Agribusiness Analysis and Forecasting

Parameter Estimation and Validation

Henry Bryant

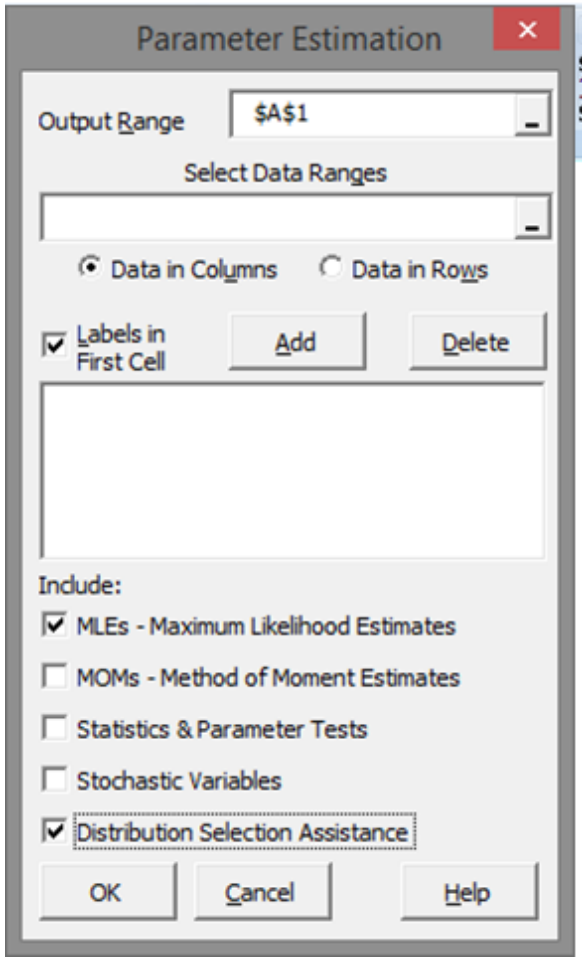
Texas A&M University

Overview

- Parametric distributions have fixed functional forms.
- Estimate their parameters with Simetar's Theta icon.
- Simulate candidate distributions and pick the distribution that, in some sense, most closely reflects the historical observations.
- Perform simulation analysis using the chosen distribution.

Parameter Estimation with Theta

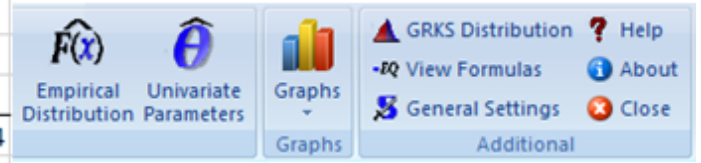
Select MLE and "Stochastic Variables" options



Parameter Estimation

Maximum Likelihood Estimates (MLEs)

Distributi	Parameter	Test
Beta	α;α>0, A≤x≤B	0.464544
	β;β>0	0.75791
Double Ex	μ; -∞<μ<∞, -∞<x<∞	12
	σ; σ>0	8
Exponenti	α; -∞<α<∞, ≤x<∞	2
	β; β>0	11.4
Gamma	α; α>0, 0≤x<∞	1.75291
	β; β>0	7.644433
Inverse G	μ; μ>0, 0≤x<∞	13.4
	σ; σ>0	0.27
Logistic	μ; -∞<μ<∞, -∞<x<∞	12.56371
	σ; σ>0	5.593188
Log-Log	μ; -∞<μ<∞, -∞<x<∞	8.968149
	σ; σ>0	7.253337
Log-Logist	μ; -∞<μ<∞, 0≤x<∞	1.947158
	σ; σ>0	10.36335
Lognorma	μ; -∞<μ<∞, 0≤x<∞	2.283671
	σ; σ>0	0.84777
Normal	μ; -∞<μ<∞, -∞<x<∞	13.4
	σ; σ>0	9.656086

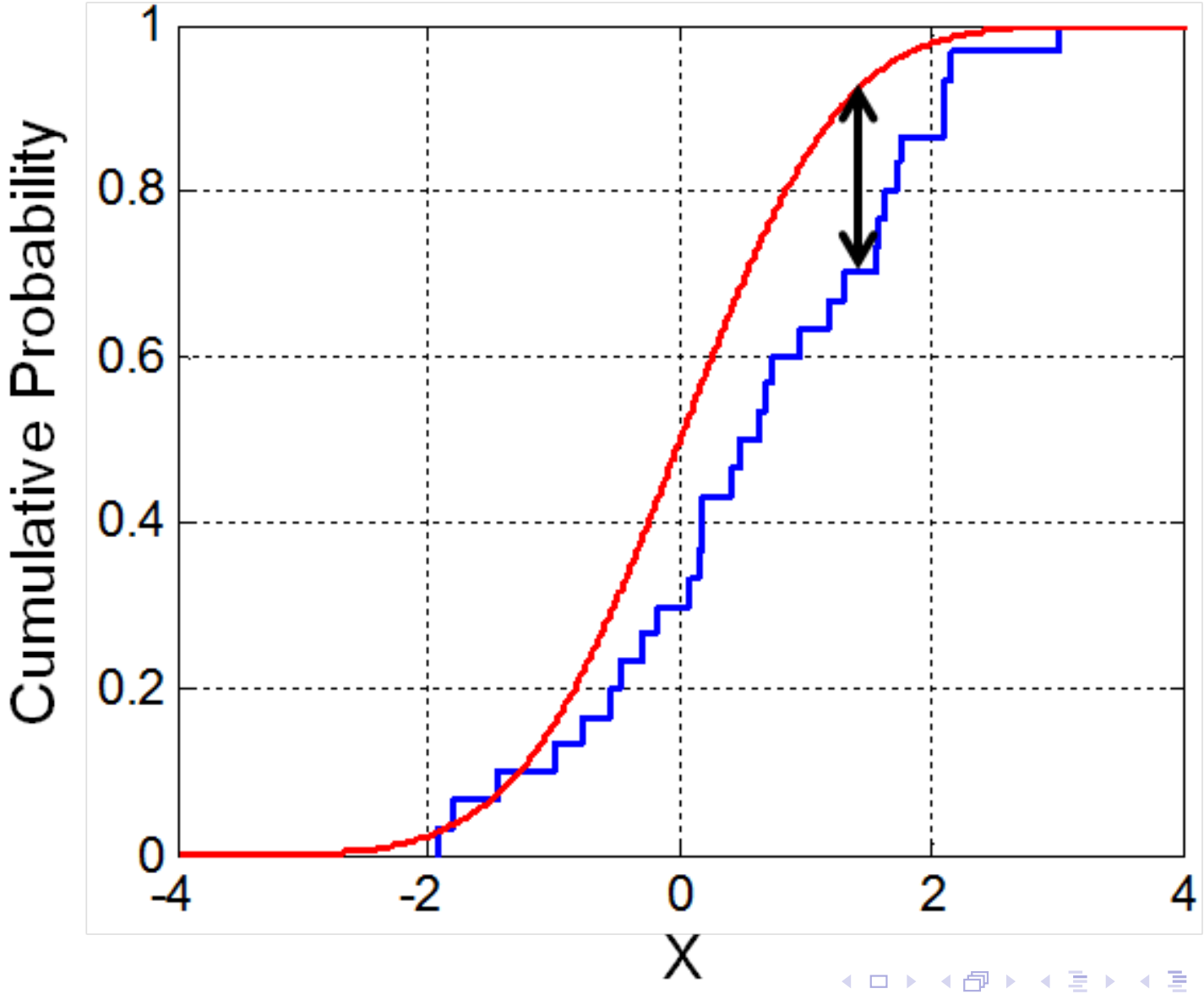


Comparing distributions

- CDFDEV is a Simetar function to compare the CDFs of two data samples
- CDFDEV calculates the integral between two distributions with a penalty for the two distributions being different.

$$\int_{-\infty}^{\infty} \left(F_1(x) - F_2(x)\right)^2 + w(x)dx$$

Comparing distributions



Calculating CDFDEV

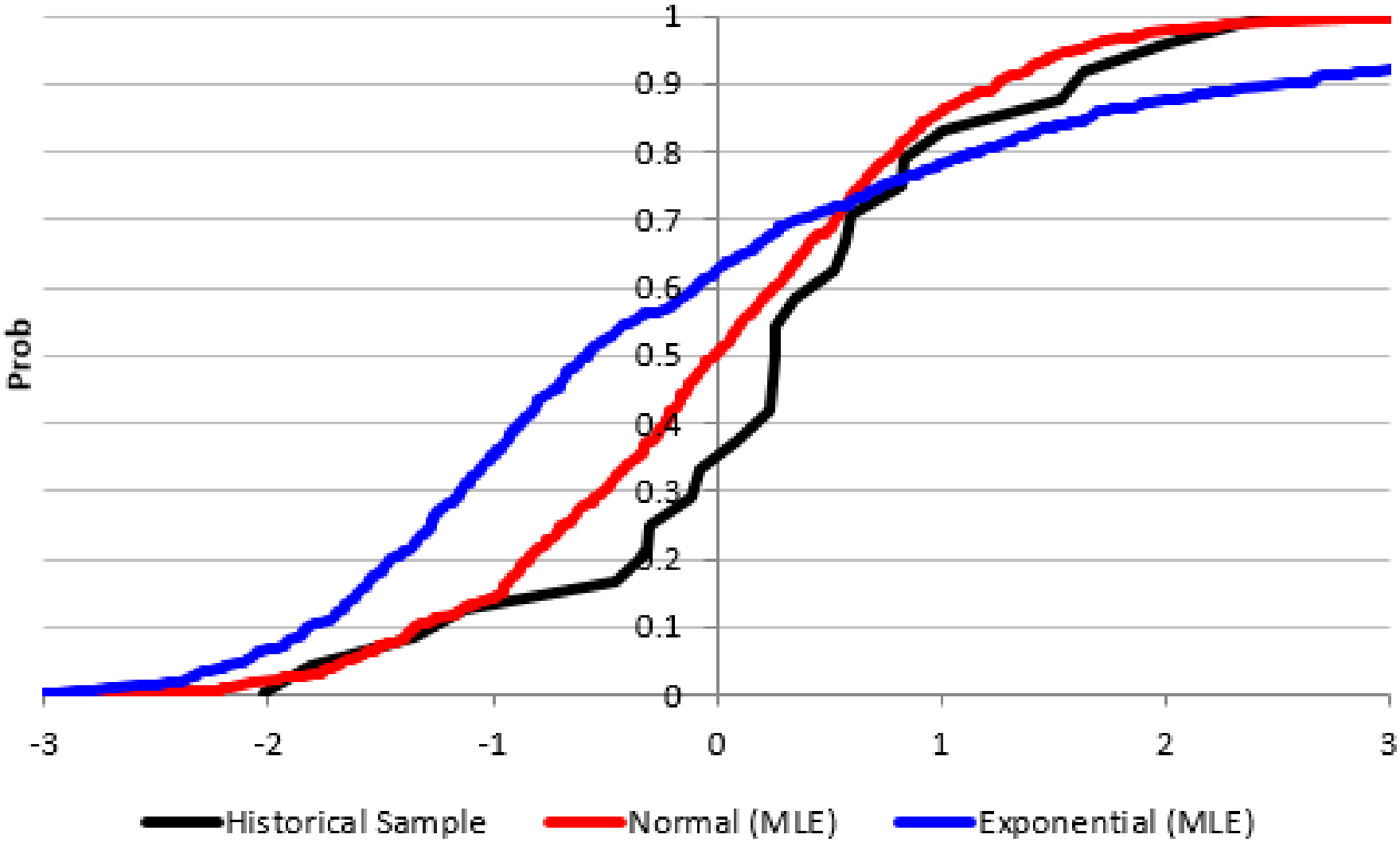
- Create simulated samples from candidate distributions
- Use CDFDEV to compare those simulated samples to the historical data sample: =CDFDEV(sample_1, sample_2)
- Select the distribution that has the lowest CDFDEV value.

Random Variables (MLE)	
Distribution	Random Va
Beta (MLE)	2.0697062
Double Exponential (ML	2.1953671
Exponential (MLE)	1.8618588
Gamma (MLE)	2.1173946
Inverse Gaussian (MLE	2.1360436
Logistic (MLE)	2.1454888
Log-Log (MLE)	2.0783888
Log-Logistic (MLE)	2.1389988
Lognormal (MLE)	2.1076268
Normal (MLE)	2.1367503
Pareto (MLE)	1.8124891
Uniform (MLE)	2.0971311
Weibull (MLE)	2.1500771
Binomial (MLE)	2
Geometric (MLE)	2
Poisson (MLE)	2
EMP	2

Distribution	CDFDEV
Beta (MLE)	0.0178771
Double Exponential (MLE)	0.3655717
Exponential (MLE)	2.7087809
Gamma (MLE)	0.067896
Inverse Gaussian (MLE)	0.1142559
Logistic (MLE)	0.1606926
Log-Log (MLE)	0.4418301
Log-Logistic (MLE)	0.3453194
Lognormal (MLE)	0.100735
Normal (MLE)	0.0497376
Pareto (MLE)	79.214829
Uniform (MLE)	0.0324191
Weibull (MLE)	0.0811444
Binomial (MLE)	0.9668741
Geometric (MLE)	59.505363
Poisson (MLE)	6.3777462
EMP	0.0003111

Comparing distributions

Another tool: visually compare emprical CDFs



What is the Next Step?

- After choosing a parametric distribution...
- It's a good idea to validate that the characteristics of the simulated data match those of the original historical data.
- Use statistical tests to check that the means variances are not significantly different from one another.
- Check if the minimum and maximum values are realistic.
- Can also visually check the shape of the CDF and PDF.

Statistical Tests for Validation

Student t test

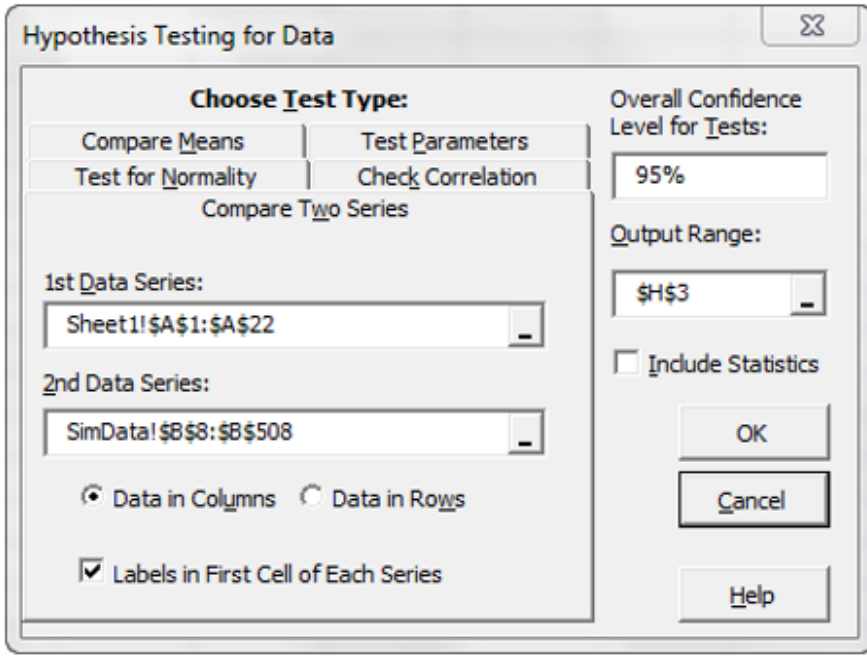
- H_o: Historical Mean $=$ Simulated Mean.
- H_a: Historical Mean $\neq$ Simulated Mean.

F test

- H_o: Historical Variance $=$ Simulated Variance
- H_a: Historical Variance $\neq$ Simulated Variance.

Validation Tests in Simetar

- Compare Two Series: Historical Data vs. Simulated Values
  - 1st Data Series is history
  - 2nd Data Series is simulated
- Simetar Icon is



Distribution Comparison of Normal Corn price & Corn price				
Confidence Level		95.0000%		
	Test Value	Critical Val	P-Value	
2 Sample t Test	0.00	2.69	0.999	*Fail to Reject the Ho that the Means are Equal*
F Test	1.00	1.90	0.437	*Fail to Reject the Ho that the Variances are Equal*