Name   **Student ID**

1. Write your name and ID on each page.

2. Numbers in brackets denote total points allotted to each question.

3. You may use calculator.

4. Laptops and phones are not allowed.

5. You are allowed to bring two pages (one A4 sheet two sided OR two A4 sheets one sided) of notes.

6. Show all your work.

Qn no	1	2	3	Total
Marks	20	10	20	50
Obtained				

1. (5+3+8+4=20) In a comparison of the visual acuity of deaf and hearing children, eye movement rates are taken on 10 deaf and 10 hearing children. The data are shown in the accompanying table. A clinical psychologist believes that deaf children have greater visual acuity than hearing children. (The larger a child's eye movement rate, the more visual acuity the child possesses.)

Deaf Children	2.75 1.95 3.14 2.17 3.23 2.45 2.30 1.83 2.64 2.23
Hearing Children	1.15 1.23 1.65 2.03 1.43 1.64 1.83 1.96 1.75 1.37

(a) Set up the problem mathematically, defining the random variables, model and the null/alternate hypotheses.

(b) We decide to do a Wilcoxon test. What is the name of the test and what are the assumptions?

(c) Calculate the value of the test statistic.

(d) In the following code to find p-value, what should be n, m and T?

```
pvalue<-function(n,m,T){
  T=T+m*(m+1)/2
x<-c(1:n)
l<-choose(n,m)
k<-rep(0,l)
y<-combn(x,m)
for(j in 1:l){
k[j]<-sum(x[y[,j]])}
  pval<-length(which(k>=T))/choose(n,m)
return(pval)
}
```

Name **Student ID**

2. (5+5=10) A continuous random variable Y is symmetric around zero if $P(Y < -a) = P(Y > a)$ for all a.

   (a) Following this definition, when can you say that a random variable X is symmetric around μ? Your response should be in terms of probabilities.

   (b) Show that your answer in part is equivalent to

$$F_{X-\mu}(a) = F_{\mu-X}(a)$$

   for all a, where $F_V(x)$ denotes the cdf of random variable V at x.

3. (10+5+5=20) A trial measured the relative efficacy of a treatment for a variety of doses.

Dose	0	3	3	4	5	6	8	8	9	10
Efficacy	2.56	0.78	4.79	2.44	1.01	1.05	-1.07	-1.37	-6.18	-12.03

We want to model the efficacy as the dependent variable in a nonparametric regression model. In particular, we'll use a Kernel Estimator with an Epanechnikov kernel with bandwidth 2.

$$K_2(t) = \frac{3}{8}\left(1 - \frac{t^2}{4}\right) \quad \text{for} \quad -2 \leq t \leq 2.$$

(a) Calculate the regression estimate this would give for efficacy when the dose is 7.

(b) Calculate the error (fitted minus actual) in fitting the efficacy at the point where the Dose= 7 if the true mean function is

$$E(Efficacy) = Dose - \frac{1}{10}Dose^2$$

(c) If instead I used a bandwidth of 3, what will be the Epanechnikov kernel?