PROBABILITY AND DISCRETE DISTRIBUTIONS

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

Axioms of Probability.

1. $P(S) = 1$
2. $P(E) \geq 0$ for any event E
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 (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. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
6. $P(A \cap B) = P(B|A)P(A)$

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

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

Note. The number of **combinations** (bottom right corner) $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ is implemented in **R** as `choose(n,k)`. It is implemented in most calculators as `nCk` or `C(n,k)`. The number of **permutations** (bottom left corner) $\frac{n!}{(n-k)!}$ is implemented in most calculators as `nPk` or `P(n,k)`.

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

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

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

Note that if $P(B|A) = P(B)$, then A and B are independent. In other words, if knowing that event A has occurred does not affect our calculation of $P(B)$, then A and B are independent.

Theorem (The Law of Total Probability)**.** If event A has probability strictly between 0 and 1, then for any event B,

$$P(B) = P(B|A)P(A) + P(B|A^C)P(A^C)$$

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

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

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

Date: February 6, 2024.

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)** defined for all real numbers: $F(x) = P(X \leq x)$.
2. If X is a discrete random variable with possible values $x_1, x_2, x_3, \ldots$, then X has a **probability mass function (PMF)**: $p(x_i) = P(X = x_i)$.

Theorem. The PMF of any (discrete) random variable with possible values $x_1, x_2, x_3, \ldots$ has the following properties:

1. $0 \leq p(x_i) \leq 1$
2. $\displaystyle\sum_i p(x_i) = 1$

Theorem. The CDF of any random variable has the following properties:

1. 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 $P(a < X \leq b) = F(b) - F(a)$

Definition. The **expected value** (or **mean**) of a random variable is a weighted average and is denoted $E(X)$ or μ_X. If X is a discrete RV with possible values $x_1, x_2, x_3, \ldots$, then

$$E(X) = \sum_i x_i P(X = x_i)$$

Definition. A random variable X has a **Bernoulli distribution** with parameter p (with $0 < p < 1$) if X has possible values 0 and 1 with $P(X = 1) = p$ and $P(X = 0) = 1 - p$. The outcome 1 is often referred to as "success" while 0 is "failure" and the experiment is often called a Bernoulli trial.

Definition. The total number of successes in n independent, identically distributed (iid) Bernoulli trials with parameter p is a random variable with a **binomial distribution**. The PMF of a random variable X having a binomial distribution with parameters n and p is

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \ldots, n$$

R Implementation. If $X \sim \text{binom}(n, p)$, then the PMF is `dbinom(x, n, p)` and the CDF is `pbinom(x, n, p)`.

Definition. Suppose n elements are to be selected without replacement from a population of size $M_1 + M_2$ consisting of M_1 successes and M_2 failures and the order of selection doesn't matter. If $n \leq M_1$, then the number of successes selected is a **hypergeometric** random variable and its PMF is

$$P(X = x) = \frac{\binom{M_1}{x}\binom{M_2}{n-x}}{\binom{M_1+M_2}{n}} \text{ for } x = 0, 1, 2, \ldots, n$$

R Implementation. If $X \sim \text{hyper}(M_1, M_2, n)$, then the PMF is `dhyper(x, M₁, M₂, n)` and the CDF is `phyper(x, M₁, M₂, n)`.

Note. A standard deck of cards has 52 cards in 4 suits: spades ♠, hearts ♡, clubs ♣, and diamonds ♢. Spades and clubs are black suits, while hearts and diamonds are red suits. Within each suit there are 13 ranks: 2, 3, 4, . . . , 10, Jack, Queen, King, and Ace. Together, a rank and a suit uniquely identify the card (cards are 2-dimensional).

1. Suppose 10 cards are dealt from a well-shuffled deck.

a) Find the probability that exactly 5 are hearts.

b) Find the probability that 5 or more are hearts.

2. Suppose now that you shuffle the deck and look a the top card 10 times (shuffling between attempts).

a) Find the probability that exactly 5 of the cards you see are hearts.

b) Find the probability that 5 or more are hearts.

3. For this problem, suppose you have created a mega deck by shuffling together 10 regular decks (for a total of 520 cards).

a) Find the probability that exactly 5 are hearts if you deal 10 cards from the mega deck.

b) Find the probability that exactly 5 cards are hearts when you shuffle and look at the top card 10 times (and shuffle between attempts).

4. Now use a super-mega deck created by shuffling together 100 regular decks (or 10 mega decks).

a) Repeat both parts a and b of the last problem for the super-mega deck of 5200 cards.

b) What are your observations?

5. Return to a regular deck of 52 cards. The goal this time is to find the PMF of a new distribution in which you count the number of cards you must look at to find a heart. For this problem, take the inefficient approach of shuffling and looking at just the top card. Return the card to the deck an repeat until you see a heart. Let X be the number of times you do this (including the time when you see the heart and stop).

a) What are the possible values for X?

b) Find the PMF for X.

6. This is the more reasonable version of the previous problem: deal cards from a well-shuffled deck until you deal a heart. Let Y be the number of cards you deal (including the heart).

a) What are the possible values for Y?

b) Find the PMF for Y.