



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

PENDEKATAN MODEL BAYESIAN BERTIERARKI UNTUK
PERBANDINGAN PRESTASI PELAJAR FAKULTI SAINS DAN
TEKNOLOGI, UNIVERSITI KEBANGSAAN MALAYSIA

NOOR AZRIN BINTI ABDUL RAZAK



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi

PROJEK PENYELIDIKAN YANG DIKEMUKAKAN UNTUK MEMENUHI
SEBAHAGIAN DARIPADA SYARAT MEMPEROLEH IJAZAH SARJANA
SAINS

PUSAT PENGAJIAN SAINS MATEMATIK
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2004



05-4506832



pustaka.upsi.edu.my



Perpustakaan Tuanku Bainun
Kampus Sultan Abdul Jalil Shah



PustakaTBainun



ptbupsi



PENGAKUAN

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang tiap-tiap satunya telah saya jelaskan sumbernya.

8 Jun 2004

NOOR AZRIN BINTI ABDUL RAZAK
P 27015





PENGHARGAAN

Bersyukur ke hadrat Allah s.w.t kerana dengan limpah kurnia dan izinNya dapat saya menyiapkan dan menyempurnakan penulisan projek penyelidikan bagi memenuhi sebahagian daripada keperluan kursus Ijazah Sarjana Statistik, Universiti Kebangsaan Malaysia (UKM).

Setinggi-tinggi penghargaan dan ucapan terima kasih ditujukan buat Prof. Madya Dr. Kamarulzaman Ibrahim selaku penyelia saya yang telah banyak membantu dan membimbing saya dari awal hingga ke akhir penulisan projek ini. Segala pandangan, tunjuk ajar, nasihat, teguran, kritikan, sokongan moral serta penyediaan bahan-bahan rujukan, telah sedikit sebanyak menambahkan keyakinan diri dan menjadi pendorong kepada saya dalam menyiapkan penulisan projek ini.

Ucapan terima kasih juga ditujukan buat Prof. Madya Dr. Abdul Aziz Jemain, Prof. Madya Dr. Ahmad Mahir Razali, Prof. Madya Najib Rafee dan semua pensyarah di Pusat Pengajian Sains Matematik, Universiti Kebangsaan Malaysia (UKM) serta yang sentiasa diingati, Pn. Noorizam Daud, pensyarah di UiTM Shah Alam, terima kasih di atas segala ilmu pengetahuan dan bimbingan yang diberikan sepanjang tempoh pengajian saya di universiti ini.

Penghargaan dan ucapan terima kasih ini juga turut ditujukan kepada Universiti Pendidikan Sultan Idris (UPSI) dan JPA(SLAB) kerana sudi membiayai dari segi kewangan sepanjang tempoh pengajian saya di Universiti Kebangsaan Malaysia (UKM).

Tidak dilupakan buat rakan-rakan seperjuangan yang banyak membantu, berkongsi idea, pendapat dan perasaan, terutamanya buat Norhaidah Mohd Asrah, Sabariah Saharan dan Norzehan Fazahiyah, ucapan terima kasih yang tidak terhingga buat kalian bertiga di atas segala bantuan dan dorongan serta semangat yang diberikan kepada saya untuk terus menyiapkan penulisan projek ini.

Seterusnya, penghargaan ini ditujukan khas buat insan yang sungguh istimewa kepada saya iaitu ibu dan abah. Hjh. Noraesah Awang dan Hj. Abdul Razak Ab. Rahman, serta adik-adik yang telah banyak memberi semangat, dorongan serta bantuan kepada saya dalam menyiapkan penulisan projek ini sehingga berjaya. Jutaan terima kasih di atas doa-doa yang telah ibu dan abah berikan. Akhir sekali sesungguhnya pengalaman, ilmu pengetahuan dan kerjasama yang telah saya peroleh amat saya hargai dan akan dikenang buat selama-lamanya.

Insyallah. Sekian, terima kasih.





ABSTRAK

Kajian yang dijalankan ini adalah mengenai analisis terhadap prestasi pelajar di Fakulti Sains dan Teknologi (FST) yang diukur melalui keputusan purata nilai gred kumulatif (PNGK). Kajian ini meliputi keseluruhan 1088 pelajar yang mengikuti pelbagai program yang ditawarkan oleh lima buah pusat pengajian di FST. Kajian ini tertumpu kepada perbandingan keputusan PNGK pelajar bagi setiap program dan pusat pengajian di fakulti ini serta memeringkatkan prestasi pelajar mengikut program dan pusat pengajian. Dua pendekatan model bayesian berhierarki digunakan iaitu model bayesian berhierarki dengan andaian satu program tidak bersandar dengan program yang lain, begitu juga dengan pusat pengajian. Seterusnya diandaikan persandaran wujud di kalangan program dan pusat pengajian. Menerusi pendekatan model bayesian berhierarki yang pertama, hasil kajian mendapati program Sains Aktuari adalah lebih cemerlang berbanding program-program lain yang terdapat di FST. Manakala, Pusat Pengajian Sains Kimia dan Teknologi Makanan (PPSK) pula merupakan pusat pengajian yang mempunyai min posterior keputusan PNGK pelajar yang tertinggi berbanding empat pusat pengajian lain yang terdapat di FST. Seterusnya, menerusi pendekatan model bayesian berhierarki yang kedua pula, keputusan yang didapati menunjukkan terdapat perbezaan yang beerti bagi nilai anggaran min posterior dan varians, dimana nilai bagi anggaran min posterior dan varians adalah lebih kecil berbanding pendekatan model bayesian berhierarki yang pertama. Manakala, selang kebarangkalian bayesian juga menunjukkan bahawa pendekatan model bayesian berhierarki yang kedua memberikan selang yang lebih kecil dan stabil berbanding pendekatan model bayesian berhierarki yang pertama.





ABSTRACT

This study involves comparison of the students' performance at Faculty of Science and Technology (FST), measured by their results of the Cumulative Grade Point Average (CGPA). This study include all 1088 students who undertake various programmes offered by five schools in FST. The objectives of the study is to compare the results of the CGPA for each programme and school and rank their performance accordingly. Two Hierarchical Bayesian Model approach was used in this study, allowing the independently assumption of one programme independent with other programmes, it goes same with schools. Based on the first assumption, result shows that programme of Actuarial Science rank first compared to the other programmes in FST. While, PPSK shows the highest mean posterior of the CGPA's result compared to the other schools. Furthermore, through the second assumption, the result shows there are significantly difference of the estimated mean posterior and variance, where is the values of the estimated mean posterior and variance are smaller compared to the first assumption model. While, bayessian credible intervals also shows that the second assumption model gives a smaller interval and quite stable compared to the first assumption model.



KANDUNGAN

	Halaman
PENGAKUAN	ii
PENGHARGAAN	iii
ABSTRAK	iv
ABSTRACT	v
KANDUNGAN	vi
SENARAI JADUAL	viii
SENARAI RAJAH	ix
BAB I LATAR BELAKANG KAJIAN	1
1.1 Pendahuluan	1
1.2 Objektif Kajian	4
1.3 Skop Kajian	4
1.4 Struktur Kajian	5
BAB II KAJIAN LAMPAU	6
2.1 Pendahuluan	6
2.2 Kaedah Bayesian	7
2.3 Kaedah Bayesian Berhierarki	12
BAB III PEMODELAN BAYESIAN	19
3.1 Pendahuluan	19
3.2 Data	20
3.3 Model Bayesian Berhierarki	21
3.3.1 Struktur Model Bayesian Berhierarki	22
3.4 Tatakaedah Pensampelan Gibbs	27

	Halaman
BAB IV	ANALISIS DAN KEPUTUSAN
4.1	Pendahuluan 31
4.2	Analisis Data 32
4.3	Analisis Model Bayesian Berhierarki 32
4.3.1	Hasil Analisis Model Bayesian Berhierarki 34
BAB V	PERBINCANGAN DAN KESIMPULAN
5.1	Pendahuluan 52
5.2	Perbandingan Model Bayesian Berhierarki 52
5.3	Kesimpulan 57
5.4	Cadangan 58
RUJUKAN	59
LAMPIRAN	62

SENARAI JADUAL

No. Jadual		Halaman
4.1	Senarai program mengikut PNGK bagi PPSM	35
4.2	Senarai program mengikut PNGK bagi PPFG	36
4.3	Senarai program mengikut PNGK bagi PPSS	36
4.4	Senarai program mengikut PNGK bagi PPSK	37
4.5	Senarai program mengikut PNGK bagi PPBB	38
4.6	Senarai pusat pengajian mengikut keputusan PNGK tertinggi ...	40
4.7	Senarai pusat pengajian mengikut keputusan PNGK tertinggi setelah program Sains Aktuari disingkirkan	40
4.8	Senarai program mengikut PNGK bagi PPSM	42
4.9	Senarai program mengikut PNGK bagi PPFG	42
4.10	Senarai program mengikut PNGK bagi PPSS	43
4.11	Senarai program mengikut PNGK bagi PPSK	44
4.12	Senarai program mengikut PNGK bagi PPBB	45
4.13	Senarai pusat pengajian mengikut keputusan PNGK tertinggi ...	46
4.14	Senarai pusat pengajian mengikut keputusan PNGK tertinggi setelah program Sains Aktuari disingkirkan	48
4.15	Senarai program-program di FST mengikut keputusan PNGK tertinggi	49
5.1	Senarai program mengikut PNGK bagi PPSM	52
5.2	Senarai program mengikut PNGK bagi PPFG	53
5.3	Senarai program mengikut PNGK bagi PPSS	53
5.4	Senarai program mengikut PNGK bagi PPSK	54
5.5	Senarai program mengikut PNGK bagi PPBB	55



SENARAI RAJAH

No. Rajah	Halaman
3.1 Carta pemeringkatan pusat pengajian dan program di FST	20





BAB I

LATAR BELAKANG KAJIAN

1.1 Pendahuluan

Merujuk kepada Berry (1996), dalam proses menganalisis data untuk tujuan pen-
taabiran, terdapat dua jenis pendekatan statistik yang boleh digunakan iaitu pen-
dekatan statistik menerusi kaedah klasik dan juga pendekatan statistik menerusi
kaedah bayesian. Secara dasarnya, kedua-dua pendekatan ini sering kali digunakan
oleh ahli statistik dalam konteks pembuatan keputusan secara berstatistik.

Kebelakangan ini, penggunaan dan pendekatan kaedah bayesian dalam men-
ganalisis data semakin meningkat, ini adalah kerana peningkatan teknologi dalam
bidang komputeran telah membuatkan kaedah ini mudah untuk diaplikasikan se-
cara praktikal. Selain itu, pendekatan bayesian juga memberi penyelesaian yang
berguna dan bermakna dalam sesetengah aplikasi apabila kaedah klasik tidak boleh
melakukannya (Trumbo 2000).

Selain dari pernyataan Trumbo (2000), Kamarulzaman (1995) juga turut
menyatakan bahawa kaedah bayesian juga sudah diaplikasikan secara meluas se-
jak 15 tahun dahulu dan sangat penting dan berguna dalam sesetengah bidang
penyelidikan terutamanya bidang perubatan dan juga bidang sains sosial.

Menurut Pollard (1986), perbezaan yang ketara di antara kaedah klasik den-



gan kaedah bayesian adalah, kaedah bayesian mengambilkira sebarang maklumat tentang parameter yang ingin dianggar sebelum sebarang data dikumpul atau direkodkan. Maklumat tentang parameter tersebut dikenali sebagai maklumat prior.

Selain itu, kaedah bayesian juga menganggap bahawa parameter populasi sebagai suatu pembolehubah rawak, bukan tetap atau anu seperti yang dipertimbangkan oleh kaedah klasik. Taburan parameter populasi yang terdahulu dikenali sebagai taburan prior. Kebiasaannya, ahli statistik memainkan peranan yang sangat penting dalam memilih taburan yang sesuai untuk dijadikan taburan prior bagi sesuatu parameter (Robert 1994).

Merujuk kepada O'Hagan *et al.* (2003), antara kelebihan pendekatan kaedah bayesian berbanding pendekatan kaedah klasik adalah kaedah bayesian mampu untuk menyelesaikan masalah yang lebih rumit dan kompleks serta berkemampuan untuk membuat pentaabiran yang lebih bererti berbanding kaedah klasik.

Menurut Iversen (1984), kelebihan lain bagi kaedah bayesian adalah kaedah ini mengambilkira anggapan tentang variasi antara parameter populasi dari satu kajian ke satu kajian lain.

Selain itu, merujuk kepada Carlin & Louis (1996), antara kelebihan kaedah bayesian adalah penggunaan maklumat terdahulu (maklumat prior), ini adalah kerana dengan menggunakan taburan prior bermaklumat, taburan posterior yang terhasil akan lebih puncak dan mempunyai varians kecil. Maka, parameter populasi yang tidak diketahui juga boleh dianggar dengan lebih tepat dan seterusnya, selang keyakinan yang didapati adalah lebih kecil berbanding dengan penggunaan taburan prior tak bermaklumat.

Seterusnya, menurut Pollard (1986), kaedah bayesian membenarkan penggunaan maklumat yang relevan (sesuai) samada maklumat tersebut datang dari sampel atau dari sumber yang lain. Apabila maklumat yang didapati itu agak sedikit perkaitannya, ianya cukup untuk diterjemahkan dari prior bermaklumat kepada prior tak bermaklumat. Walaubagaimanapun, prior tak bermaklumat ini masih sesuai untuk diaplikasikan menerusi Teorem Bayes' bagi mendapatkan hasil

taburan posterior dengan parameter yang ingin dikaji.

Selain dari kelebihan kaedah bayesian berbanding kaedah klasik seperti yang telah dinyatakan di atas, kaedah bayesian juga tetap mempunyai beberapa kelemahannya. Antaranya adalah kesukaran dalam menentukan taburan prior yang sesuai dengan taburan data yang telah dicerap. Ini kerana apabila maklumat prior mengenai parameter yang ingin dikaji kurang jelas, maka ianya perlu diterjemahkan ke dalam taburan kebarangkalian dan ini merupakan satu proses yang sukar untuk dilaksanakan (Iversen 1984).

Selain dari itu, kesukaran untuk mendapatkan taburan fungsi kebarangkalian keluarga eksponen bagi data juga merupakan salah satu faktor kelemahan kaedah bayesian. Ini adalah kerana apabila sesuatu data itu tidak mengikuti sesetengah taburan yang khusus iaitu fungsi kebarangkalian keluarga eksponen yang digunakan menerusi Teorem Bayes', maka ianya akan memberikan kesan kepada kepekaan hasil bagi taburan posterior yang juga dikenali sebagai isu keteguhan.

Antara kelemahan lain bagi kaedah bayesian adalah kesukaran untuk memanipulasi Teorem Bayes' dalam mendapatkan taburan posterior. Dengan menggunakan taburan konjugat prior menerusi penggunaan Teorem Bayes', taburan posterior mudah untuk didapati. Tetapi sekiranya taburan prior itu agak rumit dan kompleks, maka taburan posterior bagi parameter yang ingin dikaji sukar diperolehi melalui kaedah berangka kerana ianya melibatkan kamiran dimensi yang tinggi dan memerlukan teknik pemfaktoran dan pengembangan yang kompleks (Berger 1985).

Selain daripada pendekatan kaedah bayesian, penggunaan kaedah bayesian berhierarki juga semakin popular digunakan dalam menganalisis data yang melibatkan pembinaan dua atau lebih aras model yang berbeza. Sebagai contoh, aras pertama model memperihalkan taburan dengan parameter tidak diketahui untuk setiap kajian, manakala aras kedua pula menganbilkira anggapan tentang variasi antara parameter populasi dari satu kajian ke satu kajian lain. Aras-aras model ini digabungkan menerusi Teorem Bayes' (Bernado & Smith 1994).



Lantaran itu, pendekatan model bayesian berhierarki akan digunakan untuk menganalisis data dalam kajian ini. Ini adalah kerana kajian ini akan melibatkan data atau maklumat mengenai keputusan PNGK keseluruhan pelajar dari pelbagai program yang berbeza dan pusat pengajian yang berlainan di Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia.

1.2 Objektif Kajian

Objektif bagi kajian yang dijalankan ini adalah seperti berikut :

1. Membuat perbandingan keputusan PNGK pelajar bagi setiap program dan pusat pengajian di Fakulti Sains dan Teknologi menerusi pendekatan kaedah bayesian berhierarki.
2. Membuat pemeringkatan prestasi pelajar mengikut program dan pusat pengajian setelah mengambilkira keperluan seseorang pelajar dari program tertentu mengambil kursus di pusat pengajian atau program yang lain.

1.3 Skop Kajian

Kajian yang dijalankan ini akan meliputi keseluruhan pelajar prasiswazah semester 1 sesi 2003/2004 yang mengikuti program-program yang ditawarkan oleh setiap pusat pengajian di Fakulti Sains dan Teknologi sahaja. Selain itu, kajian ini hanya akan tertumpu kepada keputusan purata nilai gred kumulatif (PNGK) pelajar bagi fakulti ini serta kajian ini tidak akan mengambilkira faktor-faktor lain yang mempengaruhi pencapaian keputusan PNGK pelajar-pelajar di fakulti ini.



1.4 Struktur Laporan

Di dalam Bab Dua, ulasan akan dibuat terhadap kajian lepas yang berkaitan dengan kajian ini. Ulasan untuk kajian lepas akan dibuat secara terperinci bagi mendapatkan idea serta maklumat untuk membuat kajian ini.

Bab Tiga akan menghuraikan secara terperinci mengenai kaedah dan metodologi yang akan digunakan di dalam kajian ini. Model Bayesian berhierarki akan digunakan secara menyeluruh bagi menganalisis data dalam kajian ini.

Di dalam Bab Empat pula, model-model yang terdapat di dalam Bab Tiga akan digunakan untuk menganalisis data dalam kajian ini.

Akhirnya, ringkasan terhadap keputusan yang didapati dari Bab Empat akan dibuat di Bab Lima. Seterusnya, kesimpulan prestasi keputusan PNGK pelajar di Fakulti Sains dan Teknologi dan cadangan mengenai penstrukturan model akan diketengahkan dalam kajian ini.