MT3772 Examination Paper Solutions (May 2004)

A1. *Explain what is meant by the term stratified randomisation in the context of a randomised controlled clinical trial.*

Stratified randomisation is an adaptive method of randomly allocating treatment to patients designed to prevent treatment groups differing in important characteristics defined by each strata. For each strata a separate randomisation list is used, usually with each constructed using block randomisation, so that similar numbers of patients are allocated to each treatment in each strata thus guaranteeing that the treatment groups have a similar composition.
[Total 3 marks]

A2. *A clinical trial compared two treatments for athlete's foot, a new ointment and a placebo ointment with no active ingredient. Using randomisation 39 patients are allocated to the new ointment and 40 to the standard ointment. Patients are assessed at the end of the two week treatment period. The infection was eradicated for 29 patients in the new treatment group and 23 patients in the placebo group. A difference in the success rate of the two treatments of 10% was considered to be clinically important.*

(i) State the hypotheses you might use to compare the treatments.

The data can be tabulated as follows

	New	*Placebo*	*Total*
Success	29(74%)	23 (58%)	52
Failure	10	17	27
Total	39	40	79

If π_N π_P the population proportions of success in the new and placebo treatment groups the hypotheses that might be used to compare the two treatments is
H_0: $\pi_N = \pi_P$ vs H_0: $\pi_N \neq \pi_P$ *[2 marks]*

(ii) Carry out a statistical test to compare the treatments specifying the assumptions that are made.

In order to test this hypothesis a two sample z-test of proportions can be used, which is define as follows.

Suppose r_n and r_p are the number of success in each group**,** p_N p_P the population proportion of success in the new and placebo treatment groups, n_n n_p the in each group. The test statistic is defined as $Z = \dfrac{|p_N - p_P|}{s.e._{null}(p_N - p_P)}$ with $s.e._{null}(p_N - p_P) = \sqrt{p\ (1-p)\left(\dfrac{1}{n_N} + \dfrac{1}{n_P}\right)}$ and $p = \dfrac{r_N + r_P}{n_N + n_P}$

Z can be assumed to be Normally distributed provided $n_N p, n_N(1-p), n_P p, n_P(1-p)$ are greater than 5. The smallest of these will be $n_N(1-p)=13.3$. Hence the normal approximation can be used.

$$p_N - p_P = \frac{r_N}{n_N} - \frac{r_P}{n_P} = \frac{29}{39} - \frac{23}{40} = 0.1685, \quad p = \frac{52}{79}$$

$$s.e._{null}(p_N - p_P) = \sqrt{\frac{52}{79} \times \frac{27}{79}\left(\frac{1}{39} + \frac{1}{40}\right)} = \sqrt{0.01139} = 0.1067$$

Hence $$Z = \frac{|p_A - p_B|}{s.e._{null}(p_A - p_B)} = \frac{0.1685}{0.1067} = 1.579$$

For a 5% level two-sided test the critical value of z_α is 1.96.
Alternatively z_p=1.579. From tables p=0.1142 greater than α =0.05.
Therefore the result is not statistically significant at a 5% level. *[5 marks]*

(iii) Comment on the results of the trial.

The difference in the success rate of the new treatment compared to placebo (treatment effect) observed in this trial is 16.8%, but this is not statistically in a 5% level two-sided test. Since a 10% improvement in the success rate was considered to be clinically important, it would suggest that the study was under-powered to detect an clinically important effect.
[3 marks]

[Total 10 Marks]

A3.

(i) *Explain what is meant by an equivalence trial.*

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. *Equivalence trial* are designed to establish that the efficacy of two or more treatments is the same. Therefore in such a trial the null hypothesis is that the average outcome is different.
[3 marks]

(iii) Outline the statistical analysis one could use to establish whether a new treatment T is equivalent to a control treatment C on a continuous outcome measure Y in a parallel group trial.

Rather than using a formal significance test, statistical analysis of equivalence trials is usually based on the confidence interval of difference between treatments. Equivalence is established by demonstrating that the confidence interval of the difference lies in the specified range ($-\delta_E$, $+\delta_E$).

Suppose outcome measure Y is continuous and Normally distributed with means μ_C and μ_T for the control and new treatment respectively. Rejection of the null hypothesis that H_0: $|\mu_T-\mu_C|>\delta_E$ against the alternative hypothesis H_1: $|\mu_T-\mu_C|\leq\delta_E$ where the (1-2α) confidence interval is within the interval [$-\delta_E$, $+\delta_E$], will have a type I error of less than α *[4 marks]*

(iv) *An equivalence trial is carried out to compare a new medication with the current standard medication. To demonstrate equivalent pain relief it was felt that mean pain scores for each group should not differ by more than 5 units. The table below summarizes the pain relief scores for each treatment group with higher pain scores representing greater pain. Test whether the two treatments may be considered to be equivalent.*

	Treatment Group					
	New Medication			*Standard Medication*		
	Mean	*S.D.*	*N*	*Mean*	*S.D.*	*N*
Pain Relief	*35.2*	*18.1*	*171*	*34.1*	*17.9*	*172*

The (1-2α) confidence interval given by $\bar{y}_T - \bar{y}_C - t_\alpha(\nu)se(\bar{y}_T - \bar{y}_C)$ to $\bar{y}_T - \bar{y}_C + t_\alpha(\nu)se(\bar{y}_T - \bar{y}_C)$, with $s.e.(\bar{y}_T - \bar{y}_C) = s\sqrt{1/n_T + 1/n_C}$ and $\nu = n_T + n_C - 2$ and s is the pooled standard deviation.

For a 5% level test one needs to consider a 90% confidence interval.

$$s = \sqrt{\frac{(n_T - 1).s_T^2 + (n_C - 1).s_C^2}{n_T + n_C - 2}} = \sqrt{\frac{170 \times 18.1^2 + 171 \times 17.9}{341}} = 18.0$$

From table $t_{0.05}(341) \cong 1.65$

The 90% confidence interval of the mean is

$\bar{y}_T - \bar{y}_C \pm t_\alpha(\nu)se(\bar{y}_T - \bar{y}_C) = 35.2 - 34.1 \pm 1.65 \times 18.0 \times \sqrt{\frac{1}{171} + \frac{1}{172}}$ giving the interval(-2.11 to 4.31)

Since this is within the interval (-5 to +5) the null hypothesis that the two treatments are not equivalent can be rejected with test size less than 5%.

[4 marks]

[Total 11marks]

A4.

A randomised controlled trial of treatments for rheumatoid arthritis compared a new treatments (T) with a standard treatment (C). Outcome is measured by a continuous health activity scale. The table below summaries the outcome in terms of scale broken down by treatment group and age group (<65 years, ≥65 years).

Age Group	*<65 years*		*≥65 years*	
Treatment	*New Treatment (T)*	*Control (C)*	*New Treatment (T)*	*Control (C)*
Mean	*1.29*	*1.47*	*1.51*	*1.49*
S.D.	*0.72*	*0.74*	*0.69*	*0.79*
N	*124*	*142*	*77*	*63*

(i) State the hypothesis you might use to compare the treatment effect in patients less than 65 years of age with the treatment effect in patients of 65 years of age or more and write down a test statistic that could be used to test this.

If δ_Y is the treatment effect in the younger patients and δ_O is the treatment effect in the older patients, the hypothesis to compare the treatment effects is $H_0:\delta_O=\delta_Y$ vs $H_1:\delta_O\neq\delta_Y$

The hypothesis $H_0:\delta_A=\delta_B$ vs $H_1:\delta_A\neq\delta_B$ can be tested using the test statistic

$$T_{YO}=\frac{\left(\bar{Y}_{TY}-\bar{Y}_{CY}\right)-\left(\bar{Y}_{TO}-\bar{Y}_{CO}\right)}{SE\left(\left(\bar{Y}_{TY}-\bar{Y}_{CY}\right)-\left(\bar{Y}_{TO}-\bar{Y}_{CO}\right)\right)} \text{ where}$$

$$\hat{SE}\left(\left(\bar{Y}_{TY}-\bar{Y}_{CY}\right)-\left(\bar{Y}_{TO}-\bar{Y}_{CO}\right)\right)=\sqrt{\left(\frac{S_{TY}^{\ 2}}{n_{TY}}+\frac{S_{CY}^{\ 2}}{n_{CY}}\right)+\left(\frac{S_{TO}^{\ 2}}{n_{TO}}+\frac{S_{CO}^{\ 2}}{n_{CO}}\right)}$$

where $\bar{Y}_{CY}$, $\bar{Y}_{TY}$, $\bar{Y}_{CO}$ and $\bar{Y}_{TO}$ are the subgroup sample means, with S_{CY}, S_{TY}, S_{CO}, and S_{TO} subgroup sample standard deviations and n_{CY}, n_{TY}, n_{CO} and n_{TO} the subgroup sample sizes.
[5 marks]

(ii) Apply this test statistic to assess if there is any evidence of difference in the treatment effect according to age group, stating any assumptions you make.

Substitution into the above formula gives

$$\hat{SE}\left(\left(\bar{Y}_{TY}-\bar{Y}_{CY}\right)-\left(\bar{Y}_{TO}-\bar{Y}_{CO}\right)\right)=\sqrt{\left(\frac{0.72^2}{124}+\frac{0.74^2}{142}\right)+\left(\frac{0.69^2}{77}+\frac{0.79^2}{63}\right)}=0.155$$

Hence

$$T_{YO}=\frac{\left(1.29-1.47\right)-\left(1.51-1.49\right)}{0.155}=\frac{0.2}{0.155}=1.29$$

Assuming T_{YO} will have a standardized normal distribution asymptotically, reference to table gives p=0.20 so that the null hypothesis cannot be rejected at a 5% level . *[4marks]*

[Total 9 marks]

A5

(i) Using asthma as an example, explain the difference between cumulative incidence and prevalence rates.

The prevalence of asthma in a population at a point in time is the number of people with asthma at that time divided by the size of the population.

The cumulative incidence of asthma refers to new cases which develop in a period of time - (0,t) say – in a population who are all free of asthma at time 0. It is the number of new cases divided by the size of the this population at time 0. *[3 marks]*

(ii) An epidemiological study compared asthma prevalence rates in two geographical areas, one with high and one with low atmospheric pollution. Describe two necessary conditions for a factor to be a confounder in this study.

A confounder must be: (a) a causal independent risk factor for asthma and (b) it must be associated with pollution in the study data. *[2 marks]*

(iii) In the following logistic regression analyses using STATA software, the four variables (asthma, pollution, sclass2 and sclass3) were all coded as 0 or 1 with code 1 indicates presence of asthma, living in a high pollution area, social class 2 and social class 3 respectively. (There were three social classes altogether). Does the data indicate that social class is a confounder? Does the data suggest an effect of pollution on asthma prevalence? What is the estimated odds ratio? Justify your answers.

```
Logit estimates                                 Number of obs   =       6117
                                                LR chi2(1)      =       3.84
                                                Prob > chi2     =     0.0501
Log likelihood = -2878.4682                     Pseudo R2       =     0.0007
------------------------------------------------------------------------------
      asthma |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   pollution |       .131    .0667     1.96   0.050       -.0001       0.2614
       _cons |    -1.586    .0485   -32.73   0.000      -1.6811      -1.4912
------------------------------------------------------------------------------


Logit estimates                                 Number of obs   =       6117
                                                LR chi2(3)      =      19.36
                                                Prob > chi2     =     0.0002
Log likelihood = -2870.7086                     Pseudo R2       =     0.0034

------------------------------------------------------------------------------
      asthma |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   pollution |       .133    .0667     1.99   0.047       .0019       .2638
     sclass1 |       .127    .0827     1.54   0.125      -.0350       .2892
     sclass2 |       .315    .0796     3.97   0.000       .1595       .4714
       _cons |    -1.702    .0601   -28.31   0.000     -1.8199     -1.5843
------------------------------------------------------------------------------
```

The second analysis appears to show that social class was a predictor of asthma. From (ii) we also need it to be associated with pollution in order that it be considered a confounder. Data on this has not been shown directly; however since the estimates of the pollution effect from models with and without (first analysis) are almost identical, this suggest no relationship between social class and pollution in this data. Social class is not a confounder. *[3 marks]*

The second analysis includes pollution and social class as covariates. Therefore the estimate of pollution effect is adjusted for any confounding effect of social class – if there is any. In this analysis the size of the effect of pollution can be estimated by β=.133 which implies an odds ratio of exp(0.133) = 1.140. Alternatively one might decide to discount social class as a confounder and use the first analyses giving an odds ratio of exp(0.131) = 1.14.

The test of H_0: β=1 gives p=0.047; if using a 5% significance level, this implies that H_0 can be rejected. However one should consider the possibility of confounding by other factors before inferring that the increase is really due to pollution. *[4 marks]*
[Total 12 marks]

A6

Consider the problem of estimating $\theta_i = \lambda_{1i}/\lambda_{0i}$, i =1,..,k where λ_{1i} is the death rate in age group i of a group of miners and λ_{0i} is the corresponding rate in the general population which is already known. Data on observed deaths, a_i and corresponding person-years at risk, T_i, i =1,..,k, are available from the miners.

Assuming that the a_i are Poisson distributed and $\theta_i = \theta$ for all i, show that the maximum likelihood estimator of θ is $\dfrac{\sum a_i}{\sum \lambda_{i0} T_i}$.

Under the assumptions, $\lambda_{1i}=\theta\lambda_{0i}$, i=1,…k and the number of deaths is Poisson with parameter $\theta\lambda_{0i}T_i$. The likelihood of the data is

$$L = \prod_{i=1}^{k} \frac{e^{-\theta\lambda_{i0}T_i}(\theta\lambda_{i0}T_i)^{a_i}}{a_i!}$$

$$\Rightarrow Ln(L) = -\theta\sum_{i=1}^{k}\lambda_{i0}T_i + \sum_{i=1}^{k} a_i \ln(\theta\lambda_{i0}T_i) + const$$

$$\Rightarrow \frac{dLn(L)}{d\theta} = -\sum\lambda_{i0}T_i + \sum\frac{a_i}{\theta} \Rightarrow \hat{\theta}_{MLE} = \frac{\sum a_i}{\sum\lambda_{i0}T_i}$$

[5 marks]

B7 For an AB/BA crossover trial a model for a continuous outcome y_{ij} of the iP^{th} patient in the j^{th} period can be written as

$y_{i1} = \mu + \delta + \xi_i + \varepsilon_{i1}$ 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

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

where μ is the mean for the sequence BA in period 1 , δ 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 and give an example of how one might occur in a crossover trial.

The *carryover effect* is cause by the effect of treatment in the first period carrying over to the second period. If there is a difference in the carryover for the two sequences then this is the carryover effect. [3 marks]

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

For sequence AB $d_i = y_{i2} - y_{i1} = \phi - \delta + \gamma + \varepsilon_{i2} - \varepsilon_{i1}$

Therefore $E\left[\bar{d}_{AB}\right] = \phi + \gamma - \delta$

For sequence BA $d_i = y_{i2} - y_{i1} = \phi + \delta + \varepsilon_{i2} - \varepsilon_{i1}$

Therefore $E\left[\bar{d}_{BA}\right] = \phi + \delta$

Hence $E\left[\hat{\delta}\right] = E\left[\frac{\bar{d}_{BA} - \bar{d}_{AB}}{2}\right] = \delta - \frac{\gamma}{2}$

So the treatment effect is biased by *$\gamma/2$.* [3 marks]

(iii) *Show that* $Var\left[\hat{\delta}\right]=\frac{\sigma_\varepsilon^2}{2}\left(\frac{1}{n_{BA}}+\frac{1}{n_{AB}}\right)$ *where* n_{AB} *and* n_{BA} *are the number of patients in each sequence.*

$$Var\left[\bar{d}_{BA}\right]=Var\left[\frac{\sum d_i}{n_{BA}}\right]=\frac{Var\left[d_i\right]}{n_{BA}}=\frac{Var\left[\phi-\delta+\gamma+\varepsilon_{i2}-\varepsilon_{i1}\right]}{n_{BA}}=\frac{2\sigma_\varepsilon^2}{n_{BA}}$$

and $Var\left[\bar{d}_{AB}\right]=\frac{2\sigma_\varepsilon^2}{n_{AB}}$.

Since sequences are independent

$$Var\left[\hat{\delta}\right]=Var\left[\frac{\bar{d}_{BA}-\bar{d}_{AB}}{2}\right]=\frac{1}{4}\left(Var\left[\bar{d}_{BA}\right]+Var\left[\bar{d}_{AB}\right]\right)=\frac{\sigma_\varepsilon^2}{2}\left(\frac{1}{n_{BA}}+\frac{1}{n_{AB}}\right)$$

as required. [4 marks]

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

For sequence AB

$$a_i=y_{i2}+y_{i1}=2\mu+\phi+\delta+\nu+2\xi_i+\varepsilon_{i2}+\varepsilon_{i1}$$

Therefore

$$E\left[\bar{a}_{AB}\right]=E\left[2\mu+\phi+\delta+\gamma+2\xi_i+\varepsilon_{i2}+\varepsilon_{i1}\right]=2\mu+\phi+\delta+\gamma$$

For sequence BA

$$a_i=y_{i2}+y_{i1}=2\mu+\phi+\delta+2\xi_i+\varepsilon_{i2}+\varepsilon_{i1}$$

$$E\left[\bar{a}_{BA}\right]=E\left[2\mu+\phi+\delta+2\xi_i+\varepsilon_{i2}+\varepsilon_{i1}\right]=2\mu+\phi+\delta$$

Subtraction gives $E\left[\bar{a}_{AB}-\bar{a}_{BA}\right]=\gamma$ as required . [4 marks]

(v) *Derive an expression for* $Var\left[\bar{a}_{AB}-\bar{a}_{BA}\right]$.

$$Var\left[\bar{a}_{AB}\right]=Var\left[\frac{\sum a_i}{n_{AB}}\right]=\frac{Var\left[2\mu+\phi+\delta+\gamma+2\xi_i+\varepsilon_{i2}+\varepsilon_{i1}\right]}{n_{AB}}=\frac{4\sigma_B^2+2\sigma_\varepsilon^2}{n_{AB}}$$ *and similarly*

$$Var\left[\bar{a}_{BA}\right]=\frac{4\sigma_B^2+2\sigma_\varepsilon^2}{n_{BA}}$$

Therefore $Var\left[\bar{a}_{AB}-\bar{a}_{BA}\right]=\left(4\sigma_B^2+2\sigma_\varepsilon^2\right)\left(\frac{1}{n_{AB}}+\frac{1}{n_{BA}}\right)$ [3 marks]

(vi) *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 in a crossover trial? What are the implications of this for the design of crossover trials?*

The advantage of a crossover trial is that the between subject variance σ_B^2, which is generally larger than the within subject variance σ_ε^2, is removed from the test of the treatment effect. The weakness of the T_a test of carryover effect is that $\hat{SE}\left[\bar{a}_{AB} - \bar{a}_{BA}\right]$ includes the between subject variance σ_B^2. The statistical test T_a will therefore have low power in the circumstance when such a test might be used. [3 marks]

The implication of this for the design of crossover trials is that they are usually only advisable in circumstance where the possibility of a carryover effect can be discounted for scientific reasons or by virtue of the design. [2 marks]

(vii) In a randomised controlled crossover to compare two medications, explain how one might prevent a carryover effect?

One way of preventing a carryover effect is to have a “washout period” between the two treatments to allow any residual effect of the first treatment to be eliminated before starting the second treatment. [3 marks]

[Total 25 marks]

B8 *Give two reasons why it is important to estimate sample size in a randomised controlled trial.*

Sample calculation is important for two reasons:

- If too few patients are recruited, the trial may lack statistical power. It is likely that the study will fail to answer the question it is attempting to address.
- If more patients than the minimum required to answer the question are recruited, some patients may be unnecessarily exposed to an inferior treatment.

[4 marks]

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 δ is the treatment effect such that

$$y = \mu + \varepsilon_y, \qquad x = \mu_x + \varepsilon_x \qquad \text{Group C}$$

$$y = \mu + \delta + \varepsilon_y, \qquad x = \mu_x + \varepsilon_x \qquad \text{Group T.}$$

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* $Cov\left[\varepsilon_x, \varepsilon_y\right] = \sigma_{xy}$

(i) Let $d = y - x$ *with* $\bar{d}_C$ *and* $\bar{d}_T$ *the sample means of Group C and Group T respectively. Show that* $E\left[\bar{d}_T - \bar{d}_C\right] = \delta$ *and* $Var\left[\bar{d}_T - \bar{d}_C\right] = \lambda^2\left(\sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}\right)$ *where* $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$ *and* n_T *and* nB_C *are the number of patients in each treatment group.*

$$E\left[\bar{d}_T\right] = E\left[\frac{\sum d}{n_T}\right] = E\left[\frac{\sum \mu + \delta + \varepsilon_y - \mu_x + \varepsilon_x}{n_T}\right] = \mu + \delta - \mu_x$$

$$E\left[\bar{d}_C\right] = E\left[\frac{\sum d}{n_C}\right] = \mu - \mu_x.$$

Hence $E\left[\bar{d}_T - \bar{d}_C\right] = \delta$.

Due to independence of groups

$$Var\left[\bar{d}_T - \bar{d}_C\right] = Var\left[\bar{d}_T\right] + Var\left[\bar{d}_C\right]$$

Now $Var\left[\bar{d}_T\right] = Var\left[\bar{y}_T - \bar{x}_T\right] = Var\left[\bar{y}_T\right] - 2Cov\left[\bar{y}_T, \bar{x}_T\right] + Var\left[\bar{x}_T\right]$

Since $Var\left[\bar{y}_T\right] = Var\left[\frac{\sum y}{n_T}\right] = \frac{\sigma_y^2}{n_T}$ $Var\left[\bar{x}_T\right] = \frac{\sigma_x^2}{n_T}$ and $Cov\left[\bar{y}_T, \bar{x}_T\right] = \frac{\sigma_{xy}}{n_T}$, it follows that

$$Var\left[\bar{d}_T\right] = \frac{1}{n_T}\left(\sigma_Y^2 + \sigma_X^2 - 2\sigma_{XY}\right).$$

Similarly $Var\left[\bar{d}_C\right]=\frac{1}{n_C}\left(\sigma_Y^2+\sigma_X^2-2\sigma_{XY}\right)$ Therefore

$Var\left[\bar{d}_T-\bar{d}_C\right]=\left(\frac{1}{n_T}+\frac{1}{n_C}\right)\left(\sigma_Y^2+\sigma_X^2-2\sigma_{XY}\right)=\lambda^2\left(\sigma_Y^2+\sigma_X^2-2\sigma_{XY}\right)$ where $\lambda=\sqrt{\frac{1}{n_T}+\frac{1}{n_C}}$.

[5 marks]

(ii) For the test statistic $T=\frac{\bar{d}_T-\bar{d}_C}{\hat{SE}\left[\bar{d}_T-\bar{d}_C\right]}$*, write down an expression for the power of the test to detect a treatment effect* τ *give a two-tailed alternative hypothesis and assuming that the test statistic T has a normal distribution under the null and the alterative hypotheses.*

Assuming that $d=y-x$ is normally distributed it follows that $T=\frac{\bar{d}_T-\bar{d}_C}{SE\left[\bar{d}_T-\bar{d}_C\right]}$ has a t-distribution. Assuming *n* is sufficiently large such that a normal approximation to the central and non-central t-distribution is adequate, *T* has a distribution N(0,1) under H_o and $N\left(\frac{\tau}{SE\left[\bar{d}_T-\bar{d}_C\right]},1\right)$ under H_1.

Since H_o is rejected if $|T|>z_{\alpha/2}$, it follows that

$$Power=1-\beta\cong\left(1-\Phi\left(z_{\alpha/2}-\frac{\tau}{\lambda\sqrt{\sigma_Y^2+\sigma_X^2-2\sigma_{XY}}}\right)\right)+\Phi\left(-z_{\alpha/2}-\frac{\tau}{\lambda\sqrt{\sigma_Y^2+\sigma_X^2-2\sigma_{XY}}}\right)$$ *[5 marks]B.*

(iii) Assuming a two-tailed α *size test, show that the sample size required for each of two equal size groups is* $n=\frac{2\left(\sigma_Y^2+\sigma_X^2-2\sigma_{XY}\right)}{\tau^2}\left(z_{\alpha/2}+z_\beta\right)^2$ *to give a power* β*.*

In the expression for power the right term is negligible if $\tau>0$. Therefore

$$Power=1-\beta\cong1-\Phi\left(z_{\alpha/2}-\frac{\tau}{\lambda\sqrt{\sigma_Y^2+\sigma_X^2-2\sigma_{XY}}}\right)$$

Since $\Phi^{-1}(\beta)=-z_\beta$ it follows that

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

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

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

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

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

[7 marks]

(iv) In a trial of an anti-hypertensive medication, the patient's outcome is to be assess by systolic blood pressure. From previous studies it is known that amongst patients that would be eligible for the trial $\sigma_X^2=\sigma_Y^2=400$ *mmHg*^{2} *and* $\sigma_{XY}=300$ *mmHg*^{2}*. Calculate the minimum sample size required in each of two equal sized groups to have a power of at least 80% power to detect a difference of 5 mmHg between the two treatment where the statistical analysis is based on change in blood pressure.*

$\sigma_Y^2+\sigma_X^2-2\sigma_{XY}=\left(400+400-2\times300\right)=200$

From tables $z_{\alpha/2}=1.96, z_\beta=0.84$

$\tau=5$

Using the above formula

$$n=\frac{2\left(\sigma_Y^2+\theta^2\sigma_X^2-2\theta\sigma_{XY}\right)}{\tau^2}\left(z_{\alpha/2}+z_\beta\right)^2=\frac{2\times200\times\left(1.96+0.84\right)^2}{5^2}=16\times2.8^2=125.4$$

Hence minimum sample size per group for 80% power is 126

[4 marks]

B3.

*Consider a population of which a proportion q are exposed to a suspected carcinogenic substance, the remainder being unexposed. The cumulative incidence rates of cancer in the exposed and unexposed parts of the population over a time interval (0,t) are π_1 and π_0 respectively. Consider three different study designs to estimate ln γ, where $\gamma = \frac{\pi_1(1-\pi_0)}{(1-\pi_1)\pi_0}$. In design **A**, random samples of size m are chosen from each of the exposed and unexposed parts of the population, in design **B** random samples of size m are chosen from people with and people without cancer at time t and in design **C** a random sample of size 2m is chosen from the population as a whole. Assume that, in all designs, the variance of $\ln\hat{\gamma}$ can be approximated by $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}$ where a, b ,c, and d are the frequencies in the cells of the observed 2x2 table.*

(i) Give expressions for the proportions of the whole population in the four cells formed by cross-classifying exposure status and cancer status at t.

	E	**not E**
D	q π_1	(1-q) π_0
not D	q(1- π_1)	(1-q)(1- π_0)

[3 marks]

(ii) Use the results in (i) to estimate the numbers of subjects in the four cells under each of the three designs.

Design A : by design, total E =m and total 'not E' = m.

Since Pr{D|E}= π_1, expected number in cell (D,E) = m π_1; since Pr{D|not E}= π_0, expected number in cell (D,notE) = m π_0 etc, giving

	E	**not E**
D	m π_1	m π_0
not D	m(1- π_1)	m(1- π_0)

[2 marks]

Design B: by design, total D =m and total 'not D' = m.

From (i) $\Pr\{E \mid D\} = \frac{q\pi_1}{q\pi_1 + (1-q)\pi_0}$, hence expected no. in cell (D,E) $= \frac{mq\pi_1}{q\pi_1 + (1-q)\pi_0}$,

Similarly $\Pr\{E \mid notD\} = \frac{q(1-\pi_1)}{q(1-\pi_1) + (1-q)(1-\pi_0)}$, etc. Hence

	E	**not E**
D	$\frac{mq\pi_1}{q\pi_1+(1-q)\pi_0}$	$\frac{m(1-q)\pi_0}{q\pi_1+(1-q)\pi_0}$
not D	$\frac{mq(1-\pi_1)}{q(1-\pi_1)+(1-q)(1-\pi_0)}$	$\frac{m(1-q)(1-\pi_0)}{q(1-\pi_1)+(1-q)(1-\pi_0)}$

[4 marks]

Design C: Pr(D,E) etc same as in whole population, hence expected numbers are:

	E	**not E**
D	$2mq\,\pi_1$	$2m(1-q)\,\pi_0$
not D	$2mq(1-\pi_1)$	$2m(1-q)(1-\pi_0)$

[3 marks]

*(iii) Show that the relative efficiency of design **A** compared to **C** for estimating ln γ is given by* $\frac{1}{4q(1-q)}$ *under H_0: $\pi_1 = \pi_0 = \pi$. Assuming that H_0 is true, in what type of population will design **A** would be more efficient than design **C**?*

Under H_0 and A:

$$V(\ln\hat{\gamma}_A)=2\left(\frac{1}{m\pi}+\frac{1}{m(1-\pi)}\right)=\frac{2}{m\pi(1-\pi)}$$

Under H_0 and C:

$$V(\ln\hat{\gamma}_C)=\frac{1}{2m}\left(\frac{1}{q\pi}+\frac{1}{q(1-\pi)}+\frac{1}{(1-q)\pi}+\frac{1}{(1-q)(1-\pi)}\right)$$

$$=\frac{1}{2m}\left(\frac{1}{q\pi(1-\pi)}+\frac{1}{(1-q)\pi(1-\pi)}\right)$$

$$=\frac{1}{2m\pi(1-\pi)q(1-q)}$$

Eff of $\hat{\gamma}_A$ compared to $\hat{\gamma}_C$ is $\frac{V(\ln\hat{\gamma}_C)}{V(\ln\hat{\gamma}_A)}$

$$=\frac{m\pi(1-\pi)}{2}\frac{1}{2m\pi(1-\pi)q(1-q)}=\frac{1}{4q(1-q)}$$

[5 marks]

Design A is preferable if the latter is greater than 1 $\Rightarrow$q(1-q)< ¼. Since the quadratic function x(1-x) has a maximum at x = ½, when it becomes ¼, this condition is satisfied when q is not equal to ½. So design A is preferable in all populations except when exactly half the population is exposed.
[2 marks]

(iv) *Show that the relative efficiency of design* ***B*** *compared to* ***A*** *for estimating ln γ is given by* $\frac{q(1-q)}{\pi(1-\pi)}$ *when* $\pi_1 = \pi_0 = \pi$. *Assuming that* H_0 *is true* $q < ½$ *and* $\pi < ½$, *what does the study organiser need to know about the population before she can decide which of these two designs is more efficient?*

Under H_0 and B:

$$V(\ln\hat{\gamma}_B) = [q\pi + (1-q)\pi]\left(\frac{1}{mq\pi} + \frac{1}{m(1-q)\pi}\right) + [q(1-\pi) + (1-q)(1-\pi)]\left(\frac{1}{mq(1-\pi)} + \frac{1}{m(1-q)(1-\pi)}\right)$$

$$= \pi\left(\frac{1}{mq(1-q)\pi}\right) + (1-\pi)\left(\frac{1}{mq(1-q)(1-\pi)}\right)$$

$$= \frac{2}{mq(1-q)}$$

Eff of $\hat{\gamma}_B$ compared to $\hat{\gamma}_A$ is $\frac{V(\ln\hat{\gamma}_A)}{V(\ln\hat{\gamma}_B)}$

$$= \frac{mq(1-q)}{2}\frac{2}{m\pi(1-\pi)} = \frac{q(1-q)}{\pi(1-\pi)}$$

[4 marks]

Design **B** is more efficient than **A** when the latter is greater than 1 $\Rightarrow$ q(1-q) > π(1-π).
Since the function x(1-x) increases to a maximum at x= ½ we can conclude that design **B** is more efficient when q > π. So organiser needs to decide which is more common: disease or exposure?

[2 marks]

[Total 25 marks]