



Risk and crisis forecast, Analysis and Reduction Methods and Tools

Dr. Abdessamad AIT EL CADI

Associate Professor @ UPHF

CRENG-VALNCIENNES-February 17th, 2022



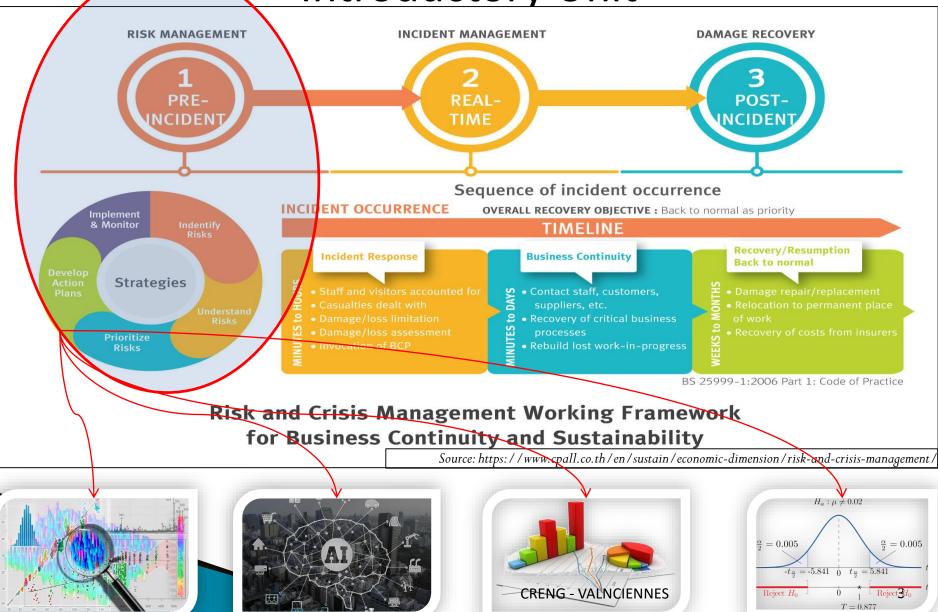


Module structure

- An Introductory unit: Risk Management & Crisis Management / definitions / motivations / issues and perspectives
- Probability & Statistics for Crisis & Risk Management
- Forecasting technics / Forecasting Process
- Aided decision
- Computing tools for Crisis & Risk Management
- Analytics / Data mining / BI / Big Data
- Risk Management / Risk Analysis
- Crisis Theory & Management



Introductory Unit





Introduction

The Ten Most Costly Catastrophes, United States

			Insured loss (\$ millions)		
Rank	Date	Peril	Dollars when occurred	in 2002 dollars ¹	
1.	Sep. 11, 2001	World Trade Center, Pentagon terrorist attacks ²	\$20,300.0	\$20,620.9 (estimate)	
2.	Aug 1992	Hurricane Andrew	15,500.0	19,874.9	
3.	Jan. 1994	Northridge, California earthquake	12,500.0	15,173.8	
4.	Sep. 1989	Hurricane Hugo	4,195.0	6,086.1	
5.	Sep. 1998	Hurricane Georges	2,900.0	3,200.7	
6.	Jun. 2001	Tropical Storm Alison	2,500.0	2,539.5	
7.	Oct. 1995	Hurricane Opal	2,100.0	2,478.9	
8.	Sep. 1999	Hurricane Floyd	1,960.0	2,116.5	
9.	Mar. 1993	20-state winter storm	1,750.0	2,178.7	
10.	Oct. 1991	Oakland, California fire	1,700.0	2,245.4	

¹Adjusted to 2002 dollars by the Insurance Information Institute. ²Property coverage only.

NOTE: Data are from Insurance Services Office, Inc. (ISO) and Insurance Information Institute.



Meaning of Risk

• Origin of the word:

Latin: Resicum, Risicul, Riscus = fate, chance, hazard Grec Rhizikon, Rhiza = reef Arabe: رزق (Rizq) = luck, fortune "all that has been given to you, and from which you make profit"

- During the Renaissance:
 - Synonymous with adventure, luck, danger, fortune, chance, danger
 - Positive connotation: one seeks to seize one's chance by taking risks
 - Negative connotation: one is exposed to a danger by facing the risk
- Nowadays, risk has evolved into a negative connotation
 - Possible danger more or less predictable
 - Possibility/probability of a fact/event considered as an evil/damage



Meaning of Risk

- <u>Risk</u>: Uncertainty concerning the occurrence of a loss
- Objective Risk vs. Subjective Risk
 - <u>Objective risk</u> is defined as the relative variation of actual loss from expected loss
 - It can be statistically calculated using a measure of dispersion, such as the standard deviation
 - declines as the number of exposure units increases
 - <u>Subjective risk</u> is defined as uncertainty based on a person's mental condition or state of mind
 - Two persons in the same situation may have different perceptions of risk
 - High subjective risk often results in conservative behavior



Chance of Loss

- <u>Chance of loss</u>: The probability that an event will occur
- Objective Probability vs. Subjective Probability
 - <u>Objective probability</u> refers to the long-run relative frequency of an event assuming an infinite number of observations and no change in the underlying conditions
 - It can be determined by deductive or inductive reasoning
 - <u>Subjective probability</u> is the individual's personal estimate of the chance of loss
 - A person's perception of the chance of loss may differ from the objective probability



Chance of Loss vs. Objective Risk

- Chance of loss is the probability that an event that causes a loss will occur.
- Objective risk is the relative variation of actual loss from expected loss

The chance of loss may be identical for two different groups, but objective risk may be quite different!

City	# homes	Average # fires	μ±σ	Chance of Fire	Objective Risk (σ/μ)
Philadelphia	10,000	100	75 – 125	1%	25%
Los Angeles	10,000	100	90 - 110	1%	10%



Peril and Hazard

- A <u>peril</u> is defined as the cause of the loss
 - In an auto accident, the collision is the peril
- A <u>hazard</u> is a condition that increases the chance of loss
 - <u>Physical hazards</u> are physical conditions that increase the chance of loss (icy roads, defective wiring)
 - <u>Moral hazard</u> is dishonesty or character defects in an individual, that increase the chance of loss (faking accidents, inflating claim amounts)
 - <u>Morale Hazard</u> is carelessness or indifference to a loss because of the existence of insurance (leaving keys in an unlocked car)
 - <u>Legal Hazard</u> refers to characteristics of the legal system or regulatory environment that increase the chance of loss (large damage awards in liability lawsuits)

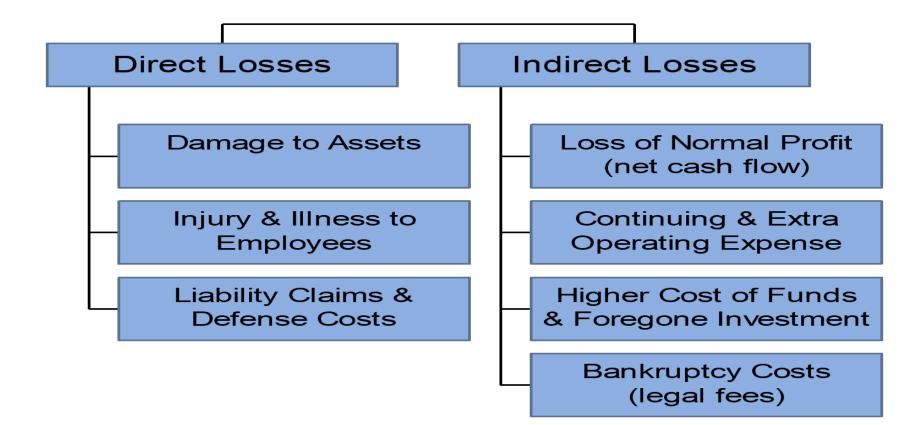


Basic Categories of Risk

- Pure and Speculative Risk
 - A <u>pure risk</u> is one in which there are only the possibilities of loss or no loss (earthquake)
 - A <u>speculative risk</u> is one in which both profit or loss are possible (gambling)
- Fundamental and Particular Risk
 - A <u>fundamental risk</u> affects the entire economy or large numbers of persons or groups (hurricane)
 - A <u>particular risk</u> affects only the individual (car theft)

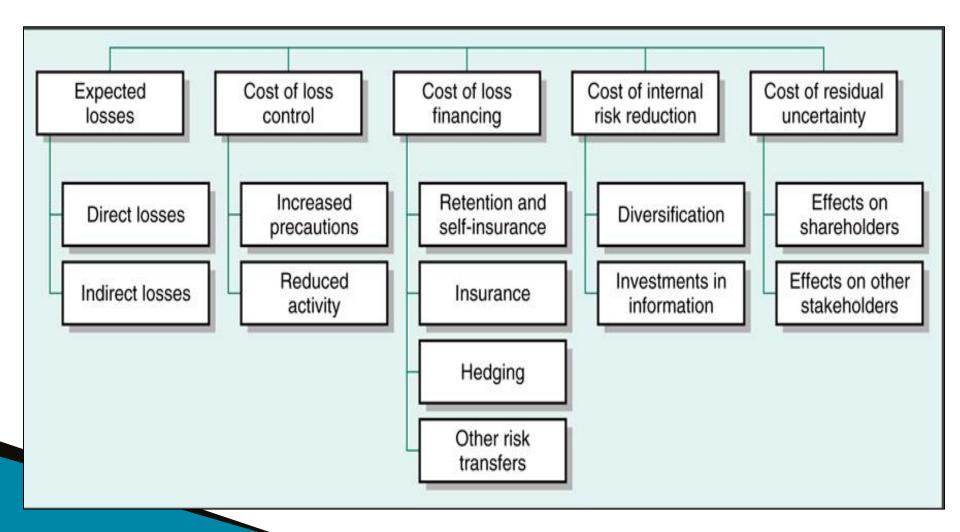


Direct loss vs. indirect loss





Components of the Cost of Risk





Burden of Risk on Society

- The presence of risk results in three major burdens on society:
 - In the absence of insurance, individuals would have to maintain large emergency funds
 - Put aside \$25,000 annually for damages (due to fire, windstorm, theft etc.) to your home valued at \$200,000?
 - The risk of a liability lawsuit may discourage innovation, depriving society of certain goods and services
 - Some 250 companies once manufactured childhood vaccines, now just a few firms do.
 - Risk causes worry and fear



Methods of Handling Risk

- Avoidance
- Loss control
 - <u>Loss prevention</u> refers to activities to reduce the frequency of losses
 - <u>Loss reduction</u> refers to activities to reduce the severity of losses
- Retention
 - An individual or firm retains all or part of a loss
 - Loss retention may be active or passive
- Noninsurance transfers
 - A risk may be transferred to another party through contracts, hedging, or incorporation
 - Insurance



CASE APPLICATION Mike's Methods of Handling Risks

- Mike is a college senior majoring in BBA.
- He owns a high-mileage 2000 Ford with current market valu \$2,500.
- The current replacement value of his personal property in a rented apartment totals \$10,000.
- He wears disposable contact lenses, which cost \$300 for a six-month supply.
- He also has a waterbed in his rented apartment that has leaked water in the past.
- As avid runner, Mike run five miles daily in a nearby public park that has reputation of being extremely dangerous due to drug dealers.
- Mike's parents both work to help him pay for his tuition.





CASE APPLICATION Mike's Methods of Handling Risks



- For each of the following risks and loss exposures, identify an appropriate RM technique that could have been used to deal with exposure.
 - Physical damage to the 2000 Ford due to a collision with another motorist.
 - Liability lawsuit against Mike arising out of the negligent operation of his car.
 - Total loss of personal property due to a fire in the kitchen of his rented apartment.
 - Disappearance of one contact lens.
 - Waterbed leak that cause property damage to the apartment.
 - Physical assault on Mike by gang members who are dealing drugs in the park where he runs.
 - Loss of tuition assistance from Mike's father who is killed by a drunk driver in an auto accident.



Meaning of Risk Management

- Risk Management is a process that identifies loss exposures faced by an organization and selects the most appropriate techniques for treating such exposures
- A <u>loss exposure</u> is any situation or circumstance in which a loss is possible, regardless of whether a loss occurs
 - E.g., a plant that may be damaged by an earthquake, or an automobile that may be damaged in a collision

RISK Management: Increased likelihood of meeting planned goals



Objectives of Risk Management

- Risk management has objectives before and after a loss occurs
- Pre-loss objectives:
 - Prepare for potential losses in the most economical way
 - Reduce anxiety
 - Meet any legal obligations



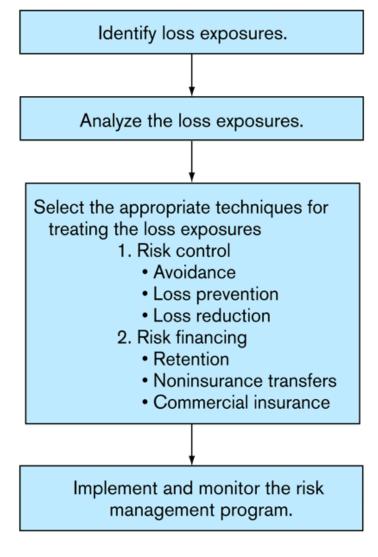
Objectives of Risk Management

- Post-loss objectives:
 - Ensure survival of the firm
 - Continue operations
 - Stabilize earnings
 - Maintain growth
 - Minimize the effects that a loss will have on other persons and on society



Risk Management Process

- Identify potential losses
- Evaluate potential losses
- Select the appropriate risk management technique
- Implement and monitor the risk management program





Identifying Loss Exposures

- Property loss exposures
- Liability loss exposures
- Business income loss exposures
- Human resources loss exposures
- Crime loss exposures
- Employee benefit loss exposures
- Foreign loss exposures
- Market reputation and public image of company
- Failure to comply with government rules and regulations



Identifying Loss Exposures

- Risk Managers have several sources of information to identify loss exposures:
 - Questionnaires
 - Physical inspection
 - Flowcharts
 - Financial statements
 - Historical loss data
- Industry trends and market changes can create new loss exposures.
 - e.g., exposure to acts of terrorism



Analyzing Loss Exposures

- Estimate the frequency and severity of loss for each type of loss exposure
 - <u>Loss frequency</u> refers to the probable number of losses that may occur during some given time period
 - <u>Loss severity</u> refers to the probable size of the losses that may occur
- Once loss exposures are analyzed, they can be ranked according to their relative importance
- Loss severity is more important than loss frequency:
 - The <u>maximum possible loss</u> is the worst loss that could happen to the firm during its lifetime
 - The <u>maximum probable loss</u> is the worst loss that is *likely* to happen

Select the Appropriate Risk Management Technique

- <u>Risk control</u> refers to techniques that reduce the frequency and severity of losses
- Methods of risk control include:
 - Avoidance
 - Loss prevention
 - Loss reduction



Risk Control Methods

- <u>Avoidance</u> means a certain loss exposure is never acquired, or an existing loss exposure is abandoned
 - The chance of loss is reduced to zero
 - It is not always possible, or practical, to avoid all losses



Risk Control Methods

- <u>Loss prevention</u> refers to measures that reduce the frequency of a particular loss
 - e.g., installing safety features on hazardous products
- <u>Loss reduction</u> refers to measures that reduce the severity of a loss after is occurs
 - e.g., installing an automatic sprinkler system

Select the Appropriate Risk Management Technique

- <u>Risk financing</u> refers to techniques that provide for the funding of losses
- Methods of risk financing include:
 - Retention
 - Non-insurance Transfers
 - Commercial Insurance



- <u>Retention</u> means that the firm retains part or all of the losses that can result from a given loss
 - Retention is effectively used when:
 - No other method of treatment is available
 - The worst possible loss is not serious
 - Losses are highly predictable
 - The retention level is the dollar amount of losses that the firm will retain
 - A financially strong firm can have a higher retention level than a financially weak firm
 - The maximum retention may be calculated as a percentage of the firm's net working capital



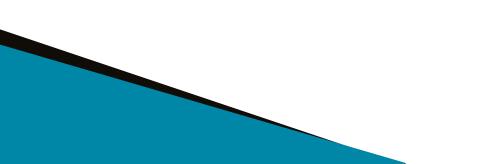
- A risk manager has several methods for paying retained losses:
 - Current net income: losses are treated as current expenses
 - Unfunded reserve: losses are deducted from a bookkeeping account
 - Funded reserve: losses are deducted from a liquid fund
 - Credit line: funds are borrowed to pay losses as they occur



- A <u>captive insurer</u> is an insurer owned by a parent firm for the purpose of insuring the parent firm's loss exposures
 - A <u>single-parent captive</u> is owned by only one parent
 - An <u>association or group captive</u> is an insurer owned by several parents
 - Many captives are located in the Caribbean because the regulatory environment is favorable
 - Captives are formed for several reasons, including:
 - The parent firm may have difficulty obtaining insurance
 - Costs may be lower than purchasing commercial insurance
 - A captive insurer has easier access to a reinsurer
 - A captive insurer can become a source of profit
 - Premiums paid to a captive may be tax-deductible under certain conditions



- <u>Self-insurance</u> is a special form of planned retention
 - Part or all of a given loss exposure is retained by the firm
 - A more accurate term would be self-funding
 - Widely used for workers compensation and group health benefits
- A <u>risk retention group</u> is a group captive that can write any type of liability coverage except employer liability, workers compensation, and personal lines
 - Federal regulation allows employers, trade groups, governmental units, and other parties to form risk retention groups
 - They are exempt from many state insurance laws



Advantages

Disadvantages

- Save money
- Lower expenses
- Encourage loss prevention
- Increase cash flow

- Possible higher losses
- Possible higher expenses
- Possible higher taxes



Risk Financing Methods: Non-insurance Transfers

- <u>A non-insurance transfer</u> is a method other than insurance by which a pure risk and its potential financial consequences are transferred to another party
 - Examples include:
 - Contracts, leases, hold-harmless agreements

Risk Financing Methods: Non-insurance Transfers

Advantages

Disadvantages

- Can transfer some losses that are not insurable
- Save money
- Can transfer loss to someone who is in a better position to control losses

- Contract language may be ambiguous, so transfer may fail
- If the other party fails to pay, firm is still responsible for the loss
- Insurers may not give credit for transfers



Risk Financing Methods: Insurance

- Insurance is appropriate for loss exposures that have a low probability of loss but for which the severity of loss is high
 - The risk manager selects the coverages needed, and policy provisions:
 - A <u>deductible</u> is a provision by which a specified amount is subtracted from the loss payment otherwise payable to the insured
 - An <u>excess insurance policy</u> is one in which the insurer does not participate in the loss until the actual loss exceeds the amount a firm has decided to retain
 - The risk manager selects the insurer, or insurers, to provide the coverages



Risk Financing Methods: Insurance

- The risk manager negotiates the terms of the insurance contract
 - A <u>manuscript policy</u> is a policy specially tailored for the firm

 Language in the policy must be clear to both parties
 - The parties must agree on the contract provisions, endorsements, forms, and premiums
- The risk manager must periodically review the insurance program



Risk Financing Methods: Insurance

Advantages

- Firm is indemnified for losses
- Uncertainty is reduced
- Insurers may provide other risk management services
- Premiums are tax-deductible

Disadvantages

- Premiums may be costly
 - Opportunity cost should be considered
- Negotiation of contracts takes time and effort
- The risk manager may become lax in exercising loss control



Risk Management Matrix

Түре of Loss	Loss Frequency	Loss Severity	Appropriate Risk Management Technique
1	Low	Low	Retention
2	High	Low	Loss prevention and retention
3	Low	High	Insurance
4	High	High	Avoidance

Implement and Monitor the Risk Management Program

- Implementation of a risk management program begins with a <u>risk</u> <u>management policy statement</u> that:
 - Outlines the firm's risk management objectives
 - Outlines the firm's policy on loss control
 - Educates top-level executives in regard to the risk management process
 - Gives the risk manager greater authority
 - Provides standards for judging the risk manager's performance
- A <u>risk management manual</u> may be used to:
 - Describe the risk management program
 - Train new employees

Implement and Monitor the Risk Management Program

- A successful risk management program requires active cooperation from other departments in the firm
- The risk management program should be periodically reviewed and evaluated to determine whether the objectives are being attained
 - The risk manager should compare the costs and benefits of all risk management activities



Benefits of Risk Management

- Pre-loss and post-loss objectives are attainable
- A risk management program can reduce a firm's <u>cost of risk</u>
 - The cost of risk includes premiums paid, retained losses, outside risk management services, financial guarantees, internal administrative costs, taxes, fees, and other expenses
- Reduction in pure loss exposures allows a firm to enact an enterprise risk management program to treat both pure and speculative loss exposures
- Society benefits because both direct and indirect losses are reduced



Probability and Statistics



UNIT: Probability and Statistics

"I only believe in statistics that I doctored myself"

- Overview:
 - Populations and samples
 - Descriptive statistics (graphical plots and numerical summaries)
 - Inferential statistics (confidence intervals and hypothesis tests)
 - Determining the relationship between two or more variables (regression analysis)
- Basics of Probability
 - Sample space and events
 - Axioms and properties of probability
 - Conditional probability
 - Independent events

- Random Variables
 - Definition and distribution function
 - Discrete random variables / Probability mass function (Bernoulli, binomial, geometric, and Poisson distributions and their applications)
 - Continuous random variables and distributions / Probability density function / Cumulative density function (Normal, exponential, gamma, and Weibull distributions and their applications)
 - Characteristics of a random variable (mean, median, variance, standard deviation)



UNIT: Probability and Statistics

- Joint Probability Distributions
 - Jointly distributed random variables
 - Marginal distributions
 - Independent random variables
 - Covariance and correlation
 - Statistics and their distributions
 - Distribution of the sample mean and the central limit theorem
- Point Estimation
 - Unbiased estimator
 - Variance of a point estimator
 - Estimators for the mean and variance
- Descriptive Statistics
 - Graphical plots (histogram, box plot, scatter plot)
 - Numerical summaries (sample mean, sample variance)

- Confidence Intervals
 - For large sample sizes
 - For normally distributed data
 - Intervals for means and proportions
- Hypothesis Tests Based on a Single Sample
 - Hypotheses and test procedures
 - Type I error, type II error, and power
 - *P*-values
 - Tests for means and proportions
- Inferences Based on Two Samples
 - Hypothesis tests and confidence intervals
 - Comparing two means / two proportions
- Regression Analysis
 - Linear regression models
 - Estimating model parameters
 - Determining the adequacy of the model

CRENG - VALNCIENNES



Random Variables

- X is a random variable if it represents a random draw from some population
- a <u>discrete</u> random variable can take on only selected values
- a <u>continuous</u> random variable can take on any value in a real interval
- associated with each random variable is a probability distribution



Random Variables – Examples

- the outcome of a coin toss a discrete random variable with P(Heads)=.5 and P(Tails)=.5
- the height of a selected student a continuous random variable drawn from an approximately normal distribution

Expected Value of X - E(X)

- The expected value is really just a probability weighted average of X
- E(X) is the mean of the distribution of X, denoted by μ_x
- Let $f(x_i)$ be the probability that $X=x_i$, then

$$\mu_X = E(X) = \sum_{i=1}^n x_i f(x_i)$$

Variance of X – Var(X)

- The variance of X is a measure of the dispersion of the distribution
- Var(X) is the expected value of the squared deviations from the mean, so

$$\sigma_X^2 = Var(X) = E\left[(X - \mu_X)^2 \right]$$

More on Variance

- The square root of Var(X) is the standard deviation of X
- Var(X) can alternatively be written in terms of a weighted sum of squared deviations, because

$$E[(X - \mu_X)^2] = \sum (x_i - \mu_X)^2 f(x_i)$$

Covariance – Cov(X,Y)

- Covariance between X and Y is a measure of the association between two random variables, X & Y
- If positive, then both move up or down together
- If negative, then if X is high, Y is low, vice versa

$$\sigma_{XY} = Cov(X, Y) = E\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right]$$

Correlation Between X and Y

- Covariance is dependent upon the units of X & Y [Cov(aX,bY)=abCov(X,Y)]
- Correlation, Corr(X,Y), scales covariance by the standard deviations of X & Y so that it lies between 1 & -1

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{Cov(X,Y)}{\left[Var(X)Var(Y)\right]_2^1}$$



More Correlation & Covariance

- If $\sigma_{X,Y} = 0$ (or equivalently $\rho_{X,Y} = 0$) then X and Y are linearly unrelated
- If $\rho_{X,Y} = 1$ then X and Y are said to be perfectly positively correlated
- If $\rho_{X,Y} = -1$ then X and Y are said to be perfectly negatively correlated
- Corr(aX,bY) = Corr(X,Y) if ab>0
- Corr(aX,bY) = -Corr(X,Y) if ab<0



Properties of Expectations

- E(a)=a, Var(a)=0
- $E(\mu_X) = \mu_X$, i.e. E(E(X)) = E(X)
- E(aX+b)=aE(X)+b
- E(X+Y)=E(X)+E(Y)
- E(X-Y)=E(X)-E(Y)
- $E(X \mu_X) = 0 \text{ or } E(X E(X)) = 0$
- $E((aX)^2)=a^2E(X^2)$



More Properties

- $Var(X) = E(X^2) \mu_x^2$
- $Var(aX+b) = a^2Var(X)$
- Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)
- Var(X-Y) = Var(X) + Var(Y) 2Cov(X,Y)
- $Cov(X,Y) = E(XY) \mu_x \mu_y$
- If (and only if) X,Y independent, then

- Var(X+Y)=Var(X)+Var(Y), E(XY)=E(X)E(Y)

The Normal Distribution

- A general normal distribution, with mean μ and variance σ^2 is written as N(μ , σ^2)
- It has the following probability density function (pdf)

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

The Standard Normal

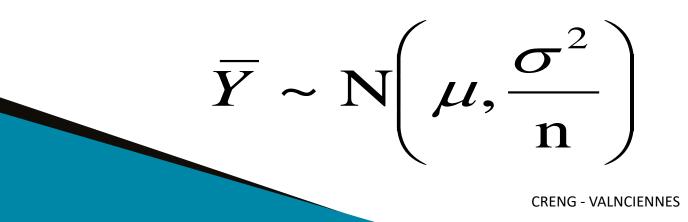
- Any random variable can be "standardized" by subtracting the mean, μ , and dividing by the standard deviation, σ , so E(Z)=0, Var(Z)=1
- Thus, the standard normal, N(0,1), has pdf

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$



Properties of the Normal

- If X~N(μ , σ^2), then aX+b ~N(a μ +b, $a^2\sigma^2$)
- A linear combination of independent, identically distributed (iid) normal random variables will also be normally distributed
- If Y_1, Y_2, \dots, Y_n are iid and $\sim N(\mu, \sigma^2)$, then





Cumulative Distribution Function

- For a pdf, f(x), where f(x) is P(X = x), the cumulative distribution function (cdf), F(x), is $P(X \le x)$; P(X > x) = 1 F(x) = P(X < -x)
- For the standard normal, $\phi(z)$, the cdf is $\Phi(z)$ = P(Z<z), so
- $P(|Z|>a) = 2P(Z>a) = 2[1-\Phi(a)]$
- $P(a \le Z \le b) = \Phi(b) \Phi(a)$

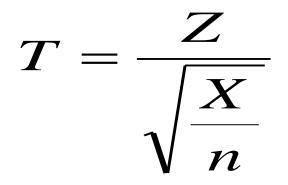


The Chi-Square Distribution

- Suppose that Z_i , i=1,...,n are iid ~ N(0,1), and $X=\sum(Z_i^2)$, then
- X has a chi-square distribution with n degrees of freedom (df), that is
- $X \sim \chi^2_n$
- If $X \sim \chi^2_n$, then E(X)=n and Var(X)=2n

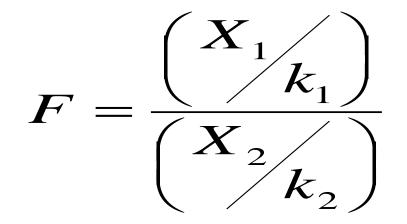
The t distribution

- If a random variable, T, has a t distribution with n degrees of freedom, then it is denoted as $T \sim t_n$
- E(T)=0 (for n>1) and Var(T)=n/(n-2) (for n>2)
- T is a function of Z~N(0,1) and X~ χ^2_n as follows:



The F Distribution

- If a random variable, F, has an F distribution with (k_1,k_2) df, then it is denoted as $F \sim F_{k1,k2}$
- F is a function of $X_1 \sim \chi^2_{k1}$ and $X_2 \sim \chi^2_{k2}$ as follows:





Random Samples and Sampling

- For a random variable Y, repeated draws from the same population can be labeled as Y_1, Y_2, \ldots, Y_n
- If every combination of n sample points has an equal chance of being selected, this is a random sample
- A random sample is a set of independent, identically distributed (i.i.d) random variables

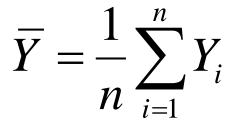


Estimators and Estimates

- Typically, we can't observe the full population, so we must make inferences base on estimates from a random sample
- An estimator is just a mathematical formula for estimating a population parameter from sample data
- An estimate is the actual number the formula produces from the sample data

Examples of Estimators

- Suppose we want to estimate the population mean
- Suppose we use the formula for E(Y), but substitute 1/n for f(y_i) as the probability weight since each point has an equal chance of being included in the sample, then
- Can calculate the sample average for our sample:



CRENG - VALNCIENNES



What Make a Good Estimator?

- Unbiasedness
 - Efficiency
 - Mean Square Error (MSE)
- Asymptotic properties (for large samples):
 Consistency

Unbiasedness of Estimator

- Want your estimator to be right, on average
- We say an estimator, W, of a Population Parameter, θ, is unbiased if E(W)=E(θ)
- For our example, that means we want

$$E(\overline{Y}) = \mu_Y$$



Proof: Sample Mean is Unbiased

$$E(\overline{Y}) = E\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right) = \frac{1}{n}\sum_{i=1}^{n}E(Y_i)$$

$$=\frac{1}{n}\sum_{i=1}^n\mu_Y=\frac{1}{n}n\mu_Y=\mu_Y$$

Efficiency of Estimator

- Want your estimator to be closer to the truth, on average, than any other estimator
- We say an estimator, W, is efficient if Var(W)< Var(any other estimator)
- Note, for our example

$$Var(\overline{Y}) = Var\left(\frac{1}{n}\sum_{i=1}^{n}Y_{i}\right) = \frac{1}{n^{2}}\sum_{i=1}^{n}\sigma^{2} = \frac{\sigma^{2}}{n}$$

MSE of Estimator

- What if can't find an unbiased estimator?
- Define mean square error as $E[(W-\theta)^2]$
- Get trade off between unbiasedness and efficiency, since MSE = variance + bias²
- For our example, that means minimizing

$$E\left[\left(\overline{Y} - \mu_Y\right)^2\right] = Var\left(\overline{Y}\right) + \left(E\left(\overline{Y}\right) - \mu_Y\right)^2$$

Consistency of Estimator

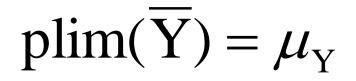
- Asymptotic properties, that is, what happens as the sample size goes to infinity?
- Want distribution of W to converge to θ,
 i.e. plim(W)=θ
- For our example, that means we want

$$P(|\overline{Y} - \mu_Y| > \varepsilon) \to 0 \text{ as } n \to \infty$$



More on Consistency

- An unbiased estimator is not necessarily consistent – suppose choose Y_1 as estimate of μ_Y , since $E(Y_1) = \mu_Y$, then $plim(Y_1) \neq \mu_Y$
- An unbiased estimator, W, is consistent if Var(W) $\rightarrow 0$ as $n \rightarrow \infty$
- Law of Large Numbers refers to the consistency of sample average as estimator for μ , that is, to the fact that:



Central Limit Theorem

- Asymptotic Normality implies that $P(Z < z) \rightarrow \Phi(z)$ as $n \rightarrow \infty$, or $P(Z < z) \approx \Phi(z)$
- The central limit theorem states that the standardized average of any population with mean μ and variance σ^2 is asymptotically ~N(0,1), or

$$Z = \frac{\overline{Y} - \mu_Y}{\sigma / \sqrt{n}} \sim N(0,1)$$

Estimate of Population Variance

- We have a good estimate of μ_{Y} , would like a good estimate of $\sigma^{2}{}_{Y}$
- Can use the sample variance given below note division by n-1, not n, since mean is estimated too – if know μ can use n

$$S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} \left(Y_{i} - \overline{Y} \right)^{2}$$



Estimators as Random Variables

- Each of our sample statistics (e.g. the sample mean, sample variance, etc.) is a random variable Why?
- Each time we pull a random sample, we'll get different sample statistics
- If we pull lots and lots of samples, we'll get a distribution of sample statistics



Forecasting

"There are two kind of forecasters: those who don't know and those who don't know they don't know"

John Kenneth Galbraith (1993)



UNIT: Forecasting

- What is forecasting / why we forecast ...
- Past vs. Future
- Methods of forecasting :
 - Qualitative: based on judgment, intuition, and informed opinions (Expert advise / Delphi / Survey ...)
 - Quantitative:
 - Intrinsic the use of historical data to create forecast

Time series: Average / (Weighted) Moving Average / Simple & Double Exponential Smoothing / Linear regression /

- Extrinsic based on external indicators that relate to the phenomena Linear regression / General regression / Classification / Clustering / RNN
- Trends
- Seasonality or other cyclicality



Introduction

- A projection of past information and/or experience into expectation in the future.
- "Those who claim to forecast the future are all lying even if, by chance, they are later proved right"
- "Forecasting is like trying to drive a car blindfolded and following directions given by a person who is looking out of the back window"



Formal vrs. informal forecasting

- Forecasting is a very common activity
- The majority of forecasting is informal
- Why do we need formal forecasting?
 - Coping with complexity
 - Coping with growth
 - Coping with change
 - Need for auditability and justification

Formal forecasting provides a vehicle for communication about the forecast and a basis for systematic improvement.

Forecasting in Action

- **Risk management**: Volatility forecasts are crucial for evaluating and insuring risks associated with asset portfolios. Volatility forecasts are also crucial for firms and investors who need to price assets such options and other derivatives.
- **Capacity planning**: Capacity planning decisions rely heavily on a variety of forecasts related both to product demand and supply.
- **Business and government planning**: Business and governments of all sorts must constantly plan and justify their expenditures. A major component of the budgeting process is the revenue forecast.
- **Demography**: Population forecasts are crucial for planning government expenditure on health care, infrastructure, social insurance, antipoverty programs, and so forth.

Basic Elements of Any Forecast

- Think on any economic variable you want to forecast. What do you need?
- Loss Function: Symmetric or Asymmetric
- Forecast Object: A time series, an event, ... etc.
- Forecast Statement: Point, Range or forecast density
- Forecast Horizon: Short, Medium or Long
- Information Set: Univariate or Multivariate
- Methods and Complexity: Model, ...etc.

Basic Elements of Any Forecast

Once you have done your forecast, someone else can come with another forecast of the same variable. *How do you compare these forecasts?*

• Forecast Evaluation: Different measures of the forecast.

IT IS IMPORTANT TO REALIZE THAT EVERY FORECAST HAS AN ERROR.

In general this error come from three different sources:

- Specification Error
- Approximation Error
- Estimation Error



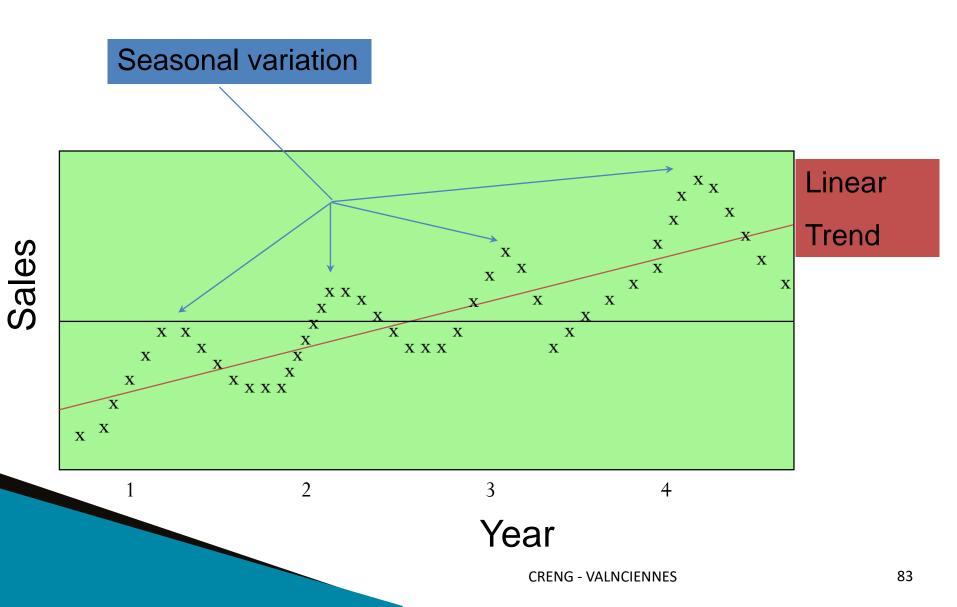
Characteristics of forecasting problems

- Time horizon
 - short-term
 - long-term
- Cost
- Complexity
- Accuracy

- Data patterns
 - Seasonality
 - Trend (upward or downward, linear or non-linear)
 - Cycles
 - Seasonality
 - Randomness
 - Autocorrelation



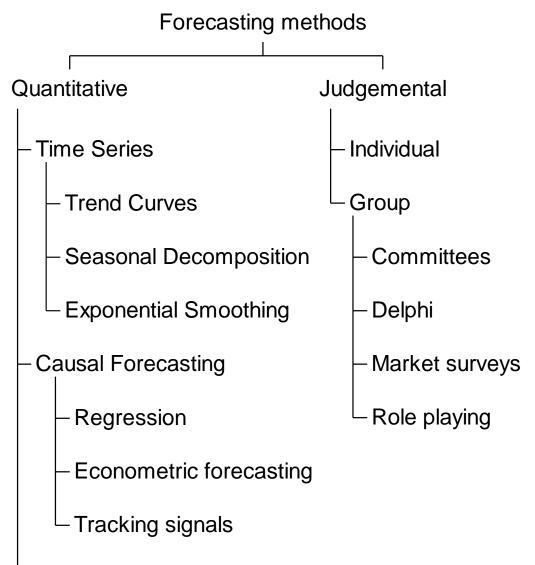
Finding Components of Data





Types of Forecasts

- Qualitative (Judgmental)
- Quantitative
 - Time Series Analysis
 - Causal
 - Relationships
 - Simulation





Delphi Method

- Choose the experts to participate. There should be a variety of knowledgeable people in different areas.
- 2. Through a questionnaire (or E-mail), obtain forecasts (and any premises or qualifications for the forecasts) from all participants.
- 3. Summarize the results and redistribute them to the participants along with appropriate new questions.
- 4. Summarize again, refining forecasts and conditions, and again develop new questions.
- 5. Repeat Step 4 if necessary. Distribute the final results to all participants.



Time Series Analysis

- Time series forecasting models try to predict the future based on past data.
- You can pick models based on:
 - 1. Time horizon to forecast
 - 2. Data availability
 - 3. Accuracy required
 - 4. Size of forecasting budget
 - 5. Availability of qualified personnel



Simple Moving Average Formula

- The simple moving average model assumes an average is a good estimator of future behavior.
- The formula for the simple moving average is:

$$F_{t} = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \dots + A_{t-n}}{n}$$

$$F_{t} = \text{Forecast for the coming period}$$

$$N = \text{Number of periods to be averaged}$$

$$A_{t-1} = \text{Actual occurrence in the past period for}$$

$$\text{up to "n" periods}$$



Simple Moving Average Problem (1)

Week	Demand
1	650
2	678
3	720
4	785
5	859
6	920
7	850
8	758
9	892
10	920
11	789
12	844

 $F_{t} = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \dots + A_{t-n}}{A_{t-1} + A_{t-1} + A_{t-1} + A_{t-n}}$ n

- Question:What are the 3-week and 6week moving average forecasts for demand?
- Assume you only have 3 weeks and 6 weeks of actual demand data for the respective forecasts

Calculating the moving averages gives us:

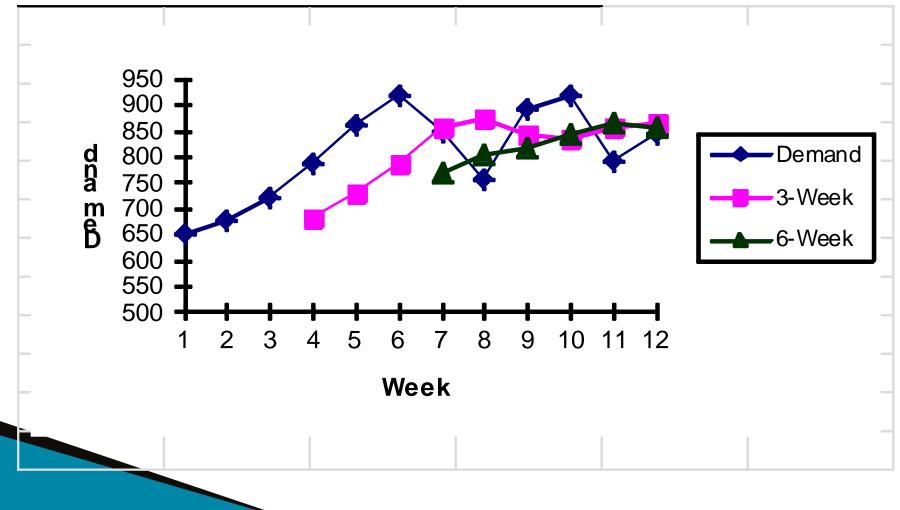


Week	Demand	3-Week	6-Week	
1	650	$F_4 = (650 + 678 + 7)$	720)/3	
2	678	=682.67		
3	720		$F_7 = (650 + 678 + 720)$	
4	785	682.67	+785+859+920	J)/6
5	859	727.67	=768.67	
6	920	788.00		
7	850	854.67	768.67	
8	758	876.33	802.00	
9	892	842.67	815.33	
10	920	833.33	844.00	
11	789	856.67	866.50	
12	844	867.00	854.83	

CRENG - VALNCIENNES



Plotting the moving averages and comparing them shows how the lines smooth out to reveal the overall upward trend in this example.





Simple Moving Average Problem (2) Data

Week	Demand
1	820
2	775
3	680
4	655
5	620
6	600
7	575

 Question: What is the 3/5 week moving average forecast for this data?



Simple Moving Average Problem (2) Solution

Week	Demand	3-Week	5-Week
1	820		
2	775		
3	680		
4	655	758.33	
5	620	703.33	
6	600	651.67	710.00
7	575	625.00	666.00



Weighted Moving Average Formula

While the moving average formula implies an equal weight being placed on each value that is being averaged, the weighted moving average permits an unequal weighting on prior time periods.

The formula for the moving average is:

$$F_{t} = W_{1}A_{t-1} + W_{2}A_{t-2} + W_{3}A_{t-3} + \dots + W_{n}A_{t-n}$$

 w_t = weight given to time period "t" occurrence. (Weights must add to one.)

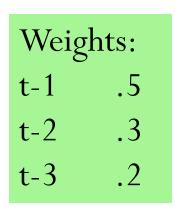
$$\sum_{i=1}^{n} w_i = 1$$



Weighted Moving Average Problem (1) Data

Question: Given the weekly demand and weights, what is the forecast for the 4th period or Week 4?

Week	Demand
1	650
2	678
3	720
4	



Note that the weights place more emphasis on the most recent data, that is time period "t-1".



Weighted Moving Average Problem (1) Solution

Week	Demand	Forecast
1	650	
2	678	
3	720	
4		693.4

 $F_4 = 0.5(720) + 0.3(678) + 0.2(650) = 693.4$



Weighted Moving Average Problem (2) Data

Question: Given the weekly demand information and weights, what is the weighted moving average forecast of the 5th period or week?

Week	Demand
1	820
2	775
3	680
4	655

Weights: .7 t-1 t-2 .2 t-3 .1



Weighted Moving Average Problem (2) Solution

Week	Demand	Forecast
1	820	
2	775	
3	680	
4	655	
5		672

 $F_5 = (0.1)(755) + (0.2)(680) + (0.7)(655) = 672$



Exponential Smoothing Model

$$F_{t} = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

 α = smoothing constant

- Premise: The most recent observations might have the highest predictive value.
- Therefore, we should give more weight to the more recent time periods when forecasting.



Exponential Smoothing Problem (1) Data

Week	Demand
1	820
2	775
3	680
4	655
5	750
6	802
7	798
8	689
9	775
10	

- Question: Given the weekly demand data, what are the exponential smoothing forecasts for periods 2-10 using α=0.10 and α=0.60?
- Assume $F_1 = D_1$



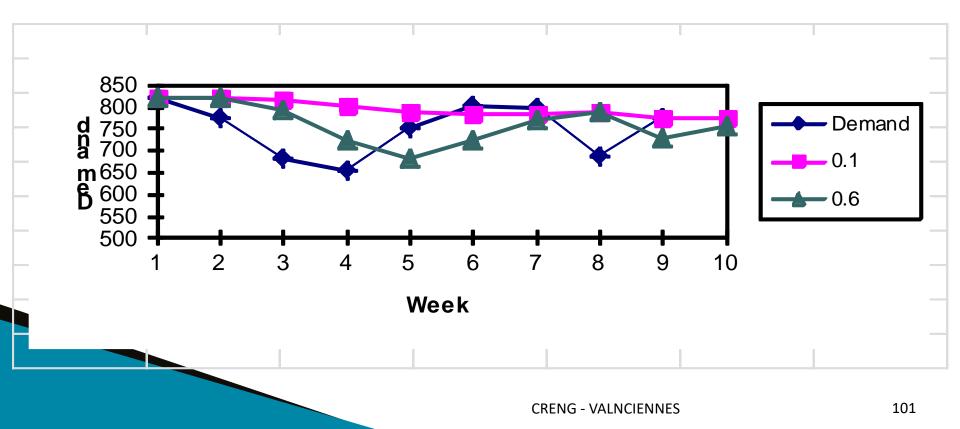
Answer: The respective alphas columns denote the forecast values. Note that you can only forecast one time period into the

f t				
future	. Week	Demand	0.1	0.6
	1	820	820.00	820.00
	2	775	820.00	820.00
	3	680	815.50	820.00
	4	655	801.95	817.30
	5	750	787.26	808.09
	6	802	783.53	795.59
	7	798	785.38	788.35
	8	689	786.64	786.57
	9	775	776.88	786.61
	10		776.69	780.77



Exponential Smoothing Problem (1) Plotting

Note how that the smaller alpha the smoother the line in this example.





Exponential Smoothing Problem (2) Data

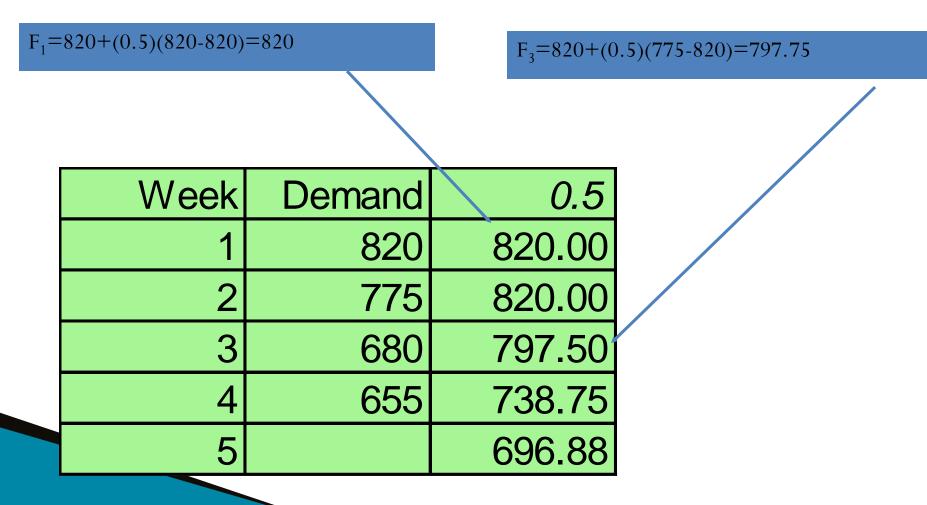
Week	Demand
1	820
2	775
3	680
4	655
5	

Question: What are the exponential smoothing forecasts for periods 2-5 using a =0.5?

Assume $F_1 = D_1$



Exponential Smoothing Problem (2) Solution





Double Exponential Smoothing Model

$$S_{t} = F_{t-1} + \alpha (A_{t-1} - F_{t-1})$$

$$T_{t} = T_{t-1} + \beta (S_{t} - F_{t-1})$$

$$F_{t} = S_{t} + T_{t}$$

 α : data smoothing factor & β : trend smoothing factor



Evaluation of Forecasts

There are many ways of making forecasts, but all of them need the following common ingredients in order to success:

- There are regularities to capture
- Such regularities are informative about the future
- They are encapsulated in the selected forecasting method
- And non-regularities are excluded.



Evaluation of Forecasts

The most common overall accuracy measures are:

Bias:

$$Bias = \frac{1}{n} \sum_{t=1}^{n} (A_t - F_t)$$
Mean absolute deviation/error:

$$MAD = \frac{1}{n} \sum_{t=1}^{n} |A_t - F_t|$$
mean squared error:

$$MSE = \frac{1}{n} \sum_{t=1}^{n} (A_t - F_t)^2$$
root mean squared error:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (A_t - F_t)^2}$$



The MAD Statistic to Determine Forecasting Error

$$\frac{\sum_{t=1}^{n} |A_t - F_t|}{n} = \frac{1 \text{ MAD} \approx 0.8 \text{ standard deviation}}{1 \text{ standard deviation}} \approx 1.25 \text{ MAD}$$

- The ideal MAD is zero. That would mean there is no forecasting error.
- The larger the MAD, the less the desirable the resulting model.



MAD Problem Data

Question: What is the MAD value given the forecast values in the table below?

Month	Sales	Forecast
1	220	n/a
2	250	255
3	210	205
4	300	320
5	325	315



MAD Problem Solution

Month	Sales	Forecast	Abs Error
1	220	n/a	
2	250	255	5
3	210	205	5
4	300	320	20
5	325	315	10
			40

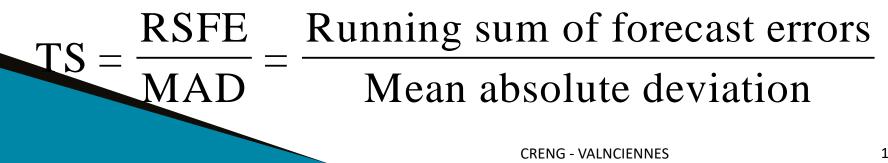
MAD =
$$\frac{\sum_{t=1}^{n} |A_t - F_t|}{n} = \frac{40}{4} = 10$$

Note that by itself, the MAD only lets us know the mean error in a set of forecasts.



Tracking Signal Formula

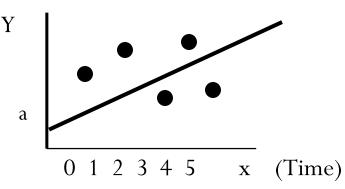
- The TS is a measure that indicates whether the forecast average is keeping pace with any genuine upward or downward changes in demand.
- Depending on the number of MAD's selected, the TS can be used like a quality control chart indicating when the model is generating too much error in its forecasts.
- The TS formula is:





Simple Linear Regression Model

The simple linear regression model seeks to fit a line through various data over time.



 $Y_t \equiv a + b_X$ Is the linear regression model.

Yt is the regressed forecast value or dependent variable in the model, **a** is the intercept value of the the regression line, and **b** is similar to the slope of the regression line. However, since it is calculated with the variability of the data in mind, its formulation is not as straight forward as our usual notion of slope.



Simple Linear Regression Formulas for Calculating "a" and "b"

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

$$b = \frac{\sum xy - n(\overline{y})(\overline{x})}{\sum x^2 - n(\overline{x})^2}$$



Simple Linear Regression Problem Data

Question: Given the data below, what is the simple linear regression model that can be used to predict sales?

Week	Sales
1	150
2	157
3	162
4	166
5	177



Answer: First, using the linear regression formulas, we can compute "a" and "b".

Week	Week*Week	Sales	Week*Sales					
1	1	150	150					
2	4	157	314					
3	9	162	486					
4	16	166	664					
5	25	177	885					
3	55	162.4	2499					
Average	Sum	Average	Sum					
$h = \sum xy$	$\frac{n(\bar{y})(\bar{x})}{-n(\bar{x})^2} = \frac{2499}{3}$	- 5(162.4)(3						
$\sum x^2$	55-5(9)	$-=\frac{10}{10}=6.3$						
a = y - bx = 162.4 - (6.3)(3) = 143.5								
a - y - 0x - 102.4 - (0.5)(5) - 143.5								

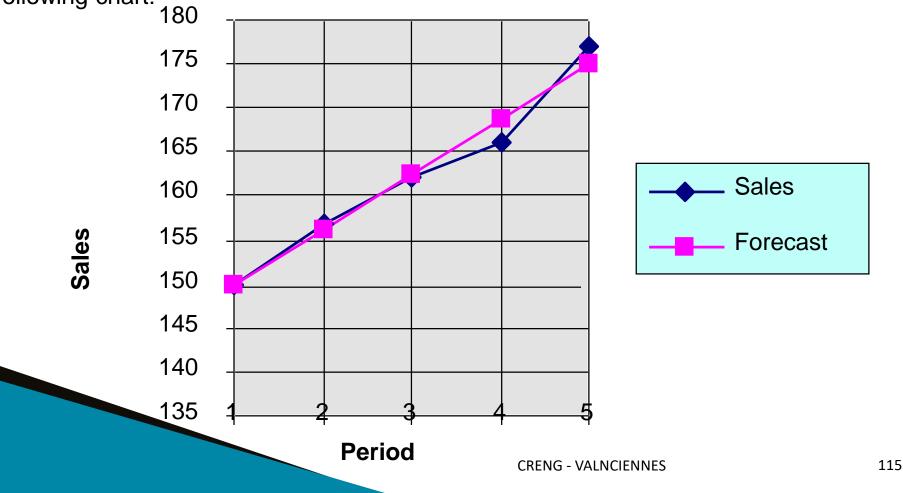
CRENG - VALNCIENNES



The resulting regression model is:

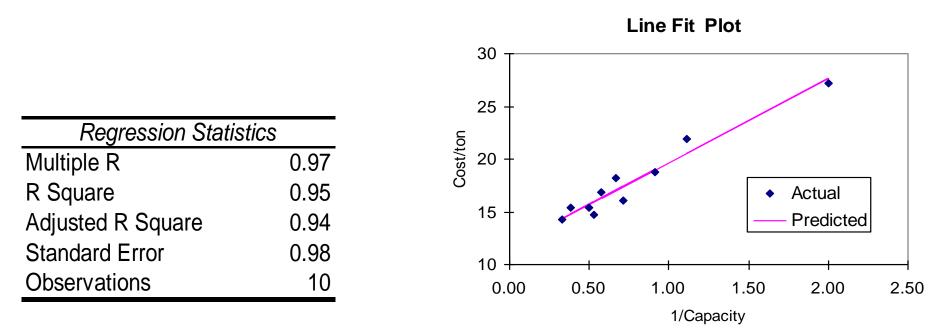
$$Y_t = 143.5 + 6.3x$$

Now if we plot the regression generated forecasts against the actual sales we obtain the following chart:





Regression



	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	11.75	0.60	19.53	0.00	10.36	13.13
1/Capacity	7.93	0.67	11.88	0.00	6.39	9.46



Regression model evaluation

- Model significant p-values
- High adjusted R²
- All coefficients significant p-values
- Low standard error
- No pattern in residuals
- Is model supported by theory?
- Does the model make sense?



Decision Support



UNIT: Aided decision

- Decision process
- The environment of decision : Certain / Uncertain / Risky
- Decision schemes : MaxMin / MaxMax / Laplace ...
- Conditional probability
- Decision trees
- Some metrics (Indicators): Actualized value / Value of the risk / Cost of the imperfect information



Support Decision

Decision support consists of making a choice between several possible options, actions or strategies. This choice will depend on the different "states of environment of decision" or upcoming events.



Environnement of decision

• **Certain**: relevant parameters such as loss, cost, capacity and demand have known values

• **Risky**: parameters with probable results (with an associated known risk)

• **Uncertain**: unable to assess the possibility of several possible events



Objective

- How to make better decisions taking advantage of modern decision and computer technology
 - Managing information
 - Addressing business problems in a structured way
 - Generating and evaluating alternatives
 - Managing uncertainty
 - Learning how to combine judgement with quantitative analysis
 - Using computer technology effectively



Decision-making process

- Step 1: Identify the decision. Identify Your Goal.
- Step 2: Gather relevant information. ...
- Step 3: Identify the alternatives. ...
- Step 4: Weigh the evidence. ...
- Step 5: Choose among alternatives. ...
- Step 6: Take action. ...
- Step 7: Review your decision & its consequences.



Criteria

- Wald (Maximin): determine the worst possible impact for each alternative and choose the alternative with the "best bad", approach or criterion of the pessimist;
- Maximax: determine the best possible impact and choose the corresponding alternative, optimist's approach or criterion;
- Laplace (Arithmetic average): determine the average payoff for each alternative and choose the alternative with the best average;
- Hurwicz (weithed min & max): Take the decision that maximizes the weighted result between maximum and minimum values of each decision = a (mini) + (1-a) (maxi). The role of a is fundamental.
- Savage (Minimax with regret): determine the worst regret for each alternative with the "best bad". This approach seeks to minimize the differences between the most realistic fallout and the best fallout for each state of nature.



Application Case

- The managers of a company wonder if they should recommend the construction of a small, medium or large factory.
- Industry experts report that building a small factory would generate profits of \$ 10 million regardless of whether demand is low, moderate or high.
- In the case of a medium-sized factory, their profits would amount to 7 million if demand for manufactured goods is low and to 12 million if demand is moderate or high.
- This same report indicates that the best profit (16 million) would be obtained if demand is high and that managers have recommended the construction of a large factory.
- However, the profit would only be 2 million if demand is moderate and the factory built is large and there is a possibility of making a loss of 4 million for this large factory if demand is relatively low.

We ask you to solve this dilemma depending on the environment considered?



Certain case

Profits' matrix :

		Future demand			
Profits (in n	nillions \$)	Low	Midium	High	
Alternatives (actions)	Small Factory	\$ 10	\$ 10	\$ 10	
	Midium Factory	\$ 7	\$ 12	\$ 12	
	Large Factory	\$(4)	\$ 2	\$ 16	



Uncertain case

Profits' matrix :

		Future demand			
Profits (in millions \$)		Low	Midium	High	
Alternatives (actions)	Small Factory	\$ 10	\$ 10	\$ 10	
	Midium Factory	\$ 7	\$ 12	\$ 12	
	Large Factory	\$(4)	\$ 2	\$ 16	



Risky case

Profits' matrix :

		Future demand			
Profits (in millions \$)		Low (P=0.3)	Midium (P=0.5)	High (P=0.2)	
Alternatives (actions)	Small Factory	\$ 10	\$ 10	\$ 10	
	Midium Factory	\$ 7	\$ 12	\$ 12	
	Large Factory	\$(4)	\$ 2	\$ 16	



Example

- Three projects whose result depends on the economic situation, identical investment, identical management
 - Project E: Expand our activities (purchase of new store)
 - Project C: develop sales by Catalog
 - Project I: invest in Real Estate
- Profits ' matrix

 VH
 H
 M
 L

 E
 20
 12
 8
 4

 C
 26
 10
 4
 -4

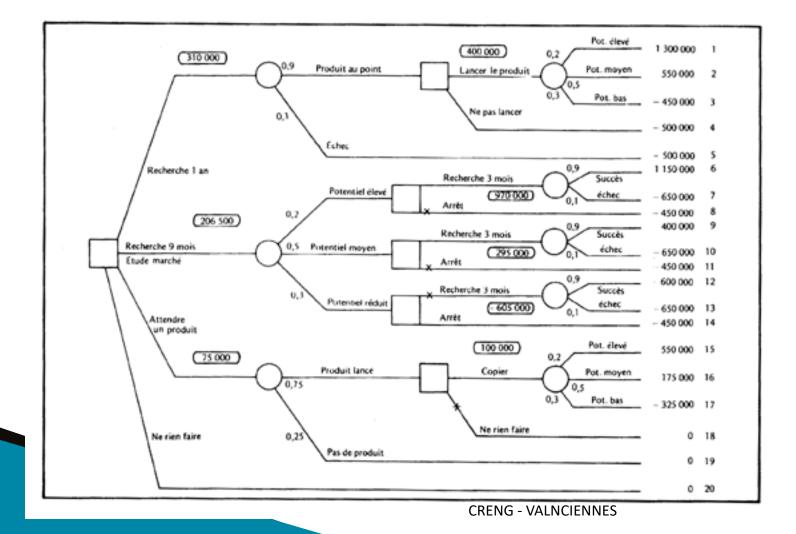
 I
 10
 8
 7
 5

Which project to chose ?



Decision trees

Decisions / Random Events



130



Unit: Data Analysis

- What is (Big) Data? Some numbers / statistics
- Where do the data come from?
- Examples of use of Big Data / Big Data Characteristics 3V's ... 5V's
- Data storage:
 - Physical solution
 - Localization of the date
 - Distributed databases
- Data modelling:
 - Relational database?
 - Data schemes
 - SQL / DQL / NoSQl?



Unit: Data Analysis

- Data analysis (mining):
 - Data colleting (Sampling, DOE, ...)
 - Data Cleaning (preprocessing)
 - Outliers / uncertainty / Quality
 - Data Visualization
 - Extracting intelligence:
 - Classification / ANOVA / PCA / AI / Deep Learning ...
- Computing Power:
 - Decisions are in real-time and the amount of data id high and changes rapidly
 - Processing streams of data!



Unit : Risk Management / Risk Analysis

- Types of risk (Fiscal, Market, Technology, Partner, Contractor, Project execution, Operability,...)
- 5 steps of risk management process
- The Objective and Key Result (OKR) method
- Risk Analysis tools
- Risk ID/Assessment
 - Risk Matrix / Risk score / Prob.
 & Conseq.





Risk Assessment Methodology

System Definition	Risk Breako	down	Structure	
Data Collection				
Hazard and Problems Scenarios	Hazard Analy	ysis		
Qualitative Risk Assessment	Likeliness	High	Î	
Quantitative Risk Assessment	& Impact	ence	Medium-Level Risk	High-Level Risk
Decision Analysis		Probability of Occurence		
Documentation & Communication	on	Proba	Low-Level Risk	Medium-Level Risk
		Low		
			Low Impac	t of Risk High



Class storyboarding / scripting

Course: Forecasting

SEQUENCE : Introduction to the Forecasting	DURATION : 2h
SPECIFIC OBJECTIVE : At the end of this sequence, participant should understand the meaning of forecasting, how forecasting is difficult,	SCHEDULE : Sept. 1th, 10:00 to 12:00
Forecaster's Dilemma (Extreme Events and Forecast Evaluation) and the link to risk and crisis management.	PRE-REQUISTES:

CONTENTE	PEDAGOGICALTECHNIQUES / ANIMATIONS	DURATION	MATERIEL S / DOCUMENTATIONS	WHAT SHOULD BE RETAINED / REMEMBERED	REMARQUES
The forecasting definition	presentation Solicit participants' experiences Immerssive pedagogy	30 mn	Lecture notes / Readings Beekast.com	Particiant soudl be able to define the forecasting.	

References



- http://faculty.salisbury.edu/~xswang/Research/Papers/SERelated/RiskManagement/PrinciplesandPractices.pdf
- George E. Rejda (2011) Principles of Risk Management and Insurance. The Prentice Hall series in finance
- (https://www.pearson.com/us/higher-education/program/Rejda-Principles-of-Risk-Managementand-Insurance-13th-Edition/PGM86481.html?tab=resources)
- http://www.eco.uc3m.es/~jgonzalo/teaching/timeseriesMA.html
- The Journal of International Crisis and Risk Communication Research (JICRCR): https://stars.library.ucf.edu/jicrcr/
- https://www.cethfrance.fr/the-words-of-crisis-forecasting
- https://felipecastro.com/en/okr/what-is-okr/
- http://statistique-et-societe.fr/
- Surowiecki, J. (2005). The wisdom of crowds. Anchor.
- Viktor, M. S., & Kenneth, C. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.
- Ratner, B. (2017). Statistical and machine-learning data mining: Techniques for better predictive modeling and analysis of big data. Chapman and Hall/CRC.
- Agiwal, S., & Mohtadi, H. (2008, April). Risk mitigating strategies in the food supply chain. In American Agricultural Economics Assocation (Annual Meeting).
- Burnson, P. (2015). Cold chain: mitigating risk in a topsy-turvy world. Logistics management (Highlands Ranch, Colo.: 2002).
- Carbonara, N., Costantino, N., & Pellegrino, R. (2014, December). Mitigating supply chain risk: A real options approach. In Industrial Engineering and Engineering Management (IEEM), 2014 IEEE International Conference on(pp. 892-896). IEEE.
- Chopra, S., & Sodhi, M. S. (2014). Reducing the risk of supply chain disruptions. MIT Sloan management review, 55(3), 73.
- Christopher, M. (2003). Understanding supply chain risk: A self-assessment workbook. Department for Transport-Cranfield University–Cranfield University.
- Marsh and RIMS. 2014. Excellence in risk management survey XI. New York: Marsh and Risk and Insurance Management Society.
- McKinsey. 2014. Enterprise-risk-management practices: Where's the evidence? McKinsey Working Papers on Risk, Number 53 CRENG - VALNCIENNES 136



References

Best Sellers in Risk Management





Tools

- R: https://rstudio.com/
- SAS: https://www.sas.com/en_us/software/on-demand-for-academics.html
- RapidMiner: https://rapidminer.com/get-started/
- Data analysis toolpack in excel: https://support.office.com/enus/article/load-the-analysis-toolpak-in-excel-6a63e598-cd6d-42e3-9317-6b40ba1a66b4
- Minitab: https://www.minitab.com/en-us/support/
- Risk and decision tools @ Palisade (add-Ins for excel): https://www.palisade.com/decisiontools_suite/
- Riscomania Game: https://www.cipe.fr/article-a-laune/riscomania-jeu-management-risques/