Pairs of Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

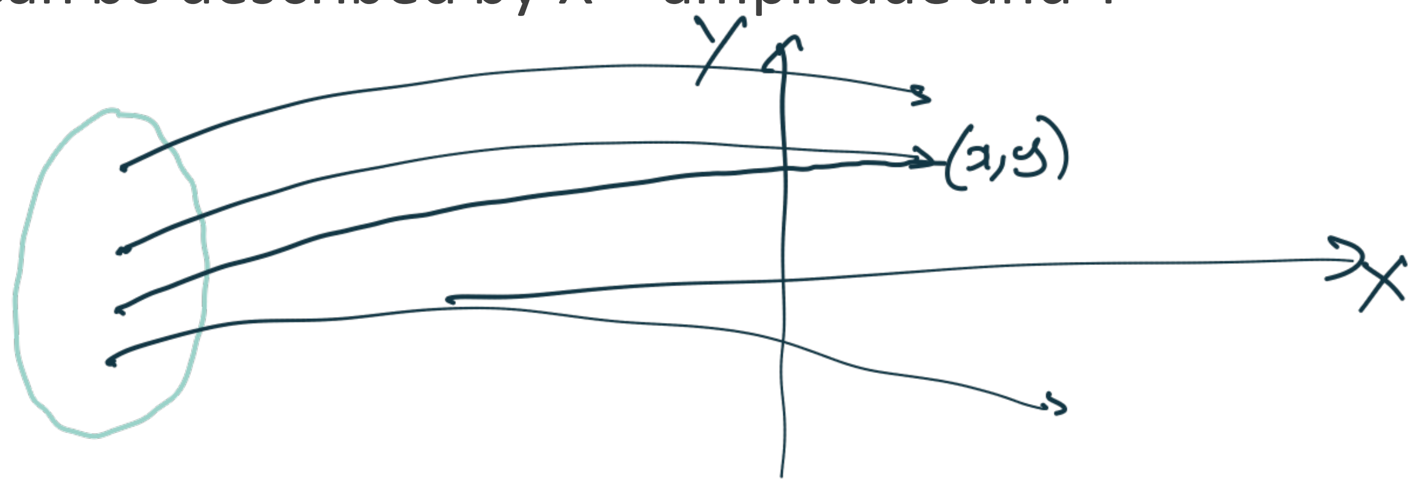
We started with…

- An Experiment that contains a
  - Procedure, Observations, and Model
- Later we mapped outcomes to numbers
  - One or more outcomes to a point on the line
- We will now map an outcome to a pair of numbers
  - One or more outcomes to a point on a 2-D plane

We started with…

- The numbers in the pair correspond to RVs X and Y

- For example, a transmitted sinusoid that is received with a random amplitude and random phase
  - Outcomes can be described by X = amplitude and Y = phase



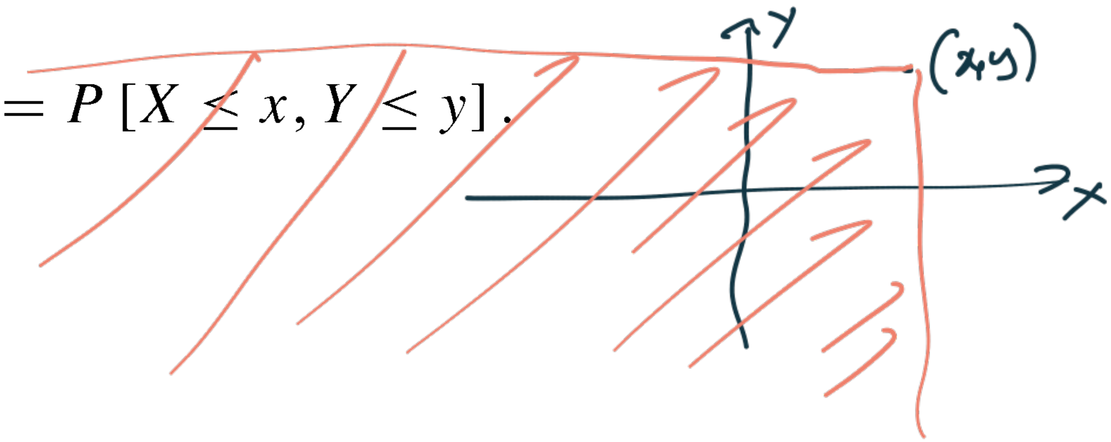
Pair of RVs

- We defined a CDF for a single RV X
- For the pair of RVs we define a ***joint*** CDF

Definition 4.1 Joint Cumulative Distribution Function (CDF)

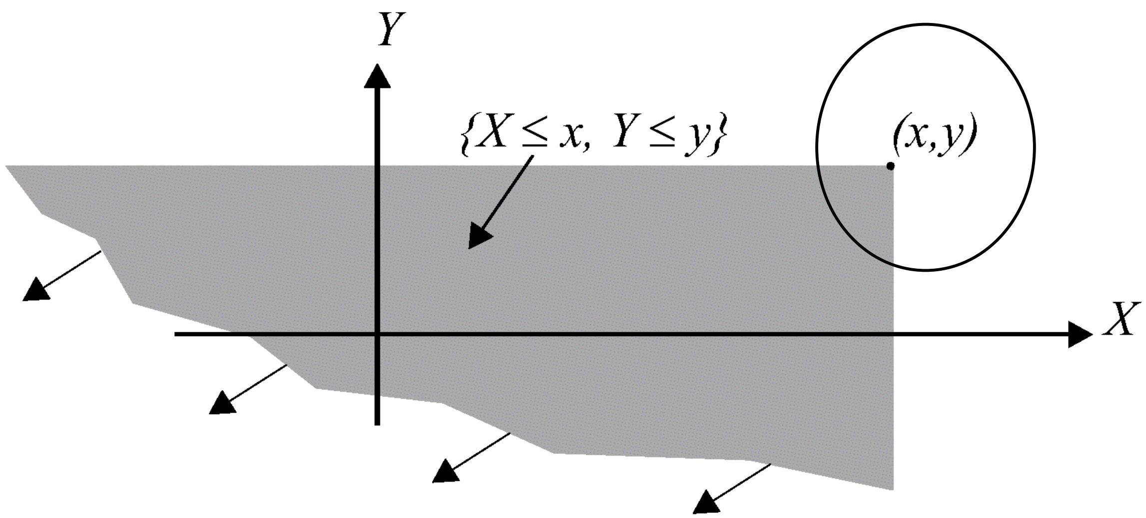
The joint cumulative distribution function *of random variables* X *and* Y *is*

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y].$$



Joint CDF

Figure 4.1



The area of the (X, Y) plane corresponding to the joint cumulative distribution function $F_{X,Y}(x, y)$.

- We are interested in the probability of the intersection of the events {X<=x} and {Y<=y}

Theorem 4.1

For any pair of random variables, X, Y,

(a) $\square \le F_{X,Y}(x, y) \le \square$

(b) $F_X(x) = \square$,

(c) $F_Y(y) = \square$,

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = \square$

(e) If $x \le x_1$ and $y \le y_1$, then $F_{X,Y}(x, y) \; \square \; F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = \square$.

(g) $F_{X,Y}$ (∞, - ∞) = ?

Theorem 4.1

For any pair of random variables, X, Y,

(a) $0 \leq F_{X,Y}(x, y) \leq 1$,

(b) $F_X(x) = F_{X,Y}(x, \infty)$,

(c) $F_Y(y) = F_{X,Y}(\infty, y)$,

(d) $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$,

(e) If $x \leq x_1$ and $y \leq y_1$, then $F_{X,Y}(x, y) \leq F_{X,Y}(x_1, y_1)$,

(f) $F_{X,Y}(\infty, \infty) = 1$.

(g) $F_{X,Y}(\infty, -\infty) = 0$.

Joint PMF

Joint Probability Mass Function

Definition 4.2 *(PMF)*

The joint probability mass function *of discrete random variables* X *and* Y *is*

$$P_{X,Y}(x, y) = P[X = x, Y = y].$$

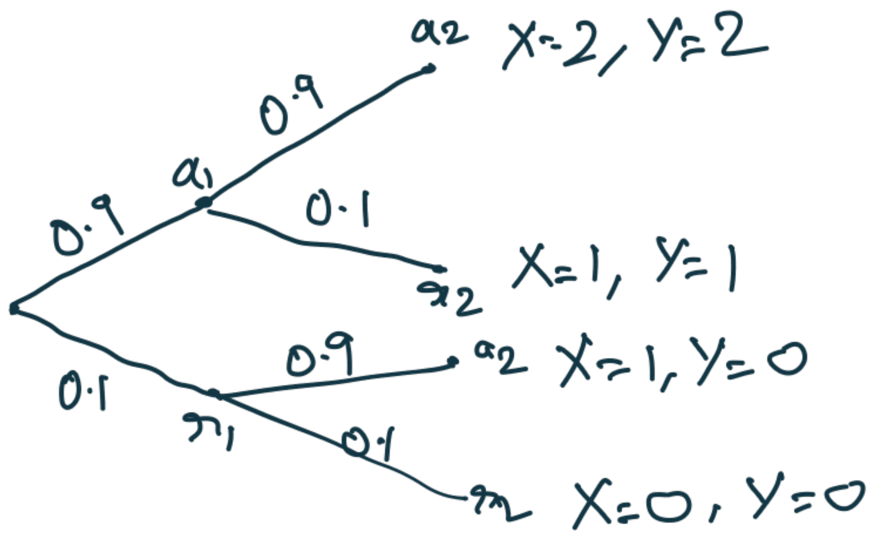
Example of a Joint PMF

Example 4.1 Problem

Test two integrated circuits one after the other. On each test, the possible outcomes are a (accept) and r (reject). Assume that all circuits are acceptable with probability 0.9 and that the outcomes of successive tests are independent. Count the number of acceptable circuits X and count the number of successful tests Y before you observe the first reject. (If both tests are successful, let $Y = 2$.) Draw a tree diagram for the experiment and find the joint PMF of X and Y.

$S_X = \{0,1,2\}$

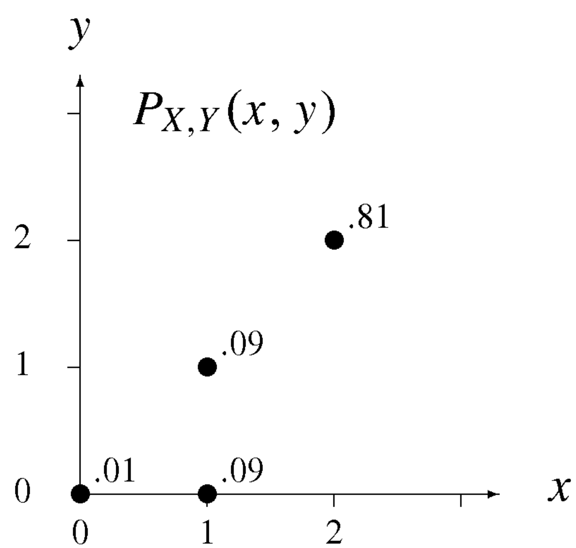
$S_Y = \{0,1,2\}$



Example of a Joint PMF

Three possible representations of a Joint PMF

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81



$$P_{X,Y}(x,y) = \begin{cases} 0.81 & x=2, y=2, \\ 0.09 & x=1, y=1, \\ 0.09 & x=1, y=0, \\ 0.01 & x=0, y=0. \\ 0 & \text{otherwise} \end{cases}$$

Joint PMF

$$\sum_{x \in S_X} \sum_{y \in S_Y} P_{X,Y}(x, y) = ?$$

Theorem 4.2

For discrete random variables X and Y and any set B in the X, Y plane, the probability of the event $\{(X, Y) \in B\}$ is

$$P[B] = \sum_{(x,y)\in B} P_{X,Y}(x, y).$$

Quiz 4.2

The joint PMF $P_{Q,G}(q,g)$ for random variables Q and G is given in the following table:

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

(4.12)

Calculate the following probabilities:

(1) $P[Q=0]$

(2) $P[Q=G]$ $= P[Q=0, G=0] + P[Q=1, G=1]$

(3) $P[G>1]$ $= P[Q=0, G>1] + P[Q=1, G>1] = P[G=2] + P[G=3]$

(4) $P[G>Q]$

$\{G>Q\} = \{(0,1), (0,2), (0,3), (1,2), (1,3)\}$

Marginal PMF

- For discrete RVs X and Y with **joint** PMF $P_{X,Y}(x,y)$, $P_X(x)$ and $P_Y(y)$ are defined as the **marginal** PMFs of X and Y respectively

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x,y)$$

Suppose we have the joint $P_{X,Y,Z}(x,y,z) = P[X=x, Y=y, Z=z]$

$$P_X(x) = \sum_{z \in S_Z} \sum_{y \in S_Y} P_{X,Y,Z}(x,y,z)$$

Quiz 4.3

The probability mass function $P_{H,B}(h, b)$ for the two random variables H and B is given in the following table. Find the marginal PMFs $P_H(h)$ and $P_B(b)$.

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$
$h=-1$	0	0.4	0.2
$h=0$	0.1	0	0.1
$h=1$	0.1	0.1	0

(4.20)

Marginal PMF

Theorem 4.3

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$,

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y), \qquad P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y).$$

Marginal PMF

Theorem 4.3

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$,

$$P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x, y), \qquad P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x, y).$$

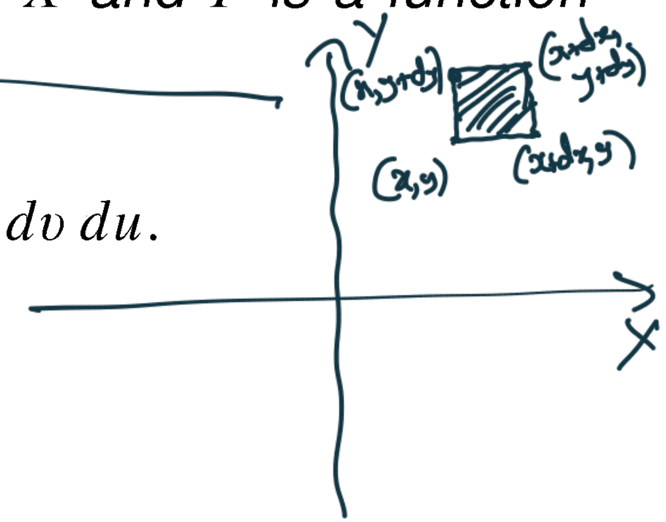
- We looked at the joint CDF and the joint PMF. What do you expect next?

Joint Probability Density

Definition 4.3 Function (PDF)

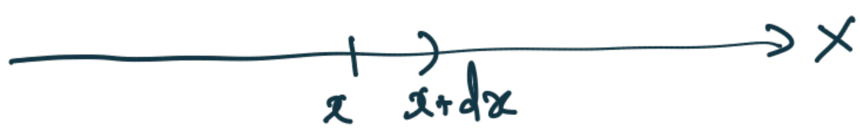
The joint PDF of the continuous random variables X and Y is a function $f_{X,Y}(x, y)$ with the property

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, dv\, du.$$



$P[\, x \le X < x+dx,\ \ y \le Y < y+dy \,]$

$\approx f_{X,Y}(x,y)\, dx\, dy.$



$P[\, x < X \le x+dx \,] \approx f_X(x)\, dx$

Joint PDF

Theorem 4.4

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \, \partial y}$$

$$z = g(x,y) = x^3 + 3x^2y^2 + y^3 + y^4 + 10$$

$$\frac{\partial^2 g(x,y)}{\partial x \partial y} = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y} g(x,y)\right]$$

$$= \frac{\partial}{\partial x}\left[6x^2y + 3y^2 + 4y^3\right]$$

$$= 12xy + 0 + 0$$

$$= 12xy$$

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

$P[X \leq \infty, Y \leq \infty] = 1$

$\Rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$

Properties of a Joint PDF

Theorem 4.6

A joint PDF $f_{X,Y}(x, y)$ has the following properties corresponding to first and second axioms of probability (see Section 1.3):

(a) $f_{X,Y}(x, y) \geq 0$ for all (x, y),

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$.

- Nonnegative
- Area under it is unity

Theorem 4.5

Theorem 4.5

$$P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)$$

Properties of a Joint PDF

- For a point (x,y) in the xy plane, we know that

$$\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)}{\Delta x \Delta y}$$

- How?

$$\frac{\partial}{\partial y} F_{X,Y}(x,y) = \lim_{\Delta y \to 0} \frac{F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x,y)}{\Delta y}$$

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} F_{X,Y}(x,y) = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \left[\frac{\left(\frac{F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y)}{\Delta x} \right)}{\Delta y} - \frac{\frac{F_{X,Y}(x+\Delta x, y) - F_{X,Y}(x,y)}{\Delta x}}{\Delta y} \right.$$

Properties of a Joint PDF

- We can write

$$
\begin{aligned}
& f_{X,Y}(x,y)\Delta x \Delta y \\
& \approx F_{X,Y}(x+\Delta x, y+\Delta y) - F_{X,Y}(x, y+\Delta y) - F_{X,Y}(x+\Delta x, y) + F_{X,Y}(x,y)
\end{aligned}
$$

$$f_{X,Y}(x,y)\Delta x \Delta y \approx P[x < X \leq x+\Delta x, y < Y \leq y+\Delta y]$$

Using a Joint PDF

- Even when the event corresponds to a rectangular area calculating probability of the event using the CDF is nontrivial!
- Using a PDF is usually more convenient

Theorem 4.7

The probability that the continuous random variables (X, Y) are in A is

$$P[A] = \iint_A f_{X,Y}(x, y)\, dx\, dy.$$

- *A* could be the rectangular region we looked at earlier
- *A* can be an area of any shape or size

VVVS Problem

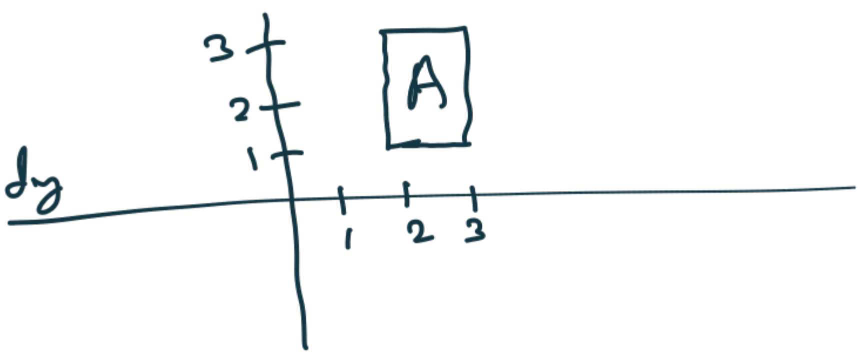
Example 4.4 Problem

Random variables X and Y have joint PDF

$$f_{X,Y}(x, y) = \begin{cases} c & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \qquad (4.22)$$

Find the constant c and $P[A] = P[2 \le X < 3, 1 \le Y < 3]$.

$$P[A] = \int_0^3 \int_0^5 f_{X,Y}(x,y)\,dx\,dy$$

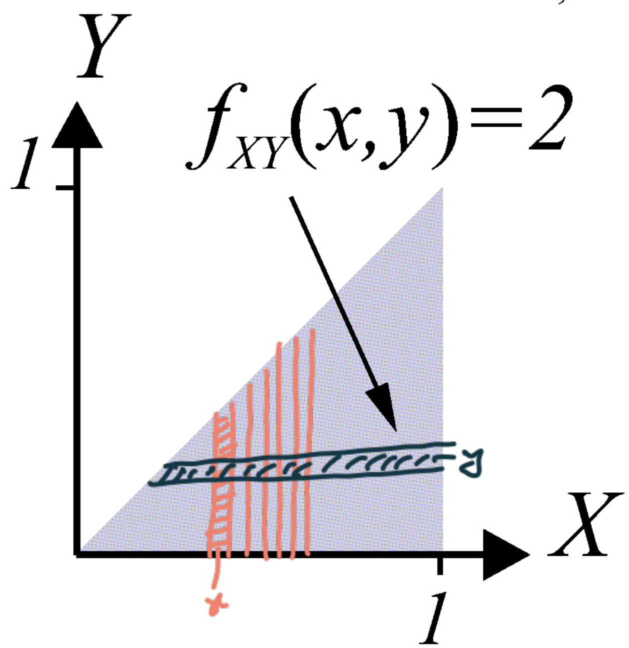


- c=1/15
- P[A] = 2/15
- Note that A is a rectangular area and so the problem is very straightforward

Calculating a CDF from a PDF!

Example 4.5 Problem

Find the joint CDF $F_{X,Y}(x,y)$ when X and Y have joint PDF



$$f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$\iint_{\{0\le y\le x\le 1\}} f_{X,Y}(x,y)\,dx\,dy = \int_0^1 \left[\int_0^x 2\,dy\right] dx$$

- All valid outcomes lie in the triangle

$$\int_0^1 \int_y^1 2\,dx\,dy$$

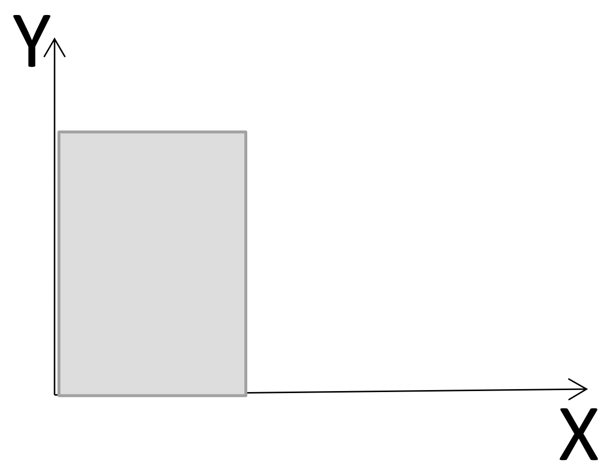
Quiz 4.4

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x,y) = \begin{cases} cxy & 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.33}$$

Find the constant c. What is the probability of the event $A = X^2 + Y^2 \le 1$?

- We want to find the probability that a pair (x,y) lies inside the unit circle
- Use polar coordinates
  - dx dy = r dr dϕ



Marginal PDF

Theorem 4.8

$$f_Z(z) = \iint_{x \in S_x, y \in S_y} f_{X,Y,Z}(x,y,z)\,dy\,dx$$

If X and Y are random variables with joint PDF $f_{X,Y}(x, y)$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx.$$

- Note that P[X <= x] = P[X<=x, Y<∞]
- We have

$$P[X \leq x] = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X,Y}(u, v)\, dv\, du$$

- Taking the derivative with respect to x (and using the fundamental theorem of calculus) we get our desired result

Quiz 4.5

The joint probability density function of random variables X and Y is

$$f_{X,Y}(x, y) = \begin{cases} 6(x + y^2)/5 & 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.40}$$

Find $f_X(x)$ and $f_Y(y)$, the marginal PDFs of X and Y.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

- You must get

$$f_X(x) = \begin{cases} \frac{6x+2}{5} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad = \int_0^1 \frac{6(x+y^2)}{5}\,dy$$

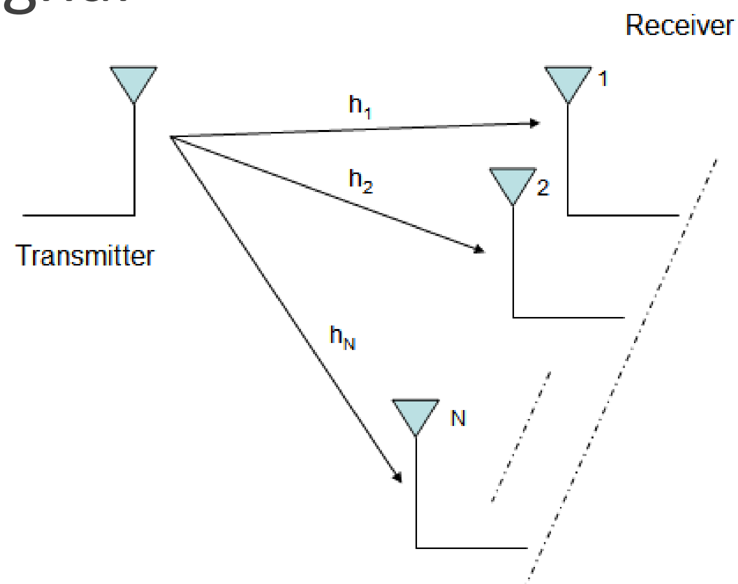
$$f_Y(y) = \begin{cases} \frac{3+6y^2}{5} & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Functions of Two Random Variables

- Assume life was easier and you had just a midterm (30% weight) and a final exam.
- Let X be the RV that denotes your midterm marks
  - $S_X = [0,30]$
- Let Y be the RV that denotes your final exam marks
  - $S_Y = [0,70]$
- Let G be the RV that denotes your grade
- $G = g(X,Y)$
  - Where g is determined by yours truly
  - Example $G = 0.3\,X + 0.7\,Y$

A More Useful Example

- More than one antenna at your receiver receives a transmitted signal



Pic courtesy http://www.dsplog.com/2008/09/06/receiver-diversity-selection-diversity/

- For N=2, we have received signals (RVs) R_1 and R_2
- The resultant received signal $R = w_1R_1 + w_2R_2$
- We could keep it simpler and only select the stronger signal
  - $R = \max(R_1, R_2)$

Functions of Two Random Variables

- We want to find the probability model that corresponds to the resulting RV
- **We know the joint probability model**
- W = X + Y
  - PMF, CDF, PDF of W?
- As always if X and Y are discrete, life is easy

Theorem 4.9

For discrete random variables X and Y, the derived random variable $W = g(X, Y)$ has PMF

$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x, y).$$

Continuous Function of Continuous RVs

- W = g(X,Y)
  - g is continuous and so are X and Y.
- W is thus a continuous RV

Theorem 4.10

For continuous random variables X and Y, the CDF of $W = g(X, Y)$ is

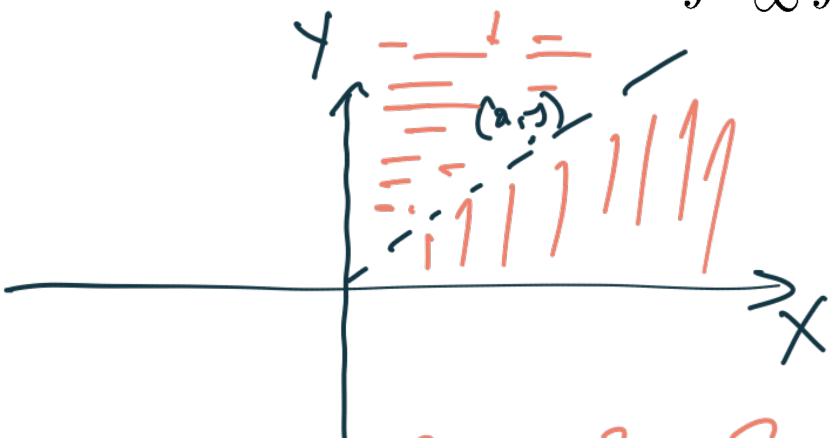
$$F_W(w) = P[W \leq w] = \iint_{g(x,y) \leq w} f_{X,Y}(x, y)\, dx\, dy.$$

Maximum of Two RVs

Theorem 4.11

For continuous random variables X and Y, the CDF of $W = \max(X, Y)$ is

$$F_W(w) = F_{X,Y}(w, w) = \int_{-\infty}^{w} \int_{-\infty}^{w} f_{X,Y}(x, y)\, dx\, dy.$$



$\{W \le w\} = \{X \le w, Y \le w\}$

What if W = min(X,Y)?

$$\{W > w\} = \{X > w, Y > w\}$$

$$P[W \le w] = 1 - P[X > w, Y > w]$$

$$= 1 - \int_w^\infty \int_w^\infty f_{X,Y}(x,y)\,dx\,dy$$

Expected Values!

Theorem 4.12

For random variables X and Y, the expected value of $W = g(X, Y)$ is

Discrete: $$E[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y}(x, y)$$

Continuous: $$E[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, dx\, dy.$$

Expected Values!

Theorem 4.13

$$E\left[g_1(X,Y)+\cdots+g_n(X,Y)\right]=E\left[g_1(X,Y)\right]+\cdots+E\left[g_n(X,Y)\right].$$

- A very useful theorem!

Theorem 4.15

The variance of the sum of two random variables is

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2E[(X-\mu_X)(Y-\mu_Y)].$$

Definition 4.4 *Covariance*

The covariance *of two random variables* X *and* Y *is*

$$\mathrm{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)].$$

Definition 4.5 Correlation

The correlation *of* X *and* Y *is* $r_{X,Y} = E[XY]$

Theorem 4.16

(a) $\mathrm{Cov}[X,Y] = r_{X,Y} - \mu_X \mu_Y$.

(b) $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y]$.

(c) If $X = Y$, $\mathrm{Cov}[X,Y] = \mathrm{Var}[X] = \mathrm{Var}[Y]$ and $r_{X,Y} = E[X^2] = E[Y^2]$.

Example 4.12 Problem

For the integrated circuits tests in Example 4.1, we found in Example 4.3 that the probability model for X and Y is given by the following matrix.

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$P_X(x)$
$x=0$	0.01	0	0	
$x=1$	0.09	0.09	0	
$x=2$	0	0	0.81	
$P_Y(y)$				

(4.73)

Find $r_{X,Y}$ and $\text{Cov}[X,Y]$.

More Definitions!

Definition 4.6 *Orthogonal Random Variables*

Random variables X *and* Y *are* orthogonal *if* $r_{X,Y} = 0$.

More Definitions!

Definition 4.7 Uncorrelated Random Variables

Random variables X *and* Y *are* uncorrelated *if* $\mathrm{Cov}[X, Y] = 0$.

Definition 4.8 *Correlation Coefficient*

The correlation coefficient *of two random variables* X *and* Y *is*

$$\rho_{X,Y} = \frac{\mathrm{Cov}\left[X, Y\right]}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}\left[X, Y\right]}{\sigma_X \sigma_Y}.$$

Correlation Coefficient

Theorem 4.17

$$-1 \leq \rho_{X,Y} \leq 1.$$

- Let W = X – aY, for any constant a
- We have Var[W] = Var[X] – 2a Cov[X,Y] + a^2 Var[Y]
- Use the fact that Var[W] >= 0
- Let a = σ_x/σ_y to get the upper bound
- Let a = - σ_x/σ_y to get the lower bound

Correlation Coefficient

- A positive correlation coefficient implies that when X is high wrt E[X], even Y tends to be high wrt E[Y]

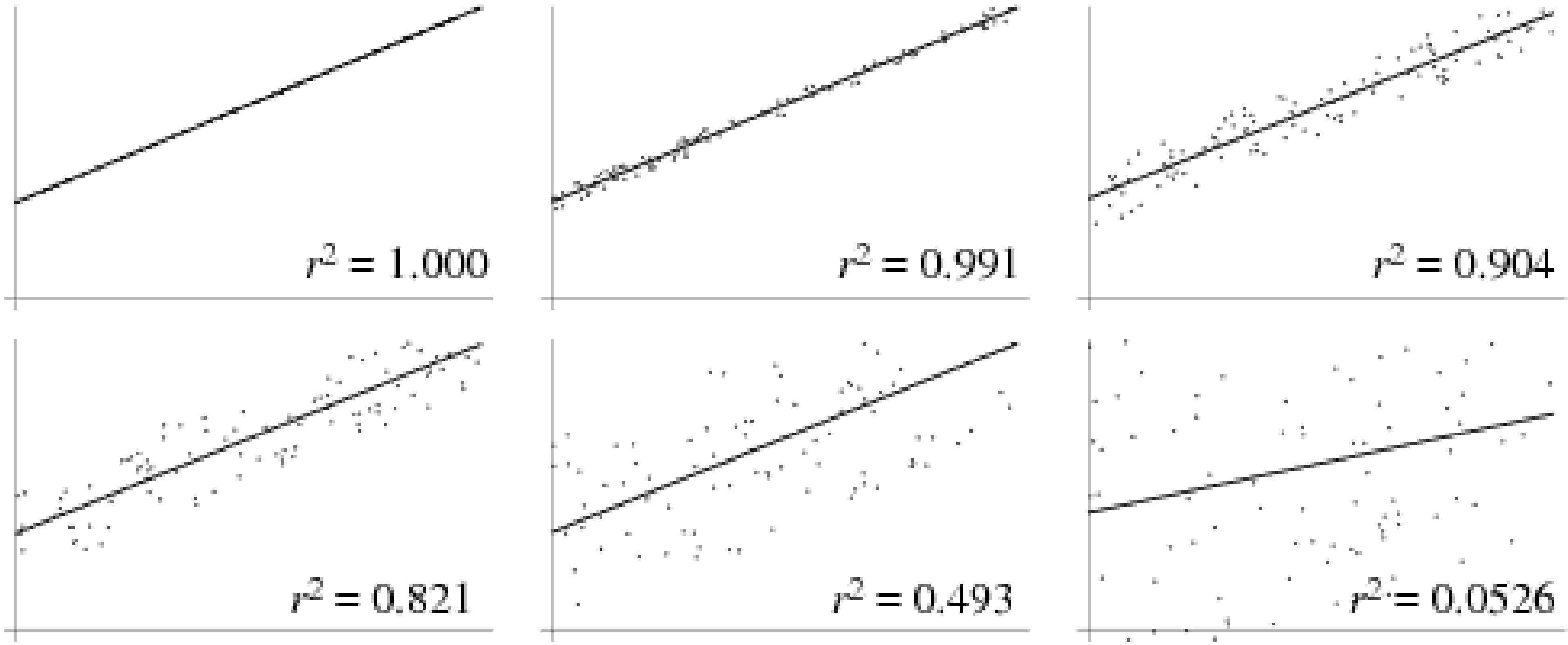
  - When we observe high values of X it is likely that Y too is high

Correlation Coefficient

- A negative coefficient implies that when X is low, Y is likely to be high and vice-versa

- If X and Y are uncorrelated then no such trend is observed

Examples (Scatter Plots)



Weisstein, Eric W. "Correlation Coefficient." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/CorrelationCoefficient.html

Y is an Affine Function of X

Theorem 4.18

If X and Y are random variables such that $Y = aX + b$,

$$\rho_{X,Y} = \begin{cases} -1 & a < 0, \\ 0 & a = 0, \\ 1 & a > 0. \end{cases}$$

Definition 4.9 Conditional Joint PMF

For discrete random variables X *and* Y *and an event,* B *with* $P[B] > 0$*, the* conditional joint PMF *of* X *and* Y *given* B *is*

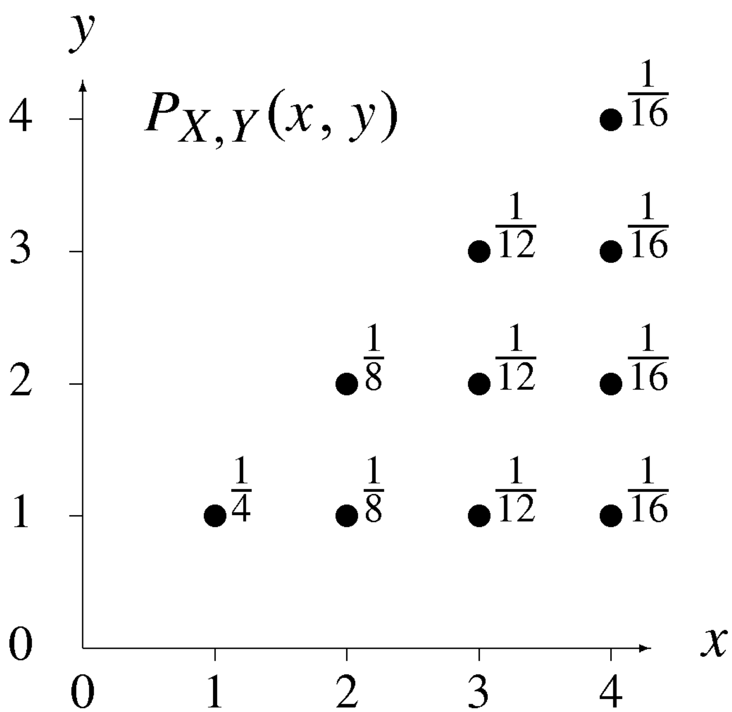
$$P_{X,Y|B}(x, y) = P[X = x, Y = y|B].$$

Theorem 4.19

For any event B, a region of the X, Y plane with $P[B] > 0$,

$$P_{X,Y|B}(x, y) = \begin{cases} \dfrac{P_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.13 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$ as shown. Let B denote the event $X + Y \leq 4$. Find the conditional PMF of X and Y given B.

Conditioning on an Event

Definition 4.10 Conditional Joint PDF

Given an event B *with* $P[B] > 0$*, the conditional joint probability density function of* X *and* Y *is*

$$f_{X,Y|B}(x, y) = \begin{cases} \dfrac{f_{X,Y}(x, y)}{P[B]} & (x, y) \in B, \\ 0 & \textit{otherwise.} \end{cases}$$

Example 4.14 Problem

X and Y are random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1/15 & 0 \le x \le 5, 0 \le y \le 3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.83}$$

Find the conditional PDF of X and Y given the event $B = \{X + Y \ge 4\}$.

Conditioning on an Event

Theorem 4.20 Conditional Expected Value

For random variables X and Y and an event B of nonzero probability, the conditional expected value of $W = g(X, Y)$ given B is

Discrete: $$E[W|B] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x, y) P_{X,Y|B}(x, y)$$

Continuous: $$E[W|B] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y|B}(x, y)\, dx\, dy.$$

- Nothing special about this

Conditioning on a RV

Definition 4.12 Conditional PMF

For any event $Y = y$ *such that* $P_Y(y) > 0$*, the* conditional PMF *of* X *given* $Y = y$ *is*

$$P_{X|Y}(x|y) = P[X = x|Y = y].$$

Conditioning on a RV

Theorem 4.22

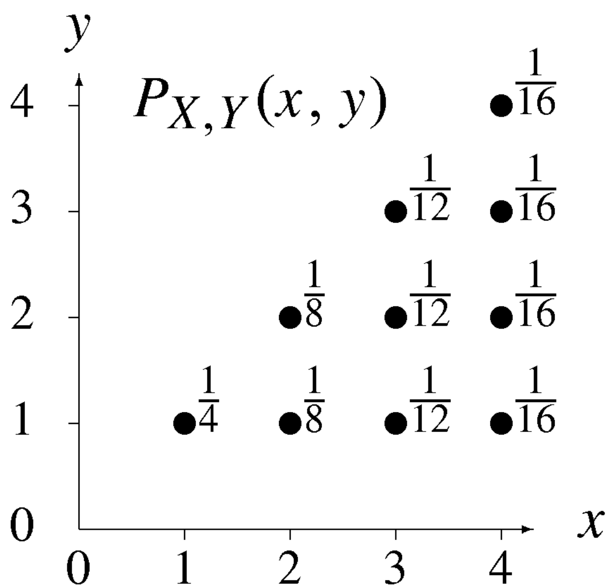
For random variables X and Y with joint PMF $P_{X,Y}(x, y)$, and x and y such that $P_X(x) > 0$ and $P_Y(y) > 0$,

$$P_{X,Y}(x, y) = P_{X|Y}(x|y)\, P_Y(y) = P_{Y|X}(y|x)\, P_X(x)\,.$$

- Why??

VS Problem

Example 4.17 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x, y)$, as given in Example 4.13 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

- How many conditional PMF(s) of Y do we have?

VVVVS Problem

- For a given *y*

$$P_{Y|X}(y|1) + P_{Y|X}(y|2) + P_{Y|X}(y|3) + P_{Y|X}(y|4) = ?$$

- For a given *x*

$$P_{Y|X}(1|x) + P_{Y|X}(2|x) + P_{Y|X}(3|x) + P_{Y|X}(4|x) = ?$$

Theorem 4.23 Conditional Expected Value of a Function

X and Y are discrete random variables. For any $y \in S_Y$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

$$E\left[g(X, Y)|Y = y\right] = \sum_{x \in S_X} g(x, y) P_{X|Y}(x|y).$$

Conditional Expectation

Theorem 4.23 Conditional Expected Value of a Function

X and Y are discrete random variables. For any $y \in S_Y$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

$$E\left[g(X, Y)|Y = y\right] = \sum_{x \in S_X} g(x, y) P_{X|Y}(x|y).$$

- The conditional expectation is a function of which variable(s)?

Conditional Expectation

Theorem 4.23 Conditional Expected Value of a Function

X and Y are discrete random variables. For any $y \in S_Y$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

$$E\left[g(X, Y)|Y = y\right] = \sum_{x \in S_X} g(x, y) P_{X|Y}(x|y).$$

- E[g(X,Y)] is a function of which variable(s)?

Now For the Continuous Case

Definition 4.13 Conditional PDF

For y *such that* $f_Y(y) > 0$, *the conditional PDF of* X *given* $\{Y = y\}$ *is*

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

Theorem 4.24

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

Problem

- You are given the conditional pdf $f_{Y|X}(y|x)$ and $f_X(x)$
- How would you find $F_Y(y)$?

Definition 4.14 Conditional Expected Value of a Function

For continuous random variables X and Y and any y such that $f_Y(y) > 0$, the conditional expected value *of $g(X, Y)$ given $Y = y$ is*

$$E\left[g(X, Y)|Y = y\right] = \int_{-\infty}^{\infty} g(x, y) f_{X|Y}(x|y)\, dx.$$

Definition 4.15 Conditional Expected Value

The conditional expected value $E[X|Y]$ *is a function of random variable* Y *such that if* $Y = y$ *then* $E[X|Y] = E[X|Y = y]$.

Conditional Expected Value

Definition 4.15 Conditional Expected Value

The conditional expected value $E[X|Y]$ *is a function of random variable* Y *such that if* $Y = y$ *then* $E[X|Y] = E[X|Y = y]$.

Note that E[X|Y] is a random variable

- It is a function of the RV ___?

Conditional Expected Value

Definition 4.15 Conditional Expected Value

The conditional expected value $E[X|Y]$ is a function of random variable Y such that if $Y = y$ then $E[X|Y] = E[X|Y = y]$.

- E[X|Y=y] is a function of ___
- E[E[X|Y]] is the expectation of a function of RV ___
- E[E[X|Y=1]] = ____ ?

Example 4.20 Problem

For random variables X and Y in Example 4.5, we found in Example 4.19 that the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.109}$$

Find the conditional expected values $E[X|Y=y]$ and $E[X|Y]$.

Example 4.20 Problem

For random variables X and Y in Example 4.5, we found in Example 4.19 that the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.109}$$

Find the conditional expected values $E[X|Y=y]$ and $E[X|Y]$.

- E[X|Y=y] = (1+y)/2
- E[X|Y] = (1+Y)/2

Independent Random Variables

Definition 4.16 Independent Random Variables

Random variables X *and* Y *are* independent *if and only if*

Discrete: $$P_{X,Y}(x, y) = P_X(x) P_Y(y)$$

Continuous: $$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

- For independent X and Y, $P_{X|Y}(x|y)$ = ?
- What about $F_{X,Y}(x,y)$?

Independent RVs

Theorem 4.27

For independent random variables X and Y,

(a) $E[g(X)h(Y)] =$,

(b) $r_{X,Y} = E[XY] =$,

(c) $\mathrm{Cov}[X, Y] =$,

(d) $\mathrm{Var}[X + Y] =$,

(e) $E[X|Y = y] =$,

(f) $E[Y|X = x] =$.

Example 4.24 Problem

$$f_{U,V}(u,v) = \begin{cases} 24uv & u \geq 0, v \geq 0, u+v \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.133}$$

Are U and V independent?

Quiz 4.10(A)

Random variables X and Y in Example 4.1 and random variables Q and G in Quiz 4.2 have joint PMFs:

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

$P_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

(1) Are X and Y independent?

(2) Are Q and G independent?

Quiz 4.10(B)

Random variables X_1 and X_2 are independentand identically distributed with probability density function

$$f_X(x) = \begin{cases} 1 - x/2 & 0 \leq x \leq 2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.144}$$

(1) What is the joint PDF $f_{X_1,X_2}(x_1, x_2)$?

(2) Find the CDF of $Z = \max(X_1, X_2)$.

Problem 4.10.13

X and Y are independent random variables with PDFs

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let $A = \{X > Y\}$.

(a) What are $E[X]$ and $E[Y]$?

(b) What are $E[X|A]$ and $E[Y|A]$?

A Shorter Exponential

Let $X \sim Exp(\lambda)$ and $Y \sim Exp(\mu)$. Assume that X and Y are independent random variables. Find $f_{X|X<Y}(x)$.

Extra Slides

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Problem 4.9.14

Suppose you arrive at a bus stop at time 0 and at the end of each minute, with probability p, a bus arrives, or with probability $1-p$, no bus arrives. Whenever a bus arrives, you board that bus with probability q and depart. Let T equal the number of minutes you stand at a bus stop. Let N be the number of buses that arrive while you wait at the bus stop.

(a) Identify the set of points (n, t) for which $P[N = n, T = t] > 0$.

(b) Find $P_{N,T}(n, t)$.

(c) Find the marginal PMFs $P_N(n)$ and $P_T(t)$.

(d) Find the conditional PMFs $P_{N|T}(n|t)$ and $P_{T|N}(t|n)$.

Problem 4.9.15

Each millisecond at a telephone switch, a call independently arrives with probability p. Each call is either a data call (d) with probability q or a voice call (v). Each data call is a fax call with probability r. Let N equal the number of milliseconds required to observe the first 100 fax calls. Let T equal the number of milliseconds you observe the switch waiting for the first fax call. Find the marginal PMF $P_T(t)$ and the conditional PMF $P_{N|T}(n|t)$. Lastly, find the conditional PMF $P_{T|N}(t|n)$.

Bivariate Gaussian Random Variables

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Joint Density of X and Y is Gaussian

Definition 4.17 Bivariate Gaussian Random Variables

Random variables X and Y have a bivariate Gaussian PDF *with parameters μ_1, σ_1, μ_2, σ_2, and ρ if*

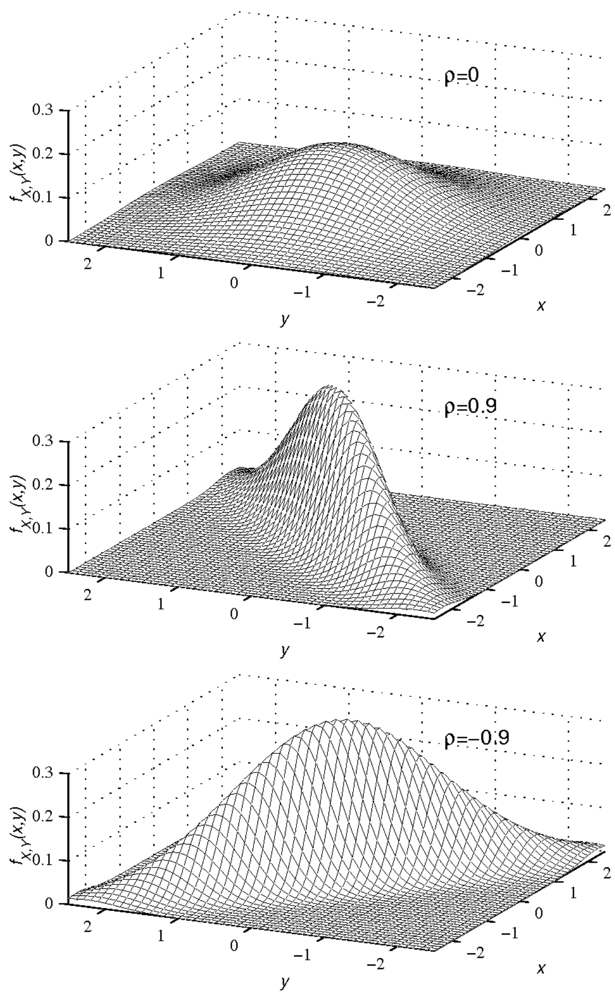
$$f_{X,Y}(x,y) = \frac{\exp\left[-\frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

where μ_1 and μ_2 can be any real numbers, $\sigma_1 > 0$, $\sigma_2 > 0$, and $-1 < \rho < 1$.

- We say that the RV(s) X and Y are $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

Examples of a Joint Gaussian PDF

Figure 4.5



The Joint Gaussian PDF $f_{X,Y}(x, y)$ for $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and three values of ρ.

If X and Y are Jointly Gaussian then they Marginally Gaussian

Theorem 4.28

If X and Y are the bivariate Gaussian random variables in Definition 4.17, X is the Gaussian (μ_1, σ_1) random variable and Y is the Gaussian (μ_2, σ_2) random variable:

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x-\mu_1)^2/2\sigma_1^2} \quad f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-\mu_2)^2/2\sigma_2^2}.$$

- **How do we prove this?**
- Marginally Gaussian may not imply jointly Gaussian

For Jointly Gaussian RVs the Conditional PDF is Gaussian

Theorem 4.29

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

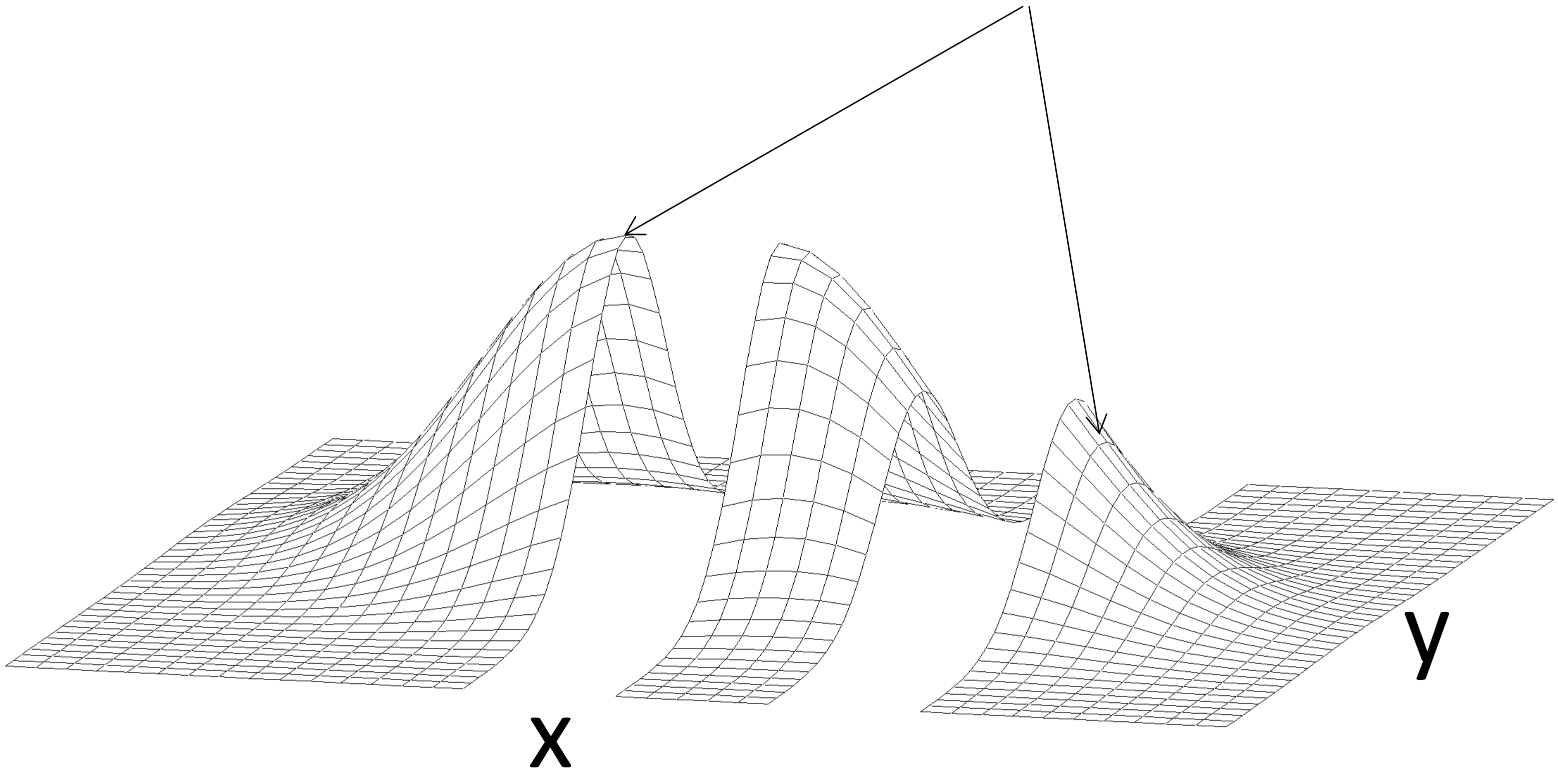
$$\underbrace{\tilde{\mu}_2(x)}_{E[Y|X=x]} = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \underbrace{\tilde{\sigma}_2^2}_{Var[Y|X=x]} = \sigma_2^2(1 - \rho^2).$$

- $E[Y|X]$ is a line that passes through (μ_1,μ_2). Also called regression line for any X and Y.
- $(x, E[Y|X=x])$ is also the point which maximizes $f_{Y|X}(y|x)$
- Following the values of $f_{X,Y}$ for points along the line implies following the maxima of all $f_{Y|X}(y|x)$

The Conditional Density

Figure 4.6

Each bell shaped curve corresponds to $f_{Y|X}(y|x)$ scaled by $f_X(x)$



Example Joint Gaussian Density

Correlation Coefficient

Theorem 4.31

Bivariate Gaussian random variables X and Y in Definition 4.17 have correlation coefficient

$$\rho_{X,Y} = \rho.$$

- We want to show that $\rho = \rho_{X,Y} = \mathrm{Cov}[XY]/(\sigma_X\ \sigma_Y) = \mathrm{Cov}[XY]/(\sigma_1\ \sigma_2)$
- $\mathrm{Cov}[XY] = E[(X - \mu_X)\ (Y - \mu_Y)] = E[(X - \mu_1)\ (Y - \mu_2)]$
- Using Iterated Expectation we get
  - $E[(X - \mu_1)\ (Y - \mu_2)] = E[(X - \mu_1)\ E[(Y - \mu_2)\,|\,X]]$
- Theorem 4.29 gave us $E[Y|X]$

Correlation Coefficient

- We have

$$E[Y-\mu_2|X=x] = E[Y|X=x]-\mu_2$$
$$= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1) - \mu_2$$

- Therefore $E[(X-\mu_1)\,E[(Y-\mu_2)|X]]$ can be written as

$$E[(X-\mu_1)\rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)] = \rho\frac{\sigma_2}{\sigma_1}E[(X-\mu_1)^2]$$

- Therefore, $\rho_{X,Y} = \mathrm{Cov}[XY]/(\sigma_1\,\sigma_2) = \rho$

Theorem 4.32

Bivariate Gaussian random variables X and Y are uncorrelated if and only if they are independent.

- If they are independent, we know that they are uncorrelated (Covariance is 0)
- Given that X and Y are jointly Gaussian and that Cov[X,Y] = 0, we have ρ=0
- How do we show that X and Y are independent?

Uncorrelated => Independent for Jointly Gaussian

We know that the joint density is given by

Random variables X *and* Y *have a* bivariate Gaussian PDF *with parameters* μ_1, σ_1, μ_2, σ_2, *and* ρ *if*

$$f_{X,Y}(x,y) = \frac{\exp\left[-\frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

where μ_1 *and* μ_2 *can be any real numbers*, $\sigma_1 > 0$, $\sigma_2 > 0$, *and* $-1 < \rho < 1$.

Substituting $\rho = 0$, we get $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Example When RVs are Uncorrelated but not Independent (Also see link)

- Consider two coin tosses. Let the outcome of the first toss be denoted by the RV X ~ Bernoulli(0.5) and the outcome of the second toss be modeled as RV Y ~ Bernoulli(0.5). Let the outcomes be independent of each other

- Clearly Cov[X,Y] = 0.
- Let Z = X – Y and W = X + Y, $S_Z = \{-1,0,1\}$ and $S_W = \{0,1,2\}$
- P[Z=0] = ½
- P[Z=0|W=2] = P[X=Y|W=2] = 1
- P[Z=1] = ¼ and P[Z=1|W=2] = 0
- Clearly W and Z are not independent
- However Cov[W,Z] = E[(X – Y) (X+Y-2μ)] = 0
- W and Z are uncorrelated but NOT independent

Variance of the Conditional Density

- The variance of the conditional density was defined in **Theorem 4.29**

If X and Y are the bivariate Gaussian random variables in Definition 4.17, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_2\sqrt{2\pi}} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

$$\tilde{\mu}_2(x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \qquad \tilde{\sigma}_2^2 = \sigma_2^2(1 - \rho^2).$$

- Var[Y|X] <= Var[Y], since ρ is the corr coeff and hence|ρ| <= 1
- If Y and X are correlated, knowing X is likely to reduce the *uncertainty* in Y

Buffon's Needle!

A fine needle of length $2a$ is dropped at random on a board covered with parallel lines distance $2b$ apart where $b > a$ as in figure. What is the probability that the needle intersects one of the lines?

Linear Combinations of Normal RV(s)

1. Let Z = aX + bY, where X and Y are **jointly normal**. Then Z is normally distributed.
   - How will you show this?
   - Find P[Z <= z], then $f_Z(z)$ and use the jointly normal distribution. Do this for at least when X and Y are independent
2. Let Z = aX + bY and W = cX + dY. X and Y are **jointly normal.** Then Z and W are jointly normal
   - Proof requires additional machinery
   - Note that 1 follows from 2