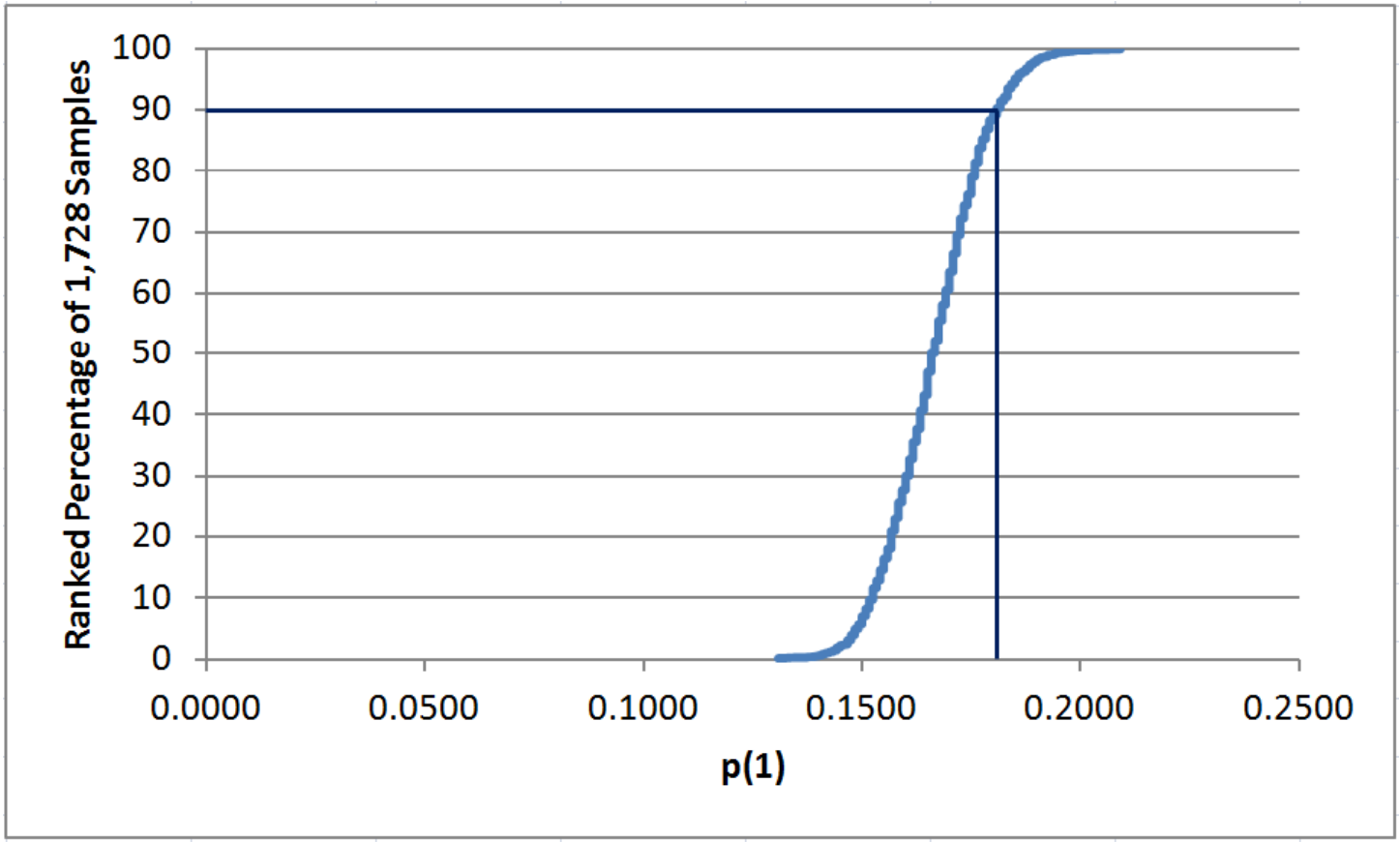


Points for Consideration when Results from Simulation Models are Analysed

Presented by Dave Kirby – Consultant, Persides Consultancy Services Ltd

Content

- What are stochastic models?
- The requirement for many replications from stochastic models.
- Potential for misinterpretation of results from many replications.
- Chaotic results v ranked results.
- Other Techniques:
  - The time element.
  - Ratios.
  - Individual quantity or total cost?
- Summary.

What are Stochastic Models?

- Deterministic models use a single value for a piece of input data; eg, time to return a failed article.

- Stochastic models (or Monte Carlo simulation) use:
  - A distribution to represent a piece of input data; eg, a triangular distribution for a transport time.
  - Each time data is required the distribution is sampled.
  - Many replications of the same scenario but potentially with different values (taken from the distribution) for each input data for each replication;
    - Replications may be conceptualised as ‘parallel universes’.
  - Results from the many replications are analysed statistically.

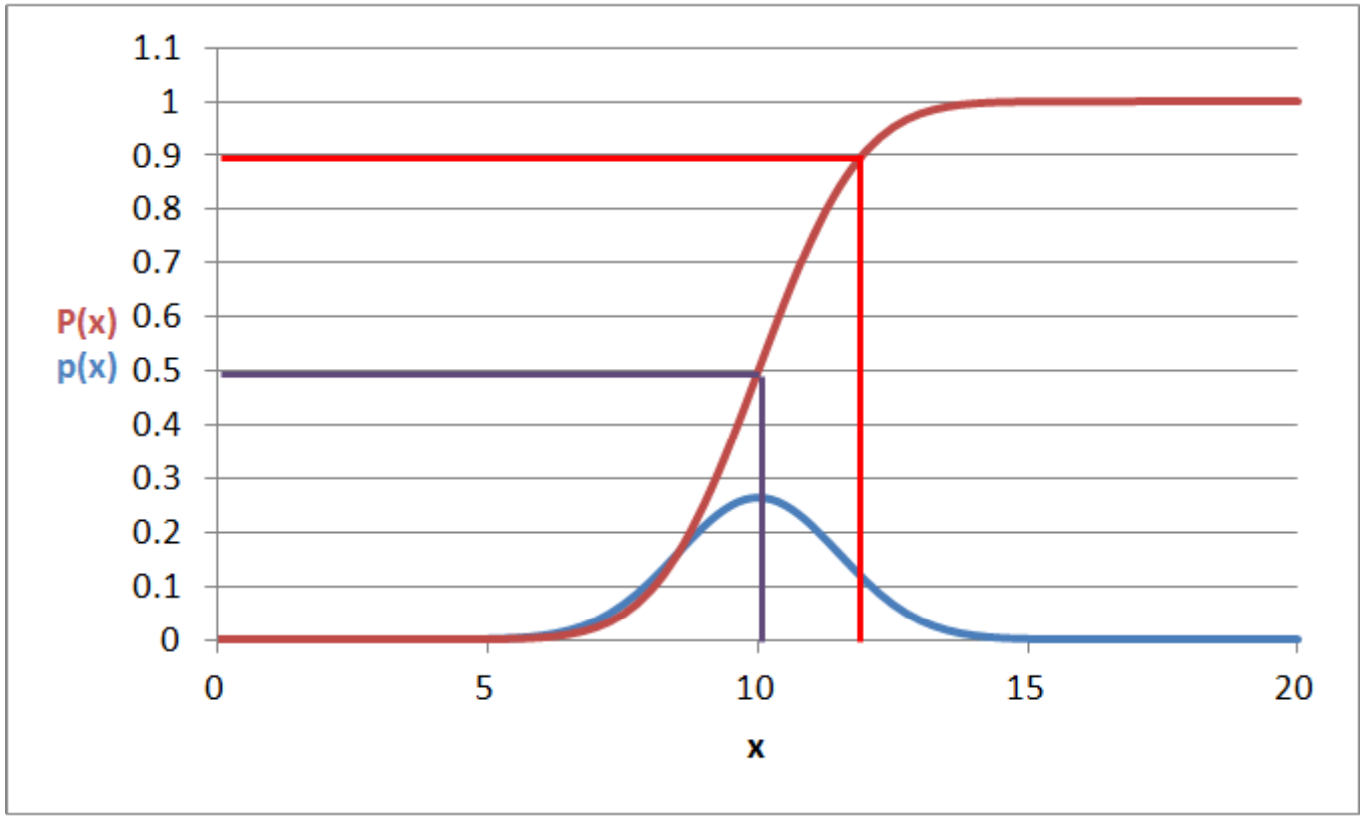
Why Stochastic Modelling?

- The time sequence of events is assessed.
- The ‘butterfly wing’ effect.
  - A minor event early in a simulation may have dramatic consequences that deterministic models could overlook.
- To assess the consequences of potentially complex interactions of known unknowns.
  - Possible but generally very time consuming with deterministic models.
- Results from many replications with potentially different outcomes are required to assess confidence levels.
  - These are unavailable from deterministic models.

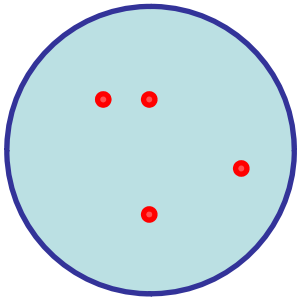
The Requirement for many Replications from Stochastic Models (1 of 4)

Definition of a Percentile

Normal Distribution
Mean = 10
Standard Deviation = 1.5



The Requirement for many Replications from Stochastic Models (2 of 4)



All die throws
- Population
Some die throws
- Samples

x	p(x)	p(x)	P(x)	Score	p(x)	P(x)	x * p(x)
1	1/6	0.1667	0.1667	195	0.1625	0.1625	0.1625
2	1/6	0.1667	0.3333	199	0.1658	0.3283	0.3317
3	1/6	0.1667	0.5000	204	0.1700	0.4983	0.5100
4	1/6	0.1667	0.6667	188	0.1567	0.6550	0.6267
5	1/6	0.1667	0.8333	203	0.1692	0.8242	0.8458
6	1/6	0.1667	1.0000	211	0.1758	1.0000	1.0550
				1,200		Mean x:	3.5317

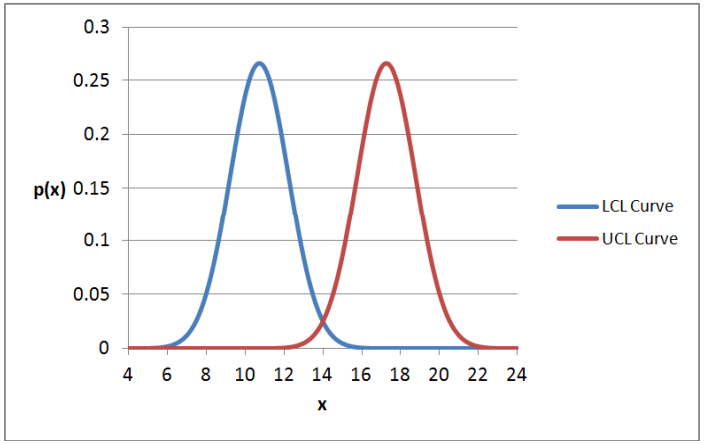
Sample mean = 14
Assumptions:
- Population distribution is Normal
- Population SD = 1.5

Requirement:
- Confidence Limit = 95%

= 1 - α

α / 2 = 2.5%

1 - α / 2 = 97.5%



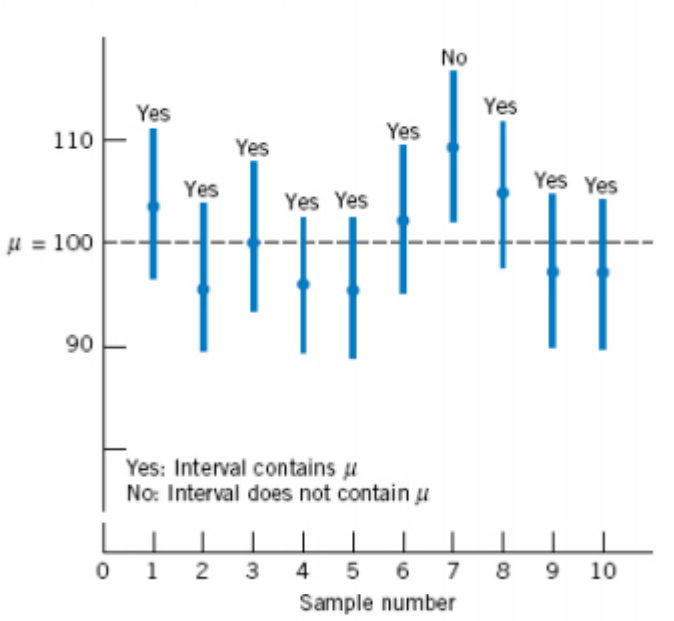
~11 ~17

95% Confidence Interval

The Requirement for many Replications from Stochastic Models (3 of 4)

95% Confidence Intervals (CI)
Samples of size 7
Normally distributed population
- Mean (μ) of 100
- SD of 10

Source - http://www-users.math.umd.edu/~nstrawn/Lecture15.pdf

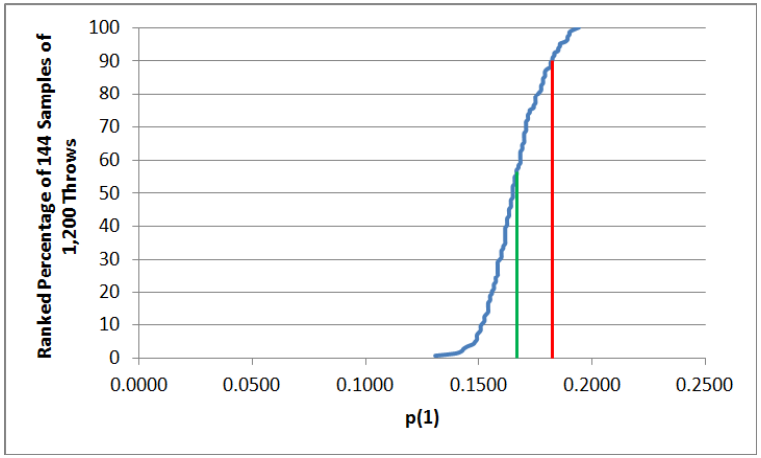


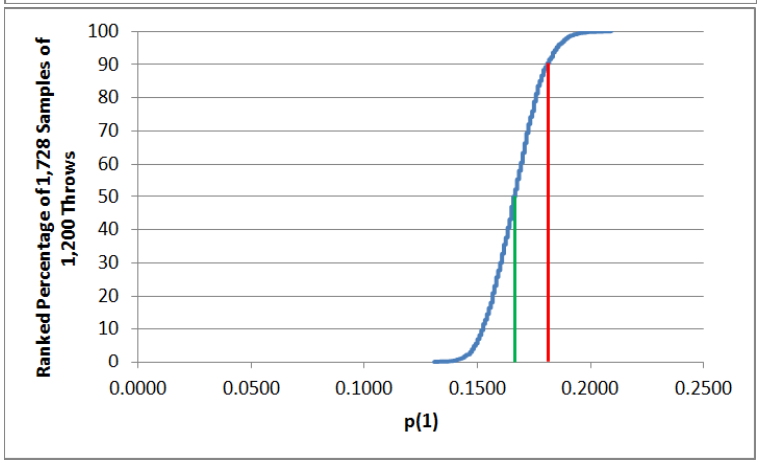
Statistical Confidence:
Specific - 95% of CIs will contain the population mean

General - Confidence Limit % of CIs will contain the population parameter

The Requirement for many Replications from Stochastic Models (4 of 4)

- Often interested in CIs bounded on a single side:
  - Eg, maximum cost with 95% confidence.
- Knowledge of samples but not population.
  - Samples





	Expected	Sample Size		Comparison
		144	1,728	
p(1)	0.1667	0.1650	0.1658	Up
90th Percentile		0.1823	0.1808	Down

- Many samples required; hence, many replications

Information on interpretation of results from many replications in Persides document PER-1928

Potential for Misinterpretation

- PER-1928 addresses 4 potential areas for misinterpretation:
  - Chaotic vs ranked results.
  - The time element.
  - Ratios.
  - Individual quantity or total cost?

Chaotic Results v Ranked Results (1 of 2)

Replication	Item			Replication Sum
	A	B	C	
	0	1	1	2
	0	6	5	11
	2	25	10	37
	4	27	10	41
	5	30	20	55
Median	2	25	10	37
Sum of Median's	37			
90th Percentile	4.6	28.8	16	49.4
Sum of 90th Percentiles	49.4			

Highest

Lowest

Median
Chaotic = 32
Sum = 37
Ranked = 37

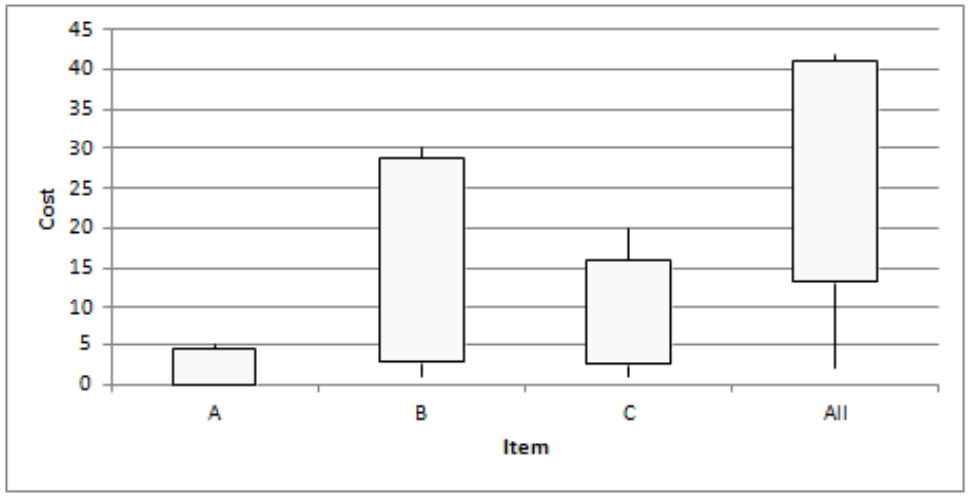
90th Percentile
Chaotic = 41.2
Sum = 49.4
Ranked = 49.4

The sum of all 90th percentile costs >> 90th percentile Whole Life Cost

perside S

Chaotic Results v Ranked Results (2 of 2)

Replication	Item			Replication Sum
	A	B	C	
1	2	30	10	42
2	4	6	20	30
3	5	25	10	40
4	0	27	5	32
5	0	1	1	2



Other Techniques - The Time Element

		Cost			
Item		D			
Month		M	N	O	P
Replication	1	10	6	18	12
	2	20	18	6	5
	3	10	18	6	7
	4	5	18	6	8
	5	1	6	18	6
90^{th} %ile		16.00	18.00	18.00	10.40
90^{th} %ile for Month		16.00	18.00	18.00	10.40
Sum of %iles		62.40			

	Cumulative Cost			
	D			
	M	N	O	P
1	10	16	34	46
2	20	38	44	49
3	10	28	34	41
4	5	23	29	37
5	1	7	25	31
	16.00	34.00	40.00	47.80
	16.00	18.00	6.00	7.80
	47.80			

Consistency between the input data and the percentile values for each period

General rule - derive values from the difference of percentile cumulative values

Other Techniques - Ratios

- PER-1928 shows 3 ways to derive ratios.
- The one recommended way is illustrated here:

		Qty of Demands			
	Period:	1	2	3	4
Replication	1	24	22	30	41
	2	28	39	28	35
	3	31	25	28	37
	4	36	32	21	30
	5	16	35	21	24

		Qty of Demands Satisfied Immediately			
	Period:	1	2	3	4
Replication	1	21	15	16	14
	2	21	23	13	13
	3	27	18	17	17
	4	30	25	14	17
	5	15	24	11	14

		Demand Satisfaction Rate (DSR)			
	Period:	1	2	3	4
Replication	1	0.88	0.68	0.53	0.34
	2	0.75	0.59	0.46	0.37
	3	0.87	0.72	0.61	0.46
	4	0.83	0.78	0.67	0.57
	5	0.94	0.69	0.52	0.58

90th &ile DSR	0.91	0.76	0.64	0.58

General rule - work with the parameter that is required as the output

perside S

Other Techniques - Individual Quantity or Total Cost

- Potential for misinterpretation when the percentile of, for example, the cost of quantities of items or activities is required.
  - Derive percentile quantity then multiply by cost ... or ✗
  - Derive percentile cost? ✓

		Qty of Demands							
Item:		A				B			
Period:		1	2	3	4	1	2	3	4
Replication	1	24	22	30	41	19	22	24	20
	2	28	39	28	35	16	25	18	23
	3	31	25	28	37	20	19	23	26
	4	36	32	21	30	27	22	25	19
	5	16	35	21	24	19	19	23	25

Item:	A	B
Cost per Arising:	24	36

		Cost of Arisings							
Item:		A				B			
Period:		1	2	3	4	1	2	3	4
Replication	1	576	528	720	984	684	792	864	720
	2	672	936	672	840	576	900	648	828
	3	744	600	672	888	720	684	828	936
	4	864	768	504	720	972	792	900	684
	5	384	840	504	576	684	684	828	900

		Cumulative Cost of Arisings											
Item:		A				B				All			
Period:		1	2	3	4	1	2	3	4	1	2	3	4
Replication	1	576	1,104	1,824	2,808	684	1,476	2,340	3,060	1,260	2,580	4,164	5,868
	2	672	1,608	2,280	3,120	576	1,476	2,124	2,952	1,248	3,084	4,404	6,072
	3	744	1,344	2,016	2,904	720	1,404	2,232	3,168	1,464	2,748	4,248	6,072
	4	864	1,632	2,136	2,856	972	1,764	2,664	3,348	1,836	3,396	4,800	6,204
	5	384	1,224	1,728	2,304	684	1,368	2,196	3,096	1,068	2,592	3,924	5,400
90th %ile Cum Cost:		816	1,622	2,222	3,034	871	1,649	2,534	3,276	1,687	3,271	4,642	6,151
90th %ile Cost:		816	806	600	811	871	778	886	742	1,687	1,584	1,370	1,510

General rule - work with the parameter that is required as the output

perside S

Summary

- Some stochastic tool input data describes distributions.
- Output from one replication is a sample (of one) from the population of outcomes.
- Output from many replications is required as input for statistical analysis.
- There is much potential for misinterpretation of output.
  - Eg, the sum of all 90th percentile costs >> 90th percentile Whole Life Cost
- Information on interpretation of results from many replications is in Persides document PER-1928

Questions