Statistical tables are attached

Two Hours

UNIVERSITY OF MANCHESTER

May 2005 Final Draft

Medical Statistics MT3772

Electronic calculators may be used provided that they cannot store text

---

Answer **ALL** six questions in **SECTION A** (50 Marks)

Answer **TWO** of the three questions in **SECTION B** (25 marks each)

The total number of marks on the paper is 100

A1.

(i) In the context of a randomised controlled trial, explain what is meant by the term *double-blind*.

(ii) Give two advantages of a trial being *double blind*?

(iii) Give an example of treatment that cannot be evaluated in a *double-blind* clinical trial.

[6 marks]

A2.

Tabulated below are summary data from randomised controlled trial comparing a surgical and medical treatment for elderly stroke patients .

Survival at 3 years	*Surgical*		*Medical*	
	Received Surgical	*Received Medical*	*Received Surgical*	*Received Medical*
No	15	15	4	26
Yes	105	45	26	124
Total	120	60	30	150

(i) Calculate the point estimates of the treatment effect of surgery compared to medical treatment measured by the proportion surviving at 3 years for

(a) *Intention-To-Treat*, (b) *Per-Protocol* and (c) *As-Treated* analyses.

(ii) Briefly explain why an *Intention-To-Treat* analysis is usually preferable to *Per-protocol* and *As-Treated* analyses.

[7 marks]

A3.

A randomised controlled crossover trial is used to assess whether *thyroxin* treatment is effective in patients with symptoms of hypothyroidism. Patients were randomised to either *thyroxin* then *placebo*(TP) or *placebo* then *thyroxin* (PT). The table below summarizes the sample mean and standard deviation (*s.d.*) of the free thyroxin (pmol/l) in the blood.

Sequence	*Period 1*		*Period 2*		*Period 2 – Period 1*		*N*
	mean	*s.d.*	*mean*	*s.d.*	*mean*	*s.d*	
TP	18.4	6.4	18.3	7.2	-0.1	3.5	20
PT	16.1	6.3	19.2	6.9	3.1	3.4	20

(i) Using the summary data for just *Period 1* test the null hypothesis that *thyroxin* treatment does not affect free thyroxin level.

(ii) Using the summary data for the difference between *Period 2* and *Period 1* test the null hypothesis that thyroxin does not affect free thyroxin levels.

(iii) Outline the advantages and disadvantages of a crossover design as compared to parallel group design.

[14 marks]

A4.

(i) Explain the difference between a *fixed effect* and a *random effect* meta-analysis.

(iii) In the context of a meta-analysis, what is a *funnel* plot?

(iv) When carrying out a meta-analysis, why might one draw a *funnel* plot?

[9 marks]

A5 .

(i) Describe three designs for sampling a population to gather data on the relationship between an exposure and disease.

(ii) Give two factors which should influence the choice between these designs, given that there is a fixed budget for the study.

[6 Mark]

A6.

Consider the problem of estimating $\theta_j = \lambda_{1j} / \lambda_{0j}$ where λ_{1j} and λ_{0j} are rate of liver cancer in age-group j of an exposed male workforce and of the male population of the UK respectively, j=1,…k, under the assumption that $\theta_j = \theta$ for all j and when the rates λ_{0j} are already known.

Data on the observed number of cancers, b_j and corresponding man-years at risk, T_j in each age-group of the workforce are to be collected.

(i) Show that the maximum likelihood estimator of θ is $\dfrac{\sum b_j}{\sum \lambda_{0j} T_j}$

(ii) Estimate θ from the following data.

Age-group	*W'force:m-years*	*W-force: LC deaths*	*UK: rate of LC per 10,000 p-years*
1	*1000*	*1*	*5*
2	*2000*	*4*	*15*
3	*1500*	*7*	*30*

[8 Marks]

B7. In a parallel group *equivalence* trial a new treatment T is being compared with a control treatment C on a continuous outcome measure Y. Let $\bar{y}_T$, $\bar{y}_C$, μ_T and μ_C be the sample and population means of Y for each treatment, n_T and n_C be the sample sizes, and s be the common within-group sample standard deviation of Y. Define the treatment effect $\tau = \mu_T - \mu_C$.

(i) Explain why a significance test of the hypothesis $H_0 : \tau = 0$ vs $H_1 : \tau \neq 0$ would be an inappropriate in an equivalence trial. [3marks]

(ii) Suppose that the null hypothesis $H_0 : |\mu_T - \mu_C| \geq \delta_E$ is rejected if the (1-2α) confidence interval, given by $\bar{y}_T - \bar{y}_C \pm z_\alpha \lambda s$ with $\lambda = \sqrt{1/n_T + 1/n_C}$, is within the interval (-δ_E, +δ_E). Show that

$$\Pr[\text{Reject } H_0 | \tau] = \Phi\left(\frac{(\delta_E - \tau)}{s\lambda} - z_\alpha\right) - \Phi\left(-\frac{(\delta_E + \tau)}{s\lambda} + z_\alpha\right)$$

where Φ is the cumulative distribution function of the standard normal distribution. [6 marks]

(iii) Show that $\Pr[\text{Reject } H_0 | \tau]$ has a maximum under H_0 when $\tau = -\delta_E$ or $\tau = \delta_E$. Hence show that this procedure has a type I error $\leq \alpha$.
[8 marks]

(iv) An equivalence trial is carried out to test whether a new cheaper generic drug with is as effective as an existing proprietary drug for controlling blood pressure. Twenty-five patients receive each treatment. The table below summarizes the systolic blood for each treatment group. A difference of 5 mmHg was considered to be clinically important.

Generic Drug Treatment			*Proprietary Drug Treatment*		
Mean	*S.D.*	*N*	*Mean*	*S.D.*	*N*
135.5	14.7	25	136.1	15.1	25

The 90% confidence interval for the mean difference between the *Proprietary Drug* group and the *Generic Drug* group is (-7.9 mmHg, 9.1 mmHg). Comment on the results.
[4 marks]

(v) Explain why compliance to treatment is particularly important in an equivalence trial.
[4 marks]
[total 25 marks]

B8 A randomised controlled trial is planned to compare a treatment (A) with the current standard therapy (B). For an outcome measure *Y* let $\bar{y}_A, \bar{y}_B$, μ_A, μ_B the sample and population means of *Y* for each treatment. Let *s* and σ be the common within-group sample and population standard deviation of *Y*. Assume that the null hypothesis of no treatment effect $H_0 : \mu_A = \mu_B$ will be tested by the statistic $T = \dfrac{\bar{y}_A - \bar{y}_B}{s\lambda}$, with $\lambda = \sqrt{1/n_A + 1/n_B}$ where n_A, n_B are the number of subjects allocated to each treatment.

(i) Write down an expression for the power (1-β) of the test for a two-tailed alternative hypothesis to detect a treatment effect τ, stating the assumptions you make.

[5 marks]

(ii) Assuming that patients are allocated in the ratio of 1: k such that $n_B = k.n_A$ show that the total sample size required to give a power (1-β) for a two-tailed α size test is

$$n = \frac{(k+1)^2}{k}\frac{\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_{\beta}\right)^2$$

.

[8 marks]

(iii) Show that the minimum sample size is obtained by using an equal allocation ratio.

[5 marks]

(iv) In a randomised controlled trial it is planned to allocate 240 patients to two treatments. The pooled within-group standard deviation is thought to be 12 units. Estimate the power of a study to detect a treatment effect of 5 units for two-tailed *0.05* size test, if an allocation ratio of 1:2 is used.

[3 marks]

(vi) Describe how block randomisation could be used to randomly allocate treatments to patients with an allocation ratio of 1:2.

[4 marks]

[total 25 marks]

Q9. An investigator wishes to estimate the effect of hormone replacement therapy (HRT) on incidence of heart disease. Assume the true ratio of the incidence rate in users of the therapy versus that in comparable unexposed people is θ_H. The investigator estimates θ_H by t_H, the ratio of the observed cumulative incidence rate in a group of users versus a group of never-users. However the groups are not comparable since proportions p_1 and p_0 of the users and never-users were smokers; smoking multiplies the incidence of heart disease by a factor θ_S.

(i) Show that $t_E = \dfrac{\theta_H(1 - p_1 + \theta_S p_1)}{(1 - p_0 + \theta_S p_0)}$.

[6 marks]

(ii) Suppose that $\theta_H = 1.5$, $\theta_S = 5$, $p_1 = 0.2$ and $p_0 = 0.7$. Calculate t_E and comment on your result in relation to the effect of HRT.

[3 marks]

(iii) Show that when $\theta_H > 1$ and $\theta_S > 1$, a necessary condition for $t_E < 1$ is that $p_0 > p_1\theta_H$.

[4 marks]

In a different study to address the same question, the investigator collects data on 4 variables: regular alcohol use (yes/no), diabetic status (Yes/No), Body Mass Index (BMI, a continuous measure) and HRT use (ever/never) from 1017 women. The incidence of heart attacks over 5 years is also recorded and recorded as ≥1 or 0. The following logistic regression analyses using STATA software are produced.

Analysis A.

```
Logit estimates                                   Number of obs   =       1007
                                                  LR chi2(4)      =      34.43
                                                  Prob > chi2     =     0.0000
Log likelihood = -354.58                          Pseudo R2       =     0.0463
------------------------------------------------------------------------------
      attack |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         HRT |      .0244   0.1979       0.12   0.902      -0.3634    0.4123
     alcohol |     -.2141   0.2402      -0.89   0.373      -0.6848    0.2566
    diabetes |     1.1681   0.2258       5.17   0.000       0.7254    1.6107
         bmi |     0.0331   0.0156       2.12   0.034       0.0025    0.0675
       _cons |    -3.0529   0.5200      -5.87   0.000      -4.0722   -2.0336
------------------------------------------------------------------------------
```

Analysis B

```
Logit estimates                                   Number of obs   =       1007
                                                  LR chi2(1)      =       0.00
                                                  Prob > chi2     =     0.9932
Log likelihood = -375.07                          Pseudo R2       =     0.0000
------------------------------------------------------------------------------
      attack |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         HRT |     -.4104   0.1923      -2.13   0.033      -0.7873   -0.0335
       _cons |    -1.9827   0.1366     -14.52   0.000      -2.2504   -1.7150
------------------------------------------------------------------------------
```

(iv) Calculate the odds ratios for alcohol, BMI and diabetes and use the results of the tests to interpret information about the relationships between each of these factors and the risk of heart attacks.

[6 marks]

(v) What is the odds ratio comparing two people whose BMI measurements differ by 5 units?

[2 marks]

(vi) Calculate the odds ratios for HRT from *Analysis A* and *Analysis B*. Why might the results for HRT be different? Which one is likely to be the better indicator of the effect of HRT and why?

[4 marks]

[total 25 marks]