# FORMAL QUANTIFICATION OF KNOWLEDGE

# INTRODUCTION

Purpose:

Describe and illustrate methods for formulating uncertainty distributions and quantifying expert information / knowledge (Phase 5)

Overview:

- Forming uncertainty distributions
- Quantification methods using probability and fuzzy logic
- Aggregation of multiple experts

# EXPERT JUDGMENT AS "DATA"

**Expert judgment shares traits with data from tests, experiments, or physical observations.**

- It is affected by the process of gathering it
- It has uncertainty, which can be characterized and subsequently analyzed.
- It can be conditioned on various factors, such as
  - the phrasing of the question,
  - the information the experts considered,
  - the experts' methods of solving the problem, and
  - the experts' assumptions.
- It can be combined with other information/data.

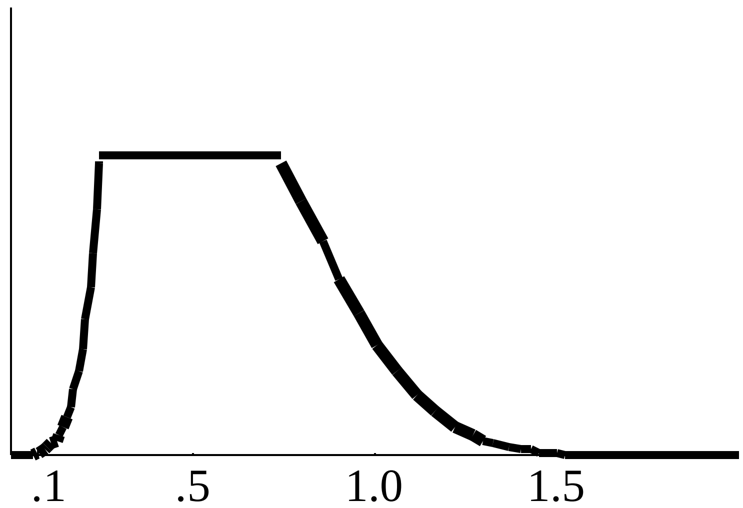
# UNCERTAINTY DISTRIBUTIONS

Distributions can be formulated by:

- Having the expert draw a distribution
- Using elicited moments, parameters, or quantiles
- Using elicited membership functions

# DRAWING DISTRIBUTIONS

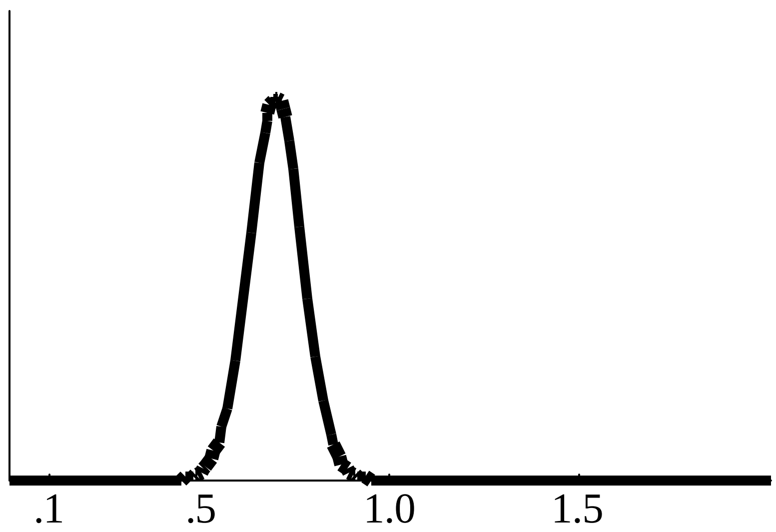
Drawing—sometimes an expert will understand what a probability density function is and be able to draw the shape of one free-hand.



# FORMULATING DISTRIBUTIONS

Moments—while an expert might be able to estimate a mean, it is extremely rare that he/she would be able to estimate a standard deviation or variance. As such, studies do not recommend this estimation.

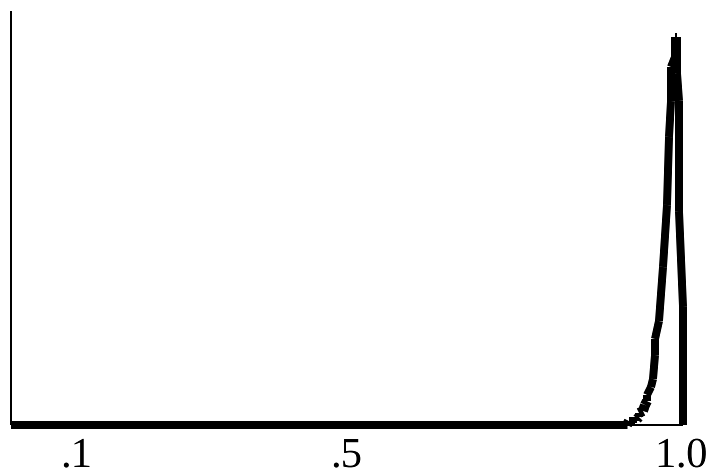
Distribution is normal with
a mean of 0.7
and a standard deviation
10% of mean



# FORMULATING DISTRIBUTIONS

Parameters—rarely can parameters be directly estimated by experts. One such possible case is with distributions whose parameters have interpretations (e.g., 1st beta parameter can be number of successes, and the 2nd parameter can be number of failures).

Beta:
1st= 98 successes
in 100 trials
2nd= 2 failures

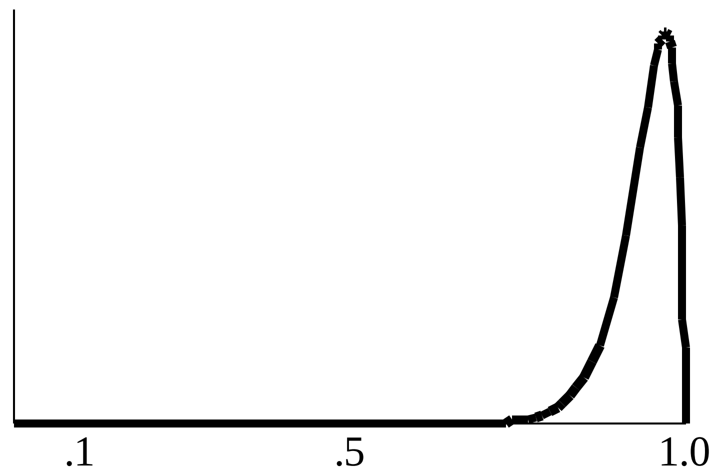


# FORMULATING DISTRIBUTIONS

Quantiles—most common. Experts do well in estimating the median as the most likely value or as their best estimate. Studies indicate if an expert provides a mean, it often is a median. Ranges of values (best/worst or max/min) are good for estimating uncertainty; however take into account the experts to underestimation of uncertainty bias.

$0 \leq p \leq 1$

$p_{max}=0.99$, $p_{min}=0.85$



# ELICITED QUANTILES

Suppose that we have elicited three "quantiles" from our expert:

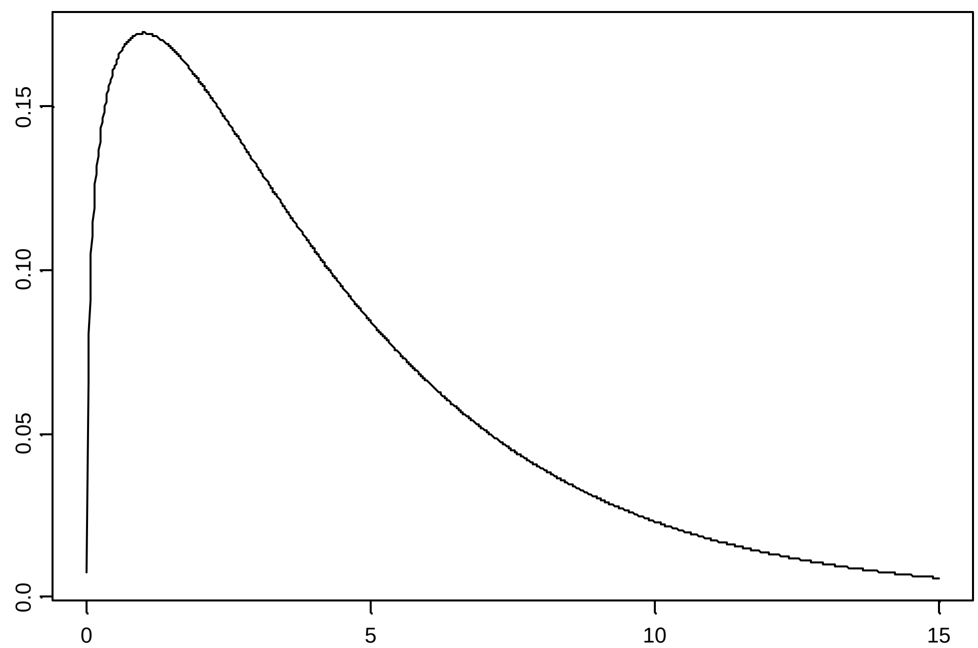
"Minimum" = 0.5
"Most likely = 3.0
"Maximum" = 12.0

Now what?

# FIT A SINGLE DISTRIBUTION

- Pick a parametric family. In this case, since the distribution is not symmetric, apparently positive, and not constrained between (0,1), we will use a gamma distribution.
- Decide what "minimum," "most likely," and "maximum" mean. In this case, we will take them to be the 5$^{th}$ percentile, median, and 95$^{th}$ percentile.
- Choose parameters for the gamma distribution (by trial and error) that come pretty close to having these quantiles.

# FIT A SINGLE DISTRIBUTION

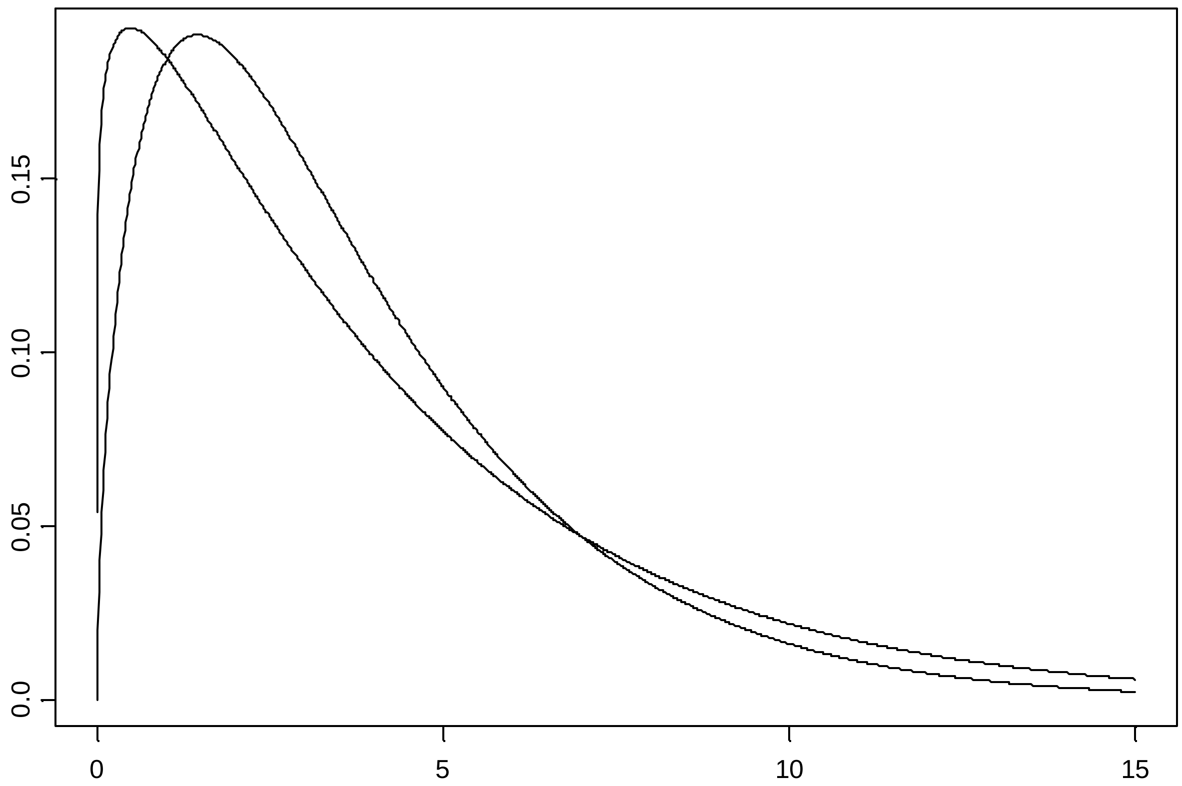



A Gamma(1.3,0.3) has 0.5 as the $7^{th}$ percentile, 3.0 as the $46^{th}$ percentile, and 12.0 as the $95^{th}$ percentile.

# FIT USING MORE DISTRIBUTIONS

- Pick a parametric family. In this case, since the distribution is not symmetric, apparently positive, and not constrained between (0,1), we will use a gamma distribution.
- Decide what “minimum,” “most likely,” and “maximum” mean. In this case, we will take them to be the 5$^{th}$ percentile, median, and 95$^{th}$ percentile.
- Fit two distributions: one to the “minimum” and “most likely” and one to the “most likely” and “maximum.” You can force the percentiles to fit, or not. (We do).

# FIT USING MORE DISTRIBUTIONS



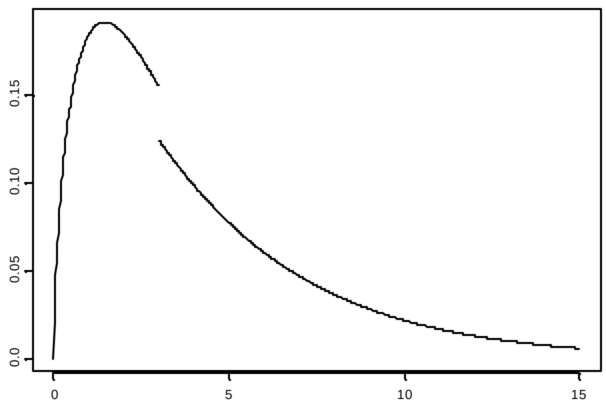

In this case, a Gamma(1.625,0.43) fits “minimum” and “most likely” and a Gamma(1.13, 0.27) fits “maximum” and “most likely”.

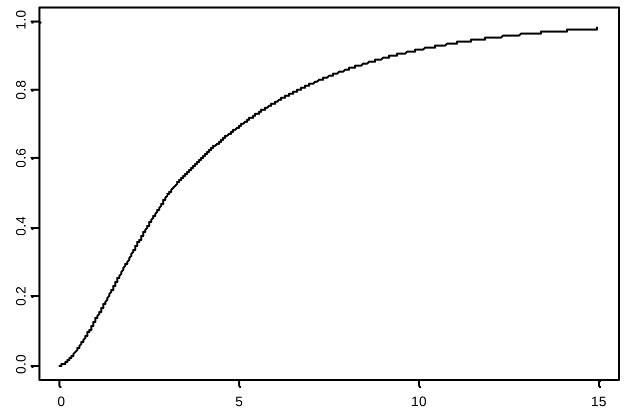
# FIT USING MORE DISTRIBUTIONS

How do we combine the two distributions?

- Derive the distribution of $p_1X + p_2Y$, where X and Y have the two distributions and $p_1 + p_2 = 1$.
- Take a weighted average of the density functions.

# HOW DO WE COMBINE THE DISTRIBUTIONS?




Use half of one distribution and half of the other. This produces the density on the left and the distribution function on the right. You can smooth the “bump” in the distribution function near 3.0 and then form another density, which may not quite be Gamma any more.

# HOW DO WE COMBINE THE DISTRIBUTIONS?

Try several of these possibilities and ask the expert you elicited the data from to pick the one that looks right. There may or may not be much difference.

# NON-PROBABILISTIC QUANTIFICATION TOOLS

Fuzzy logic control systems methods based on fuzzy set theory can help quantify such rules.

*If* the temperature is too hot, *then* this component is not going to work very well.

Map Condition (too hot) into
Performance (won't work very well)

# FORMULATING DISTRIBUTIONS

Membership functions—turns rules into numerical functions. Designed for capturing a classification type of uncertainty. From fuzzy set theory: an alternative calculus for uncertainties (imprecision) introduced by Lotfi Zadeh in 1965.

**Membership function of a subset, A, $\mu$A(*x*), is almost always (but not necessarily) a number between 0 and 1 that reflects the extent to which *x* $\in$ A. Experts assign these numbers to each *x*, for all subsets of interest. The set is called a *fuzzy set*. For (our usual) crisp sets; $\mu$A(*x*) = 0 or 1 for all *x*.**

# MEMBERSHIP FUNCTION (MF) EXERCISE

Consider the set of integers $X = \{0, 1, 2, \ldots, 10\}$.

Define a subset, of $X$, where

$A = \{x : x \in X$ and $x$ is "medium"$\}$

# MF EXERCISE

What number(s) would be “medium” ?

What number(s) would be half “medium” and half “large”?

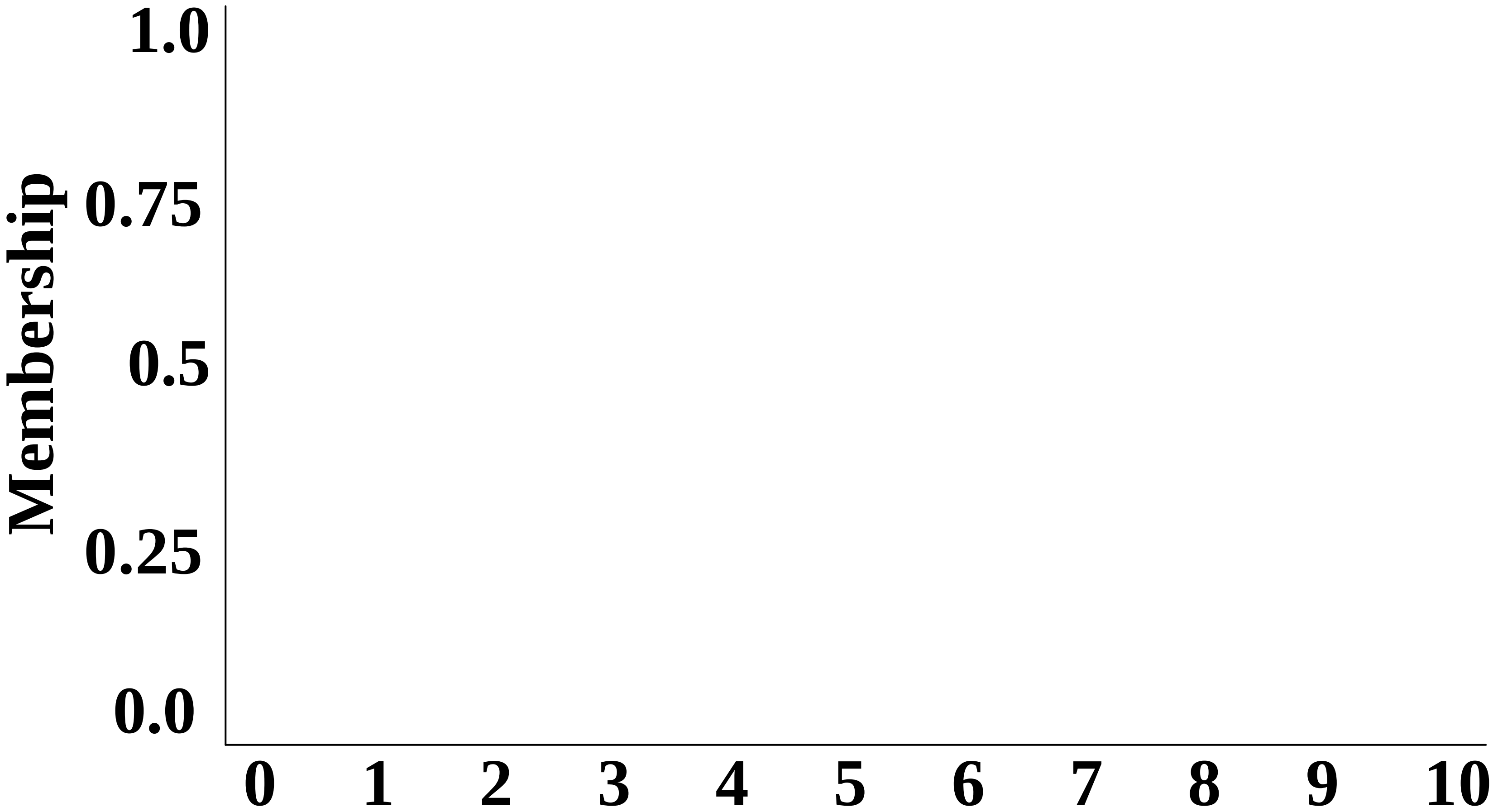
What number(s) would be half “medium” and half “small”?

What number(s) would be more “medium” than “large”?

What number(s) would be more “medium” and half “small”?

Now let’s do the same for “small” and “large”

# MF EXERCISE



# MF QUANTIFICATION EXAMPLE

In order to use the “fuzzy control” approach to quantification, you must elicit the following things:

- A discrete set of conditions
- A continuous range of values for the condition
- Membership functions relating the discrete set of conditions to the continuous range of values
- A discrete set of performance levels
- A continuous range of values for performance
- Membership functions relating the discrete set of performance levels to the continuous range of values
- If-then rules relating the performance levels to the set of conditions

# CONDITIONS

MISS DISTANCE (condition) and

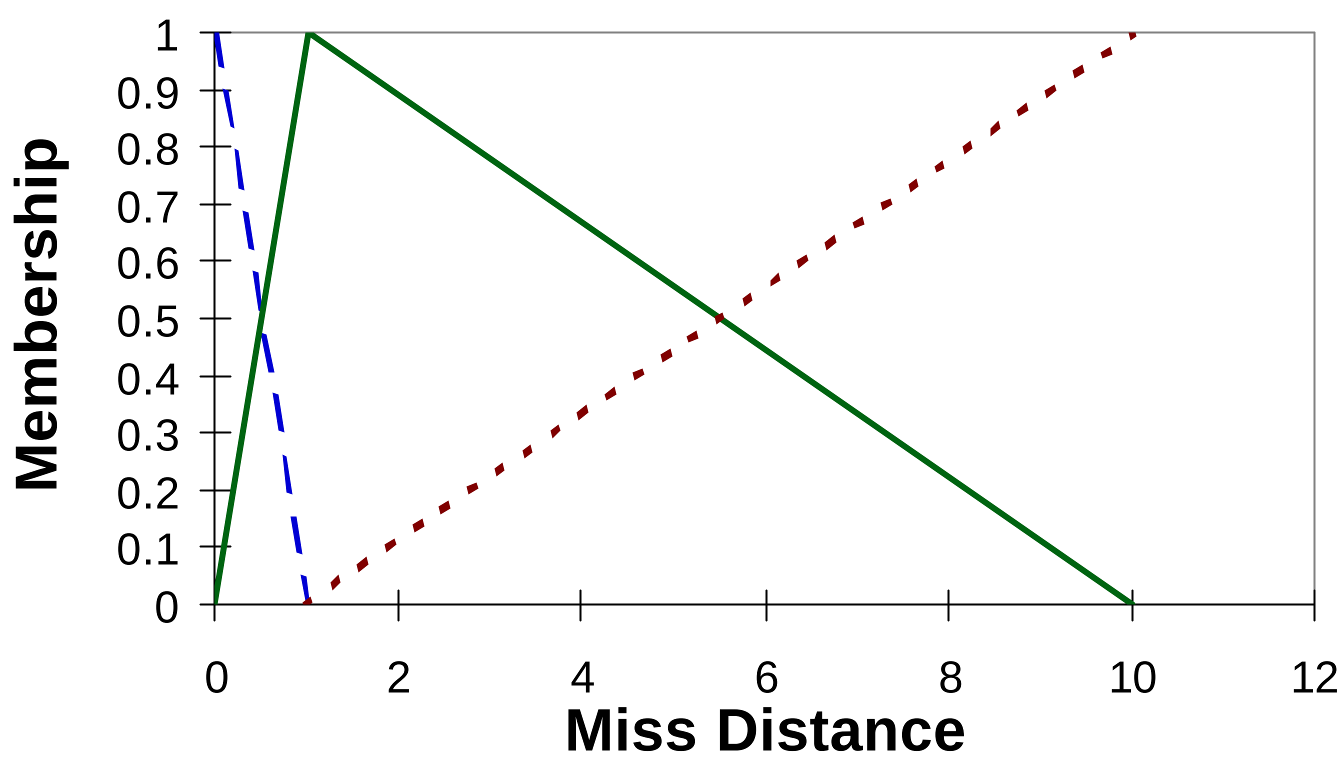
PERFORMANCE OF MISSILE (performance)

***Membership functions (input, condition)***

***A —> close***

***B —> nominal***

***C —> far***




**Assume: triangular**

# PERFORMANCE

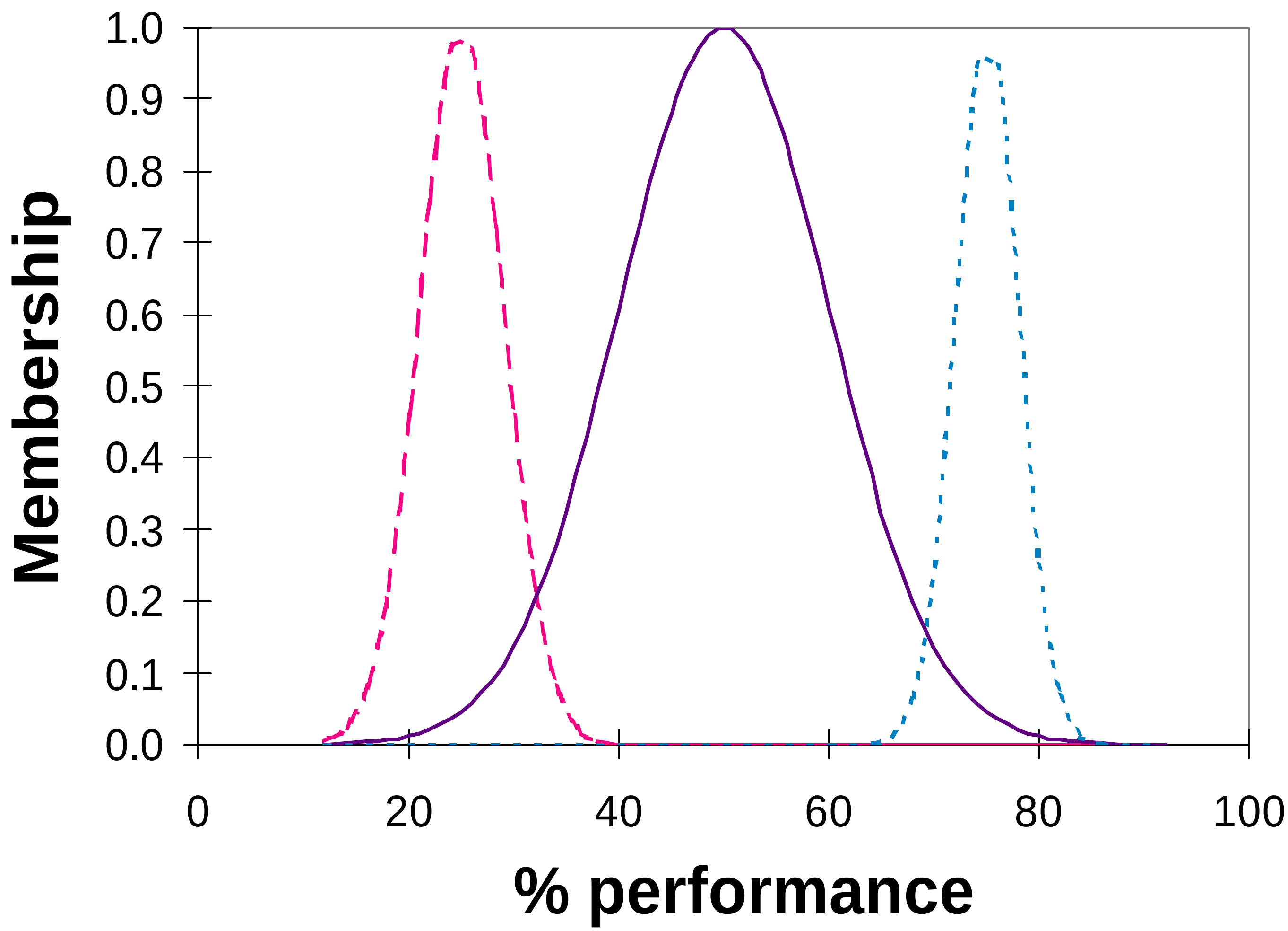
***Membership functions (for % performance)***

***a—> low***

***b —>marginal***

***c —> high***

Assume:
scaled (0 to 1) normals



# *If-Then* RULES MAP CONDITION INTO PERFORMANCE

***If*** **Condition (*x*)** ⇨ ***Then*** **Performance (*y*)**

| | |
|---|---|
| ***A=close*** | ***a=high*** |
| ***B=nominal*** | ***b=marginal*** |
| ***C=far*** | ***c=low*** |

***If*** miss distance condition (x) is *nominal* (in XXX), ***then*** % performance (y) is *marginal*.

# COMBINING PERFORMANCE MFs

## Fuzzy Combination (Mixture)

$$f_{FC}(y|x=\boldsymbol{x}) = \mathrm{m}\mathbf{A}(\boldsymbol{x})\cdot f_{\mathbf{a}}(\boldsymbol{y}_{\mathbf{a}}) + \mathrm{m}\mathbf{B}(\boldsymbol{x})\cdot f_{\mathbf{b}}(\boldsymbol{y}_{\mathbf{b}})$$

## Linear Combination (of Random Variables)

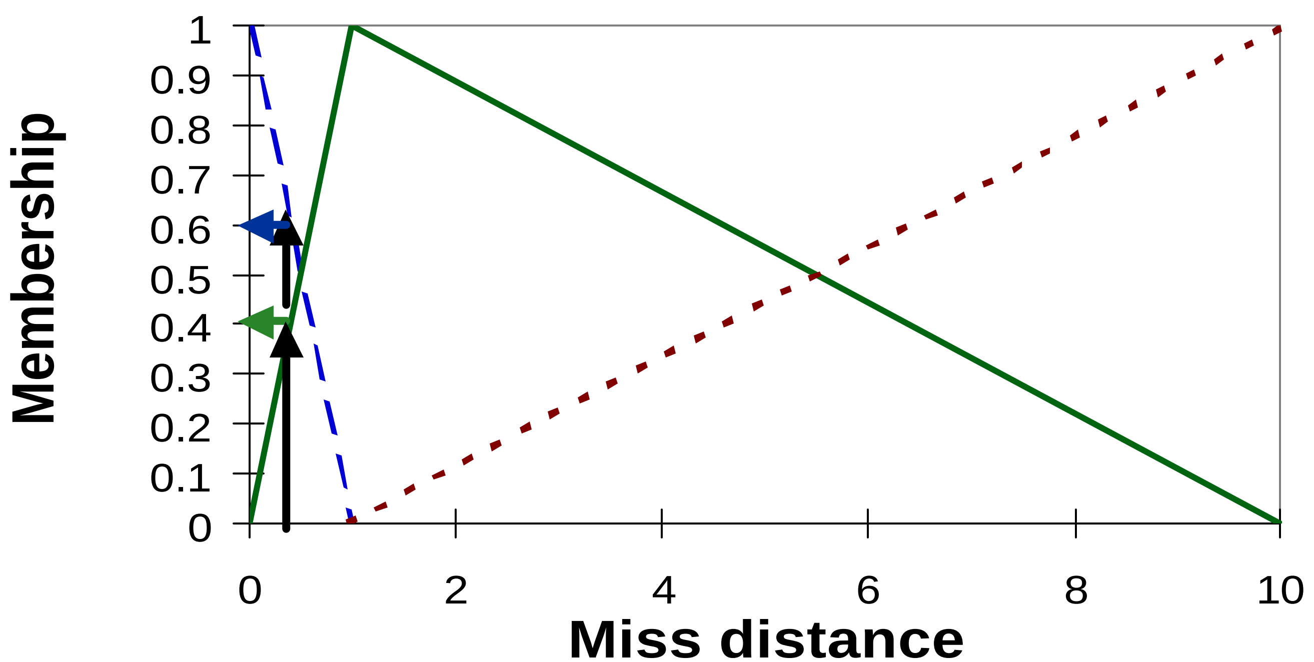
$$f_{LC}\,(y|x=\boldsymbol{x}) = f\,(\mathrm{m}\mathbf{A}(\boldsymbol{x})\cdot y_{\mathbf{a}} + \mathrm{m}\mathbf{B}(\boldsymbol{x})\cdot y_{\mathbf{b}})$$

$f_{\mathbf{a}}$, $f_{\mathbf{b}}$, $f_{\mathbf{LC}}$ **are normals**

# MAPPING CONDITION INTO PERFORMANCE FOR ONE VALUE

If miss distance is *x*=0.8 meters , then *x* has membership of 0.6 in *close* and 0.4 in *nominal*.

Using the *If-Then* rules, map into *high* performance with a weight of 0.6 and into performance set *marginal* with a weight of 0.4 .

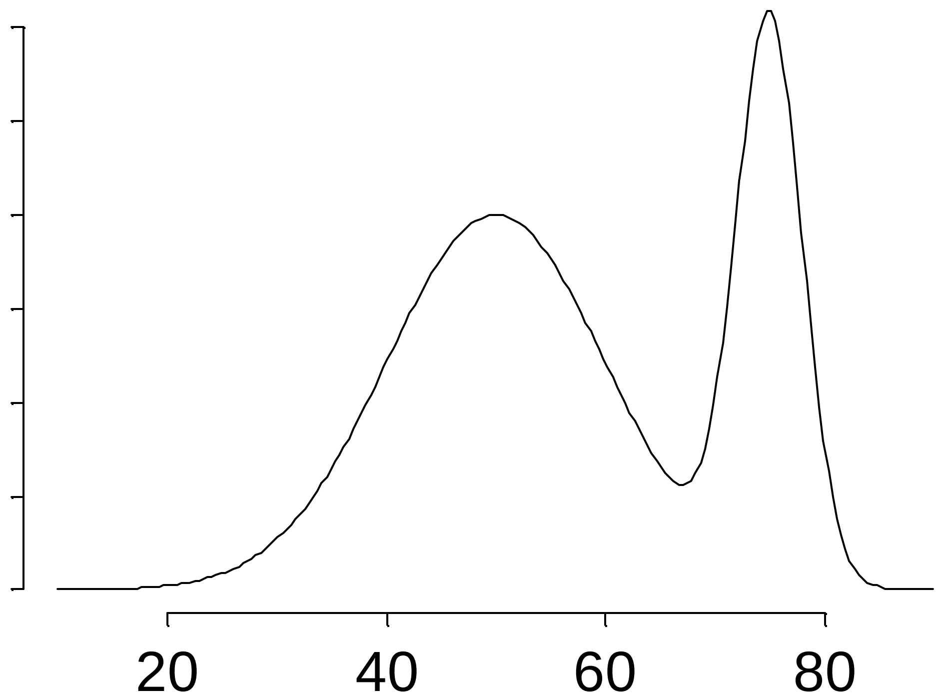


# DETERMINING THE UNCERTAINTY DISTRIBUTION

For component condition $x$=0.8: The *fuzzy combination* combines performance sets *high* and *marginal* with weights of 0.6 and 0.4. This combined function can serve as an uncertainty distribution of performance given $x$, $f(y|x)$.

In fuzzy control systems, this is reduced to a *centroid* value (*defuzzification*).

**Problem: can be bimodal.**



# COMPARING UNCERTAINTY DISTRIBUTIONS

Performance MFs

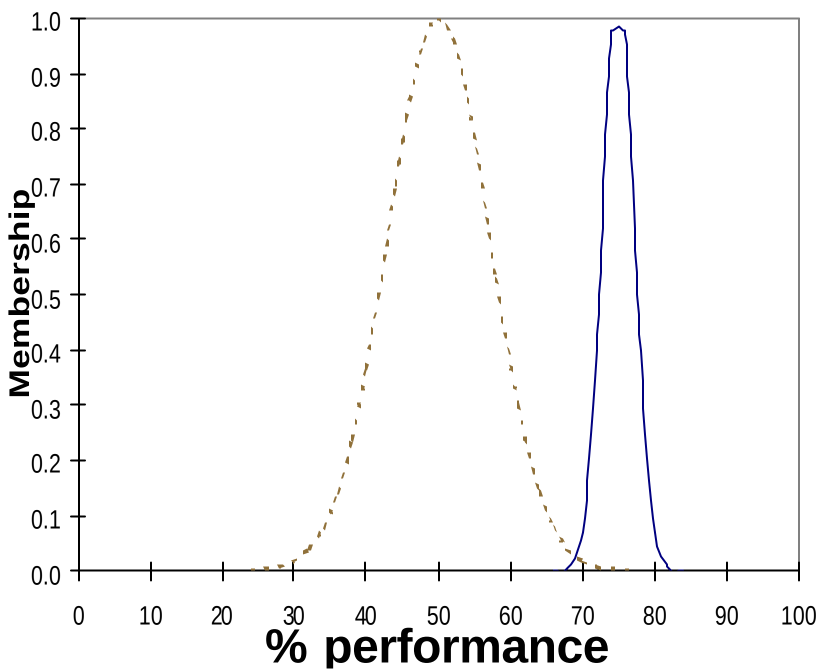
*a—> high*

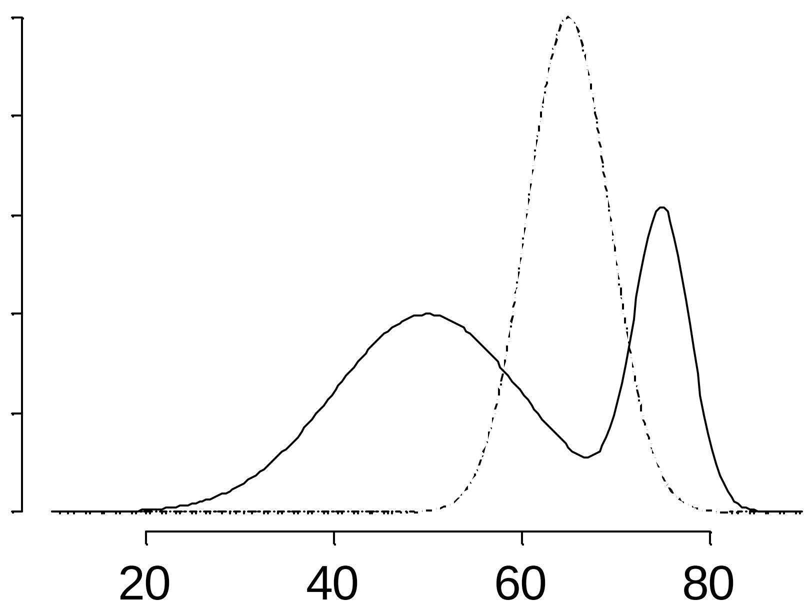
*b—> marginal*

**Combinations**

**fuzzy**

**linear**



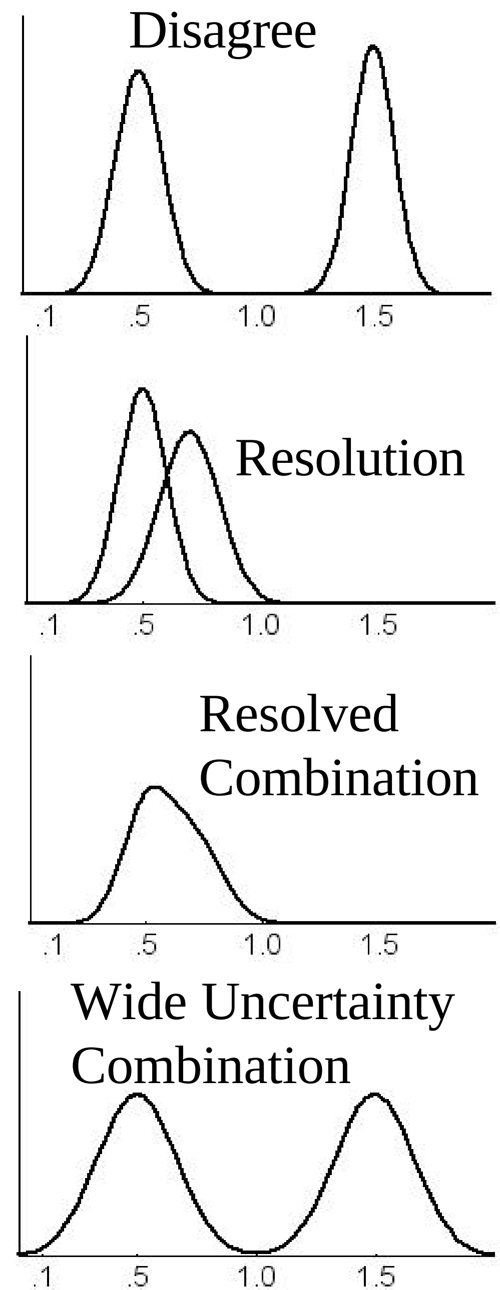
# WHICH QUANTIFICATION DO YOU USE?

- What kind of expert judgment did you elicit?
- What method is the most tractable?
- What matches the features that your expert considers important?
- Which one can you simulate from?

# WHEN EXPERTS DISAGREE

- Try to resolve differences
  - *Different questions being answered*
  - *Different assumptions or conditions*
  - *Different level of detail (granularity or resolution)*
  - *Different available information for problem solving*

- If resolved differences are not the cause of the separation, then wide uncertainties from the (combination of) experts represent the existing, wide, state of uncertainty.



# AGGREGATING MULTIPLE EXPERTS

- Reaching consensus (using elicitation techniques)
- Mathematical or weighting methods by decision maker *(default is equal weights, unless have other reasons)*
- Mathematical or weighting methods by analyst with feedback to experts (also through the Advisor) *(dependence among experts is a problem)*

*Reference list is in the notebook.*

# AGGREGATING MULTIPLE EXPERTS

Combining expert distributions is not unlike combining the diverse sources of information using information integration tools.

More on combining in next segment