Do We Look Like Our Dogs?

Some people believe that dog owners tend to look like their dogs. To investigate this claim, a group of researchers photographed a number of pure bred dogs and their owners. A group of participants were shown a picture of an owner and pictures of two different dogs (one was the owner's dog and the other was a different randomly selected dog).

1. If 25 participants were shown pictures of one owner and a pair of dogs, how many of the participants do you think were able to pair the owner to their dog correctly? Why?

2. How many different possible hypotheses could we make for this situation regarding the photos of dogs and their owners? What are they?

3. a) If participants could not really tell which dog went with which owner, what would be the most likely outcome (number of correct paired photos) when this study is conducted with 25 participants?

   (b) Still assuming participants could not tell which dog went with which owner (i.e. participants were randomly guessing) what kind of results (for number of correctly guessed pairs) would you not be surprised to see when this study is conducted with 25 participants?

4. The actual number of the 25 participants in the study who correctly matched was _____. If it is REALLY the case that the participants were randomly guessing, do you find the researchers' results surprising? Why or why not?

The key question is, "How surprising is the observed result under the assumption that participants could not really pair the photos correctly (i.e. they were randomly guessing)?" We will call this assumption of randomly guessing the **null hypothesis**. Let's simulate this situation.

5. Design a simulation to represent this experiment assuming that the null hypothesis is true. Carry out three trials of the simulation and record your results below. Be sure to get approval from your teacher before carrying out the simulation.

_____ _____ _____

Combine your results with your classmates. Do this by producing a dotplot (of the number of correctly paired owner and dog) on the board, where you contribute a total of three dots corresponding to your 3 simulations from part 5.

6. a) Does it seem like the results actually obtained by these researchers would be surprising under the null hypothesis that participants could not match the owner and dog correctly and were randomly guessing? Explain.

   b) Now, we will use technology to simulate completing this experiment many, many times under the assumption that the null hypothesis is true. Based on this simulation how surprising are the actual results of this study? Explain your reasoning.

There is **variability** between what we expected to happen assuming the null hypothesis is true (12-13 owners paired correctly) and the actual results. The question is, "Is it reasonable for us to say that as participants selected which dog was the owner's dog, random chance alone can explain obtaining a result of 16 rather than 12 or 13?" Many statisticians say that the field of statistics is primarily about explaining variability. This is what we are attempting to do in this investigation.

Terminology:

The ***probability*** of an event is the long-run proportion of times the event happens when its random process is repeatedly indefinitely.

The ***p-value*** is the probability that randomness alone would produce data as extreme (or more extreme) as the result obtained in the actual study, assuming the null hypothesis to be true.

A small p-value indicates that the observed data would be surprising to occur by randomness alone, if the null hypothesis were true. Such a result is said to be ***statistically significant***, and provides evidence against the null.

c) Based on our above simulations, what conclusion should the researchers draw? Justify your conclusions and use the above terminology in your justification (use all 3 bold terms!).

d) If the actual study had instead found that 22 of the 25 participants paired the owners correctly with their dog, then what decision should the researchers make based on this result? Justify your conclusions and use the above terminology in your justification.