

# BAYES YÖNTEMİNİN META-ANALİZİNDE KULLANILMASI

Filiz KARAMAN\*

## ÖZET

*İstatistiksel yorumlamalarda Bayes yönteminin kullanımına olan ilgi son zamanlarda giderek artmaktadır. Bu çalışmada meta analizi için Bayes yöntemi kullanılarak sonsal dağılım elde edilecektir. Eğer sonsal dağılım kolay elde edilebilir bir dağılım değilse, bu dağılım alternatif bir yaklaşım olarak simülasyon yardımıyla elde edilebilmektedir. Örnek verisi olarak propranolol ve placebo kullananlar için çok merkezli Beta-Blokör kalp krizi tecrübesi incelenecektir. Bayes yorumuna ulaşabilmek için BUGS paket programı yardımıyla MCMC tekniği kullanılarak sonsal dağılım elde edilmiştir. Sonuç olarak, sonsal dağılım istatistiklerine göre beta-Blokörler kalp krizi ölümlerini azaltır.*

**Anahtar Kelimeler:** BUGS, MCMC yöntemi, Meta analizi, Beta-blokör

## 1. GİRİŞ

Farklı disiplinler için, istatistiksel yorumlamalardaki Bayes yönteminin kullanımına olan ilgi son zamanlarda giderek artmaktadır. Örneğin, bu yöntemin tıp, eczacılık ve biyoloji alanında kullanılması ile ilgili makaleler bulunmaktadır. Berry ve Stangl (1996), Bayes yöntemini kullanarak sağlık alanında yapılan makaleleri birleştirmiştir.

Meta analizi çok sayıda yapılmış çalışma sonuçlarının toplanmasını ve bu sonuçları kullanarak istatistiksel analizin yapılmasını sağlar. Amacı, kaynakçalarda bulunan farklı çalışmalardaki uygun olan verileri birleştirerek ilgilenilen konu hakkında genel bir sonuca varmaktır. Meta analizi klinik araştırmalarda çok kullanılır. Karasoy ve Ata (2008) yaşam verilerinden elde edilen sonuçları birleştiren farklı yaklaşımları inceleyerek, farklı makalelerden elde edilen verilere Meta analizi uygulamışlardır. Sutton vd. (2008) bireysel ve tüm veri seviyeleri için Meta analizi çalışması yapmışlardır.

Meta analizi kullanılarak, amacı, çalışma yöntemi, kalitesi, örnekleme çerçevesi, anakütlesi ve raporlanan bilginin farklı olarak yapıldığı çalışmalar birleştirilir. Bu durum filozofik ve pratik problemler ortaya çıkarır. Ayrıca küçük örneklerle, yeni bilgi elde edildiğinde sonuçların güncellendiği metotlarla ve yayınlamakta, seçmekte, vb., oluşan sistematik hatalarla uğraşmak endişe vermektedir. Bu endişeleri azaltabilmek için Meta analizinde (Stojanovski ve Mengersen, 2006) Bayes yönteminden faydalanılmaktadır.

Bu çalışmada, Meta analizi için Bayes yöntemi kullanılarak sonsal (posterior) dağılıma ulaşılmaya çalışılacaktır. Bayes yorumuna ulaşabilmek için BUGS (Bayesian inference Using Gibbs sampling technique) paket programı yardımıyla Markov Chain Monte Carlo (MCMC) tekniği kullanılarak örnek üzerinde konu hakkında bilgi verilmeye çalışılacaktır.

\*Yrd. Doç. Dr, Yıldız Teknik Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, 34210 İstanbul, e-posta: [fkaraman@yildiz.edu.tr](mailto:fkaraman@yildiz.edu.tr)

## 2. YÖNTEM

Bayesci yaklaşımla veriler ve bilinmeyen parametreler arasındaki yapısal ilişki tanımlanır (Broemeling, 2007). Burada bahsedilen veriler ve parametreler belirsizlik altındaki rassal değişkenler olarak düşünülür. Eczacılık ve tıp çalışmalarında, belirli bir ilaç tedavisi uygulanan canlıların ortalama iyileşmesi, belirti seviyesi ve yaşam süresi olasılıkları alternatif ilaç tedavisi altında olan canlılara göre daha iyi olup olmadığı karşılaştırılmaya çalışılır. Meta analizi, genellikle yeni geliştirilen ilaçlarla alternatif ilaçların veya placeboların (kontrol grubu) karşılaştırılmasında kullanılan bir yöntemdir (Stojanovski ve Mengersen, 2006).

Meta analiz bazı basamaklar içerir. Bunlar; araştırma sorusunu, çıktıyı ve toplanacak çalışmaları tanımlamak, bu çalışmaları seçerek istenilen bilgiyi toplamak, model geliştirmek, hesaplama metodunu belirlemek, sonuçları yorumlamak ve modelin hassasiyetini değerlendirmektir. Meta analizi için Bayes yöntemi kullanılarak diğer yapılan Meta analizlerine ek olarak hastalık farklılıkları hakkında önsel bilgi modele eklenir. Bu da model içinde belirsizliğin farklı çeşitlerde ifadesi ve bunların önsel dağılım kurularak modele kaynaştırılmasıyla olur. Bu basamaklar aşağıdaki paragraflarda daha detaylı anlatılacaktır (Stojanovski ve Mengersen, 2006).

Meta analizine başlamadan önce ilk basamak araştırma sorusunu bulmaktır. Bu bölümün bir kısmı çıktıların tespitini içerir. Meta analiz için veri oluşturan her çalışmanın çıktıları; meta analizinden yapılacak yorumlara ve verinin tipine bağlıdır. İkili veriler için risk farkı, göreceli risk ve odds oranı ortak çıktılarıdır. Sürekli veriler için ise her çalışmadaki ortalama ve standart hata potansiyel bir çıktıdır (Stojanovski ve Mengersen, 2006).

Makalede kullanacağımız notasyonumuzu tanımlayacak olursak:  $i = 1, \dots, k$  olarak ele alınırsa, burada  $i$  çalışma numarasını ve  $k$  ise çalışma sayısını belirtir.  $\theta_i$ ,  $i$ nci çalışmadaki doğru etki ve  $\theta = (\theta_1, \dots, \theta_k)'$  da bütün çalışmalardaki doğru etkilerin vektörüdür.  $i$  çalışmasından gözlenen etki  $Y_i$  ile gösterilecek olursa,  $Y = (Y_1, \dots, Y_k)'$  gözlenen etkilerin vektörüdür.

Bayesci yaklaşımda WinBUGS programında parametrelerin dağılımı tanımlanırken varyans yerine kesinlik parametresi yazılır. Özet istatistiğin  $i$ 'inci çalışmadaki değişimi  $\sigma_i^2$  ve kesinliği  $\tau_i$  olup burada  $\tau_i = 1/\sigma_i^2$  dir. Eğer  $k$  tane çalışma yer değiştirilerek kullanılabilirse sabit etki modeli, bütün  $\theta_i$ 'lerin eşit olduğunu kabul eder. Dolayısıyla bütün çalışmalar tek bir anakütleden rassal olarak çekilmiş olur. Bir rassal etki modeli her  $\theta_i$ 'nin ortalaması  $\mu$  olan etkili anakütleden çekilmiş bir örneklemdir (Bkz. tablo 1).

**Tablo 1. Üç seviyeli rassal değişkenlerle basit sıralı Bayesci yaklaşım modeli.**

Parametre Adları	Parametrelerin Simgeleri	Dağılımlar
Genel Parametreler	$\mu, \tau^2$	$P(\mu), P(w^2)$
Özel Çalışma Parametreleri	$\theta_1, \dots, \theta_k$	$P(\theta_i   \mu, w^2)$
Veriler	$Y_1, \dots, Y_k$	$P(Y_i   \theta_i, \sigma_i^2)$

Tablo 1’de Global parametreler  $\mu$ : genel (overall) ortalama ve  $\tau^2$  onun varyansını ifade eder. Bayes analizi veri verilmişken  $\theta_i$  ve  $\mu$  için ortak sonsal dağılım bulmaya çalışır.  $E(\theta_i) = \mu$  ve varyansı  $w^2$  çalışmalar arası heterojenliktir.

Önceki çalışma m gözleme sahipse ( $x = x_1, \dots, x_m$ ) ve sürekli bir dağılımdan geliyorsa, önsel bilgi  $\mu$  ve  $\sigma^2$  için (1) ile ifade edilir:

$$f(x|\mu, \sigma) \propto \left[ \sqrt{m} / \sqrt{2\pi\sigma} \right] \exp - (m/2\sigma^2) (\bar{x} - \mu)^2 \\ \left[ (2\pi)^{-(n-1)/2} \sigma^{-(n-1)} \right] \exp - (1/2\sigma^2) \sum_{i=1}^{i=m} (x_i - \bar{x})^2 \quad (1)$$

Bu iki parametrelili normal ailesi için eşlenik yoğunluk fonksiyonudur ve normal-inversegama yoğunluk fonksiyonu olarak adlandırılır (Broemeling, 2007).

Dağılımı normal( $\mu, 1/\tau$ ) anakütleden gelen n tane rassal örneklem olan ( $x = x_1, \dots, x_n$ )’i ele alalım. Olabilirlik fonksiyonu (2) ile ifade edilir:

$$L(\mu, \tau/x) \propto \tau^{n/2} \exp - \left( \tau/2 \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + n(\mu - \bar{x})^2 \right] \right) \quad (2)$$

Önsel bilginin belirsiz olduğunu,  $f(\mu, \tau) \propto 1/\tau$  ’nun uygun olduğunu varsayarsak ve yukarıdaki olabilirlik fonksiyonu ile birleştirecek olursak parametrelerin sonsal yoğunluk fonksiyonu (3) ile ifade edilir:

$$f(\mu, \tau/x) \propto \tau^{n/2-1} \exp - \left( \tau/2 \left[ n(\mu - \bar{x})^2 + \sum_{i=1}^n (x_i - \bar{x})^2 \right] \right) \quad (3)$$

Buradan da  $\tau$  ve  $\mu$  için ayrı ayrı integral alınarak marjinal yoğunluk fonksiyonları bulunabilir. Ortak yoğunluk fonksiyonlarında integral alınması zor olduğu durumlarda Gibbs algoritması diye adlandırılan algoritmayı kullanmak iyi bir alternatiftir (Broemeling, 2007).

Sonsal dağılım, simülasyon yardımıyla elde edilebilir. Temel hesaplama tekniği MCMC’dir. Bu teknik, simülasyon yardımıyla şartlı sonsal dağılım oluşturur. Bunlardan bazı ortak algoritmalar Gibbs, Metropolis-Hastings, Slice ve Tam örneklemeleridir. Kullanıcının olabilirlik ve önsel dağılım olarak belirlediği modeli verilerek, BUGS gerekli şartlı sonsal dağılımı oluşturur ve MCMC örnekleme ile sonsal dağılımın özet istatistiklerinin elde edilmesini sağlar (Ntzoufras, 2009).

### 3. BULGULAR

Meta analizinin WinBUGS programında uygulanması için Böhning vd. (2008) deki veriler kullanılmıştır. Veriler aylık bazda toplanmış olup hastalar 12. aydan 40. aya kadar takip edilerek (ortalama 25 aylık) toplanmıştır. Araştırma bilgileri Propranolol ve placebo (kontrol grubu) kullananlar için 31 merkezdeki çok merkezli Beta-Blokör kalp krizi tecrübesi içerir. Veriler, gözlem yapılan süre boyunca her bir merkezde bulunan

Propranolol kullanan toplam kişi sayısını, bu kişiler arasından ölen toplam kişi sayısını ve aynı zamanda placebo kullanan toplam kişi sayısını, bu kişiler arasından ölen toplam kişi sayısını göstermektedir. WinBUGS 1.4 versiyonunun Cilt 1 yardım bölümünden alınan BUGS kodu Şekil 1’de verilmiştir.

BUGS kodunu inceleyecek olursak;  $rc[i]$  i’inci denemedeki placebo kullananların ve  $rt[i]$  i’inci denemedeki Propranolol kullananların ölüm sayılarını gösterir ve binom dağılımına sahiptirler. Placebo kullananların risk oranı  $pc[i]$  ve Propranolol tedavisi olanların risk oranı  $pt[i]$  ile gösterilmiştir. Bu oranları başarı olasılığı olarak kabul edersek, her bir denemedeki kontrol ve tedavi grupları için farklı başarı olasılıkları verilmiş ve bu olasılıklar logit ölçeğine çevrilmiştir. Yapılan Gibbs uygulamasında normal ve inversegama dağılımları eşlenikden dolayı seçilmiştir. Ayrıca gama dağılımı, normal dağılıma sahip kesinlik parametresi için (varyansın tersi) eşlenik önsel dağılımdır. Logit parametreleri için bilgilendirici olmayan normal önsel dağılım verilmiştir (Broemeling, 2007).

```

Model {
  for( i in 1 : Num ) {
    rc[i] ~ dbin(pc[i], nc[i])
    rt[i] ~ dbin(pt[i], nt[i])
    logit(pc[i]) <- mu[i]
    logit(pt[i]) <- mu[i] + delta[i]
    mu[i] ~ dnorm(0.0,1.0E-5)
    delta[i] ~ dnorm(d, tau)
    d ~ dnorm(0.0,1.0E-6)
    tau ~ dgamma(0.001,0.001)
    sigma <- 1 / sqrt(tau)
  }
}

```

Şekil 1. BUGS kodu

Kodlardan anlaşılacağı gibi parametrelerin üç seviyesi vardır. Bunlar; başarı olasılıkları, logit parametreleri olan  $\mu$ ’lar, delta’lar ve son olarak ikinci seviyedeki delta parametrelerinin ortalaması olan  $d$ ’dir. Burada  $\mu$ ’lar ortalaması 0.0 ve kesinliği 0.00001 olan bağımsız normal dağılıma ve  $d$  ortalaması 0.0 ve kesinliği 0.000001 olan bağımsız normal dağılıma sahiptir. Deltaların dağılımı da normaldir Bir diğer üçüncü seviye parametresi, delta parametrelerinin kesinliğini ifade eden  $\tau$ ’dur ve dağılımı bilgilendirici olmayan parametreleri 0.001 olan Gama’dır.  $\tau$  parametresi varyansın tersidir.  $\sigma$  standart sapmayı ifade eder (Broemeling, 2007).

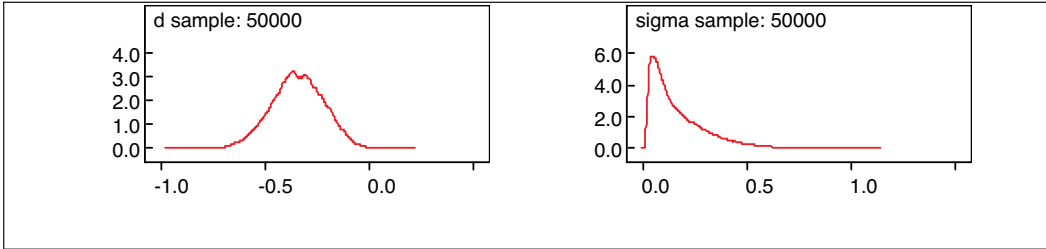
Burada ilgilenilen ana parametre  $d$ ’dir. Eğer  $d$  değeri sıfır ise kontrol ve tedavi grubu ölüm oranı arasında fark yoktur denir. Burn-in periyodu 1000 olarak alınmış ve tüm tahmin için 50000 iterasyon yapılmıştır. Sonsal dağılım için elde edilen istatistikler Tablo 2’de, grafikler Tablo 3-6’da verilmiştir.

Tablo 2. Sonsal dağılım istatistikleri

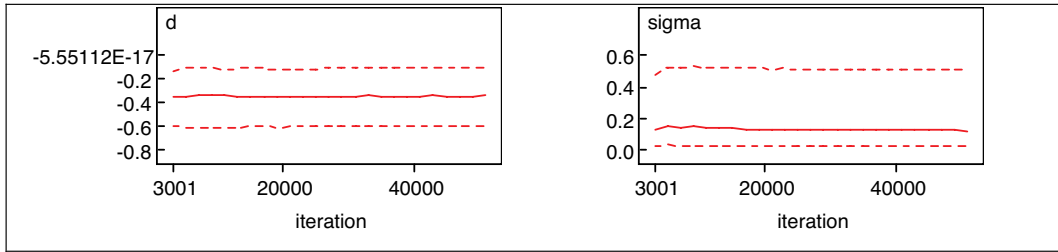
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
d	-0.3423	0.1251	0.003851	-0.5919	-0.3416	-0.107	1001	50000
sigma	0.1671	0.1331	0.004156	0.02687	0.1244	0.5107	1001	50000

d beta-blokör etkisini ölçen sonsal dağılım parametresidir. Sigma, standart sapmayı ifade eder.

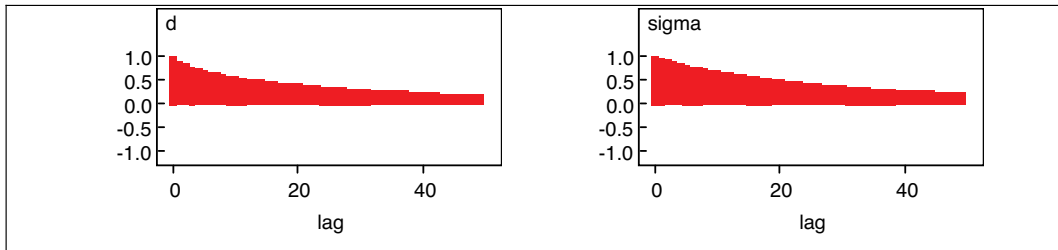
Tablo 3. Yoğunluk fonksiyonu



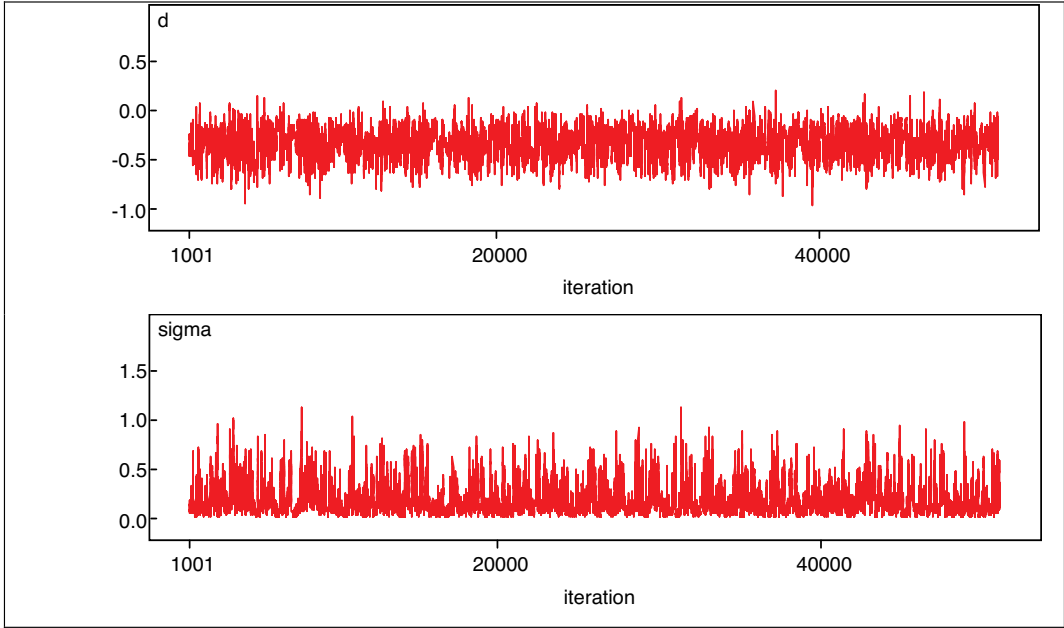
Tablo 4. Sıklık derecesi (Running Quantiles)



Tablo 5. Otokorelasyon fonksiyonları



Tablo 6. Zaman serileri



Tablo 2'deki düğüm  $d$  değerine bakıldığında negatif aralıkta bulunduğu ve sıfır değerini içermediğinden logit ölçeğinde delta sıfır değildir. Diğer bir ifade ile beta-blokör kalp krizi ölümlerini azaltır.

Bütün istatistik değerler (Tablo 2) ve grafikler (Tablo 3-6) model parametrelerinin yakınsak bir dağılımdan geldiğini desteklemektedir: Model parametreleri için MC error ve sd değerlerine bakarsak MC error sd'nin %5'inden küçüktür (Ntzoufras, 2009).

#### 4. SONUÇ

Günümüzde, bilgisayar yardımıyla simülasyon kullanarak hesaplaması zor ve zaman alıcı işlem sonuçlarına oldukça pratik ve kısa zamanda ulaşılabilir. Bayes yöntemi kullanılarak sonuç elde edilmesi, istatistiksel paket programların uygulamada sağladığı kolaylıklar sebebiyle giderek artmaktadır. Bu çalışmada, meta analizinde Bayes yöntemi kullanılmış ve bilgilendirici olmayan önsel dağılım verildiğinde MCMC tekniği yoluyla WinBUGS istatistiksel paket programı kullanılarak sonuca ulaşılmaya çalışılmıştır. Ele alınan tedavi yönteminin kontrol grubu karşısında etkin olup olmadığı test edilmiştir. Sonuçta, bu yöntemin etkili olduğu görülmüştür. Bu program yardımıyla başka veriler için kullanılan yöntemlerin tedavi için etkin olup olmadığı kolayca görülebilir.

## 5. KAYNAKLAR

Berry, D. A., Stangl, D. K., Eds., 1996. Bayesian Biostatistics. New York: Marcel Dekker.

Böhning, D., Kuhnert, R., Rattanastiri, S., 2008. Meta-Analysis of Binary Data Using Profile Likelihood. Boca Raton: Chapman&Hall/CRC.

Broemeling, L. D., 2007. Bayesian Biostatistics and Diagnostic Medicine. Boca Raton: Chapman&Hall/CRC.

Karasoy, D., Ata, N. 2008. Yaşam Verilerinin Meta Analizi. SDÜ Fen-Ed. Fak. Fen Dergisi (e-dergi), 3(2), 211-218.

Ntzoufras, I., 2009. Bayesian Modeling Using WinBUGS. New Jersey: John Wiley and Sons.

Sutton, A.J., Kendrick, D., Coupland, C. A. C., 2008. Meta-Analysis of individual- and aggregate- level data. Statistics in Medicine, 27, 651-669.

Stojanovski, E., Mengersen, K. L., 2006. Bayesian Methods in Meta-Analysis. Encyclopedia of Biopharmaceutical Statistics, 1, 1, 1-7.

## USAGE OF BAYESIAN METHOD IN META ANALYSIS

### ABSTRACT

*Usage of Bayesian methods in statistical inferences takes improved attention for the last decades. In this study, Bayesian method to meta analysis will be used to get posterior distribution. If the posterior distribution is not easy to have, as an alternative approach posterior distribution will be gathered by simulation. As an example, data from the multicenter Beta-Blocker Heart attack trial comparing propranolol with placebo will be used. To reach Bayesian inference, MCMC techniques will be used to get posterior distribution by the help of BUGS package programming language. As a result, based on posterior distribution statistics, Beta-Blockers lower the mortality of a heart attack.*

**Keywords:** BUGS, MCMC method, Meta analysis, Beta-blocker.