"Creating Customer-Focused Success"

Correlations in Cost Risk Analysis

Raymond P. Covert
Technical Director
MCR, LLC
rcovert@mcri.com

Presented at 2006 Annual SCEA Conference
Tysons Corner, VA
June 13-16, 2006

Agenda

- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

Introduction

- **What is Correlation?** Ref. 1
  - **A statistical measure of association between two variables.**
  - **It measures how strongly the variables are related, or change, with each other.**
    - **If two variables tend to move up or down together, they are said to be positively correlated.**
    - **If they tend to move in opposite directions, they are said to be negatively correlated.**
  - **The most common statistic for measuring association is the Pearson (linear) correlation coefficient, ρ_P.**
  - **Another is the Spearman (rank) correlation coefficient, ρ_S, which is used in Crystal Ball and @Risk**

"Creating Customer-Focused Success"

- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

Pearson and Spearman Correlation

- **Pearson "Product-Moment" Correlation**
- **Measures Extent of LINEARITY of a relationship between two random variables**

$$\rho_P = \rho_{Pxy} = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right) * \left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2 * \left(y_i - \bar{y}\right)^2}}$$

$\rho_P = \rho_{Pxy}$, the observed Pearson correlation coefficient between two sets of variables, *x* and *y*

***i* = an index variable**

***N* = the number of variables in either set**

= the mean of data set x

= the mean of data set y

x_i = The ith element of data set x

y_i = The ith element of data set y

- **Spearman Rank Correlation**
- **Measures Extent of MONOTONICITY of a relationship between two random variables**

$$\rho_S = \rho_{Sxy} = \frac{\sum_{i=1}^{N}\left(R_i - \bar{R}\right) * \left(S_i - \bar{S}\right)}{\sqrt{\sum_{i=1}^{N}\left(R_i - \bar{R}\right)^2} * \sqrt{\sum_{i=1}^{N}\left(S_i - \bar{S}\right)^2}}$$

$\rho_S = \rho_{Sxy}$, the observed Spearman correlation coefficient between two sets of variables, *x* and *y*

***i* = an index variable**

***N* = the number of variables in either set**

= the mean of the ranks of the set of variables in set *x*

= the mean of the ranks of the set of variables in set *y*

R_i = the rank order of variable x_i in data set *x*

S_i = the rank order of variable y_i in data set *y*

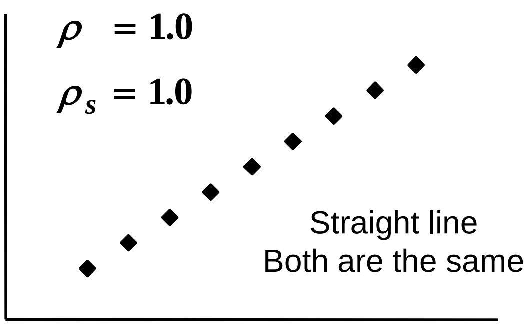
R = the mean of the ranks of the set of variables in data set *x*

S = the mean of the ranks of the set of variables in data set *y*

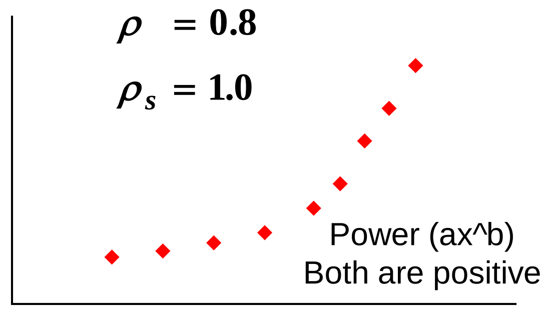
Statistics Theorem: Spearman Rank Correlation Coefficient Equals Pearson (Linear) Correlation Coefficient Calculated Between the Two Sets of Ranks

Product Moment (Linear) vs. Rank Correlation

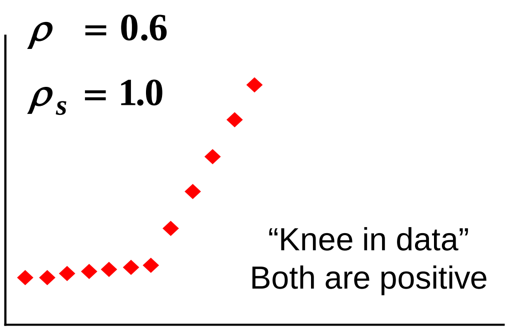
LINEAR



POWER

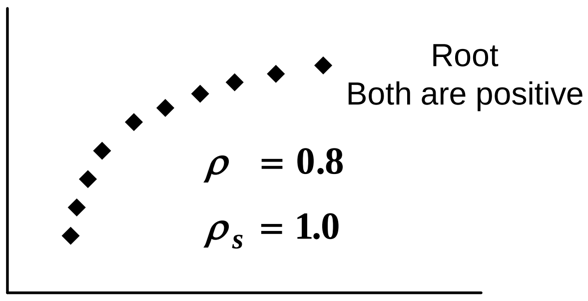


"KNEE"

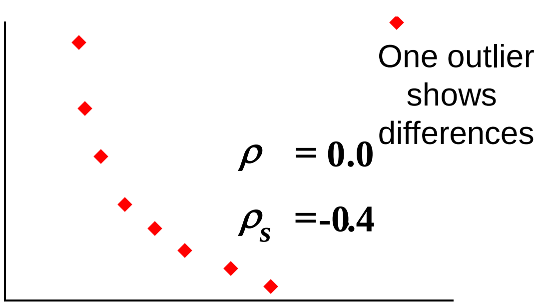


More Nonlinear

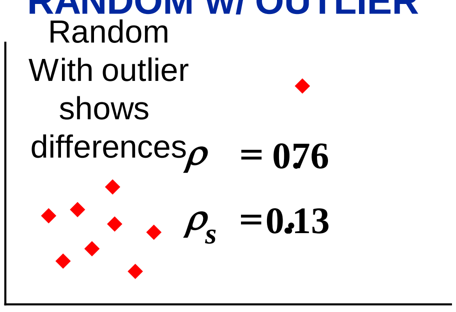
ROOT



DECAY w/ OUTLIER



RANDOM w/ OUTLIER



Linear Data gives similar ρ and ρ_s

- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

The Different Ways to Correlate Random Variables

- **Correlation is a statistic – so all correlation is statistical**

CORRELATION
All correlation is statistical

It can be a purely statistical artifact

It can be causal (a relationship exists)

PURELY STATISTICAL
No causal relationship exists

CAUSAL
Something causes A and B to co vary

We may be able to functionally describe the relationship

The equation for the relationship may not be known or modeled with an equation

CAUSAL STATISTIC
Relationship between A and B is statistically modeled using correlation coefficients

FUNCTIONAL
Relationship between A and B is modeled by an equation

Purely Statistical Correlation

- How accurate are your cost models?
- The percentage errors of the Aerospace Small Satellite Cost Model (SSCM1998) subsystem CERs have percentage errors between 30% and 40% (1σ)
- When we take the RSS (square root of sum of squares) of the model errors, the model reports a 1σ error of about 13%

$$\sigma_{Total}^2 = \sum_{k=1}^{n} \sigma_k^2 + \cancelto{0}{2\sum_{k=2}^{n} \sum_{j=1}^{k-1} \rho_{jk} \sigma_j \sigma_k}$$

- However, when we plug-in the actual database cost drivers into the model, SSCM estimated the database to 24% (That's how accurate the model is)
- The missing piece is the correlation between the errors, effectively 10%

$$\rho_{eff} = \frac{\sigma_{Total}^2 - \sum_{k=2}^{n} \sigma^2}{2\sum_{k=2}^{n} \sum_{j=1}^{k-1} \sigma_j \sigma_k}$$

Functional Correlation

- **Functional correlation exists:** **[Ref.2 - Coleman]**
  - **Between cost drivers (Power and Weight)**
  - **Between CERs and their cost drivers (Cost = a*Weight^b)**
  - **Between certain pairs of CER (SEITPM = a*PMP^b)**
  - **Between CERs using the same cost driver**

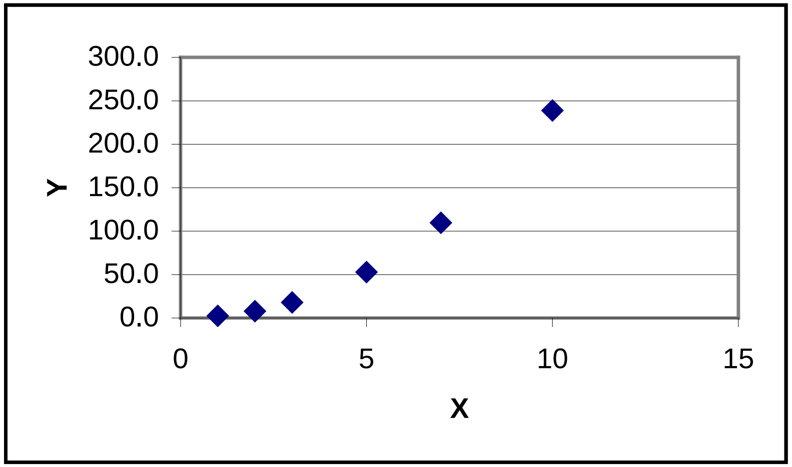
- **This happens when:**
  - **Random variables are transformed (scaled and/or distorted) by a function**
  - **Random variables are "reused"**

Functional Correlation: Transformations

- Y is correlated to X through a function (transformation)
  - X and Y are correlated through the transformation
  - Suppose $Y = a + b*x^c$: $Y = 1 + 1.5*X^{2.2}$
  - And if X varies from 1 to 10, then:
    - $\rho_{xy} = 0.964$

X	Y
1	2.5
2	7.9
3	17.8
5	52.7
7	109.5
10	238.7

ρ(Y1,Y2) 0.964



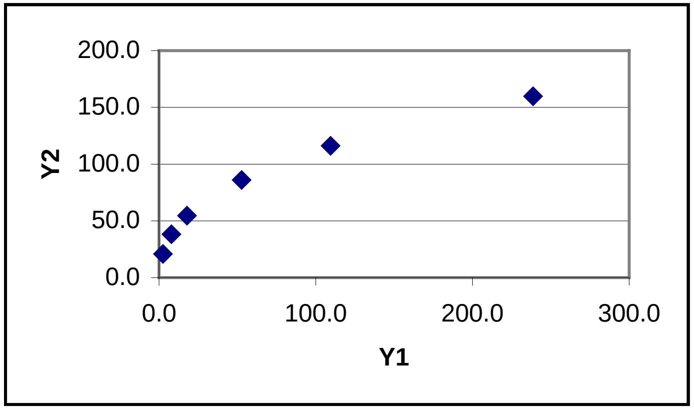
Functional Correlation: Reusing Random Variables

- Y_1 and Y_2 are correlated to each other by reusing a random variable

- Suppose $Y_1 = a_1 + b_1 * x^{c1}$ and $Y_2 = a_2 + b_2 * x^{c2}$
  - $Y1 = 1.0 + 1.5 * X^{2.2}$
  - $Y2 = 0.9 + 20 * X^{0.9}$
  - And if X Varies from 1 to 10, then:
  - $\rho_{xy} = 0.956$

X	Y1	Y2
1	2.5	20.9
2	7.9	38.2
3	17.8	54.7
5	52.7	86.0
7	109.5	116.1
10	238.7	159.8

r(Y1,Y2) 0.956



Causal Statistical Correlation

- **Correlation does not imply direct causation**
- **Consider this statement based on some statistical data:**
  - **"Shark attacks are correlated to ice cream sales".**

- **This <u>DOES NOT</u> mean that ice cream sales increase because of shark attacks (or vice versa).**

- **More shark attacks happen in warm weather, and more ice cream is consumed in warm weather, therefore ice cream sales and shark attacks are positively correlated.**
  - **So as temperatures increase, more people swim in the ocean and more people eat ice cream.**
  - **Something (the temperature) causes A (shark attacks) and B (ice cream sales) to co vary, but we do not know the exact equation.**

Correlation tells us the degree to which two variables covary, not why

Causal Correlation

- Causal Statistical correlation is correlation that is causal in nature without a functional relationship defined in our model
- What does this mean?
  - We know that two random variables covary, but we have not modeled the relationship with an equation
  - We use correlation coefficients to mimic their behavior
- An example
  - Spacecraft Weights and Powers
    - We know that as we increase the mass of a spacecraft component, inertia and weight increase
    - This will drive both structural rigidity requirements and attitude control torque requirements
    - Structure and ADACS weight may go up as component mass increases
- We can mimic this behavior by correlating the variance on the masses (but it is better to use equations!)

"Creating Customer-Focused Success"

Causal Correlation of Weight?

- Weight (as a cost driver) correlation statistics <u>should not be used</u> explicitly to correlate weights as cost drivers in probabilistic simulations.
- They do not tell you how your weights for your spacecraft will covary, only how those weights covary in the databases they came from.
- Weight Correlation from USCM 7 and Small Satellite Cost Model

…and they differ between models

USCM7	ADCSWT	AKMWT	COMMWT	EPSWT	STRCWT	THERWT	TT_CWT
ADCSWT	1.000	-0.128	0.772	0.869	0.441	0.193	0.447
AKMWT	-0.128	1.000	0.721	-0.011	0.438	0.000	-0.029
COMMWT	0.772	0.721	1.000	0.779	0.672	0.230	0.645
EPSWT	0.869	-0.011	0.779	1.000	0.552	0.374	0.222
STRCWT	0.441	0.438	0.672	0.552	1.000	0.157	0.443
THERWT	0.193	0.000	0.230	0.374	0.157	1.000	-0.119
TT_CWT	0.447	-0.029	0.645	0.222	0.443	-0.119	1.000

Avg Correl
0.595

Pretty close!

SSCM	ACSWT	PROPWT		EPSWT	STRWT		CDHWT
ACSWT	1.000	0.295		0.469	0.163		0.553
PROPWT	0.295	1.000		0.461	0.561		0.757
EPSWT	0.469	0.461		1.000	0.828		0.860
STRWT	0.163	0.561		0.828	1.000		0.903
CDHWT	0.553	0.757		0.860	0.903		1.000

Avg Correl
0.670

- Hu has derived weight correlations for USCM-8 **[Ref. 3- Hu]**

Finding Causal Statistical Correlation

- **What to do:**
  - **Ask an engineer to determine these functions and error correlations using a model of your system**
- **What not to do**
  - **Use Weight and power data from multiple satellites to determine these correlations** [Ref. 3- Hu]
  - **Why?**
    - **Because data from multiple missions, orbits, manufacturers and requirements will give poor, misleading results**

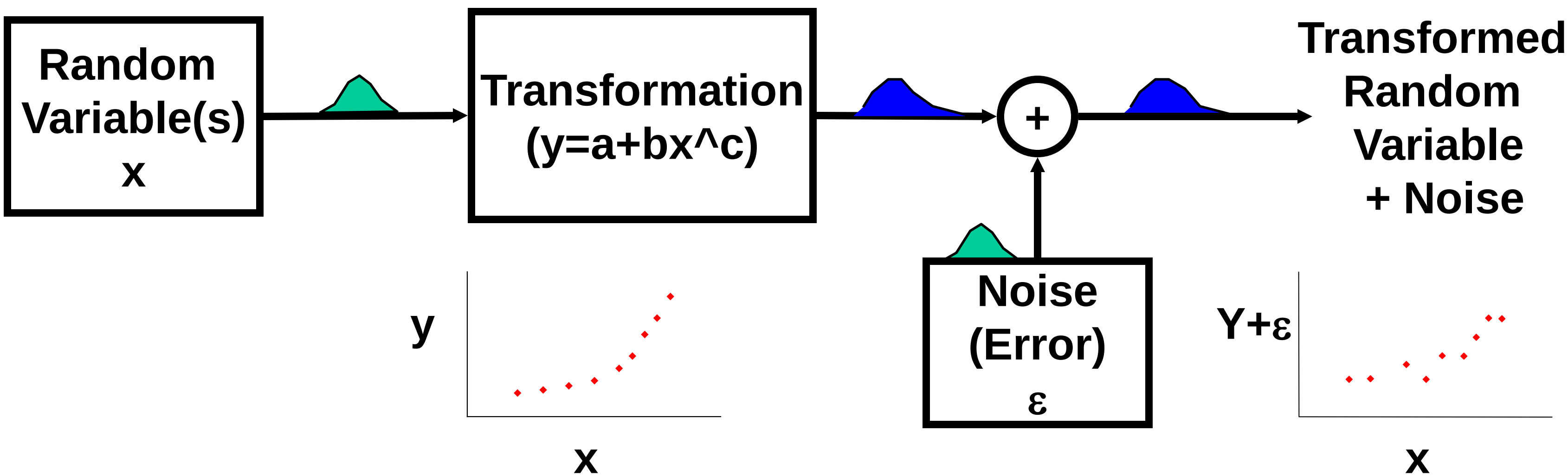
- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

Mix of Correlations

- Remember there are (at least) the following ways to correlate random variables:
  - Purely statistical (through correlation coefficients)
  - Causally:
    - Statistical basis (through correlation coefficients)
    - Functional basis (through equations)
- What if an estimate has all of these?
  - What type is most important?
  - Can you ignore some of them?
- The answer depends on the model…

Signal To Noise Ratio (SNR)

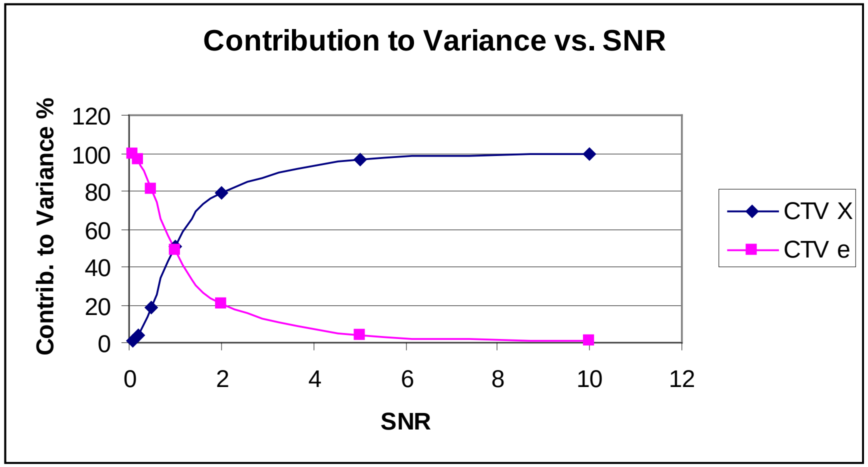
- The transformation function ($Y=a+bx^c$) correlates the output (Y) to the input (X)
  - If the (random) noise, ε, added to the transformed "signal", Y, is small, then Y+ε and X will be strongly correlated because ε has such a small contribution to the variance of Y+ε
  - If ε is large, the correlation magnitude will decrease



Impacts of Functional Correlation with Varying SNR

- We can model the ratio of the variances $\sigma^2_Y/\sigma^2_\varepsilon$ as our "Signal to Noise Ratio", SNR,
  - We will see the impacts on contribution to variance of X (the system input) and ε (the noise input)
- In our example, X and ε are identical symmetric distributions, so we will model the ratio of Y to ε as the SNR (in this case only)

In this example:
X = Normal Dist, μ=1, σ = 0.3
Y = 1 + 1* X ^ 2
E = Normal Dist, μ=k*1, σ = k*0.3
K = (0.2,0.4,1,2,4,10,20)



- This tells us that functional correlation dominates when the added noise magnitude is less than the signal's and vice versa.

Interpreting the SNR Demonstration

- **The SNR demonstration tells us that functional correlation only dominates when its contribution overpowers the noise contribution**
- **If there is functional correlation and "noise" in a cost estimate, we should see which one dominates and why**
- **Remember added noise <u>may be</u> the statistical sum of all of the non functionally correlated terms in our estimate (other errors).**
  - **If there are a large number of added noise terms, their correlation will be an important factor**
  - **Correlation of the noise will have a big impact if the number of terms is large**

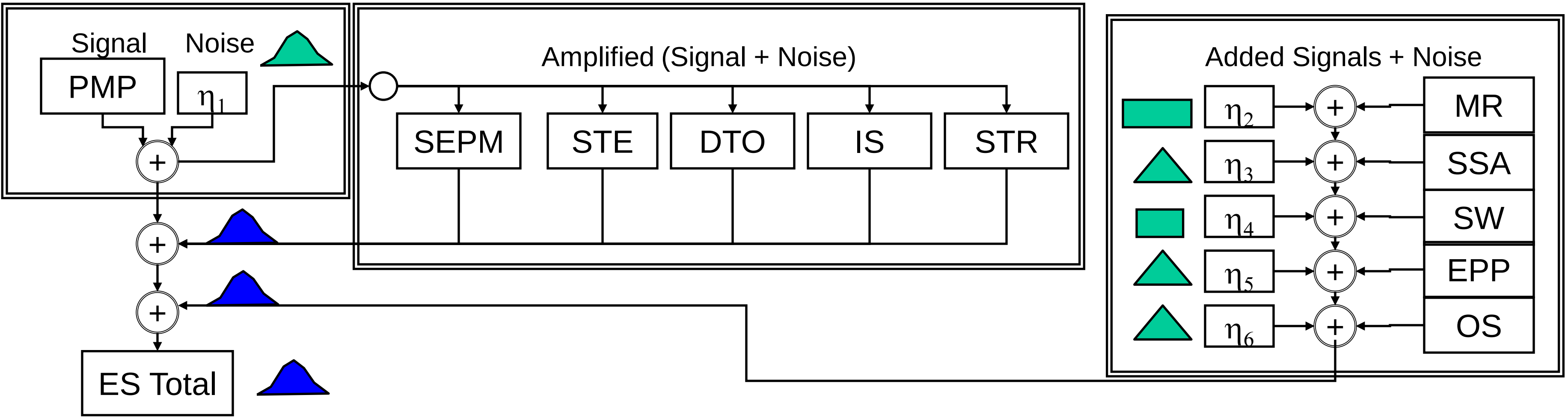
SCEA Training Example

- **The first example can be found in the SCEA Training Manual Case Studies Page CE V – 80**
- **Hu and Smith have shown this example has:** **[Ref 4.]**
  - **A combination of throughput and factor relationships**
  - **No risk applied to the factors**
  - **PMP drives about 70% of the model result, so 70% of the risk is modeled with a normal distribution making it reasonable that the total cost is likely to be normally distributed.**
- **We can model this example using an approach similar to our SNR example**

Case Study Page CE V – 80 SCEA Training Manual

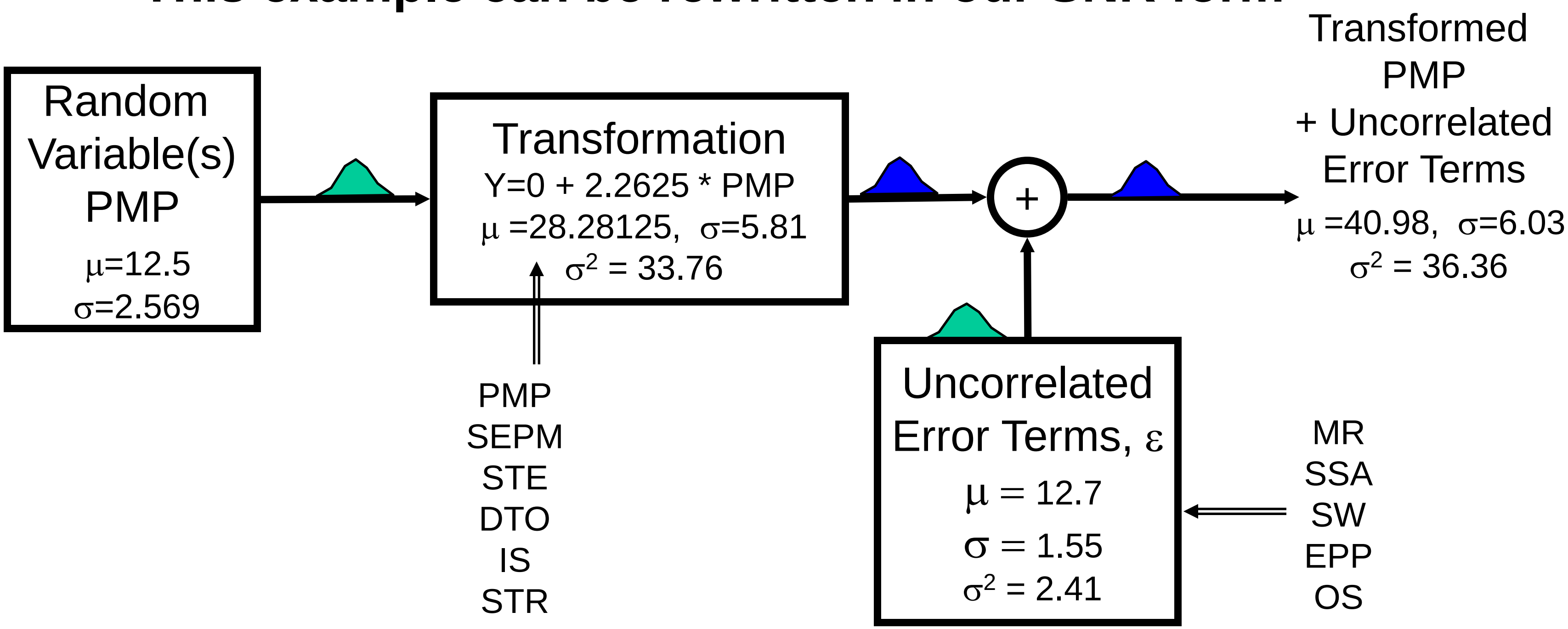
- A Signal (PMP) to Noise (η) Ratio Problem

WBS	Equation/ Throughput	Distrn	Lower	Point Estimate	Upper	Analytic Stdev	ACE Stdev	CB Stdev	@Risk Stdev
Electronic System						6.015	6.013	6.026	5.998
PMP	12.50	Normal		12.500		2.569	2.570	2.569	2.569
SEPM	0.5*PMP			6.250		1.285	1.285	1.284	1.285
Sys Test & Evaluation				4.706		0.811	0.811	0.812	0.809
Sys Test & Eval	0.3125*PMP			3.906		0.803	0.803	0.803	0.803
Management Reser	0.80	Uniform	0.6	0.800	1.0	0.115	0.116	0.115	0.115
Data and Tech Orders	0.1*PMP			1.250		0.257	0.257	0.257	0.257
Site Survey & Activatio	6.60	Tiangular	5.1	6.600	12.1	1.505	1.505	1.505	1.505
Initial Spares	0.1*PMP			1.250		0.257	0.257	0.257	0.257
System Warranty	1.10	Uniform	0.9	1.100	1.3	0.115	0.116	0.115	0.115
Early Prototype Phase	1.50	Triangular	1.0	1.500	2.4	0.290	0.290	0.290	0.290
Operations Supt	1.20	Triangular	0.9	1.200	1.6	0.143	0.143	0.143	0.143
System Training	0.25*PMP			3.125		0.642	0.643	0.642	0.642



SCEA Example in SNR form

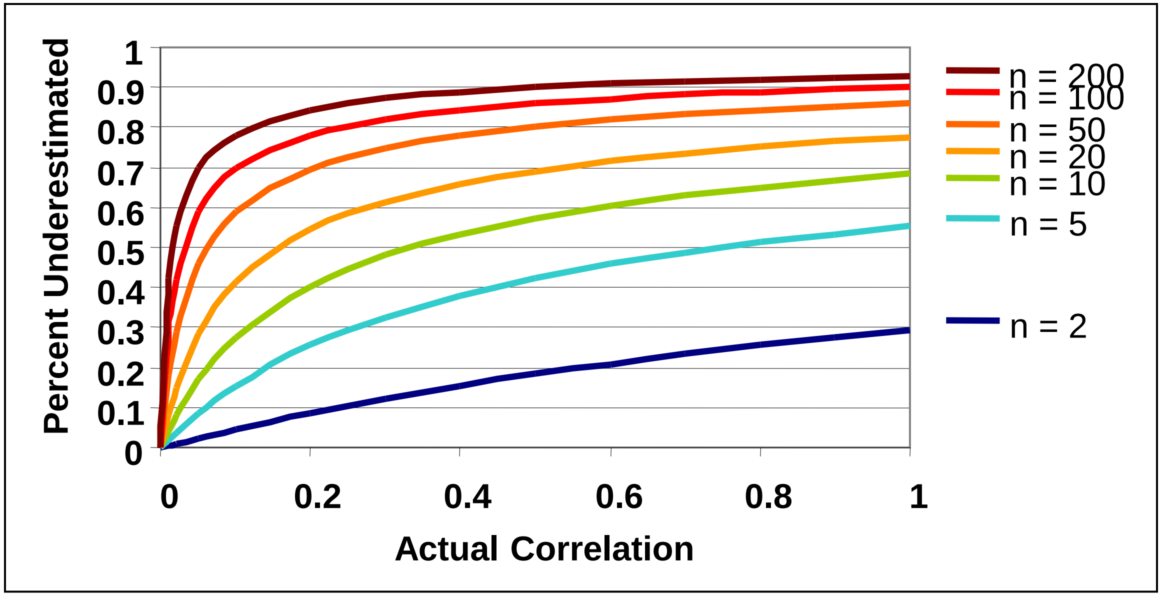
- **This example can be rewritten in our SNR form**



- **So the Ratio of variances is (SNR) 33.76/ 2.41 = 14.01**
  - **The signal dominates the noise by a significant degree**
- **If we correlate all of the noise terms to $\rho = 1$, it still doesn't make a difference SNR = 33.76 / 4.69 = 7.2**
  - **The signal still dominates the noise by a significant degree**

But when does it make a difference?

- **The sigma of the sum of the error terms in our example is sensitive to the number of terms, n**
  - **Correlation has a greater impacts as n increases**
  - **In the relationship:** $\sigma^2_{Total} = \sum_{k=1}^{n} \sigma_k^2 + 2\sum_{k=2}^{n} \sum_{j=1}^{k-1} \rho_{jk}\sigma_j\sigma_k$
    - **The number of covariance terms = n*(n-1)**
    - **The number of variance terms = n**



Example: The "Big" WBS

Suppose a risk analyst diligently applies distributions to all costs at the "level of estimating" – this is good.

- **Assume that:**
  - **There are 300 cost elements (N=300)**
  - **There are about four cost elements in each subsystem (n=4)**
  - **There are (N/n = 75 subsystems)**
  - **Correlation is defined between all elements within a subsystem using a grouping technique**
- **This means:**
  - **That** $$\frac{(N/n)\left[n*(n-1)\right]}{N*(N-1)} = \frac{(300/4)\left[4*3\right]}{300*299} = \frac{900}{89700} = 0.010033$$
  - **WBS elements are correlated**
  - **Only about 1% of the cost elements are correlated**
  - **Risk is very narrow and understated**
- **The correlation appears "just-off-the diagonal" of the correlation matrix – This is bad.**

"Just-Off-Diagonal" Correlation

1	0.5	0.5																	
	1	0.5																	
		1	0.5	0.5	0.5														
			1	0.5	0.5														
				1	0.5														
					1	0.5	0.5	0.5	0.5										
						1	0.5	0.5	0.5										
							1	0.5	0.5										
								1	0.5										
									1	0.5	0.5	0.5	0.5						
										1	0.5	0.5	0.5						
											1	0.5	0.5						
												1	0.5						
													1	0.5	0.5	0.5			
														1	0.5	0.5			
															1	0.5			
																1	0.5	0.5	0.5
																	1	0.5	0.5
																		1	0.5
																			1

These Inter-Subsystem WBS Elements are Effectively Uncorrelated

These Intra-Subsystem WBS Elements are Correlated

- **Some tools cannot support this function**
- **Some nominal statistical correlation does exist**
- **Even a few percent makes a big difference with a big WBS**

When Functional Correlation Dominates

- **Functional correlation could be the dominant type of correlation affecting the total cost variance if:**
  - **There are few WBS elements (less than about 30)**
    - **Allow the central limit theorem to dominate**
  - **The cost estimates of WBS elements are related to each other (applying a factor)**
    - **Functional relationships**
  - **The cost estimates are driven by few random variables that are “reused” directly or indirectly throughout the estimate**
    - **Lots of reuse**

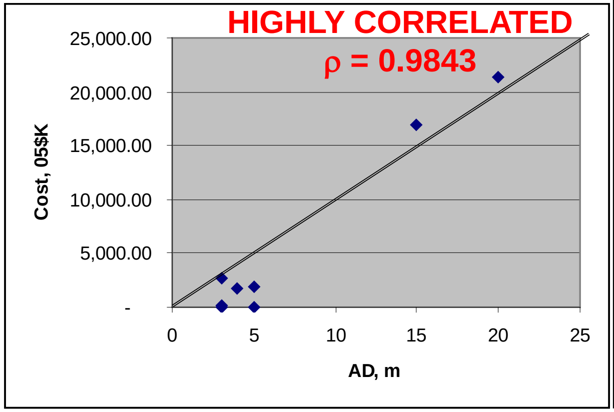
When Purely Statistical Correlation Dominates

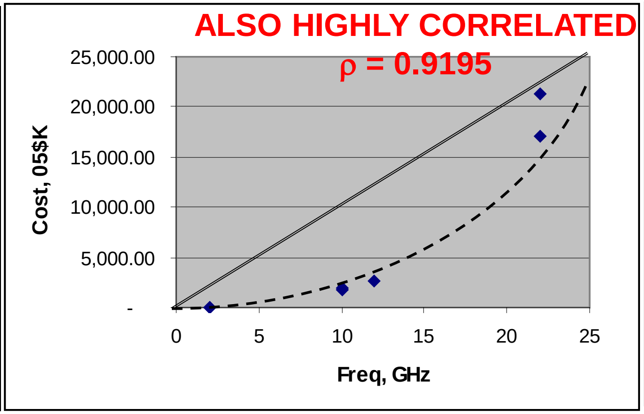
- **Purely Statistical Correlation will be the dominant source of correlation if these conditions are met:**
  - **There are many WBS elements (more than 30)**
  - **The WBS elements are grouped in parent WBS elements that are not causally related**
  - **The cost estimates for WBS elements are driven by several different random variables that are not related by functions in the model**
  - **There is little variance on the cost drivers**
  - **The cost estimating relationships are independent of each other (i.e. they are not functionally correlated) and little causal relationship is known between the variables**

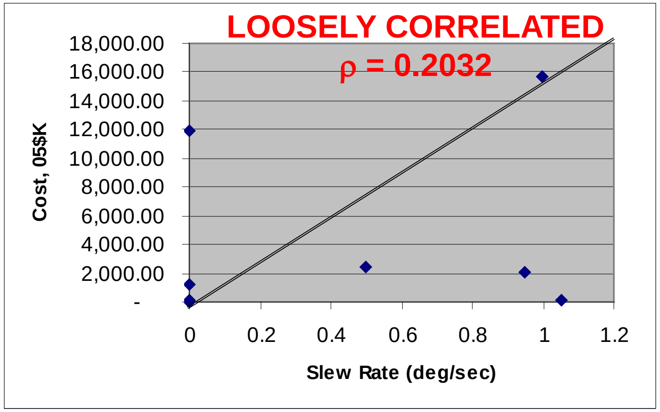
Find the Missing Causal Correlation

- **Correlation starts when we develop CERs**
  - **Find independent variables that are correlated with cost**
  - **Here is an example using fictitious antenna data:**

Program	Antenna Diameter, m	Frequency, GHz	Slew Rate, deg/sec	Cost BY05$K
1	3	2	0	45.40
2	3	2	1.05	83.72
3	5	2	0	78.56
4	15	22	0	11,882.00
5	20	22	1	15,656.33
6	4	10	0	1,243.68
7	3	12	0.95	2,087.28
8	5	10	0.5	2,436.85
Correl with Cost	0.9843	0.9195	0.2032	1.0000

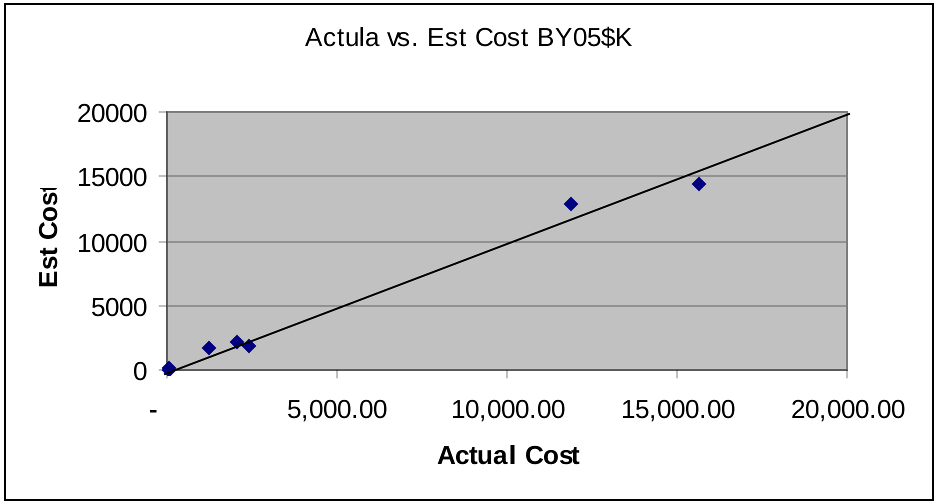






Regression Analysis

- **Pick antenna diameter (AD) and frequency (Freq) as cost drivers for a regression of the form Y = a*x1^b*x2^c**
- **The regression results from a zero percent bias, minimum percentage error are**
  - **Y = 11.239 * AD ^ 0.3931 * Freq ^ 1.935**
  - **Bias = 0%**
  - **% Std Error = 22.8%**



**We showed that cost was correlated to two variables:
Antenna Diameter , m
Frequency, GHz**

But, We Left Something Out

- **All our independent variables are (pretty much) uncorrelated with the percentage error (as they should be)**
  - **We didn't include the slew rate in the regression**
  - **Remember it was loosely correlated with cost, ρ = 0.2032**

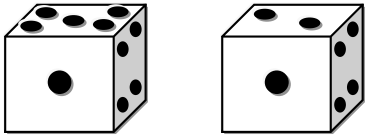
Program	Antenna Diameter, m	Frequency, GHz	Slew Rate, deg/sec	Cost BY05$K	Est	(Act-Est) /Est
1	3	2	0	45.40	66.17	(0.31)
2	3	2	1.05	83.72	66.17	0.27
3	5	2	0	78.56	80.89	(0.03)
4	15	22	0	11,882.00	12,884.91	(0.08)
5	20	22	1	15,656.33	14,427.77	0.09
6	4	10	0	1,243.68	1,667.15	(0.25)
7	3	12	0.95	2,087.28	2,118.54	(0.01)
8	5	10	0.5	2,436.85	1,820.01	0.34
Correl with Cost	0.9843	0.9195	0.2032	1.0000	% Bias	0.0000
Correlation with %Error	0.0847	0.0512	0.6565	0.0924	%SE	0.227617

- **But it has a correlation with the percentage error of cost, ρ = 0.6565, which can be used to drive your risk model!**

Hidden correlations exist

Effects of Correlation

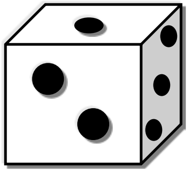
- **Why is correlation used?**
  - **To quantify the effects of statistical dependence when performing algebra on random variables.**
  - **It has a large impact on the statistical properties of the results, particularly when many random variables are involved.**

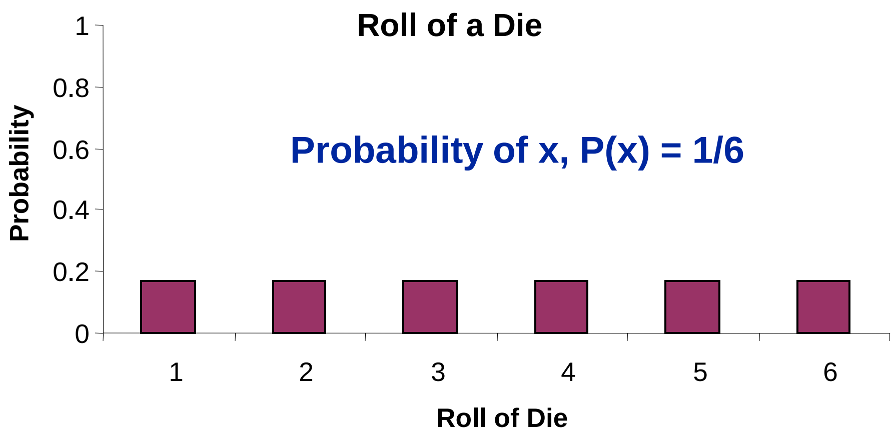


- **Example: Dice Roll.**
  - **What happens when we roll 2 dice and add their result?**
  - **Assume 3 cases:**
    - **Case 1: Uncorrelated. Outcome of 1 die is independent from the other.**
    - **Case 2: Negatively correlated. Outcome of 1 die relate to the outcome of the other. If one die is a "6", the other must be "1".**
    - **Case 3: Positively correlated. Outcome of 1 die is same as the other.**

Example: Dice Roll

- **Roll of the die gives an equal chance of getting an outcome (1,2,3,4,5 or 6)**
  - **Equal, discrete probability**
  - **Uniform discrete distribution of probabilities**
  - **Variance, $\sigma^2 = 3.5$**





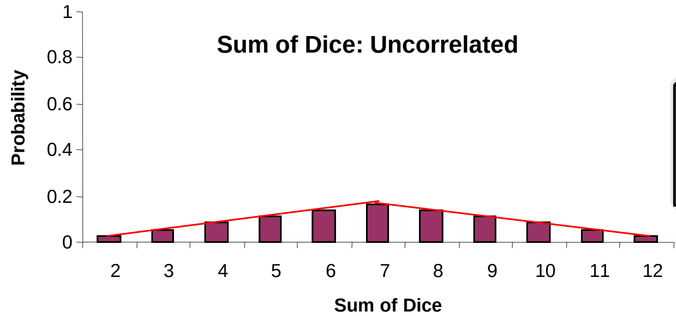
- **What happens when we sum 2 correlated dice?**

Example: Dice Roll

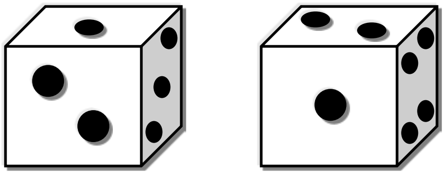
"Creating Customer-Focused Success"

Case 1: ρ = 0

Triangular, discrete shape
Moderate variance, $\sigma^2=6$
Mean = 7

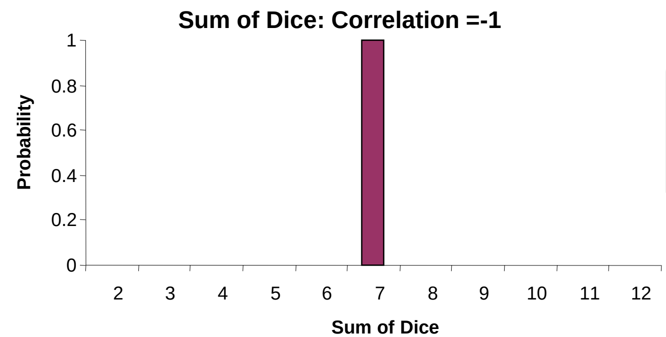


2 + 1 = 3

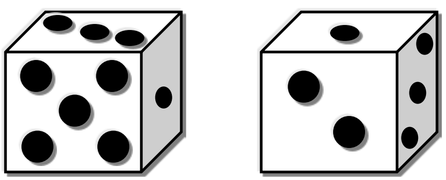


Case 2: ρ = -1

P(7) = 1
P(<>7)=0
No variance, $\sigma^2=0$
Mean = 7

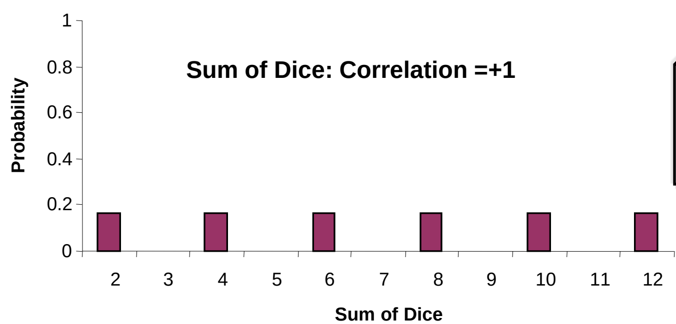


5 + 2 = 7

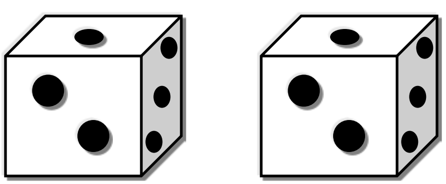


Case 2: ρ = +1

Uniform, discrete shape
P(each even)=1/6, P(odd) =0
Wide variance, $\sigma^2=14$
Mean = 7



2 + 2 = 4



Dice Roll Results*

- **What happens when we increase correlation from –1 to +1:**

Correlation	Mean	Variance	Kurtosis (Est)
-1	7	0	+Inf.
0	7	6	2.36
+1	7	14	1.73

- **Mean stays the same**
- **Variance increases with increasing correlation**
- **Kurtosis (measure of peakedness of the distribution) goes down with increasing correlation**
- **What we learned about the effects of correlation on sums of dice:**
  - **It affects the variance and shape**
  - **It doesn't affect the mean**
    - **ρ= 0 changes shape to a discrete triangular distribution**
    - **ρ=-1 changes shape and removes all variance**
    - **ρ=+1 preserves shape, adds the most variance, and is the same as multiplying by 2**

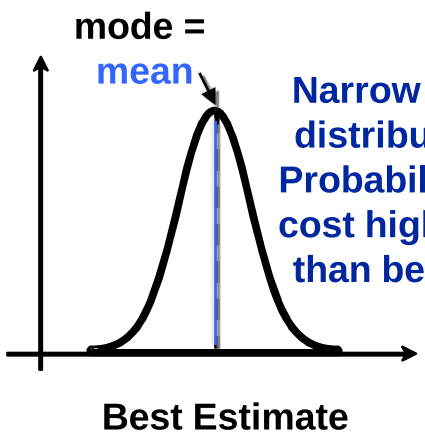
***Theoretical, of course**

What We Learned From Dice Roll

- **The sum of dice example used a discretely distributed random variable, but the same rules apply for continuously distributed random variables.**
  - **Uniform, Triangular, Normal, Lognormal, Weibull, Gamma, etc.**
- **You should care because correlation can be a huge contributor to the amount of risk in probabilistic cost estimates**

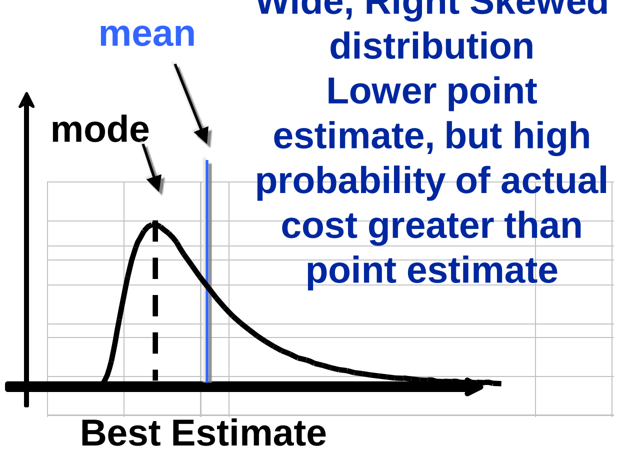
Cost-element Probability Distributions

Low Risk



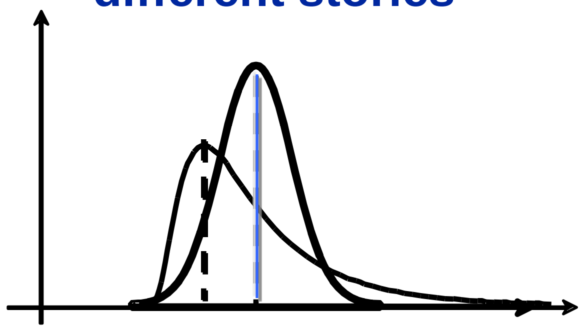
Narrow Symmetric distribution: equal Probability of actual cost higher or lower than best estimate

High Risk



Wide, Right Skewed distribution
Lower point estimate, but high probability of actual cost greater than point estimate

These curves tell two very different stories



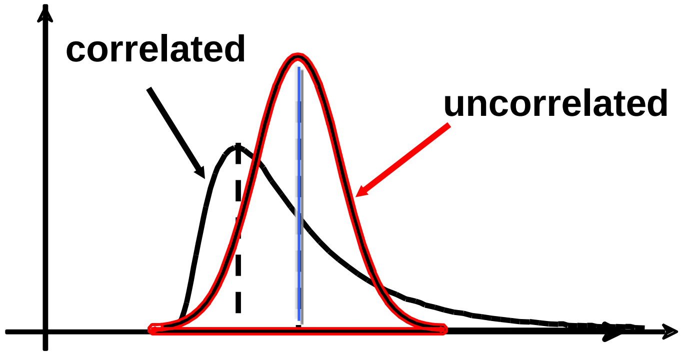
Low Cost, High Risk vs.

High Cost, Low Risk

Would you believe both could come from the same estimate?
Well, they do.

Two Distributions

- Remember these two curves?
- They were formed from the same estimate, but the skewed distribution includes correlation and the other assumes no correlation
  - Means are the same
  - Variance and skewness are different



Low Cost, High Risk vs.

High Cost, Low Risk

- Introduction
- The Different Types of Correlation
- Different Ways to Correlate Random Variables
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

Representing Correlation Matrices

- **Single value shorthand:** $\begin{bmatrix} 1 & & & & \\ & 1 & & \rho & \\ & & 1 & & \\ \rho & & & 1 & \\ & & & & 1 \end{bmatrix} = \rho$
  - **This means all of the off diagonal terms are the same value**
- **Correlation Matrix**
  - **Contains all inter element correlations**
- **The Rules:**
  - **Always positive definite**
  - **Diagonal terms always 1.0**
  - **Off diagonal terms are correlation values**
  - **Columns and rows are transposed,** $\rho_{j,k} = \rho_{k,j}$

- **Now for some practical examples**

Spacecraft Bus: USCM7 Correlation Coefficients

- **Correlation coefficients for USCM7 Weight based, Mean Unbiased Percentage Error (MUPE) CERs**
  - **Average correlation coefficient = 0.160**

	ADCSNR	AGENR	COMMNR	EPSNR	IATNR	PROGNR	STRCNR	THERNR	TT CNR	ADCST1	AKMT1	COMMT1	EPST1	IATT1	LOOST1	PROGT1	STRCT1	THERT1	TT CT1
ADCSNR	1.000	-0.067	-0.096	-0.035	0.035	0.012	0.413	0.605	0.121	-0.095	0.983	-0.122	0.099	0.564	0.139	0.089	-0.047	-0.057	0.092
AGENR		1.000	-0.028	0.525	-0.079	0.127	0.091	-0.230	-0.125	0.416	0.001	0.085	-0.043	-0.163	-0.189	0.033	0.146	0.151	0.232
COMMNR			1.000	0.888	0.884	0.966	0.762	0.281	0.850	-0.166	0.305	-0.176	0.157	0.368	0.884	-0.158	0.109	0.037	-0.004
EPSNR				1.000	0.265	0.604	0.409	0.003	0.337	0.237	0.011	-0.275	0.076	0.342	0.021	-0.049	0.465	0.123	0.035
IATNR					1.000	0.721	0.615	0.331	0.747	-0.037	0.391	-0.133	-0.028	0.501	0.265	-0.145	0.113	-0.014	-0.189
PROGNR						1.000	0.697	0.222	0.868	-0.065	0.145	-0.191	-0.044	0.444	0.329	-0.191	-0.000	-0.125	0.019
STRCNR							1.000	0.837	0.761	-0.001	0.117	-0.214	-0.113	0.418	0.173	-0.018	0.220	-0.103	0.069
THERNR								1.000	0.077	-0.200	0.662	-0.171	-0.053	0.514	0.102	-0.010	-0.063	-0.165	0.092
TT CNR									1.000	-0.149	0.475	-0.118	-0.071	0.519	0.294	-0.178	-0.111	-0.095	0.022
ADCST1										1.000	-0.100	0.614	0.421	-0.262	-0.354	0.543	0.676	-0.029	0.655
AKMT1											1.000	-0.006	0.292	0.855	0.286	0.176	-0.003	-0.027	0.052
COMMT1												1.000	0.266	-0.454	-0.088	0.777	0.729	0.126	0.391
EPST1													1.000	-0.150	-0.145	0.381	0.388	-0.007	0.520
IATT1														1.000	0.448	-0.144	-0.224	-0.014	-0.320
LOOST1															1.000	-0.336	-0.097	-0.074	-0.169
PROGT1																1.000	0.421	-0.039	0.481
STRCT1																	1.000	-0.175	0.285
THERT1																		1.000	-0.140
TT CT1																			1.000

These correlation coefficients should not be used for all spacecraft cost models

How to Use Correlation Matrices

- Typically, we wouldn't want to define all of the correlation coefficients for a big WBS (>10 elements)

- We can break it up into parts, get the statistics and then sum at higher levels
  - Reduces the size of correlation matrices
  - Provides Risk Breakout by WBS Summary Level

- Lets use an example of a "Big" WBS with 40 elements

40 Element WBS

- **40 Individual WBS Elements and the correlation matrix**

- SEITPM
  - Systems Engineering
  - Integration & Test
  - Program Management
  - Configuration Management
  - Data
- Space
  - Space Vehicle SEITPM
  - Space Vehicle
    - Spacecarft Bus
      - Bus Systems Engineering
      - Bus I&T
      - Bus PM
      - Bus Data
      - Structures & Mechanisms
      - Thermal Control
      - Attitude Determination & Control
      - TTC / C&DH
      - Propulsion
      - Elecrical Power
      - LOOS
      - AGE
    - Payload
      - PL SEITPM
      - Optical Telescope
      - Panchomatic Sensor
      - Multispectral Sensor
      - Spectrometer
      - Magnetometer
      - Gravitometer
      - UV Sensor
- Ground
  - Ground SEITPM
  - Ground Terminal
  - Mission Planning
  - Satellite OPS / Control
  - Data Archive and Dissemination
- Launch
  - Launch Vehicle
  - Launch Systems Integration
  - Launch Vehicle Integration
  - On-Orbit Checkout
  - Launch Vehicle SE
- Operations
  - SEITPM
  - Maintenance
  - Mission Planning
  - Mission Ops
  - Data Archive and Dissemination

1	0.2	0.4	0.3	0.3	0.1	0.1	0.3	0.2	0.1	0.1	0.1	0	0.2	0.3	0.1	0	0.2	0.4	0.5	0.5	0.1	0.1	0.4	0.4	0.1	0.1	0.4	0.2	0	0.1	0.3	0.2	0.3	0.3	0.4	0.3	0.2
	1	0.1	0.2	0.3	0	0.2	0.4	0.2	0.4	0.3	0.3	0.2	0.5	0.1	0.1	0.1	0.3	0.5	0.3	0	0.4	0.1	0.1	0.4	0.1	0.1	0.5	0	0.3	0.1	0.1	0.4	0.2	0.5	0.2	0.5	0
		1	0.3	0.5	0.3	0.3	0.2	0.1	0.4	0.3	0.2	0.5	0	0.1	0.5	0.4	0.4	0.4	0.4	0.5	0.3	0.4	0.5	0.2	0	0	0.4	0.1	0.3	0.5	0.3	0	0.1	0.2	0.4	0.5	0.4
			1	0.5	0.1	0.3	0.5	0.2	0.1	0.1	0.3	0.1	0.2	0.5	0.3	0.5	0.3	0.3	0.3	0.2	0.5	0.3	0.1	0.4	0.3	0.1	0.5	0.4	0.4	0.5	0.4	0.4	0.4	0.2	0.4	0.5	0.2
				1	0.2	0.4	0.1	0	0.3	0.3	0.4	0	0.2	0.4	0.5	0.3	0.2	0.1	0.2	0.5	0.2	0.4	0.5	0.2	0.2	0.4	0.4	0.4	0.5	0.3	0.4	0.5	0	0.2	0.3	0.3	0.2
					1	0.1	0.2	0	0.1	0.2	0.3	0	0.5	0.2	0.1	0.3	0.1	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.4	0.1	0.1	0.2	0.4	0.4	0.2	0	0.5	0.5	0	0.1	0.1
						1	0.3	0.2	0.3	0.3	0.1	0.1	0.1	0.4	0	0.1	0.3	0.4	0.2	0.4	0.4	0.2	0.1	0.1	0.3	0	0.4	0.3	0.5	0.1	0.4	0.5	0.5	0.1	0.1	0.2	0.4
							1	0.4	0.4	0.2	0.2	0.5	0.5	0	0.4	0	0.2	0.2	0.3	0.1	0.4	0.2	0	0.1	0.4	0.1	0.1	0.1	0	0.1	0.4	0.4	0	0.3	0.5	0	0.3
								1	0.1	0	0.1	0	0.2	0.5	0.4	0.3	0.2	0	0.5	0	0.4	0.3	0.5	0.1	0.1	0	0.4	0.2	0.1	0.4	0.2	0.1	0.1	0.1	0.4	0.3	0.1
									1	0.4	0	0.2	0.5	0.1	0.3	0.5	0.4	0.3	0.1	0.1	0.5	0.4	0	0.5	0.1	0.5	0.1	0.2	0.1	0.2	0.1	0.1	0.2	0.3	0.5	0.1	0.2
										1	0.4	0.3	0.3	0.1	0.1	0.1	0.3	0.5	0.3	0.1	0.5	0.1	0.3	0.4	0.2	0.2	0.2	0.5	0.1	0.5	0.2	0.3	0	0.1	0.1	0.2	0.5
											1	0.2	0.1	0	0.1	0.3	0	0.3	0.3	0.4	0.2	0.5	0.1	0	0	0.2	0.2	0.1	0.4	0.3	0	0.3	0.2	0.3	0.4	0.3	0.3
												1	0.2	0.1	0.1	0.2	0.4	0.5	0.4	0	0.3	0.5	0.2	0.2	0.3	0	0.4	0.1	0.1	0.2	0.1	0.2	0.4	0.3	0.2	0.1	0.1
													1	0.2	0.1	0.2	0.3	0.1	0.5	0.2	0.4	0.4	0.5	0.5	0.4	0.4	0.2	0.3	0.2	0.4	0.3	0.1	0	0.1	0.5	0.3	0.3
														1	0.2	0.1	0.1	0.2	0.4	0.2	0.2	0.3	0.5	0.4	0.2	0.4	0.4	0	0.3	0.4	0.4	0.2	0.3	0	0.4	0.2	0.3
															1	0.3	0.1	0	0	0.2	0.2	0.3	0.4	0	0.1	0.4	0.1	0.3	0.1	0.4	0.1	0.3	0.2	0.2	0.1	0.4	0.1
																1	0.3	0.4	0.3	0.4	0.1	0.4	0	0.4	0.1	0.4	0.1	0.4	0.3	0.5	0.2	0.2	0.3	0.4	0.5	0.1	0.4
																	1	0.3	0.3	0.2	0	0.3	0.5	0.4	0.4	0.4	0.4	0.2	0.4	0.2	0.1	0.3	0.3	0.5	0.4	0.5	0.4
																		1	0	0.5	0.3	0	0.3	0.3	0.2	0.5	0.4	0.1	0.3	0.4	0.4	0.4	0.5	0	0.2	0.5	0.3
																			1	0.3	0.1	0.5	0.3	0.1	0.2	0.3	0.1	0.1	0.4	0.4	0.2	0.1	0.1	0.3	0.1	0.2	0.1
																				1	0.1	0.1	0.1	0.2	0.4	0.3	0.5	0.4	0.1	0	0.5	0.2	0.4	0.1	0.2	0.5	0.3
																					1	0.3	0.3	0.2	0.5	0.3	0.1	0.4	0.3	0.3	0.2	0.2	0.3	0	0.5	0.3	0.4
																						1	0.1	0.1	0.5	0.3	0.5	0.4	0.1	0.2	0.3	0.2	0.5	0	0.3	0.1	0.4
																							1	0.2	0.1	0.4	0.3	0.3	0.3	0.1	0.4	0.2	0.2	0.2	0.2	0.5	0.3
																								1	0.5	0.5	0.2	0.3	0.1	0.3	0.4	0.3	0.4	0.3	0.3	0.2	0.3
																									1	0.3	0.2	0.5	0.3	0.5	0.1	0.4	0.3	0.1	0.3	0.1	0.2
																										1	0.5	0.3	0.1	0.1	0.2	0.1	0.4	0.2	0.2	0.5	0.2
																											1	0.2	0.4	0.5	0.4	0.1	0	0.5	0.3	0.3	0.1
																												1	0.1	0.2	0.3	0.1	0	0.4	0.2	0.3	0.5
																													1	0.5	0.2	0.2	0.3	0.1	0.4	0.4	0.4
																														1	0.1	0.1	0.3	0.2	0	0	0.4
																															1	0.2	0.2	0.4	0.2	0.1	0.4
																																1	0.4	0.1	0.2	0.3	0.2
																																	1	0.4	0.3	0.3	0.1
																																		1	0.1	0	0.2
																																			1	0	0.4
																																				1	0.2
																																					1

Imagine 300 WBS Elements!

Big Correlation Matrix Layout

"Creating Customer-Focused Success"

A= SETPM
(5 elements)

B= Space Segment
(20 elements)

C= Ground Segment
(5 elements)

D= Launch Segment
(5) elements)

E= Operations Segment
(5 elements)

AA	AB	AC	AD	AE
BA	BB	BC	BD	BE
CA	CB	CC	CD	CE
DA	DB	DC	DD	DE
EA	EB	EC	ED	EE

Each Block Represents a group of inter element correlations
The full matrix requires (40*39)/2=780 different correlations

Multilevel Risk

- **Look at the problem one small set of pieces at a time**

	Mean	Sigma
SEITPM	**11.22**	**2.70**
Systems Engineering	1.2	0.24
Integration & Test	1.8	0.72
Program Management	0.9	0.27
Configuration Management	7.2	2.16
Data	0.12	0.048

ρ_{SEITPM}

1	0.3	0.3	0.3	0.3
0.3	1	0.3	0.3	0.3
0.3	0.3	1	0.3	0.3
0.3	0.3	0.3	1	0.3
0.3	0.3	0.3	0.3	1

SEITPM Mean = Sum(Means)
SEITPM Sigma =SQRT(MMULT(MMULT(TRANSPOSE(sigma),correl_matrix),sigma))

Space Element Risk

- **In the Space Element, first break-out the Bus calculation**

	Mean	Sigma
Spacecarft Bus	**40**	**7.48**
Bus Systems Engineering	3.2	1.12
Bus I&T	3.3	1.32
Bus PM	1	0.25
Bus Data	0.5	0.2
Structures & Mechanisms	1	0.3
Thermal Control	1	0.3
Attitude Determination & Control	8	2.4
TTC / C&DH	10	3
Propulsion	6	1.8
Elecrical Power	3	0.9
LOOS	2	0.6
AGE	1	0.3

$\rho_{\text{S/C Bus}}$

1	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
0.25	1	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	1	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	1	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	1	0.25	0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	1	0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25	1	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1	0.25
0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1

- **Space Vehicle SEITPM is one line item,**

	Mean	Sigma
Space Vehicle SEITPM	**43.1**	**19.38**

- **Let's assume we already calculated mean and sigma for the Payload, like we did for the Bus**

	Mean	Sigma
Payload	**83.0**	**17.74**

Space Element Risk

- **Now we roll-up 3 Items:**

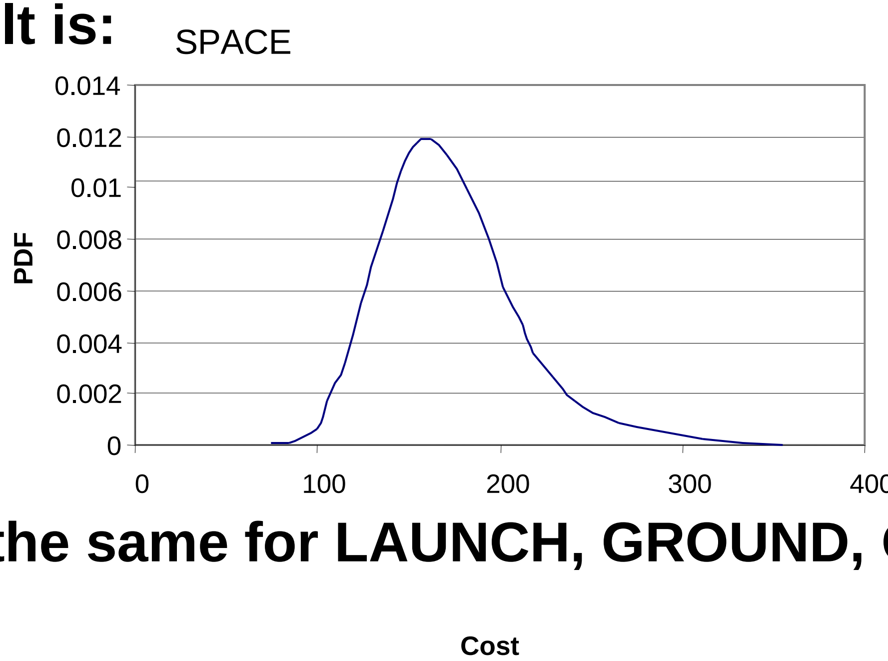
	Mean	**Sigma**
Space Vehicle SEITPM	43.1	19.38
Spacecarft Bus	40.0	7.48
Payload	83.0	17.74

- **Use a small correlation matrix:**

1	0.4	0.4
0.4	1	0.4
0.4	0.4	1

- **The result is:**

Mean	**Sigma**
166.1	**35.26**



- **Now do the same for LAUNCH, GROUND, O&M**

Economics of Multi-Level Risk

- **After summing all of the elements, we used :**
  - **144 Correlation Coefficients vs. 780 (one big matrix)**

	# Elements	# Rhos
Total System	5	10
SEITPM	5	10
Space Element	0	0
	12	66
	8	28
Ground	5	10
Launch	5	10
Ops	5	10
Number of Rhos		144

- **Views of Risk at all roll-up levels**
- **Easier to obtain values for correlation coefficients**
  - **We will discuss this in the next part**

What We Just Did

A= SETPM
(5 elements)
B= Space Segment
(20 elements)

C= Ground Segment
(5 elements)

D= Launch Segment
(5) elements)

E= Operations Segment
(5 elements)

AA	AB	AC	AD	AE
BA	BB	BC	BD	BE
CA	CB	CC	CD	CE
DA	DB	DC	DD	DE
EA	EB	EC	ED	EE

Relied on AA, BB, CC, DD, and EE correlation

Mathematically

Step1: Calculate σ_A, σ_B

$$\sigma = \begin{bmatrix} \sigma_A \\ \sigma_B \end{bmatrix} = \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \end{bmatrix}$$

; Where $\sigma_A = \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}$ **, and** $\sigma_B = \begin{bmatrix} \sigma_3 \\ \sigma_4 \\ \sigma_5 \end{bmatrix}$

Step 2: Need correlation coefficients of partition AA, BB and all σs to calculate σ_A, σ_B

$$\sigma_{Tot}^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2 + 2\begin{bmatrix} \rho_{12}\sigma_1\sigma_2 + \rho_{13}\sigma_1\sigma_3 + \rho_{14}\sigma_1\sigma_4 + \rho_{15}\sigma_1\sigma_5 + \\ \rho_{23}\sigma_2\sigma_3 + \rho_{24}\sigma_2\sigma_4 + \rho_{25}\sigma_2\sigma_5 + \\ \rho_{34}\sigma_3\sigma_4 + \rho_{35}\sigma_3\sigma_5 + \\ \rho_{45}\sigma_4\sigma_5 \end{bmatrix}$$

Mathematically

- **Step 3: Calculate total variance using** ρ_{AB}
- $\sigma_{Tot}^2 = \sigma_A^2 + \sigma_B^2 + 2\rho_{AB}\sigma_A\sigma_B$
- $\rho_{AB}\sigma_A\sigma_B = \rho_{13}\sigma_1\sigma_3 + \rho_{14}\sigma_1\sigma_4 + \rho_{15}\sigma_1\sigma_5 + \rho_{23}\sigma_2\sigma_3 + \rho_{24}\sigma_2\sigma_4 + \rho_{25}\sigma_2\sigma_5$
-
- $$\rho_{AB} = \frac{\rho_{13}\sigma_1\sigma_3 + \rho_{14}\sigma_1\sigma_4 + \rho_{15}\sigma_1\sigma_5 + \rho_{23}\sigma_2\sigma_3 + \rho_{24}\sigma_2\sigma_4 + \rho_{25}\sigma_2\sigma_5}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2}\sqrt{\sigma_3^2 + \sigma_4^2 + \sigma_5^2 + 2\left[\rho_{34}\sigma_3\sigma_4 + \rho_{35}\sigma_3\sigma_5 + \rho_{45}\sigma_4\sigma_5\right]}}$$

- **This is useful when :**
- **We know the correlation between subsystem elements**
  - **But not the correlation between subsystems from different elements to each other (i.e., thermal control SS in spacecraft to ground Command and control CSCIs)**
  - **But do have an idea of correlation between higher-level elements like space to ground.**

Mathematically

$$\rho_{AB}\sigma_A\sigma_B = \underbrace{\rho_{13}\sigma_1\sigma_3 + \rho_{14}\sigma_1\sigma_4 + \rho_{15}\sigma_1\sigma_5 + \rho_{23}\sigma_2\sigma_3 + \rho_{24}\sigma_2\sigma_4 + \rho_{25}\sigma_2\sigma_5}$$

ρ_{AA}

ρ_{BA}

11	12	13	14	15
21	22	23	24	25
31	32	33	34	35
41	42	43	44	45
51	52	53	54	55

ρ_{AB}

ρ_{BB}

Correlation between two larger blocks include the inter-element correlation coefficients from the large matrix

•Introduction

•The Different Types of Correlation

•Different Ways to Correlate Random Variables

•Impact of Correlation on Risk Analysis

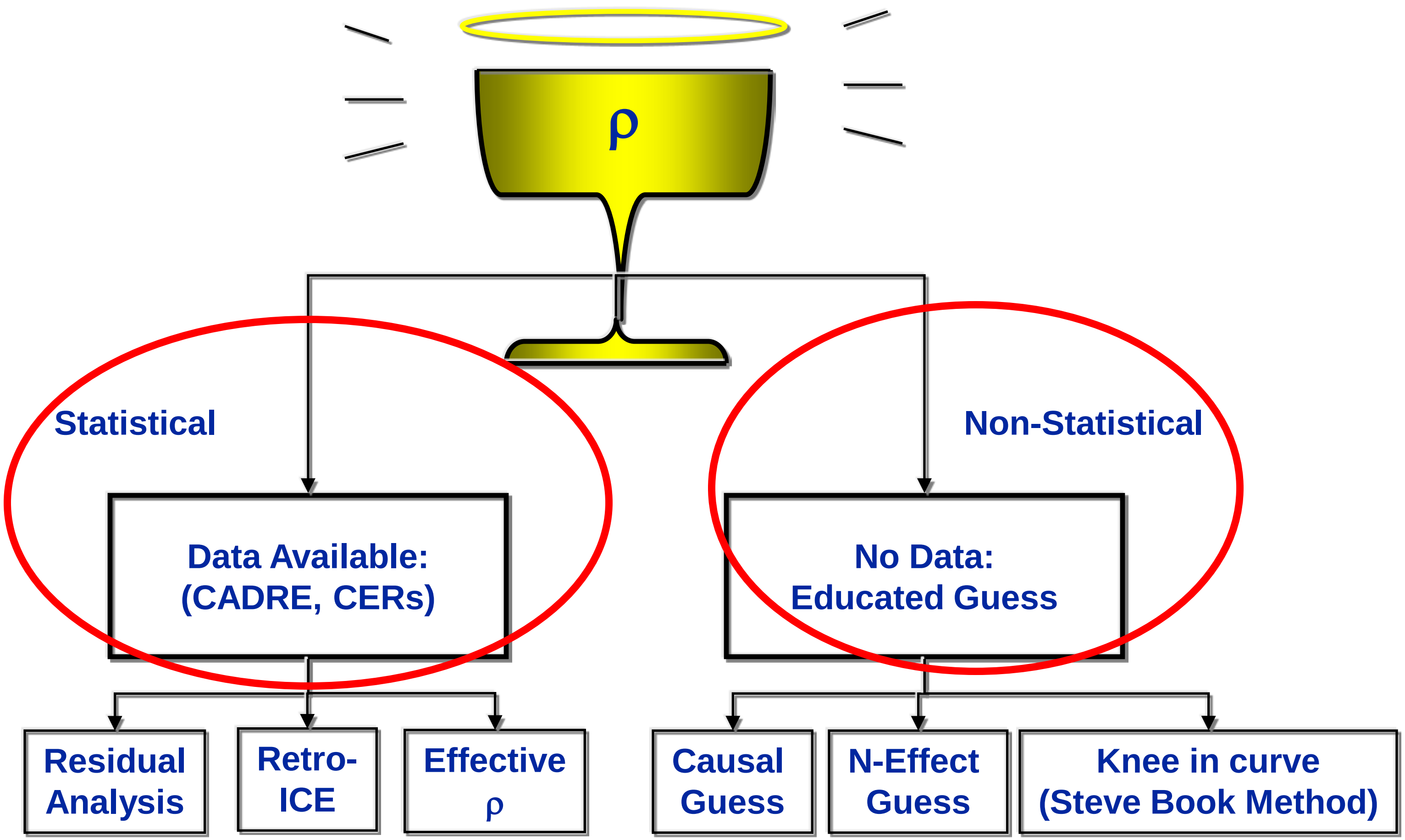
•Modeling Correlation

•Deriving Correlation Coefficients

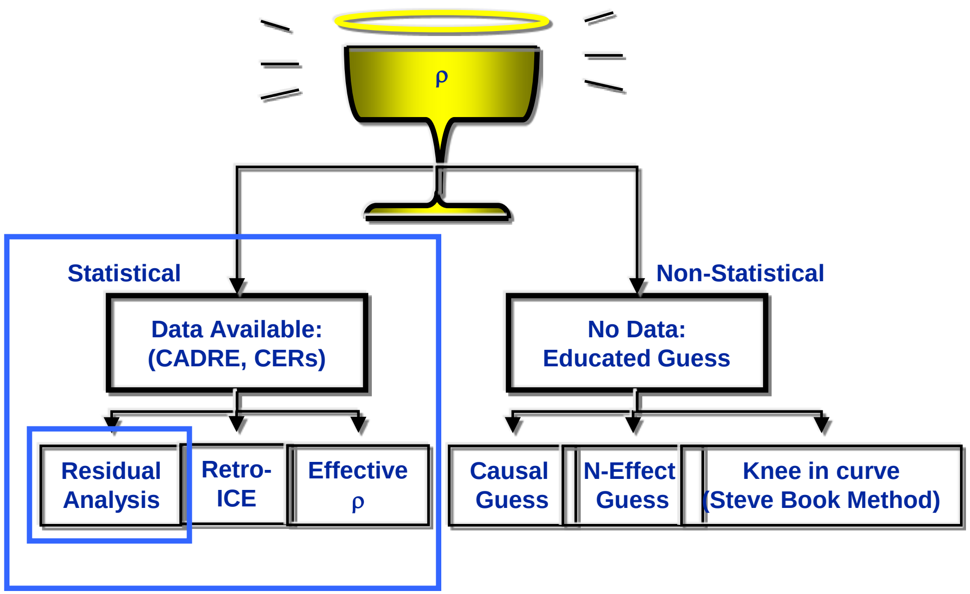
•Summary

Deriving Correlation Coefficients

- **2 Ways to derive correlation coefficients**

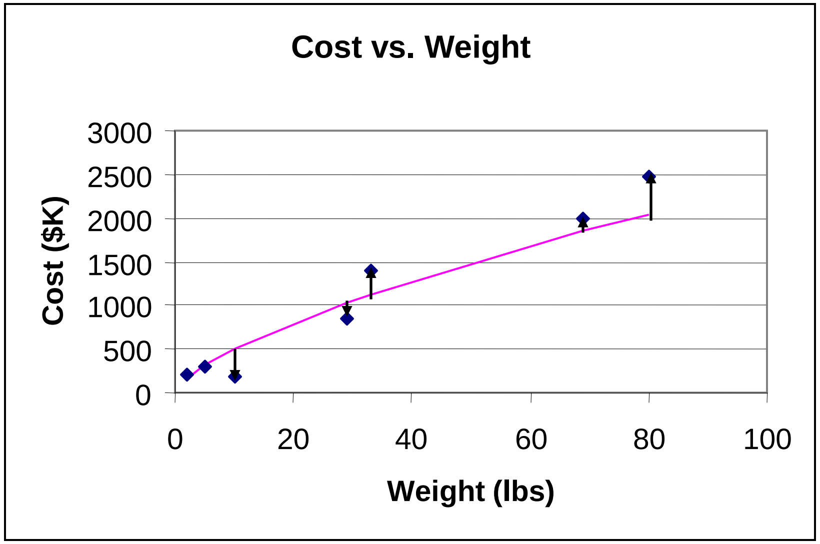


One Example: Determining Correlation When Data is Available using the Residual Analysis Method



Statistical Correlation From Residual Analysis

- **Percentage error or standard error are a measure of residual errors**
- **Uncertainty and risk calculations**
  - **Use residual errors to represent uncertainty**
  - **Correlation between residuals**



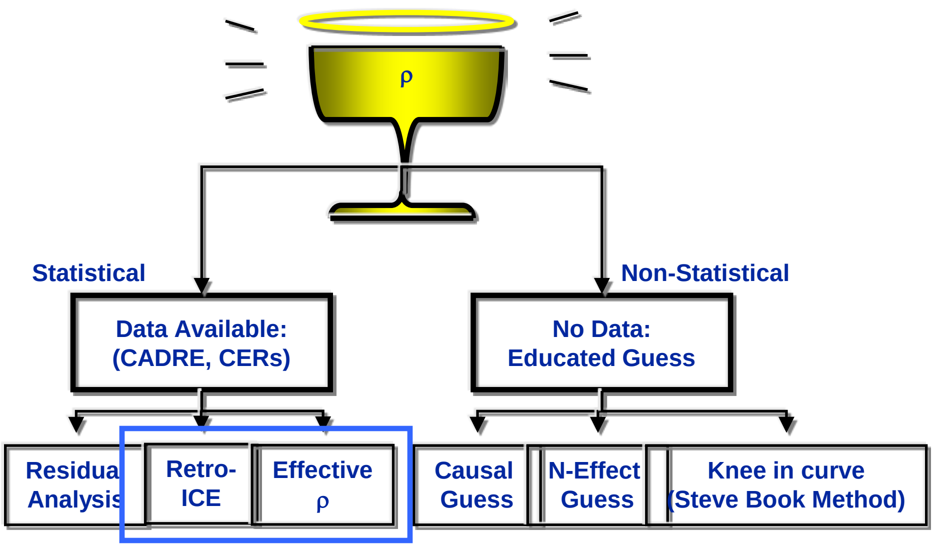
Deriving Correlation Coefficients

- **Sample calculation using randomly generated numbers**
  - **Error X_i and Error Y_i represent regression residuals for 2 CERs (X and Y) for 8 programs**

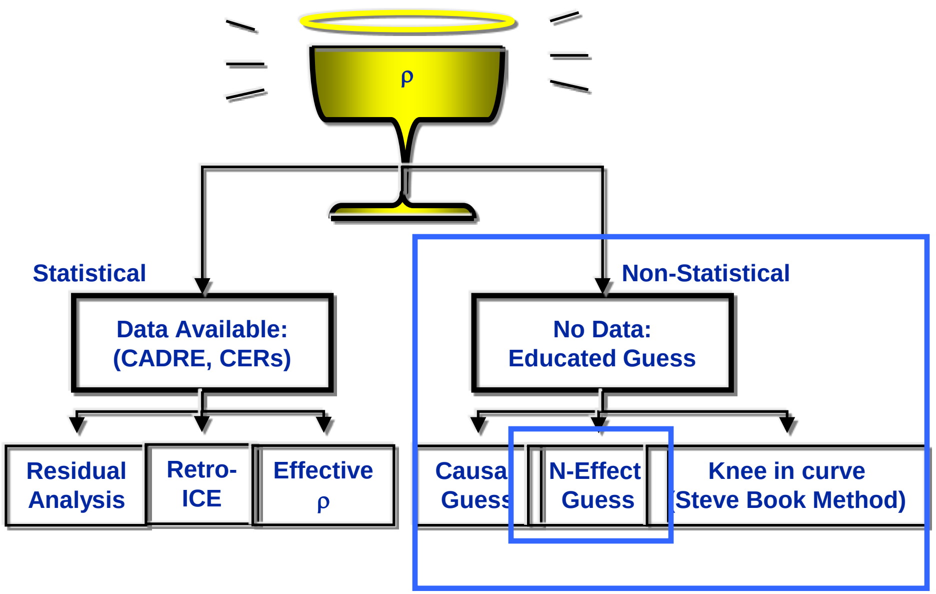
PROGRAM	Error, Xi	Error, Yi	(Xi-Xm)	(Yi-Ym)	(Xi-Xm)(Yi-Ym)	(Xi-Xm)^2	(Yi-Ym)^2
1	0.5404	0.4224	0.1102	0.0167	0.0018	0.0121	0.0003
2	0.4943	0.3719	0.0641	-0.0339	-0.0022	0.0041	0.0011
3	0.4496	0.4340	0.0194	0.0282	0.0005	0.0004	0.0008
4	0.0088	0.2598	-0.4214	-0.1460	0.0615	0.1776	0.0213
5	0.5679	0.4291	0.1377	0.0234	0.0032	0.0190	0.0005
6	0.4486	0.5126	0.0184	0.1069	0.0020	0.0003	0.0114
7	0.7960	0.5357	0.3659	0.1300	0.0475	0.1339	0.0169
8	0.1359	0.2804	-0.2943	-0.1253	0.0369	0.0866	0.0157
SUM					0.1513	0.4340	0.0681
MEAN	0.4302	0.4057					
RHO	0.8804	= 0.151 / SQRT(0.434 * 0.068)					

$$\rho_{jk} = \frac{\sum (x_i - x_m)(y_i - y_m)}{\sqrt{\sum (x_i - x_m)^2 \sum (y_i - y_m)^2}}$$

Other examples available on request…



Determining Correlation When Data is Not Available using the “N-effect” Correlation Method



"Creating Customer-Focused Success"

The Problem

- It is not always possible to calculate statistical correlation between WBS elements.
  - May be insufficient data to determine statistical correlation.
  - May be no known functional relationship between WBS elements.
- Yet, there may be reason to believe increases or decreases in the cost of a certain WBS element are likely to cause corresponding increases or decreases in the cost of another WBS element.
- In cases such as these, it is still desirable to construct a correlation matrix in order to ensure a truer picture of the total cost variance.

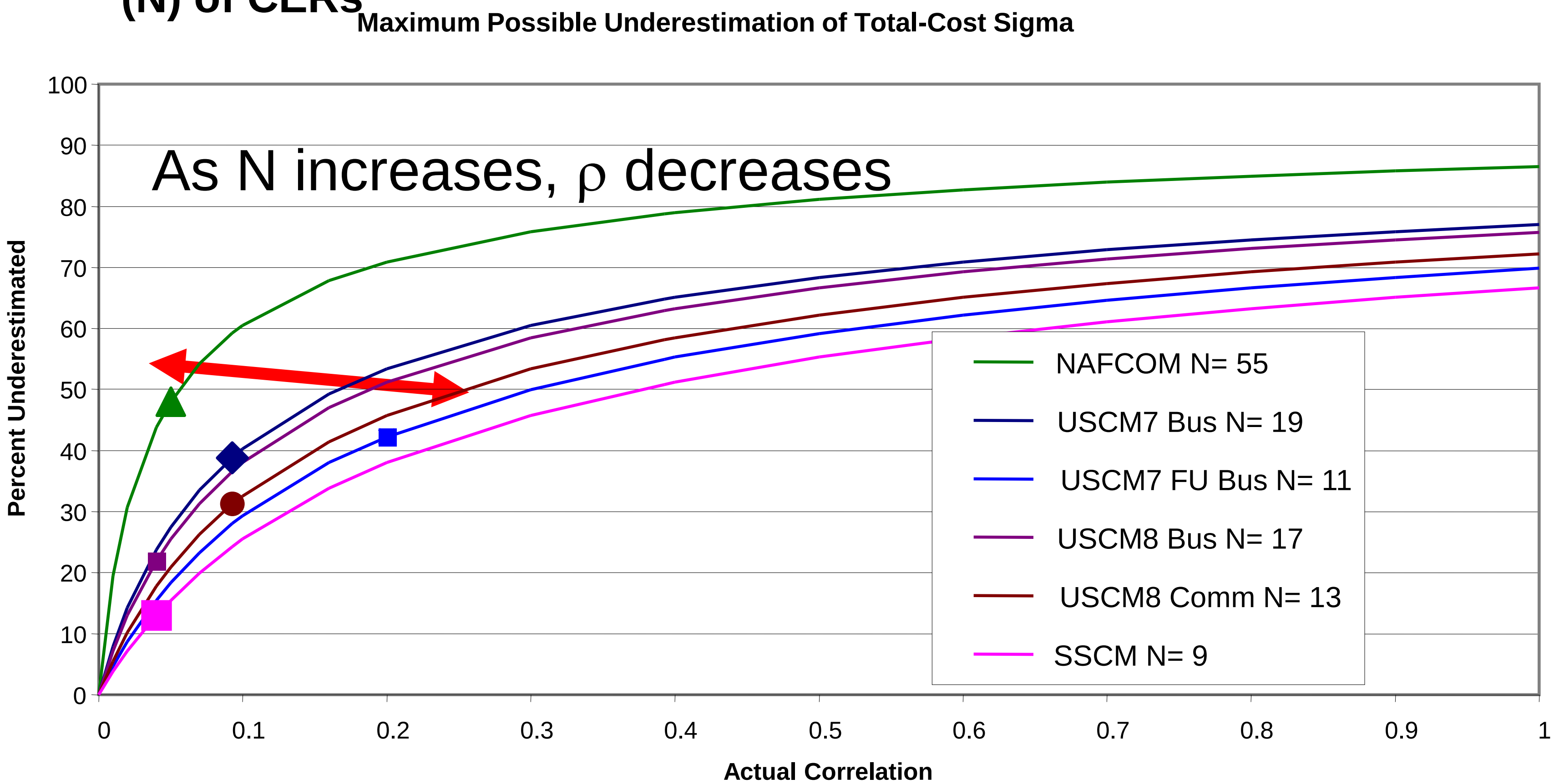
“N-effect” Correlation (1)

- As N increases, the effective correlation (ρ_{eff}) will decrease in reaction to the central limit theorem. This is the “N-effect”

- Why? There is a fundamental limit to the predictive capability of our CERs. Just by breaking the WBS up into more pieces doesn’t improve our estimates.

“N-effect” Correlation (2)

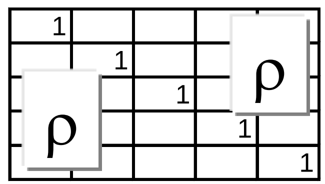
- **Average Correlation* in models seem to be sensitive to number (N) of CERs**



•The average correlation is different from the effective correlation, but the effect is similar

Determining Correlation from the Number of WBS Items

- **There appears to be a trend between the number of WBS Elements (N) in a cost model and the derived average correlation coefficient (ρ_{AVG}) and effective correlation ρ_{EFF}**
- **ρ_{EFF} is a single number used to fill the correlation matrix**



- **As N increases, ρ_{EFF} decreases**
- **We looked at the following models:**
  - **NAFCOM (NASA/ Air Force Cost Model)**
  - **USCM7 (Unmanned Space Vehicle Cost Model, Ver. 7)**
  - **USCM8 (Unmanned Space Vehicle Cost Model, Ver. 8)**
  - **SSCM (Small Satellite Cost Model)**

Determining Correlation from Number of WBS Items

- **If we see a trend in the chart of percent under-estimation of sigma vs. effective correlation, we have a sound basis for determining correlations when the number of WBS elements grows.**
- **If the actual percent underestimated is *k* then the N-effect correlation ρ_N for a model with *N* CERs would be:**

$$\rho = \frac{\dfrac{1}{\left(1-\dfrac{k}{100}\right)^2}-1}{N-1}$$

- **So, for k=50%:** $$\rho = \frac{\dfrac{1}{\left(1-\dfrac{k}{100}\right)^2}-1}{N-1} = \frac{\dfrac{1}{(1-.5)^2}-1}{N-1} = \frac{4-1}{N-1} = \frac{3}{N-1}$$

N	ρ
10	0.333
15	0.214
20	0.158
30	0.103
50	0.061
100	0.030
150	0.020
200	0.015
300	0.010
500	0.006

- Introduction
- The Different Types of Correlation
- Impact of Correlation on Risk Analysis
- Modeling Correlation
- Deriving Correlation Coefficients
- Summary

Summary

- **Two types of correlation**
  - **Spearman (Rank) = Monotonicity**
  - **Pearson (Product Moment) = Linearity**
- **Different Ways to Correlate Random Variables**
  - **Purely Statistical**
  - **Functional**
  - **Causal Statistical**
- **Impact of Correlation on Risk Analysis**
  - **Affects shape and variance**
- **Modeling Correlation**
  - **Multilevel risk**
- **Deriving Correlation Coefficients**
  - **Many choices available with and without available data**

References

1. www.statlets.com/usermanual/glossary.htm
2. Coleman, R.L, and Gupta, S.S., “An Overview of Correlation and Functional Dependencies in Cost Risk and Uncertainty Analysis”, DoDCAS, 1994.
3. Hu, S., “Correlation Analysis for Weight Variables in the USCM8 Database”, 2003 ISPA/SCEA Conference, Orlando, FL, June 2003.
4. Hu, S, and Smith, A.,“Cost Risk Analysis ‘Made Simple’, AAIA Conference, San Diego, September 2004.
5. Covert, R., "Comparison of Spacecraft Cost Model Correlation Coefficients", Presented at 74th SSCAG, San Pedro, CA, February 12-13, 2002.