Sign test

If X is cont.
$P[X = M_0] = 0$.

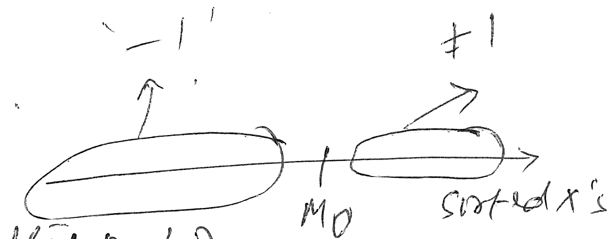
Data: $X_1, \ldots, X_n$ — i.i.d. sample from X.

pop. median → null value.

Hypotheses: $H_0 : M = M_0$ vs. H_1 : one of three possibilities, $M > M_0$ or $M < M_0$ or $M \neq M_0$

Signs: Remove the X's that are equal to M_0 and reduce the sample size accordingly. — $n^* = \#$ obs. that are not equal to M_0.

sign of $X_i = \begin{cases} +1, & X_i > M_0 \\ -1, & X_i < M_0. \end{cases}$



Test statistic:

$S = \#$ positive signs.

Null distribution:

If H_0 is true: sign of $X_i \sim \text{Bernoulli}(p = \frac{1}{2})$

$\Rightarrow S \sim \text{Bin}(n^*, p = \frac{1}{2})$.

When to reject H_0?

- $H_1 : M > M_0$: S is too large (compared with $\frac{n^*}{2}$) p-value: $P[\text{Bin}(n^*, \frac{1}{2}) \geq S_{obs}]$
- $H_1 : M < M_0$: S is too small (" " $\frac{n^*}{2}$) $P[\text{Bin}(n^*, \frac{1}{2}) \leq S_{obs}]$
- $H_1 : M \neq M_0$: S is either too large or too small.

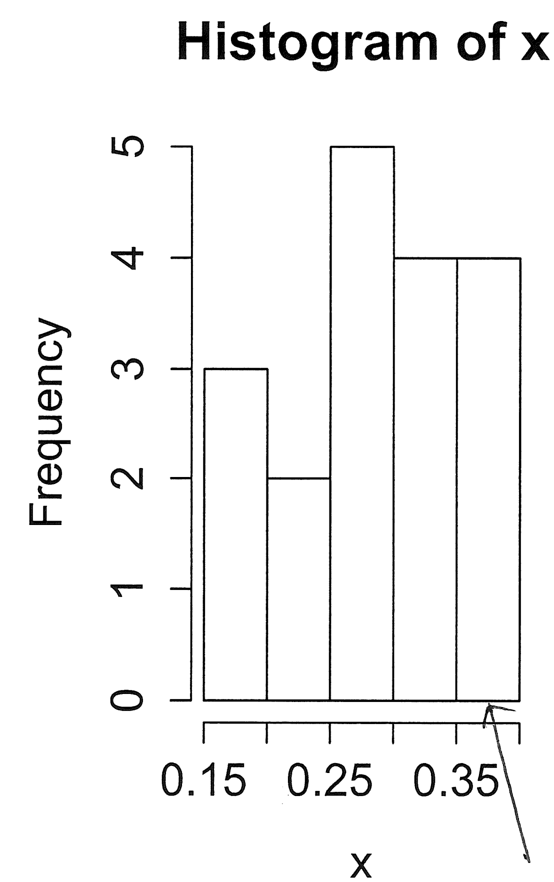
$2 \min\{$the two one-sided p-values$\}$.

R code:

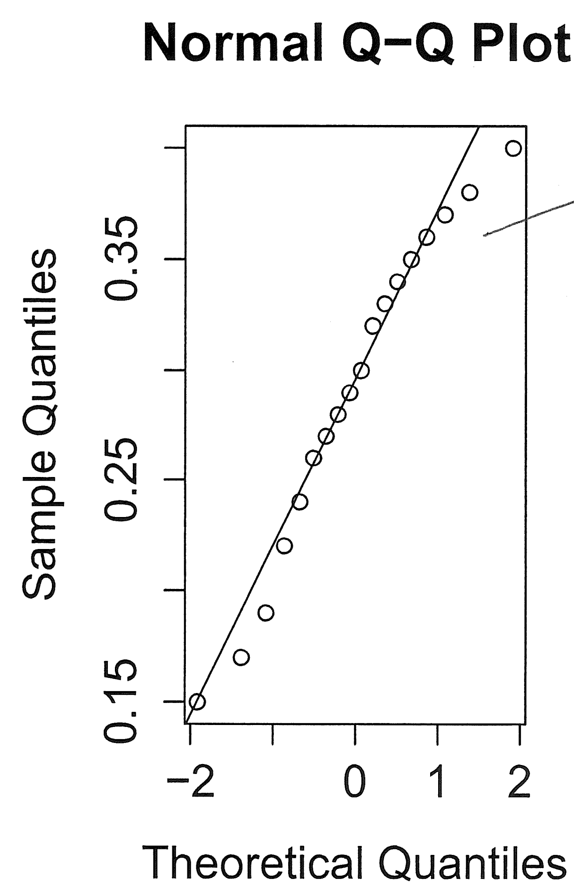
```
# Time between keystrokes data from Example 10.9

x <- c(0.24, 0.22, 0.26, 0.34, 0.35, 0.32, 0.33, 0.29,
       0.19, 0.36,
       0.30, 0.15, 0.17, 0.28, 0.38, 0.40, 0.37,0.27)


# Histogram and boxplot

par(mfrow=c(1,2)) # 2 plots in 1 row

hist(x)
qqnorm(x)
qqline(x)

library(nortest)
```

```
> shapiro.test(x)

Shapiro-Wilk normality test

data:  x
W = 0.9611, p-value = 0.6233
>
```

Normality appears reasonable from this Q-Q plot as well



suggests uniform distribution.



normality appears reasonable.

NB Both models may be reasonable for these data. With a larger sample size, we may be able to better distinguish b/w the two distributions.

```
# Sign test of H0: M = 0.2 vs M is not equal to 0.2

sign.stat <- sum(x > 0.2)

> sign.stat
[1] 15
>

> binom.test(sign.stat, n=sum(x != 0.2), p = 0.5,
alt="two.sided", conf.level=0.95)

Exact binomial test

data:  sign.stat and sum(x != 0.2)
number of successes = 15, number of trials =
  18, p-value = 0.007538
alternative hypothesis: true probability of
success is not equal to 0.5
```

Median of population

$M \neq 0.2$ (sec)

~~median~~ Null value.

$p = P[+ \text{ sign}]$
$= P[X > M_0]$.

$H_0: M = M_0 \Leftrightarrow p = \frac{1}{2}$
$H_1: M \neq M_0 \Leftrightarrow p \neq \frac{1}{2}$

$\Rightarrow$ Reject H_0, conclude: attempt is unauthorized.

```
95 percent confidence interval:
 0.5858225 0.9642149
sample estimates:
 probability of success
0.8333333
>
```

$\longrightarrow$ CI for p $\uparrow$

success — p.

$\equiv$ (+)ve sign

H_0: Q_γ

100 ~~pth~~ qu

$$H_0: Q_\gamma = Q_0 \quad \text{vs} \quad H_1: \begin{cases} Q_\gamma \neq Q_0 \\ Q_\gamma > Q_0 \\ Q_\gamma < Q_0. \end{cases}$$

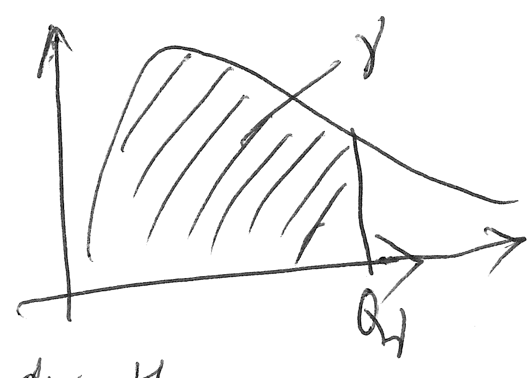
$$\text{sign of } x_i = \begin{cases} +1, & x_i > Q_0 \\ -1, & \text{otherwise.} \end{cases}$$

$$p = P[\oplus\text{ve sign} \mid Q_\gamma = Q_0] = P[X_i > Q_0 \mid Q_\gamma = Q_0]$$

$$= 1-\gamma.$$

Sign test stat = ~~Sum~~ # of $\oplus$ ve signs

$\sim Bin(n^*, p = 1-\gamma)$ under H_0.



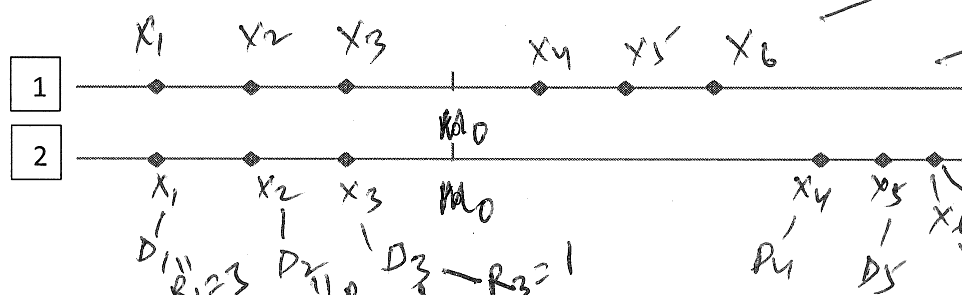
Wilcoxon Signed Rank Test

$R_1 = R_6 = 5.5$ $R_2 = R_5 = 3.5$

Issue: Sign test does ignores a lot of information in the data — only uses whether an observation is above M_0 or not.

$D_1 = D_6$, $D_2 = D_5$, $D_3 = D_4$ $R_3 = R_4 = 1.5$

Ex: Consider two datasets for testing H_0: $M = M_0$.



$S = 3$, $W = R_4 + R_5 + R_6 = 1.5 + 3.5 + 5.5 = 10.5$

$S = 3$.

$W = 4 + 5 + 6 = 15$.

The sign test statistic is the same for both datasets. Does one dataset have stronger evidence that true $M > M_0$?

Yes, 2nd one.

Ranks: Sort the n observations in increasing order. The smallest observation has rank 1, the next smallest has rank 2, . . ., the largest observation has rank n.

rank: 1, 6, 3.5, 5, 3.5, 2

Ex: Find ranks of the following observations: 3, 7, 5, 6, 5, 4.

sorted: 3, 4, 5, 5, 6, 7

Avg the ranks in case of a tie

Test statistic:

- Step 1: Find the distances $D_i = |X_i - M_0|$ between the observations and M_0.
- Step 2: Compute the ranks R_i of the D_i (not the X_i)
- Step 3: Consider only the ranks for those X_i that are greater than M_0, and sum them. In other words, compute positive signs -

$$W = \sum_{i:X_i>m} R_i$$

When to reject H_0?

- $H_1 : M > M_0$: when W is too large $\Rightarrow$ +ve ranks tend to be higher
- $H_1 : M < M_0$: when W is too small.
- $H_1 : M \neq M_0$: W is either too large or too small.

Null distribution: Assume X is symmetric. Use formula 10.6 in textbook for small n. For $n \geq 15$, W is approximately normal with mean $n(n+1)/4$ and variance $n(n+1)(2n+1)/24$.

R code:

```
# Example 10.12 on page 319

# Wilcoxon signed rank test H0: M = 0.2 vs
# M is not equal to 0.2

x <- c(7, 5.5, 9.5, 6, 3.5, 9)

> wilcox.test(x,alternative = "greater",
              mu = 5, conf.level = 0.95)

Wilcoxon signed rank test

data:  x
V = 18, p-value = 0.07813
alternative hypothesis: true location is greater than 5
>
```

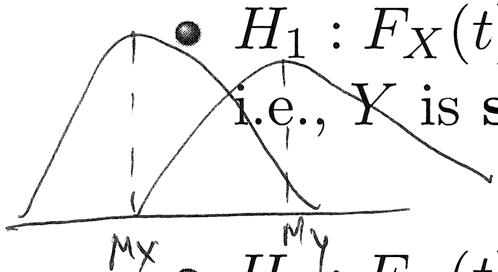
Ex 10.13. keystrokes data
wilcox.test(x, alternative = "two.sided", mu = 0.2, conf.level = 0.95)
gives: V = 1[illegible], p-value = <0.001

Rank Sum Test

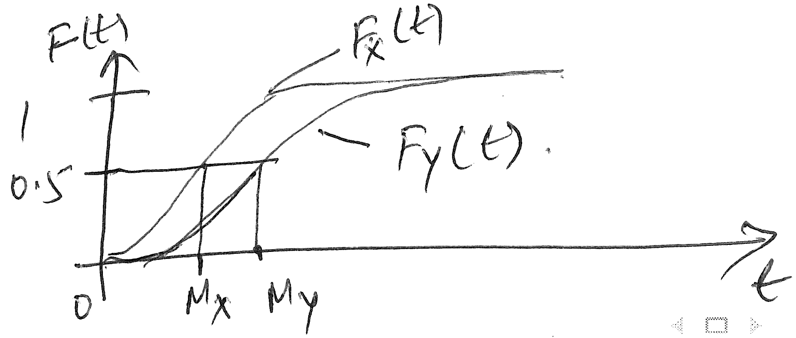
Set up: Have two independent samples: $X_1, \ldots, X_n$ — i.i.d. as X with cdf F_X, and $Y_1, \ldots, Y_m$ — i.i.d. as Y with cdf F_Y.

Hypotheses: H_0: $\underbrace{F_X(t) = F_Y(t)}_{\Rightarrow M_X = M_Y}$ for all t. Three possibilities for H_1:

⇒ $M_X = M_Y$, but the converse is NOT true.



- $H_1 : F_X(t) \geq F_Y(t)$, with strict inequality for at least one t, i.e., Y is **stochastically larger** than X. In this case, $M_X < M_Y$
  (this relation hold for more general quantiles)
- $H_1 : F_X(t) \leq F_Y(t)$, with strict inequality for at least one t, i.e., Y is **stochastically smaller** than X. In this case, $M_Y < M_X$
  ( ).
- $H_1 : F_X(t) \neq F_Y(t)$ for at least one t. In this case,



Test statistic:

- Step 1: Combine all the X and Y data into one sample
- Step 2: Rank observations of the combined sample. The ranks are from 1 to $(n+m)$.
- Step 3: Find U as the sum of all X-ranks.

When to reject H_0?

- $H_1 : Y$ is stochastically larger than X — U is too small
- $H_1 : Y$ is stochastically smaller than X — U is too large
- $H_1 : F_Y(t) \neq F_X(t)$ for at least one t — U is too large or too small.

Null distribution: Use formula 10.7 in textbook for small m and n. For $m, n > 10$, U is approximately normal with mean $n(n+m+1)/2$ and variance $nm(n+m+1)/12$.

Also, known as Mann-Whitney-Wilcoxon rank sum test.

Ex: (Exercise 10.22) Fifteen email attachments were classified as benign and malicious. The seven benign attachments had sizes 0.4, 2.1, 3.6, 0.6, 0.8, 2.4, and 4.0 MB. The eight malicious attachments had sizes 1.2, 0.2, 0.3, 3.3, 2.0, 0.9, 1.1, and 1.5 MB. Is there a significant difference in the distribution of sizes of benign and malicious attachments? (If so, the size could help classify email attachments and warn about possible malicious code.)

Here are the 15 observations sorted in increasing order:
0.2, 0.3, 0.4, 0.6, 0.8, 0.9, 1.1, 1.2, 1.5, 2.0, 2.1, 2.4, 3.3, 3.6, 4.0

verify:

R code:

```
> x <- c(0.4, 2.1, 3.6, 0.6, 0.8, 2.4, 4.0)
> y <- c(1.2, 0.2, 0.3, 3.3, 2.0, 0.9, 1.1, 1.5)
>
> wilcox.test(x, y, alternative="two.sided")

Wilcoxon rank sum test

data:  x and y
W = 36, p-value = 0.3969
alternative hypothesis: true location shift is
not equal to 0
```

$\Rightarrow$ Accept H_0

Q: Why does value of the test statistic computed by R differs from ours?

R substracts $\frac{n(n+1)}{2}$, from W, where n = size of X sample.