Learning Curve Analysis

- Developed as a tool to estimate the recurring costs in a production process
  - Recurring costs: those costs incurred on each unit of production. There are no learning curves associated with overhead costs, just manufacturing costs.

- Dominant factor in learning theory is direct labor
  - Based on the common observation that as a task is accomplished several times, it can be completed in shorter periods of time
    - “Each time you perform a task, you become better at it and accomplish the task faster than the previous time”
Cost Progress

- Cost progress (or “getting better at the task”) comes from production workers’ learning their tasks better, but also from:
  - Re-design of product for lower-cost production
  - Improved production facility and production lines
  - Better layout / better efficiencies
  - Management improvements
  - Lower-cost suppliers
  - Better “make or buy” decisions (produce in-house or out-source)

Learning Theory

![Graph showing decreasing unit cost with increasing quantity]

Unit Cost

$0.00 $20.00 $40.00 $60.00 $80.00 $100.00 $120.00

Qty

1 3 5 7 9 11 13 15
Log-Log Plot of Linear Data

Linear Plot of Log Data
Learning Theory

• Two variations:
  – Cumulative Average Theory
  – Unit Theory

  – Note: Can only use one theory throughout an estimation. You must be consistent! Pick one and stick with it. We will discuss why soon.

Cumulative Average Theory

• “If there is learning in the production process, the cumulative average cost of some doubled unit equals the cumulative average cost of the un-doubled unit times the slope of the learning curve”

• Historical Facts: Described by T. P. Wright in 1936
  – Based on examination of WW I aircraft production costs

• Aircraft companies and DoD were interested in the regular and predictable nature of the reduction in production costs that Wright observed
  – Implied that a fixed amount of labor and facilities would produce greater and greater quantities in successive periods
Unit Theory

“If there is learning in the production process, the cost of some doubled unit (say, unit #100) equals the cost of the undoubled unit (= unit #50) times the slope of the learning curve”

- Credited to J. R. Crawford in 1947
  - Led a study of WWII airframe production commissioned by USAF to validate learning curve theory

Basic Concept of Unit Theory

- As the quantity of units produced doubles, the cost** to produce a unit is decreased by a constant percentage.
  - For an 80% learning curve, there is a 20% decrease in cost each time that the number of units produced doubles
    - the cost of unit 2 is 80% of the cost of unit 1
    - the cost of unit 4 is 80% of the cost of unit 2
    - the cost of unit 8 is 80% of the cost of unit 4, etc.

- One of the few times when 80% is better than 90%. Why is 80% better?

** The Cost of a unit can be expressed in dollars, labor hours, or other units of measurement.
80% Unit Learning Curve

Unit Theory

- Defined by the equation $Y_x = A \times x^b$

  where
  
  $Y_x$ = the cost of unit $x$ (dependent variable)
  $A$ = the theoretical cost of unit 1 (a.k.a. T1)
  $x$ = the unit number (independent variable)
  $b$ = a constant representing the slope
    (slope = $2^b$)

- This is a Power model (from Ch 9, next slide)
Some Linear Transformations

<table>
<thead>
<tr>
<th>Unit Space</th>
<th>Model</th>
<th>Log Space</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><strong>Logarithmic</strong>&lt;br&gt;$y = a + b \ln x$</td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Graph" /></td>
<td><strong>Exponential</strong>&lt;br&gt;$y = a e^{b x}$&lt;br&gt;$\ln y = \ln a + b x$</td>
<td><img src="image4.png" alt="Graph" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Graph" /></td>
<td><strong>Power</strong>&lt;br&gt;$y = a x^{b}$&lt;br&gt;$\ln y = \ln a + b \ln x$</td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

### Learning Parameter

- In practice, $-0.5 \leq b \leq -0.05$
  - corresponds roughly with learning curves between 70% (entirely manual operations) and 96% (slope = $2^b$)
  - learning parameter largely determined by the type of industry and the degree of automation
  - for $b = 0$, the equation simplifies to $Y = A$, which means any unit costs the same as the first unit. In this case, the learning curve is a horizontal line and there is no learning. Not good in the business world! This is referred to as a 100% learning curve....

9 - 14
Learning Curve Slope versus the Learning Parameter

“As the number of units doubles, the unit cost is reduced by a constant percentage which is referred to as the slope of the learning curve”

Cost of unit 2n = (Cost of unit n) x (Slope of learning curve)

\[
\text{Slope of learning curve} = \frac{\text{Cost of unit 2n}}{\text{Cost of unit n}} = \frac{A(2n)^b}{A(n)^b} = 2^b
\]

Taking the natural log of both sides:

\[
\ln (\text{slope}) = b \times \ln (2)
\]

\[
b = \ln (\text{slope})/\ln (2)
\]

For a typical 80% learning curve:

\[
\ln (.8) = b \times \ln (2)
\]

\[
b = \ln (.8)/\ln (2) = -.3219 = \text{slope coefficient}
\]

General Guidelines for Slopes

- If an operation is 75% manual and 25% automated, slopes are generally in the 80% vicinity.
- If it is 50% manual and 50% automated, slopes can be expected to be about 85%.
- If it is 25% manual and 75% automated, slopes can be expected to about 90%.

- The average slope for the aircraft industry is about 85%. But departments within can vary greatly from that.

- Shipbuilding slopes run between 80 and 85%.
General Guidelines for Slopes

• Assuming repetitive operations within an industry, typical slopes include:

  – Electrical: 75-85%
  – Electronics: 90-95%
  – Machining: 90-95%
  – Welding: 88-92%

Slope and 1st Unit Cost

• To use a learning curve for a cost estimate, a slope and 1st unit cost are required

  – Slope may be derived from analogous production situations, industry averages, historical slopes for the same production site, or historical data from previous production quantities

  – 1st unit costs may be derived from engineering estimates, CERs, or historical data from previous production quantities
Notes on A (aka T1)

• T1 is the *theoretical cost of unit 1*, or the cost where the learning curve crosses the Y-axis at unit n=1.

• As in Unit Theory, the *theoretical first unit cost* is usually different than the *actual* cost. This is because the learning curve is a quantitative model, representing the general behavior of the actual cost in an aggregate fashion, but one that does not follow each data point exactly.

Slope and 1st Unit Cost from Historical Data

• When historical production data is available, slope and first unit cost can be calculated by using the learning curve equation:

\[ Y_x = A \times x^b \]

  – take the natural log of both sides:

  \[ \ln(Y_x) = \ln(A) + b \ln(x) \]

  – rewrite as \( Y' = A' + b X' \) and solve for \( A' \) and \( b \) using simple linear regression

  – \( A = \exp(A') \)

  – no transformation for \( b \) required
Example #1

- Given the following historical data, find the Unit Theory learning curve equation which describes this production environment. Use this equation to predict the cost (in hours) of the 150th unit and find the slope of the curve (Note: same numbers as in Chapter 9.)

<table>
<thead>
<tr>
<th>Unit #</th>
<th>Hours</th>
<th>ln(Unit #)</th>
<th>ln(Hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>60</td>
<td>1.6094</td>
<td>4.0943</td>
</tr>
<tr>
<td>12</td>
<td>45</td>
<td>2.4849</td>
<td>3.8067</td>
</tr>
<tr>
<td>35</td>
<td>32</td>
<td>3.5553</td>
<td>3.4657</td>
</tr>
<tr>
<td>75</td>
<td>26</td>
<td>4.3175</td>
<td>3.2581</td>
</tr>
<tr>
<td>125</td>
<td>21</td>
<td>4.8283</td>
<td>3.0445</td>
</tr>
</tbody>
</table>

Original Data Scatter Plot
Y vs. X (note nonlinear)
Transformed Data Scatter Plot
\[
\ln(Y) \text{ vs. } \ln(X): \text{ (now linear)}
\]
\[
y = -0.3192x + 4.6062
\]

Example #1 (cont)

- Or, using the Regression Add-In in Excel...

<table>
<thead>
<tr>
<th>SUMMARY OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression Statistics</td>
</tr>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>p-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.606188433</td>
<td>0.02915617</td>
<td>157.982287</td>
<td>5.5922E-07</td>
<td>4.513370488</td>
</tr>
<tr>
<td>ln(hours)</td>
<td>-0.319218362</td>
<td>0.008190526</td>
<td>-38.97406994</td>
<td>3.71633E-05</td>
<td>-0.345284277</td>
</tr>
</tbody>
</table>

ICEAA 2016 Bristol – TRN 03
Example #1

The equation which describes this data can be written:

\[ Y_x = A x^b \]

\[ A = e^{4.606} = 100.083 \]

\[ b = -0.3192 = \text{slope coefficient} \]

\[ Y_x = 100.083(x)^{-0.3192} \]

Solving for the cost (hours) of the 150th unit…

\[ Y_{150} = 100.083(150)^{-0.3192} \]

\[ Y_{150} = 20.22 \text{ hours} \]

Slope of this learning curve = \(2^b\)

\[ \text{slope} = 2^{-0.3192} = .8015 = 80.15\% \]

Example Conclusions

- Time to complete first unit: 100.083 hours
- Time to complete 150th unit: 20.22 hours
- So we have gotten “better” by approximately 80 hours!
Estimating Lot Costs

• After finding the learning curve which best models the production situation, the estimator must now use this learning curve to estimate the cost of future units. But rarely are we asked to estimate the cost of just one unit! Rather, we usually need to estimate lot costs.

– This is calculated using a cumulative total cost equation:

\[ CT_N = A(1)^b + A(2)^b + \cdots + A(N)^b = A \sum_{x=1}^{N} x^b \]

– where \( CT_N \) = the cumulative total cost of \( N \) units

– \( CT_N \) may be approximated using the following equation:

\[ CT_N \approx \frac{A(N)^{b+1}}{b+1} \]

Estimating Lot Costs

• Compute the cost (in hours) of the first 150 units from the Example #1:

\[ CT_{150} = A * (150)^{b+1} = (100.083)(150)^{-0.3192 + 1} = 4454.75 \text{ hrs} \]

• To compute the total cost of a specific lot with first unit # \( F \) and last unit # \( L \):

\[ CT_{F,L} = A \left[ \sum_{x=1}^{L} x^b - \sum_{x=1}^{F-1} x^b \right] \]

• and this equation is approximated by:

\[ CT_{F,L} \approx \frac{AL^{b+1}}{b+1} = \frac{A(F-1)^{b+1}}{b+1} \]
Estimating Lot Costs: Example

• Compute the cost (in hours) of the lot containing units 26 through 75 using the numbers from Example #1:

\[ CT_{F,L} \approx \frac{AL^{b+1}}{b+1} - \frac{A(F-1)^{b+1}}{b+1} \]

CT\textsubscript{26,75} = \frac{100.083 \cdot (75)^{-0.3192+1}}{-0.3192 + 1} - \frac{100.083 \cdot (26-1)^{-0.3192+1}}{-0.3192 + 1} = 1463.56 \text{ hrs} 

Fitting a Curve Using Lot Data

• Cost data is generally reported for production lots (i.e., lot cost and units per lot), not individual units

  – Lot data must be adjusted since learning curve calculations require a unit number and its associated unit cost (i.e., needs an x and y)

  – Unit number and unit cost for a lot are represented by algebraic lot midpoint (LMP) and average unit cost (AUC)
Lot Cost Example: How Data is Usually Received

<table>
<thead>
<tr>
<th>Lot #</th>
<th># Units</th>
<th>First Unit</th>
<th>Last Unit</th>
<th>Lot Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>1</td>
<td>50</td>
<td>$10M</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>51</td>
<td>100</td>
<td>$8M</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>101</td>
<td>200</td>
<td>$14M</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>201</td>
<td>250</td>
<td>$6M</td>
</tr>
</tbody>
</table>

So How Do We Fit that Data to a Learning Curve??

- Lot Midpoint (LMP) = x-axis
- Average Unit Cost (AUC) = y-axis
Fitting a Curve Using Lot Data

• The Algebraic Lot Midpoint (LMP) is defined as the theoretical unit whose cost is equal to the average unit cost for that lot on the learning curve.
• Calculation of the exact LMP is an iterative process. If learning curve software is unavailable, solve by approximation:
• For the first lot (the lot starting at unit 1):
  – If lot size < 10, then LMP = Lot Size/2
  – If lot size ≥ 10, then LMP = Lot Size/3
• For all subsequent lots:

\[
LMP = \frac{F + L + 2\sqrt{FL}}{4}
\]

where

F = the first unit number in a lot, and
L = the last unit number in a lot.

Fitting a Curve Using Lot Data

• The LMP then becomes the independent variable (X) which can be transformed logarithmically and used in our simple linear regression equations to find the learning curve for our production situation.
• The dependent variable (Y) to be used is the AUC which can be found by:

\[
AUC = \frac{Total\ Lot\ Cost}{Lot\ Size}
\]

• The dependent variable (Y) must also be transformed logarithmically before we use it in the regression equations. We then regress ln(AUC) vs ln(LMP).
Unit Theory Lot Cost Example

- Given the following historical production data on a tank turret assembly, find the Unit Theory Learning Curve equation which best models this production environment and estimate the cost (in man-hours) for the seventh production lot of 75 assemblies which are to be purchased in the next fiscal year.

<table>
<thead>
<tr>
<th>Lot #</th>
<th>Lot Size</th>
<th>Cost (man-hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>36,750</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>19,000</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>90,000</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>39,000</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>60,000</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>in process, no data available</td>
</tr>
</tbody>
</table>

Solution

<table>
<thead>
<tr>
<th>Lot #</th>
<th>Lot Size</th>
<th>Cost</th>
<th>Cum Qty</th>
<th>LMP</th>
<th>AUC</th>
<th>ln (LMP)</th>
<th>ln (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>36,750</td>
<td>15</td>
<td>5.00</td>
<td>2450</td>
<td>1.609</td>
<td>7.804</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>19,000</td>
<td>25</td>
<td>20.25</td>
<td>1900</td>
<td>3.008</td>
<td>7.550</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>90,000</td>
<td>85</td>
<td>51.26</td>
<td>1500</td>
<td>3.937</td>
<td>7.313</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>39,000</td>
<td>115</td>
<td>99.97</td>
<td>1300</td>
<td>4.605</td>
<td>7.170</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>60,000</td>
<td>165</td>
<td>139.42</td>
<td>1200</td>
<td>4.938</td>
<td>7.090</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>?</td>
<td>215</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ b = -0.217 \]
\[ A' = 8.17 \quad \Rightarrow \quad A = 3533.22 \]
\[ \text{slope} = 2^b = 2^{-0.217} = 0.8604 = 86.04\% \]

The Unit Learning Curve equation:

\[ Y_x = 3533.22x^{-0.217} \]
Regression Printout of Ln(AUC) vs Ln(LMP)

<table>
<thead>
<tr>
<th>SUMMARY OUTPUT</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple R</td>
<td>0.997900729</td>
<td></td>
</tr>
<tr>
<td>R Square</td>
<td>0.995805865</td>
<td></td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.994407819</td>
<td></td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.021841242</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>ANOVA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>df</td>
<td>SS</td>
</tr>
<tr>
<td>Regression</td>
<td>1</td>
<td>0.33978808</td>
</tr>
<tr>
<td>Residual</td>
<td>3</td>
<td>0.00143112</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>0.3412192</td>
</tr>
<tr>
<td>Coefficients</td>
<td>Standard Error</td>
<td>t Stat</td>
</tr>
<tr>
<td>Intercept</td>
<td>8.170231838</td>
<td>0.030986711</td>
</tr>
<tr>
<td>ln(LMP)</td>
<td>-0.216840315</td>
<td>0.008124811</td>
</tr>
</tbody>
</table>

Solution

- To estimate the cost (in hours) of the tank turret assembly 7th production lot:
  - Note: the units included in the 7th lot are 216 – 290.

\[
CT_{F,L} = \frac{AL^{b+1}}{b+1} - \frac{A(F-1)^{b+1}}{b+1}
\]

\[
A = 3533.22
\]

\[
F = 216
\]

\[
L = 290
\]

\[
b = -0.217
\]

\[
CT_{216,290} = \frac{(3533.22)(290)^{-0.217+1}}{-0.217 + 1} - \frac{(3533.22)(216 - 1)^{-0.217+1}}{-0.217 + 1} = 79,866.7 \text{ hrs}
\]
**Cumulative Average Learning Curves**

ICEAA 2016 International Training Symposium
Bristol, 17th to 20th October 2016

Alan R Jones
Estimata Limited
Promoting TRACeability in Estimating

Wright’s Cumulative Average Learning Curve Model:

- The formula for the Cumulative Average version of a Learning Curve is very similar to that of Crawford’s Unit Learning Curve
- The difference being that the learning function applies to the Cumulative Average time of cost rather than Unit

\[ T_N = T_1 N^\varepsilon \]

where \( \varepsilon \) is the learning exponent: \( \varepsilon = \log p / \log 2 \)

with \( p \) = the learning percentage expressed as a decimal

and \( T_N \) is the Cumulative Average Time after \( N \) Units

- Unsurprisingly, the function also displays as a straight line in Log space, and follows the same characteristic steady state percentage reduction whenever the number of units produced is doubled (or is subjected to any other constant multiplier)

… except this time it applies to the Cumulative Average values
Cumulative Average Learning Curve: Unit Values

Unit Values for a Cumulative Average Learning Curve Model:

- Just as we had to calculate the Cumulative Average of a Unit Learning Curve by aggregating the Unit Values and dividing by the Number of Units
- The procedure for deriving the Unit Values from a Cumulative Average Curve is the reverse … one of multiplication and disaggregation:
  - Multiply the Cumulative Average Values by the Number of Units to get the Cumulative Values
  - Take the difference between successive Cumulative Values to get the Unit Values

<table>
<thead>
<tr>
<th>Unit No</th>
<th>Cum Ave</th>
<th>Cum Value</th>
<th>Unit Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000.0</td>
<td>1000.0</td>
<td>1000.0</td>
</tr>
<tr>
<td>2</td>
<td>800.0</td>
<td>1600.0</td>
<td>600.0</td>
</tr>
<tr>
<td>3</td>
<td>702.1</td>
<td>2106.3</td>
<td>506.3</td>
</tr>
<tr>
<td>4</td>
<td>640.0</td>
<td>2560.0</td>
<td>453.7</td>
</tr>
<tr>
<td>5</td>
<td>595.6</td>
<td>2978.2</td>
<td>418.2</td>
</tr>
</tbody>
</table>

Disaggregation of Cumulative Average

Giving the characteristic sharper initial unit reduction relative to the Unit Learning Curve

... then Asymptotically parallel in Log Space
Calibrating a Cumulative Average Learning Curve

Cumulative Average Data – A Useful Corollary

• The Cumulative value, $C_N$, of Wright’s Cumulative Average Learning Curve is also a straight line in Log space:

  Cum Ave Curve: \[ T_N = T_1 N^E \]

  By definition: \[ T_N = \frac{C_N}{N} \] \[ \Rightarrow \] \[ N T_N = C_N \]

  Giving: \[ C_N = T_1 N^{E+1} \]

Why is this useful? When it comes to calibrating a Cumulative Average Learning Curve, we have two choices:

1. Perform a Linear Regression on the Log of the Cumulative Average data
2. Perform a Linear Regression on the Log of the Cumulative data

Example

<table>
<thead>
<tr>
<th>Lot</th>
<th>No of Units</th>
<th>Lot Total Hrs</th>
<th>Cum Hrs</th>
<th>Cum Units</th>
<th>Cum Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>1675</td>
<td>1675</td>
<td>5</td>
<td>335</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1925</td>
<td>3600</td>
<td>15</td>
<td>240</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>2160</td>
<td>5760</td>
<td>30</td>
<td>192</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>1800</td>
<td>7560</td>
<td>45</td>
<td>168</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>3315</td>
<td>10875</td>
<td>75</td>
<td>145</td>
</tr>
</tbody>
</table>
Calibrating a Cumulative Average Learning Curve

Example – Cumulative Average Regression

<table>
<thead>
<tr>
<th>Lot</th>
<th>No of Units</th>
<th>Lot Total Hrs</th>
<th>Cum Hrs</th>
<th>Cum Units</th>
<th>Cum Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>1675</td>
<td>1675</td>
<td>5</td>
<td>335</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1925</td>
<td>3600</td>
<td>15</td>
<td>240</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>2180</td>
<td>5760</td>
<td>30</td>
<td>192</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>1800</td>
<td>7560</td>
<td>45</td>
<td>168</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>3315</td>
<td>10875</td>
<td>75</td>
<td>145</td>
</tr>
</tbody>
</table>

Adapted from

Let's Regress the Cum Ave Data …

SUMMARY OUTPUT

**Regression Statistics**

- Multiple R: 0.99985278
- R Square: 0.999705581
- Adjusted R Square: 0.999607441
- Standard Error: 0.002819859
- Observations: 5

**ANOVA**

- Regression: df=1, SS=0.080999489, MS=0.080999489, F=10186.55715, Significance F=2.14425E-06
- Residual: df=3, SS=2.38548E-05, MS=7.95161E-06
- Total: df=4

**Coefficients**

- Intercept: 2.743776701, Standard Error: 0.004431096, t Stat: 619.2094378, P-value: 9.28868E-09
- Log Cum Units: -0.311556218, Standard Error: 0.003086901, t Stat: -100.9284754, P-value: 2.14425E-06

The result is statistically significant on all 3 measures: R-Square, F and t Statistics

Calibrating a Cumulative Average Learning Curve

Example – Cumulative Average Regression

SUMMARY OUTPUT

**Regression Statistics**

- Multiple R: 0.99985278
- R Square: 0.999705581
- Adjusted R Square: 0.999607441
- Standard Error: 0.002819859
- Observations: 5

**ANOVA**

- Regression: df=1, SS=0.080999489, MS=0.080999489, F=10186.55715, Significance F=2.14425E-06
- Residual: df=3, SS=2.38548E-05, MS=7.95161E-06
- Total: df=4

**Coefficients**

- Intercept: 2.743776701, Standard Error: 0.004431096, t Stat: 619.2094378, P-value: 9.28868E-09
- Log Cum Units: -0.311556218, Standard Error: 0.003086901, t Stat: -100.9284754, P-value: 2.14425E-06

The result is statistically significant on all 3 measures: R-Square, F and t Statistics
Calibrating a Cumulative Average Learning Curve

Same Example – Cumulative Regression

<table>
<thead>
<tr>
<th>Lot</th>
<th>No of Units</th>
<th>Lot Total Hrs</th>
<th>Cum Hrs</th>
<th>Cum Units</th>
<th>Cum Ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>1675</td>
<td>1675</td>
<td>5</td>
<td>335</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>1925</td>
<td>3600</td>
<td>15</td>
<td>240</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>2160</td>
<td>5760</td>
<td>30</td>
<td>192</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>1800</td>
<td>7560</td>
<td>45</td>
<td>168</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>3315</td>
<td>10875</td>
<td>75</td>
<td>145</td>
</tr>
</tbody>
</table>

Now let's Regress the Cumulative Data ...

SUMMARY OUTPUT

Regression Statistics

Multiple R: 0.999969844
R Square: 0.999939688
Adjusted R Square: 0.999919584
Standard Error: 0.002819859
Observations: 5

ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>0.395499915</td>
<td>0.395499915</td>
<td>49738.3692</td>
<td>1.98793E-07</td>
</tr>
<tr>
<td>Residual</td>
<td>3</td>
<td>0.002323846</td>
<td>0.000774289</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>0.39552377</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.743776701</td>
<td>0.004431096</td>
<td>619.2094378</td>
<td>9.28886E-09</td>
</tr>
<tr>
<td>Log Cum Units</td>
<td>0.688443782</td>
<td>0.003086901</td>
<td>223.0210062</td>
<td>1.98793E-07</td>
</tr>
</tbody>
</table>

The result is statistically significant on all 3 measures: R-Square, F and t Statistics

Calibrating a Cumulative Average Learning Curve

Example – Cumulative Average Regression

The result is statistically significant on all 3 measures: R-Square, F and t Statistics

Adapted from...
Calibrating a Cumulative Average Learning Curve

By regressing the Cumulative Average Values or Regressing the Cumulative Values …

We get exactly the same result … or do we?

Whilst the calculated parameter values may be the same the statistical test results are not!

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Cum Ave Regression</th>
<th>Cumulative Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-Square</td>
<td>0.999705581</td>
<td>&lt; 0.999939688</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.002819859</td>
<td>= 0.002819859</td>
</tr>
<tr>
<td>Significance F</td>
<td>2.14425E-06</td>
<td>&gt; 1.98793E-07</td>
</tr>
<tr>
<td>Intercept P-value</td>
<td>9.28868E-09</td>
<td>= 9.28868E-09</td>
</tr>
<tr>
<td>Slope P-value</td>
<td>2.14425E-06</td>
<td>&gt; 1.98793E-07</td>
</tr>
</tbody>
</table>

In this case the difference is immaterial, but if the decisions conflict in terms of acceptance/rejection for the same data and same parameters, then it is probably not a Cumulative Average Curve anyway!

Choice of Learning Curve: Unit or Cum Average

Cumulative Average or Unit Learning: Which should we use?
- The choice of Learning Curve Type is often one of Organisational Policy, or simply custom and practice
- Really, the analyst should decide which we should use based on the data available, and the purpose for which we intend to use the analysis

What difference does it make which we choose?
- Cumulative Average Curves are inherently smoother than Unit Curves … but they are slower to respond to changes in the underlying Learning Curve Drivers
- The main difference between them is in the early units

What if we make the wrong choice?
- To some degree, the choice of Learning Curve type is somewhat forgiving, providing that we have sufficient quantity going forward …
**Choice of Learning Curve: Unit or Cum Average**

**Conway & Schultz Cumulative Approximation Formula for Unit Learning**
(Basis of the Unit Learning Curve’s Lot Average Approximation):

\[ C_N \approx \frac{t_1}{E+1} \left( (N + 0.5)^{E+1} - 0.5^{E+1} \right) \]

… which is asymptotic to:

\[ C_N \rightarrow \frac{t_1}{E+1} N^{E+1} \]

… which in turn gives a Cumulative Average Approximation which is asymptotic to:

\[ T_N \rightarrow \frac{t_1}{E+1} N^E \]

**Compare this with Wright's Cumulative Average Curve:**

\[ T_1 = \frac{t_1}{E+1} N^E \]

Source:
Conway, R.W. and Schultz, A.Jr., 'The Manufacturing Progress Function', Journal of Industrial Engineering, Jan-Feb 1959, pp.39-54

---

**Choice of Learning Curve: Unit or Cum Average**

**How precisely inaccurate do we want to be?**

The main difference occurs in first few unit values but the knock-on Cumulative and Cumulative Average effects last a little longer.

Cumulative Average Curve runs parallel to the Unit Curve for larger quantities

80% Learning Rate Example

---

ICEAA 2016 Bristol – TRN 03
Choice of Learning Curve: Unit or Cum Average

How precisely inaccurate do we want to be?

<table>
<thead>
<tr>
<th>Unit No</th>
<th>Crawford</th>
<th>Wright</th>
<th>Difference</th>
<th>Crawford</th>
<th>Wright</th>
<th>Difference</th>
<th>Crawford</th>
<th>Wright</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000</td>
<td>1475</td>
<td>475</td>
<td>1000</td>
<td>1475</td>
<td>475</td>
<td>1000</td>
<td>1475</td>
<td>475</td>
</tr>
<tr>
<td>2</td>
<td>800</td>
<td>885</td>
<td>85</td>
<td>900</td>
<td>1180</td>
<td>280</td>
<td>1800</td>
<td>2360</td>
<td>560</td>
</tr>
<tr>
<td>3</td>
<td>702</td>
<td>747</td>
<td>45</td>
<td>834</td>
<td>1035</td>
<td>201</td>
<td>2502</td>
<td>3106</td>
<td>604</td>
</tr>
<tr>
<td>4</td>
<td>640</td>
<td>669</td>
<td>29</td>
<td>786</td>
<td>944</td>
<td>158</td>
<td>3142</td>
<td>3775</td>
<td>633</td>
</tr>
<tr>
<td>5</td>
<td>596</td>
<td>617</td>
<td>21</td>
<td>748</td>
<td>878</td>
<td>131</td>
<td>3738</td>
<td>4392</td>
<td>654</td>
</tr>
<tr>
<td>6</td>
<td>562</td>
<td>578</td>
<td>16</td>
<td>717</td>
<td>828</td>
<td>112</td>
<td>4299</td>
<td>4970</td>
<td>671</td>
</tr>
<tr>
<td>7</td>
<td>534</td>
<td>546</td>
<td>13</td>
<td>691</td>
<td>788</td>
<td>98</td>
<td>4834</td>
<td>5518</td>
<td>684</td>
</tr>
<tr>
<td>8</td>
<td>512</td>
<td>523</td>
<td>11</td>
<td>668</td>
<td>705</td>
<td>37</td>
<td>5346</td>
<td>6041</td>
<td>695</td>
</tr>
<tr>
<td>9</td>
<td>493</td>
<td>502</td>
<td>9</td>
<td>649</td>
<td>727</td>
<td>78</td>
<td>5839</td>
<td>6543</td>
<td>704</td>
</tr>
<tr>
<td>10</td>
<td>477</td>
<td>485</td>
<td>8</td>
<td>632</td>
<td>703</td>
<td>71</td>
<td>6315</td>
<td>7027</td>
<td>712</td>
</tr>
<tr>
<td>15</td>
<td>418</td>
<td>423</td>
<td>5</td>
<td>567</td>
<td>617</td>
<td>49</td>
<td>8511</td>
<td>9251</td>
<td>741</td>
</tr>
<tr>
<td>20</td>
<td>381</td>
<td>384</td>
<td>3</td>
<td>524</td>
<td>562</td>
<td>38</td>
<td>10485</td>
<td>11244</td>
<td>759</td>
</tr>
<tr>
<td>25</td>
<td>355</td>
<td>357</td>
<td>2</td>
<td>492</td>
<td>523</td>
<td>31</td>
<td>12309</td>
<td>13081</td>
<td>772</td>
</tr>
<tr>
<td>30</td>
<td>335</td>
<td>336</td>
<td>1</td>
<td>467</td>
<td>493</td>
<td>26</td>
<td>14020</td>
<td>14802</td>
<td>782</td>
</tr>
<tr>
<td>40</td>
<td>305</td>
<td>306</td>
<td>1</td>
<td>430</td>
<td>450</td>
<td>20</td>
<td>17193</td>
<td>17990</td>
<td>797</td>
</tr>
<tr>
<td>50</td>
<td>284</td>
<td>285</td>
<td>1</td>
<td>402</td>
<td>419</td>
<td>16</td>
<td>20122</td>
<td>20929</td>
<td>807</td>
</tr>
<tr>
<td>75</td>
<td>249</td>
<td>250</td>
<td>1</td>
<td>356</td>
<td>367</td>
<td>11</td>
<td>26727</td>
<td>27552</td>
<td>825</td>
</tr>
<tr>
<td>100</td>
<td>227</td>
<td>227</td>
<td>0</td>
<td>327</td>
<td>335</td>
<td>8</td>
<td>32651</td>
<td>33486</td>
<td>836</td>
</tr>
<tr>
<td>200</td>
<td>182</td>
<td>182</td>
<td>0</td>
<td>264</td>
<td>268</td>
<td>4</td>
<td>52720</td>
<td>53578</td>
<td>858</td>
</tr>
<tr>
<td>500</td>
<td>135</td>
<td>135</td>
<td>0</td>
<td>168</td>
<td>199</td>
<td>2</td>
<td>98847</td>
<td>99729</td>
<td>881</td>
</tr>
<tr>
<td>1000</td>
<td>108</td>
<td>108</td>
<td>0</td>
<td>159</td>
<td>160</td>
<td>1</td>
<td>158671</td>
<td>159566</td>
<td>895</td>
</tr>
</tbody>
</table>

80% Learning Rate Example

Cum Average Learning

Questions?