Multivariate Distributions

Concerns the distribution of multiple RVs.
Particularly the relationship between them.

Bivariate Distributions (two RVs).

Discrete $\xleftrightarrow[\text{CDF}]{}$ Continuous

Joint Distribution.

Discrete Case Let X, Y be two discrete RVs.

the joint distribution

$$p(x,y) := \mathbb{P}(X = x, Y = y)$$

Note: $p(x,y) \neq 0$ for a discrete set of (x,y).

Properties

① $0 \leq p(x,y) \leq 1$ for all x, y.

② $\sum_{x} \sum_{y} p(x,y) = 1$

↳ sum over all x, y s.t. $p(x,y) \neq 0$.

Probability Table

$X \backslash Y$	y_1	y_2	$\cdots$	y_m
x_1	$p(x_1,y_1)$	$p(x_1,y_2)$		$p(x_1,y_m)$
x_2	$p(x_2,y_1)$	$p(x_2,y_2)$		
$\vdots$				
x_n	$p(x_n,y_1)$	$p(x_n,y_2)$		$p(x_n,y_m)$

Multivariate Case

$X_1, X_2, X_3, \ldots X_n$ discrete RVs.

Joint distribution

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

① $0 \leq p(x_1, \ldots x_n) \leq 1$

② $\sum p(x_1, x_2, \ldots, x_n) = 1$.

$x_1, x_2, \dots x_n$

Continuous Case

Let X, Y be two continuous RVs.

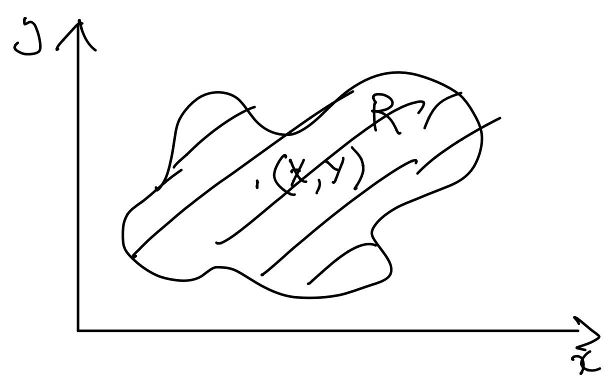
$$P(X = x, Y = y) = 0$$

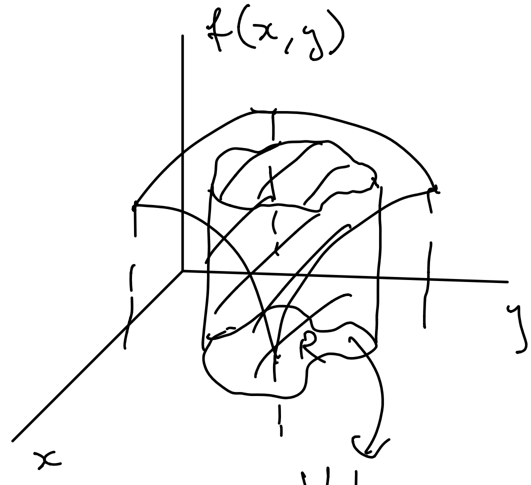
Need a probability density.

Joint probability density.

$f(x,y)$. such that, for any region $R \subseteq \mathbb{R}^2$.

$$P\big((X, X) \in R\big) = \iint_R f(x,y)\, dA.$$





Volum of this shape

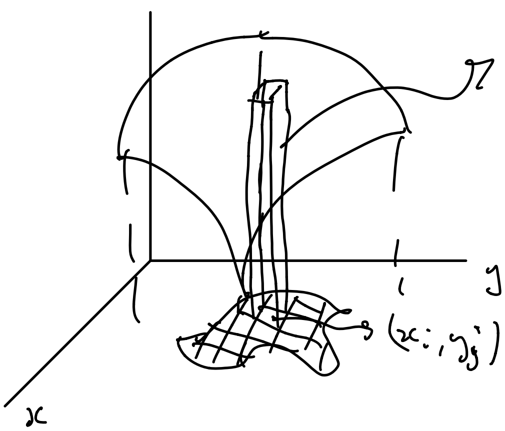
Properties

① $0 \le f(x,y)$

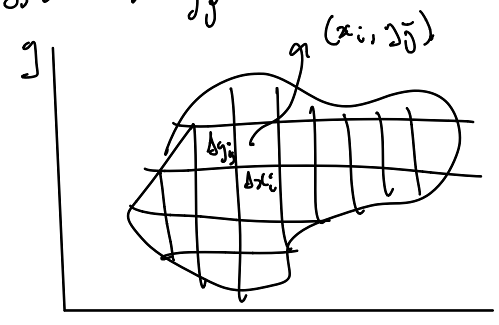
② $\displaystyle\iint_{\mathbb{R}^2} f(x,y)\,dA = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$

Recall (Area Integrals).

$$\iint_R f(x,y)\,dA \simeq \sum_{i,j} \underbrace{f(x_i, y_j)\,\Delta x_i\,\Delta y_j}_{\text{"}P(x_i,y_j)\text{)}}$$



Volume $= f(x_i, y_j)\,\Delta x_i\,\Delta y_j$

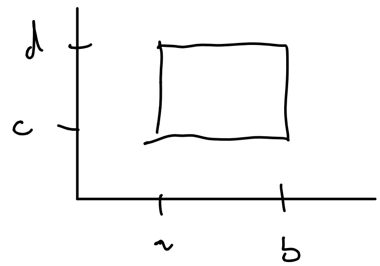


How to compute?

- Use iterated integrals.

1. Rectangular region

$R = [a,b] \times [c,d]$



$$\iint_R f(x,y)\,dA = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx$$

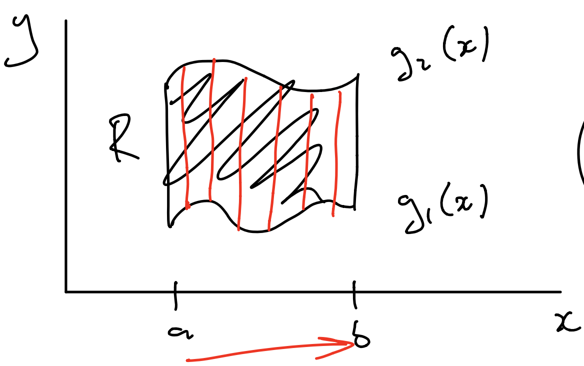
$$= \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy$$

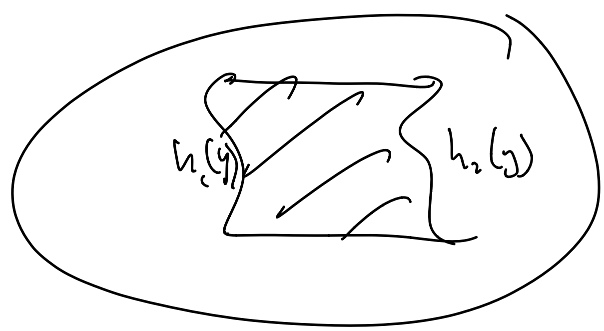
Fubini's Theorem

---

2. Region bounded by two curves.

$R = \{(x,y) : a \le x \le b \,;\, g_1(x) \le y \le g_2(x)\}$.





then

$$\iint_R f(x,y)\,dA = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \right) dx.$$

Example Suppose X, Y have joint pdf

$$f(x,y) = \begin{cases} cxy, & 0 \le x, y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

What value of c makes f a well-defined pdf?

$f(x,y) \geq 0$ if $c \geq 0$.

Normality.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^1 \left(\int_0^1 cxy \, dx \right) dy.$$

$$= \int_0^1 c \tfrac{1}{2} x^2 y \Big|_0^1 dy = \int_0^1 \tfrac{c}{2} y \, dy.$$

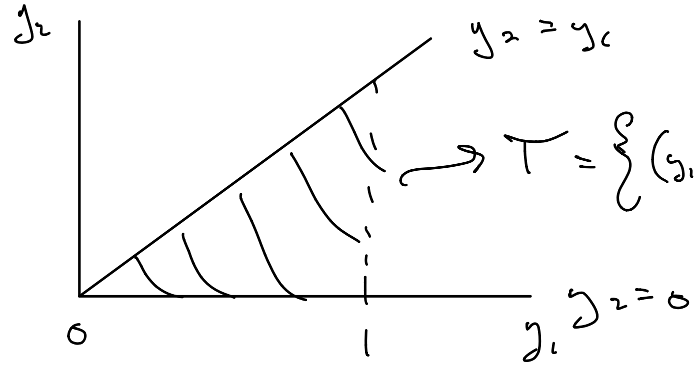
$$= \tfrac{c}{4} y^2 \Big|_0^1 = \tfrac{c}{4} \underset{\text{want}}{=} 1$$

$$\Rightarrow c = 4.$$

Example

Suppose

$$f(y_1, y_2) = \begin{cases} c y_1 & \text{if } 0 \le y_2 \le y_1 \le 1 \\ 0 & \text{otherwise} \end{cases}$$



$$T = \{(y_1, y_2) \in \mathbb{R}^2 ; 0 \le y_1 \le 1, \ 0 \le y_2 \le y_1\}$$

Normality.

$$\iint_{\mathbb{R}^2} f(y_1, y_2)\, dA = \iint_T c y_1 \, dA.$$

$$= \int_0^1 \left(\int_0^{y_1} c y_1 \, dy_2 \right) dy_1$$

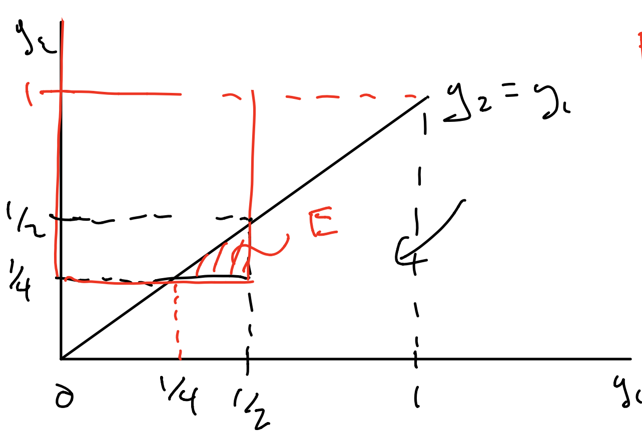
$$= \int_0^1 c y_1 y_2 \Big|_0^{y_1} dy_1 = \int_0^1 c y_1^2 \, dy_1$$

$$= \frac{c}{3} y_1^3 \Big|_0^1 = c/3 = 1 \implies c = 3,$$

to want.

How can we find probabilities?

$$P\left(0 \le Y_1 \le 1/2,\ Y_2 \ge 1/4\right) = ?$$



$E = \{(y_1, y_2) \in \mathbb{R}^2 ;\ 1/4 \le y_1 \le 1/2,$
$1/4 \le y_2 \le y_1\}$

$$P\left(0 \le Y_1 \le 1/2,\ Y_2 \ge 1/4\right)$$

$$= \iint_E 3y_1 \, dA = \int_{1/4}^{1/2} \left(\int_{1/4}^{y_1} 3y_1 \, dy_2 \right) dy_1$$

$$= \int_{1/4}^{1/2} \left(3y_1 y_2 \Big|_{1/4}^{y_1} \right) dy_1 = \int_{1/4}^{1/2} 3y_1 (y_1 - 1/4) \, dy_1$$

$$= y_1^3 - \tfrac{3}{8} y_1^2 \Big|_{1/4}^{1/2} = \frac{5}{128}.$$

Joint Cumulative distributions.

Def Let X, Y by two RVs (cont or discrete).

The joint cumulative distribution (joint cdf).

is

$$F(x,y) = P(X \le x, Y \le y).$$

Discrete

$$F(x,y) = \sum_{u \le x} \sum_{v \le y} p(u,v).$$

Continuous

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,dv\,du.$$

Can relate $F(x,y)$ to $f(x,y)$. by partial derivatives.

$$\boxed{f(x,y) = \frac{\partial^2 F(x,y)}{\partial x\, \partial y}.}$$

Properties

① Non-decreasing. if x or y increase. then $F(x,y)$ can't decrease.

② $F(-\infty,-\infty) = F(x,-\infty) = F(-\infty,y) = 0$

"$F(-\infty) = \lim_{x \to -\infty} F(x)$".

③ $F(\infty,\infty) = 1$.

Marginal Distributions

X, Y also have their own distributions, called marginal distributions.

Discrete X, Y are discrete R.V's.

$$p_X(x) = \sum_y p(x,y) \quad , \quad p_Y(y) = \sum_x p(x,y).$$

$$X \sim p_X(x) \quad , \quad Y \sim p_Y(y).$$

$X \backslash Y$	y_1	y_2	$--$	y_m	$p_X(x)$
x_1	.	.		.	$p_X(x_1)$
x_2	.	.		.	
$\vdots$	.	.		.	
x_n	.	.	.	.	
$p_Y(y)$	$p_Y(y_1)$				1

Continuous Case

Marginal probability density.

X, Y continuous RVs, with joint pdf $f(x,y)$.

$$\begin{cases} f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy. \\ f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \end{cases}$$

Example Let the joint pdf.

$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \le y_2 \le y_1 \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

$\int^{\infty}$

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2.$$

$$= \int_0^{y_1} 3y_1 \, dy_2 = 3y_1 y_2 \Big|_0^{y_1} = 3y_1^2.$$

$$f_{Y_2}(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1$$

$$= \int_{y_2}^{1} 3y_1 \, dy_1 = \frac{3}{2} y_1^2 \Big|_{y_2}^{1}$$

$$= \frac{3}{2}(1 - y_2^2).$$

$$\begin{cases} f_{Y_1}(y_1) = 3y_1^2, & 0 \le y_1 \le 1 \quad (0 \text{ otherwise}) \\ f_{Y_2}(y_2) = \frac{3}{2}(1 - y_2^2), & 0 \le y_2 \le 1 \end{cases}$$

Marginal CDFs.

Can calculate marginal cdf from $F(x,y)$.

$$F_X(x) = P(X \le x) = \lim_{y \to \infty} F(x, y)$$

$$F_Y(y) = P(Y \le y) = \lim_{x \to \infty} F(x,y).$$

Conditional Distributions

Want to understand how the distribution of one RV changes given knowledge of another.

Discrete

Recall

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \neq 0$$

Can take $A = \{X = x\}$, $B = \{Y = y\}$.

$$P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Def X, Y discrete R.V. with joint distribution $p(x,y)$ and marginals $p_X(x)$, $p_Y(y)$.

Then

$$p(x \mid y) = \frac{p(x,y)}{p_Y(y)}\ \neq 0$$

and

$$p(y \mid x) = \frac{p(x,y)}{p_X(x)}\ \neq 0.$$

Continuous Case

$$P(X=x \mid Y=y) = 0 = ?, = \frac{0}{0}$$

Condition CDF.

Can we define.

$$P(X \le x \mid Y=y) = \lim_{h\to 0} P(X \le x \mid y \le Y \le y+h)$$
?

Theorem

Let X, Y have joint pdf $f(x,y)$ and joint cdf $F(x,y)$. then

$$F(x \mid y) = P(X \le x \mid Y=y) = \lim_{h\to\infty} P(X \le x \mid y \le Y \le y+h)$$

$$= \frac{\frac{\partial F}{\partial y}(x,y)}{\ } \qquad \int_{-\infty}^{x} f(u,y)\,du$$

$$\frac{\quad}{f_Y(y)} = \frac{\quad}{f_Y(y)}.$$

Proof

$$P(X \le x \mid y \le Y \le y+h) = \frac{P(X \le x,\ y \le Y \le y+h)}{P(y \le Y \le y+h)}$$

$$= \frac{F(x, y+h) - F(x,y)}{F_Y(y+h) - F_Y(y)} = \frac{(F(x,y+h) - F(x,y))/h}{(F_Y(y+h) - F_X(y))/h}$$

$$\lim_{h \to 0} P(X \le x \mid y \le Y \le y+h) = \frac{\frac{\partial F}{\partial y}(x,y)}{F_Y'(y)}$$

$$F_Y'(y) = f_Y(y). \quad \triangledown$$

$$F(x \mid y) = \frac{\frac{\partial F}{\partial y}(x;y)}{f_X(y)}$$

Now we can define the conditional density.

Conditional density.

Let X, Y be two continuous RV with joint pdf $f(x,y)$ and joint cdf $F(x,y)$.

Then the conditional density is

$$f(x|y) = \frac{\partial F(x|y)}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\frac{\partial F}{\partial y}(x,y)}{f_Y(y)} \right)$$

$$= \frac{\frac{\partial^2 F(x,y)}{\partial x \partial y}}{f_Y(y)} = \frac{f(x,y)}{f_Y(y)}.$$

$$\boxed{f(x|y) = \frac{f(x,y)}{f_Y(y)}.}$$

Ex Suppose

$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \le y_2 \le y_1 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $f(y_1|y_2)$?

$$f_{Y_2}(y_2) = \frac{3}{2}(1-y_2^2) \quad 0 \le y_2 \le 1$$

Therefore

$$f(y_1 | y_2) = \begin{cases} \dfrac{2y_1}{(1-y_2^2)} & 0 \le y_2 \le y_1 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

<u>Independence of Random Variables</u>

<u>Recall</u> X, Y are independent if

$$P(X \le x, Y \le y) = P(X \le x) P(Y \le y).$$

<u>Def</u> Let X, Y be two R.V. with joint cdf $F(x,y)$. Then they are independent if and only if

$$F(x,y) = F_X(x) F_Y(y).$$

Discrete Case

This simplifies to.

$$p(x,y) = p_X(x)\, p_Y(y).$$

for all x, y.

Continuous Case

This simplifies to.

$$f(x,y) = f_X(x)\, f_Y(y).$$

Note: If there is one (x,y) which violates the condition, then X, Y are not independent (dependent).