Lecture Notes for Math 210 – 26 October 2007

Shannon Starr

26 October 2007

Chapter 5: Tools in Probability Theory

Conditional Probability and Conditional Expectation for Continuous Random Variables

Suppose X is a random variable.

The c.d.f. of X is a function $F : \mathbb{R} \to [0,1]$, defined as

$$F_{\mathsf{X}}(a) \,=\, \mathbf{P}\{\mathsf{X} \leq a\}\,.$$

(I'm changing notation to $F_{\mathsf{X}}(a)$ instead of $G_{\mathsf{X}}(a)$.)

Then X is called a continuous random variable if

$$f_{\mathsf{X}}(x) \,=\, \frac{d}{dx} F_{\mathsf{X}}(x)\,,$$

exists at "almost all" points $x \in \mathbb{R}$, and if

$$F_{\mathsf{X}}(a) \,=\, \int_{-\infty}^{a} f_{\mathsf{X}}(x)\, dx\,,$$

at *all* points $a \in \mathbb{R}$.

<u>*Example 1.*</u> The exponential random variable with "rate" $\lambda > 0$ is defined as X such that

$$F_{\mathsf{X}}(a) \;=\; \begin{cases} 0 & \text{if } a < 0 \;, \\ 1 - e^{-\lambda a} & \text{if } a \geq 0 \;. \end{cases}$$

Then note that

$$\frac{d}{dx} F_{\mathsf{X}}(x) \;=\; \begin{cases} 0 & \text{if } x < 0 \;, \\ \text{D.N.E.} & \text{if } x = 0 \;, \\ \lambda e^{-\lambda x} & \text{if } x > 0 \;. \end{cases}$$

So it is impossible to say the value of $f_{\mathsf{X}}(0)$ from the derivative of $F_{\mathsf{X}}(x)$.

Related to this, there is a jump discontinuity in $f_{\mathsf{X}}(x)$ of size λ at $x = 0$.

But, no matter what value we plug in for $f_{\mathsf{X}}(0)$, we *do* have

$$\int_{-\infty}^{a} f_{\mathsf{X}}(x)\, dx \;=\; 0 \;=\; F_{\mathsf{X}}(a) \qquad \text{if } a \leq 0 \;,$$

and, for $a \geq 0$,

$$\int_{-\infty}^{a} f_{\mathsf{X}}(x)\, dx \;=\; \int_{0}^{a} \lambda e^{-\lambda x}\, dx \;=\; -e^{-\lambda x}\Big|_{0}^{a} \;=\; -e^{-\lambda a} + 1 \;=\; F_{\mathsf{X}}(a)\,.$$

The reason is that 1 single value of $f_{\mathsf{X}}(x)$ does not change the integral.

So the fact that $f_{\mathsf{X}}(0)$ is not clearly specified is not a problem.

For a continuous random variable, $f_{\mathsf{X}}(x)$ is called the probability density function, p.d.f.

Joint distributions:

Remember that random variables X and Y are defined as functions from some sample space Ω to $\mathbb{R}$.

Suppose that there is an event space $\mathcal{F}$ such that both X and Y are $\mathcal{F}$-measurable.

Also suppose there is a probability "measure" $P : \mathcal{F} \to \Omega$.

Then we define the joint c.d.f. of X and Y as

$$F_{\mathsf{X},\mathsf{Y}}(a,b) \;=\; \mathbf{P}\{\mathsf{X} \le a\,,\; \mathsf{Y} \le b\}\,.$$

We say that X and Y are jointly continuous if there is a function $f_{\mathsf{X},\mathsf{Y}}(x,y)$ such that

$$F_{\mathsf{X},\mathsf{Y}}(a,b) \;=\; \int_{-\infty}^{a}\int_{-\infty}^{b} f_{\mathsf{X},\mathsf{Y}}(x,y)\,dx\,dy\,,$$

for all $a, b \in \mathbb{R}$.

(Then one knows that $f_{\mathsf{X},\mathsf{Y}}(x,y) = \frac{\partial}{\partial x}(\frac{\partial}{\partial y}(F_{\mathsf{X},\mathsf{Y}}(x,y)))$ at "almost all" points x and y.)

One defines the marginal distributions as

$$f_{\mathsf{X}}(x) \;=\; \int_{-\infty}^{\infty} f_{\mathsf{X},\mathsf{Y}}(x,y)\,dy\,,$$

which is a p.d.f. for X, alone, and similarly for Y.

The conditional p.d.f. for X, given Y, is then defined as

$$f_{\mathsf{X}|\mathsf{Y}}(x|y) \;=\; \frac{f_{\mathsf{X},\mathsf{Y}}(x,y)}{f_{\mathsf{Y}}(y)}\,,$$

which is defined everywhere that $f_{\mathsf{Y}}(y) \neq 0$.

Then, at least heuristically, one should expect that

$$\mathbf{E}[\mathsf{X} \,|\, \mathsf{Y}] \;=\; g(\mathsf{Y})\,,$$

where

$$g(y) \;=\; \int_{-\infty}^{\infty} x \, f_{\mathsf{X}|\mathsf{Y}}(x|y)\, dx\,.$$

We just write this as

$$\mathbf{E}[\mathsf{X} \,|\, \{\mathsf{Y} = y\}] \;=\; \int_{-\infty}^{\infty} x \, f_{\mathsf{X}|\mathsf{Y}}(x|y)\, dx\,.$$

More generally, we expect that for any function $h(x)$

$$\mathbf{E}[h(\mathsf{X}) \,|\, \{\mathsf{Y} = y\}] \;=\; \int_{-\infty}^{\infty} h(x)\, f_{\mathsf{X}|\mathsf{Y}}(x|y)\, dx\,.$$

This formula is basically true, although one has to be somewhat careful if $f_{\mathsf{Y}}(y)$ is ever 0.

We will skip the technicalities necessary to make this rigorous.

Example. A preliminary definition of Brownian motion is a family of random variables Z_t, defined for all $t \geq 0$, such that $\mathsf{Z}_0 = 0$, and for every $0 \leq s < t$, the increment $[\mathsf{Z}_t - \mathsf{Z}_s]$ is a normal random variable, with mean 0 and variance equal to $t - s$, which is independent of all Z_r with $r \leq s$.

Q: Find $f_{\mathsf{Z}_s|\mathsf{Z}_t}(x, y)$, for $s < t$.

Let us start by calculating $f_{\mathsf{Z}_s,\mathsf{Z}_t}(x, y)$.

It will actually be easier to start by calculating $f_{\mathsf{Z}_s,\mathsf{Z}_t}(x, x+y)$, and then substitute

$y - x$ for y.

The reason is that we know Z_s by itself is a normal random variable with variance s:

$$f_{\mathsf{Z}_s}(x) \;=\; \frac{e^{-\frac{x^2}{2s}}}{\sqrt{2\pi s}}\;.$$

But also, $\mathsf{Z}_t - \mathsf{Z}_s$ is independent of Z_s and it is normal with variance $t - s$:

$$f_{\mathsf{Z}_t - \mathsf{Z}_s}(y) \;=\; \frac{e^{-\frac{y^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}}\;.$$

From this we can deduce that

$$f_{\mathsf{Z}_s,\mathsf{Z}_t}(x, x+y) \;=\; f_{\mathsf{Z}_s}(x)\cdot f_{\mathsf{Z}_t-\mathsf{Z}_s}(y) \;=\; \frac{1}{2\pi\sqrt{s(t-s)}}\, e^{-\frac{x^2}{2s}-\frac{y^2}{2(t-s)}}\;.$$

So,

$$f_{\mathsf{Z}_s,\mathsf{Z}_t}(x, y) \;=\; \frac{1}{2\pi\sqrt{s(t-s)}}\, e^{-\frac{x^2}{2s}-\frac{(y-x)^2}{2(t-s)}}\;.$$

We also know that Z_t by itself is normal with variance t:

$$f_{\mathsf{Z}_t}(y) \;=\; \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}}\;.$$

So,

$$\begin{aligned} f_{\mathsf{Z}_s|\mathsf{Z}_t}(x|y) \;&=\; \frac{f_{\mathsf{Z}_s,\mathsf{Z}_t}(x,y)}{f_{\mathsf{Z}_t}(y)} \\ &=\; \frac{\sqrt{2\pi t}}{2\pi\sqrt{s(t-s)}}\, e^{-\frac{x^2}{2s}-\frac{(y-x)^2}{2(t-s)}+\frac{y^2}{2t}}\;. \end{aligned}$$

By expanding the quadratic functions in the exponent, you get

$$f_{\mathsf{Z}_s|\mathsf{Z}_t}(x|y) \;=\; \frac{1}{\sqrt{2\pi[s(t-s)/t]}}\, e^{-\frac{t}{2s(t-s)}\left[x-\frac{s}{t}y\right]^2}\,.$$

In other words, if we condition Brownian motion at time s on Brownian motion Z_t at time t, then the conditioned random variable is normally distributed but with mean $\frac{s}{t}\mathsf{Z}_t$ and variance $\frac{s(t-s)}{t}$.

Or to put it yet another way: if we want to simulate a finite dimensional approximation to Brownian motion on the interval $[0, T]$, we can start by simulating Z_T by a normal random variable, then we can fill in $\mathsf{Z}_{T/2}$ by taking the conditional probability as stated above. Then we can iteratively fill-in the picture by taking $\mathsf{Z}_{T/4}$ and $\mathsf{Z}_{2T/4}$ and so, on, splitting each subinterval in half at each step of the iteration.