PROBABILITY BOUNDS ANALYSIS FOR S-CURVES

9 June 2010

Presented by:

Christopher Mehl, Ph.D.
Omnitec Solutions, Inc.
E-mail: christopher.mehl.ctr@navy.mil
Phone: (301) 342-6110

Michael Popp
NAVAIR 4.2 Cost Department
E-mail: michael.popp@navy.mil
Phone: (301) 995-7945

WHAT'S IN YOUR P-BOX?

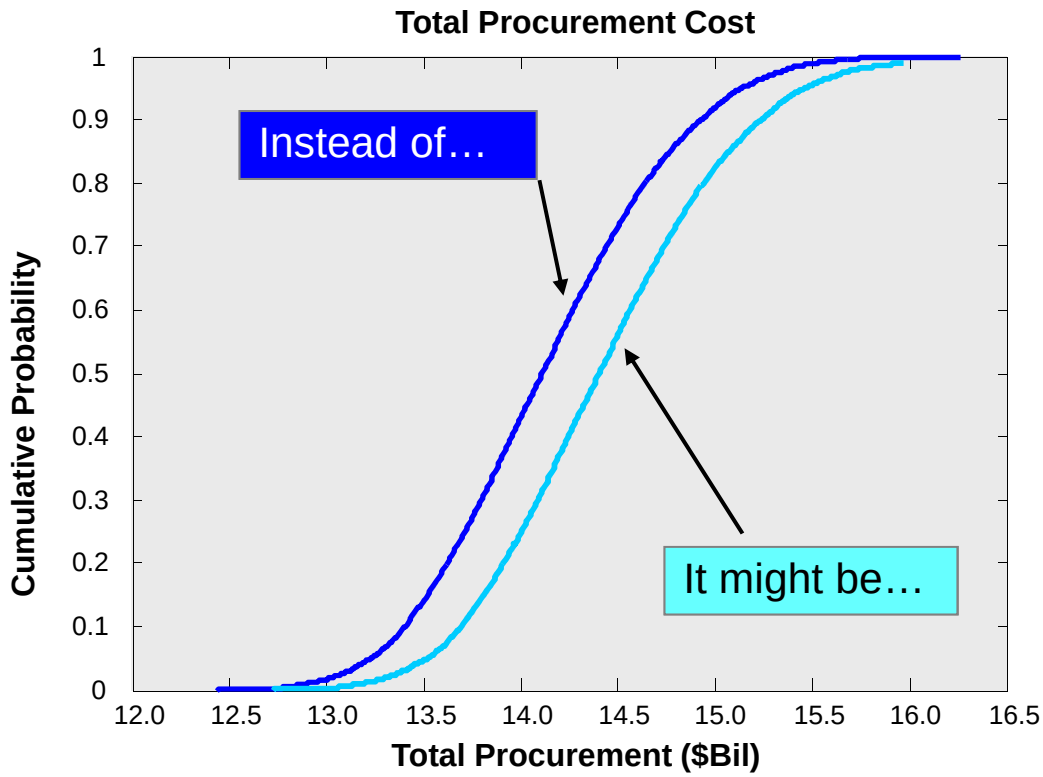
INTRODUCTION

- UNCERTAINTY/RISK ANALYSIS IN COST ESTIMATION WAS MEANT TO CONVEY IMPRECISENESS IN AN ESTIMATE
- THIS HAS MORPHED INTO AN ABSOLUTE PROBABILITY ASSOCIATED WITH SPECIFIC DOLLAR VALUES IN THE MINDS OF DECISION MAKERS
- MOST RECIPIENTS OF THE COST MODEL OUTPUT DON'T UNDERSTAND FULLY WHAT IS PRESENTED
- INSTEAD OF A POINT ESTIMATE WE NOW HAVE A POINT ESTIMATE WITH A MONTE CARLO SIMULATION MAKING IT A BETTER POINT ESTIMATE
- A POTENTIAL TOOL TO HELP DIFFUSE SOME OF THIS CERTAINTY IS THE P-BOX

P-BOXES: OUTLINE

- S-CURVES CAPTURE ONLY ONE TYPE (THERE ARE 3) OF UNCERTAINTY (*STATISTICAL UNCERTAINTY*) PRESENT IN THE COST ESTIMATE
- ALSO PRESENT IS *EPISTEMIC UNCERTAINTY* OR *TYPE 2*
- P-BOXES ARE CURVES THAT ARE UPPER AND LOWER BOUNDS FOR THE S-CURVE, SHOWING THE EPISTEMIC UNCERTAINTY
- INSTEAD OF THE NUMBER OF OBSERVATIONS (A COMMON PARAMETER IN STATISTICAL TOOLS) WE USE THE PROGRAM'S AGE
- STEPS TO CONSTRUCT BOUNDS:
  - KOLMOGOROV-SMIRNOV BOUNDS
  - QUANTILE BOUNDS BASED ON ORDER STATISTICS
  - COMBINING BOUND TYPES BY ENVELOPING
- EXAMPLE

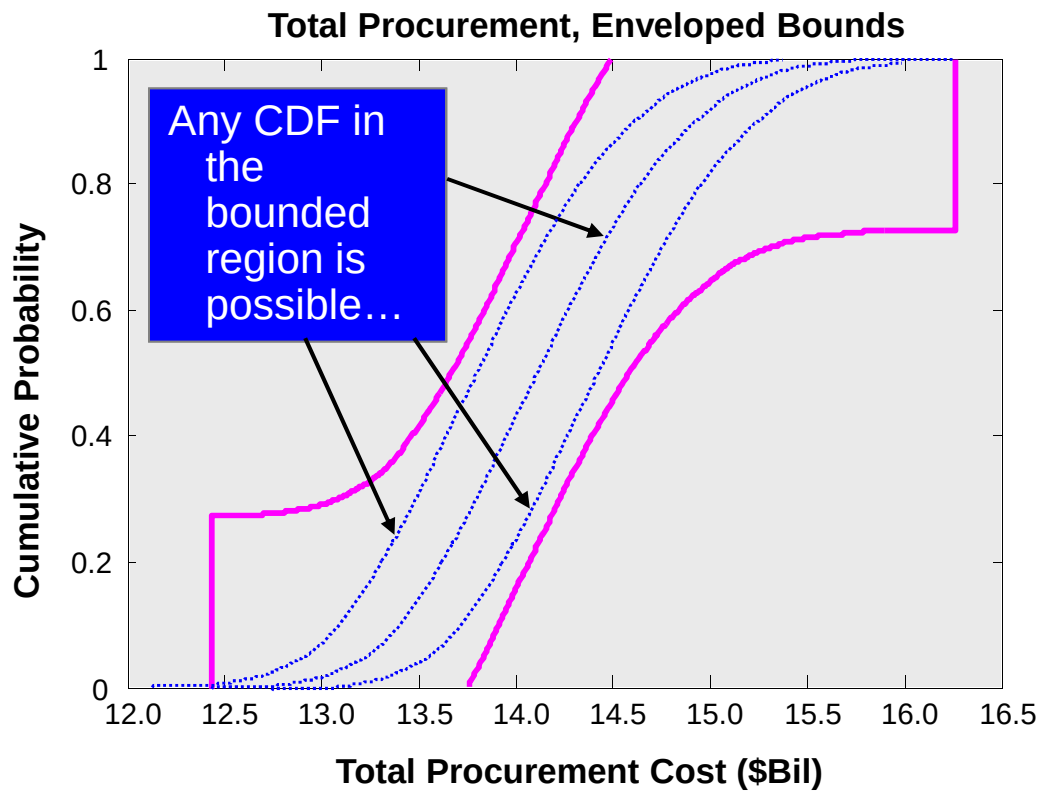
S-CURVES:
PART OF THE UNCERTAINTY PICTURE



- The S-curve is a cumulative distribution function for an unknown random variable ***Cost***, the statistical uncertainty inherent in the cost estimate
- The S-curve fully represents the uncertainty involved ONLY IF all sources of uncertainty are:
  - Known
  - Modeled correctly
  - Fixed, unchanging over time
- BUT…

- …**NONE** of those assumptions are true
  - Most sources of uncertainty aren't known; there are "unknown unknowns"
  - Those that are known may not be modeled correctly; different analysts choose different distributions for the same cost drivers
  - Requirements, material and labor rates, and other factors change during the life of the program

S-CURVE BOUNDS



- By bounding the S-curve using a tool called a *p-box*, we can capture more of the uncertainty
- A p-box consists of an *upper and lower bound* for the S-curve
- With some specified probability, the "true" S-curve, which we would get if we accurately knew and modeled all sources of uncertainty, lies somewhere between these bounds
- The bounds narrow as the program progresses (and uncertainty decreases)

- Rather than a single curve representing the uncertainty, we get a region that we know contains the true S-curve with a specified probability (90% for example)
- Bounds capture both the fact that the cost is uncertain, and that the processes driving the uncertainty and how it should be modeled are unknown
- The bounds show this epistemic or type 2 uncertainty

S- CURVE AS EDF

EDF

Empirical Distribution Function

- Most often formed from a random sample
- Converges to true CDF as the number of observations increases

- **We treat the S-curve as a kind of *EDF*, an approximation to the "true" cumulative distribution function**
  - Different analysts or information could result in a different S-curve
    -- a different approximation to the same "true" CDF
- **Usually, the EDF is formed from a random sample; it converges to the true CDF as the number of observations increases**
  - We use program age in lieu of number of observations
- **We construct bounds from the EDF for a region which contains the true CDF with some probability**

HOW DO WE CONSTRUCT BOUNDS?

Types of P-boxes

1) Kolmogorov-Smirnov bounds
2) Quantile bounds

- **We construct two types of bounds from the EDF for a region which contains the true CDF with some probability:**
  - Kolmogorov-Smirnov bounds: vertical bounds for the probability (risk) at fixed quantiles (cost)
  - Quantile bounds: horizontal bounds for quantiles (cost) at fixed probabilities (risk)
- **Both types of bounds are non-parametric**
- **Both bound types have number of observations as a parameter**
  - Exploiting the relationship between information, uncertainty, and observations, we derive a proxy value for observations

OBSERVATIONS AND UNCERTAINTY

Statistical Applications

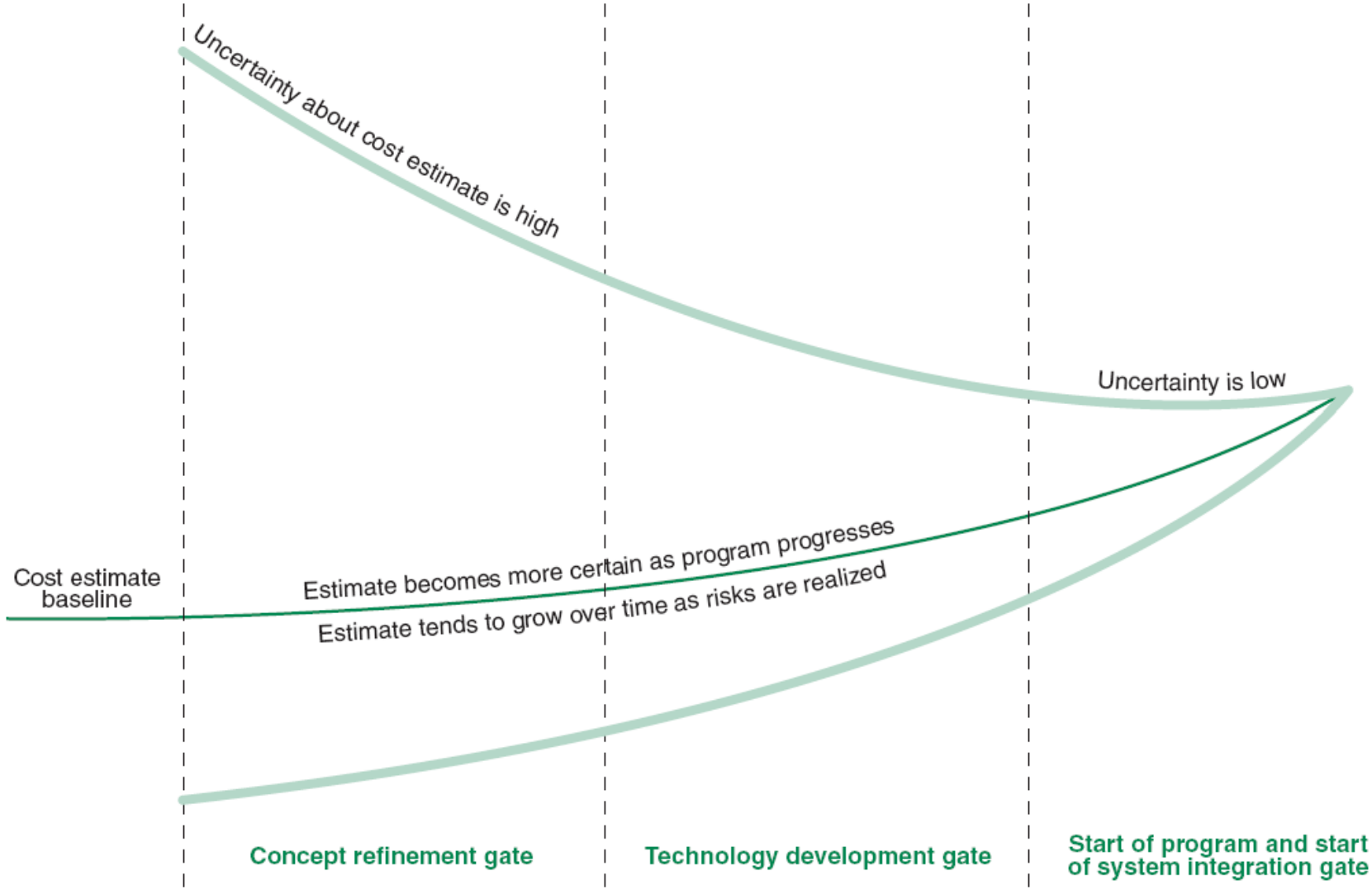
- Each data point is an observation in a random sample

Cost Models: no samples

- Information from disparate sources

- **In most statistical applications, we have a population or process from which we collect data; each data point is an observation in a random sample**
  - The more observations, the more information
  - The more information we have, the less uncertainty
- **For a cost model, no random sample, instead information from disparate sources, like expert opinion or analogy with past programs**
  - Distributions and associated parameters chosen to model the uncertainty
  - Data becomes available and uncertainty decreases with life of program
- **In lieu of random observations, we use the program's age to determine a proxy value for observations**

UNCERTAINTY DECREASES, INFORMATION INCREASES



Source: GAO.

FINDING THE EFFECTIVE OBSERVATIONS VALUE

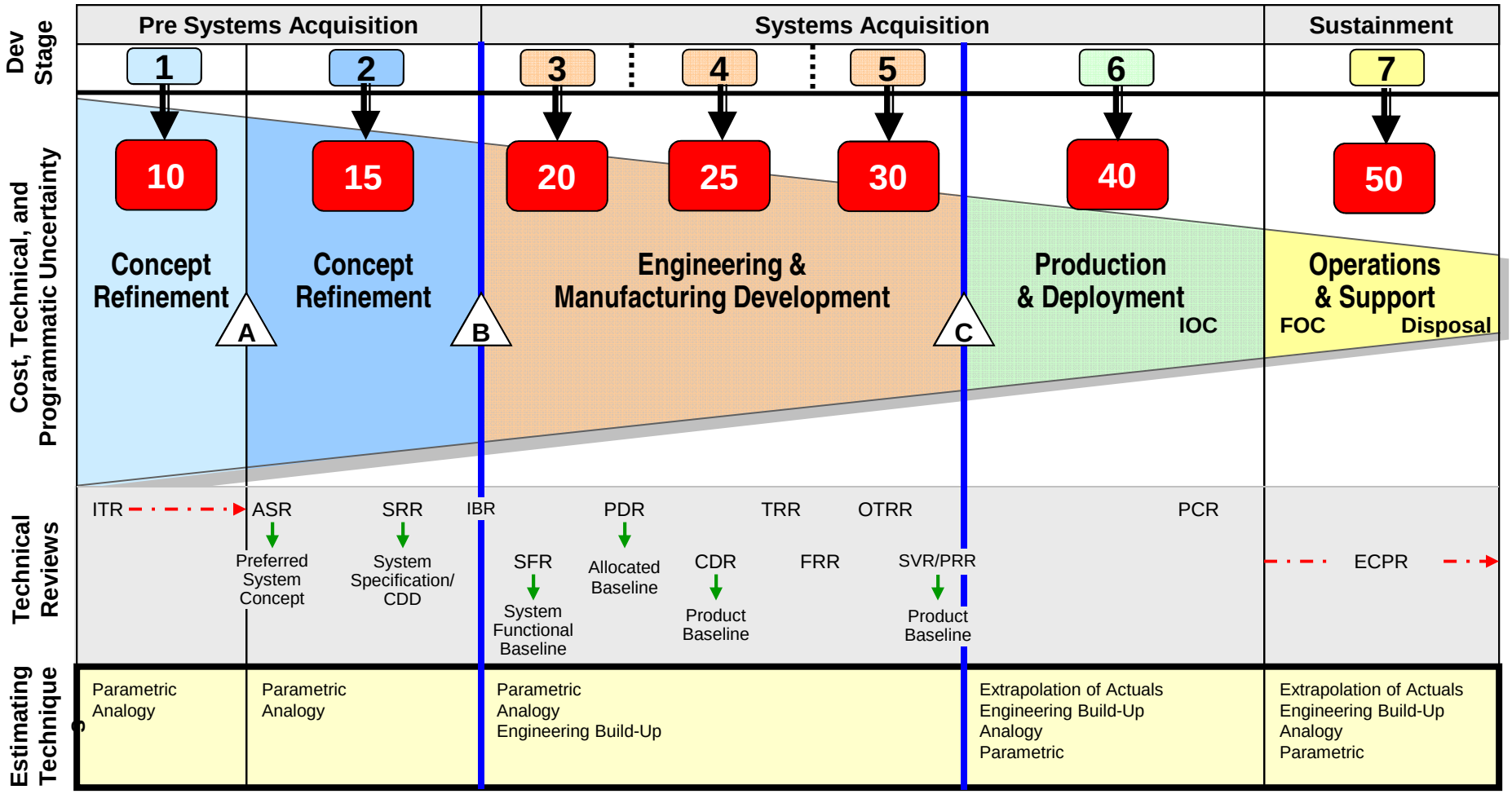
Effective Observations

- "Stabilization point" around 25
- Magnitude effect diminished over time
- Baseline of 25 effective observations in stage 4 (near CDR)
- Obs>25 in later stages; Obs<25 in earlier stages

- **Used cost data from several existing programs**
  - Ranging from pre-milestone B to programs in production
  - Program life cycle divided into 7 stages, pinned to technical reviews and decision milestones
- **We calculate the percent range (bound width), ranging effective observations from 10 to 50 to quantify the effect**
- **Analysis revealed following key points about the effect of number of observations:**
  - At all stages of program development, there was a "stabilization point" around 25 beyond which the effect on bounds width is small, relative to effect before this point
  - Magnitude of effect diminished as programs progress; choice of observations less important for older programs
  - Most data for programs in stage 4 (near CDR), giving baseline of 25 observations
  - Programs in later stages have Obs>25, those in earlier stages have Obs<25
- **The resulting equivalence between program life cycle and number of effective observations is shown on the next slide**

PROGRAM UNCERTAINTY VS. TIME

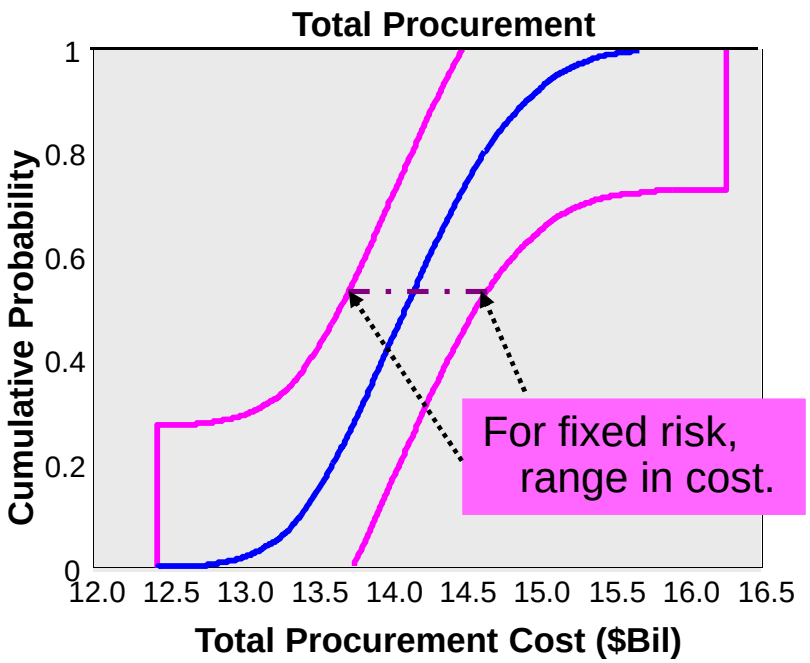
O_{eff} = f(DEVELOPMENT STAGE)



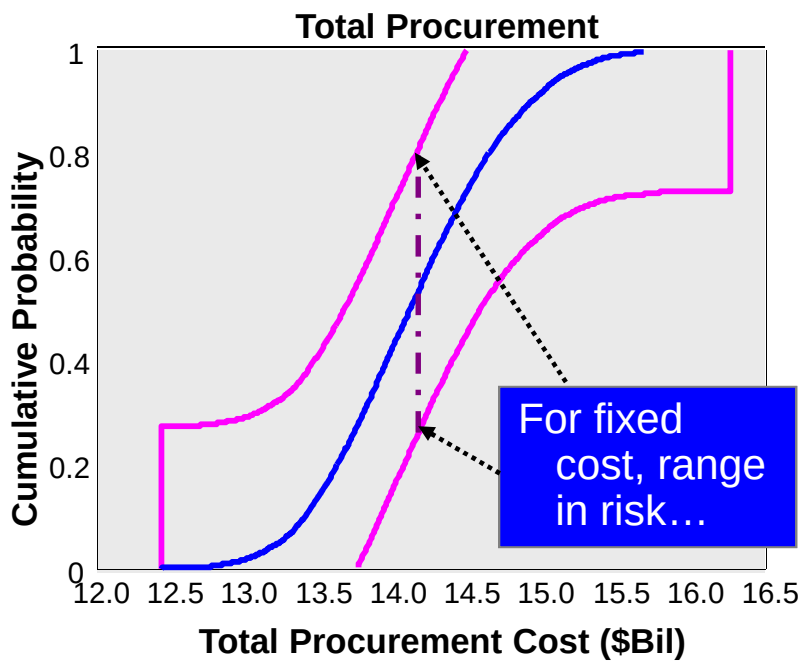
UNCERTAINTY LESSENS AND INFORMATION INCREASES OVER TIME, O_{eff} INCREASES

CONSTRUCTING BOUNDS FOR S-CURVES

- **We construct bounds from the EDF for a region which contains the true CDF with some probability. Two ways of constructing non-parametric bounds:**
  - **Kolmogorov-Smirnov bounds:** vertical bounds for the probability (risk) at fixed quantiles (cost)
  - **Quantile bounds:** horizontal bounds for quantiles (cost) at fixed probabilities (risk)
- **Bound width is a function of a program's stage in it's life cycle.**



Quantile Bounds



Kolmogorov-Smirnov Bounds

KOLMOGOROV-SMIRNOV P-BOX

- **Treating the S-curve as an empirical distribution function (EDF) enables use of the EDF's convergence properties**
- **In particular we use convergence for continuous CDFs *F(x)* to the Kolmogorov Distribution of:**

$$\sqrt{n}\left\|\hat{F}_n(x) - F(x)\right\|_\infty$$

- **Using the γ variate from the Kolmogorov Distribution we can construct Kolmogorov-Smirnov Bounds:**

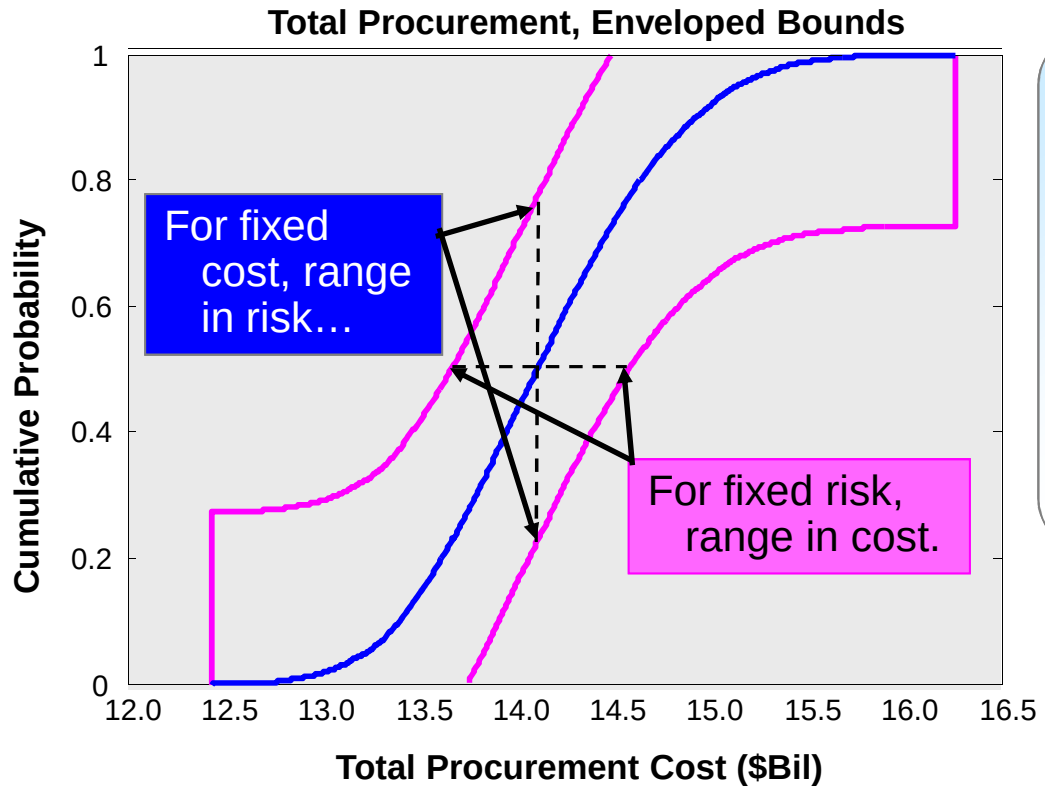
$$\Pr[\max(0, \hat{F}_n(x) - \frac{k_\gamma}{\sqrt{n}}) \le F(x) \le \min(1, \hat{F}_n(x) + \frac{k_\gamma}{\sqrt{n}})] \approx \gamma$$

QUANTILE INTERVAL P-BOX

- QUANTILE INTERVAL BASED P-BOXES ARE CREATED BY DIVIDING THE S-CURVE INTO O_{EFF} EQUIPROBABLE REGIONS, PULLING A COST VALUE FROM EACH REGION TO CREATE A PSEUDO-SAMPLE, AND USING THE ORDER STATISTICS TO FIND INTERVALS FOR EACH QUANTILE

- FOR O_{EFF} EFFECTIVE OBSERVATIONS, THE NUMBER OF VALUES FROM THE PSEUDO-SAMPLE LESS THAN THE Q^{TH} QUANTILE WILL FOLLOW A BINOMIAL DISTRIBUTION WITH PARAMETERS O_{EFF} AND Q

THE BENEFITS OF S-CURVE BOUNDS



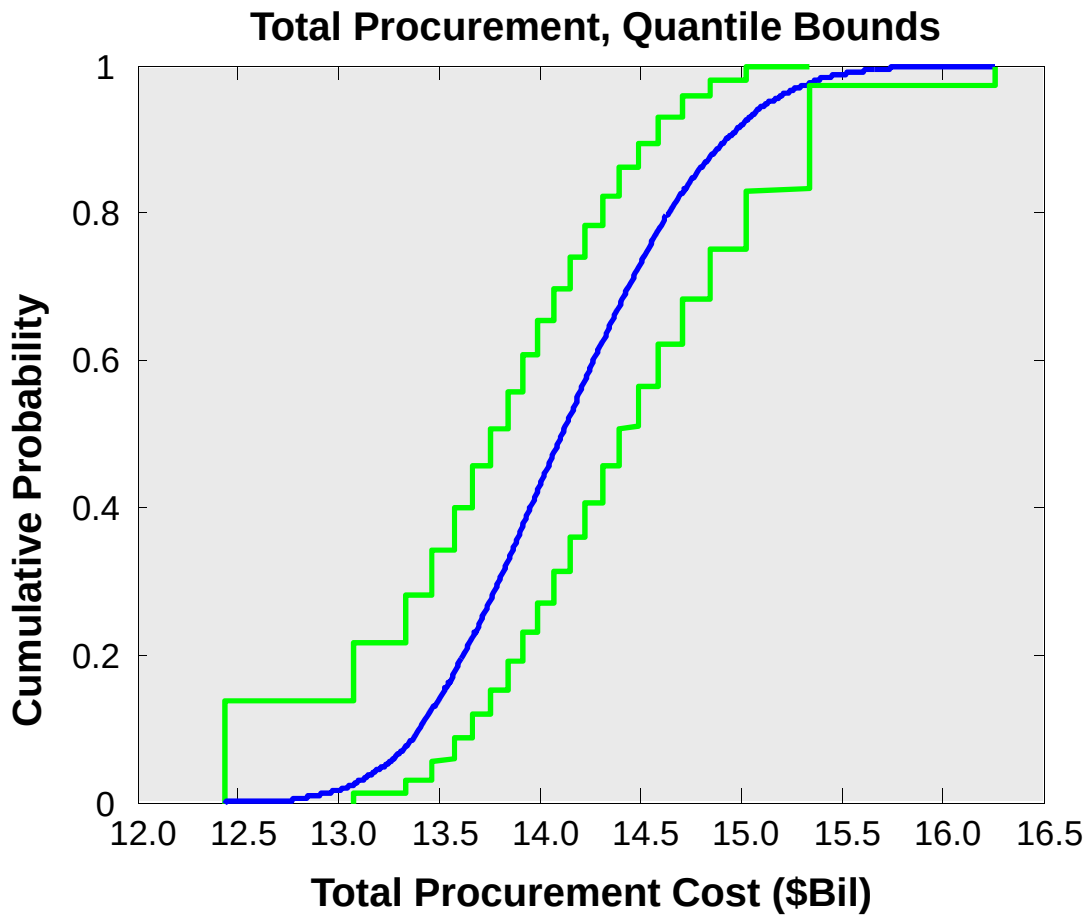
- Recent acquisition reforms (e.g. ACQ Reform 2009) call for disclosing confidence levels with the cost estimate.
- In particular, there is a focus on determining a "confidence level" for cost estimates of 80%.
- Bounds convey the uncertainty in both cost and risk.

- The bounds help us answer two distinct, key questions about uncertainty:
  - For a fixed dollar value (cost) what is the range in probability (or risk)?
  - For a fixed level of risk (probability) what is the range in cost?
- The bounds and derived ranges show the upside risk to the cost; i.e., how bad do we think it could get?

BOUNDS APPLIED TO AN EXAMPLE

QUANTILE BOUND P-BOX

TOTAL PROCUREMENT COST ($TY)



- Confidence Level = 90%
- Effective Observations = 20

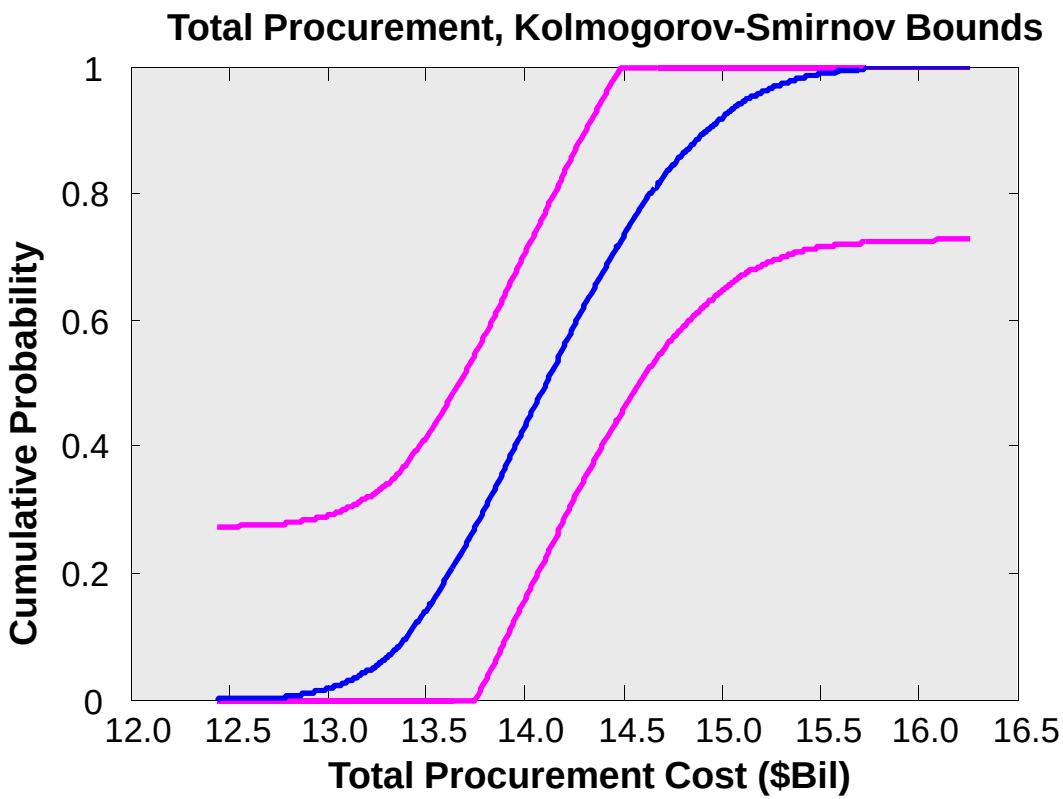
- Cost Ranges (for fixed P, range in cost):

Bound/ Percentile	Lower ($Bil)	Upper ($Bill)
20th	13.08	13.91
50th	13.75	14.40
80th	14.31	15.03

For a fixed level of risk, what is the range in cost?

KOLMOGOROV-SMIRNOV P-BOX

TOTAL PROCUREMENT COST ($TY)

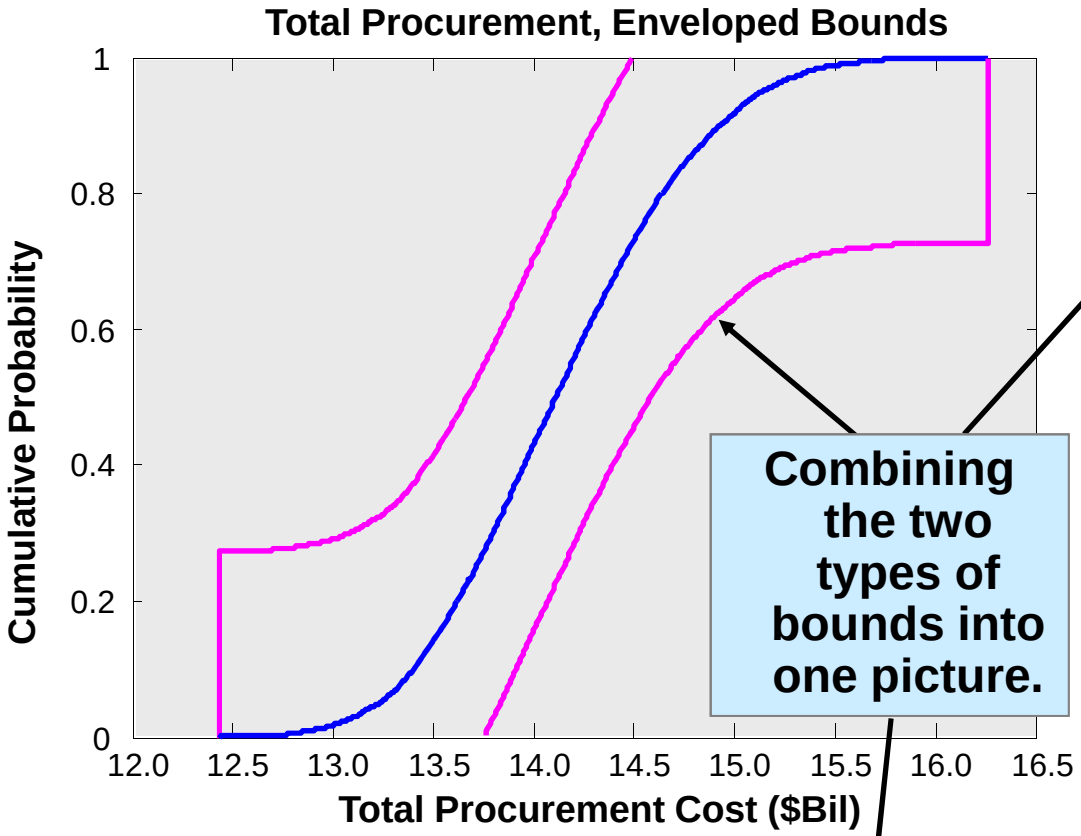


- Confidence Level = 90%
- Effective Observations = 20

For a fixed cost, what is the range in risk?

Cost ($Bill)\Range	Lower	Upper	Risk Percentile
13.62	0%	47.37%	20%
14.11	22.63%	77.37%	50%
14.64	52.63	100%	80%

ENVELOPED P-BOX:
COMBINING KS AND QUANTILE



Combining the two types of bounds into one picture.

- Confidence Level = 90%
- Effective Observations = 20

Bound/ Percentile	Lower ($Bil)	Upper ($Bill)
20th	12.43	14.06
50th	13.66	14.58
80th	14.15	16.25

Range/Cost ($Bill)	Lower	Upper	Risk Percentile
13.62	0%	47.37%	20%
14.11	22.63%	77.37%	50%
14.64	52.63	100%	80%

CONCLUSION

- S-CURVES ALONE DO NOT CAPTURE ALL OF THE UNCERTAINTY PRESENT IN THE COST ESTIMATE
- P-BOXES HELP CAPTURE THE *EPISTEMIC UNCERTAINTY* THAT IS ALSO PRESENT
- CHANGES IN THE TECHNIQUES USED FOR DEVELOPING COST MODELS, I.E. PRODUCING "BETTER" S-CURVES, WILL NOT ELIMINATE UTILITY OF P-BOXES
  - PERIODIC UPDATING OF THE EFFECTIVE OBSERVATIONS ANALYSIS WILL ENSURE MORE ACCURATE BOUNDS AS COST ESTIMATING TECHNIQUES PROGRESS
  - EPISTEMIC UNCERTAINTY WILL ALWAYS BE PRESENT

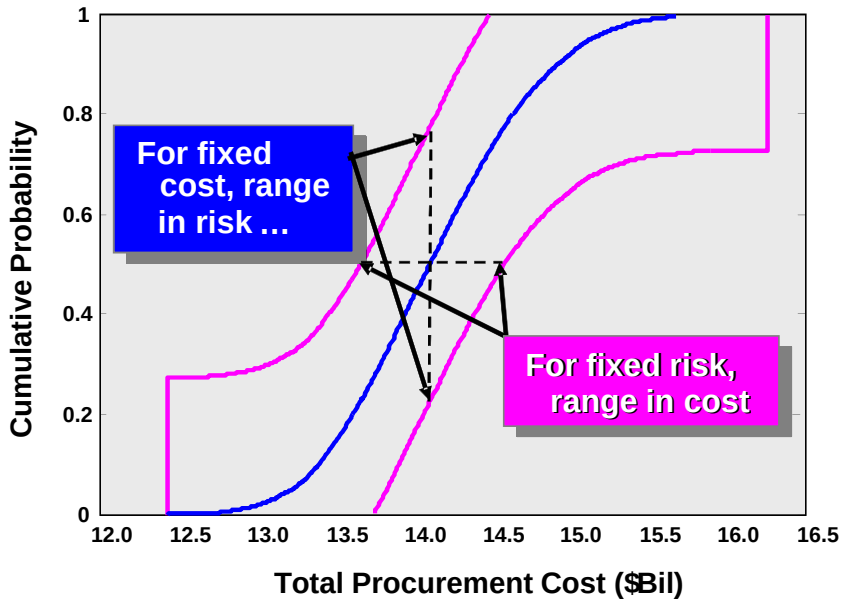
OTHER

THE BENEFITS OF S-CURVE BOUNDS

COST RISK ANALYSIS

- Developed a method called the P-Box to measure uncertainty around a models "S-Curve" output
  - Utilizes Quantile and Kolmogorov-Smirnov measures to develop horizontal and vertical bounds around a risk output
- Has a surrogate time function based on program maturity to adjust range of bounds
- Attempts to counter the implication of absolute certainty in the dollar value at a given percent probability

Total Procurement, Enveloped Bounds



- **Recent acquisition reforms (e.g. ACQ Reform 2009) call for disclosing confidence levels with the cost estimate**
- **In particular, there is a focus on determining a "confidence level" for cost estimates of 80%**
- **Bounds convey the uncertainty in both cost and risk**

- **The bounds help us answer two distinct, key questions about uncertainty:**
  - **For a fixed dollar value (cost) what is the range in probability (or risk)?**
  - **For a fixed level of risk (probability) what is the range in cost?**
- **The bounds and derived ranges show the upside risk to the cost; i.e., how bad do we think it could get?**