STAT 311: Introduction to Probability

Y. Samuel Wang

Summer 2016

Logistics

- Practice Exam Posted today

Probability

What do we mean when we say use the word "probability?"

- There is a 10% chance of rain tomorrow
- There is a 50% chance the coin flip is heads
- There is a 80% chance that I accidentally left the stove on

Probability Distributions

Generally, probability refers to patterns which occur in the long run

- Any single event may be unpredictable
- Repeating the same process many times allows us to see long run behavior

Probability Distributions

Generally, probability refers to patterns which occur in the long run

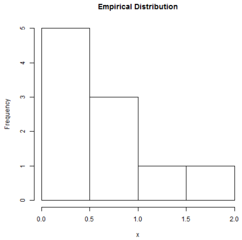
- Any single event may be unpredictable
- Repeating the same process many times allows us to see long run behavior

For Example-

- Any single flip of the coin is unpredictable, but flipping a coin 1 million time will result in roughly 1/2 million heads and 1/2 million tails
- How many emails I receive on any given day is unpredictable, but over many days it will be 10.3 emails

Probability Distributions

- **Empirical Distributions** describe the distribution of observations in some set of data we have seen. This is mostly what we've talked about so far



- **Theoretical Distributions** can be fully described by some mathematical model.
- Standard Normal Distribution-

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Probability Terminology

- The **Sample Space** refers to the set of all possible **outcomes**
- A **trial** is a single occurrence of a random event. Each trial results in an **outcome**. A collection of outcomes is called an **event**.

Probability Terminology

- The **Sample Space** refers to the set of all possible **outcomes**
- A **trial** is a single occurrence of a random event. Each trial results in an **outcome**. A collection of outcomes is called an **event**.

For example-

- When rolling a die, the **Sample Space** is comprised of the **outcomes**

$$\{1, 2, 3, 4, 5, 6\}$$

- Each roll is a **trial**. An **event** we might be interested in is the set of even numbers $\{2, 4, 6\}$.

Probability Terminology

The sample space can be-

- **Discrete**: Outcomes which can be enumerated
- Drawing cards from a deck, number of individuals at a Mariners game, rolls of dice
- **Continuous**: sample space is continuous with uncountable number of outcomes
- Time until your next email, height of individual

Probability of an event

Each possible event is assigned a probability. We can assign probabilities in several ways

- Observing the long run frequency over repeated observations
- Personal beliefs or assumptions about the physical world

Probability of an event

Each possible event is assigned a probability. We can assign probabilities in several ways

- Observing the long run frequency over repeated observations
- Personal beliefs or assumptions about the physical world

Regardless of how they are assigned, probabilities must follow several rules

- The probability of outcome A is denoted $P(A)$
- $0 \leq P(A) \leq 1$
- $P(S) = 1$ where S denotes the entire sample space. In other words, the probability of all possible outcomes must be 1.

Combining events

Outcomes can be combined in two ways to form outcomes

- **Intersection**: Events A and B, can be written as $A \bigcap B$
- The Sounders lose and Clint Dempsey gets hurt
- **Union**: Events A or B, can be written as $A \bigcup B$. Note that this includes the case where A and B happen
- The Mariners win or Kyle Seager hits a home run
- On a single roll of a dice, rolling a number less than 3 or greater than 5

How to relate different events

- **Complement**: Any event which is not A, can be written as A^c
- The temperature is above 80 degrees vs the temperature is less than or equal to 80 degrees
- **Mutually exclusive events**: Two events which cannot happen simultaneously. Note that complements are always mutually exclusive, but mutually exclusive events are not always complements
- For a single roll, outcomes 6 and 4 are mutually exclusive, but not complements since 6 and 4 do not make up the entire sample space

Indepdendent Events

If the result of a trial effects the result of another trial, we say the events are **dependent**. If the result of a trial does not effect the result of another trial, the events are **independent**

- If two events are independent, the conditional probabilities are the same as the marginal probabilities
- If A and B are independent $P(A|B) = P(A)$ and $P(B|A) = P(B)$

Indepdendent Events

If the result of a trial effects the result of another trial, we say the events are **dependent**. If the result of a trial does not effect the result of another trial, the events are **independent**

- If two events are independent, the conditional probabilities are the same as the marginal probabilities
- If A and B are independent $P(A|B) = P(A)$ and $P(B|A) = P(B)$

The color of my shirt is independent of whether or not the Mariners win. So

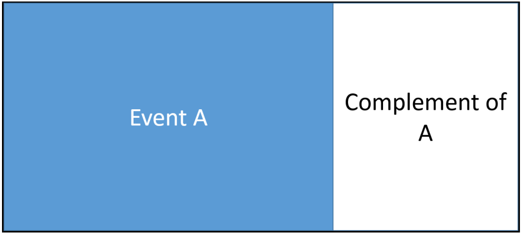
$$P(Win|\text{Blue Shirt}) = P(Win)$$

But whether or not the Mariners win is dependent on whether or not Felix Hernandez is pitching. So

$$P(Win|\text{Felix}) \neq P(Win)$$

Basic Probability rules

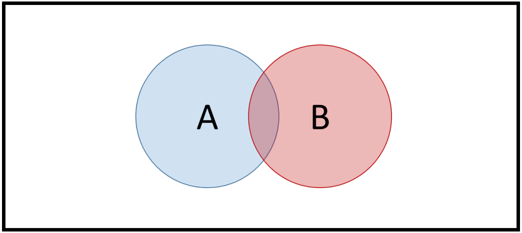
Complement Rule: $P(A^c) = 1 - P(A)$.



If the probability of rain is 25%, the probability of no rain is 75%

Basic Probability rules

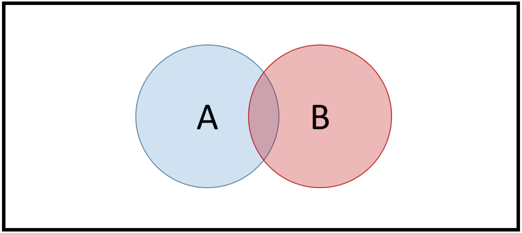
Union Rule: $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$
If A and B are mutually exclusive the $P(A \bigcap B) = 0$



If the probability of being born in July is $1/12$, and the probability of being born on a Thursday is $1/7$ and the probability of being born on a Thursday in July is $4/365$, then the probability of being born in July or a Thursday is $1/12 + 1/7$ - $4/365$

Basic Probability rules

Intersection Rule: $P(A \bigcap B) = P(A|B)P(B) = P(B|A)P(A)$
If A and B are independent then $P(A \bigcap B) = P(A)P(B)$



If the probability Felix Hernandez pitching is $P(\text{Felix}) = 1/5$, and the probability of the Mariners winning when Felix is pitching is $P(\text{Win}|\text{Felix}) = 2/3$, then the probability of Felix pitching and the Mariners winning is $P(\text{Win} \bigcap \text{Felix}) = 2/3 \times 1/5$

Relationship between probabilities

For any events A and B

$$0 \leq P(A \bigcap B) \leq P(A) \leq P(A \bigcup B)$$

Law of total probability

The law of total probability gives us a way to find marginal probabilities by summing across joint probabilities

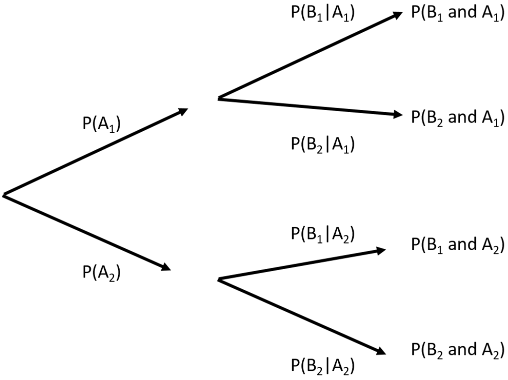
$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$$

We've done this without even thinking in two way tables

	Death	Survive	Total
Guinea	2536	1268	3804
Liberia	4806	5860	10666
Sierra Leone	3955	10167	14122
Total	11297	17295	28592

Law of total probability

When thinking about the law of total probability it can sometimes be helpful to draw our a probability tree.



Bayes Rule

Bayes Rule gives us a way to calculate conditional probabilities from joint and marginal probabilities.

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

Bayes Rule

Bayes Rule gives us a way to calculate conditional probabilities from joint and marginal probabilities.

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

We can better understand the intuition when looking at two way tables again

	Death	Survive	Total
Guinea	2536	1268	3804
Liberia	4806	5860	10666
Sierra Leone	3955	10167	14122
Total	11297	17295	28592