Conditional Distributions: Jointly Continuous

X, Y jointly continuous.

Note: Since Y is continuous, we have P(Y = y) = 0 for all y, so we cannot condition on Y = y exactly as above. But it turns out, we can define conditional pdf in analogy with conditional pmf.

Conditional Probability Density Function of X Given Y = y

\[ f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{for all } x, y \text{ with } f_Y(y) > 0. \]

Conditional Probability That X ∈ A Given Y = y

\[ P(X \in A | Y = y) = \int_A f_{X|Y}(x|y) \, dx \quad \text{when } f_Y(y) > 0. \]

Conditional Expectation of X Given Y = y

\[ E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) \, dx \quad \text{when } f_Y(y) > 0. \]

LOTUS: \[ E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X,Y}(x,y) \, dx \quad \text{when } f_Y(y) > 0. \]

Law of Total Probability and Total Expectation

\[ P(X \in A) = \int_{-\infty}^{\infty} P(X \in A | Y = y) f_Y(y) \, dy \]

\[ E[X] = \int_{-\infty}^{\infty} E[X|Y = y] f_Y(y) \, dy \]

Proof: Exercise (similar to discrete)
Solution to Exercise:

- \( P(\mathbf{X} \in A) = \int_A f_X(x) \, dx \)

  \[ = \int_A \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dy \, dx \]

  \[ = \int_{-\infty}^{\infty} \int_A f_{X,Y}(x,y) \, dx \, dy \]

  \[ = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx \right) f_Y(y) \, dy \]

  \[ = \int_{-\infty}^{\infty} P(\mathbf{X} \in A \mid Y = y) f_Y(y) \, dy \]

- \( E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx \)

  \[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X,Y}(x,y) \, dy \, dx \]

  \[ = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} x f_{X,Y}(x,y) \, dy \right) f_Y(y) \, dx \]

  \[ = \int_{-\infty}^{\infty} \left[ E[X \mid Y = y] \right] f_Y(y) \, dy \]
Independence

- If $X, Y$ are independent, then $f_{X,Y}(x,y) = f_X(x)$ for all $x,y$.
- If $X, Y$ are independent, then $E[X|Y=y] = E[X]$ for all $y$ with $P(Y=y) > 0$.

Conditional Expectation as a Random Variable

$E[X|Y]$ is the random variable defined by

$$E[X|Y](w) = E[X|Y=y]$$

when $w$ is such that $Y(w) = y$.

In other words,

$$E[X|Y] = h(Y)$$

where $h: \mathbb{R} \to \mathbb{R}$ is defined by $h(y) = E[X|Y=y]$.

Averaging Identity

$$E[E[X|Y]] = E[X]$$

Properties

1. If $X, Y$ independent, then $E[X|Y] = E[X]$ (the random variable $E[X|Y]$ is constant).
2. $E[g(X)h(Y)|Y=y] = E[g(X)h(Y)|Y=y] = h(y)E[g(X)|Y=y]$ constant
3. $E[g(X)h(Y)|Y] = h(Y)E[g(X)|Y]$
Example

\((X, Y)\) uniformly distributed on the triangle \(A\) with vertices \((0,0), (1,0), (0,1)\)

Find the conditional density of \(X\) given \(Y = y\).

Identity the distribution of \(X\) given \(Y = y\).

The joint pdf is

\[
\begin{align*}
    f_{X,Y}(x,y) &= \begin{cases} 
    \frac{1}{\text{area}(A)} = 2 & \text{if } (x,y) \in A \\
    0 & \text{if } (x,y) \notin A
    \end{cases}
\end{align*}
\]

The conditional pdf of \(X\) given \(Y = y\) is

\[
    f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{when } f_Y(y) > 0
\]

Need \(f_Y(y)\).

\[
    f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx = \begin{cases} 
    0 & \text{if } y \notin [0,1] \\
    \int_0^{1-y} 2 \, dx = 2(1-y) & \text{if } y \in [0,1]
    \end{cases}
\]
If $y \notin (0, 1)$, then $f_{X\mid Y}(x\mid y)$ is undefined.

If $y \in (0, 1)$, then

$$f_{X\mid Y}(x\mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{1-y} & \text{if } x \in (0, 1-y) \\ 0 & \text{if } x \notin (0, 1-y) \end{cases}$$

The conditional distribution of $X$ given $Y=y$ is uniform on $(0, 1-y)$ when $y \in (0, 1)$.

Briefly, $X\mid Y=y \sim \text{Unif}(0, 1-y)$ for $y \in (0, 1)$.

Makes sense geometrically. If we fix $y \in (0, 1)$, if we choose $(x, y)$ uniformly from the triangle $A$ but require a fixed $y$-coordinate $y_0$, then we are choosing uniformly from the line segment $\{(x, y_0) : 0 \leq x \leq 1-y_0\}$. 

\[ 
\]
Example

The joint density function of $(X,Y)$ is

$$f(x,y) = \frac{x+y}{4} \quad \text{if } 0 < x < y < 2$$

($f(x,y) = 0$ otherwise)

Find the conditional pdf of $X$ given $Y = y$.
Find $P(X < \frac{1}{2} \mid Y = 1)$ and $P(Y < \frac{3}{2} \mid Y = 1)$.
Find $E[X^2 \mid Y = y]$ and $E[X^2 \mid Y]$.

Need $f_Y(y)$ first.

For $0 < y < 2$,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) \, dx = \frac{1}{4} \int_{0}^{y} (x+y) \, dx = \frac{1}{4} \left( \frac{1}{2} x^2 + xy \right) \bigg|_{x=0}^{x=y} = \frac{3}{8} y^2$$

For $y \notin (0,2)$, $f_Y(y) = 0$.

For $0 < y < 2$,

$$f_{X \mid Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \left\{ \begin{array}{ll} \frac{1}{4} \frac{x+y}{y^2} & \text{if } 0 < x < y < 2 \\ \frac{2}{3} \frac{x+1}{y^2} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{array} \right.$$
\[ P(X < \frac{1}{2} \mid Y = 1) = \int_0^{1/2} f_{X|Y}(x|1) \, dx = \frac{2}{3} \int_0^{1/2} (x+1) \, dx = \frac{17}{48} \]

\[ P(X < \frac{3}{2} \mid Y = 1) = \int_0^{3/2} f_{X|Y}(x|1) \, dx = \frac{2}{3} \int_0^{1} (x+1) \, dx = 1 \]

\[ E[X^2 \mid Y = y] = \int_0^{y} x^2 f_{X|Y}(x|y) \, dx = \int_0^{y} x^2 \cdot \frac{2}{3} \cdot \frac{x+1}{y^2} \, dx \]

\[ = \frac{2}{3y^2} \left[ \int_0^{y} (x^3 + x^2) \, dx \right] \]

\[ = \frac{2}{3y^2} \left[ \left( \frac{1}{4} x^4 + \frac{1}{3} x^3 \right) \right]_{x=0}^{x=y} \]

\[ = \frac{7}{18} y^2 \]

\[ \therefore E[X^2 \mid Y] = \frac{7}{18} Y^2 \]
Example

Poisson process with rate $\lambda$.

$T_1 =$ time to first arrival $\sim \text{Exp}(\lambda)$

$N([a,b]) =$ # arrivals in time interval $[a,b]$

$\sim \text{Poisson} ((b-a)\lambda)$

Identify the conditional distribution of $T_1$ given $N([0,s]) = 1$.

We compute the conditional cdf

$P(T_1 \leq t \mid N([0,s]) = 1)$

Easy case: $t < 0$

$P(T_1 \leq t \mid N([0,s]) = 1) = 0$

Easy case: $t > s$

$P(T_1 \leq t \mid N([0,s]) = 1) = 1$

Hard case: $0 \leq t \leq s$

$P(T_1 \leq t \mid N([0,s]) = 1) = \frac{P(T_1 \leq t, N([0,s]) = 1)}{P(N([0,s]) = 1)}$

$= \frac{P(1 \text{ arrival in } [0,t], 0 \text{ arrivals } (t,s])}{P(N([0,s]) = 1)}$

$= \frac{P(N([0,t]) = 1, N((t,s]) = 0)}{P(N([0,s]) = 1)}$

$= \frac{P(N([0,t]) = 1)P(N((t,s]) = 0)}{P(N([0,s]) = 1)}$ (independent intervals)

$= \frac{e^{-t\lambda} (t\lambda)^1 \cdot e^{-(s-t)\lambda} ((s-t)\lambda)^0}{e^{-s\lambda} (s\lambda)^1 \cdot e^{-s\lambda} (s\lambda)^1} = \frac{t}{s}$
Summary:
\[
P(T_1 \leq t \mid N([0,s]) = 1) = \begin{cases} 
0 & \text{if } t < 0 \\
\frac{t}{s} & \text{if } 0 \leq t \leq s \\
1 & \text{if}
\end{cases}
\]

This is the cdf of a Unit $[0,s]$ random variable.

\[
\therefore T_1 \text{ given } N([0,s]) = 1 \text{ is uniformly distributed on } [0,s]
\]

In words, given that there is exactly 1 arrival in the time interval $[0,s]$, the time (point) of that arrival is uniformly distributed on $[0,s]$. So the arrival is equally likely to occur at any time point between 0 and s.