Statistical tables are attached

Two Hours

UNIVERSITY OF MANCHESTER

19 January 2015

14:00 – 16:00

Medical Statistics MATH38071

Electronic calculators may be used provided that they cannot store text

---

Answer **ALL** five questions in **SECTION A** (40 marks in total).

Answer **TWO** of the three questions in Section B (40 marks in total). If more than two questions from Section B are attempted, then credit will be given for the two best answers.

The total number of marks on the paper is 80.

A1

(i) In studies investigating the effect of an exposure on health, what is the difference between observational and experimental studies? [2 marks]

(ii) What is a confounding variable? [2 marks]

(iii) Consider an epidemiological study investigating whether high fat consumption causes heart disease. For such a study suggest an example of each of the three types of variable

a) Exposure

b) Outcome

c) Confounding.

[3 marks]

[Total 7 marks]

A2.

A randomized controlled trial is carried out to compare a new treatment regime (*N*) with the existing standard treatment (S) for patients. The effectiveness of treatment is assessed by whether the patient is still *infectious after 2 weeks*. The results are summarized in the frequency table below.

		Treatment	
		Standard (S)	*New (N)*
Infectious after 2 weeks	*Yes*	100	70
	No	400	430
Total		500	500

(i) Calculate the rate ratio (RR) of the patient still being *infectious after 2 weeks* with the new treatment (*N*) compared to the standard treatment (*S*). [2 marks]

(ii) Suppose n_T patients are randomized to treatment (*T*) and n_C to the control (*C*). Suppose that the number of events in each of the two treatment groups are r_T and r_C. The standard error of the $\log_e\left[\hat{RR}\right]$ is $\hat{SE}\left[\log_e\left[\hat{RR}\right]\right] = \sqrt{\frac{1}{r_T} - \frac{1}{n_T} + \frac{1}{r_C} - \frac{1}{n_C}}$.

Calculate the 95% confidence interval for rate ratio of being *infectious after 2 weeks*.

[5 marks]

(iii) Is there evidence that the new treatment (*N*) is better than the standard treatment (*S*)?

[2 marks]

[Total 9 marks]

A3.

(i) A researcher designing a randomised control trial considers patient's age and gender to be prognostic. Explain how you would carry out stratified randomisation with stratification by age and sex. [5 marks]

(ii) What are the advantages and disadvantages of stratified randomisation? [3 marks]

[Total 8 marks]

A4.

The results for a randomised controlled trial comparing a *New* treatment with a *Control* treatment for binary outcome measure (*Recovered after 4 weeks*) are tabulated below. Some patients randomised to the *New* treatment receive the *Control* treatment, but no patients randomised to the *Control* treatment receive the *New* treatment.

Recovered after 4 weeks	*Randomised group*		
	New		*Control*
	Received New	*Received Control*	
Yes	110	5	120
No	40	45	80
Total	150	50	200

(i) Calculate the point estimates of the treatment effect of the *New* treatment compared to the *Control* treatment define by the difference in the proportion of patients who have *Recovered after 4 weeks* for

a) an *Intention-To-Treat* analysis

b) a *Per-Protocol* analysis.

c) an As Treated analysis

[4 marks]

(ii) Drawing on the example above, explain why an *Intention-To-Treat* analysis is preferable to *Per-protocol* and *As-treated* analyses in a superiority trial.

[4 marks]

[Total 8 marks]

A5.

In a meta-analysis of *k* trials, suppose $\hat{\theta}_i$ is an estimate of the treatment effect for the i^{th} study and let $Var\left[\hat{\theta}_i\right]$ be its sampling variance. The minimum variance estimate is defined by $\hat{\theta}_{MV} = \dfrac{\sum_i^k w_i \hat{\theta}_i}{\sum_i^k w_i}$,

where $w_i = 1/Var\left[\hat{\theta}_i\right]$, and $Var\left[\hat{\theta}_{MV}\right] = \dfrac{1}{\sum_{i=1}^{k} \dfrac{1}{Var\left[\hat{\theta}_i\right]}}$.

The table below summarizes the outcome of two trials comparing a new drug to prevent high blood cholesterol with the current standard drug treatment. The treatment effect for each study $\left(\hat{\theta}_i, i = 1, 2\right)$ is the difference in mean cholesterol for the two treatments.

(i) Compute the minimum variance estimate of the overall treatment effect, $\hat{\theta}_{MV}$, and determine its 95% confidence interval stating any assumptions you make.

Study (Date of Publication)	Reduction in cholesterol $\hat{\theta}_i$	$Var\left[\hat{\theta}_i\right]$
Rahman (2001)	0.36	0.2916
Chung (2008)	0.68	0.1156

[6 marks]

(ii) What can you conclude from the meta-analysis regarding the performance of the new drug?

[2 marks]

[Total 8 marks]

B1.

(i) For a binary measure Y let π and p be the population and sample proportions respectively.

Suppose $\gamma = \arcsin\left(\sqrt{\pi}\right)$ and $\hat{\gamma} = \arcsin\left(\sqrt{p}\right)$. Given that $\frac{d\gamma}{d\pi} = \frac{1}{2\sqrt{\pi(1-\pi)}}$, use the approximation $Var\left[f(x)\right] \cong f'(x)^2\Big|_{x=E[x]} Var[x]$, to show that $Var[\hat{\gamma}] \cong \frac{1}{4n}$ where n is the sample size. [6 marks]

(ii) Consider a parallel group trial with two treatment groups of size n_T and n_C, and a binary outcome measure. Suppose π_T, π_C, p_T and p_C are the population and sample proportions for each treatment. With the treatment effect defined by $\hat{\tau} = \arcsin\left(\sqrt{p_T}\right) - \arcsin\left(\sqrt{p_C}\right)$, show that $SE[\hat{\tau}] = \sqrt{\frac{1}{4n_T} + \frac{1}{4n_C}}$. [4 marks]

(iii) For a normally distributed test statistic $T = \frac{\hat{\tau}}{SE[\hat{\tau}]}$ the power to detect a difference τ_D with a two-sided α size test is given by the expression $Power = 1 - \Phi\left(z_{\alpha/2} - \frac{\tau_D}{SE[\hat{\tau}]}\right)$. For a test of $H_0 : \pi_T = \pi_C$ vs $H_1 : \pi_T \neq \pi_C$, show that the power to detect a difference between two proportions π_T and π_C can be estimated by $1 - \Phi\left(z_{\alpha/2} - \frac{\arcsin\left(\sqrt{\pi_T}\right) - \arcsin\left(\sqrt{\pi_C}\right)}{\sqrt{\frac{1}{4n_T} + \frac{1}{4n_C}}}\right)$, stating any required assumptions. [4 marks]

(iv) Suppose it is required to design a trial with power $(1\text{-}\beta)$ to detect a difference between π_T and π_C with a two-sided α level test. Show that with two equal size groups, the sample size in each group must be

$$n = \frac{\left(z_{\alpha/2} + z_{\beta}\right)^2}{2\left(\arcsin\left(\sqrt{\pi_T}\right) - \arcsin\left(\sqrt{\pi_C}\right)\right)^2}$$

to achieve the desired power. [6 marks]

[Total 20 marks]

B2.

In a parallel group *non-inferiority* trial a new treatment *T* is being compared with a control treatment *C* using a normally distributed outcome measure *Y*. Assume that large values of *Y* represent a worse outcome for the patient. Let μ_T, $\bar{y}_T$, μ_C $\bar{y}_C$, and n_C be the population means, the sample mean and the sample size for each treatment. Suppose σ is the population standard deviation of both treatments. The treatment effect is defined as $\tau = \mu_T - \mu_C$.

(i) Explain why a significance test of the hypothesis $H_0 : \tau = 0$ vs $H_1 : \tau > 0$ would not be appropriate in a *non-inferiority* trial. [3 marks]

(ii) Suppose that the null hypothesis $H_0 : \mu_T - \mu_C \geq \tau_N$ is rejected if the upper (1-α) single sided confidence interval for $\hat{\tau}\left(= \bar{y}_T - \bar{y}_C\right)$ is less than the limit of non-inferiority τ_N. Stating any assumptions, show that $\Pr\left[\text{Reject } H_0 \middle| \tau\right] = \Phi\left(\left(\frac{\tau_N}{\sigma\lambda} - z_\alpha\right) - \frac{\tau}{\sigma\lambda}\right)$,

where Φ is the cumulative distribution function of the standard normal distribution and $\lambda = \sqrt{1/n_T + 1/n_C}$. [5 marks]

(iii) Show that $\Pr\left[\text{Reject } H_0 \middle| \tau\right]$ has a maximum under H_0 when $\tau = \tau_N$. Hence, show that this procedure has a type I error less than or equal to α. [7 marks]

(iv) A randomised controlled non-inferiority trial is carried out to test whether a *New* drug is as effective as a current *Standard* drug for controlling pain. Outcome is measured on a continuous scale with high scores representing greater pain. Eighty patients are randomised to the *New* treatment and 78 to the *Standard* treatment. The statistical output is given below.

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     New |      80        45.9     2.62738        23.5    40.67033    51.12967
Standard |      78        46.1    2.581592        22.8    40.95939    51.24061
---------+--------------------------------------------------------------------
    diff |                -.2     3.68486              -7.478658    7.078658
------------------------------------------------------------------------------
    diff = mean(New) - mean(Standard)                             t =  -0.0543
Ho: diff = 0                                     degrees of freedom =      156

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.4784         Pr(|T| > |t|) = 0.9568          Pr(T > t) = 0.5216
```

A difference of 5 mm was considered by researchers to be clinically important difference between treatments. Use the procedure described in part (ii) with a 5% significance level to test whether the *New* drug is non-inferior to the *Standard* drug. [5 marks]

[Total 20 marks]

B3.

For an AB/BA crossover trial a model for a continuous outcome y_{ij} of the i^{th} patient in the j^{th} period can be written as

$y_{i1} = \mu + \tau + \xi_i + \varepsilon_{i1}$ for a patient in sequence AB in period 1,

$y_{i2} = \mu + \phi + \gamma + \xi_i + \varepsilon_{i2}$ for a patient in sequence AB in period 2,

$y_{i1} = \mu + \xi_i + \varepsilon_{i1}$ for a patient in sequence BA in period 1,

$y_{i2} = \mu + \tau + \phi + \xi_i + \varepsilon_{i2}$ for a patient in sequence BA in period 2,

where μ is the mean for the sequence BA in period 1 , τ is the treatment effect of A relative to B, ϕ is the effect of the second period relative to the first, γ is the carryover effect, ξ_i is a random variable representing patient *i* with mean zero and variance σ_B^2, and ε_{ij} is the error term for patient *i* in period *j* assumed to be normally distributed with mean zero and variance σ_ε^2. Let $d_i = y_{i2} - y_{i1}$ and let $\bar{d}_{AB}$, μ_{AB}, $\bar{d}_{BA}$ and μ_{BA} be the sample and population means of these for sequences AB and BA respectively.

(i) Explain what is meant by the term *carryover* effect.

[2 marks]

(ii) In a crossover trial the treatment effect τ is estimated by $\hat{\tau} = \left(\bar{d}_{BA} - \bar{d}_{AB}\right)/2$. Show that this will be biased if there is a carryover effect.

[4 marks]

(iii) Let $a_i = y_{i2} + y_{i1}$ and $\bar{a}_{AB}, \mu_{AB}^A, \bar{a}_{BA}$ and μ_{BA}^A be the sample and population means for sequences AB and BA respectively. Show that $E\left[\bar{a}_{AB} - \bar{a}_{BA}\right] = \gamma$. [4 marks]

(iv) Show that $Var\left[\bar{a}_{BA} - \bar{a}_{AB}\right] = \left(4\sigma_B^2 + 2\sigma_\varepsilon^2\right)\left(\frac{1}{n_{BA}} + \frac{1}{n_{AB}}\right)$ [4 marks]

(v) The test statistic T_a, defined as $T_a = \frac{\bar{a}_{AB} - \bar{a}_{BA}}{\hat{SE}\left[\bar{a}_{AB} - \bar{a}_{BA}\right]}$, has been suggested as a test of the hypothesis H_0: $\gamma = 0$ vs H_1: $\gamma \neq 0$. What is the limitation of this as a test for carryover effect?

[2 marks]

(vi) What are the implications of parts (ii) and (iv) for the use of the crossover trial design?

[2 marks]

(vii) Give ONE possible means of preventing a carryover effect in a randomised controlled crossover trial. [2 marks]

[Total 20 marks]

END OF EXAMINATION PAPER