STOR 435.001 Lecture 15

Jointly distributed Random Variables - III

Jan Hannig

UNC Chapel Hill

Before we dive in

Contents of this lecture

1. Conditional pmf/pdf: definition and simple properties.
2. Functions of **two random variables**. Finding joint distributions.

Brief aside

1. How many of you have heard of the statistical technique "regression"
2. `https://idc9.github.io/stor390/notes/linear_regression/linear_regression.html`
3. Also see `https://www.coursera.org/learn/machine-learning/lecture/Cf8DF/multivariate-gaussian-distribution`
4. Then see: `https://www.coursera.org/learn/machine-learning/lecture/DnNr9/anomaly-detection-using-the-multivariate-gaussian-distribution`
5. We will see connections between the above and the next topic "Conditional pmf and pdfs".

Conditional distributions: discrete case: X and Y are discrete with joint p.m.f. $p(x, y)$:

Conditional p.m.f. of X given $Y = y$ (y fixed):

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}$$

for all y with $p_Y(y) > 0$.

Note: If X and Y are independent, then $p_{X|Y}(x|y) = p_X(x)$.

Conditional d.f. of X given $Y = y$ (y fixed):

$$F_{X|Y}(a|y) = P(X \leq a|Y = y) = \sum_{x \leq a} p_{X|Y}(x|y)$$

for all y with $p_Y(y) > 0$.

Important note

In many cases it is easy to calculate marginal pmf of one random variable (say Y) and conditional pmf of the other (say $p_{X|Y}$). Then we can get joint pmf by

$$p_{XY}(x, y) = p_Y(y) p_{X|Y}(x|y)$$

Example: Let X be the number of claims submitted to a life-insurance company in April and let Y be the corresponding number but for May. Suppose the joint pmf of the two random variables is given by

$$p_{X,Y}(x, y) = \frac{1}{2}\left(\frac{1}{2}\right)^{x-1} e^{-x}(1 - e^{-x})^{y-1}, \qquad x = 0, 1, \ldots, y = 1, 2, \ldots.$$

Find the conditional pmf of Y given that there were 2 claims in April.

Citation: Motivated by a problem from

`http://faculty.atu.edu/mfinan/actuarieshall/Pbook.pdf`

☼

Solution contd

Conditional distributions: continuous case: X and Y are jointly continuous with density $f(x, y)$:

Conditional density of X given $Y = y$ (y fixed):

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

for all y with $f_Y(y) > 0$.

Note: If X and Y are independent, then $f_{X|Y}(x|y) = f_X(x)$.

Conditional probabilities given $Y = y$ (y fixed):

$$P(X \in A|Y = y) = \int_A f_{X|Y}(x|y)dx.$$

Conditional d.f. of X given $Y = y$ (y fixed):

$$F_{X|Y}(a|y) = P(X \leq a|Y = y).$$

Example: Suppose that the joint density of X and Y is given by

$$f(x,y) = \begin{cases} \frac{e^{-x/y}e^{-y}}{y}, & 0 < x < \infty, 0 < y < \infty, \\ 0, & \text{otherwise.} \end{cases}$$

Find $f_{X|Y}(x|y)$ and $P(X > 1|Y = y)$.

☼

Bivariate Normal distribution: Jointly continuous random variables X and Y are bivariate normal if their density is

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left((\frac{x-\mu_X}{\sigma_X})^2 + (\frac{y-\mu_Y}{\sigma_Y})^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} \right)},$$

for $-\infty < x, y < \infty$, where $\sigma_X > 0, \sigma_Y > 0$, $\rho \in (-1,1)$, $-\infty < \mu_X, \mu_Y < \infty$. Find $f_{X|Y}(x|y)$.

Notation:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} \right).$$

Here, $\mu_X = EX$, $\mu_Y = EY$, $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$. The parameter ρ accounts for dependence and will be clarified in the next chapter.
This and generalizations: multivariate normal are one of the foundation pieces of modern statistics. Linear Regression based on this fundamental object!

Example

Suppose we wanted to plot the pdf of the above function. Let us say $\mu_X = 0$, $\mu_Y = 0$, $\sigma_X^2 = 1$, $\sigma_Y^2 = 1$ and $\rho = .5$. Using Mathematical one can plot the above pdf on the region $\{(x, y) : -2 \leq x \leq 2, -2 \leq y \leq 2\}$

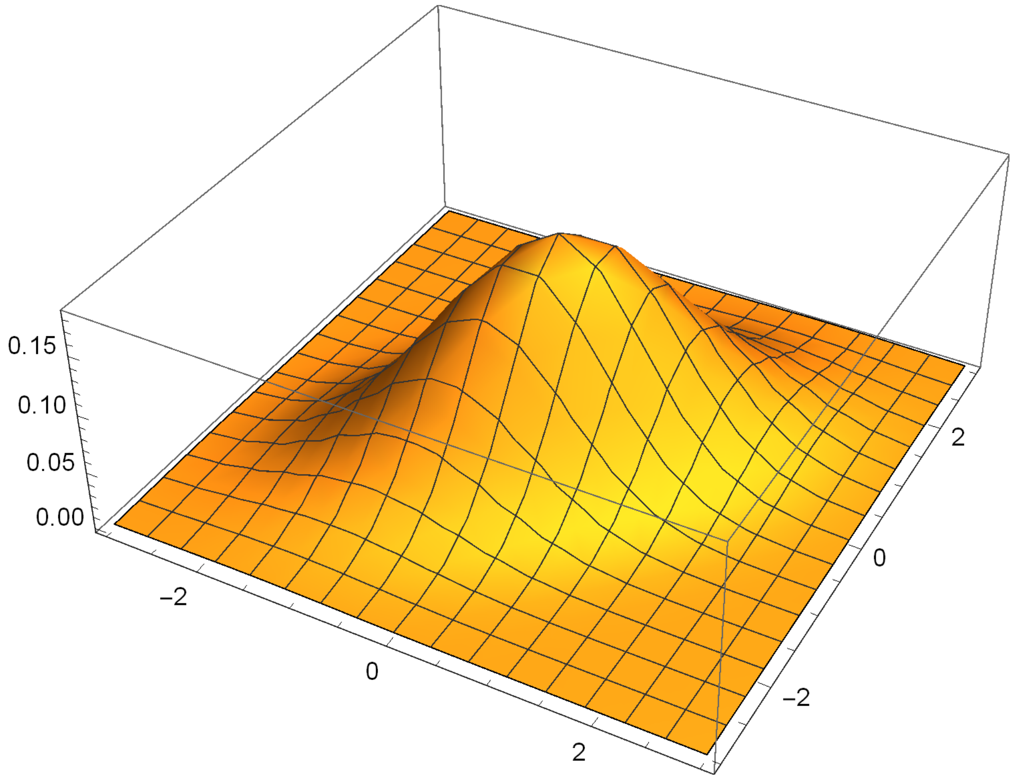

Figure: Using Mathematica

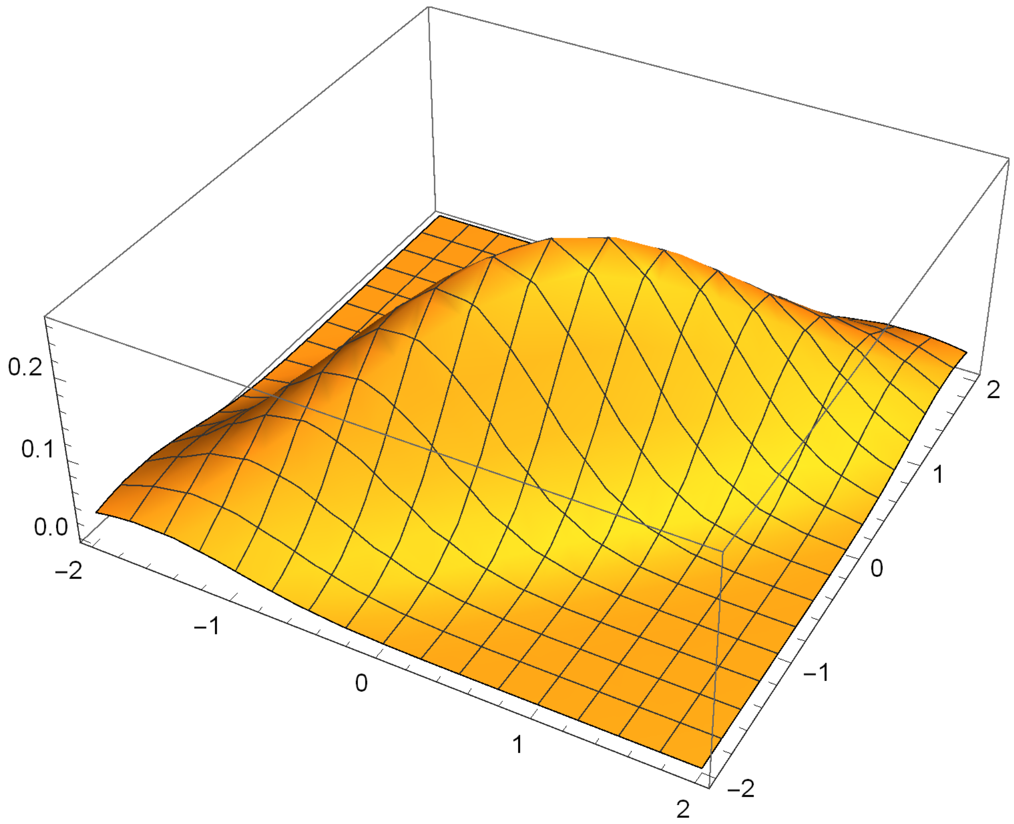

Figure: Using Mathematica. Same parameters as above with only change $\rho = .8$. Higher "correlation" between X and Y. In both cases, the univariate distributions of X and Y turn our to be standard normal. In the second case: X and Y tend to be "closer" whatever that means. Will see more in the next chapter.

(Bivariate Normal distribution) contd:

☼

Look at the exponent in the formula of the Bivariate normal

$$\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)$$

Messy algebra implies that one can write this as

$$=\frac{(x-\mu_X)^2}{2\sigma_X^2}+\frac{1}{2\sigma_Y^2(1-\rho^2)}\left(y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2$$

Thus

$$f_{XY}(x,y)=\underbrace{\frac{1}{\sigma_X\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu_X)^2}{\sigma_X^2}}}_{=\text{pdf of}}\cdot\underbrace{\frac{1}{\sigma_Y\sqrt{1-\rho^2}\sqrt{2\pi}}e^{-\frac{1}{2}\frac{\left(y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2}{\sigma_Y^2(1-\rho^2)}}}_{=\text{pdf of}}$$

(Bivariate Normal distribution) contd:

Thus marginal of X

$$f_X(x) =$$

Conditional pdf of $Y|X$

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

Thus the distribution of Y given that $X = x$ is

1. As I said huge parts of statistics are based on extensions of the above to higher dimensions.
2. For our class: Suppose someone came and asked you to simulate
$$\begin{pmatrix} X \\ Y \end{pmatrix} = N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right).$$
3. Step 1: Simulate Z_1, Z_2 standard Normal random variables.
4. Step 2: $X = \mu_X + \sigma_X Z_1$.
5. Step 3:
$$Y = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X - \mu_X) + Z_2\sigma_Y\sqrt{(1-\rho^2)}$$

A related question

How do you simulate standard normal random variables? Assume there are algorithms to simulate $U(0,1)$ random variables. This is one motivation for the next topic.

Joint probability distribution of functions of random variables

Suppose X_1, X_2 are jointly continuous with density $f(x, y)$. Consider $Y_1 = g_1(X_1, X_2)$, $Y_2 = g_2(X_1, X_2)$, for example, $g_1(x_1, x_2) = x_1 + x_2$ and $g_2(x_1, x_2) = x_1 - x_2$. Suppose the map

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} g_1(x_1, x_2) \\ g_2(x_1, x_2) \end{pmatrix}$$

is continuous, differentiable and invertible. What is the joint density of Y_1, Y_2?

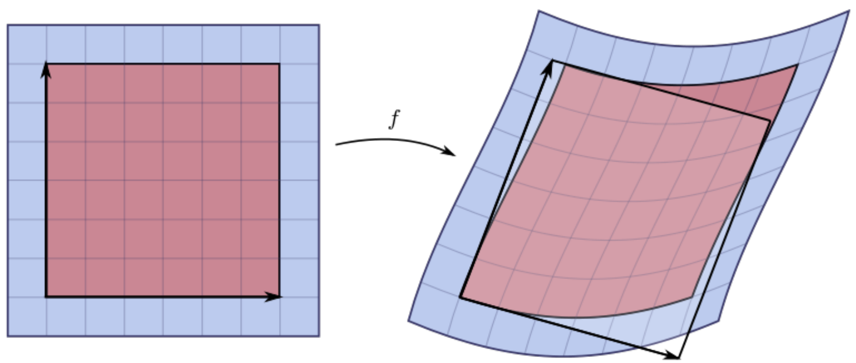

Figure: Jacobian demonstration. Picture from wikipedia.

Joint probability distribution of functions of random variables contd

☼

Jointly distributed random variables (Box Muller transformation)

Let (U, T) be two independent random variables with $U \sim$ Unif$(0, 2\pi)$ and $T \sim \exp(1)$. Consider the transformation

$$X = \sqrt{2T}\cos(U), \qquad Y = \sqrt{2T}\sin(U)$$

What is the joint distribution of (X, Y)?

☼

Example: Let X, Y be independent uniform $(0, 1)$ random variables. Find the distribution (pdf) of $U = XY, V = Y$.

☼