PROBABILITY

The set of all **outcomes** for an **experiment** is called the **sample space** (usually S, but our book uses Ω). An **event** is a set of outcomes. The assignment of probabilities to events must obey the following rules:

Axioms of Probability. *Axioms for probability:*

(1) $P(E) \geq 0$ *for any event* E
(2) $P(S) = 1$
(3) *If* $E_1, E_2, E_3, \ldots$ *are disjoint events, then* $P\left(E_1 \cup E_2 \cup E_3 \cup \ldots\right) = \sum_{i=1}^{\infty} P\left(E_i\right)$

Theorem (Basic theorems of probability)**.** *Let* A *and* B *be events.*

(1) $P(\emptyset) = 0$
(2) *If* $A \subseteq B$*, then* $P(A) \leq P(B)$
(3) $P(A) = 1 - P(A^C)$
(4) $P(A) = P(A \cap B) + P(A \cap B^C)$
(5) $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$

Note that part 5 can be rearranged: $P(A \cap B) = P(A) + P(B) - P(A \cup B)$.

Theorem. *If an experiment has* N *equally likely outcomes, then* $P(E) = \dfrac{\textit{number of outcomes in } E}{N}$.

This means that counting is one of our basic tools for calculating probabilities.

Method. (Multiplication rule for counting) If a process occurs in two steps and there are m options for the first step and n options for the second, then there are mn total possibilities.

Method. The number of ways to select k elements from an n-element set is...

	Order matters	Order doesn't matter
With replacement	n^k	$\binom{n+k-1}{k}$
Without replacement	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{(n-k)!k!}$

Definition. Let A and B be events with $P(B) \neq 0$. The **conditional probability of** A **given** B is

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

.

Definition. Events A and B are **independent** if and only if $P(A \cap B) = P(A)P(B)$.

Note that it follows that A and B are independent if $P(A|B) = P(A)$, that is, knowing that event B has occurred does not affect our calculation of $P(A)$.

Theorem (Multiplication rule for probabilities)**.** *Let* A *and* B *be events with* $P(B) \neq 0$*. Then*

$$P(A \cap B) = P(A|B)P(B)$$

Theorem (The Law of Total Probability)**.** *If event* B *has probability strictly between* 0 *and* 1*, then for any event* A*,* $P(A) = P(A|B)P(B) + P(A|B^C)P(B^C)$.

Theorem (Bayes' Law)**.** *If* A *and* B *are events with positive probability, then*

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

Note: often the law of total probability is used to calculate $P(A)$*.*

Date: February 4, 2020.

Theorem (Extended multiplication rule for probabilities)**.** *Let $E_1, E_2, \ldots, E_n$ be events with $P(E_i) \neq 0$. Then*

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1)\, P(E_2|E_1)\, P(E_3|E_1 \cap E_2) \ldots P(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1})$$

Definition. Events $E_1, E_2, \ldots, E_2$ form a **partition** of S if

(1) $E_1 \cup E_2 \cup \cdots \cup E_2 = S$ and
(2) $E_i \cap E_j = \emptyset$ if $i \neq j$ (the events are pairwise disjoint).

Theorem (The Law of Total Probability Extended)**.** *If events $E_1, E_2, \ldots, E_n$ each have probability strictly between 0 and 1 and form a partition of S, then for any event A,*

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_n)P(E_n)$$

Definition. A **random variable** X assigns a number to each outcome in the sample space S.

(1) All random variables have a **cumulative distribution function (CDF)**: $F(x) = P(X \leq x)$.
(2) A discrete random variable has a **probability mass function (PMF)**: $p(x) = P(X = x)$.

Example. Suppose you roll a fair 6-sided die until the first 6. Let X be the total number of rolls. Find the PMF of X.

Theorem. *The CDF of any random variable satisfies the following:*

(1) *it is non-decreasing: if $a \leq b$, then $F(a) \leq F(b)$*
(2) $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to \infty} F(x) = 1$
(3) *If $a < b$, then $P(a < X \leq b) = F(b) - F(a)$*

Theorem. *The PMF of any random variable satisfies the following:*

(1) $0 \leq p(x) \leq 1$ *for all x*
(2) $\sum_x p(x) = 1$ *(where the sum is over all possible values of X)*

Definition. The **expected value** (or **mean**) of a random variable is a kind of weighted average and is denoted $E(X)$ or μ.

(1) If X is a discrete RV with PMF $p(x)$, then $E(X) = \sum_x xp(x)$.