Math 425
Intro to Probability
Lecture 25

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

March 16, 2009

Jointly distributed random variables

☞ In many experiments we have several random variables that we are interested at the same time.

- A study records the age, weight, blood pressure, and cholesterol levels. Each of these items is a different random variable over the sample space of people in study. Are some of the random variables dependent on others? (Up to now we only have a notion of independence of events.)
- Two normally distributed variables X and Y are normally distributed with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2. What does the distribution of their sum $X + Y$ look like?

Example

Example. Choose a point $P = (x, y)$ "at random" in the unit square $[0, 1] \times [0, 1]$. We can think of P as a random variable with values in the plane $\mathbb{R}^2$.

Let $X = x$ and $Y = y$ be the random variables on $\mathbb{R}$ which give the first and second coordinates of P. What is the probability $\mathbf{P}\{X < Y\}$?

Solution. It seems reasonable to take "at random" to mean that for any subset $A \subseteq [0, 1] \times [0, 1]$, the probability

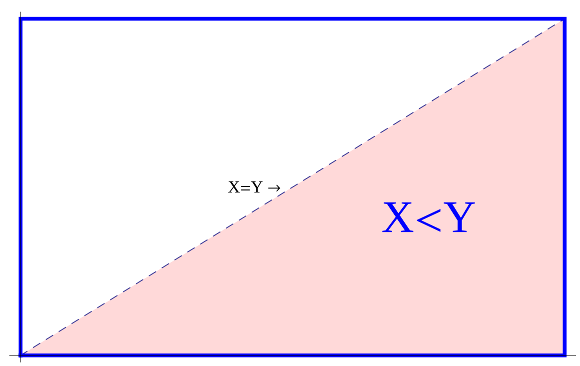
$$\mathbf{P}\{P \in A\} = \frac{\text{area of } A}{\text{area of } [0,1] \times [0,1]} = \text{area of } A\,.$$

This is analogous to choosing a point in an interval of $\mathbb{R}$ "at random".

Example

☞ The area of the region $\{X < Y\}$ is half that of the unit square. So,

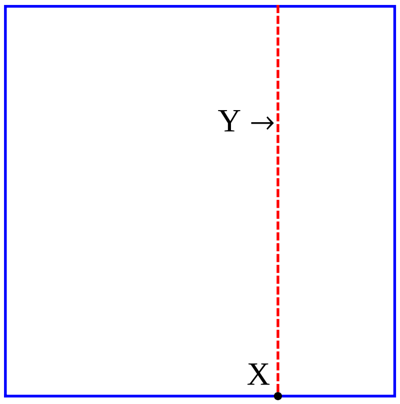
$$\mathbf{P}\{X < Y\} = \frac{1}{2}.$$



Example – continued

☞ Suppose $P = (X, Y)$ is chosen "at random" in the unit square? Are X and Y unformly distributed when (X, Y) is? Are X and Y independent?

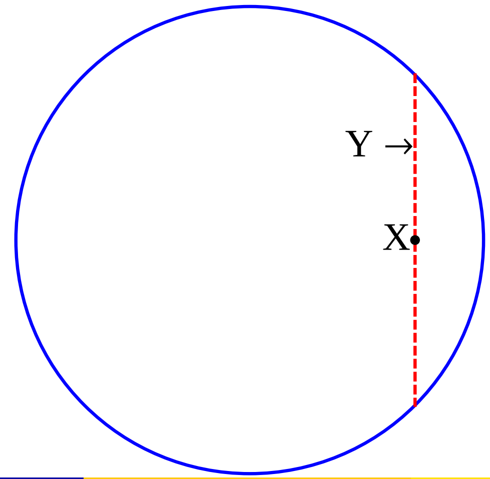
Are X and Y independent? Are they uniformly distributed on the line $[0, 1]$? That is, would you get the same distribution for P in the square if you choose X and Y uniformly at random and let $P = (X, Y)$?



Example – continued

☞ Suppose $P = (X, Y)$ is chosen "at random" in the unit circle?

Are X and Y independent? In the picture I have chosen X first; how are the values of Y constrained?



.

Joint Cumulative Distribution

☞ The definition of joint cumulative distribution is "abstract" – it does not yet tell you HOW to compute the probabilities.

Definition

Let X and Y be two random variables over the same sample space.

The joint cumulative distribution function on X and Y is defined for $-\infty \leq a, b \leq \infty$

$$F(a, b) = \mathbf{P}\{X \leq a, Y \leq b\} = \mathbf{P}\left(\{X \leq a\} \cap \{Y \leq b\}\right)$$

If $X_1, X_2, \ldots, X_n$ are random variables over the same sample space, their joint cumulative distribution function is defined for $-\infty \leq a_1, a_2, \ldots, a_n \leq \infty$

$$F(a_1, a_2, \ldots, a_n) = \mathbf{P}\{X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n\}.$$

Marginal Probabilities

Theorem

Let X and Y be random variables over the same sample space with cumulative distributions F_X and F_Y.

The joint distribution F is related to F_X and F_Y by

$$\begin{aligned} F_X(a) &= F(a, \infty) = \mathbf{P}\{X \le a, Y \le \infty\} \qquad -\infty \le a \le \infty \\ F_Y(b) &= F(\infty, b) = \mathbf{P}\{X \le \infty, Y \le b\} \qquad -\infty \le b \le \infty \end{aligned}$$

for all

The distribution F_X and F_Y are called marginal distributions.

The reason for the name is explained in Ross, Example 6.1.a.

Proof

Proof.

☞ To compute $\mathbf{P}\{X \le a\}$, the key is to observe that

$$\{Y < \infty\} = S \quad \text{so} \quad \{X \le a, Y < \infty\} = \{X \le a\} \cap S = \{X \le a\}.$$

Thus,

$$\mathbf{P}\{X \le a\} = \mathbf{P}\{X \le a, Y \le \infty\}.$$

☞ Similarly, to compute $\mathbf{P}\{Y \le b\}$, the key is to observe that

$$\{X < \infty\} = S \quad \text{so} \quad \{X \le \infty, Y < b\} = \{Y \le b\} \cap S = \{Y \le b\}.$$

Thus,

$$\mathbf{P}\{Y \le b\} = \mathbf{P}\{X \le \infty, Y \le b\}.$$

□

Basic Events

☞ All joint basic events for random variables X and Y can be computed using the joint cumulative distribution F for X and Y.

Theorem

Let X and Y be random variables over the same sample space.

For any $-\infty \le a, b, c, d \le \infty$ with $a \le b$ and $c \le d$

$$\mathbf{P}\{a < X \le b,\ c < Y \le d\} = F(b, d) + F(a, c) - F(a, d) - F(b, c)$$

Examples

Examples. To apply the previous theorem, use the identities:

$$\begin{gathered} F(a, \infty) = F_X(a) \qquad F(\infty, b) = F_Y(b) \\ F(\infty, \infty) = 1 \qquad F(a, -\infty) = F(-\infty, b) = 0. \end{gathered}$$

X and Y are random variables from the same sample space.

$$\begin{aligned} \mathbf{P}\{X > a, Y > b\} &= \mathbf{P}\{a < X < \infty,\ b < Y < \infty\} \\ &= F(a, b) + F(\infty, \infty) - F(a, \infty) - F(b, \infty) \\ &= 1 + F(a, b) - F_X(a) - F_Y(b) \end{aligned}$$

$$\begin{aligned} \mathbf{P}\{X \le a, Y > b\} &= \mathbf{P}\{-\infty < X \le a,\ b < Y < \infty\} \\ &= F(a, \infty) + F(-\infty, b) - F(a, b) - F(-\infty, \infty) \\ &= F_X(a) - F(a, b). \end{aligned}$$

Joint probability mass functions

Definition

Suppose that each of X and Y are discrete random variables from the same sample space.

The joint probability mass function is defined for real a and b by

$$p(a,b) = \mathbf{P}\{X = a,\ Y = b\}.$$

The definition extends naturally to any number of variables.

Properties Joint Density Function

☞ The joint probability mass function for X and Y satisfies the following properties.

$$(a) \quad p(a,b) \geq 0 \qquad \text{for all reals } a \text{ and } b,$$

$$(b) \quad 1 = \sum_{a,b:p(a,b)>0} p(a,b)$$

$$(c) \quad \mathbf{P}\{X \in A,\ Y \in B\} = \sum_{a \in A} \sum_{b \in B} p(a,b) \qquad \text{when } A, B \subseteq \mathbb{R}$$

The last property requires that A and B be integrable.

Marginal probability mass functions

☞ The individual probability mass functions p_X and p_Y for X and Y can be obtained from their joint probability mass function p by

$$p_X(a) = \sum_{y:p(a,y)>0} p(a,y)$$

$$p_Y(b) = \sum_{x:p(x,b)>0} p(x,b)$$

p_X and p_Y are also called the marginal distributions of X and Y.

Example

Example

A pair of dice bear the numbers 1, 2, 3 twice each. Let X and Y denote the face value of each die.

Let $U = X + Y$ and $V = X - Y$. So,

$$2 \leq V \leq 6 \qquad -2 \leq V \leq 2.$$

Calculate the joint probability mass function of U and V.

☞ The joint probability mass function of X and Y is easy:

$$p(j,k) = \frac{1}{9} \qquad 1 \leq j,k \leq 3.$$

Example – continued

☞ Joint mass function p for U and V. Note the the marginal p_U is obtained by summing the probabilities across the row, and the marginal p_V is obtained by summing the probabilities down the column.

		V					
		-2	1	0	1	2	
	6			$\frac{1}{9}$			1/9
	5		$\frac{1}{9}$		$\frac{1}{9}$		2/9
U	4	$\frac{1}{9}$		$\frac{1}{9}$		$\frac{1}{9}$	3/9 p_U
	3		$\frac{1}{9}$		$\frac{1}{9}$		2/9
	2			$\frac{1}{9}$			1/9
	p_V	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	

Example: Bernoulli trials

Example

In n Bernoulli trials with probability of success p, let X count the number of successes and Y the number of failures.

What is the joint probability mass function?

☞ For $0 \le j, k \le n$

$$p(j,k) = \begin{cases} \frac{n!}{j!k!} p^j (1-p)^k & \text{if } j+k=n \\ 0 & \text{otherwise} \end{cases}$$

The marginals are determined by

$$p_X(j) = p(j, n-j) \qquad p_Y(k) = p(n-k, k) \qquad 0 \le j, k \le n.$$

$p(j,k) = 0$ when $k \ne n-j$ and when $j \ne n-k$.

Example: DeMoivre trials

Example

A DeMoivre trial has three possible outcomes. Let X, Y, Z count the three possible outcomes in n independent DeMoivre trials with probabilities p_1, p_3 and p_3.

What is the joint probability mass function?

☞ For $0 \le i, j, k$

$$p(i,j,k) = \begin{cases} \frac{n!}{i!j!k!} p_1^i p_2^j p_3^k & \text{if } i+j+k=n \\ 0 & \text{otherwise} \end{cases}$$

The marginals are determined by

$$p_X(i) = \sum_{j,k:n-j-k=i} p(i,j,k) \qquad p_Y(j) = \sum_{i,k:n-i-k=j} p(i,j,k)$$

$$p_Z(k) = \sum_{i,j:n-i-j=k} p(i,j,k)$$

Example

☞ The marginals do not necessarily determine the joint distribution.

(A) Flip a coin; let X count heads and Y count tails.

$$p_X(0) = p_X(1) = p_Y(0) = p_Y(1) = \frac{1}{2}$$

with

$$p(0,1) = p(1,0) = \frac{1}{2} \qquad p(0,0) = p(1,1) = 0.$$

(B) Flip two coins; let X count heads on the first coin and Y count heads on the second coin.

$$p_X(0) = p_X(1) = p_Y(0) = p_Y(1) = \frac{1}{2}$$

with

$$p(0,0) = p(1,0) = p(0,1) = p(1,1) = \frac{1}{4}$$

☞ (A) and (B) have the same marginals, but different joint distributions.

Joint Density Function

Definition

Random variables X and Y over the same sample space are jointly continuous with joint density function $f(x,y)$ if
for all $-\infty < a < b < \infty, \ -\infty < c < d < \infty$

$$\mathbf{P}\{a \leq X \leq b, \ c \leq Y \leq d\} = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

The definition generalizes to n random variables using n integrals, see Ross p. 266.

Properties Joint Density Function

☞ The joint density function for X and Y satisfies the following properties.

(a) $f(x,y) \geq 0$ for all reals x and y,

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

(c) $\mathbf{P}\{X \in A, \ Y \in B\} = \int_B \int_A f(x,y)\,dx\,dy$ when $A, B \subseteq \mathbb{R}$

The last property requires that A and B be integrable.

Example

Example. Let X and Y have joint density

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

Determine c?

☞

$$\begin{aligned} 1 = \int_0^1 \int_x^1 cxy\,dy\,dx &= \int_0^1 cx \int_x^1 y\,dy\,dx \\ &= \int_0^1 \frac{cx}{2}(1-x^2)\,dx = \frac{c}{8} \end{aligned}$$

So, $c = 8$ is required for $f(x,y)$ to be a joint density function.

Joint Density and Distribution

☞ Let X and Y be random variables over the same sample space.
The joint cumulative distribution F for X and Y is determined by the joint density f by

$$F(a,b) = \mathbf{P}\{X \leq a, \ Y \leq b\} = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)\,dx\,dy$$

☞ The joint density f for X and Y is determined by the joint cumulative distribution F by

$$f(a,b) = \frac{\partial^2}{\partial a \partial b} F(a,b)$$

whenever the partials exist.

Marginals

- If X and Y are jointly continuous, they are individually continuous. Let $f(x, y)$ be a joint density for X and Y. Then

$$\begin{aligned} \mathbf{P}\{X \le a\} &= \mathbf{P}\{X \le a,\ Y < \infty\} \\ &= \int_{-\infty}^{a} \int_{-\infty}^{\infty} f(x, y)\, dy\, dx \\ &= \int_{-\infty}^{a} f_X(x)\, dx \end{aligned}$$

where $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$.

- Similarly,

$$\mathbf{P}\{Y \le b\} = \int_{-\infty}^{b} f_Y(y)\, dy$$

where $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$.

Example

Example. Let X and Y have joint density

$$f(x, y) = c(x + y), \qquad 0 \le x, y \le 1.$$

- Determine c.

$$1 = \int_0^1 \int_0^1 c(x + y)\, dx\, dy = \int_0^1 cx\, dx + \int_0^1 cy\, dy = c$$

- Determine the joint distribution F for X and Y.

$$\begin{aligned} 0 \le a, b \le 1: &\quad F(a, b) = \int_0^b \int_0^a x + y\, dx\, dy = \frac{1}{2} ab(a + b) \\ 0 \le a \le 1,\ b > 1: &\quad F(a, b) = \int_0^1 \int_0^a x + y\, dx\, dy = \frac{1}{2} a(a + 1) \\ a > 1,\ 0 \le b \le 1 &\quad F(a, b) = \int_0^b \int_0^1 x + y\, dx\, dy = \frac{1}{2} b(1 + b) \\ a, b > 1, &\quad F(a, b) = 1 \end{aligned}$$

Example – continued

- Determine the marginal densities $f_X(x)$ and $f_Y(y)$.

$$\begin{aligned} f_X(x) &= \int_0^1 x + y\, dy = \frac{1}{2} + x \\ f_Y(y) &= \int_0^1 x + y\, dx = \frac{1}{2} + y \end{aligned}$$

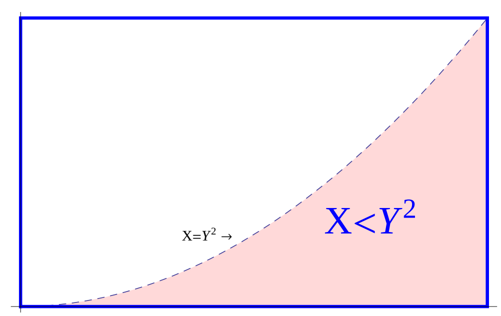
- Determine the marginal distributions $F_X(a)$ and $F_Y(b)$.

$$\begin{aligned} F_X(a) &= \int_0^a \frac{1}{2} + x\, dx = \frac{a^2 + a}{2} \\ F_Y(b) &= \int_0^1 \frac{1}{2} + y\, dy = \frac{b^2 + b}{2} \end{aligned}$$

Example – continued

$$\begin{aligned} \mathbf{P}\{X < Y^2\} &= \int_0^1 \int_0^{y^2} x + y\, dx\, dy \\ &= \int_0^1 \frac{1}{2} y^4 + y^3\, dy = \frac{7}{20} \end{aligned}$$

- Region of integration: $\{X < Y^2\}$:



Example

Example. Let (X, Y) have a joint density given by

$$f(x,y) = \lambda\mu e^{-\lambda x - \mu y}, \qquad 0 \le x, y < \infty.$$

☞ Determine the marginal densities $f_X(x)$ and $f_Y(y)$.

$$\begin{aligned} f_X(x) &= \int_0^\infty \lambda\mu e^{-\lambda x-\mu y}\,dy \\ &= \lambda e^{-\lambda x}\int_0^\infty \mu e^{-\mu y}\,dy = \lambda e^{-\lambda x} \\ f_Y(y) &= \int_0^\infty \lambda\mu e^{-\lambda x-\mu y}\,dx = \mu e^{-\mu y}. \end{aligned}$$

Example – continued

☞ (X, Y) has a joint density given by

$$f(x,y) = \lambda\mu e^{-\lambda x - \mu y}, \qquad 0 \le x, y < \infty.$$

☞ What is the probability that $X < Y$?

$$\begin{aligned} \mathbf{P}\{X<Y\} &= \int_0^\infty\int_0^y \lambda\mu e^{-\lambda x-\mu y}\,dx\,dy \\ &= \int_0^\infty \mu e^{-\mu y}\int_0^y \lambda e^{-\lambda x}\,dx\,dy \\ &= \int_0^\infty \mu e^{-\mu y} - \mu e^{-(\mu+\lambda)y}\,dy \\ &= 1 - \frac{\mu}{\mu+\lambda} \\ &= \frac{\lambda}{\mu+\lambda} = \frac{E[X]}{E[Y]+E[X]} \end{aligned}$$

Example – continued

☞ What is the distribution and density function for $\frac{X}{Y}$?

$$\begin{aligned} F_{X/Y}(a) &= \mathbf{P}\{\frac{X}{Y} \le a\} = \mathbf{P}\{X \le aY\} \\ &= \int_0^\infty\int_0^{ay} \lambda\mu e^{-\lambda x-\mu y}\,dx\,dy \\ &= \int_0^\infty \mu e^{-\mu y}\int_0^{ay} \lambda e^{-\lambda x}\,dx\,dy \\ &= \int_0^\infty \mu e^{-\mu y} - \mu e^{-(\mu+\lambda a)y}\,dy \\ &= -e^{-y} + \frac{\mu e^{-(\mu+\lambda a)y}}{\mu+\lambda a}\Big|_0^\infty \\ &= 1 - \frac{\mu}{\mu+\lambda a} \\ f_{X/Y}(a) &= \frac{d}{da}F_{X/Y}(a) = \frac{\mu\lambda}{(\mu+\lambda a)^2} \end{aligned}$$

when $0 < a < \infty$.

Example – continued

☞ Region of integration: $\{X < aY\}$ (here, $a = 2$).
Note that the distribution $F_{X/Y}(a)$ is the volume above this region and below the joint density curve $f(x,y) = \lambda\mu e^{-\lambda x-\mu y}$.