0130

Efficient Learning and Working Methods

Man Nguyen– HCMUT

Ngày 16 tháng 9 năm 2014

Introduction

Audiences: Undergraduates and first year graduates
Time Distribution: Theory + Practice+ Discussion,

be respectively spanned in 6+ 9+ 12 + 3h (30h in 10 weeks);

Theory part I briefs theoretic concepts collective advices on soft skills in 6h, then part II provides basic principles of `Computational Mathematics & Design of Experiments` in 9h.

Practice: self-working in team at home, learning ways

- to effectively learn in the universities or
- to carry out specific case studies / jobs in working places;

in total working up to 18h per semester, or 2 hours / week on team project.
Discussion and presentation in 12 h: interactive talks in class

Disclaimer

This curriculum was adapted from materials inspired from
* the Statistical, Mathematical and Computational Sciences,
* plus Economic and Environmental Perspectives, as well as
* other texts about *Methods and Models in Scientific Research*.

1. *Collapse- How have societies failed or succeeded?*
   Jared Diamond, 2005
2. *Five Equations That Changed The World*,
   Michael Guillen, 1995
3. *Tip Collection on how to survive in academic world*, Dianne Prost O'Leary, University of Maryland (online texts)

Part 0: An Overview about Scientific Method in Our World.

Part I: Understanding of Learning Process- a background and related principles

- Bloom Taxonomy
- Some rules to have a Career Success

Part II: What Modern Workers nowadays need

(but academic sector could not fully provide?)

Part II-A: Soft Skills for Modern Workers

Part II-B: Useful personalities for Modern Workers

Part III: Valuable approaches for Graduates

Part III: Valuable approaches for Graduates

A/ Models and modeling

B/ Computational Mathematics

C/ Design of Experiments

D/ Statistical Science and Statistical Thinking

E/ Ethical matters in R & D

Why do we need Scientific Investigations?

Nowadays Scientific investigations are important not only in the academic laboratories of research universities but also in the engineering laboratories of industrial manufacturers.

1. Computer & Computational Science,
2. Mathematics and
3. Statistics

are key components of Scientific Investigation!

Hence we first discuss a few major connected topics of
✓the Scientific Method and
✓✓the **Role of Statistics in Science Investigation and Development**

Why is Computational Science and Statistics?

At least the followings

1. Promising careers: **Statistical Consultant**, Biostatistician, Statistical Data Analyst, Actuarial scientists ...

   (The New York Times 05-08- 2009,
   For Today's Graduate, Just One Word: Statistics
   by Steve Lohr)

2. Statistics provides The Scientific Methodology in the context of the scientific process and discovery

3. Statistics sets fundamental background for obtaining optimal solutions under uncertainty!

A. The Scientific Method

In 18th Century we had:

- Newtonian Mechanics. Universal Gravitation.
- The problem of Navigation. Astronomy and the chronometer.
- Biology:
  - Development of individual organisms.
  - Circulation of the Blood.
  - Cell theory.
- Chemistry: Composition of matter.

The scientific method In 19th Century

- Darwin and the theory of Evolution.
- Development of Geology
- The atomic theory of Matter
- The Wave theory of Light.
- Electricity and Magnetism.
- Mathematics: Geometry, **Statistics/Probability**. [Poincare, Hilbert, ...]
- Engineering and Science.

The Scientific Method In 20th Century

VERY SHORT LIST.

- Modern Biology.
- Relativity Theory.
- Quantum Mechanics [Dirac, Bohr, Heisenberg, Hawking ...]
- Data Science: **Data Analysis and Statistics**.
- Social Sciences....

under developing by ourselves! Particularly

* Social Networks, Media, Intelligence and related things;
* Traffic and Communication Engineering
* Life sciences ....

==================

Definition

Statistics is the science of problem-solving in the presence of variability or uncertainty

This identifies *Statistics* as a scientific discipline, which demands the same type of rigor and adherence to basic principles as Physics or Chemistry.

The definition also implies that when problem solving involves the collection of data, the **science of statistics** should be an integral component of the process.

B. The Development of Statistics

- Monte Carlo casino gamblers $\Rightarrow$ MC simulation algoritthm
- Pascal and Fermat established some of the foundations of **Probability** (1654).
- Early developments in Data Collection and Processing in the Seventeenth Century.
  * Tables of Mortality in England 1662.
  * Tables of Annuities: Johan De Witt (1671).
- Normal distribution. Quetelet (1835).
- Poisson, The law of large numbers.

B. The Development of Statistics

- Galton 1879. Introduced the concepts of correlation and regression (towards the mean). He used the Quetelet-Gaussian distribution.
- Pearson followed Galton and established the foundations of modern statistics.
- **R A Fisher. The design of experiments**, 95% confidence.
- Connection between Lung Cancer and smoking $\Rightarrow$ Bayesian net and Bayesian statistics

C. Why Computational Science & Statistics for the 21st Century?

You already known

- 'Scientific investigations are important not only in the academic laboratories of *research universities* but also in the engineering laboratories of *industrial manufacturers*.'

- Statistical methods are applied in an enormous diversity of problems in fields as?

Computational Science & Statistics for the 21st Century

- *Agriculture* (which varieties grow best?)
- *Genetics, Biology* (selecting new varieties, species)
- *Economics* (how are the living standards changing?)
- *Market Research* (comparison of advertising campaigns)
- *Education* (what is the best way to teach small children reading?)
- *Environmental Studies* (do strong electric or magnetic fields induce higher cancer rates?)
- **Quality engineering** ...

Part I: Understanding of Learning Process- a background and related principles

Why we need to understand the Learning Process?

To be:
- a useful citizen, a productive worker
- a famous singer or fashion model,
- a well known people, ...
- or to have a Career Success?

But how? **First step**: to understand basic rules of learning and how to exploit.

Contents of Part I

A/ Bloom Taxonomy

B/ Some rules to have a Career Success

A/ Bloom Taxonomy

What is the Original Bloom Taxonomy?

A proposed scheme about six key levels in learning things.

Knowledge
$\rightarrow$ Comprehension
$\rightarrow$ Application
$\rightarrow$ Analysis
$\rightarrow$ Synthesis
$\rightarrow$ Evaluation

The Revised Bloom's Taxonomy

What do you know about the Revised Bloom's Taxonomy: Two distinct dimensions

I/ The Knowledge Dimension
and

II/ THE COGNITIVE PROCESS DIMENSION: formed by new six phases

Remember
$\rightarrow$ Understand
$\rightarrow$ Apply
$\rightarrow$ Analyze
$\rightarrow$ Evaluate
$\rightarrow$ Create

On the new facets of Cognitive Dimension

Introducing,
Questioning and
Discussing
the **new constituents of Cognitive Dimension**

1/ Remember
$\rightarrow$ Understand
$\rightarrow$ Apply
$\rightarrow$ Analyze
$\rightarrow$ Evaluate
$\rightarrow$ Create

The Knowledge Dimension revisited

This one can be classified into at least four types:
Factual Knowledge
Conceptual Knowledge
Procedural Knowledge
Meta-cognitive Knowledge

Why should we know about Bloom Taxonomy?

Open Discussion

B/ Some rules to have a Career Success

1. Be passionate about what you do
2. Establish early goals of your life
3. Demand reasons for / when doing somethings
4. Set priorities which one is most important to do
5. Network up and down getting to know or interacting with others
6. Communicate well and often how to be a good speaker or leader
7. Hone your skills practice whenever you get chance
8. Get a life making friends, mates and having collaborators

Rule 1: Be passionate about what you do

Passion is

- a luxury,
- but essential to high success.

Tell the class some examples of passion in your life?

Rule 2: Establish early goals

Goals: things that you plan to do.

Short term: finish project, organize a meeting, teach child to swim, take a course, plant a garden, . . .

Long term: find better job, pay off mortgage, write a book, ...

What are not goals?!

These are not goals:

- I want to be a grandmother.
- I want to be promoted.
- I want to win the Nobel Prize.

They are just wishes!

Better recognition:
Not
I want to win the Nobel Prize,

but
I want my work to be good enough to merit an award.

Goals and Wishes

Not

I want to be promoted.

but

I want ...

Rule 3: Demand reasons

Reasons:

Know why your supervisor wants a task done.
More importantly, know why you are doing the task!

Your reason should be drawn from a short list:

- I want to.
- It is a step toward my goals.
- It gives back to the community.

Rule 4: Set priorities

Setting priorities
means **time management**.

- Put the big rocks in first.
- If you care about your family, save *some energy* for them!
- If anything is important to you, schedule time for yourself.

Rule 5: Network up and down

Networking:

- Get to know the senior people. Right now, you need to be known by them.
- Get to know the junior people.

Right now, they need you.
Later, you will need them!

Rule 6: Communicate well and often

Make sure that your supervisor and your colleagues know what you do.

Keep your resume updated.

Rule 7: Hone your skills

Hone your skills.

Never stop learning.

Be appropriately confident of your abilities.

Rule 8: Get a life

Keep a balance in your life.

Make sure all of you self-esteem eggs are not in one basket.

Rule 8: Get a life

Support system:

- faith
- family
- friends
- fitness
- fun

What else can you propose?

Next week topic:
Part II: Soft Skills for University undergraduates

Part II-A: Soft Skills for modern workers

1. Reading
2. Writing

---

3. Listening
4. Talking

---

5. Searching
6. Arguing- Reasoning

---

7. Effective Presenting
8. Effective Negotiating

Soft Skill 1: Reading

Soft Skill 2: Writing

Soft Skill 1 and 2: Practices

Student groups brief articles

- Brief History of Silicon Valley in Tia Sang, pages 22-27 number 15, 2007
- India: Economic Development on the Science and Technology Foundation, pages 25-30 number 14, 2007
- 35 un-solvable problems, pages 30-33 number 11, 2006
- Fight Climate Change: the war to protect the Human-being future, pages 20-23 number 16, 2009

Request: reduce each article to your own essay that has 1/3 volume of the original.

Soft Skill 3: Listening Skills

During Socializing

- Before talking/replying to someone, (try to) hear them first.
- Rephrase what you have heard in your own words in brief sentences
- Ask politely the speaker to repeat points that you did not catch

or Working

- Checking and clarifying (complicated) points during conversation in your own words
- Express your positive/ optimistic opinions if the speaker's saying is interseting.
- (*) Express politely your opinions/ comments if the speaker's saying is not right

See more at Soft Skill 6: Arguing- Reasoning
and Soft Skill 8: Effective Negotiating

Soft Skill 4: Talking

Type A: Talking- Socializing

Type B: Talking- Working

Type A: Talking- Socializing would include
Basic Level

- welcoming- introducing yourself or someone
- responding to introductions
- opening small talks
  (travel, weather, accommodation)
- offering (tea,...)

Soft Skill 4: Talking- Socializing

Advanced Level

- Request (favours) and Responses
- Topics to talk (jobs, family, spare time, origin, history, childhood)
- Inviting and Responses (accepting or declining)
- Changing - Ending a topic
- Checking and clarifying: to make sure you fully and correctly get the points

Soft Skill 4: Type B Talking at Work Place

Checking and clarifying : especially important in work places, in meetings, where the background of the speakers and the listeners often are different...

Talk with decent speed / emotion : not too fast, not feel inferior (i.e. confidently) to boss / native speaker; not use idiom so often, use simple, well-articulated language

Showing appreciation : tell honestly/ frankly what you know best

Offering opinions in meetings : in a polite way with soft voice, backed by cheked information

Making arrangements - promises : be on time- try your best to fulfill the promises

See more at Soft Skill 6: Arguing- Reasoning
and Soft Skill 8: Effective Negotiating

Soft Skill 3 and 4: Practice

* The class suggests a policy on something to make, then

* split classmates into two groups A and B, and

* simulate a meeting about making that decision / policy

in which

- group A promotes necessary actions to bring policy to life, and
- group B evidently provides backwards of the policy

Soft Skill 5: Searching Information

How to efficiently search useful information and discover knowledge in our Internet Age?

Soft Skill 6: Arguing- Reasoning

Meaning: To debate or discuss; to treat by reasoning

- How to logically argue to protect your opinions?
- How to convincingly reason to get the right conclusion?

Soft Skill 5 and 6: Practice

Soft Skill 7: Effective Presenting

- Suitable fontsizes when writing
- Max 7 key points in one slide
- Keep consistency in using fontsizes and types (time new roman, arial, courrier ...)
- Use short sentences, telling audiences ideas is more useful than providing right syntax

- Employ max 3 colors in slides, don't change colors so often
- Images, pictures, short movies are allowed, but not often
- Pay attention to background color and text color

Soft Skill 8: Effective Negotiating

Crucial for your career! Proposed steps include the Setting up phase

1. Set the ground
2. Set the agenda
3. Establish positions
4. Clarify positions

Soft Skill 8: Effective Negotiating

and then the Negotiating phase

5. Managing conflict
6. Making and responding to proposals
7. Bargaining
8. Conclusion and agreement

Soft Skill 8: Effective Negotiating

Step	**Commu. skills**	**Culture & tactics**	**Language**	**Practice**
5. Managing conflict	avoiding personal criticism	conflict versus collaboration	downtoning your language	handling conflict
6. Making, responding to proposals	encouraging responses	emotional versus neutral behaviour	making suggestions	making, responding to proposals

Soft Skill 8: Effective Negotiating

Step	Commu. skills	Culture & tactics	Language	Practice
7. Bargaining	maintaining positive communi.	high-context vs. low-context cultures	exerting pressure and making conditions	bargaining practice
8. Conclusion and agreement	summarizing and agreeing follow-up	win-win vs. win-lose	concluding and closing	closing the negotiation

Part II-B: Useful personalities for Undergraduates

Teamwork most important all times!

Honest (integrity or fairness and straightforwardness in conduct, thought, speech)

Competence (ability; adequacy)

Open-minded (ready to entertain new ideas)

Teamwork: 3 important questions

1/ What is an ideal team? The team could provide good / best performance, as a whole, with given resources.

2/ What are major components of a team?
a) individuals and their technical skills
b) the communication between them

3/ How to improve team performance?

Teamwork- How to improve team performance?

Can you think and tell the class the key requirements for an 'ideal team':

1. Team members are active
2. Role assignment is clear
3. Advice, indication, and suggestion can be performed freely
   So: every team member has an opportunity to understand each other's character!
4. There is a team atmosphere that backs up and supports individual team members.
5. What else could you think??

How to improve team performance?

Brain- storming discussion (10 minutes)

⋮

Could you group requirements that we have just found into dimensions of teamwork? (10 minutes)

Dimensions of Teamwork

1. Communications
2. Openness
3. Task coordination
4. Team spirit
5. Leadership (maintenance of tasks, ...); and
6. Adaptability

Practical Questions:
a/ Which one is most useful?
b/ How to achieve / get that dimension?

Summary: Good Rules for Your Daily Life and ...

Open reviewing

Part III: Valuable approaches for Graduates

- The objective of a masters in **Computer and Computational Science** (as mathematics, engineering, statistics, CS, MIS ...) is
  1. to expose you (graduate students) in to challenging and exciting real-world problems, and
  2. to train you problem solving methods for/ through those problems.

- Real-world problems arise in industrial and government laboratory research. Students get experience in the **team approach** to problem solving.

Valuable approaches for Graduates- Contents

A/ Models and modeling

B/ Computational Mathematics

C/ Design of Experiments

D/ Statistical Science and Statistical Thinking

E/ Ethical matters in R & D

Part III. A: Models and modeling

1. **Models and model building** are commonplace in the engineering and physical sciences.
2. A research engineer or scientist generally has some basic knowledge about the phenomenon under study and seeks to use this information to obtain a plausible model of the data-generating process.
3. **Experiments** are conducted to characterize, confirm, or reject models–in particular, through hypotheses about those models.
4. Models take many shapes and forms, but in general they all **seek to characterize one or more response variables**, perhaps through relationships with one or more factors.

Part III.A/ Models and modeling- General Key characteristics

All mini research projects share some common points:

1. use mostly some structures, methods and algorithms of **Discrete Math** for investigation of that problem,
2. (specifically some cases) can be reduced to a certain **(polynomial) system of equations**,
3. various techniques of **Operations Research** have to be employed during formulation and solving processes,
4. employ **Computing software** at some computing stages, and
5. in various cases, Statistical methods and thinking are extremelly useful, particularly if you have realistics dada!

Part III. A: Models and modeling- concepts

Mathematical model and Statistical model.

1. A model is termed *mathematical* if it is derived from theoretical considerations that represent exact, error-free assumed relationships among the variables.

2. A model is termed *statistical* if it is derived from data that are subject to various types of specifications, observations, experimental, and/or measurement errors.

Part III. A: Models and modeling- concepts

3. A suitable **statistical model** for our data will often be of the form

$$\text{Observed data } = f(\mathbf{x};\mu) + error;$$

where $\mathbf{x}$ are variables measured and μ are parameters of our model.

4. **Observation**. The collection of information in an experiment, or actual values obtained on variables in an experiment. Response variables are outcomes or observed values of an experiment.

A/ Models and modeling- example in Economics

- **Consumption Theory** tells us that generally people increase their consumption expenditure C as their after-tax (disposable) income Y_d increases,
  but not by as much as the increase in their disposable income.
- Mathematically in explicit linear equation form, as:

$$C = b_0 + b_1 Y_d$$

  where b_0, b_1 are unknown constants called parameters.
- But **different people** having the same disposable income are likely to have somewhat **different consumption expenditures**, the above deterministic relationship must be modified to include a random disturbance or error term, u, making it stochastic or statistical model:

$$C = b_0 + b_1 Y_d + u.$$

Part III. B: Computational Mathematics

See **Logistical and Transportation Planning** text.

Part III. C: Design of Experiments- History

Sir Ronald Aylmer Fisher

- The DOE's history goes back to the 1930s of the 20-*th* century, when Sir R. A. Fisher in England used Latin squares to randomize the plant varieties before planning at his farm, among other activities.
- The goal was to get high productivity havests. The mathematical theory of combinatorial designs was developed by R.C. Bose in the 1950s in India and then in the US.
- Nowadays, DOE is extensively studied and employed in virtually any engineering and scientific investigation, and the Mathematics for DOE is extremely rich.

Part III. C: Design of Experiments- Key contents

The main aim of experimental design is
to identify an unknown function

$$\phi : \ D \to \mathbb{R}$$

on a certain design D

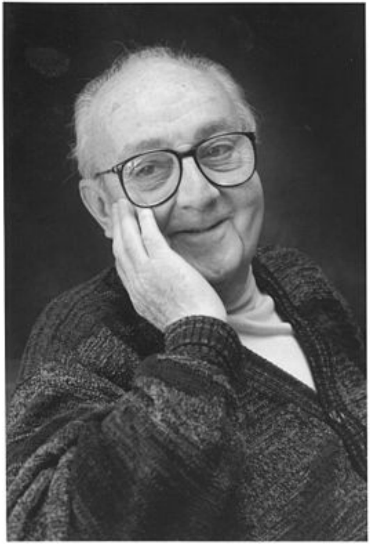
ϕ is a mathematical model of some quantity of interest
(*efficiency, quality, customer loyalty, bearing hardness* ...),
that is be computed or optimized.

Design of Experiments- History from 1950s

- Giai đoạn thứ hai hay còn gọi là giai đoạn công nghiệp, khi DOE được áp dụng nhiều vào ngành công nghiệp, được đánh dấu bằng công trình của Box và Wilson về **phương pháp luận mặt đáp ứng** (*response surface methodology – RSM*) vào năm 1951.

- Hai ông đã chỉ ra hai đặc tính khác biệt đặc trưng giữa các thí nghiệm trong công nghiệp so với nông nghiệp là
  * **tính tức thì** (*immediacy*) của kết quả thí nghiệm mà không phải chờ một thời gian dài như trong nông nghiệp và
  ** **tính tuần tự** (*sequentiality*) khi các kỹ sư có thể tìm hiểu rất nhanh các thông tin quan trọng từ một số ít thí nghiệm để từ đó thiết kế các thí nghiệm sau.

DOE nhờ đó dần được áp dụng nhiều hơn trong ngành hóa chất và công việc nghiên cứu phát triển (*R & D*) trong công nghiệp.



George Edward Pelham Box

Design of Experiments- studied directions

Some mostly studied directions together with appropriate mathematical tools include:

The study of factorial designs mathematically described as matrices which consist many factors, each has several settings (levels), and these settings are arranged in a regular way. The main tools are combinatorics, algebra, geometry.

The study of optimal designs the main tools are matrix and probability theory.

The response surface methodology the main tools are combinatorial designs and approximation theory... Application in chemical engineering, ...

Design of Experiments- studied directions

- **Thiết kế thí nghiệm tối ưu** (*optimal DOE*) cũng được khai sinh nhờ công của Kiefer và Wolfowitz.
- Kiefer đã đề xuất phương pháp lựa chọn thiết kế dựa trên một tập hợp các tiêu chí tối ưu. Ý tưởng của ông là chọn loại thiết kế sao cho **các tham số của mô hình có thể được ước lượng** với độ chính xác cao nhất.
- Nghiên cứu này không được ứng dụng nhiều vào thời điểm đó do thiếu thiết bị tính toán đủ mạnh để hiện thực giải thuật mà ông đề xuất. Đến nay, lĩnh vực này phát triển rất sôi động nhờ máy tính điện tử và được ứng dụng rộng khắp.

Design of Experiments- History from 1970s

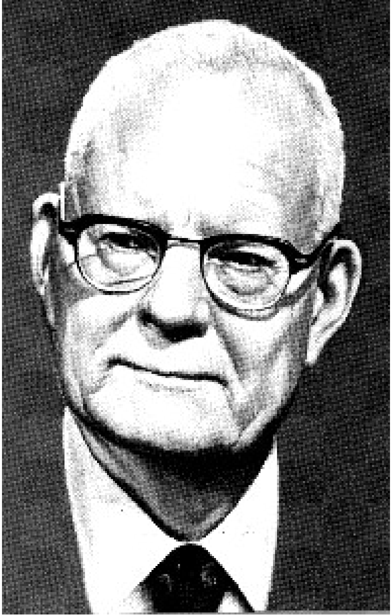
- Giai đoạn thứ ba được đánh dấu bằng sự quan tâm của các hãng sản xuất quy mô lớn đến DOE khi nó được áp dụng vào quy trình kiểm soát chất lượng vào cuối những năm 1970, được gọi là Kiểm soát chất lượng thống kê (*Statistical Quality Control – SQC*).
- Việc áp dụng DOE vào quy trình tại Mỹ chỉ dừng lại ở ngành công nghiệp hóa chất mà không được các hãng công nghiệp sản xuất xem trọng. Genichi Taguchi, một kỹ sư người Nhật, đã phát triển **thiết kế tham số mạnh** (*robust parameter design*) với mục tiêu:
    1. Làm cho quy trình bớt phụ thuộc vào các yếu tố môi trường hay các yếu tố khó kiểm soát khác
    2. Làm cho sản phẩm bớt phụ thuộc vào sự khác biệt từ các thành phần cấu tạo
    3. Tìm ra các cấp độ của nhân tố tác động sao cho vừa làm cho giá trị trung bình đạt tới giá trị mong muốn vừa làm giảm sai số.

Design of Experiments- Taguchi's paradigm- Comments

1. Ông sử dụng các thiết kế nhân tố từng phần các ma trận trực giao cùng với các kiến thức thống kê truyền thống để giải quyết bài toán này.
2. Các hãng sản xuất Nhật nhờ đó đã tạo nên bước nhảy thần kỳ cho đất nước này bằng những sản phẩm có chất lượng cao và ổn định.
3. Cách tiếp cận từ nhu cầu của một nhà công nghiệp nhưng chưa được kiểm chứng từ góc độ lý thuyết thống kê đã tạo ra sự tranh cãi rất lớn.
4. Ý tưởng và mục tiêu của Taguchi về mặt kỹ thuật là rất rõ ràng, có nhiều sai sót trong thiết kế và phương pháp phân tích dữ liệu.
5. DOE đã bước vào trường học và trở thành một môn học bắt buộc của nhiều chương trình đào tạo kỹ sư ở các trường đại học.

The innovation of quality as a strategic weapon in Japan after WW2

- What Toyota and other Japanese companies had invented was a **manufacturing paradigm** that was superior to the century-old mass fabrication paradigm of the West.
- The new paradigm: **quality versus cost versus responsiveness versus flexibility**.
  The Japanese manufacturing virtuosity made it possible to have them all rather than choose.
- In the late 1970s the Japanese appeared to have huge competitive advantages: they sold similar products to those of Western companies, but
  * at **lower cost, with fewer defects**,
  ** more **reliability and better durability**.



Dr. W. Edwards Deming

Design of Experiments- quality by design

- Quality engineers around the world nowadays demonstrate the impact of statistical methods on process and product improvements and the competitive position of organisations.
- They use a systematic approach to the evaluation of benefits from process improvement and quality by design (QbD) that can be implemented within and across organisations.
- The different approaches to the management of industrial organisations can be summarised and classified using a four-step quality ladder (Kenett and Zacks, 1998).
- The four approaches are: (1) fire fighting , (2) inspection, (3) process control, and **(4) quality by design**.

The connection between four management approaches and set of suitable statistical methods is shown in Table .

Design of Experiments (DOE)- the quality ladder

Statistical methods	Management approach
DOE, TQM Six-sigma, ...	Quality by Design (1980s- now)
Statistical Process- Quality Control	Process Improvement (1960-70's)
Sampling	Inspection (1950-60's)
Data Accumulation	Fire Fighting (before WW 2)

Bảng : Quality ladder

Statistically designed experiments in R & D and in manufacturing

- *A few well-known rules* in any economy say that industries and services nowadays need help from academia, asking for high quality R & D.
- The problem practically comes down to a matter of cost: conducting R & D activities costs money, but this spending is worthy to do in the pre-production period.
- As a result, a few key *practical demands*, be proposed by Malcolm Balrige, a former US Secretary of Commerce.

DOE = Design of experiments plays an extremely important role in Quality Engineering, and moreover very useful for

* scientific investigation and
** technological innovation!

Statistically designed experiments in R & D and in manufacturing

In the article ‘Designing for productivity’ (Design News, 1982), Malcolm Balrige said that:

- for managers, the challenge is to create an organizational environment that fosters *creativity, productivity* and *quality consciousness*; that
- 40 % of all costs in getting a product to the market place are in **the design cycle**;
- and that top management must emphasize prevention, rather than correction.

Prevention means **conducting designed experiments** in the so-called *design cycle* or *off-line manufacturing*

What is a design cycle?

A design cycle, also called 'off-line manufacturing phase', informally is a period in which you (a manager or engineer)

- use all kind of information of an expected commodity
- to choose the best raw ingredients and components,
- to tune their combination

so that when the commodity is produced and sold to the market place, it meets **society demands**, or more precisely, satisfies **expectation of clients**.

Design of Experiments- Basic concept

A natural and powerful remedy for the problem is **experimentation**, specifically, **statistically designed experiments**.

Definition

Statistically designed experiments are
* a sequence of trials or tests
* performed under controlled conditions
* which produces measurable outcomes.

Design of Experiments- a systematic process

The common point of these directions are the research targets:
- look at a specific problem in science or industry,
- find a suitable design, mathematically investigate it, and then
- use it in assembly lines...

Briefly, we have the following phases:

1. planning phase determine problem and select responses of process/product
2. designing & constructing phase learn how to find (construct) those experiments, given the scope of expected commodities and the parameters of components, that have been determined in planning phase

Design of Experiments- key steps

3. exploring & selecting phase investigate design characteristics (proposed by researchers) to choose good designs. For instance, in factorial designs we learn how to detect *interactions between factors* (components), and if they exist, calculate how strongly they could affect on outcomes
4. conducting phase where the planned experiment is carried out; and, finally
5. analyzing & consulting phase study how to use them (ie., conduct experiments in a particular application, measure outcomes, analyze data obtained, and consult clients what they should do).

A golden rule in the field of quality management and improvement: the earlier you make the change, the greater the effect down the road (ie., the commodity get in the market place).

Design of Experiments- A golden rule

Designing experiments specifically help us achieves the followings key aims:

(a) **Find & Perform experiments** to evaluate the effects the factors have on the characteristics of interest, and also discover possible relationship among the factors (which could affect the characteristics). The goal is to use these new understanding to improve product.

(b) **Answers to questions** such as:

- What are the key factors in a process?
- At what settings would the process deliver acceptable performance?
- What are the key, main and *interaction effects* in the process?
- What settings would bring about less *variation* in the output?

DOE and beyond

We understand DOE belongs to a broader catergory called *Total Quality Management*; in which we concern three things: quality planning, quality control an quality improvement. What actions should we conduct to improve products and processes in TQM context? We could study some related topics

- quality planning: organizing for improvement, inustrial R & D strategies, and planning for quality
- quality control: DOE, project execution, measurement assurance, and verification
- quality improvement: the improvement process, measures of effectiveness.

DOE and beyond

1. C.F. Jeff Wu, Michael Hadamard (2000),
   Experiments: Planning, Analysis and Parameter Design Optimization, Wiley, 630 pp
2. Madhav, S. P. (1989),
   Quality Engineering using robust design, Prentice Hall.
3. Raymond H. Myers, Douglas C. Montgomery and Christine M. Anderson-Cook (2009),
   Response Surface Methodlogy: Process and Product Optimization Using Designed Experiments, Wiley

Calyampudi Radhakrishna Rao

Part III.D: Statistical Science and Statistical Thinking

The definition of the discipline of statistics: **Statistics is the science of problem-solving in the presence of variability**.

Statistical methods nowadays are applied in an enormous diversity of problems in such fields as:

- *Agriculture* (which varieties grow best?)
- *Genetics, Biology* (selecting new varieties, species)
- *Economics* (how are the living standards changing?)
- *Market Research* (comparison of advertising campaigns)
- *Education* (what is the best way to teach small children reading?)
- *Environmental Studies* (do strong electric or magnetic fields induce higher cancer rates?)
- **Quality engineering**

Part III.D: Statistical Science - Quality engineering

- Quality and productivity are characteristic goals of *industrial and service processes*, which are expected to result in goods and services that are highly sought by consumers and that yield profits for the firms that supply them.
- The setting just described provides one motivation for examining the role of Statistics in scientific and engineering investigations.

- Nowadays it is no longer satisfactory just to monitor *on-line industrial processes* and to ensure that products are within desired specification limits. Competition demands that a *better product be produced within the limits of economic realities*.
- How to achieve this with least cost?

Part III.D: Statistical Science- Branches

- DOE and Response Surface Methodlogy (RSM)
- MDA in Statistical modeling and inference ...
- Bayesian Statistics

- Statistical simulation (including MCMC simulation, rare event simulation ...)
- Spatial statistics (applications in Environment, Ecology, Petroleum engineering ...)
- Big Data and Data mining
- Process mining
- Statistical shape analysis

Part III.D: Statistical Science- RSM

- **Why is RSM?** In Conventional Optimization, a solution is not explicitly obtained in a nonlinear large problem. When using response surface creation, a function is approximated, then optimization calculation using the response surface is easier to obtain, thanks to the Least Square method and employing designed experiments.

- **Mathematically, what is RSM?** This is about an approximation-based optimization, where the approximation function is called response surface.

In general, suppose that the scientist or engineer (refer to as the *experimenter*) is concerned with a product, process or system involving

- a response y that depends on
- a finite number of controllable input variables (called independent variables) $\xi_1, \xi_2, \xi_3, \ldots, \xi_k$.

The relationship is

$$y = f(\xi_1, \xi_2, \xi_3, \ldots, \xi_k) + \varepsilon$$

- where the form of the true response function f is unknown and perhaps very complicated, and
- ε is a term that represents other sources of variability not accounted for in f.

In practice several assumptions are used,

- that the *statistical error* ε has a normal distribution with mean 0, and
- that the natural variables $\xi_1, \xi_2, \xi_3, \ldots, \xi_k$ should be transformed to coded variables $x_1, x_2, x_3, \ldots, x_k$, then the response mean now is

$$\eta = \mathbf{E}(y) = \mathbf{E}(f(x_1, x_2, x_3, \ldots, x_k)) + \mathbf{E}(\varepsilon) = f(x_1, x_2, x_3, \ldots, x_k).$$

In most of practical settings, low-order polynomials are used for f.

- GS George Box, năm 1987 đã viết rằng "*essentially, all models are wrong, but some are useful*" trong quyển 'Empirical Model-Building and Response Surfaces', p. 424, Wiley; nói lên tầm quan trọng, sự uyển chuyển và hữu ích của mô hình hóa thống kê (khi mô tả thế giới so với mô hình hóa toán học).
- Từ quan điểm của các nhà công nghiệp quá trình, các phương pháp này
  - giúp họ rút giảm sự biến thiên bất định (*variablity*) của quá trình bằng cách chọn được các mức phù hợp của một số yếu tố điều khiển được,
  - theo đó cực tiểu hóa sự biến thiên của hàm đáp ứng - kết xuất đang quan tâm,
  - khi cho phép các yếu tố nhiễu (khó hay không điều khiển được) thay đổi tự do (nhưng không có ảnh quan trọng lên kết xuất)

A Case Study in Statistical Quality Control

We now briefly describe an application of our ideas to the industrial problem of camera tuning. This experiment was carried out by Sanyo DI in HCMC Vietnam.

Description of our case study. In quality engineering, evaluating the precision of product's features is economically necessary.

- OK (good) products that evaluate as NG (malfunctioning): will require unnecessary repairs leading to a cost increase.
- NG products that evaluate as OK: will lead to compensation claims and extra service costs.

The goal is to keep the resolution parameters within certain specification limits, while reducing the time taken for the adjustment.

Setting of experiments

Four controllable factors in our process of resolution adjusting are:

- Tilt of lens & chart to vertical axis (named factor A, unit ^{o}).
- Tilt of lens and chart to horizontal axis (factor B, unit ^{o}). We use an angle meter to measure value of factors A and B.
- Brightness on surface chart (C, unit EV, measured by illuminometer), where

$$EV = \text{ Exposure Value (or Light Value)} = \log_2 \left[\frac{F^2}{T}\right],$$

  where F is Lens iris speed (s), and T is the exposured time (s).
- Distance between Lens and chart (D, unit cm).

The aims of conducting experiments

By Sanyo's industrial standards, there are 13 interest process responses, (i.e. resolution data of picture chart that are measured by computer software), for our multi-response optimization. We define

1. $f(A, B, C, D)$ to be adjustment time of the process;
2. $Y_i = y_i(A, B, C, D)$ to be a regression model for the mean of checking point i, and
3. $\sigma_i = \sigma_i(A, B, C, D)$ to be a regression model for the standard deviation of checking point i, for $i = 1, 2, \ldots, 13$.

Now, our desired goal is:

$$\begin{cases} \text{Max } Y_i, & \text{for each } i = 1, 2, \ldots, 13 \\ \text{Min } \sigma_i(A, B, C, D), & \text{for each } i = 1, \ldots, 13, \text{ and} \\ \text{Min } f(A, B, C, D). & \end{cases}$$

Use of DOE and Computing

Consider that large-than-two factor interactions are negligible, for any mean response Y on the four factors $X_k \in \{A, B, C, D\}$, our model is

$$Y = \beta_0 + \sum_k \beta_k X_k + \sum_{k<l} \beta_{kl} X_k X_l + e.$$

We employ the statistical software R to analyse the data obtained, for all responses (from Y_1, σ_1 to Y_{13}, σ_{13}, and the adjustment time response f). As an illustration, for determining response Y_1, since key aims are to find out which main effects and two-factor interactions are significant, after screening we employ the following R code:

```
LinearModel.1 <- lm(Y1 ~ A+B+C+D+A*B+A*C+A*D+B*D,
data=Design.withresp)
```

A proposed algorithm using factorial designs

However, determining linear models Y_i, σ_i and the adjusting time response f in terms of predictors A, B, C, D is just the first step in the following procedure of 4 steps:

- Step1 Build up regression model of the means for all quantitative responses with replication.
- Step 2 Convert quality objectives (regression model of mean and regression model of standard deviation) to one equation by using capability index. Specifically we calculate process capability indices $CP_{k,i}$ by formula:

$$CP_{k,i} = \frac{\text{specification width}}{\text{process width}} = \text{Min}\left\{\frac{USL_i - Y_i}{3\sigma_i}, \frac{Y_i - LSL_i}{3\sigma_i}\right.$$

in which USL_i is the upper specification limit and LSL_i the lower specification limit of the i-th quality feature.

A proposed algorithm using factorial designs

Step 3 Convert $CP_{k,i}$ to Sigma level using equation:

$$\text{Sigma Level } = 3CP_k + k\sigma$$

with $k = 1.5$ (see below).

Step 4 Solve a non-linear multi objective problem to determine values of A, B, C, D that provides optimal response functions.

6-sigma quality level

- In Step 3, to address typical maximum shifts of the process mean, Sanyo, following accepted industry standards, uses the shift value $\pm 1.5\sigma$ for the process mean [Park].
- When using this shifting ratio, the manufacturer would theoretically achieve 6-sigma quality level, corresponding to the ideal error ratio 3.4 PPM (parts per million) [Park].

====

Sung H. Park, *Six-Sigma for Quality and Productivity Promotion*. Asian Productivity Organization, 1-2-10 Hirakawacho, Chiyoda-ku, Tokyo 102-0093, Japan, pp.12-23 (2003)

Numerical outcomes and discussion

- Experimental data was collected from 10 adjustment runs with a master (i.e. a standard sample camera) at the equilibrium levels $(A_0, B_0, C_0, D_0) = (0, 0, 10, 95)$.
- Some responses were out of the specification range – they are colored red in Figure 1a.
- Noise factors of the industrial manufacturing environment (E and F) had significant impact on the adjustment system.

Numerical outcomes and discussion - 1

- The objective of the case study is to find how factor values (A, B, C, D) other than the equilibrium condition (0, 0, 10, 95) affect to the output, while remaining within specification.
- After optimizing adjustment to the master camera, we found that the parameters

$$(A, B, C, D) = (2.252, 3.878, 8.969, 94.930)$$

gave the resolution values obtained as shown Figure 1b.

Clearly, the outputs are now better than the previous data in Figure 1a. The time of resolution adjustment is now 232.90 seconds.

	Factor				Respose													
Rep#	A(°)	B(°)	C(EV)	D(cm)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Time_Adj
1					69.68	55.73	53.42	53.39	55.38	50.12	50.38	50.13	49.14	46.08	45.32	42.75	39.96	244.10
2					75.06	49.79	48.25	49.91	57.22	46.14	51.15	46.07	43.96	40.02	39.35	44.69	43.30	236.35
3					72.91	52.32	52.51	51.78	52.67	44.08	47.13	49.83	48.20	41.58	40.95	40.06	42.16	248.75
4					75.87	56.78	53.89	57.48	55.35	48.35	45.26	44.75	50.51	47.01	46.34	42.80	46.12	240.17
5	0	0	10	95	68.75	52.92	50.84	48.93	56.23	50.47	50.44	51.90	46.55	43.21	42.61	43.65	38.68	247.35
6					71.08	52.16	49.75	56.59	49.37	44.79	48.73	50.87	44.42	42.36	41.69	36.70	40.33	223.18
7					68.30	54.62	55.22	53.30	55.69	48.08	49.80	43.57	50.92	44.89	44.22	43.08	39.55	238.10
8					74.63	56.29	55.24	53.58	53.04	50.43	49.97	50.38	45.97	46.49	45.82	40.44	44.88	229.78
9					70.74	47.80	49.80	49.98	56.19	52.59	50.55	48.56	49.86	44.08	43.33	43.61	38.93	241.46
10					75.41	52.23	57.23	56.06	51.92	44.67	47.73	50.69	51.94	42.46	41.79	39.39	44.66	226.10
				Spec(≥)	70.00	50.00	50.00	50.00	50.00	45.00	45.00	45.00	45.00	40.00	40.00	40.00	40.00	--
				ULS	75.87	56.78	57.23	57.48	57.22	52.59	51.15	51.90	51.94	47.01	46.34	44.69	46.12	248.75
				LLS	68.30	47.80	48.25	48.93	49.37	44.08	45.26	43.57	43.96	40.02	39.35	36.70	38.68	223.18
				Mean	72.24	53.06	52.62	53.10	54.31	47.97	49.11	48.68	48.15	43.82	43.14	41.72	41.86	237.53

(a)

	Factor				Respose													
Rep#	A(°)	B(°)	C(EV)	D(cm)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Time_Adj
1					70.08	54.23	53.18	52.65	54.68	49.35	49.31	49.25	48.39	45.38	44.69	42.52	41.35	235.30
2					74.35	53.13	52.24	51.48	55.89	47.35	50.16	47.35	45.87	41.36	42.35	41.89	42.35	238.23
3					75.92	51.18	53.13	50.79	53.84	46.73	48.32	48.97	49.16	42.38	41.09	41.38	43.05	239.45
4					74.70	55.45	49.19	56.34	54.38	47.86	49.35	46.39	48.93	46.04	45.72	39.94	45.38	239.34
5	2.252	3.878	8.969	94.93	71.04	51.48	51.09	51.48	54.85	49.39	44.89	50.13	49.35	42.97	43.56	43.28	42.38	239.86
6					70.14	53.28	52.34	54.75	51.78	48.32	49.76	49.89	47.38	43.09	42.63	42.56	41.21	237.22
7					69.87	56.15	54.26	52.38	53.61	47.97	48.65	46.72	50.01	43.76	43.58	41.78	40.56	240.16
8					73.57	53.17	55.78	49.79	51.35	51.06	47.38	51.38	48.39	45.68	44.67	40.58	43.79	236.43
9					71.54	51.75	52.34	52.34	55.49	50.78	49.78	49.45	49.38	44.72	45.89	42.38	43.59	238.75
10					74.31	53.24	55.78	56.06	53.26	48.72	46.35	48.72	50.87	41.89	42.35	41.87	45.08	235.15
				Spec(≥)	70.00	50.00	50.00	50.00	50.00	45.00	45.00	45.00	45.00	40.00	40.00	40.00	40.00	--
				ULS	75.92	56.15	55.78	56.34	55.89	51.06	50.16	51.38	50.87	46.04	45.89	43.28	45.38	240.16
				LLS	69.87	51.18	49.19	49.79	51.35	46.73	44.89	46.39	45.87	41.36	41.09	39.94	40.56	235.15
				Mean	72.55	53.31	52.93	52.81	53.91	48.75	48.40	48.83	48.77	43.73	43.65	41.82	42.87	237.99

(b)

Hình : The actual data set before and after optimizing

Part III.E: Ethical matters in R & D

Discussion with guidances.