Probability

PSYC 573

University of Southern California
January 18, 2022

History of Probability

- A mathematical way to study uncertainty/randomness
- Origin: To study gambling problems

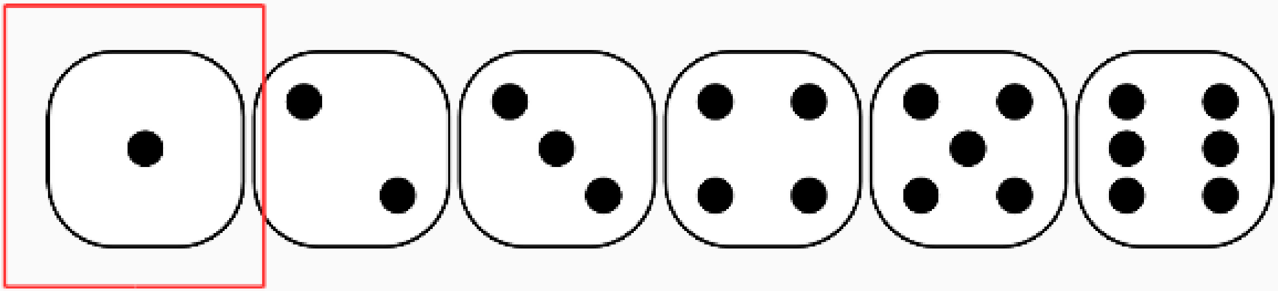
> Someone asks you to play a game. The person will flip a coin. You win $10 if it shows head, and lose $10 if it shows tail. Would you play?

Kolmogorov Axioms

For an event A_i (e.g., getting a "1" from throwing a die)

- $P(A_i) \geq 0$ [All probabilities are non-negative]
- $P(A_1 \cup A_2 \cup \cdots) = 1$ [Union of all possibilities is 1]
- $P(A_1) + P(A_2) = P(A_1 \textbf{ or } A_2)$ for mutually exclusive A_1 and A_2 [Addition rule]

Throwing a Die With Six Faces



A_1 = getting a one, . . . A_6 = getting a six

- $P(A_i) \geq 0$
- $P(\text{the number is 1, 2, 3, 4, 5, or 6}) = 1$
- $P(\text{the number is 1 or 2}) = P(A_1) + P(A_2)$

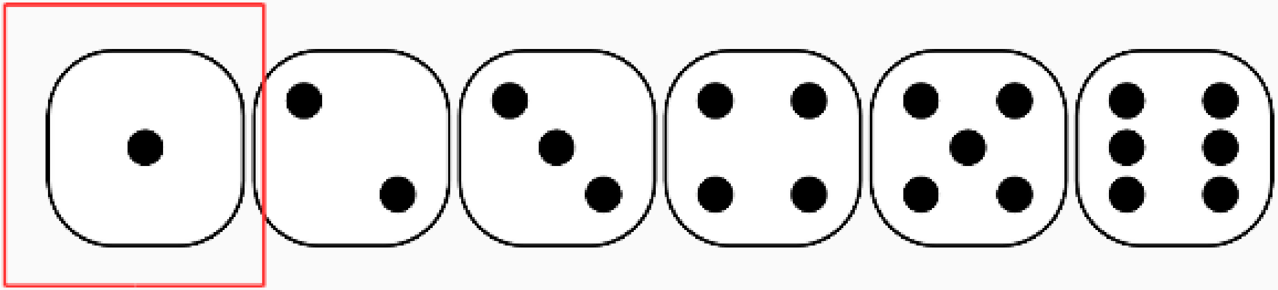
Interpretations of Probability

Ways to Interpret Probability

- **Classical:** Counting rules
- **Frequentist:** long-run relative frequency
- **Subjectivist:** Rational belief

Note: there are other paradigms to interpret probability. See https://plato.stanford.edu/entries/probability-interpret/

Classical Interpretation

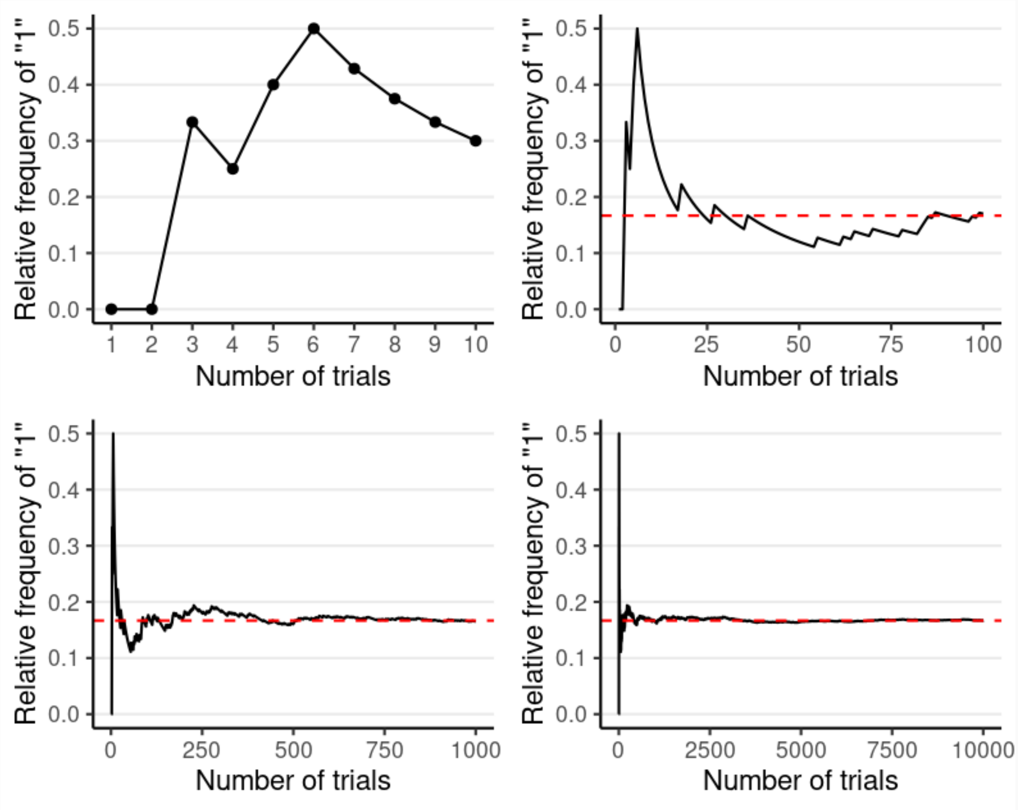


- Number of target outcomes / Number of possible "indifferent" outcomes
  - E.g., Probability of getting "1" when throwing a die: 1 / 6

Frequentist Interpretation

- Long-run relative frequency of an outcome

Trial	Outcome
1	2
2	3
3	1
4	3
5	1
6	1
7	5
8	6
9	3
10	3



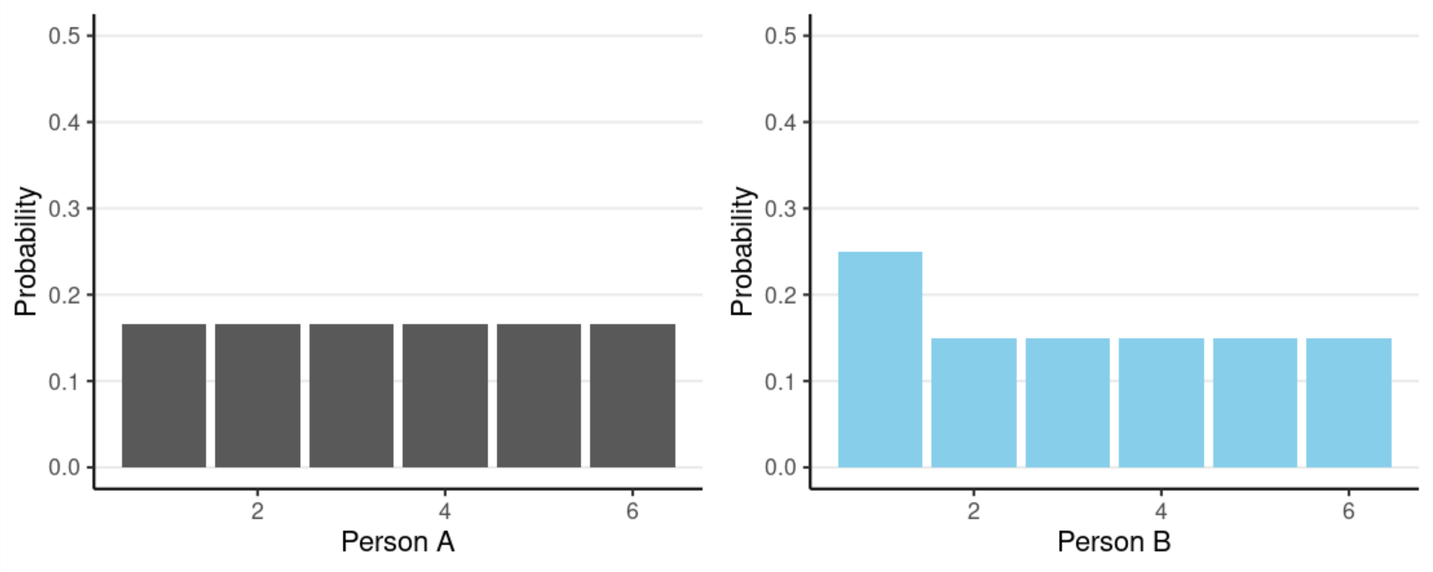
Problem of the single case: Some events cannot be repeated

- Probability of Democrats/Republicans "winning" the 2022 election
- Probability of the LA Rams winning the 2022 Super Bowl
- Probability that the null hypothesis is true

Frequentist: probability is not meaningful for these

Subjectivist Interpretation

- State of one's mind; the belief of all outcomes
  - Subjected to the constraints of:
    - Axioms of probability
    - That the person possessing the belief is rational



Describing a Subjective Belief

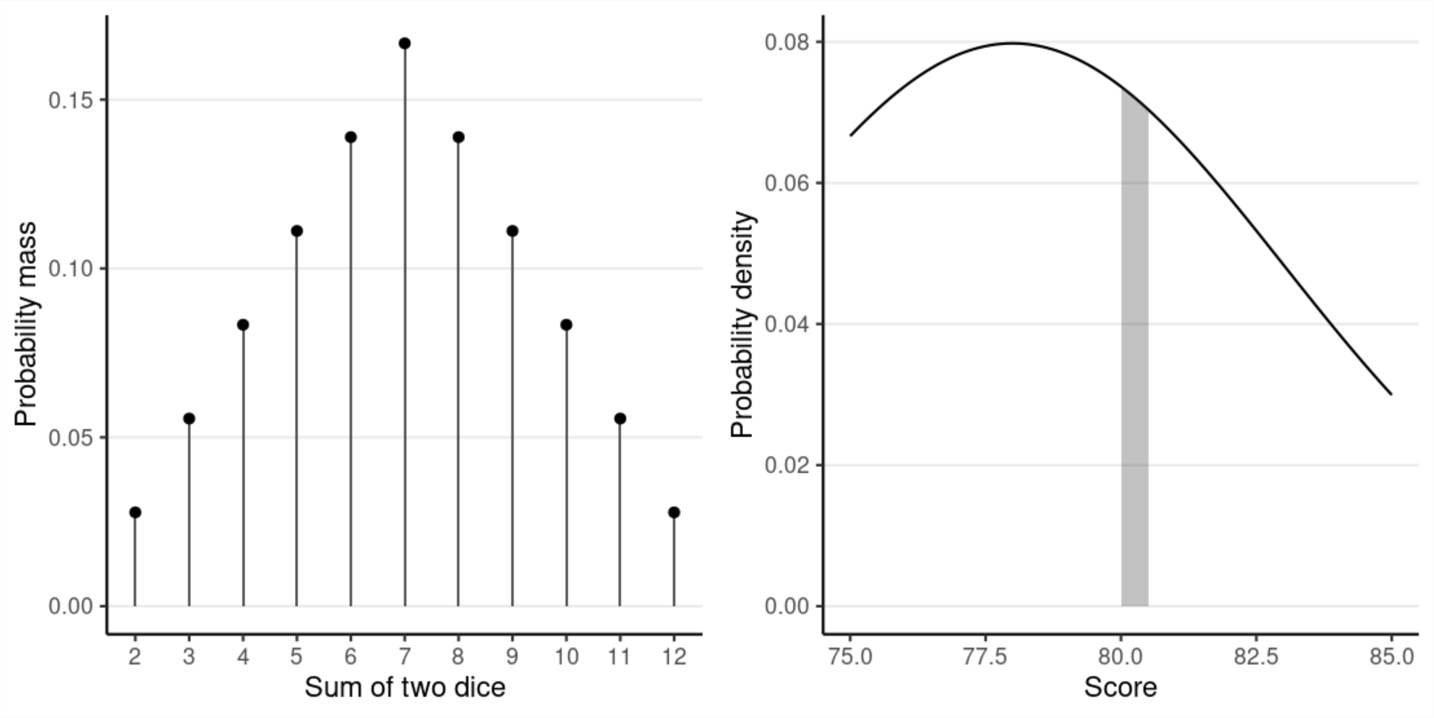
- Assign a value for every possible outcome
    - Not an easy task
- Use a *probability distribution* to approximate the belief
    - Usually by following some conventions
    - Some distributions preferred for computational efficiency

Key to forming *prior* distributions

Probability Distribution

Probability Distributions

- Discrete outcome: Probability **mass**
- Continuous outcome: Probability **density**



Probability Density

- If X is continuous, the probability of X having any particular value $\to 0$
  - E.g., probability a person's height is 174.3689 cm

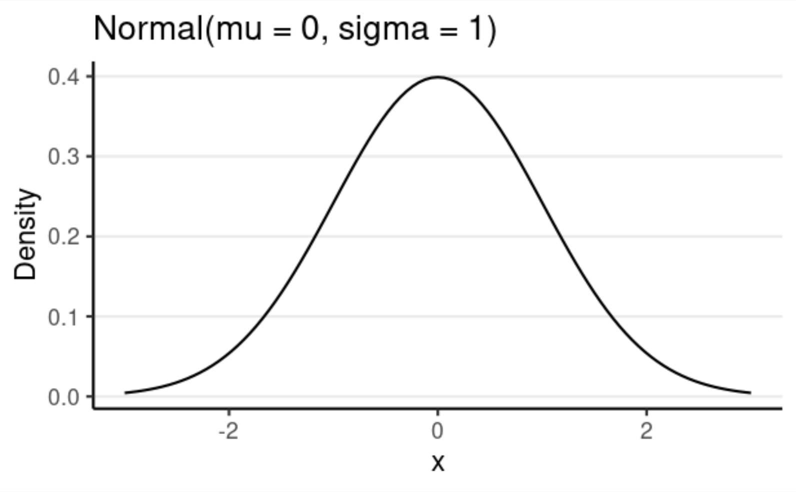
Density:

$$P(x_0) = \lim_{\Delta x \to 0} \frac{P(x_0 < X < x_0 + \Delta x)}{\Delta x}$$

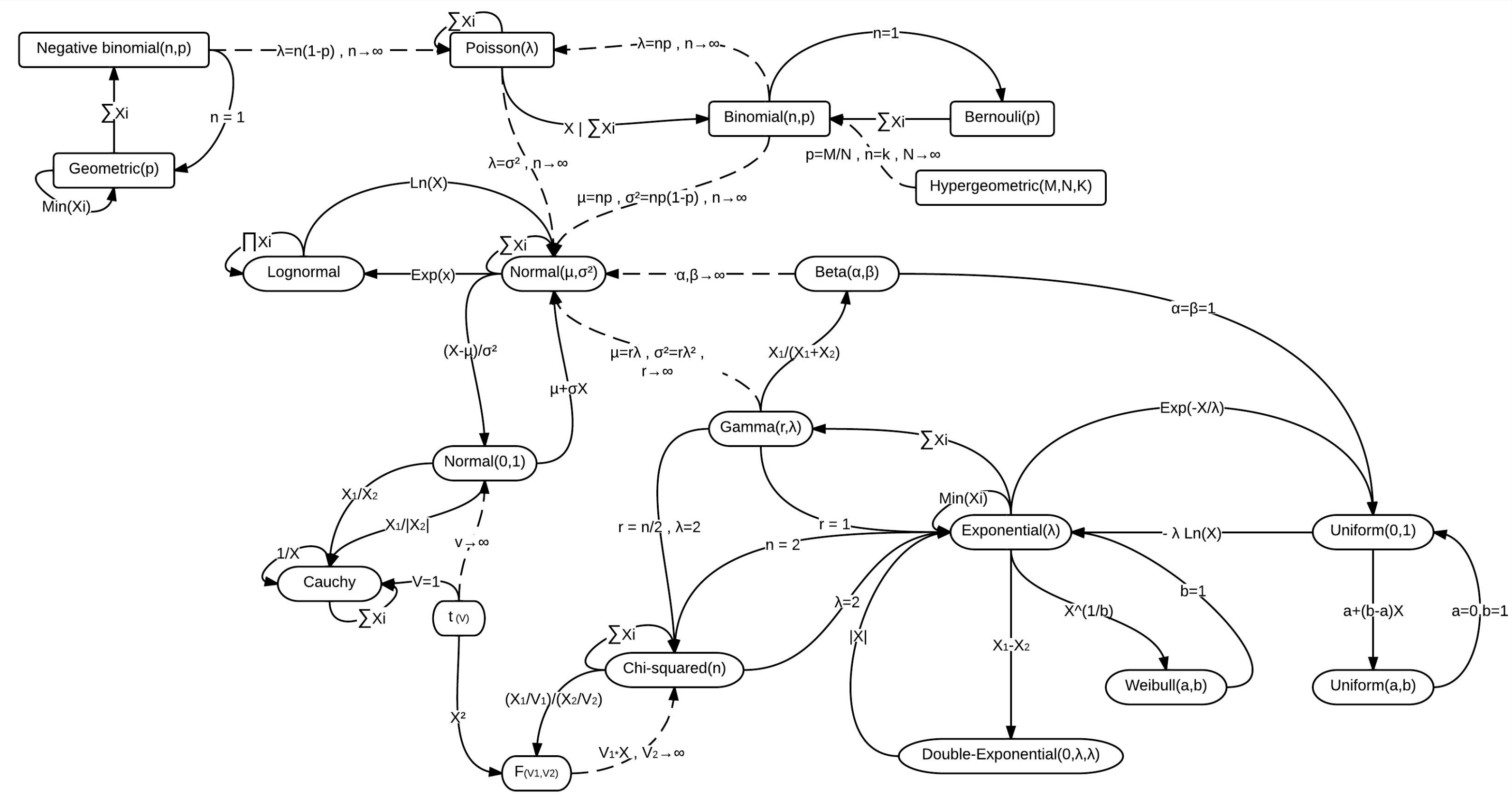
Normal Probability Density

Math   R Code

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right)$$



Some Commonly Used Distributions



Summarizing a Probability Distribution

Central tendency

The center is usually the region of values with high plausibility

- Mean, median, mode

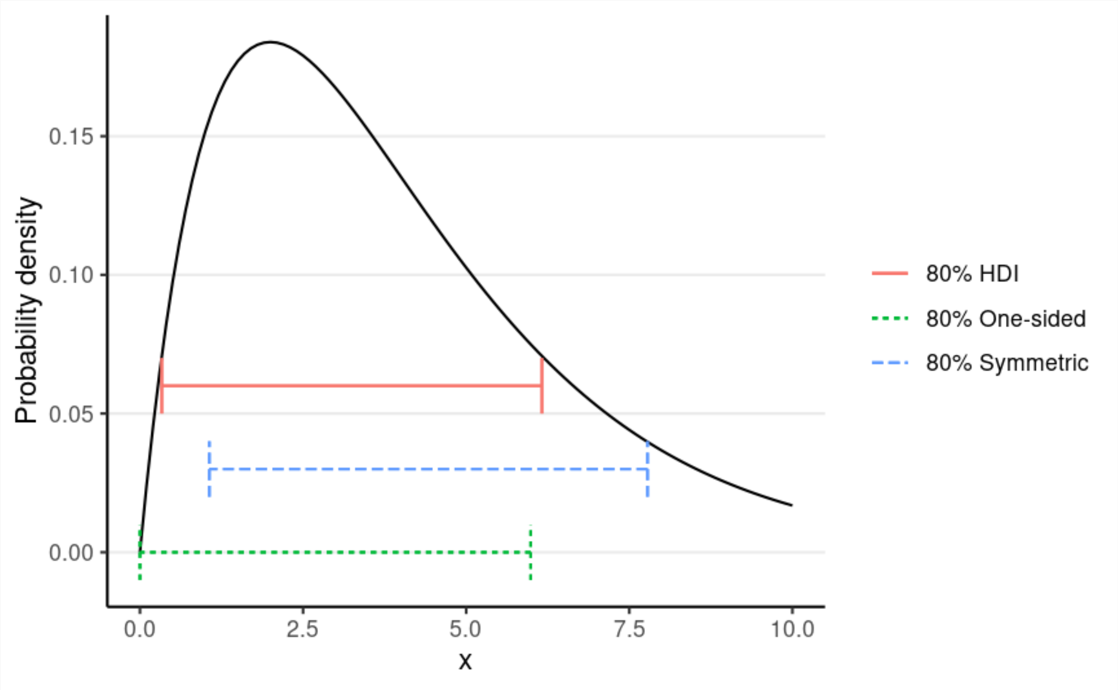
Dispersion

How concentrated the region with high plausibility is

- Variance, standard deviation
- Median absolute deviation (MAD)

Interval

- One-sided
- Symmetric
- Highest density interval (HDI)



Probability with Multiple Variables

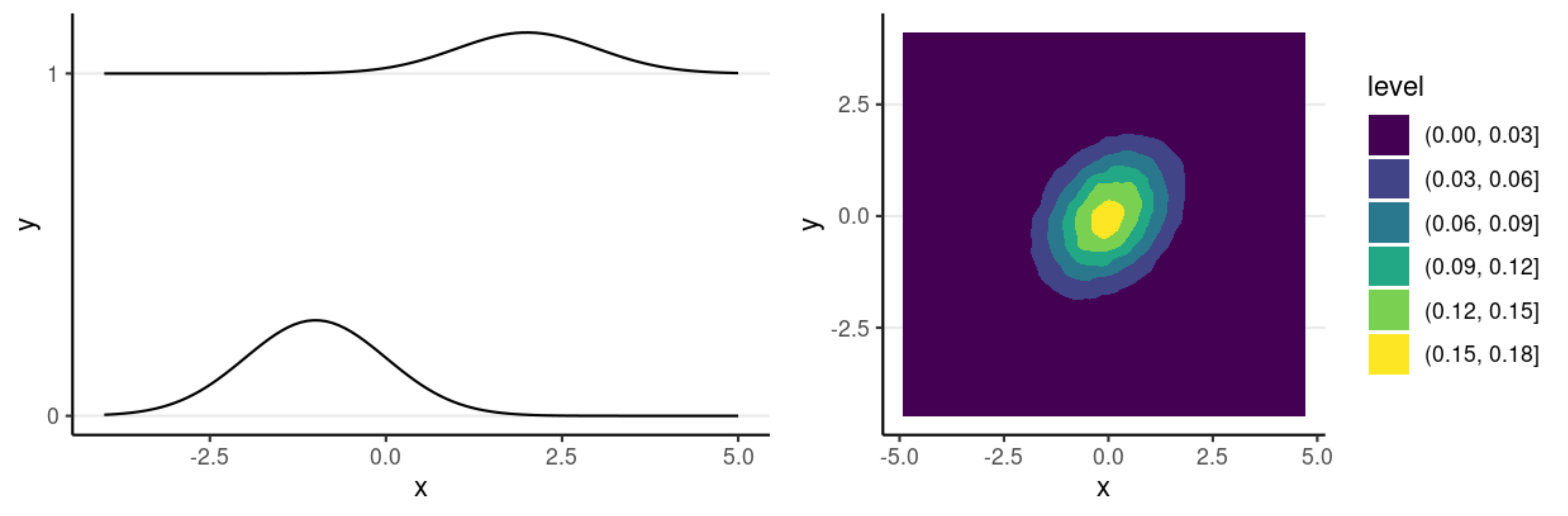
Multiple Variables

- Joint probability: $P(X, Y)$
- Marginal probability: $P(X)$, $P(Y)$

	>= 4	**<= 3**	**Marginal (odd/even)**
odd	1/6	2/6	3/6
even	2/6	1/6	3/6
Marginal (>= 4 or <= 3)	3/6	3/6	1

Continuous Variables

- Left: Continuous X, Discrete Y
- Right: Continuous X and Y



Conditional Probability

Knowing the value of B, the relative plausibility of each value of outcome A

$$P(A \mid B_1) = \frac{P(A, B_1)}{P(B_1)}$$

E.g., P(Alzheimer's) vs. P(Alzheimer's | family history)

E.g., Knowing that the number is odd

	>= 4	**<= 3**
odd	1/6	2/6
~~even~~	~~2/6~~	~~1/6~~
Marginal (>= 4 or <= 3)	3/6	3/6

Conditional = Joint / Marginal

	>= 4	<= 3
odd	1/6	2/6
Marginal (>= 4 or <= 3)	3/6	3/6
Conditional (odd)	(1/6) / (3/6) = 1/3	(1/6) / (2/6) = 2/3

$P(A \mid B) \neq P(B \mid A)$

- P(number is six | even number) = 1 / 3
- P(even number | number is six) = 1

Another example: P(road is wet | it rains) vs. P(it rains | road is wet)

- Problem: Not considering other conditions leading to wet road: sprinkler, street cleaning, etc

Sometimes called the *confusion of the inverse*

Independence

A and B are independent if

> $$P(A \mid B) = P(A)$$

E.g.,

- A: A die shows five or more
- B: A die shows an odd number

P(>= 5) = 1/3. P(>=5 | odd number) = ? P(>=5 | even number) = ?

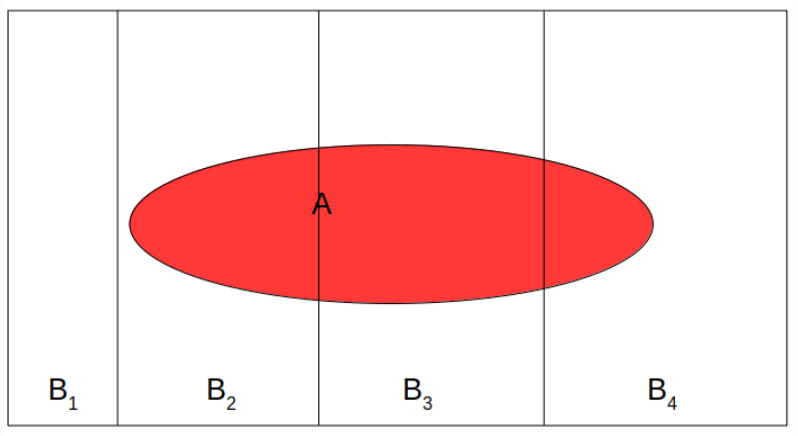
P(<= 5) = 2/3. P(<=5 | odd number) = ? P(>=5 | even number) = ?

Law of Total Probability

From conditional $P(A \mid B)$ to marginal $P(A)$

- If $B_1, B_2, \cdots, B_n$ are all possibilities for an event (so they add up to a probability of 1), then

$$\begin{aligned} P(A) &= P(A, B_1) + P(A, B_2) + \cdots + P(A, B_n) \\ &= P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n) \\ &= \sum_{k=1}^{n} P(A \mid B_k)P(B_k) \end{aligned}$$



> Example: Consider the use of a depression screening test for people with diabetes. For a person with depression, there is an 85% chance the test is positive. For a person without depression, there is a 28.4% chance the test is positive. Assume that 19.1% of people with diabetes have depression. If the test is given to 1,000 people with diabetes, around how many people will be tested positive?

Data source: https://doi.org/10.1016/s0165-0327(12)70004-6, https://doi.org/10.1371/journal.pone.0218512