Multivariate distributions

1 Properties of a joint PDF

Continuous random variables X and Y have the following joint probability density function (PDF):

$$f_{XY}(x,y) = \begin{cases} kx^2y^3 & \text{where } 0 < x, y < 6 \\ 0 & \text{otherwise} \end{cases}$$

Note: $0 < x, y < 6$ means that both x and y are between 0 and 6; it does not mean that x is greater than 0 and y is less than 6. This notation is not uncommon, so keep it in mind.

a. Find k.

Solution:

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy &= 1 \\
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} kx^2y^3\,dx\,dy &= 1 \\
\int_{0}^{6}\int_{0}^{6} kx^2y^3\,dx\,dy &= 1 \\
k\int_{0}^{6}\int_{0}^{6} x^2y^3\,dx\,dy &= 1 \\
k\int_{0}^{6} y^3 \cdot \frac{x^3}{3}\Big|_0^6\,dy &= 1 \\
k\int_{0}^{6} y^3 \cdot \left(\frac{6^3}{3} - 0\right)\,dy &= 1 \\
72k\int_{0}^{6} y^3\,dy &= 1 \\
72k \cdot \frac{y^4}{4}\Big|_0^6 &= 1 \\
72k\left(\frac{6^4}{4} - 0\right) &= 1 \\
72k \cdot 324 &= 1 \\
23328k &= 1 \\
k &= \frac{1}{23328}
\end{aligned}$$

b. Find the marginal PDF of X, $f_X(x)$.

Solution:

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \\
&= \int_{-\infty}^{\infty} \frac{x^2y^3}{23328}\,dy \\
&= \int_{0}^{6} \frac{x^2y^3}{23328}\,dy \\
&= \frac{x^2}{23328}\int_{0}^{6} y^3\,dy \\
&= \frac{x^2}{23328}\cdot \frac{y^4}{4}\bigg|_0^6 \\
&= \frac{x^2}{23328}\cdot \frac{6^4}{4} \\
&= \frac{x^2}{72}
\end{aligned}
$$

c. Find the marginal PDF of Y, $f_Y(y)$.

Solution:

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \\
&= \frac{1}{23328}\int_{0}^{6} x^2y^3\,dx \\
&= \frac{1}{23328}\cdot \frac{x^3}{3}\bigg|_0^6 y^3 \\
&= \frac{1}{23328}\left(\frac{216}{3}-0\right)y^3 \\
&= \frac{1}{23328}\,(72)\,y^3 \\
&= \frac{y^3}{324}
\end{aligned}
$$

d. Find $\mathrm{E}[X]$.

Solution:

$$
\begin{aligned}
\mathrm{E}[X] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,dx\,dy \\
&= \frac{1}{23328}\int_{0}^{6}\int_{0}^{6} x\cdot x^2y^3\,dx\,dy \\
&= \frac{1}{23328}\times \int_{0}^{6} x^3\,dx \times \int_{0}^{6} y^3\,dy \\
&= \frac{1}{23328}\times \frac{x^4}{4}\bigg|_0^6 \times \frac{y^4}{4}\bigg|_0^6 \\
&= \frac{1}{23328}\left(\frac{1296}{4}-0\right)\left(\frac{1296}{4}-0\right) \\
&= \frac{1}{23328}\,(324)\,(324) \\
&= 4.5
\end{aligned}
$$

e. Find $\mathrm{E}[Y]$.

Solution:

$$\begin{aligned}
\mathrm{E}[Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dx\,dy \\
&= \frac{1}{23328}\int_0^6\int_0^6 y\cdot x^2y^3\,dx\,dy \\
&= \frac{1}{23328}\times\int_0^6 x^2\,dx\times\int_0^6 y^4\,dy \\
&= \frac{1}{23328}\times\frac{x^3}{3}\bigg|_0^6\times\frac{y^5}{5}\bigg|_0^6 \\
&= \frac{1}{23328}\left(\frac{216}{3}-0\right)\left(\frac{6^5}{5}-0\right) \\
&= \frac{1}{23328}\,(72)\left(\frac{7776}{5}\right) \\
&= 4.8
\end{aligned}$$

f. Find $\mathrm{Var}(X)$.

Solution: To find $\mathrm{Var}(X)$, we first need to find $\mathrm{E}[X^2]$.

$$\begin{aligned}
\mathrm{E}[X^2] &= \frac{1}{23328}\int_0^6\int_0^6 x^2\cdot x^2y^3\,dx\,dy \\
&= \frac{1}{23328}\times\int_0^6\int_0^6 x^4y^3\,dx\,dy \\
&= \frac{1}{23328}\times\int_0^6 x^4\,dx\times\int_0^6 y^3\,dy \\
&= \frac{1}{23328}\times\frac{x^5}{5}\bigg|_0^6\times\frac{y^4}{4}\bigg|_0^6 \\
&= \frac{1}{23328}\left(\frac{7776}{5}-0\right)\left(\frac{1296}{4}-0\right) \\
&= \frac{1}{23328}\left(\frac{7776}{5}\right)(324) \\
&= 21.6
\end{aligned}$$

With $\mathrm{E}[X^2]$ determined, we can now calculate $\mathrm{Var}(X)$.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}[X^2]-\mathrm{E}[X]^2 \\
&= 21.6-(4.5)^2 \\
&= 21.6-20.25 \\
&= 1.35
\end{aligned}$$

g. Find $\mathrm{Var}(Y)$.

Solution: Again, to find $\mathrm{Var}(Y)$, we first need $\mathrm{E}[Y^2]$.

$$
\begin{aligned}
\mathrm{E}[Y^2] &= \frac{1}{23328} \int_0^6 \int_0^6 y^2 \cdot x^2 y^3 \, dx \, dy \\
&= \frac{1}{23328} \times \int_0^6 \int_0^6 x^2 y^5 \, dx \, dy \\
&= \frac{1}{23328} \times \int_0^6 x^2 \, dx \times \int_0^6 y^5 \, dy \\
&= \frac{1}{23328} \times \left. \frac{x^3}{3} \right|_0^6 \times \left. \frac{y^6}{6} \right|_0^6 \\
&= \frac{1}{23328} \left(\frac{216}{3} - 0 \right) (7776 - 0) \\
&= \frac{1}{23328} (72)(7776) \\
&= 24
\end{aligned}
$$

With $\mathrm{E}[Y^2]$ in hand, we can now calculate $\mathrm{Var}(Y)$.

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}[Y^2] - \mathrm{E}[Y]^2 \\
&= 24 - (4.8)^2 \\
&= 24 - 23.04 \\
&= 0.96
\end{aligned}
$$

h. Find $\mathrm{Cov}(X, Y)$.

Solution: To find $\mathrm{Cov}(X, Y)$, we first need $\mathrm{E}[XY]$.

$$
\begin{aligned}
\mathrm{E}[XY] &= \frac{1}{23328} \int_0^6 \int_0^6 xy \cdot x^2 y^3 \, dx \, dy \\
&= \frac{1}{23328} \times \int_0^6 \int_0^6 x^3 y^4 \, dx \, dy \\
&= \frac{1}{23328} \times \int_0^6 x^3 \, dx \times \int_0^6 y^4 \, dy \\
&= \frac{1}{23328} \times \left. \frac{x^4}{4} \right|_0^6 \times \left. \frac{y^5}{5} \right|_0^6 \\
&= \frac{1}{23328} \left(\frac{1296}{4} - 0 \right) \left(\frac{7776}{5} - 0 \right) \\
&= \frac{1}{23328} (324) \left(\frac{7776}{5} \right) \\
&= 21.6
\end{aligned}
$$

Now, we calculate $\mathrm{Cov}(X, Y)$.

$$
\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= 21.6 - (4.5)(4.8) \\
&= 21.6 - 21.6 \\
&= 0
\end{aligned}
$$

i. Are X and Y independent? Why?

Solution: X and Y are independent because $f_{XY}(x,y) = f_X(x)f_Y(y)$ (definition of independence), or in other words, the product of the marginal densities of X and Y is equal to the joint density of X and Y:

$$f_X(x)f_Y(y) = \frac{x^2}{72}\left(\frac{y^3}{324}\right) = \frac{x^2y^3}{23328} = f_{XY}(x,y)$$

We **cannot** say that X and Y are independent because the covariance is zero. While it is true that independent variables have a covariance of zero, it is not necessarily true that variables with a covariance of zero are independent.

j. What is the PDF of X conditional on Y, $f_{X|Y}(x|y)$?

Solution: We've previously shown that X and Y are independent. This implies that $f(x) = f(x|y)$ so the answer is the same as the marginal distribution of x from part (b),

$$\begin{aligned} f(x|y) &= f(x) \\ &= \frac{x^2}{72} \end{aligned}$$

k. What is the PDF of Y conditional on X, $f_{Y|X}(y|x)$?

Solution: Again since we've already shown that X and Y are independent we can just refer back to the answer to (c).

$$\begin{aligned} f(y|x) &= f(y) \\ &= \frac{y^3}{324} \end{aligned}$$

2 Properties of joint random variables

- $\mathrm{E}[D] = 10$
- $\mathrm{E}[F] = 4$
- $\mathrm{E}[DF] = 8$
- $\mathrm{Var}(D) = 60$
- $\mathrm{Var}(F) = 60$

a. What is $\mathrm{Cov}(D,F)$?

$$\begin{aligned} \mathrm{Cov}(D,F) &= \mathrm{E}[DF] - \mathrm{E}[D]\mathrm{E}[F] \\ &= 8 - (4 \times 10) \\ &= -32 \end{aligned}$$

b. What is the correlation between D and F?

$$\begin{aligned} \mathrm{Cor}(D,F) &= \frac{\mathrm{Cov}(D,F)}{\sqrt{\mathrm{Var}(F)\mathrm{Var}(D)}} \\ &= \frac{-32}{\sqrt{60 \times 60}} \\ &= -0.5333333 \end{aligned}$$

c. Suppose you multiplied F by 2 to generate a new variable, H. What is $\text{Cov}(D, H)$?

Solution: Multiplying F by 2 increases the magnitude of the covariance between D and H.

$$\begin{aligned}
\text{Cov}(D,H) &= \text{E}[DH] - \text{E}[D]\text{E}[H] \\
\text{E}[DH] &= \text{E}[D \times 2F] = \int 2DFf(D,F)d(D,F) = 2\int DFf(D,F)d(D,F) = 2\text{E}[DF] = 16 \\
\text{E}[H] &= \text{E}[2F] = \int 2Ff(F)dF = 2\int Ff(F)dF = 2\text{E}[F] = 8 \\
\text{Cov}(D,H) &= 16 - (8 \times 10) = -64
\end{aligned}$$

d. What is $\text{Cor}(D, H)$? How does this compare to your answer to Part (b) of this question?

$$\begin{aligned}
\text{Var}(H) &= \text{Var}(2F) = 2^2\text{Var}(F) = 4 \times 60 = 240 \\
\text{Cor}(D,H) &= \frac{\text{Cov}(D,H)}{\sqrt{\text{Var}(F)\text{Var}(H)}} \\
&= \frac{-64}{\sqrt{60 \times 240}} \\
&= -0.5333333
\end{aligned}$$

This is the same as $\text{Cor}(D, F)$. In other words, multiplying one of the variables by a constant leaves the correlation between the two variables unchanged. This occurs despite the covariance changing.

e. Suppose instead that $\text{Var}(D) = 30$. How would this change $\text{Cor}(D, F)$?

Solution: The magnitude of the correlation between the variables increases as Var(D) decreases:

$$\begin{aligned}
\text{Cor}(D,F) &= \frac{\text{Cov}(D,F)}{\sqrt{\text{Var}(F)\text{Var}(D)}} \\
&= \frac{-32}{\sqrt{60 * 30}} \\
&= -0.7542472
\end{aligned}$$

3 Calculating conditional PDF

Let $f(x, y) = 15x^2y$ for $0 \le x \le y \le 1$. Find $f(x|y)$.

Solution:

$$
\begin{aligned}
f(y) &= \int_0^y f(x,y)\,dx \\
&= \int_0^y 15x^2 y\,dx \\
&= 15y \int_0^y x^2\,dx \\
&= 15y \frac{x^3}{3}\Big|_0^y \\
&= \frac{15y^4}{3} \\
f(x|y) &= \frac{f(x,y)}{f(y)} \\
&= \frac{15x^2 y}{15y^4/3} \\
&= \frac{3x^2}{y^3}
\end{aligned}
$$

4 Deriving a joint PDF

We start with a stick of length l. We break it at a point which is chosen according to a uniform distribution and keep the piece, of length Y, that contains the left end of the stick. We then repeat the same process on the piece that we were left with, and let X be the length of the remaining piece after breaking for the second time.

a. Find the joint PDF of Y and X

Solution: We have

$$f_Y(y) = \frac{1}{l}, \forall\, 0 \le y \le l$$

Furthermore, given the value y of Y, the random variable X is uniform in the interval $[0, y]$. Therefore

$$f_{X|Y}(x|y) = \frac{1}{y}, \forall\, 0 \le x \le y$$

We conclude that

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y) = \begin{cases} \frac{1}{l} \times \frac{1}{y} & , 0 \le x \le y \le l, \\ 0 & , \text{otherwise} \end{cases}$$

b. Find the marginal PDF of X

Solution: We have

$$f_X(x) = \int f_{X,Y}(x,y)\,dy = \int_x^l \frac{1}{ly}\,dy = \frac{1}{l} \log\left(\frac{l}{x}\right), \forall\, 0 \le x \le l$$

c. Use the PDF of X to evaluate $\mathrm{E}[X]$

Solution: We have

$$\mathrm{E}[X] = \int_0^l x f_X(x)\, dx = \int_0^l \frac{x}{l} \log\left(\frac{l}{x}\right)\, dx = \frac{l}{4}$$

d. Evaluate $\mathrm{E}[X]$, by exploiting the relation $X = Y \times \frac{X}{Y}$

Solution: The fraction $\frac{Y}{l}$ of the stick that is left after the first break, and the further fraction $\frac{X}{Y}$ of the stick that is left after the second break are independent. Furthermore, the random variables Y and $\frac{X}{Y}$ are uniformly distributed over the sets $[0, l]$ and $[0, 1]$, respectively, so that $\mathrm{E}[Y] = \frac{l}{2}$ and $\mathrm{E}\left[\frac{X}{Y}\right] = \frac{1}{2}$. Thus,

$$\mathrm{E}[X] = \mathrm{E}[Y]\mathrm{E}\left[\frac{X}{Y}\right] = \frac{l}{2} \times \frac{1}{2} = \frac{l}{4}$$

5 Continuous Bayes' theorem

Previously, we used Bayes' theorem to link the conditional probability of discrete events A given B to the probability of B given A. There is an analogous Bayes' theorem that relates the conditional densities of random variables X and θ:

$$f(\theta \mid X) = \frac{f(X \mid \theta) f(\theta)}{\int f(X \mid \theta) f(\theta) d\theta}$$

Prove the continuous Bayes' theorem.

Solution:

Recall the definition of the conditional distribution of two random variables:

$$f_{\theta|X}(\theta \mid X) = \frac{f(\theta, X)}{f_X(X)}$$

Remember via the "chain rule" of probability that $f(\theta, X) = f(X \mid \theta) f_\theta(\theta)$, and via our rule for marginalization, $f_X(X) = \int f_{X|\theta}(X \mid \theta) f_\theta(\theta) d\theta$. Substitute these equalities in and we have proven the statement:

$$\begin{aligned} f_{\theta|X}(\theta \mid X) &= \frac{f(\theta, X)}{f_X(X)} \\ &= \frac{f(X \mid \theta) f_\theta(\theta)}{\int f_{X|\theta}(X \mid \theta) f(\theta) d\theta} \end{aligned}$$