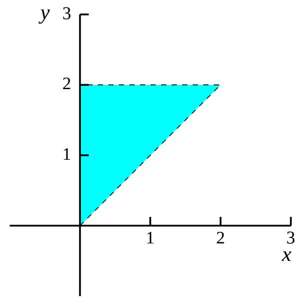
Math 180A, Fall 2005, Prof. Tesler – November 21, 2005
Joint Densities for continuous random variables

1. Example: continuous pdf defined in a triangular region

Let $f_{X,Y}(x,y)$ denote the joint pdf of continuous random variables (vs. $p_{X,Y}(x,y)$ for discrete). We will work with this example:

$$f_{X,Y}(x,y) = \begin{cases} xy/2 & \text{if } 0 < x < y < 2; \\ 0 & \text{otherwise.} \end{cases}$$



The marginal pdfs are defined as $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$.
In our example, if $x \le 0$ or $x \ge 2$ then $f_{X,Y}(x,y) = 0$ for all y so $f_X(x) = \int_{-\infty}^{\infty} 0\,dy = 0$.
If $0 < x < 2$ then $f_{X,Y}(x,y) \ne 0$ only for y between x and 2, giving

$$f_X(x) = \int_{-\infty}^{x} 0\,dy + \int_{x}^{2} \frac{xy}{2}\,dy + \int_{2}^{\infty} 0\,dy = 0 + \left(\left.\frac{xy^2}{4}\right|_{y=x}^{2} \right) + 0 = \frac{x}{4}(2^2 - x^2) = x - \frac{x^3}{4}$$

Likewise, if $y \le 0$ or $y \ge 2$ then $f_Y(y) = 0$, while if $0 < y < 2$ then

$$f_Y(y) = \int_{-\infty}^{0} 0\,dy + \int_{0}^{y} \frac{xy}{2}\,dy + \int_{y}^{\infty} 0\,dy = 0 + \left(\left.\frac{x^2 y}{4}\right|_{x=0}^{y} \right) + 0 = \frac{(y^2 - 0^2)y}{4} = \frac{y^3}{4}$$

So $$f_X(x) = \begin{cases} x - x^3/4 & \text{if } 0 < x < 2; \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} y^3/4 & \text{if } 0 < y < 2; \\ 0 & \text{otherwise.} \end{cases}$$

2. Independence

Continuous random variables X, Y are independent when $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ for all values of x and y. Here, when $0 < x < y < 2$, we have $f_X(x) f_Y(y) = (x - \frac{x^3}{4})(\frac{y^3}{4}) \ne \frac{xy}{2} = f_{X,Y}(x,y)$, so they are not independent.

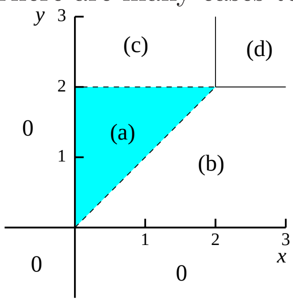
Incidentally, if $0 < y < x < 2$ then $f_X(x) f_Y(y) = (x - \frac{x^3}{4})(\frac{y^3}{4})$ and $f_{X,Y}(x,y) = 0$, which also does not obey the factor rule so that's another way to demonstrate it's not independent.

3. Cumulative distribution function

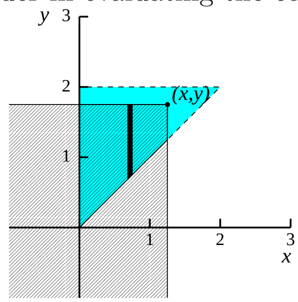
$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)\,du\,dv$$

Integrate in whichever order is easiest. Since x and y are already being used as variables, the integration variables cannot also be called x and y, so instead I used u and v.

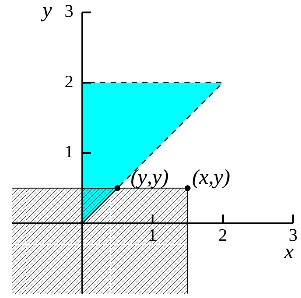
There are many cases to consider in evaluating the cdf:



All regions



Domain when (x,y) is in region (a)



Domain when (x,y) is in region (b)

Region (a) $0 < x < y < 2$: $F_{X,Y}(x,y) = \int_0^x \int_u^y \frac{uv}{2}\,dv\,du$ (the rest of $u \le x$, $v \le y$ contributes 0)

The inside integral is $\int_u^y \frac{uv}{2}\,dv = \left.\frac{uv^2}{4}\right|_{v=u}^{y} = \frac{u(y^2-u^2)}{4} = \frac{uy^2-u^3}{4}$.

The outside integral is $F_{X,Y}(x,y) = \int_0^x \frac{uy^2-u^3}{4}\,du = \left.\left(\frac{u^2y^2}{8} - \frac{u^4}{16}\right)\right|_{u=0}^{x} = \frac{(x^2-0^2)y^2}{8} - \frac{x^4-0^4}{16} = \frac{x^2y^2}{8} - \frac{x^4}{16}$.

Region (b) $0 < y < 2$ and $x > y$:

The part of the region $u \le x$, $v \le y$ with nonzero pdf is the same as for $u \le y$, $v \le y$. So $F_{X,Y}(x,y) = F_{X,Y}(y,y) = \frac{y^4}{8} - \frac{y^4}{16} = \frac{y^4}{16}$.

Region (c) $y > 2$ and $0 < x < 2$: $F_{X,Y}(x,y) = F_{X,Y}(x,2) = \frac{x^2 \cdot 2^2}{8} - \frac{x^4}{16} = \frac{x^2}{2} - \frac{x^4}{16}$

Region (d) $x > 2$ and $y > 2$: $F_{X,Y}(x,y) = F_{X,Y}(2,2) = \frac{2^2 \cdot 2^2}{8} - \frac{2^4}{16} = 2 - 1 = 1$

Other $x < 0$ or $y < 0$: $f_{X,Y}(u,v) = 0$ for all $u \le x$ and $v \le y$, so $F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} 0\,dv\,du = 0$.

Summary:
$$F_{X,Y}(x,y) = \begin{cases} \frac{x^2y^2}{8} - \frac{x^4}{16} & \text{if } 0 < x < y < 2 & \text{(region (a));} \\ \frac{y^4}{16} & \text{if } 0 < y < 2 \text{ and } x > y & \text{(region (b));} \\ \frac{x^2}{2} - \frac{x^4}{16} & \text{if } 0 < x < 2 < y & \text{(region (c));} \\ 1 & \text{if } x > 2 \text{ and } y > 2 & \text{(region (d));} \\ 0 & \text{otherwise.} \end{cases}$$

$$F_X(x) = F_X(x,\infty) = \begin{cases} 0 & \text{if } x \le 0; \\ x^2/2 - x^4/16 & \text{if } 0 \le x \le 2 \text{ (c);} \\ 1 & \text{if } x \ge 2 \text{ (d).} \end{cases} \qquad f_X(x) = F_X{}'(x) = \begin{cases} x - x^3/4 & \text{if } 0 < x < 2; \\ 0 & \text{if } x < 0 \text{ or } x > 2. \end{cases}$$

$$F_Y(y) = F_Y(\infty,y) = \begin{cases} 0 & \text{if } y < 0; \\ y^4/16 & \text{if } 0 < y < 2 \text{ (b);} \\ 1 & \text{if } y > 2. \end{cases} \qquad f_Y(y) = F_Y{}'(y) = \begin{cases} y^3/4 & \text{if } 0 < y < 2; \\ 0 & \text{if } y < 0 \text{ or } y > 2. \end{cases}$$

4. Computing pdf from cdf

$$f_{X,Y}(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y} F_{X,Y}(x,y) = \frac{\partial^2}{\partial x\,\partial y} F_{X,Y}(x,y)$$

In our example, if $0 < x < y < 2$ then $\frac{\partial}{\partial x}\frac{\partial}{\partial y}\left(\frac{x^2y^2}{8} - \frac{x^4}{16}\right) = \frac{\partial}{\partial x}\left(\frac{x^2y}{4}\right) = \frac{xy}{2}$

Indeed, this does equal $f_{X,Y}(x,y)$. In all other cases, $\frac{\partial}{\partial x}\frac{\partial}{\partial y} F_{X,Y}(x,y) = 0$ since it's either a function of x alone, y alone, or a constant, and indeed $f_{X,Y}(x,y) = 0$ in all other cases.

5. Expected values for joint pdfs of continuous random variables

Let $g(X,Y)$ be a function of continuous random variables X, Y. The expected value is $E(g(X,Y)) = \iint_D g(x,y)\, f_{X,Y}(x,y)\, dA$ where $D = \mathbb{R}^2$ or just the subset of it where the pdf is nonzero.

In the current example, D is the triangular region $0 < x < y < 2$. The mean value of X is

$$\mu_X = E(X) = \iint_D x \cdot (xy/2)\, dA = \int_0^2 \int_x^2 \frac{x^2 y}{2}\, dy\, dx$$

The inside integral is $(x^2y^2/4)|_{y=x}^2 = x^2(2^2 - x^2)/4 = x^2 - x^4/4$.

The outside integral is $\int_0^2 (x^2 - x^4/4)\, dx = (x^3/3 - x^5/20)|_{x=0}^2 = (2^3 - 0^3)/3 - (2^5 - 0^5)/20 = 16/15$ so $\mu_X = E(X) = 16/15$.

Likewise, $E(X^2) = \iint_D x^2 \cdot (xy/2)\, dA = \iint_D (x^3y/2)\, dA = \cdots = 4/3$, so the variance of X is ${\sigma_X}^2 = \mathrm{Var}(X) = E(X^2) - (E(X))^2 = (4/3) - (16/15)^2 = 44/225$. The standard deviation of X is $\sigma_X = \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{44/225} = 2\sqrt{11}/15$.

The mean of Y is $\mu_Y = E(Y) = \iint_D y \cdot (xy/2)\, dA = \iint_D (xy^2/2)\, dA = \cdots = 8/5$.

$E(Y^2) = \iint_D y^2 \cdot (xy/2)\, dA = \iint_D (xy^3/2)\, dA = \cdots = 8/3$

${\sigma_Y}^2 = \mathrm{Var}(Y) = E(Y^2) - (E(Y))^2 = (8/3) - (8/5)^2 = 8/75$

$\sigma_Y = \mathrm{SD}(Y) = \sqrt{8/75}$

6. Conditional probability for joint distributions

(i) Compute $P(X < 1 | Y < 3/2)$:

$$P(X < 1 | Y < 3/2) = \frac{P(X < 1, Y < 3/2)}{P(Y < 3/2)} = \frac{F_{X,Y}(1^-, \frac{3}{2}^-)}{F_{X,Y}(\infty, \frac{3}{2}^-)} = \frac{7/32}{81/256} = \boxed{\mathbf{\frac{56}{81}}}$$

Note: $P(X < 1, Y < 3/2) = F_{X,Y}(1^-, \frac{3}{2}^-)$ is a general formula that works for both discrete and continuous pdfs. In the discrete case, there could be a jump at some values of x and y. In this example, it's a continuous pdf, and there are no jumps, so $P(X < 1, Y < 3/2) = F_{X,Y}(1^-, \frac{3}{2}^-) = F_{X,Y}(1, \frac{3}{2})$.

(ii) Compute $P(X < 1 | Y = 3/2)$:

Wrong way: $P(X < 1 | Y = 3/2) = \frac{P(X<1, Y=3/2)}{P(Y=3/2)} = \frac{0}{0}$

Right way: For discrete pdfs, the *conditional probability density function* is $P(X = x | Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$, where y is fixed and it's viewed as a function of x. The conditional probability density function for a continuous pdf is defined analogously:

$$f_X(x | Y = y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad f_Y(y | X = x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

In our triangle example, $f_X(x|Y=y) = \begin{cases} \frac{xy/2}{y^3/4} = \frac{2x}{y^2} & \text{if } 0 < x < y < 2; \\ 0 & \text{otherwise.} \end{cases}$

So $P(X < 1 | Y = 3/2) = \int_{-\infty}^1 f_X(x | Y = \frac{3}{2})\, dx = \int_0^1 \frac{2x}{(3/2)^2}\, dx = \left.\frac{4x^2}{9}\right|_{x=0}^1 = \frac{4}{9}(1^2 - 0^2) = \boxed{\mathbf{\frac{4}{9}}}$

(Since $f_X(x|Y = \frac{3}{2})$ is nonzero only for $0 < x < 3/2$, the integration domain was shortened.)

7. Conditional expected values for discrete random variables

Let A be an event. Conditional expectation is defined as

$E(X|A) = \sum_x xP(X = x|A)$

$E(g(X)|A) = \sum_x g(x)P(X = x|A)$

$E(g(X,Y)|A) = \sum_x \sum_y g(x,y)P(X = x, Y = y|A)$

The expected value of the roll of a fair die, given that it shows an even value:

$P(X = x|X \text{ is even}) = 1/3$ for $x = 2, 4, 6$ and is 0 otherwise.

$E(X = x|X \text{ is even}) = (2)(1/3) + (4)(1/3) + (6)(1/3) = 12/3 = 4.$

8. Conditional expected values for continuous random variables

(i) Compute $E(X|Y < 3/2)$:

The event $Y < 3/2$ has nonzero probability, so we compute

$$f_{X,Y}(x, y|Y < \tfrac{3}{2}) = \begin{cases} \frac{f_{X,Y}(x,y)}{P(Y<\frac{3}{2})} & \text{if } y < \frac{3}{2}; \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{xy/2}{81/256} = \frac{128}{81}xy & \text{if } 0 < x < y < \frac{3}{2}; \\ 0 & \text{otherwise} \end{cases}$$

Then $E(X|Y < 3/2) = \int_0^{3/2} \int_0^y x \cdot (128/81)xy\,dx\,dy$.

Inside: $\int_0^y (128/81)x^2y\,dx = (128/81)(x^3/3)y\big|_{x=0}^{y} = (128/243)(y^3 - 0^3)y = (128/243)y^4$.

Outside: $\int_0^{3/2} (128/243)y^4\,dy = (128/1215)y^5\big|_{y=0}^{3/2} = (128/1215)((3/2)^5 - 0^5) = \boxed{\mathbf{4/5}}$

(ii) Compute $E(X|Y = 3/2)$:

Define $E(g(X)|Y = y) = \int_{-\infty}^{\infty} g(x)f_X(x|Y = y)\,dx$.

In our earlier example,

$$E(X|Y = 3/2) = \int_0^{3/2} x \cdot \frac{2x}{(3/2)^2}\,dx = \frac{8x^3}{27}\bigg|_{x=0}^{3/2} = \frac{8((3/2)^3 - 0^2)}{27} = 1.$$

More generally, for $0 < y < 2$ we had $f_X(x|Y = y) = 2x/y^2$ when $0 < x < y < 2$, $f_X(x|Y = y) = 0$ otherwise. This gives

$$E(X|Y = y) = \int_0^y x \cdot \frac{2x}{y^2}\,dx = \frac{2x^3}{3y^2}\bigg|_{x=0}^{y} = \frac{2(y^3 - 0^3)}{3y^2} = \frac{2y}{3}.$$

This may also be written $E(X|Y) = 2Y/3$. Since it's still a function of Y, we may take the expected value of this as a function of the random variable Y whose pdf is $f_Y(y) = y^3/4$ if $0 < y < 2$, $f_Y(y) = 0$ otherwise:

$$E(E(X|Y)) = \int_0^2 (2y/3)(y^3/4)\,dy = \int_0^2 (y^4/6)\,dy = (y^5/30)\big|_{y=0}^{2} = (2^5 - 0^5)/30 = 32/30 = 16/15 = E(X)$$

In general,

$$\begin{aligned} E(E(X|Y)) &= \int_{-\infty}^{\infty} E(X|Y = y)f_Y(y)\,dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xf_X(x|Y = y)\,dx \right) f_Y(y)\,dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y)\,dx\,dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X,Y}(x,y)\,dx\,dy = E(X) \end{aligned}$$