Human Communication
Methods for studying communication phenomena

Role of Experiments

Experimental questions are defined in order to help answer research questions.
To be useful, they should be: specific, tightly focused, preferably quantitative, and the answer should be the same whoever does the experiment.
Science proceeds by:
- selecting research questions,
- devising hypotheses (i.e. possible answers to these questions), and
- experimentally seeking evidence to support or refute the hypotheses.

What research questions are interesting, and the general class of methods appropriate for answering them, are defined by a scientific paradigm.

Hypotheses

When reading scientific papers some hypotheses may seem obvious, but are actually tentative conclusions not yet shown to be wrong.
Some endure for hundreds of years without contradiction;
- does not prove their truth, but can be confident they work in most situations we are likely to encounter.

Where do hypotheses come from - not obvious once a research question is posed how scientists find them?
- they guess, make them up, use their imaginations, mess around for a while to see what's going on, or just have make accidental discoveries…..

Scientific process or method deals with weeding out the incorrect hypotheses --- we are free to find hypotheses where we can.

Two broad categories of experiments

Exploratory studies: not certain what you are looking for
- generate data which raises interesting questions,
- suggest relationships between aspects of the theory being investigated,
- allow important parameters to be quantified.

Results are:
- preliminary models of what causes what, or what influences what,
- more precise questions to ask to narrow further the range of possibilities;
- indications of what aspects of the behaviour of humans (or system) are interesting and of where to look next.

Confirmatory Studies

Try to answer a carefully formulated experimental question relevant to your research question.

Confirmatory experiments follow from exploration.
- precise and specific question formulated based on results of exploratory work and hypotheses they suggest;
- experiment is designed to answer that question;
- serious thought given to whether answer provided by the experiment is unique (may be more than one explanation for a set of results)

The best-designed experiments produce results consistent with just one explanation.

Manipulation Experiments

Another useful way of categorising experiments, independent of the exploratory-confirmatory division, is as observation or manipulation experiments.

Manipulation experiments
- look at influence of some factor on output
- proceed by altering the factor in a controlled way
- measures the resulting changes of the output.

The factor is manipulated by the experimenter.

Observation Experiments

Observation experiments: when the factor cannot be manipulated, when participants are people and it is impossible or unethical to manipulate them in the necessary way

e.g. if believe exposure to language through parents affects development, could remove exposure and isolate child

In these experiments:

- population is divided into groups using the factor(s) of interest
- each group's members have the same combination of factors (other criteria for selection may apply)
- look for differences in the output between the groups.

These differences, we hope, are the result of the groups' differing constitution.

Attempting to disprove hypotheses

Formulate a precise experimental question or hypothesis

Test whether evidence supports the hypothesis or not

e.g. children born Oct-Dec show more advanced language development than those born Jan -Mar;

Or STANDUP can improve social communication skills of children with complex communication needs (CCN)

Design an experiment to disprove the hypothesis

- positive result could be caused by something not thought of
  e.g. children born Oct-Dec start school sooner
- but a single negative results disproves the hypothesis

This means finding a way to answer the question:

"Are measurements of X and Y related?"

When to use which?

Observation experiments:

- Necessary when cannot directly manipulate X
- Group subjects based on measurement of X

e.g. 2 groups, 1 of children born Jan -Mar and 1 of those born Oct - Dec; see variation in language development (use standardised tests)

Manipulation experiments:

- When factor of interest is directly manipulable

e.g. two groups of children with CCN; one uses STANDUP and one uses other tools; assess social communication before and after

Examine whether there is any relationship (e.g. correlation) between values of measurements of the factor and those of the behaviour

Hypothesis Formation

Typical hypothesis: factor X affects behaviour Y

Typical Null hypothesis: no effect of X on Y

What will we measure about X and Y?

Observation v Manipulation

- Observation studies: look at the population to see if X correlates with Y
- Manipulation experiments: change X and see what happens to Y

But we need to be sure that any change in Y is due only to the differences in X…

Influence of other factors

How do we know that the effects that we see (variations in measured behaviour) are due only to the changes in the factor of interest?

- other factors may influence behaviour of interest and may contaminate our experiments

Consider this during the experimental design:

- ***well designed experiment allows us just one explanation for effects we see in data it produces***
- ***while a poor design may allow many***

When you look at data, and consider the conclusions drawn, you need **always to ask what else might account for the effects described**….

What to do with Data I

The results that we collect by the various methods we use are data, often numerical representations of the behaviour of the system we are trying to investigate.

What can we do with such data?

What form can the data take?

How can we analyse it?

Data is a representation of reality:

- coloured by our presuppositions about the nature of reality
- encodes the behaviour of reality we consider significant.

What to do with Data II

If the following sequence was obtained:
4 7 9 3 4 11 12 1304 10 15 12 13 17 ...
We can look for trends
- early data around 3--9, later data 12--17.
But what about 1304? an anomaly? ignore it?

If no reason to ignore it, must account for it some other way.

The point is that the data ***just is***;
- what we think it means depends on what we believe about what it represents
- and the encoding process by which it was obtained

Feb-25-11 Human Communication 13

Types of Data

1. **Nominal (or Categorial)** data falls into classes, e.g.
- **weather** data classed as '***bright', 'cloudy', 'wet'***;
- categorisation by **treatment condition**

2. **Ordinal** data can be put in order or ranked e.g.
- **days** can be classified as '***cold***', '***warm***' or '***hot***', ranked as ***'cold' < 'warm' < 'hot'***
- **assignment performance** can be graded ***A, B, C, D***
- **ease of use** of a system graded as ***very easy, easy, difficult, very difficult***

3. **Numerical** data comprises numbers, for instance actual temperature - probably the most common type, but is also the most abused

When interpreting numbers we need to know what kinds of comparison are valid and where the origin of the scale is

Feb-25-11 Human Communication 14

Nominal (or Categorial) data
(the categories matter......)

1. Different **categories** of **user** of the joke generation system would be one of the set:
   {CCN, TD, CCN adult, SLT}
2. Language learners' **errors** can be divided into **types**, the frequency of each type is represented as nominal data, e.g.
   {use of definite article;
   incorrect placement of adjective}
3. Nominal data often identifies categories of object:
   {the category of all things coloured green}
   It is not possible to order the elements of these sets

Feb-25-11 Human Communication 15

Ordinal data *(the order matters...)*

The **usability** of an interface might be evaluated, using a questionnaire, with users rating it as:
{easy to use, average to use, difficult to use}
This would be **ordinal data** as:
easy to use < average to use < difficult to use
Users could be:
{children under 12, teenagers aged 13-18, adults over 18}
Again, we can order the values:
{children < teenagers < adults}
(We could decide to just use these as nominal categories)

But **there is no scale** on which we can place these values, so we do not know how much easier ***easy*** is than ***average***

e.g. ***one < a few < hundreds*** is another scale

Feb-25-11 Human Communication 16

Numerical Data: Interval
(distance between points matters...)

A weather forecasting system might take as input the **temperature** in ***degrees Celsius***
(e.g. say today was ***8*** *degrees Celsius*)

It makes sense to reason about the intervals between data points - so if yesterday was ***14*** it was ***6*** *Celsius units* above today's temperature.

This makes it sensible sometimes to plot one interval variable against another, such as the **variation in temperature** (in ***degrees Celsius***) against **time** (in ***days***).

Feb-25-11 Human Communication 17

Numerical Data: Ratio
(there is an absolute zero....)

a. Temperature in **degrees Kelvin** rather **Celsius** this gives us ***a scale with an absolute zero*** across ***days***
   - **today** is ***281.15 degrees Kelvin***
   - **yesterday** was ***287.15 degrees Kelvin***
b. Count **goals** in a football match ***across games***: scoring **4 goals** is ***100% more*** than scoring **2**

BUT scoring 60% in a test may not mean ***knowing twice as much*** as **scoring 30%**
it may be harder to score more marks the further up the scale you go

Feb-25-11 Human Communication 18

Transforming and missing data

We may want to transform data:
- to change the basic data type
- to obtain more uniform coverage (in clustered data)
- to avoid outliers

e.g. if exploring the efficiency of word prediction on mobile phone, measure no. of key presses, then group by frequency (10 people took 1-3 key presses, 12 took 4-8), ***but not equal intervals***, so it is **ordinal** rather than **numerical** data

What to do with missing data?
- skip it
- re-measure (if you can replicate the conditions)
- invent "expected values" (from the mean of the value under similar conditions)

Feb-25-11 Human Communication 19

Tools for Analysing Data

Data normally comes in sets - single experiment may involve repeating a test a number of times

Visualisation techniques used for **exploratory data**:
- display relationships between variables visually to make patterns in dataset apparent
- tools for this include:

MATLAB*, a matrix manipulation system with excellent graphical display abilities*
SPSS - Statistics Package for Social Science
Microsoft tools

Statistical tests used for **confirmatory experiments**:
- to determine extent to which an anticipated effect is present in the data from the experiment
- visualisation plays a much less significant role here, but may be a good starting point for "***eyeballing***" data

Feb-25-11 Human Communication 20

Looking for Effects

1. ***Population***: set of all instances of items of interest, e.g.
   - the set of all university students
   - the set of all runs of a program with certain parameters

   Typically too large for us to study exhaustively.
2. ***Sample***: subset of the population, small enough to work with, to draw conclusions about the population as a whole:
   - **set of all Human Comms students** as sample of the population of Edinburgh University students (which are a sample of the set of all University students)
   - **100 runs of a program** with given set of parameters as a sample of all possible runs of that program with those parameters

Feb-25-11 Human Communication 21

Sampling

Sample must be representative of the population it is taken from

e.g. ***Informatics 1 students*** not representative of ***students at Edinburgh University***
But ***random*** sample of ***20 Inf1B students*** could be ***representative of the Inf1 class***
Children from schools in different areas not representative of all children

How is the sample selected?
- must ensure a representative sample, for the purposes of the experiment at hand

How large must the sample be?
- the larger the sample the more work is involved but the more secure the conclusions will be

Feb-25-11 Human Communication 22

Statistic(s)

A ***statistic*** is a ***numerical encoding of some property of a population or sample***:
- hopefully characteristic of that population or sample, which we can use instead of the sample for reasoning about the properties of the sample.

e.g. **the average age of children in a class** summarizes certain properties of the population of such children.

Statistics is the subject that studies the properties of such encodings

Feb-25-11 Human Communication 23

Summary statistics

Summary Statistics express a property of the data set in a single number or set of a few numbers.

Most common are ***mean***, ***mode*** (most frequent value), ***median*** (middle value), ***variance*** and ***standard deviation***

Mean (μ) gives the centre of mass of the set:

$$\mu = \frac{\sum \{x1,...xn\}}{n}$$

So mean of {2, 3, 6, 1, 5, 1} = 18/6 = 3

Feb-25-11 Human Communication 24

Summary statistics

{6, 2, 3, 6, 1, 5, 1, 7, 2, 5, 6}

Mean = $\mu = \frac{\sum \{x1,...xn\}}{n}$

= (6+ 2+ 3+ 6+1+5+1+7+2+5+6)/11
= 44/11 = **4**

Median = middle value
1,1,2,2,3,5,5,6,6,6,7 = **5**

Mode = most frequent value = **6**

Feb-25-11 Human Communication 25

Variance

Variance is the mean deviation from the centre:

$$\text{Variance} = \frac{\sum \{(x1 - \mu)^2,...(xn - \mu)^2\}}{N}$$

$$= (\sum (X - \mu)^2)/N$$

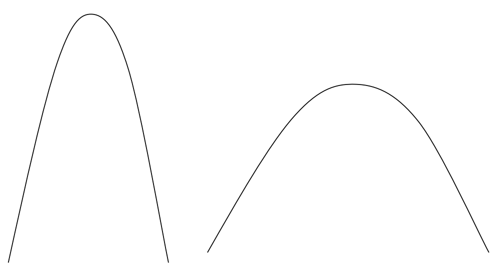
The Standard deviation σ is the square root of the variance:

$$\sigma = \sqrt{(\sum (X - \mu)^2)/N}$$

Feb-25-11 Human Communication 26

Normal distributions

Left curve shows less variance than right curve



Feb-25-11 Human Communication 27

General References

Cohen, P. (1995) *Empirical Methods for Artificial Intelligence*, MIT Press, 1995.
Dix, A., Finlay, J., Abowd, R. and Beale, R. (2004) *Human-Computer Interaction.* Prentice Hall
Preece, J., Rogers, Y., Sharp, H., Benyon, D. Holland, S. and Carey, T. (1994). *Human-Computer Interaction*. Addison-Wesley

Feb-25-11 Human Communication 28