Course material

Course:

ENG606 / PHYS 442

Video:

DOE_lesson1_part2

Concepts (extracted from automatically generated subtitles):

Karl pearson. Origin of the design of experiments. Much things. Important persons. Last book of michelle deville. Yield of the fields. Basic elements of statistics. Pierre simon de la place. Francis galton. Small explanation of traffic jams. Complicated interaction. Causes of traffic jams. Mind map of a case. Pearson galton. Data analysis.

to video sequence search
(within ENG606 / PHYS 442.)

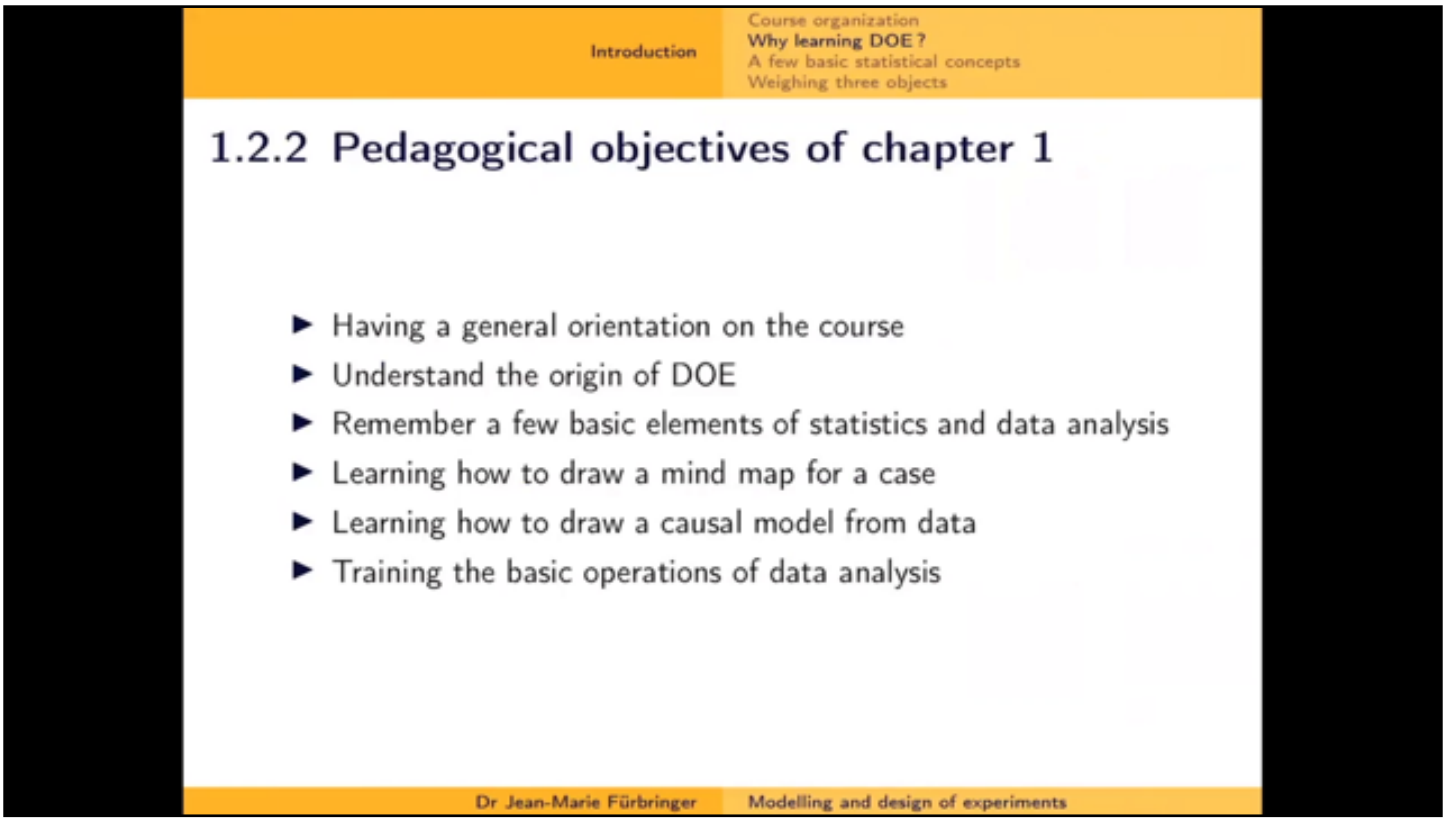
to video

...

notes

summary

0m 0s

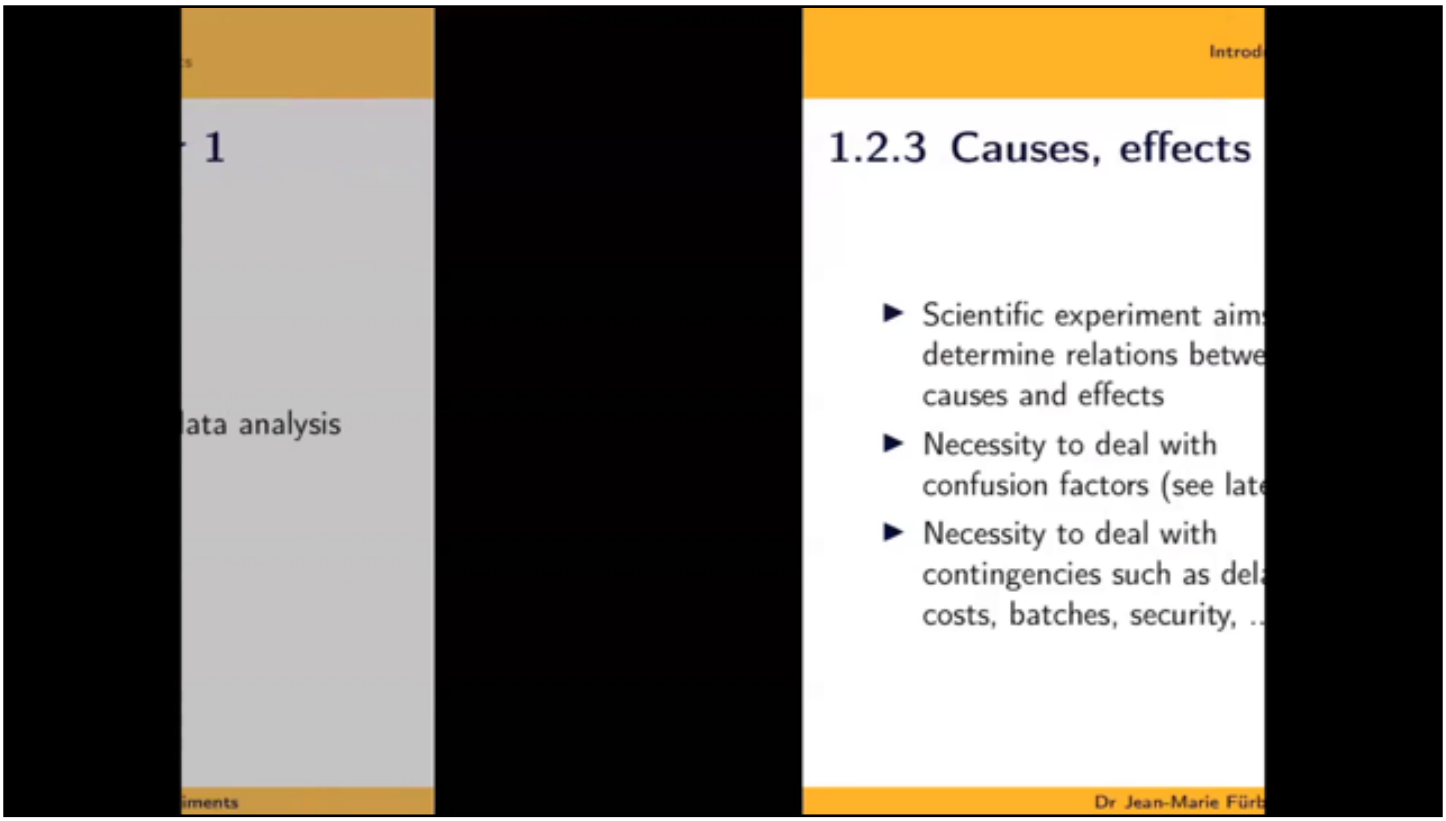


These subtitles have been generated automatically First chapter, what is the pedagogical objective of my chapter?

notes

summary

0m 1s

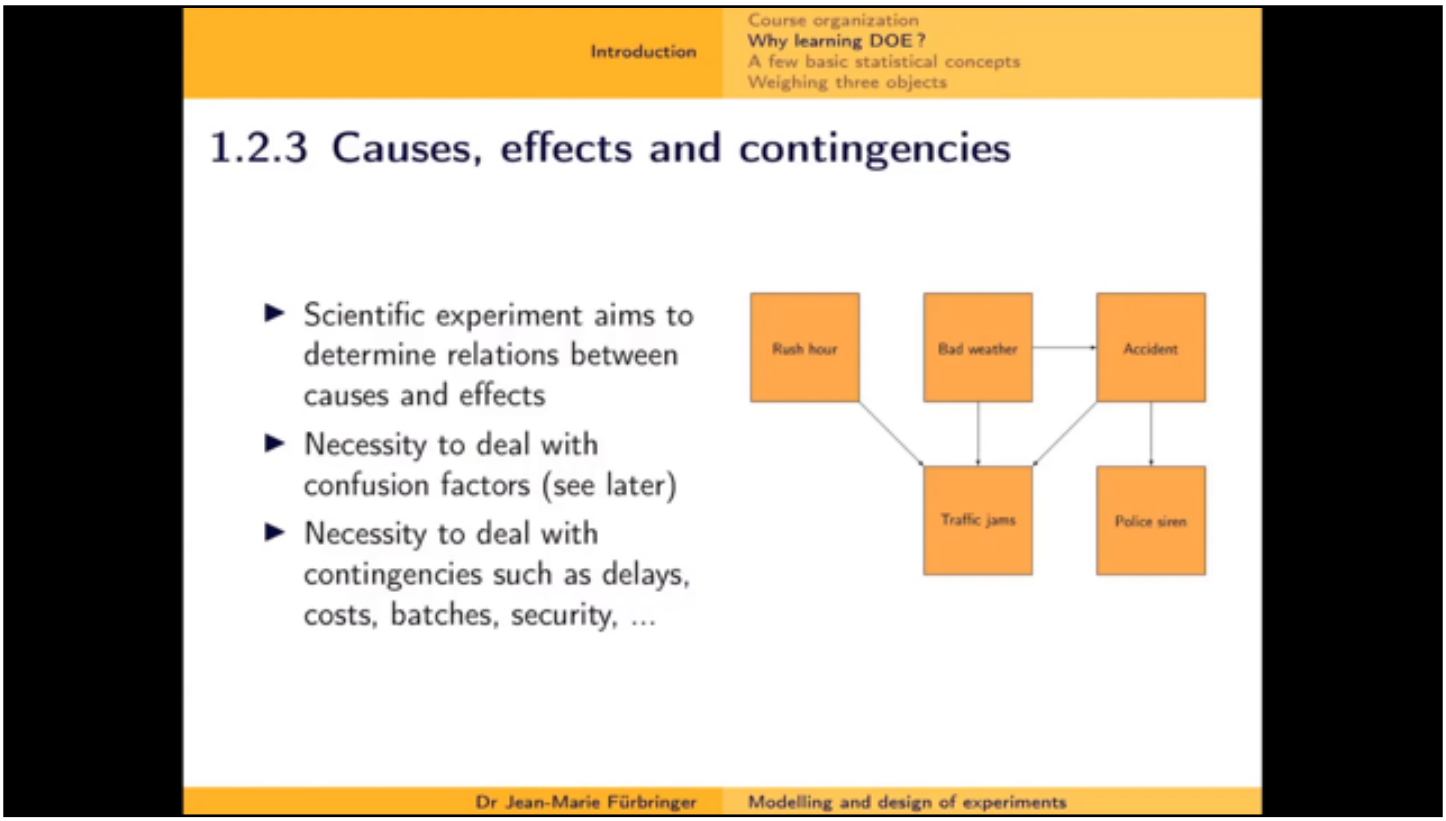


Having a general orientation of the course, so this is done, understand the origin of the design of experiments, remember a few basic elements of statistics and data analysis, learning how to draw a mind map of a case, it will be really either the center of our question because we would like, I don't want only to transmit you techniques, I also would like to transmit you some strategic thinking about experiments. So we need a map for understanding how to work it. I will also talk about causal model. We all have in mind, but sometimes we forget to draw them and to consider them and also training the basic operation in data analysis is more the MATLAB or Python aspect, so how

notes

summary

0m 5s

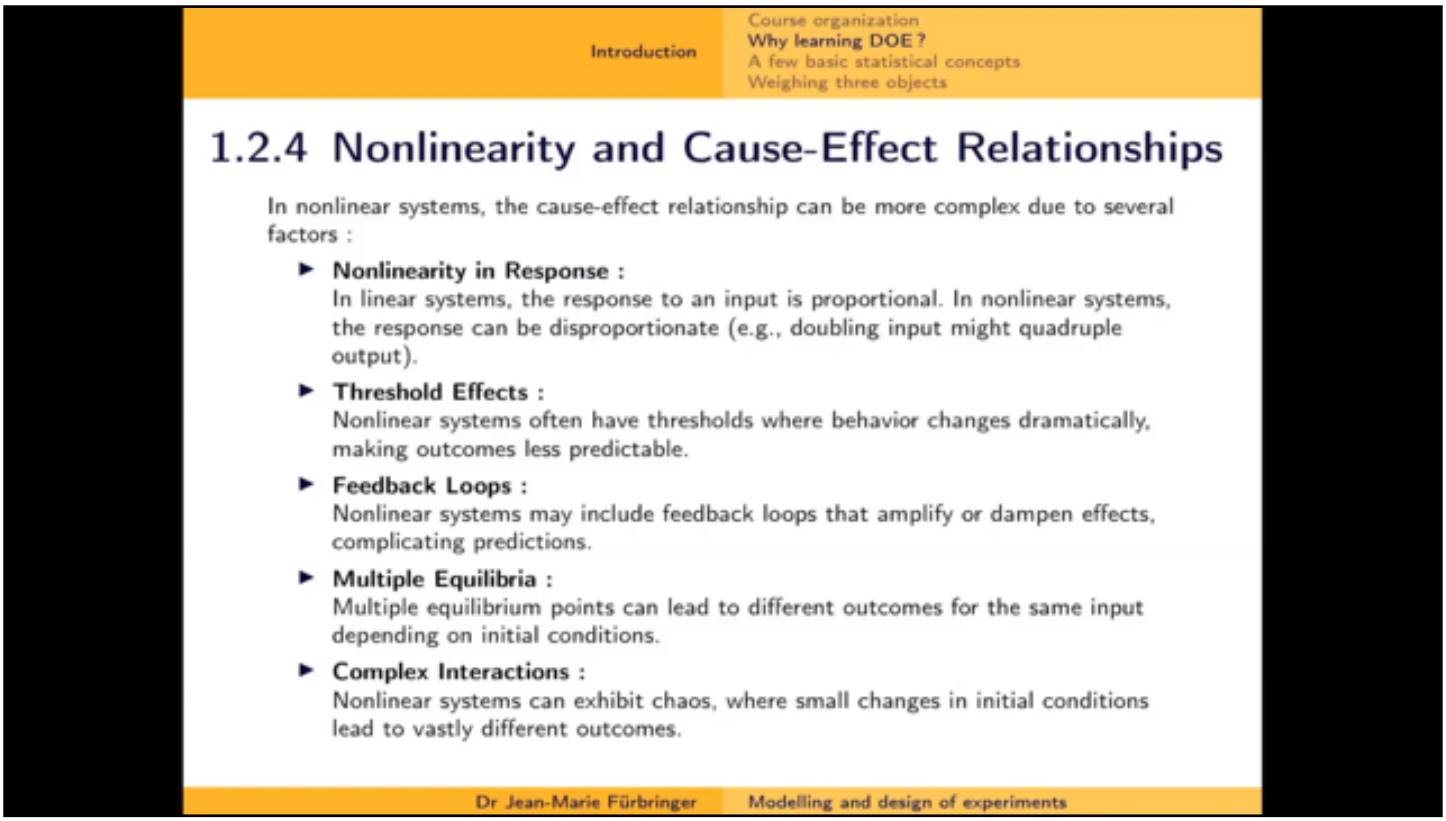


to manage a data base. Okay, so making models, working in engineering, we are all dealing with what I call causals, effects and contingencies. So if you see this very small explanation of traffic jams, we have to think what could be the causes of traffic jams. And some people in engineering, although they work in other in traffic, it could be because it's rush hour, it could be a cause of a traffic jam, it could be because the weather is bad or it could be if you are going back in ballet or going home, it could be an accident. I do actually. But if you are in the highway, it could be an accident. But well, but there are also instructions between your factors. So finally, the accident would be also more probable if you have bad weather. So our factor has a relation between them. And also we can have factors that has all the consequence that only the consequence that we have in mind. So this is what I have in mind when I'm talking about cause and effect, it could be more or less complicated, but making models, it's all the time what we are doing. We would like to obtain something of a product, quality of the product. And what are the factors that are managing that? And we need to make some order and to really understand the influence of the different factors on it. But we are in the real world. We have families, we have colleagues, we have the jets.

notes

summary

1m 9s



So design of experiment is really a general theory that had to solve one of those, some of those problems and deal with them, try to optimize the thing in terms of pure statistical quality, but also taking in consideration the contingencies. I have such a GPT to prepare this slide. Okay. I was reading during the summer, the last book of Michelle Deville, I know people in mechanical engineering, you'll know it was a professor, it's a retired professor of civil engineering, and he said, okay, cause and effect work very well in linear, but non-linear situation like with mechanics, it could be complicated. So I asked Chad GPT, okay, I explain to me why some people are questioning the cause and effect relation in non-linear situation Usually it's not the problem. And the stress-hold effects, this could be interesting. Imagine that you want to evaporate the liquid and you put some eats, so it doesn't start from scratch to evaporate. At least you have to get evaporation temperature depending on the pressure and after it's happened. So you kind of have stress-hold effects After a while, it could be quite complicated. I was talking about traffic. If you are interested by this mobility problems, you know, the big discussion, we have to make more highways, but if we make more highways, we have in fact more traffic, if we have more traffic, more traffic jumps, et cetera. This is a relation of feedback loop.

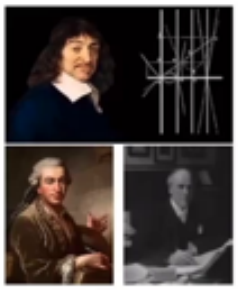
notes

summary

3m 13s

1.2.5 A (very) brief time line of statistics

- Sumer (III millennium BC)
  - Livestock control (list, coding)
  - Prediction of sunrises and sunsets, tides and floods
- Aristotle (384-322 BC)
  - Things that change all the time can not be the objects of science
- René Descartes (1596-1650), Pierre de Fermat (1607-1665) et Blaise Pascal (1623-1662)
  - Theory of probability
- Thomas Bayes (1702-1761) et Pierre Simon de Laplace (1749-1827)
  - Conditional probability, prior vs posterior information
- Francis Galton (1822-1911), Karl Pearson (1857-1936) and Ronald Fisher (1890-1962)
  - Genetics, eugenics, correlations

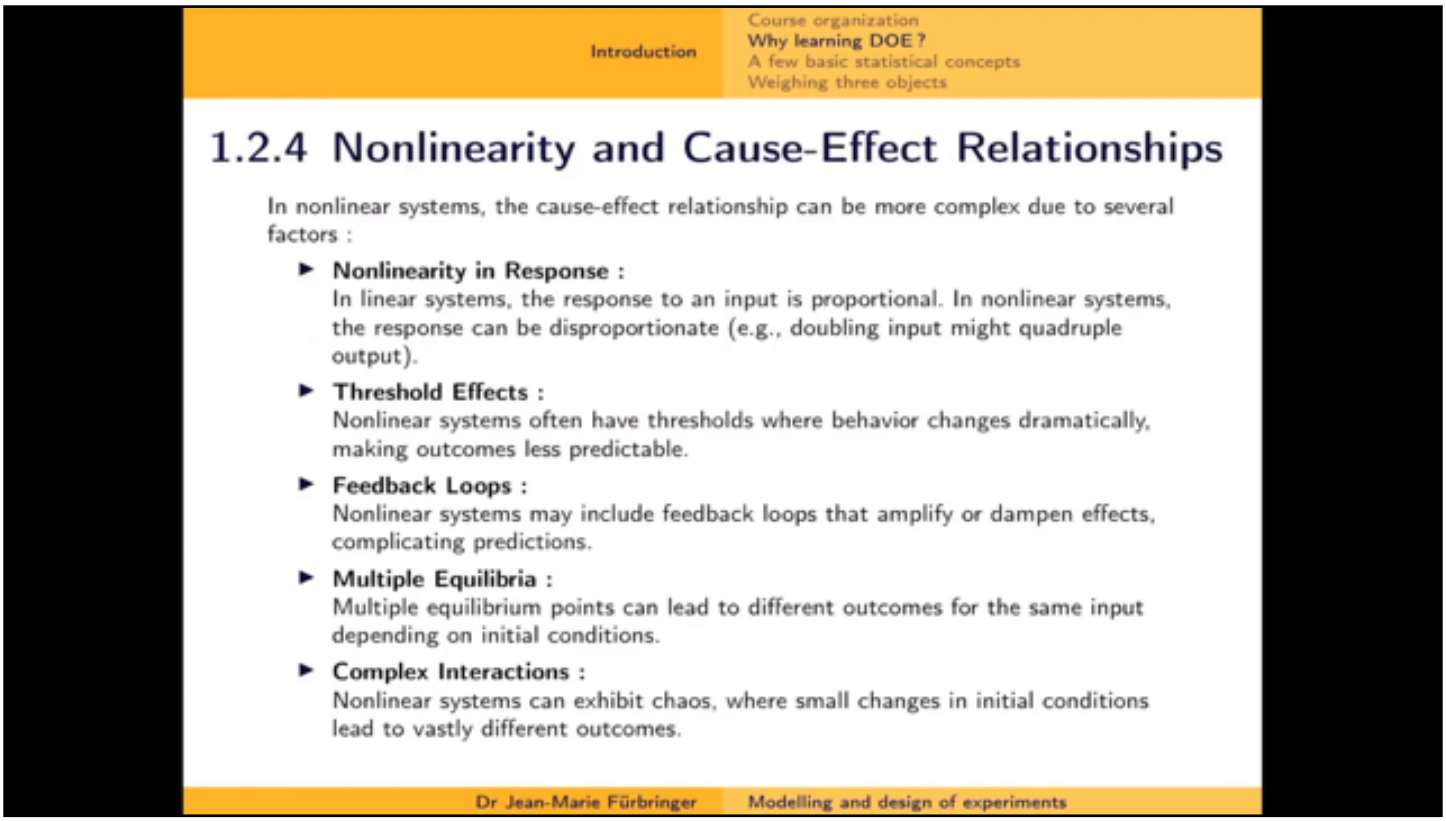


And if you follow a little bit about also the problem. Sustainability, time actually sees a lot of loop of interrelations feedback loops are not easy to present present to people and to act on because we never know if you are really acting on the real factor or in fact, you are just acting on the consequence. In medicine, what is you a small ball that is a part of one of the pin and it could go in any direction and depending on what you have it could go in some way so the equilibria could make also that the cause and effect would be and not so easy to put in evidence and after we can have quite complex instructions in this course we just worked very simple instructions that are very very complicated interaction I'm reading a book from Mr. Ale about the plants and the trees it's incredible in the nature the size of complex a web of cause and effect that happened in the middle of what was first now and so these are some elements yes

notes

summary

6m 25s



But this is a linear and a linear or nonlinear answer.

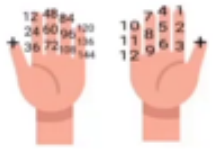
notes

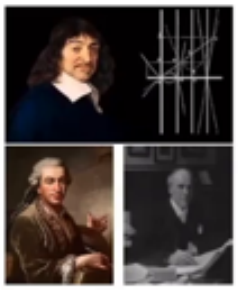
summary

9m 37s

1.2.5 A (very) brief time line of statistics

- Sumer (III millennium BC)
  - Livestock control (list, coding)
  - Prediction of sunrises and sunsets, tides and floods
- Aristotle (384-322 BC)
  - Things that change all the time can not be the objects of science
- René Descartes (1596-1650), Pierre de Fermat (1607-1665) et Blaise Pascal (1623-1662)
  - Theory of probability
- Thomas Bayes (1702-1761) et Pierre Simon de Laplace (1749-1827)
  - Conditional probability, prior vs posterior information
- Francis Galton (1822-1911), Karl Pearson (1857-1936) and Ronald Fisher (1890-1962)
  - Genetics, eugenics, correlations





But we have linear systems and linear systems are systems that you can represent with a matrix and some coefficient. So we will stay in this world. The nonlinear, very complex nonlinear systems, we will not see that. It exists, design of experiment for nonlinear systems. This is also a design of experiment for biasing approach, etc. So we take this order, we will stay the situation where we treat linear system. Now already quite a lot of sense. Okay, a very, very brief timeline of statistics. I don't know what is your general culture in the domain. So statistics is quite... Statistics is a subject that have been developed quite late, 17th century. But people have made statistics without saying it for a long time. So I just quoting the Sumerian civilization, 3 millennium before our era. People used to check the birth rate in their livestock. We are for the yield of the fields. They also look sunrise, sunset, tides, etc. So we are making statistics in fact for a long time, but without saying it. So for illustrating that, just as a joke or just as an element, I don't know if you know somebody, understand what is this? What this represents? So Sumerians have a basic of 12. Because you can count very easily big numbers with your hands. So you count your knees. One, two, three, four, five, six, seven, eight, nine, 10, 11, 12. And after you continue. So it's a way of calculating with the hands. Okay, so in the antiquity, the Greek didn't really appreciate so much things changing. So they said, we pretened... Aristotle has said, but I don't know. But pretending things that change all the time cannot be object of science. So it's why also in this reflection, mathematical, geometrical reflections, they come from antiquity. We see nothing about statistics. Very surely they do something. They don't make concepts.

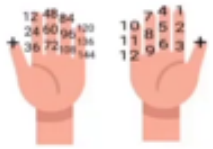
notes

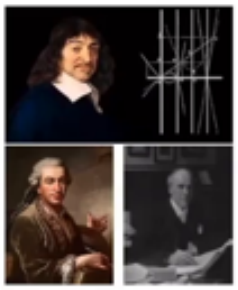
summary

9m 45s

1.2.5 A (very) brief time line of statistics

- Sumer (III millennium BC)
  - Livestock control (list, coding)
  - Prediction of sunrises and sunsets, tides and floods
- Aristotle (384-322 BC)
  - Things that change all the time can not be the objects of science
- René Descartes (1596-1650), Pierre de Fermat (1607-1665) et Blaise Pascal (1623-1662)
  - Theory of probability
- Thomas Bayes (1702-1761) et Pierre Simon de Laplace (1749-1827)
  - Conditional probability, prior vs posterior information
- Francis Galton (1822-1911), Karl Pearson (1857-1936) and Ronald Fisher (1890-1962)
  - Genetics, eugenics, correlations





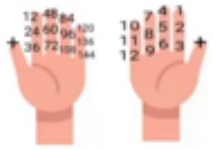
They don't make cell rent on the earth. And they let it outside of what they call science. Since change in the end of the 16th and 17th century, especially between Fermat, Descartes and Blaise Pascal, they start to make a theory about dice and about cards. Two important persons, Thomas Pius, makes the Piusian scarring of Pius. So it is an empty conditional probability and all the same. But it was not the scientists, it was the priests. I don't know what I'm looking at. It was in England. I don't know what I'm looking at. I'm looking at all the same. But his interest was intellectual interest. So he made his theorem, published it in the New York Times. After he died by a friend, and it stayed there. Nobody did nothing with that. That's one very polymath person, Pierre Simon de la Place, is the one who changed our lives, because he is the one who have introduced statistics in science. Before, when people make measurements, they decided which one is the best measurement and they left the parts to rest. In the revolution, people who wanted to simplify the measurements, the metric distance, weight, surfaces, and decided to develop the measurements. So they make measurements from Dunkirk to... Ah, in Spain. Barcelona, because it's quite a long same line. But some battle is the border of Spain. And the problem is they have to continue the triangulation from another point, and when they try to put the two triangulations, it doesn't work. They try to find the solution, and finally invent the linear regression. And before, it was not like that to make measurements. And you go, yeah, it was nice, it's good. It was, so this was the best measurement I could... And Pierre Simon de la Place discovers a theorem of bi and integrates statistics within science. And so the first statistic was biogenic.

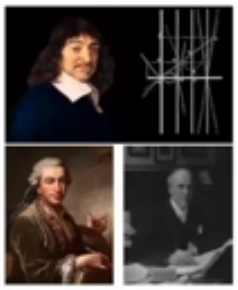
notes

summary

1.2.5 A (very) brief time line of statistics

- Sumer (III millennium BC)
  - Livestock control (list, coding)
  - Prediction of sunrises and sunsets, tides and floods
- Aristotle (384-322 BC)
  - Things that change all the time can not be the objects of science
- René Descartes (1596-1650), Pierre de Fermat (1607-1665) et Blaise Pascal (1623-1662)
  - Theory of probability
- Thomas Bayes (1702-1761) et Pierre Simon de Laplace (1749-1827)
  - Conditional probability, prior vs posterior information
- Francis Galton (1822-1911), Karl Pearson (1857-1936) and Ronald Fisher (1890-1962)
  - Genetics, eugenics, correlations





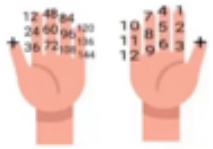
And the problem, I don't know if you are familiar with the biogenic statistics, we need a prior probability, you need to start with something, and after you have information coming, and that seems very complicated for the colleague of bi. So when he died, he put everything in the rubbish bin and starts what we have today, most of the time the frequentist way of looking at statistics. And after we rediscovered, I don't know exactly when, we rediscovered the biogenic statistics. After we arrived to design a website, the modern statistic, like the modern biology and the modern genetics have been invented, set up by 3%. Francis Galton, Karl Pearson, and Ronald Fischer. Really, they quite made quite everything. They did linear regression mathematically. They made mathematical theory of probability and statistics. They invented Ronald Fischer's design of experiments, the ANOVA person developed the concept of correlations. They invented quite everything. Pearson Galton is a very important person. It was in a few of Darwin, and was in fact those three persons differently, but to have the same interest understanding of the genetics, the inheritance, etc. So Galton has a few of Darwin to prove that we see the family. You have very important intelligent persons. We want to get some of them. Yeah, the geniuses are coming from the same family, etc. And so if you read me into the life of this, I think that Francis arrived to something quite different because of the geniuses. These three persons were very convinced first that the British race was the race, that the British race was the best of the universe and after you have the other. Well, perhaps you look, Karl Pearson. Karl Pearson was baptized by his parents, Karl, C-A-R-L. It's too much time in Germany. I hope you are German among us, but you will not be offended by what I will say. And so

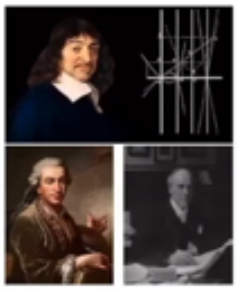
notes

summary

1.2.5 A (very) brief time line of statistics

- Sumer (III millennium BC)
  - Livestock control (list, coding)
  - Prediction of sunrises and sunsets, tides and floods
- Aristotle (384-322 BC)
  - Things that change all the time can not be the objects of science
- René Descartes (1596-1650), Pierre de Fermat (1607-1665) et Blaise Pascal (1623-1662)
  - Theory of probability
- Thomas Bayes (1702-1761) et Pierre Simon de Laplace (1749-1827)
  - Conditional probability, prior vs posterior information
- Francis Galton (1822-1911), Karl Pearson (1857-1936) and Ronald Fisher (1890-1962)
  - Genetics, eugenics, correlations

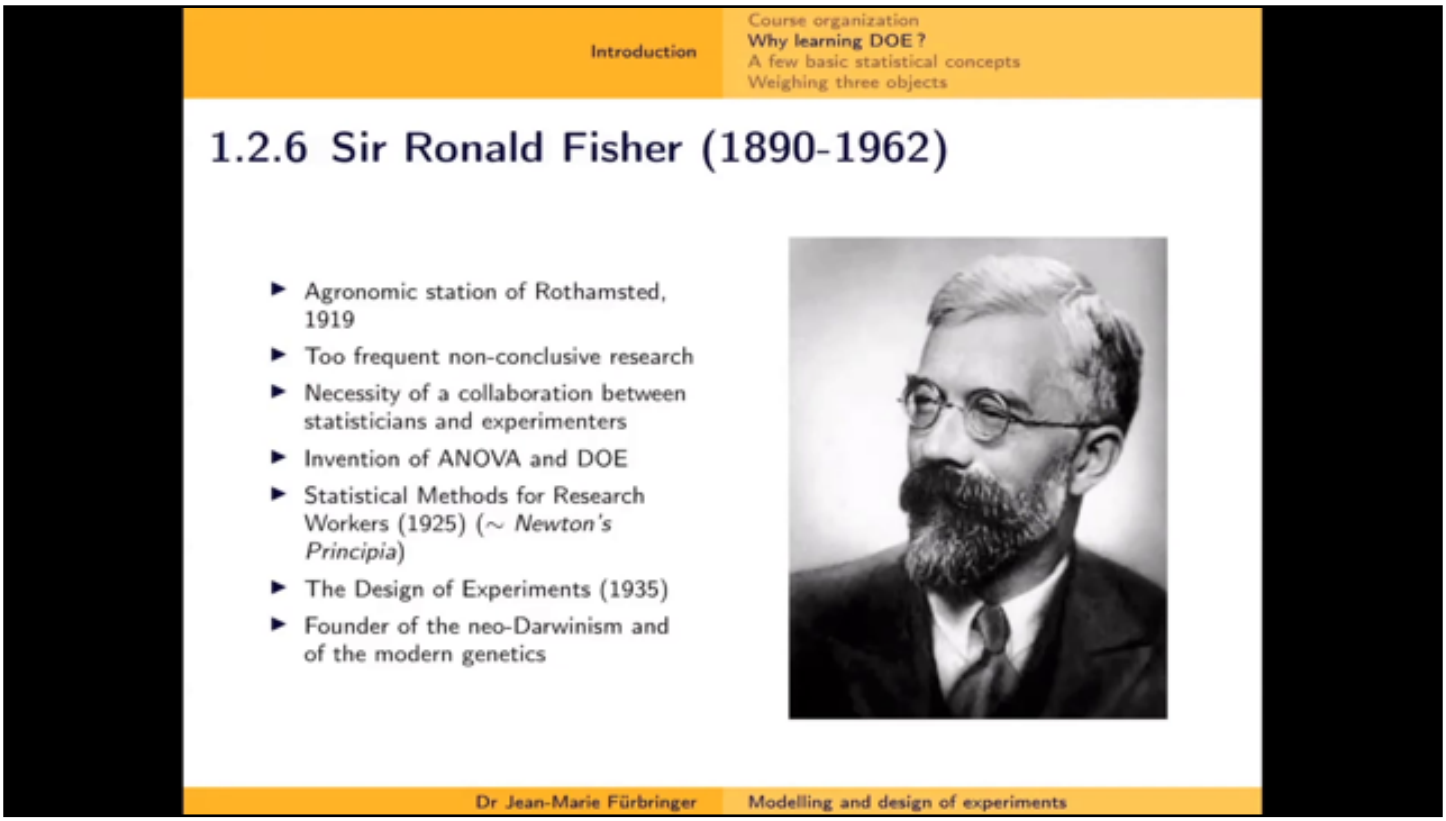




he decided to change the sea by a character. Those people have developed something we call the scientific races. They founded Galton, probably, and it's a very important term in terms of genius for improving the race. There were people who can say, don't accept Jewish people in England. We are better than just British. People say, okay, let's intelligent people to get married, but people are not so much intelligent. Try to do things for them, etc. They didn't make concentration camps. They were like that, but they have quite strange idea about genetics and about genes.

notes

summary



So today, it's quite complicated in a written, I don't know if you hear, you read something. Pearson was from the UCL University London. There are some buildings. They were called in, and the director said, I will never teach again in this building. No, we are re-evaluating. Nevertheless, there were even if they had some strange ideas, and at that time, a lot of people were having quite the same idea. Scientifically, I found a sense of behavior because the genius was pretending to be also in science, but statistics really made quite average. It really made the basis. For physicists, I was looking also at the history about Carl Pearson. Carl Pearson's right to write an important book is the Grammar of Science. I'm a young student of a jury called Albert Einstein. I read this book, I appreciate it a lot. And if you look in this book, there are a lot of ideas of the relativity. So, Carl Pearson developed a few ideas, not in physics, but let's say in philosophy. It seems that it's him who's saying, what's happening to travel with the velocity of light and things like that? So it was quite very, very interesting questions. Design of experiment, admit inventors, by this person, Ronald Fisher. He was a PhD student. I could continue his career in academic, but he wanted to go in the real world. So he got a job in the agronomic station in Wolftham Stats, in England, in the 1920s. And so his job was to look at the work by his colleagues, agronomists, biologists, and to make the end of the report, checking the confidence interval, checking the conclusions, etc. And unfortunately, many times, he said, no, the data is not really proving what to pretend. It's not so clear. The confidence interval are too high. And imagine that in agronomy, it's not the experiment during one afternoon.

notes

summary

19m 25s

1.2.6 Sir Ronald Fisher (1890-1962)

- Agronomic station of Rothamsted, 1919
- Too frequent non-conclusive research
- Necessity of a collaboration between statisticians and experimenters
- Invention of ANOVA and DOE
- Statistical Methods for Research Workers (1925) (∼ *Newton's Principia*)
- The Design of Experiments (1935)
- Founder of the neo-Darwinism and of the modern genetics

It's usually experiment over many years, five years, six years, ten years. I did an experiment for 20 or more years on how to get the best potatoes with which type of soil, which type of salt, which type of water, which type of... ...angre. So very difficult situation when people have worked for many years and you say, sorry, the job is not clear. We cannot go through it on what measurement you have made. So two-frequent, non-conclusive research. So you really think to that and want to propose something for solving that. And to do things before the experiment to be sure that in the end, you are able to conclude. You cannot predict what will be the answer. Sometimes when you are looking for money, for research, you have the impression that the fund is asking you the result already. But not that. But just to be sure that in the end, the confidence interval of your measurement will be sufficient for bringing the conclusion. And for that, he invented the analysis of variance. I know that, but you have... You've got it. I don't know if you are using it. It was one chapter in the crystal statistics. But it is very important. It is in fact the scientific truth. And I'm sometimes surprised. I will not make the test today. But if you are, people who are in the scientific world here, the Swiss Federal Institute of Technology, and it's impressive how many people are not very clear about what is the scientific truth, for some reason. Okay, very interesting. We'll spend a little bit of time to understand better and to use it. And he invented DOA, Design of Experiences. We used some mathematical things. It was not already in mass, but he used it for that. He wrote a book published in 1925. Statistical Method for Research Worker. And this book is sometimes

notes

summary

1.2.6 Sir Ronald Fisher (1890-1962)

- Agronomic station of Rothamsted, 1919
- Too frequent non-conclusive research
- Necessity of a collaboration between statisticians and experimenters
- Invention of ANOVA and DOE
- Statistical Methods for Research Workers (1925) ($\sim$ *Newton's Principia*)
- The Design of Experiments (1935)
- Founder of the neo-Darwinism and of the modern genetics

compared to the book by Lupin, the Transigna of Luton. Really, it was a change in the amount of science, in the world of science. Already Pearson, but Pearson and him are not really friends. But... The Pearson book was a grammar of science. More generally, more field fields, because I think there are more epistemology and design of experiments. So in the 20s, the 30s, the world of science, as we know it today, was really developing a lot. And he worked on other things. So I mentioned it already. He worked on Neo Darwin. So Darwin has, which is a book about the origin of species, has developed this way that biology is evolving. That it has not integrated the gene. The gene was just a discovery of the turn of the century. So they integrate with also Pearson and Galton. They integrate this knowledge within the evolution of what we call the Neo Darwinism. They reform the evolution, the theory of evolution, taking into account the genes. And the model of genetics, sometimes is more seen as a genetician, one of the fathers of the genetic, than the father of the statistic.

notes

summary