PROBABILITY SO FAR

The set of all **outcomes** for an **experiment** is called the **sample space** and we'll usually call it S. An **event** is a set of outcomes for a given experiment. 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* $P(E_1 \cup E_2 \cup E_3 \cup \ldots) = \sum_{i=1}^{\infty} P(E_i)$

Theorem. *Some properties of sets:*

(1) $(A \cup B)^C = A^C \cap B^C$
(2) $(A \cap B)^C = A^C \cup B^C$
(3) $A \cup A^C = S$
(4) $A = (A \cap B) \cup (A \cap B^C)$ *and the sets* $A \cap B$ *and* $A \cap B^C$ *are disjoint*

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

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	$\dfrac{n!}{(n-k)!}$	$\binom{n}{k} = \dfrac{n!}{(n-k)!k!}$

Theorem. *Some properties of* **binomial coefficients***:*

(1) $\binom{n}{k} = \binom{n}{n-k}$
(2) $\binom{n}{1} = \binom{n}{n-1} = n$
(3) $\binom{n}{2} = \dfrac{n(n-1)}{2}$
(4) $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$
(5) $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
(6) $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Date: September 15, 2019.

Theorem. *Properties of (all) probabilities:*

(1) $P(\emptyset) = 0$
(2) $P(E) = 1 - P\left(E^C\right)$
(3) *If* $A \subseteq B$*, then* $P(A) \leq P(B)$
(4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Note that property 4 can be rearranged: $P(A \cap B) = P(A) + P(B) - P(A \cup B)$.

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

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

.

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

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

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

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,*

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

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 $P(A)$.

Example. Suppose 0.01% of a population carry a genetic disease (and have no obvious symptoms). Doctors screen for the disease using a test that correctly identifies carriers 99.9% of the time. The test also gives false positives 0.2% of the time. What is the probability that a random person who tests positive actually carries the disease?

Example. There are two taxi companies in Extremistan: Crimson Cab (which has 80% of the cars) and Tangerine Taxi (which owns the remaining 20%). A taxi from one of these companies sideswiped some parked cars. An eyewitness claims to have seen a Tangerine Taxi, but given the conditions (dusk) and distance (1 block) the eyewitness is considered to be only 75% reliable. Should Tangerine Taxi be found liable?

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

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)**: $p(x) = P(X = x)$.
(3) A continuous random variable has a **probability density function (PDF)** $f(x)$ such that for any numbers a and b (with $a \leq b$)

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Example. Find the PMF of a random variable having the CDF

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{6} & \text{if } 0 \le x < 1 \\ \frac{1}{2} & \text{if } 1 \le x < 3 \\ 1 & \text{if } 3 \le x \end{cases}$$

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 $a \le b$, then $F(a) \le F(b)$
(2) $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$
(3) If $a < b$, then $P(a < X \le b) = F(b) - F(a)$

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

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

Theorem. *A function f may be the PDF of a random variable if and only if*

(1) $f(x) \ge 0$ for all x[1] and
(2) $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1.$

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 $E(X) = \sum_x xp(x)$.
(2) If X is a continuous RV with PDF $f(x)$, then $E(X) = \int_{-\infty}^{\infty} xf(x)dx$.

[1] this is not technically required