ECE 307 – Techniques for Engineering Decisions

10. Basic Probability Review

George Gross

Department of Electrical and Computer Engineering

University of Illinois at Urbana-Champaign

OUTLINE

- Definitions
- Axioms on probability
- Conditional probability
- Independence of events
- Probability distributions and densities
  - discrete
  - continuous

SAMPLE SPACE

- Consider an experiment with uncertain outcomes but with the entire set of all possible outcomes known
- The *sample space* $\mathcal{S}$ is the *set of all possible outcomes*, *i.e.*, every outcome is an element of $\mathcal{S}$

SAMPLE SPACE

- Examples of sample spaces

  - flipping a coin: $\mathcal{S} = \{ H, T \}$

  - tossing a die: $\mathcal{S} = \{ 1, 2, 3, 4, 5, 6 \}$

  - flipping two coins: $\mathcal{S} = \{ (H, H), (H, T), (T, H), (T, T) \}$

  - tossing two dice: $\mathcal{S} = \{ (i, j) : i, j = 1, \ldots, 6 \}$

  - hours of life of a device: $\mathcal{S} = \{ x : 0 \leq x < \infty \}$

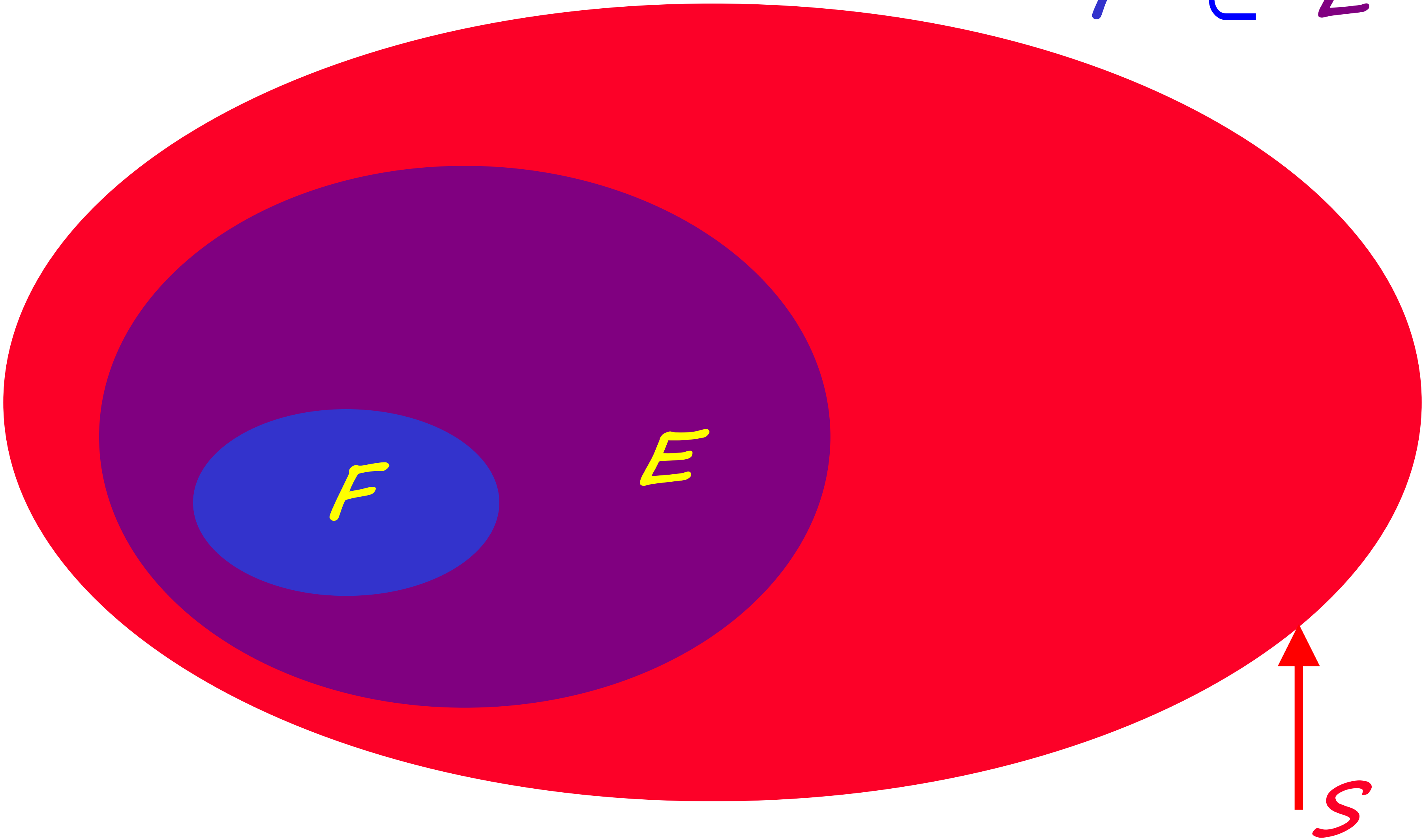
SUBSETS

- We say a set E is a subset of a set F if E is contained in F and we write $E \subset F$ or $F \supset E$
- If E and F are sets of events, then $E \subset F$ implies that each event in E is also an event in F
- Theorem

$$E \subset F \text{ and } F \supset E \Leftrightarrow E = F$$

SUBSETS

$F \subset E$



EVENTS

- An *event* $\mathcal{E}$ is an element or a subset of the *sample space* $\mathcal{S}$

- Some examples of events are:
  - flipping a coin: $\mathscr{E} = \{ H \}, \ \mathscr{F} = \{ T \}$
  - tossing a die: $\mathscr{E} = \{ 2, 4, 6 \}$ is the event that the die lands on an even number

EVENTS

- **flipping two coins:** $\mathcal{E} = \{(H,H),(H,T)\}$ **is the event of the outcome** H **on the first coin**
- **tossing two dice:**

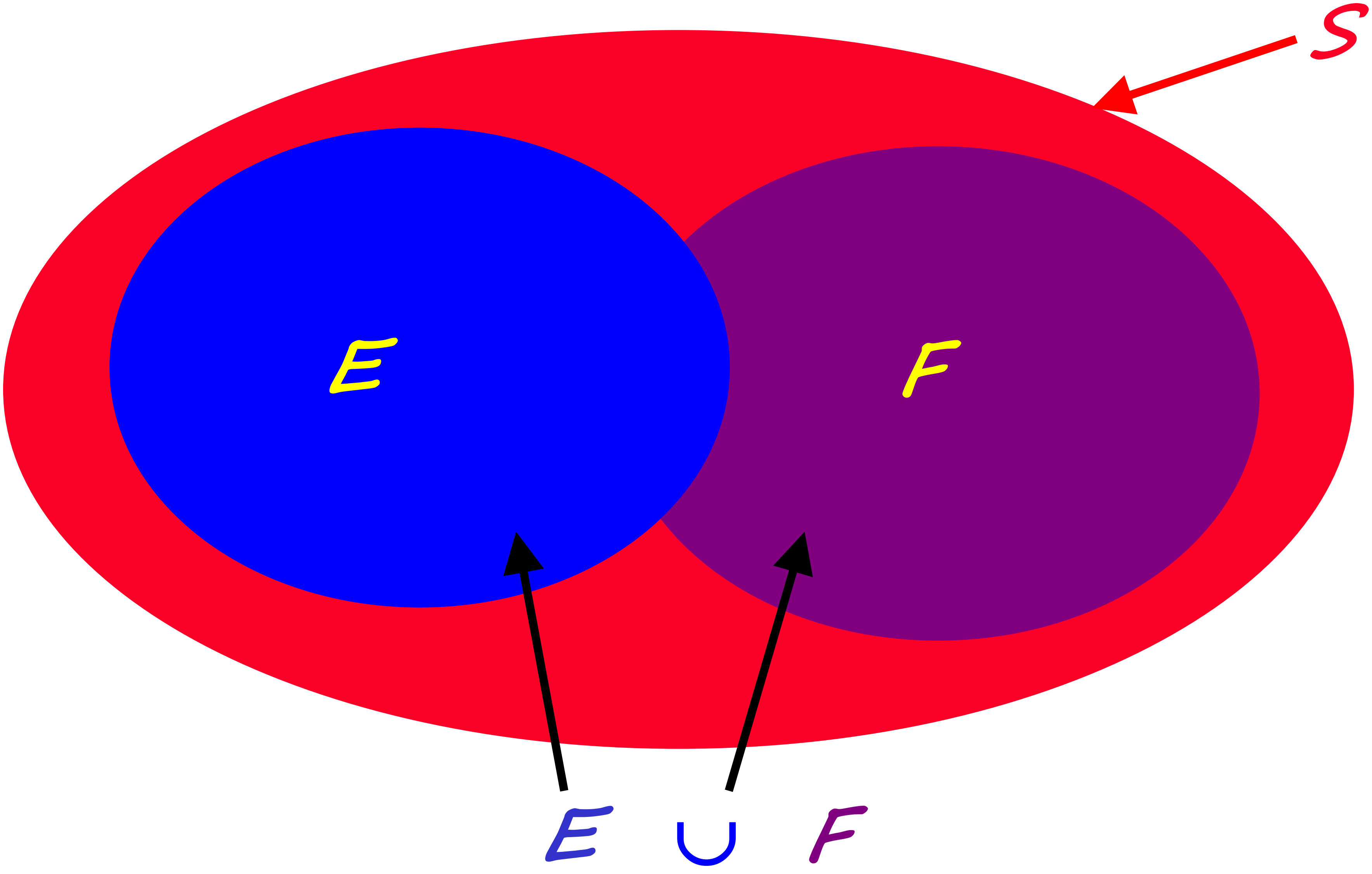
$$\mathcal{E} = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$

  is the event of sum of the two tosses is 7
- **hours of life of a device:** $\mathcal{E} = \{5 < x \leq 10\}$ **is the event that the life of a device is greater than 5 and at most** 10 ***hours***

UNION OF SUBSETS

- We consider two subsets E and F ; the *union* of E and F is denoted by $E \cup F$ and is the set of all the elements that are either in E *or* in F *or* in both E and F
- If E and F represent subsets of events, then set $E \cup F$ occurs only if either E *or* F or both occur
- $E \cup F$ is equivalent to the logical *or*

UNION OF SUBSETS



UNION OF SUBSETS

- **Examples:**
  - $\mathcal{E} = \{2, 4, 6\}, \ \mathcal{F} = \{1, 2, 3\} \Rightarrow \mathcal{E} \cup \mathcal{F} = \{1, 2, 3, 4, 6\}$
  - $\mathcal{E} = \{H\}, \ \mathcal{F} = \{T\} \Rightarrow \mathcal{E} \cup \mathcal{F} = \{H, T\} = \mathcal{S}$
  - **$\mathcal{E}$ = set of outcomes of tossing two dice with sum being an even number**

    $\mathcal{F}$ = set of outcomes of tossing two dice with sum being an odd number

    $$\Longrightarrow \quad \mathcal{E} \cup \mathcal{F} = \mathcal{S}$$

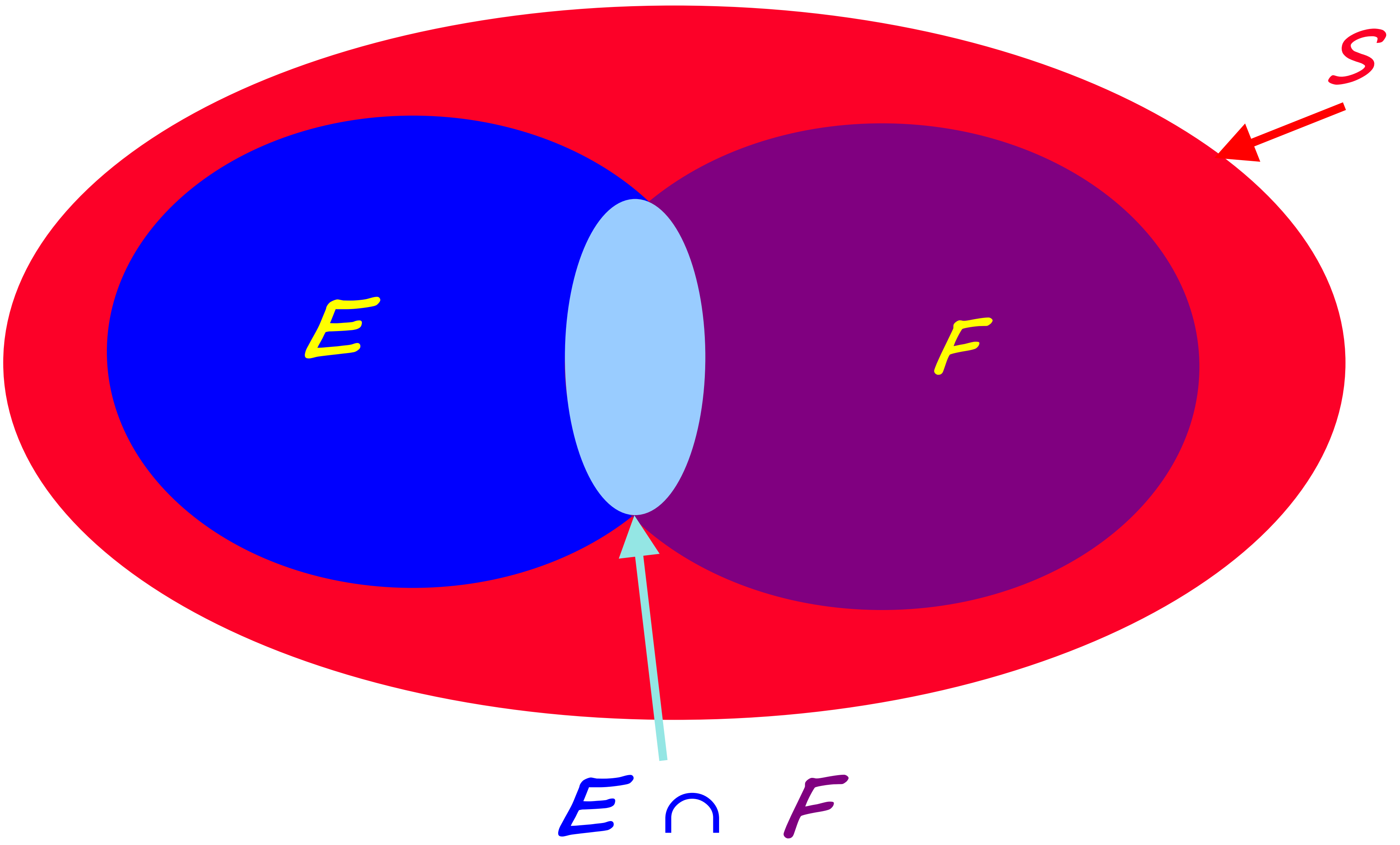
INTERSECTION OF SUBSETS

- We consider two subsets E and F ; the intersection of E and F, denoted by $E \cap F$, is the set of all the elements that are both in E *and* in F
- E and F represent subsets of events, then the events in $E \cap F$ occur only if both E and F occur

INTERSECTION OF SUBSETS

- We define $\varnothing$ to be the *empty* set, i.e., the set consisting of no elements
- For event subspaces E and F, if $E \cap F = \varnothing$ if and only if E and F are *mutually exclusive* events
- Examples:
  - $\mathscr{E} = \{H\}, \mathscr{F} = \{T\} \Rightarrow \mathscr{E} \cap \mathscr{F} = \varnothing$
  - $\mathscr{E} = \{1,3,5\}, \mathscr{F} = \{1,2,3\} \Rightarrow \mathscr{E} \cap \mathscr{F} = \{1,3\}$

INTERSECTION OF SUBSETS



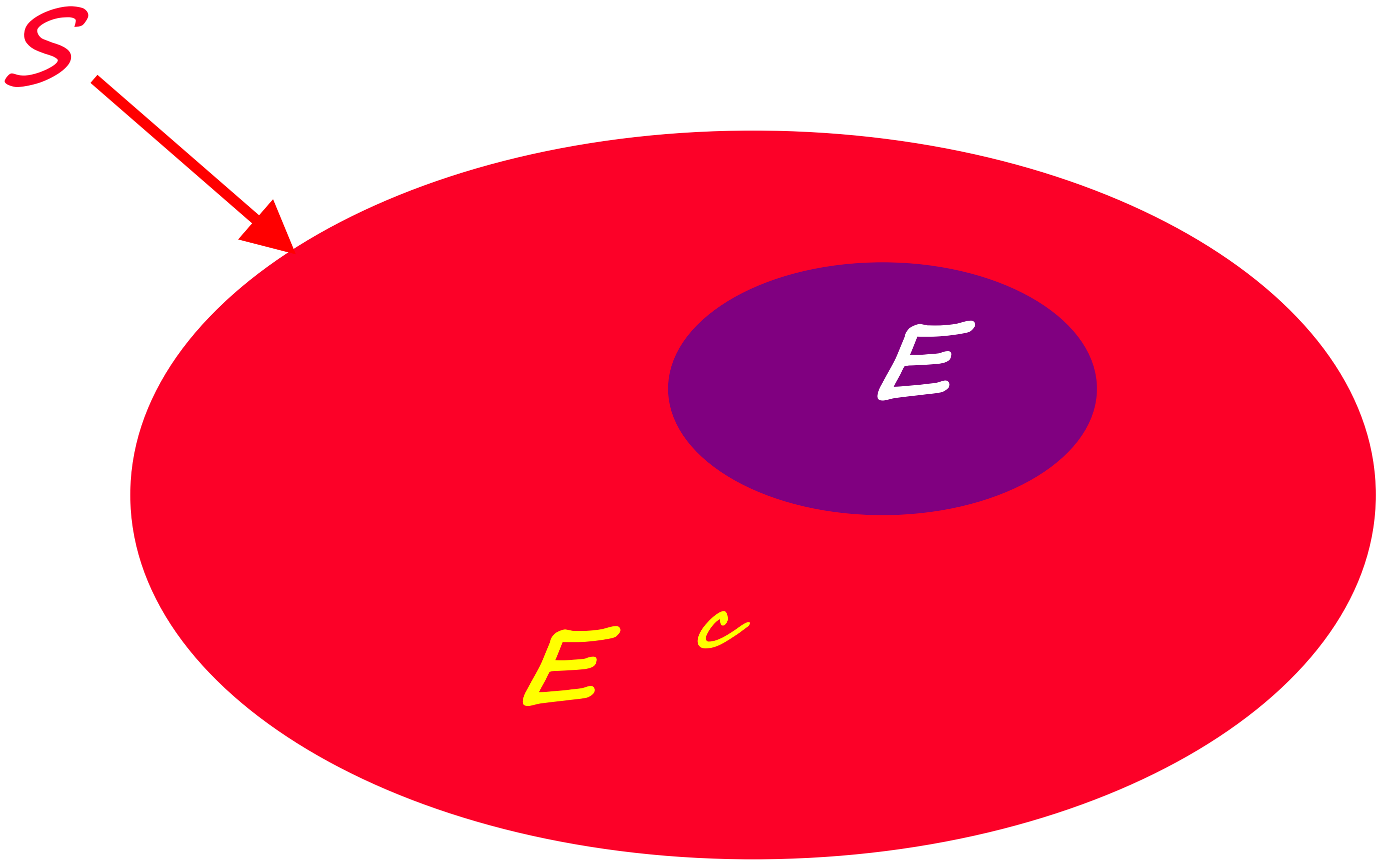
GENERALIZATION OF CONCEPTS

- We consider the countable subsets $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \ldots$ in the state space $\mathcal{S}$
- The term $\bigcup_i \mathcal{E}_i$ is defined to be that subset consisting of those elements that are in $\mathcal{E}_i$ for *at least one value* of $i = 1, 2, \ldots$
- The term $\bigcap_i \mathcal{E}_i$ is defined to be the subset consisting of those elements that are *in every subset* $\mathcal{E}_i, i = 1, 2, \ldots$

COMPLEMENT OF A SUBSET

- The complement $\mathcal{E}^c$ of a set $\mathcal{E}$ is the set of all elements in the sample space $\mathcal{S}$ not in $\mathcal{E}$
- By definition, $\mathcal{S}^c = \varnothing$
- For the example of tossing two dice, the subset $\mathcal{E} = \{(1,6),\ (2,5),\ (3,4),\ (4,3),\ (5,2),\ (6,1)\}$ is the collection of events that the sum of dice is 7; then, $\mathcal{E}^c$ is the collection of events that the sum of dice is **not** 7

COMPLEMENT OF A SUBSET



DE MORGAN'S LAWS

- De Morgan's laws establish some important relationships between $\bigcup$, $\bigcap$ and ^{c}
- The first De Morgan law states:

$$\left(\bigcup_{i=1}^{n} \mathcal{E}_i \right)^{c} = \bigcap_{i=1}^{n} \mathcal{E}_i^{c}$$

- The second De Morgan law states:

$$\left(\bigcap_{i=1}^{n} \mathcal{E}_i \right)^{c} = \bigcup_{i=1}^{n} \mathcal{E}_i^{c}$$

DEFINITION OF PROBABILITY

- Consider an event $\mathcal{E}$ in the sample space $\mathcal{S}$ and let us denote by $n(\mathcal{E})$ the number of times that the event $\mathcal{E}$ occurs in a total of n random draws
- We define the *probability* $P\{\mathcal{E}\}$ for the sample space of the event $\mathcal{E}$ by

$$P\{\mathcal{E}\} = \lim_{n \to \infty} \frac{n(\mathcal{E})}{n}$$

PROBABILITY AXIOMS

- **Axiom 1:**

$$0 \leq P\left\{\mathcal{E}\right\} \leq 1$$

the probability that the outcome of the experiment is the event $\mathcal{E}$ lies in $[0, 1]$

- **Axiom 2:**

$$P\left\{\mathcal{S}\right\} = 1$$

the probability associated with all the events in the sample space $\mathcal{S}$ is 1 as $\mathcal{S}$ is the collection of all the events of the sample space

PROBABILITY AXIOMS

- Axiom 3: For any collection of mutually exclusive events $\mathcal{E}_1, \mathcal{E}_2, \ldots$ with $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset, i \neq j$,

$$P\left\{\bigcup_i \mathcal{E}_i\right\} = \sum_i P\{\mathcal{E}_i\},$$

i.e., for a collection of mutually exclusive events, the probability that at least one of the events of the collection occurs is the sum of the probabilities of all the events in the collection

APPLICATIONS OF THE AXIOMS

- In a coin tossing experiment, we assume that a head is equally likely to appear as a tail so that:

$$P\{\{H\}\} = P\{\{T\}\} = 0.5$$

- If the coin is biased and we have the situation that the head is twice as likely to appear as the tail, then

$$P\{\{H\}\} = 0.66\dot{6} \quad \text{and} \quad P\{\{T\}\} = 0.33\dot{3}$$

EXAMPLE

- In a die tossing experiment, we assume that each of the six sides is equally likely to appear so that

$$P\{\{1\}\} = P\{\{2\}\} = P\{\{3\}\} = P\{\{4\}\} = P\{\{5\}\} = P\{\{6\}\} = 0.16\dot{6}$$

- The probability of the event that the toss results in an even number is:

$$P\{\{2,4,6\}\} = P\{\{2\}\} + P\{\{4\}\} + P\{\{6\}\} = \left(0.16\dot{6}\right)3 = 0.5$$

SIMPLE PROBABILITY THEOREMS

- The theorem on a complementary set states that the probability of the complement of the event $\mathcal{E}$ is 1 minus the probability the event itself

$$P\left\{\mathcal{E}^{c}\right\} = 1 - P\left\{\mathcal{E}\right\}$$

- For example, if the probability of obtaining an outcome $\left\{H\right\}$ on a biased coin is 0.375, then the probability of obtaining an outcome $\left\{T\right\}$ is 0.625

SIMPLE PROBABILITY THEOREMS

- The theorem on a subset considers two subsets $\mathcal{E}$ and $\mathcal{F}$ of $\mathcal{S}$ and states

$$\mathcal{E} \subset \mathcal{F} \Rightarrow P\left\{\mathcal{E}\right\} \leq P\left\{\mathcal{F}\right\}$$

- For example, the probability of rolling a 1 with a die is less than or equal to the probability of rolling an odd value with that same die
- Theorem on the union of two subsets concerns two subsets $\mathcal{E}$ and $\mathcal{F}$ of $\mathcal{S}$ and states that

$$P\left\{\mathcal{E} \cup \mathcal{F}\right\} = P\left\{\mathcal{E}\right\} + P\left\{\mathcal{F}\right\} - P\left\{\mathcal{E} \cup \mathcal{F}\right\}$$

SIMPLE PROBABILITY THEOREMS

- For example, in the experiment of tossing two fair coins

$$\mathcal{S} = \left\{ \left\{ H,H \right\}, \left\{ H,T \right\}, \left\{ T,H \right\}, \left\{ T,T \right\} \right\}$$

and the four outcomes are equally likely; the subset of the events that either the first or the second coin falls on H is the union of the subsets of events

$$\mathcal{E} = \left\{ \left\{ H,H \right\}, \left\{ H,T \right\} \right\}$$

that the first coin is H and the subset of events

SIMPLE PROBABILITY THEOREMS

$$\mathscr{F} = \left\{ \left\{ H,H \right\}, \left\{ T,H \right\} \right\}$$

represents the event second coin toss is *H*; so

$$P\left\{ \mathscr{E} \cup \mathscr{F} \right\} = P\left\{ \mathscr{E} \right\} + P\left\{ \mathscr{F} \right\} - P\left\{ \mathscr{E} \cap \mathscr{F} \right\}$$

$$= 0.5 + 0.5 - \underbrace{P\left\{ \left\{ H,H \right\} \right\}}_{0.25}$$

$$= 0.75$$

CONDITIONAL PROBABILITY

- A conditional event $\mathcal{E}$ is one that occurs given that some other event $\mathcal{F}$ has already occurred
- The *conditional probability* $P\{\mathcal{E} \mid \mathcal{F}\}$ is the probability that event $\mathcal{E}$ occurs given that event $\mathcal{F}$ has occurred and is defined by

$$P\left\{\mathcal{E} \mid \mathcal{F}\right\} = \frac{P\left\{\mathcal{E} \cap \mathcal{F}\right\}}{P\left\{\mathcal{F}\right\}}$$

CONDITIONAL PROBABILITY

- As an example, consider that a coin is flipped twice and assume that each of the events in

$$\mathcal{S} = \{ \{ H,H \}, \{ H,T \}, \{ T,H \}, \{ T,T \} \}$$

is equally likely to occur; then, $\{ H \}$ and $\{ T \}$ are equally likely to occur

- The conditional probability that both flips result in $\{ H \}$, given that the first flip is $\{ H \}$ is obtained as follows:

CONDITIONAL PROBABILITY

$$\mathcal{E} = \{\{H,H\}\}$$

$$\mathcal{F} = \{\{H,H\},\{H,T\}\}$$

$$P\{\mathcal{E} \mid \mathcal{F}\} = \frac{P\{\mathcal{E} \cap \mathcal{F}\}}{P\{\mathcal{F}\}} = \frac{\overbrace{P\{\{H,H\}\}}^{0.25}}{\underbrace{P\{\{H,H\},\{H,T\}\}}_{0.5}} = 0.5$$

CONDITIONAL PROBABILITY APPLICATION

- Bev must decide whether to select either a *French* or a *Chemistry* course
- She estimates to have probability of 0.5 to get an *A* in a *French* course and that of 0.333 in a *Chemistry* course, which she actually prefers
- She decides by flipping a fair coin and determines the probability she can get *A* in *Chemistry*:

CONDITIONAL PROBABILITY APPLICATION

- $\mathcal{C}$ is the event that she takes *Chemistry*
- $\mathcal{A}$ is the event that she receives an A in whichever course she takes
- then $P\{\mathcal{C} \cap \mathcal{A}\}$ is the probability she gets A in *Chemistry*

$$P\{\mathcal{C} \cap \mathcal{A}\} = P\{\mathcal{C}\}\ P\{\mathcal{A} | \mathcal{C}\} = (0.5)\,(0.333) = 0.166$$

BAYES' THEOREM

- Consider two subsets of events $\mathscr{E}$ and $\mathscr{F}$ in $\mathscr{S}$; then,

$$P\left\{\mathscr{E}\mid\mathscr{F}\right\}=\frac{P\left\{\mathscr{F}\mid\mathscr{E}\right\}P\left\{\mathscr{E}\right\}}{P\left\{\mathscr{F}\mid\mathscr{E}\right\}P\left\{\mathscr{E}\right\}+P\left\{\mathscr{F}\mid\mathscr{E}^{c}\right\}P\left\{\mathscr{E}^{c}\right\}}$$

- The proof of this theorem makes use of the definition of conditional probability

$$P\left\{\mathscr{E}\middle|\mathscr{F}\right\}=\frac{P\left\{\mathscr{E}\cap\mathscr{F}\right\}}{P\left\{\mathscr{F}\right\}}=\frac{P\left\{\mathscr{F}\middle|\mathscr{E}\right\}P\left\{\mathscr{E}\right\}}{P\left\{\mathscr{F}\right\}}$$

BAYES' THEOREM

and of the fact that any subset $\mathcal{F}$ is the union of

two nonintersecting subsets

$$\mathcal{F} = \left\{ \mathcal{F} \cap \mathcal{E} \right\} \cup \left\{ \mathcal{F} \cap \mathcal{E}^{\,c} \right\}$$

- These expressions are derived from the relation

$$P\left\{ \bigcup_{i} \mathcal{E}_{\,i} \right\} = \sum_{i} P\left\{ \mathcal{E}_{\,i} \right\}$$

APPLICATION OF BAYES' THEOREM TO DIAGNOSIS

- A laboratory test is 95 % effective in correctly detecting a certain disease when it is present, but the test yields a false positive result for 1 % of the healthy persons tested, *i.e.*, with probability 0.01, the test result incorrectly concludes that a healthy person has the disease
- We are given that 0.5 % of the population actually has the disease

APPLICATION OF BAYES' THEOREM TO DIAGNOSIS

- We compute the probability that a person has the disease given that his test result is positive
- *D* is the event that the tested person actually has the disease and

$$P\{D\} = 0.005$$

- *E* is the event that the test result is positive

A DIAGNOSIS EXAMPLE COMPUTATION

- We evaluate the

$$P\left\{\mathcal{D}\middle|\mathcal{E}\right\}=\frac{P\left\{\mathcal{E}\middle|\mathcal{D}\right\}P\left\{\mathcal{D}\right\}}{P\left\{\mathcal{E}\middle|\mathcal{D}\right\}P\left\{\mathcal{D}\right\}+P\left\{\mathcal{E}\middle|\mathcal{D}^{c}\right\}P\left\{\mathcal{D}^{c}\right\}}$$

$$=\frac{(0.95)\cdot(0.005)}{(0.95)\cdot(0.005)+(0.01)\cdot(0.995)}$$

$$=\ 0.323$$

MULTIPLE CHOICE EXAM APPLICATION

- In answering a question on a multiple choice test, a student either knows the answer or he guesses: the probability is p that the student knows the answer and so $(1-p)$ is the probability that he guesses; a student who guesses has a probability of $1/m$ to be correct where m is the number of multiple choice alternatives

MULTIPLE CHOICE EXAM APPLICATION

- We wish to compute the conditional probability that a student knows the answer to a question which he answered correctly
- To evaluate we define
  - $\mathcal{C}$ is the event that the student answers the question correctly
  - $\mathcal{K}$ is the event that he actually knows the answer with $P\{\mathcal{K}\} = p$

MULTIPLE CHOICE EXAM APPLICATION

$$P\left\{\mathcal{K} \mid \mathcal{C}\right\} = \frac{P\left\{\mathcal{K} \cap \mathcal{C}\right\}}{P\left\{\mathcal{C}\right\}}$$

$$= \frac{P\left\{\mathcal{C} \mid \mathcal{K}\right\} P\left\{\mathcal{K}\right\}}{P\left\{\mathcal{C} \mid \mathcal{K}\right\} P\left\{\mathcal{K}\right\} + P\left\{\mathcal{C} \mid \mathcal{K}^{c}\right\} P\left\{\mathcal{K}^{c}\right\}}$$

$$= \frac{(1)(p)}{(1)(p) + \left[(1/m)(1-p)\right]} = \frac{mp}{1+(m-1)p}$$

- If $m = 5$ and $p = 0.5$, the probability that a student knew the answer to a question he correctly answered is 5/6

CONDITIONAL PROBABILITY GENERALIZATION

- Consider three events A, B and C in the sample space $\mathcal{S}$

- We apply the conditional probability definition repeatedly to evaluate $P\left\{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}\right\}$

$$P\left\{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}\right\} = P\left\{\mathcal{A} \,\middle|\, \mathcal{B} \cap \mathcal{C}\right\} \cdot P\left\{\mathcal{B} \cap \mathcal{C}\right\}$$

$$= P\left\{\mathcal{A} \,\middle|\, \mathcal{B} \cap \mathcal{C}\right\} \cdot P\left\{\mathcal{B} \middle| \mathcal{C}\right\} \cdot P\left\{\mathcal{C}\right\}$$

CONDITIONAL PROBABILITY GENERALIZATION

- However, we also have that

$$P\{\mathcal{A} \cap \mathcal{B} \mid \mathcal{C}\} \cdot P\{\mathcal{C}\} = P\{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}\}$$

$$= P\{\mathcal{A} \mid \mathcal{B} \cap \mathcal{C}\} P\{\mathcal{B} \mid \mathcal{C}\} \cdot P\{\mathcal{C}\}$$

and therefore

$$P\{\mathcal{A} \cap \mathcal{B} \mid \mathcal{C}\} = P\{\mathcal{A} \mid \mathcal{B} \cap \mathcal{C}\} \cdot P\{\mathcal{B} \mid \mathcal{C}\}$$

INDEPENDENT EVENTS

- Two events *E* and *F* are said to be independent if and only if:

$$P\left\{\mathscr{E} \cap \mathscr{F}\right\} = \left[P\left\{\mathscr{E}\right\}\right]\left[P\left\{\mathscr{F}\right\}\right]$$

- Equivalently, *E* and *F* are independent if and only if:

$$P\left\{\mathscr{E} \,\middle|\, \mathscr{F}\right\} = P\left\{\mathscr{E}\right\}$$

- We give an example concerning picking cards from an ordinary deck of 52 playing cards

INDEPENDENT EVENTS

- $\mathcal{E}$ is the event that the selected card is an ace

- $\mathcal{F}$ is the event that the selected card is a spade

- $\mathcal{E}$ and $\mathcal{F}$ are independent since

$$P\left\{\mathcal{E} \cap \mathcal{F}\right\} = \frac{1}{52} \text{ and so } P\left\{\mathcal{E}\right\} = \frac{4}{52} \text{ and } P\left\{\mathcal{F}\right\} = \frac{13}{52}$$

INDEPENDENT EVENTS

- Two coins are flipped and all 4 distinct outcomes are assumed to be equally likely
- $\mathcal{E}$ is the event that the first coin is H and $\mathcal{F}$ is the event that the second coin is T
- Then, $\mathcal{E}$ and $\mathcal{F}$ are independent events with

$$P\{\mathcal{E}\} = P\{\{H,H\},\{H,T\}\} = 0.5$$

$$P\{\mathcal{F}\} = P\{\{H,T\},\{T,T\}\} = 0.5$$

and

$$P\{\mathcal{E} \cap \mathcal{F}\} = P\{\{H,T\}\} = (0.5)(0.5) = 0.25$$

PROBABILITY DISTRIBUTIONS

- **A *probability distribution* describes mathematically the set of probabilities associated with each possible outcome of a random variable (*r.v.*)**
- **A *discrete probability distribution* is a distribution characterized by a random variable that can assume a *finite* set of possible values**
- **A *continuous probability distribution* is a distribution characterized by a random variable that can assume infinitely many values**

DISCRETE PROBABILITY DISTRIBUTIONS

- ***Discrete probability distribution specification***: the probability distribution of a discrete *r.v.* $\underset{\sim}{Y}$ with n discrete possible values may be expressed in terms of either a
  - a ***probability mass function*** that provides the list of the probabilities for each possible outcome

DISCRETE PROBABILITY DISTRIBUTIONS

$$P\{\ \underset{\sim}{Y} = y_i \},\quad i = 1, 2, \ldots, n\ ;$$

or,

- a ***cumulative distribution function (c.d.f.)*** that gives the probability that a *r.v.* is less than or equal to a specific value

$$P\{\ \underset{\sim}{Y} \leq y_i \},\quad i = 1, 2, \ldots, n$$

DISCRETE PROBABILITY DISTRIBUTIONS

- As an example consider a set of chocolate chip cookies with at most 5 chips
- Assume that the probability that one of them has *0*, 1, 2, 3, 4 or 5 chips is 0.02, 0.05, 0.2, 0.4, 0.22, and 0.11, respectively
- The *probability mass function* of the *r.v.* $\underset{\sim}{Y}$, defined to be the random number of chips on a cookie, can be given either in tableau format or as a graph

DISCRETE PROBABILITY DISTRIBUTIONS

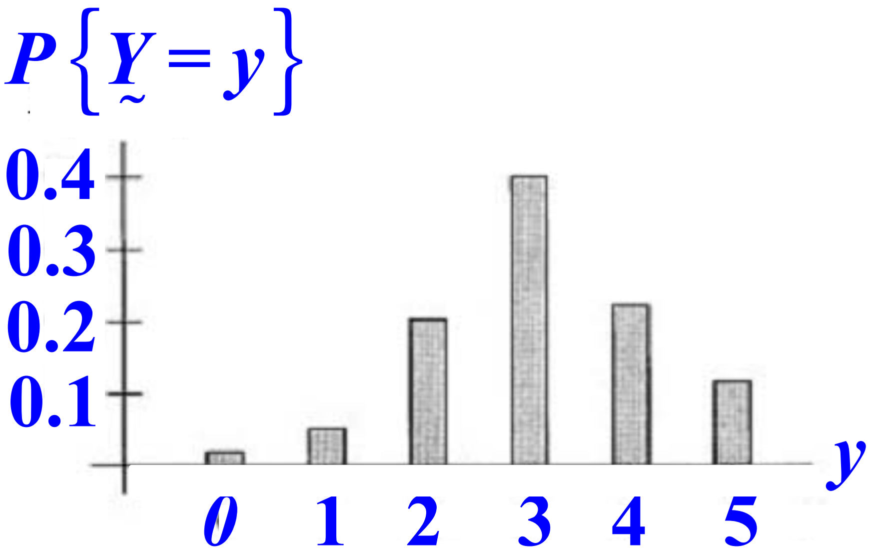
probability mass function → column 2

cumulative distribution function (c.d.f.) → column 3

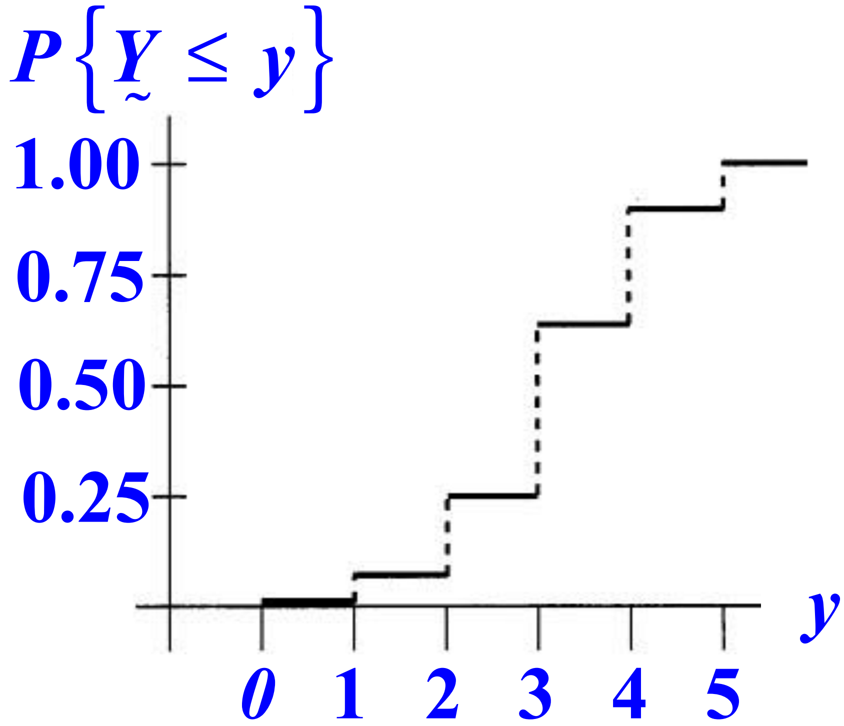
y	$P\{\underset{\sim}{Y} = y\}$	$P\{\underset{\sim}{Y} \leq y\}$
0	0.02	0.02
1	0.05	0.07
2	0.20	0.27
3	0.4	0.67
4	0.22	0.89
5	0.11	1.00

DISCRETE PROBABILITY DISTRIBUTIONS

probability mass function



cumulative distribution function (c.d.f.)



THE EXPECTED VALUE

- The expected value $E\{\underset{\sim}{X}\}$ of the random variable $\underset{\sim}{X}$ is the probability–weighted average of all its possible values: for the set of possible values $\{x_1, x_2, \ldots, x_n\}$ for the variable $\underset{\sim}{X}$

$$\mu_{\underset{\sim}{X}} = E\{\underset{\sim}{X}\} = \sum_{i=1}^{n} x_i \, P\{\underset{\sim}{X} = x_i\}$$

- The expectation operator $E\{\cdot\}$ is also defined for any function $f(\cdot)$ of the *r.v.* $\underset{\sim}{X}$

THE EXPECTED VALUE

- Let

$$\underset{\sim}{Y} = f\left(\underset{\sim}{X}\right)$$

then

$$E\left\{\underset{\sim}{Y}\right\} = E\left\{f\left(\underset{\sim}{X}\right)\right\}$$

- In general , for an arbitrary function f

$$E\left\{f\left(\underset{\sim}{X}\right)\right\} \neq f\left(E\left\{\underset{\sim}{X}\right\}\right)$$

THE EXPECTED VALUE

- If $f\{\underset{\sim}{X}\}$ is affine, then,

$$E\left\{f\left(\underset{\sim}{X}\right)\right\} = f\left(E\left\{\underset{\sim}{X}\right\}\right)$$

and we have some special cases:

  - for $\underset{\sim}{Y} = a + b\,\underset{\sim}{X}$, we have

$$E\left\{\underset{\sim}{Y}\right\} = a + bE\left\{\underset{\sim}{X}\right\}$$

  - for $\underset{\sim}{Y} = \underset{\sim}{X}_1 + \ldots + \underset{\sim}{X}_n$, we have

$$E\left\{\underset{\sim}{Y}\right\} = E\left\{\underset{\sim}{X}_1\right\} + \ldots + E\left\{\underset{\sim}{X}_n\right\}$$

THE VARIANCE

- The *variance* $var\{\underset{\sim}{X}\}$ of the random variable $\underset{\sim}{X}$ is the expected value of the squared difference between the uncertain quantities and their expected value $E\{\underset{\sim}{X}\}$:

$$var\{\underset{\sim}{X}\} \triangleq E\left\{\left[\underset{\sim}{X} - E\{\underset{\sim}{X}\}\right]^2\right\} = \sum_{i=1}^{n}\left(x_i - \mu_{\underset{\sim}{X}}\right)^2 P\{\underset{\sim}{X} = x_i\}$$

THE VARIANCE

- for $\underset{\sim}{Y} = a + b\underset{\sim}{X}$

$$
\begin{aligned}
var\{\underset{\sim}{Y}\} &= var\{a + b\underset{\sim}{X}\} \\
&= E\left\{\left[\left(a + b\underset{\sim}{X}\right) - \left(a + bE\{\underset{\sim}{X}\}\right)\right]^2\right\} \\
&= E\left\{\left[b\underset{\sim}{X} - bE\{\underset{\sim}{X}\}\right]^2\right\} \\
&= \left(b^2\right)\underbrace{E\left\{\left[\underset{\sim}{X} - E\{\underset{\sim}{X}\}\right]^2\right\}}_{var\{\underset{\sim}{X}\}} \\
&= \left(b^2\right)var\{\underset{\sim}{X}\}
\end{aligned}
$$

THE VARIANCE

- for

$$\underset{\sim}{Y} = \underset{\sim}{X}_1 + \ \ldots \ + \underset{\sim}{X}_n \text{ and } P\left\{\underset{\sim}{X}_i \,\middle|\, \underset{\sim}{X}_j\right\} = P\left\{\underset{\sim}{X}_i\right\} \ \forall \ i \neq j$$

  then

$$var\left\{\underset{\sim}{Y}\right\} = var\left\{\underset{\sim}{X}_1\right\} + \ \ldots \ + var\left\{\underset{\sim}{X}_n\right\}$$

- The standard deviation $\sigma_{\underset{\sim}{X}}$ is given by

$$\sigma_{\underset{\sim}{X}} = \sqrt{var\left\{\underset{\sim}{X}\right\}}$$

COVARIANCE AND CORRELATION COEFFICIENT

- The *covariance* $cov\{\underset{\sim}{X}, \underset{\sim}{Y}\}$ is defined by

$$cov\{\underset{\sim}{X},\underset{\sim}{Y}\} \triangleq E\{(\underset{\sim}{X} - E\{\underset{\sim}{X}\})(\underset{\sim}{Y} - E\{\underset{\sim}{Y}\})\}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m}\left[x_i - E\{\underset{\sim}{X}\}\right]\left[y_j - E\{\underset{\sim}{Y}\}\right] P\{\underset{\sim}{X} = x_i, \underset{\sim}{Y} = y_j\}$$

- The *correlation* $\rho_{\underset{\sim}{X}\underset{\sim}{Y}}$ is defined by

$$\rho_{\underset{\sim}{X}\underset{\sim}{Y}} = \frac{cov\{\underset{\sim}{X},\underset{\sim}{Y}\}}{\sigma_{\underset{\sim}{X}}\sigma_{\underset{\sim}{Y}}}$$

APPLICATION EXAMPLE

- A company is selling a product G with different net profits corresponding to different levels of product sales

level of sales	*probability*	*net profits* [*M $*]
high	0.38	8
medium	0.12	4
low	0.50	*0*

- The standard deviation and variance of the net profits $\underset{\sim}{X}$ for the product are given by

APPLICATION EXAMPLE

$$E\{\underset{\sim}{X}\} = \sum_{i=1}^{n} x_i \, P\{\underset{\sim}{X} = x_i\} = 8(0.38) + 4(0.12) + 0(0.50)$$

$$= 3.52 \ M\$$$

$$var\{\underset{\sim}{X}\} = \sum_{i=1}^{n} \left[x_i - E\{\underset{\sim}{X}\}\right]^2 P\{\underset{\sim}{X} = x_i\}$$

$$= 0.38(8-3.52)^2 + 0.12(4-3.52)^2 + 0.5(0-3.52)^2$$

$$= 13.8496 \ (M\$)^2$$

$$\sigma_{\underset{\sim}{X}} = \sqrt{var\{\underset{\sim}{X}\}} = \sqrt{13.8496} = 3.72 \ M\$$$

ANOTHER EXAMPLE

- Consider the following probabilities:

$$P\{\underset{\sim}{Y}=10\,|\,\underset{\sim}{X}=2\}=0.9$$

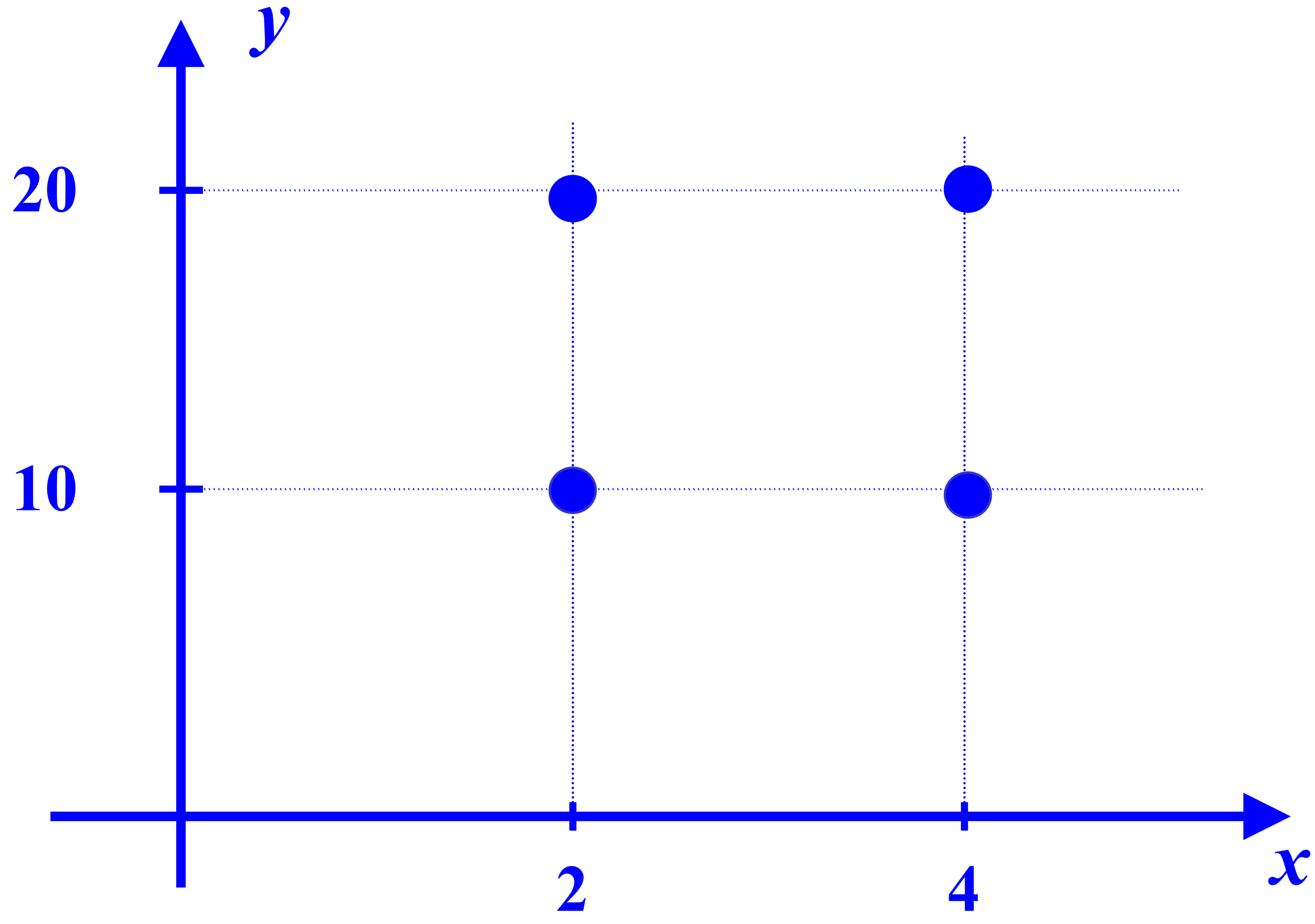
$$P\{\underset{\sim}{X}=2\}=0.3 \qquad P\{\underset{\sim}{Y}=20\,|\,\underset{\sim}{X}=2\}=0.1$$

$$P\{\underset{\sim}{X}=4\}=0.7 \qquad P\{\underset{\sim}{Y}=10\,|\,\underset{\sim}{X}=4\}=0.25$$

$$P\{\underset{\sim}{Y}=20\,|\,\underset{\sim}{X}=4\}=0.75$$

and compute the covariance and correlation between $\underset{\sim}{X}$ and $\underset{\sim}{Y}$

ANOTHER EXAMPLE



ANOTHER EXAMPLE

- Using the definition of conditional probability:

$$P\{\underset{\sim}{X}=2,\underset{\sim}{Y}=10\} = P\{\underset{\sim}{Y}=10 | \underset{\sim}{X}=2\} P\{\underset{\sim}{X}=2\}$$

$$= (0.9)(0.3) = 0.27$$

$$P\{\underset{\sim}{X}=2,\underset{\sim}{Y}=20\} = P\{\underset{\sim}{Y}=20 | \underset{\sim}{X}=2\} P\{\underset{\sim}{X}=2\}$$

$$= (0.1)(0.3) = 0.03$$

$$P\{\underset{\sim}{X}=4,\underset{\sim}{Y}=10\} = P\{\underset{\sim}{Y}=10 | \underset{\sim}{X}=4\} P\{\underset{\sim}{X}=4\}$$

$$= (0.25)(0.7) = 0.175$$

$$P\{\underset{\sim}{X}=4,\underset{\sim}{Y}=20\} = P\{\underset{\sim}{Y}=20 | \underset{\sim}{X}=4\} P\{\underset{\sim}{X}=4\}$$

$$= (0.75)(0.7) = 0.525$$

ANOTHER EXAMPLE

$$P\{\underset{\sim}{Y}=10\} = P\{\underset{\sim}{Y}=10 \mid \underset{\sim}{X}=2\}P\{\underset{\sim}{X}=2\} +$$

$$P\{\underset{\sim}{Y}=10 \mid \underset{\sim}{X}=4\}P\{\underset{\sim}{X}=4\}$$

$$= 0.27 + 0.175 = 0.445$$

$$P\{\underset{\sim}{Y}=20\} = 1-(0.445) = 0.555$$

$$E\{\underset{\sim}{X}\} = (0.3)2 + (0.7)4 = 3.4$$

$$\sigma_{\underset{\sim}{X}} = \sqrt{(0.3)(-1.4)^2 + (0.7)(0.6)^2} = 0.917$$

$$E\{\underset{\sim}{Y}\} = (0.445)10 + (0.555)20 = 15.55$$

$$\sigma_{\underset{\sim}{Y}} = \sqrt{(0.445)(-4.45)^2 + (0.555)(14.45)^2} = 11.17$$

EXAMPLE

x_i	y_j	$x_i - E\{\underset{\sim}{X}\}$	$y_j - E\{\underset{\sim}{Y}\}$	$\left[x_i - E\{\underset{\sim}{X}\}\right] \cdot \left[y_j - E\{\underset{\sim}{Y}\}\right]$	$P\left\{\underset{\sim}{X},\underset{\sim}{Y}\Big\|_{x_i,y_i}\right\}$
2	10	-1.4	4.45	$-$ 6.23	0.27
2	20	-1.4	14.45	-20.23	0.03
4	10	0.6	4.45	2.67	0.175
4	20	0.6	14.45	8.67	0.525

EXAMPLE

$$cov\{\underset{\sim}{X},\underset{\sim}{Y}\} = (0.27)(-6.23)+(0.03)(-20.23)+(0.175)2.67$$

$$= 2.73$$

$$\rho_{\underset{\sim}{X}\underset{\sim}{Y}} = \frac{cov\{\underset{\sim}{X},\underset{\sim}{Y}\}}{\sigma_{\underset{\sim}{X}}\sigma_{\underset{\sim}{Y}}} = \frac{2.73}{(0.917)(4.970)} = 0.60$$

CONTINUOUS PROBABILITY DISTRIBUTIONS

- The *continuous probability distribution* specification of a continuous *r.v.* $\underset{\sim}{X}$ may be expressed either in terms of a
  - a *probability density function (p.d.f.)* $f_{\underset{\sim}{X}}(\cdot)$

$$f_{\underset{\sim}{X}}(x)\,dx \approx P\left\{x < \underset{\sim}{X} \leq x + dx\right\}$$

  - or, a *cumulative distribution function (c.d.f.)* $F_{\underset{\sim}{X}}(\cdot)$ which expresses the probability that the value of $\underset{\sim}{X}$ is less or equal to a given value x

$$F_{\underset{\sim}{X}}(x) = P\left\{\underset{\sim}{X} \leq x\right\} = \int_{-\infty}^{x} f_{\underset{\sim}{X}}(\xi)\,d\xi$$

EXPECTED VALUE, VARIANCE, STANDARD DEVIATION

- The *expected value* $\mu_{\underset{\sim}{X}}$ is given by

$$E\{\underset{\sim}{X}\} = \int_{-\infty}^{+\infty} \xi \, f_{\underset{\sim}{X}}(\xi)\, d\xi$$

- The *variance* $var\{\underset{\sim}{X}\}$ of $\underset{\sim}{X}$ is defined by

$$var\{\underset{\sim}{X}\} = \int_{-\infty}^{+\infty} \left[\xi - E\{\underset{\sim}{X}\}\right]^2 f_{\underset{\sim}{X}}(\xi)\, d\xi$$

- The *standard deviation* $\sigma_{\underset{\sim}{X}}$ of $\underset{\sim}{X}$ is

$$\sigma_{\underset{\sim}{X}} = \sqrt{var\{\underset{\sim}{X}\}}$$

THE COVARIANCE AND THE CORRELATION

- The *covariance* $cov\{\underset{\sim}{X},\underset{\sim}{Y}\}$ of the two continuous *r.v.*s $\underset{\sim}{X}$ and $\underset{\sim}{Y}$

$$cov\{\underset{\sim}{X},\underset{\sim}{Y}\}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\xi-E\{\underset{\sim}{X}\}\right]\left[\eta-E\{\underset{\sim}{Y}\}\right]f_{\underset{\sim}{X},\underset{\sim}{Y}}(\xi,\eta)\,d\xi\,d\eta$$

  where $f_{\underset{\sim}{X},\underset{\sim}{Y}}(\bullet,\bullet)$ is the joint density function of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$

- The *correlation coefficient* $\rho_{\underset{\sim}{X},\underset{\sim}{Y}}$ is computed by

$$\rho_{\underset{\sim}{X},\underset{\sim}{Y}}=\frac{cov\{\underset{\sim}{X},\underset{\sim}{Y}\}}{\sigma_{\underset{\sim}{X}}\sigma_{\underset{\sim}{Y}}}$$

APPLICATION

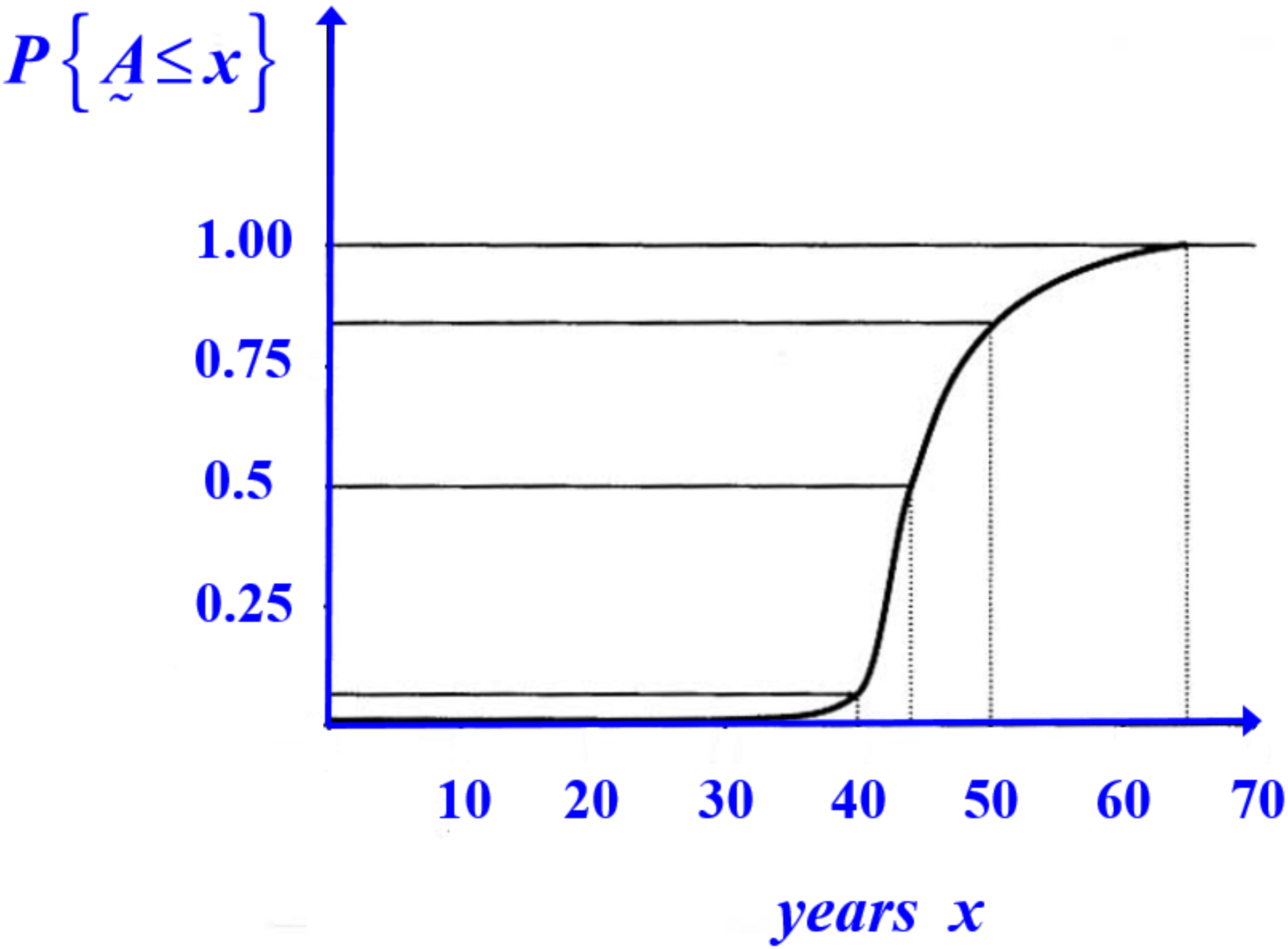
- We wish to guess the age $\underset{\sim}{A}$ of a movie star based on the following data:
  - we are sure that she is older than 29 and not older than 65
  - we assume the probability that she is between 40 and 50 is 0.8 and $P\left\{\underset{\sim}{A} > 50\right\} = 0.15$
  - we also estimate that $P\left\{\underset{\sim}{A} \leq 40\right\} = 0.05$ and $P\left\{\underset{\sim}{A} \leq 44\right\} = P\left\{\underset{\sim}{A} > 44\right\}$

APPLICATION

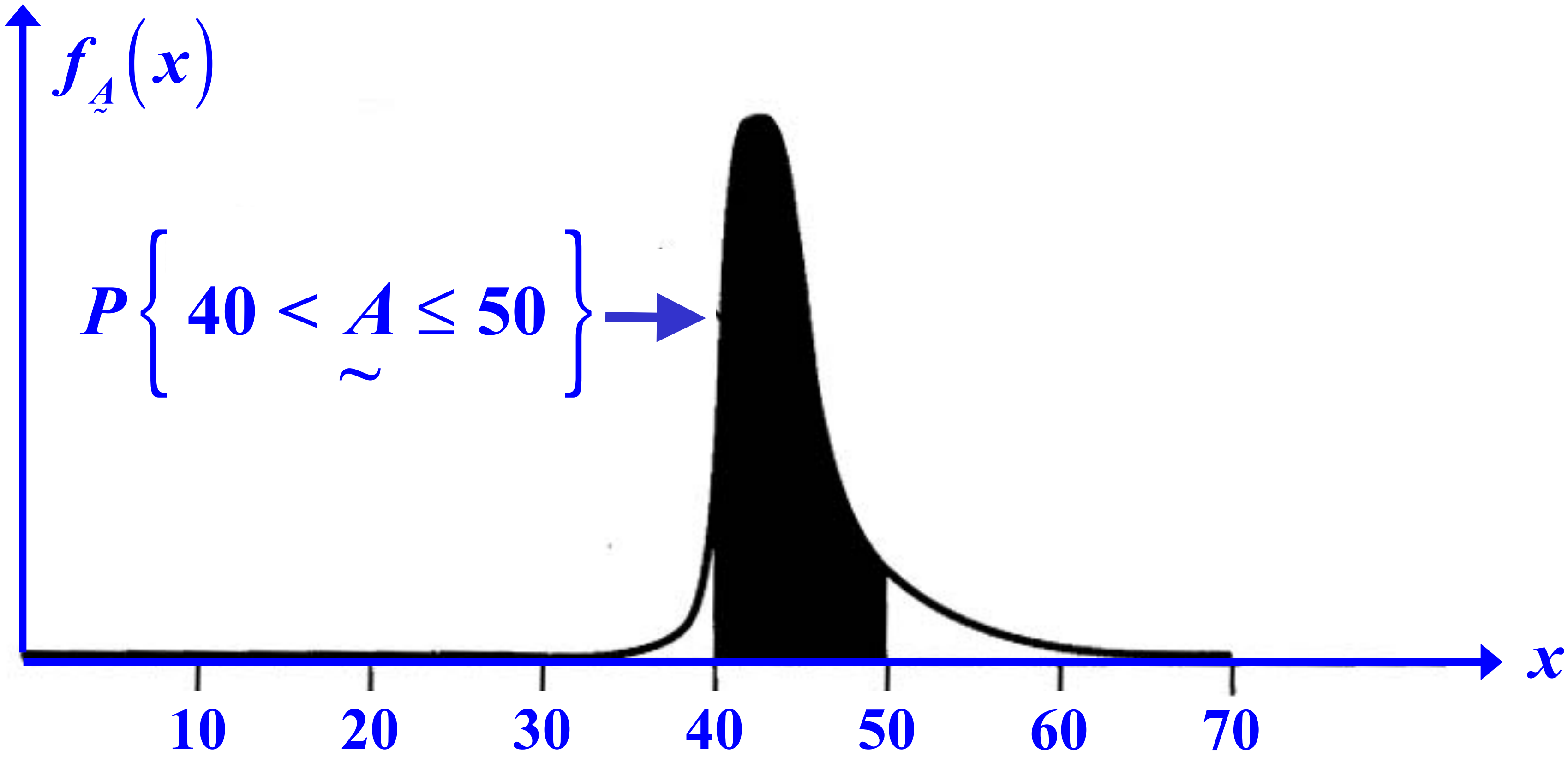
- We construct the table of cumulative probability

$P\{\underset{\sim}{A} \leq 29\} = 0.00$
$P\{\underset{\sim}{A} \leq 40\} = 0.05$
$P\{\underset{\sim}{A} \leq 44\} = 0.50$
$P\{\underset{\sim}{A} \leq 50\} = 0.85$
$P\{\underset{\sim}{A} \leq 65\} = 1.00$

APPLICATION



APPLICATION



years x