Cardiff University School of Mathematics

MA2500

Foundations of Probability and Statistics

Reading Material

2014-15

Version: 3/12/2014

Contents

Lecture 20 The Bivariate Normal Distribution

To be read in preparation for the **11.00** lecture on **Wed 03 Dec** in **Physiology A**.

20.1 Bivariate transformations

Definition 20.1
Let $h : \mathbb{R}^2 \to \mathbb{R}^2$ and let $(u, v) = h(x, y)$. The *Jacobian determinant* of the transformation h is the determinant of its 2×2 matrix of partial derivatives:

$$J = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$$

Theorem 20.2
Let U and V be jointly continuous random variables, let $f_{U,V}$ be their joint PMF, let $g : \mathbb{R}^2 \to \mathbb{R}^2$ be an injective transform over the support of $f_{U,V}$ and let $(X, Y) = g(U, V)$. Then the joint PMF of X and Y is given by

$$f_{X,Y}(x, y) = |J| f_{U,V}\left[g^{-1}(x, y)\right] \quad \text{where} \quad J = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} \quad \text{with} \quad (u, v) = g^{-1}(x, y).$$

Remark 20.3
The absolute value $|J|$ is a scale factor, which ensures that $f_{X,Y}(x, y)$ integrates to one.

Example 20.4
Let U and V be continuous random variables, and let $X = U + V$ and $Y = U - V$.

(1) Find the joint PDF of X and Y in terms of the joint PDF of U and V.

(2) If $U, V \sim$ Exponential(1) are independent, find the joint PDF of X and Y.

Solution:

(1) • The transformation $g : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $g(u, v) = (u + v, u - v)$.

- To compute the inverse transformation $g^{-1}: \mathbb{R}^2 \to \mathbb{R}^2$, we solve the equations

$$x = u + v \quad \text{and} \quad y = u - v.$$

- This yields $u = \frac{1}{2}(x+y)$ and $v = \frac{1}{2}(x-y)$.
- Thus the inverse transformation is

$$(u,v) = g^{-1}(x,y) = \left[\frac{1}{2}(x+y), \frac{1}{2}(x-y)\right].$$

The Jacobian determinant is given by

$$J = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}.$$

Hence the joint PDF of X and Y is

$$\begin{aligned} f_{X,Y}(x,y) &= |J| f_{U,V}(u,v) \\ &= \left|-\frac{1}{2}\right| f_{U,V}\left[\frac{1}{2}(x+y), \frac{1}{2}(x-y)\right] \\ &= \frac{1}{2} f_{U,V}\left[\frac{1}{2}(x+y), \frac{1}{2}(x-y)\right]. \end{aligned}$$

(2) Let U and V be independent with $U, V \sim \text{Exponential}(1)$.
By independence, the joint PDF of U and V is

$$f_{U,V}(u,v) = \begin{cases} e^{-(u+v)} & u,v > 0 \\ 0 & \text{otherwise.} \end{cases}$$

To compute the support of $f_{X,Y}$, since $u > 0$ and $v > 0$ we have $x > 0$, so

- $\min(y) = \min(u - v) = -x$ (which occurs when $u = 0$ and $v = x$), and
- $\max(y) = \max(u - v) = x$ (which occurs when $u = x$ and $v = 0$).

Thus, substituting for $u + v = \frac{1}{2}(x+y) + \frac{1}{2}(x-y) = x$, we obtain

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2}e^{-x} & \text{for } x > 0 \text{ and } -x < y < x, \\ 0 & \text{otherwise.} \end{cases}$$

20.2 The bivariate normal distribution

Theorem 20.5
if $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, then

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

[*Proof omitted.*]

Corollary 20.6
If $U, V \sim N(0,1)$ are independent, then $aU + bV \sim N(0, a^2 + b^2)$ for all $a, b \in \mathbb{R}$.

Definition 20.7
A pair of random variables U and V have the *standard bivariate normal distribution* if their joint PMF $f : \mathbb{R}^2 \to [0,\infty)$ can be written as

$$f_{U,V}(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\big(u^2 - 2\rho uv + v^2\big)\right)$$

where ρ is a constant satisfying $-1 < \rho < 1$.

Definition 20.8
A pair of random variables X and Y are said to have *bivariate normal distribution* with means μ_1 and μ_2, variances σ_1^2 and σ_2^2 and correlation ρ, if their joint PMF can be written as

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right)$$

The following lemma can be used to derive many properties of the bivariate normal distribution.

Lemma 20.9
Let $U, V \sim N(0,1)$ be independent, let $\rho \in (-1,+1)$. Then the random variables

$$\begin{aligned} X &= \mu_1 + \sigma_1 U, \\ Y &= \mu_2 + \sigma_2\big(\rho U + \sqrt{1-\rho^2}V\big) \end{aligned}$$

have bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ.

Proof: To find the joint PDF of X and Y, let $g(u,v)$ denote the transformation:

$$g(u,v) = \big[\mu_1 + \sigma_1 u, \mu_2 + \sigma_2(\rho u + \sqrt{1-\rho^2}v)\big].$$

The inverse transformation is

$$g^{-1}(x,y) = \left(\frac{x-\mu_1}{\sigma_1}, \frac{1}{\sqrt{1-\rho^2}}\left[\left(\frac{y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right]\right)$$

The joint PDF of X and Y is $f_{X,Y}(x,y) = |J| f_{U,V}(u,v)$, where J is the Jacobian determinant of the inverse transformation:

$$J = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{\sigma_1} & 0 \\ \dfrac{1}{\rho\sigma_1} & \dfrac{1}{\sigma_2\sqrt{1-\rho^2}} \end{vmatrix} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

Because U and V are independent,

$$f_{U,V}(u,v) = f_U(u) f_V(v) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(u^2+v^2)\right) \qquad u,v \in \mathbb{R}.$$

and since

$$\begin{aligned} u^2 + v^2 &= \left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \frac{1}{1-\rho^2}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \rho^2\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \\ &= \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right] \end{aligned}$$

it follows that

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_1\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right)$$

as required.

The following theorem shows that if X and Y have bivariate normal distribution, then any linear combination of X and Y is normally distributed.

Theorem 20.10
If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ then

$$aX + bY \sim N\big(a\mu_1 + b\mu_2, a^2\sigma_1^2 + 2ab\sigma_1\sigma_2\rho + b^2\sigma_2^2\big)$$

Proof: Let $Z = aX + bY$, let U and V be independent standard normal random variables, and let

$$\begin{aligned} X' &= \mu_1 + \sigma_1 U \\ Y' &= \mu_2 + \sigma_2\big(\rho U + \sqrt{1-\rho^2}V\big) \end{aligned}$$

By Lemma 20.9, X and Y have the same joint distribution as X' and Y', so $Z = aX + bY$ has the same distribution as

$$Z' = aX' + bY' = (a\mu_1 + b\mu_2) + (a\sigma_1 + b\sigma_2\rho)U + b\sigma_2\sqrt{1-\rho^2}V$$

Because $U, V \sim N(0,1)$ are independent, it follows by Corollary 20.6 that

$$Z' \sim N\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + 2ab\sigma_1\sigma_2\rho + b^2\sigma_2^2\right),$$

so $Z = aX + bY$ has normal distribution, as required.

20.3 Properties of the bivariate normal distribution

Theorem 20.11
Let X and Y have bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ. Then

(1) $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$,

(2) ρ is the correlation coefficient of X and Y, and

(3) X and Y are independent if and only if $\rho = 0$.

Proof: Let $U, V \sim N(0,1)$ and define

$$\begin{aligned} X &= \mu_1 + \sigma_1 U \\ Y &= \mu_2 + \sigma_2\big(\rho U + \sqrt{1-\rho^2}V\big) \end{aligned}$$

(1) In the proof of Theorem 20.10:

- taking $a = 1$ and $b = 0$ yields $X \sim N(\mu_1, \sigma_1^2)$, and
- taking $a = 0$ and $b = 1$ yields $Y \sim N(\mu_2, \sigma_2^2)$.

(2) Using the fact that $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$ for all $a, b, c, d \in \mathbb{R}$,

$$\begin{aligned} \text{Cov}(X,Y) &= \text{Cov}\big[\mu_1 + \sigma_1 U, \mu_2 + \sigma_2(\rho U + \sqrt{1-\rho^2}V)\big] \\ &= \sigma_1\sigma_2\text{Cov}(U, \rho U + \sqrt{1-\rho^2}V) \\ &= \sigma_1\sigma_2\big[\rho\mathbb{E}(U^2) + \sqrt{1-\rho^2}\mathbb{E}(UV)\big] \\ &= \sigma_1\sigma_2\rho. \end{aligned}$$

Thus $\rho = \dfrac{\text{Cov}(X)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$ as required.

(3) If X and Y are independent, they are uncorrelated. If X and Y are uncorrelated then $\rho = 0$, so

the joint PDF of X and Y satisfies

$$\begin{aligned}
f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) \times \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right) \\
&= f_X(x)f_Y(y).
\end{aligned}$$

Because this holds for all $x, y \in \mathbb{R}$, it follows that X and Y are independent.

20.4 Conditional distributions

Theorem 20.12
Let X and Y have bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ. Then the conditional distribution of Y given $X = x$ is also normal, with conditional mean and variance given by

$$\begin{aligned}
\mathbb{E}(Y|X=x) &= \mu_2 + \rho\left(\frac{\sigma_2}{\sigma_1}\right)(x-\mu_1), \\
\mathrm{Var}(Y|X=x) &= \sigma_2^2(1-\rho^2),
\end{aligned}$$

and the conditional mean and variance of Y given X is

$$\begin{aligned}
\mathbb{E}(Y|X) &= \mathbb{E}(Y) + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\big[X - \mathbb{E}(X)\big], \\
\mathrm{Var}(Y|X) &= \mathrm{Var}(Y)(1-\rho^2).
\end{aligned}$$

Proof: Let $U, V \sim N(0,1)$ be independent, and define the random variables

$$\begin{aligned}
X &= \mu_1 + \sigma_1 U, \\
Y &= \mu_2 + \sigma_2\big[\rho U + \sqrt{1-\rho^2}V\big] \\
&= \mu_2 + \sigma_2\left[\rho\left(\frac{X-\mu_1}{\sigma_1}\right) + \sqrt{1-\rho^2}V\right].
\end{aligned}$$

If X is fixed at x, then Y is a linear transformation of V, so the conditional distribution of Y given that $X = x$ is a normal distribution. Furthermore, since $\mathbb{E}(V) = 0$ and $\mathrm{Var}(V) = 1$ we have

$$\begin{aligned}
\mathbb{E}(Y|X=x) &= \mu_2 + \rho\left(\frac{\sigma_2}{\sigma_1}\right)(x-\mu_1) \\
\mathrm{Var}(Y|X=x) &= \sigma_2^2(1-\rho^2)
\end{aligned}$$

as required.

20.5 Exercises

Exercise 20.1

1. (a) Let X and Y be jointly continuous random variables, and let $f_{X,Y}$ be their joint PDF. Show that the PDF of the random variable $X+Y$ can be written as

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

   (b) Show that if $U, V \sim N(0,1)$ are independent, then $U+V \sim N(0,2)$. (This is a special case of Theorem 20.5.)

2. Let $U, V \sim N(0,1)$. Show that the random variables $X = U+V$ and $Y = U-V$ are independent.