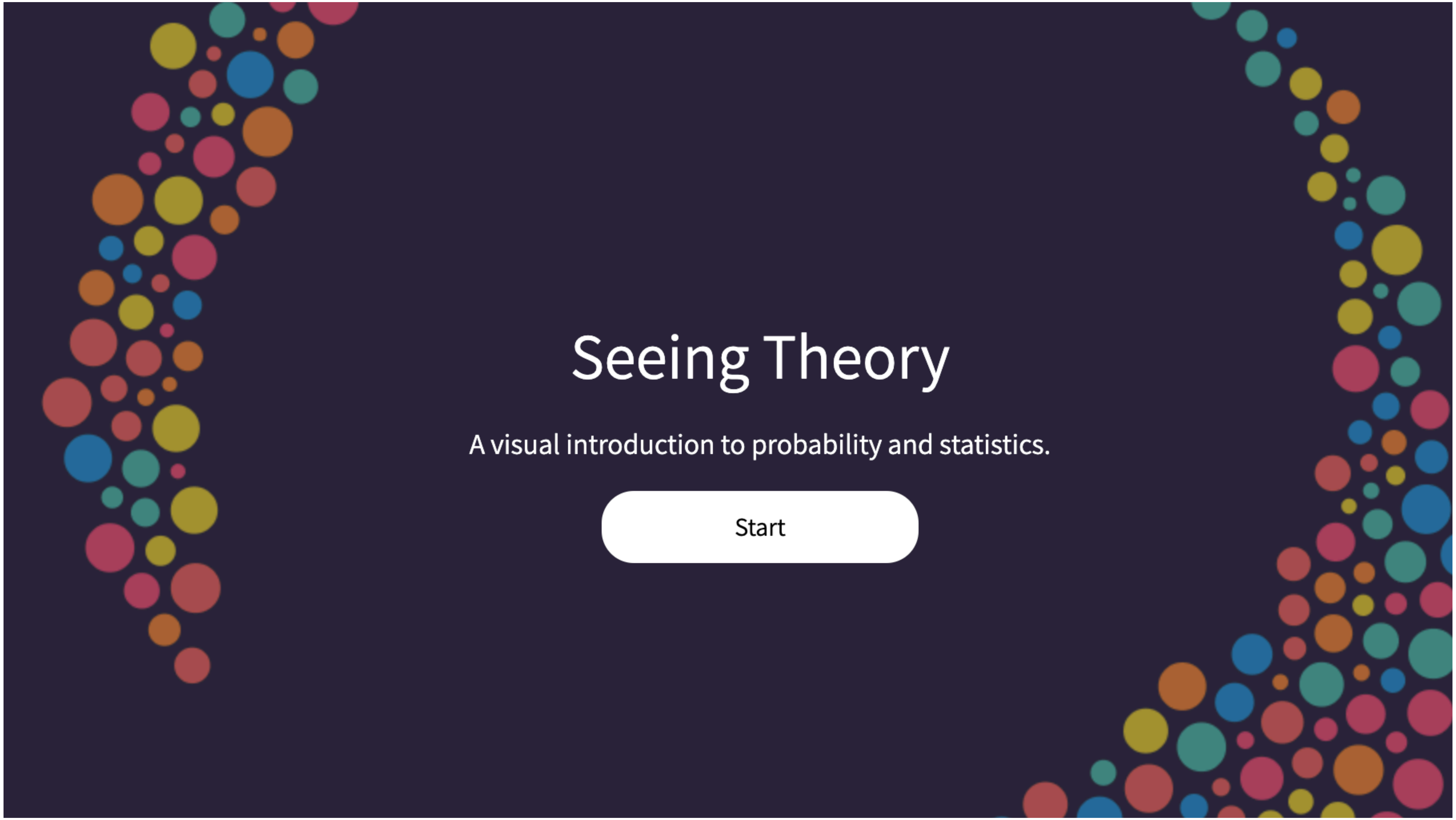
Lab IX: Modern Statistical Methods

Bootstrapping, permutation testing, and cross-validation

Aaron Lanz

Shoutout to Lyndon Duong



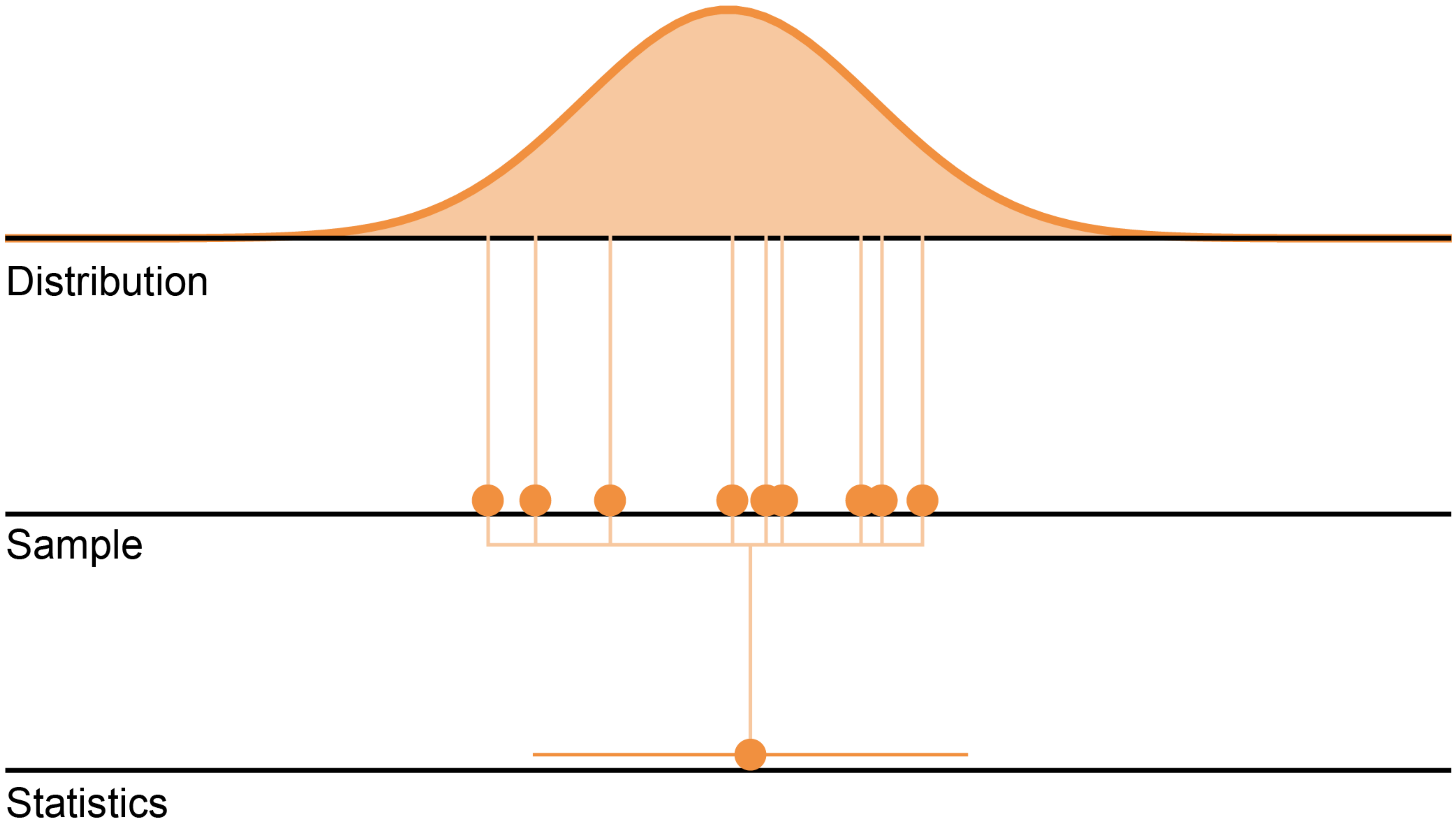
https://seeing-theory.brown.edu/index.html

Basic Probability	Compound Probability	Probability Distributions	Frequentist Inference	Bayesian Inference	Regression Analysis
Chance Events	Set Theory	Random Variables	Point Estimation	Baye’s Theorem	Ordinary Least Square
Expectation	Counting	Discrete and Continuous	Interval Estimation	Likelihood	Correlation
Variance	Conditional Probability	Central Limit Theorem	The Bootstrap	Prior to Posterior	Analysis of Variance

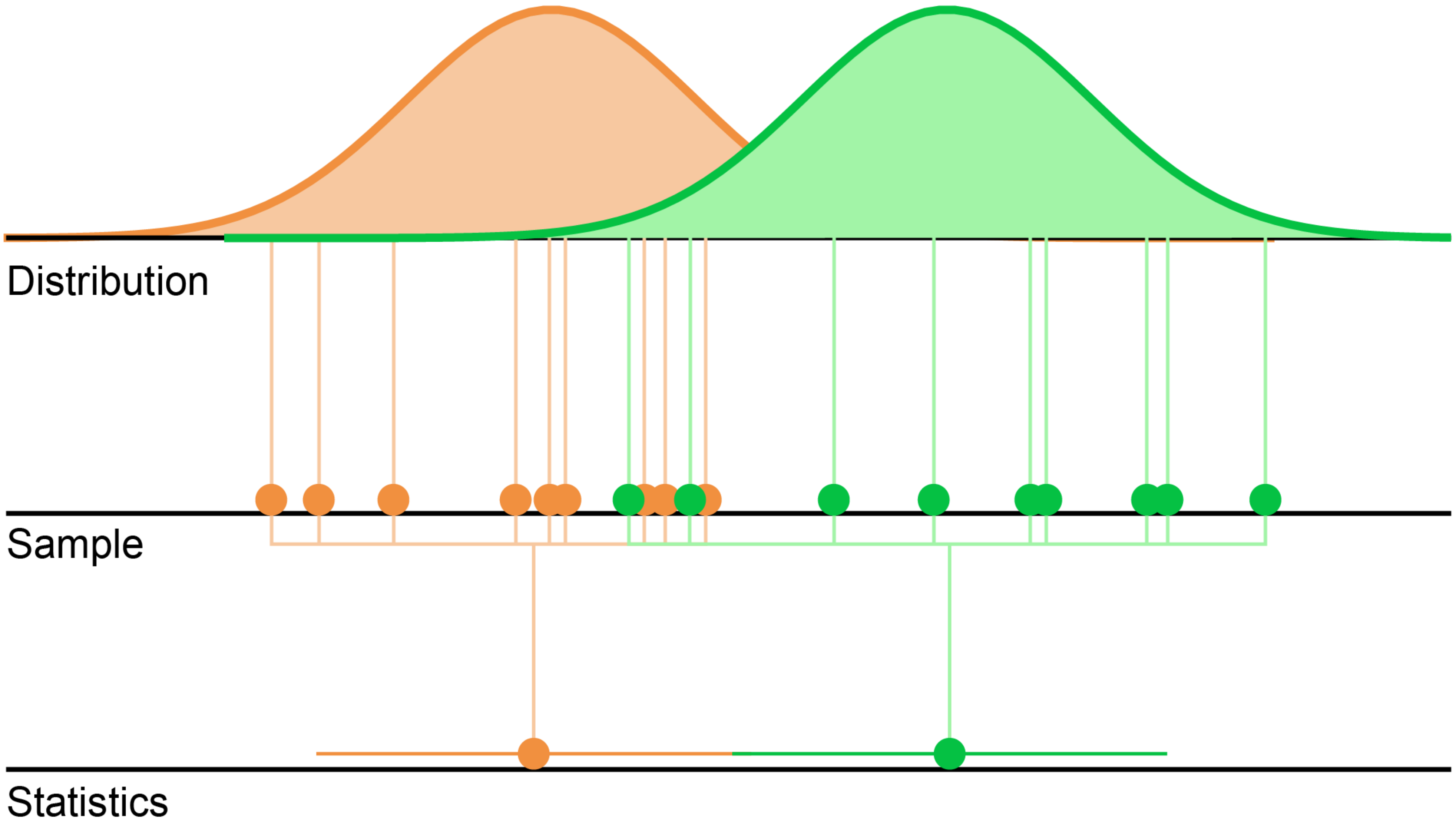
https://seeing-theory.brown.edu/index.html

1) What does the distribution of my statistic look like?

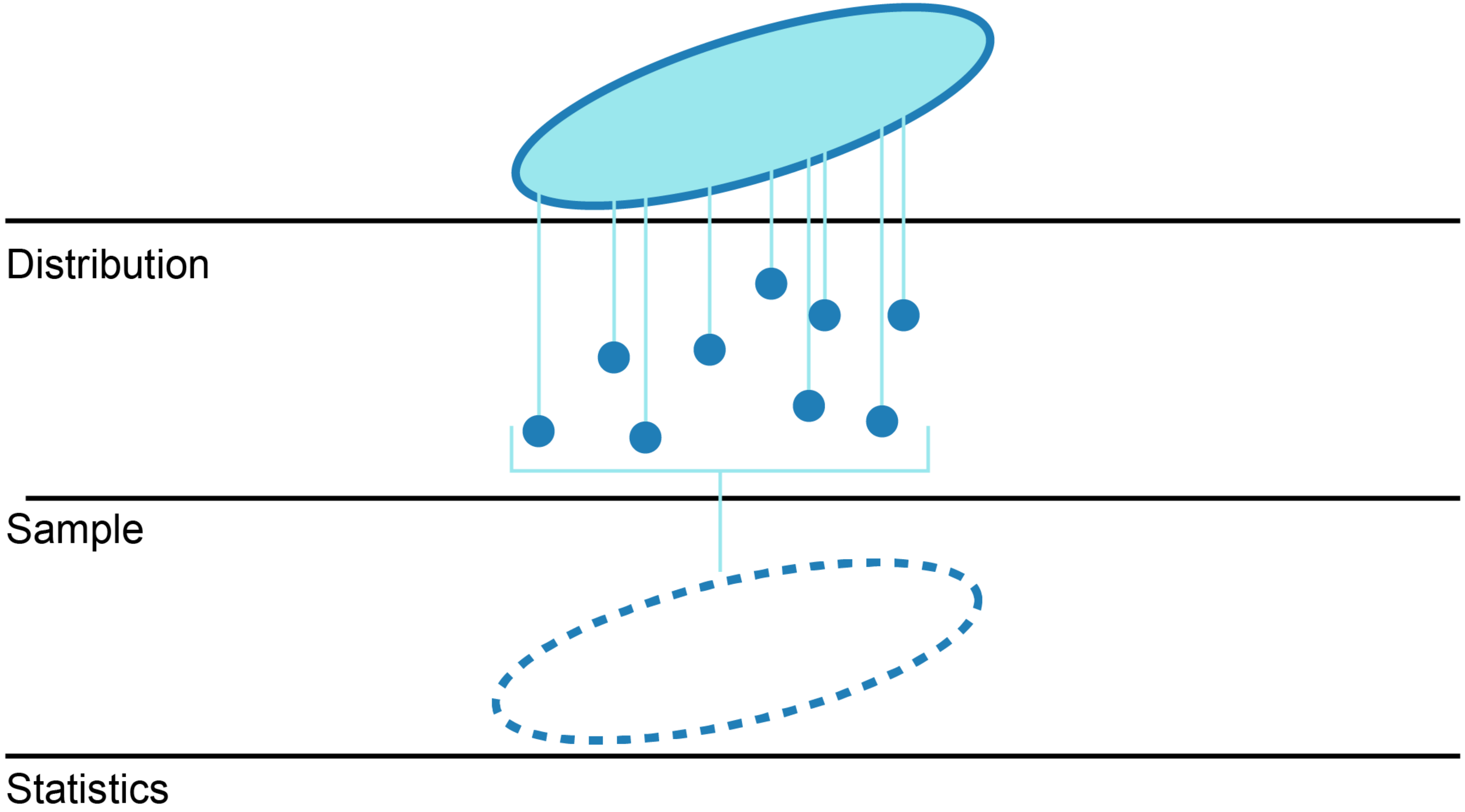
e.g. Does it include 0?



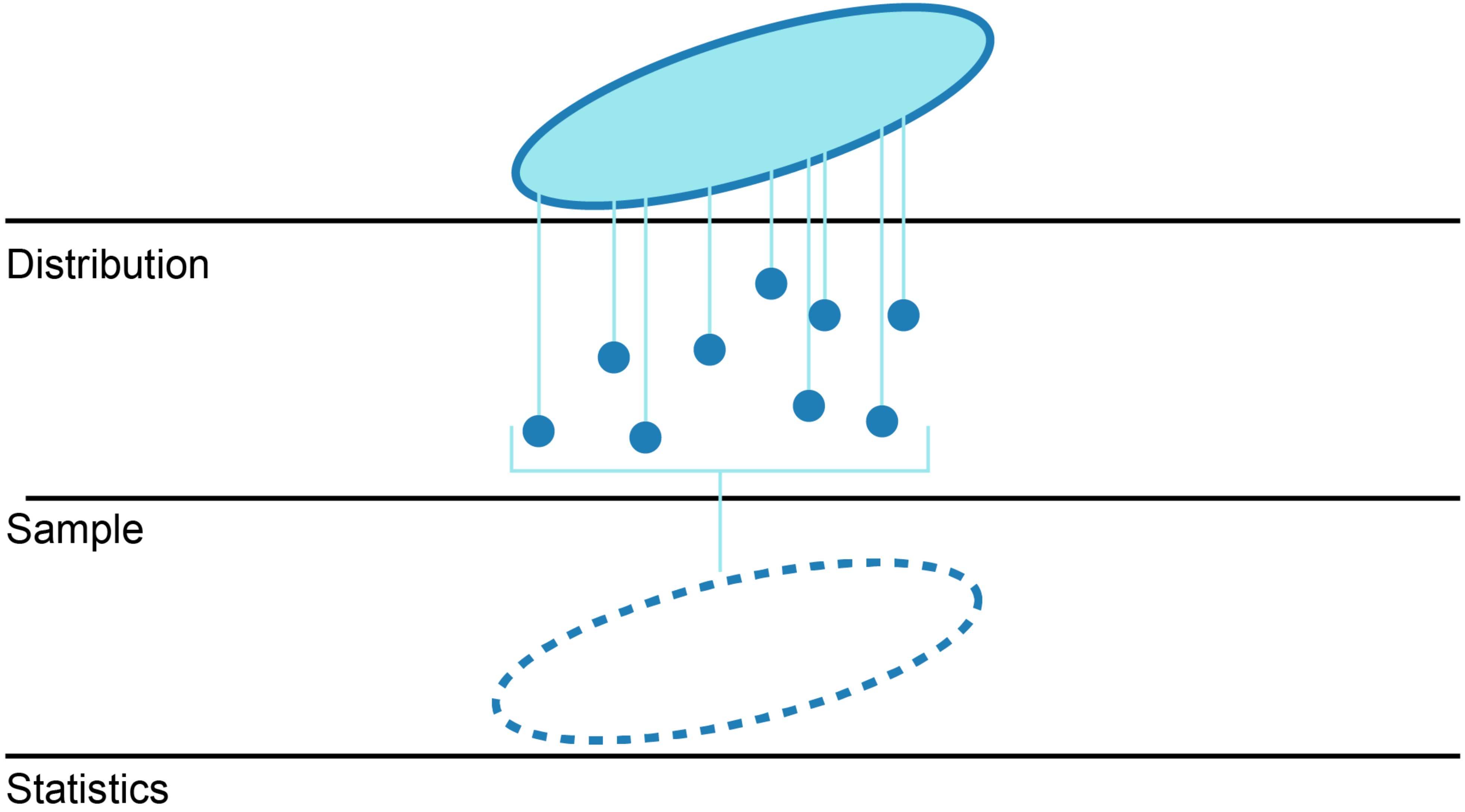
2) Are the groups different?



3) Does a pattern exist in the data?



4) Is my model a good fit?



Outline

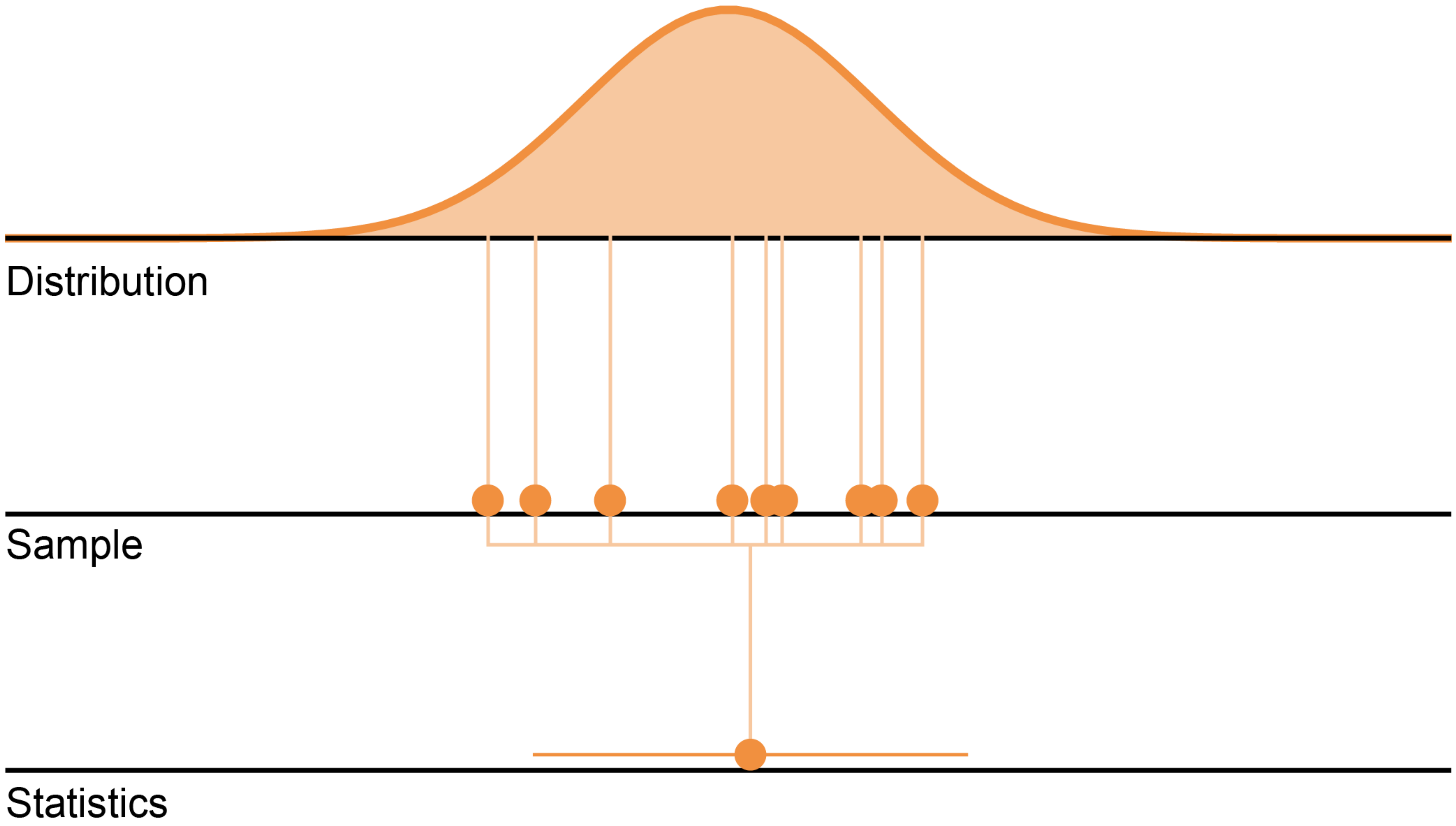
1. What does the distribution of my statistic look like? - **Bootstrapping**

2. Does an effect or pattern exist? - **Permutation testing**

3. Is my model a good fit? – **Cross Validation**

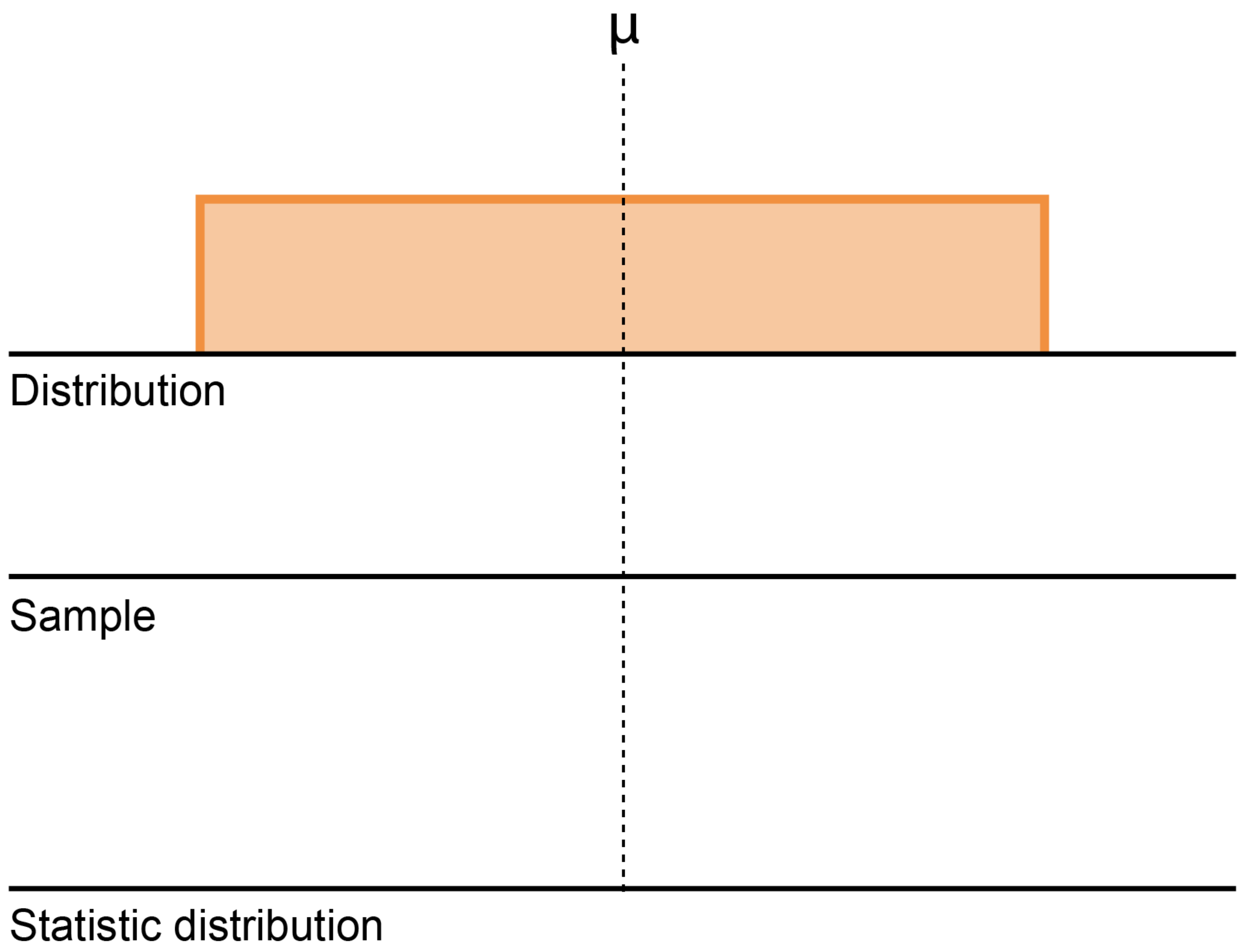
1) What does the distribution of my statistic look like?

e.g. Does it include 0?



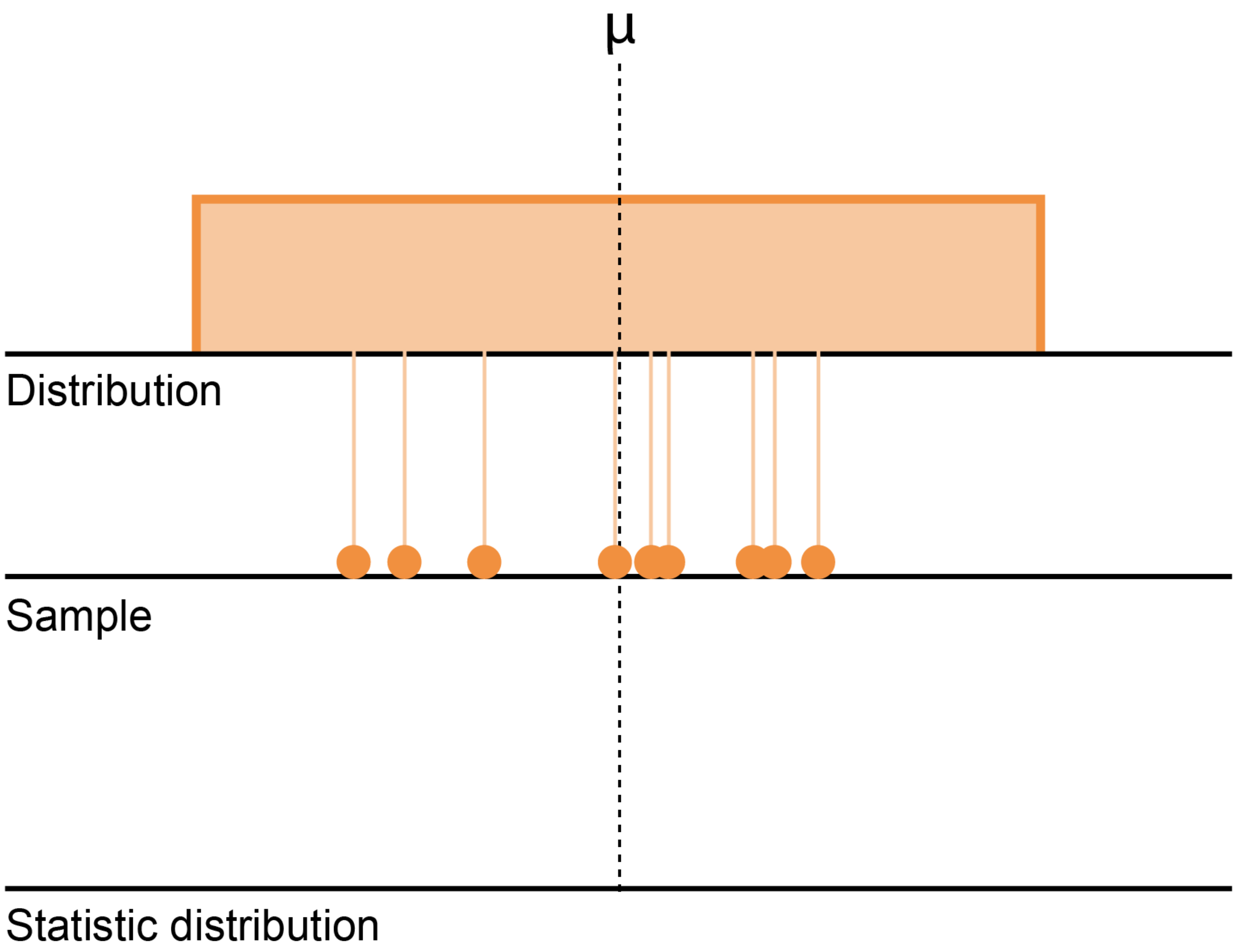
1) What does the distribution of my statistic look like?

What we typically learn



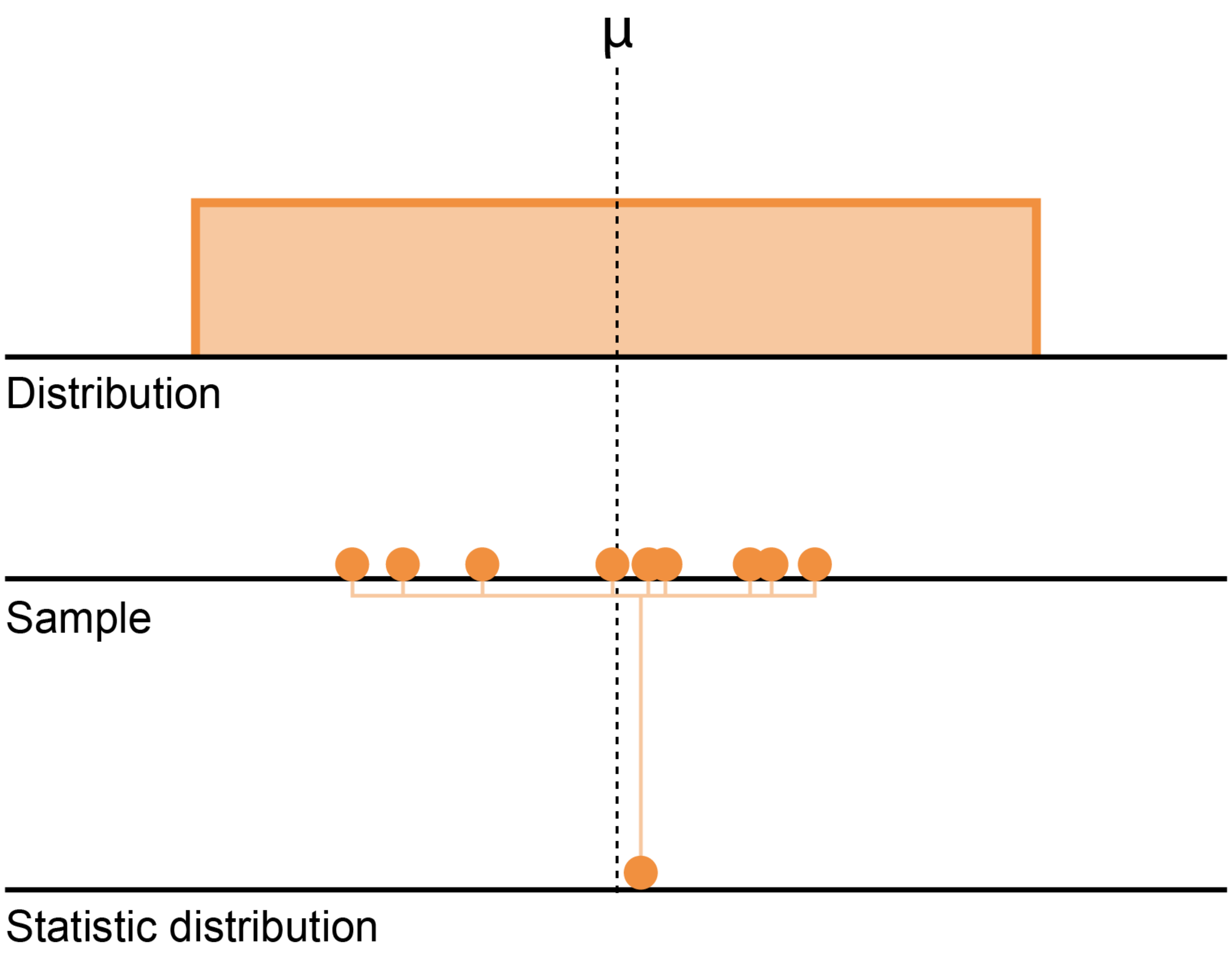
1) What does the distribution of my statistic look like?

What we typically learn



1) What does the distribution of my statistic look like?

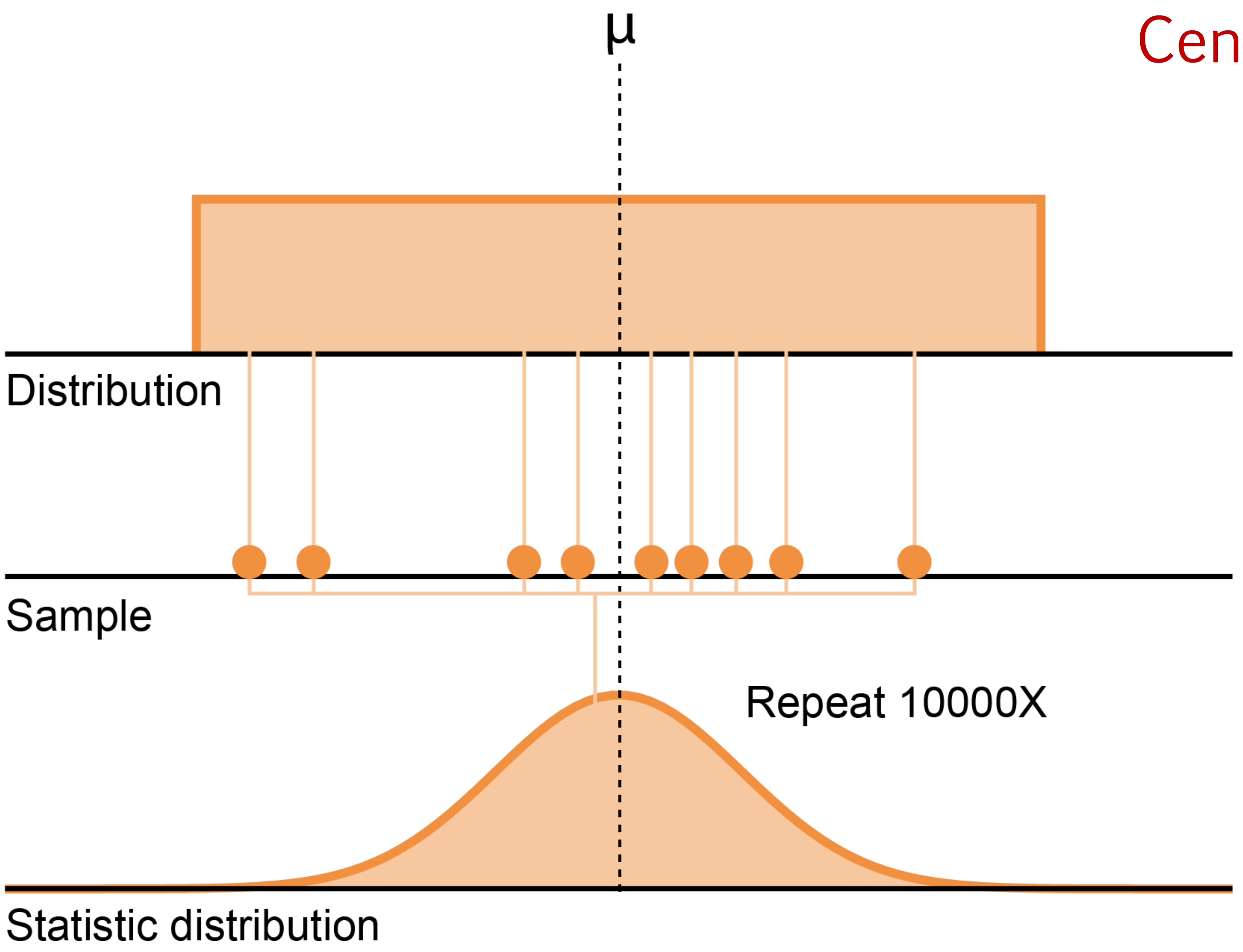
What we typically learn



1) What does the distribution of my statistic look like?

What we typically learn

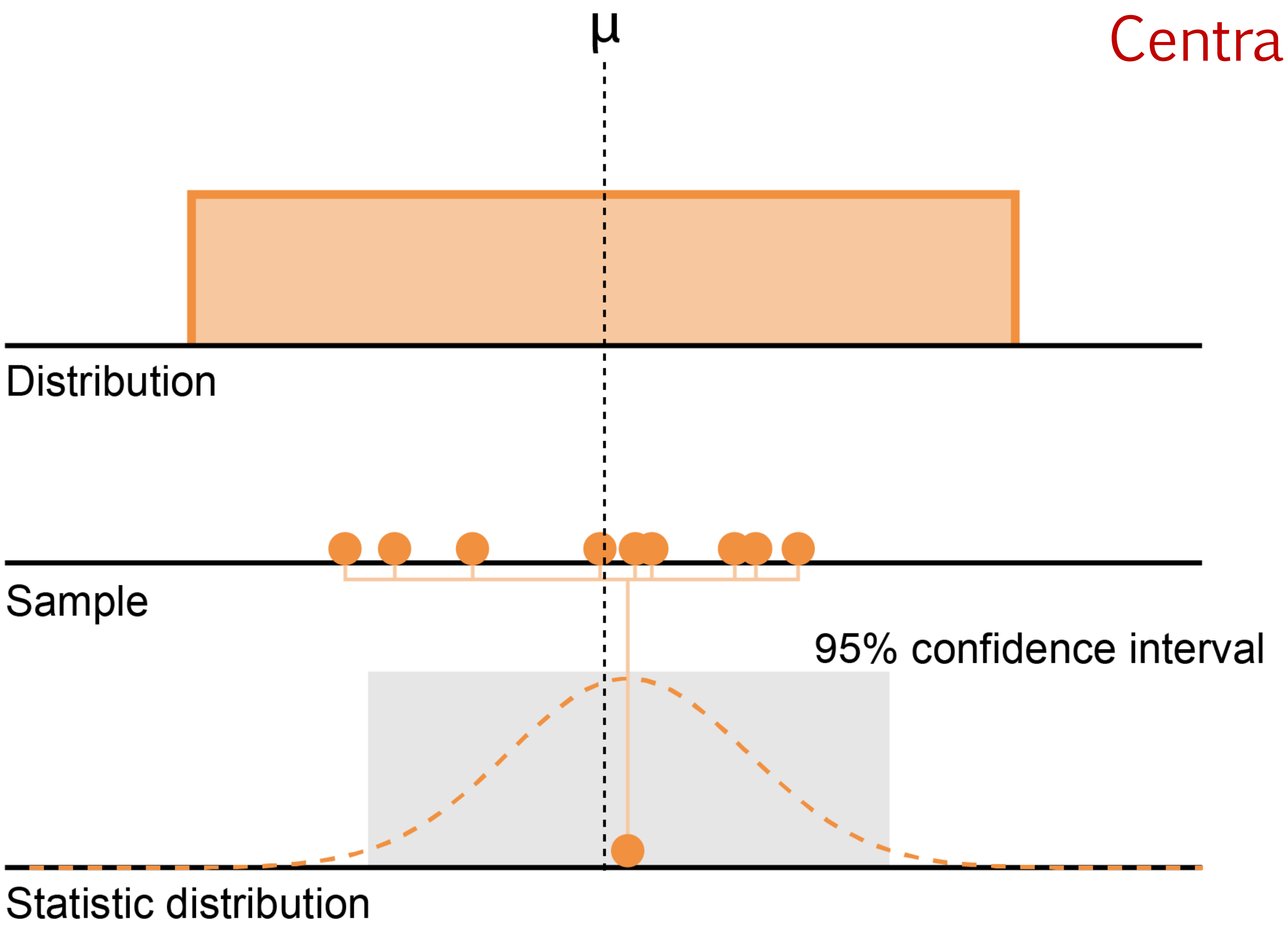
Central Limit Theorem



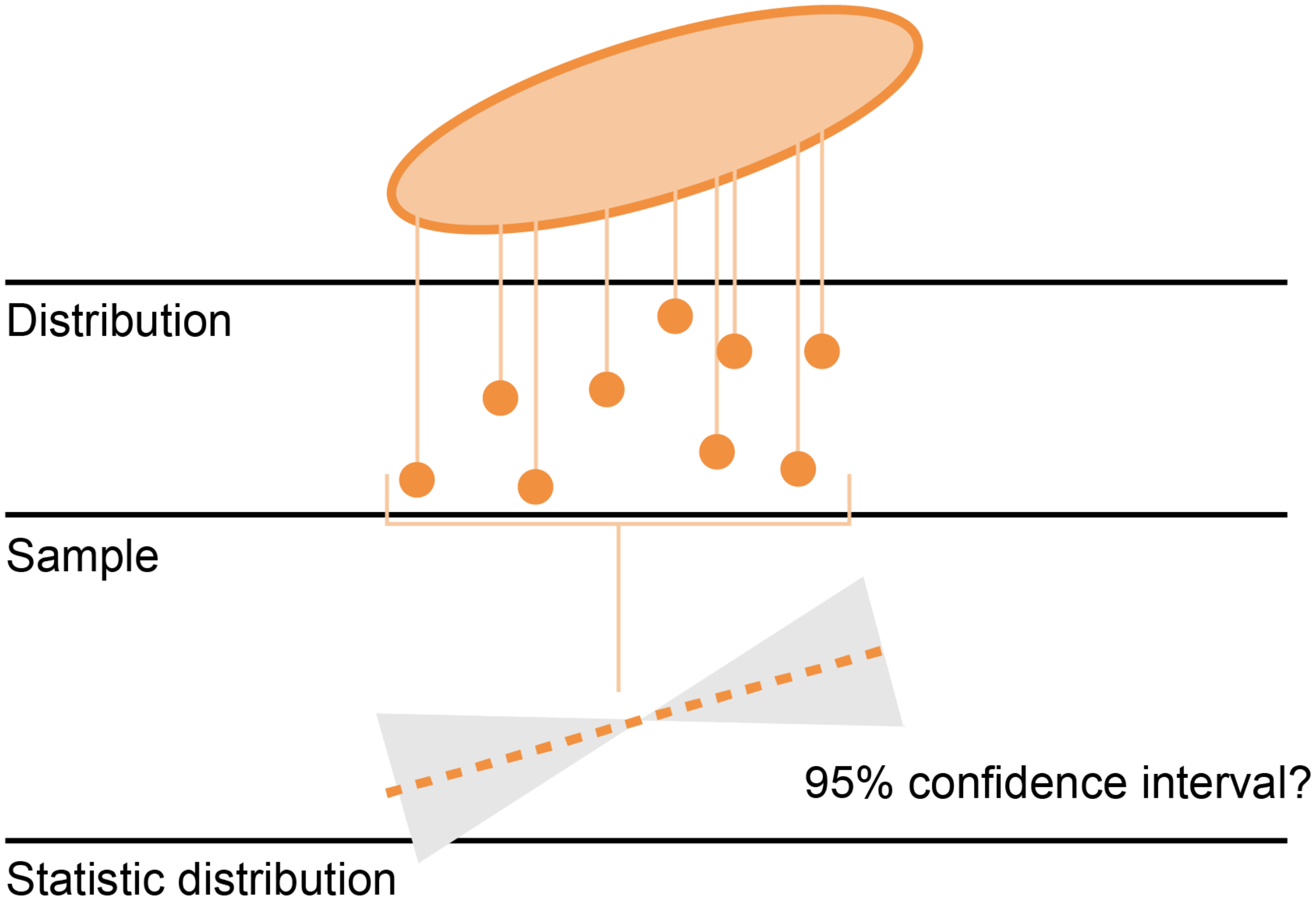
1) What does the distribution of my statistic look like?

What we typically learn

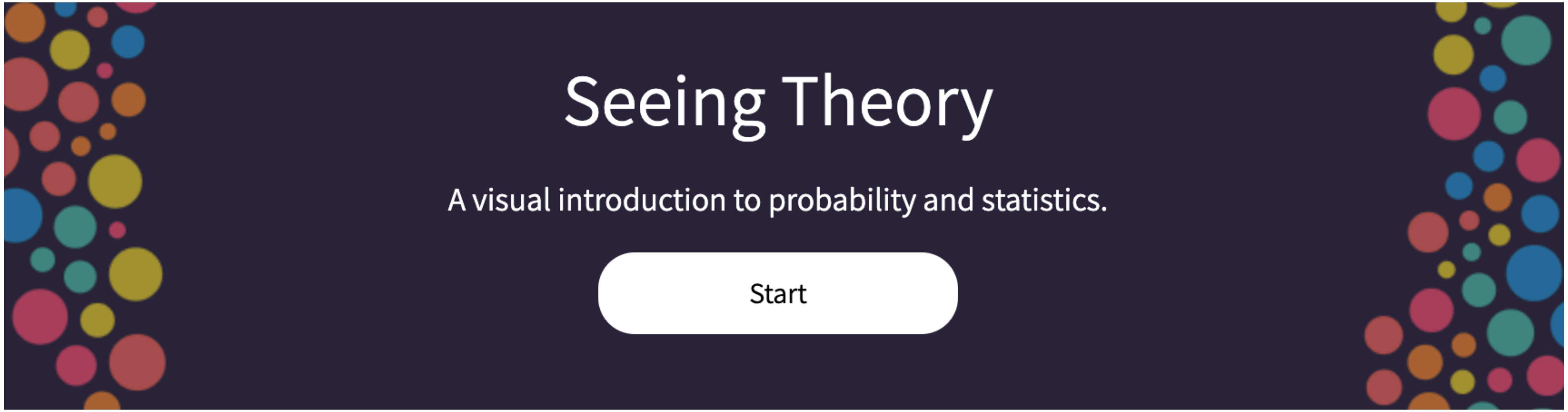
Central Limit Theorem



1) What does the distribution of my statistic look like?



Bootstrapping



https://seeing-theory.brown.edu/frequentist-inference/index.html#section3

Bootstrapping

	(fatal plus non-fatal) heart attacks	subjects
aspirin group:	104	11037
placebo group:	189	11034

Is this effect significant?

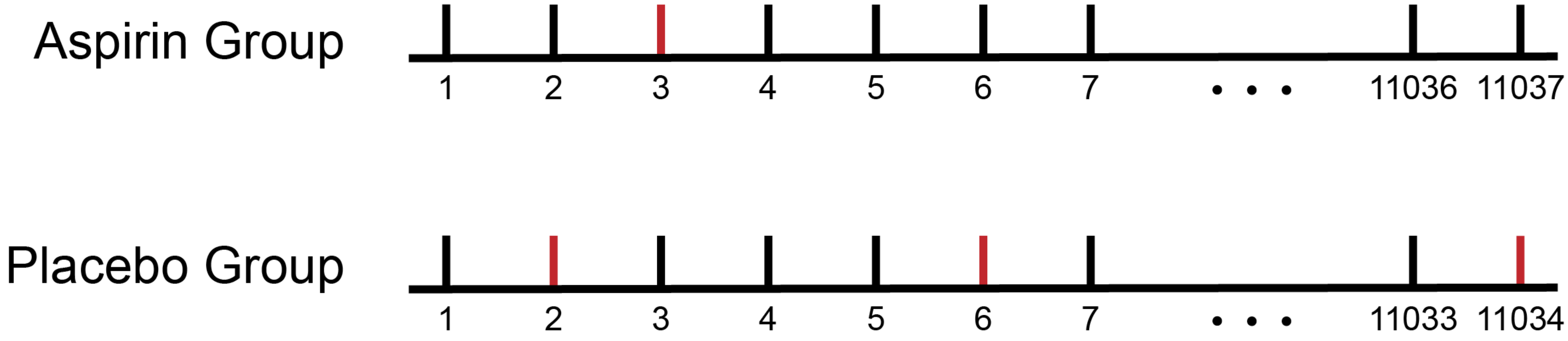
Bootstrapping

	(fatal plus non-fatal) heart attacks	subjects
aspirin group:	104	11037
placebo group:	189	11034

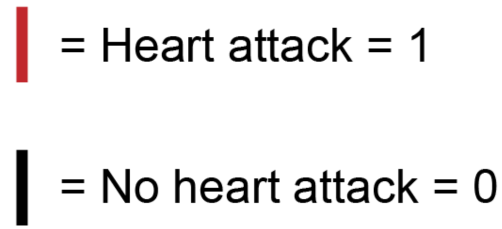
$$Ratio = \frac{H_{Aspirin}}{Tot_{Aspirin}} \Big/ \frac{H_{Placebo}}{Tot_{Placebo}} = \frac{\frac{104}{11037}}{\frac{189}{11034}} = 0.5501$$

Aspirin group had 55% the number of heart attacks compared to the placebo group

How sure are we of this estimate? – Does the 95% confidence interval include 1?



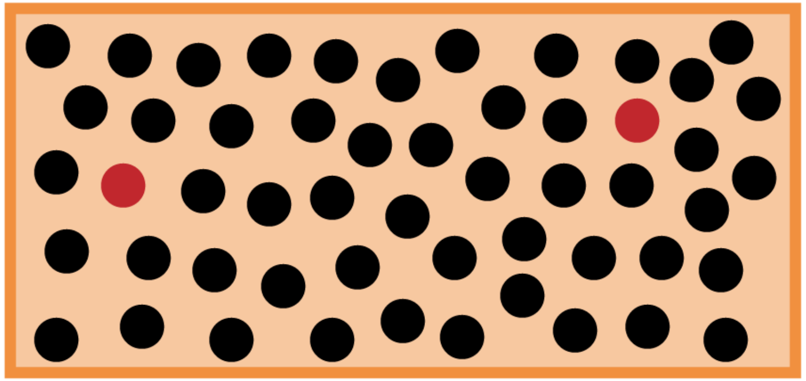
	(fatal plus non-fatal) heart attacks	subjects
aspirin group:	104	11037
placebo group:	189	11034



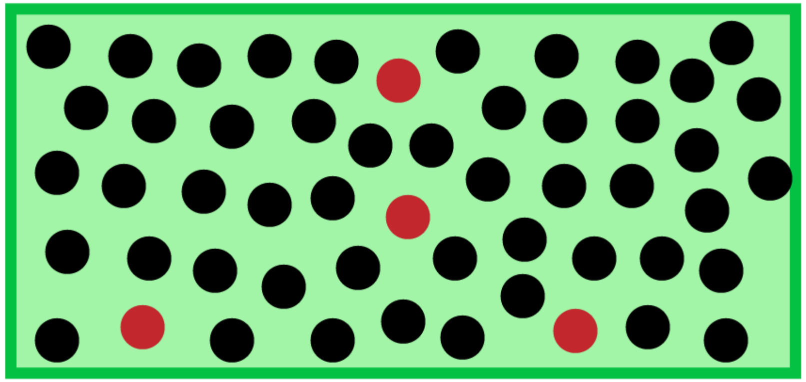
```
aspirin_heart = 104;
aspirin_total = 11037;

placebo_heart = 189;
placebo_total = 11034;

% like flipping a coin…
asp_data = zeros(aspirin_total,1); % 11037 total subjects, all zero
asp_data(1:aspirin_heart) = 1; % set 104 subjects to 1 (1 = heart attack)

placebo_data = zeros(placebo_total, 1); % 11034 tot subjects, all zero
placebo_data(1:placebo_heart) = 1; % 189 had heart attacks
```

From Lyndon Duong

Aspirin Group



Placebo Group



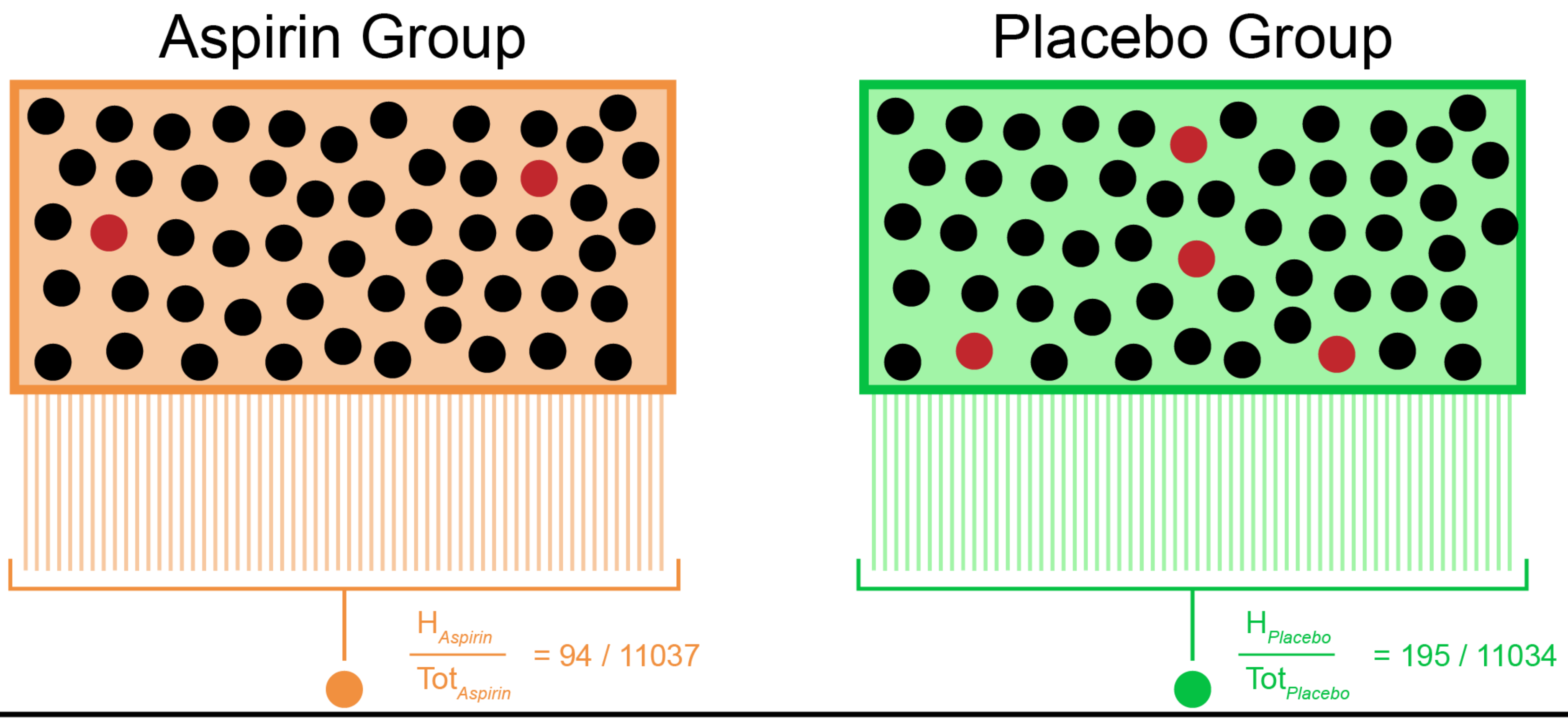
Distribution

(Re-)Sample

(with replacement!)

Statistic Distribution

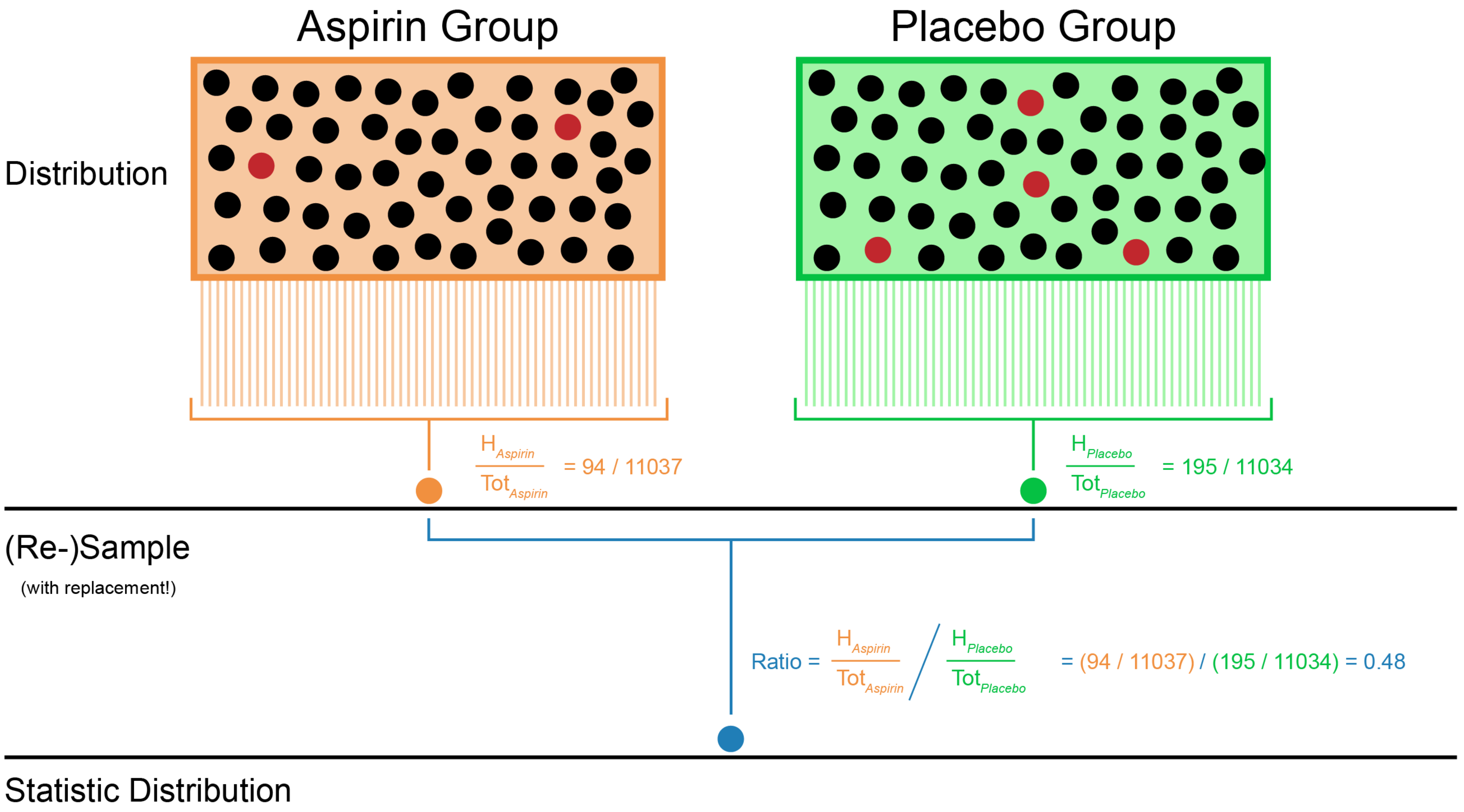
Distribution



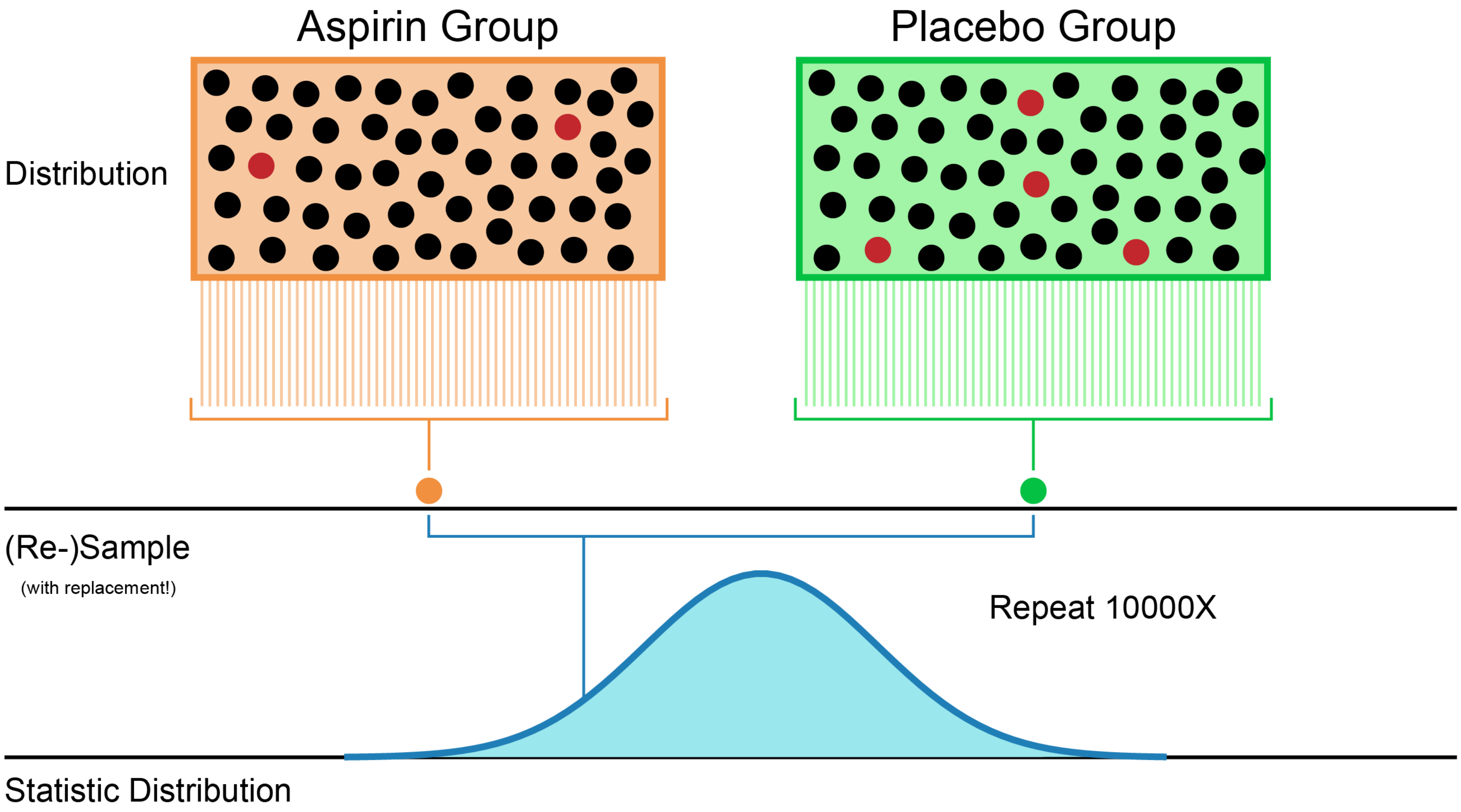
(Re-)Sample

(with replacement!)

Statistic Distribution



Aspirin Group
Placebo Group
Distribution
$\frac{H_{Aspirin}}{Tot_{Aspirin}}$ = 94 / 11037
$\frac{H_{Placebo}}{Tot_{Placebo}}$ = 195 / 11034
(Re-)Sample
(with replacement!)
Ratio = $\frac{H_{Aspirin}}{Tot_{Aspirin}} \Big/ \frac{H_{Placebo}}{Tot_{Placebo}}$ = (94 / 11037) / (195 / 11034) = 0.48
Statistic Distribution

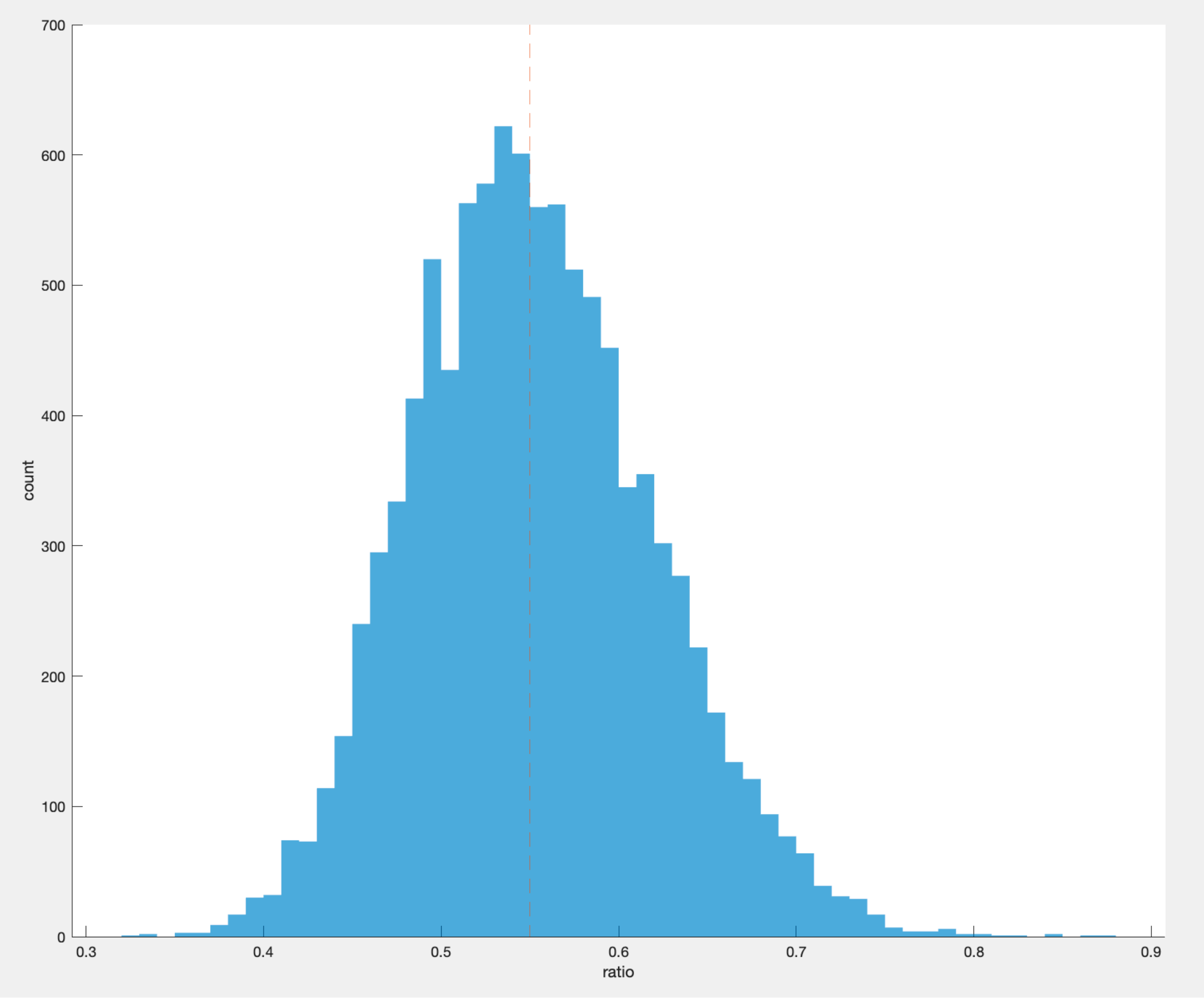

Aspirin Group
Placebo Group
Distribution
(Re-)Sample
(with replacement!)
Repeat 10000X
Statistic Distribution

```matlab
aspirin_heart = 104;
aspirin_total = 11037;

placebo_heart = 189;
placebo_total = 11034;

% like flipping a coin…
asp_data = zeros(aspirin_total,1); % 11037 total subjects, all zero
asp_data(1:aspirin_heart) = 1; % set 104 subjects to 1 (1 = heart attack)

placebo_data = zeros(placebo_total, 1); % 11034 tot subjects, all zero
placebo_data(1:placebo_heart) = 1; % 189 had heart attacks
```

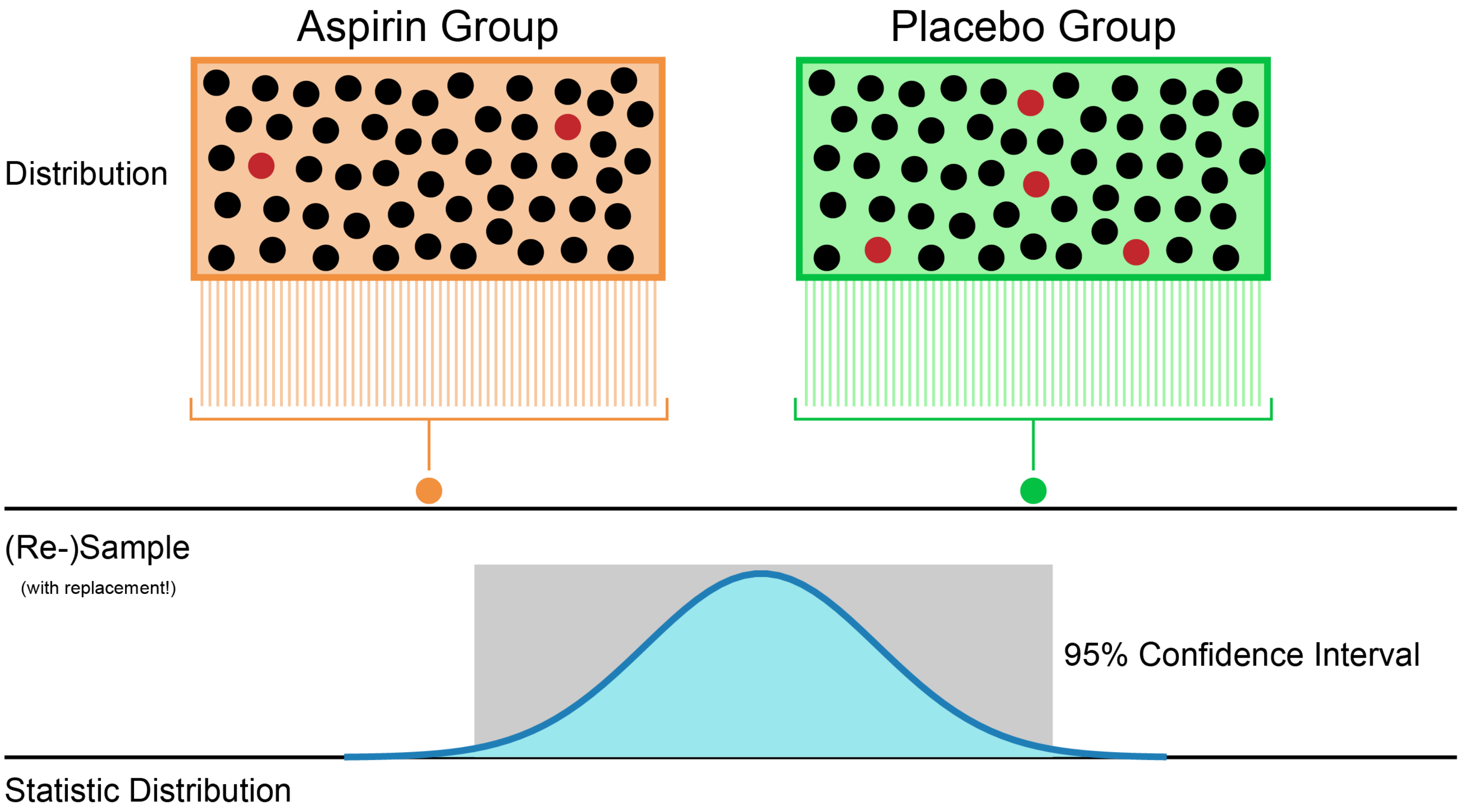
```matlab
ratio_empirical = (aspirin_heart/aspirin_total)/(placebo_heart/placebo_total);
n_boot = 10000;
ratio_boot = zeros(n_boot, 1);

for boot = 1:n_boot
    boot_asp = randsample(asp_data, aspirin_total, 'true');
    boot_placebo = randsample(placebo_data, placebo_total, 'true');
    n_boot_asp = sum(boot_asp);
    n_boot_placebo = sum(boot_placebo);
    ratio_boot(boot) = (n_boot_asp/aspirin_total)/(n_boot_placebo/placebo_total);
end
```

From Lyndon Duong

N = 10,000 bootstrapped ratios

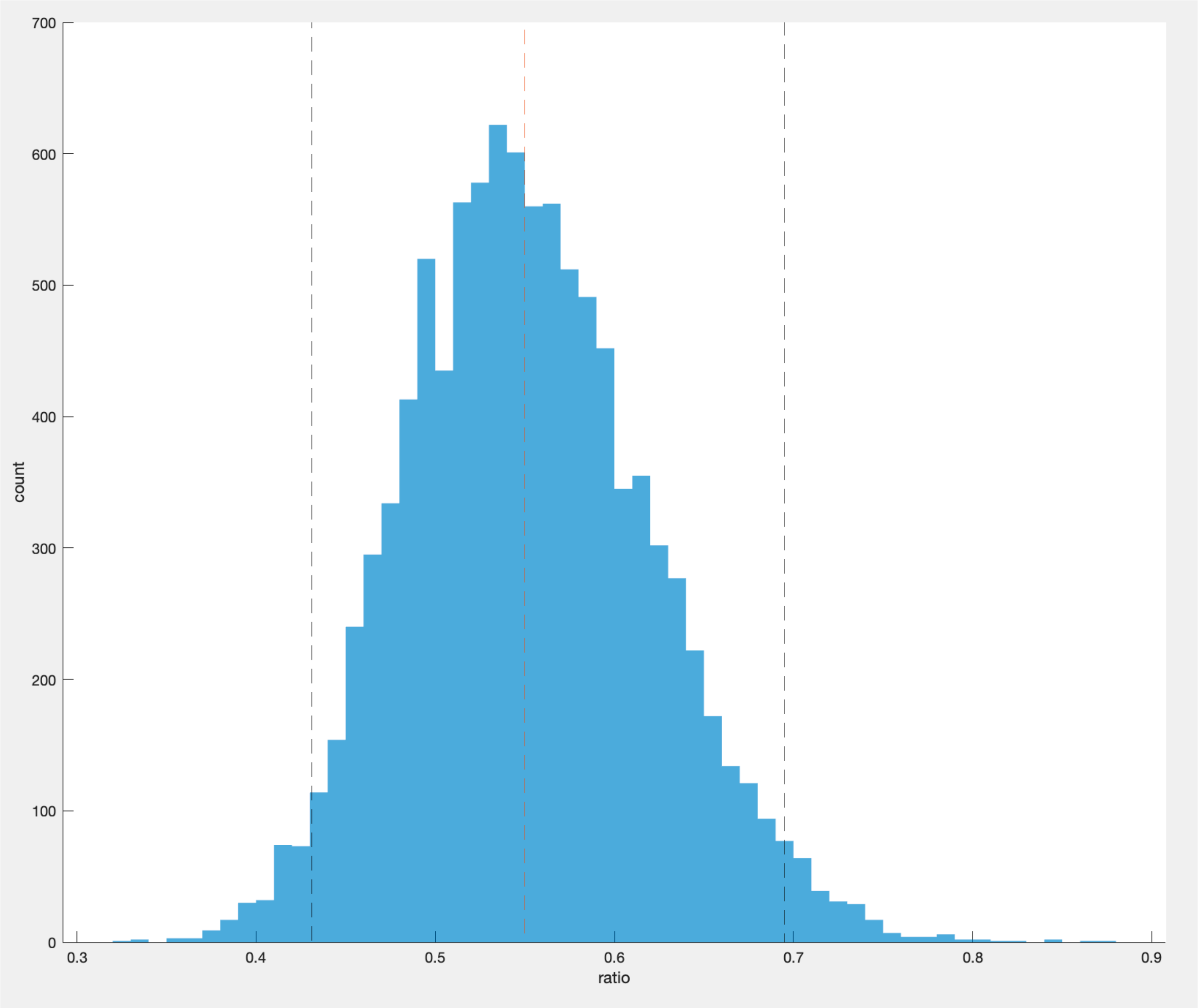


From Lyndon Duong

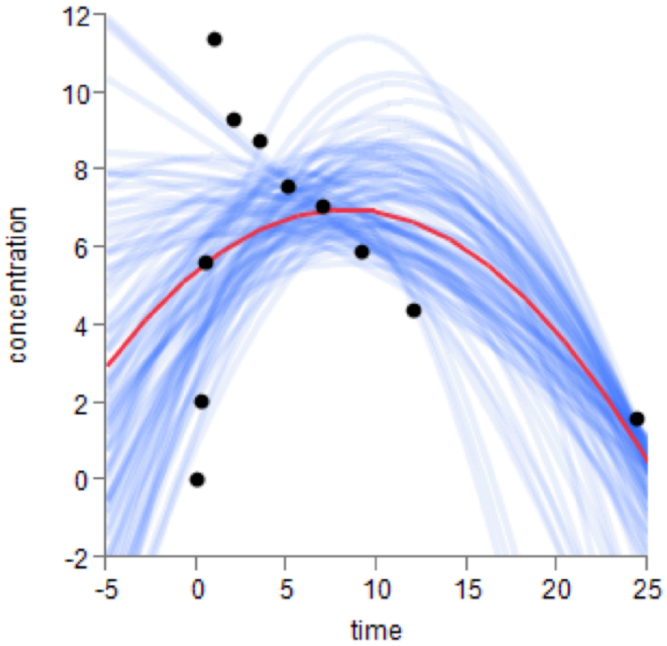

Aspirin Group
Placebo Group
Distribution
(Re-)Sample
(with replacement!)
95% Confidence Interval
Statistic Distribution

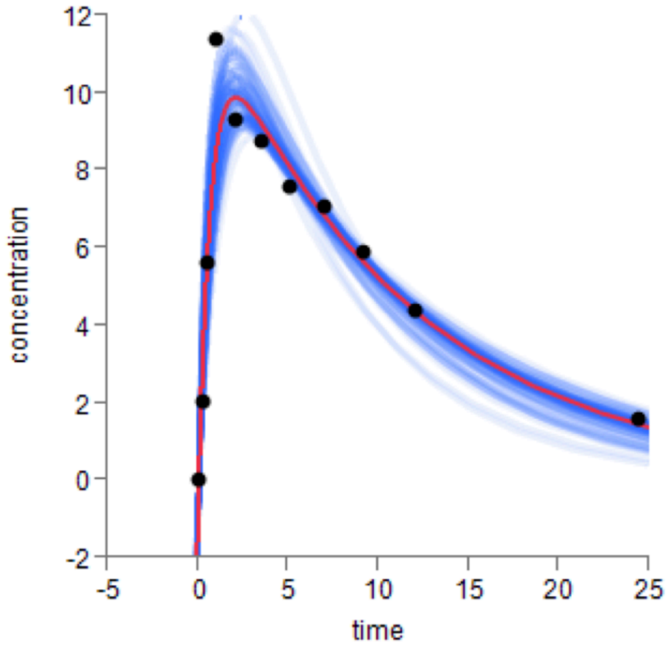
```
% Find 95% confidence interval
ratio_boot_sorted = sort(ratio_boot); % Sort ratio_boots from lowest to highest
lower_bound_ind = floor(0.025 * n_boot); % Find index of 2.5% value
upper_bound_ind = floor(0.975 * n_boot); % Find index of 97.5% value

lower_bound = ratio_boot_sorted(lower_bound_ind); % Use lower index to find value corresponding to 2.5% position
upper_bound = ratio_boot_sorted(upper_bound_ind); % Use higher index to find value corresponding to 97.5% position
```

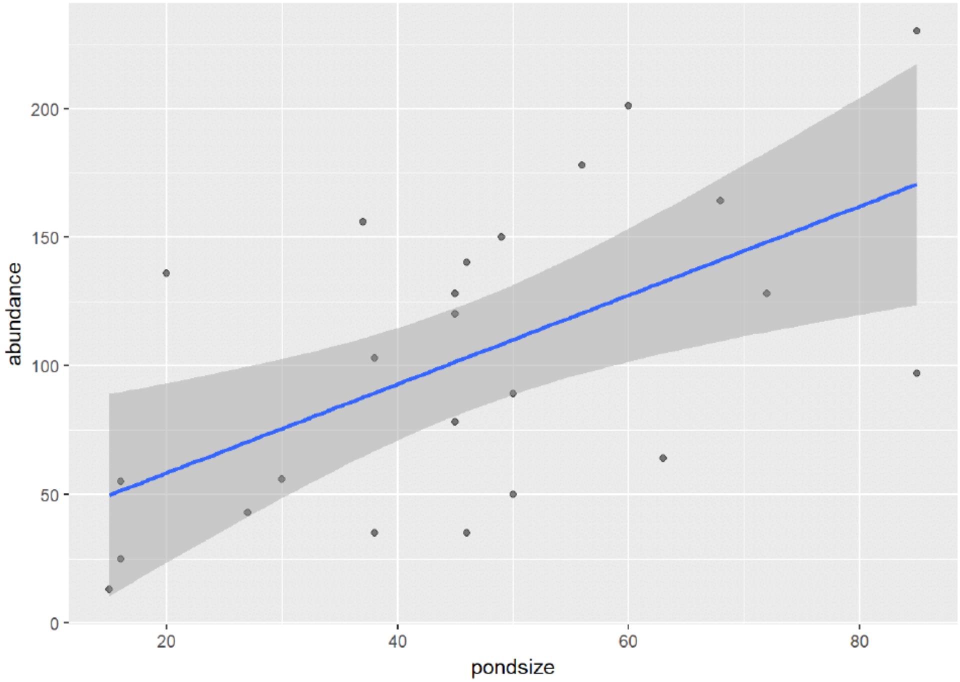
N = 10,000 bootstrapped ratios



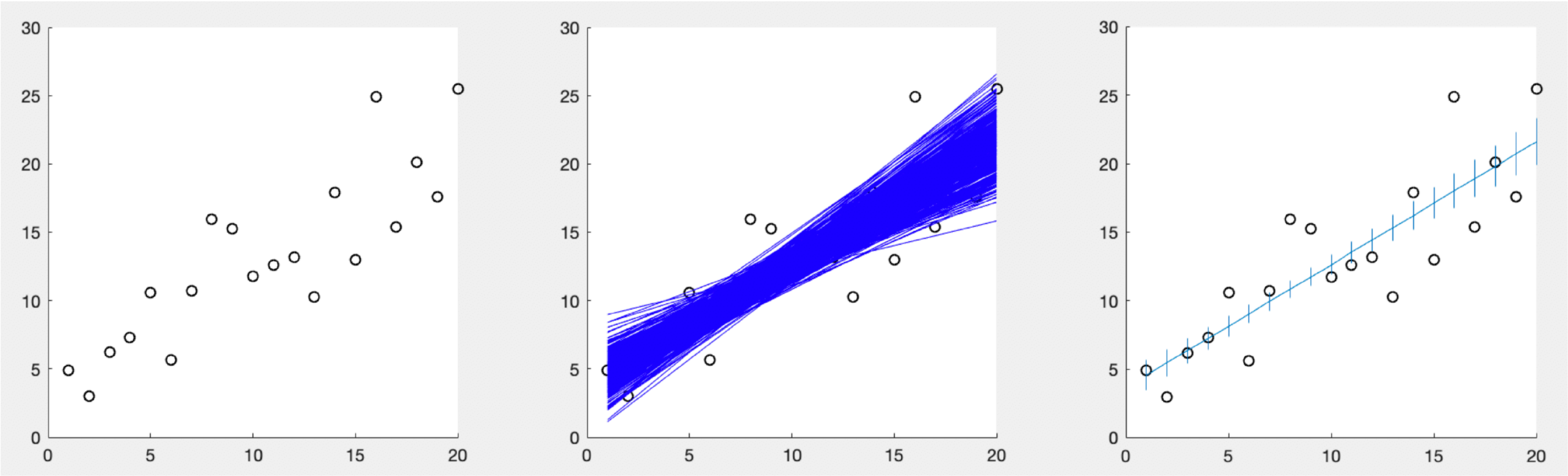
Bootstrapping



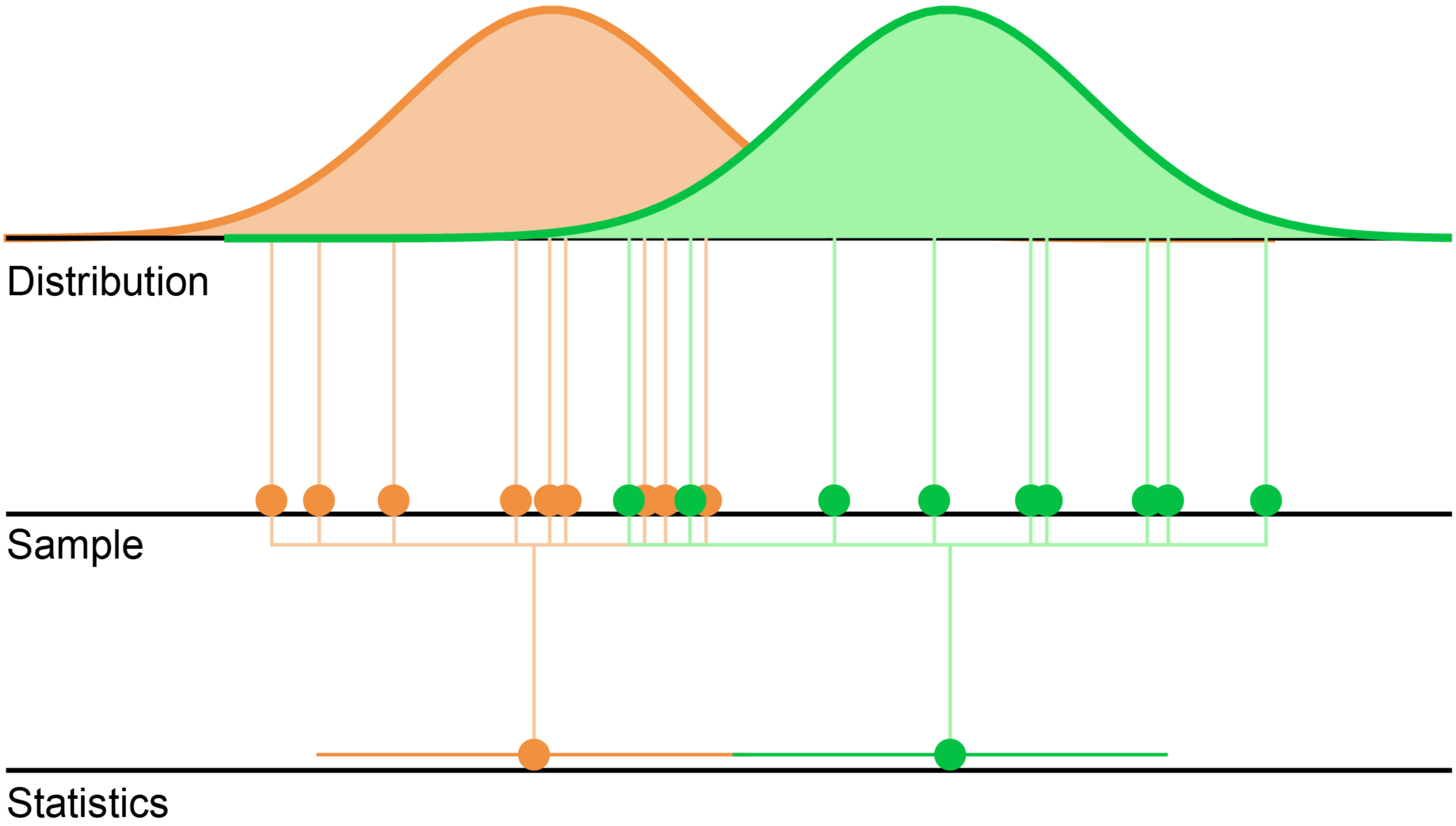




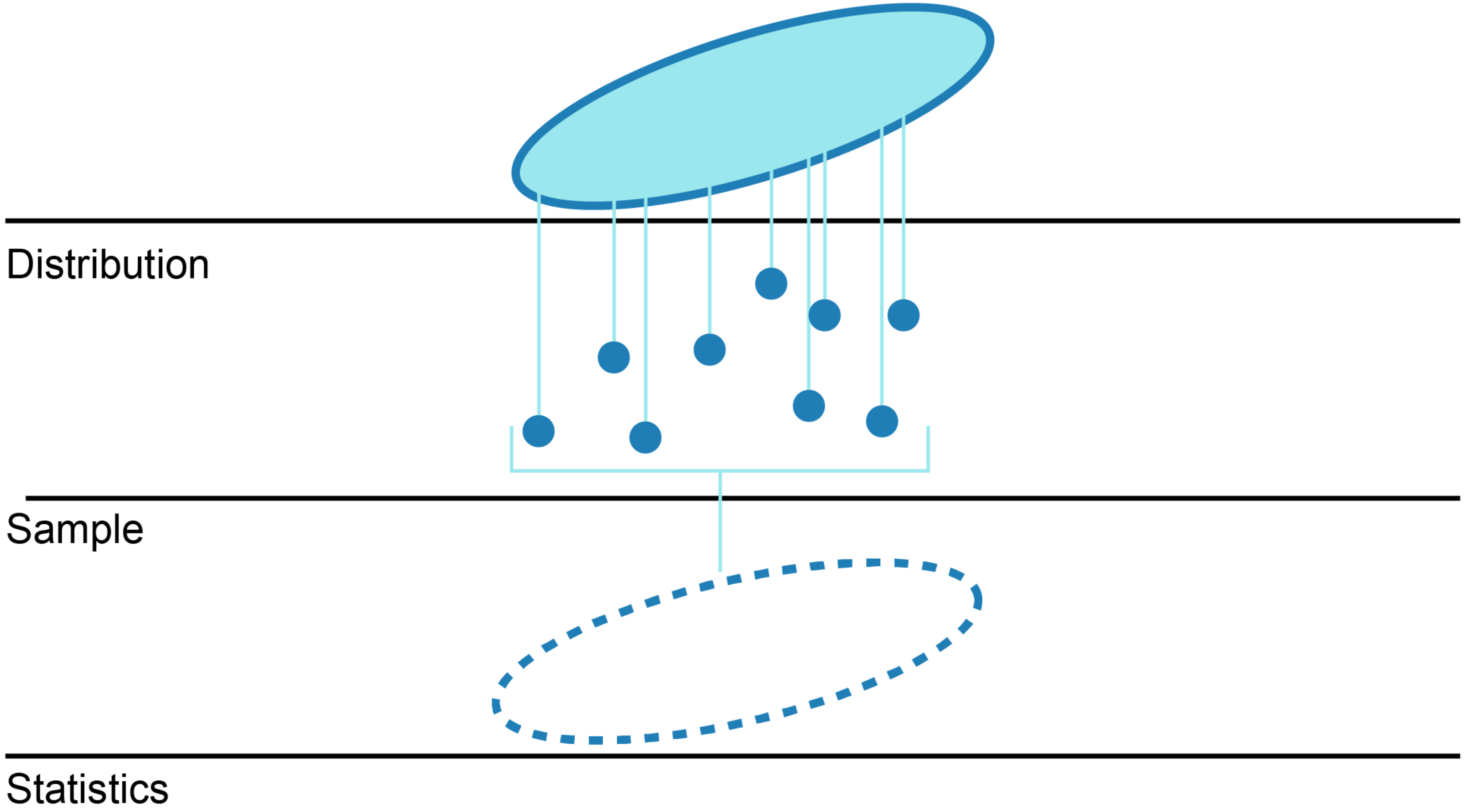
Exercise 1: Model fitting using bootstrapping



2) Are the groups different?



3) Does a pattern exist in the data?



Permutation testing (randomization)

Bootstrapping: generally used to generate confidence intervals about an estimate

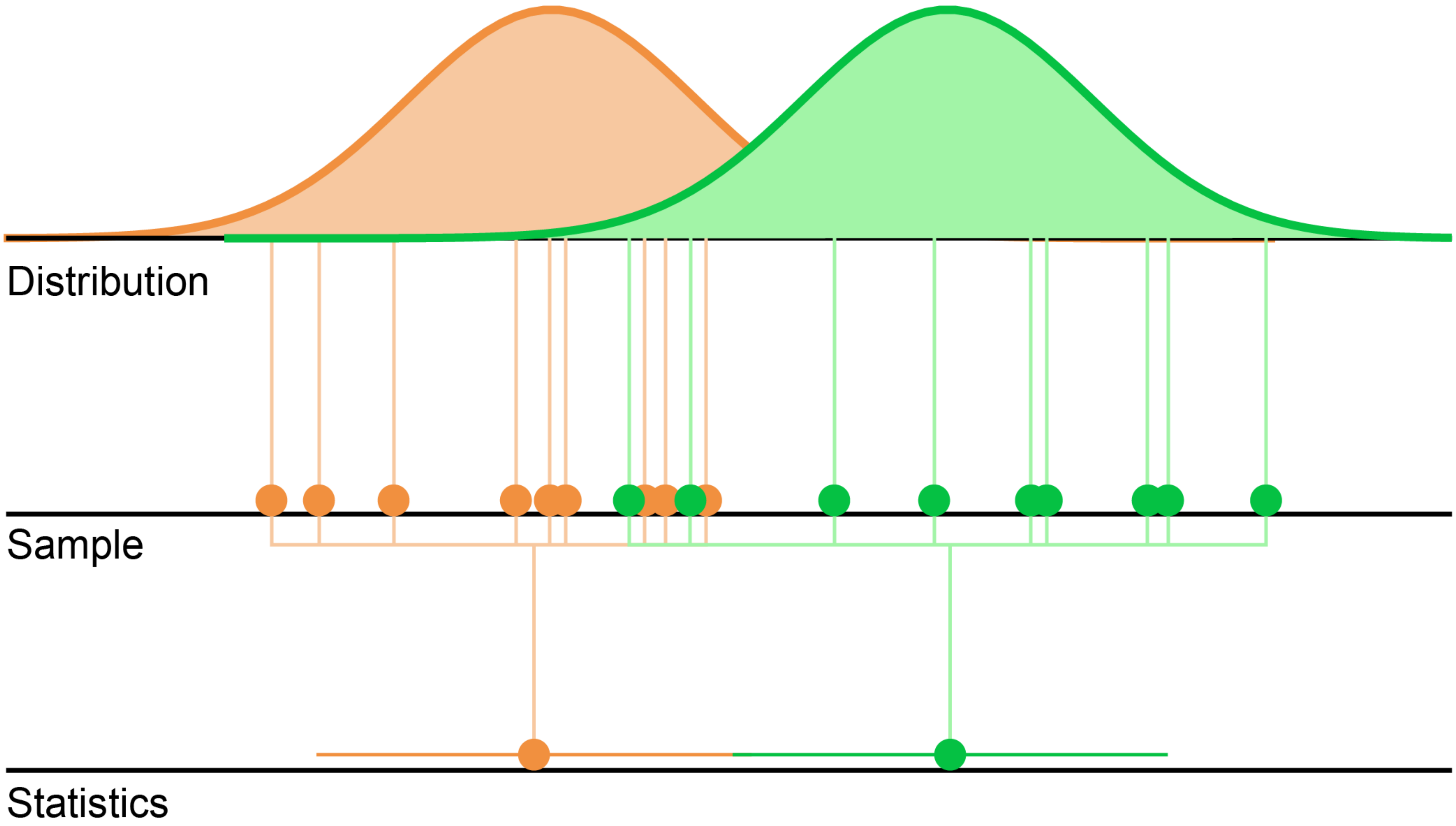
- "How certain am I about my observed estimate (e.g. mean, median, ratio, etc)"

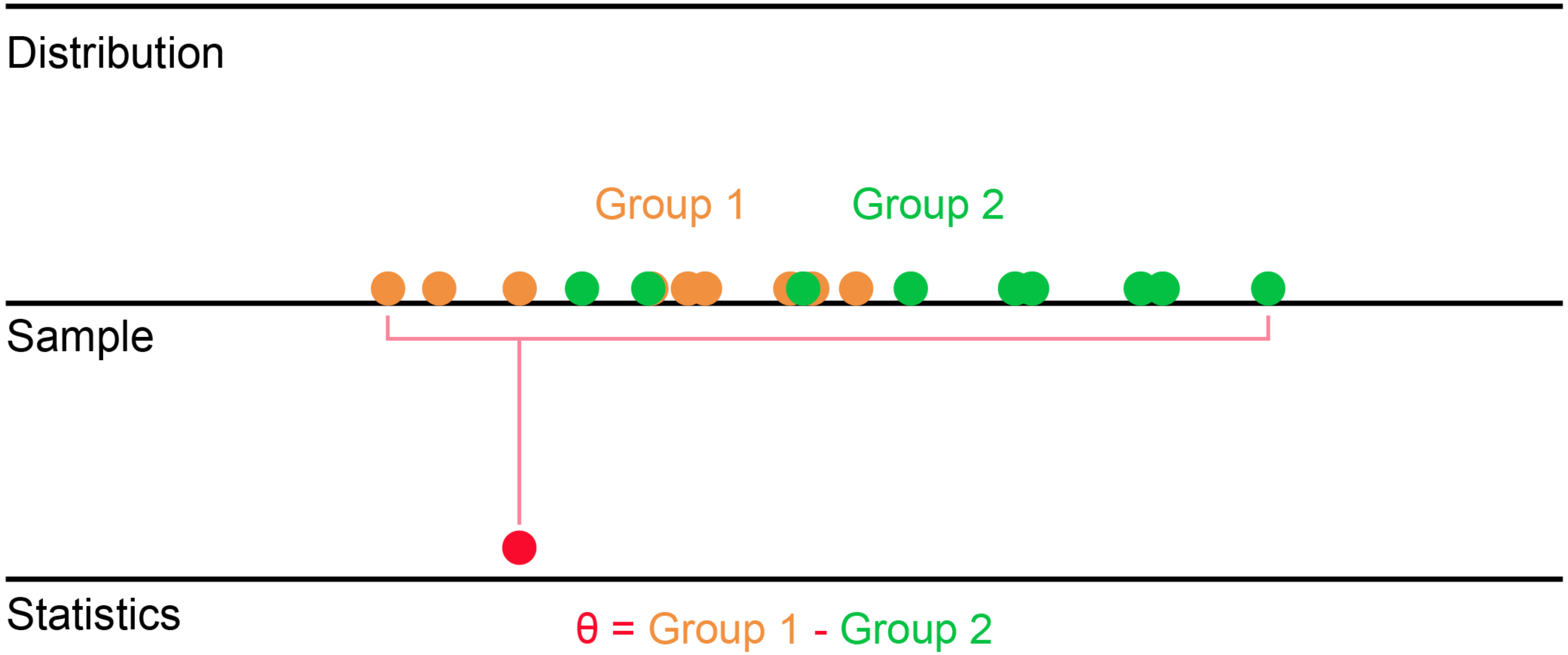
Permutation test: generate a null distribution of estimates with which to test your estimate against

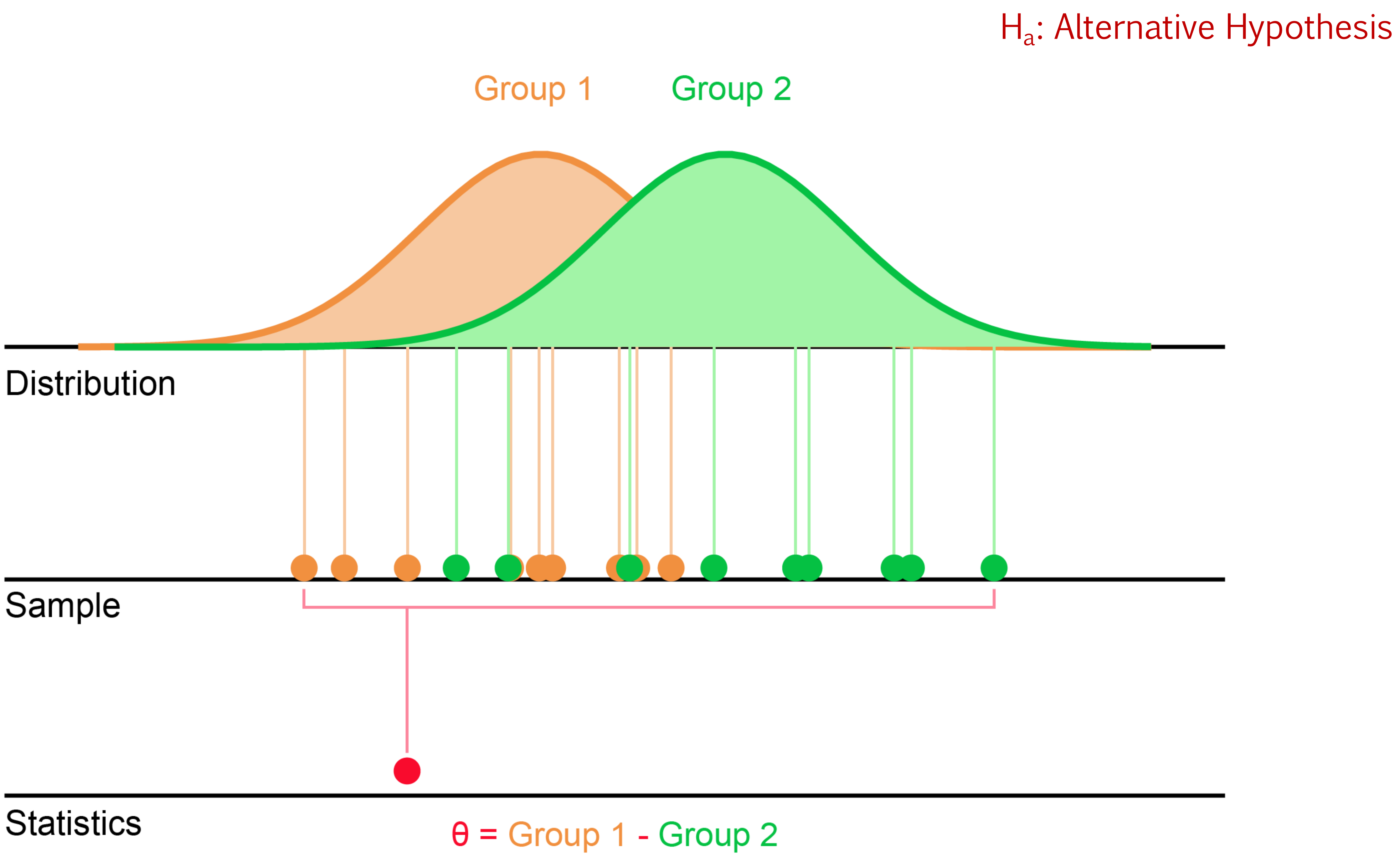
- "Is my observed estimate difference from chance?"

From Lyndon Duong

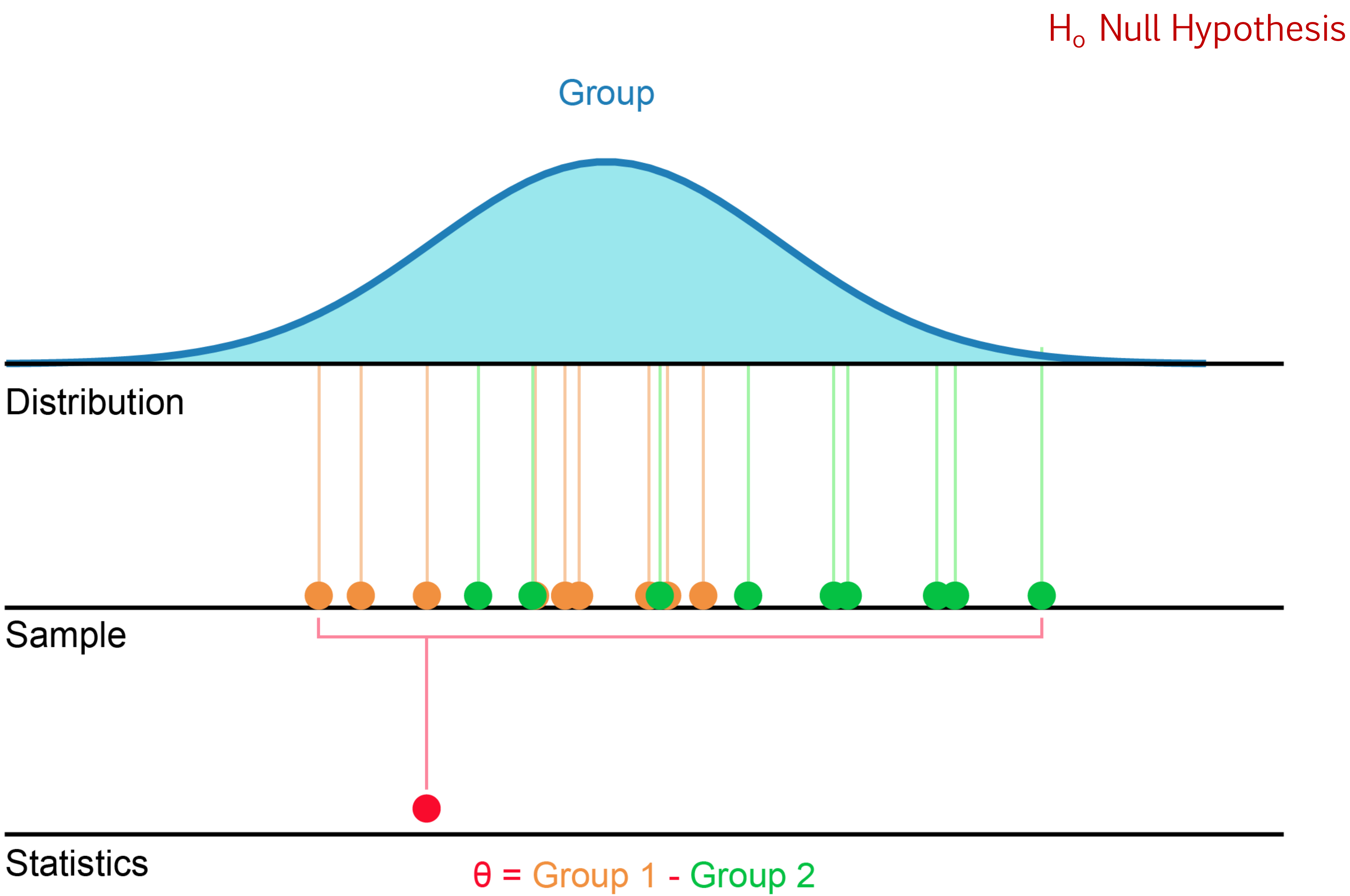
2) Are the groups different?



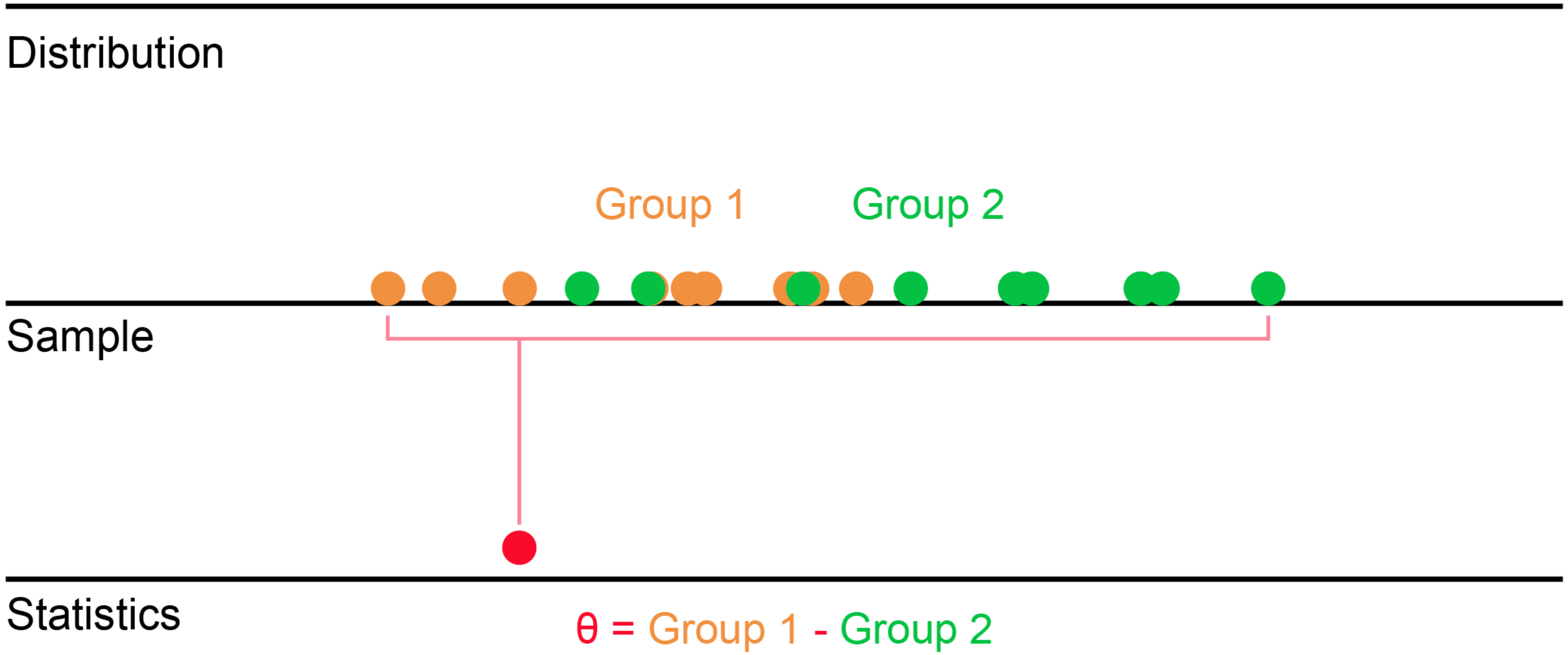

Distribution
Group 1
Group 2
Sample
Statistics
θ = Group 1 - Group 2



H_a: Alternative Hypothesis
Group 1
Group 2
Distribution
Sample
Statistics
θ = Group 1 - Group 2



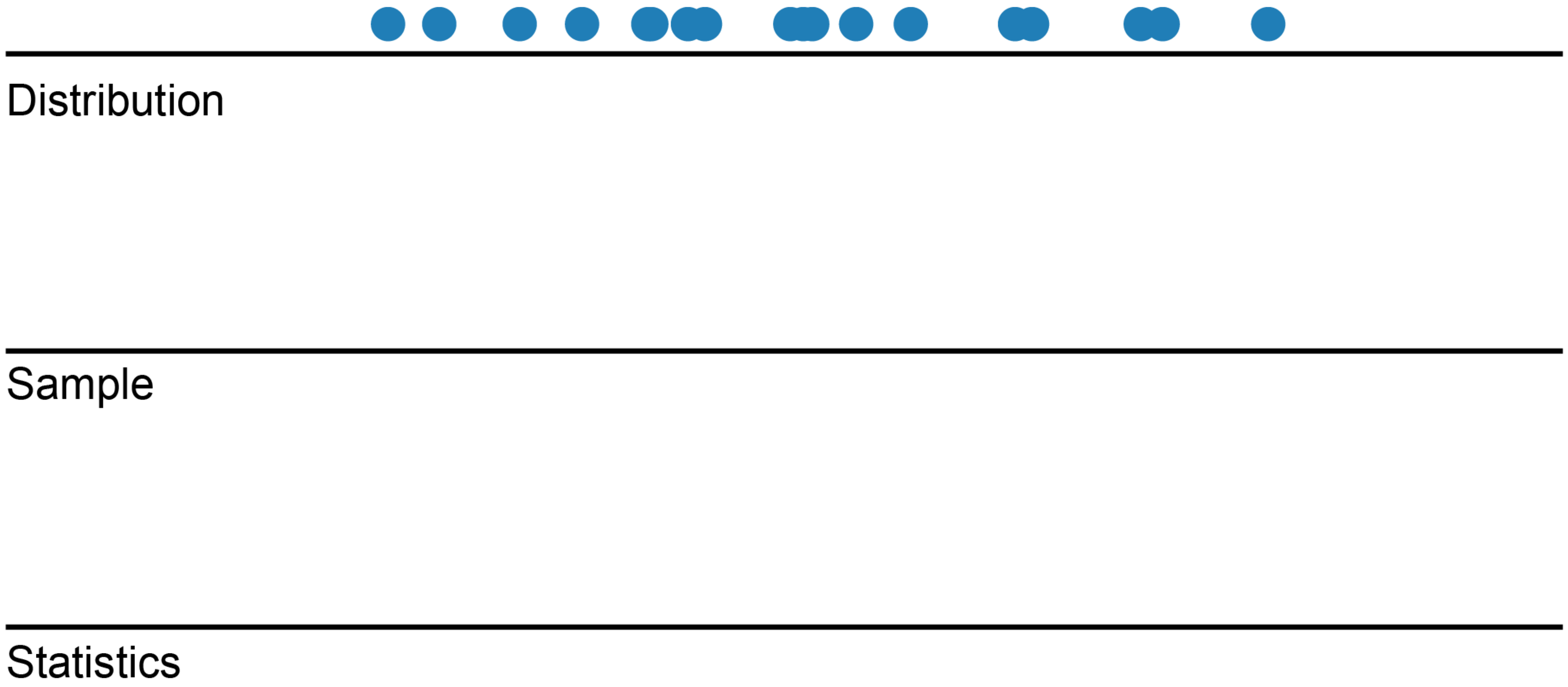
H_o Null Hypothesis
Group
Distribution
Sample
Statistics
θ = Group 1 - Group 2


Distribution
Group 1
Group 2
Sample
Statistics
θ = Group 1 - Group 2

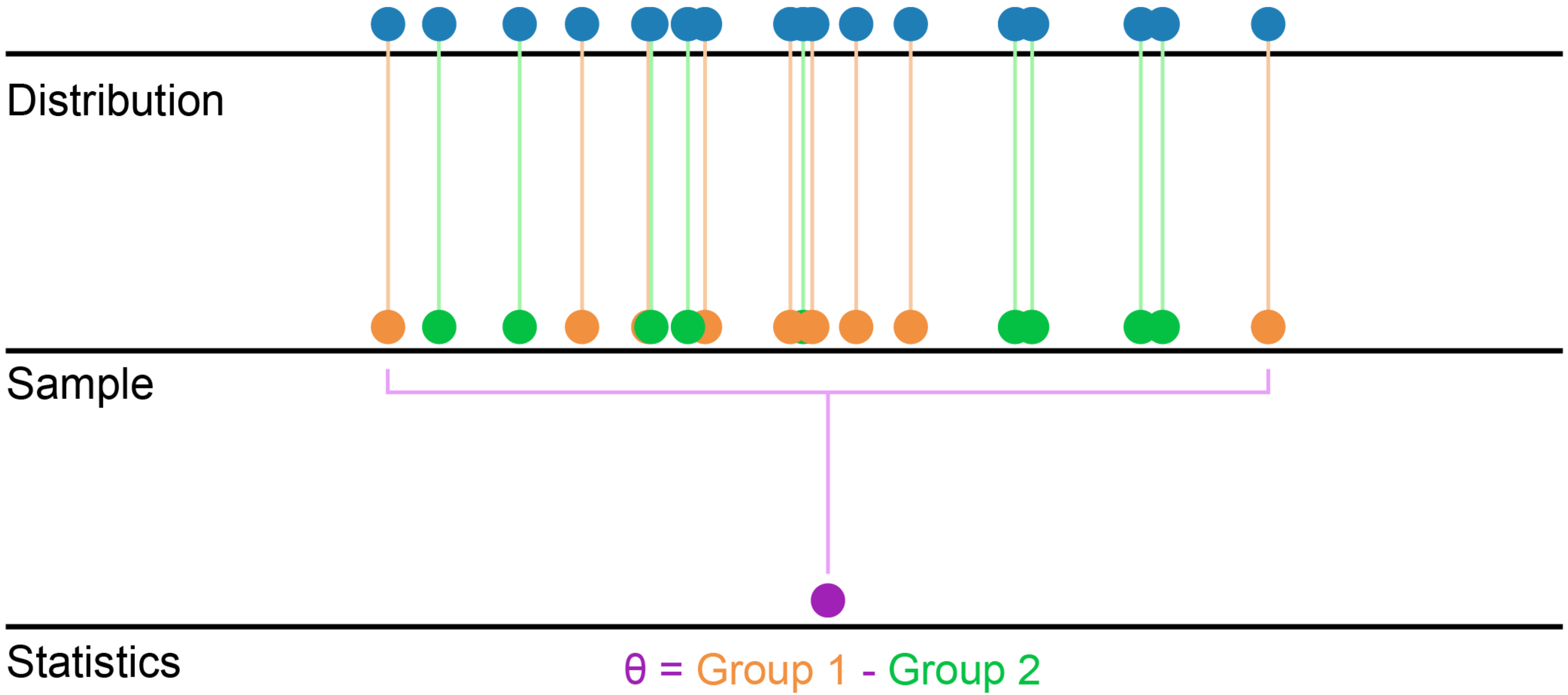
Distribution

Sample

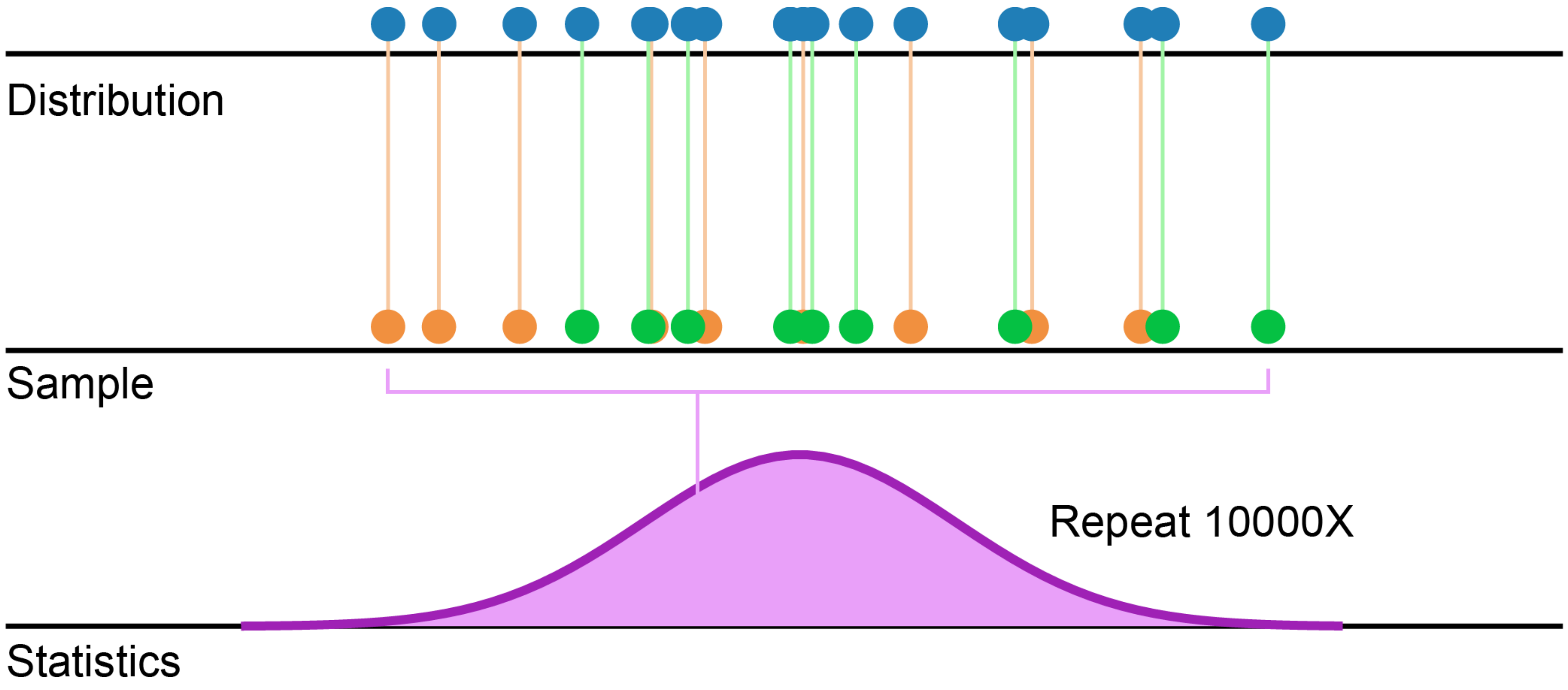
Statistics


Distribution
Sample
Statistics

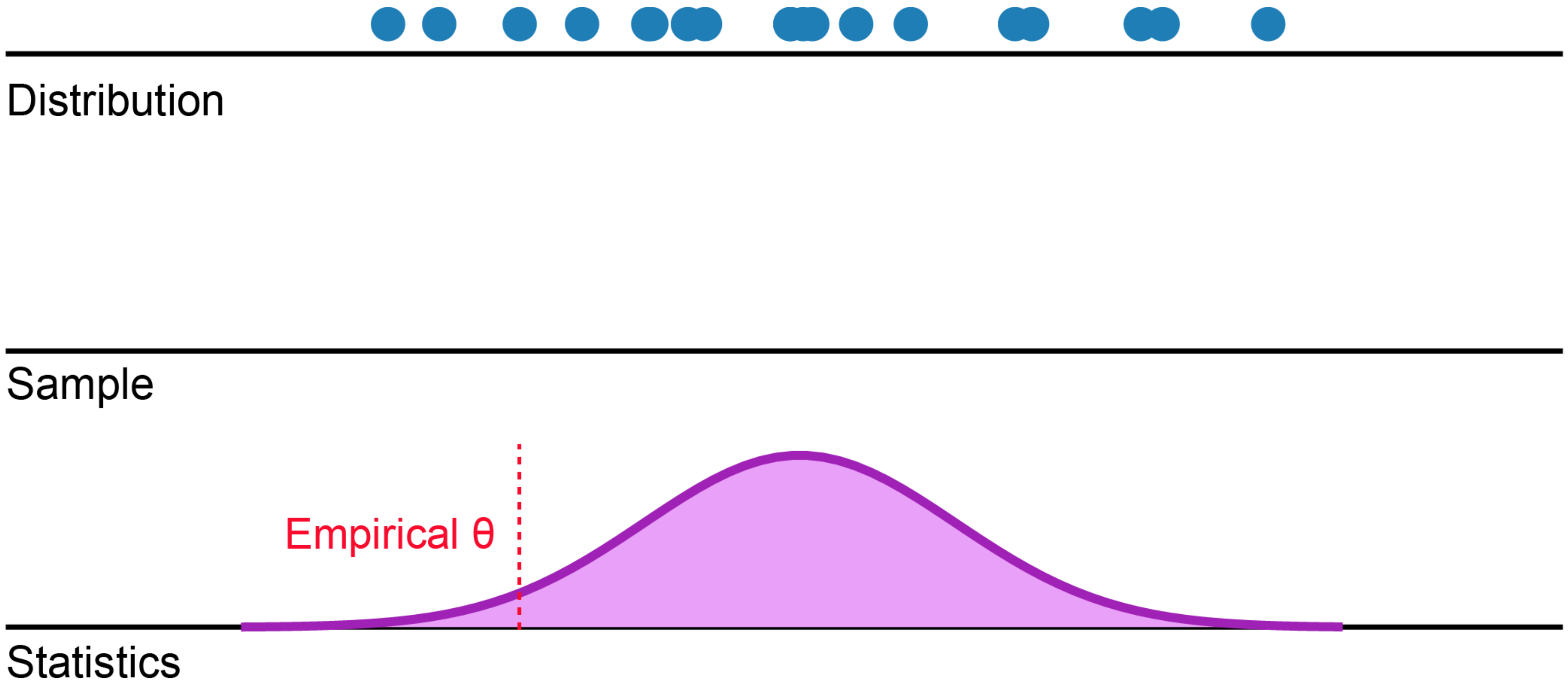
Permutation / shuffle / re-sample without replacement

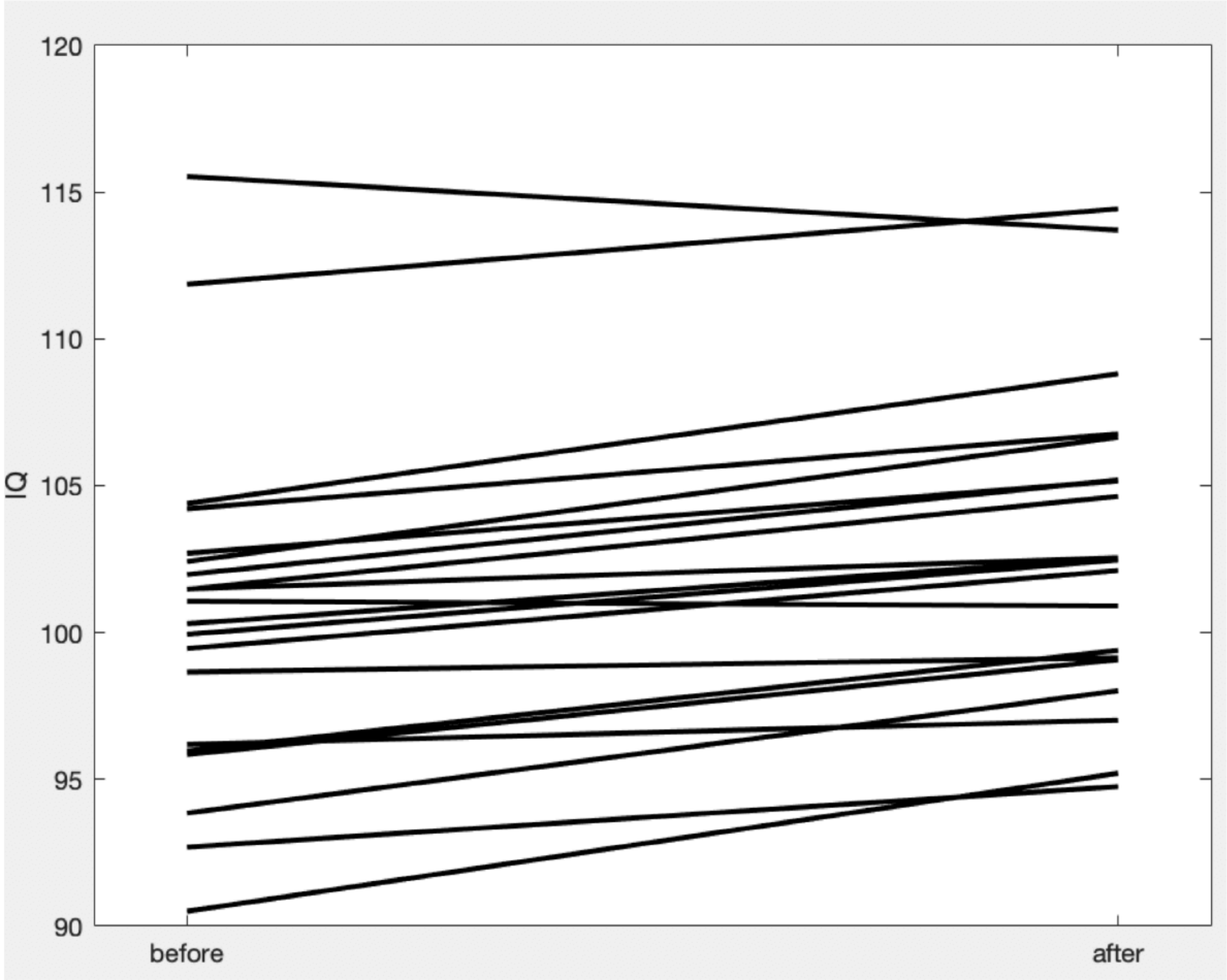


Permutation / shuffle / re-sample without replacement



Permutation test creates a null distribution of estimates that we can use to test the null hypothesis





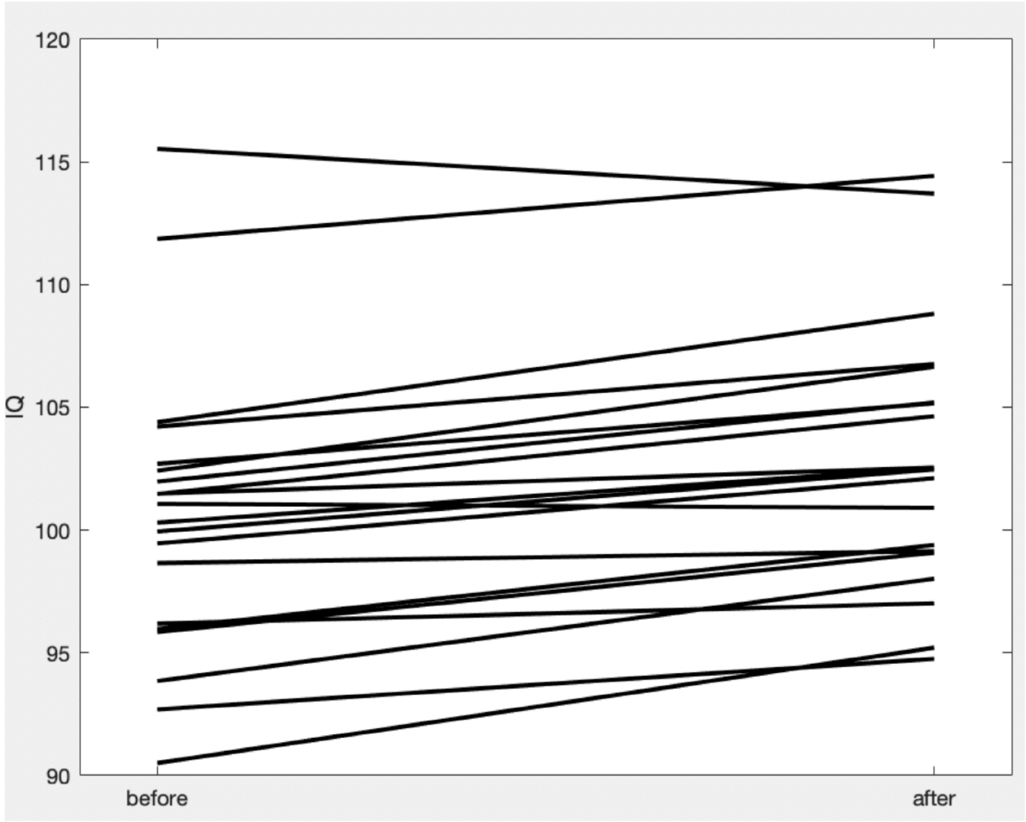
Empirical mean difference = 2.55 IQ pts

Is this change statistically significant?

From Lyndon Duong

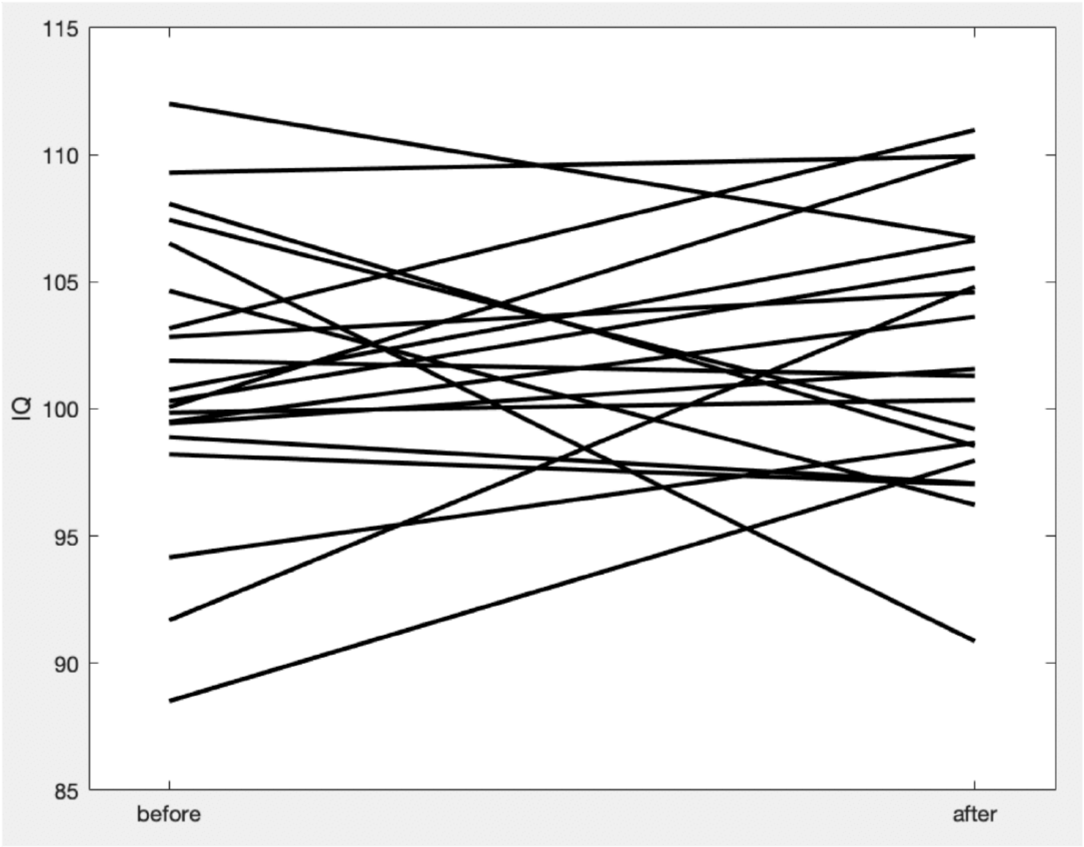
```
% Perform permutation / randomization
n_perm = 10000;
all_data = [drug_before, drug_after]; % concatenate the data
perm_diffs = zeros(n_perm,1);
for perm = 1:n_perm
    perm_ind = randperm(2*n_subj); % Obtain a random permuation of indices from
    % 1 to 2*n_pts (2*n_pts = total data points)

    perm_before= all_data(perm_ind(1:n_subj))'; % generate new pre-drug group
    perm_after = all_data(perm_ind(n_subj+1:end))'; % generate new post-drug group

    perm_diffs(perm) = mean(perm_after-perm_before); % calculate new statistic

end
```

From Lyndon Duong



Original data

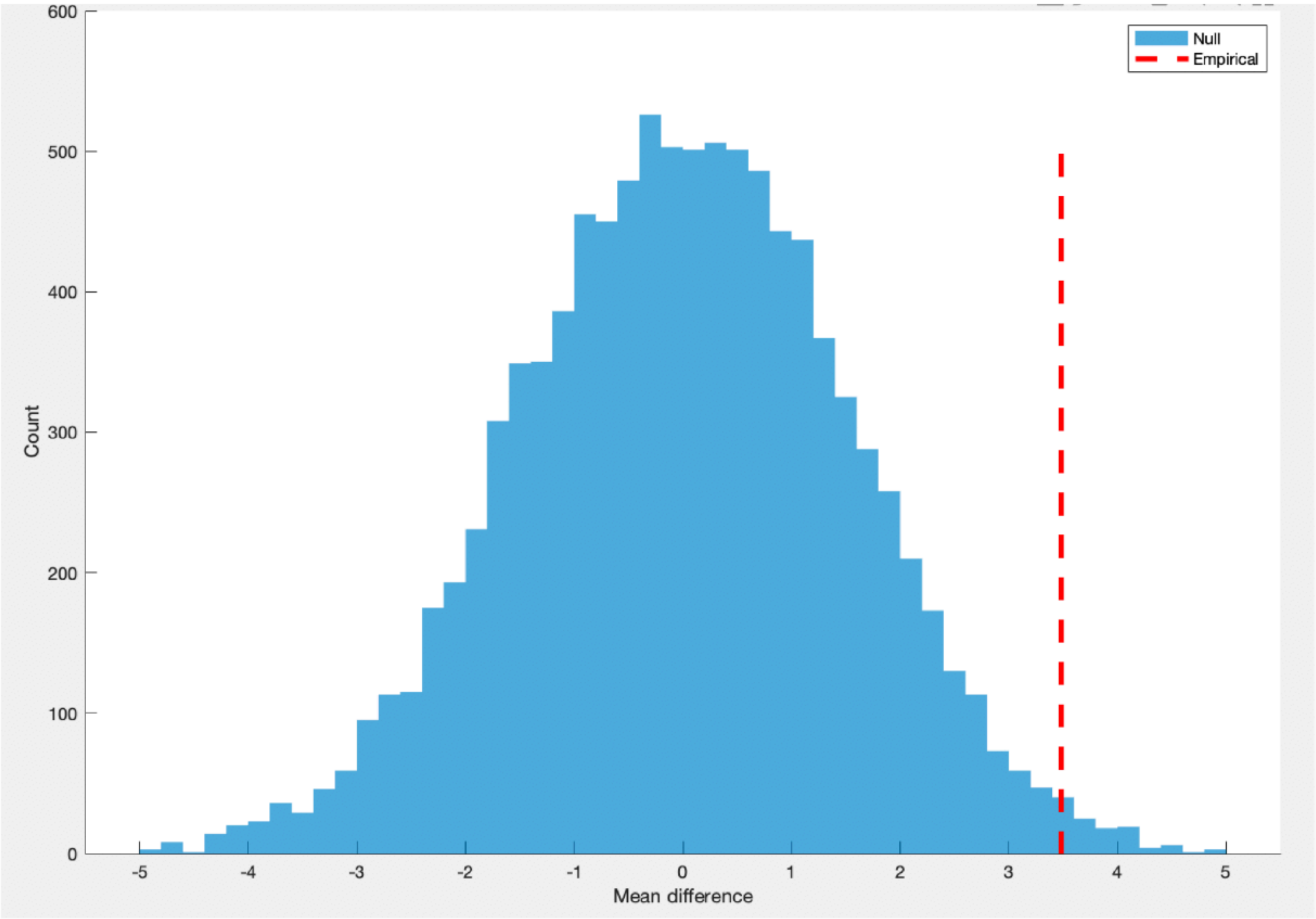
Empirical mean difference = 2.55 IQ pts



Example shuffled data

Shuffled mean difference = 0.7161 IQ pts

From Lyndon Duong



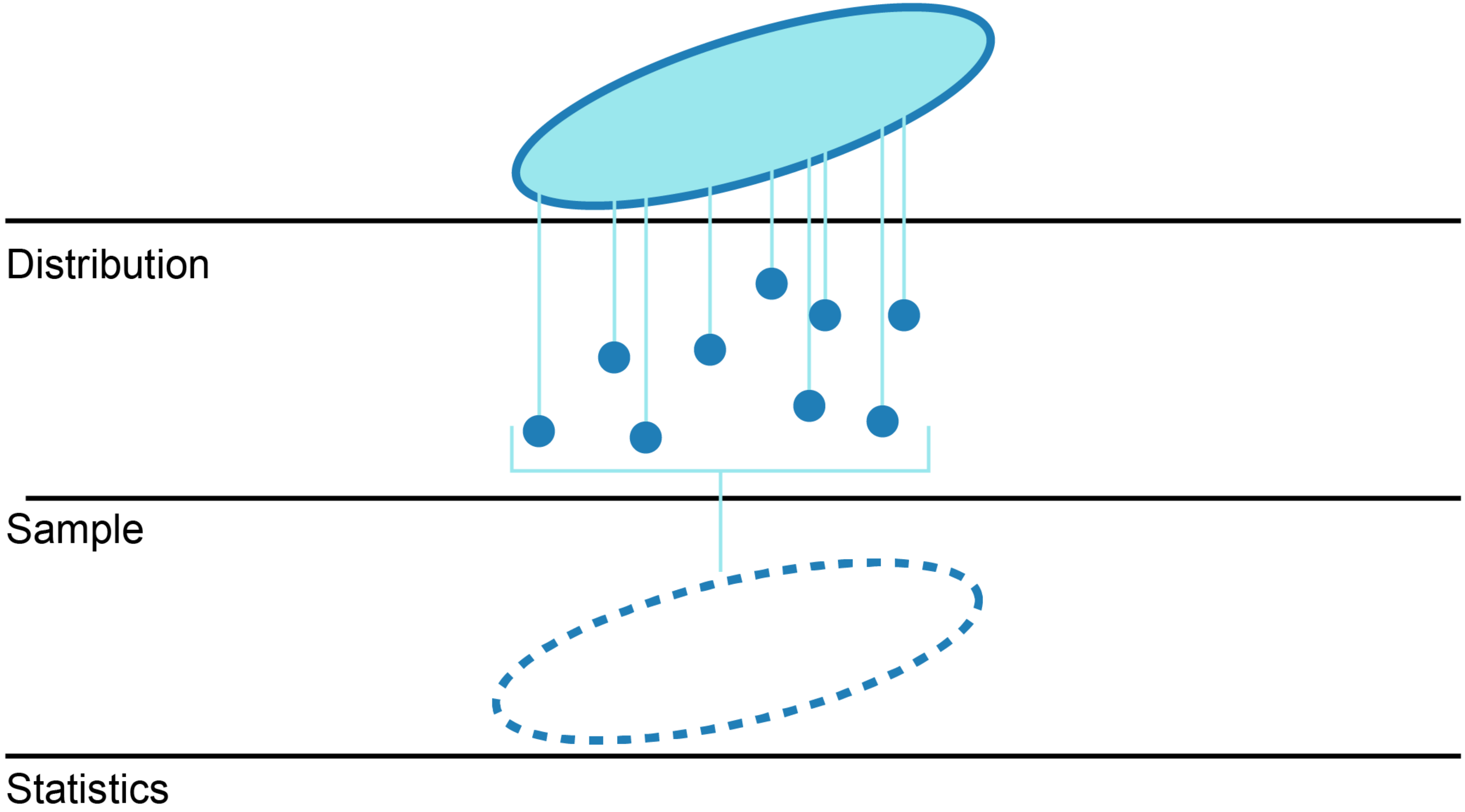
10,000 shuffles and their mean differences

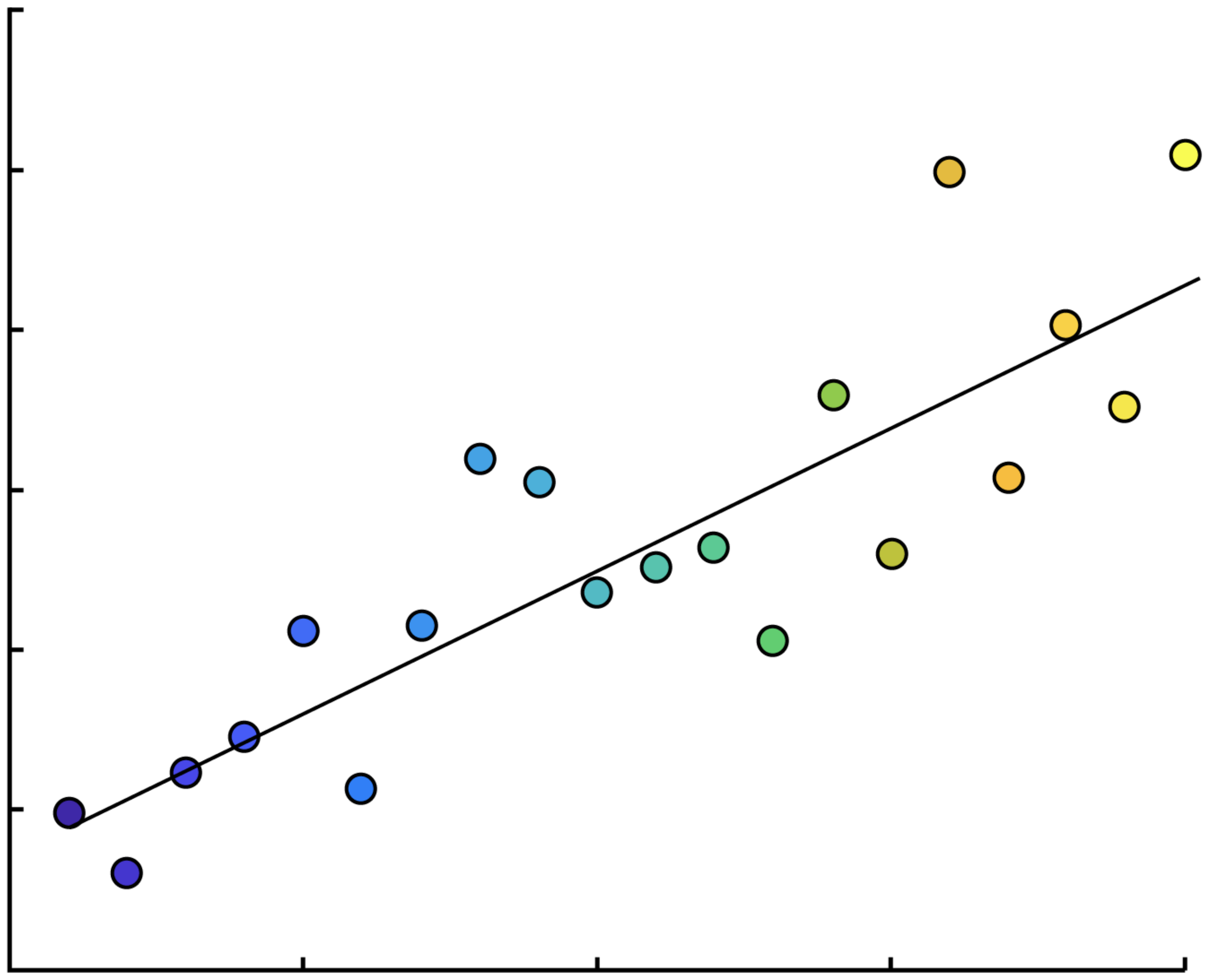
From Lyndon Duong

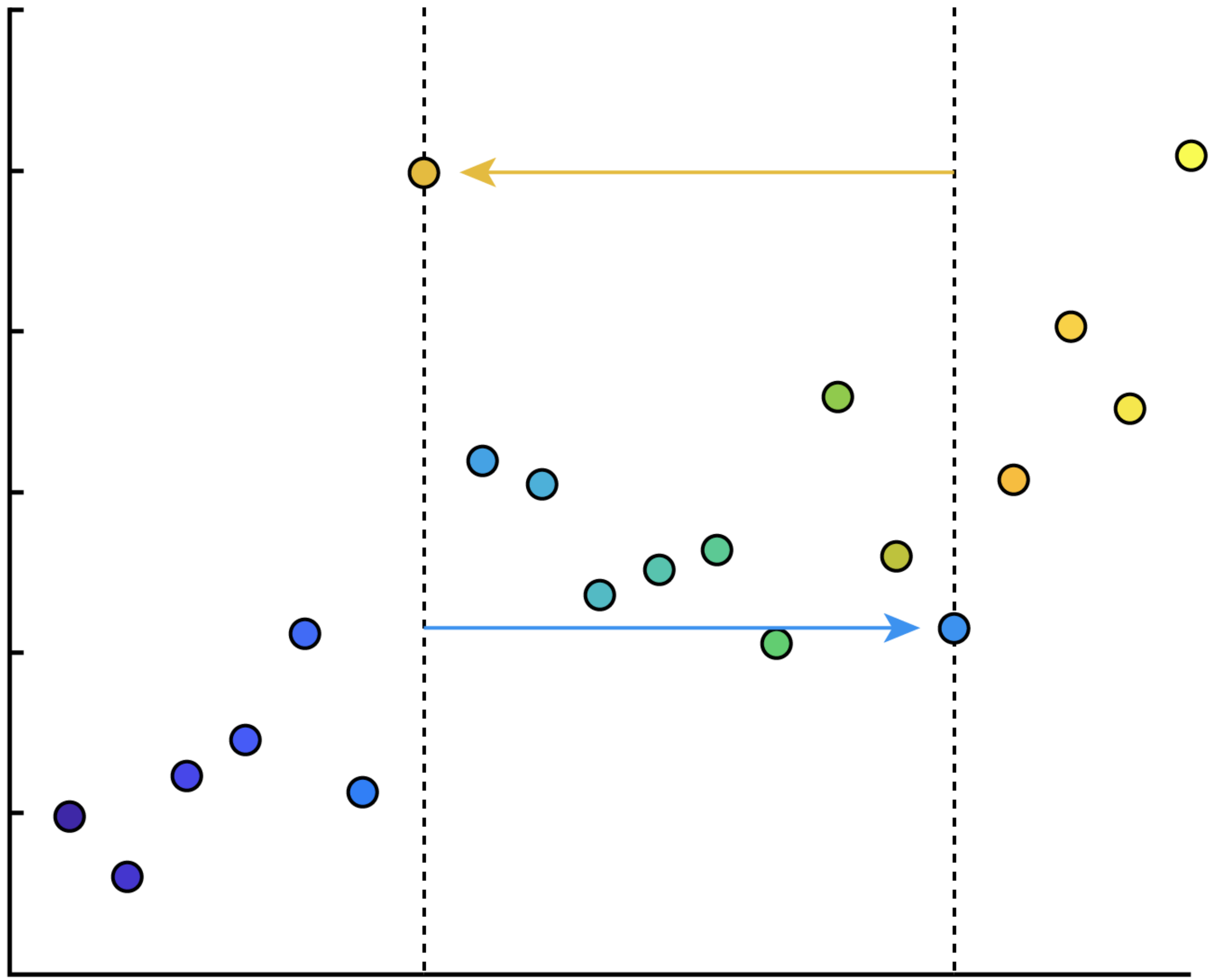
```
p_val = 1 - sum(emp_mean_diff>perm_diffs)/n_perm;
```

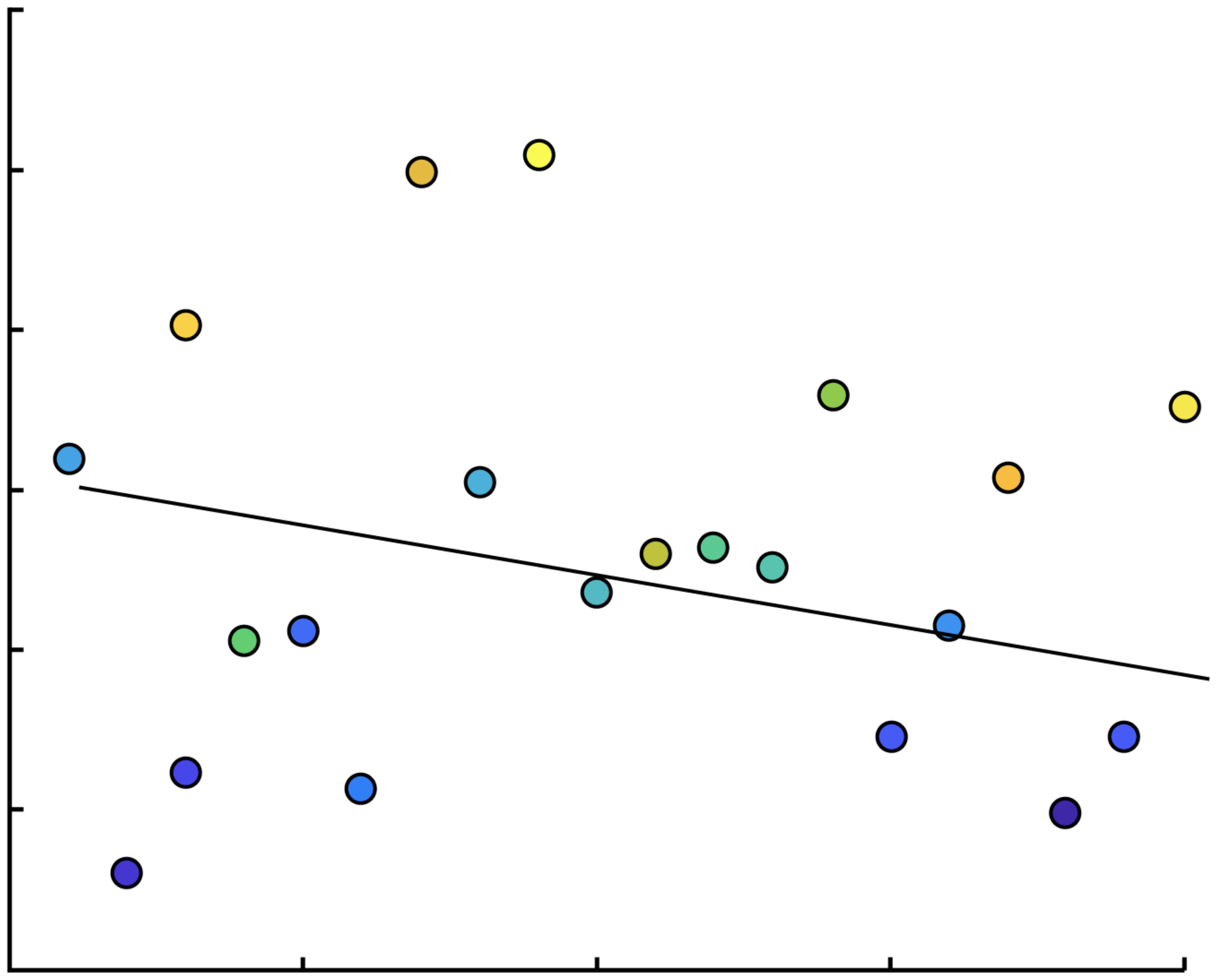
From Lyndon Duong

3) Does a pattern exist in the data?

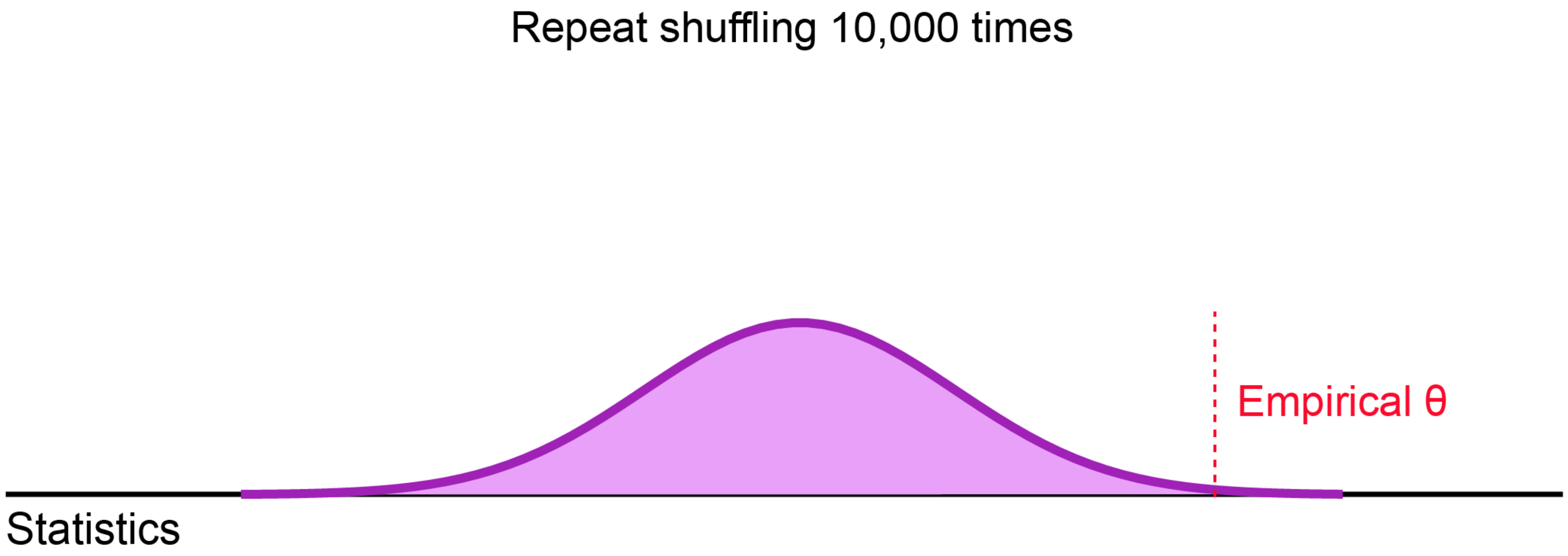




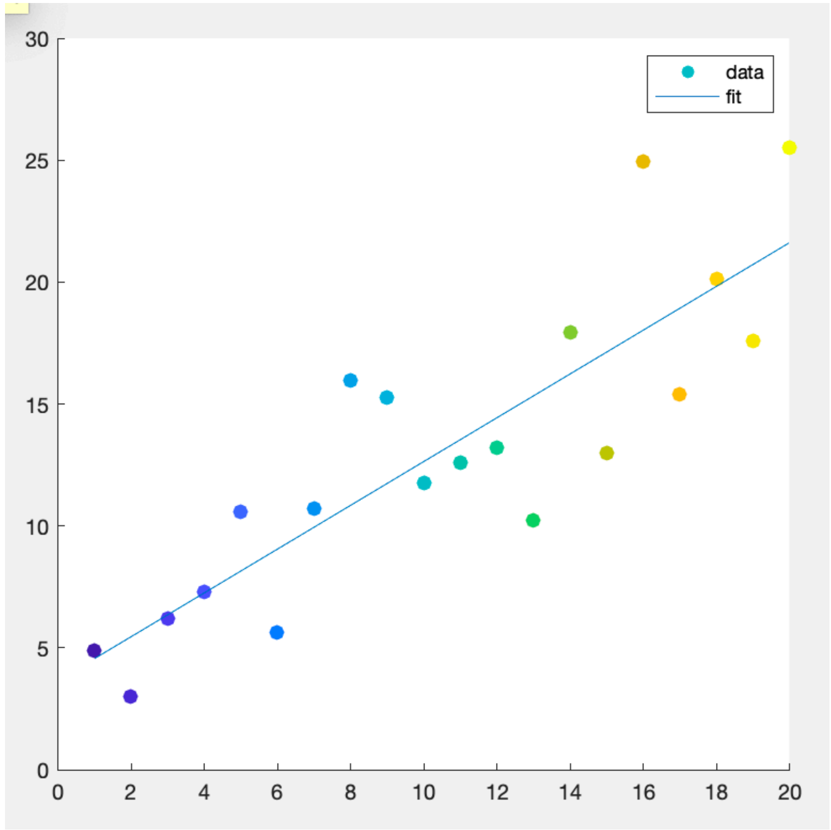


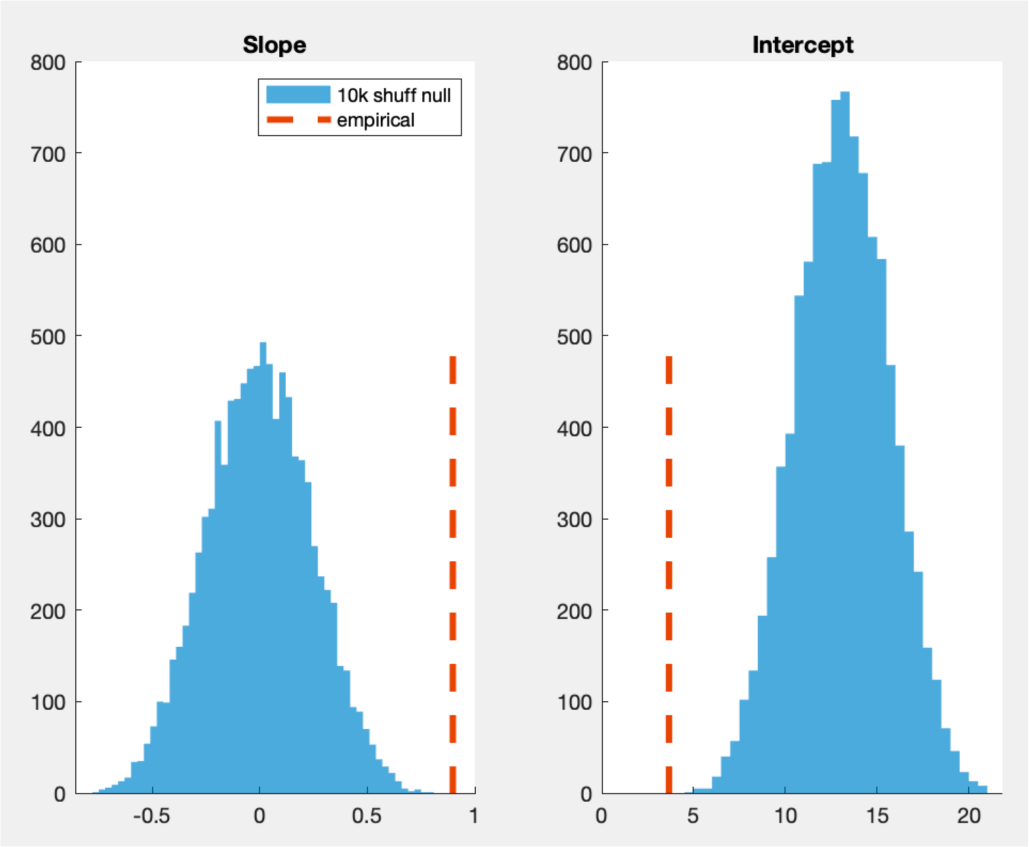


Permutation test creates a null distribution of estimates that we can use to test the null hypothesis

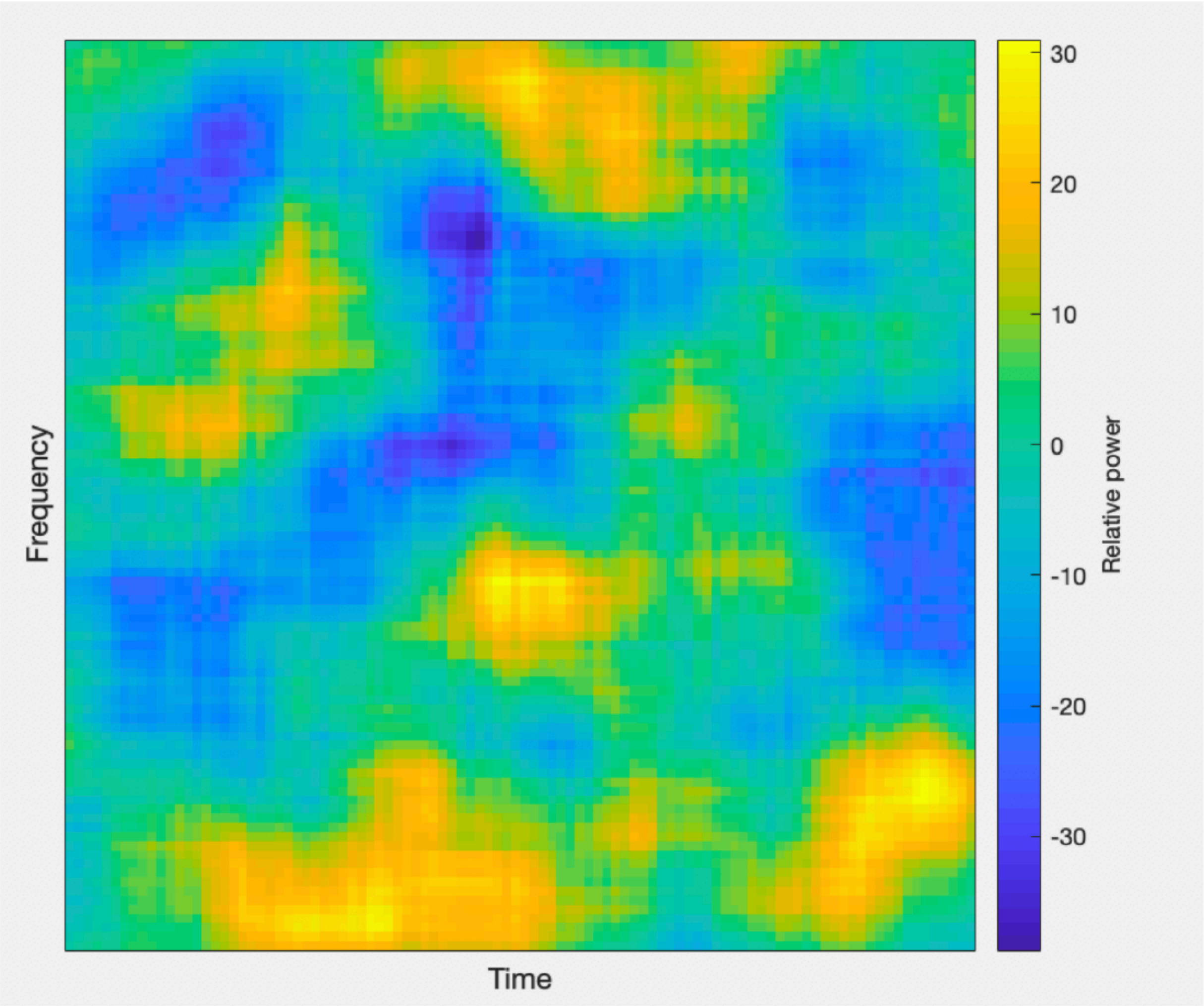


Exercise 2: Testing patterns using permutation





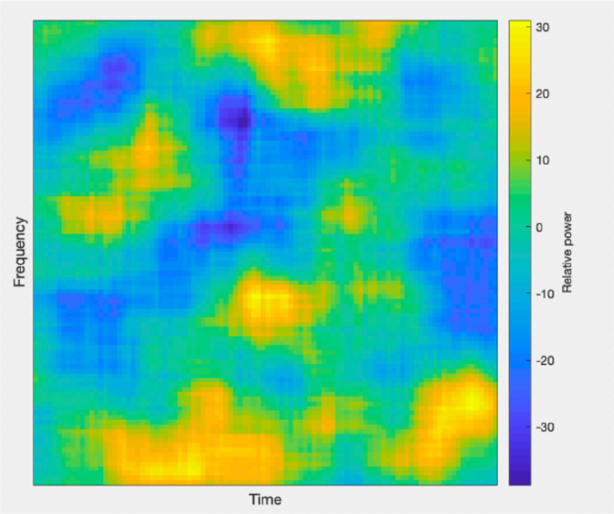
More permutation test applications...



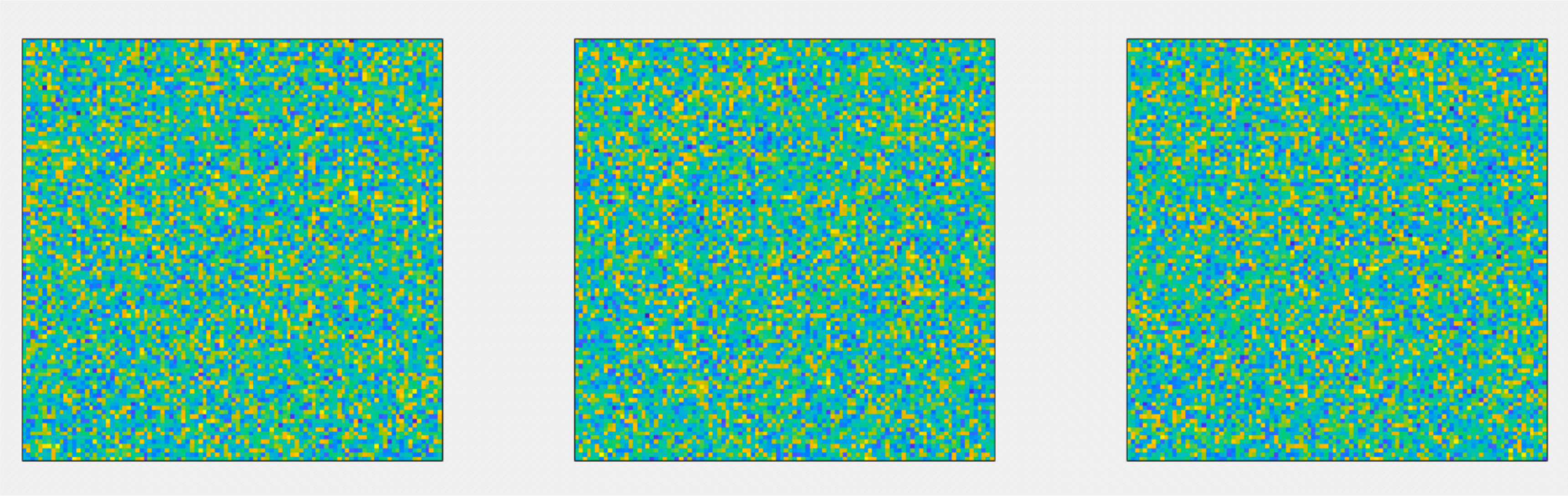
What at which times and frequency is the power different than chance?

From Lyndon Duong

More permutation test applications…



Original



Shuffle #1

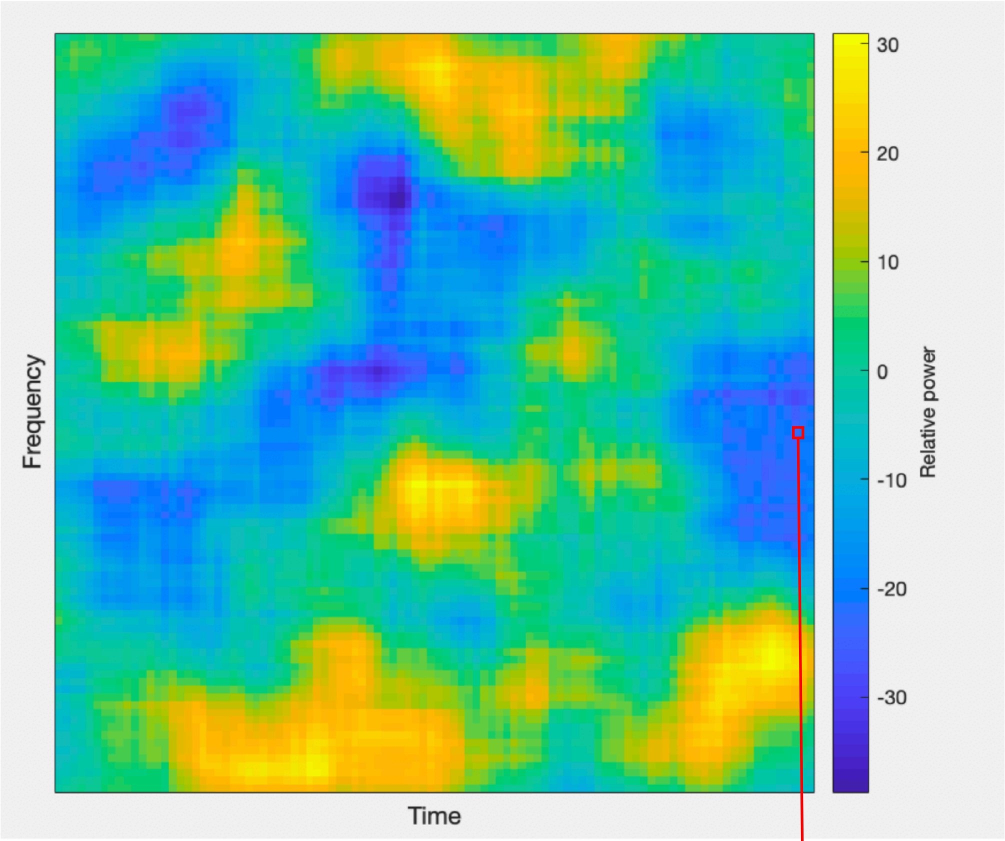
Shuffle #2

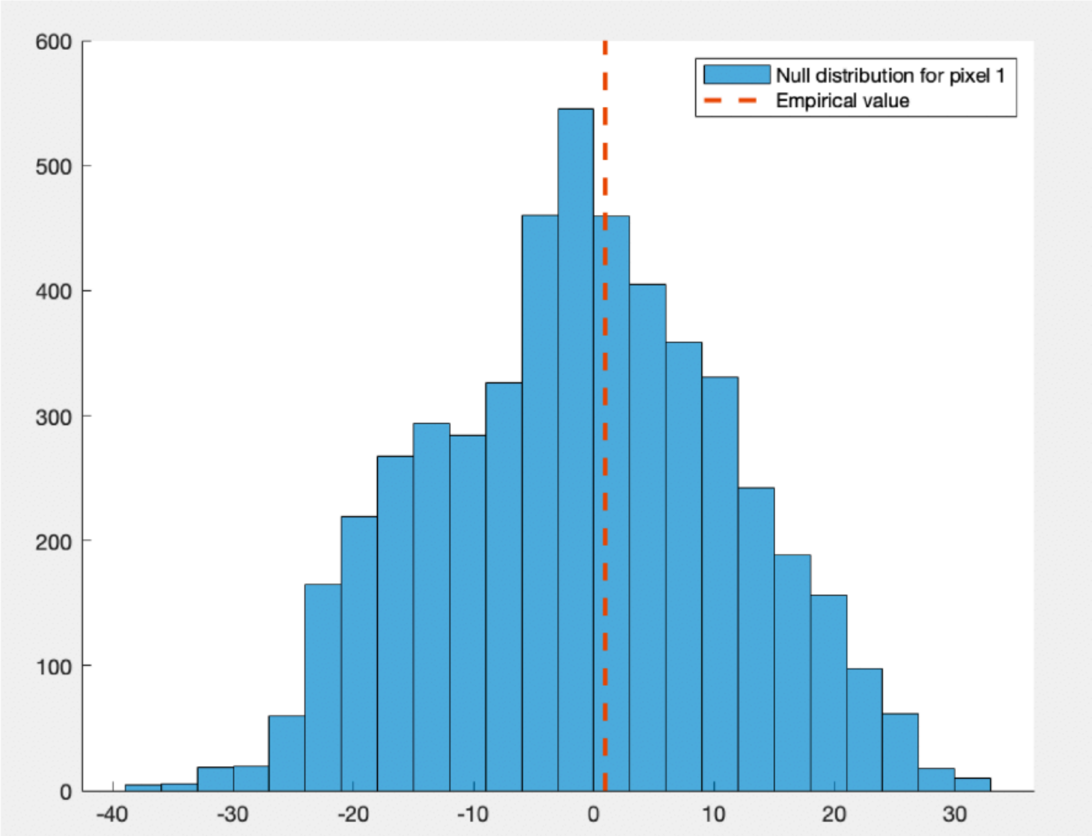
…

Shuffle N

From Lyndon Duong

More permutation test applications...

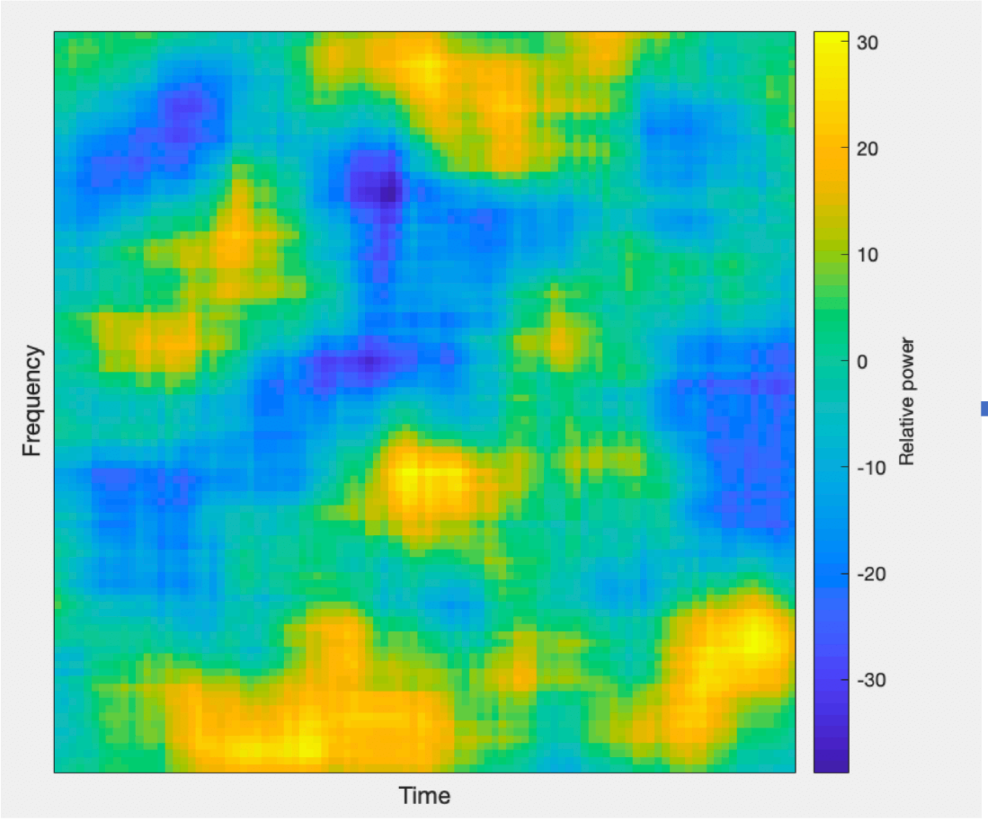


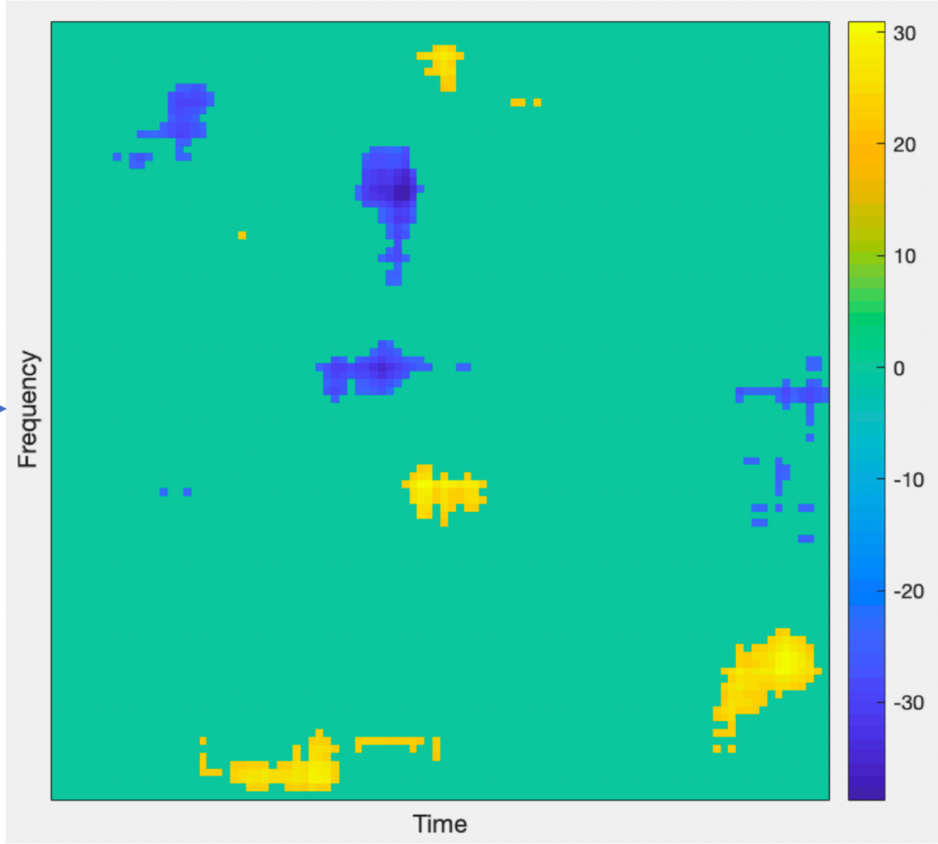


Each pixel has a null distribution

From Lyndon Duong

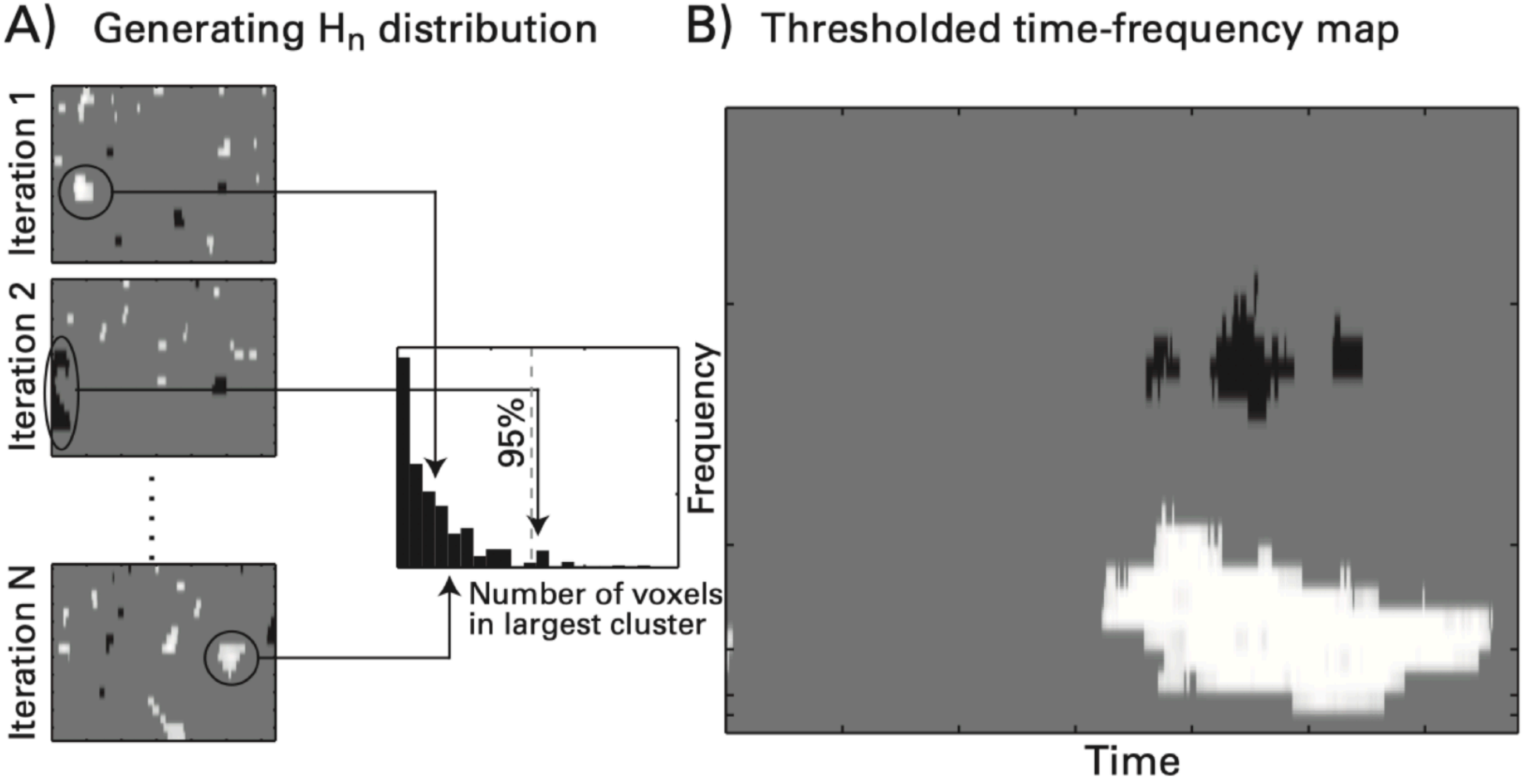
More permutation test applications...





From Lyndon Duong

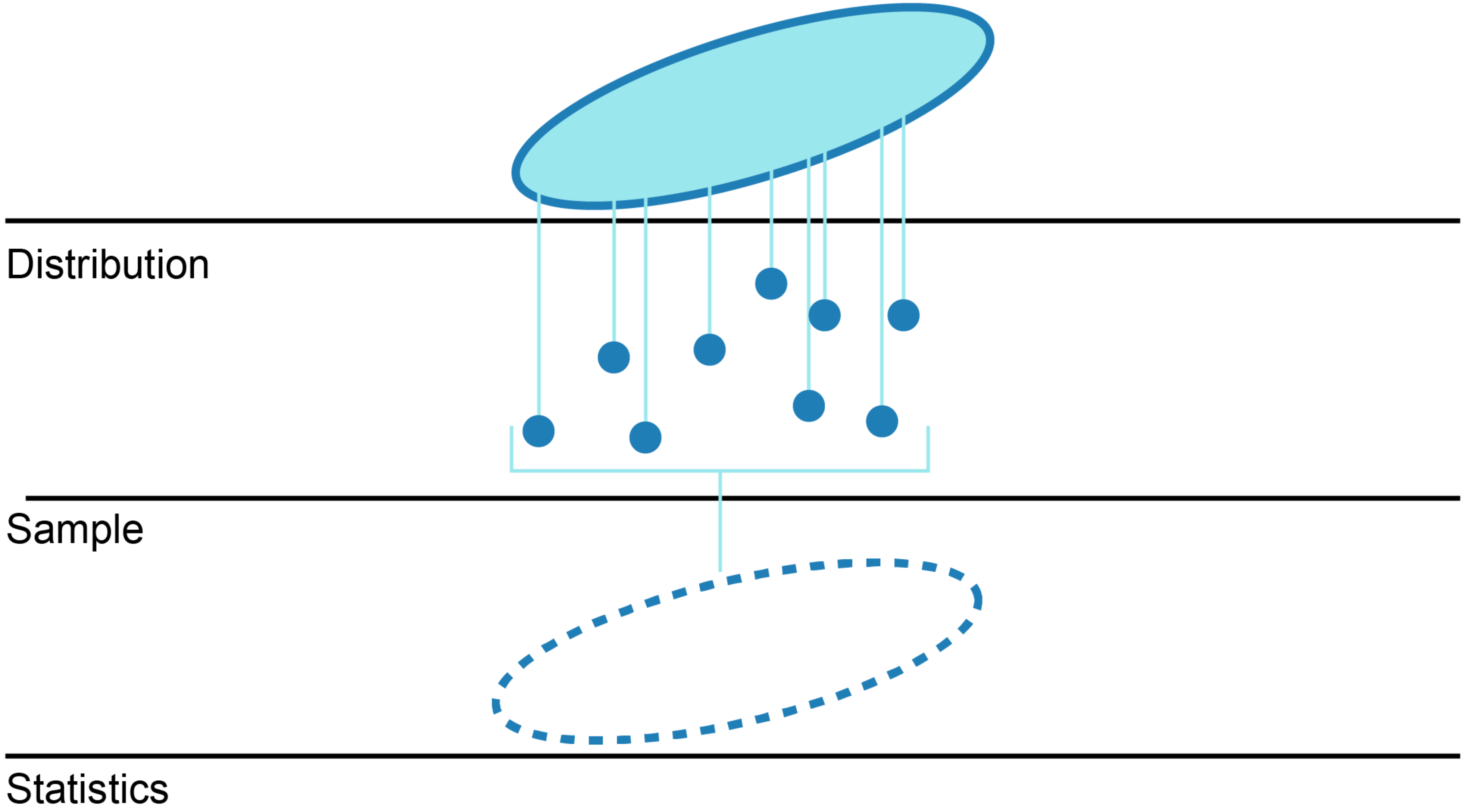
More permutation test applications…



Analyzing neural time series data, Mike X. Cohen

From Lyndon Duong

4) Is my model a good fit?

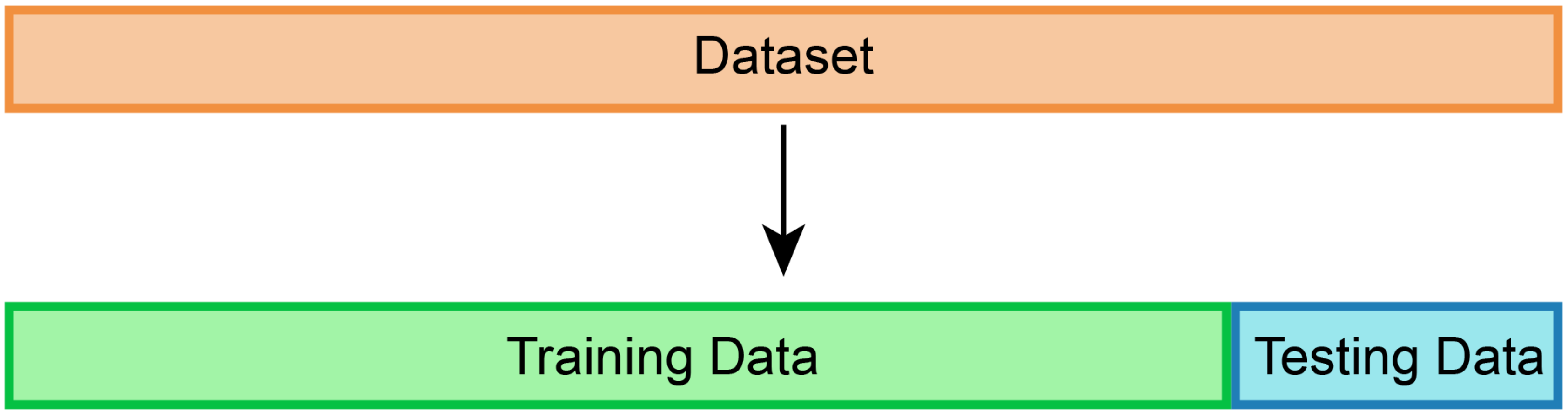


4) Is my model a good fit?

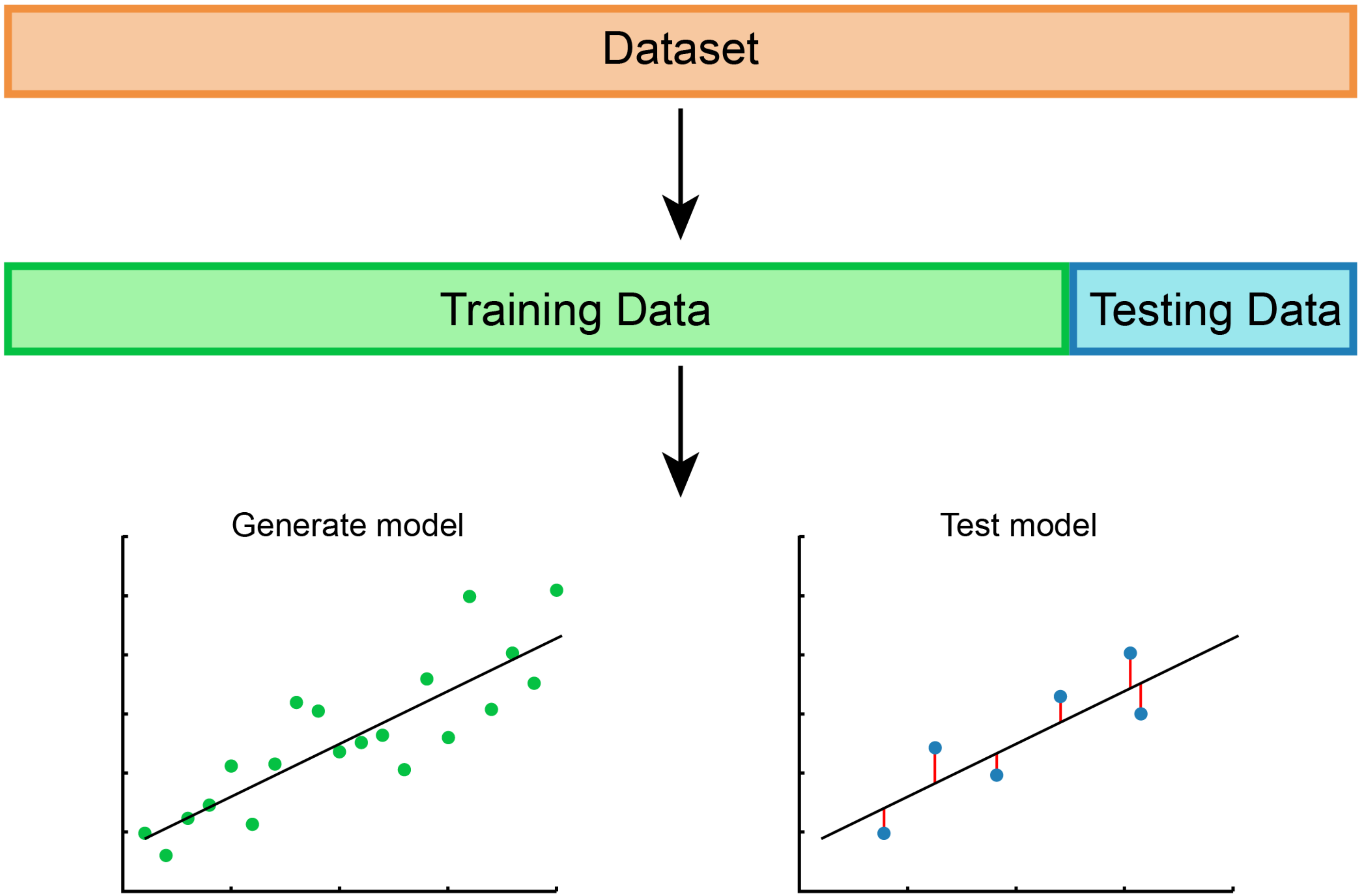
- 99.999% of the time, we only have part of the complete dataset
- We create models with the goal to describe some general underlying process
- The models we fit should also fit out-of-sample data

From Lyndon Duong

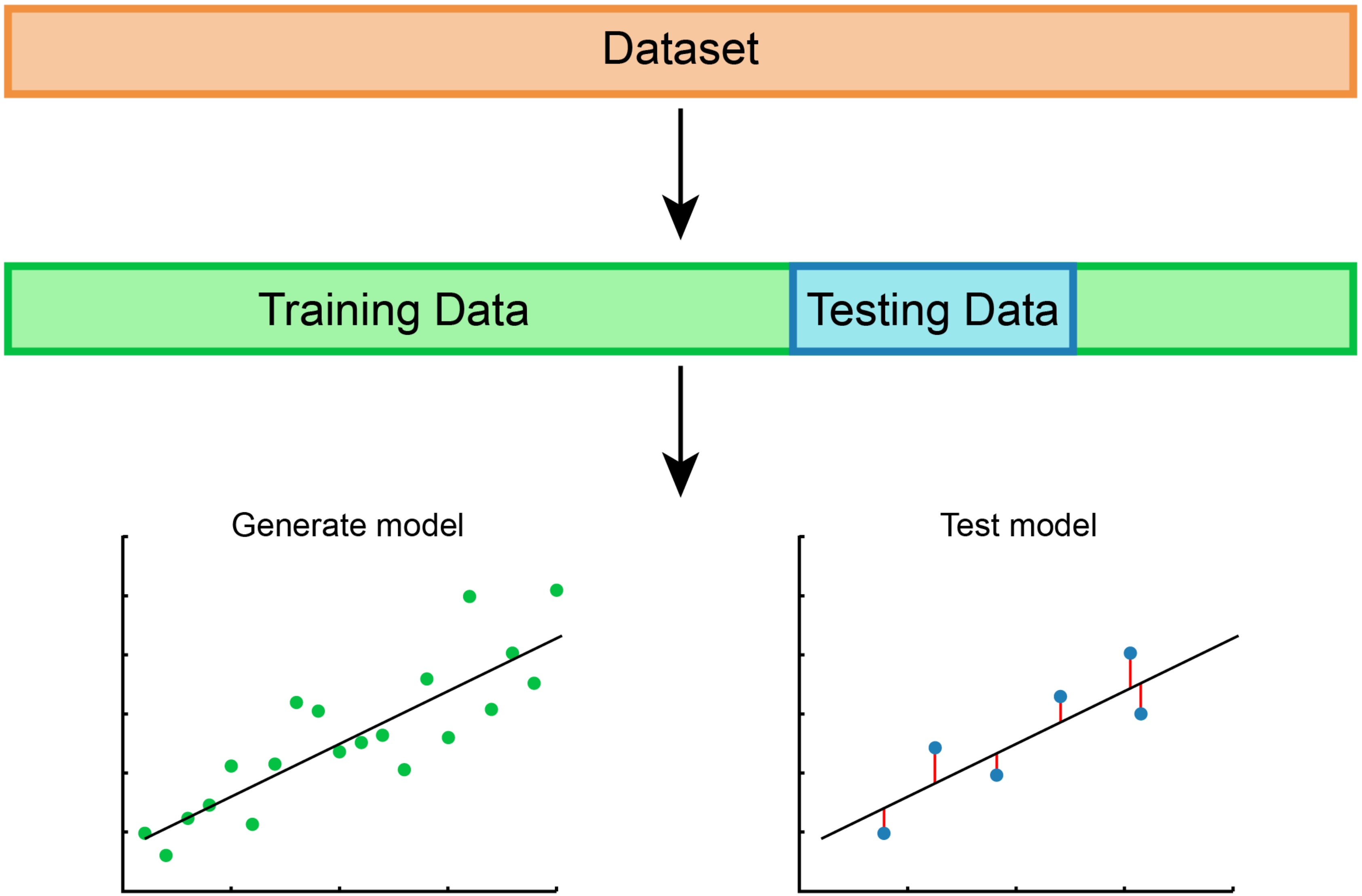
Cross Validation



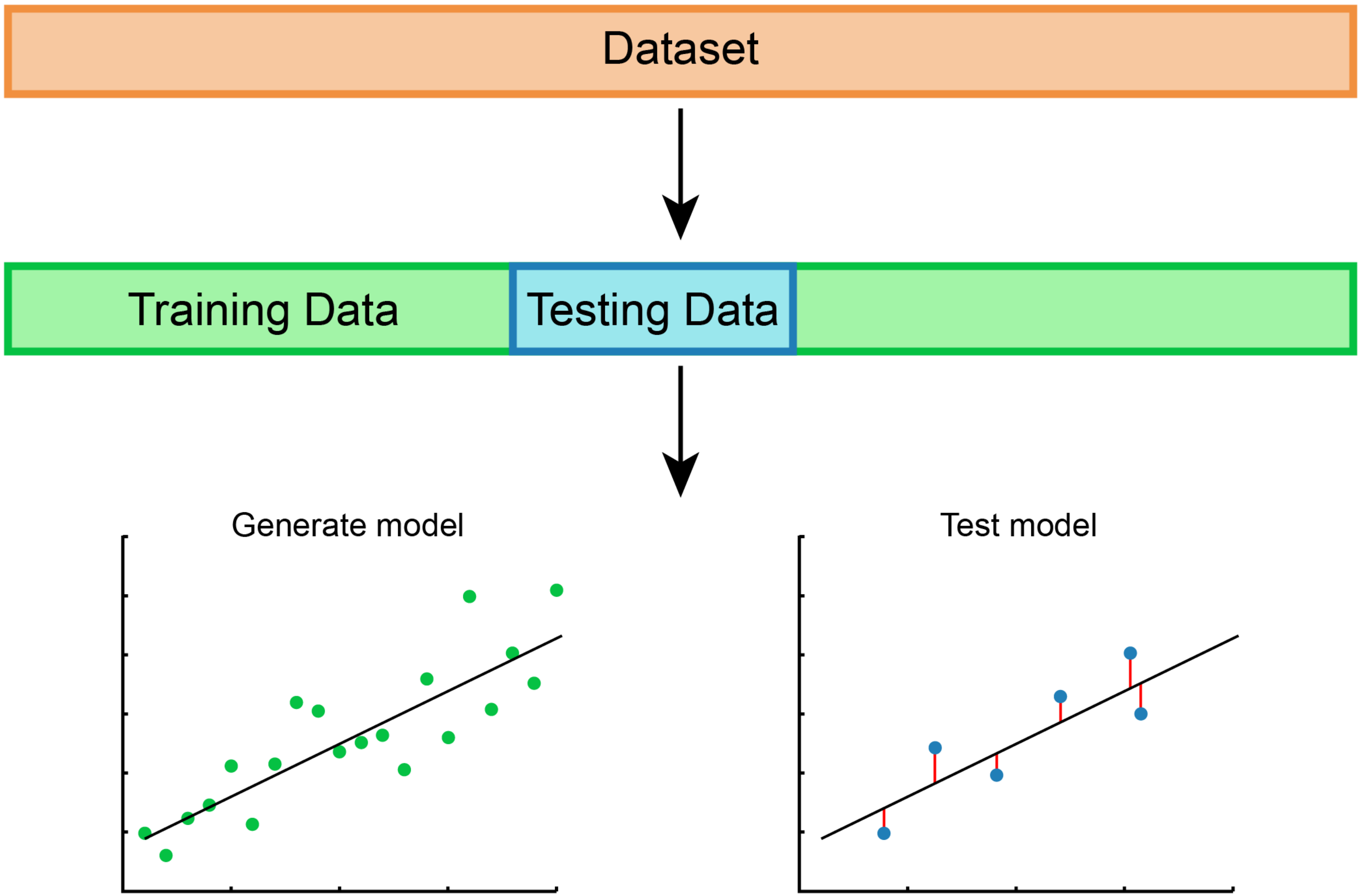
Cross Validation



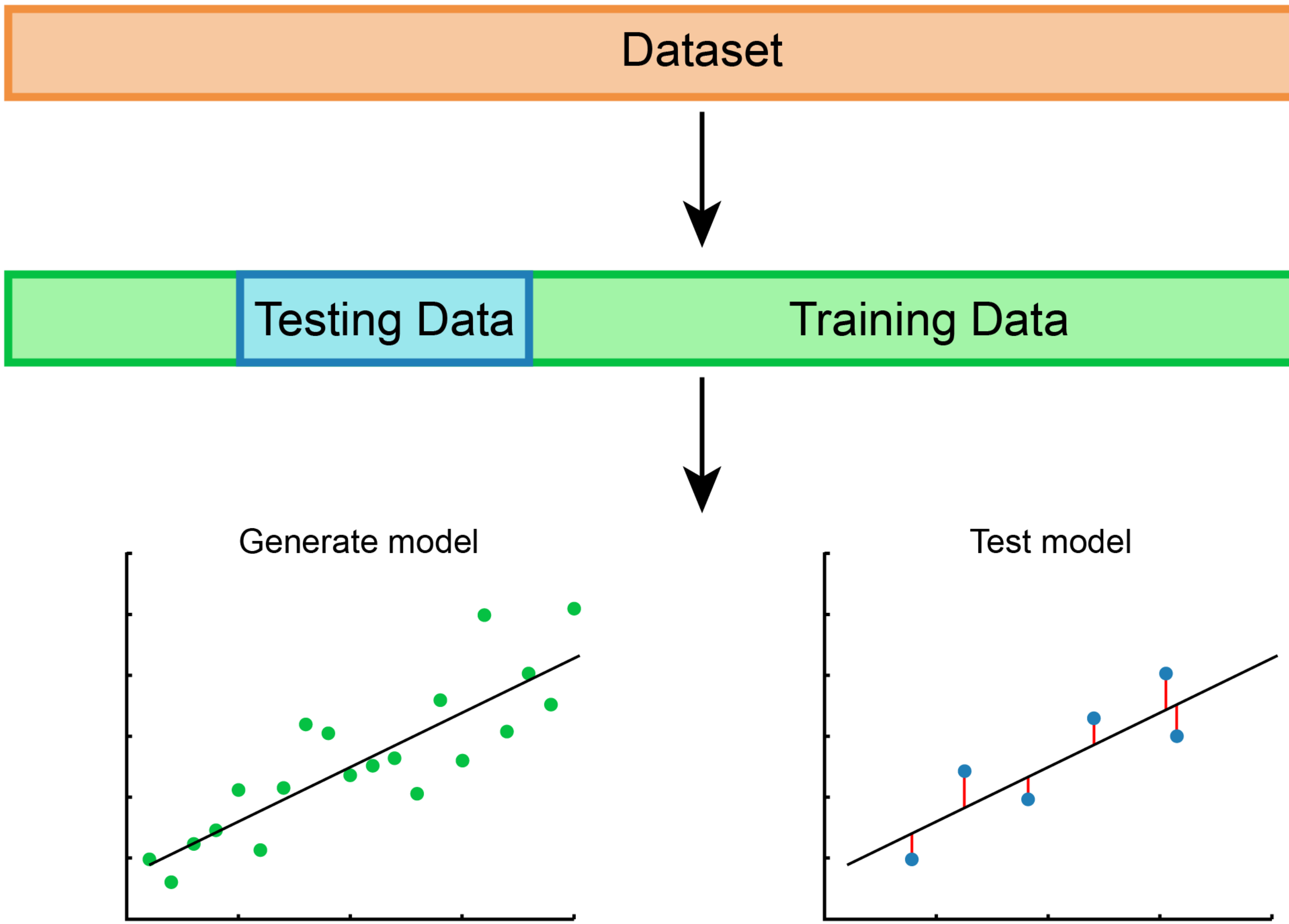
Cross Validation



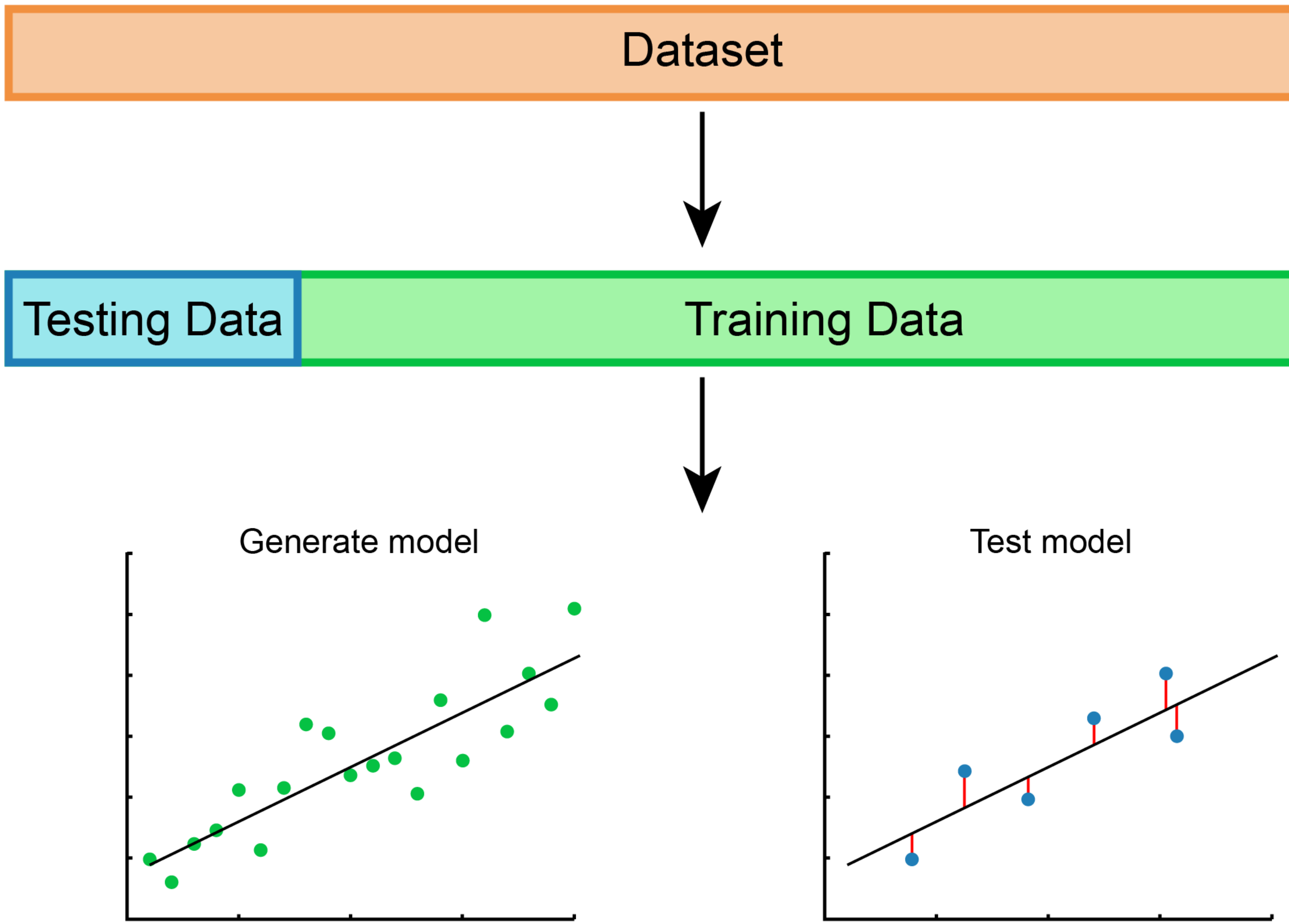
Cross Validation



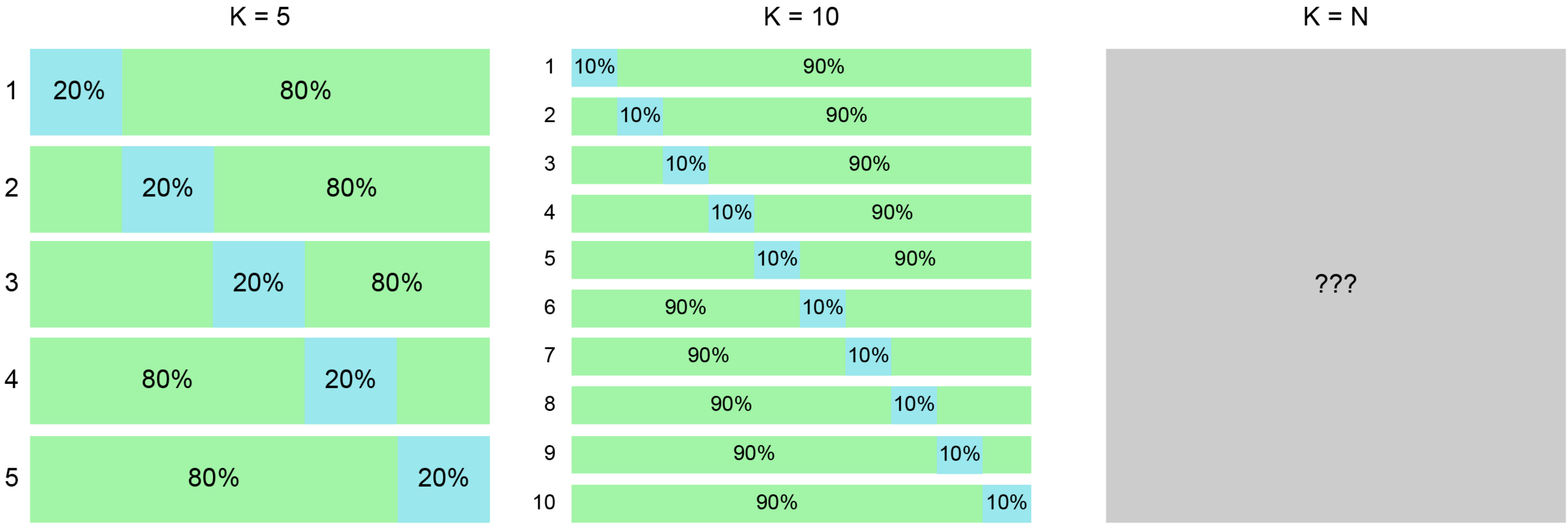
Cross Validation



Cross Validation



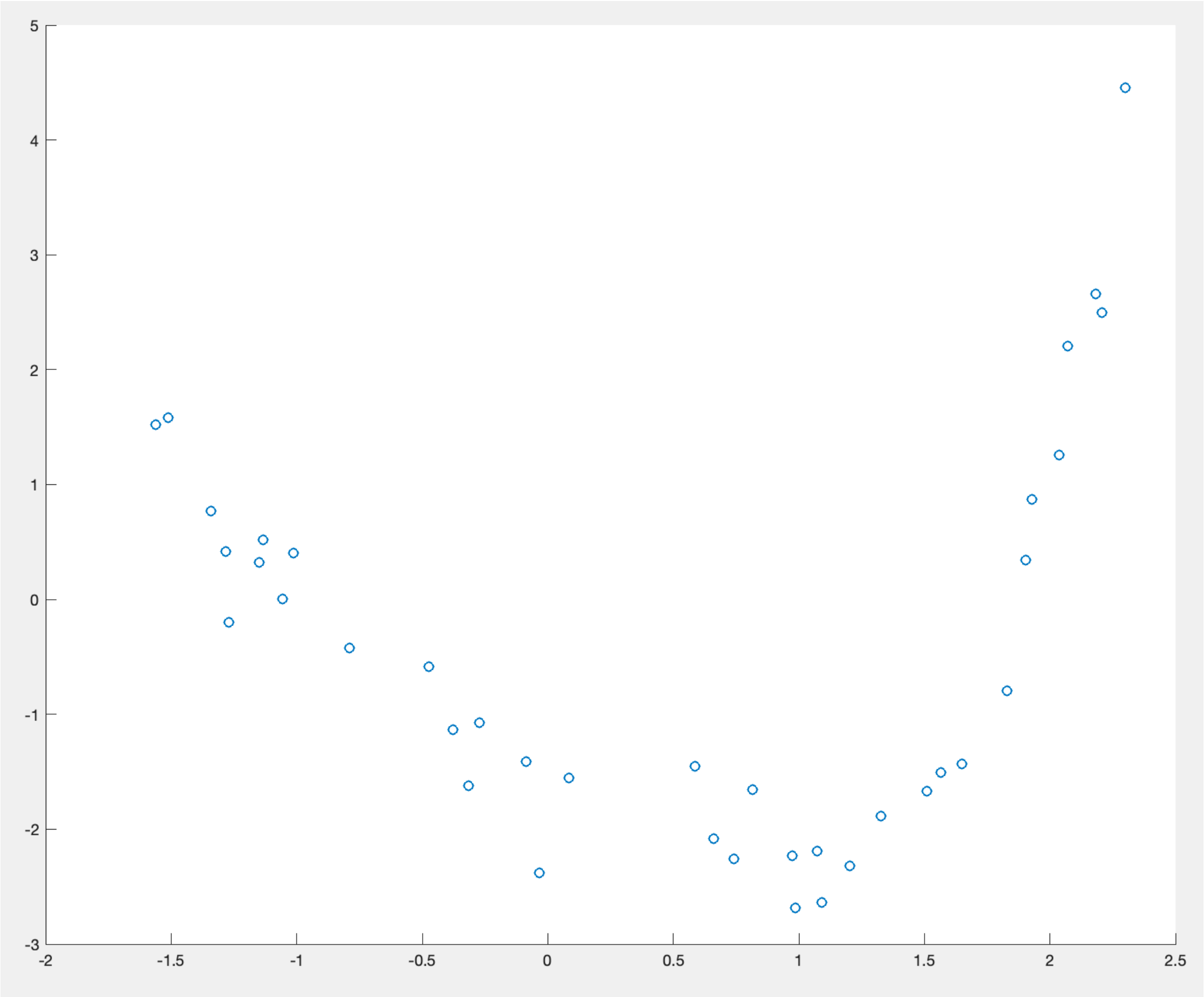
Cross Validation



Leave-one-out cross validation

(K=N)-fold

Exercise 3: Leave-one-out cross validation



Order 0: $\vec{1} * \beta_0$

Order 1: $\vec{1} * \beta_0 + \vec{x} * \beta_1$

Order 2: $\vec{1} * \beta_0 + \vec{x} * \beta_1 + \vec{x}^2 * \beta_2$

Order 3: $\vec{1} * \beta_0 + \vec{x} * \beta_1 + \vec{x}^2 * \beta_2 + \vec{x}^3 * \beta_3$

Order 4: $\vec{1} * \beta_0 + \vec{x} * \beta_1 + \vec{x}^2 * \beta_2 + \vec{x}^3 * \beta_3 + \vec{x}^4 * \beta_4$

Order 5: $\vec{1} * \beta_0 + \vec{x} * \beta_1 + \vec{x}^2 * \beta_2 + \vec{x}^3 * \beta_3 + \vec{x}^4 * \beta_4 + \vec{x}^5 * \beta_5$

Questions?

Next lab is December 4th (last lab!!)