9-2014

Generalized Least-Squares Regressions IV: Theory and Classification Using Generalized Means

Nataniel Greene
CUNY Kingsborough Community College

Recommended Citation
N. Greene. "Generalized Least-Squares Regressions IV: Theory and Classification Using Generalized Means," in Mathematics and Computers in Science and Industry, 2014, pp. 19-35.

Generalized Least-Squares Regressions IV: Theory and Classification Using Generalized Means

Nataniel Greene
Department of Mathematics and Computer Science
Kingsborough Community College, CUNY
2001 Oriental Boulevard, Brooklyn, NY 11235, USA
Email: ngreene.math@gmail.com

***Abstract*— The theory of generalized least-squares is reformulated here using the notion of generalized means. The generalized least-squares problem seeks a line which minimizes the average generalized mean of the square deviations in x and y. The notion of a generalized mean is equivalent to the generating function concept of the previous papers but allows for a more robust understanding and has an already existing literature. Generalized means are applied to the task of constructing more examples, simplifying the theory, and further classifying generalized least-squares regressions.**

***Keywords*—Linear regression, least-squares, orthogonal regression, geometric mean regression, generalized least-squares, generalized mean square regression.**

I. Overview

Ordinary least-squares regression suffers from a fundamental lack of symmetry: the regression line of y given x and the regression line of x given y are not inverses of each other. Ordinary least-squares $y|x$ regression minimizes the average square deviation between the data and the line in the y variable and ordinary least-squares $x|y$ minimizes the average square deviation between the data and the line in the x variable. A theory of generalized least-squares was described by this author for minimizing the average of a symmetric function of the square deviations in both x and y variables [7,8,9]. The symmetric function was referred to as a generating function for the particular regression method.

This paper continues the development of the theory of generalized least-squares, reformulated using the notion of generalized means. The generalized least-squares problem described here seeks a line which minimizes the average generalized mean of the square deviations in x and y. The notion of a generalized mean is equivalent to the generating function concept of the previous papers but allows for a more robust understanding and has an already existing literature.

It is clear from the name that geometric mean regression (GMR) seeks a line which minimizes the average geometric mean of the square deviations in x and y. Orthogonal regression seeks a line which minimizes the average harmonic mean of the square deviations in x and y. Therefore it is also called harmonic mean regression (HMR). Arithmetic mean regression (AMR) seeks a line which minimizes the average arithmetic mean of the square deviations in x and y and was called Pythagorean regression previously. Here, logarithmic, Heronian, centroidal, identric, Lorentz, and root mean square regressions are described for the first time. Ordinary least-squares regression is shown here to be equivalent to minimum or maximum mean regression. Regressions based on weighted arithmetic means of order α and weighted geometric means of order β are explored. The weights α and β parameterize all generalized mean square regression lines lying between the two ordinary least-squares lines.

Power mean regression of order p offers a particularly simple framework for parameterizing all the generalized mean square regressions previously described. The p-scale has fixed numerical values corresponding to many known special means. All the symmetric regressions discussed in the previous papers are power mean regressions for some value of p. Ordinary least-squares corresponds to $p = \pm\infty$. The power mean is one example of a generalized mean whose free parameter unites a variety of special means as subcases. Other generalized means which do the same include: the Dietel-Gordon mean of order r, Stolarsky's mean of order s, and Gini's mean of order t. There are also two-parameter means due to Stolarsky and Gini. Regression formulas based on all these generalized means are worked out here for the first time.

II. Regressions Based on Generalized Means

Generalized means are applied to the task of constructing more examples, simplifying the theory, and further classifying generalized least-squares regressions.

A. Axioms of a Generalized Mean

The axioms of a generalized mean presented here are drawn from Mays [13] and also from Chen [2].

Definition 1: A function $M(x, y)$ defines a generalized mean for $x, y > 0$ if it satisfies Properties 1-5 below. If it satisfies Property 6 it is called a homogenous generalized mean. The properties are:

1. (Continuity) $M(x, y)$ is continuous in each variable.
2. (Monotonicity) $M(x, y)$ is non-decreasing in each variable.

3. (Symmetry) $M(x,y) = M(y,x)$.
4. (Identity) $M(x,x) = x$.
5. (Intermediacy) $\min(x,y) \le M(x,y) \le \max(x,y)$.
6. (Homogeneity) $M(tx,ty) = tM(x,y)$ for all $t > 0$.

All known means are included in this definition. All the means discussed in this paper are homogeneous. The reader can verify that the weighted arithmetic mean or convex combination of any two generalized means is a generalized mean. The weighted geometric mean of any two generalized means is a generalized mean. More generally, the generalized mean of any two generalized means is itself a generalized mean.

The equivalence of generalized means and generating functions is now demonstrated.

Theorem 1: Let $M(x,y)$ be any generalized mean, then $\psi(x,y) = M(x^2,y^2)$ is the generating function for a corresponding generalized symmetric least-squares regression. Let $\psi(x,y)$ be any generating function, then $M(x,y) = \psi\left(\sqrt{|x|},\sqrt{|y|}\right)$ defines a generalized mean. The weight function is given by $g(b) = \psi\left(1,\frac{1}{b}\right) = M\left(1,\frac{1}{b^2}\right)$.

From here it is clear that the theory of generalized least-squares can be reformulated using generalized means. The general symmetric least-squares problem is re-stated as follows.

Definition 2: (The General Symmetric Least-Squares Problem) Values of a and b are sought which minimize an error function defined by

$$E = \frac{1}{N}\sum_{i=1}^{N} M\left((a+bx_i-y_i)^2, \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2\right) \quad (1)$$

where $M(x,y)$ is any generalized mean. The solution to this problem is called generalized mean square regression.

Definition 3: (The General Weighted Ordinary Least-Squares Problem) Values of a and b are sought which minimize an error function defined by

$$E = g(b)\cdot\frac{1}{N}\sum_{i=1}^{N}(a+bx_i-y_i)^2 \quad (2)$$

or using Ehrenberg's formula

$$\begin{aligned} E &= g(b)\left(\sigma_y^2\left(1-\rho^2\right)+\sigma_x^2\left(b-\rho\frac{\sigma_y}{\sigma_x}\right)^2\right. \\ &\quad \left. + \left(a+b\mu_x-\mu_y\right)^2\right) \end{aligned} \quad (3)$$

where $g(b)$ is a positive even function that is non-decreasing for $b < 0$ and non-increasing for $b > 0$.

The next theorem states that every generalized mean square regression problem is equivalent to a weighted ordinary least-squares problem with weight function $g(b)$.

Theorem 2: The general symmetric least-squares error function can be written equivalently as

$$E = g(b)\cdot\frac{1}{N}\sum_{i=1}^{N}(a+bx_i-y_i)^2 \quad (4)$$

or using Ehrenberg's formula

$$\begin{aligned} E &= g(b)\left(\sigma_y^2\left(1-\rho^2\right)+\sigma_x^2\left(b-\rho\frac{\sigma_y}{\sigma_x}\right)^2\right. \\ &\quad \left. + \left(a+b\mu_x-\mu_y\right)^2\right) \end{aligned} \quad (5)$$

where

$$g(b) = M\left(1,\frac{1}{b^2}\right). \quad (6)$$

Proof: Substitute $\frac{a}{b}+x_i-\frac{1}{b}y_i$ with $\frac{1}{b}(a+bx_i-y_i)$ and then use the homogeneity property:

$$\begin{aligned} E &= \frac{1}{n}\sum_{i=1}^{n} M\left((a+bx_i-y_i)^2, \frac{(a+bx_i-y_i)^2}{b^2}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}(a+bx_i-y_i)^2 M\left(1,\frac{1}{b^2}\right). \end{aligned}$$

Define

$$g(b) = M\left(1,\frac{1}{b^2}\right),$$

factor $g(b)$ outside of the summation and replace using Ehrenberg's formula. ■

The theory now continues unchanged with all the same theorems and formulas involving the weight function $g(b)$. It is reviewed now. To find the regression coefficients a and b take first order partial derivatives of the error function E_a and E_b and set them equal to zero. The result is $a = \mu_y - b\mu_x$ and the First Discrepancy Formula:

$$b-\rho\frac{\sigma_y}{\sigma_x} = -\frac{1}{2}\frac{g'(b)}{g(b)}\left(\left(b-\rho\frac{\sigma_y}{\sigma_x}\right)^2+\left(\frac{\sigma_y}{\sigma_x}\right)^2\left(1-\rho^2\right)\right). \quad (7)$$

To derive the slope equation for any generalized regression of interest, begin with the First Discrepancy Formula and substitute the specific expression for $g(b)$ into the formula, simplify and reset the equation equal to zero. What emerges is the specific slope equation. This is the procedure employed for all the slope equations presented in this paper.

Solving this equation for the discrepancy $b-\rho\frac{\sigma_y}{\sigma_x}$ using the quadratic formula yields the Second Discrepancy Formula:

$$b-\rho\frac{\sigma_y}{\sigma_x} = -\frac{g(b)}{g'(b)}\left(1-\sqrt{1-\kappa^2\left(\frac{\sigma_y}{\sigma_x}\right)^2\left(\frac{g'(b)}{g(b)}\right)^2}\right). \quad (8)$$

In order for the y-intercept a and slope b to minimize the error function the Hessian determinant $E_{aa}E_{bb}-(E_{ab})^2$ must be positive. Calculation of this determinant in the general case yields a function

$$G(b) = 2g'(b)/g(b) - g''(b)/g'(b) \quad (9)$$

called the indicative function. The Hessian determinant is positive provided that $G(b)\left(b-\rho\frac{\sigma_y}{\sigma_x}\right) > -1$. This differential equation for the weight function is solved to yield

$$g(b) = 1/(c+k\int\exp(-\int G(b)db)db). \quad (10)$$

B. The Exponential Equivalence Theorem Revisited

In the third paper of this series [9] the weight function

$$g_0(b) = \exp(-\gamma p_0 |b|) \tag{11}$$

was shown to generate all generalized least-squares regression lines as γ varies from 0 to 1. Here $p_0 = 1/\left(|\kappa| \frac{\sigma_y}{\sigma_x}\right)$ and $\kappa = \operatorname{sgn}\rho\sqrt{1-\rho^2}$ is the coefficient of scatter. The regression lines have the simple form described in the next theorem.

Theorem 3: (Fundamental Generalized Least-Squares Formula) The generalized least-squares regression line is given by $y = a + bx$ where

$$a = \mu_y - b\mu_x \tag{12}$$

and

$$b = (\rho + \lambda\kappa)\frac{\sigma_y}{\sigma_x}. \tag{13}$$

The parameters γ and λ are related by the equations

$$\lambda = \frac{\gamma}{1+\sqrt{1-\gamma^2}} \tag{14}$$

and

$$\gamma = \frac{2\lambda}{\lambda^2+1}. \tag{15}$$

Proof: Substitute $g_0(b)$ into the Second Discrepancy Formula and simplify. ∎

Trigonometric formulas for the parameters γ and λ are now derived. It is shown that γ and λ are functions of a more fundamental parameter ϕ in $\left[0, \frac{\pi}{2}\right]$ or an equivalent parameter δ in $[0,1]$ satisfying $\phi = \delta \cdot \frac{\pi}{2}$.

Lemma 4: If $\phi = \delta \cdot \frac{\pi}{2}$ with δ in $[0,1]$ then

$$\gamma = \sin\phi = \sin\left(\delta \cdot \frac{\pi}{2}\right) \tag{16}$$

and

$$\lambda = \tan\frac{\phi}{2} = \tan\left(\delta \cdot \frac{\pi}{4}\right). \tag{17}$$

Proof: Substitute $\lambda = \tan\frac{\phi}{2}$ and $\gamma = \sin\phi$ into the formula relating γ and λ and verify that $\sin\phi = 2\tan\frac{\phi}{2}/\left(\tan^2\frac{\phi}{2}+1\right)$. ∎

Theorem 5: The parameters γ and λ are related by the trigonometric equations

$$\lambda = \tan\left(\frac{1}{2}\sin^{-1}\gamma\right) \tag{18}$$

and

$$\gamma = \sin\left(2\tan^{-1}\lambda\right). \tag{19}$$

Proof: Solve for ϕ in each of the two equations and substitute into the other equation. ∎

Corollary 6: The following inequality holds between the parameters.

$$0 \le \lambda \le \delta \le \gamma \le 1 \tag{20}$$

with equality only when all three parameters equal 0 or 1.

Proof: This is due to the concavity of the functions $\gamma(\delta)$ and $\lambda(\delta)$ with $\gamma(\delta)$ lying above the diagonal line $\gamma = \delta$ and $\lambda(\delta)$ lying below the diagonal line $\lambda = \delta$. ∎

Corollary 7: Let $\rho = \cos\theta$ for θ in $[0,\pi]$. Then GMR corresponds to $\phi = \theta$ with $\gamma = \sin\theta$ and $\lambda = \tan\frac{\theta}{2}$. OLS $x|y$ corresponds to $\phi = 2\theta$ with $\gamma = \sin 2\theta$ and $\lambda = \tan\theta$.

Proof: It was shown [8] that for GMR, $\gamma = \kappa$ and $\lambda = \kappa/(1+|\rho|)$. For OLS $x|y$, $\gamma = 2\kappa\rho$ and $\lambda = \kappa/\rho$. ∎

To summarize, γ and λ are interrelated using simple trigonometric formulas. The parameter δ (or equivalently ϕ) generates both γ and λ using simple trigonometric formulas. As δ varies between 0 and 1, ϕ varies between 0 and $\frac{\pi}{2}$ and γ and λ vary between 0 and 1. All possible generalized least-squares lines are thereby generated. The case $\delta = 0$ corresponds to the ordinary least-squares line and $\delta = 1$ corresponds to the extremal line.

C. An Admissibility Criterion for Generalized Least-Squares

In order for the pair (a,b) to minimize the error function, the expression inside the radical of the Second Discrepancy Formula must be non-negative. This means that the following inequality must be satisfied:

$$\left(\frac{g'(b)}{g(b)}\right)^2 \le 1/\left(\left(\frac{\sigma_y}{\sigma_x}\right)^2\kappa^2\right). \tag{21}$$

For specific choices of $g(b)$ and fixed ρ, this inequality can be solved for b. The result is an admissibility condition on the slope b: slopes which exceed a certain magnitude become inadmissible in the sense that they no longer lie on the spectrum between OLS $y|x$ and the extremal line and they no longer minimize the error function.

According to the exponential equivalence theorem, $g(b) = \exp(-\gamma p_0 |b|)$ for some γ in $[0,1]$. This inequality is therefore equivalent to the statement $\gamma \le 1$, which is in turn equivalent to $\lambda \le 1$ and $\delta \le 1$. Numerically, it suffices to compute any one of these parameters and check that it does not exceed 1. If it exceeds 1 then the line is inadmissible for the given data: the slope and y-intercept will fail to minimize the error function.

D. A Restriction on Ordinary Least-Squares $x|y$

It is known that as ρ^2 tends to zero, the OLS $x|y$ line becomes increasingly perpendicular to both the OLS $y|x$ line and to the data cloud itself. This suggests that for some critical value of ρ^2, OLS $x|y$ ceases to be a model for the data. The above admissibility criterion when applied to ordinary least-squares $x|y$ produces this critical value which may not have been known. It is stated in the next theorem.

Theorem 8: The ordinary least-squares $x|y$ line

$$y = \left(\frac{1}{\rho}\frac{\sigma_y}{\sigma_x}\right)x + \left(\mu_y - \frac{1}{\rho}\frac{\sigma_y}{\sigma_x}\mu_x\right) \tag{22}$$

is admissible provided that

$$\frac{1}{2} \le \rho^2 \le 1. \tag{23}$$

In this case, the slope and y-intercept of the line minimize the error function.

ISBN: 978-1-61804-247-7

Proof: Substitute $g(b) = \frac{1}{b^2}$ and simplify. Replace b with $\frac{1}{\rho}\frac{\sigma_y}{\sigma_x}$ and solve the inequality for ρ^2. ∎

This restriction was already stated in a chart in the previous paper [9] but without elaboration. To summarize, for data with $\rho^2 < \frac{1}{2}$ the OLS $x|y$ line is inadmissible: its slope and y-intercept will fail to minimize the error function and its graph no longer passes through the data cloud sufficiently for it to model the data.

E. A Symmetry Property of Least-Squares Lines

Without loss of generality, all the slope equations presented in this paper are for least-squares lines with positive slope, corresponding to data with a positive correlation coefficient. This restriction makes the slope equations less cumbersome to write. Least-squares lines with negative slope, corresponding to data with a negative correlation coefficient can be computed from a positive slope equation as follows.

1. Reflect the data across the x-axis: $(x_i, y_i) \to (x'_i, y'_i)$ where $x'_i = x_i$ and $y'_i = -y_i$.

2. Compute the regression line for the reflected data. This corresponds to replacing ρ with $|\rho|$ and b with $b \operatorname{sgn} b = b \operatorname{sgn} \rho$ in the original slope equation and solving for the relevant *positive* solution for b'. Note that $\mu_{x'} = \mu_x$, $\mu_{y'} = -\mu_y$, $\sigma_{x'} = \sigma_x$, and $\sigma_{y'} = \sigma_y$. Therefore the y-intercept for the reflected data is given by $a' = \mu_{y'} - b'\mu_{x'}$.

3. The regression line for the original data has slope $b = -b'$ and y-intercept $a = -a'$, which is the same as writing $a = \mu_y - b\mu_x$.

In practice, to write a general slope equation which is true for all b and ρ, take the positive slope equation presented here, replace b with $(\operatorname{sgn} b)(b) = (\operatorname{sgn} \rho)(b)$ and replace ρ with $|\rho|$ and simplify. For example, in the case of square perimetric regression (SPR), use Formula 77 with $p = \frac{1}{2}$ and obtain the positive slope equation:

$$b^3 - \rho\frac{\sigma_y}{\sigma_x}b^2 + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \tag{24}$$

Replace b with $(\operatorname{sgn} b)(b) = (\operatorname{sgn} \rho)(b)$ and replace ρ with $|\rho|$ in this equation and obtain the general form which appears in the previous work [7], namely:

$$b^3 - \rho\frac{\sigma_y}{\sigma_x}b^2 + |\rho|\frac{\sigma_y}{\sigma_x}b - \operatorname{sgn}\rho\left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \tag{25}$$

Similarly, for squared harmonic mean regression (SHR), again use Formula 77 with $p = -\frac{1}{2}$ and obtain the positive equation:

$$b = \frac{\sigma_y}{\sigma_x} \cdot \frac{\frac{\sigma_y}{\sigma_x} + \rho}{\rho\frac{\sigma_y}{\sigma_x} + 1} \tag{26}$$

Again make the above replacements and obtain the form which appears in the previous work:

$$b = \operatorname{sgn}\rho\frac{\sigma_y}{\sigma_x} \cdot \frac{\frac{\sigma_y}{\sigma_x} + |\rho|}{|\rho|\frac{\sigma_y}{\sigma_x} + 1}. \tag{27}$$

The focus of this paper is to derive slope equations for all the known special means, as well as the generalized means with free parameters which appear in the literature. The free parameters are used to classify the known special cases.

F. A Naming Convention for Generalized Regressions: XMR

A naming convention is adopted here for generalized mean regression following what is already done for geometric mean regression: take the letter already in use for denoting the given mean and add 'MR' after it. If X is the letter already in use to denote a particular generalized mean, then XMR is the corresponding regression. For example, the geometric mean is denoted by 'G', therefore the corresponding geometric mean regression is denoted GMR. What was called Pythagorean regression earlier, is now called AMR here since it minimizes the arithmetic mean of the square deviations in x and y. Similarly, orthogonal regression is also denoted here as HMR since it minimizes the average harmonic mean of the square deviations in x and y. Where the name is already established, as in the case of orthogonal regression, the convention here simply provides an alternative name.

G. Arithmetic, Geometric and Harmonic Mean Regression

The weight functions and slope equations are reviewed for arithmetic, geometric and harmonic regression.

For AMR, which was called Pythagorean regression previously, the mean is

$$M(x, y) = \frac{x + y}{2}, \tag{28}$$

the weight function is

$$g(b) = \frac{1}{2}\left(1 + \frac{1}{b^2}\right), \tag{29}$$

and the slope equation is

$$b^4 - \rho\frac{\sigma_y}{\sigma_x}b^3 + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \tag{30}$$

Here the user selects the unique real solution that is greater than $\rho\frac{\sigma_y}{\sigma_x}$ when ρ is positive or less than $\rho\frac{\sigma_y}{\sigma_x}$ when ρ is negative.

For GMR, the mean is

$$M(x, y) = \sqrt{xy}, \tag{31}$$

the weight function is

$$g(b) = \frac{1}{|b|}, \tag{32}$$

and the slope equation is

$$b^2 - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0, \tag{33}$$

so that

$$b = \operatorname{sgn}\rho\frac{\sigma_y}{\sigma_x}. \tag{34}$$

For HMR, which is the same as orthogonal regression, the mean is

$$M(x,y) = \frac{2xy}{x+y}, \tag{35}$$

the weight function is

$$g(b) = \frac{2}{1+b^2}, \tag{36}$$

and the slope equation is

$$\rho\frac{\sigma_y}{\sigma_x}b^2 + \left(1 - \left(\frac{\sigma_y}{\sigma_x}\right)^2\right)b - \rho\frac{\sigma_y}{\sigma_x} = 0. \tag{37}$$

The quadratic formula yields the well-known solutions.

H. Ordinary Least-Squares is Minimum or Maximum Mean Regression

It is now shown that ordinary least-squares regression is also generated by a generalized mean.

Theorem 9: Let $y = a + bx$ be a regression line generated by the minimum mean or the maximum mean. If $|b| \le 1$, then $M(x,y) = \min(x,y)$ corresponds to OLS $y|x$ regression and $M(x,y) = \max(x,y)$ corresponds to OLS $x|y$ regression. If $|b| > 1$, then $M(x,y) = \min(x,y)$ corresponds to OLS $x|y$ regression and $M(x,y) = \max(x,y)$ corresponds to OLS $y|x$ regression.

Proof: It is a geometric fact that the vertical distance between a point and a line is always less than or equal to the horizontal distance when $|b| \le 1$ and greater than the horizontal distance when $|b| > 1$. The error function is given by

$$E = \frac{1}{N}\sum_{i=1}^{N} M\left((a+bx_i-y_i)^2, \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2\right) \tag{38}$$

For minimum mean regression $M(x,y) = \min(x,y)$ and

$$\min\left((a+bx_i-y_i)^2, \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2\right) = \begin{cases} (a+bx_i-y_i)^2, & |b| \le 1 \text{ (OLS } y|x) \\ \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2, & |b| > 1 \text{ (OLS } x|y) \end{cases}.$$

Similarly, for maximum mean regression, $M(x,y) = \max(x,y)$ and

$$\max\left((a+bx_i-y_i)^2, \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2\right) = \begin{cases} \left(\frac{a}{b}+x_i-\frac{1}{b}y_i\right)^2, & |b| \le 1 \text{ (OLS } x|y) \\ (a+bx_i-y_i)^2, & |b| > 1 \text{ (OLS } y|x) \end{cases}.$$

Alternatively, the result is obtained by examining the weight functions. Recall that OLS $y|x$ corresponds to $g(b) = 1$ and OLS $x|y$ corresponds to $g(b) = \frac{1}{b^2}$. For minimum mean regression the weight function is given by

$$g_{\min}(b) = \min\left(1, \frac{1}{b^2}\right) = \begin{cases} 1, & |b| \le 1 \text{ (OLS } y|x) \\ \frac{1}{b^2}, & |b| > 1 \text{ (OLS } x|y) \end{cases}.$$

For maximum mean regression the weight function is given by:

$$g_{\max}(b) = \max\left(1, \frac{1}{b^2}\right) = \begin{cases} \frac{1}{b^2}, & |b| \le 1 \text{ (OLS } x|y) \\ 1, & |b| > 1 \text{ (OLS } y|x) \end{cases}.$$

■

III. More Regression Examples Based on Known Special Means

This section examines regressions based on several special means that are already known from the references. They are: the logarithmic, Heronian, centroidal, contraharmonic, root mean square and identric means. For each case, the weight function and the corresponding slope equation are derived.

A. Logarithmic Mean Regression (LMR)

The logarithmic mean is given by

$$L(x,y) = \begin{cases} (y-x)/(\ln y - \ln x), & x \ne y \\ x, & x = y. \end{cases} \tag{39}$$

Therefore, the weight function is given by

$$g(b) = L\left(1, \frac{1}{b^2}\right) = \frac{b^2-1}{b^2\ln b^2}. \tag{40}$$

To find b one must solve the slope equation:

$$\begin{aligned} \ln\left(b^2\right) = {} & \left\{b^4 - 2\rho\frac{\sigma_y}{\sigma_x}b^3 + \left(\left(\frac{\sigma_y}{\sigma_x}\right)^2 - 1\right)b^2 \right. \\ & \left. +2\rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2\right\} / \\ & \left\{b^4 - \rho\frac{\sigma_y}{\sigma_x}b^3 - \rho\frac{\sigma_y}{\sigma_x}b + \left(\frac{\sigma_y}{\sigma_x}\right)^2\right\}. \end{aligned} \tag{41}$$

It will be seen further on that it is useful to define here a second logarithmic mean

$$L_2(x,y) = \begin{cases} xy(\ln y - \ln x)/(y-x), & x \ne y \\ x, & x = y \end{cases} \tag{42}$$

and when necessary the first logarithmic mean is denoted by $L_1(x,y) = L(x,y)$. The second logarithmic mean satisfies the relation $L_2(x,y) = (G(x,y))^2/L(x,y)$. The weight function for the second logarithmic mean is given by

$$g(b) = \frac{\ln\left(b^2\right)}{b^2-1}, \tag{43}$$

and the slope equation is given by

$$\begin{aligned} \ln\left(b^2\right) = {} & -\left\{b^4 - 2\rho\frac{\sigma_y}{\sigma_x}b^3 + \left(\left(\frac{\sigma_y}{\sigma_x}\right)^2 - 1\right)b^2 \right. \\ & \left. +2\rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2\right\} / \\ & \left\{\rho\frac{\sigma_y}{\sigma_x}b^4 - \left(1 + \left(\frac{\sigma_y}{\sigma_x}\right)^2\right)b^2 + \rho\frac{\sigma_y}{\sigma_x}b\right\}. \end{aligned} \tag{44}$$

B. Heronian Mean Regression (NMR)

The Heronian mean is given by

$$N(x,y)=\frac{1}{3}\left(x+\sqrt{xy}+y\right). \quad (45)$$

The weight function is given by

$$g(b)=N\left(1,\frac{1}{b^2}\right)=\frac{1}{3}\left(1+\frac{1}{|b|}+\frac{1}{b^2}\right). \quad (46)$$

To find b one must solve a quartic polynomial equation:

$$\begin{aligned}0 &= b^4+\left(\frac{1}{2}-\rho\frac{\sigma_y}{\sigma_x}\right)b^3 \\ &+\left(\rho\frac{\sigma_y}{\sigma_x}-\frac{1}{2}\left(\frac{\sigma_y}{\sigma_x}\right)^2\right)b-\left(\frac{\sigma_y}{\sigma_x}\right)^2. \quad (47)\end{aligned}$$

C. Contraharmonic Mean Regression (CMR)

The contraharmonic mean is given by

$$C(x,y)=\frac{x^2+y^2}{x+y}. \quad (48)$$

The weight function is given by

$$g(b)=1+\frac{1-b^2}{b^4+b^2}, \quad (49)$$

and the slope equation is a polynomial equation of degree eight:

$$\begin{aligned}0 &= b^8-\rho\frac{\sigma_y}{\sigma_x}b^7+2b^6-3\rho\frac{\sigma_y}{\sigma_x}b^5 \\ &+\left(\frac{\sigma_y}{\sigma_x}-1\right)b^4+3\rho\frac{\sigma_y}{\sigma_x}b^3-2\left(\frac{\sigma_y}{\sigma_x}\right)^2b^2 \\ &+\rho\frac{\sigma_y}{\sigma_x}b-\left(\frac{\sigma_y}{\sigma_x}\right)^2. \quad (50)\end{aligned}$$

D. Centroidal Mean Regression (TMR)

The centroidal mean is given by

$$T(x,y)=\frac{2}{3}\left(\frac{x^2+xy+y^2}{x+y}\right). \quad (51)$$

The centroidal mean is a convex combination of the contraharmonic mean and the harmonic mean: $T(x,y)=\frac{2}{3}C(x,y)+\frac{1}{3}H(x,y)$. It also satisfies the relation $T(x,y)=N\left(x^2,y^2\right)/A(x,y)$.

The weight function is given by

$$g(b)=T\left(1,\frac{1}{b^2}\right)=\frac{2}{3}\left(1+\frac{1}{b^2\left(b^2+1\right)}\right). \quad (52)$$

The slope equation is a polynomial equation of degree eight:

$$\begin{aligned}0 &= b^8-\rho\frac{\sigma_y}{\sigma_x}b^7+2b^6-2\rho\frac{\sigma_y}{\sigma_x}b^5 \\ &+2\rho\frac{\sigma_y}{\sigma_x}b^3-2\left(\frac{\sigma_y}{\sigma_x}\right)^2b^2-\left(\frac{\sigma_y}{\sigma_x}\right)^2. \quad (53)\end{aligned}$$

E. Root-Mean-Square Regression (RMR)

The root-mean-square is given by

$$R(x,y)=\left(\frac{x^2+y^2}{2}\right)^{1/2}. \quad (54)$$

The weight function is given by

$$g(b)=\frac{1}{\sqrt{2}}\left(1+\frac{1}{b^4}\right)^{1/2}. \quad (55)$$

The slope equation is a polynomial equation of degree six:

$$b^6-\rho\frac{\sigma_y}{\sigma_x}b^5+\rho\frac{\sigma_y}{\sigma_x}b-\left(\frac{\sigma_y}{\sigma_x}\right)^2=0. \quad (56)$$

F. Identric Mean Regression (IMR)

The identric mean is given by

$$I(x,y)=\begin{cases}\frac{1}{e}\cdot\left(\frac{x^x}{y^y}\right)^{1/(x-y)}, & x\neq y\\ x, & x=y.\end{cases} \quad (57)$$

The weight function is given by

$$g(b)=\frac{1}{e}\cdot b^{2/\left(b^2-1\right)}, \quad (58)$$

and the slope equation is given by

$$\begin{aligned}\ln\left(b^2\right) &= \left\{b^6-\rho\frac{\sigma_y}{\sigma_x}b^5-b^4\right. \\ &\left.+\left(\frac{\sigma_y}{\sigma_x}\right)b^2+\rho\frac{\sigma_y}{\sigma_x}b-\left(\frac{\sigma_y}{\sigma_x}\right)^2\right\}/ \\ &\left\{b^4-2\rho\frac{\sigma_y}{\sigma_x}b^3+\left(\frac{\sigma_y}{\sigma_x}\right)^2b^2\right\}. \quad (59)\end{aligned}$$

IV. Classifying Generalized Regressions Using Weighted Arithmetic Means

Perhaps the simplest generalized mean is a weighted arithmetic mean of order α.

Definition 4: The weighted arithmetic mean of order α is defined for $x<y$ by

$$M_\alpha(x,y)=(1-\alpha)x+\alpha y. \quad (60)$$

The standard arithmetic mean corresponds to $\alpha=\frac{1}{2}$. The weight function is given by

$$g_\alpha(b)=(1-\alpha)+\alpha\left(\frac{1}{b^2}\right) \quad (61)$$

and the corresponding slope equation is given by

$$(1-\alpha)b^4-(1-\alpha)\rho\frac{\sigma_y}{\sigma_x}b^3+\alpha\rho\frac{\sigma_y}{\sigma_x}b-\alpha\left(\frac{\sigma_y}{\sigma_x}\right)^2=0. \quad (62)$$

The indicative function is given by

$$G_\alpha(b)=\frac{(1-\alpha)3b^2-\alpha}{b\left((1-\alpha)b^2+\alpha\right)}. \quad (63)$$

It is clear that α continuously parameterizes a spectrum of generalized regression lines with OLS $y|x$ corresponding to

ISBN: 978-1-61804-247-7

$\alpha = 0$ and OLS $x|y$ corresponding to $\alpha = 1$. Due to intermediacy, every generalized mean lies between $x = \min(x, y)$ and $y = \max(x, y)$, corresponding to OLS $y|x$ and OLS $x|y$ respectively. Therefore every generalized mean-square regression can be assigned an effective parameter α in $[0, 1]$, classifying the line on the spectrum between OLS $y|x$ and OLS $x|y$. This is formalized in the next theorem.

Theorem 10: (Weighted Arithmetic Mean Equivalence Theorem) Let b be the slope of a generalized least-squares regression line lying between the two ordinary least-squares lines. Then the regression line is generated by an equivalent weighted arithmetic mean of order α in $[0, 1]$ where

$$\alpha(b) = \frac{b^3\left(b - \rho\frac{\sigma_y}{\sigma_x}\right)}{b^3\left(b - \rho\frac{\sigma_y}{\sigma_x}\right) - \left(\rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2\right)}. \quad (64)$$

Proof: Begin with the slope equation and solve for α in terms of b. ■

V. Classifying Generalized Regressions Using Weighted Geometric Means

Weighted geometric means allow for a similar parameterization of generalized least-squares lines.

Definition 5: The weighted geometric mean of order β is defined for $x < y$ by

$$M_\beta(x, y) = x^{1-\beta}y^\beta. \quad (65)$$

The standard geometric mean corresponds to $\beta = \frac{1}{2}$. The weight function is given by

$$g_\beta(b) = \frac{1}{b^{2\beta}} \quad (66)$$

and the corresponding slope equation is given by

$$(1 - \beta)b^2 + (2\beta - 1)\rho\frac{\sigma_y}{\sigma_x}b - \beta\left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \quad (67)$$

The indicative function is given by

$$G_\beta(b) = \frac{1 - 2\beta}{b}. \quad (68)$$

It is again clear that β continuously parameterizes a spectrum of generalized regression lines with OLS $y|x$ corresponding to $\beta = 0$ and OLS $x|y$ corresponding to $\beta = 1$. Again, due to intermediacy, every generalized mean lies between $x = \min(x, y)$ and $y = \max(x, y)$ corresponding to OLS $y|x$ and OLS $x|y$ respectively. Therefore every generalized mean-square regression can be assigned an effective parameter β in $[0, 1]$, classifying the line on the spectrum between OLS $y|x$ and OLS $x|y$. This is formalized in the next theorem.

Theorem 11: (Weighted Geometric Mean Equivalence Theorem) Let b be the slope of a generalized least-squares regression line lying between the two ordinary least-squares lines. Then the regression line is generated by an equivalent weighted geometric mean of order β in $[0, 1]$ where

$$\beta(b) = \frac{b\left(b - \rho\frac{\sigma_y}{\sigma_x}\right)}{b\left(b - \rho\frac{\sigma_y}{\sigma_x}\right) - \left(\rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2\right)}. \quad (69)$$

Proof: Begin with the slope equation and solve for β in terms of b. ■

VI. Classifying Generalized Regressions Using Power Means

It is now shown that the subset of lines lying between the two ordinary least-squares lines is also naturally parameterized using power means. The power mean parameter p is either an increasing or decreasing function of the slope.

A. Power Mean Regression

Definition 6: The power mean of order p is defined by

$$M_p(x, y) = \left(\frac{x^p + y^p}{2}\right)^{1/p} \quad (70)$$

where $-\infty < p < \infty$ and $p \neq 0$. The following special cases are defined separately:

$$M_{-\infty}(x, y) = \min(x, y), \quad (71)$$

$$M_\infty(x, y) = \max(x, y), \quad (72)$$

and

$$M_0(x, y) = \sqrt{xy}. \quad (73)$$

For $p > 0$, a power mean of order p is the arithmetic mean of x^p and y^p raised to the $\frac{1}{p}$ power. For $p < 0$, a power mean of order p is the harmonic mean of $x^{|p|}$ and $y^{|p|}$ raised to the $\frac{1}{|p|}$ power. The following conjugacy relation therefore holds:

$$M_{-p}(x, y)\, M_p(x, y) = (G(x, y))^2 \quad (74)$$

and therefore

$$G(M_{-p}(x, y), M_p(x, y)) = G(x, y). \quad (75)$$

The general weight function for power mean regression when $p \neq 0$ and $p \neq \pm\infty$ is given by

$$g_p(b) = M_p\left(1, \frac{1}{b^2}\right) = 2^{-1/p}\left(1 + \frac{1}{b^{2p}}\right)^{1/p}. \quad (76)$$

To find b one must solve the slope equation

$$b^{2p+2} - \rho\frac{\sigma_y}{\sigma_x}b^{2p+1} + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \quad (77)$$

The equation is a polynomial when p is any whole number multiple of $\frac{1}{2}$. When $p = 0$ the power mean formula becomes undefined. Nevertheless, substituting $p = 0$ into the slope equation and solving for the slope yields $b = \pm\frac{\sigma_y}{\sigma_x}$ which corresponds to geometric mean regression. This explains the separate definition $M_0(x, y) = G(x, y)$.

The indicative function for power mean regression has a simple form:

$$G_p(b) = \frac{(2p+1)\,b^{2p} - 1}{b\,(b^{2p}+1)}. \tag{78}$$

Note that if $|b| < 1$, then $p = -\infty$ corresponds to OLS $y|x$ and $p = \infty$ corresponds to OLS $x|y$. If $|b| > 1$, then $p = -\infty$ corresponds to OLS $x|y$ and $p = \infty$ corresponds to OLS $y|x$.

B. Power Mean Equivalence Theorem

Ordinary least-squares and geometric mean regression were already observed to be specific cases of power mean regression. The harmonic mean $H(x, y)$ is equivalent to $M_{-1}(x, y)$ and the arithmetic mean $A(x, y)$ is equivalent to $M_1(x, y)$. Square perimeter regression (SPR) (also called least-perimeter squared regression previously) is generated by $M_{1/2}(x, y)$ and squared harmonic mean regression (SHR) is generated by $M_{-1/2}(x, y)$. This suggests that it is natural to use the power mean parameter p to parameterize generalized least-squares regression lines. The next table summaries these cases.

Case	**Mean** $M(x,y)$ $x < y$	**Weight Function** $g(b) = M\left(1, \frac{1}{b^2}\right)$	**Order** p
OLS $y\mid x$	x	1	$\pm\infty$
OLS $x\mid y$	y	$\frac{1}{b^2}$	$\pm\infty$
HMR (Orthogonal)	$\frac{2xy}{x+y}$	$\frac{2}{1+b^2}$	-1
SHR	$\left(\frac{2\sqrt{xy}}{\sqrt{x}+\sqrt{y}}\right)^2$	$\frac{4}{(1+\|b\|)^2}$	$-\frac{1}{2}$
GMR	$\sqrt{xy}$	$\frac{1}{\|b\|}$	0
SPR	$\left(\frac{\sqrt{x}+\sqrt{y}}{2}\right)^2$	$\frac{1}{4}\left(1+\frac{1}{\|b\|}\right)^2$	$\frac{1}{2}$
AMR (Pythagorean)	$\frac{x+y}{2}$	$\frac{1}{2}\left(1+\frac{1}{b^2}\right)$	1

It is observed further on that Heronian mean regression (NMR) has a value for p that is slightly below $\frac{2}{3}$ and identric mean regression (IMR) has a value for p that is slightly above $\frac{2}{3}$. First and second logarithmic mean regressions have values for p that are close to $\pm\frac{1}{3}$ respectively. Power mean regression of order $\frac{1}{3}$ is also called Lorentz mean regression (ZMR).

Theorem 12: (Power Mean Equivalence Theorem) Let b be the slope of a generalized least-squares regression line. Then the regression line is generated by an equivalent power mean of order p where

$$p(b) = \frac{\ln\left(\frac{\rho\frac{\sigma_y}{\sigma_x}\left(\frac{1}{\rho}\frac{\sigma_y}{\sigma_x} - b\right)}{b\left(b - \rho\frac{\sigma_y}{\sigma_x}\right)}\right)}{\ln(b^2)}. \tag{79}$$

Proof: Begin with the slope equation and solve for p in terms of b. ■

The next theorem describes how the power mean order continuously parameterizes a spectrum of generalized least-squares regression lines as p increases from $-\infty$ to $+\infty$.

Theorem 13: (The Power Mean Spectrum)

Case 1: For all b satisfying $0 < \rho\frac{\sigma_y}{\sigma_x} < b < \frac{1}{\rho}\frac{\sigma_y}{\sigma_x} \le 1$, $p(b)$ is an increasing function and the inverse $b(p)$ is an increasing function over $(-\infty, \infty)$. OLS $y|x$ corresponds to $p = -\infty$ and OLS $x|y$ corresponds to $p = +\infty$.

Case 2: For all b satisfying $1 \le \rho\frac{\sigma_y}{\sigma_x} < b < \frac{1}{\rho}\frac{\sigma_y}{\sigma_x}$, $p(b)$ is a decreasing function and the inverse $b(p)$ is a decreasing function over $(-\infty, \infty)$. OLS $x|y$ corresponds to $p = -\infty$ and OLS $y|x$ corresponds to $p = +\infty$.

Case 3: For all b satisfying $-1 \le \frac{1}{\rho}\frac{\sigma_y}{\sigma_x} < b < \rho\frac{\sigma_y}{\sigma_x} < 0$, $p(b)$ is a decreasing function and the inverse $b(p)$ is a decreasing function over $(-\infty, \infty)$. OLS $y|x$ corresponds to $p = -\infty$ and OLS $x|y$ corresponds to $p = +\infty$.

Case 4: For all b satisfying $\frac{1}{\rho}\frac{\sigma_y}{\sigma_x} < b < \rho\frac{\sigma_y}{\sigma_x} \le -1$, $p(b)$ is an increasing function and the inverse $b(p)$ is an increasing function over $(-\infty, \infty)$. OLS $x|y$ corresponds to $p = -\infty$ and OLS $y|x$ corresponds to $p = +\infty$.

Case 5: For all b satisfying $0 < \rho\frac{\sigma_y}{\sigma_x} < b < 1 < \frac{1}{\rho}\frac{\sigma_y}{\sigma_x}$ and $\frac{\sigma_y}{\sigma_x} < 1$, $p(b)$ is an increasing function over $\left(\rho\frac{\sigma_y}{\sigma_x}, 1\right)$. The inverse $b(p)$ is an increasing function over $(-\infty, \infty)$. OLS $y|x$ corresponds to $p = -\infty$ and the regression line with $b = 1$ corresponds to $p = +\infty$.

Case 6: For all b satisfying $\frac{1}{\rho}\frac{\sigma_y}{\sigma_x} < -1 < b < \rho\frac{\sigma_y}{\sigma_x} < 0$ and $\frac{\sigma_y}{\sigma_x} < 1$, $p(b)$ is a decreasing function over $\left(-1, \rho\frac{\sigma_y}{\sigma_x}\right)$. The inverse $b(p)$ is a decreasing function over $(-\infty, \infty)$. OLS $y|x$ corresponds to $p = -\infty$ and the regression line with $b = -1$ corresponds to $p = +\infty$.

Proof: In all cases begin with the derivative

$$\begin{aligned} p'(b) \;=\; & \frac{1}{(\ln b^2)^2}\left\{\left(\frac{1}{b - \frac{1}{\rho}\frac{\sigma_y}{\sigma_x}} - \frac{1}{b - \rho\frac{\sigma_y}{\sigma_x}}\right)\ln b^2 \right. \\ & \left. - \ln\left(\frac{\rho\frac{\sigma_y}{\sigma_x}\left(\frac{1}{\rho}\frac{\sigma_y}{\sigma_x} - b\right)}{\left(b - \rho\frac{\sigma_y}{\sigma_x}\right)}\right)^2 \left(\frac{1}{b}\right)\right\}. \end{aligned} \tag{80}$$

The critical values of the $p(b)$ occur at $b = \pm 1, b = \rho\frac{\sigma_y}{\sigma_x} = b_{\text{OLS Y|X}}$, $b = \frac{1}{\rho}\frac{\sigma_y}{\sigma_x} = b_{\text{OLS X|Y}}$ and when $p'(b) = 0$. The first derivative test is now applied between the critical values to determine whether $p(b)$ is increasing or decreasing. Choose the test value b_* to be the arithmetic mean of the two ordinary least-squares slopes. That is, choose

$$b_* = \frac{1}{2}\left(\rho + \frac{1}{\rho}\right)\frac{\sigma_y}{\sigma_x} \tag{81}$$

and obtain

$$p'(b_*) = -\frac{\ln\left(\rho\frac{\sigma_y}{\sigma_x}\right)^2}{b_*\left(\ln b_*^2\right)^2}. \tag{82}$$

For this choice, $|b_*| < 1$ whenever $\frac{\sigma_y}{\sigma_x} < 1$. According to the first derivative test when $p'(b_*)$ is positive, $p(b)$ is increasing over the interval. It follows that its inverse $b(p)$ is an increasing function over $(-\infty, \infty)$. When $p'(b_*)$ is negative, $p(b)$ is decreasing over the interval. Its inverse $b(p)$ is therefore a decreasing function over $(-\infty,\infty)$.

In Case 1, $p'(b_*) > 0$, in Case 2, $p'(b_*) < 0$, in Case 3, $p'(b_*) < 0$, in Case 4, $p'(b_*) > 0$, in Case 5, b_* is in $\left(\rho\frac{\sigma_y}{\sigma_x},1\right)$ and $p'(b_*) > 0$, and in Case 6, b_* is in $\left(-1,\rho\frac{\sigma_y}{\sigma_x}\right)$ and $p'(b_*) < 0$. ■

Corollary 14: Generalized least-squares lines follow the same order relations as their corresponding means. The slopes increase or decrease in the following order: HMR, SHR, GMR, SPR, AMR.

Proof: The power mean is a increasing function of p and the slope b is always either an increasing function or a decreasing of p. ■

Thus the order of the generalized least-squares regression lines is explained by the corresponding inequalities of the underlying special means. The reader can observe this order among the lines in the numerical examples below.

The case of $|b| = 1$ plays a special role in the previous theorem; therefore it is given a name.

Definition 7: Call a generalized regression line with $|b| = 1$ unitary least-squares (ULS). If $b = 1$ then $a = \mu_y - \mu_x$ and if $b = -1$ then $a = \mu_y + \mu_x$.

C. Power Mean Regressions of Order $-\frac{1}{3}$, $\frac{1}{3}$ and $\frac{2}{3}$

It is known and observed here that the Lorentz mean or power mean of order $\frac{1}{3}$, approximates the logarithmic mean well [12]. The power mean of order $-\frac{1}{3}$ approximates both the second logarithmic mean and the geometrically weighted harmonic-geometric mean $H^{1/3}G^{2/3}$, usually written as $\left(HG^2\right)^{1/3}$ [3,13]. The power mean of order $\frac{2}{3}$ approximates both the Heronian mean and the identric means well. Therefore the regression formulas for these power means deserve special consideration and are described now.

For the Lorentz mean regression (ZMR),

$$Z(x,y) = M_{1/3}(x,y) = \left(\frac{\sqrt[3]{x}+\sqrt[3]{y}}{2}\right)^3, \qquad (83)$$

the weight function is given by

$$g_{1/3}(b) = \frac{1}{8}\left(1+b^{-2/3}\right)^3, \qquad (84)$$

and the slope equation is

$$b^{8/3} - \rho\frac{\sigma_y}{\sigma_x}b^{5/3} + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \qquad (85)$$

The slope equation is solved as a polynomial equation in q after setting $q = b^{1/3}$. The result is

$$q^8 - \rho\frac{\sigma_y}{\sigma_x}q^5 + \rho\frac{\sigma_y}{\sigma_x}q^3 - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0 \qquad (86)$$

and $b = q^3$.

For the power mean with $p = -\frac{1}{3}$,

$$M_{-1/3}(x,y) = \left(\frac{2\sqrt[3]{xy}}{\sqrt[3]{x}+\sqrt[3]{y}}\right)^3, \qquad (87)$$

the weight function is given by

$$g_{-1/3}(b) = 8\left(1+b^{2/3}\right)^{-3}, \qquad (88)$$

and the slope equation is

$$b^{4/3} - \rho\frac{\sigma_y}{\sigma_x}b^{1/3} + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \qquad (89)$$

The slope equation is again solved as a polynomial equation in q after setting $q = b^{1/3}$. The result is

$$q^4 + \rho\frac{\sigma_y}{\sigma_x}q^3 - \rho\frac{\sigma_y}{\sigma_x}q - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0 \qquad (90)$$

and $b = q^3$.

For the power mean with $p = \frac{2}{3}$,

$$M_{2/3}(x,y) = \left(\frac{\sqrt[3]{x^2}+\sqrt[3]{y^2}}{2}\right)^3, \qquad (91)$$

the weight function is given by

$$g_{2/3}(b) = \frac{1}{\sqrt{8}}\left(1+b^{-4/3}\right)^{3/2}, \qquad (92)$$

and the slope equation is

$$b^{10/3} - \rho\frac{\sigma_y}{\sigma_x}b^{7/3} + \rho\frac{\sigma_y}{\sigma_x}b - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0. \qquad (93)$$

The slope equation is again solved as a polynomial equation in q after setting $q = b^{1/3}$. The result is

$$q^{10} - \rho\frac{\sigma_y}{\sigma_x}q^7 + \rho\frac{\sigma_y}{\sigma_x}q^3 - \left(\frac{\sigma_y}{\sigma_x}\right)^2 = 0 \qquad (94)$$

and $b = q^3$.

VII. Classifying Regressions Using Alternative Generalized Means

Power means of order p were just shown in certain cases to naturally parameterize the regression lines lying between the two ordinary least-squares lines. This section explores the use of other known generalized means having free parameters which give alternative ways to parameterize generalized regression lines.

A. Dietel and Gordon's Generalized Mean

It was shown by Dietel and Gordon [3] that if $(x, f(x))$ and $(y, f(y))$ are two distinct points on the graph of a function f, then the tangent lines of f at these two points intersect at the arithmetic mean of x and y if $f(x) = x^2$, the geometric mean of x and y if $f(x) = \sqrt{x}$ and the harmonic mean of x and y if $f(x) = \frac{1}{x}$. The authors generalize this result to means $S_f(x,y)$ that are generated by an arbitrary function f satisfying certain conditions. When $f(x) = x^r$, the resulting means are parametrized by r and are given by

$$S_r(x,y) = \begin{cases} \frac{r-1}{r}\cdot\frac{y^r-x^r}{y^{r-1}-x^{r-1}}, & x \neq y, r \neq 0,1 \\ x, & x = y. \end{cases} \qquad (95)$$

The function $S_r(x,y)$ is called here the Dietel-Gordon mean of order r. The case $S_1(x,y)=\lim_{r\to 1} S_r(x,y)$ is given by

$$S_1(x,y)=(y-x)/(\ln y-\ln x) \quad (96)$$

and is the logarithmic mean. It is parameterized by $f(x)=x\ln x$. The case $S_0(x,y)=\lim_{r\to 0} S_r(x,y)$ is given by

$$S_0(x,y)=xy(\ln y-\ln x)/(y-x) \quad (97)$$

and is the second logarithmic mean. It is parameterized by $f(x)=\ln x$. The authors show that $\lim_{r\to -\infty} S_r(x,y)=x=\min(x,y)$ and $\lim_{r\to +\infty} S_r(x,y)=y=\max(x,y)$ and that $S_r(x,y)$ is an increasing function of r. Dietel and Gordon show that many of the special means mentioned here are specific cases of this generalized mean for particular values of r. The next chart summarizes this.

Mean Type	**Formula**	**Order** r
Minimum	$\min(x,y)$	$-\infty$
Harmonic	$H(x,y)=\dfrac{2xy}{x+y}$	-1
Second logarithmic	$L_2(x,y)=\dfrac{xy(\ln y-\ln x)}{y-x}$	0
Geometric	$G(x,y)=\sqrt{xy}$	$\frac{1}{2}$
Logarithmic	$L(x,y)=\dfrac{y-x}{\ln y-\ln x}$	1
Heronian	$N(x,y)=\dfrac{1}{3}(x+\sqrt{xy}+y)$	$\frac{3}{2}$
Arithmetic	$A(x,y)=\dfrac{x+y}{2}$	2
Centroidal	$T(x,y)=\dfrac{2}{3}\cdot\dfrac{x^2+xy+y^2}{x+y}$	3
Maximum	$\max(x,y)$	∞

Since $S_r(x,y)$ is an increasing function of r, it follows that

$$H\le L_2\le G\le L\le N\le A\le T. \quad (98)$$

Chen defines a related generalized mean $F_t(x,y)=S_{r-2}(x,y)$ in his paper and presents a similar classification [2]. Because of its simple relation to the already existing Dietel-Gordon mean, the Dietel-Gordon parameterization is preferred here.

B. Regression Formulas for the Dietel-Gordon Generalized Mean

The generalized mean $S_r(x,y)$ is now applied to the task of classifying and further understanding generalized least-squares regressions. The weight function is given by

$$g_r(b)=\frac{r-1}{r}\left(1+\frac{b^2-1}{b^{2r}-b^2}\right), r\neq 0,1. \quad (99)$$

As in the case of power mean regression, since $S_{-\infty}(x,y)=\min(x,y)$, it follows that when $|b|<1$, $r=-\infty$ corresponds to OLS $y|x$ and when $|b|>1$, $r=-\infty$ corresponds to OLS $x|y$. Similarly, since $S_\infty(x,y)=\max(x,y)$, it follows that when $|b|<1$, $r=\infty$ corresponds to OLS $x|y$ and when $|b|>1$, $r=\infty$ corresponds to OLS $y|x$.

The slope equation is given by

$$\begin{aligned} 0 \;=\;& b^{4r}-\rho\frac{\sigma_y}{\sigma_x}b^{4r-1}-rb^{2r+2} \\ &+(2r-1)\rho\frac{\sigma_y}{\sigma_x}b^{2r+1} \\ &+(r-1)\left(1-\left(\frac{\sigma_y}{\sigma_x}\right)^2\right)b^{2r} \\ &-(2r-1)\rho\frac{\sigma_y}{\sigma_x}b^{2r-1}+r\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^{2r-2} \\ &+\rho\frac{\sigma_y}{\sigma_x}b-\left(\frac{\sigma_y}{\sigma_x}\right)^2. \end{aligned} \quad (100)$$

Unlike power mean regression, the parameter r cannot be solved in terms of b explicitly. Nevertheless, it can always be done numerically. The slope profile $b=b(r)$ exhibits similar behavior to $b=b(p)$ and is plotted numerically for the examples below. Generalized least-squares lines follow the same order relations as the inequalities of their corresponding means. The slope b is either a strictly increasing function of r or a strictly decreasing function of r. Therefore the slopes increase or decrease in the following order: HMR, $L2MR$, GMR, LMR, SPR, NMR, AMR, TMR. The reader can observe this in the numerical examples presented further on.

C. Stolarsky's Generalized Logarithmic Mean

Mays [13] describes how the integral and differential forms of the mean value theorem of calculus can be used to define generalized means. The integral mean value theorem states that for f continuous and strictly monotone on $(0,\infty)$ there is a unique c in (x,y) such that

$$f(c)(y-x)=\int_x^y f(w)\,dw. \quad (101)$$

Therefore one can define the generalized mean

$$U_f(x,y)=f^{-1}\left(\int_x^y f(w)\,dw/(y-x)\right). \quad (102)$$

For $f(x)=x^{s-1}$ the generalized mean becomes Stolarsky's generalized logarithmic mean [3,13,14,15]:

$$U_s(x,y)=\left(\int_x^y w^{s-1}dw/(y-x)\right)^{1/(s-1)} \quad (103)$$

which can be written explicitly as

$$U_s(x,y)=\begin{cases}\left(\frac{1}{s}\cdot\frac{y^s-x^s}{y-x}\right)^{1/(s-1)}, & x\neq y, s\neq 0,1\\ x, & x=y.\end{cases} \quad (104)$$

The case $s=0$ corresponds to the logarithmic mean and the case $s=1$ corresponds to the identric mean. The weight function for Stolarsky's mean is given by

$$\begin{aligned} g_s(b) &= U_s\left(1,\frac{1}{b^2}\right) \\ &= s^{-1/(s-1)}\left(1+\frac{b^{2s-2}-1}{b^{2s}-b^{2s-2}}\right)^{1/(s-1)} \end{aligned} \tag{105}$$

and the slope equation is given by

$$\begin{aligned} 0 &= (s-1)\,b^{2s+4}-(s-1)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+3}-sb^{2s+2} \\ &\quad +(s+1)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+1}-\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^{2s} \\ &\quad +b^4-(s+1)\,\rho\frac{\sigma_y}{\sigma_x}b^3+s\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^2 \\ &\quad +(s-1)\,\rho\frac{\sigma_y}{\sigma_x}b-(s-1)\left(\frac{\sigma_y}{\sigma_x}\right)^2. \end{aligned} \tag{106}$$

Again, the parameter s cannot be solved in terms of b explicitly, however it can always be done numerically. The slope profile $b=b(s)$ exhibits similar behavior to $b=b(p)$ and $b=b(r)$ and is plotted numerically in the examples below.

D. Gini's Mean

Gini's mean [6,11,13] is given by

$$G_t(x,y)=\frac{x^t+y^t}{x^{t-1}+y^{t-1}}. \tag{107}$$

The weight function for Gini's mean is:

$$g_t(b)=G_t\left(1,\frac{1}{b^2}\right)=1+\frac{1-b^2}{b^{2t}+b^2} \tag{108}$$

and the slope equation is given by

$$\begin{aligned} 0 &= b^{4t}-\rho\frac{\sigma_y}{\sigma_x}b^{4t-1}+tb^{2t+2} \\ &\quad -(2t-1)\,\rho\frac{\sigma_y}{\sigma_x}b^{2t+1} \\ &\quad +(t-1)\left(\left(\frac{\sigma_y}{\sigma_x}\right)^2-1\right)b^{2t} \\ &\quad +(2t-1)\,\rho\frac{\sigma_y}{\sigma_x}b^{2t-1}-t\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^{2t-2} \\ &\quad +\rho\frac{\sigma_y}{\sigma_x}b-\left(\frac{\sigma_y}{\sigma_x}\right)^2. \end{aligned} \tag{109}$$

A related mean is Moskovitz's mean $M_t(x,y)=G_{1-t}(x,y)$ [13], given explicitly by

$$M_t(x,y)=\frac{xy^t+yx^t}{x^t+y^t}. \tag{110}$$

Because of its simple relation to Gini's mean and because Gini's mean parameterizes the specific cases in the same order as the power mean, the Dietel-Gordon mean, and Stolarsky's logarithmic mean, Gini's mean is preferred here. Moskovitz's mean parameterizes the specific cases in the reverse order. Again, the parameter t cannot be solved in terms of b explicitly, however it can always be done numerically. The slope profile $b=b(t)$ will be analyzed in greater detail in future work.

E. Two-Parameter Means

The two-parameter mean of Stolarsky [10,14] is given by

$$E_{r,s}(x,y)=\begin{cases}\left(\frac{r(x^s-y^s)}{s(x^r-y^r)}\right)^{1/(s-r)}, & x\neq y\\ x, & x=y.\end{cases} \tag{111}$$

Some properties include: $E_{r,s}(x,y)=E_{s,r}(x,y)$, $E_{r,-r}(x,y)=G(x,y)$. The power mean is a specific case: $M_p(x,y)=E_{p,2p}(x,y)$, as is Stolarsky's one-parameter mean: $U_s(x,y)=E_{s,1}(x,y)$. The weight function is given by

$$\begin{aligned} g_{r,s}(b) &= E_{r,s}\left(1,\frac{1}{b^2}\right) \\ &= \left(\frac{r}{s}\right)^{1/(s-r)}\left(1+\frac{b^{2(s-r)}-1}{b^{2s}-b^{2(s-r)}}\right)^{1/(s-r)} \end{aligned} \tag{112}$$

and the slope equation is given by

$$\begin{aligned} 0 &= (s-r)\,b^{2s+2r+2}-(s-r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+2r+1} \\ &\quad -sb^{2s+2}+rb^{2r+2}+(s+r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+1} \\ &\quad -(s+r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2r+1}-\left(\frac{\sigma_y}{\sigma_x}\right)^2 rb^{2s} \\ &\quad +\left(\frac{\sigma_y}{\sigma_x}\right)^2 sb^{2r}+\rho\frac{\sigma_y}{\sigma_x}(s-r)\,b \\ &\quad -(s-r)\left(\frac{\sigma_y}{\sigma_x}\right)^2. \end{aligned} \tag{113}$$

Gini's two-parameter mean [6,13] is given by

$$G_{r,s}(x,y)=\left(\frac{x^s+y^s}{x^r+y^r}\right)^{1/(s-r)}. \tag{114}$$

Gini's two-parameter mean includes as specific cases the power means: $M_p(x,y)=G_{p,0}(x,y)$ and Gini's one-parameter mean: $G_s(x,y)=G_{s,s-1}(x,y)$. The weight function is given by

$$\begin{aligned} g_{r,s}(b) &= G_{r,s}\left(1,\frac{1}{b^2}\right) \\ &= \left(1+\frac{1-b^{2(s-r)}}{b^{2s}+b^{2(s-r)}}\right)^{1/(s-r)} \end{aligned} \tag{115}$$

and the slope equation is given by

$$\begin{aligned} 0 &= (s-r)\,b^{2s+2r+2}-(s-r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+2r+1} \\ &\quad +sb^{2s+2}-rb^{2r+2}-(s+r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2s+1} \\ &\quad +(s+r)\,\rho\frac{\sigma_y}{\sigma_x}b^{2r+1}+r\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^{2s} \\ &\quad -s\left(\frac{\sigma_y}{\sigma_x}\right)^2 b^{2r}+(s-r)\,\rho\frac{\sigma_y}{\sigma_x}b \\ &\quad -(s-r)\left(\frac{\sigma_y}{\sigma_x}\right)^2. \end{aligned} \tag{116}$$

F. Relating the Mean Parameters Using Linear Approximations

The p, r, s, and t scales offer different ways of parameterizing and classifying the known special means. The actual functions $p = p(r)$, $r = r(p)$, etc. depend on the data (i.e. on ρ and $\frac{\sigma_y}{\sigma_x}$) and cannot be written explicitly. However, linear approximations are possible which give further insight and understanding. The linear relations presented here reveal that the Heronian and identric means are always approximate power means of order $\frac{2}{3}$, regardless of the data. The first and second logarithmic means are always approximate power means of order $\pm\frac{1}{3}$ respectively, regardless of the data. The geometrically weighted harmonic-geometric mean $H^{1/3}G^{2/3}$, usually written $\left(HG^2\right)^{1/3}$, has an exact Stolarsky parameter of $s = -2$, which corresponds approximately to a power mean of order $-\frac{1}{3}$ and to the second logarithmic mean.

Theorem 15: (Approximate Linear Relations) The power mean parameter p, the Dietel-Gordon parameter r, the Stolarsky parameter s and the Gini parameter t are interrelated according to the following pairs of linear approximations:

$$\begin{aligned}
p &\approx \frac{2}{3}r - \frac{1}{3},\ r \text{ in } [-1,2] \text{ and} \\
r &\approx \frac{3}{2}p + \frac{1}{2},\ p \text{ in } [-1,1], \qquad (117) \\
s &\approx 2r - 2,\ r \text{ in } \left[\frac{1}{2}, 2\right] \text{ and} \\
r &\approx \frac{1}{2}s + 1,\ s \text{ in } [-1,2], \qquad (118) \\
s &\approx 3p - 1,\ p \text{ in } [0,1] \text{ and} \\
p &\approx \frac{1}{3}s + \frac{1}{3},\ s \text{ in } [-1,2], \qquad (119) \\
t &\approx \frac{1}{2}p + \frac{1}{2},\ p \text{ in } [-1,1] \text{ and} \\
p &\approx 2t - 1,\ t \text{ in } [0,1], \qquad (120) \\
t &\approx \frac{1}{3}r + \frac{1}{3},\ r \text{ in } [-1,2] \text{ and} \\
r &\approx 3t - 1,\ t \text{ in } [0,1], \qquad (121) \\
t &\approx \frac{1}{6}s + \frac{2}{3},\ s \text{ in } [-1,2] \text{ and} \\
s &\approx 6t - 4,\ s \text{ in } \left[\frac{1}{2}, 1\right]. \qquad (122)
\end{aligned}$$

Proof: In the (p,r) plane, regardless of the data, the harmonic mean corresponds to $(p,r) = (-1,-1)$, the geometric mean corresponds to $(p,r) = \left(0, \frac{1}{2}\right)$, and the arithmetic mean corresponds to $(p,r) = (1,2)$. There is a single straight line passing though these three points which approximates the values of p and r lying between the harmonic, geometric and arithmetic means.

In the (s,r) plane, the geometric mean corresponds to $(s,r) = \left(-1, \frac{1}{2}\right)$, the logarithmic mean corresponds to $(s,r) = (0,1)$ and the arithmetic mean corresponds to $(s,r) = (2,2)$. The linear approximation is the single straight line which passes through these points.

In the (s,p) plane, the geometric mean corresponds to $(s,p) = (-1,0)$, the power mean of order $\frac{1}{2}$ corresponds to $(s,p) = \left(\frac{1}{2}, \frac{1}{2}\right)$, and the arithmetic mean corresponds to $(2,1)$. The linear approximation is again the single straight line which passes through these points.

In the (p,t) plane, the harmonic mean corresponds to $(-1,0)$, the geometric mean corresponds to $\left(0, \frac{1}{2}\right)$, and the arithmetic mean corresponds to $(1,1)$. The linear approximation is again the single straight line which passes through these points.

In the (r,t) plane, the harmonic mean corresponds to $(-1,0)$, the geometric mean corresponds to $\left(\frac{1}{2}, \frac{1}{2}\right)$, and the arithmetic mean corresponds to $(2,1)$. The linear approximation is again the single straight line which passes through these points.

Finally, in the (s,t) plane, the harmonic mean does not have a fixed value, but the geometric mean corresponds to $\left(-1, \frac{1}{2}\right)$, and the arithmetic mean corresponds to $(2,1)$. The linear approximation is the straight line which passes through these points. ■

The next chart summarizes the approximate correspondences for all the notable cases. An asterisk by the number indicates that it is an approximate value obtained from the linear formulas. Two asterisks indicate that the approximate value is outside the domain of the linear formulas. Therefore the approximation will be less accurate than the values with single asterisks.

Mean Type	**Power** $M_p(x,y)$	**Dietel-Gordon** $S_r(x,y)$	**Stolarsky** $U_s(x,y)$	**Gini** $G_t(x,y)$
Order	p	r	s	t
Harmonic	-1	-1	-4**	0
Power	$-\frac{1}{2}$	$-\frac{1}{4}$*	-2.5**	$\frac{1}{4}$*
Second Logarithmic	$-\frac{1}{3}$*	0	-2*	$\frac{1}{3}$*
$(HG^2)^{1/3}$	$-\frac{1}{3}$*	0*	-2	$\frac{1}{3}$*
Geometric	0	$\frac{1}{2}$	-1	$\frac{1}{2}$
Logarithmic	$\frac{1}{3}$*	1	0	$\frac{2}{3}$*
Lorentz	$\frac{1}{3}$	1*	0*	$\frac{2}{3}$*
Power	$\frac{1}{2}$	$\frac{5}{4}$*	$\frac{1}{2}$	$\frac{3}{4}$*
Heronian	$\frac{2}{3}$*	$\frac{3}{2}$	1*	$\frac{5}{6}$*
Identric	$\frac{2}{3}$*	$\frac{3}{2}$*	1	$\frac{5}{6}$*
Arithmetic	1	2	2	1

Although there is no fixed value of the Stolarsky parameter s corresponding to the harmonic mean [13], there is always a variable value of s which depends in the data corresponding to orthogonal regression which is HMR. The linear approximation may no longer be a good fit for s outside the domain $[-1,2]$. Nevertheless, the equations suggest taking $s=-4$ as an initial guess. Similarly, root-mean-square regression corresponds to $p=2$. Since this value is outside the domain $[-1,1]$, the linear equations may no longer give a good fit. Nevertheless, the linear approximations suggest taking $r=\frac{7}{2}, s=5$ and $t=\frac{3}{2}$ as initial guesses.

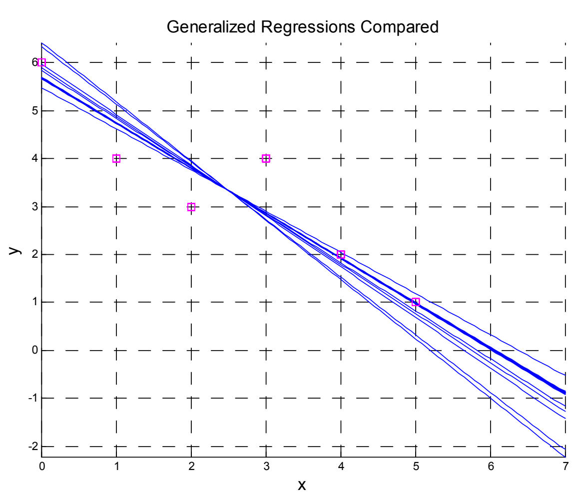
To summarize, the inequalities involving the known special means are due to the fact that they are parameterized by generalized means which are increasing functions of the parameter. The corresponding regression lines have the same order since their slopes are either increasing or decreasing functions of the parameter.

The approximate symmetry of all the special means about $p=0$ on the power mean scale is striking. Power mean regression has a slope equation where the free parameter p can be solved explicitly in terms of the slope. Because of this, the power mean spectrum may be a preferred scale in which to work.

VIII. Numerical Examples

This section revisits the examples explored previously. Regressions corresponding to all the special means are computed and the corresponding numerical values for the generalized mean parameters are computed for each regression line as well.

Example 1 Six data values are given: $(0,6)$, $(1,4)$, $(2,3)$, $(3,4)$, $(4,2)$, and $(5,1)$. The reader can verify that $\rho=-0.9157$, $\kappa=-0.4019$, $\mu_x=2.5000$, $\mu_y=3.3333$, $\sigma_x=1.7078$, and $\sigma_y=1.5986$. The generalized regression lines are plotted together with the extremal line thereby displaying the region containing all admissible generalized regression lines.



The equation of each line is presented along with the scatter parameter λ, the weighted arithmetic mean parameter α, and the weighted geometric mean parameter β. The parameters ϕ and δ defined earlier can be computed by the reader using $\phi=2\tan^{-1}\lambda$ and $\delta=\frac{4}{\pi}\tan^{-1}\lambda$.

Method	$y=a+bx$	λ	α	β
Extremal	$y=6.4166-1.2333x$	1.0000	—	—
CMR	$y=6.3335-1.2001x$	0.9117	—	—
TMR	$y=5.9729-1.0558x$	0.5282	—	—
OLS $x\mid y$	$y=5.8889-1.0222x$	0.4389	1.0000	1.0000
USR	$y=5.8333-1.0000x$	0.3798	0.8824	0.8824
RMR	$y=5.6958-0.9450x$	0.2335	0.5283	0.5563
AMR	$y=5.6855-0.9409x$	0.2226	0.5000	0.5304
IMR	$y=5.6817-0.9394x$	0.2185	0.4897	0.5209
PMR $p=\frac{2}{3}$	$y=5.6817-0.9394x$	0.2185	0.4897	0.5208
NMR	$y=5.6817-0.9394x$	0.2185	0.4896	0.5208
SPR	$y=5.6797-0.9386x$	0.2164	0.4842	0.5159
ZMR $p=\frac{1}{3}$	$y=5.6777-0.9377x$	0.2143	0.4786	0.5107
LMR	$y=5.6777-0.9377x$	0.2143	0.4786	0.5107
GMR	$y=5.6735-0.9361x$	0.2098	0.4670	0.5000
L2MR	$y=5.6690-0.9343x$	0.2050	0.4548	0.4887
PMR $p=-\frac{1}{3}$	$y=5.6690-0.9343x$	0.2050	0.4548	0.4887
SHR	$y=5.6667-0.9333x$	0.2026	0.4484	0.4828
HMR	$y=5.6593-0.9304x$	0.1947	0.4283	0.4640
OLS $y\mid x$	$y=5.4762-0.8571x$	0.0000	0.0000	0.0000

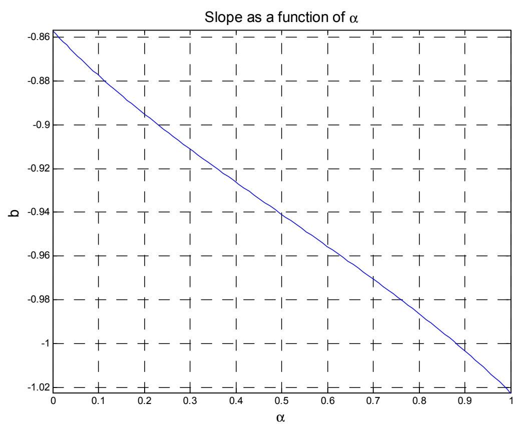
Method	p	r	s	t
USR	$+\infty$	$+\infty$	$+\infty$	$+\infty$
RMR	2.0000	3.4990	5.0077	1.4984
AMR	1.0000	2.0000	2.0000	1.0000
IMR	0.6667	1.5001	1.0000	0.8335
PMR $p=\frac{2}{3}$	0.6667	1.5001	1.0000	0.8335
NMR	0.6666	1.5000	0.9998	0.8334
SPR	0.5000	1.2501	0.5000	0.7501
ZMR $p=\frac{1}{3}$	0.3333	1.0000	0.0000	0.6668
LMR	0.3333	1.0000	0.0000	0.6668
GMR	0.0000	0.5000	-1.0000	0.5000
L2MR	-0.3333	0.0000	-2.0012	0.3332
PMR $p=-\frac{1}{3}$	-0.3333	0.0000	-2.0012	0.3332
SHR	-0.5000	-0.2501	-2.5029	0.2499
HMR	-1.0000	-1.0000	-4.0125	0.0000
OLS $y\mid x$	$-\infty$	$-\infty$	$-\infty$	$-\infty$

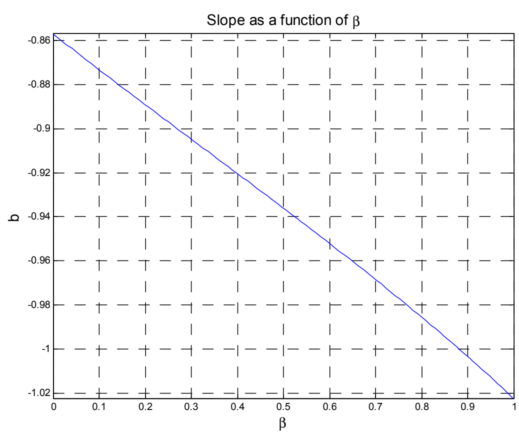
The parameter p is also computed using the explicit formula

ISBN: 978-1-61804-247-7

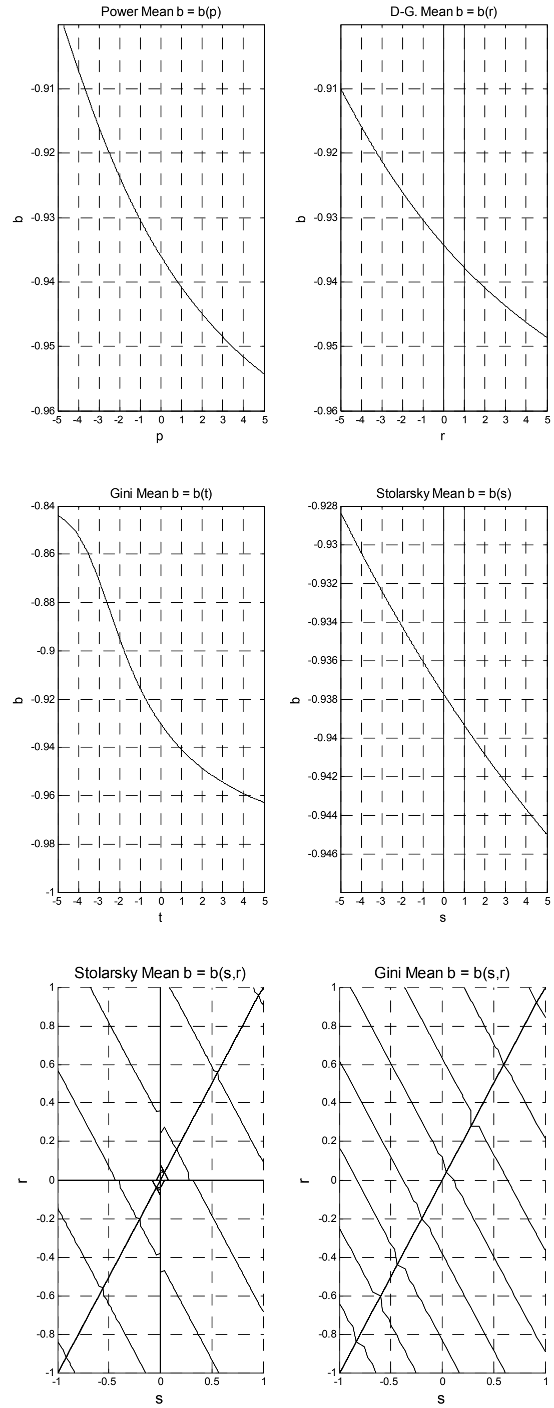
for $p(b)$ and the value for the slope. The parameters r, s, and t are computed by solving the slope equations numerically using the linear approximations $r_0 = \frac{3}{2}p + \frac{1}{2}$, $s_0 = 3p - 1$, and $t_0 = \frac{1}{2}p + \frac{1}{2}$ as initial guesses. For these data, the actual values for the parameters are in excellent agreement with the linear approximations.

The reader will observe that for this data set and level of accuracy, LMR and ZMR are indistinguishable as are NMR, PMR with $p = \frac{2}{3}$, and IMR. The slope profiles for weighted arithmetic mean regression and weighted geometric mean regression are displayed next.





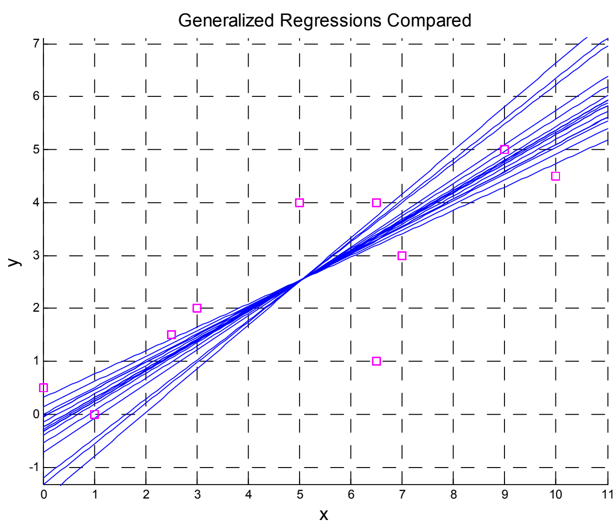
This example corresponds to Case 6 of the Power Mean Spectrum: the slope decreases continuously from the unitary least-squares slope -1 down to the ordinary least-squares slope -0.8571 as p varies from $-\infty$ up to ∞. The Dietel-Gordon mean, the Stolarsky mean, and the Gini mean, display similar behavior as is suggested by the next two pairs of slope profiles. In the Dietel-Gordon plot, $s = 0$ and $s = 1$, correspond to the two logarithmic mean regressions. In the Stolarsky plot, $s = 0$ corresponds to LMR and $s = 1$ corresponds to IMR.



In the third pair of plots, slope contours are plotted in the (s, r) parameter plane for Stolarsky's two-parameter mean $E_{r,s}(x, y)$ and Gini's two-parameter mean $G_{r,s}(x, y)$. The line $r = s$ which appears in the graphs is a removable singularity. The contours are decreasing as s and r increase. For the Stolarsky slope equation, six contours $b = -0.9327, ..., -0.9392$ are plotted in increments of -0.0013. For the Gini equation, eight contours $b =$

$-0.9262, ..., -0.9444$ are plotted in increments of -0.0026. The jaggedness of the contours is an artifact of the plotting algorithm.

Example 2 Ten data values are given: $(0, 0.5)$, $(1, 0)$, $(2.5, 1.5)$, $(3, 2)$, $(5, 4)$, $(6.5, 4)$, $(6.5, 1)$, $(7, 3)$, $(9, 5)$, and $(10, 4.5)$. The reader can verify that $\rho = 0.8268$, $\kappa = 0.5625$, $\mu_x = 5.0500$, $\mu_y = 2.5500$, $\sigma_x = 3.1737$, and $\sigma_y = 1.6948$. The generalized regression lines are plotted together with the extremal line thereby displaying the region containing all admissible generalized regression lines.



The equation of each line is presented along with the scatter parameter λ, the weighted arithmetic mean parameter α, and the weighted geometric mean parameter β. The parameters ϕ and δ defined earlier can be computed by the reader using $\phi = 2\tan^{-1}\lambda$ and $\delta = \frac{4}{\pi}\tan^{-1}\lambda$. The parameter p is also computed using the explicit formula and the value for the slope. The parameters r, s, and t are computed by solving the slope equations numerically using the linear approximations $r_0 = \frac{3}{2}p + \frac{1}{2}$, $s_0 = 3p - 1$, and $t_0 = \frac{1}{2}p + \frac{1}{2}$ as initial guesses. This example corresponds to Case 1 of the Power Mean Spectrum: the slope increases continuously from the OLS $y|x$ slope to the OLS $x|y$ slope as p increases from $-\infty$ to ∞. TMR has the equation $y = -1.6144 + 0.8246x$ and CMR has the equation $-1.321 + 0.7665x$ Their slopes exceed the extremal line and their corresponding λ values exceed 1. These lines are therefore inadmissible by the criteria described earlier.

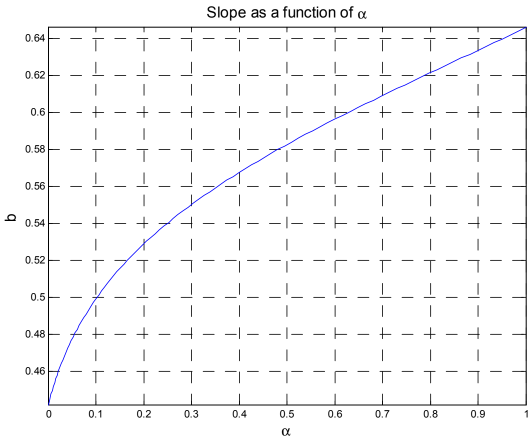
Method	$y = a + bx$	λ	α	β
Extremal	$y = -1.1966 + 0.7419x$	1.0000	—	—
OLS $x\|y$	$y = -0.7116 + 0.6459x$	0.6803	1.0000	1.0000
RMR	$y = -0.5431 + 0.6125x$	0.5691	0.7272	0.8766
AMR	$y = -0.3924 + 0.5827x$	0.4698	0.5000	0.7466
IMR	$y = -0.3227 + 0.5688x$	0.4238	0.4079	0.6804
PMR $p = \frac{2}{3}$	$y = -0.3220 + 0.5687x$	0.4233	0.4070	0.6797
NMR	$y = -0.3211 + 0.5685x$	0.4228	0.4059	0.6789
SPR	$y = -0.2824 + 0.5609x$	0.3972	0.3593	0.6407
ZMR $p = \frac{1}{3}$	$y = -0.2397 + 0.5525x$	0.3692	0.3119	0.5976
LMR	$y = -0.2390 + 0.5523x$	0.3686	0.3110	0.5967
GMR	$y = -0.1468 + 0.5340x$	0.3079	0.2219	0.5000
L2MR	$y = -0.0478 + 0.5144x$	0.2426	0.1459	0.3923
PMR $p = -\frac{1}{3}$	$y = -0.0465 + 0.5142x$	0.2417	0.1451	0.3909
SHR	$y = 0.0041 + 0.5041x$	0.2084	0.1136	0.3352
HMR	$y = 0.1406 + 0.4771x$	0.1184	0.0493	0.1854
OLS $y\|x$	$y = 0.3202 + 0.4416x$	0.0000	0.0000	0.0000

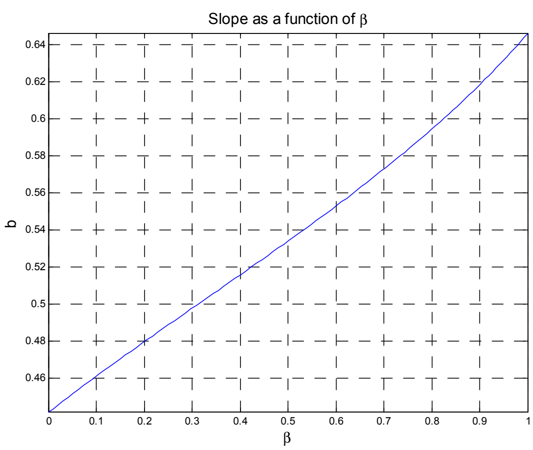
Method	p	r	s	t
OLS $x\|y$	$+\infty$	$+\infty$	$+\infty$	2.0113
RMR	2.0000	3.4396	5.6917	1.3952
AMR	1.0000	2.0000	2.0000	1.0000
IMR	0.6698	1.5103	1.0000	0.8445
PMR $p = \frac{2}{3}$	0.6667	1.5057	0.9906	0.8430
NMR	0.6628	1.5000	0.9795	0.8411
SPR	0.5000	1.2562	0.5000	0.7604
ZMR $p = \frac{1}{3}$	0.3333	1.0052	0.0101	0.6754
LMR	0.3299	1.0000	0.0000	0.6736
GMR	0.0000	0.5000	−1.0000	0.5000
L2MR	−0.3291	0.0000	−2.1243	0.3247
PMR $p = -\frac{1}{3}$	−0.3333	−0.0065	−2.1404	0.3224
SHR	−0.5000	−0.2586	−2.8197	0.2355
HMR	−1.0000	−1.0000	−5.9780	0.0000
OLS $y\|x$	$-\infty$	$-\infty$	$-\infty$	−0.3737

In the next diagrams slope profiles are plotted. The slope

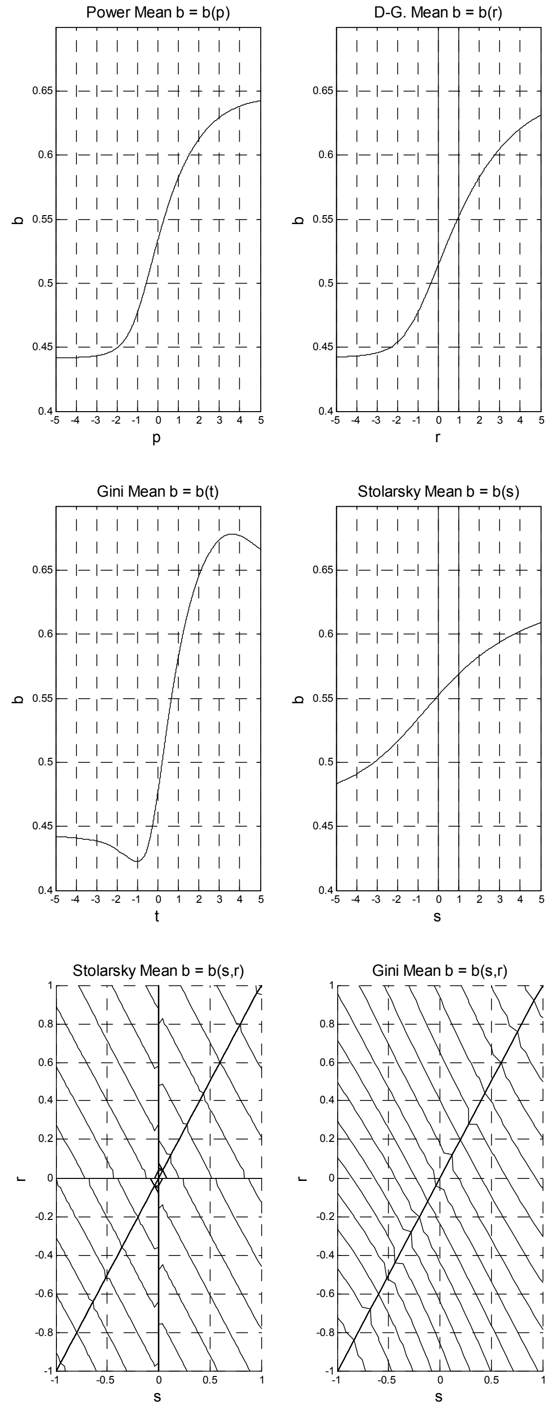
ISBN: 978-1-61804-247-7

profiles for weighted arithmetic mean regression and weighted geometric mean regression are displayed next.





For the power mean, the Dietel-Gordon mean and the Stolarsky mean, the slope equations are solved numerically, and the slope profile is plotted as a function of the parameter. For these cases, the slope is a strictly increasing function of the parameter with parameter values $\pm\infty$ corresponding to ordinary least-squares regression. The Gini mean has the same asymptotic behavior as the other generalized means, but it experiences an undershoot and an overshoot. This means that it is not one-to-one over $(-\infty, \infty)$. However, one can restrict the parameter t to the subinterval $[-0.3737, 2.0113]$ over which the slope is one-to-one and lies between the two OLS lines. Again, in the Dietel-Gordon plot, $s = 0$ and $s = 1$, correspond to the two logarithmic mean regressions. In the Stolarsky plot, $s = 0$ corresponds to LMR and $s = 1$ corresponds to IMR.



Again, in the third pair of plots, slope contours are plotted in the (s, r) parameter plane for Stolarsky's two-parameter mean and Gini's two-parameter mean. The line $r = s$ which appears in the graphs is a removable singularity. The contours are increasing as s and r increase. For the Stolarsky slope equation , thirteen contours $b = 0.496, ..., 0.568$ are plotted in increments of 0.006. For the Gini equation, sixteen contours $b = 0.436, ..., 0.616$ are plotted in increments of 0.012. The jaggedness of the contours is again an artifact of the plotting algorithm.

IX. Summary

Least-squares regressions arising from generalized means are explored. The notion of a generalized mean is equivalent to the generating function concept of the previous work but allows for a more robust understanding of what generalized least-squares is about. Generalized least-squares seeks a line which minimizes the average generalized mean of the square deviations in x and y.

The theory is reviewed and trigonometric formulas are also derived relating the parameters γ and λ in the fundamental generalized least-squares formula. An admissibility condition for least-squares lines in general and for the OLS $x|y$ line in particular is discussed. The OLS $x|y$ line is admissible, in the sense that its coefficients minimize the error function, provided that $\frac{1}{2} \leq \rho^2 \leq 1$.

The specific cases of arithmetic, geometric and harmonic mean regression were already explored in the previous papers with AMR called Pythagorean regression and HMR noted to be equivalent to orthogonal regression. Here, logarithmic, Heronian, centroidal, identric, Lorentz, and root mean square regressions are described for the first time. Ordinary least-squares regression is also shown here to be equivalent to minimum or maximum mean regression. Regressions based on weighted arithmetic means of order α and weighted geometric means of order β are explored. The weights α and β parameterize all generalized mean square regression lines lying between the two ordinary least-squares lines.

Power mean regression of order p is shown to be another particularly simple framework for parameterizing all the generalized mean square regressions previously described. The p-scale has fixed numerical values corresponding to many known special means. Since the power mean is an increasing function of p, the parameter p concisely explains the inequalities of the special means and why the same order is found among the corresponding regression lines. The parameter p has an explicit formula in terms of b, ρ, and $\frac{\sigma_y}{\sigma_x}$. Under certain conditions, every regression line lying between the two ordinary least-squares lines is generated by a power mean of order p for some p in $(-\infty, \infty)$ with the ordinary least-squares lines corresponding to $p = \pm\infty$.

The Dietel-Gordon mean of order r, Stolarsky's logarithmic mean of order s, Gini's mean of order t, Gini's two-parameter mean and Stolarsky's two-parameter mean each provide an alternative parameterization and classification. Linear approximations relating the parameters are derived. The parameter scales again have fixed numerical values corresponding to many known special means. The parameters again concisely explain the inequalities of the known special means and why the same order is found among the corresponding regression lines. Again under certain conditions, every regression line lying between the two ordinary least-squares lines is generated by these generalized means for some value of the parameters.

References

[1] B. C. Carlson. "The Logarithmic Mean." *American Mathematical Monthly*, vol. 79, pp. 615-618, 1972.

[2] H. Chen. "Means Generated by an Integral." *Mathematics Magazine*, vol. 82, p. 370, Dec. 2005.

[3] B. C. Dietel and R. A. Gordon. "Using Tangent Lines to Define Means." *Mathematics Magazine*, vol. 76, pp.52-62, Feb. 2003.

[4] S. C. Ehrenberg. "Deriving the Least-Squares Regression Equation." *The American Statistician*, vol. 37, p.232, Aug. 1983.

[5] H. Eves. "Means Appearing in Geometric Figures." *Mathematics Magazine*, vol. 76, No. 4, pp.292-294, Oct. 2003.

[6] C. Gini. "Di una formula comprensiva della medie." *Metron*, vol. 13, pp. 3-22, 1938.

[7] N. Greene. "Generalized Least-Squares Regressions I: Efficient Derivations," in *Proceedings of the 1st International Conference on Computational Science and Engineering (CSE'13)*, 2013, pp. 145-158.

[8] N. Greene. "Generalized Least-Squares Regressions II: Theory and Classification," in *Proceedings of the 1st International Conference on Computational Science and Engineering (CSE '13)*, 2013, pp. 159-166.

[9] N. Greene. "Generalized Least-Squares Regressions III: Further Theory and Classification," in *Proceedings of the 5th International Conference on Applied Mathematics and Informatics (AMATHI '14)*, 2014, pp. 34-38.

[10] E. B. Leach and M. C. Sholander. "Extended Mean Values." *American Mathematical Monthly*, vol. 85, pp. 84-90, 1978.

[11] D. H. Lehmer. "On the compounding of certain means." *J. Math. Anal. Appl.*, vol. 36, pp. 183-200, 1971.

[12] T.-P. Lin. "The power mean and the logarithmic mean." *American Mathematical Monthly*, vol. 81, pp. 879-883, 1974.

[13] M. E. Mays. "Functions Which Parametrize Means." *American Mathematical Monthly*, vol 90, pp. 677-683, 1983.

[14] K. B. Stolarsky. "Generalizations of the logarithmic mean." *Mathematics Magazine*, vol. 48, pp. 87-92, 1975.

[15] K. B. Stolarsky. "The power mean and generalized logarithmic means." *American Mathematical Monthly*, vol. 82, pp. 545-548, 1980.