

# Chapter 12: Nonparametric Inference

DSCC 462

Computational Introduction to Statistics

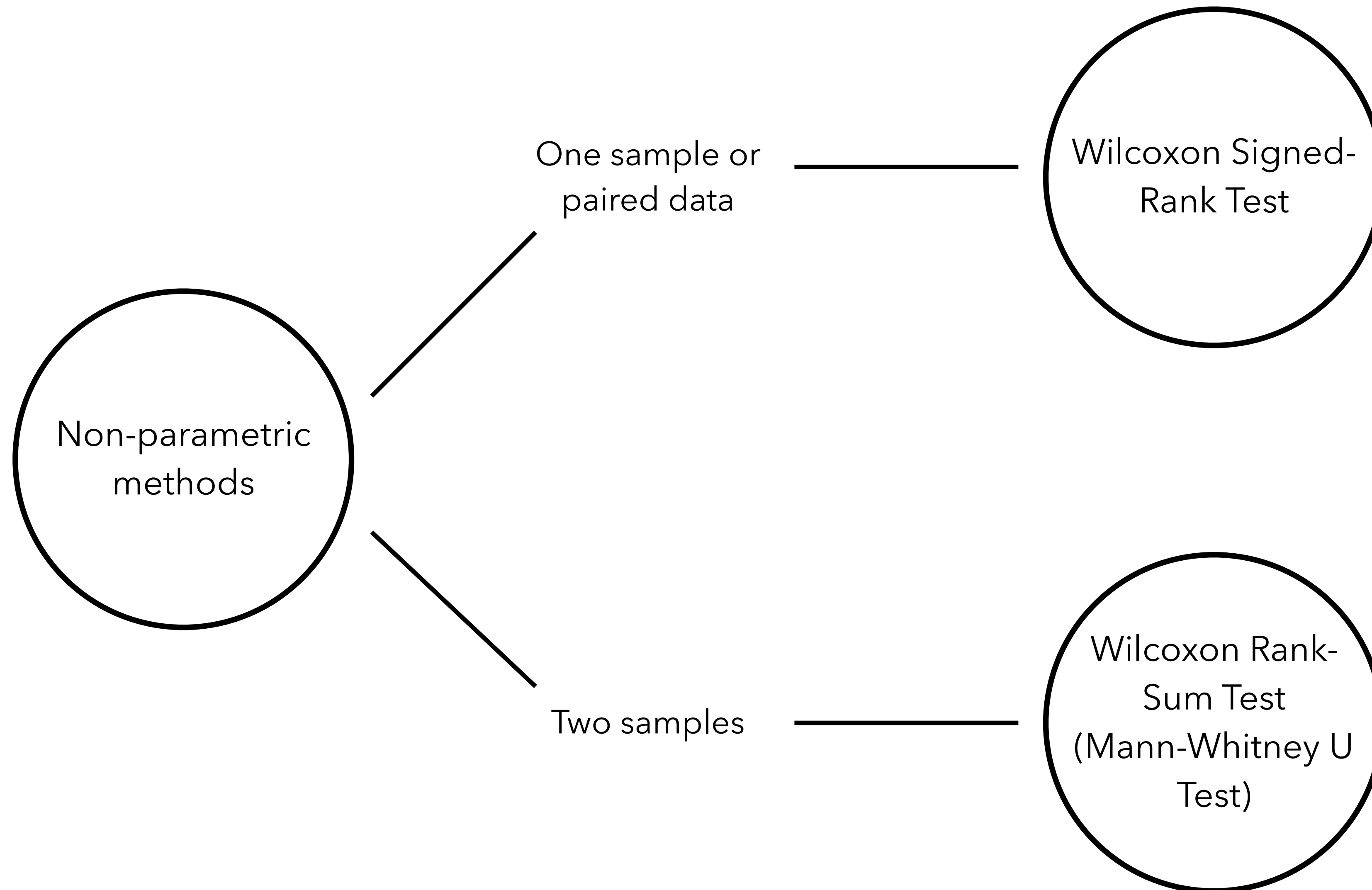
Anson Kahng

Fall 2022

# Plan for Today

- Introduce nonparametric analogues to hypothesis tests
- *Wilcoxon Signed-Rank Test*
  - Nonparametric analog to the one-sample or paired t-test
- *Wilcoxon Rank-Sum Test (Mann-Whitney U Test)*
  - Nonparametric analog to the two-sample t-test

# Plan for Today, Visualized



# Nonparametric Methods

- Think about the statistical tests we've done so far (z-test, t-test,  $\chi^2$ -test, F-test, etc.)
- In all of these, we knew the distribution of the population and we only needed to perform inference on the unknown parameters
  - *Parametric methods*
  - We knew what distribution the population followed
- What if we don't know the distribution of the population?
  - *Nonparametric methods*

# Nonparametric Methods

- When do we use nonparametric methods?
  - When we don't know the underlying population distribution
  - Or the data do not meet the assumptions needed for particular parametric techniques (e.g., CLT doesn't hold, normal approximation for proportions doesn't hold, etc.)
- In this case, we use *nonparametric methods*, which make fewer assumptions regarding the underlying distribution
  - Also known as *distribution-free methods*

# Nonparametric Methods

- Although nonparametric testing procedures make different assumptions, they still follow the same general setup as all hypothesis tests we have discussed so far
  - Make a claim, develop hypotheses, state significance level
  - Calculate a test statistic based on a random sample of data
  - Determine whether to reject or fail to reject the null hypothesis based on the test statistic and significance level

# Motivating Example #1

- Suppose we want to determine whether a new drug changes tumor size
- We cannot assume that tumor sizes are normally distributed
- Let's say that we have a sample of  $n$  pairs of observations (tumor size before drug vs. tumor size after drug), where  $n = 13$
- Can we apply the CLT here?
- This means that we need to use a nonparametric method
  - *Wilcoxon Signed-Rank Test*

# Wilcoxon Signed-Rank Test

- Used to compare two samples from populations that are not independent
  - Nonparametric analog to the paired t-test
- Because we are considering paired data, we may look at the difference in values for each pair of observations
- Does not require populations to be normally distributed
- Takes into account both the magnitudes of the differences and their signs
- Null hypothesis: In the underlying population differences among pairs, the median difference is equal to 0
  - Note that we consider medians for nonparametric tests as opposed to means



# Wilcoxon Signed-Rank Test: Back to Example #1

- Suppose we want to determine whether a new drug changes tumor size
- We cannot assume that tumor sizes are normally distributed
- Let's say that we have a sample of  $n$  pairs of observations (tumor size before drug vs. tumor size after drug), where  $n = 13$
- $H_0$  : The median difference in tumor size equals 0
- $H_1$  : The median difference in tumor size is different from 0
- Test at the  $\alpha = 0.05$  significance level

# Wilcoxon Signed-Rank Test: Steps

- Next, take the difference for each pair of observations
- Ignoring the sign of these observations, rank their absolute values from smallest to largest
  - A difference of 0 is not ranked
  - Remove pair from data set and reduce number of pairs by 1
- Tied observations are assigned an average rank
- Finally separate the ranks by sign to either  $+$  or  $-$

# Wilcoxon Signed-Rank Test: Data Table

Subject	Tumor Size (mm)		Difference	Rank	Signed Rank	
	Before	After			+	-
1	36.3	27.1	9.2			
2	21.7	17.4	4.3			
3	45.1	33.1	12.0			
4	27.8	32.1	-4.3			
5	5.1	8.3	-2.2			
6	23.4	22.1	1.3			
7	25.0	31.2	-6.2			
8	12.6	16.4	-3.8			
9	19.9	12.5	7.4			
10	22.1	22.1	0			
11	18.6	4.8	13.8			
12	8.9	22.6	-13.7			
13	12.7	6.4	6.3			
14	29.3	18.3	9.0			
15	26.4	21.8	4.6			

# Wilcoxon Signed-Rank Test: Data Table

Subject	Tumor Size (mm)		Difference	Rank	Signed Rank	
	Before	After			+	-
1	36.3	27.1	9.2	11		
2	21.7	17.4	4.3	4.5		
3	45.1	33.1	12.0	12		
4	27.8	32.1	-4.3	4.5		
5	5.1	8.3	-2.2	2		
6	23.4	22.1	1.3	1		
7	25.0	31.2	-6.2	7		
8	12.6	16.4	-3.8	3		
9	19.9	12.5	7.4	9		
10	22.1	22.1	0	-		
11	18.6	4.8	13.8	14		
12	8.9	22.6	-13.7	13		
13	12.7	6.4	6.3	8		
14	29.3	18.3	9.0	10		
15	26.4	21.8	4.6	6		

# Wilcoxon Signed-Rank Test: Data Table

Subject	Tumor Size (mm)		Difference	Rank	Signed Rank	
	Before	After			+	-
1	36.3	27.1	9.2	11	11	
2	21.7	17.4	4.3	4.5	4.5	
3	45.1	33.1	12.0	12	12	
4	27.8	32.1	-4.3	4.5		4.5
5	5.1	8.3	-2.2	2		2
6	23.4	22.1	1.3	1	1	
7	25.0	31.2	-6.2	7		7
8	12.6	16.4	-3.8	3		3
9	19.9	12.5	7.4	9	9	
10	22.1	22.1	0	-		
11	18.6	4.8	13.8	14	14	
12	8.9	22.6	-13.7	13		13
13	12.7	6.4	6.3	8	8	
14	29.3	18.3	9.0	10	10	
15	26.4	21.8	4.6	6	6	

# Wilcoxon Signed-Rank Test

- Calculate the sum of the positive ranks,  $T^+$ , and the sum of the negative ranks,  $T^-$
- Calculate  $T = T^+ - T^-$
- Under the null hypothesis, the median of the underlying population differences is equal to 0
- Thus, we expect approximately equal numbers of positive and negative ranks
- Additionally, the sum of the positive ranks should be approximately equal to the sum of the negative ranks, so  $T$  should be approximately 0

# Wilcoxon Signed-Rank Test

- Evaluate the null hypothesis using the test statistic:

$$z_T = \frac{T - \mu_T}{\sigma_T}$$

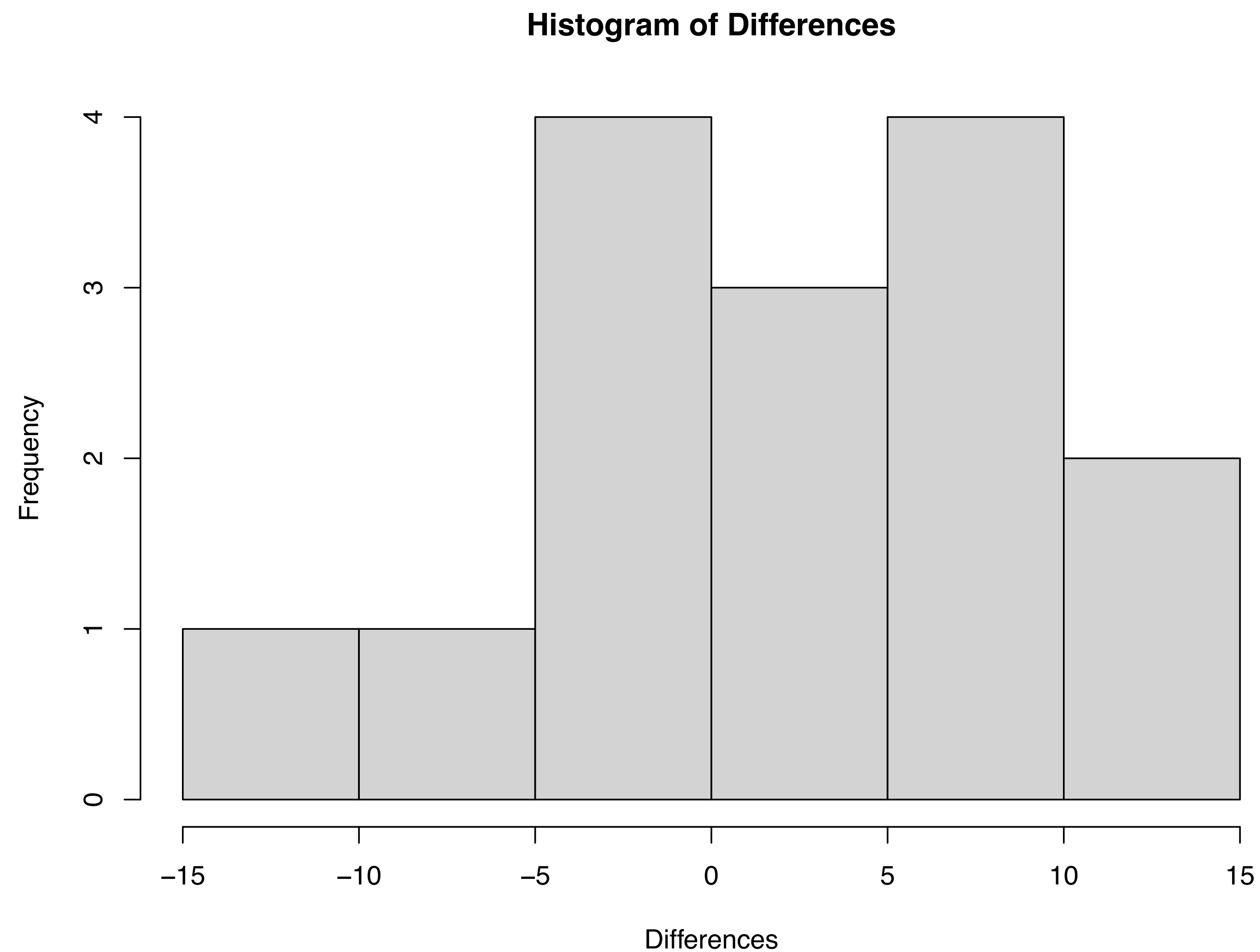
- Note that

$$\mu_T = 0$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{6}}$$

- $Z_T \sim N(0,1)$  given that  $n$  is large enough (typically  $n > 12$ )

# Histogram of Differences





# Wilcoxon Signed-Rank Test

- Based on the histogram, the difference in tumor size does not appear to be normally distributed, so we want to use the Wilcoxon signed-rank test
- $T^+ =$
- $T^- =$
- $T =$
- $n =$

Subject	Signed Rank	
	+	-
1	11	
2	4.5	
3	12	
4		4.5
5		2
6	1	
7		7
8		3
9	9	
10		
11	14	
12		13
13	8	
14	10	
15	6	

# Wilcoxon Signed-Rank Test

- Based on the histogram, the difference in tumor size does not appear to be normally distributed, so we want to use the Wilcoxon signed-rank test
- $T^+ = 11 + 4.5 + 12 + 1 + 9 + 14 + 8 + 10 + 6 = 75.5$
- $T^- = 4.5 + 2 + 7 + 3 + 13 = 29.5$
- $T = 46$
- $n = 14$

Subject	Signed Rank	
	+	-
1	11	
2	4.5	
3	12	
4		4.5
5		2
6	1	
7		7
8		3
9	9	
10		
11	14	
12		13
13	8	
14	10	
15	6	

# Wilcoxon Signed-Rank Test

- Given  $T = 46$ , we then have the following:

$$\mu_T = 0$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{6}} =$$

- Thus,

$$z_T = \frac{T - \mu_T}{\sigma_T} =$$

# Wilcoxon Signed-Rank Test

- Given  $T = 46$ , we then have the following:

$$\mu_T = 0$$

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{6}} = 31.86$$

- Thus,

$$z_T = \frac{T - \mu_T}{\sigma_T} = 1.44$$

# Wilcoxon Signed-Rank Test

- Calculating the p-value, we have
- Conclusion:

# Wilcoxon Signed-Rank Test

- Calculating the p-value, we have
$$p = 2 \cdot \Pr(Z > 1.44) = 2 * (1 - \text{pnorm}(1.44)) = 0.149$$
- Conclusion: Since the p-value of 0.149 is greater than  $\alpha = 0.05$ , we fail to reject the null hypothesis and have insufficient evidence to conclude that the median difference is not equal to 0
- There is not enough evidence to conclude that the new drug significantly changes median tumor size

# Wilcoxon Signed-Rank Test

- If the sample size is  $n \leq 12$ , we cannot use the normal approximation
- In that case, we can use `psignrank(T, n)` in R to calculate the exact p-value
  - `2 * (1 - psignrank(75.5, n=14)) = 0.135`
- R requires  $T = T^+$  for this to work correctly!

# Wilcoxon Signed-Rank Test: R Code

```
> wilcox.test(before, after, paired=T, exact=F,correct=F)
```

Wilcoxon signed rank test

```
data: before and after  
V = 64, p-value = 0.1961  
alternative hypothesis: true location shift is not equal to 0
```

```
> wilcox.test(before, after, paired=T, exact=T,correct=F)
```

Wilcoxon signed rank test

```
data: before and after  
V = 64, p-value = 0.2163  
alternative hypothesis: true location shift is not equal to 0
```



# Motivating Example #2

- Suppose that we want to determine whether having a certain disease is associated with a raised body temperature
- Cannot assume that body temperature is normally distributed
- Take a sample of  $n_1 = 12$  people who do not have the disease
- Take a sample of  $n_2 = 15$  people who do have the disease
- How can we compare the median body temperature for these two populations?
  - *Wilcoxon Rank-Sum Test (Mann-Whitney U Test)*

# Wilcoxon Rank-Sum Test

- Used to compare samples from independent populations
  - Nonparametric analog to the two-sample t-test
- Does not require populations to be normally distributed
- Requires the two populations to have the same general shape
- $H_0$  : The medians of the two populations are identical

# Wilcoxon Rank-Sum Test: Back to Example #2

- Suppose that we want to determine whether having a certain disease is associated with a raised body temperature
- Cannot assume that body temperature is normally distributed
- Take a sample of  $n_1 = 12$  people who do not have the disease
- Take a sample of  $n_2 = 15$  people who do have the disease
- $H_0$  : The median body temperature for those without the disease is greater than or equal to those with the disease
- $H_1$  : The median body temperature for those without the disease is less than those with the disease
- Test at the  $\alpha = 0.05$  significance level

# Wilcoxon Rank-Sum Test: Steps

- Combine all data from the two samples and rank the observations from smallest to largest
- If ranks are tied, we assign the average rank to those values
- We then find the sum of ranks for each of the two original samples, denoted  $W_1$  and  $W_2$ , and then let  $W = \min(W_1, W_2)$
- Under  $H_0$ , the underlying populations have the same median, so we would expect ranks to be randomly distributed between the two groups
- Thus, the average ranks for the two samples (i.e.,  $W_1/n_1$  and  $W_2/n_2$ ) should be approximately equal

# Data Table

No Disease		Disease	
Temp	Rank	Temp	Rank
98.1		99.3	
98.5		99.4	
98.6		99.4	
98.8		99.5	
98.9		99.5	
99.0		99.6	
99.2		99.7	
99.5		99.7	
99.6		100.0	
99.7		100.0	
100.5		100.1	
101.0		100.1	
		100.1	
		101.1	
		101.9	

# Data Table

No Disease		Disease	
Temp	Rank	Temp	Rank
98.1	1	99.3	8
98.5	2	99.4	9.5
98.6	3	99.4	9.5
98.8	4	99.5	12
98.9	5	99.5	12
99.0	6	99.6	14.5
99.2	7	99.7	17
99.5	12	99.7	17
99.6	14.5	100.0	19.5
99.7	17	100.0	19.5
100.5	24	100.1	22
101.0	25	100.1	22
		100.1	22
		101.1	26
		101.9	27

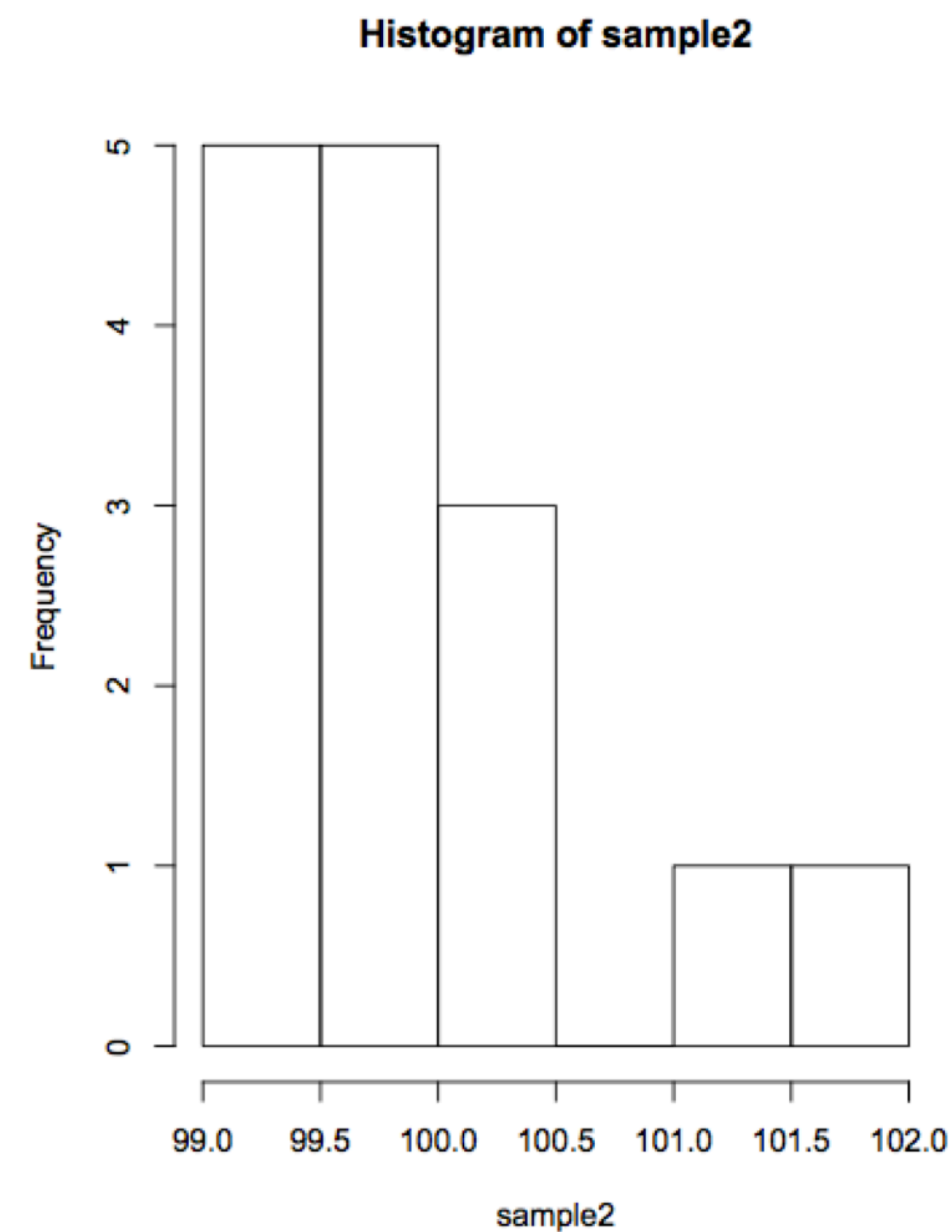
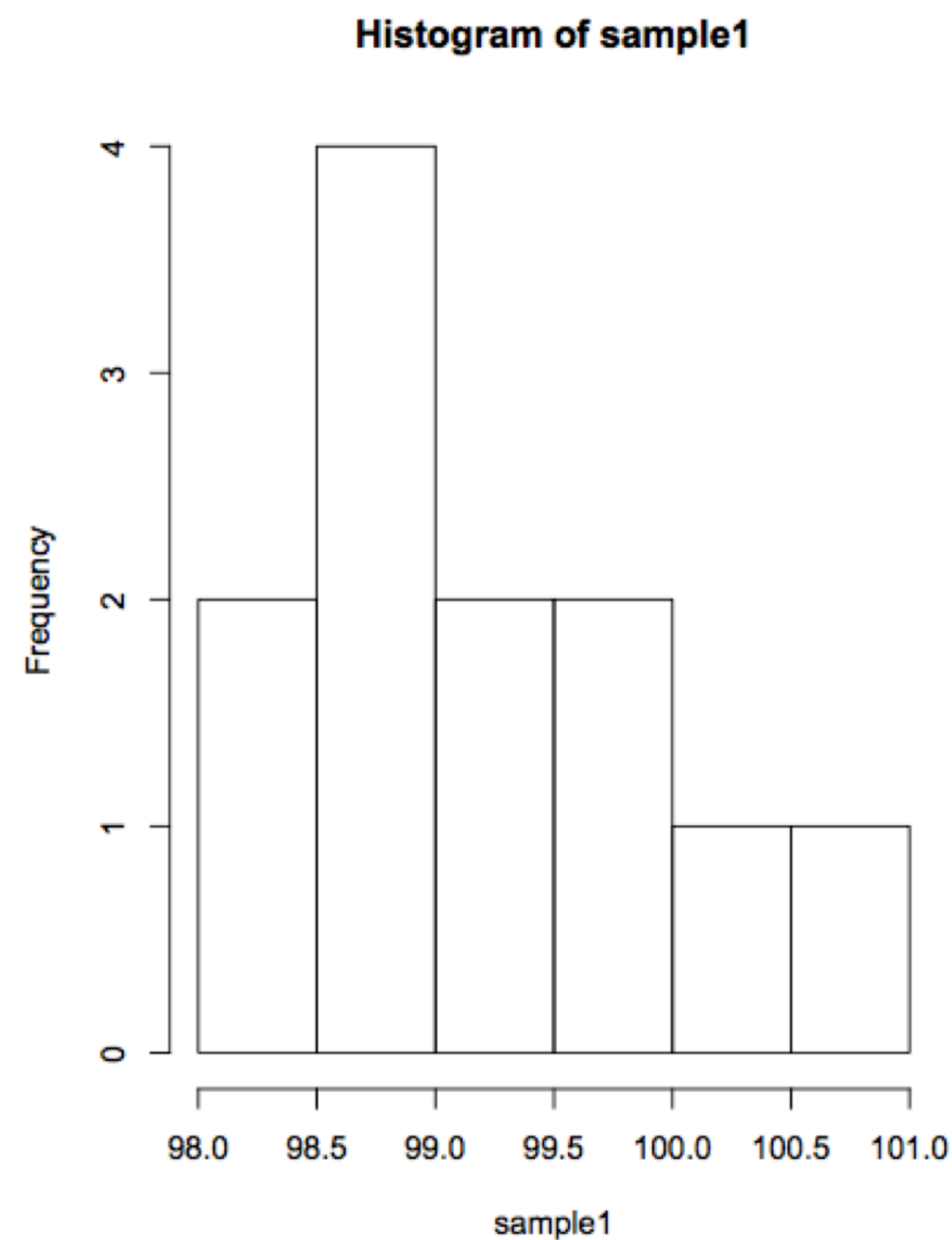
# Wilcoxon Rank-Sum Test: Steps

- Evaluate the null hypothesis using the test statistic  $z_W = \frac{W - \mu_W}{\sigma_W}$
- Let  $n_1$  be the number of observations in the sample with the smaller sum of ranks
- Let  $n_2$  be the number of observations in the sample with the larger sum of ranks
- Then,

- $\mu_W = \frac{n_1(n_1 + n_2 + 1)}{2}$  and  $\sigma_W = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$

- $z_W \sim N(0,1)$  when  $n_1$  and  $n_2$  are large enough ( $n_1, n_2 > 10$ )

# Histograms of Samples





# Wilcoxon Rank-Sum Test

- Based on the histograms, the two samples do not appear to be coming from normally distributed populations, so we want to use the Wilcoxon rank sum test
- Wilcoxon rank sum test is appropriate since both populations have similar shapes
- The sum of ranks for sample 1 is 120.5
- The sum of ranks for sample 2 is 257.5
- Thus,  $W = 120.5$ ,  $n_1 = 12$ , and  $n_2 = 15$

# Wilcoxon Rank-Sum Test

- Given  $W = 120.5$ ,  $n_1 = 12$ , and  $n_2 = 15$ , we then have

- $\mu_W = \frac{n_1(n_1 + n_2 + 1)}{2} =$

- $\sigma_W = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} =$

- Thus, we have  $z_W = \frac{W - \mu_W}{\sigma_W} =$

# Wilcoxon Rank-Sum Test

- Given  $W = 120.5$ ,  $n_1 = 12$ , and  $n_2 = 15$ , we then have

- $\mu_W = \frac{n_1(n_1 + n_2 + 1)}{2} = 168$

- $\sigma_W = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = 20.49$

- Thus, we have  $z_W = \frac{W - \mu_W}{\sigma_W} = -2.318$

# Wilcoxon Rank-Sum Test

- Calculating the p-value:
- Conclusion:

# Wilcoxon Rank-Sum Test

- Calculating the p-value:  
 $\Pr(Z < -2.318) = \text{pnorm}(-2.318) = 0.010$
- Conclusion: Since the p-value of 0.01 is less than  $\alpha = 0.05$ , we reject the null hypothesis and conclude that the median body temperature for those with the disease is greater than the median body temperature for those without the disease

# Wilcoxon Rank-Sum Test

- If  $n_1$  and  $n_2$  are very small (i.e., either is less than or equal to 10), we cannot use the normal approximation
- When sample sizes are small, we can use the exact distribution to calculate p-values
- In R, we use `pwilcox( $W_{\text{obs}}$ ,  $n_1$ ,  $n_2$ )`
  - In this case,  $W_{\text{obs}} = W - \frac{n_1(n_1 + 1)}{2}$
  - `pwilcox(120.5-78, 12, 15) = 0.0093`

# Wilcoxon Rank-Sum Test: R Code

```
> wilcox.test(sample1,sample2, exact=F, correct=F, alt="less")
```

```
Wilcoxon rank sum test
```

```
data: sample1 and sample2
```

```
W = 42.5, p-value = 0.01009
```

```
alternative hypothesis: true location shift is less than 0
```

# Nonparametric Methods: Pros and Cons

- Advantages:
  - Do not impose restrictive assumptions
  - Do not require normally distributed populations
  - Are sometimes easier to compute by hand
  - Ranks are less sensitive to measurement error
  - Permits the use of ordinal data
- Disadvantages:
  - If a parametric test can be used, it is more powerful than its nonparametric counterpart
  - Hypotheses tend to be less specific for nonparametric tests
  - Variances are typically overestimated



# Nonparametric Methods: Summary

- Sometimes we want to run tests on variables but do not know their distributions
- Nonparametric tests are a flexible but sometimes underpowered way of doing so
- Nonparametric analog to the one-sample or paired t-test: Wilcoxon Signed-Rank Test
- Nonparametric analog to the two-sample t-test: Wilcoxon Rank-Sum Test (Mann-Whitney U Test)

# Review Sessions Next Week

- TA review sessions (Wegmans 1201):
  - Tuesday, November 1, from 2 - 3 pm (Lucinda's normal OH)
  - Wednesday, November 2, from 6 - 8 pm
- Instructor: Tuesday and Thursday during class
  - Submit specific requests via Google Form