

# Math 362: Mathematical Statistics II

Le Chen

le.chen@emory.edu

Emory University  
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

# Chapter 11. Regression

§ 11.1 Introduction

§ 11.4 Covariance and Correlation

§ 11.2 The Method of Least Squares

§ 11.3 The Linear Model

§ 11.A Appendix Multiple/Multivariate Linear Regression

§ 11.5 The Bivariate Normal Distribution

# Chapter 11. Regression

§ 11.1 Introduction

§ 11.4 Covariance and Correlation

§ 11.2 The Method of Least Squares

§ 11.3 The Linear Model

§ 11.A Appendix Multiple/Multivariate Linear Regression

§ 11.5 The Bivariate Normal Distribution

# Regression analysis

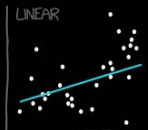
FITS A STRAIGHT LINE TO THIS MESSY SCATTERPLOT.  $x$  IS CALLED THE INDEPENDENT OR PREDICTOR VARIABLE, AND  $y$  IS THE DEPENDENT OR RESPONSE VARIABLE. THE REGRESSION OR PREDICTION LINE HAS THE FORM

$$y = a + bx$$

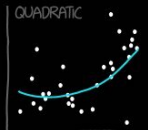


<https://madhureshkumar.wordpress.com/>

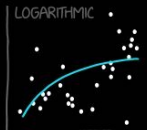
## CURVE-FITTING METHODS AND THE MESSAGES THEY SEND



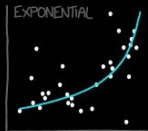
"HEY, I DID A  
REGRESSION."



"I WANTED A CURVED  
LINE, SO I MADE ONE  
WITH MATH."



"LOOK, IT'S  
TAPERING OFF!"



"LOOK, IT'S GROWING  
UNCONTROLLABLY!"



"I'M SOPHISTICATED, NOT  
LIKE THOSE BUMBLING  
POLYNOMIAL PEOPLE."



"I'M MAKING A  
SCATTER PLOT BUT  
I DON'T WANT TO."

<https://xkcd.com/>

## Three ways to view the same thing

$$(x_1, y_1), \dots, (x_n, y_n)$$

1. Purely data, no probability structure assumed.

$$(x_1, Y_1), \dots, (x_n, Y_n)$$

2. A random sample of size  $n$ , where  $Y_i$  follows a distribution depending on  $x_i$  which is deterministic.

$$(X_1, Y_1), \dots, (X_n, Y_n)$$

3. A random sample of size  $n$ , where  $(X_i, Y_i)$  follow some joint distribution.