Medical Statistics (MATH38071) Exercise Sheet 7

(Equivalence and Non-Inferiority Trials)

1. Quinn et al [Critical Appraisal Exercise 2] compared conservative treatment with suturing for the treatment of minor hand wounds. The primary outcome was *cosmetic appearance* assessed by a 100mm visual analogue scale with 0 and 100 representing the worst and best outcomes respectively. The analysis of this presented in the paper is tabulated below.

	Treatment						*Mean difference*	
	Suture			*Conservative*				
	Mean (mm)	*s.d. (mm)*	*N*	*Mean (mm)*	*s.d. (mm)*	*n*	τ *(mm)*	*95% c.i.*
Cosmetic Appearance	83	10.0	41	80	11.3	40	-3	(-8 to 1)

(i) Suppose that a difference of 5 mm on the visual analogue scale is clinically important. If τ is the treatment effect, test the hypothesis $H_0 : |\tau| \geq \tau_E$ vs $H_1 : |\tau| < \tau_E$ with $\tau_E = 5$ using a 5% significance level.

(ii) Comment on the results of (i).

(iii) Using the pooled within-group standard deviation calculated in (i), estimate the sample size required to have 90% power to reject $H_0 : |\tau| \geq \tau_E$ vs $H_1 : |\tau| < \tau_E$ with $\tau_E = 5$ for a 5% level test assuming that $\tau = 0$ under the alternative hypothesis.

(iv) Calculate the sample size required to test $H_0 : \tau \leq -5$ vs $H_1 : \tau > -5$ for a 5% level test, again assuming that $\tau = 0$ under the alternative hypothesis.

2. A randomised controlled trial is carried out to compare a standard vaccine with a new vaccine that is easier to deliver. The vaccines are allocated to two equal size groups of 200 children each. One week after inoculation the immune response is measured in each child. An immune response is found in 86%(172/200) children with the standard vaccine and of 84%(168/200) children in the new vaccine. The researchers have decided that the vaccines would be considered to have equivalent efficacy if the proportion of children with an immune response in the two treatments differs by less than 5%. Assuming that a $(1-2\alpha)$ confidence interval of proportions corresponds to a test of proportions with a significance level α, test whether the immune response of the new vaccine can be considered to be equivalent to the existing vaccine using a 5% significance level.

[Continued]

3. A randomised controlled non-inferiority trial compared *cognitive behavioural therapy* (CBT) delivered by *telephone* with the same therapy delivered *face-to-face* for the treatment of patients with obsessive compulsive disorder. The trial was testing whether telephone treatment was non-inferior to standard face-to-face CBT. The outcome measure, called the Yale-Brown Obsessive Compulsive Scale (YBOCS) was measured at 3 months after the completion of treatment. Lower scores on the YBOCS scale represent a better outcome. The limit for non-inferiority was 3 units on the YBOCS scale. Using the information in the STATA output below, test whether *telephone* delivery could be considered to be non-inferior to *face-to-face* delivery using :

   (i) a 2.5% significance level.

   (ii) a 5% significance level.

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
Telephone|      34    12.64706    1.280515    7.466622    10.04183    15.25229
Face-to-F|      29    12.93103    1.428019    7.690119    10.00587     15.8562
---------+--------------------------------------------------------------------
    diff |           -.2839757    1.913504               -4.110264    3.542313
------------------------------------------------------------------------------
    diff = mean(Telephone) - mean(Face-to-Face)                   t =  -0.1484
Ho: diff = 0                                     degrees of freedom =       61
    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.4413         Pr(|T| > |t|) = 0.8825          Pr(T > t) = 0.5587
```

4. Consider a continuous and normally distributed outcome measure *Y* for which lower scores correspond to a better outcome. It is suggested in the notes that the null hypothesis $H_0 : \tau \geq \tau_N$ in a non-inferiority trial is rejected if the $(1-\alpha)$ single-sided confidence interval is in the region $(-\infty, +\tau_N)$.

   (i) Assuming that the variance σ^2 is known enabling a normal approximation to the t-distribution, derive an expression for probability of rejection of H_0.

   (ii) Show that the Type I error rate has a maximum when $\tau = \tau_N$.

   (iii) Show that the Type I error is $\leq \alpha$.

5. Consider a continuous and normally distributed outcome measure Y for which low scores correspond to a better outcome. Suppose the null hypothesis $H_0 : \tau \geq \tau_N$ in a non-inferiority trial is rejected if the $(1-\alpha)$ single-sided confidence interval is in the region $(-\infty, +\tau_N)$.

   (i) Suppose that $\tau = \tau_A$ under the alternative hypothesis. Write down an expression for the power of the trial.

   (ii) Show that the sample size per group required in a trial with two equal size treatment groups to test the hypothesis $H_0 : \tau \geq \tau_N$ vs $H_1 : \tau < \tau_N$ is $n = 2\sigma^2 \left(z_\beta + z_\alpha\right)^2 \Big/ \left(\tau_N - \tau_A\right)^2$, when $\tau = \tau_A$ under the alternative hypotheses.

   (iii) Sketch a plot of the sample size per group (*n*) against $\tau_A \in (-\infty, \tau_N)$ marking the intercept $\tau_A = 0$.