Probabilistic Approach to Localisation

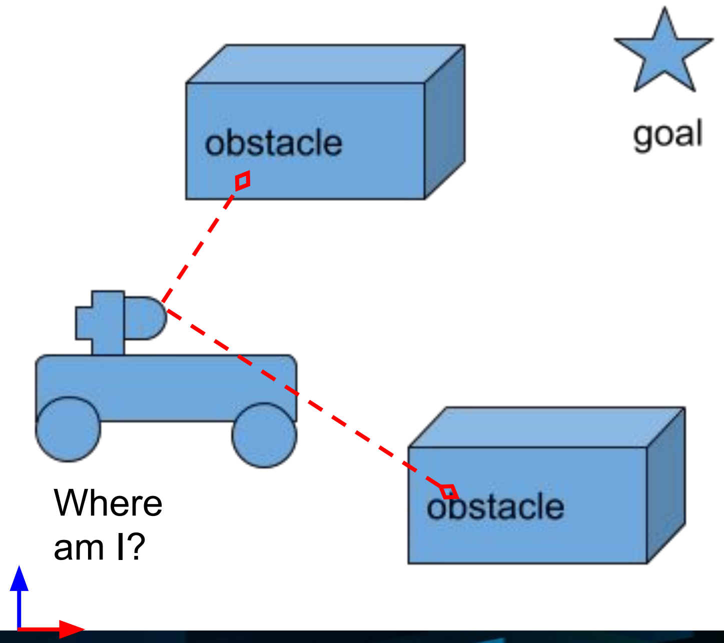
Feras Dayoub

Learning Objectives

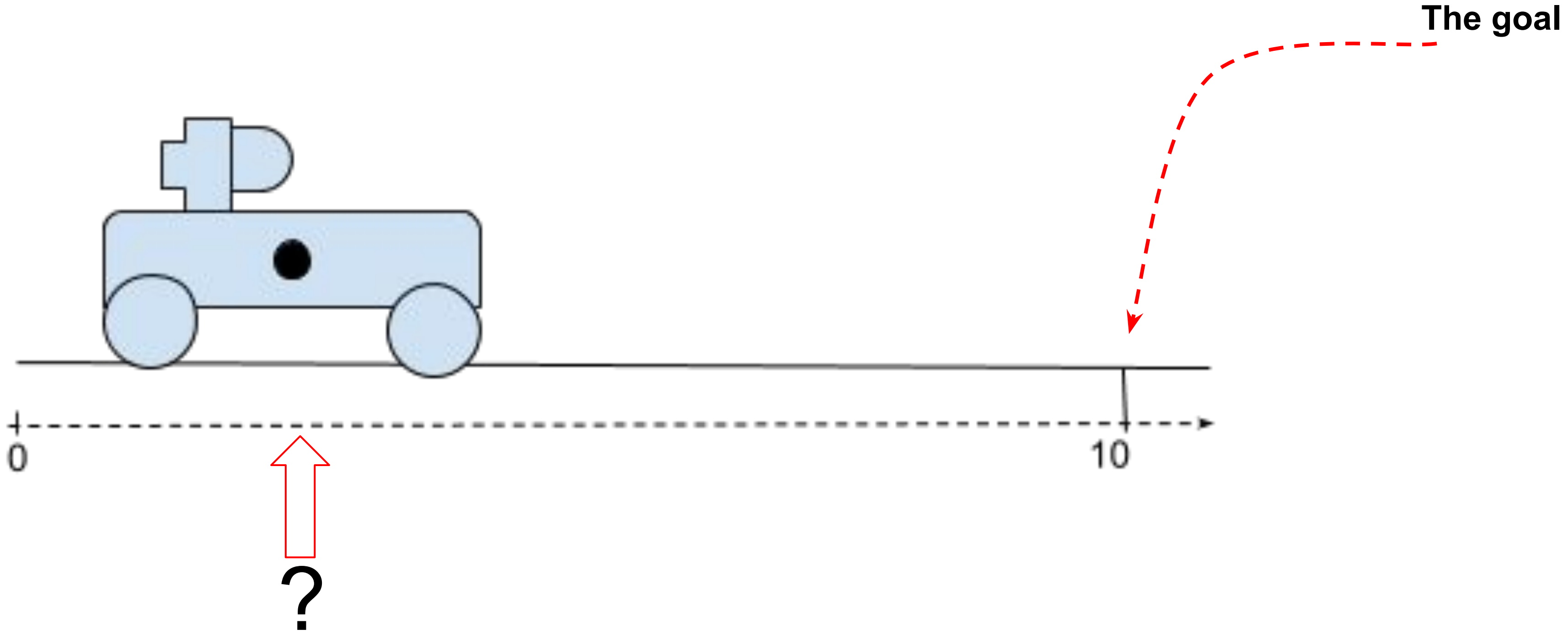
- Probabilistic approach to localisation.
- Introduction to Bayesian inference.
  - Bayesian inference is a method by which Bayes' rule is used to update the probability for a hypothesis as more evidence or measurements become available.

Why does the robot need to localise itself?

- The robot needs to know its state (e.g position and orientation) and the state of its working place (e.g a map of a landmarks) in order to perform its task successfully.



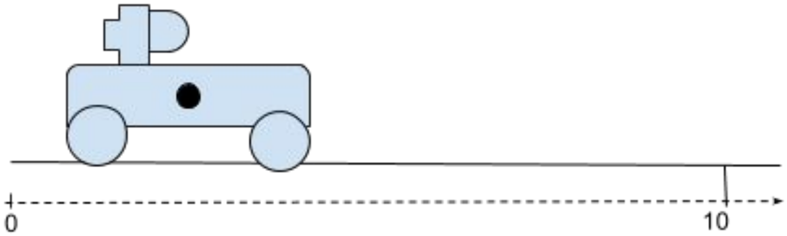
How does a robot localise itself?



How does a robot localise itself?

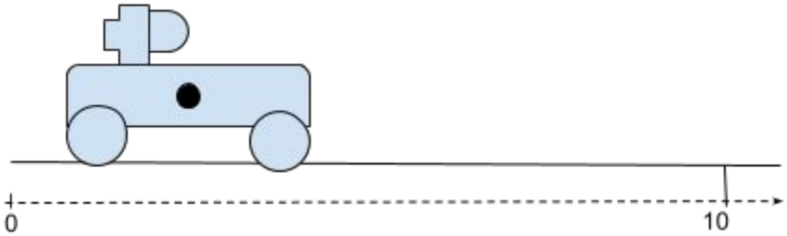
Why wheel encoders are not enough?

- Limited resolution.
- Unequal floor contact and variable friction can lead to slipping.



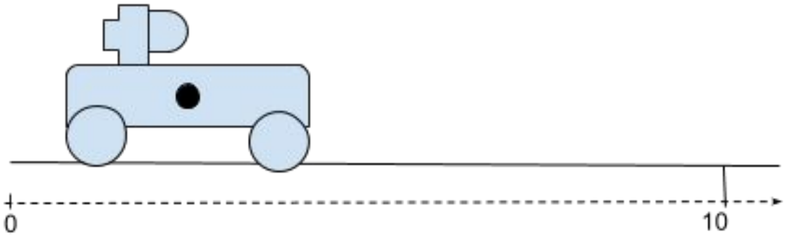
How does a robot localise itself?

- Can we use another sensors?



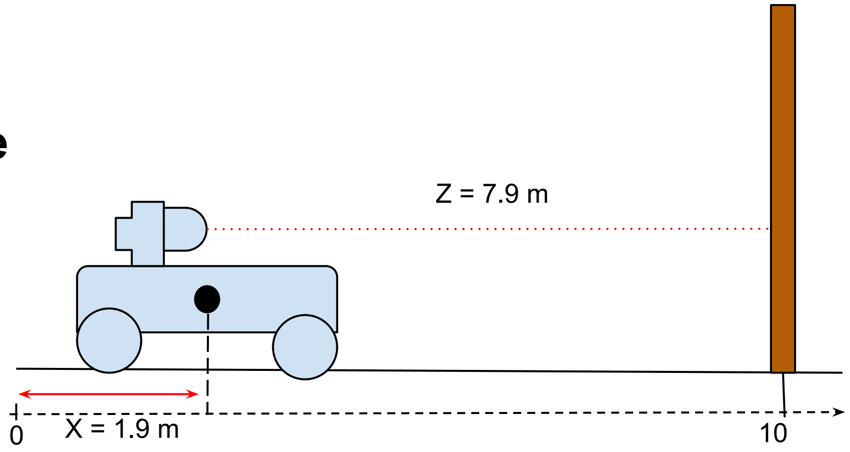
How does a robot localise itself?

- Internal sensors
  - Accelerometers (spring-mounted masses)
  - Gyroscopes (spinning mass, laser light)
  - Compasses, inclinometers (earth magnetic field, gravity)
- Proximity sensors
  - Sonar
  - Radar
  - Laser range-finders
- **Visual sensors: Cameras**
- Satellite-based sensors: GPS



Using another sensor to improve the localisation

- An external sensor measures distance to a landmark with known location.
- The measurement is not perfect.
- The readings of the sensor give us an idea about the **error in our current pose estimate**.



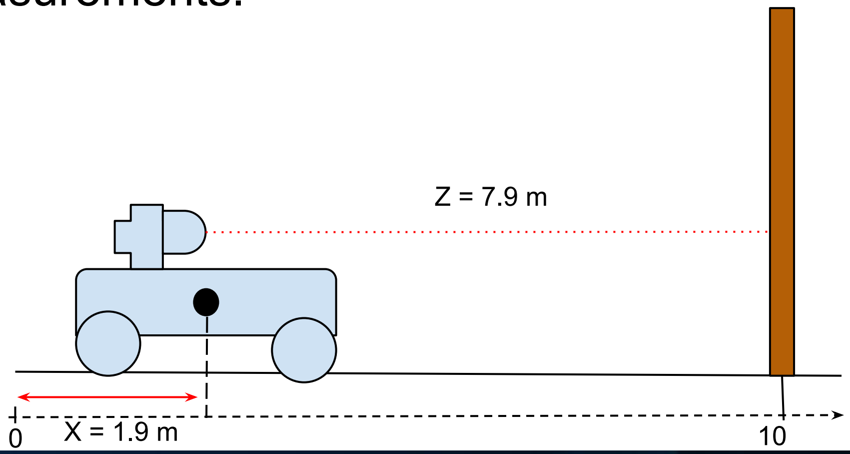
State estimation

The state of the robot is a set of quantities that, if known fully, describe the robot's motion over time.

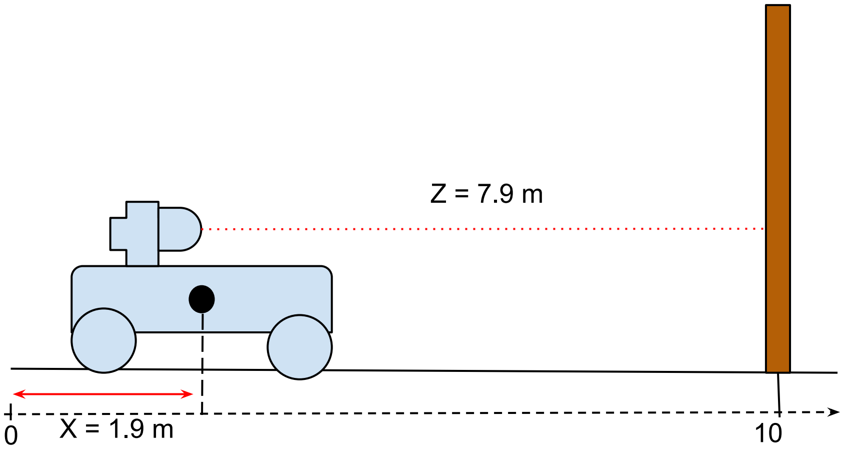
- The robot poes (i.e. position and orientation)
- Velocity.

Why do we need to use probability theory in robotics?

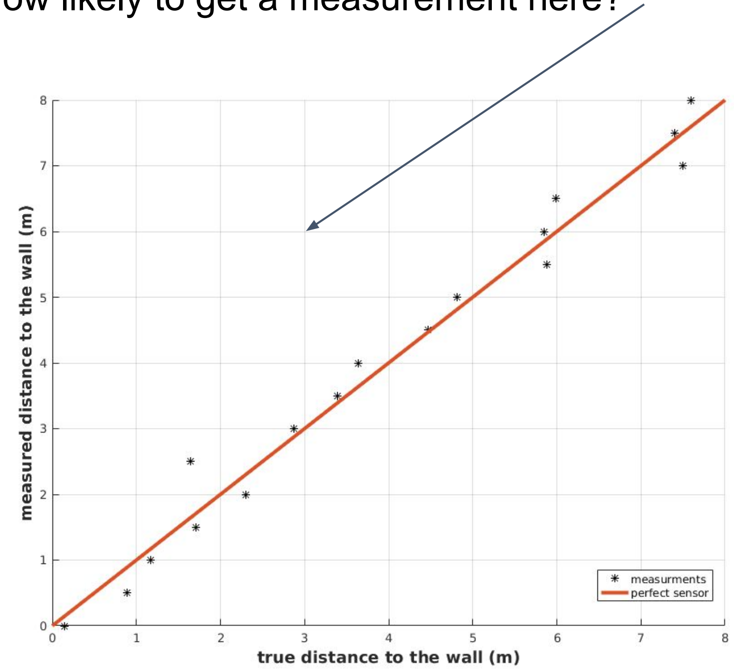
- The state of the robot is estimated using the wheel encoders and measurements from the robot's sensors.
- Wheel slippage and noisy sensor data make knowing the **true** state of the robot and its workspace **unattainable**.
- Probability theory provides a framework to handle the uncertainty caused by all the noise in the measurements.



Modeling sensor uncertainty



How likely to get a measurement here?



Modeling sensor uncertainty

In robotics, we model uncertainty using **probability distributions**.

A probability distribution tell us how likely a particular measurement is given the information we do know about the accuracy of the sensor and the robot's state.

Also, sensor measurements, the control inputs, the state of the robot and its environment all are treated as **random variable.**

Probabilistic Approach

Quick Refresher

- Frequentist vs Bayesian interpretation of probability
- Axioms of Probability Theory.
- Random Variables.
- Conditional Probabilities.
- Bayes Rule.

Frequentist vs Bayesian interpretation of probability

"the probability that a fair coin will land heads is 0.5"

<u>In the frequentist view:</u> probabilities represent long run frequencies of events. For example, the above statement means that, if we flip the coin many times, we expect it to land heads about half the time.

<u>In the bayesian view:</u> probabilities represent measures of *uncertainty* or degrees of belief. The above statement means we think the coin is equally likely to land heads or tails on the next toss.

In this unit we are <u>bayesians.</u>

Axioms of Probability Theory

$$1) \quad P(A) \geq 0 \quad \forall \quad A \subset S$$

$$2) \quad P(S) = 1$$

$$3) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

S **Sample space: the set of all possible outcomes of an experiment.**

A B **Events: subsets of the sample space.**

Random Variables

A random variable X is the mathematical expression of the output of an experiment.

- X might take on any value from the set of possible outcomes.
    - $X \in \{heads, tails\}$ if the experiment was flipping a coin.
        - $P(X=heads) = P(heads) = 0.5$
    - $X \in \{1,2,3,4,5,6\}$ if the experiment was rolling a die.
        - $P(1) = \frac{1}{6}$
- We will be working with continuous random variables.

Conditional probabilities

- **P(a|b) is the probability of X = a given that Y = b**

$$P(a|b) = P(a,b)/P(b)$$

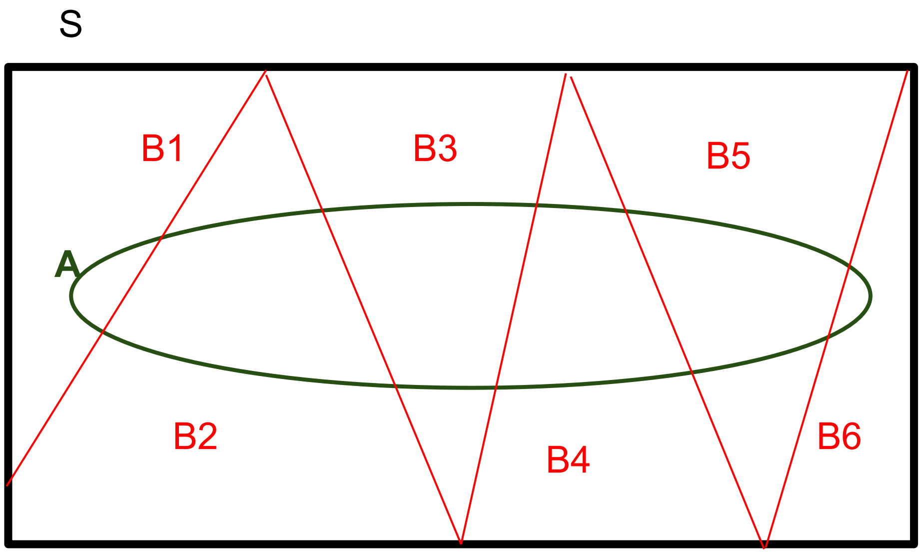
$$P(a,b) = P(a|b) \times P(b)$$

$$P(X = a \text{ and } Y = b) = P(a,b)$$

Law of total probability

$$P(A) = \sum_{i}^{n} P(A \cap B_i)$$

$$P(A) = \sum_{i}^{n} P(A|B_i) \times P(B_i)$$



Bayes Rule

$$P(a|b) = P(a,b)/P(b)$$

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

$$P(X = a \text{ and } Y = b) = P(a,b)$$

Bayes Rule

Suppose a robot with a door detector obtains a door measurement $Z \in \{0, 1\}$

What is the probability that the door is open given that the detector is showing **1**?

$$P(X = open | Z = 1)$$

Bayes Rule

Suppose a robot with a door detector obtains a door measurement $Z \in \{0, 1\}$

What is the probability that the door is open given that the detector is showing **1**?

$$P(X = open|Z = 1) = \frac{P(Z = 1|X = open) \times P(X = open)}{P(Z = 1)}$$

$$P(Z = 1|X = open) = 0.8$$
$$P(Z = 1|X = closed) = 0.1$$
$$P(X = open) = 0.1$$

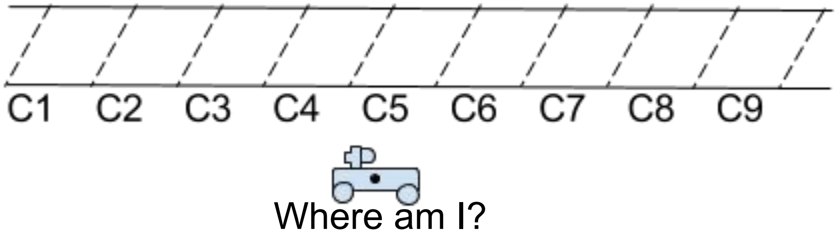
What about this?

Bayes Rule

Suppose a robot with a door detector obtains a door measurement $Z \in \{0, 1\}$

What is the probability that the door is open given that the detector is showing **1**?

$$P(X = open|Z = 1) = \ 47\%$$



Toy example

Our robot lives on a grid of 9 cells.

Let **X** be a random variable representing the location of the robot on the grid.

Without having any prior knowledge about the robot,

What is P(X=C5)?

What is the probability distribution for X?

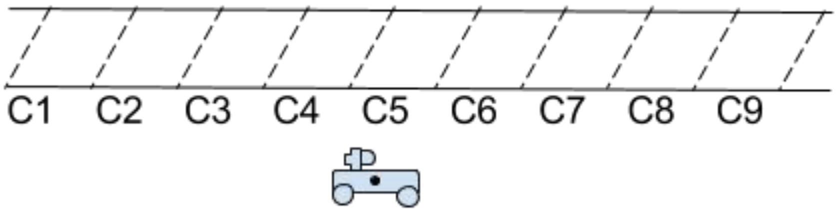
Toy example

X at time 0 has a discrete uniform distribution.

The probability mass function (**PMF**) of X is:

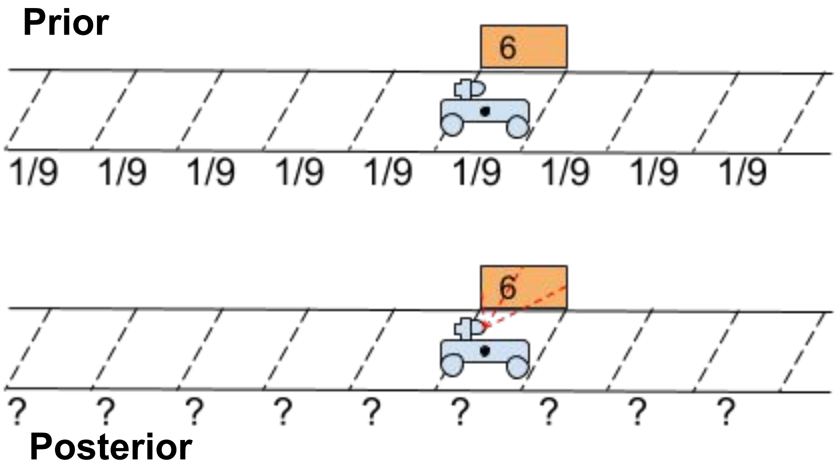
$$f(C_i) = P(X = C_i) = \frac{1}{9}$$

$$\sum_{i=1}^{9} f(C_i) = 1$$



Toy example

- The robot has a sign detector onboard, and it knows that there is only one sign located in cell 6.
- The sign detector is not perfect, so we model its measurements as a random variable (Z). $Z \in$ {sign, no_sign}
- The company who sold us the sensor tell us that
  - P(Z=sign|sign) = 0.65
  - P(Z=sign|no_sign) = 0.1



What is the probability distribution of X after detecting a sign (the posterior)?

Toy example

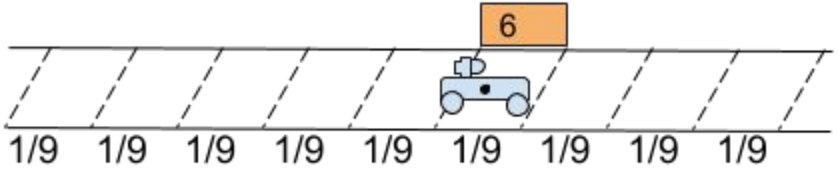
$$P(X = C_i | Z = sign) = \frac{P(Z = sign | X = C_i) P(X = C_i)}{P(Z = sign)}$$

How we calculate the denominator in this simple case? Remember that:

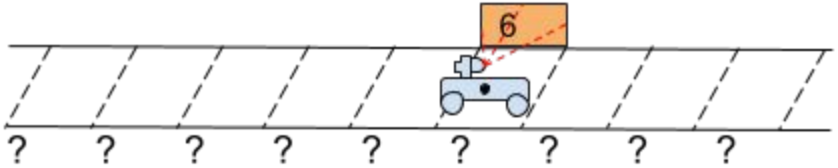
Z ∈ {sign, no_sign}

P(Z=sign|sign) = 0.65, **reads as, detecting a sign given that we are on a cell with a sign and it is referred to as the <u>likelihood</u>.**

P(Z=sign|no_sign) = 0.1



Prior distribution



Posterior

Toy example

The **law of total probabilities** and the **marginal** PMF.

$$P(Z = sign) = \sum_{i=1}^{9} P(Z = sign|X = C_i)P(X = C_i)$$

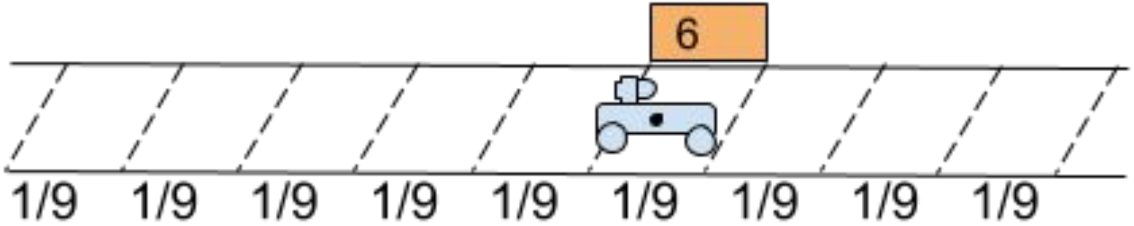
$$P(Z = sign) = P(sign|C_1)P(C_1)+...+P(sign|C_6)P(C_6)+...+P(sign|C_9)P(C_9)$$

$$P(Z = sign) = 0.1 \times \frac{1}{9} + ... + 0.65 \times \frac{1}{9} + ... + 0.1 \times \frac{1}{9}$$

Toy example

$$P(X = C_i | Z = sign) = \frac{P(Z = sign | X = C_i) P(X = C_i)}{P(Z = sign)}$$

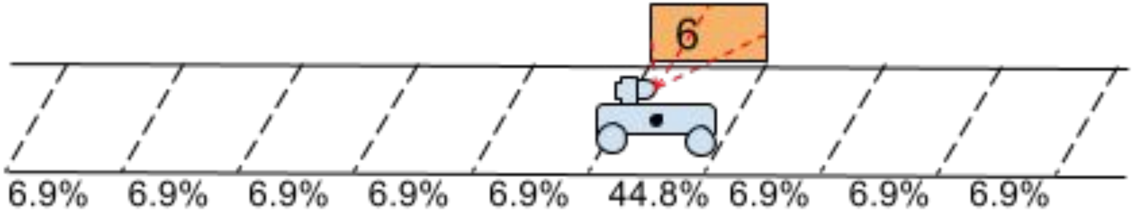
Prior



11%

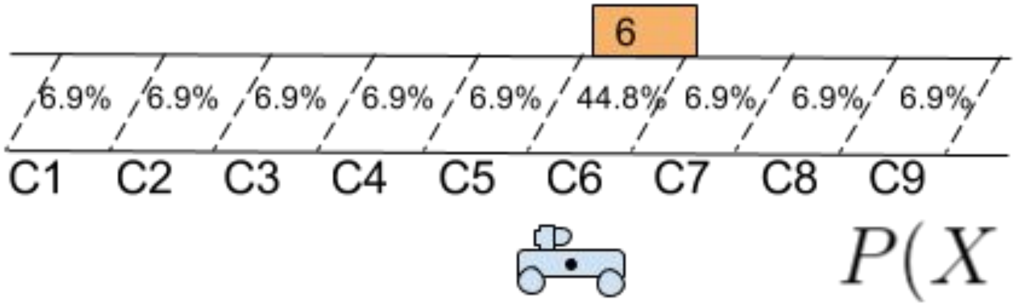
before detection

Posterior



after detection

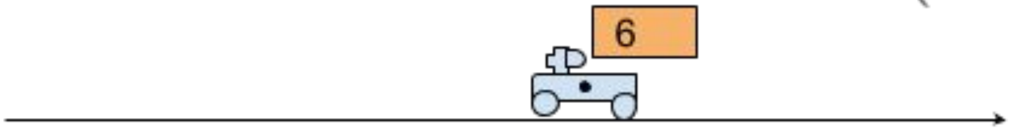
We live in a continuous world



discrete world

$$P(X = C_6) = 44.8\ \%$$

$$P(X = 5.22\ cm) =?$$



continuous world

Continuous random variables

- Discrete random variables can take only a countable number of possible values (the robot can be in one of the 9 cells).
- In reality, the robot is moving in a continuous space and can be in any of infinite values (i.e $\mathbf{x \in R, P(X = x) = 0}$). Therefore we have to use **intervals** instead.
- We use probability density function (**PDF**) instead of the probability mass function (PMF):

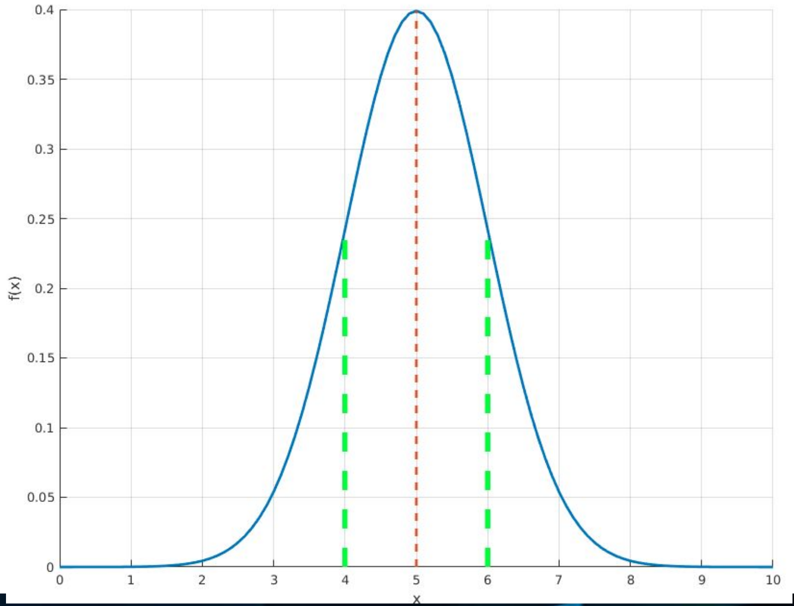
The normal distribution

$$f(x) = \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{1}{2}(x-\mu)^2 \frac{1}{\sigma^2}}$$

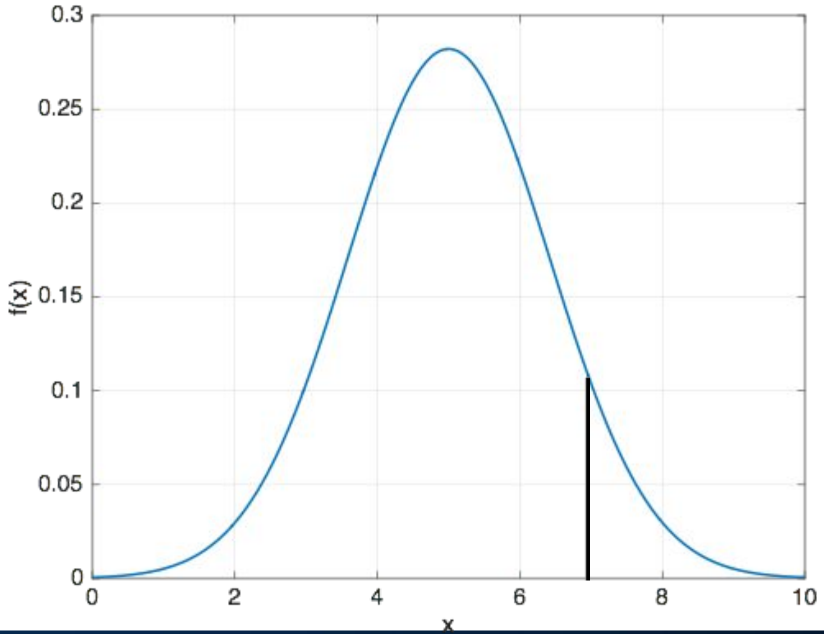
$$X \sim N(\mu_X, \sigma_X^2)$$

```
x = [0:0.1:10];
mu = 5
sigma = 1
norm = normpdf(x,mu,sigma)
hold on
grid on
plot(x,norm)
plot([mu,mu],[0,normpdf(mu,mu,sigma)],'--')
plot([mu+sigma,mu+sigma],[0,normpdf(mu+sigma,mu,sigma)],'g--')
plot([mu-sigma,mu-sigma],[0,normpdf(mu-sigma,mu,sigma)],'g--')
```

Only 2 parameters:
mean μ and standard deviation σ
variance is σ2



Probability from density

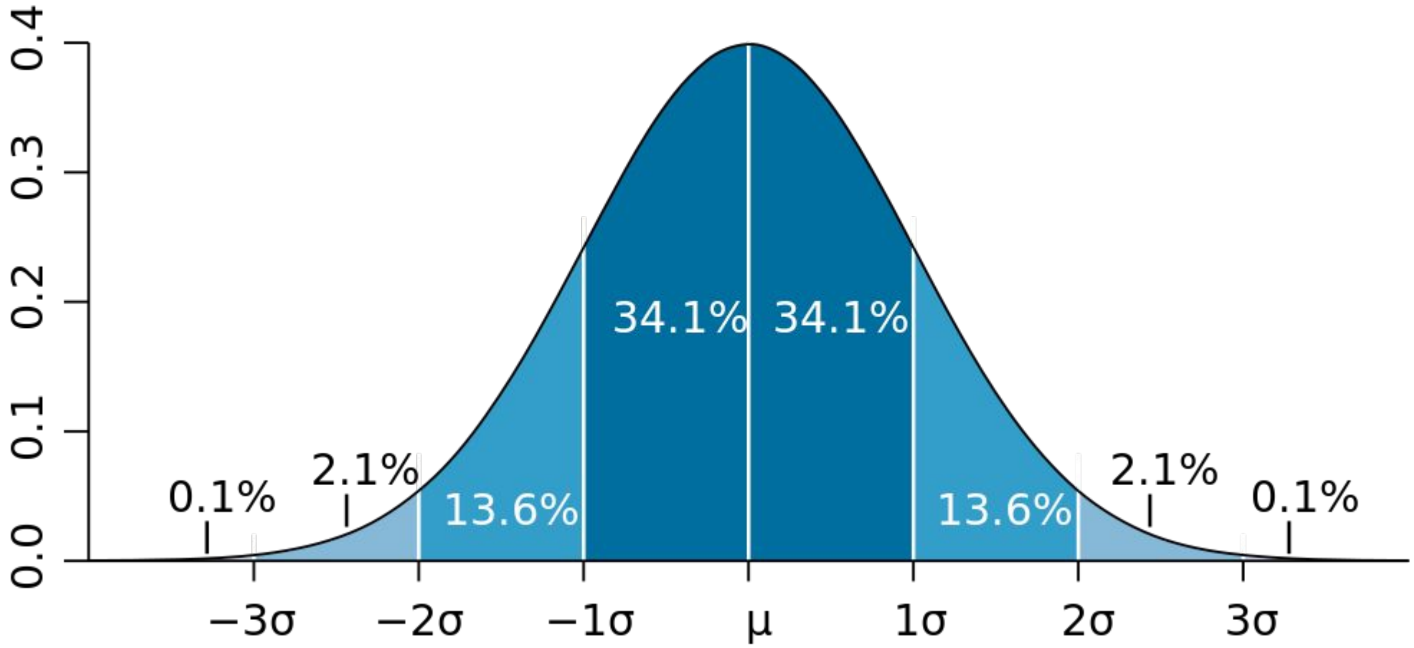


$$P(X \approx x) = f(x)dx$$

$$P(X = 7.00) = P(6.995 \leq X < 7.005)$$

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Normal/Gaussian distribution



Expectation and variance of a random variable

- Like when we use the average of a set of numbers to describe the whole set,
- the average of a random variable X, or as it is called, its expected value is:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

- If X has a normal distribution, the above integral gives: $E(X) = \mu$
- And its variance is:

$$\mathrm{Var}(X) = E\left[(X - \mu)^2\right] = \sigma^2$$

Linear Operations on random variables

A Gaussian random variable $X \sim N(\mu_X, \sigma_X^2)$

If $Y = aX + b$

Then: $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$

$$\begin{aligned} \mu_Y &= a\mu_X + b \\ \sigma_Y^2 &= a^2\sigma_X^2 \end{aligned}$$

Linear Operations on random variables

- Sum of two independent Gaussian random variables $Z = X + Y$

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2) \qquad Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

- is a Gaussian random variable

- where: $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$

$$\mu_Z = \mu_X + \mu_Y$$
$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$$

The proof

In case you are curious.

For independent random variables X and Y, the distribution f_Z of $Z = X + Y$ equals the convolution of f_X and f_Y:

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x)\, dx$$

Given that f_X and f_Y are normal densities,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x-\mu_X)^2/(2\sigma_X^2)}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-(y-\mu_Y)^2/(2\sigma_Y^2)}$$

Substituting into the convolution:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(z-x-\mu_Y)^2}{2\sigma_Y^2}} \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_X^2+\sigma_Y^2}} \exp\left[-\frac{(z-(\mu_X+\mu_Y))^2}{2(\sigma_X^2+\sigma_Y^2)}\right] \frac{1}{\sqrt{2\pi}\frac{\sigma_X\sigma_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}} \exp\left[-\frac{\left(x-\frac{\sigma_X^2(z-\mu_Y)+\sigma_Y^2\mu_X}{\sigma_X^2+\sigma_Y^2}\right)^2}{2\left(\frac{\sigma_X\sigma_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right)^2}\right] dx$$

$$= \frac{1}{\sqrt{2\pi(\sigma_X^2+\sigma_Y^2)}} \exp\left[-\frac{(z-(\mu_X+\mu_Y))^2}{2(\sigma_X^2+\sigma_Y^2)}\right] \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\frac{\sigma_X\sigma_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}} \exp\left[-\frac{\left(x-\frac{\sigma_X^2(z-\mu_Y)+\sigma_Y^2\mu_X}{\sigma_X^2+\sigma_Y^2}\right)^2}{2\left(\frac{\sigma_X\sigma_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right)^2}\right] dx$$

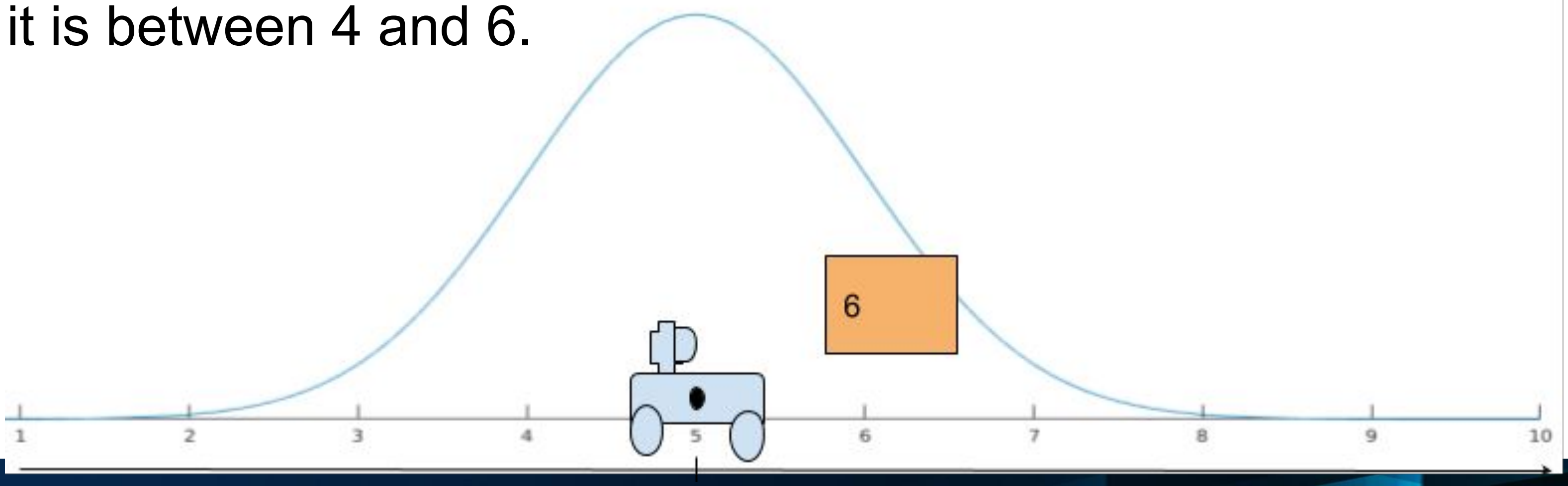
The expression in the integral is a normal density distribution on x, and so the integral evaluates to 1. The desired result follows:

$$f_Z(z) = \frac{1}{\sqrt{2\pi(\sigma_X^2+\sigma_Y^2)}} \exp\left[-\frac{(z-(\mu_X+\mu_Y))^2}{2(\sigma_X^2+\sigma_Y^2)}\right]$$

Toy example

We represent the state (in this case the position) of the robot at time **t** as a random variable $X_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$

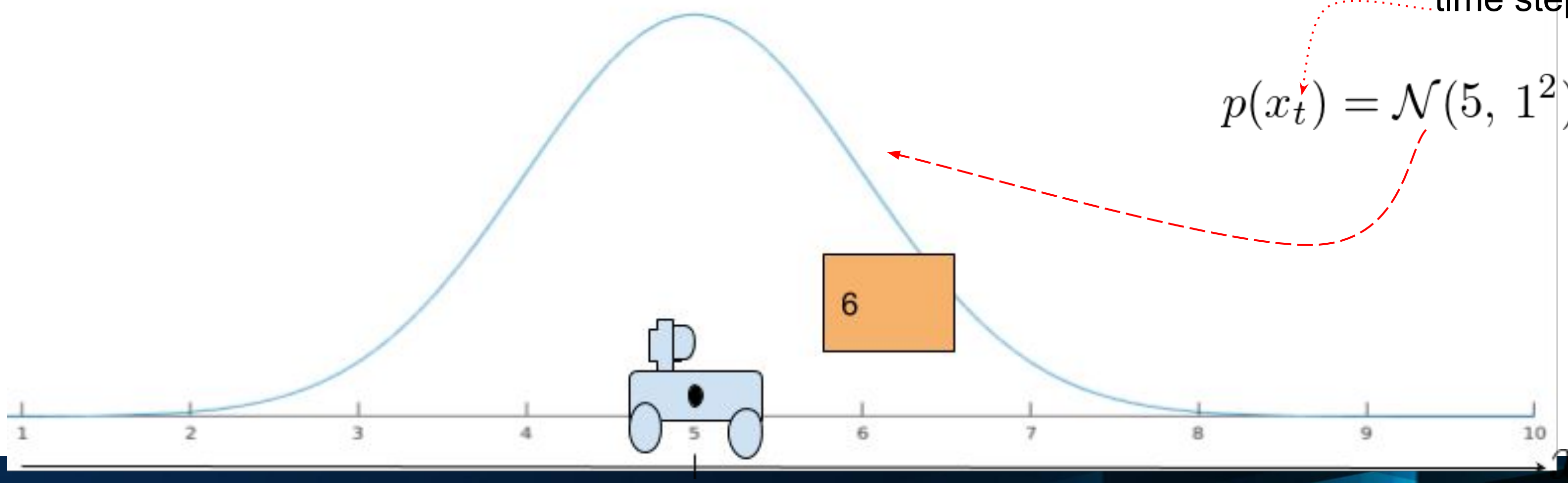
The most likely position of the robot is 5 and there is 68.2% chance that it is between 4 and 6.



Notation

Small **p** to denote density functions.

$p(X = x) = \mathcal{N}(\mu_X, \sigma_X^2)$ is usually written shorter: $p(x) = \mathcal{N}(\mu_X, \sigma_X^2)$.



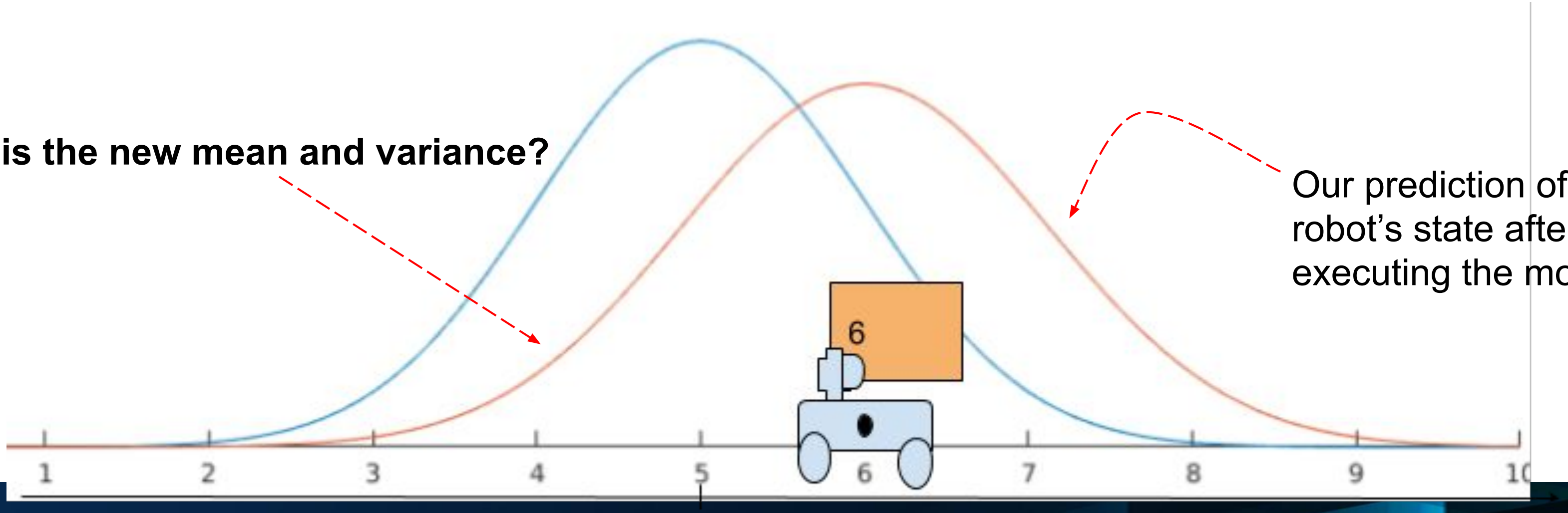
$$p(x_t) = \mathcal{N}(5,\ 1^2)\,.$$

Toy example - **prediction step**

Let's command the robot to move 1 m forward. Its motion is not accurate so we model the motion as a random variable $U_t \sim \mathcal{N}(\mu_U = 1, \sigma_U^2 = (0.5)^2)$

After the motion we get, $X_t = X_{t-1} + U_t$ which is normally distributed as well.

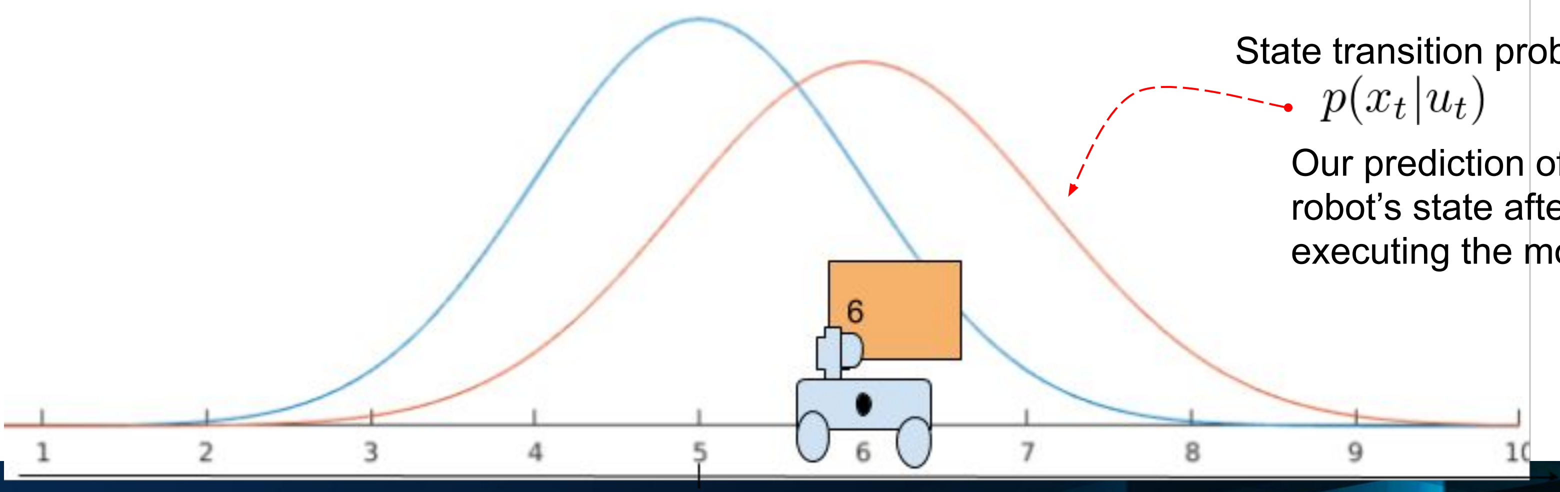
What is the new mean and variance?



Our prediction of the robot's state after executing the motion.

Toy example - **prediction step**

Given our previous estimate of the state and the motion command, we can predict the current state of the robot: $p(x_t|u_t) = \mathcal{N}(\mu_{X_{t-1}} + \mu_{U_t}, \sigma^2_{X_{t-1}} + \sigma^2_{U_t})$



State transition probability

$p(x_t|u_t)$

Our prediction of the robot's state after executing the motion.

Toy example - **measurement probability**

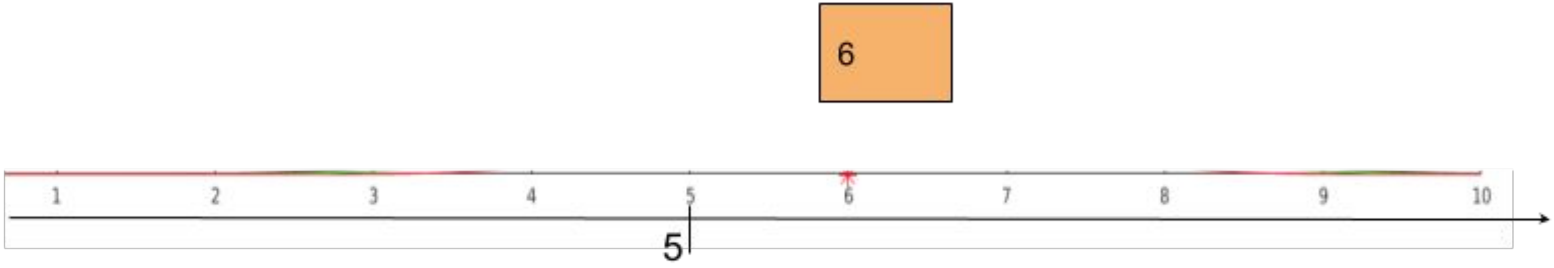
Our robot has a different sign sensor. It detects only the sign located at 6 m. However, the sensor is not perfect so we model it as follows:

$$p(z|x) = \frac{1}{\sqrt{2\pi\sigma_z}} \exp\left\{-\frac{1}{2}\left(\frac{z-x}{\sigma_z}\right)^2\right\} \qquad \sigma_z = 0.2, z = 6$$

- As a function of z: The probability of obtaining the measurement z given x.
- As a function of x: The likelihood of x given that it generated the measurement z.

The likelihood distribution

If the sensor on the robot detected the 6 m sign, what is the most likely position of the robot? What does the distribution looks like?



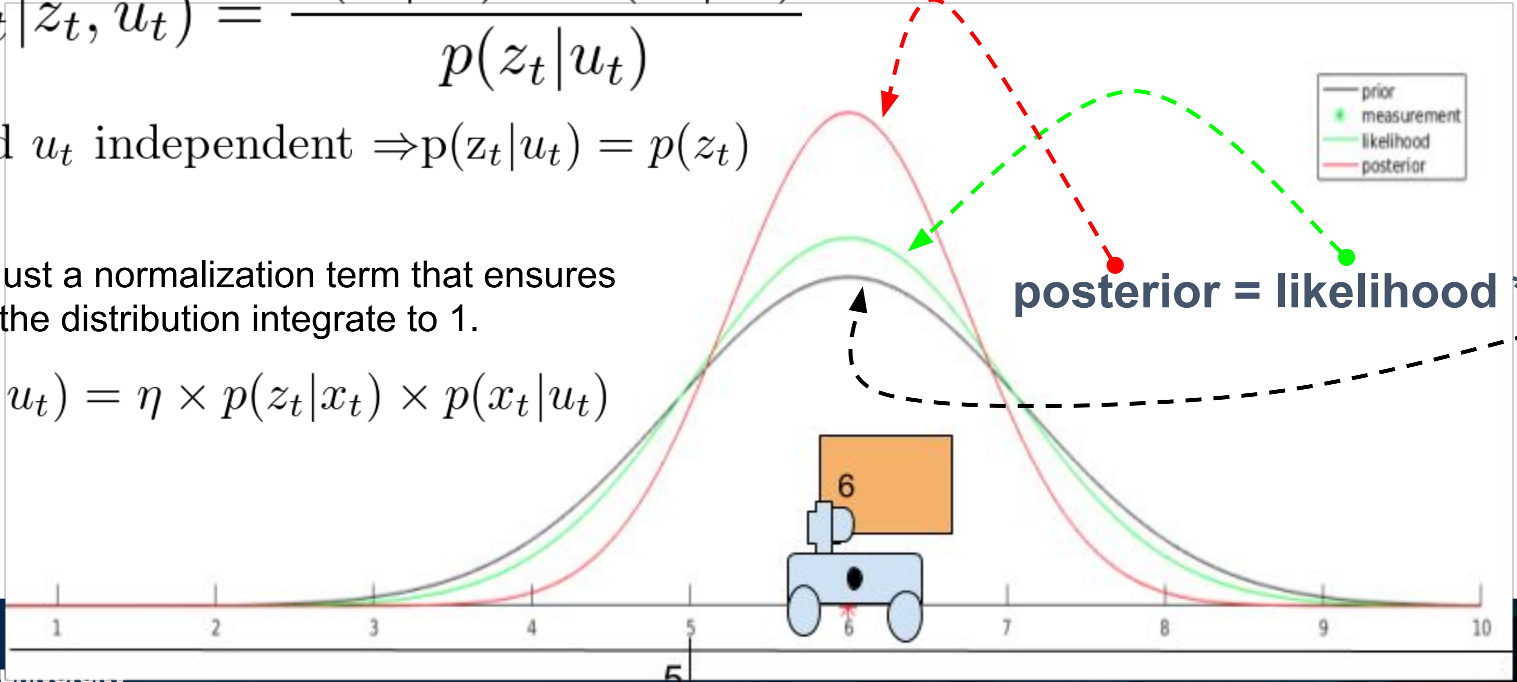
Toy example (**update step**)

$$p(x_t|z_t, u_t) = \frac{p(z_t|x_t) \times p(x_t|u_t)}{p(z_t|u_t)}$$

z_t and u_t independent $\Rightarrow p(z_t|u_t) = p(z_t)$

$p(z_t)$ is just a normalization term that ensures that the distribution integrate to 1.

$$p(x_t|z_t, u_t) = \eta \times p(z_t|x_t) \times p(x_t|u_t)$$



posterior = likelihood * prior

Toy example (**the posterior**)

We obtain the posterior (red) which is narrower than the other two.

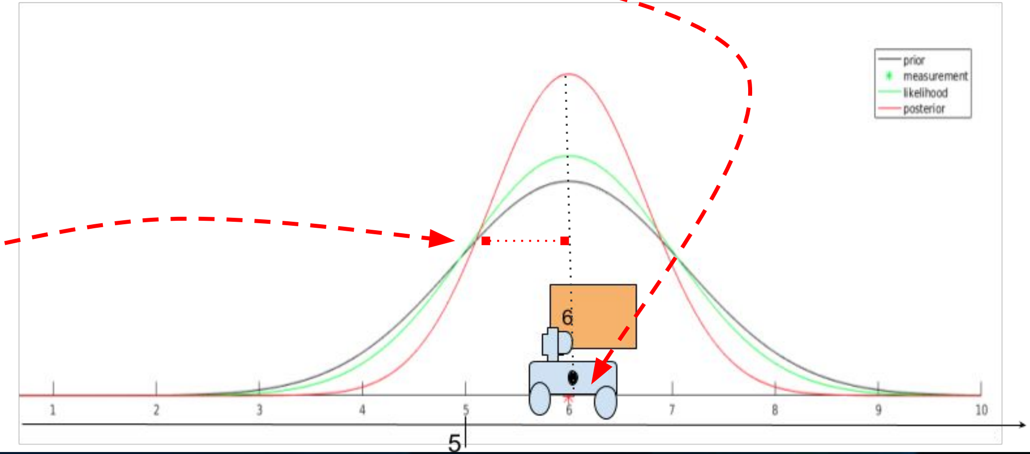
$$\mu = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2} z + \frac{\sigma_Z^2}{\sigma_X^2 + \sigma_Z^2} \mu_X$$

With some rearrangement:

$$\mu = \mu_X + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}(z - \mu_X)$$

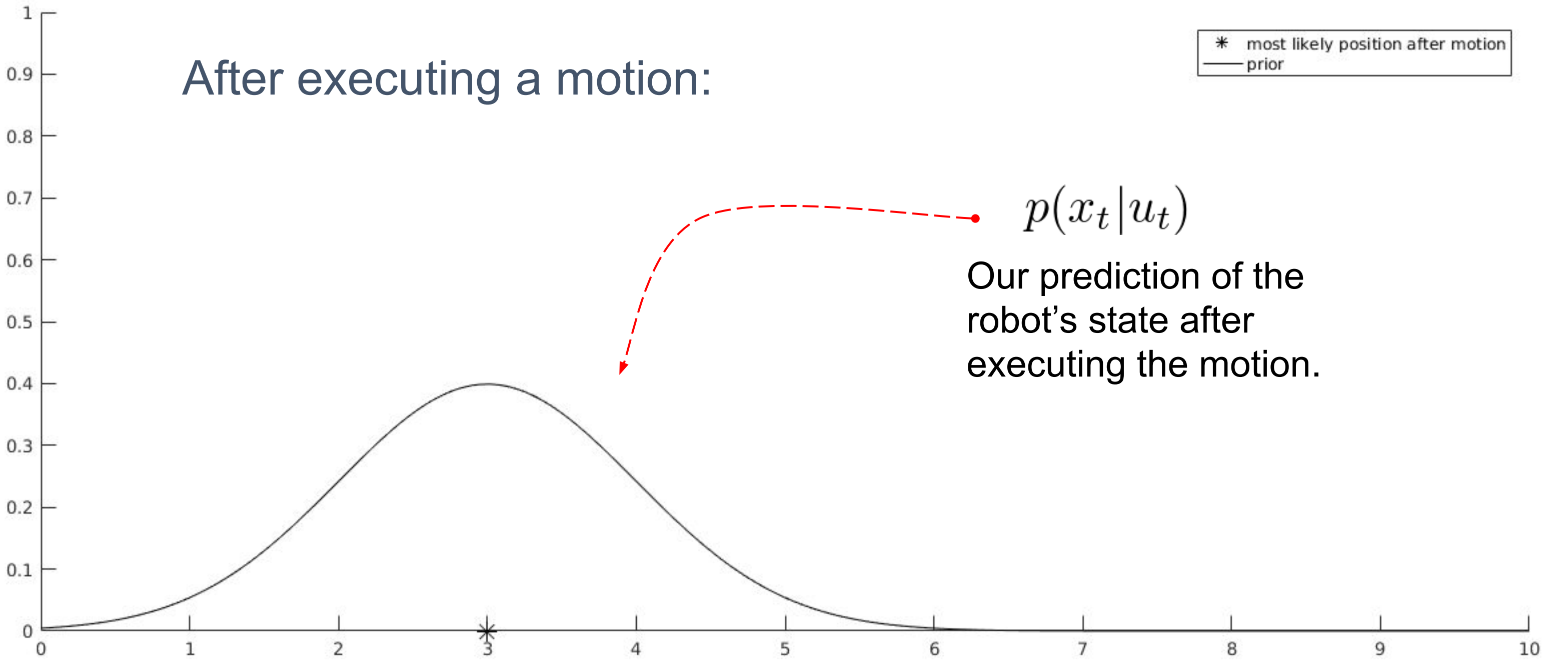
And the variance:

$$\sigma^2 = \frac{1}{\left(\frac{1}{\sigma_X^2} + \frac{1}{\sigma_Z^2}\right)}$$



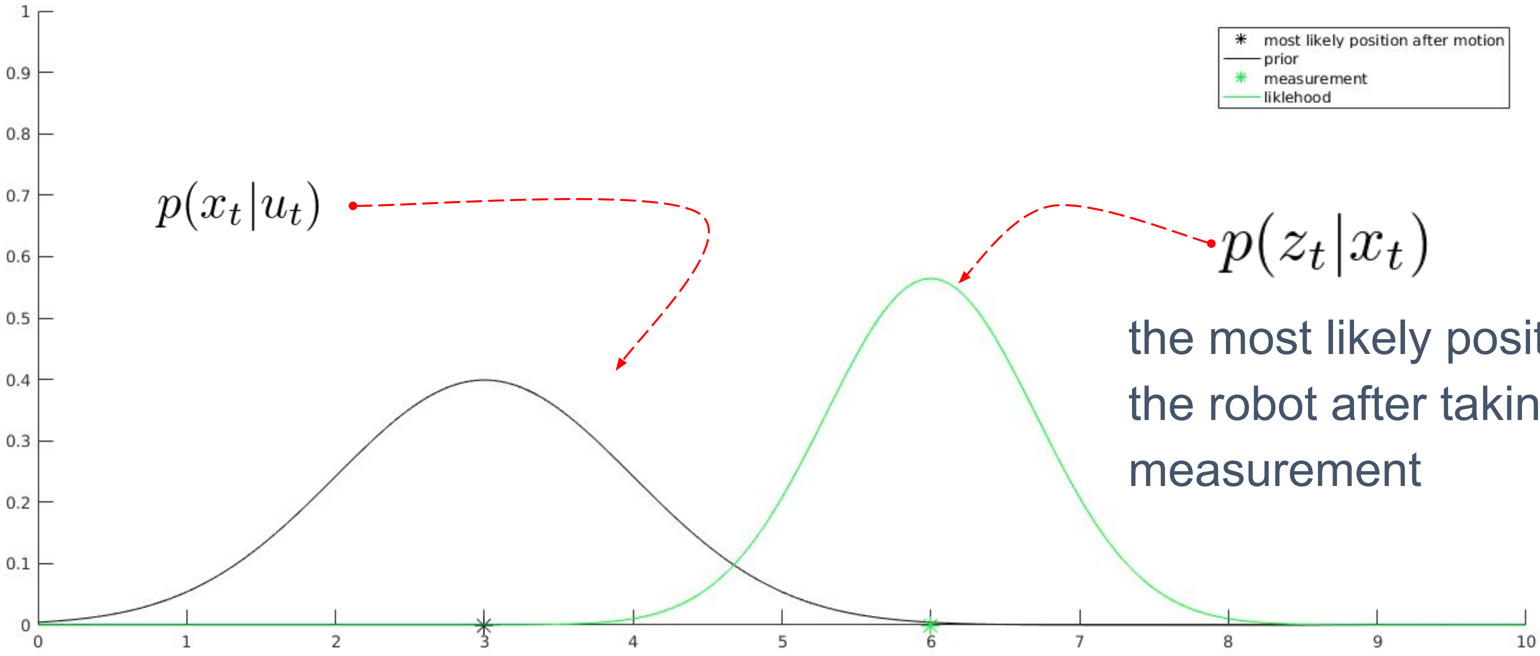
Let’s go over it again

After executing a motion:



$p(x_t|u_t)$

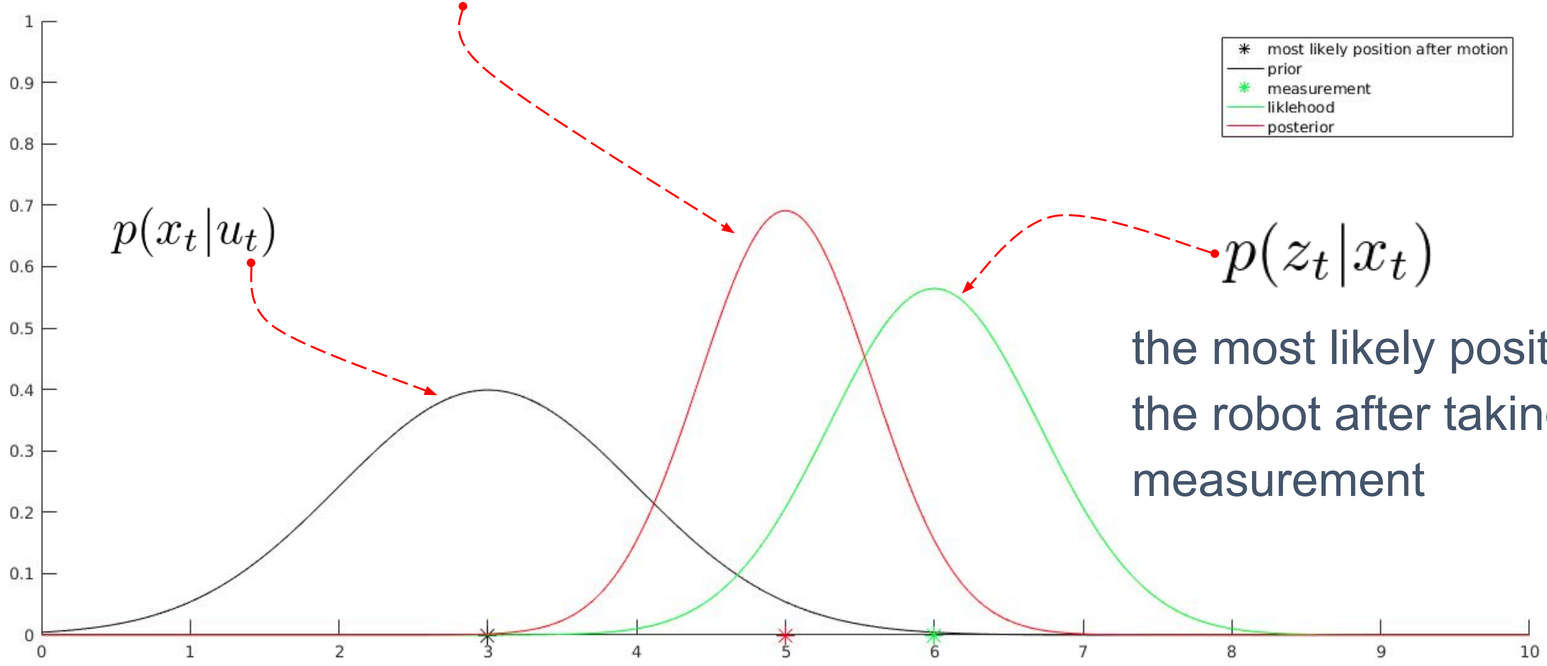
Our prediction of the robot's state after executing the motion.



the most likely position of the robot after taking a measurement

posterior = likelihood * prior

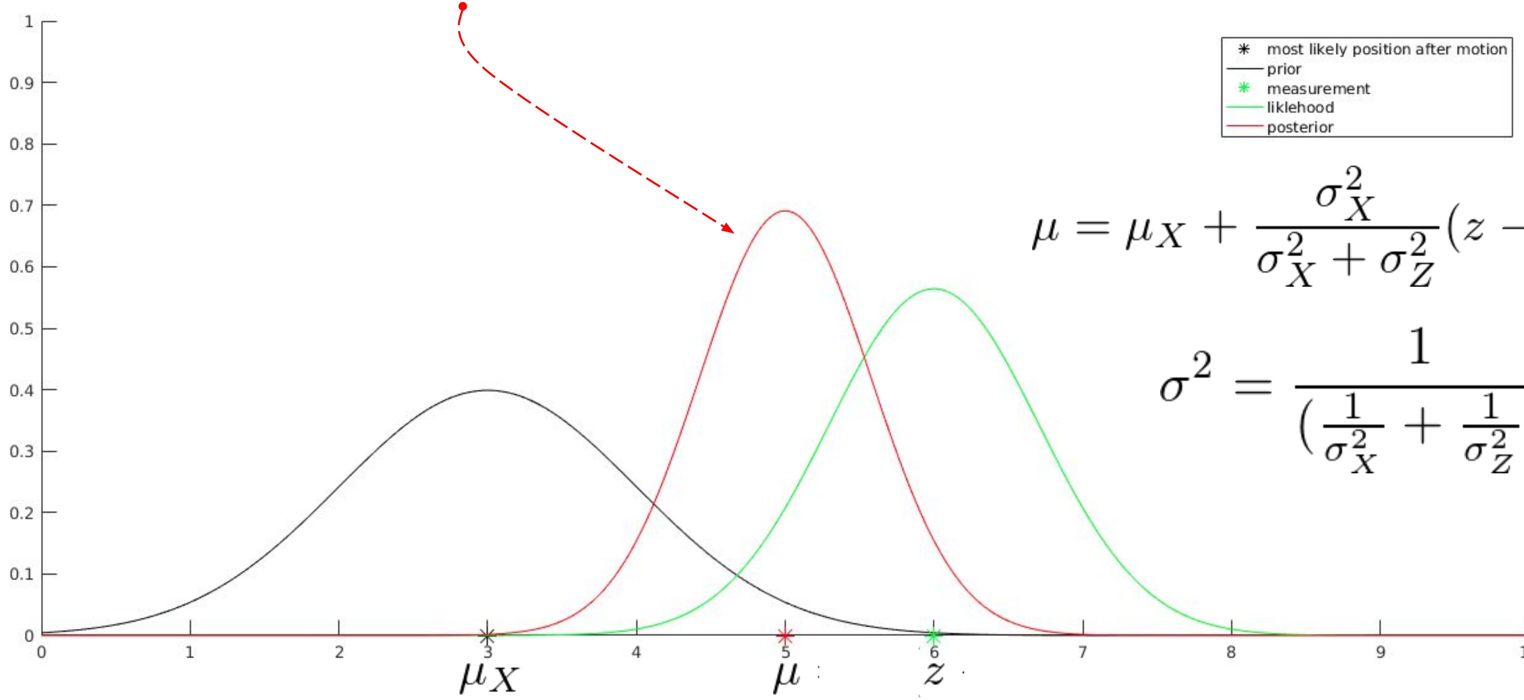
$$\mathrm{p}(\mathrm{x}_t|z_t, u_t) = \eta \times p(z_t|x_t) \times p(x_t|u_t)$$



the most likely position of the robot after taking a measurement

posterior = likelihood * prior

$$\mathrm{p}(\mathrm{x}_t|z_t, u_t) = \eta \times p(z_t|x_t) \times p(x_t|u_t)$$



$$\mu = \mu_X + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}(z - \mu_X)$$

$$\sigma^2 = \frac{1}{\left(\frac{1}{\sigma_X^2} + \frac{1}{\sigma_Z^2}\right)}$$

What we just did is called a 1D Kalman filter.

Next Lecture will be its formal introduction!