David Varodayan
CS109

Lecture Notes #7
January 22, 2020

Bernoulli and Binomial Random Variables

Based on a chapter by Chris Piech

There are some classic random variable abstractions that show up in many problems. At this point in the class you will learn about several of the most significant discrete distributions. When solving problems, if you are able to recognize that a random variable fits one of these formats, then you can use its precalculated probability mass function (PMF), expectation, variance, and other properties.

Random variables of this sort are called "parametric" random variables. If you can argue that a random variable falls under one of the studied parametric types, you simply need to provide parameters. A good analogy is a *class* in programming. Creating a parametric random variable is very similar to calling a constructor with input parameters.

Bernoulli Random Variable

A **Bernoulli random variable** is the simplest kind of random variable. It can take on two values, 1 and 0. It takes on a 1 if an experiment with probability p resulted in success and a 0 otherwise. Some example uses include a coin flip, a random binary digit, whether a disk drive crashed, and whether someone likes a Netflix movie.

If X is a Bernoulli random variable, denoted $X \sim \text{Ber}(p)$:

$$\text{Probability mass function:} \quad P(X = 1) = p$$
$$P(X = 0) = (1 - p)$$

$$\text{Expectation:} \quad E[X] = p$$

$$\text{Variance:} \quad \text{Var}(X) = p(1 - p)$$

Bernoulli random variables and **indicator variables** are two aspects of the same concept. As a review, a random variable I is called an indicator variable for an event A if $I = 1$ when A occurs and $I = 0$ if A does not occur. $P(I = 1) = P(A)$ and $E[I] = P(A)$. Indicator random variables *are* Bernoulli random variables, with $p = P(A)$.

Binomial Random Variable

A **binomial random variable** is random variable that represents the number of successes in n successive independent trials of a Bernoulli experiment. Some example uses include the number of heads in n coin flips, the number of disk drives that crashed in a cluster of 1000 computers, and the number of advertisements that are clicked when 40,000 are served.

If X is a Binomial random variable, we denote this $X \sim \text{Bin}(n, p)$, where p is the probability of success in a given trial. A binomial random variable has the following properties:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{if } k \in \mathbb{N}, 0 \leq k \leq n \text{ (0 otherwise)}$$
$$E[X] = np$$
$$\text{Var}(X) = np(1 - p)$$

Example 2

Let X = number of heads after a coin is flipped three times. $X \sim \text{Bin}(3, 0.5)$. What is the probability of each of the different values of X?

$$P(X = 0) = \binom{3}{0} p^0 (1 - p)^3 = \frac{1}{8}$$

$$P(X = 1) = \binom{3}{1} p^1 (1 - p)^2 = \frac{3}{8}$$

$$P(X = 2) = \binom{3}{2} p^2 (1 - p)^1 = \frac{3}{8}$$

$$P(X = 3) = \binom{3}{3} p^3 (1 - p)^0 = \frac{1}{8}$$

Example 3

When sending messages over a network, there is a chance that the bits will be corrupted. A Hamming code allows for a 4 bit code to be encoded as 7 bits, with the advantage that if 0 or 1 bit(s) are corrupted, then the message can be perfectly reconstructed. You are working on the Voyager space mission and the probability of any bit being lost in space is 0.1. How does reliability change when using a Hamming code?

Image we use error correcting codes. Let $X \sim \text{Bin}(7, 0.1)$.

$$P(X = 0) = \binom{7}{0}(0.1)^0(0.9)^7 \approx 0.468$$

$$P(X = 1) = \binom{7}{1}(0.1)^1(0.9)^6 = 0.372$$

$$P(X = 0) + P(X = 1) = 0.850$$

What if we didn't use error correcting codes? Let $X \sim \text{Bin}(4, 0.1)$.

$$P(X = 0) = \binom{4}{0}(0.1)^0(0.9)^4 \approx 0.656$$

Using Hamming Codes improves reliability by about 30%!