Medical Statistics (MATH38071) Solutions Exercise Sheet 6
(Analysis with Baseline Data)

1.

(i) By examining the output suggest two biases that may affect the results of the trial.

Solution

There is evidence of bias in the allocation as there would appear to be some difference between the two treatment groups in their level of pain at base line (mean leaflet group = 51.6 , mean physiotherapy group= 44.9). Assuming randomisation was conducted correctly this can be taken as being *chance* bias. There is also evidence of *follow-up* bias as the percentage of patient in the physiotherapy group followed up to 9 months (91%, 105/116) is higher than in the leaflet group (83%, 98/118).

(ii) Suggest two ways in which the design could have been improved to prevent these biases.

Solution

a) The trial used simple randomisation. If instead the trial had used stratified randomisation or minimisation controlling for baseline pain, the imbalance at baseline would have been reduced, making the results of the trial more convincing.

b) More rigorous follow-up of patient might reduce the differential follow-up rate thereby reducing follow-up bias.

(iii) Write down an estimate of the treatment effect of physiotherapy as compared to self-help booklet based on (a) an unadjusted analysis (b) an analysis adjusted for baseline pain, giving the 95% confidence interval and the p-value for the test of the null hypothesis of no treatment effect .

Solution

The point estimate, the confidence interval and p-value of the treatment effect adjusted for baseline are –8.93(95% c.i. -16.12 to -1.73 ,p=0.015)

The point estimate, the confidence interval and p-value of the treatment effect adjusted for baseline are -4.18(95% c.i. -10.94 to 2.59,p=0.225).

(iv) Briefly comment on the results of these analyses?

Solution

Two estimates of the treatment effect have been determined and are contradictory. First, the unadjusted analysis that suggests a benefit of a reduction in pain(- 8.93mm , 95% c.i. -16.12 to -1.73 ,p=0.015), and secondly an adjusted analysis that fails to show a benefit (-4.18,95% c.i. -10.94 to 2.59,p=0.225). There is more than five points difference between the two estimates that can be explained by the difference at baseline. As discussed in the notes the two analyses estimate the same effect on average, but an analysis adjusted for baseline can be expected to be more precise. Provide the assumption of the procedure are

satisfied I would favour just presenting the adjusted analysis from which the conclusion would be that there is insufficient evidence to justify rejecting the null hypothesis of no treatment effect.

2. A researcher has carried out a randomised controlled trial to compare a new treatment (A) with a standard treatment (B) for patients with depression. A depression score has been recorded at baseline (**bprsbase**) and at follow-up (**bprsfu**). Lower scores represent less depression. The researcher generates the printout output listed below from the data.

Results for group = A

Paired T-Test and CI: bprsbase, bprsfu

```
Paired T for bprsbase - bprsfu

                  N      Mean     StDev   SE Mean
bprsbase         23     24.35      6.87      1.43
bprsfu           23     20.87      8.45      1.76
Difference       23      3.48      7.10      1.48

95% CI for mean difference: (0.41, 6.55)
T-Test of mean difference = 0 (vs not = 0):
T-Value = 2.35  P-Value = 0.028
```

Results for group = B

Paired T-Test and CI: bprsbase, bprsfu

```
Paired T for bprsbase - bprsfu

                  N      Mean     StDev   SE Mean
bprsbase         24     24.33      7.28      1.49
bprsfu           24     22.67      7.63      1.56
Difference       24      1.67      6.86      1.40

95% CI for mean difference: (-1.23, 4.56)
T-Test of mean difference = 0 (vs not = 0):
T-Value = 1.19  P-Value = 0.246
```

Because there is a statistically significant change at the 5% level from baseline to follow-up for group A but not in group B, the researcher concludes that treatment A is more effective than treatment B for treating depression.

(i) Explain the flaw in this conclusion.

Solution

Tests of within group change may not measure the effect of treatment. The statistically significant change observed for treatment A may not be due to factors other than treatment. For example, it may have occurred because the condition naturally resolves. Even if change within patient was due to treatment the investigators reasoning is flawed for because the p-values relate to two separate hypotheses test. Failure to reject the null hypothesis does not make the null hypothesis true and so does not imply no treatment effect.

(ii) Based on the information in the output two statistical analyses are possible (a) an unadjusted analysis using `bprsbase` and (b) a change score analysis using `bprsbase – bprsfu` . By inspecting the output explain why the analysis based on change should have greater power than the unadjusted analysis.

Solution

Both methods estimate the same treatment effect on average. If we examine the standard deviations the followup (`bprsbase`) and change (`bprsbase – bprsfu`) the latter has a smaller standard deviation for both groups and so will have a smaller standard error giving greater power to detect the same treatment effect.

(iii) Use the output to test whether treatment A is superior to treatment B using an analysis based on the change score, stating any assumption you make.

Solution

Note: Ideally one would use a linear model as this is the most efficient analysis but this would need the raw data rather than just the summary statistics.

The analysis based on the change score using a two sample t-test is below. It assumes

- patient outcomes are independent,
- that differences are normally distributed, and
- equal population standard deviation.

Summary data from the computer output for the two-sample t-test.

	n	$\bar{d}$	s.d.
Treatment A	23	3.48	7.10
Treatment B	24	1.67	6.86

$$T = \frac{\bar{d}_A - \bar{d}_B}{SE\left[\bar{d}_A - \bar{d}_B\right]} \text{ where } SE\left[\bar{d}_A - \bar{d}_B\right] = s\sqrt{1/n_A + 1/n_B} \text{ and } s = \sqrt{\frac{(n_A - 1).s_A^2 + (n_B - 1).s_B^2}{n_A + n_B - 2}}$$

$$s = \sqrt{\frac{(n_A - 1).s_A^2 + (n_B - 1).s_B^2}{n_A + n_B - 2}} = 6.978$$

$$SE\left[\bar{d}_A - \bar{d}_B\right] = s\sqrt{1/n_A + 1/n_B} = 2.036$$

$$T = \frac{\bar{d}_A - \bar{d}_B}{SE\left[\bar{d}_A - \bar{d}_B\right]} = \frac{(3.48 - 1.67)}{2.036} = 0.89$$

For a 5% size test H_0: $\tau = 0$ vs H_1: $\tau \neq 0$ compare T with $t_{0.025}(45)$. From statistical table $t_{0.025}(45) \approx 2.015$.

Hence, there is no evidence to reject the null hypothesis of no treatment effect.

3. For a parallel group trial comparing a control treatment (Group C) with a new intervention (Group T) suppose *y* is a continuous, normally distributed outcome variable and *x* is the value of the same variable recorded at baseline prior to randomisation. Suppose that τ is the treatment effect such that

$y = \mu + \varepsilon_y$, $x = \mu_x + \varepsilon_x$ Group C

$y = \mu + \tau + \varepsilon_y$, $x = \mu_x + \varepsilon_x$ Group 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 $Cov\left[\varepsilon_x, \varepsilon_y\right] = \sigma_{xy}$

Define $\hat{\tau}(\theta) = \left(\bar{Y}_T - \theta \bar{X}_T\right) - \left(\bar{Y}_C - \theta \bar{X}_C\right)$.

(i) Write down an expression for $Var\left[\hat{\tau}(\theta)\right]$

Solution

From the notes $Var\left[\hat{\tau}(\theta)\right] = \lambda^2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)$ where $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$.

(ii) Assume that $T = \hat{\tau}(\theta) / SE\left[\hat{\tau}(\theta)\right]$ has a normal distribution. A general expression for the power to detect a difference τ_S with a normally distributed test statistic with a two-sided α -size test is

$$\text{Power} = (1\text{-}\beta) = 1 - \Phi\left(z_{\alpha/2} - \frac{\tau_S}{SE[\tau]}\right),$$

Write down an expression for the power of a test statistic $T = \hat{\tau}(\theta) / SE\left[\hat{\tau}(\theta)\right]$ to detect a treatment effect τ_S.

$$1 - \beta \cong 1 - \Phi\left(z_{\alpha/2} - \frac{\tau_S}{\lambda\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}}\right).$$

Solution

Since $Var\left[\hat{\tau}(\theta)\right] = \lambda^2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)$ where $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$,

$SE\left[\hat{\tau}(\theta)\right] = \lambda\sqrt{\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)}$

Substitution into $\text{Power} = (1\text{-}\beta) = 1 - \Phi\left(z_{\alpha/2} - \frac{\tau_S}{SE[\tau]}\right)$ gives

$$\text{Power} = 1 - \beta \cong 1 - \Phi\left(z_{\alpha/2} - \frac{\tau_S}{\lambda\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}}\right).$$

(iii) Assuming that $n_T = n_C = n$, show that the formula for estimating sample size to detect an effect τ_S with the test statistic $T = \hat{\tau}(\theta) \big/ SE\left[\hat{\tau}(\theta)\right]$ and power $(1\text{-}\beta)$ using a two-sided α -size test is

$$n = \frac{2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

Solution

Rearrangement gives $\beta \cong \Phi\left(z_{\alpha/2} - \frac{\tau_S}{\lambda\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}}\right)$

Since $\Phi^{-1}(\beta) = -z_\beta$ it follows that $-z_\beta = z_{\alpha/2} - \frac{\tau_S}{\lambda\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}}$

giving $\frac{\tau_S}{\lambda\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}} = z_{\alpha/2} + z_\beta$.

If equal sized groups are assumed so that $n_T = n_C = n$, then $\lambda = \sqrt{2/n}$.

Substitution into [2] gives $\frac{\tau_S}{\sqrt{\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}}}\sqrt{\frac{n}{2}} = z_{\alpha/2} + z_\beta$.

Rearrangement gives $n = \frac{2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$ as required.

(iv) Show that the formula for estimating sample size for with an analysis based on change Y – X is

$$n = \frac{2\left(\sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

and for analysis based on a linear model adjusting for a baseline covariate *X* is

$$n = \frac{2\left(\sigma_Y^2\left(1 - \rho_{XY}{}^2\right)\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

Solution

$\theta = 1$ corresponds to change. Substitution into $n = \frac{2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$ with $\theta = 1$ gives

the result for change, $n = \frac{2\left(\sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$

$\theta = \beta$ corresponds to a a linear model. Substitution with $\beta = \sigma_{XY} \big/ \sigma_x^2$

$$n = \frac{2\left(\sigma_Y^2 + \theta^2\sigma_X^2 - 2\theta\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

$$= \frac{2\left(\sigma_Y^2 + \left(\frac{\sigma_{XY}}{\sigma_x^2}\right)^2\sigma_X^2 - 2\left(\frac{\sigma_{XY}}{\sigma_x^2}\right)\sigma_{XY}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

$$= \frac{2\left(\sigma_Y^2 - \frac{\sigma_{XY}^2}{\sigma_x^2}\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2 = \frac{2\sigma_Y^2\left(1-\rho^2\right)}{\tau_S^2}\left(z_{\alpha/2} + z_\beta\right)^2$$

4. In a trial of a new dietary intervention to reduce the blood cholesterol a new treatment is compared against a standard treatment. A 10mg/dl reduction (improvement) in cholesterol levels for the new treatment is considered to be the minimum that would be clinically important for patients. Baseline cholesterol data is being collected on each patient. From a previous trial the *within-group* standard deviation for cholesterol is estimated to be 60mg/dl at baseline and 50mg/dl at follow-up. The within-group correlation between baseline and follow-up measurements for cholesterol has been estimated to be 0.6. Assuming a 5% two-sided significance level, determine the minimum sample size per group to have 80% power for:

   (i) an unadjusted analysis,

   (ii) an analysis based on change scores, and

   (iii) an analysis based on linear model adjusting for baseline cholesterol.

Solution

The formula from question 3 may be applied. Assuming a 5% two tailed significance level $z_{\alpha/2} = 1.96$. If we chose a one-tailed test, we use $z_\alpha = 1.64$. $z_\beta = 0.84$. δ_M = 10mg/dl

(i) The basic formula can be used with $\sigma_Y = 50$. Assuming a two tailed test substitution into

$$n = \frac{2\left(\sigma_Y^2\right)}{\tau_M^2}\left(z_{\alpha/2} + z_\beta\right)^2 = \frac{2\left(50^2\right)}{100}\left(1.96 + 0.84\right)^2 = 392$$

giving a minimum sample size for the required power of 392 per group.

(ii) $\sigma_Y = 50$, $\sigma_X = 60$ and $r_{XY} = 0.6$. Hence $\sigma_{XY} = 1800$. Assuming a two tailed test substitution into $n = \frac{2\left(\sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}\right)}{\delta_M^2}\left(z_{\alpha/2} + z_\beta\right)^2 = \frac{2\left(50^2 + 60^2 - 2\times 1800\right)}{100}\left(1.96 + 0.84\right)^2 = 392$

giving a minimum sample size for the required power of 392 per group.

(iii) $n = \frac{2\left(50^2\left(1 - 0.36\right)\right)}{10^2}\left(1.96 + 0.84\right)^2 = 250.88$ giving a minimum sample size for the required power of 251 per group.