ECE302: Probability and Applications

Arnav Patil

University of Toronto

Contents

1 Random Experiments and Foundations of Probability

There are deterministic models that give the same results every time (e.g. Ohm's and Kirchoff's Laws) and random models (Internet traffic, gait analysis) whose results we cannot predict. Random experiments are experiments in which the outcome is unpredictable when repeated under the same conditions

$$S = \{k_1, k_2, k_3, \ldots, k_N\}$$

S is the sample set of all possible outcomes.
k_i for $i = 1, 2, \ldots, N$ are all of the possible outcomes.

Relative frequency represents how often something happens divided by all outcomes.

$$f_k(n) = \frac{N_k(n)}{n}$$

$$\lim_{n \to \infty} f_k(n) = p_k$$

Three Axioms of Probability

- Axiom 1: $0 \leq P[A] \leq 1$
- Axiom 2: $P[S] = 1$
- Axiom 3: If A and B cannot occur simultaneously then $P[A \text{ } or \text{ } B] = P[A] + P[B]$

Some more notes:

- U is the universal set that contains all possible objects and outcomes in a given application.
- Every set contains the empty set – the set with no elements {}
- If $A \cap B = 0$ then the sets are disjoint, or mutually exclusive.

Sample Spaces

If S is discrete:

$$P[A] = \frac{\text{\# elements in A}}{N}$$

If S is continuous:

$$P[A] = \frac{\text{area of A}}{\text{area of S}}$$

Regardless of mutual exclusivity, we find that:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

Intro to Conditional Probability

$$P[B|A] = \frac{P[A \cap B]}{P[A]}$$

2 Conditional Probability, Bayes' Rule, and Independence

Conditional probability is a measure of the probability of an event happening given another event having already occurred.

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

The low of total probability expresses the total probability of an outcome as a sum of several distinct events.

$$\begin{aligned} P[A] &= P[A|B_1] \cdot P[B_1] + P[A|B_2] \cdot P[B_2] + \ldots + P[A|B_n] \cdot P[B_n] \\ &= \sum_{i=1}^{N} P[A|B_i] \cdot P[B_i] \end{aligned}$$

Bayes' Rule describes probability of an event based on prior knowledge of conditions that might be related to that event.

$$P[A|B] = \frac{P[B|A] \cdot P[A]}{P[B]}$$

Two variables are independent if the occurrence of one does not affect the probability of the other. For two independent variables:

$$P[A \cap B] = P[A] \cdot P[B]$$

Remember that independent does not mean mutually exclusive.

$$\begin{aligned} P[A|B] &= \frac{P[A \cap B]}{P[B]} \\ &= \frac{P[A] \cdot P[B]}{P[B]} \\ &= P[A] \end{aligned}$$

3 Discrete Random Variables and Expected Value

A Bernoulli trial involves performing an experiment once and noting whether a particular event A occurs. We consider the experiment a success if the event occurs, and a failure otherwise. If k is the number of successes in n independent trials then the probability of k is given by the binomial probability law.

$$p_n(k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ where } \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

A random variable X is a function that assigns a real number to each outcome in the sample space of a random experiment. Discrete random variable assumes values from a countable set.

A probability mass function (pmf) of a discrete random variable is defined as:

$$p_X(x) = P[X = x] = P[\{\beta : X(\beta) = x\}] \text{ for } x \in \mathbb{R}$$

An expected value is the average value of a random variable X is defined as:

$$m_x = E[X] = \sum_{x \in S_X} x p_x(x)$$

From the above, we see that:

$$\text{For } Z = ag(X) + bh(X) + c,$$
$$E[Z] = aE[g(X)] + bE[h(X)] + c$$

4 Variance and Common Discrete Distributions

Variance represents a statistical measure of the spread between numbers in a data set, or, how far a set of numbers is spread out from its mean value.

$$\begin{aligned}\sigma_x^2 &= \text{VAR}[X] \\ &= E[(X - m_x)^2] \\ &= E[X^2] - m_x^2\end{aligned}$$

$E[X^2]$ is called the second moment of X, and we say that in general, $E[X^n]$ is the nth moment of X.

Variance and Conditional Probability

If X is a random variable and event B has occured, then the conditional expected value of $X|B$ is defined as:

$$m_{X|B} = \sum_k x_k p_x(x_k|B)$$

Thus the conditional variance of X given B is:

$$\text{VAR}[X|B] = E[(X - m_{X|B})^2|B] = E[X^2|B] - m_{X|B}^2$$

Random Variables

1. The Bernoulli random variable is the value of the indicator function I_A for the event A. We say $X = 1$ is A occurs and 0 otherwise. Binomial random variables is the number of successes in n Bernoulli trials. The variance of the Bernoulli random variable is given by:

$$\sigma_x^2 = np(1 - p)$$

2. Geometric random variables are the number of failures before the first success in a sequence of independent Bernoulli trials.

    (a) Probability mass function:

    $$p_k = p(1 - p)^k$$

    (b) Expected value function:

    $$E[X] = \frac{1 - p}{p}$$

    (c) Variance function:

    $$\text{VAR}[X] = \frac{1 - p}{p^2}$$

3. Negative binomial random variables are the number of trials until the $n$he success in a sequence of independent Bernoulli trials.

    (a) Probability mass function:

    $$p_k = \binom{k-1}{r-1} p^r (1 - p)^{k-r}$$

    (b) Expected value function:

    $$E[X] = \frac{r}{p}$$

    (c) Variance function:

    $$\text{VAR}[X] = \frac{r(1 - p)}{p^2}$$

4. Uniform random variables:
   (a) Probability mass function:
   $$p_k = \frac{1}{L}$$
   (b) Expected value function:
   $$E[X] = \frac{L+1}{2}$$
   (c) Variance function:
   $$\text{VAR}[X] = \frac{L^2-1}{12}$$
5. Zipf random variable:
   (a) Few outcomes occur frequently but most occur rarely, seen in applications like Huffamn decoding.
   (b) Probability mass function:
   $$p_k = \frac{1}{kc_L} \text{ where } c_L = \sum_{j=1}^{L} \frac{1}{j}$$
   (c) Expected value function:
   $$E[X] = \frac{L}{c_L}$$
   (d) Variance function:
   $$\text{VAR}[X] = \frac{L(L+1)}{2cL} - \frac{L^2}{c_L^2}$$