STAT 333
Section 3.4: The Sign Test

• The **sign test,** as we will typically use it, is a method for analyzing paired data.

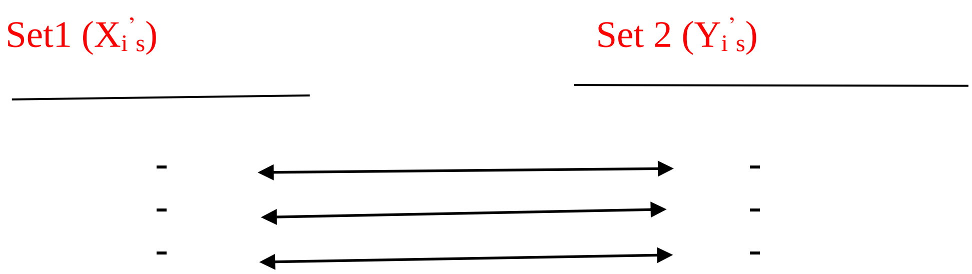
Examples of Paired Data:

• Similar subjects are paired off and one of two treatments is given to each subject in the pair.

or

• We could have two observations on the same subject.

The key: With paired data, the pairings cannot be switched around without affecting the analysis.



• We might label one of the variables X_i and the other variable Y_i.

• Our entire bivariate data set for *n*' **individual pairs** is:

(X_1, Y_1) , (X_2, Y_2), (X_3, Y_3),……… (X_n, Y_n)

• The bivariate random vectors are assumed to be independent across observations.

• The goal may be to determine whether the *X* variable tends to be larger than or smaller than the corresponding *Y* variable.

• Assuming the data are at least ordinal, we could classify each pair as "+" if $X_i < Y_i$ or "–" if $X_i > Y_i$.

• If $X_i = Y_i$ then the pair is classified as "0" or "tie".

• We further assume internal consistency: If P(+) > P(–) for one pair, then P(+) > P(–) for all pairs, and same holds for P(+) < P(–) and P(+) = P(–).

Test Statistic:

T = Total number of +'s in the sample

• **The null distribution of *T* is**

Binomial (n, p=0.5)

Where , *n* = number of non-tied pairs. (n=n - tie pairs)

• The **hypotheses of the sign test** can be stated in a variety of ways.

• Most generally, we **can test** any one of:

H_0: P(+) = P(-)	H_0: P(+) ≥ P(-)	H_0: P(+) ≤ P(-)
H_1: P(+) ≠ P(-)	H_1: P(+) < P(-)	H_1: P(+) > P(-)

• These could be stated in terms of comparing the population medians of *X* and *Y*:

H_0: Med(Y) = Med(X)	H_0: Med(Y) ≥ Med(X)	H_0: Med(Y) ≤ Med(X)
H_1: Med(Y) ≠ Med(X)	H_1: Med(Y) < Med(X)	H_1: Med(Y) >Med(X)

• The corresponding **rejection rules in each case are:**

H_1: Med(Y) ≠ Med(X)	H_1: Med(Y) < Med(X)	H_1: Med(Y) >Med(X)
Reject H_0 if $T \leq t$ or $T \geq n-t$	Reject H_0 if $T \leq t$	Reject H_0 if $T \geq n-t$
Where $P(T \leq t) \leq \alpha/2$ Under H_0	Where $P(T \leq t) \leq \alpha$ Under H_0	Where $P(T \leq t) \leq \alpha$ Under H_0

• The **P-values** for each case are:

H_1:	Med(Y) ≠ Med(X)	Med(Y) < Med(X)	Med(Y) >Med(X)
P-value	2x[min{ $P(T \leq t_{obs})$, $P(T \geq t_{obs})$}]	$P(T \leq t_{obs})$	$P(T \geq t_{obs})$

where *T* ~ Binomial(*n*, 0.5).

• The exact critical region and P-value are found using Table A3 (or the computer) similarly to the other binomial-based tests.

Example 1:

Six students are given two tests, one after being fed, and one on an empty stomach. Is there evidence that students perform better on a full stomach? (Use $\alpha = 0.05$)

	Student					
Score	1	2	3	4	5	6
X (with food)	74	71	82	77	72	81
Y (without food)	68	71	86	70	67	80
+/-	-	0	+	-	-	-

Solution :

$n = 5$, $T = 1$ (number of +'s in the sample)

Hypothesis:

H_0: $Med(Y) \geq Med(X)$ H_1: $Med(Y) < Med(X)$

Decision rule :

Reject H_0 if $T \leq t$

Reject H_0 if $T \leq 0$

So, we accept H_0 since $T = 1$

Note: can get the value of t from Table A3:Bin(5,0.5) $P(Y \leq 0) = 0.0312 \leq 0.05$

P-value = $P(T \leq 1) = 0.1875$ (Accept H_0)

We conclude that median score (without food) is more than or equal the median score (with food)

Example 2:

18 boy/girl sets of twins were scored for "empathy" on a personality test. In ten sets, the girl scored higher; in 7 sets, the boy scored higher, and in one set, the scores were equal. Can we conclude a difference in median empathy between the boy and girl populations? (Use $\alpha = 0.05$)

Solution :

Let X =boy , Y = girl

H_0: Med(Y) = Med(X) **H_1:** Med(Y) ≠ Med(X)

T = number of +'s = 10

n = 17 (n= 18 – tie pairs)

Decision rule:

Reject H_0 if $T \leq t$ or if $T \geq n - t$

Reject H_0 if $T \leq 4$ or if $T \geq 17 - 4 = 13$

Note: can get the value of t from Table A3:Bin(17,0.5)

$\alpha/2 = 0.05/2 = 0.025$

$P(Y \leq 4) = 0.0245 \leq 0.025$

Since,

$T = 10$, we accept H_0

P-value $= 2 \text{ x}[\min\{P(T \leq 10), P(T \geq 10)]$

$= 2 \text{ x } [1- 0.6855] = 0.629$ (P-value > α)

We **cannot conclude** any difference in median empathy between boys and girls.(Median for boys is equal the median for girls)

Some Notes

• One way to view the sign test is simply as the binomial test, where $p^* = 0.5$.

• We classify each trial as "+" or "–" and determine whether the probability of "+" is different from/greater than/ less than 0.5.

• Note that performing the quantile test about the median is essentially performing the sign test, where the second variable is simply the constant number x^*.

• The sign test is appropriate for any data measured on an ordinal or stronger scale.

• If the paired differences $Y_i - X_i$ are continuous with a symmetric distribution, the **Wilcoxon signed-rank test** (we will see it in Chapter 5) may be more powerful than the sign test.

• If the paired differences $Y_i - X_i$ have **a normal distribution**, the **paired t-test** is the most powerful option.

Efficiency of the Sign Test

Population	A.R.E (Sign vs signed- rank)	A.R.E (Sign vs paired-t)
Normal	0.667	0.637
Uniform(light tail)	0.333	0.333
Double exponential (heavy tail)	1.333	2.00

- Sign test works well for heavy-tailed distributions.

R code :

Example 1: (with raw paired data)

```
scores.with.food <- c(74,71,82,77,72,81)

scores.without.food <- c(68,71,86,70,67,80)
```

 # We can perform the sign test with the sign.test function:

```
sign.test<-function(x=0,y=NULL,alternative="two.sided"){
  n<-sum((x-y)!=0)
  T<-sum(x<y)
  if (alternative=="less") {
    p.value<-pbinom(T,n,0.5)}
  if (alternative=="greater"){
    p.value<- 1-pbinom(T-1,n,0.5)}
  if (alternative=="two.sided"){
    p.value<-2*min(1-pbinom(T-1,n,0.5),pbinom(T,n,0.5))}
  list(n=n,alternative=alternative,T=T,p.value=p.value)}

# Copy and paste this function into R, and then it can be implemented
# as in the following examples:

# When calculating the differences, be careful which variable is labeled "y"
and which variable labeled "x":

sign.test(x=scores.with.food, y=scores.without.food, alternative="less")
```

output:

```
$n
[1] 5

$alternative
[1] "less"

$T
[1] 1

$p.value
[1] 0.1875
```

Example 2: (with counts given)

We can perform the sign test with the sign.test.counts function:

```
sign.test.counts<-
function(plus.count,minus.count,zero.count=0,alternative="two.sided"){
  n<-plus.count+minus.count
  T<-plus.count
  if (alternative=="less") {
    p.value<-pbinom(T,n,0.5)}
  if (alternative=="greater"){
    p.value<- 1-pbinom(T-1,n,0.5)}
  if (alternative=="two.sided"){
    p.value<-2*min(1-pbinom(T-1,n,0.5),pbinom(T,n,0.5))}
  list(n=n,alternative=alternative,T=T,p.value=p.value)}
```

Using this function on the twin data set:

```
sign.test.counts(plus.count=10,minus.count=7,zero.count=1,alternative="two.sided")
```

output:

```
$n
[1] 17

$alternative
[1] "two.sided"

$T
[1] 10

$p.value
[1] 0.6290588
```