Bayesian Analysis

Introduction

Karthik Sriram

karthiks@iima.ac.in

Indian Institute of Management Ahmedabad
Production & Quantitative Methods Area

Motivating Example 1

A Bank is providing a loan to an up coming business. For risk assessment and pricing, it wants to estimate the

recovery rate = percentage of loan that can be recovered in case the company **defaults**.

How could the bank go about doing this ?

The problem…

- To estimate

  Recovery Given Default

  =1- Loss Given Default (LGD)

An Excerpt from Basel II

" Internal estimates of PD, **LGD**, and EAD must incorporate all relevant, material and available data, information and methods."

"**A bank may utilise internal data** and data from external sources (including pooled data)."

"Where internal or external data is used, the bank must demonstrate that its estimates are representative of long run experience"

Consider this situation….

- Bank has past data on 5 such deals

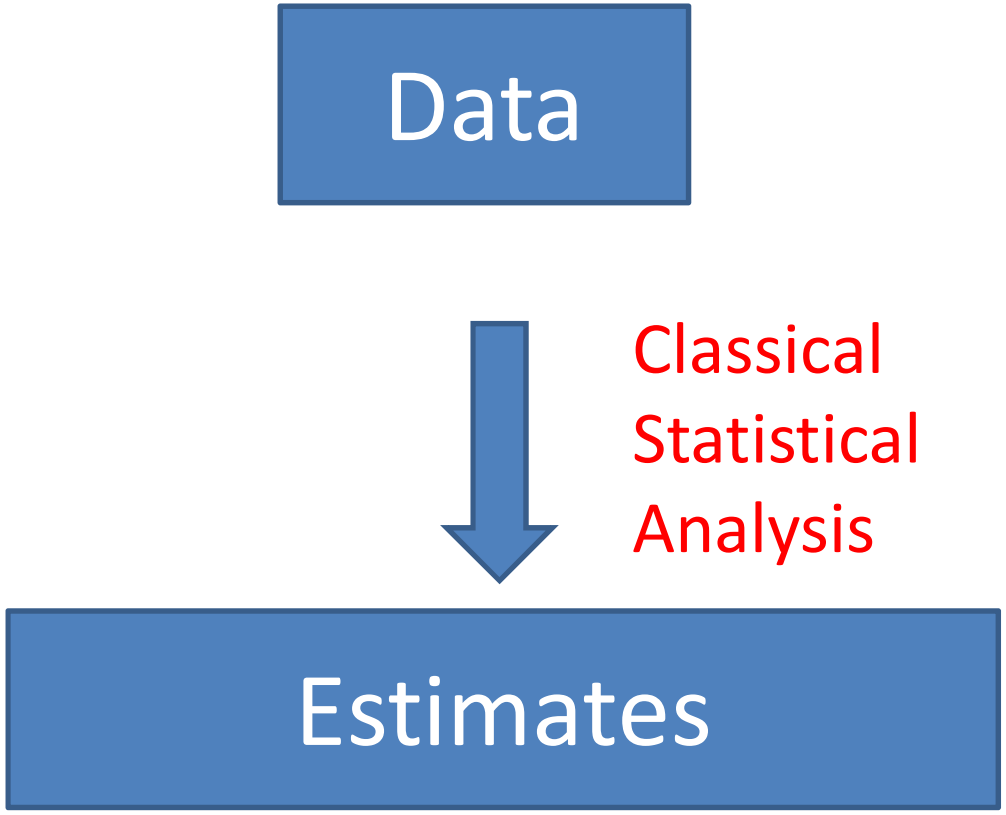
  Average Recovery rate is over 95%.

- Bank believes that their risk processes are superior to that of industry.

Excerpt from Basel II

"The bank's own workout and collection expertise significantly influences their recovery rates and must be reflected in their " [recovery] "estimates",

The Classical Statistical Approach



Excerpt from Basel II

"The bank's own workout and collection expertise significantly influences their recovery rates and must be reflected in their " [recovery] "estimates",

" but adjustments to estimates for such expertise must be conservative until the bank has sufficient internal empirical evidence of the impact of its expertise".

Excerpt from Basel II

"A bank must have procedures for human review of model-based rating assignments. Such procedures should focus on finding and limiting errors associated with known model weaknesses……"

An Expert Opinion

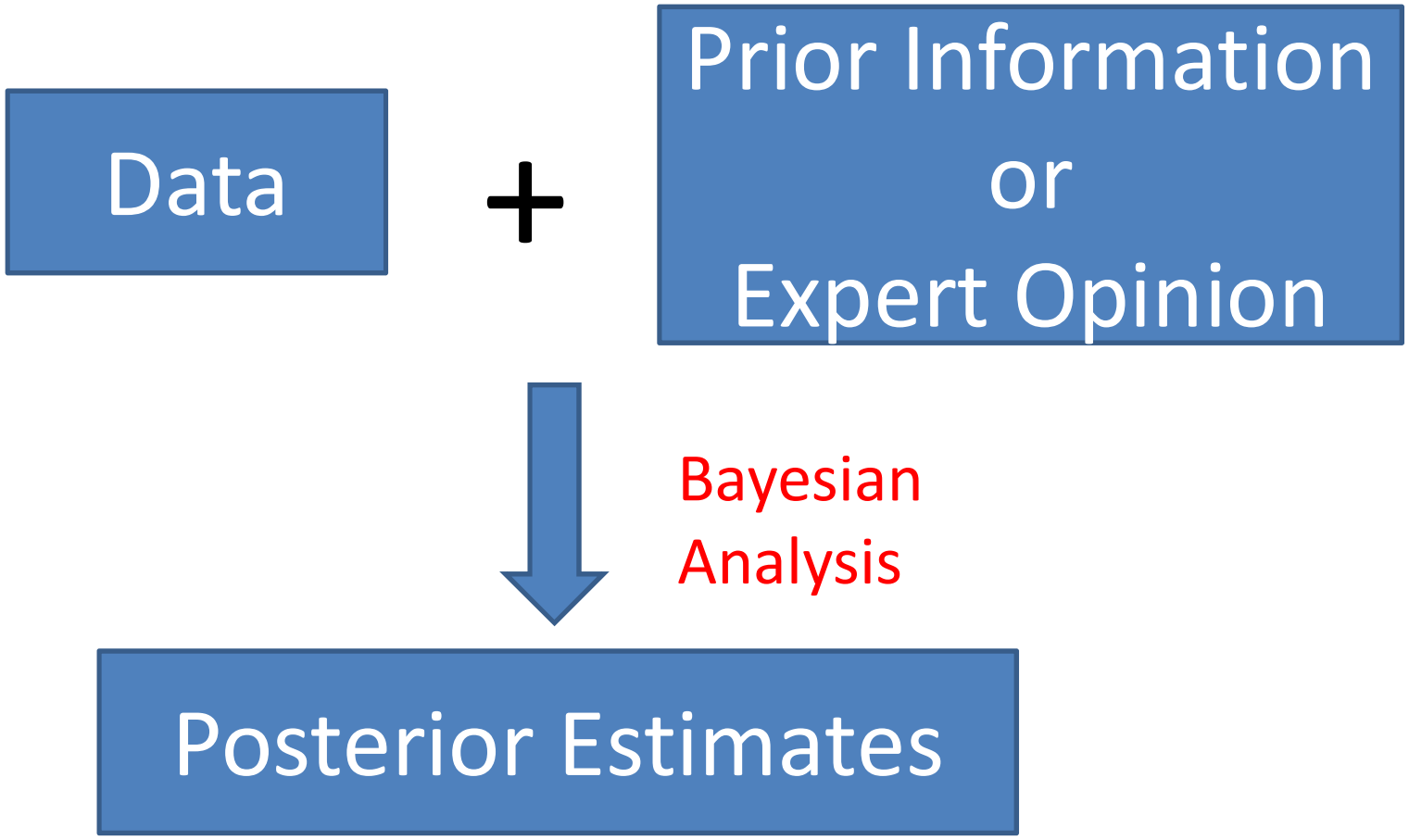
- A risk manager has dealt with lending to similar new businesses in his past jobs.

- Based on his prior experience he believes that the such deals are actually risky and recoveries from such deals can be easily as low as 40% (plus or minus 5%).

Excerpt from Basel II

" …When Combining model results with human judgement, the judgement must take into account relevant and material information not considered by the model. The bank must have written guidance describing how human judgement and model results are to be combined. "

The Bayesian Framework



Example 2- Search for AF 447

- June 1 2009 , Air France flight 447 went missing over the Atlantic
- After 6 days, found some debris and bodies –38NM north of last known position
- But subsequent searches remained unsuccessful

(Reference: Stone et al (2014), Statistical Science)

Last known position on map

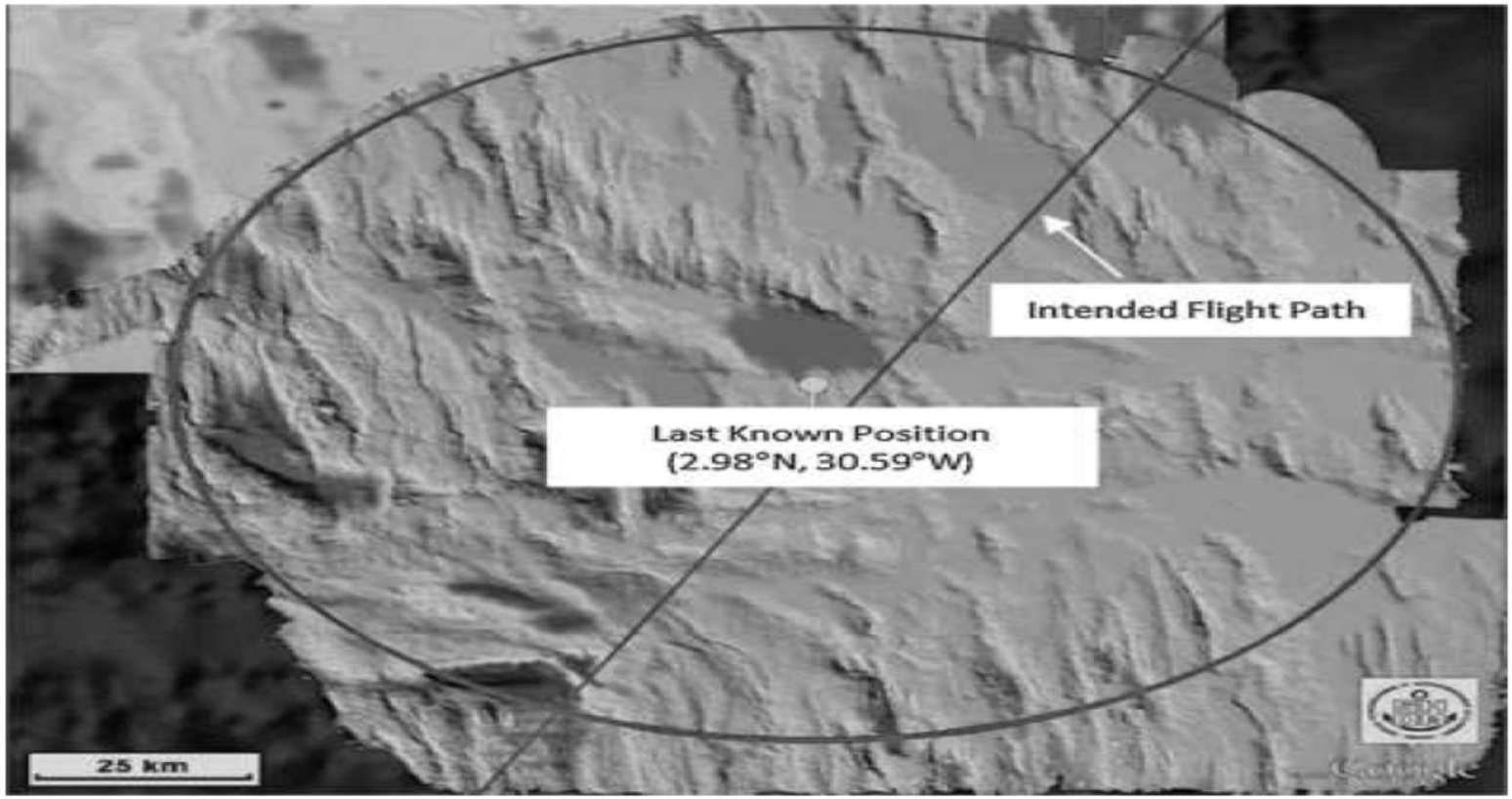
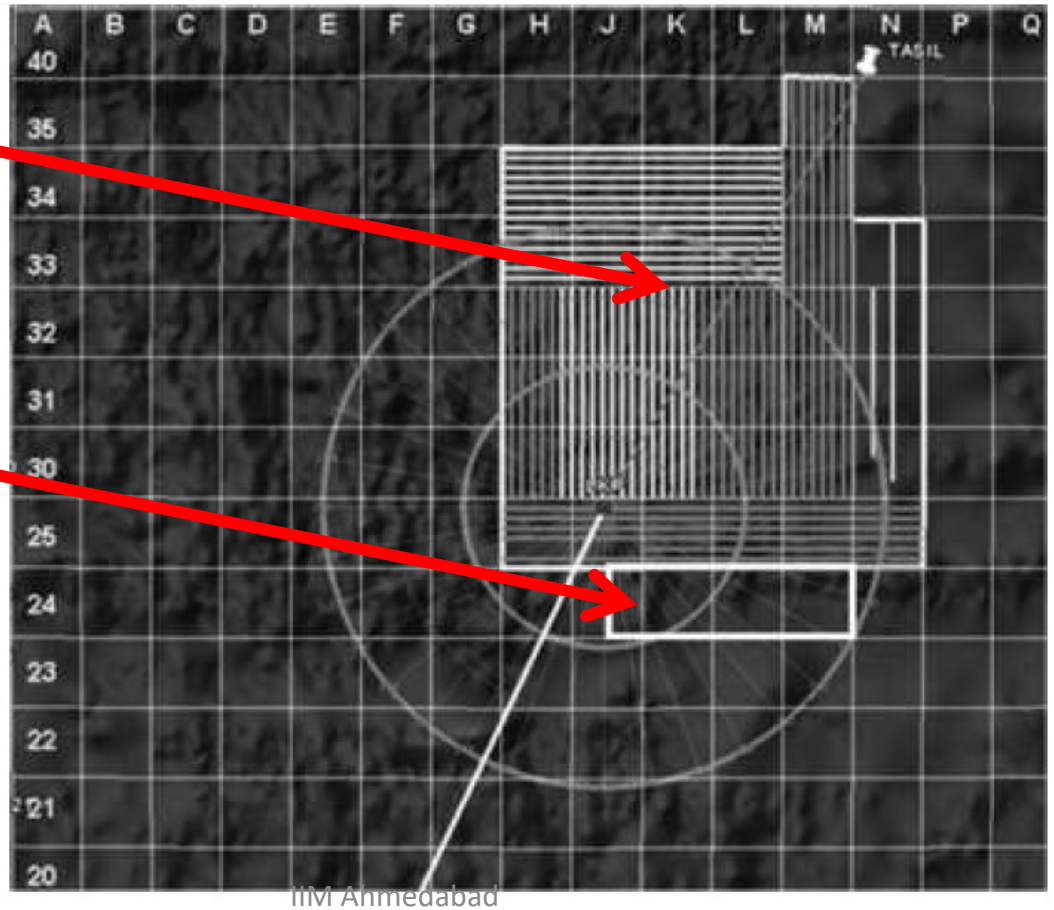


FIG. 1. *Last known position of the aircraft, intended flight path and the 40 NM circle.*

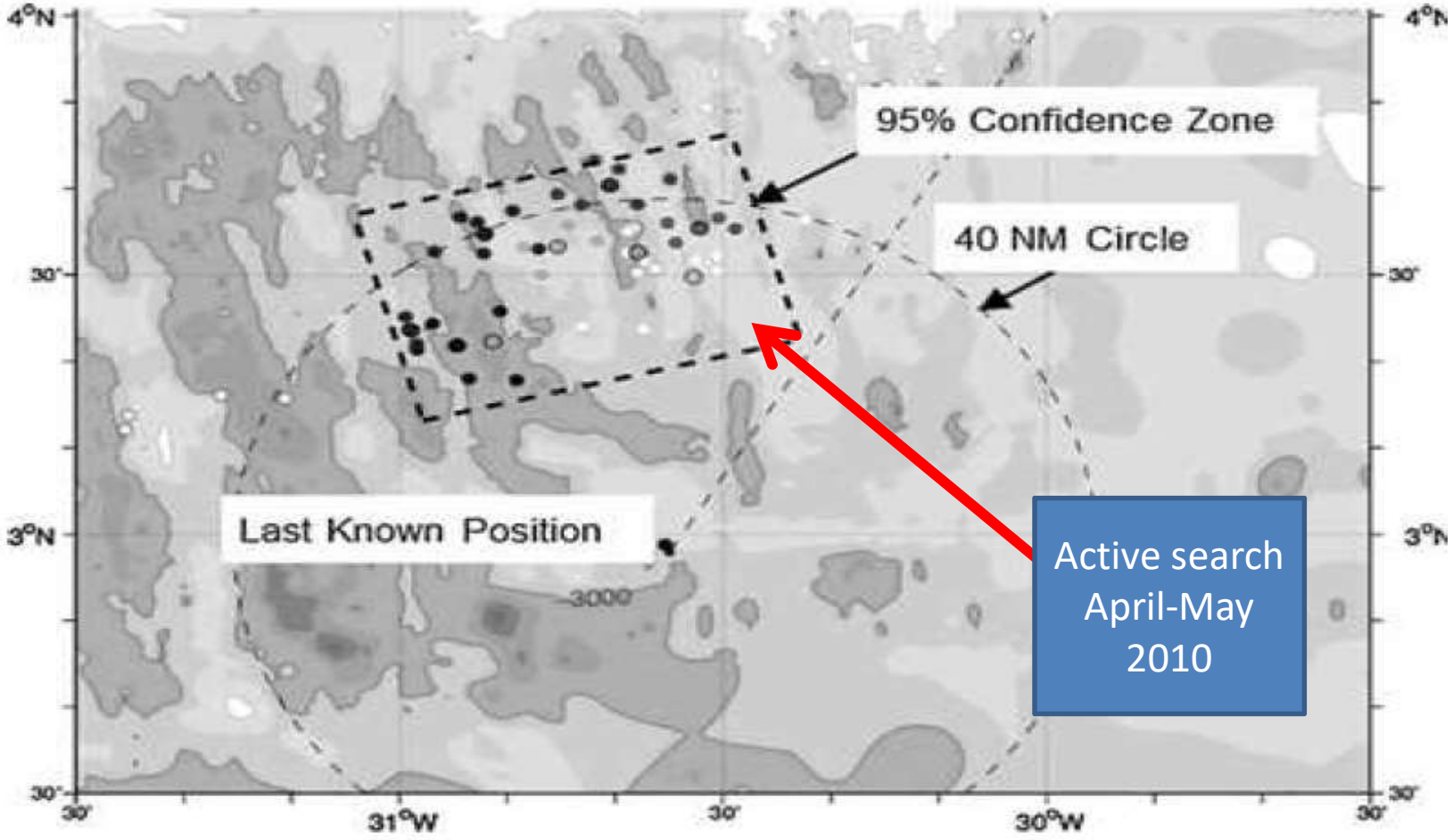
Example 2- Search for AF 447 (ctd..)

Passive acoustic search- June 10 2009 + 31 days

Active acoustic search- August 2009



95% confidence zone based on back tracking possible trajectories of debris



Example 2- Search for AF 447 (ctd..)

- In July 2010, Stone, Keller, Kratzke and Strumper were assigned the task of developing a probability distribution for the location of the wreckage.

Example 2- Search for AF 447 (ctd..)

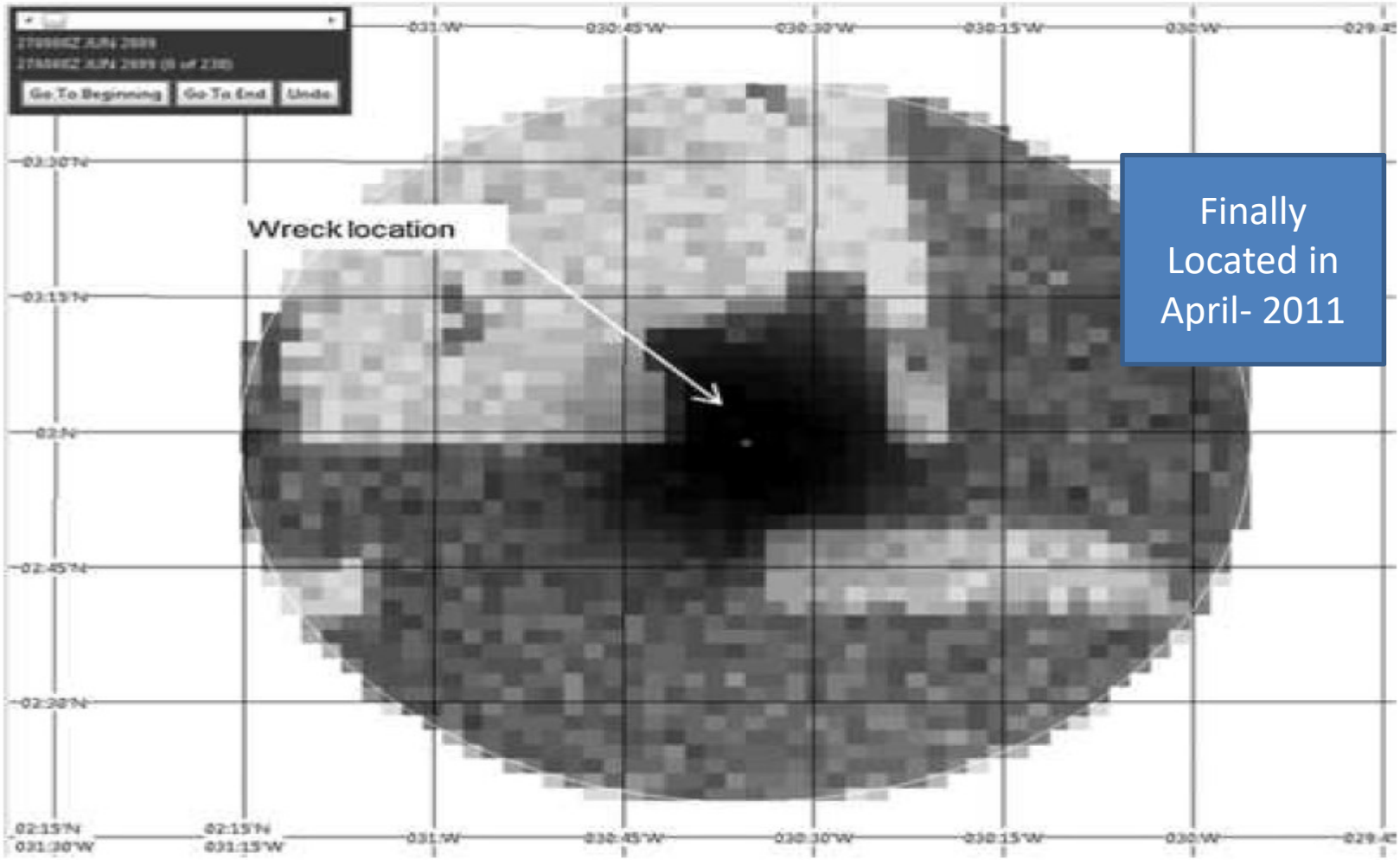
Bayesian Approach

Prior – Uniform [<=40NM]
+ Past incidents
+ Reverse Drift

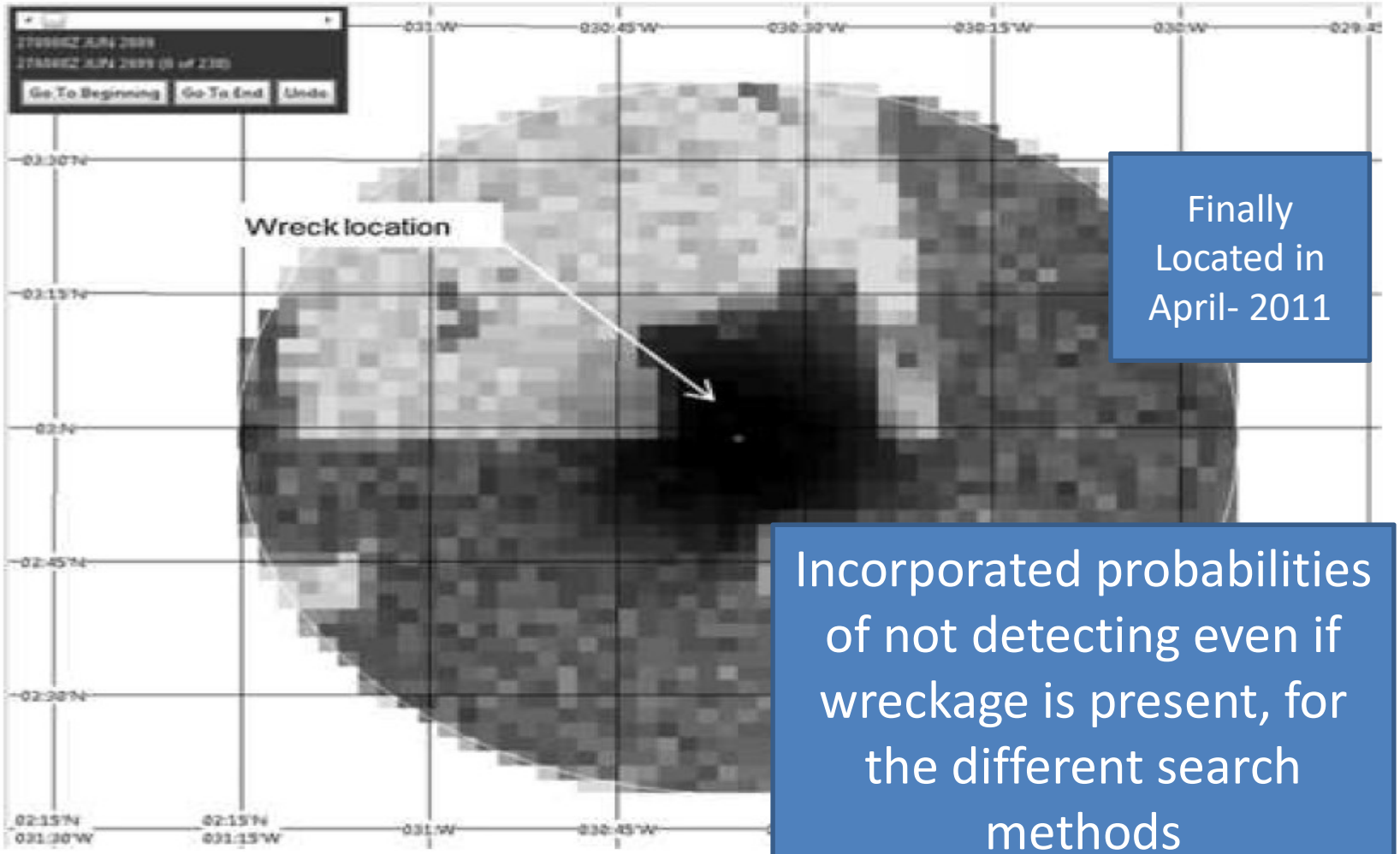
Data-
Surface search + Passive search + Active 2009 + Active 2010

Posterior - Search and then update the distribution

The posterior distribution of wreckage location



The posterior distribution of wreckage location



Key take aways…..

- Data gives objectivity

…… but may not always exactly represent the situation we are trying to handle.

- Expert opinion or Prior Information is valuable

……. but is subjective.

- Bayesian Approach: A systematic way to balance.

In this course….

What we will do

Bayesian formulations, solutions and Ideas, Implementation

Theory (algebra at class 11 /12 level)

Practical (R, Excel)

Evaluation:

- 2 Exams (take-home)
- CP = "correctly answering my questions "

Some guidelines

- No SMS, Surfing, Messaging, whatsapp, FB …
  No unnecessary distraction in class
- **Please take down notes** ….
- Participate in class exercises
- Feel free to seek clarifications inside or outside class
- Work out any suggested exercises.

Essential Concepts Involved

- Conditional Probability and Bayes Rule.
- Likelihood based on Statistical Model for Data
- Posterior:= “Prior” X “Data Likelihood”
- Posteriors Summaries and Inference

CONDITIONAL PROBABILITY

Simple Example

- I throw a six faced fair die.
- I do not reveal the outcome to you.
- How do you summarize the information you have so far ?

Simple Example ctd

- Sample space: {1,2,3,4,5,6}
- Each outcome has same probability = 1/6

  e.g. P({3})= 1/6

Simple Example ctd

Additional Information:

If we knew the number is divisible by 3

Question: Now, What is P({3}) ?

Simple Example ctd

All possibilities : {1,2,3,4,5,6}

Random Event : A= {divisible by 3}={3,6}

Probability that 3 occurs **given** that A has occurred?

$$P(\{3\}|A) = \frac{1}{2}$$

Closer look …

$$P(\{3\}|A) = \frac{1}{2} = \frac{\left(\frac{1}{6}\right)}{\left(\frac{2}{6}\right)} = \frac{P(\{3\})}{P(\{3,6\})} = \frac{P(\{3\} \cap A)}{P(A)}$$

Another Calculation …

$$P(\{1\}|A) = 0 = \frac{P(\{1\} \cap A)}{P(A)}$$

Definition: Conditional Probability

Probability of event B given A

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

BAYES RULE

HIV Test Example

- A test is 95 % accurate in detecting the existence or non-existence of HIV.
- A person tests positive.
- What is chance that this person has HIV?

Conditional Probability Formulation

- A = { Tests Positive}
- B= {Has HIV}
- **Given** (i) P(A | B) = 95%
  (ii) P(Not A | Not B)= 95%
- **Question**: What is P(B|A) ?

Bayes Rule ---Step 1

- $P(B|A) = \frac{P(A \cap B)}{P(A)}$ and $P(A|B) = \frac{P(A \cap B)}{P(B)}$

So…… $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B) * P(B)}{P(A)}$

Bayes Rule ---Step 2

- $P(A) = P(A \cap B) + P(A \cap Not\ B\ )$

- $P(A)$

$= P(A|B) * P(B) + P(A|Not\ B\ ) * P(Not\ B)$

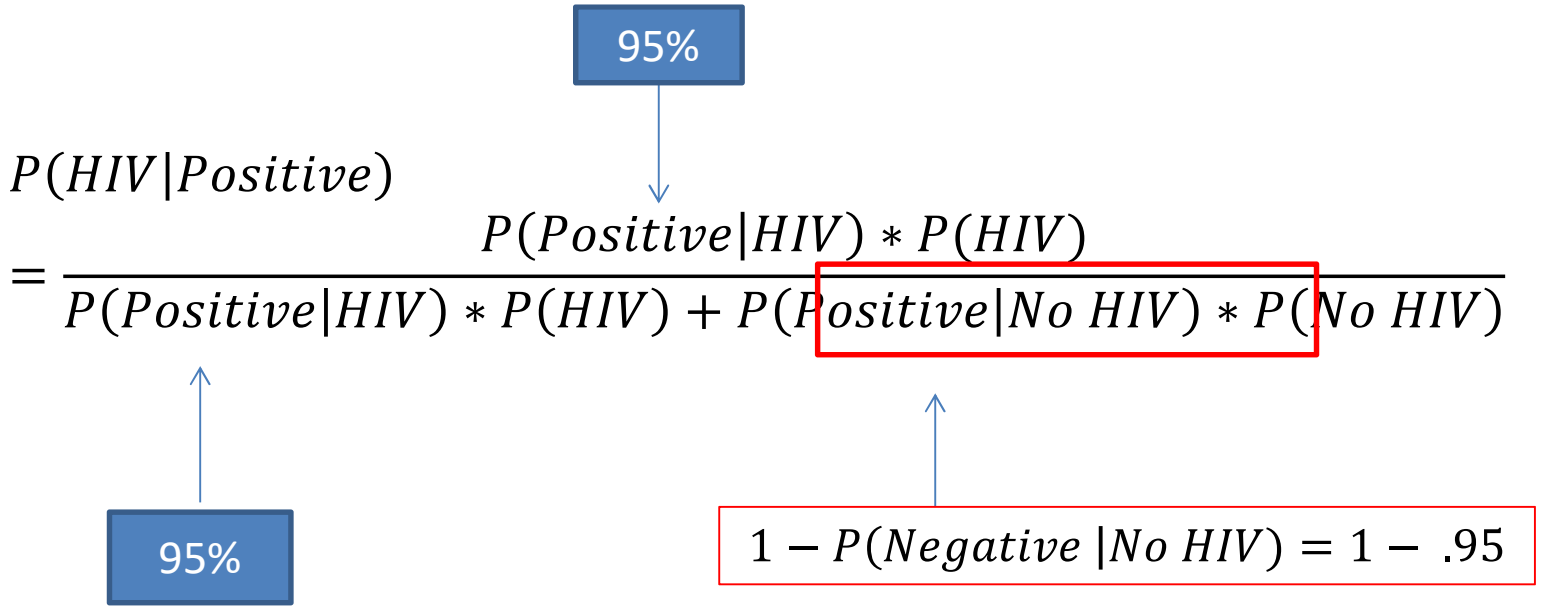
Bayes Rule

Putting the two together….

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A|B) * P(B) + P(A|Not\ B) * P(Not\ B)}$$

Denominator is called "marginal probability" of A

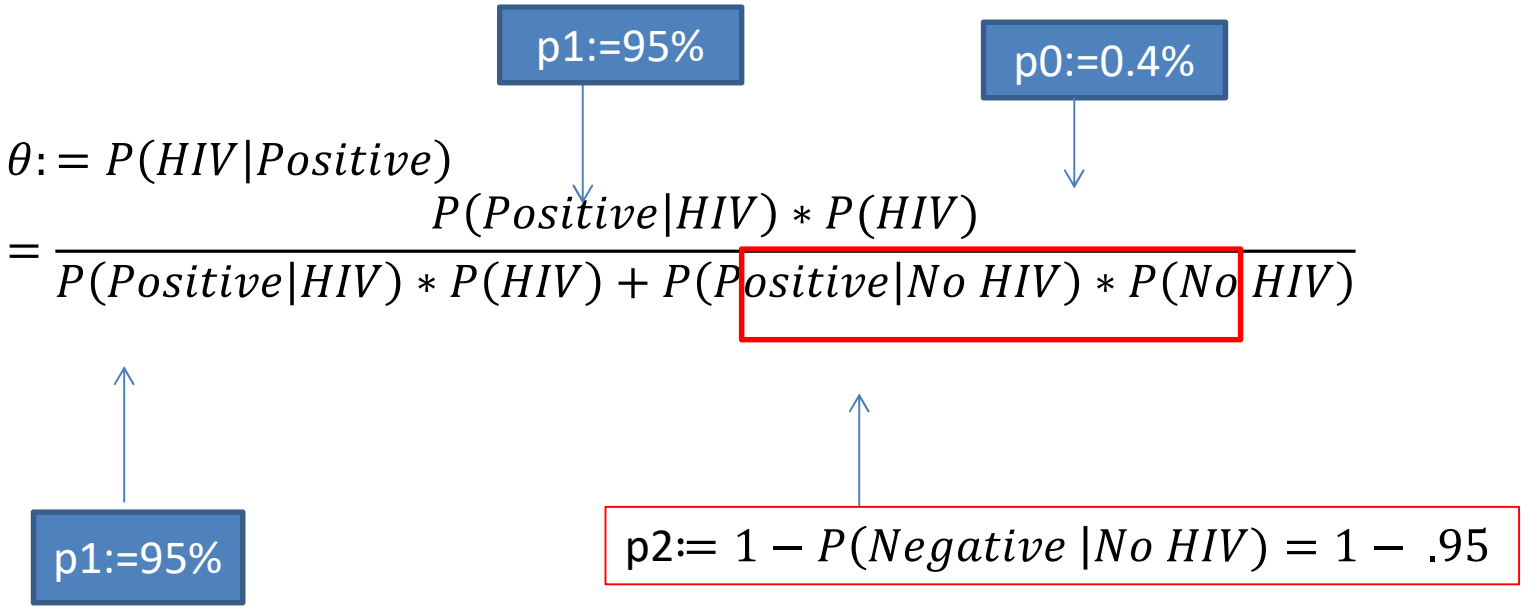
HIV Example



Suppose that $P(HIV) = .4\ \%$, then …..

P(HIV |Positive)= ?

HIV Example



Suppose that $P(HIV) = .4\ \%$, then .....

P(HIV |Positive)= $\theta = \frac{p_1 p_0}{p_1 p_0 + p_2(1-p_0)}$ =7.1%

Bayesian Idea

- P(HIV) = .4%

  --- "Prior Belief" about incidence of HIV

- A= {Test Positive}

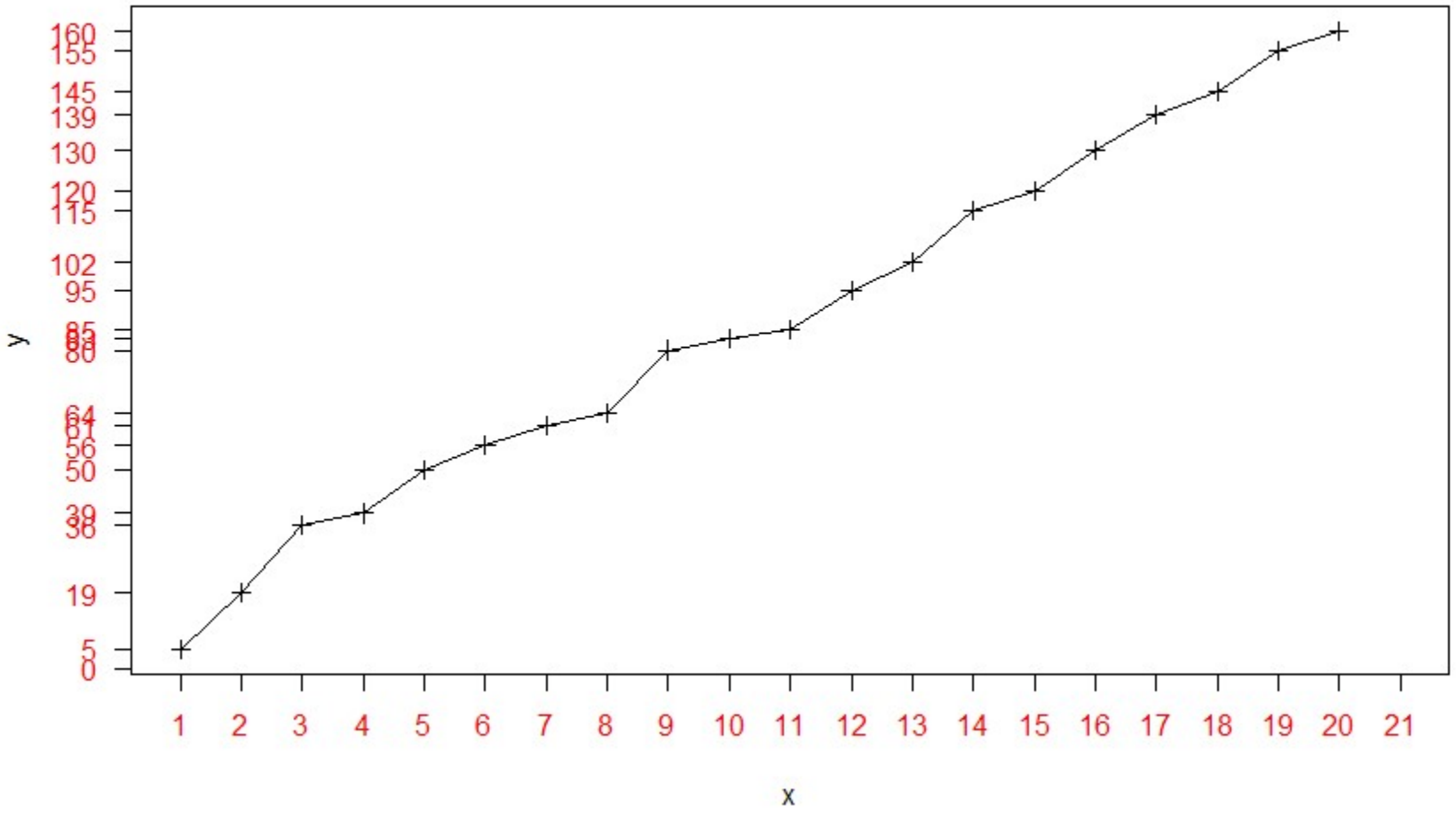
  --- Additional Information "Data"

- P(HIV|Positive):

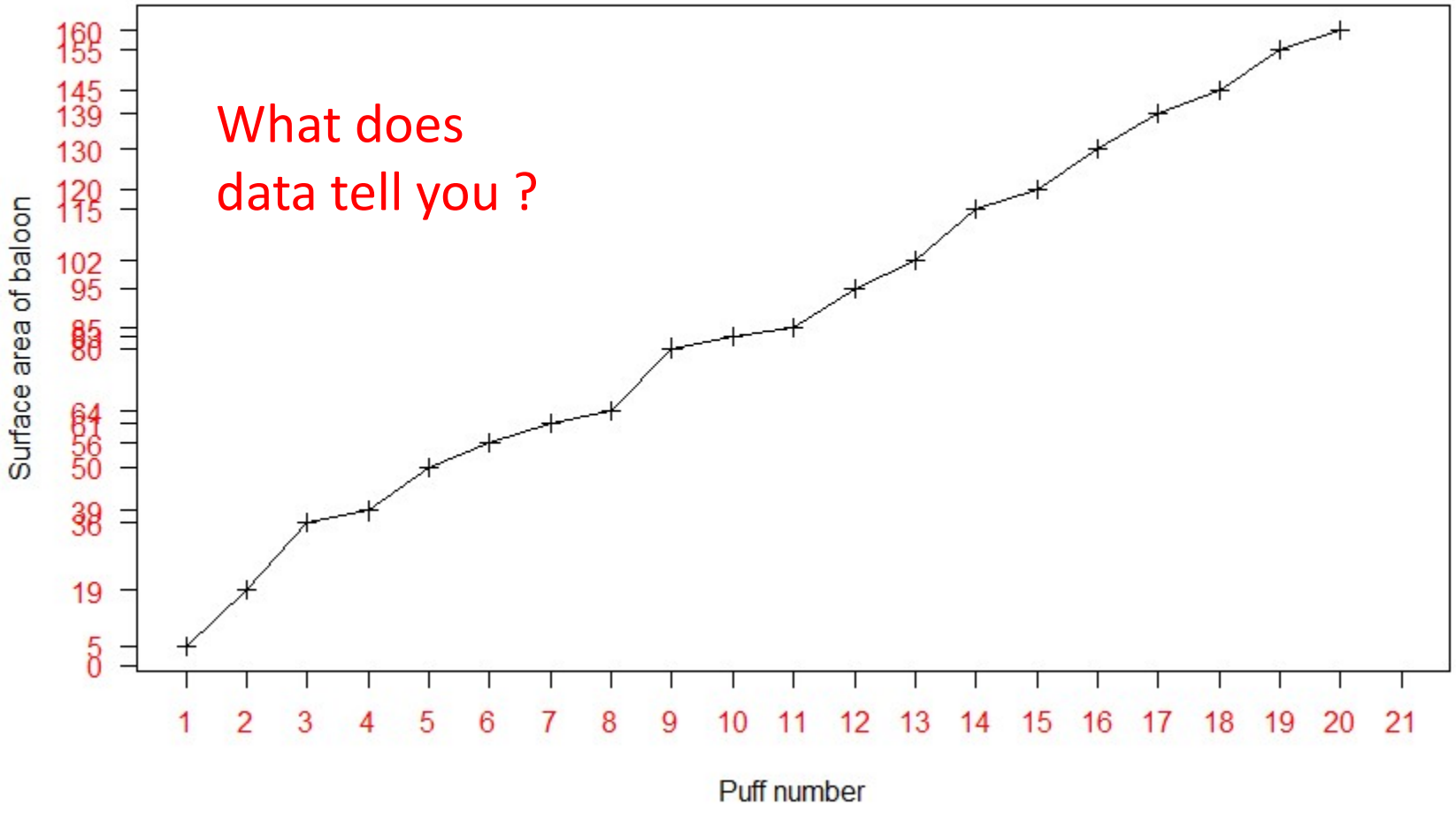
  --- "Updated Belief or Posterior given Data"

IMPORTANCE OF COMBINING PRIOR OR EXPERT KNOWLEDGE WITH DATA

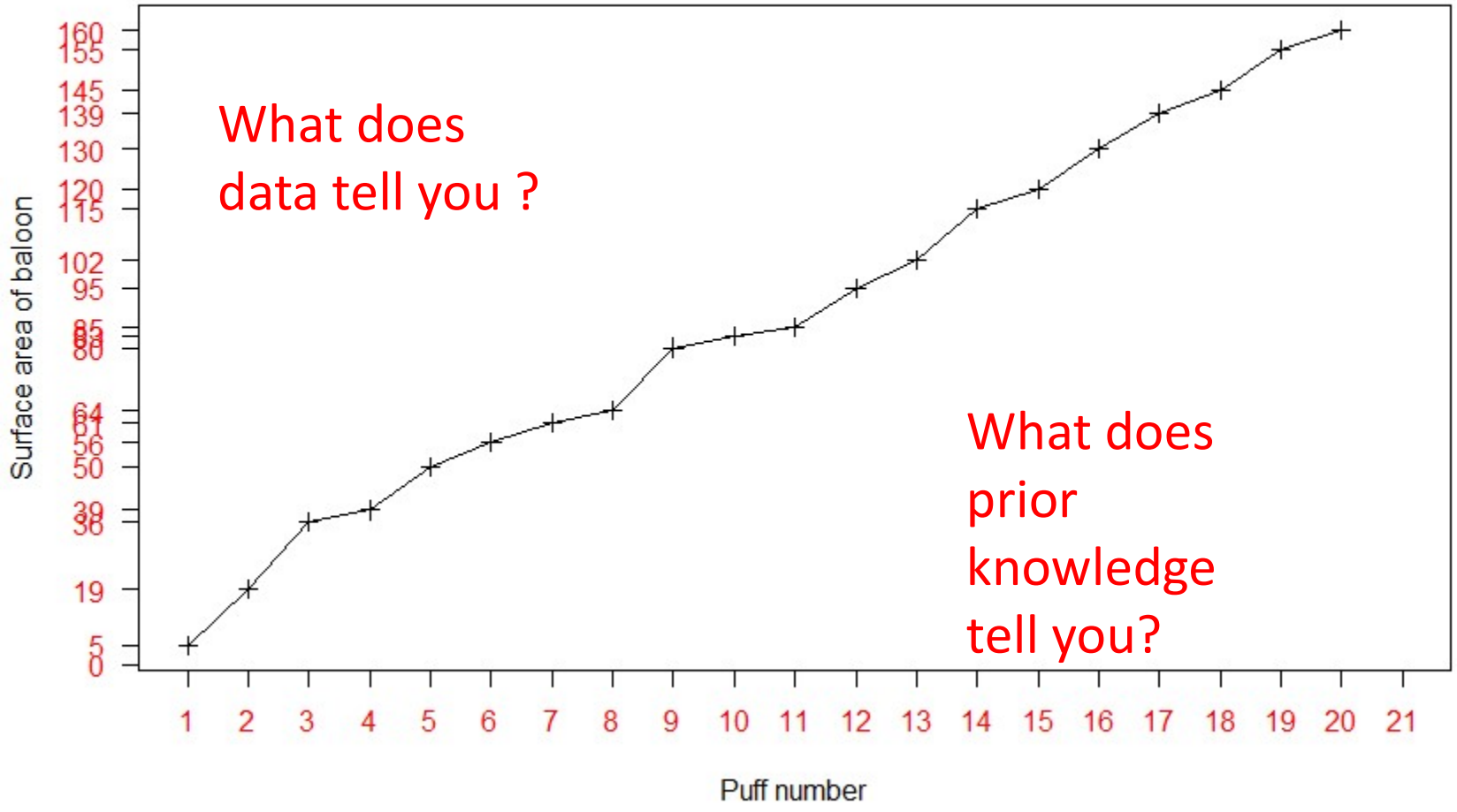
Regression: Can you predict y for x=21



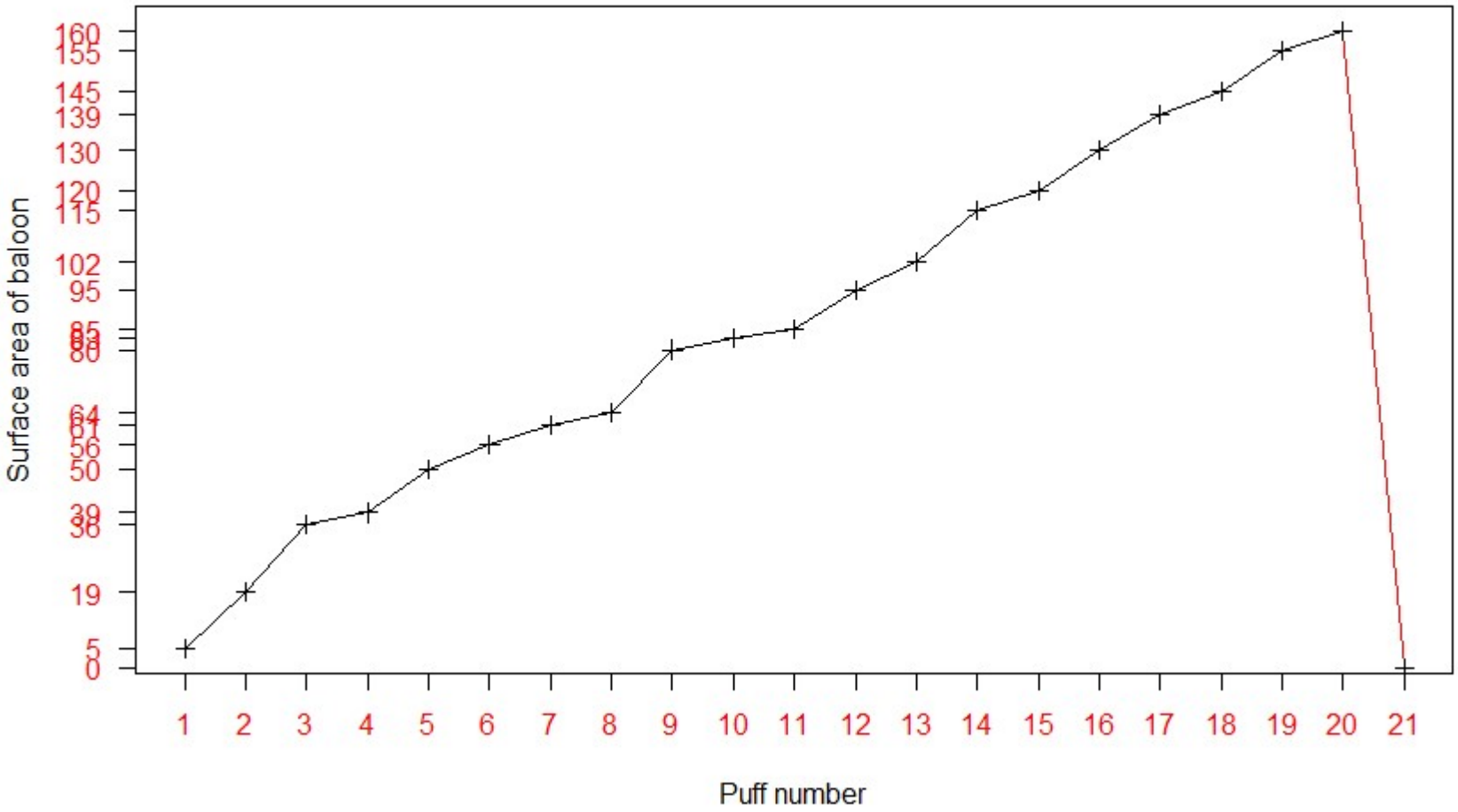
Context ! : Puffing of balloon.. ..Now?



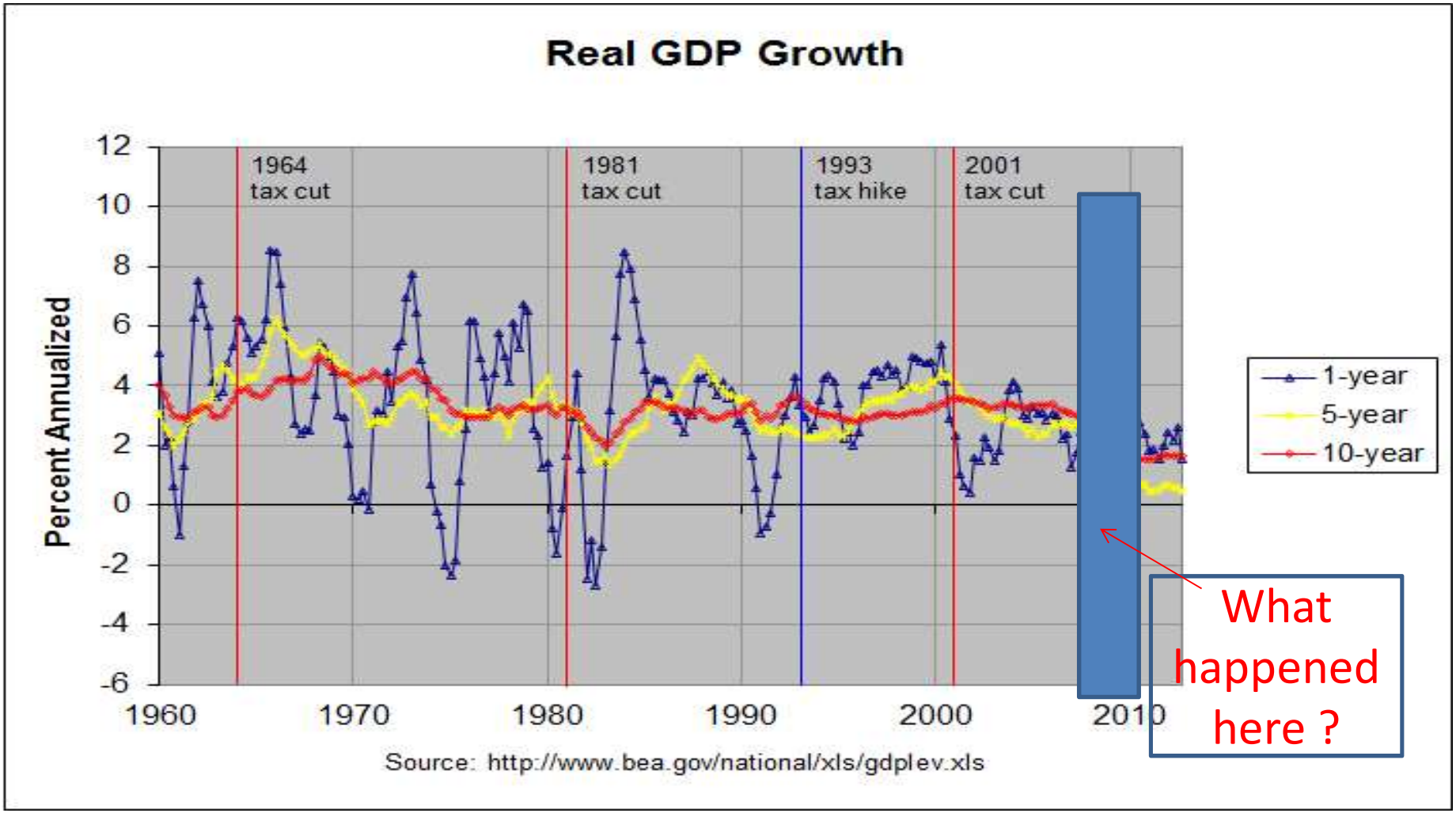
Context ! : Puffing of balloon.. ..Now?



What may happen!



U.S GDP annualized quarterly growth rates



Expert opinion

- **Nouriel Roubini, chairman of Roubini Global Economics**

As early as August 2006, Roubini **had commented** that slumping house prices were "enough to trigger a US recession" and, also that month, that "sub-prime lending institutions may thus be the proverbial canary in the mine" signalling an ugly housing bust that "will be associated with a broader economic recession".

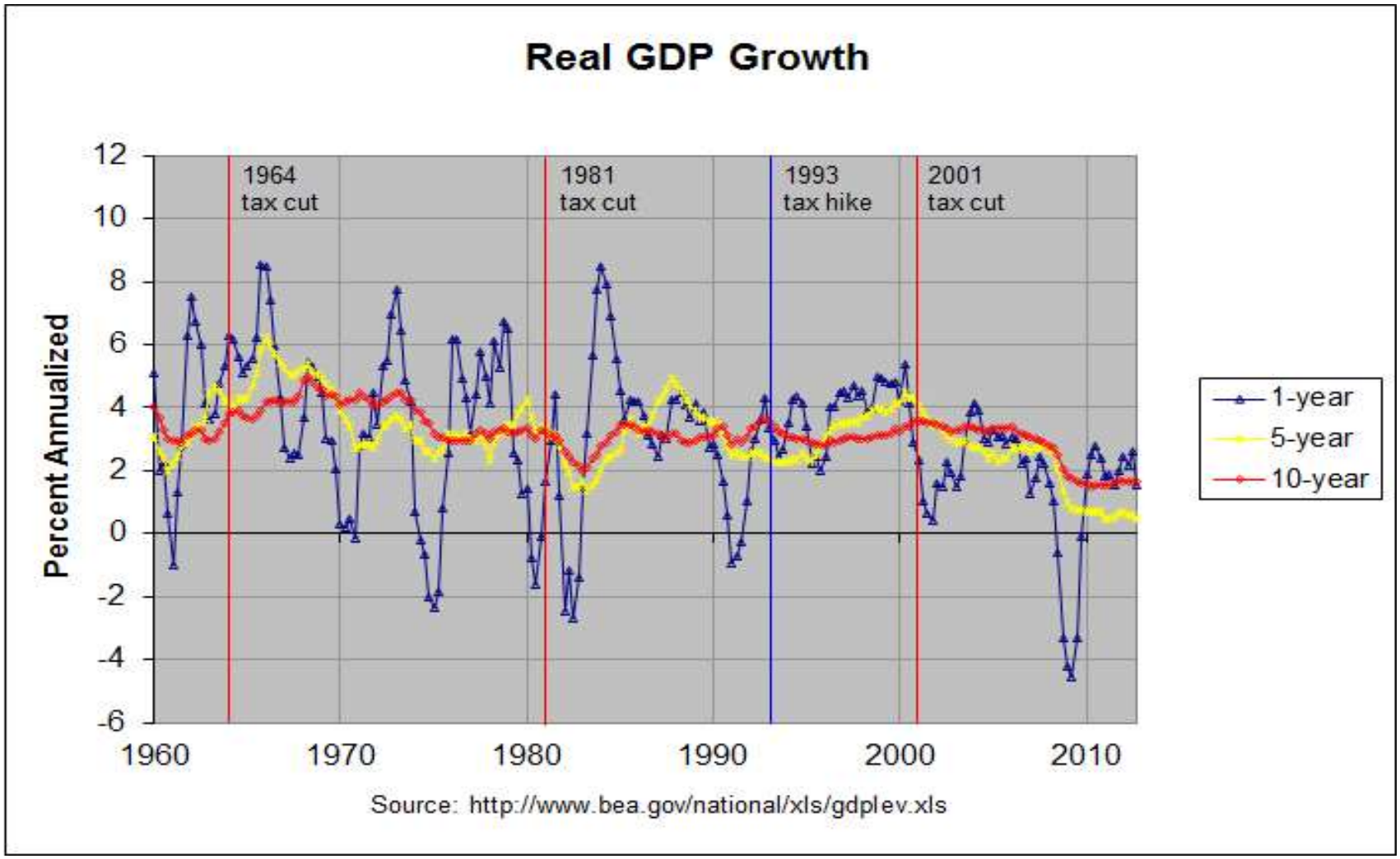
https://intheblack.com/articles/2015/07/07/6-economists-who-predicted-the-global-financial-crisis-and-why-we-should-listen-to-them-from-now-on

Expert opinion

- **Dean Baker, co-director of the Centre for Economic and Policy Research**
- In November 2006, Baker published his paper ***Recession Looms for the US Economy in 2007***, in which he predicted a "downturn in consumption spending, which together with plunging housing investment, will likely push the economy into recession."

https://intheblack.com/articles/2015/07/07/6-economists-who-predicted-the-global-financial-crisis-and-why-we-should-listen-to-them-from-now-on

U.S GDP growth rates



Bioassay- Experiment

	Control Group	Treament Group	
Developed Tumor	0	3	
Did not develop Tumor	50	47	
	50	50	100

- Testing Effect of a certain radiation on rodents
- Classical Statistical Test would conclude:
  radiation had no significant effect.
- But…the doctor believes 3 infected is Biologically Significant!

Bridging statistical and contextual significance

Crucial:

Doctor's prior belief based on his experience

……"Occurrence of tumor is rare"

In summary

Bayesian ideas matter where

- Expert opinions or Prior Information matters
- Data may be insufficient or not fully relevant
- We need to actively consider context/domain in the formulation of solution
- (Also….) where Computations are more difficult with classical formulations