지수분포 모수함수 간의 다중비교에 관한 연구

김대황1) 김혜중2)

요 약

본 연구에서는 확률모형의 모수로부터 얻어지는 여러형태의 함수간의 크기를 다중비교하는 방법을 제안하고자 한다. 이 방법은 비교대상인 모수 함수간의 선호확률을 베이지안 방법으로 추정하고, 이들로부터 얻어지는 선호행렬을 이용한 새로운 다중비교법이다. 이러한 방법의 제안에 필요한 이론과 비교기준을 고안하였으며, 응용 예로, 제안된 방법을 s개의 독립인 지수분포 모수의 기하평균 크기비교에 적용하였다.

주요용어 : 선호행렬, 다중비교, 기하평균, 지수분포

1. 서 론

실생활에서 우리는 어떤 동일한 상품에 대해 어느 회사의 상품을 살 것인가를 결정하는 문제에 종종있다. 이러한 경우는 여러 모집단의 모수의 크기를 비교하는 문제로서, 이를 해결하기 위해 다중비교와 모수들의 동시신뢰구간 문제(Bauer, 1997), 모수의 중요도의 관점에서 최적의 모집단을 선택(Bechhofer, Santner, Goldsman; 1995참조, Kim과 Nelson; 2001)하는 방법과 같은 여러 가지가 있다. 특히, 일반적으로 이용되어지는 다중비교는 모형에 대한 정규성이나 등분산성을 가정해야고, 모수들의 함수 형태인 경우는 사용하지 못한다는 단점이 있다. 따라서, 본 연구에서는 확률모형의 모수로부터 얻어지는 여러형태의 함수간의 크기를 다중비교 하는 방법을 제안하고자 한다. 이 방법은 비교대상인 모수 함수간의 선호확률을 베이지안 방법으로 추정하고, 이들로부터 얻어지는 선호행렬로부터 행-합 점수를 이용한 새로운 다중비교법이다. 이러한 방법의 제안에 필요한 이론과 비교기준을 고안하였으며, 응용 예로, 제안된 방법을 s개의 독립인 지수분포 모수의 기하평균 크기비교에 적용하였다.

본 연구의 구성은 다음과 같다. 2절에서는 Tibshirani(1989)의 방법을 이용하여 관심모수에 대한 무정보적사전분포를 유도하고, 3절에서는 선호확률을 이용한 다중비교 방법과 그에 따른 이론을 제안하고, 4절에서는 베이지안 추론을 위한 계산방법을 소개하였다.

2. 사전분포

여러 가지 무정보적사전분포중 s개 모수에 대한 기하평균에 대한 사전분포로서 Tibshirani(1989)가 제안한 방법을 이용한 확률대응사전분포를 유도하고자 한다. 이를 위해 $X_l(k)$, $l=1,\cdots,s$, $k=1,\ldots,K$를 평균이 $\lambda_l(k)$를 따르는 확률변수라하고, 관심모수를 각 모집단의 기하평균 $\delta_k=\left(\prod_{l=1}^{s}(\lambda_l(k)\right)^{1/s}$라 하자.

관심모수인 기하평균 δ와 장애모수 ζ를 각각 다음과 같이 정의하자.

---

1) (100-715) 서울 중구 필동 동국대학교 대학원 통계학과, 대학원생
2) (100-715) 서울 중구 필동 동국대학교 통계학과 교수

$$\delta = \left(\prod_{i=1}^{s}\lambda_i\right)^{1/s}, \quad \zeta_i = \zeta_i(\lambda), \quad i = 2,3,\cdots,s.$$

그리고, $\zeta_i^j = \partial\zeta_i(\lambda)/\partial\lambda_j$, $\eta_{(j)} = 1/s\left(\prod_{i=1}^{s}\lambda_i\right)^{1/s}\lambda_j^{-1}$라고 가정하면 Jacobian행렬은 다음과 같아진다.

$$\frac{\partial(\delta,\zeta)}{\partial(\lambda)} = \begin{pmatrix} \eta_{(1)} & \eta_{(2)} & \cdots & \eta_{(s)} \\ \zeta_2^1 & \zeta_2^2 & \cdots & \zeta_2^s \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_s^1 & \zeta_s^2 & \cdots & \zeta_s^s \end{pmatrix}.$$

따라서, Fisher의 기대정보행렬의 역행렬은

$$I^{-1}(\delta,\zeta) = \left(\frac{\partial(\delta,\zeta)}{\partial(\lambda)}\right)\left(\frac{\partial(\delta,\zeta)}{\partial(\lambda)}\right)^T = \begin{pmatrix} \sum_{j=1}^{s}\eta_{(j)}^2 & \phi^T \\ \phi & A \end{pmatrix}$$

이 된다. 여기서, $\phi = \left(\sum_{j=1}^{s}\eta_{(j)}\zeta_2^j, \cdots, \sum_{j=1}^{s}\eta_{(j)}\zeta_s^j\right)^T$이고, A는 $(s-1)\times(s-1)$인 정칙행렬(nonsingular matrix)이다.

만약 $\phi = 0$이라면 δ와 ζ는 서로 직교하게 된다. 이러한 가정으로부터 $s-1$개의 동질적인 선형인 편미분방정식을 유도할 수 있다. $\psi(\lambda_i^2-\lambda_j^2,\ i<j)$의 형식을 가진 어떤 함수가 방정식의 해가 될 수 있다. 예를 들어, $\zeta_i(\lambda) = \nu_i = (\lambda_1^2-\lambda_i^2)/2$, $i = 2,3,\cdots,s$를 가정할 수 있다. 그러면 δ와 ζ들은 직교하고, Jacobian 행렬은 다음

$$\frac{\partial(\delta,\zeta)}{\partial(\lambda)} = \begin{pmatrix} \eta_{(1)} & \eta_{(2)} & \eta_{(3)} & \cdots & \eta_{(s)} \\ \lambda_1 & -\lambda_2 & 0 & \cdots & 0 \\ \lambda_1 & 0 & -\lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \lambda_1 & 0 & 0 & \cdots & -\lambda_s \end{pmatrix},$$

이 되며, 이 행렬의 행렬식(Determinant)은 $1/s\left(\prod_{i=1}^{s}\lambda_i\right)^{1+1/s}\left(\sum_{i=1}^{s}\lambda_i^{-2}\right)$이 된다. 위의 행렬을 이용하면 Fisher의 정보행렬은 다음과 같다.

$$I(\delta,\zeta) = \begin{pmatrix} \sum_{i=1}^{n}\eta_{(i)}^2 & 0 & \cdots & 0 \\ 0 & \lambda_1^2+\lambda_2^2 & \cdots & \lambda_1^2 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & \lambda_1^2 & \cdots & \lambda_1^2+\lambda_s^2 \end{pmatrix}^{-1}$$

Tibshirani의 방법(Berger; 1992 참조)을 이용하면 다음과 같은 사전분포를 구할 수 있다.

$$\pi(\delta,\zeta) \propto g(\zeta)\left(\sum_{j=1}^{s}\eta_{(j)}^2\right)^{-1/2}.$$

여기서, $g(\zeta)$는 임의의 양의 함수이다. 위의 사전확률분포를 변환시켜 원래의 모수 λ의 사전확률분포를 구하면

$$\begin{aligned} \pi(\lambda) &\propto g(\zeta(\lambda))\left(\sum_{j=1}^{s}\eta_{(j)}^2\right)^{-1/2}\left|\frac{\partial(\lambda,\zeta)}{\partial(\lambda)}\right| \\ &\propto g(\zeta(\lambda))\frac{\frac{1}{s}\left(\prod_{i=1}^{s}\lambda_i\right)^{1+1/s}\sum_{i=1}^{s}\lambda_i^{-2}}{\frac{1}{s}\sqrt{\left(\prod_{i=1}^{s}\lambda_i\right)^{2/s}(\lambda_1^{-2}+\lambda_2^{-2}+\cdots+\lambda_s^{-2})}} \\ &\propto g(\zeta(\lambda))\left(\prod_{i=1}^{s}\lambda_i\right)\sqrt{\sum_{i=1}^{s}\lambda_i^{-2}} \end{aligned}$$

이다. 만약, $g(\zeta(\lambda))=1$ 이면 사전확률분포는 다음과 같이 유도되어 진다.

$$\pi(\lambda)\propto\left(\prod_{i=1}^{s}\lambda_i\right)\sqrt{\sum_{i=1}^{s}\lambda_i^{-2}} \tag{1}$$

3. 다중비교를 위한 이론 및 비교기준

$\Theta=\{\theta_{ij}\},\ (i,j=1,\ldots,K)$는 $\theta_{ij}+\theta_{ji}=1$, $\theta_{ii}=1/2$을 만족하는 $\theta_{ij}=\Pr(i\rightarrow j)$는 j번째 모집단의 δ_j보다 i번째 모집단의 δ_i를 선호할 선호확률로 이루어진 선호행렬이라고 하자. 만약, $\{1,\ldots,K\}$에서 작은 것을 선호한다면, 선호순서(preference order)는 $P=(p_1,\ldots,p_K)$로 정의된다. 즉, 선호순서는 δ_{p_i}를 δ_{p_j}보다 선호할 때, $p_i<p_j$와 같이 오름차순으로 정렬하여 얻어진다. Θ로부터 선호순서 정하기 위한 두 가지 이행(transitivity)조건을 소개한다.

약확률이행성(weak stochastic transitivity; C_1) : 모든 세쌍 (i,j,l)에 대해,

$$\theta_{ij}\geq 1/2,\ \theta_{jl}\geq 1/2\text{이면 } \theta_{il}\geq 1/2\text{이다.} \tag{2}$$

강확률이행성(strong stochastic transitivity; C_2) : 모든 세쌍 (i,j,l)에 대해,

$$\theta_{ij}\geq 1/2,\ \theta_{jl}\geq 1/2\text{이면 } \theta_{il}\geq \max(\theta_{ij},\theta_{jl})\text{을 만족한다.} \tag{3}$$

조건 C_2를 다음과 같은 형태로 표현할 수 있다.

$$\text{모든 쌍 } (i,j)\text{에 대해서, } \theta_{ij}\geq 1/2\text{는 } \theta_{il}\geq\theta_{jl}\text{를 의미한다.} \tag{4}$$

여기서, $l=1,\ldots,K$이다.

정리. *만약 강확률이행조건 C_2를 만족하면, $\{1,\ldots,K\}$ 집단에 대한 최적상체비교순서 $P=(p_1,\ldots,p_K)$는 선호확률행렬 Θ의 행합점수에 의한 순서와 일치한다.*

증명. *일반성을 잃지않고, $P=(1,2,\ldots,K)$라 하자. $G(X,R,\Theta)$에 대한 조건 C_2는 다음의 관계를 의미한다. (3.3)으로부터 $l>k$, $k=1,\ldots,K-1$, $l=1,\ldots,K$에 대하여 $\theta_{kl}\geq\theta_{k+1l}$, 정의로부터 $l=k$에 대해 $\theta_{kl}=1/2$, (3.2)에 의해 $l<k$에 대해 $\theta_{lk+1}\geq\theta_{lk}$임을 알 수 있다. 그리고, $\theta_{kl}=1-\theta_{lk}$와 $\theta_{k+1l}=1-\theta_{lk+1}$이므로 $\theta_{kl}\geq\theta_{k+1l}$이다. Θ의 k번째와 $k+1$번째 행합을 분해하면*

$$\sum_{l=1}^{K}\theta_{kl}=\sum_{l<k}\theta_{kl}+\theta_{kk}+\sum_{l>k}\theta_{kl}$$

와

$$\sum_{l=1}^{K}\theta_{k+1l}=\sum_{l<k}\theta_{k+1l}+\theta_{k+1k}+\sum_{l>k}\theta_{k+1l}$$

을 각각 구할 수 있다. 따라서, $P=(1,2,\ldots,K)$에 대해 $\theta_{kk}>\theta_{k+1k}$이므로 $k=1,\ldots,K-1$에 대해

$$\sum_{l=1}^{K}\theta_{kl}>\sum_{l=1}^{K}\theta_{k+1l},$$

을 만족한다.

4. 지수분포 모수함수 간의 다중비교

4.1 사후선호확률

사전분포 (1)을 이용하면, $D=(x_{l1}(k),\ldots,x_{lN_l}(k))$가 주어졌을 때, $\lambda_l(k)$, $l=1,\ldots,s;k=1,\ldots,K$의 결합사후분포를 다음과 같이 구할 수 있다.

$$\pi(\boldsymbol{\lambda}(1),\ldots,\boldsymbol{\lambda}(K)\mid D)\propto\prod_{k=1}^{K}\frac{\sqrt{\sum_{l=1}^{s}\lambda_l(k)^{-2}}}{\left(\prod_{l=1}^{s}\lambda_l(k)\right)^{N_l-1}}\exp\left\{-\sum_{l=1}^{s}\frac{\sum_{j=1}^{N_l}x_{lj}(k)}{\lambda_l(k)}\right\}. \tag{5}$$

여기서 관심의 대상은 K 모집단의 모수 s개에 대한 기하평균의 상대적 크기에 대한 순서를 구하는 것이다. 특히, 사후선호확률 θ_{ij}로 이루어진 선호확률행렬, $\Theta=\{\theta_{ij}\}$를 구하는 것이 목적으로 다음을 이용하여 선호확률 $\theta_{ij}=\Pr(\delta_i\rightarrow\delta_j\mid data)$를 구할수 있다.

$$\theta_{ij}=\frac{\pi_i}{\pi_i+\pi_j} \tag{6}$$

여기서, $\pi_i=pr(\delta_i-\delta_m<0\mid Data)$, $\pi_j=pr(\delta_j-\delta_m<0\mid Data)$, $\delta_m=\sum_{k=1}^{K}\delta_k/K$이다.

따라서, 식(6)의 사후선호확률은 $\pi_k=E_p[I\{\delta_l-\delta_m<0\}\mid Data]$, $k=1,\ldots,K$를 이용하여 구할 수 있다. 여기서, $I(\cdot)$는 지시함수이다.

선호확률 θ_{ij}를 계산하기 위한 π_k의 계산은 사후분포 (5)이 복잡하므로 유용하지 않다. 이 점에서 Chen과 Shao(1999)의 가중몬테칼로방법이 이 확률을 계산하는데 대안으로 사용되고 있다. 다음 절에서는 이 접근법에 대해 논의 할 것이다.

4.2 가중몬테칼로방법

사후분포가 (5)와 같이 복잡하여 이 분포로부터 표본추출이 용이하지 않으면 마코브체인 몬테칼로 방법을 적용하는 대신 주표본 함수를 이용한 가중 몬테칼로 방법이 사용된다.

주표본 함수 $g(\boldsymbol{\lambda})$를 아래와 같이 가정하여 보자.

$$g(\boldsymbol{\lambda}(1),\ldots,\boldsymbol{\lambda}(K))=\prod_{k=1}^{K}\frac{1}{\left(\prod_{l=1}^{s}\lambda_l(k)\right)^{N_l-1}}\exp\left\{-\sum_{l=1}^{s}\frac{\sum_{j=1}^{N_l}x_{lj}(k)}{\lambda_l(k)}\right\}.$$

위 식에서 k번째 모집단의 $\lambda_{-l}(k)=(\lambda_1(k),\cdots,\lambda_{l-1}(k),\lambda_{l+1}(k),\cdots,\lambda_s)'$ 와 D가 주어졌을 때 $\lambda_l(k)$ 의 조건부 분포를 구하면 $l=1,2,\cdots,s$, $k=1,\ldots,K$에 대해 아래와 같아진다.

$$g(\lambda_l(k)\mid\lambda_{-l}(k),D)\propto\frac{1}{\lambda_l(k)^{N_l-1}}\exp\left\{-\frac{\sum_{j=1}^{N_l}x_{lj}(k)}{\lambda_l(k)}\right\}. \tag{7}$$

그러므로, $\lambda_l(k)^{-1}$ 의 조건부 분포는 모수가 $\gamma=N_l-2$, $\beta=1/(\sum_{j=1}^{N_l}x_{lj}(k))$ 인 감마분포를 따른다.

$\{\lambda_l^{(t)}(k),t=1,\ldots,m;k=1,\ldots,K;l=1,\ldots,s\}$를 주표본함수로부터 추출한 확률표본이라고 한다면, 주표본가중치는

$$w_k^{(t)}=\frac{p(\lambda_l^{(t)}(k);l=1,\ldots,s\mid Data)}{g(\lambda_l^{(t)}(k);l=1,\ldots,s)}, \tag{8}$$

이 되며, $\delta_k^{(t)} = \left(\prod_{l=1}^{s} \lambda_l^{(t)}(k) \right)^{1/s}$, $k=1,\ldots,K$에 대해

$$\pi_k = E_p[I\{\delta_k - \delta_m < 0\} | Data]$$

$$= E_g\left[I\{\delta_k - \delta_m < 0\} \frac{p(\lambda_l(k); l=1,\ldots,s|Data)}{g(\lambda_l(k); l=1,\ldots,s)} \right],$$

이 만족하므로 π_k의 가중몬테칼로 추정치는 다음과 같이 구할 수 있다.

$$\hat{\pi}_k = \hat{p}(\delta_i - \delta_m < 0 \mid Data) = \frac{\sum_{t=1}^{m} w_k^{(t)} I\{\delta_k^{(t)} - \delta_m^{(t)} < 0\}}{\sum_{t=1}^{m} w_k^{(t)}}$$

따라서, 사후선호확률의 추정치는

$$\hat{\theta}_{ij} = \frac{\hat{\pi}_i}{\hat{\pi}_i + \hat{\pi}_j}$$

이 된다.

Geweke(1989)은 $m \to \infty$일 때, $k=1,\ldots,K$에 대해

$$\hat{\pi}_k \xrightarrow{a.s.} \pi_k$$

임을 보였다. 식 (8)는 $p(\lambda_l(k); l=1,\ldots,s|Data)$와 $g(\lambda_l(k); l=1,\ldots,s)$의 비가 상수가 되도록 만드는 주표본 함수로부터 추출된 표본을 이용하여 사후확률의 몬테칼로추정치를 구한 것이다. $\hat{\pi}_k$의 실험의 정확도를 측정하는 도구로서 $\hat{\pi}_k$의 실험표준오차는 매우 중요하다. 대수의 법칙에 의하면

$$m^{1/2}(\hat{\pi}_k - \pi_k) \to N(0, \sigma^2),$$

여기서, $\sigma^2 = \sigma_V^2 / S^2$이며,

$$\hat{\sigma}_V^2 = \frac{\sum_{t=1}^{m} (w_k^{(t)} I\{\delta_k^{(t)} - \delta_m^{(t)} < 0\} - w_k^{(t)} \hat{\theta}_k)^2}{m.} \text{ 과 } \hat{S} = 1/m \sum_{t=1}^{m} w_{ij}^{(t)}$$

$\hat{\sigma}^2 = \hat{\sigma}_V^2 / \hat{S}^2$으로 추정할 수 있다.

참고문헌

Bauer, P.(1997). A note on multiple testing procedure in dose finding. *Biometric* **53**, 1125-1128.

Bechhofer, R. E., Santner, T. J., and Goldsman, D. M.(1995). *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*, New York: Wiley.

Chen, M. H. and Shao, Q. M.(1999) "Monte Carlo estimation of Bayesian credible and HPD intervals", *Journal of Computational and Graphical Statistics* 8, 69-92.

Geweke, J.(1989). Bayesian inference in econometrics models using Monte Carlo integration. *Econometrica* **57**, 1371-1340.

Huzurbazar, S and Butler, R. W.(1998), "Importance Sampling for p-value Computations in Multivariate Tests", *Journal of Computational and Graphical Statistics*, Vol. 7,

342-355.

Kim, S. and Nelson, B. L. (2001). A fully sequential selection procedure for indifference-zone selection in simulation. *Transactions on Modeling and Computer Simulation* **11**, 251-273.

Tibshirani, R. (1989), "Non-Informative Priors for One Parameter of Many", *Biometrika*, 76, 604-608