Journal of Korean
Data & Information Science Society
2002, Vol. 13, No.2 pp 201 207

Noninformative Priors for the Ratio of Means of Two Poisson Distributions

Sang Gil Kang[1] · Woo Dong Lee[2] · Dal Ho Kim[3]

Abstract

In this paper, Jeffrey's and reference priors are derived when the parameter of interest is the ratio of means of two independent Poisson distribution. To achieve the parameter orthogonality in the sense of Cox and Reid (1987), non-trivial orthogonal transformation is provided. The orthogonal transformation makes to find noninformative priors easy. Our simulation study indicates that the reference prior meet very well the target coverage probabilities in a frequentist sense. Using the real data, we compute Bayes estimator and MLE for the ratio of means based on the reference prior.

Keywords : 제프리스 사전분포, 기준사전분포, 포아송 분포의 평균비

1. 서론

베이지안 통계적 추론에서 사전정보가 거의 없을 때 무정보사전분포(noninformative prior)를 사용한 베이지안 분석이 가장 일반적이라는 것이 알려져 있다. 가장 널리 사용되어지는 사전분포로는 제프리스 사전분포 (Jeffrey's prior)가 있다.

제프리스 사전분포는 가장 오래된 무정보사전분포이며 객관적인 사전분포라는 사실은 잘 알려져 있다. 이 사전분포는 피셔 정보행렬식의 제곱근에 비례하도록 사전분포를 주는 방법이다. 모수가 한 개인 모형에서 제프리스 사전분포는 좋은 성질을 가지고 있음이 잘 알려져 있다. 그러나 관심의 대상인 흥미모수 (parameter of interest)와 장애모수 (nuisance parameter)가 함께 존재하는 모형인 경우에는 심각한 문제점을 갖게된다. 예를 들면, Neyman-Scott 문제, Marginalization 패러독스, Stein 패러독스, 구간추정시에 신뢰계수를 만족시키지 못하는 확률일치성 등의 문제가 있음은 잘 알려

1. 부산광역시 남구 대연동 110-1 경성대학교 교양과정부 전임강사
2. 경상북도 경산시 점촌동 산75 경산대학교 정보과학부 부교수
3. 대구광역시 북구 산격동 1370번지 경북대학교 통계학과 부교수

진 사실이다.

이러한 문제점을 해결하기 위하여 Berger와 Bernardo (1989, 1992)는 Bernardo (1979)의 기준사전분포 (reference prior) 유도방법을 개선하여 추론의 중요성에 따라 모수를 순서화하여 기준사전분포를 유도하는 일반적인 알고리즘을 개발하였다. 이 사전분포는 결측정보 (missing information: 엔트로피 측면)를 최대화함으로써 얻어지게 된다. 많은 연구에 의하면 여러 가지 모형들에서 개발된 기준사전분포가 성공적이라는 것이 밝혀졌다.

평균비 (ratio of means)에 대한 연구로는 Datta와 Ghosh (1995)가 지수분포의 평균비에 대하여 기준사전분포를 개발하였고, Bernardo (1977)는 정규분포의 평균비에 대하여 연구하였다. Liseo (1993)는 포아송 분포의 평균비에 대하여 제프리스 사전분포와 기준사전분포를 개발하였는데, 그의 연구는 두 모집단의 표본의 크기가 동일한 경우에서의 연구이다.

본 논문에서는 독립인 두 개의 포아송 분포의 평균비에 대한 무정보사전분포의 개발이다. 이 연구는 두 모집단의 표본크기가 다른 경우에서의 연구이므로 Liseo (1993)의 연구를 포함하는 일반적인 경우가 된다. 2절에서는 모수의 직교화를 이루기 위한 직교변환를 찾고, 변환된 모수하에서 피셔 정보행렬 (information matrix)을 유도한다. 그런 다음 이 정보행렬을 이용하여 평균비에 대한 제프리스 사전분포와 기준사전분포를 유도한다. 3절에서는 제프리스 사전분포와 기준사전분포를 포함하는 일반적인 사전분포하에서 사후확률분포 (posterior distribution)의 적절성 (propriety)을 밝히고, 평균비의 사후확률분포를 계산한다. 4절에서는 모의실험을 통하여 제안된 사전분포하에서의 프리컨티스트 포함확률 (frequentist coverage probability)을 계산하고, 실제 자료를 이용하여 베이즈 추정량 (Bayes estimator)과 최우추정량 (MLE)을 계산하는 예를 보인다.

2. 평균비의 무정보사전분포

X는 평균이 θ_1, Y는 평균이 θ_2인 서로 독립인 포아송 확률변수라고 두자. 그리고 $X_1, X_2, \cdots, X_m$을 X로부터의 확률표본이라고 하고, $Y_1, Y_2, \cdots, Y_n$을 Y로부터의 확률표본이라고 하자. 또한 $\underline{x} = (x_1, x_2, \cdots, x_m)$과 $\underline{y} = (y_1, y_2, \cdots, y_n)$을 두 모집단으로부터의 확률표본의 관찰값이라고 한다면, 모수 θ_1과 θ_2에 대한 우도함수 (likelihood function)는 다음과 같다.

$$L(\theta_1, \theta_2 \mid \underline{x}, \underline{y}) = \frac{\theta_1^{\sum_{i=1}^m x_i} \theta_2^{\sum_{j=1}^n y_j} \exp\{-m\theta_1 - n\theta_2\}}{\prod_{i=1}^m x_i! \prod_{j=1}^n y_j!}.$$

위 우도함수에서 관심있는 모수는 두 모수의 평균비인 $\omega = \theta_1/\theta_2$이다. 모수 직교화 변환을 이용하면 ω에 대한 기준사전분포가 쉽게 유도되어진다. 이를 위하여

$$\omega = \theta_1/\theta_2, \quad \eta = m\theta_1 + n\theta_2$$

라 두면, 우도함수는 다음과 같이 주어지게 된다.

$$L(\omega,\eta \mid \underline{x},\underline{y}) \propto \left(\frac{\omega\eta}{n+\omega m}\right)^{T_1}\left(\frac{\eta}{n+\omega m}\right)^{T_2}\exp\left\{-m\frac{\omega\eta}{n+\omega m}-n\frac{\eta}{n+\omega m}\right\}.$$

여기서 $T_1=\sum_{i=1}^{m}x_i$, $T_2=\sum_{j=1}^{n}y_j$ 이다. 위의 우도함수에서 피셔 정보행렬은

$$I_{\omega,\eta}=\begin{bmatrix}\frac{nm\eta}{\omega(n+\omega m)^2} & 0\\ 0 & \frac{1}{\eta}\end{bmatrix}.$$

가 됨을 알 수 있다. 피셔 정보행렬 $I_{\omega,\eta}$는 대각행렬로서 흥미모수 ω와 장애모수 η는 Cox와 Reid (1987)가 정의한 모수직교성을 만족함을 알 수 있다.

Cox와 Reid (1987)의 모수직교성이 만족되는 경우 Berger와 Bernardo (1989, 1992)의 기준사전분포들은 Datta와 Ghosh (1995)의 방법에 의하여 관심있는 모수들의 추론의 중요성에 따라서 아래와 같이 쉽게 유도되어 진다.

모수공간이 하나의 그룹 $(\{\omega,\eta\})$인 경우에 기준사전분포는

$$\pi_1^R(\omega,\eta) \propto \frac{1}{\sqrt{\omega}(n+\omega m)},\quad \omega>0,\eta>0$$

이 된다는 것을 알 수 있으며, 이 경우 기준사전분포는 제프리스 사전분포가 된다. 다음으로 두 개의 그룹 $(\{w\},\{\eta\}$ 또는 $\{\eta\},\{w\})$으로 나누어지는 경우에 기준사전분포는

$$\pi_2^R(\omega,\eta) \propto \frac{1}{\sqrt{\omega}\sqrt{\eta}(n+\omega m)},\quad \omega>0,\eta>0.$$

이다.

Remark 1. 만일 m과 n이 같다면 기준사전분포는

$$\pi_2^R(\omega,\eta) \propto \frac{1}{\sqrt{\omega}\sqrt{\eta}(1+\omega)},\quad \omega>0,\eta>0$$

가 되며, 이것은 Liseo (1993)가 개발한 기준사전분포와 동일하다.

Remark 2. 제프리스 사전분포와 기준사전분포를 원래의 모수 θ_1과 θ_2로 표현하면 각각 아래와 같이 주어지게 된다.

$$\pi_1^R(\theta_1,\theta_2) \propto (\theta_1\theta_2)^{-1/2},\quad \omega>0,\eta>0,$$

$$\pi_2^R(\theta_1,\theta_2) \propto (\theta_1\theta_2)^{-1/2}(m\theta_1+n\theta_2)^{-1/2},\quad \omega>0,\eta>0.$$

3. 사후분석

먼저 2절에서 제안한 두 개의 무정보사전분포를 포함하는 일반적인 사전분포를 생각해보자. 그 사전분포는 다음과 같이 표현 할 수 있다.

$$\pi^G(\omega,\eta) \propto \omega^{\alpha}(n+\omega m)^{\beta}\eta^{\gamma}.$$

위의 사전분포 π^G에서 $\alpha=-1/2$, $\beta=-1$, $\gamma=0$을 대입한다면 이는 제프리스 사전분포와 동일하며, $\alpha=-1/2$, $\beta=-1$, $\gamma=-1/2$를 대입한다면 기준사전분포가 됨을

확인할 수 있다. 이제 사전분포 π^G를 이용하여 모수 ω와 η에 대한 결합사후확률분포 (joint posterior distribution)를 유도하고, 그 결합사후분포가 적절분포 (proper distribution)를 가지는 지를 조사해 보자.

변환된 모수 ω와 η에 대한 우도함수는 다음과 같다.

$$L(\omega,\eta \mid \underline{x},\underline{y}) \propto \left(\frac{\omega\eta}{n+\omega m}\right)^{T_1}\left(\frac{\eta}{n+\omega m}\right)^{T_2}\exp\left\{-m\frac{\omega\eta}{n+\omega m}-n\frac{\eta}{n+\omega m}\right\}.$$

여기에서 $T_1=\sum_{i=1}^{m}x_i$, $T_2=\sum_{j=1}^{n}y_j$ 이다. 이 우도함수와 사전분포인 π^G와 결합하여 사후분포를 유도하면 다음과 같다. $\omega>0$, $\eta>0$에 대하여,

$$\begin{aligned}\pi^G(\omega,\eta \mid \underline{x},\underline{y}) &\propto L(\omega,\eta \mid \underline{x},\underline{y})\pi^G(\omega,\eta)\\ &\propto \eta^{T+\gamma}\exp\{-\eta\}(n+\omega m)^{\beta-T}\omega^{T_1+\alpha}.\end{aligned}$$

여기서 $T=T_1+T_2$를 나타낸다.

정리 1. $T+\gamma+1>0$, $T_1+\alpha+1>0$ 그리고 $T_2-\beta-\alpha-1>0$을 만족하는 경우, ω와 η에 대한 결합사후확률밀도함수 $\pi^G(\omega,\eta \mid \underline{x},\underline{y})$는 적절확률분포이다.

증명. ω와 η의 결합사후확률분포는 다음과 같이 주어진다.

$$\pi^G(\omega,\eta \mid \underline{x},\underline{y}) \propto \eta^{T+\gamma}\exp\{-\eta\}(n+\omega m)^{\beta-T}\omega^{T_1+\alpha}.$$

이 사후확률분포로부터 만일 $T+\gamma+1>0$을 만족한다면 ω의 주변사후확률분포 (marginal posterior distribution)는

$$\pi^G(\omega \mid \underline{x},\underline{y}) \propto (n+\omega m)^{\beta-T}\omega^{T_1+\alpha}$$

이다. $Z=(1+\frac{m}{n}\omega)^{-1}$라 두면 위의 주변사후확률분포에 대한 적분은

$$\int_0^\infty (n+\omega m)^{\beta-T}\omega^{T_1+\alpha}d\omega \propto \int_0^1 Z^{T_2-\beta-\alpha-2}(1-Z)^{T_1+\alpha}dZ$$

이 된다. 따라서 $T_1+\alpha+1>0$, $T_2-\beta-\alpha-1>0$의 조건을 만족하게 된다면 적분값은 유한하게 된다. □

위의 증명과정에서 흥미모수 ω에 대한 주변사후확률분포는

$$\pi^G(\omega \mid \underline{x},\underline{y}) \propto (n+\omega m)^{\beta-T}\omega^{T_1+\alpha},\quad \omega>0$$

이 된다는 사실을 알 수 있다. 따라서 ω에 대한 정확한 사후확률밀도함수를 유도하면 다음과 같이 주어진다.

정리 2. ω에 대한 주변사후확률밀도함수는

$$\pi^G(\omega \mid \underline{x},\underline{y}) = \frac{m^{T_1+\alpha+1}n^{T_2-\beta-\alpha-1}}{B(T_1+\alpha+1,\,T_2-\beta-\alpha-1)}(n+\omega m)^{\beta-T}\omega^{T_1+\alpha},\quad \omega>0$$

이다. 여기서

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

이다.

증명. 먼저 $\pi^G(\omega \mid \underline{x},\underline{y})$에 대한 정규화상수 (normalizing constant) K를 구하면

$$\begin{aligned} K &= \int_0^\infty (n+\omega m)^{\beta-T}\omega^{T_1+\alpha}d\omega \\ &= n^{\beta+\alpha+1-T_2}m^{-T_1-\alpha-1}B(T_1+\alpha+1,\ T_2-\beta-\alpha-1) \end{aligned}$$

이다. 따라서 ω에 대한 사후확률밀도함수는

$$\pi^G(\omega \mid \underline{x},\underline{y})=\frac{m^{T_1+\alpha+1}n^{T_2-\beta-\alpha-1}}{B(T_1+\alpha+1,\ T_2-\beta-\alpha-1)}(n+\omega m)^{\beta-T}\omega^{T_1+\alpha},\quad \omega>0$$

이 된다. □

Remark 3. 흥미로운 결과는 정리 2에서 계산되어진 ω의 사후분포는 α와 β에만 의존하고 γ값에는 의존하지 않음을 알 수 있다. 즉, 평균비에 대한 제프리스 사전분포나 기준사전분포는 서로 다르더라도 이 두 사전분포로부터 유도된 ω의 주변사후분포는 같다는 사실이다. 따라서 이 사후분포로부터 계산되어진 ω에 대한 적률(moment)들은 같게 된다. 예를 들면, 제곱오차손실함수 (squared error loss function)를 가정하는 경우에 흥미모수 ω에 대한 베이즈 추정량은 사후분포의 평균이 되며, 이는 정리 2로부터

$$\begin{aligned} \hat{\omega}_G &= E(\omega \mid \underline{x},\underline{y}) \\ &= \frac{nB(T_1+\alpha+2,\ T_2-\beta-\alpha-2)}{mB(T_1+\alpha+1,\ T_2-\beta-\alpha-1)} \end{aligned}$$

이 된다는 것을 알 수 있다. 이 베이즈 추정량은 4장에서 실제자료를 사용하여 계산될 것이다.

4. 모의실험 및 예제

(θ_1,θ_2)과 (n,m)의 여러 가지 값들하에서 ω의 주변사후분포의 신뢰구간을 조사함으로서 프리컨티스트 포함확률을 계산할 것이다. 이 계산은 수치적으로 행하여지는데 표 1에 수치적으로 계산되어진 프리컨티스트 포함확률 값들이 주어져있다. 이 값들은 (θ_1,θ_2)과 (n,m)의 주어진 값 하에서 아래의 알고리즘을 통하여 계산되어졌다. $\omega^\pi(\alpha \mid X_1,X_2)$를 ω의 α-사후분위수라 두자. 여기서 α는 0.05(0.95)이다. 즉, $F(\omega^\pi(\alpha \mid X_1,X_2) \mid X_1,X_2)=\alpha$이고, $F(\cdot \mid X_1,X_2)$는 ω의 주변사후분포함수이다. 그러면 ω의 단측신뢰구간의 프리컨티스트 포함확률은

$$P_{\theta_1,\theta_2}(\alpha;\omega)=P_{\theta_1,\theta_2}(0<\omega<\omega^\pi(\alpha \mid X_1,X_2))$$

이다. $\alpha = 0.05(0.95)$일 때 추정된 $P_{\theta_1,\theta_2}(\alpha;\omega)$값이 표 1에 주어져 있다. 이 값들은 주어진 포아송 모형으로부터 추출된 10,000개의 독립인 확률표본들로부터 계산되어진 값이다.

포아송분포가 이산분포임에도 불구하고, 표 1의 결과에서 알 수 있듯이 개발되어진 사전분포가 목표포함확률에 잘 일치한다는 것을 알 수 있다. 그리고 θ_1과 θ_2의 값의 변화에 따라 그 결과가 그렇게 민감하지 않다는 것을 또한 알 수 있다.

표 1: ω의 0.05 (0.95) 사후분위수의 프리컨티스트 포함확률

(θ_1, θ_2)	(n, m)	π
0.2, 1	3, 3	0.0540 (0.9998)
	3, 5	0.0478 (1.0000)
	5, 5	0.0524 (0.9952)
	5 ,7	0.0613 (0.9977)
	7, 7	0.0513 (0.9741)
	7, 10	0.0512 (0.9878)
	10, 10	0.0480 (0.9447)
0.5, 1	3, 3	0.0702 (0.9576)
	3, 5	0.0616 (0.9702)
	5, 5	0.0653 (0.9398)
	5 ,7	0.0461 (0.9524)
	7, 7	0.0584 (0.9413)
	7, 10	0.0520 (0.9521)
	10, 10	0.0474 (0.9430)
1, 1	3, 3	0.0661 (0.9313)
	3, 5	0.0489 (0.9488)
	5, 5	0.0574 (0.9490)
	5 ,7	0.0515 (0.9486)
	7, 7	0.0531 (0.9491)
	7, 10	0.0574 (0.9484)
	10, 10	0.0500 (0.9500)
3, 1	3, 3	0.0577 (0.9430)
	3, 5	0.0482 (0.9473)
	5, 5	0.0466 (0.9410)
	5 ,7	0.0537 (0.9469)
	7, 7	0.0542 (0.9463)
	7, 10	0.0490 (0.9451)
	10, 10	0.0518 (0.9492)
5, 1	3, 3	0.0578 (0.9499)
	3, 5	0.0504 (0.9449)
	5, 5	0.0528 (0.9509)
	5 ,7	0.0522 (0.9461)
	7, 7	0.0495 (0.9484)
	7, 10	0.0474 (0.9499)
	10, 10	0.0489 (0.9469)

다음의 자료는 양배추 자벌레의 조절에 대한 여러 가지 처리로부터 얻어진 자료이다. 각 처리는 4개의 구획에서 실험되었고, 처리에 대한 각 구획에서의 자벌레의 수를 나타낸다 (Snedecor와 Cochran, 1980).

X	14	27	8	18
Y	11	4	4	5

위의 자료로부터 $T_1 = 67$이며, $T_2 = 24$, $m = 4$, $n = 4$이므로, $\hat{\omega}_{MLE} = 2.7917$, $\hat{\omega}_{Bayes} = 2.8723$가 된다. 고전적 방법이나 베이즈 추정법 모두 거의 비슷한 추정값을 준다. 그러나 고전적인 추정에서 추정량의 분산은 근사적인 방법으로 계산되어지고, 소표본하에서는 분산의 추정치는 문제점을 가지고 있다. 그러나 베이즈 추정법에서는 평균비의 주변사후분포를 알고 있으므로 분산을 직접적으로 계산할 수 있다.

위의 모의실험결과와 예제를 통하여 개발되어진 기준사전분포가 바람직한 성질을 가지고 있다는 것을 알 수 있다. 그러므로 평균비의 추론의 연구에 개발되어진 사전분포가 사용될 수 있으리라 생각된다.

참고문헌

1. Berger, J. O. and Bernardo, J. M. (1989). Estimating a Product of Means : Bayesian Analysis with Reference Priors, *Journal of the American Statistical Association*, 84, 200-207.
2. Berger, J. O. and Bernardo, J. M. (1992). On the Development of Reference Priors (with discussion), *Bayesian Statistics IV*, J. M. Bernardo, et. al., Oxford University Press, Oxford, 35-60.
3. Bernardo, J. M. (1977). Inferences about the Ratio of Normal Means; A Bayesian Approach to the Fieller-Creasy Problem, In *Recent Developments in Statistics* (J. R. Barra et al. eds.). Amsterdam: North-holland, 345-349.
4. Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference (with discussion), *Journal of Royal Statistical Society*, B, 41, 113-147.
5. Cox, D. R. and Reid, N. (1987). Orthogonal Parameters and Approximate Conditional Inference (with discussion), *Journal of Royal Statistical Society*, B, 49, 1-39.
6. Datta, G. S. and Ghosh, M. (1995). Some Remarks on Noninformative Priors. *Journal of the American Statistical Association*, 90, 1357-1363.
7. Liseo, B. (1993). Elimination of Nuisance Parameters with Reference Noninformative Priors, *Biometrika*, 80, 295-304.
8. Snedecor, G. W. and Cochran, W. G. (1980). *Statistical Methods*, Ames, IA: Iowa State University.

[2002년 9월 접수, 2002년 10월 채택]