Lecture 7: Probability

Chapter 2.x

Outcomes

Probability forms the theoretical backbone of statistics. We use probability to characterize randomness.

We often frame probability in terms of a random process giving rise to an outcome.

Typical examples

- Die roll: 6 outcomes
- Coin Flip: 2 outcomes

Disjoint AKA Mutually Exclusive Outcomes

Two outcomes are disjoint (AKA mutually exclusive) if they cannot both occur at the same time.

Die example:

- Rolling a 1 and a 2 are disjoint.
- Rolling a 1 and rolling "an odd number" are not disjoint.

Addition Rule of Probability

If A_1 and A_2 are disjoint outcomes, then

$$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$$

Ex: Rolling 1 and 2 are disjoint, so:

$$P(\text{rolling 1 or 2}) = P(\text{rolling 1}) + P(\text{rolling 2}) = \frac{1}{6} + \frac{1}{6}$$

General Addition Rule of Probability

If A_1 and A_2 are two outcomes (not necessarily disjoint), then

$$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2) - P(A_1 \text{ and } A_2)$$

Venn diagram:

General Addition Rule of Probability

Events are just combinations of outcomes. Ex: Deck of cards

- A_1 = event we draw a diamond
- A_2 = event we draw a face card

These two events are not disjoint, as there are 3 diamond face cards. Venn diagram:

General Addition Rule of Probability

$$
\begin{aligned}
P(A_1 \text{ or } A_2) &= P(\text{diamond or a face card}) \\
&= P(\text{diamond}) + P(\text{face card}) - \\
&\quad P(\text{diamond AND face card}) \\
&= \frac{13}{52} + \frac{3 \times 4}{52} - \frac{3}{52} = \frac{22}{52} = 42.3\%
\end{aligned}
$$

Sample Space and the Complement of Events

A die has 6 possible outcomes. The sample space is the set of all possible outcomes $S = \{1, 2, \ldots, 6\}$.

Say event A is the event of rolling an even number i.e $A = \{2, 4, 6\}$. The complement of event A is $A^c = \{1, 3, 5\}$ i.e. getting an odd number.

We have

$$P(A) + P(A^c) = 1$$

Independence

Two processes are independent if knowing the outcome of one provides no useful information about the outcome of the other. Otherwise they are dependent.

Consider:

1. Die rolls
2. You get a movie recommendation from your friend Robin, but then their significant other Sam also recommends it.
3. You compare test scores from two Grade 9 students in the same class. Then same school. Then same school district. Then same city. Then same state.

Independence

We say that events A and B are independent if

$$P(A \text{ and } B) = P(A) \times P(B)$$

Ex: Dice rolls are independent:

$$\begin{aligned} P(\text{rolling 1 and then 6}) &= P(\text{rolling 1}) \times P(\text{rolling 6}) \\ &= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \end{aligned}$$

Conditional Probability

The conditional probability of an event A given the event B, is defined by

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

Example

Let's suppose I take a random sample of 100 Reed students to study their smoking habits.

	Smoker	Not Smoker	Total
Male	19	41	60
Female	12	28	40
Total	31	69	100

- What is the probability of a randomly selected male smoking?

$$P(S|M) = \frac{P(S \text{ and } M)}{P(M)} = \frac{19/100}{60/100} = \frac{19}{60}$$

- What is the probability that a randomly selected smoker is female?

$$P(F|S) = \frac{P(F \text{ and } S)}{P(S)} = \frac{12/100}{31/100} = \frac{12}{31}$$

Put It Together! Independence and Conditional Prob.

If A and B are independent events, then

$$P(A \text{ and } B) = P(A) \times P(B)$$

then

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

i.e. $P(A|B) = P(A)$: the event B occurring has no bearing on the probability of A

Gambler's Fallacy: Roulette

You can bet on individual numbers, sets of numbers, or red vs black. Let's assume no 0 or 00, so that $P(R) = P(B) = \frac{1}{2}$.

Gambler's Fallacy: Roulette

One of the biggest cons in casinos: spin history boards.

Let's ignore the numbers and just focus on what color occurred.
Note: the white values on the left are black spins.

Gambler's Fallacy: Roulette

Let's say you look at the board and see that the last 4 spins were red.

You will always hear people say "Black is due!"

Ex. on the 5th spin people think:

$$P(B_5 \mid R_1 \text{ and } R_2 \text{ and } R_3 \text{ and } R_4) >$$
$$P(R_5 \mid R_1 \text{ and } R_2 \text{ and } R_3 \text{ and } R_4)$$

Gambler's Fallacy: Roulette

But assuming the wheel is not rigged, spins are independent i.e. $P(A|B) = P(A)$. So:

$$
\begin{aligned}
P(B_5|R_1 \text{ and } R_2 \text{ and } R_3 \text{ and } R_4) = P(B_5) &= \frac{1}{2} \\
P(R_5|R_1 \text{ and } R_2 \text{ and } R_3 \text{ and } R_4) = P(R_5) &= \frac{1}{2}
\end{aligned}
$$

Next Time

Discuss the Normal Distribution