Notes: Joint Probability and Independence for Continuous RV's

CS 3130 / ECE 3530: Probability and Statistics for Engineers

October 24, 2017

Joint Probability Density Functions. Remember that we can use a continuous random variable X to define events such as $\{a \leq X \leq b\}$, which is the event "X landed somewhere between a and b". Also, remember that probability of such an event is computed by integrating the pdf for X, $f(x)$:

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

Just as in the discrete case, we can extend this concept to the case where we consider the joint probability of two continuous random variables. Let X and Y be two continuous random variables. Now an event for both random variables might be something of the form: $\{a \leq X \leq b\} \cap \{c \leq Y \leq d\}$, meaning "the pair (X, Y) fell inside the box $[a, b] \times [c, d]$". The joint pdf for X and Y is a function $f(x, y)$ satisfying

1. $f(x, y) \geq 0, \quad$ for all x, y
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$
3. $P(a \leq X \leq b,\, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)\, dx\, dy$

Double Integrals. Computing probabilities of events for joint random variables requires double integrals like the one in rule #3 above. Double integrals are not that scary (if you can integrate once, you can integrate twice!). Here is the procedure for evaluating the integral

$$P(a \leq X \leq b,\, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

1. First evaluate the "inside" integral: $\int_a^b f(x, y)\, dx$. Treat the variable y as if it were a *constant* (like you would the number 2). Evaluating the integral over the interval $[a, b]$ will result in a function $F(y)$, where the variable x does not appear.
2. Next evaluate the "outside" integral: $\int_c^d F(y)\, dy$. This will result in the answer you are looking for.

<u>Example:</u> Consider random variables X, Y with joint probability density function:

$$f(x, y) = \begin{cases} 2y \sin(x) & \text{for } 0 \leq x \leq \frac{\pi}{2},\, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $P(0 \leq X \leq \frac{\pi}{4},\, 0.5 \leq Y \leq 1)$?

Let's go through the two steps above to evaluate this probability as a double integral:

$$
\begin{aligned}
&P(0 \le X \le \tfrac{\pi}{4},\, 0.5 \le Y \le 1) = \\
&= \int_{0.5}^{1} \int_{0}^{\pi/4} 2y \sin(x)\, dx\, dy && \text{notice how the integration bounds are set up} \\
&= \int_{0}^{1} -2y \cos(x) \Big|_{x=0}^{x=\pi/4} dy && \text{integrating in } x\text{, and considering } y \text{ constant} \\
&= \int_{0.5}^{1} -2y \left(\sqrt{2}/2 - 1\right) dy && \text{evaluating } \cos(x) \text{ from 0 to } \pi/4 \\
&= -y^2 \left(\sqrt{2}/2 - 1\right) \Big|_{y=0.5}^{y=1} && \text{integrating in } y \\
&= -\left(1 - \frac{1}{4}\right)\left(\sqrt{2}/2 - 1\right) && \text{evaluating } y^2 \text{ from 0.5 to 1} \\
&= \frac{6 - 3\sqrt{2}}{8} \approx 0.2197 && \text{simplifying final answer}
\end{aligned}
$$

Marginal Probabilities. Remember that for joint discrete random variables, the process of "marginalizing" one of the variables just means to sum over it. For continuous random variables, we have the same process, just replace a sum with an integral. So, to get the pdf for X or the pdf for Y from the joint pdf $f(x, y)$, we just integrate out the other variable:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx.$$

Example: The marginal pdf's for the above example are:

$$f_X(x) = \int_0^1 f(x, y)\, dy = \int_0^1 2y \sin(x)\, dy = y^2 \sin(x) \Big|_{y=0}^{y=1} = \sin(x)$$

$$f_Y(y) = \int_0^{\frac{\pi}{2}} f(x, y)\, dx = \int_0^{\frac{\pi}{2}} 2y \sin(x)\, dx = -2y \cos(x) \Big|_{x=0}^{x=\frac{\pi}{2}} = 2y$$

Conditional Probability. Conditional probability works much like the discrete case. For random variables X, Y with joint pdf $f(x, y)$ and marginal pdf's $f_X(x)$ and $f_Y(y)$, we define the **conditional density function**:

$$f(x|Y = y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & \text{for all values of } y \text{ where } f_Y(y) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Now, conditional probabilities are found by integrating $f(x|Y = y)$:

$$P(a \le X \le b \,|\, Y = y) = \int_a^b f(x \,|\, Y = y)\, dx.$$

Example: Again, using the joint pdf in the example above, what is the conditional density $f\,(x \,|\, Y = y)$?

We just need to use the formula we found above for $f_Y(y)$:

$$f\,(x \,|\, Y = y) = \frac{f(x, y)}{f_Y(y)} = \frac{2y \sin(x)}{2y} = \sin(x) \quad \text{(for } 0 \le x \le \pi/2\text{, 0 otherwise).}$$

Notice that this turned out to be just $f_X(x)$. Let's try a more interesting example.

In-class Exercise: Given the joint density of X, Y:

$$f(x,y) = \begin{cases} x^2 + \frac{4}{3}xy + y^2 & \text{for } (x,y) \in [0,1] \times [0,1], \\ 0 & \text{otherwise.} \end{cases}$$

What are the marginal densities $f_X(x)$, $f_Y(y)$, and what are the conditional densities $f(x \,|\, Y = y)$ and $f(y \,|\, X = x)$?

Independence. Again, independence is just the same as in the discrete case, we just have pdf's instead of pmf's. The three equivalent definitions for independence of X and Y are:

For all possible $x \in \mathbb{R}$ and $y \in \mathbb{R}$:

$$f(x,y) = f_X(x) f_Y(y)$$
$$f(x \,|\, Y = y) = f_X(x)$$
$$f(y \,|\, X = x) = f_Y(y)$$

In-class Exercise: For the two joint densities in the examples above, determine if X and Y are independent.