Probability Tutorial

Building Blocks

- Sample space (finite set); for example,

$$\Omega = \{Paul, Brian, Bill\}$$

  is a set of three names, and

$$\Omega = \{2, 3, 5, 7, 11, 13\}$$

  is the set of the first 6 prime numbers.

- Often we don't need to specify what's in the sample space, but it's still useful to have placeholders.

- For a set with N elements we write

$$\Omega = \{x_1, x_2, \ldots, x_N\}$$

- In the preceding first example, $N = 3$ and $x_1 = Paul, x_2 = Brian, x_3 = Bill$ or N=6, etc.

Probability Measure or Distribution

- A probability on a finite set is a collection of numbers

$$P(x_1), P(x_2), \ldots, P(x_N)$$

with $P(x)$ non-negative–we write $P(x) \geq 0$ and in fact $P(x) = 0$ is allowed–and with total sum 1, that is,

$$P(x_1) + P(x_2) + \cdots + P(x_N) = 1.$$

- For example, $P(Paul) = P(Brian) = P(Bill) = 1/3$ or $P(Paul) = 1/10, P(Brian) = 1/5, P(Bill) = 7/10$.

Probability Measure or Distribution: Examples

- A three-volume work is placed in random order on a bookshelf.
  - What is the probability of the volumes being in proper (increasing) order from left to right?

Basic Question

- Given a sample space $\Omega = \{x_1, x_2, \ldots, x_N\}$, with N elements and $P(x_1), P(x_2), \ldots, P(x_N)$, a probability on Ω; and a subset A[a] of Ω, calculate $P(A)$,

$$P(A) = \sum_{x \in A} P(x)$$

- Often it is required to compute, evaluate, or approximate $P(A)$.[b]

[a] We often call such a set A and *event*.

[b] We will often use the notation $\Pr(A)$ when the probability distribution P is clear from the context.

Basic Question: Example

- If $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $P(x) = 1/10$ for all x, and $A = \{2, 3, 5, 7\}$ is the set of primes in X, then $P(A) = 4/10$.
- Let A denote the set of outcomes for which the sum of two standard dice equals 7. What is the probability of the event A?
- I.e., someone tells us Ω, $P(x)$, and A and our job is to calculate $P(A)$.
- Computing $P(A)$ can be difficult if N is large or $P(x)$ is given in an indirect way, or A is a complicated set.

Basic Rules/Facts

- Sum rule for disjoint events
- Product rule for independent events
- Conditional probability
- Bayes' theorem
- Law of total probability
- Random variables
- Expectation
- Conditional expectation
- Wald's Identity.

Sum Rule

- If A and B are subsets of X with no element in common, we denote with $A \cup B$ the set $\{x : x \in A \text{ or } x \in B\}$. We have

$$P(A \cup B) = P(A) + P(B).$$

Sum Rule: Examples

1. In throwing a pair of dice, let A be the event that "the total number (sum) of spots is even," A_1 the event that "both dice turn up even," and A_2 the event that "both dice turn up odd"
   - Show that $A = A_1 \cup A_2$
   - Show that A_1 and A_2 are mutually exclusive.
2. One shooter has an 80% probability of hitting a target, while another has only a 70% probability of hitting the target.
   - What is the probability of the target being hit (at least once) if both shooters fire at it simultaneously?

   Hint: Do not use the sum rule!

Independence and Product Rule

- A and B are independent if

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

- Independence depends both on A, B, and the probability P.

Conditional Probability

- We define conditional probability for subsets A and B (with $P(B) > 0$) as

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

- The left-hand side is read as the probability of A given that B has occurred.

- It is defined by taking the underlying probability P, restricting it to B–that's $P(A \cap B)$–and then renormalizing so that the total sum is 1.

Conditional Probability: Examples

1. Prove that if $P(A|B) > P(A)$, then $P(B|A) > P(B)$.
2. Two events A and B with positive probabilities are incompatible (disjoint). Are they dependent?
3. If events A and B are independent the $P(A|B) = P(A)$.

Bayes' Theorem

- If A and B are any subsets, both with positive probability, then

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

- This follows easily from

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

Bayes' Theorem: Examples

1. A coin is flipped twice. If we assume that all four points in the sample space $S = \{(H,H),(H,T),(T,H),(T,T)\}$ are equally likely, what is the conditional probability that both flips result in heads, given that the first flip does?

   Hint. If $E = \{(H,H)\}$ denotes the event that both flips land heads, and $F = \{(H,H),(H,T)\}$ the event that the first flip lands heads, then the desired probability is given by $P(E|F)$.

Law of Total Probability

- Let $B_1, B_2, \ldots, B_k$ be a decomposition of the probability space X into disjoint subsets with $P(B_i) > 0$ for all i. Then, for any set A,

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

Random Variables and Expectation

- A random variable[a] is a variable whose possible values are outcomes of a random phenomenon. (Thus, an r.v. is a numerical measurement of the outcome of a random process.)
- The expectation of an r.v. X is the average with points weighted by $\Pr[X = x]$,

$$E[X] = \sum_x x \Pr[X = x].$$

- A simple consequence of the preceding definitions is the most useful linearity property: If $X_1, X_2, \ldots, X_n$ are random variables, then

$$E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] \qquad (1)$$

[a] We use the abbreviation r.v. to mean random variable.

Useful Formula for the Expectation

- In many applications the following formula is very useful:

$$E[X] = \sum_x \Pr[X > x]$$

Random Variables and Expectation: Examples

1. Two dice are rolled independently at random. What is the expected sum of the two dice?

Conditional Expectation

- If X and Y are random variables, define the conditional expectation of Y given $X = x$ as

$$E(Y|X = x) = \sum_{z} Y(z)P(z|X = x).$$

  On the right, $P(z|X = x) = P(z)/P(B)$ if z is in B and zero if z is not in B, where $B = \{y : X(y) = x\}$.

- Like the expectation, the "conditional expectation" is still linear, as a function of Y.

- If X and Y are independent,

$$E(Y|X = x) = E(Y)$$

Wald's Identity

- **Theorem 1** *Let $X_1, X_2, \ldots$ be independent, identically distributed random variables with finite mean $E[X]$. Let N be a random variable with finite mean and non-negative integer values such that N is independent of X_i for $i \leq N$. Then the following identity holds*

$$E\left[\sum_{k=1}^{N} X_k\right] = E[X] \cdot E[N]. \tag{2}$$

- Note the difference between Equations (1) and (2)!

Exercises[a]

1. For any subsets A and B in any probability space, verify the rule $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$, where $A \cap B$, the intersection of A and B.

2. Prove the law of total probability.

3. Prove the rule of linearity of expectation.

4. Prove the rule of linearity of conditional expectation.

[a] Do not hand in!

Sources

- Persi Diaconis and Brian Skyrms, Ten Great Ideas about Chance. Princeton University Press, 2017.
- W. Feller, An Introduction to Probability Theory and its Applications, 3d ed. (New York: Wiley, 2008).
- Evangelos Kranakis, Useful Ideas in Elementary Probability (Notes), Download from course web page.
- Laszlo. Mero, Logic of Miracles: Making Sense of Rare, Really Rare, and Impossibly Rare Events. Yale University Press, 2018.
- Sheldon M. Ross, Probability Models for Computer Science, Harcourt/Academic Press, 2001