# Week 1: Descriptive Analytics

- An Operational Decision Problem

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

- Thinking about Trends and Seasonality

- Forecasting for new Products
- Fitting distributions

# Week 1: Descriptive Analytics

- An Operational Decision Problem **Session 1**

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

- Thinking about Trends and Seasonality

- Forecasting for New Products
- Fitting distributions

# Week 1: Descriptive Analytics

- An Operational Decision Problem

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

**Session 2**

- Thinking about Trends and Seasonality

- Forecasting for New Products
- Fitting distributions

# Week 1: Descriptive Analytics

- An Operational Decision Problem

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

- Thinking about Trends and Seasonality

**Session 3**

- Forecasting for New Products
- Fitting distributions

# Week 1: Descriptive Analytics

- An Operational Decision Problem

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

- Thinking about Trends and Seasonality

- Forecasting for New Products
- Fitting distributions

**Session 4**

# Week 1: Descriptive Analytics

- An Operational Decision Problem **Session 1**

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

- Thinking about Trends and Seasonality

- Forecasting for New Products
- Fitting distributions

# Descriptive Analytics

- Before we dive into analyzing data, let us a look at a fundamental problem that firms face

- An Operations problem:
  - How much to produce?

  - We need to know or estimate the cost of the product, price of the product, and some data on the demand of the product.

- Let us explore a problem to get started.

# A Fundamental Operations Problem: An example

- Suppose that you are making operations decisions for a retailer who orders a product from a supplier and sells it to customers.

- The ordered product items are received and placed on store shelf.

- There is a large customer population
  - Each customer may choose to buy or not buy the product.
  - If the customer chooses to buy, he arrives at the store to buy the product.
  - He buys it as long as it is available on the shelf.

- However, you have to order the product before you see the customer demand, since you have to have the items available on shelf.

- You get only one chance to order (i.e., you cannot change your purchase order after your decision).

# An Operations Problem:  Costs

- You order the product  from the supplier at cost = 3 talers/item. (Talers are the currency units).
- After your order is received and placed on shelves, demand occurs.
- The product on the shelf sells at price  = 12 talers/item.
- All unsold items are salvaged. Salvage value =0 talers/item.
- Let us look at timeline of events.

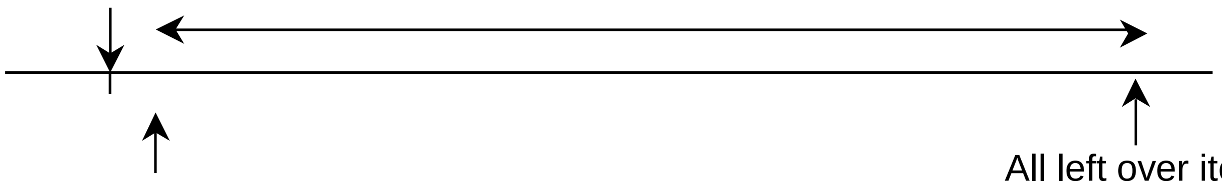
# Timeline of Events

Submit an order
to your supplier.

cost = 3 talers/item

Uncertain Demand occurs.
Items on shelves sell as long as they are available

Selling price = 12 talers/item.



Receive all ordered
items and shelve them.

All left over items that are unsold
are salvaged.

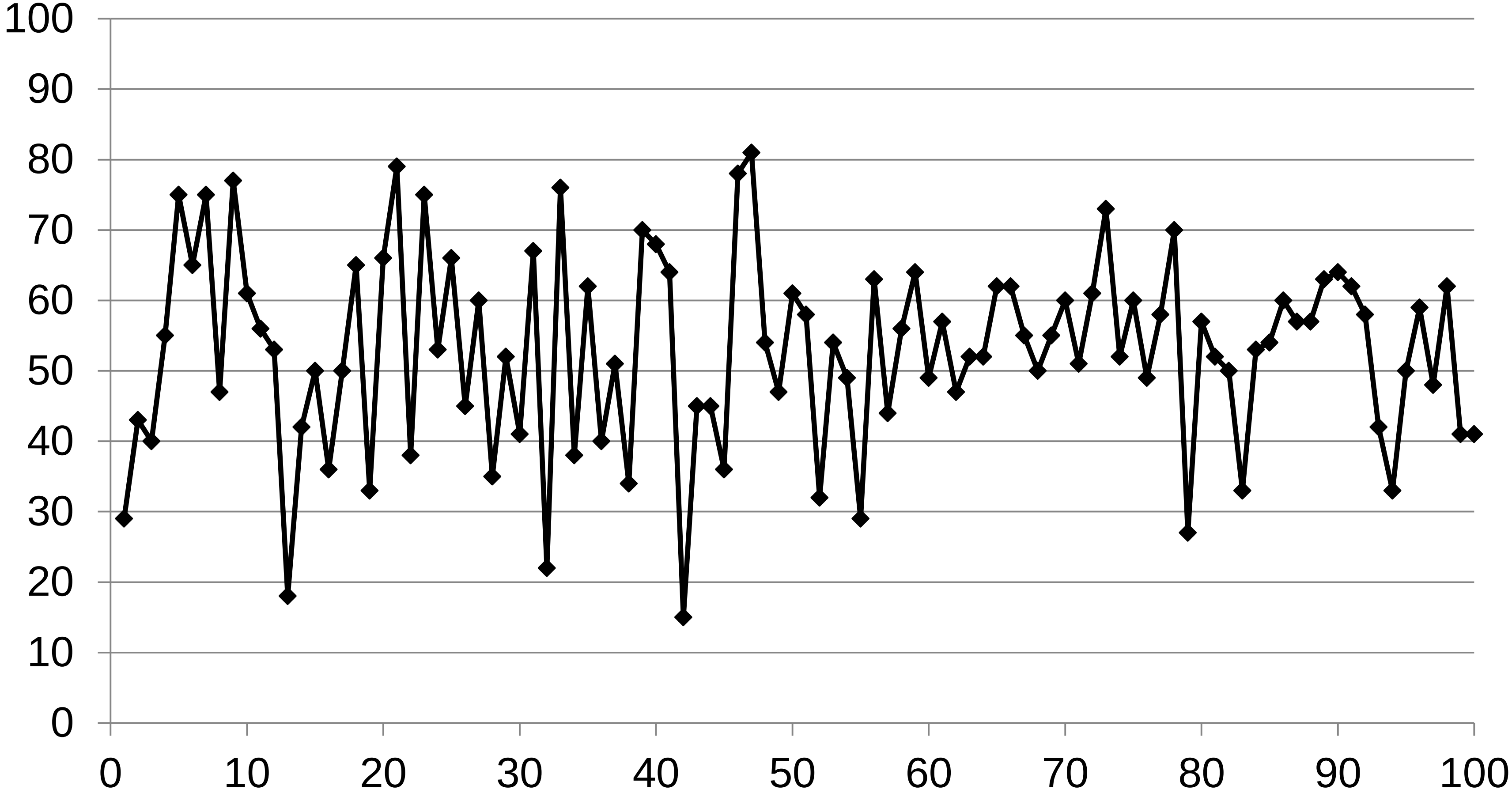
Salvage value = 0 talers/item.

- Demand is uncertain. Suppose you bought 10 items
  - A High Demand Scenario: Demand is 100. You will sell all 10 items, and make a profit of 10*(12-3)=90 talers.
  - A Low demand Scenario: No demand (i.e., demand = 0). You sell nothing and lose 10*3=30 talers.

# Problem Recap

- You don’t know what the demand is going to be…
- You have to decide on the number of units to order from supplier before seeing the customer demand.
- What could help?
  - Past demand data…
  - Fortunately, we have the demand data from past 100 periods.

# Past Demand Data




- The chart shows the demands (y-axis) observed in past 100 periods (x-axis).

# Past Demand Data

- Some more information from past demand data
- From the observations over the past 100 such periods.
  - Maximum Demand observed was 81.
  - Minimum Demand observed was 15.
  - The arithmetic average of those 100 observations is 52.8
- Based on the data, I am going to ask you to go through an exercise
  - on deciding how much to order.

# Before you make your decision

- There is no penalty for a wrong answer, or conversely, no extra course credit for the right answer.
- You get one attempt at making your decision.
- The objective of the exercise is not to test or grade you, but to set a baseline “initial thinking” as we start the course.
- Write down your answer on a sheet of paper and keep the sheet through the course.
- We will see the best answer and you will then get a chance to compare your answers and calibrate learning progress.

# Question: How much would you order…

- Suppose you are a manager contemplating the question of how many items to order from the supplier.
- Choose the quantity (Q) that you will order.
- Once you select Q, the market will produce 50 random demand instances from the distribution of demand similar to the Figure I showed you.
- Each random demand instance will correspond to the demand value you may face in the coming selling season.
- Your objective is to select Q to maximize total profit that you will earn when faced with these 50 random demand values.

# Newsvendor Problem

- The problem you just saw is called a Newsvendor problem.
  - Its characteristics are:
    - » You have an objective (usually maximize profits, minimize costs, improve market share, etc.)
    - » You have to make one decision (usually, how much to buy, or plan for).
    - » … before you see the future demand
    - » Demand occurs, and profits and costs are realized.

- This is called the newsvendor problem:
  - because it is similar to a vendor who sells newspapers:
    - » Buy too much and you may be left with unsold newspapers,
    - » or buy too little, and you will forgo revenue opportunity.

- In this course, we will show you how to think about and analyze this problem.

# A Business Application at *Time Inc.*

- Time Magazine Supply chain:
  - Stores were either selling out inventories (too little inventory)
  - or sold only a small fraction of allocation (too much inventory).

- Time Magazine evaluated and adjusted for every issue:
  - National print order (total number of copies printed and shipped),
  - Wholesale allotment structure (How those copies are allotted to wholesalers).
  - Store distribution (Final distribution to stores).

- Note: above three decisions are made before the actual demand is realized
  - Need to analyze past data
  - Forecast future demand.

- Time Magazine reports saving $3.5M annually from tackling the newsvendor problem.
  - Koschat et al, *Interfaces*, Volume 33, No 3. May-June 2003, pages 72-84.

# Broader applications of the Newsvendor problem

- Governments order flu vaccines before the flu season begins, and before the extent or the nature of the flu strain is known
  - How many vaccines to order?

- Smart phone users buy mobile data plans before they know their actual future usage
  - What is the right plan for you?

- Consumers buy health insurance plans, before they know their actual health expenditures
  - How to think about the right plans?

- For all the above examples: some forecast of future demand is essential

# Introduction to Forecasting

- What is forecasting?
  - Primary Function is to Predict the Future
- Why are we interested?
  - Dictates the decisions we make today
- Examples: who uses forecasting in their jobs?
  - forecast demand for products and services
  - forecast inventory and capacity needs daily
- What makes a good forecast?
  - It should be timely, reliable.
  - It should be as accurate as possible, and
  - It should be in meaningful units
  - The method should be easy to use and be understood in practice.

# Characteristics of Forecasts

- Point forecasts are usually wrong! Why?
  - Examples:  In December 2015, there will be 37cms of snow.
  - We will sell 314 umbrellas during the rains next week.
  - Demand could be a random variable.
- Therefore, a good forecast should be more than a single number
  - mean and standard deviation
  - range (high and low) (e.g. weather forecasts).

# Modeling Uncertain Future: Probability Distributions

- We often do not control purchasing behavior – as a result, we cannot predict future demand with certainty
- How do we describe uncertain future demand?
- We can try to decide what future demand scenarios are possible, for each scenario, estimate the likelihood of its realization
- Where do scenarios come from?
  - Past data
  - Expert estimates

# An Example of a Model of Future Demand

- Let's start by looking at a small number of scenarios, say, three: "high demand", "ordinary demand" and "low demand".

- Let's say that "high demand" scenario corresponds to the demand value of 80, "ordinary demand" scenario – to the value of 50, and "low demand" scenario - to a value of 20

- For each scenario, a likelihood of its occurring must be estimated

# Example of a Model of Future Demand: How Likely is Each Scenario

- For each scenario, a likelihood of its coming true must be estimated

- Where do estimates of likelihood come from?
  - Statistical analysis of past data

- Suppose that after analyzing the past data and using subjective inputs, we estimate that scenarios have the following likelihoods of being realized in the next selling season:

  - Likelihood of “high” demand is 20%
  - Likelihood of “normal” demand is 70%
  - Likelihood of “low” demand is 10%

# Three Scenarios and Probability Distribution

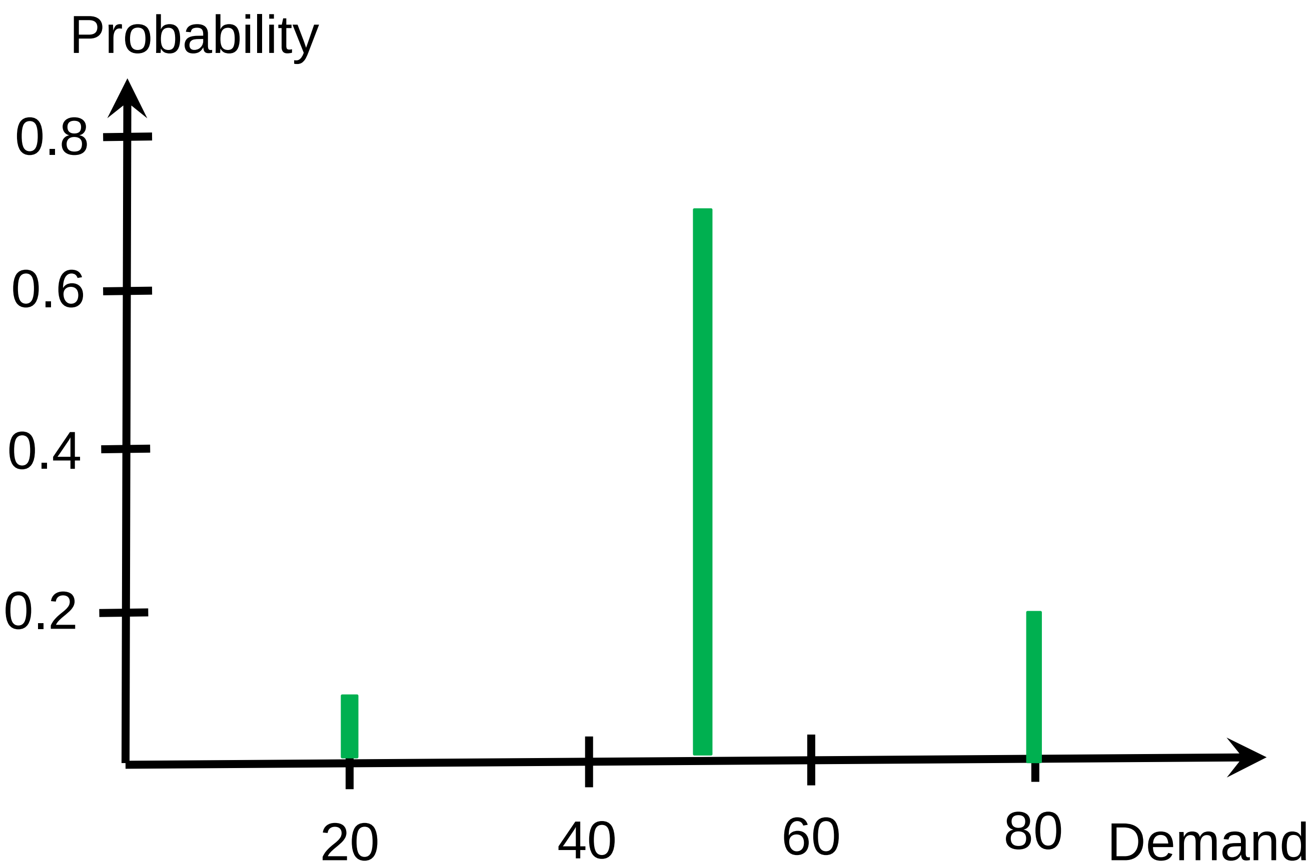
- In other words, we project that the demand is not equal to a certain number with probability 1, but, rather can take one of three values with those probabilities

- We have just created a probability distribution for the future demand:
  - $D_1$ = 80 with probability $p_1 = 0.2$
  - $D_2$ = 50 with probability $p_2 = 0.7$
  - $D_3$ = 20 with probability $p_3 = 0.1$

- Probability distributions like that one, described by a number of distinct scenarios with attached probabilities, are called **discrete**

- Note that the probabilities are
  - greater than zero, and
  - they sum upto 1.

# Three Scenarios Probability Distribution: Scenarios and Their Probabilities



# Describing Probability Distribution: Mean and Standard Deviation

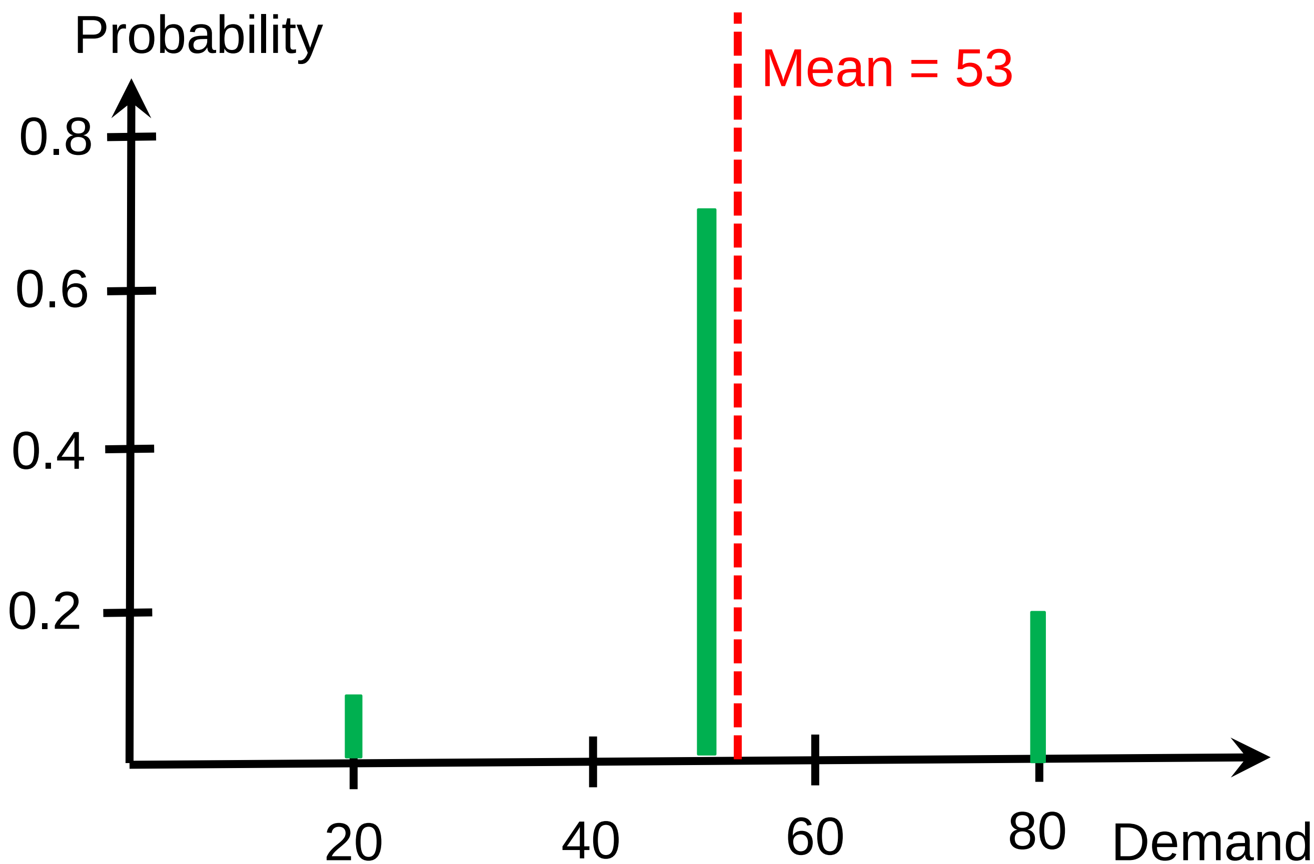
- For any probability distribution, including a simple one reflecting three demand scenarios, two useful descriptive quantities are often calculated: **mean** (also called **expected value**) and **standard deviation**

- For a discrete probability distribution, the mean is defined as a sum of the products of scenario values and their probabilities

- For our demand distribution, the mean $\overline{\mathrm{D}} = p_1 \mathrm{D}_1 + p_2 \mathrm{D}_2 + p_3 \mathrm{D}_3 = 0.2 * 80 + 0.7 * 50 + 0.1 * 20 = 53$.

- Mean reflects the demand value that we will get, on average, in a selling season, if we keep observing the demand realizations over infinite number of selling seasons

# Three Scenarios Probability Distribution: Mean



# Describing Probability Distribution: Mean and Standard Deviation

- Standard deviation describes, roughly speaking, how far away actual random variable values are from the mean, on average. In other words, it describes how, in a colloquial sense, “spread out” the distribution is around its mean

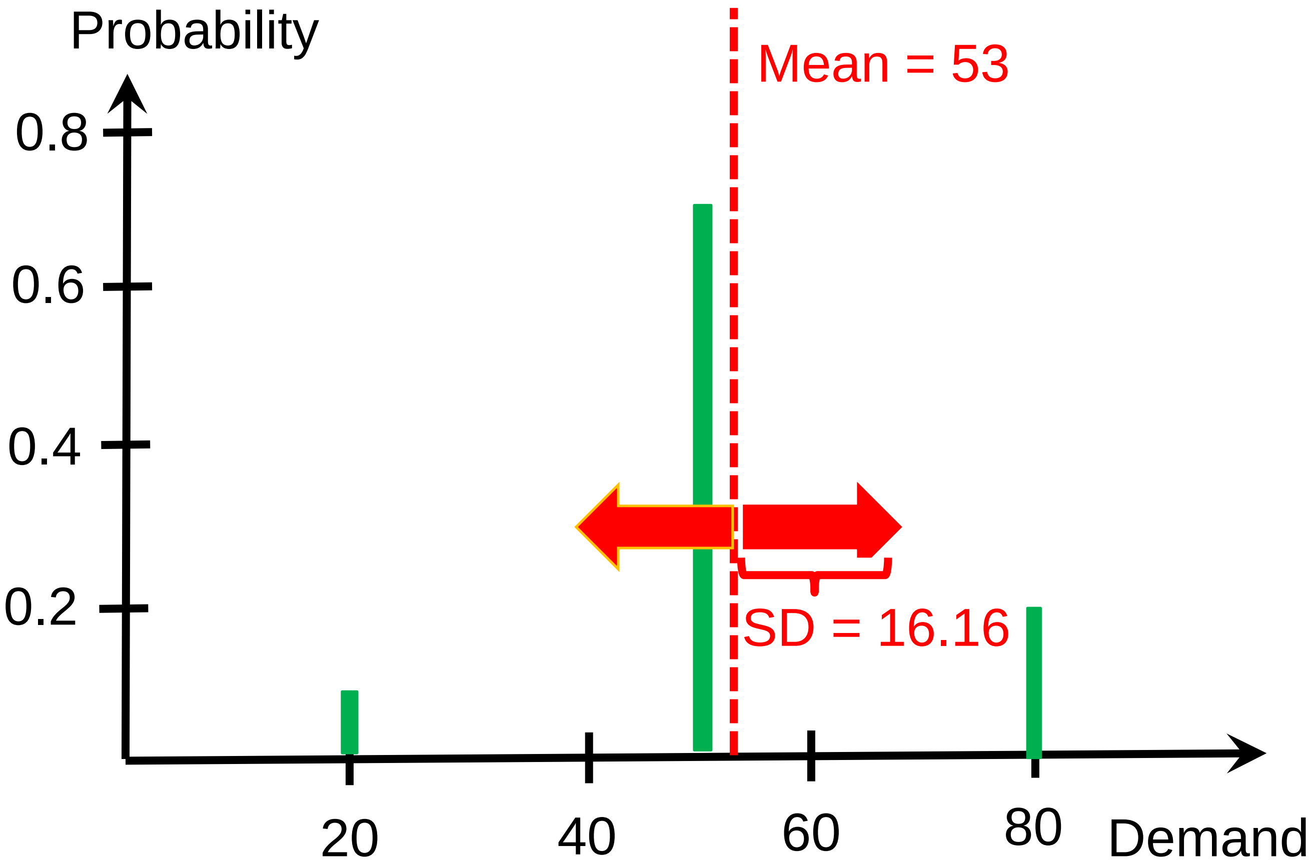
- Standard deviation is defined as a square root of the sum of products of scenario probabilities and the squares of the difference between scenario value and the mean value

- For example, for the three-scenario demand probability distribution we consider, the standard deviation is calculated as

$$\mathrm{SD} = \sqrt{p_1 * (\mathrm{D}_1 - \overline{D})^2 + p_2 * (\mathrm{D}_2 - \overline{D})^2 + p_3 * (\mathrm{D}_3 - \overline{D})^2}$$
$$= \sqrt{0.2 * (80 - 53)^2 + 0.7 * (50 - 53)^2 + 0.1 * (20 - 53)^2} \approx 16.16$$

# Three Scenarios Probability Distribution: Mean and Standard Deviation

- Knowledge of mean and standard deviation values helps to support a general intuition about the nature of a random variable



# Mean and Standard Deviation: More than three scenarios

- What if we have more than three scenarios?
  - $D_1$ with probability $p_1$
  - $D_2$ with probability $p_2$
  - $D_3$ with probability $p_3$

  ……
  - $D_n$ with probability $p_n$

  and $p_1 + p_2 + p_3 + \cdots + p_n = 1$

- What about mean and standard deviation of this demand distribution for *n* scenarios?

  Mean = $\overline{\mathrm{D}} = p_1\mathrm{D}_1 + p_2\mathrm{D}_2 + p_3\mathrm{D}_3 + \cdots + p_n\mathrm{D}_\mathrm{n}$

  Standard Deviation =

  $\sqrt{p_1 * (\mathrm{D}_1 - \overline{\mathrm{D}})^2 + p_2 * (\mathrm{D}_2 - \overline{\mathrm{D}})^2 + \cdots + p_n * (\mathrm{D}_n - \overline{\mathrm{D}})^2}$

# Discrete vs. Continuous Probability Distributions

- So far, we have looked at a discrete probability distributions with a number of future scenarios with "attached" probability for each scenario

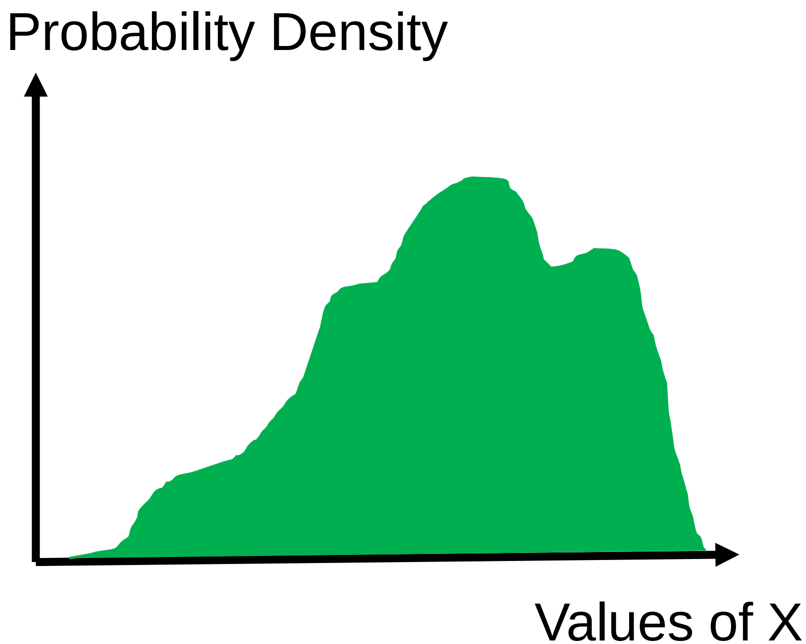
- But what will happen to a discrete probability picture when
  - a) random variable being modeled has a really large number of scenarios on any small interval of possible interval of values and
  - b) the probability that any one scenario is realized is really small

- Think of examples such as stock prices, or the amount of rainfall in a region.

- In such cases, it makes sense to describe such probability distribution using groups of scenarios rather than focusing on individual scenarios

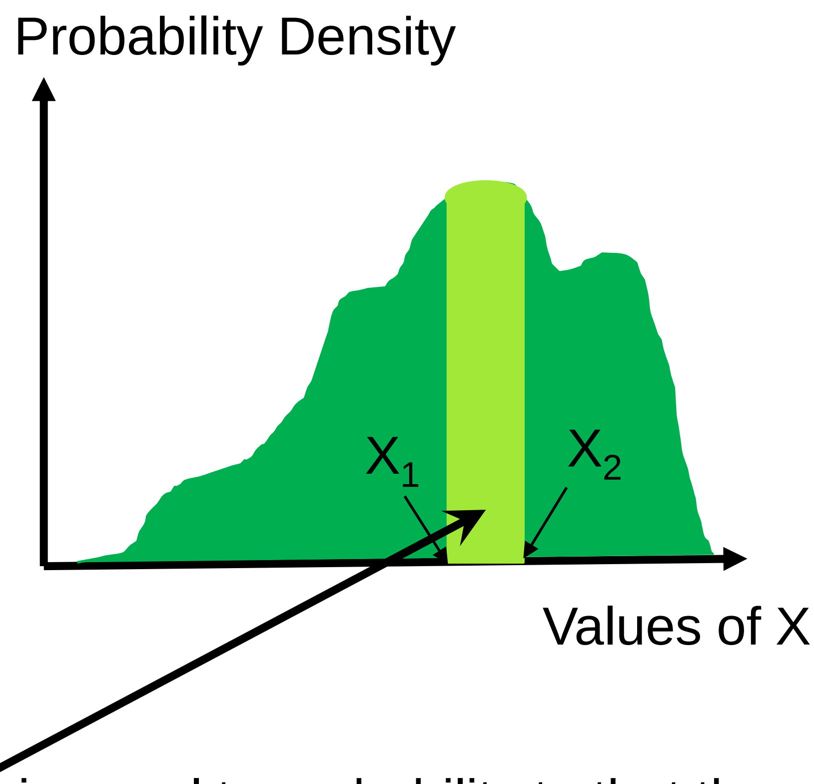
# Continuous Distribution: Random Variable X

- Distributions like this are called continuous



# Continuous Distribution: Random Variable X

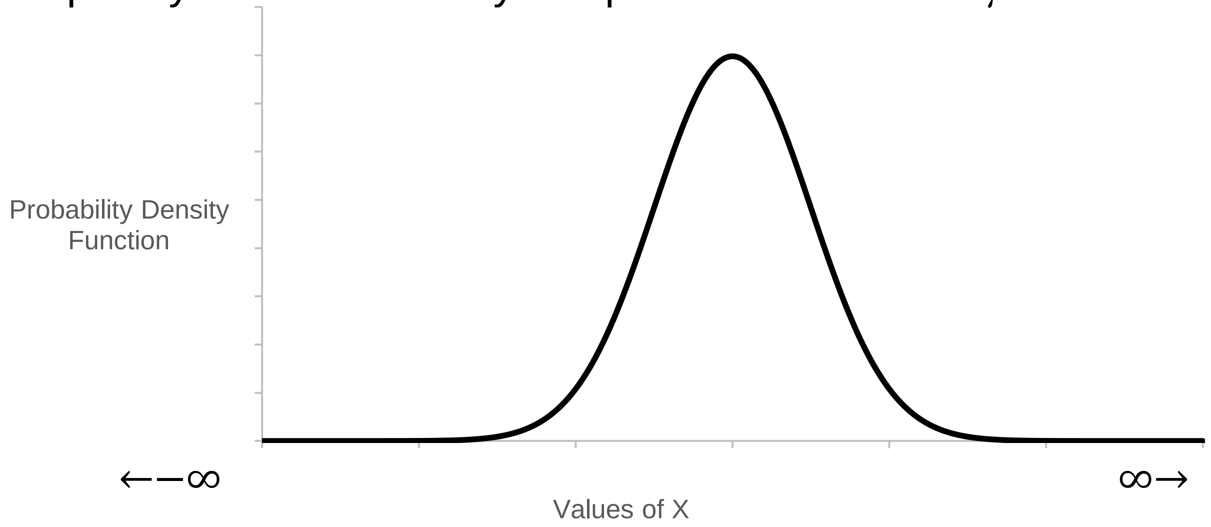
- Distributions like this are called continuous




- The area is equal to probability to that the random variable X takes values in the interval between $X_1$ and $X_2$
- The area under the entire curve is equal to 1

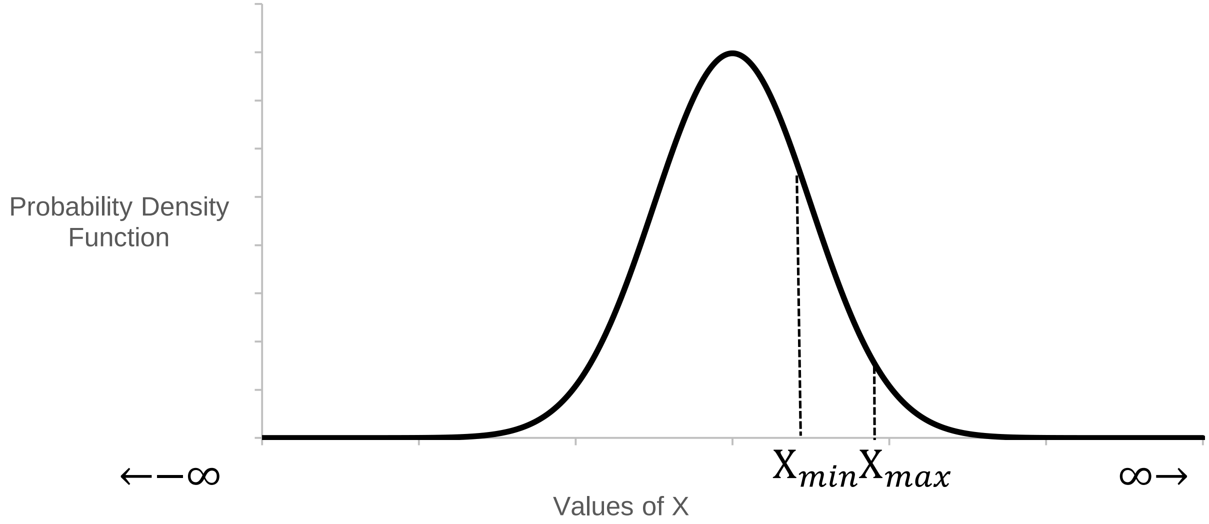
# Normal Distribution

- One of the most popular examples of a continuous probability distribution is normal distribution
- Normal distribution:
  - Allows the underlying random variable to take any value from negative infinity to positive infinity, and
  - is completely characterized by two parameters – mean $\mu$ and standard deviation $\sigma$.



# Normal Distribution

- There exist statistical formulas (also implemented in Excel) that calculate a probability that a normal random variable X with given mean $\mu$ and standard deviation $\sigma$ produces a value within a specified interval of values $[X_{min}, X_{max}]$



# Other Continuous Probability Distributions

- There exist a large number of other “popular” continuous probability distribution: exponential, beta, etc. with easily computable mean and variance/standard deviation

- Each of those distributions is often used to describe specific uncertain setting/quantity

- For example, normal distribution is used to describe a distribution of a future relative (percentage) changes in the values of stocks, FX rates

- Another example: exponential distribution can be used in characterizing time between successive arrivals of customers in service systems (e.g. call centers).

# Returning back: Characteristics of Forecasts

- Point forecasts are usually wrong! Why?
  - Demand could be a random variable
- Therefore, a good forecast should be more than a single number
- Forecasts should include some distribution information
  - mean and standard deviation
  - range (high and low)
- Aggregate forecasts are usually more accurate
- Accuracy of forecasts erodes as we go further into the future
- Don’t exclude known information

# Subjective Forecasting Methods

- Composites
  - Sales Force Composites: Aggregation of sales personnel estimates.
  - Election Polling Composites: sites that aggregate polls.
- Customer Surveys
- Jury of Executive Opinion
- The Delphi Method
  - Individual opinions are compiled and reconsidered. Repeat until overall group consensus is (hopefully) reached.
- We will return to subjective forecasting methods at the end of the Week 1 (Last Session).

# How to forecast with past data, *objectively*?

- We can leverage past data to come up with forecasts:
  - Two primary methods: causal models and time series methods

- Causal Models
  - Let $D$ be the demand or future outcome to be predicted and assume that there are $n$ variables (or root causes) that influence the demand.
  - A causal model is one which demand $D$ is formulated as a theoretical function of all those $n$ causes.
  - Causal models are generally intricate and complex, and need advanced tools in addition to domain expertise.
  - In this course, we will focus mainly on time series based models.

# Time Series Methods

- A time series is just collection of past values of the variable being predicted.
- Can be considered as a “naïve” method. Goal is to isolate patterns in past data.
- Past data might have characteristics such as:
  - Trend
  - Seasonality/Cycles
  - Randomness

# Next…

- An Operational Decision Problem

- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

**Session 2**

- Thinking about Trends and Seasonality

- Forecasting for new products
- Fitting distributions