Proceedings of the Autumn Conference in 2001
and the 30th Anniversary of the Korean Statistical Society

베이지안 통계학의 과거, 현재와 미래

김용대1), 김혜중2), 오만숙3), 오현숙4), 정윤식5)

1. 머리말

본 논문은 한국통계학회 창립 30주년을 기념하기 위하여 베이지안 통계학의 역사, 현재 및 앞으로의 연구방향을 소개하면서 우리나라 베이지안 통계학의 현황과 미래를 조명하여 논의하고자 한다. 특히 베이지안 통계학의 소개는 한국 통계학회 창립이후 처음 시도되는 것이라 더욱 의미가 있으리라 할 것이다.

우선 "베이지안 통계학"이란 관심 있는 모든 것(모수, 결측치, 예측치 등)들에 대한 불확실성(uncertainty)을 확률분포로써 나타낸다는 가정에서 출발한다. 따라서, 전통적인 통계학의 기본 가정인 "고정된 미지"라는 관점과는 근본적으로 출발점이 다르다. 따라서, 베이지안 추론의 근간은 사후분포 또는 사후밀도함수이다. 사후밀도함수는 우도함수와 사전밀도함수의 곱에 비례하는 함수로서 우도함수에 압축된 표본정보와 사전밀도함수에 압축된 사전정보를 베이즈 정리에 의하여 합성한 것이다. 따라서 베이지안 패러다임은 개념적으로 간단하고, 직관적, 확률적 타당성을 지닌다고 볼 수 있다.

그러나 베이지안 접근 방법이 통계학의 워낙 넓은 분야에서 이루어져 있으므로 이를 체계적으로 열거하여 소개하는 것은 쉬운 일이 아니다. 특히 베이지안 통계학은 국내에서의 연구보다는 외국에서의 연구가 보다 활발히 전개되어 그에 대한 영향을 받아오고 있으므로 우선 국제적인 학문의 흐름을 소개하고자 한다. 우선 1970년대부터의 JASA(Journal of American Statistical Association)와 AS(Annal of Statistics)를 중심으로 그 흐름을 간략히 살펴보고자 한다. 그 다음으로 한국 통계학자들의 연구현황을 알아보기 위하여 한국통계학회에서 발간하는 통계학연구, 응용통계연구 및 한국통계학회 논문집을 중심으로 분석한다. 마지막으로, 현재 활발히 연구되어지는 분야들 중, 사전분포함수, 베이지안 계산, 베이지안 검정 및 모형선택, 베이지안 다변량연구 및 비모수적 베이지안 생존자료 분석에 대한 연구 동향을 소개한다.

2. 베이지안 통계학의 연구 현황

2.1 국제적인 현황

베이지안 활동이 급진전한 추세는 실제 숫자들을 통해서 부분적으로 나타내어질 수 있다. JASA 논문집을 중심으로 살펴보면 1973년에서 1980년까지는 매년 6편 정도가 게재되었으며(총 55편), 1981년부터 1989년까지는 매년 평균 7편(총 68편)으로 약간 증가하였다. 그러나 1990년 Gelfand and Smith가 Gibbs sampler를 소개한 후 논문 수는 매우 큰 폭으로 증가하였다. 1990년에 10편을 시작으로 1991년에는 20편이 게재되었으며 그 이후로 매년 15편 이상씩을 발표하고 있다. 더욱이 1996, 1997, 1998년도에 베이지안 논문이 수적으로나 비율적으로 현격한

1) 이화여자대학교, 통계학과
2) 동국대학교, 통계학과
3) 이화여자대학교, 통계학과
4) 경원대학교, 응용통계학과
5) 부산대학교, 통계학과

증가가 있어왔다. (표 2.1 참조). 이와 같은 현상은 MCMC방법들이 복잡한 통계적 모형들에 다양하게 응용되었으며, 더욱이 비모수적 베이지안 접근법들도 폭넓게 사용된 것이다. 또한 이론적인 측면을 강조하는 AS에서도 1990년도 이후에 베이지안 논문 편수가 급증함을 볼 수 있다. 이는 MCMC의 수렴성에 대한 연구와 비모수적인 베이지안 접근 방법에 대한 이론적 연구들이 기존의 베이지안 연구에 첨부되어 연구 영역이 확장되어 베이지안 논문 편수가 증가한 것이다. 이러한 현상들은 베이지안 통계학에 관한 책들의 수에 대하여서도 보여질 수 있다. 처음 200년 동안(1769 - 1969) 베이지안 통계학에 관한 책은 오직 15권 정도였다. 또한, 그 다음 20년간(1970-1989) 제작된 베이지안 통계학 책들은 30권 정도이다. 그 후 10년(1990-1999)간 베이지안 통계학 책들은 대체로 60권이다.

또한 그간에 있던 베이지안들의 연구모임도 1990년대에 접어들면서 확대되었다. 이들 중에서 현재 활동이 활발한 것은 1992년 창립된 Intenational Society for Bayesian Analysis (ISBA)와 같은 해 S. J. Press교수의 주도로 북미주지역의 베이지안들을 위해 결성된 미국통계협회의 베이지안 통계 연구회(Section on Bayesian Statistical Science; SBSS)가 있다. 이 외에도 베이지안들의 연구발표모임으로는 A. Zellner교수가 1970년에 시작한 Seminar on Bayesian Inference in Statistics and Econometrics (SBIES)가 매년 개최되고 있고, Bernardo교수의 주도로 스페인에서 3년 또는 4년마다 개최되는 “Valencia Conference”, E. Jane교수의 업적을 기리기 위해 1981년부터 세계 각지에서 매년 번갈아 열리는 ”Max Ent Workshop”, 그리고 W. Edward교수에 의해 구성된 Bayesian Method in Decision Theory 연구모임 등이 있어 베이지안들 간에 활발한 연구교류가 이루어지고 있다.

표 2.1. JASA 와 AS에서의 베이지안 논문 편수 비율

년	JASA			Annals of Statistics		
	총편수	Bayesian	비율(%)	총편수	Bayesian	비율(%)
1973-1980	1147	55	4.8	920	69	7.5
1981-1989	1181	68	5.6	1030	78	7.6
1990-1995	833	102	12.3	680	62	9.1
1996	166	31	18.7	143	17	11.9
1997	161	28	17.4	106	9	8.5
1998	119	21	17.7	101	11	10.9
1999	107	13	12.1	88	9	10.2
2000	143	20	14	75	8	10.7
2001	58	18	13.8	9	0	0

2.2. 국내연구 현황

국내 베이지안 통계학의 연구 동향을 살펴보기 위하여 한국통계학회에서 발간되는 세 가지 논문집인, 통계학연구(JKSS), 응용통계연구 및 한국통계학회 논문집들을 조사하였다. 이들을 베이지안 추론에 관한 논문 편수와 논문집에 게재된 총 논문 편수를 연도별로 정리하여 표2.2에 나타내었다. 이 표에 의하면 세 가지 논문집에 실린 베이지안 논문 편수는 전체 1593편중 133편으로 약 8.4%를 차지한다. 이를 논문집 별로 나누어 보면, 통계학 연구(JKSS)에서는 전체(618편)의 9%(56편), 응용통계연구에서는 6.5%, 그리고 한국통계학회 논문집에서는 9%를 차지하고 있어서, 방법론에 비해 응용분야에 대한 연구가 활발하지 않은 것으로 나타났다. 또한 각 논문집들의 논문 수가 90년대 중반이후로 두 배 이상 증가됨을 볼 수 있다. 이는 베이지안 통계학자의 증가이기도 하고, 90년에 깁스 샘플러가 소개된 후 어렵게 느껴지던 계산 방법이 보다 쉽게 접근할 수 있다는 이점에 의하여 보다 복잡한 모형에서의 해석에서 많은 통계학자들이 베이지안 방법을 선호하기 때문이다.

표 2.2. 한국통계학회 발간 논문집에서의 베이지안 논문 비율

년	통계학연구(JKSS)			응용통계연구			한국통계학회 논문집		
	총편수	베이지안	비율(%)	총편수	베이지안	비율(%)	총편수	베이지안	비율(%)
1973-1986	174	9	5.2	0	0	0	0	0	0
1987-1993	113	9	8	124	8	6.5	0	0	0
1994	37	3	8.1	41	3	7.3	24	1	4.2
1995	42	1	2.4	29	0	0	102	9	8.8
1996	60	4	6.7	27	1	3.7	59	4	6.8
1997	43	4	9.3	31	2	6.6	62	7	11.3
1998	38	5	13.2	36	3	8.3	88	5	5.7
1999	39	8	20.5	57	3	5.3	93	11	11.8
2000	38	9	23.7	47	4	8.5	87	9	10.3
2001	34	4	11.8	39	4	10.3	29	3	10.3
합	618	56	9	431	28	6.5	544	49	9

1990년까지 국내 베이지안 통계학자들의 수는 매우 적은 수준이었으나, 그 이후로 많은 통계학자들이 베이지안 방법론에 많은 관심을 보여 왔다. 더욱이 외국에서 연구하고 온 베이지안 학자들의 증가로 인하여 베이지안 통계학자들 모임의 필요성이 대두되어 1996년 1월에 첫 번째 모임을 충남대에서 갖게 되었다. 그 후 비정기적으로 춘・추계 학회 학술 논문 발표회에서 베이지안 section을 구성하여 베이지안 통계학자들간의 연구 교환을 활발히 하였다. 그러나 많은 베이지안 통계학자들의 욕구를 충족하기에는 비정기적인 모임으로는 부족하여 1999년 2월 가칭 베이지안 통계 연구회를 동국대에서 발족하였으며, 2000년 6월 한국통계학회의 정식연구회로 가입하여 현재 약 40여명의 회원으로 매년 2월과 8월에 정기 논문 발표회를 갖고 있다.

3. 베이지안 통계학의 연구동향

3.1. 사전분포함수

Bayes(1763)와 Fisher(1922)이래 베이지안 추론에 대한 논쟁이 활발하였다. 이러한 논쟁의 초점은 사전분포함수들의 선택의 임의성이었다. Laplace(1812)는 전 모수 공간에서 무정보 사전분포로써 균일한(flat) 사전분포를 이용할 것을 제안하였다. 그러나 제안된 사전분포는 일대일 재모수화에서 불변성을 유지하지 못한다. 따라서 Jeffreys(1961)는 일대일 재모수화에서 불변을 유지하는 사전 분포를 제안하였다. 이 분포는 피셔 정보 행렬식에 양의 제곱근으로 유도된다. 이러한 불변성에 불구하고, Jeffreys 사전분포는 장애(nuisance) 모수가 있는 경우에 비판을 받고 있다. 이러한 Jeffreys 사전분포의 문제점을 극복하고자 Bernardo (1979)는 참조 사전분포(reference prior)로 알려진 무정보 사전분포를 제안하였다. 참조 사전 분포는 모수 벡터를 관심 있는 모수와 장애 모수로 나누어 적당한 entropy distance 를 최대화함으로써 얻어지며 이때 장애 모수가 없는 경우는 Jeffreys 사전분포와 같다. 이러한 개념을 바탕으로 Berger and Bernardo(1989, 1992)는 참조 사전 분포를 찾는 알고리즘을 개발하였고, 그 후로 Berger와 그 제자들에 의하여 많은 응용분야에서 연구되었다.

또한, 무정보 사전 분포를 발전시키는 다른 형태는 베이지안 신뢰구간의 사후 포함 확률과 그에 대응되는 프리퀀티스트 포함 확률를 근사적 $o(n^{-u})$으로 일치하도록 하는 사전분포를 확률대응 사전분포(probability matching prior)라 한다. 이는 Welch and Peers(1963)가 처음 제안하였으며, 여기서 그들은 장애 모수가 존재하지 않는 경우인 단변량 모수에 국한하였다. 그 후로 장애 모수가 존재한 경우에 관심 있는 모수에 대한 확률대응 사전분포를 구하는 연구가 Peers(1965), Stein(1985)등에 의하여 진행되었다. 여기서 $u=1/2$이면 일차 대응 사전 분포,

$u=1$이면 2차 대응 사전 분포라 한다. 이러한 사전분포함수를 구하는 것들은 적당한 미분방정식의 해를 구하는 문제이다. 특히 Tibshirani(1989)는 관심 있는 모수가 장애 모수에 직교인 경우 일차 대응사전분포의 완전한 해를 제시하였다. 이러한 연구들은 Datta and Ghosh(1995), Sun and Ye(1996), Mukerjee and Ghosh(1996), Garavn and Ghosh(1998)등이 의하여 계속되었다. 또한 Mukerjee and Dey(1993)는 장애 모수가 단변량인 경우에 $u=1$를 만족시키는 이차 대응사전분포를 구하는 방법을 제시하였다. 이러한 결과들을 Mukerjee and Ghosh(1996)는 여러 개의 장애 모수들이 있는 경우로 확장하였다.

마지막으로 소개될 정보적 사전 분포는 연구자가 현재연구에서와 같이 같은 반응 값과 설명변수를 측정할 수 있는 과거연구에 접근 할 수 있는 의학자료 등에서 이용될 수 있다. 베이지안 관점에서 유사한 연구를 수행한 과거 자료들이 현재 연구의 결과들을 설명하는데 매우 많은 도움을 줄 수 있다. 이러한 관점에서 historical data에서 사전분포를 구할 수 있으며, 이를 Power prior라 한다. Chen, Manatunga and Williams(1998)이 인간 쌍둥이 자료에서 유전성 추정치 연구를 위한 power prior를 처음 시도하였다. 그 후로 Chen, Ibrahim(1999, 2000, 2001)등에 의하여 일반화 선형모형과 다양한 생존자료 모형에서 연구되었다.

3.2. 베이지안 계산

베이지안의 실제적용은 단순하지 않은 경우가 종종 발생하는데 이는 근간이 되는 사후밀도함수가 수리적으로 주어지지 않고 단지 그 함수형태만 알 수 있는 경우가 많기 때문이다. 따라서 사후밀도함수와 나아가서는 사후추론을 위한 사후통계량을 구하는 베이지안 계산기법이 요구되므로 몇 가지의 베이지안 계산기법을 알아보기로 한다. 첫 번째 방법은, 사후밀도함수를 모수 θ의 MLE를 평균으로 그리고 헤시안 행렬의 역함수에 부호를 바꾼 것을 분산으로 하는 정규밀도함수로 근사 시킨다. 이는 매우 간단하나 표본의 크기가 작으면 근사가 정확하지 않은 단점이 있다. 두 번째로, θ의 함수 $u(\theta)$의 사후기대치에 대한 근사식을 구하는 것으로 앞의 정규근사에 보정치를 추가하여 정확도를 높이는 방법으로 Lindley(1980)에 의해 제안되었고 Tierney and Kadane(1986)은 Lindley의 근사식을 개선하여 정확도가 더 높은 근사식을 제안하였다. 위의 두 방법들은 수리적 근사를 이용한 것으로 표본의 크기가 충분히 크지 않을 경우에는 상당히 오차가 클 수 있다. 따라서, Naylor and Smith (1982)는 이점을 지적하고 Gaussian quadrature 방법을 이용하여 수치적으로 사후기대치를 추정할 것을 제안하였다. 이 방법은 사후기대치 계산에 필요한 적분의 차원이 작은 경우에는 효율이 좋으나 차원이 높아지면 효율이 매우 떨어지는 단점이 있다.

다차원 적분을 효율적으로 수행하기 위한 수치적 기법으로 몬테칼로 기법이 1970년 후반부터 통계학자들의 관심을 끌기 시작하였다. 몬테칼로 기법 중 주 표본기법(importance sampling)은 초기에 Kloek and van Dijk(1978), Van Dijk and Kloek (1980, 1983, 1984), Stewart (1983), Geweke (1988) 등에 의하여 연구되었다. 주 표본기법은 알고리즘이 단순하고 다차원 적분에도 효율이 뛰어난 장점이 있으나 기법의 효율과 정확도가 표본생성함수에 크게 의존한다는 것이 단점이다.

1990년 Gelfand and Smith (1990)은 그 동안 물리학계에서 사용되어지던 마코브 체인 몬테칼로 기법 (Markov chain Monte Carlo, MCMC)이 베이지안 추론의 여러 분야에서 매우 유용하게 사용되어질 수 있음을 보여주었다. MCMC 기법은 주 표본기법에 비해 효율은 다소 떨어지나 전문가가 아니라도 사용이 용이하고 그 적용범위가 매우 넓다는 장점을 가지고 있어 90년대 들어서 폭발적인 관심을 끌었으며 베이지안 패러다임의 적용 범위를 크게 확장시켰다는데 그 의의가 있다. 특히 모형이 복잡한 다차원 문제에도 적용이 용이하며, 필요에 따라 잠재변수의 사용으로 문제를 단순화시킬 수 있고 결측치의 처리 또한 용이하다. 그러나 깁스표본기법은

각 조건부 사후분포가 편리한 형태로 주어져야 한다는 조건이 필요한데 이를 위하여 보통 짝사전분포의 사용이 요구된다는 단점이 있다. 이를 해결하기 위한 방법으로 Metropolis-Hastings 기법(Hastings, 1970)이 있다. 최근 들어 모형선택에 사용되어질 수 있는 MCMC기법이 관심을 끌고 있는데 Green(1995)의 역점프 MCMC (Reversible jump MCMC) 기법은 차원이 서로 다른 모형들 사이에 마코브 체인을 형성함으로써 각 모형의 사후확률을 추정할 수 있는 장점이 있어 관심을 끌고 있다.

3.3. 베이지안 가설 검정(베이지안 모형 선택)

베이지안 가설검정의 기본 원리는 설정된 가설, $H_i\ (i=0,1)$의 사후확률(posterior probability of H_i)값을 비교하여 큰 쪽을 선택하는 것이다. 따라서 두 개 이상의 여러 개의 가설검정에도 좀 더 쉽게 적용할 수 있다. 계층적 모형이나 합성모형의 경우, 초기단계의 모형설정에서 고전적 가설검정은 결과에 따라 채택 또는 기각을 함으로써 오차를 무시하지만 베이지안 방법에서는 베이즈 요인을 이용하여 가중치를 줌으로써 유연성을 보인다. 베이지안 가설검정이라는 용어보다는 흔히 베이지안 모델 선택 또는 베이지안 모델 비교라고 불리는 이유가 여기에 있다고 본다. 다른 베이지안 분석과 마찬가지로, 베이즈 요인도 모델의 사전확률(prior probability)의 영향을 받는데, 각 모델의 사전확률에 따라 P-값을 이용한 고전적 가설검정의 결과와 불일치할 수 있다는 사실이 Lindley(1957)에 의해 최초로 발표되었으며 "Lindley's paradox" 라고 불리고 있다. 이로부터 베이지안과 고전적 가설검정간의 결과의 불일치성에 대한 연구가 활발히 이루어졌다. Pratt(1965), DeGroot(1973), Dempster(1973), Dickey(1977), Zellner와 Siow(1980), Hill(1982)은 한쪽 검정에서는 P-값이 무정보 사전확률(vague prior probability)에 대한 H_0의 사후확률과 근사적으로 일치한다는 것을 제시했다. Casella와 Berger(1987)는 한쪽검정에서 H_0의 사후확률의 infimum을 구하여 P-값과 비교하여 이 둘이 근사적으로 일치하거나 P-값이 더 큼을 보였다. 여기서 H_0 사후확률의 infimum은 타당한 사전확률분포들의 군집 하에서 구한 것으로 Dickey(1977)에 의해 처음으로 시도되었다. Berger와 Sellke(1987), Berger와 Delampady(1987), Delampady와 Berger(1990)는 Dickey(1977)가 접근한 방법과 유사하게 여러 개의 사전확률분포들의 군집을 고려하여 H_0의 사후확률의 infimum을 P-값과 비교한 결과 H_0의 사후확률의 infimum이 P-값보다 큰 경우가 많음을 보였다. 이것은 양쪽검정에서 P-값에 의한 판별의 결과를 믿을 수 없음을 의미한다. 최근 베이지안과 고전적 방법이 일치될 수 있는 분석방법에 대한 연구가 계속되고 있다 (Good(1992)). Berger, Brown과 Wolpert(1994)는 단순 대 단순 검정에서 conditional frequentist 방법이 베이지안 방법과 오류율에서 일치함을 보였으며 Berger, Boukai와 Wang(1997)은 이를 단순 대 복합검정으로의 확장을 시도하였다.

3.4. 베이지안 다변량통계

다변량 해석은 19세기 후반 F. Galton이 상관개념을 처음으로 도입한 후 다양한 분야에서 자료분석의 기법으로 개발되었다. 교육학과 심리학 분야에서 개발된 C. Spearman의 인자분석, 유전학적 현상을 설명하기 위해 R. A. Fisher가 개발한 여러 다변량 해석기법, 천문학분야에서 개발된 Gauss의 회귀모형과 경제학분야에서 이를 일반화 시켜 개발한 여러 종류의 다변량 회귀모형들, 농학분야에서 개발한 다변량 실험계획모형; 교육학, 심리학 및 사회학분야에서 개발한 잠재구조모형과 다차원 척도모형, 화학, 경제학, 공학분야에서 개발한 조절모형; 재정학분야에서 사용되는 다변량 안정분포(stable distribution)이론, 그리고 인류학, 분류학 및 마케팅분야에서 사용되는 분류 및 판별 분석 등의 기법들이 다변량 분포이론들과 어우러져 응용성이 높은

통계학의 한 분야로 자리 메김하고 있다.

다변량 해석의 추론 및 의사결정기법에는 크게 표본이론에 근거한 전통적인 방법과 베이지안 방법의 두 가지 방법으로 나눌 수 있다. 베이지안 방법이란 서로 상관을 가진 확률변수들의 벡터 또는 행렬의 분포 및 표본분포이론을 바탕으로, 분석하고자 하는 현상에서 수집한 다변량 자료정보와 이에 대한 개인적 또는 주관적인 정보를 베이즈 정리를 통해 결합시켜 그 현상을 추론하는 기법들의 집합을 말한다.

다변량 해석에서 위 두 가지 방법의 우열에 관한 뚜렷한 결론은 없으나, 이들을 이용하여 추론 및 의사결정을 함에 있어 각 추론법이 주관적 그리고 기술적인 어려움을 안고 있다는 것은 이미 밝혀진 사실이다. 예를 들면, 다변량 해석에서 전통적 방법을 사용할 경우 표본의 크기와 제 1종 및 제 2종 오류간의 승산 (trade-off)에 관해 분석자의 주관적인 결정이 필요하며, 제 1종 및 제 2종 오류의 개념을 사용하지 않는 베이지안 방법에서는 주관적인 사전확률분포 설정 문제를 지니고 있다. 한편, 전통적인 방법에서 최적의 분석법이 없는 다변량 회귀분석과 이분산 정규판별분석 등을 베이지안 방법으로 간단하게 분석할 수 있는데 반하여, 복잡한 사후확률분포(zonal polynomials 분포)를 가진 주성분모형이나 정준분석모형 등은 전통적 방법이 간단한 분석법을 제공한다. 그러므로, 이 두 방법이 현재 다변량 해석의 기술적인 측면에서 서로 보완적인 관계를 가진다고 볼 수 있다. 20세기 후반부터 급진전된 베이지안 통계계산법, 로버스트 베이지안 분석법, 그리고 비모수적 방법에 대한 연구는 그 동안 큰 발전이 없었던 다변량 베이지안 통계에 대한 연구에 활로를 열어주었다.

3.5. 베이지안 생존분석

Ferguson (1973)이 비모수 사전분포로써 Dirichlet process를 제안한 후 많은 통계학자에 의하여 중도절단 자료에서의 사후분포의 유도가 시도되었다. Susarla and van Ryzin (1976)에 의하여 우절단 자료에서 생존함수의 베이즈 추정량이 계산되었다. Doksum (1974)은 Dirichlet process 사전분포를 일반화한 neutral to right process 사전분포를 개발하였으며, Ferguson and Phadia (1978)에 의하여 우절단 자료에서 neutral to right process 사전분포가 공액류가 됨이 보여졌다. 하지만, neutral to right process 사전분포는 그 범위가 너무 커서 자료분석에 필요한 실제적인 의미는 거의 없다. Hjort (1990)는 누적분포함수대신 누적위험함수를 이용하였고, 이를 위하여 Beta process 사전분포를 개발하였다. 사후분포의 유도와 함께, 사후분포의 효율적인 계산을 위한 연구가 MCMC 알고리즘의 개발 후에 많이 수행되었다. Doss (1994)에 의하여 중도절단 자료에서 Dirichlet process 사전분포를 사용한 경우에 사후분포를 계산하는 효율적인 Gibbs sampling방법이 개발되었으며, Damien et al. (1996)에 의하여 beta process를 이용한 MCMC 알고리즘이 개발되었다. 최근에는 사후분포의 이론적 성질에 대한 연구가 활발히 진행되고 있는데, Kim and Lee (2001,a)은 사후분포가 점근적으로 불편의(posterior consistency)할 충분조건을 제안하였다. 실제문제에 많이 쓰이는 사전분포들(Dirichlet process, beta process, gamma proces)은 모두 점근적으로 불편의한 사후분포를 갖음이 증명되었다. 비례위험모형에서의 베이지안 분석방법은 Kalbfleisch (1978)와 Hjort (1990)에 의하여 시도되었으며, Laud et al. (1998)에 의하여 MCMC 알고리즘이 개발되었고, Kim and Lee (2001,b)에 의하여 사후분포의 이론적 성질뿐 아니라, left truncation이 포함된 자료로 확장되었다. 향후 베이지안 생존분석의 과제로는 무정보 사전분포의 개발, 다변량 생존분석에서의 베이지안 추론방법의 개발, 임의변량을 포함하는 모형에서의 효율적인 사후분포 계산 알고리즘의 개발 등이 있다.

참고문헌

[1] Berger, J. O, Boukai, B. and Wang, Y. (1997). Unified Frequentist and Bayesian Testing of a Precise Hypothesis, *Statistical Science*, **12**, 133-160.

[2] Berger, J. O. and Sellke, T. (1987). Testing a Point Null Hypothesis: the Irreconcilability of P-values and Evidence, *Journal of American Statistical Association*, **82**, 112-122.

[3] Bickel, P.J. and Ghosh, J.K.(1990) "A decomposition for the likelihood ratio statistic and the bartlett correction - A Bayesian argument", *Annals of Statistics*, 18, 3, 1070-1090.

[4] Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. London: Addison-Wesley Publishing Company.

[5] Chen, M-H., Ibrahim,J.G. and Yiannoutsos, C.(1999) "Prior elicitation, variable selection and Bayesian compuatation for logistic regression models." JRSS-B, 61, 223-242.

[6] Damien, P., Laud, P.W. and Smith, A.F.M. (1996). Implementation of Bayesian nonparametric inference based on beta processes. *Scandinavian Journal of Statistics*, **23**, 27-36.

[7] Doksum, K.A. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Prob.* **2**, 183-201.

[8] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, **1**, 209-230.

[9] Gelfand, A.E. and Smith, A.F.M. (1990), Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, vol. 85, pp. 398-409.

[10] Ghosh, J.K. and Mukerjee, R.(1992) "Bayesian and frequentist Bartlett corrections for likelihood ratio and conditional likelihood ratio tests", *Journal of Royal Statistical Society, Ser. B*, 54, 867-875.

[11] Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika, vol. 82, pp. 711-732.

[12] Hjort, N.L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.*, **18**, 1259-1294.

[13] Kim, Y. and Lee, J. (2001). On posterior consistency of survival models. *Ann Stat.*,(to appear).

[14] Lindley, D. V. (1957). A Statistical Paradox, *Biometrika*, **44**, 187-192.

[15] Naylor, J.C. and Smith, A.F.M. (1982), Applications of a method for efficient computation of posterior distributions, *Applied Statistics*, vol. 31, pp. 214-225.

[16] Tierney, L. and Kadane, J.B. (1986), Accurate approximations for posterior moments and marginals, *Journal of the American Statistical Association*, vol. 81, pp. 82-86.