CS 215- Data Interpretation and Analysis
(Post Midsem)

Fuzzy cntd, ML and hypothesis testing

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Lecture-6

6nov23

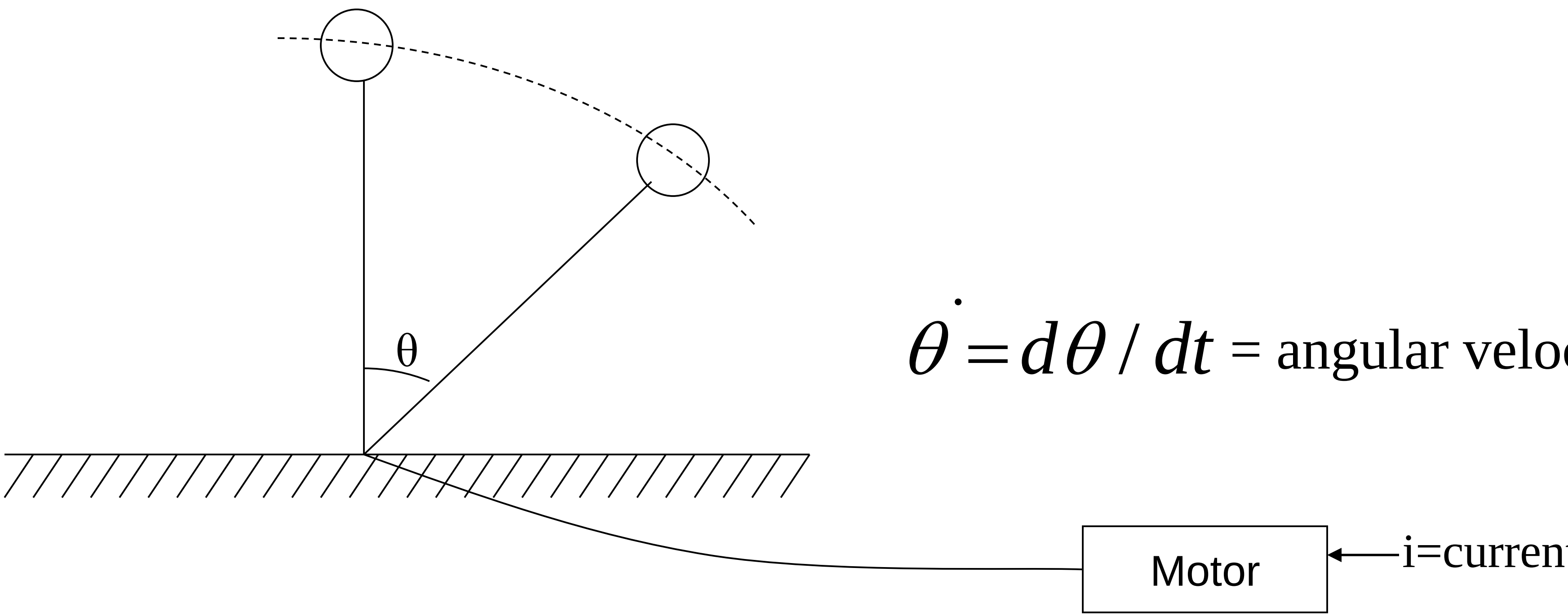
Recap

Application of Fuzzy Logic

An example

An Example

Controlling an inverted pendulum:



$$\dot{\theta} = d\theta / dt$$ = angular velocity

The goal: To keep the pendulum in vertical position ($\theta=0$) in dynamic equilibrium. Whenever the pendulum departs from vertical, a torque is produced by sending a current ‘i’

Controlling factors for appropriate current

Angle θ, Angular velocity $\dot{\theta}$

<u>Some intuitive rules</u>

If θ is +ve small and $\dot{\theta}$ is –ve small

then current is zero

If θ is +ve small and $\dot{\theta}$ is +ve small

then current is –ve medium

Control Matrix

$\dot{\theta}$ \ θ	-ve med	-ve small	Zero	+ve small	+ve med
-ve med					
-ve small		+ve med	+ve small	Zero	
Zero		+ve small	Zero	-ve small	
+ve small		Zero	-ve small	-ve med	
+ve med					

Region of interest

Each cell is a rule of the form

If θ is <> and $\dot{\theta}$ is <>

then i is <>

4 “Centre rules”

1. if θ = = Zero and $\dot{\theta}$ = = Zero then i = Zero
2. if θ is +ve small and $\dot{\theta}$ = = Zero then i is –ve small
3. if θ is –ve small and $\dot{\theta}$ = = Zero then i is +ve small
4. if θ = = Zero and $\dot{\theta}$ is +ve small then i is –ve small
5. if θ = = Zero and $\dot{\theta}$ is –ve small then i is +ve small

End Recap

ML and Hypothesis Testing

Bias detection

	English				
	Muril	**XLMR**	**Mbert**	**Bernice**	**IndicBERT**
Gender	47.08	**50.26**	47.61	56.61	52.38
Socioeconomic	58.77	64.03	**53.50**	54.38	63.15
Age	**49.15**	44.06	47.45	45.76	54.23
Physical-appearance	58.13	58.13	**53.48**	62.79	62.79
Disability	65.51	68.96	58.62	72.41	**51.72**
Crows Pair	52.87	55.17	**50.75**	56.09	56.55

From our AAAI submitted paper on Bias detection

Understanding comparison of means

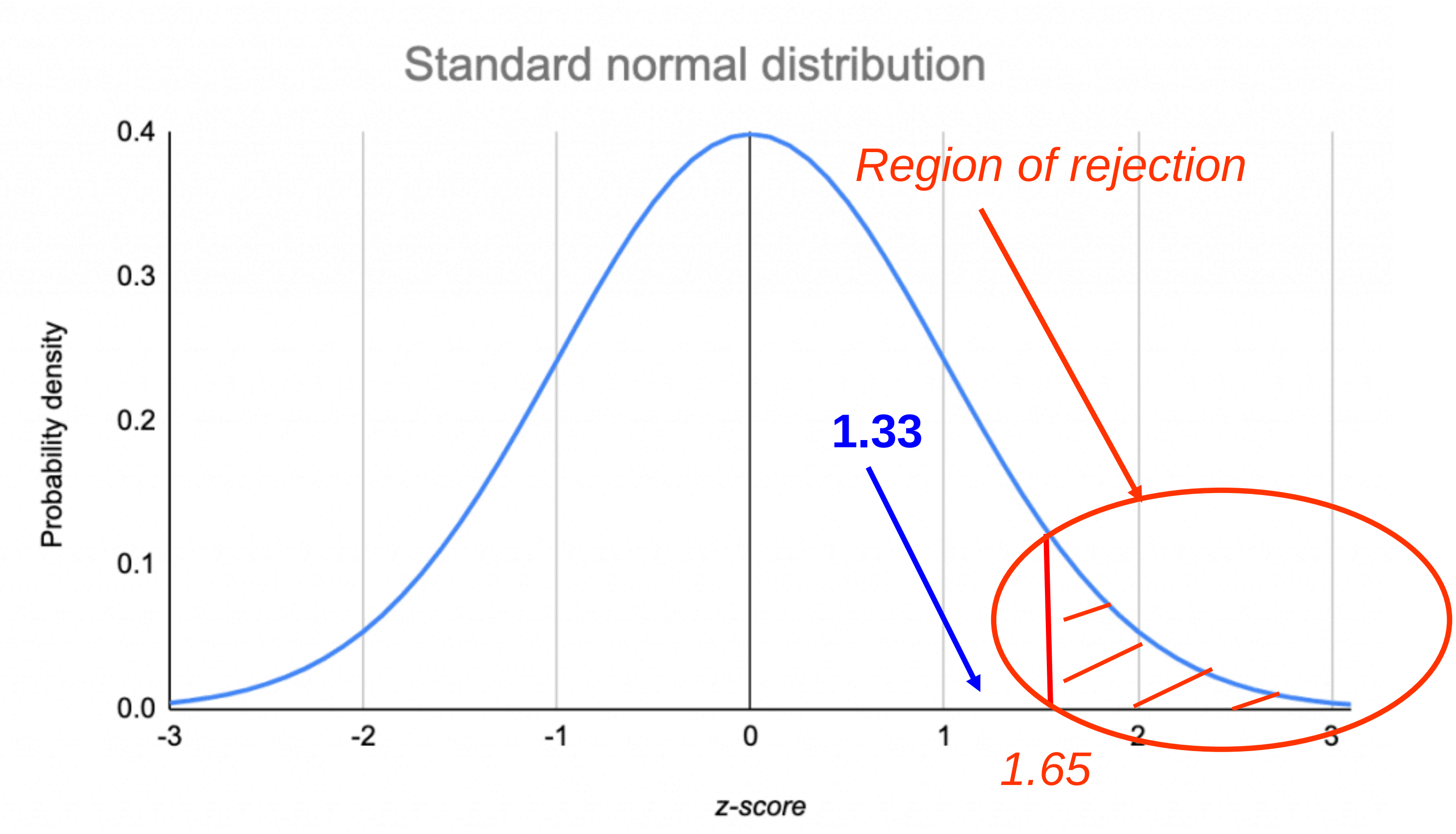
Allied Problem (comparison of means)

The mean height of 50 male students who showed above average participation in college athletics was 173.3 cm with a standard deviation of 6.4 cm, while 50 male students who showed no interest in such participation had a mean height of 171.5 cm with a standard deviation of 7.1cm. Is it the case that male students who participate in college athletics are taller than other male students?

Solution

- H_0: $\mu_1 = \mu_2$ (there is no diff in mean heights)
- H_A: $\mu_1 > \mu_2$
- Under H_0, $\mu_{\theta 1-\theta 2}=0$, θ_1 and θ_2 are sample means
- $\sigma_{\theta 1-\theta 2}= [(6.4)^2/50+ (7.1)^2/50]^{1/2}=1.35$
- Z
  - **$=[\{(\theta_1-\theta_2)-0\}/ \sigma_{\theta 1-\theta 2}]$**
  - **$=(173.3-171.5/1.35)=1.33$**

1-sided confidence interval (upper/right)



Conclusion from the test

- Cannot reject at 95% confidence level
- What about 99% confidence interval?
- $Z_c = 2.33$
- So cannot reject

What about 90% confidence interval?

- 1.33 > 1.28
- Can reject H_0

- What does this mean?
  - Cannot reject at 95% and 99% CI, but can reject at 90% (more permissive)

Interpretation

- What does 90% confidence interval mean?
- Linked with Type-I error
- Probability of Type-I error is 10%
- Type-I error ➔ reject the hypothesis when it should be accepted

Level of significance

- Given a hypothesis, the maximum probability with which we would be willing to risk a Type-I error is called the level of significance
- In case of “participation” in athletics what can we say wrt the levels of significance

α= 10%

- In 10% cases, 1-in-10, 10-in-100, 100-in-1000, 1000-in-10000, we will make an error, if we reject H_0 when we should have accepted it

- Recall, H0: there is no difference in mean height of those who do athletics and those who do not.

Interpretation cntd.

- If we are <u>less</u> strict (α=0.10), i.e., prepared to tolerate 10 errors in 100 case, we will reject the stand that that there is no diff in the mean height
- If we are <u>more</u> strict (α=0.05)- prepared to tolerate only 5 errors in 100- we will not reject the stand
- If we are <u>very</u> strict (α=0.01)- prepared to tolerate only 1 errors in 100- we will not reject the stand

Back to Bias detection

	English				
	Muril	**XLMR**	**Mbert**	**Bernice**	**IndicBERT**
Gender	47.08	**50.26**	47.61	56.61	52.38
Socioeconomic	58.77	64.03	**53.50**	54.38	63.15
Age	**49.15**	44.06	47.45	45.76	54.23
Physical-appearance	58.13	58.13	**53.48**	62.79	62.79
Disability	65.51	68.96	58.62	72.41	**51.72**
Crows Pair	52.87	55.17	**50.75**	56.09	56.55

Parameters for MURIL data

- (drop decimels) 47, 58, 49, 58, 65, 52
- Mean = (47 + 58 + 49 + 58 + 65 + 52)/6 = 329/6= 54.83
- Variance=Mean of squared deviations = 37.84
- Standard deviation=$\sqrt{37.84}$=6.16

Parameters for INDICBERT data

- (drop decimels) 52, 63, 54, 62, 51, 56
- Mean = (52 + 63 + 54 + 62 + 51 + 56)/6 = 338/6=56.33
- Variance=Mean of squared deviations = 21.56
- Standard deviation=$\sqrt{21.56}$=4.65

Hypothesis testing on bias

- H_0: $\mu_1 = \mu_2$ (there is no diff in mean biasedness)
- H_A: $\mu_1 <> \mu_2$
- Under H_0, $\mu_{\theta 1-\theta 2}=0$, θ_1 and θ_2 are sample means
- $\sigma_{\theta 1-\theta 2}= [37.84/6+ 21.56/6]^{1/2}=3.15$
- Z

$=[\{(\theta_1-\theta_2)-0\}/ \sigma_{\theta 1-\theta 2}]$

$=[(54.83-56.33)/3.15]=-0.48$

Examination of H_0

test-type (col)			
vs.	Two-Tail	1 sided to +inf	1 sided from -inf
Confidence Interval (significance level)			
90% (0.10)	(- and +) 1.65	-1.28 to +inf	-inf to +1.28
95% (0.05)	(- and +) 1.96	-1.65 to +inf	-inf to +1.65
99% (0.01)	(- and +) 2.58	-2.33 to +inf	-inf to 2.33

- Cannot reject H_0: for 95% CI, since -1.96<-0.48<+1.96
- Nor for 99%, nor for 90%

Conclusion for relative biasedness of MURIL and INDICBERT

- There is 95% probability that MURIL and INDICBERT are equally biased

	English				
	Muril	**XLMR**	**Mbert**	**Bernice**	**IndicBERT**
Gender	47.08	**50.26**	47.61	56.61	52.38
Socioeconomic	58.77	64.03	**53.50**	54.38	63.15
Age	**49.15**	44.06	47.45	45.76	54.23
Physical-appearance	58.13	58.13	**53.48**	62.79	62.79
Disability	65.51	68.96	58.62	72.41	**51.72**
Crows Pair	52.87	55.17	**50.75**	56.09	56.55

Back to control of inverted pendulum

Control Matrix

$\dot{\theta}$ \ θ	-ve med	-ve small	Zero	+ve small	+ve med
-ve med					
-ve small		+ve med	+ve small	Zero	
Zero		+ve small	Zero	-ve small	
+ve small		Zero	-ve small	-ve med	
+ve med					

Region of interest

Each cell is a rule of the form

If θ is <> and $\dot{\theta}$ is <>

then i is <>

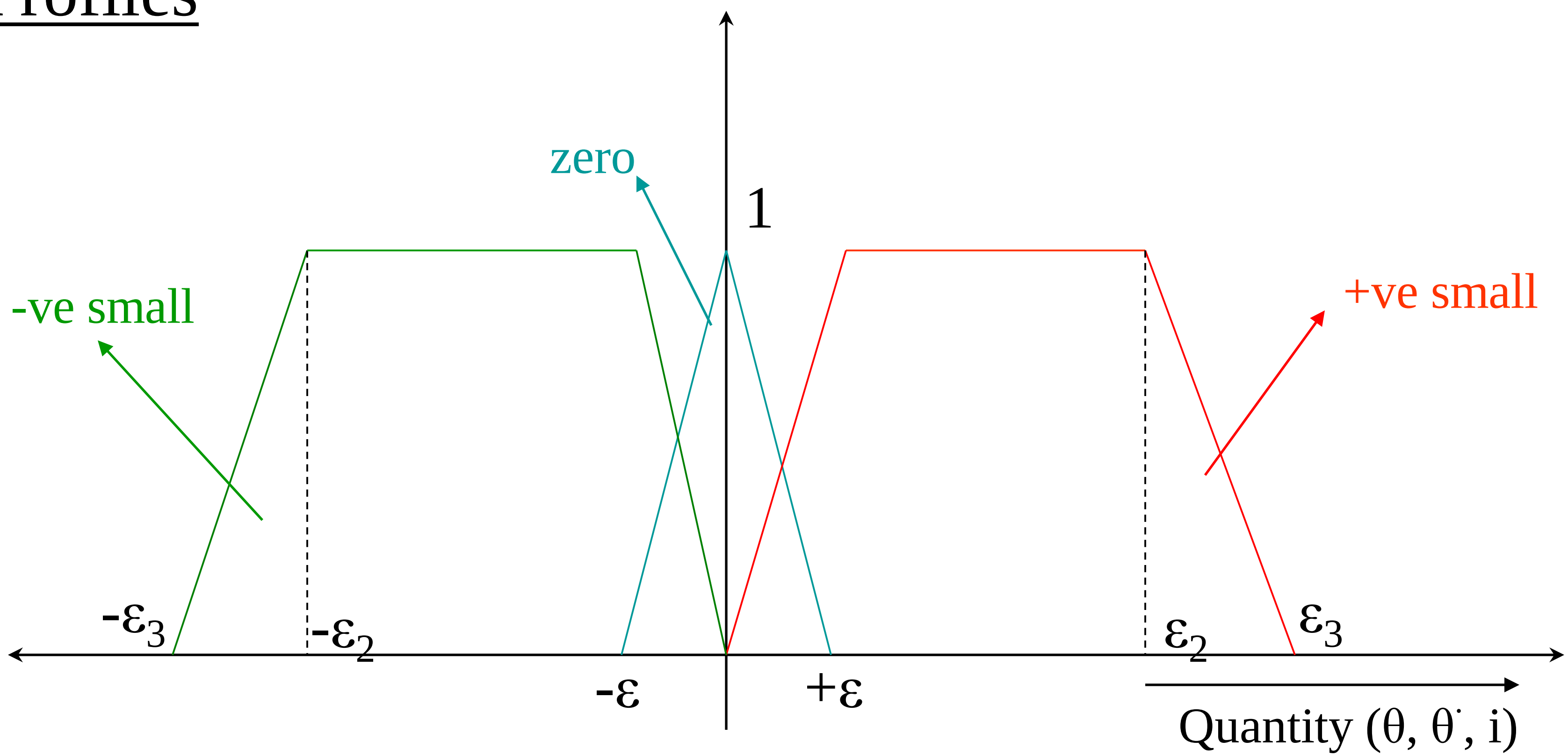
<u>4 “Centre rules”</u>

1. if θ = = Zero and $\dot{\theta}$ = = Zero then i = Zero
2. if θ is +ve small and $\dot{\theta}$ = = Zero then i is –ve small
3. if θ is –ve small and $\dot{\theta}$ = = Zero then i is +ve small
4. if θ = = Zero and $\dot{\theta}$ is +ve small then i is –ve small
5. if θ = = Zero and $\dot{\theta}$ is –ve small then i is +ve small

Linguistic variables

1. Zero
2. +ve small
3. -ve small

Profiles

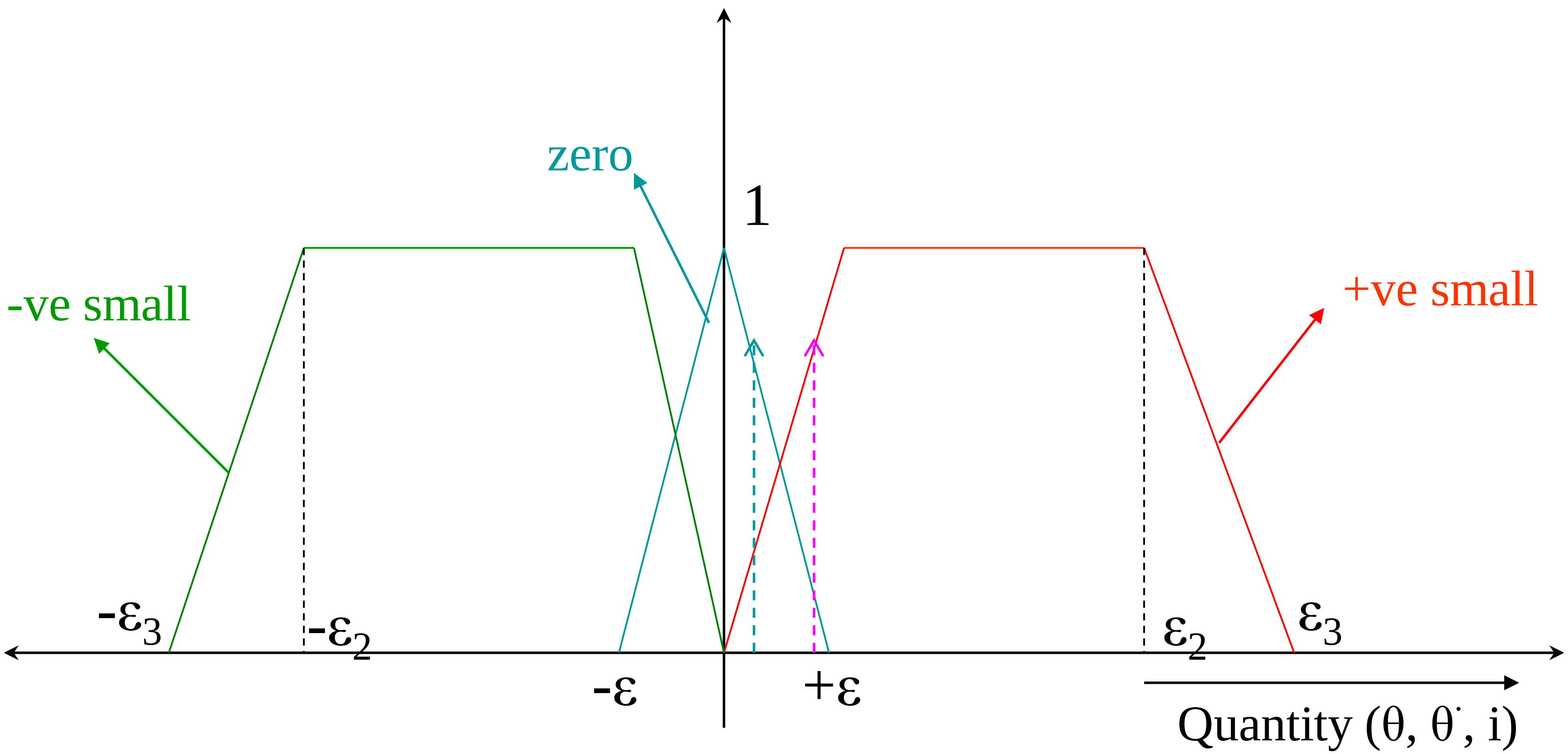


Inference procedure

1. Read actual numerical values of θ and $\dot{\theta}$
2. Get the corresponding μ values μ_{Zero}, $\mu_{(+ve\ small)}$, $\mu_{(-ve\ small)}$. This is called FUZZIFICATION
3. For different rules, get the fuzzy I-values from the R.H.S of the rules.
4. “Collate” by some method and get ONE current value. This is called DEFUZZIFICATION
5. Result is one numerical value of ‘i’.

Rules Involved

if θ is Zero and dθ/dt is Zero then i is Zero
if θ is Zero and dθ/dt is +ve small then i is –ve small
if θ is +ve small and dθ/dt is Zero then i is –ve small
if θ +ve small and dθ/dt is +ve small then i is -ve medium



Fuzzification

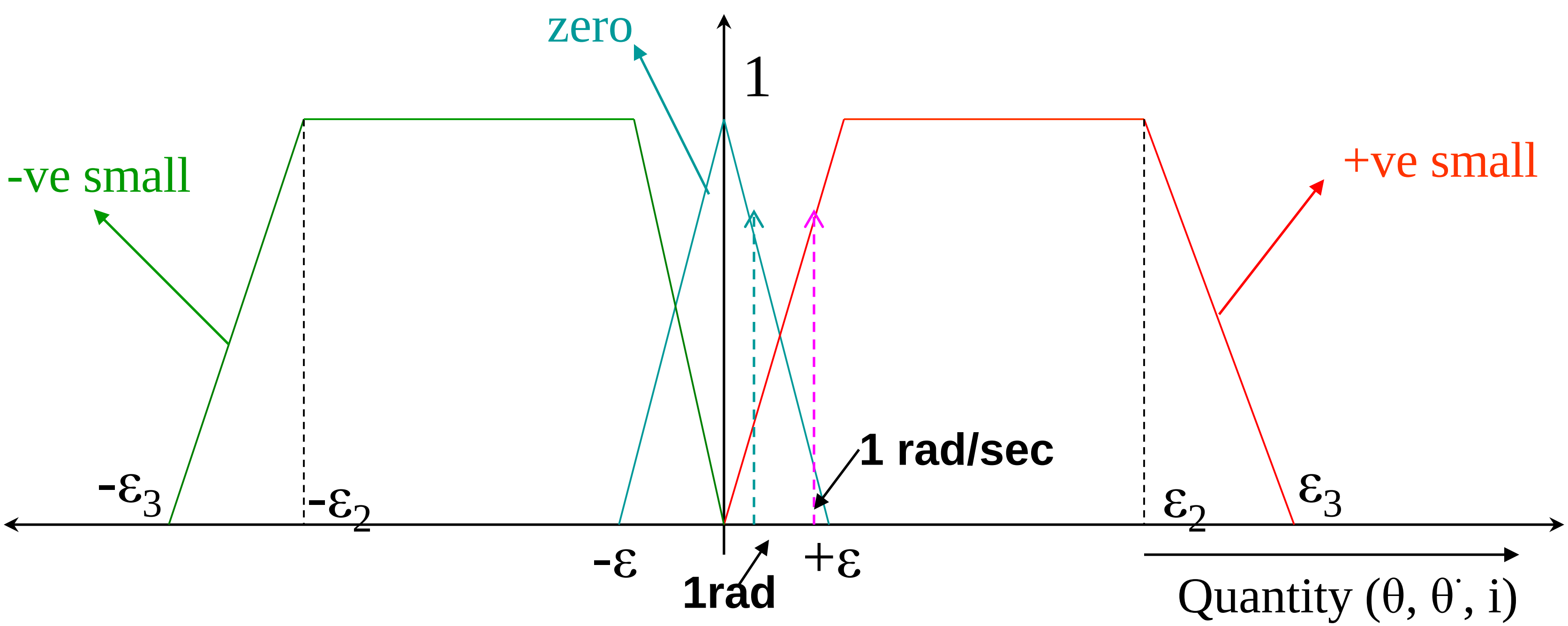
Suppose θ is 1 radian and dθ/dt is 1 rad/sec

μ_{zero}(θ =1)=0.8 (say)

$M_{+ve\text{-}small}$(θ =1)=0.4 (say)

μ_{zero}(dθ/dt =1)=0.3 (say)

$\mu_{+ve\text{-}small}$(dθ/dt =1)=0.7 (say)



Fuzzification

Suppose θ is 1 radian and dθ/dt is 1 rad/sec

$\mu_{zero}(\theta = 1) = 0.8$ (say)

$\mu_{+ve\text{-}small}(\theta = 1) = 0.4$ (say)

$\mu_{zero}(d\theta/dt = 1) = 0.3$ (say)

$\mu_{+ve\text{-}small}(d\theta/dt = 1) = 0.7$ (say)

if θ is Zero and dθ/dt is Zero then i is Zero

min(0.8, 0.3)=0.3

hence $\mu_{zero}(i) = 0.3$

if θ is Zero and dθ/dt is +ve small then i is –ve small

min(0.8, 0.7)=0.7

hence $\mu_{-ve\text{-}small}(i) = 0.7$

if θ is +ve small and dθ/dt is Zero then i is –ve small

min(0.4, 0.3)=0.3

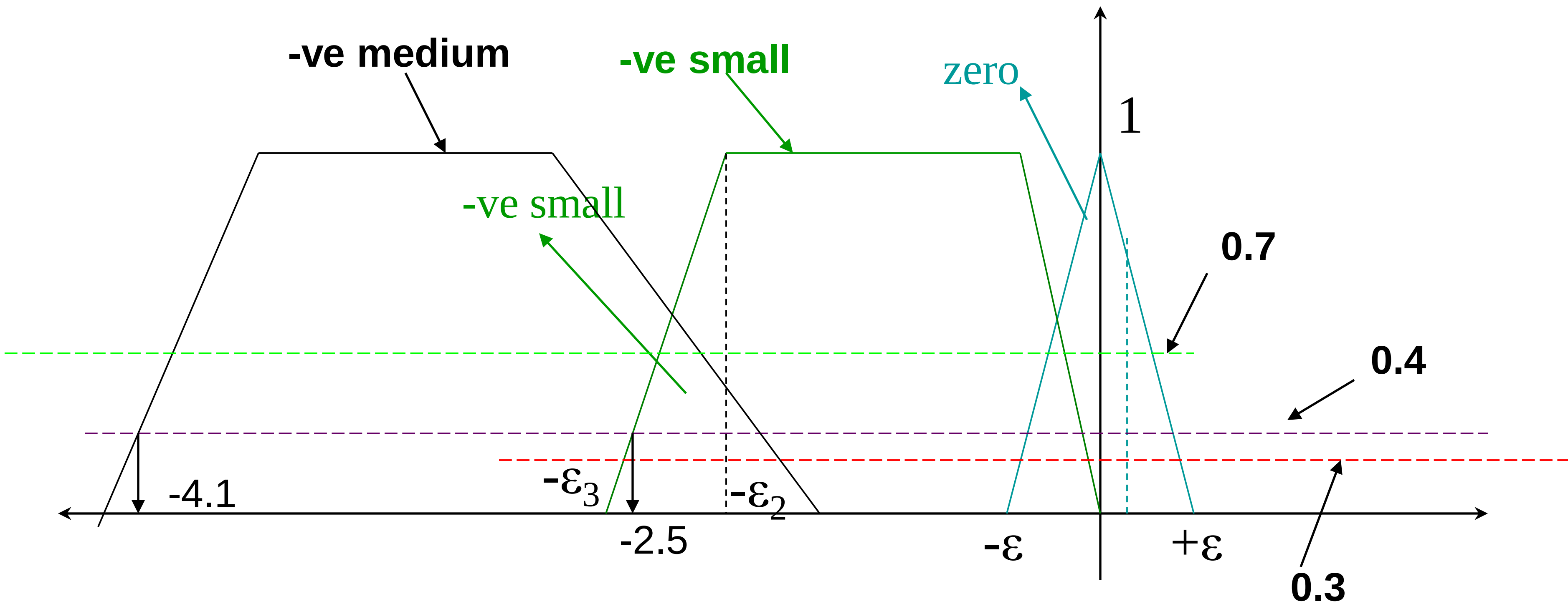
hence μ-ve-small(i)=0.3

if θ +ve small and dθ/dt is +ve small then i is -ve medium

min(0.4, 0.7)=0.4

hence $\mu_{-ve\text{-}medium}(i) = 0.4$

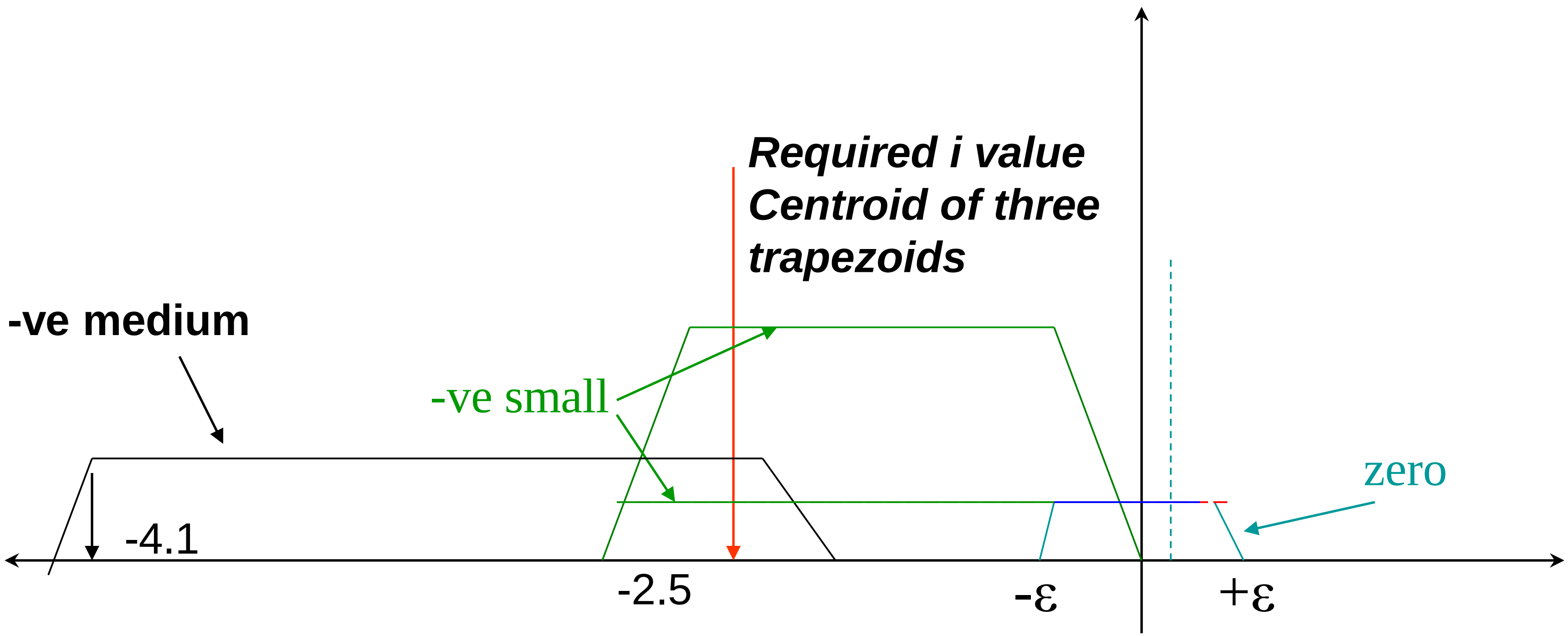
Finding *i*



Possible candidates:

i=0.5 and -0.5 from the "zero" profile and μ=0.3
i=-0.1 and -2.5 from the "-ve-small" profile and μ=0.3
i=-1.7 and -4.1 from the "-ve-small" profile and μ=0.3

Defuzzification: Finding *i* by the *centroid* method



Possible candidates:

i is the x-coord of the centroid of the areas given by the blue trapezium, the green trapeziums and the black trapezium

How to define subset hood?

Meaning of fuzzy subset

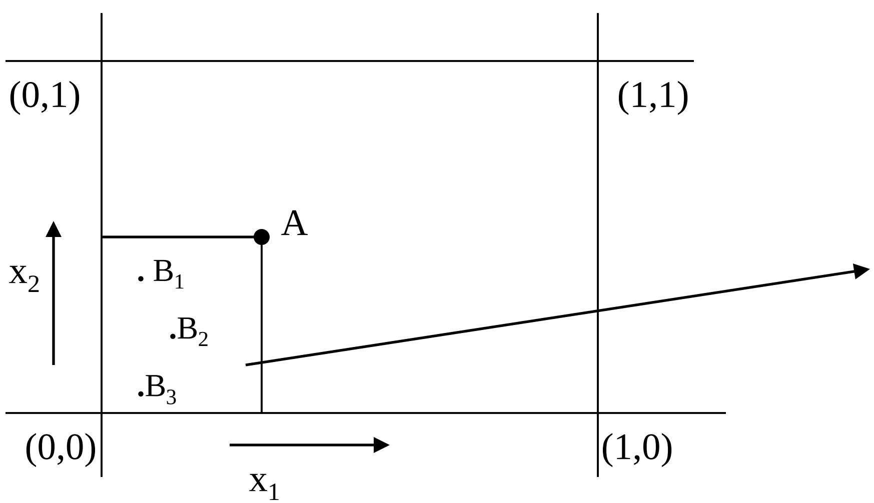
Suppose, following classical set theory we say

$$B \subset A$$

if

$$\mu_B(x) \le \mu_A(x) \forall x$$

Consider the n-hyperspace representation of A and B



Region where $\mu_B(x) \le \mu_A(x)$

This effectively means

$B \in P(A)$ CRISPLY

$P(A)$ = Power set of A

Eg: Suppose

A = {0,1,0,1,0,1,……………….,0,1} – 10^4 elements

B = {0,0,0,1,0,1,…………………..,0,1} – 10^4 elements

Isn't $B \subset A$ with a degree? (only differs in the 2nd element)

Subset operator is the “odd man” out

- AUB, A∩B, A^c are all “Set Constructors” while $A \subseteq B$ is a Boolean Expression or predicate.
- According to classical logic
  - In Crisp Set theory $A \subseteq B$ is defined as

    $$\forall x \quad x \in A \Rightarrow x \in B$$

  - So, in fuzzy set theory $A \subseteq B$ can be defined as

    $$\forall x \quad \mu_A(x) \Rightarrow \mu_B(x)$$

Zadeh's definition of subsethood goes against the grain of fuzziness theory

- Another way of defining $A \subseteq B$ is as follows:

$$\forall x \quad \mu_A(x) \le \mu_B(x)$$

But, these two definitions imply that $\mu_{P(B)}(A)=1$
where P(B) is the power set of B

Thus, these two definitions violate the fuzzy principle that every belongingness except Universe is fuzzy