

Perform Power Regression in R (Step-by-Step)

Authored by
Mohammed loot

November 5, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Perform Power Regression in R (Step-by-Step)*.
PSYCHOLOGICAL STATISTICS. Retrieved from
<https://statistics.arabpsychology.com/?p=10487>

Modeling complex relationships that exhibit exponential growth or decay often requires specialized statistical techniques beyond simple linearity. Understanding and applying [Power regression](#) is critical in these scenarios. This method, a foundational type of [non-linear regression](#), proves indispensable across disciplines such as physics, biological sciences, and econometrics, where relationships frequently adhere to a power law rather than a straight line.

The core mathematical expression that defines the power relationship is straightforward yet powerful:

$$y = ax^b$$

This formulation allows us to quantify a curvilinear trend. It is essential for analysts to recognize the statistical role of each component within this equation to ensure correct interpretation and application of the resulting model coefficients.

y: Represents the [response variable](#), sometimes called the dependent variable, which is the outcome being modeled or predicted.

x: Denotes the [predictor variable](#), or independent variable, whose influence on y is being measured.

a: This is the scaling coefficient (analogous to an intercept in linear models) that sets the baseline magnitude of the relationship.

b: Known as the power exponent, this coefficient quantifies the rate and curvature of the non-linear relationship between x and y.

Power regression is utilized specifically when the magnitude of the response variable is theorized to change proportionally to the predictor variable raised to a constant power. The following comprehensive, step-by-step guide demonstrates the precise methodology for executing a power regression analysis on a typical dataset using the statistical programming environment, [R](#).

Setting Up the R Environment and Generating Sample Data

Before any model fitting can commence, the initial and most critical step involves ensuring the data is correctly structured and loaded into the R environment. For this instructional demonstration, we will forego loading an external file and instead generate a synthetic dataset. This approach offers clarity by allowing us to work with a dataset where we know, or strongly suspect, that a power relationship exists.

Our synthetic data will consist of 20 observations for the predictor variable (x) and the corresponding response variable (y). Defining the data this way establishes the foundation for accurate statistical analysis, emphasizing the importance of correct variable assignment and data integrity prior to modeling.

The following R commands create and assign the vectors for our working dataset:

```
#create data
```

```
x=1:20
```

```
y=c(1, 8, 5, 7, 6, 20, 15, 19, 23, 37, 33, 38, 49, 50, 56, 52, 70, 89, 97, 115)
```

Executing this code in the R console prepares our environment. This structured data, with its inherent non-linear characteristics, is now ready for the vital diagnostic stage of visualization, which helps confirm the initial choice of a power regression model.

Diagnostic Visualization of the Curvilinear Relationship

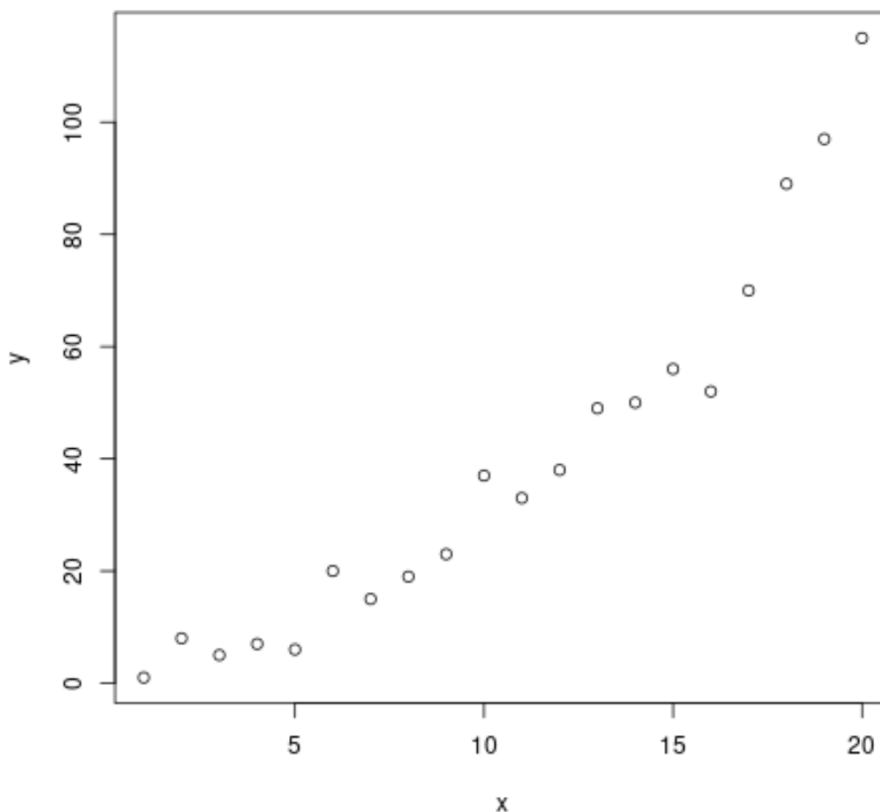
A fundamental requirement in any robust regression analysis is the graphical inspection of the raw data. By generating a scatterplot, the analyst can visually ascertain the functional form of the relationship between the predictor and response variables, which in turn dictates the appropriate modeling technique.

We utilize R's standard `plot()` function to produce this visualization, passing the predictor (x) and response (y) variables:

```
#create scatterplot
```

```
plot(x, y)
```

The resulting plot provides immediate insight into the underlying data structure:



A careful review of the scatterplot confirms that the relationship is distinctly **curvilinear**. We observe a pattern where the slope increases as x increases, characteristic of an accelerating, non-linear trend. This visual evidence provides compelling justification for rejecting a standard **linear regression** model, as such a model would violate the core assumption of linearity and result in biased parameter estimates. Consequently, fitting a power regression equation is the statistically rigorous method required to accurately capture this specific non-linear trajectory.

Linearizing the Model: The Logarithmic Transformation

The primary methodological challenge in solving the power regression equation ($y = ax^b$) lies in its inherently non-linear structure, which prevents the direct application of common techniques like **Ordinary Least Squares (OLS)** estimation. To overcome this, we must transform the equation into a linear form that the `lm()` function in R can efficiently solve.

This crucial linearization is achieved by applying the natural logarithm (\ln) to both sides of the original equation. This mathematical manipulation converts the multiplicative relationship into an additive one:

Original Equation: $y = ax^b$

Logarithmic Transformation: $\ln(y) = \ln(a) + b * \ln(x)$

The transformed equation now mirrors the structure of a simple linear model: $Y' = A + B * X'$, where $Y' = \ln(y)$, $X' = \ln(x)$, $A = \ln(a)$, and $B = b$. This allows us to use R's standard linear modeling function, applying the `log()` transformation to both the response and predictor variables directly within the model specification:

```
#fit the model
```

```
model <- lm(log(y)~ log(x))
```

```
#view the output of the model
```

```
summary(model)
```

Call:

```
lm(formula = log(y) ~ log(x))
```

Residuals:

```
Min 1Q Median 3Q Max
```

```
-0.67014 -0.17190 -0.05341 0.16343 0.93186
```

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 0.15333 0.20332 0.754 0.461
```

```
log(x) 1.43439 0.08996 15.945 4.62e-12 ***
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.3187 on 18 degrees of freedom

Multiple R-squared: 0.9339, Adjusted R-squared: 0.9302

F-statistic: 254.2 on 1 and 18 DF, p-value: 4.619e-12

Interpreting Statistical Diagnostics and Coefficients

The output generated by `summary(model)` provides essential diagnostic statistics that confirm the successful transformation and fitting of the model. Evaluating these metrics is crucial for determining the statistical significance and overall quality of the linear relationship established between the log-transformed variables.

We first focus on the **Coefficients** table. The estimates--0.15333 for the intercept and 1.43439 for $\log(x)$ --are the parameters of the linearized equation, $\ln(y) = \ln(a) + b * \ln(x)$. The coefficient for $\log(x)$ is highly significant, evidenced by its extremely low **p-value** (4.62e-12). This robust

significance confirms a strong, non-random relationship between the logarithm of the predictor variable and the logarithm of the [response variable](#).

Next, we assess the overall explanatory power of the model. The **Multiple R-squared** value of 0.9339 indicates that approximately 93.39% of the variability in $\ln(y)$ is successfully accounted for by $\ln(x)$. This high value suggests an exceptionally good fit for the linearized data. The **Adjusted R-squared** (0.9302) provides a slightly more conservative estimate, adjusting for the single predictor used. Furthermore, the substantial **F-statistic** (254.2) and its associated minute **p-value** ($4.619e-12$) collectively affirm that the model is statistically significant and performs far better than any null model lacking the [predictor variable](#).

Deriving the Final Equation and Making Predictions

While the linearized equation is necessary for calculation via OLS, the final objective is to express the relationship in its original, untransformed power form ($y = ax^b$) for practical use and interpretation. We begin by substituting the estimated coefficients from the R output into the logarithmic equation:

$$\ln(y) = 0.15333 + 1.43439\ln(x)$$

To isolate y , we must apply the inverse of the natural logarithm--the exponential function (e)--to both sides of the equation. This reverse transformation returns the relationship to its intended non-linear state:

The systematic derivation involves standard exponential and logarithmic rules:

$$y = e^{(0.15333 + 1.43439\ln(x))}$$
 (Applying the exponential function to the entire right side.)

$$y = e^{0.15333} * e^{1.43439\ln(x)}$$
 (Separating the additive terms using the exponent rule $e^{A+B} = e^A * e^B$.)

$$y = e^{0.15333} * x^{1.43439}$$
 (Simplifying the second term using the logarithm rule $e^{k\ln(x)} = x^k$.)

Calculating the constant a ($e^{0.15333} \approx 1.1657$) and recognizing that the exponent b is directly derived from the coefficient of **log(x)** (1.43439), we arrive at the final fitted power regression equation:

$$y = 1.1657x^{1.43439}$$

This derived equation is now fully functional for practical predictive applications. For instance, if we need to predict the value of the [response variable](#), y , when the [predictor variable](#), x , is equal to 12, we substitute this value into the equation:

$$y = 1.1657(12)^{1.43439}$$

The resulting predicted value for y is approximately **41.167**. This demonstrates how the fitted power model provides a statistically robust method for forecasting outcomes governed by power law dynamics, effectively bridging the gap between the necessary linearized estimation process and the actual non-linear reality of the data.

For individuals requiring immediate results without undertaking the manual logarithmic transformation and derivation steps, numerous specialized statistical software packages and online regression tools are available to streamline the process of fitting complex non-linear models like power regression.

Additional Resources for Advanced Modeling

The principles demonstrated here form the basis for many advanced statistical modeling techniques. Further exploration into specialized [non-linear regression](#) methods and statistical software utilization is highly recommended for analysts seeking a deeper analytical capability.