

# Permutation Hypothesis Tests

Robert Williams

Summer Statistics Workshop

6/18/2020

# Permutation Hypothesis Tests

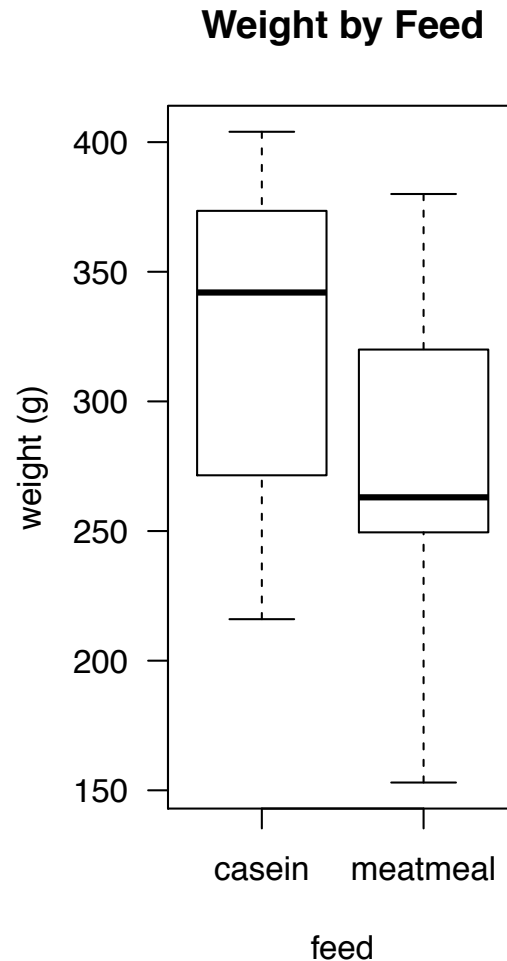
- General concept
  - Determine statistical significance of test statistic between two groups (i.e. the mean value of treatment vs control)
  - Calculate all possible values of the test statistic under all possible rearrangements of the observed data points
- When to use?
  - When you have small sample sizes
  - Test something other than classic approaches comparing means and medians
  - Assumptions of other approaches are not met
- Adapted from MarinStatsLectures YouTube lesson

# Is there a difference in chick weight between casein and meat meal diets?

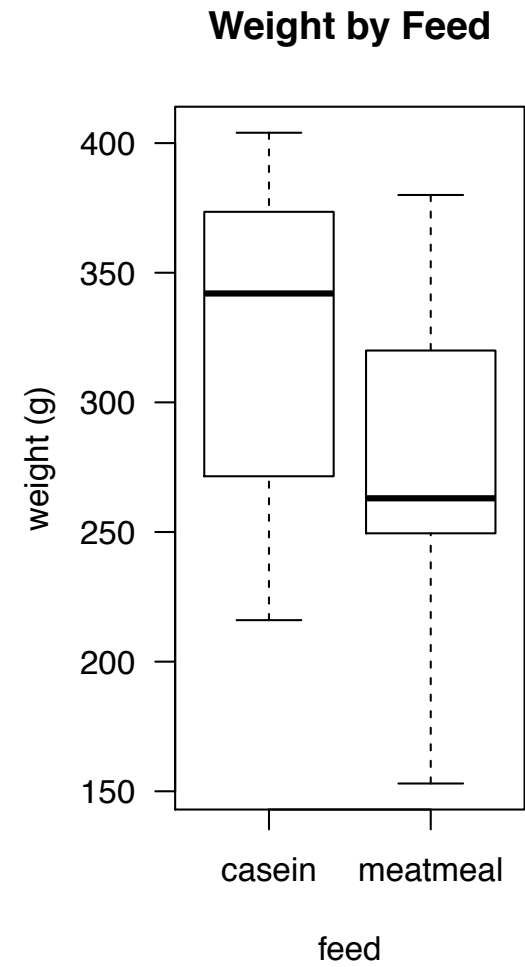
12 Casein Fed Chicks



11 Meat meal Fed Chicks



- Specify  $H_0$ 
  - $H_0 = \text{Weight change same under both feed types}$
  - $|mean(C) - mean(M)| = 0$
- Choose test statistic (TS)
  - $TS = |mean(C) - mean(M)|$
- Determine distribution of test statistic
  - Randomize labels and data points
  - Compute TS
- Convert test statistic to a p-value
  - How many times is the randomized TS greater than the measured TS?



Sample	Observed Mean (g)
casein	323.58
meatmeal	276.90
Test Statistic	
Observed	46.67

Test Statistic =  $|mean(C) - mean(M)|$

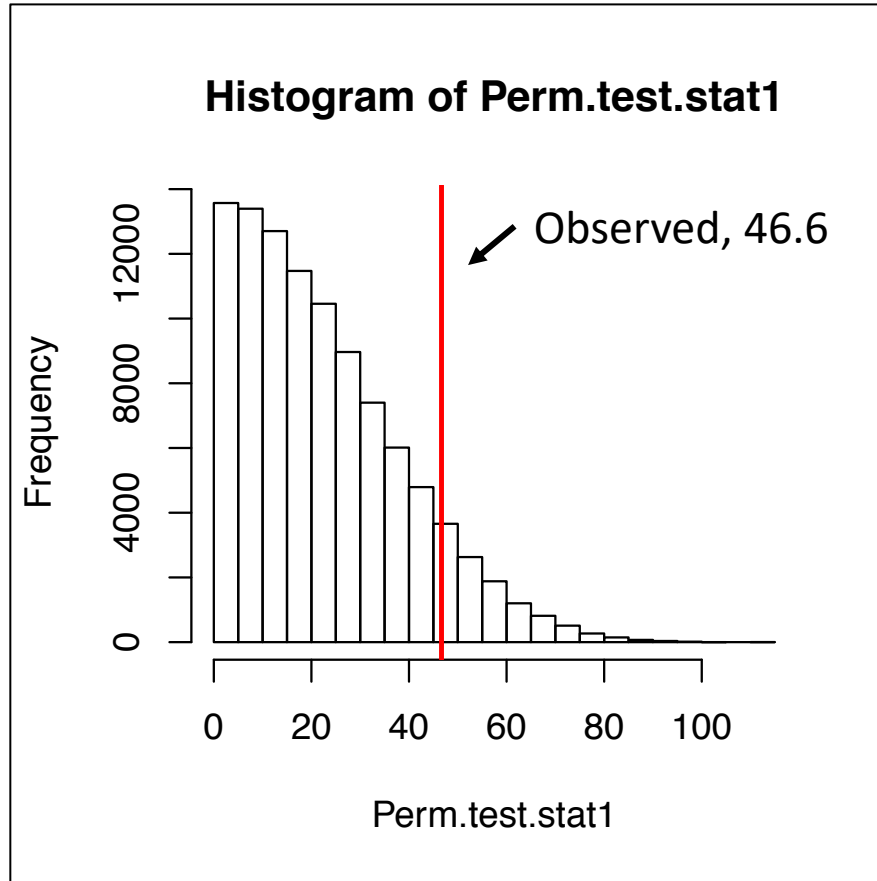
Perform Permutations

	perm1	perm2	perm3	perm4	perm5
meatmeal_1	379	283	380	352	206
meatmeal_2	380	303	258	260	380
meatmeal_3	257	206	379	380	153
meatmeal_4	283	242	222	404	359
meatmeal_5	222	260	325	258	258
meatmeal_6	315	352	153	379	263
meatmeal_7	352	263	263	325	325
meatmeal_8	153	325	315	359	216
meatmeal_9	368	379	344	242	260
meatmeal_10	344	258	368	368	257
meatmeal_11	359	257	206	257	315
casein_1	206	153	404	222	303
casein_2	404	344	303	390	390
casein_3	325	318	318	303	352
casein_4	242	404	332	263	404
casein_5	390	380	257	206	379
casein_6	260	332	216	315	318
casein_7	303	359	352	344	368
casein_8	263	222	242	283	222
casein_9	332	368	260	332	344
casein_10	318	315	283	318	283
casein_11	216	390	390	153	332
casein_12	258	216	359	216	242

... x 10,000

	Perm1	Perm2	Perm3	Perm4	perm5
Mean(M)	310.18	284.36	292.09	325.81	272
Mean(C)	293.0833	316.75	309.6	278.75	328.0833
Test Statistic	17.09	32.3	17.5	47.06	56.08

Calculate Test Statistic for each permutation



$$\text{Test Statistic} = |\text{mean}(C) - \text{mean}(M)|$$

Calculate P-value

$$\text{p-value} = \frac{\# \text{ of times Permutation Test Stat} > \text{Observed Test Stat}}{\text{Total number of Permutations}}$$

$$\text{p-value} = 0.099$$