9.3. Hata Olasılıkları ve Güç Fonksiyonu

Bir hipotez (ya da iddia) ya doğrudur ya da değildir. Bir deneyin sonuçlarından elde edilen gözlem değerlerine göre  yokluk hipotezi red edilir ya da red edilemez. Ayrıca, gerçekte  doğru olmasına rağmen, gözlem değerlerine göre  red edilebilir. Bu durumda bir hata yapılmış olur. Bu hataya *birinci tür hata* denir. Tersine, gerçekte  doğru olmamasına rağmen, verilere göre  yokluk hipotezi red edilemez. Burada da bir hata yapılmış olur. Bu hataya da *ikinci tür hata* denir. Örneğin, bir paranın düzgün (hilesiz) olup olmadığını araştırmak için para 100 defa atılsın. Gelen turaların (veya yazıların) sayısı gözlenerek paranın düzgün (hilesiz) olup olmadığı araştırılmak istendiğinde, eğer gelen turaların oranı  ye yakın ise paranın düzgün olduğu söylenebilir. Gelen turaların sayısı 40 dan az ve 60 dan fazla ise paranın düzgün olduğu yokluk hipotezi red edilir şeklinde bir kural da oluşturulabilir. O zaman, örneğin 60 dan fazla tura gözlendiğinde paranın düzgün olmadığı sonucuna varılır. Oysa, para düzgün olmasına rağmen, 100 defa tura gözlenmesi olasıdır. Dolayısı ile, burada da kontrol edilemeyen bir hata yapılmıştır. Bu hata birinci tür hatadır. Yani, gerçekte,  doğru olmasına rağmen gözlem değerlerine göre  yokluk hipotezi red edilmiştir.

Bir sınıftaki istatistik dersinin başarısı ile ilgilendiğimizi düşünelim. Bir grup öğrenci dersi veren öğretim üyesinden şikayetçi olmak amacı ile sınıf ortalamasının oldukça düşük olduğunu iddia ediyor olsun. Gerçekte sınıf ortalaması yüksek olabilir. Oysa şikayete gelen öğrenciler düşük not alan öğrenciler olacağından, bu öğrencilerin notları dikkate alındığında öğrencilerin iddiası red edilemez. Dolayısı ile, gerçekte sınıf ortalaması yüksek olmasına rağmen, şikayete gelen öğrencilerin notları dikkate alındığında öğrencilerin iddiası red edilemez. O halde, bir hata yapılmıştır. Ancak, burada örnek değerler olarak şikayete gelen öğrencilerin notları dikkate almak istatistiki olarak anlamlı değildir. Bu durumda yapılan hata değil, yanlıştır. Oysa, rasgele bir örnekleme yapıldığında, örnekleme bu öğrenciler de çıkabilir. Bu hatalar aşağıda tablo halinde verilmiştir.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Karar | | |
| Doğru Hipotez |  | red edilemez | red edilir |
| doğrudur | Doğru Karar | *Birinci Tür Hata* |
| doğrudur | *İkinci Tür Hata* | Doğru Karar |

Hipotez testlerinde amaçlarından biri, bu hatalar minimum olacak şekilde yöntemlerin ya da test kurallarının geliştirilmesidir. Hataları minimum yapmak yerine hata olasılıkları minimum olacak yöntemlerin geliştirilmesi üzerinde durulur. Genel olarak birinci tür hata olasılığı sabit tutularak ikinci tür hata olasılığı en küçük (minimum, yani gücü en fazla) olacak şekilde yöntemler geliştirilmeye çalışılır.

**Tanım 9.3.1** Birinci tür hata olasılığına *testin anlam düzeyi* denir ve  ile gösterilir

**Tanım 9.3.2**  yokluk hipotezinin red edilmesi olasılığına *testin gücü* denir

Bir testin gücü parametrenin fonksiyonu olup, genelklikle  ile gösterilir. Yani,

 karşı 

hipotez testi problemi için  (İkinci tür hata) olmak üzere testin güç fonksiyonu,



şeklinde ifade edilir.

**Örnek 9.3.1**  beklenen değeri , varyansı  olan normal dağılımdan bir örneklem olsun.  hipotezini  alternatif hipotezine karşı test etmek isteyelim. Birinci tür hata olasılığının en fazla  ve  için ikinci tür hata olasılığının da en fazla  (yani  da testin gücünün en az  olması arzu edildiğinde  örneklem hacmini yaklaşık olarak bulmak isteyelim.

Böyle bir problem için test istatistiğinin,



şeklinde olduğunu biliyoruz (Örnek (9.2.1b)). Buna göre, testin birinci tür hata olasılığı için

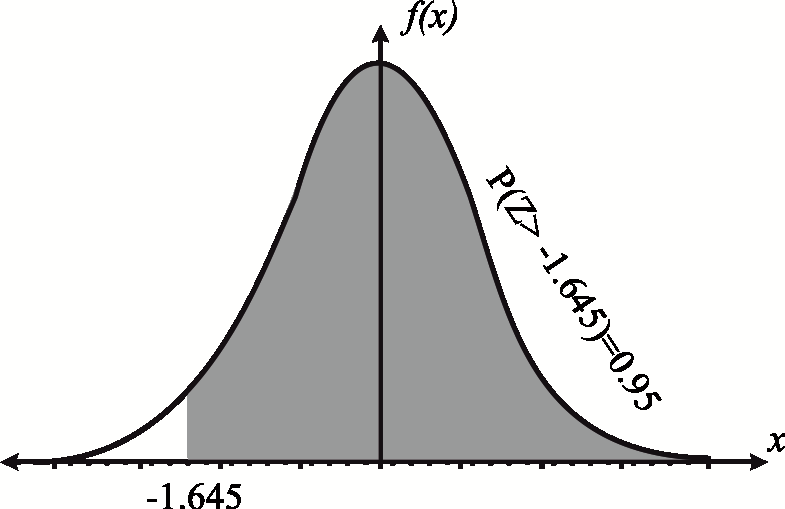


eşitliğinden  bulunur. Yani, dir.

Ayrıca, testin gücünün en az  olması beklendiğinden,  olması için,



olup, normal dağılım tablosundan  ya da  bulunur. Buna göre,  eşitliğinden  ya da  bulunur. Ancak,  örneklem hacmi olduğundan bir tamsayı olmak zorundadır. Yani,  alınmalıdır.



*Şekil 9.3.1 Örnek (9.3.1) deki hipotez testi problemine ait bölgenin alanı*

O halde, birinci tür hata olasılığının en fazla  ve ikinci tür hata olasılığının da en fazla  olması için deneyin en az  defa tekrarlanması gerekir. Yani,  dür

Hata olasılıkları ve testin gücü hipotezlere ve parametrelere bağlı olduğu gibi, örneklem hacmine de bağlıdır. Bilindiği gibi, test kuralının belirlenmesi önemlidir. Bunun için de, önceden belirlenen bir  değeri testin anlam düzeyi olarak alınır ve  sabitinin değeri seçilen  değerine (birinci tür hata olasılığına) bağlı olarak belirlenmeye çalışılır. Buradaki  sayısına *testin kritik değeri* denir. Bu kritik değerler, normal dağılımlı kitleler için tablolardan bulunur. Bazen, örneklem hacmi yeterince büyük ve merkezi limit teoreminin koşulları geçerli ise, yine dağılım tabloları (beklenen değer için normal ve dağılım tabloları, varyans için ki-kare dağılım tablosu, varyansların karşılaştırılması gibi durumlarda da  dağılım tablosu) kullanılır. Diğer durumlarda bu kritik değerler ayrıca belirlenmelidir.

**Örnek 9.3.2**  parametresi  olan ötelenmiş üstel dağılımdan (Örnek 9.2.3)) bir örneklem olsun.  lerin olasılık yoğunluk fonksiyonu,



olup önceden belirlenen  sayısı için  yokluk hipotezini  alternatif hipotezine karşı test etmek isteyelim.  olmak üzere test fonksiyonunu



olarak yazmıştık (Örnek(9.2.3)). Buradaki  değerini belirlemek için, testin anlam düzeyi olan belirlenen bir  sayısı (birinci tür hata olasılığı) önceden belirlenir. Buradan da,



eşitliğinden,  bulunur. Görüldüğü gibi, kritik değer  örneklem hacmi ile önceden belirlenen  sayısına bağlıdır

Testin gücü, parametrenin ve örneklem hacminin bir fonksiyonudur. Bu fonksiyon, bazen analitik olarak elde edilebilir. Bazen, parametrenin monoton artan veya azalan bir fonksiyonu olarak da karşımıza çıkabilir. Örneklem hacmi büyüdükçe testin gücü de artar.

**Örnek 9.3.3**  dağılımından bir örneklem  ve  biliniyor olsun.  yokluk hipotezinin  alternatif hipotezine karşı test edilmesi problemini ele alalım. Bu problem için olabilirlik oran testi,



şeklindedir. Birinci tür hata olasılığı (testin anlam düzeyi)  önceden belirlenmiş olsun. Buna göre,

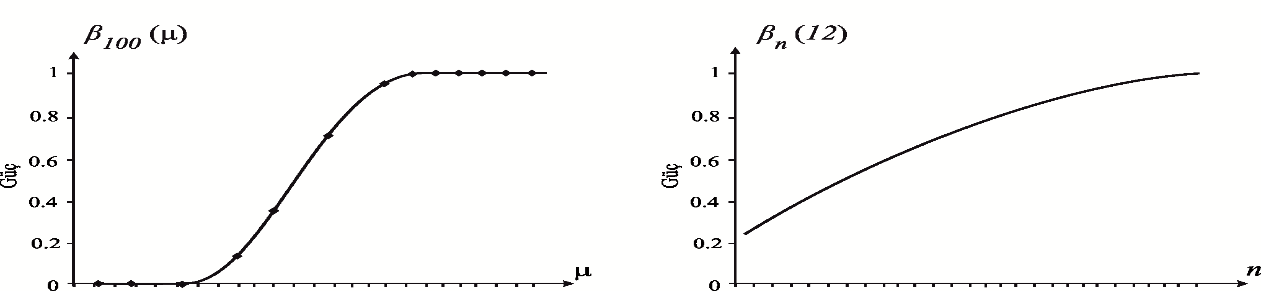


olup normal dağılım tablosundan  olacak şekilde  sayısı () belirlenir. Bu hipotez testi problemi için güç fonksiyonu,



eşitliğinden  şeklindedir. Önceden belirlenen  ve  değerleri yerine konarak güç fonksiyonunun değerleri, normal dağılım tablosundan bulunarak aşağıdaki tabloda verilmiştir. Testin güç fonksiyonunun grafiği de  ve  ye göre aşağıdadır (Şekil (9.3.2)). Tablo ve grafiklerden de görüldüğü gibi, testin güç fonksiyonu  ve  ye göre artandır. Güç fonksiyonu, örneklem hacmine göre artan olmasına rağmen, alternetif hipotezin yönüne bağlı olarak  ye göre artan ya da azalan olabilir. Tablo değerleri ,  ,  ve  için hazırlanmıştır.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|  | 0.00013 | 0.00408 | 0.04998 | 0.25946 | 0.63871 | 0.91229 | 0.99074 |
|  | 15 | 16 | 17 | 18 | 19 | 20 |  |
|  | 0.9996 | 0.99999 | 1.000 | 1.000 | 1.000 | 1.000 |  |



|  |  |
| --- | --- |
| *sabit,  ye göre güç fonksiyonu* | *sabit,  ye göre güç fonkaiyonu* |

*Şekil 9.3.2 Güç Fonksiyonunun Grafiği*

Açıkça görüldüğü gibi,

 olup, ve 

dir.  için  olup,  ve  için  örneklem hacmi sabit tutularak  ye göre,  sabit tutularak da  ye göre güç fonksiyonunun grafikleri Şekil (9.3.2) de verilmiştir. Her iki grafikte de güç fonksiyonu  ve  nin artan bir fonksiyonudur.

Aynı örneklem dikkate alınarak,  yokluk hipotezinin  alternatif hipotezine karşı test edilmesi problemini ele alalım. Bu durumda test fonksiyonu,



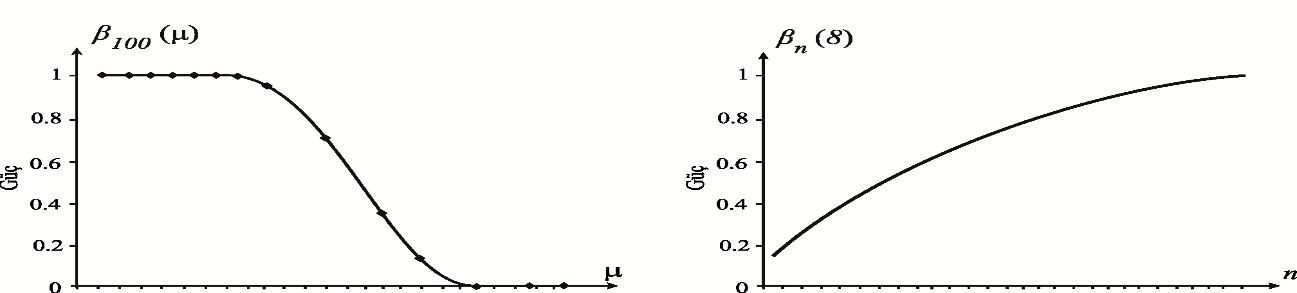
olup birinci tür hata olasılığı  nın önceden belirlendiğinde  kritik değeri,



eşitliğinden  olarak belirlenir. Güç fonksiyonu ise,



eşitliğinden  şeklinde olur. Yine,  için  olup,  ve  için,  örneklem hacmi sabit tutularak  ye göre,  sabit tutularak  ye göre testin güç fonksiyonunun grafikleri (Şekil (9.3.3)) aşağıdadır. Burada güç fonksiyonu  ye göre azalan,  örneklem hacmine göre artandır. Yani her iki durumda da testin güç fonksiyonu,  örneklem hacminin artan bir fonksiyonudur.



|  |  |
| --- | --- |
| *sabit,  ye göre güç fonksiyonu* | *sabit,  ye göre güç fonkaiyonu* |

*Şekil 9.3.3 Güç Fonksiyonunun Grafiği*

Aynı örneklem kullanılarak,  yokluk hipotezinin  alternatif hipotezine karşı test edilmesi problemi için test fonksiyonunun,



şeklinde olduğunu biliyoruz.  birinci tür hata olasılığı önceden belirlendiğinde kritik değer normal dağılım tablosundan  olarak bulunur.  standart normal rasgele değişkenin dağılım fonksiyonunu göstermek üzere testin güç fonksiyonu,





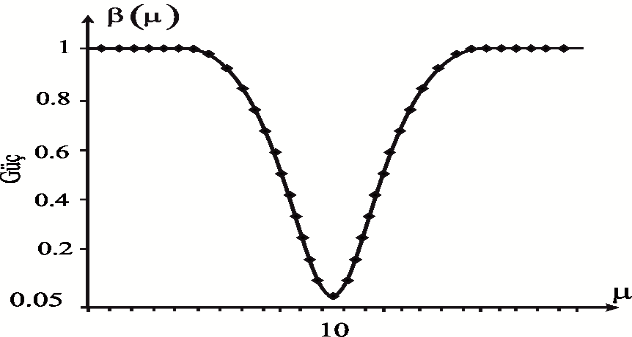




eşitliğinden

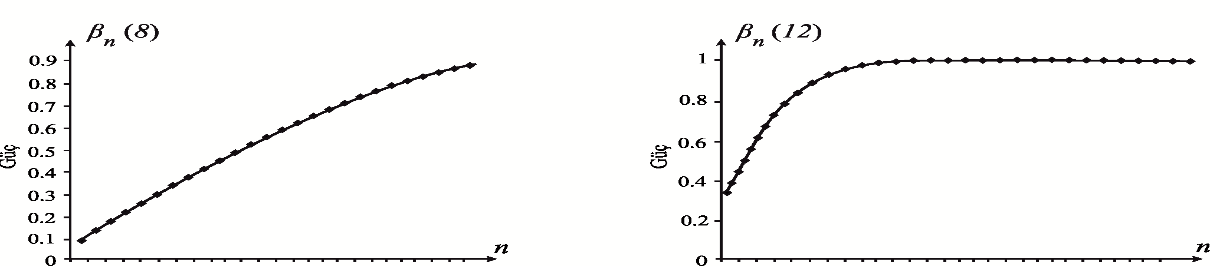


olarak elde edilir.



*Şekil 9.3.4 Çift yölü hipotez testi problemi için güç fonksiyonu*

Testin güç fonksiyonunun grafiği yukarıda Şekil (9.3.4) de verilmiştir. Grafikten de görüldüğü gibi, güç fonksiyonu belli bir yere kadar azalmakta ( noktasına kadar) daha sonra artmaktadır. Fonksiyonun  noktasındaki değeri  dir



|  |  |
| --- | --- |
| Testin gücü ( sabit) | Testin gücü ( sabit) |

*Şekil 9.3.5 Güç fonksiyonunun grafiği*

 için  olup, ,  ve  için güç fonksiyonu  şeklindedir. Alternatif hipotez ister  isterse  olsun, örneklem hacmi arttıkça testin gücü artar. Yukarıda, iki özel durumda için grafikler verilmiştir

Hipotez testlerinde, testin anlam düzeyi yani birinci tür hata olasılığı çok önemlidir. Hemen hemen bütün istatistiki sonuç çıkarımlar testin anlam düzeyine bağlıdır. Ayrıca testin anlam düzeyi güç fonksiyonunun  hipotezi altındaki değeridir. Fonksiyonun  hipotezi altındaki değerine göre  düzeyli ve ölçekli testler tanımlanır.

**Tanım 9.3.3** Güç fonksiyonu  olan bir teste,  ise *ölçekli* (size-) *test*,  ise *düzeyli* (level-) *test* denir

**Örnek 9.3.4** a) Bir istatistik sınavından öğrencilerin aldığı notlar beklenen değeri , varyansı  olan normal dağılıma sahiptir. Aşağıda küçükten büyüğe sıralanmış veriler bulunmaktadır.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Rasgele seçilen 68 öğrencinin sınavda aldığı notlar | | | | | | | | | | | | | |
| 18 | 18 | 20 | 20 | 21 | 21 | 21 | 25 | 26 | 27 | 28 | 29 | 30 | 30 |
| 30 | 30 | 30 | 30 | 30 | 31 | 31 | 32 | 33 | 33 | 33 | 35 | 35 | 35 |
| 36 | 36 | 36 | 36 | 38 | 38 | 38 | 38 | 38 | 39 | 40 | 40 | 40 | 40 |
| 40 | 41 | 45 | 45 | 45 | 45 | 46 | 46 | 47 | 47 | 48 | 49 | 50 | 50 |
| 51 | 51 | 51 | 53 | 56 | 56 | 56 | 59 | 59 | 63 | 64 | 67 |  |  |

Verilerden,

 ve 

değerleri hesaplanmıştır.  yokluk hipotezi  alternatif hipotezine karşı  anlam düzeyinde testi problemi için test kuralı,



şeklinde olup  ise  yokluk hipotezi red edilir.

Buradan,  olup,  olduğundan  yokluk hipotezi  alternatif hipotezine karşı red edilir. Testin güç fonksiyonu,



şeklinde olup bazı  değerleri için güç fonksiyonun değerleri aşağıdadır. Bu verilere göre, sınıf ortalamasının  den küçük ya da eşittir yokluk hipotezi  anlam düzeyinde  alternatif hipotezine karşı red edilmiştir. Hesaplanan ortalama  olup testin bu noktadaki *(empirik) gücü*  dir.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 30 | 32 | 34 | 35 | 36 | 38 | **38.39** | 39 | 40 |
|  | 0.00 | 0.00 | 0.0098 | 0.04995 | 0.169 | 0.66 | ***0.85*** | 0.87 | 0.96. |

b)  parametresi  olan Poisson dağılımından bir örneklem olmak üzere  yokluk hipotezi  alternatif hipotezine karşı test edilmek istensin. Bu hipotez testi problemi için test kuralı da,  ve  nin gözlem değeri de  olmak üzere



şeklinde verilmiş olsun. Yani,  ise  hipotezi red edilecektir. Ayrıca,  olup testin anlam düzeyi,



olarak hesaplanır. Bu test yerine,



alınmış olsaydı, testin anlam düzeyi bu defa



olacaktı. Böyle bir hipotez için  anlam düzeyinde bir test oluşturmak istendiğinde, bu testlerden ikisi de kullanılamaz.  olduğundan  nin değerleri kesirli sayılar olarak da belirlenemez. Bu durumda,  örneklemden bağımsız başarı olasılığı  olan Bernoulli rasgele değişkeni olmak üzere test kuralını,



olarak tanımlayalım. Burada,



olup  testinin anlam düzeyi,



olarak hesaplanır. Buna göre,  gözlendiğinde  yokluk hipotezi red edilir ya da red edilemez. Yani,  gözlendiğinde  olasılıkla  hipotezi red edilir. Bu tür testlere *rasgeleleştirilmiş* (*randomized*) testler denir

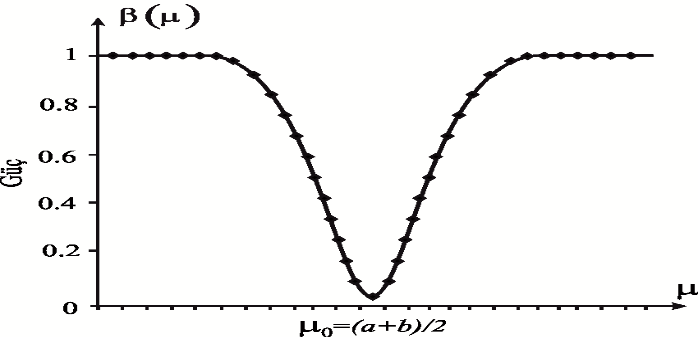
**Örnek 9.3.5**  dağılımından bir örneklem  olsun.  yokluk hipotezini  alternatif hipotezine karşı test etmek için test fonksiyonu,



şeklinde verilsin. Testin güç fonksiyonunun grafiği Şekil (9.3.6) de verilmiştir. Yani, güç fonksiyonu  noktasında en küçük değere ulaşır. Bu değer de testin anlam düzeyidir. Buradaki,  ve  sayıları önceden belirlenmiş olabilir. Böyle bir durumda,  değerini  ve  sayılarına bağlı olarak bulmak isteyelim.  standart normal rasgele değişkenin dağılım fonksiyonunu göstermek üzere testin güç fonksiyonu,



şeklindedir.



*Şekil 9.3.6 Çitf yölü hipotez testi için güç fonksiyonunun grafiği*

Daha açık olarak testin güç fonksiyonu,



şeklinde hesaplanmıştır. Bu fonksiyonu minimum yapan değeri bulabilmek için fonksiyonun birinci türevi sıfıra eşitlenir.  olmak üzere dağılım fonksiyonunun türevi olasılık yoğunluk fonksiyonudur. Dolayısı ile güç fonksiyonunun türevi,



olup  yazılarak  fonksiyonunu minimum (belki maksimum) yapan değer bulunur. Buradan,



elde edilir. Dolayısı ile,



denkliğindeki  ifadesi düzenlenirse  bulunur. Yani,  güç fonksiyonunu minimum yapan değer  dir. Testin güç fonksiyonu  noktasında minimum olduğundan  alınır (Bkz. Şekil (9.3.6)) 

9.4. Bayes Testleri

Bayes çalışmalarının en belirgin özelliği, kitle parametresinin bir rasgele değişken olarak ele alınmasıdır. Burada da, kitle parametresi bir rasgele değişken olarak ele alınacaktır.

 verildiğinde,  parametre değeri  olan kitleden bir örneklem,  de  nın yeterli bir tahmin edicisi olsun.  nın olasılık veya olasılık yoğunluk fonksiyonu (önsel dağılım)  olmak üzere  gözlendiğinde,  nın koşullu dağılımını (sonsal dağılım)  ile gösterelim.  yokluk hipotezinin  alternatif hipotezine karşı test edilmesi problemi için sonsal dağılımın olasılık veya olasılık yoğunluk fonksiyonundan,

“ve ”

olasılıkları hesaplanabilir. Bu olasılıklar yerine  ve 

koşullu olasılıkları da hesaplanır. Buna göre Bayes test kuralı,



şeklinde oluşturulur. Yani,  ise  yokluk hipotezi red edilir.  için yeterli bir tahmin edici  ise Bayes test kuralı,



olarak da verilir. Bayes testleri ile ilgili olarak aşağıdaki örnekleri inceleyelim.

**Örnek 9.4.1** (Hogg, McKean ve Craig (2005))  verildiğinde  beklenen değeri  olan Poisson dağılımından bir örneklem,  nın önsel dağılımı da  ve  (rasgele değil) parametreli Gamma olsun. Yani,  olasun.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 11 | 7 | 11 | 6 | 5 | 9 | 14 | 10 | 9 | 5 |
| 8 | 10 | 8 | 10 | 12 | 9 | 3 | 12 | 14 | 4 |

Yukarıdaki verilere göre  hipotezini  alternatif hipotezine karşı test etmek isteyelim. Önsel dağılımın olasılık yoğunluk fonksiyonu  için,



olup,  de  için yeterlidir. Ayrıca,  verildiğinde  olup  nin koşullu olasılık fonksiyonu  için,



şeklindedir. Buradan,  ve  nın ortak olasılık fonksiyonu,  ve  için



şeklinde yazılabilir. Buna göre,  nin marjinal olasılık fonksiyonu ve  verildiğinde  nın koşullu olasılık yoğunluk fonksiyonu (sonsal dağılımın olasılık yoğunluk fonksiyonu) sırası ile,

 ve 

şeklindedir. Buradan, sonsal dağılım  ve  için parametreleri  ve  olan Gammadır. Yani,  dır. ,  ve  nin gözlem değeri  olmak üzere,

, 

şeklinde sonsal dağılımın parametre değerleri hesaplanmıştır. Sonsal dağılım  olup  hipotezi altında



ve  alternatif hipotez altında,



koşullu olasılıkları hesaplanmıştır. Buradan,



olduğundan  yokluk hipotezi red edilemez

**Örnek 9.4.2**  başarı olasılığı  olan Bernoulli dağılımından bir örneklem olsun.  tahmin edicisi  için yeterli olup  nin önsel dağılım  seçildiğinde sonsal dağılımın da  dır (Casella ve Berger, 2002, sayfa 324).

Bir paranın  defa atılması deneyinde  gelen turaların sayısı olsun. Yani,  ve  olmak üzere aşağıdaki sonuçlar gözlenmiş olsun.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |

Bayes yöntemini kullanarak,  hipotezini  alternatifine karşı test etmek isteyelim. Bu durumda,  nin gözlenen değeri  olup sonsal dağılım,  dür.  hipotezi altındaki olasılık da,



integralinin sonucudur.  ve  nın değişik değerleri için integralin sonuçları aşağıdadır.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | 0.3229 | 0.2744 | 0.2314 | 0.1938 |
|  | 0.3824 | 0.3302 | 0.2829 | 0.2406 |
|  | 0.4418 | 0.3872 | 0.3366 | 0.2905 |
|  | 0.5000 | 0.4441 | 0.3915 | 0.3424 |
|  | **0.5558** | 0.5000 | 0.4462 | 0.3953 |

Alternatif hipotez altında ise olasılıklar 

eşitliğinden hesaplanır. Buna göre,  ve  için



olduğundan,  hipotezi red edilemez. Diğer durumlarda ise,  yokluk hipotezi red edilir

Bu örnekte de görüldüğü gibi, parametre hakkındaki herhangi bir sonuç çıkarım önsel dağılımın seçimine bağlıdır. Önsel dağılım veya önsel dağılımın parametreleri değiştikçe parametre hakkındaki sonuç çıkarım değişir. Örneğin, Örnek (9.4.2) de önsel dağılım,



olarak seçilmiş olsaydı,  verildiğinde  nin sonsal dağılımı  için



şeklinde olup sonsal dağılım,  ve  için  olurdu. Bu durumda,  hipotezi altındaki olasılık,



olarak hesaplanacaktı. Bu olasılıklar da  nın bazı değerleri için,

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

şeklinde hesaplanmış ve  olduğundan  hipotezi red edilmiştir.