

Course Notes: Statistics 1401

Notebook 10: Probability

UNG Mathematics

October 2024

Please open **Notebook 10** which is called:

Probability

We begin some basic rules of probability which is likely review for you. Then we discuss the difference between **Theoretical** vs **Empirical** probabilities.

Empirical probabilities are quite easy when we have a machine that can run a simulation ten thousand times (or ten million times) in less than a second.

Key new material: **.sample** method

1 Examples

First Set of Examples

We use the **.sample** method to roll dice. We can calculate probabilities using this simulation.

Second Set of Examples

We use the **stats.binom.pmf** function to calculate probabilities.

2 Python and Empirical Probabilities

Probability Calculations. We have a tool that will calculate exact probabilities for us:

scipy.stats.binom.pmf

In our initial code block, we used **import scipy.stats as stats**.

Thus, the code we can use to call the above function work is:

stats.binom.pmf

We need to enter 3 values:

1st Number of **Successes**

2nd Number of Trials

3rd Probability of Success

3 The first Hypothesis Test

In the classic story of the **Lady Tasting Tea**, we find that Ronald Fisher's acquaintance claimed she could tell, just by tasting, whether the tea or the milk had been added to the cup first. Dr. Muriel Bristol claimed tea tasted better if the milk was added first. Fischer, the most influential pioneer of statistical inquiry, proposed a blind taste test. Dr. Bristol would be served 8 cups of tea, each randomly chosen to have either milk or tea added first. If she identified 7 or more correctly, her claim would be established beyond question.

Fischer's grand idea was that of a **null hypothesis**, a claim that could *not* be established but only disproved using probability. In this experiment, the null hypothesis was that Dr. Bristol was **guessing at random** and thus had a 50% chance of success at each guess.

3.1 The Lady Tasting Tea: Questions

Using Python, we can determine some probabilities. Calculate those probabilities and determine whether:

- If Dr. Bristol correctly identified 7 or more cups of tea correctly, would you believe her claim?
- What if she identified all 8 correctly?
- What if she only identified 6 of 8 correctly?

3.2 Conclusion

Every statistics-based research study has a hypothesis (Lady Tasting Tea example shown below):

$$H_0 : \pi = \frac{1}{2}$$
$$H_a : \pi > \frac{1}{2}$$

Why do we study probability in statistics?

Each hypothesis test calculates a p -value which is a probability. Based upon this p -value, we make a scientific determination whether our hypothesis was true or false. When

$$p < \alpha = 0.05$$

our conclusion is that we **REJECT THE NULL**. We reject the null because, based upon the probability, it is quite likely to be false.