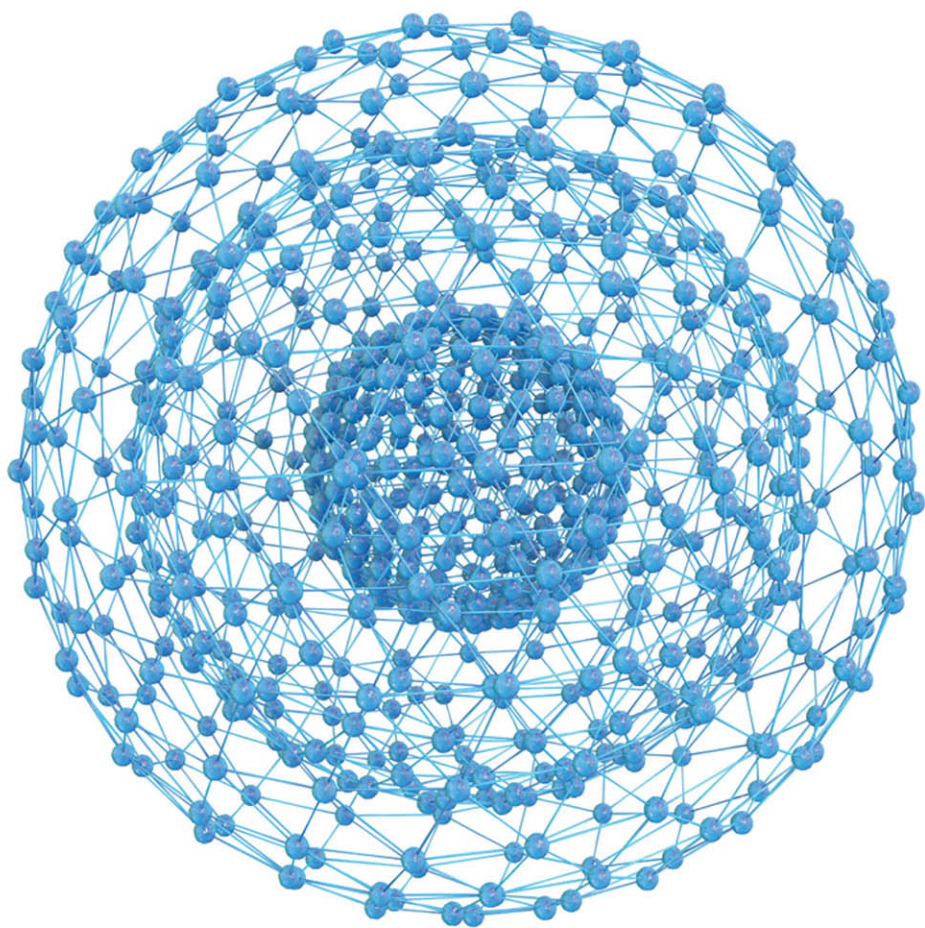


# REGRESSION ANALYSIS

Microsoft® Excel®



# Regression Analysis Microsoft® Excel®

*Conrad Carlberg*

**que®**

800 East 96th Street,  
Indianapolis, Indiana 46240 USA

## C o n t e n t s   a t   a   G l a n c e

Introduction.....	1
1 Measuring Variation: How Values Differ.....	5
2 Correlation .....	29
3 Simple Regression.....	59
4 Using the LINEST() Function.....	103
5 Multiple Regression .....	151
6 Assumptions and Cautions Regarding Regression Analysis.....	199
7 Using Regression to Test Differences Between Group Means .....	245
8 The Analysis of Covariance .....	295
Index.....	337

# Regression Analysis Microsoft Excel

## Table of Contents

Cover

Title Page

Copyright Page

Contents

Introduction

### 1 Measuring Variation: How Values Differ

How Variation Is Measured

Sum of Deviations

Summing Squared Deviations

From the Sum of Squares to the Variance

Using the VAR.P ( ) and VAR.S ( ) Functions

The Standard Deviation

The Standard Error of the Mean

About z-Scores and z-Values

About t-Values

### 2 Correlation

Measuring Correlation

Expressing the Strength of a Correlation

Determining a Correlations Direction

Calculating Correlation

Step One: The Covariance

Watching for Signs

From the Covariance to the Correlation Coefficient

# Table of Contents

Using the CORREL( ) Function

Understanding Bias in the Correlation

Checking for Linearity and Outliers in the Correlation

Avoiding a Trap in Charting

## Correlation and Causation

Direction of Cause

A Third Variable

## Restriction of Range

## 3 Simple Regression

### Predicting with Correlation and Standard Scores

Calculating the Predictions

Returning to the Original Metric

Generalizing the Predictions

### Predicting with Regression Coefficient and Intercept

The SLOPE( ) Function

The INTERCEPT( ) Function

Charting the Predictions

## Shared Variance

The Standard Deviation, Reviewed

More About Sums of Squares

Sums of Squares Are Additive

$R^2$  in Simple Linear Regression

Sum of Squares Residual versus Sum of Squares Within

## The TREND( ) Function

Array-entering TREND( )

TREND( )s new xs Argument

TREND( )s const Argument

Calculating the Zero-constant Regression

## Partial and Semipartial Correlations

# Table of Contents

Partial Correlation

Understanding Semipartial Correlations

## 4 Using the LINEST( ) Function

Array-Entering LINEST( )

Understanding the Mechanics of Array Formulas

Inventorying the Mistakes

Comparing LINEST( ) to SLOPE( ) and INTERCEPT( )

The Standard Error of a Regression Coefficient

The Meaning of the Standard Error of a Regression Coefficient

A Regression Coefficient of Zero

Measuring the Probability That the Coefficient is Zero in the Population

Statistical Inference as a Subjective Decision

The t-ratio and the F-ratio

Interval Scales and Nominal Scales

The Squared Correlation,  $R^2$

The Standard Error of Estimate

The t Distribution and Standard Errors

Standard Error as a Standard Deviation of Residuals

Homoscedasticity: Equal Spread

Understanding LINEST( )'s F-ratio

The Analysis of Variance and the F-ratio in Traditional Usage

The Analysis of Variance and the F-ratio in Regression

Partitioning the Sums of Squares in Regression

The F-ratio in the Analysis of Variance

The F-ratio in Regression Analysis

The F-ratio Compared to  $R^2$

The General Linear Model, ANOVA, and Regression Analysis

Other Ancillary Statistics from LINEST( )

# Table of Contents

## 5 Multiple Regression

### A Composite Predictor Variable

Generalizing from the Single to the Multiple Predictor

Minimizing the Sum of the Squared Errors

### Understanding the Trendline

### Mapping LINEST( )s Results to the Worksheet

### Building a Multiple Regression Analysis from the Ground Up

Holding Variables Constant

Semipartial Correlation in a Two-Predictor Regression

Finding the Sums of Squares

$R^2$  and Standard Error of Estimate

F-Ratio and Residual Degrees of Freedom

Calculating the Standard Errors of the Regression Coefficients

Some Further Examples

### Using the Standard Error of the Regression Coefficient

Arranging a Two-Tailed Test

Arranging a One-Tailed Test

### Using the Models Comparison Approach to Evaluating Predictors

Obtaining the Models Statistics

Using Sums of Squares Instead of  $R^2$

### Estimating Shrinkage in $R^2$

## 6 Assumptions and Cautions Regarding Regression Analysis

### About Assumptions

Robustness: It Might Not Matter

Assumptions and Statistical Inference

### The Straw Man

### Coping with Nonlinear and Other Problem Distributions

### The Assumption of Equal Spread

# **Table of Contents**

Using Dummy Coding

Comparing the Regression Approach to the t-test Approach

Two Routes to the Same Destination

## **Unequal Variances and Sample Sizes**

Unequal Spread: Conservative Tests

Unequal Spread: Liberal Tests

Unequal Spreads and Equal Sample Sizes

Using LINEST( ) Instead of the Data Analysis Tool

Understanding the Differences Between the T.DIST( ) Functions

Using Welchs Correction

The TTEST( ) Function

## **7 Using Regression to Test Differences Between Group Means**

### **Dummy Coding**

An Example with Dummy Coding

Populating the Vectors Automatically

The Dunnett Multiple Comparison Procedure

### **Effect Coding**

Coding with -1 Instead of 0

Relationship to the General Linear Model

Multiple Comparisons with Effect Coding

### **Orthogonal Coding**

Establishing the Contrasts

Planned Orthogonal Contrasts Via ANOVA

Planned Orthogonal Contrasts Using LINEST( )

### **Factorial Analysis**

Factorial Analysis with Orthogonal Coding

Factorial Analysis with Effect Coding

### **Statistical Power, Type I and Type II Errors**

# **Table of Contents**

Calculating Statistical Power

Increasing Statistical Power

Coping with Unequal Cell Sizes

Using the Regression Approach

Sequential Variance Assignment

## **8 The Analysis of Covariance**

Contrasting the Results

ANCOVA Charted

Structuring a Conventional ANCOVA

Analysis Without the Covariate

Analysis with the Covariate

Structuring an ANCOVA Using Regression

Checking for a Common Regression Line

Summarizing the Analysis

Testing the Adjusted Means: Planned Orthogonal Coding in ANCOVA

ANCOVA and Multiple Comparisons Using the Regression Approach

Multiple Comparisons via Planned Nonorthogonal Contrasts

Multiple Comparisons with Post Hoc Nonorthogonal Contrasts

Index