Joint Distribution

- How can we model two rv's using probability models? For example, if we are interested in both weight and height.

- Is it enough if we just use a normal model for weight and another normal model for height?

- We need to introduce joint probability distribution in order to model multiple rv's.

Joint PMF

- Let X and Y be two discrete rv's defined on the sample space. The joint probability mass function $p(x, y)$ is defined for each pair of numbers (x, y) by

$$p(x, y) = P(X=x, Y=y).$$

- As in the single rv case, we must have $p(x, y) \geq 0$ and $\sum_x \sum_y p(x, y) = 1$.

Example

Ex. We randomly put two different balls into 3 numbered (numbered as {1,2,3}) boxes. Let X be the number of empty boxes left; let Y be the minimum of the box number that has balls in it. What is the joint distribution of (X, Y)?

X can take values from {1, 2};

Y can take values from {1, 2, 3};

It's not hard to see we have the following (why?):

$p(2, j) = P(X=2, Y=j) = 1/9$, for $j = 1, 2, 3$.

$p(1, 3) = P(X=1, Y=3) = 0$.

$p(1, 1) = P(X=1, Y=1) = 4/9$.

$p(1, 2) = P(X=1, Y=2) = 2/9$.

p_{ij}	1	2	3
1	$4/9$	$2/9$	0
2	$1/9$	$1/9$	$1/9$

Marginal PMF

- The marginal probability mass functions of X and Y, denoted by $p_X(x)$ and $p_Y(y)$, respectively, are given by

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

Ex.

p_{ij}	1	2	3	$p_X(x)$
1	4/9	2/9	0	2/3
2	1/9	1/9	1/9	1/3
$p_Y(y)$	5/9	1/3	1/9	

- Notice that the marginal probability mass functions are automatically proper pmf's. (why?)

Two continuous rv's

- We would like to extend the same ideas to the continuous case. Let X and Y be continuous rv's. A joint probability density function $f(x, y)$ for these two variables is a function satisfying $f(x, y) \geq 0$ and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

- The marginal probability density function of X and Y, denoted by $f_X(x)$ and $f_Y(y)$, respectively, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{for} \quad -\infty < x < \infty$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad \text{for} \quad -\infty < y < \infty$$

Remarks

- In the continuous case, roughly speaking, $f(x, y)\mathrm{d}x\mathrm{d}y$ can be treated as $\mathrm{P}(\mathrm{X}=x,\mathrm{Y}=y)$.

- $\mathrm{P}(a < \mathrm{X} < b, c < \mathrm{Y} < d) = \int_a^b \int_c^d f(x, y)dxdy$

- As in the discrete case, $f_\mathrm{X}(x)$ and $f_\mathrm{Y}(y)$ calculated from the joint distribution are automatically proper pdf's.

- Marginal distributions are, in fact, the distributions of the marginal random variables when they are treated as univariate random variables.

Example

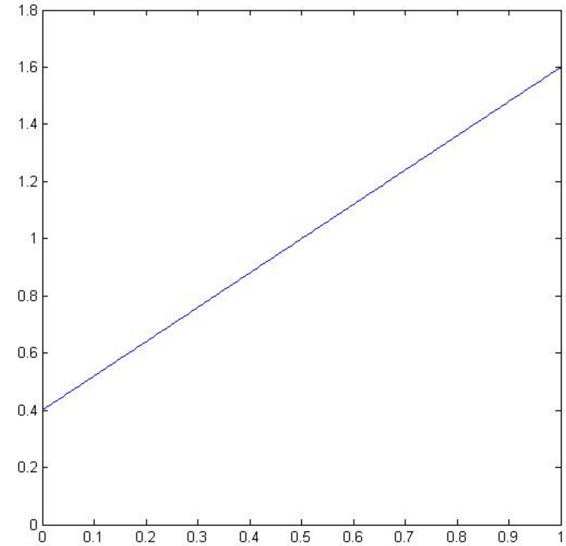
<u>Ex.</u> Suppose the joint pdf of the pair (X, Y) is given by

$$f(x,y) = \begin{cases} \frac{6}{5}(x+y^2) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

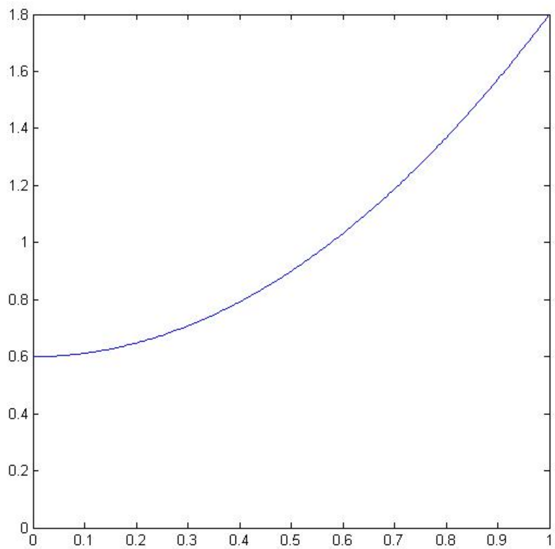
1. Show that this is a proper joint pdf.
2. What is $P(0 \leq X \leq 1/4, 0 \leq Y \leq 1/4)$?
3. What is $P(0 \leq Y \leq 1/4)$

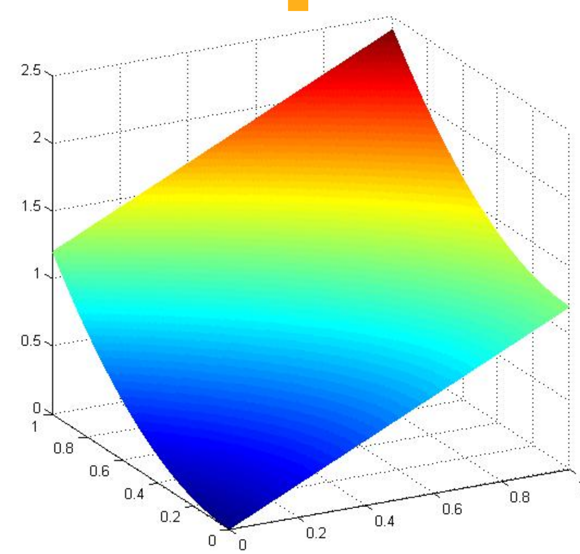
Example cont.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 \frac{6}{5}(x+y^2)dy = \frac{6}{5}x + \frac{2}{5}$$



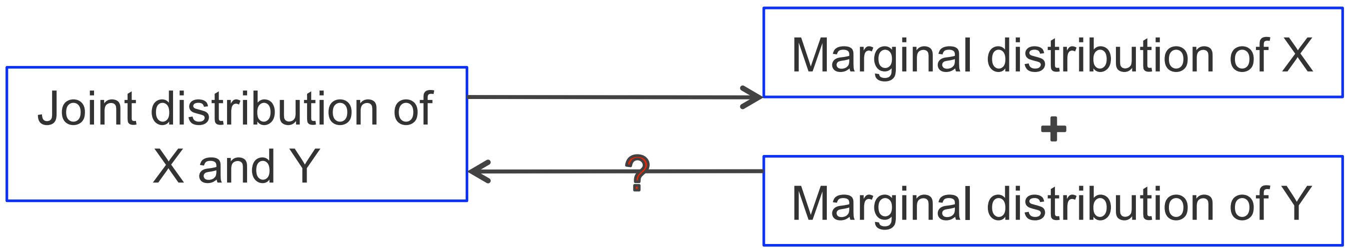
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 \frac{6}{5}(x+y^2)dx = \frac{6}{5}y^2 + \frac{3}{5}$$





Joint and Marginal

- Now we have



- In general, we CANNOT go the other way around. Further information about the dependence structure of X and Y is needed to determine the joint distribution.

Example

Ex. Consider the following two joint distributions of X and Y.

p_{ij}	0	1
0	$3/10$	$3/10$
1	$3/10$	$1/10$

p_{ij}	0	1
0	$9/25$	$6/25$
1	$6/25$	$4/25$

It is easy to see that the marginal distributions of X and Y are the same in both cases. $P(X=0) = P(Y=0) = 3/5$; $P(X=1) = P(Y=1) = 2/5$.

This is the example that *different* joint distributions may have the *same* marginal distributions.

Independent rv's

- Recall the definition of independence of two random events A and B.

$$P(A \cap B) = P(A)\,P(B)$$

- We say two random variables X and Y are independent if and only if

$$P(X=x, Y=y) = P(X=x)\,P(Y=y), \text{ for any } x \text{ and } y.$$

- More specifically, two random variables X and Y are said to be independent if for every pair x and y values,

$$p(x, y) = p_X(x)\,p_Y(y), \text{ when X and Y are discrete;}$$

or

$$f(x, y) = f_X(x)\,f_Y(y), \text{ when X and Y are continuous.}$$

Ex. The second case of the previous example.

Multiple Random Variables

- If $X_1, X_2, \ldots, X_n$ are all discrete random variables, the joint pmf of the variables is the function

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

  If the variables are continuous, the joint pdf of $X_1, X_2, \ldots, X_n$ is the function $f(x_1, x_2, \ldots, x_n)$ such that for any n intervals $[a_1, b_1], \ldots, [a_n, b_n]$,

  $$P(a_1 \leq X_1 \leq b_1, \ldots, a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

- What should be the regularity conditions for $p(x_1, x_2, \ldots, x_n)$ and $f(x_1, x_2, \ldots, x_n)$?

- How do get the marginal distributions of $X_1, X_2, \ldots$ by using $p(x_1, x_2, \ldots, x_n)$ and $f(x_1, x_2, \ldots, x_n)$?

Independence

- Proposition:

  The random variables $X_1, X_2, \ldots, X_n$, are said to be independent if for every subset $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$, of the variables (each pair, each triple, and so on), the joint pmf or pdf of the subset is equal to the product of the marginal pmf's or pdf's.

- $$p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$

- $$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

Example

Ex. Two people each arrive independently at the station at some random time between 5:00 am and 6:00 am (arrival time for either person is uniformly distributed). They stay exactly five minutes and then leave. What is the probability they will meet on a given day.

Conditional dist.

- Using the marginal distributions, one can calculate the conditional distribution of one rv given the other.

- Let X and Y be two conditional rv's with joint pdf $f(x, y)$ and marginal X pdf $f_X(x)$. Then for any X value x for which $f_X(x)>0$, the conditional probability density function of Y given that X=x is

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

- If X and Y are discrete, replace pdf's by pmf's in this definition gives the conditional probability mass function of Y when X=x.

Example

Ex. Let (X, Y) have the joint density

$$f(x, y) = 24y(1 - x - y),\ x,\ y \geq 0,\ x+y < 1.$$

1. What is the conditional density of X given Y=1/2?
2. What is the conditional density of Y given X=1/2?

Example

<u>Ex.</u> For some $\lambda>0$, random variable X has the density function

$$f(x) = \lambda^2 x e^{-\lambda x}, \; x > 0,$$

and given X, Y is a uniform random variable on the interval [0, X].

1. What is the joint distribution of X and Y?
2. What is the distribution of Y?

Expectation of Functions

- Recall how we compute E[*h*(X)]. A similar result also holds for a function *h*(X, Y) of two jointly distributed rv's.

- Let X and Y be jointly distributed rv's with pmf *p*(*x*, *y*), if they are discrete; or pdf *f* (*x*, *y*), if they are continuous. The expected value of a function *h*(X, Y), denoted by E[*h*(X, Y)] is given by

$$\mathrm{E}[h(\mathrm{X},\mathrm{Y})] = \begin{cases} \sum_x \sum_y h(x,y)\cdot p(x,y) & \text{if X and Y are discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\cdot f(x,y)dxdy & \text{if X and Y are continuous} \end{cases}$$

- This result can also be extended to multiple (>2) rv case.

Examples

Ex. (Important! Linearity of expectations) Show that for any two random variables X and Y, E(X+Y) = E(X) + E(Y).

Example

Ex. If two random variables X and Y are independent, what is E(XY)? What about E (g(X)h(Y))?

Covariance

- When two random variables X and Y are not independent, it is often of interest to assess how strongly they are related to one another.

- A popular measurement to characterize the dependence of two rv's is called correlation. To calculate correlation of two rv's, we'll have calculate the covariance of the two rv's.

- The covariance between two rv's X and Y is

$$\begin{aligned}\mathrm{Cov}(\mathrm{X}, \mathrm{Y}) &= \mathrm{E}[(\mathrm{X} - \mu_X)(\mathrm{Y} - \mu_Y)] \\ &= \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y) \cdot p(x, y) & \mathrm{X},\ \mathrm{Y} \text{ discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) \cdot f(x, y)\, dx dy & \mathrm{X},\ \mathrm{Y} \text{ continuous} \end{cases}\end{aligned}$$

Short cut

- Proposition:

$$\boxed{\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)}$$

- What happens if we set Y=X?

Example

Ex. Suppose the joint distribution of X and Y are

$$f(x, y) = \begin{cases} 24xy & 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the covariance of X and Y?

$$f_X(x) = \int_y f(x, y)dy = \int_0^{1-x} 24xydy = 12x(1-x)^2$$

$$f_Y(y) = 12y(1-y)^2$$

$$\mathrm{E(X)} = \int_0^1 x \cdot 12x(1-x)^2 dx = \frac{2}{5} = \mathrm{E(Y)}$$

$$\mathrm{E(XY)} = \int \int_{x,y} xyf(x, y)dxdy = \int_0^1 \int_0^{1-y} 24x^2y^2 dxdy = \frac{2}{15}$$

$$\mathrm{Cov(X, Y)} = \mathrm{E(XY)} - \mathrm{E(X)E(Y)} = \frac{2}{15} - \left(\frac{2}{5}\right)^2 = -\frac{2}{75}$$

Correlation

- The correlation coefficient of X and Y, denoted by Corr(X, Y) or $\rho_{X,Y}$ is defined by

$$\rho_{X,Y} = \frac{\mathrm{Cov}(\mathrm{X}, \mathrm{Y})}{\sigma_X \cdot \sigma_Y}$$

- Because of Cauchy-Schwarz inequality, we have

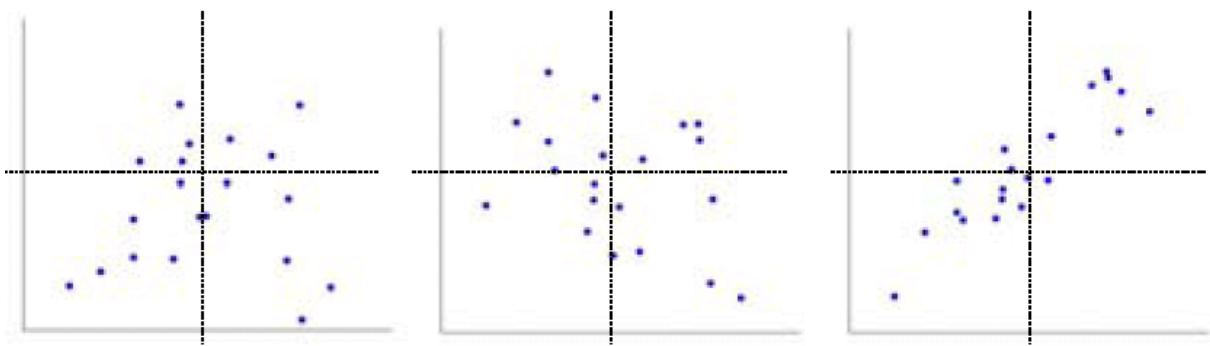
$$\mathrm{Cov}^2(\mathrm{X}, \mathrm{Y}) \leq \mathrm{Var}(\mathrm{X})\mathrm{Var}(\mathrm{Y}) \Longrightarrow |\rho_{X,Y}| \leq 1$$

- The correlation coefficient $\rho_{X,Y}$ is NOT a completely general measure of the strength of a relationship. $\rho_{X,Y}$ is actually a measure of the degree of *linear* relationship between X and Y.

Remarks

- If X and Y are independent, then $\rho_{X,Y} = 0$ (why?). But $\rho_{X,Y} = 0$ does NOT imply independence.

- $\rho_{X,Y} = 1$ or -1 iff $Y = aX+b$ for some numbers a and b with $a \neq 0$.

- 

$r \approx 0$ $r \approx -0.3$ $r \approx 0.9$