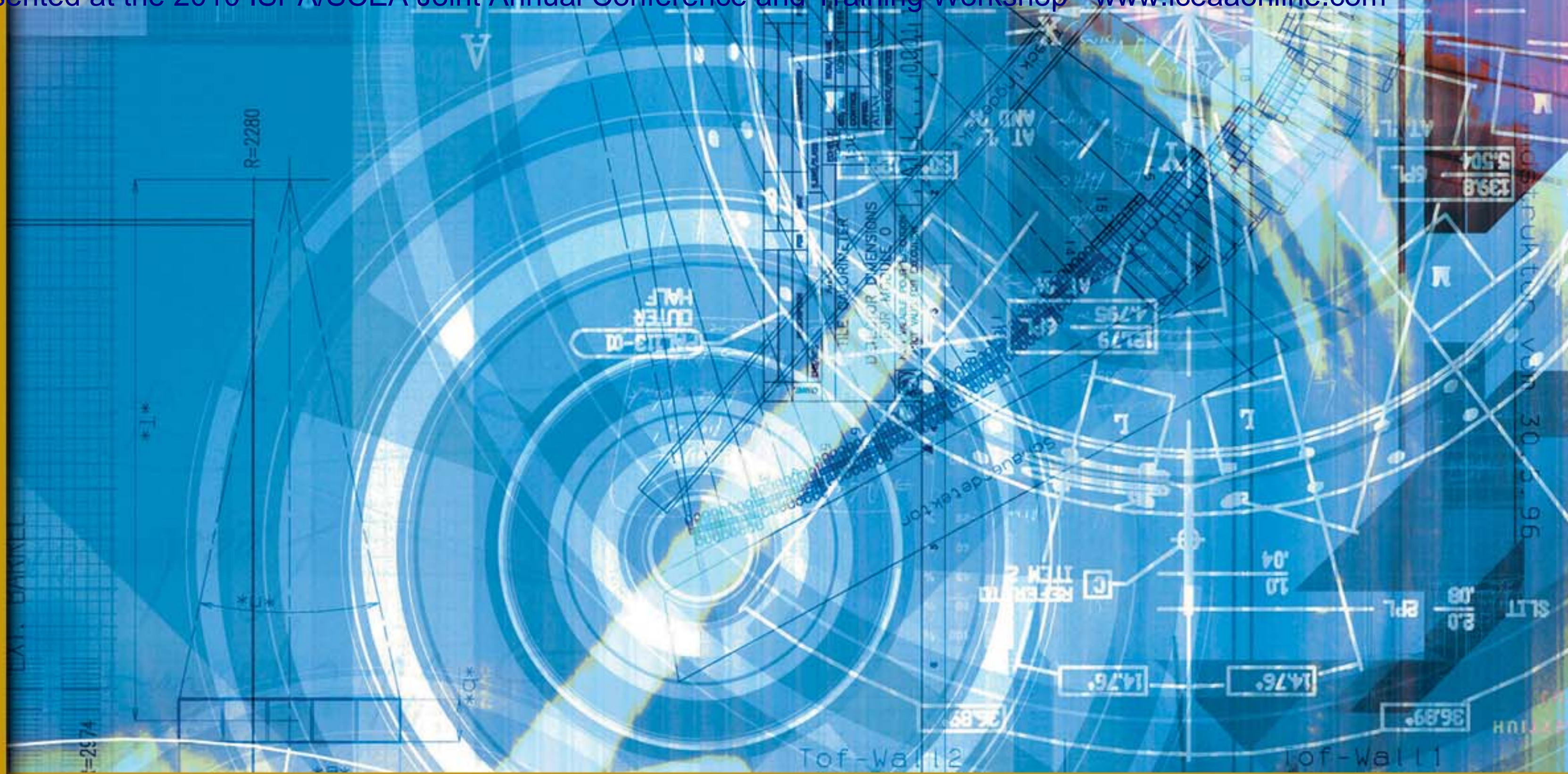


Testing for the Significance of Cost Drivers Using Bootstrap Sampling

ISPA/SCEA Joint Conference

presented by:

Daniel I. Feldman, MCR LLC
390 No. Sepulveda Blvd.,
El Segundo CA 90245
(424) 218-1618

dfeldman@mcri.com

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Abstract

Most regression-based statistical modeling aims to derive, on the basis of historical data, an algebraic expression for estimating a quantity of interest, such as cost. Such an algebraic expression is characterized by one or more "fit parameters," which are typically coefficients or exponents. Once these fit parameters are determined, it is a logical next step to test for their accuracy by a process called "statistical inference." The fit parameters of a cost-estimating relationship (CER) are each directly linked to a candidate cost driver. The inference process may lead to some fit parameters being judged "significant" and others "not significant," where "significance" refers to the extent to which their respective associated cost drivers impact the cost estimates made using the CER. Cost drivers associated with fit parameters that are judged to be not significant can (and should) be removed from the algebraic expression for the CER without sacrificing any estimating capability. Furthermore, the ability to remove insignificant cost drivers from a CER can be valuable, because fewer fit parameters means more degrees of freedom, and that can be important when only a small number of data points is available. Thus it is important to be able to eliminate the insignificant cost drivers from a CER.

Because of the strict mathematical assumptions underlying ordinary least squares (OLS) linear regression, explicit formulas for fit parameters, significance testing and confidence bounds can be established. However, if any of the OLS assumptions do not apply, inferences based on those formulas are unreliable. This situation has led cost analysts to understand the need for general-error regression. Up to now, analysts have been concerned primarily with the quality of the estimating process itself, and have been ignoring the need to assess the significance of the fit parameters and cost drivers. Recently, however, researchers have begun to slowly realize this shortcoming and have started to explore avenues for filling the gap. In a paper entitled, "A Distribution-Free Measure of the Significance of CER Regression Fit Parameters Established Using General Error Regression Methods," (*Journal of Cost Analysis and Parametrics*, Vol. 2, Issue 1, Summer 2009, pages 7-22), T.P. Anderson proposed some "heuristic" techniques of assessing the significance of the fit parameters.

Anderson calls his method heuristic, because it is not determined by strict mathematical calculations. Instead, we will apply bootstrap sampling, a method that can approximate the mathematics of statistical inference procedures and that was recently applied to the case of prediction bounds for CERs derived by general-error regression, another situation where explicit formulas exist for OLS CERs but not for general-error CERs. (See the briefing by S.A. Book, "Prediction Bounds for General-Error-Regression CERs," 39th Annual DoDCAS, Williamsburg VA, 14-17 February.) The bootstrap is a data-based method that does not require any OLS-like distributional assumptions or explicit formulas. In this paper, we specifically look at constructing approximate bootstrap-based statistical tests to assess the significance of CER fit parameters and their associated cost drivers, allowing conclusions to be drawn regarding which candidate cost drivers are significant and which are not.

Contents

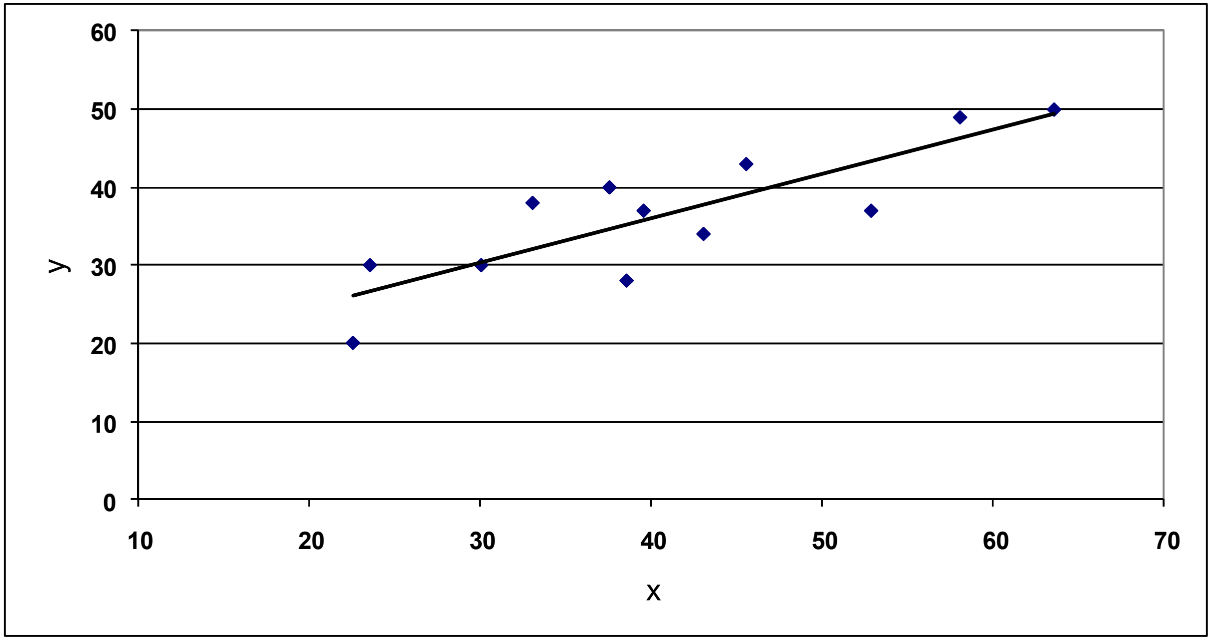
CRITICAL THINKING. SOLUTIONS DELIVERED.

Introduction

- **In Ordinary Least Squares (OLS) regression, there exist formulas for making inferences about the regression coefficients and independent variables (cost drivers), evaluating their statistical significance**
- **Details can be found in most introductory statistical textbooks – [Ref. 1] for example**

$$y = 0.57x + 13.186$$

Both parameters (β_0 and β_1) in this case were found to be significant at significance level α = 0.05



CRITICAL THINKING. SOLUTIONS DELIVERED.

A Caveat

- However, if any of the strict assumptions of OLS are violated, these and other inferences from OLS formulas become unreliable
- This has led to the need and growing popularity of General Error Regression Methods (GERM)
- Despite the growing use of GERM, it seems most analysts have been ignoring the issue of the need to determine the significance of the parameters and cost drivers
- T.P. Anderson has been doing work on addressing this very shortcoming through something he calls the "SIG Test" – [Ref. 2]

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

General Regression

- General Regression refers to a method of establishing a regression relationship that is either
  - Nonlinear (such as exponential, power or square)
  - A linear form that violates one or more of the OLS assumptions (such as possessing multiplicative or non-normally distributed errors)
- It is not realistic to expect one expression to be sufficient for inferences about the parameters/cost drivers or predictions for every regression form, and in fact such formulas are not available for general regression
- The Bootstrap has been suggested as a method for approximating inference procedures
- We will specifically look at constructing bootstrap-derived approximate hypothesis tests to examine the significance of the estimated CER parameters and the linked cost driver

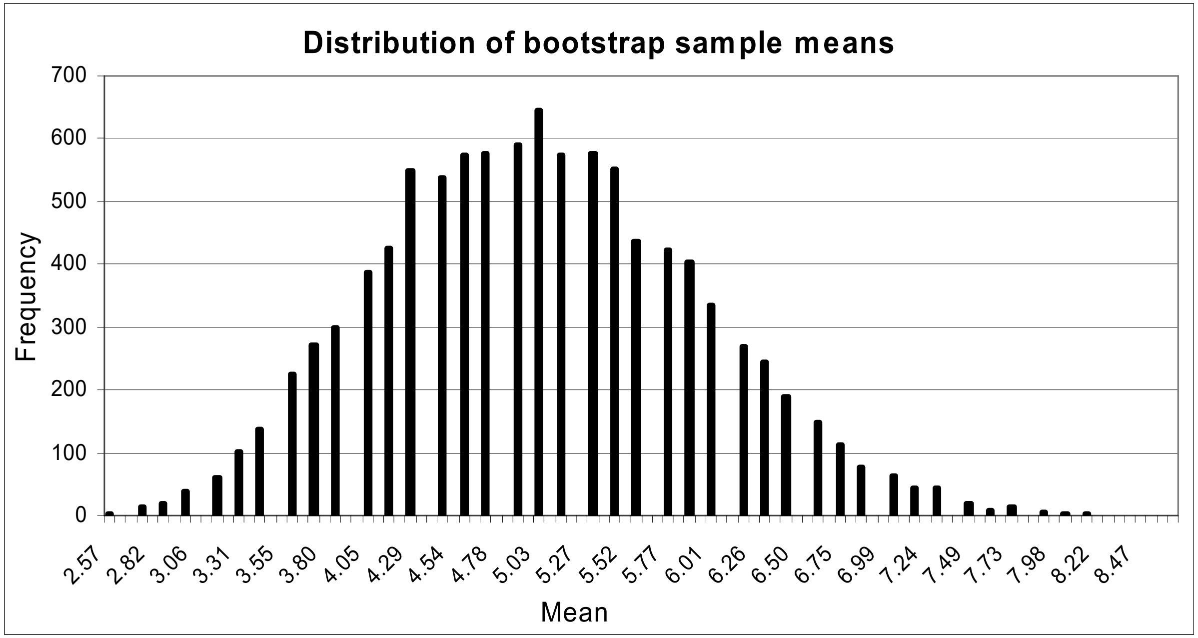
MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Introduction to the Bootstrap

- **Bootstrapping involves resampling with replacement from sample data numerous times in order to generate an empirical distribution of a statistic**
- **It is a data-based simulation method that is used for making statistical inferences**
- **It is a distribution-free method that has no need for any distributional assumptions or analytic formulas**
- **This freedom allows the use of the bootstrap in situations where traditional theory fails to provide an answer**

Basic Bootstrap Illustration (Sample Mean)

- **Suppose our sample is (2, 5, 4, 9, 8, 4, 3), with the sample mean, $\mu = 5$**
- **1st Bootstrap sample = (4, 9, 8, 3, 3, 2, 2), $\mu_1^* = 4.43$**
- **2nd Bootstrap sample = (4, 3, 3, 9, 8, 3, 8), $\mu_2^* = 5.43$**
- **… 10,000th Bootstrap sample = (3, 2, 2, 2, 3, 8, 5), $\mu_{10,000}^* = 3.57$**



Note: A superscript '*' indicates a bootstrap quantity/statistic

Bootstrap Methods

- The bootstrap is a family of methods, many variants of which are available, and, depending on the problem at hand, some are more relevant than others
- One such bootstrap technique is called the Bootstrap-t (also called “Percentile-t”) method
  - It is a valuable bootstrap technique for constructing hypothesis tests, structured very similarly to classical hypothesis tests
  - This method estimates the error distribution directly from the data, therefore there is no need to make assumptions about the underlying error distribution
  - The first step for the bootstrap-t is to construct a “t-like” table based on the data

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 1 Dataset

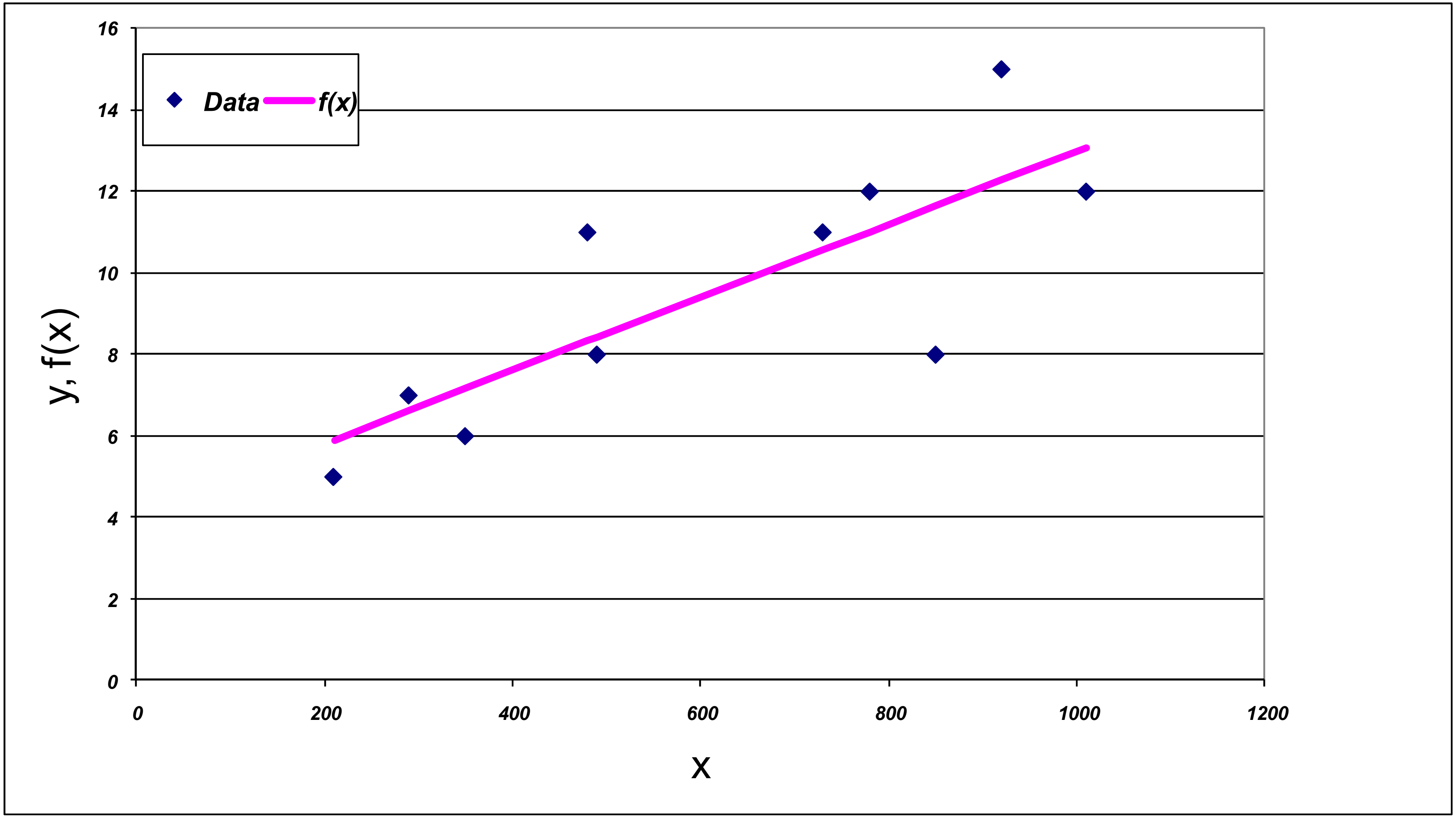
Weight (X)	Cost (Y)	f(x)	Residual e = Y - f(x)
210	5	5.902	-0.902
290	7	6.620	0.380
350	6	7.158	-1.158
480	11	8.325	2.675
490	8	8.414	-0.414
730	11	10.568	0.432
780	12	11.016	0.984
850	8	11.644	-3.644
920	15	12.273	2.727
1010	12	13.080	-1.080

$$f(x) = 4.018 + 0.00897x$$

Result of OLS regression

MCR Critical Thinking. Solutions Delivered.

Graph of Example 1



Before Bootstrap-t

- **We are interested in assessing the statistical significance of the fit parameters**

  $\hat{\beta}_0 = 4.018$ **and** $\hat{\beta}_1 = 0.00897$

  and the cost driver Weight (x)
- **Bootstrapping of the data is done by resampling the residuals from the regression model (Full details of this process can be found in [Ref. 5, pp. 111-112], but will be touched on in the next slide)**
- **Let's say 500 iterations are run, then the result will be 500 "new" equations based on these 500 bootstrap resamples**

Bootstrap Residuals

- Instead of resampling the data as in the 'Basic Bootstrap Illustration' on Chart 8, fit a regression model to the original data and resample the residuals
- Add those resampled residuals to the estimated y values to get the bootstrap samples (or multiply them by the estimated y values if the model has multiplicative error)
- Then regression equations are fit to each bootstrap sample
- Now, back to the construction of the "t-like" table

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap Pivoting

- First we will standardize the $\hat{\theta}^*_{[b]}$ s using the approximate pivot (where the particular statistic of interest, say $\hat{\beta}_0$ or $\hat{\beta}_1$, is substituted for $\hat{\theta}$):

$$t^*_b = \frac{(\hat{\theta}^*_{[b]} - \hat{\theta})}{\hat{\sigma}^*_{\hat{\theta}_{[b]}}}$$

where:

$\hat{\theta}^*_{[b]}$ is the θ bootstrap coefficient estimate from the b^{th} bootstrap sample,

$\hat{\theta}$ is the coefficient estimate from the original sample,

$\hat{\sigma}^*_{\hat{\theta}_{[b]}}$ is the standard error (SE) of θ from the b^{th} bootstrap sample

CRITICAL THINKING. SOLUTIONS DELIVERED.

Double Bootstrap

- When no formula for the SE of θ exists, one can use the “double bootstrap” to approximate the SE
- The double bootstrap involves resampling from each bootstrap resample R times
- Bootstrapping the bootstrap sample is a way to investigate the variability in the bootstrap sample (just as bootstrapping the original sample does for the original sample)
- For example, if the sample is bootstrapped 500 (B) times and each bootstrap sample is resampled 500 (R) times, we need to draw a total of 500 x 500 = 250,000 samples

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

A Look at the Double Bootstrap

- **Let the original sample be:**

$$X = (x_1, x_2, K , x_{10})$$

- **Then an example of a bootstrap sample is:**

$$X^* = (x_7, x_{10}, x_8, x_7, x_6, x_{10}, x_6, x_6, x_9, x_4) = (x_1^*, x_2^*, K , x_{10}^*)$$

- **And then bootstrapping that bootstrap sample gives the double bootstrap sample:**

$$X^{**} = (x_7, x_4, x_4, x_4, x_6, x_{10}, x_6, x_{10}, x_6, x_{10}) = (x_1^{**}, x_2^{**}, K , x_{10}^{**})$$

- **How can this be used to approximate the SE of θ?**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap Standard Error (σ*)

- **The SE for any coefficient can be approximated using the bootstrap by:**

$$\hat{\sigma}^*_{\hat{\theta}_{[b]}} = \sqrt{\frac{\sum_r \left(\hat{\theta}^*_{[b,r]} - \bar{\hat{\theta}}^*_{[b,(.)]}\right)^2}{(R-1)}}$$

where $\hat{\theta}^*_{[b,r]}$ **is the parameter estimate for the** ***rth*** **bootstrap resample from the** ***bth*** **bootstrap sample, and**

$$\bar{\hat{\theta}}^*_{[b,(.)]} = \frac{\sum_r \hat{\theta}^*_{[b,r]}}{R}$$

CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap Standard Error of β_0

- In the case of β_0, the SE can be approximated by:

$$\hat{\sigma}^*_{\hat{\beta}_{0,[b]}} = \sqrt{\frac{\sum_r \left(\hat{\beta}^*_{0,[b,r]} - \bar{\hat{\beta}}^*_{0,[b,(.)]}\right)^2}{(R-1)}}$$

where

$$\bar{\hat{\beta}}^*_{0,[b,(.)]} = \frac{\sum_r \hat{\beta}^*_{0,[b,r]}}{R}$$

CRITICAL THINKING. SOLUTIONS DELIVERED.

Constructing the "t-like" Table

- The result is that the t_b^* s are distributed as a standardized bootstrap distribution of the estimator (based on slides 14, 17, and 18)
- These values are then rank-ordered from smallest to largest
- Then they are listed in a "t-like" table form (different tables for both $\hat{\beta}_0$ and $\hat{\beta}_1$)
- These tables can be utilized in a similar way as the standard t-table
- The next slide will display an example of how the tables are established

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Quick t_b* Calculation for Example 1

- **Let B = 500 and R = 500**
- **Recall that** $\hat{\beta}_0 = 4.018$

$$t_5* = \frac{(\hat{\beta}^*_{0,[5]} - \hat{\beta}_0)}{\hat{\sigma}^*_{\hat{\beta}_{0,[5]}}}$$

b	$\hat{\beta}^*_0$	σ*	t_b*
1	4.257	1.802	0.133
2	2.990	1.459	-0.705
3	5.668	1.421	1.161
4	3.645	1.724	-0.216
5	2.746	0.952	-1.335
6	3.136	1.109	-0.795
7	4.036	1.102	0.016
8	5.699	1.736	0.969
9	4.340	1.514	0.213
10	3.429	1.693	-0.348
⋮	⋮	⋮	⋮
500	2.047	1.524	-1.293

$$t^*_5 = \frac{(2.746 - 4.018)}{0.952} = -1.335$$

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Setting Up SE(β_0) of the 5th Bootstrap Sample

- **Why is the bootstrap approximation to the SE of β_0 for the 5th bootstrap sample = 0.952?**
- **Suppose the 5th Bootstrap Sample is:**

$$Y^* = (4.822,\ 5.462,\ 7.59,\ 7.91,\ 8.795,\ 9.488,\ 12,\ 14.372,\ 11.858,\ 14.064)$$

(with the associated x values) Giving

$$\hat{\beta}^*_{0,[b]} = \hat{\beta}^*_{0,[5]} = 2.746$$

- **Using the double bootstrap, the first resample of the bootstrap sample might be:**

$$Y^{**} = (6.334, 9.347, 6.078, 7.166, 8.795, 9.488, 10.602, 10.564, 12.653, 12.666)$$

giving $\hat{\beta}^{**}_{0,[b,r]} = \hat{\beta}^{**}_{0,[5,1]} = 4.879$

- **Repeat this R times (500 times in this case)**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Calculating SE(β_0) of the 5th Bootstrap Sample

- **Based on the equations in slides 17 and 18:**

$$\overline{\hat{\beta}}^{*}_{0,[5,(.)]} = \frac{\sum_{r=1}^{500} \hat{\beta}^{*}_{0,[5,r]}}{500} = \frac{2053.19}{500} = 4.106$$

so that

$$\hat{\sigma}^{*}_{\hat{\beta}_{0,[5]}} = \sqrt{\frac{\sum_{r=1}^{500} \left(\hat{\beta}^{*}_{0,[5,r]} - \overline{\hat{\beta}}^{*}_{0,[5,(.)]}\right)^2}{(500-1)}} = \sqrt{\frac{(4.879 - 4.106)^2 + \ldots}{(500-1)}}$$

$$= \sqrt{\frac{452.205}{499}} = 0.952$$

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Formulas for SE(β_0) of the 5th Bootstrap Sample

- **However, formulas exist to calculate the SE of β_0 in this OLS case**

$$\mathrm{SE}(\hat{\beta}_0) = \sqrt{s^2\left(\frac{1}{n} + \frac{\bar{x}^2}{s_{xx}}\right)}$$

where

$$s^2 = \frac{\sum (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2}{n-2} \quad \textbf{and} \quad s_{xx} = \sum (x_i - \bar{x})^2$$

- **Had we chosen to utilize them instead of approximating using the bootstrap, then SE($\hat{\beta}_{0,[5]}$) = 0.868, reasonably close to our bootstrap result of 0.952**

CRITICAL THINKING. SOLUTIONS DELIVERED.

Hypothesis Test

- **Now we have "t-like" tables, so suppose we want to test:**

$$H_0: \theta = \hat{\theta}_0$$
$$H_A: \theta \neq \hat{\theta}_0$$

- **Based on standard theory, the T-value used in a standard hypothesis test is:**

$$T = \frac{\hat{\theta} - \hat{\theta}_0}{\hat{\sigma}_\theta}$$

- **Using that idea and what was introduced earlier, a "T-like" value can be calculated using the bootstrap:**

$$T^* = \frac{\hat{\theta} - \hat{\theta}_0}{\hat{\sigma}_\theta^*}$$

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

P-Value

- **Definition:** $\mathbf{p-value} = \mathbf{P}\left(|t| > |T|\right)$
- **This expression can be approximated by:**

$$\mathrm{p-value}_{\theta} = \frac{\left\lfloor \#\quad \mathrm{of}\quad |t_i^*| \geq |T^*| \right\rfloor}{B}$$

where B = number of bootstrap runs, and θ is the coefficient (cost driver) being tested

- **If p-value$_\theta$ ≤ α, we can reject the null hypothesis (H$_0$) at significance level α and conclude that $\theta \neq \hat{\theta}_0$**

- **Thus our coefficient and associated cost driver are significant at the prescribed α value**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Summary Overview of Example 1

- B = 500 and R = 500
- Testing the two coefficients β_0 and β_1 where

$y = \beta_0 + \beta_1 x$:

H_0: $\beta_0 = 0$ H_0: $\beta_1 = 0$

H_a: $\beta_0 \neq 0$ H_a: $\beta_1 \neq 0$

- The Bootstrap approximates the SE of the original sample in order to calculate T^*, and the double bootstrap approximates the SE of each bootstrap sample

CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 1: Classical Hypothesis Tests

	Coefficients	*Standard Error*	*t Stat*	*P-value*
β_0	4.0177	1.5909	2.5254	0.0355
β_1	0.0090	0.0024	3.7630	0.0055

- **Based on the classical methods of constructing the hypothesis tests, both β_0 and β_1 are significant at significance level α = 0.05:**
  - **β_0 has p-value = 0.0355**
  - **β_1 has p-value = 0.0055**

CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 1: Bootstrap for Testing β_0

Fit parameter =	4.018
σ* of parameter =	1.446
T* =	2.778
p-value =	0.002

$$T^* = \frac{4.018 - 0}{1.446} = 2.778$$

- **Of the 500 values in the "t-like" table, the absolute value of only one of them was found to be greater than *T** (approximate p-value = 1/500 = 0.002)**
- **Thus we can reject the null hypothesis at significance level α = 0.05**
- **Therefore β_0 is significant at α value 0.05, just as in the classical theory**

Example 1: Bootstrap for Testing β_1

Fit parameter =	0.00897
σ* of parameter =	0.0021
T* =	4.299
p-value =	0

- The results above give us an approximate p-value = 0/500 = 0
- Thus we can reject the null hypothesis at significance level $\alpha = 0.05$
- This shows that β_1 is significant at α value 0.05, just as in the classical theory

Example 1: P-value Comparison

	Classical Theory	Bootstrap Theory
β_0	0.0355	0.002
β_1	0.0055	0

- We have reached similar conclusions using the bootstrap approach for the OLS linear regression example
- However, the bootstrap approach apparently shows the fit parameters to be “more” significant
- Why would the bootstrap result be more significant?
- Recall that the standard error (σ^*) for β_0 based on the bootstrap was 1.446 (Chart 28), while, based on the classical equation, it was 1.591 (Chart 27)

Optimistic P-values

- Statistical research has shown that the bootstrap is typically optimistic (meaning it favors rejecting the null hypothesis) in small sample sizes, because it tends to underestimate the error rate [Ref. 8]
- This effect is apparent in the lower p-values associated with the bootstrap, implying that the fit parameters (and corresponding cost drivers) are judged to be more significant than what the classical theory says
- One suggestion to partially help correct for this is to make an adjustment to the residuals (see next chart)

OLS Residual Modification

- **It is known that $\mathbf{var}(e_i) = \sigma^2(1-h_i)$ [Ref. 8], where the "leverage" of x_i is defined as:**

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_j - \bar{x})^2}$$

- **The raw residuals can therefore be modified in order that they have a constant variance:**

$$r_i = \frac{e_i}{\sqrt{1-h_i}}$$

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Implementing OLS-Adjusted Residuals

- **Then we define the adjusted residuals (errors) as** $e_i(adj) = r_i - \bar{r}$
- **This definition ensures that the average of the adjusted residuals is zero (i.e., estimates of the points in the data base are unbiased)**
- **Now use the adjusted residuals in the resampling bootstrap process (described on Chart 13) and the double bootstrap process (if necessary)**
- **Then implement the bootstrap-t just as before**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Reexamination of Example 1: P-value Comparison

	Classical Theory	Bootstrap Theory	Bootstrap Theory (Modified Residuals)
β_0	0.0355	0.002	0.034
β_1	0.0055	0	0

- The p-values coming from the bootstrap theory with modified residuals are closer to the classical theory (approximate p-value = 17/500 = 0.034)
- This shows the coefficients to be significant at an α value similar to that of the classical theory
- This appears to help correct for some of the inherent optimism of the bootstrap

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

General-Regression Residual Modification

- However, in non-OLS regression should the adjustment be the same?
- One simple suggested residual modification is to use the average leverage [Ref. 4]

$$1 - \overline{h} = 1 - \frac{2}{n}$$

- This leads to adjusting the raw residuals by:

$$r_i = \frac{e_i}{\sqrt{1 - \frac{2}{n}}}$$

(this adjustment can also be used for OLS cases)
- However, more appropriate adjustments may exist

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 2 Dataset

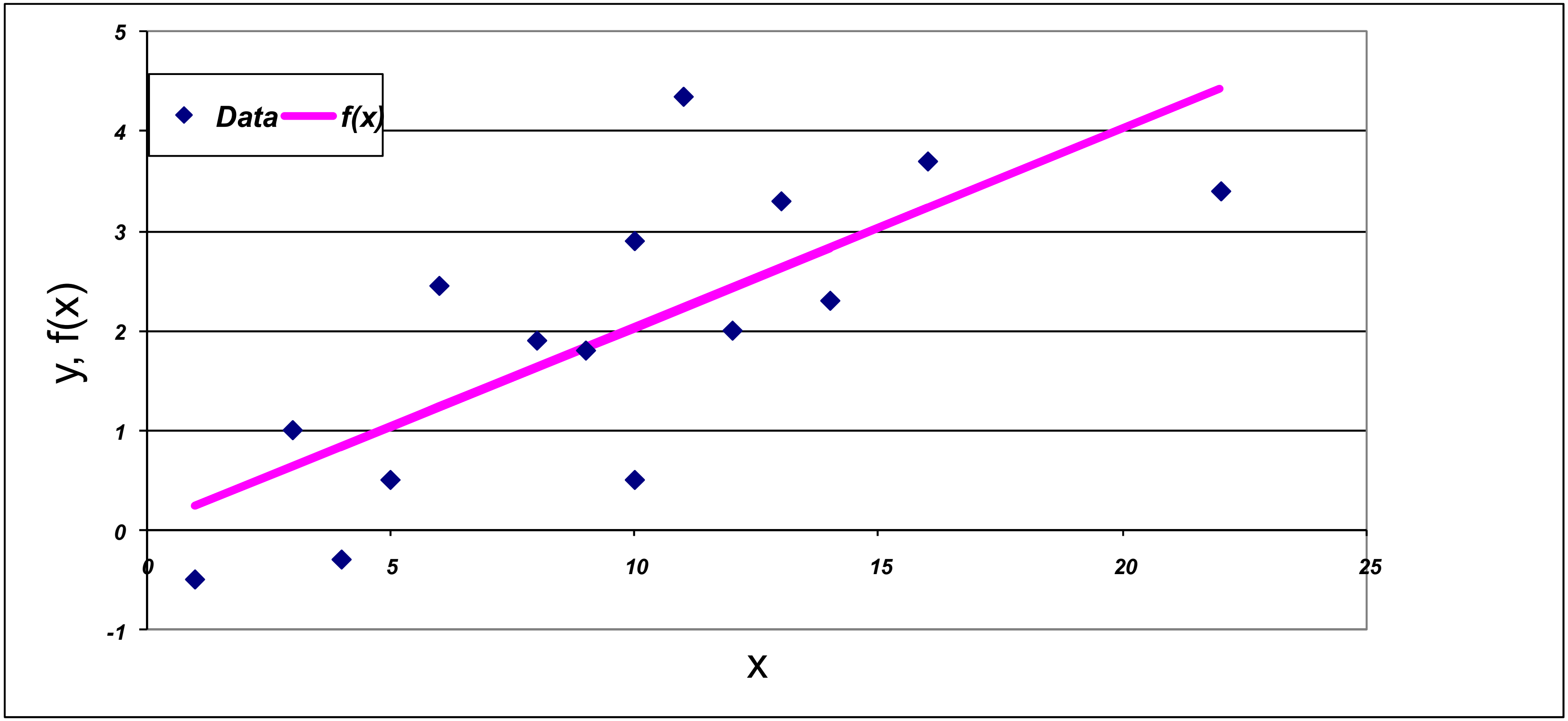
X	Y	f(x)	Y - f(x)
4	-0.3	0.838	-1.138
9	1.8	1.834	-0.034
10	0.5	2.033	-1.533
14	2.3	2.830	-0.530
12	2.0	2.431	-0.431
22	3.4	4.423	-1.023
1	-0.5	0.241	-0.741
3	1.0	0.639	0.361
8	1.9	1.635	0.265
5	0.5	1.037	-0.537
6	2.5	1.236	1.214
10	2.9	2.033	0.867
11	4.4	2.232	2.118
16	3.7	3.228	0.472
13	3.3	2.630	0.670

$$f(x) = 0.0413 + 0.1992x$$

Result of OLS regression

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Graph of Example 2



CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 2: Classical Theory

	Coefficients	Standard Error	t Stat	P-value
β_0	0.0413	0.5468	0.0756	0.9409
β_1	0.1992	0.0499	3.9920	0.0015

- **Based on the classical methods of constructing the hypothesis tests, β_0 was found to be insignificant**
- **β_0 has a p-value = 0.9409**
- **β_1 has a p-value = 0.0015**

CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 2: Testing β_0 Using Bootstrap

Standard Residuals Resampled:

fit parameter =	0.041341
σ* of parameter =	0.501267
T* =	0.082473
p-value =	0.938124

Modified Residuals Resampled:

fit parameter =	0.041341
σ* of parameter =	0.515104
T* =	0.080257
p-value =	0.94012

- **Let B = 501 and R = 500, using modified residuals**
- **Resampling the standard residuals gave the absolute value of 470 of the values greater then T^* (470/501 = 0.938), while the modified gave 471 greater then T^* (471/ 501 = 0.94012)**
- **As sample size increases, modification of residuals turns out to be not as crucial**
- **Bootstrap theory compares well with classical theory in this case, showing β_0 to be highly insignificant**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 3

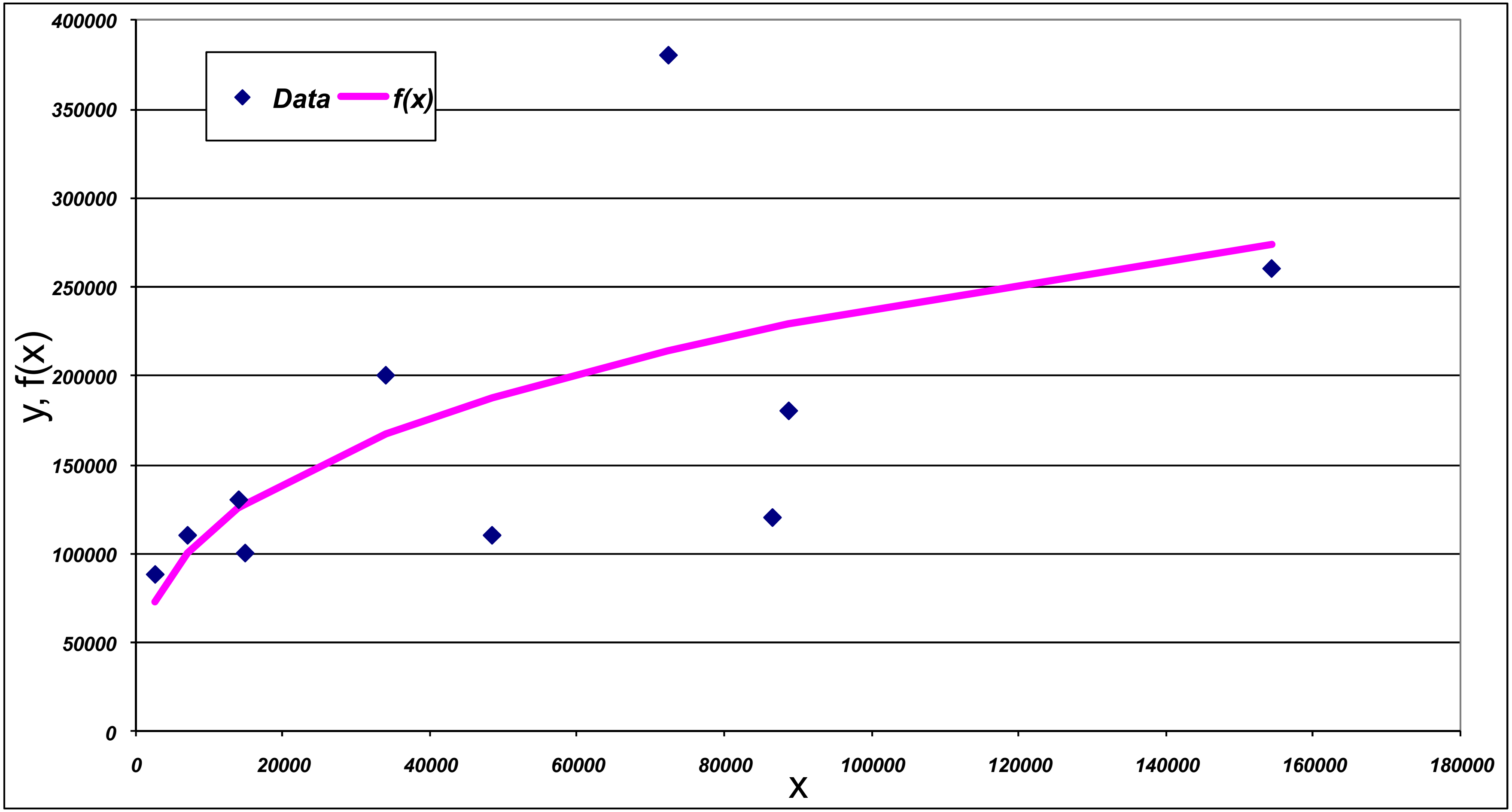
Weight (X)	Cost (Y)	f(x)	Y / f(x)
86507	120000	226942.59	0.529
6974	110000	100208.57	1.098
2562	88000	72396.92	1.216
33919	200000	167462.15	1.194
14800	100000	127934.92	0.782
48375	110000	187917.94	0.585
72360	380000	214160.64	1.774
154378	260000	273887.90	0.949
88718	180000	228809.56	0.787
13929	130000	125440.46	1.036

$$f(x) = 5664.7\, x^{0.3246}$$

Result of Zero Percentage Bias, Minimum Percentage Error (ZMPE) regression

CRITICAL THINKING. SOLUTIONS DELIVERED.

Graph of Example 3



MCR Critical Thinking. Solutions Delivered.

Overview of Example 3

- From the graph on the previous chart, the CER appears to be of the power form, $y = \beta_0 x^{\beta_1}$, with multiplicative error
- Let B = 400 and R = 200
- Testing the two parameters (β_0 and β_1):

H_0: $\beta_0 = 1$	H_0: $\beta_1 = 0$
H_a: $\beta_0 \neq 1$	H_a: $\beta_1 \neq 0$

- Notice that the null hypothesis $\beta_0 = 1$ (not $\beta_0 = 0$) is being tested
- This is due to the fact that β_0 is a multiplier and does not effect the functional form if it equals 1 (and, if so, is insignificant)

CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap: Testing β_0

Standard Residuals Resampled:

fit parameter =	5664.678
σ* of parameter =	5030.367
T* =	1.125898
p-value =	0.18

$$T^* = \frac{\hat{\beta}_0 - 1}{\sigma^*_{(\beta_0)}} = \frac{5664.7 - 1}{5030.367} = 1.126$$

Modified Residuals Resampled:**

fit parameter =	5664.678
σ* of parameter =	12576.82
T* =	0.450406
p-value =	0.415

$$T^* = \frac{\hat{\beta}_0 - 1}{\sigma^*_{(\beta_0)}} = \frac{5664.7 - 1}{12576.82} = 0.45$$

- **These calculations give us an approximate p-value = 0.18 (unmodified) and p-value = 0.415 (modified)**
- **Thus we cannot reject the null hypothesis, showing that β_0 is not significant at $\alpha = 0.05$**
- **Therefore is weight not a significant cost driver?**

** **'Modified Residuals' are modified by the method suggested on chart 35**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap: Testing β_1

Standard Residuals Resampled:

fit parameter =	0.324635
σ* of parameter =	0.065985
T* =	4.919831
p-value =	0.0025

$$T^* = \frac{\hat{\beta}_1 - 0}{\sigma^*_{(\beta_1)}} = \frac{0.3246}{0.06599} = 4.92$$

Modified Residuals Resampled:**

fit parameter =	0.324635
σ* of parameter =	0.101231
T* =	3.206876
p-value =	0.0225

$$T^* = \frac{\hat{\beta}_1 - 0}{\sigma^*_{(\beta_1)}} = \frac{0.3246}{0.1012} = 3.207$$

- **These calculations give us an approximate p-value = 0.0025 (unmodified) and p-value = 0.0225 (modified)**
- **In either case, we can reject the null hypothesis, which shows that β_1 is significant at an α value of 0.05**

** **'Modified Residuals' are modified by the method suggested on chart 35**

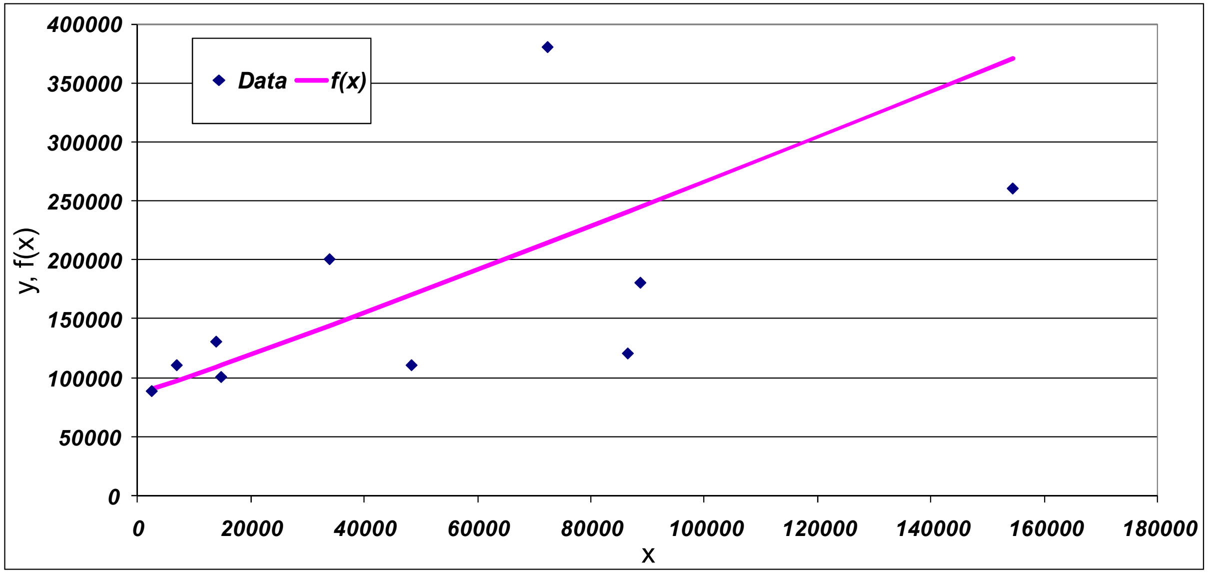
MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Another Form

- **Since only one of the coefficients was found to be significant, we should try a different equation form, such as:**

$$y = \beta_0 + x^{\beta_1} = 86211.5 + \mathrm{x}^{1.05}$$

- **Bootstrapping is run again, showing that both these parameters are significant at an α value of 0.05, and, given this relationship, weight is a significant cost driver**



CRITICAL THINKING. SOLUTIONS DELIVERED.

SIG Test

- T.P. Anderson [Ref. 2] has been exploring this same problem
- There are some similar characteristics of both the SIG test and the bootstrap approach
  - Anderson intended to examine the "significance" of fit parameters independent of the underlying error distribution
  - His "SIG Test" works on any functional form

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Example 4: Data From Reference [2]

AUC ($K)	Weight	Qty	Perf	Type
$124	209	3	5.2	0
$126	106	4	16.7	0
$105	35	10	9.2	0
$337	255	6	9.6	0
$281	206	7	2.9	0
$131	109	12	4.9	0
$155	120	2	12.7	0
$46	27	5	3.4	0
$224	286	1	9.8	0
$233	257	4	36.8	0
$357	134	4	21.1	1
$240	149	5	3.3	1
$97	10	44	5.0	1
$134	32	1	25.9	1
$149	46	18	5.5	1
$422	253	4	18.3	1
$262	81	20	13.7	1
$126	37	18	4.7	1
$152	28	18	6.8	1
$331	183	2	5.6	1
$204	24	37	2.7	1
$59	6	20	6.5	1
$134	46	4	1.1	1
$232	108	3	20.8	1
$48	5	28	3.2	1
$281	80	28	11.7	1
$272	149	9	1.2	1
$142	39	8	8.8	1

$$AUC = a + b(Wt)^c (Qt)^d (Perf)^e f^{(Type)}$$

- **CER has a multiplicative error term**
- **Coefficients are found to be the following:**
  - ***a*** **= -79.65**
  - ***b*** **= 31.35**
  - ***c*** **= 0.366**
  - ***d*** **= 0.109**
  - ***e*** **= 0.0576**
  - ***f*** **= 1.44**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

What Do the SIG Test Results Mean?

- Details can be found in [Ref. 2]
- Anderson ran through the SIG test for all 6 parameters
- The author suggests the rule of thumb that if the $|SIG_{Mean}|$ or $|SIG_{SPE}|$ is less then 5% and the SIG_{Total} (defined in [Ref. 2]) is less then 10%, then the parameter and its corresponding cost driver are "insignificant"
- However, regarding the interpretation of the results of his test, he states, "This is still an open question, subject to interpretation, and more research is needed in order to answer it."

CRITICAL THINKING. SOLUTIONS DELIVERED.

Anderson's Results from SIG Test

Parameter	Full CER	Minus '*a*'	Minus '*b*'	Minus '*c*'	Minus '*d*'	Minus 'e'	Minus 'f'
a	-79.65	0.00	42.62	151.96	-59.99	-69.32	-58.03
b	31.35	3.33	1.00	9.03	40.37	31.60	38.25
c	0.37	0.67	0.86	0.00	0.35	0.39	0.35
d	0.11	0.23	0.24	-0.37	0.00	0.09	0.12
e	0.06	0.13	0.09	0.56	0.01	0.00	0.06
f	1.44	1.89	2.20	3.60	1.52	1.47	1.00
$f_{\bar{Y}}(x)$	245.08	266.68	240.51	179.69	215.70	235.55	246.32
SIG_{Mean}		8.81%	-1.86%	-26.68%	-11.99%	-3.89%	0.51%
SPE	21.36%	24.44%	24.70%	52.71%	26.04%	22.51%	32.21%
SIG_{SPE}		14.42%	15.63%	146.74%	21.92%	5.37%	50.79%
SIG_{Total}		23.23%	17.49%	173.42%	33.91%	9.26%	51.30%

- **The only parameter found to be insignificant is e**
- **$|SIG_{Mean}|$ is less then 5%, and SIG_{Total} is less then 10%**
- **$|SIG_{Mean}|$s of both *b* and *f* are less then 5%, but SIG_{Total} is far greater then 10%; therefore not all the "requirements" for insignificance are met**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Nullifying Coefficient *a*

- **Various seed values for Excel Solver give different SIG_{Mean} and SIG_{Total} values**

Parameter	Full CER	Trial 1	Trial 2	Trial 3	Trial 4
a	-79.65	0.00	0.00	0.00	0.00
b	31.35	26.36	1.65	2.57	3.33
c	0.37	0.39	0.75	0.69	0.67
d	0.11	0.01	0.32	0.27	0.23
e	0.06	-0.01	0.20	0.16	0.13
f	1.44	1.44	2.03	1.96	1.89
$f_{\bar{Y}}(x)$	245.08	204.71	300.25	278.22	266.68
SIG_{Mean}		-16.47%	22.51%	13.52%	8.81%
SPE	21.36%	29.33%	30.26%	26.27%	24.44%
SIG_{SPE}		37.28%	41.67%	22.97%	14.42%
SIG_{Total}		53.76%	64.18%	36.50%	23.23%

- **However, despite some weaknesses, the SIG test is a very quick and easy way to begin to examine the significance of CER fit parameters and cost drivers**

CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrapping Example 4

- **Going back to the bootstrap approach to evaluate the significance of the CER fit parameters of Example 4, set *B* = 494 and *R* = 200**
- **All parameters but *e* turn out to be significant**

Fit Parameters	Approximate p-value (Unmodified)	Approximate p-value (Modified)
a	0	0
b	0	0.0020
c	0	0
d	0.0081	0
e	0.1215	0.1883
f	0.0061	0.0020

- **Thus, the cost driver ‘Performance (Perf)’ can be concluded to be insignificant**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Bootstrap: Testing Coefficient e

Standard Residuals Resampled:

parameter =	0.057603
SE of parameter* =	0.03873
T* =	1.487302
p-value =	0.121457

Modified Residuals Resampled:**

parameter =	0.057603
SE of parameter* =	0.043065
T* =	1.337584
p-value =	0.188259

- **These results give an approximate p-value = 60/494 = 0.1215 (unmodified) or an approximate p-value = 93/494 = 0.1883 (modified)**
- **Thus we cannot reject the null hypothesis, supporting the result of the SIG test**
- **This example provides the ability to describe the probability that random sampling would give similar results**

** **'Modified Residuals' are modified by the method suggested on chart 35**

CRITICAL THINKING. SOLUTIONS DELIVERED.

Drawbacks of the Bootstrap

- **One notable weakness is that running a large number of simulations can be time-consuming**
  - **Solution: Initially perform a small number of bootstrap runs**
  - **If results are close to your desired value threshold (α), you can do more runs to determine a more "accurate" result**
- **Also, keep in mind that higher levels of confidence require percentiles in the tails of the distribution of the "t-like" table, where bootstrap approximations may not be as adequate – although they seemed to perform well at a significance level of 0.05**

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Summary

- The bootstrap is applicable no matter the CER form, and it also doesn't make any assumptions about the error distribution
- It appears to have performed well in OLS examples
- Bootstrap conclusions about the statistical inferences can be mapped to results reached by OLS methods, thus tracking to the results of classical theory
- Once OLS tracking was established, the bootstrap approach was applied to a few non-OLS examples

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Conclusions

- **The bootstrap approach to hypothesis testing is another method that can be utilized in determining the significance of fit parameters and their corresponding cost drivers**
- **The bootstrap approach offers more statistical rigor than the SIG Test by allowing approximations to statistical probabilities**
- **What remains inconclusive is the nature of a more appropriate adjustment to the residuals**

References

[1] Abraham, Bovas and Johannes Ledolter. Introduction to Regression Modeling, Duxbury Applied Series, 2006.

[2] Anderson, Timothy. "A Distribution-Free Measure of the Significance of CER Regression Fit Parameters Established Using GERM (General Error Regression Methods)," 2009 SCEA/ISPA Joint Annual Conference & Training Workshop, St. Louis, MO (June 2009): Department of the Navy Cost Analysis Symposium, Quantico VA (Sept 2008): 432 charts; Journal of Cost Analysis and Parametrics, Vol. 2, Issue 1 (Summer 2009): 7-22.

[3] Boos, Dennis. "Introduction to the Bootstrap World," Statistical Science, Vol. 18, No. 2 (2003): 168-174.

[4] Davison, A.C. and D.V. Hinkley. Bootstrap Methods and their Application, Cambridge University Press, 1997.

[5] Efron, Bradley and Robert Tibshirani. An Introduction to the Bootstrap, Chapman and Hall/CRC, 1993.

[6] Hall, Peter and Susan Wilson. "Two Guidelines for Bootstrap Hypothesis Testing," Biometrics, Vol. 47, No. 2 (1991): 757-762.

[7] Higgins, James. Introduction to Modern Nonparametric Statistics, Duxbury Advanced Series, 2004.

[8] Stine, Robert. "Bootstrap Prediction Intervals for Regression," Journal of the American Statistical Association, Vol. 89, No. 392 (December 1985):1026-1031.

MCR CRITICAL THINKING. SOLUTIONS DELIVERED.

Acronyms

- CER = Cost Estimating Relationship
- GERM = General Error Regression Method
- OLS = Ordinary Least Squares
- SE = Standard Error
- SIG = Significance
- ZMPE = Zero Percentage Bias, Minimum Percentage Error