

Figure 12: The Borrowed Slope method

In the above figure, an object with weight of $weight_o$ and cost of $\$Y_o$ is the analogy. This object lies off the best-known CER for reasons that are sensible, in accord with the direction of the offset, and for reasons that are shared by the system being estimated. For example, suppose the CER is based upon industry-wide data, but the analogy system was made by a factory that has known higher costs, and that this factory will make the system being estimated (the reader is requested to accept the example as reasonable, and for purposes of the illustration). Given that the estimator accepts these beliefs, the estimator would revise the CER so as to make it pass through the analogy point, retaining the slope of the CER.

Adjusting by borrowed slope is compared to adjusting by ratio in the Figure 13. As can be seen in Figure 13, there can be considerable difference between a borrowed slope adjustment and a ratio adjustment. In general we develop “bigger, faster” and the like, and the values of parameters are usually above those of the analogy, so we tend to over estimate with the ratio method.

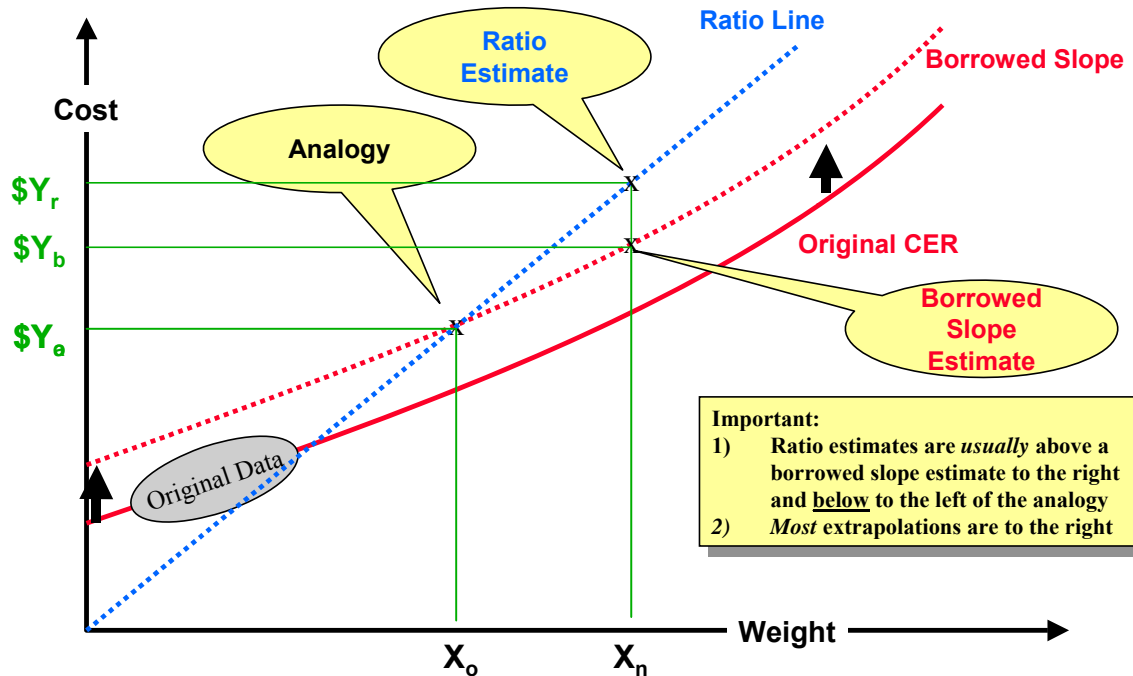


Figure 13: The borrowed slope method compared to the ratio method

Conclusions

We have come to believe that the y-intercept is a poorly understood part of cost estimation, and have observed a number of significant problems that arose because of failure to consider them. We discussed the 3 schools of thought with regard to y-intercepts. We discussed CERs, rates, metrics (thumb rules), and analogies with regard to the y-intercept. We discussed the implications of ignoring the y-intercept, and of including it in each of these areas, as well as the confusion that the y-intercept can cause in metrics and ratios. We discussed a method of adjusting analogies that takes the y-intercept into account.

We urge you to think about this. We are less offended at your holding beliefs (after all, we do!) than by sailing by this issue all unawares. Of course, if we had our way, we'd hope you believed in y-intercepts⁷, but in any event, think about it!

⁷ Clap if you believe in y-intercepts: In the second act of Walt Disney's "Peter Pan", Tinkerbell drinks poison that Peter is about to drink in order to save him. Peter turns to the audience and says, "Tinkerbell is going to die because not enough people believe in fairies. But if all of you clap your hands real hard to show that you do believe in fairies, maybe she won't die." We all started to clap. I clapped so long and so hard that my palms hurt. Then suddenly the actress playing Peter Pan turned to the audience and she said, "That wasn't enough. You did not clap hard enough. Tinkerbell is dead." We all started to cry. The actress stomped off stage and refused to continue the production. They had to lower the curtain. The ushers had to come help us out of the aisles and into the street. You hear that? CLAP LOUDER!

Appendix B

The Relational Correlation Method

A much more esoteric method is available, which borrows from bivariate normality and the geometry of regression. This method is available when there is no “trusted slope” to borrow.

Bivariate Normality

Let us first consider the case of bivariate normality.

Suppose X and Y are distributed $N(\mu_x, \sigma_x)$ and $N(\mu_y, \sigma_y)$

Suppose X and Y are jointly bivariate normal with correlation ρ

Then the graph of X and Y will appear as follows:

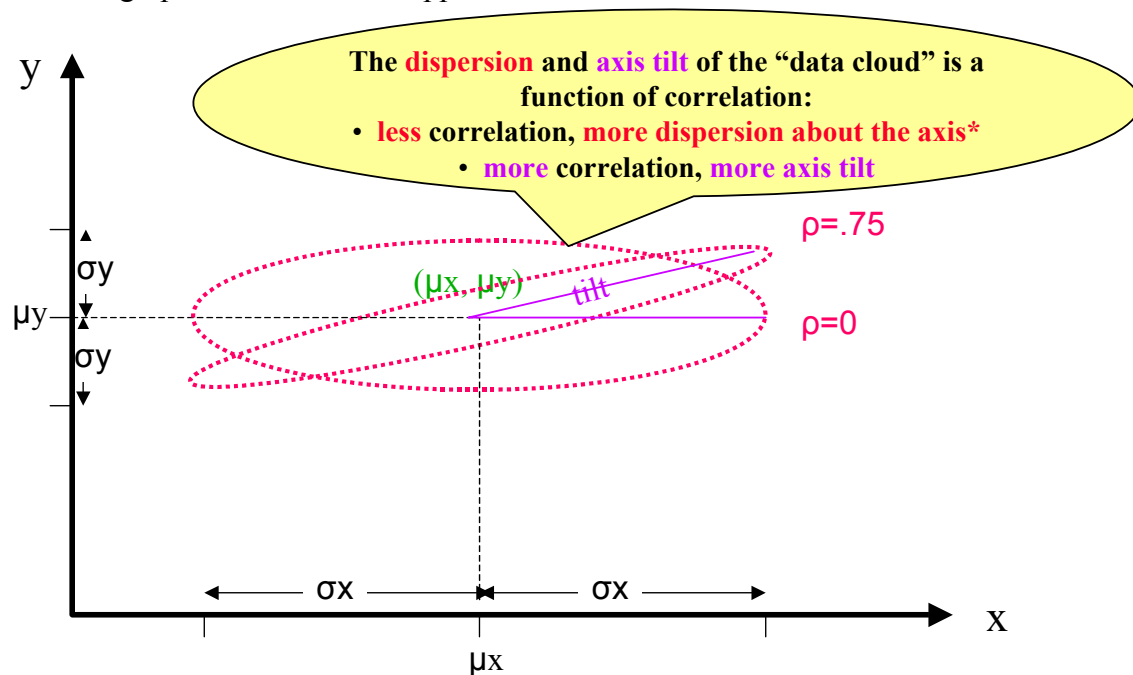


Figure 14: The borrowed slope method compared to the ratio method

We note that the “data cloud” will be shaped something like one of the two red dotted ovals, with 68.3% of the mass of the joint probability distribution inside the ovals which mark the 1-sigma curve, centered at the means of the two variables. The degree of correlation will affect the tilt of the oval. As noted on the illustration, the “fatness” of the data cloud is also connected to correlation.

For further background, we will now consider “the geometry of regression.” The below facts are known to mathematicians, but obscure, and not remembered in cost analysis:

For any two jointly distributed variables, there is a regression line

The slope is:

$$m = \rho * (\sigma_y / \sigma_x)$$

The y intercept is:

$$b = \mu_y - \rho (\sigma_y / \sigma_x) * \mu_x$$

If the variables are joint bivariate normal, then ρ is the correlation coefficient. This is best seen by a series of graphics, which follow:

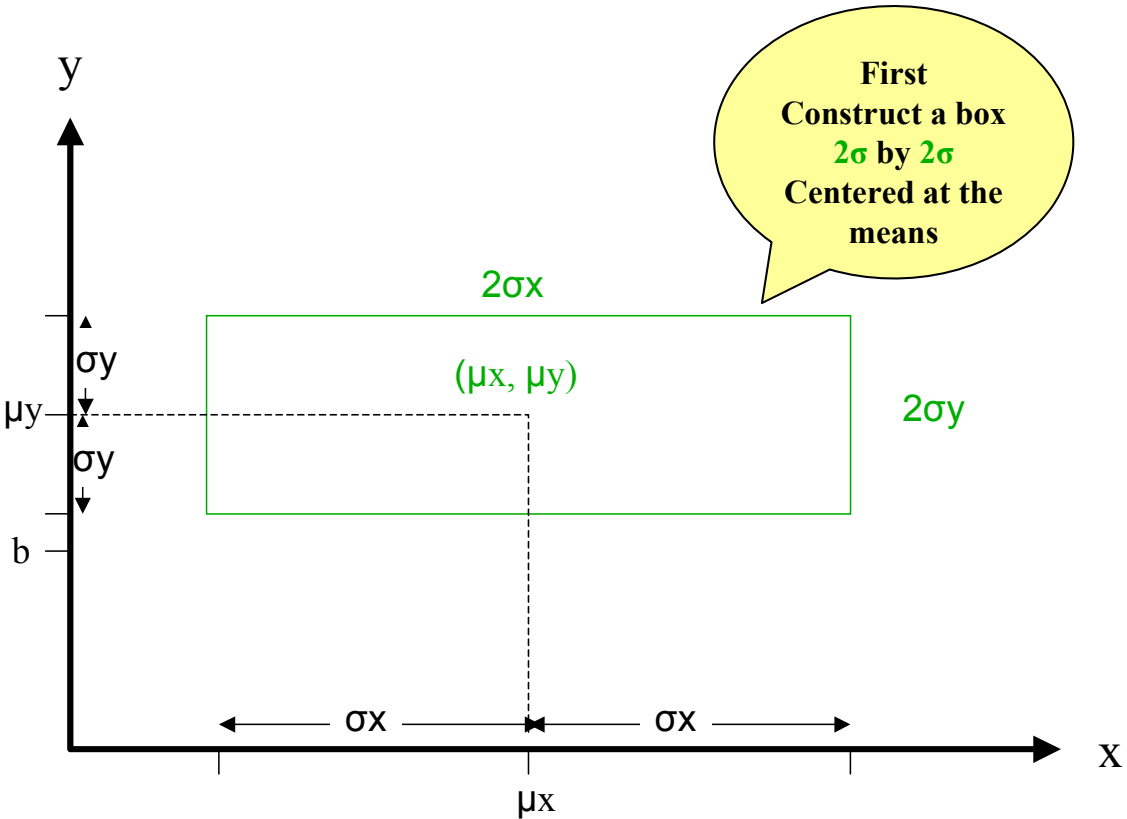


Figure 15: Bivariate normality – constructing the box

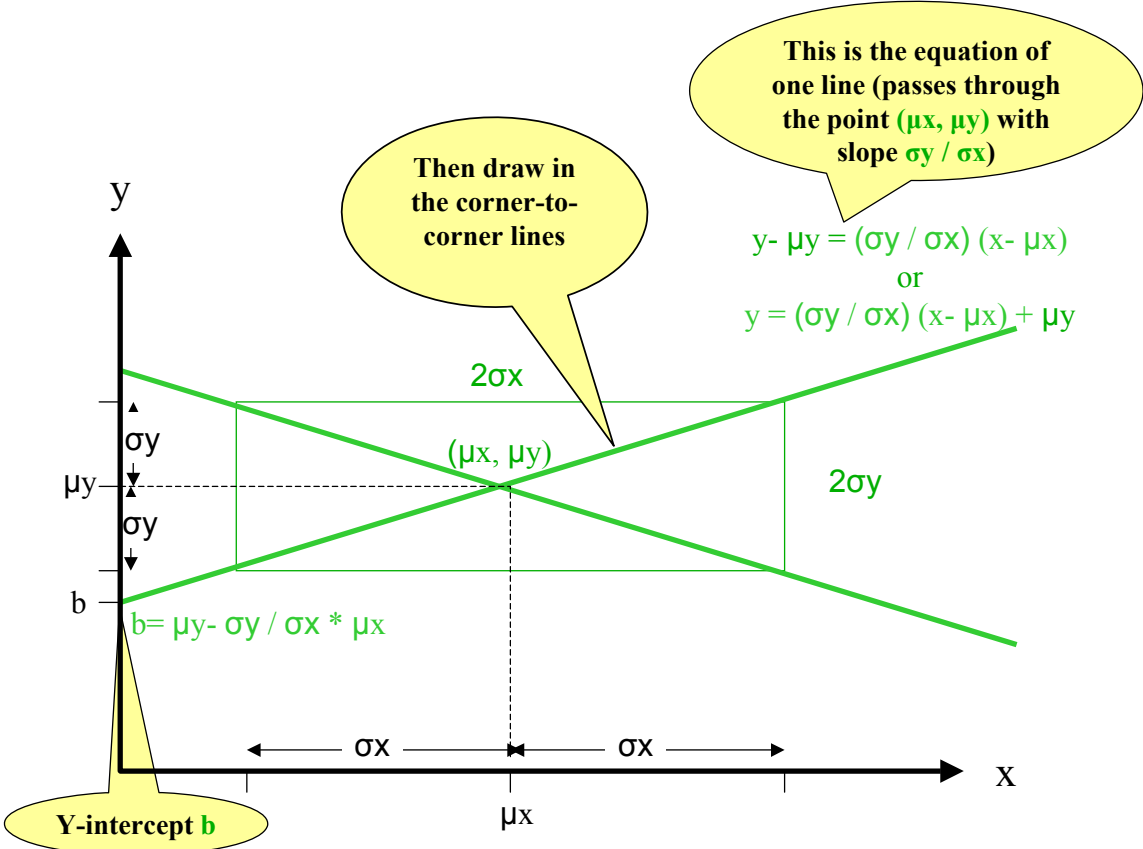


Figure 16: Bivariate normality – inserting the diagonals

Now, populate the box with data, shown here as the already illustrated ‘data clouds.’

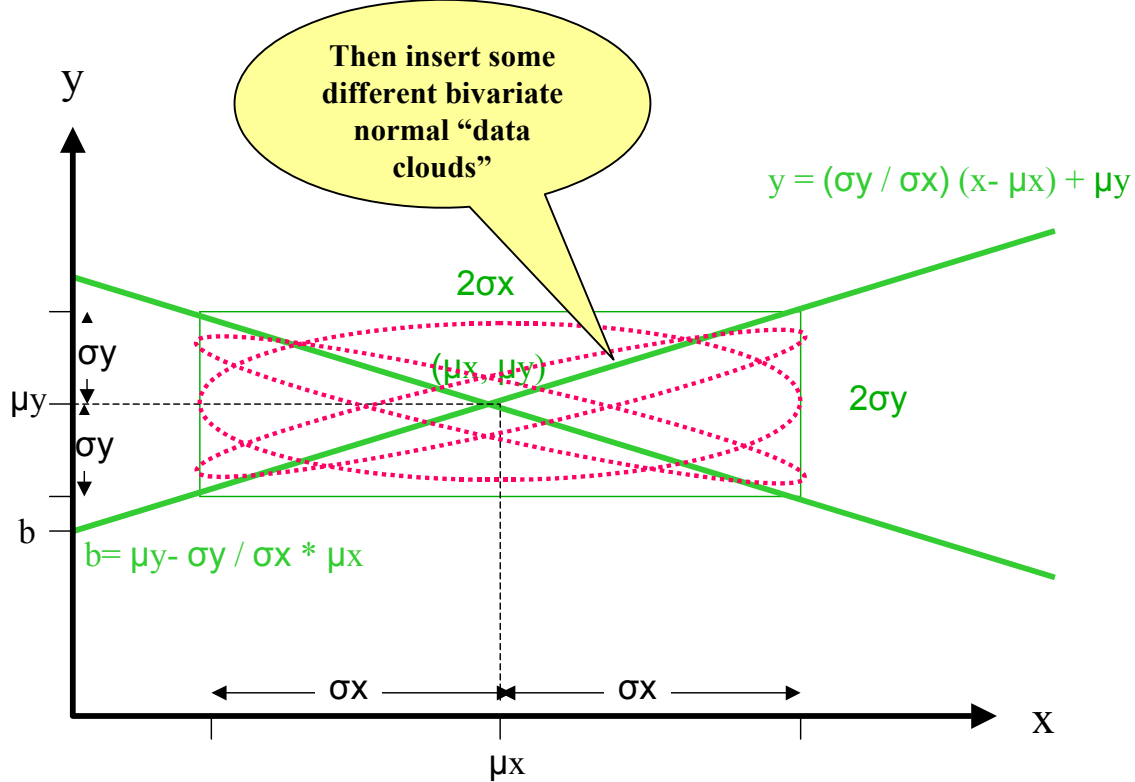


Figure 17: Bivariate normality – inserting the data

Now, let us look at the meaning of what we have constructed. And consider the geometry of regression.

The Geometry of Regression

We will look at the picture we have constructed and see what the geometry of regression tells us. We should note that two variables need not be jointly bivariate for the regression line to exist – the only addition to our ‘picture’ is that the slope of the regression line is affected by a parameter called ρ , and if the variables are jointly-distributed bivariate normals, this parameter ρ is their correlation.

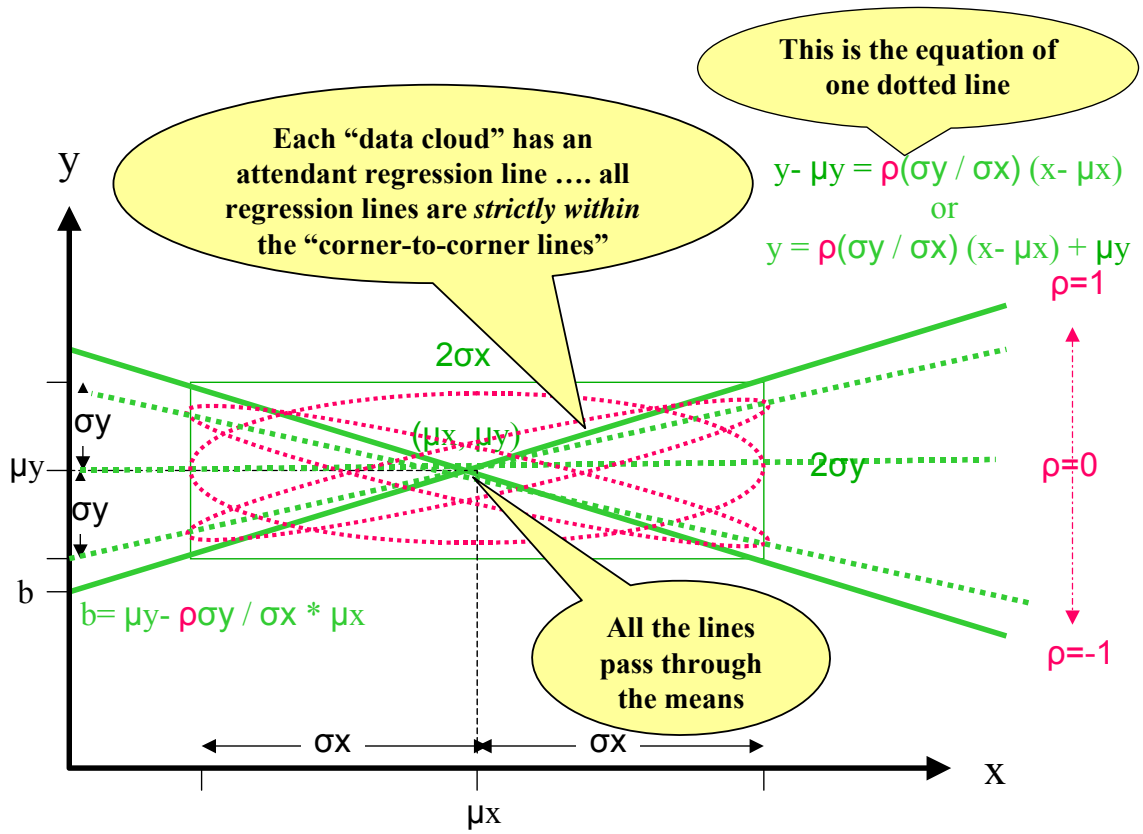


Figure 18: The geometry of regression - the implications

Now let us look at how the parameters ρ , σ_y and σ_x affect the regression line.

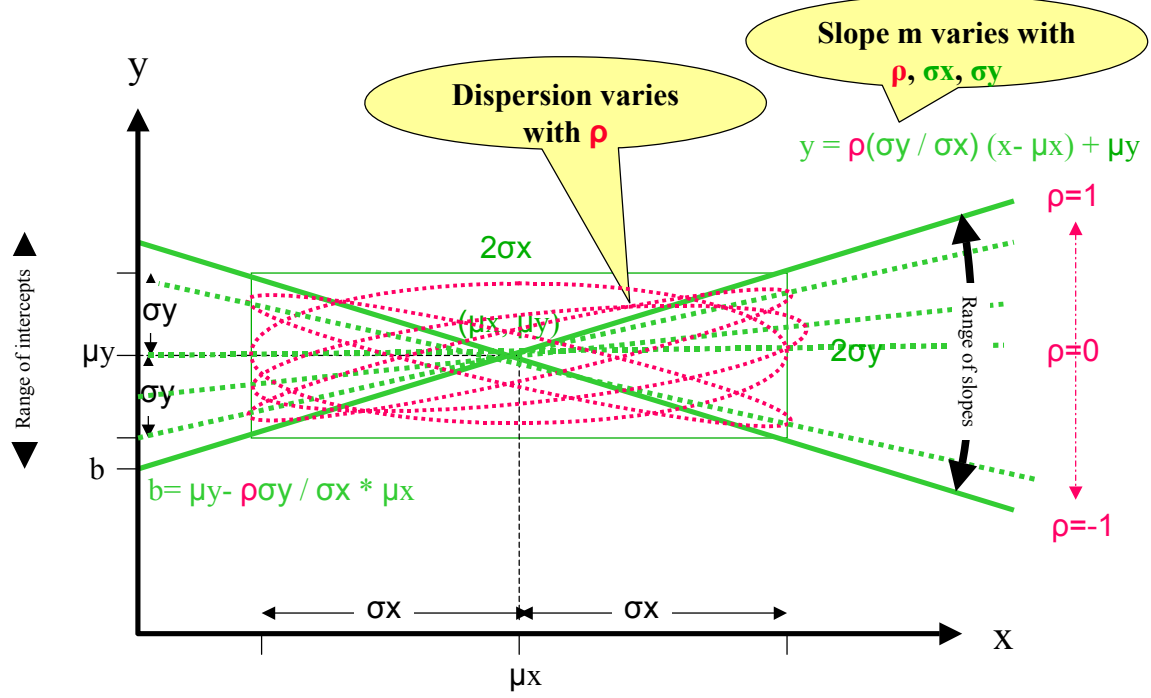


Figure 19: The geometry of regression - the effect of the parameters ρ , σ_y and σ_x

Now we will depart briefly from the case we are building and just look at the meaning of r^2 . We do this simply because we are already well-versed in the meaning of the geometry of regression, and we can see this important parameter with little additional work. We do not need it for our development, but it's a good thing to know and we may as well know it!

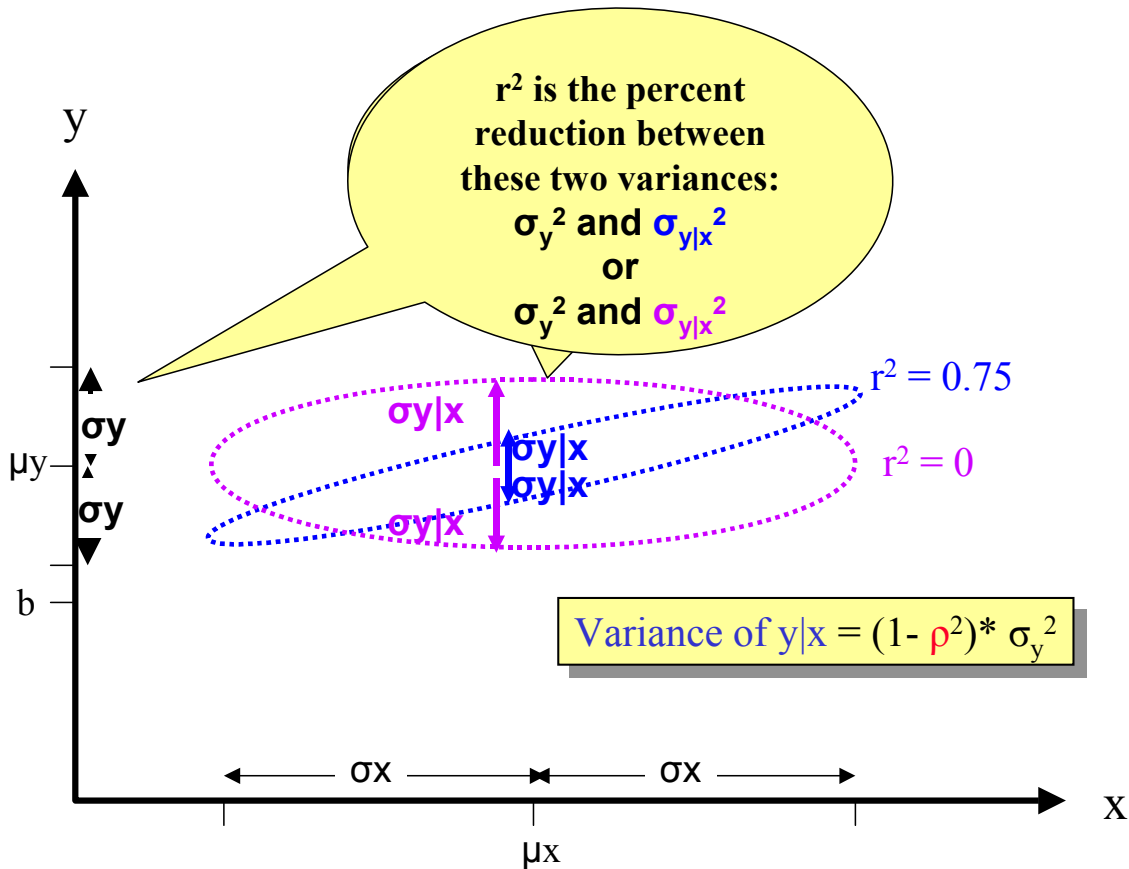


Figure 20: The meaning of r^2

Implications of the Geometry of Regression

For every regression with apparent slope m , there is an unseen equation with steeper slope m/ρ which is the unseen slope of the two variables, and with an unseen accompanying y intercept. Once we decide upon the means and the variances of x and y , the unseen line is fixed. Once we pick ρ , the regression line is fixed.

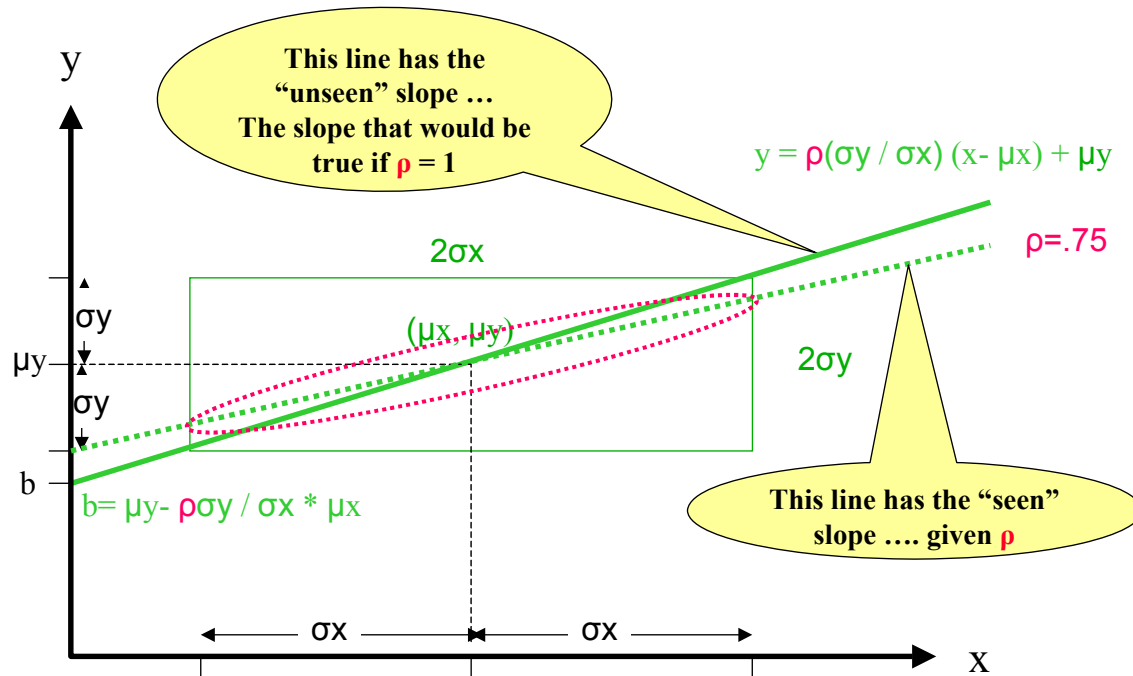


Figure 21: The implications of the geometry of regression

Implementing Relational Correlation for Analogies (and buildups)

For Single Point Analogies

- 1) Determine a reasonable (preferably historically-based) standard deviation for the x and y variable, e.g., to estimate ship repair parts as a function of tonnage you'll need:
 - a. The standard deviation for the analogy ship class repair parts cost
 - b. The standard deviation for the tonnage within the ship class
 - c. The standard deviation of repair parts for a single ship of the class
- 2) The ratio of 1 and 2 gives you the unseen slope
- 3) The relationship of 3 and 1 will yield r^2 (Variance of $y|x = (1 - \rho^2) * \sigma_y^2$)

For buildups

For buildups, do as above, but use an analogy for the values of the standard deviations, and apply it to your buildup using percents

We will now look at the next figure to see what this looks like.

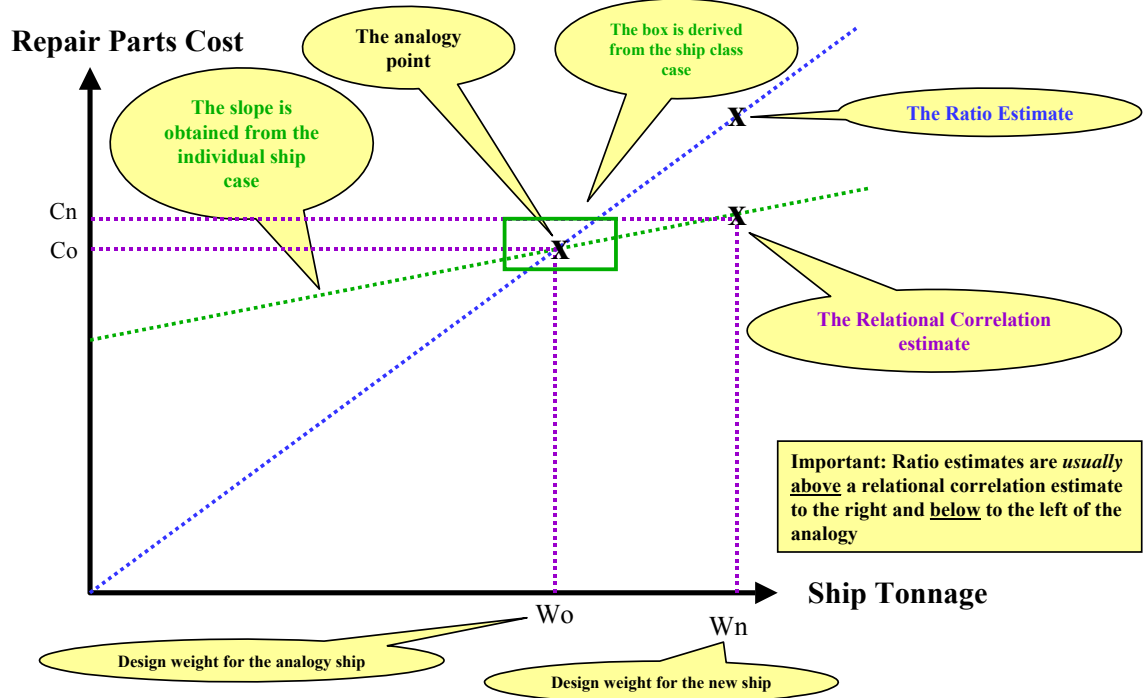


Figure 22: The relational correlation method pictorially

Appendix B

Definitions

Adjustments: Scaling of a cost by some physical, performance, or other such attribute. Scaling is usually (in practice) directly proportional to the attribute. Scaling parameters are usually countable or measurable and intuitively tied to cost.

Analogies: Estimation by assuming that the costs of a new system will be equal to (or similar to) the costs of a system that is similar in form. “Adjustments” are almost always made.

Buildups: Costed-out physical Bill of Materials (BOMs) and CAD-generated material lists and the like. We do not mean “buildups” consisting entirely of Staffing levels multiplied times duration. Such estimating techniques are little more than “engineering judgment” in fine detail. Buildups often include “adjustments” to allow for size differences.

Composite methods: A method that involves at least two of the three other types.

Parametric Estimates: Estimates made by developing statistical “Cost Estimating Relationships” (CERs) based on one or more parameter and cost Estimates involving parameters but not based on statistical analysis are more properly called either “adjusted analogies” or “adjusted buildups”.