HW1- Advanced Data Analysis

1. (5pt) Let η denote the median of a random variable X. Consider testing $H_0 : \eta = 0$ against $H_a : \eta \neq 0$ using $X_1, X_2, \ldots, X_{25}$, a random sample of size $n = 25$ from the distribution of X

   (a) (2.5pt) Let S denote the sign statistic. Determine the level of the test that rejects H_0 is $S \geq 16$. (Hint: Compute $P(S \geq 16)$ under H_0.)

   (b) (2,5pt) Determine the power of the test in a) if X has N(0.5, 1) distribution (Hint: compute $P(S \geq 16)$. For this you will need the distribution of S when $X \sim N(0.5, 1)$)

2. (10pt) The data in in the table below gives the pretest and posttest scores on the MLA listening test in Spanish for 20 high school teachers who attended an intensive course in Spanish.

subject	pretest	posttest	subject	pretest	posttest
1	30	20	11	30	32
2	28	30	12	29	22
3	31	32	13	31	34
4	26	30	14	29	32
5	20	16	15	34	32
6	30	25	16	20	27
7	34	31	17	26	28
8	15	18	18	25	29
9	28	33	19	31	32
10	20	25	20	29	32

   Assume that the differences between these scores (pretest scores $-$posttest) constitute a random sample from a distribution F with mean μ and variance σ^2

   (a) (2.5pt) Use a t-test and $\alpha = 0.05$ to test $H_0 : \mu = 0$ against $H_a : \mu \neq 0$. What is the p-value of the test? What assumption you need to make. Use a graphical technique to check this assumption.

   (b) (2,5pt) Obtain a 95% confidence interval for the mean in a)

   (c) (2,5pt) If the median of F is η, use the sign test and $\alpha = 0.05$ to test $H_0 : \eta = 0$ against $H_a : \eta \neq 0$. What is the p-value of this test?

(d) (2,5pt) Obtain a 95% confidence interval for η and compare use answer the answer in b)

3. (5pt) Twelve one week old infants were randomly assigned into two groups of six infant each. One group participated in an experimental active-exercise to learn to walk and the other was used as a control group. The following are the ages at which these infants first walked alone

Active-execise group	No-exercise group
9.00	11.50
9.50	12.00
9.75	9.00
10.00	11.50
13.00	13.25
9.50	13.00

Call the no-exercise group Y sample and the active-exercise group X sample.

(a) Compare the two groups using two tests (one parametric and one nonparametric) (take $\alpha = 0.05$). State all the assumptions that you make in carrying out these tests.

(b) Define

$$D = \begin{cases} 1 & \text{infant in active-exercise group} \\ -1 & \text{otherwise} \end{cases}$$

let $y =$ at which an infant worked. Fit the regression model

$$y = \beta_0 + \beta_1 D + \epsilon$$

to this data and test $H_0 : \beta_1 = 0$ against $H_a : \beta_1 \neq 0$. Compare your answer with the t-test result