Probabilistic Models

DASC 512

Probabilistic Models

- Goal: Compare quantitative variables using a model in which the population mean of one variable is a function of other variables
- This week, we'll be talking about models with one response variable and one factor variable

- Examples:
- Automobile fuel economy (mpg) is dependent on horsepower
- Rocket propellant strength is dependent on the age of the propellant
- Weight is dependent on height

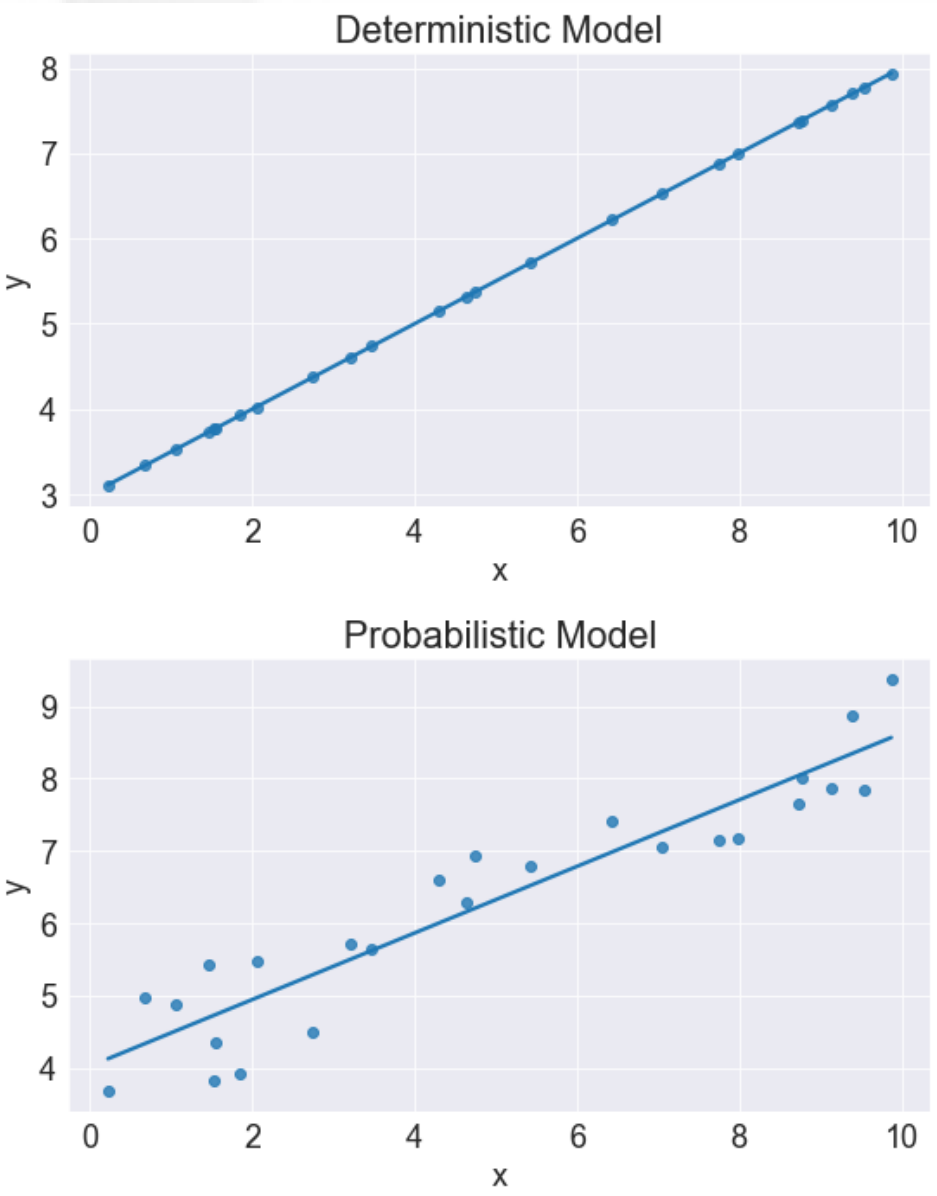
Deterministic vs. Probabilistic

- If an exact relationship exists between two variables, then the model of this relationship is called a deterministic model

$$y = \frac{x}{2} + 3$$

- Often, random phenomena cause the relationship to be inexact, instead having some random error present. This results in a probabilistic model

$$y = \frac{x}{2} + 3 + \epsilon, \qquad \epsilon \sim N(0, \sigma^2)$$



Deterministic vs. Probabilistic

Deterministic models:

- Position of a falling object after falling for a specified time
- Degree of lunar eclipse in a location by time and date
- Area of a square by length of a side
- Circumference of a circle by diameter

Probabilistic models:

- Waiting time at the CAC office by time of day
- School bus arrival time by ambient air temperature
- Bomb drop miss distance by angle of attack
- Outdoor temperature per hour over the next week

First Order Models

First Order (i.e., straight line) Probabilistic Model:

$$y = \text{deterministic component} + \text{random error}$$
$$y = \beta_0 + \beta_1 x + \text{random error}$$
$$y = \beta_0 + \beta_1 x + \epsilon$$

y is the actual response variable value

β_0 is the y-intercept of the line (the expected value of y for $x = 0$)

β_1 is the slope of the line (relationship between x and y)

ϵ is the random error present in the model. $\mu_\epsilon = 0$

First Order Models

The mean (expected value) of y is equal to the deterministic portion

$$E(y) = \beta_0 + \beta_1(x)$$

β_0 and β_1 must be estimated from the data in order to describe the relationship and to make inference on the mean response

Steps for Regression Modeling

1. Hypothesize a deterministic model
2. Estimate unknown parameters in the model using sample data
3. Infer the distribution of the remaining error and check assumptions.
4. Statistically evaluate the usefulness of the model
5. If the model is useful, use it for prediction or estimation.

Next time…

Least Squares Estimation