Introduction

- Making decisions is one of the most
  - fruitful
  - enjoyable
  - empowering
  - aggravating
  - enervating

  human activities
- Probabilists and statisticians study the methodology for making rational decisions

An example from radar systems

- A radar transmitter emits electromagnetic pulses in the direction of a target that may (or may not) be “out there”
- The radar receiver listens for the echoes
- The length of time between transmission of the pulse and reception of the echo is used to estimate the range of the target
- Other information (e.g. Doppler shift in frequency, antenna orientation) is used to estimate velocity and compass bearing

I hate the inverse square law…

- The signal (if any) that is the input to the radar receiver is very weak, and is also contaminated by background noise
- The echo (if present at all) of a pulse is weak, and difficult to distinguish from the background noise
- If n transmitted pulses are reflected from a target, it is not certain that all n echoes will be heard clearly by the radar receiver

I shot an arrow into the air…

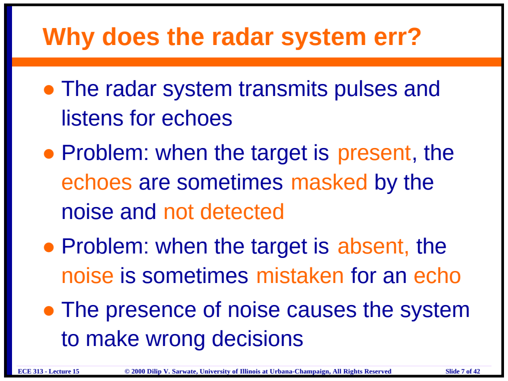
- A radar system does not know if a target is present or not
- The radar system transmits pulses and listens for echoes
- Based on the presence (or absence) of echoes, the radar system must decide whether a target is present or not
- This decision-making process is the subject of this lecture

It fell to the earth I know not where

- System decides if a target is present or not
- The decision may be correct or incorrect
- If the decision is that a target is present
  - then this decision is correct
    if in fact, there is a target present
  - but this decision is incorrect
    if in fact, there is no target present
- Such an incorrect decision is called a false alarm

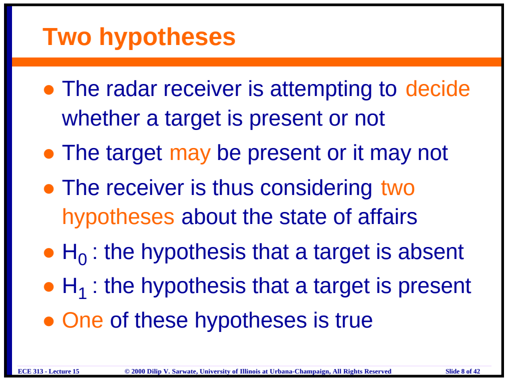
Being wrong the other way

- System decision as to whether a target is present or not may be correct or incorrect
- If the decision is that a target is absent
  - then this decision is correct
    if in fact, the target is absent
  - but this decision is incorrect
    if in fact, the target is present
- Such an incorrect decision is called a false dismissal or missed detection



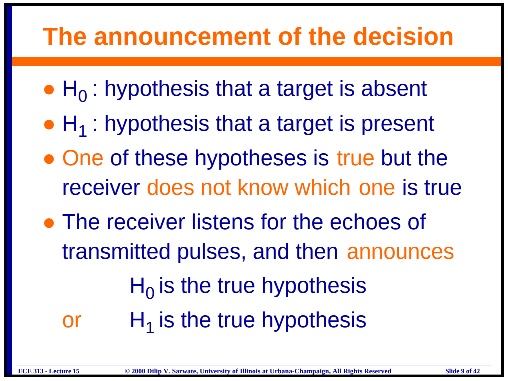
Why does the radar system err?

- The radar system transmits pulses and listens for echoes
- Problem: when the target is present, the echoes are sometimes masked by the noise and not detected
- Problem: when the target is absent, the noise is sometimes mistaken for an echo
- The presence of noise causes the system to make wrong decisions



Two hypotheses

- The radar receiver is attempting to decide whether a target is present or not
- The target may be present or it may not
- The receiver is thus considering two hypotheses about the state of affairs
- H_0 : the hypothesis that a target is absent
- H_1 : the hypothesis that a target is present
- One of these hypotheses is true

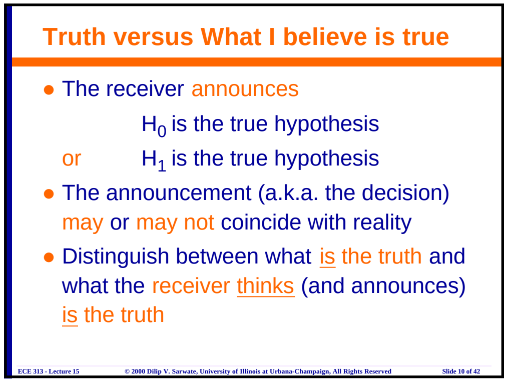


The announcement of the decision

- H_0 : hypothesis that a target is absent
- H_1 : hypothesis that a target is present
- One of these hypotheses is true but the receiver does not know which one is true
- The receiver listens for the echoes of transmitted pulses, and then announces

H_0 is the true hypothesis

or H_1 is the true hypothesis



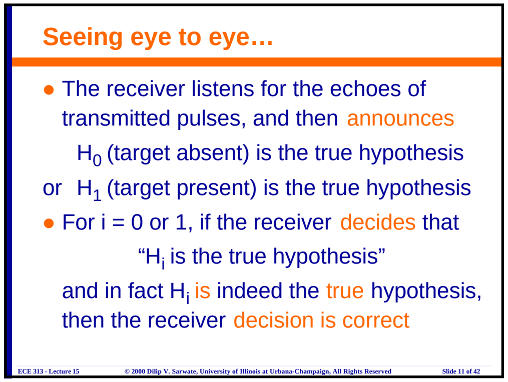
Truth versus What I believe is true

- The receiver announces

H_0 is the true hypothesis

or H_1 is the true hypothesis

- The announcement (a.k.a. the decision) may or may not coincide with reality
- Distinguish between what is the truth and what the receiver thinks (and announces) is the truth



Seeing eye to eye…

- The receiver listens for the echoes of transmitted pulses, and then announces

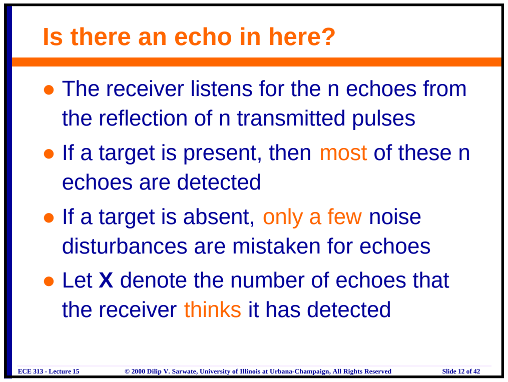
H_0 (target absent) is the true hypothesis

or H_1 (target present) is the true hypothesis

- For i = 0 or 1, if the receiver decides that

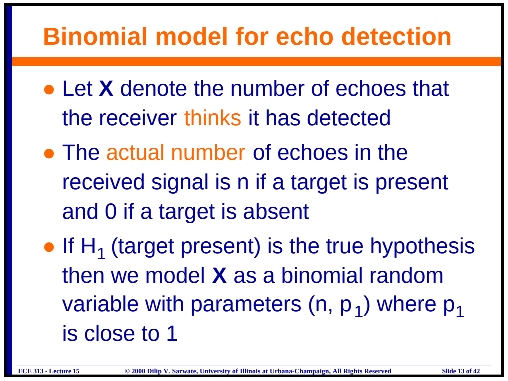
“H_i is the true hypothesis”

and in fact H_i is indeed the true hypothesis, then the receiver decision is correct



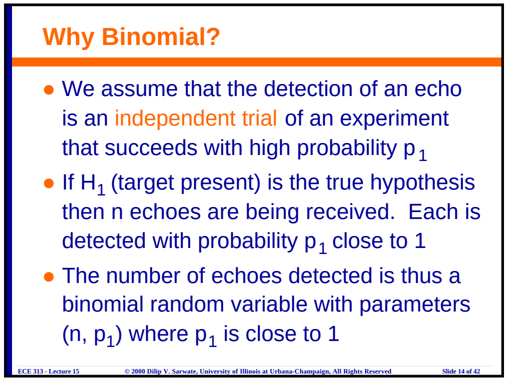
Is there an echo in here?

- The receiver listens for the n echoes from the reflection of n transmitted pulses
- If a target is present, then most of these n echoes are detected
- If a target is absent, only a few noise disturbances are mistaken for echoes
- Let **X** denote the number of echoes that the receiver thinks it has detected



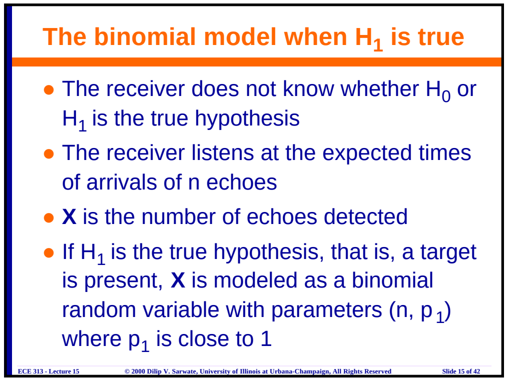
Binomial model for echo detection

- Let **X** denote the number of echoes that the receiver thinks it has detected
- The actual number of echoes in the received signal is n if a target is present and 0 if a target is absent
- If H_1 (target present) is the true hypothesis then we model **X** as a binomial random variable with parameters (n, p_1) where p_1 is close to 1



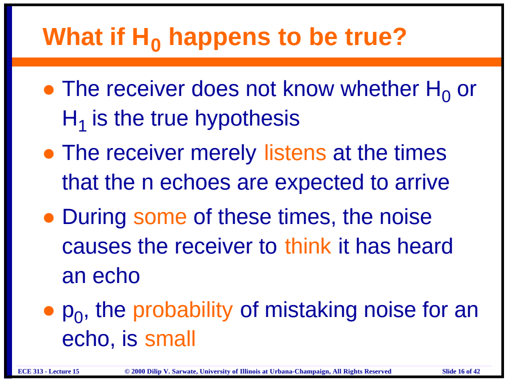
Why Binomial?

- We assume that the detection of an echo is an independent trial of an experiment that succeeds with high probability p_1
- If H_1 (target present) is the true hypothesis then n echoes are being received. Each is detected with probability p_1 close to 1
- The number of echoes detected is thus a binomial random variable with parameters (n, p_1) where p_1 is close to 1



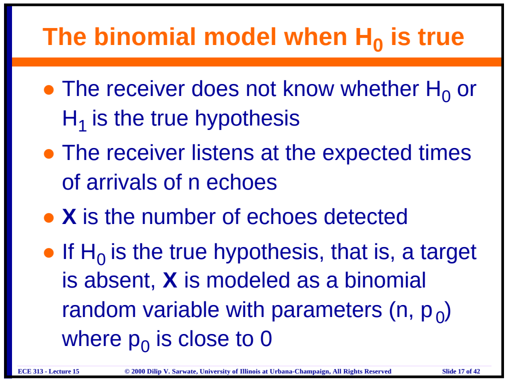
The binomial model when H_1 is true

- The receiver does not know whether H_0 or H_1 is the true hypothesis
- The receiver listens at the expected times of arrivals of n echoes
- **X** is the number of echoes detected
- If H_1 is the true hypothesis, that is, a target is present, **X** is modeled as a binomial random variable with parameters (n, p_1) where p_1 is close to 1



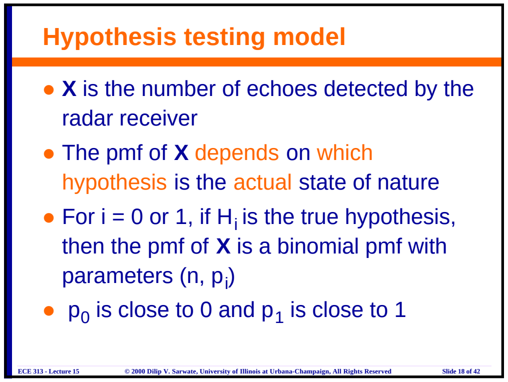
What if H_0 happens to be true?

- The receiver does not know whether H_0 or H_1 is the true hypothesis
- The receiver merely listens at the times that the n echoes are expected to arrive
- During some of these times, the noise causes the receiver to think it has heard an echo
- p_0, the probability of mistaking noise for an echo, is small



The binomial model when H_0 is true

- The receiver does not know whether H_0 or H_1 is the true hypothesis
- The receiver listens at the expected times of arrivals of n echoes
- **X** is the number of echoes detected
- If H_0 is the true hypothesis, that is, a target is absent, **X** is modeled as a binomial random variable with parameters (n, p_0) where p_0 is close to 0



Hypothesis testing model

- **X** is the number of echoes detected by the radar receiver
- The pmf of **X** depends on which hypothesis is the actual state of nature
- For i = 0 or 1, if H_i is the true hypothesis, then the pmf of **X** is a binomial pmf with parameters (n, p_i)
- p_0 is close to 0 and p_1 is close to 1

Standard notation

- **X** is the number of echoes detected by the radar receiver
- We write

  $H_o : X \sim \text{Binom}(n, p_0)$

  $H_1 : X \sim \text{Binom}(n, p_1)$

  to mean that if H_i is the true hypothesis then the pmf of **X** is a binomial pmf with parameters (n, p_i), $i = 0, 1$

What makes sense here?

- **X** is the number of echoes detected by the radar receiver
- $H_o : X \sim \text{Binom}(n, p_0)$
- $H_1 : X \sim \text{Binom}(n, p_1)$ $0 \approx p_0 << p_1 \approx 1$
- Since $p_1 >> p_0$, on average, the number of echoes detected when H_1 is true is much larger than the average number of echoes when H_0 is true

Do right and fear no man

- **X** is the number of echoes detected by the radar receiver
- $H_o : X \sim \text{Binom}(n, p_0)$
- $H_1 : X \sim \text{Binom}(n, p_1)$ $0 \approx p_0 << p_1 \approx 1$
- If **X** has a large value (close to n), the receiver should decide in favor of H_1
- If **X** has a small value (close to 0) , the receiver should decide in favor of H_0 as the true hypothesis

Where do we draw the line?

- **X** is the number of echoes detected by the radar receiver
- $H_o : X \sim \text{Binom}(n, p_0)$
- $H_1 : X \sim \text{Binom}(n, p_1)$ $0 \approx p_0 << p_1 \approx 1$
- If **X** has value not close to n or to 0, what should the receiver decide?
- A threshold test such as "If **X** > k, decide that H_1 is the true hypothesis; else decide that H_0 is the true hypothesis" is desirable

A general discussion

- Instead of completing the solution of the radar receiver problem right now, we take up a general formulation of the hypothesis testing problem
- The results will be applied later to the radar receiver problem and a complete solution provided
- Don't change the channel; we will be right back!

Hypotheses

- One of M mutually exclusive hypotheses $H_o, H_1, \ldots , H_{M-1}$ is true
- **X** is a random variable whose value we can observe, and use, to decide which of the hypotheses is true
- If H_i happens to be the true hypothesis, then the pmf of **X** is $P_i(u)$
- To avoid trivialities, we assume that $M \geq 2$

Binary hypothesis tests

- The case M = 2 with hypotheses H_o and H_1 is a very important special case
- H_o is called the null hypothesis
- H_1 is called the alternative hypothesis
- The alternative hypothesis is the one that we are trying to establish or prove, e.g. this drug cures cancer or target is present
- The null hypothesis is the default condition

It all depends on your viewpoint…

- In the binary hypothesis testing problem, statisticians usually speak of accepting the alternative or rejecting the alternative
- Accepting the alternative is the same as rejecting the null hypothesis, e.g. this drug does cure cancer or a target is present
- However, rejecting the alternative does not always mean accepting the null hypothesis

What say you? Guilty or not guilty?

- Rejecting the alternative may not be the same as accepting the null hypothesis
- The alternative is rejected if there is not sufficient evidence to support it
- It may be that neither hypothesis is true and the truth lies elsewhere
- People are often careless in formulating the statements of the hypotheses

The fairness doctrine

- Example of testing a coin to determine if it is a biased coin with P(Head) = 0.6
- The hypotheses should be

  H_o : H_1 is not true

  H_1 : Coin is a biased coin; P(Head) = 0.6
- The hypotheses are often stated as

  H_o : Coin is a fair coin; P(Head) = 0.5

  H_1 : Coin is a biased coin; P(Head) = 0.6

What's the difference?

- Example of testing a coin to determine if it is a biased coin with P(Head) = 0.6

  H_o : H_1 is not true

  H_1 : Coin is a biased coin; P(Head) = 0.6
- You are given one of two coins, a fair one or a biased coin with P(Head) = 0.6

  H_o : Coin is a fair coin; P(Head) = 0.5

  H_1 : Coin is a biased coin; P(Head) = 0.6

In some cases, it's all the same

- Some binary hypothesis testing problems are symmetrical by nature
- Distinction between the null and the alternative hypotheses as being two different types of creatures is unimportant
- In statistical testing, it is very important to know what exactly are the hypotheses
- Careful modeling and careful interpretation of the results is paramount

Speaking carelessly

- In binary hypothesis testing, we shall often carelessly say

  “announce H_0 is true” instead of the more precise “reject H_1”

or

  “announce H_1 is true” instead of the more precise “accept H_1”

- In many engineering situations, it does not matter

Back to the general case

- We have a random variable **X** whose pmf depends on which one of M hypotheses happens to be true
- Let **X** take on values u_1, u_2, … u_N
- We write the pmfs in an likelihood array (or likelihood matrix) of M rows (one for each hypothesis) and N columns (one for each of the possible values of **X**)

The likelihood matrix

- The likelihood matrix is an array of M rows and N columns
- Each row corresponds to a hypothesis and each column to one of the values taken on by **X**
- The entry in the i-th row and j-th column is $P_i(u_j)$, the probability that **X** = u_j when H_i is the true hypothesis

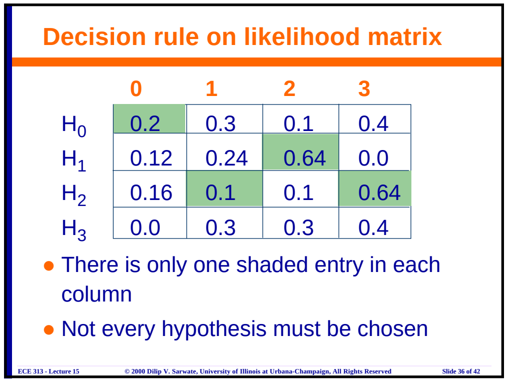
Example of a likelihood matrix

	0	1	2	3
H_0	0.2	0.3	0.1	0.4
H_1	0.12	0.24	0.64	0.0
H_2	0.16	0.1	0.1	0.64
H_3	0.0	0.3	0.3	0.4

- The values of **X** are shown in orange above the top row
- Each row is a pmf; some values are 0

Decisions, decisions, …

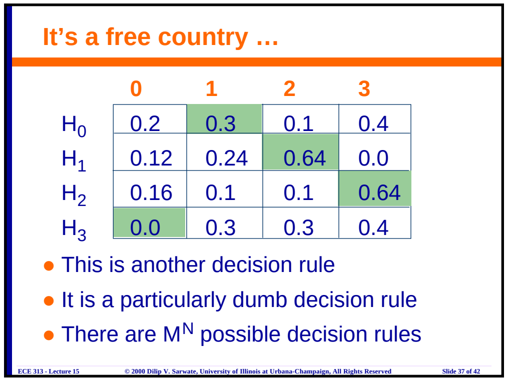
- After observing the value of **X**, a decision is made as to which hypothesis is true
- For each of the N values of **X**, there is a corresponding decision choosing one H_i
- The set of decisions is called a decision rule
- It is convenient to describe the decision this on the likelihood matrix by shading the corresponding entry in the matrix



Decision rule on likelihood matrix

	0	1	2	3
H_0	🟩 0.2	0.3	0.1	0.4
H_1	0.12	0.24	🟩 0.64	0.0
H_2	0.16	🟩 0.1	0.1	🟩 0.64
H_3	0.0	0.3	0.3	0.4

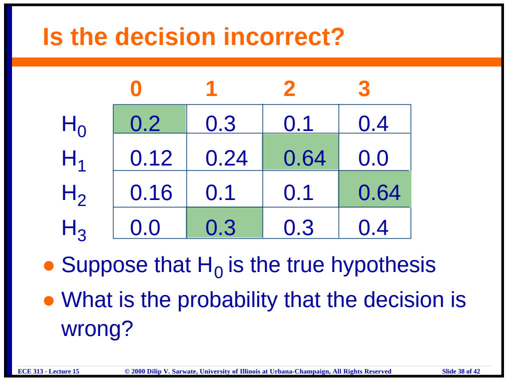
- There is only one shaded entry in each column
- Not every hypothesis must be chosen



It's a free country …

	0	1	2	3
H_0	0.2	**0.3**	0.1	0.4
H_1	0.12	0.24	**0.64**	0.0
H_2	0.16	0.1	0.1	**0.64**
H_3	**0.0**	0.3	0.3	0.4

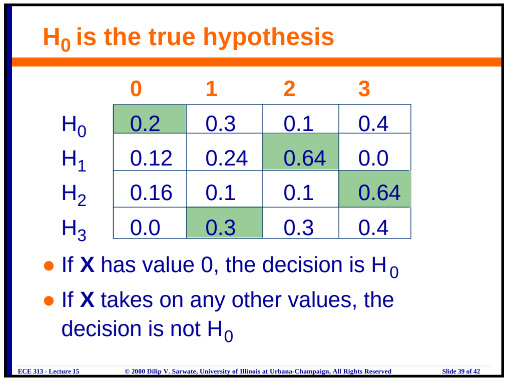
- This is another decision rule
- It is a particularly dumb decision rule
- There are M^N possible decision rules



Is the decision incorrect?

	0	1	2	3
H_0	**0.2**	0.3	0.1	0.4
H_1	0.12	0.24	**0.64**	0.0
H_2	0.16	0.1	0.1	**0.64**
H_3	0.0	**0.3**	0.3	0.4

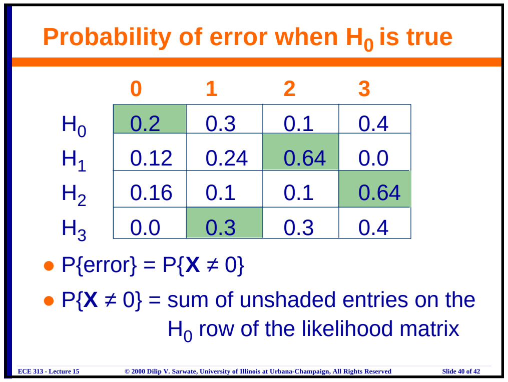
- Suppose that H_0 is the true hypothesis
- What is the probability that the decision is wrong?



H_0 is the true hypothesis

	0	1	2	3
H_0	**0.2**	0.3	0.1	0.4
H_1	0.12	0.24	**0.64**	0.0
H_2	0.16	0.1	0.1	**0.64**
H_3	0.0	**0.3**	0.3	0.4

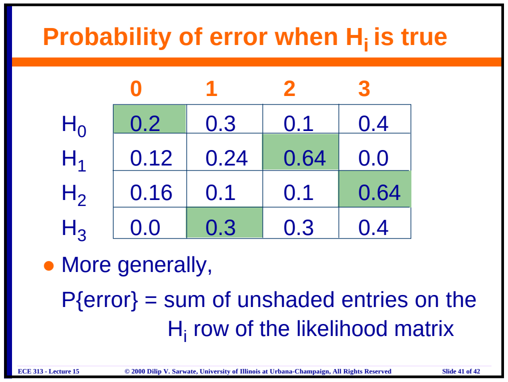
- If **X** has value 0, the decision is H_0
- If **X** takes on any other values, the decision is not H_0



Probability of error when H_0 is true

	0	1	2	3
H_0	**0.2**	0.3	0.1	0.4
H_1	0.12	0.24	**0.64**	0.0
H_2	0.16	0.1	0.1	**0.64**
H_3	0.0	**0.3**	0.3	0.4

- P{error} = P{**X** ≠ 0}
- P{**X** ≠ 0} = sum of unshaded entries on the H_0 row of the likelihood matrix

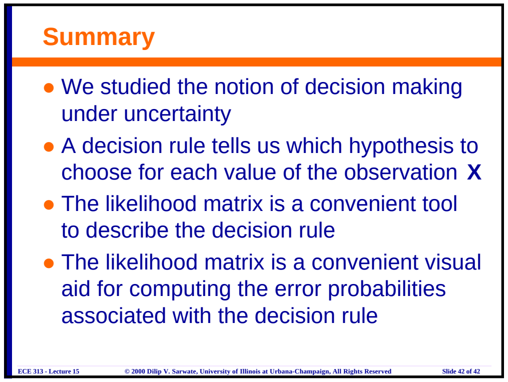


Probability of error when H_i is true

	0	1	2	3
H_0	**0.2**	0.3	0.1	0.4
H_1	0.12	0.24	**0.64**	0.0
H_2	0.16	0.1	0.1	**0.64**
H_3	0.0	**0.3**	0.3	0.4

- More generally,

  P{error} = sum of unshaded entries on the H_i row of the likelihood matrix



Summary

- We studied the notion of decision making under uncertainty
- A decision rule tells us which hypothesis to choose for each value of the observation **X**
- The likelihood matrix is a convenient tool to describe the decision rule
- The likelihood matrix is a convenient visual aid for computing the error probabilities associated with the decision rule