ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BAYESCİ HİPOTEZ TESTLERİ VE UYGULAMALARI ÜZERİNE BİR ÇALIŞMA

Tuğba KARACAOĞLU UZUN

İSTATİSTİK ANABİLİM DALI

ANKARA
2020

Her hakkı saklıdır

TEZ ONAYI

Tuğba KARACAOĞLU UZUN tarafından hazırlanan "Bayesci Hipotez Testleri ve Uygulamaları Üzerine Bir Çalışma" adlı tez çalışması 31/01/2020 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Doç. Dr. Esin KÖKSAL BABACAN
Ankara Üniversitesi / İstatistik Anabilim Dalı

Jüri Üyeleri:

Başkan: Doç. Dr. Cenker BİÇER
Kırıkkale Üniversite / İstatistik Anabilim Dalı

Üye : Prof. Dr. Yılmaz AKDİ
Ankara Üniversitesi / İstatistik Anabilim Dalı

Üye : Doç. Dr. Esin KÖKSAL BABACAN
Ankara Üniversitesi / İstatistik Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. Özlem YILDIRIM
Enstitü Müdürü

ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

31.01.2020



Tuğba KARACAOĞLU UZUN

ÖZET

Yüksek Lisans Tezi

BAYESCİ HİPOTEZ TESTLERİ VE UYGULAMALARI ÜZERİNE BİR ÇALIŞMA

Tuğba KARACAOĞLU UZUN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Doç. Dr. Esin KÖKSAL BABACAN

Kitle parametresi ile ilgili öne sürülen iddiaya hipotez denir. Klasik yaklaşım ile hipotez testi yapılırken örneklem istatistikleri ve onların dağılımlarından yararlanılır. Bayesci yaklaşımda ise parametreler birer rasgele değişken olarak düşünülür. Bu nedenle kendilerine ait dağılımları vardır. Parametre ile ilgili istatistiki sonuç çıkarımı yapılmak istendiğinde, parametreye ilişkin dağılım bilgisi ve veriye ait örneklem bilgisi kullanılarak sonsal dağılım elde edilir. Sonsal dağılım elde edildikten sonra tüm çıkarım bu dağılım kullanılarak yapılır. Sonsal dağılımın elde edilmesi bazı durumlarda kolay değildir. Bu tür durumlarda, sonsal dağılımdan örneklem çekme yöntemiyle çalışan algoritmalar kullanılarak sonuç çıkarımına gidilir.

Bu çalışmanın amacı, son yıllarda kullanımı artan Bayesci sonuç çıkarımının istatistiğin önemli bir konusu olan hipotez testlerine uygulanışını araştırmaktır. Bu amaçla, ilk olarak Bayesci çıkarımın adımları anlatılmış daha sonra kitle oranı, kitle ortalaması, iki kitle oranı farkı ve iki kitle ortalaması farkı için Bayesci hipotez testinin nasıl yapılacağı ve güvenilir aralığın nasıl oluşturulacağı konularına değinilmiştir. Konu anlatımları yapılan örneklerle pekiştirilmiş ve elde edilen sonuçların karşılaştırılması için çeşitli programlardan (Matlab, R, SPSS) yararlanılmıştır.

Son olarak, gerçek EEG verileri kullanılarak ortaya atılan iddialar hem klasik hem de Bayesci hipotez testi kullanılarak test edilmiştir.

Ocak 2020, 112 sayfa

Anahtar Kelimeler: Bayesci yaklaşım, Bayes faktörü, Bayesci hipotez testi, Önsel dağılım, Sonsal dağılım.

ABSTRACT

Master Thesis

A STUDY ON BAYESIAN HYPOTHESIS TESTING AND APPLICATIONS

Tuğba KARACAOĞLU UZUN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Assoc. Prof. Dr. Esin KOKSAL BABACAN

Hypothesis is a statement about the population parameter. When we test hypothesis in classical approach, sampling statistics and their distributions are used. In Bayesian approach, parameters are considered as random variables. So, they have their own distributions. This is called prior distribution. When we want to make Bayesian statistical inference about the population parameter, first we obtain the posterior distribution of the parameter by using the prior knowledge and the knowledge coming from the observation. Obtaining the posterior distribution, all statistical inferences are made by using this distribution. However, in some cases it is not easy to compute this posterior distribution. In such cases, inferences are obtained by using algorithms which are working to draw sample from the posterior distribution.

The aim of this study is to investigate the application of Bayesian inference to hypothesis testing which is an important subject of statistics. For this aim, first, the steps of Bayesian inference are explained and then, Bayesian hypothesis testing for population proportion, population mean, two population proportion difference and two population mean difference are investigated and credible intervals are obtained. Some examples are given to emphasize the results and various programs (Matlab, R, SPSS) are used to compare the results.

Finally, the hypotheses about the real EEG observations are tested by using with both classical and Bayesian methods.

January 2020, 112 Pages

Key Words: Bayesian approach, Bayesian factor, Bayesian hypothesis testing, , Prior distributions, Posterior distribution,

ÖNSÖZ VE TEŞEKKÜRLER

Lisansüstü eğitiminin tez dönemimin her aşamasında bilgi ve deneyimleriyle bana yol göstererek destekleyen, moral olan, bana ışık olan, ilgi ve özverisiyle gelişimime büyük katkı sağlayan çok kıymetli hocam ve danışmanım Sayın Doç. Dr. Esin Köksal Babacan'a teşekkür ve saygılarımı sunarım.

Evlilik hayatım boyunca bana olan inancını hiç kaybetmeyen, mutluluğum için fedakarlığı esirgemeyen, eğitimim için en az benim kadar caba sarfeden sevgili eşim Gökan Uzun'a ve bana anneliği tattıran, hayatımı anlamlandıran biricik oğlum Göktuğ Uzun'a teşekkür ederim.

Tuğba KARACAOĞLU UZUN
Ankara, Ocak 2020

İÇİNDEKİLER

TEZ ONAY SAYFASI
ETİK ..... **i**
ÖZET ..... **ii**
ABSTRACT ..... **iii**
ÖNSÖZ VE TEŞEKKÜR ..... **iv**
SİMGELER VE KISALTMALAR DİZİNİ ..... **vii**
ŞEKİLLER DİZİNİ ..... **viii**
ÇİZELGELER DİZİNİ ..... **ix**
1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR **1**
1.1 Bayes Teoremi **5**
2. BAYESCİ YAKLAŞIM **9**
2.2 Önsel Dağılımın Seçimi **19**
2.2.1 Önsel dağılım türleri **20**
3. BAYESCİ HİPOTEZ TESTLERİ **23**
4. ORAN İÇİN BAYESCİ SONUÇ ÇIKARIMI **27**
4.1 Oran İçin Düzgün Önsel Kullanılarak Sonuç Çıkarımı **28**
4.2 Oran için Beta Önseli Kullanılarak Sonuç Çıkarımı **29**
4.3 Oran için klasik ve Bayesci sonuç çıkarımına göre elde edilen tahminlerin karşılaştırması **30**
4.4 Oran İçin Bayesci Güvenilir Aralık **33**
4.5 Oran için hipotez testi **37**
4.5.1 Tek yönlü Bayesci hipotez testi **38**
4.5.2 İki yönlü Bayesci hipotez testi **40**
5. NORMAL DAĞILIMIN ORTALAMASI İÇİN BAYESCİ SONUÇ ÇIKARIMI **44**
5.1 Sürekli Önsel Kullanarak Normal Dağılım Ortalaması İçin Sonuç Çıkarımı **44**
5.1.1 μ için Düz önsel alındığında, ... **45**
5.1.2 μ için Normal önsel alındığında, ... **46**
5.1.3 Normal ailelerde kullanılan temel güncelleme kuralı **48**
5.2 μ İçin Bayesci Güvenilir Aralık **53**
5.2.1 μ için Tek Yanlı Bayes Hipotez Testi **55**
5.2.2 μ için iki yanlı Bayes hipotez testi **58**
6. İKİ KİTLE ORTALAMASI FARKI İÇİN BAYESCİ SONUÇ ÇIKARIMI **59**
6.1 Bağımsız Örneklemler İçin Sonuç Çıkarımı **59**
6.1.1 Varyansların eşit olduğu durum **60**
6.1.1.1 Varyansların biliniyor **60**
6.1.1.2 Tek yanlı Bayes hipotez testi **64**
6.1.1.3 İki yanlı Bayes hipotez testi **66**
6.1.1.4 Varyanslar bilinmiyor **66**
6.1.2 Varyansların eşit olmadığı durum **68**

6.1.2.1 Varyansların biliniyor 68
6.1.2.2 Varyanslar bilinmiyor 69
6.2 Normal Yaklaşım Kullanılarak İki Oran Arasındaki Fark için Bayesci Sonuç Çıkarımı 72
6.2.1 Oranlar arasındaki farkın tek yanlı Bayes hipotez testi 74
6.2.2 Oranlar arasındaki fark için iki yanlı Bayes hipotez testi 74
6.3 Bağımlı Normal İki kitle Ortalaması Arasındaki Fark İçin Sonuç Çıkarımı ... 78
7. GERÇEK VERİ İLE UYGULAMA 86
8. SONUÇ VE TARTIŞMA 93
KAYNAKLAR 96
EKLER 98
EK 1 Örnek 4.4 için sonuçlar 99
EK 2 Örnek 6.1 için frekansçı sonuçlar 100
EK 3 Örnek 6.1 için Bayesci yaklaşım sonuçları 101
EK 4 Örnek 6.2 için sonuçlar 102
EK 5 Örnek 6.3 için sonuçlar 103
EK 6 Örnek 6.4 için sonuçlar 104
EK 7 Uygulama 1 Normal dağılıma uygunluk sonuçları 105
EK 8 Uygulama 1 için SPSS frekansçı sonuçlar 106
EK 9 Uygulama 1 için R programında frekançı sonuçlar 107
EK 10 Uygulama 1 için R programı ile Bayesci sonuç çıkarımı 108
EK 11 Uygulama 2 Normal dağılıma uygunluk 109
EK 12 Uygulama 2 için SPSS ile frekançı sonuç çıkarımı 110
EK 13 Uygulama 2 için R ile Bayesci güvenilir aralık sonuçları 111
ÖZGEÇMİŞ 112

SİMGELER DİZİNİ

σ^2	Varyans
μ	Beklenen değer
∞	Sonsuz
$\propto$	Orantılıdır
$\in$	Elemanıdır
β	Beta
α	Alfa
θ	Teta
π	Oran
σ	Sigma
Γ	Gama
Θ	Teta Parametre Kümesi
Φ	Fi
$\sum$	Toplam
$\int$	İntegral
$\prod$	Çarpım

Kısaltmalar

HKO	Hata Kareler Ortalaması
H.P.D.	Son yoğunluğun en yüksek olduğu bölge veya aralık anlamında (Highest Posterior Density Region or Interval) için "En Yüksek Son Yoğunluk Bölgesi veya Aralığı" kullanılmıştır.
MCMC	Markov Chain Monte Carlo
NHST	Null Hypothesis Significance Testing

ŞEKİLLER DİZİNİ

Şekil 1.1 Örnek uzayın bir parçalanması 5
Şekil 2.1 Önsel ve Sonsal Beta Dağılımları 19
Şekil 2.2 Önsel ve Sonsal Beta Dağılımları 19
Şekil 2.3 Önsel Dağılım Türleri 20
Şekil 4.1 Oran tahmini için Bayesci ve Klasik hata kareler ortalamaları karşılaştırılması 33
Şekil 4.2 R programında sonuç grafiği 43
Şekil 5.1 Alınan önsellerin olasılık yoğunluk fonksiyonları 51
Şekil 5.2 Sonsal dağılımların olasılık yoğunluk fonksiyonları 51
Şekil 6.1 MCMC yöntemi kullanılarak elde edilen örneklemlere ilişkin grafikler 73
Şekil 6.2 R programında sonuç grafikleri 77
Şekil 6.3 R programında sonuç grafikleri 83
Şekil 6.4 R programında sonuç grafikleri 85
Şekil 7.1 Gerçek veri ile R programında sonuç grafikleri 89
Şekil 7.2 Gerçek veri ile R programında sonuç grafikleri 92

ÇİZELGELER DİZİNİ

Çizelge 2.1 Önsel için α ve β değerleri ve örneklem bilgisi 16
Çizelge 2.2 α=1, β=3 ve n=5 için sonuçlar 17
Çizelge 2.3 α=1, β=3 ve n=25 için sonuçlar 17
Çizelge 2.4 α=10, β=30 ve n=5 için sonuçlar 18
Çizelge 2.5 α=10 , β=30 ve n=25 için sonuçlar 18
Çizelge 2. 6 Eşlenik önsel dağılımlar 22
Çizelge 3.1 Klasik yaklaşımda karşılaşılan durumlar 23
Çizelge 3.2 Thumbkuralına göre Bayes Faktör değerinin yorumlanması 26
Çizelge 4.1 Bayesci güvenilir aralık sınırları 37
Çizelge 5.1 Bayesci güvenilir aralık sınırları 55
Çizelge 5.2 Hipotez testi sonuçları 57
Çizelge 6.1 100 km birimiyle lastik ömürleri 62
Çizelge 6.2 Hastaların yaş ve hastalık baz alınarak sınıflandırılması 75
Çizelge 6.3 İneklerin yıllık süt verimi 80
Çizelge 6.4 Can'ın önseli için köşe noktaları 81
Çizelge 6.5 Bayesci güven aralığı hesaplama sonuçları 82
Çizelge 6.6 Deneklerin önce ve sonra nabız sayıları 83
Çizelge 6.7 Temporal bölgedeki Teta gücü 87
Çizelge 6.8 Oksipital bölgenin Alfa gücü 90

1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR

"(…) önsel ve sonsal, verilerle ilgili göreceli terimlerdir. Bugünkü sonsal yarının önselidir. "(Lindley, 2000).

Klasik yaklaşıma alternatif olan Bayes yaklaşımının temeli Bayes teoremine dayanır. Bu teoremin isim babası, rahip ve matematikçi olan Thomas Bayes'dir (1702-1761). Yaptığı çalışmalar içinde kendini gösteren asıl araştırması ani ölümü nedeniyle Bayes tarafından yayınlanamamıştır. Bayes'in ölümünün üzerinden geçen kısa bir süre sonra yakın arkadaşı Richard Price (1723-1791) bu çalışmayı bulmuştur. John Canton ise, o dönemin çok saygın bir fizikçisidir ve Kraliyet ailesiyle bağları kuvvetlidir. Bu nedenlerden dolayı Price, Canton'ın dikkatini Bayes'in bu çalışmasına çekmek için çok çaba harcamıştır. Sonuç olarak Kraliyet Cemiyeti Üyesi adayı olmasının da yardımı ile "An Essay Towards Solving a Problem in the Doctrine of Chances" isimli Thomas Bayes'in denemesini yayınlatmıştır(Anonymous, 2019). Bu sayede, Bayes yaklaşımının temeli olan ***Bayes teoremi*** 1764'de ortaya çıkmıştır. Bu tarihten sonra Bayes teoremi etkin rol oynarken, 19. yy'in sonlarına kadar Laplace'ın, 19 yy sonlarında Edgeworth, A. Wald, Galton ve Pearson'un katkıları ve 20.yy'in başında Neyman ile Pearson'un birlikte yaptıkları araştırmalar ışığında yaklaşımın teorisi daha iyi desteklenerek geliştirilmiştir. Bu arada klasik yaklaşımı savunanların Bayes yaklaşımına olumsuz bakışları sonucu Bayes yaklaşımı geri planda kalmıştır. Yaklaşım, 20. yy'ın başlarında F. P. Ramsey'in "Gerçeklik ve Olasılık (Truth and Probability)" ismini taşıyan denemesi ile yeniden ortaya çıkmıştır. 20. yüzyılın ortalarında H. Jeffreys, "Ters Olasılık (Inverse Probability)" adlı eserinde; Bayesci regresyon analiziyle ilgili ilk çalışmayı yapmış, yaklaşımın mantıksal eksikliklerini gidermiş, Bayesci hipotez testini ve objektif bakış açılı Bayesci analizi geliştirmiştir. Ayrıca; I.J.Good, L.J. Savage, B. de Finetti, D.V. Lindley ve R. Schlaifer gibi bilim adamlarının da klasik teknikte gözlenen eksiklikleri gideren çalışmaları ile Bayes yaklaşımı tekrar gündeme gelmiştir. 20. yüzyılın ortalarından sonra Bayesci hipotez testi ve model seçimi konusunda literatür genişlemiştir (Ekici, 2005).

Fakat, bazı matematiksel kısıtlılıklar, özellikle sonsal dağılımın hesaplanmasında yüksek dereceden integral hesabı gerektirmesi Bayesci yaklaşımın dezavantajlarındandır. Bu konuda, N. Metropolis, Markov Zinciri Monte Carlo (MCMC) tekniğinin temel taşlarını oluşturup geliştirmişken, W.K. Hastings ise, Markov Zinciri Monte Carlo yöntemi ile istatistik alanındaki uygulamaların geliştirilmesine katkıda bulunmuştur. Monte Carlo yöntemi kullanılarak, istenilen bir olasılık dağılımından birbirinden bağımsız, simülasyon ile örneklem değerleri üretilir. Yani, seçilen dağılımdan rasgele çok sayıda gözlem oluşturulur. MCMC yönteminde ise sadece bir önceki örneklem değerine bağlı olarak zincir değerleri üretilir. Zincir kullanılarak elde edilen değerlerden özet istatistikler oluşturulur. Yani, bu örnekler kullanılarak sonsal dağılımın ortalamasına, medyanına, HPD (Highest Posterior Density Region or Interval) Bayesci güven aralığına ulaşılabilir (Öztürk, 2004).

P.H. Peskun ve S. Geman'ın çalışmalarıyla, A.E. Gelfand ve A.F.M. Smith'in katkılarıyla ve bilgisayar teknolojisinin gelişmesi sayesinde, kullanılan Winbugs, Bugs, R vb simülasyon tekniklerini içeren programlama dilleriyle Bayesci yaklaşımdaki bahsedilen dezavantaj da ortadan kaldırılmıştır (Ekici, 2005; Demirci, 2016). WinBUGS yazılımının temeli BUGS'a (Bayesian inference Using Gibbs Sampling) dayanmaktadır. BUGS 1996'da yayınlanmasının ardından, biraz daha geliştirilerek ve grafik arayüzlü hale dönüştürülerek, yeni bir isimle 2003'te WinBUGS olarak kullanıma sunulmuştur. Winbugsta, komut satırları daha kısadır, değişkenlerin model dışında ayrıca bildirimine gerek yoktur, dizilerin boyutu sınırlandırılmamıştır, deyimlerin noktalı virgülle sonlandırılmasına gerek yoktur. Yazılım, MCMC kullanarak karmaşık istatistiksel modellerin Bayesci analizini gerçekleştirmek üzere oluşturulmuştur. Yüksek seviyeli bir programlama dilidir (Şen, 2018).

Ortega ve Navarrete (2017), Bayesci hipotez testinin, mevcut NHST (Null Hypothesis Significance Testing)'ye karşı faydalı bir tamamlayıcıyı hatta bir alternatif olabileceğini belirtmişlerdir. Psikoloji ve sosyal bilimler alanında Bayesci hipotez testini Bayes faktörü ile açıklamayı amaçlamışlardır.

Van der ve Chryst (2017), yaptıkları çalışmada Bayes faktörünün kanıt için tutarsız bir çerçeve sağladığını ifade etmişlerdir. Bayesci yaklaşımın, önsele karşı yüksek hassasiyetinden dolayı, Bayes faktörünün önsel bilgiyi göz önüne almadığı durumlar için "Bayesci çıkarımın direk bir anlamı yoktur" ifadesinin üzerinde durmuşlardır.

Vasilyev ve Dobrovidov (2016), Bayesci hipotez testi ile H_1 alternatif hipotezinin koşullu dağılımı ve önsel dağılımının bilinmediği durumda oluşan problemi ele almışlardır. Birçok uygulamada (örneğin, radar ve sonar sinyal işlemede) kaynaklardan gelen gerçek sinyaller genellikle gürültü olmadan gözlemlenemez. Bu nedenle, belirsizliklerin altında tespit sorunlarının çözülmesine yönelik uyarlanabilir (denetimsiz) sınıflandırma prosedürünü kullanmışlardır.

Olivetti vd. (2012), beyin aktivitesi ile uyaranlar arasında anlamlı bir bağlantı olup olmadığını test etmek istemişlerdir. Bunun için gerçek bir fonksiyonel manyetik rezonans görüntüleme (fMRI) deneyinden dokuz beyin kod çözme araştırmaları üzerine bir uygulama yapmışlardır. Çalışmalarında, klasik hipotez testlerine ve çeşitli yaklaşımlara dayanan beyin kod çözme çalışmalarına Bayesci hipotez testi çerçevesinde daha doğru bir istatistiksel test önermişlerdir.

Gu vd. (2016), Bayesci t-testi için oluşturulan Bayes faktörünün klasik tip I ve tip II hata olasılıklarını araştırmışlar ve ortaya çıkan hata olasılıklarının klasik yöntemle elde edilenlere ortalama olarak eşit ve daha küçük bulmuşlardır.

Tezin birinci bölümünde Bayes teoremi açıklanmış, Bayesci sonuç çıkarımının bel kemiği olan önsel dağılım kavramı verilmiş, önsel dağılımı araştırmacının nasıl seçeceği sorununa değinilerek önsel dağılım türleri tablolanarak açıklanmıştır.

İkinci bölümde Bayesci yaklaşım anlatılmış ve konu, açıklayıcı örneklerle aydınlatılmıştır.

Üçüncü bölümde Bayesci hipotez testlerine giriş yapılmış, Örneklerin çoğunda hipotez testleri bayes faktörü kullanılarak da yorumlanmıştır. Bayesci hipotez testleri için önemli bir kavram olan Bayes faktörünün tanımı verilmiştir.

Dördüncü bölümde oran için hipotez testi yapılmak istendiğinde; önsel olarak beta önseli ya da düzgün önsel şeçildiğinde sonsalın elde edilmesi anlatılmış, oran için Bayesci güvenilir aralık bulunmuştur. Oran için, tek ve iki yönlü hipotez testi ayrı ayrı açıklanıp anlatılmıştır. Ayrıca oran için klasik ve Bayesci sonuç çıkarımına göre elde edilen tahmin ediciler karşılaştırılmış, Bayes yaklaşımının üstünlüğü örneklerle gösterilmiştir.

Beşinci bölümde Normal dağılımın ortalaması için Bayesci yaklaşım göz önüne alınmıştır. Sürekli önsel olarak, düz ve Normal önsel alınması durumları incelenmiş ve sonsalların nasıl elde edildiği gösterilmiştir. Normal ailelerde kullanılan basit güncelleme kuralı açıklanmış ve gerekli denklemler verilmiştir. Ortalama için, tek ve iki yönlü hipotez testi ayrı ayrı açıklanıp anlatılmıştır.

Altıncı bölümde iki kitle ortalaması farkı için Bayesci sonuç çıkarımı açıklanmıştır. Bağımsız örneklemler için varyansların eşit olduğu ve varyansların eşit olmadığı durumlarda hipotez testinin nasıl yapılacağı ile ilgili formüller ve açıklamalar verilmiştir. Normal yaklaşımla iki oran arasındaki farkın güven aralığı, tek ve iki yanlı hipotez testi Bayesci yöntemle anlatılmıştır. Normal dağılıma sahip bağımlı iki örneklem göz önüne alındığında, kitle ortalaması farkı için sonuç çıkarımı eşleştirilmiş deney kullanılarak yapılmıştır.

Yedinci bölümde ise gerçek veriler ile Bayesci hipotez testi yapılmıştır. Çocukların tedavi öncesi ve tedavi sonrası EEG dalgalarının ölçümleri alınarak sonuçlar karşılaştırılmıştır. Yani bağımlı iki normal rasgele örneklem için eşleştirilmiş deneyden Bayesci sonuç çıkarımı yapılmıştır.

Özetle, çalışmada ön bilgi kullanılan Bayesci yaklaşım ile hipotez testinin nasıl sonuçlar verdiği çeşitli uygulama örnekleri ile incelenmiştir.

1.1 Bayes Teoremi

(Ω, U, P) bir olasılık uzayı, A ve B, bu olasılık uzayında tanımlı olaylar ve B olayının gerçeklenme olasılığı 0 dan farklı olmak üzere, B olayı gerçekleştiğinde A olayının gerçekleşmesi olasılığı,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

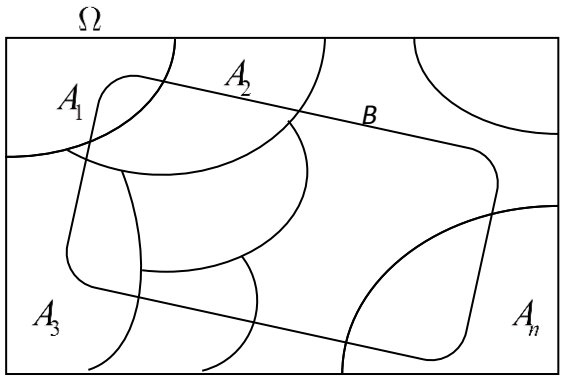
biçiminde ifade edilir ve bu olasılığa A nın B ye göre **koşullu olasılığı** denir (Öztürk, 2006).

Tanım. (Ω, U, P) bir olasılık uzayı ve $A_1, A_2, ..., A_n \in U$ olmak üzere aşağıdaki koşullar sağlanırsa $A_1, A_2, ..., A_n$ olaylarına Ω nın bir **sonlu parçalanması** denir.

a) $\forall\, i \in 1,2, \dots, n$ için A_i lerin olasılığı 0'dan farklıdır.

b) $A_1, A_2, \dots, A_n$' ler ayrık olaylardır.

c) $A_1, A_2, \dots, A_n$'lerin birleşimi Ω ya eşittir.



Şekil 1.1 Örnek uzayın bir parçalanması

Kolayca görüleceği gibi herhangi bir B olayı,

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (B \cap A_i) = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$$

Biçiminde yazılabilir ve B olayının olasılığı da,

$$P(B) = P(\bigcup_{i=1}^{n} (B \cap A_i)) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)\, P(B/A_i)$$

olarak hesaplanır.

Bayes Teoremine göre, (Ω, U, P) bir olasılık uzayı, $A_1, A_2, ..., A_n$ olayları Ω nın bir sonlu parçalanması ve $B \in U$, $P(B) \neq 0$ olmak üzere,

$$P(A_j / B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j)P(B/A_j)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)} \quad , \quad j = 1, 2, ..., n$$

dır (Öztürk, 2006).

Örnek 1.1 Bayesci Yaklaşımla İlgili Gebelik Örneği

Kadınlar hamile olduklarından şüphelendiklerinde, beta hcg ölçümü yapan gebelik testi uygulanır. Ancak bu gebelik testi ile alınan sonuç yüzde yüz doğru değildir. Bu örnekte, sonucun %90 doğru olma ihtimali olduğunu düşünülsün. Yani, hamile olan birinin hcg testinin pozitif olması olasılığı %90olsun.

$$P(\text{hcg} + | \text{gebe}) = 0.90$$

Böyle bir durumda hamile olmayan birinin beta hcg testinde pozitif sonuç elde etmesi de mümkündür. Bu olasılık 0.50 olsun. Buna göre,

$$\mathrm{P}(\mathrm{hcg}+|\text{gebe değil})=0.50$$

dir. Burada kullanılacak bir bilgi de hamilelik şüphesi olan normal sağlıklı bir kadının hamile kalabilme olasılığıdır. Bu olasılığın daha önceki araştırmalardan,

$$\mathrm{P}(\mathrm{B}) \equiv \mathrm{P}(\text{gebe})$$

$\mathrm{P}(\text{gebe}) = 0.15$ olarak bilindiği varsayılsın.

Bu durumda testi yapan bir kadın beta hcg testini pozitif okuduktan sonra gerçekten doğru bir sonuç mu? sorusuyla karşı karşıya kalır.

Bu bilgiler ışığında test pozitif çıktığında bir kadının hamile olma olasılığı,

$$p(B|A) \equiv p(\text{gebe}|\mathrm{hcg}+)$$

ile ifade edilir ve bu olasılık Bayes teoremi kullanılarak bulunabilir. Buna göre,

$$\mathrm{P}(\text{gebe}|\mathrm{hcg}+) = \frac{\mathrm{P}(\,\mathrm{hcg}+|\text{gebe}) \times \mathrm{P}(\text{gebe})}{\mathrm{P}(\,\mathrm{hcg}+|\text{gebe}) \times \mathrm{P}(\text{gebe}) + \mathrm{P}(\mathrm{hcg}+|\text{gebe değil}) \times \mathrm{P}(\text{gebe değil})}$$

olmak üzere

$$\mathrm{P}(\text{gebe}|\mathrm{hcg}+) = \frac{(0.90 \times 0.15)}{(0.90 \times 0.15) + (0.50 \times 0.85)} = \frac{0.135}{0.135 \times 0.425} = 0.241$$

biçiminde hesaplanır. Buradan, test pozitifken kadının hamile olma olasılığının 0.241 yani çok düşük olduğu görülür. Bulunan bu değer sonsal olasılıktır. Çünkü kadının hamile olma olasılığı kadın testi pozitif gözlemlendikten sonra hesaplanmıştır. Bu olasılığın bu kadar küçük çıkmasının sebeplerinden birincisi kadının hamile kalma

olasılığının çok düşük (%15) olmasıdır. İkincisi ise, hcg testinin hamile olmayan birini hamile olarak yanlış sonuçlandırma olasılığının aşırı derecede yüksek bir değer (% 50) dır.

Kadın testteki olasılığın düşüklüğünün farkında ise, ikinci defa test yapmaya karar verebilir. Bu testi yeniden yaptığında, artık $P(gebe) = 0.241$ ve $P(gebe\ değil) = 1 - 0.241 = 0.759$ olarak alınır. Bu sefer test pozitif çıktığında hamile olmasının sonsal olasılığı

$$P(gebe|hcg\ +) = \frac{(0.90 \times 0.241)}{(0.90 \times 0.241) + (0.50 \times 0.759)} = \frac{0.2169}{0.2169 + 0.3795} = 0.364$$

olarak bulunur. Görüldüğü gibi, bu sonuçta hala kadının hamile olduğuna inanması için ikna edici değildir. Bu durumda kadın testi tekrar yaparsa ve hcg testinin sonucu tekrar pozitif çıkarsa kadının hamile olma olasılığı,

$$P(gebe|hcg\ +) = \frac{(0.90 \times 0.364)}{(0.90 \times 0.364) + (0.50 \times 0.636)} = 0.507$$

olur. Bu şekilde devam edildiğinde sonuçlar test4=0.649, test5= 0.769, test6=0.857, test7=0.915, test8=0.951, test9=0.972, test10=0.984 olarak bulunur. Dikkat edilirse burada elde dilen sonsal olasılık değerleri bir sonraki denemede önsel olasılık olarak alınmışlardır.

Bayesci istatistik yöntemi uygulandığında, önsel bilgiler yeni bilgiler ile güncellenerek ***sonsal olasılıklar*** hesaplanır. Bu sonsal olasılıklar sonraki analizlerde ***önsel olasılık*** olarak kullanılır. Bu işlemler, Bayesci sonuç çıkarımının doğal adımlarıdır. Her defasında elde edilen sonsal olasılıklar daha sonraki çalışmalar için önsel bilgiler sağlarlar (Lynch, 2006).

2. BAYESCİ YAKLAŞIM

Bayesci yaklaşımın temeli, 1763 yılında Thomas Bayesin tanımladığı ***Bayes teoremine*** dayanır. Bayes teoremi matematiksel istatistiğin önemli bir teoremidir.

Bayesci yaklaşımda, olasılık ya da olasılık yoğunluk fonksiyonu $p(y|\theta)$ olan istatistiksel model ile ilgili olarak alınan $y = y_1, y_2, \dots, y_n$ gözlemleri kullanılarak θ parametresi tahmin edilmek istenir. Bu yaklaşımda θ parametresi bir rasgele değişken olarak düşünülür. Bu nedenle kendisine ait dağılımı vardır. Bu dağılım ***önsel dağılım (prior)*** olarak bilinir. Önsel dağılım veri bilgisi ile güncellenerek ***sonsal dağılım (posterior)*** elde edilir. Bayesci yaklaşıma göre parametre ile ilgili tüm sonuç çıkarımı sonsal dağılım kullanılarak elde edilir. Bayesci istatistikte bir çözümlemede ulaşılan sonsal dağılım, bir başka çözümlemede önsel dağılım olarak kullanılabilir (Demirhan, 2004).

Aşağıdaki kısımda Bayes yaklaşımını kullanan bir örnek verilmiştir.

Örnek 2.1 Bayesci Yaklaşımla İlgili Binom Oranı Örneği

Ölümcül bir hastalıkla ilgili bu hastalığa yakalanan hastaların yeni uygulanan, zor ve ciddi bir tedaviye nasıl cevap verecekleri araştırılmak istenmiştir. Eğer p, her bir hastanın tedaviye cevap verme olasılığı ve Y de n hasta içinden tedaviye cevap verenlerin sayısı ise Y rasgele değişkeninin dağılımı Binom dağılımı olur. Gözlem yapılmadan önce, hastalığın ölümcül olma bilgisi, p nin değerinin nispeten küçük olacağını düşündürür. p nin değerinin nispeten küçük olacağı bilindiğinden örneğin 0.25 komşuluğunda alınabilir.

Bu bilgiler kullanılarak, p nin çıkarım sürecinde nasıl bir yol izlenebilir?

Bu yollardan biri Bayesci yaklaşım kullanarak p ile ilgili önsel bilginin değerlendirilmesidir. Bu yaklaşımda, p bilindiğinde Y rasgele değişkeninin koşullu olasılık fonksiyonu P(y|p),

$$\mathrm{P(y|p)} = \binom{\mathrm{n}}{\mathrm{y}} \mathrm{p^y q^{n-y}}\ , \qquad \mathrm{y} = 0,1,\dots,\mathrm{n}.$$

şeklindedir.

p için önsel dağılım, veriler gözlemlenmeden önce belirlenmelidir. Önsel dağılım belirlenirken, p nin $0 < p < 1$ aralığında olduğu bilgisi kullanılabilir. Beta dağılımına sahip bir rasgele değişkenin $(0,1)$ aralığında değerler aldığı bilindiğinden p için önsel dağılım olarak Beta dağılımı alınabilir. Fakat burada hangi Beta dağılımının kullanılacağı problemi ortaya çıkar. Parametreleriα ve β olan Beta dağılımına sahip rasgele değişkenin ortalaması $\mu = \frac{\alpha}{(\alpha+\beta)}$ dür. Problemde, p nin 0.25 komşuluğunda olabileceği düşünülmüştü. Buradan, $\alpha = 1$ ve $\beta = 3$ olan Beta dağılımı alındığında $\mu = 0.25$ olarak bulunur ve bu da çalışmada önsel bilgi olarak kullanılabilir. Buna göre p nin olasılık yoğunluk fonksiyonu;

$$\mathrm{g(p)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{(\beta-1)}$$

$$= \frac{\Gamma(4)}{\Gamma(1)\Gamma(3)} p^0(1-p)^2 = \frac{\Gamma(3)3}{\Gamma(3)}(1-p)^2$$

$$g(p) = \frac{1}{3}(1-p)^2, \qquad 0 < p < 1$$

olur (Wackerly vd. 2008). p nin dağılımı ve $(Y|p)$ nin koşullu dağılımı belirlendikten sonra, Y nin marjinal dağılımı ve $(p|\mathrm{y})$ nin koşullu olasılık yoğunluk fonksiyonu

belirlenebilir. $\mathrm{Y} = \mathrm{y}$ verisi gözlemlendikten sonra, p nin sonsal olasılık yoğunluk fonksiyonu $\mathrm{g}^*(p|\mathrm{y})$ ile gösterilebilir. Buna göre $\mathrm{g}^*(\mathrm{p}|\mathrm{y}\ )$

$$\mathrm{g}^*(\mathrm{p}|\mathrm{y}\ ) = \frac{\Gamma(\mathrm{n}+4)}{\Gamma(\mathrm{y}+1)\Gamma(\mathrm{n}-\mathrm{y}+3)}\mathrm{p}^{\mathrm{y}}(1-\mathrm{p})^{(\mathrm{n}-\mathrm{y}+2)}, \quad 0 < p < 1$$

biçiminde elde edilir. Dikkat edilecek olursa elde edilen bu sonsal dağılım parametreleri,

$$\alpha^* = \mathrm{y}+1 \text{ ve } \beta^* = \mathrm{n}-\mathrm{y}+3$$

olan Beta dağılımıdır (Wackerly vd. 2008). Bu sonsal olasılık yoğunluk fonksiyonu p nin, alınan veri kullanılarak elde edilen güncellenmiş olasılık yoğunluk fonksiyonudur ve Bayes çıkarımının temelidir.

Bayesci sonuç çıkarımının genel hali şu şekilde verilebilir. $Y_1, Y_2, \ldots, Y_n$ birbirinden bağımsız aynı f(y; θ) olasılık (yoğunluk) fonksiyonuna sahip kitleden alınan rasgele örneklem ve olabilirlik fonksiyonu $L(y_1, y_2, \ldots, y_n|\theta)$ olsun. θ sürekli parametresi için önsel olasılık yoğunluk fonksiyonu $g(\theta)$ olmak üzere, θ parametresine ilişkin önsel dağılım bilgisi ve verinin olabilirliği kullanılarak $Y_1, Y_2, \ldots, Y_n, \theta$ için ortak olabilirlik fonksiyonu,

$$\mathrm{f}(\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n};\ \theta) = \mathrm{L}(\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n}|\ \theta)\ \mathrm{g}(\theta)$$

biçiminde yazılır. Buradan, $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_\mathrm{n}$ nin marjinal dağılımı sürekli durum için,

$$\mathrm{f}(\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n}) = \int_{-\infty}^{\infty} \mathrm{L}(\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n}|\theta)\ \mathrm{g}(\theta)\mathrm{d}\,\theta$$

olarak elde edilir. Buradan, $Y_1, Y_2, \dots, Y_n$ verildiğinde θ nın koşullu olasılık yoğunluk fonksiyonu yani, $\theta|y_1, y_2, \dots, y_n$ sonsal olasılık yoğunluk fonksiyonu,

$$g^*(\theta|y_1, y_2, \dots, y_n) = \frac{f(y_1, y_2, \dots, y_n; \theta)}{f(y_1, y_2, \dots, y_n)} = \frac{L(y_1, y_2, \dots, y_n|\theta)\, g(\theta)}{\int_{-\infty}^{\infty} L(y_1, y_2, \dots, y_n|\theta)\, g(\theta) d\,\theta}$$

olarak bulunur. Bayesci yaklaşımda; θ parametresi hakkında tüm sonuç çıkarımı bu sonsal olasılık yoğunluk fonksiyonu kullanılarak yapılır.

Burada, paydadaki ifade $\int_{-\infty}^{\infty} L(y_1, y_2, \dots, y_n|\theta)g(\theta)d\theta$ sonsal dağılımın normalleştirme sabiti olarak bilinir. Bu integral sonlu olduğu sürece sonsal dağılım için herhangi bir ilave bilgi sunmaz. Bundan dolayı, $g^*(\theta|y_1, y_2, \dots, y_n)$ sonsal olasılık (yoğunluk) fonksiyonu

$$g^*(\theta|y_1, y_2, \dots, y_n) \propto L(y_1, y_2, \dots, y_n|\theta)\, g(\theta) \qquad (2.1)$$

biçiminde orantısal olarak yazılabilir (Lynch, 2006). Yani,

$$Sonsal \propto Olabilirlik \times Önsel$$

olarak ifade edilir. Sonsal dağılımın olasılık (yoğunluk) fonksiyonu, olabilirlik fonksiyonu ile önsel olasılık (yoğunluk) fonksiyonunun çarpımının orantısal bir fonksiyonudur.

Buna göre, ***Bayesci yaklaşım*** ile var olan bilginin yeni gelen bilgi ile nasıl güncelleneceği ifade edilmiş olur. Bir parametre için önsel dağılım, veriyi toplamaya başlamadan önce parametre hakkında kesin olmayan bilgileri içeren olasılık dağılımının kendisidir. Bayesci çıkarımda, önsel dağılım ve örneklem bilgisi kullanılarak sonsal dağılım elde edilir. Parametreye ilişkin tüm çıkarım bu sonsal dağılım kullanarak yapılır. Bayesci çıkarımda Bayes tahmin edicisi elde edilen sonsal dağılımın beklenen değeridir.

Örnek 2.2. $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_\mathrm{n}$, $P(Y_i = 1) = p$ ve $P(Y_i = 0) = 1 - p$ olan Bernoulli dağılımndan n birimlik bir örneklem olsun ve p parametresi için önsel dağılımın Beta(α, β) olduğu varsayılsın. Buna göre p nin sonsal dağılımı elde edilsin.

p nin sonsal dağılımını elde edebilmek için yukarıda anlatılan işlemler sırayla uygulanmalıdır. Buna göre ilk olarak ortak olasılık yoğunluk fonksiyonu,

$$\mathrm{f}(\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n}, \mathrm{p}) = L(y_1, y_2, \ldots, y_n|p)\ g(p)$$

$$= p^{\sum y_i}(1-p)^{n-\sum y_i}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{(\beta-1)},\ y_i = 0{,}1\ ve\ 0 < p < 1$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{(\sum y_{i+}\alpha-1)}(1-p)^{(n-\sum y_i+\beta-1)}$$

ve buradan marjinal olabilirlik fonksiyonu;

$$f(y_1, y_2, \ldots, y_n) = \int_0^1 L(y_1, y_2, \ldots, y_n|p) \times\ g(p)dp$$

$$= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{(\sum y_{i+}\alpha-1)}(1-p)^{(n-\sum y_i+\beta-1)}d\,p$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\sum y_i+\alpha)\Gamma(n-\sum y_i+\beta)}{\Gamma(n+\beta+\alpha)}$$

olarak bulunur. Buradan sonsal dağılımın olasılık yoğunluk fonksiyonu;

$$\mathrm{g}^*(\mathrm{p}|\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n}) = \frac{\Gamma(\mathrm{n}+\alpha+\beta)}{\Gamma(\sum \mathrm{y}_\mathrm{i}+\alpha)\Gamma(\mathrm{n}-\sum \mathrm{y}_\mathrm{i}+\beta)}\mathrm{p}^{(\sum \mathrm{y}_\mathrm{i}+\alpha-1)}\ (1-\mathrm{p})^{(\mathrm{n}-\sum \mathrm{y}_\mathrm{i}+\beta-1)}$$

biçiminde elde edilir. Dikkat edilecek olursa elde edilen bu sonsal dağılım parametreleri,

$$\alpha^* = \sum y_i + \alpha \text{ ve } \beta^* = n - \sum y_i + \beta$$

olan Beta dağılımıdır (Wackerly vd. 2008).

Yeterli istatistik örneklem içinde parametre hakkındaki bilgileri özetleyen istatistiktir (Akdi, 2011). Sonsal dağılım yeterli istatistik kullanılarak bulunabilir. Buna göre, $Y_1, Y_2, \dots, Y_n$ parametresi p olan Bernoulli dağılımından bir örneklem olduğunda $Y = \sum_{i=1}^{n} Y_i$, p için yeterli istatistir. $Y \sim Binom(n, p)$ olmak üzere bu yeterli istatistik kullanılarak p için sonsal dağılım aşağıdaki gibi elde edilir. İlk olarak Y nin olasılık fonksiyonu ve p nin olasılık yoğunluk fonksiyonu yazılırsa,

$$f(y, p) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0,1, \dots, n$$

ve

$$g(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}, 0 < p < 1$$

olmak üzere ortak olasılık fonksiyonu,

$$f(y; p) = \binom{n}{y} p^y (1-p)^{n-y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

$$= \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1}, y = 0,1, \dots, n \text{ ve } 0 < p < 1$$

biçiminde elde edilir. Buradan Y nin marjinal fonksiyonu,

$$f(y) = \int_0^1 f(y;p)dp$$

$$= \int_0^1 \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1} d\,p$$

$$= \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+y)\Gamma(n-y+\beta)}{\Gamma(\alpha+\beta+n)}$$

şeklinde olup sonsal dağılımın olasılık yoğunluk fonksiyonu,

$$g^*(p|y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} p^{(y+\alpha-1)} (1-p)^{(n-y+\beta-1)}$$

olarak bulunur. Buna göre sonsal dağılım, parametreleri $y+\alpha$ ve $n-y+\beta$ olan Beta dağılımıdır. Görüldüğü gibi, yukarıda elde edilen sonuç tekrarlanmıştır. Buna göre Bayes tahmin edicisi,

$$E(p|y) = \int_0^1 p \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} p^{(y+\alpha-1)} (1-p)^{(n-y+\beta-1)} dp$$

$$= \frac{\sum_{i=1}^n Y_i + \alpha}{n+\alpha+\beta} = \frac{\sum_{i=1}^n Y_i}{n+\alpha+\beta} + \frac{\alpha}{n+\alpha+\beta}$$

$$= \frac{n}{n+\alpha+\beta}\overline{Y} + \frac{\alpha+\beta}{n+\alpha+\beta}\mu$$

olarak elde edilir. Burada, μ önsel dağılımın beklenen değeridir. Görüldüğü gibi, Bayes tahmin edicisi örneklem ortalaması ile önsel dağılımın beklenen değerinin lineer birleşimidir. Bayesci çıkarım için bu genellikle doğrudur.

Farklı gözlem değerleri için, önsel dağılımın parametreleri olan α ve β değerleri değiştiğinde sonsal dağılımların değişimini görebilmek için farklı α, β ve n değerlerinde elde edilen sonsal dağılım parametreleri aşağıda özetlenmiş ve grafikleri verilmiştir.

Çizelge 2.1 Önsel için α ve β değerleri ve örneklem bilgisi

	α	**β**	$\boldsymbol{n}$	$\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} y_i$
a)	1	3	5	2
b)	1	3	25	10
c)	10	30	5	2
d)	10	30	25	10

Alınan Beta önsel dağılımına ilişkin ortalama ve varyanslar,

a ve b için $\mu = \frac{\alpha}{(\alpha+\beta)} = \frac{1}{1+3} = 0.25 \sigma^2 = \frac{\alpha \times \beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = 0.375$

c ve d için $\mu = \frac{\alpha}{(\alpha+\beta)} = \frac{10}{30} = 0.25 \sigma^2 = \frac{\alpha \times \beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = 0.046$

biçimindedir. Klasik yöntem ile eldeki verilere göre p için en çok olabilirlik tahmini,

a *ve c için* $\hat{p} = \frac{1}{n}\sum y_i = \frac{2}{5} = 0.40$

b ve d için $\hat{p} = \frac{1}{n}\sum y_i = \frac{10}{25} = 0.40$

olarak bulunur. Şimdi Bayesci yöntemle elde edilen sonuçlara bakılsın.

$Y_1, Y_2, \ldots, Y_n$ rasgele değişkenleri, $P(Y_i = 1) = p$ ve $P(Y_i = 0) = 1 - p$ olan Bernoulli dağılımına sahip ve p nin önsel dağılımı Beta(α, β) olduğunda, p nin sonsal dağılımı da yine bir Beta dağılımıdır ve bu Beta dağılımının parametreleri

$$\alpha^* = \sum y_i + \alpha \text{ ve } \beta^* = n - \sum y_i + \beta \text{ dır.}$$

Buna göre her bir şık için sonsal dağılımın α^* ve β^* parametreleri bulunsun.

a) $\alpha^* = \sum y_i + \alpha = 2 + 1 = 3$ ve $\beta^* = n - \sum y_i + \beta = 5 - 2 + 3 = 6$

Çizelge 2.2 α=1, β=3 ve n=5 için sonuçlar

	α	**β**	**n**	$\sum_{i=1}^{n} y_i$	**α^***	**β^***
a)	1	3	5	2	3	6

Buna göre sonsal dağılımın ortalaması (yani Bayes tahmini) ve varyansı,

$\mu_{sonsal} = \frac{\alpha^*}{\alpha^*+\beta^*} = \frac{3}{3+6} = 0.0333$ $\sigma^2_{sonsal} = \frac{\alpha^* \times \beta^*}{(\alpha^*+\beta^*)^2(\alpha^*+\beta^*+1)} = 0.0222$

olarak bulunur.

b) $\alpha^* = \sum y_i + \alpha = 10 + 1 = 11$ ve $\beta^* = n - \sum y_i + \beta = 25 - 10 + 3 = 18$

Çizelge 2. 3 α=1, β=3 ve n=25 için sonuçlar

	α	**β**	**n**	$\sum_{i=1}^{n} y_i$	**α^***	**β^***
b)	1	3	25	10	11	18

Buna göre sonsal dağılımın ortalaması ve varyansı,

$\mu_{sonsal} = \frac{\alpha^*}{\alpha^*+\beta^*} = \frac{11}{11+18} = 0.4074$ $\sigma^2_{sonsal} = \frac{\alpha^* \times \beta^*}{(\alpha^*+\beta^*)^2(\alpha^*+\beta^*+1)} = 0.0078$

olarak bulunur.

c) $\alpha^* = \sum y_i + \alpha = 2 + 10 = 12$ *ve* $\beta^* = n - \sum y_i + \beta = 5 - 2 + 30 = 33$

Çizelge 2.4 α=10, β=30 ve n=5 için sonuçlar

	α	**β**	***n***	$\sum_{i=1}^{n} y_i$	α^*	β^*
c)	10	30	5	2	12	33

Buna göre sonsal dağılımın ortalaması ve varyansı,

$\mu_{sonsal} = \frac{\alpha^*}{\alpha^*+\beta^*} = \frac{12}{12+33} = 0.2667 \sigma^2_{sonsal} = \frac{\alpha^* \times \beta^*}{(\alpha^*+\beta^*)^2(\alpha^*+\beta^*+1)} = 0.0043$

olarak bulunur.

d) $\alpha^* = \sum y_i + \alpha = 10 + 10 = 20$ ve $\beta^* = n - \sum y_i + \beta = 25 - 10 + 30 = 45$

Çizelge 2.5 α=10 , β=30 ve n=25 için sonuçlar

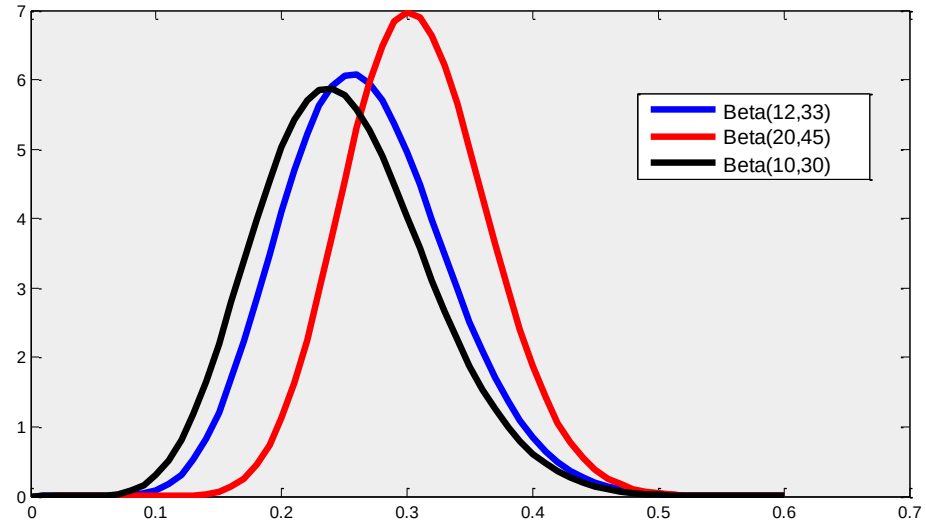
	α	**β**	***n***	$\sum_{i=1}^{n} y_i$	α^*	β^*
d)	10	30	25	10	20	45

Buradan sonsal dağılımın ortalaması ve varyansı,

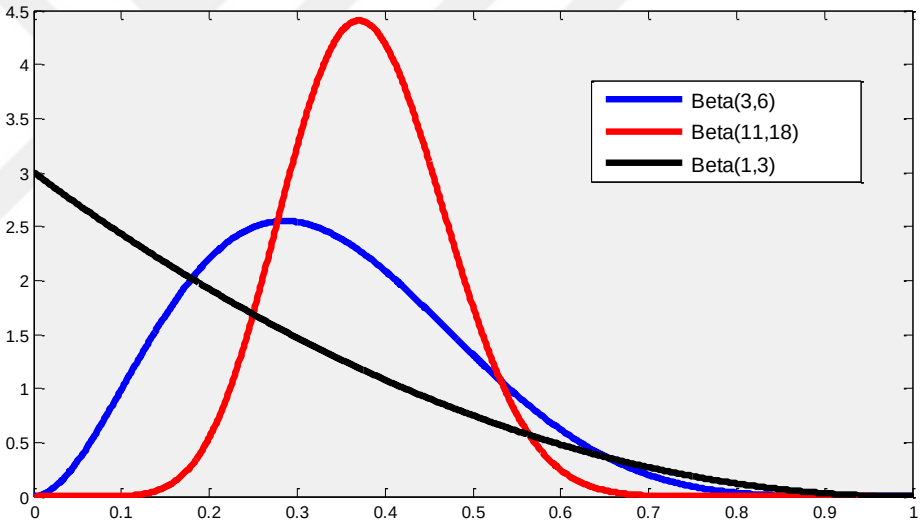
$\mu_{sonsal} = \frac{\alpha^*}{\alpha^*+\beta^*} = \frac{20}{20+45} = 0.3077 \quad \sigma^2_{sonsal} = \frac{\alpha^* \times \beta^*}{(\alpha^*+\beta^*)^2(\alpha^*+\beta^*+1)} = 0.0032$

olarak bulunur (Wackerly vd. 2008).

Buna göre, önsel ve sonsal dağılımların olasılık yoğunluk fonksiyonlarının grafiği aşağıdaki gibi olur.



Şekil 2.1. Önsel ve Sonsal Beta Dağılımları



Şekil 2.2 Önsel ve Sonsal Beta Dağılımları

2.2 Önsel Dağılımın Seçimi

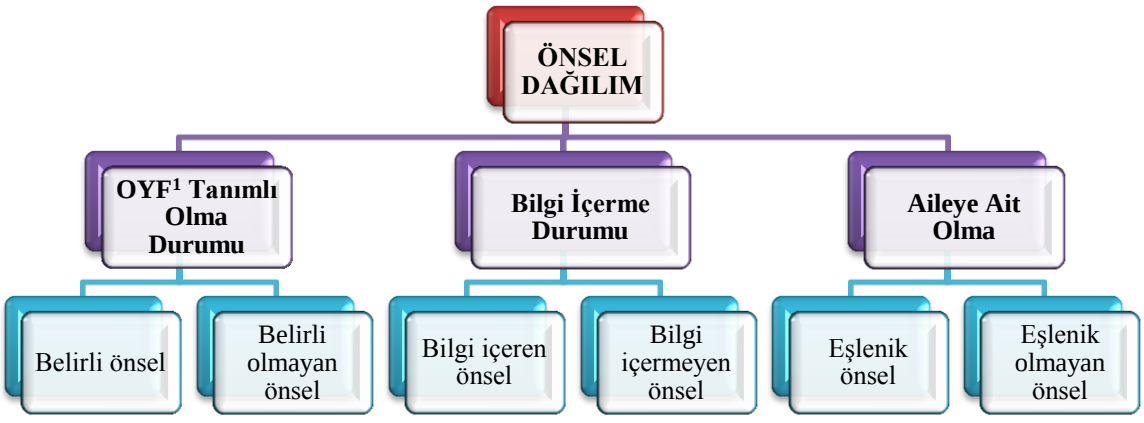
Bayesci sonuç çıkarımında, araştırmacılar, benzer veya geçmiş verilere ait yapılan çıkarımlardan veya diğer kaynaklardan parametre ile ilgili ön bilgiye sahip olabilirler. Bilinmeyen parametreler hakkında daha önceki çalışmalardan elde edilen bu bilgiler

önsel (prior) olarak adlandırılır. Bayesci yaklaşımda sonsal dağılımın (posterior distribution) elde edilebilmesi için önsel dağılımın belirtilmesi gerekir. Bayesci yaklaşımda, model parametreleri için önsel dağılımın belirlenmesi ve olasılıkları hakkında doğrudan çıkarım yapılabilmesi klasik yaklaşıma göre ayırt edici bir özelliktir. Bunu yapabilmek için araştırmanın başında model parametreleri için hangi önsel dağılımın kullanacağı belirlenmelidir. Bayesci sonuç çıkarımı açısından en zorlayıcı kısımlardan biri budur.

Bayesci analizde karar verilmesi gereken esas nokta önsel seçimidir. Çünkü araştırmacının bilinmeyen parametre için seçtiği önsel, analiz öncesi yaptığı araştırma kalitesine, inanç düzeyine, eldeki bilginin ne kadarını analize yansıtmak istediğine göre değişir. Şimdi bu önsel türleri incelensin.

2.2.1 Önsel dağılım türleri

Önceki çalışmalara ve kaynak bilgisine göre belirlenecek olan önsel dağılım aşağıdaki gibi sınıflandırılabilir (Güner, 2014).



Şekil 2.3 Önsel Dağılım Türleri

Belirsiz önsel olasılıklar, kesikli rasgele değişkenler için olasılık fonksiyonunun toplam değeri 1'e eşit olmayan veya sürekli rasgele değişkenler için olasılık yoğunluk fonksiyonunun integrali hesaplandığında sonsuz bir değere eşit olan önsel olasılıklardır (Ekici, 2005). Tanımlı olduğu aralıkta integrali veya toplamı 1'e eşit çıkan önsel dağılımlara ***belirli önsel dağılım***, aksi haldeki önsel dağılımlara ***belirsiz önsel dağılım*** denir (Tektaş, 2006).

Eğer önsel dağılımı elde etmek için parametre hakkındaki önsel bilgi, alanında uzman kişilerin bilgisine, daha önceden yapılan çalışmalara ve gözlem sonuçlarına, literatür bilgisine ve olası diğer veri kaynaklarına aitse ***bilgi içeren önsel dağılım*** kullanılmış olur.

Eğer Bayesci yaklaşım kullanılarak bir araştırma yapılacak fakat parametreler hakkındaki önsel bilgi için yeterli bir kaynağa ulaşılamamış ise ya da araştırma yapılacak veriler yeterli ve makul olduğunda ayrı bir önsel araştırmasına ihtiyaç duyulmadığında veya elindeki önsel bilgi veri bilgisinden çokta farklı değil ise araştırmacı ***bilgi içermeyen önsel*** dağılımı kullanır. Bilgi içermeyen önsel olasılık objektiftir. Yani araştırmacı kişisel bakış açısına dayanarak önsel seçimi yapmaz. Bilgi içermeyen önsel dağılımlara örnek olarak düzgün (tekbiçimli-uniform) önsel dağılımlar, Jeffreys'in önsel dağılımları verilebilir.

Bilgi içermeyen önsel olasılık, araştırmacının görüşlerine dayanan sübjektif bir bakış açısıyla elde edilmez. Bu sayede Bayesci yaklaşımın sübjektif olduğuna dair yapılan eleştiriler azaltılmış olur.

Düzgün önsel dağılımlar, belirli önsel dağılım ya da belirsiz önsel dağılım olabilirler. Örneğin, $P(\theta) = k, k < 1$ şeklindeki önsel dağılım da parametre değerinin ne olduğu bilinmediği için belirsiz düzgün dağılım iken; $P(\theta) = 1, 0 < \theta < 1$ olduğu durumda önsel dağılımın integrali 1 değerini aldığı için belirli önsel dağılım olur. Önsel dağılım belirsiz dahi olsa sonsal dağılım belirli olur (Myung, 2006).

Parametrelerle ilgili önsel bilginin belirsiz olmadığı durumda, önsel dağılım belli bir aileye ait iken sonsal dağılım da aynı ailenin üyesi ise bu tür dağılımlara ***eşlenik önsel dağılım*** denir. Yani, eşlenik önsel kullanılması durumunda, önsel dağılım ile sonsal dağılım aynı ailenin üyesidirler. Eşlenik önseller hesaplamalarda kolaylık sağladığı için çoğunlukla tercih edilirler. Örneğin, Üstel dağılım için eşlenik önsel dağılım Gama dağılımıdır. Eşlenik önseller, sonsal dağılımların kapalı formda elde edilmesine olanak sağladıkları için oldukça kullanışlıdırlar (Ando, 2010). Bazı eşlenik aileler Çizelge 2 ile verilmiştir (Press,1989).

Çizelge 2.6 Eşlenik önsel dağılımlar

Önsel	*Sonsal*	*Olabilirlik*
• *Beta*	• *Beta*	• *Binom*
• *Beta*	• *Beta*	• *Bernoulli*
• *Normal*	• *Normal*	• *Normal*
• *Gamma*	• *Gamma*	• *Poisson*
• *Gamma*	• *Gamma*	• *Üstel*
• *Gamma*	• *Gamma*	• *Gama*

3. BAYESCİ HİPOTEZ TESTLERİ

İlk olarak klasik bakış açısı ile hipotez testi göz önüne alınsın. Parametre kümesi, $\Theta \subset R^r$, $r > 1$ olmak üzere $X_1, X_2, \ldots, X_n$ olasılık ya da olasılık yoğunluk fonksiyonu $f(.;\theta), \theta \epsilon \Theta$ olan dağılımdan bir örneklem olsun. Parametre kümesinin, boş olmayan iki alt kümeye,

$$\Theta = \Theta_0 \cup \Theta_1, \Theta_0 \cap \Theta_1 = \emptyset$$

gibi bir parçalanması için, $\Theta \epsilon \Theta_0$ gibi bir ifadeye (önermeye, iddiaya) ***parametrik hipotez*** denir ve genellikle H_0 ile gösterilir. H_0 hipotezine ***sıfır*** yada ***yokluk hipotezi*** denir. $\Theta \epsilon \Theta_1$ ifadesi veya Θ_1 alt kümesi de θ hakkında bir hipotezdir. Bu hipoteze, H_0 hipotezinin ***karşıt (alternatif) hipotezi*** denir ve genellikle H_1 ile gösterilir. Klasik anlamda hipotez testi, ortaya atılan bu iki hipotezden hangisinin gözlenen örneklem değerleri ile desteklendiğinin ortaya çıkarılmasıdır (Öztürk, 2006).

Hipotez testinde parametre kümesi, hipotezleri oluşturan ayrık iki kümenin birleşimi olmak üzere, bir hipotez bir elemanlı olduğunda ***basit hipotez***, birden çok elemanlı olduğunda ***karmaşık hipotez*** adını alır.

Yapılan hipotez testi sonucunda aşağıdaki durumlar ortaya çıkabilir.

Çizelge **3.1** Klasik yaklaşımda karşılaşılan durumlar

		Gerçek	
		H_0 **doğru**	H_1 **doğru**
Sonuç	H_0 **red**	1.Tip hata(α)	Doğru karar
	H_0 **red edilememesi**	Doğru karar	2.Tip hata(β)

Hipotez testi probleminde 1. Tip hata yapma olasılığı için göze alınan ve önceden belirlenen sınıra ***testin anlam düzeyi*** denir ve α ile gösterilir. α aslında bir koşullu olasılık değeridir ve

$$P(I.tip\ hata) = P(H_0\ red \mid H_0 \text{ doğru}) = \alpha$$

ile ifade edilir. $1-\alpha$ ya testin ***güven düzeyi*** denir ve

$$1-\alpha = Güven\ Düzeyi = 1-P(I.tip\ hata) = P(H_0\ red\ edilemiyor \mid H_0 \text{ doğru})$$

ile ifade edilir. Hipotezin red edildiği bölgeye ***red bölgesi*** ya da ***kritik alan*** denir. $p-$değeri, H_0 doğru olduğunda test istatistiğinin hesaplanan değerine eşit ya da uç değerler alması olasılığıdır ve $p-$değeri, α değerinden büyük olursa H_0 yokluk hipotezi reddedilemez aksi takdirde H_0 reddedilir.

Klasik yaklaşımda hipotez testleri genellikle 3 aşama ile ifade edilir. 1.Aşamada, hipotez kurulur. 2.Aşamada, test istatistiği oluşturulur ve verilen örneklem değerlerine göre test istatistiğinin değeri H_0 ın doğruluğu koşulu altında hesaplanır. 3. aşamada ise test istatistiğinin hesaplanan değerinin kritik bölgeye düşüp düşmemesine göre H_0 hipotezi reddedilir veya reddedilemez.

Şimdi Bayesci yaklaşım ile hipotez testi ele alınsın.θparametresine ilişkin,

$$H_0: \theta \epsilon \Theta_0$$
$$H_1: \theta \epsilon \Theta_1$$

Hipotezi ile ilgilenilsin. Bayes kuralına göre; θ parametresine ilişkin önsel olasılıkların bilindiği varsayılsın. Yani,

$P(\theta \epsilon \Theta_0) = p_0$, H_0 ın önsel olasılığı

$P(\theta \epsilon \Theta_1) = p_1$, H_1 in önsel olasılığı

olsun. $X_1, X_2, \dots, X_n$rasgele örneklemi alındığında θ parametresine ilişkin sonsal olasılık,

$$P(\theta\epsilon\Theta|X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \frac{P(\theta\epsilon\Theta)P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n|\theta\epsilon\Theta)}{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)}$$

ile ifade edilir. T, θ parametresi için yeterli bir istatistik olmak üzere, T=t gözlemlendiğinde H_1 ve H_0 hipotezlerinin desteklenme olasılıkları θ nın sonsal olasılığı kullanılarak,

$$P(\theta\epsilon\Theta_1|T = t) = \frac{P(\theta\epsilon\Theta_1)P(T = t|\theta\epsilon\Theta_1)}{P(T = t)}$$

$$P(\theta\epsilon\Theta_0|T = t) = \frac{P(\theta\epsilon\Theta_0)P(T = t|\theta\epsilon\Theta_0)}{P(T = t)}$$

biçiminde elde edilir. Her iki sonsal olasılık birbirine oranlandığında,

$$\underbrace{\frac{P(\theta\epsilon\Theta_0|T = t)}{P(\theta\epsilon\Theta_1|T = t)}}_{\text{Sonsal odds}} = \underbrace{\left[\frac{P(\theta\epsilon\Theta_0)}{P(\theta\epsilon\Theta_1)}\right]}_{\text{Önsel odds}} \underbrace{\left[\frac{P(T = t|\theta\epsilon\Theta_0)}{P(T = t|\theta\epsilon\Theta_1)}\right]}_{\text{Bayes faktörü}}$$

Sonsal odds= Önsel odds ×Bayes faktörü

$$Bayes\ Faktörü\ = BF_{01} = \frac{Sonsal\ Odds}{Önsel\ Odds} = \frac{P((T = t|\theta\epsilon\Theta_0)}{P(T = t|\theta\epsilon\Theta_1)} \qquad (3.1)$$

olarak tanımlanır. Eğer önsel olasılıklardan biri diğerine göre daha tercih edilebilir değilse $P(H_0) = P(H_1)$ alınır. Bu durumda ***Bayes Faktörü*** sonsal oddsa eşit olur. ***Bayes faktörü***, Bayesci hipotez testlerinin en önemli noktalarından biridir. Çünkü, H_0 hipotezinin doğruluğuna ilişkin kanıtın gücünü verir (Terzi, 2008). H_0 ve H_1 hipotezlerinin seçiminde Bayes faktörü şu şekilde kullanılır, eğer BF_{01}>1 ise H_0

hipotezinin H_1 hipotezine tercihi olarak yorumlanır, aynı şekilde $BF_{01}<1$ ise H_1 hipotezinin H_0 hipotezine tercihi olarak yorumlanır.

Bayes faktörü, yokluk ve alternatif hipotezlerin model parametrelerine ilişkin önsel yoğunluklarına bağlıdır, hesaplamalarda yüksek dereceden integral gerektirebilir. Bilinen istatistik programlar da bu işlemler kolaylıkla yapılamaz ancak geliştirilen Winbugs, SPLUS ve R programları kullanılarak bu hesaplamalar yapılabilir. Bayes Faktörü aldığı değerlere göre [1]Thumb kuralı kullanılarak yorumlanabilir. Bu yorumlar aşağıdaki tabloda verilmiştir (Terzi, 2008).

Çizelge 3.2 Thumb $kuralına$[1]göre Bayes Faktör değerinin yorumlanması

Bayes Faktör değeri	**Yorum**
$BF_{01} \geq 1$	H_0 kabul edilir
$0.3162 \leq BF_{01} < 1$	H_0 a karşı az delil var
$0.1 \leq BF_{01} < 0.3162$	H_0 a karşı sağlam delil var
$0.01 \leq BF_{01} < 0.1$	H_0 a karşı güçlü delil var
$BF_{01} < 0.01$	H_0 a karşı kesin delil var

[1]Thumb kuralı: Aralık için Minimum Normal Değer= Ortalama – (2 × Standart sapma)
Aralık için Maksimum Normal Değer = Ortalama + (2 × Standart sapma)

4. ORAN İÇİN BAYESCİ SONUÇ ÇIKARIMI

Bir kasabada yasayan seçmenler ***kitle*** olarak alınsın. Burada belediye başkanlık seçimi için A ve B gibi iki aday olduğu düşünülsün. Bu kitleden rasgele bir örneklem alınsın ve Y rasgele değişkeni ile belediye başkanlığı için A adayına oy vereceğini ifade edenlerin sayısı gösterilsin. Bu durumda, π ile, seçilen bir bireyin A adayına oy verme olasılığı ifade edilir (Bolstad, 2004)..

Başarı olasılığı π olan bir deneyde Y_i ile i. denemedeki başarı ifade edilirse Y_i, 0 ve 1 değerlerini alan bir Bernoulli rasgele değişkeni olur. Y ile n denemedeki toplam başarı sayısı gösterilmek üzere Y rasgele değişkeninin dağılımı Binom(n, π) olur. Y yeterli olduğundan parametre hakkında örneklemdeki tüm bilgi Y ile özetlenir. Buna göre, $Y = y$ nin olasılık fonksiyonu π koşulu altında;

$$\mathrm{f(Y = y|\pi)}=\binom{\mathrm{n}}{\mathrm{y}} \pi^{\mathrm{y}}(1-\pi)^{\mathrm{n-y}},\ \mathrm{y} = 0,1,\dots,\mathrm{n}$$

dır (Bolstad, 2004).

Bayesci sonuç çıkarımı için, veri bilgisi kullanılmadan önce π parametresine ilişkin önsel dağılım belirlenir $(g(\pi))$. Önsel dağılım veriler gözlenmeden önceki inanç derecesini yansıtır. Buna göre oran ile ilgili Bayes teoremi özetlenecek olursa; sonsal dağılımın olasılık (yoğunluk) fonksiyonu önsel dağılımın olasılık (yoğunluk) fonksiyonu ile olabilirliğin çarpımıyla orantılı olarak elde edilebileceğinden,

$$\mathrm{g}(\pi|y) \ \propto\ g(\pi) \times \mathrm{f(y}|\pi)$$

yazılabilir. Bu oran, sonsal dağılımın yapısını verir fakat birebir sonsal olasılık (yoğunluk) fonksiyonunun kendisi değildir. Gerçek sonsala ulaşmak için, ortak

olabilirlik fonksiyonu, marjinal olabilirlik fonksiyonuna bölünür. Bu sayede elde edilen fonksiyon olasılık (yoğunluk) fonksiyonu olur. Yani tanımlı olduğu aralıkta toplamı (integrali) 1 değerini verir. Genel olarak sonsal dağılımın olasılık (yoğunluk) fonksiyonu

$$\mathrm{g}(\pi|y) = \frac{g(\pi) \times \mathrm{f}(\mathrm{y}|\pi)}{\int_0^1 g(\pi) \times \mathrm{f}(\mathrm{y}|\pi)\mathrm{d}\pi}$$

biçiminde elde edilir (Bolstad, 2004). Önselin seçimine bağlı olarak integralin kapalı formu her zaman elde edilemeyebilir. Bu yüzden bu tarz integralleri hesaplayabilmek için sayısal integrasyon yöntemleri kullanılabilir.

4.1 Oran İçin Düzgün Önsel Kullanılarak Sonuç Çıkarımı

π oranı hakkında ön bir bilgiye sahip olunmadığında, tanımlanan olayı eşit olasılıkla destekleyen bir önsel seçilebilir. Başka bir ifadeyle, önselin mümkün olduğunca objektif olması istenir. Kişisel fikirden çıkarım yapılmamalıdır. Bu koşullarda düz önsel kullanılabilir. Çünkü, düz önsel π başarı olasılığı için bütün olası değerlere eşit ağırlık verir.

π başarı olasılığı için önsel dağılım olarak Düzgün önsel kullanıldığında olasılık yoğunluk fonksiyonu,

$$g(\pi) = 1, 0 \leq \pi \leq 1$$

ile verilir. Buna göre sonsal dağılımın olasılık yoğunluk fonksiyonu

$$g\big((\pi|y)\big) \propto \binom{n}{y} \pi^{y}(1-\pi)^{n-y},\ 0 \leq \pi \leq 1$$

biçiminde elde edilir. Buradan da düzgün önselin kullanılması ile elde edilen sonsal dağılımın olabilirlikle orantılı olduğu görülür. Dikkat edilirse elde edilen sonsal dağılım yine bir Beta dağılımıdır. Elde edilen Beta dağılımının parametreleri, $\alpha = y + 1$ ve $\beta = n - y + 1$ dir (Bolstad, 2004).

4.2 Oran için Beta Önseli Kullanılarak Sonuç Çıkarımı

π bilinmeyen parametresi $(0,1)$ aralığında değer aldığı için önsel dağılım olarak $Beta(a,b)$ dağılımı alınabilir. $Beta$ dağılımının parametreleri, incelenen örneğe göre parametreyle ilgili ilk inancı yansıtacak şekilde seçilebilir. Bu durumda önsel olasılık yoğunluk fonksiyonu,

$$g(\pi; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \pi^{a-1}(1-\pi)^{b-1} \text{ için } 0 \leq \pi \leq 1$$

biçiminde verilir. Sonsal olasılık yoğunluk fonksiyonun, olabilirlik ve önsel olasılık yoğunluk fonksiyonunun çarpımı ile orantılı olduğu daha önce ifade edilmişti. Buna göre,

$$\mathrm{g}(\pi|y) \propto g(\pi) \times \mathrm{f}(\mathrm{y}|\pi)$$

olmak üzere

$$g((\pi|y) \propto \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \pi^{a-1}(1-\pi)^{(b-1)} \times \binom{n}{y} \pi^{y}(1-\pi)^{n-y}$$

olur. Denklemdeki önsel ve olabilirliğin sabit değerle çarpımı sonucu etkilemeyecektir. Buna göre bu çarpımdaki parametreye bağlı olmayan sabit değerler göz ardı edilerek sonsal dağılımın olasılık (yoğunluk) fonksiyonu

$$g((\pi|y) \propto \pi^{(y+a-1)}(1-\pi)^{(n-y+b-1)}, 0 \leq \pi \leq 1$$

biçiminde orantısal olarak yazılabilir. Buradan da olasılık (yoğunluk) fonksiyonu

$$g((\pi|y) = \frac{\Gamma(n+a+b)}{\Gamma(a+y)\Gamma(n-y+b)}\pi^{(y+a-1)}(1-\pi)^{(n-y+b-1)}, \quad 0 \leq \pi \leq 1$$

olarak elde edilir. Dikkat edilirse bu dağılım yine bir Beta dağılımıdır ve parametreleri, $\alpha^{\iota} = \alpha + y$ ve $b^{\iota} = b + n - y$ dir. Bu şekilde, integral hesaplamaya gerek kalmadan sonsal dağılım kolaylıkla elde edilmiş olur (Bolstad, 2004).

4.3 Oran için klasik ve Bayesci sonuç çıkarımına göre elde edilen tahminlerin karşılaştırması

Klasik yaklaşıma göre π nin en çok olabilirlik tahmin edicisi,

$$\hat{\pi}_f = \frac{Y}{n} \tag{4.1}$$

olarak bulunur. $\hat{\pi}_f$ tahmin edicisi yansızdır ve varyansı

$$\text{Var}(\hat{\pi}_f) = \frac{\pi \times (1-\pi)}{n} \tag{4.2}$$

ile verilir. Buna göre, $\hat{\pi}_f$ için hata kareler ortalaması;

$$\text{HKO}(\hat{\pi}_f) = 0^2 + var(\hat{\pi}_f)$$

$$= \frac{\pi \times (1-\pi)}{n}$$

olarak bulunur.

π için Bayes tahmin edicisi olarak sonsal dağılımın ortalaması alınır. Önsel olarak $Beta(1,1)$ olan düz önsel alındığında sonsal dağılımın ortalaması

$$\hat{\pi}_B = m^{\iota} = \frac{a^{\iota}}{a^{\iota} + b^{\iota}}$$

$$a^{\iota} = y + 1, \quad b^{\iota} = 1 - y + n$$

olarak bulunur. y, n denemede elde edilen başarıların sayısı olmak üzere denklem yeniden yazılırsa;

$$\hat{\pi}_B = \frac{y+1}{n+2} = \frac{y}{n+2} + \frac{1}{n+2}$$

olur. Buna göre tahmin edicinin örnekleme dağılımının ortalaması;

$$\hat{\pi}_B = \frac{n\pi}{n+2} + \frac{1}{n+2} \tag{4.3}$$

ve örneklem dağılımının varyansı;

$$\left[\frac{1}{n+2}\right]^2 \times n\pi(1-\pi) \tag{4.4}$$

olarak bulunur. Buna göre hata kareler ortalaması;

$$\text{HKO}(\widehat{\theta}) = \text{bias}(\widehat{\theta})^2 + var(\widehat{\theta})$$

formülünde yerine konularak

$$HKO(\hat{\pi}_B) = \left[\left[\frac{n\pi + 1}{n+2} - \pi\right]\right]^2 + \left[\frac{1}{n+2}\right]^2 n\pi(1-\pi)$$

$$= \left[\frac{1-2\pi}{n+2}\right]^2 + \left[\frac{1}{n+2}\right]^2 n\pi(1-\pi) \quad (4.5)$$

biçiminde elde edilir.

Elde edilen sonuçlara göre Bayes tahmin edicisinin hata kareler ortalamasının klasik yöntemle elde edilenden daha küçük olduğu söylenebilir. Bir başka deyişle, ortalamada, Bayes yaklaşımı gerçek değere daha yakındır. Bu nedenle, Bayes tahmin edicisinin hata kareler ortalaması anlamında klasik yöntemle elde edilen tahmin ediciden daha iyi olduğu sonucuna ulaşılır (Bolstad, 2004).

Örnek 4.1 Bu sonucu sayısal olarak görebilmek için, $\pi = 0.4$ ve örneklem genişliği n=10 olarak alınsın. Buna göre her iki yöntemle elde edilen hata kareler ortalamaları,

$$HKO(\hat{\pi}_f) = \frac{0.4 \times 0.6}{10} = 0.024$$

$$\text{HKO}(\hat{\pi}_B) = \left[\frac{1 - 2 \times 0.4}{10+2}\right]^2 + \left[\frac{1}{12}\right]^2 \times 4 \times 0.6$$

$$= 0.0169$$

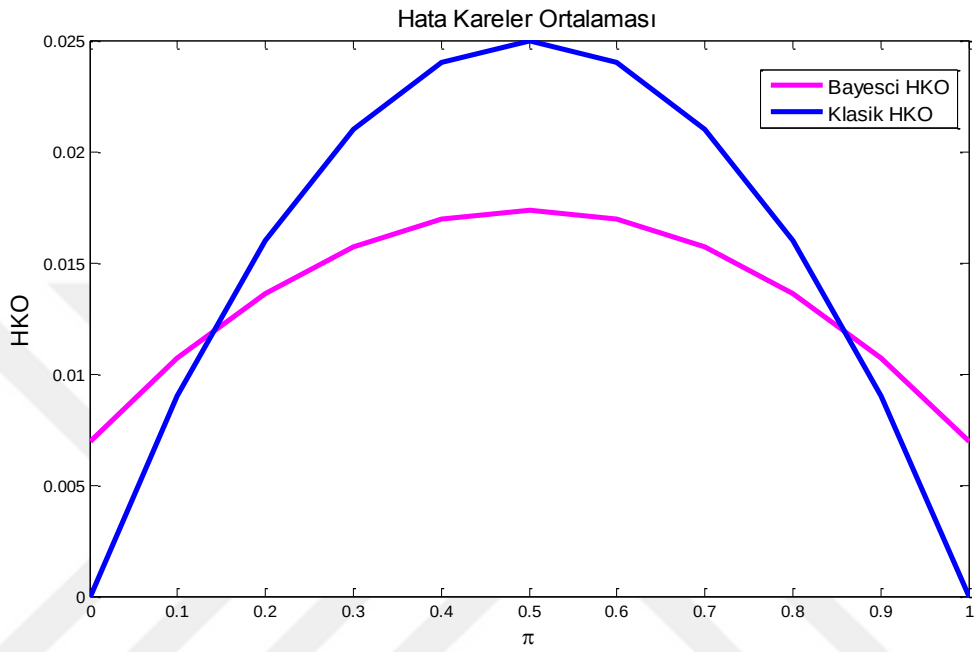
olarak bulunur. Benzer şekilde, örneklem genişliği $n = 10$için $\pi = 0.5$ olarak alındığında hata kareler ortalamaları,

$$HKO(\hat{\pi}_f) = \frac{0.5 \times 0.6}{10} = 0.025$$

$$\text{HKO}(\hat{\pi}_B) = \left[\frac{1 - 2 \times 0.5}{10+2}\right]^2 + \left[\frac{1}{12}\right]^2 \times 5 \times 0.5$$

$$= 0.01736$$

olarak bulunur. Bu şekilde $n = 10$ için değişik π değerlerine göre hata kareler ortalamaları hesaplanarak bunlara ait grafik çizdirildiğinde aşağıdaki sonuç elde edilir.



Şekil 4.1 Oran tahmini için Bayesci ve Klasik hata kareler ortalamaları karşılaştırılması

Elde edilen bu sonuca göre Bayes yöntemi kullanılarak elde edilen tahminin klasik yöntemle elde edilen tahmine göre hata kareler ortalaması anlamında daha iyi olduğu söylenebilir.

4.4 Oran İçin Bayesci Güvenilir Aralık

Klasik yaklaşıma göre θ parametresi için (1-α) güven düzeyinde güven aralığı,

$$P(l \leq \theta \leq u) = 1 - \alpha$$

ile verilir. Klasik yaklaşımda bu güven aralığının yorumu, bu aralığın θ parametresini içeren aralıklardan bir tanesi olması olasılığı $(1-\alpha)$'dır biçiminde yapılır. Çünkü burada rasgele olan l ve u sınırlarıdır. Bu sınırlar örneklemden örnekleme değişkenlik gösterir.

Örneklem oranı,

$$\hat{\pi}_f = \frac{Y}{n}$$

biçiminde ifade edilmek üzere n büyük olduğunda dağılımı ortalaması π, standart sapması $\sqrt{\frac{\pi(1-\pi)}{n}}$ olan yaklaşık normaldir.

Buradan da π için yaklaşık $(1-\alpha)$ güven düzeyli güven aralığı:

$$\hat{\pi}_f \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{\pi}_f(1-\hat{\pi}_f)}{n}} \tag{4.6}$$

biçiminde yazılabilir.

Bayesci yaklaşıma göre parametre ile ilgili her türlü sonuç çıkarımı sonsal dağılım kullanılarak yapılır. Parametreye ilişkin sonsal dağılım elde edildiği için bu parametreyi belli bir olasılıkla içeren aralık kolaylıkla belirlenebilir. Buna göre örneklem gözlendiğinde sonsal dağılım kullanılarak θ parametresi için$(1-\alpha)$ güven katsayılı güvenilir bölge,

$$P(\theta|x) = \int_{l}^{u} f(\theta|x) d\,\theta = 1 - \alpha$$

biçiminde belirlenebilir. Bayesci yaklaşımda güven aralığı yerine ***güvenilir (inanılır) aralık*** kavramı kullanılır (Bolstad, 2004).

Örnek 4.2 Belediye başkanlığı seçimleri için yapılan anket çalışması sonucunda $n = 100$ kişiden 36 tanesinin A adayına oy vereceğinisöylediği görülsün. Buna göre $\alpha = 0.05$ anlam düzeyinde π parametresi için güven aralığı klasik ve Bayesci yöntemlerle elde edilsin.

Yukarıda verilen sonuçlara göre klasik yöntemle güven aralığı,

$$\hat{\pi}_f \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{\pi}_f(1-\hat{\pi}_f)}{n}}$$

$$0.36 \pm 1.96 \times \sqrt{\frac{0.36 \times (1-0.36)}{100}}$$

$$(0.2659, 0.4541)$$

olarak elde edilir.

Bayesci yöntem için, düz önsel kullanıldığı durumda sonsal dağılım $\alpha = y + 1$ ve $\beta = n - y + 1$ olan Beta dağılımı olarak bulunmuştu. Buna göre, $(1 - \alpha)$ güven katsayılı

güven bölgesi Matlab programı yardımıyla, betainv(0.025,36,64) ve betainv(0.975,36,64) komutları kullanılarak,

$$(0.2693, 0.4560)$$

biçiminde elde edilir.

Klasik yöntemle elde edilen aralık uzunluğu 0.1882 iken Bayesci yöntemle elde edilen aralık uzunluğu 0.1867 olarak bulunmuştur. Bu değerler birbirine oldukça yakın olmakla birlikte Bayesci yöntemle elde edilen aralık klasik yönteme göre biraz daha dardır. Bunun sebebi Bayesci yöntemle bulunan aralığın parametrenin kendi dağılımı kullanılarak elde edilmesidir.

Parametre için önsel dağılım Beta(a, b) dağılımı olarak alındığında sonsal dağılım yine bir Beta$(\alpha' = \alpha + y,\ b' = b + n - y)$ dağılımı olarak elde edilmişti. Buna göre önsel dağılımın farklı parametrelerine göre elde edilen sonsal dağılım için güven aralıkları,

Beta$(a = 1, b = 1)$	$betainv(0.025,37,65) =$ 0.2727
	$betainv(0.975,37,65) =$ 0.4579
$Beta(a = 2, b = 1)$	$betainv(0.025,38,65) =$ 0.2788
	$betainv(0.975,38,65) =$ 0.4639
$Beta(a = 1, b = 2)$	$betainv(0.025,37,66) =$ 0.2698
	$betainv(0.975,37,66) =$ 0.4538

biçiminde bulunmuştur. Bu sonuçlar Çizelge 4.2 ile özetlenmiştir.

Çizelge 4.1 Bayesci güvenilir aralık sınırları

Önsel	**Güven sınırları**	**Aralık uzunluğu**
Düz Önsel	(0.2693,0.4560)	0.1867
Beta($a = 1, b = 1$)	$(0.2727, 0.4579$	0.1852
$Beta(a = 2, b = 1)$	$(0.2788, 0.4639)$	0.1851
$Beta(a = 1, b = 2)$	$(0.2698, 0.4538)$	0.1840

Sonuçlardan görüldüğü gibi önsel dağılımın parametreleri farklı olarak da alınsa oran için oluşturulan Bayesci güvenilir aralık klasik yöntemle elde edilenden daha dardır.

4.5 Oran için Hipotez Testi

Bayesci yöntemle hipotez testi yapıldığında bakılması gereken iki durum vardır. Birincisi sadece tek yanlı etkiyi tespit etmekle ilgilenilen tek yanlı hipotez testidir. Tek yanlı hipotez testi için

$$H_0: \pi \leq \pi_0$$
$$H_1: \pi > \pi_0$$

H_0 hipotezinin sonsal olasılık değeri H_1 hipotezinin sonsal olasılık değerinden küçük ise H_0 yokluk hipotezi reddedilir. Fakat bu durum Bayesci yöntemde hipotezin reddi olarak yorumlanmaz. Hangi hipoteze ait sonsal olasılık değeri daha büyükse o hipotezin daha olası olduğu söylenir ve o hipotez yönünde karar verilir.

İkinci durum ise; her iki yönde de bir kaymanın tespit edilmek istendiği iki yanlı hipotez testidir. Burada iki yanlı alternatif hipoteze karşı nokta hipotezi (yokluğun etkisi) test edilir. Bu tür hipotez testlerinde, eğer parametre için sürekli önsel dağılım kullanılırsa elde edilecek sonsal dağılım da sürekli olacaktır. Buna göre, sonsal dağılım kullanılarak hesaplanacak olasılık değeri sıfır olacaktır. Bu nedenle, tek noktaya eşitlik

için hipotez testi yapılmak istendiğinde hipotezin ***sonsal olasılığı kullanılamayacaktır.*** Bu gibi durumlarda hipotez testinde karar verebilmek için parametreye ilişkin güvenilir aralık kullanılır. Yokluk hipotezinde verilen değerin güvenilir aralıkta olup olmadığına bakılarak test gerçeklenir. Eğer yokluk hipotezinde verilen değer güvenilir aralıkta ise, yokluk hipotezi lehine karar verilir. Çünkü yokluk değeri güvenilir bir değer olarak kalır. Yani, iki yanlı hipotez testi için,

$$H_0: \pi = \pi_0$$
$$H_1: \pi \neq \pi_0$$

H_0 hipotezinin sonsal olasılığını hesaplamak yerine güvenilir aralık kullanılır. π parametresi için (1-α) olasılıklı güvenilir aralık bulunur. Eğer yokluk hipotezi altında verilen π_0 değeri güvenilir aralığın içindeyse yokluk hipotezi reddedilmez (Bolstad, 2004).

4.5.1 Tek yanlı Bayesci hipotez testi

$$H_0: \pi \leq \pi_0$$
$$H_1: \pi > \pi_0$$

hipotezi α anlam düzeyinde Bayesci yaklaşım kullanılarak test edilmek istensin. Burada yokluk hipotezinin doğruluğu koşulu altında sonsal olasılık

$$P(H_0: \pi \leq \pi_0 | \mathrm{y}) = \int_0^{\pi_0} g(\pi|y) d\pi$$

ile hesaplanır. Eğer hesaplanan bu sonsal olasılık değeri H_1 hipotezinin sonsal olasılık değerinden büyük ise H_0 yokluk hipotezinin daha olası olduğu söylenir ve H_0 hipotezi yönünde karar verilir.

Örnek 4.3 Yeni bulunan bir tedavinin, standart uygulanan tedaviden daha iyi olup olmadığı araştırılmak istensin. Eğer yeni tedavi daha iyi ise, yeni tedaviden fayda gören hastaların oranı (π), standart uygulanan tedaviden fayda görenlerin oranından (π_0) daha

iyi olmalıdır. Geçmiş kayıtlardan π_0= 0.6 olarak bilinmektedir. Rasgele seçilen 10 hastaya yeni tedavi uygulanmıştır. Y, tedaviden fayda görenlerin sayısı olmak üzere dağılımı binom(n, π) olur. Gözlem sonuçlarına göre 8 hastanın yeni tedaviden fayda gördüğü görülmüştür (y=8). Bu sonuç önceki elde edilen 0.6 dan daha iyi gözükmektedir. π>0.6 olması hipoteziBayesci hipotez testi kullanılarak aşağıdaki gibi yapılır,

$$H_0: \pi < 0.6$$
$$H_1: \pi \geq 0.6.$$

Parametre için önsel dağılım olarak $Beta(1,1)$ alındığında, $y = 8$ gözlemi için sonsal dağılım $Beta(9,3)$ dağılımıdır. Buna göre yokluk hipotezinin sonsal olasılığı,

$$P(H_0: \pi \leq 0.6 | y = 8) = \int_0^{0.6} g(\pi|y)d\pi$$

$$= \int_0^{0.6} \frac{\Gamma(12)}{\Gamma(3)\Gamma(9)} \pi^{(2)}(1-\pi)^{(8)} d\pi$$

integralinin çözümü olacaktır. Bu integralin çözümü için Matlab programından faydalanılırsa,

$$betacdf(0.6,9,3) = 0.1189$$

olarak hesaplanır. H_0 hipotezinin sonsal olasılık değeri 0.1189, H_1 hipotezinin sonsal olasılık değeri olan 0.8811(=1-0.1189) den küçük olduğundan H_0 yokluk hipotezinin daha olası olmadığı söylenir ve H_1 hipotezi yönünde karar verilir. Yani, yeni tedavinin başarı oranının 0.6 dan daha büyük olduğu söylenebilir.

Bayes faktörü kullanılarak yorum yapılmak istendiğinde, önsel olarak düzgün önsel alındığı durumda $\mathrm{P}(H_0) = P(H_1) = 1$ olduğu için;

$$\frac{P(\theta \epsilon \Theta_0 | T = t)}{P(\theta \epsilon \Theta_1 | T = t)} = \frac{P(H_0 : \pi \leq 0.6 | \mathrm{y} = 8)}{P(H_0 : \pi > 0.6 | \mathrm{y} = 8)} = \frac{0.1189}{0.8811} = 0.134$$

elde edilir. Buna göre, Thumb kuralında, 0.134 değeri $0.3162 \leq \mathrm{BF}_{01} < 0.1$ aralığına düştüğü için H_0 a karşı sağlam delil var olduğu görülür. Hatta H_0'a karşı güçlü delilin var olduğu aralığa daha yakındır. Bu yüzden, Bayes faktörüne göre de alternatif hipotezin gerçeklenmesi olasılığı daha fazladır.

4.5.2 İki yanlı Bayesci hipotez testi

$$H_0 : \pi = \pi_0$$
$$H_1 : \pi \neq \pi_0$$

hipotezi test edilmek istensin.

Eğer parametre için sürekli önsel dağılım kullanılırsa elde edilecek sonsal dağılım da sürekli olacaktır. Buna göre, yukarıda da ifade edildiği gibi sonsal dağılım kullanılarak hesaplanacak olasılık değeri sıfır olacaktır. Bu nedenle, tek noktaya eşitlik için hipotez testi yapılmak istendiğinde ***sonsal olasılık kullanılamayacaktır.*** Bu durumda sonsal olasılığı hesaplamak yerine, güvenilir aralık kullanılır.π parametresi için (1-α) olasılıklı güvenilir aralık bulunur. Eğer yokluk hipotezi altında verilen π_0 değeri güvenilir aralığın içindeyse yokluk hipotezi lehine karar verilir. Aksi halde yokluk hipotezi aleyhine karar verilir (Bolstad, 2004).

Örnek 4.4 Atılan bir madeni paranın hileli olup olmadığı ile ilgilenilsin. Buna göre kurulacak hipotez testi,

$$H_0 : \pi = 0.5$$

$$H_1: \pi \neq 0.5$$

biçiminde olur.

Parametre için önsel dağılım olarak düz önsel kullanıldığında, sonsal dağılım Beta$(\alpha' = a + y, b' = b + n - y)$ olarak elde edilmişti.

İlk simülasyon çalışmasında gözlemler hilesiz bir paradan 10 kez alınmış (yani binom dağılımının başarı olasılığı 0.5 olarak veri üretilmiştir) ve deneme sonucu 5 tane Tura gözlenmiştir. Buna göre sonsal dağılım $Beta(6,6)$ olacaktır. Bayesci sonuç çıkarımına göre parametre için 0.95 güvenilir aralık Matlab programı yardımıyla,

betainv(0.025,6,6) = 0.2338
betainv(0.975,6,6) = 0.7662

olmak üzere $(0.2338,0.7662)$ biçiminde elde edilir. $\pi_0 = 0.5$ değeri güvenilir aralığın içinde yer aldığı için H_0 hipotezi reddedilemez sonucuna ulaşılır.

Klasik yaklaşımla bu hipotez test edilmek istendiğinde,

1) Hipotez kurulur.

$H_0: \pi = 0.5$
$H_1: \pi \neq 0.5$

2) Test istatistiği

$$Z = \frac{\pi - \hat{\pi}}{\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}} \sim N(0,1)$$

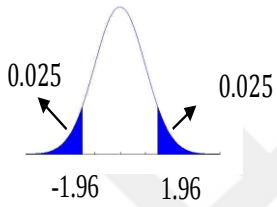
olmak üzere test istatistiğinin aldığı değer,

$$z_h = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.5 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{10}}} = 0$$

biçiminde hesaplanır. π_0 değeri H_0 yokluk hipotezinde verilen değerdir.

3) Kritik bölgeye göre hipotez reddedilir ya da reddedilemez.

$H_0: \pi = \pi_0$
$H_1: \pi \neq \pi_0$



$z_h = 0$ olduğundan kritik bölgeye düşmez dolayısı ile H_0 hipotezi reddedilemez. Klasik yöntemle %95 güven düzeyinde güven aralığı Matlab yardımı ile

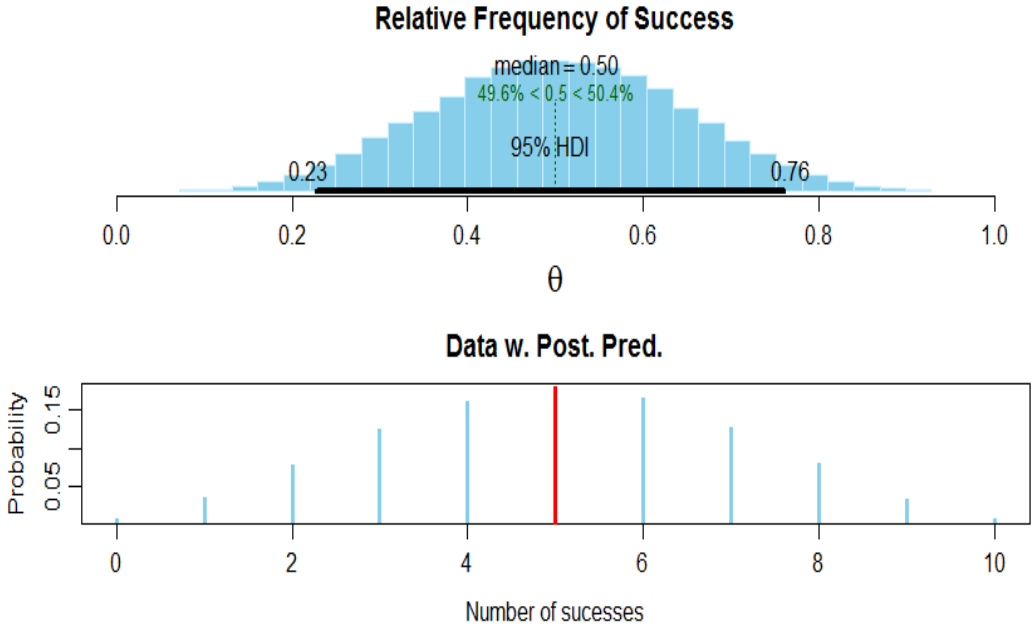
0.5-1.96*sqrt(0.5*0.95/10) = 0.0728
0.5+1.96*sqrt(0.5*0.95/10) = 0.9272

(0.0728, 0.9272) olarak bulunur. Elde edilen bu aralık Bayesci yöntemle elde edilen güvenilir aralığa göre oldukça geniştir.

Aynı test R programında veri seti $n = 10 < 30$ olduğundan binom.test komutu kullanılarak yapıldığında, $\alpha = 0.05$ için elde edilen güven aralığı, (0.187086,0.812914) biçiminde olup π_0 =0.5 değerini içerdiği için yokluk hipotezi reddedilemez sonucu tekrarlanmış olur.

Yaklaşık Normal dağılım kullanan prop.test komutu ile R programında işlem yapıldığında ise, güven aralığı (0.2365931,0.7634069) olarak elde edilmiştir.

R programı kullanılarak Bayesci sonuç çıkarımı yapılmak istendiğinde bayes.prop.test komutu kullanılır. Bayesci sonuç çıkarımı ile R da elde edilen sonuca göre %95 lik güvenilir aralık (0.23,0.77) olarak elde edilmiştir. R programında oran için Bayesci sonuç çıkarımı yapılmak istendiğinde önsel bilgi olarak Beta(1,1) dağılımı alınmaktadır. Bu da aslında, Düzgün(0,1) dağılımına denk gelmektedir. R programında Bayesci sonuç çıkarımı için sonsal dağılımdan örneklem çekilir. Aşağıda elde edilen bu örnekleme ilişkin histogram ve başarı sayılarına ilişkin çubuk grafiği verilmiştir. Histogramda %95 lik güvenilir aralığın alt ve üst sınırları gösterilmiştir (EK 1).



Şekil 4.2 R programında sonuç grafiği

5. NORMAL DAĞILIMIN ORTALAMASI İÇİN BAYESCİ SONUÇ ÇIKARIMI

5.1 Sürekli Önsel Kullanarak Normal Dağılım Ortalaması İçin Sonuç Çıkarımı

$Y_1, Y_2, \dots, Y_n$ bilinen σ^2 varyanslı Normal dağılımdan alınan bir örneklem, μ için önsel olasılık yoğunluk fonksiyonu $g(\mu)$ olmak üzere sonsal olasılık yoğunluk fonksiyonu,

$$g(\mu|y_1, y_2, \dots, y_n) \propto g(\mu) f(y_1, y_2, \dots, y_n|\mu)$$

önsel olasılık (yoğunluk) fonksiyonu ile olabilirliğin çarpımıyla orantılı olarak yazılabilir ya da tam olarak

$$g(\mu|y_1, y_2, \dots, y_n) = \frac{g(\mu) f(y_1, y_2, \dots, y_n|\mu)}{\int g(\mu) f(y_1, y_2, \dots, y_n|\mu)\, d\mu} \tag{5.1}$$

ile ifade edilir. Normal dağılım için rasgele örneklemin olabilirliği, örneklem ortalamasının olabilirliği ($\bar{y}$) ile orantılıdır. Buna göre herhangi bir sürekli $g(\mu)$ önsel dağılımı için,

$$g(\mu|y_1, y_2, \dots, y_n) = \frac{g(\mu) e^{\frac{-(\bar{y}-\mu)^2}{2\sigma^2/n}}}{\int g(\mu) e^{\frac{-(\bar{y}-\mu)^2}{2\sigma^2/n}}\, d\mu}$$

biçiminde yazılabilir. Burada integral almak zor olacağı için nümerik yöntemler kullanılmalıdır. Bazı özel durumlarda orantısal ifadeden integrale gerek kalmadan sonuca ulaşılabilir (Bolstad, 2004).

5.1.1 μ için Düz önsel alındığında

Düz önselde parametre için seçilen aralıkta parametreye aynı olasılık değerleri atanır. Yani düz önsel, μ nün her olası değeri için eşit ağırlık verir. Düz önsel herhangi bir değere, başka bir değere göre üstünlük vermez. Aralık dahilindeki her bir olayın gerçekleşme olasılığı, başka bir olayın gerçekleşme olasılığından büyük değildir. g(μ)=1 düz önseli seçildiğinde gerçekteμ, $-\infty \leq \mu \leq \infty$ aralığında değer aldığından uygun bir önsel dağılım değildir. Yani, bu aralıktaki integrali bir değerini vermez. Ancak, önsel olarak bu belirsiz önselin alınması sonsal dağılımın hesaplanmasında bir sorun yaratmaz. Önsel belirsiz alınmasına rağmen, sonsalın integrali 1 sonucunu verir. Yani, uygun bir dağılım olur(Bolstad, 2004).

Tek Normal gözlem alındığı durum,

Y gözlemi Normal(μ,σ^2) dağılımından alınmış olsun. Olabilirlik fonksiyonu sabitler göz ardı edildiğinde;

$$f(y|\mu) \propto e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

Gibi orantılı biçimde verilebilir. Düz önsel kullanıldığından sonsal olasılık yoğunluk fonksiyonu da olabilirlik fonksiyonuyla orantılı olmuş olur. O zaman sonsal olasılık yoğunluk fonksiyonu;

$$g(\mu|y) \propto e^{-\frac{1}{2\sigma^2}(\mu-y)^2}$$

olarak yazılır ve buradan sonsalın $Normal(y, \sigma^2)$ dağılımına sahip olduğu görülür (Bolstad, 2004).

$Y_1, \ldots, Y_n$ nin Normal rasgele örneklem olarak alındığı durum,

Normal dağılımdan alınan rasgele örneklemin olabilirliği, örneklem ortalamasının ($\bar{Y}$) olabilirliği ile orantılıdır. ($\bar{Y}$) dağılımının Normal(μ,$\frac{\sigma^2}{n}$) olduğu biliniyor. Buna göre olabilirlik fonksiyonu;

$$f(\bar{y}|\mu) \propto e^{\frac{-1}{2\sigma^2/n}(\bar{y}-\mu)^2}$$

biçiminde orantısal olarak yazılabilir. Düz önsel kullanıldığından sabit değerler göz ardı edilerek sonsal dağılımın olasılık yoğunluk fonksiyonu

$$g(\mu|\bar{y}) \propto e^{\frac{-1}{2\sigma^2/n}(\mu-\bar{y})^2}$$

biçiminde yazılır. Buna göre sonsal dağılım ortalaması $\bar{Y}$, varyansı $\frac{\sigma^2}{n}$ olan Normal dağılım olur (Bolstad, 2004).

5.1.2 μ için Normal önsel alındığında,

Tek Normal gözlem alındığı durum,

y değeri N(μ, σ^2) dağılımından alınmış rasgele bir gözlem değeri olsun. μ parametresi için önsel dağılım olarak ortalaması m, varyansı S^2 olan normal dağılım ($Y \sim (N(m, S^2))$ alınırsa önselin yoğunluk fonksiyonu orantısal olarak;

$$g(\mu) \propto e^{-\frac{1}{2S^2}(\mu-m)^2}$$

biçiminde yazılabilir. Önselin herhangi bir sabit ile çarpımı sonsal dağılımda etkisiz olacaktır. Bu nedenle olabilirlik fonksiyonu,

$$f(\mathrm{y}|\mu) \propto e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

biçiminde yazılabilir. μ ye bağlı olmayan kısımlar denklemden çıkarılmıştır. Buna göre önsel ve olabilirlik fonksiyonlarının,

$$g(\mu)f(\mathrm{y}|\mu) \propto e^{-\frac{1}{2S^2}(y-\mu)^2} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$$\propto e^{-\frac{1}{2S^2}(\mu-m)^2-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$$\propto e^{\frac{-1}{2}\left[\frac{(y^2-2\,\mu y+\mu^2)\times\sigma^2+(\mu^2-2\mu^2 m+m^2)\times S^2}{\sigma^2 S^2}\right]}$$

olur. Burada, üstel ifade $(\sigma^2+S^2)\Big/(\sigma^2S^2)$ ifadesi ile çarpılıp düzenlenirse;

$$g(\mu)f(\mathrm{y}|\mu) \propto e^{\frac{-1}{2\sigma^2S^2/\sigma^2+S^2}\left[\mu^2-2\frac{(\sigma^2m+S^2y)\mu}{(\sigma^2+S^2)}+\left(\frac{(\sigma^2m+S^2y)}{(\sigma^2+S^2)}\right)^2\right]}$$

elde edilir. μ ye bağlı olmayan terimler, iki kare farkının kapalı formu şeklinde ayrı biçimde yazılırsa;

$$g(\mu)f(\mathrm{y}|\mu) \propto e^{\frac{-1}{2\sigma^2S^2/\sigma^2+S^2}\left[\mu-\frac{(\sigma^2m+S^2y)}{(\sigma^2+S^2)}\right]^2}$$

olur. Elde edilen bu sonsal dağılım, parametreleri

$$m^{\iota} = \frac{(\sigma^2 m + S^2 y)}{(\sigma^2+S^2)} \quad ve\ (S^{\iota})^2 = \frac{\sigma^2 S^2}{\sigma^2+S^2} \tag{5.2}$$

olan Normal dağılım olur. Yani, sonsal dağılım Normal($\frac{(\sigma^2 m+S^2 y)}{(\sigma^2+S^2)}, \frac{\sigma^2 S^2}{\sigma^2+S^2}$) dir.

Görüldüğü gibi önsel olarak $Normal(m, S^2)$ dağılımı ile başlanıp, sonsal Normal($m^{\iota}, (S^{\iota})^2$) dağılımı elde edildi. Buda gösterir ki, Normal(m,S^2) dağılımı, varyansı bilinen normal gözlemlerden elde edilen eşlenik ailesindendir. Bu sayede sonsala ulaşmak için integral hesabına gerek duyulmaz. İhtiyaç duyulan tek şey parametreleri güncelleme kuralını uygulamaktır (Bolstad, 2004).

5.1.3 Normal ailelerde kullanılan temel güncelleme kuralı

Bayesci çıkarımda $1/Varyans$ kesinlik olarak bilinir. Burada sonsal kesinlik, önsel kesinlik ile gözlem kesinliğinin toplamına eşittir. (5.2) ile verilen S^{ι} göz önüne alınarak,

$$\frac{1}{(S^{\iota})^2} = \left(\frac{\sigma^2 S^2}{\sigma^2 + S^2}\right)^{-1} = \frac{\sigma^2 + S^2}{\sigma^2 S^2} = \frac{1}{S^2} + \frac{1}{\sigma^2} \tag{5.3}$$

kolayca yazılır. Ayrıca yine (5.2) ile verilen m^{ι} nu göz önünde bulundurulmasıyla sonsal dağılımın ortalaması;

$$m^{\iota} = \frac{(\sigma^2 m + S^2 y)}{\sigma^2 + S^2} = \frac{\sigma^2}{\sigma^2 + S^2} \times m + \frac{S^2}{\sigma^2 + S^2} \times y$$

olur. Bu denklem basitleştirilirse;

$$m^{\iota} = \frac{1/S^2}{1/\sigma^2 + 1/S^2} \times m + \frac{1/\sigma^2}{1/\sigma^2 + 1/S^2} \times y \tag{5.4}$$

elde edilir.

Buna göre, elde edilen sonsal ortalama önsel ortalama ve gözlemin ağırlıklı ortalamasıdır.

Güncelleme kuralı düz önsel için de kullanılabilir. Düz önsel sonsuz varyansa sahiptir. Bundan dolayı sıfır kesinliği vardır. Bu durumda sonsal kesinlik gözlem kesinliğine eşittir. Yani,

$$1/\sigma^2 = 0 + 1/\sigma^2$$

dır. Ayrıca sonsal varyans gözlem varyansı σ^2 ye eşittir ve düz önsel iyi bir şekilde tanımlanmış önsel ortalamaya da sahip değildir. Önselin ortalaması herhangi bir reel değer olabilir. Bu durumda sonsal dağılımın ortalaması,

$$m^{\iota} = \frac{0}{1/\sigma^2} \times reel\ sabit + \frac{1/\sigma^2}{1/\sigma^2} \times y$$

$$m^{\iota} = y$$

biçiminde yazılabilir. Yani, düz önsel kullanılarak elde edilen sonsal ortalama, y gözlemine eşit olmuş olur (Bolstad, 2004).

$Y_1, \dots, Y_n$ nin Normal rasgele örneklem olarak alındığı durum,

$Y_1, \dots, Y_n$ rasgele örneklemi $Normal(\mu, \sigma^2)$ olan dağılımdan alınsın. Önsel dağılım ise $Normal(m, S^2)$ olsun. Buna göre;

$$g(\mu) \propto e^{-\frac{1}{2S^2}(\mu - m)^2}$$

biçimindedir. μ yü içermeyen kısım denklemden çıkarılır çünkü önseli herhangi bir sabitle çarpmanın sonsalda etkisi yoktur.

Örneklem ortalamasının ($\bar{Y}$) dağılımı,$normal(\ \mu,\ \frac{\sigma^2}{n})$ dir. Buna göre $\bar{Y}$' nin kesinliği $\frac{n}{\sigma^2}$ dir. Buna göre, rasgele örneklem için kesinliğin bütün gözlemlerin kesinliklerinin toplamı olduğu görülür.

Problemin çözümü tek gözlem verileri için elde edilen sonuçların $\bar{Y}$ için güncellenmesine indirgenir. Sonsal kesinlik, önsel kesinlik ile$\bar{Y}$'nin kesinliğinin toplamıdır. Yani,

$$\frac{1}{(S^{\prime})^2} = \frac{1}{S^2} + \frac{n}{\sigma^2} = \frac{\sigma^2 + nS^2}{\sigma^2 S^2} \tag{5.5}$$

dır. Sonsalın varyansı, sonsalın kesinliğinden elde edilen sonucun çarpmaya göre tersine eşittir. Sonsal ortalama, ağırlıkları sonsal kesinlikle orantılı olan önsel ortalama ve $\bar{Y}$ nun ağırlıklı ortalamasına eşittir,

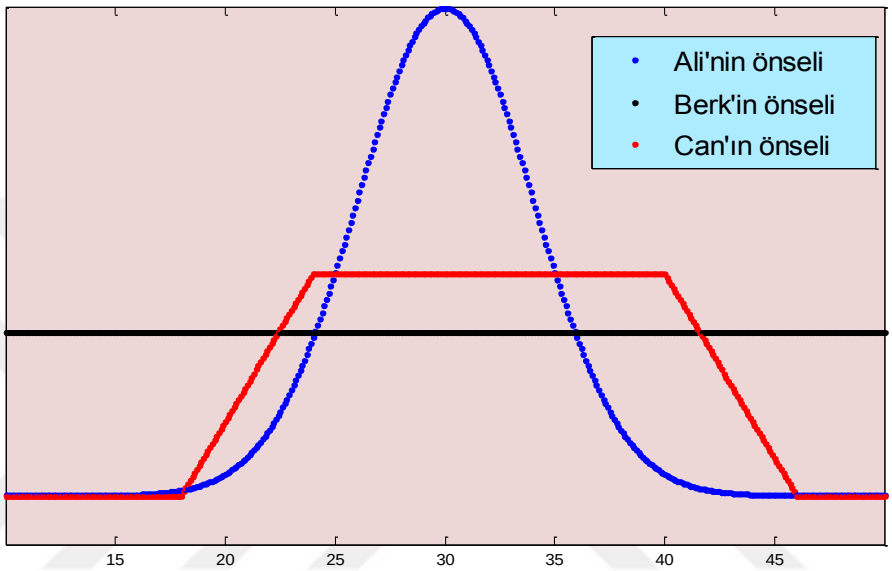
$$m^{\prime} = \frac{1/S^2}{n/\sigma^2 + 1/S^2} \times m + \frac{n/S^2}{n/\sigma^2 + 1/S^2} \times \bar{y} \tag{5.6}$$

Örnek 5.1 Ali, Berk ve Can bir derede yasayan bir yaşındaki gökkuşağı alabalığının ortalama boy uzunluğunu tahmin etmek istemektedirler. Önceki çalışmalardan başka bir derede yaşayan bir yaşındaki gökkuşağı alabalığının boy uzunluğunun normal dağılıma uyduğu ve standart sapmasınında 2cm olduğu biliniyor (Bolstad, 2004).

Ali, bu balık türü için bir yaşındaki balıkların ortalama boy uzunluğunun 30cmcivarında değiştiğini öne sürmüştür. Ayrıca, bir yaşındaki balıkların boyunun 18 ile 42 cm arasında olacağını düşündüğünden önselin standart sapmasını 4 cm olarak alıp önseli $Normal(30, 16)$ olarak belirlemiştir. Berk, hiçbir şey öngöremediğinden düz önsel kullanmıştır. Can ise önselin normal olmadığını iddia etmiş ve önselin yamuk şeklinde olduğunu öne sürmüştür. Can yamuk şeklindeki önselini tanımlamak için (x,y)

düzleminde köşe noktalarının değerlerini vermiştir. Bu değerler (18,0), (24,1), (40,1) ve (46,0) dir.

Buna göre üç önsele ait olasılık yoğunluk fonksiyonlarının grafikleri aşağıdaki gibi verilebilir.



Şekil 5.1 Alınan önsellerin olasılık yoğunluk fonksiyonları

Bu kişiler, deredeki bir yaşındaki gökkuşağı balıklarından $n = 12$ tanesini rasgele örnek olarak almış ve örneklem ortalamasını $\bar{y}$=32 cm olarak hesaplamışlardır.

Bir önceki kısımda elde edilen sonuçlara göre, Ali ve Berk'in sonsal dağılımları eşlenik aile kavramından aşağıdaki gibi elde edilir. Ali için

$$\frac{1}{(S^{\prime})^2} = \frac{1}{S^2} + \frac{n}{\sigma^2} = \frac{\sigma^2 + nS^2}{\sigma^2 S^2} = 3.0625$$

$$S^{\prime} = 0.5714$$

Denklem (5.6) kullanılarak $m^{\prime} = 31.96$ olarak bulunur. Yani sonsal dağılım ortalaması 31.69 ve standart sapması 0.5714 olan Normal dağılımdır.

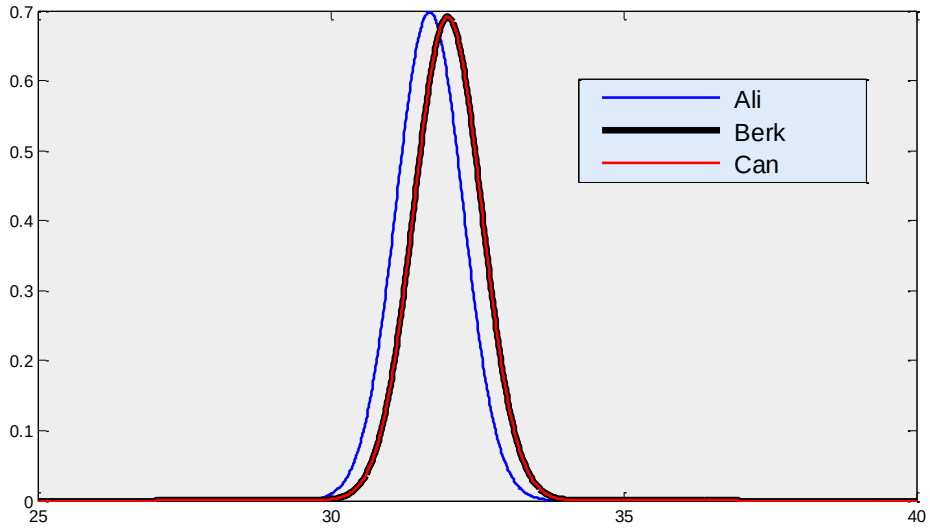
Berk düz önsel kullandığı için; sonsal dağılımın varyansını

$$(S^{\prime})^2 = \frac{12}{2^2} = 0.333$$

olarak hesaplamıştır. Sonsal dağılımın ortalaması da örneklem ortalaması olan $m^{\prime} = 32$' ye eşittir. Buna göre Berk'in sonsalı ortalaması 32 standart sapması 0.5774 olan normal dağılımdır.

Can sürekli önsel kullanmıştır. Can'ın sonsal dağılımını elde etmek için Bayes teoremine göre (4.1) eşitliğini kullanmalıdır. Buna göre elde edeceği sonsal dağılım düz önsel kullanımıyla elde edilen sonsal dağılıma çok yakın çıkacaktır.

Sonuç olarak, Ali bilgi veren önsel olarak Normal dağılımı kullanmıştır. Berk önsel bilgiye fazla önem vermeyen düz önseli kullanmıştır. Can ise düz önsele çok yakın olan yamuk şeklinde bir önsel kullanmıştır. Kullanılan önsellere göre elde edilen sonsal dağılımların olasılık yoğunluk fonksiyonlarının grafikleri aşağıda verilmiştir.



Şekil 5.2 Sonsal dağılımların olasılık yoğunluk fonksiyonları

Farklı önseller kullanılmasına rağmen elde edilen sonsalların oldukça benzer olduğu çizilen grafikle de görülmektedir.

5.2 μ İçin Bayesci Güvenilir Aralık

Sonsal dağılım, $g(\mu|y_1, \dots, y_n)$ veriler gözlendikten sonra μ için elde edilen olasılık dağılımıdır. Bu sonsal dağılım kullanılarak μ parametresi için bir güvenilir aralık belirlenebilir.

Varyans bilindiğinde

$Y_1, \dots, Y_n$, rasgele örneklemi Normal(μ, σ^2) dağılımından ise, $\bar{Y}$ nin örnekleme dağılımı Normal($\mu, \sigma^2/n$) alındığındaki ortalamaya eşittir. Dağılımın varyansı ise, tek gözlem alındığındaki varyansın örneklem genişliğine bölünmesiyle elde edilir. Önsel "düz" veya Normal (μ, σ^2) kullanıldığında $\bar{y}$ verildiğinde μ nun sonsal dağılımı normaldir, $N(m', (S')^2)$. Bu güncelleme aşağıdaki kurallara göre yaplır:

1. kesinlik, varyansın çarpmaya göre tersidir.

2. Sonsalın kesinliği, önselin kesinliği ile örnekleme ortalamasının kesinliğinin toplamına eşittir.
3. Sonsalın ortalaması, önsel ortalaması ile örnekleme ortalamasının ağırlıklı toplamıdır; ağırlıklar, kesinliklerin sonsal kesinliğe oranıdır.

μ nün $(1 - \alpha) \times 100\%$ Bayesci güvenilir aralığı;

$$m' \pm Z_{\alpha/2} \times S' \tag{5.7}$$

olarak elde edilir. Yani sonsal ortalama $(\mp)z - değeri$ ile sonsalın standart sapmasının çarpımıdır. $z - değeri$ standart normal dağılım tablosundan bulunur. Gerçek ortalama μ nün, güvenilir aralığın dışında değer alması sonsal olasılığı α dır. Sonsal dağılım normal olduğu için bu simetrik bir dağılımdır. Bu nedenle elde edilen güvenilir aralık en kısa aralıktır ve eşit kuyruk olasılıklarına sahiptir (Bolstad, 2004).

Varyans bilinmediğinde

Varyans bilinmiyorsa, kesinlikte bilinmez, dolayısıyla doğrudan güncelleme kuralı kullanılamaz. Yapılacak şey, örneklem varyansını hesaplamaktır. Örneklem varyansı,

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

ile hesaplanır. $(S')^2$ ve m'nü bulmak için (5.5) ve (5.6) denklemleri kullanılır.

Denklemlerde σ^2 bilinmeyen varyans yerine örnekleme varyansı $\hat{\sigma}^2$ kullanılır. Burada extra bir belirsizlik vardır bu da σ^2'yi tahmin etmedeki belirsizliktir. Bu eklenen belirsizlikte hesaba katılınca güvenilir aralık genişletilmelidir. Bunun için Standart normal tablo yerine student t dağılımına ilişkin tablo kullanılır. Bayesci güvenilir aralık,

$$m' \pm t_{\alpha/2,n-1} \times S' \tag{5.8}$$

ile elde edilir.

Buna göre bahsedilen örnekte elde edilen sonuçlara göre her bir araştırıcı için elde edilen sonsal dağılımlar ve bunlar kullanılarak elde edilen Bayesci güvenilir aralıklar,

Çizelge 5.1 Bayesci güvenilir aralık sınırları

İsim	**Sonsal**	**Güvenilir aralık**
Ali	N(31.69,0.3265)	(30.4323, 32.9477)
Berk	N(32,0.333)	(30.7299, 33.2701)
Can	Sayısal	(30.8,33.1)

biçiminde elde edilir.

5.2.1 μ için Tek Yanlı Bayes Hipotez Testi

Verileri gözlemledikten sonra, parametre hakkındaki tüm bilgi sonsal dağılım kullanılarak elde edilir.μdeğerinin verilen herhangi bir μ_0 değerinden büyük olduğu iddia edildiğinde iddiayı test etmek için hipotez kurulması gerekir. μ_0 genellikle bir önceki deneyimlerden ileri gelir.

$$H_0: \mu \leq \mu_0$$
$$H_1: \mu > \mu_0$$

Yokluk hipotezinin sonsal olasılığı ise;

$$\mathrm{P}(H_0: \mu \leq \mu_0 | y_1, \dots, y_n) = \int_{-\infty}^{\mu_0} g(\mu | y_1, \dots, y_n) d\mu$$

biçiminde hesaplanır. Sonsal dağılım, Normal($(m^{\prime}, S^{\prime\,2})$ olduğundan bu olasılık standart normal tablosu kullanılarak hesaplanır. Yani,

$$\mathrm{P}(H_0: \mu \leq \mu_0 | y_1, \ldots, y_n) = P\left(\frac{\mu - m^{\prime}}{S^{\prime}} \leq \frac{\mu_0 - m^{\prime}}{S^{\prime}}\right) \tag{5.9}$$

$$= \mathrm{P}(\mathrm{Z} \leq \frac{\mu_0 - m^{\prime}}{S^{\prime}})$$

ile gösterilebilir. Burada Z, standart normal rastgele değişkenidir. Eğer elde edilen bu olasılık, H_0 hipotezinin sonsal olasılık değeri, H_1 hipotezinin sonsal olasılık değerinden küçük ise H_0 aleyhine karar verilir ve $\mu > \mu_0$ olduğu söylenir.

Yukarıda verilen örneğe dönülecek olursa, Ali, Berk ve Can bir derede yasayan bir yaşındaki gökkuşağı alabalığının boyunun ortalama ne kadar olduğunu tahmin etmek istemişlerdi. Önceki çalışmalardan başka bir derede yaşayan bir yaşındaki gökkuşağı alabalığının ortalama boy uzunluğu 31cm olarak bilinmektedir. Araştırmacılar bu deredeki balıkların ortalama boy uzunluğunun bu değerden büyük olup olmadığını test etmek istemektedirler. Buna göre,

$$H_0: \mu \leq 31$$
$$H_1: \mu > 31$$

tek yanlı hipotezini kurmuşlar ve α =0.05 olarak belirlemişlerdir.
Ali ve Berk için sonsal dağılım normal($m^{\prime}, S^{\prime}$) ile gösterilirse bu sonsal dağılımlar altında H_0'ın olasılığı

$$\mathrm{P}(H_0: \mu \leq \mu_0 | y_1, \ldots, y_n) = P(\frac{\mu - m^{\prime}}{S^{\prime}} \leq \frac{\mu_0 - m^{\prime}}{S^{\prime}})$$

$$= \mathrm{P}\left(\mathrm{Z} \leq \frac{\mu_0 - m^{\prime}}{S^{\prime}}\right)$$

ile bulunur. Matlab programı kullanılarak bu olasılıklar,

normcdf(31, 31.96, 0.5714) = 0.0465
normcdf(31, 32, 0.5774) = 0.0416

biçiminde hesaplanır. Hesaplanan H_0 hipotezinin sonsal olasılık değeri 0.0465 ve 0.0416, H_1 hipotezinin sonsal olasılık değerinden küçük olduğundan her iki araştırmacı da H_0 aleyhine karar verir.

Can için sonsal dağılım bir önceki kısımda nümerik olarak hesaplanmıştı. Yokluk hipotezinin sonsal olasılığı;

$$\mathrm{P}(H_0: \mu \leq \mu_0 | y_1, \dots, y_n) = \int_{-\infty}^{\mu_0} g(\mu \leq \mu_0 | y_1, \dots, y_n) d\mu$$

biçiminde yine nümerik olarak hesaplanabilir. Bunun için Matlab programı yardımı ile 0.0489 değeri bulunmuştur. Benzer şekilde hesaplanan H_0 hipotezinin sonsal olasılık değeri, H_1 hipotezinin sonsal olasılık değerinden küçük olduğu için Can da H_0 hipotezi aleyhine kararını verir.

Buna göre elde edilen sonuçlar aşağıdaki tabloda özetlenebilir.

Çizelge 5.2 Hipotez testi sonuçları

Araştırmacı	**Sonsal**	$\mathbf{P(\mu \leq 31 | y_1, \dots, y_n)}$	**Sonuç**
Ali	$N(31.96, 0.5714^2)$	$P(Z \leq \frac{31-31.96}{0.5714}) = 0.0465$	H_1 daha olasıdır.
Berk	$N(32, 0.5774^2)$	$P(Z \leq \frac{31-32}{0.5774}) = 0.0416$	H_1 daha olasıdır.
Can	Sayısal	$\int_{-\infty}^{31} g(\mu \leq \mu_0 | y_1, \dots, y_n) d\mu = 0.0489$	H_1 daha olasıdır.

Aynı örnek klasik yaklaşımla çözülmek istenseydi, eldeki örnekleme göre $\bar{y} = 32$ olarak bulunmuştu. Buna göre $p - değeri$,

$$\text{p} - \text{değeri} = P(Z > \frac{\bar{y} - 31}{\sigma/\sqrt{n}})$$
$$= P(Z > 1.732)$$

olarak bulunur. Bu değer de Matlab yardımıyla,

1-normcdf(1.732, 0,1) = 0.0416

biçiminde hesaplatılmıştır. Hesaplanan $p - değeri$ α değerinden küçük olduğu için klasik yaklaşımla da H_0 hipotezinin reddedileceği sonucuna ulaşılır.

5.2.2 μ için iki yanlı Bayesci hipotez testi

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Eğer önsel dağılım olarak sürekli önsel kullanılırsa, iki yanlı hipotez testi için yokluk hipotezinin sonsal olasılığı hesaplanamaz. Sürekli rasgele değişkenin tek noktada değer alma olasılığı sıfır olduğu için, yokluk hipotezinin sonsal olasılığı sıfıra eşit olacaktır. Bu durumda, bir önceki kısımda da ifade edildiği gibi yokluk hipotezinin sonsal olasılığı hesaplanarak, hipotez test edilemez. Bunun yerine, μ için sonsal dağılım kullanarak (1- α)×100% lık güvenilir aralık hesaplanır. Eğer μ_0 elde edilen bu güvenilir aralığın içindeyse, olası değer olarak μ_0 hala güvenilirdir. Bu durumda yokluk hipotezilehine karar verilir. Eğer μ_0 güven aralığının dışındaysa, olası değer olarak μ_0 güvenilir değildir. Yokluk hipotezinin aleyhine karar verilir (Bolstad, 2004).

6. İKİ KİTLE ORTALAMASI FARKI İÇİN BAYESCİ SONUÇ ÇIKARIMI

6.1 Bağımsız Örneklemler İçin Sonuç Çıkarımı

Uygulanan bir tedavinin kuzuların büyüme hızında etkili olup olmadığı araştırılmak istensin. Bu amaçla ilk olarak, sürüdeki her bir kuzu rasgele tedavi grubu ve kontrol grubuna atansın. Sürüdeki kuzular deneme ve kontrol gurubuna rasgele atandığı için her iki gruptaki varyans ve ortalama eşittir. Her iki gruptaki kuzuların büyüme oranları ortalama ve varyansları aynı olan normal dağılıma sahip olsun.

Buna göre, tedavi grubundaki kuzuların ortalama büyüme hızı μ_1, temel büyüme hızının ortalaması ile tedavinin kuzulara etkisinin toplamı biçiminde, kontrol grubundaki kuzuların ortalama büyüme hızı μ_2 ise temel büyüme hızının ortalaması biçiminde olacaktır.

Tedavinin kuzuların gelişimine olumlu etkisi olduğu düşünüldüğünde ve bu etki sabit olarak varsayıldığında iki grup için varyanslar eşit olacaktır. Çünkü etki sabit olarak kabul edildiğinden rasgele değişkene bir sabit eklenmesi varyansı değiştirmeyecektir. Buna katkı modeli denir. Eğer, etkinin kuzudan kuzuya değiştiği yani rasgele olduğu varsayılırsa iki grubun varyansı farklı olacaktır. Buna katkısal olmayan model denir.

Eğer, tedavi etkiliyse, deneme grubundaki kuzuların büyüme hızlarının ortalaması μ_1, kontrol grubundaki kuzuların büyüme hızlarının ortalaması μ_2'den büyük olacaktır. Bu bölümde hem katkı hem de katkısal olmayan modelle μ_1-μ_2 ortalamaları arasındaki fark Bayesci yöntemle incelenecektir (Bolstad, 2004).

6.1.1 Varyansların eşit olduğu durum

Yukarıda bahsedildiği gibi tedavinin etkisi uygulanan bütün deneklerde aynı olabilir. Bu durum göz önüne alındığında,

Tedavi verilen birimin gözlenen değeri =birim için ortalama+ sabit deneme etkisi

biçiminde olup eklenen sabitler varyansı değiştirmez kuralına göre, tedavi grubunun varyansı, kontrol grubunun varyansına eşittir. Bu, katkı modelini ortaya koyar (Bolstad, 2004).

6.1.1.1 Varyanslar biliniyor

İki örneklem birbirinden bağımsız olduğu durumda, her iki ortalama içinde bağımsız önsel kullanılabilir. Bu önseller $N(m_1, S_1^2)$ ve $N(m_2, S_2^2)$ olsun. Ya da ortalamaların biri veya ikisi için düz önsel kullanılabilir. Çünkü önseller bağımsız ve örneklemler bağımsız dolayısıyla sonsallarda bağımsız olur.

Sonsal dağılımlar ise;

$$\mu_1 | y_{11}, \dots, y_{n_{11}} \sim N(m_1', (S_1^{\iota})^2)$$

ve

$$\mu_2 | y_{12}, \dots, y_{n_{22}} \sim N(\mathrm{m}_2', (\mathrm{S}_2^{\iota})^2)$$

biçiminde olur. Burada, m_1', $(\mathrm{S}_1^{\iota})^2$, $\mathrm{m}_2{}^{\iota}$ ve $(\mathrm{S}_2^{\iota})^2$ parametreleri (5.5) ve (5.6) ile verilen temel güncelleme kuralı uygulanarak bulunur.

$\mu_1 | y_{11}, \dots, y_{n11}$ ve $\mu_2 | y_{12}, \dots, y_{n22}$ birbirinden bağımsız olduğu için bağımsız rasgele değişkenlerin farkı için sonsal dağılım, ortalaması ve varyansı

$$m'_d = m'_1 - m'_2 \qquad ve \qquad (S'_d)^2 = (S'_1)^2 + (S'_2)^2$$

olan Normal dağılım olur. Yani, $\mu_d|y_{11}, \dots, y_{n11}, y_{12}, \dots, y_{n22} \sim \mathrm{N}(m'_d, (S'_d)^2)$ biçiminde ifade edilir. Bu durumda, $\mu_1 - \mu_2$ ortalamaları arasındaki farkla ilgili sonuç çıkarımı yapmak için bu sonsal dağılım kullanılır (Bolstad, 2004).

Ortalamalar arasındaki fark için güvenilir aralık,

Sonsal dağılım $\mathrm{N}(m', (S')^2)$ olduğunda $(1 - \alpha) \times 100\%$ Bayesci güvenilir aralık $sonsalortalama \pm kritik\ değer \times sonsal\ standartsapma$ formülü ile oluşturulabilir. Gözlemlerin standart sapması ya da varyansı bilindiğinde, kritik değer standart normal tablodan bulunur. $\mu_d = \mu_1 - \mu_2$ nin (1- α)×100% Bayesci güvenilir aralığı,

$$\mathrm{m}'_d \pm z_{\frac{\alpha}{2}} \times \mathrm{S}'_d$$

dir. Ayrıca,

$$m'_1 - m'_2 \pm z_{\frac{\alpha}{2}} \times \sqrt{(S'_1)^2 + (S'_2)^2}$$

biçiminde de yazılabilir (Bolstad, 2004).

Örnek 6.1 Lastik üretmede iki yeni yöntem önerilmektedir. Hangisinin daha üstün olduğunu belirlemek için birinci yöntemle 10, ikinci yöntemle 8 lastik üretiliyor. Birinci grup ve ikinci grup lastikler bir yol testine tabi tutuluyor. Daha önceki tecrübelerden, yol testinde ortalama ömür lastik çeşidine ve varyans bölgeye bağlı olmak üzere normal dağılımlı olduğu biliniyor. Bu bölgede test edilen lastiklerin ortalama ömrünün, standart sapması $\sigma = 6000$ km olan normal dağılıma sahip olduğu biliniyor.

Çizelge 6.1 100 km birimiyle lastik ömürleri

1. grup lastik ömrü× 100	**2. grup lastik ömrü × 100**
61.1	62.2
58.2	56.6
62.3	66.4
64	56.2
59.7	57.4
66.2	58.4
57.8	57.6
61.4	65.4
62.2	
63.6	

Buna göre iki lastik türü ortalaması arasındaki fark için güvenilir aralık elde edilsin.

Örneklemlerden,

$$\bar{x}_1 = 6165$$
$$\bar{x}_2 = 6002.5$$

olarak elde edilir.

$\sigma = 6000$ olarak biliniyor. Burada önsel dağılım olarak Normal dağılım alınarak Bayesci sonuç çıkarımı yapılacaktır. Her iki lastik türü için seçilen önsel Normal dağılım ortalaması $m = 5900$ ve varyansı $S^2 = 300^2$ olarak alınmıştır.

Buna göre ilk olarak her iki lastik türü içinde sonsal dağılımlar normal dağılım olacaktır ve ortalama ve varyansları aşağıdaki gibi hesaplanacaktır.

$$\frac{1}{(s_1^t)^2} = \frac{1}{300^2} + \frac{10}{6000^2}$$

$$(s_1^t)^2 = 87804.88$$

$$m_1^l = \frac{\frac{1}{300^2}}{\frac{1}{300^2}+\frac{10}{6000^2}} \times 5900 + \frac{\frac{10}{6000^2}}{\frac{1}{300^2}+\frac{10}{6000^2}} \times (6165) = 5906.46$$

ve

$$\frac{1}{(s_2^l)^2} = \frac{1}{300^2} + \frac{8}{6000^2}$$

$$(s_2^l)^2 = 88235.29$$

$$m_2^l = \frac{\frac{1}{300^2}}{\frac{1}{300^2}+\frac{8}{6000^2}} \times 5900 + \frac{\frac{8}{6000^2}}{\frac{1}{300^2}+\frac{8}{6000^2}} \times 6002.5 = 5901.96$$

İki ortalama farkı için sonsal dağılım $\mu_d = \mu_1$-μ_2 olan normal$(m_d^l, (S_d')^2)$ olacaktır. Buna göre,

$$m_d^l = \mu_1 - \mu_2 = 5906.46 - 5901.96 = 4.5$$

$$(S_d')^2 = 87804.88 + 88235.29 = 176040.17$$

$$S_d' = 419.57$$

$\mu_d = \mu_1$-μ_2 için %95 lik Bayesci güvenilir aralık;

$$4.51 \pm 1.96 \times 419.57 = (-817.85, 826.87)$$

biçiminde elde edilir.

6.1.1.2 Tek yanlı Bayes hipotez testi

Eğer iki kitle ortalaması için, birinci kitle ortalaması μ_1 in, ikinci kitle ortalaması μ_2 den daha büyük olup olmadığı test edilmek istenirse, tek yanlı hipotez testi kullanılır. Buna göre hipotez testi

$$H_0: \mu_d \leq 0$$
$$H_1: \mu_d > 0$$

biçiminde verilir.

μ_d=$\mu_1 - \mu_2$ iki ortalama arasındaki farktır. Yokluk hipotezinin sonsal olasılığı P($\mu_d \leq 0$|gözlem), iki örneklemden de ($y_{11}, \dots, y_{n11}$ ve $y_{12}, \dots, y_{n22}$) gözlem verileri içerir.

$$P(\mu_d \leq 0|\text{veri}) = \text{P}(\frac{\mu_d - m_d^{\iota}}{S_d'} \leq \frac{0 - m_d^{\iota}}{S_d'})$$

$$= \text{P}(Z \leq \frac{0 - m_d^{\iota}}{S_d'})$$

Z, standart normal dağılıma sahip rasgele değişkendir. Bu olasılık standart normal dağılım tablosu kullanılarak bulunur. Eğer bu değer H_1 için hesaplanan olasılık değerinden küçük ise, yokluk hipotezinin daha olası olmadığı söylenir ve H_0 aleyhine karar verilir ve $\mu_d > 0$ olduğu söylenir. Buna göre, α- anlam düzeyinde μ_1'in μ_2'den daha büyük olduğu söylenir.

Bir önceki örnek gözönüne alınarak,

$$H_0: \mu_d \leq 0$$
$$H_1: \mu_d > 0$$

tek yanlı hipotez test edilsin.

$$P(\mu_d \leq 0|\text{veri}) = \text{P}(\frac{\mu_d - m_d^l}{S_d'} \leq \frac{0 - m_d^l}{S_d'})$$

$$= \text{P}(Z \leq \frac{0 - 4.5}{419.57})$$

$$= \text{P}(Z \leq -0.0107)$$

$$= normcdf(-0.0107,0,1)$$

$$= 0.4957$$

H_0 hipotezinin sonsal olasılığı 0.4957 olarak hesaplanmıştır. Buna göre, H_1 hipotezinin sonsal olasılığı 1-0.4957=0.5043 olup H_1 hipotezinin sonsal olasılığı ile H_0 hipotezinin sonsal olasılığı birbirine oldukça yakındır. Burada, H_1lehine karar vermekle birlikte sonuçlar birbirine çok yakın olduğu için H_0' a karşı delilin az olduğu söylenebilir. Yani, iki ortalama arasındaki farkın sıfırdan büyük olduğu yani birinci lastik türünün ortalama ömrünün daha büyük olduğu söylenebilir. Bu sonuç çıkarımı seçilen önseller eşit olarak alınıp elde edildi, önseller değiştirilerek tekrar sonuç çıkarımı yapıldığında sonuçlar ona göre değişip yorum yapılacaktır.

Bayes faktörüne göre yorum yapılmak istendiğinde, önsellerin birbirine eşit alındığı durumda;

$$Bayes\ Faktörü\ \ = BF_{01} = Sonsal\ Odds = \frac{P(\theta \epsilon \Theta_0|T = t)}{P(\theta \epsilon \Theta_1|T = t)}$$

$$BF_{01} = \frac{\text{P}(\mu_d \leq 0|\bar{X}_{gözlenen})}{\text{P}(\mu_d > 0|\bar{X}_{gözlenen})} = \frac{0.4957}{1 - 0.4957} = 0.983$$

olarak elde edilir. Thumb kuralına göre bu değer, $10^{-1/2} \leq BF_{01} < 1$ aralığına düştüğü için H_0 yokluk hipotezine karşı az delil olduğu tekrarlanmış olur.

6.1.1.3 İki yanlı Bayes hipotez testi

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

hipotez testi göz önüne alınsın. Daha önceki bölümlerde de ifade edildiği gibi iki yanlı hipotez testi, yokluk hipotezinin sonsal olasılığını hesaplayarak test edilemez. Sürekli önsel kullanıldığında, sürekli sonsal elde edilir ve sürekli rasgele değişkenin tek noktada aldığı olasılık sıfırdır. Bunun yerine μ_d için güvenilir aralık bulunarak kullanılır. Eğer sıfır aralığın içindeyse, yokluk hipotezi reddedilmez ve ortalamalar arasındaki fark için sıfır güvenilir bir değer olur. Ancak sıfır aralığın dışındaysa, sıfır artık α anlam seviyesinde güvenilir değildir.

Bir önceki örnek için iki yanlı hipotez testi yapılacak olursa hipotezler;

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

biçiminde olup $\mu_d = \mu_1 - \mu_2$için %95 lik Bayesci güvenilir aralık

$$4.51 \pm 1.96 \times 419.57 = (-817.85, 826.87)$$

olarak hesaplanmıştı. Bu aralık sıfırı içerdiği için yokluk hipotezi lehine karar verilir. Yani, iki ölçüm grubunun ortalaması arasında fark yoktur (Bolstad, 2004).

6.1.1.4 Varyanslar bilinmiyor

Önsel olarak düz önselin kullanıldığı durum

μ_1 ve μ_2 için bağımsız düz önseller kullanılsın. Her iki örneklemde normal dağılıma sahip kitlelerden alındığı için, $\bar{Y}_1$ örneklem ortalamasının dağılımı da normal olur.

Ortalaması kitle ortalaması ile eşit, μ_1 olur ve kitle varyansı bilinmediği için varyans yerine örneklemden tahmini ${S_1}^2/\,n_1$ alınır. Aynı şekilde, $\bar{Y}_2$ örneklem ortalamasının dağılımı da normal ortalaması μ_2 ve varyansı yine örneklemden tahmini olacaktır. Bu durumda μ_1 ve μ_2 nin sonsal dağılımı sırasıyla $Normal\left(\bar{y}_1, {S_1}^2/\,n_1\right)$, $Normal\left(\bar{y}_2, {S_2}^2/\,n_2\right)$ olur. Buradan, iki kitle ortalaması arasındaki fark için sonsal dağılım varyanslar eşit olduğu için varyansı,

$$(\mathrm{S}_d^{\iota})^2 = \frac{(n_1 - 1)(S_1^{\iota})^2 + (n_2 - 1)(S_2^{\iota})^2}{n_1 + n_2 - 1}$$

ve ortalaması

$$\mathrm{m}_d^{\iota} = \bar{y}_1 - \bar{y}_2$$

olan Normal dağılım olur. Yani, μ_1-μ_2 kitle ortalamaları farkı için sonsal dağılım Normal $(\bar{y}_1 - \bar{y}_2,\ (\mathrm{S}_d^{\iota})^2)$ olarak bulunur.

Düz önsel kullanıldığında ortalamalar arasındaki fark için güvenilir aralık

σ^2 varyans bilindiğinde Bayesci güvenilir aralık;

$$\bar{y}_1\ - \bar{y}_2 \pm\ \mathrm{z}_{\frac{\alpha}{2}} \times \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

biçiminde ifade edilmişti. Ancak σ^2 bilinmiyorsa, yukarıda da ifade edildiği gibi bu değer verilerden tahmin edilmelidir. Bu değer her iki örneklem kullanılarak tahmin edilir. Bunun en iyi yolu ortak varyans tahmini olan;

$${S_p}^2 = \frac{\sum_{i=1}^{n_1}(y_{i1} - \bar{y}_1)^2 + \sum_{j=1}^{n_2}(y_{j2} - \bar{y}_2)^2}{n_1 + n_2 - 2}$$

kullanmaktır. Burada bilinmeyen gerçek σ^2 varyansı yerine ${S_p}^2$ tahmini değeri kullanılır. Bu durumda tablo değeri olarak, Student t tablosunda $n_1 + n_2 - 2$ serbestlik derecesine karşılık gelen kritik değer alınır. Buna göre; yaklaşık (1- α)×100% Bayesci güvenilir aralık $\mu_1 - \mu_2$ için;

$$\bar{y}_1 - \bar{y}_2 \pm t_{\frac{\alpha}{2}} \times S_p \sqrt{\frac{1}{\mathrm{n1}} + \frac{1}{\mathrm{n2}}}$$

olarak yazılır.

6.1.2 Varyansların eşit olmadığı durum

6.1.2.1 Varyanslar biliniyor

Bu durumda katkısal olmayan model kullanılır. Her değişken için uygulanan çalışmanın sonucu farklıdır, varyanslar bilinir fakat eşit değildir. Buna göre, $Y_{11}, \dots, Y_{n11}$ örneklemi $\mathrm{N}(\mu_1, \sigma_1^2)$ dağılımından rasgele örneklem ve $Y_{12}, \dots, Y_{n22}$ örneklemi $\mathrm{N}(\mu_2, \sigma_2^2)$ dağılımından rasgele örneklem olsun. İki rasgele örneklem birbirinden bağımsızdır.

μ_1 ve μ_2 için bağımsız rasgele önseller kullanılır. Bu önsel normal ya da düz önsel olabilir. Örneklem bağımsız, önsel bağımsız olduğundan her sonsal da diğerinden bağımsız olarak bulunabilir. Buna göre daha önce denklem (5.5) ve (5.6) verilen güncelleme kuralı kullanılarak sonuca ulaşılabilir. Buna göre parametreler için sonsal dağılım;

$$\mu_1 \big| y_{11}, \dots, y_{n_1 1} \sim \mathrm{N}(m_1', (S_1')^2)$$

ve

$$\mu_2 \big| y_{12}, \dots, y_{n_2 2} \sim \mathrm{N}(m_2', (S_2')^2)$$

olur. $\mu_d = \mu_1 - \mu_2$ için sonsal dağılım, ortalaması sonsal ortalama farkına eşit ve varyansı sonsal varyanslar toplamına eşit olan Normal dağılımdır. Yani,

$$\mu_d \left| y_{11}, \dots, y_{n_1 1}, y_{12}, \dots, y_{n_2 2} \sim \mathrm{N}(\mathrm{m}_d^{\prime}, (\mathrm{S}_d^{\prime})^2) \right.$$

biçiminde olmak üzere beklenen değer ve varyans

$$\mathrm{m}_d^{\prime} = m_1^{\prime} - m_2^{\prime},\ (\mathrm{S}_d^{\prime})^2 = (S_1^{\prime})^2 + (S_2^{\prime})^2$$

ile hesaplanır (Bolstad, 2004).

Ortalamalar arasındaki fark için Bayesci güvenilir aralık

$\mu_d = \mu_1 - \mu_2$ nin (1- α)×100% lık Bayesci güvenilir aralığında

$$\mathrm{m}_d^{\prime} \pm z_{\frac{\alpha}{2}} \times \mathrm{S}_d^{\prime}$$

dir. Yani,

$$m_1^{\prime} - m_2^{\prime} \pm z_{\frac{\alpha}{2}} \times \sqrt{(S_1^{\prime})^2 + (S_2^{\prime})^2}$$

biçiminde de yazılabilir.

6.1.2.2 Varyanslar bilinmiyor

Varyanslar eşit değil ve bilinmiyor ise, her bir örnek varyansı verilerden tahmin edilmelidir.

$$\hat{\sigma}_1^{\,2} = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (y_{i1} - \bar{y}_1)^2 \qquad \text{ve} \qquad \hat{\sigma}_2^{\,2} = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_{i2} - \bar{y}_2)^2$$

Bu tahminler gerçek varyans değerleri yerine basit güncelleme formülünde kullanıldığında ayrıca bir belirsizlik katarlar. Bu durumda kritik değeri bulmak için student t tablosu kullanılır. Ancak serbestlik derecesinin ne olacağı açık değildir. Satterthwaite

$$\mathrm{sd} = \frac{(\frac{\hat{\sigma}_1^{\,2}}{n_1} + \frac{\hat{\sigma}_2^{\,2}}{n_2})^2}{\frac{(\hat{\sigma}_1^{\,2}/n_1)^2}{n_1+1} + \frac{(\hat{\sigma}_2^{\,2}/n_2)^2}{n_2+1}}$$

düzeltilmiş serbestlik derecesini önerir. Serbestlik derecesi tam bir değer olacağından elde edilen bu değer en yakın tam sayıya yuvarlanarak kullanılır (Bolstad, 2004).

Ortalamalar arasındaki fark için güvenilir aralık

Bilinmeyen varyansın yerine varyansın örneklemden tahmini (5.5) ve (5.6)' daki denklemde kullanıldığında, $\mu_d = \mu_1 - \mu_2$ için yaklaşık (1-α)×100% güvenilir aralık;

$$m_1^l - m_2^l \pm t_{\frac{\alpha}{2}} \times \sqrt{(S_1')^2 + (S_2')^2}$$

biçiminde verilir. Serbestlik derecesi, Satterhwaites'in formülü kullanılarak bulunur. μ_1 ve μ_2 için bağımsız düz önsel kullanılırsa güvenilir aralık;

$$m_1^l - m_2^l \pm t_{\frac{\alpha}{2}} \times \sqrt{\frac{(\hat{\sigma}_1')^2}{n_1} + \frac{(\hat{\sigma}_2')^2}{n_2}}$$

olarak elde edilir (Bolstad, 2004).

Bayesci Hipotez Testi

$$H_0: \mu_1 - \mu_2 \leq 0$$
$$H_1: \mu_1 - \mu_2 > 0$$

$$P(\mu_d \leq 0|\text{veri}) = \text{P}(\frac{\mu_d - m_d^{\iota}}{S_d'} \leq \frac{0 - m_d^{\iota}}{S_d'})$$
$$= \text{P}(Z \leq \frac{0 - m_d^{\iota}}{S_d'})$$

Bilinmeyen varyans yerine tahmin edilen varyans kullanılır. Yokluk hipotezinin sonsal olasılığı bulunur. Eğer bu değer H_1 için hesaplanan olasılık değerinden küçükse yokluk hipotezi aleyhine karar verilir yani $\mu_1 > \mu_2$ olduğu söylenir. Aksi halde yokluk hipotezi lehine karar verilir. Satterhwaites yaklaşımı ile elde edilen serbestlik dereceli t −dağılım tablosu kullanılarak hesaplanır.

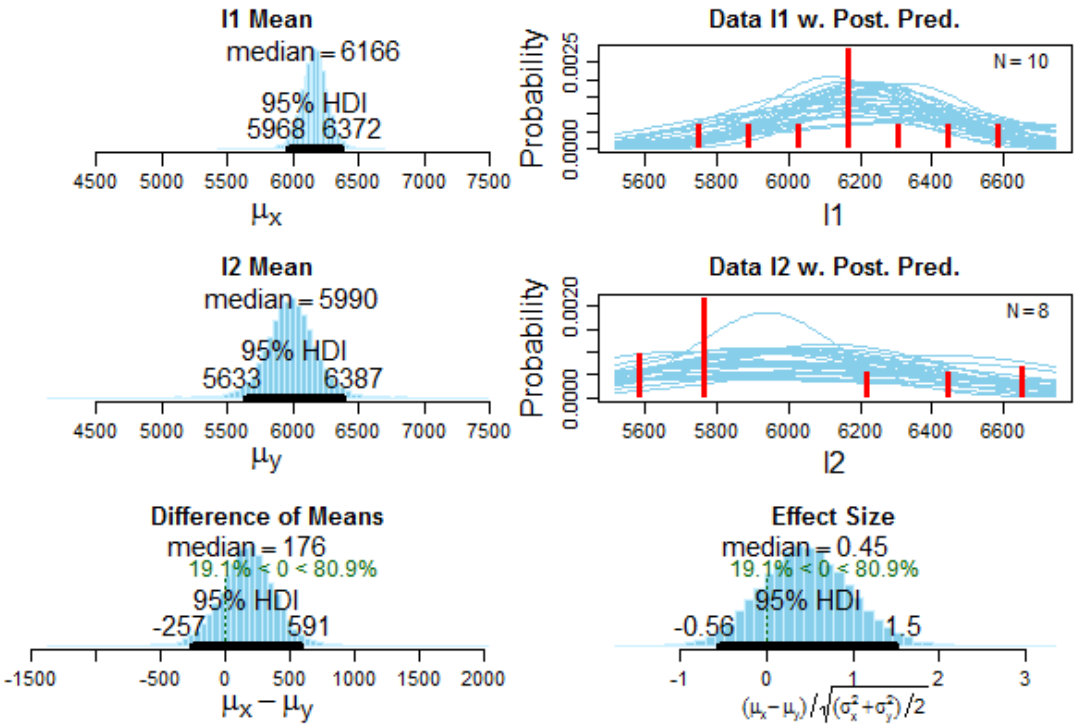
Örnek 6.1 için, frekansçı yaklaşımla varyanslar bilinmiyor iken; varyanslar birbirine eşit olarak alındığında R programında %95 lik güven aralığı (-172.5943,497.5943) biçiminde, varyanslar birbirinden farklı alındığı durumda aralık (-201.3154, 526.3154) biçiminde elde edilmiştir. Klasik bakış açısıyla, her iki aralıkta sıfır değerini içerdiği için $\alpha = 0.05$ anlam düzeyinde iki kitle ortalaması arasında fark yoktur sonucuna ulaşılır (EK 2)

Örnek 6.1 ile verilen lastik örneği için,

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

hipotezi test edilmek istensin. R programı kullanılarak Bayesci sonuç çıkarımı yapılmak istendiğinde, bayes.t.test komutu yardımıyla %95 lik güvenilir aralık (-257, 591) olarak elde edilmiştir. Buna göre iki kitle ortalaması arasında fark olmadığı Bayesci sonuç çıkarımına göre söylenebilir(EK 3)

R programı, Bayesci sonuç çıkarımı yaparken MCMC yöntemini kullanarak sonsal dağı lımdan örneklem çeker. Buna göre program çalıştırıldığında iki kitle ortalaması arasında ki fark için Bayesci sonuç çıkarımında aşağıdaki grafikler elde edilir. Birinci satırdaki ilk grafik ile birinci lastik türünün ortalaması için elde dilen sonsal dağılımdan çekilen ö rnekleme ilişkin histogram; ikinci satırdaki ilk grafik ile ikinci lastik türü ortalaması için elde edilen sonsal dağılımdan çekilen örnekleme ilişkin histogram ve üçüncü satırdaki il k grafik ile birinci lastik türü ortalaması ile ikinci lastik türü ortalaması arasındaki fark i çin elde dilen sonsal dağılımdan çekilen örnekleme ilişkin histogram yer almaktadır.



Şekil 6.1 MCMC yöntemi kullanılarak elde edilen örneklemlere ilişkin grafikler

6.2 Normal Yaklaşım Kullanılarak İki Oran Arasındaki Fark için Bayesci Sonuç Çıkarımı

Birbirinden bağımsız iki kitle göz önüne alınsın. Birinci kitlede belli bir özelliğe sahip olanların oranı π_1; ikinci kitlede aynı belli özelliğe sahip olanların oranı π_2 olmak üzere her iki kitleden rasgele örneklem alındığı ve belli özelliğe sahip olanların sayısı ile ilgilenildiği varsayılsın.Y_1ve Y_2 ile sırası ile birinci ve ikinci kitlede bahsedilen özelliğe

sahip bireylerin sayısı gösterildiğinde, $(\mathrm{Y}_1|\pi_1)$ in dağılımı Binom(n_1, π_1) ve $(\mathrm{Y}_2|\pi_2)$ in dağılımı Binom(n_2, π_2) olur.

$\pi_1$ve π_2' nin önsel dağılımları bağımsız olduğundan sonsal dağılımlar da bağımsız olacaktır.

π_1' in önseli Beta(a_1, b_1)ve π_2' in önseli Beta(a_2, b_2) alındığında, sonsal dağılımları da yine birbirinden bağımsız Beta dağılımı olacaktır.

$\pi_1$in sonsalı Beta(a_1^l, b_1^l)ve $a_1^l = a_1 + \mathrm{y}_1$, $b_1^l = b_1 + n_1 - \mathrm{y}_1$

$\pi_2$in sonsalı Beta(a_2^l, b_2^l) ve $a_2^l = a_2 + \mathrm{y}_2$, $b_2^l = b_2 + n_2 - \mathrm{y}_2$

Her bir sonsal için normal dağılım yaklaşımı kullanılırsa π_1, ortalaması $\frac{a_1^l}{a_1^l+b_1^l}$ve varyansı $\frac{a_1^l b_1^l}{(a_1^l+b_1^l)^2(a_1^l+b_1^l+1)}$, π_2, ortalaması $\frac{a_2^l}{a_2^l+b_2^l}$ ve varyansı $\frac{a_2^l b_2^l}{(a_2^l+b_2^l)^2(a_2^l+b_2^l+1)}$ olan normal dağılım olarak yazılabilirler.

Bu durumda, iki oran arasındaki fark $\pi_d = \pi_1 - \pi_2$ için sonsal dağılım yaklaşık Normal$(m_d^l, (S_d')^2)$ dır. Sonsalın ortalaması

$$m_d^l = \frac{a_1^l}{a_1^l + b_1^l} - \frac{a_2^l}{a_2^l + b_2^l}$$

sonsalın varyansı

$$(S_d')^2 = \frac{a_1^l b_1^l}{(a_1^l + b_1^l)^2(a_1^l + b_1^l + 1)} + \frac{a_2^l b_2^l}{(a_2^l + b_2^l)^2(a_2^l + b_2^l + 1)}$$

biçiminde elde edilir (Bolstad, 2004).

Oranlar arasındaki fark için Bayesci güvenilir aralık

Yaklaşık normal sonsal dağılım için güvenilir aralığın genel elde edilme kuralı kullanılarak $\pi_d = \pi_1 - \pi_2$ nin (1- α)×100% lik Bayesci güvenilir aralık;

$$m_d^{\prime} \pm z_{\frac{\alpha}{2}} \times S_d^{\prime} \tag{5.1}$$

biçiminde elde edilir.

6.2.1 Oranlar arasındaki farkın tek yanlı Bayesci hipotez testi

$\pi_d = \pi_1 - \pi_2 > 0$ olup olmadığını test etmek için;

$$H_0: \pi_d \leq 0$$
$$H_1: \pi_d > 0$$

hipotezi gözönüne alınsın. Sonuca ulaşmak için yokluk hipotezinin yaklaşık sonsal olasılığı hesaplanır.

$$P(\pi_d \leq 0) = \mathrm{P}(\frac{\pi_d - m_d^{\prime}}{S_d^{\prime}} \leq \frac{0 - m_d^{\prime}}{S_d^{\prime}})$$

$$= P(Z \leq \frac{0 - m_d^{\prime}}{S_d^{\prime}})$$

Bu olasılık H_1 için hesaplanan olasılık değerinden küçük ise, yokluk hipotezi aleyhine karar verilir. Yani, $\pi_1 > \pi_2$ dir. Aksi durumda, yokluk hipotezi lehine karar verilir (Bolstad, 2004).

6.2.2 Oranlar arasındaki fark için iki yanlı Bayes hipotez testi

$$H_0: \pi_1 - \pi_2 = 0$$

$$H_1: \pi_{1-}\pi_2 \neq 0$$

Bu durumda, (5.1) denkleminde π_d için verilen aralık sıfır değerini içeriyorsa yokluk hipotezi lehine karar verilir ve α anlam düzeyinde güvenilirdir. Eğer aralığın dışındaysa, yokluk hipotezi α anlam düzeyinde güvenilir değildir ve H_0 aleyhine karar verilir.

Örnek 6.2. Pulmoner rehabilitasyon ünitesine gelen hastaların 1 yıllık kayıtları yapılmıştır. Bu kayıtlı hastaların içinden rasgele 200 hasta seçilmiştir. Koahlı olan 60 hastadan 40'i 45 yaşın üstündedir. Koahlı olmayan 140 hastadan 60'ı ise yine 45 yaşın üstündedir. 45 yaşın üstünde olanlarda koah olma olasılığı 45 yaşın altında koah olma olasılığından farklı mıdır? Buna göre kurulacak hipotez,

$$H_0: \pi_{>45} - \pi_{\leq 45} = 0$$
$$H_1: \pi_{>45} - \pi_{\leq 45} \neq 0$$

olur. Elde edilen gözlem sonuçları aşağıdaki tablo ile özetlenebilir.

Çizelge 6.2 Hastaların yaş ve hastalık baz alınarak sınıflandırılması

	Koah-pzt	**Koah-ngtf**	**Toplam**
yaş>45	40	60	100
yaş≤45	20	80	100
Toplam	60	140	200

$\pi_{>45}$ ve $\pi_{\leq 45}$ in önsel dağılımı bağımsız Beta(1,2) olarak şeçildiğinde; $\pi_{>45}$ in sonsal dağılımı Beta(41,62) ve $\pi_{\leq 45}$ in sonsal dağılımı Beta(21,82) olarak bulunur. Buradan oranlar arasındaki farkın sonsal dağılımı, $\pi_d = \pi_{>45} - \pi_{\leq 45}$ yaklaşık Normal$(m_d^{\iota}, (S_d')^2)$ dır. Sonsalın ortalaması,

$$m_d^{\iota} = \frac{41}{103} - \frac{21}{103} = 0.2365$$

ve sonsalın varyansı,

$$(S_d')^2 = \frac{41 \times 62}{(41+62)^2(41+62+1)} + \frac{21 \times 82}{(21+82)^2(21+82+1)}$$

$$= 0.00386464$$

$$S'_d = 0.062166$$

olarak hesaplanır. Buna göre, değerler $m'_d \pm z_{\frac{\alpha}{2}} \times S'_d$ förmülünde yerine konulduğunda π_d için %95 lik Bayesci güvenilir aralık

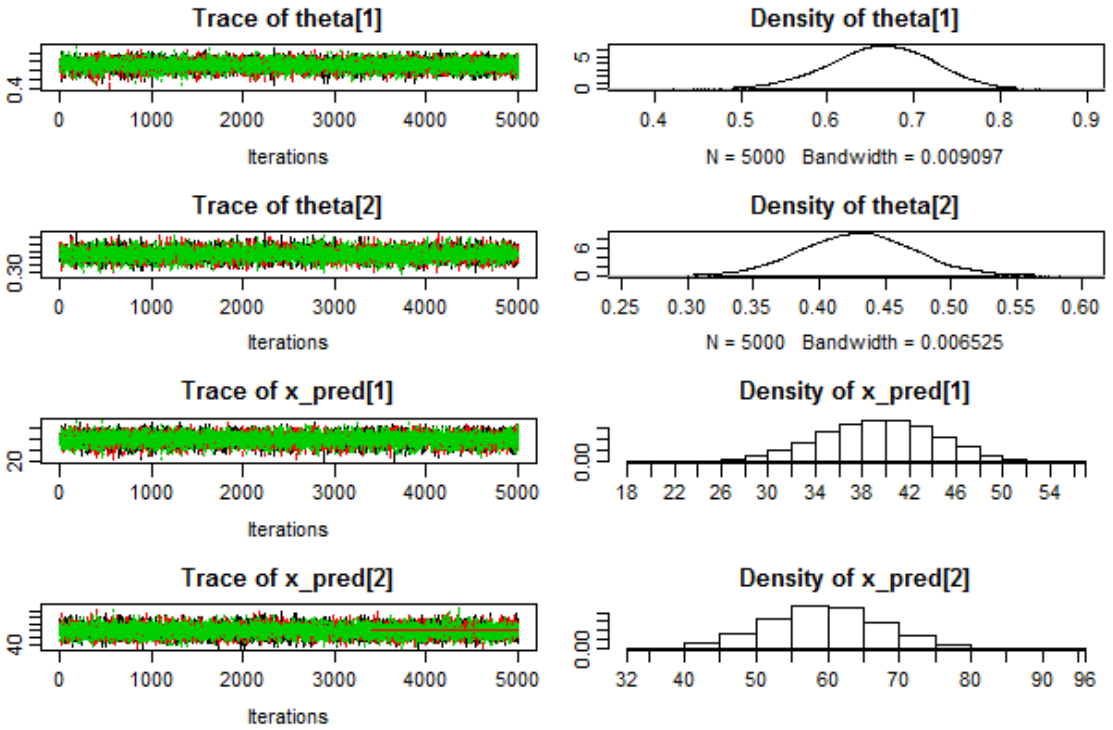
$$0.2365 \pm 1.96 \times 0.062116 = (0.114654, 0.358345)$$

biçiminde elde edilir. Görüldüğü gibi aralık sıfır değerini içermez. Buna göre, yokluk hipotezi aleyhine karar verilir, H_1 hipotezi daha olasıdır. Yani veriler doğrultusunda koah hastası yaş>45 ile yaş≤45 olanlarınarasında fark vardır, yaş koah hastalığı için anlamlı bir parametredir.

R programı ile ilk olarak klasik yöntemle sonuç çıkarımında iki oran arasındaki farkın güven aralığı (0.08145838, 0.39473209) olarak elde edilir. Görüldüğü gibi aralık sıfır değerini içermez. Buna göre, 45 yaşın üstünde olanlarda koah olma olasılığı farklıdır.

Bayesci Sonuç çıkarımı için R programı kullanıldığında iki oran arasındaki fark için Bayesci güvenilir aralık, (0.091,0.374) olarak elde edilir. Aralık sıfır değerini içermediği için 45 yaşın üstündekilerde koah olma riskinin farlı olduğu sonuçu tekrarlanmış olur. Bu elde edilen sonuç, yapılan diğer sonuç çıkarımlarını desteklemiştir (EK 4).

Bayesci sonuç çıkarımı ile R programından elde edilen grafikler,



Şekil 6.2 R programında sonuç grafiği

biçimindedir. Birinci satır birinci sütun da verilen grafik ilk kitle oranı için elde edilen sonsal dağılımdan çekilen örneklem değerlerine ait grafiktir. İkinci sütunda ise birinci kitle oranı için elde dilen sonsal dağılıma ait olasılık yoğunluk fonksiyonunun grafiği yer almaktadır. İkinci satır birinci sütun da verilen grafik ikinci kitle oranı için elde dilen sonsal dağılımdan çekilen örneklem değerlerine ait grafiktir. İkinci sütunda ise ikinci kitle oranı için elde edilen sonsal dağılıma ait olasılık yoğunluk fonksiyonunun grafiği yer almaktadır. Üçüncü satır birinci sütunda birinci grup için elde edilen tahmini başarıların sayısının grafiği ve ikinci sütunda bunlara ait histogram grafiği yer almaktadır. Benzer şekilde dördüncü satır birinci sütunda ikinci grup için elde edilen tahmini başarıların sayısının grafiği ve ikinci sütunda bunlara ait histogram grafiği yer almaktadır.

6.3 Bağımlı Normal İki Kitle Ortalaması Arasındaki Fark İçin Sonuç Çıkarımı

Verinin varyasyonuna değiştiren asıl etmen, verideki değişkenler arasındaki farklılıktır. Yapılan bir deneyde iki farklı deneme birbirinden bağımsız rasgele değişkenlere uygulanırsa, bu varyasyona etki eden asıl etmeni belirlemeyi ya da tedavinin etkileri arasındaki farkı belirlemeyi zorlaştırır.

Tasarlanan eşleştirilmiş deneyler, tedavinin etkileri arasındaki farkı tespit etmeyi daha da kolaylaştırır. Eşleştirilmiş deneyde, veriler çift halinde benzer öğelerle eşleştirilmelidir. Her eşten bir öğe birinci denemeye ve diğer öğe ikinci denemeye atanmalıdır. Bu randomize (rastgele) blok deney tasarımıdır. Burada çiftler bloklanır.

Örneğin süt endüstrisinde deney için özdeş ikiz buzağılar kullanılır. İkiz buzağılar aynı genetiğe sahiptir. Her çiftin birine birinci deneme, diğerine ikinci deneme uygulanır.

Eşleştirilmiş veri bir başka yolla da uygulanabilir. Örneğin iki deneme de aynı deneysel öğelere uygulanabilir (farklı zamanda). İlk denemenin etkisi geçtikten sonra ikinci deneme uygulanır. Denemeden önceki ya da denemeden sonraki deneysel öğeler arasındaki değişkenlikten dolayı aynı çift iki gözlem almak daha mantıklıdır. Çünkü farklı çift öğelerden alınan iki gözleme göre daha iyi sonuçlar elde edilir.

Aynı çiftler için; uygulanan A ve B denemelerinin gözlemleri arasındaki tek fark, deneme etkisi artı ölçüm hatasıdır.

Farklı çiftler için; uygulanan A ve B denemesinin gözlemleri arasındaki fark, deneme etkisi artı deneysel öğelerin (değişkenlerin)etkisi artı ölçüm hatasıdır. Çünkü eşleşmiş rasgele örneklere, birbirinden bağımsızmış gibi davranılamayacaktır (ikizler bağımlıdır).

Deneme grubuna atanan değişkenlerin ortalaması, deneme grubunun ortalaması artı denemenin etkisidir. Sırasıyla ortalamaları μ_1 ve μ_2 olan normal kitleden iki rasgele örneklem alınsın. Eğer denemenin etkisi sabit ise, buna katkı modeli denir. Her iki

gözlem verisi aynı varyansa sahiptir. Katkı modeli seçildiğinde, kitlelerin varyansları eşit olacaktır. Varyans iki parçadan oluşur. Bunlar; ölçüm hatası ve deneysel öğeler (değişkenler) arasındaki rasgele varyasyondur.

Dikkat edilmesi gereken bazı durumlar,

- ✓ Eğer varyans bilinmiyor ise, her bir örnek varyansı verilerden tahmin edilmelidir. Bu durumda kritik değeri bulmak için standart normal tablo yerine serbestlik derecesi $(n_1 + n_2 - 2)\ olan$ Student t tablosu kullanılır.

- ✓ Eğer varyans bilinmiyor ve eşit değil ise, örnek tahmin değerlerini gerçek değerlermiş gibi kullanılır. Kritik değeri bulmak için Student t tablosu kullanılır. Serbestlik derecesi ise Satterthwaite yaklaşımından elde edilir.

- ✓ $\mu_d = \mu_1 - \mu_2$ nin (1- α)×100% Bayesci güvenilir aralık,

$$\mathrm{m}_d^{\iota} \pm z_{\frac{\alpha}{2}} \times \mathrm{S}_d^{\iota}$$

dir.

- ✓ 1.Denemeden y_{i1}, 1.deneme i.çiftingözlemi
  2.Denemeden y_{i2},2.deneme i.çiftin gözlemi

bağımlı örneklemlerde gözlemler eşleştirildiğinde,$d_i = y_{i1} - y_{i2}$ her çift için fark hesaplanır. Buna göre, d_i' ler varyansı σ_d^2 ortalaması $m_d = \mu_1 - \mu_2$olan normal rasgele örneklem olur. Normal dağılımlar arasındaki fark tedavi etkisi hakkında çıkarım yapmak için kullanılır.

Yani, d_i=y_{i1} -y_{i2} olmak üzere $m_d = \mu_1 - \mu_2$ ve dağılımları N(m_d, σ_d^2)' dir.

Örnek 6.3 Yıllık süt verimini artırmak için mineral takviyesinin etkili olup olmadığı araştırılmak isteniyor. Bu amaçla, 15 çift aynı tür ikiz süt inekleri kullanılıyor. Her çift

inekten biri tedavi grubuna atanıyor ve mineral takviyesi veriliyor, 15 çiftin geride kalan inekleri kontrol grubuna atanıyor ve mineral takviyesi verilmiyor. Aşağıdaki tabloda ineklerin yıllık süt verimi gösterilmiştir (Bolstad, 2004).

Çizelge 6.3 İneklerin yıllık süt verimi

İkiz çift	**Kontrol grubunun yıllık verimi(lt)**	**Tedavi grubunun yıllık verimi(lt)**
1	3525	3340
2	4321	4279
3	4763	4910
4	4899	4866
5	3234	3125
6	3469	3680
7	3439	3965
8	3658	3849
9	3385	3297
10	3226	3124
11	3671	3218
12	3501	3246
13	3842	4245
14	3998	4186
15	4004	3711

Genetik kodu aynı olan bu doğal ikizlerin normal dağılımdan alındığı kabul ediliyor. Genetiği aynı olan inekler bağımsız olarak düşünülemez ve farkları alınır.

Fark, d_i=y_{i1}-y_{i2} olup N(m_d, σ_d^2) dağılımlıdır. Burada, $\sigma_d^2 = 270^2$ bilindiği varsayılıyor ve $m_d = \mu_1 - \mu_2$ olarak alınıyor.

Üç farklı araştırmacının aldığı önsellere bakılsın. Ali m_d için düz önsel kullanıyor. Berk m_d için $N(m, S^2)$ ve $m = 0$ ve $S = 200$ olan normal dağılım alıyor. Can m_d için üç köşeli bir önsel düşünüyor. Bunun için kullandığı noktaları aşağıdaki tablo ile veriyor. Can sürekli önselini, belirlediği değerler için lineer interpolasyon yaparak buluyor.

Çizelge 6.4 Can'ın önseli için köşe noktaları

x	y
-300	0
0	3
300	0

Ali düz önsel kullandığı için sonsalı N$(m', (S')^2)$ dağılım olarak buluyor. Burada, $m'=\bar{y}=7.07$ ve $(S')^2=\frac{270^2}{15}=4860$ olarak hesaplıyor. Buna göre, Ali sonsalının standart sapmasını $S' = \sqrt{4860} = 69.71$ olarak hesaplıyor.

Berk N$(0,200^2)$ önseli kullandığı için sonsalı da N$(m', (S')^2$) oluyor. Ortalama ve varyansı,

$$\frac{1}{(S^\iota)^2} = \frac{1}{S^2} + \frac{n}{\sigma^2} = \frac{\sigma^2 + nS^2}{\sigma^2 S^2} \quad \text{ve} \quad m^\iota = \frac{1/S^2}{n/S^2 + 1/S^2} \times m + \frac{n/S^2}{n/S^2 + 1/S^2} \times \bar{y}$$

$$\frac{1}{(S^\iota)^2} = \frac{1}{200^2} + \frac{15}{270^2} = 0.000230761 S^\iota = 65.83$$

ve

$$m^\iota = \frac{1/200^2}{0.000230761} \times 0 + \frac{15/270^2}{0.000230761} \times 7.07$$

$$= 6.33$$

olarak hesaplıyor.

Can ise $g(\mu|y_1, \dots, y_n) = \frac{g(\mu) \times f(y_1, \dots, y_n|\mu)}{\int g(\mu) \times f(y_1, \dots, y_n|\mu) d\mu}$ denklemini kullanarak sonsala ulaşıyor. Tedavinin süt protein üretimini artırmada etkili olup olmadığını öğrenmek için kurulacak hipotez

$$H_0: \mu_d \leq 0$$
$$H_1: \mu_d > 0$$

biçiminde olacaktır. $\alpha = 0.05$ anlam düzeyinde Ali ve Berk sonsal dağılım olarak normal dağılım elde ettiği için

$$P(H_0: \mu \leq \mu_0 | y_1, \dots, y_n) = P(\frac{\mu - m^{\iota}}{S^{\iota}} \leq \frac{\mu_0 - m^{\iota}}{S^{\iota}})$$

$$= \mathrm{P}(\mathrm{Z} \leq \frac{\mu_0 - m^{\iota}}{S^{\iota}})$$

denklemi ile yokluk hipotezinin sonsal olasılığını hesaplarlar. Can'ın sonsalı sayısal olarak elde edildiği için yokluk hipotezinin olasılığı, $\mathrm{P}(H_0: \mu \leq \mu_0 | y_1, \dots, y_n) = \int_{-\infty}^{\mu_0} g(\mu \leq \mu_0 | y_1, \dots, y_n) d\mu$ ile hesaplanır. Buna göre sonuçlar,

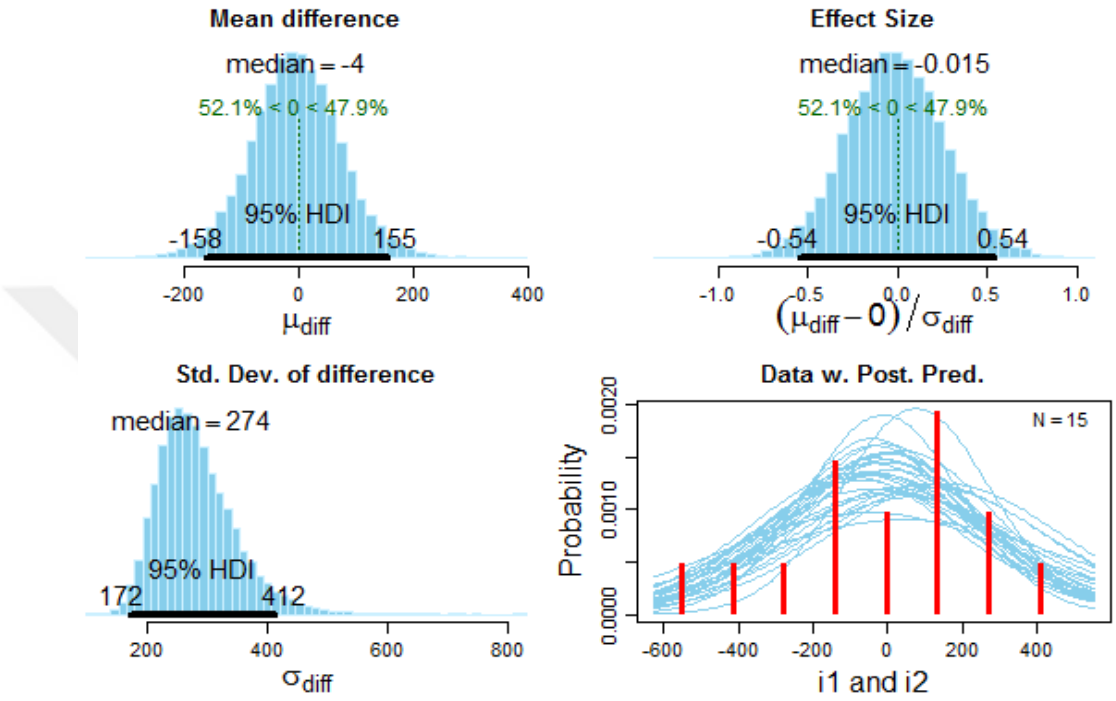
Çizelge 6.5 Bayesci güven aralığı hesaplama sonuçları

İsim	**Sonsal**	$\mathbf{P(\mu \leq 0 | d_1, \dots, d_n)}$	
Ali	N(7.07,69.71^2)	P(Z$\leq \frac{0-7.07}{69.71}$)=0.4596	H_0 lehine karar verilir
Berk	N(6.33,65.83^2)	P(Z$\leq \frac{0-6.33}{65.83}$)=0.4619	H_0 lehine karar verilir
Can	Sayısal	$\int_{\infty}^{0} g(\mu_d | d_1, \dots, d_n) d\mu = 0.4684$	H_0 lehine karar verilir

biçiminde elde edilir. Farklı önseller kullanılarak hesaplanan her üç durum içinde H_0 hipotezi lehine karar verildi. Yani, mineral takviyesinin ineklerin süt verimine herhangi bir etkisi olmadığı sonucuna ulaşılır.

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

hipotezlerinin testi, R programı ile yapıldığında Bayesci sonuç çıkarımına göre %95 lik güvenilir aralık, R'da yapılan hesaplamada (-158.215,154.534) biçiminde olup 0 değerini içerdiğinden ortalamalar arasında fark olmadığı söylenebilir (EK 5)
Sonuçlara ilişkin grafikler aşağıda verilmiştir.



Şekil 6.3 R programında sonuç grafikleri

Örnek 6.4 Jogging (kır koşusu) yapmanın kalp atış sayısı üzerine etkisi olup olmadığı tıp dünyasında süregelen bir tartışma konusudur. Bu analizi yapmak için önceden jogging aktivitesi olmayan 8 gönüllü seçilmiş ve 1 ay süren düzenli jogging programına alınmıştır. Teste başlamadan önce ve 1 aylık sürenin sonunda nabız sayıları ölçülerek aşağıdaki tabloda verilen değerler elde edilmiştir (Ross, 2004).

Çizelge 6.6 Deneklerin jogging programı öncesi ve sonrası nabız sayıları

Kişi no	1	2	3	4	5	6	7	8
Önceki nabız sayısı	74	86	98	102	78	84	79	70
Sonraki nabız sayısı	70	85	90	110	71	80	69	74
$d_i = y_{i1} - y_{i2}$	4	1	8	-8	7	4	10	-4

$$\sum_{i=1}^{8} d_i = \sum_{i=1}^{8} y_{i1} - y_{i2} = 22$$

$$\bar{d_\iota} = \frac{22}{8} = 2.75$$

$$\mu_1 - \mu_2 \sim \mathrm{N}(m_d, \sigma_d^2)$$

Kalp atış sayısının dağılımının normal olduğu varsayımı altında aynı kişiler üzerinden elde edilen ölçümler arasındaki farkın dağılımı da normal olacaktır. Farkların dağılımı, $\mu_d \sim \mathrm{N}(m_d, \sigma_d^2)$ olmak üzere $\sigma_d^2 = 6^2$ olduğu varsayılsın.

Kalp atış sayısı ortalaması için önsel dağılım olarak normal dağılım alınsın ve önselin parametreleri $m_d = 3\,, S^2 = 3^2$ olarak alınsın. Buna göre, μ_d için sonsal dağılım da bir normal dağılım olur. Bu dağılımın varyansı ve ortalaması,

$$\frac{1}{(s^\iota)^2} = \frac{1}{3^2} + \frac{8}{6^2} = 0.333$$

$$s^\iota = 3$$

$$m^\iota = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{8}{6^2}} \times 3 + \frac{\frac{8}{6^2}}{\frac{1}{3} + \frac{8}{6^2}} \times (2.75) = 2.9$$

olarak elde edilir. Yani, sonsal dağılım $N(2.9{,}3)$ olur. Buna göre, $\mu_d = \mu_1\text{-}\mu_2$ için %95 lik Bayesci güvenilir aralık standart normal dağılım tablosu kullanılarak,

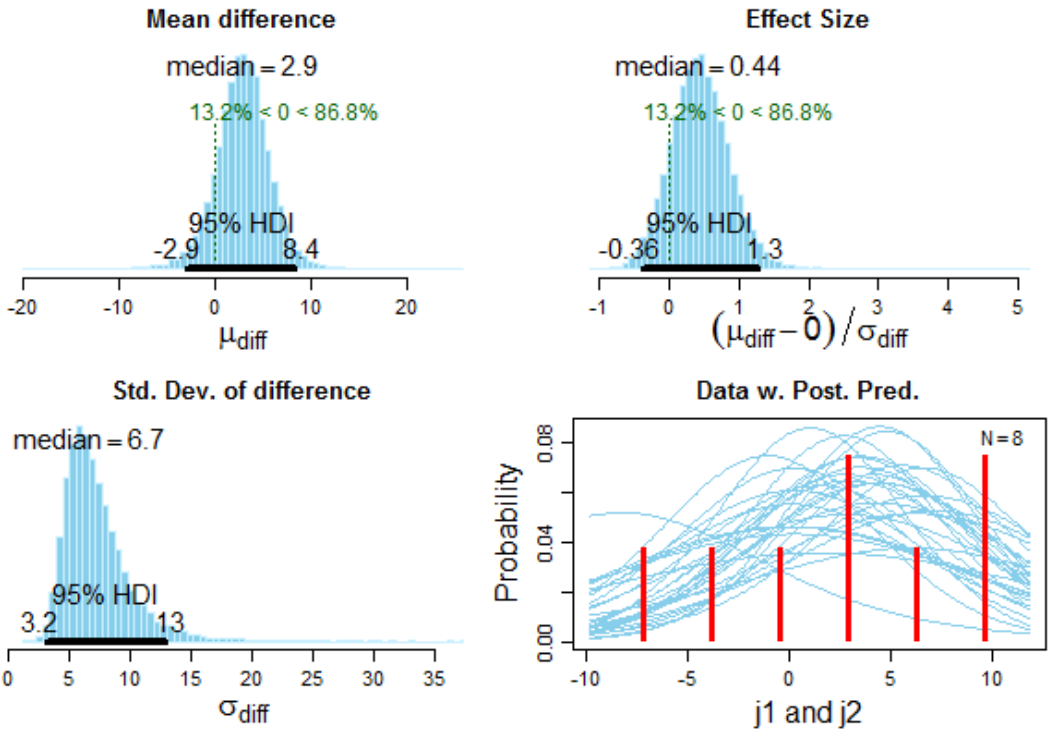
$$2.9 \pm 1.96 \times \sqrt{3} = (-0.49482,\ 6.29482)$$

biçiminde elde edilir. Bu aralık sıfırı içerdiği için iki ölçüm grubunun ortalaması arasında fark olmadığı söylenebilir.

R da klasik yöntemle μ_d için % 95 güven düzeyinde güven aralığı (-2.398733,7.898733) olarak elde edilir. Yani, ortalamalar arası fark yoktur sonucuna ulaşılır.

R'da Bayesci yöntemle sonuç çıkarımı için gerekli kodlar yazıldığında Bayesci sonuç çıkarımına göre %95 lik Bayesci güvenilir aralık (-3.034,8.602) olup ortalamalar arası fark yoktur sonucu ortaya çıkar (EK 6)

Sonuçlara göre grafikler,



Şekil 6.4 R programında sonuç grafikleri

biçiminde elde edilir.

7. GERÇEK VERİ İLE UYGULAMA

Uygulama çalışmasında, bir üniversitenin çocuk psikiyatri bölümüne dikkat eksikliği ve hiperaktivite sorunu ile başvuran 22 hasta göz önüne alınmıştır. Bu hastaların, uygulanan tedavi öncesi ve tedavi sonrası EEG çekimi yapılmıştır. Problem detayına inmeden önce EEG ile ilgili gerekli ön bilgiler aşağıda verilmiştir.

EEG rahat bir konumda 4 er dakika gözler kapalı ve açık pozisyonda kaydedilmiştir.

Yöntem: EEG kaydı10-20 yerleştirme yöntemiyle 19 kafa derisi elektrot başlığı kullanılarak hesaplanmıştır. Montaj olarak kulak referans montajı kullanılmıştır. EEG, psikiyatri alanında zihinsel bozukluğun kesin yöntemler kullanılarak belirlenmesi amacıyla, organik nedenli beyin patolojisinin var olup olmadığının belirlenmesinde kullanılmaktadır. EEG dalgalarından uygulamada kullanılacak dalgalar ve beyin loblarıyla olan ilişkisi söyledir:

- Alfa Dalgaları, 8-12 Hz arasındaki beyin dalgasıdır.. Uyanıklık, rahat ve sakinlik halinde belirir. En çok oksipital bölgede görülür.
- Gamma Dalgaları,25-140 Hz arasındaki beyin dalgalarıdır. Başın merkez bölgesinde, genlikleri normalden daha fazladır.
- Teta Dalgaları,4-7 Hz arasında yer alan beyin dalgalarıdır, belirgin olarak çocuklarda parietal ve temporal bölgelerde ortaya çıkarlar.
- Beta Dalgaları, frekansları 13-32 Hz arasındaki beyin dalgalarıdır. Beynin parietal ve frontal bölgelerinden daha yoğun elde edilirler (Güler, 2003).
- Teta/Beta oranı kişinin dikkat yeteneği düzeyini gösteren orantısal bir ifadedir. Bu oran yaşla ters orantılıdır, büyük çocuklarda daha çok olması beklenirken, yetişkinlerde daha küçüktür (Monastra et.al., 1999). Teta/Beta oranı parietal kortekste artmaktadır (Anonymous. 2018)

Uygulama 1 Elde edilen verilere göre, tedavinin Alfa dalga güçleri arasında anlamlı bir fark yaratıp yaratmadığına bakılsın.

Çizelge 6.7 Temporal bölgenin Teta gücü

Temporal bölge tedavi öncesi Teta gücü							
26.6436	31.2839	29.1631	26.2222	33.8823	28.2907	30.6711	28.1375
26.8213	26.2729	26.6278	25.3562	27.3727	27.1762	29.1241	27.8773
25.5633	27.5944	24.6113	31.5630	28.9088	25.2543		
Temporal bölge tedavi sonrası Teta gücü							
28.1797	30.5817	29.1452	25.5782	30.2320	26.8521	27.4939	25.4960
27.4804	26.6482	27.1865	25.9828	26.6730	28.2817	26.5932	26.9376
25.8754	27.1233	24.6888	30.6834	28.8000	25.3425		

İlk olarak verilerin normallik testi için,

H_0: Verilerin dağılmı Normal dağılıma uygundur
H_1:Verilerin dağılmı Normal dağılıma uygun değildir

Hipotezleri yazılmış ve bu hipotez R programı kullanılarak test edilmiştir.

Her iki durumda da $p-$değerleri $\alpha = 0.05$ değerinden büyük çıktığı için H_0 hipotezi reddedilememiştir. Yani, verilerin dağılımı Normal dağılıma uygun bulunmuştur (EK 7).

μ_1, ile tedaviden önce temporal bölgedeki Teta güçlerinin ortalaması, μ_2, ile tedaviden sonra temporal bölgeki Teta güçlerinin ortalaması gösterilir ise istenilen hipotez, μ_d için iki yanlı,

$$H_0: \mu_1\text{-}\mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

hipotez testidir.

Bayesci yaklaşımda; kitle varyansı bilinmediği için elimizdeki veriden örneklem varyansı (σ_d^2=1.9028) hesaplandı. Örneklem ortalaması $\bar{d}_i$=0.5710 olarak bulundu.

Önsel dağılım olarak Normal dağılım alındı ve önsel dağılım için parametreler $m_d = 1\,, S^2 = 1^2$ olarak alındığında, μ_d için sonsal dağılımın da bir normal dağılım olduğu bilgisi ile varyansı ve ortalaması,

$$\frac{1}{(s_p{}^{\iota})^2} = \frac{1}{1^2} + \frac{22}{\sigma_d^2} = 12.56$$

$$s_p{}^{\iota} = 0.28$$

$$m_p{}^{\iota} = \frac{1}{1 + \frac{22}{0.28}} \times 1 + \frac{\frac{22}{0.28}}{1 + \frac{22}{0.28}} \times 0.57$$

$$m_p{}^{\iota} = 0.013 + 0.563 = 0.576$$

olarak elde edildi. Buna göre, sonsal dağılım $N(0.576{,}0.28)$ olarak bulundu.

$\mu_d = \mu_1 - \mu_2$ farkı için %95 lik Bayesci güvenilir aralık;

$$m^{\iota} \pm t_{21,0.025} \times S^{\iota}$$

$$0.576 \pm -2.0796 * 0.28 = (-0.00628, 1.158)$$

biçiminde elde edildi. Sıfır değeri güvenilir aralığın içinde olduğu için H_0 yokluk hipotezi lehine karar verilir. Yani, "Uygulanan tedavinin dikkat eksikliği üzerinde etkisi yoktur" sonucuna ulaşılır.

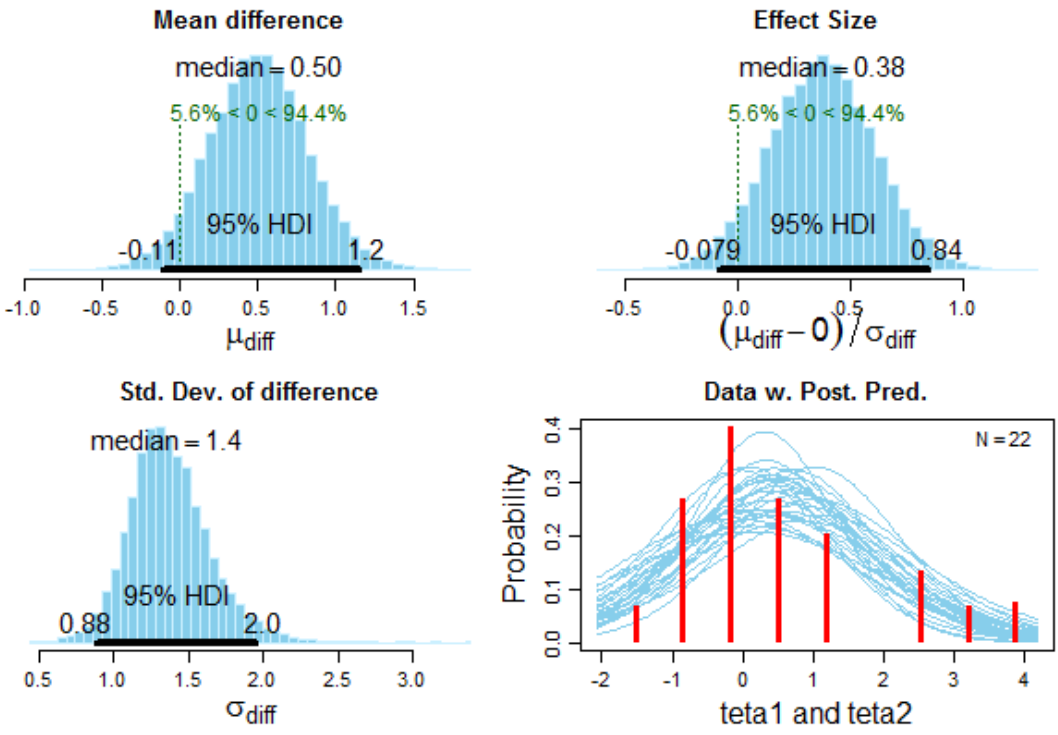
Bu testi klasik yaklaşımla test etmek için SPSS paket programı kullanılmıştır. SPSS paket programında veriler girilip Paired Samples Tests seçeneği seçildiğinde klasik

yaklaşım $\mu_d = \mu_1 - \mu_2$ ortalama farkı için %95 lik güven aralığı (-0.04059, 1.18263) olarak bulunmuştur. Bu aralık sıfırı içerdiği için yokluk hipotezi reddedilemez. Yani "Uygulanan tedavinin dikkat eksikliği üzerinde etkisi yoktur" sonucuna ulaşılır. (EK 8)

Klasik sonuç çıkarımı için R programında gerekli komutlar yazıldığında iki ortalama ara sındaki fark için R'da güven aralığı, (-0.04058908, 1.18262545) olarak elde edilir (EK 9). Buna göre klasik yaklaşım için SPSS ve R sonuçları birbirleriyle uyumludur.

Bayesci Sonuç çıkarımı için R programı kullanılarak hipotez test edildiğinde iki ortalam a arasındaki fark için %95 lik güvenilir aralık (-0.114, 1.154) olarak elde edilmiştir. Ara lık sıfır değerini içerdiği için; "Uygulanan tedavinin dikkat eksikliği üzerinde etkisi yokt ur" sonucuna ulaşılır (EK 10).

Bayesci sonuç çıkarımı için R programı kullanıldığında, elde edilen grafikler,



Şekil 7.1 Gerçek veri ile R programında sonuç grafikleri

biçimindedir.

Uygulama 2 Veriler de ilgilenilen konulardan biride Oksipital bölgedeki alpha dalga güçleri için tedavi öncesi ve sonrası arasında fark olup olmadığı sorusudur. Buna göre verilerden, tedavinin Alfa dalga güçleri arasında anlamlı bir fark yaratıp yaratmadığı incelensin. Veriler Çizelge 6.8 de verilmiştir.

Çizelge 6.8 Oksipital bölgenin Alfa gücü

Oksipital bölgenin tedavi öncesi Alfa gücü							
37.0198	27.3710	27.5910	33.3577	27.3600	34.4786	29.0937	26.5923
34.7495	29.9598	27.1966	34.8735	31.4470	33.7161	33.4669	32.2268
27.9788	31.4389	32.0812	30.0786	34.6649	34.4212		
Oksipital bölgenin tedavi sonrası Alfa gücü							
41.384	*30.4026*	*27.1416*	*32.9351*	*25.1505*	*32.4975*	*27.4409*	*21.9390*
30.2107	*29.5698*	*26.8306*	*34.9172*	*30.4709*	*30.6255*	*31.2510*	*28.8457*
30.6388	*30.9884*	*32.1294*	*28.9253*	*30.1088*	*33.4649*		

İlk olarak verilerin normal dağılıma uyumları için;

H_0: Verilerin dağılmı Normal dağılıma uygundur
H_1:Verilerin dağılmı Normal dağılıma uygun değildir

hipotezleri R programı kullanarak test edilsin.

Her iki durumda da $p-$değeri $\alpha = 0.05$ değerinden büyük çıktığı için H_0 hipotezleri reddedilemez yani verilerin dağılımı Normal dağılıma uygundur (EK 11).

μ_1, ile tedaviden önce oksipital bölgedeki Alfa güçlerinin ortalaması, μ_2, ile tedaviden sonra oksipital bölgedeki Alfa güçlerinin ortalaması gösterilir ise istenilen hipotez, μ_d için iki yanlı,

$$H_0: \mu_1\text{-}\mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

hipotez testidir.

Bayesci yaklaşım ile hipotez test edilmek istendiğinde, ilk olarak önsel dağılım olarak Normal dağılım alınsın ve önselin parametreleri $m_d = 1\,, S^2 = 4^2$olsun. Buna göre, μ_d için sonsal dağılım da bir normal dağılım olur ve varyansı ve ortalaması,

$$\frac{1}{(s_{o1}{}')^2} = \frac{1}{4^2} + \frac{22}{5.48} = 4.07415$$

$$s_{o1}{}' = 0.4954$$

$$m_{o1}{}' = \frac{\frac{1}{4^2}}{4.07415} \times 1 + \frac{\frac{22}{5.48}}{4.07415} \times (1.0589) = 1.05797$$

olarak elde edilir. Yani, sonsal dağılım $N(1.05797, 0.4954^2)$ olarak elde edilir.

$\mu_d = \mu_1\text{-}\mu_2$ için %95 lik Bayesci güvenilir aralık;

$$1.05797 \pm t_{21,0.025} \times 0.49543 = (0.02767, 2.08827)$$

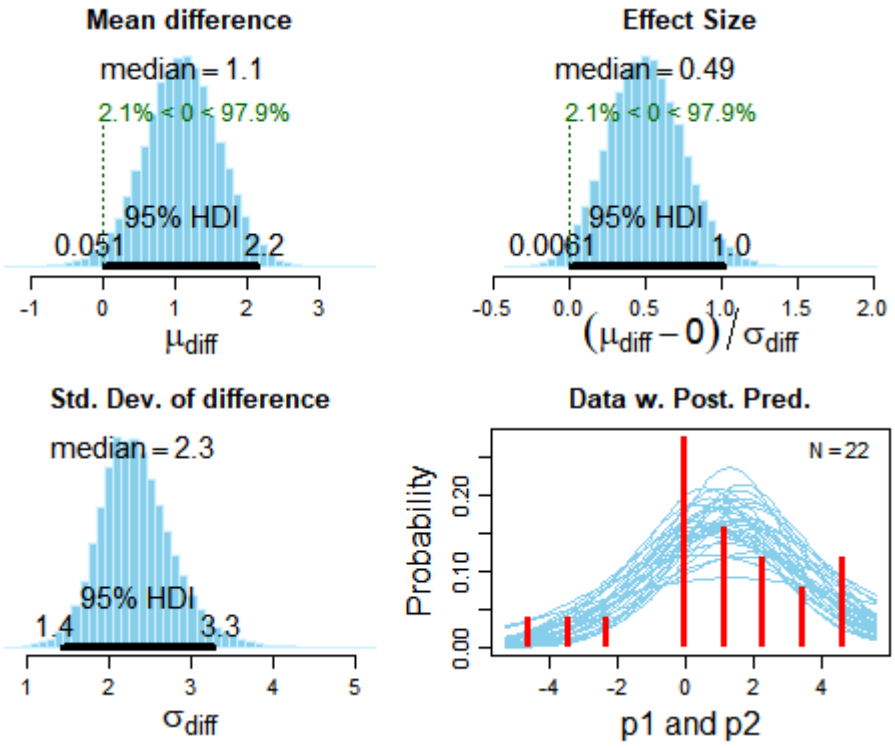
olarak elde edilir. Bu aralık da sıfırı içermediği için yokluk hipotezi aleyhine karar alınır . Yani, "Uygulanan tedavi ile oksipital bölgedeki alfa dalga güçleri farklılaşmıştır" sonu cuna bir kez daha ulaşılmış olur. Aynı dağılımdan farklı önsellerin kullanımının test son ucunu değiştirmediği görülmüştür. μ_d için iki yanlı hipotez testi,

$$H_0: \mu_1\text{-}\mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

biçiminde alınsın. Bir önceki örnekte olduğu gibi klasik yöntemle hipotezi test etmek için SPSS paket programı kullanıldığında $\mu_d = \mu_1 - \mu_2$ için %95 lik güven aralığı (0.02058, 2.09717) olarak bulunmuştur. Bu aralık sıfırı içermediği için yokluk hipotezi reddedilir. Yani, "Uygulanan tedavi oksipital bölgedeki alpha dalga güçlerini farklılaştırmıştır." sonucuna ulaşılır (EK 12)

Bayesci sonuç çıkarımı ile R'da iki ortalama arasındaki fark için güvenilir aralık (0.051, 2.156) olarak elde edilir. Aralık sıfır değerini içermediği için; "Uygulanan tedavinin beynin alpha dalga aktivitesi üzerinde etkisi vardır" sonucuna ulaşılır (EK 13).

R ile elde edilen grafikler,



Şekil 7.2 Gerçek veri ile R programında sonuç grafikleri

biçimindedir.

8. SONUÇ VE TARTIŞMA

Bayesci yaklaşımın temeli Bayes teoremidir. Bayesci yaklaşım ile parametreler de birer rasgele değişken olarak ele alınır. Bu nedenle kendilerine ait dağılımları vardır. Bu dağılım önsel dağılm olarak bilinir. Önsel dağılım bilgisi ve alınan örneklem bilgisi ile bu dağılım güncellenerek sonsal dağılım elde edilir. Bayesci sonuç çıkarımında tüm çıkarımlar elde edilen bu sonsal dağılm kullanılarak yapılır. Bu nedenle, önsel dağılım seçimi Bayesci çıkarım için oldukça önemlidir. Önsel dağılım olarak eşlenik aile kullanımı sonsala ulaşmada kolaylık sağlamaktadır.

Bayesci çıkarımın önemli noktalarından biri de hipotez testidir. Bu tez çalışmasında, hipotez testleri konusu ele alınmış ve Bayesci yöntem kullanılarak hipotez testlerinin nasıl yapıldığı araştırılmıştır. Her bir konu için yapılan örneklerle Bayesci hipotez testi uygulanmıştır. Parametreler için güvenilir aralıklar elde edilmiştir. Bayesci yöntem ile güvenilir aralık elde edilirken parametrenin kendi dağılımı kullanılır. Bu nedenle, elde edilen Bayesci güvenilir aralık klasik yöntemle elde edilene göre daha dardır.

Bayesci yöntem ile hipotez testi yapılmak istendiğinde,H_0 ve H_1 hipotezlerine önsel olasılıklar atanır. Daha sonra örneklem bilgisi ile bu olasılıklar Bayes teoremi kullanılarak güncelleştirilir ve H_0 ve H_1 hipotezlerinin sonsal olasılıkları elde edilir. Daha sonra bu olasılıkların karşılaştırılması ile sonuca ulaşılır. Eğer, H_0 hipotezinin sonsal olasılık değeri H_1 hipotezinin sonsal olasılık değerinden büyük ise H_0 yokluk hipotezinin daha olası olduğu söylenir. Ayrıca, Bayesci hipotez testlerinde sonuç çıkarımı yapmak için Bayes faktörü de kullanılır. Bayes Faktörü hipotezin kanıt gücünü değerlendirir.

Bu çalışmada ilk olarak oran için Bayesci hipotez testinin nasıl olduğu konusuna değinilmiş ve elde edilen sonuçlar yapılan örnekler ile değerlendirilmiştir. Daha sonra oran için güvenilir aralık elde edilmiş ve klasik yöntem kullanılarak elde edilen sonuçlarla karşılaştırması yapılmıştır.

Ortalama için Bayesci hipotez testinde normal ailelerde sonsal olasılığa ulaşmak için kolaylık sağlayan basit güncelleme kuralı uygulanır. Bayesci hipotez testinde örneklerde de görüldüğü gibi farklı dağılımlı önseller kullanımı sonsal dağılımı çok fazla etkilememiştir. Bayesci güvenilir aralık hesabında varyans biliniyor ise klasik yöntemde olduğu gibi z değeri kullanılır; varyans bilinmiyor ise bilinmeyen varyans yerine örneklem varyansı $\hat{\sigma}^2$ hesaplanır ve Student-t dağılımına ilişkin tablo değeri kullanılır. Ortalama için yapılan örnekler klasik yöntem ve Bayesci yaklaşım kullanılarak test edilmiş ve elde edilen sonuçlar karşılaştırılmıştır. Bulunan sonuçlar birbiriyle uyumludur.

Bağımsız ve bağımlı örneklemlerde iki kitle ortalaması farkı için Bayesci güvenilir aralık elde edilirken temel güncelleme formülü kullanılır. Bağımsız örneklemlerde iki kitle ortalaması farkı için güvenilir aralık varyanslar eşit ve biliniyor ve varyanslar eşit ve bilinmiyor durumları için elde edilmiştir. Örneklerde, klasik yaklaşım için SPSS ve R programlarında hesaplamalar yapılıp, Bayesci yöntem ile elde edilen test sonuçları ile karşılaştırılmıştır. Yine, benzer sonuçlar elde edilmiştir.

Bağımsız örneklemlerde iki kitle ortalaması farkı için varyansların eşit olmadığı ve bilinmediği durumda ortalamalar arasındaki farkın güven aralığı elde edilmiştir. Serbestlik derecesi, Satterhwaites'in formülü kullanılarak bulunmuştur.

Oranlar arasındaki fark için Bayesci hipotez testi örnekle desteklenerek anlatılmış çıkan sonuçlar R programı kullanılarak elde edilen sonuçlarla karşılaştırılmıştır.

Bayesci yöntem ile, bağımlı Normal iki kitle ortalaması arasındaki fark için sonuç çıkarımı yapılmak istendiğinde klasik yöntemde olduğu gibi gözlemlerin her çifti için $d_i = y_{i1} - y_{i2}$ farkı hesaplanır. Normal dağılımlar arasındaki bu fark tedavi etkisi hakkında çıkarım yapmak için kullanılır. Bu konuda farklı örnekler ile hesaplamalar yapılıp klasikçi yaklaşımla tutarlılığı göz önüne serilmiştir.

En son olarak Normal dağılıma sahip olan EEG verileri kullanılmıştır. Elde edilen, EEG sinyalleri ile uygulanan tedavinin etkisinin olup olmadığı Bayesci hipotez testi ile test edilmiştir. Klasik yaklaşım ile hipotez testi sonuçları için SPSS ve R programları

kullanılmıştır. Bayesci hipotez testi sonuçları için R programlama dili kullanılarak elde edilen sonuçlar karşılaştırılmıştır. Karşılaştırma sonucuna göre, elde edilen sonuçlar benzerdir. Buna göre, klasik hipotez testlerine alternatif olarak Bayesci yaklaşımın kullanılabileceği söylenebilir.

Sonuç olarak, tez kapsamında yapılan çalışmalarda Bayesci yöntem uygulanarak elde edilen hipotez testi sonuçlarının klasik yöntemlerle elde edilen sonuçlarla benzer olduğu görülmüştür. Bayesci yöntemlerde klasik yöntemlerden farklı olarak işin içine parametreler ile ilgili önsel bilgiler dahil edilir. Bu bilgilerin dahil edilmesi ile elde edilen güvenilir aralıkların klasik yöntemle elde edilenlerden daha dar olduğu söylenebilir.

Tez boyunca yapılan örneklerden elde edilen sonuçlara göre, Bayesci hipotez testinin klasik yaklaşıma iyi bir alternatif olduğu söylenebilir.

KAYNAKLAR

Akdi, Y. 2011, Matematiksel İstatistiğe Giriş. Gazi Kitabevi. 317, Ankara

Ando, T., 2010, Bayesian Model Selection and Statistical Modeling, Chapman & Hall, NewYork, 286 p.

Anonymous. 2018. Web Sitesi: https://www.maximumbrain.com.tr/wp-content/uploads /2017/07/QEEG-Report-Xxxx-XXXX.pdf, Erişim Tarihi: 15.11.2019.

Anonymous. 2019. Web Sitesi: http://www.rssb.be/bsn57/bsn57-6.pdf, Erişim Tarihi: 15.11.2019.

Berger, J. 2006. The case for objective bayesian analysis, bayesian analysis, 3, 385-402..

Bolstad, W. 2004. Introduction to Bayesian Statistics. Wiley Interscience, 354, New Zealand.

Demirci, M. 2016. Bayes Teoremi ve Teoremin İşletme Bölümünde Uygulamaları, vol. 43, pp. 439-462.

Demirhan, H., 2004, Logaritmik Doğrusal Modellerde Parametrelerin ve Beklenen Göze Sıklıklarının Bayesci Kestirimi, Yayınlanmamıs Bilim Uzmanlığı Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Ekici, O. 2005. Bayesyen Regresyon ve WinBUGS ile Bir Uygulama İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, Ekonometri Anabilim Dalı, 115, İstanbul.

Goldstein, A. 2014. Subjective bayesian analysis: principles and practice. Bayesian Analysis,3, 403-420.

Gu, X., Hoijtink, H. & Mulder, J.. 2016. Error probabilities in default Bayesian hypothesis testing. Journal of Mathematical Psychology, 72, 130-143. https://doi.org/10.1016/j.jmp.2015.09.001.

Güler, İ.. 2003. Web sitesi: https://slideplayer.biz.tr/slide/14035926/.

Güner, A.. 2014. Bayesci Yaklaşımda Eşlenik Aileleri Önseli ile Jeffrey's Önselinin Karşılaştırılması. Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 31-40, Eskişehir.

Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association.

Lindley, DV. 2000. The philosophy of statistics. Journal of the Royal Statistical Society: Series D (The Statistician).;49:syf 301.

Lindley DV. 2000. The philosophy of statistics. Journal of the Royal Statistical Society: Series D (The Statistician).;49:293-337.1995;90(430):773-795.

Lynch ,Scott M. 2006. Introduction to Applied Bayesian Statistics and Estimation for Social Scientists, Springer Science+Business Media, Llc.

Monastra, V. J., Lubar, J. F., Linden, M. K., et al. (1999). Assessing attention deficithyperactivity disorder via quantitative electroencephalography: An initial validation study.Neuropsychology,13,424–433.

Myung, J., 2006, Bayesian Methods for Social and Behavioral Scientists: Handson Bayes using WinBUGS, http://quantrm2.psy.ohio-state.edu/injae/ course/83108Bayes/priors.pdf, erisim tarihi: Mart, 2006.

Olivetti, E., Veeramachaneni, S., & Nowakowska, E.. 2012 Bayesian hypothesis testing for pattern discrimination in brain decoding. Pattern Recognition, 45(6), 2075-2084. https://doi.org/10.1016/j.patcog.2011.04.025.

Ortega, A and Navarrete, G. 2017. Bayesian Hypothesis Testing: An Alternative to Null Hypothesis Significance Testing (NHST) in Psychologysy . In Tejedor, J. P. (Ed.), Bayesian inference. 23525.http://dx.doi.org/10.5772/intechopen.70230.

Öztürk, L. 2004. Monte-Carlo simulasyon metodu ve bir işletme uygulaması. Doğu Anadolu bölgesi araştırmaları. syf 116-120.

Öztürk. F. 2006. Web sitesi: https://fikriozturk.wordpress.com/dersler/ist-101/. (Erişim tarihi: 0.9.09.2019)

Press, S.J., 1989, Bayesian Statistics, Principles, Models, and Applications, John Wiley and Sons, NewYork.

Radder, H. 2003. Toward a more developed philosophy of scientific experimentation. Pittsburgh: University of Pittsburgh Press; pp. 1-18.

Şen, S. 2018. An Implementation of the Gibbs Sampling Method Under the Rasch Model. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 9(3), 258-276.

Terzi, Y., Murat, N. ve Cengiz, M.A., Bayesci Hipotez Testleri ve Bayes Faktörü e-Journal of New World Sciences, 2008. Academy Natural and Applied Sciences, 3, (2), A0073, 321-329.

Tektaş, D., 2006. İki Düzeyli Lojit ve Probit Modellerde ParametreTahminlerine Bayesci Bir Yaklaşım. Hacettepe Üniversitesi, FenBilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara.

Van der Linden, S.,& Chryst, B. (2017). No need for Bayes Factors: A fully Bayesian evidence synthesis. Frontiers in Applied Mathematics and Statistics, 3. http://journal.frontiersin.org/article/10.3389/fams.2017.00012/full.

Vasilyev V.O. & Dobrovidov, A.V., 2016. Unsupervised Bayesian Hypothesis Testing. IFAC-PapersOnLine. 49(12), 592-597. https://doi.org/10.1016/j .ifacol. 2016.07.739.

Wackerly, D., Mendenhall, W., Scheaffer, R. 2008. Mathematical Statistics with Applications. Thomson Learning, 939,USA.

EKLER

EK 1 Örnek 4.4 için sonuçlar
EK 2 Örnek 6.1 için frekansçı sonuçlar
EK 3 Örnek 6.1 için Bayesci yaklaşım sonuçları
EK 4 Örnek 6.2 için sonuçlar
EK 5 Örnek 6.3 için sonuçlar
EK 6 Örnek 6.4 için sonuçlar
EK 7 Uygulama 1 Normal dağılıma uygunluk sonuçları
EK 8 Uygulama 1 için SPSS frekansçı sonuçlar
EK 9 Uygulama 1 için R programında frekançı sonuçlar
EK 10 Uygulama 1 için R programı ile Bayesci sonuç çıkarımı
EK 11 Uygulama 2 Normal dağılıma uygunluk
EK 12 Uygulama 2 için SPSS ile frekançı sonuç çıkarımı
EK 13 Uygulama 2 için R ile Bayesci güvenilir aralık sonuçları

EK 1 Örnek 4.4 için sonuçlar

```
        Exact binomial test

data:  5 and 10
number of successes = 5, number of trials = 10, p-value = 1
alternative hypothesis:true probability of success is not equal to 0.5
95 percent confidence interval:
 0.187086 0.812914
sample estimates:
probability of success
                   0.5
```

```
1-sample proportions test without continuity correction

data:  5 out of 10, null probability 0.5
X-squared = 0, df = 1, p-value = 1
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.2365931 0.7634069
sample estimates:
p
0.5
```

```
Bayesian First Aid binomial test

data: x and n
number of successes = 5, number of trials = 10
Estimated relative frequency of success:
  0.50
95% credible interval:
  0.23 0.77
The relative frequency of success is more than 0.5 by a probability of
0.494 and less than 0.5 by a probability of 0.506
```

EK 2 Örnek 6.1 için frekansçı sonuçlar

```
t.test(las1,las2,var.eq=TRUE)
Varyansların eşit alındığı durum:

        Two Sample t-test

data:  las1 and las2
t = 1.028, df = 16, p-value = 0.3192
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -172.5943  497.5943
sample estimates:
mean of x mean of y
   6165.0    6002.5
```

```
>t.test(las1,las2,var.eq=FALSE)
Varyansların farklı alındığı durum:

        Welch Two Sample t-test

data:  las1 and las2
t = 0.97867, df = 11.422, p-value = 0.348
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -201.3154  526.3154
sample estimates:
mean of x mean of y
   6165.0    6002.5
```

EK 3 Örnek 6.2 için sonuçlar

```
Data
l1, n = 10
l2, n = 8

  Model parameters and generated quantities
mu_x: the mean of l1
sigma_x: the scale of l1 , a consistent   estimate of SD when nu is la
rge.
mu_y: the mean of l2
sigma_y: the scale of l2
mu_diff: the difference in means (mu_x - mu_y)
sigma_diff: the difference in scale (sigma_x - sigma_y)
nu: the degrees-of-freedom for the t distribution fitted to l1 and l2
eff_size: the effect size calculated as
  (mu_x - mu_y) / sqrt((sigma_x^2 + sigma_y^2) / 2)
x_pred: predicted distribution for a new datapoint generated as l1
y_pred: predicted distribution for a new datapoint generated as l2

  Measures
mean       sd    HDIlo    HDIup %<comp %>comp
mu_x       6165.472 101.003 5968.111 6372.148  0.000  1.000
sigma_x     296.802  91.146  153.212  476.813  0.000  1.000
mu_y       5991.472 192.528 5633.048 6386.586  0.000  1.000
sigma_y     493.880 184.465  219.175  846.187  0.000  1.000
mu_diff     174.000 217.245 -257.235  591.193  0.191  0.809
sigma_diff -197.078 204.798 -611.995  171.502  0.880  0.120
nu           34.929  30.331    1.040   95.662  0.000  1.000
eff_size      0.456   0.527   -0.559    1.509  0.191  0.809
x_pred     6170.320 380.172 5453.017 6847.760  0.000  1.000
y_pred     5996.172 605.658 4807.797 7185.838  0.000  1.000

'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
'%<comp' and '%>comp' are the probabilities of the respective paramete
r being smaller or larger than 0.

  Quantiles
              q2.5%     q25%   median     q75%   q97.5%
mu_x       5963.924 6102.269 6166.103 6228.154 6369.385
sigma_x     172.492  234.366  279.685  339.003  522.890
mu_y       5623.186 5875.723 5990.010 6104.902 6377.648
sigma_y     262.395  371.359  455.509  570.465  957.504
mu_diff    -257.184   40.833  175.807  308.815  591.355
sigma_diff -674.897 -293.311 -171.090  -71.509  132.577
nu            3.597   13.471   26.119   46.676  116.893
eff_size     -0.558    0.102    0.449    0.803    1.511
x_pred     5470.727 5965.354 6170.662 6373.684 6867.452
y_pred     4821.897 5656.049 5993.672 6334.315 7201.729
```

EK 4 Örnek 6.2 için sonuçlar

```
2-sample test for equality of proportions with continuity correction

data:  c(40, 60) out of c(60, 140)
X-squared = 8.5952, df = 1, p-value = 0.00337
alternative hypothesis: two.sided
95 percent confidence interval:
 0.08145838 0.39473209
sample estimates:
   prop 1    prop 2
0.6666667 0.4285714
```

```
Bayesian First Aid proportion test

data: c(40, 60) out of c(60, 140)
number of successes:   40,  60
number of trials:      60, 140
Estimated relative frequency of success [95% credible interval]:
  Group 1: 0.66 [0.54, 0.77]
  Group 2: 0.43 [0.35, 0.51]
Estimated group difference (Group 1 - Group 2):
  0.23 [0.091, 0.37]
The relative frequency of success is larger for Group 1 by a probabili
ty
of >0.999 and larger for Group 2 by a probability of <0.001 .
```

```
>summary(res1)
  Data
number of successes:   40,  60
number of trials:      60, 140
  Model parameters and generated quantities
theta[i]: the relative frequency of success for Group i
x_pred[i]: predicted number of successes in a replication for Group i
theta_diff[i,j]:the difference between two groups (theta[i] -theta[j])
  Measures
mean    sd  HDIlo  HDIup %<comp %>comp
theta[1]          0.662 0.059  0.545  0.774  0.004  0.996
theta[2]          0.430 0.042  0.347  0.513  0.951  0.049
x_pred[1]        39.732 5.066 30.000 49.000  0.000  1.000
x_pred[2]        60.143 8.276 43.000 75.000  0.000  1.000
theta_diff[1,2]   0.232 0.072  0.091  0.374  0.001  0.999

'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
'%<comp' and '%>comp' are the probabilities of the respective paramete
r being smaller or larger than 0.5 (except for the theta_diff paramete
rs where the comparison value comp is 0.0).

  Quantiles
                  q2.5%   q25% median   q75% q97.5%
theta[1]          0.543  0.623  0.664  0.703  0.772
theta[2]          0.348  0.401  0.430  0.458  0.514
x_pred[1]        29.000 36.000 40.000 43.000 49.000
x_pred[2]        44.000 55.000 60.000 66.000 77.000
theta_diff[1,2]   0.087  0.184  0.234  0.282  0.371
```

EK 5 Örnek 6.3 için sonuçlar

```
Data
 i1, n = 15
 i2, n = 15

  Model parameters and generated quantities
 mu_diff: the mean pairwise difference between i1 and i2
 sigma_diff: the scale of the pairwise difference, a consistent
  estimate of SD when nu is large.
 nu: the degrees-of-freedom for the t distribution fitted to the pairw
 ise difference
 eff_size: the effect size calculated as (mu_diff - 0) / sigma_diff
 diff_pred: predicted distribution for a new datapoint generated
 as the pairwise difference between i1 and i2

  Measures
 mean       sd    HDIlo   HDIup %<comp %>comp
 mu_diff     -3.440  78.584 -158.215 154.534  0.521  0.479
 sigma_diff 283.041  64.307  172.330 412.362  0.000  1.000
 nu          33.831  29.804    1.032  92.277  0.000  1.000
 eff_size    -0.011   0.275   -0.541   0.538  0.521  0.479
 diff_pred   -0.197 479.495 -647.246 641.346  0.503  0.497

 'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
 '%<comp' and '%>comp' are the probabilities of the respective paramet
 er being
 smaller or larger than 0.

  Quantiles
              q2.5%     q25%  median    q75%  q97.5%
 mu_diff    -159.757  -53.851  -4.013  46.830 153.267
 sigma_diff  183.462  237.976 273.871 318.632 433.126
 nu            3.402   12.748  25.111  45.611 113.793
 eff_size     -0.545   -0.195  -0.015   0.173   0.533
 diff_pred  -644.218 -197.626  -2.029 192.701 646.669
```

EK 6 Örnek 6.4 için sonuçlar

```
Paired t-test Çift yönlü için

data:  j1 and j2
t = 1.263, df = 7, p-value = 0.247
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.398733  7.898733
sample estimates:
mean of the differences
                   2.75
```

```
Data
j1, n = 8
j2, n = 8

  Model parameters and generated quantities
mu_diff: the mean pairwise difference between j1 and j2
sigma_diff: the scale of the pairwise difference, a consistent
  estimate of SD when nu is large.
nu: the degrees-of-freedom for the t distribution fitted to the pairwi
se difference
eff_size: the effect size calculated as (mu_diff - 0) / sigma_diff
diff_pred: predicted distribution for a new datapoint generated
as the pairwise difference between j1 and j2

  Measures
mean     sd    HDIlo  HDIup %<comp %>comp
mu_diff     2.930  2.919  -3.206  8.389  0.133  0.867
sigma_diff  7.372  2.902   3.165 12.794  0.000  1.000
nu         32.147 28.959   1.003 90.496  0.000  1.000
eff_size    0.456  0.422  -0.347  1.299  0.133  0.867
diff_pred   2.863  9.401 -16.293 20.475  0.345  0.655

'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
'%<comp' and '%>comp' are the probabilities of the respective paramete
r being
smaller or larger than 0.

  Quantiles
             q2.5%   q25% median   q75%  q97.5%
mu_diff     -3.034  1.244  2.970  4.674   8.602
sigma_diff   3.819  5.513  6.751  8.506  14.427
nu           2.612 11.344 23.556 43.644 109.782
eff_size    -0.337  0.176  0.444  0.723   1.312
diff_pred  -15.799 -2.160  2.959  7.937  21.053
```

EK 7 Uygulama 1 Normal dağılıma uygunluk sonuçları

>shapiro.test(teta1) Shapiro-Wilk normality test data: teta1 W = 0.9388, p-value = 0.1869	>shapiro.test(teta2) Shapiro-Wilk normality test data: teta2 W = 0.94173, p-value = 0.2149

EK 8 Uygulama 1 için SPSS frekansçı sonuçlar

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	VAR1	27,9281	22	2,32075	,49479
	VAR2	27,3571	22	1,69778	,36197

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	VA1 & VAR2	22	,808	,000

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
d=x-y	,57102	1,37944	,29410	-,04059	1,18263	1,942	21	,066

EK 9 Uygulama 1 için R programında frekançı sonuçlar

```
Paired t-test

data:  teta1 and teta2
t = 1.9416, df = 21, p-value = 0.06572
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.04058908  1.18262545
sample estimates:
mean of the differences
              0.5710182
```

EK 10 Uygulama 1 için R programı ile Bayesci sonuç çıkarımı

```
summary(sont)
  Data
teta1, n = 22
teta2, n = 22

  Model parameters and generated quantities
mu_diff: the mean pairwise difference between teta1 and teta2
sigma_diff: the scale of the pairwise difference, a consistent
  estimate of SD when nu is large.
nu: the degrees-of-freedom for the t distribution fitted to the pairwi
se difference
eff_size: the effect size calculated as (mu_diff - 0) / sigma_diff
diff_pred: predicted distribution for a new datapoint generated
as the pairwise difference between teta1 and teta2

  Measures
mean     sd  HDIlo  HDIup %<comp %>comp
mu_diff     0.506  0.321 -0.114  1.154  0.056  0.944
sigma_diff  1.375  0.271  0.883  1.953  0.000  1.000
nu         31.954 29.714  1.001 91.253  0.000  1.000
eff_size    0.375  0.235 -0.079  0.844  0.056  0.944
diff_pred   0.498  1.652 -2.586  3.713  0.360  0.640

'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
'%<comp' and '%>comp' are the probabilities of the respective paramete
r being
smaller or larger than 0.

  Quantiles
            q2.5%   q25% median   q75%  q97.5%
mu_diff    -0.120  0.294  0.505  0.714   1.149
sigma_diff  0.902  1.195  1.352  1.530   1.980
nu          2.633 10.950 22.978 43.145 113.319
eff_size   -0.087  0.218  0.375  0.532   0.837
diff_pred  -2.639 -0.452  0.497  1.463   3.672
>plot(sont)
```

EK 11 Uygulama 2 Normal dağılıma uygunluk

shapiro.test(e1) Shapiro-Wilk normality test data: e1 W = 0.92827, p-value = 0.1129	>shapiro.test(e2) Shapiro-Wilk normality test data: e2 W = 0.93036, p-value = 0.1248

EK 12 Uygulama 2 için SPSS ile frekançı sonuç çıkarımı

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	x	31,4165	22	3,14291	,67007
	y	30,3577	22	3,79271	,80861

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	x& y	22	,788	,000

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
d=x-y	1,05888	2,34180	,49927	,02058	2,09717	2,121	21	,046

EK 13 Uygulama 2 için R ile Bayesci güvenilir aralık sonuçları

```
Measures
  mean     sdHDIlo  HDIup %<comp %>comp
  mu_diff     1.109  0.535  0.051  2.156  0.021  0.979
  sigma_diff  2.327  0.460  1.429  3.250  0.000  1.000
  nu             31.205 28.836 1.015 89.624 0.000  1.000
  eff_size 0.497  0.256  0.006  1.003  0.021  0.979
  diff_pred   1.129  2.747 -4.168  6.443  0.318  0.682

  'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
  '%<comp' and '%>comp' are the probabilities of the respective parameter being
  smaller or larger than 0.


   Quantiles
     q2.5%   q25% median q75%  q97.5%
  mu_diff     0.040  0.761  1.111  1.462   2.148
  sigma_diff  1.513  2.021  2.289  2.593   3.353
  nu2.645 10.665 22.432 42.202 109.373
  eff_size    0.016  0.324  0.491  0.663   1.018
  diff_pred  -4.162 -0.500  1.134  2.749   6.456
```

ÖZGEÇMİŞ

Adı Soyadı : Tuğba KARACAOĞLU UZUN

Doğum Yeri : Mamak

Doğum Tarihi : 17/07/1987

Medeni Hali : Evli

Yabancı Dili : İngilizce

Eğitim Durumu (Kurum ve Yıl)

Lise : Ahmet Yesevi Süper Lisesi (2005)

Önlisans : Üsküdar Üniversite Elektronöröfizyoloji Bölümü (2016)

Lisans : Ankara Üniversite Fen Fakülte Astronomi ve Uzay Bilimleri (2011)

Lisans Yandal : Ankara Üniversitesi İstatistik Anabilim Dalı (2009-2011)

Yüksek Lisans: Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı (Şubat 2020)

Çalıştığı Kurum/Kurumlar ve Yıl

Hacettepe Hastanesi EEG teknikeri (2016-)