Section 3.1 & 3.2 - Design of Experiments

Statistics 104

Autumn 2004

Producing Data

Where can get data to answer a question of interest?

- Anecdotal Evidence: Based on haphazardly selected individual cases, which often come to our attention because they are striking in some way. These cases need not be representative of any larger group of cases.

  They usually tend to be biased towards one part of the group of cases and this leads to overestimated effects.

- Available Data: Data that were produced in the past for some other purpose but that may answer a present question.

  These might be government databases, such as those available from the census bureau, databases available online or a library. An example is the Harvard-MIT Data Center <http://www.hmdc.harvard.edu/jsp/index.jsp>

- Sampling

  A observational study designed to answer specific questions, where the study group is sampled from a larger population. Examples of sample surveys include political polls and Nielsen TV ratings. A sample survey observes individuals and measures variables of interest, but does not attempt to influence responses.

- Experiments

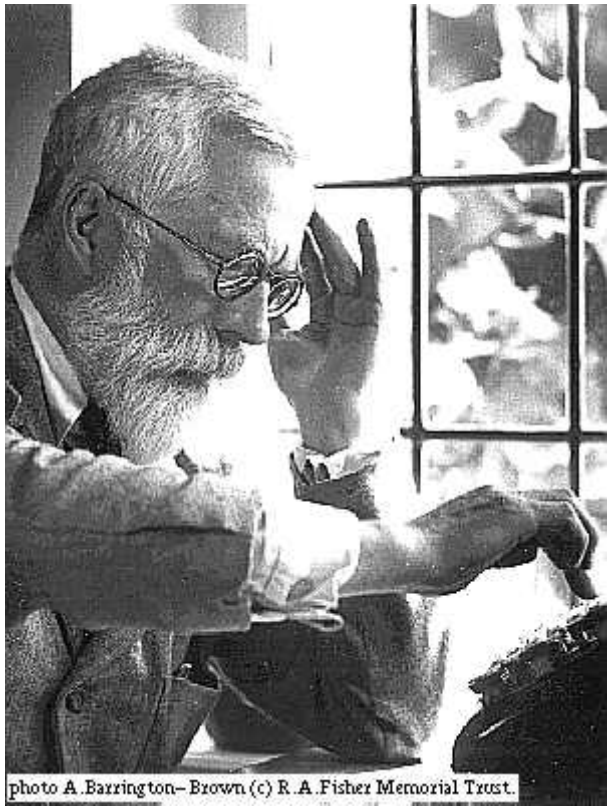
  A designed study where researchers impose treatments on individuals to observe their responses. Examples of experiments include clinical trials, agricultural trials, studies to improve product quality, etc.

Design of Experiments

Example: Plastic Parts Experiment

- X: time spent in mold (assume possible times are 10, 20, 30, and 40 seconds)
- Y: strength of the part
- Z: temperature of mold. Seen to be a confounding factor.

How can we improve the experiment to avoid the confounding problems? Remember in the original experiment, all the 10 second parts were made first, then the 20 second parts, and so on.



R.A. Fisher

Terminology

- Experimental units: Objects on which the experiment is performed. Referred to as subjects in people, as in a clinical trial.

- Factor: Explanatory variables in an experiment. Can be set by the experimentor.

- Level: Specific value of a factor.

- Treatment: Specific experimental condition applied to an experimental unit.

Example: Effect of antibiotics and vitamin B12 on weight gain in cattle.

Could give a steer an antibiotic (or not). Similarly with the vitamin B12 supplement

- Experimental Units:
  - Steers
- Factors:
  - Antibiotic – Levels: Yes, No
  - Vitamin B12 – Levels: Yes, No
- Treatments:

Treatment	1	2	3	4
Antibiotic	No	No	Yes	Yes
Vitamin B12	No	Yes	No	Yes

Example: Effects of Biotene Toothpaste

- Suggested that Biotene can control plaque and gum disease in people with reduced saliva levels
- System can be retarded by radiation therapy in people with tumours of the head and neck
- Works by stimulating the salivary system
- Lab studies have provided direct and indirect evidence that Biotene toothpaste should help patients with reduced and normal saliva level.

The clinical trial preformed in discussed in

Toljanic JA et al (1996). An Evaluation of a Dentifrice Containing Salivary Peroxidase Elements for the Control of Gingival Disease in Irradiated Head and Neck Cancer Patients. Journal of Prosthetic Dentistry, **76**: 292-296.

Question: Does Biotene toothpaste lower plaque levels?

Study Design (simplified):

- 40 normal patients, 60 cancer patients
- 3 visits to dentist, 3 months apart
- At each visit, hygienist clean teeth.
- At each visit, before cleaning, Plaque Index (PI) is measured on 4 faces of each tooth (up to 128 ($= 32 \times 4$) measurements).

Proposal 1:

> Give all 100 participants Biotene toothpaste. Compare average PI score after 3 months with initial average PI score

Observation 1 ⟶ Treatment ⟶ Observation 2

Lets assume that for most people, the average PI level decreases. PI levels could have possibly decreased due to

- Biotene toothpaste
- Better brushing habits
- Cleaning by the hygienist
- **Placebo effect**

Placebo: a dummy treatment, in the case, regular toothpaste.

Placebo effect: many people respond positively to any treatment, even a sugar pill.

Solution: need a control group

Controlling the effects of outside variables is the first principle of statistical design of experiments.

In a clinical trials setting, the control group would receive the current standard treatment, assuming one exists. Otherwise a placebo control would be used, as in the gastric freezing example in the text.

In other settings, the control group could be a standard or reference dose. For example, in the Antibiotic/Vitamin B12 example, the levels could be

Antibiotic – 0, 200, 400 mg / day Vitamin B12 – 0, 600, 1200 mg / day

This would lead to 9 possible treatments. The control group could receive the 0/0 treatment combination.

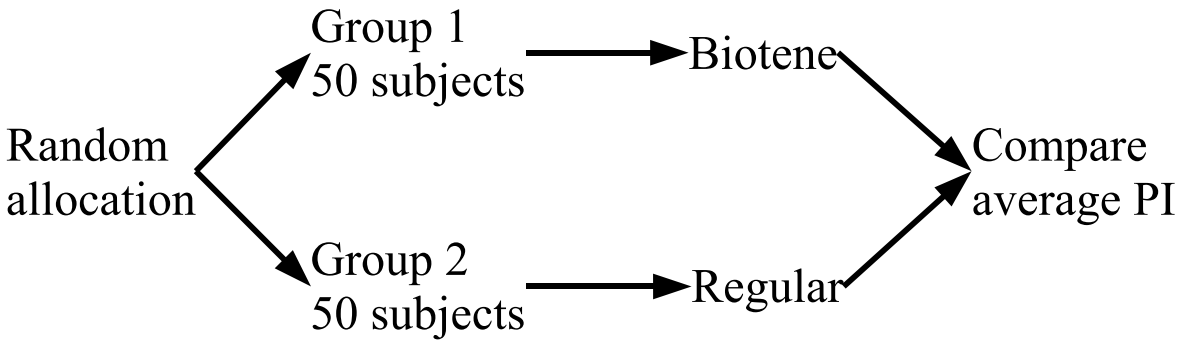
Proposal 2:

Give 50 people Biotene toothpaste

Give 50 people regular toothpaste

Allocate people by **Randomization**

Compare changes in PI between both groups



Randomization is the second major principle of statistical design of experiments.

Replication is the third major principle of statistical design of experiments.

Problems:

- People may have wide range of PI, possibly due to lurking variables
- It is expected that cancer subjects could have quite different response compared to normal subjects

Comparison of several treatments is the simplest form of comparison

Without comparison of treatments, experimental results can be dominated by such influences as the details of experimental arrangements, selection of subjects, and the placebo effect.

Bias:

> The design of an experiment is biased if it systematically favours certain outcomes.

Want a scheme to give some subjects Biotene toothpaste and the rest regular toothpaste that avoids bias.

Want to avoid, for example, the situation where mainly normal subjects get Biotene and mainly cancer patients get normal toothpaste

Could match people based on age, sex, medical conditions, etc.

However this could bias the study since there may be non-experimental variables which may affect the outcome.

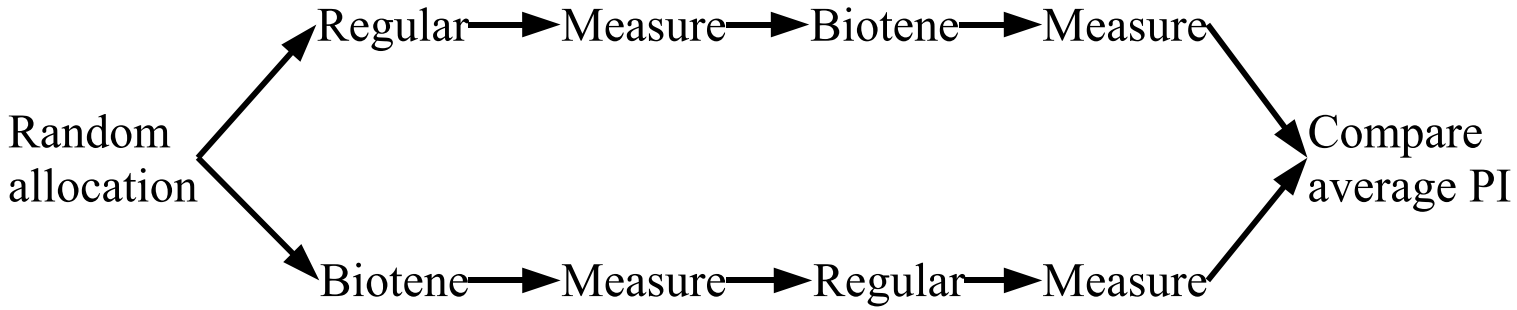
Want an allocation method that doesn't depend on the experimental units of judgement of the experimenter.

Proposal 3:

Give each subject both toothpastes, half starting with Biotene and switching to regular after the second visit, with the rest getting the opposite.

Allocate normal and cancer patients separately (blocking).

Within each block (cancer or normal), compare average PI when using Biotene with average PI when using regular toothpaste.



This is an example of what is known as a crossover trial.

Advantages:

- Each person is his own control, reducing variation
- Similar subjects are compared, also reducing variation

Logic Behind Randomization

Produces groups which should be similar in all respects before treatments are applied.

There will be some differences since each unit is different.

Averaging over groups helps cancel out these differences, leaving only the treatment differences.

Allows for standard statistical inference procedures to be used.

How to randomize - Using Table B

1) Each entry in the table is equally likely to be 0 to 9.

2) Entries are independent.

3) These imply that any pair of digits has the same chance of being any of the 100 possible pairs: $00, 01, 02, \ldots, 98, 99$. Also for triplets, quadruplets, etc of digits.

TABLE B Random digits

Line								
101	19223	95034	05756	28713	96409	12531	42544	82853
102	73676	47150	99400	01927	27754	42648	82425	36290
103	45467	71709	77558	00095	32863	29485	82226	90056
104	52711	38889	93074	60227	40011	85848	48767	52573
105	95592	94007	69971	91481	60779	53791	17297	59335
106	68417	35013	15529	72765	85089	57067	50211	47487
107	82739	57890	20807	47511	81676	55300	94383	14893
108	60940	72024	17868	24943	61790	90656	87964	18883
109	36009	19365	15412	39638	85453	46816	83485	41979
110	38448	48789	18338	24697	39364	42006	76688	08708
111	81486	69487	60513	09297	00412	71238	27649	39950
112	59636	88804	04634	71197	19352	73089	84898	45785
113	62568	70206	40325	03699	71080	22553	11486	11776
114	45149	32992	75730	66280	03819	56202	02938	70915
115	61041	77684	94322	24709	73698	14526	31893	32592
116	14459	26056	31424	80371	65103	62253	50490	61181
117	38167	98532	62183	70632	23417	26185	41448	75532
118	73190	32533	04470	29669	84407	90785	65956	86382
119	95857	07118	87664	92099	58806	66979	98624	84826
120	35476	55972	39421	65850	04266	35435	43742	11937
121	71487	09984	29077	14863	61683	47052	62224	51025
122	13873	81598	95052	90908	73592	75186	87136	95761
123	54580	81507	27102	56027	55892	33063	41842	81868
124	71035	09001	43367	49497	72719	96758	27611	91596
125	96746	12149	37823	71868	18442	35119	62103	39244

This table is a set of realizations from the discrete uniform distribution.

Lets suppose we want to allocate 100 units, as for proposal 2 in the Biotene study.

In general label units from 1 to N (or from 0 to $N-1$)

Want n_1 units for treatment 1 (50 with Biotene)
Want n_2 units for treatment 2 (50 with regular)

Group digits in block (pairs for Biotene case as units can be numbered from 0 to 99)

Go along line. If a number hasn't come up before, allocate to the current group.

Stop when enough are allocated for group.

If more groups need to be allocated, repeat procedure.

Assume the following line is in the table

$45149 \;\; 34572 \;\; 84441$

$45 \mid 14 \mid 93 \mid 45 \mid 72 \mid 84 \mid 44 \mid 1?$

Biotene subjects: 45, 14, 93, 72, 84, 44, etc

Regular toothpaste subject: the remaining 50.

When using the table, you should start at a random location, otherwise all your studies will have similar treatment allocations.

You can also use the computer to generate treatment randomization assignments using random number generators

Assume there are N units total in the experiment, with n_1 units getting treatment 1, n_2 units getting treatment 2, and n_3 units getting treatment 3.

Generate N observations from a continuous $Unif(0,1)$ distribution. (Actually any continuous distribution can be used)

Calculate the ranks of the N random numbers

rank $= 1$ if smallest

rank $= 2$ if 2nd smallest

etc

Units with ranks 1 to n_1 get treatment 1, rank $n_1 + 1$ to $n_1 + n_2$ get treatment 2, etc.

Why does this work?

The rank corresponding to the first unit is equally like to be 1 to N. Thus

$$\begin{aligned} P[1 \leq Rank_1 \leq n_1] &= \frac{n_1}{N} \\ P[n_1 + 1 \leq Rank_1 \leq n_1 + n_2] &= \frac{n_2}{N} \\ P[n_1 + n_2 + 1 \leq Rank_1 \leq n_1 + n_2 + n_3] &= \frac{n_3}{N} \end{aligned}$$

Similarly for the other units. This occurs since every set of N generated random numbers containing the same list of numbers has the same likelihood (e.g. the pair 0.2453 and 0.6532 is just as likely as 0.6532 and 0.2453).

Cautions When Performing Experiments

- Hidden bias:

  All units must be treated as similarly as possible, or bias may occur. In the Biotene trial, the two types of toothpaste could not be easily identified.

- Double-blind experiments:

  Subjects and evaluators should not know which treatment was given.

  Subjects knowing could influence general outcome.

  Evaluators may act differently if they know which treatment (e.g. subconsciously score Biotene patients better).

  The Biotene study was a double blind study.

- Realism:

If the experimental conditions don't match a realistic situation, conclusions may not follow.

Other study considerations

- Paired comparisons:

  Subject to subject variability may be large, within subject to subject variability may be small.

  Suggests self control (as done in the Biotene study)

- Blocks:

  Groups of units known ahead of time to be similar in some way which is expected to influence the response (e.g. cancer vs. regular subjects)

- Block Design:

  Separate randomizations are carried out separately in each block.

Plastic Parts Experiment

Suggested changes from a previous class

1. Monitor mold temperature
2. Adjust mold temperature
3. Try different molds at the same temperature
4. Try other times
5. Preliminary trials before taking measurements
6. Mix up order of times