Math 159, Spring 2005
Professor Hardin
HW #4b
due Tuesday 3/1/05

Additional Problem: The sign test of section 3.4 is a nonparametric version of the paired t-test. Using the steps below, compare the size and power of the paired t-test vs. the sign test for 3 data sets: normal, t_3 (heavy tailed), and $chisq_3$ (skewed right).

1. Consider the distribution of the **paired** data; generate 6 data sets. 3 different types (above); 2 different centers (the null center should be obvious, choose the alternative center to be slightly larger); all std. dev. should be sqrt(3); let your sample size be n=50.

2. You should have two different (types of) test statistics, what are they?

3. For what values would you reject the null hypothesis?

4. Describe how you would calculate probabilities of type I and type II errors (α and β).

5. Plot histograms of your test statistics for the two different data sets (null data set and alternative data set.)

6. Calculate a simulated α and β value for each of the 6 tests (3 pairs of data sets, 2 different tests.)

7. What can you say about the relationship of the paired t-test to the sign test for different types of data? Explain.

The R code for the plots comparing the t-test with the median test is:

```
#############
# t-test vs. median Ho: mu=2, H1: mu=3
#############

>mu1<-2
>mu2<-3
>n<-50
>alpha<-.05
>reps<-10000

>z1.t.ts<-c()
>z2.t.ts<-c()
>z1.t.sd<-c()
>z2.t.sd<-c()
>z1.m.ts<-c()
>z2.m.ts<-c()
```

```
>t1.t.ts<-c()
>t2.t.ts<-c()
>t1.t.sd<-c()
>t2.t.sd<-c()
>t1.m.ts<-c()
>t2.m.ts<-c()

>de1.t.ts<-c()
>de2.t.ts<-c()
>de1.t.sd<-c()
>de2.t.sd<-c()
>de1.m.ts<-c()
>de2.m.ts<-c()

>csq1.t.ts<-c()
>csq2.t.ts<-c()
>csq1.t.sd<-c()
>csq2.t.sd<-c()
>csq1.m.ts<-c()
>csq2.m.ts<-c()

>out1.t.ts<-c()
>out2.t.ts<-c()
>out1.t.sd<-c()
>out2.t.sd<-c()
>out1.m.ts<-c()
>out2.m.ts<-c()

>for(i in 1:reps){
z1 <- rnorm(n,mean=mu1,sd=sqrt(3))
z2 <- rnorm(n,mean=mu2,sd=sqrt(3))

t1 <- rt(n,df=3) + mu1
t2 <- rt(n,df=3) + mu2

de1<-ifelse(runif(n)>.5,mu1+ rexp(n)*sqrt(3),mu1 - rexp(n)*sqrt(3))
de2<-ifelse(runif(n)>.5,mu2+ rexp(n)*sqrt(3),mu2 - rexp(n)*sqrt(3))

csq1<-(rchisq(n,df=3)-3)/sqrt(6)*sqrt(3)+mu1
csq2<-(rchisq(n,df=3)-3)/sqrt(6)*sqrt(3)+mu2

out1<-c(rnorm(n-1,mean=mu1,sd=sqrt(3)),c(47))
out2<-c(rnorm(n-1,mean=mu2,sd=sqrt(3)),c(47))
```

```
z1.t.ts<-c(z1.t.ts,mean(z1))
z2.t.ts<-c(z2.t.ts,mean(z2))
z1.t.sd<-c(z1.t.sd,sd(z1))
z2.t.sd<-c(z2.t.sd,sd(z2))
z1.m.ts<-c(z1.m.ts,sum(z1<=mu1))
z2.m.ts<-c(z2.m.ts,sum(z2<=mu1))

t1.t.ts<-c(t1.t.ts,mean(t1))
t2.t.ts<-c(t2.t.ts,mean(t2))
t1.t.sd<-c(t1.t.sd,sd(t1))
t2.t.sd<-c(t2.t.sd,sd(t2))
t1.m.ts<-c(t1.m.ts,sum(t1<=mu1))
t2.m.ts<-c(t2.m.ts,sum(t2<=mu1))

de1.t.ts<-c(de1.t.ts,mean(de1))
de2.t.ts<-c(de2.t.ts,mean(de2))
de1.t.sd<-c(de1.t.sd,sd(de1))
de2.t.sd<-c(de2.t.sd,sd(de2))
de1.m.ts<-c(de1.m.ts,sum(de1<=mu1))
de2.m.ts<-c(de2.m.ts,sum(de2<=mu1))

csq1.t.ts<-c(csq1.t.ts,mean(csq1))
csq2.t.ts<-c(csq2.t.ts,mean(csq2))
csq1.t.sd<-c(csq1.t.sd,sd(csq1))
csq2.t.sd<-c(csq2.t.sd,sd(csq2))
csq1.m.ts<-c(csq1.m.ts,sum(csq1<=(qchisq(.5,3)-3)/sqrt(6)*sqrt(3)+mu1))
csq2.m.ts<-c(csq2.m.ts,sum(csq2<=(qchisq(.5,3)-3)/sqrt(6)*sqrt(3)+mu1))

out1.t.ts<-c(out1.t.ts,mean(out1))
out2.t.ts<-c(out2.t.ts,mean(out2))
out1.t.sd<-c(out1.t.sd,sd(out1))
out2.t.sd<-c(out2.t.sd,sd(out2))
out1.m.ts<-c(out1.m.ts,sum(out1<=mu1))
out2.m.ts<-c(out2.m.ts,sum(out2<=mu1))
}

>hist(z1.t.ts,xlim=c(0,6),ylim=c(0,5000),xlab="sample mean",main="z / t",
breaks=10,density=10)
>hist(z2.t.ts,add=T,breaks=6,density=12,angle=-60)
>legend(3,2750,list(paste("alpha =",sum(z1.t.ts>=mu1+
qt(1-alpha,n-1)*z1.t.sd/sqrt(n))/reps),
paste("beta =",sum(z2.t.ts<=mu1+qt(1-alpha,n-1)*
z2.t.sd/sqrt(n))/reps)),bty="n")
```

```
>hist(z1.m.ts/n,xlim=c(0,1),ylim=c(0,4000),xlab="prop <= median",
main="z / median", breaks=10,density=10)
>hist(z2.m.ts/n,add=T,breaks=10,density=12,angle=-60)
>legend(.5,3500,list(paste("alpha =",sum(z1.m.ts <= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps),
paste("beta =", sum(z2.m.ts >= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps)),bty="n")

>hist(t1.t.ts,xlim=c(0,8),ylim=c(0,5000),xlab="sample mean",main="t / t",
breaks=15,density=10)
>hist(t2.t.ts,add=T,breaks=15,density=12,angle=-60)
>legend(3,2750,list(paste("alpha =",sum(t1.t.ts>=mu1+
qt(1-alpha,n-1)*t1.t.sd/sqrt(n))/reps),
paste("beta =",sum(t2.t.ts<=mu1+qt(1-alpha,n-1)*
t2.t.sd/sqrt(n))/reps)),bty="n")

>hist(t1.m.ts/n,xlim=c(0,1),ylim=c(0,5000),xlab="prop <= median",
main="t / median", breaks=10,density=10)
>hist(t2.m.ts/n,add=T,breaks=10,density=12,angle=-60)
>legend(.5,3500,list(paste("alpha =",sum(t1.m.ts <= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps),
paste("beta =", sum(t2.m.ts >= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps)),bty="n")

>hist(de1.t.ts,xlim=c(0,6),ylim=c(0,5000),xlab="sample mean",main="doub exp / t",
breaks=8,density=10)
>hist(de2.t.ts,add=T,breaks=5,density=12,angle=-60)
>legend(3,2750,list(paste("alpha =",sum(de1.t.ts>=mu1+
qt(1-alpha,n-1)*de1.t.sd/sqrt(n))/reps),
paste("beta =",sum(de2.t.ts<=mu1+
qt(1-alpha,n-1)*de2.t.sd/sqrt(n))/reps)),bty="n")

>hist(de1.m.ts/n,xlim=c(0,1),ylim=c(0,4000),xlab="prop <= median",
main="doub exp / median", breaks=10,density=10)
>hist(de2.m.ts/n,add=T,breaks=10,density=12,angle=-60)
>legend(.5,3500,list(paste("alpha =",sum(de1.m.ts <= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps),
paste("beta =", sum(de2.m.ts >= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps)),bty="n")
```

```
>hist(csq1.t.ts,xlim=c(0,5),ylim=c(0,3500),xlab="sample mean",main="chisq / t",
breaks=10,density=10)
>hist(csq2.t.ts,add=T,breaks=15,density=12,angle=-60)
>legend(2.5,2750,list(paste("alpha =",sum(csq1.t.ts>=mu1+
qt(1-alpha,n-1)*csq1.t.sd/sqrt(n))/reps),
paste("beta =",sum(csq2.t.ts<=mu1+
qt(1-alpha,n-1)*csq2.t.sd/sqrt(n))/reps)),bty="n")

>hist(csq1.m.ts/n,xlim=c(0,1),ylim=c(0,4000),xlab="prop <= median",
main="chisq / median",breaks=10,density=10)
>hist(csq2.m.ts/n,add=T,breaks=10,density=12,angle=-60)
>legend(.5,3500,list(paste("alpha =",sum(csq1.m.ts <= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps),
paste("beta =", sum(csq2.m.ts >= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps)),bty="n")

>hist(out1.t.ts,xlim=c(1.5,6),ylim=c(0,4000),xlab="sample mean",main="out / t",
breaks=10,density=10)
>hist(out2.t.ts,add=T,breaks=10,density=12,angle=-60)
>legend(3.7,3250,list(paste("alpha =",sum(out1.t.ts>=mu1+
qt(1-alpha,n-1)*out1.t.sd/sqrt(n))/reps),
paste("beta =",sum(out2.t.ts<=mu1+
qt(1-alpha,n-1)*out2.t.sd/sqrt(n))/reps)),bty="n")

>hist(out1.m.ts/n,xlim=c(0,1),ylim=c(0,4000),xlab="prop <= median",
main="out / median", breaks=10,density=10)
>hist(out2.m.ts/n,add=T,breaks=10,density=12,angle=-60)
>legend(.5,3500,list(paste("alpha =",sum(out1.m.ts <= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps),
paste("beta =", sum(out2.m.ts >= .5*n -
qnorm(1-alpha)*sqrt(n*.25))/reps)),bty="n")
```