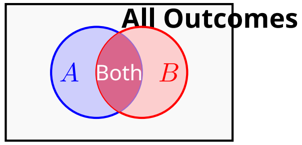
Probability Ultimate Cheat Sheet

Understanding Chance & Uncertainty with Python

1 Probability Foundations

Simple Definitions

- **Sample Space (S)**: The list of everything that *could* happen (e.g., Heads or Tails).
- **Event (A)**: The specific thing we are interested in.
- **Probability $P(A)$**: How likely it is. 0 = Impossible, 1 = Certain.



The Rules of Chance

- **OR Rule (Union)**: Chance of A OR B happening. Add them up, subtract overlap.

$$P(A \cup B) = P(A) + P(B) - \text{Overlap}$$

- **AND Rule (Intersection)**: Chance of A AND B happening together.

$$P(A \cap B) = P(A) \times P(B) \text{ (if independent)}$$

- **NOT Rule**: Chance of A *not* happening.

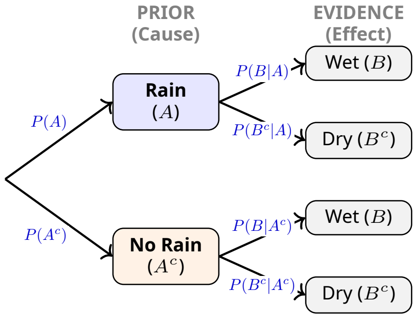
$$P(\text{Not } A) = 1 - P(A)$$

2 Conditional Probability

The "Given" Concept How probability changes when we know something new.

$$P(A|B) = \text{Chance of A, knowing B happened}$$

Bayes' Theorem (The Update Rule) Updating your beliefs when you get new evidence.



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Example: Probability of having a disease (A) given a positive test (B).

3 Random Variables

Turning outcomes into numbers.

Expectation (Mean) $E[X]$ The "long-run average" outcome. If you played the game a million times, what's the average score?

Variance (Spread) How consistent are the results?

- **Low Variance**: Consistent, predictable results.
- **High Variance**: Wild swings, risky.

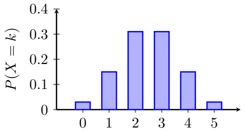
$$\text{Standard Deviation } (\sigma) = \sqrt{\text{Variance}}$$

Python (Scipy Stats)

```
from scipy import stats
# The 'Average' (Location)
mean = stats.norm.mean(loc=0, scale=1)
# The 'Spread' (Scale)
std = stats.norm.std(loc=0, scale=1)
```

4 Discrete Outcomes

Things you can count (1, 2, 3...)



1. Bernoulli (The Coin Flip) One try. Success or Fail. *Example: Will I pass this test?*

2. Binomial (Repeated Flips) How many successes in n tries? *Example: Number of heads in 10 flips.*

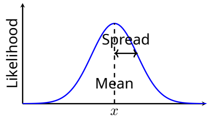
```
# Chance of exactly 2 heads in 5 flips
p = stats.binom.pmf(k=2, n=5, p=0.5)
```

3. Poisson (Rate/Time) How many events happen in a time period? *Example: Number of emails received in 1 hour.*

```
# Chance of 3 emails (avg is 5)
p = stats.poisson.pmf(k=3, mu=5)
```

5 Continuous Outcomes

Things you measure (height, time, speed).



1. Uniform (Random) Every outcome is equally likely. *Example: Rolling a fair die (discrete equivalent).*

2. Normal (The Bell Curve) Most things cluster around the average. Extremes are rare. *Example: Heights, Test Scores, Errors.*

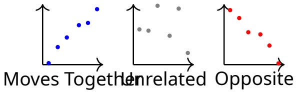
- 68% of data is within 1 standard deviation.

```
# Probability of being below 1.96
p = stats.norm.cdf(1.96) # -> 97.5%
```

3. Exponential (Wait Time) Time until the next event happens. *Example: Time until the bus arrives.*

6 Stats & Theorems

Correlation Do two things move together?



Law of Large Numbers More data = Better Average. If you flip a coin 10 times, you might get 80% heads.

If you flip it 1,000,000 times, you will get very close to 50%.

Central Limit Theorem (The Magic Rule) If you take enough samples from *any* weird distribution, the averages of those samples will look like a Normal Distribution (Bell Curve). *Why it matters: Allows us to use Normal math on messy real-world data.*

7 Testing Ideas

The Setup

- **Null Hypothesis (H_0)**: "Nothing interesting is happening." (Default).
- **p-value**: The "Surprise Factor". If $p < 0.05$, the result is so weird that H_0 is probably wrong.

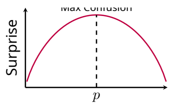
Mistakes We Can Make

		TRUTH (Reality)	
		Null True (H_0)	Null False (H_1)
DECISION	Reject H_0	False Alarm (Type I) α	Correct (Power) $1 - \beta$
	Keep H_0	Correct $1 - \alpha$	Missed It (Type II) β

```
# Compare two groups (A/B Test)
t_stat, p_val = stats.ttest_ind(a, b)
if p_val < 0.05:
    print("Significant Difference!")
```

8 Information Theory

Entropy (Uncertainty) How surprised are we?



A coin flip (50/50) has max entropy (max surprise). A trick coin (100% heads) has 0 entropy.

Cross-Entropy (The Loss Function) Used in AI training. Measures how far our prediction is from the truth. Lower is better.