MAT 3701: Random Variables

February 5, 2013

Random variables assign (real) numerical values to the outcomes of a probability space. They can be thought of as variables whose values we don't choose or solve for, but which take values "at random." (Here the words "at random" do not mean that all values are equally likely, but only that they are not determined.)

It is crucial that the probability of any interval of values can be measured, hence the following definition. (Note that for simplicity we refer to a probability space S, although the probability space consists not only of the sample space, S, but also a σ-algebra of subsets (the events) and a probability measure on this σ-algebra, both of which are implied by this terminology.)

Definition. A *random variable* X on a probability space S is a function from S to $\mathbb{R}$, such that for every $x \in \mathbb{R}$, $\{\omega \in S : X(\omega) \le x\}$ is an event.

Exercise: Prove from this definition that, for any real number x, $\{\omega \in S : X(\omega) < x\}$ is an event, and also that $\{\omega \in S : X(\omega) = x\}$ is an event. (Note: the probability of either of these events may be zero.)

Notation: For brevity we usually write $P(X \le x)$ for the probability of the event $\{\omega \in S : X(\omega) \le x\}$. Similarly, we write $P(X = x)$ for the probability of the event $\{\omega \in S : X(\omega) = x\}$, etc.

Definition. Let X be a random variable. The function on $\mathbb{R}$ defined by $F(x) = P(X \le x)$ is called the *(cumulative) distribution function* of X.

1 Discrete Random Variables

We begin with random variables that take only countably many values, which are known as *discrete* random variables. To do calculations with discrete random variables, we will need the calculus of infinite series. To deal with continuous random variables, we will need integral and differential calculus.

Definition. A random variable X is *discrete* if the image of X is finite or countably infinite.

Exercise: Prove that if X is a discrete random variable, then $\{x \in \mathbb{R} : P(X = x) > 0\} \neq \emptyset$.

The assignment of a probability to each value of a discrete random variable can be thought of as giving it a density with respect to the probability measure, with various amounts of probability concentrated at different values. This point of view is exactly analogous to assigning mass density to the points along some dimension of a physical object, such as a metal rod of varying thickness. The places where the rod is thicker have more mass density, that is, mass per unit of length. The accumulation of mass along a rod is continuous, but if we imagine instead a sequence of point masses along a line, we have a situation analogous to the probability densities at the values of a discrete random variable. Note that a point on the line with no mass has density equal to zero.

Definition. Let X be a discrete random variable. The function with domain $\mathbb{R}$ defined by $f(x) = P(X = x)$ is called the *density function* of X. If $f(x) > 0$, then x is called a *possible value* of X.

Remark. Unfortunately, the terminology of probability theory is not consistent and can be confusing. Some authors refer to the density function of a random variable as its *probability distribution*, especially in the discrete case. Since this terminology is rather natural – after all, this function describes how the probability is distributed among the outcomes – we will sometimes use it as well. The term *probability density* is also used (although this should not cause any confusion). Our text uses "probability distribution" in the discrete case, "probability density" in the continuous case.

In order to avoid confusion, we will always use the phrase "distribution function" when we mean the function F defined by $F(x) = P(X \leq x)$, and I will often add the qualifier "cumulative" in parentheses for further emphasis, as I did in the definition above. I addition, we will always notate a distribution function by an upper-case letter and a density function by a lower-case letter.

For a discrete random variable, the distribution function is readily computed from the density function by summation, since the number of possible values is countable. Let $\{x_1, x_2, x_3, \ldots\}$ be the distinct possible values of a discrete random variable X (where the index terminates if this set is finite). Then $F(x) = \sum_{i:x_i \leq x} f(x)$. The distribution function of a discrete random variable is a step function. It is constant except at the possible values of the random variable, at which it "jumps." Note that all distribution functions are continuous from the right, since if $(x_i)_{i=1}^{\infty}$ is a decreasing sequence of numbers converging to x from above, then the sets $E_i = \{\omega \in S : X(\omega) \leq x_i\}$ are nested: $E_1 \supseteq E_2 \supseteq E_3 \ldots$. With a little analysis work and one of the results from our first problem set, it follows that $\lim_{h \to 0^+} F(x + h) = F(x)$. But the distribution functions of discrete random variables are *not* continuous from the left at the possible values of the random variable. (See Figure 1.)

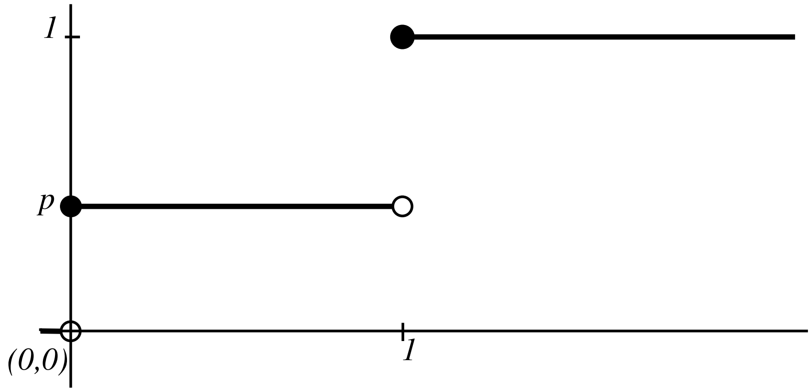

Figure 1: The distribution function of a Bernoulli random variable.

2 Continuous Random Variables

At the other extreme from discrete random variables are *continuous random variables*. Rather than taking a discrete, countable set of values, they take a continuum of values, with zero probability of assuming any particular one. An example is the time at which a radioactive atom decays. It can decay at any moment, but its probability of decaying at any specific moment, as we have seen, must be zero.

The precise definition of a continuous random variable is simply that its distribution function is continuous at every point (that is, from both left and right). There is no reason it needs to be smooth, just continuous. For example, suppose a point in the plane is picked at random from the union of the rectangles $[0,1] \times [0,2]$ and $[2,3] \times [0,3]$. Let X be the x-coordinate of the point chosen. Then $F_X(x) = 0$ for $x < 0$, increases linearly with slope $\frac{2}{5}$ on $[0,1]$, is constant from 1 to 2, and then increases linearly again from 2 to 3, this time with slope $\frac{3}{5}$. It reaches its maximum value - one - at $x = 3$ and is constant on $[3, \infty)$. So F_X is not differentiable (it has “corners”) at $x = 0, 1, 2$, or 3. (See Figure 2.)

Exercise: Calculate and sketch the distribution function of the y-coordinate of the point chosen.

Definition. A random variable X is *continuous* if its distribution function, $F_X(x) = P(X \leq x)$, is continuous at every point of $\mathbb{R}$.

A random variable does not have to be either discrete or continuous. For example, suppose I pick a point at random from the interval $[0,1]$, but then I also flip a fair coin, and if the coin comes up heads I change my choice to $\frac{\pi}{5}$. The random variable that models the number chosen

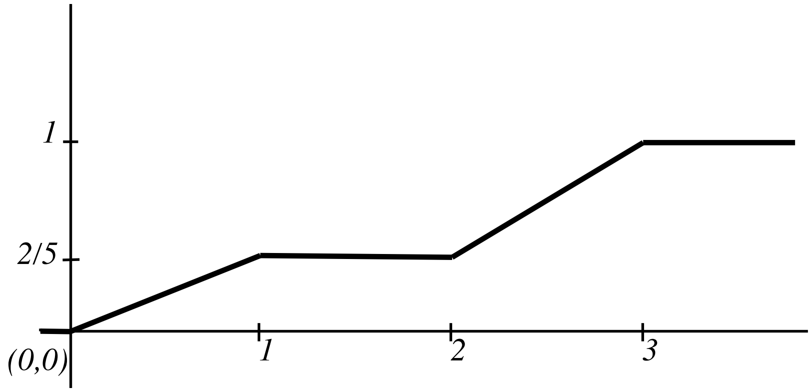

Figure 2: The distribution function of a continuous (but not differentiable) random variable.

takes the value $\frac{\pi}{5}$ with probability $\frac{1}{2}$, so the value of its distribution function rises linearly with slope $\frac{1}{2}$ from 0 to $\frac{\pi}{10}$ as x goes from 0 to $\frac{\pi}{5}$, jumps by $\frac{1}{2}$ at $x = \frac{\pi}{5}$, and then continues to rise linearly with slope $\frac{1}{2}$ from $x = \frac{\pi}{5}$ to $x = 1$. The probability that I pick a number between 0 and $\frac{1}{2}$ is $\frac{1}{4}$, but the probability that I pick a number between $\frac{1}{4}$ and $\frac{3}{4}$ is $\frac{3}{4}$.

Random variables that are neither discrete nor continuous are harder to study. (To understand them requires an advanced area of analysis known as *measure theory.*) We will restrict our attention to those that are discrete or continuous, except in simple cases. Just as we require the theory of infinite series to define and study the expectation, variance, and other properties of discrete random variables, we require integral calculus to define and study the corresponding properties of continuous random variables. Just as calculating with joint distributions of discrete random variables involves multiple sums, calculating with joint distributions of continuous random variables involves multiple integrals. *You must review calculus, including infinite series, single integrals, vectors, and multiple integrals! You should own or otherwise have ready access to a good calculus reference!*