MATH38071

Two Hours

UNIVERSITY OF MANCHESTER

Statistical tables are attached

Medical Statistics

24 January 2013

14:00 – 16:00

Answer **ALL** five questions in **SECTION A** (40 Marks)

Answer **TWO** of the three questions in **SECTION B** (20 marks each)

The total number of marks on the paper is 80.

Electronic calculators may be used provided that they cannot store text

---

SECTION A

Answer **ALL** seven questions

SECTION A

A1.

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

(ii) Describe two ways in which a trial being *double blind* might reduce bias.

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

[5 marks]

A2.

A randomized controlled trial is carried out to compare a new treatment regime (N) with the existing standard treatment (S) for patients with tuberculosis. 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	Yes	50	30
	No	200	220
Total		250	250

(i) Estimate the odds ratio for the patient being infectious after 2 weeks with the new treatment (N) compared to the standard treatment (S).

(ii) Calculate the 95% confidence interval of this odds ratio.

(iii) Compare the two treatments using a z-test of proportions calculating the p-value.

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

[14 marks]

A3.

A randomised controlled trial is being designed to compare two treatments with a normally distributed primary outcome measure. The power to demonstrate a treatment effect τ with a two-sided two sample t-test can be estimated by the approximation $1-\Phi\left(z_{\alpha/2}-\frac{\tau\sqrt{n}}{\sigma\sqrt{2}}\right)$ where σ is the within treatment group standard deviation, n is the sample size of each of two equal size groups, and Φ is the cumulative density function of a standardised normal distribution. Suppose one wishes to detect a treatment effect of 2 units with a 5% significance level and the within treatment group standard deviation has been estimated to be 7 units,

(i) Estimate the power of a trial with 98 subjects in each treatment group.

(ii) Assuming equal size groups determine the minimum total sample size required to obtain 80% power.

[9 marks]

A4.

In a trial comparing an acupuncture treatment (A) with a homeopathic treatment (H) for patients suffering chronic headaches, patients are allocated to treatment using deterministic minimization controlling for sex and type of head ache (migraine, tension). The numbers of patients with each characteristic for each treatment are given in the table below after twenty-five patients have entered the trial.

Patient Characteristic	Male		Female		Migraine		Tension	
Treatment	(A)	(H)	(A)	(H)	(A)	(H)	(A)	(H)
Number of Patients	8	5	5	7	7	6	6	6

(i) How many patients have been allocated to each treatment?

(ii) The characteristics of the next two patients to enter the trial are:

26^{th} (Male, Migraine)

27^{th} (Female, Migraine)

Determine the treatment allocation of each patient.

[5 marks]

A5.

(i) Explain what is meant by the term *publication bias*.

(ii) Give two possible causes of *publication bias*.

(iii) How might one investigate *publication bias* graphically?

[7 marks]

B1.

For a parallel group randomised controlled trial comparing a control treatment (C) with a new treatment (T) suppose Y is a continuous normally distributed outcome variable and X is the value of the same variable recorded prior to randomisation. Suppose that τ is the treatment effect such that:

$Y = \mu_y + \varepsilon_y$ and $X = \mu_x + \varepsilon_x$ for treatment C

$Y = \mu_y + \tau + \varepsilon_y$ and $X = \mu_x + \varepsilon_x$ for treatment T

with $E[\varepsilon_x] = E[\varepsilon_y] = 0$, $Var[\varepsilon_y] = \sigma_y^2$, $Var[\varepsilon_x] = \sigma_x^2$, and $Cov[\varepsilon_x, \varepsilon_y] = \sigma_{xy}$.

Suppose that $\bar{x}_T$, $\bar{x}_C$, $\bar{y}_T$, and $\bar{y}_C$, are the sample means of X, and Y for each treatment.

Define $\hat{\tau}(\theta) = (\bar{y}_T - \theta\bar{x}_T) - (\bar{y}_C - \theta\bar{x}_C)$.

(i) Show that $E[\hat{\tau}(\theta)] = \tau$.

[4 marks]

(ii) Show that $Var[\hat{\tau}(\theta)] = \lambda^2(\sigma_y^2 + \theta^2\sigma_x^2 - 2\theta\sigma_{xy})$ where $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$, n_T is the numbers of patients allocated to the new treatment and n_C is the number allocated to the control treatment.

[7 marks]

(iii) Show that $Var[\hat{\tau}(\theta)]$ has a minimum when $\theta = \frac{\sigma_{XY}}{\sigma_X^2}$.

[4 marks]

(iv) In this setting three statistical analyses might be used to estimate and test the treatment effect:

a) an unadjusted analysis using just the outcome variable Y,

b) an analysis based on the change score Y-X or

c) a linear model of the outcome variable Y with treatment group and X as covariates.

What are the implications of the results in (i) and (iii) for the choice between the three statistical analyses?

[3 marks]

(v) Why is it important for randomised controlled trials to have a statistical analysis plan?

[2 marks]

[Total 20 marks]

B2.

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-treatment group sample standard deviation of Y. Define the treatment effect $\tau = \mu_T - \mu_C$ and $\hat{\tau} = \bar{y}_T - \bar{y}_C$.

Suppose that the null hypothesis $H_0 : |\tau| \geq \tau_E$ is rejected if the (1-2α) confidence interval given by $(\hat{\tau} - t_\alpha(\nu)s\lambda, \hat{\tau} + t_\alpha(\nu)s\lambda)$ is within the interval $(-\tau_E, \tau_E)$, where $\lambda = \sqrt{1/n_T + 1/n_C}$, and $t_\alpha(\nu)$ is the value of the t-distribution with $\nu = n_T + n_C - 2$ degrees of freedom having cumulative probability equal to $(1-\alpha)$.

(i) Show that $\Pr[\text{Reject } H_0] \cong \Phi\left(\frac{\tau_E - z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right) - \Phi\left(\frac{-\tau_E + z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right)$, where Φ is the cumulative distribution function of the standard normal distribution.

[6 marks]

(ii) Show that the sample size required in each treatment group to demonstrate equivalence with a power $(1-\beta)$ is

$$n = \frac{2\sigma^2}{\tau_E^2}\left(z_\alpha + z_{\beta/2}\right)^2.$$

[7 marks]

(iii) Suppose the interval [-2, 2] is to be used as the range of equivalence and the within treatment group standard deviation has been estimated to be 4. Determine the sample size per group required to obtain 90% power.

[4 marks]

(iv) Explain why patient compliance to treatment is particularly important in an equivalence trial.

[3 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

$$y_{i1} = \mu + \tau + \xi_i + \varepsilon_{i1} \quad \text{for a patient in sequence AB in period 1,}$$
$$y_{i2} = \mu + \phi + \xi_i + \varepsilon_{i2} \quad \text{for a patient in sequence AB in period 2,}$$
$$y_{i1} = \mu + \xi_i + \varepsilon_{i1} \quad \text{for a patient in sequence BA in period 1,}$$
$$y_{i2} = \mu + \tau + \phi + \xi_i + \varepsilon_{i2} \quad \text{for a patient in sequence BA in period 2,}$$

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 period 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. Defining $d_i = y_{i2} - y_{i1}$ let $\bar{d}_{AB}$, μ_{AB}^d, $\bar{d}_{BA}$ and μ_{BA}^d be the sample and population means for sequences AB and BA respectively.

(i) Show that $\left(\bar{d}_{BA} - \bar{d}_{AB}\right)/2$ is an unbiased estimator of the treatment effectτ.

[3 marks]

Two drugs used to treat chronic heart-burn were compared in a randomised controlled crossover trial. Eight patients were allocated to the sequence drug A then drug B and eleven patients were allocated to the sequence drug B then drug A. Outcome is assessed at the end of each period using a continuous normally distributed measure of acid-reflux with higher scores representing a worse outcome for the patient. The computer output below gives the sample mean and standard deviation for each sequence and period and the results of a two-sample t-test based on the difference in outcome d_i defined above.

```
-------------------------------------------------
Sequence    Period 1          Period 2
            mean  s.d.  n     mean  s.d.  n
-------------------------------------------------
AB          4.73  0.67  8     4.51  0.65  8
BA          4.92  0.79  11    4.41  0.82  11
-------------------------------------------------
Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
      AB |       8       -0.22    0.1555635    0.44      -0.2478492   0.4878492
      BA |      11       -0.51    0.1537708    0.51      -0.7426227  -0.0573773
---------+--------------------------------------------------------------------
    diff |              -0.28    0.2241559               -0.7529276   0.1929276
------------------------------------------------------------------------------
    diff = mean(BA) - mean(AB)                                    t =  -1.2491
Ho: diff = 0                                     degrees of freedom =       17

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.1143         Pr(|T| > |t|) = 0.2285          Pr(T > t) = 0.8857
------------------------------------------------------------------------------
```

[Continued]

(ii) Using the computer output determine the treatment effect (τ) and its 95% confidence interval.

[3 marks]

(iii) Define $c_i = y_{i1} - y_{i2}$ for sequence AB and $c_i = y_{i2} - y_{i1}$ for sequence BA. Let μ^c_{AB} μ^c_{BA}, $\bar{c}_{AB}$ and $\bar{c}_{BA}$ be the population and sample means of these for sequences AB and BA respectively. Show that a test of the null hypothesis $H_0 : \mu^c_{AB} = \mu^c_{BA}$ is the same as a test of the period effect, $H_0 : \phi = 0$.

[3 marks]

(iv) Using the computer output, estimate the period effect and test the null hypothesis $H_0 : \phi = 0$

[4 marks]

(v) Briefly comment on the result of the trial.

[3 marks]

(vi) It is sometimes suggested that the treatment effect in a cross-over trial can be estimated by the overall sample mean of the differences c_i, say $\bar{c} = \frac{\sum_{i=1}^{N} c_i}{N}$, where N is the total number of subjects in the trial. Using the computer output estimate the treatment effect of drug A as compared to drug B using this method. Why does this estimate differ from that obtained in part (ii)?

[4 marks]

[Total 20 marks]

END OF EXAMINATION PAPER