STAT 518 --- Section 3.4: The Sign Test

• The <u>sign test</u>, as we will typically use it, is a method for analyzing <u>paired</u> data.

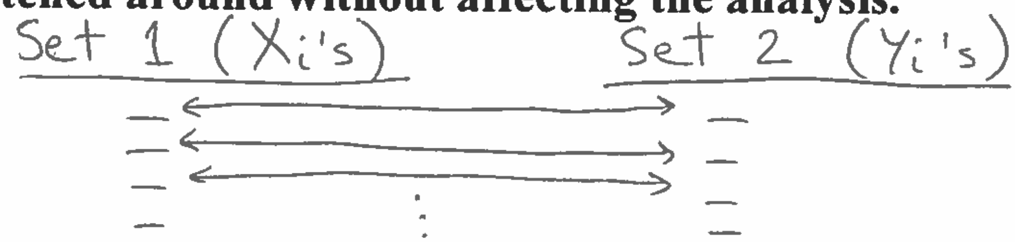
Examples of Paired Data:

• 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:

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_{n'}, Y_{n'})$$

• 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 nontied pairs

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

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

H₀: $P(+) = P(-)$	**H₀:** $P(+) \geq P(-)$	**H₀:** $P(+) \leq P(-)$
H₁: $P(+) \neq P(-)$	**H₁:** $P(+) < P(-)$	**H₁:** $P(+) > P(-)$

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

H₀: $Med(Y) = Med(X)$	**H₀:** $Med(Y) \geq Med(X)$	**H₀:** $Med(Y) \leq Med(X)$
H₁: $Med(Y) \neq Med(X)$	**H₁:** $Med(Y) < Med(X)$	**H₁:** $Med(Y) > Med(X)$

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

Reject Ho if $T \le t$ or $T \ge n-t$ where $P(T \le t) \le \alpha/2$ under Ho	Reject Ho if $T \le t$ where $P(T \le t) \le \alpha$ under Ho	Reject Ho if $T \ge n-t$ where $P(T \le t) \le \alpha$ under Ho

• **The P-values for each case are:**

$2[\min\{P(T \le t_{obs}), P(T \ge t_{obs})\}]$	$P[T \le t_{obs}]$	$P[T \ge t_{obs}]$

where $T \sim$ Binom(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 $\alpha = .05$.)

Scores \ **Student**	**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	+	−	−	−

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

H_0: med(Y) $\ge$ med(X) H_1: med(Y) $<$ med(X)

Decision rule: Reject Ho if $T \le 0$

Note: $P(Y \le 0) = .0312 \le .05$

So we fail to reject Ho since $T = 1$.

P-value $= P(T \le 1) = 0.1875$

– We do not conclude the median score without food is less than 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 $\alpha = .05$.) Let X = boy Y = girl

H_0: med(Y) = med(X) H_1: med$(Y) \neq$ med(X)

T = # of +'s = 10, $n = 17$

Decision rule:
Reject H_0 if $T \le 4$ or if $T \ge 17 - 4 = 13$

$\boxed{\alpha/2 = .025 \text{ and } P(Y \le 4) = .0245}$

Since $T = 10$, we fail to reject H_0.

P-value $= 2[\min\{P(T \le 10), P(T \ge 10)\}]$

$= 2[1 - .6855] = \boxed{.629}$

– We cannot conclude any difference in median empathy between boys and 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 <u>Wilcoxon signed-rank test</u> (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 <u>paired t-test</u> 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 tails)	0.333	0.333
Double exponential (heavy tails)	1.333	2.00

– Sign test works well for heavy-tailed distributions.