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 $S$ is the set of all possible outcomes, i.e., every outcome is an element of $S$. 
SAMPLE SPACE

- Examples of sample spaces

  - Flipping a coin: \( S = \{ H, T \} \)
  - Tossing a die: \( S = \{ 1, 2, 3, 4, 5, 6 \} \)
  - Flipping two coins: \( S = \{ (H, H), (H, T), (T, H), (T, T) \} \)
  - Tossing two dice: \( S = \{ (i, j) : i, j = 1, \ldots, 6 \} \)
  - Hours of life of a device: \( S = \{ x : 0 \leq x < \infty \} \)
We say a set $E$ is a subset of a set $F$ if $E$ is contained in $F$ and we write $E \subseteq F$ or $F \supseteq E$.

If $E$ and $F$ are sets of events, then $E \subseteq F$ implies that each event in $E$ is also an event in $F$.

Theorem

$E \subseteq F$ and $F \supseteq E$ $\iff$ $E = F$
SUBSETS

$F \subseteq E$

© 2006 – 2019 George Gross, University of Illinois at Urbana-Champaign, All Rights Reserved.
EVENTS

- An event $E$ is an element or a subset of the sample space $S$.

- Some examples of events are:
  - flipping a coin: $E = \{H\}$, $F = \{T\}$
  - tossing a die: $E = \{2, 4, 6\}$ 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
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

$E \cup F$
UNION OF SUBSETS

Examples:

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

$\Rightarrow \mathcal{E} \cup \mathcal{F} = \mathcal{S}$
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 \( \emptyset \) to be the *empty* set, i.e., the set consisting of no elements.

- For event subspaces \( E \) and \( F \), if \( E \cap F = \emptyset \) if and only if \( E \) and \( F \) are *mutually exclusive* events.

- Examples:
  - \( E = \{ H \}, \mathcal{F} = \{ T \} \Rightarrow E \cap \mathcal{F} = \emptyset \)
  - \( E = \{ 1, 3, 5 \}, \mathcal{F} = \{ 1, 2, 3 \} \Rightarrow E \cap \mathcal{F} = \{ 1, 3 \} \)
INTERSECTION OF SUBSETS

\[ E \cap F \]
We consider the countable subsets $E_1, E_2, E_3, \ldots$ in the state space $S$.

The term $\bigcup_i E_i$ is defined to be that subset consisting of those elements that are in $E_i$ for at least one value of $i = 1, 2, \ldots$.

The term $\bigcap_i \tilde{E}_i$ is defined to be the subset consisting of those elements that are in every subset $E_i$, $i = 1, 2, \ldots$. 
COMPLEMENT OF A SUBSET

- The complement $E^c$ of a set $E$ is the set of all elements in the sample space $S$ not in $E$.

- By definition, $S^c = \emptyset$.

- 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, $E^c$ is the collection of events that the sum of dice is not 7.
COMPLEMENT OF A SUBSET

$S$

$E$

$E^c$
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 $E$ in the sample space $S$ and let us denote by $n(E)$ the number of times that the event $E$ occurs in a total of $n$ random draws.

We define the probability $P\{E\}$ for the sample space of the event $E$ by

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

- **Axiom 1:**

  \[ 0 \leq P(\mathcal{E}) \leq 1 \]

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

- **Axiom 2:**

  \[ P(\mathcal{S}) = 1 \]

  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.
Axiom 3: For any collection of mutually exclusive events $E_1, E_2, \ldots$ with $E_i \cap E_j = \emptyset$, 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.

$$P\left\{ \bigcup_{i} E_i \right\} = \sum_{i} P\left\{ E_i \right\},$$
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

\[ P\{\{H\}\} = 0.666 \text{ and } P\{\{T\}\} = 0.333 \]
EXAMPLE

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

\[ P\{\{1\}\} = P\{\{2\}\} = P\{\{3\}\} = P\{\{4\}\} = P\{\{5\}\} = P\{\{6\}\} = 0.166 \]

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

\[ P\{\{2,4,6\}\} = P\{\{2\}\} + P\{\{4\}\} + P\{\{6\}\} = \left(0.166\right)3 = 0.5 \]
The theorem on a complementary set states that the probability of the complement of the event $E$ is 1 minus the probability the event itself.

$$P\{E^c\} = 1 - P\{E\}$$

For example, if the probability of obtaining an outcome $\{H\}$ on a biased coin is 0.375, then the probability of obtaining an outcome $\{T\}$ is 0.625.
The theorem on a subset considers two subsets $E$ and $F$ of $S$ and states

$$E \subset F \Rightarrow P\{E\} \leq P\{F\}$$

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 $E$ and $F$ of $S$ and states that

$$P\{E \cup F\} = P\{E\} + P\{F\} - P\{E \cup F\}$$
For example, in the experiment of tossing two fair coins

\[\mathcal{S} = \left\{ \{H, H\}, \{H, T\}, \{T, H\}, \{T, T\} \right\}\]

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

\[\mathcal{E} = \left\{ \{H, H\}, \{H, T\} \right\}\]

giving that the first coin is \(H\) and the subset of events...
SIMPLE PROBABILITY THEOREMS

\[ \mathcal{F} = \left\{ \{H,H\}, \{T,H\} \right\} \]

represents the event second coin toss is \( H \); so

\[
P\{\mathcal{E} \cup \mathcal{F}\} = P\{\mathcal{E}\} + P\{\mathcal{F}\} - P\{\mathcal{E}\cap \mathcal{F}\}
\]

\[
= 0.5 + 0.5 - P\{\{H,H\}\}
\]

\[
= 0.75
\]

\[
= 0.25
\]
A conditional event $E$ is one that occurs given that some other event $F$ has already occurred.

The conditional probability $P\{E \mid F\}$ is the probability that event $E$ occurs given that event $F$ has occurred and is defined by

$$P\{E \mid F\} = \frac{P\{E \cap F\}}{P\{F\}}$$
As an example, consider that a coin is flipped twice and assume that each of the events in
\[ S = \{ \{ H, H \}, \{ H, T \}, \{ T, H \}, \{ T, T \} \} \]
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:
\[ \mathcal{E} = \{ \{H,H\}\} \]

\[ \mathcal{F} = \{ \{H,H\}, \{H,T\}\} \]

\[
P\{\mathcal{E} | \mathcal{F}\} = \frac{P\{\mathcal{E} \cap \mathcal{F}\}}{P\{\mathcal{F}\}} = \frac{0.25}{P\{\{H,H\}\} \text{\quad 0.25}} = \frac{P\{\{H,H\}\}}{P\{\{H,H\}, \{H,T\}\} \text{\quad 0.5}} = 0.5 \]

\[
= \frac{0.25}{0.5} = 0.5
\]
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

○ $C$ is the event that she takes Chemistry

○ $A$ is the event that she receives an $A$ in whichever course she takes

○ then $P\{C \cap A\}$ is the probability she gets $A$ in Chemistry

$$P\{C \cap A\} = P\{C\} \cdot P\{A|C\} = (0.5)(0.333) = 0.166$$
BAYES’ THEOREM

1. Consider two subsets of events $E$ and $F$ in $S$; then,

\[
P\left\{ E \mid F \right\} = \frac{P\left\{ F \mid E \right\} P\left\{ E \right\}}{P\left\{ F \mid E \right\} P\left\{ E \right\} + P\left\{ F \mid E^c \right\} P\left\{ E^c \right\}}
\]

2. The proof of this theorem makes use of the definition of conditional probability

\[
P\left\{ E \mid F \right\} = \frac{P\left\{ E \cap F \right\}}{P\left\{ F \right\}} = \frac{P\left\{ F \mid E \right\} P\left\{ E \right\}}{P\left\{ F \right\}}
\]
BAYES’ THEOREM

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

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

These expressions are derived from the relation

$$P\left\{ \bigcup_{i} \mathcal{E}_i \right\} = \sum_{i} P\{ \mathcal{E}_i \}$$
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.
We evaluate the

\[
P\left\{ D \mid E \right\} = \frac{P\left\{ E \mid D \right\} P\left\{ D \right\}}{P\left\{ E \mid D \right\} P\left\{ D \right\} + P\left\{ E \mid D^c \right\} P\left\{ D^c \right\}}
\]

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

\[
= 0.323
\]
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

- $C$ is the event that the student answers the question correctly.
- $K$ is the event that he actually knows the answer with $P \{ K \} = p$.
If $m = 5$ and $p = 0.5$, the probability that a student knew the answer to a question he correctly answered is $\frac{5}{6}$.
Consider three events $A$, $B$ and $C$ in the sample space $S$

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

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

$$= P\{A \mid B \cap C\} \cdot P\{B \mid C\} \cdot P\{C\}$$
However, we also have that

\[ P\{A \cap B \mid C\} \cdot P\{C\} = P\{A \cap B \cap C\} \]

\[ = P\{A \mid B \cap C\} P\{B \mid C\} \cdot P\{C\} \]

and therefore

\[ P\{A \cap B \mid C\} = P\{A \mid B \cap C\} \cdot P\{B \mid C\} \]
INDEPENDENT EVENTS

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

\[ P\{E \cap F\} = P\{E\} P\{F\} \]

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

\[ P\{E|F\} = P\{E\} \]

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

- $E$ is the event that the selected card is an ace

- $F$ is the event that the selected card is a spade

- $E$ and $F$ are independent since

\[ P\{E \cap F\} = \frac{1}{52} \quad \text{and so} \quad P\{E\} = \frac{4}{52} \quad \text{and} \quad P\{F\} = \frac{13}{52} \]
INDEPENDENT EVENTS

Two coins are flipped and all 4 distinct outcomes are assumed to be equally likely.

\( E \) is the event that the first coin is \( H \) and \( F \) is the event that the second coin is \( T \).

Then, \( E \) and \( F \) are independent events with

\[
P\{E\} = P\left\{\{H,H\},\{H,T\}\right\} = 0.5
\]

\[
P\{F\} = P\left\{\{H,T\},\{T,T\}\right\} = 0.5
\]

and

\[
P\{E \cap F\} = P\left\{\{H,T\}\right\} = (0.5)(0.5) = 0.25
\]
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 distribution specification: the probability distribution of a discrete r.v. \( Z \) 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

\[ P \{ Y = y_i \}, \quad i = 1, 2, \ldots, n; \]

or,

\[ P \{ Y \leq y_i \}, \quad i = 1, 2, \ldots, n \]

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

\[ P \{ Y \leq y_i \}, \quad i = 1, 2, \ldots, n \]
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. $Y$, 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**

<table>
<thead>
<tr>
<th>$y$</th>
<th>$P{Y = y}$</th>
<th>$P{Y \leq y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>1</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
<td>0.27</td>
</tr>
<tr>
<td>3</td>
<td>0.4</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>0.22</td>
<td>0.89</td>
</tr>
<tr>
<td>5</td>
<td>0.11</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**cumulative distribution function (c.d.f.)**
**DISCRETE PROBABILITY DISTRIBUTIONS**

**probability mass function**

\[ P \{ Y = y \} \]

- 0.1
- 0.2
- 0.3
- 0.4

<table>
<thead>
<tr>
<th>( y )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P { Y = y } )</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**cumulative distribution function (c.d.f.)**

\[ P \{ Y \leq y \} \]

- 0.00
- 0.25
- 0.50
- 0.75
- 1.00

<table>
<thead>
<tr>
<th>( y )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P { Y \leq y } )</td>
<td>0.0</td>
<td>0.25</td>
<td>0.50</td>
<td>0.75</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
The expected value $E\{X\}$ of the random variable $X$ 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 $X$

$$\mu_X = E\{X\} = \sum_{i=1}^{n} x_i \ P\{X = x_i\}$$

The expectation operator $E\{\cdot\}$ is also defined for any function $f(\cdot)$ of the r.v. $X$. 
THE EXPECTED VALUE

Let

\[ Y = f(X) \]

then

\[ E\{Y\} = E\{f(X)\} \]

In general, for an arbitrary function \( f \)

\[ E\{f(X)\} \neq f(E\{X\}) \]
THE EXPECTED VALUE

☐ If $f\{X\}$ is affine, then,

$$E\{f(X)\} = f(E\{X\})$$

and we have some special cases:

☐ for $Y = a + bX$, we have

$$E\{Y\} = a + bE\{X\}$$

☐ for $Y = X_1 + \ldots + X_n$, we have

$$E\{Y\} = E\{X_1\} + \ldots + E\{X_n\}$$
THE VARIANCE

The variance \( \text{var}\{X\} \) of the random variable \( X \) is the expected value of the squared difference between the uncertain quantities and their expected value \( E\{X\} \):

\[
\text{var}\{X\} \triangleq E\left\{ \left[ X - E\{X\} \right]^2 \right\} = \sum_{i=1}^{n} \left( x_i - \mu_{X} \right)^2 P\{X = x_i\}
\]
for $Y = a + bX$

\[
\text{var}\{Y\} = \text{var}\{a + bX\}
\]

\[
= E \left\{ \left[ (a + b\tilde{X}) - (a + bE\{\tilde{X}\}) \right] \right\}^2
\]

\[
= E \left\{ \left[ b\tilde{X} - bE\{\tilde{X}\} \right] \right\}^2
\]

\[
= \left( b^2 \right) E \left\{ \left[ \tilde{X} - E\{\tilde{X}\} \right] \right\}^2
\]

\[
= \left( b^2 \right) \text{var}\{\tilde{X}\}
\]

\[
= \left( b^2 \right) \text{var}\{\tilde{X}\}
\]
THE VARIANCE

for

\[ Y = X_1 + \ldots + X_n \quad \text{and} \quad P \{ X_i \mid X_j \} = P \{ X_i \} \quad \forall \ i \neq j \]

then

\[ \text{var} \{ Y \} = \text{var} \{ X_1 \} + \ldots + \text{var} \{ X_n \} \]

The standard deviation \( \sigma_{\tilde{X}} \) is given by

\[ \sigma_{\tilde{X}} = \sqrt{\text{var} \{ \tilde{X} \}} \]
COVARIANCE AND CORRELATION COEFFICIENT

- The covariance \( \text{cov}\{X, Y\} \) is defined by

\[
\text{cov}\{X,Y\} \triangleq E \left\{ (X - E\{X\})(Y - E\{Y\}) \right\}
\]

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

- The correlation \( \rho_{\tilde{X}\tilde{Y}} \) is defined by

\[
\rho_{\tilde{X}\tilde{Y}} = \frac{\text{cov}\{X,Y\}}{\sigma_X \sigma_Y}
\]
A company is selling a product $G$ with different net profits corresponding to different levels of product sales.

<table>
<thead>
<tr>
<th>level of sales</th>
<th>probability</th>
<th>net profits [M $]</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>0.38</td>
<td>8</td>
</tr>
<tr>
<td>medium</td>
<td>0.12</td>
<td>4</td>
</tr>
<tr>
<td>low</td>
<td>0.50</td>
<td>0</td>
</tr>
</tbody>
</table>

The standard deviation and variance of the net profits $X$ for the product are given by.
**APPLICATION EXAMPLE**

\[ E\{X\} = \sum_{i=1}^{n} x_i P\{X = x_i\} = 8(0.38) + 4(0.12) + 0(0.50) \]

\[ = 3.52 \text{ M$\$$} \]

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

\[ = 0.38(8 - 3.52)^2 + 0.12(4 - 3.52)^2 + 0.5(0 - 3.52)^2 \]

\[ = 13.8496 \text{ (M$\$$)^2} \]

\[ \sigma_X = \sqrt{var\{X\}} = \sqrt{13.8496} = 3.72 \text{ M$\$$} \]
Consider the following probabilities:

\[ P \{ Y = 10 \mid X = 2 \} = 0.9 \]

\[ P \{ X = 2 \} = 0.3 \quad P \{ Y = 20 \mid X = 2 \} = 0.1 \]

\[ P \{ X = 4 \} = 0.7 \quad P \{ Y = 10 \mid X = 4 \} = 0.25 \]

\[ P \{ Y = 20 \mid X = 4 \} = 0.75 \]

and compute the covariance and correlation between \( X \) and \( Y \).
ANOTHER EXAMPLE

\[ y \]

\[ x \]

\[ 20 \]

\[ 10 \]

\[ 2 \]

\[ 4 \]
ANOTHER EXAMPLE

- Using the definition of conditional probability:

\[
P \{ \tilde{X} = 2, \tilde{Y} = 10 \} = P \{ \tilde{Y} = 10 | \tilde{X} = 2 \} P \{ \tilde{X} = 2 \}
= (0.9)(0.3) = 0.27
\]

\[
P \{ \tilde{X} = 2, \tilde{Y} = 20 \} = P \{ \tilde{Y} = 20 | \tilde{X} = 2 \} P \{ \tilde{X} = 2 \}
= (0.1)(0.3) = 0.03
\]

\[
P \{ \tilde{X} = 4, \tilde{Y} = 10 \} = P \{ \tilde{Y} = 10 | \tilde{X} = 4 \} P \{ \tilde{X} = 4 \}
= (0.25)(0.7) = 0.175
\]

\[
P \{ \tilde{X} = 4, \tilde{Y} = 20 \} = P \{ \tilde{Y} = 20 | \tilde{X} = 4 \} P \{ \tilde{X} = 4 \}
= (0.75)(0.7) = 0.525
\]
ANOTHER EXAMPLE

\[ P \{ Y = 10 \} = P \{ Y = 10 \mid X = 2 \} P \{ X = 2 \} + P \{ Y = 10 \mid X = 4 \} P \{ X = 4 \} \]

\[ = 0.27 + 0.175 = 0.445 \]

\[ P \{ Y = 20 \} = 1 - (0.445) = 0.555 \]

\[ E \{ X \} = (0.3)2 + (0.7)4 = 3.4 \]

\[ \sigma_X = \sqrt{(0.3)(-1.4)^2 + (0.7)(0.6)^2} = 0.917 \]

\[ E \{ Y \} = (0.445)10 + (0.555)20 = 15.55 \]

\[ \sigma_Y = \sqrt{(0.445)(-4.45)^2 + (0.555)(14.45)^2} = 11.17 \]
### EXAMPLE

| $x_i$ | $y_j$ | $x_i - E\{X\}$ | $y_j - E\{Y\}$ | $P\{X, Y | x_i, y_j\}$ |
|-------|-------|-----------------|-----------------|-----------------|
| 2     | 10    | −1.4            | 4.45            | 0.27            |
| 2     | 20    | −1.4            | 14.45           | 0.03            |
| 4     | 10    | 0.6             | 4.45            | 0.175           |
| 4     | 20    | 0.6             | 14.45           | 0.525           |
EXAMPLE

\[ \text{cov}\{X, Y\} = (0.27)(-6.23) + (0.03)(-20.23) + (0.175)(2.67) \]

\[ = 2.73 \]

\[ \rho_{XY} = \frac{\text{cov}\{X, Y\}}{\sigma_X \sigma_Y} = \frac{2.73}{(0.917)(4.970)} = 0.60 \]
The continuous probability distribution specification of a continuous r.v. $X$ may be expressed either in terms of a

- a probability density function (p.d.f.) $f_X(\cdot)$

$$f_X(x) \, dx \approx P\{x < X \leq x + dx\}$$

- or, a cumulative distribution function (c.d.f.) $F_X(\cdot)$

which expresses the probability that the value of $X$ is less or equal to a given value $x$

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(\xi) \, d\xi$$
EXPECTED VALUE, VARIANCE, STANDARD DEVIATION

- The **expected value** $\mu_X$ is given by

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

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

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

- The **standard deviation** $\sigma_X$ of $X$ is

$$\sigma_X = \sqrt{\text{var}\{X\}}$$
THE COVARIANCE AND THE CORRELATION

- The covariance $\text{cov}\left\{\tilde{X},\tilde{Y}\right\}$ of the two continuous random variables $\tilde{X}$ and $\tilde{Y}$ is given by

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

where $f_{\tilde{X},\tilde{Y}}(\cdot,\cdot)$ is the joint density function of $\tilde{X}$ and $\tilde{Y}$.

- The correlation coefficient $\rho_{\tilde{X},\tilde{Y}}$ is computed by

$$\rho_{\tilde{X},\tilde{Y}} = \frac{\text{cov}\{\tilde{X}, \tilde{Y}\}}{\sigma_X \sigma_Y}$$
We wish to guess the age $\tilde{A}$ 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 $P\{\tilde{A} > 50\} = 0.15$
- we also estimate that $P\{\tilde{A} \leq 40\} = 0.05$ and $P\{\tilde{A} \leq 44\} = P\{\tilde{A} > 44\}$
We construct the table of cumulative probability

<table>
<thead>
<tr>
<th>Event</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A \leq 29$</td>
<td>0.00</td>
</tr>
<tr>
<td>$A \leq 40$</td>
<td>0.05</td>
</tr>
<tr>
<td>$A \leq 44$</td>
<td>0.50</td>
</tr>
<tr>
<td>$A \leq 50$</td>
<td>0.85</td>
</tr>
<tr>
<td>$A \leq 65$</td>
<td>1.00</td>
</tr>
</tbody>
</table>
APPLICATION

$$P\left\{ A \leq x \right\}$$

years $x$

10 20 30 40 50 60 70

0.25 0.5 0.75 1.00

© 2006 – 2019 George Gross, University of Illinois at Urbana-Champaign, All Rights Reserved.
APPLICATION

\[ f_A(x) \]

\[ P\left\{ 40 < A \leq 50 \right\} \]

years \( x \)