Continuous random variables

Math 477

December 19, 2014

1 Introduction

Imagine you're dropping a needle onto the interval $[0, 1]$ so that its sharp point falls straight down. Suppose you choose your drop point uniformly randomly between 0 and 1 (we'll clarify precisely what we mean by uniformly later on, for now let's just say without bias to any particular location). Let X be the location of the sharp point of the needle when it hits. Then X is a random variable taking values between 0 and 1.

What is $P(X = 1/2)$? Intuitively the probability is small. Let us ask instead the probability of the needle falling in between the interval $(1/2 - 1/n, 1/2 + 1/n)$ for some $n \geq 2$. Since you drop the needle uniformly, it is reasonable to conclude that

$$P(X \in (1/2 - 1/n, 1/2 + 1/n)) = 2/n.$$

But since the event $\{X = 1/2\}$ implies the event $\{1/2 - 1/n < X < 1/2 + 1/n\}$ we must conclude

$$P(X = 1/2) \leq 2/n,$$

for all n. Let $n \to \infty$ we get $P(X = 1/2) = 0$.

At first this result might seem surprising to some, because it seems *possible* for us to hit the point 1/2 with the needle point. This is true, but only if you intentionally drop the needle right on top of the point 1/2. Otherwise you may be convinced that you never hit the point 1/2 if you choose a random drop point of the needle.

Another example may help you to see the idea. Suppose the bus arrives at the Hill bus stop uniformly from 0 to 10 minutes, counting from the moment you arrive at the bus stop. Then the probability that the bus arrives at exactly 5 minutes is 0 (It may arrive at 5 minutes 1 sec, but it would never arrive at exactly 5 minutes).

On the other hand, observe that the probability that your wait time is between 4.5 minutes to 5.5 minutes is positive (and equals to 1/10 if the bus arrives uniformly).

In these examples, we call the random variables described continuous random variables.

2 Representation of the probability distribution of a continuous random variable

2.1 The density function

To capture the idea of a random variable whose probability of hitting an exact value is 0, but whose probability of belonging in an interval is positive, we use a function f, defined on the real line with two properties:

$$\begin{aligned} f(x) &\geq 0, \forall x \\ \int_{-\infty}^{\infty} f(x)dx &= 1. \end{aligned}$$

You don't have to look very hard for such a function f. Here's an example:

$$\begin{aligned} f(x) &= \frac{1}{b-a}, a \leq x \leq b \\ f(x) &= 0 \text{ otherwise } . \end{aligned}$$

Once we have such a function f, then we just define a continuous random variable X corresponding to f as followed:

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

The fact that $f(x) \geq 0$ assures that $P(a \leq X \leq b) \geq 0$ for all a, b. The fact that $\int_{-\infty}^{\infty} f(x)dx = 1$ assures that $P(a \leq X \leq b) \leq 1$, as we would expect from any probability distribution.

We will refer to such function f as the probability density function (or p.d.f, or just density) of X. It serves as the "signature" of the RV X.

2.2 Some implications

We see that if $a = b$

$$P(a \leq X \leq b) = \int_a^a f(x)dx = 0.$$

Moreover,

$$P(a \le X < b) = P(a \le X \le b) - P(X = b) = P(a \le x \le b) = \int_a^b f(x)dx.$$

Similarly,

$$P(a < X \le b) = P(a < X < b) = P(a \le X < b) = P(a \le X \le b) = \int_a^b f(x)dx.$$

Thus less than or equal to or strictly less than give you the same probability with continuous random variable (this is NOT the case with discrete random variable). That is, if X is discrete and $P(X = k) > 0$ then

$$P(X \le k) = P(X < k) + P(X = k) > P(X < k).$$

Finally, letting $a \to -\infty$ in the above expression we get

$$P(X \le b) = P(-\infty < X \le b) = \int_{-\infty}^b f(x)dx.$$

Letting $b \to \infty$ we get

$$P(a \le X) = P(a \le X < \infty) = \int_a^\infty f(x)dx.$$

2.3 Example

Example 2.1. *The life time in years of the Iphone is a random variable having the pdf*

$$\begin{aligned} f(x) &= 0, x \le 2 \\ &= 100/x^2, x > 100. \end{aligned}$$

What's the probability that a particular Iphone dies before the 3rd year?

2.4 From probability to density

From the above, we have

$$P(X \le x) = \int_{-\infty}^x f(v)dv.$$

Thus by the Fundamental theorem of calculus (since $f(-\infty) = 0$):

$$f(x) = f(x) - f(-\infty) = d/dx P(X \le x).$$

3 Function of continuous RV

Just as the discrete case, if X is a continuous RV, $f(X)$ for some function f is a continuous RV. To determine this new RV, we need to know its pdf. The idea essentially relies on finding "f^{-1}" (in quote because usually f^{-1} is not a function, nevertheless one can still identify a map from the range to domain of f, called f^{-1}).

3.1 A review of the FTC

As you have seen from above, figuring out the pdf requires us to apply the Fundamental theorem of calculsus. In the case of a function of a RV, we will need to apply the FTC with the chain rule.

$$d/dx \int_a^{g(x)} f(v)dv = f(g(x))g'(x).$$

If $g_1(x) \leq g_2(x)$:

$$d/dx \int_{g_1(x)}^{g_2(x)} f(v)dv = f(g_2(x))g_2'(x) - f(g_1(x))g_1'(x).$$

3.2 Some examples

Example 3.1. *Let X be a RV with density $f(x)$. What is the density of $2X$?*

Ans:

$$P(2X \leq x) = P(X \leq x/2) = \int_{-\infty}^{x/2} f(v)dv.$$

Therefore,

$$f_{2X}(x) = 1/2 f(x/2).$$

Example 3.2. *Let X be a RV with density $f(x)$. What is the density of $2X+1$?*

Ans:

$$P(2X+1 \leq x) = P(X \leq (x-1)/2) = \int_{-\infty}^{(x-1)/2} f(v)dv.$$

Therefore,

$$f_{2X+1}(x) = 1/2 f((x-1)/2).$$

Example 3.3. *Let X be a RV with density $f(x)$. What is the density of X^2?*

Ans: For $x \geq 0$

$$P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f(v)dv.$$

Thus,

$$f_{X^2}(x) = \frac{f(\sqrt{x}) + f(-\sqrt{x})}{2\sqrt{x}}.$$

4 Expectation

Similar to the discrete case, we define

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx.$$

So the question is, how to compute $E(g(X))$? We can use the technique listed in the previous section, that is to find the density of $g(X)$:

$$E(g(X)) = \int_{-\infty}^{\infty} x f_{g(X)}(x)dx.$$

But it can be cumbersome to find $f_{g(X)}$. Instead, we have a similar theorem as the discrete case:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)dx.$$

The proof of the above is a bit more involved. First we state the following result: For a non-negative RV X:

$$E(X) = \int_0^{\infty} P(X > x)dx.$$

Reason:

$$\begin{aligned} E(X) &= \int_0^{\infty} x f_X(x)dx = \int_0^{\infty} [\int_0^x dy] f_X(x)dx \\ &= \int_0^{\infty} \int_y^{\infty} f_X(x)dxdy = \int_0^{\infty} P(X > y)dy. \end{aligned}$$

Now we can show the theorem for the case $g(x) \geq 0$:

$$\begin{aligned} E(g(X)) &= \int_0^\infty P(g(X) \geq x)dx \\ &= \int_0^\infty P(g(X) \geq x)dx \\ &= \int_0^\infty \int_{\{y:g(y)\geq x\}} f(y)dydx \\ &= \int_{-\infty}^\infty \int_0^{g(y)} dxf(y)dy \\ &= \int_{-\infty}^\infty g(y)f(y)dy. \end{aligned}$$

5 Variance

Now that we have define $E(g(X))$ and show how to compute it, we can define $Var(X)$, similar to the discrete case:

$$Var(X) = E(X^2) - E^2(X).$$