EE 590
Scientific Research Methods and Ethics for Engineers

Week 2 Course Notes

Bilge Karaçalı, PhD

Topics

- Scientific body of knowledge
  - Disciplines
  - Scientific knowledge
- Falsifiability
  - Science vs. non-science
- Hypothesis testing
  - Weighing empirical evidence towards a claim

Science

- Objective: to form a collection of **consistent and verifiable claims** on the nature of things
  - Q: What things?
- Consistent and verifiable claims add to the scientific body of **knowledge**
  - History of science begins with history
    - Q: Why?

Scientific Disciplines

- Basic scientific disciplines:
  - Mathematics
  - Physics
  - Biology
  - Chemistry
- Applied scientific disciplines:
  - Engineering
  - Astronomy
  - Economics
  - …

Scientific Knowledge

- Hypothesis
  - Represents a specific **claim** on a specific correlation or cause and effect relationship
  - Is subject to extensive experimental validation
- Model
  - Describes the underlying working of a topic of interest through an **explanatory mechanism** when the corresponding hypothesis is known to be valid
- Theory and law
  - Represents a collection of hypotheses that have been **extensively validated** in experimental tests

Experimental Validation

- In science, all claims are subject to testing
  - Suppose a claim is made according to observations
    - most of the time in the form of a prediction:

    "If {such and such conditions are met}, then {such and such results will materialize}."
  - The claim will be evaluated in terms of whether the predictions hold true
    - Every time the conditions are met, the expectations must be observed
    - If there are instances where the conditions do not produce the expectations, the claim is **falsified**

Falsifiability

- Scientific body of knowledge is a collection of **falsifiable** claims
  - Falsifiability is the basic requirement for a scientific claim
  - It implies that a claim is scientific if and only if it provides ways in which it can be discredited
    - The claim must identify the conditions under which the validity of the claim will be compromised
      - If the conditions are met and the expected results are not materialized, the claim will be invalidated
  - Testing of a given scientific claim is carried out in **controlled experiments**
    - If a claim is not testable, then it cannot be scientific
  - Q: What is the difference between astronomy and astrology?

Case Study: The *Secret* to Success

- Observation:

All successful people have wanted to accomplish their goals in life very much.

- Claim:

If {one wants to achieve their goals in life very much}, then {they become successful}.

- Controlled experiment:
  1. Interview many people asking them if they wanted to achieve their goals in life very much.
  2. Ask also if they have become successful.
- Result and conclusion:

??

The Method

- Given that you are interested in a topic
  - Growth of crops in dry climates
  - Relationship between age and the onset of Alzheimer’s Disease
  - Mechanism linking certain environmental pollutants to cancer
  - …
- Combining factual information from various observations and the current scientific knowledge, you form a **hypothesis**
- You set up a series of **controlled experiments** to evaluate the validity of the hypothesis
  - Every hypothesis involves **predictions** on events that will occur whenever a series of conditions are met
  - The experiments create the **conditions** and allow making observations on the predictions
- Based on the results of the experiments, either **establish or reject the hypothesis**
  - Rejecting a hypothesis means it will either be revised or completely scrapped

Example

- Observation:

Schizophrenia patients exhibit reduced cognitive abilities compared to healthy individuals.

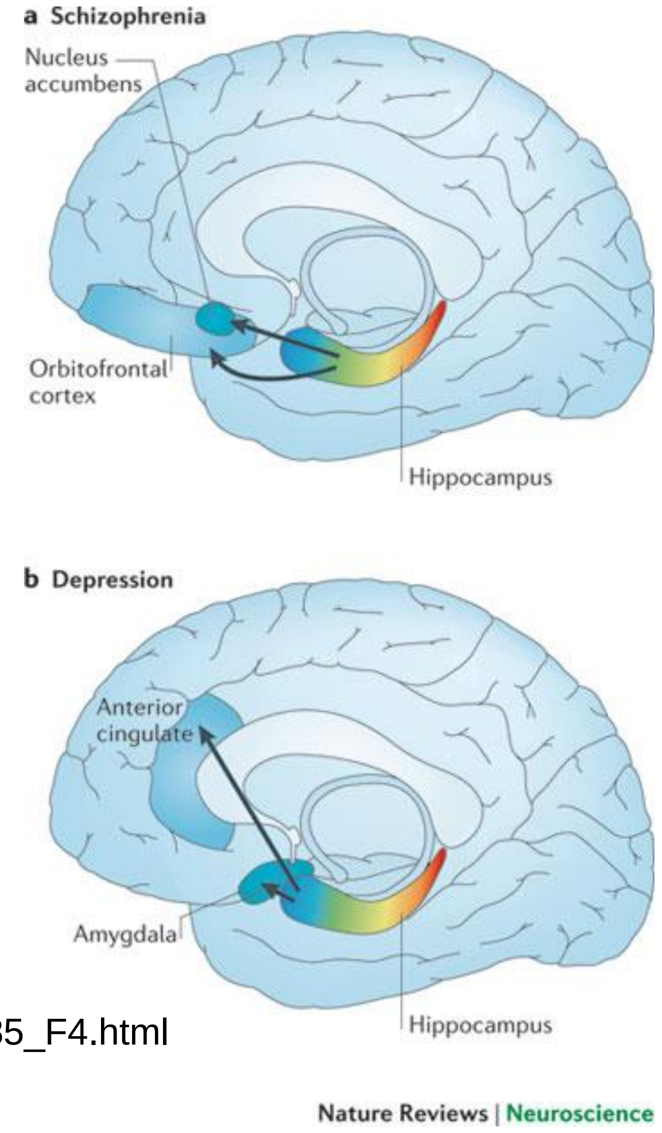
Cognitive capacity is linked to the amount of brain tissue available for cognitive tasks.

A reduced cognitive ability may be indicative of a smaller brain volume in brain regions responsible for cognitive function.

Hippocampus is such a region.

- Claim:

If {a person is schizophrenic}, then {the volume of their hippocampus would be smaller}.



Source: http://www.nature.com/nrn/journal/v12/n10/fig_tab/nrn3085_F4.html

Example (cont'd)

- Controlled experiment:
  - Sign up an equal number of schizophrenia patients and healthy control individuals to the study
  - Acquire 3D MR images of their brains
  - Use computational algorithms to measure the volume of their hippocampi
  - Carry out a **statistical comparison** of the hippocampal volumes between the schizophrenics and the healthy controls against a predetermined **statistical significance level**
- Result and conclusion :

On the average, the hippocampal volumes do **differ** between schizophrenics and healthy individuals. The lesser amount of brain tissue indicates a degraded capacity in cognitive function.

On the other hand, this **correlation** does not indicate **causality** and by no means dooms the individuals with a smaller hippocampus to schizophrenia.

Statistical Hypothesis Testing

- Suppose you have made a lot of measurements divided into two disjoint sets
  - Some measurements made under a given condition, and others under a complementary conditions
    - Hippocampal volumes
      - Some from schizophrenia patients, others from healthy individuals
    - Calculus grades of freshman students
      - Some from engineering, others form humanities
    - …
- The objective is to determine if these two sets are different from each other
  - Note that all numeric measurements will be different from each other unless they are from a finite/discrete number set
  - Comparison can be made in terms of the underlying distributions of the measurements under the two conditions
    - The (probability) distribution of measurements when the first condition is true
    - The (probability) distribution of measurements when the second condition is true

Statistical Hypothesis Testing

- Example: Newborn weights
  - Observation:
    - In humans, on the average, men are heavier than women
    - The question of interest is whether such a difference exists at birth between baby boys and baby girls
  - Procedure:
    - Collect newborn weight data for a certain period of time
    - Compare the weights of the boys to the weights of the girls using hypothesis testing to see if there is a statistically significant difference between the average weights

Source: http://i.telegraph.co.uk/multimedia/archive/02380/babies_2380929b.jpg

Statistical Hypothesis Testing

- Example (continued):
  - Data

  A record total of 44 babies were born in a 24-hour period at a hospital in Australia. The dataset contains the times of birth, genders and weights of these babies.

  Source: http://www.amstat.org/publications/jse/jse_data_archive.htm

Time	Gender (1:g, 2:b)	weight (gr)	Time	Gender (1:g, 2:b)	weight (gr)	Time	Gender (1:g, 2:b)	weight (gr)	Time	Gender (1:g, 2:b)	weight (gr)
0:05	1	3837	8:12	2	3294	14:07	1	3480	19:47	2	3630
1:04	1	3334	8:14	1	2576	14:33	1	3116	19:49	2	3406
1:18	2	3554	9:09	1	3208	14:46	1	3428	19:51	2	3402
1:55	2	3838	10:35	2	3521	15:14	2	3783	20:10	1	3500
2:57	2	3625	10:49	1	3746	16:31	2	3345	20:37	2	3736
4:05	1	2208	10:53	1	3523	16:57	2	3034	20:51	2	3370
4:07	1	1745	11:33	2	2902	17:42	1	2184	21:04	2	2121
4:22	2	2846	12:09	2	2635	18:07	2	3300	21:23	2	3150
4:31	2	3166	12:56	2	3920	18:25	1	2383	22:17	1	3866
7:08	2	3520	13:05	2	3690	18:54	2	3428	23:27	1	3542
7:35	2	3380	14:06	1	3430	19:09	2	4162	23:55	1	3278

Statistical Hypothesis Testing

- Example (continued):
  - The hypothesis to be tested:
    - **The null hypothesis**:

      average weight of newborn girls ≥ average weight of newborn boys
    - **The complementary hypothesis**:

      average weight of newborn girls < average weight of newborn boys
  - The *t*-test procedure:
    - Calculate the sample averages and standard deviations for the baby boy weights and the baby girl weights
    - Calculate the ***T* statistic** (as defined by the statistics books)
    - Determine the ***P* value** for the calculated *T* statistic
    - If the *P* value is smaller than a statistical significance threshold, say 0.05,

    ➔Reject the null hypothesis (Cheers!! Yippee! Etc.)

    Otherwise,

    ➔The test is inconclusive ☹

the uninteresting hypothesis that we expect will be rejected

the explanation that we are rooting for

Statistical Hypothesis Testing

- Example (continued):
  - Results:
    - Average weight of 18 newborn baby girls is 3132.44gr
    - Average weight of 26 newborn baby boys is 3375.31gr
    - The difference is 242.87gr in favor of the boys
    - However, the *t*-test determines a *P* value of 0.0676 for this difference
  - Conclusion:
    - The test reveals that such a difference is to be expected in 6.76 of 100 cases (about 1 in 15) by pure chance under the null hypothesis
      - that we do not think is true
    - This is a small number, but it is <u>not small enough</u>!
      - Had to be smaller than 0.05 for us to reject the null hypothesis
  - What to do now:
    - Collect more data for greater **statistical power**

Statistical Correlation and Causation

- Science is all about determining the reasons for the observed phenomena
  - Hypotheses act as explanations for the observations
- This necessitates determining relationships between the observations and possible causes through a statistical **correlation** analysis
- The problem here is that while good correlation does indicate the existence of some sort of a link between the observation and the possible cause, it does not indicate **causality**!
  - If events A and B are correlated,
    - A may be causing B, or
    - B may be causing A, or
    - A and B may both be caused by an entirely different event C

Statistical Correlation and Causation

- Example: Busy intersections
  - Observation:
    - Every time a bus driver approaches a busy intersection, he sees that there is a traffic police officer trying to guide the traffic
    - He then forms an opinion that the cause of the traffic jam is the police officer
  - Statistical correlation analysis:
    - Every time there is a traffic jam, a police officer is observed at the junction
    - In rare instances when the traffic flows, no officer is in sight
    - This indicates good correlation between the presence of a police officer and the traffic jam

Q: So, is the bus driver right? Is it the police officer causing the traffic jam?

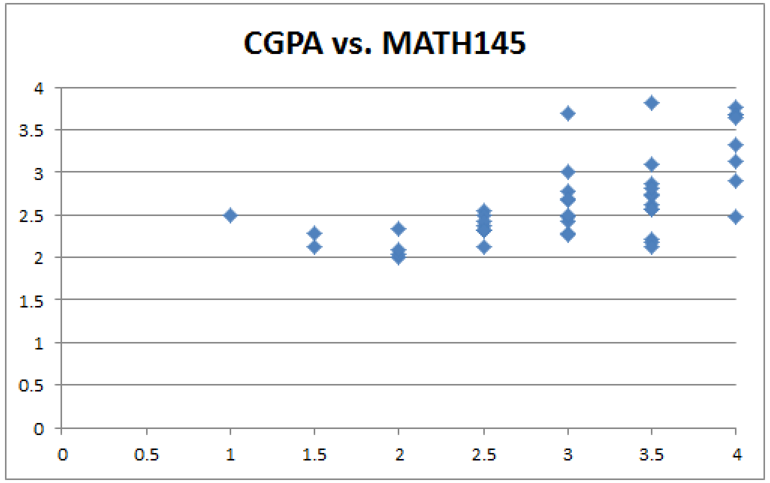
Statistical Correlation and Causation

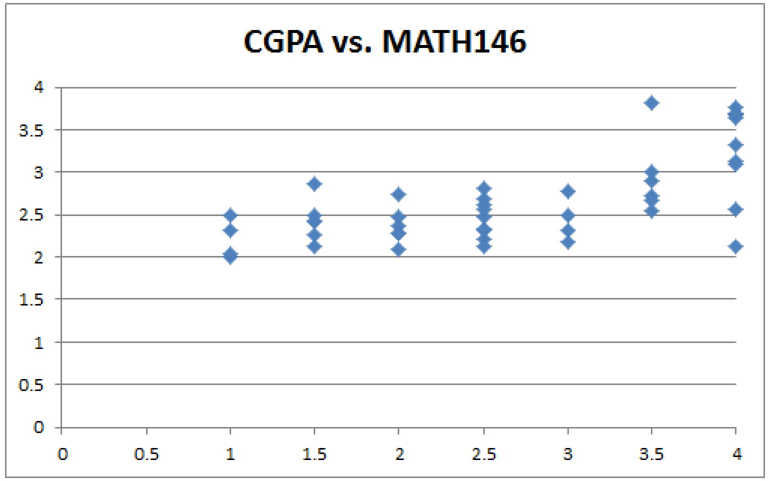
- Another example:
  - Consider the statistical relationships between

    the passing grade from MATH145 vs. graduation CGPA

    and

    the passing grade from MATH146 vs. graduation CGPA
    - The correlation between the first pair is 0.614362257
    - The correlation between the second pair is 0.660799039
  - Observations:
    - Generally speaking, the higher the MATH145 and MATH146 grades, the higher the graduation CGPA
      - Though the 2.00-2.50 graduates may have failed these courses the first time they took them
    - The correlation is better for MATH146
      - Probably because some students recover from the culture shock in their second semester





Summary

- The scientific method rests on
  - Observations on the nature of things
  - Hypotheses formulated to capture the correlation or cause and effect relationships
  - Models that explain inner workings based on validated hypotheses
  - Theories and laws formed by the collection of hypotheses related to a topic of interest that have survived all experimental tests
- In order to be classified as scientific, all explanatory processes must follow the scientific method
  - Everything in science is subject to validation by testing
  - If predictions are not testable, they are not scientific
    - Testability most often implies measurability – cannot test what cannot be measured