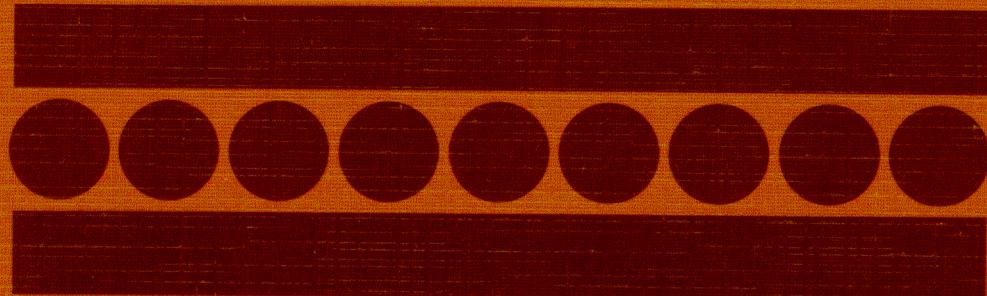


NAZIOARTEKO ESTATISTIKA
MINTEGIA EUSKADIN

1983

SEMINARIO INTERNACIONAL
DE ESTADÍSTICA EN EUSKADI



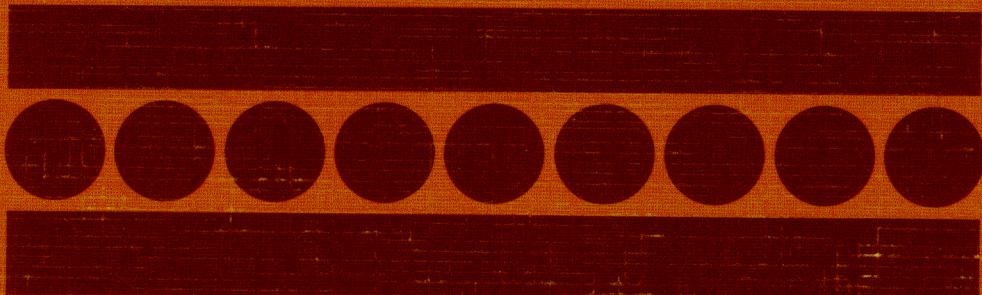
ESTATISTIKA-IRAKASKUNTZA

ENSEÑANZA DE LA ESTADÍSTICA

ENSEIGNEMENT DE LA STATISTIQUE

STATISTICAL EDUCATION

V. BARNETT



EUSKO JAURLARITZA
Estatística Zuzendaritza



GOBIERNO VASCO
Dirección de Estadística

NAZIOARTEKO ESTATISTIKA
MINTEGIA EUSKADIN

1983

SEMINARIO INTERNACIONAL
DE ESTADISTICA EN EUSKADI

ESTATISTIKA-IRAKASKUNTZA



ENSEÑANZA DE LA ESTADISTICA



ENSEIGNEMENT DE LA STATISTIQUE



STATISTICAL EDUCATION

V. BARNETT

CUADERNO **3** KOADERNOA

EUSKO JAURLARITZA
Estatística Zuzendaritza



GOBIERNO VASCO
Dirección de Estadística

Zuzendaritza/Dirección: José Pérez Vilaplana
Lankidetza/Colaboración: Anjeles Iztueta Azkue

© Eusko Jaurlaritza - Gobierno Vasco

Depósito Legal: S.S. 18-85
ISBN: 84-7542-127-10. Obra Completa.
ISBN: 84-7542-158-X. Tomo III.
Impreso en Itxaropena, S.A. - Errikobarra kalea, 2 - Zarautz.

AURKEZPENA

Estatistikako Mintegi Internazionalak sustatzean, hainbat xederekin bete nahi luke Eusko Jaurlaritzaren Estatistika Zuzendaritzak, hala nola:

- Unibertsitatearekiko eta, bereziki, Estatistika Sailarekiko lankidetza bultzatu.
- Funtzionari, irakasle, ikasle eta estatistikaren alorrean interesaturik leudekeen guztien birziklapen profesionala erraztu.
- Estatistfkako alorrean eta mundu-mailan irakasle prestu eta abangoardiako ikerlari diren pertsonaiak Euskadira ekarri, guzti horrek zuzeneko harremanei eta esperientzien ezagupenei dago-kienez soposatzen duen ondorio positiboarekin.

Iharduketa osagarri bezala eta interesaturik leudekeen ahalik eta pertsona eta Erakunde gehienetara iristearren, Ikastaro hauetako txostenak argitaratzea erabaki da, beti ere txostenemaiaren jatorrizko hizkuntza errespetatuz, horrela gure Herrian gai honi buruzko ezagutza zabaltzen laguntzeko asmoarekin.

PRESENTACION

Al promover los Seminarios Internacionales de Estadística, la Dirección de Estadística del Gobierno Vasco pretende cubrir varios objetivos:

- Fomentar la colaboración con la Universidad y en especial con los Departamentos de Estadística.
- Facilitar el reciclaje profesional de funcionarios, profesores, alumnos y cuantos puedan estar interesados en el campo estadístico.
- Traer a Euskadi a ilustres profesores e investigadores de vanguardia en materia estadística, a nivel mundial, con el consiguiente efecto positivo en cuanto a relación directa y conocimiento de experiencias.

Como actuación complementaria y para llegar al mayor número posible de personas e Instituciones interesadas, se ha decidido publicar las ponencias de estos Cursos, respetando en todo caso la lengua original del ponente, para contribuir así a acrecentar el conocimiento sobre esta materia en nuestro País.

PRESENTATION

La Direction de Statistique du Gouvernement Basque se propose d'atteindre plusieurs objectifs par la promotion des Séminaires Internationaux de Statistique:

- Encourager la collaboration avec l'université et spécialment avec les départements de statistique.
- Faciliter le recyclage professionnel des fonctionnaires, professeurs, élèves, et tous ceux qui pourraient être intéressés par la statistique.
- Inviter en Euskadi des professeurs mondialement renommés et des chercheurs de premier ordre en matière de Statistique avec tout ce que cela pourrait entraîner comme avantage dans les rapports et l'échange d'expériences.

En outre, il a été décidé de publier les exposés de ces rencontres afin d'atteindre le plus grand nombre de personnes et d'institutions intéressées, et pour contribuer ainsi à développer dans notre pays les connaissances sur cette matière. Dans chaque cas la langue d'origine du conférencier sera respectée.

PRESENTATION

In promoting the International Seminars on Statistics, the Statistics Office of the Basque Government is attempting to achieve a number of objectives:

- Encourage joint working with the Basque University and, in particular, with its Department of Statistics.
- Facilitate the in-training of civil servants, teachers and students and of all those interested in the field of statistics.
- Bring to Euskadi distinguished academics and researchers in the front line of statistics work, at a world-wide level, with all the benefits that this will bring through direct contacts and the interchange of experiences and ideas.

As an additional step this year, it has been decided to publish in advance the papers to be presented at these courses, respecting the native language of the speaker, in each case. This is in order that as many interested people and institutions as possible are made aware. In this way we hope to contribute to the growth and awareness concerning this topic in our country.

Vitoria-Gazteiz, Diciembre 1984 Abendua

JOSE IGNACIO GARCIA RAMOS
Estatistikako Zuzendaria
Director de Estadística

SARRERA

Koaderno honek, Eusko Jaurlaritzaren Estatistika-Zuzendaritzak eta Euskal Herriko Unibertsitatearen Matematika Aplikatuko Departamentuak antolaturik, Vic BARNETTek Euskadiko Estatistikako I. Nazioarteko Mintegien barruan "Statistical Education" gaiari buruz eman duen ikastaroa laburbiltzen du. I. Mintegi honek kontatzen du baita USAko Pittsburgh-eko Unibertsitateko C.R. RAO Jaunaren partaidetzarekin ere, "Lineal Statistical Inference" gaiari buruzko ikastaro batekin, Txileko Santiacagoko Unibertsitateko E. CANSADO Jaunaren partaidetzarekin, "Muestreo y Aplicaciones" gaiari buruzko ikastaroarekin, eta orobat Pariseko CREDOCeko P. CLAPIER Jaunaren partaidetzarekin, "Analyse des Données" gaiari buruzko ikastaroarekin

INTRODUCCION

Este cuaderno resume el curso que sobre "Statistical Education" ha impartido V. BARNETT dentro del I Seminario Internacional de Estadística en Euskadi, organizado por la Dirección de Estadística del Gobierno Vasco y el Departamento de Matemática Aplicada de la Universidad del País Vasco. Este I Seminario cuenta con la participación de C.R. RAO de la Universidad de Pittsburgh, USA, con un curso sobre "Lineal Statistical Inference", E. CANSADO de la Universidad de Santiago de Chile con un curso sobre "Muestreo y Aplicaciones" y P. CLAPIER de CREDOC París con un curso sobre "Analyse des Données".

INTRODUCTION

Ce cahier résume le cours donné par V. BARNETT par "Statistical Education" dans le cadre du I Séminaire International de Statistique d'Euskadi, organisé par le Conseil de Statistique du Gouvernement Basque et par le Département de Mathématique Appliquée de l'Université du Pays Basque. Ce I Séminaire compte avec la participation de C.R. RAO de l'Université de Pittsburgh (USA), avec un cours sur "Lineal Statistical Inference", et de E. CANSADO de l'Université de Santiago de Chile avec un cours sur "Muestreo y Aplicaciones" et de P. CLAPIER du CREDOC de Paris avec sur "Analyse des Données".

INTRODUCTION

This paper recapitulates the lecture on "Statistical Education" presented by V. BARNETT during the I International Seminar on Statistics in Euskadi, organised by the Statistics Office of the Basque Government and the Department of Applied Mathematics at the University of the Basque Country. This I Seminar included the participation of R.A. RAO of Pittsburgh University (USA) with a theme titled "Lineal Statistical Inference" of E. CANSADO of the University of Santiago in Chile with a course on "Muestreo y Aplicaciones", and of P. CLAPIER of CREDOC, Paris with a lecture on "Analyse des Données".

BIOGRAFIA

Vic Barnett B.H.ko Inglaterrako Shefield-eko Unibertsitatean Probabilitate eta Estatistikako irakasle dugu. Manchester-eko Unibertsitatean egin zituen bere ikasketak, bertan Matematikan bere B.Sc. eta Estatistika Matematikoan bere M. Sc. eta Ph. D. lortuz, eta Birmingham-eko Unibertsitatetik D.Sc. delarik. Royal Statistical Society eta Institut of Statisticans erakundeetako partaidea da eta International Statistical Institute eta Biometrics Society-ko baxkide. B.H.ko, I.S.I. Task Force on School-Level Statistical Education-eko Lehendakaria da, Center for Statistics Education at Shefield University/Shefield City Polytechnic-eko Lehendakarikidea eta Statistics Sub-Committee of the Joint Matriculation Board of the Departement of Education and Science-eko Lehendakari.

Argitaratzaile elkartua da Royal Statistical Society Journal, Journal of Applied Probability eta Teaching Statistics aldizkarietan. Estatistikako artikulu ugariren eta hainbat libururen autore, bestalde.

Vic Barnett es profesor de Probabilidad y Estadística en la Universidad de Shefield (Inglaterra, U.K.). Estudió en la Universidad de Manchester donde obtuvo su B.Sc. en Matemáticas y sus M.Sc. y Ph.D. en Estadística Matemática, siendo D.Sc. por la Universidad de Birmingham. Es fellow de la Royal Statistical Society y del Institute of Statisticians y miembro del International Statistical Institute y de la Biometrics Society. Es Presidente de la I.S.I. Task Force on School-Level Statistical Education, Co-Presidente del Center for Statistical Education at Shefield University/Shefield City Polytechnic y Presidente del Statistics Sub-Committee of the Joint Matriculation Board of the Department of Education and Science, U.K.

Editor asociado de la Royal Society Journal, Journal of Applied Probability y Teaching Statistics Statistics. Autor de numerosos artículos y libros de Estadística.

Vic Barnett est Professeur de Probabilité et Statistique à l'Université de Shefield (Angleterre U.K.). Il fait ses études à l'Université de Manchester où il obtient son B.Sc. en Mathématiques et ses M.Sc. et Ph.D. en Statistique Mathématique, et est nommé D.Sc. par l'Université de Birmingham. Il est Fellow de la Royal Statistical Society et de l'Institute of Statisticians ainsi que membre de l'International Statistical Institute et de la Biometrics Society. Il est président de l'I.S.I. Task Force on School-Level Statistical Education, Co-Président du Center for Statistical Education à l'Université de Shefield/Shefield City Polytechnic Board of the Department of Education and Science (U.K.9.).

Editeur associé du Royal Statistical Society Journal, Journal of Applied Probability and Teaching Statistics. Auteur de nombreux articles et ouvrages de Statistique.

Vic Barnett is Professor of Probability and Statistics at the University of Sheffield in England. He studied at Manchester University where he obtained his B.Sc. in Mathematics and his M.Sc. and Ph. D in Statistical Mathematics. D. Sc. at Birmingham University. Fellow of the Royal Statistical Society and of the Institute of Statisticians and Member of the International Statistical Institute and of the Biometrics Society. President of the I.S.I. Task Force on School-Level Statistical Education, Co-President of the Centre for Statistical Education at Sheffield University/Sheffield City Polytechnic and President of the Statistics Sub-Committee of the Joint.

Matriculation Board of the Department of Education and Science in the U.K. Associate Editor of the Royal Statistical Society Journal, of the Journal of Applied Probability and Teaching Statistics. Author of numerous articles and books on statistics.



AURKIBIDEA/ INDICE

Aurkezpena/Presentación/Presentation	
Presentation	5
Sarrera/Introducción/Introduction/Introduction	
.....	7
Biografia	9
ESTATISTIKA-IRAKASKUNTZA	13
Laburpena	13
ESTATISTICAL EDUCATION	15
1. Introduction	15
1.1. General comments	15
1.2. Teaching Statistics in Schools	16
2. Teaching Statistics in Universities and Colleges	17
2.1. A degree course in Statistics	19
2.2. Undergraduate Mathematics Student	21
2.3. Service Courses in Statistics	21
2.4. Taught Masters Courses in Statistics	22
3. Statistical Education for Statistical Practice	23
APPENDIX I	26
APPENDIX II	26
APPENDIX III	27
APPENDIX IV	29
APPENDIX V	30
APPENDIX VI	31
APPENDIX VII	34
ENSEÑANZA DE LA ESTADISTICA	35
1. Introducción	35
1.1. Comentarios generales	35
1.2. Enseñanza de la Estadística a nivel Secundario	36
2. La Enseñanza de la Estadística en la Universidad	37
2.1. Graduación en Estadística	39
2.2. Estudiantes de Matemáticas	41
2.3. Cursos de Aplicaciones de la Estadística	41
2.4. Enseñanza de Master en Estadística	42
3. Educación Estadística para la práctica Estadística	43
APENDICE I	45
APENDICE II	45
APENDICE III	46
APENDICE IV	48
APENDICE V	49
APENDICE VI	50
APENDICE VII	53

ESTATISTIKA-IRAKASKUNTZA

V. Barnett

Sheffield University

LABURPENA

Ondoko galdera oinarrizkoak kontsideratzen dira: Zergatik irakatsi behar da Estatistika? Nori iraka_{tsi} behar zaio Estatistika? Estatistikako metodo eta printzipoetatik zeintzu azpimarratu behar dira? / Nola motiba daiteke ikaslea? Nola adieraz dakitzioke ideia estatistikoak? Nolako adibideak eman? Lanera_{ko} programa eta eskemak. Kondizio orokorrak azaltzeko adibide xehekatuak. Nola afektatzen dute profesio_{naria} eta hiritar arrunta heziketa estatistikoaren alderdi orokorrekin?.

Konkretuago tratatuko diren gaiak hauek izango dira:

1. Atala.- Heziketa estatistikoaren eta helburuari eta beronen premiari buruzko kontsiderazio oroko-//rrak: Estatistika, egoera errealkak estudiatzeko tresna moduan ikusita. Gizartea egiten duen estatistikeri kualifikatuak, profesionari konpetenteen eta hiritar kultoen eskaria.

Estatistikaren irakaskuntza bigarren mailan: estatistikaren irakaskuntzaren problema bereziak eta helburuak maila honetan. Estatistikaren ulerpena 11 urtetik 16ra bitartean. Tkasketa-planaren garapena. Diferentzia nazionalak. Erresuma Batuko egoeraren adibideak. Estatistika 16 urtetik 18ra bitartean. Estatistikaren ezagutza orokor, praktika, kontzeptu eta metodoeentzako premia desberdinak.

2. Atala.- Heziketa estatistikoa hirugarren mailan: Matematika, Estatistika eta beste disiplina batzueko ikasleak. Estatistika-maila. Programa eta heziketa-materialaren adibideak. Estatistikaren aplikazio-kurtsoak. Maisutza.

3. Atala.- Praktika estatistikorako heziketa estatistikoa: Estatistikariaren eta ikertzailearen arteko/elkarrengina. Konsultagintzako problema tipikoak. Estatistikako sail unibertsitario baten aholkulari/ eta zerbitza-funtzioa.

STATISTICAL EDUCATION

1.- INTRODUCTION

I shall be considering the following basic matters:

- why we should teach Statistics.
- to whom we should teach Statistics.
- what emphasis should be placed on different aspects of statistical method and principle
- how we should motivate, present and illustrate statistical ideas to the student.
- syllabus and schemes of work.
- specific detailed examples to illustrate the general conditions.
- the broader aspects of statistical education as they affect the professional worker and the ordinary citizen.

Whilst I shall seek to represent the wide range of / different opinions on these matters it is inevitable // that I shall concentrate mainly on my own personal attitudes to them. I have written quite extensively on this/ theme and I shall be drawing widely (with appropriate // attribution) on much material published by myself and /// others.

The particular topics I intend to cover in the different sections will be as follows.

Section 1.- General considerations of principle and aim. The need for statistical education. Statistics as a tool for study of real situations: the requirement of society for qualified statisticians, competent practitioners and enlightened citizens. Teaching general Statistics at // school-level. Special problems and aims of teaching Statistical awareness in the 11-16 year olds. Curriculum development. National distinctions. Illustrations from the situation in the UK: Statistics for 16-18 year olds. Different needs for general knowledge, practical technique, formal concept and method.

Section 2.- Statistical education in Colleges and Universities. Students of Mathematics, Statistics and of other disciplines. A degree course in Statistics. Illustrations of syllabus and teaching material. Service courses in Statistics. Masters degrees.

Section 3.- Statistical education for statistical practice. Statistical consultancy: interaction of statistician and research worker. Typical consultancy problems. The / service and advisory role of a University Statistics Department.

1.1.- GENERAL COMMENTS

I recall as a young university lecturer in Statistics a visit by a very eminent Russian probabilist renowned for his research but also for his skill in teaching/ probability and Statistics to all ages of students from 9 to 90 and from a variety of backgrounds. I asked him / how he could be so successful as a teacher. He replied / in a rather puzzled manner that there was no difficulty in teaching Statistics (or indeed any subject); all / you had to do was to play interesting games, chosen to/ suit the stage of development and special knowledge of / the students, and where the student had a chance to win/ the game! There was much in this response that is crucial -interest, challenge, success, practical relevance- to my view of the teaching of statistics.

Statistics nowadays cannot be avoided. It is part / of the society in which we exist. Radio, TV, newspapers/ make constant statistical demands on the citizen, to which he must be able to respond. Every schoolchild needs some exposure to informal ideas of the statistical reasoning set within his social everyday experience. He needs to / be 'statistically numerate' (i.e. to be able to interpret tables and draw broad inferences, not to know about set theory and event spaces). At another level anyone concerned with quantitative matters (the sociologist, engineer, doctor) needs some knowledge and experience of statistical methods. Finally, we need the professional statistician: as educator, innovator and solver of hard statistical problems. This trichotomy of

Enlightened citizen
Practical statistical data analyst
Professional statistician

should fashion the process of statistical education. // Great strides have been made in the last 30 years from a time when a tiny minority studied a small amount of Mathematical Statistics at University to a situation in some (but not all countries) where Statistics is beginning to feature with all three emphases throughout the education spectrum.

We shall consider later in some detail the needs of the latter two groups: statistical data analyst and professional statistician. This concerns mainly the university and college sectors, although some statistical method should be encountered at the 16-18 year levels in / schools.

Statistical numeracy at the secondary school level/ (say 11-15 years) will become more and more essential. // But it is not only the school-child who needs educating; society at large needs educating to the need for such a process. Administrators, governments, professional societies must all play their role, and they are // slowly beginning to do so. The International Statistical Institute has two task forces -one at School level/ (chairman: Prof. R.M. Laynes also of Sheffield University). The recent TOTSAS report [BARNETT (1)] clearly/ illustrates the progress (or lack of progress) in different countries. The recent ISI First International / Conference on Teaching Statistics (ICOTS I) was highly/ successful and informative. See GREY et al (2). ICOTS / II is planned for August 1986 in Victoria, BC, Canada.// We have conducted major curriculum development programmes in Sheffield for the age ranges 11-15 and 16-18. / But do not underestimate the cost in time and money to/ develop, at a cost of about 65,000.000 pesetas.

The journal *Teaching Statistics* and the new "Centre for Statistical Education" in the UK are also welcome / initiatives.

1.2.- TEACHING STATISTICS IN SCHOOLS

Space does not permit a detailed discussion of // this matter even in a single national context. The TOTSAS report [BARNETT (1)] provides much detail. Also // Chapter 1 of the TOTSAS report, HOLMES et al [(4),(5)], and BARNETT et al (3) give extensive coverage of the UK situation. We shall pick up just a few points here for/ illustration. Statistics has featured in school syllabuses in the UK for about 25 years: as material in public examinations at Ordinary level (O-level; 16 year olds)/ and Advanced Level (A-level; 18 year olds). In the // past there has been a tendency to include too much material and to present it in too formal (mathematical) a manner. Supporting text-books were not good and teachers not well-prepared for teaching Statistics. We are seeing improvements on all fronts, although teacher // training remains a real problem. There is increasing / practical emphasis in O-level and A-level syllabuses // and teaching: in some cases substantial practical projects are required. We see this clearly if we look at / some recent A-level questions and some recent practical project topics at this level BARNETT (1). Two questions from a Joint Matriculation Board A-level examination paper in Statistics are shown in Appendix I. Appendix II shows a selection of practical project topics for the /

Joint Matriculation Board A-level Statistics syllabus. // Contrast these with the following question which is a // slightly disguised version of one which once occurred at A-level for another board.

A continuous random variable X has probability density function

$$f(x) = a + bx + cx^2 \quad (0 < x < 2)$$

The mean and standard deviation of X are $8/5$ and $1/5$ respectively. Determine a , b and c .

This is quite straightforward. We have to solve 3 equations

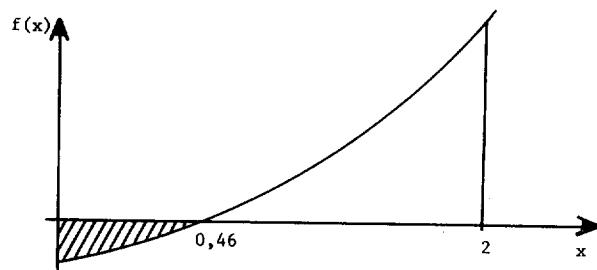
$$2a + 2b + \frac{8c}{3} = 1$$

$$2a + \frac{8b}{3} + 4c = 8/5$$

$$\frac{8a}{3} + 4b + \frac{32c}{5} = 13/5$$

$$a = 3/20 \quad b = 3/20 \quad c = 15/40$$

But what sort of statistics question is this? It involves no probabilistic or statistical concepts. In what/ circumstances would such an $f(x)$, or such a knowledge, provide a realistic model. It is simply a pure mathematics / exercise. Indeed suppose we were to draw $f(x)$ (as clearly the examiners did not!). It is negative over about the / first 25% of its range!!



But, what of the younger age range (11-16 years) where we want to introduce basic notions of data handling / and statistical reasoning? This is not easy. The popular image of Statistics as a means of misleading or 'probing everything' is against us, as are the constant errors of interpretation in the media.

Furthermore, there was until recently no published / material as support for the type of extended, practical-project-oriented, approach needed at this stage. In // 1975, I was successful in persuading one of the government supported agencies (the Schools Council) to fund a curriculum development project for 11-16 year olds. The aim was to produce about 30 units of teaching material, / each taking about 5 hours to teach, with about 7 for // each year group covering various types of practical context (Social Sciences, Sciences, Humanities, Mathematics etc.) and a wide range of probabilistic and statistical/ ideas. The statistical concepts and methods were 'submerged' in the practical examples so that they were learnt/ implicitly rather than directly in the sense of

'a probability distribution is...'

or

'the sample mean is defined by...'

Instead the problems guided pupils to self-discovery of the concepts. Each unit had a published set of 'pupil's notes and of teachers' notes- they needed to inform the teachers (to an extent) as well as to instruct/ the pupils. The units were prepared, printed tested in / about 60 schools, validated, rewritten, etc. The full material of the project is now published:

HOLMES et al: (1980): *The structure of the project // approach.*

HOLMES et al (1981): *A review of the UK situation in schools and*

SCHOOLS COUNCIL PROJECT ON STATISTICAL EDUCATION // (1980-81): *The units.*

To conclude this brief discussion of teaching Statistics at school level Appendix III shows extracts from some of the units prepared by the Schools Council Project/ on Statistical Education (11-16).

REFERENCES

- 1.- BARNETT, VIC (Ed) (1982) *Teaching Statistics in //*

Schools throughout the World. Voorburg, The Netherlands: International Statistical Institute.

- 2.- GREY, D.R., HOLMES, P., BARNETT, V. AND CONSTABLE, G. M. (eds.), (1983). *Proceedings of the First International Conference on Teaching Statistics.* Vols I and II. Sheffield: Teaching Statistics Trust.
- 3.- BARNETT V., CABLE, P., CONWAY, D.A., HARRIS, J., HOLMES, P., MANN, E.J., MORLEY, S.A., OLDFIELD, S.J., // WILKINSON, R. (Eds) (1979). *Statistical Education and Training for 16-19 year olds.* Sheffield: Teaching Statistics Trust.
- 4.- HOLMES, P. ET AL (Eds) (1980). *Teaching Statistics / 11-16.* Slough, UK: Foulsham Educational.
- 5.- HOLMES; P., KAPADIA, R., RUBRA, G:N: AND TURNER, D. / (Eds) (1981). *Statistics in Schools 11-16: A Review.* London: Methuen Educational.
- 6.- SCHOOLS COUNCIL (1980-81). *Statistics in your World / (Pupils Notes and Teachers Notes)* Slough, UK:Foulsham Educational.

2.- TEACHING STATISTICS IN UNIVERSITIES AND COLLEGES

The situation in universities and colleges is intrinsically more straightforward than in schools in many respects:

Students should be committed and motivated usually / they choose to study Statistics (although service // courses for other-subject specialists e.g. biologists, engineers etc. may not be so fortunate in // their clientele).

Lecturers should be professionally trained in Statistics -although this does not guarantee skill, or wide practical experience.

Material is often well-structured and documented: we/ know what we want to teach and good text books are // (usually) available for background support -but I do/ not favour text-oriented teaching and some of the // practical and applied topics do not even now have // good published coverage.

We need to distinguish different types of course and/ student-group. At undergraduate level we have:

- i) (Short) Service courses for students of Agriculture, Biology, Economics, Engineering, Medicine, Social Science, etc.: so called 'Service Courses'.
- ii) Probability and Statistics courses for Mathematics students.
- iii) Probability and Statistics courses for specialist/ Statistics students.

At postgraduate level there are taught Masters Degree courses, and support courses for research students.

The last 30 years has seen a vast proliferation of / Statistics teaching in universities in the UK -from a situation where individual courses in elementary Statistics were occasionally given (apologetically perhaps) as light relief in Mathematics degree courses, to the present comprehensive presentation of the subject as a discipline in its own right or as the experimental methods/ and analysis basis of subjects such as Psychology, Geology, Sociology, Economics and so on. Most universities have Statistics Departments (or substantial groupings in / Mathematics Departments)- most offer first, and postgraduate, degrees in Statistics or play a major role within Mathematics degrees. Service courses in Psychology etc., may be given by specialists in the Statistics Department (with, one hopes, a firm interest and commitment to the applications subject area)- or by subject specialist /// (with, one hopes, wide experience and a sound interest / in Statistics).

In the main, the Universities are not substantially/ dependent on the school, or examinations, background of their students. Degrees in Statistics 'start from /// scratch', likewise service courses: either because of a belief that no common background knowledge even of basic principles can be assumed or because of a dissatisfaction with teaching in the schools, allied to inappropriate examinations syllabuses.

But is the schools, and Examining Board, situation/ entirely irrelevant to, or unaffected by, teaching at / the University level? I believe that the Schools and Examination Boards have little influence on the Universities, but the reverse is not true. A somewhat 'holier // than thou' attitude, fostered in earlier days (but still persisting to some extent) by mathematical snobbery has/ had its influence on the non-university-level teaching/ and syllabus construction. Earlier university attitudes/ may be partly to blame for the over-formal, superficial,

unreal, characteristics of many school syllabuses. In // spite of some honest individual attempts to 'bridge the/ communication gap' between university Statistics, school Statistics and Applied Statistics (in industry, research and commerce) there is still too much reluctance to 'mix it'. University staff must do more to help the subject / at school level and in the commercial world. Whilst they / take the attitude that non-university work is 'academically inferior', 'wrongly conceived' or 'badly taught' and therefore irredeemable they do nothing to remedy any /// truth there might be in such beliefs.

A review of the present situation in UK universities is given by BARNETT et al (1). Some extracts are relevant to general considerations.

"The growth in the teaching of Statistics in the higher education sector has been dramatic. For example, it was only in 1949 that the first University professorship in Statistics was filled outside Oxbridge and London,... In less than 30 years we have reached a stage // where there are about 70 chairs in Statistics in British Universities, and over 700 staff employed specifically / to teach Statistics in British Universities and Polytechnics (the vast majority within the Universities). Statistics is taught as a specialist subject of study or as / ancillary material in a wide range of applied disciplines... The pattern of teaching, and its level and extent, varies widely from one institution to another..."

"The emphasis and style of courses has also ranged / widely from formal, highly mathematical, concern for probability, stochastic process, and statistical inference, to informal, 'cook-bookish' presentation of statistical/ method. There are pleasing signs however of improvement/ of balance with increasing attention to practical motivation and illustration from real-life problems in diverse areas. Sound practical Statistics is vitally important / but it is not easy to teach; most universities now seem/ committed to doing so, with some rewarding results. Modern attitudes to the analysis of data demand a problem/ oriented approach..."

"Any hope of developing statistical consciousness // (through our teaching of the students) rest in meaningful demonstration that most problems of importance are / statistical. That must take place over a long period /// with a deal of practical reinforcement. (Particularly at school level) Statistics has to be seen as a thread // through all aspects of human endeavour, not as a dry component of the Mathematics lesson- it is part of the geo-

graphy lesson - it is part of metalwork... Such matters / are not easily taught. They require long-term exposure / to real problems, rather than a few stylised classrrrom / illustrations. They depend on teaching technique that is leisurely and problem-based rather than in the usual mathematics mould of placing one small brick on top of another until the house is built for all time...

"And if not a 'brick on brick' approach, then what is the method of instruction to be? It is essentially as follows. We first inspect the statistical world out of the window. We then look at the same view through binoculars, then through a microscope. Let the schools do the window gazing and do it well. The Universities and Polytechnics may then be better placed to properly employ the binoculars and the microscope".

I shall not spend time discussing the inadequacies / of the present situation. Instead I shall give my views/ on what I regard as aspects of sound course construc- // tion, syllabus and detailed content -with examples to // illustrate the views- based on my experience at Sheffield and elsewhere.

I will start with the most comprehensive undergraduate programme.

2.1.- A DEGREE COURSE IN STATISTICS

On the English model, students enter university at age 18 usually to study one (or two) subjects in depth / over a three-year period. Some students wish to study Mathematics with specialisation in Statistics, or Statistics specifically, to obtain a degree in Statistics. There is much they must be taught in statistical and probabilistic topics over the three years. It is crucial that they learn not only the mathematics of the subject -but/ most important that they are trained to operate as professional statisticians: to solve real problems. This poses two distinct needs

a)A coordinated programme of statistical concept // and method study keeping pace with the appropriate level of Mathematics.

b)A comprehensive, and progressively more sophisticated scheme of practical work throughout the three/ years.

The required level of Mathematics under a) is inevitably as demanding as that for mathematics students per/ se over the relevant topics. This is best covered by the

Mathematics staff in liaison with the statisticians with/ regard to times, and depths of coverage. The Mathematics/ and Statistics students will study together over the to-/ pics of common interest. This requires a good inter-departmental relationship. Ideally a 'school structure' -with/ departments of Statistics, Pure Mathematics, Applied Ma-/ thematics, Computing having independent status and con- / trol, but being closely linked in their teaching func- / tions and pastoral roles. As well as easing the coordination of different topics there is a further advantage in/ that a common admissions policy can be employed, with a // common first-year programme of study, delaying choice of specialisation to a more informed stage in the second // year.

The programme of practical work under b) should keep/ pace with (but not slavishly follow) the concept and me- / thod components of the course in the first two years. By/ the third year however the student should have been // brought to a stage where he has developed a real feel for data modelling, scrutiny and analysis (not an easy task)/ and can be exposed to more substantial real problems whe- re no immediately obvious existing technique is indicated or even relevant. The consulting environment of the de- / partment can assist here- see comments later on this sche me

So how would a three-year statistics degree look? I / would propose, and have operated, the following scheme.

Time	Year		
	1	2	3
1/3	Probability and Statistics	Probability and	Probability
1/3	Pure Mathematics. Applied Mathematics	Statistics Mathematical	and
1/3	and Computational Methods	Methods and Principles	Statistics Mathematics

Year 1. A compulsory course for all Statistics and Mathematics students (ensuring some contact with Statistics

for all Mathematics students).

Course details About 150 hours made up of 80 hour of lectures, 25 hours of examples classes and 45 hours of practical sessions (in units of 2-3 hours).

Syllabus: A typical example of syllabus and support/reading list for this type of first course is shown in / Appendix IV (A recent version of the first-year course / for Special Honours Mathematics students at Sheffield // University).

Such a syllabus could, of course, be taught with a / variety of emphasis -from purely formal and mathematical, to non-rigorous rote application of 'recipes'. Neither extreme is reasonable. Lectures should develop concepts and methods at a sound mathematical (conditional / on the student's current ability) but major emphasis // should be given to developing statistical consciousness/ and understanding, and skill in applying ideas to real / problems.

The examples classes discuss difficulties on sets / of about 5 exercises which students have worked on each week and have handed in for correction. Exercises are designed to be as realistic and useful as possible. We // avoid dull and unmotivated problems (sadly typical of many introductory texts) such as:

Q1 Suppose that in 8 throws of a die we obtain 6, 5, 1, 1, 6, 4, 2.

Calculate the sample mean, the sample variance and the sample standard derivation.

Q2 A continuous variable X has probability density // function.

$$f(x) = a^2 - x^2 \quad \text{for } -\sqrt{a} \leq x \leq \sqrt{a}$$

If the variance of X is 0.5 determine a , and // $P(0 \leq x \leq 0.4)$.

I would be quite sympathetic to the answer, 'Why?', / in each case.

What have these to do with Statistics? They are // again unmotivated numerical and mathematical exercises./ Some idea of how we can approach the real need for reinforcement, understanding and application is seen in the/ five examples in Appendix V, from the first-year course/ just described.

The practical sessions provide opportunity for more detailed modelling, data handling and interpretation. There is also an extended practical where students work for two weeks on a project and write a report for assessment. Appendix VI shows a practical exercise on simulation of a doctors waiting room: the extended practical / might ask students to use the complete set of data from/ the class, to analyse it and to write a reasoned technical report and a letter giving recommendations for the / doctor.

Year 2. Here the Statistics student will be spending about half his time on Probability and Statistics (the/ Mathematics student can still study some Statistics for interest, but for much less than half his time). A course of similar size and structure to the first year course (with examples classes and practical sessions) covers:

Distribution theory
Statistical principles and
Statistical methods.

This is a key course in the three year programme, covering most of the central concepts and essential methodology needed by the professional statistician -normal / distribution theory, estimation and testing, linear models, regression, association and some sequential, finite-population and non-parametric methods of analysis.

The remainder (one-third) of the second years programme (with 2 lectures and an examples class each week) is concerned with intermediate Probability theory and an introduction to Stochastic processes.

Year 3. In the third year the Statistics student // spends about 80% of his time on topics in Probability // and Statistics. A course on a mathematical topic can also be taken for support or interest. Here quite a wide / extent of choice is possible. Certain topics are regarded as central and the student is strongly encouraged to take.

Practical and Applied Statistics
Inference and Decision Theory
Time Series
Design of Experiments
Multivariate Analysis

These represent about 50% of the statistics material. The remainder is chosen to reflect personal interest and career aims from the following areas:

Probability, Measure and Integral
Mathematical Probability
Medical Statistics
Stochastic and Statistical Methods in Industry
Operations Research
Stochastic Processes
Advanced Stochastic Process
Statistical model building
Sampling Theory
Non-parametric Methods and Robustness
Spatial Processes
Categorical data Analysis

At this stage some courses are highly technical or / mathematical, others present detailed methodology, others are vocational. The Practical and Applied course makes quite sophisticated demands on the student.

Appendix VII describes the aims of this course. Experience of computing, using calculators, micro-computers, mainframe machines, packages etc. is of course included/ here; but is not allowed to dominate the statistical component. For some references on Statistics Teaching and / Computing see e.g. EVANS (2), MEAD and STERN (3).

2.2.- UNDERGRADUATE MATHEMATICS STUDENTS

The question of what is appropriate teaching emphasis and material for Mathematics students is answered to a large extent by the comments above on what is needed / by Statistics students. Any mathematician should be conscious of the need to apply his Mathematics and should be aware of the breadth of applicable Mathematics. Thus I / believe all should have an introductory course of the type described above (the first year component of a Statistics Degree programme). They need to have concepts well-motivated and rooted in applications. But beyond the // first year, the Probability and Statistics topics for Mathematics students can be chosen to suit individual tastes from the Statistics programme. Many of the courses / in the areas of Probability and Stochastic Processes /// will have mathematical appeal (as will Design of Experiments, Multivariate Analysis, Operational Research etc.) and will also serve to broaden the horizon of the mathematician to enhance his professional competence.

2.3.- SERVICE COURSES IN STATISTICS

We read in BARNETT et al (1) that many "degree // courses in applied subjects now require service courses/ in Statistics, usually, for their first year students. / These include Agriculture, Economics, Social Sciences, //

Medicine, Psychology, Librarianship, Engineering, Biology, Geography, Archaeology, etc. It is important that // such subject specialists have some relevant knowledge // about statistical reasoning and technique. One wonders / sometimes, however, whether the typical first-year 10 // lecture course including Markov chains and cluster analysis quite meets the bill in terms of relevance, scope // and immediacy! There is a need to properly re-assess the style, timing and content of service courses. Many such/ courses are given by staff from Statistics Departments / or Groups, and if soundly based require the member of // staff to thoroughly immerse himself in aspects of the // main discipline in order to make Statistics appear relevant and interesting. There are also demands on teaching technique -the Biology student is not a mathematician // manqué! Some service courses are preserved by the // applied discipline -Psychology seems the extreme example of the attitude 'we know what our students need'. Such/ an attitude may have some merit, if it really happens to be true and if there is sound enough knowledge of statistics for good teaching. The danger is of arbitrary and // fashionable choice of statistical topics and of terminological isolation".

The most important factor is that the student should find the Statistics material interesting and relevant, / rather than feeling that he is asked to attend some dull set of lectures with no contact with his subject and for no good reason. The material needs to be developed at // leisure with copious practical examples from the appropriate subject area. This means that even the briefest / coverage requires a minimum of contact of about 20 hours (say 10 lectures and 10 examples sessions). Examples sessions must be adequately staffed for all students to // discuss their difficulties (perhaps 1 supervisor to at / most 12-15 students). The applied department should // stress the importance of the Statistics course to its // students.

It is tempting to believe that practical relevance / is not crucial if the statistical development is sound./ This is wrong from the psychological viewpoint. Consider the following exercise in a short service course for Civil Engineers.

Q1. Let p denote the probability that in any one // year high tides that year will overtop the Thames Embankment and cause flooding in Central London.

If the Thames Barrage, which will remove the danger, had not been completed for another 4 years, and if the sizes of the high tides in each of these 4 years are/

independent, find, in terms of p , the probabilities that

- i) there will be no tidal flooding before completion/ of the Barrage,
- ii) there will be tidal flooding in at least one year/ before completion of the Barrage,
- iii) there will be tidal flooding in 2 of the next 4 // years.

Evaluate these probabilities if (a) $p = 0.05$, (b) // $p = 0.10$, discussing whether such p values are realistic in this context.

This is expressed in terms which will be relevant to the Civil Engineering student. The following exercise is essentially the same statistically.

Q2. Let p denote the probability that a standard volume of a laboratory prepared culture contains external/ contamination (the growth of bacteria). If four independent samples of culture are prepared find, in terms of / p , the probabilities that

- i) no samples are contaminated,
- ii) at least one sample is contaminated,
- iii) precisely 2 samples are contaminated.

Evaluate these probabilities if (a) $p = 0.10$, (b) / $p = 0.05$, discussing whether these p values are realistic in this context.

Exercise 2 will be appropriate for a service course for Biology students; the civil engineers would probably derive far less motivation and benefit from it. Note that even the specification of p -values needs care. / Set $p = 0.85$ in exercise 1 and the civil engineers will/ lose confidence in the lecturers care or knowledge in / his choice of material. Now contrast exercises 1 and 2 / with the following one.

Q3. Let p denote the probability that a ball drawn / at random from a box containing a large number of balls/ is black. If four balls are drawn independently find, in terms of p , the probabilities that

- i) no balls are black,

ii) at least one ball is black,

iii) precisely 2 balls are black.

Evaluate these probabilities if (a) $p = 0.3$, (b) // $p = 0.5$.

It is likely that biologists and civil engineers // would fall asleep if given this exercise!

2.4.- TAUGHT MASTERS COURSES IN STATISTICS

It has been common practice in the UK, US and elsewhere to offer a one-year programme of study leading/ to a Masters degree in Statistics. Such courses aim primarily at Mathematics graduates (sometimes others from/ quantitative areas of study) with a view to their 'conversion' to Statistics. The typical pattern is of two / terms coursework with formal examinations followed by / four months of individual work and the writing of a dissertation on that work. It is my view, and my practice, for the individual work to comprise a detailed statistical analysis from some applied field, usually computer-assisted, with a full written report on it.

The coursework might involve 250 hours of class contact; topics of study will be not dissimilar to those / in the Statistics Degree there is about 700 hours of // class contact in Statistics alone, and it is clear that only a proportion of this material could be covered in the Masters course. It looks likely that the proliferation of full first-degree programmes in Statistics will limit the usefulness of taught Masters courses and perhaps make us change our views on what they should be designed to achieve.

REFERENCES

- 1.- BARNETT, V., CABLE, P., CONWAY, D.A., HARRIS, J., / HOLMES, P., MANN, E.J., MANN, E.J., MORLEY, S.A., / OLDFIELD, S.J., WILKINSON, R. (eds.) (1979). Statistical Education and Training for 16-19 year olds. / Centre for Statistical Education,
- 2.- EVANS, D.A., (1973). The influence of computers on/ the teaching of statistics. J. Roy. Statist. Soc. / A., 136, 153-190.
- 3.- MEAD, R. and STERN, R.D. (1973). The use of the computer in the teaching of statistics. J. Roy. Statist. Soc. A., 191-205.

3.- STATISTICAL EDUCATION FOR STATISTICAL PRACTICE

I enjoy solving mathematical problems; I enjoy doing cross-words! It is most rewarding if a statistical problem leads to some interesting mathematics. But Statistics is not just Mathematics. Statistics as a subject // needs those interested in following the mathematical pathways to Probability Theory or Statistical Inference (very much the pattern of academic work in parts of Europe/ and the US). But my attitude is based on the British tradition of empiricism. I believe that statistical methods are designed to solve real problems (whether in Archaeology, Industry, Medicine or Spanish Literature). The stimulus should come from such problems. This is not to deny the importance of fundamental philosophical or mathematical study which often arises from such stimulus.

I recall being approached by a local doctor to 'analyse a simple set of data -just fitting a straight line'. He wanted to compare four different instruments for measuring human lung function: to assess relative calibration and relative accuracy. Details of this (and some related) problems are described in BARNETT (2), (3), (4).

What do these examples tell us about statistical education? Essentially three things.

1. Communication between the statistician and research worker is part of the wider educational process // for each. The statistician gains in experience and ability to employ his art by contact with diverse disciplines. This also stimulates new developments within the armoury of statistical methods. The research worker gains/ in the realisation of what Statistics has to offer for / the understanding of his own subject area. Thus this role of the statistician as a consultant and collaborator/ has vital educative overtones. Many have written on this e.g. BARTHOLOMEW (5), BENJAMIN (7), SPRENT (20), CHANTER (9), FREEMAN (13), GREENFIELD (15), MARQUARDT (18), HOOKE (16), BARNETT (4) considers the ubiquitous role of the / statistician under the heading of 'The Statistician:Jack of all trades, master of one?'.

'the statistician must be a translator and a communicator: he needs to understand enough of other people's disciplines to appreciate their problems. He / must express these in statistical terms, in cooperation with the experimentalists develop and use appropriate tools and, most important, communicate answers in an understandable way. So it would seem that

he has a somewhat wider brief than many -as a mathematician/statistician, a computer, a lay philologist (physician, nuclear physicist, you name it) and not least a communicator (this latter facility is not // usually regarded as the stock in trade of the scientist.) All in all he needs to be master of his own / statistical trade, but Jack of many others.'

The point is well illustrated by a problem in the // oil industry which required the resources of many different disciplines BARNETT AND LEWIS, (1); BARNETT, (4).

2. The student needs to be trained to do practical / statistics, and consultancy work. This is of course obvious as far as specific applied statisticians are concerned. The agricultural, industrial, medical statistician, the biometrician or psychometrician, all need to be/ trained for their speciality. See e.g. BISHOP (8), DEANE/ (11), SPICER (19), FINNEY (12), FREEMAN (13), GREENFIELD/ (15). We have considered some of the problems in lecture 2. But I would claim that the professional (non-subject-specific) statistician, mathematical statistician or even mathematician must also be trained to apply his knowledge widely in real problems to benefit others and to advance statistical knowledge.

I have stressed throughout the role of practical // work in education: particularly in lecture 2 in relation to University statistics majors. This should be structured as part of a training in statistical consultancy. The consulting experience of lecturers and professors gives/ a vital input of material for practical classes; it is / even better if the student, at undergraduate or postgraduate level, can become involved in real consulting work (if only as an observer). See COX (10), BASKERVILLE (6).

But this latter point has perhaps the most important implications for a university Statistics Department.

3. The service, and statistical advisory, role, of a university Statistics Department is vital for creating / the proper educational environment. I believe that any / Statistics Department must be seen as providing a university-wide practical facility, both in the teaching of // students from other disciplines but, more important, in/ offering a full statistical advisory service to all university staff, and postgraduate students (and to outside organisations, if logically feasible). See TUKEY, BOX, HARTLEY and KEMPTHORNE (1968), GIBBONS and FREUND // (1980). This requires more than just a casual attitude / of being prepared to help if anyone happens to ask.

The advisory service should be widely advertised. Staff/ (not necessarily all) must be committed to providing the service, and have the confidence to do so. The university needs to recognise and welcome the facility, to the / extent of providing adequate resources, including computing equipment and a staff member to conduct computational work. There is nothing worse than offering a service and saying that someone's problem is soluble in principle, but that lack of time, computing power etc. pre- // vents it from actually being solved. In the Sheffield Department of Probability and Statistics, about 8 out of / 15 staff are involved in the Statistical Advisory Service with one as supervisor of the service. Over 100 major consulting tasks are carried out each year, covering // about 40 departments. Every two weeks a one-hour Statistics Clinic is manned and answers given to simple pro- / blems or contacts established for more substantial ones. A statistical consulting assistant does the university // mainframe computers. Some recent typical examples from / the log-book read as follows:

Dr. Sherwell, Dentistry 'Trials of a new tooth pas-/te'.

Dr. Brownton, Education '3-way classification on binary responses use of logistic regression.'

Interesting work comes from unusual areas. We have / done much in the fields of Archaeology and Pre-History./ Dr. N.R.H. Fieller was approached with some data on //// lengths and breadths of horse teeth from the Devensian / era (15000 years ago) and asked if the horses had died / at particular seasons of the year or uniformly through- / out the year. This lead to fascinating and sophisticated analyses and developments (e.g. using cosine quantogram methods) over a long period to examine data for 'pe-riodicities'. Other examples included Viking feet (was / there a basic unit of measurement), Troy weights (again/ seeking a unit) and numbers of tree-rings in Hazel trees from 5000 BC (indicative of husbandry).

I started by stressing the need in statistical edu-

cation for the subject to be 'alive' and interesting -to pose problems to answer, and a chance to answer, and a / chance to actually solve them. Success in teaching Statistics depends on persuading others that Statistics is important (even entertaining). This point is well made (if somewhat ironically) in a well-known quotation that /// bears repeating-yet again!.

It serves as the frontspiece to the second volume of Kendall and Stuart (17) and is attributed to the novel //

The Undoing of Lamia Curdleneck by K.A.C. Manderville.

"You haven't told me yet", said Lady Nuttall, "what/ it is your fiancé does for a living".

"He's a statistician", replied Lamia, with an anno- / ying sense of being on the defensive.

Lady Nuttall was obviously taken aback. It had not // occurred to her that statisticians entered into nor-/ mal social relationships. The species, she would have surmised, was perpetuated in some collateral manner,/ like mules.

"But, Aunt Sara, it's a very interesting profession", said Lamia warmly.

"I don't doubt it", said her aunt, who obviously doubted it very much. "To express anything important in / mere figures is so plainly impossible that there must be endless scope for well-paid advice on how to do // it. But don't you think that life with a statistician would be rather, shall we say, humdrum?".

Lamia was silent. She felt reluctant to discuss the / surprising depth of emotional possibility which she / had discovered below Edward's numerical veneer.

"It's not the figures themselves", she said finally,/ "it's what you do with them that matters."!

REFERENCES

- 1.- BARNETT, V. and LEWIS, T. (1967). A study of low-temperature probabilities in the context of an industrial problem. *J. Roy. Statist. Soc., A*, 131, 410- / 433.
- 2.- BARNETT, V. (1969). Simultaneous pairwise linear // structural relationships. *Biometrics*, 22, 129-142.
- 3.- BARNETT, V. (1971). Fitting straight linear functional relationship. *Applied Statistics*, 19, 135-144.
- 4.- BARNETT, V. (1976). The statistician: Jack of all // trades, master of one? *The Statistician*, 25, 261-279.
- 5.- BARTHOLOMEW, D. J. (1973). Post-experience training/ for statisticians. *J. Roy. Statist. Soc, A*, 133, 65-70.
- 6.- BASKERVILLE, J.C. (1981). A systematic study of the/ consulting literature as an integral part of applied training in statistics. *American Statistician*, 35, / 121-123.
- 7.- BENJAMIN, B. (1971). The Statistician and the manager. *J. Roy Statist. Soc. A*, 134, 1-13.
- 8.- BISHOP, H.E. (1964=). The training of government statisticians. *J. Roy. Statist. Soc. A*, 127, 211-215.
- 9.- CHANTER, H.E. (1978). Communication and the statistician. *BIAS*, 5, 49-56.
- 10.- COX, C.P. (1968). Some observations on the teaching/ of statistical consultancy. *Biometrics*, 24, 789- // 802.
- 11.- DEANE, MARJORIE (1964). The training of statisticians for economics and business. *J. Roy Statist. / Soc., A*, 127, 216-218.
- 12.- FINNEY, D.J. (1968). Teaching biometry in the University. *Biometrics*, 24, 1-12-
- 13.- FREEMAN, G.H. (1978). Consultation and research in a statistics section. *BIAS*, 5, 22-32.
- 14.- GIBBONS, J.D. and FREUND, R.J. (1980). Organisations for statistical consulting at colleges and universities. *American Statistician*, 34, 140-145.
- 15.- GREENFIELD, A.A. (1979). Statisticians in industrial research: the role and training of an industrial consultant. *The Statistician*, 28, 71-82.
- 16.- HOOKE, R. (1980). Getting people to use statistics / properly. *American Statistician*, 34, 39-42.
- 17.- KENDALL, M.G. and STUART, A. (1973). *The Advanced // Theory of Statistics* (3rd Ed.) Griffin, London.
- 18.- MARQUARDT, D.W. (1979). Statistical consulting in industry. *American Statistician*, 33, 102-107.
- 19.- SPICER, C.C. (1964). The training of medical statisticians. *J. Roy Statist. Soc. A*, 127, 219-221.
- 20.- SPRENT, P. (1970). Some problems with statistical // consultancy (with Discussion). *J. Roy. Statist. Soc. A*, 133, 139-164.
- 21.- TUKEY, J.W., BOX, G.E.P., HARTLEY, H.O. and KEMPTHORNE, O. (1968). The future of departments of statistics. In Watts, D.G. (Ed.) (1968) *The Future of Statistics*. Academic Press, NY. pp 103-137.

APPENDIX I*

6. In an investigation into the effectiveness of a particular course in speed reading a group of 500 students was split into two groups, A and B, of sizes 300 and/200 respectively, thought to have been chosen at random.

Those in group A were given no special instruction; / those in group B were given a course in speed reading. //

Each student was asked to read the same passage and / the time taken was measured. The results were

Group A: mean time 78.4 s, variance 14 s²,

Group B: mean time 77.4 s, variance 15 s².

Carry out a significance test to see if there is evidence that the course has improved reading speed. State carefully your null hypothesis, alternative hypothesis and final conclusion.

You learn later that, of the original 500 students, / 200 students had decided for themselves that they wanted to take the course in speed reading and that these students became group B. Discuss briefly how this/ might affect your previous conclusion.

7. A loom is used to produce cloth. Occasionally, at random, the machine falters and a flaw appears across // the cloth. In the past it has been found that the distance between flaws in the cloth follows an exponential distribution with mean 1 metre. State briefly // why an exponential distribution would be expected to serve as a reasonable model in this situation.

The 100 distances between successive flaws on a given day are shown in the following table.

Distance between flaws	Frequency
0 to 0.5 m	44
0.5 to 1.0 m	20
1.0 to 2.0 m	22
2.0 to 3.0 m	10
more than 3 m	4

(i) Plot these data as a histogram. On the same diagram draw an appropriately scaled graph of the density / function of the exponential distribution with mean/ 1.

(ii) Calculate the corresponding expected frequencies // for an exponential distribution with mean 1 and // test whether this model gives a reasonable fit to the data. (Two significant figure accuracy for expected frequencies is accepted).

APPENDIX II*

(a) Leisure Activities & Sport

Effect of a sight on the accuracy of a rifle.
Predicting football results.

Playing and watching sport.

Comparison of distribution of points obtained by football teams with that obtained by chance.
Leisure activities of sixth form students.

(b) Work

Commerce

The number of tills required in a particular sports / shop.

Investigation into pharmaceutical services.

Number of pumps required by a self-service garage.
Efficiency of checkouts at a local supermarket.
Comparison of national newspaper circulation with // that occurring locally.

Industry

Breakdown of lorries belonging to a particular haulage company.

The effect of an overhaul on a canning machine.

Is 52 the average number of matches in a box?

Fill height and content of 33 centilitre bottles.

Lifetime of light bulbs compared with manufacturers / claims.

Transport

Arrival of traffic at a junction in 2 different time/ periods.

Local road accidents 1974 and 1975.

How long do cars stay roadworthy?

* Reprinted with authorization of The Joint Matriculation Board, Manchester, U.K.

Pupils' modes of transport to school.
Effects and efficiency of a set of suburban traffic/lights.

(c) Sociological Phenomena

Trends and changes in consumer spending.
Road accidents and casualties.
Distribution over the country of people with a particular surname.
Relation between reading ability and report scores.
Analysing crime statistics.
Marriage and divorce.
Death rates in industrial and non-industrial areas.
Relationship between examination results and final / achievements.
Factors associated with people smoking.
Age distribution of male and female unemployment.

(d) Others

Do blood relationships increase the effectiveness of telepathy?
Prediction.

Other subjects

Mathematics

Are random numbers random?
Testing for random numbers.
Principles of estimation.
Markov chains.
Effect of number of throws on the distribution of faces showing on the throw of a die.

Science

Effect of soil conditions on the distribution of wild plants.
Breeding mice and survival.
Effect of water levels on germination of cress seeds.
Effect of light on growth of cress seeds.
The effect of smoking on lung cancer.
Size of pebbles and position on storm beaches.
Odour preference.
Factors affecting reaction times.
Weight and ages of different pupils.
Hair and eye colour of children and their parents.

Economics

The principle of maximising profits and the distribution of land use near the city centre.
Increasing the labour force.
Gold price and the exchange rate.
International Economic Development analysis.
Price rises in the local chemist's shop.

Languages

Sentence length in English and French novels.
Did Bacon write Shakespeare?
A grammatical analysis of the spoken and written news.

Geography

Do meander wave lengths vary with distance?
Population trends.
Land use variations.

History

Population survey of Stockport 1931 to 1971.
Age group populations of Huddersfield 1901 to 1971.

APPENDIX III*

LEVEL 2

E GOALS GALORE**

E1 Two seasons

Older people sometimes say that they used to see more goals at football matches. Try to find out what adults you know think. How could you decide if more goals used to be scored?

One way is to look at the goals scored in the First/Division 20 years ago and recently.

Look at Table 6. It shows the goals scored by the 22 teams in the First Division over two seasons. Brian Clough was a leading goal-scorer in the 1950s, scoring about 40 goals in a season. In the mid-1970s Macdonald was a leading scorer with 25 goals in the season.

a) Compare the goals scored by teams in the 1957/8 season. Do you think more goals were scored in the 1950s?

* Reprinted with authorization by W. Foulsham & Co., Ltd., Slough, Berkshire, U.K.

** From "on the ball", pp 12-14.

- b) You may need more information. Discuss what you need/ to know to compare goals scored in the 1950s and the/ 1970s.

It may help to find the total number of goals scored/ in each season. You could also look at the goals scored/ by the highest and the lowest scoring teams.

For the two seasons shown in Table 6 find:

- c) the total number of goals scored,
- d) the average (mean) number of goals, scored per team / (divide by 22)
- e) the lowest number of goals scored by a team,
- f) the highest number of goals scored by a team.
- g) Write a few sentences comparing the number of goals / scored in the 1950s and the 1970s. Mention any differences or similarities in the results and anything/ else you notice.

TABLE 6 GOALS SCORED BY TEAMS IN DIVISION 1 IN TWO SEASONS

Team	1957/8 goals	Team	1977/8 goals
Wolves	103	Notts F	69
Preston	100	Liverpool	65
Tottenham	93	Everton	76
WBA	92	Man City	74
Man City	104	Arsenal	60
Burnley	80	WBA	62
Blackpool	80	Coventry	75
Luton	69	Aston V	57
Man Utd	85	Leeds	63
Notts F	69	Man Utd	67
Chelsea	83	Birmingham	55
Arsenal	73	Derby	54
Birmingham	76	Norwich	52
Aston V	73	Middlesbro	42
Bolton	65	Wolves	51
Everton	65	Chelsea	46
Leeds	51	Bristol C	49
Leicester	91	Ipswich	47
Newcastle	73	OPR	47
Portsmouth	73	West Ham	52

Team	1957/8 goals	Team	1977/8 goals
Sunderland	54	Newcastle	42
Sheff Wed	69	Leicester	26



LEVEL 4

OTHER PROBABILITIES **

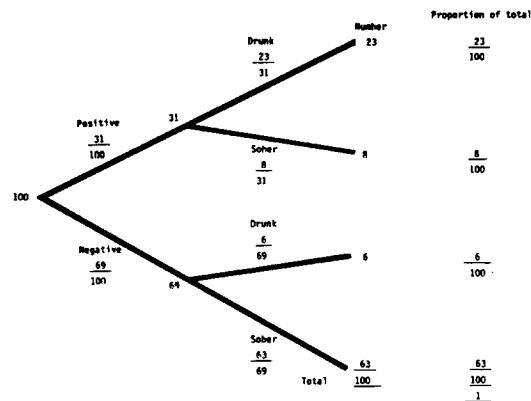
What does a positive reaction mean?

- a) How many drivers gave a positive reaction?
- b) Of these, how many were drunk?
- c) Of those who gave a negative reaction, how many were // drunk?

It is possible to draw a different tree diagram from // the information contained in the table.

** From "Testing Testing", pp. 14-15

Figure 2 Tree diagram showing further results



Using Figure 2, estimate the following:

- d) Probability (a driver who gives a positive reaction is sober)
- e) Probability (a driver who gives a positive reaction / is drunk).
- f) Probability (a driver who gives a negative reaction / is drunk).
- g) Probability (a driver who gives a negative reaction is sober).
- h) Probability (a driver is both drunk and gives a positive reaction)
- j) Probability (a driver gives a positive reaction).
- k) Can you find a relationship some of the answers to d/ j. If so, write it down first as a multiplication sum involving fractions and then in words, using probabilities.
- l) Make a copy of Figure 2, but insert the figures obtained from your version of Table 1.
- m) Work out the proportions based on your figures and insert them in the boxes.
- n) Use your tree diagram to find the answers to d to j // from your simulation.

APPENDIX IV*

FIRST YEAR SPECIAL (IS SYLLABUS)

INTRODUCTION TO PROBABILITY AND STATISTICS

Introduction to Data. Random variation with examples. // Methods of data representation: Frequency distribution, / histogram, pie chart, stem and leaf plots. Descriptive / measures: mean, median, quartiles, range, variance, standard deviation, interquartile distance, box plot, skewness and Kurtosis coefficients, data coding, correlation coefficient. Inference: sample, population.

Probability. Sample space and events. Definition of probability: classical, frequency. Axioms of probability: / long-run relative frequencies as motivation, addition // law. Conditional probability: definition, multiplication rule, independence, Bayes's theorem. Permutations and combinations and calculation of probabilities.

Random variables and Probability Distributions. Random / variables: definition, induced sample space, discrete // and continuous random variables. Probability distribution function, probability density function. Expectation: definition, expectation of a function of random variables, moments, variance, expected value of linear combination/ of a random variable. Important discrete distributions:/ Bernoulli, geometric, binomial, Poisson, hypergeometric. Important continuous distributions: uniform, gamma, exponential, normal.

Further distribution theory. Transformations. Bivariate/ and multivariate random variables: introduction in discrete case, marginal and conditional probability functions, expectation and variance of sum of two random variables, outline of extensions to continuous bivariate / and general multivariate cases. Generating functions: probability generating function, moment generating func- // tion, linear combination of normal random variables. Poisson process: definition, distribution of time between / events and time to first event, distribution of number / of events in unit time.

Statistical inference. General aim of inference. Random / sample. Distribution of $\bar{X}, \frac{1}{n} \sum (X_i - \bar{X})^2$ from normal populations: chi-square distribution. Distribution of \bar{X} from non-normal populations: central limit theorem, normal / approximation to binomial.

Estimation. Definition of estimator: statistic, estimator, sampling distribution. Unbiased estimation: minimum

* Reprinted with authorization of Department of Probability and Statistics, The University, Sheffield, U.K.

variance unbiased estimation (simple examples), relative efficiency. Methods of obtaining estimators: least squares, maximum likelihood. Fitting distributions to data. Interval estimation: confidence intervals and their interpretation, t-distribution.

Hypothesis testing. Framework of test: definitions of // simple and composite hypotheses test, acceptance and rejection regions, type I and type II errors, power. Tests on mean of normal distribution: practical considerations as motivation for forms of tests, tests with σ^2 known / and σ^2 unknown. Tests on variances of normal distribution. Interpretation of test results: significance level. Theoretical background: Neyman-Pearson lemma with / simple examples, most powerful test. Connection between/ hypothesis tests and confidence intervals.

Two sample test and extensions. Two normal populations:// t-test for means, F-test for variances, paired 2 sample/ t-test. One way analysis of variance.

Regression and correlation. Linear regression. Correlation: test of $\rho = 0$ in normal populations. Fisher z-transformation, connection with linear regression. Extensions of simple linear regression: brief discussion of multiple regression, polynomial regression and general linear model.

Inferences with discrete data and chi-square goodness of fit tests. Single sample tests for binomial and poisson. Two sample tests: chi-square test in 2×2 tables. Multinomial distribution. The chi-square goodness of fit // test: contingency tables (test of association and homogeneity).

Nonparametric test. General concepts. Sign test: single/ sample, paired samples. Rank sum tests: Wilcoxon signed/ rank test, Wilcoxon rank sum test.

Bayes methods and decision theory. Definition of probability as degree of belief. Prior and posterior probabilities and use of Bayes's theorem. Brief discussion of // approaches to point and interval estimation and hypothesis testing. Decision theory: loss function, risk function, Bayes approach via Bayes risk.

COURSE MATERIAL

HOEL P.G.: Introduction of Mathematical Statistics. 4th/ Ed. Wiley.

ADDITIONAL REFERENCES

Huntsberger D.V. & Billingeley P.: Elements of Statistical Inference. 4th Ed. Allyn and Bacon.

Mendenhall H.: Introduction of Probability and Statistics. 4th Ed. Buxbury.

Wetherill G.: Elementary Statistical Methods. Methuen.

TABLES

Neave M.R.: Statistical Tables. George Allan and Unwin.

APPENDIX V*

FIRST YEAR STATISTICS EXAMPLES

1.2. Comment briefly on the following (just slightly modified) published reports:

- a) From a catering journal: "In a recent survey on/ the use of works canteens in different sized companies it was found that of those workers who // had access to canteen facilities, 43% used a // small canteen, 38% used a medium-sized canteen / and 35% used a large canteen".
- b) "In 1971 there were 7443 road deaths. The Ministry of Transport reports that 5926 of these // deaths occurred during daylight hours and only / 1517 occurred at night. The Government is considering a campaign to persuade people to drive at night".

5. Suppose there were a diagnostic test for cancer // with the properties that $P(A/C) = 0.95$, $P(\bar{A}/\bar{C}) = // = 0.95$, in which C denotes the event that a person/ tested has cancer and A denotes the event that the/ test states that the person tested has cancer. The/ test is applied to people chosen at random from a / population where $p(C) = 0.0005$. If the test is positive, what is the probability that the individual / actually has cancer? Discuss the implication of the result for the introduction of a screening test for cancer in the population.

Suppose instead that $p(A/C) = p(\bar{A}/\bar{C}) = r$. What value must r take to ensure that $p(C/A) = 0.95$?

* Reprinted with authorization of Department of Probability and Statistics, The University, Sheffield, U.K.

7.3. Mice are injected with organisms, each mouse being/given a dose consisting of equal proportions of two cultures A, B. It may be assumed that the numbers / effective organisms per dose in each of the two cultures vary independently in Poisson distributions / with mean 2. The mouse will survive if and only if/ there are no effective organisms in the dose. Dead/mice are examined to find whether they contain orga/nisms of one or both cultures. Prove that the probability that an examined mouse contains organisms of only one culture is 0.24..

10.3. A sample x_1, x_2, \dots, x_n is drawn from the uniform // distribution on the interval $(0, \beta)$. Show that \bar{x} is an unbiased estimator of β . Show that the cumulative distribution function of $y = \max(x_i)$ is $(y/\beta)^n$ / for $0 \leq y \leq \beta$. Hence show that $(1 + 1/n)y$ is also/ an unbiased estimator of β . Evaluate the efficiency of \bar{x} with respect to $(1 + 1/n)y$.

16.1. Two types of petrol (A, B) were compared for 'mi-// leage' on the 1979 model of a certain make of car./ Of 15 such cars, chosen at random, 7 were given a / standard mileage test using petrol A and the other 8 were given the same test with B. The mileage (mi- / les/gallon) for the 2 groups were:

A: 22.3 23.1 20.4 21.4 22.9 22.6 20.9
B: 28.7 20.3 22.6 24.5 22.6 23.8 21.9 23.1

(i) Do the mileages appear to vary more widely for/ one type of petrol than the other. Calculate a/ 95% confidence interval for σ_A^2/σ_B^2 : the ra-/ tio of the variances for the two types of pe-/ trol.

(ii) Is there significant evidence that the mean mi- / leage is different for the two types of petrol/ on the assumption that $\sigma_B^2 = 2\sigma_A^2$?

APPENDIX VI *

IS STATISTICS

ASSESSED PRACTICAL

On the attached sheet, the accumulated results for / the simulation of the doctor's surgery from Practical 5 are given.

As a statistical consultant, you have been given the

* Reprinted with authorization of Department of Probability and Statistics, The University, Sheffield, U.K.

se results and asked what recommendations you would make. Analyse the data and write a report on your findings. The report should include sufficient detail on the basic problem, on the nature of the data, on the methods and re- / sults of analyses and on your findings, for the reader to understand what has gone on without further information./ There is of course a practical aim in the exercise: to ad- vise the doctor on how he should run his practice. As well as producing your technical report, you should prepare/ a letter to the doctor informing him of your findings and recommendations in terms he is likely to understand (avo- ding technical jargon).

The report and letter should be handed in before the/ last day of term.

IS STATISTICS

PRACTICAL 5

Simulation

A doctor is considering whether or not it is worthw-/ hile introducing an appointments system to his evening // surgery. The surgery is scheduled to operate from 5.30 to 7.00 p.m. and the doctor arrives at 5.30 p.m. prompt.

At present the surgery doors are open from 5.20 p.m./ and patients arrive according to a Poisson process, such/ that the mean time between the arrival of each patient is 5 minutes. At 6.50 p.m. the surgery doors are shut and no more patients are allowed into the doctor's waiting room, but all patients who arrive before this time are seen by/ the doctor. Each consultation takes an average of five mi- nutes, the distribution being Gamma with parameters(5,1). The doctor sees the patients in the order of their arri-/ val.

Under the proposed appointments systems, 18 appoint-/ ments would be made each evening at 5 minute intervals / from 5.30 to 6.55 p.m. A patient's arrival time is assu-/ med to be distributed uniformly from 7 minutes earrly to/ 3 minutes late, relative to his appointment time. The doc- tor would then see patients in order of appointment time/ except where this would mean an unnecessarily idle pe- // riod. We would still assume distribution of the consulta- tion time would not change.

Using the random digits and the exponential random va- riables (with mean 1) provided, simulate a complemente // evening surgery for each of the systems, working to an // accuracy of 0.1 minutes. Illustrate your results graphica

lly and record the following observations for each system.

- (i) The number of patients seen by the doctor.
- (ii) The time at which the surgery terminates.
- (iii) The total time for which the doctor is idle.
- (iv) The longest wait for any patient before he sees // the doctor.
- (v) The arrival time of the patient in (iv).

(vi) The number of chairs needed in the waiting room.

Compare the two systems and on the basis of your results make suitable recommendations to the doctor. Assume that the doctor's time is more valuable than that of his/ patients. Also comment on the reality of any assumptions/ and distributions used in the above models.

N.B. Choose your starting point for both the random / digits and the exponential random variables at random and then proceed row-wise.

IS STATISTICS

DATA for PRACTICAL 5

(Groups A & B)

WITHOUT APPOINTMENT						WITH APPOINTMENT					
<u>Number of Patients</u>	<u>End Surgery</u>	<u>Doctor Idle</u>	<u>Longest Wait</u>	<u>When?</u>	<u>Chairs</u>	<u>Number of Patients</u>	<u>End Surgery</u>	<u>Doctor Idle</u>	<u>Longest Wait</u>	<u>When?</u>	<u>Chairs</u>
17	92.5	2.3	16.5	48.9	4	18	97.9	0.0	12.1	53.1	3
16	89.2	14.4	9.7	10.0	4	18	110.9	7.0	24.7	75.6	4
12	77.7	32.2	6.1	-6.1	2	18	98.9	21.2	11.1	52.2	3
20	94.2	0.0	16.2	53.1	5	18	108.6	2.1	19.5	68.1	5
14	80.1	9.9	21.5	0.7	4	18	91.3	5.0	16.5	35.2	4
16	101.6	11.7	30.4	53.7	6	18	96.2	0.0	17.3	8.0	4
15	88.8	13.8	11.7	41.2	3	18	86.8	2.3	15.2	33.4	4
23	109.8	0.0	30.7	74.1	6	18	104.3	7.8	18.5	81.2	4
18	108.6	0.0	36.5	72.1	5	18	99.4	8.8	15.3	49.5	4
23	106.5	3.4	34.1	54.4	5	18	95.7	10.5	15.4	78.8	4
22	116.9	0.0	46.6	54.6	10	18	93.3	16.3	10.1	65.9	2
19	104.5	0.0	29.6	64.9	5	18	105.2	2.4	19.9	78.7	3
21	105.7	14.8	35.0	67.5	6	18	98.5	3.9	16.2	49.3	3
15	89.8	23.3	9.8	39.7	3	18	93.7	16.2	11.3	34.5	3
17	91.3	14.7	14.5	2.1	3	18	99.9	6.5	15.1	39.4	3
18	108.6	0.4	29.7	75.9	6	18	104.8	0.1	21.1	69.8	3
18	98.3	14.9	15.6	77.9	3	18	99.6	8.7	16.5	71.6	4
15	79.6	24.2	9.5	37.6	2	18	100.9	7.4	11.6	34.7	4
16	88.5	19.9	15.9	39.5	6	18	88.4	4.3	14.2	45.3	3
17	89.7	6.8	15.2	74.5	4	18	99.4	9.4	12.9	65.4	3
19	96.9	12.7	17.0	37.5	4	18	99.3	4.1	14.2	78.1	3
22	122.2	0.0	50.6	69.6	9	18	106.6	0.0	15.4	55.1	3
18	104.7	18.7	27.0	76.2	5	18	105.1	0.0	20.1	63.3	4
18	103.4	0.7	25.4	71.6	5	18	95.8	12.3	12.9	62.8	3
31	151.0	0.0	71.5	79.5	16						
20	121.9	16.7	47.0	69.1	5						
30	158.5	0.0	76.3	68.6	14						

APPENDIX VII **

PRACTICAL AND APPLIED STATISTICS

This will be a continuous assessment course, consisting of a series of 12 projects spread over 2 terms, each one being a practical problem with real data (perhaps reduced in size sometimes, to avoid unnecessary tediousness in coding and data handling). The overall aim of the course will be to allow students to take a problem through the stages of: objective definition, preliminary examination of data, modelling, analysis, computation and interpretation and communication of results. There will be a deadline for the handing in of each project and late projects will be severely downgraded. This is necessary to ensure the necessary feedback by which the students should (with reasonable application) be able to make a worthwhile attempt at them without devoting more of his time to this course than to any other. In more detail, the course will aim to:

1. Encourage and teach the use of graphical and other simple methods of presentation of data, for preliminary examination and analysis; the use and effect of transformations of variables.
2. Introduce to the student the techniques of report writing and the communication to non-statistical results; (the student will give a verbal and visual presentation of one project).
3. Provide an opportunity for the student to look at and find information from statistical sources (e.g. Social Trends).
4. Introduce the student to the critical reading of journal articles, extracting the essential points and summarising them (e.g. four articles on 'statistics and football' -from Moroney's chapter on 'goals, horse-kicks and floods' to Reep and Benjamin's paper in JRSS).
5. Train the student to 'think statistically'; to assess 'pseudo-statistical' statements and to study implied assumptions (e.g. 'an individual is five times more likely to die of a heart attack during or immediately after violent exercise than if he were sitting in an armchair').

6. Allow the student to use a variety of calculating and computing facilities (e.g. hand calculators, multiple linear regression program on the Apple, remote terminals to the 1906S). The first 6 projects will be given out and expected to be completed at a rate of one per week. These will be relatively straightforward in terms of the statistical techniques required (no new techniques introduced -using only those covered in IS and 2S). Later projects will be given out at a rate of 1 every two weeks* and will, in general, be rather more open-ended and the 'appropriate' statistical technique to be used will not necessarily be immediately obvious (in fact, there may well be alternative lines of approach). Each student will be required to give a verbal and visual presentation of one of these projects, and, to allow sufficient time for the preparation of this, 4 weeks rather than 2, will be allocated. The total number of projects will, therefore, be 6 (1 week each) + 5 (2 weeks each) + 1 (4 weeks) = 12. For the presentation project the class will be grouped into pairs or triples and the different groups will work on different projects. Each student will be expected to submit a personal written account of his group's project as well as take part in its oral presentation. In the assessment rather more weight (about 3/4) will be attached to the written account as compared with the oral presentation.

There will inevitably be some computing content, since the projects are to involve practical problems with real data, but it is hoped that this will be kept to a fairly minimal level, because (1) the writing and debugging to programs can be a time-consuming process, particularly for the student who is less able and interested in computing, and this will mean that the computing aspects can dominate the statistical aspects and (2) there will be the Statistical Modelling and Computing option available to those interested in that.

For the assessment of performance in the course as a whole the individual projects will be weighted in proportion to their lengths.

* At Christmas one project will extend over the vacation, but occupy in total two weeks of term. In the Spring term another project will be based on an experiment which will last for several weeks.

ENSEÑANZA DE LA ESTADÍSTICA

1.- INTRODUCCIÓN

Voy a considerar las cuestiones básicas siguientes:

- ¿Por qué debemos enseñar Estadística?
- ¿A quién debemos enseñar Estadística?
- ¿Qué aspectos de los métodos y principios estadísticos debemos poner de relieve?
- ¿Cómo debemos motivar al estudiante?
- ¿Cómo exponerle las ideas estadísticas?
- ¿Qué ejemplos darle?
- Programas y esquemas de trabajo.
- Ejemplos detallados para explicar las condiciones/ generales.
- ¿Cómo afectan al profesional y al ciudadano medio/ los aspectos más generales de la educación estadística?

Al intentar presentar el amplio espectro de opiniones sobre estas cuestiones es inevitable que me concentre principalmente en mi propia actitud con respecto a / ellas. He escrito mucho sobre este tema y haré uso ampliamente del material publicado por mi mismo y otros autores, indicando su procedencia.

Los temas que intento abarcar en las diferentes secciones serán los siguientes:

Sección 1.- Consideraciones generales acerca del principio y del fin de la educación estadística y su necesidad. La Estadística como herramienta para el estudio de situaciones reales. La demanda por parte de la sociedad/ de estadísticos cualificados, de profesionales competentes y de ciudadanos cultos. La enseñanza de la Estadística a nivel secundario: Problemas especiales a este nivel. Comprensión estadística de los 11 a los 16 años. Desarrollo del plan de estudios. Diferencias nacionales. / Ejemplos de la situación en el Reino Unido. La Estadística de los 16 a los 18 años. Necesidades diferentes para/ el conocimiento general, la práctica, el concepto y los métodos de la Estadística.

Sección 2.- Educación estadística a nivel terciario: Estudiantes de Matemáticas, de Estadística y de otras disciplinas. Grado en Estadística. Ejemplos de programas y material educativo. Cursos de aplicaciones de la Estadística. Maestría.

Sección 3.- Educación estadística para la práctica estadística. Consultoría estadística: Interacción entre el estadístico y el investigador. Problemas típicos de consultoría. El papel asesor y de servicio de un departamento universitario de Estadística.

1.1.- COMENTARIOS GENERALES

De cuando yo era un joven profesor universitario de Estadística recuerdo la visita de un eminentísimo probabilista ruso. No solo famoso por sus investigaciones, sino también por su habilidad para enseñar Probabilidad y / Estadística a estudiantes de cualquier edad entre 9 y 90/ años de edad y procedentes de diferentes niveles de formación. Le pregunté cómo podía tener él tanto éxito como // profesor. Me contestó de una forma bastante enigmática // que no había dificultad en enseñar Estadística (o quizás otra materia). Todo lo que se tenía que hacer era / proponer juegos atractivos adecuados al grado de desarrollo y conocimiento especial de los estudiantes y donde el estudiante tuviese posibilidad de ganar el juego. Había / en esta respuesta muchas cosas que son cruciales desde mi punto de vista: interés, reto, éxito, aplicabilidad de la Estadística.

Hoy día no se puede dejar de lado la Estadística. Es/ parte de la sociedad en que vivimos. La radio, la TV., los periódicos, plantean constantemente al ciudadano medio // cuestiones estadísticas a las que él debe ser capaz de // responder. Todo escolar necesita una orientación acerca / de las ideas informales del razonamiento estadístico dentro de su experiencia social cotidiana. Necesita estar // "alfabetizado estadísticamente", (es decir, saber interpretar tablas y hacer inferencias rudimentarias, no saber teoría de conjuntos ni espacios muestrales). En otro nivel, cualquier persona que esté interesada en temas cuantitativos (sociólogo, ingeniero, médico) necesita algún conocimiento y experiencia en los métodos estadísticos. Por último, es necesario el estadístico profesional: como educador, innovador y capaz de resolver problemas estadísticos difíciles. Esta tricotomía de

Ciudadano culto

Analista práctico de datos estadísticos

Estadístico profesional

formará el proceso de la educación estadística. En los últimos 30 años, se han dado grandes pasos desde el momento en que una pequeña minoría estudiaba una pequeña do/ sis de Estadística Matemática en la Universidad a una situación en que en algunas universidades (pero no en todos los países) la Estadística comienza a dar importancia a / estos tres aspectos a lo largo de todo el espectro educativo.

Más adelante, consideraremos con algún detalle la necesidad de los dos últimos grupos: analistas de datos estadísticos, y estadísticos profesionales. Aunque esto atañe

especialmente al sector universitario, se podría dar algún método estadístico en los niveles escolares de los 16 a los 18 años.

La alfabetización estadística a nivel secundario va a ser cada vez más esencial. Pero no es sólo el escolar / quién necesita ser educado: la sociedad como un todo necesita ser educada en la necesidad de tal proceso. Los ministros, los funcionarios, las sociedades profesionales deben interpretar su propio papel y están comenzando lentamente a hacerlo así. El Instituto Internacional de Estadística tiene dos grupos de trabajo uno a nivel secundario, (TOTALSAS, del que soy presidente) y otro a nivel terciario (presidente: Prof. R.M. LOYNES, también de la Universidad de Sheffield). El reciente informe TOTALSAS [BARNETT, (1)] expone claramente el progreso o falta de progreso en diferentes países. El reciente "First International Conference On Teaching Statistics" (ICOTS I) / del ISI tuvo mucho éxito y fue muy informativo. Veáñse / las Actas del Congreso [GREY et al (2)]. El ICOTS II se celebrará en Agosto 1986 en Victoria, BC, Canadá. En Sheffield, hemos dirigido programas de desarrollo del plan / de estudios principal para edades entre 11-15 y 16-18 // años. Pero no debemos subestimar el coste de tiempo y dinero para desarrollar, ensayar, validar y publicar dicho material. Esto ha ocupado aproximadamente 25 años-hombre a un coste de unos 65.000.000 de pesetas.

La revista "Teaching Statistics" y el nuevo "Center for Statistical Education" en el Reino Unido son también iniciativas muy bien acogidas.

1.2.- ENSEÑANZA DE LA ESTADÍSTICA A NIVEL SECUNDARIO

El tiempo de que disponemos no nos permite una exposición detallada de esta cuestión, ni siquiera a nivel nacional. El informe TOTALSAS [BARNETT (1)], lo explica / detalladamente. Puede también encontrarse una extensa // descripción de la situación en el Reino Unido en el Capítulo 1 del informe TOTALSAS, en HOMES et al. [(4), (5)] y en BARNETT et al. (3). Citaremos aquí unas cuantas cuestiones como ejemplo. La Estadística tiene protagonismo / en los programas escolares del Reino Unido desde hace // unos 25 años: es parte esencial en los exámenes del Ordinary Level (Nivel 0; 16 años) y en el Advanced Level // (Nivel A; 18 años). En el pasado hubo tendencias a incluir demasiado material y a presentarlo de una manera / matemática demasiado formal. Los libros de texto de apoyo no eran muy buenos y los profesores no estaban bien preparados para enseñar Estadística. Aunque se observan progresos en todos los frentes, existe el problema de la // formación de los profesores. Ahora, en la enseñanza y en

los programas del Nivel 0 y del Nivel A se está aumentando el interés por la parte práctica: en algunos casos se exigen proyectos verdaderamente prácticos.

Podemos comprobarlo al leer algunos de los temas de / los proyectos prácticos y de las cuestiones propuestas recientemente en el Nivel A [BARNETT (1)]. En el Apéndice / aparecen dos cuestiones del examen de Estadística del Nivel A propuesto por el Joint Matriculation Board. En el // Apéndice II aparece una selección de temas de proyectos / prácticos para el programa de Estadística del Nivel A propuestos por el Joint Matriculation Board. Comparemos éstas con la siguiente cuestión, que es una versión ligeramente modificada de una que propuso otra Comisión a Nivel A.

Una variable aleatoria X tiene como función de densidad de probabilidad

$$f(x) = a + bx + cx^2 \quad (0 < x < 2)$$

La media y la desviación típica de X son $8/5$ y un $1/5$ respectivamente. Determinar a , b y c .

Salta a la vista que tenemos que resolver tres ecuaciones

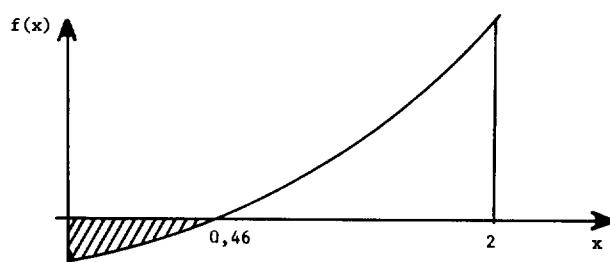
$$2a + 2b + \frac{8c}{3} = 1$$

$$2a + \frac{8b}{3} + 4c = 8/5$$

$$\frac{8a}{3} + 4b + \frac{32c}{5} = 13/5$$

$$a = 3/20 \quad b = 3/20 \quad c = 15/40$$

Pero, ¿qué clase de cuestión estadística es ésta? No contiene ni conceptos estadísticos ni probabilísticos. // ¿En qué circunstancias tal $f(x)$, o tal conocimiento, nos proporcionaría un modelo real?. Es sencillamente un ejercicio matemático. Ahora, supongamos que hay que dibujar / $f(x)$ (¡como evidentemente no hicieron los examinadores!). ¡Es negativa a lo largo de aproximadamente el primer 25% de todo su dominio de definición!



Pero, ¿qué ocurre con los más jóvenes (de 11 a 16 // años) a los cuales queremos introducir en las nociones / básicas del manejo de datos y del razonamiento estadístico?. Esto no es fácil. La imagen popular de la Estadística como un medio de mentir o de "demostrar todo" está en contra nuestra, como lo están los errores constantes de interpretación de la media.

Además, hasta hace poco, no se había publicado material de apoyo para el enfoque de tipo abierto, orientado al proyecto práctico, necesario para esta etapa. En 1975, tuve la suerte de persuadir a una de las agencias gubernamentales (el Schools Council) para que facilitase los fondos necesarios para el desarrollo de un proyecto de / plan de estudios para los estudiantes de 11 a 16 años. / El objeto era editar unas 30 unidades de material didáctico, cubriendo cada una aproximadamente 5 horas de clase, y correspondiendo unas 7 unidades a cada grupo de / edad. Ciencias Sociales, Ciencias, Humanidades, Matemáticas, etc.) y una amplia gama de ideas probabilísticas y estadísticas. Los conceptos y métodos estadísticos están "sumergidos" en los ejemplos de modo que se aprendan implicitamente mejor que directamente en el sentido de

una distribución de probabilidad es ...

o
la media muestral se define como...

Así los problemas guían a los alumnos al autodescubrimiento de los conceptos. Cada unidad consta de un conjunto de notas para el alumno y de otro de notas para el profesor con el propósito tanto de informar a los profesores (en extenso) como de instruir a los alumnos. Las // unidades fueron preparadas, impresas, ensayadas en unos/ 60 centros educativos, validadas, escritas de nuevo, etc. El material completo del proyecto está ahora publicado / en:

HOLMES et al (1980): *The structure of project // approach.*

HOLMES et al (1981): *A review of the UK situation in schools y*

SCHOOLS COUNCIL PROJECT ON STATISTICAL EDUCATION // (1980-81): *The units.*

Esta breve exposición de la enseñanza de la Estadística a nivel secundario acaba con el Apéndice III en el/ que aparecen fragmentos de algunas de las unidades preparadas por el Schools Council Project on Statistical Education (11-16).

2.- LA ENSEÑANZA DE LA ESTADÍSTICA EN LAS UNIVERSIDADES

La situación en las Universidades es intrínsecamente/ más clara que en el nivel secundario en muchos aspectos:

Los estudiantes pueden ser comprometidos y motivados, generalmente eligen estudiar Estadística (aunque los cursos de aplicaciones de la Estadística para especialistas de otras ramas, por ejemplo, Biólogos, Ingenieros, etc., pueden no ser tan adecuados para los que / asistan a lo mismo).

Los profesores pueden ser formados profesionalmente / en Estadística aunque esto no garantiza habilidad, o mucha experiencia práctica.

El material está a menudo bien estructurado y documentado; se sabe lo que se quiere enseñar y se dispone / (generalmente) de buenos libros de texto como ayuda / aunque, personalmente no soy partidario de utilizar / un texto único y además algunos de los temas prácticos y aplicados no disponen todavía de muchas publicaciones buenas.

Necesitamos distinguir diferentes tipos de cursos y / diferentes grupos de estudiantes. A nivel de pregrado/ ción tenemos:

i) Cursos de aplicaciones de la Estadística (cortos)/ para estudiantes de Agricultura, Biología, Econo-/ mía, Ingeniería, Medicina, Ciencias Sociales, etc.: llamados "Cursos de Aplicaciones".

ii) Cursos de Estadística y probabilidad para estudiantes de Matemáticas.

iii) Cursos de Estadística y Probabilidad para estudiantes de Estadística (especialistas).

A nivel de postgrado se imparten Cursos para el Grado de Maestría y cursos para estudiantes dedicados a la investigación.

Los últimos 30 años han visto una gran proliferación/ de estadísticos en las Universidades del Reino Unido -de/ una situación en que los cursos individuales de Estadística Elemental eran dados ocasionalmente quizás apologeticamente como leve ayuda en cursos para graduados en Matemáticas, a la actual presentación comprehensiva de la Estadística como una disciplina por derecho propio o como / método experimental y base del análisis de materias tales como Psicología, Geología, Economía, etc. La mayoría de /

las universidades tienen departamentos de Estadística (o son un grupo importante en los Departamentos de Matemáticas), la mayoría ofrecen grados y estudios de postgrado/ en Estadística u ocupan un lugar importante en los grados de Matemáticas. Los cursos de aplicaciones a la Psicología, etc, pueden ser impartidos por especialistas // del Departamento de Estadística (de quienes se espera un gran interés y compromiso con respecto a la materia) o / por especialistas en la materia (de quienes se espera // gran interés y mucha experiencia en Estadística).

Por lo general, las universidades no se atienen sustancialmente a los expedientes escolares de los estudiantes. Los cursos de Estadística "parten de los comienzos", así como los cursos de aplicaciones: bien por creer que no puede asumirse un conocimiento común de principios básicos ó a causa del descontento con la enseñanza secundaria, unido a programas de exámenes poco adecuados. Pero, / ¿son la enseñanza secundaria y el Tribunal de Examen totalmente irrelevantes, o no afectan, a la enseñanza a // nivel universitario? Creo que la enseñanza secundaria // tiene poca influencia en las universidades, pero lo contrario no es verdad. Una actitud como "más listo que tú" fué fomentada al principio (aunque todavía persiste hasta cierto punto) por el esnobismo matemático que ha influido en la enseñanza y en la preparación del programa/ a nivel no universitario. Antes, la postura de la universidad era en cierto modo, la de censurar las características irrealistas, superficiales y demasiado formales de // los programas de estudio de la mayor parte de la enseñanza secundaria. A pesar de algunos intentos honestos individuales para tender un puente entre la Estadística a nivel universitario, a nivel secundario y la Estadística / aplicada a la industria, a la investigación y al comercio, hay todavía mucha renuencia a mezclarlas. Los profesores universitarios deben hacer un esfuerzo para ayudar a la Estadística a nivel secundario y a nivel comercial/ e industrial. Mientras mantenga la actitud de que el trabajo no universitario es "académicamente inferior", "mal pensado" y "peor enseñado" y por lo tanto, irremisible,/ no hacen nada para remediar cualquier verdad que hubiese en tales creencias.

En BARNETT et al (1) se hace una revisión de la situación actual en el Reino Unido. Algunos de sus párrafos son muy importantes desde un punto de vista general.

"El crecimiento de la Estadística en el sector de la enseñanza superior ha sido dramático. Por ejemplo, fué / sólo en 1949 cuando se cubrió la primera cátedra de Estadística aparte de las de Oxbridge y Londres... En menos/

de 30 años hemos llegado a unas 70 cátedras de Estadística en las Universidades Británicas y a más de 700 profesores dedicados específicamente a enseñar Estadística en // las Universidades y Politécnicos Británicos (la mayoría / en las universidades). La Estadística se enseña a todos / los niveles, desde el Certificado del Politécnico (o HNC) a los cursos de postgrado de Maestría, se enseña como especialidad o como material auxiliar en una amplia gama de ciencias aplicadas... La forma de enseñar, en su nivel y extensión, varía mucho de una institución a otro..."

"El estilo de los cursos varía también mucho desde // una presentación formal, altamente matemática para el cálculo de Probabilidades, Procesos Estocásticos, e Inferencia Estadística, a una informal "tipo libro de cocina" del método estadístico. Sin embargo, hay signos de mejora del equilibrio con atención creciente a la motivación y explicación práctica de problemas de la vida real en diversas/ áreas. La Estadística verdaderamente práctica es de vital importancia, pero no es fácil de enseñar; la mayoría de / las universidades parecen dispuestas a hacerlo ahora, con algunos resultados gratificantes. La tendencia actual respecto al análisis de datos exige un enfoque orientado a / los problemas.

"Cualquier esperanza de desarrollar la conciencia estadística en los estudiantes a través de la enseñanza desencansa en la demostración significativa de que la mayoría/ de los problemas de importancia son estadísticos. Eso debe hacerse durante un largo periodo con una gran cantidad reforzamiento práctico. Particularmente a nivel secundario, la Estadística tiene que verse como un hilo a través de todos los aspectos del esfuerzo humano, no como un componente seco de la lección de Matemáticas, es parte de la lección de Geografía; incluso de F.F. en Metalistería... Tales materias no se enseñan fácilmente, requieren una exposición detallada de los problemas reales más que unas/ cuantas explicaciones someras en clase. Dependen de la experiencia del profesor y de la técnica de enseñanza que / es pausada y basada en problemas en vez de en el sistema/ matemático habitual de colocar un pequeño ladrillo encima de otro hasta que la casa esté construida del todo..."

..."y si no es un enfoque de ladrillo sobre ladrillo, entonces, ¿cuál es el método de instrucción a seguir? Es esencialmente como sigue. Primero, miramos el mundo de la / Estadística desde la ventana con ayuda de los prismáticos y finalmente a través del microscopio. Si la enseñanza secundaria es como contemplar el paisaje desde la ventana y lo hace bien, la Universidad y el Politécnico pueden entonces estar en mejor posición para emplear los prismáticos

cos y el microscopio".

No voy a pasar el tiempo discutiendo lo *inadecuado* de la situación presente. En lugar de eso, expondré mis/ puntos de vista sobre los aspectos de lo que considero / que debe ser el plan de estudios, el programa del mismo/ y su contenido detallado -con ejemplos para explicar mis puntos de vista- basados en mi experiencia en Sheffield/ y en otros lugares.

Voy a comenzar con una exposición comprehensiva del/ plan de estudios para no graduados.

2.1.- GRADUACION EN ESTADISTICA

En el modelo inglés, los estudiantes entran en la Universidad a los 18 años generalmente para estudiar una (o dos) materias en profundidad durante un periodo de // tres años. Algunos estudiantes desean estudiar Matemáticas con especialización en Estadística, o específicamente Estadística, para obtener un grado en Estadística.

Hay muchos temas sobre Probabilidad y Estadística // que deben ser enseñados a lo largo de los tres años. Es crucial que aprendan no solamente las matemáticas de la/ materia -sino lo que es más importante que sean entrenados para trabajar como estadísticos profesionales: para resolver problemas reales. Para esto se necesita:

- a) un plan de estudios coordinado del concepto y de los métodos estadísticos que mantenga un nivel matemático apropiado.
- b) un esquema de trabajo práctico, comprehensivo y / progresivamente más sofisticado a lo largo de los tres años.

El nivel matemático necesario para a) es inevitablemente tan exigente respecto a los temas importantes como el de los estudiantes de Matemáticas per se.

Este es el que mejor dominan los profesores de Matemáticas en coordinación con los estadísticos con respecto a horas y profundización en los temas. Los estudiantes de Matemáticas y Estadística estudiarán juntos las / materias de interés común, lo que requiere una buena relación interdepartamental. El ideal sería una "estructura escolar" con Departamentos de Estadística, Matemática Pura, Matemática Aplicada y Computación, que tuvieran // status y control independientes, pero estrechamente vinculados en sus funciones docentes.

El que pueda llevarse a cabo una política de admisiones en común, con un programa común de estudios para el primer año es una gran ventaja que además de facilitar / la coordinación de las diferentes materias, pospone la / elección de especialización hasta el segundo año cuando/ ya se posee una mayor información.

El programa de trabajo práctico bajo b) debería co-/ rrer parejas (pero no seguir servilmente) con los componentes del concepto y de los métodos estadísticos en los primeros años. Sin embargo, en el tercer año el estudiante debe ser conducido a una etapa en la que desarrolle / una verdadera inclinación por la construcción de mode- / los, el recuento y análisis (tarea no fácil) de los da-/ tos y se le pueden plantear problemas reales más sustan- ciales, en los que no se indica una técnica directamente o ni siquiera pertinente. Aquí puede ayudarle el ambien- te de consultoría del departamento -Veáñse los comenta- rios sobre este esquema más adelante.

¿Cómo es un grado en Estadística de tres años? Pro-/ pongo el siguiente esquema:

Tiempo	Años		
	1	2	3
1/3	Probabilidad y Estadística	Probabilidad y	Probabilidad
1/3	Matemática Pura Matemática Aplicada	Estadística	y Estadística
1/3	y Métodos de cálculo	y Principios Matemáticos	Matemáticas

1º Año. Curso obligatorio para todos los estudiantes de Estadística y de Matemáticas (que garantiza algún contacto con la Estadística a todos los estudiantes de / Matemáticas).

Detalles del Curso De las aproximadamente 150 horas, 80 son de clases teóricas, 25 horas de aplicaciones y 45 horas de prácticas (en unidades lectivas de 2-3 horas).

Programa de Estudios: En el Apéndice IV aparece un / ejemplo típico de programa y una lista de libros de referencia para este tipo de primer curso (una versión reciente del curso de primer año Special Honours de estudiantes de Matemáticas de la Universidad de Sheffield).

Dicho programa podría, sin embargo, enseñarse de diversas maneras -desde la formal puramente matemática, hasta la aplicación memorística no rigurosa de "recetas".

Ninguno de estos extremos es razonable. Las clases / deben desarrollar conceptos y métodos a un nivel matemático sólido (según la capacidad media del estudiante), pero debe ponerse especial empeño en desarrollar la comprensión y conciencia estadística y la habilidad para aplicar las ideas a problemas reales.

En las clases de aplicaciones se discuten las dificultades de un conjunto de unos 5 ejercicios en los cuales los estudiantes han trabajado durante la semana. Los ejercicios deben ser tan reales y útiles como sea posible. Para evitar los problemas aburridos y faltos de interés (típicos, por desgracia, en muchos libros de introducción) tales como:

Q1 Supongamos que en 8 tiradas de un dado obtenemos/ 6, 5, 1, 1, 2, 6, 4, 2.

Calcular la media, la varianza y la desviación típica de la muestra.

Q2 Una variable continua X tiene una función de densidad de probabilidad

$$f(x) = a^2 - x^2 \text{ para } -\sqrt{a} < x < \sqrt{a}$$

Si la varianza de X es 0,5. Calcular a y $P(0 < x < 0,4)$.

En los dos casos yo estaría completamente de acuerdo con la pregunta ¿Para qué? ¿Qué tienen que ver con la Estadística? Son ejercicios matemáticos y numéricos faltos de interés. Los cinco ejemplos del Apéndice V del curso/ de primer año al que acabamos de referirnos pueden darnos una idea de como aproximarnos a la necesidad real, / para reforzarla, comprenderla y aplicarla.

En las clases prácticas se tiene la oportunidad de / estudiar con más detalle la construcción de modelos y el manejo e interpretación de datos. Hay también un Laboratorio de Estadística donde los estudiantes trabajan durante dos semanas en un proyecto y redactan un informe / con su opinión personal. El Apéndice VI muestra un ejercicio práctico de simulación de una sala de espera de un

médico. Se pide a los estudiantes que utilicen el conjunto completo de datos de la clase para analizarlo y escribir un informe técnico razonado y una carta dando recomendaciones al doctor.

2º Año. Aquí el estudiante de Estadística dedicará / la mitad de su tiempo a estudiar Probabilidad y Estadística (el de Matemáticas estudiará todavía algo de Estadística, pero mucho menos de la mitad de su tiempo). Es un // curso semejante al de primer año en estructura y tamaño/ (con clases de aplicaciones y prácticas) que comprende:

Teoría de distribuciones
Principios estadísticos
y Métodos estadísticos

Este es un curso clave en el plan de estudios de // tres años, que abarca la mayoría de los conceptos centrales y la metodología esencial que necesita el estadístico profesional -teoría de la distribución normal, estimación y dígitas, modelos lineales, regresión, asociación/ y algo de análisis secuencial, poblaciones finitas y métodos no paramétricos.

Lo que queda (un tercio) del segundo año (con dos // clases teóricas y una de aplicaciones por semana) se refiere a la Teoría de Probabilidad, a nivel intermedio, y a una Introducción a los Procesos Estocásticos.

3º Año. En el tercer año, el estudiante de Estadística dedica el 80% del tiempo a temas de Probabilidad y Estadística. Puede también seguir un curso sobre algún tema matemático como complemento o por interés. Aquí hay / mucho donde elegir. Ciertos temas se consideran como // centrales y se le anima al estudiante a escoger.

Estadística Aplicada y Práctica
Inferencia y Teoría de la Decisión
Series Cronológicas
Diseño de Experimentos
Análisis Multivariante

Esto representa el 50% del material estadístico. El/ resto se elige por interés personal y necesidades de la/ carrera entre las siguientes:

Probabilidad, Medida e Integral
Teoría Matemática de la Probabilidad
Estadística Médica
Métodos Estadísticos y Estocásticos en
la Industria.
Investigación Operativa

Curso Superior en Procesos Estocásticos
Construcción de Modelos Estadísticos.
Teoría del Muestreo
Métodos No paramétricos y Métodos Robustos.
Procesos Especiales
Análisis Categórico de Datos

En esta etapa, algunos cursos son muy técnicos o matemáticos, otros presentan una metodología detallada, // otros corresponden a formación profesional. El curso de Estadística Aplicada y Práctica es muy duro para los estudiantes.

El Apéndice VII describe los fines de este último // curso. También se incluyen aquí, experiencia en computación, utilizando calculadoras, microcomputadores, ordenadores, paquetes, etc. aunque no se llega a dominar el / componente estadístico. Para algunas referencias sobre / Estadística, Enseñanza y Computación ver: EVANS (2), // MEAD y STERN (3).

2.2.- ESTUDIANTES DE MATEMÁTICAS NO GRADUADOS

La cuestión del material y de lo que es apropiado/enseñar a los estudiantes de Matemáticas queda contestado hasta cierto punto por los comentarios acerca de lo / que es necesario para los estudiantes de Estadística. Todo matemático debería ser consciente de la necesidad de/ aplicar sus conocimientos matemáticos y debería conocer/ la envergadura de la matemática aplicable. Por lo tanto//, creo que todos deben tener un curso de introducción/ del tipo de los descritos anteriormente (el componente / del primer año de un grado en Estadística). Necesitan tener conceptos bien motivados y enraizados de las aplicaciones. Pero a partir del primer año los temas de Probabilidad y Estadística para estudiantes de Matemáticas // pueden elegirse a gusto de cada uno dentro del programa/ de Estadística. Muchos de los cursos en las áreas de Probabilidad y Procesos Estocásticos tendrán atractivo matemático (como el Diseño de Experimentos, Análisis Multivariante, Investigación Operativa, etc.) y servirán también para ensanchar el horizonte de los matemáticos y para aumentar su competencia profesional.

2.3.- CURSOS DE APLICACIONES DE LA ESTADÍSTICA

Leemos en BARNETT et al(1)" que muchos estudios // universitarios en materias aplicadas requieren ahora cursos de aplicaciones de la Estadística, generalmente para los estudiantes de primer año. Entre estos se incluyen / Agricultura, Economía, Ciencias Sociales, Medicina, Psicología, Bibliotecas, Ingeniería, Biología, Geografía, /

Arqueología, etc. Es importante que los especialistas de estas materias tengan algún conocimiento acerca del razonamiento y de la técnica estadística. Sin embargo, uno // se pregunta algunas veces si el típico curso de 10 lecciones de primer año, que incluye Cadenas de Markov y // Análisis por Conglomerados, cumple este objetivo en términos de pertinencia, campo de aplicación y urgencia. // Existe la necesidad de evaluar adecuadamente el estilo, / tiempo y contenido de los cursos de aplicaciones. La mayoría de tales cursos los dictan miembros de los Grupos o Departamentos de Estadística, y si se desea fundamentar/ con firmeza es necesario que el profesor se sumerja por/ completo en temas de la disciplina básica para hacer que la Estadística aparezca importante e interesante. Existen también demandas sobre la metodología -el estudiante de Biología no es un matemático frustrado!-. Algunos/ cursos de aplicaciones están protegidos por la disciplina aplicada -Psicología parece ser el ejemplo extremo / de la actitud "nosotros conocemos lo que nuestros estudiantes necesitan". Tal actitud puede ser de algún mérito, si realmente ocurre que es verdad y si existe un conocimiento bastante verdadero de la Estadística, para // una buena enseñanza. El peligro está en una elección arbitraria y a la moda de los temas estadísticos y en la / incomunicación tecnológica".

El factor más importante es que el estudiante debe / encontrar interesante e importante el material estadístico, mejor que considerar que se le pide asistir a un conjunto de clases obtusas sin contacto con su asignatura / y sin ninguna justificación aceptable. El material necesita ser desarrollado despacio, con gran cantidad de // ejemplos prácticos procedentes del área adecuada de la / asignatura. Esto significa que aún el más breve campo de aplicación exige un mínimo de unas 20 horas (p.e 10 clases teóricas y 10 clases prácticas). Las clases prácticas deben estar bien dotadas de personal de modo que todos los estudiantes pueden exponer sus dudas (quizás 1 / supervisor por cada 12-15 estudiantes). El departamento/ involucrado debería poner de relieve la importancia de / los cursos de Estadística para sus estudiantes.

Si el desarrollo estadístico es firme es tentador // considerar que la aplicabilidad práctica no es crucial./ Dede el punto de vista psicológico es un error. Considerese el siguiente ejercicio de un curso breve de aplicaciones para Ingenieros Civiles.

Q1. Sea p la probabilidad de que en un año cualquiera la más alta de las riadas de dicho año sobrepase el malecón del Támesis (Thames Embankment) y origine inundaciones en el centro de Londres. Si la presa de con-/

tención del Támesis (Thames Barrage), que evitará dicho peligro no se ha terminado en los 4 años próximos, y // las magnitudes de las más altas riadas en cada uno de estos 4 años son independientes, hallar, en función de p , las probabilidades de que:

- i) no haya ninguna inundación por riada antes de la terminación de la presa,
- ii) haya inundación por riada por lo menos una vez antes de la terminación de la presa.
- iii) haya inundación por riada en 2 de los 4 años próximos.

Calcular estas probabilidades si (a) $p = 0.05$, (b) $p = 0.10$, discutiendo si tales valores de p son realistas en este contexto.

Esto está expresado en términos que son importantes/ para el estudiante de Ingeniería Civil. El siguiente // ejercicio es en esencia el mismo estadísticaicamente.

Q2. Sea p la probabilidad de que un volumen-patrón / de un cultivo preparado en laboratorio contenga contaminación externa (desarrollo de bacterias). Si se preparan cuatro muestras independientes de cultivo, hallar, en / función de p , las probabilidades de que

- i) no haya muestras contaminadas
- ii) haya por lo menos una muestra contaminada
- iii) haya exactamente 2 muestras contaminadas.

Calcular estas probabilidades si (a) $p = 0.01$, (b) $p = 0.5$, discutiendo si estos valores de p están de // acuerdo con la realidad de este contexto.

El ejercicio 2 sería adecuado para un curso de aplicaciones para estudiantes de Biología; los ingenieros civiles obtendrán muy poca motivación y beneficio del mismo. Observense que aún la especificación de los valores de p exigen atención. Sea $p = 0.85$ en el ejercicio 1 y los ingenieros civiles perderán confianza en el cuidado o conocimiento en la elección de material de su profesor. Ahora contrastemos los ejercicios 1 y 2 con el siguiente

Q3. Sea p la probabilidad de que una bola extraída / al azar de una urna, que contiene un gran número de bolas, sea negra. Si se extraen independientemente cuatro/ bolas, hallar, en función de p , las probabilidades de /

que

- i) ninguna bola sea negra
- ii) al menos una bola negra
- iii) exactamente 2 bolas sean negras

Calcular estas probabilidades si (a) $p = 0.3$, (b) $p = 0.5$.

¡Probablemente los biólogos y los ingenieros civiles/ caerían dormidos si se les propusiese este ejercicio!

2.4.- ENSEÑANZA DE MASTER EN ESTADÍSTICA

Es práctica corriente en el Reino Unido, en Estados Unidos y en otros lugares ofrecer programas de un año de estudio conducentes a la obtención de un grado de Maestría en Estadística. Tales cursos se dirigen principalmente a graduados en Matemáticas (a veces a otros procedentes de áreas cuantitativas de estudio) con vistas a su // "conversión" a la Estadística. Su configuración típica es ta constituida por dos cuatrimestres con exámenes finales seguidos por cuatro meses de trabajo individual y la redacción de una tesis sobre dicho trabajo. Es mi opinión, y mi experiencia, que el trabajo individual comprenda análisis estadístico detallado de algún campo aplicado, coorientemente con la ayuda de un computador, junto con un informe escrito completo sobre el mismo.

El cuatrimestre debe implicar 250 horas de clases, los temas de estudio no serán muy distintos de los del curso/ de graduado de Estadística (pero dados por separado en lo fundamental). Sin embargo, en el grado en Estadística hay aproximadamente 700 horas de clases sólo de Estadística, / y es evidente que sólo una parte de este material puede / cubrirse en los cursos de Maestría. Parece, pues, que la proliferación de programas completos de primer grado en/ Estadística limitarán la utilidad de enseñar cursos de // Maestría y quizás que cambiemos nuestras opiniones acerca de cómo deben enseñarse para llevarlos a cabo.

3.- EDUCACION ESTADÍSTICA PARA LA PRACTICA ESTADÍSTICA

Disfruto resolviendo problemas matemáticos; disfruto/ haciendo crucigramas. Si un problema estadístico lleva a/ un problema matemático algo más interesante es más remunerador. Pero la Estadística no es solo Matemáticas. La Estadística como tema de estudio exige a los interesados seguir la senda matemática hacia la Teoría de la Probabilidad o hacia la Inferencia (como se enseña en la mayor par

te de Europa y de los Estados Unidos). Pero mi actitud / se basa en la tradición británica del empirismo. Creo // que los métodos estadísticos están diseñados para resolver// problemas reales (sean en la Arqueología, en la Industria, en la Medicina o en la Literatura Española). Esto no es negar la importancia del estudio matemático o / filosófico fundamentales que a menudo surgen de tales estímulos.

Recuerdo que el médico de mi localidad me propuso // "analizar un conjunto sencillo de datos -simplemente // ajustar una línea recta". Deseaba comparar cuatro instrumentos diferentes para medir la función pulmonar humana: evaluar la calibración y la precisión relativas: En BARNETT (2, 3, 4) se dan detalles de este y de otros problemas parecidos ¿Qué nos dicen estos ejemplos acerca de la educación estadística? Principalmente tres cosas:

1. La comunicación entre el estadístico y el investigador es parte del proceso educativo de ambos. El estadístico gana en experiencia y habilidad al ponerse en // contacto con otras disciplinas. Esto también estimula // nuevos desarrollos dentro del arsenal de los métodos estadísticos. El investigador gana al verificar lo que la Estadística brinda para comprender mejor su propia área/ de estudio. Por consiguiente, este papel del estadístico como consultor y colaborador es de vital importancia para su formación. Se ha escrito mucho sobre esto, por ejemplo, BARTHOLOMEW (5), BENJAMIN (7), SPRENT (20), CHANTER (9), FREEMAN (13), GREENFIELD (15), MARQUARDT (18), HOOKE (16). BARNETT (4) opina del papel omnipresente del estadístico bajo el encabezamiento de "El Estadístico: Comodín de todos los oficios, maestro de alguno?".

"el estadístico debe ser un traductor y un comunicador: necesita saber bastante de las disciplinas de los demás para conocer sus problemas, que debe expresar en términos estadísticos en cooperación con/ el desarrollo de los investigadores y utilizar las herramientas adecuadas y lo más importante, el comunicar las respuestas de manera comprensible. Así parecerá que tiene un papel más amplio que otros -como estadístico-matemático, calculista, filólogo, (físico, físico-nuclear y otros muchos) y no menos importante como comunicador. (Esta última destreza no se considera generalmente como una habilidad innata al científico). En conjunto, necesita ser un profundo conocedor de la Estadística, a la vez que un comodín de muchos otros saberes".

Este punto está bien explicado en un problema de la /

industria del petróleo que necesitó de las fuentes de muchas disciplinas diferentes BARNETT y LEWIS (1); BARNETT, (4).

2. El estudiante necesita ser adiestrado en la práctica estadística y en el trabajo de consultoría. Desde / luego, esto es evidente en la medida que concierne a determinados estadísticos aplicados. El estadístico agronómico, industrial, ó médico, el biómetro ó el psicómetra, todos necesitan ser adiestrados en su especialidad. Véase BISHOP (8), DEANE (11), SPICER (19), FINNEY (12), FREE MAN (13), GREENFIELD (15). Algunos de los problemas se / han expuesto en la segunda conferencia. Pero yo afirmaría que el profesional estadístico, el estadístico matemático o el simple matemático también deben ser adiestrados/ para aplicar ampliamente sus conocimientos a los problemas reales -para beneficiar a otros y para acrecentar el conocimiento estadístico.

Hasta aquí, he puesto de relieve la importancia de / la práctica en la educación: en particular en la segunda conferencia con respecto a los estudios universitarios / de Estadística. Estos deberían ser estructurados como // parte de un adiestramiento en consultoría estadística. La experiencia en consultoría de todos los profesores proporciona un material vital para las clases prácticas; es mejor aún si el estudiante, a nivel pre-ó post-graduado,/ puede ser implicado en trabajos reales de consultoría // (aunque sólo sea como mero observador). Ver COX(10); BARKERVILLE (6).

Este último punto tiene quizás las implicaciones más importantes para un Departamento de Estadística universitario.

3. El papel de asistencia y asesoría estadística de/ un Departamento de Estadística universitario es fundamental para crear el ambiente educativo adecuado. Creo que/ cualquier Departamento de Estadística debe considerarse como proveedor de servicios de tipo práctico para toda/ la Universidad, tanto en la enseñanza de estudiantes de otras disciplinas como en la más importante de ofrecer/ un servicio completo de consultoría estadística a todo/ el personal universitario y estudiantes de postgrado (y/ a organizaciones externas, si es lógicamente factible). Véase TUKEY, BOX, HARTLEY y KEMPTHORNE (21), GIBBONS y FREUND (14). Esto requiere algo más que sólo una actitud casual de estar preparado para ayudar si a alguien se le ocurre preguntar. El servicio de consultoría debe estar ampliamente difundido. El personal (no// necesariamente todo) debe encargarse de proporcionar este servicio y contar con la confianza para hacerlo así.

La Universidad necesita reconocer y apreciar estos servicios, debe proporcionarle los recursos necesarios, incluido equipo de cálculo y un miembro de su personal para que dirija el trabajo de cálculo. No hay nada peor que ofrecer un servicio y decir que el problema de alguien tiene solución en principio, pero que la falta de tiempo, la potencia de cálculo, etc. impide que sea resuelto en el momento. En el Departamento de Probabilidad y Estadística de Sheffield, 8 de 15 personas están en el servicio de consultoría estadística, con uno como supervisor del servicio. Cada año se llevan a cabo unos 100 trabajos de consultoría mayor, cubriendo unos 40 departamentos. Cada dos semanas se dedica una hora a Consultorio de Estadística y se da respuesta a problemas más complicados. Un ayudante de consultoría estadística hace el análisis de datos adecuado utilizando micro-computadores y terminales del computador de la Universidad. Algunos de los ejemplos típicos más recientes son:

Dr. Sherwell, Odontología, "Pruebas de una nueva pasta dentífrica".

Dr. Brownston, Educación, "Clasificación Triple sobre el uso de respuestas binarias de regresión logística".

El trabajo más interesante viene de las áreas más raras. Hemos hecho mucho en los campos de la arqueología y de la Prehistoria. El Dr. N.R.J. Fieller se acercó con algunos datos sobre la longitud y anchura de los dientes de caballo de la Era Devensian (hace 1500 años) y preguntó si los caballos habían muerto en determinadas estaciones o uniformemente a lo largo de todo el año. Esto condujo a fascinantes y sofisticados análisis y desarrollos (por ejemplo, usando los métodos de los cuantógramas del coseno) durante un largo periodo de tiempo para examinar los datos por "periodicidades". Otros ejemplos incluyen los pies Vikingo (en busca de una unidad de medida básica), sistema de pesos Troy (otra vez buscando una unidad) y el número de tocónes de árboles de Hazel de 5000 años antes de Cristo (indicativo de producción agrícola).

Comencé subrayando la necesidad en la educación estadística de que la asignatura esté viva e interese -plantear problemas para resolver y una oportunidad para resolverlos realmente. El éxito en la enseñanza de la Estadística depende de saber convencer a los demás de que la Estadística es importante (hasta entretenida). Este punto está bien claro (si bien algo irónicamente) en una bien conocida cita que merece la pena repetirse aquí de nuevo.

Sirve como portada al segundo volumen de KENDALL y STUART (17) y se atribuye a la novela *The undoing of Lamia Gurdeneck* de K.A.C. Manderville.

"No me has dicho todavía" dijo Lady Nutall, "a lo que se dedica tu novio". "Es estadístico" replicó Lamia con un fastidioso sentido de estar a la defensiva. Evidentemente, Lady Nuttal estaba desconcertada. No se le había ocurrido que los estadísticos entrasen dentro de las relaciones sociales formales. Había supuesto que la especie se perpetuaba de alguna manera colateral, como las mulas.

"Pero tía Sara, es una profesión muy interesante" dijo Lamia acaloradamente.

"No lo dudo" dijo su tía que evidentemente lo dudaba mucho. Expresar algo importante en meras cifras es tan evidentemente imposible que debe haber un campo de acción sin fin para un consejo bien pagado de como hacerlo. Pero, ¿no crees que la vida con un estadístico serfa, diremos, aburrida?

Lamia estaba silenciosa, se sentía renuente a discutir el sorprendente fondo de potencia emocional que ella había descubierto bajo la apariencia numérica de Edward.

"No son las cifras en sí mismas" dijo ella finalmente, "es lo que se hace con ellas lo que cuenta".

APENDICE I *

6. En una investigación acerca de la efectividad de un / determinado curso de lectura rápida, se dividió un // grupo de 500 estudiantes en dos, A y B, de 300 y 200/ estudiantes, respectivamente, elegidos al azar.

A los del grupo A no se les dió instrucción especial; a los del grupo B se les dió un curso de lectura rápi da.

A cada estudiante se le pidió que leyese el mismo pa saje y se midió el tiempo. Los resultados fueron

Grupo A: tiempo medio 78,4 seg. varianza 14 seg²

Grupo B: tiempo medio 77,4 seg. varianza 15 seg²

Efectúese una décima de significación para ver si hay evidencia de que el curso ha mejorado la rapidez la / lectura. Establézcase cuidadosamente la hipótesis nu la, la hipótesis alternativa y la conclusión final.

Se tuvo conocimiento posteriormente de que, de los / 500 estudiantes, 200 habían decidido por sí mismos // que necesitaban tomar el curso de lectura rápida y // que estos estudiantes se integraron en el grupo B. Dis cútase brevemente como ésto podría afectar a la con clusión anterior.

7. Se utiliza un telar para producir tela. Ocasionalmen te, al azar, la máquina oscila y aparece un defecto / a lo ancho de la tela. Con anterioridad se ha encon trado que la distancia entre defectos en la tela si gue una distribución exponencial con media 1 metro. Es tablecer brevemente por qué podría esperarse que una/ distribución exponencial sirviera como modelo razonable en este caso. En la tabla siguiente aparecen 100 distancias entre defectos consecutivos de cierto dfa.

Distancia entre defectos	Efectivos
0 a 0.5 m	44
0.5 a 1.0 m	20
1.0 a 2.0 m	22
2.0 a 3.0 m	10
Más de 3 m	4

(i) Construir un histograma para estos datos. Sobre / el mismo diagrama dibujar un gráfico, a escala ade cuada, de la función de densidad de la distribu ción exponencial de media 1.

(ii) Calcular los efectivos esperados correspondientes a una distribución exponencial con media 1 y doc imar si este modelo proporciona un ajuste razonable para los datos (Se aceptará una exactitud de/ dos cifras significativas para los efectivos espe rados).

APENDICE II *

(a) Actividades en Horas Libres y Deportes

Efecto de una alza en la exactitud de un rifle
Predicción de los resultados de futbol

Viendo y practicando deportes

Comparación de la distribución de puntos obtenidos / por los equipos de futbol y los obtenidos por casua lidad.

Actividades en horas libres de los estudiantes de / sexto.

(b) Trabajo

Comercio

Número de cajas necesarias en un determinado almacen de artículos deportivos

Investigación en los servicios farmacéuticos

Número de bombas necesarias en un garaje en régimen de autoservicio

Eficiencia de las verificaciones en un supermercado

Comparación de la circulación nacional de periódicos con la local

Industria

Averías en una de las furgonetas pertenecientes a // una compañía privada de transporte

Efecto de una parada para mantenimiento en una máqui na enlatadora de alimentos

¿Es 52 el promedio de cerillas en una caja?

Altura y volumen del contenido de botellas de 33 cen tilitros

Duración de vida de lámparas incandescentes compara da con las afirmaciones de los fabricantes.

* Traducido con la autorización de The Joint Matriculation Board, Manchester, U.K.

Transporte

Llegada de tráfico a una intersección en 2 períodos/
de tiempo diferentes
Accidentes en las carreteras locales en 1974 y 1975
¿Cuántos años puede durar en circulación un coche?
Modos de transporte de alumnos a una escuela
Efectos y eficiencia de un conjunto de semáforo su-/
urbanos.

(c) Fenómenos sociológicos

Tendencias y cambios en el gasto consumtivo.
Accidentes de carreteras y víctimas
Distribución en el país de personas con un apellido
determinado
Relación entre habilidad lectora y puntuaciones
Análisis de estadística criminales
Matrimonio y divorcio
Tasas de mortalidad en zonas industriales y no indus-
triales
Relación entre las puntuaciones de los exámenes par-
ciales y el final
Factores asociados con fumadores
Distribución según la edad de desempleo femeninos y
masculinos.

(d) Otros

¿Las relaciones consanguíneas incrementan la efecti-
vidad de la telepatía?
Predicción

Otras materias

Matemáticas

¿Son aleatorios los números aleatorios?
Docimasfa de números aleatorios
Principios de estimación
Cadenas de Markov
Efecto del número de tiradas en la distribución
del número de puntos que aparecen en el lanzamiento/
de un dado.

Ciencia

Efecto de las condiciones del suelo en la distribu-/
ción de plantas salvajes
Reproducción y supervivencia de ratones
Efectos de los niveles de agua en la germinación de/
semillas de lepidio

El efecto de fumar en el cáncer de pulmón
Tamaño de los guijarros y posición en playas turbona-
das
Preferencias de aromas
Factores que afectan a los tiempos de reacción
Pesos y edades de distintos alumnos
Color de los ojos y del pelo de los niños y sus pa- /
dres

Economía

El principio del máximo beneficio y la distribución /
de uso del suelo en el centro de la ciudad
Incremento en la población activa
Precio del oro y tasa del cambio
Análisis del Desarrollo Económico Internacional
Incremento de precios en la droquería local

Lingüística

Longitud de las frases en las novelas inglesas y fran-
cesas
¿Escribió Bacon las obras de Shakespeare?
Análisis gramatical de las noticias habladas y escri-
tas

Geografía

¿Varían la longitud de las ondas de los meandros con/
la distancia?
Tendencias demográficas
Variaciones en el uso de la tierra

Historia

Estudio de la población de Stockport de 1931 a 1971
Poblaciones de Huddersfield, por grupos de edad, de /
1901 a 1971

APÉNDICE III*

NIVEL 2

E GOLES A PORRILLO **

El Dos Temporadas

Las personas mayores dicen algunas veces que antes so-
ñan ver más goles en los partidos de fútbol. Trata de /
averiguar lo que piensan los adultos que conoces. ¿Cómo /
se podría determinar si se marcaban más goles antes?...

* Traducido con la autorización de W. Foulsham & Co., Ltd., Slough, Berkshire, U.K.

** De la obra "on the ball", páginas 12-14

Una forma sería mirar los goles marcados en la Primera División hace 20 años y en la actualidad.

La Tabla 6 nos muestra los goles marcados por los 22 equipos de Primera División en dos temporadas. Brian Clough fué el máximo goleador en la década de los 50, marcando unos 40 goles por temporada. A mediados de la década de los 70 Macdonald era el máximo goleador con unos 25 goles por temporada.

a) Compara los goles marcados por los equipos en la temporada 1957/8 con los de la temporada 1977/8. ¿Crees que se marcaron más goles en la década de los 50?

b) Puedes necesitar más información. Estudia lo que necesitas conocer para comparar los goles marcados en los 50 y en los 70.

Esto puede ayudarte a hallar el número total de goles marcados en cada temporada. Puedes también mirar los goles marcados por los equipos más y menos goleador.

Para las dos temporadas indicadas en la Tabla 6 calcula

c) el número total de goles marcados,

d) el promedio de goles (media), marcado por el equipo/ (divide por 22) ,

e) el menor número de goles marcados por un equipo,

f) el mayor número de goles marcados por un equipo,

g) redacta unas cuantas frases comparando el número de goles marcados en los 50 y en los 70. Cita cualesquier diferencias o semejanzas que observes en los resultados, como asimismo cualquier otra observación.

TABLA 6 GOLES MARCADOS POR LOS EQUIPOS DE 1^a DIVISION EN DOS TEMPORADAS

Equipo	Goles 1957/8	Equipo	Goles 1977/8
Wolves	103	Notts F	69
Preston	100	Liverpool	65
Tottenham	93	Everton	76
WBA	92	Man City	74
Man City	104	Arsenal	60
Burnley	80	WBA	62
Blackpool	69	Coventry	75
Luton	69	Aston V	57

Equipo	Goles 1957/8	Equipo	Goles 1977/8
Man Utd	85	Leeds	63
Notts F	69	Man Utd	67
Chelsea	83	Birmingham	55
Arsenal	73	Derby	54
Birmingham	76	Norwich	52
Aston V	73	Middlesbro	42
Bolton	65	Wolves	51
Everton	65	Chelsea	46
Leeds	51	Bristol C	49
Leicester	91	Ipswich	47
Newcastle	73	QPR	47
Portsmouth	73	West Ham	52
Sunderland	54	Newcastle	42
Sheff Wed	69	Leicester	26



NIVEL 4

OTRAS PROBABILIDADES **

¿Qué significa una Reacción Positiva?

a) ¿Cuántos conductores dieron una reacción positiva?

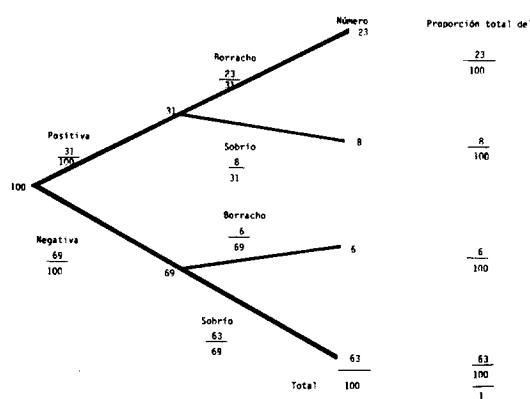
b) De éstos, ¿cuántos estaban borrachos?

c) De los que dieron una reacción negativa, ¿cuántos estaban borrachos?

Es posible dibujar un diagrama de árbol distinto a partir de la información contenida en la tabla.

** De la obra "Testing Testing", páginas 14-15.

Figura 2 Diagrama de árbol que muestra resultados



Utilizando la Figura 2, calcula:

- d) Probabilidad (un conductor que dió reacción positiva esté sobrio).
- e) Probabilidad (un conductor que dió reacción positiva esté borracho).
- f) Probabilidad (un conductor que dió reacción negativa esté borracho).
- g) Probabilidad (un conductor que dió reacción negativa esté sobrio).
- h) Probabilidad (un conductor esté borracho y de reacción positiva).
- i) Probabilidad (un conductor esté sobrio y de reacción positiva).
- j) Probabilidad (un conductor de reacción positiva).
- k) Pueden encontrar relaciones entre algunas de las respuestas para d a j. Si es así, escribe las primero como una suma de multiplicaciones que implican fracciones y luego exprésalo verbalmente, utilizando probabilidades.
- l) Reproduce la figura 2, insertando las cifras obtenidas en tu versión de la Tabla 1.

das en tu versión de la Tabla 1.

- m) Calcula las proporciones basadas en tus cifras y colócalas en los huecos existentes.
- n) Utiliza tu diagrama de árbol para encontrar las respuestas para d a j en tu simulación.

APENDICE IV*

PRIMER AÑO ESPECIAL (Programa IS)

INTRODUCCION A LA PROBABILIDAD Y A LA ESTADISTICA

Introducción a los Datos. Variación aleatoria con ejemplos. Métodos de representación de datos: Distribución / de frecuencias, histograma, gráfico de sectores, diagramas en forma de hoja y en forma de espiga. Medidas descriptivas: mediana, cuartiles, amplitud o recorrido, varianza, desviación estandar, distancia intercuartílica, / diagrama de urnas, coeficientes de asimetría y curtosis, codificación de datos, coeficiente de correlación. Inferencia: muestra, población.

Probabilidad. Espacio muestral y eventos. Definición de probabilidad: clásica, frecuencial. Axiomas de la probabilidad: estabilidad de las frecuencias como motivación/ ley de la adición. Probabilidad condicional: definición, / regla de la multiplicación, independencia, teorema de Bayes. Permutaciones y combinaciones y cálculo de probabilidades.

Variables aleatorias y Distribuciones de Probabilidad. / Variables aleatorias: definición espacio muestral inducido, variables aleatorias discretas y continuas. Distribución de probabilidad de una variable aleatoria: función/ de probabilidad, función de distribución, función de densidad de probabilidad. Esperanza: definición, esperanza/ de una función de variables aleatorias, momentos, varianza, valor esperado de una combinación lineal de una variable aleatoria. Distribuciones importantes de tipo discreto: Bernoulli, geométrica, binomial, Poisson, hipergeométrica. Distribuciones importantes de tipo continuo: rectangular, gamma, exponencial, nor-

* Traducidos con la autorización del Department of Probability and Statistics, The University, Sheffield, U.K.

mal.

Teoría adicional de las distribuciones. Transformaciones. Variables aleatorias bivariantes y multivariantes: introducción al caso discreto, funciones de probabilidad marginal y condicional, esperanza y varianza de sumas de dos variables aleatorias, bosquejo de su extensión a los casos bivariante continuo y multivariante general. Funciones generatrices: función generatriz de probabilidad, función generatriz de momentos, combinación lineal de variables aleatorias normales. Procesos de Poisson: definición, distribución del tiempo entre eventos y tiempo para el primer evento, distribución del número de eventos en unidades de tiempo.

Inferencia estadística. Fines generales de la inferencia. Muestra aleatoria. Distribución de $\bar{X}, \frac{1}{n} \sum (X_i - \bar{X})^2$ de // poblaciones normales: distribución chi-cuadrado. Distribución de \bar{X} de poblaciones no normales: teorema central del límite, aproximación normal para la binomial.

Estimación. Definición de estimador: estadígrafo, estimador, distribución muestral. Estimación insesgada: estimación insesgada de varianza mínima (ejemplos sencillos), eficiencia relativa. Métodos para obtener estimadores: mínimos cuadrados, máxima verosimilitud. Ajuste de distribución para los datos. Estimación por intervalos: intervalos de confianza y su interpretación, distribución/ t.

Dósimas de hipótesis. Esquema de la dósimia: definiciones de dósimas de hipótesis simple y compuesta, regiones de aceptación y exclusión, errores de tipo I y tipo II, potencia. Dósimas sobre medias de distribución normal: reflexiones de tipo práctico como motivación para la formación de dósimas, dósimas con σ^2 conocida y σ^2 desconocida. Dósimas sobre varianzas de distribución normal. Interpretación de los resultados de la dósimia: nivel de significación. Fundamentos teóricos: lema de Neyman-Pearson con ejemplos sencillos, dósimia más potente. Co_nexión entre dósimas de hipótesis e intervalos de confianza.

Dósimia bimuestrales y extensiones. Dos poblaciones normales: dósimia t para medias, dósimia F para varianzas, dósimia t para 2 muestras emparejadas. Análisis. Análisis de varianza de clasificación simple.

Regresión y correlación. Regresión lineal. Correlación: / docimas de $\rho = 0$ en poblaciones normales, transformación z de Fisher, conexión con la regresión lineal. Ex-/

tensiones de la regresión lineal simple: breve exposición de la regresión múltiple, de la regresión polinómica y / del modelo lineal general.

Inferencias con datos discretos y dósimas chi-cuadrado de bondad de ajuste. Dósimia de una muestra única de distribución binomial y de distribución de Poisson. Dósimia chi- cuadrado de bondad de ajuste: tablas de contingencia (dósimas de asociación y homogeneidad).

Dósimas no paramétricas. Conceptos generales. Dósimia de / los signos: muestra única, muestras emparejadas. Dósimas de las sumas de los rangos: dósimia de los rangos signados de Wilcoxon, dósimia de la suma de los rangos de Wilcoxon.

Métodos bayesianos y teoría de la decisión. Definición de la probabilidad como grado de creencia. Probabilidades a priori y a posteriori y uso del teorema de Bayes. Breve / exposición de los enfoques a la estimación puntual y por intervalos y la docimia de hipótesis. Teoría de la decisión: función de pérdida, función de riesgo, enfoque bayesiano via riesgo de Bayes.

BIBLIOGRAFIA BASICA PARA EL CURSO

HOEL, P.G.: Introduction to Mathematical Statistical Statistics. 4^a Edición, Wiley.

LIBROS DE CONSULTA

HUNTSBERGER D.V. & BILLINGSLEY P.: Elements of Statistical Inference. 4^a Edición, Allyn and Bacon.

MENDENHALL W.: Introduction to Probability and Statistics. 4^a Edición. Buxbury.

WETHERILL G.: Elementary Statistical Methods. Methuen

TABLAS

NEAVE M.R.: Statistical Tables. George Allen and Unwin.

APENDICE V*

EJEMPLOS DEL PRIMER AÑO DE ESTADISTICA

1.2. Hágase un breve comentario acerca de los siguientes/ (ligeramente modificados) informes publicados:

a) De una revista de hostelería: "En una encuesta re ciente sobre el uso de los comedores en las fábrí

* Traducidos con la autorización del Department of Probability and Statistics, The University, Sheffield, U.K.

- cas, en empresas de distintos tamaños, se ha encontrado que entre los trabajadores que disponen de acceso a comedores, 43% hacen uso de comedores pequeños, 38% usan comedores de tamaño medio y 35% hacen uso de comedores grandes".
- b) "En 1971 hubo 7443 accidentes mortales en carretera. El Ministerio de Transporte informa que 5926 de estas muertes ocurrieron durante las horas diurnas y sólo 1517 ocurrieron durante la noche. El Gobierno está considerando hacer una campaña para persuadir al público de que conduza de noche".
- 5.5. Supóngase que se hizo una prueba de diagnóstico del cáncer con las propiedades de que $P(A/C) = 0.95$, $P(\bar{A}/\bar{C}) = 0.95$, en donde C representa el evento de que una persona analizada tenga cáncer y \bar{A} representa el evento de que la prueba establezca que la persona analizada tenga cáncer. La prueba se aplica a personas elegidas aleatoriamente de una población con $P(C) = 0.0005$. Si la prueba es positiva, ¿Cuál es la probabilidad de que la persona tenga realmente cáncer? Discútase la complicación del resultado al implantar una prueba exhaustiva para el diagnóstico de cáncer en la población.
- Supóngase ahora que $P(A/C) = P(\bar{A}/\bar{C}) = r$. ¿Cuánto debe valer r para asegurarse que $P(C/A) = 0.95$?
- 7.3. Se inyectan varios ratones con organismos, dándose a cada ratón una dosis formada a partes iguales por dos cultivos A y B . Puede suponerse que el número de organismos efectivos por dosis en cada uno de los dos cultivos varía independientemente con distribuciones de Poisson de media 2. El ratón sobrevivirá si y sólo si no hay organismos efectivos en la dosis. Se examinan los ratones muertos para hallar si contienen organismos de uno o de los dos cultivos. Demostrar que la probabilidad de que un ratón examinado contenga organismos de un solo cultivo es 0.24.
- 10.3. Una muestra x_1, x_2, \dots, x_n se extrae de una distribución uniforme en el intervalo (α, β) . Demostrar que $2\bar{x}$ es un estimador insesgado de β . Demostrar que la función de distribución de $y = \max(x_i)$ es $(y/\beta)^n$ para $0 \leq y \leq \beta$. De aquí, demostrar que $(1 + 1/n)y$ es también un estimador insesgado de β . Evaluar la eficiencia de $2\bar{x}$ con respecto a $(1 + 1/n)y$.
- 16.1. Dos tipos de gasolina (A , B) se comparan con respecto a su "millaje" sobre el modelo 1979 de cierta marca de automóviles. De 15 de tales automóviles, elegidos al azar, a 7 se les sometió a la misma prueba con gasolina B . El millaje (millas/galón) para los dos grupos fué:
- | | | | | | | |
|---------|------|------|------|------|------|------|
| A: 22.3 | 23.1 | 20.4 | 21.4 | 22.9 | 22.6 | 20.9 |
| B: 28.7 | 20.3 | 22.6 | 24.5 | 22.6 | 23.8 | 21.9 |
- (i) El millaje parece variar más ampliamente para un tipo de gasolina que para el otro. Calcular un intervalo de confianza del 95% para σ_A^2 / σ_B^2 : razón de las varianzas para ambos tipos de gasolina.
- (ii) ¿Hay evidencia significativa de que el millaje medio es diferente para los dos tipos de gasolina sobre la base de que $\sigma_B^2 = 2\sigma_A^2$?

APENDICE VI *

ESTADISTICA IS

EJERCICIO PRACTICO OBLIGATORIO

En la hoja adjunta aparecen los resultados acumulados para la simulación del consultorio médico del Ejercicio Práctico Núm.5.

Como consultor estadístico, le han sido entregados estos resultados y preguntado que recomendaciones haría. Analice los datos y redacte un informe con sus conclusiones. El informe debe ser bastante detallado sobre el problema básico, sobre la naturaleza de los datos, sobre el método y resultados de los análisis y sobre sus conclusiones, de modo que el lector lo comprenda sin necesidad de información adicional. Existe, desde luego, un propósito práctico en el ejercicio: aconsejar al doctor sobre cómo debería organizar su trabajo. Al mismo tiempo que redacta su informe técnico, debe redactar una carta para el doctor informándole sobre sus conclusiones y recomendaciones de forma que le sea comprensible (evite los términos técnicos).

El informe y la carta deben entregarse antes de la última clase del cuatrimestre.

ESTADISTICA IS

PRACTICA 5

* Traducido con la autorización del Department of Probability and Statistics, The University, Sheffield, U.K.

Simulación

Un médico está considerando si merece la pena o no / adoptar un sistema de citas previas en su consulta de la tarde. La consulta funciona de 5,30 a 7,00 de la tarde / y el médico llega a las 5,30 en punto.

Actualmente, la puerta de la consulta se abre a las/ 5,20 y los pacientes llegan según un proceso de Poisson, de modo que el tiempo medio entre la llegada de cada paciente es de 5 minutos. A las 6,50 se cierran las puertas de la consulta y no se permite entrar a más pacientes en la sala de espera, pero el doctor ve a todos los pacientes que llegaron antes de esa hora. Cada consulta dura por término medio 5 minutos, según una distribución Gamma de parámetros (5,1). El médico atiende a los pacientes por orden de llegada.

De acuerdo con el sistema de citas previas propuesto, se harían 18 citas cada tarde con intervalos de 5 minutos, desde las 5,30 hasta las 6,55. El tiempo de llegada de un paciente se supone que está distribuido uniformemente desde 7 minutos antes a 3 minutos después de la hora de su cita. El doctor vería entonces a los pacientes por orden de cita, excepto cuando ésto representase un periodo de inactividad innecesario. Para este sistema propuesto, supondríamos no obstante que el médico llegaría a las 5,30 y que la distribución del tiempo de consulta no cambiaría.

Utilizando una tabla de números aleatorios y admis-

tiendo una variable exponencial (con media 1), simular // una tarde de consulta completa para cada uno de los sistemas, con una exactitud de 0.1 minutos. Representar los resultados gráficamente y anotar las observaciones siguientes para cada sistema.

- (i) Número de pacientes vistos por el médico.
- (ii) Tiempo en que finaliza la atención del médico.
- (iii) Tiempo total durante el cual el doctor está inactivo.
- (iv) Espera más larga para cualquier paciente antes de / ver al médico.
- (v) Tiempo de llegada del paciente en (iv).
- (vi) Número de sillas necesarias en la sala de espera.

Comparar los dos sistemas y basándose en los resultados hacer las recomendaciones apropiadas al médico. Suponer que el tiempo del doctor es más valioso que el de sus pacientes. Hacer también comentarios sobre la realidad de la hipótesis y distribuciones utilizadas en los modelos / anteriores.

N.B. Elegir al azar el punto de partida para los números aleatorios y para las variables aleatorias exponenciales y luego seguir fila a fila.

ESTADISTICA IS

DATOS PARA EL TRABAJO PRACTICO 5

(Grupos A & B)

SIN CITA PREVIA						CON CITA PREVIA					
Número de pacientes	Fin de atención	Doctor inactivo	Espera más larga	Cuando?	Número de sillas	Número de pacientes	Fin de atención	Doctor inactivo	Espera más larga	Cuando?	Número de sillas
17	92.5	2.3	16.5	48.9	4	18	97.9	0.0	12.1	43.1	3
16	89.2	14.4	9.7	10.0	4	18	110.9	7.0	24.7	75.6	4
12	77.7	32.2	6.1	-6.1	2	18	98.9	21.2	11.1	52.2	3
20	94.2	0.0	16.2	53.1	5	18	108.6	2.1	19.5	68.1	5
14	80.1	9.9	21.5	0.7	4	18	91.3	5.0	16.5	35.2	4
16	101.6	11.7	30.4	53.7	6	18	96.2	0.0	17.3	8.0	4
15	88.8	13.8	11.7	41.2	3	18	86.8	2.3	15.2	33.4	4
23	109.8	0.0	30.7	74.1	6	18	104.3	7.8	18.5	81.2	4
18	108.6	0.0	36.5	72.1	5	18	99.4	8.8	15.3	49.5	4
23	106.5	3.4	34.1	54.4	5	18	95.7	10.5	15.4	78.8	4
22	116.9	0.0	46.6	54.6	10	18	93.3	16.3	10.1	65.9	2
19	104.5	0.0	29.6	64.9	5	18	105.2	2.4	19.9	78.7	3
21	105.7	14.8	35.0	67.5	6	18	98.5	3.9	16.2	49.3	3
15	89.8	23.3	9.8	39.7	3	18	93.7	16.2	11.3	34.5	3
17	91.3	14.7	14.5	2.1	3	18	99.9	6.5	15.1	29.4	3
18	108.6	0.4	29.7	75.9	6	18	104.8	0.1	21.1	69.8	3
18	98.3	14.9	15.6	77.9	3	18	99.6	8.7	16.5	71.6	4
15	79.6	24.2	9.5	37.6	2	18	100.9	7.4	11.6	34.7	4
16	88.5	19.9	15.9	39.5	6	18	88.4	4.3	14.2	45.3	3
17	89.7	6.8	15.2	74.5	4	18	99.4	9.4	12.9	65.4	3
19	96.9	12.7	17.0	37.5	4	18	99.3	4.1	14.2	78.1	3
22	122.2	0.0	50.6	69.6	9	18	106.6	0.0	15.4	55.1	3
18	104.7	18.7	27.0	76.2	5	18	105.1	0.0	20.1	63.3	4
18	103.4	0.7	25.4	71.6	5	18	95.8	12.3	12.9	62.8	3
31	151.0	0.0	71.5	79.5	16						
20	121.9	16.7	47.0	69.1	5						
30	158.5	0.0	76.3	68.6	14						

APENDICE VII

ESTADISTICA APLICADA (Programa 3S)T (40 Conferencias Magistrales).

Fundamentos del Curso

Es un curso de evaluación continua, que consta de 12 proyectos distribuidos a lo largo de los dos cuatrimeses/tres. Cada proyecto es un problema práctico con datos reales (quizás de tamaño reducido algunas veces, para evitar que sean muy pesados de manejar y codificar los datos). El objetivo principal del curso es permitir a los estudiantes resolver un problema a través de las etapas/siguientes: Definición de objetivos, examen preliminar / de los datos, modelaje, análisis, cálculo, interpretación y comunicación de resultados. Se dará un plazo para cada proyecto y aquellos que se presentan tarde serán severamente penalizados. Es imprescindible garantizar la / realimentación necesaria a través de la cual los estudiantes aprenden de sus propios errores. La extensión y/ dificultades de los proyectos serán adecuados para que el estudiante medio sea capaz (con una dedicación razonable al estudio) de hacer un buen trabajo sin dedicar más tiempo a este curso que el que dedica a cualquier otro./ Más detalladamente, el fin del curso será:

1. Enseñar y estimular el uso de métodos gráficos de representación de datos y similares, para examen y análisis previos; el uso y efecto de transformación de / variables.
2. Introducir al estudiante en las técnicas de redacción informes y en la comunicación de resultados estadísticos a no estadísticos (el estudiante hará presentación oral y visual de un proyecto).
3. Dar al estudiante una oportunidad para buscar y encontrar información a partir de fuentes estadísticas (p. e. tendencias sociales).
4. Introducir al estudiante en la lectura crítica de artículos técnicos, extractando y resumiendo los puntos esenciales (p.e. cuatro artículos sobre "Estadística y Fútbol" -del capítulo del libro de Monroney sobre / "Goles, Cocos y Riadas" para el artículo de Reep y // Benjamin en Journal of Royal Statistical Society).
5. Enseñar al estudiante a "pensar estadísticamente", evaluar frases "pseudo-estadísticas" y estudiar hipótesis implícitas (p.e."es cinco veces más probable que /

una persona muera de un ataque cardiaco durante o inmediatamente después de un ejercicio violento que si/ estuviese sentado en un sillón").

6. Permitir al estudiante usar diversos equipos de cálculo y computación (e.p. calculadoras de bolsillo, programa de regresión lineal múltiple en un Apple, terminales remotos de 1906S). Los primeros 6 proyectos se distribuirán y se espera que sean completados a razón de uno por semana. Estarán descritos en relación directa a las técnicas estadísticas exigidas (no introducen nuevas técnicas -sólo usarán aquellas expuestas en IS 2S). Los proyectos posteriores se distribuirán/ a razón de 1 cada dos semanas * y, en general, serán / más abiertos y la técnica estadística "adecuada" a // utilizar no será necesariamente inmediata (en efecto, podrá haber líneas de enfoque alternativas). Cada estudiante tendrá que hacer una presentación oral y visual de uno de estos proyectos y para darle tiempo suficiente para prepararlo le serán asignados 4 semanas, en vez de 2. El número total de proyectos, por consiguiente, serán 6 (1 cada semana) + 5 (2 semanas/cada uno) + 1 (4 semanas) = 12. Para el proyecto de / presentación la clase se agrupará en pares o ternas y los diferentes grupos trabajarán sobre proyectos diferentes. Cada estudiante se espera que presente un informe escrito personal del proyecto de su grupo, como asimismo que participe en su presentación oral. En la evaluación se asignará bastante más peso (aproximadamente 3/4) al informe escrito cuando se compare con / la presentación oral.

Sera inevitable realizar algunos trabajos de computación, puesto que los proyectos se basan en problemas/prácticos con datos reales, pero se espera mantenerlos a un nivel bajo, a causa de que (1) el desarrollo y escritura de programas puede ser un proceso que consuma tiempo, particularmente cuando el estudiante es poco hábil y no muestra interés por la computación, lo que significará que los factores de la computación podrían dominar sobre los factores estadísticos, y (2)/ habrá disponible una opción de Computación y Modelaje Estadístico para los interesados en el tema.

Para la evaluación del rendimiento del curso como un/ todo, los proyectos individuales serán ponderados en/ proporción a su extensión.

* En Navidades se dará un proyecto para las vacaciones,/ que ocupara en total dos semanas del cuatrimestre. En/ Primavera se hará otro proyecto basado en un experimento, que ocupará las últimas semanas.

