Department of the Navy Cost Analysis Symposium

Notional Case Study

Estimating the Cost of the Intra-theater Fast Ship

17 September 2009
Quantico, Virginia

presented by:

Marc Greenberg, NCCA
Kathlyn Loudin, DAU

Background

- **Disclaimer**
- **DoDI 5000.02 (8 Dec 08)**
- **Weapon Systems Acquisition Reform Act**
- **Why We're Here (Objective)**

Background: Disclaimer

Disclaimer

- All data used to develop cost estimates and cost estimating relationships (CERs) were collected (and normalized) using public-domain sources.
- Platforms in this case study are fictitious.
  - The Intra-Theater Fast Ship (IFS)
  - The American National Defense Ship (ANDS)
    - Used as an analogy for the Threshold IFS
  - The Bluewater Unmanned Technology Ship (BUTS)
    - Used as an analogy for the Objective IFS

To put it simply, Navy Acquisition Programs no longer accept IFS, ANDS or BUTS!!!

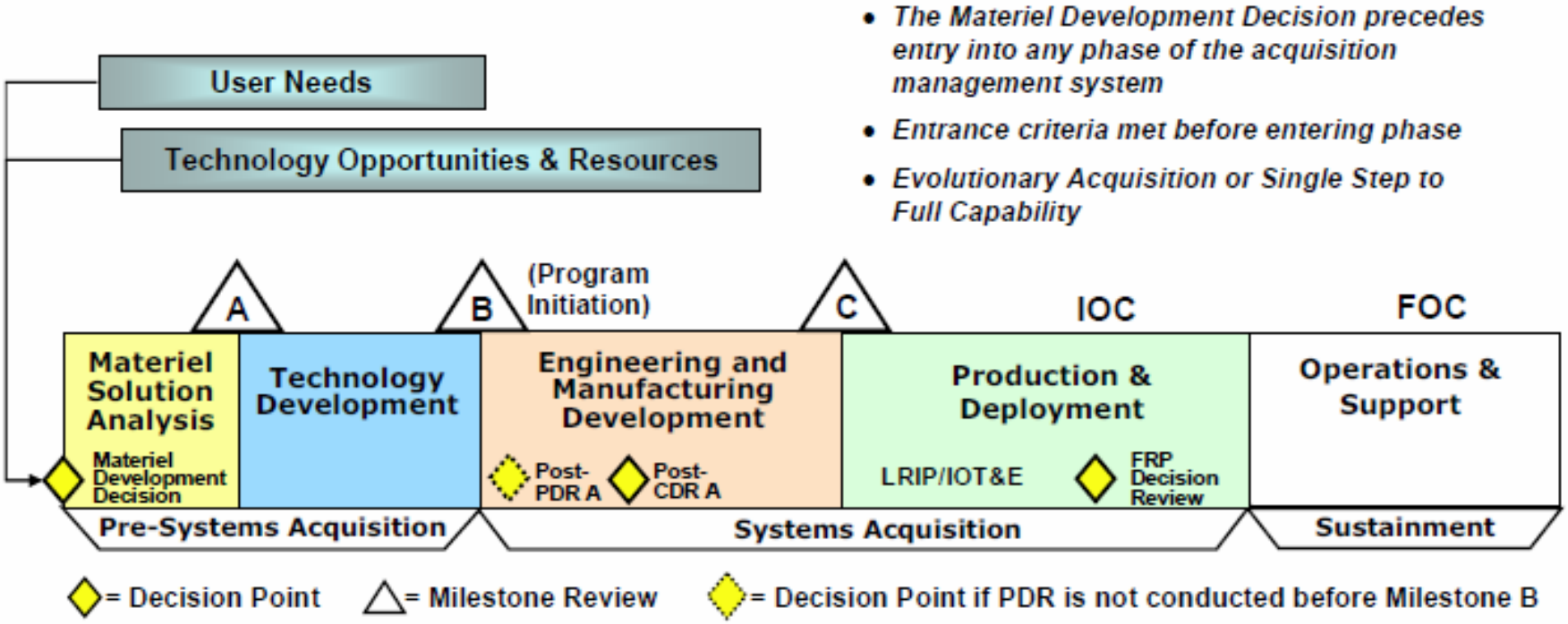
Background: DoDI 5000.02 (8 Dec 08)

Policy flows from revisions to Public Law since 2003, numerous policy memos, six NDAAs, and DoD responses to GAO, IG, and Congress.

Importance of Sound Cost Estimating…

- **MDA Certification at Milestones A & B**
- **Nunn-McCurdy Breach**
  - **25% over current APB cost**
  - **50% over original APB cost**

Background: Defense Acquisition Management Framework (8 Dec 2008)



Analysis of Alternative (AoA) Materiel Solutions shall be completed prior to the Milestone A Decision Point

Background: Weapon Systems Acquisition Reform Act of 2009 (Public Law 111-23, 22 May 2009)

Sec. 101. Cost Assessment & Program Evaluation

Sec. 103. Root Cause Analysis

Sec. 201. Trade-Offs in Cost, Schedule, Performance

Sec. 204. Milestone A Certification

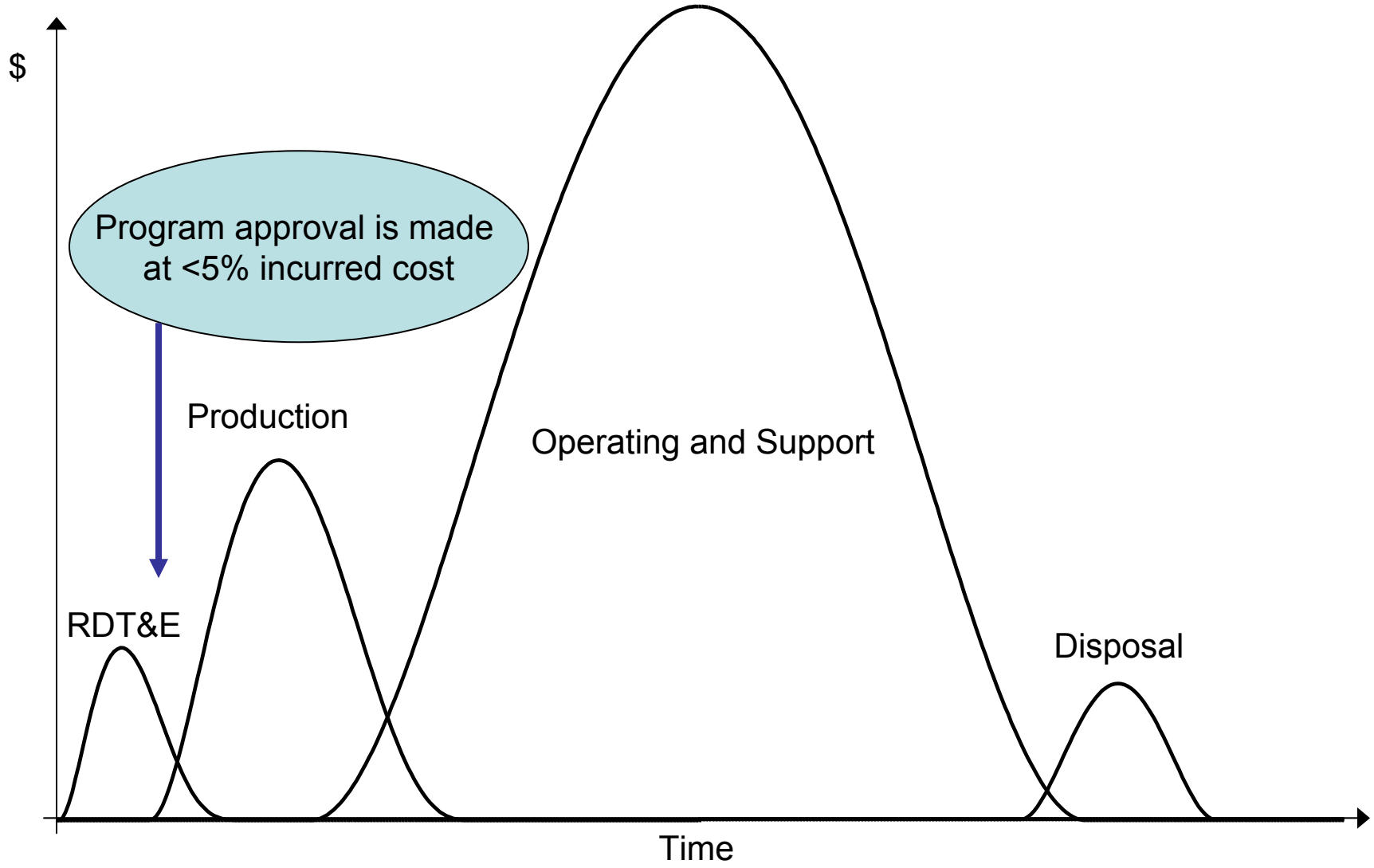
- Cost Estimate
- Adequate Resource Level (RDT&E/Procurement)

Sec. 205. Milestone B Certification

- Business Case Analysis (Affordability Assessment, Cost /Schedule Estimate)
-Funding available for RDT&E / Procurement in the FYDP

Increased Confidence – Programs Executed on Time and Within Budget!

Background: Life Cycle Cost vs. Time



So what might be a cost estimator's paradox prior to program approval? (Hint: Pre-Milestone A)

Background: Why we're here

The Case Study is NOT intended to provide...

***Detailed* instruction on common methods of data collection, data normalization, basic statistics, regression or uncertainty analyses.**

- We will address these topics. However, we will not go too deeply into all the "nuts and bolts" of each
  - Example: Teaching JUST statistics can easily take up a full week!!!

- Such detailed instruction is provided by SCEA, Defense Acquisition University, and other academic institutions
  - SCEA Lessons: *CostProf and CEBoK* (*www.sceaonline.org*)
  - DAU Courses: *BCF 106 and BCF 204* (*www.dau.mil*)

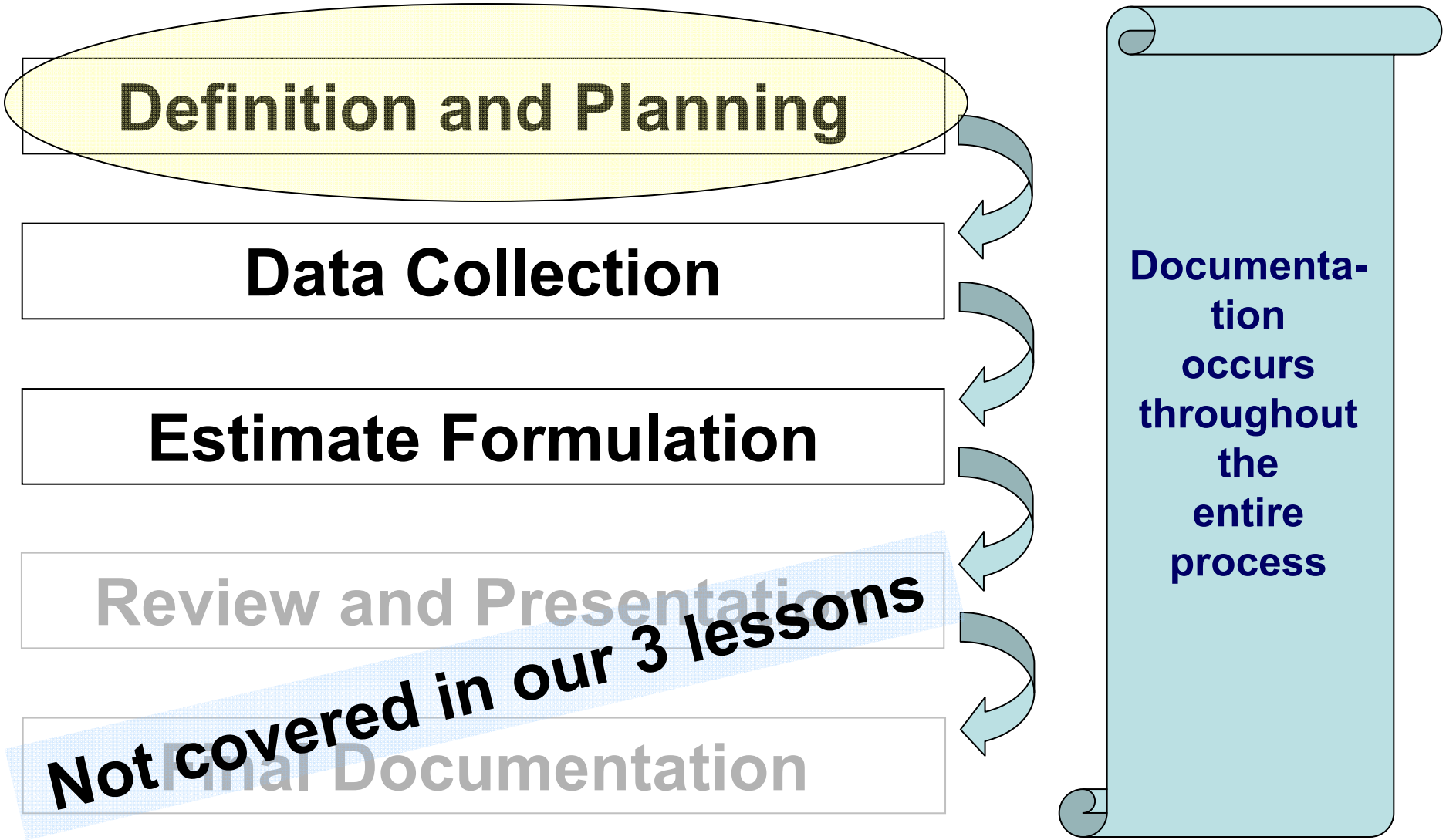
Background: Why we're here

<u>Learning Objective</u>. ***The intent of this Case Study is:***

... to discuss ways the cost analyst can develop a cost estimate with minimal guidance or data by

- Framing the task by working with experts / asking questions
- Collecting data when little or no data are readily available
- Evaluating data critically (prior to analyzing it)
- Analyzing data
  - Applying one or more cost estimating methods to produce an estimate.
  - Applying cost risk methods to account for uncertainty in the estimate
- Communicating the process, best CER and cost estimate.
- Communicating cost estimating uncertainty and risk

The Cost Estimating Process



Notional Case Study

- **Lesson 1 – Definition & Planning**
  - **Scenario for the Cost Estimator (the task & status)**
  - **Identify experts; meet w/them to elicit information**
  - **Create requirements sheet**
- **Lesson 2 – Developing the Cost Estimate**
  - **Data Collection**
  - **Statistics-based Cost Estimate**
  - **Regression-based Cost Estimate**
- **Lesson 3 – Estimating & Describing Cost Risk**
  - **Prediction Interval about the Cost Estimate**
  - **Changes in Cost due to Requirements Uncertainty**
    - **Method 1: Sensitivity Analysis**
    - **Method 2: Monte Carlo Simulation**

We also set aside 20 minutes after Lesson 3 to answer any questions related to Lessons 1 - 3

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Task

- User-needs are <u>currently</u> being determined for a Navy fast ship concept called the Intra-theater Fast Ship (IFS)
- You've agreed to support the Intra-theater Fast Ship (IFS) Program Office as their sole cost analyst.
- Within your 1st week of supporting the Program Office, the IFS Program Manager (PM), Wyatt Zomuchinkos, asks you one question:

How much should I request to fund the first IFS?

Before you attempt to respond to his question, you scramble to collect available information on the IFS ...

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Status

- *The Intra-theater Fast Ship (IFS) will be the Navy's premier 21st century platform to support forward presence in regions of conflict. [Website under construction; more information coming soon!!!]*
  - *Source: the IFS official website*

This is it? You're kidding me, right? Nope ...
Indeed, this is **all** the **"readily available"** information!
So **NOW** what do you do?

For starters: You find the right experts and start asking the right questions ...

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Experts

Wyatt Zomuchinkos

Program Manager, IFS Program Office

Stanley Schooner

Lead Naval Architect, IFS Program Office

Wendy Kounts

Head of Machinery Division, NAVSEA Carderock

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: Initial Questions

Q1. Will the IFS replace an existing Navy class of ships?
Q2. What will be its maximum speed?
Q3. What type of propulsion system will it have?
Q4. What type of hull form will it likely have?
Q5. What materials will be used for hull structures?
Q6. Do you know its draft, beam, LOA, etc?
Q7. Do you have any weight estimates for the IFS?
Q8. How many crew members are needed?
Q9. What is the range of the IFS?
Q10. When is the expected award date? IOC date?
Q11. How many IFS will be delivered?
Q12. What's the expected service life of the IFS?

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Answers

Q1. Will the IFS replace an existing Navy class of ships?

- **Yes.** The IFS is the replacement for the American National Defense Ship (ANDS). Source: IFS PM

Q2. What will be its maximum speed?

- Have a max speed **equal to the ANDS** but preferably **25% faster** than the ANDS. Source: W. Kounts, NAVSEA

Q3. What type of propulsion system will it have? **Not sure.**

Q4. What type of hull form will it be? **Monohull**

Q5. What materials will be used for hull structures?

- Structures will be **steel** and/or **aluminum** only

Q6. Do you have any weight estimates for the IFS? **Not yet.**

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Answers (cont'd)

Q7. Do you know its draft, beam, LOA, etc.

- IFS will have a draft **no more than FFG7** but preferably **equal** to the Bluewater Unmanned Technology Ship

Q8. How many crew members are needed?

- **Equal** to the crew of the **ANDS** but preferably **10% fewer** crew members than the ANDS

Q9. What is the range of the IFS?

- **At least** that of the BUTS but preferably **equal** to FFG7 range

Q10. When is the expected award date? IOC date?

- The IFS will be awarded in Dec 2011 with an IOC by Nov 2014

Q11. How many IFS will be delivered? **TBD.**

Q12. What's the expected service life of the IFS? **30 years.**

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: The Answers (cont'd)

Additional Information Gleaned from Discussions:

- The PM wants an answer in less than 8 weeks
  - This is after you negotiated up from his original 4 week deadline
  - He wants recurring costs broken out in budget format
- IFS weight and power data will be provided soon
- The CARD is not yet available; will be available soon
- Builders will be given a great deal of latitude in design
- "Objective" IFS data will be available in 8 weeks
- Throughout process, you found (and made corrections to) inconsistent information provided by the experts
  - You are responsible to check realism of experts' information

Based upon the top-level inputs we received on the IFS Program, we can develop a "Requirements" sheet ...

Lesson 1: Definition & Planning

Scenario for the Cost Estimator: Ship Design Specs

Input Parameter	Options or Description	Threshold Monohull	Rationale, source or additional notes
Qualitative Design Input:			
Platform Standards	Commercial, More Robust, Military	**More Robust**	Navy mil-spec with some commercial & USCG specs
Hull Form	Monohull or Catamaran	**Monohull**	Expert opinion of Stanley Schooner, IFS Prog Office
Hull Material	Steel, Aluminum or Composite	**Steel**	Expert opinion of Stanley Schooner, IFS Prog Office
Superstructure Material	Steel, Aluminum or Composite	**Steel or Alum**	Expert opinion of Stanley Schooner, IFS Prog Office
Propulsion Plant	Diesel, Diesel-Electric, GT, CODAG	**Not sure**	Expert opinion of Wendy Kounts, NAVSEA Carderock
Propulsion Method	Fixed pitch, CR propeller, or waterjets	TBD	
Quantitative Design Input:			
Quantity	# of hulls delivered	TBD	
Hull Service Life	years	**30**	Based upon American National Defense Ship
Speed	Maximum Speed in knots	**32**	Based upon American National Defense Ship
Draft at full load displacement	meters	**5.8**	Threshold value based on FFG7
Beam	meters	TBD	
Length Overall LOA	meters	TBD	
Installed SHP	same as plant rating	TBD	
Number of Shafts or Waterjets	Waterjets	TBD	
Number of Crew	Live aboard	**70**	Based upon American National Defense Ship
Range @ 20 knots	Nautical Miles at economical speed	**3200**	Thresh based on Bluewatr Unman Tech Ship
Helo Support? (Yes or No)		**Yes**	Based upon American National Defense Ship
Weights by Navy SWBS:			
1-digit weights (100 - 700)	metric tons	TBD	
2-digit weights (110 - 790)	metric tons	TBD	
Subtotal	metric tons	TBD	
Weight Margin	Mil Allowance as % of (Sum 100-700)	TBD	
Lightship Weight (WLS)	metric tons	TBD	
Full Load Displacement (WFL)	metric tons	TBD	
Programmatic Input:			
Initial Operating Capability	Yr when Platform is fully operational	**2015**	
Year Dollars for AoA	Depicted as Constant FY$10M	**2010**	For comparative analyses purposes only
Year Dollars for Budget	Depicted as Then-Year $12M	**2012**	For budgeting purposed only

What could be unintended "utilities" of this spec sheet?

Lesson 1: Definition & Planning

End Cost (Breakout of a typical P-5 Exhibit)

Shipbuilder Related
- Plans
- **Basic Construction** Conversion
- Change Orders
- Escalation

Combat Systems / GFE / CFE
- Electronics
- HM&E
- Ordnance
- Propulsion

Other Costs
- Other

Summation = End Cost

SWBS 100, Hull Structure
SWBS 200, Propulsion Plant
SWBS 300, Electric Plant
SWBS 400, Command & Control
SWBS 500, Auxiliary Systems
SWBS 600, Outfit & Furnishing
SWBS 700, Armament

SWBS 800, Engineering
SWBS 900, Assembly

Facility Cost of Money

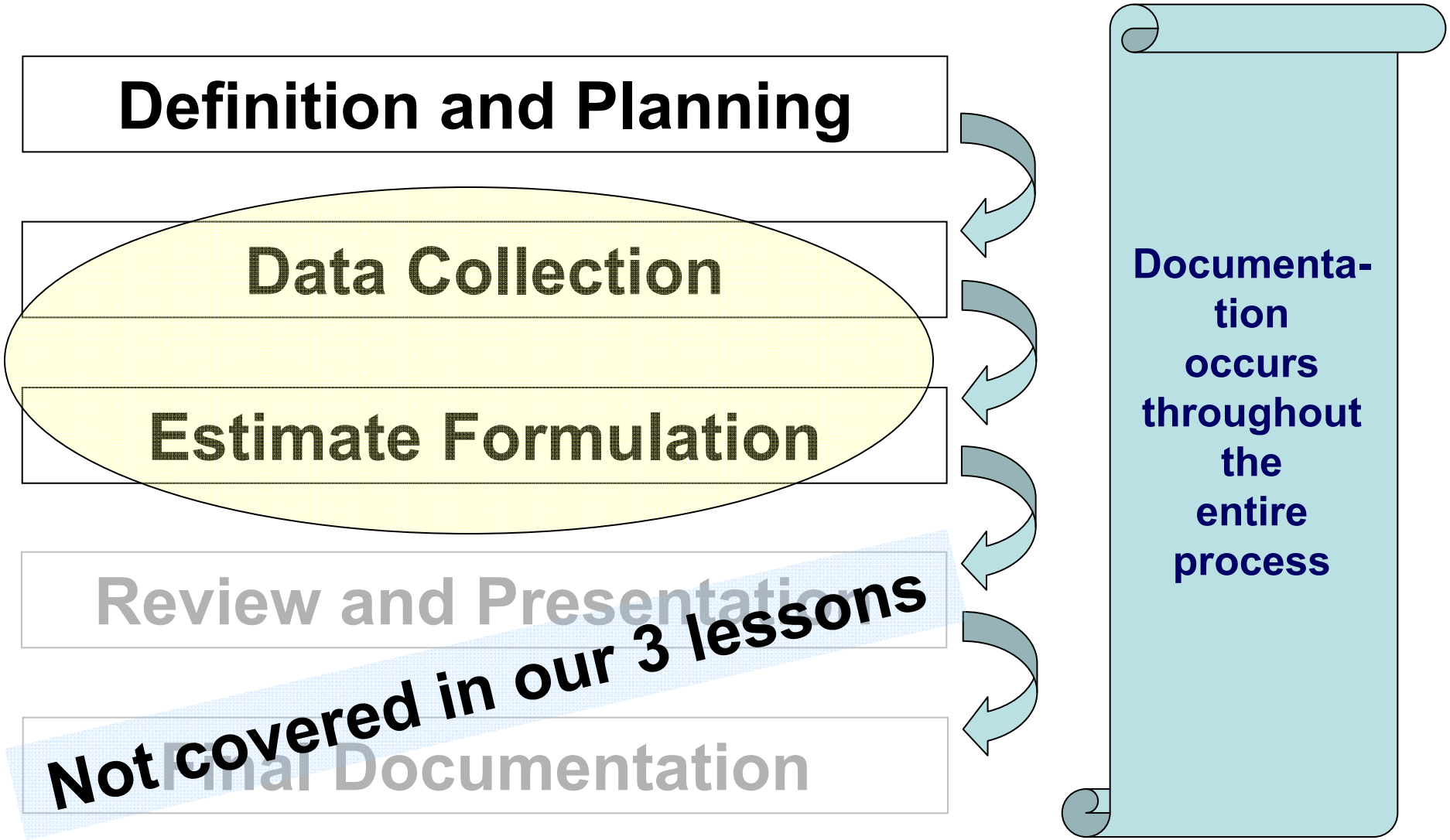
Profit

We will only estimate Basic Construction for this study

Notional Case Study

- Lesson 1 – Definition & Planning
  - Scenario for the Cost Estimator (the task & status)
  - Identify experts; meet w/them to elicit information
  - Create requirements sheet
- **Lesson 2 – Developing the Cost Estimate**
  - **Data Collection**
  - **Statistics-based Cost Estimate**
  - **Regression-based Cost Estimate**
- Lesson 3 – Estimating & Describing Cost Risk
  - Prediction Interval about the Cost Estimate
  - Changes in Cost due to Requirements Uncertainty
    - Method 1: Sensitivity Analysis
    - Method 2: Monte Carlo Simulation

The Cost Estimating Process



Lesson 2 - Developing the Cost Estimate

Data Collection

- What are data?
  - Cost
    - Contractor Cost Data Reports (CCDR), Cost Performance Reports (CPR), Contracts, Contractor Accounting Data, Forward Pricing Rates, Labor Rates, Material Costs, etc.
  - Technical
    - Weight, power, and size budgets, drawing counts, source lines of code (SLOC), Reports Interfaces Conversions Extensions (RICE objects), function points, etc.
  - Programmatic
    - Program Management, System Test, Budget, co-dependencies, head-counts
  - Schedule
    - Detailed Integrated Master Schedule
- Commonly used databases & authoritative sources:
  - PBIS, VAMOSC, DAMIR, NVR, DoN Budgets, IMS, etc.

Lesson 2 - Developing the Cost Estimate

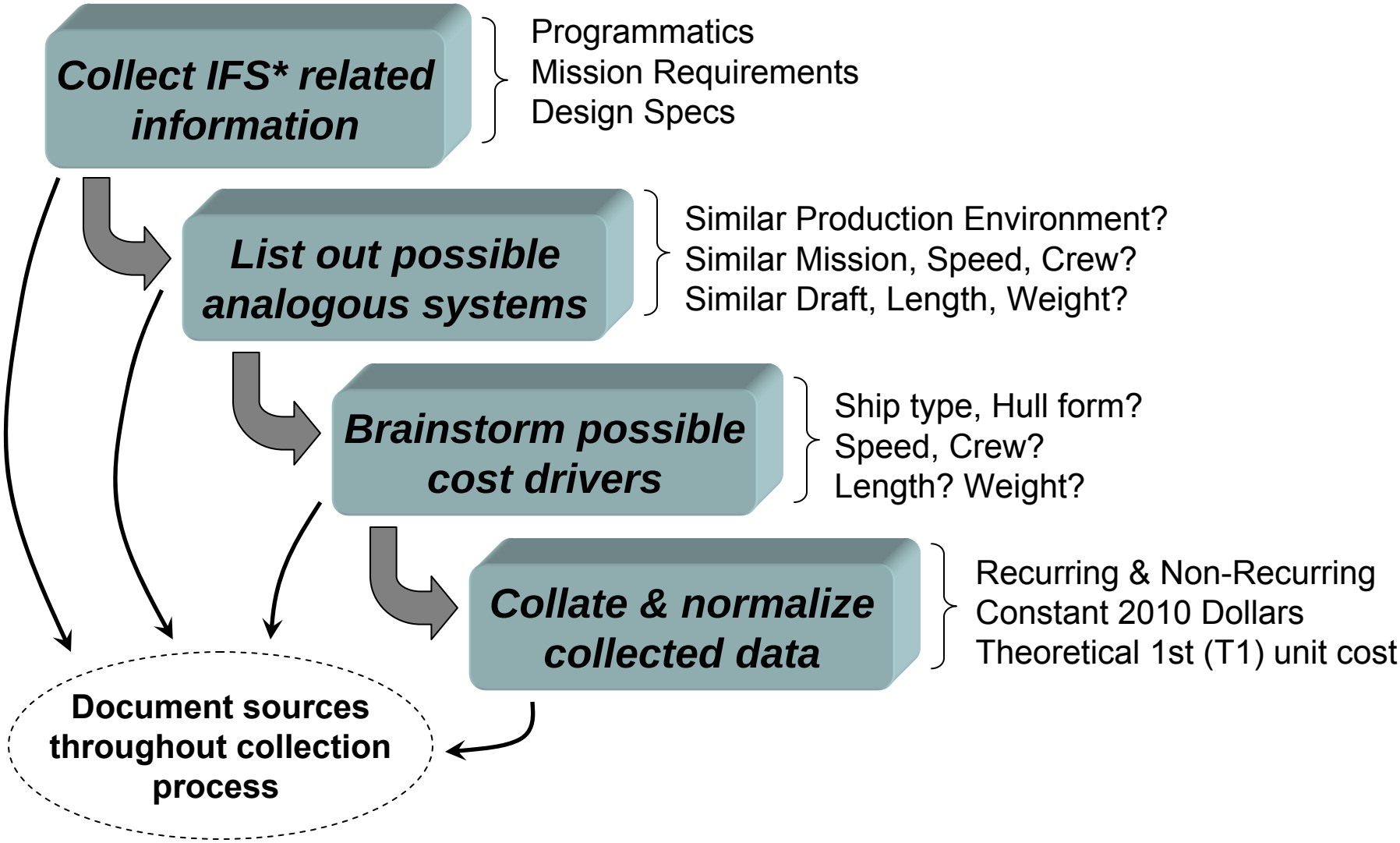
Data Collection

- Where do we get data?
  - Program Office
  - Prime Contractor
  - Major Sub-contractors
  - Industry
  - Similar program offices
  - Other cost estimating organizations
  - Websites

Because this is a hypothetical study, presented in an open forum, all data were collected from open sources

Lesson 2: Developing the Cost Estimate

Data Collection



*** Note: Intra-theater Fast Ship (IFS) information was described and collected in Lesson 1**

Lesson 2 - Developing the Cost Estimate

Data Collection

- Normalizing Data
  - Data are “mapped” in the correct category
  - Cost is normalized to a defined base year
  - Convert hours to cost using actual contractor and/or analogous rates
  - Technical data are normalized to Meters, Metric Tons, Pounds, Watts, Effective Source Lines of Code, and other common technical bases
  - Analysis of outliers and sample populations
- Data Collection and Normalization often consumes up to 80% of our estimate development timeline

It is important to (a) be aware of major design characteristics of the ship and (b) identify performance parameters that may impact cost

Lesson 2 - Developing the Cost Estimate

Data Requirements (per discussions with experts)

- All ships must be US-built for Navy or USCG
- All ships must have been awarded on or after 1970
- Hull forms are monohull only
- Hull materials are either steel or aluminum
- Max speed must exceed 15 knots but not exceed 40 knots
  - *Based on Wendy Kounts follow-up input on IFS speed*
- Lightship weight can not exceed 6,000 metric tons
  - *Based on Dr. Schooner's follow-up input on IFS sizing*
- Costs will be theoretical first unit (T1) Basic Construction Cost
  - In Constant Year FY$10M
  - Using analogy, "Basic" will later be used to estimate "End" cost

If all ships meet the data requirements, we will describe the data as "homogeneous"

Lesson 2 - Developing the Cost Estimate

10 Ship Classes (or Flights) Considered

	Class or Flight	Class Name	Hull Type
Combatant & Frigate	DD 963	Spruance	Guided Missile Destroyer
	FFG 7	Oliver Hazard Perry	Guided Missile Frigate
Patrol-like ships & Icebreakers	PC 14	Cyclone	Coastal Patrol
	WMEC 901	Bear	Medium Endurance Cutter
	WAGB 10	Polar Icebreaker	Icebreaker
	WLBB	Mackinaw	Icebreaker
	FRC	Sentinal	Fast Response Cutter
	WMSL 750	National Security Cutter	National Security Cutter
Mine-related ships	MCM 1	Avenger	Mine Countermeasure
	MHC 51	Osprey	Coastal Minehunter

Now let's look more closely at the data ...

Lesson 2 - Developing the Cost Estimate

Original Data (v.1): Sample Size = 10

Tech = Award - 1970

Class or Flight	Award	IOC	Basic$_T1	Draft	LOA	Beam	WLS	WFL	SHP	Range	Speed_ Max	Crew	Tech
DD 963	1970	1978	$ 459.60	5.8	171.7	16.8	5,034	8,040	80,000	6,000	33	334	0
FFG 7	1974	1978	$ 284.54	5.8	135.6	13.9	2,892	3,705	41,000	4,200	29	225	4
MCM 1	1985	1993	$ 156.87	4.0	68.3	11.9	1,272	1,333	2,400	2,500	14	84	15
MHC 51	1993	2000	$ 156.25	3.4	57.3	11.6	659	851	1,600	2,500	12	52	23
PC 14	1997	2000	$ 94.19	2.6	54.6	7.6	358	393	14,400	2,500	35	39	27
WMEC 901	1978	1985	$ 113.43	4.3	82.3	11.6	1,322	1,809	7,200	10,250	20	100	8
WAGB 10	1993	2000	$ 396.73	8.9	128.0	25.0	11,215	16,663	30,000	16,000	17	137	23
WLBB	2002	2006	$ 105.18	4.9	73.2	17.8	2,394	3,556	13,518	4,000	15	55	32
FRC	2008	2011	$ 95.40	2.9	46.8	7.7	273	353	14,600	2,950	28	22	38
WMSL 750	2005	2009	$ 344.45	6.4	127.4	16.5	3,536	4,572	40,220	12,000	28	148	35
		Units:	CY$10M	m	m	m	mtons	mtons	shp	naut mi	knots	billets	yrs

Is the data set homogeneous? No. Why?

1. **WAGB 10's lightship weight > 6,000 mtons**
2. **MCM 1 & MHC 51 each have max speeds < 15 knots**
3. **MCM 1 & MHC 51 hulls are not steel or aluminum**

Lesson 2 - Developing the Cost Estimate

Updated Data (v.2): Sample Size = 7

Class or Flight	Award	IOC	Basic$_T1	Draft	LOA	Beam	WLS	WFL	SHP	Range	Speed_ Max	Crew	Tech
DD 963	1970	1978	$ 460	5.8	171.7	16.8	5,034	8,040	80,000	6,000	33.0	334	0
FFG 7	1974	1978	$ 285	5.8	135.6	13.9	2,892	3,705	41,000	4,200	29.0	225	4
PC 14	1997	2000	$ 94	2.6	54.6	7.6	358	393	14,400	2,500	35.0	39	27
WMEC 901	1978	1985	$ 113	4.3	82.3	11.6	1,322	1,809	7,200	10,250	20.0	100	8
WLBB	2002	2006	$ 105	4.9	73.2	17.8	2,394	3,556	13,518	4,000	15.0	55	32
FRC	2008	2011	$ 95	2.9	46.8	7.7	273	353	14,600	2,950	28.0	22	38
WMSL 750	2002	2009	$ 344	6.4	127.4	16.5	3,536	4,572	40,220	12,000	28.0	148	35

- What are some pros and cons of this data?
  - Cons
    - Sample size = 7. Optimal size at least 30. Sub-optimal sample size appx 12 - 20.
    - Missions of each class vary, sometimes significantly
    - All ships built by different builders, with different production rates, in different competitive environments
  - Pros
    - We have more than one or two data points (e.g. regression vs analogy)!
    - The data have reasonable variation (range) in cost, design and performance
    - Three ships reflect design and costs of current shipbuilding environment

Question 1 (Topic: Normalizing Data)

1. ***Which are required for a regression dataset (for developing a Navy Ship CER)?***

   A. All ships are analogous to what you want to estimate

   B. All costs are in the same constant year dollars

   C. All systems are built for the US Navy

   D. All costs are Theoretical First Unit Cost

   E. A and B only

   F. All of the above

Answer: ***E. A and B only***

Lesson 2 - Developing the Cost Estimate

Updated Data (v.2) Dummy Variables

- A dummy variable indicates whether or not something is true
- "Dummy" is also called an "indicator" or "categorical" variable

<u>Example:</u> Our sample data contains 4 ships with a helo deck.

The dummy variable introduced takes the value of 1 if the ship has a helo deck, 0 if it does not have a helo deck

Class or Flight	Helo_Deck	USCG	Combat	Icebrker	Patrol	Non-Combat	St_Superst	Al_Superst	Diesel	Dies_Elec	GT	CODAG
DD 963	1	0	1	0	0	0	1	0	0	0	1	0
FFG 7	1	0	1	0	0	0	0	1	0	0	1	0
PC 14	0	0	0	0	1	1	0	1	1	0	0	0
WMEC 901	1	1	0	0	1	1	1	0	1	0	0	0
WLBB	0	1	0	1	0	1	1	0	0	1	0	0
FRC	0	1	0	0	1	1	0	1	1	0	0	0
WMSL 750	1	1	0	0	1	1	1	0	1	0	1	1

Notional Case Study

- Lesson 1 – Definition & Planning
  - Scenario for the Cost Estimator (the task & status)
  - Identify experts; meet w/them to elicit information
  - Create requirements sheet
- **Lesson 2 – Developing the Cost Estimate**
  - **Data Collection**
  - ⇨ **Statistics-based Cost Estimate**
  - **Regression-based Cost Estimate**
- Lesson 3 – Estimating & Describing Cost Risk
  - Prediction Interval about the Cost Estimate
  - Changes in Cost due to Requirements Uncertainty
    - Method 1: Sensitivity Analysis
    - Method 2: Monte Carlo Simulation

Lesson 2 – Developing the Cost Estimate

Types of Cost Estimates

- Analogy
  - Similar system – similar cost
- Parametric
  - Cost Estimating Relationships (CERs)
    - Derived using basic statistics or regression analyses
- Engineering / Detailed Bottoms Up
  - Detailed list of products and activities
    - Example: work packages
- Extrapolation from Actual Cost
  - Based on actual costs from current contract(s)

Lesson 2 - Developing the Cost Estimate

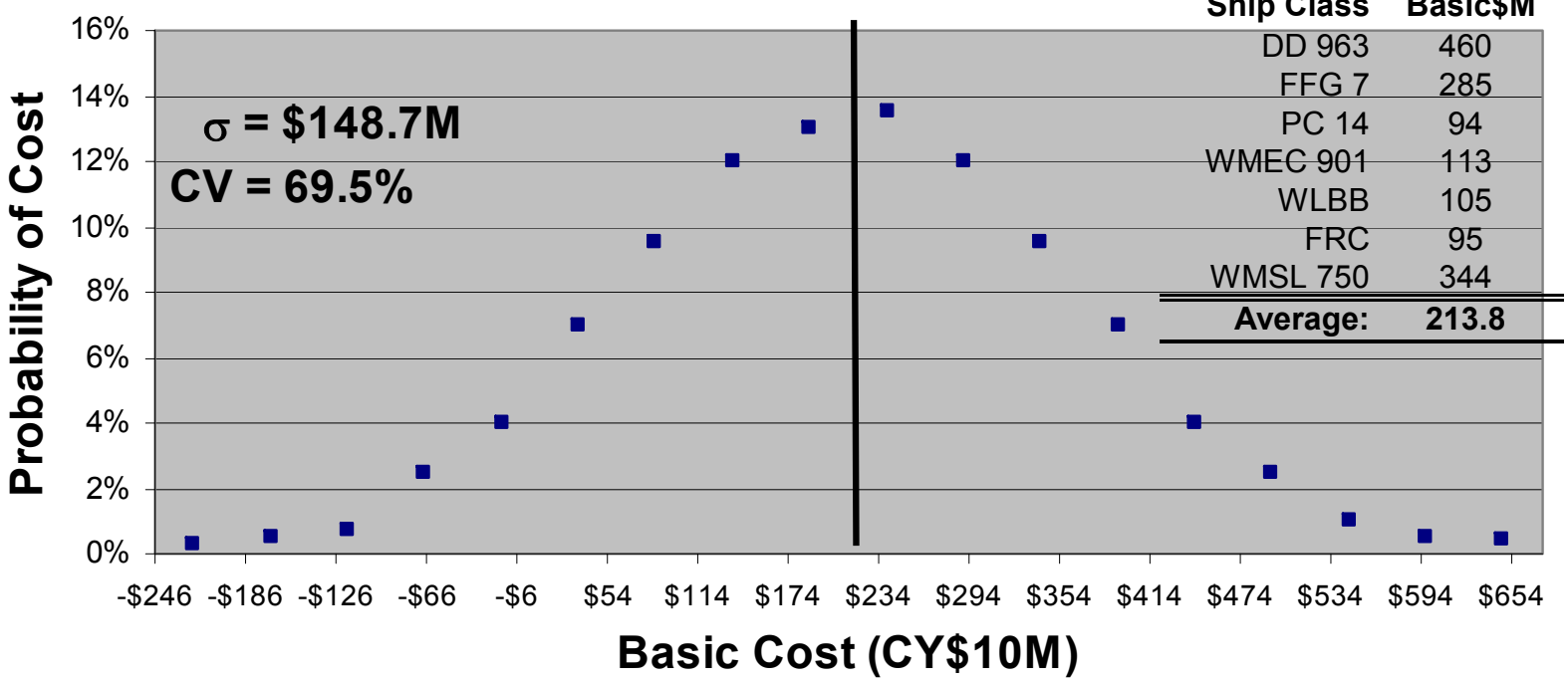
Basic Statistics: The Mean and Standard Deviation

Normal PDF

PDF of Average Basic Cost of $213.8M

(assumes our ship cost data are normally distributed)

Ybar= $213.8



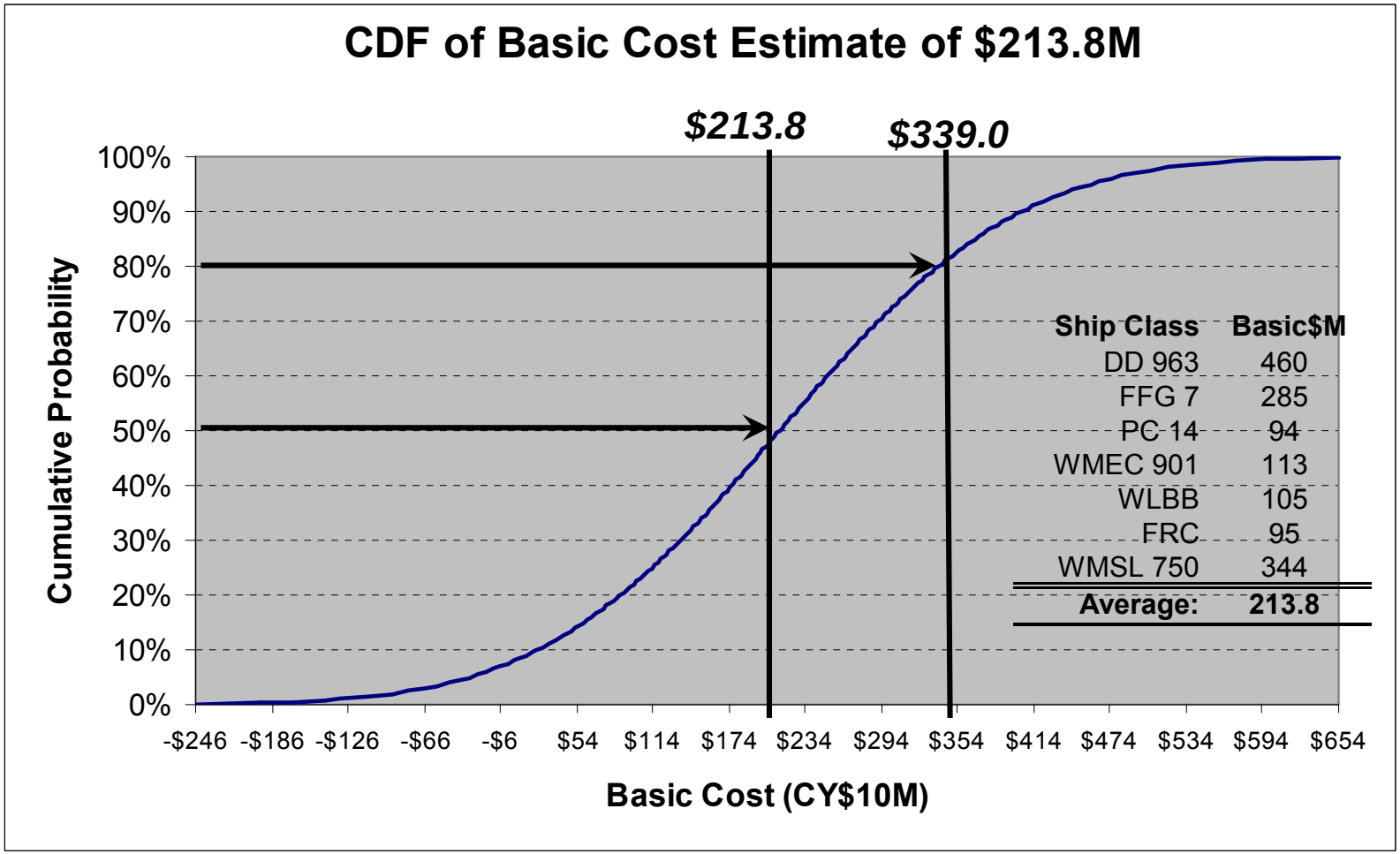
Ship Class	Basic$M
DD 963	460
FFG 7	285
PC 14	94
WMEC 901	113
WLBB	105
FRC	95
WMSL 750	344
Average:	**213.8**

- **Is there significant dispersion about the mean? Explain.**
- **What might be tough to explain to decision-makers?**

Herein, we'll assume our risk-adjusted estimates at p = 0.8. Such probabilities typically vary from program to program.

Lesson 2 - Developing the Cost Estimate

Basic Statistics: The Estimate at p = 0.80



Our risk-adjusted estimate at ($Cost_{p=0.8}$ = $339M) is much higher than our mean estimate ($Cost_{p=0.5}$ = $213.8M).

Lesson 2 - Developing the Cost Estimate

Estimating by Parametrics

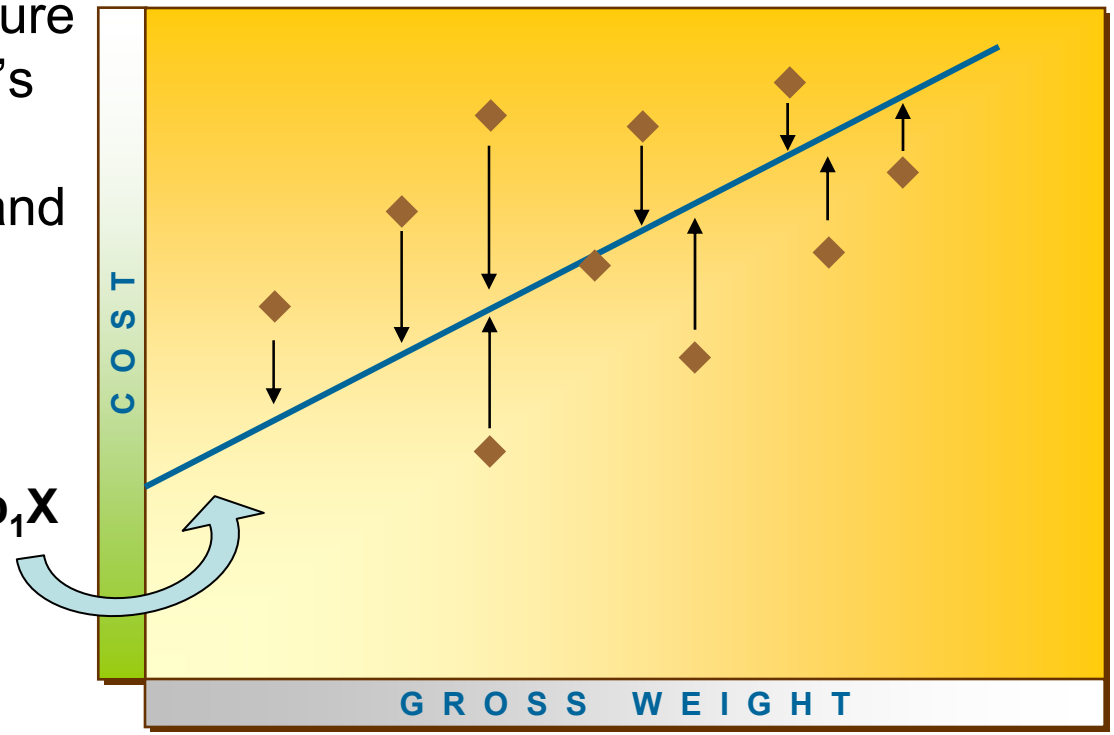
- What? A Cost Estimating Relationship (CER) based on historical trends across programs or systems
  - Cost of a system as a function of its physical or performance characteristics
- When? When we have sufficient data
  - Variables like range, speed and weight are available while more detailed information is limited
  - Useful prediction w/o drawings and standards
- How? Apply statistical methods to the costs of two or more analogous systems
  - Data normalized for economic year & quantity effects
  - The more data, the better (if reliable sources)

Lesson 2 - Developing the Cost Estimate

Estimating by Parametrics

Regression

... a calculation that predicts future occurrences based on the data's correlation, standard deviation, standard error of the estimate and the value of the independent variable.



Regression Formula: $Y = b_0 + b_1X$

Where:

Y = Dependent Variable
X = Independent Variable

b_0 **=** Y intercept
b_1 **=** Slope of the line

Lesson 2 - Developing the Cost Estimate

Parametric Cost Estimating Steps

1. Create a database
2. Adjust all costs to a common base year
3. Select variables that may influence cost

   Cost = function of SHP, Speed, Weight, etc
4. "Hypothesize" most-likely functional forms

   Cost = $b_0 + b_1 * (Weight)$ or … $b_0 + b_1 * (SHP)$
5. Apply regression methods to calculate CERs
6. Use regression statistics to select best CER

   *Cost = 0.017 * (SHP) + 58*
7. Ensure that selected CER seems logical

   Ex: As SHP increases, cost should also increase
8. To estimate cost, replace CER variable(s) with proposed system specifications

   *IFS Cost = 0.017 * (36,000) + 58 = $568 in FY$10M*

Lesson 2 - Developing the Cost Estimate

Step 3: Select Variables that May Influence Cost

- We have a sample size of 7 ships, n = 7
- Degrees of Freedom (df) = n - k - 1, where k = # of independent variables
  - *"Illustrating degrees of freedom in terms of sample size and dimensionality," by Dr. Chong Ho (Alex) Yu (2009)*
- **Our minimum df = 3**. Therefore we can use up to 3 independent variables for the given sample set
  - df = n - k - 1 = 7 - 3 - 1 = 3 df
- **Our maximum df = 5**. This is when just one independent variable is used to explain the variation in Basic $ (the dependent variable)
  - df = n - k - 1 = 7 - 1 - 1 = 5 df

All things being equal, how many independent variables are most preferred in our case? Why?

Lesson 2 - Developing the Cost Estimate

Step 3: Select Variables that May Influence Cost

Correlation Matrix

	Basic$ _T1	Draft	LOA	Beam	WLS	WFL	SHP	Range	Speed_ Max	Crew	Tech
Basic$_T1	1.00										
Draft	**0.80**	1.00									
~~LOA~~	~~**0.96**~~	~~**0.86**~~	~~1.00~~								
~~Beam~~	~~0.61~~	~~**0.88**~~	~~0.69~~	~~1.00~~							
~~WLS~~	~~**0.92**~~	~~**0.90**~~	~~**0.95**~~	~~**0.86**~~	~~1.00~~						
~~WFL~~	~~**0.91**~~	~~**0.83**~~	~~**0.92**~~	~~**0.83**~~	~~**0.99**~~	~~1.00~~					
~~SHP~~	~~**0.96**~~	~~0.66~~	~~**0.92**~~	~~0.53~~	~~**0.88**~~	~~**0.90**~~	~~1.00~~				
Range	0.37	0.58	0.38	0.40	0.39	0.32	0.14	1.00			
Speed_Max	0.43	-0.12	0.28	-0.38	0.09	0.10	0.50	-0.22	1.00		
Crew	**0.93**	**0.74**	**0.97**	0.57	**0.89**	**0.89**	**0.93**	0.24	0.36	1.00	
Tech	-0.50	-0.38	-0.66	-0.23	-0.49	-0.52	-0.54	-0.10	-0.18	**-0.77**	1.00

We also have 12 dummy variables!

- Shaded cells have high correlations
- Except for correlations with Basic$, |R| > 0.70 not recommended
- However, there are only a few combinations of independent variables that have an |R| < 0.70

Assuming linear and log linear forms, we have 3,306 possible CERs! How could we "narrow the field?"

Lesson 2 - Developing the Cost Estimate

Step 3: Select Variables that May Influence Cost

Correlation Matrix (of parameters specified for the IFS)

	Basic$ _T1	Draft	Range	Speed_ Max	Crew	Tech
Basic$_T1	1.00					
Draft	**0.80**	1.00				
Range	0.37	0.58	1.00			
Speed_Max	0.43	-0.12	-0.22	1.00		
Crew	**0.93**	**0.74**	0.24	0.36	1.00	
Tech	-0.50	-0.38	-0.10	-0.18	**-0.77**	1.00

We also have 12 dummy variables!

- Shaded cells have high correlations
- Except for correlations with Basic$, |R| > 0.70 not recommended

Due to limited data on the IFS, we pared the possible CERs to 1,666. How else can we "narrow the field?"

Question 2 (Topic: Multicollinearity)

Which of the following combinations of variables are least preferred? Why?

A. **Speed_ Max & Draft**

B. **Draft & Range**

C. **Crew & Tech**

D. **Draft & Tech**

E. **Draft, Crew & Helo Dummy Variable**

F. **None of the above (none are "least preferred")**

Answer: *C. Crew & Tech*

$|R|_{\text{Crew, Tech}} = 0.77$ is > 0.70

Lesson 2 - Developing the Cost Estimate

Step 6. Select the Best CER

1. Group by form. Linear, Power, etc.
2. Exclude equations that don't "violate" thresholds of t-test, F-test, correlation, and outlier measures
3. Of remaining equations, select the one in each group that has highest R^2_a*
4. Select equation with lowest Unit Space SE.* Also consider complexity, data range etc.
5. Ensure the equation is logical

* You can skip Step 4 by selecting lowest adj-R^2 if the adj-R^2 is in Unit Space

Lesson 2 - Developing the Cost Estimate

Steps 6 & 7: Select the Best CER & Check its logic

Model Type:	*Linear 1*	*Linear 2*	*Power 1*	*Power 2*
Equation	T1 Basic $ = $-489 + 86.47\ X_1 + 11.16\ X_2$	T1 Basic $ = $-365 + 73.4\ X_1 + 8.11\ X_2 + 132\ D$	T1 Basic $ = $0.179\ X_1^{1.771}\ X_2^{1.302}$	T1 Basic $ = $0.355\ X_1^{1.64}\ X_2^{1.139}\ 1.35^{D}$
Logical	Yes	Yes	Yes	Yes
adj-R^2 (Unit Space)	86.89%	98.45%	90.70%	99.55%
SE (Unit Space)	$53.85	$18.50	$45.36	$9.98
CV	25.18%	8.65%	21.21%	4.67%
degrees of freedom (df)	4	3	4	3
Complexity	2	3	3	4
Range	X_1 = 2.6 ---> 6.4 (m) X_2 = 15 ---> 35 (knots)	X_1 = 2.6 ---> 6.4 (m) X_2 = 15 ---> 35 (knots)	X_1 = 2.6 ---> 6.4 (m) X_2 = 15 ---> 35 (knots)	X_1 = 2.6 ---> 6.4 (m) X_2 = 15 ---> 35 (knots)
Legend	X_1 = Draft (m) X_2 = Speed_Max (knots)	X_1 = Draft (m) X_2 = Speed_Max (knots) D = Combatant (dummy)	X_1 = Draft (m) X_2 = Speed_Max (knots)	X_1 = Draft (m) X_2 = Speed_Max (knots) D = Combatant (dummy)

- Linear CERs tend to be simpler to explain. However, you may find it difficult to explain (a) an extremely negative intercept and (b) that any combatant you estimate would be $132M more than a non-combatant.
- "Power 2" is our preferred equation because it (a) can calculate cost based on given requirements, (b) has highest adj-R^2 (or lowest SE) and (c) appears to be logical (as each X_i increases, so should Basic $)

Lesson 2 - Developing the Cost Estimate

Describing selected CER to Decision-makers

$$\textbf{Basic\$_T1} = 0.355\ \textbf{Draft}^{1.64}\ \textbf{Speed_Max}^{1.139} * 1.35^{\textbf{Combat}}$$

This is a Power Model: $Y = AX^B$ where, in this case study ...
For every 1% change in Draft or Speed, there's some fixed % change in Cost

"B" depicts this ratio of % changes as the "elasticity of Y with respect to X":

$$B = \frac{\textit{percentage change in Y}}{\textit{percentage change in X}}$$

If all other independent variables were held constant ...

A 1% increase in Draft = 1.640% increase in cost
A 1% increase in Speed_Max = 1.139% increase in cost
If the ship is "Combatant-like," the cost increases by 35%

Notional Case Study

- Lesson 1 – Definition & Planning
  - Scenario for the Cost Estimator (the task & status)
  - Identify experts; meet w/them to elicit information
  - Create requirements sheet
- Lesson 2 – Developing the Cost Estimate
  - Data Collection
  - Statistics-based Cost Estimate
  - Regression-based Cost Estimate
- ⇨ Lesson 3 – Estimating & Describing Cost Risk
  - Prediction Interval about the Cost Estimate
  - Changes in Cost due to Requirements Uncertainty
    - Method 1: Sensitivity Analysis
    - Method 2: Monte Carlo Simulation

Lesson 3 - Estimating & Describing Cost Risk

Step 1: Use CER to Estimate Basic T1 Cost

$$\text{Basic\$_T1} = 0.355\ \text{Draft}^{1.64}\ \text{Speed_Max}^{1.139} * 1.35^{\text{Combat}}$$

IFS Draft = 5.8 meters

IFS Maximum Speed = 32 knots

Combatant Dummy = 0

Basic$_T1 = ??

$$\text{Basic\$_T1} = 0.355\ (\text{Draft})^{1.64}\ (\text{Speed_Max})^{1.139} * 1.35^{\text{Combat}}$$

$$\text{Basic\$_T1} = 0.355 * 17.866 * 51.804 * 1 = 328.6$$

Basic$_T1 = $328.6 in Constant FY$10M

Question 3 (Topic: Your CER Estimate)

Assuming that the distribution of residuals follows a normal distribution, what does our CER's Basic T1 estimate of $328.6M represent?

A. The most-likely estimate

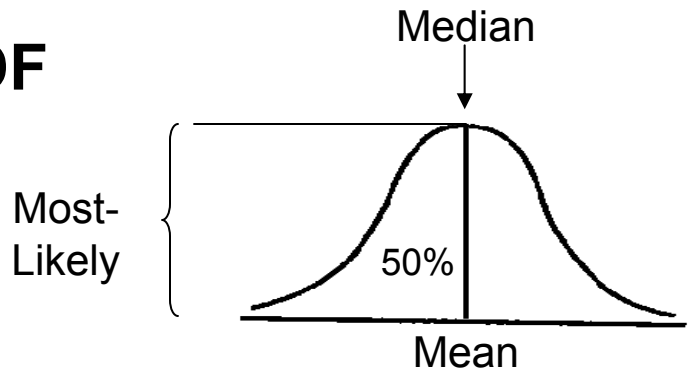
B. The mean estimate (i.e. average)

C. The estimate at p = 0.5 on CDF

D. The median estimate

E. All of the above

F. B & D only



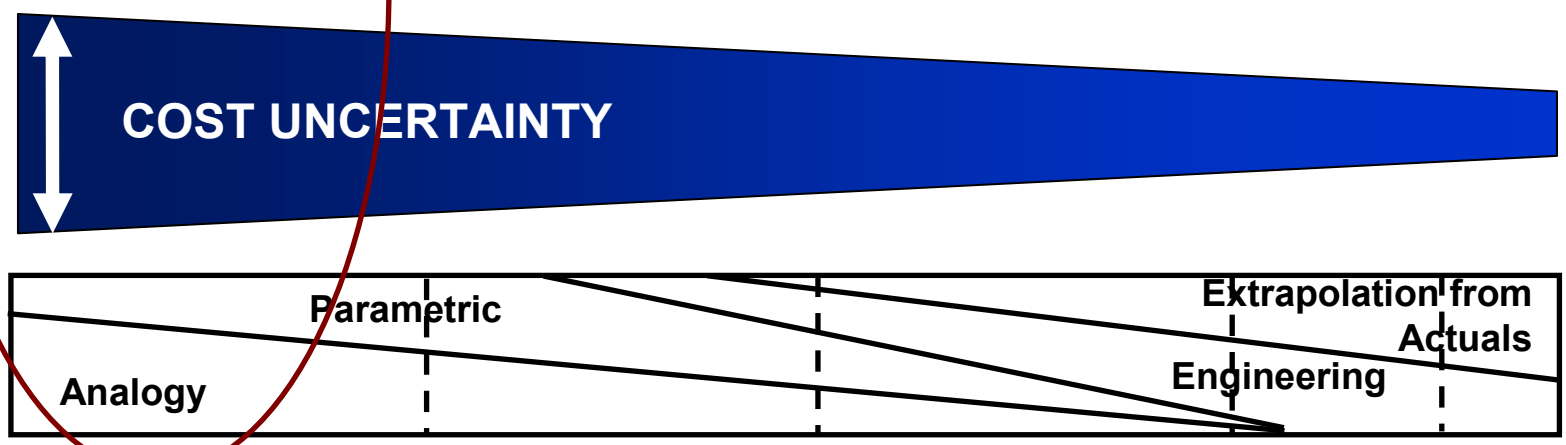
Answer: *E. All of the above*

Lesson 3 - Estimating & Communicating Risk

Background: Cost Estimating Uncertainty

PROGRAM TIMELINE

	CONCEPT	MS A	MS B	MS C	O&S
PERFORMANCE	COST/REQMTS TRADE-OFFS	OPERATIONAL REQUIREMENTS	FIRM SPECS		
DESIGN	MULTIPLE CONCEPTS	2 OR 3 CONCEPTS	SINGLE CONCEPTS SUBSYSTEM TRADES	LEVEL III DWGS	
PROGRAMMATICS	ALTERNATIVES	SMALL FAMILY OF ALTERNATIVES	STABLE ACQUISITION PLAN		
FINANCIAL	8-10 YEARS TO PRODUCTION	6-8 YEARS TO PRODUCTION	4-5 YEAR TO PRODUCTION		

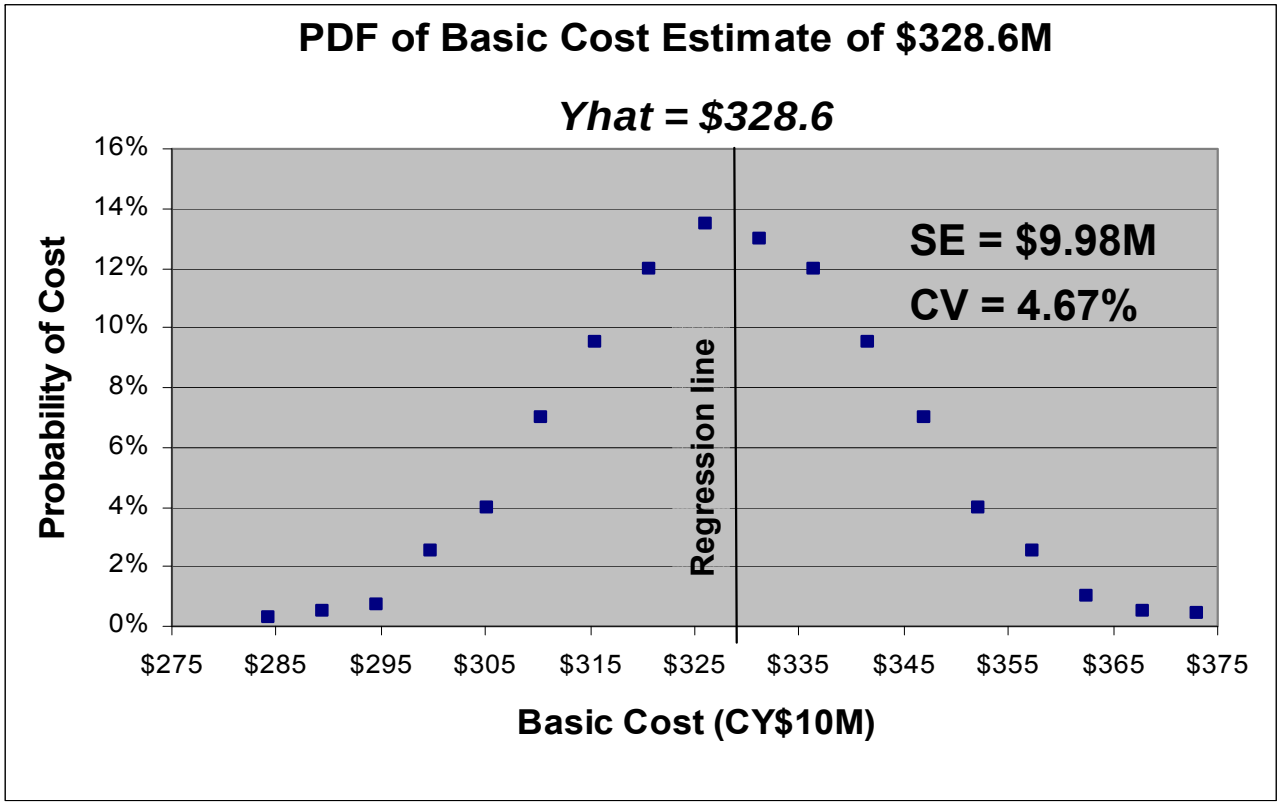


We are in the most uncertain phase of the program!

Lesson 3 - Estimating & Describing Cost Risk

Prediction Interval about the Cost Estimate (Forecast Error)

$$\text{Basic\$_T1} = 0.355\ \text{Draft}^{1.64}\ \text{Speed_Max}^{1.139} * 1.35^{\text{Combat}}$$



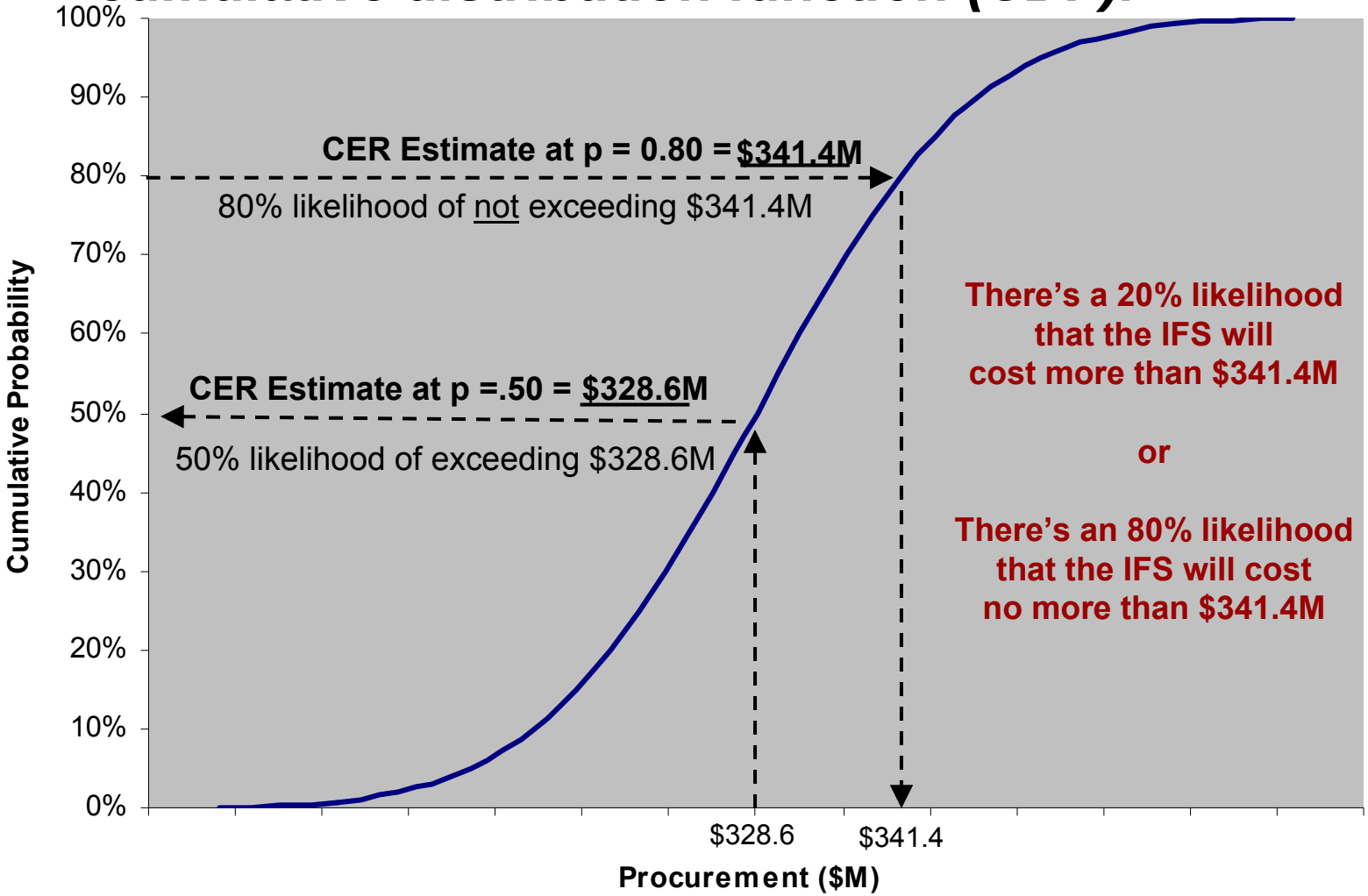
This is a "section" view of the prediction interval for Draft = 5.8 m & Speed = 32 knots

Would you consider the dispersion about the estimate to be significant? Why or Why not?

Lesson 3 - Estimating & Describing Cost Risk

Prediction Interval about the Cost Estimate (Forecast Error)

Two cost estimates are depicted along the following cumulative distribution function (CDF):



Notional Case Study

- Lesson 1 – Definition & Planning
  - Scenario for the Cost Estimator (the task & status)
  - Identify experts; meet w/them to elicit information
  - Create requirements sheet
- Lesson 2 – Developing the Cost Estimate
  - Data Collection
  - Statistics-based Cost Estimate
  - Regression-based Cost Estimate
- **Lesson 3 – Estimating & Describing Cost Risk**
  - **Prediction Interval about the Cost Estimate**
  - ⇨ **Changes in Cost due to Requirements Uncertainty**
    - **Method 1: Sensitivity Analysis**
    - **Method 2: Monte Carlo Simulation**

Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty (Elicitation)

- In an unlikely scenario, could Draft and Max Speed go up?
  - Answer: Yes. (Then expert provides possible reasons ...)
- How much could Draft and Max Speed go up?
  - Answer: Draft could go up, but no more than the BUTS Concept
  - Answer: There's a slight chance we could get squeeze another 0.5 knot
- Is there some chance these could go even higher?
  - Answer: Hmmm, I doubt the Draft would exceed that of the BUTS
  - Answer: If we tweek stern design, Max Speed could get up to 33 knots
- In an unlikely scenario, could Draft and Max Speed go down?
  - Answer: Yes. (Then expert provides possible reasons ...)
- How much could Draft and Max Speed go down?
  - Answer: Draft could decrease; below 5.7m would require new hullform
  - Answer: If weight creeps up, Max Speed could decrease by 2 knots
- What's the chance these could go even lower?
  - Answer: I doubt Max Speed will go below 30 knots
  - Answer: But Draft can not decrease below 5.7m!

Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Sensitivity

The elicitation produced a range of possible draft and max speed values:

- Draft could be as low as 5.70m and as high as 6.50m
- Maximum speed could be as low as 28.5 knots and as high as 33 knots

Sensitivity (Data Table)

Max_Speed \ Draft	5.70	5.80	5.90	6.00	6.10	6.20	6.30	6.40	6.50
28.5	279.9	288.0	296.2	304.5	312.8	321.3	329.8	338.5	347.2
29.0	285.5	293.8	302.1	310.6	319.1	327.7	336.4	345.2	354.1
29.5	291.1	299.5	308.1	316.7	325.4	334.2	343.1	352.0	361.1
30.0	296.8	305.3	314.0	322.8	331.7	340.6	349.7	358.8	368.1
30.5	302.4	311.1	320.0	328.9	338.0	347.1	356.3	365.6	375.1
31.0	308.0	317.0	326.0	335.1	344.3	353.6	363.0	372.5	382.1
31.5	313.7	322.8	332.0	341.2	350.6	360.1	369.7	379.3	389.1
32.0	319.4	328.6	338.0	347.4	357.0	366.6	376.3	386.2	396.1
32.5	325.1	334.5	344.0	353.6	363.3	373.1	383.1	393.1	403.2
33.0	330.8	340.3	350.0	359.8	369.7	379.7	389.8	400.0	410.3

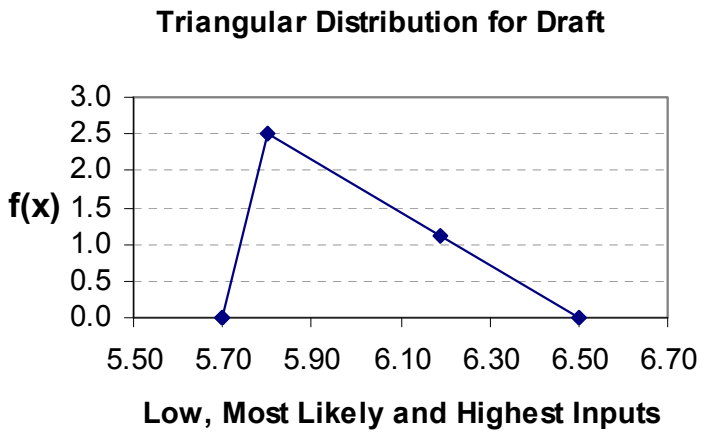
What are pros & cons of presenting costs of select pairs of draft & speed?

The elicited uncertainties could, <u>instead</u>, be used to create distributions that are "input parameters" to our CER ...

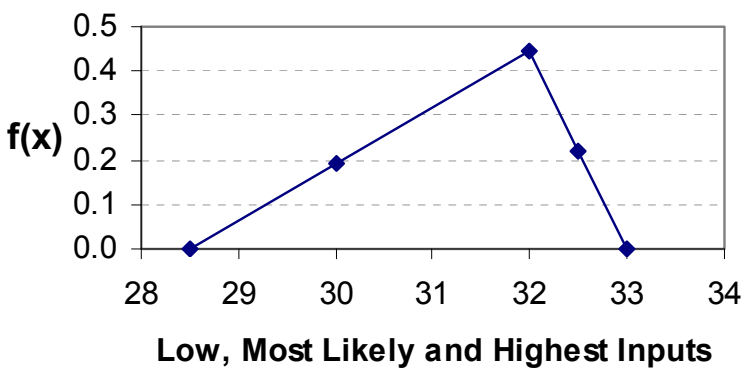
Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Input Parameters

Estimate:	Draft	f(x)	Notes:
Lowest	5.70	*0.00*	Expert-provided (<5.7 = new hull form)
Low	5.70	*0.00*	Expert-provided (No less than 5.7m)
Most Likely	5.80	*2.50*	Req't from Program Office
High	6.19	*1.11*	Expert-provided (No more than BUTS)
Highest	6.50	*0.00*	5% adjustment for Expert Bias



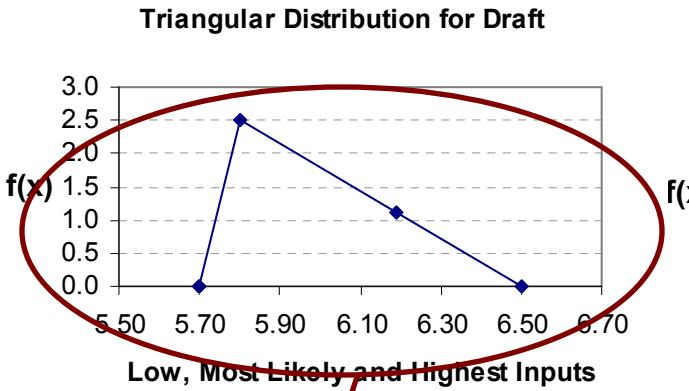
Estimate:	Speed_ Max	f(x)	Notes:
Lowest	28.5	*0.00*	5% adjustment for Expert Bias
Low	30.0	*0.19*	Expert-provided
Most Likely	32.0	*0.45*	From Program Office
High	32.5	*0.22*	Expert-provided
Highest	33.0	*0.00*	Expert-provided (if hull form adjusted)

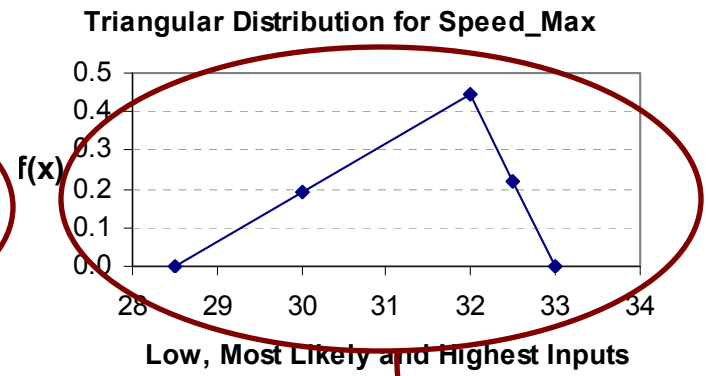


There are many other ways to characterize cost-driver uncertainties: some simpler, some more complex

Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Monte Carlo



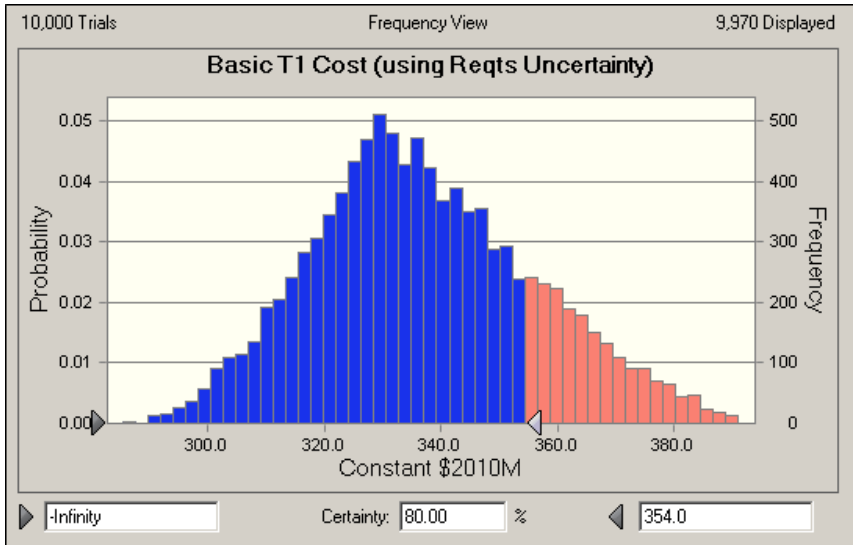


0

Correlation = -0.12

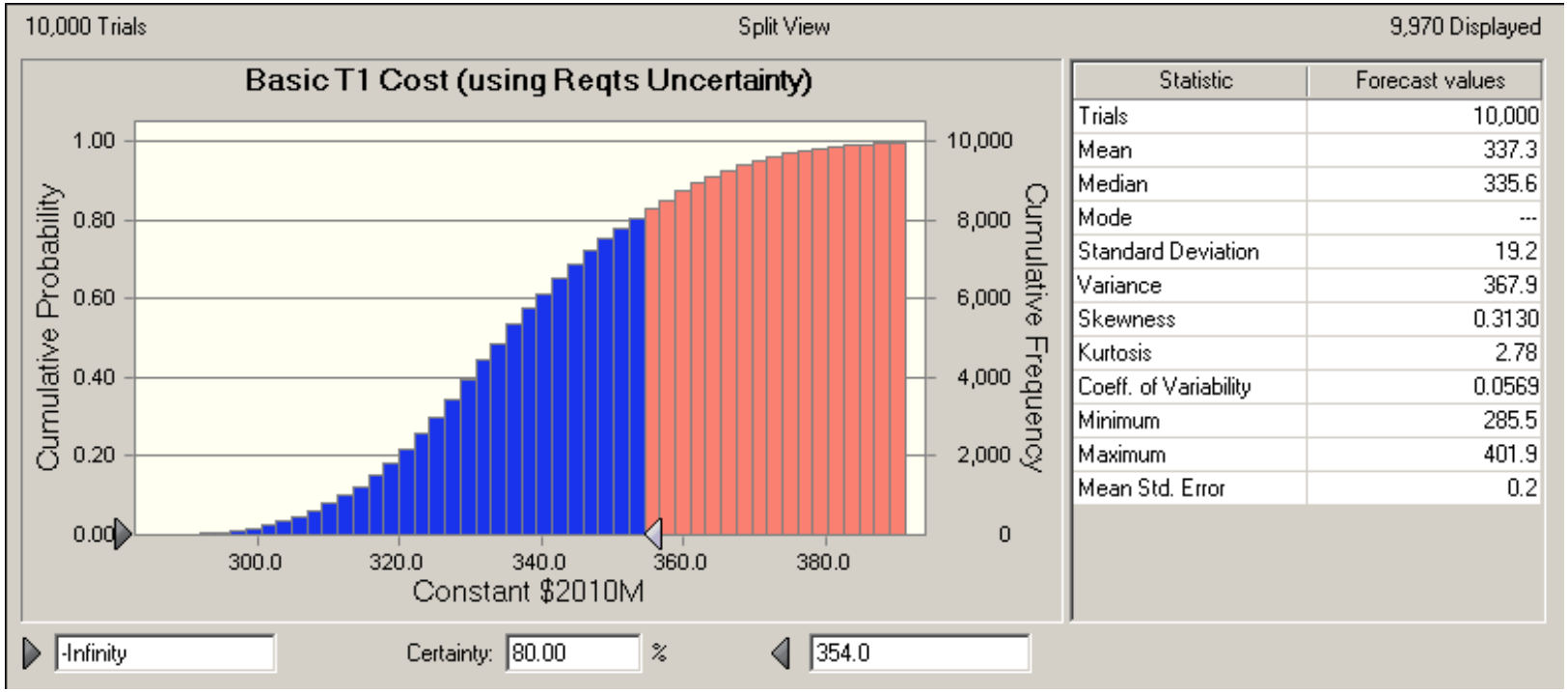
$$\text{Basic\$_T1} = 0.355\ \text{Draft}^{1.64}\ \text{Speed_Max}^{1.139} * 1.35^{\text{Combat}}$$

Monte Carlo Simulation



Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Monte Carlo



Statistic	Forecast values
Trials	10,000
Mean	337.3
Median	335.6
Mode	---
Standard Deviation	19.2
Variance	367.9
Skewness	0.3130
Kurtosis	2.78
Coeff. of Variability	0.0569
Minimum	285.5
Maximum	401.9
Mean Std. Error	0.2

- **The mean estimate went from $328.6M to $337.3M**

  - Increase driven by right-skewness of input parameters

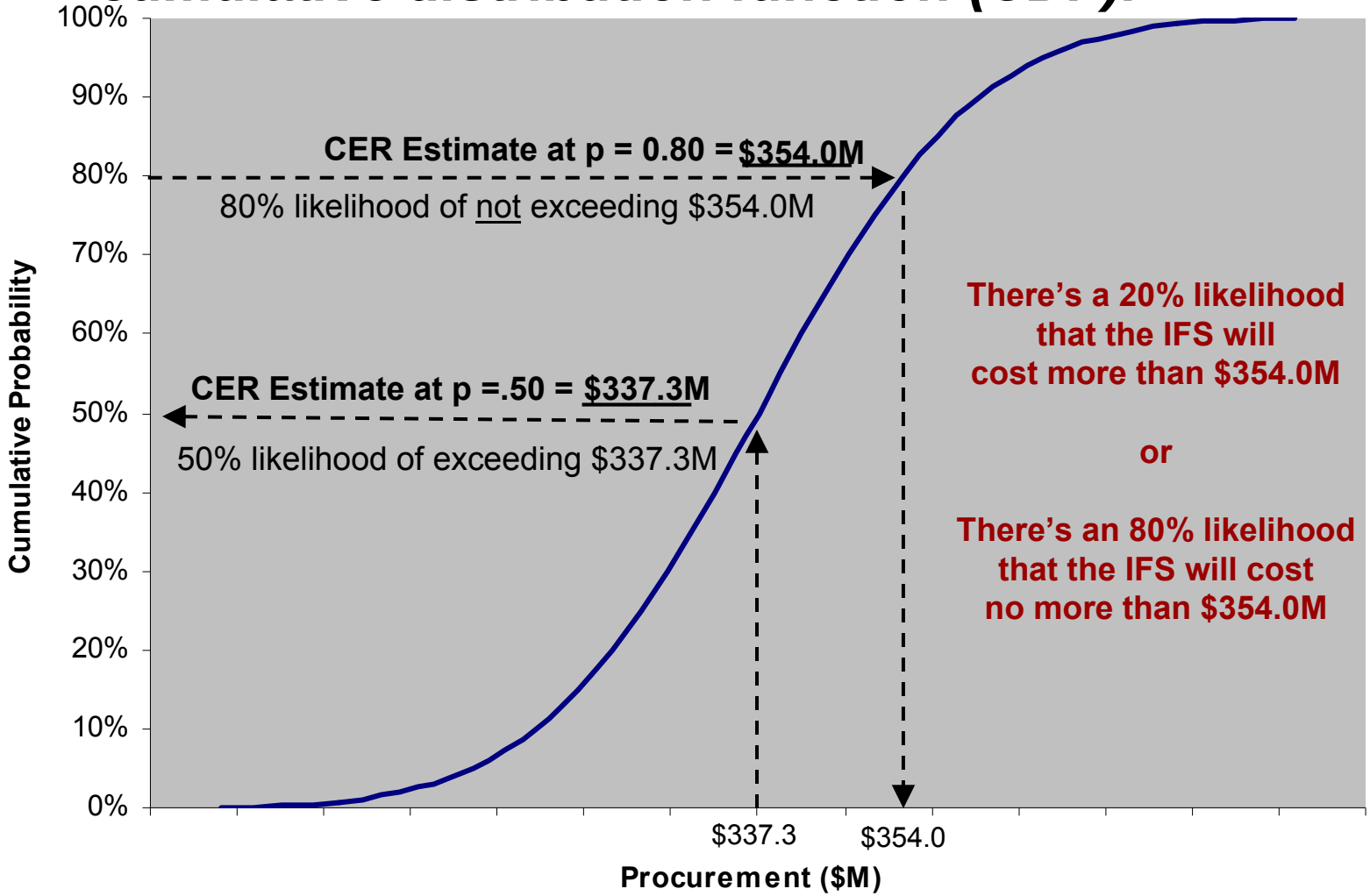
- **The estimate at p = 0.8 went from $329.8M to $354.0M**

  - Increase driven by increase in dispersion about the mean

Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Monte Carlo

Two cost estimates are depicted along the following cumulative distribution function (CDF):



Lesson 3 - Estimating & Describing Cost Risk

Changes in cost due to requirements uncertainty: Monte Carlo

Methodology	Mean (Cost at p = 0.50)	Risk-Adjusted (Cost at p = 0.80)	Notes
Sample Data Statistics	*$213.8M*	*$339.0M*	*Based on Average and Standard Deviation of Basic Cost of 7 ships*
Parametric Analysis (CER)	*$328.6M*	*$341.4M*	*Draft=5.8m, Max Speed=32 knots Extremely low standard error (SE)*
CER Sensitivity Analysis of Draft and Max Speed	*$337.3M*	*$354.0M*	*Draft Uncertainty: 5.70 - 6.50m Speed Uncertainty: 28.5 - 33 knots*

We have 6 estimates, based upon 3 methodologies. Which, if any, do you provide to the Wyatt Zomuchinkos?

Question 4: (Topic: Providing Best Estimate)

PM Wyatt Zomuchinkos wants to go ahead & present the CER cost estimate of $328.6M for his MS A review. Based on the analysis, what is your best response?

A. I was originally going to recommend the average of the data ($213.8M) but, because the data are so disperse, I recommend you use the risk-adjusted estimate of the data = $339M.

B. Because there's little difference between the average and my CER risk-adjusted estimate ($329.8M), a cost overrun should not cause any problems. So, the $328.6M should be fine.

C. The CER cost estimate of $328.6M does not account for requirements uncertainty; I recommend using the "requirements" risk-adjusted estimate of $354.0M

D. The program office plans to address all requirements risks. Budget $341.4M, the value that accounts only for the CER-risk.

Answer: *C*

Final Thoughts and Wrap-up

So do you <u>have</u> to go with $354.0 (CY$10M). Why or why not?

Not necessarily ...

Being the sole estimator, you need to have an estimate that you not only believe is the most realistic, but also an estimate that you can defend to decision-makers

Is our work done with the Threshold IFS estimate? Why or why not?

No. The PM asked for Budget Format = End Cost.

Could derive End Cost factors from the analogous ANDS P-5 Exhibit, then apply these factors to our recommended IFS Basic Cost:

Intra-theater Fast Ship (IFS) Costs x 000

ELEMENT OF COST	Qty	CY 2010 TOT COST	% of Basic based on American National Defense Ship (ANDS)
PLAN COSTS	1	19,301	5.45%
BASIC CONST/CONVERSION		**354,000**	
CHANGE ORDERS		31,003	8.76%
ELECTRONICS		21,687	6.13%
HULL,MECH,ELECT		2,522	0.71%
OTHER COST		17,108	4.83%
ORDNANCE		9,212	2.60%
NET P-1 LINE ITEM:		**454,833**	

Final Step: Convert $454.8M to FY12 Then-Year = $488.2M

Resources: Policies and Guidelines

- DoDD 5000.1, DoDI 5000.2 “Defense Acquisition System,” 23 October 2000 (incl Change 1, 4 Jan 2001)
- DoD 5000.02, “Operation of the Defense Acquisition System,” 8 December 2008. http://www.dtic.mil/whs/directives/corres/pdf/500002p.pdf
- DoD 5000.2-R “Mandatory Procedures for MDAPs and MAIS Acquisition Programs,” 11 Jun 2001
  - http://www.acq.osd.mil/ar/library.htm#directives
- DoDD 5000.4 “OSD Cost Analysis Improvement Group”
- DoD 5000.4-M “Cost Analysis Guidance and Procedures,” 11 Dec 92
  - http://www.dtic.mil/whs/directives/corres/html/50004m.htm
- SECNAV MEMO (02 May 08), SECNAV INST (16 Dec 08)

Further Study: DAU Courses

- BCF 106 – Fundamentals of Cost Analysis (online)
- BCF 107 – Applied Cost Analysis (classroom)
- BCF 204 – Intermediate Cost Analysis (classroom)
- BCF 206 – Cost Risk Analysis (classroom)
- BCF 208 - Software Cost Estimating (classroom)
- BCF 215 – Operations & Support Cost Analysis (classroom)
- BCF 302 – Advanced Cost Estimating (classroom)
- CLB 007 – Cost Analysis (online)
- CLB 012 – Cost As an Independent Variable (online)
- CLB 023 – Software Cost Estimating (online)
- CLB 024 – Cost Risk Analysis (online)
- CLM 016 – Cost Estimating (online)
- CLM 013 – Work Breakdown Structure (online)
- CLB 029 - Rates (Coming in FY10) (online)
- CLB 026 - Forecasting Techniques (online)
- CLB 030 - Cost Data Sources (Coming in FY10) (online)

Case Study Discussion and Questions (14:00 – 14:15 hrs)

Backup Slides

The Cost Analysis Requirements Description (CARD)

- Cost Analysis Requirements Description (CARD)
  - Produced by the Program Office
  - Draft CARD required 180 days before the decision; Final CARD required at least 60 days before the decision
  - Critical to definition of the program and therefore the cost analysis
  - Quality of the CARD directly impacts the quality of the cost estimate
- The CARD:
  - Collects, integrates, and describes the technical, programmatic, and schedule information necessary to estimate the costs of a program.
  - Includes manpower requirements, operational concept, logistics support concept, acquisition strategy and system test & evaluation plan
  - Is a living document. As more is known about the program, and as actual costs emerge, the CARD is updated

The CARD ensures cost estimates are based upon a common definition of the system & its acquisition strategy

Lesson 2 - Developing the Cost Estimate

20 Ship Classes (or Flights) Considered

	Class or Flight	Class Name	Hull Type
Combatants & Frigates	DDG 2	Charles F Adams	Destroyer
	DD 963	Spruance	Guided Missile Destroyer
	FFG 7	Oliver Hazard Perry	Guided Missile Frigate
	SAAR	Lahav	Corvette
Patrol-like ships & Icebreakers	WLM 685	Red Wood	Coastal Buoy Tender
	PC 14	Cyclone	Coastal Patrol
	PHM 1	Pegasus	Guided Missle Patrol Boat
	PG / PGM 84	Ashville	Patrol Gunboat
	WPB 110	Island	Patrol Boat
	WMEC 901	Bear	Medium Endurance Cutter
	WAGB 10	Polar Icebreaker	Icebreaker
	WLBB	Mackinaw	Icebreaker
	FRC	Sentinal	Fast Response Cutter
	WMSL 750	National Security Cutter	National Security Cutter
Mine-related, Ocean Research & Concepts	MCM 1	Avenger	Mine Countermeasure
	MHC 51	Osprey	Coastal Minehunter
	TAGOS 23	Victorious	Ocean Surveillance
	TAGS 60	Pathfinder	Ocean Surveillance
	SWATH	6000 ton SWATH Concept	SWATH
	3KSES	3000 ton SES Concept	Surface Effect Ship

Lesson 2 - Developing the Cost Estimate

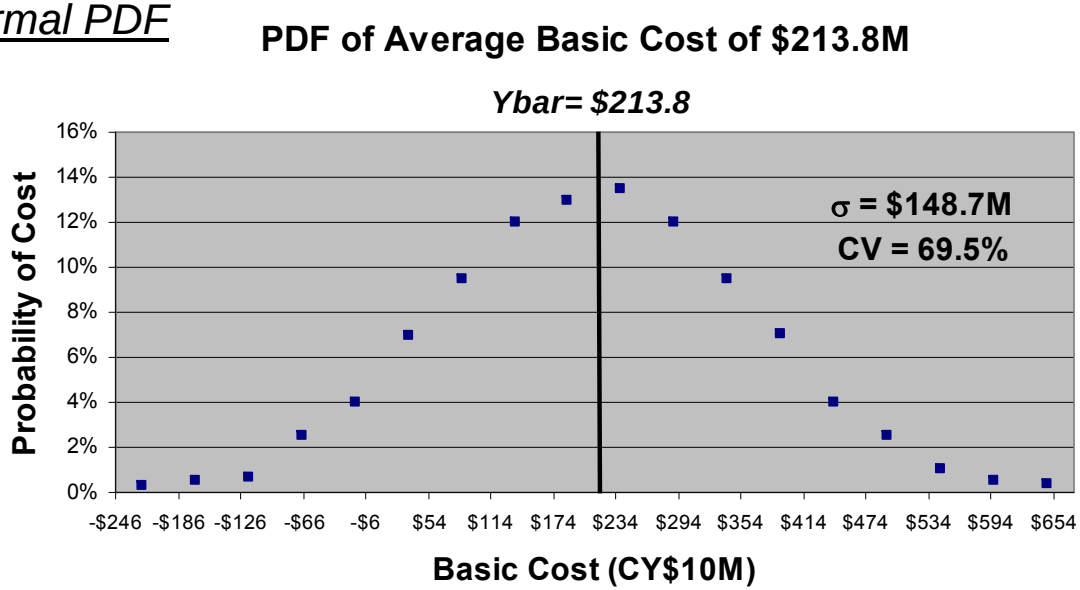
10 Ship Classes (or Flights) No Longer Considered

	Class or Flight	Class Name	Hull Type	Exclude from Data Set?
Combatants & Frigates	DDG 2	Charles F Adams	Destroyer	x
	DD 963	Spruance	Guided Missile Destroyer	
	FFG 7	Oliver Hazard Perry	Guided Missile Frigate	
	SAAR	Lahav	Corvette	x
Patrol-like ships & Icebreakers	WLM 685	Red Wood	Coastal Buoy Tender	x
	PC 14	Cyclone	Coastal Patrol	
	PHM 1	Pegasus	Guided Missle Patrol Boat	x
	PG / PGM 84	Ashville	Patrol Gunboat	x
	WPB 110	Island	Patrol Boat	x
	WMEC 901	Bear	Medium Endurance Cutter	
	WAGB 10	Polar Icebreaker	Icebreaker	
	WLBB	Mackinaw	Icebreaker	
	FRC	Sentinal	Fast Response Cutter	
	WMSL 750	National Security Cutter	National Security Cutter	
Mine-related, Ocean Research & Concepts	MCM 1	Avenger	Mine Countermeasure	
	MHC 51	Osprey	Coastal Minehunter	
	TAGOS 23	Victorious	Ocean Surveillance	x
	TAGS 60	Pathfinder	Ocean Surveillance	x
	SWATH	6000 ton SWATH Concept	SWATH	x
	3KSES	3000 ton SES Concept	Surface Effect Ship	x

Lesson 2 - Developing the Cost Estimate

Basic Statistics: The Mean and Standard Deviation

Normal PDF

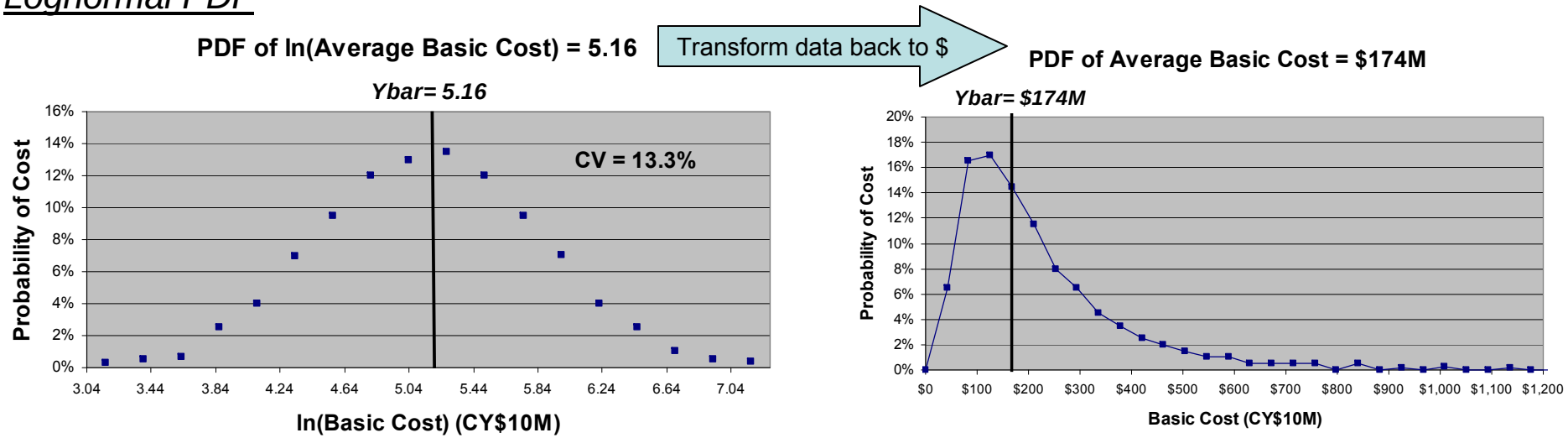


This is our PDF when we assume our ship cost data are normally distributed.

What might be tough to explain?

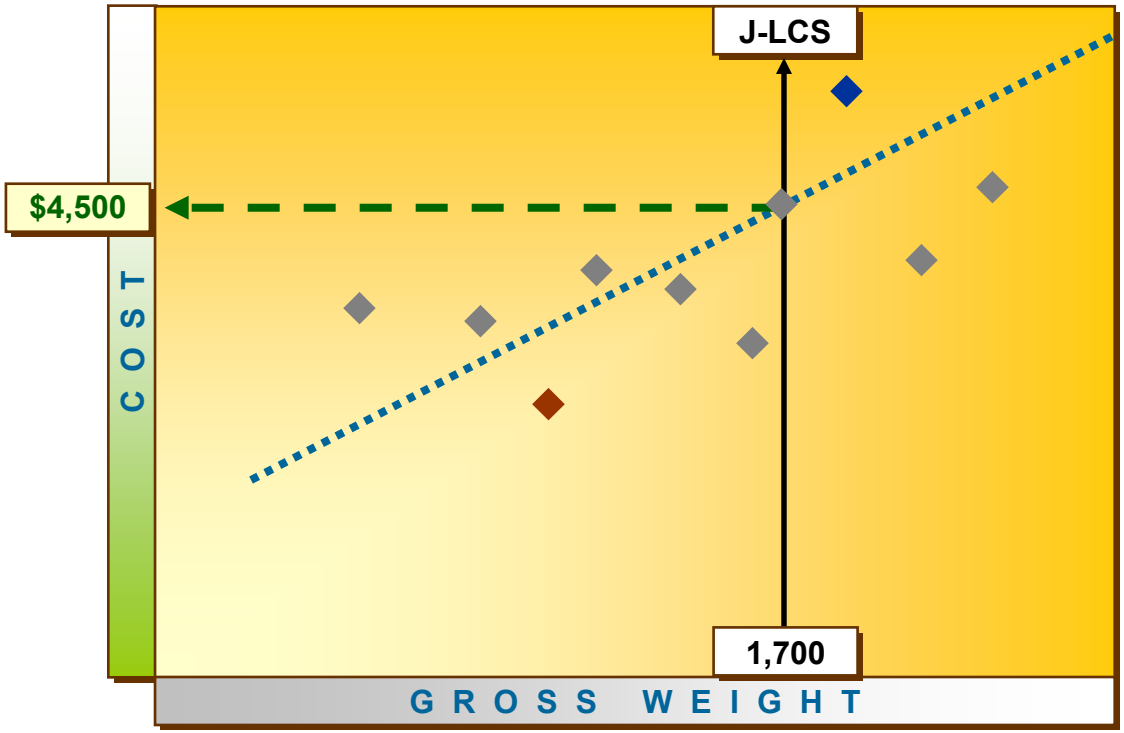
What could we do to avoid seeing such values on our PDF?

Lognormal PDF



Parametric Cost Estimating

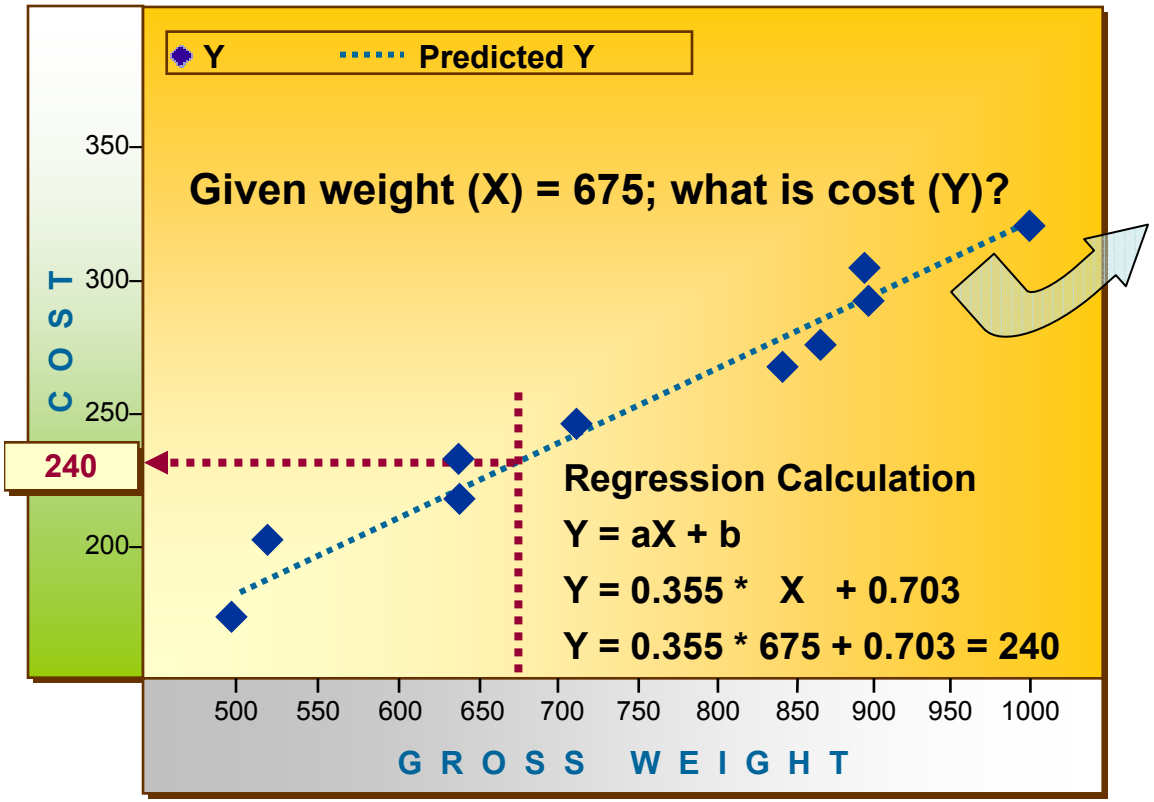
Example



Cost Estimating Relationship (CER): Can a weapon system's variation in cost (dependent parameter) be explained by variation in gross weight (independent parameter)?

Parametric Cost Estimating

Example: Estimate cost of 675 lb 'smart' missile



The *Method of Least Squares* was used to calculate the following parameters of the regression line:

slope (a) = 0.355

Y-intercept (b) = 0.703

The *coefficient of determination* (r^2), aka 'goodness of fit' = 0.96
That is, 96% of the variation in cost can be explained by the variation in gross weight

The regression calculation produced a CER of Y = 0.355 * X + 0.703

To apply the CER, we substitute 'X' with the weight of the proposed missile (X = 675). Thus, Y = 0.355 * <u>675</u> + 0.703 = $240 (in thousands)

Lesson 2 – Developing the Cost Estimate

- Collect all relevant information available for proposed system
  - Requirements, specifications and overall program information
- Brainstorm possible cost drivers of proposed system
- List out possible analogous systems
- Identify and document data sources for analogous systems
- Collate and normalize the data
  - Separate recurring and non-recurring data
  - Convert all data to Constant Year 2010 dollars
  - Adjust data to reflect Theoretical First (T1) Unit Cost in 2010 Dollars
- Quantify statistics of normalized data by getting measures of:
  - Central Tendency (Mean, Median and Mode)
  - Dispersion (Range, Standard Deviation and Coefficient of Variation)
  - Basic Construction Cost falling on the 80th percentile

Lesson 2 - Developing the Cost Estimate

Step 3: Select Variables that May Influence Cost

- We've collected 22 variables that can possibly "explain" the dependent variable, "T1 Basic Cost."
  - Note: 12 of these are dummy variables

Using $\frac{n!}{r!*(n-r)!}$ *, we can get the # of possible combinations of independent variables*

1 for Linear
1 for Log Linear

For a CER having 1 independent variable: ***22 x 2***

For a CER having 2 independent variables: $\frac{22!}{2!*20!} = \frac{22*21}{2} =$ ***231 x 2***

For a CER having 3 independent variables: $\frac{22!}{3!*19!} = \frac{22*21*20}{3*2} =$ ***1,400 x 2***

There are 3,306 possible linear & log linear combinations!
Is there a way to reduce the number of possible CERs?

Lesson 2 - Developing the Cost Estimate

Step 3: Select Variables that May Influence Cost

- We've have 17 variables that can possibly "explain" the dependent variable, "T1 Basic Cost."
  - Note: 12 of these are dummy variables

Using $\frac{n!}{r!*(n-r)!}$ *, we can get the # of possible combinations of independent variables*

1 for Linear
1 for Log Linear

For a CER having 1 independent variable: ***17 x 2***

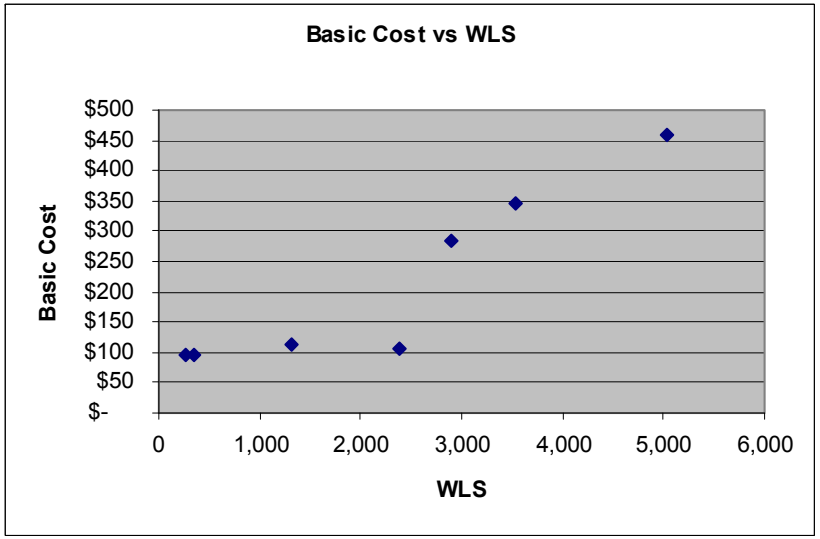
For a CER having 2 independent variables: $\frac{17!}{2!*15!} = \frac{17*16}{2} =$ ***136 x 2***

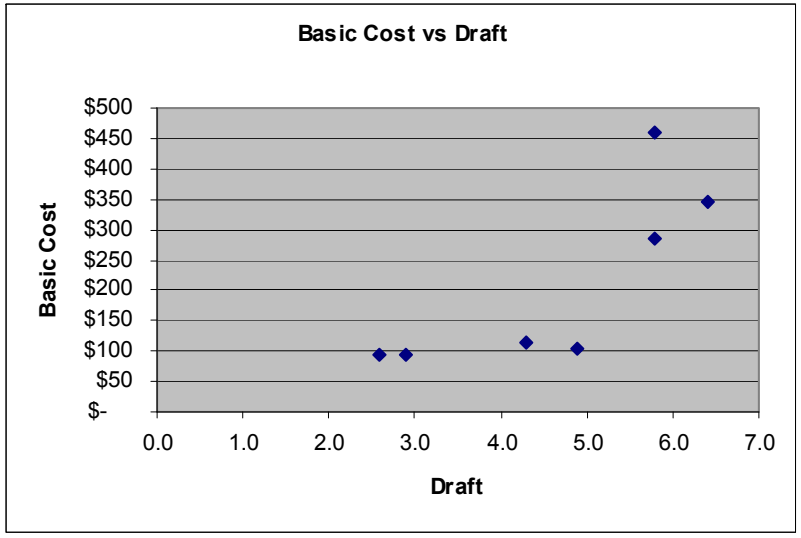
For a CER having 3 independent variables: $\frac{17!}{3!*14!} = \frac{17*16*15}{3*2} =$ ***680 x 2***

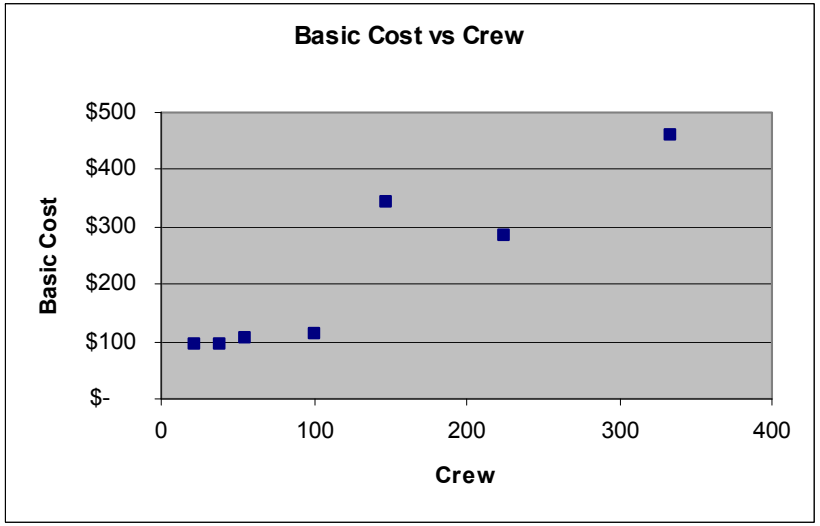
There are 1,666 possible linear & log linear combinations!
Is there a way to reduce the number of possible CERs?

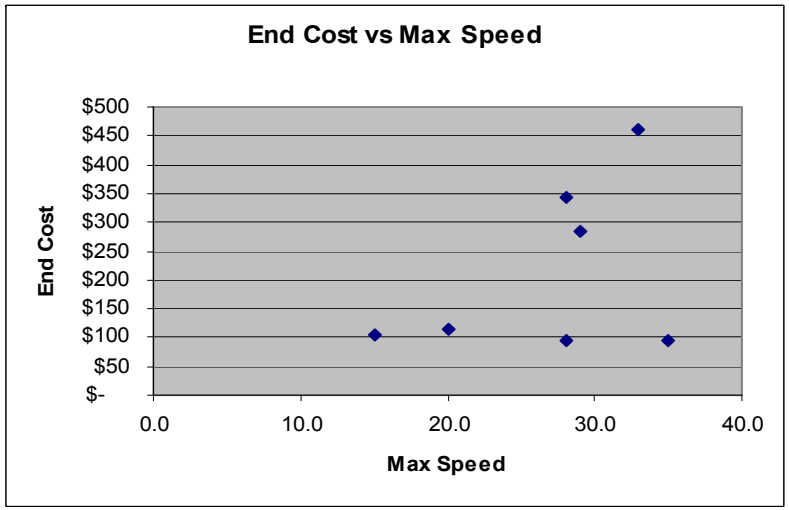
Lesson 2 - Developing the Cost Estimate

Step 4: "Hypothesize" most likely functional forms









Cost Model Diagnostics I

Logic - Does the relationship, or lack thereof, make logical sense? (Example: In MILCON, Cost and Size? Yes. In satellites, Cost and Size? Not necessarily.)

R - Coefficient of Correlation - indicates strength (-1 to 1) and direction of relationships between variables.

R^2 - Coefficient of Determination. R^2 measures the strength (0 to 1) of a relationship between variables, but does not indicate causality. (In simple linear regression, it's the proportion of SST accounted for by regression.)

R^2_a - Coefficient of Determination (adjusted). R^2_a permits comparisons of models with different numbers of variables by accounting for degrees of freedom in each case.

Cost Model Diagnostics II

SE - Standard Error. SE is the average estimating error (in terms of dollars, hours, or other units).

CV - Coefficient of Variation. CV measures in relative terms how far off the estimate will typically be. CV = SE divided by the mean of the dependent variable.

T-stat - T-stat is required for each independent variable (IV) to indicate whether it's a good predictor. In equations with only one IV, T-stat and F-stat will be equivalent.

F-stat - (a.k.a. F-ratio). F-stat indicates whether the equation overall is a good model. The F-test is needed only for equations with multiple IVs. (If the model passes the F-test, but not all T-stats are significant, you'll get a better model by removing the insignificant IVs)

Calculating Standard Error (SE) and Coefficient of Variation (CV)

Standard Error (SE) can be calculated in unit space

Allows direct comparison of linear and non-linear power models

Use: $$SE = \sqrt{\frac{\Sigma(Y - \hat{Y})^2}{n - k - 1}} = \sqrt{\frac{SSE}{\mathbf{n - k - 1}}} = \sqrt{MSE}$$

Once we have unit space SE, we can also get:

$$CV = \frac{SE}{\overline{Y}}$$

Lesson 2 - Developing the Cost Estimate

When is a High Correlation (R) Acceptable

- When you expect current relationship(s) between independent variables to continue on throughout the life of the system
- When multiple independent variables truly provide the best cost model (best explanatory power)

Defining & Dealing with Outliers

For outliers with respect to X:

- Calculate the mean and the standard deviation of the Xi values in the dataset
- Divide the difference between each Xi and X by the sx

$$\#\ \text{Std Deviations} = \frac{(x_i - \overline{x})}{s_x}$$

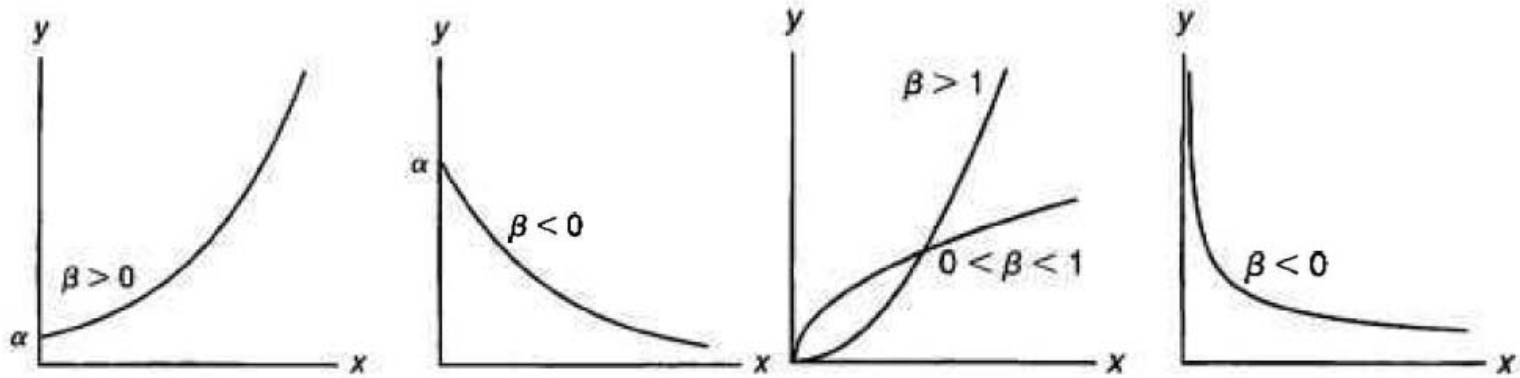
Identify observations that fall *more than 2 standard deviations* from the mean. (For small data sets, this is about 10% of your observations; for larger data sets, about 5%.)

If some observations are identified as potential outliers, it does NOT mean the there is necessarily a problem with the observations or the data set…need closer investigation

Backup

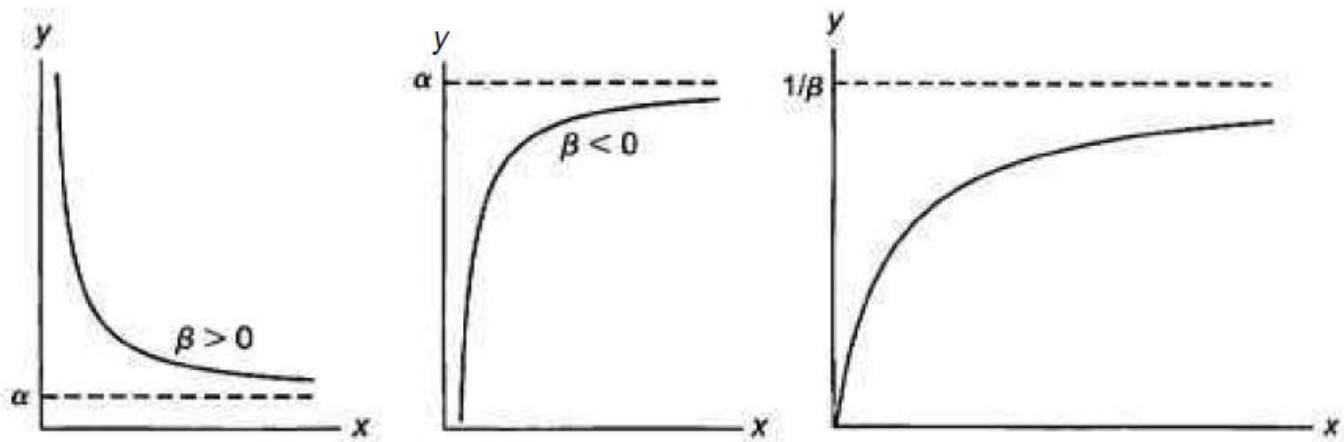
Table 11.6: Some Useful Transformations to Linearize

Functional Form Relating y to x	Proper Transformation	Form of Simple Linear Regression
Exponential: $y = \alpha e^{\beta x}$	$y^* = hxy$	Regress y^* against x
Power: $y = \alpha x^{\beta}$	$y^* = \log y;\quad x^* = \log x$	Regress y^* against x^*
Reciprocal: $y = \alpha + \beta\left(\frac{1}{x}\right)$	$x^* = \frac{1}{x}$	Regress y against x^*
Hyperbolic function: $y = \frac{x}{\alpha+\beta x}$	$y^* = \frac{1}{y};\quad x^* = \frac{1}{x}$	Regress y^* against x^*



(a) Exponential function

(b) Power function



(c) Reciprocal function

(d) Hyperbolic function

Lesson 2 - Developing the Cost Estimate

Steps 6 & 7: Select the Best CER & Check its logic

Model Type:	*Linear*	*Linear*	*Power*	*Power*
Equation	T1 Basic $ per WFL ton = $-48.2 + 4.17\ X_1 + 728\ X_2 + 38.8\ D$	T1 Basic $ per WFL ton = $20.36 + 2349\ X_1 + 30.2\ D$	T1 Basic $ per WFL ton = $3377\ X_1^{-0.818}\ X_2^{0.615}\ 2.11^{D}$	T1 Basic $ per WFL ton = $40.44\ X_1^{-0.5197}\ X_2^{1.404}$
Logical				
adj-R^2 (Unit Space)	99.08%	98.96%	99.52%	91.76%
SE (Unit Space)	$9.29	$9.84	$6.67	$27.75
CV	8.01%	8.49%	5.76%	23.94%
degrees of freedom (df)				
Complexity	3	2	4	3
Range	X_1 = 5.4 ---> 53.5 (shp/mt) X_2 = 0.023 ---> 0.11 (ppl/mt)	X_1 = 0.006 ---> 0.103 (knots per metric ton)	X_1 = 273 ---> 5034 (mt) X_2 = 15 ---> 35 (knots)	X_1 = 273 ---> 5034 (mt) X_2 = 15 ---> 35 (knots)
Comments				
Legend	X_1 = SHP / WLS (shp per mt) X_2 = Crew / WLS (people per mt) D = Non-combatant (dummy)	X_1 = Max Speed / WLS (knots/mt) D = Gas Turbine (dummy)	X_1 = Lightship Weight (mtons) X_2 = Speed_Max (knots) D = Gas Turbine (dummy)	X_1 = Lightship Weight (mtons) X_2 = Speed_Max (knots)

Lesson 2 - Developing the Cost Estimate

Describing selected CER to Decision-makers

What if we selected the linear model, $Y = b_0 + b_1X$?

Basic$_T1 = - 63.5 + 35.7 Draft + 0.004 SHP - 59 * Icebreaker

If all other independent variables were held constant ...

A 1 meter increase in Draft = $35.7M increase in cost

A 1,000 shp increase in SHP = $4M increase in cost

If the ship is an icebreaker, the cost decreases by $59M

What are some pros and cons of this CER?

Lesson 2 - Developing the Cost Estimate

Describing selected CER to Decision-makers

What if we selected the exponential model, $Y=Ae^{Bx}$?

Basic$_T1 = (82.27 e^{.002993 CubicNo}) 0.6524 ^ Icebreaker

* Note: "CubicNo" is Cubic Number in thousands) = [(LOA x Beam x Draft) / 100] / 1,000

For $Y=Ae^{Bx}$, a unit change in CubicNo results in some fixed % change in Cost

- This fixed % change in Cost equals (e^B – 1)

If all other independent variables were held constant ...

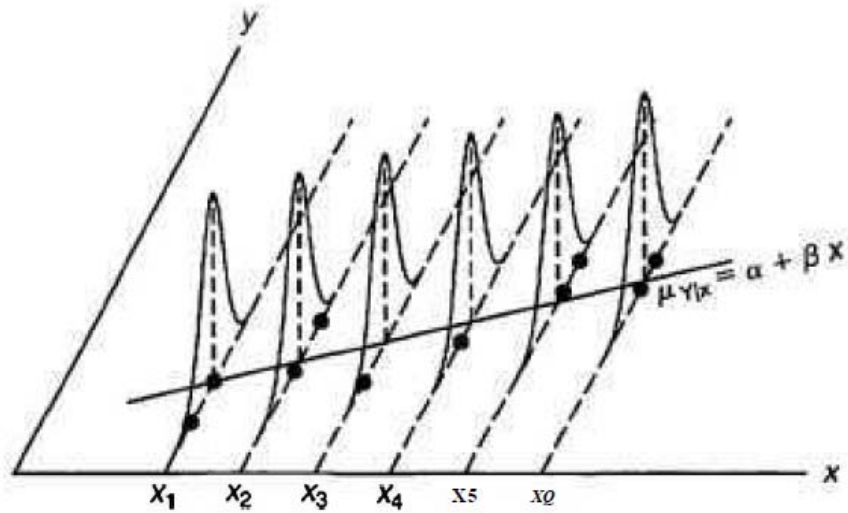
A unit increase in Cubic Number = 2.993% increase in cost

If the ship is an icebreaker, the cost decreases by 34.76%

What are some pros and cons of this CER?

Lesson 3 - Estimating & Describing Cost Risk

Prediction Interval about the Cost Estimate (Forecast Error)



- The line in the graph is that of the true regression line.
- Points plotted are actual (y,x) points, scattered above and below the line.
  - Where each point is on its own normal distribution
- Recall that the center of a normal distribution falls on the regression line.
  - As a result, the true regression line goes through the means of the response
- Actual observations are on the distribution around the means.
- The deviation between an individual y & the point on the line is the residual.

Lesson 3 - Estimating & Describing Cost Risk

Prediction Interval about the Cost Estimate (Forecast Error)

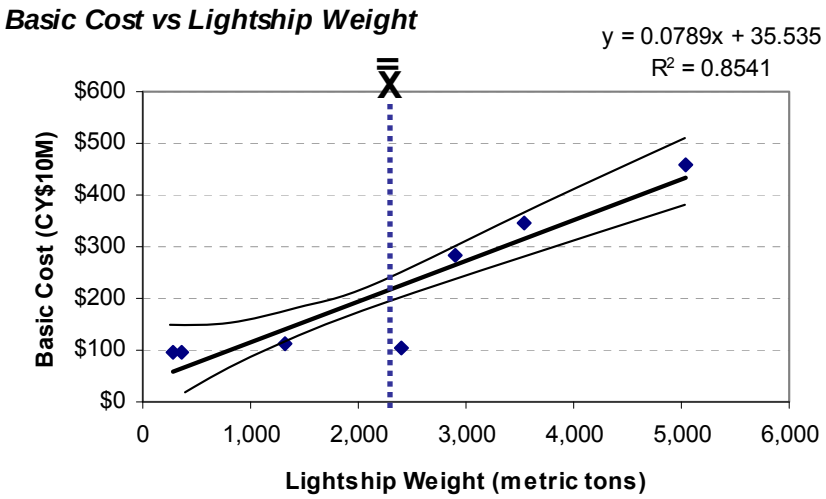
The Prediction Interval (PI) for any given value of x_i is estimated by:

$$PI = \hat{Y}_0 \pm (SE)\, t_p \sqrt{\frac{n+1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

where:

$\hat{Y}_0$ = Cost estimated using CER
SE = Standard Error of estimating equation
t_p = value obtained from table based on α and df
$\bar{\bar{x}}$ = mean value of all sample X's
x_i = value of sample X (used to develop CER)
x_0 = specified value of X used in the CER

Example regression line, with 80% PI:
Basic$_T1 = 0.0789 WLS + 35.535



The PI becomes wider ...

- For x_0 values that are farther from $\bar{\bar{x}}$
- As the level of confidence increases (e.g. 95% PI is wider than a 90% PI)
- As degrees of freedom decreases, especially for n < 30

What happens to PI as SE approaches 0?

Lesson 3 - Estimating & Describing Cost Risk

Converting Basic Estimate to End Cost

Historical End Cost of Most Analogous Ship (Costs x 000)

American National Defense Ship (ANDS)

ELEMENT OF COST	Qty	FY 2009 TOT COST	% of Basic
PLAN COSTS	1	22,000	5.45%
BASIC CONST/CONVERSION		403,500	
CHANGE ORDERS		35,338	8.76%
ELECTRONICS		24,720	6.13%
HULL,MECH,ELECT		2,875	0.71%
OTHER COST		19,500	4.83%
ORDNANCE		10,500	2.60%
NET P-1 LINE ITEM:		**518,433**	

Intra-theater Fast Ship (IFS)

ELEMENT OF COST	Qty	CY 2010 TOT COST	% of Basic based on American National Defense Ship (ANDS)
PLAN COSTS	1	19,301	5.45%
BASIC CONST/CONVERSION		**354,000**	
CHANGE ORDERS		31,003	8.76%
ELECTRONICS		21,687	6.13%
HULL,MECH,ELECT		2,522	0.71%
OTHER COST		17,108	4.83%
ORDNANCE		9,212	2.60%
NET P-1 LINE ITEM:		**454,833**	

And yet still ... there is one final step ...

Lesson 3 - Estimating & Describing Cost Risk

Converting Basic Estimate to End Cost

Intra-theater Fast Ship (IFS)

ELEMENT OF COST	Qty	CY 2010 TOT COST	% of Basic based on American National Defense Ship (ANDS)
PLAN COSTS	1	19,301	5.45%
BASIC CONST/CONVERSION		**354,000**	
CHANGE ORDERS		31,003	8.76%
ELECTRONICS		21,687	6.13%
HULL,MECH,ELECT		2,522	0.71%
OTHER COST		17,108	4.83%
ORDNANCE		9,212	2.60%
NET P-1 LINE ITEM:		**454,833**	

Intra-theater Fast Ship (IFS) converted to Then-Year Dollars for P-1 Budget

For P-1 Budget, converted CY$10 to TY$12 using SCN = Shipbuilding and Conversion, Navy (1611)
per NCCA Inflation calculator using Feb 2009 OSD Guidance CY$10 to TY$12 = 1.0733

Intra-theater Fast Ship (IFS)

APPROPRIATION: SHIBUILDING AND CONVERSION **WEAPON SYSTEM COST ANALYSIS EXHIBIT (P-5)**

ELEMENT OF COST	Qty	FY 2012 TOT COST
PLAN COSTS	1	20,715
BASIC CONST/CONVERSION		379,941
CHANGE ORDERS		33,275
ELECTRONICS		23,277
HULL,MECH,ELECT		2,707
OTHER COST		18,361
ORDNANCE		9,887
NET P-1 LINE ITEM:		**488,163**