CS 107, Probability, Spring 2020
Lecture 30

Michael Poghosyan
mpoghosyan@aua.am

AUA

17 April 2020

Content

- Joint Distribution
  - Jointly continuous r.v.s and Joint PDFs

Some Notes and Reminders

Note on notations

Note: For points in $\mathbb{R}^2$, we will usually write

$$(x, y) \in \mathbb{R}^2 \qquad \text{or} \qquad x, y \in \mathbb{R}.$$

Also, if we want to describe the point of a rectangle $[a, b] \times [c, d]$, we will write

$$\text{either } (x, y) \in [a, b] \times [c, d] \qquad \text{or} \qquad x \in [a, b],\ \ y \in [c, d].$$

But please note that if we want to describe, say, the open unit disc $D = B\big((0, 0), 1\big)$ centered at the origin, we will write

$$(x, y) \in D \qquad \text{or} \qquad x^2 + y^2 < 1,$$

but writing $x \in [-1, 1]$ and $y \in [-1, 1]$ is incorrect.

Reminder on Level Curves

Assume we have a function $f(x, y)$. Sometimes, drawing the surface of f is not so easy or is not helping us to draw conclusions or do calculations. And in some cases the level-curve map or the contour map is helping us.

Let ℓ be any number. The ℓ-level curve of f is the set of all points (x, y) such that

$$f(x, y) = \ell.$$

Geometrically, the ℓ-level curve is the set of all points (x, y) of the intersection of the surface $z = f(x, y)$ and the horizontal plane $z = \ell$.

Reminder on Level Curves

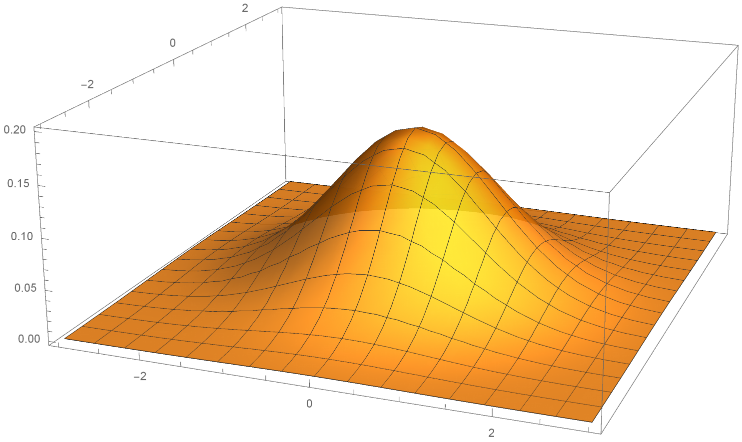


Figure: The surface $z = f(x, y)$

Reminder on Level Curves

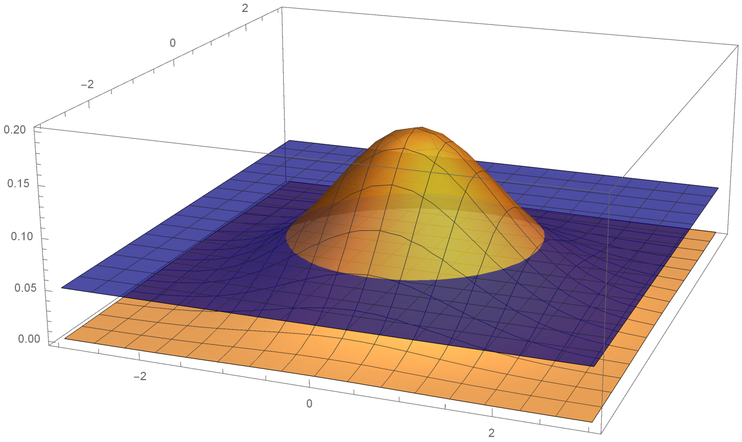


Figure: The surface $z = f(x, y)$ and the plane $z = \ell$

Reminder on Level Curves

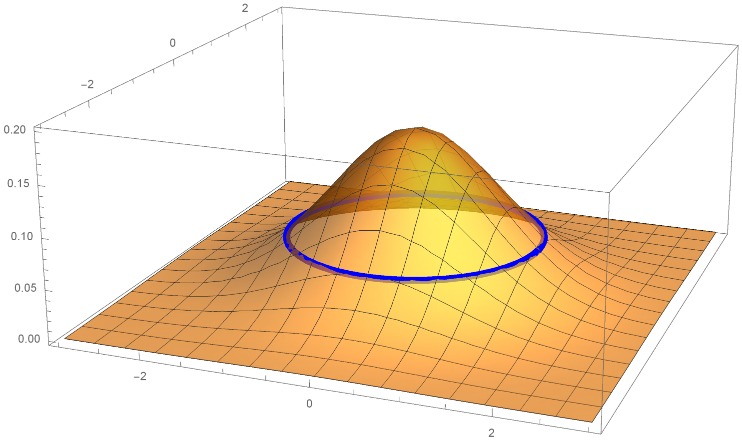


Figure: The intersection of $z = f(x, y)$ and the plane $z = \ell$

Reminder on Level Curves

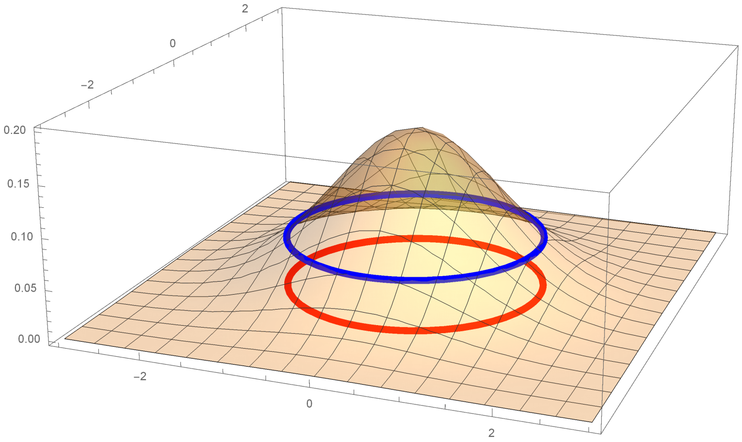


Figure: The intersection curve projected to OXY

Reminder on Level Curves

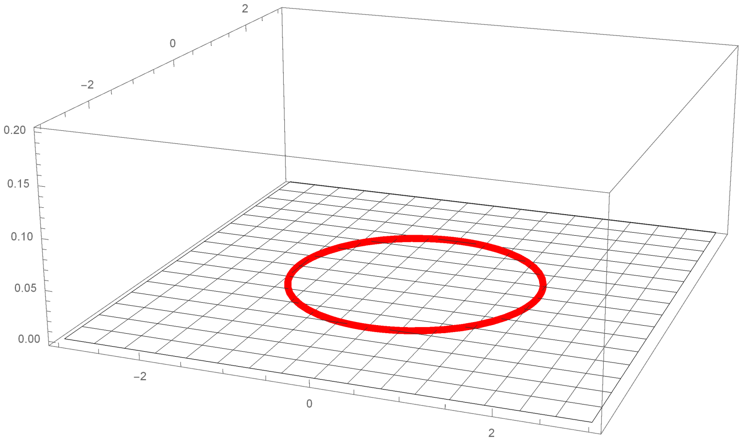


Figure: The intersection curve on OXY

Reminder on Level Curves

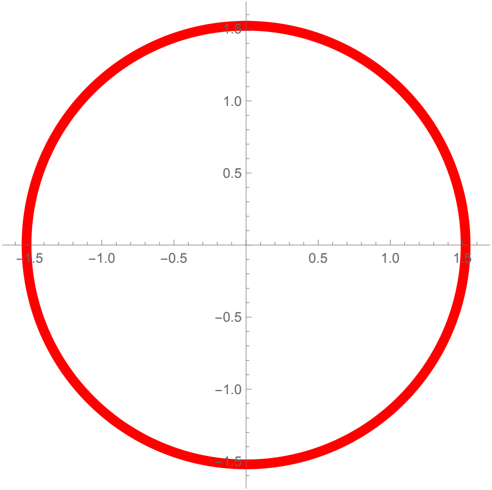


Figure: The ℓ-level curve on OXY

Reminder on Double Integration

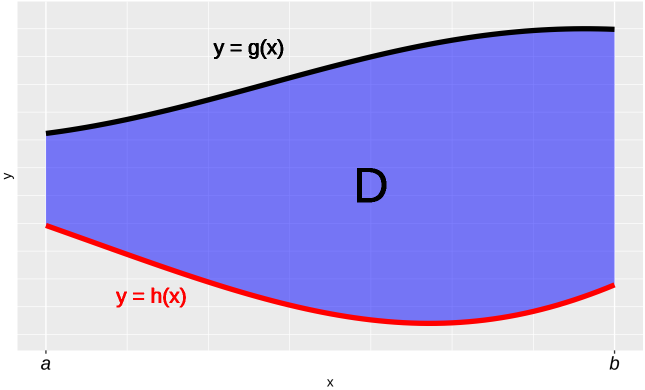
When dealing with 2D continuous r.vectors, we will use double integration to calculate Probabilities (and later, to calculate Expected values and Variances). This is a very quick reminder how to reduce double integrals to iterated ones for some regions, and also how to make a change of variables to pass to polar coordinates.

Assume $D \subset \mathbb{R}^2$ is a 2D region, $f : D \to \mathbb{R}$ is an integrable function, and we want to calculate

$$\iint_D f(x, y)\, dxdy.$$

Reminder on Double Integration

If D has the following shape:

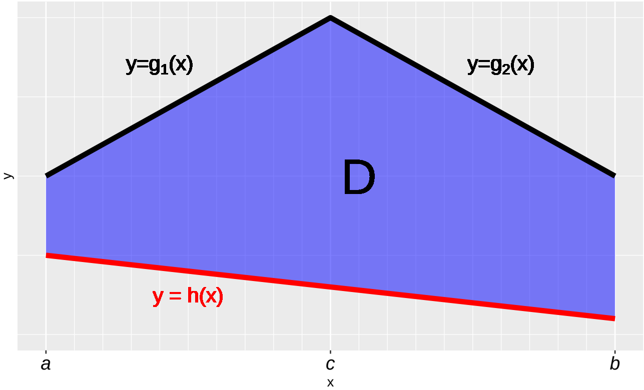


then

$$\iint_D f(x,y)\,dxdy = \int_a^b dx \int_{h(x)}^{g(x)} f(x,y)\,dy.$$

Reminder on Double Integration

Or, if D has the following shape:



then

$$\iint_D f(x, y)\, dxdy = \int_a^c dx \int_{h(x)}^{g_1(x)} f(x, y)\, dy + \int_c^b dx \int_{h(x)}^{g_2(x)} f(x, y)\, dy.$$

Reminder on Double Integration

Now, if the region D and/or function $f(x,y)$ is easy-to-handle-with using polar coordinates, then we can make a change of variables in our double integral. To that end

- we use $x = r \cdot \cos\varphi$ and $y = r \cdot \sin\varphi$;
- describe our integration region D in terms of r and φ - let $\widetilde{D}$ be that region in the (r,φ) coordinate system;
- calculate the absolute value of the Jacobian, which is r in the polar case;
- make the change of variables:

$$\iint_D f(x,y)\,dxdy = \iint_{\widetilde{D}} f(r\cos\varphi, r\sin\varphi) \cdot r\,drd\varphi.$$

Jointly continuous random vectors

Joint distribution of 2 r.vs

Again we will talk about the 2D case, about the distribution of 2 r.v.s X and Y, and later give the generalization ideas for the general, n-dim, case. Last time we considered that the following cases can happen:

- X and Y are discrete r.v.s - so they are Jointly Discrete;
- X and Y are Jointly Continuous r.v.s;
- One of X and Y is Discrete, the other one is Continuous;
- Other

In our last lecture we were talking about the Joint Distribution of Discrete r.v.s. Now we consider the case when X and Y are Jointly continuous.

Reminder: PDF of a Continuous Random Variable

Recall the definition of the Continuous r.v. X: we say that X is (absolutely) continuous, if there exists a non-negative integrable function f such that

$$F(x) = \int_{-\infty}^{x} f(t)\,dt,$$

where F is the CDF of X.

Also, we have the following properties for the PDF f:

- $f(x) \geq 0$ for all $x \in \mathbb{R}$;
- $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Joint PDF of a Continuous Random Vector

In the analogy:

Jointly Continuous R.V.s

We say the the r.v.s X and Y are **Jointly (Absolutely) Continuous**, if there exists a non-negative integrable function $f(u, v)$ defined on $\mathbb{R}^2$ such that

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, dudv, \qquad \forall (x, y) \in \mathbb{R}^2,$$

where F is the Joint CDF of (X, Y).

The function f is called **the Joint PDF of X and Y**, and, if necessary, is denoted by

$$f(u, v) = f_{X,Y}(u, v).$$

Properties of a Joint PDF

It can be proved that

Properties of a Joint PDF

A (double) integrable function $f : \mathbb{R}^2 \to \mathbb{R}$ is a Joint PDF of some r.v. X and Y iff

- $f(u,v) \geq 0$ for any $u, v \in \mathbb{R}$;
- $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u,v)\, dudv = 1.$

In particular, the volume under the Joint PDF surface is 1.

Relationship between the Joint PDF and CDF

Now, the relationship between the Joint CDF and Joint PDF of X and Y:

- If the Joint PDF f is given, then

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, dudv, \qquad \forall (x, y) \in \mathbb{R}^2,$$

- Now, if the Joint CDF F is given, then

$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y}(x, y) \qquad \text{for almost all} \quad (x, y) \in \mathbb{R}^2.$$

Calculating Probabilities using the Joint PDF

Now, having the Joint PDF f of Continuous r.v.s X and Y, we can calculate all probabilities we want for X and Y. In fact, it can be proven that for any $A \subset \mathbb{R}^2$,

$$\mathbb{P}\big((X, Y) \in A\big) = \iint_A f(u, v)\, dudv.$$

In particular,

$$\mathbb{P}(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(u, v)\, dudv.$$

Geometric Interpretation: The Probability that $(X, Y) \in A$ is equal to the Volume of the body under the surface of the Joint PDF graph, over the region A.

Geometric Interpretation

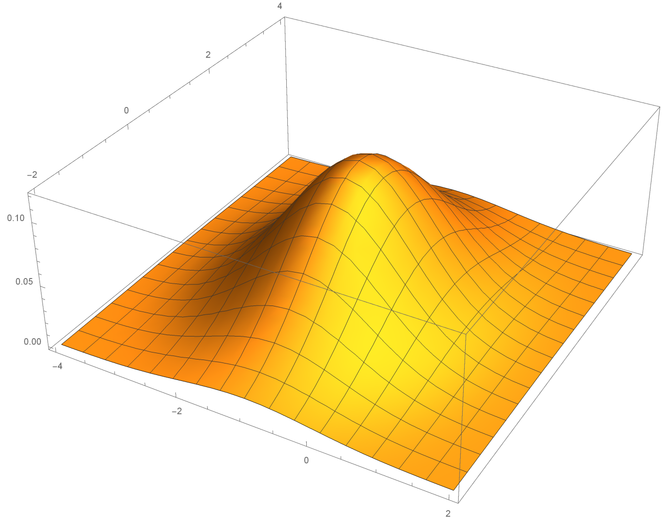


Figure: Joint PDF

Geometric Interpretation

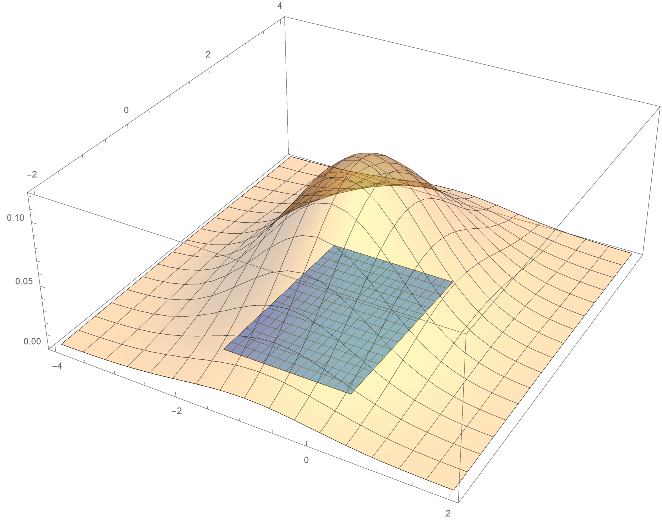


Figure: The region A in OXY plane

Geometric Interpretation

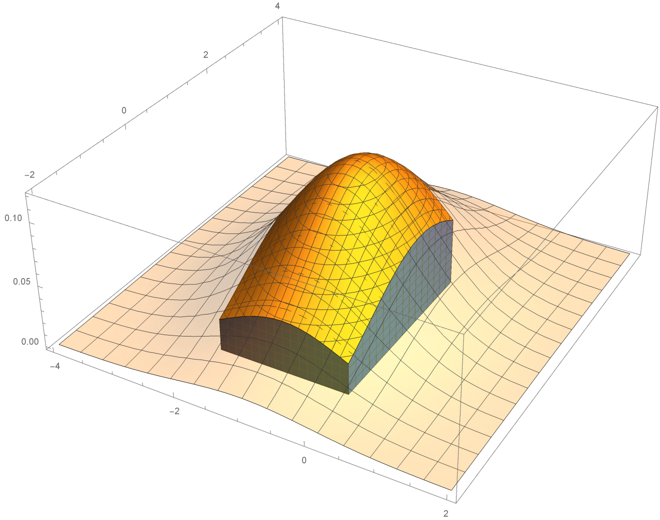


Figure: The volume of this region is $\mathbb{P}\big((X, Y) \in A\big)$

Joint PDF: Note

Note: Recall that, for the 1D case, with 1 continuous r.v. X, for any $a \in \mathbb{R}$, we had

$$\mathbb{P}(X = a) = 0.$$

Now, if X and Y are jointly continuous, then for any fixed point $(a, b) \in \mathbb{R}^2$,

$$\mathbb{P}\Big((X, Y) = (a, b)\Big) = \mathbb{P}(X = a, Y = b) = 0.$$

But in this 2D (and more-D) case, we can state more: if the set $A \subset \mathbb{R}^2$ has an area (measure) equal to 0, i.e., $Area(A) = 0$, then

$$\mathbb{P}\Big((X, Y) \in A\Big) = 0.$$

Joint PDF: Note, Cont'd

For example, if A is some (nice) curve in $\mathbb{R}^2$, then

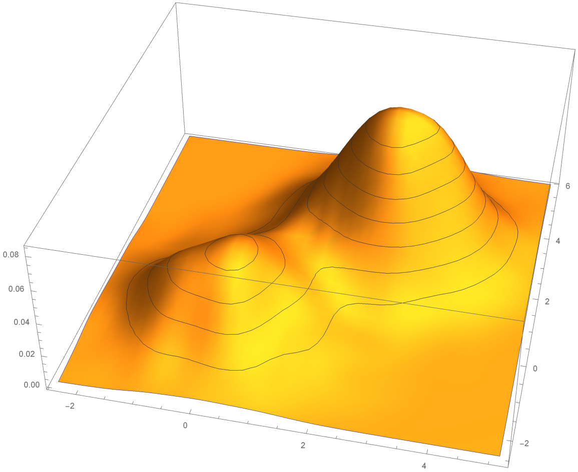
$$\mathbb{P}\Big((X, Y) \in A\Big) = 0.$$

In particular,

$$\mathbb{P}\Big(X = Y\Big) = 0.$$

Joint PDF: Random Variate and PDF

The figure on the next slide is showing one random sample generated from the Joint distribution with the following Joint PDF:



Joint PDF: Random Variate and PDF

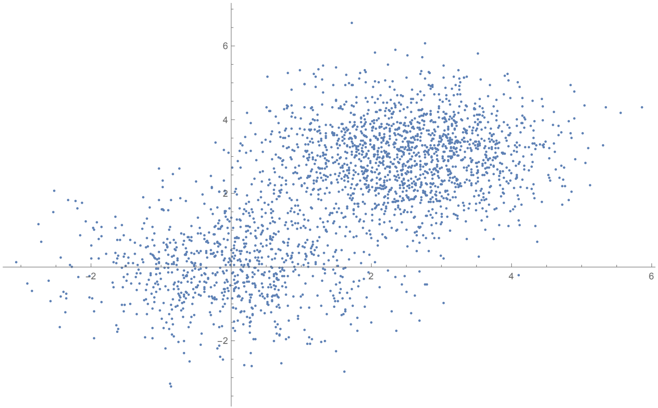


Figure: Dataset generated from the previous Joint PDF

Joint PDF: 2D Histogram and PDF

The figure on the next slide is showing the 2D density histogram of one random sample generated from the Joint distribution with the following Joint PDF:

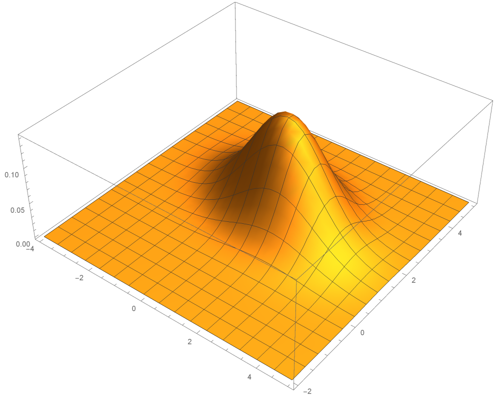


Figure: Joint PDF

Joint PDF: 2D Histogram and PDF

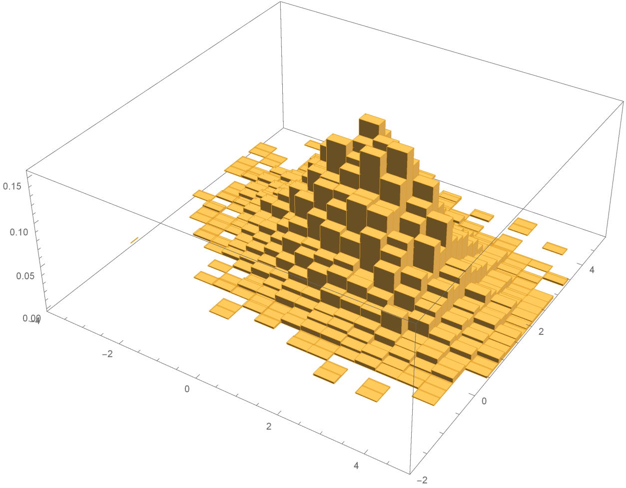


Figure: 2D Density Histogram

Joint PDF: Example

Example 30.1: Let T be the triangle with the vertices at $(0,0), (1,2)$ and $(4,0)$. And let

$$f(x,y) = K, \quad \text{if } (x,y) \in T, \quad \text{and} \quad f(x,y) = 0, \quad \text{if } (x,y) \notin T.$$

be a Joint PDF for some random vector (X, Y).

a. Find K;
b. Calculate $\mathbb{P}(1 \le X \le 2)$;
c. Calculate $\mathbb{P}(X + Y \le 2 \,|\, X < 2)$.

Joint PDF: Example

Example 30.2: Assume the Joint PDF of r.v. X and Y is given by:

$$f(x,y) = \begin{cases} 0.2, & \text{if } x^2+y^2 \leq 1 \\ K, & (x,y) \in [2,3] \times [2,4] \\ 0, & \text{otherwise} \end{cases}$$

a. Find K;
b. Calculate $\mathbb{P}(X^2 + Y^2 = 0.5^2)$;
c. Calculate $\mathbb{P}(X > 0,\ Y > 0)$;
d. Describe where the values of (X, Y) can be;
e. Generate random points from this distribution in **R**.

Note: Please note that, in general,

$$\mathbb{P}(X > a,\ Y > b) \neq 1 - \mathbb{P}(X \leq a,\ Y \leq b).$$

Joint PDF: Example

Example 30.3: Assume

$$f(x,y) = \begin{cases} \frac{1}{3} \cdot \left(1 + 4x^2 y\right), & \text{if } (x,y) \in [0,1] \times [1,2] \\ 0, & \text{otherwise.} \end{cases}$$

a. Show that f is a legitimate Joint PDF;
b. Calculate $\mathbb{P}(Y \leq 1.2)$;
c. Calculate $\mathbb{P}(Y > 1 + X^2)$;
d. On the next few slides you will find the graph of f, random sample generated from this distribution, and some visualizations. Generate, in **R**, a random sample from this distribution.

Joint PDF surface

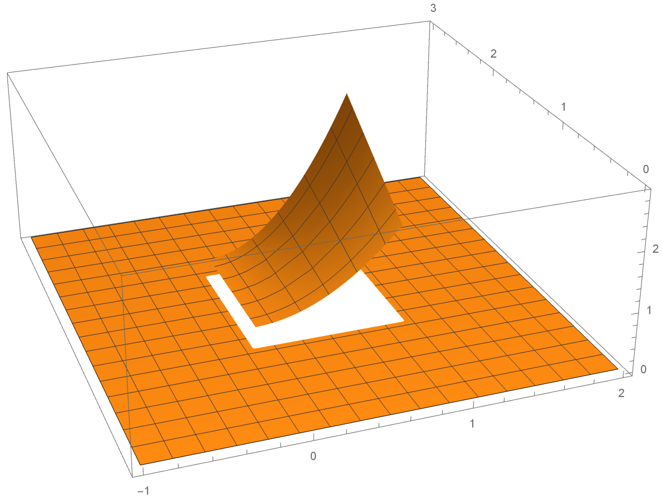


Figure: Joint PDF, $f(x, y) = \frac{1}{3} \cdot \left(1 + 4x^2 y\right)$

Joint PDF: random sample

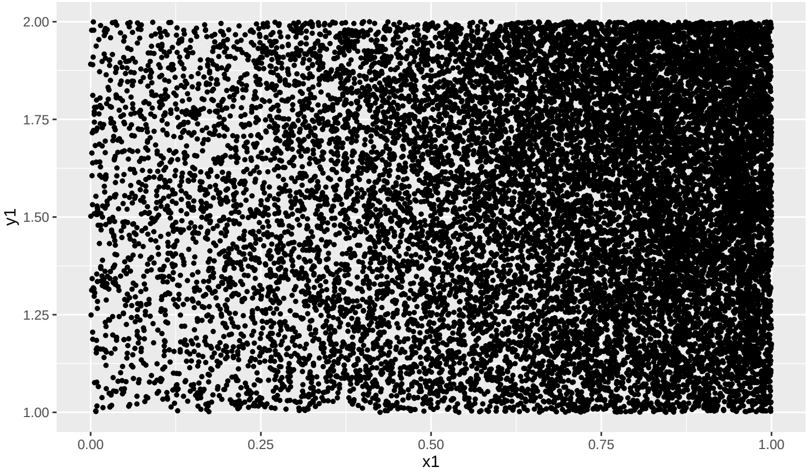


Figure: Random sample

Joint PDF: random sample, heatmap

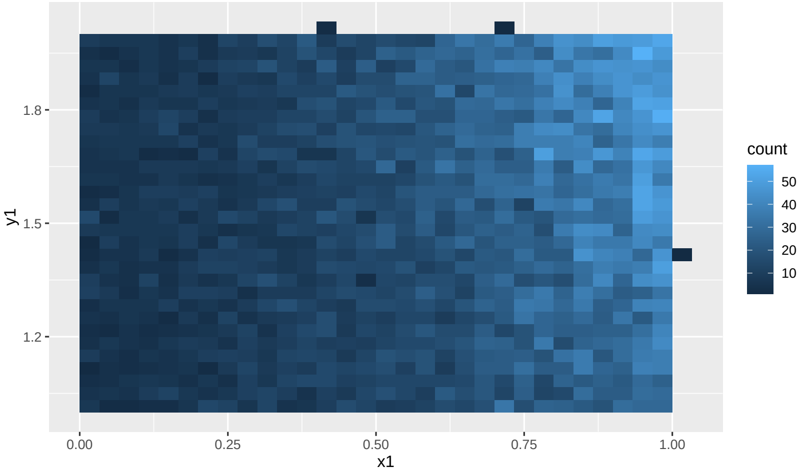


Figure: Random sample, 2D HeatMap

Joint PDF: random sample, heatmap

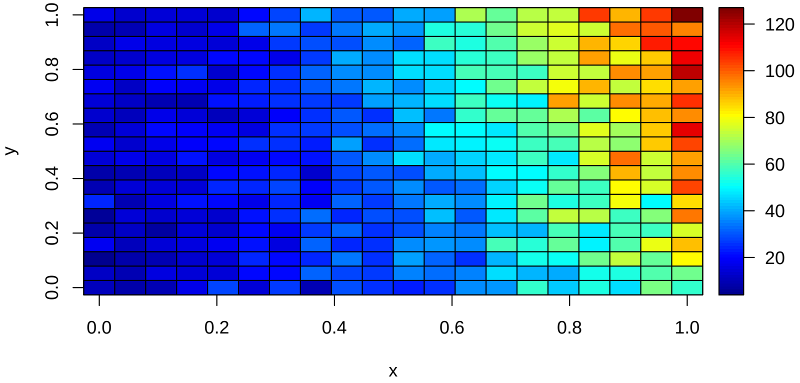


Figure: Random sample, another colored 2D HeatMap

Joint PDF: random sample, heatmap

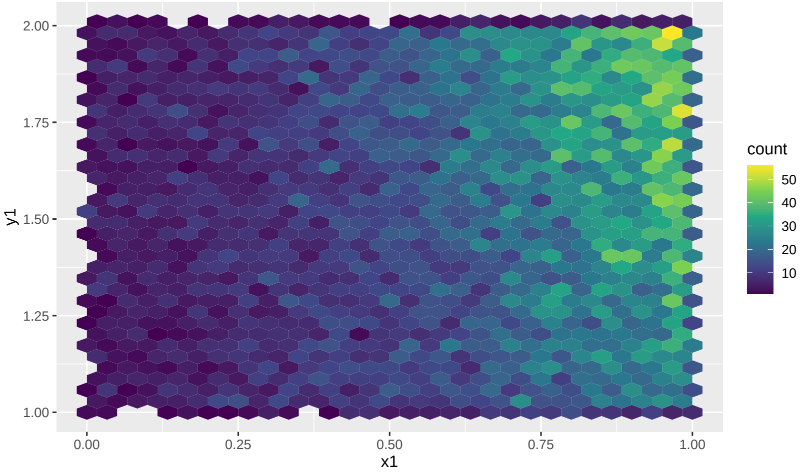


Figure: Random sample, 2D Hexagonal HeatMap

Joint PDF: random sample, 3D Histogram

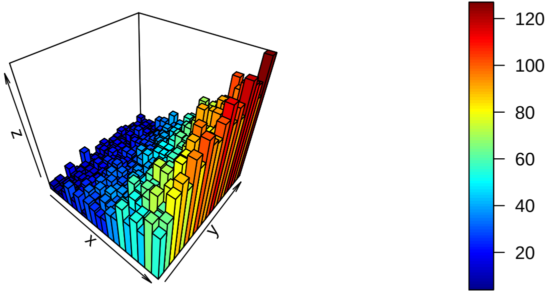


Figure: Random sample, 3D Histogram

Joint PDF: Example

Example 30.4: Assume (X, Y) is a r. vector with the Joint PDF

$$f(x, y) = \begin{cases} K \cdot x \cdot y, & x^2 + y^2 \leq 1, \; x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

a. Find K;
b. Calculate $\mathbb{P}(\,Y < 1 - X | X < 0.5)$.

Obtaining the Marginal PDFs from the Joint PDF

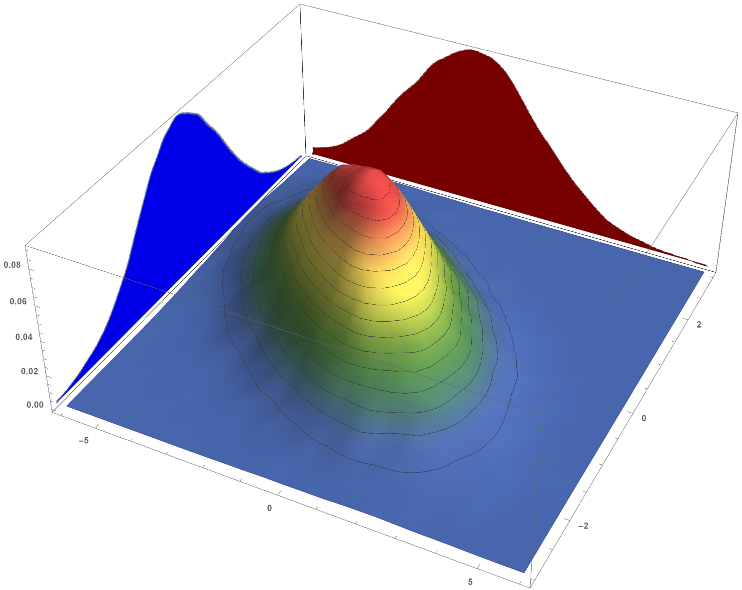
Assume that we know the Joint PDF of X and Y, so we know the Joint Distribution of the random vector (X, Y). Our aim here is to find individual PDFs of X and Y, called Marginal PDFs.

Let (X, Y) is a continuous r.vector with a PDF $f(x, y)$. Then X and Y are both continuous r.v.s and for their PDFs $f_X(x)$ and $f_Y(x)$ we have

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \qquad \forall x \in \mathbb{R},$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx, \qquad \forall y \in \mathbb{R}.$$

Marginal PDFs: Graphically



Marginal PDFs: Graphically

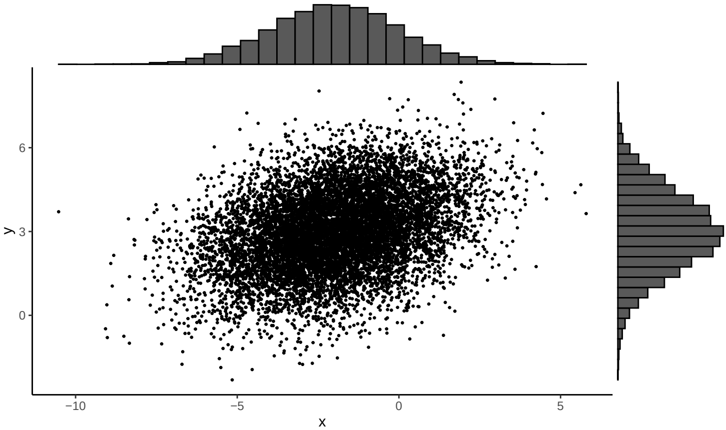


Figure: 2D random numbers and marginal histograms

Marginal PDFs: Example

Example 30.5: Assume (X, Y) is a continuous r.vector with a Joint PDF

$$f(x, y) = \begin{cases} K \cdot (x + y), & x, y \in [0, 2] \\ 0, & \text{otherwise} \end{cases}$$

a. Find K;
b. Calculate $\mathbb{P}(Y = X^2)$;
c. Calculate $\mathbb{P}(Y > X)$;
d. Calculate $\mathbb{P}(X^2 + Y^2 \leq 1)$;
e. Find the Marginal PDFs of X and Y;
f. Calculate $\mathbb{P}(X \in [0, 1])$;
g. Calculate the PDF of the r.v. $Z = Y - X$.

Joint PDF: Example

Example 30.6: Assume the Joint PDF of X and Y is given by

$$f(x,y) = \begin{cases} C \cdot x \cdot e^{-y}, & x \in [0,1], y \geq 0; \\ 0, & \text{otherwise} \end{cases}$$

a. Find the value of C;
b. Calculate $\mathbb{P}(X + Y \leq 4)$;
c. Show how to find the distribution of the r.v. $Z = X^2 + Y^2$.

Joint PDF: Example

Example 30.7: Assume the lifetime of my desktop computer HDD and CPU, in years, are Jointly distributed with the Joint PDF

$$f(x, y) = x \cdot e^{-x \cdot (1+y)}, \qquad x, y \geq 0$$

(and 0 otherwise, of course; here x is for the HDD, and y for CPU). What is the probability that my desktop will work no more that 1 years?

Joint PDF: Example

Example 30.8: Assume $f(x, y)$ is the Joint PDF of X and Y. Find the PDF of the r.v.

$$Z = X + Y.$$

Note

Note (18+): From our last lecture we know that, if X and Y are Jointly Discrete r. v.s, then both X and Y are Discrete r.v.s, and we have calculated their (Marginal) PMFs from the Joint one. Inversely, if X and Y are Discrete r.v.s, then (X, Y) will be a Discrete r. vector.
Now, above, we have stated that if X and Y are jointly continuous r.v.s, then X and Y, separately, are both continuous r.v.s. It is remarkable that the inverse is not true, in general: if X is a continuous r.v. and Y is a continuous r.v., that it can happen that (X, Y) is not jointly continuous r. vector. For example, if X is a continuous r.v., then the r.vector (X, X) is not continuous (try to prove this!).

n-dim continuous r. vectors

Assume now that

$$\mathbf{X} = [X_1, X_2, ..., X_n]^T$$

is an n-dim random vector. Recall that we have defined its CDF by

$$F(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}),$$

where $\mathbf{x} = [x_1, x_2, ..., x_n]^T$, and $\mathbf{X} \leq \mathbf{x}$ means

$$X_1 \leq x_1, \quad X_2 \leq x_2, \quad ..., X_n \leq x_n.$$

Now, we will say that $\mathbf{X}$ is Jointly Continuous r.vector, if there exists a nonnegative integrable function $f(\mathbf{x}) = f(x_1, x_2, ..., x_n)$, called the Joint PDF of $\mathbf{X}$, such that

$$F(\mathbf{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} f(x_1, x_2, ..., x_n)\, dx_1\, dx_2 ... dx_n.$$

n-dim continuous r. vectors

Now, if $\mathbf{X}$ is a Jointly Continuous r.vector with the Joint PDF $f(\mathbf{x})$, then for any (measurable) set $A \subset \mathbb{R}^n$,

$$\mathbb{P}(\mathbf{X} \in A) = \int \cdots \int_{A} f(\mathbf{x})\, d\mathbf{x}.$$

Also, having the Joint PDF of $\mathbf{X}$, we will have that the coordinate r.v.s X_k and all combinations like pairs, triples etc. are continuous, and we can calculate their PDFs by integrating out all other variables. Say, the formula to calculate the PDF of X_1 is

$$f_{X_1}(x_1) = \int \cdots \int_{\mathbb{R}^{n-1}} f(x_1, x_2, ..., x_n)\, dx_2\, dx_3 \ldots dx_n.$$

Supplementary: Conditional Distributions

Conditional Distribution: Discrete case

We have already talked a little bit about the conditional distribution of a discrete r.v., given the value of the other discrete r.v.. Let us formalize that: assume we have Jointly distributed r.v.s X and Y, with values, respectively, x_i and y_j, and assume their Joint PMF is

$$p_{i,j} = \mathbb{P}(X = x_i,\, Y = y_j), \qquad \forall i, j.$$

Now, the Conditional PMF of Y given $X = x_i$ is defined by

$$\mathbb{P}(Y = y_j \,|\, X = x_i) = \frac{\mathbb{P}(X = x_i,\, Y = y_j)}{\mathbb{P}(X = x_i)}, \qquad \forall j.$$

I was denoting this as the distribution of $Y|X = x_i$.

Conditional Distribution: Continuous case

Now, assume that X and Y are Jointly Continuous r.v.s with the Joint PDF $f(x, y)$, and let $f_X(x)$ be the (Marginal) PDF of X. Then we define the Conditional PDF of Y given $X = x$ by

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}, \qquad \forall y \in \mathbb{R},$$

This Conditional PDF is defined for all x such that $f_X(x) > 0$. I am using the following notation for $f_{Y|X}(y|x)$:

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

and interpret that as the PDF of the r.v. $Y|X = x$, or the PDF of Y given that $X = x$.

Conditional Distribution: Continuous case

In that case, it is easy to prove that

$$\int_{\mathbb{R}} f_{Y|X=x}(y)\, dy = 1,$$

so $f_{Y|X=x}$ is a legitimate PDF. Now, we can calculate conditional probabilities about Y: say, for $A \subset \mathbb{R}$,

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

Conditional Distribution: Bayes Rule

In the analogy of the above, we can define also the conditional PDF $f_{X|\,Y=y}(x) = f_{X|\,Y}(x|y)$:

$$f_{X|\,Y=y}(x) = f_{X|\,Y}(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad x \in \mathbb{R}.$$

Then, clearly,

$$f_{X|\,Y}(x|y) = \frac{f_{Y|X}(y|x) \cdot f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x) \cdot f_X(x)}{\int_{\mathbb{R}} f_{Y|X}(y|t) \cdot f_X(t)\,dt}, \quad x, y \in \mathbb{R},$$

which is called the Bayes Rule for continuous r.v.s.

Joint PDF: Example

Example 30.9: Assume the Joint PDF of X and Y is given by

$$f(x,y) = \begin{cases} K, & x^2+y^2 \le 1, \;\; y \ge 0; \\ 0, & \text{otherwise} \end{cases}$$

a. Find the value of K;
b. Find the Marginal PDF of X, $f_X(x)$;
c. Find the Conditional Distribution of $Y|X=x$;
d. Generate a random sample (variate) from this distribution in **R**.

Joint PDF: Example

Example 30.10: Person A is choosing a real number at random from $[0, 1]$ (uniformly). If the number chosen is x, then the second person is choosing at random a number from $[-x, x]$. What is the Joint distribution of the chosen pair of numbers?

Conditional Distribution: Mixed case

Assume now X and Y are Jointly distributed r.v.s, and assume X is discrete with values $x_1, x_2, \dots$ and PMF $\mathbb{P}(X = x_i)$. Next, assume, for each i, $Y|X = x_i$ is continuous with PDF $f_i(y) = f_{Y|X=x_i}(y)$. Then, say, we can calculate the following type probabilities: $\mathbb{P}(Y \in A)$.

$$\mathbb{P}(Y \in A) \overset{TPF}{=\!=\!=} \sum_i \mathbb{P}(Y \in A \,|X = x_i) \cdot \mathbb{P}(X = x_i) =$$

$$= \sum_i \mathbb{P}(X = x_i) \cdot \int_A f_i(t)\, dt.$$

Conditional Distribution: Mixture distribution

In particular, if, for any y, we will take $A = (-\infty, y]$, then we'll get

$$F_Y(y) = \mathbb{P}(Y \leq y) = \sum_i \mathbb{P}(X = x_i) \cdot \int_{-\infty}^{y} f_i(t)\,dt.$$

And calculating the derivatives of both sides w.r.t. y, we'll obtain

$$f_Y(y) = \sum_i \mathbb{P}(X = x_i) \cdot f_i(y) = \sum_i \mathbb{P}(X = x_i) \cdot f_{Y|X=x_i}(y).$$

Joint PDF: Example

Example 30.11: Assume the proportion of females in some country is 0.6. Assume also that the distribution of female heights in that country is Normal with the mean 160 (cm) and standard deviation 5 cm, and the distribution of male heights is Normal with the mean 170 (cm) and standard deviation 6 cm . Find the distribution of the person height in that country, and calculate the probability that the height of a person will be between 150 and 180 cm.

Conditional Distribution: Mixed case

Now, the other case: assume X and Y are Jointly distributed r.v.s, and assume X is continuous with the PDF $f_X(x)$. Next, assume, for each possible value x of X, $Y|X=x$ is discrete with PMF $\mathbb{P}(Y=y_j|X=x)$. Then, say, we want to calculate the following type probabilities: $\mathbb{P}(Y=y_j)$. Then, in the analogy of the discrete case, we will have

$$\mathbb{P}(Y=y_j) \stackrel{TPF}{=\!=\!=} \int_{\mathbb{R}} \mathbb{P}(Y=y_j|X=x)\cdot f_X(x)\,dx.$$

Joint PDF: Example

Example 30.12: Assume p is a r.v. with the PDF $f(x) = 6x(1-x)$, $x \in [0,1]$, and 0 otherwise, and assume that $Y \sim Bernoulli(p)$. That is, Y is Bernoulli r.v. with uncertain probability p, which is, most probably, close to 0.5. Calculate

$$\mathbb{P}(Y=0) \qquad \text{and} \qquad \mathbb{P}(Y=1).$$