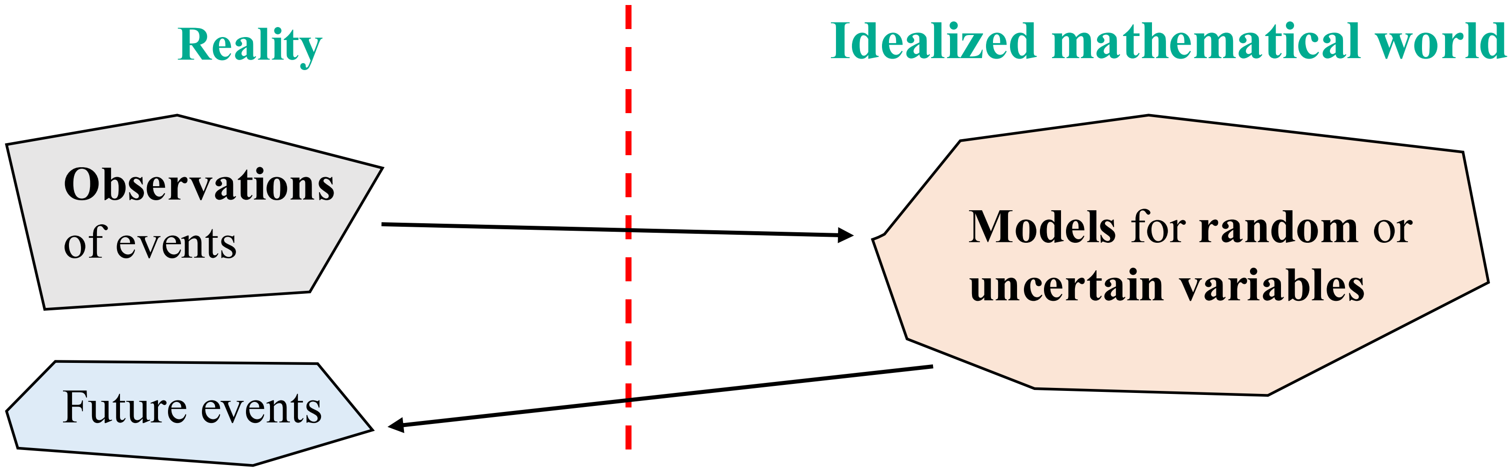
CEE 6040: Reliability Analysis & Prob. Modeling

2. Probability review II

Dr. Zaker Esteghamati (he/his/him)

Utah State University
Department of Civil and Environmental Engineering

Motivation: how to model the real world using mathematics?



- So far, we have only talked about random phenomena using symbols representing events
- Most engineering problems are better addressed using random variables
- A *random variable* is a numerical variable whose specific value cannot be predicted with certainty before an experiment

Random variables

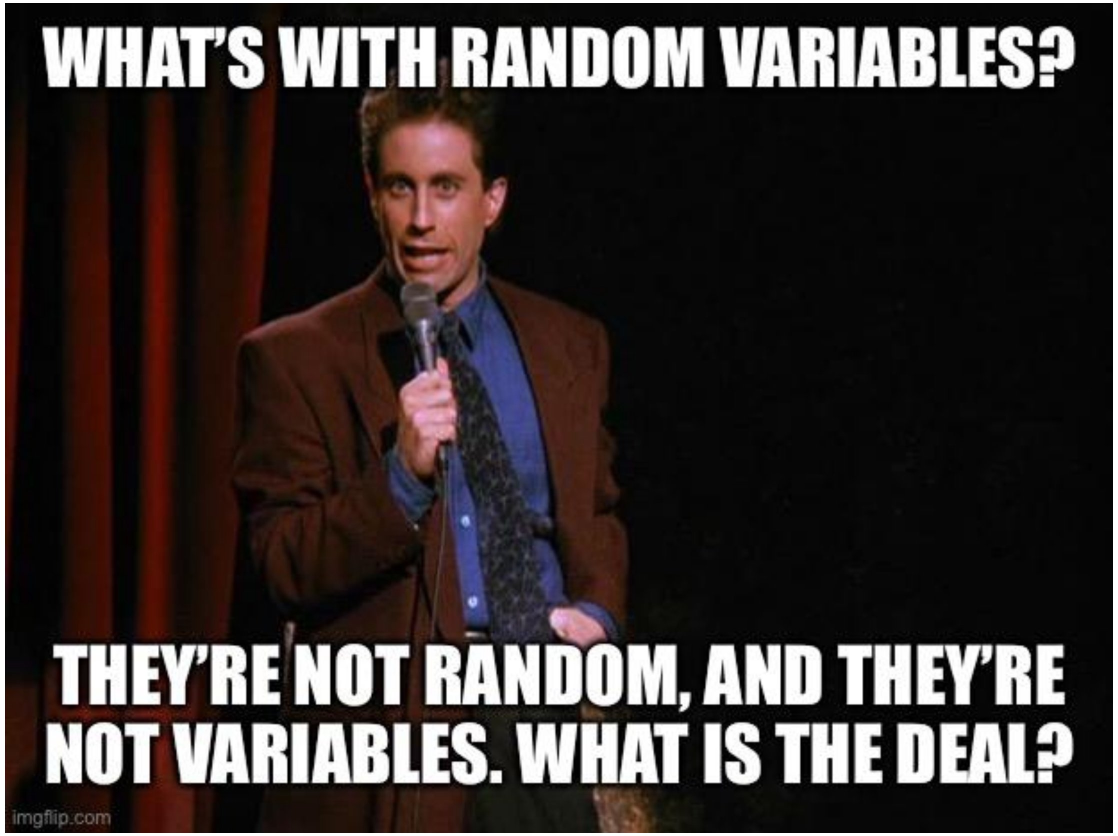
Random variables are often easy to define
– X = the magnitude of a future earthquake
– Y = the yield stress of a material
– Z = the peak wind pressure during a given year

We denote a random variable with an uppercase letter, and the values it can take on will be represented by the same letter in lowercase

- That is, x_1, x_2, x_3, … denote specific outcomes of X
- We can then talk about terms such as $P(X=x_1)$

Random variables:

1. Discrete
2. Continuous

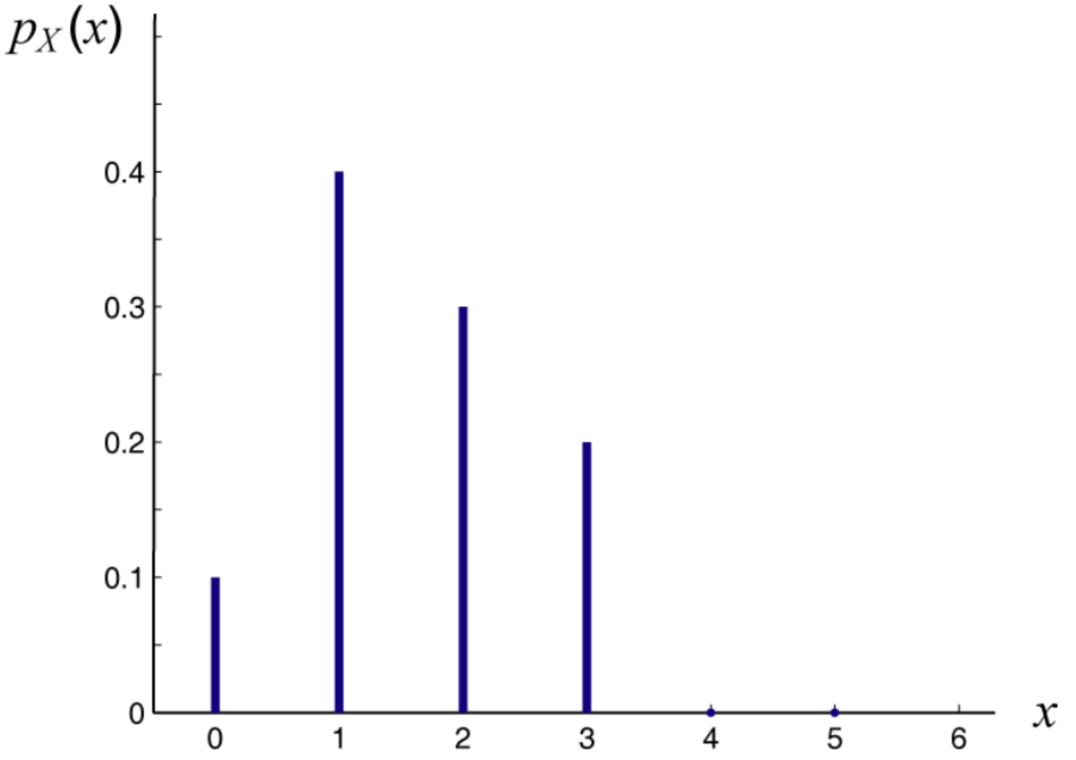


Discrete random variables

If the number of values a random variable can take on are countable

Examples:

- Number of hurricanes in a year
- Number of vehicles crossing a bridge
- Number of delays in a construction project



The probability distribution for a discrete random variable can be quantified by a probability mass function (PMF), defined as:

$$p_X(x) = P(X = x)$$

Examples:

$$p_X(1) = P(X = 1) = 0.4$$

$$p_X(2) = P(X = 2) = ?$$

Cumulative distribution function (CDF)

The cumulative distribution function is defined as the probability of the event that the random variable takes a value less than or equal to the value of the argument

$$F_x(x) = P(X \leq x)$$

The probability mass function and cumulative distribution function have a one-to-one relationship:

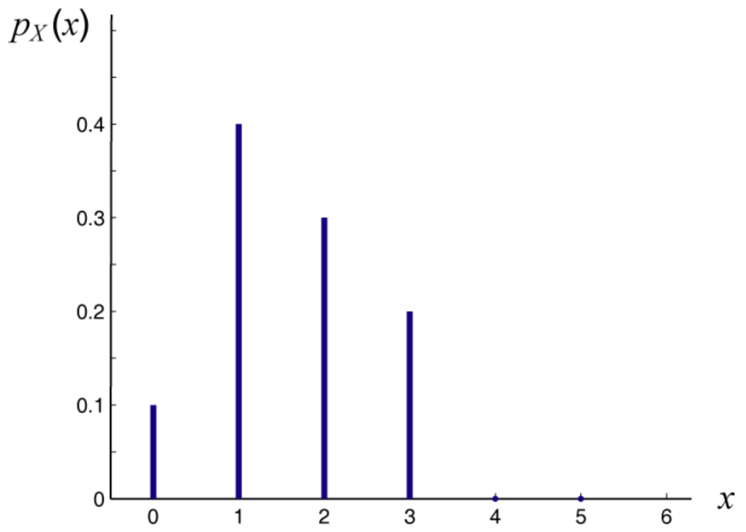
$$F_x(a) = \sum_{all\ x_i \leq a} p_x(x_i)$$

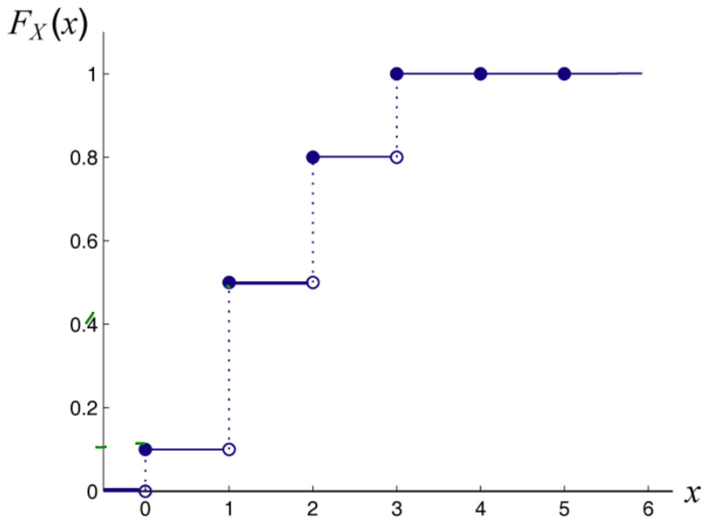
Examples: $F_x(1) = P(X \leq 1) = 0.5$ $F_x(2) = P(X \leq 2) = ?$

Related: inverse CDF

$$F_X^{-1}(u) = x \leftrightarrow F_X(x) = u$$

Examples: $F_X^{-1}(0.5) = 1$





Rules for discrete random variables

1) $0 \leq p_x(x) \leq 1$

2) $\sum_{all\ i} p_x(x_i) = 1$

✓ The CDF must follow certain rules as well

$$F_x(-\infty) = 0$$

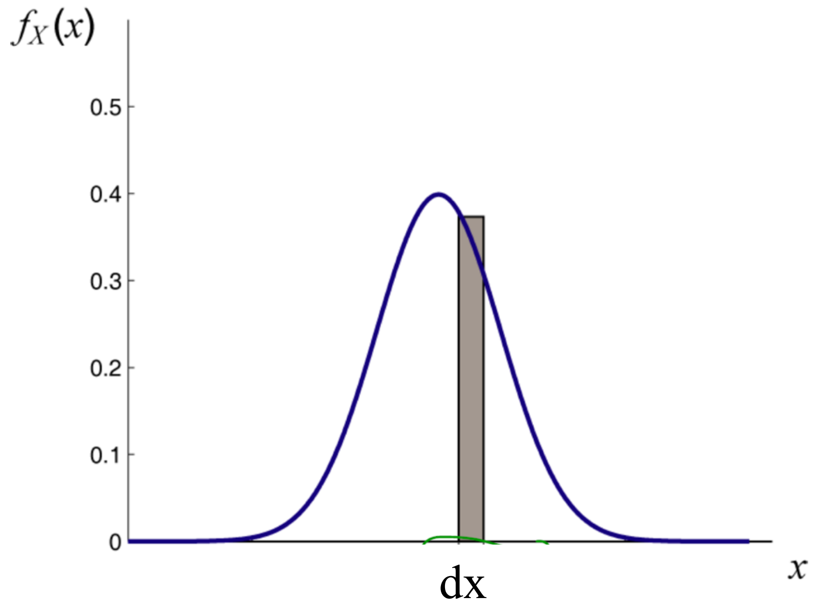
$$F_x(+\infty) = 1$$

$$F_x(b) \leq F_x(a) \text{ if } b \leq a$$

"I wish we hadn't learned probability 'cause I don't think our odds are good."

Continuous random variables

Continuous random variables can take any value on the real axis (although they don't have to). Because there are an infinite number of possible realizations, the probability that a continuous random variable X will take on any single value x is zero



Probability density function for continuous random variables

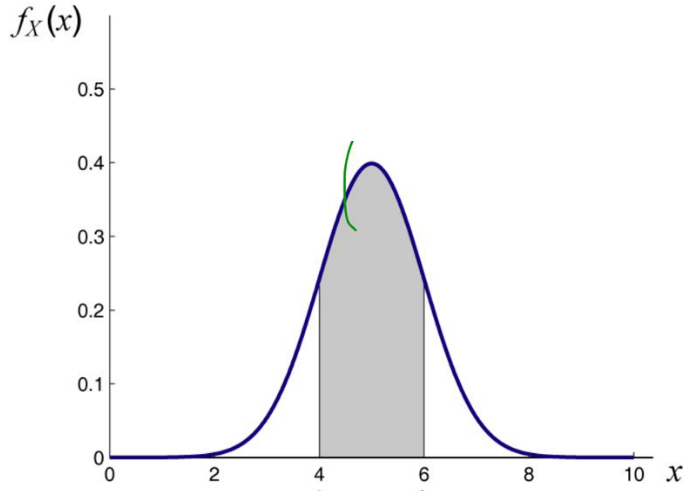
We define a *probability density function* (PDF) as follows

$$f_X(x)dx = P(x < X \leq x + dx)$$

where dx is a differential element of infinitesimal length

We can compute the probability that the outcome of X is between a and b by integrating the probability density function over the interval of interest

$$P(a < X \leq b) = \int_a^b f_X(x)dx$$



Another way to describe a continuous random variable is with a *cumulative distribution function* (CDF)

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)du$$

The PDF is related to CDF as follows:

$$f_x(x) = \frac{d}{dx} F_x(x)$$

The properties of continuous random variables:

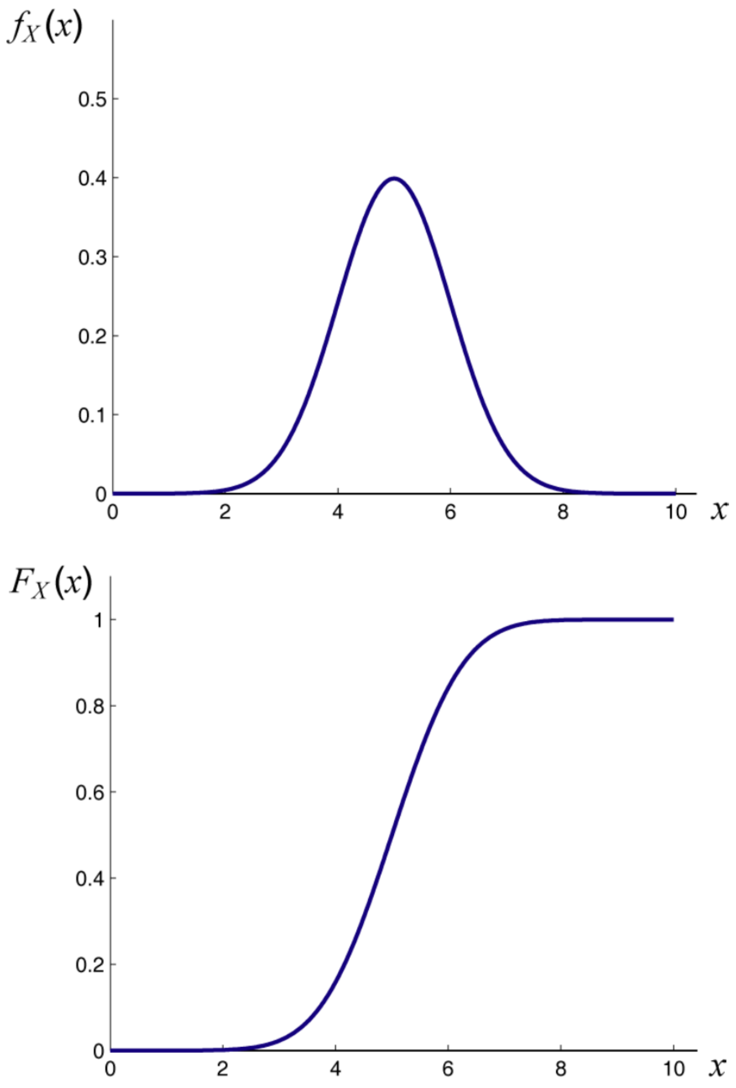
$$f_X(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f_X(x)dx = 1$$

$$F_X(-\infty) = 0$$

$$F_X(+\infty) = 1$$

$$F_X(b) \geq F_X(a) \quad if\ b \geq a$$



Jointly distributed random variables

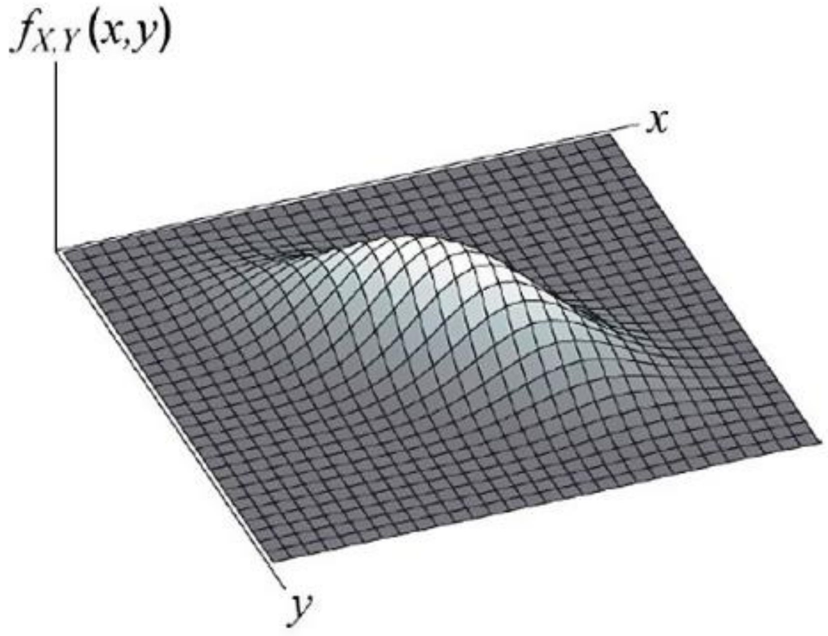
- In most engineering problems, we are worried about more than one random variable
- In this case, we need additional information that quantifies the probabilistic relationship between the random variables
- Joint probability distributions are used to define this information

Joint probability density functions

$$f_{x,y}(x,y)dxdy = P(x < X \leq x + dx \cap y < Y \leq y + dy)$$

Similarly to the single variable distribution, the probability of X, Y being in the square region ($a < X \leq b \cap c < Y \leq d$) is

$$P(a < X \leq b \cap c < Y \leq d) = \int_a^b \int_c^d f_{x,y}(u,v)dvdu$$



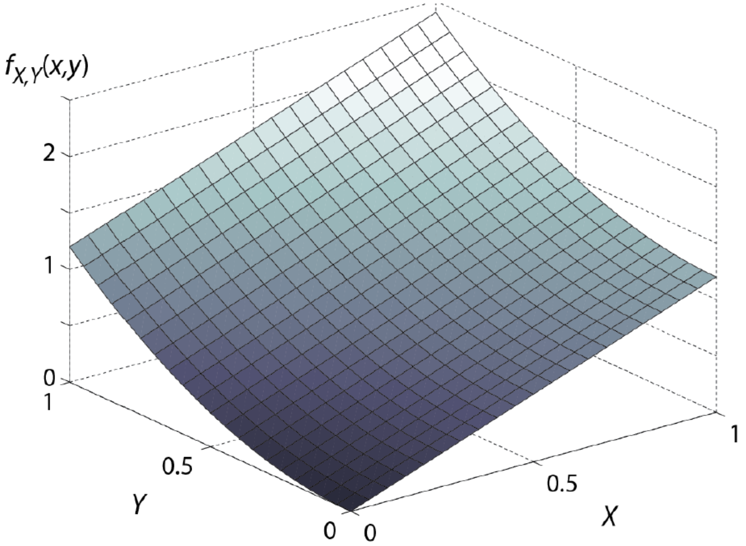
$$f_{X,Y}(x,y) \geq 0$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy = 1$$

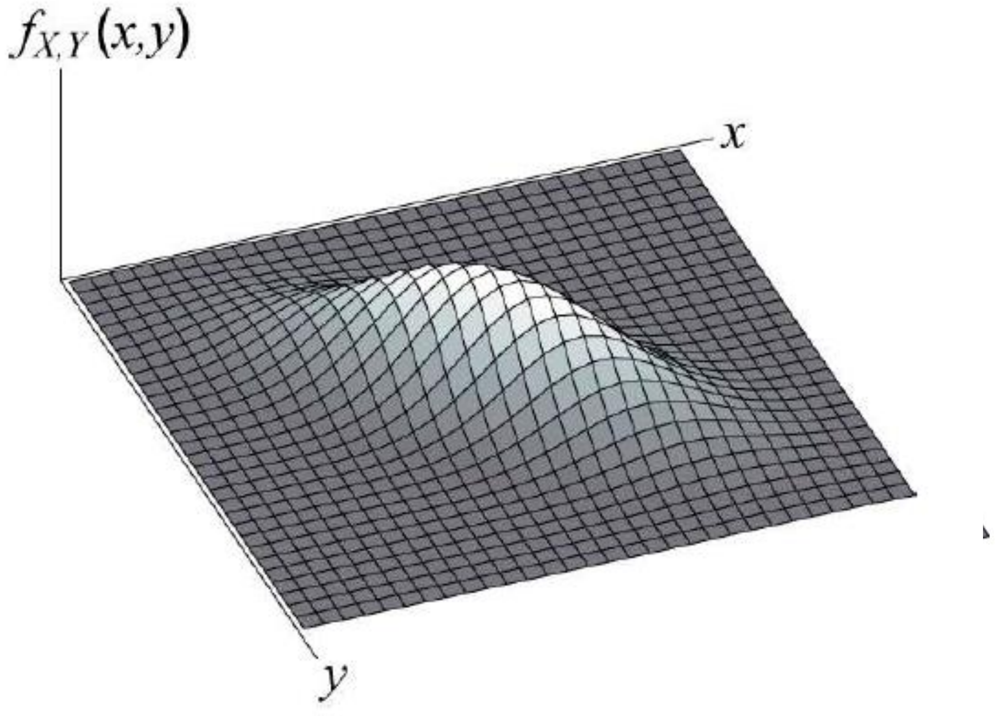
Example:. During a storm, the flows to first and second dam of Logan (relative to their capacities) is given by the following joint PDF:

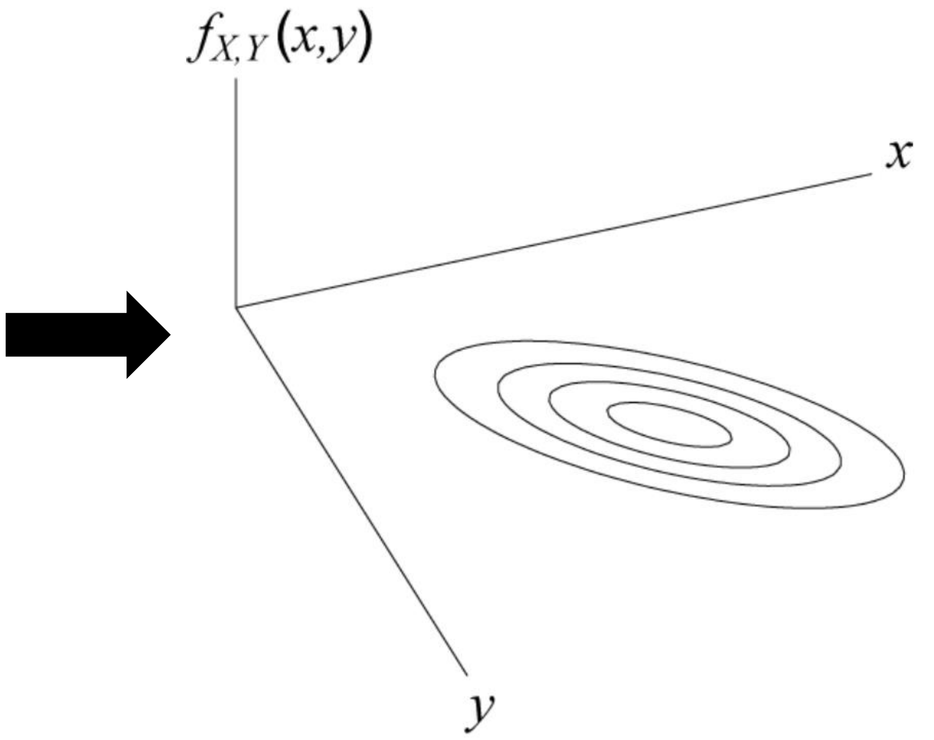
$$f_{X,Y}(x,y) = \frac{6}{5}(x + y^2) \qquad 0 \leq x, y \leq 1$$

What is the probability that flows to both dams will be more than half of their capacities?



Plotting joint distributions with contours

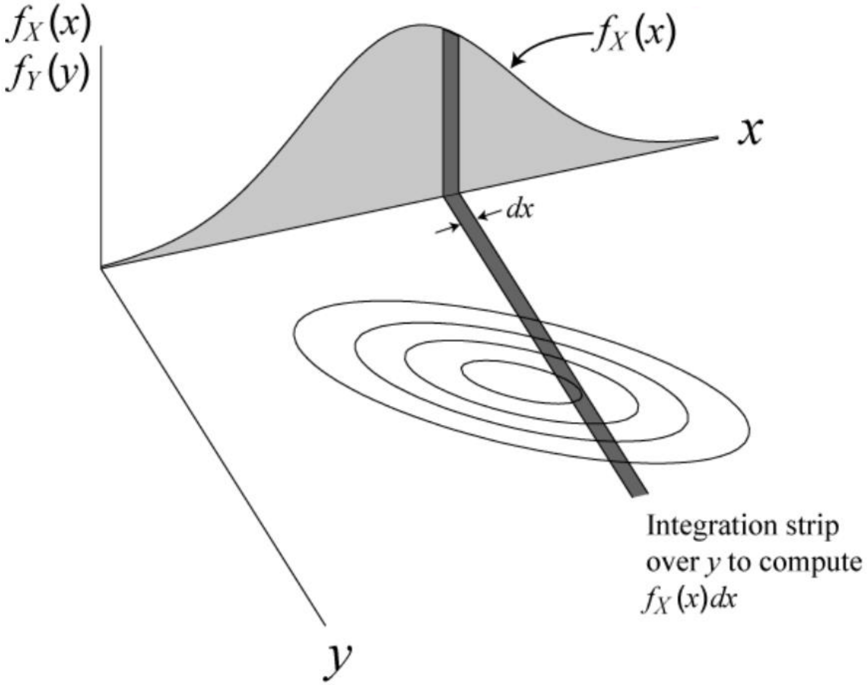


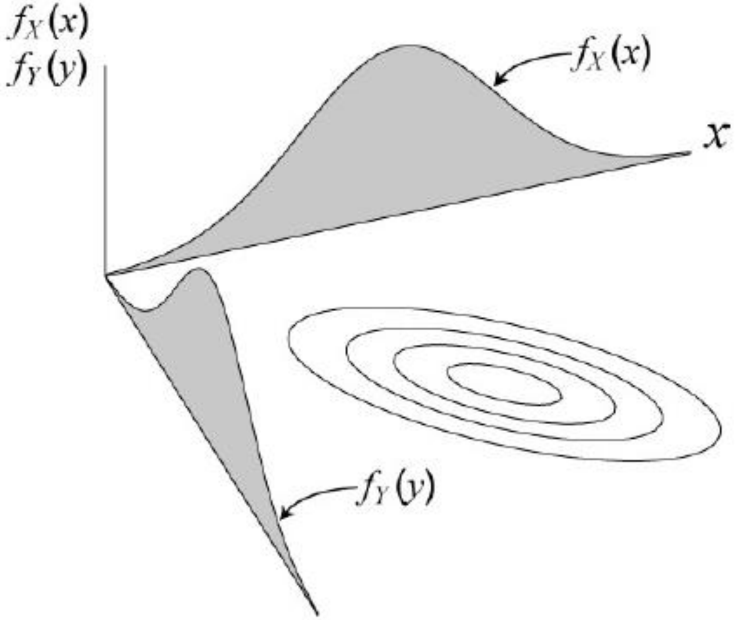


Marginal distributions

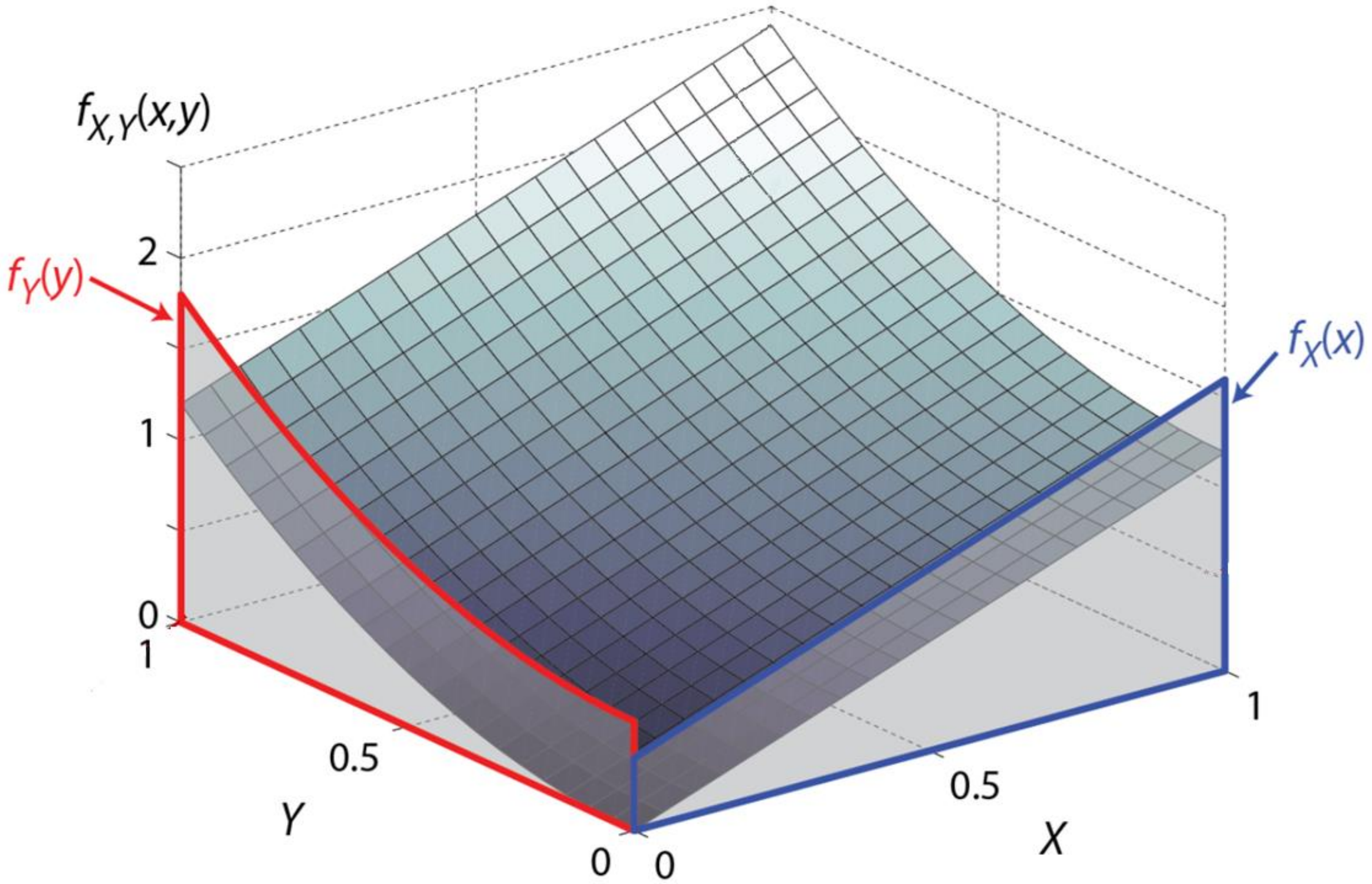
Given the joint distribution of *X* and *Y*, one can easily obtain the distribution of *X* (or *Y*) alone. This is called the *marginal distribution* of *X*

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





Example: In Logan dams example, what is the marginal PDFs of X and Y?


$f_{X,Y}(x,y)$
$f_Y(y)$
$f_X(x)$
2
1
0
1
0.5
Y
0
0
0.5
1
X

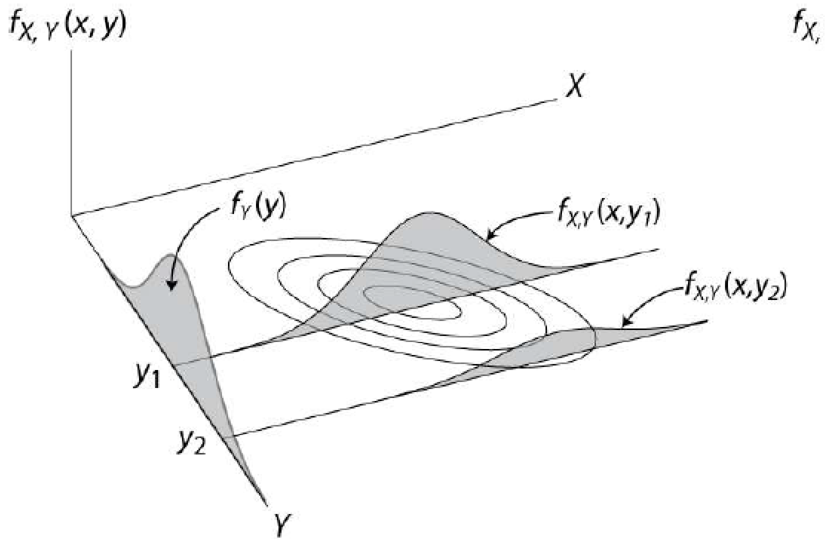
Conditional probability distributions

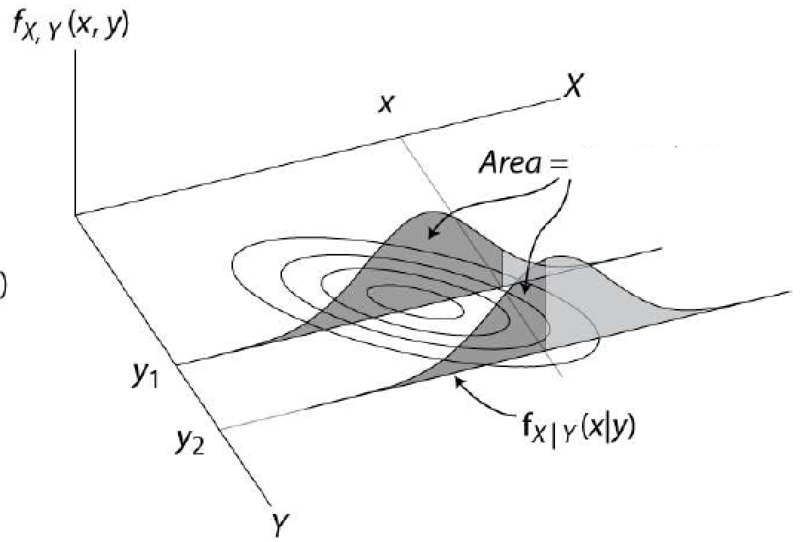
- ✓ Conditional probability distributions tell us, given that $Y = y$, what is the probability distribution of X?
- ✓ Recall the conditional probability calculation for events

$$P(E_1 | E_2) = \frac{P(E_1 E_2)}{P(E_2)}$$

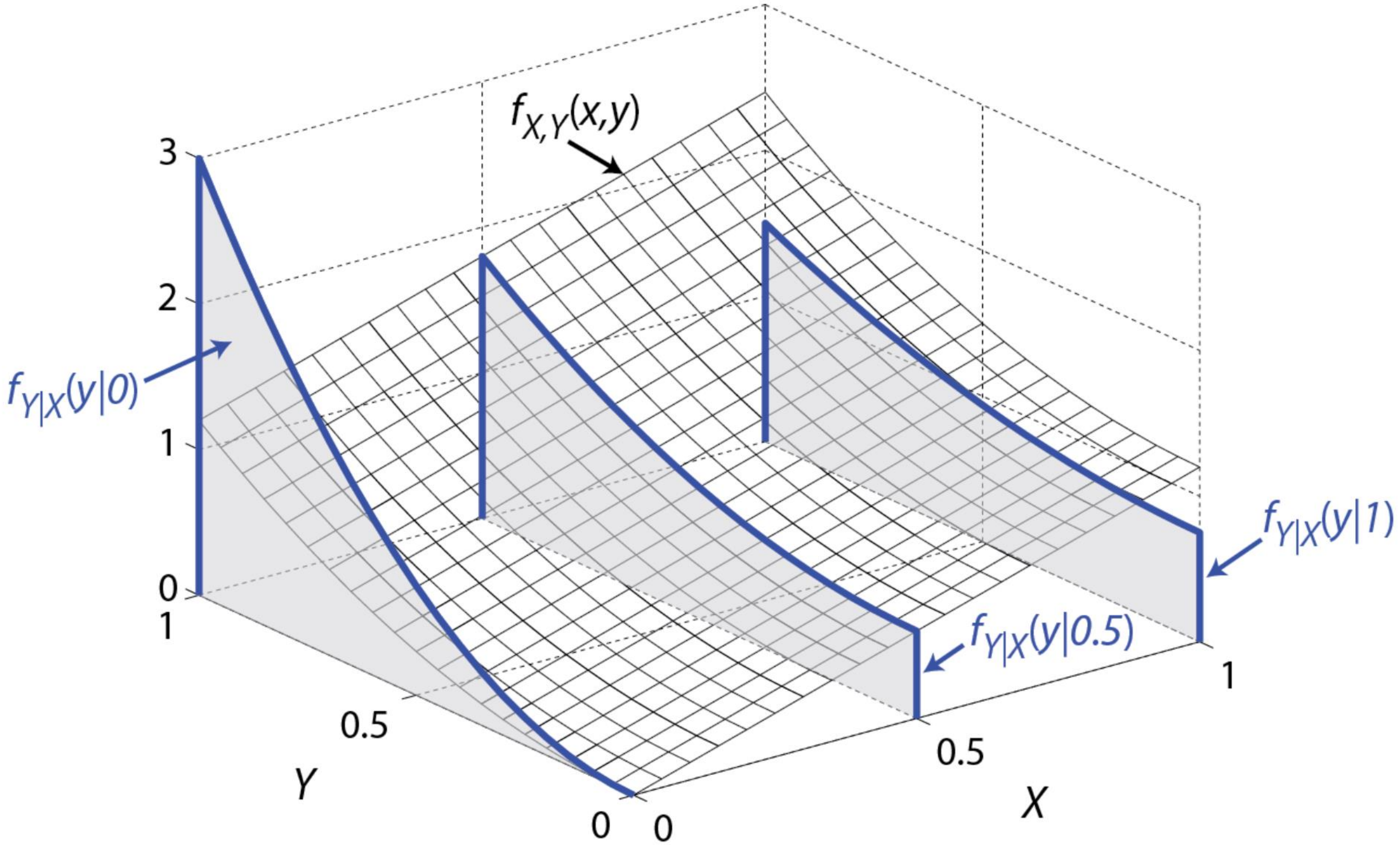
- ✓ For continuous random variables, the conditional distribution of X given Y is:

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





Example: Consider the Logan dams example again. What is the conditional PDF of Y given X?


$f_{X,Y}(x,y)$
$f_{Y|X}(y|0)$
$f_{Y|X}(y|0.5)$
$f_{Y|X}(y|1)$
3
2
1
0
1
0.5
Y
0
0
0.5
X
1

Independence

- ✓ X and Y are said to be independent random variables if for all y

$$f_{X|Y}(x \mid y) = f_X(x)$$

- ✓ This is also equivalent to the following statements

$$f_{Y|X}(y|x) = f_Y(y)$$

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

$$F_{X|Y}(x|y) = F_X(x)$$

$$F_{Y|X}(y|x) = F_Y(y)$$

$$F_{X,Y}(x,y) = F_X(x) F_Y(y)$$