

## Probability

- A full description of a random phenomenon requires:

1. Listing of all the possible outcomes.
2. List the probability associated with each outcome.

- Example: Tossing a fair coin.

1. Possible outcomes: Heads or tails
2. Probability of heads $=0.5$; probability of tail $=0.5$

## Sample Space

- Sample Space: Set of all possible outcomes, usually denoted by $\mathbf{S}$.
- Example: Tossing a fair coin

$$
\text { S = \{Heads, Tails }\}
$$

- Example 2: Tossing a fair coin three times and counting the number of heads.

$$
S=\{0,1,2,3\}
$$

## Sample Space and Events

- Event: An event is an outcome or a set of outcomes and is a subset of the sample space $S$.
- Events (often denoted with the letters A, B, C,...) have associated probabilities.
- Example: Outcome of tossing a fair coin three times:
$\mathrm{S}=\{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{TH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$
- Let A denote the event of 2 heads;
$\mathrm{A}=\{\mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\}$


## Probability

- The probability if an event is its long-term relative frequency of occurrence.
- If an experiment is repeated $\mathbf{n}$ times under essentially identical conditions and the event $\mathbf{A}$ occurs $\mathbf{m}$ times, then as $n$ gets large,
$P(A)=m / n$


## Probability

- An event that always occurs: $\mathrm{n} / \mathrm{n}=1$ $\mathrm{P}($ "'sure thing") $=1$
- An event that never occurs: $0 / \mathrm{n}=0$ P ("impossible") $=0$


## Probability

- For any event A,
$\mathrm{m} \leq \mathrm{n}$, so
$0 \leq \mathrm{P}(\mathrm{A}) \leq 1$
- Complement
$P(A)=m / n$
$\mathrm{P}\left(\mathrm{A}^{\mathrm{c}}\right)=(\mathrm{n}-\mathrm{m}) / \mathrm{n}=1-\mathrm{P}(\mathrm{A})$
$\mathrm{P}(\mathrm{A})+\mathrm{P}\left(\mathrm{A}^{\mathrm{c}}\right)=1$


## Disjoint or Mutually Exclusive Events

- Any two events A and B are disjoint if they have no outcomes in common and so cannot ever occur simultaneously.
- If A and B are disjoint:
$\mathrm{P}(\mathrm{A}$ or B$)=\mathrm{P}(\mathrm{A})+\mathrm{P}(\mathrm{B})$
This is the addition rule for disjoint events.


## Overlapping Events

- If two events A and B are not disjoint, they have outcomes in common and the addition rule cannot be applied.
- Instead:
$\mathrm{P}(\mathrm{A}$ or B$)=\mathrm{P}(\mathrm{A})+\mathrm{P}(\mathrm{B})-\mathrm{P}(\mathrm{A}$ and B$)$


## Conditional Probability

- Conditional probability is the probability of $B$ given, or knowing that the event A has happened.
$\mathrm{P}(\mathrm{B} \mid \mathrm{A})=\mathrm{P}(\mathrm{A}$ and B$) / \mathrm{P}(\mathrm{A})$


## General Multiplication Rule

- $\mathrm{P}(\mathrm{A}$ and B$)=\mathrm{P}(\mathrm{A}) \mathrm{P}(\mathrm{B} \mid \mathrm{A})$
- Note:
- $\mathrm{P}(\mathrm{B} \mid \mathrm{A})=\mathrm{P}(\mathrm{A}$ and B$) / \mathrm{P}(\mathrm{A})$
- $\mathrm{P}(\mathrm{A} \mid \mathrm{B})=\mathrm{P}(\mathrm{A}$ and B$) / \mathrm{P}(\mathrm{B})$
- $\mathrm{P}(\mathrm{A}$ and B$)=\mathrm{P}(\mathrm{A}) \mathrm{P}(\mathrm{B} \mid \mathrm{A})$

$$
=\mathrm{P}(\mathrm{~B}) \mathrm{P}(\mathrm{~A} \mid \mathrm{B})
$$

## Multiplication Rule for Independent Events

- A and B are said to be "independent" if $\mathrm{P}(\mathrm{A}$ and B$)=\mathrm{P}(\mathrm{A}) \mathrm{P}(\mathrm{B})$
- Also, since in general:
$\mathrm{P}(\mathrm{A}$ and B$)=\mathrm{P}(\mathrm{A}) \mathrm{P}(\mathrm{B} \mid \mathrm{A})$
independence implies that $\mathrm{P}(\mathrm{B} \mid \mathrm{A})=\mathrm{P}(\mathrm{B})$.
- Similarly, independence implies that $\mathrm{P}(\mathrm{A} \mid \mathrm{B})=\mathrm{P}(\mathrm{A})$


## Example

- A mandatory drug test has a false-positive rate of $1.2 \%$ (or 0.012).
- Given 150 employees who are drug free, what is the probability that at least one will (falsely) test positive?
- $\mathrm{P}($ at least one positive $)=\mathrm{P}(1$ or 2 or $3 \ldots$ or 150 positive $)$
$=1-\mathrm{P}$ (none positive)
$=1-\mathrm{P}(150$ negative $)$
$\mathrm{P}(150$ negative $)=(0.988)^{150}=0.16$
$\mathrm{P}($ at least 1 positive $)=0.84$


## Summary

1. For any event $\mathrm{A}, 0 \leq \mathrm{P}(\mathrm{A}) \leq 1$
2. $\mathrm{P}(\mathrm{S})=1$, where S is the sample space
3. For any event $\mathrm{A}, \mathrm{P}\left(\mathrm{A}^{\mathrm{c}}\right)=1-\mathrm{P}(\mathrm{A})$
4. Addition Rule: $\mathrm{P}(\mathrm{A}$ or B$)=\mathrm{P}(\mathrm{A})+\mathrm{P}(\mathrm{B})-\mathrm{P}(\mathrm{A}$ and B$)$
5. Conditional Probability: $\mathrm{P}(\mathrm{B} \mid \mathrm{A})=\mathrm{P}(\mathrm{A}$ and B$) / \mathrm{P}(\mathrm{A})$
6. Multiplication Rule for Independent Events:

If A and B are independent, $\mathrm{P}(\mathrm{A}$ and B$)=\mathrm{P}(\mathrm{A}) \mathrm{P}(\mathrm{B})$

## Why Intuition Can Be BAD

- Using intuition when it comes to probability has caused many people some BIG problems.
- Consider the Gambler's Fallacy:
- A gambler feels like he/she is on "a roll." The gambler continues to wager more and more money assuming that the streak of luck will continue.
- Ultimately, the gambler will lose.
- Law of small numbers:
- Small samples do not represent the population well. Thus, laws of probability may not appear to hold.


## Why Intuition Can Be BAD

- Confusion of the inverse:
- The Eddy Scenario (1982).
- Doctors asked if a patient has a positive mammogram, and the test is $90 \%$ for those with benign lumps (and $80 \%$ accurate for those with malignancies), what is the probability the patient has breast cancer?
- Prevalence of breast cancer is 0.01 .
- Most doctors said 0.75.
- In truth, it's 0.075 .


## Why Intuition Can Be BAD

- The Birthday Problem
- If there are 27 of you, what is the probability that at least two of you share the same birthday? 0.627
- In a class of 50 , the probability is 0.970 .
- These numbers seem counterintuitive.
- How can we calculate this?

