

MINUTES OF THE STATISTICS SUBCOMMITTEE 84TH MEETING, MAY 16 – 18, 2005

TUESDAY MAY 16, 2005

Present: Veda Abu-Bakare (Langara College and Thompson Rivers University-Open Learning Institute), Susan Chen (Camosun College), Bruce Dunham (University of British Columbia), Bevan Ferreira (Selkirk College), Claude Hurtubise (College of New Caledonia), Kevin Keen (University of Northern British Columbia), Colin Lawrence (British Columbia Institute of Technology), Alex Liu (Kwantlen University College), Geoff Salloum (Camosun College), Ken Towson (Capilano College), Min Tsao (University of Victoria), Larry Weldon (Simon Fraser University), Julie Zhou (University of Victoria)

Acting Chair: Kevin Keen

Acting Secretary: Claude Hurtubise

1. Approval of Agenda

Motion: That the agenda as proposed be adopted with the addition of items: 4.4 Institutional Reports; 8.1 Comments on the Plenary Presentation by Mr. Richard DeMerchant (BC Ministry of Education), and; 8.2 Province-Wide Minitab Licence (Veda Abu-Bakare, Langara & TRU-OL). Moved by Susan Chen and seconded by Colin Lawrence. **Carried unanimously.**

2. Approval of Notes and Minutes of the Statistics Subcommittee from the 83rd Articulation Meeting

Motion: That the notes and minutes of the Statistics Subcommittee from the 83rd Articulation Meeting as circulated be adopted. Moved by Colin Lawrence and seconded by Alex Liu. **Carried unanimously.**

3. Articulation and Transfer Issues for Courses in Statistics

3.1. General Items.

It was noted that policies regarding transfer credits varied among the four research intensive universities for the same course. It was noted that a general rule for articulation of two courses was 80% of overlap in content. It was remarked that the content of a course as taught does not always match calendar descriptions. It was recommended that statistics courses carry the prefix *STAT* as opposed to labelling the course as *MATH*.

3.2. Response from the Canadian Engineering Accreditation Board Concerning Requirements in Introductory Probability and Statistics Courses for Engineers.

With respect to item 6 of the minutes of Statistics Subcommittee at the 83rd Meeting, it was moved that the Chair write a letter to the Canadian Engineering Accreditation Board (CEAB) to enquire about content required in statistics courses for engineering students. The Chair received a response from Mr. Lynn Villeneuve that members of the CEAB Executive had considered the question posed by the Chair. The response conveyed by Mr. Villeneuve was that the “CEAB avoids being prescriptive.... Probability and Statistics courses would be expected to be calculus based, and would generally include probability theory to a depth (e.g., probability density function) that could not be done without calculus. The result of our finding a case, including—or perhaps especially—in a transferred credit, where the examiner found a course to be of inadequate depth, would be to raise a flag and probably require a correction. Several such courses would be serious.”

4. Issues relating to Existing or Future Accredited Undergraduate Statistics Programmes

4.1. Changes in Promotional Criteria at UBC (Bruce Dunham, UBC).

There are to be some notable changes in the progression/promotion rules within the Science faculty, coming into effect in the 2006/07 academic year. Specifically, for the Statistics major programme, for promotion to year 3 the following must occur.

- (i) Students must achieve an average of at least 60% on their first attempt in their best three of five "named" courses in year 2. For Statistics these five are Math 200 (Calculus III), Math 220 (Mathematical Proof), Math 221 (Matrix Algebra), Stat 200 (Elementary Statistics) and Math/Stat 302 (Introduction to Probability).
- (ii) Stat 200 must be passed.

For promotion to year 4, Stat 305 and Stat 306 must be passed.

4.2. New Course at SFU: STAT 300 Communications (Larry Weldon, SFU)

This is a writing course in statistics. Requirements to be met by students include writing a scientific report, providing a description to a lay audience, and critiquing scientific articles. The course will module-based. A course outline will be available on the SFU website for anyone to view in September 2006. (Course Description attached in Annex 1.)

4.3. Statistics Courses at the University of Victoria (Min Tsao, UVic)

It was noted that the Department of Mathematics & Statistics at the University of Victoria offers services statistics courses for [1] computer sciences, mathematics and engineering students (Stat 254, 260 and 261), [2] Life sciences and environmental studies students (Stat

255 and 256) and Business students (Stat 252). Other faculties/departments have their own statistics courses.

Details of the services courses, including the coverage, the computer software and the number of students each year, were discussed. Of particular interest to participants from the colleges are the fact that Stat 254, 260 and 261 are calculus based whereas Stat 255, 256 and 252 are not calculus based. Also, Stat 254 and Stat 252 are stand-alone courses, Stat 255 and 256 form a sequence, and Stat 260 and 261 form a sequence.

4.4. Institutional Reports

Simon Fraser University (Larry Weldon)

1. The WQB requirements to graduate from SFU will apply to all students registering in the Fall Semester 2006 or later. Details may be gleaned from

http://www.sfu.ca/ugcr/For_Faculty/WQB_Criteria_and_Definitions/

There are provisions for upgrading students in quantitative and writing skills for those students who need it. The courses FAN 099 (Numeracy) and FAL 099 (Language), are 4-credit courses that have been created for this purpose. Students successfully completing one of these courses will be better prepared to satisfy the Q or W required courses. The handout SFU Curric.doc attached as Annex 2 was handed out at the meeting.

2. STAT 300 is W- and Q- certified course. This new course *Statistics Communication* is scheduled to be taught for the first time in Fall 2006. Its objective is to teach students how to write reports of data-based studies to both scientific and lay audiences. Course Description attached in Annex 1.
3. SFU has a new Faculty of Health Sciences. There is already graduate activity in biostatistics and interest in creating an undergraduate program too.

4. Enrolments: Number of Majors in Statistics, Actuarial Science, Management and Systems Science, and Mathematics courses. (attached)
5. The Management and Systems Science undergraduate program at SFU has been ongoing for 25 years. The program mainly requires courses in Math, Business and Computing, and a few courses in Statistics and Economics. The program graduates only 10-15 students per year. There has been discussion of moving the program to the Surrey Campus.
6. SFUs STAT 270 and STAT 285 as the first two courses that a statistics major must take in order to declare a major in statistics, and prepare for the UD STAT courses. STAT 270 and STAT 285 both contain statistics and probability, as described in the SFU calendar. This may differ from the perception of some that STAT 270 is probability while STAT 285 is statistics.
7. The current instructor for STAT 100 has lots of resources at <http://www.sfu.ca/~ablejec/STAT100/>, including some Excel applications, and Dr. Weldon still has his own 2002 STAT 100 notes (and MINITAB applications) at <http://www.stat.sfu.ca/~weldon/stat100-02-3.html>.
8. Enrolled Majors for Calendar years 1997-2005, Department of Statistics and Actuarial Science at SFU.

Year	Act Sci	Stat	MSsc	Math
97	0	12	15	77
98	0	21	17	95
99	13	27	20	104
00	28	34	13	102
01	53	39	26	116
02	73	42	46	122
03	54	53	36	108
04	54	57	25	97
05	45	53	15	79

Source :

<http://www.sfu.ca/analytical-studies/FactBook/d-student/sectd.ug.html>

University of British Columbia (Bruce Dunham)

At the institutional level, UBC recently announced its new President as Stephen Toope, a law scholar, who will take office from 1st July 2006. Furthermore the Science Faculty, which houses Mathematics and Statistics, will be appointing a new dean shortly - at the time of writing a shortlist of three had been announced and the selection process underway.

The Statistics department has enjoyed a good year, with student numbers showing a modest increase and two new faculty members being appointed. Prof. Jim Zidek, a long-standing member of the department and former head, retired in 2005, but is still active in his research activities within the department. Plans are in progress for the creation of a new building to accommodate Statistics, along with Earth and Ocean Sciences, PIMS and the Dean of Science office.

University of Northern British Columbia (Kevin Keen)

The grades D-, D, and D+ now replace the D letter grade at UNBC. MATH 242 Statistics for Social and Health Sciences will be offered in the 2007 January semester and might then be offered regularly in both September and January semesters. MATH 242 is being taught in the 2006 May Intersession and was also taught in 2005 May Intersession.

Camosun College

Math 116, an introductory statistics course for students with Grade 11 math, had a 2005 summer offering, which seemed to reduce fall enrolment. It has not been scheduled for summer 2006. We have lost a section of Math 116, but have gained a section of Math 216, which is an introductory statistics course for students with Grade 12 math. Enrolment in Math 218, the calculus-based introductory statistics course, is down slightly. We suspect this drop is due to the enrolment decline in Computer Science. Geoffrey Salloum has been hired as a continuing instructor to replace Bill Calver who has retired.

Langara College

Langara has developed a new STAT course, STAT 3223 Intermediate Quantitative Methods which is designed to meet the needs of students enrolled in a business administration degree program and those planning to achieve a professional accounting designation such as a Certified General Accountant (CGA). Topics include statistical inference for one and two populations, simple linear regression multiple linear regression, index numbers and time series analysis, decision theory, and linear programming.

Selkirk College

Our recently-introduced second-year, calculus-based introductory statistics course, STAT 206, continues to hold steady enrolment, and has thus far had some appeal beyond its intended audience of science and engineering majors.

In the fall semester of this year we will introduce a 1½ -hour lecture format to replace our previous 2-1 (with 2-hour lab) and 2-2 structure. It was felt that this would help us cover certain courses more effectively, would bring us more into line with current university practice, and also help ease timetabling restrictions.

Thompson Rivers University, Open Learning Division (TRU-OL)

STAT 102 is being revised for Intro Stats by DeVeaux and Velleman . We have in the past used Minitab but have run into technical difficulties with supporting the software at a distance. The course writer has chosen to use the TI83+.

5. Update on the Activities of the Education Committee of the Statistical Society of Canada (Min Tsao, UVic)

Current activities of the Committee center on the following objectives.

- Organization of a session or presentations on statistical education at the Society's annual meeting (subject to approval by the program chair).

- Development and maintenance of a website for the purposes of reporting the activities of the committee and providing links to useful on-line education resources.
- Development of an internet discussion site and/or list serve could be considered for the purpose of fostering communication among teachers of statistics regarding new developments in textbooks, teaching methods, computer hardware and software useful in teaching, interesting databases and teaching examples, videotapes, etc.
- Service to teaching institutions: provide resources and information that could be used to present different aspects of statistics to educators and students.
- Service to students: provide information about career paths in statistics, student awards available in statistics, databases and case studies that could be used for statistical projects.
- Facilitating project judging in the area of statistics and awarding the SSC prizes at the annual Canada-Wide Youth Science Fair.
- In particular, the Statistical Education Session at the upcoming 2006 SSC annual meeting sponsored by the Education Committee was noted and participants were invited to attend the session.

6. Information Items

No information items.

7. Recommendations for the Statistical Subcommittee at Future Articulation Meetings

The Chair noted that at the 83rd Articulation Meeting the previous year in Prince Rupert there were only two statistics representatives in attendance (one each from UBC and UNBC). Attendance at the Statistics Subcommittee sessions of the 84th Articulation Meeting has been good with 8 institutions sending statisticians. However, the 85th Articulation Meeting will be held in the Whitehorse, Yukon. Because of travel costs, this meeting might be as poorly attended as the 83rd in Prince Rupert. The Chair raised the notion of the Statistics Subcommittee meeting in the Lower Mainland apart from the Mathematics Subcommittee when the latter in alternate years meets outside the metropolitan southwest. Ms. Abubakare spoke against this proposal. No motion emerged from discussions to depart from the practice of meeting every year with the Mathematics Subcommittee.

8. New Business

8.1. Comments on the Plenary Presentation by Mr. Richard DeMerchant (BC Ministry of Education)

Ms. Abu-Bakare encouraged members of the Statistics Subcommittee to consider the recommendations by Mr. DeMerchant regarding changes to the secondary grades 10-12 mathematics curriculum and to provide their comments in writing to Mr. DeMerchant before September 2006.

8.2. Province-Wide Minitab Licence (Veda Abu-Bakare, Langara & TRU-OL)

Ms. Abu-Bakare is working with Minitab, Inc. to obtain a single licence of Minitab for all interested post-secondary institutions in BC. If your institution would like to participate, please contact Ms. Abu-Bakare

9. Motion to Adjourn

Colin Lawrence moved to adjourn. **Carried unanimously.**

THURSDAY MAY 18, 2006

9:00 – 9:50

Roundtable Discussion: A Roundtable Discussion on the Role of Technology in Introductory Statistics Courses

Presenter Susan Chen, Ph.D. (Camosun College)

Acting Secretary: Kevin Keen

Present: Veda Abu-Bakare (Langara College and Thompson Rivers University-Open Learning Institute), Susan Chen (Camosun College), Bruce Dunham (University of British Columbia), Bevan Ferreira (Selkirk College), Claude Hurtubise (College of New Caledonia), Costa Karavas (Vancouver Community College), Kevin Keen (University of Northern British Columbia), Colin Lawrence (British Columbia Institute of Technology), Alex Liu (Kwantlen University College), Geoff Salloum (Camosun College), Ken Towson (Capilano College), Min Tsao (University of Victoria), Larry Weldon (Simon Fraser University), Patricia Wrean (Camosun College), Julie Zhou (University of Victoria)

On the topic of statistical software used in introductory statistics courses, representatives of postsecondary institutions replied and commented as follows.

Simon Fraser University (Larry Weldon)

Minitab, R, DataDesk, ActivStats, Excel, and JMP have all been used. SPSS has been used exclusively in STAT 203. The advantage of Minitab is that it is simple to use while R is free and Excel is widely available.

Langara College (Veda Abu-Bakare)

StatGraphics Plus is used with free downloads for student-owned computers.

Thompson Rivers University—Open Learning (Veda Abu-Bakare)

Minitab has been used in the past but there will be a switch to the TI-83+ calculator

Camosun College (Geoff Salloum)

SPSS is bundled with 1/3 of the textbooks for the statistics course for the social sciences but Minitab is used for all other courses. All statistics courses have a 1-hour schedule lab every two weeks, and a lab final examination worth 10% of the course grade.

Camosun College (Patricia Wrean)

The TI-89 calculator is used for the statistics course for the Engineering Technology and bridging programs at the Interurban campus. Minitab was considered but determined to be too expensive and so was not chosen.

University of Northern British Columbia (Kevin Keen)

A variety of statistical software packages have been used in MATH 242 and MATH 342: Excel, Minitab, SAS, and SPSS. In the future, R might be used in MATH 242. However, MAPLE has been used consistently for MATH 340 and MATH 341 which have calculus prerequisites. MAPLE is used in calculus courses at UNBC.

University of Victoria (Julie Zhou)

The second year statistics courses use Minitab. The third year statistics courses use S-Plus or R. SPSS has been used in a fourth year course in survival analysis.

University of British Columbia (Bruce Dunham)

STAT 200 has used Minitab and JMP, but now uses Excel. Upper division statistics courses use R, as do STAT 241 and 251 which are statistics courses for engineering students. Other courses have used Excel, JMP, and Minitab, the latter no longer being supported at UBC.

Capilano College (Ken Towson)

The statistics courses use the TI-83+ graphing calculator

College of New Caledonia (Claude Hurtubise)

MATH 104 uses no statistical analysis software.

Selkirk College (Bevan Ferreira)

StatDesk, bundled with the Triola *et al.* textbook, is used. DD Excel has also been used. Consideration will be given to use MAPLE in the calculus-based statistics course in the future as MAPLE is already available.

Kwantlen University College (Alex Liu)

Excel is currently used. Minitab has been used in the distance past. However, regression calculations in Excel have been found not to be accurate.

British Columbia Institute of Technology (Colin Lawrence)

Individual instructor decides which statistical software package to use. Excel, Minitab, and TI-83+ calculators have all been used.

Vancouver Community College (Costa Karavas)

VCC has one statistics course for Arts students and it uses Excel. However, Minitab or MAPLE could be used.

On the topic of computer lab hours for statistics courses, representatives of postsecondary institutions replied and commented as follows.

Simon Fraser University (Larry Weldon)

SFU has no scheduled lab hours but does have an open lab.

Camosun College (Patricia Wrean)

There are no scheduled computer lab hours on the Interurban campus.

Camosun College (Geoff Salloum)

One hour scheduled lab for every 2 weeks, and there is laboratory final examination worth 10% of the course grade.

University of Victoria (Julie Zhou)

STAT 260/261 have a preparatory lab and lab tests. STAT 254/256 have Minitab assignments. Other 200-level statistics courses do not have regularly scheduled lab hours.

University of Northern British Columbia (Kevin Keen)

MATH 242 and 342 each have one scheduled lab hour per week. MATH 340 and 341 have no scheduled lab hours.

University of British Columbia (Bruce Dunham)

UBC does have scheduled lab hours for introductory statistics courses but they are poorly attended.

Capilano College (Ken Towson)

Capilano College has weekly drop-in Excel labs for MATH 204.

College of New Caledonia (Claude Hurtubise)

CNC statistics courses have no lab component

Selkirk College (Bevan Ferreira)

MATH 105 has a 4 hours of class and 1 hour lab weekly. MATH 206 has a 3 hours of class and 2 hours of lab (which can also be used for classes) weekly.

Kwantlen University College (Alex Liu)

There are no lab classes.

British Columbia Institute of Technology (Colin Lawrence)

Each instructor has tutorial hours and can use them as he or she sees fit.

Vancouver Community College (Costa Karavas)

The introductory statistics course has 3 hours of classes and 1 hour of lab.

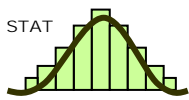
10:00 – 10:50

- Seminar:** How to Use R in a First Course in Statistics
- Presenters:** Larry Weldon, Ph.D. (Simon Fraser University)
Kevin Keen, Ph.D., P.Stat. (UNBC)
- Handouts:** See Annex 3 for handouts of Powerpoint slides.

11:00 – 11:50

- Seminar:** Quality Assurance in University Teaching: A View from Both Sides of the Atlantic
- Presenter:** Bruce Dunham, Ph.D. (UBC)
- Summary:** Having recently arrived in Canada from the UK, Dr. Dunham illustrated some key differences in practices in university teaching between the two countries. Starting with a short historical background to the role of quality assurance in higher education in the UK, the talk focused on the different expectations that would be placed on a new academic teaching a course for the first time in UK or Canada, the objective being to highlight and motivate some of the procedures which while now commonplace in Britain could be deemed unwieldy or unnecessary over here. The aim of the talk was to stimulate thought and discussion regarding the appropriateness of academic quality assurance procedures in higher education establishments in BC, considerations which might in turn lead to some changes in practice.
- Handouts:** See Annex 4 for handouts of slides.

See pages following for Course Description of STAT 300 at Simon Fraser University.



Students requiring accommodations as a result of disability, must contact the Centre for Students with Disabilities 604-291-3112 or csdo@sfu.ca

Instructor: [Dr. Larry Weldon](#)

Prerequisite:

Admission to major or honors programs in statistics or actuarial science at SFU.

Calendar Description:

Guided experience in written and oral communication of statistical ideas and results with both scientific and lay audiences.

Outline:

The course will involve seminars for three hours per week.

The objectives of the course are:

1. To provide students with a vocabulary which facilitates written verbal communication of statistical ideas and statistical results.
2. To provide guided experience in writing technical reports of data-based studies.
3. To provide guided experience in writing about statistical concepts to a non-technical audience.

The course would give students the opportunity to receive feedback from the instructor, a TA, and peer students in the class, on their writing. Students would be assigned writing tasks involving reports of data-based studies. Sources for the studies would include statistical and scientific journals or small scale data-based studies performed by the students themselves. Assigned writing tasks would include summary reports of published studies, or critiques of studies.

The first draft of such reports would be critiqued but not marked. Students would be expected to respond to the critiques and resubmit a final version for marking. Each student would complete approximately six such reports during the semester. Students would also be marked on their contribution to seminar-type discussions led by the instructor.

The choice of data-based study to present each week will be up to the instructor. However, data sources could include any of the following:

1. Data is collected by the students themselves as directed by the instructor, directed at a particular question.
2. A video describing a scientific project could be presented.
3. Data extracted from public data bases: Statistics Canada, Canadian Institute of Actuaries
4. Students could be asked to extract descriptive information from individual masters projects in statistics or actuarial science.
5. Published scientific studies could be used as a basis for a summary report.

The task in reports could be either to draft a formal report (such as might be prepared for an applications journal or technical magazine) or it could be a less formal report aimed at making objectives and results clear to a lay audience. An assignment might involve two reports of the same study, the difference being the degree of familiarity of the audience to statistical jargon.

Grading Scheme:

20% - Timely submission of draft reports for critiquing.

20% - Participation in class discussions.

60% - Response to critique and the quality of the final submission of the reports.

Quality of reports will involve assessment of report layout, scientific logic and language clarity.

There will be no final examination.

Students should be aware that they have certain rights to confidentiality concerning the return of course papers and the posting of marks. Please pay careful attention to the options discussed in class at the beginning of the semester. Students are reminded that Academic Honesty is a cornerstone of the acquisition of knowledge. Scholarly integrity is required of all members of the University. Please consult the General Guidelines of the calendar for more details.

See pages following for the handout SFU Curric.doc as handed out at the meeting.

SFU's Undergraduate Curriculum Initiative

Items for consideration by Articulation Committees

Spring 2006

Introduction

Simon Fraser University has approved changes to its degree requirements that will apply to all undergraduate students admitted for September 1 2006 and thereafter. This document is a summary of these changes and of implications to sending institutions.

The changes being implemented include the requirement that students complete 6 credits in courses identified as "writing intensive" (W), 6 credits in courses identified as "quantitative-reasoning" (Q) and 24 credits identified as "breadth" (B) including

- a) 18 credits of Designated Breadth (DB) (6 credits in Social Sciences, 6 credits in Humanities, and 6 credits in Sciences); and
- b) 6 credits of Undesignated Breadth (UB) (courses taken outside the student's major program).

We have made significant progress in reviewing both our courses and courses from sending institutions and certifying appropriate courses as W, Q and/or B. Note that it is possible for a course to hold more than one designation.

These innovations to the SFU undergraduate curriculum will enrich the quality of an SFU education and prepare students for more promising futures.

Certification of College Courses

We hope that students who transfer to SFU will be able to satisfy some of these requirements using courses that they complete at sending institutions. Ideally, students transferring into second year should bring 1 W, 1 Q and 2 B courses. Students transferring into third year should bring 1 W, 1 or 2Q and 4 B courses.

We have worked with BCCAT to add W, Q and B evaluation to the regular transfer articulation process so that this process should be as seamless as possible. Many college courses have now been certified. We will summarize the process here – for details please see the UCI web page (www.sfu.ca/ugcr). There are two important things to keep in mind:

1. Some courses will automatically be certified based on current transferability, but colleges may need to present other courses and supporting information for certification.
2. The procedure for certifying W courses is different from the procedure for certifying Q and B courses. This is because W-designation depends mainly on pedagogy, while Q and B designation depends mainly on course content.

Articulation Process for College courses:

- a) Courses that articulate to SFU for assigned (Type 1) credit will automatically be designated as Q or B if the SFU courses are certified as Q or B. This is because Q and B designations are assigned based on content, as are transfer/articulation evaluations.
- b) All courses seeking W certification, and courses seeking Q or B certification that have been articulated as unassigned (Type 2 or 3) credit require that the sending institutions submit a College WQB Certification Request form (available at the UCI web page www.sfu.ca/ugcr).
- c) Courses that have not yet been articulated require that the sending institutions complete the usual material for transfer requests **plus** the College WQB Certification Request form.

For more information

Please feel free to contact the Undergraduate Curriculum Implementation office if you need clarification on any of these issues, and encourage your colleagues to contact us with questions or requests for help in articulating courses.

More information can be found on the UCI web page : www.sfu.ca/ugcr

Dr. Barbara Frisken, Chair, Undergraduate Curriculum Implementation Task Force
Phone: (604) 291-5767
E-mail: frisken@sfu.ca

Ms. Sarah Dench, Director, University Curriculum
Phone: (604) 268-6854
E-mail: sjdench@sfu.ca

Ms. Susan Rhodes, Program Coordinator, Undergraduate Curriculum Project
Phone: (604) 291-3312
E-mail: slrhodes@sfu.ca

See pages following for handouts of Powerpoint slides for *How to Use R in a First Course in Statistics* by Larry Weldon, Ph.D. (Simon Fraser University) and Kevin Keen, Ph.D., P.Stat. (University of Northern British Columbia).

How to Use R in a First Course in Statistics

Larry Weldon, Ph.D.
Simon Fraser University

and

Kevin Keen, Ph.D., P.Stat. (#43)
Univ. of Northern British Columbia

What is R?

- S was developed by John Chambers at Bell Labs, Murray Hill, NJ and others in the 1970's for the UNIX operating system
- S-Plus is a commercial version of S
- R was developed by Ross Ihaka and Robert Gentleman in the 1990's

2

What else about R?

- R is open-source software
 - GNU project licence
- Operating systems
 - UNIX
 - Windows XP
 - Mac OS X
- To download, go to
www.r-project.org
(use the SFU mirror)

3

Introductory Textbooks using R

- Verzani, John (2005) *Using R for introductory statistics*, Chapman Hall / CRC Press: Boca Raton.
 - ISBN: 1-58488-450-9
 - Can be used as a standalone textbook
 - I would use it as an R manual until an R supplement is available for Moore & McCabe (IPS)

4

Let's look at R

R

5

Isn't there a spreadsheet version of R?

- Yes, there are several
- On the basis of a number of ease of use features, the best appears to be:

Rcmdr

- Author: John Fox, McMaster University

6

Can I cut and paste from Rcmdr into a Powerpoint presentation?

- Yes!
- Examples slides follow

7

2.1 Scatterplots

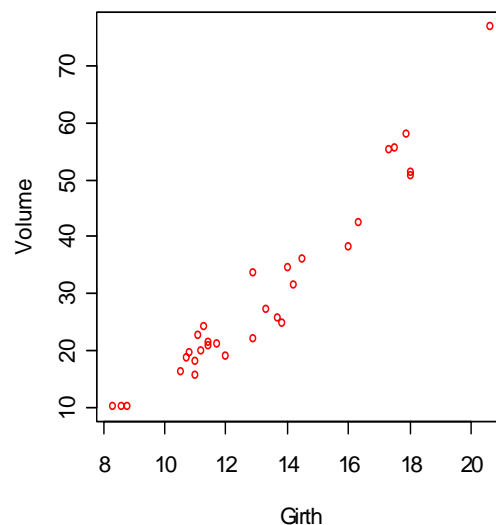
- A **scatterplot** shows the relationship between two quantitative variables measured on the same individuals
- Plot the explanatory variable x on the horizontal axis of the scatterplot
- Plot the response variable y on the vertical axis of the scatterplot
- Each individual in the data appears as a point in the scatterplot fixed by the values of both variables

8

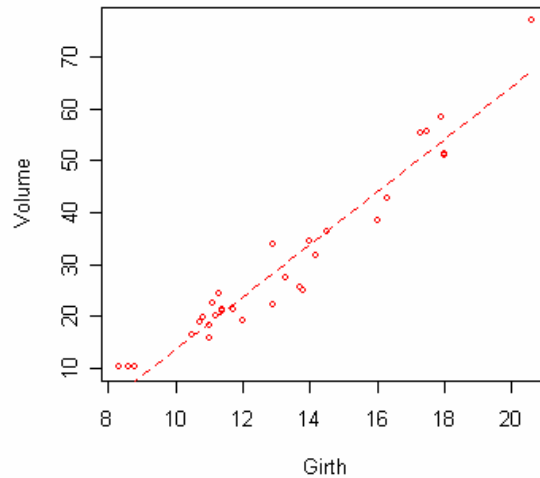
Example

- This data set provides measurements of the girth and volume of timber in 31 felled black cherry trees
- Black cherry trees are used in the manufacturer of wooden furniture
- Note that girth is the diameter of the tree (in inches) measured at 4 ft 6 in above the ground—diameter at breast height (DBH)

9



10



11

Statistical Package Output

```
Call:
lm(formula = Volume ~ Girth, data = trees)

Residuals:
    Min       1Q   Median       3Q      Max
-8.0654 -3.1067  0.1520  3.4948  9.5868

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -36.9435     3.3651  -10.98 7.62e-12 ***
Girth         5.0659     0.2474   20.48 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.252 on 29 degrees of freedom
Multiple R-Squared:  0.9353,    Adjusted R-squared:  0.9331
F-statistic: 419.4 on 1 and 29 DF,  p-value: < 2.2e-16
```

12



14

R for elementary stats

Is R for the novice?

On day 1, students can ...

- download and install R
- follow simplest commands
- run complex simulation programs
- re-run with new parameters

and

- understand what they have done!

Go to R ...

Is R easy to obtain?

- Free
- Easy to download (cran.r-project.org)

Does R help to teach stats?

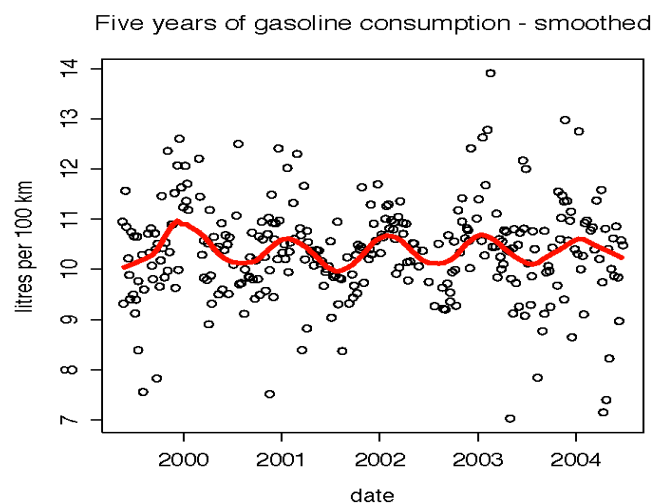
Some examples

- Time series - fuel consumption example
- Unbiasedness - theory
- Modeling - traffic example
- Graphical Summary - most examples!

Instructor's Software?

- Some programming is needed to demo concepts
- Use R
- Some programming is needed so students can do stats
- Use Minitab

Ex 1: A time series



Example 2: Unbiasedness

Normal Model

$$\hat{s}^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n + k)$$

n=10

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Expo Model

$$\hat{s}^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n + k)$$

n=10

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Ex 3: Traffic Accordion

- Simple Rule

Adjust speed to allow 2.5 seconds gap

(and add a little noise)

Uses only simple models.

[Go to R ...](#)

```
x=1:10
y=x^2
plot(x,y)
y=y+10*rnorm(10)
plot(x,y)
lm(y~x)
fit=-36.59+13.41*x
lines(x,fit)
plot(lm(y~x))
out=lm(y~x)
plot(out)
summary(out)

help(lm)

hist(rgamma(100,1))
hist(rgamma(100,2))
hist(rgamma(100,3))
hist(rgamma(100,10))

plot(density(rgamma(100,5)))

tse()
tse.run()
```



Instructions For Making 'Unit Roadmap' PowerPoint Slides

A 'Roadmap' is a type of conceptual map for structuring the content of a unit, where students first see the overall structure of the material and they drill down to find more detail as needed.

These slides explain how use PowerPoint to create a 'Roadmap' of a unit.

Run the Slide Show and click on the image below to view the contents of this file



1

Some macros are provided to assist in the creation of Roadmaps.

To use:

Run the Slide Show and click the button

MakeMenu

or

Tools > Macro > Macros > MakeMenu

Contents

- 1 [Title](#)
- 2 [Contents](#)
- 3 [Showing structure using Roadmaps: examples and suggestions](#)
- 6 [Slide layouts](#)
- 9 [Colours](#)
- 10 [The RoadmapTools menu](#)
 - Link Slides
 - Action Setting
 - Update Images
 - Insert New Slide
 - Make a new Roadmap file
 - Make a Roadmap master
 - Add Home Buttons
 - Add Next Slide Button
 - Add Previous Slide Button
 - Apply Title Layout
 - Apply Slide Layout
 - Apply Template to selected slides
 - Align Left
 - Evenly Space Slide Images
 - Promote Text
 - Demote Text
 - Move Below Title
 - Create Reference
 - Append Files
 - UnMake Menu
- 33 [Hints and Ideas](#)

This tool was developed by
Rodney Carr, Faculty of Business
and Law.

© Deakin University, 2005

2

Showing Structure Using Roadmaps

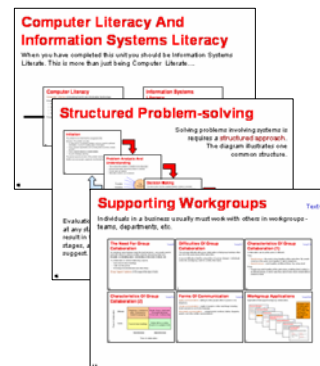
The normal linear arrangement of PowerPoint slides is the most common method for providing a structure of the material for students and in lectures.

However, hyperlinking of slides allows other structures that work just like normal web pages.

The simplest is an hierarchical structure. You can add visual cues such as clickable images to guide navigation.

The examples show the idea.

Examples: Run this slide show and click on the image of a slide. The actual slides start at slide 42.



[Suggestions for building Roadmaps](#)

3

Suggestions For Building Roadmaps (1)

1. If possible, start with a normal set of slides for a topic in your unit.
2. Make a Roadmap Master (see RoadmapTools menu) for the slides. This will probably mess up some of the formatting, but that can be fixed up later.
3. Most sets of slides are comprised of short sequences of slides that are related and have a common theme. Group these into a linked structure as follows:
 - Select the first slide of the sequence and RoadmapTools > Insert New Slide. This new slide will be the parent slide for the ones in the sequence.
 - Using the Slide viewer on the left of the screen in Normal View mode, or Slide Sorter View, Ctrl-click on each of the slides in the sequence and then on the parent slide and RoadmapTools > Link Slides.
 - Add a title to the parent and maybe some introductory text and arrange the images of the child slides as you want. It often helps to add arrows or other information to indicate how the slides (and the ideas) develop and to assist in navigation.

You now have a simple tree-structure that allows you to navigate from parent to any of the child slides when the Slide Show is run.

4

Suggestions For Building Roadmaps (2)

4. Repeat 3 throughout the file.
5. Now treat the parent slides as child slides – as appropriate, group them and link them to their own parents.
6. As you are roadmapping, look for themes and other structures. It might help to have another person with you as you are doing the roadmapping, so you can discuss possibilities. You will usually find that themes emerge as you think and explore the structure, so don't worry if you get the structure wrong first time. (When you do it wrong, simply delete any parent slides you don't want, and re-link the child slides to new parents.
7. When you have done all the separate topics, RoadmapTools > Make a new Roadmap file and create a parent for all the topics. This can be as simple as a table of contents, but you may find other themes to link the topics together.
8. You may wish to request assistance for editorial or graphics designer help when it comes to checking copyright or formatting the slides. You may find, however, that it is just as quick to do it much of the formatting yourself, especially since you will need to provide navigation aids like a particular layout of slides or use of arrows, etc, yourself. Some formatting tools are provided on the RoadmapTools menu.

Slide Layouts

<Ref>

This shows the layout of the standard title slide (Slide 7)



This shows the layout of a normal slide (Slide 8)



These are based on master slides that can be modified if required (View > Master).

It is often necessary to modify slides based on the above templates by adding textboxes or other objects. See the slides in the Examples.



Title Slide Layout

This format is used for the title slide in a topic, usually the first slide.

The format can be applied to any slide by selecting RoadmapTools > Apply
Title Layout



Normal Slide Layout

Text goes here



Colours

The slide title is red

Regular text is black

Comments are green

References are blue

Key words are brown and **bold**

Examples are orange

Questions are pink

These colours can be applied using the font colour icon in PowerPoint's Formatting toolbar



The RoadmapTools Menu

- [Link Slides](#)
- [Action Setting](#)
- [Update Images](#)
- [Insert New Slide](#)
- [Make a new Roadmap File](#)
- [Make Roadmap Master](#)
- [Add Home Buttons](#)
- [Add Next Slide Button](#)
- [Add Previous Slide Button](#)
- [Apply Title Layout](#)
- [Apply Slide Layout](#)
- [Apply Layout to selected slides](#)
- [Align Left](#)
- [Evenly Space Slide Images](#)
- [Promote Text](#)
- [Demote Text](#)
- [Move Below Title](#)
- [Create Reference](#)
- [Append Files](#)
- [UnMake Menu](#)

Link Slides

This is the most important tool in the package - it is the tool that allows you to build a roadmap by linking slides, with images of one or more child slides appearing on a parent slide to build up an hierarchical map of the unit or topic or module.

Link Slides (Cont)

When slides have been linked using the Link Slides tool, an image of the child slides appear on the parent, and a Return navigation button is added to each of the child slides, in the bottom right-hand corner of the slide.

When first linked, the images of the child slides are placed near in the middle of the parent. These should be moved (using click-and-drag) to a suitable location on the slide. They can also be resized by using one of the corner resize handles (don't use the resize handles on the edge - that alters the aspect ratio). Use Ctrl-click to select more than one slide at a time; they can then be moved and resized together.

If you wish to use a text-based hyperlink to another slide (Roadmap Tools > Action Settings > Hyperlink to > Slide), you can still use the Link slides tool to place the Return navigation button on the child slides, then delete any unwanted images of the child slides on the parent.

To use

- View the slides in the PowerPoint file in Slide Sorter mode, or use the slide viewer on the left of the screen in Normal mode
- Hold down the Ctrl key and click on each child slide in turn and then on the slide that is to be the parent of the child slides
- Select RoadmapTools > Link Slides

Link Slides (Cont)

When slides have been linked using the Link Slides tool, an image of the child slides appear on the parent, and a Return navigation button is added to each of the child slides, in the bottom right-hand corner of the slide.

When first linked, the images of the child slides are placed near in the middle of the parent. These should be moved (using click-and-drag) to a suitable location on the slide. They can also be resized by using one of the corner resize handles (don't use the resize handles on the edge - that alters the aspect ratio). Use Ctrl-click to select more than one slide at a time; they can then be moved and resized together.

If you wish to use a text-based hyperlink to another slide (Roadmap Tools > Action Settings > Hyperlink to > Slide), you can still use the Link slides tool to place the Return navigation button on the child slides, then delete any unwanted images of the child slides on the parent.

Create Link

Used to set a hyperlink to the selected object or some text on a slide. Same as Slide Show > Action Settings...

To use

View a slide in Normal mode and click on an object or select some text to which you wish to apply a hyperlink

RoadmapTools > Action Setting

To link to another slide in the current file:

- click 'Place in This Document'
- select the slide to link to
- OK

To link to another file

- click 'Existing File or Web Page'
- select the file to link to
- if you want, click Bookmark to see the places



Update Images

Use Update Images to refresh the images of the child slides if you alter one or more of the child slides in a roadmap. This actually refreshes all the images of slides starting at end of the file and working backwards, so if you arrange the slides so that child slides appear after any parent slide, then any parents-of-parents in the roadmap will also be appropriately updated.

To use

View in Normal mode the file you wish to update

RoadmapTools > Update Images



Insert New Slide

Same as Insert > New Slide

15



Make A New Roadmap File

This creates a new blank file with two slides, a title slide and one normal slide. This is used when you are starting a new file rather than basing it on a previous PowerPoint file.

To use

RoadmapTools > Create Roadmap file

16



Make A Roadmap Master

Slides used for roadmaps normally need to have a plain background and you normally need to make the most of the screen real estate. The 'Create Roadmap master' tool creates a master slide with a simple structure that can be applied to each slide in an existing PowerPoint file, ready for it to be 'Roadmapped'. The tool also provides you with the standard colours for a Roadmap described on Slide 7 (see Contents). It also removes all animation effects from each slide. The 'Create Roadmap template' tool should usually be applied to an existing PowerPoint file before linking any slides.

To use

View the file to whose slides you wish to roadmap

RoadmapTools > Create Roadmap file

17

Add Home Buttons

Most roadmap files are multilayered, with images of child slides appearing on parent slides, which then are themselves child slides appearing on other parent slides. If you construct such a multi-layered file, it is sometimes useful to add a home button to each slide that links back to a Home slide. The Home slide can be in another PowerPoint file, thus allowing you to create a Home page for use in a whole unit with many separate topics.

To use

View the file to whose slides you wish to add home buttons

RoadmapTools > Add Home Buttons

You will be prompted to select the name of the file containing the Home slide (default is the current file), and the number of the home slide in that file (default is slide 1)

18

Add Next Slide Button

If you have a parent slide with a number of child slides, usually it is best to navigate from one child slide to the next by returning to the parent using the Return button and then clicking on the image of the next child slide. However, in some situations you might wish to allow users to go from directly from one child slide to another. This might be appropriate if you have two child slides that are closely related, maybe where the information you wish to provide will not all fit on the first slide. Users can use the normal navigation tools in PowerPoint (eg PgDn to view to the next slide) but you can also provide a 'Next slide' navigation button to guide users.

To use

View in Normal mode the slide to which you wish to add the next button

RoadmapTools > Add Next Slide Button

Add Previous Slide Button

This tool places a 'Previous slide' navigation button. These can be used to link child slides - see 'Add Next Slide Button'.

To use

View in Normal mode the slide to which you wish to add the Previous slide button

RoadmapTools > Add Previous Slide Button

Apply Title Layout

Sometimes you want to apply the Title slide master layout to an individual slide – this tool allows you to do that.

To use

View in Normal mode the slide to which you wish to apply the Title Layout

RoadmapTools > Apply Title Layout

21

Apply Slide Layout

This tool allows you to apply the layout for a normal slide.

To use

View in Normal mode the slide to which you wish to apply the normal slide Layout

RoadmapTools > Apply Slide Layout

22

Apply Layout To Selected Slides

This tool applies the Roadmap Layout to a number of selected slides. It is typically used where an existing PowerPoint file is to be roadmapped. The 'Apply Template to Selected Slides' tool will apply the normal slide master template to all selected slides other than Slide 1.

To use

Select the slides you wish to apply the Roadmap layout to in Normal mode, or select a number of slides from the Slide Sorter.

RoadmapTools > Apply Layout to Selected Slides

23

Align Left

Aligns the selected object to the left margin

24

Evenly Space Slide Images

Places the selected slides so they are spaced evenly between the first and last selected ones.

To use

Select the images of the slides you wish to align (use Ctrl-click to select one after another). The first and last ones you select are not moved, and the ones in between are spaced evenly between these.

RoadmapTools > Evenly Space Slide Images

25



Move Below Title

This moves the selected object (usually a text box) to a position just below the title.

To use

Select a textbox containing the text to use for the reference

RoadmapTools > Move Below Title

26



Promote Text

This undents the selected text slightly, usually for a sublist using dotpoints or numbering. Same as Promote on the PowerPoint's Outlining Toolbar.

To use

Select the text to promoted

RoadmapTools > Promote Text

Demote Text

This indents the selected text slightly, usually for a sublist using dotpoints or numbering. Same as Demote on the PowerPoint's Outlining Toolbar.

To use

Select the text to promoted

RoadmapTools > Demote Text

Create Reference

You may wish to provide a reference to a textbook on a particular slide. Such references appear near the top right on the slide. The 'Create Reference' tool will create a textbox, placed and formatted appropriately, for you to provide such information. If you select a textbox first, the tool will also copy the text in the box and place it into the new reference textbox.

To use

If you wish, select a textbox containing the text to use for the reference

RoadmapTools > Create Reference

29

Create Index

Not completed yet, sorry

30

Append Files

Sometimes you may want to combine a number of separate roadmapped files into one larger file. For example, if you have a multi-topic set of files that you want to convert to web pages, or to some other format, you will need to first combine all the files otherwise links between the files will not work. The 'Append Files' achieves this by appending files to an existing file, usually the Home file.

To use

View the Home file (to which you wish to append other files)

RoadmapTools > Append Files. You will be prompted to select the file to append.

You can only append one file at a time. When appending files it is a good idea to save the merged file after each append operation. The file can end up quite big, and PowerPoint has a tendency to freeze up.

If you have a slide with an activeX control, make it have some code, even if it does nothing (there is a bug in PowerPoint that will cause the macro to stop).

31

Unmake Menu

When you open RoadmapTools and run 'Make menu', the RoadmapTools menu appears on PowerPoint's main menu. Unmake menu removes the menu item. Closing PowerPoint and re-opening removes it, too.

To use

RoadmapTools > Unmake Menu

32

Hints And Ideas

- [Creating a link to another slide](#)
- [Use of action buttons for navigation](#)
- [Creating new slides](#)
- [Use of textboxes](#)
- [Paste Special > Unformatted text](#)
- [Resizing text](#)
- [Use of Dotpoints](#)
- [Create a customized toolbar](#)

33

Creating A Link To Another Slide

Select the object that you wish to link from (so that when the user clicks the object the link is followed). It can be the image of a slide or a picture, a word or any other object on a slide.

- Roadmap Tools > Action Settings (or Right-click, select 'Action Settings').
- Select 'Hyperlink to'.
- Use the drop-down box to select 'Slide'
- Click on the slide to which you wish to link.

You can use the Link slides tool in the RoadmapTools menu to create the link back to the parent.

34

Use Of Action Buttons For Navigation

The action buttons on this slide are created by the following Roadmap Tools:



(Return button) – Link Slides



(Next Slide) - Add Next Slide Button



(Previous Slide) - Add Previous Slide Button



(Home) - Add Home Buttons

You can also copy a button from any slide, paste it onto another slide (it'll appear in the same position). If you set a link and then copy the button, the links will be copied, too.

Creating New Slides

To create a new slide, it is often best to copy an existing slide and modify that (rather than use Insert > New slide):

Slide Sorter

Click on a slide that has similar features to the one you want to create

Edit > Copy

Click where you want the new slide to go

Edit > Paste

Double-click on the new slide and edit it.

Use Of Textboxes

The default slide has only one textbox, with set margins and indenting, etc. You can create other textboxes that you can format as desired, and that can be placed on the slide as you wish.

It is often best to copy an existing textbox rather than insert a new one from the drawing toolbar. If you select one on a slide and Ctrl-Shift drag it'll create a copy that is lined up with the old one.

It is often a good idea to place text (maybe different paragraphs) into a number of different textboxes on a slide. That way you can move them around, resize as required, set formatting, etc without affecting all the text.

37

Paste Special > Unformatted Text

If you copy text from somewhere, use Paste Special > Unformatted text when pasting it into a textbox on a slide to avoid having PowerPoint apply the default formatting (which is often not what you want).

38

Resizing Text

Text will sometimes automatically resize in a textbox, but if you want to set the size, select the text and use Ctrl-Shift > or Ctrl-Shift <

Use of Dotpoints

Usually, don't use dotpoints on the top level paragraphs.

If you want to use dotpoints for items following a paragraph:

- Type the items on separate rows
- Select the rows
- Click the Roadmap Tools > Demote Text

With the normal textboxes in a Roadmap file, the demoted text will have dotpoints. If it doesn't, click on dotpoints icon on PowerPoint's formatting toolbar, or Format > Bullets and Numbering.

Create A Customized Toolbar

There are some operations that you will carry out many times while roadmapping, and you might wish to create a customised toolbar containing some common tools, or add them to an existing toolbar (eg Formatting).

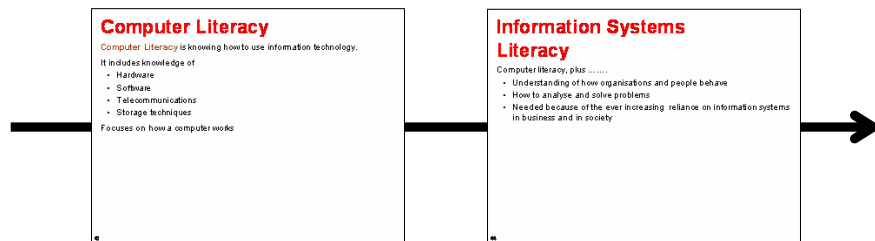
Tools > Customize (> New if you wish to create a new toolbar)

Commands > Edit

- Locate Paste Special and drag it to your toolbar
- Locate the 'Bring to front' and 'Send to back' and drag onto your toolbar
- Click on Macros and add LinkSlides, AddNextButton, AddPreviousButton, etc

Computer Literacy And Information Systems Literacy

When you have completed this unit you should be Information Systems Literate. This is more than just being Computer Literate....



Computer Literacy

Computer Literacy is knowing how to use information technology.

It includes knowledge of

- Hardware
- Software
- Telecommunications
- Storage techniques

Focuses on how a computer works

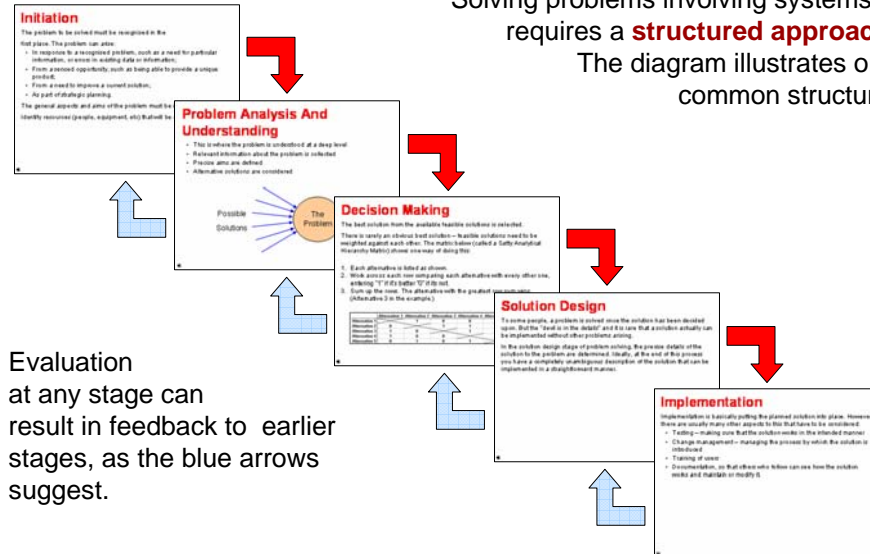
Information Systems Literacy

Computer literacy, plus

- Understanding of how organisations and people behave
- How to analyse and solve problems
- Needed because of the ever increasing reliance on information systems in business and in society

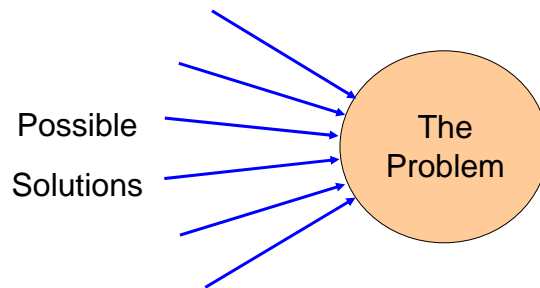
Structured Problem-solving

Solving problems involving systems is requires a **structured approach**.
The diagram illustrates one common structure.



Problem Analysis And Understanding

- This is where the problem is understood at a deep level
- Relevant information about the problem is collected
- Precise aims are defined
- Alternative solutions are considered



47

Decision Making

The best solution from the available feasible solutions is selected.

There is rarely an obvious best solution – feasible solutions need to be weighted against each other. The matrix below (called a Satty Analytical Hierarchy Matrix) shows one way of doing this:

1. Each alternative is listed as shown.
2. Work across each row comparing each alternative with every other one, entering "1" if it's better "0" if its not.
3. Sum up the rows. The alternative with the greatest row sum wins. (Alternative 3 in the example.)

	Alternative 1	Alternative 2	Alternative 3	Alternative 4	Alternative 5	Row sum	Rank
Alternative 1		1	0	0	1	2	2
Alternative 2	0		1	1	0	2	2
Alternative 3	1	0		1	1	3	1
Alternative 4	1	0	0		0	1	5
Alternative 5	0	1	0	1		2	2

48

Solution Design

To some people, a problem is solved once the solution has been decided upon. But the “devil is in the details” and it is rare that a solution actually can be implemented without other problems arising.

In the solution design stage of problem solving, the precise details of the solution to the problem are determined. Ideally, at the end of this process you have a completely unambiguous description of the solution that can be implemented in a straightforward manner.

Implementation

Implementation is basically putting the planned solution into place. However, there are usually many other aspects to this that have to be considered:

- Testing – making sure that the solution works in the intended manner
- Change management – managing the process by which the solution is introduced
- Training of users
- Documentation, so that others who follow can see how the solution works and maintain or modify it.

Supporting Workgroups

Text Ch9

Individuals in a business usually must work with others in workgroups - teams, departments, etc.

The Need For Group Collaboration

As individuals cannot always make the best decisions – we usually need to work with others to achieve a better outcome, discussing ideas and thoughts, coordinating plans, commenting on the work of others, etc.

To collaborate on common tasks they need to:

- Have face-to-face meetings
- Talk over the phone
- Exchange documents (text and other files)

Group Support Systems (GSS) support this type of work.

Difficulties Of Group Collaboration

The principal difficulty with group collaboration is that group members often are not in the same place at the same time.

A second difficulty is that the composition of a group changes – individuals come into a workgroup, work for a while, then leave.

Characteristics Of Group Collaboration (1)

Collaboration can be at the same or different:

Time

- Synchronous – the work is done together at the same time. Eg. people could be in the same room together or talk by telephone.
- Asynchronous – work together at different times. Eg. using email.

Place

- People may work together at the same place, working direct contact, or at different places, in which case they communicate direct contact without extensive travel.

Characteristics Of Group Collaboration (2)

Place of collaboration	Others	Telephone, Conference calls, Chatrooms, Instant messaging	Email, Fax, Mail, Overnight deliveries, Voice mail, Conferencing
	Same	Face-to-face meetings	Notes left on a notice board or in a pigeon hole
		Synchronous	Asynchronous

Time of collaboration

Forms Of Communication

Audio communication – talking to other people either in person or by telephone.

Visual communication – sights of people or other real things including facial expressions and body language.

Document communication – containing text, numbers, tables, diagrams, graphs, and other written representations.

Workgroup Applications

Applications that support workgroup collaboration:

Example: See Collaborative Applications for a list of these applications in a table.

51

The Need For Group Collaboration

Text p276

An individual cannot always make the best decisions – we usually need to work with others to achieve a better outcome, discussing ideas and thoughts, coordinating plans, commenting on the work of others, etc.

To collaborate on common tasks they need to

- Have face-to-face meetings
- Talk over the phone
- Exchange documents (text and other files)

Group Support Systems (GSS) support this type of work.

52

Difficulties Of Group Collaboration

Text p276-7

The principal difficulty with group collaboration is that group members often are not in the same place at the same time.

A second difficulty is that the composition of a group changes – individuals come into a workgroup, work for a while, then leave.

53

Characteristics Of Group Collaboration (1)

Text p277-8

Collaboration can be at the same or different:

Time

- **Synchronous** – the work is done together at the same time. Eg: people could be in the same room together or talk by telephone.
- **Asynchronous** – work together at different times. Eg: using email

Place

- People may work together at the same place, enabling direct contact, or at different places, in which case they cannot have direct contact without extensive travel.

54

Characteristics Of Group Collaboration (2)

Text p277-8

Place of collaboration	Different	Telephone, Conference calls, Chat session, Instant messaging	Email, Faxes, Snail mail, Overnight deliveries, Voice mail, Conferencing
	Same	Face-to-face meetings	Notes left on a notice board or in a pigeon hole
		Synchronous	Asynchronous
Time of collaboration			

55

Forms Of Communication

Text p279

Audio communication – talking to other people either in person or by telephone.

Visual communication – sights of people or other real things including facial expressions and body language.

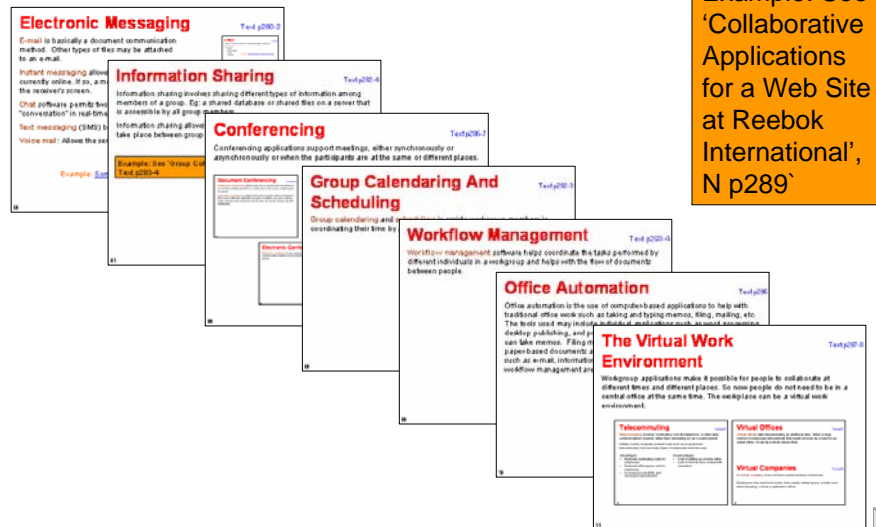
Document communication - containing text, numbers, tables, diagrams, graphs, and other written representations.

56

Workgroup Applications

Text p279-96

Applications that support workgroup collaboration:



57

Electronic Messaging

Text p280-2

E-mail is basically a document communication method. Other types of files may be attached to an e-mail.

Instant messaging allows the sender to check if the intended recipient is currently online. If so, a message can be sent that immediately appears on the receiver's screen.

Chat software permits two or more people to have an electronic "conversation" in real-time.

Text messaging (SMS) between two mobile phones is instant.

Voice mail: Allows the sender to leave short audio messages.

Example: [Some Text messaging \(txt_messages.doc\)](#)



58

e-Mail

Text p280-2

e-Mail is a common method for communicating within a workgroup.

Some issues:

- Spam
- Global emails
- e-Mail overload

Example: [eMail Overload Is a Myth, Study Says](#)



Cc
Bcc
Subject
Attachments

59

e-Mail Overload

"e-Mail Overload Is a Myth, Study Says"

Washington Post (12/09/02) P. E5; Henry, Shannon

The Pew Internet & American Life Project says not as many U.S. workers feel overwhelmed by email as is widely thought. In all, around 60% of 1,003 workers surveyed by telephone reported getting 10 or fewer email messages per day. Just 4% of all people who use email at work reported receiving too many messages every day, while 65% of those interviewed said email was not a problem whatsoever. Senior Pew researcher Deborah Fallows said she was surprised by the results, and surmised that a small but vocal minority of email users are what she calls "power emailers." According to the study, 11% of those heavy users felt their inboxes were deluged with email everyday. Gartner analyst Maureen C. Grey said the Pew study was suspect, given the high profile of email growth and the success of email management software, such as the anti-spam software from Brightmail. That company reports that spam email totaled 35% of all incoming messages as of September 2002, but that it was just 8% of incoming messages one year earlier. However, spammers target individual users of free email accounts such as those from Yahoo!, AOL, and Microsoft, since addresses are easy to guess. Corporate email accounts have become harder targets for spammers because of increasingly spam-savvy CIOs and better email management software, according to Pew.

<http://www.washingtonpost.com/wp-dyn/articles/A24684-2002Dec7.html>

60

Information Sharing

Text p282-4

Information sharing involves sharing different types of information among members of a group. Eg: a shared database or shared files on a server that is accessible by all group members.

Information sharing allows audio, visual, and document communication to take place between group members at different times and different places.

Example: See 'Group Collaboration for the Sable Offshore Energy Project', Text p283-4

Conferencing

Text p286-7

Conferencing applications support meetings, either synchronously or asynchronously or when the participants are at the same or different places.

Document Conferencing

Textp286-8

Whitboard conferencing allows each user to see the same document on an electronic whiteboard which is a white area on the screen containing the document.

Application conferencing allows each user to see the same document on their screen within the application program. In addition, any user is able to make changes to the document, and all users can see the change. Eg: MS NetMeeting.

Audioconferencing

Textp286-7

Computer telephony allows telephone-like communication to be carried out using a computer over a computer network (eg the internet).

Audioconferencing is like a conference call on a telephone, except that it uses computer telephony.

Videoconferencing

Textp287-90

Videoconferencing allows audio and visual communication between group members at the same time from different places.

Electronic Conferencing

Textp289

Electronic conferencing can combine audio, visual, and document communication between group members at the same time from different places.

Electronic Meeting Support

Textp289-92

Electronic Meeting Systems (EMS) permit using computer systems to facilitate face-to-face meetings or meetings that involve people in different places. An advantage of EMS is that contributions can be anonymous. They can also allow users to vote on particular issues, thus providing a group decision support system (GDSS).

Document Conferencing

Text p284-6

Whiteboard conferencing allows each user to see the same document on an electronic whiteboard which is a white area on the screen containing the document.

Application conferencing allows each user to see the same document on their screen within the application program. In addition, any user is able to make changes to the document, and all users can see the change. Eg: MS NetMeeting.

Audioconferencing

Text p286-7

Computer telephony allows telephone-like communication to be carried out using a computer over a computer network (eg the internet).

Audioconferencing is like a conference call on a telephone, except that it uses computer telephony.

Videoconferencing

Text p287-90

Videoconferencing allows audio and visual communications between group members at the same time from different places.

Electronic Conferencing

Text p290

Electronic conferencing can combine audio, visual, and document communication between group members at the same time from different places.

Electronic Meeting Support

Text p291-2

Electronic Meeting Systems (EMS) permit using computer systems to facilitate face-to-face meetings or meetings that involve people in different places. An advantage of EMS is that contributions can be anonymous. They can also allow users to vote on particular issues, thus providing a group decision support system (GDSS).

67

Group Calendaring And Scheduling

Text p292-3

Group calendaring and **scheduling** is assists workgroup members in coordinating their time by allowing people to see each others schedules.

68

Workflow Management

Text p293-4

Workflow management software helps coordinate the tasks performed by different individuals in a workgroup and helps with the flow of documents between people.

Office Automation

Text p296

Office automation is the use of computer-based applications to help with traditional office work such as taking and typing memos, filing, mailing, etc. The tools used may include individual applications such as word processing, desktop publishing, and presentation graphics. Voice-recognition software can take memos. Filing might be done electronically by scanning any paper-based documents and/or using a database. Workgroup applications such as e-mail, information sharing, calendaring and scheduling, and workflow management are also used.

The Virtual Work Environment

Text p297-8

Workgroup applications make it possible for people to collaborate at different times and different places. So now people do not need to be in a central office at the same time. The workplace can be a virtual work environment.

Telecommuting

Text p297

Telecommuting involves "commuting" over the telephone, or other data communications channel, rather than commuting by car or public transit.

Initially, mainly computer professionals such as programmers telecommuted, but now many types of employees work this way.

Advantages:

- Reduced commuting costs for employees
- Reduced office space cost for employers
- Increased productivity and decreased absenteeism

Disadvantages:

- Cost of setting up a home office
- Lack of face-to-face contact with coworkers

Virtual Offices

Text p297

Virtual offices take telecommuting an additional step. When a large number of employees telecommute, their might not even be a need for an actual office — it can be a whole virtual office.

Virtual Companies

Text p298

A virtual company does not have a physical place of business.

Employees may work from home, from nearby rented space, a hotel room when traveling, or from a customer's office.

Telecommuting

Text p297

Telecommuting involves "commuting" over the telephone, or other data communications channel, rather than commuting by car or public transit.

Initially, mainly computer professionals such as programmers telecommuted, but now many types of employees work this way.

Advantages:

- Reduced commuting costs for employees
- Reduced office space cost for employers
- Increased productivity and decreased absenteeism

Disadvantages:

- Cost of setting up a home office
- Lack of face-to-face contact with coworkers

Virtual Offices



Text p297

Virtual offices take telecommuting an additional step. When a large number of employees telecommute their might not even be a need for an actual office – it can be a whole virtual office.

Virtual Companies

Text p298

A **virtual company** does not have a physical place of business.

Employees may work from home, from nearby rented space, a hotel room when traveling, or from a customer's office.





Chance and Data Analysis

Case Studies



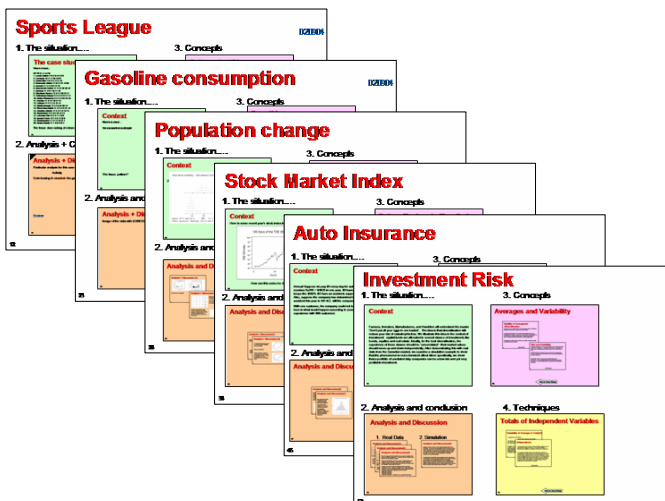
Concepts



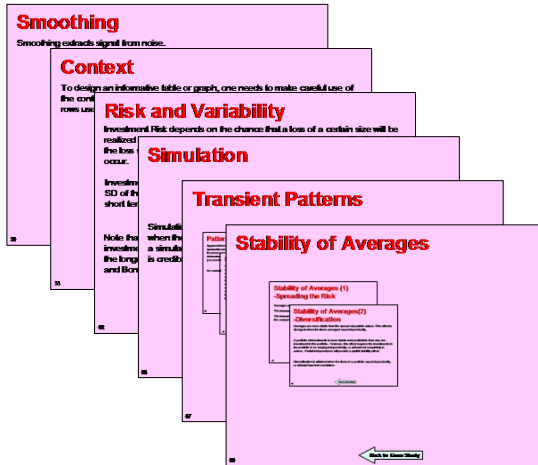
Techniques



Case Studies



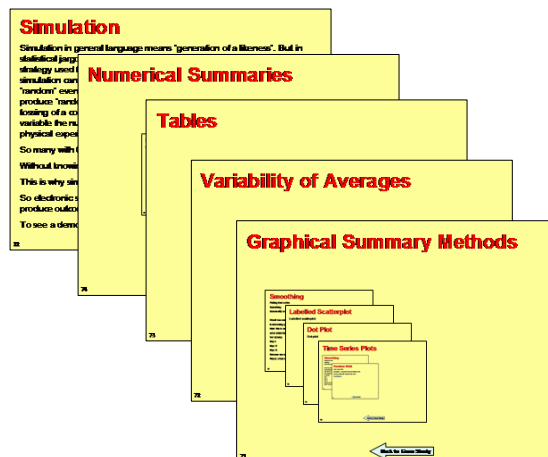
Concepts



3



Techniques



4



Stock Market

1. The situation....

Context

Here is some recent year's stock index levels for the Toronto Stock Market.



How can this series be described?

2. Analysis and conclusion

Analysis and Discussion



3. Concepts

Pattern Illusions in Time Series

Apparent trends can be useless for prediction, as is the case in the symmetric random walk – the trend may be useful to guide your actions, but the trend up or down may not persist. It takes a long time series to determine whether trends are real or illusions, and even in a long time series you need some stability in the mechanism to infer anything.

An example of this is in [cars.html.xls](#)

4. Techniques

Random Walk

Time Series Plot

Simulation – Symmetric Random Walk ($p=0.5$)

Random Walk with Variable Step Sizes

[RandomWalk.xls](#)

5

Effects can be illusions

Patterns that Disappear

Some patterns in data are transient, and some are persistent. The transient ones are called illusions. The cause of these illusions is unexplained variation, and it can lead to misinterpretation of data by the statistically naïve. Patterns in data can appear that seem too regular to be transient, but this can be an illusion. By learning what can happen when we create a model that has no useful pattern, we can guard against being fooled by an apparent pattern. Of course, if they are transient and persistent techniques of statistics

Pattern Illusions in Time Series

Apparent trends can be useless for prediction, as is the case in the symmetric random walk – the trend may be useful to guide your actions, but the trend up or down may not persist. It takes a long time series to determine whether trends are real or illusions, and even in a long time series you need some stability in the mechanism to infer anything.

An example of this is in [cars.html.xls](#)

6

Probability

Stability of Averages (1) -Spreading the Risk

Averages are more stable than the spread of possible values.
The insurance company can spread the risk across many policy holders.
The insurance company sells this service to policy holders — which is why
the company takes in more than it pays out.

59

Back to Course Menu

7



Smoothing

Smoothing

Smoothing extracts signal from noise.

60

Back to Course Menu

8



Make use of the Context of the problem

Context

To design an informative table or graph, one needs to make careful use of the context of the data. (labelling plotted points by country, ordering table rows usefully).

38

Back to Class Weekly

9



Dot Plots

Dot Plot

Dot-plot

36

Back to Class Weekly

10



Tables

Ordering Rows or Columns in a Table

Table
The general idea is that if a feature of a display is arbitrary, it may sometimes be re-organized to advantage.

17

Back to Table of Contents

11



Sports League

020904

1. The situation....

The case study

West & Rutter

1. Lovers (11 6 5 8 15 129)
2. Lovers (11 6 5 8 15 129)
3. Lovers (11 6 5 8 15 129)
4. Lovers (11 6 5 8 15 129)
5. Lovers (11 6 5 8 15 129)
6. Lovers (11 6 5 8 15 129)
7. Lovers (11 6 5 8 15 129)
8. Lovers (11 6 5 8 15 129)
9. Lovers (11 6 5 8 15 129)
10. Lovers (11 6 5 8 15 129)
11. Lovers (11 6 5 8 15 129)
12. Lovers (11 6 5 8 15 129)
13. Lovers (11 6 5 8 15 129)
14. Lovers (11 6 5 8 15 129)
15. Lovers (11 6 5 8 15 129)
16. Lovers (11 6 5 8 15 129)
17. Lovers (11 6 5 8 15 129)
18. Lovers (11 6 5 8 15 129)
19. Lovers (11 6 5 8 15 129)
20. Lovers (11 6 5 8 15 129)

The focus: does ranking of a team in a sports league reflect the quality of the team?

18

3. Concepts

Patterns that Disappear

Some patterns in data are transient, and some are persistent. The transient ones are called *illusions*. The cause of these illusions is unexplained variation, and it can lead to misinterpretation of data by the statistical model. Patterns in data can appear that seem too regular to be transient, but this can be an illusion. By learning what can happen when we create a model that has no useful pattern, we can guard against being fooled by an apparent pattern. Of course, we do want to find those patterns that last, are not transient, if they exist. But it can be very difficult to distinguish between transient and persistent patterns. This dilemma motivates many of the techniques of statistics.

19

Back to Table of Contents

2. Analysis + Conclusion

Analysis + Discussion

Particular analysis for this case

Discussion about mean & sd

Activity

Coin tossing to simulate the games

Review

20

4. Techniques

Means and Standard Deviations

Calculation of the SD

Using the Mean and SD

Mean and Standard Deviation

Mean and Standard Deviation

21

Back to Table of Contents

12



The case study

What is it about...

GP W D L F A Pts

1. Leeds United 11 6 5 0 15 5 23
2. Liverpool 10 7 1 2 20 10 22
3. Aston Villa 11 6 3 2 17 11 21
4. Newcastle United 11 6 2 3 21 14 20
5. Arsenal 11 5 4 2 24 13 19
6. Manchester United 11 5 3 3 28 20 18
7. Chelsea 11 4 6 1 16 11 18
8. Blackburn Rovers 12 4 5 3 20 15 17
9. Tottenham Hotspur 12 5 2 5 19 17 17
10. Bolton Wanderers 12 4 4 4 16 16 16
11. Everton 11 4 3 4 18 17 15
12. Fulham 11 3 5 3 12 12 14
13. Middlesbrough 12 4 2 6 16 20 14
14. West Ham United 11 4 2 5 12 21 14
15. Charlton Athletic 11 3 4 4 12 13 13
16. Sunderland 12 3 4 5 10 14 13
17. Leicester City 12 2 3 7 7 23 9
18. Ipswich Town 12 1 5 6 13 20 8
19. Southampton 11 2 1 8 10 21 7
20. Derby County 11 1 4 6 9 22 7

The Issue: does ranking of a team in a sports league reflect the quality of the team?

13



Analysis + Discussion

Particular analysis for this case

Discussion about mean & stdev

Activity

Coin tossing to **simulate** the games

[Review](#)

14



Analysis + Discussion

Analysis + Discussion

Particular analysis for this case

Activity

Coin tossing to simulate the games

[Review](#)

1

chjsdak cjdakjdxsa hdjksl

15



Review

1. Influence of Unexplained Variation (UV) on Interpretation of Data
2. UV can make temporary effects seem like permanent ones (illusions)
3. Graphing of Data is an essential first step in data analysis
4. Need for summary measures when UV present

MCQ

What is first thing you do with data?

- a. Collect it ☐
- b. Graph it ☐
- c. Decide what type of data it is ☐

16



Not the best answer

This is not the best answer....

17



Good answer

You got the answer....

Have you thought about...

☐ CheckBox1

☒ CheckBox1

☐ CheckBox1

18



Patterns that Disappear

Some patterns in data are transient, and some are persistent. The transient ones are called illusions. The cause of these illusions is unexplained variation, and it can lead to misinterpretation of data by the statistically naïve. Patterns in data can appear that seem too regular to be transient, but this can be an illusion. By learning what can happen when we create a model that has no useful pattern, we can guard against being fooled by an apparent pattern. Of course, we do want to find those patterns that last, are not transient, if they exist. But it can be very difficult to distinguish between transient and persistent patterns. This dilemma motivates many of the techniques of statistics.

19

[Back to Case Study](#)



Using the Mean and SD

Using mean & standard deviation (SD)

The “mean” is just a synonym for “arithmetic average” – the usual one found by adding up a batch of numbers and dividing the total by the number of numbers in the batch. It gives a reasonable one-number summary of the batch. Of course, it is not a great summary of the batch! We need at least to describe the spread of the batch of numbers. The usual measure of spread is the “standard deviation” or SD. Think of it as a typical deviation of the numbers from their mean. Before we give the formula, here an example: your male classmates probably average about 178cm in height and the SD is about 6 cm. Although the two numbers 178 and 6 do not say exactly what the collection of heights is, it does give a rough idea.

So the numbers mean and SD do give a convenient numerical description of a batch of numbers.

20

[Back to Case Study](#)



Calculation of the SD

Standard Deviation (SD) – how to compute it.


Suppose you want the SD of n numbers:

The SD is based on deviations of the n numbers from the mean. What you do is take these n deviations, square them, sum them up, divide the sum by n, and finally take the square root.

Example: 1,2,3,5,9 is our batch. Mean

is $20/5=4$. SD is $\sqrt{[(1-4)^2 + (2-4)^2 + (3-4)^2 + (5-4)^2 + (9-4)^2]/5} = 2.83$

21

 Back to Case Study



Simulation

Simulation in general language means “generation of a likeness”. But in statistical jargon it is short for Monte Carlo simulation which is a particular strategy used to explore the implications of probability models. This simulation can be physical (making use of coins or dice or cards to produce “random” events) or electronic (making use of a computer algorithm to produce “random” numbers.) An example of a physical model would be the tossing of a coin 10 times, many groups of 10 tosses, to find out how variable the number of heads in 10 tosses is. The result of 100 such physical experiments would be a distribution of the number of heads:

So many with 0 heads, so many with 1 headso many with 10 heads.


Without knowing any theory of probability, you could actually get the result.

This is why simulation is useful. However, the tossing of coins is laborious.

So electronic simulation is a very welcome alternative. The computer can produce outcomes with the same properties as the physical experiment.

To see a demonstration of this, click on [coin.xls](#).

22

 Back to Case Study



Gasoline consumption

020904

1. The situation....

Context

What is it about...

The unsmoothed scatterplot

The Issue: pattern?

3. Concepts

Smoothing

Smoothing extracts signal from noise.

2. Analysis and conclusion

Analysis + Discussion

Image of the data with LOWESS smoothing

Possible causes

Year	Year	Year
1984	1985	1986
1987	1988	1989
1990	1991	1992
1993	1994	1995
1996	1997	1998
1999	2000	2001
2002	2003	2004
2005	2006	2007
2008	2009	2010
2011	2012	2013
2014	2015	2016
2017	2018	2019
2020	2021	2022

4. Techniques

Smoothing

Plotting Time series

Smoothing

Seasonality vs trends

Plotting Gas consumption data. The graph suggested the seasonality, and this led to identifying specific years when the consumption of gas was unusually high. This is an example of a scatter plot. Two "variables" are plotted for a data set to which the rows of the data are listed:

Year	Gas (liters per kg)
1984	1984.5
1985	1985.7
1986	1986.5
1987	1987.5

Because one of the variables is "Time", this kind of data is called a time series. Why is a time series different from other kinds of data?

23



Context

What is it about...

The unsmoothed scatterplot

The Issue: pattern?

24



Analysis + Discussion

Image of the data with LOWESS smoothing

Possible causes

jfkdsajfkldjask fjkdasljklksa

25



Smoothing

Smoothing extracts signal from noise.

← Back to Case Study

26



- Plotting time series
- Smoothing
- Seasonality vs trends

Recall Gas consumption data. The graph suggested the seasonality, and this led to interesting questions about the cause of the seasonality.

Note: this is an example of a scatter plot. Two “variables” are plotted for a data set in which the rows of the data are linked:

Var 1(Date)	Var 2(Miles per Km)
May 5	1999 6.5
May 12	1999 6.7
May 15	1999 5.9

Because one of the variables is "Time", this kind of data is called a time series.

Why is a time series different from other kinds of data?

k,mmk

k,mmk

Back to Case Study

27

Population change

1. The situation....

Context

More about variability: vital statistics of 69 countries - rates per 1000 population

2

Birth-rt

death-rt

means: 23.3 and 10.3 SDs: 11.7 and 4.8
Other measures: Medians (23.0 and 9.1) Range (45.0 and 29.1)

3. Concepts

Context

To design an informative table or graph, one needs to make careful use of the context of the data. (labeling plotted points by country, ordering table rows *unusually*).

2. Analysis and conclusion

Analysis and Discussion

4. Techniques

Techniques

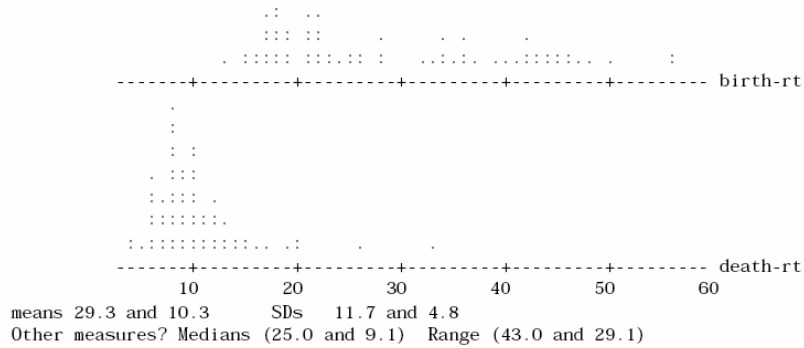
- Dot Plot
- Labelled Scatterplot
- Ordering Rows or Columns in a Table
- Crosstabs

28

Context

More about variability: vital statistics of 69 countries – rates per 1000 population

2.

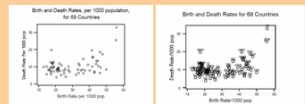


29



Analysis and Discussion

Analysis + Discussion (1)



Can't read the second graph?

Analysis + Discussion (2)

Row	Birth-rt	Death-rt	country	Continent
1	30.4	11.0	AFG	AFRICA
2	37.3	9.9	COM	AFRICA
12	20.4	11.0	MDG	AFRICA
14	20.9	9.9	MLI	AFRICA
27	24.6	7.9	MDG	AFRICA
28	49.9	6.3	COM	AFRICA
41	16.9	12.9	MDG	AFRICA
46	17.1	12.7	MDG	AFRICA
68	21.6	9.7	MDG	AFRICA
69	21.5	9.9	MDG	AFRICA

But is this a good way to arrange the table? Note that sorting the rows often helps.

Sort by birth rate, or by death rate, or by ratio birth/death rate, for example.

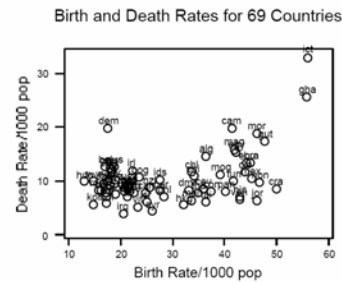
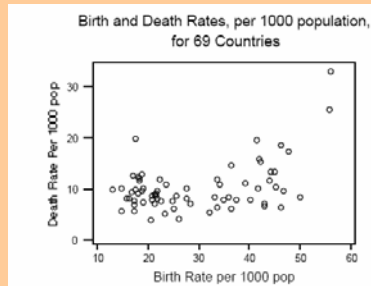
Homework: (not to hand in, yet)
Propose a method for numerical summary of this data. By eye-balling the table, anticipate what your summary would show, and express this in words.



30



Analysis + Discussion (1)



Can't read the second graph!

31



Analysis + Discussion (2)

Row	birth-rt	death-rt	country	Contin.
1	36.4	14.6	alg	AFRICA
2	37.3	8.0	con	AFRICA
13	32.1	5.5	hkg	ASIA
14	20.9	8.8	ind	ASIA
27	24.8	7.8	can	AMERICA
28	49.9	8.5	cra	AMERICA
45	18.8	12.8	aus	EUROPE
46	17.1	12.7	bel	EUROPE
68	21.6	8.7	ast	OCEANIA
69	25.5	8.8	nzl	OCEANIA

But is this a good way to arrange the table? Note that sorting the rows often helps.

Sort by birth rate, or by death rate, or by ratio birth/death rate, for example.

Homework: (not to hand in, yet)
Propose a method for numerical summary of this data. By eye-balling the table, anticipate what your summary would show, and express this in words.

32



Context

To design an informative table or graph, one needs to make careful use of the context of the data. (labelling plotted points by country, ordering table rows usefully).

33

Back to Case Study



Techniques

Dot Plot

Dot plot

Labelled Scatterplot

Labelled scatterplot

Ordering Rows or Columns in a Table

Table:
The general idea is that if a feature of a display is arbitrary, it may sometimes be re-organized to advantage.

Back to Case Study

34



Dot Plot

Dot-plot

35

[Back to Case Study](#)



Labelled Scatterplot

Labelled scatterplot

36

[Back to Case Study](#)



Ordering Rows or Columns in a Table

Table

The general idea is that if a feature of a display is arbitrary, it may sometimes be re-organized to advantage.

Back to Case Study

37



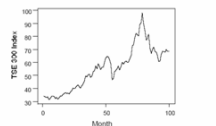
Stock Market Index

1. The situation....

Context

Here is some recent year's stock index levels for the Toronto Stock Market.

100 Days of the TSE 300 Stock Index



How can this series be described?

2. Analysis and conclusion

Analysis and Discussion



3. Concepts

Pattern Illusions in Time Series

Apparent trends can be useful for prediction, as is the case in the symmetric random walk – the trend may be useful to guide your actions, but the trend up or down may not persist. It takes a long time series to determine whether trends are real or illusions, and even in a long time series you need some stability in the mechanism to infer anything.

An example of this is in [randomwalks](#)

Back to Home Study

4. Techniques

Random Walk

Time Series Plot

Simulation – Symmetric Random Walk ($p=0.5$)

Random Walk with Variable Step Sizes

[RandomWalks](#)

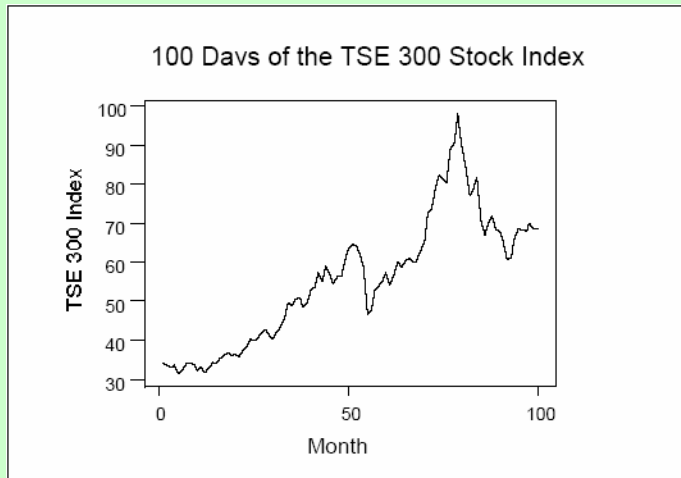
Back to Home Study

38



Context

Here is some recent year's stock index levels for the Toronto Stock Market.



How can this series be described?

39



Analysis and Discussion

Analysis + Discussion(1)

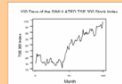
Coin Fl
Miles Ma

Analysis + Discussion (2)

The fact that the trend in the series
is accidental, but the variability
does seem similar.

What does this tell us?

That the TSE trend could have
occurred when the series had no
predictable trend at all - because
the simulated series was
designed to have no predictable
trend.

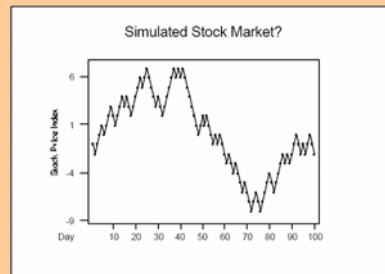


40

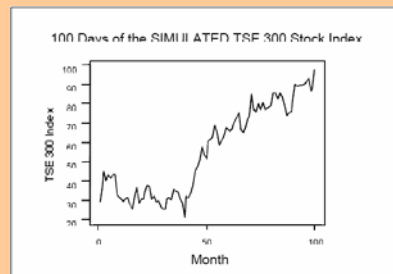


Analysis + Discussion(1)

Coin Flipping reproduces a trend a little like the market. $H = +1$, $T = -1$



A slight modification to allow steps of varying size, but still equally likely up or down.



41

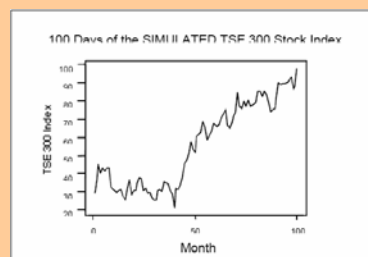


Analysis + Discussion (2)

The fact that the trend is the same is accidental. But the variability does seem similar.

What does this tell us?

That the TSE trend could have occurred when the series had no predictable trend at all - because the simulated series was designed to have no predictable trend.



42




Pattern Illusions in Time Series

Apparent trends can be useless for prediction, as is the case in the symmetric random walk – the level may be useful to guide your actions, but the trend up or down may not persist. It takes a long time series to determine whether trends are real or illusions, and even in a long time series you need some stability in the mechanism to infer anything.

An example of this is in [randwalk.xls](#)

43

 Back to Case Study



Random Walk


Time Series Plot

Simulation – Symmetric Random Walk ($p=0.5$)

Random Walk with Variable Step Sizes

[RandWalk.xls](#)

44

 Back to Case Study



Auto Insurance

1. The situation....

Context

Annual Suppose we pay \$5 every day for auto insurance. The company receives $5 \times 365 = \$1825$ in one year. If I have no accident the company keeps the \$1825. If I have an accident, suppose the average cost is \$6,000. Also, suppose the company has determined that my probability of having an accident this year is $1/5=0.2$. Will the company make money?

With one customer, the company could not be sure. But with 100 customers, here is what would happen according to several simulations of a years experience with 100 customers:

3. Concepts

Stability of Averages (1) -Spreading the Risk

Averages are more stable than the spread of possible values.
The insurance company can spread the risk across many policy holders.
The insurance company sells this service to policy holders -- which is why the company takes in more than it pays out.

2. Analysis and conclusion

Analysis and Discussion



4. Techniques

Variability of Averages & Totals(2)

Averages have an SD that is =
(the SD of the original measurement) / \sqrt{n}
The total of n measurements is just n times the average measurement.
If the average varies by ± 10 , then a total of the n things averaged will vary by $\pm 10 n$. It is assumed here that the things averaged vary independently.
Example: The "Risky Company" has an average return of \$1.28 and an SD of 1.2. The average return for 100 such companies (acting independently) is $1.28 \pm 1.2 / \sqrt{100}$ or $1.28 \pm .12$. In other words, typical returns are in the range 28% to 56% -- not bad for risky companies! (But note the independence assumption -- which is hard to achieve completely in practice.)

45



Context

Annual Suppose we pay \$5 every day for auto insurance. The company receives $5 \times 365 = \$1825$ in one year. If I have no accident the company keeps the \$1825. If I have an accident, suppose the average cost is \$6,000. Also, suppose the company has determined that my probability of having an accident this year is $1/5=0.2$. Will the company make money?

With one customer, the company could not be sure. But with 100 customers, here is what would happen according to several simulations of a years experience with 100 customers:

46



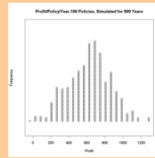
Analysis and Discussion

Analysis and Discussion(2)

Analysis and Discussion(1)

Even though the outcome will be 0000-1 of \$4.12, expect almost.

The graph shows what would happen if the insurance company had 100 such policies. Each square is the outcome of one year's simulated experience. The simulation suggests that the insurance company will only make a profit in 499 out of 500 years' experience.



47

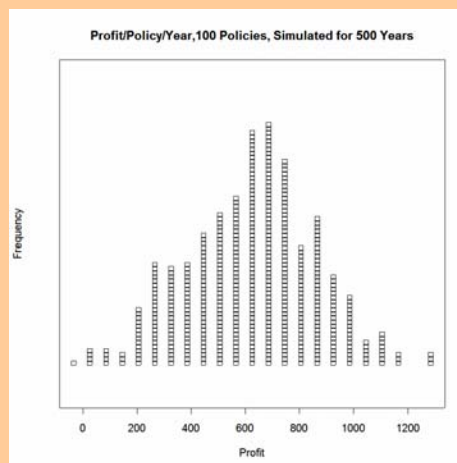


Analysis and Discussion(1)

The graph shows what would happen if the insurance company had 100 such policies.

Each square is the outcome of one year's simulated experience.

The simulation suggests that the insurance company will only make a profit in 499 out of 500 years' experience.



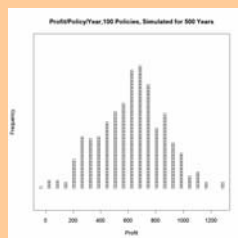
48



Analysis and Discussion(2)

Even though a single policy in which an accident must be paid out will cost the company 6000-1825, i.e. result in a loss of \$4175, the aggregated experience of the 100 policies is almost always profitable.

Although the 100 policies could possibly have lost the company \$417,500, and could possibly have profited the company \$182,500, the realistic range of outcomes is a profit in the range (\$200, \$1150). This is the range of outcomes 95% of the time.



The insurance company is happy.

The policy holder may not be happy with this expense, but must have decided the benefit is worth the cost.

49



Stability of Averages (1) -Spreading the Risk

Averages are more stable than the spread of possible values.

The insurance company can spread the risk across many policy holders.

The insurance company sells this service to policy holders – which is why the company takes in more than it pays out.

50

← Back to Case Study



Variability of Averages & Totals(1)

Averages have an SD that is =

(the SD of the original measurements) $/\sqrt{n}$

The total of n measurements is just n times the average measurement.
If the average varies by ± 10 , then a total of the n things averaged will vary by $\pm 10 n$.

Example: Policy outcomes average \$625 and have an SD of \$2683.
Then the average outcome 100 policies is still \$625 but the SD of this average is $2683/\sqrt{100} = \$268.30$

The total of the 100 policy outcomes averages \$62,500 and has an SD of 100 times \$268.30 = \$26,830.

[Back to Case Study](#)

51



Investment Risk

1. The situation....

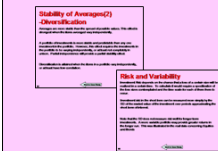
Context

Farmers, Investors, Manufacturers, and Gamblers all understand the maxim "Don't put all your eggs in one basket". The idea is that diversification will reduce your risk of catastrophe. We illustrate this idea in the context of investment: capital funds are allocated to several classes of investment, like bonds, equities and real estate. Ideally, for the best diversification, the experiences of these classes should be "uncorrelated": their market values should move up and down independently. After demonstrating this with real data from the Canadian market, we continue a simulation example to show that the phenomenon is not a transient effect. More specifically, we show that a portfolio of unrelated risky companies can be a low-risk and yet very profitable investment.

52

3. Concepts

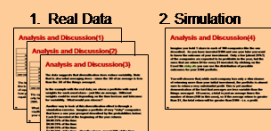
Averages and Variability



53

2. Analysis and conclusion

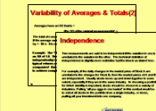
Analysis and Discussion



54

4. Techniques

Totals of Independent Variables



55



Context

Farmers, Investors, Manufacturers, and Gamblers all understand the maxim “Don’t put all your eggs in one basket”. The idea is that diversification will reduce your risk of catastrophic loss. We illustrate this idea in the context of investment: capital funds are allocated to several classes of investment, like bonds, equities and real estate. Ideally, for the best diversification, the experience of these classes should be “uncorrelated”- their market values should move up and down independently. After demonstrating this with real data from the Canadian market, we examine a simulation example to show that the phenomenon is not a transient effect. More specifically, we show that a portfolio of unrelated risky companies can be a low-risk and yet very profitable investment.

53



Analysis and Discussion

1. Real Data

Analysis and Discussion(1)

Analysis and Discussion(2)

Analysis and Discussion(3)

The data suggests that diversification does reduce variability. Note that is also total averaging down - since the SD of an average is less than the SD of the things averaged.

In the example with the real data, we chose a portfolio with equal weights for each asset class - just like an average. Different weights could be used depending on the time horizon and tolerance for variability. What would you choose?

Another way to look at this diversification effect is through a simulation exercise. Imagine a portfolio of very “risky” companies that have a one year prospect described by the probabilities below. Each \$1 invested at the beginning of the year returns

\$0.00	25% of the time
\$0.50	25% of the time
\$1.00	25% of the time
\$4.00	25% of the time

Usually a loser, except 25% of the time.

2. Simulation

Analysis and Discussion(4)

Imagine you hold 1 share in each of 100 companies like the one described. So you have invested \$100 and one year later you want to know the outcome of your investment. Only a few (about 25%) of the companies are expected to be profitable in the year, but the ones that are return \$4 for every \$1 invested. By clicking on the Excel file (only.xls) you can see the distribution of possible outcomes for your \$100 portfolio.

You will observe that, while each company has only a slim chance of returning more than your initial investment, the portfolio is almost sure to return a very substantial profit. This is yet another demonstration of the fact that averages are less variable than the things averaged. (Of course, a total is just an average times the number of items (100 in this case), so if the average return is greater than \$1, the total return will be greater than \$100 - i.e. a profit.

54



Analysis and Discussion(1)

	Cash	Bonds	Cdn Equities	US Equities	Fgn Equities	Real Estate	Combo
1995	7.30%	18.79%	13.57%	14.77%	7.93%	4.92%	11.21%
1996	4.90%	11.56%	24.96%	11.98%	6.20%	6.78%	11.06%
1997	3.15%	9.19%	13.96%	6.85%	6.02%	17.25%	9.40%
1998	4.63%	8.77%	-1.59%	27.47%	25.08%	14.84%	13.20%
1999	4.55%	-1.15%	27.54%	18.86%	18.01%	10.12%	12.99%
1995-1999	4.90%	9.25%	15.21%	15.78%	12.39%	10.68%	11.37%

The table above shows the annual return for 6 Asset Classes. The bottom line is the annualized returns for the 5 year period. The "Combo" is a portfolio of equal amounts of each asset class invested January 1995. Consider now how you would invest in January 2000 for the next five year period. Perhaps equal amounts of the three equity classes?

55



Analysis and Discussion(2)

	Cash	Bonds	Cdn Equities	US Equities	Fgn Equities	Real Estate	Combo
2000	5.34%	9.77%	7.15%	32.32%	-11.41%	10.80%	8.99%
2001	4.61%	7.75%	-13.43%	30.96%	-18.15%	8.80%	3.42%
2002	2.47%	8.37%	-13.28%	10.68%	-18.20%	8.34%	-0.27%
2003	2.87%	6.49%	23.68%	-8.85%	12.58%	7.88%	7.44%
2004	2.26%	6.90%	13.52%	-7.38%	11.35%	12.13%	6.47%
2000-2004	3.50%	7.85%	2.46%	10.12%	-5.75%	9.58%	4.63%

The equities have not done as well! In fact that equities-only portfolio would only have made 2.28% annualized rate of return over the five years 2000-2004. Compare with the Combo at 4.63% - not great, but better than 2.28%

What about the period 1995-2004 inclusive? Combo is 7.9% while equities are 8.2%. For longer periods still, equities have done better. So your choice depends on your time horizon. How long can you wait?

56



Analysis and Discussion(3)

The data suggests that diversification does reduce variability. Note that is also what averaging does – since the SD of an average is less than the SD of the things averaged.

In the example with the real data, we chose a portfolio with equal weights for each asset class – just like an average. Different weights could be used depending on the time horizon and tolerance for variability. What would you choose?

Another way to look at this diversification effect is through a simulation exercise. Imagine a portfolio of very “risky” companies that have a one year prospect described by the probabilities below: Each \$1 invested at the beginning of the year returns

\$0.00 25% of the time

\$0.50 25% of the time

\$1.00 25% of the time

\$4.00 25% of the time Usually a loser, except 25% of the time.

57



Analysis and Discussion(4)

Imagine you hold 1 share in each of 100 companies like the one described. So you have invested \$100 and one year later you want to know the outcome of your investment. Only a few (about 25%?) of the companies are expected to be profitable in the year, but the ones that are return \$4 for every \$1 invested. By clicking on the Excel file [risky.xls](#) you can see the distribution of possible outcomes for your \$100 portfolio.

You will observe that, while each company has only a slim chance of returning more than your initial investment, the portfolio is almost sure to return a very substantial profit. This is yet another demonstration of the fact that averages are less variable than the things averaged. Of course, a total is just an average times the number of items (100 in this case), so if the average return is greater than \$1, the total return will be greater than \$100 – i.e. a profit.

58



Variability of Averages & Totals(2)

Averages have an SD that is =

(the SD of the original measurements)/ \sqrt{n}

The total of n measurements is just n times the average measurement.

If the average varies by ± 10 , then a total of the n things averaged will vary by $\pm 10n$. It is assumed here that the things averaged vary independently.

Example: The “Risky Company” has an average return of \$1.38 and an SD of 1.8. The average return for 100 such companies (acting independently) is $1.38 \pm 1.8/\sqrt{100}$ or $1.38 \pm .18$. In other words, typical returns are in the range 20% to 56% - not bad for risky companies! But note the independence assumption – which is hard to achieve completely in practice.

59

[Back to Case Study](#)



Independence

Two measurements are said to be independent if the variation in one is unrelated to the variation in the other. The technical definition of independence is slightly more restrictive but the idea is as stated here.

For example, if the daily changes in the market price of Stock A are unrelated to the changes for Stock B, then the market prices of A and B are independent. Usually stocks move up and down together to some extent, especially if they are in the same industry. In choosing a portfolio for which stability is important, stocks should be chosen in a variety of industries. Putting “all your eggs in one basket” in this context would be to select all stocks in the portfolio from a single industry, or worse, putting all your investment into one company.

60

[Back to Case Study](#)



Stability of Averages(2)


-Diversification

Averages are more stable than the spread of possible values. This effect is strongest when the items averaged vary independently.

A portfolio of investments is more stable and predictable than any one investment in the portfolio. However, this effect requires the investments in the portfolio to be varying independently, or at least not completely in unison. Partial independence will provide a partial stability effect.

Diversification is attained when the items in a portfolio vary independently, or at least have low correlation.

61

 Back to Case Study




Risk and Variability

Investment Risk depends on the chance that a loss of a certain size will be realized in a certain time. To calculate it would require a specification of the loss sizes contemplated and the time scale for each of them to occur.

Investment risk in the short term can be measured more simply by the SD of the market value of the investment over periods approximating the short term of interest.

Note that the SD does not measure risk well for longer term investments. A more variable portfolio may provide greater returns in the longer run. This was illustrated in the real data concerning Equities and Bonds

62

 Back to Case Study



Context

63



Analysis and Discussion

64



Averages and Variability

Stability of Averages(2)

-Diversification

Averages are more stable than the spread of possible values. This effect is strongest when the items averaged vary independently.

A portfolio of investments is more stable and predictable than any one investment in the portfolio. However, this effect requires the investments in the portfolio to be varying independently, or at least not completely in unison. Partial independence will provide a partial stability effect.

Diversification is attained when the items in a portfolio vary independently, or at least have low correlation.

Back to Case Study

Risk and Variability

Investment Risk depends on the chance that a loss of a certain size will be realized in a certain time. To calculate it would require a specification of the loss sizes contemplated and the time scale for each of them from to occur.

Investment risk in the short term can be measured more simply by the SD of the market value of the investment over periods approximating the short term of interest.

Note that the SD does not measure risk well for longer term investments. A more variable portfolio may provide greater returns in the longer run. This was illustrated in the real data concerning Equities and Bonds.

Back to Case Study

Back to Case Study

65

Stability of Averages

Stability of Averages (1)

-Spreading the Risk

Averages are
The insurance
The insurance
The company

Stability of Averages(2)

-Diversification

Averages are more stable than the spread of possible values. This effect is strongest when the items averaged vary independently.

A portfolio of investments is more stable and predictable than any one investment in the portfolio. However, this effect requires the investments in the portfolio to be varying independently, or at least not completely in unison. Partial independence will provide a partial stability effect.

Diversification is attained when the items in a portfolio vary independently, or at least have low correlation.

Back to Case Study

Back to Case Study

66

Transient Patterns

Pattern Illusions in Time Series

Apparent trends can be useless for prediction, as is the case in the apparent random walk — the level may be useful to guide your actions, but the trend determines you need.

Patterns that Disappear

Some patterns in data are transient, and some are persistent. The transient ones are called illusions. The cause of these illusions is unexplained variation, and it can lead to misinterpretation of data by the statistically naive. Patterns in data can appear that seem too regular to be transient, but this can be an illusion. By learning what can happen when we create a model that has no useful pattern, we can guard against being fooled by an apparent pattern. Of course, we do want to find those patterns that last, are not transient, if they exist. But it can be very difficult to distinguish between transient and persistent patterns. This dilemma motivates many of the techniques of statistics.

An example

Back to Case Study

67

Back to Case Study



Simulation

Simulation provides a way to explore the variability of a measurement when the model is known or assumed. We can compare the outcome of a simulation with known data to infer whether or not the assumed model is credible.

68

Back to Case Study



Means and Standard Deviations

Calculation of the SD

Using the Mean and SD

Standard
compute

Support
The SD
number
into the
with the
study is
Example
is 200-

So the numbers mean and SD do give a convenient statistical description of a batch of numbers.

Back to Case Study

69

Back to Case Study



Totals of Independent Variables

Variability of Averages & Totals(2)

Averages have an SD that is =

(the SD of the original measurement) / \sqrt{n}

The total of n measures
if the average varies
by $\pm \sqrt{n}$ it is as

Example: The "N"
an SD of 1.5. The
independently in
typical releases as
companies? But
to achieve comp

Independence

Two measurements are said to be independent if the variation in one is unrelated to the variation in the other. The technical definition of independence is slightly more restrictive but the idea is as stated here.

For example, if the daily changes in the market price of Stock A are unrelated to the changes for Stock B, then the market prices of A and B are independent. Usually stocks move up and down together to some extent, especially if they are in the same industry. In choosing a portfolio for which stability is important, stocks should be chosen in a variety of industries. Putting "all your eggs in one basket" in this context would be to select all stocks in the portfolio from a single industry, or worse, putting all your investments into one company.

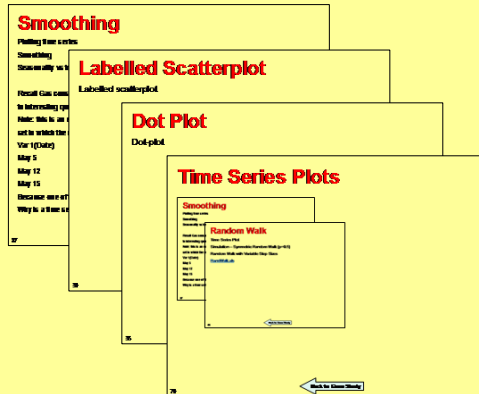
Back to Case Study

70

Back to Case Study

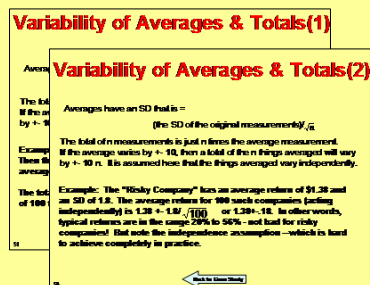


Graphical Summary Methods



71

Variability of Averages



72

Tables

Ordering Rows or Columns in a Table

Table

The general idea is that if a feature of a display is arbitrary, it may sometimes be re-organized to advantage.

Apply to Case Study

Back to Case Study

73



Numerical Summaries

Using the Mean and SD

Using mean & standard deviation (SD)

The "mean" found by summing all the values of x and dividing by n . The mean of x is \bar{x} .
The SD is the square root of the average squared deviation from the mean.
So the SD is the average distance from the mean.

Standard deviation
Suppose y is the average of n numbers. The SD of y is the SD of the numbers divided by \sqrt{n} .
Example: The SD of 200-4 is 10.

Calculation of the SD

Variability of Averages & Totals(1)

Average
The total of n things is $n\bar{x}$.
Example: The total of 200-4 is 2000.

Variability of Averages & Totals(2)

Averages have an SD that is $\frac{1}{\sqrt{n}}$ times the SD of the original measurements.
(The SD of the original measurements is \sqrt{n} times the SD of the averages.)
The total of n measurements is just n times the average measurement. If the average varies by ± 10 , then a total of the n things averaged will vary by $\pm 10n$. It is assumed here that the things averaged vary independently.
Example: The "Risky Company" has an average return of \$1.20 and an SD of 1.8. The average return for 100 such companies (acting independently) is $1.20 \pm 1.8/\sqrt{100}$ or 1.20 ± 0.18 . In other words, typical returns are in the range 10% to 50% - not bad for risky companies! But note the independence assumption - which is hard to achieve completely in practice.

Apply to Case Study

Back to Case Study

74



Scatter Plot

75

Back to Case Study



Time Series Plots

Smoothing

Plotting time series
Seasonality
Seasonally adjusted

Recall Gas came
to interesting ques
Note: this is an ex
not to which the r
Var 1 (Date)

May 5
May 12
May 15
Discussion was effe
Why is a time ser

37

Random Walk

Time Series Plot

Simulation - Symmetric Random Walk ($p=0.5$)

Random Walk with Variable Step Sizes

[RandomWalk.xls](#)

46

Back to Case Study

76

Back to Case Study



See pages following for handouts of slides for *Quality Assurance in University Teaching: A View from Both Sides of the Atlantic* Bruce Dunham, Ph.D. (University of British Columbia).

QUALITY ASSURANCE IN UNIVERSITY TEACHING: A VIEW FROM BOTH SIDES OF THE ATLANTIC

BCCUPMS 2006

Dr. Bruce Dunham
Department of Statistics
University of British Columbia

May 2006

Abstract

Higher education teaching in the UK is regulated and audited in a way that is not evident in North America. An overview will be given regarding the history, motivation and impact of the quality assurance procedures in Britain, with the aim of highlighting some of the so-called “good practices” that might beneficially be adopted in the teaching of Mathematics and Statistics in Canada. The speaker, having recently arrived in Canada from the UK, does not consider himself a “quality guru” in any sense, but has been struck by notable differences in attitudes and practices between the two countries.

... but previously of

- University of Nottingham (1992 – 1997)
- University of Derby (1998 – 2005)

Some history

1992: Higher Education Funding Council for England (HEFCE, www.hefce.ac.uk) formed – similar bodies for Wales, Scotland and N. Ireland.

1997: A separate body, Quality Assurance Agency for Higher Education (QAA, www.qaa.ac.uk), created to regulate and assess quality procedures in HE in UK, contracted to HEFCE.

1995 – 2001: Subject reviews.

2002 – 2005: Institutional audits in England and N. Ireland.

The motivation

Why do we need QA?

1. Accountability
2. Liability
3. LEAGUE TABLES!
4. FUNDING

So how do things differ?

Lecturer A starts work at University of X, and is told he will be teaching Stat 123 in his first term. He locates the lecturer who taught it the previous academic year.

Getting started

At University of X in Canada, he is informed:

- “We use Bloggs and Smith as the course book – generally chapters 1 to 8, then 10, then maybe 12 and 14 if you can. But you can do what you like.”
- “I gave two mid-terms, three assignments and a final, but you can do what you like.”

- “You can decide how to weight the pieces of assessment, and change it at the end if you prefer.”

At University of X in England, he is

- Given the “course box”, and told to look at the validated course specification, including learning outcomes, assessment strategy and indicative content. The box contains all resources for the previous run of the course, including assessments, samples of student work, summary of student feedback etc.
- Instructed to give the students a course guide at the first class, indicating clearly the assessment strategy, including weightings, related learning outcomes and dates.
- (possibly even) told that his course guide must be agreed with his head of subject.

During the course

At University of X in Canada, in Stat 123

- Students hand in work after a lecture, or put in a box outside the department.
- Student work is marked and returned – minimal feedback is given or expected.
- Students missing a mid-term (with “good reason”) have the weighting moved to the final exam.
- Lecturer A leaves setting the exam until a week before it is to be sat. There is no expectation that the paper will be seen by anyone else prior to the day.

- The exam papers are marked the day after the exam by Lecturer A and a team of TAs. Marks then are entered directly into the University of X's grade management system.

At University of X in England, in Stat 123

- Students hand in coursework, and are given dated receipts.
- Lecturer A completes feedback sheets for each student on their work. A sample of the work is internally moderated.
- A student missing a component of assessment has missed a learning outcome, so is deferred on that component.

- The exam is set three months before it is sat. It is internally moderated, then sent to the external examiner, from University of Y. The paper is then revised in the light of any comments.
- After the exam papers are marked, they (or at least a sample) are internally moderated. They are then moderated by the external examiner. Provisional course grades for all subjects within a degree programme are considered at the exam board, which includes all relevant academic staff and external examiners.

After the course

At University of X in Canada,

- Lecturer A has numerous queries regarding grades, requests to re-consider marks, and students wishing to see their final exam paper.
- Marks can be changed, via some paperwork.
- Exam papers are stored for two years, other work not yet returned is re-cycled.
- Failing students would repeat the course.

At University of X in England,

- Students have long since disappeared before their marks are available.
- Grades could only be changed via a “chair’s action” at the next exam board.
- Exam papers are stored for several years, samples of all assessment copied for course box.
- Failing students have referrals on failed components.
- Lecturer A completes a course leader’s report form, summarising results, aspects of how the course went and student feedback. This is considered at a subject team meeting, along with other reports for that year.

Does improved quality assurance improve courses?

Well arguably, but the QAA audits could be accused of ...

- Focusing on certain issues *ad infinitum*, while not touching others.
- Ignoring some key issues, particularly inter-establishment differences.
- Failing to detect some heinous crimes.