CIE 482: Path Planning & Navigation for Mobile Robots

Lecture 14: Basics of Probability Theory

King Fahd University of Petroleum & Minerals

Ramy Rashad
Assistant Professor
CIE Department

Outline

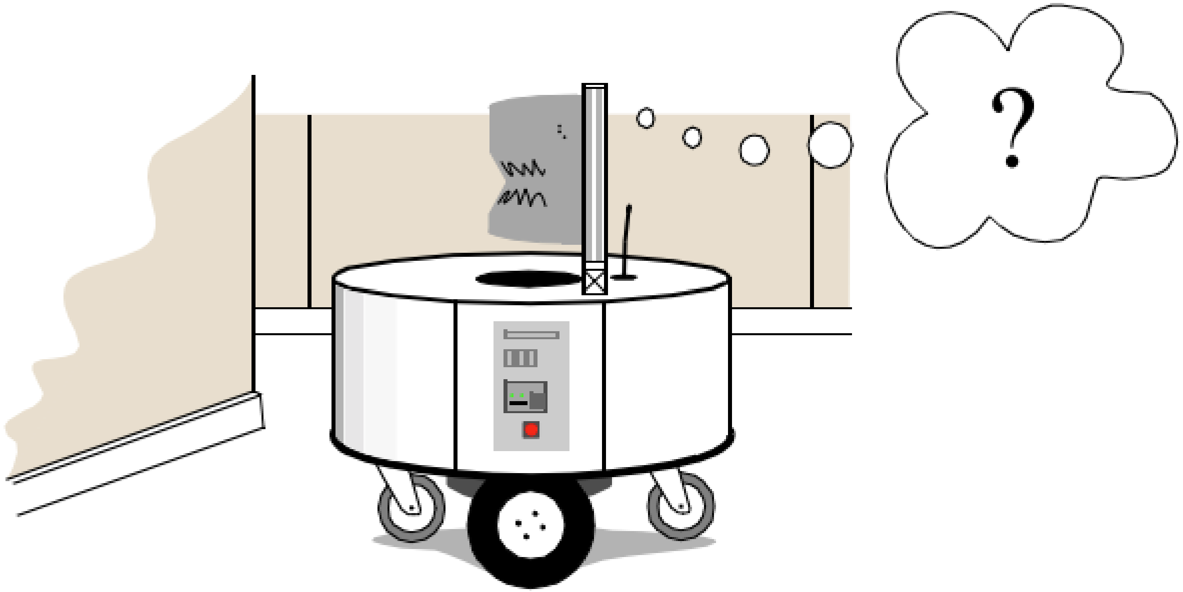
- Recap last lecture
- Basics of probability theory
  - Gaussian Distribution
  - Joint & Conditional Probability
  - Theorem of Total Probability

Outline

- Recap last lecture
- Basics of probability theory
  - Gaussian Distribution
  - Joint & Conditional Probability
  - Theorem of Total Probability

Recap: Robot Localization

- Localization is more than knowing the robot’s absolute position in Earth’s reference frame (using a GPS for instance)
- In general, localization involves:
  1. Identifying the robot’s pose relative to a map, which is a model of the (unknown) environment.
  2. Identifying the robot’s pose relative to targets, either static or dynamic (for example human targets).



Recap: Challenges of localization

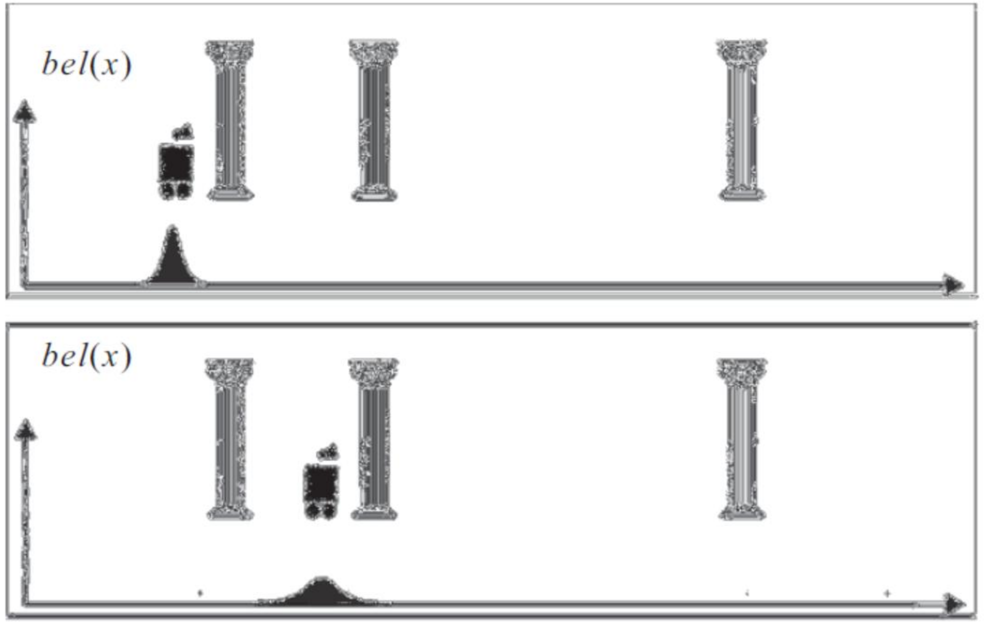
- The robot’s sensors and effectors (actuators) play an integral role in any robot localization framework.
- It is because of the inaccuracy and incompleteness of these sensors and effectors that localization poses difficult challenges.
- Challenges:
  - Sensor uncertainty
  - Actuator uncertainty

Recap: Belief Representation

- In the case of multiple-hypothesis beliefs regarding position, the robot tracks not just a single possible position but also a possibly infinite set of positions.
- In this approach, the belief is modeled by a probability distribution function $bel(x)$.
- In this approach, beliefs are

represented by conditional probability distributions:

$bel(x) = p(x|y)$



The belief function $bel(x)$ is our best guess of the robot's state x.

Recap: Probability Density Functions

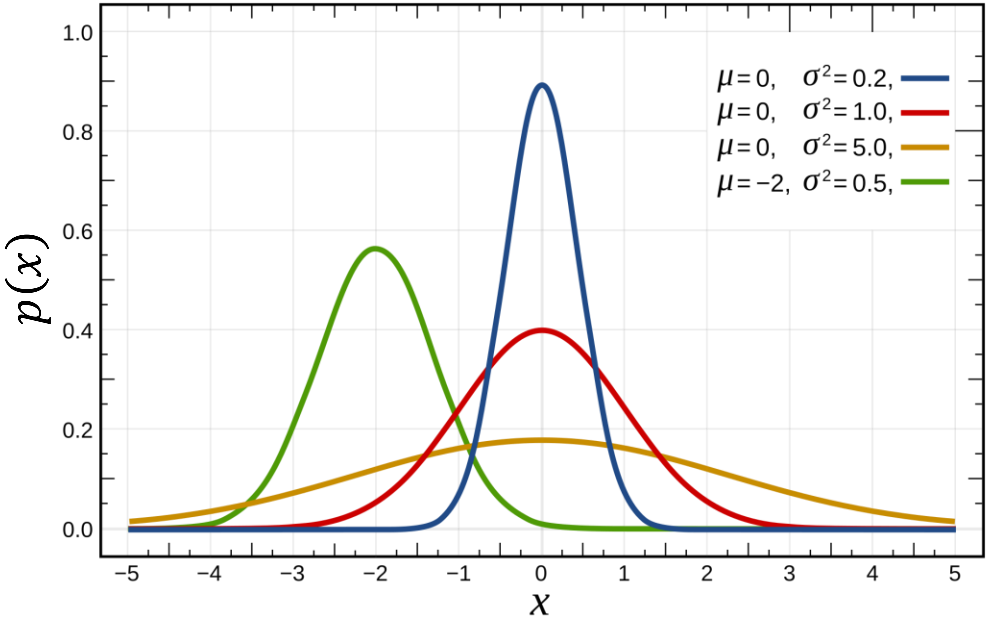
- A common probability density function is the Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(\frac{(x-\mu)^2}{2\,\sigma^2}\right)$$

also called normal distribution, which is commonly abbreviated with

$$p(x) = N(\mu, \sigma^2)$$

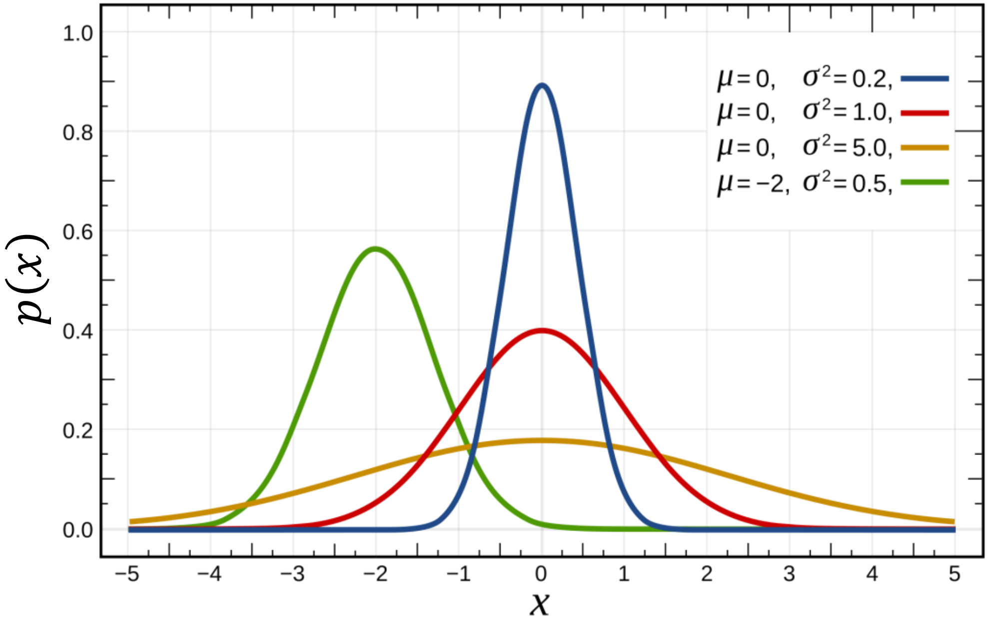
where $\mu \in \mathbb{R}$ & $\sigma^2 \in \mathbb{R}$ denote the mean & variance of the random variable $x \in \mathbb{R}$.



The expression $\boldsymbol{X} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$ means that the random variable X has a normal distribution.

Recap: Probability Density Functions

- The standard deviation σ is a measure of how dispersed the data is in relation to the mean.
- Small σ indicates data are clustered tightly around the mean
- Large σ indicates data are more spread out.



Question: What happens when $\sigma = 0$?

Outline

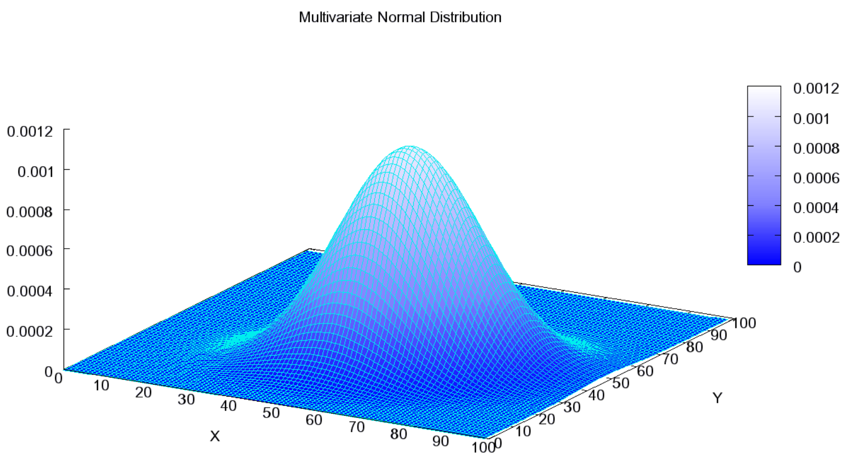
- Recap last lecture
- Basics of probability theory
  - Gaussian Distribution
  - Joint & Conditional Probability
  - Theorem of Total Probability

Multivariate Gaussian Distribution

- In case the random variable $x \in \mathbb{R}^k$ is a k-dimensional vector, we will have a multivariate normal distribution characterized by the density function:

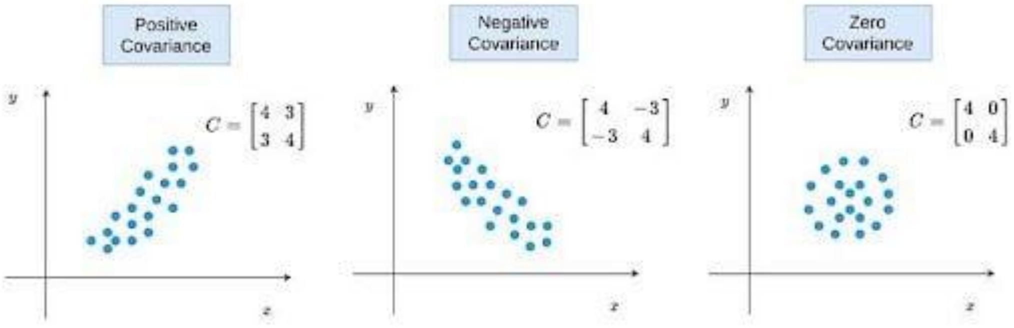
$$p(x) = \frac{1}{(2\pi)^{k/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$$

where $\mu \in \mathbb{R}^k$ & $\Sigma \in \mathbb{R}^{k \times k}$ denote the mean & covariance matrix.



Covariance

- Covariance is calculated between two different variables $\mathrm{cov}(x, y)$.
- Its purpose is to find the value that indicates how these two variables vary together.
  - $\mathrm{cov}(x, y) > 0$ means positive relationship (i.e. variate in the same direction)
  - $\mathrm{cov}(x, y) < 0$ means negative relationship
  - $\mathrm{cov}(x, y) = 0$ means no relationship



Positive, negative and near-zero covariance. | Image: Sergen Cansiz

The covariance between the same variables equals the variance

$$\mathrm{cov}(x, x) := \sigma_x^2$$

Covariance Matrix

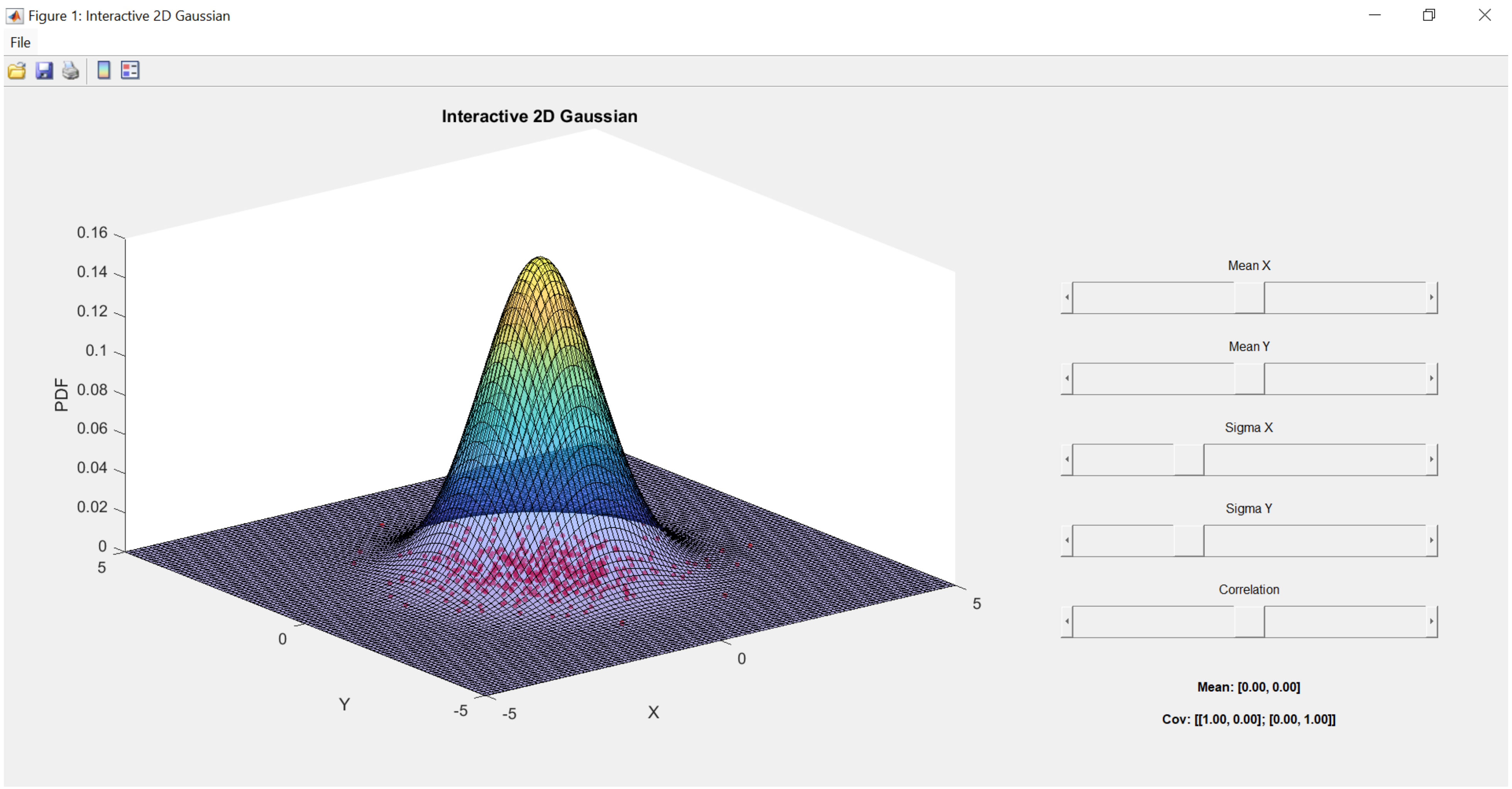
- The covariance matrix $\Sigma \in \mathbb{R}^{k \times k}$ is a symmetric positive definite matrix representing covariance values of each pair of variables in multivariate data.
- Example $k = 2, k = 3$

$$\begin{array}{c c} & \begin{array}{cc} x & y \end{array} \\ \begin{array}{c} x \\ y \end{array} & \begin{bmatrix} var(x) & cov(x,y) \\ cov(x,y) & var(y) \end{bmatrix} \end{array} \qquad \begin{array}{c c} & \begin{array}{ccc} x & y & z \end{array} \\ \begin{array}{c} x \\ y \\ z \end{array} & \begin{bmatrix} var(x) & cov(x,y) & cov(x,z) \\ cov(x,y) & var(y) & cov(y,z) \\ cov(x,z) & cov(y,z) & var(z) \end{bmatrix} \end{array}$$

Two- and three-dimensional covariance matrices. | Image: Sergen Cansiz

Question: What about eigenvalues of Σ ?

Gaussian Distribution Interactive Code



Available on Blackboard !

Outline

- Recap last lecture
- Basics of probability theory
  - Gaussian Distribution
  - Joint Probability & Conditional Probability
  - Theorem of Total Probability

Joint Probability

- The joint distribution of two random variables X and Y is given by

$$p(X = x, Y = y) \quad \text{or simply} \quad \boldsymbol{p(x, y)}$$

which describes the probability that the random variable X takes on the value x and that Y takes on the value y.

- Discrete case, the probabilities must sum to 1:

$$\sum_x \sum_y p(x, y) = 1$$

- Continuous case, the density must integrate to 1 over all possible x and y:

$$\int \int p(x, y) dx\, dy = 1$$

Example 1

- Consider a robot which can be in one of three possible locations: $X \in \{A, B, C\}$ while its sensor reading can be $Z = \{\mathrm{LOW}, \mathrm{HIGH}\}$.
- Suppose we have a joint probability table $p(X, Z)$ where X is the location and Z is the sensor reading:

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

- $p(X = A, Z = \mathrm{HIGH})$ =0.15 means there is a 15% probability that the robot is in location A **&** the sensor reads "HIGH."
- All the values together must sum to 1:

Marginal Probability

- Marginal Probability refers to the probability of a single event occurring, without consideration of any other events.
- It is derived from a joint probability distribution and represents the likelihood of an event happening in isolation.

- Discrete case:

$$p(x) = \sum_y p(x, y) \qquad\qquad p(y) = \sum_x p(x, y)$$

- Continuous case:

$$p(x) = \int p(x, y) dy \qquad\qquad p(y) = \int p(x, y) dx$$

Example 2

- What is the probability we get a HIGH reading from the sensor ?

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Example 2

- What is the probability we get a HIGH reading from the sensor ?

$$p(Z = \mathrm{HIGH}) = \sum_{x} p(X = x, Z = \mathrm{HIGH})$$

$$= p(X = A, Z = \mathrm{HIGH}) + p(X = B, Z = \mathrm{HIGH}) + p(X = C, Z = \mathrm{HIGH})$$

$$= 0.15 + 0.05 + 0.25$$

$$= 0.45$$

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Example 3

- What is the probability the robot is at position C ?

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Example 3

- What is the probability the robot is at position C ?

$$p(X = \mathrm{C}) = \sum_{z} p(X = C, Z = \mathrm{z})$$

$$= p(X = C, Z = \mathrm{HIGH}) + p(X = C, Z = \mathrm{LOW})$$

$$= 0.25 + 0.25$$

$$= 0.5$$

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Conditional Probability

- The conditional probability describes the probability that the random variable X takes on the value x conditioned that Y takes on the value y (i.e., we already know that $Y = y$).
- We denote this conditional probability by

$$p(X = x|Y = y) \quad \text{or simply} \quad \boldsymbol{p(x|y)}$$

- Intuitively, conditional probability tells us how our belief about X changes once we have observed (or “condition on”) Y.

Conditional Probability

- It is related to the joint probability by the formula:

$$\boxed{p(x|y) = \frac{p(x,y)}{p(y)}}$$

which reads as:

> The conditional probability of x given y, is equal to their joint probability divided by the marginal probability of y.

Example 4

- If we get a HIGH reading from our sensor, what is the probability that the robot is now in location A ?

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Example 4

- If we get a HIGH reading from our sensor, what is the probability that the robot is now in location A ?

$$p(X = A \mid Z = \mathrm{HIGH}) = \frac{p(X = A\,, Z = \mathrm{HIGH})}{p(Z = \mathrm{HIGH})}$$

- From the table, we have that $p(X = A\,, Z = \mathrm{HIGH}) = 0.15$
- From Example 2, we have that $p(Z = \mathrm{HIGH}) = 0.45$
- Therefore,

$$p(X = A \mid Z = \mathrm{HIGH}) = \frac{0.15}{0.45} = \frac{1}{3} = 0.333\ldots$$

	Z = LOW	Z = HIGH
X = A	0.10	0.15
X = B	0.20	0.05
X = C	0.25	0.25

Special case: Independent variables

- Two random variables X and Y are said to be independent if and only if the occurrence (or value) of one does not affect the occurrence (or value) of the other.
- Formally,

$$p(x, y) = p(x)p(y)$$

This means the joint probability of $X = x$ and $Y = y$ **factorizes** into the product of the individual (marginal) probabilities $p(X = x)$ and $p(Y = y)$.

Special case: Independent variables

- Two random variables X and Y are said to be independent if and only if the occurrence (or value) of one does not affect the occurrence (or value) of the other.
- Formally,

$$p(x, y) = p(x)p(y)$$

- In this case, the conditional probability $p(x|y)$becomes:

$$p(x|y) = \frac{p(x, y)}{p(y)} = p(x)p(y) = p(x)$$

This means that when X and Y are independent, the knowledge of Y does not provide any extra information about the value of X.

Example 5

- Let X be the outcome of the first coin flip and Y be the outcome of the second coin flip.
- Each coin flip has two possible outcomes: $X \in \{H, T\}, Y \in \{H, T\}$.

where H =Heads and T= Tails.

Example 5

- Let X be the outcome of the first coin flip and Y be the outcome of the second coin flip.
- Each coin flip has two possible outcomes: $X \in \{H, T\}, Y \in \{H, T\}$.

where $H =$Heads and T= Tails.

- For a fair coin, we have that

$$P(X = H) = 0.5, \qquad P(X = T) = 0.5.$$
$$P(Y = H) = 0.5, \qquad P(Y = T) = 0.5.$$

Example 5

- Let X be the outcome of the first coin flip and Y be the outcome of the second coin flip.
- Each coin flip has two possible outcomes: $X \in \{H, T\}, Y \in \{H, T\}$.

where $H =$Heads and T= Tails.

- For a fair coin, we have that

$$P(X = H) = 0.5, \qquad P(X = T) = 0.5.$$
$$P(Y = H) = 0.5, \qquad P(Y = T) = 0.5.$$

- Compute the joint probability $P(X = H, Y = T)$ and the conditional probability $P(X = H | Y = T)$.