The random variables X, Y and Z are independent and are distributed as

$$X \sim N(0,1), \quad Y \sim N(1,4), \quad Z \sim N(-1,2).$$

Let $U = 2X + 3Z$ and $V = X + Y - 2Z$. What is the joint distribution of (U,V)?

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix} \right)$$

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \begin{bmatrix} U \\ V \end{bmatrix} \sim N\left(\begin{bmatrix} -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 22 & -10 \\ -10 & 13 \end{bmatrix} \right)$$

$$a \quad + \quad B \begin{bmatrix} \end{bmatrix}$$

mean vector of (U,V) is $\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$

covariance matrix of (U,V) $\begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 3 & -2 \end{bmatrix}$

$$= \begin{bmatrix} 2 & 0 & 6 \\ 1 & 4 & -4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 22 & -10 \\ -10 & 13 \end{bmatrix},$$

Compute the correlation between U and V.

$$\mathrm{corr}(U,V) = \frac{\mathrm{cov}(U,V)}{\sqrt{\mathrm{Var}(U)\,\mathrm{Var}(V)}}$$

$\mathrm{cov}(U,V) = -10$

$\mathrm{Var}(V) = 13$

$\mathrm{Var}(U) = 22$

$$= \frac{-10}{\sqrt{22 \times 13}} \simeq -0.59\ldots$$

$X \sim$ Normal(μ, Σ)

$\Sigma_{ij} = \Sigma_{ji}$ since $\mathrm{cov}(X_i, X_j) = \mathrm{cov}(X_j, X_i)$

where

$$\mathbb{E}(X_i) = \mu_i, \qquad \text{and} \qquad \mathrm{Cov}(X_i, X_j) = \Sigma_{ij}$$

In particular, if Σ is diagonal, then the $X_1, \ldots, X_n$ are independent random variables with $X_i \sim \mathsf{Normal}(\mu_i, \Sigma_{ii})$.

For us, the most important property of the multivariate Normal distribution is its behaviour under linear transformations.

very important

Suppose $\mathbf{X} := (X_1, \ldots, X_n)'$ has a multivariate Normal distribution. Let $\boldsymbol{a} \in \mathbb{R}^m$ and B is an $(m \times n)$ matrix (with $m \leqslant n$). If $\boldsymbol{X} \sim \mathsf{Normal}(\boldsymbol{\mu}, \Sigma)$, then the random vector $Y := \boldsymbol{a} + B\boldsymbol{X}$ has a $\mathsf{Normal}(\boldsymbol{a} + B\boldsymbol{\mu}, B\Sigma B^T)$.

Example: Suppose that $X_1 \sim \mathsf{Normal}(-1, 2)$ and $X_2 \sim \mathsf{Normal}(1, 3)$ are independent. What is the distribution of $Y = 3 + 2X_1 - X_2$?

Observe

$$Y = \underset{a}{3} + \underset{B}{\begin{bmatrix} 2 & -1 \end{bmatrix}} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

and

$$\mathbf{X} \sim \text{Normal}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \right)$$

so

$$Y \sim \text{Normal}\left(\underset{a}{3} + \underset{B}{\begin{bmatrix} 2 & -1 \end{bmatrix}} \underset{\mu}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}, \underset{B}{\begin{bmatrix} 2 & -1 \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}} \underset{B^T}{\begin{bmatrix} 2 \\ -1 \end{bmatrix}} \right)$$

$$\begin{bmatrix} 4 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$Y \sim$ Normal$(0, 11)$.

Conditional probability density functions and conditional expectation

Recall the definition of conditional probability mass function for discrete random variables;

$$f_{X|Y}(x|y) = \mathbb{P}(X = x \,|\, Y = y) = \frac{\mathbb{P}(X = x,\, Y = y)}{\mathbb{P}(Y = y)} = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

provided $f_Y(y) = \mathbb{P}(Y = y) > 0$.

For continuous random variables (X, Y) we can similarly define the **conditional probability density function** of X given $\{Y = y\}$, denoted by $f_{X|Y}(x|y)$, ;

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

Rearrange

$$f_{X,Y}(x,y) = f_{X|Y}(x,y) \cdot f_Y(y)$$

when $f_Y(y) > 0$.

Note that if X and Y are independent, then

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

when $f_Y(y) > 0$.

Exercise. Write $f_{X|Y}$ in terms of f_X, f_Y and $f_{Y|X}$.

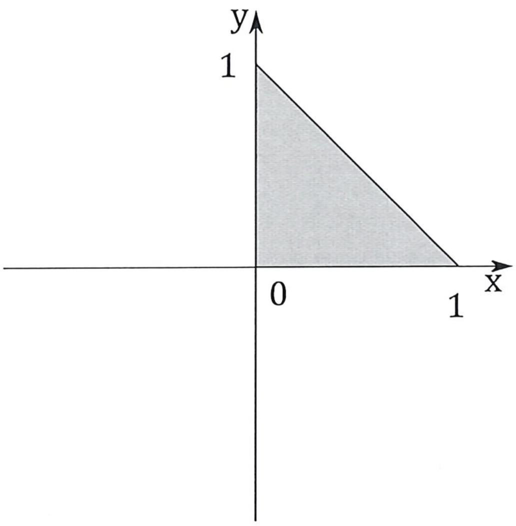
$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)\, f_X(x)}{f_Y(y)}$$

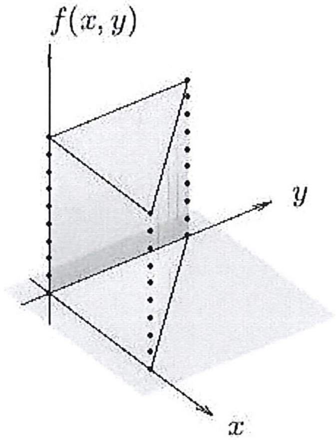
when $f_Y(y) > 0$.

Bayes Rule

Example. We draw a random vector (X, Y) uniformly from the triangle $(0,0)$–$(0,1)$–$(1,0)$–$(0,0)$ (see figure). This pdf is only nonzero when both $0 \leqslant x \leqslant 1$ and $0 \leqslant y \leqslant 1-x$. You can also write those conditions as $0 \leqslant y \leqslant 1$ and $0 \leqslant x \leqslant 1-y$.

What is the joint pdf of X and Y? (Clearly specify where it is zero.)





The triangle has area 1/2. As the joint pdf (of a uniformly-chosen point) must be constant over the support and $\mathbb{P}(\Omega) = 1$, the joint pdf of X and Y is

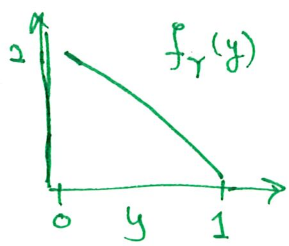
$$f_{X,Y}(x, y) = \begin{cases} 2, & 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1-x. \ // \ 0 \leqslant y \leqslant 1,\ 0 \leqslant x \leqslant 1-y \\ 0, & \text{else} \end{cases}$$

What is the marginal pdf of Y and the conditional pdf of X given $\{Y = y\}$ for this example?

The marginal pdf of Y is

$y \in (0,1)$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx$$
$$= \int_0^{1-y} 2 \, dx$$
$$= 2(1-y).$$



The conditional pdf of X given $\{Y = y\}$ is

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

$$= \begin{cases} \frac{2}{2(1-y)} = \frac{1}{1-y}, & 0 \leq x \leq 1-y \\ 0, & \text{else} \end{cases}$$

Example. Suppose that (X, Y) has a standard bivariate normal distribution, that is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\varrho^2}} \exp\left(-\frac{1}{2(1-\varrho^2)}\left(x^2 - 2\varrho xy + y^2\right)\right), \quad (x,y) \in \mathbb{R}^2,$$

where $\varrho \in (-1, 1)$. What is the conditional pdf of Y given $\{X = x\}$?

Using the same trick as before, we can write the joint pdf as

$f_X(x) \rightarrow$

$$f_{X,Y}(x,y) = \boxed{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)} \frac{1}{\sqrt{2\pi(1-\varrho^2)}} \exp\left(-\frac{1}{2(1-\varrho^2)}(y - \varrho x)^2\right).$$

The marginal distribution of X is Standard Normal $N(0,1)$ So

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

$y \in \mathbb{R}$.

$$= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2(1-\rho^2)}(y - \rho x)^2\right)$$

That is, conditional on $\{X = x\}$, Y has a $N(\rho x, 1 - \rho^2)$ distribution.

As we did in the case of discrete random variables, we can take expectations conditional on events such as $\{X = x\}$ or conditional on random variables. We now adapt our definitions to continuous random variables.

Definition. (Conditional expectation given the event $\{X = x\}$) The conditional expectation of Y given $\{X = x\}$ is

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y \, f_{Y \mid X}(y|x) \, dy.$$

We can relate the conditional expectation $\mathbb{E}[Y \mid X = x]$ to the unconditional expectation $\mathbb{E}[Y]$ as follows: For any random variables X and Y defined in the same random experiement

$$\int_{-\infty}^{\infty} \mathbb{E}[Y \mid X = x]\, f_X(x)\, \mathrm{d}x = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dy \right\} f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, \underline{f_{Y|X}(y|x)\, f_X(x)}\, dy\, dx$$

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

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

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

Standard Normal $N(0,1)$ mean variance

Example. We saw that if (X, Y) has a standard bivariate normal distribution, then $\mathbb{E}[Y \mid X = x] = \varrho x$. Using the above property of conditional expectation we see

$$\mathbb{E}Y = \int_{-\infty}^{\infty} \rho x\, f_X(x)\, dx = \rho \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \rho \times 0 = 0$$

We do not need to change the definition of conditional expectation given a random variable.

Definition. (Conditional expectation given a random variable) The expression $\mathbb{E}[Y \mid X]$ is a random variable $g(X)$ that takes the value $\mathbb{E}[Y \mid X = x]$ when $X = x$.

This conditional expectation has the same properties as in the discrete case.

- Property 1: $\mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}(Y)$
- Property 2: Let $Y_1, Y_2, \ldots, Y_n$ and X be random variables defined in the same random experiment. Then $\mathbb{E}\left[\sum_{i=1}^{n} Y_i \middle| X\right] = \sum_{i=1}^{n} \mathbb{E}[Y_i \mid X]$
- Property 3: If X and Y are independent, then $\mathbb{E}[Y \mid X] = \mathbb{E}[Y]$.

Example. Suppose that $X \sim \mathsf{Uniform}(0, 1)$ is chosen, and then $Y \sim \mathsf{Uniform}(X, 1)$. What is the expected value of Y?

The conditional pdf $f_{Y|X}$ is given by