# PSY 201: Statistics in Psychology

## Lecture 16
## Underlying distributions
## *Can you read my mind?*

Greg Francis

Purdue University

Fall 2023

# DISTRIBUTION

- representation of all possible outcomes
- area under the curve represents relative frequency of events
- completely describes an aspect of a situation relative to a particular variable
- often theoretical curves (but not always)

# DICE ROLES

| Die 1 | Die 2 | Sum | Difference |
|---|---|---|---|
| 1 | 1 | 2 | 0 |
| 1 | 2 | 3 | 1 |
| 1 | 3 | 4 | 2 |
| 1 | 4 | 5 | 3 |
| 1 | 5 | 6 | 4 |
| 1 | 6 | 7 | 5 |
| 2 | 1 | 3 | 1 |
| 2 | 2 | 4 | 0 |
| 2 | 3 | 5 | 1 |
| 2 | 4 | 6 | 2 |
| 2 | 5 | 7 | 3 |
| 2 | 6 | 8 | 4 |
| 3 | 1 | 4 | 2 |
| 3 | 2 | 5 | 1 |
| 3 | 3 | 6 | 0 |
| 3 | 4 | 7 | 1 |
| 3 | 5 | 8 | 2 |
| 3 | 6 | 9 | 3 |
| 4 | 1 | 5 | 3 |
| 4 | 2 | 6 | 2 |
| 4 | 3 | 7 | 1 |
| 4 | 4 | 8 | 0 |
| 4 | 5 | 9 | 1 |
| 4 | 6 | 10 | 2 |
| 5 | 1 | 6 | 4 |
| 5 | 2 | 7 | 3 |
| 5 | 3 | 8 | 2 |
| 5 | 4 | 9 | 1 |
| 5 | 5 | 10 | 0 |
| 5 | 6 | 11 | 1 |
| 6 | 1 | 7 | 5 |
| 6 | 2 | 8 | 4 |
| 6 | 3 | 9 | 3 |
| 6 | 4 | 10 | 2 |
| 6 | 5 | 11 | 1 |
| 6 | 6 | 12 | 0 |

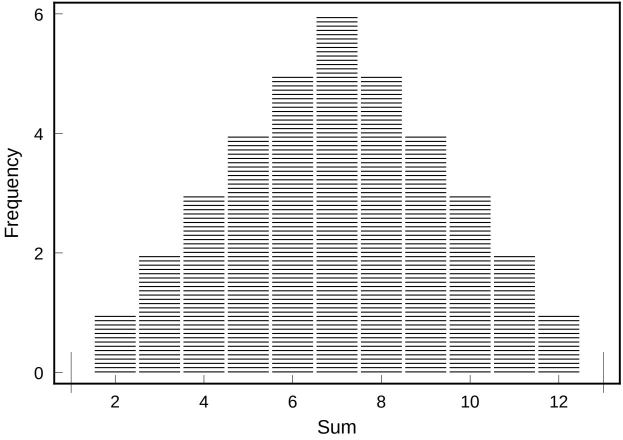
# DISTRIBUTION

- we can identify the underlying distribution of the sum of dice variable

| Sum | $f$ |
|---|---|
| 1 | 0 |
| 2 | 1 |
| 3 | 2 |
| 4 | 3 |
| 5 | 4 |
| 6 | 5 |
| 7 | 6 |
| 8 | 5 |
| 9 | 4 |
| 10 | 3 |
| 11 | 2 |
| 12 | 1 |
| 13 | 0 |

# DISTRIBUTION

- same type of stuff we did earlier




- frequency of every possible outcome of the variable Sum

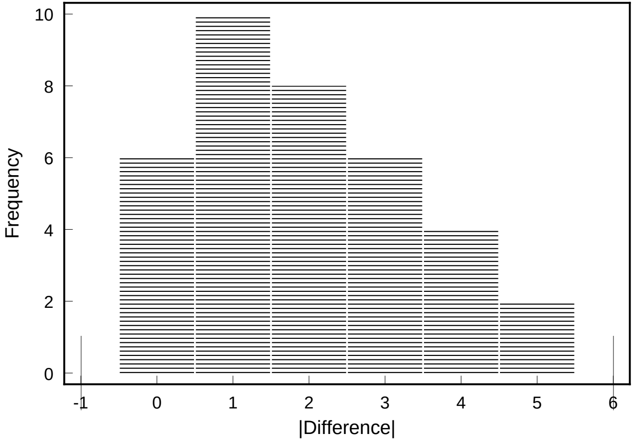
# VARIABLE

- a distribution is specific to a variable ($x$-coordinate)
- suppose instead of the sum of dice roles, we look at the distribution of the absolute value of the difference of dice roles

| $\|$Difference$\|$ | $f$ |
|---|---|
| 0 | 6 |
| 1 | 10 |
| 2 | 8 |
| 3 | 6 |
| 4 | 4 |
| 5 | 2 |
| 6 | 0 |

# DISTRIBUTION

- the underlying distribution is different because we are considering a different variable



# USE

- once we have the underlying distribution we can calculate probabilities

$$P(A) = \frac{\text{Number of outcomes that include } A}{\text{Total number of possible outcomes}}$$

- you better believe a casino cares about this!
- so does the government
- in practice statisticians generally work with theoretical distributions

# BINOMIAL DISTRIBUTION

- suppose you have a situation where there are only two possible outcomes from an action
- e.g., flip a coin: H or T
- each flip is **independent** of the other flips
- how many H's do you get if you flip the coin over and over (or flip many identical coins at once)?

# BINOMIAL DISTRIBUTION

- suppose you flip the coin twice
- the possible outcomes are

| First coin | Second Coin | Number H |
|---|---|---|
| H | H | 2 |
| H | T | 1 |
| T | H | 1 |
| T | T | 0 |

- can produce a frequency distribution table

| Number H's | $f$ |
|---|---|
| 0 | 1 |
| 1 | 2 |
| 2 | 1 |

# BINOMIAL

- from Webster
  1. a mathematical expression consisting of two terms connected by a plus sign or minus sign
  2. a biological species name consisting of two terms

$$(H+T)$$

- to find out how many H's and how many T's, for two coin flips, square the binomial

$$(H+T)^2 = H^2 + 2HT + T^2$$

- or

$$(H+T)^2 = HH + 2HT + TT$$

- coefficient in front identifies how many of each combination

# BINOMIAL DISTRIBUTION

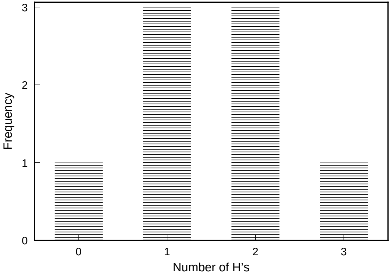
- suppose you flip the coin thrice
- the possible outcomes are

| First coin | Second Coin | Third Coin | Number H |
|---|---|---|---|
| H | H | H | 3 |
| H | T | H | 2 |
| T | H | H | 2 |
| T | T | H | 1 |
| H | H | T | 2 |
| H | T | T | 1 |
| T | H | T | 1 |
| T | T | T | 0 |

# BINOMIAL DISTRIBUTION

- can produce a frequency distribution table

| Number H's | $f$ |
|---|---|
| 0 | 1 |
| 1 | 3 |
| 2 | 3 |
| 3 | 1 |



# BINOMIAL

- for three flips, cube the binomial

$$(H+T)^3 = H^3 + 3H^2T + 3HT^2 + T^3$$

- or

$$(H+T)^3 = HHH + 3HHT + 3HTT + TTT$$

- coefficient of each term indicates number of occurrences!
- this approach works in general (combinatorics)

# BINOMIAL DISTRIBUTION

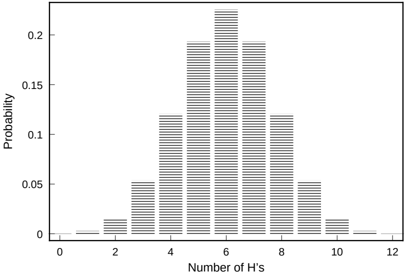
- in turns out that for $m$ flips of the coin the **probability** of getting $x$ number of H's is

$$P(x \text{ number of H's}) = \frac{m!}{x!(m-x)!}(0.5)^m$$

- where $x! = (x)(x-1)(x-2)...(2)(1)$
  is "$x$-factorial"
  (don't worry about it)
- works for probabilities other than 0.5 too (slightly more complicated)
- Your textbook provides an on-line calculator for any probability

# BINOMIAL AND NORMAL

- for $m = 12$



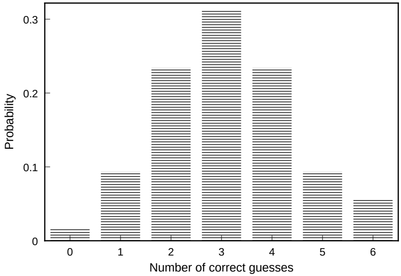

- looks a lot like a normal distribution
- for $m > 20$, the difference is very small

# USE

- suppose you have a friend who always drinks Sprite, claiming it is better than 7-Up
- you test your friend's ability to distinguish between Sprite and 7-Up
- Your friend sips two glasses of soda, one containing Sprite and the other 7-Up. Your friend must decide which is the Sprite. You do this 6 times. (Glasses are identical, randomized for tasting first,...)
- Your friend identifies the glass containing Sprite every time. Now you need to decide if your friend really knows his stuff or is just lucky.

# USE

- You need to know the probability of **guessing** 6 correct identifications out of 6 trials
- binomial distribution gives us exactly what we want to know



# USE

- the probability of guessing correctly 6 out of 6 times is very small (0.0156)
- most likely your friend can tell Sprite from 7-Up
- we will be using distributions like this a lot!
- compare performance to guessing performance
- performance we get from experimentation
- guessing performance we get through tables and calculations (can be complicated)

# MIND READING?

- Suppose I flip a coin and look at the upward side.
- Can people read my mind?
- Suppose we took 10 people and asked them to guess which side I saw.
- Some will guess correctly, just by luck.
- How often must people guess correctly before we decide they can read my mind?

# MIND READING?

- Each person guessing has a 1 in 2 chance of being correct. So if each person was guessing, how many would we expect to guess correctly?
- What is the probability for each number of guessing correctly?: its a binomial distribution
- can produce a table of probabilities
- It would be surprising (rare) if people were correct 8, 9, or 10 times out of 10.

| Number correct | $p$ |
|---|---|
| 0 | 0.0010 |
| 1 | 0.0098 |
| 2 | 0.0439 |
| 3 | 0.1172 |
| 4 | 0.2051 |
| 5 | 0.2461 |
| 6 | 0.2051 |
| 7 | 0.1172 |
| 8 | 0.0439 |
| 9 | 0.0098 |
| 10 | 0.0010 |

# MIND READING?

- Let's see if people can read my mind:
- measure ability to read my mind
- get number correct
- see if it is "rare enough" for us to conclude they can read minds
  - By using the on-line Binomial distribution calculator

# CONCLUSIONS

- underlying distributions
- binomial distribution
- started hypothesis testing

# NEXT TIME

- sampling distribution of the mean
- properties of sampling distributions

*Marvel at my predictive powers!*