 |
| 7 | -1.130 | 1.432 | -1.618 | 1.197 |
| 8 | -0.169 | 3.294 | -0.558 | 1.815 |

Pooled log odds ratio $\hat{y}_p = -1.935/57.535 = -0.034$.

Value is negative, suggesting that odds of a 'response' on Treatment A are lower than on Treatment B, so that Treatment A is better. Difference appears to be very small.

[5 marks for $\hat{y}_i$, $w_i$ values; 1 mark for $\hat{y}_p$; 2 marks for interpretation.]

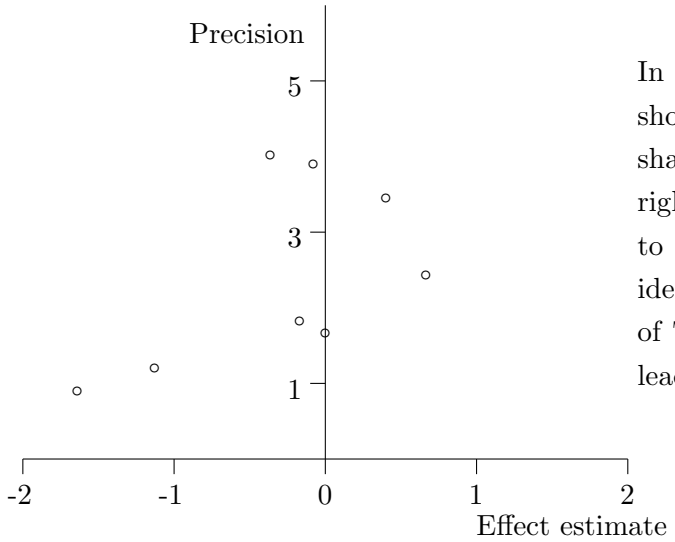
(b) Value of $Q$ to be compared with $\chi^2_7$; from tables, $\chi^2_7(0.05) = 14.07$; given value of $Q$ not as large as this, so no evidence of heterogeneity. (In fact, $\chi^2_7(0.20) = 9.803$, $\chi^2_7(0.10) = 12.02$, so that $0.1 < p < 0.2$.)

Seems OK to combine results and use pooled treatment effect.

[3 marks for test and conclusion;
1 mark for saying OK to pool.]

(c) *Publication bias* is the fact that trials which produce statistically significant results are more likely to get published than those which don't.

To investigate the possibility of publication bias, plot precision against treatment effect estimate for the 8 trials. In this case, treatment effect estimate is the log odds ratio $\hat{y}_i$ and precision is 1/s.d. $(\hat{y}_i)$.




In the absence of bias, the plot should be symmetrical and funnel-shaped. In this case, the bottom right corner of the funnel seems to be missing, suggesting some evidence of publication bias in favour of Treatment A being less likely to lead to myocardial infarction.

[2 marks for meaning of publication bias;
4 marks for precision values and graph;
2 marks for saying evidence of publication bias, with reason.]

6. (a)
- **Phase I trials** are done during the early development of a drug, and include clinical pharmacology and toxicology, dose-finding studies and perfection of technique. These studies may be done on animals or healthy volunteers. They can include both randomised and un-controlled studies.
- **Phase II trials** are the initial clinical studies to look at treatment effect, including pilot studies. They involve patients, but can include those for whom other treatments have failed. They are typically un-controlled.
- **Phase III trials** are the full-scale evaluation of treatment, done on patients prior to registration of a new drug. These are typically large randomised trials of two or more groups.
- **Phase IV trials** are done for post-marketing surveillance to identify rare adverse reactions. These are usually huge, and may be of uncontrolled, non-randomised controlled or randomised design.

[2 marks for each phase, up to a maximum of 7 marks]

(b) $n$ is the number of patients required per group;
$\sigma$ is the standard deviation of the outcome measure, assumed to be the same for the two groups;
$d$ is the magnitude of the difference between means which would be clinically significant;
$\alpha$ is the significance level of the test;
$\beta$ is the probability of Type II error, $\beta = 1 - \text{Power}$.

[1 mark for defining each term]

(c) Have $\sigma = 4$, $d = 3$, $\alpha = 0.05$, $\beta = 0.1$, so

$$n = \frac{2 \times 4^2}{3^2} \left( \Phi^{-1}(0.9) + \Phi^{-1}(0.975) \right)^2 = \frac{32}{9} (1.2816 + 1.9600)^2 = 37.36$$

That is, require 38 patients per group, or 76 patients in total.

Re-calculating with $d = 5$ gives $n = 13.45$, so require 28 patients in total.

A bigger difference will be easier to detect, so with $d = 5$ don't need such a large sample as with $d = 3$.

[4 marks for sample size with $n = 3$;
2 marks for sampe size with $n = 5$;
2 marks for comment.]

7. (a) Cox regression model supposes that the hazard rate for an individual with covariate values $x_1, \ldots, x_n$ may be written in the form

$$\lambda(t) \quad = \quad \lambda_0(t) \exp\{\beta_1 x_1 + \cdots + \beta_n x_n\}$$

where $\beta_1, \ldots, \beta_n$ are parameters to be estimated from the data and $\lambda_0(t)$ is some unknown baseline hazard rate. Proportional hazards assumption implicit in this is that for any two individuals A and B, the hazard ratio $\lambda_A(t)/\lambda_B(t)$ remains constant over time.

[4 marks]

(b) (i) Fitted model: Hazard rate is

$$\lambda(t) \quad = \quad \lambda_0(t) \exp\{-1.055(\text{if treated}) + 1.955(\text{if male}) + 0.852 \times (\text{thickness in mm})\}$$

[2 marks]

(ii) Treatment: $p$-value of 0.05 means there is some evidence that treatment affects survival chances, but the evidence isn't very strong. Only just significant at the 5% level. Relative hazard is $\exp(-1.055) = 0.348$, so that at any given time, the risk of death for a patient in the treatment group is about 35% that of a patient of the same sex and tumour thickness in the control group.

Sex: $p$-value of 0.002 means there's strong evidence that sex affects survival chances. Relative hazard is $\exp(1.955) = 7.064$, so that at any given time, the risk of death for a male is about 7 times that of a female in the same group with the same tumour thickness.

Thickness: $p$-value of $< 0.001$ means very strong evidence that tumour thickness affects survival. Relative hazaed is $\exp(0.852) = 2.344$, so that at any given time, for a person of given sex and group, each extra mm of tumour thickness increases the risk of death by a factor of around 2.3.

Conclude that treatment does appear to be effective, reducing risk of death to 35% what it would be without treatment, but given that $p$-value is 0.05 the evidence of such an effect isn't totally convincing.

[2 marks for each covariate;
2 marks for conclusion about effectiveness of treatment.]

(iii) Risk score $= -1.055(\text{if treated}) + 1.955(\text{if male}) + 0.852 \times (\text{thickness in mm})$.
Patient A: Risk score $= -1.055 + 1.955 + 0.852 \times 0.6 = 1.411$.
Patient B: Risk score $= -1.055 + 0.852 \times 1.7 = 0.393$.
Patient A is at greater risk.

[1 mark for definition;
2 marks for the two calculations;
1 mark for conclusion.]

(iv) Multiple regression inappropriate because survival data is never Normally distributed and because the data include censored observations.

[2 marks]