FORECASTING IN ECONOMICS & BUSINESS ECOM90024

LECTURE 1: INTRODUCTION & REVIEW

1

LECTURE OUTLINE

- Welcome!
- Introduction and motivation
- Course details
- Review of statistics and probability
- Defining a time series

2

CONTACT INFORMATION

LECTURER

Jonathan Thong
Room 354, FBE Building
Office Hours: Tuesdays 16:00 – 17:00 (Week 3 onwards)
jonathan.thong@unimelb.edu.au

TUTOR

Heidi Quah
heidi.quah@unimelb.edu.au

3

COURSE INTRODUCTION

- The practice of ***forecasting***, at its core, involves the utilisation of available information (i.e. *data*) to form predictions about the future.

- In this course, we will build statistical models which will allow us to use historical data to learn about the future in a rigorous and logical way. Specifically:
  - We will use ***time series analysis*** to build models of historical data
  - We will use these models to ***construct forecasts*** into the future
  - We will see how to ***evaluate the performance*** of our forecasts
  - We will learn how to ***refine our models*** to improve their performance

4

MOTIVATION (i.e., WHY SHOULD I TAKE THIS COURSE?)

- Forecasting and time series analysis can be applied in many different settings:
  - Economic variables
  - Financial indicators
  - Marketing and advertising metrics
  - Public opinion and politics
  - Firm revenues and industry conditions
  - Behavioral indicators
  - Natural phenomena
- While we will be focusing mainly on economic and financial data in this course, the techniques we will be using can be applied to many other contexts!

5

MOTIVATION (i.e., WHY SHOULD I TAKE THIS COURSE?)

- Forecasts matter and can have deep and wide-ranging implications!


Bloomberg
Asia Edition
● Live Now Markets Economics Industries Technology Politics Wealth Pursuits Opinion Businessweek Equality

Economics
Indicators

US Producer Prices Exceed Forecast in Biggest Gain Since June

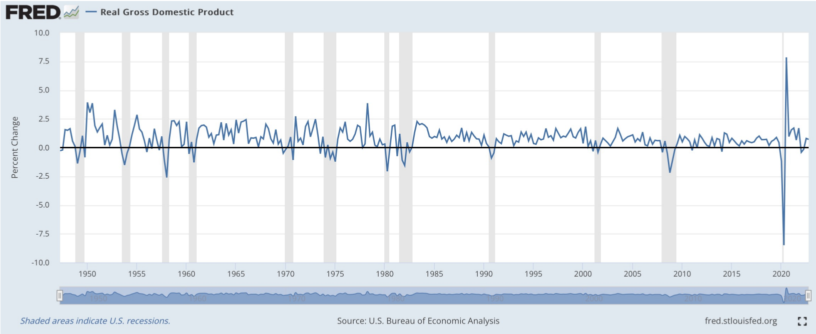
- Index of wholesale costs rose 0.7% in January on monthly basis
- Goods prices climbed 1.2%, while measure of services rose 0.4%

- Developing a fluency in time series analysis and forecasting techniques will be extremely useful for future courses and any future applied research that you may wish to undertake.

6

MOTIVATION (i.e., WHY SHOULD I TAKE THIS COURSE?)

- It is a super interesting and challenging time for forecasters! Consider the following plot of quarterly real GDP growth for the United States. How should we construct forecasts of the next few quarters?



When the world changes, how can we use the past to forecast or make predictions about the future?

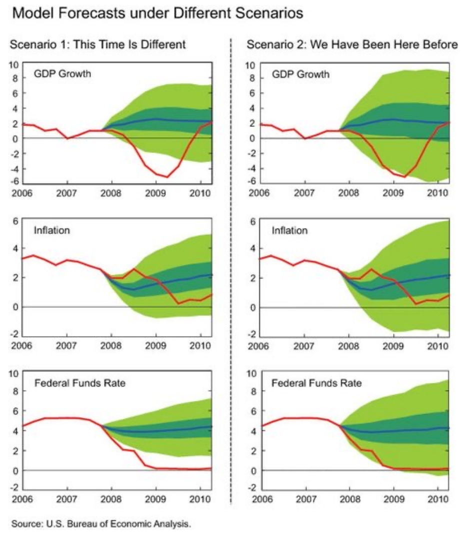
Will things will be the same when the crisis passes?

These are very interesting and very hard questions!

7

MOTIVATION (i.e., WHY SHOULD I TAKE THIS COURSE?)

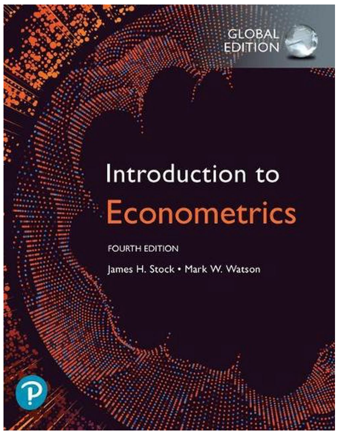
- Time series and forecasting models give us a way of thinking about and depicting uncertainty.
- A good model should not only provide point forecasts! We need to consider the range of possible values that could likely occur.
- Forecasting models allow us to connect our assumptions about the underlying data generating process to the predictions that we make.
- Read: https://libertystreeteconomics.newyorkfed.org/2012/05/the-great-moderation-forecast-uncertainty-and-the-great-recession.html



8

ASSUMED KNOWLEDGE & STATISTICAL BACKGROUND

- You will be assumed to have encountered concepts such as random variables, probability distributions, expected values, estimation and hypothesis testing and simple/multiple linear regression, among other topics in your prior undergraduate work.
- We will review some of these topics briefly and discuss their application to time series models and forecasting over the next couple of lectures.
- A good reference for these topics should you need to review them is Stock, J. and Watson, M. (2015), Introduction to Econometrics (4th Ed) – Chapters 2-7. You can access an e-book version of this text via the university library website.
- There will be a moderate amount of mathematical manipulation and notation throughout the course, including some rudimentary matrix algebra. The focus of the course, however, will be practical.



9

OTHER USEFUL REFERENCES

- There is no required text for this course beyond the lecture notes and additional materials provided on the subject LMS on Canvas.
- The following texts may be useful as supplementary or further reading:
  - Stock, J. and Watson, M. (2015), ***Introduction to Econometrics***, 4th Edition, Pearson.
  - Elliot E. and Timmermann A. (2016), ***Economic Forecasting***, Princeton University.
  - Tsay, Ruey. (2010), ***Analysis of Financial Time Series***, 3rd Edition, Wiley.
  - Tsay, Ruey. (2013), ***An introduction to analysis of financial data with R***, Wiley. (ebook available through university library page)
- **Only the content covered in lectures, tutorials and assignments will be examinable.**

10

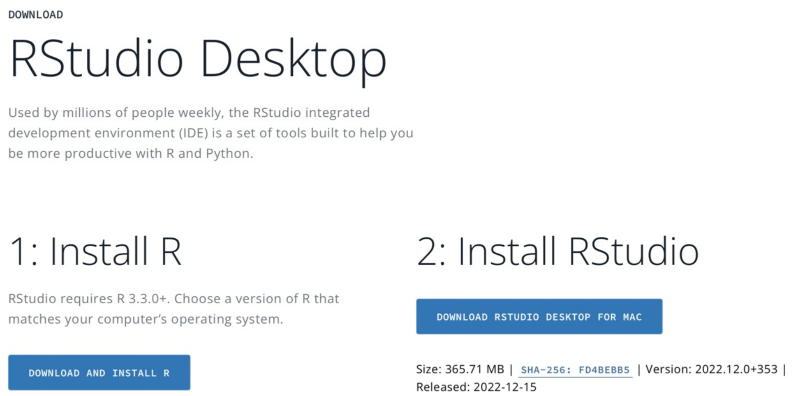
COMPUTING

- We will be using the ***RStudio*** statistical environment to perform all of our computations and analysis for this course. ***Rstudio*** is a front end for ***R***, an open source, high level programming language.
- Some good introductory guides to ***Rstudio*** and ***R*** include:
  - https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf
  - http://www.r-tutor.com/r-introduction
  - https://colinfay.me/intro-to-r/
- I will be devoting time in each lecture to using ***R*** to illustrate the concepts.
- Fluency in ***R*** is a valuable skill! It is something worthwhile developing!

11

COMPUTING

- You will have to go to https://posit.co/download/rstudio-desktop/ to download and install R and Rstudio (these are two distinct items!):



12

LECTURE TOPICS

WEEK STARTING	WEEK	TOPIC
Feb 26	1	Introduction & Review of Linear Regression
March 4	2	Multiple Regression & Simple Trend Models
March 11	3	Decomposing Time Series: Smoothing & Seasonality
March 18	4	Covariance Stationary Time Series
March 25	5	Autoregressions and Moving Averages
April 1	6	**EASTER BREAK - NO CLASS**
April 8	7	The General Linear Model -ARMA
April 15	8	Estimating & Forecasting ARMA Models
April 22	9	Stochastic Trends & Unit Root Tests
April 29	10	Volatility Models: ARCH
May 6	11	Volatility Models: GARCH
May 13	12	Realized Volatility
May 20	13	Review & Exam Practice

13

ASSESSMENT

- Your assessment for this subject comprises of the following tasks:

TASK	DUE DATE	WEIGHTING
Assignment 1	Friday of Week 5 at 5:00PM	10%
Assignment 2	Friday of Week 9 at 5:00PM	10%
Assignment 3	Friday of Week 13 at 5:00PM	10%
Final Exam	Exam Period	70%

- The final exam is a ***hurdle requirement*** for the subject. You will need to pass the exam (i.e. obtain a score of 50% and above) in order to pass the subject.

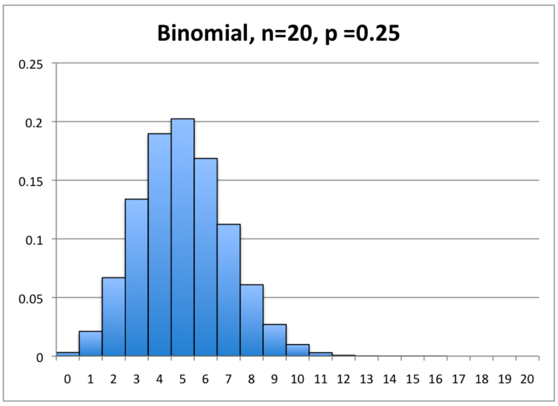
14

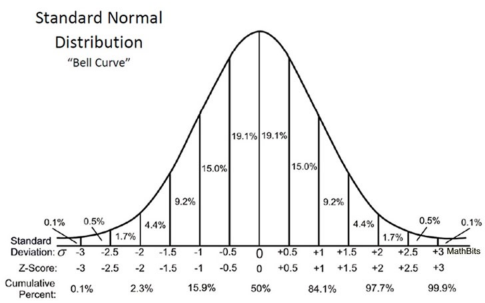
REVIEW OF STATISTICS & PROBABILITY

- A ***random variable***, X, is defined as a set of numerical outcomes that are generated by an underlying random process. Attached to each numerical outcome is a quantity called a ***probability*** that represents its likelihood of occurrence.
- Random variables can either be ***discrete*** or ***continuous***.
- A discrete random variable is completely characterized by its probability distribution function.
- A continuous random variable is completely characterized by its probability density function.

15

REVIEW OF STATISTICS & PROBABILITY





16

REVIEW OF STATISTICS & PROBABILITY

- The ***mean*** and ***variance*** of a random variable describe the locus and spread of the probability mass across its set of possible values and can be calculated using the expectations operator.
- The mean, μ, is calculated as:

$$\mu = E[X] = \sum_{i=1}^{k} x_i P(x_i) = \sum_{x} xP(x)$$

For a ***discrete random*** variable X that can take one of k possible values.

$$\mu = E[X] = \int x\, f(x)$$

For a ***continuous random*** variable X that can take one of an infinite number of possible values along a continuum.

17

REVIEW OF STATISTICS & PROBABILITY

- The variance, σ^2 is calculated as:

$$\sigma^2 = E[(X-\mu)^2] = \sum_{i=1}^{k} (x_i - \mu)^2 P(x_i)$$

$$\sigma^2 = E[(X-\mu)^2] = \int (x-\mu)^2 f(x)$$

- Sometimes we will also use the **variance operator $\mathbf{V}ar$( )** in our derivations

$$Var(X) = E[(X-\mu)^2] = E[X^2] - \mu^2$$

18

REVIEW OF STATISTICS & PROBABILITY

- The expectations operator is a linear operator. Therefore, for constants a and b (i.e., non-random quantities), we have that

$$E[aX+b] = aE[X] + b$$

- The variance operator on the other hand, for constants a and b:

$$Var(aX) = \ a^2 Var(X)$$

$$Var(X+b) = Var(X)$$

- If random variables X and Y are independent (i.e. $P(X,Y) = P(X)P(Y)$, or $f(X,Y) = f(X)f(Y)$), then:

$$E[XY] = E[X]E[Y]$$

$$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y)$$

19

REVIEW OF STATISTICS & PROBABILITY

- If random variables X and Y are ***NOT*** independent:

$$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X,Y)$$

$$Var(aX-bY) = a^2 Var(X) + b^2 Var(Y) - 2abCov(X,Y)$$

- Where the covariance of X and Y is given by

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$$

- Where for constants a and b:

$$Cov(a,X) = 0$$

$$Cov(aX,bY) = abCov(X,Y)$$

$$Cov(X+a,Y+b) = Cov(X,Y)$$

20

REVIEW OF STATISTICS & PROBABILITY

- Also recall that the correlation coefficient ρ (also known as Pearson's correlation coefficient) is a rescaling of the covariance into the [-1,1] interval.

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

- Note that if X and Y are independent, then $\rho_{X,Y} = 0$. However, $\rho_{X,Y} = 0$ **DOES NOT** imply independence!

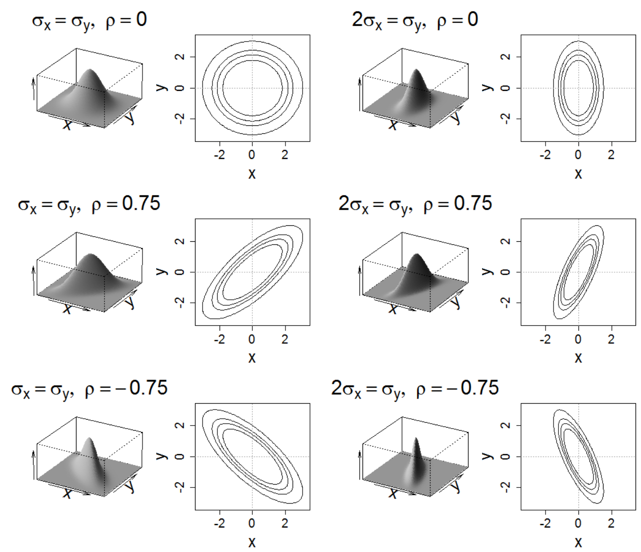
- Only in the special case when X and Y are ***jointly normal*** does $\rho_{X,Y} = 0$ imply independence.

21

REVIEW OF STATISTICS & PROBABILITY

- The mean and variance tell us something about the shape characteristics of the marginal distribution (density) functions of random variables.

- The covariance and correlation tell us something about the shape characteristics of the joint distribution (density) function of random variables.



22

REVIEW OF STATISTICS & PROBABILITY

- To illustrate using a discrete random variable, consider a fair coin that is tossed three times. The number of heads, x obtained can be 0, 1, 2 or 3. The probabilities of each of these possibilities can be tabulated as shown:

x	0	1	2	3
$P(x)$	$^1/_8$	$^3/_8$	$^3/_8$	$^1/_8$

- Then,

$$E[X] = 0(^1/_8) + 1\,(^3/_8) + 2(^3/_8) + 3(^1/_8) = {}^{12}/_8 = 1.5$$

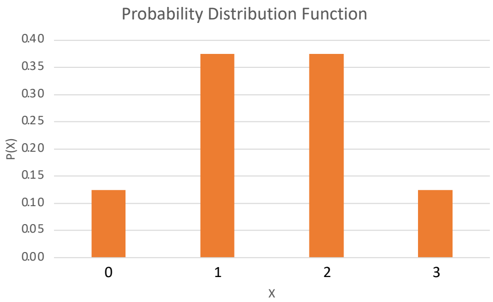
$$E[(X-\mu)^2] = E[X^2] - E[X]^2 = 0(^1/_8) + 1\,(^3/_8) + 4(^3/_8) + 9(^1/_8) - 1.5^2$$

$$= 0.75$$

23

REVIEW OF STATISTICS & PROBABILITY

- Visually,

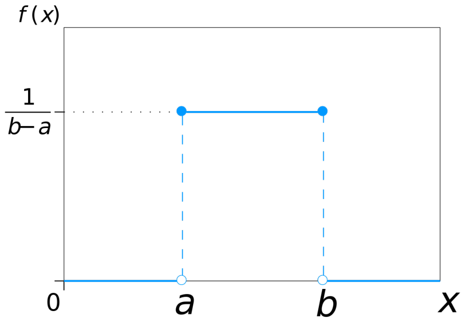


- Also note that this is a binomial random variable with parameter values $n = 3$ and $p = 0.5$

24

REVIEW OF STATISTICS & PROBABILITY

- To illustrate using a continuous random variable, consider a uniform random variable, defined over the interval $[a, b]$. The density function is given by



25

REVIEW OF STATISTICS & PROBABILITY

- We can derive the mean of the uniform random variable in the following way,

$$E[X] = \int_a^b x\left(\frac{1}{b-a}\right) d(x)$$

- The anti-derivative is given by,

$$E[X] = \frac{x^2}{2}\left(\frac{1}{b-a}\right)\Bigg|_a^b = \frac{b^2 - a^2}{2}\left(\frac{1}{b-a}\right) = \frac{b+a}{2}$$

26

REVIEW OF STATISTICS & PROBABILITY

- Again, we use the fact that

$$\sigma^2 = E[(X-\mu)^2] = E[X^2] - E[X]^2$$

- Where

$$E[X^2] = \int_a^b x^2\left(\frac{1}{b-a}\right)d(x) = \frac{x^3}{3}\left(\frac{1}{b-a}\right)\Bigg|_a^b = \frac{b^3-a^3}{3(b-a)}$$

- So that

$$E[X^2] - E[X]^2 = \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}$$

27

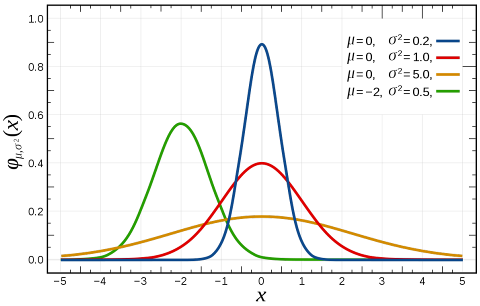
REVIEW OF STATISTICS & PROBABILITY

- We will be working with the normal random variable throughout the course. Its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- The shape of the normal is determined entirely by its parameters, μ and σ^2

- A normal random variable will always have a skewness of zero (i.e. it is symmetric around its mean) and a kurtosis of 3 (or an excess kurtosis of 0).



28

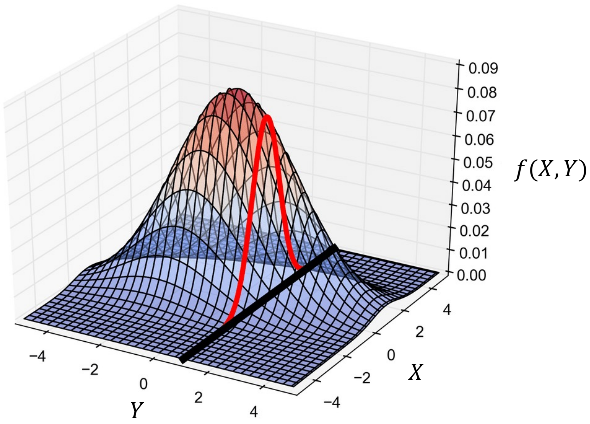
REVIEW OF STATISTICS & PROBABILITY

- We will also be working a lot with conditional expectations. Recall that the conditional probability of $X = x$ given that $Y = y$ is given by

$$P(X = x|Y = y) = P_{X|Y}(x|y)$$
$$= \frac{P(X = x, Y = y)}{P(Y = y)}$$

- When working with density functions, we have that

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$



29

REVIEW OF STATISTICS & PROBABILITY

- It follows that the conditional expectation of X given Y is given by

$$E[X|Y = y] = \sum_{i=1}^{k} x_i P_{X|Y}(x_i|y)$$

$$E[X|Y = y] = \int x\, f_{X|Y}(x|y)$$

- We can always recover the unconditional expectation from the conditional expectation by using the **Law of Iterated Expectations** (**LIE**)

$$E[X] = E_Y\big[E[X|Y]\big] = E\big[E[X|Y]\big]$$

- Where E_Y denotes the expectation taken with respect to Y.

30

REVIEW OF STATISTICS AND PROBABILITY

- To illustrate, let X and Y be random variables where Y is a discrete random variable that can take one of m possible values $\{y_1, \ldots, y_m\}$. Then we have that

$$E_Y[E[X|Y]] = E[X|Y = y_1]P(y_1) + \cdots + E[X|Y = y_m]P(y_m) = E[X]$$

- According to the Law of Iterated Expectations, this weighted average of conditional expectations is equal to the unconditional expectation of X. This will hold for both discrete and continuous variables.

- Proof of the discrete case will be a tutorial exercise!

31

REVIEW OF STATISTICS & PROBABILITY

- It is extremely important that you understand the structure of statistical inference.

- For instance, given a sequence of independent and identically distributed random variables, $\{X_1, X_2, \ldots, X_k\}$ with mean μ and variance σ^2 we have,

Parameter	Estimator	Estimate
μ	$X = \frac{1}{n}\sum_{i=1}^{n} X_i$	$x = \frac{1}{n}\sum_{i=1}^{n} x_i$
σ^2	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X)^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)^2$

32

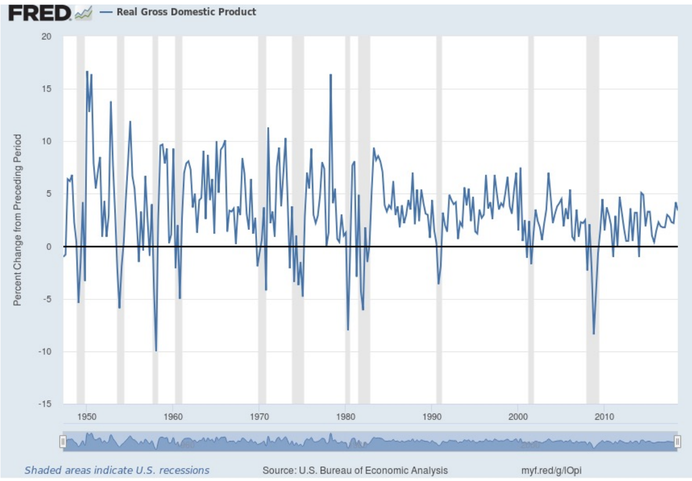
REVIEW OF STATISTICS & PROBABILITY

- It is important that you are able to do the following before progressing further into the course:
  - Compute and interpret ***population moments*** and their ***sample analogues*** (e.g., mean, median, standard deviation, variance, correlation coefficients, etc.).
  - Define and distinguish between a ***parameter***, ***estimator*** and ***estimate***.
  - Conducting and interpreting ***hypothesis tests***.
  - Understand and compute ***joint***, ***marginal*** and ***conditional probabilities***.
  - Compute and interpret ***conditional*** and ***unconditional expectations***.
  - Forming and interpreting ***confidence intervals***.
  - Reading and using ***common distributional tables*** (e.g. z, t, χ^2, F)

33

DEFINING A TIME SERIES MODEL

- A ***time series*** is a sequence of data points that have been observed through repeated measurements over time, typically at fixed intervals.
- A ***time series model*** specifies the random process that generates the time series, otherwise known as the ***data generating process***.



34

DEFINING A TIME SERIES MODEL: AN EXAMPLE

- Consider the following random process:

$$\varepsilon_t \sim_{iid} N(0,1)$$

- The subscript t is known as a ***time index*** and typically takes values $t: \{1,2,3,4, \ldots\}$.

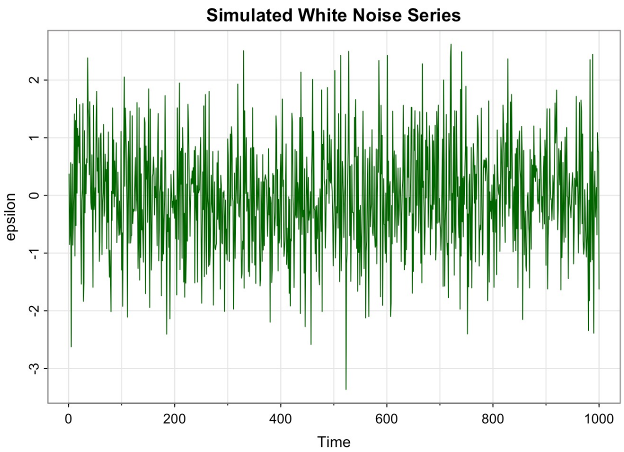
- This particular random process is known as ***white noise*** since the above notation states that every observation is ***independently*** and ***identically*** drawn from a normal random variable with a mean of zero and a variance of one.

35

DEFINING A TIME SERIES MODEL: AN EXAMPLE

- We can use ***R*** to generate such a series (see R Markdown file):

- If the data generating process of a variable of interest was white noise, would we be able to compute good forecasts using the data?



36

DEFINING A TIME SERIES MODEL: AN EXAMPLE

- Now consider the following random process:

$$X_t = 0.95X_{t-1} + \varepsilon_t$$

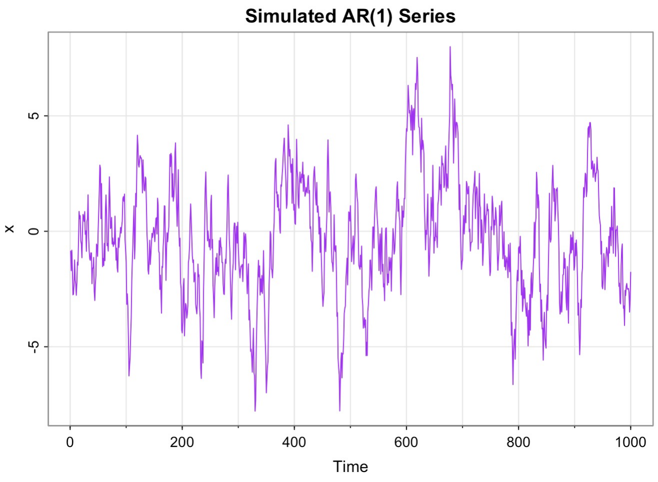
$$\varepsilon_t \sim_{iid} N(0,1)$$

- This is known as an autoregressive model of order 1, or $AR(1)$.
- According to this specification, the current realization of the random variable is the sum of a fraction of the realization in the previous period, plus a white noise error term.
- For such a data generating process, past observations contain information about future observations.

37

DEFINING A TIME SERIES MODEL: AN EXAMPLE

- Visually,



38

DEFINING A TIME SERIES MODEL: AN EXAMPLE

- In practice, quantities such as the autoregressive ***parameter*** and the variance of the innovations are unobservable and thus unknown. In general, we would write an $AR(1)$ model as:

$$x_t = \beta x_{t-1} + \varepsilon_t$$

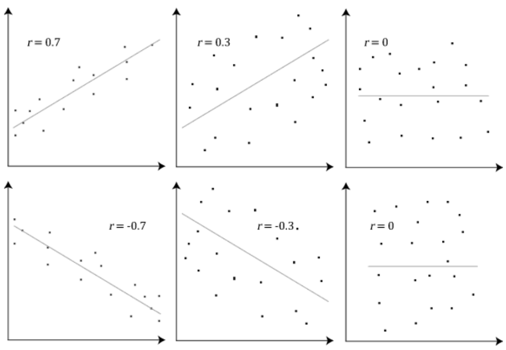
$$\varepsilon_t \sim_{iid} N(0, \sigma^2)$$

- We will use regression analysis as a primary means to estimating these parameters!

39

A REVIEW OF SIMPLE LINEAR REGRESSION

- Recall that the coefficient of correlation r measures the degree to which two variables are linearly related.



40

A REVIEW OF SIMPLE LINEAR REGRESSION

- If two variables (x, y) are perfectly linearly related (i.e. $r \pm 1$), then we can represent their relationship using a linear function.

$$y_i = a + bx_i$$

- More generally, we can determine the ***degree*** to which two variables are linearly related by finding a line of best fit. To do this, we specify:

$$y_i = a + bx_i + e_i$$

- Intuitively, the greater the degree to which (x, y) are linearly related, the smaller the magnitude of each e_i!

41

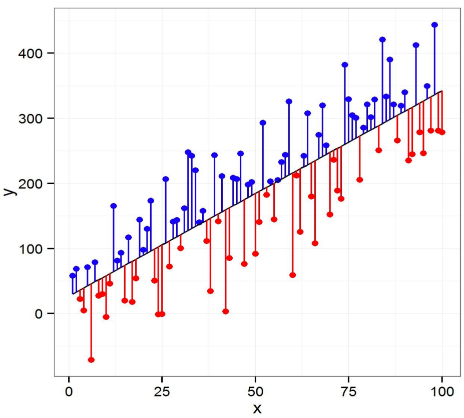
A REVIEW OF SIMPLE LINEAR REGRESSION

- Fitting a line to a bivariate data set $\{x_i, y_i\}_{i=1}^{n}$ involves choosing values a and b so that the fitted line is given by

$$\widehat{y_i} = a + bx_i$$

- Computing a line of best fit involves choosing values a and b such that the sum of squared differences are minimized.

$$min \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 = min \sum_{i=1}^{n} e_i^2$$



42

A REVIEW OF SIMPLE LINEAR REGRESSION

- Given a set of n observations on variables x and y, we obtain a line of best fit by:

1. First, computing $\bar{x}$ and $\bar{y}$
2. Then, computing the sample covariance of (x, y) and the sample variance of x
3. Finally, computing,

$$b = \frac{cov(x,y)}{var(x)}$$

$$a = \bar{y} - b\bar{x}$$

These are the solutions to the minimization problem that we stated in the previous slide! We can verify this using calculus!

43

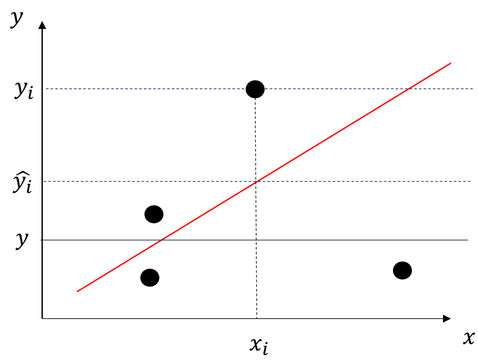
A REVIEW OF SIMPLE LINEAR REGRESSION

- Simple linear regression allows us to decompose the total variation of the dependent variable into two distinct components. First, let's show that,

$y_i = a + bx_i + e_i$ (observed value)

$\hat{y}_i = a + bx_i$ (fitted value)

$\bar{y} = a + b\bar{x} + \bar{e}$ (average value)



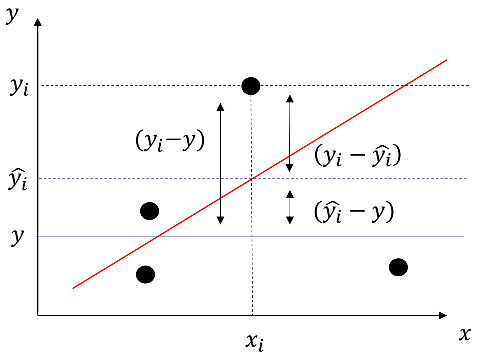
44

A REVIEW OF SIMPLE LINEAR REGRESSION

- The difference between y_i and $\bar{y}$ (i.e. the total variation in y) can be decomposed into two parts:

$$y_i - \bar{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})$$

- We can see that $|y_i - \hat{y}_i|$ will be small, the greater the degree to which x and y are linearly related.



45

A REVIEW OF SIMPLE LINEAR REGRESSION

- To apply this idea over all observations, squaring each $(y_i - \bar{y})$ and summing over n we obtain:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$

$\boldsymbol{SST} = Total\ sum\ of\ squares$

$\boldsymbol{SSE} = Sum\ of\ squares\ of\ erros\ (i.e. the\ amount\ of\ variation\ that\ is\ unexplained\ by\ the\ explanatory\ variable\ x)$

$\boldsymbol{SSR} = Sum\ of\ squares\ for\ regression\ (i.e. the\ amount\ of\ variation\ that\ is\ explained\ by\ the\ explanatory\ variable\ x)$

46

A REVIEW OF SIMPLE LINEAR REGRESSION

- The decomposition allows us to define what is known as the ***coefficient of determination***, more commonly known as R^2

$$R^2 = \frac{Explained\ variation\ in\ y}{Total\ variation\ in\ y} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

- Since R^2 is a proportion of the total variation, it must lie in the range: $0 \leq R^2 \leq 1$

47

A REVIEW OF SIMPLE LINEAR REGRESSION

- Now that we know how to draw a line of best fit, we can formally define a simple regression model. Given two random variables X and Y, we say that a regression of Y on X is defined as,

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

- This is known as the ***population regression function*** as it specifies a relationship between random variables.
- The quantities α and β are ***parameters*** which have to be estimated.
- Given a set of observations (x_i, y_i), drawing a line of best fit produces estimates (a, b) of the regression parameters (α, β).

48

A REVIEW OF SIMPLE LINEAR REGRESSION

- Recall that given a sequence of independent and identically distributed random variables, $\{X_1, X_2, \ldots, X_k\}$ with mean μ and variance σ^2

Parameter	Estimator	Estimate
μ	$X = \frac{1}{k}\sum_{i=1}^{k} X_i$	$x = \frac{1}{k}\sum_{i=1}^{k} x_i$
σ^2	$S^2 = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - X)^2$	$s^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_i - x)^2$

- Remember, ***estimators*** are functions of random variables and are thus random variables themselves. They are rules that tell us how to compute estimates when we have observations.

49

A REVIEW OF SIMPLE LINEAR REGRESSION

- Similarly, in the context of a simple linear regression, given n independent and identically distributed draws $\{X_i, Y_i\}_{i=1}^{n}$ and the population regression function

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Parameter	Estimator	Estimate
α	$\hat{\alpha} = Y - \hat{\beta}X$	$a = y - bx$
β	$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - X)(Y_i - Y)}{\sum_{i=1}^{n}(X_i - X)^2}$	$b = \frac{\sum_{i=1}^{n}(x_i - x)(y_i - y)}{\sum_{i=1}^{n}(x_i - x)^2}$

50

A REVIEW OF SIMPLE LINEAR REGRESSION

- Recall that the probability distribution of an estimator is called a ***sampling distribution***. It allows us to perform ***inference*** on the unknown population parameters.

- For instance, we know that the ***central limit theorem*** tells us that given a sequence of $i.i.d$ random variables $\{X_1, X_2, \dots, X_n\}$ with mean μ and variance σ^2, the sample mean,

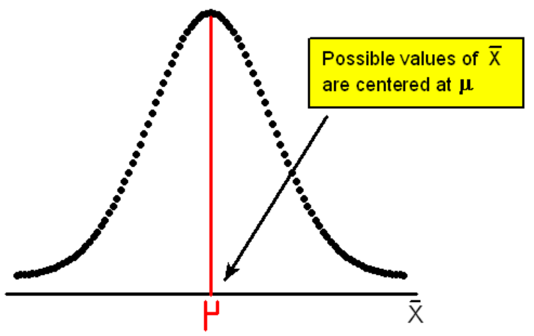
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Will be distributed as a ***normal random variable*** with a mean μ and a variance of σ^2/n.

51

A REVIEW OF SIMPLE LINEAR REGGRESSION

- The sampling distribution of the sample mean allows us to construct confidence intervals and perform hypothesis tests.



52

A REVIEW OF SIMPLE LINEAR REGRESSION

- Given n independent and identically distributed pairs $\{X_i, Y_i\}_{i=1}^{n}$ and the population regression function

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

- If we make the additional assumption that

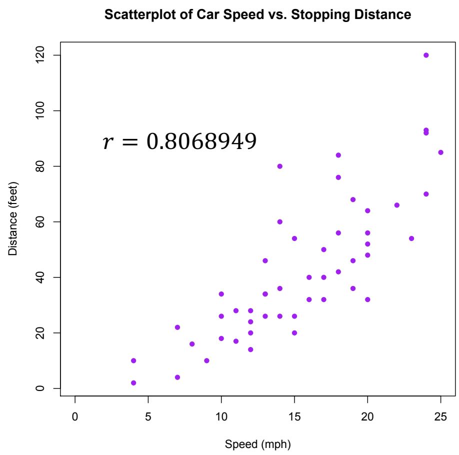
$$E\left[\varepsilon_i \middle| X_i\right] = 0$$

- Then the estimators $\hat{\alpha}$ and $\hat{\beta}$ will be jointly normally distributed, $\hat{\alpha} \sim N\left(\alpha, \sigma_{\hat{\alpha}}^2\right)$, $\hat{\beta} \sim N\left(\beta, \sigma_{\hat{\beta}}^2\right)$

53

A REVIEW OF SIMPLE LINEAR REGRESSION

- To illustrate, let's consider the following dataset:



54

A REVIEW OF SIMPLE LINEAR REGRESSION

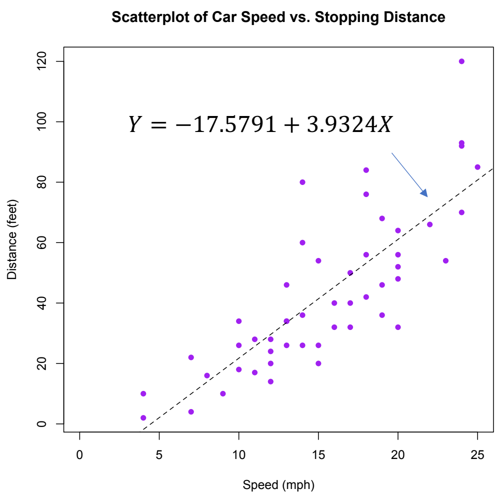
- Let's now compute a linear regression in which we set the stopping distance as our dependent variable and car speed as our explanatory variable. We use the lm() function in ***R*** to compute our estimates and associated statistics.

```
Call:
lm(formula = dist ~ speed, data = cardata)

Residuals:
    Min      1Q  Median      3Q     Max 
-29.069  -9.525  -2.272   9.215  43.201 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -17.5791     6.7584  -2.601   0.0123 *  
speed         3.9324     0.4155   9.464 1.49e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.38 on 48 degrees of freedom
Multiple R-squared:  0.6511,	Adjusted R-squared:  0.6438 
F-statistic: 89.57 on 1 and 48 DF,  p-value: 1.49e-12
```

55

A REVIEW OF SIMPLE LINEAR REGRESSION

- Visually,



56

NEXT WEEK

- We will briefly discuss multiple regression and then jump into building some simple trend models!

57