ISSN:1306-3111
e-Journal of New World Sciences Academy
2008, Volume: 3, Number: 2
Article Number: A0073

Yüksel Terzi
Naci Murat
Mehmet Ali Cengiz
University of Afyon Kocatepe
yukselt@aku.edu.tr
Afyonkarahisar-Turkiye

NATURAL AND APPLIED SCIENCES
STATISTICS
Received: December 2007
Accepted: March 2008

BAYESCİ HİPOTEZ TESTLERİ VE BAYES FAKTÖRÜ

ÖZET

Bayes faktörü Bayesci hipotez testlerinin en önemli noktasıdır. Klasik p değerine karşın, hipotezin doğru olup olmadığını test etmede Bayes faktörü direk yoruma sahiptir. Sonsal olasılıklarının oranını elde etmek için, eldeki verinin iki hipoteze ait önsel olasılıkların oranlarıyla güncelleştirilmesidir. Ancak Bayes faktörü yokluk ve alternatif hipotezlere ilişkin model parametrelerine ait önsel yoğunluklara bağlıdır. Ayrıca hesaplaması yüksek dereceden integral hesabı gerektirir. Bu nedenlerden ötürü, Bayes faktörü Klasik hipotez testlerinden daha az sıklıkla kullanılır. Bu çalışma Bayesci hipotez testlerinin kısa bir özetini sunar.

Anahtar Kelimeler: Bayesci Yaklaşım, Hipotez Testi, Bayes Faktör

BAYESIAN HYPOTHESIS TEST AND BAYES FACTOR

ABSTRACT

Bayes factors are the cornerstone of Bayesian hypothesis testing. In contrast to classical p values, the value of a Bayes factor has a direct interpretation in terms of whether or not a hypothesis is true: It represents the factor by which data modify the prior odds of two hypotheses to give the posterior odds. Unfortunately, the values of Bayes factors often depend on the prior densities assigned to the model parameters inherent to null and alternative hypotheses. In addition, the calculation of Bayes factors usually involves the evaluation of high dimensional integrals. For these reasons, Bayes factors are employed less frequently than Classic hypotheses are. This paper provides a brief review of Bayesian hypothesis testing.

Keywords: Bayesian Approach, Hypothesis Test, Bayes Factor

1. GİRİŞ (INTRODUCTION)

İstatistiksel hipotezi test etme problemi Bayesci çalışanlar ile klasik istatistik çalışanları arasında önemli bir tartışma noktası olmuştur. Klasik istatistik yaklaşımı bir ret bölgesi oluşturur ve ilişkili olasılıkları hesaplar. H_0 boş hipotezinin yanlış olarak ret edilmesi α olasılığına sahiptir bu da I. tip hatadır. H_0 boş hipotezinin yanlış olarak kabul edilmesi β olasılığına sahiptir bu da II. tip hatadır. Bu tür α ve β'nın kullanımı bir çok araştırmacı tarafından eleştirilmiştir. Eleştirinin temel nedeni bu hata olasılıklarının bilinen veri tarafından sağlanan veriyi yansıtmamasıdır. Bu durumda H_0hipotezine karşı delilin gücünü gösteren veriye bağlı bir ölçüm olan p-değerinin kullanımı ortaya atılmıştır. Ancak p değeri gerçek bir ölçüm değeri değildir ve delilin bir ölçüsü olarak bazı noksanlıklara sahiptir (Edwards, W., Lindman, H. ve Savage, L., 1963; Berger, J.O. ve Delampady, M., 1987; Delampady, M. ve Berger, J.O., 1990). p-nin kullanımındaki tartışmaları ve p değeri ile delilin veriye bağlı diğer ölçümleri arasındaki farklılıkları ifade ettiler. Genel olarak yokluk hipotezine karşı veri tarafından sağlanan delilin bir ölçüsü olarak p-nin kullanımının yanıltıcı olabileceği vurgulanır (Berger, J.O., Boukai, B. ve Wang, Y., 1997).

Son 30 yıldaki bilgisayar teknolojisindeki inanılmaz gelişimin aslı (Bayes, T.R., 1963) de ortaya konan Bayes teoremine dayanan Bayesci yaklaşımını klasik istatistik yaklaşımına alternatif bir çıkarım tekniği olarak sunmuştur. Günümüzde her türlü istatistik tahmin probleminde alternatif bir yaklaşım olarak kullanılan ve dünya literatüründe yer alan Bayesci yaklaşım hipotez testlerinde de çok yaygın olmasa bile kullanılmaktadır. Maalesef dünyada çok yaygın kullanıma karşın bu yaklaşımla ilgili çalışmalar Türkçe literatürde fazla yer almamaktadır. Bu çalışmanın amaçlarından biri Bayesian hipotez testlerini uygulamacılara sunmaktır.

Bayesci metotlar hesaplamalarda çoğu zaman nümerik yöntemlerle çözülebilen integraller ve sonsal olasılıkları güncelleştiren önsel olasılıkların hesaplanmasını gerektirir. Her iki nokta için pek çok çalışma yapılmış ve uygun yöntemler sunulmuştur (O'Hagan, A., 1994; Bernardo, J.M. ve Smith, A.F.M., 1994). Bayesci yaklaşımın hipotez testlerinde kullanım ilk olarak (Jeffreys, H., 1961) tarafından ortaya konulmuştur. Jeffreys sunduğu metodu istatistiksel testler olarak isimlendirse de kullanılan yaklaşım Bayesci idi. Bayesci hipotez testlerinde en önemli unsurlardan bir tanesi Bayes Faktörüdür. Klasik p değerinin yorumlanmasındaki eleştirilere karşın Bayes Faktörünün hesaplanması zor ancak hipotezin doğru ya da yanlış olduğuna ilişkin yorumlanması kolay ve kesindir. Yokluk hipotezi için p değeri ile Bayes Faktörünü (BF) karşılaştırılmıştır (Berger, J.O. ve Sellke, T., 1987). Aynı karşılaştırmayı tek taraflı hipotez testi için (Casella , G. ve Berger, R.L., 1987) tarafından yapılmıştır. Berger, J.O., Delampady, M. (1987) ile Berger, J.O. ve Berry, D.A.(1988) BF ile alakalı karşılaştırmalı çalışmalar sunmuştur. Kass, R.E. ved Raftery, A.E. (1995) BF'nin ve alternatiflerinin kullanımını örneklerle sunmuştur. Ayrıca Sellke, T. , Bayarri, M.J. ve Berger, J.O. (2001) hipotez için p değerinin kalibrasyonuyla ilgili çalışmayı yapmıştır.

Bu çalışmada geniş bir literatür çalışmasından sonra Bayesci hipotez testleri ve Bayes Faktörü kısaca özetlenmiştir. Uygulamada farklı basit örnekler üzerinde durulmuştur. Ayrıca hesaplamamalar için

Visual Basic 6.0 ortamında yazılmış MACNAMstat programı hazırlanmış ve program çıktıları Ek sunulmuştur.

2. ÇALIŞMANIN ÖNEMİ (RESEARCH SIGNIFICANCE)

Bayes faktörü Bayesci hipotez testlerinin en önemli noktasıdır. Klasik p değerine karşın, hipotezin doğru olup olmadığını test etmede Bayes faktörü direk yoruma sahiptir. Sonsal olasılıklarının oranını elde etmek için, eldeki verinin iki hipoteze ait önsel olasılıkların oranlarıyla güncelleştirilmesidir. Ancak Bayes faktörü yokluk ve alternatif hipotezlere ilişkin model parametrelerine ait önsel yoğunluklara bağlıdır. Ayrıca hesaplaması yüksek dereceden integral hesabı gerektirir. Bu nedenlerden ötürü, Bayes faktörü Klasik hipotez testlerinden daha az sıklıkla kullanılır.

3. YÖNTEM (METHOD)

Bayes hipotez testinde, temel olarak bilinen eldeki veriler kullanılarak iki hipotezin olasılıklarının hesaplanmasıyla H_0 ile H_1 arasında karar verilir.

$$\begin{cases} \dfrac{P(H_0/D)}{P(H_1/D)} > 1 & \Rightarrow H_0 \ seçilir \\ aksidurumda & \Rightarrow H_1 \ seçilir \end{cases} \quad (1)$$

Kararlar $P(Hipotez/Veri)$ ile kullanılarak verilir, oysa klasik istatistikte aynı işlem $P(Veri/Hipotez)$ olasılığı kullanılarak verilir. Bayes kuralına göre D ile gösterilen veri bilindiğinde H_0'ın sonsal olasılığı;

$$P(H_0/D) = \frac{P(D/H_0)P(H_0)}{P(D/H_0)P(H_0) + P(D/H_1)P(H_1)} \quad (2)$$

Benzer biçimde; D bilindiğinde H_1' in sonsal olasılığı;

$$P(H_1/D) = \frac{P(D/H_1)P(H_1)}{P(D/H_0)P(H_0) + P(D/H_1)P(H_1)} \quad (3)$$

dır. Burada $P(H_0)$, H_0'ın ve $P(H_1)$, H_1 önsel olasılılıklarıdır. Her iki sonsal olasılığı birbirine oranlarsak,

$$\left[\frac{P(H_0/D)}{P(H_1/D)}\right] = \left[\frac{P(H_0)}{P(H_1)}\right]\left[\frac{P(D/H_0)}{P(D/H_1)}\right] \quad (4)$$

elde ederiz. Burada $BF = \left[\dfrac{P(D/H_0)}{P(D/H_1)}\right]$ Bayes faktörü olarak isimlendirilir. (4) deki ifadeyi sözel olarak aşağıdaki gibi ifade edebiliriz:

$$[Sonsal\ odds\ oranı] = [Önsel\ odds\ oranı] \times [Bayesfaktörü]$$

Thumb kuralı kullanılarak elde edilen Tablo 1'e göre Bayes Faktörü değeri kullanılabilir.

Tablo 1. Thumb kuralına göre Bayes Faktör değerinin yorumlanması
(Table 1. Interpretation of Bayes factor values according to Thumb rule)

Bayes Faktör değeri	Yorum
$BF \geq 1$	H_0 kabul edilir
$1 > BF \geq 10^{-1/2}$	H_0 'a karşı az delil var
$10^{-1/2} > BF \geq 10^{-1}$	H_0'a karşı sağlam delil var
$10^{-1} > BF \geq 10^{-2}$	H_0'a karşı güçlü delil var
$10^{-2} > BF$	H_0'a karşı kesin delil var

4. BULGULAR (FINDINGS)

Bu bölümde dört basit uygulama ile hesaplamaları ve yorumlamalarını sunacağız.

Uygulama 1:

Farklı ortalamalı ancak eşit varyanslı $(\sigma^2 = 1)$ normal dağılıma sahip iki ortalamayı Bayesci yöntemle karşılaştırmak istiyoruz. Hipotezlerimiz;

$$H_0 : \mu = 0$$
$$H_1 : \mu = 1$$

şeklindedir. (4) eşitliğini,

$$\frac{P\left(H_0 / \overline{X}\, gözlenen\right)}{P\left(H_1 / \overline{X}\, gözlenen\right)} = \left(\frac{P(H_0)}{P(H_1)}\right)\left[\frac{P\left(\overline{X}\, gözlenen / H_0\right)}{P\left(\overline{X}\, gözlenen / H_1\right)}\right] \quad (5)$$

şeklinde yazalım. N örneklem genişliği için (5) nolu ifade

$$= \left(\frac{P(H_0)}{P(H_1)}\right)\left[\frac{\frac{\sigma}{\sqrt{2.\pi N}} e^{-\frac{N}{2.\sigma^2}\left(\overline{X}_{gözlenen} - \mu_{H_0}\right)}}{\frac{\sigma}{\sqrt{2.\pi N}} e^{-\frac{N}{2.\sigma^2}\left(\overline{X}_{gözlenen} - \mu_{H_1}\right)}}\right] \quad (6)$$

olur. $P(H_0)$ = $P(H_1)$=0,5 alınırsa (6) nolu eşitlik

$$= e^{-\frac{N}{2.\sigma^2}\left(\overline{X}_{gözlenen} - \frac{\mu_{H_0} + \mu_{H_1}}{2}\right)} \quad (7)$$

olur. Örneğin N=5 ve $\overline{X}_{gözlenen} = 1$ için; $P\left(H_0 / \overline{X}_{gözlenen}\right)$=0.076 ve $P\left(H_1 / \overline{X}_{gözlenen}\right)$=0.924 ve $\frac{P\left(H_0 / \overline{X}_{gözlenen}\right)}{P\left(H_1 / \overline{X}_{gözlenen}\right)} = e^{-5(1-0,5)} = 0.082$ bulunur. Sonuç olarak H_1 hipotezi kabul edilir. Benzer şekilde $BF =$0,082 olduğundan Thumb kuralına göre $10^{-1} > BF = 0.082 \geq 10^{-2}$ olduğundan H_0 'a karşı güçlü delil vardır denir. Farklı N ve ortalama değerleri için olası sonuçlar Tablo 2, Tablo3 ve Tablo 4 de sunulmuştur.

Tablo 2. N=5 için örneklem genişliği sonuçları
(Table 2. Sample size results for N=5)

Örnek ortalaması	$P\left(H_0 / \overline{X}_{gözlenen}\right)$	$P\left(H_1 / \overline{X}_{gözlenen}\right)$	Karar
0	0.924	0.076	H_0
0.25	0.777	0.223	H_0
0.50	0.500	0.500	kararsız
0.75	0.223	0.777	H_1
1.00	0.076	0.924	H_1

Tablo 3. N=10 için örneklem genişliği sonuçları
(Table 3. Sample size results for N=10)

Örnek ortalaması	$P\left(H_0 / \overline{X}_{gözlenen}\right)$	$P\left(H_1 / \overline{X}_{gözlenen}\right)$	Karar
0	0.993	0.007	H_0
0.25	0.924	0.076	H_0
0.50	0.500	0.500	kararsız
0.75	0.076	0.924	H_1
1.00	0.007	0.993	H_1

Tablo 4. N=20 için örneklem genişliği sonuçları
(Table 4. Sample size results for N=20)

Örnek ortalaması	$P\left(H_0 / \overline{X}_{gözlenen}\right)$	$P\left(H_1 / \overline{X}_{gözlenen}\right)$	Karar
0	0.999	0.001	H_0
0.25	0.993	0.007	H_0
0.50	0.500	0.500	kararsız
0.75	0.007	0.993	H_1
1.00	0.001	0.999	H_1

Uygulama 2:

Sabit varyanslı (σ^2 =1) normal dağılımlı tek bir kitle için ortalamanın sıfırdan farklı olup olmadığını test etmek istiyoruz. Çift yönlü hipotezlerimiz

$$H_0 : \mu = 0$$
$$H_1 : \mu \neq 0$$

dir. $P(H_0)$ = $P(H_1)$ alınırsa N örneklem genişliği için,

$$\frac{P\left(H_0 / \overline{X}_{gözlenen}\right)}{P\left(H_1 / \overline{X}_{gözlenen}\right)} = \frac{P\left(\overline{X}_{gözlenen} / \mu = 0\right)}{\int_{\mu \neq 0} P\left(\overline{X}_{gözlenen} / \mu\right) P(\mu) d\mu}$$

elde edilir. Örnek olarak N=5 ve $\overline{X}_{gözlenen}$ =1 için $P\left(H_0 / \overline{X}_{gözlenen}\right)$=0,0147 ve $P\left(H_1 / \overline{X}_{gözlenen}\right)$ = 0,9853 bulunur. Sonuç olarak H_1 kabul edilir. Benzer şekilde $BF = \frac{P\left(H_0 / \overline{X}_{gözlenen}\right)}{P\left(H_1 / \overline{X}_{gözlenen}\right)}$=0,0149 hesaplaması ve Thumb kuralı gereği $10^{-1} > BF = 0.0149 \geq 10^{-2}$ olduğunda H_0 'a karşı güçlü delil vardır denir. Farklı N ve ortalama değerleri için olası sonuçlar Tablo 5, Tablo 6 ve Tablo 7 de sunulmuştur.

Tablo 5. N=5 için örneklem genişliği sonuçları
(Table 5. Sample size results for N=5)

Örnek ortalaması	$P(H_0 / \overline{X}_{gözlenen})$	$P(H_1 / \overline{X}_{gözlenen})$	Karar
0	0.1785	0.8215	H_1
0.25	0.0956	0.9044	H_1
0.50	0.0511	0.9489	H_1
0.75	0.0274	0.9726	H_1
1.00	0.0147	0.9853	H_1

Tablo 6. N=10 için örneklem genişliği sonuçları
(Table 6. Sample size results for N=10)

Örnek ortalaması	$P(H_0 / \overline{X}_{gözlenen})$	$P(H_1 / \overline{X}_{gözlenen})$	Karar
0	0.1262	0.8738	H_1
0.25	0.0362	0.9638	H_1
0.50	0.0104	0.9896	H_1
0.75	0.003	0.997	H_1
1.00	0.0008	0.9992	H_1

Tablo 7. N=20 için örneklem genişliği sonuçları
(Table 7. Sample size results for N=20)

Örnek ortalaması	$P(H_0 / \overline{X}_{gözlenen})$	$P(H_1 / \overline{X}_{gözlenen})$	Karar
0	0.0892	0.9108	H_1
0.25	0.0073	0.9927	H_1
0.50	0.0006	0.9994	H_1
0.75	0.0001	0.9999	H_1
1.00	0.0000	1.0000	H_1

Uygulama 3:

Kabul edelim ki hayat sigortası sahiplerinin yaşları X-tesadüfi değişkeni ile gösterilsin ve X / μ ~ N(μ , 51.84) normal dağılımlı, n=36 ve $\overline{X} = 39,22$ olduğu bilinsin. Bu durumda; $H_0 : \mu \leq 37$, $H_1 : \mu > 37$ hipotezlerini test etmek istiyoruz..

$$\frac{P(H_0 / X)}{P(H_1 / X)} = \left(\frac{P(H_0)}{P(H_1)}\right)\left[\frac{P(X / H_0)}{P(X / H_1)}\right]$$ ve $P(H_0) = P(H_1) = 1$

düzgün önsel alınırsa

$$P(H_0 / X) = P(\mu \leq 37 / \overline{X}) = Z\left(\frac{37 - \overline{X}}{\sigma / \sqrt{n}}\right) \cong Z(-1,852) \cong 0,032$$ ve

$P(H_1 / X) = 1 - P(H_0 / X) = 1 - 0{,}032 \cong 0{,}968$ bulunur. Sonuç olarak H_1 kabul edilir.

Uygulama 4:

Kabul edelim ki $Y_i = X_i + 1$ olsun ve

$$H_0 : X_i / \mu \sim Po(\mu)$$
$$H_1 : Y_i / \pi \sim Ge(\pi)$$

hipotezlerini test etmek istiyoruz. Bunun için μ / H_0 ve π / H_1 için referans önselleri kabul ederek Bayes faktörünü hesaplayalım. $BF = \left[\frac{P(D/H_0)}{P(D/H_1)}\right]$ ifadesini ve μ parametreli Poisson dağılımının π parametreli Geometrik dağılımın olasılık fonksiyonlarını kullanarak,

$$BF = \frac{\int_0^{\infty}\left[\frac{\mu^{\sum X}}{\pi X!} exp(-n,\mu)\right]\left[\mu^{-\frac{1}{2}}\right] d\mu}{\int_0^{1}\left[\pi^n (1-\pi)^{\sum X}\right]\left[\pi^{-1}(1-\pi)^{-\frac{1}{2}}\right] d\pi}$$

$$= \frac{\frac{1}{\pi X!}\int_0^{\infty} \mu^{\sum X - \frac{1}{2}} \exp(-n\mu)\, d\mu}{\int_0^{1} \pi^{n-1}(1-\pi)^{\sum X - \frac{1}{2}} d\pi}$$

$$= \frac{\Gamma(n,\sum X + \frac{1}{2})}{\pi X!\, n^{\sum X + \frac{1}{2}} B(n,\sum X + \frac{1}{2})} = \frac{\Gamma(n + \sum X + \frac{1}{2})}{\pi X!\, n^{\sum X + \frac{1}{2}} (n-2)}$$

elde ederiz. Sayısal örnek olarak, n=2 , $X_1 = X_2 = 0$ alırsak

$$BF = \frac{\Gamma(21/2)}{1 \times \sqrt{2} \times 1} \cong 0{,}94$$

olur. Thumb kuralına $1 > BF = 0.94 \geq 10^{-1/2}$ olduğundan H_0 'a karşı az delil vardır denir. Bir başka sayısal örnek olarak; n=2 , $X_1 = X_2 = 2$ alırsak

$$BF = \frac{\Gamma(61/2)}{4 \times 2^{5/2} \times 1} \cong 3{,}181$$

bulunur. $BF = 3.181 \geq 1$ olduğundan H_0 kabul edilir.

5. SONUÇ VE ÖNERİLER (CONCLUSION AND SUGGESTIONS)

Bayesci yaklaşım incelenen modelin parametre değerleri hakkındaki inanışların yani önsel olasılıkların verideki bilgiye nasıl ekleneceği sorusuna tam olarak cevap verir. Klasik çıkarım yaklaşımın da mümkün olmayan önsel bilginin kullanımına imkan tanır.Herhangi bir konu dışı fakat anlamlı bir bilgi, örneğin geçmişteki benzer sonuçlar veya başka bir benzer araştırma çalışmalarının sonuçları önsel formülasyon içinde Bayesci metodunda birleştirilebilir. Karşıt olarak Klasik istatistikte böyle bilginin ihmal edilmeye daha uygundur ve meta analizi gibi ileri tekniklerle bu birleştirmeler yapılır. Avantaj gibi gözüken bu durum önsellerin seçimi ve formülüze edilmesi konusunda bir takım zorluklar ve subjektiflik de getirir.

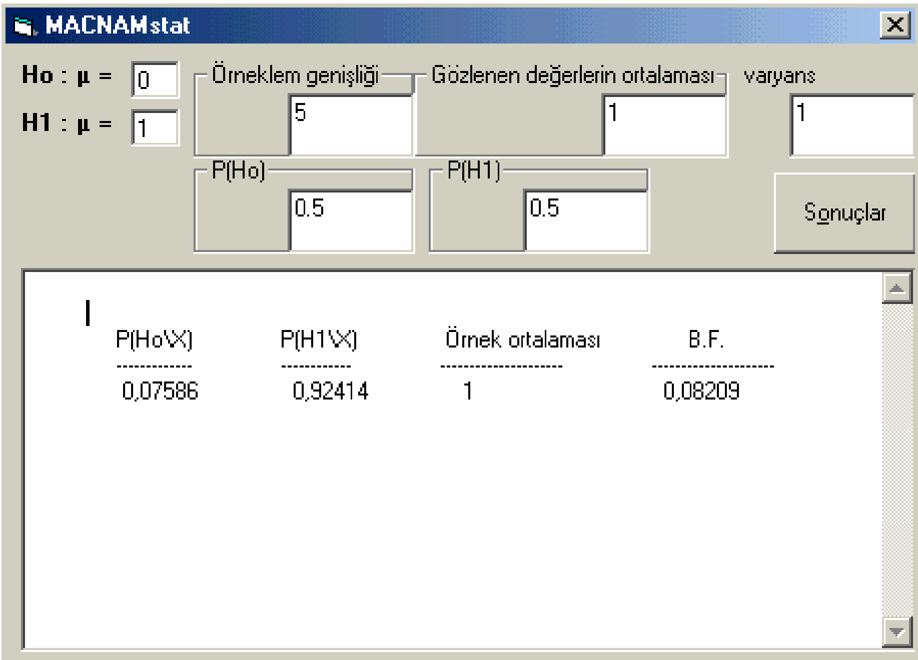
Hipotez testleri gibi analizlerde direkt yorumlama avantajları sağlamasına rağmen sonsal olasılıkların hesaplamasında gereken integrallerin bulunmasında analitik çözümler yeterli olmamaktadır. Dolayısıyla nümerik yöntemlere ihtiyaç duyulmaktadır. Özellikle bilinen istatistik paket programlarında bu işlemler kolaylıkla yapılamamaktadır. Ancak son yıllarda geliştirilen WinBugs gibi programlarda ve SPLUS makro yazılımıyla bu hesaplamalar yapılabilmektedir.

KAYNAKLAR (REFERENCES)

1. Edwards, W., Lindman, H., and Savage, L., (1963). Bayesian Statistical Inference for Psychological Research. Psychological Review: Vol:70, pp:193-242.
2. Berger, J.O., Delampady, M.,(1987).Testing Precise Hypotheses. Statist. Sci. J.: pp:317-352.
3. Delampady, M. and Berger, J.O., (1990). Lower Bounds on Bayes Factors for the Multinomial Distribution with Application to Chi-Squared Tests of Fit. Ann. Statist.: Vol:18, pp:1295-1316.
4. Berger, J.O., Boukai, B., and Wang, Y., (1997). Unified Frequentist and Bayesian Testing of a Precise Hypothesis. Statistical Science: Vol:12 (3), pp:133-160,
5. Bayes, T.R., (1963). An Essay Towards Solving a Problem in the Doctrine of Chances. London: Phil Trans.Roy.Soc.: pp:53.
6. O'Hagan, A, (1994). Bayesian Inference. New York: Edward Arnold.
7. Bernardo, J.M. and Smith, A.F.M., (1994). Bayesian Theory. New York: Wiley.
8. Jeffreys, H., (1961). Theory of Probability (3rd ed.). Oxford, U.K.: Oxford University Pres.
9. Berger, J.O. and Sellke, T., (1987). Testing a Point Null Hypothesis. The irreconcilability of P values and evidence (with discussion). Journal of the American Statistical Association: Vol:82(397), pp:112-139.
10. Casella, G. and Berger, R.L., (1987). Reconciling Bayesian and Frequentist Evidence in the One-Sided Testing Problem (with discussion). Journal of the American Statistical Association: Vol:82(397), pp:106-111,123-139 (discussion).
11. Berger, J.O. and Berry, D.A., (1988). The Relevance of Stopping Rules in Statistical Inference in Statistical Decision Theory and Related Topics IV, Vol:1, eds. S. S. Gupta and J.O. Berger, New York: Springer verlang, pp:29-47.
12. Kass, R.E. and Raftery, A.E., (1995). Bayes Factors. Journal of the American Statistical Association: Vol:90, pp:773-795.
13. Sellke, T., Bayarri, M.J., and Berger, J.O., (2001). Calibration of P Values for Testing Precise Null Hypotheses. The American Statistician: Vol:55, pp:62-71.

EK (APPENDIX)

1. Uygulama 1'de yapılan hesaplamalar için hazırlanan Visual Basic 6.0 programının çıktısı aşağıdaki gibidir.



2. Uygulama 2'de yapılan hesaplamalar için hazırlanan Visual Basic programının çıktısı aşağıdaki gibidir.