Notes on random variables, density functions, and measures

September 29, 2010

1 Probability density functions and probability measures

We defined a probability density function (pdf for short) as a function $f : \mathbb{R} \to [0, \infty)$ satisfying

$$\int_{\mathbb{R}} f(x)dx \ = \ 1.$$

(Of course, this means that f must be Lebesgue-integrable, and hence measurable.) Likewise, we defined a "probability mass function", which I will also refer to as a pdf, as a function $p : \{a_1, a_2, a_3, ...\} \to [0, 1]$ satisfying

$$\sum_i p(a_i) \ = \ 1.$$

So, mass functions are analogues of density functions for discrete (finite, or possibly countably infinite) sets of real numbers.

What a pdf gives us is a probability measure for subsets of the real numbers. Specifically, suppose $A \in \mathcal{B}(\mathbb{R})$. Then, we can define

$$\mathbb{P}(A) \ := \ \int_{\mathbb{R}} 1_A(x) f(x)dx \ = \ \int_A f(x)dx. \quad (1)$$

And we can do the same in the discrete case, for if $A \subseteq \{a_1, a_2, ...\}$, then we can define

$$\mathbb{P}(A) \ := \ \sum_{a_i \in A} p(a_i).$$

There is a way to unify both of these types of measures: Basically, let $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty)$. Then, for $A \in \mathcal{B}(R)$ we use the notation

$$\int_{\mathbb{R}} 1_A(x) d\mu(x), \quad \text{or, alternatively} \quad \int_R 1_A(x)\mu(dx).$$

to denote $\mu(A)$. Once we have a way to define integrals for such indicator functions, we can do so for linear combinations of indicator functions; and then we can generalize further, producing a kind of integral "weighted by μ". To see that this unifies both the probability density function and the probability mass function, first note that if $f(x)$ is a pdf, then we can just define μ using (1), letting $\mu(A) = \mathbb{P}(A)$. The discrete case is more difficult, because we have to go from a mass function defined on a discrete (but possibly countably infinite) set to a measure defined on $\mathcal{B}(\mathbb{R})$: Well, what we can do is just let, for $A \in \mathcal{B}(\mathbb{R})$,

$$\mu(A) \ := \ \sum_{a_i \in A} p(a_i).$$

(which is weird, since we are relating discrete sums to measure on arbitrary subsets of $\mathcal{B}(\mathbb{R})$.)

2 Random variables

In general, sample spaces are not collections of numbers that we can measure using pdf's. What we need is a bridge from arbitrary sample spaces to the real numbers; and so-called "random variables" provide that bridge. Basically, a random variable is a *mapping*

$$X \ : \ \mathcal{S} \ \to \ \mathbb{R},$$

such that X is measurable, where here $\mathcal{S}$ denotes a sample space for which we have a sigma algebra Σ. Basically, the measurability requirement means that for each $A \in \mathcal{B}(\mathbb{R})$ we have that $X^{-1}(A) \in \Sigma$.

Why would we want X to be measurable? Well, we want to be able to compute probabilities like, say,

$$\mathbb{P}(X = 4), \text{ or } \mathbb{P}(X < -1), \text{ or, ingeneral, } \mathbb{P}(X \in A), \text{ where } A \in \mathcal{B}(\mathbb{R}).$$

So we need these events like "$X = 4$" and "$X < -1$" to correspond to some things that we can measure with $\mathbb{P}$; and the only things that $\mathbb{P}$ *can* measure are those elements of Σ (which, recall, are subsets of $\mathcal{S}$) – hence the need to have $X^{-1}(A) \in \Sigma$.

What *is* a random variable? It may seem a little confusing at first, but random variables are not really "variables" at all, nor are they "random" for that matter. Basically, probabilists have removed 'time' from the picture, so that r.v.'s, like every other mathematical object, reside in some eternal Platonic realm: As we said, r.v.'s are maps or functions; but in many cases to wrap our heads around them it's best to think of them as being like "input data from some experiment, that can change each time we perform the experiment."

It is customary to associate to X a pdf (or more generally, a measure) $f(x)$, so that

$$\mathbb{P}(X \in A) \ = \ \int_A f(x)dx.$$

3 Passing from one r.v. to another

Consider the following problem: Suppose that X is a r.v. having pdf

$$f(x) \ = \ e^{-x^2/2}/\sqrt{2\pi}.$$

Let $Y = 2X - 3$. Find the pdf for Y.

To address this, we introduce the notion of a "cumulative distribution function". Usually, we denote this with CAPITAL LETTERS; i.e. if $f(x)$ is the pdf, then $F(x)$ is the cdf – but, because $f(x)$ in the above problem happens to correspond to a special distribution called a Gaussian or "normal distribution", we use the symbol $\Phi(x)$ instead. In general, though, a cdf is a function $F(x)$ defined by

$$F(x) \ := \ \int_\infty^x f(x)dx.$$

Or, if we have an r.v. to begin with, where no mention is made of $f(x)$, we define it as

$$F(x) \ := \ \mathbb{P}(X \leq x).$$

Now, if we are in the former case (we *have* $f(x)$), then it turns out that

$$F'(x) \; = \; f(x).$$

In other words, to work out a pdf, you can start with a cdf and take its derivative.

Let's now apply this to the above problem: Let $G(y)$ denote the cdf for Y, so that

$$G(y) \; = \; \mathbb{P}(Y \leq y).$$

Next, rewrite this in terms of X, so that

$$\begin{aligned} G(y) \; = \; \mathbb{P}(Y \leq y) \; = \; \mathbb{P}(2X - 3 \leq y) \; = \; \mathbb{P}(X \leq (y+3)/2) \; &= \; \Phi((y+3)/2) \\ &= \; \int_{-\infty}^{(y+3)/2} e^{-x^2/2}/\sqrt{2\pi} dx. \end{aligned}$$

Now, taking a derivative, we find that the pdf for Y, denote by $g(y)$, say, satisfies

$$g(y) \; = \; G'(y) \; = \; \frac{dG(y)}{dy} \; = \; e^{-(y+3)^2/8}/2\sqrt{2\pi}.$$

The extra factor of 2 here in the denominator comes from the chain rule when we differentiate $(y+3)/2$ w.r.t. y.