4. Random Variables

Chris Piech and Mehran Sahami

May 2017

A Random Variable (RV) is a variable that probabilistically takes on different values. You can think of an RV as being like a variable in a programming language. They take on values, have types and have domains over which they are applicable. We can define events that occur if the random variable takes one values that satisfy a numerical test (eg does the variable equal 5, is the variable less than 8). We often think of the probabilities of such events.

As an example, let's say we flip three fair coins. We can define a random variable Y to be the total number of "heads" on the three coins. We can ask about the probability of Y taking on different values using the following notation:

- $P(Y = 0) = 1/8$ (T, T, T)

- $P(Y = 1) = 3/8$ (H, T, T), (T, H, T), (T, T, H)

- $P(Y = 2) = 3/8$ (H, H, T), (H, T, H), (T, H, H)

- $P(Y = 3) = 1/8$ (H, H, H)

- $P(Y \geq 4) = 0$

Using random variables is a convenient notation technique that assists in decomposing problems. There are many different types of random variables (indicator, binary, choice, Bernoulli, etc). The two main families of random variable types are discrete and continuous.

Probability Mass Function

Probability mass functions (PMF) is a function that maps possible outcomes of a random variable to the corresponding probabilities. We can plot PMF graphs:
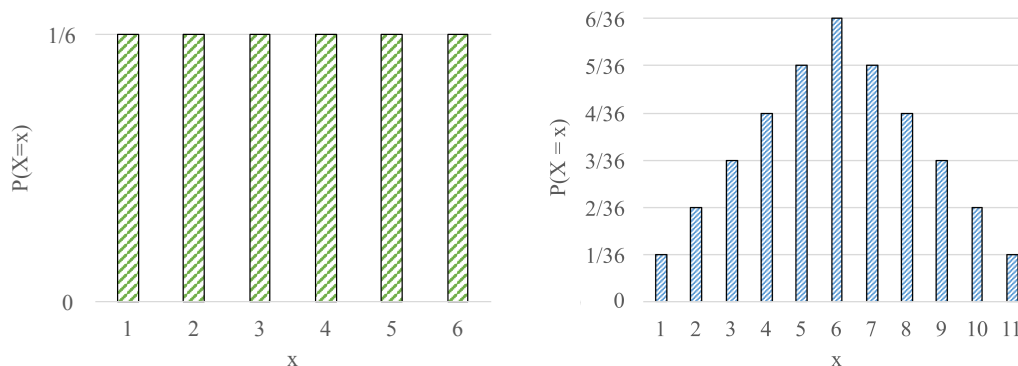

Figure: On the left, the PMF of a single 6 sided die roll. On the right, the PMF of the sum of two dice rolls.

Cumulative Distribution Function

For a random variable X, the Cumulative Distribution Function (CDF) is defined as:

$$F(a) = P(X \leq a) \text{ where } -\infty < a < \infty$$

Expected Value

The Expected Value for a discrete random variable X is defined as:

$$E[X] = \sum_{x:P(x)>0} xP(x)$$

It goes by many other names: Mean, Expectation, Weighted Average, Center of Mass, 1st Moment.

Example 1

Lets say you roll a 6-Sided Die and that a random variable X represents the outcome of the roll. What is the $E[X]$? This is the same as asking what is the average value.

$$E[X] = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6) = 7/2$$

Example 2

Lets say a school has 3 classes with 5, 10, and 150 students. If we randomly choose a class with equal probability and let X = size of the chosen class:

$$E[Y] = 5(1/3) + 10(1/3) + 150(1/3)$$
$$= 165/3 = 55$$

If instead we randomly choose a student with equal probability and let Y = size of the class the student is in

$$E[X] = 5(5/165) + 10(10/165) + 150(150/165)$$
$$= 22635/165 = 137$$

Example 3

Consider a game played with a fair coin which comes up heads with p = 0.5. Let n = the number of coin flips before the first "tails". In this game you win \$$2^n$. How many dollars do you expect to win? Let X be a random variable which represents your winnings.
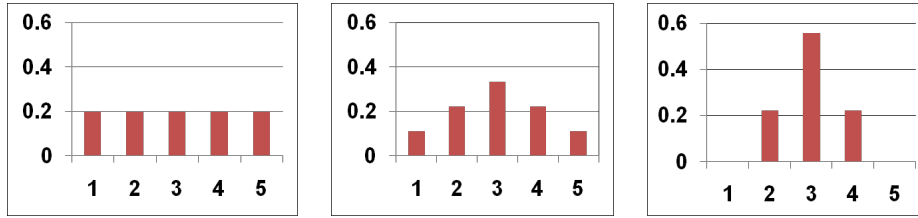
$$E[X] = \left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + \left(\frac{1}{2}\right)^4 2^3 + \cdots = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i+1} 2^i$$

$$= \sum_{i=0}^{\infty} \frac{1}{2} = \infty$$

Linearity

Expectations preserve linearity which means that $E[aX + b] = aE[X] + b$

Variance

Consider the following 3 distributions (PMFs)

All three have the same expected value, $E[X] = 3$ but the "spread" in the distributions is quite different. Variance is a formal quantification of "spread".

If X is a random variable with mean μ then the variance of X, denoted $Var(X)$, is: $Var(X) = E[(X-\mu)^2]$. When computing the variance often we use a different form of the same equation: $Var(X) = E[X^2] - E[X]^2$. Intuitively this is the weighted average distance of a sample to the mean.

Here are some useful identities for variance:

- $Var(aX + b) = a^2 Var(X)$
- Standard deviation is the root of variance: $SD(X) = \sqrt{Var(X)}$

Example 1

Let X = value on roll of a 6 sided die. Recall that $E[X] = 7/2$. First lets calculate $E[X^2]$

$$E[X^2] = (1^2)\frac{1}{6} + (2^2)\frac{1}{6} + (3^2)\frac{1}{6} + (4^2)\frac{1}{6} + (5^2)\frac{1}{6} + (6^2)\frac{1}{6} = \frac{91}{6}$$

Which we can use to compute the variance:

$$Var(X) = E[X^2] - (E[X])^2$$
$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Indicator Random Variable

A 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 variables are useful for cleaning up notation.