Statistics 351 (Fall 2015) November 13, 2015
Prof. Michael Kozdron

Lecture #23: The Bivariate Normal Density Function

Reference. §5.5 pages 125–126

Last lecture we derived the density function for the multivariate normal by starting with independent and identically distributed $\mathcal{N}(0,1)$ random variables, performing a linear transformation, and using the change-of-variables formula from Chapter 1.

Definition III. A random vector $\mathbf{X}$ with $\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$ and $\operatorname{Cov}(\mathbf{X}) = \boldsymbol{\Lambda}$ where $\det[\boldsymbol{\Lambda}] > 0$ is $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ if and only if its density function is given by

$$f_{\mathbf{X}}(\boldsymbol{x}) = f_{\mathbf{X}}(x_1, \ldots, x_n) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \frac{1}{\sqrt{\det[\boldsymbol{\Lambda}]}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Lambda}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}.$$

The case of the bivariate normal is of particular importance. Let $\mathbf{X} = (X_1, X_2)' \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ with

$$\boldsymbol{\mu} = \begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix} \quad \text{and} \quad \boldsymbol{\Lambda} = \begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}.$$

In particular, $X_j \in \mathcal{N}(\mu_j, \sigma_j^2)$, $j = 1, 2$, with $\operatorname{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$. Notice that

$$\operatorname{Corr}(X_1, X_2) = \frac{\operatorname{Cov}(X_1, X_2)}{\sqrt{\operatorname{Var}(X_1)\operatorname{Var}(X_2)}} = \frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2} = \rho.$$

Furthermore,

$$\det[\boldsymbol{\Lambda}] = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1-\rho^2).$$

Provided $\sigma_1 > 0$, $\sigma_2 > 0$, and $-1 < \rho < 1$, then $\det[\boldsymbol{\Lambda}] > 0$ so that $\mathbf{X}$ has a density. Since

$$\boldsymbol{\Lambda}^{-1} = \frac{1}{\det[\boldsymbol{\Lambda}]}\begin{bmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{bmatrix} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{bmatrix}$$
$$= \frac{1}{1-\rho^2}\begin{bmatrix}1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2\end{bmatrix},$$

we conclude that

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right\}. \tag{$*$}$$

In the special case that $\boldsymbol{\mu} = \mathbf{0}$, $\sigma_1^2 = \sigma_2^2 = 1$, we find

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(x_1^2 - 2\rho x_1 x_2 + x_2^2\right)\right\}.$$

Remark. The density is nicely typed on page 126. At the top of page 126, a general formula for $\boldsymbol{\Lambda}^{-1}$ is used (and given in equation (1.6) on page 119). We will not require this fact, and will only be concerned with the bivariate density as derived above.

Conditional Distributions for the Bivariate Normal

Reference. §5.6 pages 127–130

Suppose that $\mathbf{X} = (X, Y)' \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Lambda} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}.$$

In particular, $X \in \mathcal{N}(\mu_x, \sigma_x^2)$, $Y \in \mathcal{N}(\mu_y, \sigma_y^2)$, and $\operatorname{Corr}(X, Y) = \rho$. Suppose further that $\det[\boldsymbol{\Lambda}] > 0$ so that $\mathbf{X}$ has a density given by Definition III.

Goal. To compute the conditional density function for $Y|X = x$.

By definition,

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

Since $X \in \mathcal{N}(\mu_x, \sigma_x^2)$ we know that

$$f_X(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_x^2}(x - \mu_x)^2\right\}.$$

We also know from $(*)$ that the joint density $f_{X,Y}(x, y)$ is given by

$$f_{X,Y}(x,y)$$
$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right)\right\}.$$

Dividing $f_{X,Y}(x, y)$ by $f_X(x)$ and simplifying gives

$$f_{Y|X=x}(y) = \frac{1}{\sigma_y\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2\sigma_y^2(1-\rho^2)}\left(y - \mu_y - \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right)^2\right\}.$$

In other words, we see that

$$Y|X = x \in \mathcal{N}\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1-\rho^2)\right).$$

In particular,

$$\mathbb{E}(Y|X) = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(X - \mu_x)$$

and

$$\operatorname{Var}(Y|X) = \sigma_y^2(1-\rho^2).$$