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

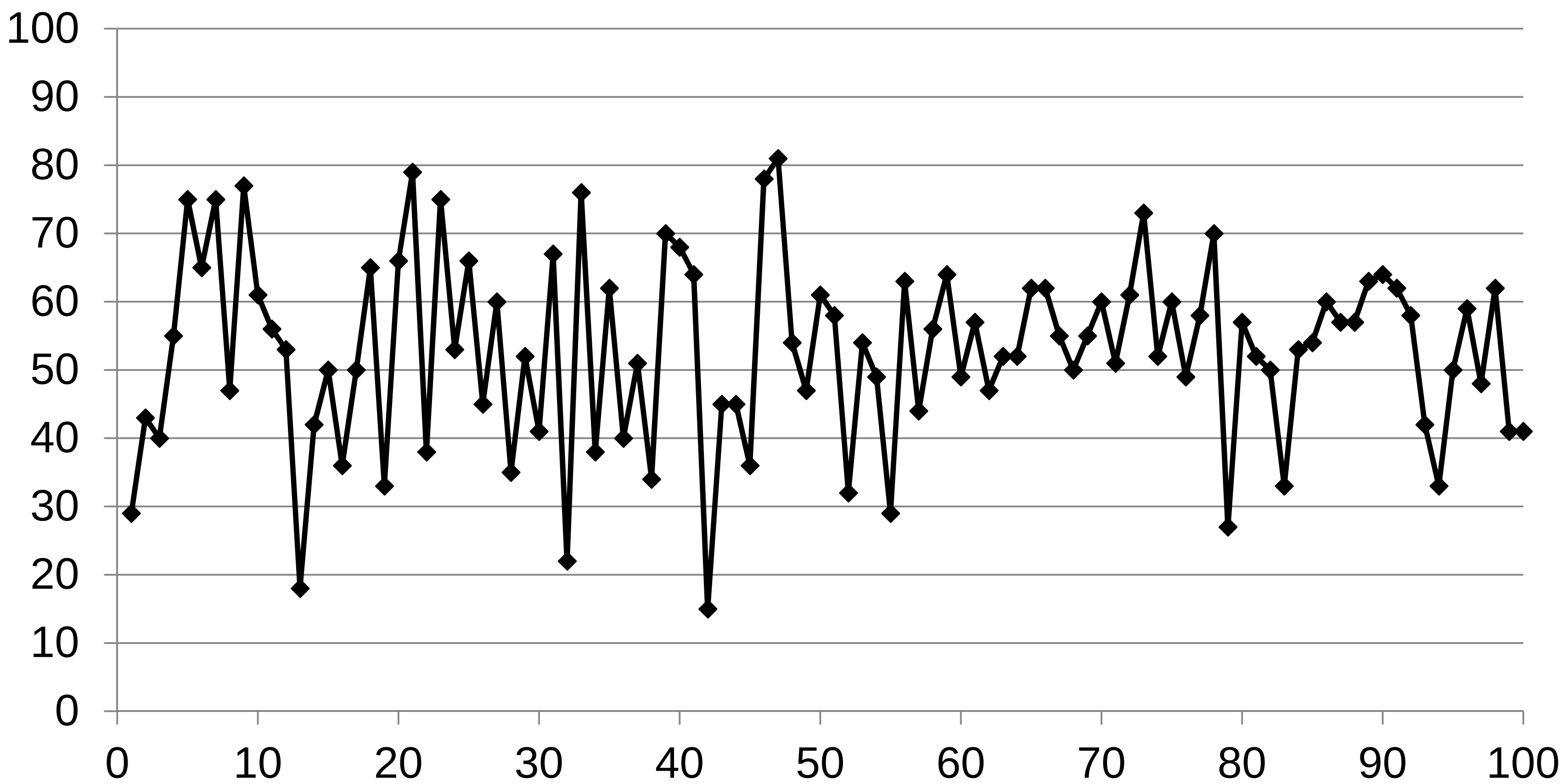
- Forecasting with Past Historical Data
- Moving Averages
- Exponential Smoothing

Session 2

- Thinking about Trends and Seasonality

- Forecasting for new Products
- Fitting distributions

Recall our past demand data



- Let D_t denote demand observed in period t
- From the past data, we have D_1, D_2, . . . D_{100}

Descriptive Statistics: Mean and Standard Deviation

- Sample Mean or Sample Average:
  - Arithmetic average of all the data points $\mu = (D_1 + D_2 + \ldots + D_n)/n$
  - Tells us (roughly) what you expect the next observation to be
  - Can be calculated using excel average() function.
  - Whatever your average is, the demand in future will deviate from the average.

- Sample Standard Deviation
  - A measure of how much noise or variation (from the average) there is in your data.
  - Standard deviation $s = \sqrt{\sum (D_t - \mu)^2 / (n-1)}$, $\mathrm{t} = 1, \ldots, \mathrm{n}$
  - Can be calculated using excel stdev( ) function.

Descriptive Statistics for our Data

- Let past data be $D_1, D_2, \ldots D_{99}, D_{100}$
- Sample Average: 52.81
- Sample Standard deviation: 13.73

- In the excel file *DemandData.xlsx*, I show how to calculate these two for our data on an excel sheet.

- These are two descriptive statistics of our data.

- Note 1: If our data were normally distributed, these two statistics would be sufficient to describe the demand.

- Predictive Statistics: We need to adjust our sample standard deviation for forecasting purposes. More on this later.

Notation: Look Ahead Forecasts

- Recall that $D_1, D_2, \ldots D_t$ are the past values (demands observed).
  - When we are making a forecast in period t, we have demands up to period t.
- We call $F_{t,\, t+\tau}$ = forecast made in period t for demand in future period $t + \tau$ where $\tau = 1, 2, 3, \ldots$
  - E.g. $F_{t,\, t+3}$ is the forecast made in period t for 3 periods ahead.
  - E.g. $F_{100,100+3}$ is the forecast made in period 100, for the period 103. This is called a 3-step forecast. Why?
  - Because, we are looking and forecasting for three periods ahead.

One-Step Forecast

- Typically, we are interested in the next outcome, or simply one-step ahead forecasts.
  - $F_{t,\, t+1}$ is the forecast made at *t* for period *t+1*
  - $F_{t-1,\, t}$ is the forecast made at *t-1* for the next period *t*

- We will use the shorthand notation F_{t+1} for one-step forecast made at t for period t+1.

  - Simply, F_{t+1} stands for $F_{t,\, t+1}$

Forecasting for Stationary Series

- Stationary data shows no trend behavior.
  - Roughly speaking the future resembles the past.
  - Example: Our Newsvendor Demand data (DemandData.xls) which shows no perceptible trend.

- A stationary time series has the form:

  $D_t = \mu + \varepsilon_t$ where μ is a constant and ε_t is a random variable with mean 0 and some standard deviation σ.

- We use past data for forecasting.

- Two common methods for forecasting stationary series are moving averages and exponential smoothing (advanced material slides).

Moving Averages

- A moving averages forecast is the arithmetic average of the n most recent observations.
- We will denote the Moving Averages method that uses n data points as
  - MA (n)
- For a one-step-ahead forecast for period t:
  - $F_t = (D_{t-1} + D_{t-2} + \ldots + D_{t-n})/n$
- For moving averages, a multi-step forecast is the same as the one-step forecast.

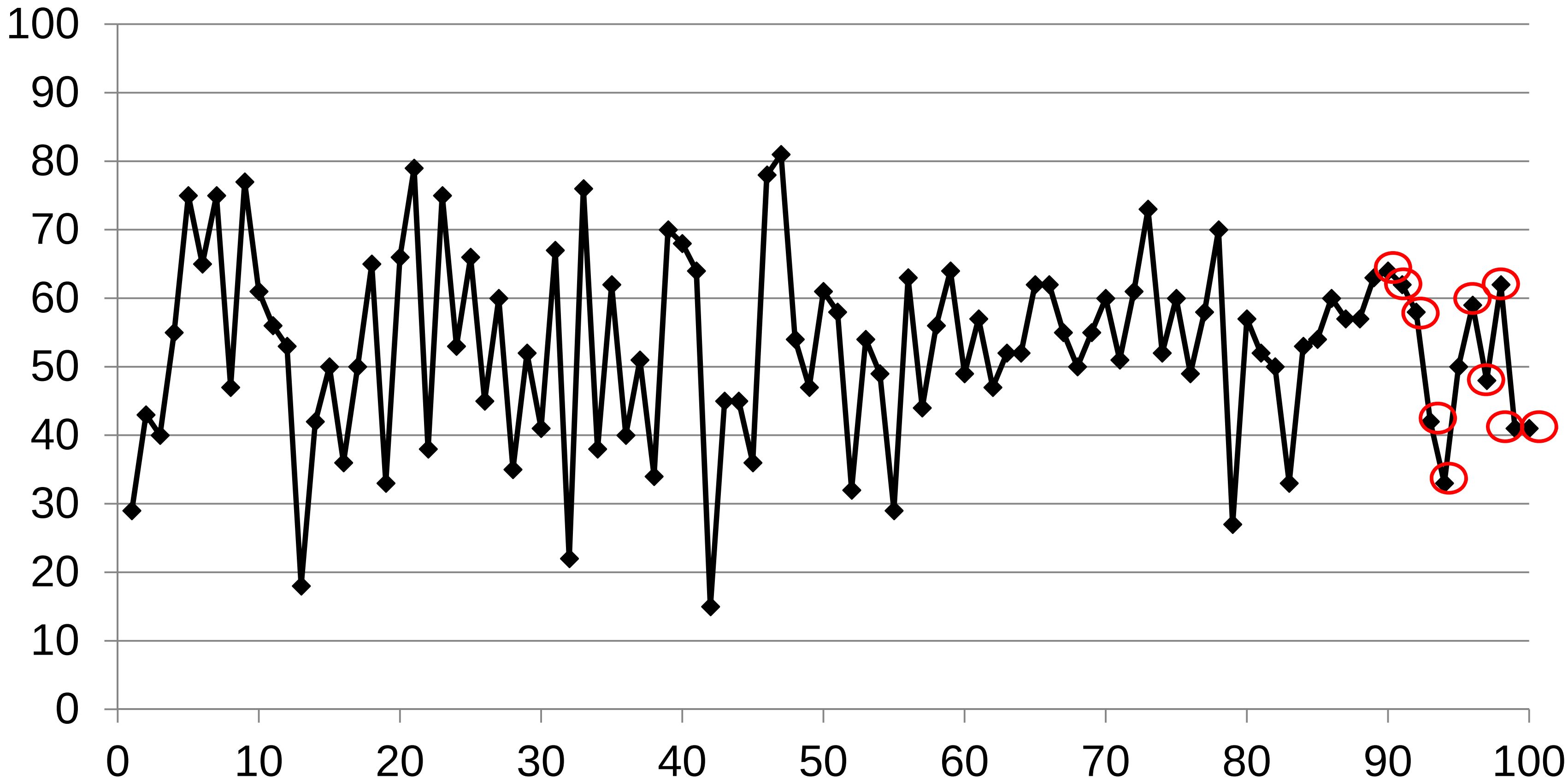
Moving Averages

- One-step-ahead forecast for period t:
  - $F_t = (D_{t-1} + D_{t-2} + \ldots + D_{t-(n-1)} + D_{t-n}) / n$
- It is called "moving" average because the chosen data points "move" and are always the most recent n data points.
- Forecasting for period t+1,
  - $F_{t+1} = (D_t + D_{t-1} + \ldots + D_{t-(n-1)}) / n$
  - We have moved the data used by 1 period (D_t is added and D_{t-n} is discarded).

Moving Averages Example 1:

- Let’s calculate the Moving Average for past 10 periods
- I simply call this MA(10).
- First we look at Descriptive Statistics.
- Then, I’ll show you Predictive Statistics which you can use for Forecasting. (You will see more of the predictive analysis in Week 4).

Last 10 Data points.



- Last 10 data points are marked in red.

Example 1: Descriptive Statistics

- Sample Mean: Average of data from last 10 points.

$$\mu = 49.60$$

- Sample Standard Deviation:

$$s = 10.28$$

- Again, I'll show you how to calculate the descriptive statistics using the template Week1MATemplate.xlsx

- If our data were normally distributed (like a bell curve) the above statistics - mean and standard deviation – would be sufficient to describe the demand distribution.

Predictive Statistics

- We can use the statistics calculated for prediction or forecasting.

- Mean for Prediction = Descriptive Sample mean
  - In other words, descriptive sample mean is an unbiased estimator for mean of the true demand distribution, and hence can be used for prediction.

- However, the standard deviation for prediction needs to be adjusted because of insufficient data.

- When the data is normally distributed,

  Standard Deviation for Prediction = $s + s/\sqrt{n}$

  Recall that s is the descriptive standard deviation that we calculated and n is the total number of data points used for calculation.

Predictive Statistics (Example 1 continued..)

- When we use n=10 data points:
- When the data is Normally distributed,
- Mean for Prediction $= \mu = 49.60$
- Standard Deviation for Prediction,

$$\sigma = s + \frac{s}{\sqrt{n}} = 10.28 + \frac{10.28}{\sqrt{10}} = 13.53$$

Example 1 (continued): Descriptive Statistics

- Suppose we use n = 20 data points. MA (20)
- Sample Mean: Average of data from last 20 points.

$$\mu = 51.95$$

- Sample Standard Deviation:

$$s = 9.62$$

- We have sufficient descriptive statistics for normally distributed data.

Predictive Statistics (Example 1 continued..)

- We have n = 20 data points.
- When the data is Normally distributed,
- Mean for Prediction $= \mu = 51.95$
  - We expect that the demand in the period would be 51.95 on average.
- Standard Deviation for Prediction,

$$\sigma = s + \frac{s}{\sqrt{n}} = 9.62 + \frac{9.62}{\sqrt{20}} = 11.77$$

  - For forecasting, we can assume that the actual demand will deviate from the average with the above standard deviation.

Using More Data

- As we have more data for forecasting
  - Descriptive statistics approach predictive statistics.
  - As we have more data, we gain more confidence for prediction.
  - Note as the number of data points increases, the descriptive standard deviation approaches standard deviation for prediction.
  - $\sigma = s + \frac{s}{\sqrt{n}}$

Statistics for our entire data

- Using all data available, i.e. n = 100

- Descriptive statistics:

$$\mu = 52.81$$
$$s = 13.73$$

- Predictive Statistics (for a normal distribution)

$$\mu = 52.81$$

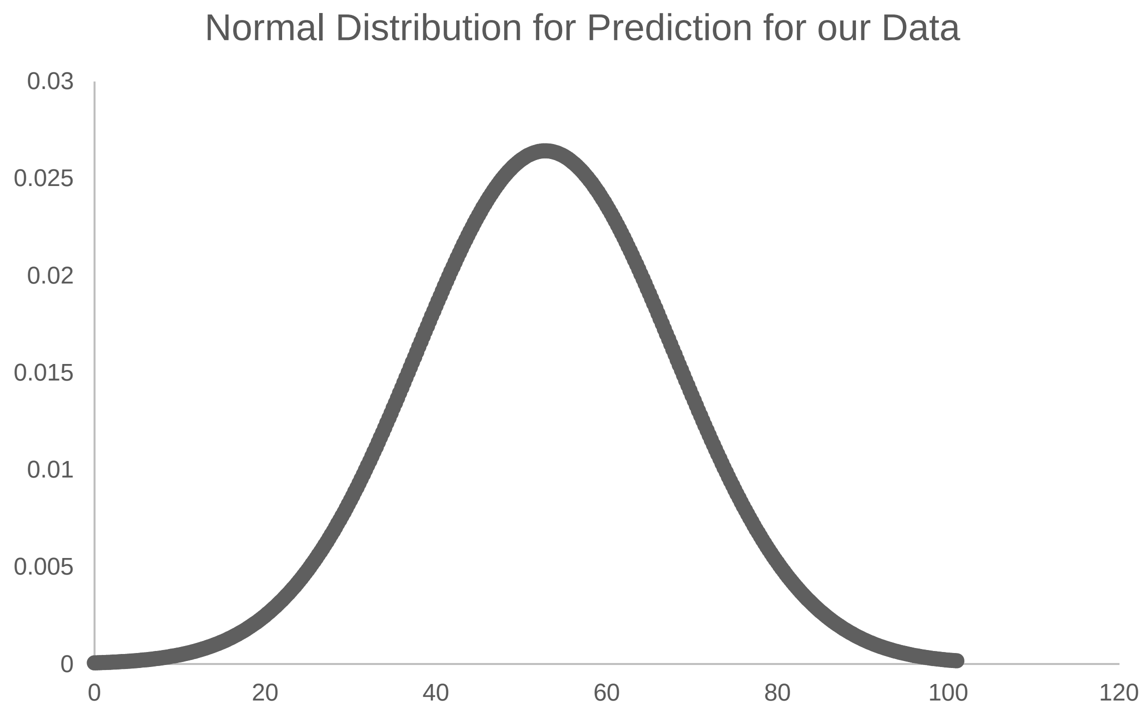
$$\sigma = s + \frac{s}{\sqrt{n}} = 13.73 + \frac{13.73}{\sqrt{100}} = 15.10$$

- We will use the above mean as 52.81 and standard deviation of 15.10 for prediction in the coming weeks.

- All Solutions documented in the excel file Week1MASolution.xlsx

Data Visualization

- We can generate a Normal Distribution graph to visualize how our demand data is distributed…
- Using mean as 52.81 and standard deviation of 15.10 for prediction



Moving Averages: A Discussion

- Advantages of Moving Average Method
  - Easy to understand.
  - Easy to compute.
  - Provides stable forecasts.

- Disadvantages of Moving Average Method
  - Lags behind a trend (as we will see in Session 3).
  - It is not a causal model, i.e., it won't explain *why* realizations in the future behave in a certain way.

- Note that Moving Average method "drops" all data older than the *n* data points you use.
  - How do you think about how to choose *n*?

Moving Averages: What data to use?

- If you choose to use moving average method of last 10 data points,
  - all the older data is ignored
    - (e.g. data from 12 periods back is not used at all).
  - all the recent 10 data points are weighed the same.
    - (e.g. yesterday's data has the same weight as the data from a week before).
- You may want to give more weight to more recent data and less weight to older data.
- Exponential smoothing is based on this precise idea.
  - Advanced slides.

Evaluation of Forecasts

- The forecast error in period t is denoted by e_t,
  - The difference between the forecast for demand in period t and the actual value of demand realized in t.
  - For one step ahead forecast: $e_t = F_t - D_t$.

- Three ways to measure errors:
  - Mean Absolute Deviation $MAD = (1/n) \Sigma | e_t |$
  - Mean Squared Error $MSE = (1/n) \Sigma e_t^2$
  - Mean Absolute Percentage Error $MAPE = (1/n) \Sigma | e_t/D_t| \times 100$.

- Lower the errors, better the forecasting process is.
- Biases in Forecasts:
  - A bias occurs when the *average* value of a forecast error tends to be positive or negative.

Measuring Errors: An Example

- Using our dataset again, we will go through different methods of calculating errors in our forecasts.
- Assume we have data up to 80 periods..
- We use Moving Averages of 10 periods to calculate our forecasts.
- Once we have forecast for period 81, demand for 81 occurs.
- We then forecast for 82, and demand for period 82 occurs and so on..

Errors Example

- This allows us to calculate the errors from using MA(10)
  - By comparing demands and forecasts over periods 81 through 100.

- Using, Week1ErrorsTemplate.xlsx, I'll walk through an example exercise.

- For MA(10) moving averages of n=10, we get:
  - Mean Absolute Deviation $MAD = (1/n)\, \Sigma\, | e_i |$ = 8.9
  - Mean Squared Error $MSE = (1/n)\, \Sigma\, e_i^{\,2}$ =113.15
  - Mean Absolute Percentage Error $MAPE = (1/n)\, \Sigma\, | e_i/D_i| \times 100.$ =19.72%

- For MA(20) moving averages of n=20, we get:
  - Mean Absolute Deviation $MAD = (1/n)\, \Sigma\, | e_i |$ = 7.66
  - Mean Squared Error $MSE = (1/n)\, \Sigma\, e_i^{\,2}$ =92.61
  - Mean Absolute Percentage Error $MAPE = (1/n)\, \Sigma\, | e_i/D_i| \times 100.$ =17.29%

Errors Example: Continued…

- The solution is available in Week1ErrorsSolution.xlsx
- It may be preferable to use MA(20) over MA(10) in this case
  - By comparing corresponding error terms in the data
- Measuring errors allows us to understand better the choice of which method to use.
- Finally, in the data there is no evidence of any bias.

Wrapping up

- We saw how to forecast using the Moving Averages method
- We saw how to measure errors and biases in Forecasting.
- We learned about Descriptive Statistics…
- … and how to adapt the Descriptive Statistics for Prediction.

See you in the next Session.

Next…

- An Operational Decision Problem

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

Session 3

- Thinking about Trends and Seasonality

- Forecasting for new products
- Fitting distributions