Contents

1 The City and Surroundings

1.1 Glasgow

"Glasgow: Scotland with style. Let Glasgow Flourish!" Whoever dreamt up the city's motto got it in one. Flourishing is what Glasgow does best.

The panache with which the city has shaken off its reputation for post industrial decay, to seize the UK high ground in cosmopolitan chic is already legendary, an object lesson in reinvention for Britain's great Victorian cities.

With its commercial palaces reborn as sparkling shops, cafes and galleries, its once derelict waterfronts framing stunning new architectural vistas, and its thriving arts and design scene at once breeding grounds and magnets for international talent, Scotlands largest city has become a byword for creative energy and style.

Glasgow is unique among UK cities in boasting an identifiable style all its own, the creation, at the turn of the 20th century, of one of the world's most famous architects and three art school cronies. Just as Gaudi in Barcelona and Frank Lloyd Wright in Chicago left those cities forever indebted to their genius, so Charles Rennie Mackintosh brought international recognition to Glasgow with his unique, Scots-flavoured fusion of Modernist and Art Nouveau architecture and design.

Mackintosh's buildings: churches, public buildings and houses; bring pilgrims to Glasgow in their thousands. Mackintosh was both heir and inspiration to a built-in flair for innovation that is once again transforming this high-octane city. The medieval monastic settlement that became a great mercantile centre and, later, the shipbuilding capital of the world is emerging as a centre of creative diversity and one of the UK's top tourism destinations.

And it's done it in style. Visitors and shoppers stroll a largely pedestrianised centre to browse shops of a quality unseen outside London. The likes of Ralph Lauren and Escada, following in the footsteps of Armani and Versace, have been falling over themselves to acquire Victorian palaces whose architecture plunders a wealth of classical European styles.

Accommodation supply has boomed to match, its 38 per cent increase over the last five years representing another style explosion. Glasgow was one of the first cities in Europe to boast a boutique hotel and that tradition continues at numerous establishments including the quirky Arthouse and sumptuous Saint Judes.

The restaurant scene changes at a breathtaking rate to keep pace with Glasgow's insatiable appetite for the new. Established favourites, majoring on a traditional French treatment of the peerless Scots larder, keep company with a host of edgy newcomers, many of them reflecting Glasgow's status as Scotland's only genuinely multicultural city.

Each of Glasgow's incarnations has brought huge stylistic dividends. The 18th century tobacco barons who dominated the UK trade with Virginia left us their Georgian mansions disposed around an elegant city centre grid of streets.

Mackintosh himself owed a debt to the artisans and craftsmen who built the Clyde ships and adorned their interiors, says Pamela Robertson of The Hunterian Art Gallery at the University of Glasgow. In 1944, the shipping magnate Sir William Burrell gifted his lifetime's art collection to the city. Housed in an airy, purpose-built showcase of glass, steel and pink stone the Burrell Collection opened in 1983 to become an instant draw to art lovers the world over.

After the Second World War, the furnaces, the railway workshops, the mills and shipyards fell gradually silent, leaving grim urban deprivation in their wake. But Glasgow's indomitable, can-do spirit remained constant, fuelled by a self-belief that has always enabled Glaswegians to get out of reverse gear and find a way forward.

Nomination as Cultural Capital of Europe in 1990 not only allowed Glasgow to showcase her thriving artistic assets, but also gave the old girl the confidence she needed to let her got-it-flaunt-it tendencies rip. In 1999 the city celebrated its reign as UK City of Architecture and Design, both events embracing Mackintosh's achievements as central to the city's image of its creative, design-propelled self.

It is where Glasgow's stylistic legacy meshes with today's burst of creativity that the sparks fly brightest. A healthy synergy between Glasgow's mushrooming community of artists and designers and the multi-disciplinary media based in the city is bringing a spate of talent across the border. Glasgow has blossomed into the UK's centre for the creative industries reckons Stuart MacDonald, Director of The Lighthouse, Scotland's Centre for Architecture, Design and the City: "Glasgow can hold its own today with Barcelona and Brussels".

The same cross-fertilisation flourishes in the Centre for Contemporary Arts, where film and television makers, computer games designers, visual and performance artists, musicians and dancers all rub shoulders. "A cutting edge commitment to risk" is how the CCA's head of development Tracey Kelly defines the centre's commitment to fostering new young talent and to bringing cross-arts programmes to the widest possible audience.

Not that Glasgow, home to 200 arts organisations, has ever needed much telling on that front. Its thriving museums and galleries are the most visited civic collections of any cities in Britain. And a host of theatres boast performances by the national opera and ballet companies, as well as innovative drama and 'West End' shows.

Giant rock bands of the likes of Travis and Texas have followed a trail blazed by fellow Glaswegians Simple Minds and Wet Wet Wet. Meanwhile, a new generation, spearheaded by Indie sensation Franz Ferdinand and Glasgow-based Snow Patrol is well on its way to conquering the world. With live music, rock to dance, the city's cellars foster a respected underground scene; out of this have grown labels such as

Glasgow-based record companies Slam and Chemikal Underground and live acts such as Mogwai, Arab Strap and Brit award winners Belle & Sebastian.

"Ninety percent of the international artists who have ever played Scotland will tell you Glasgow audiences are the best in the world", says Dave Corbet of DF Concerts, the company responsible for promoting some 75 per cent of Scotland's live music events. "Warm, friendly, up for it, loud and appreciative. Glaswegians really know how to have a good time".

And none better than the city's 100,000-strong student population whose universities and colleges continue to lead the field in fostering some of Scotland's greatest inventive talent.

Focus of Glasgow's meteoric economic growth, the Clyde is coming back on stream. The gleaming plates of Norman Foster's Clyde Auditorium and the dazzling titanium Science Centre opposite reflect the shape of things to come. A ten-year project, providing 20,000 new jobs is currently transforming the old Pacific Quay into a media village.

Meanwhile, US architects the Richard Rodgers Partnership are soon to throw a stunning ellipse of a bridge over the river, where the glossy towers of a new international financial services district are reaching skywards above the Anderston and Hydepark Quays.

As one visiting journalist said: "When you hear of post-industrial cities claiming born again status you tend to think 'Oh yeah?' With Glasgow, it's 'Oh Wow!'

Adapted from 'Glasgow: Scotland with Style' by Juliet Clough at www.seeglasgow.com: the website of Glasgow City Marketing Bureau.

`source: http://www.gla.ac.uk/visitors/glasgow.html`

1.2 The University of Glasgow

The University of Glasgow dates from the middle of the fifteenth century, a time of critical change in Europe.

In 1451, the Scottish King James II persuaded Pope Nicholas V to grant a bill authorising Bishop Turnbull of Glasgow to set up a university. Thus, 40 years after the creation of St Andrew's University, Scotland, like England, could boast two universities. Modelled on the University of Bologna, Glasgow was, and has remained, a University in the great European tradition.

Initially, the young institution operated from Glasgow Cathedral and temporary accommodation nearby but in 1460 it was given property by Lord Hamilton on the east side of the High Street immediately north of Blackfriars which remained its home until 1870. In the seventeenth century, as the intellectual activity foreshadowing the Enlightenment took root, the University replaced the Hamilton buildings and created one of the finest buildings of the period in Scotland and which was described by contemporaries as 'the chief ornament of the city'. This building was known as the "Old College".

The University played its distinguished part in the Enlightenment and in fostering the research and inquiry which prepared the ground for the Industrial Revolution in which the city of Glasgow was to play a world role. Ironically it was the encroaching overcrowding and squalor of factories and railways, fruits of the industrial expansion it had helped to shape, which forced the University to move to its present site in what was then suburban Gilmorehill, a location it has occupied since 1870. Here the University celebrated its 550th anniversary in 2001.

Today, the University of Glasgow is one of the UK's leading universities with an international reputation for its research and teaching and an important role in the cultural and commercial life of the country.

With almost 16,000 undergraduate and 4,000 postgraduate students, it is one of the country's largest universities. Employing 5,700 staff, it is a major employer in the city and, with an annual turnover of £285M, it makes a substantial contribution to the local economy.

Firmly rooted in the West of Scotland from where it recruits 50% of its students, the University of Glasgow is nevertheless an international institution, attracting students from 80 countries and sending large numbers of students on study periods abroad. Today's research projects are typically international, with academics from every continent working in Glasgow while the University's own staff make valued contributions to collaborative work with some 200 institutions around the world.

Most of the University's 100 departments are to be found on the Gilmorehill campus, centred on Sir George Gilbert Scott's neo-Gothic main building. Its spire, added by his son John Oldrid Scott, is a landmark across the city. Glasgow's campus has more

listed buildings than any other and reflects a vast range of styles. Pearce Lodge and the Lion and Unicorn Staircase are relics of the old University, moved stone by stone to the new site. The circular Reading Room is a listed building from the 1930s while the Library, Boyd Orr and Adam Smith Buildings reflect post-war fashions in public building design. The new Wolfson Medical School Building, which provides state-of-the-art facilities for medical students and staff, was opened in 2003.

The University Veterinary School is located three miles away at the Garscube Campus which is also home to the new outdoor sports facilities. The University's Crichton Campus is located on the outskirts of Dumfries, in South West Scotland.

The University is a member of the Russell Group of major research-led universities and a founder member of Universitas 21, an international grouping of universities dedicated to setting world-wide standards for higher education.

`source: http://www.gla.ac.uk/general/welcome/past_present.html`

1.3 The Department of Statistics

Welcome to the Department of Statistics, University of Glasgow

It is a great pleasure to welcome the Research Students Conference back to Glasgow, and I hope that you will have an exciting and stimulating meeting. It is highly appropriate that you are here, in this the department's 40th birthday year, since Glasgow over those 40 years has been pleased to play host to the RSC on several occasions and many of the staff have memories of their own attendance as students. We place great importance on the role that our research students play in, and the contribution they make to, the life of the Department.

The Department of Statistics is one of the largest and oldest statistical groups in the UK. It was established in 1966, starting life with two professors (David Silvey and John Aitchison) one whose research was optimal design, the other whose interest lies in Medical Statistics. These two topics remain research interests in a department where the number of staff and students has grown steadily since those early days. Statistics is now part of the Faculty of Information and Mathematical Sciences, which also includes Mathematics, Computing Science and Psychology. It also has strong links with the Robertson Centre for Biostatistics which was established to carry out research in biostatistical methodology and to promote its application to practical problems addressing major medical and biological issues.

The Department has very wide and diverse research interests, just as the RSC programme has and this is one of the most rewarding aspects of involvement in Statistics. Statisticians have an opportunity to work with scientists in many other disciplines including Medicine, Astronomy, Environmental Science, Biology and Life Sciences to mention only a few and this work requires statistical research in many areas including topics such as nonparametric regression modelling; spatial statistics; multivariate analysis of data, including incomplete data problems; theory of optimum design; applications of Bayesian principles; bioinformatics; statistical genetics. I am sure that over the next few days, some, if not all, of these topics as well as many others will be mentioned.

Within this thriving atmosphere, the Department also has a substantial population of around twenty research students from the UK and beyond, many of whom are involved in organising your conference. I would like to take this opportunity to thank them for all their hard-work in organising, what I am sure will be, a most enjoyable and scientifically rewarding conference.

Marian Scott
Head of Department

2 Travel and Maps

2.1 Travel within Glasgow

- **UNDERGROUND:** The main station for the University is Hillhead. Kelvinhall and Kelvinbridge stations are more convenient for some locations (see map of Glasgow).
- **LOCAL TRAINS:** The nearest suburban rail station is Partick, about one mile to the west of the University. It has an interchange with the underground and with bus services on Dumbarton Road.
- **TAXIS:** Glasgow Taxis (black cabs, 0141 429 7070) can be hailed at most times in the university area and city centre and can also be booked. Pacific Cars (0141 429 4040) is a private hire cab company and can only be booked but are often cheaper than black cabs.
- **CYCLE RACKS** are provided on campus at A21, B3, B4, E1 and E2 (see Campus map).

2.2 Getting to the Conference Centre (Garscube)

- Two ways from **City Centre to Garscube:**
    1. If you want to get a taxi, at Central Station, the Gordon Street entrance (straight ahead as you exit station), there is a taxi rank directly across the road with taxis there all hours of the day. If you exit the station at either side exit then there is a taxi rank directly outside again. At Queen Street Station, at either side of the station next to the exits there are taxi ranks.
    2. If you want to get a bus, then go to Hope Street where the 40 and/or 59 stop. Both of these will take you to Garscube Sports Complex (the driver should give you a heads up if you need it). The route 40 or 59 bus can also be taken from Buchanan Street Bus Station to Garscube Sports Complex.
- From **The University to Garscube:**
    1. Alternatively, you can get the underground from the City Centre (St. Enochs or Buchanan Street) to the University (Hillhead). You would then have to take a taxi or bus to Garscube.

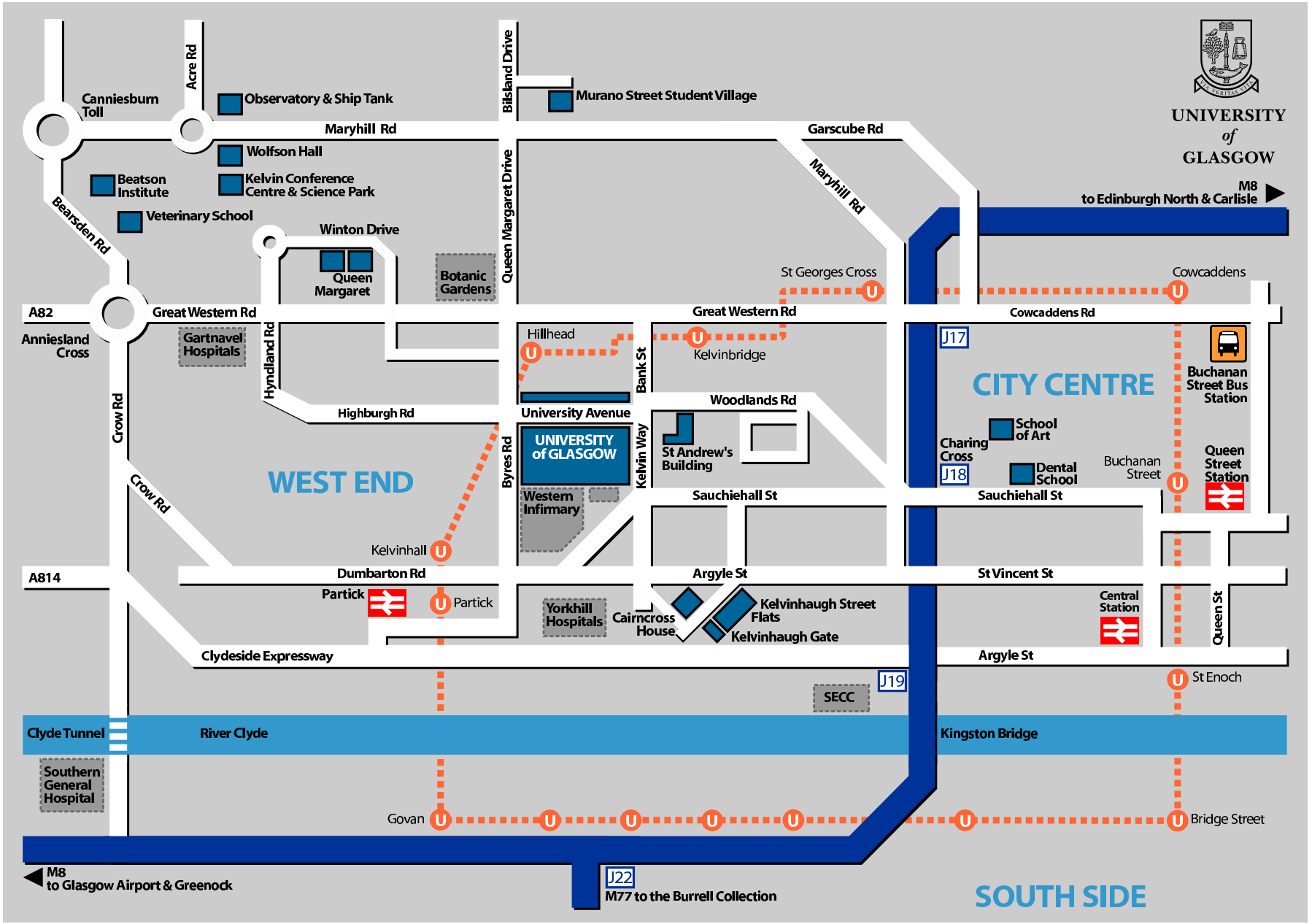

Figure 1: Map of Glasgow

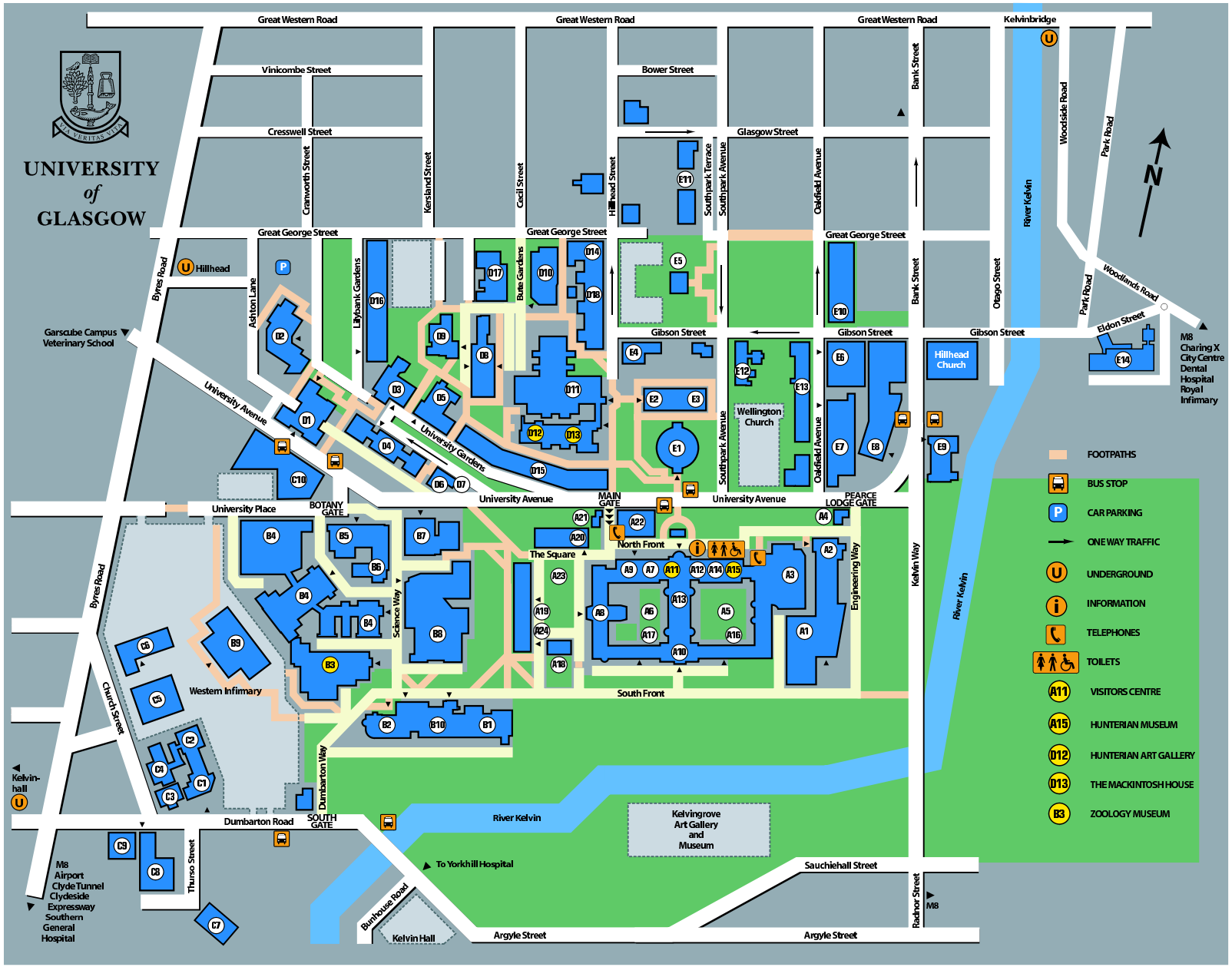

Figure 2: Campus Map

KEY TO BUILDINGS AND DEPARTMENTS:

Accommodation Services	D14 - 73 Great George Street
Accounting & Finance	E12 - 65-73 Southpark Avenue
Adam Smith Building	D8
Adult & Continuing Education	E14 - St Andrews Building 11 Eldon S
Aerospace Engineering	A1 - James Watt South Building
Anderson College	C1
Archaeology	D2 - Gregory Building
Arts Faculty Office	D15 - 6 University Gardens
Bedellus Office	A12 - Main Building
Bower Building	B7
Boyd Orr Building	D1
Business & Management	A6 - West Quadrangle
Business School	A6 - West Quadrangle
Bute Hall	A13
Careers Service	D15 - 3 University Gardens
Carnegie Room	A10 - Main Building
Catering Service	A9 - 1A The Square
Catholic Chaplaincy	E11 - 15 Southpark Terrace
Celtic	D5 - Modern Languages Building
Central & East European Studies	D10 - Hetherington Building
Central Room Bookings	B6 - Estates & Buildings Office
Central Services	A21 - Main Gatehouse
Centre for Drugs Misuse Research	C9 - 89 Dumbarton Road
Centre for Exercise Science & Medicine	B2 - West Medical Building
Centre for Medical History	D15 - 5 University Gardens
Chapel	A8
Chaplaincy	A24 - 11 The Square
Chemistry	B4 - Joseph Black Building
Children's Panel Training Group	E14 - St Andrew's Building 11 Eldon S
Civil Engineering	E7 - Rankine Building
Classics	A6 - West Quadrangle
Clinical Physics	Western Infirmary
College Club	A9
Computing Science	D16 - 17 Lilybank Gardens
Computing Service	A2 - James Watt North Building
Concert Hall	A12
Conference & Visitor Services	A20 - 3 The Square
Court Office	A10 - Main Building
Davidson Building	B1
Dermatology	C2 - Robertson Building
Development & Alumni Office	A20 - 3 The Square
Dining Rooms	A9
Divinity (School of)	A20 - 4 The Square
Earth Sciences	D2 - Gregory Building
East Quadrangle	A5
Economic & Social History	D15 - 4 University Gardens
Economics	D8 - Adam Smith Building
Education Faculty	E14 - St Andrew's Building 11 Eldon St
Education Faculty Office	E14 - St Andrew's Building 11 Eldon St
Electronics & Electrical Engineering	E7 - Rankine Building
Engineering Faculty Office	A1 - James Watt South Building
English as a Foreign Language (EFL)	D18 - 52 Hillhead Street
English Language	D15 - 12 University Gardens
English Literature	A6 - West Quadrangle
Estates & Buildings	B6 - Botany Gate
External Planning & Marketing	A20 - 2 The Square
External Relations & Marketing	A20 - 2 The Square
Finance Office	A16 - Main Building
Financial Studies (School of)	E12 - 65-73 Southpark Avenue
Florentine House	E4
Fore Hall	A8 - Main Building adjoining Chapel
Forensic Medicine & Science	B4 - Joseph Black Building
French Language & Literature	D5 - Modern Languages Building
Gardiner Institute	C6 - Western Infirmary
GAUT	E13 - 68 Oakfield Avenue
Geography & Geomatics	A5 - East Quadrangle
George Service House	D7
German	D5 - Modern Languages Building
Gilmorehill Centre	E9 - 9 University Avenue
Glasgow University Sports Association	E13 - 62 Oakfield Avenue
Glasgow University Union	E8 - 32 University Avenue
Graham Kerr Building	B3
GUIDE	E14 - St Andrew's Building 11 Eldon St
HATII - Faculty of Arts	D7 - George Service House
Hetherington Building	D10
Hetherington House	D6
Hispanic Studies	D10 - Hetherington Building
History	D15 - 10 University Gardens
History (Mediaeval)	D15 - 10 University Gardens
History (Modern)	D15 - 1-2 University Gardens
History (Scottish)	D15 - 9 University Gardens
History of Art	D15 - 8 University Gardens
History of Medicine	D15 - 5 University Gardens
Hospitality Services	A23 - 1a The Square
Human Resources	A10 - Main Building
HUB	E2
Hunter Halls	A14 - Main Building
Hunterian Art Gallery	D12 - Hillhead Street
Hunterian Museum	A15 - Main Building
Institute of Biomedical & Life Sciences (IBLS)	
- Administration	B2 - West Medical Building
- Biochemistry & Molecular Biology	B3 - Bower Building
- Environmental & Evolutionary Biology	B3 - Graham Kerr Building
- Graduate School	B2 - West Medical Building
- Infection & Immunity	B4 - Joseph Black Building
- Molecular Genetics	C1 - Anderson College
	C3 - Pontecorvo Building
- Neuroscience & Biomedical Systems	B2 - West Medical Building
- Undergraduate School	B10 - Wolfson Building
- Virology	C4 - Church Street
Immunology and Bacteriology	Western Infirmary
Information Services Dept	A10 - Main Building
Institute for Art History	A17 - Main Building
Internal Audit Services	A10 - Main Building
IT Education Unit	E1 - McMillan Reading Room &
	E14 - St Andrew's Building 11 Eldon St
Italian	D10 - Hetherington Building
James Watt North Building	A2
James Watt South Building	A1
John McIntyre Building	A22
Kelvin Building	B8
Kelvin Gallery	A7 - Hunterian Museum
Laboratory of Human Anatomy	A3 - Thomson Building
Language Centre	D10 - Hetherington Building
Law (School of)	A19 - Stair Building
Learning & Teaching	A10 - Main Building
Learning Works	E14 - St Andrew's Building 11 Eldon St
Library	D11 - Hillhead Street
Lilybank House	D9
Mackintosh House	D13 - Hunterian Art Gallery
McMillan Reading Room	E1
Mail Room	D1 - Boyd Orr Building
Main Building	A10
Main Gatehouse (Security)	A21 - Main Gate, University Avenue
Management Information Services	A10 - Main Building
Mathematics	D4 - Maths Building
Mechanical Engineering	A1 - James Watt South Building
Media Services	E5 - Southpark House
Medical Faculty Office	C10 - Wolfson Medical School Building
Modern Languages Building	D5 - 16 University Gardens
MRC Medical Sociology Unit	D16 - 4-7 Lilybank Gardens
MSF Union	E13 - 68 Oakfield Avenue
Music	D15 - 14 University Gardens
Naval Architecture & Marine Engineering	100 Montrose Street
Glasgow G4 0LZ	
Nursery	D19 - 28 Hillhead Street
Nursing & Midwifery School	E10 - 57-61 Oakfield Avenue
Officer Training Corps	B5 - University Place
Pearce Lodge	A4
Philosophy	E10 - Henry Jones Building
Photographic Unit	D11 - University Library
Physics & Astronomy	B8 - Kelvin Building
Planning Office	A10 - Main Building
Politics	D8 - Adam Smith Building
Pontecorvo Building	C3
Postgraduate Club	D6 - 13 University Gardens
Principal's Lodgings	A18 - 12 The Square
Principal's Office	A10 - Main Building
Print Unit (Gilmorehill)	A1 - James Watt South Building
Psychology	D18 - Hillhead Street Terrace
Public Health	D16 - 1/2 Lilybank Gardens
Publicity Services	A20 - 2 The Square
Purchasing Office	A10 - Main Building
Queen Margaret Union	D3 - 22 University Gardens
Radiation Protection Service	B8 - Kelvin Building
Randolph Hall	A10 - Main Building
Rankine Building	E7
Rector's Office	A22 - John McIntyre Building
Refectory	E3
Registry	A10 - Main Building
Research & Enterprise	A19 - 10 The Square
Robert Clark Centre for	
Technological Education	E14 - St Andrew's Building 11 Eldon St
Robertson Building	C2
Robing Room	A12 - Main Building
	(enter via Hunterian Museum stair)
Safety & Environmental	
Protection Services	A4 - Pearce Lodge
SATRO	B8 - Kelvin Building
Science Faculties Support Unit	D1 - Boyd Orr Building
Scottish Literature	D15 - 7 University Gardens
Security (Central Services)	A21 - Main Gatehouse
Senate Office	A10 - Main Building
Slavonic Studies	D10 - Hetherington Building
Social Policy & Social Work	D8 - Adam Smith Building
Social Sciences Faculty Office	D8 - Adam Smith Building
Sociology & Anthropology	D8 - Adam Smith Building
Southpark House	E5
Special Needs Service	A22 - John McIntyre Building
Sport & Recreation Service	E6 - Stevenson Building
Staff Development Service	A10 - Main Building
Stair Building	A19
St. Andrew's Building	E14 - Eldon Street
Statistics	D4 - Maths Building
Stevenson Building	E6
Student & Staff Support Division	A18 - 13 Professor Square
Student Counselling & Advisory Service	E10 - 65 Oakfield Avenue
Student Information Desk	E1 - McMillan Reading Room
Student Recruitment &	
Admissions Service	A20 - 1 The Square
Students' Representative Council (SRC)	A22 - John McIntyre Building
Teaching & Learning Service	E4 - Florentine House
Telephone Exchange	A10 - Main Building
TGWU	E13 - 68 Oakfield Avenue
Theatre Film & TV Studies	E9 - Gilmorehill Centre
Theology & Religious Studies	A20 - 4 The Square
Thomson Building	A3
Turnbull Room	A10 - Main Building
UNISON	E13 - 68 Oakfield Avenue
University Archive Services	C7 - 13 Thurso Street
	C8 - 77-87 Dumbarton Rd
University Health Service	E10 - 63 Oakfield Avenue
University Transport Services	C8 - 11 Thurso Street
Urban Studies	D17 - 25 Bute Gardens
Visitor Centre	A11 - Main Building
West Medical Building	B2
West Quadrangle	A6
Western Clinical Research &	
Education Centre	C5 - Western Infirmary
Western Infirmary Lecture Theatre	B5
Widening Participation Service	E11 - 12 Southpark Terrace
Wolfson Building	B10
Wolfson Medical School Building	C10
Zoology Museum	B3 - Graham Kerr Building

3 Accommodation at Wolfson Hall

Accommodation will be in Wolfson Hall which is a part of the Kelvin Conference Centre.

Set in three acres of quiet woodland, Wolfson Hall offers a city location with a get-away from it all atmosphere. Offering both en-suite and standard Bed & full Scottish Breakfast accommodation, Wolfson Hall is the perfect base if you are visiting Glasgow for business and/or pleasure.

All rooms are tastefully furnished, and include tea & coffee making facilities. Guests also have access to shared kitchens and all bedrooms have telephone access (phone cards available from Reception).

Other facilities on offer include common rooms, TV rooms and games room along with free of charge laundry facilities and free off-street parking.

Based in the same complex as Wolfson Hall is the Garscube Sports and Recreation facility offering superb leisure facilities for your use. From a quiet run along the riverbank to a game of rugby or football, the Garscube facility can accommodate most tastes. Along with outdoor pursuits, a full programme of aerobic activities is on offer and the fully equipped Cardio-Vascular Suite can be used at a specially discounted rate.

Full details of Garscube facility are available from Wolfson Hall on arrival.

`source:`
`http://www.cvso.co.uk/accommodation/residences/res_wolf.htm`

4 Kelvin Conference Centre and Conference Details

The conference opening will take place in Kelvin Conference Centre. On Monday 20th, delegates should arrive at Wolfson Hall between 2:00 - 3:30pm to check-in to their accommodation and collect conference packs containing all the information needed during the conference.

On Tuesday 21st and Wednesday 22nd, presentations will take place in the Mathematics Building at the University of Glasgow. On both days, coaches will depart from Wolfson Hall after breakfast to transport delegates to the Mathematics Building. Coaches will then return delegates to Wolfson Hall at the end of the day.

Breakfast on the 21st - 23rd and dinner on Monday 20th will be in Wolfson Hall. Lunches will be available in the Food Factory of Queen Margaret Union opposite the Mathematics Building. Dinner on Tuesday 21st will be in Gong restaurant close to the Mathematics Building. If you haven't already done so, please let us know if you have any special dietary requirements.

For detailed instructions on how to get to Wolfson Hall, please see Section 2.

WOLFSON HALL

1. Amenities Block: Snooker, Gym
2. Library
3. Computer Room
4. Common Room
5. TV Room
6. Shop/Bar
7. Public Toilets
8. Public Telephones
9. Main Entrance Hall (Wolfson Hall)
10. Reception
11. Warden/Manager's Office
12. Boardroom
13. Dining Room
14. Guest Wing
15. Deputy Warden's House
16. Warden's House
17. Manager's House
18. Kelvin Conference Centre
19. Bike Sheds x 2
20. Laundries x 4
21. Kelvin Conference Centre Reception
22. Front Door 2 (KCC)
23. Front Door 3 (KCC)
24. E Block TV Room
25. Fraser Room
26. Glasgow Room
27. Coffee Shop
28. Music Room
29. Student Committee Office
30. Night Security Desk

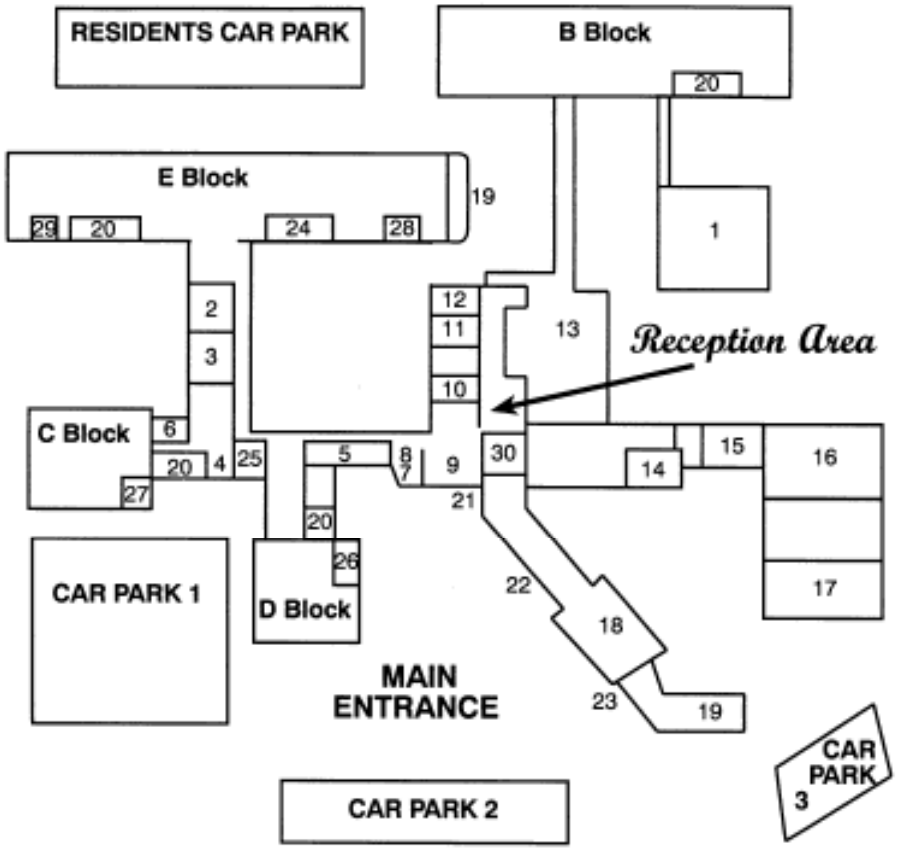

Figure 3: Floor Plan of Conference Centre and Accommodation

5 Meals

5.1 Dining at Wolfson Hall

Dinner on Monday 20th March and breakfast on Tuesday 21st, Wednesday 22nd and Thursday 23rd March will be in the dining room at Wolfson Hall.

5.2 Breaks and Lunch in Queen Margaret Union

Morning refreshment breaks on Tuesday 21st March and Wednesday 22nd March will be in Qudos in Queen Margaret Union.

Lunch on Tuesday 21st and Wednesday 22nd March will be in the Food Factory in Queen Margaret Union where you will be issued with £5 lunch vouchers for both days. This should easily be enough to cover a two course meal with a drink, but should the value of your lunch exceed this amount you will be asked to pay the difference. Lunch in the Food Factory is self service.

To find your way to Queen Margaret Union see the University campus map which can be found in Section 2. The porter at the main entrance will be able to direct you to the Food Factory.

5.3 Dining at Gong Restaurant

Dinner on Tuesday 21st March will be in Gong restaurant on Vinicombe Street, off Byres Road.

Housed within the former Salon Cinema, Gong is the West End's best restaurant & bar, specialising in an eclectic mix of flavours & styles from around the globe. Gong provides a stunning space to eat & drink. The balcony and stage bars specialise in fantastic cocktails and are the haunts of the West End's coolest crowd.

`source: http://www.g1group.co.uk/content/default.asp?page=s18_2`

5.4 Conference Dinner at St Andrews in the Square

On Wednesday 22nd March there will be a wine reception with a display from the sponsors to learn more about some possible career opportunities. The conference dinner will follow and both these events will take place at St Andrews in the Square, a beautiful 18th century restored church.

As is tradition in Scotland, after the dinner there will be a ceilidh (Scottish dancing). Kilts are optional, dancing is not!!! (well dancing really is optional but you won't be able to help yourself!)

6 Things to Do

6.1 Food and Drink

The West End

In the heart of the West End of Glasgow, Byres Road is buzzing with entertainment, bars, restaurants, delicatessens, student coffee bars and interesting shops.

Byres Road is the perfect location for a day of exploration, shopping, dining and drinking. You could start the day off with some self-indulgent shopping: Byres Road is filled with loads of great little eclectic shops selling just about everything.

You can take your pick from all the great restaurants on Byres Road. Whether you enjoy Chinese, Belgian, bar food, or Italian, there's something here for all tastes. The pick of the crop include :

- Antipasti (Italian cafe/restaurant)
- Brel (specialising in Belgian food and beers)
- Cul de Sac (cheap pasta, crepe and burgers)
- The Ubiquitous Chip (traditional and original Scottish recipes)

Afterwards, you may want to take in a few bars, of which there are many on Byres Road! Ashton Lane off Byres Road is a great place to go for a few drinks any time of the year.

`source:`
`http://www.myglasgow.org/glasgow/restaurants-byres_road.htm`

The City Centre

It used to be known primarily for its Indian food, but Glasgow now has a great variety of places to eat, drink and make merry. Contemporary cuisine, Euro-food and organic menus are all on offer for the hungry client. There's no shortage of pubs and bars either, so dining and drinking in this Scottish city are a real pleasure, whatever your tastes.

The Merchant City has a healthy collection of bars and restaurants, which are all conveniently located near to each other. Feeling spicy? If you fancy fajitas or chilli, Pancho Villas' lively atmosphere and tasty menu will add sizzle to your evening, or try Khublai Khan for a taste of Mongolian magic. If this sounds too hot, pop into the

Candleriggs branch of Oblomov and cool down with some icy cold vodka and a big plate of goulash. Mao, an export from Dublin, offers delicious yet healthy Oriental fusion, dished up in huge bowls. Local food is also well represented in this area. In the heart of Merchant City, Schottische and Rab Ha's serve excellent Scottish fare, while the City Merchant specialises in seafood and local cuisine. On the same street as the latter, try Granny Black's for a traditional pub atmosphere. Merchant City is a popular night-time haunt for Glasgow's beautiful people so there's no shortage of fashionable bars to be seen in. Try Bargo if you think you're cool enough, while Bacchus and Bar 91 have a more relaxed atmosphere. Corinthian and Arta attract a civilised and slightly older clientele who feel at home in the equally palatial settings.

The city centre, unsurprisingly, has a greater selection of eateries than any of the other districts. All the popular menus are on offer here, from Chinese to Indian, French to Italian. Curry with good music is the dish of the day at Bombay Blues and Kama Sutra puts the spice into Baltis. Malmaison Brasserie and 78 St Vincent offer fine French fare in opulent surroundings. If you're a fan of pasta and pizza, you'll love the enormous amount of Italian restaurants in the city centre. Fratelli Sarti has a lively, vibrant atmosphere and Rico's is a top place to eat. La Tasca, just around the corner, is popular for munching on tasty tapas, no matter what the hour. When it comes to seafood, you can't beat Rogano for quality or luxury, although this is a restaurant best visited when somebody else is paying. Bars to check out include Strata, Spy Bar, Budda and the Bier Halle Republic. The latter is representative of an East European trend amongst Glasgow's newer drinking holes. For a more traditional atmosphere, admire the interiors in The Counting House or The Drum and Monkey, both of which are housed in former bank buildings, or call into The Horseshoe to discover why it merits an entry in the Guinness Book of Records.

Most of the options for a late drink or meal are to be found around Charing Cross on either side of the motorway, which separates the city centre from the west end. Canton Express, Glasgow Noodle Bar and Pattaya all serve food until the early hours of the morning while Insomnia never closes. For an alcoholic beverage when it's past normal closing time, try Cas(Bah) or the Baby Grand.

```
source: http://travel.yahoo.com/...
        ...p-travelguide-2681921-glasgow_restaurants_and_bars-i
```

6.2 Going Out

Nightclubs

In the West End, The Viper and Oran Mor are the only nightclubs. However, there are a variety of clubs in the city centre especially on Sauchiehall Street. Some of the more popular nightclubs for students are:

- The Garage
- ABC
- Campus
- WigWam

If you would like more information please ask one of the conference organisers.

Theatres

The two main theatres in Glasgow are The Kings (Tel: 0141 240 1111) and The Theatre Royal (Tel: 0141 240 1133) .

Cinemas

The closest cinema is The Grosvenor on Ashton Lane (Tel: 0141 339 8444) and the main cinema in the city is Cineworld at 7 Renfrew Street, Glasgow, G2 3AB (Tel: 0871 200 2000).

Tourist Information

Glasgow Tourist Information Centre: 11 George Square, Glasgow, G2 1DY (Tel: 0141 204 4400).

```
Email: enquiries@seeglasgow.com
Web: www.seeglasgow.com
```

7 Help, Information and Telephone Numbers

Accommodation

Kelvin Conference Centre
West of Scotland Science Park
Maryhill Road
Glasgow
G20 0TH
Tel: 0141 330 2586

Department of Statistics, University of Glasgow

Department of Statistics
15 University Gardens
University of Glasgow
G12 8QW
Tel: 0141 330 5024

Emergency Numbers

Emergency Conference Organiser (Helen Parker)	07747 885599
Local Police	0141 532 3700
Western Infirmary	0141 211 2000

Transport

Public Transport:

Traveline Scotland (all public transport information)	0870 608 2608

Taxis:

Black Cabs	0141 429 7070
Pacific Cars	0141 429 4040

8 Computing and Internet Access

Free computer and internet access will be available to all delegates in Computing Laboratory 215, opposite Lecture Theatre 203, in the Mathematics Building of the University of Glasgow. To access the Laboratory type the code below into the key pad on the door and turn the silver handle to the right.

Opening Hours:	9:00am-5:00pm
Door Code:	C734XZ
username:	rscuser
password:	rsc&2006

9 Timetable of Events

9.1 Monday 20th March

14:00-15:30 Arrival and registration of delegates (Reception, Wolfson Hall)

15:30-16:00 Tea and coffee (Foyer, Kelvin Conference Centre)

16:00-16:15 Conference Opening (Lecture Theatre, Kelvin Conference Centre)

Welcome talk delivered by the Head of Department of Statistics, Professor Marian Scott
Opening address from the Vice-Principal of Physical Sciences & Engineering, Professor Robin Leake

16:15-17:50 Plenary Session (Lecture Theatre, Kelvin Conference Centre)

16:15-17:00	Professor Mike Titterington (Chair of Statistics, University of Glasgow)
17:00-17:50	Professor Stephen Senn (Professor, University of Glasgow)

18:30-19:30 Dinner (Dining Room, Wolfson Hall)

20:00-23:00 First Evening Entertainment
Pub Quiz (Dining Room, Wolfson Hall)
Movie (Lecture Theatre, Kelvin Conference Centre)

9.2 Tuesday 21st March

07:30-08:30 Breakfast (Dining Room, Wolfson Hall)

08:40 Coaches depart accommodation for Mathematics Building

09:25-13:00 Morning Session
- 09:25-11:00 Session 1 (Rooms 203, 214, 325, 326)
- 11:00-11:25 Break for refreshments (Qudos, Queen Margaret Union)
- 11:25-13:00 Session 2 (Rooms 203, 214, 325, 326)

13:00-14:15 Lunch (Food Factory, Queen Margaret Union)

14:15-15:50 Session 3 (Rooms 203, 214, 325, 326)

15:50-18:30 Afternoon Break
- 15:50-18:30 Football (Garscube Sports Complex)
  - 16:00 Coach departs Mathematics Building (for those playing football) to return to accommodation
  - 18:30 Coach departs accommodation (for those playing football) for dinner
- 17:15-18:30 RSS Guest Speaker (Room 203)
  Professor Tim Holt (University of Southampton)

18:45-20:15 Dinner (Gong Restaurant, West End)

20:30-23:00 Second Evening Entertainment
- Whisky Tasting (Reading Room, Glasgow University Union)
- Sports Bar (Playing Fields, Glasgow University Union)
- Cinema (Grosvenor Cinema, Ashton Lane)

23:00 Coaches depart to return to accommodation

9.3 Wednesday 22nd March

07:30-08:30	Breakfast (Dining Room, Wolfson Hall)
08:40	Coaches depart accommodation for Mathematics Building
09:25-13:00	Morning Session
	09:25-11:00 Session 4 (Rooms 203, 214, 325, 326)
	11:00-11:25 Break for refreshments (Qudos, Queen Margaret Union)
	11:25-13:00 Session 5 (Rooms 203, 214, 325, 326)
13:00-14:15	Lunch (Food Factory, Queen Margaret Union)
14:15-15:25	Session 6 (Rooms 203, 214, 325, 326)
15:30	Coaches depart Mathematics Building for accommodation
16:00-16:45	Refreshments and Poster Session (Foyer, Kelvin Conference Centre)
18:00	Coaches depart accommodation for evening's entertainment
18:30-00:00	Third Evening Entertainment
	18:30-19:30 Sponsor's wine reception (St Andrews in the Square)
	19:30-21:00 Dinner (St Andrews in the Square)
	21:30 Early coach departs to return to accomodation
	21:30-00:00 Ceilidh (St Andrews in the Square)
00:00	Coaches depart to return to accommodation

9.4 Thursday 23rd March

07:30-08:30	Breakfast (Dining Room, Wolfson Hall)
08:30-	Delegates depart

10 Organised Entertainment

There will be a series of events and activities organised during the conference to help you relax after the sessions and allow you to socialise with other delegates while enjoying your Glasgow experience.

10.1 Monday 20th March

On the evening of Monday 20th there will be the traditional pub quiz at Wolfson Hall. Alternative entertainment will also be provided, e.g. a movie screening.

10.2 Tuesday 21st March

On Tuesday 21st there will be a choice of activities. There will be a break after the afternoon session and you are welcome to wander around the university or explore Glasgow's West End. This year the conference will incorporate an RSS Glasgow Local Group Meeting which will take place before dinner.

During the afternoon break, a football match will take place in the late afternoon. A coach will transport players to the pitch and players will be able to change into their sports kit at the accommodation before play. A coach will then transport players to dinner. Those people wishing to play football will, however, miss the special invited RSS guest speaker.

Dinner will be in Gong restaurant on Vinicombe Street, just off Byres Road. After dinner there will be a whisky tasting session in the Glasgow University Union. Alternatively you can relax in the Playing Fields (Glasgow University Union's bar) or watch a movie at the local cinema.

If none of these scheduled activities appeals to you then you are also free to explore the night life of Glasgow's West End or make your way into the city centre on the Clockwork Orange (Glasgow's circular subway).

Ashton Lane in the West End is a popular cobble stone street with many of Glasgow's trendiest bars. You will also find a cinema (http://www.grosvenorcinema.co.uk/) there.

There will be a limited number of people who can participate in each activity and you were required to specify a chosen activity when you registered. Places for the evening's social activities were allocated on a first come first serve basis.

10.3 Wednesday 22nd March

On Wednesday 22nd there will be a wine reception with a display from the sponsors to learn more about some possible career opportunities. The conference dinner will follow and both these events will take place in a beautifully restored grand cathedral at St Andrews in the Square. As is tradition in Scotland, after the dinner there will be a ceilidh (Scottish dancing).

11 Instructions

11.1 For Chairs

- Please arrive in the appropriate seminar room five minutes before the start of each session. It is important to familiarise yourself with the set up of the digital projector for those presenters intending to give an electronic presentation.
- There will be packs left in each of the seminar rooms. Do not remove the packs or any of their contents from the seminar room. If you think something might be missing from the pack, please let us know.
- Before and between the talks, please display the schedule information for the ongoing session.
- You should clearly introduce yourself and each speaker in turn.
- It is very important that we stick to the schedule. Therefore please start the session on time, use the time remaining cards, and make sure that questions are not allowed to delay the rest of the session.
- If a speaker fails to show for their scheduled talk please advise the audience to attend a talk in an alternative seminar room. Do not move the next talk forward.
- After each talk, thank the speaker, encourage applause and then open the floor to questions.

11.2 For those Doing Talks

- Each lecture room will contain a laptop, data projector, overhead projector, white board/black board, overhead and white board pens and chalk.
- Please arrive at least 5 minutes before the start of your session, introduce yourself to the chair of the session and put your presentation onto the laptop (if necessary).
- Talks are strictly 15 minutes plus 5 minutes for questions. Unfortunately, anyone going over this time will be asked to stop by the chair.
- Your chair will let you know when you have 5 minutes and then 1 minute remaining for your presentation.

11.3 For those Displaying Posters

- The poster session will be held in the Foyer of the Kelvin Conference Centre.
- Posters will be displayed at 4pm on the third day of the conference.
- Please submit your posters upon registration on Monday 20th March.
- Posters will be erected by the conference organisers.
- During the session, please be close to your poster at all times to field any questions from interested delegates.
- Please ensure that your poster is removed promptly after the end of the allocated session.

12 Talks schedule in chronological order

12.1 Monday 20th March

Session Plenary Session
Start Time: 16:15

Room	Speaker	Title	Page
KCC	Mike Titterington	Pearson the Elder - A Statistical Pioneer	39

Session Plenary Session
Start Time: 17:00

Room	Speaker	Title	Page
KCC	Stephen Senn	Has modelling killed randomisation inference?	39

12.2 Tuesday 21st March

Session 1
Start Time: 09:25

Room	Speaker	Title	Page
203	Kelly Handley	Classification of Proteomic Cell-line Data	41
214	Andrew Rose	Bayesian Designs for Model Discrimination	45
325	Katharina Gruenberg	Homicide in England and Wales	48
326	Cheran Vithanage	Approximate inference for hidden Markov models using iterative belief selection	51

Session 1
Start Time: 09:50

Room	Speaker	Title	Page
203	Malcolm Price	Modelling Treatment effects in discrete time Markov data from Clinical Trials of Asthma	41
214	Shan Lin	Simultaneous confidence bands for linear regression models with restricted predictor variables	45
325	Theodore Papamarkou	Statistical dependence of chaotic maps	48
326	Alicia Huntriss	Bayesian Calibration in Luminescence Dating	51

Session 1
Start Time: 10:15

Room	Speaker	Title	Page
203	Oarabile Molaodi	Comparing Measles maps of Scotland over time	42
214	Gavin Hardman	Bayesian methods for the design of inspection routines for large systems	46
325	Andrew Golightly	Biochemical Networks: Stochastic Simulation and Bayesian Inference	49
326	Ana-Maria Magdalina	A Bayesian Approach to Regime Change in Linear Regression	52

Session 1
Start Time: 10:40

Room	Speaker	Title	Page
203	Robert Mastrodomenico	A Comparison of Single Locus and Haplotype Association based methods for Gene Mapping	43
214	Oleg Volkov	Optimization of Disease Screening Ages	47
325	Tarek Medkour	A variance equality test for two correlated complex gaussian variables and its application to spectral power comparison	50
326	Despoina Vasileiou	Bayesian Analysis of human Isochores	53

Session 2
Start Time: 11:25

Room	Speaker	Title	Page
203	Elizabeth Williamson	Propensity scores and M-estimators	54
214	Limin Wang	Maximum Entropy Sampling of Gaussian Processes	57
325	Alex Goodwin	Wavelet regression for classification	60
326	Ali Al-Alwan	Bayesian Discrimination for Pesticide Contamination	62

Session 2
Start Time: 11:50

Room	Speaker	Title	Page
203	Andrew Thomson	Measures of between cluster variability in cluster randomised trials with binary outcomes	54
214	Zhen Liu	Online Inference for Multiple Changepoints Problem	57
325	Jeffrey Samuel	Data Modelling for Combinatorial Materials Development for Lithium Batteries	60
326	Demetris Lamnisos	Bayesian Variable Selection in Discrimination Problems	62

Session 2
Start Time: 12:15

Room	Speaker	Title	Page
203	Anastasia Lykou	Predictive scoring using categorised variables to mediate response and explanatory variables	55
214	Solange Correa	Estimation of Multilevel Model Parameters from Complex Survey Data	58
325	Patrick Rubin-Delanchy	Simulating Complex-valued Time Series	61
326	Chris Sherlock	0.234 and all that	63

Session 2
Start Time: 12:40

Room	Speaker	Title	Page
203	Emma O'Connor	Statistical Issues in Medical Imaging	56
214	Leonardo Trujillo	Benchmarking, Contemporaneous Disaggregation and Reconciliation Using State Space Models	58
325	Billy Wu	Graphical Modelling for Time Series	61
326	Gundula Behrens	Measuring and Improving Accuracy in MCMC	63

Session 3
Start Time: 14:15

Room	Speaker	Title	Page
203	Miland Joshi	Decision theory in diagnostic testing (provisional)	64
214	Matthew Killeya	Want an interesting and well paid job?	68
325	Elaine Wilson	Is it better to give information, receive it or be ignorant in a two-player game?	71
326	Hassan Khalil	Efficient Particle Filters for Ergodic Stationary Processes	74

Session 3
Start Time: 14:40

Room	Speaker	Title	Page
203	Elisa Bonvini	Investigating statistical methods for analysing randomised trials in Type 2 diabetes	64
214	Lucas Carbonaro	Studies of financial time series analysis	68
325	Mercedes Andrade Bejarano	Monthly Average Temperature Modelling	71
326	Richard Everitt	Graphical Models for the Internet	74

Session 3
Start Time: 15:05

Room	Speaker	Title	Page
203	Raquel Granell	Longitudinal analysis of a birth cohort in determination of discrete phenotypes of asthma	65
214	John Fry	The mathematics of financial crashes	69
325	Duncan Lee	Time-varying coefficient models for the analysis of air pollution and health outcome data	72
326	Amber Tomas	Measures of Between Cluster Variability in Cluster Randomised Trials with Binary Outcomes	75

Session 3
Start Time: 15:30

Room	Speaker	Title	Page
203	Guy Freeman	Bayesian Value of Information Methods in Clinical Research	66
214	Pauline Sculli	Counterparty Default Risk in Affine Processes with Jump Decay	69
325	Emma Eastoe	An Extreme Value Analysis of Air Pollution in the UK	73
326	Chiara Mazzetta	Bayesian Dynamic Generalised Linear Models to monitor population trends in ecology	75

Session RSS Guest Speaker
Start Time: 17:15

Room	Speaker	Title	Page
203	Tim Holt	Estimating International Migration	77

12.3 Wednesday 22nd March

Session 4
Start Time: 09:25

Room	Speaker	Title	Page
203	Yugang Jia	Gaussian approximation based mixture reduction for near optimal symbol detection in MIMO systems	78
214	Bander Al-Zahrani	Properties of a class of life distribution involving half-normal mixtures	81
325	Maria Costa	Penalties in Nonproportional Regression Survival Models	83
326	Graciela Muniz Terrera	How can we best estimate missing observations in longitudinal studies of ageing?	86

Session 4
Start Time: 09:50

Room	Speaker	Title	Page
203	Dan Bailey	Feature Selection using Discriminant Analysis	78
214	Michailina Siakalli	The moments of a Levy process	81
325	Hayley Jones	Automatic identification of 'interesting' longitudinal series of counts: application to monitoring MRSA rates in NHS Trusts	83
326	Andrew Titman	Hidden Markov models with time dependent misclassification	86

Session 4
Start Time: 10:15

Room	Speaker	Title	Page
203	Chris Brignell	Shape and symmetry analysis of the human brain	79
214	Eleni Bakra	The simplex sampler	82
325	Kerry Dwan	Within Study Selective Reporting in Meta Analysis	84
326	Vilda Purutcuoglu	Stochastic modelling of the MAPK signalling pathway	87

Session 4
Start Time: 10:40

Room	Speaker	Title	Page
203	Jinghao Xue	Generative vs Discriminative Modelling	79
214	Erik Baurdoux	Game options	82
325	Lauren Rodgers	Effect of Missing Values on the Analysis of the AB/ BA Crossover Trial	85
326	Anne Presanis	Bayesian evidence synthesis for estimating HIV prevalence and incidence	88

Session 5
Start Time: 11:25

Room	Speaker	Title	Page
203	Naventhan Mahadevan	Statistical Visualization of Wireless Sensor Networks	89
214	Simon Spencer	Stochastic Epidemic Models for Emerging Diseases Incorporating Interventions in Real-time	92
325	Michalis Kolossiatis	Bayesian Nonparametric Modelling of Spatial Data	95
326	Patrick Ho	Mixture Cure Models and Misspecification	98

Session 5
Start Time: 11:50

Room	Speaker	Title	Page
203	Brian Sullivan	Mixed Interaction Modelling with applications	89
214	David Jenkinson	Eliciting People's Probabilities Using Computer Software	92
325	Judy Tang	Nonparametric Smoothing Methods and the Multi-Resolution Criterion	95
326	Hathaikan Chootrakool	Meta-analysis of Multi-arm Trial	98

Session 5
Start Time: 12:15

Room	Speaker	Title	Page
203	Kim Evans	Looking for similar sites on the surfaces of proteins	90
214	Brett Houlding	Sequential Decision Making with Adaptive Utilities	93
325	Irene Kaimi	Spatial Modeling of Line Transect Survey Data	96
326	Guangquan Li	A stochastic model of carcinogenesis incorporating various types of genetic instability	99

Session 5
Start Time: 12:40

Room	Speaker	Title	Page
203	Ioanna Manolopoulou	MCMC methods on estimating the genetic and geographical history of individuals	91
214	I-Ding Wu	Optimal customers management strategies	94
325	Mark Kelly	Derivation of synthetic area boundaries and their use in the analysis of mental health	96
326	Richard Jacques	Statistical Analysis of High Content Screening Data	99

Session 6
Start Time: 14:15

Room	Speaker	Title	Page
203	Efthimios Motakis	Variance Stabilization for One-colour Microarray data	101
214	Theresa Cain	Bayesian inference for health state utilities using pair-wise comparison data	104
325	Keith Harris	Statistical Modelling and Inference for Radio-Tracking	106
326	Isabel Sassoon	Using SAS Enterprise Guide for data manipulation, analysis and presentation of results	109

Session 6
Start Time: 14:40

Room	Speaker	Title	Page
203	Fay Hosking	Statistical Methodology for genomewide association scans	102
214	Gillian Golden	Gambling and Financial Markets	104
325	Paul Birrell	Modelling the Survival of Juvenile Salmon on a Northern Scottish Rivers	106
326	Peter Colman	The Role of Statisticians at Pfizer	109

Session 6
Start Time: 15:05

Room	Speaker	Title	Page
203	Ximin Zhu	Bayesian analysis of designed microarray experiment	102
214	Tso-Jung Yen	Computational methods for extended jump processes with applications to Bayesian nonparametric statis	105
325	Lynsey McColl	Bayesian techniques used to investigate issues of contemporaneity between archaeological and environmental records	107
326	Matt Coates	Towards Multivariate Methods in Industrial Applications	109

13 Chairs in chronological order

Monday 20th March

Chair Person	Room	Start Time	Page
Professor Marian Scott	KCC	16:00	39

Tuesday 21st March

Chair Person	Room	Start Time	Page
Robert Mastrodomenico	203	09:25	41
Leonardo Trujillo	214	09:25	45
Alex Goodwin	325	09:25	48
Despoina Vasileiou	326	09:25	51
Emma O'Connor	203	11:25	54
Limin Wang	214	11:25	57
Billy Wu	325	11:25	60
Ana-Maria Magdalina	326	11:25	62
Miland Joshi	203	14:15	64
Erik Baurdoux	214	14:15	68
Duncan Lee	325	14:15	71
Richard Everitt	326	14:15	74
Dr. Ben Torsney	203	17:15	77

Wednesday 22nd March

Chair Person	Room	Start Time	Page
Jinghao Xue	203	09:25	78
Michailina Siakalli	214	09:25	81
Andrew Thomson	325	09:25	83
Brett Houlding	326	09:25	86
Chris Brignell	203	11:25	89
David Jenkinson	214	11:25	92
Mark Kelly	325	11:25	95
Michael Fahey	326	11:25	98
Chiara Mazzetta	203	14:15	101
Cian Reynolds	214	14:15	104
Keith Harris	325	14:15	106
Neil Henderson	326	14:15	109

14 Talk Abstracts by Session

14.1 Monday 20th March

14.1.1 Plenary Session

Session Room KCC Chair: Professor Marian Scott

Start time 16:15

Pearson the Elder - A Statistical Pioneer

Prof. Mike Titterington
University of Glasgow, UK

Keywords: *Karl Pearson; History of Statistics*

Pearson is perhaps most widely known nowadays for his eponymous goodness-of-fit statistic and the sample correlation coefficient, but the story of his life reveals him as a man of many interests and activities, typically pursued with great passion.

Start time 17:00

Has modelling killed randomisation inference?

Prof. Stephen Senn
University of Glasgow, UK

Keywords: *randomisation; additivity; unit-treatment interaction; general balance; history of statistics; clinical trials*

The Fisherian approach to analysing designed experiments reached its apotheosis in John Nelder's 1965 theory of generalised balance. The analysis of any of a very wide class of experiments would follow automatically from three inputs: a statement describing the block structure of the experimental material, a statement describing the treatment structure to be applied to that material and a data matrix of treatment and block factors, together with the outcome, reflecting the experiment that was run. This approach is reflected even today in the way that GenStat analyses experiments and the implication is that there is only one correct way to analyse a given randomised experiment. The modelling approach to inference which, somewhat ironically, also received a great impetus from work of Nelder's (in particular his paper

of 1972 with Wedderburn) has now all but displaced the randomisation approach. An interesting example of the clash of the two traditions is given by Yates's 1982 criticism in Biometrics of split-plot analyses of repeated measures data. A further illustration is given by differences between the R/S-Plus and GenStat communities to analysing Yates's (1935) split-plot oats data included as a standard data-set part of these packages.

A difficulty with the randomization tradition is that the class of statistical problems to which it can be applied is far narrower than that which may be tackled by the modelling approach and that hardly any statisticians would be prepared to concede that statistical inference is only applicable to this restricted class. This raises the question as to whether the fact that certain particular models can be justified in terms of design and randomisation is not a fundamental fact of statistical inference but merely an interesting coincidence and whether, therefore, the randomisation approach has not simply been a misleading historical distraction. In this talk I will consider whether there can be any reconciliation between these two approaches and propose (tentatively) that it may be found by supposing that the model justifies the design rather than vice versa.

References

[1] Nelder, J.A. The analysis of randomised experiments with orthogonal block structure I. Block structure and the null analysis of variance, Proceedings of the Royal Society of London. Series A 1965; 283: 147-162.

[2] Nelder, J.A. The analysis of randomised experiments with orthogonal block structure II. Treatment structure and the general analysis of variance, Proceedings of the Royal Society of London. Series A 1965; 283: 163-178.

[3] Nelder, J.A. & Wedderburn, R.W.M. Generalized Linear Models, Journal of the Royal Statistical Society A 1972; 132: 107-120.

[4] Yates, F. Complex experiments (with discussion). Supplement to the Journal of the Royal Statistical Society, 2: 181-247.

[5] Yates, F. Reader Reaction - Regression-Models for Repeated Measurements, Biometrics 1982; 38: 850-853.

14.2 Tuesday 21st March

14.2.1 Session 1a

Session Room 203 Chair: Robert Mastrodomenico

Start time 09:25

Classification of Proteomic Cell-line Data

Kelly Handley
University of Nottingham, UK

The development of surface-enhanced laser desorption/ionisation time-of-flight (SELDI-TOF) mass spectrometry has enabled the analysis of complex protein samples over a large range of molecular weights. This presentation concerns the correct classification of breast cancer and melanoma mass spectra obtained through the SELDI-TOF method.

Due to the high-dimensional nature of the data, methods need to be employed to reduce the number of variables used in classification. A single mass spectrum in this study consists of around 14,000 datapoints, with each datapoint comprising a relative intensity of proteins at a particular mass over charge (m/z) value. A deterministic peak finding algorithm is described which models the spectra as a series of Gaussian kernels. We model common peak locations across spectra from the same dataset but use differing heights to indicate the relative intensity at each location. The use of this algorithm reduces the size of the dataset from 14,000 to around 50 variables per spectrum.

Various classification methods including linear discriminant analysis and support vector machines are used on a training subset of the data to predict the class of a test subset. To determine the effectiveness of the peak finding method we use principal components analysis to classify the complete spectra and compare with the results we obtain through a cross validation experiment.

Start time 09:50

Modelling Treatment effects in discrete time Markov data from Clinical Trials of Asthma

Malcolm Price, Dr. Nicky Welton and Prof. Tony Ades
University of Bristol, Department of Social Medicine, UK

Keywords: *Markov, Meta-analysis, Multiple treatment Comparison, Asthma*

Purpose: To develop a structured approach to compare the relative treatment effects between the different arms of Randomised Controlled Trials (RCTs) where data are reported in last observation carried forward format, and give information on state transitions in a discrete-time Markov Model. The methods are then extended to allow a Multiple Treatment Comparison (MTC) meta-analysis.

Methods: We analysed data taken from five RCTs comparing eight different treatments for asthma. Here data are collected for subjects moving between four discrete health states with nine possible transitions. The state transition rates are modelled using log-linear regressions and these transition rates are converted to transition probabilities using Kolmogorovs Forward Equations. Bayesian inferential techniques are used to generate posterior distributions for these transition probabilities. We sample from these distributions using Markov Chain Monte Carlo simulation techniques via WinBUGS and WBDiff. We considered modelling differing numbers of transitions, structured and unstructured baselines and numerous plausible treatment structures including: no treatment effect, all forward transitions only, all backward transitions only, transitions to treatment failure only, transitions from successful treatment only, and combinations of the above. /newpar Results: The preferred (best fitting) model included an unstructured (saturated) baseline and a structured treatment effect. The treatment model found to have the best fit (lowest Dic) was the all backwards (health improvement) transitions only structure. /newpar Conclusions: Modelling transition rates, rather than probabilities, allows the estimation of relative treatment effects in a generalisable form, Meta-analysis with MTC can then readily be performed on the data.

Start time 10:15

COMPARING MEASLES MAPS OF SCOTLAND OVER TIME

Oarabile Ruth Molaodi, Dr. Alison Gray and Prof. Chris Robertson
University of Strathclyde, UK

Keywords: *Bayesian modelling, Disease maps, Log-normal model, Measles data, Relative risk*

This work uses data obtained from Health Protection Scotland (HPS). This consists of two sets of measles data, for pre-school children and primary 1 and 2 children for the 56 districts of Scotland for each year in the period 1999-2004. To compare measles over time, Bayesian modelling of relative risks is used to smooth the relative risk and obtain Bayesian estimates. The model used is the log-normal model proposed by Besag *et al.* [1] (usually referred to as the Besag, York and Mollié (BYM) model). For each data set, the log-normal model was fitted for each year using the package WinBUGS, and measles maps produced based on Bayesian estimates, enabling visual comparison. The plots of parameters versus time were also produced

and used to compare the measles data over time. It was found that the results for the pre-school data and primary 1 and 2 school data were different. More formal methods for comparison of disease maps are also being considered.

References

[1] Besag J, York J and Mollié A (1991), Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**, 1-59.

Start time 10:40

A Comparison of Single Locus and Haplotype Association based methods for Gene Mapping

Robert Mastrodomenico[1], Karen Ayres[1] and Andrew Morris[2]

[1] *School of Applied Statistics, The University of Reading, UK*
[2] *Welcome Trust Centre, The University of Oxford, UK*

The function of many genes is now reasonably well understood due to the advances in functional and gene expression studies. Consequently biologists are able to provide candidate genes that may be responsible for a disease of interest. Association studies focus on genotyping markers (usually SNPs) in samples of unrelated affected cases and controls.

It has been suggested that typing haplotypes would provide greater power in association studies. However some consider the gain in power insufficient relative to the extra cost of typing these haplotypes. Others suggest that single locus methods provide power equivalent to haplotype based methods.

Here we compare the power of four association based methods for localising disease susceptibility loci.

- A single locus chi-square test of allele frequencies for cases and controls.
- Main effects logistic regression using genotypes.
- Main effects logistic regression using haplotypes resolved from genotypes via and EM algorithm.

- Main effects logistic regression using actual haplotypes.

Data are simulated under varying disease models, and the effect of marker spacing , number of markers, sample size and disease allele frequency on each of the methods is investigated.

14.2.2 Session 1b

Session Room 214 Chair: Leonardo Trujillo

Start time 09:25

BAYESIAN DESIGN OF EXPERIMENTS FOR MODEL DISCRIMINATION

Andrew Rose
University of Southampton, UK

Keywords: *Design of Experiments, Bayesian Methods*

We take a decision theoretic approach to model selection, using a loss function dependent on the true model and the model selected. The model selected is hence that with the lowest expected loss, given the posterior model probabilities. A criterion for experimental design is introduced, equal to the average expected loss of models selected. We use simulation to evaluate this criterion, and a Modified Federov Exchange Algorithm to search for designs. The criterion is applied to several situations where model selection is of interest. In simple cases, the designs produced may be similar to those chosen by other design criteria, such as orthogonal designs. More complicated situations lead to less usual designs with aliasing of effects common to all models, or use of a combination of prior information and aliasing to distinguish between models.

Start time 09:50

SIMULTANEOUS CONFIDENCE BANDS FOR LINEAR REGRESSION MODELS WITH RESTRICTED PREDICTOR VARIABLES

Shan Lin
University of Southampton, UK

Keywords: *Simultaneous confidence band, Tube formula*

This talk first introduces simultaneous confidence bands, basically including Scheffé-type confidence bands, constant-width confidence bands and some other shaped easy-to-construct confidence bands, for normal-error linear regression models with predictor variables restricted in intervals to achieve the purpose of solving real

problems. And then some different methods are presented to show how to construct simultaneous confidence bands for both simple and multiple regression functions with constraints on predictor variables. These methods include exact methods for simple linear regression in Uusipaikka (1983) and Wynn & Bloomfield (1971), Naiman's method of approximating the coverage probability of conservative simultaneous confidence bands, Sun's method which is an approximation to the tube formula, and simulation-based exact method, respectively. A comparison of the involved methods is discussed after that by comparing the results computed by Matlab programmes correspondingly for polynomial regression with various orders and multiple regression. Reasonable conclusion follows finally.

Start time 10:15

EFFICIENT BAYESIAN INSPECTION DESIGN FOR LARGE INDUSTRIAL SYSTEMS

Gavin Hardman[1], Michael Goldstein[1] and Philip Jonathan[2]

[1] *University of Durham, UK*
[2] *Shell Research, Chester, UK*

Identifying optimal inspection and maintenance routines is a challenging problem faced by many industries. An inspection and maintenance routine can be considered optimal for a fixed quantity of inspection resource if it provides the best quality of information for this budget. Our particular example is that of inspecting petrochemical processing systems for corrosion damage. Our aim is to provide an efficient method for choosing between inspection designs, and thus recommending optimal inspection and maintenance routines.

The problem can be split into a modelling problem and a design problem. We adopt a Bayesian approach which allows both the incorporation of expert information into our modelling and also the development of appropriate decision theoretic criteria for design identification commensurate with the inspection objectives. We develop models for the behaviour of corrosion within a system, and also for the inspection procedures. We are intending to deal with systems consisting of very large numbers of varying components, and this constrains our model. We use a simple model for each component which has enough complexity to allow us to take into account some covariate information, but remains sufficiently simple so as not become computationally intractable for higher dimensions. The design problem tackles how we choose between inspection and maintenance routines. Typically, evaluating the 'value' of a single inspection design can be computationally expensive. We outline a rational criterion for design comparison and a method for comparing many designs involving minimal computation.

Start time 10:40

Optimization of Disease Screening Ages

Oleg Volkov
Queen Mary, University of London, UK

Keywords: *Disease Screening, Bayesian Approach, Optimal Design*

The objective of medical screening is to detect an asymptomatic disease while the disease is most curable. Although screening has reduced mortality – notably in some cancers – many current tests are invasive for the individuals and costly for the health care provider. Therefore, a key design issue in population screening is determination of the optimal screening frequency. As shown by the recent change in the recommended frequency of Pap smears for women aged 45 – 64, from annual to once every three or five years, this task can be difficult.

In this talk I focus on designing pilot studies which help to determine optimal screening ages. The goal of such studies is to estimate the probability of a true positive screen as a function of age. To build the experimental model I assume different lags between the incidence of the clinical disease and the probability of a true positive screen. A Bayesian D-optimal design – with ages and numbers of study subjects – is found for parameter priors that are based on the incidence data. The practical benefits of the proposed pilot studies are demonstrated using examples of existing screening programmes.

14.2.3 Session 1c

Session Room 325 Chair: Alex Goodwin

Start time 09:25

THE HOMICIDE INDEX OF ENGLAND AND WALES

Katharina Gruenberg and Prof. Brian Franics
Lancaster University, UK

Keywords: *latent class*

According to the Bible, apart from deception, homicide is the oldest crime in the world and has since intrigued the world. Much research has been done since but so far the philosopher's stone has not been found. This might be due to the fact that all research so far has been theoretical, and posterior, everything else would be unethical. Few research has been backed up by statistics, all of which have been confirmatory analysis though. So that maybe one or the other researcher has been tempted to find exactly what he wanted to see. This dissertation takes a different approach. We will attempt to be data detectives and let the data speak for itself. That means, that we will undertake an analysis using the correct statistical without any prior assumptions on the outcome, for which current methods have to be adapted and extended. The data at hand has some features that need to be kept in mind throughout the analysis. First of all, although we have data that has been collected over a period of time, it is not time series data, as the observations here are not repeated observations. Each of the victim has only been killed once, it is hence comparable with situations in which threshold levels are passed once, yet we are interested in the development over time. Two more issues with this data, that occur frequently in all applied data bases are missing information and imbalancedness, due to the presence of frequent and rare events. An example here for missing information would be, the lack of knowledge of the murder weapon while the victim, the suspect and the circumstance are known, whereas an example for imbalancedness is the varying frequency of the weapon, sharp instruments are used 200 times as often as explosives. We will propose several ways forward and determine the most suitable one.

Start time 09:50

STATISTICAL DEPENDENCE OF CHAOTIC MAPS

Theodore Papamarkou
University of Warwick, UK

This talk is mainly concerned with the statistical dependence of chaotic maps and several associated implications in the communications field. The linear and adjusted (for the mean) quadratic autocorrelation functions (ACF) are employed as the main measures of statistical dependence of the maps under consideration.

We start from some established results such as the ACF of the tent and Bernoulli shift maps and then generalize to the family of fully stretching piece-wise linear maps. The linear ACF of this collection of maps is known; we further provide our result, the adjusted quadratic ACF of these linear maps.

From the communications point of view, the less the first lag of the adjusted quadratic ACF of a chaotic map the better; we invoke the concept of bit energy to justify our argument. The family of fully stretching piece-wise linear maps has the disadvantage that the lag 1 adjusted quadratic ACF is non-negative and therefore far away from the minimum value -1. We present the first-order circular map, which approximates better the required statistical dependence compared to the above-mentioned linear maps. Furthermore, we extend the circular map to the new family of squared circle maps, which we recently introduced, and comment on the benefits in terms of the reduction of the lag 1 adjusted quadratic ACF.

As a side-effect of the intention to reduce the lag 1 quadratic dependence of the maps involved, other statistical aspects of them have been illuminated; explicit expressions for the their invariant densities are provided.

Start time 10:15

Biochemical Networks: Stochastic Simulation and Bayesian Inference

Andrew Golightly and Dr. Darren J. Wilkinson
University of Newcastle, UK

Keywords: *Bayesian inference, Nonlinear diffusion, Markov chain Monte Carlo*

Traditionally, the time evolution of a biochemical network is described by a set of coupled differential equations derived using the law of mass action and the concentrations of each species. This widely used approach, however, assumes that the system is both continuous and deterministic. In reality, chemical reactions are intrinsically stochastic and occur as discrete events resulting from random molecular collisions. Although relatively little work has addressed the stochasticity of biochemical networks, it is clear that many important intra-cellular processes such as signal transduction and gene expression can only be effectively described by stochastic processes.

There are three common types of stochastic Markov process models used to simulate biochemical networks: 1) discrete models commonly solved by the Gillespie

algorithm, 2) diffusion or stochastic differential equation (SDE) models in which the variables are approximated as continuous and a white noise term models stochastic behaviour and 3) hybrid models where some chemical species are treated as discrete and others are treated with a continuous approximation. This talk will focus on the first two methods, paying particular attention to the diffusion approximation. Although the latter approach is often inadequate for simulation purposes, it appears to be quite satisfactory when used as the basis of a Bayesian inference algorithm.

Start time 10:40

A VARIANCE EQUALITY TEST FOR TWO CORRELATED COMPLEX GAUSSIAN VARIABLES AND ITS APPLICATION TO SPECTRAL POWER COMPARISON

Tarek Medkour and Andrew T. Walden
Department of Mathematics, Imperial College London, UK

Keywords: *complex bivariate gaussian distribution, hypothesis test, power of test, spectrum estimation*

Complex-valued Gaussian distributions occur very frequently in signal processing. We derive a simple statistic, independent of any complex-valued correlation, for testing for the equality of variances using a sample drawn from such a bivariate distribution. The percentage points of the distribution are easy to compute. The power of the test is determined and is shown to be high even for small sample sizes when the variables are highly correlated. The new test is used to determine whether the spectral power associated with an ultra low frequency wave in the solar magnetic field is equal at two different observing spacecraft.

14.2.4 Session 1d

Session Room 326 Chair: Despoina Vasileiou

Start time 09:25

APPROXIMATE INFERENCE FOR HIDDEN MARKOV MODELS USING ITERATIVE BELIEF SELECTION

C. M. Vithanage
Department of Mathematics, University of Bristol, UK.

Keywords: *Hidden Markov models, Approximate inference, expectation propagation*

The inferential task of computing the marginal posterior probability mass functions of state variables and pairs of consecutive state variables of a hidden Markov model (HMM) is considered. This can be exactly and efficiently performed using a message passing scheme such as the BCJR algorithm. We present a novel iterative reduced complexity variation of the BCJR algorithm which uses reduced support approximations for the forward and backward messages as in the M-BCJR algorithm. Forward/backward message computation is based on the concept of Expectation Propagation, which results in an algorithm similar to the M-BCJR algorithm with the active state selection criterion being changed from the filtered distribution of state variables to beliefs of state variables. By allowing possibly different supports for the forward and backward messages we derive identical forward and backward recursions which can be iterated. Simulation results of application for trellis based equalization of a wireless communication system confirm the improved performance over the M-BCJR algorithm.

Start time 09:50

BAYESIAN CALIBRATION IN LUMINESCENCE DATING

Alicia Huntriss
Durham University, UK

Keywords: *Luminescence Dating, Palaeodose, Bayesian Methods*

Optically stimulated luminescence dating is an archaeological technique which can be used for constructing sediment chronologies. It exploits structural properties of crystals which accumulate energy over time by trapping free electrons from ionising radiation. These electrons are released by optical stimulation producing light

(the luminescence signal). The signal intensity indicates the amount of energy absorbed in the material, and the radiation dose that needs to be applied to produce this strength of signal is known as the palaeodose. The rate of radiation the sample has been exposed to, the annual dose, is also estimated. The age can then be calculated through

$$Age = \frac{Palaeodose}{Annual\ dose}$$

The energy obtained from exposure to heat or light will result in electrons being released from their traps. Hence luminescence dating can measure the time span since the electrons were last detrapped. This can occur, for example, by heat when a clay pot is fired or by sunlight when sediment is deposited and subsequently buried. This is known as resetting the luminescence clock.

Luminescence is an important method of dating, although it can be prone to a large degree of uncertainty. Developments in both experimental procedures and knowledge of the process have allowed these to be reduced, but there is still scope to describe the variance of the estimated age in a more meaningful manner. Bayesian methods have long been established in the carbon dating arena, and there have been some initial developments in applying such an approach to luminescence dating. A routine method of palaeodose evaluation involves a back-interpolation from a least squares regression line. An alternative Bayesian calibration model is proposed, this is being developed with a view to be able to produce estimates of palaeodose which combine all of the relevant sources of information in an effective manner.

Start time 10:15

A Bayesian Approach to Regime Change in Linear Regression

Ana-Maria Magdalina and Prof. Don Barry
University of Limerick, Ireland

Keywords: *Change points*

The change points problem deals with a sequence of n consecutive independent observations $X_1, X_2, ..., X_n$ whose distribution is dependent on a changing parameter. The parameter changes at unknown times and at unknown values.

We are concerned with the case where the observations are assumed to have a normal distribution $Xi \sim N(\theta_i, \sigma^2)$, for $i = 1..n$, the mean θ_i being the parameter which changes at unknown times while σ^2, the variation around the different values of the

mean, remains constant. Given only the observations X_i, we estimate when these changes occur in the underlying parameter θ_i, as well as its values.

Further, we extend the change point problem in the context of simple linear regression model. In linear regression analysis it is sometimes the case that the relationship between the two variables changes. We want to identify when these changes occur and estimate the changing values of the regression parameters. For a Bayesian approach we must specify sensible prior distributions for the unknown parameters of the problem.

Start time 10:40

BAYESIAN ANALYSIS OF HUMAN ISOCHORES

Paul Fearnhead and Despina Vasileiou
Lancaster University, UK

Keywords: *Bayesian analysis,isochores,direct simulation*

The human genome consists of long DNA segments which are homogeneous in G+C content, and are called isochores. There is currently much interest in detecting the isochores in the human and other organism's genomes as G+C content has been shown to be correlated with various biological processes, such as mutation and recombination, and understanding these correlations and their causes is of great scientific interest. Existing statistical methods, operating in a classical framework, aim to segment the genome into isochores using for example binary segmentation. We present a new approach based on a Bayesian analysis, based on directly modelling features of the G+C content across the genome. Inference for our model is made by simulating directly from the posterior distribution of the number, position and CG content of the isochores. Simulation studies have shown that this Bayesian approach can signicantly improve the inference for how C+G content varies across the genome.

14.2.5 Session 2a

Session Room 203 Chair: Emma O'Connor

Start time 11:25

Propensity scores and M-estimators

Elizabeth Williamson
London School of Hygiene & Tropical Medicine, UK

Keywords: *Observational studies, confounding*

When treatment assignment is not randomised, systematic differences between the treated population and the non-treated population will usually be present. Unadjusted estimates of treatment effect may then be biased. Regression models are the standard method used to adjust for this confounding. An alternative solution is to use propensity-score methods. This presentation considers two methods of adjustment: stratification and matching by the propensity score. The variance of the resulting treatment effect estimators is currently unknown. M-estimation methods will be used to calculate variance formulae for these treatment effect estimators. Simulation studies will then be used to determine whether there are any situations in which propensity score methods are better able to adjust for confounding than traditional regression methods.

Start time 11:50

Measures of Between Cluster Variability in Cluster Randomised Trials with Binary Outcomes

Andrew Thomson and Richard Hayes
London School of Hygiene and Tropical Medicine, UK

Keywords: *Cluster Randomised Trials*

Cluster randomised trials are increasingly being used to evaluate the effectiveness of health-care interventions. It is well known that cluster randomised trials differ from individually randomised trials, as observations on individuals within clusters tend to be correlated due to the presence of between-cluster variability. Such correlations have to be accounted for in the design and analysis of these trials. When the

outcome is dichotomous, there are two common measures of between-cluster variability that have been used in sample size formulae: the coefficient of variation, k, and the intracluster correlation coefficient, ρ. For the design of cluster randomised trials it is assumed that one has an estimate of the between-cluster variability for the control arm of the study, either from expert opinion, previous published values or a pilot study. Two common formulae are used to calculate the sample size, one based on k, the other on ρ. The two formulae give different sample size estimates as they make different assumptions about the between-cluster variability in the intervention arm relative to that in the control arm. We describe the relationship between these two alternative measures of between-cluster variability and explore how assumptions of constant k or ρ across arms of trials relate to assumptions about the effect of the intervention. In particular, we wish to answer questions such as if the intervention effect is fixed across all clusters, what does this imply for the constancy of k or ρ across arms of the trial? We then give an example of the magnitude in differences one can expect using sample size calculations based on possible parameters one may observe. We also discuss the factors that influence the magnitude of the difference.

Start time 12:15

PREDICTIVE SCORING USING CATEGORISED VARIABLES TO MEDIATE RESPONSE AND EXPLANATORY VARIABLES

Anastasia Lykou and Joe Whittaker
Lancaster University, UK

Keywords: *cluster analysis, heuristic optimisation*

In our electronic age scientific investigations with statistical interest have seen a proliferation of the number of measured variables. Many investigations lead to the consideration of variables which naturally separate into response and explanatory subsets. A recent paper by Hand et al gives several examples. Statistical analysis requires some decisions as to the structure of the relationships between these subsets. Clearly this can be particular and complex. However in some applications a simple structure is to assume that the multiple response variables depend on the multiple covariates through one, or at most a few, additional mediating variables. Examples of this structure include both structural equation models with mediating latent variables, and canonical covariate analysis with intervening scores.

The aims of such analyses usually include simplification of the data structures with a broad understanding of the inter-relationships between variables, the provision of data summaries, clustered groups of units, and prediction from partial information of future values of the response and of the mediating variables. We compare

and contrast different metbods of dealing with this structure, that include building latent variable models with GLM type distributions, and taking deterministic linear combinations of response and explanatory variables. While these methods reflect different approaches to statistical modelling they all lead to the specification of objective functions to be optimised over certain parameter spaces subject to constraints.

Within this general framework a particular technique, of interest in the social and medical sciences, is to replace a variable by its categorised version, for instance, by recording only the interval rather than the value of a variable. This too leads to an optimisation problem, and we examine a variety of heuristics based on the analysis of single variables in turn

The paper compares methods of predicting mediating variables and extracting scores, in a variety of simulation experiments. Two applications are discussed in detail, a medical application concerning neonatal health, and the analysis of information held on creditcards.

Start time 12:40

Advancements In The Analysis Of Medical Images

Emma O'Connor
University of Sheffield, UK

Keywords: *MRI, Image Analysis, Multivariate*

Medical imaging provides a non-destructive method of direct investigation of effects of treatments on target tissues. This allows tissue to be examined on several occasions during the course of treatment, thus avoiding inter-individual variability. This presentation investigates the advances in methodology for assessing the statistical differences between images and hence the effectiveness (or otherwise) of the treatment. The particular study described here involved collection of three-dimensional images by MRI before and after treatment on individuals receiving one of a range of doses. Data extracted from the images were the separate voxel values of a parameter of interest. Statistical analysis focuses on the frequency distributions of these voxel values and looks at different methods of density estimation and subsequent multivariate techniques to assess the change post-treatment.

14.2.6 Session 2b

Session Room 214 Chair: Limin Wang

Start time 11:25

MAXIMUM ENTROPY SAMPLING

Limin Wang and Henry P. Wynn
London School of Economics, UK

Keywords: *Sampling, Gaussian Processes, Shannon Entropy*

Maximum entropy sampling (MES) was introduced as a choice of experiments by Shewry and Wynn (1987). Define $X_S = \{X_i\}_{i \in S}$, $X_s = \{X_i\}_{i \in s}$ and $X_{s^c} = \{X_i\}_{i \in s^c}$, where the index set $S = \{1, 2, \cdots, N\}$ represents the whole "population" and $s \subset S,\ |s| = n$ is a purposively chosen sample. A theorem of Shannon suggests that the maximum information can be obtained by maximizing Ent(X_s) with respect to the choice of s, where Ent(.) is Shannon Entropy. The current research is to extend these ideas to Gaussian processes by letting $n = |s|$ and $N = |S|$ tend to infinity in a special way that preserves the notion of MES. The theory of Karhunen-Loeve expansions is useful in carrying out this programme and a type of duality can be set up with the theory of D-optimality from classical optimal design. Brownian motion, the Brownian bridge and some non-standard Gaussian processes are studied. Numerical methods play an important role when no analytical solution exists.

Start time 11:50

ON-LINE INFERENCE FOR MULTIPLE CHANGEPOINT PROBLEMS

Paul Fearnhead and Zhen Liu
Lancaster University, UK

Keywords: *Particle Filtering; Direct simulation; Stratified sampling; Optimal resampling; Rejection control; DNA segme*

We propose an on-line algorithm for exact filtering of multiple changepoint problem. This algorithm enables simulation from the true joint posterior distribution of the number and position of the changepoints for a class of changepoint models. The computational cost of this exact algorithm is quadratic in the number of observations. We further show how resampling ideas from particle filters can be used

to reduce the computational cost to linear in the number of observations, at the expensed of introducing small errors; and propose two new, optimum resampling algorithms for this problem. One, a version of rejection control, allows the particle filter to automatically choose the number of particles required at each time-step. The new resampling algorithms substantially out-perform standard resampling algorithms on examples we consider; and we demonstrate how the resulting particle filter is practicable for segmentation of human CG content.

Start time 12:15

Estimation of Multilevel Model Parameters from Complex Survey Data

Solange Correa
University of Southampton, UK

Keywords: *Multilevel models, Survey Sampling, Resampling Methods*

Despite the availability of several methods that account for the effects of the sampling design when fitting multilevel models, these methods either restrict to simple models or that they do not account for the complexity of the data in a proper way. As a result, the use of these methods may induce nonnegligible bias, depending on the sample sizes, the sampling schemes and the magnitude and variability of the sampling weights.

In this talk I shall present some initial ideas for reducing the bias and improving the efficiency of probability weighted estimators in multilevel model fitting. A preliminary simulation study assesses the use of resampling techniques for bias correction.

Start time 12:40

Benchmarking, Contemporaneous Disaggregation and Reconciliation Using State Space Models

Leonardo Trujillo
Social Statistics Department, University of Southampton, UK

Keywords: *Benchmarking, State Space Models, Kalman Filtering*

Benchmarking is the process of constraining a high frequency (quarterly or monthly) indicator series to a lower frequency (annual, for example) benchmark series. The purpose is to create a new series that preserves, as far as possible, the short-term movements of the original series but being consistent with the benchmarks.

Other two problems related with benchmarking are: a) how to prepare tabular data when they are available only in an aggregated form and b) how to modify the values when the aggregated values do not correspond to the sum of the disaggregated values because, for instance, they come from different sources of information. An alternative solution for all these situations is presented using state space models. An application with simulated and real data is presented. The real data corresponds to ONS turnover data from two sources: MPI (Monthly Production Inquiry) and ABI (Annual Business Inquiry) from 1998-2003 in the UK.

14.2.7 Session 2c

Session Room 325 Chair: Billy Wu

Start time 11:25

WAVELET REGRESSION FOR CLASSIFICATION

Alex Goodwin
University of Leeds, UK

Keywords: *wavelets, regression*

Wavelets have proven to be highly effective at extracting frequency information from data. Their multi-scale nature enables the efficient description of both transient and long-term signals. Furthermore, only a small number of wavelet coefficients are needed to describe complicated signals and the wavelet transform is computationally efficient.

In problems where frequency properties are known to be important, it is proposed that a modelling approach which attempts to explain the response time-series in terms of a multi-scale wavelet representation of the explanatory time-series will be an improvement on standard regression techniques. The problem with classical regression is that differing frequency characteristics are not exploited and only confuse the model parameters. The proposed modelling method is presented with application to examples from seismology and tomography.

Start time 11:50

DATA MODELLING FOR COMBINATORIAL MATERIALS DEVELOPMENT FOR LITHIUM BATTERIES

Jeffrey J. Samuel
University of Southampton, UK

Keywords: *Combinatorial Chemistry, Time Series Analysis, Gibbs Sampling, Bayesian Inference*

Advances in technological devices have created a demand for efficient battery power sources. An example of which is the development of mobile phones over the past twenty-five years. If battery power sources had not been improved, mobile phones would still be of a similar size as those in the 1980s, and with the same short lifetime.

Key to enhancing the current generation of batteries is the understanding of their chemical compositions. To further their knowledge, chemists conduct experiments to test key hypothesis. The data from the experiments contain the time, current and potential, together with the chemical composition. It is usually presented as a voltammogram, which is a plot of current against potential. The aim is to produce an empirical model that predicts the voltammogram together with the attributes of interest. This talk will discuss the statistical approach and the model built so far. In particular, the focus will be on Bayesian time series analysis.

Start time 12:15

Simulating Complex-valued Time Series

Patrick Rubin-Delanchy
Imperial College London, UK

Start time 12:40

Graphical Modelling for Time Series

Billy Wu
London School of Economics, UK

14.2.8 Session 2d

Session Room 326 Chair: Ana-Maria Magdalina

Start time 11:25

FULLY BAYESIAN METHODS TO ANALYZE PESTICIDE CONTAMINATION

Ali Al-Alwan
University of Durham, UK

Keywords: *Pesticides, Discrimination, MCMC, WinBUGS.*

Two chemical properties of pesticides are thought to control environmental fate. These are the adsorption coefficient k_{oc} and soil half-life $t_{1/2}^{soil}$. This study aims to demonstrate the use of Bayesian methods in exploring whether or not it is possible to discriminate pesticides as leachers or non-leachers on the basis of the above mentioned chemical properties, when the monitored values of these properties are uncertain in the sense that there are a range of values reported for both k_{oc} and $t_{1/2}^{soil}$. The developed method uses a logistic regression model, and prior information was extracted from the United States Department of Agriculture (USDA) pesticide properties database, where a single value is reported for both k_{oc} and $t_{1/2}^{soil}$. The study starts by considering the main effects model and goes on to investigate it with the interaction term. The proposed Bayesian models are implemented by means of Markov Chain Monte Carlo (MCMC) simulation techniques via the freely available WinBUGS software.

Start time 11:50

BAYESIAN VARIABLE SELECTION IN DISCRIMINATION PROBLEMS.

Demetris Lamnisos
University of Warwick, UK

One problem in functional genomics is to use genes expression profiling data to discriminate between two or more classes. We want to both classify future samples and identify predictors that differ among the classes.Both can simultaneously be achieved by combining a multinomial probit model for classification with Bayesian variable selection method.The standard Bayesian approach makes use of mixture

priors for the model parameters and Markov Chain Monte Carlo techniques to select set of variables that characterize the different classes.My recent work is focused in the effect of priors for those model parameters.

Start time 12:15

0.234 AND THE METROPOLIS-HASTINGS RANDOM WALK

Chris Sherlock
Lancaster University, Lancaster, UK

Keywords: *MCMC, Metropolis-Hastings, optimal scaling*

Existing results on optimal scaling for the Metropolis-Hastings random walk mainly cover independent identically (up to a possible re-scaling) distributed targets with Gaussian proposals, and are obtained through a limiting diffusion approximation to the random walk as dimension $d \to \infty$. Results to be detailed are based around exact closed forms for the expected acceptance rate and mean square jump distance, which apply whatever the dimension of the problem. They are obtained using a simpler approach and apply to unimodal spherically and elliptically symmetric targets. The method provides an intuition as to when the limiting acceptance rate of 0.234 is applicable, and to possible outcomes when it is not. It also yields a simple formula for limiting relative efficiencies between elliptical and spherical proposals on elliptical targets.

Start time 12:40

MEASURING AND IMPROVING ACCURACY IN MCMC

Gundula Behrens
University of Bath, UK

Keywords: *MCMC, Accuracy, Integrated autocorrelation time, Antithetic variables*

MCMC samples are used for estimation. As MCMC methods come with a high computational cost, it is important to know how many samples are needed for a reliable estimation. This can be assessed by the accuracy of estimation. The better the accuracy, the less samples are needed. We will present and illustrate ways of measuring and improving accuracy in MCMC based on earlier work by Geyer (Practical Markov chain Monte Carlo, Statistical Science 1992, Vol. 7 No. 4, pp. 473-511) and Green and Han (Metropolis methods, Gaussian proposals, and antithetic variables, Lecture Notes in Statistics 1992, Vol. 74, pp. 142-164).

14.2.9 Session 3a

Session Room 203 Chair: Miland Joshi

Start time 14:15

THE USE OF DECISION THEORY IN DIAGNOSTIC TESTING

Miland Joshi
University of Warwick, UK

Keywords: *diagnostic tests decisions costs utilities QALYs*

Researchers like Margaret Sullivan Pepe have made it clear that diagnostic testing involves decisions about costs, to the patient or the health services. Costs are usually regarded in financial terms, e.g. what people (or agencies) are prepared to pay (or receive in compensation). However the actual value or utility of money doesn't rise linearly with the amount but follows a utility curve. This necessitates a translation of results involving costs into utilities. Furthermore the utility for Health problems may be better represented in terms of Quality-Adjusted Life Years or other measures. We therefore need to develop a utility function for QALYs which can be usefully applied to the problem of diagnostic testing for individuals or groups.

Start time 14:40

INVESTIGATING STATISTICAL METHODS FOR ANALYSING RANDOMISED TRIALS IN TYPE 2 DIABETES

Elisa Bonvini
The University of Reading, UK

Keywords: *open label extension, study design, switching treatment, missing values*

In recent years, Open Label Extension (OLE) studies have become increasingly popular in the design of clinical trials. OLE studies are those protocols in which patients, who were randomised to different treatments in the main study, are now re-assigned to one (or more) active treatments to assess long-term safety and tolerability. Their use prompts many interesting questions relating to design and analysis. In particular, it is desirable to find a suitable statistical method to analyse the complete data from both the main study and the OLE phase, rather than presenting separate analyses.

A Type 2 diabetes trial has been conducted, followed by an OLE phase. The data collected from both parts provides a dataset for use in exploring the issues. It is important to note that subjects drop out randomly during the OLE and, therefore individuals have unequal numbers of repeated measurements. Also, some of the subjects switch treatment when entering the OLE.

In this talk, the design of the diabetes study is described, and then a simplified version, which allows easier analysis of results, is proposed. For this simpler design, different effects of possible interest, such as treatment effect, carry-over effect, duration effect and order effect, are discussed. It is highlighted which effects can and cannot be interpreted with the design implemented. Results for the Type 2 diabetes trial are presented and discussed. Finally topics for further work are examined.

Start time 15:05

Longitudinal analysis of a birth cohort in determination of different phenotypes of asthma: A finite growth mixture modeling approach

Raquel Granell
ALSPAC,University of Bristol, UK

Keywords: *Asthma, wheezing, allergy, lung function, phenotypes, multilevel modeling, growth mixture modeling*

Asthma remains one of the most important public health problems in the UK. There is good evidence that most asthma begins in childhood and that factors in the environment before and soon after birth are important determinants of who will develop asthma. Research in this area is complicated by the large number of potential factors involved and the fact that they may have different effects depending on when an individual is exposed to them. It is also recognised that asthma is a disorder that is characterised by several different forms that vary in their onset and natural history. The Avon Longitudinal Study of Parents and Children (ALSPAC) is a population-based birth-cohort that followed 14,000 children from birth to 8.5 years. Existing data consists of measurements repeated at 7 regularly spaced time intervals between 6 and 81 months of age. The repeated measurements are made up of binary variables- did the child wheeze -, as well as ordinal variables ?how many attacks, how long did they last- and continuous variables related to allergy, lung function and severity of asthma. We want to use these data to design a mathematical model that will come out with different groups of children who we suspect to be suffering from different types of asthma and tracing different trajectories through

time. Recent developments in finite mixture modeling allow for the identification of different developmental processes in distinct but unobserved subgroups within a population. A new approach, described within the general growth mixture modeling framework, extends conventional growth models to incorporate a categorical latent trajectory variable representing latent classes or mixtures (i.e., the subgroups in the population whose membership must be inferred from the data). I will provide a didactic example of how this new methodology can be applied with asthma data to identify different phenotypes.

Start time 15:30

Bayesian Value of Information Methods in Clinical Research

Guy Freeman
University of Birmingham, UK

Keywords: *Bayesian methods, Value of Information, Cost-effectiveness, Trial design*

Outline: The Value of Information in a decision problem is defined as the increase in expected utility that will occur when that information is incorporated into the model. One way information has value is to update the parameters in a model using Bayes' Theorem. The worth of the trial can therefore be assessed by calculating its Expected Value of Sample Information (EVSI) - the expected increase in expected utility from the resultant reduction in uncertainty. But this is a complicated procedure. A simpler method to assess the value of a trial is to assume it will eliminate all uncertainty surrounding the parameters. This value, the Expected Value of Perfect Information (EVPI), and is much easier to work out. If a trial is expected to cost more than the EVPI, then it can be proven that the trial is definitely not worth carrying out. Hence the EVPI can be used as a threshold value which proposed trials have to overcome to be worthy of an EVSI calculation. Applications of both of these methods will be demonstrated with proposed trials for Left Ventricular Assist Devices and Tissue Engineering.

Sometimes uncertainty is so great that the scale of the uncertainty is uncertain. The information needed is nigh-on impossible to get. The solution here is similar to the one above: use a simpler model as a threshold for further analysis. This approach is evinced with a cost-effectiveness analysis of tissue-engineered urethra, a technology appearing prima facie to be unlikely to be cost-effective. Instead of setting up a complicated decision model, I assumed a very optimistic case for the technology over current best treatment, and showed that even then cost-effectiveness is not achieved. A more formal calculation was thus not required.

Conclusion: Sometimes it is not necessary to gather as much detail as possible about

a problem before making a decision. A Value of Information approach, applied as appropriate for the level of uncertainty present in the model, is potentially both a pragmatic and rigorous way of optimising the process of decision-making.

14.2.10 Session 3b

Session Room 214 Chair: Erik Baurdoux

Start time 14:15

WANT AN INTERESTING AND WELL PAID JOB?

Matthew Killeya
Winton Capital Management, UK

Keywords: *Hedge Fund, efficient market hypothesis, statistical modelling*

The existence of hedge funds challenges one of the cornerstones of modern econometric theory; the *efficient market hypothesis*. The theory basically says that all relevant information is contained within current prices so that it should be impossible to "beat the market". Systematic funds, such as Winton, provide evidence to the contrary through consistent profitable speculation on the markets using statistical models. Characterising and profiting from market inefficiencies gives rise to several interesting and difficult statistical issues, such as non-stationarity and judgments of exchangeability, particularly as effects are typically masked by a large noise component. The talk will touch on some of these issues and should be of interest to PhD students interested in joining a research group of scientists with a substantial and growing number of statisticians.

Start time 14:40

STUDIES OF FINANCIAL TIME SERIES ANALYSIS

L.J. Carbonaro and Prof. B.D. Ripley
University of Oxford, UK

Keywords: *Time-Varying Volatility, Stochastic Volatility, Recurrent Neural Networks, Penalised Quasi-Likelihood, Foreign Exchange Market, Basel Accords.*

This talk will describe two of the most important volatility models in finance: the ARCH model by Robert F. Engle, Nobel Price winner in 2003, and the SV model by Stephen J. Taylor. Then, it will consider the role of neural networks as statistical tools and the concept of penalised quasi-likelihood methods as possible techniques to overcome the limitations of stochastic volatility processes. In respect of these two considerations, a novel model will be introduced together with a new estimation procedure for financial time series analysis. Results will be provided by using foreign exchange market data and simulation methods. The talk concludes with the measuring of market risk in accordance to the recommendation of the Basel Accords.

Start time 15:05

THE MATHEMATICS OF FINANCIAL CRASHES

John Fry
University of Sheffield, UK

Keywords: *Mathematical finance, financial crashes, log-periodic precursors, econophysics*

As the stock market came to the attention of increasing numbers of physicists, one of the ideas that has recently emerged is that it might be possible to develop a theory of stock market crashes. Central to this talk is the somewhat controversial subject of Log-Periodic Precursors to financial crashes. The original driving-force behind it all was the idea that so-called complex systems might exhibit universal fingerprints prior to failure. We motivate the Sornette-Johansen method, with a brief overview discussion of statistical mechanics in general and the 2-d Ising model in particular. We then go on to describe the method, introduce the Johansen-Ledoit-Sornette power-law model, and detail some of the criticisms of the method that have been made in the literature. If time permits, we will go on to discuss the Sornette-Johansen method in the light of empirical statistical analysis, and further developments and results obtained. We close with a brief summary of the likely flavour of future research.

Start time 15:30

COUNTERPARTY DEFAULT RISK IN AFFINE PROCESSES WITH JUMP DECAY

Angelos Dassios and Pauline Sculli
London School of Economics, UK

Keywords: *Shot Noise Process, Piece-wise Deterministic Markov Process, Counterparty Risk, Credit Derivatives*

We study a piece-wise deterministic Markovian default intensity with a deterministic jump decay rate in a bivariate setting. The framework originates from Dassios and Jang (2003) where a Cox process with shot noise intensity is used to model claim arrivals for the pricing of catastrophe reinsurance contracts.

Recall Jarrow and Yu (2001) where counterparty default is studied in a two firm

setting. Our approach generalises Jarrow and Yu to allow for a much richer dependency structure. We allow arbitrary jump shocks to the credit quality of name i, not necessarily the event of default itself, to trigger default for counterparty name j. Additionally, we allow jump shocks to credit quality to persist, but dampen through time. Jumps to the default intensity of counterparty i that do trigger default to counterparty j are governed by arbitrary user-defined distributions, which we refer to as *contagion distributions*.

We solve the infinitesimal generator of the system to arrive at an exponential-affine martingale and find it is recursively defined through Laplace transforms. Explicit solutions are derived for the case where we restrict default dependency in one direction, by setting one of the contagion distributions to be degenerate. From here we can exploit the Laplace transform representation to arrive at the probability generating function of the bivariate distribution of the two default counting processes. The joint survival probabilities are then used for the risk neutral pricing of a variety of simple default contingent claims. We consider the commonly traded credit default swap, and compare values obtained using different contagion distribution assumptions.

14.2.11 Session 3c

Session Room 325 Chair: Duncan Lee

Start time 14:15

Is it better to give information, receive it or be ignorant in a two-player game?

Elaine M. K. Wilson, John M. McNamara and Alasdair I. Houston
University of Bristol, UK

Keywords: *game, Stackelberg solution, simultaneous solution, evolutionarily stable strategy, information*

The standard approach in a biological two-player game is to assume both players choose their actions independently of one another, having no information about their opponent's action (simultaneous game). However, this approach is not realistic in some circumstances. In many cases, one player chooses his action first, and then the second player chooses her action with information about his action (Stackelberg game). We compare these two games, which can be mathematically analyzed into two types, depending on the direction of the best response function (BRF) at the evolutionarily stable strategy in the simultaneous game (ESS_{sim}). We sub-categorize each type of game into two cases, depending on the change in payoff to one player, when both players are at the ESS_{sim}, and the other player increases his action. Our results show that in cases where the BRF is decreasing at the ESS_{sim}, the first player in the Stackelberg game receives the highest payoff, followed by both players in the simultaneous game, followed by the second player in the Stackelberg game. In these cases, it is best to be the first Stackelberg player. In cases where the BRF is increasing at the ESS_{sim}, both Stackelberg players receive a higher payoff than players in a simultaneous game. In these cases, it is better for both players to play a Stackelberg game rather than a simultaneous game. However, in some cases the first Stackelberg player receives a higher payoff than the second Stackelberg player, and in some cases the opposite is true.

Start time 14:40

Monthly Average Temperature Modelling

Mercedes Andrade Bejarano
University of Reading, UK

This research is associated with the goal of the horticultural sector of Colombian southwest, which is to obtain climatic information, specifically, to estimate monthly average temperature in sites where it has not been measured. Three areas of the county of Valle del Cauca are included in the research: Western Range Mountain, the valley of the Cauca River and Central Range Mountain. Because the influence of altitude on the temperature, this variable is included in the monthly average temperature estimation. ”El Niño” and ”La Niña” phenomena have affected the climate in the study zone. Particularly, in temperature, increase in the values during ”El Niño” years occurring in the period 1971 to 2002 and decrease in the temperature values during ”La Niña” (period 1999-2000) were recorded.

Two components are identified in the data of this research: 1) A component due to the temporal aspects, determined by characteristics of the time series, distribution of the monthly average temperature through the months and the temporal phenomena, which increased (”El Niño”) and decreased (”La Niña”) the monthly average temperature values 2) a component due to the sites, which is determined for the clear differentiation of two populations: The valley and the mountains, which is associated with the behaviour of monthly average temperature with the altitude, and finally, due to the closeness between meteorological stations is possible to find spatial correlation between data from nearby sites.

Three models are analysed in the Monthly Average Temperature Modelling: Random Coefficient Model without Spatial Modelling, Random Coefficient Model with spatial structure in the errors (Spherical and Gaussian models) and Meteorological model. Random Coefficient Model without Spatial Modelling and Random Coefficient Model with spatial structure in the errors are capturing the influence of ”El Niño” and ”La Niña” phenomena, which indicates that the inclusion of the random part in the model is appropriate.

Start time 15:05

USING TIME-VARYING COEFFICIENT MODELS TO ESTIMATE THE EFFECTS OF AIR POLLUTION ON PUBLIC HEALTH

Duncan Lee and Gavin Shaddick
University of Bath, UK

Keywords: *Time-varying coefficient models, Bayesian hierarchical models, penalised splines, epidemiology, MCMC simulation*

The effects of air pollution on public health are commonly estimated using time series regression methods, such as generalised linear models (GLM) and generalised

additive models (GAM). The former estimates a constant effect of pollution, while the latter models the effect as a non-linear function of the level of air pollution. We propose a time-varying coefficient model (TVCM) for this analysis, which allows the effects of air pollution to change over time. Such temporal changes may be due to an interaction with temperature, or from a change in the composition of pollutants, such as particulate matter, over time. The model estimates a smooth time-varying effect which does not follow an *a-priori* fixed parametric form. This is achieved using penalised natural cubic splines within a Poisson regression model. Estimation is performed using both quasi-likelihood and Bayesian techniques, using an iteratively re-weighted least squares procedure and Markov chain monte carlo (MCMC) simulation respectively. The models are applied to data from the National Morbidity, Mortality, and Air Pollution Study (NMMAPS).

Start time 15:30

An Extreme Value Analysis of Air Pollution in the UK

Emma Eastoe
University of Lancaster, UK

Keywords: *Air Pollution data, Extreme Value Theory, Generalised Pareto distribution, Sub-asymptotic thresholds*

The motivation for this work comes from attempts to model surface-level air pollution datasets using the peaks over threshold method, a technique common in Extreme Value Analysis. Essentially, this involves fitting a generalised Pareto (GP) distribution to the excesses of some sufficiently high (asymptotic) threshold. We introduce the data and describe some of the modelling problems, which include non stationarity and clustering. The talk focuses on the need to use sub-asymptotic, rather than asymptotic, thresholds due to the short data series available. We suggest a way in which one might derive a distribution for cluster maxima of these sub-asymptotic threshold excesses. Time permitting, some simulation study results comparing the distribution obtained with the GP distribution will be shown and some fitted models for our original data presented.

14.2.12 Session 3d

Session Room 326 Chair: Richard Everitt

Start time 14:15

Efficient Particle Filters for Ergodic Stationary Processes

Hassan K. Khalil
Department of Mathematics, University of Bristol, UK

Keywords: *Sequential Monte Carlo, Particle Filters, Markov Chains, Partially Observed Stochastic Systems*

Stochastic filtering or Bayesian filtering is the theory of recursively estimating the underlying state of a dynamical system, using observations obtained over time.

Particle filtering algorithms are sequential Monte Carlo methods that allow one to numerically solve the problem of stochastic filtering for dynamic models with hidden states.

Key to the efficiency of these algorithms is the appropriate design of the so-called "importance distribution". The optimal importance distribution is a well known theoretical quantity, but rarely usable directly in practice. Analytical approximations are usually sought.

In this work we take a different approach, and exploit the stationarity property of some dynamical system, which allows us to quasi-automatically design near optimal importance distribution in an off-line manner. A direct by-product of our approach is that we can directly and efficiently implement the auxiliary particle filtering principle and extensions, which leads to even more efficient algorithms with only marginal computational overhead.

We demonstrate the efficiency of our approach through numerical simulations, and discuss extensions of our methodology to certain types of non-stationary processes.

Start time 14:40

Graphical Models for the Internet

Richard Everitt
University of Bristol, UK

Keywords: *Graphical models, Bayesian networks, random graphs, social networks, internet, relational models.*

State-of-the-art search engines have proved themselves to be useful portals to the information contained in the internet. However, their deterministic nature can be restrictive. This talk describes some initial ideas about how Bayesian modelling may be used in order to make more sophisticated queries of the available information.

Start time 15:05

Implementation of Model Averaging for Dynamic Classification Problems

Amber Tomas
University of Oxford, UK

Keywords: *Model Averaging, Dynamic Classification*

Model averaging has proved to be a very useful technique for improving the predictive performance of classifiers. From a Bayesian standpoint it also results in more accurate estimation of the variance of our estimates, as it accounts for the uncertainty in model selection.

In the past, most classification algorithms have been built to be used on stationary data. However, in more and more fields of application it is becoming clear that often this assumption of stationarity is violated. Hence methods for classification need to be developed that can applied to problems where the underlying populations are dynamic rather than stationary.

In this talk I will briefly introduce both model averaging and dynamic classification, and then discuss the main issues that arise when trying to apply model averaging to dynamic classification. This discussion will be focused on the computational difficulties which arise, and finding suitable approximations to allow such methods to be implemented.

Start time 15:30

A Bayesian Dynamic Generalised Linear Model to monitor population trends in ecology

Chiara Mazzetta[1], Steve Brooks[1] and Steve Freeman[2]

[1] *Statistical Laboratory, University of Cambridge, UK*
[2] *British Trust for Ornithology, Thetford, UK*

Keywords: *Dynamic GLM; MCMC; Population Trends; State space; Zero-inflated Negative Binomial*

We consider the UK Common Birds Census (CBC) counts and their use in monitoring the species. We use a state space modelling approach within a Bayesian framework to describe the population trend over time and the alert system used by the British Trust for Ornithology. We account for potential overdispersion and excess zero counts, by modelling the observation process with a Zero-Inflated Negative Binomial, while the system process is described by second-order polynomial growth models. In order to provide a biological motivation for the amount of smoothing applied to the observed series, the system variance is related to the demographic characteristics of the species, so as to help the specification of its prior distribution

14.2.13 RSS Guest Speaker

Session Room 203 Chair: Dr. Ben Torsney

Start time 17:15

ESTIMATING INTERNATIONAL MIGRATION

Prof. Tim Holt
School of Social Sciences, University of Southampton, UK

Estimates of international migration into and out of the UK are of great importance for a number of reasons: they contribute the main uncertainty into total population estimates which are used for a variety of purposes and are central to any national statistical system, they have high political profile in their own right given the political sensitivity of the level of migration and its sources into the UK.

These estimates are based in the main on the International Passenger Survey and have been subject to very large revisions following the recent Census.

The seminar will give a brief account of the current system and then focus on whether there is an alternative approach that could give better quality statistics either as a replacement for the IPS in this regard or as a means of assessing the quality of the estimates coming from the IPS.

14.3 Wednesday 22nd March

14.3.1 Session 4a

Session Room 203 Chair: Jinghao Xue

Start time 09:25

GAUSSIAN APPROXIMATION BASED MIXTURE REDUCTION FOR NEAR OPTIMUM DETECTION IN MIMO SYSTEMS

Y. Jia[1], C. Andrieu[1], R. J. Piechocki[1] and M. Sandell[2]
[1] *University of Bristol, UK*
[2] *Toshiba Research Europe Ltd.*

Keywords: *Space-time processing, Multiuser Detection, Gaussian approximation*

The optimal "soft" symbol detection for spatial multiplexing multiple input multiple output (MIMO) system with known channel information requires knowledge of the marginal posterior symbol probabilities for each antenna. The calculation of these quantities requires the evaluation of the likelihood function of the system for all possible symbol combinations, which is prohibitive for large systems. It is however most often the case that most of the transmitted symbol combinations contribute only very little to these marginal posterior probabilities. We propose in this paper a suboptimal procedure which identifies the most significant symbol combinations via a sequential algorithm with Gaussian approximation (SGA). Simulation results show that our method can approach the optimal a posteriori probability detector (APP) performance while being less complex than comparable suboptimal algorithms, such as the sphere decoder (SD). We further demonstrate that as opposed to the SD the complexity and memory requirements of our algorithm are fixed, therefore easing practical implementation.

Start time 09:50

FEATURE SELECTION USING DISCRIMINANT ANALYSIS

Dan Bailey
University of Bristol, UK

Keywords: *Early Day Motions, scaling solutions, linear discriminant analysis, genetic optimisation*

Early Day Motions (EDMs) are unwhipped motions tabled and signed by Members of Parliament (MPs) to give an opinion on any subject matter. Despite there being no pressure for an MP to sign according to their party, the data set shows strong party structure and has been used previously to investigate similarities between MPs and political parties.

Having classified EDMs into different types, methodology will be explained to select features of the data (i.e. particular issues) which differentiate the three main political parties. Drawing together various mathematical tools used in this application, initial results will be discussed.

Start time 10:15

SURFACE SHAPE ANALYSIS, WITH AN APPLICATION TO BRAIN CORTICAL SURFACE ANALYSIS IN SCHIZOPHRENIA

Christopher J. Brignell
School of Mathematical Sciences, University of Nottingham, UK

Keywords: *Cortical surface, High-dimensional, Likelihood, Neuroscience, Shape, Symmetry*

In many application areas it is of interest to compare the shapes and sizes of high-dimensional surfaces, and to investigate symmetry. We focus on a particular application in neuroscience, investigating large scale cortical shape differences between control and schizophrenia patients. We introduce an automatic maximum likelihood method for brain registration, identifying the inter-hemispherical join, the anterior commissure, and the posterior commissure. Likelihood based inference is considered, and significant differences between the two groups are observed. The model is extended to account for curvature in the inter-hemispherical join. General practical issues in high-dimensional data analysis will be discussed.

Start time 10:40

GENERATIVE VS. DISCRIMINATIVE MODELLING

Jing-Hao Xue and Mike Titterington
Department of Statistics, University of Glasgow, UK

Keywords: *Generative Models, Discriminative Models*

Generative models, like linear discriminant analysis, and discriminative models, like logistic regression, are two sides of a coin of statistical modelling for classification. This presentation will introduce a pedigree of different generative models and their available counterpart of discriminative models. We will also discuss methods that represent compromise between these two approaches.

14.3.2 Session 4b

Session Room 214 Chair: Michailina Siakalli

Start time 09:25

PROPERTIES OF A CLASS OF LIFE DISTRIBUTION INVOLVING HALF-NORMAL MIXTURES

Bander Al-Zahrani
University of Newcastle upon Tyne, UK

We study a class of lifetime distributions with decreasing, increasing, bathtub shaped, or upside-down bathtub shaped mean remaining life (MRL) and variance remaining life (VRL). These characteristics used to model various lifetime data. Most populations of components are heterogeneous consisting of "good" components with long lives and "defective" components with short lives so it would of interest to study mixtures of such distributions. In particular, we concentrate and derive several results for mixtures of half-normal distributions.

Start time 09:50

THE MOMENTS OF A LEVY PROCESS

Michailina Siakalli
University of Sheffield, UK

Keywords: *Levy processes, existence of moments, compound Poisson process*

Lévy processes are stochastic processes with stationary and independent increments. They are the simplest class of process that have continuous paths that are interrupted with jump discontinuities of random size that appear at random times.

The main purpose of this talk is to investigate necessary and sufficient conditions under which a Lévy process has finite moments of all orders. In published literature proofs exist but in my talk an alternative simpler approach will be given.

The problem is reduced to finding necessary and sufficient conditions for existence of the moments of a compound Poisson process, which is naturally associated to a Lévy process since it describes its "large jumps".

Start time 10:15

THE SIMPLEX SAMPLER

Nial Friel and Eleni Bakra
Department of Statistics, University of Glasgow, UK

In this talk we present a new MCMC sampler based on ideas from population MCMC. Usual MCMC methods seek to gather a single sample based on moving a single point x around the sample space. However the performance of the sampler often depends on appropriate choice of scale of the proposal distribution. For example, in the Metropolis-Hastings algorithm, if the proposal variance is not large enough, then the sample might tend to get trapped in local maxima. Here we aim to avoid these problems by instead defining a Markov chain on a population of points from the sample space, and where proposal moves are based losely on ideas from the Nelder-Meade simplex method. In this way, moves are proposed taking into account local information based on the current state of the population. We outline the performance of the algorithm on real and simulated data.

Start time 10:40

GAME OPTIONS

Andreas Kyprianou and Erik Baurdoux
Heriot Watt University, UK

Keywords: *Stochastic game, Lévy processes*

A game option is a contract between a writer and a holder. The value of the contract depends on an underlying stochastic process. Both the writer and the holder can terminate the contract at any time, inferring a payment from the writer to the holder. We consider the game version of the put option and take a spectrally negative Lévy process as the underlying process. We show that there exist optimal stopping strategies for the writer and the holder and find the value of this game option.

14.3.3 Session 4c

Session Room 325 Chair: Andrew Thomson

Start time 09:25

Penalties in Nonproportional Regression Survival Models

Maria João Costa
Department of Statistics, University of Warwick, UK

Keywords: *Penalized Methods, Survival analysis, Time-Varying Coefficients*

Recent years have seen an increase in interest in the use of penalized spline methods, especially in the context of regression and generalized linear models. In this talk some of the penalty functionals proposed in the literature will be discussed. A parameterization based on the values and slopes of the spline function at the knots is presented which allows the decomposition of the penalty term into interpretable quantities. Emphasis will be given to the implementation of penalized methods in the survival analysis setting in the particular case where the usual Cox proportional hazards model is extended to accommodate dynamic covariate effects. In this context, penalized spline methods can be used not only to measure departures from the proportional hazards assumption but also to obtain an estimate of how such effects vary with time.

Start time 09:50

Automatic identification of 'interesting' longitudinal series of counts: application to monitoring MRSA rates in NHS Trusts

Hayley Jones
MRC Biostatistics Unit, University of Cambridge, UK

Keywords: *Longitudinal, Performance monitoring, MRSA, Risk-adjusted EWMA, FDR*

Data measuring the 'performance' of institutions is often collected at regular intervals over time, with interest lying in temporal trends as well as cross-sectional comparisons at any one time point. One example which currently has a high profile is rates of Methicillin Resistant *Staphylococcus Aureus* (MRSA) bacteraemia infections

in NHS Trusts. Government targets call for each Trust to reduce its rate by 20% each year.

I will discuss simple methods for automatically identifying 'odd' changes in individual Trusts over the first four and a half years of MRSA surveillance. Firstly, Trusts with apparent *overall* increases and decreases in rate are picked out by fitting independent Poisson regressions and examining the fitted gradients. Secondly, Trusts with 'significant' changes in the last six months are identified, using a simple test based on observed performance in the final two periods. An alternative, more powerful, test for recent change is then carried out based on comparisons of 'smoothed' estimates of recent performance, using a risk-adjusted version of an exponentially weighted moving average (EWMA). The false discovery rate (FDR) approach to thresholding is outlined and applied throughout as a means of accounting for multiple testing.

Start time 10:15

Outcome Selection Bias in Meta Analysis

Kerry Dwan
University of Liverpool, UK

Keywords: *Meta-analysis, Publication bias, Outcome selection bias*

Important new findings in medical research have become evident due to the extensive use of Meta-analysis in systematic reviews. An example of this is the increased risk of death following administration of human albumin solution, which was once standard practice.

However, there are several types of bias that are recognised as a potential threat to the validity of a meta-analysis; Publication bias, which arises due to trials being published or not dependent on their results. Empirical research suggests that published work is more than twice as likely to be statistically significant ($p < 0.05$) than unpublished research. Evidence is also gathering for the existence of outcome selection bias, which is the selection of a subset of the original variables recorded for inclusion in publication of trials.

Therefore, readers of systematic reviews should be aware of the problems of bias in meta-analyses and interpret the results with caution.

Different methods of meta-analysis are compared and different approaches in assessing the existence of bias are investigated. These include; tests of asymmetry, methods for estimating the number of missing studies, methods that allow the pooled meta-analysis estimates to be adjusted for the bias and sensitivity analyses. We are currently looking at extending a recent model for quantifying the size of outcome

selection bias to randomised controlled trials.

Long term solutions to the problem of bias include; protocol registration and an outcome amnesty to complement the trial amnesty and improving the quality of published research by reviewing the report and protocol together. Where-as, in the short term, missing data should be obtained when possible and a thorough search strategy to identify relevant research should always be completed.

Start time 10:40

Effects of Missing Values on the Analysis of the AB/ BA Crossover Trial

Lauren Rodgers and Prof. JNS Matthews
University of Newcastle, UK

Keywords: *Crossover Randomised Clinical Trials, Missing Data, Medical Statistics*

It is known that the omission of trial subjects from the final analysis can lead to bias this is an inevitable difficulty when the subject defaults and no response is observed. This project looks specifically at this aspect of trials in the context of crossover randomised clinical trial, concentrating on the effect missing values have on parameter estimates and consequent conclusions. The crossover design can have either fixed or random subject effects and their approaches to missing values differ. The fixed subject effect discards all data from a subject if they have any missing values whereas the random subject effect utilises all available data in the analysis. A simulation study with values missing completely at random (MCAR in Rubins terminology) compares the outcomes of the fixed and random subject effect methods. The differences and similarities of the outcomes of the two models will be discussed.

14.3.4 Session 4d

Session Room 326 Chair: Brett Houlding

Start time 09:25

How can we best estimate missing observations in longitudinal studies of ageing?

Graciela Muniz Terrrera
MRC Biostatistics Unit, University of Cambridge, UK

Keywords: *Missing at random, Non Ignorable Missing Data,Growth Mixture Models*

In longitudinal studies of ageing, missing observations are usually considered to be missing at random, though a non ignorable missing data mechanism might model reality better. Subject specific trajectories of cognitive decline were simulated with increasing number of monotone missing observations under a missing at random and non ignorable missing data mechanism. Bayesian models under both missing data mechanisms were run and their results compared against results obtained from a Growth Mixture Model obtained in MPLUS when individuals were examined at fixed interview times.

Start time 09:50

Hidden Markov models with time dependent misclassification

Andrew Titman
MRC Biostatistics Unit, University of Cambridge, UK

Keywords: *hidden Markov model, multi-state models, disease progression*

Markov models are a convenient way of modelling data that are characterised by repeated measurements of a discrete state stochastic process. In particular they are used in the modelling of chronic diseases. Where disease state is not directly observed, or is measured with error, it is necessary to incorporate misclassification via a hidden Markov model (HMM). A HMM makes the assumption that conditional on the underlying true state, the observations are independent and identically distributed. For HMMs with a discrete observed disease marker, this implies

the observed states for an individual should only be correlated through the underlying state and not beyond. The assumption of independence in the misclassification mechanism is violated for disease processes in which the disease marker also depends on an explanatory variable. Explanatory variables may affect the disease marker over a number of observations causing a series of misclassified states. Where there are sufficient data about the explanatory variables they can be modelled as covariates on misclassification probabilities. A more sophisticated model in which the misclassification process is governed by a second, unobserved Markov process, independent of the disease process is proposed.

The method is applied to the 4-levels (absent, mild, moderate and severe) of bronchiolitis obliterans syndrome in recipients of lung transplantation, which are diagnosed with error, and the probability of misclassification is dependent on the presence of acute events, such as infection or rejection of the transplanted lungs. In this application the two states of the secondary, unobserved Markov process represent 'acute event', resulting in inflated misclassification probabilities, and 'no acute event'. Limitations of the model and possible extensions will be discussed.

Start time 10:15

STOCHASTIC MODELLING OF THE MAPK SIGNALLING PATHWAY

Prof. Ernst Wit and Vilda Purutçuoğlu
University of Lancaster, UK

Keywords: *MAPK pathway, Stochastic simulation*

The MAPK (mitogen-activated protein kinase) or its synonymous ERK (extracellular signal regulated kinase) pathway whose components are Ras, Raf, and MEK proteins with many biochemical links, is one of the major signalling systems involved in cellular growth control of all eukaryotes including cell proliferation, transformation, differentiation, and apoptosis. The structure of the MAPK/ERK pathway includes a number of phosphorylation on the protein level directed by positive and negative feedback loops. Current knowledge about the MAPK/ERK pathway allows us to represent it as a list of (quasi) biochemical reactions. A novelty of our approach is to use multiple parametrizations in order to deal with molecules for which localization in the cell is an intricate part of the dynamic process.

The common way to model gene regulations is to use the ordinary differential equations (ODE), which can describe several reactions such as linear production and degradation realistically but can not explain the stochastic nature of the actual interactions. We have implemented the MAPK/ERK pathway via a stochastic Markov process. Under some weak assumptions, the system is simulated by the Gillespie

algorithm. However due to its inefficiency for developing realistic complex models, we also use diffusion approximation for simulating the MAPK/ERK pathway. In this study we aim to compare the diffusion approximation with the exact algorithm for MAPK/ERK pathway.

Start time 10:40

BAYESIAN EVIDENCE SYNTHESIS FOR ESTIMATING HIV PREVALENCE AND INCIDENCE

A. M. Presanis
Medical Research Council Biostatistics Unit, Cambridge, UK

HIV prevalence and incidence are estimated by synthesising routine surveillance data and data from ad-hoc surveys in a Bayesian framework. We first model HIV prevalence in England and Wales in 2001 among five groups, sex between men, injecting drug users, genito-urinary clinic attenders, men and women born in Sub-Saharan Africa, and lower risk individuals. We simultaneously estimate the proportion of the population of England and Wales in each risk group (ρ_g), HIV prevalence in each group (π_g) and the proportion of HIV positives diagnosed in each group (δ_g), by expressing each data source as functions of these parameters.

We extend this model to incorporate HIV incidence through a Markov multi-state model. Individuals move through states representing HIV negatives, HIV positives who are undiagnosed and diagnosed HIV positives. Entry into the model is at age 15 into the lower risk group, or else through immigration into any of the risk groups. Individuals may also move into higher risk groups, and exit the model either at age 45, through emigration, or through death. In each of the years 1997 - 2001, state membership is expressed in terms of the three parameters ρ_g, π_g and δ_g. The earlier prevalence model is hence a snapshot in each year of the Markov process governing movement between states. Differential equations in terms of the transition rates (incidence, diagnosis rates, mortality rates, entry rates into higher risk groups, migration into and out of the population, entry at age 15 and exit at age 45) describe the change in population in each state.

These models are implemented through Markov Chain Monte Carlo simulation using WinBUGS 1.4.1 and the WBDiff differential equation solver.

14.3.5 Session 5a

Session Room 203 Chair: Chris Brignell

Start time 11:25

STATISTICAL VISUALIZATION OF WIRELESS SENSOR NETWORKS (WSN)

Naventhan Mahadevan, Professor Guy Nason and Professor Alistair Munro
University of Bristol, UK

Keywords: *wireless sensor netwoks, information visualization, multiscale visualization*

Recent advances in the areas of microprocessor and wireless communication technologies have enabled development of small, low cost, low power distributed sensor nodes capable of sensing, processing and communicating. The main focus thus far has been given to the development of routing and medium access control protocols to meet the unique challenges of wireless sensor networks but less emphasis on information visualization. The challenge of managing a large-scale self-organising network remains unsolved. Wavelet-based distributed data processing hold much promise for wireless sensor networks. However irregular sensor node placement limits the direct application of standard wavelet techniques. We propose a multiscale visualization technique for tree-based sensor network data based on the lifting scheme. We also expect to improve the network performance using the information predicted by this multiscale technique.

Start time 11:50

MIXED INTERACTION MODELLING WITH APPLICATIONS

Brian Sullivan
University of Limerick, Ireland

Keywords: *Mixed Interaction Models, Graphical Models, Delete = Replace*

Consider the mixed interaction model for the set I of p discrete variables and the set Y of q continuous variables

$$f(i,y) = p_i|2\pi\Sigma_i|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(y-\mu_i)'\,\Sigma_i^{-1}\,(y-\mu_i)\}$$

where i is a p-tuple containing the values of the discrete variables, p_i is the probability that $I = i$ and μ_i and Σ_i are the corresponding mean vector and covariance matrix for each level of I.

This distribution underpins the modelling of discrete and continuous random variables in terms of graphical models. This is implemented in the software package MIM [1]. This talk will focus on a graphical modelling approach to the analysis of large health based surveys. In particular, we will focus on the application of the 'delete equals replace' paradigm [2] for addressing issues of missing data in this context.

References

[1] David Edwards, Introduction to Graphical Modelling, 2000, 2nd Edition, Springer-Verlag, New York.

[2] Haslett, J. and D. Dillane, Application of 'delete = replace' to deletion diagnostics for variance component estimation in the linear mixed model, Royal Statistical Society, Series B, 2004, vol.66, 131-143.

Start time 12:15

COMPARING THE SOLVENT-ACCESSIBLE SURFACES OF PROTEINS

Kim Evans

School of Mathematical Sciences, University of Nottingham, UK

Keywords: *proteins, MCMC, active sites*

Here we consider the problem of how to establish whether two protein surfaces share a region of geometric and physico-chemical similarity. One of the challenges is the large number of atoms (several thousand) defining the surfaces of the proteins. We describe an MCMC-based approach which enables us to optimally match unlabelled point sets with up to a couple of hundred atoms in each set. The main parameter of interest is a so-called match matrix that codes for a particular matching of points between the two point sets. The original algorithm often gets stuck in local maxima, so we extend the method by occasionally allowing very big jumps in parameter space. This makes the algorithm much more likely to converge on the true region of interest.

An important stage is to find a way to select some potential (small) sites from which to begin the MCMC algorithm. We use distance-based methods to compare pairs of balls of atoms. Specifically we look at the nearest n atoms to atom i in protein 1 and the nearest n atoms to atom j in protein 2 and use a simple criterion based on radial distances from the central atoms to decide whether a pair of balls of atoms is a suitable starting point for the MCMC algorithm. The results are encouraging and generate a series of possible sites of similarity between the protein surfaces which could in theory be tested by biologists.

Start time 12:40

MCMC METHODS ON ESTIMATING THE GENETIC AND GEOGRAPHICAL HISTORY OF INDIVIDUALS

Prof. Stephen P. Brooks and Ioanna Manolopoulou
University of Cambridge, UK

We are given a section of the DNA sequences of a number of individuals from different locations. Our aim is to understand the history behind our data: who has the "oldest" sequence, who is a mutant of whom, which population colonized where? A number of difficulties arise in solving this problem: Missing sequences (i.e. uncertainty in the intermediate mutations between two sequences) and homoplasy (i.e. a specific amino acid position mutating repeatedly) being two of the main ones. Using appropriate assumptions for the mutation process, we use Reversible Jump Markov Chain Monte Carlo methods (RJMCMC) to overcome these difficulties in order to estimate the underlying gene tree and associate it with the geographical data. We propose methods to update all our parameters (e.g. the underlying gene tree, mutation process parameters, significant clusters and their means) simultaneously. We apply our methods to a dataset from beetles on La Palma island of the Canaries and compare our output to the results of methods currently used.

14.3.6 Session 5b

Session Room 214 Chair: David Jenkinson

Start time 11:25

Stochastic Epidemic Models for Emerging Diseases Incorporating Interventions in Real-time

Simon Spencer
University of Nottingham, UK

Keywords: *Stochastic epidemic models*

Epidemic models are used to predict the behaviour of emerging diseases and also to assess the effectiveness of proposed control measures or interventions. This work develops continuous time stochastic models that incorporate a real-time delay before the intervention is applied. The threshold parameter or reproduction number is derived in order to quantify the reduction in the effectiveness of the intervention caused by its delay.

Start time 11:50

Using Computer Graphics in Probability Elicitation - Priors for Generalized Linear Models

David J. Jenkinson and Paul H. Garthwaite
Open University, UK

Keywords: *Bayesian, prior, computer graphics, probability elicitation, generalized linear model*

Bayesian statistics requires the use of prior distributions. There are differences of opinion as to the type of prior that should be used, specifically whether you should include personal opinion (subjective probabilities) or not. If you allow personal probabilities you need to know about how to determine an individual's "degree of belief" about uncertainty.

There is a lot of research into such *probability elicitation*, within the psychology field

as well as within statistics. Many methods have been tried, with varying degrees of success. An under-explored method is that of using graphs. People like pictures, so perhaps asking them to plot a graph of their opinions is better than asking them for some numbers that represent them.

Eliciting opinion about a generalized linear model is a complex process. We might model the prior as a multivariate normal distribution, thus requiring a mean vector and a covariance matrix. It is not straightforward to ask questions that an individual can easily understand and answer that also have the required complexity to ensure that the resulting covariance matrix is positive definite. The theory about one way to do this is given by Garthwaite and Al-Awadhi (currently unpublished). The technique involves modelling the sampling distribution as piecewise-linear.

Graphical elicitation becomes much easier if we take advantage of computer software. In this talk I will demonstrate some software that I have written which can be used to elicit a prior distribution for a generalized linear model, using the above theory.

Start time 12:15

Sequential Decision Making with Adaptive Utilities

Brett Houlding
Department of Mathematical Sciences, University of Durham, UK

Keywords: *Sequential Decision Making, Uncertain Utility, Bayesian Learning*

The use of the Bayesian paradigm coupled with acceptance of Bernoulli's expected utility hypothesis provides a powerful and philosophically compelling methodology for decision making in situations involving uncertainty. In the case of sequential decision making a decision maker may use observations of past results to update her beliefs over the future likelihood of outcomes given selection of a particular decision. However, it is traditionally assumed that the decision maker is able to fully state her utility function for any possible consequence realisation. As such, once the distribution of various consequences are known for each possible decision, classical Bayesian decision theories do not permit a decision maker to ever learn or be surprised about the amount of utility gained.

This talk presents a brief overview of the progress made in the development of *Adaptive Utility Theory*. Such a theory generalises classical Bayesian decision making by permitting the case that the decision maker is initially unable to fully quantify her current and future preferences. Assuming a parametric form for the utility function, it is envisioned that learning takes place in a similar manner to that which allows

the decision maker to update beliefs over parameter values for consequence distributions, *i.e.*, the decision maker may update her beliefs over unknown parameter values following data observation. These unknown parameters and observations may take many forms, but interesting examples include vectors of trade-off weights or measures of risk aversion, with categorical data being comprised of self-reflection into how surprising the actually experienced utility of a particular consequence was. The important distinction from classical theory, h! owever, is that in a sequential setting, the optimal initial decision need no longer correspond to that which would be determined if it were assumed that learning of utilities did not take place.

Start time 12:40

CHOOSING OPTIMAL CUSTOMER MANAGEMENT STRATEGIES

I-Ding Wu
Department of Mathematics, Imperial College London, UK

Scorecards are widely used in the personal financial services sector to guide choices between different ways of treating a customer; that is between 'actions'. For simplicity, in this paper we assume that there are just two possible courses of action. A common problem is that the data available from which to estimate the comparative effects of the two actions is not a random sample from the population: the customers assigned to the two actions have differences. We describe a strategy for tackling this, which begins with the selection models of Heckman and others, and adaptively improves the action assignment rule.

14.3.7 Session 5c

Session Room 325 Chair: Mark Kelly

Start time 11:25

BAYESIAN NONPARAMETRIC MODELLING OF SPATIAL DATA

Michalis Kolossiatis and Dr. James Griffin
University of Warwick, UK

An area of Bayesian Statistics that has attracted much attention recently is non- or semi-parametric modelling. In many of those applications, a non-parametric prior process is expressed by a Dirichlet Process (and more recently a Normalised Inverse-Gaussian Process) for an unknown distribution. These models seem to be the natural choice in a hierarchical model. Finally, these models take great advantage of the computational benefits of MCMC methodology.

Taking the Dirichlet Process as an example, we can see that a way of constructing a random probability measure is by normalising a random measure (the Dirichlet Process can be seen as a normalised Gamma Process). Also, the Normalised Inverse-Gaussian Process is a normalised Inverse-Gaussian Process. The aim of my project will be to exploit the infinite divisibility of the underlying random measure (in our examples, the Gamma and the Inverse-Gaussian Processes, for which the infinite divisibility holds), in order to construct random probability measures that are identically distributed, but not independent. I will study the properties of those models, and in the future, try to apply them in spatial modelling.

Start time 11:50

NONPARAMETRIC SMOOTHING METHODS AND THE MULTI-RESOLUTION CRITERION

Judy Tang
University of Bristol, UK

Keywords: *Nonparametric Regression, Minimisation Problem, Curve Estimation, Taut String*

There are many situations where it would be useful to summarise trends, patterns and relationships that may occur within any given data set or to make predictions

on future events from past information. However, there is often noise clouding the data and finding the underlying pattern of the data can be difficult.

This is often represented by the classic problem of how to distinguish the best curve passing through a set of data points. I will be looking at several smoothing methods to find a good approximation to a data set; to find the underlying function or signal within the noisy data in the hope of producing a clearer picture of what is happening. Among the Nonparametric Regression methods explained, there will be a detailed description of the Taut String Algorithm. I will also investigate an automated procedure to find the optimal smoothing parameters often required in regression techniques, called The Multi-Resolution Criterion. Both 1 and 2-Dimensional data sets are used to test the smoothing methods and conclusions are drawn upon their effectiveness in light of other work on the same subject area.

Start time 12:15

SPATIAL MODELING OF LINE TRANSECT SURVEY DATA

Irene Kaimi and Carmen Fernandez
Lancaster University, UK

Keywords: *Line transect data, Voronoi tessellations, Reversible Jump MCMC*

A Bayesian formulation is given for the problem of modeling spatial point patterns arising from line transect surveys of wildlife. Our data consist of the locations of observed animals, but only incomplete information is available because not all animals present in the region surveyed are detected. Our interest is in estimating the spatially varying animal density surface. We consider a spatial Poisson process for the locations of animals, where the intensity surface is assumed to be a piece-wise constant, using Voronoi tessellations of constant intensity in each partition set of the space (tile). The location and number of tile centers is unknown and prior distributions are chosen. The posterior distribution is computed using a Markov Chain Monte Carlo algorithm with Reversible Jump. By treating the tessellation as an unknown on which we conduct inference in a Bayesian context, we obtain an adaptive and flexible spatial estimate of the density surface, which can capture both smoothness and abrupt changes.

Start time 12:40

DERIVATION OF SYNTHETIC AREA BOUNDARIES AND THEIR USE IN THE ANALYSIS OF MENTAL HEALTH

Mark Kelly
Department of Epidemiology, Statistics and Public Health, Cardiff University

Keywords: *hierarchical modelling, mental health, spatial modelling*

It has long been known that individual characteristics such as, age, gender and marital status affect a person's mental health. A more complicated question involves the effect of where a person lives on their mental health. There is some evidence to show that where one lives can affect one's physical health, however investigations on the effect of place of residence on mental health are fraught with complications. In recent times larger computing power and better software have helped by providing researchers with the means to implement complicated hierarchical models. The ability to model such situations has perhaps progressed faster than the methodologies surrounding their use. Much work has yet to be done to assess the effect of assumption violation on the validity of the results of hierarchical modelling. Similarly, the suitability of administrative boundaries to act as proxies for areas of residence has not been properly explored. In this talk, an alternative approach will be outlined, whereby new, or"synthetic" areas will be created by grouping regions with similar resident composition together. These new, homogenous areas will improve the efficiency of the analysis. The method will be illustrated with reference to a dataset collected in Caerphilly, an area of Wales just north of Cardiff. Hierarchical models will be fitted using both the administrative and synthetic boundaries and the results compared.

14.3.8 Session 5d

Session Room 326 Chair: Michael Fahey

Start time 11:25

MIXTURE CURE MODELS AND MISSPECIFICATION

Patrick Ho
University of Warwick, UK

Start time 11:50

META-ANALYSIS OF MULTI-ARM TRIAL AND PUBLICATION BIAS

Hathaikan Chootrakool
University of Newcastle upon Tyne, UK

Keywords: *Meta-Analysis, Multi-arm trial, Publication bias*

Some meta-analysis studies assume that there is one treatment and one control arm while other studies have considered several treatment arms and possibly more than one control arm. The aim of multi-arm trial analysis is to compare the effectiveness of different treatments. The treatment might be from different studies or the same study and may be compared directly or indirectly. These lead to the problems of multi-arm trial.

One of the major problems in meta-analysis for multi-arm trials is publication bias. The studies are usually chosen through a literature review. Systematic reviews aim to find and assess for inclusion all high quality studies, but finding all studies is not always possible and there is no way of knowing what might have been missed. Studies with significant or positive results are more likely to be chosen than those studies with non-significant or 'negative' results. There are several methods to detect publication bias such as a funnel plot, or the "Trim and Fill" method. However there is no simple method to correct for publication bias without making the strong assumptions. Copas and Shi (2001 and 2002) proposed a method of sensitivity analysis using a stochastic selection model to represent the underlying publication process. This method is used to address the problem of publication bias in meta-analysis for multi-arm trials.

Start time 12:15

A STOCHASTIC MODELING APPROACH TO INTEGRATE DIFFERENT TYPES OF GENETIC INSTABILITY

Guangquan Li, Dr. Mark Little and Prof. Paolo Vineis
Department of Epidemiology and Public Health, Imperial College London, UK

Keywords: *Stochastic modeling, Carcinogenesis, Genetic instability*

Carcinogenesis, the progression of cancer, is a complex biological process where heredity, ageing and environmental exposures acting upon the human genome. Mathematical models attempt to connect various risk factors to the onset of cancers in order to understand the underlying neoplastic mechanisms. A number of carcinogenesis models have been proposed assuming cancers develop from a single pathway through a multistage stochastic process. Recent biological evidence suggests that carcinogenesis may involve genetic instability (GI) which allows cells to progress into cancers through different carcinogenic pathways. Furthermore, the emergence of GI accelerates the cancer process by enhancing the mutation rates. At least two types of GI have been observed in colorectal cancers: chromosomal instability (CIN) and microsatellite instability (MIN). We proposed a model that preserves features of the existing models, yet it goes beyond in being capable of describing various types of genetic instability. The model assumes different forms of GI to be mutually exclusive, as it is known to be the case for CIN and MIN. Preliminary model assessment was performed using a set of colon cancer data; the model provided satisfactory fits to the data. It also predicted that the mutation rates associated with GI were higher than those of the background rates, which is exactly the characteristic of genetic instability.

Start time 12:40

STATISTICAL ANALYSIS OF HIGH CONTENT SCREENING DATA

Richard Jacques
University of Sheffield, UK

The current paradigm for the identification of candidate drugs within the pharmaceutical industry typically involves the use of high throughput screens. A high

throughput screen with automated imaging platform allows a large number of compounds to be tested in a biological assay in order to identify any activity inhibiting or activating a biological process. From each of the assays run through a high throughput screen a high content screen image is produced which can be analysed using advanced imaging algorithms to produce a set of variables which reflect the observed activity of the cells within the image.

Statistical approaches have been developed that enable identification of biologically active compounds from high throughput screens using a single parameter. However, approaches for multi-parametric selection are still in their infancy. Furthermore, proper exploitation of the information contained within each high content screen image will enable more refined compound selection. This work is concerned with applying multivariate techniques to develop a multi-parametric approach.

14.3.9 Session 6a

Session Room 203 Chair: Chiara Mazzetta

Start time 14:15

VARIANCE STABILIZATION AND NORMALIZATION FOR ONE-COLOR MICROARRAY DATA

Efthimios Motakis
University of Bristol, UK

Keywords: *Generalized Logarithm Transformation, Haar-Fisz Transformation, Variance Stabilization*

Working on cDNA microarrays, Rocke and Durbin (2001) and Huber et al (2002) observed that the variance of replicated microarray data increases as their mean increases, thus resulting in heterogeneous spot intensities. Currently, this mean-variance dependence is modelled by using the two-component model. The authors suggested a logarithmic-based transformation scheme that stabilizes the variance of one-color microarrays intensities by simply estimating a set of parameters. Instead of considering a logarithmic-based transformation, we perform variance stabilization by a Multiscale approach that is based on Haar wavelets and on Fisz (1955) work in the distribution of a function of two independent random variables. Our method is called Data-Driven Haar-Fisz transformation for Microarrays (DDHFm) and is an extension of a previous work of Fryzlewicz and Nason (2004). Application of the DDHFm algorithm on simulated and real cDNA data proves the superiority of our method over the existing variance stabilization approaches.

References

[1] Fisz, M. (1955). *The limiting distribution of a function of two independent random variables and its statistical application*, Colloquium Mathematicum, 3, 138−146.

[2] Fryzlewicz, P. and Nason, G.P. (2004). *A Haar-Fisz algorithm for Poisson intensity estimation*, J. Comp. Graph. Stat., 13, 621−638.

[3] Huber, W., Von Heydebreck, A., Sultmann, H., Poustka, A. and Vingron, M. (2002). *Variance stabilization applied to microarray data calibration and to the quantification of differential expression*, Bioinformatics, 18, S96−S104.

[4] Rocke, D.M. and Durbin, B.P. (2001). *A model for measurement error for gene expression arrays*, J. Comp. Biol., 8, 557−569.

Start time 14:40

STATISTICAL METHODOLOGY FOR WHOLE GENOME ASSOCIATIONS SCANS FOR COMPLEX DISEASES

Fay Hosking, Prof Peter Green, Dr. Jonathan Sterne and Prof. George Davey Smith
University of Bristol, UK

Keywords: *Whole genome association scans; genome wide association scans; Bayesian modelling*

Many human diseases are known to "run in the family" and are thus likely to have genetic factors. Genetic causes of some human diseases have been found, but generally only for rare, so called "simple" diseases. In order to find the genetic causes of many other "complex" diseases large case-control studies will be required. Such studies will involve several thousand cases and controls and with tens of thousands of genetic variables in the form of Single Nucleotide Polymorphisms (SNPs) being measured on each individual. Due to the rapid increase of biotechnical knowledge and simultaneous decrease in cost such "genome wide" studies are now possible.

This presentation will give a basic introduction to human genetics before using existing analysis methods: 2×2 tables yielding log odds ratios estimates as a starting point to explore the challenges of both complex diseases and whole genome association studies. Simple graphical analysis will be shown as well as current methodological work. This current work is Bayesian in nature and involves the utilisation of hidden Markov Models to obtain inference about the position of the causal variant in a short section of genetic data.

Start time 15:05

BAYESIAN ANALYSIS OF DESIGNED MICROARRAY EXPERIMENT

Ximin Zhu[1] and Prof. Ernst Wit[2]
[1] *University of Glasgow, UK*
[2] *Lancaster University, UK*

Keywords: *Microarray, Bayesian Analysis, Gene Expression, Loop and Reference Design, MCMC, Simulation*

Nowadays some practitioners are beginning to perform more complex cDNA microarray experiments with larger number of samples and conditions and this brings with it a challenge how to design the experiments. Reference and loop design are two available choices. Although the latter has obvious theoretical advantages the former is more popular. The main reason for it is that it is relatively easy and intuitive to analyze reference design but not loop design. We propose a general framework for a Bayesian analysis of designed microarray experiment to solve the problem. We illustrate our method by means of performing a logistics regression within a loop designed experiment. MCMC technique utilizing the Metropolis-Hastings algorithm is used for computational implementation. We show that by using our framework a loop design gives more accurate result than a reference design if both experiments involves the same number of microarrays and generally a loop design needs less number of microarrays to achieve a similar accurate result than a reference design.

14.3.10 Session 6b

Session Room 214 Chair: Cian Reynolds

Start time 14:15

BAYESIAN INFERENCE FOR HEALTH STATE UTILITIES USING PAIR-WISE COMPARISON DATA

Theresa Cain
University of Sheffield, UK

Keywords: *Bayesian Inference, Health Economics, Utility*

This topic concerns medical decision making by health care providers. When evaluating a treatment, it is important to consider cost-effectiveness and hence it is necessary to value each possible health outcome that could result from the treatment. A utility function would provide such a value and the utility is required for each of the possible health states. To define all the health states, a descriptive system such as the EuroQol is used and the associated utilities are usually evaluated by methods such as the Standard Gamble and Time Trade-off. There are issues regarding what questions can realistically be answered using these methods and how precise a value a respondent can provide. A possible solution is to ask respondents to make pair-wise comparisons which involves easier questions but more complicated analysis. This project develops a Bayesian approach for the estimation and reporting of uncertainty about the population utilities given pair-wise comparison data.

Start time 14:40

GAMBLING AND FINANCIAL MARKETS

Gillian Golden
Trinity College, Dublin, Ireland

In recent years there has been exponential growth in the gaming industry worldwide. More and more people are gambling on a professional and semi-professional level. This has let to the use by gamblers of techniques and strategies that were tradionally the forte of professionals trading in financial markets. An overview is given of the different types of financial and gambling products available to the investor, and some similarities and differences between gambling and financial markets are discussed. The question of how one may place a value on a bet is considered. An example illustrates some of the ideas presented, and finally a possible risk management strategy for gamblers is outlined.

Start time 15:05

Simulation Methods for the Extended Pure Jump Processes with Applications to Bayesian Nonparametric Statistics

Tso-Jung Yen
Department of Mathematics, Imperial College London, UK

In this report, we discuss some simulation methods for the extended pure jump processes. The extended pure jump processes can be seen as an extension of the pure jump processes driven by Lévy jumps. From this point of view, simulation methods can be established either via the compound Poisson process construction, or the infinite series construction. We apply the methods to simulate gamma processes and make some comparisons. We also give two case studies by applying these methods to simulate posterior distributions of the pure jump processes under Bayesian nonparametric framework.

14.3.11 Session 6c

Session Room 325 Chair: Keith Harris

Start time 14:15

STATISTICAL MODELLING AND INFERENCE FOR RADIO-TRACKING

Keith Harris
University of Sheffield, UK

Keywords: *Radio-tracking, animal movement, bivariate Ornstein-Uhlenbeck processes, continuous time threshold autoregressive models, Bayesian inference, hitting time densities*

Radio-tracking is a well-established tool in ecological research for collecting data about the location over time of animals. Hence, statistical models of animal movement are needed to help us interpret data collected in this manner. There are many possible approaches to analysing such data when the observations are sufficiently separated in time that they can be regarded as statistically independent. However, when successive observations are dependent, which is usually the case with radio-tracking data, there is only one standard approach. This method models animal movement using a bivariate Ornstein-Uhlenbeck diffusion process. However, this model has been criticised for being an unrealistic description of animal movement behaviour. In particular, it fails to take into account the fact that animals may move differently in different types of habitat. My research is focussing on how this problem can be overcome, that is, how spatial heterogeneity can be incorporated into models for animal movement. The approach I am pursuing is to extend a class of models called continuous time threshold autoregressive models into two dimensions. In my talk I will define and illustrate my proposed model for animal movement. I will also discuss how the statistical inference for such models can be tackled in the simplest case of there being a single linear threshold. Even in this case an approximation is needed to calculate the likelihood contribution from observations that are not in the same region as the previous observation. An approximation is needed because the likelihood contribution in these cases involves two expressions that are closely related to hitting time densities of bivariate Ornstein-Uhlenbeck processes to a linear boundary, which cannot be evaluated explicitly. The approach will be illustrated using data on the movements of wood mice.

Start time 14:40

TOWARDS MODELLING SURVIVAL AND PREDATION OF

Wild Atlantic Salmon on a Northern Scottish River

Paul Birrell
University of Cambridge, UK

Keywords: *Cormack-Jolly-Seber models, ecosystem modelling, MCMC*

Over the last twenty years, the numbers of wild adult salmon returning to Scottish rivers has fallen markedly. Whilst there is some thought this may be part of a cyclical pattern, the risks incurred in such an assumption are too great as the salmon fisheries are of huge economic importance to regions such as the Moray Firth.

One way to tackle such a problem is through the management of predator populations in such a way that these populations are done no lasting harm. This motivates studies into salmon mortality, whose findings fisheries managers can utilise in the formulation of effective predator management strategies.

The focus here is on a tag-recapture study of migrating juvenile salmon. Through a re-paremeterisation of a Cormack-Jolly-Seber model with integrated recoveries, we can fit and select optimal models which incorporate uncertainty regarding how the salmon time their migrations and which fish are selected by predators, carrying out our analysis from within the Bayesian paradigm due to the ease with which results can be interpreted by fisheries managers.

Start time 15:05

Bayesian techniques used to investigate issues of contemporaneity between archaeological and environmental records

Lynsey J. McColl
University of Sheffield, UK

Keywords: *Bayesian chronology building, radiocarbon dating, temporal uncertainty*

Answering questions such as whether two events in the past occurred at the same time are fundamental to both the archaeological and environmental communities. The evidence to date these events is often in the form of radiocarbon determinations. Due to the uncertainties involved in radiocarbon dating, the true calendar age of a sample being dated cannot be ascertained. Instead, a probability density which represents an estimate of the calendar age is produced. Such estimates are on

a continuous time scale, thus the probability that an object dates to a specific moment in time is zero. This means that all questions relating to the date of a single object or to the relationship between the dates of multiple events, must focus on intervals of time rather than moments in time.

The focus of this talk will be to describe the types of Bayesian statistical models that can be used to investigate questions of contemporaneity. Although I aim to develop generic tools that can be used throughout both disciplines, I will illustrate the concepts using a case study on ancient hunting implements discovered in Southern Yukon, Canada. These were found to be examples of throwing-dart (atlatl) and bow-and-arrow technology. Using the radiocarbon dates obtained from these implements, it has been hypothesized that atlatl technology was abruptly replaced by bow-and-arrow technology in the region. In addition to this archaeological data, the White River volcanic eruption that occurred in this region has been recently redated, highlighting a possible concordance between the timing of the eruption and the technological transition. Both hypotheses will be tested by building formal statistical models to estimate the timings of the technological transition and the eruption, and using these estimates, the probability of contemporaneity can be investigated.

14.3.12 Session 6d

Session Room 326 Chair: Neil Henderson

Start time 14:15

USING SAS ENTERPRISE GUIDE FOR DATA MANIPULATION, ANALYSIS AND PRESENTATION OF RESULTS

Isabel Sassoon
Senior Consultant, SAS, UK & Ireland

Start time 14:40

THE ROLE OF STATISTICIANS AT PFIZER

Peter Colman
Pfizer Global Research & Development, Sandwich, CT13 9NJ, UK

Keywords: *Pfizer, The role of statisticians, Statistics, Drug development, Pharmaceutical industry*

The aim of this presentation is to introduce students to the role of statisticians within the pharmaceutical industry, specifically at Pfizer. The reasons why statisticians are becoming increasingly important in the drug development process will be described along with the major roles and responsibilities of statisticians at Pfizer. The key skills required for such a role will be outlined and the career progression opportunities available will be summarised. The talk will highlight the many other disciplines that statisticians at Pfizer liaise with and the areas within the drug-development process in which statistics is involved. The talk will conclude with an overview of the role.

Start time 15:05

TOWARDS MULTIVARIATE METHODS IN INDUSTRIAL APPLICATIONS

Matt Coates
Senior Statistical/Technical Consultant, StatSoft Ltd, UK

Keywords: *Statistical methods, Quality improvement, Manufacturing industry*

For many years, the application of statistical methods to quality improvement in manufacturing industry has been restricted largely to relatively simple univariate methods. Progress can be notoriously slow and many companies are still only just beginning to apply these traditional methods. However, others are finding a range of applications for more sophisticated methods including PCA/PLS, Neural Networks and Multivariate SPC. An overview of some of the methods that are becoming increasingly common in industry will be presented, with case-study based examples.

15 Poster abstracts by author

THE INDIRECT ESTIMATION OF ELDERLY MIGRANT FLOWS IN ENGLAND AND WALES

Guy Abel
University of Southampton, UK

Keywords: *Internal Migration, Indirect Estimation, Expectation Maximisation, Iterative Proportional Fitting*

Internal migration statistics in England and Wales may be obtained from the Census and the Office of National Statistics (ONS), whose figures are derived from the NHSCR and Health Authority Patient Registers. The regularity and detail of these data sources differ. Indirect estimation techniques illustrated in this poster are applied to bridge these differences allowing the regular total in-and-outflow data of the ONS to duplicate the detailed desegregation of flows found in the Census. Two indirect estimation techniques are compared; Iterative Proportional Fitting (IPF) and the Expectation Maximisation (EM) algorithms to calculate migration counts by origin and destination of Local Authority Districts grouped according to Office of National Statistics categorisations. Both sets of estimates are shown to be equivalent. These techniques are then expanded to predict migration counts by a third discrete variable, Limiting Long Term Illness. These estimates are shown to produce different results due to a lack of information regarding a sufficient statistics for a third discrete variable in the ONS in-and-outflow data.

TRIPLE SYSTEM ESTIMATION OF CENSUS UNDER-ENUMERATION

Bernard Baffour
Division of Social Statistics, University of Southampton, UK

Keywords: *capture-recapture, dual system estimation, triple system estimation, log-linear models*

The theoretical framework of triple system estimation is based on the capture-recapture methodology employed in the measurement of biological populations. This estimates the unknown size of a population by firstly capturing a set of individuals. Further subsequent captures are taken at later periods. The possible capture histories are represented by the cells of a 2^k contingency table, where k is the number of captures. The contingency table will have one cell missing the population missed in all k captures. An estimate of this cell is ascertained by log-linear modelling and hence the total population can be estimated.

In the 2001 UK Census, dual system estimation was used to estimate the total population size including those missing. The two systems used were the Census and the Census Coverage Survey. One of the assumptions underlying dual system estimation is that the Census and Survey are independent; correlation bias occurs when there is dependence. Thus data from a third source Administrative Records has been proposed as a means of augmenting for correction bias. In a census context, this is synonymous with a three-system capture-recapture approach. Data from the three lists the Census, Census Post-enumeration Survey and Administrative Records are assumed to form an incomplete 2^3 contingency table and the expected values of the observed cells can be estimated using log-linear models.

The work presented in this poster seeks to estimate the unobserved cell, corresponding to the individuals missed by all three lists, from the observed cells within the log-linear framework. This is demonstrated for a range of dependency assumptions as well as different coverage rates.

On The Relationship Between Quasi-Stationarity and Hazard Rates for Continuous-Time Markov Chains

Crossman, R and Dr. Coolen-Schrijner, P
Durham University, UK

Keywords: *Hazard Rate, Quasi-Stationary Distribution, Birth-Death Process, Stochastically Greater*

With this poster we study the relationship between hazard rates (a mathematical concept useful in several areas, perhaps most obviously reliability theory) and the quasi-stationary distribution, with specific reference to birth-death processes. The conditions on existance of a quasi-stationary distribution (QSD) for a finite-space Markov chain with one absorbing state, and how to find it if it does exist, are well known results. So too is the knowledge that this distribution leads to a hazard

rate which is constant over time. Following lines of enquiry posited by Aalen and Gjessing([1]) we consider the possibility that the behaviour of the hazard rate of a process over time is dictated in some way by the relationship between the QSD and the initial distribution of a specific case. Focussing on birth-death processes, we consider how choosing initial distributions that are stochastically greater (respectively less than) the QSD appears to lead to hazard rates that are bounded from above (respectively below) by the value of the constant hazard rate associated with the QSD.

References

[1] Aalen, O.O. and Gjessing, K. (2001) Understanding the Shape of the Hazard Rate: A Process Point of View. *Statistic. Science. Vol. 16.* **1**, 1-22.

A Class of Bayesian Space-Time State-Space Models for Environmental Forecasting

Swarup De and Alvaro E.Faria
The Open University, UK

A spatio-temporal state-space Bayesian model is proposed for the statistical modelling of radioactivity deposition on the ground. A Gaussian hierarchical form of this model is implemented to produce fast estimates of ground contamination levels which combine information from atmospheric dispersal with ground measurements of deposited radioactivity. It includes the treatment of multiple nuclides as well as wet and dry components of radioactive deposition, total energy and isotope decay. The proposed model is appropriate for Markov random field processes in space and handles uncertainties associated with the predictions from an atmospheric dispersal model, the measurements and the spatial interpolation.

Real data of radioactivity deposition from the 1986 Chernobyl accident in southern Germany are assimilated by the model and the effects of various isotropic spatial correlation structures are investigated.

Safety and Efficacy: Which Outcomes Should Independent Data Monitoring Committees Assess?

Laura J Gray and Philip MW Bath
Institute of Neuroscience, University of Nottingham, UK

Keywords: *Clinical trials, Data monitoring committees*

Introduction: The primary role of a trials Data Monitoring Committee (DMC) is to ensure the safety of enrolled patients. Conventionally, safety is monitored by comparing death and stroke specific events such as deterioration, recurrence and intracranial haemorrhage between treatment groups. Although the DMC also may have the remit for monitoring efficacy (end of trial functional outcome), trialists may choose for this to be excluded from assessment by the DMC. We hypothesised that functional outcome is a powerful measure of safety and tested this in a semi-quantitative research synthesis.

Methods: Acute stroke trials with a negative outcome (active treatment worse than control) or which were stopped prematurely on the grounds of safety were sought from electronic searches of The Cochrane Library, PUBMED and www.strokecenter.org. Information on early and late death, impairment and functional outcome were recorded for each trial, as was the presence of a DMC. The results of significance tests for each outcome were ranked within each trial to determine which was most efficiently statistically in detecting hazard.

Results: 11 negative trials were included (AASI, ASK, ASSIST, EAST, INWEST, MAST-E, MAST-I, DCLHb, sipatrigine, STIPAS, TESS II); insufficient data were present for 2 other trials. The most efficient outcomes were: early death, 4 trials; late death, 1 trial; late death or disability, 5 trials; late death or impairment, 1. Early death was insensitive to hazard in all 5 trials where late death or dependency was most sensitive. Death was the most sensitive measure of hazard in the three trials of thrombolysis. Two trials (both phase II) did not report the presence of a DMC.

Conclusions: All multicentre trials should have an independent DMC. End of trial functional outcome should be included in all assessments of safety, whether or not efficacy itself is being assessed.

Circular time series with application to protein conformations

Gareth Hughes, Kanti V. Mardia and Charles C. Taylor

University of Leeds, UK

Keywords: *Circular time series, Von Mises distribution*

The 3-dimensional conformation of proteins is an active research area that has received much attention. Much of the statistical modelling of these conformations has been investigated with bivariate circular models (see for example Mardia et al. [1], Singh et al. [2]). In this poster we consider three first order autoregressive time series models for univariate directional data which have the potential for extension to higher orders and, ultimately, bivariate directional time series models.

References

[1] Mardia, K. V., Taylor, C. C. and Subramaniam, M (2003). Application of circular distributions to conformational angles in protein. *Proceedings in Stochastic Geometry, Biological Structure and Images*, 149–152. Edited by R.G. Aykroyd, K.V. Mardia and M.J. Langdon. Leeds University Press.

[2] Singh, H., Hnizdo, V. and Demchuk, E. (2002). Probabilistic model for two dependent circular variables. *Biometrika*, **3**, 89, 719–723.

[3] Downs, T. D. and Mardia, K. V. (2002). Circular regression. *Biometrika*, **3**, 89, 683–697.

[4] Mardia, K. V. (1975). Statistics of Directional Data. *J. R. Statist. Soc. B*, **3**, 37, 349–393.

[5] Rivest, L. P. (1987). A distribution for dependent unit vectors. *Communications in Statistics A*, **17**, 461–483.

RELATIONSHIP AMONG RKHS, SPLINES AND KRIGING

Ali Mohammadian Mosamam and John T Kent
The Department of Statistics, The University of Leeds, UK

Keywords: *Intrinsic process, Conditionally Positive Definite, Reproducing Kernel Hilbert Space(RKHS), Kriging, Splines*

The relationship among intrinsic kriging, splines and RKHS is described and similarities and differences are considered. We also have extended the theory of RKHS from normed spaces to semi-normed spaces which corresponds to an intrinsic rather than ordinary process.

CAN THE INTERNET BE UNSTABLE?

Cian Reynolds
Heriot-Watt University, UK

Keywords: *Internet Traffic, File Transfers, Probability Theory*

We look at the mathematics of bandwidth allocation and control in networks with simultaneous resource requirements (as is required in communications networks which require file transfers like part the Internet). We are interested in the problem of stability. A control strategy (bandwidth allocation) is stable if the corresponding network is positive recurrent.

A network with given capacities and input rates is feasible if each resource has sufficient capacity for those call types which use it, and a control strategy is Pareto efficient (PE) if no call type may be allocated more bandwidth without decreasing that allocated to other call types. For every feasible network there exists at least one PE control strategy which is stable. However, it is well-known that the simultaneous resource requirement for calls of each type means that, for all but the simplest networks, not every PE control strategy is stable. This phenomenon is known as entrainment.

We attempt to develop some characteristics of stable control strategies.

References

[1] Clarke B. and Gustafson P. (1998) On the overall sensitivity of the posterior distribution to its inputs. *Journal of Statistical Planning and Inference*, **71**, 137-150.

[2] Mugglin A.S., Cressie N. and Gemmel I. (2002) Hierarchical statistical modelling of influenza epidemic dynamics. *Statistics in Medicine*, **21**, 2703-2721.

Assessing prior sensitivity of a hierarchical space-time model using MCMC

Jennifer Richardson
Durham University, UK

Keywords: *MCMC, prior, sensitivity*

The introduction of MCMC methods has made Bayesian analysis of more complex models possible. However, with complex models, prior distributions can often be decided on in a casual way. Since there is no guarantee that the prior distributions we start with are reasonable, there is a need to see if making changes to the prior has an effect on the outcome. Clarke and Gustafson [1] develop a method for quantifying the local sensitivity of the posterior distribution to simultaneous changes in the three inputs - the model, the data and the prior distribution. This poster uses the method of Clarke and Gustafson [1] and the hierarchical space-time model of Mugglin *et al.* [2] and considers the sensitivity of the posterior distribution to changes in the prior for situations where the posterior distribution is not known exactly but instead we have a sample from it - i.e. MCMC output.

Statistical Models of SNP Variation Applied to Admixture Analysis

Maarya Sharif and Dr Vincent Macaulay
University of Glasgow, UK

Keywords: *Admixture, Single Nucleotide Polymorphism(SNPs), SNP ascertainment, Genetic Drift, Markov Chain Monte Carlo (MCMC), HapMap*

Recently admixed populations provide a potential resource for mapping the mutations responsible for common diseases. Thus methods to characterise such populations are urgently required. Large scale, international projects have identified $> 10^6$ single nucleotide polymorphisms (SNPs) which are single base polymorphisms across the genome. We develop a hierarchical model of admixture for such data that incorporates the process of SNP discovery/ascertainment and estimates the parameters of the model which describe genetic drift and the level of admixture via Markov Chain Monte Carlo sampling. Crucially, we find that estimates are sensitive to the ascertainment process and the way in which it is modelled. A cost study of data from the International HapMap project reveals that where low genetic drift occurs between 'parental' populations prior to admixture, investing in allele copies

rather than SNP loci provides better admixture estimates. Whereas in the presence of high genetic drift between parental populations prior to admixture, irrespective of where cost is directed, little difference occurs in estimates.

Factor Analysis of Urological Data

Graham Smith, Dr Alison Gray and Prof. Chris Robertson
University of Strathclyde, UK

Keywords: *Urological Data, Factor Analysis*

This work concerns the epidemiology of Lower Urinary Tract Symptoms (LUTS), which are distressing conditions affecting the quality of life. There are many unanswered questions on the prevalence and impact of LUTS in different populations. Data was provided by Health Protection Scotland (HPS), and arises from a questionnaire used in a multi-centre survey of men and their partners from four cities: Boxmeer (Netherlands), Auxerre (France), Birmingham (UK) and Seoul (Republic of Korea). The questionnaire is split into seven different sections on different topics, and different blocks of questions were included from different sources.

The aim of the work was to validate the questionnaire using factor analysis. Questions on urological symptoms and their impact were analysed to determine whether the questions were in the appropriate block of questions and if there was redundancy in questions from different sources. The analysis was carried out for each country separately and also for all countries combined. The results of this analysis will be presented.

Estimating treatment causal effect under non-random assignment

Maria Vazquez
Department of Statistics, University of Warwick, UK

Keywords: *Observational studies, propensity score, Horvitz-Thompson estimator, Marginal Structural method*

The estimation of treatment causal effect in observational studies must compensate for the treatment assignment mechanism. The Horvitz-Thompson estimator does it

by weighting by the inverse of the propensity score. However, this estimator is not invariant under translations in the observations. An imputational estimator is also available for the treatment causal effect, but it makes strong assumptions about the fitting of a linear regression. We show how an estimator called Marginal-Structural overcomes both problems. And we present a variance focused comparison of the three estimators through an analysis of a data set.

MODELLING THE BRAZILIAN MONTHLY EMPLOYMENT SURVEY

Alinne Veiga, Peter Smith and James Brown
University of Southampton, UK

Keywords: *Panel Survey, Labour Force Survey, Longitudinal Data, Clustering, Multilevel Modelling*

The Brazilian Monthly Employment Survey (PME) is a household sample survey created in 1980 and recently revised. With a stratified two-stage cluster design, it is representative of six Brazilian metropolitan areas. It has a panel design, where each selected household remains in the sample in four consecutives waves returning for another set of four consecutive interviews after eight months not in the sample. Households in all eight waves are included in the sample every month; 75% of the sample are matched in two consecutive months; and 50% in two consecutive years. The latest revision of this survey included redefinition of certain economic concepts, inclusion of some additional variables and weight adjustment for non-response. This poster presents some of the main issues to be accounted for when analysing data from the 2004 Brazilian Monthly Employment Survey with the main objective to identify appropriate methods to be applied when modelling these data. The analysis of longitudinal data should include methods that account for complex data structure such as those presented in the Brazilian labour force survey. One approach to be considered is to model longitudinal data in a multilevel framework, where repeated measures are level one units and individuals the level two units. Sample design are accounted for by the inclusion of cluster variables representing higher levels in the hierarchy with random effects, and dependency of the repeated individual measurements are accounted for by the inclusion of a serial correlation component in the models. Other issues to be considered include different estimation procedures, assumptions on the drop-out mechanism, methods for unbalanced data, weighting and stratification.

Epidemiology of Mycobacterium Bovis in a Badger Population

Neil J. Walker, Dr. Richard J. Delahay and Dr. Chris L. Cheeseman
Central Science Laboratory, York, UK

Keywords: *M. Bovis, social group, incidence, spatio-temporal*

This Veterinary Epidemiology study is based on data between 1982 and 2000 from a mark-recapture study on a wild badger population, located in Woodchester Park, Gloucestershire. This is a high-density population where badgers live in clearly defined social groups. Every time an animal is captured, a suite of clinical samples is submitted for M. Bovis (bovine TB) culture testing along with a blood sample (for antibody detection) and a number of demographics recorded. This provides valuable information on individual life histories and aids our understanding of M. Bovis Epidemiology in the study population.

In order to represent progressive states of infection, 3 different definitions of M. Bovis incidence are used. Exposure (first positive antibody test), Excretor (first positive culture test), Super-Excretor (more than one positive culture test). Mixed models (GLMM) were used to model the probability of incidence (as defined above) as a function of a number of demographics and social-group level variables.The significant explanatory factors were primarily individual level rather than group-level. Cubs were significantly more likely to become exposure incident, whereas adults were more likely to become super-excretor. Marked seasonal differences were observed with incidence generally higher in spring and lower in summer. A strong temporal trend was observed over the study period.

The random effect parameter estimates for individual badgers and social group suggested significant differences at these levels. The next step is to model the spatio-temporal movement of M. Bovis at the social group level.

An adaptive empirical Bayesian thresholding procedure for analysing microarray experiments

R.E. Walls, S. Barber, J.T. Kent and M.S. Gilthorpe

University of Leeds, UK

Keywords: *Bayesian, microarray experiment, thresholding*

A typical microarray experiment attempts to ascertain which genes display significant differential expression between various samples, often comparing the levels of expression for many thousands of genes simultaneously. The issues of multiple testing associated with such high-dimensional data sets, often combined with insufficient replication, makes the conventional method of significance testing infeasible. Here we demonstrate a highly effective empirical Bayesian thresholding technique for the detection of a few pieces of useful information 'hidden' in a sequence of noisy data, most of which contains no useful information, and explore the application of this method to microarray data, where the problem is to detect the few 'interesting' genes amongst many 'uninteresting' ones. Assuming that most genes will not be differentially expressed between the samples, we apply the method to the sequence of log-ratios, $\log(T_i/C_i)$ for $i = 1, \ldots, n$, where T_i and C_i are the levels of expression for gene i in samples T and C respectively. The method works by estimating a threshold value: any genes with an absolute log-ratio below the threshold are set to zero and concluded to be not differentially expressed; those genes with an absolute log-ratio above the threshold value are said to be truly differentially expressed. We illustrate the results for various data sets and compare to the common fold change approach and multiple significance testing.

STATISTICS IN CHEMOMETRICS

Carmen Ybarra-Moncada and Alan Kimber
The University of Reading, UK

Chemometrics provides methodologies to produce not only rapid and accurate predictions but also noninvasive and nondestructive quality assessments. This research focuses on the most common calibration methods in chemometrics and the algorithm Least Square Support Vector Machines (LS-SVM), with particular emphasis on application in real and simulated data to predict quality attributes of each individual fresh fruit.

The application of these procedures centres on noninvasive evaluation of total soluble solids (TSS), by means of Near infra Red (NIR) spectroscopy data provided by the "Plant Sciences Group" at Central Queensland University Rockhampton, Australia. One case is based on NIR measurements of twenty intact fresh fruit samples of peaches. The spectra for each are measured in absorbance mode scanned from 309 to 1150 nanometers (nm) in 3 nm steps, giving 256 independent variables.

Owing to the availability in a short time of thousands of spectral predictors, the

problems to solve are essentially: variable selection, collinearity, non-linearity, outliers, over-fitting and under-fitting.
In order to deal with some of these issues, methodologies from machine learning have been proposed such as LS-SVM, largely investigated as regression and classification computing problems of supervised learning. The application of new methodologies like LS-SVM may be limited if they have not been validated to meet some statistical requirements, and requirements from the areas where the novel procedures are applied. Therefore, LS-SVM deserves particular analysis.

Name	University	E-mail Address
Mr Guy Abel	University of Southampton	gja104@soton.ac.uk
Mr John Aitken	University of Glasgow	johna@stats.gla.ac.uk
Mr Ali Al-Alwan	Durham University	a.a.s.al-alwan@durham.ac.uk
Mrs Wafa Al-Fwzan	University of Strathclyde	wafa@stams.strath.ac.uk
Mr Bander Al-Zahrani	University of Newcastle Upon Tyne	bander.al-zahrani@ncl.ac.uk
Miss Mercedes Andrade Bejarano	University of Reading	m.andrade@reading.ac.uk
Mr Bernard Baffour	University of Southampton	bb1304@soton.ac.uk
Mr Dan Bailey	University of Bristol	dan.bailey@bristol.ac.uk
Miss Eleni Bakra	University of Glasgow	eleni@stats.gla.ac.uk
Miss Sarah Barry	University of Glasgow	sarah@stats.gla.ac.uk
Dr Erik Baurdoux	Heriot-Watt University, Edinburgh	erikb@ma.hw.ac.uk
Ms Gundula Behrens	University of Bath	g.r.behrens@maths.bath.ac.uk
Mr Miguel Belmonte	University of Warwick	m.a.g.belmonte@warwick.ac.uk
Mr Paul Birrell	University of Cambridge	pjb51@statslab.cam.ac.uk
Miss Elisa Bonvini	University of Reading	E.Bonvini@reading.ac.uk
Mr Partha Bose	Winton Capital Management	p.bose@wintoncapital.com
Mr Chris Brignell	University of Nottingham	pmxcjb@nottingham.ac.uk
Miss Theresa Cain	University of Sheffield	T.Cain@Sheffield.ac.uk
Mr Lucas Carbonaro	University of Oxford	carbon@stats.ox.ac.uk
Miss Vanessa Cave	University of Cambridge	V.Cave@statslab.cam.ac.uk
Mr Kan Chen	University of Strathclyde	kan@stams.strath.ac.uk
Miss Hathaikan Chootrakool	University of Newcastle Upon Tyne	hathaikan.chootrakool@ncl.ac.uk
Ms Norma Coffey	University of Limerick	Norma.Coffey@ul.ie
Miss Solange Correa	University of Southampton	scorrea@soton.ac.uk
Miss Maria Costa	University of Warwick	M.J.Costa@warwick.ac.uk
Mr Richard Crossman	Durham University	r.j.crossman@durham.ac.uk
Mr Fraser Daly	University of Nottingham	pmxfad@nottingham.ac.uk
Mr Swarup De	Open University	S.De@open.ac.uk
Miss Kerry Dwan	University of Liverpool	kerry.dwan@liv.ac.uk

Name	University	E-mail Address
Miss Emma Eastoe	Lancaster University	e.eastoe@lancaster.ac.uk
Ms Entisar Elgmati	University of Newcastle Upon Tyne	entisar.elgmati@ncl.ac.uk
Mr Kim Evans	University of Nottingham	pmxke@maths.nottingham.ac.uk
Mr Richard Everitt	University of Bristol	rich@redimps.com
Mr Jude Eze	University of Glasgow	jude@stats.gla.ac.uk
Mr Michael Fahey	University of Cambridge	mtf29@cam.ac.uk
Mr Ally Fearnside	University of Glasgow	alastair@stats.gla.ac.uk
Miss Claire Ferguson	University of Glasgow	claire@stats.gla.ac.uk
Mr Guy Freeman	University of Birmingham	freemang@medmail2.bham.ac.uk
Mr John Fry	University of Sheffield	J.Fry@sheffield.ac.uk
Mr Roger Gill	Winton Capital Management	r.gill@wintoncapital.com
Mr Rainer-Georg Goeldner	University of Reading	r.g.goeldner@reading.ac.uk
Miss Gillian Golden	Trinity College, Dublin	goldeng@tcd.ie
Mr Andrew Golightly	University of Newcastle Upon Tyne	a.golightly@ncl.ac.uk
Mr Alex Goodwin	University of Leeds	alexg@maths.leeds.ac.uk
Miss Lindsay Govan	University of Glasgow	lindsay@stats.gla.ac.uk
Miss Raquel Granell	University of Bristol	Raquel.Granell@bristol.ac.uk
Miss Laura Gray	University of Nottingham	laura.gray@nottingham.ac.uk
Miss Katharina Gruenberg	Lancaster University	K.Gruenberg@Lancaster.ac.uk
Miss Caroline Haig	University of Glasgow	caroline@stats.gla.ac.uk
Miss Lisa Hampson	University of Bath	ma1lvh@maths.bath.ac.uk
Miss Kelly Handley	University of Nottingham	pmxkh@nottingham.ac.uk
Mr Gavin Hardman	Durham University	g.a.hardman@durham.ac.uk
Mr Keith Harris	University of Sheffield	stp03kjh@sheffield.ac.uk
Mr Neil Henderson	University of Glasgow	neil@stats.gla.ac.uk
Miss Weang Kee Ho	University of Newcastle Upon Tyne	w.k.ho@ncl.ac.uk
Mr Patrick Ho	University of Warwick	P.H.K.Ho@warwick.ac.uk
Ms Fay Hosking	University of Bristol	fay.hosking@bristol.ac.uk
Mr Brett Houlding	Durham University	brett.houlding@durham.ac.uk

Name	University	E-mail Address
Mr Gareth Hughes	University of Leeds	ghughes@maths.leeds.ac.uk
Miss Alicia Huntriss	Durham University	alicia.huntriss@durham.ac.uk
Miss Gail Hutton	University of Strathclyde	gail@stams.strath.ac.uk
Miss Adriana Irawati Nur Ibrahim	University of Bath	aini20@maths.bath.ac.uk
Mr Richard Jacques	University of Sheffield	stp05rmj@sheffield.ac.uk
Mr David Jenkinson	Open University	d.j.jenkinson@open.ac.uk
Mr Chris Jewell	Lancaster University	c.jewell@lancs.ac.uk
Mr Yugang Jia	University of Bristol	yugang.jia@bristol.ac.uk
Miss Jill Johnson	University of Newcastle Upon Tyne	J.S.Johnson@ncl.ac.uk
Miss Hayley Jones	MRC Biostatistics Unit, University of Cambridge	hej27@cam.ac.uk
Mr Miland Joshi	University of Warwick	m.a.i.joshi@warwick.ac.uk
Miss Irene Kaimi	Lancaster University	i.kaimi@lancaster.ac.uk
Miss Kimberley Kavanagh	University of Strathclyde	kim@stams.strath.ac.uk
Mr Mark Kelly	Cardiff University	kellymj1@cf.ac.uk
Mr Hassan Khalil	University of Bristol	h.k.khalil@bristol.ac.uk
Dr Matthew Killeya	Winton Capital Management	m.killeya@wintoncapital.com
Mr Adam Kinnison	University of Bath	adamkinnison@yahoo.co.uk
Mr Michalis Kolossiatis	University of Warwick	michaelk_cy@yahoo.com
Mr Demetris Lamnisos	University of Warwick	D.S.Lamnisos@warwick.ac.uk
Mr Duncan Lee	University of Bath	mapdpl@bath.ac.uk
Mr GuangQuan Li	Imperial College, London	guang.li@imperial.ac.uk
Mr Shan Lin	University of Southampton	s.lin@soton.ac.uk
Mr Zhen Liu	Lancaster University	z.liu@lancaster.ac.uk
Miss Jiayi Liu	Lancaster University	j.liu9@lancaster.ac.uk
Miss Elisa Loza Reyes	University of Bath	e.loza.reyes@bath.ac.uk
Miss Anastasia Lykou	Lancaster University	a.lykou@lancaster.ac.uk
Miss Ana-Maria Magdalina	University of Limerick	ana-maria.magdalina@ul.ie
Mr Naventhan Mahadevan	University of Bristol	M.Naventhan@bristol.ac.uk
Mr Boikanyo Makubate	University of Glasgow	boikanyo@stats.gla.ac.uk

Name	University	E-mail Address
Miss Ioanna Manolopoulou	University of Cambridge	im246@cam.ac.uk
Mr Tristan Marshall	Lancaster University	t.marshall2@lancaster.ac.uk
Mr Robert Mastrodomenico	University of Reading	snr03rm@rdg.ac.uk
Ms Chiara Mazzetta	University of Cambridge	C.Mazzetta@statslab.cam.ac.uk
Miss Lynsey McColl	University of Sheffield	l.mccoll@sheffield.ac.uk
Mr Tarek Medkour	Imperial College, London	tarek.medkour@imperial.ac.uk
Mr Sang Taphou Mendy	University of Liverpool	taphou@liv.ac.uk
Mr James Miller	University of Glasgow	jamesm@stats.gla.ac.uk
Dr Rishot John Minty	Lancaster University	r.minty@lancaster.ac.uk
Ms Giuseppina Moffa	University of Bristol	giusi.moffa@bristol.ac.uk
Mr Mehdi Mohammadi	University of Newcastle Upon Tyne	mahdi.mohammadi@ncl.ac.uk
Mr Ali Mohammadian	University of Leeds	moham_am@yahoo.com
Mrs Oarabile Molaodi	University of Strathclyde	oarabile@stams.strath.ac.uk
Miss Victoria Montgomery	Durham University	v.j.montgomery@durham.ac.uk
Mr Efthimios Motakis	University of Bristol	e.motakis@bris.ac.uk
Mrs Graciela Muniz Terrera	MRC Biostatistics Unit, University of Cambridge	gm299@cam.ac.uk
Miss Emma O'Connor	University of Sheffield	e.oconnor@sheffield.ac.uk
Miss Rebecca O'Neil	Durham University	r.k.oneil@durham.ac.uk
Mr Mathias Onabid	University of Glasgow	mathias@stats.gla.ac.uk
Mr Theodore Papamarkou	University of Warwick	T.Papamarkou@warwick.ac.uk
Ms Helen Parker	University of Glasgow	helen@stats.gla.ac.uk
Miss Anne Presanis	MRC Biostatistics Unit, University of Cambridge	amp62@cam.ac.uk
Mr Malcolm Price	University of Bristol	malcolm.price@bristol.ac.uk
Mr Gareth Pritchard	University of Glasgow	gareth@stats.gla.ac.uk
Miss Vilda Purutcuoglu	Lancaster University	v.purutcuoglu@lancaster.ac.uk
Miss Stamatiki (Matina) Rassias	University of Strathclyde	matina@stams.strath.ac.uk
Mr Cian Reynolds	Heriot-Watt University, Edinburgh	cian@ma.hw.ac.uk
Mrs Jennifer Richardson	Durham University	j.l.richardson@durham.ac.uk
Miss Lauren Rodgers	University of Newcastle Upon Tyne	l.r.rodgers@ncl.ac.uk

Name	University	E-mail Address
Mr Andrew Rose	University of Southampton	ar@maths.soton.ac.uk
Mr Patrick Rubin-Delanchy	Imperial College, London	patrubdel@hotmail.com
Mr Efstathios Sakkas	University of Nottingham	pmxes1@maths.nottingham.ac.uk
Mr Jeffrey Samuel	University of Southampton	js1@maths.soton.ac.uk
Miss Pauline Sculli	London School of Economics	p.k.sculli@lse.ac.uk
Ms Maarya Sharif	University of Glasgow	maarya@stats.gla.ac.uk
Mr Chris Sherlock	Lancaster University	c.sherlock@lancaster.ac.uk
Miss Michailina Siakalli	University of Sheffield	stp05ms@sheffield.ac.uk
Mr Graham Smith	University of Strathclyde	grahams@stams.strath.ac.uk
Mr Simon Spencer	University of Nottingham	pmxss@nottingham.ac.uk
Miss Christiana Spyrou	University of Cambridge	C.Spyrou@statslab.cam.ac.uk
Miss Catherine Stewart	University of Glasgow	catherine-s@msoc.mrc.gla.ac.uk
Mrs Patricia Stone	University College London	pat@stats.ucl.ac.uk
Mr Brian Sullivan	University of Limerick	brian.sullivan@ul.ie
Ms Judy Tang	University of Bristol	j.tang@bris.ac.uk
Mr Noel Thomson	University of Glasgow	noel@stats.gla.ac.uk
Mr Andrew Thomson	London School of Hygiene & Tropical Medicine	andrew.thomson@lshtm.ac.uk
Mr Andrew Titman	MRC Biostatistics Unit, University of Cambridge	act41@cam.ac.uk
Miss Amber Tomas	University of Oxford	tomas@stats.ox.ac.uk
Mr Michael Tracy	University of Glasgow	michael@stats.gla.ac.uk
Mr Edward Tricker	Imperial College, London	etricker@gmail.com
Dr Leonardo Trujillo	University of Southampton	trujillo@soton.ac.uk
Ms Elena Tsoy	University of Warwick	E.Tsoy@warwick.ac.uk
Miss Lesley Turnbull	University of Strathclyde	lesley@stams.strath.ac.uk
Miss Despoina Vasileiou	Lancaster University	d.vasileiou@lancaster.ac.uk
Mrs Maria Vazquez	University of Warwick	strdaf@gwmail.warwick.ac.uk
Miss Alinne Veiga	University of Southampton	adcv@soton.ac.uk
Dr Jaco Vermaak	Winton Capital Management	j.vermaak@wintoncapital.com
Mr Cheran Vithanage	University of Bristol	c.vithanage@bristol.ac.uk

Name	University	E-mail Address
Mr Oleg Volkov	Queen Mary, University of London	volkovoleg@hotmail.com
Mr Adam Wagner	University of Strathclyde	adam@stams.strath.ac.uk
Mr Neil Walker	University of Bristol	n.walker@csl.gov.uk
Miss Rebecca Walls	University of Leeds	pmt1rew@leeds.ac.uk
Mr Limin Wang	London School of Economics	l.wang7@lse.ac.uk
Miss Yaan Wei	University of Bristol	maxyw@bristol.ac.uk
Miss Amanda Weir	University of Strathclyde	amanda@stams.strath.ac.uk
Mr Simon White	University Of Nottingham	pmxsw@nottingham.ac.uk
Miss Elizabeth Williamson	London School of Hygiene & Tropical Medicine	Elizabeth.Williamson@lshtm.ac.uk
Ms Elaine Wilson	University of Bristol	elaine.wilson@bristol.ac.uk
Miss I-Ding Wu	Imperial College, London	i.wu@ic.ac.uk
Mr Billy Wu	London School of Economics	wub@lse.ac.uk
Mr David Wyncoll	Lancaster University	d.wyncoll@lancaster.ac.uk
Mr Jinghao Xue	University of Glasgow	jinghao@stats.gla.ac.uk
Ms Ying Yan	Queen Mary, University of London	y.yan@qmul.ac.uk
Mr Haojie Yan	University of Bath	hy204@bath.ac.uk
Mrs Carmen Ybarra Moncada	University of Reading	m.c.ybarramoncada@rdg.ac.uk
Mr Tso-Jung Yen	Imperial College, London	tso-jung.yen@imperial.ac.uk
Mr Gang Yi	University of Newcastle Upon Tyne	gang.yi@ncl.ac.uk
Mr Chun Yin Yip	University of Southampton	cyy102@soton.ac.uk
Miss Jiangjiang Yu	University of Bath	jjy20@bath.ac.uk
Mrs Xiaohui Zhao	University of Newcastle Upon Tyne	xiao.zhao@ncl.ac.uk
Mr Ximin Zhu	University of Glasgow	ximin@stats.gla.ac.uk

Cover Picture Acknowledgements

- University of Glasgow

  `www.argyllhotelglasgow.co.uk/do_see.htm`

- City Chambers

  `www.cgcs.org.uk`

- Glasgow Science Centre (night)

  `homepages.tesco.net/~ben.craven/gsc.html`

Durham University

RSC 2007

Durham

30th Research Students' Conference in Probability and Statistics

27th–29th March 2007

rsc2007@durham.ac.uk

www.durham.ac.uk/rsc2007