SECTION A

Answer **ALL** five questions

A1.

In the context of a clinical trial

(i) What is *a potential* outcome?

Solution (i)

Where a patient can be offered one of several treatments in a clinical trial, a potential outcome is the outcome should they be offered a particular treatment

[1 marks]

(ii) What is *a counter factual* outcome?

Solution(ii)

A counter factual outcome is the potential outcome for a treatment that the patient does not receive.

[1 marks]

(iii) By considering potential outcomes demonstrate why randomization justifies causal inference.

Solution (iii)

$Y_i(T)$ and $Y_i(C)$ are the potential outcomes for the i[th] patient. Let $\tau_i = Y_i(T) - Y_i(C)$ be the treatment effect for the i[th] patient. The expected effect of treatment is therefore,

$$\tau = E[\tau_i] = E[Y_i(T) - Y_i(C)]$$
$$= E\left[Y_i(T) - Y_i(C) \middle| i \in T\right].\Pr[i \in T] + E\left[Y_i(T) - Y_i(C) \middle| i \in C\right].\Pr[i \in C]$$
$$= \left(E\left[Y_i(T) \middle| i \in T\right].\Pr[i \in T] - E\left[Y_i(C) \middle| i \in T\right].\Pr[i \in T]\right)$$
$$+\left(E\left[Y_i(T) \middle| i \in C\right].\Pr[i \in C] - E\left[Y_i(C) \middle| i \in C\right].\Pr[i \in C]\right)$$

Randomisation means that

$$E\left[Y_i(T) \middle| i \in C\right] = E\left[Y_i(T) \middle| i \in T\right]$$

and $E\left[Y_i(C) \middle| i \in T\right] = E\left[Y_i(C) \middle| i \in C\right]$.

Therefore

$$\tau = \left(\mu_T.\Pr[i \in T] + \mu_T.\Pr[i \in C]\right) - \left(\mu_C.\Pr[i \in T] + \mu_C.\Pr[i \in C]\right)$$
$$= \mu_T - \mu_C$$

[5 Marks]

[Total 7 Marks]

A2.

In a published report of a randomised trial comparing a drug for reducing blood pressure in patients with a high systolic blood pressure, two groups of 25 patients were compared. The estimated mean difference in systolic blood pressure between the new treatment and the standard treatment was -7 mmHg (95% confidence interval -19.8 to 5.8 mmHg). The p-value for a two-sample t-test is 0.277. A 5 mmHg reduction in systolic blood pressure is considered to be a clinically worthwhile benefit.

(i) Comment on the results of the trial.

Solution (i)

The study was underpowered as it failed to show that a clinically important effect of 5 mmHg reduction was statistically significant conventional levels of significance or the 95% confidence interval includes both no effect and a clinically important effect. A larger sample size would be needs for a confirmatory trial.

[2 marks]

(ii) Use the data above to estimate the pooled within group standard deviation.

Solution (ii)

The formula for a 95% confidence interval for the difference of two means is given by $\bar{y}_1 - \bar{y}_2 - t_{\alpha/2}(\nu)\,se(\bar{y}_1 - \bar{y}_2)$ to $\bar{y}_1 - \bar{y}_2 + t_{\alpha/2}(\nu)\,se(\bar{y}_1 - \bar{y}_2)$ where $s.e.(\bar{y}_1 - \bar{y}_2) = s\sqrt{1/n_1 + 1/n_2}$ and $\nu = n_1 + n_2 - 2$. Considering the difference between the point estimate and either the upper or lower confidence interval $t_{\alpha/2}(\nu)\,se(\bar{y}_1 - \bar{y}_2) = 12.8$. From tables $t_{0.025}(48) = 2.0106$.

Therefore $se(\bar{y}_1 - \bar{y}_2) = 12.8 / 2.0106 = 6.366$. Now

$s = s.e.(\bar{y}_1 - \bar{y}_2) / \sqrt{1/n_1 + 1/n_2} = 6.366 / \sqrt{1/25 + 1/25} = 6.366 \times 5/\sqrt{2} = 22.5$ as an estimate of σ.

[Calculation 4 mins]

[5 marks]

(iii) A new trial is planned to compare the same two drugs. Using the value of the pooled within group standard deviation determined in (ii) as an estimate of the within group standard deviation σ, calculate the sample size required in each group to have a power equal to 80% to detect a reduction of 5mmHg in systolic blood pressure assuming a 5% two-sided significance level using the formula $n = \dfrac{2\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_{\beta}\right)^2$.

Solution(iii)

For α=0.05 from tables $z_{\alpha/2}$ =1.96.

For 80% power (1-β)=0.8. giving z_β=0.842 τ=5mg/dl σ=22.5.

Therefore sample size per group, $n = \frac{2\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_\beta\right)^2 = \frac{2\times 22.5^2}{25}\left(1.960 + 0.842\right)^2 = 317.97$.

Hence the minimum sample size per group is 318

[Calculation 4 mins]

[4 marks]

(iv) It is thought that about 20% of patients randomised will be lost to follow-up, and that only 60% of the patients that are screened for the study will be eligible of whom 50% will consent to join the trial. Estimate the total numbers of patients that need to be screened to obtain the sample size determined in (iii).

Solution(iv)

Total number of patients that needs to be randomized = 2 x 317.97/0.8=794.9

Total number that need to be screened =794.9/0.6/.5=2649.667Hence the number that need to be screened is 2650.

[Calculation 2 mins]

[2 marks]

[Total 13 marks]

A3.

(i) Show how you might prepare a randomisation list for first twenty patients in a trial with two treatments using *Block randomization*.

Solution (i)

With two treatments, say A and B, one could choose a block size of 4. With this block size there are 6 possible blocks (1) AABB (2) ABAB (3) ABBA (4) BBAA (5) BABA (6) BAAB

To assemble a randomization list for twenty subjects one would select 5 random numbers between 1- 6 with replacement in sequence, say the numbers 2, 6, 3, 1, 3 from which one could assemble the following list for the first 20 allocations in order A,B,A,B| B,A,A,B| A,B,B,A|A,A,B,B|A,B,B,A

[5 marks]

(ii) Gender is thought to be strongly predictive of the success of treatment with women having a better outcome than men. How you might you use block randomisation to improve balance in this characteristic between two treatment groups?

Solution (ii)

Block randomisation can be used in conjunction with stratification to obtain balance in a categorical prognostic factor. Separate block randomisation lists could be prepared for men and women.

[2 marks]

[Total 7 marks]

A4.

(i) Explain the difference between a non-inferiority trial and a superiority trial.

Solution (i)

Usually the aim of a trial is to detect a difference between the treatments under study, testing whether a new treatment is superior to the existing standard treatment or a placebo. Such trials are called *Superiority Trials*. In such a trial the null hypothesis is that the average outcome is the same. *Non-inferiority trials* are designed to establish that the efficacy of a test treatments is as good or better than the control. Therefore in such a trial the null hypothesis is that the test treatment is worse than the control.

[3 marks]

(ii) A randomised controlled non-inferiority trial is carried out to test comparing a *new* drug with the current *standard* drug for controlling pain. At follow-up this is measured by a 100 mm analogue scale with higher scores representing greater pain. Forty-nine patients are randomised to the new treatment and fifty one to the standard treatment. The statistical computer package output is given below. A difference of 5 mm was considered by researchers to be the minimum that was clinically important. Using the results in the output below, test whether the *new* drug is non-inferior to the *standard* drug specifying the significance level used.

```
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [90% Conf. Interval]
---------+--------------------------------------------------------------------
NEW      |      49        45.1    2.171429        15.2    41.45803    48.74197
STANDARD |      51        46.9    2.282457        16.3    43.07482    50.72518
---------+--------------------------------------------------------------------
combined |     100      46.018      1.5717      15.717    43.40836    48.62764
---------+--------------------------------------------------------------------
    diff |               -1.8    3.154794               -7.038697    3.438697
------------------------------------------------------------------------------
    diff = mean(NEW) - mean(STANDARD)                             t =  -0.5706
Ho: diff = 0                                     degrees of freedom =       98
    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2848         Pr(|T| > |t|) = 0.5696          Pr(T > t) = 0.7152
```

Solution (ii)

The printer out gives a 90% 2-sided confidence interval. The question states that a 5mm difference on the visual analogue scale was considered to be the minimum clinically important difference. Therefore + 5 mm can be used

as used as then limits of non-inferiority . Where higher values of the outcome measure mean a worse outcome for patients, the null hypothesis of inferiority is rejected if the upper (1-α) confidence interval is below limits of non-inferiority with a significance level less than or equal to From the printout the 90% confidence interval is (-7.038697 to 3.438697). The upper 95% confidence limit is therefore 3.438697. Since this is less than 5, it is possible to reject the null hypothesis that the NEW treatment inferior in a 5% level test.

[6 marks]
[Total 9 marks]

A5.

Inferential Statistical tests comparing treatment groups at baseline are quite often seen in published reports of randomised controlled trials. Why are such tests not recommended when reporting the results of randomised controlled trials?

Solution

Statistical test of baseline variable test the null hypothesis that the two treatment groups are the same at baseline. Randomisation means that this hypothesis will be true. Hence any statistically significant differences between treatment groups at baseline will be type I errors. If this type of analysis is carried out and finds a significant difference between randomised groups at baseline, one is left with the choice between disregarding the analysis or concluding that randomisation has not been carried out properly.

The answer might also refer to multiplicity.

[4 marks]

SECTION B

Answer **TWO** of the three questions

B6.

Consider a parallel group trial with treatment (*T*) and control (*C*). For an outcome measure *Y* let $\overline{y}_T$, $\overline{y}_C$, μ_T, and μ_C be 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 $H_0 : \mu_T - \mu_C = 0$ will be tested by the statistic $T = \dfrac{\overline{y}_T - \overline{y}_C}{s\lambda}$, with $\lambda = \sqrt{1/n_T + 1/n_C}$ where n_T and n_C are the number of subjects allocated to the respective treatments.

(i) Assuming a normal approximation to the *t*-distribution, what is the distribution of T when $\mu_T - \mu_C = \tau$

Solution (i)

Assuming a normal approximation to the t-distribution, the distribution of T when $\mu_T - \mu_C = \tau$ is

$$N\left[\frac{\tau}{\sigma\lambda}, 1\right]$$

[2 marks]

(ii) Show that $\Pr[\text{Reject } H_0 |\ \tau\ ] \cong \left(1 - \Phi\left(z_{\alpha/2} - \dfrac{\tau}{\sigma\lambda}\right)\right)$.

Solution (ii)

The null hypothesis will be rejected if $|T| > z_{\alpha/2}$. Hence

$$\Pr[\text{Reject } H_0 |\ \tau] = Pr[|T| > z_{\alpha/2} |\ \tau] = \Pr[T < -z_{\alpha/2} |\ \tau] + \Pr[T > z_{\alpha/2} |\ \tau]$$

$$= \Phi\left(-z_{\alpha/2} - \frac{\tau}{\sigma\lambda}\right) + \left(1 - \Phi\left(z_{\alpha/2} - \frac{\tau}{\sigma\lambda}\right)\right)$$

Without loss of generality we can assume that $\tau > 0$. Hence $\Phi\left(-z_{\alpha/2} - \dfrac{\tau}{\sigma\lambda}\right)$ is negligible.

Hence $\Pr[\text{Reject } H_0 |\ \tau\ ] \cong \left(1 - \Phi\left(z_{\alpha/2} - \dfrac{\tau}{\sigma\lambda}\right)\right)$.

Bookwork

[3marks]

(iii) Assuming equal size groups $(n_T = n_C)$, show that the total sample size required to give a power $(1-\beta)$ for a two-tailed α size test is

$$N = 4\frac{\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_{\beta}\right)^2.$$

Solution (iii)

$= \Pr[\text{Reject } H_0 |\ \tau] = \left(1 - \Phi\left(z_{\alpha/2} - \frac{\tau}{\sigma\lambda}\right)\right) = 1 - \beta$

There $\beta = \Phi\left(z_{\alpha/2} - \frac{\tau}{\sigma\lambda}\right)$

Since $\Phi^{-1}(\beta) = -z_{\beta}$ it follows that $-z_{\beta} = z_{\alpha/2} - \frac{\tau}{\sigma\lambda}$

giving $\frac{\tau}{\sigma\lambda} = z_{\alpha/2} + z_{\beta}$.

If $n_T = n_C = n$ then $\lambda = \sqrt{2/n}$

Therefore $\frac{\tau}{\sigma}\sqrt{\frac{n}{2}} = z_{\alpha/2} + z_{\beta}$. Rearrangement gives

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

Hence the total sample size $N(k) = 4\frac{\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_{\beta}\right)^2$

Bookwork

[6 marks]

(iv) Suppose the total sample size to give a power $(1-\beta)$ using a test size α assuming equal group size is equal to N. Following treatment allocation there is an imbalance in group sizes with $n_T = kn_C$ and $n_T + n_C = N$. Show that the power of the trial equals

$$1 - \Phi\left(z_{\alpha/2} - \left(\frac{2\sqrt{k}}{k+1}\right)\left(z_{\alpha/2} + z_{\beta}\right)\right).$$

Solution (iv)

From (ii) $Power = 1 - \Phi\left(z_{\alpha/2} - \frac{\tau}{\sigma\lambda}\right)$ (8)

Due to imbalance $\lambda = \sqrt{1/n_T + 1/n_C} = \sqrt{1/kn_C + 1/n_C} = \sqrt{\frac{1+k}{kn_C}}$.

From (iv) $N = 4\frac{\sigma^2}{\tau^2}\left(z_{\alpha/2} + z_\beta\right)^2$.

Since $N = (k+1)n_c$, it rearrangement gives $\frac{\tau}{\sigma} = \frac{2}{\sqrt{(1+k)n_c}}\left(z_{\alpha/2} + z_\beta\right)$.

Substitution into (*) gives

$$Power = 1 - \Phi\left(z_{\alpha/2} - \frac{2}{\sqrt{(1+k)n_c}}\left(z_{\alpha/2} + z_\beta\right)\frac{1}{\sqrt{\frac{1+k}{kn_C}}} \right) = 1 - \Phi\left(z_{\alpha/2} - \left(\frac{2\sqrt{k}}{k+1}\right)\left(z_{\alpha/2} + z_\beta\right) \right)$$ as required.

[6 marks]

(v) Suppose samples size for a trial has been estimated to give 80% power assuming a two-tailed test size of 5% and equal size groups $(n_T = n_C)$. Determine the power the trial will have if the allocation ratio k = 1.2, and comment briefly on the result.

$$Power = 1 - \Phi\left(z_{\alpha/2} - \left(\frac{2\sqrt{k}}{k+1}\right)\left(z_{\alpha/2} + z_\beta\right) \right) = 1 - \Phi\left(1.96 - \left(\frac{2\sqrt{1.2}}{2.2}\right)(1.96 + 0.84) \right)$$

$$= 1 - \Phi\left(1.96 - \left(\frac{2\sqrt{1.2}}{2.2}\right)(1.96 + 0.84) \right) = 1 - \Phi(-0.83) = \Phi(0.83) = 0.797$$

The power has reduced only very slightly compared to equal allocation. Small imbalances in allocation will have only a very small affect on power.

[3 marks]

[Total 20 marks]

B7.

Consider a randomised controlled trial. Suppose the patient population can be divided into three sub-groups as follows:

Compliers: patients who will comply with the allocated treatment,

*Always control treatmen*t: patients who will receive control treatment regardless of allocation,

Always new treatment: patients who will receive the new treatment regardless of allocation.

This assumes that there are no defiers, that is patient who will always receive the opposite of the treatment to which they are randomised. Assume that the proportion and characteristics of *compliers*, *always control treatment*, *always new treatment* is the same in both arms and that randomization can only affect the outcome through the receipt of treatment.

(i) Show that an *Intention-To-Treat* estimate of the treatment effect is biased towards the null hypothesis of no treatment effect.

Solution(i)

Table of expected means under assumptions of model

	Type	**Control Group**	**New Treatment Group**	**Proportion In Latent Class**
As Randomized	**A**	μ	$\mu + \tau$	$\theta_A = 1-\theta_B - \theta_C$
Always Control	**B**	$\mu+\gamma_B$	$\mu+\gamma_B$	θ_B
Always New Treatment	**C**	$\mu+\gamma_C+\tau$	$\mu+\gamma_C+\tau$	θ_C

τ is the causal effect of treatment

For Intention-to-Treat Estimate

$$\tau_{ITT} = \left[\theta_A(\mu+\tau)+\theta_B(\mu+\gamma_B)+\theta_C(\mu+\gamma_C+\tau)\right]-\left[\theta_A\mu+\theta_B(\mu+\gamma_B)+\theta_C(\mu+\gamma_C+\tau)\right]$$
$$= \theta_A\tau$$

as second and third terms in each bracket cancel.

Hence $|\hat{\tau}_{ITT}| \le \tau$ which means $\hat{\tau}_{ITT}$ is biased towards zero if $\theta_A < 1$ i.e. if some patients do not comply with treatment.

[4 Marks] [Bookwork]

(ii) Show that a *As-Treated* estimate of the treatment effect may be biased either towards or away from the null hypothesis of no treatment effect.

Solution (ii)

For the As-Treated Estimate

$$\tau_{AT} = \left[\frac{\theta_A(\mu+\tau)+2\theta_C(\mu+\gamma_C+\tau)}{\theta_A+2\theta_C}\right] - \left[\frac{\theta_A\mu+2\theta_B(\mu+\gamma_B)}{\theta_A+2\theta_B}\right]$$

$$= \left[\frac{(\theta_A+2\theta_C)\mu+2\theta_C\gamma_C+(\theta_A+2\theta_C)\tau}{\theta_A+2\theta_C}\right] - \left[\frac{(\theta_A+2\theta_B)\mu+2\theta_B\gamma_B}{\theta_A+2\theta_B}\right]$$

$$= \tau+\mu+\left[\frac{2\theta_C\gamma_C}{\theta_A+2\theta_C}\right]-\mu-\left[\frac{2\theta_B\gamma_B}{\theta_A+2\theta_B}\right]$$

$$= \tau+\left[\frac{2\theta_C\gamma_C}{1-\theta_B-\theta_C+2\theta_C}\right]-\left[\frac{2\theta_B\gamma_B}{1-\theta_B-\theta_C+2\theta_B}\right]$$

$$= \tau+\left[\frac{\theta_C\gamma_C}{1-\theta_B+\theta_C}\right]-\left[\frac{\theta_B\gamma_B}{1-\theta_C+\theta_B}\right]$$

The values of γ_C and γ_B can be positive or negative so that the second and third terms can be either positive or negative. Hence the bias can be away from or towards the null hypothesis.

[5 Marks]

The table below summarizes the outcome of patients from randomised controlled trial comparing two treatments according to randomized group and treatment received. Some patients allocated to the *New* treatment received the *Control* treatment and some patients allocated to *Control* treatment received the *New* treatment.

	Randomised Group			
Recovered after 6 weeks	*New Treatment*		*Control*	
	Received New	*Received Control*	*Received New*	*Received Control*
Yes	360	72	48	360
No	120	48	12	180
Total	480	120	60	540

(iii) For the outcome measure Recovered after 6 weeks, calculate the point estimate of the treatment effect of the *New* treatment compared to the *Control* treatment assuming

(a) an *Intention-To-Treat* analysis,

(b) an*As-Treated analysis*. **Solution(iii)**

(a) Intention-To-Treat = 432/600 – 408/600= 0.72 - 0.68 = 0.04

(b) *Per-Protocol* = 408/540-432/660=0.75555 - 0.654545 = 0.1010101

[Calculation time 2 mins]

[2 marks]

(iv) Briefly explain why an *Intention-To-Treat* estimate is preferable to the *As-Treated* estimate in a superiority trial.

Solution (iv)

As we have seen in (i) an intention-to-treat analyses will bias an estimate of the treatment effect towards the estimate of no effect. This means that any effect will be bias towards the null hypothesis of a superiority trial. Hence if we reject the null hypothesis using an intention-to-treat analysis, we can be more confident that the true treatment effect is at least as large as that observed. In contrast a per-protocol analyses may bias the estimate of the treatment effect either away or towards the null hypothesis of no effect as seen in (ii). [3 marks]

(v) What are the implications of this for the conduct of randomised controlled trials?

Solution (v)

The implications of intention-to-treat analysis are that researchers running trials should endeavour to get outcome data on all patients, irrespective of whether they receive the treatment to which they are randomized. [2 marks]

(vi) Calculate the point estimates of the C*ompliance Average Causal Effect* of *New* treatment compared to *Control*.

Solution (vi)

From above $\tau_{ITT} = \theta_A \tau$ where θ_A is the proportion of subjects that accept randomized treatment and τ is the causal effect of treatment or the C*ompliance Average Causal Effect* . $\theta_A = 1-\theta_B - \theta_C$ where θ_B is the proportion of patients that will always receive the control and θ_C the proportion who will always receive the new active treatment.

θ_B can be estimated from the new treatment arm and θ_C always the control arm.

θ_B=120/600

θ_C =60/600

Hence θ_A =1-60/600 -120/600=0.7

Hence the C*ompliance Average Causal Effect* of *New* treatment compared to *Control* treatment, τ=0.04/0.7=0.057. [4 marks]

[Total mark 20]

B8.

(i) Suppose that in a clinical trial n_T patients are randomized to treatment (*T*) and n_C to the control (*C*). The outcome measure is binary. Suppose that the number of events in each of the two treatment groups are r_T and r_C with probability parameters π_T and π_C, respectively.

Consider the odds ratio for the treatment effect of *T* compared to *C* defined by $OR = \dfrac{\pi_T(1-\pi_C)}{(1-\pi_T)\pi_C}$.

Using the approximate relationship $Var\left[f(X) \right] \cong f'(X)^2_{X=E[X]} Var\left[X \right]$, show that

$$Var\left[\log_e \left(\widehat{OR} \right) \right] \approx \frac{1}{n_T \pi_T} + \frac{1}{n_T(1-\pi_T)} + \frac{1}{n_C \pi_C} + \frac{1}{n_C(1-\pi_C)}. \qquad \text{[7 marks]}$$

Solution (i)

$$Var\left[\log_e \left(\widehat{OR} \right) \right] = Var\left(\ln \left[\frac{p_T(1-p_C)}{(1-p_T)p_C} \right] \right) = Var\left[\ln \frac{p_T}{1-p_T} - \ln \frac{p_C}{1-p_C} \right]$$

$$= Var\left(\ln \frac{p_T}{1-p_T} \right) + Var\left(\ln \frac{p_C}{1-p_C} \right)$$

as treatment groups are independent.

$$\left. \frac{d}{d\pi} \left(\ln \frac{p}{1-p} \right) \right|_{p=\pi} = \frac{(1-\pi)}{\pi} . \frac{1}{(1-\pi)^2} \text{ and } Var(\pi) = \frac{\pi(1-\pi)}{n}.$$

Using delta method approximation,

$$Var\left(\ln \frac{\pi}{(1-\pi)} \right) \approx \left[\frac{d}{d\pi} \left(\ln \frac{\pi}{1-\pi} \right) \right]^2 . Var(\pi) = \frac{1}{n\pi(1-\pi)} = \frac{1}{n\pi} + \frac{1}{n(1-\pi)}.$$

Hence

$$Var\left[\log_e \widehat{OR} \right] \approx \frac{1}{n_T \pi_T} + \frac{1}{n_T(1-\pi_T)} + \frac{1}{n_C \pi_C} + \frac{1}{n_C(1-\pi_C)}$$

[Bookwork]

[7 marks]

(ii) Hence show that the $(1-\alpha)$ confidence interval for the odds ratio is given by the values of

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

Solution(ii)

$n_T\pi_T, n_C\pi_C, n_T(1-\pi_T)$ and $n_C(1-\pi_C)$ are the expected cell frequencies. To estimate $SE\left[\log_e \widehat{OR}\right]$ replace with the observed cell frequencies r_T, r_C, $(n_T - r_T)$ and $(n_C - r_C)$ giving

$$\widehat{SE}\left[\log_e \widehat{OR}\right] \approx \sqrt{\frac{1}{r_T} + \frac{1}{n_T - r_T} + \frac{1}{r_C} + \frac{1}{n_C - r_C}}$$

The $(1-\alpha)$ confidence interval for the log odds ratio is therefore

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

Taking antilogs the $(1-\alpha)$ confidence interval for the odds ratio is therefore

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

[Bookwork] [4 marks]

A systematic review of trials of vaccines to prevent influenza has identified two randomized trials that compare the *New* vaccine with a *Standard* vaccine. The table below summarizes the data from the two trials giving the number of subjects who had and not had influenza in the 12 months following vaccination by intervention group for each of the two trials.

	New Vaccine			*Standard Vaccine*		
	Influenza		N	Influenza		N
Trial	Yes	No		Yes	No	
A	50	500	550	100	450	550
B	30	200	230	50	180	230

(iii) For each trial determine $\log_e\left[\widehat{OR}\right]$ and $\hat{SE}\left[\log_e\left[\widehat{OR}\right]\right]$.

Solution (iii)

For trial A $\log_e\left[\widehat{OR}\right] = \log_e\left[\frac{50\times450}{500\times100}\right] = \log_e[0.45] = -0.7985$ and

$$\widehat{Var}\left[\log_e\left[\widehat{OR}\right]\right] = \frac{1}{50} + \frac{1}{500} + \frac{1}{100} + \frac{1}{450} = 0.03422$$

For trial A $\log_e\left[\widehat{OR}\right] = \log_e\left[\frac{30\times180}{200\times50}\right] = \log_e[0.54] = -0.61619$ and

$$\widehat{Var}\left[\log_e\left[\widehat{OR}\right]\right] = \sqrt{\frac{1}{30} + \frac{1}{200} + \frac{1}{50} + \frac{1}{180}} = 0.06389$$ [4 marks]

(iv) Suppose θ_i is the treatment effect for the i^{th} trial. The inverse-variance pooled estimate is given by $\hat{\theta} = \dfrac{\sum_i w_i \hat{\theta}_i}{\sum_i w_i}$ where $w_i = 1/Var\left[\hat{\theta}_i\right]$ with $Var[\hat{\theta}] = \dfrac{1}{\sum_i w_i}$. By setting $\hat{\theta}_i = \log_e\left[\widehat{OR}_i\right]$, compute the pooled estimate of the odds ratio for *New* vaccine as compared to *Control*.

Solution (iv)

For A $w_A = 1/Var\left[\widehat{OR}_A\right]$=29.223, $\log_e\left[\widehat{OR}_A\right] = -0.7985$

For B $w_B = 1/Var\left[\widehat{OR}_B\right]$=15.6522, $\log_e\left[\widehat{OR}_B\right] = -0.61619$

Hence

$$\log_e\left[\widehat{OR}_{pool}\right](\ ) = \frac{w_A \log_e\left[\widehat{OR}_A\right] + w_B \log_e\left[\widehat{OR}_B\right]}{w_A + w_B}$$

$$= \frac{29.223(-0.7985) - 15.6522(0.61619)}{44.8752}$$

$$= -0.73488$$

Hence pooled Odds Ratio = $\exp(-0.73488) = 0.4795$.

[5 marks]

[Total 20 marks]

END OF EXAMINATION PAPER