KDP-B Range Estimates

Unbiased Range Estimation

Abstract: One of the most common methods for estimating the parameters of a population from data, the method of maximum likelihood (MLE) can yield biased estimators when only a small number of data points are available. We look at one such case and try the method of maximum spacing estimation (MSE).

Will Jarvis, IPAO

This document is for NASA internal distribution and use only.

KDP-B Range Estimates

Unbiased Range Estimation

Abstract: One of the most common methods for estimating the parameters of a population from data, the method of maximum likelihood (MLE) can yield biased estimators when only a small number of data points are available. We look at one such case and try the method of maximum spacing estimation (MSE).

Will Jarvis, IPAO

This document is for NASA internal distribution and use only.

Range for Normal Data

An estimator of Range sometimes reported by Projects is the *Sample Standard Deviation* – a familiar MLE - denoted by s_N and defined as follows:

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2}.$$

This estimator, when applied to a small or moderately sized sample, tends to be too low: it is a <u>biased estimator</u>. The correction (Bessel's correction) is the following:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2},$$

Range for Normal Data

Tall

$$s_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2}.$$

N	N-1/N
2	0.50
3	0.67
4	0.75
5	0.80
6	0.83
7	0.86
8	0.88
9	0.89
10	0.90
100	0.99

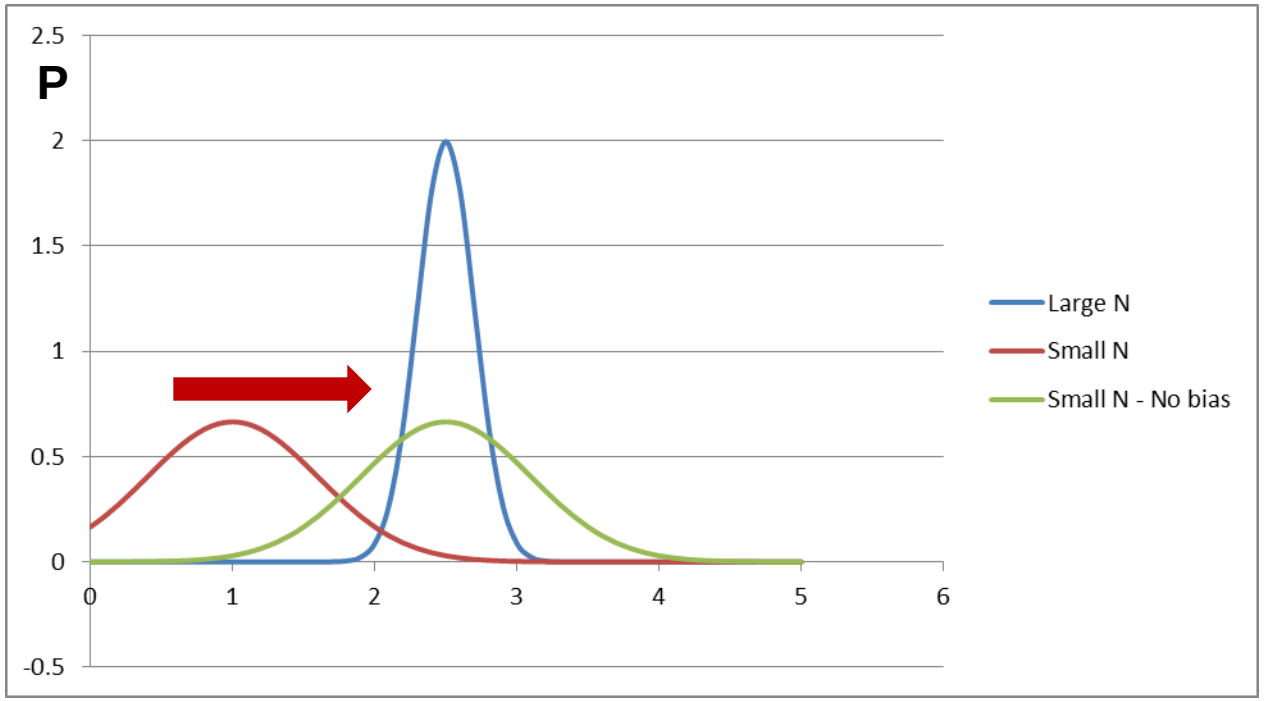
Grande

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2},$$

Large ***N***:

- More precise
- Less biased

Our Goal



We would like to eliminate bias in estimates that are based on small sample sizes.

Range Estimates

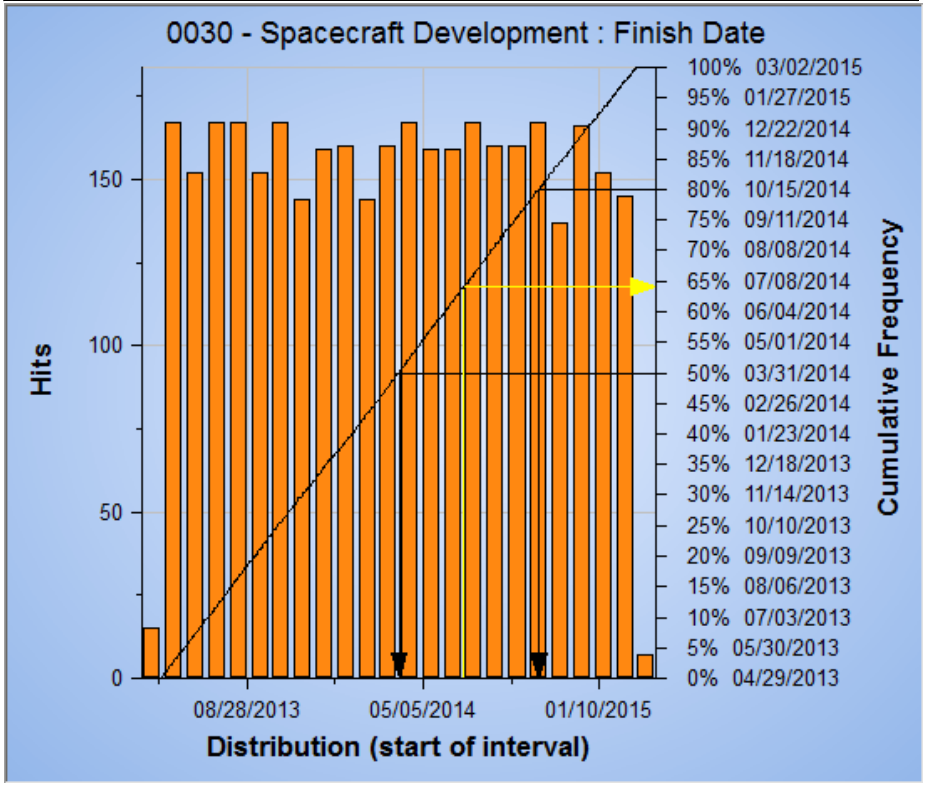
Estimate of distribution's *Range* is biased. Small sample sizes introduce more bias than large sample sizes.

Element Distribution for Spacecraft

- Distribution based on analogous historic RSDO actuals
- Utilized uniform distribution with low as 24 months and high as 47 months
  - Mean is 36 months
- Spacecraft planned duration is 37 months (no change based on revised LRD)

Spacecraft	Months***
NPP	36
GLAST (RSDO)	47
Swift (RSDO)	32
Coriolus (RSDO)	24
ICESat	41
Average	36

*** Schedule is ATP to bus complete
Data received from IPAO (backup material available)



Range for Schedule Duration

Suppose $\{x_{(1)}, \ldots, x_{(n)}\}$ is the ordered sample from a uniform distribution $U(a,b)$ with unknown endpoints a and b. Sometimes the endpoints of the sample are used to estimate the endpoints of the population from which they were drawn. This is biased…

Unbiased estimators for the endpoints of the uniform distribution are the following:

$$\hat{a} = \frac{nx_{(1)} - x_{(n)}}{n-1}, \quad \hat{b} = \frac{nx_{(n)} - x_{(1)}}{n-1}.$$

For 2 Data Points:

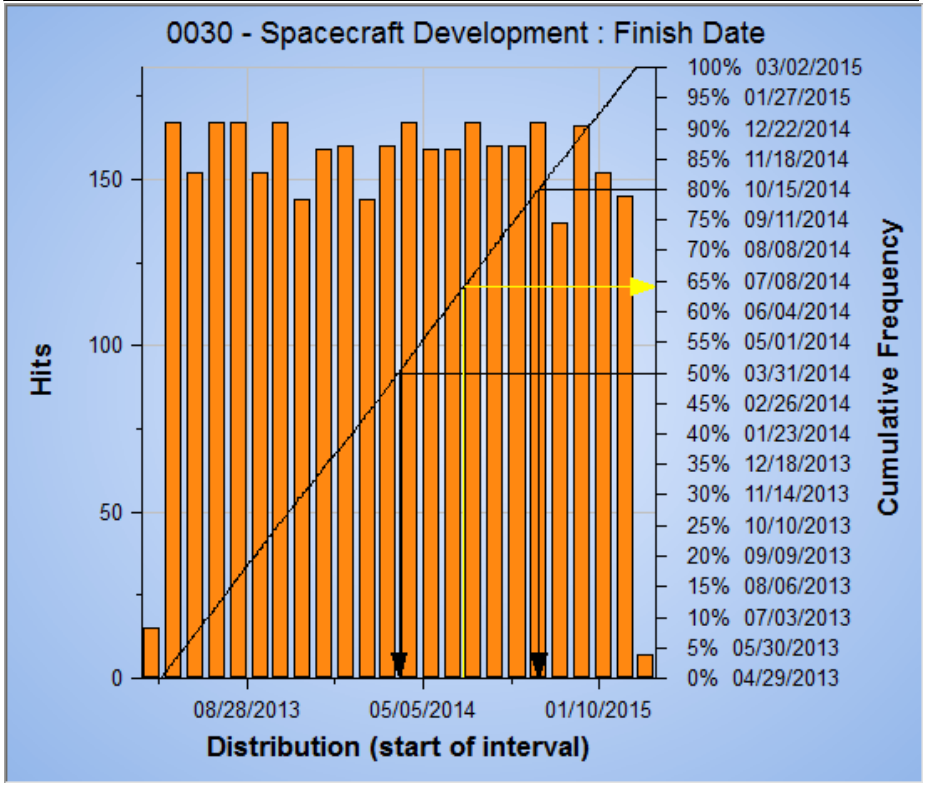
Unbiased Lower Limit = 2 * Smallest Value – Largest Value
Unbiased Upper Limit = 2 * Largest Value – Smallest Value

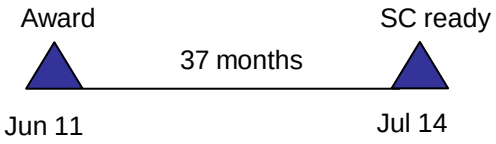
Element Distribution for Spacecraft

- Distribution based on analogous historic RSDO actuals
- Utilized uniform distribution with low as 24 months and high as 47 months
  - Mean is 36 months
- Spacecraft planned duration is 37 months (no change based on revised LRD)

Spacecraft	Months***	
NPP	36	
GLAST (RSDO)	47	<= High
Swift (RSDO)	32	
Coriolus (RSDO)	24	<= Low
ICESat	41	
Average	36	

*** Schedule is ATP to bus complete
Data received from IPAO (backup material available)





Small Sample Bias Correction (N=5)
Low: (5*24 – 47) / 4 = 18 months
High: (5*47 – 24) / 4 = 53 months

Element Distribution for Instrument

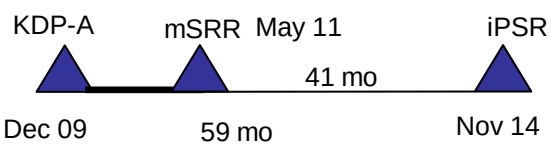
- Distribution based on analogous historic instruments
- Utilized uniform distribution with low as 35 months and high as 66 months
  - Mean is 50 months
- ATLAS planned duration is 59 months (increased 4 months based on revised LRD schedule)

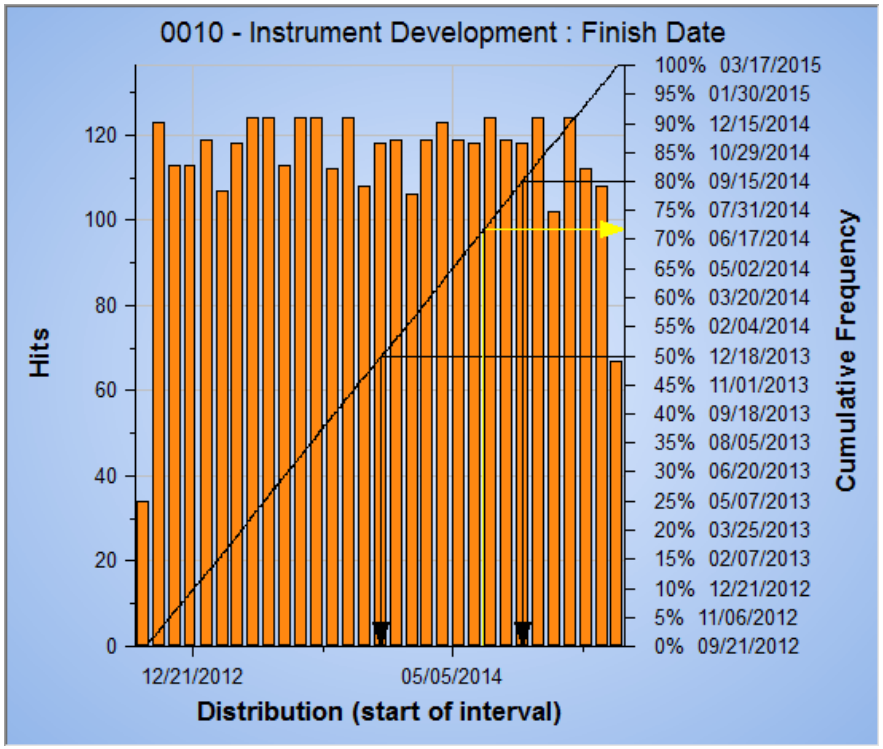
Instrument	Months	KDP A - Devilery
GLAS	66	<= High
MLA	48	
MOLA	35	<= Low
LOLA	42	
CALIOP	59	Appropriate Data Points?
Average	**50**	

Source: Provided by ICESAT via Aerospace Corp.

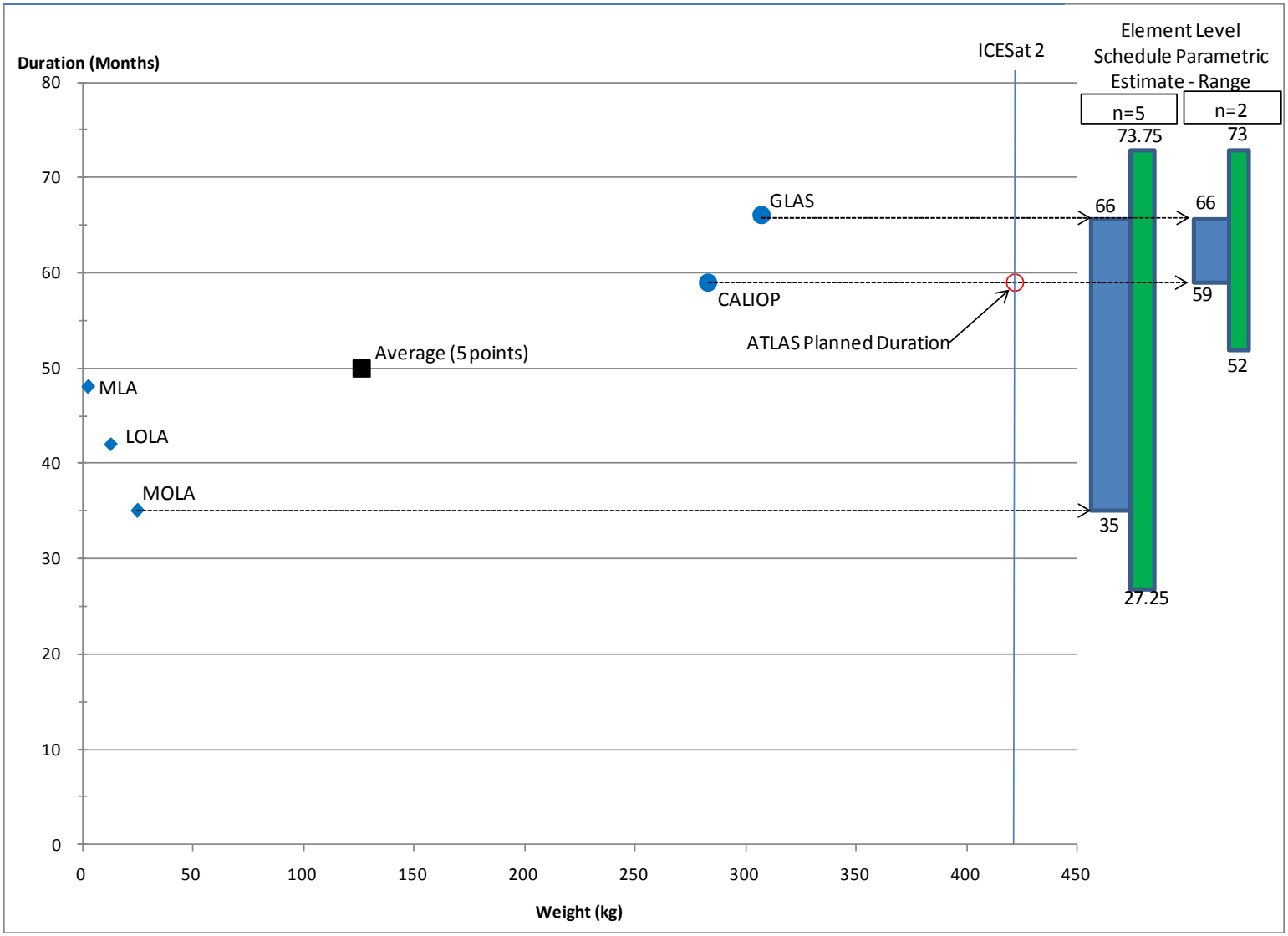
Small Sample Bias Correction (N=5)
Low: (5*35 – 66) / 4 = 27 months
High: (5*66 – 35) / 4 = 74 months

Small sample bias correction (N=2)
Low: (2*59 – 66) = 52 months
High: (2*66 – 59) = 73 months





Bias in Instrument Schedule Estimate (SRB)



Estimate of Range - Based on 2 Data Points

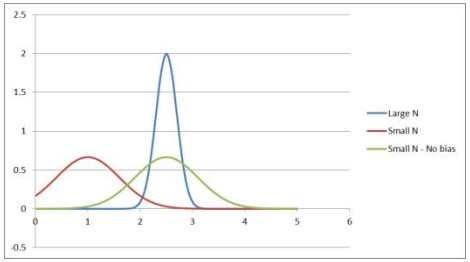
1,000 trials

actual LOW	actaual HIGH
0	**5**

sample 1	sample 2	ordered 1	ordered 2	Est LOW	Est HIGH
3.35	3.34	3.34	3.35	3.33	3.36
2.83	2.96	2.83	2.96	2.69	3.10
3.85	4.58	3.85	4.58	3.13	5.30
3.61	3.93	3.61	3.93	3.28	4.25
3.33	3.37	3.33	3.37	3.30	3.40
3.07	4.24	3.07	4.24	1.90	5.42
1.23	3.80	1.23	3.80	-1.33	6.37
0.12	1.19	0.12	1.19	-0.96	2.27
3.74	4.43	3.74	4.43	3.06	5.12
0.70	2.88	0.70	2.88	-1.48	5.07
4.10	1.84	1.84	4.10	-0.43	6.37
1.28	0.58	0.58	1.28	-0.13	1.99
1.67	2.39	1.67	2.39	0.95	3.11
2.25	1.69	1.69	2.25	1.13	2.81
2.86	1.09	1.09	2.86	-0.68	4.63
1.47	2.20	1.47	2.20	0.73	2.93
4.93	1.76	1.76	4.93	-1.42	8.11
2.03	1.33	1.33	2.03	0.63	2.73
1.09	0.25	0.25	1.09	-0.60	1.93
4.42	1.11	1.11	4.42	-2.21	7.73

Average (Biased)		*Average (Unbiased)*	
Low	High	Low	High
1.64	**3.31**	**-0.03**	**4.98**

Not Bad!



Conclusion: Bias is eliminated. Precision is still weak.

Google “Maximum Spacing Estimation” for Theory.

See @RISK Spreadsheet for Demo.

- File: *Small Sample Bias in Project Schedule Estimate.xlsx*