MATH38071 EXAMINATION SOLUTION (JANUARY 2013)

SECTION A

A1.

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

Solution

This is method to reduce bias in a randomised controlled trial, where neither the study participant nor the experimenters knows which of two treatments the participant is receiving

[2 marks]

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

Solution

Two reasons from: It is advantageous for a trial to be double blind as knowledge of treatment allocation may influence (i) the behaviour of the patient, (ii) the behaviour of the treating health professional or (iii) the assessor of outcome. (i) For example, if the patient knows which treatment they are receiving, it may motivate them to default from treatment or seek alternative treatments. (ii) If the treating health professional know the allocation, it may influence choice of secondary treatments. (iii) If the outcome assessor is aware of treatment allocation, their judgement may be bias.

[2 marks]

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

Solution

An treatment involving surgery , talking or physical therapies.

[1 mark]

[Total 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).

Solution

The odds ratio for being infectious on the new treatment compared to the standard treatment is

OR= (30/220)/(50x200)= 0.5454545

[Calculation 1 Minutes]

[2 marks]

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

Solution

Confidence interval for log odds ratio of being infectious are calculated using

$$\log_e\left[\frac{r_T\left(n_C-r_C\right)}{\left(n_T-r_T\right)r_C}\right]\pm z_{\alpha/2}\sqrt{\frac{1}{r_T}+\frac{1}{n_T-r_T}+\frac{1}{r_C}+\frac{1}{n_C-r_C}}$$

log(OR)= -0.606

Substitution into the formula with cell frequencies gives

$$\mathrm{SE}\left[\log\left[\mathrm{OR}\right]\right]=\sqrt{\frac{1}{r_T}+\frac{1}{n_T-r_T}+\frac{1}{r_C}+\frac{1}{n_C-r_C}}$$

$$=\sqrt{\frac{1}{200}+\frac{1}{50}+\frac{1}{220}+\frac{1}{30}}=0.25075$$

95% c.i. $\log_e\left[OR\right]$ is therefore (-1.098,-0.117).

Taking exponentials gives the 95% c.i. for the odds ratio (OR) is (0.334, 0.889)

[Calculation 3 Minutes]

[5 marks]

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

Z can be assumed to be normally distributed provided $n_N p, n_N\left(1-p\right), n_P p$, $n_P\left(1-p\right)$ are greater than 5. From the table the smallest of these is 250 x 30 /500 =15. Hence the normal approximation can be used.

$$p_T-p_C=\frac{30}{250}-\frac{50}{250}=-0.08=0.2,\ \ p=\frac{80}{500}=0.16$$

$$SE_{null}\left[p_T-p_C\right]=\sqrt{p\left(1-p\right)\left(\frac{1}{n_T}+\frac{1}{n_C}\right)}=\sqrt{\frac{0.16\times0.84}{125}}=0.0328$$

$$Z=\frac{\left|p_T-p_C\right|}{SE_{null}\left[p_T-p_C\right]}=\frac{0.08}{0.0328}=2.439750$$

From tables p=0.0146. This is statistically significant at a 5% level.

[Calculation 4 Minutes]

[5 marks]

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

The trial suggests that the new treatment is more effect than the standard treatment as the odds ratio for being infectious is 0.545 with 95% confidence interval (0.334, 0.889). The null hypothesis that the odds ratio equals 1 has also been rejected (p=0.0146).

[2 marks]

[Total 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.

Solution

$\tau = 2 \quad \sigma = 7 \quad z_{\alpha/2} = 1.96$

$$1-\Phi\left(z_{\alpha/2}-\frac{\tau\sqrt{n}}{\sigma\sqrt{2}}\right)=1-\Phi\left(1.96-\frac{2\sqrt{98}}{7\sqrt{2}}\right)=1-\Phi(-0.04)=\Phi(0.04)=\Phi(0.54)=0.5160$$

The power of the study is 51.6%

[Calculation 3 Minutes]

[4 marks]

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

Solution

$$0.8=1-\Phi\left(z_{\alpha/2}-\frac{\tau\sqrt{n}}{\sigma\sqrt{2}}\right)$$

$$\Phi\left(z_{\alpha/2}-\frac{2\sqrt{n}}{7\sqrt{2}}\right)=0.2$$

$$z_{\alpha/2}-\frac{2\sqrt{n}}{7\sqrt{2}}=-z_{0.2}$$

$$\sqrt{n}=(0.842+1.960)\frac{7}{\sqrt{2}}$$

$$n=2.802^2\times\frac{49}{2}=192.35$$

The number of subjects in each group to achieve 80% power is 193

[Calculation 4 Minutes]

[5 marks]

[Total 9 marks]

A4. In a trial comparing an acupuncture treatment (A) with a homeopathic treatment (H) for patients suffering from 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?

13 Acupuncture, 12 Homeopathic patients

[1 mark]

(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.

Table with updated totals

		Male		Female		Migraine		Tension		total		Allocated treatment
		(A)	(H)	(A)	(H)	(A)	(H)	(A)	(H)	(A)	(H)	
25		8	5	5	7	7	6	6	6	13	12	
26	Male,Migraine	8	**6**	5	7	7	**7**	6	6	8+7=15	5+6=11	(H)
27	Female, Migraine	8	7	**6**	7	**8**	7	6	7	5+7=12	7+6=13	(A)

The 26^{th} patient is allocated to <u>Homeopathic</u>,

The 27^{th} patient is allocated to <u>Acupuncture.</u>

[Calculation 3 Minutes]

[4 marks]

[Total 5 marks]

A5.

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

Solution

Studies in which an intervention is not found to have a statistically significant effect are sometimes never published whereas studies that demonstrate an effect may be more likely to be published. This means that the published literature may be unrepresentative leading to bias in meta-analysis estimates of treatment effects.

[2 marks]

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

Solution

Possible cause of publication bias could include:

Selective publications: Journal more likely to accept paper with statistically significant results as they may be perceived as more interesting.

Selective reporting: Where studies have multiple outcomes measured, statistically significant results may be emphasized in reports whereas non-significant results may be given less prominence or left out.

Identification: Studies in which results are statistically significant are likely to be published in more prestigious, and hence more easily accessible.

[2 marks]

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

Solution

Publication bias is more likely for small studies than large studies. One way of investigating whether publication bias is present is to construct a funnel plot in which the treatment effect of each study is plotted against the sample size. If there is no publication bias the plot will be symmetric about the line representing the overall treatment effect. If there is publication bias, the funnel plot will be asymmetric about this line with treatment effects in smaller studies tending to be larger than in bigger studies. Answer will probably give a sketch graph illustration.

[3 marks]

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

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 $E\left[\hat{\tau}(\theta)\right] = \tau$.

Solution

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

Therefore $E\left[\hat{\tau}(\theta)\right] = E\left[\bar{y}_T - \bar{y}_C\right] - \theta E\left[\bar{x}_T - \bar{x}_C\right]$

Randomisation means that $E\left[\bar{x}_T\right] = E\left[\bar{x}_C\right]$.

Therefore $E\left[\hat{\tau}(\theta)\right] = E\left[\bar{y}_T - \bar{y}_C\right] = \mu_T - \mu_C = \tau$ as required.

[Book Work]

[4 marks]

(ii) Show that $Var\left[\hat{\tau}(\theta)\right] = \lambda^2\left(\sigma_y^2 + \theta^2\sigma_x^2 - 2\theta\sigma_{xy}\right)$ 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.

Solution

Consider $\hat{\tau}(\theta) = (\bar{y}_T - \theta\bar{x}_T) - (\bar{y}_{\boldsymbol{C}} - \theta\bar{x}_C)$

$Var\left[\hat{\tau}(\theta)\right] = Var\left[(\bar{y}_T - \bar{y}_C) - \theta(\bar{x}_T - \bar{x}_C)\right]$

$= Var\left[\bar{y}_T - \bar{y}_C\right] + Var\left[\theta\left((\bar{x}_T - \bar{x}_C)\right)\right] - 2Cov\left[\bar{y}_T - \bar{y}_C, \theta(\bar{x}_T - \bar{x}_C)\right]$

$Var\left[\bar{y}_T - \bar{y}_C\right] + \theta^2 Var\left[\bar{x}_T - \bar{x}_C\right] - 2.\theta.Cov\left[\bar{y}_T - \bar{y}_C, \bar{x}_T - \bar{x}_C\right]$ [1]

Considering the first term $Var[\bar{y}_T - \bar{y}_C] = Var[\bar{y}_T] + Var[\bar{y}_C] - 2Cov[\bar{y}_T, \bar{y}_C]$

Since treatment groups are independent, $Cov[\bar{Y}_T, \bar{Y}_C] = 0$.

Therefore $Var[\bar{y}_T - \bar{y}_C] = Var[\bar{y}_T] + Var[\bar{y}_C]$.

Since observations are independent $Var[\bar{y}_T] = \frac{\sum_{i \in T} Var[y_T]}{n_T^2} = \frac{\sum_{i \in T} \sigma_Y^2}{n_T^2} = \frac{\sigma_Y^2}{n_T}$.

Similarly $Var[\bar{y}_C] = \frac{\sigma_Y^2}{n_C}$.

Therefore $Var[\bar{y}_T - \bar{y}_C] = \lambda^2 \sigma_Y^2$ where $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$.

Similarly $Var[\bar{x}_T - \bar{x}_C] = \lambda^2 \sigma_X^2$ and $Cov[\bar{y}_T - \bar{y}_C, \bar{x}_T - \bar{x}_C] = \lambda^2 \sigma_{XY}$.

Substitution into [1] gives $Var[\hat{\tau}(\theta)] = \lambda^2 (\sigma_y^2 + \theta^2 \sigma_X^2 - 2\theta \sigma_{XY})$.

[Book Work]

[7 marks]

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

Solution

Differentiation with respect to θ gives

$$\frac{\partial}{\partial \theta} Var[\hat{\tau}(\theta)] = \lambda^2 (2\theta \sigma_x^2 - 2\sigma_{xy}).$$

This equals zero when $\theta = \sigma_{xy}/\sigma_x^2$.

The second derivative $\frac{\partial^2}{\partial \theta^2} Var[\hat{\tau}(\theta)] = 2\lambda^2 \sigma_x^2$.

As this is positive , it follows that $Var[\hat{\tau}(\theta)]$ has a minimum when $\theta = \frac{\sigma_{XY}}{\sigma_X^2}$.

[4 marks]
[Book Work]

(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?

Solution

Values of θ equal to 0 , 1 and when $\theta = \frac{\sigma_{XY}}{\sigma_X^2}$ correspond to the treatment effect in an unadjusted, change and linear adjusted model analyses. All three estimates are unbiased, but an estimate of the treatment effect based on a linear model smaller variance compared to an unadjusted analysis or a change analysis. Reducing the variance of the treatment effect estimate increases the power of the analysis. As a consequence if a baseline variable is thought to be correlated with outcome, an analysis adjusting for baseline is recommended, and where the baseline value of the outcome is recorded a linear model analysis is superior to an analysis based on change.

[3 marks]

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

Solution

A statistical analysis should be prepared prior to beginning the analysis. There can be many different ways in which the outcome from a randomised controlled trial can be analysed. For example 3 possible analysis were considered in (iv). These analyses may give results that are more or less supportive of the investigators' opinion . Unless the analysis is pre-specified the investigators may present that which is closest to their opinion, instead of the best analysis.

[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 $\left(\hat{\tau} - t_\alpha(\nu)s\lambda, \hat{\tau} + t_\alpha(\nu)s\lambda\right)$ 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.

Solution

The distribution of $\hat{\tau}/s\lambda$ has a non-central t-distribution if $\tau \neq 0$. Assuming the variance σ, is known, so $\hat{\tau}$ has a distribution $N\left[\tau, \sigma^2\lambda^2\right]$ and one can replace $t_\alpha(\nu)$ with z_α.

Type 1 error = $\Pr[\text{Reject } H_0]$ under H_0.

$$\begin{aligned}
&\Pr[\text{Reject } H_0] = \Pr\left[\left(\hat{\tau} - t_\alpha(\nu)SE[\hat{\tau}], \hat{\tau} + t_\alpha(\nu)SE[\hat{\tau}]\right) \subseteq (-\tau_E, \tau_E)\right] \\
&\cong \Pr\left[\left(\hat{\tau} - z_\alpha\sigma\lambda, \hat{\tau} + z_\alpha\sigma\lambda\right) \subseteq (-\tau_E, \tau_E)\right] \\
&= \Pr\left[\left(\hat{\tau} - z_\alpha\sigma\lambda > -\tau_E\right) \bigcap \left(\hat{\tau} + z_\alpha\sigma\lambda < \tau_E\right)\right] \\
&= \Pr\left[\left(\hat{\tau} > -\tau_E + z_\alpha\sigma\lambda\right) \bigcap \left(\hat{\tau} < \tau_E - z_\alpha\sigma\lambda\right)\right] \\
&= \Pr[\hat{\tau} < \tau_E - z_\alpha\sigma\lambda] - \Pr[\hat{\tau} < -\tau_E + z_\alpha\sigma\lambda] \\
&= \Phi\left(\frac{\tau_E - z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right) - \Phi\left(\frac{-\tau_E + z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right)
\end{aligned}$$

[Book Work]

[6 marks]

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

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

Solution

Using $\Pr[\text{Reject } H_0] \cong \Phi\left(\frac{\tau_E - z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right) - \Phi\left(\frac{-\tau_E + z_\alpha\sigma\lambda - \tau}{\sigma\lambda}\right)$

Under the alternate hypothesis assume τ=0, so the power

$$1-\beta=\Phi\left(\tau_E/\sigma\lambda-z_\alpha\right)-\Phi\left(-\tau_E/\sigma\lambda+z_\alpha\right)$$

Since $\Phi\left(-\tau_E/\sigma\lambda+z_\alpha\right)=1-\Phi\left(\tau_E/\sigma\lambda-z_\alpha\right)$ it follows that

$$1-\beta=2\Phi\left(\tau_E/\sigma\lambda-z_\alpha\right)-1.$$

Rearrangement gives

$$1-\beta/2=\Phi\left(\tau_E/\sigma\lambda-z_\alpha\right).$$

Since $\Phi^{-1}\left(1-\beta/2\right)=z_{\beta/2}$, it follows that $z_{\beta/2}=\dfrac{\tau_E}{\sigma\lambda}-z_\alpha$.

Hence $\dfrac{\tau_E}{\sigma\lambda}=z_\alpha+z_{\beta/2}$

Assuming equal sample size $n_T = n_C = n$, then $\lambda=\sqrt{\dfrac{2}{n}}$.

Hence $\sqrt{\dfrac{n}{2}}=\dfrac{\sigma}{\tau_E}\left(z_\alpha+z_{\beta/2}\right).$

Rearrangement gives

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

[Book Work]

[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.

Solution

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

Substitution with δ=2 σ=4 and α=0.05 β=0.1.

From table $z_\alpha = z_{0.05}$=1.645 and $z_{\beta/2} = z_{0.05}$=1.645

Substitution gives

$$n=\frac{2\sigma^2}{\tau_E^2}\left(z_\alpha+z_{\beta/2}\right)^2=\frac{2\times4^2}{2^2}\left(1.645+1.646\right)^2=86.59$$

Therefore the minimum sample size is 87 in each group.

[Calculation 3 Minutes]
[4 marks]

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

Solution

If patients do not comply with treatment or switch treatments in a clinical trial, the mean outcome for the treatment groups will tend to closer together than if patients complied with treatment. In a superiority trial this is a bias against rejecting the null hypothesis. If the null hypothesis is still rejected, one can be confident of the result. In an equivalence trial this is a bias towards the alternative hypothesis and may lead to erroneous rejection of the null.

[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

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

$y_{i2} = \mu + \phi + \xi_i + \varepsilon_{i2}$ 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 period effect, ξ_i is a random variable representϕ is the period effect, ing 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 effect τ.

Solution

Let $\bar{d}_{AB} = \dfrac{\sum d_i}{n_{AB}}$, μ_{AB} , $\bar{d}_{BA} = \dfrac{\sum d_i}{n_{BA}}$ and μ_{BA}, the sample and population means for sequences AB and BA.

$E\left[\bar{d}_{AB}\right] = E\left[\dfrac{\sum d_i}{n_{AB}}\right] = \phi - \tau$ and $E\left[\bar{d}_{BA}\right] = E\left[\dfrac{\sum d_i}{n_{BA}}\right] = \phi + \tau$.

Therefore $E\left[\bar{d}_{BA} - \bar{d}_{AB}\right] = 2\tau$.

Hence $\hat{\tau} = \dfrac{\bar{d}_{BA} - \bar{d}_{AB}}{2}$ is an unbiased estimator of τ.

[Book Work]

[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
------------------------------------------------------------------------------
```

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

Solution

From the printout $\bar{d}_{BA} - \bar{d}_{AB}$ equals -0.28

$$\hat{\tau} = \frac{\bar{d}_{BA} - \bar{d}_{AB}}{2} = \frac{-0.28}{2} = -0.14 \ .$$

From the printout the 95% c.i. is -0.7529276/2 to 0.1929276/2, which give the 95% c.i. of the treatment effect as -0.376 to 0.096.

[Calculation 2 Mins]

[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, $H_0 : \phi = 0$.

Solution

$$\mu^c_{AB} = E\left[y_{i1} - y_{i2}\right] = E\left[\left(\mu + \tau + \varepsilon_{i2}\right) - \left(\mu + \phi + \varepsilon_{i1}\right)\right] = \tau - \phi$$

$$\mu^c_{BA} = E\left[y_{i2} - y_{i1}\right] = E\left[\left(\mu + \phi + \tau + \varepsilon_{i2}\right) - \left(\mu + \varepsilon_{i1}\right)\right] = \phi + \tau$$

Therefore $\mu^C_{BA} - \mu^C_{AB} = 2\phi$.

Hence the test H_0: $\mu^c_{AB} = \mu^c_{BA}$ is equivalent to a test of the period effect $H_0 : \phi = 0$.

[Book Work]

[3 marks]

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

Solution

The hypothesis H_0: ϕ=0 vs. H_1: $\phi \neq 0$ can be tested using a two-sample t-test of the means of the differences $H_0 : \mu^c_{AB} = \mu^c_{BA}$.

Now $\mu^c_{AB} = -\mu^d_{AB}$ =0.22 and $\mu^c_{BA} = \mu^d_{BA}$ =-0.51

$$\hat{SE}\left[\bar{c}_{BA} - \bar{c}_{AB}\right] = \hat{SE}\left[\bar{d}_{BA} - \bar{d}_{AB}\right] = 0.229$$

$$\phi = \frac{\mu_{BA}^{C} - \mu_{AB}^{C}}{2} = \frac{-0.51-(-0.22)}{2} = -0.365$$

Period effect =-0.365

Therefore $T_C = \dfrac{\left|-0.51-(-0.22)\right|}{0.229} = 3.19$

From tables $t_{0.025}(16) = 2.12$ Therefore one can reject the null hypothesis that H_0: ϕ=0 at the 5% level.

[Calculation 3 Minutes]
[4 marks]

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

Solution

The treatment effect of drug A compared to drug B is -0.14 suggesting that patients receiving drug B have a slightly better outcome than those receiving drug A, but from the test of the hypothesis H_0: τ = 0 (p=0.2285) there is no evidence of a treatment effect. In contrast there is evidence of a period effect of -0.365 which was statistically significant (p<0.05) .

[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 $\overline{c} = \dfrac{\sum_{i=1}^{N} c_i}{N}$, 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)?

Solution

$$\overline{c} = \frac{\sum_{i=1}^{N} c_i}{N} = \frac{n_{AB} \times \overline{c}_{AB} + n_{BA} \times \overline{c}_{BA}}{n_{AB} + n_{BA}} = \frac{11 \times -0.51 + 8 \times 0.22}{19} = \frac{-3.85}{19} = -0.202\,.$$

[Calculation 2 Minutes]

The estimate differs from that in part (ii) due to the period effect which biases this estimator.

[4 marks]

[Total 20 marks]