Random variables — a review

- A random variable assigns a real number to each outcome in the sample space Ω
- The random variable **X** is said to map the outcome ω Ω to the real number **X**(ω)
- The function **X** from Ω to is fixed; ω is always mapped to the same number **X**(ω)
- Randomness arises from not knowing which outcome will occur on a trial and hence not knowing what numerical value of **X** will be observed on that trial

Probabilistic description

- The probabilistic behavior of any random variable **X** can be described via its CDF

$$F_X(u) = P\{X \leq u\}$$

- CDFs are cumbersome to use, but are an important concept in problem-solving
- Discrete random variables are usually discussed in terms of their pmfs
- Continuous random variables are usually discussed in terms of their pdfs

Multiple random variables

- We want to study the joint probabilistic behavior of many random variables defined on the same sample space
- Different random variables correspond to different physical parameters; their joint behavior is often of great interest
- Some experimental observations result in vectors or sequences of numbers — thus it is important to have a theory for dealing with random vectors

The joint behavior is important!

- Example: The time of arrival of requests for files at a server is a Poisson process with given arrival rate. The sizes of the files requested are random
- Large files are more likely to be requested at some particular times of the day than at other times
  - What is the distribution of the average output traffic per second?
  - What is P{buffer overflow}?

How many RVs did you say?

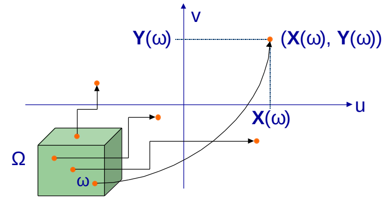
- The generalization from one random variable to two random variables is the most challenging intellectual concept
- Once the two random variable case is understood, the extension of the ideas to many random variables is easy
- We discuss the two-random-variable case in detail and indicate briefly how the theory extends to more than two variables

The random point in the plane I

- Let **X** and **Y** denote two random variables defined on the same sample space Ω
- The outcome ω Ω is mapped to the real number **X**(ω) by the random variable **X**, and to the real number **Y**(ω) by the random variable **Y**
- Jointly, the random variables **X** and **Y** are said to map the outcome ω Ω to the point (**X**(ω), **Y**(ω)) in the plane

The random point in the plane II

- Jointly, the random variables **X** and **Y** are said to map the outcome ω Ω to the point (**X**(ω), **Y**(ω)) in the plane



The random point in the plane III

- The random variables **X** and **Y** jointly map ω Ω to the point (**X**(ω), **Y**(ω)) in the plane
- We do not know which outcome ω will be occur on the next trial of the experiment
- We do not know which point in the plane will be observed as a result of the next trial of the experiment
- As the trials are repeated, we will observe different points in the plane occurring at random

The random point in the plane IV

- The random point (**X**, **Y**) in the plane is the model for our observation
- Once we know which outcome occurred on the trial, we also know which point in the plane was observed
- Before the trial is performed, we can only discuss the probability that the random point (**X**, **Y**) will be in a specified region of the plane, or will have a specific value

The random point in the plane V

- The random point (**X**, **Y**) in the plane has more information than **X** and **Y** separately
- Example: **X** is Bernoulli with parameter 0.5 and **Y** is Bernoulli with parameter 0.4
- What is P{(**X**, **Y**) = (1,1)} = P{{**X** = 1} ∩ {**Y** = 1}}?
- Knowing that P{**X** = 1} = P(A) = 0.5 and P{**Y** = 1} = P(B) = 0.4 is insufficient to tell us what P{(**X**, **Y**) = (1,1)} = P(A∩B) is

The random point in the plane VI

- The individual probabilistic descriptions of **X** and **Y** are insufficient to determine the probabilistic behavior of the random point (**X**, **Y**) in the plane
- The random point (**X**, **Y**) is also called
  - the bivariate random variable (**X**, **Y**)
  - the joint random variable (**X**, **Y**)
  - the random vector (**X**, **Y**)

The joint CDF of X and Y

- The joint CDF of **X** and **Y**, also called the
  - CDF of the random point (**X**, **Y**)
  - CDF of bivariate random variable (**X**, **Y**)
  - CDF of the random vector (**X**, **Y**)

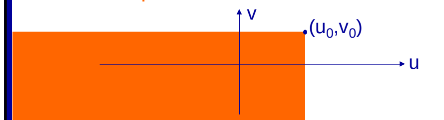
  is a function of two real variables u and v and is defined to be

  $F_{X,Y}(u,v) = P\{\{X \le u\} \cap \{Y \le v\}\}, -\infty < u, v < \infty$

  $= P\{X \le u, Y \le v\}, -\infty < u, v < \infty$
- Convention: commas mean intersections

What does the joint CDF mean?

- The joint CDF of **X** and **Y** is
  $F_{X,Y}(u,v) = P\{\{X \le u\} \cap \{Y \le v\}\}, -\infty < u, v < \infty$
  $= P\{X \le u, Y \le v\}, -\infty < u, v < \infty$
- $F_{X,Y}(u_0,v_0) = P\{X \le u_0, Y \le v_0\}$ is the total probability mass in the shaded region of the plane, i.e. the probability that the random point lies in the shaded area



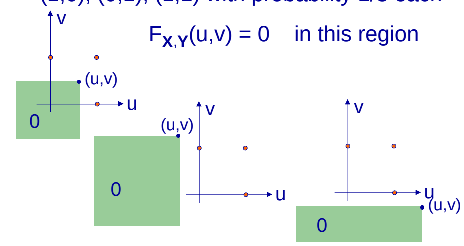
What does the joint CDF look like?

- $F_{X,Y}(u,v) = P\{X \le u, Y \le v\}$, the joint CDF of **X** and **Y**, is a function of two real variables u and v
- The value of $F_{X,Y}(u,v)$ is plotted along a third axis (say, the w axis) and defines a surface above the u-v plane
- Exercise: Why must the surface be above the plane?

An example

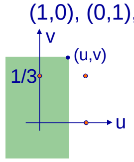
- Example: (**X**,**Y**) takes on the three values (1,0), (0,1), (1,1) with probability 1/3 each

$F_{X,Y}(u,v) = 0$ in this region



An example (continued)

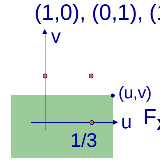
- Example: (**X**,**Y**) takes on the three values (1,0), (0,1), (1,1) with probability 1/3 each



$F_{X,Y}(u,v) = 1/3$ in this region

An example (continued some more)

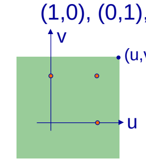
- Example: (**X**,**Y**) takes on the three values (1,0), (0,1), (1,1) with probability 1/3 each



$F_{X,Y}(u,v) = 1/3$ in this region too

An example (nearly done now)

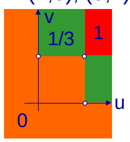
- Example: (**X**,**Y**) takes on the three values (1,0), (0,1), (1,1) with probability 1/3 each



$F_{X,Y}(u,v) = 1$ in this region

The CDF, finally!

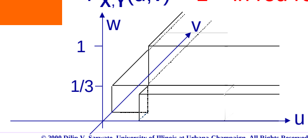
- Example: (**X**,**Y**) takes on the three values (1,0), (0,1), (1,1) with probability 1/3 each



$F_{X,Y}(u,v) = 0$ in orange region

$F_{X,Y}(u,v) = 1/3$ in green region

$F_{X,Y}(u,v) = 1$ in red region

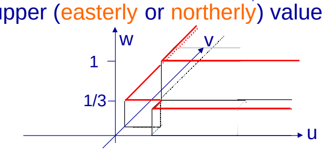


Properties of the joint CDF I

- Most of the properties of the joint CDF are analogous to the properties of the CDFs of single random variables
- Proofs are very similar, and are omitted
- $0 \le F_{X,Y}(u,v) \le 1$ for all u and v
- $F_{X,Y}(u,v)$ is a non-decreasing function of u, and a non-decreasing function of v, that is, if a > 0, then $F_{X,Y}(u+a,v) \ge F_{X,Y}(u,v)$ and $F_{X,Y}(u,v+a) \ge F_{X,Y}(u,v)$

Properties of the joint CDF II

- $F_{X,Y}(u,v)$ is a right-continuous function of u, and a right-continuous function of v, that is,
  $F_{X,Y}(u,v) = F_{X,Y}(u^+,v)$
  and $F_{X,Y}(u,v) = F_{X,Y}(u,v^+)$
- At any cliff, the value of $F_{X,Y}(u,v)$ is the upper (easterly or northerly) value



Properties of the joint CDF III

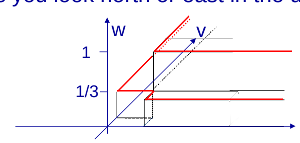
- $\lim_{u\to\infty}\lim_{v\to\infty} F_{X,Y}(u,v) = F_{X,Y}(\infty,\infty) = 1$
- $\lim_{u\to-\infty}\lim_{v\to-\infty} F_{X,Y}(u,v) = F_{X,Y}(-\infty,-\infty) = 0$
- In fact, for fixed real number v, $\lim_{u\to-\infty} F_{X,Y}(u,v) = F_{X,Y}(-\infty,v) = 0$ while for fixed real number u, $\lim_{v\to-\infty} F_{X,Y}(u,v) = F_{X,Y}(u,-\infty) = 0$
- The individual CDFs of **X** and **Y** can be obtained from the joint CDF of **X** and **Y**

Properties of the joint CDF IV

- For fixed real number u,
  $\lim_{v\to\infty} F_{X,Y}(u,v) = F_{X,Y}(u,\infty)$
  $= P\{\{X \le u\} \cap \{Y \le \infty\}\}$
  $= P\{\{X \le u\} \cap \Omega\} = P\{X \le u\} = F_X(u)$!!
- Similarly, $F_{X,Y}(\infty, v) = F_Y(v)$
- Nomenclature: When $F_X(u)$ and $F_Y(v)$ are obtained in this fashion from the joint CDF $F_{X,Y}(u,v)$, they are referred to as the marginal CDFs of **X** and **Y** respectively

Properties of the joint CDF V

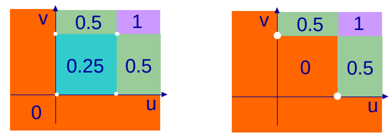
- Marginal CDFs are obtained by "setting the unwanted variable to ∞"
- Marginal ≠ barely making the grade!
- Marginal CDF is the "silhouette of the hills" as you look north or east in the u-v plane



Properties of the joint CDF VI

- The marginal CDFs $F_X(u)$ and $F_Y(v)$ can be determined from the joint CDF $F_{X,Y}(u,v)$
- This transformation is many-to-one — different joint CDFs can give rise to the same pair of marginal CDFs
- Therefore, it is generally not possible to determine the joint CDF from the marginal CDFs

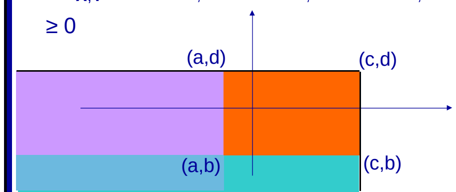
You can't go home again…



- Both joint CDFs have the same marginal CDFs — both **X** and **Y** are Bernoulli random variables with parameter 0.5
- Left picture: **X** and **Y** are independent RVs
  Right picture: **Y** = 1 – **X**

Properties of the joint CDF VII

- For all a < c and b < d,
  $P\{a < \mathbf{X} \le c, b < \mathbf{Y} \le d\}$
  $= F_{X,Y}(c,d) - F_{X,Y}(c,b) - F_{X,Y}(a,d) + F_{X,Y}(a,b)$
  ≥ 0



Exercises

- Prove that $F_{X,Y}(u,v)$ is
  - a non-decreasing function of u
  - a non-decreasing function of v
- Express
  - $P\{\{\mathbf{X} \le u\} \quad \{\mathbf{Y} \le v\}\}$
  - $P\{\{\mathbf{X} \le u\} \quad \{\mathbf{Y} > v\}\}$
  - $P\{\mathbf{X} > u, \mathbf{Y} > v\}$

  in terms of the joint CDF $F_{X,Y}(u,v)$

Generalization to many RVs

- Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$, be n random variables defined on a sample space
- Let $\underline{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$ be a random vector
- Let $\underline{u} = (u_1, u_2, \ldots, u_n)$ be a real vector
- The notation $\{\underline{\mathbf{X}} \le \underline{u}\}$ denotes the event $\{\mathbf{X}_1 \le u_1, \mathbf{X}_2 \le u_2, \ldots, \mathbf{X}_n \le u_n\}$, where, as before, the commas denote intersections, that is,

$\{\underline{\mathbf{X}} \le \underline{u}\} = \{\mathbf{X}_1 \le u_1\} \cap \{\mathbf{X}_2 \le u_2\} \cap \ldots \cap \{\mathbf{X}_n \le u_n\}$

The joint CDF of many RVs

- The joint CDF of $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ or the CDF of the random vector $\underline{\mathbf{X}}$ is defined as
  $F_{\underline{X}}(\underline{u}) = P\{\underline{\mathbf{X}} \le \underline{u}\}$
  $= P\{\mathbf{X}_1 \le u_1, \mathbf{X}_2 \le u_2, \ldots, \mathbf{X}_n \le u_n\}$
- $F_{\underline{X}}(\underline{u})$ is a real-valued function of n real variables (or of the n-vector $\underline{u}$)
- $F_{\underline{X}}(\underline{u})$ always has value between 0 and 1
- $F_{\underline{X}}(\underline{u})$ is a non-decreasing right-continuous function of each argument u_i

More on the joint CDF of many RVs

- $\lim_{u_i \to -\infty} F_{\underline{X}}(\underline{u}) = 0$
- If some of the $u_i \to +\infty$, the corresponding random variables $\mathbf{X}_i$ disappear and we get the joint CDF of the remaining variables
- Example: If $F_{X,Y,Z}(u,v,w)$ is the joint CDF of **X**, **Y**, **Z**, then $F_{X,Y,Z}(u,\infty,w) = F_{X,Z}(u,w)$ is the joint CDF of **X** and **Z**
- Even though this is a joint CDF, it is called a marginal (in the sense of derived) CDF

Thoughts on joint CDFs

- Joint CDFs are generally cumbersome to work with even in the two variable case
- Fortunately, since **X** and **Y** being discrete random variables or continuous random variables are the two most important cases, it is possible to use simpler descriptions of the probabilistic behavior
- We shall discuss both cases with special emphasis on the case when **X** and **Y** are jointly continuous random variables

Summary

- We have studied the notion of a random point in the plane as describing the joint behavior of two random variables
- We have described the probabilistic joint behavior of the two random variables in terms of their joint CDF
- We have generalized these notions to multiple random variables or random vectors