Creating a randomization Distribution:

**Null Hypothesis and Alternative Hypothesis.**

* Alternative hypothesis (Ha or H1): claim you are testing
* Null hypothesis (Ho): assumption you are beginning with
  + Opposite of what you are testing

**Sampling Distribution**

* A ***sampling distribution*** is a collection of statistics from samples simulated assuming the null hypothesis is true
* The sampling distribution shows what types of statistics would be observed, just by random chance, if the null hypothesis were true

**Let’s look at an example:**

**Paul the Octopus**

**In 2008, Paul the Octopus predicted 8 World Cup games, and predicted them all correctly. Is this evidence that Paul’s chance of guessing correctly, *p*, is really greater than 50%? The observed result is calculated from the sample.**

* **To create the randomization distribution (Null model)**
  + **Need to find the null hypothesis**
    - **If Paul is just guessing we would expect him to get half right or a proportion of 0.5 right.**
    - **The null hypothesis is Ho: p=0.5**

***Simulate many samples of size n = 8 with p = 0.5 (probability ½)***

* ***We can simulate this with a spinner! (Heads = correct, Tails = incorrect)***
* ***Spin 8 times, count the number of successes, and calculate the sample proportion of successes***
  + ***This is one sample based on the null hypothesis***
* ***Repeat this at least 1000 times***

***Show example with Web App using Paul’s info:***

***One dot on the plot is created by sampling with replacement 8 times from 8 cards where ½ are labeled as right and ½ as wrong. The proportion of right is calculated. The proportion of right cards would be plotted on the randomization distribution. To create the randomization distribution this would be repeated at least 1000x.***