YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

bayesgil istatistik yöntemleri ile pırl.

Yüksek Lisans Tezi

can cılız

YILDIZ TEKNİK ÜNİVERSİTESİ
KÜTÜPHANE VE DOKÜMANTASYON
DAİRE BAŞKANLIĞI

Yer No (DDC) : R360/17

Kayıt No : 3711
Geldiği Yer : Fen Bilimleri
Enstitüsü
Tarih : 18/12/2007
Fiyat : 5,00 YTL
Fatura No :
Ayniyat No :
Ek :

YILDIZ TEKNİK ÜNİVERSİTESİ
Demirbaş No:............................
25507020399993300700001
XI-101

YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BAYESGİL İSTATİSTİK YÖNTEMLERİ İLE PIRLANTA FİYATININ ANALİZİ

İSTATİSTİKÇİ CAN CILIZ

İstatistik Ana Bilim Dalı'nda
Hazırlanan

YÜKSEK LİSANS TEZİ

Yrd. Doç. Dr. Atıf Evren

Yard. Doç. Dr.
Atıf Evren

Yrd. Doç. Dr.
Osman Zeki Çıtıkaya

M. Genceli

İSTANBUL ,2007

İÇİNDEKİLER

SİMGE LİSTESİ

b	Parametre tahminci vektörü
b_0	β_0'ın tahmincisi
b_1	β_1'in tahmincisi
I	Birim Matrisi
L	Olabilirlik fonksiyonu
r^2	Belirlilik katsayısı
R^2	Çoklu belirlilik katsayısı
$s(b_0)$	b_0'ın standart hatası
$s(b_1)$	b_1'in standart hatası
$S(\theta)$	Kalıntı kareleri toplamı
$Var(b_0)$	b_0'ın varyansı
$Var(b_1)$	b_1'in varyansı
X	Bağımsız veya açıklayıcı değişken
X	Veri matrisi
Y	Bağımlı veya açıklanan değişken
$\hat{Y}$	Y'nin tahmin değeri
Y	Gözlem vektörü
β_0	Sabit terim parametresi
β_1	Eğim parametresi
β	Parametre vektörü
L(θ)	en çok olabilir tahmin edicisi
X	boyutunda bağımsız değişken matrisi
Y	boyutunda bağımlı değişken vektörü
β	boyutunda parametreler vektörü
U	boyutunda hata terimleri vektörüdür.
Y_i	Veri matrisi
θ	parametre vektörü
ε	Hata terimleri
ε	Hata terimleri vektörü
μ	Anakütle ortalaması
γ	Doğrusal olmayan regresyonda parametre vektörü
σ^2	Hata payı varyansı
$\hat{\sigma}^2$	σ^2'nin EKK tahmincisi
integral (∫)	integral
$\in$	eleman
$\cap$	kesişim
$\varnothing$	boşküme
|	fark
Ω	parametreler uzayı
ξ (θ)	prior olasılık yoğunluk fonksiyonu

$\xi(\theta/x)$	posterior olasılık yoğunluk fonksiyonu
$g_n(x)$	son dağılım
B’	sıklık tahmini
$\sum$	toplama
Θ_1	tarafsız tahminci
$f(x/\theta)$	olasılık fonksiyonu
$f_n(x/\theta)$	olasılık yoğunluk fonksiyonu
$\delta^*(X)$	bayes tahmin edicisi
P_1	hiyerarşik dağılım
$f(\theta)$	ön dağılım
$f(\theta/y)$	son dağılım
$f(y/\theta)$	benzerlik fonksiyonu
$P(\theta)$	kesikli fonksiyon

KISALTMA LİSTESİ

BLUE	en iyi doğrusal sapmasız tahmincileri (Best Linear Unbiased Estimators)
VAR	varyans
COV	kovaryans
EÇO	en çok olabilirlik tahmin edicileri

ŞEKİL LİSTESİ

ÖNSÖZ

Bu tezi yazarken bana destek olan danışman hocam Yrd.Doç.Dr Atıf Evren'e şükranlarımı sunarım.Ayrıca Yrd.Doç.Dr Doğan Yıldız'a bana olan desteğinden dolayı teşekkür ederim.Özel olarak teşekkür edeceğim diğer bölüm hocalarıma da bende olan emeklerinden dolayı şükranlarımı borç bilirim Elif Öztürk ,Fatma Noyan,Gülder Kemalbay ve diğer asistanlara bana olan yardımlarından dolayı teşekkür ederim.Ayrıca Okay Şimşek'e de teşekkür ederim..Evde hem babalık hem de annelik yapan anneme Selva Cılız'a ve bana zaman ayırmama izin veren kardeşlerime teşekkür ederim.

Can Cılız

ÖZET

Bu çalışmanın amacı Bayesgil regresyon analizi ile pırlanta fiyatlarının incelenmesidir. Çalışmada diğer parametre tahmin yöntemlerinden de bahsedilmektedir. En çok olabilirlik metodu üzerinde özellikle durulmuş olup Bayes yaklaşımı ile olan ilgisinden söz edilmiştir. Ayrıca prior ve posterior dağılım kavramları üzerlerinde de durulmuştur. Prior dağılımların posterior dağılım üzerine olan etkileri anlatılmaya çalışılmıştır.

Çalışmanın uygulama kısmında pırlanta fiyatlarının belirlenmesinde etkin olan faktörlerden yola çıkılarak bir regresyon modeli kurulmuştur. Prior dağılım bilgileri ile örnekten gelen bilgiler birleştirilerek yeni bir regresyon modeli oluşturulmuştur. Böylelikle regresyon modelinin parametreleri için elde edilen güven aralıklarının uzunluklarının azaldığı gözlenmiştir.

Anahtar Kelimeler: prior dağılım, posterior dağılım, bayesgil regresyon

ABSTRACT

The aim of this study is to analyse the factors which determine the levels of the prices of diamonds by a Bayesian regression analysis. Some parameter estimation techniques are also mentioned in this study. The importance of the maximum likelihood method in parameter estimation is emphasized and the relation between the likelihood function and Bayesian approach is put forward. Besides some concepts of prior and posterior distributions are discussed to some extent. Some of the affects of prior distributions on posterior distributions are exposed. On the empirical part of this study, a regression model is constructed by using the factors which determine the price levels of diamonds. By unifying the information obtained from prior assumptions and the data at hand a new regression model is developed. Thus it is observed that the lengths of the confidence intervals for the parameters have been reduced.

Key Words: prior distribution , posterior distribution , bayesian analysis

GİRİŞ

Tezin ilk bölümlerinde olasılık kavramları olasılık dağılımlarını sürekli kesikli dağılımları ve özel dağılımlarını anlattım. Daha sonra istatistikte kullanılan bazı parametre tahmin yöntemlerinden söz ettim. Bayesgil metodolojinin prior dağılım, posterior dağılım, eşlenik dağılımlar , kayıp fonksiyonu gibi temel kavramlarını tartışmaya çalıştım.
Çalışmanın üçüncü bölümünde ise klasik regresyon analizinin bazı varsayımlarından ve metodolojisinden bahsettim.

Uygulama olarak pırlanta fiyatının meydana gelmesini sağlayan parametreleri bayesgil regresyon metoduyla analiz ettim.Bayesgil regresyon sonuçlarını klasik regresyon sonuçları ile karşılaştırdım. Bunu yaparken Minitab ve Excel programlarından yararlandım.
Taşın fiyatıyla özellikleri arasındaki ilişkiyi önce klasik regresyonla analiz ettim . Çıkan sonucu modele ekleyerek bayesgil regresyon modeli oluşturdum. Sonuç olarak pırlanta taşında tecrübeli insanların görüşlerinin a priori bir bilgi olarak modele dahil edilmesinin daha uygun bir regresyon modeli oluşturmak açısından yararlı olacağı sonucuna vardım.

BÖLÜM 2: OLASILIK DAĞILIMLARI VE OLASILIK YOĞUNLUKLARI

2.1. Olasılık Tanımı

Olasılık, bir belirsizlik ölçütüdür. Olasılık kavramı, bir olayın gerçekleşebilmesinin sayısal bir ölçüsünü verir. 0 ile 1 arasındaki bir ölçekle gösterilir. 0 olasılığı olayın gerçekleşmesinin imkansız olduğunu, 1 olasılığı ise olayın gerçekleşmesinin kesin olduğunu gösterir.

Olasılıkları tanımlamanın en eski yolu olan "klasik olasılık kavramı", olanaklı bütün sonuçların şansının eşit olduğu zaman uygulanır. Bu durumda, aralarından birinin gerçekleşmek zorunda olduğu N tane eşit şanslı durum varsa ve bunlardan n tanesi lehte ya da başarı olarak görülüyorsa, o zaman bir başarının olasılığı “n/N” oranıyla gösterilir.

Klasik olasılık kavramının temel aksaklığı kullanım alanının sınırlı olmasıdır, çünkü pek çok durumda ortaya çıkan olanaklar eşit şanslı değildir.

Olasılık kavramlarından en çok benimsenenler arasında “sıklık yorumu” yer alır.

2.2. Örneklem Uzayları

Olasılık kavramının daha iyi anlaşılabilmesi için öncelikle deney, sonuç, örneklem uzayı ve olay terimlerinin açıklanması gerekir.

İstatistikte gözleme ya da ölçme süreçlerinden her birine deney denmektedir. Bir deneyden elde edilen bulgular ise deneyin sonuçları diye adlandırılır.

Bir deneyin bütün olanaklı sonuçlarından oluşan kümeye örneklem uzayı denir ve genellikle S harfiyle gösterilir. Bir örneklem uzayındaki her sonuç örneklem uzayının öğesi ya da kısaca örneklem noktası adını alır. Bir deneyi betimlemek için farklı örneklem uzayları kullanılabilir.

Örneklem uzayları çoğunlukla içerdikleri öğe sayısına göre sınıflanırlar. Örneğin bir deneme bir zarın bir kez atılmasından ibaretse ve kaç geleceğiyle ilgileniliyorsa, örneklem uzayı, S={ 1, 2, 3, 4, 5, 6 } olur. Bu örneklem uzayı sonlu sayıda öğe içerir. Diğer bir örnek olarak, madeni bir para yazı gelene kadar atılmaya devam ederse, bu ilk atışta, ikinci atışta, üçüncü atışta, ... olabilir yani sonsuz sayıda durum vardır. Bu durumda örneklem uzayı, S={ Y, TY, TTY, TTTY, ...} olur. Ama buradaki öğelerin sayıları sayma sayıları kümesinin elamanları ie birebir eşleşebilir ve örneklem uzayına sayılabilir sonsuz denir.

Bir örneklem uzayının sonlu sayıda öğesi varsa ya da öğeleri sonsuz sayıda ama sayılabilir ise bu uzaya ayrık ya da kesikli denir.

Bazı deneylerin sonuçlan ne sonludur, ne de sayılabilir sonsuz. Örnek olarak iki kimyasal maddenin tepkime süresi ölçülmek isteniyorsa, örneklem uzayını oluşturan sonuçlar sonsuzdur ve sayılamaz.Örnek olarak zaman,ağırlık,hacim süreklidir. Bir doğru parçası üzerindeki bütün noktalar ya da bir düzlem üzerindeki bütün noktalar gibi bir süreklilik oluşturan örneklem uzayı süreklidir denir.

Bir örneklem uzayının her bir altkümesine olay denir. Bir olay sadece bir örneklem noktası içeriyorsa basit olay, birden fazla örneklem noktası içeriyorsa bileşik olay olarak adlandırılır. Bir deneyin sonucunda gerçekleşme olasılığı 1 olan olaylara kesin olay, gerçekleşme olasılığı 0 olan olaylara ise imkansız olay denir.

2.3. Bir Olayın Olasılığı

A kümesi, S örnek uzayı içerisindeki herhangi bir olay olsun. A olayının olasılığı da P(A) ile gösterilsin. P(A)'nın tanımlanabilmesi ile ilgili şu durumlar geçerlidir.:

• Bir olayın olasılığı negatif değerler almayan reel bir sayıdır, yani S'in herhangi bir A altkümesi için $P(A) \geq 0$'dır.

• $P(S) = 1$

Örnek uzayının gerçekleşme olasılığı 1'dir.

• A_1, A_2, A_3, ... olayları, S'nin bağdaşmayan olaylarının sonlu ya da sonsuz bir dizisiyse,

$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2)\ P(A_3) + \ldots)$

2.4. Bazı Teoremler

Olasılığın üç önermesine dayanarak aşağıdaki önermeler türetilmiştir.

Teorem 1: A ile A', S örneklem uzayında tamamlayıcı(complementary) olaylarsa;

$P(A') = 1 - P(A)$

Teorem 2: Herhangi bir örneklem uzayı S' de $P(\varnothing) = 0$ olur.

Teorem 3: A ile B, S örneklem uzayında iki olaysa ve$A \subset B$ ise,

$P(A) \leq P(B)$ olur.

Teorem 4: Herhangi bir A olayı için $0 \leq P(A) \leq 1$'dir.

Teorem 5: A ve B, S örneklem uzayında herhangi iki olaysa;

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$'dir.

Teorem 6: A, B ve C, S örneklem uzayındaki herhangi üç olay ise;

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

gerçekleşir.Bu durum aşağıdaki gibi genellenebilir:

Tümevarım yardımı ile aynı örnek uzayı S içerisinde yer alan $E_1, E_2, ..., E_n$ olayları için

$$P(E_1 \cup E_2 \cup ... \cup E_n)$$
$$= \sum_{i=1}^{n} P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \sum_{i<j<k<l} P(E_i \cap E_j \cap E_k \cap E_l) +$$
$$...+(-1)^{n+1} P(E_i \cap E_j \cap ... \cap E_n)$$

bulunur. Sözel bir ifade ile **n** olayın birleşiminin olasılığı olayların tek tek olasılıklarının toplamı eksi ikili kesişimlerinin olasılıklarının toplamı artı üçlü kesişimlerinin toplamı … biçiminde hesaplanır.

2.5. Koşullu Olasılık

Bazen, bir olayın gerçekleşme şansı, başka olayların gerçekleşip gerçekleşmemesine bağlı olabilir. Bu gibi durumlarda koşullu olasılık kavramının kullanılması gerekmektedir.

Koşullu olasılığın tanımı şu şekilde yapılabilir; A ile B, S örneklem uzayında iki olaysa ve $P(A) \neq 0$ ise, A verilmişken B' nin koşullu olasılığı

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad ; \quad (P(A) \neq 0 \text{ için}) \text{ olur.}$$

Bazı Teoremler

1. A ile B, S örneklem uzayında birer olaysa ve $P(A) \neq 0$ ise,

$P(A \cap B) = P(A) \cdot P(B \backslash A)$

2. A, B ve C, S örneklem uzayı içinde üç olaysa ve $P(A \cap B) \neq 0$ ise,

$P(A \cap B \cap C) = P(A) \cdot P(B \backslash A) \cdot P(C \backslash (A \cap B))$

2.6. Bağımsız Olaylar

A ile B gibi iki olaydan birinin gerçekleşmesi ya da gerçekleşmemesi ötekinin gerçekleşme olasılığını etkilemiyorsa, bu olaylar birbirinden bağımsızdır. Simgelerle gösterilirse, P(B\A) = P(B) ve P(A\B) = P(A) ise A ve B olayları bağımsızdır.

Bunun sonucunda bağımsızlıkla ilgili şu durum ortaya çıkar;

$P(A \cap B) = P(A).P(B)$ sağlanırsa A ve B bağımsızdır.

Teorem: A ile B bağımsızsa A ile B' de bağımsızdır. Genellemek gerekirse

$$\boxed{\text{A ve B bağımsız} \Rightarrow \begin{cases} A \quad ve \quad B' \\ A' \quad ve \quad B \quad \text{de bağımsız} \\ A' \quad ve \quad B' \end{cases}}$$

olur.

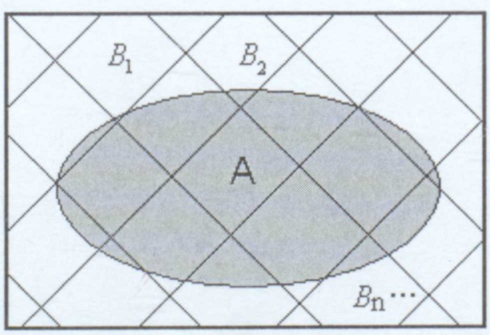
2.7. Bayes Teoremi

Bir deneyin sonucu çoğu zaman ara aşamalarda neler olup bittiğine bağlıdır. $B_1, B_2, \ldots, B_k$ olayları S örneklem uzayının bir parçalanışını oluşturuyorsa ve i = 1, 2, ..., k için $P(B_j) \neq 0$ ise, S içindeki herhangi bir A olayı için;

$$P(A) = \sum_{i=1}^{k} P(B_i).P(A/B_i) \text{ dir.}$$

Bu aşamadan sonra Bayes Teoremine geçilebilir: $B_1, B_2, \ldots, B_k$ olayları S örneklem uzayının bir ayrışımını(partition)[1] oluşturuyorsa, $P(B_i) \neq 0$ $(i = 1, 2, \ldots, k$ için) ise S içinde $P(A) \neq 0$ olan herhangi bir A olayı için (r= 1, 2, ..., k iken)

$$P(B_r / A) = \frac{P(B_r).P(A/B_r)}{\sum_{i=1}^{k} P(B_i).P(A/B_i)}$$



---

[1] $F_1, F_2, \ldots, F_n$ olayları dizisi ; $\bigcup_{i=1}^{n} F_i = S$ ve $F_i \cap F_j = \varnothing$ **i≠j** için koşullarını yerine getiriyorlarsa **S'**in bir **parçalanışı** (partition) olarak adlandırılır.

Şekil.2.1 Bayes Teoremi

$B_i \cap B_j = \varnothing$ ve $\bigcup_{1}^{n} B_i = S$ olursa $B_1, B_2, \ldots, B_n$, S'in bir parçalanışı olarak adlandırılır.

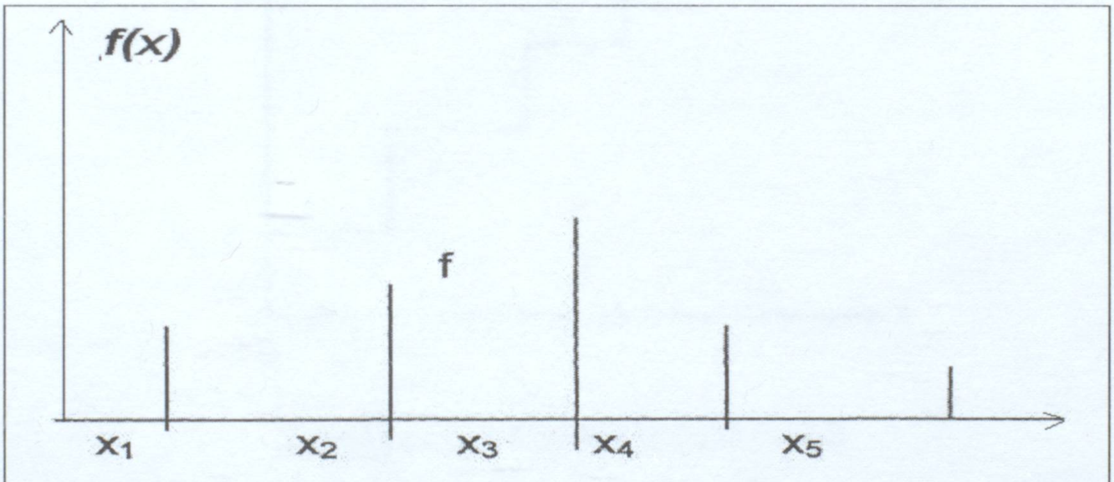
BÖLÜM 3 : KESİKLİ VE SÜREKLİ DAĞILIMLAR

Bir rassal değişken, yalnızca sayılabilir sayıda değerler alabiliyorsa kesiklidir. Örneğin bir zarın atılması olayının yalnızca 6 olası sonucu vardır ve bu sonuçların her birine bir olasılık atanabilir ya da bir paranın sonsuz kez atılmasından doğan yazı sayısı da kesikli bir rassal değişkendir. Çünkü atış sayısı sonsuz olsa da olası sonuçlar sayılabilir[2]. Bir rassal değişken bir aralıktaki bütün değerleri alabiliyorsa süreklidir. Buna da hava sıcaklığı, bir ailenin yıllık geliri gibi örnekler verilebilir Sürekli rassal değişkenlerde belli değerlere olasılıklar verilemez. Hava sıcaklığı değeri tam olarak bilinemez ancak belli bir aralık içinde verilebilir.

3.1. Kesikli Rassal Değişkenlerin Olasılık Dağılımı

Kesikli bir örneklem uzayında tanımlanmış olasılık ölçüsü, bir rassal değişkenin kendi aralığı içindeki herhangi bir değeri alma olasılığını verir. Rassal bir değişkenin değerlerine ilişkin olasılıkları, yani X rassal değişkeninin aralığı içindeki her x değerinin P(X = x)'e eşit olan olasılıkları f(x) fonksiyonuyla gösterilir.

X kesikli bir rassal değişkense, X' in aralığı içindeki her bir x için f(x) = P(X = x) ile verilen fonksiyona X' in olasılık fonksiyonu denir. Rassal bir değişkenin olasılık dağılımı, olası bütün sonuçları için olasılıkların bir gösterimidir.



Şekil 3.2 Tipik Bir Kesikli Olasılık Fonksiyonunun Grafiği

[2] Bir kümenin elemanları sayısı sayma sayıları kümesi ile birebir eşlenebiliyorsa bu küme sayılabilirdir. Sayılabilirlik de kendi içinde sayılabilir ve sayılabilir sonsuz olmak üzere iki grupta incelenebilir. Örneğin sayma sayıları kümesi sayılabilir sonsuzdur.

Bir fonksiyon, ancak ve ancak şu koşullar sağlanırsa kesikli bir X rassal değişkeninin olasılık fonksiyonu olur:

a. Tanım aralığı içindeki her değer için f(x) $\geq$ 0;

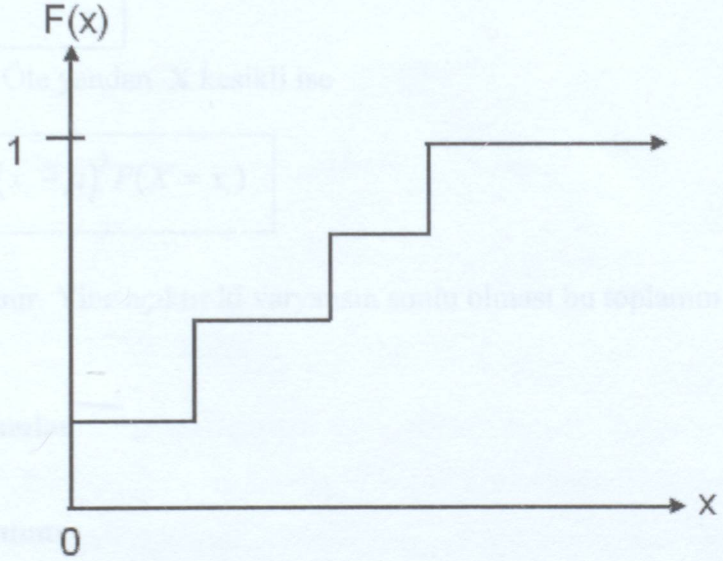
b. $\sum_x f(x) = 1$

X 'in bir x değerine eşit ya da ondan küçük olma olasılığını bulmak için dağılım fonksiyonu denen bir diğer olasılık fonksiyonundan yararlanmak mümkündür. Bu fonksiyon F(x) ile gösterilsin. . Bu durumda, X' in x' e eşit ya da ondan küçük bir değer alma olasılığı

F(x) = P(X < x) ile gösterilsin. Olasılık fonksiyonu f(x), dağılım fonksiyonu F(x) ise iki fonksiyon arasındaki ilişki aşağıdaki gibidir:

$$F(X) = \sum_{\forall t \leq x \quad için}^{x} f(t)$$

Kesikli rassal değişkenler için birikimli olasılık fonksiyonu her zaman 0' dan başlayıp 1' de sona eren bir basamak fonksiyonu biçimindedir.



Şekil 3.3 Kesikli Birikimli Dağılım Grafiği

Kesikli bir X rassal değişkeninin dağılım fonksiyonu F(x) şu koşulları yerine getirmektedir:

1) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$

2) $F(\infty) = \lim_{x \to \infty} F(x) = 1$

3) Herhangi iki gerçek sayı **a** ile **b** için **a<b** ise **F(a)** $\leq$ **F(b)** olur.

Bir rassal değişken olan X' in aralığı $\mathbf{x_1 < x_2 < x_3 < ... < x_n}$ değerlerinden oluşuyorsa,

$\mathbf{f(x_1) = F(x_1)}$

ve $\mathbf{i = 2,3,...,n}$ için $\mathbf{f(x_i) = F(x_i) - F(x_{i-1})}$ olur.

Kesikli Bir Rassal Değişkenin Beklenen Değeri

X kesikli bir rassal değişken, f(x) de bunun olasılık dağılımının x' teki değeriyse X' in beklenen değeri:

$$E(X) = \sum_x x.f(x)$$ olur

Burada her değişkenin zorunlu olarak bir beklenen değeri olduğu söylenemez. Beklenen değerin sonlu olması yukarıdaki toplamın yakınsak olmasına bağlıdır.

Kesikli Bir Rassal Değişkenin Varyansı

X'in ortalaması **μ** olsun. Bu halde **X**'in varyansı **Var(X)** ya da σ^2 ile gösterilir ve

$$\boxed{Var(X) = \sigma^2 = E(X-\mu)^2} \qquad 3.1$$

olarak tanımlanır. Öte yandan **X** kesikli ise

$$\boxed{E(X-\mu)^2 = \sum_i (x_i - \mu)^2 P(X = x_i)} \qquad 3.2$$

toplamı ile hesaplanır. Yine açıktır ki varyansın sonlu olması bu toplamın **yakınsak** olmasına bağlıdır.

Bazı Kesikli Dağılımlar

3.1. Bernoulli Dağılımı

Bir rassal deneyin sadece iki sonucu olsun. Bu sonuçlardan bir tanesi "başarı" durumuna; bir diğeri de "başarısızlık" durumuna karşılık gelsin. Her denemede sözkonusu başarı olasılığı (ve dolayısıyla başarısızlık olasılığı) sabit kalsın. Bu durumda bu deneye Bernoulli deneyi denir. Her denemedeki başarı olasılığı p ise başarısızlık olasılığı da 1-p olmaktadır.

Bernoulli deneyinde başarı sonucunu X=1 durumu ile özdeşleştirelim.. Diğer sonuç da X=0 şeklinde ifade edilecektir. Bu durumda, X rassal değişkenine Bernoulli değişkeni denir. Başka bir deyişle tek bir denemede gözlenen başarı sayısı Bernoulli rastlantı değişkeni olarak adlandırılacaktır. Bernoulli dağılımı binom dağılımının özel bir halidir.

$$P(x)=\begin{cases} p^x(1-p)^{1-x} & , \text{ x=0,1 için} \\ 0 & , \text{ diğer durumlarda} \end{cases}$$

Not etmek gerekirse Bernoulli dağılımının tek bir parametresi vardır; o da p dir.

3.2. Binom Dağılımı

n bağımsız Bernoulli deneyi tekrar edildiğinde karşımıza Binom dağılımı çıkmaktadır. . Burada da Bernoulli dağılımındaki gibi iki sonucu olan deneylerde tek bir denemedeki başarı olasılığı p denemeden denemeye sabit kalmaktadır. Ve denemeler birbirinden bağımsızdır. Başka bir ifade ile N denemedeki başarı sayısı aranıyorsa bu problem Binom olasılık dağılımı ile çözülebilir. X'in olasılık fonksiyonu

$$p(x)=\begin{cases} \binom{n}{x} p^x q^{n-x} & x=0,1,2,... \\ 0 & aksi \quad durumlarda \end{cases} \qquad 3.3$$

şeklindedir.

Tanım 3.2. Binom Dağılımı

X rassal değişkeninin olasılık fonksiyonu aşağıdaki gibi olsun.

$$p(x)=\begin{cases} \binom{n}{x} p^x q^{n-x} & , \text{ x=0,1,2,…,n} \\ 0 & , \text{ diğer durumlarda} \end{cases}$$

Aslında n ve p nin alabileceği değerlere göre sayısız binom dağılımı vardır. Bu nedenle binom dağılımını belirleyen n ve p değerleri, aynı zamanda bu dağılımın parametreleridir. $\mathrm{P} \cong 1-\mathrm{p}$ durumunda binom dağılımı simetrik olup, $\mathrm{P} \neq \mathrm{p}$ için simetriden uzaklaşır. N sabit kaldığında $\mathrm{p} \rightarrow 0,5$ için ve p sabit kaldığında $\mathrm{n} \rightarrow \infty$ için dağılım simetriye yaklaşır[3].

Binom dağılımında yer alan binom katsayıları $\binom{\mathrm{n}}{\mathrm{x}}$ in deney sayısı n artıkça hesaplanması zorlaşır. Bu nedenle, n ve p'nin değişik değerlere göre hazırlanmış çizelgeleri vardır ve onlardan yararlanılır. Bu dağılımın beklenen değeri ve varyansı şu şekildedir.

$$E(X)=\sum x.f(x)=np$$

[3] Bu gibi durumlarda Binom olasılıkları, normal dağılım yardımı ile de kolaylıkla bulunabilir.

$Var(X) = np(1-p)$

3.3. Genelleştirilmiş Binom (Multinomial) Dağılımı

Binom dağılımı sadece iki olanaklı sonuca dayanan birbirinden bağımsız n tane deney için geçerli idi. Eğer bir deney ikiden fazla sonuç içerirse ve bu deney aynı koşullar altında n defa tekrarlanırsa istenen sonuç, genelleştirilmiş (çok terimli) binom dağılımı kullanılırak elde edilir.

Tanım 3.3. Multinomial (Genelleştirilmiş Binom) Dağılımı

Bir deneyin birbiri ile bağdaşmayan s_1, s_2,..., s_k gibi k tane olanaklı sonucu varsa ve bunların olasılıkları $p_1, p_2, ..., p_k$ da denemeden denemeye değişmeden kalıyorsa ve $X_1, X_2, ..., X_k$ rastlantı değişkenleri sırasıyla birinci olanaklı, ikinci olanaklı,…, k. olanaklı sonucun gerçekleşme sayıları ise

$x_j \geq 0, \sum_{j=1}^{k} x_i = n$ ve $\sum_{j=1}^{n} p_j = 1$ olmak üzere, x_1 kez X_1, x_2 kez X_2 s_2,...,x_k kez X_k elde etme olasılığı aşağıdaki gibidir:

$$P(x_1, x_2, ..., x_k) = n! \prod_{j=1}^{k} \left(\frac{p_j^{x_i}}{x_j!} \right) = \frac{n!}{x_1! x_2! ... x_k!} (p_1)^{x_1} (p_2)^{x_2} (p_k)^{x_k} \qquad 3.4$$

Bu tür dağılımlara genelleştirilmiş binom dağılımı denir. k=2 için binom dağılım elde edilir.

Bu dağılımın beklenen değeri ve varyansı şu şekildedir.

$E(X_i) = np_i \qquad i = 1, 2, ..., r$

$$Var(X_i) = np_i(1 - p_i) \qquad i = 1, 2, ..., r \qquad 3.5$$

3.4. Geometrik Dağılım

Arka arkaya n kez tekrarlanan bir Bernoulli deneyi ele alınsın ve ilk istenen sonucun elde edilmesi için gereken deney sayısı X olsun.Bu durumda X'e geometrik rassal değişken denir. Binom dağılımında deney sayısının sabit istenen sonuçların sayısı bir rassal değişken iken; geometrik dağılımda istenen sonucun sayısı bire eşit olmak üzere bir sabit, deneylerin sayısı ise bir rassal değişkendir. X'in olasılık fonksiyonu aşağıdaki gibidir:

$$p(x)=\begin{cases} q^{x-1}.p & , \text{ x= 1,2,3…için} \\ 0 & , \text{ diğer durumlarda} \end{cases} \qquad 3.6$$

Bu dağılımın beklenen değeri ve varyansı şu şekildedir

$$\mathrm{E(X)}=\frac{1}{\mathrm{p}}$$

$$\mathrm{Var(X)}=\mathrm{E(X^2)}-\left[\mathrm{E(X)}\right]^2$$

3.5. Pascal (Negatif Binom) Dağılımı

Geometrik dağılımda önemli olan ilk başarının elde edilmesi için gerekli deney sayısının olasılık dağılımını belirlemektir. Eğer ilk başarı değil de k tane başarı elde etmek söz konusu ise, geometrik dağılımın genelleştirilmiş hali olan Pascal dağılımından yararlanmak gerekir.

Geometrik dağılıma uyan bir deney ele alınsın ve deneye k sayıda başarı elde edilinceye kadar devam edilsin. k. başarının elde edilmesi için gerekli deneylerin sayısı X rassal değişkeni ile gösterilsin. Aynen geometrik dağılımda olduğu gibi negatif binom dağılımında da deneylerin sayısı bir rassal değişken , başarıların sayısı ise sabittir. Olasılık fonksiyonu aşağıdaki gibidir:

$$p(x)=\begin{cases} \binom{x-1}{k-1} p^k (1-p)^{x-k} & , \text{ x=k,k+1,k+2,…} \\ 0 & , \text{ diğer durumlarda} \end{cases} \qquad 3.7$$

Bu olasılık fonksiyonu k ve p değerlerine göre değiştiğinden dolayı, k ve p negatif binom dağılımın parametreleridir. Bu dağılımın beklenen değeri ve varyansı şu şekildedir

$$\mathrm{E(X)}=\mathrm{k}.\frac{1}{\mathrm{p}} \qquad \mathbf{3.8}$$

$$\mathrm{Var(X)}=\mathrm{k}.\frac{\mathrm{q}}{\mathrm{p}^2} \qquad 3.9$$

3.6. Hipergeometrik Dağılım

Kısım (3.2.)'de binom dağılımı incelenirken, bu dağılımın iadeli örnekleme varsayımına dayandığını ve bu sayede olasılıkların denemeden denemeye sabit kaldığı ve örneklem çekimlerinin de biri birinden bağımsız olduğu belirtilmişti. Bazı durumlarda sonlu bir kütleden iadesiz örnekleme gerekli olabilir. Bu durumda hipergeometrik dağılıma başvurmak gerekir. Hipergeometrik dağılımı kullanabilmek için aşağıdaki üç koşulun ortaya çıkması gerekir.

i)Bir deney iki olanaklı sonuca sahipse,

ii)Deneyin tekrarlanma sayısı n sabitse,

iii)Deneylerden birinin sonucu diğer deneyler etkiliyorsa (deneyler bağımlı ise),

N birimlik bir sonlu örneklem uzayının, N_1 tanesi ilgilenilen sonucu N_2 tanesi de ilgilenilmeyen sonucu göstersin. $N = N_1 + N_2$ olur. Bu örneklem uzayından örneklem hacmi n olan birimler grubu iadesiz yöntemle çekilsin ve X rassal değişkeni de n birimlik örneklem içindeki ilgilenilen sonuçların sayısını göstersin. Bu durumda, iadesiz örneklem birimlerinin seçimi birbirleri ile bağımlı olacaktır. Örneğin birinci çekimde istenen sonucun sağlanması $P = N_1 / N$ iken, ikinci örneklem çekiminin olasılığı; istenen sonucun sağlanması veya diğer sonucun sağlanamaması durumuna bağlı olarak $N_1 - 1 / N - 1$ veya $N_1 / N - 1$ olur…

Hipergeometrik dağılımın olasılık fonksiyonu incelenirken olasılığın klasik tanımını kullanmak mümkündür. Bu durumda

$$P(X = x) = \frac{\text{İlgilenilen Sonuçların Sayısı}}{\text{Karşılaşılabilir Sonuçların Sayısı}} \tag{3.10}$$

Örneklem uzayında N tane birim olduğuna ve bunlardan n tanesi seçilebileceğine göre, ayrıca bu seçilişte sıranın önemi de olmadığına göre karşılaşılabilir sonuçların sayısı N'nin n'e göre kombinasyonu yani $\binom{N}{n}$ dir. İlgilenilen sonuçların sayısı x ve örneklem uzayındaki ilgilenilen sonuçlar içeren birimlerin toplamı da N_1 olduğundan; N_1 birimden x tanesinin seçimi $\binom{N_1}{x}$ kombinasyonuna eşittir. Örneklem hacmi n'den x tanesi çıkartılırsa, (n-x) de $N - N_1$ veya N_2 den $\binom{N_2}{n-x}$ kombinasyonu kadar farklı şekilde seçilecektir. O zaman ilgilenilen sonuçların sayısı, istatistikteki çarpım kurallarına göre

$\binom{N_1}{x}\binom{N_2}{n-x}$ olacaktır.

Hipergeometrik olasılık fonksiyonu bu durumda aşağıdaki gibi tanımlanır:

$$P(X=x)=\begin{cases}\dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} & \textbf{\textit{, x= 0, 1, 2, …, n için}}\\ 0 & \textbf{\textit{, diğer durumlarda}}\end{cases}$$

3.11

Hipergeometrik olasılık dağılımı N, N_1 ve n olmak üzere toplam üç tane parametreye sahiptir. Bu dağılımın beklenen değeri ve varyansı şu şekildedir

$$E(X)=\frac{N_1 n}{N}\sum_{y=0}^{n-1}\frac{\binom{N_1-1}{y}\binom{N_2}{n-y-1}}{\binom{N-1}{n-1}}$$ **3.12**

$$\mathrm{Var(X)}=\frac{\mathrm{(N-n)}}{\mathrm{(N-1)}}\mathrm{np(1-p)}$$ 3.13

3.7 Poisson Dağılımı

İstatistik'te olasılık dağılımlarının türetildiği temelde iki tane süreç vardır. Bunlardan bir tanesi Bernoulli Süreci, bir diğeri de Poisson Sürecidir. Poisson dağılımı , Poisson sürecinden yola çıkılarak türetilmiş bir olasılık dağılımıdır.

Poisson dağılımı nadir olayların gerçekleşme olaslılklarının bulunmasında da kullanıldığından "küçük olasılıklar dağılımı" olarak da bilinmektedir. Belli ve çok dar bir zaman aralığında az rastlanan olaylar bu tür bir dağılım yardımı ile modellenebilmektedir. Örneğin, Boğaziçi Köprüsü'nde meydana gelen günlük kazaların sayısı, bir havaalanından her saat kalkan veya inen uçakların sayısı İstanbul Boğazı'ndan bir saatte geçen yabancı gemilerin sayısı vb. gibi.

Poisson dağılımında zaman öyle küçük parçalara bölünür ki bu küçük zaman parçalarında birden fazla olayın gerçekleşmesi beklenmez. Başka bir ifade ile belirlenen o dar zaman birimi içerisinde olay ya gerçekleşir ya da gerçekleşmez. Bu nedenden dolayı binom dağılımı n tane deneydeki başarı sayısı ile ilgilenirken, Poisson dağılımı da belirli bir zaman aralığındaki ilgilenilen sonucun sayısı ile uğraşır. (Bener,2003)

Araştırıcının Poisson dağılımını kullanabilmesi için iki ayrık zaman aralığında ortaya çıkan olayların sayısının birbirlerinden bağımsız kabul edilmesi gerekmektedir.

X rassal değişkeni yukarıdaki özellikleri taşıyorsa, ona Poisson rassal değişkeni ve X'in fonksiyonuna da Poisson dağılımı adı verilir ve X'in olasılık fonksiyonu (λ; λ>0 koşulunu sağlayan bir sabit olmak üzere) aşağıdaki gibi gösterilir:

$$P(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & \text{, x=0,1,2,3… için} \\ 0 & \text{, diğer durumlarda} \end{cases} \qquad 3.14$$

Bu dağılımın beklenen değeri ve varyansı şu şekildedir

$$E(X) = \lambda e^{-\lambda}.e^{\lambda} = \lambda$$

3.15

Var(x)= λ

3.9. Sürekli Dağılımlar

3.9.1. Sürekli Rassal Değişkenlerin Olasılık Dağılımları

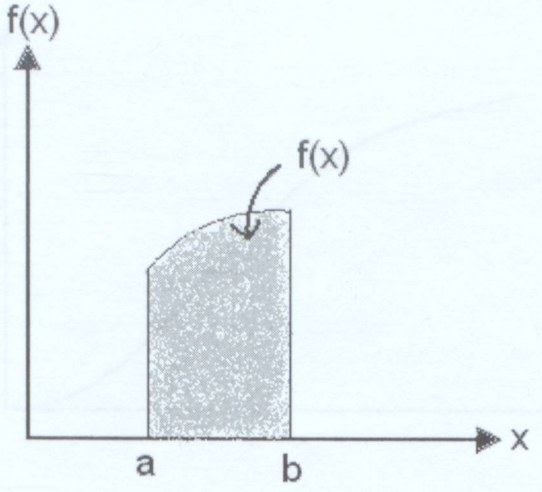
Rassal değişkenlerin sürekli bir ölçek üzerinde her değeri alabildikleri süreklilik durumunda işlemler kesikli durumdakiyle benzerdir. Deney sonuçları doğrular ya da doğru parçaları üzerindeki nokta kümeleri ile gösterilir. Rassal değişkenlerin değerleri de bu noktalar kurallar kümesine ait reel sayılar olarak düşünülür.

Sürekli hallerde P(X=a) olasılığı, kesikli hallerden farklı olarak sıfırdır. $P(a \le X \le b)$ olasılığı

$P(a \le X \le b) = \int_a^b f(x)dx$ integrali ile hesaplanır. Burada f(x) X'in olasılık yoğunluk fonksiyonu olarak adlandırılır.

Bir f(x) fonksiyonu, aşağıdaki koşulları sağlıyorsa sürekli bir X rassal değişkeninin bir olasılık yoğunluğu olarak işlev görebilir.(Hsu ,H.,1997)

1. $f(x) \ge 0$ $\qquad -\infty < x < \infty$ için;
2. $\int_{-\infty}^{\infty} f(x)dx = 1$ $\qquad$ 3.16



Şekil 3.4 Sürekli Olasılık Dağılımı Grafiği ($P(a \le X \le b) = \int_a^b f(x)dx$ olmaktadır.)

Daha önceden de değinildiği gibi X' in c' deki olasılık yoğunluk değerini gösteren f(c), kesikli durumdakinin tersine, P(X=c) yi vermez. Sürekli rassal değişkenlere ilişkin olasılıklar her zaman aralıklara atanır ve her gerçek değerli c sabiti için P(X=c)=0' dır.

X sürekli bir rassal değişken ve bunun olasılık yoğunluğunun t_1 deki değeri f(t) ise;

$$F(x) = P(X \le x) = \int_{-\infty}^{\infty} f(t)dt \qquad -\infty < x < \infty \text{ için} \qquad 3.17$$

fonksiyonuna X' in dağılım fonksiyonu ya da birikimli dağılım fonksiyonu denir.

Tanımdan da anlaşılabileceği gibi dağılım fonksiyonu X' in verilen bir değere eşit ya da küçük olma olasılığını verir.

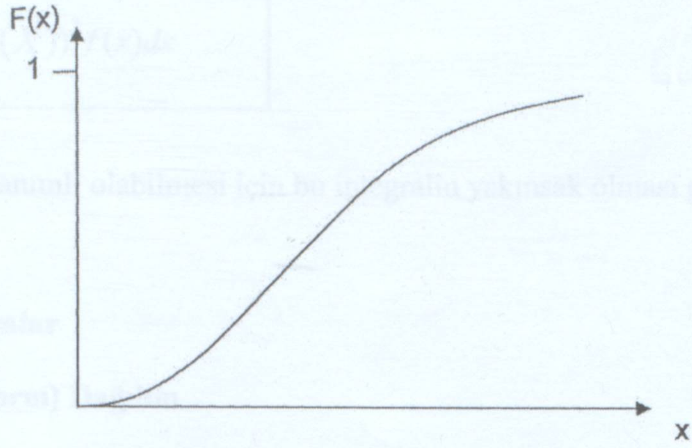
Sürekli hallerde, a < b koşulunu sağlayan a ile b gibi iki sabit için

$$P(a \le X \le b) = F(b) - F(a) \qquad 3.18$$

şeklinde dağılım fonksiyonundan hareketle de bulunabilir. Yine olasılık yoğunluk fonksiyonu ile dağılım fonksiyonu arasında aşağıdaki eşitlikler geçerlidir:

i) $f(x) = \dfrac{dF(x)}{dx}$

ii) $F(x) = \int_{-\infty}^{x} f(t)dt$ 3.19



Şekil 3..5 Sürekli Birikimli Dağılım Grafiği

Sürekli dağılımlarda da dağılım fonksiyonu F(x) şu koşulları yerine getirecektir:

1) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$

2) $F(\infty) = \lim_{x \to \infty} F(x) = 1$

3) Herhangi iki gerçek sayı **a** ile **b** için **a<b** ise **F(a)** $\leq$ **F(b)** olur.

Sürekli Dağılımların Beklenen Değeri

Benzer biçimde, X sürekli bir rassal değişken, f(x) de bunun olasılık yoğunluğunun X' teki değeriyse X' in beklenen değeri:

$E(X) = \int_{-\infty}^{\infty} x.f(x)dx$ olur. 3.20

Bir kez daha beklenen değerin tanımlı olabilmesi için yukarıdaki integralin yakınsak olması gerektiğini vurgulamak gerekir. Örneğin Cauchy dağılımı (1 serbestlik dereceli student-t dağılımı) için beklenen değer mevcut değildir.

Sürekli Dağılımlarda Varyans

X bir sürekli rastlantı değişkeni olsun. bu durumda **X**'in varyansı **Var(X)** biçiminde sembolize edilir ve şu şekilde tanımlanır:

$Var(X) = E(X - E(X))^2$ 3.21

$$Var(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx \qquad 3.22$$

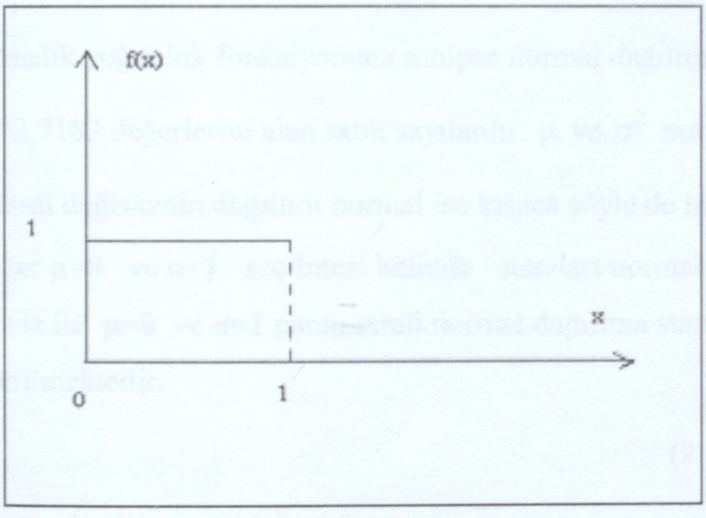
Şüphesiz varyansın tanımlı olabilmesi için bu integralin yakınsak olması gerekir.

Bazı Sürekli Dağılımlar

3.9.2.Düzgün (Uniform) Dağılım

X sürekli rastlantı değişkeni **(α,β)** aralığında aşağıdaki olasılık yoğunluk fonksiyonuna sahip ise düzgün dağılan bir rastlantı değişkeni olarak adlandırılır:

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \alpha < x < \beta \\ 0 \quad Aksi & halde \end{cases} \qquad 3.23$$



Şekil 3.5: alfa=0 ve beta=1 parametreli düzgün dağılım

Düzgün Dağılımın Birikimli Dağılım Fonksiyonu

$$F(x)=\begin{cases} 0 & x \le \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & \alpha < x < \beta \\ 1 & x \ge \beta \end{cases} \qquad 3.24$$

biçimindedir.

Düzgün dağılımın beklenen değeri ve varyansı aşağıdaki gibidir:

$$E(X)=\frac{\alpha+\beta}{2}$$

$$Var(X)=\frac{(\beta-\alpha)^2}{12} \qquad 3.25$$

Tanım. 3.9. Normal Dağılım

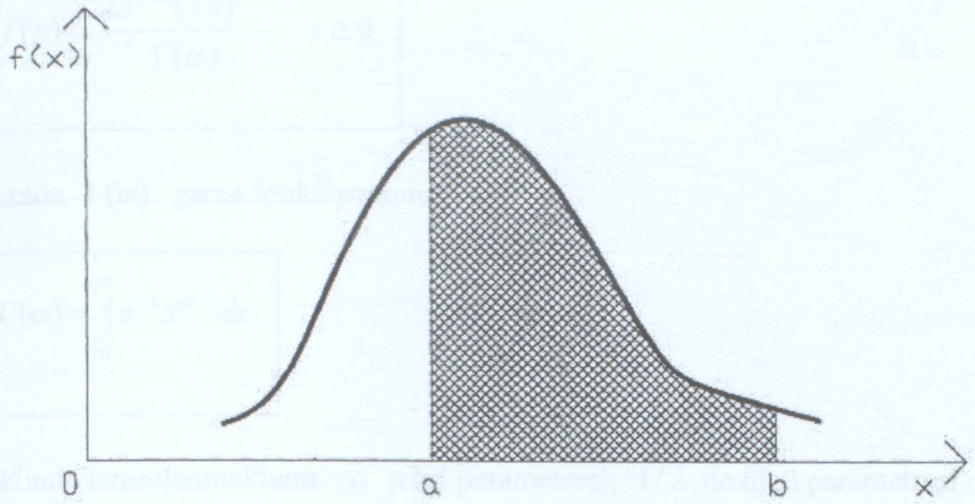
X rasal değişkeni, gerçel sayılar uzayında tanımlanmak üzere,

$$f(x)=\begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & , \sigma > 0, -\infty < x < +\infty \\ 0 & , \textit{diğer durumlarda} \end{cases} \qquad 3.27$$

olasılık yoğunluk fonksiyonuna sahipse normal dağılmıştır denir. Burada $\pi = 3{,}1415$ ve e=2,7183 değerlerini alan sabit sayılardır. μ ve σ^2 normal dağılımın parametreleridir. X rassal değişkenin dağılımı normal ise kısaca şöyle de ifade edilebilir: $X \sim N(\mu, \sigma^2)$

Eğer **μ=0** ve **σ=1** seçilmesi halinde standart normal dağılım elde edilmektedir. Başka bir ifade ile **μ=0** ve **σ=1** parametreli normal dağılıma standart normal dağılım adı verilmektedir.

(2.1)



Şekil3.6:Normal Dağılım

Hem Binom, hipergeometrik, Poisson, Ki-Kare, student-t dağılımı gibi çok kullanılan bazı dağılımların limit halleri normal dağılıma yakınsayacağı için hem de Merkezi Limit Teoremi uyarınca bağımsız dağılan n tane rastlantı değişkeninin toplamının (ya da ortalamasının) olasılık dağılımı, standart normal dağılımla modellenebileceği için normal dağılım İstatistik Teorisi'nde merkezi bir öneme sahiptir

.Normal Dağılımın Beklene Değer Ve Varyansı

$E(x)=\mu$ 3.28

$Var(x)=\sigma^2$

3.9.3 Gama Dağılımı

Olasılık teorisinde önemli rol oynayan dağılımlardan da bir tanesi gama dağılımıdır. Aslında gama dağılımı bir dizi olasılık dağılımının genel halini veren **zarf** niteliğinde bir dağılımdır. Örneğin İstatistik'te özelikle normal dağılan bir anakütleden çekilen örnek varyansının dağılımı Ki-kare dağılımı özel bir gama dağılımıdır. Yine üstel dağılım gama dağılımının özel bir halidir. Bu bakımdan gama dağılımı sözkonusu olasılık dağılımlarının belirli özelliklerinin tartışılmasında da yararlı olmaktadır.

$\lambda > 0, \quad \alpha > 0$ olmak üzere **X** sürekli rastlantı değişkeni aşağıdaki olasılık yoğunluk fonksiyonuna sahip ise gama rastlantı değişkeni olarak adlandırılır:

$$f(x)=\frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \quad x\geq 0 \qquad 3.29$$

Burada $\Gamma(\alpha)$ gama fonksiyonudur ve

$$\Gamma(\alpha)=\int_0^\infty e^{-x}x^{\alpha-1}dx \qquad 3.30$$

şeklinde tanımlanmaktadır. α şekil parametresi; $1/\lambda$ de ölçü parametresi olarak adlandırılır. Tamsayı değerli α parametreli Gama dağılımına **Erlang dağılımı** da denmektedir.

$\alpha=n$ için $\Gamma(\alpha)=(n-1)!$ olduğu gösterilebilir.

Gama Dağılımının Beklenen Değeri ve Varyansı

$$E(X)=\frac{\alpha}{\lambda} \qquad 3.31$$

$$Var(X)=\frac{\alpha}{\lambda^2} \qquad 3.32$$

şeklindedir.

3.9.4. Üstel Dağılım

X rastlantı değişkeni aşağıdaki olasılık dağılımına sahip olsun. Bu durumda **X** üstel dağılıyor denir.

$$f(x)=\lambda e^{-\lambda x} \quad x\geq 0 \qquad 3.33$$

Burada λ pozitif bir sabittir.

Üstel Dağılımın Beklenen Değeri ve Varyansı

$$\mathbf{E(X)}=\frac{1}{\lambda} \qquad \mathbf{3.34}$$

$$Var(X)=\frac{1}{\lambda^2}$$ 3.35

$\alpha=1$ için gama dağılımı üstel dağılıma dönüşmektedir.

3.9.4 Beta Dağılımı

X sürekli rastlantı değişkeni aşağıdaki olasılık yoğunluk fonksiyonuna sahip ise beta dağılımına uyar:

$$f(x)=\begin{cases}\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1} & 0<x<1\end{cases}$$ 3.36

Burada

$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx$$ 3.37

olarak tanımlanmaktadır. **a=b** için beta dağılımı **x=1/2** etrafında simetrik olmaktadır. **b>a** için sola çarpık ve **b<a** için sağa çarpık bir dağılım elde edilmektedir.

Yine **B(a,b)** beta fonksiyonu olarak adlandırılmaktadır. Gama fonksiyonu ile beta fonksiyonu arasında şu ilişki vardır:

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$ 3.38

Beta Dağılımının Beklenen Değeri ve Varyansı

$$E(X)=\frac{a}{a+b}$$ 3.39

$$Var(X)=\frac{ab}{(a+b)^2(a+b+1)}$$ 3.40

3.10. Çok Değişkenli Dağılımlar

Bir rassal değişken, olasılık ölçüsü mevcut olan bir örneklem uzayında tanımlanmış gerçek değerli bir fonksiyondur. Aynı tek örneklem uzayında çok sayıda değişik rassal değişken de tanımlanabilir. X ile Y kesikli rassal değişkenler ise, X' in x değerini, Y' nin de y değerini alma olasılıkları P(X=x, Y=y) olur. Bu durumda P(X=x, Y=y), X=x ile Y=y olaylarının kesişiminin (birlikte) gerçekleşme olasılığıdır.

X ile Y kesikli rassal değişkenlerse, X ile Y' nin aralıklarındaki her (x,y) değer çifti için f(x,y)=P(X=x, Y=y) ile gösterilen fonksiyona X ile Y' nin ortak olasılık dağılımı denir.

İki değişkenli bir fonksiyon, ancak ve ancak bu dağılımın f(x,y) değerleri şu koşulları sağlarsa, X ile Y kesikli rassal değişkenler çiftinin ortak olasılık dağılımı işlevini görür:

1. Fonksiyonun tanım aralığındaki her (x,y) değer çifti için

f(x,y) > 0' dır.

2. $\sum_x \sum_y f(x,y) = 1$, burada çifte toplama işlemi, tanım aralığındaki her

(x,y) değer çiftini kapsar.

X ile Y kesikli rassal değişkenler ise, $-\infty < x < \infty$, $-\infty < y < \infty$ iken

$$F(x,y) = P(X \le x, Y \le y) = \sum_{s \le x} \sum_{t \le y} f(s,t) \qquad 3.41$$

fonksiyonuna, X ile Y' nin ortak dağılım fonksiyonu ya da ortak birikimli dağılım fonksiyonu denir. Burada f(s,t), X ile Y ortak olasılık dağılımının (s,t) noktasındaki değeridir. Buraya kadar bahsedilen çok değişkenli dağılımlar hep kesikli rassal değişkenler içindir. Tek değişkenli durumda olduğu gibi çok değişkenli durumda da kesikli rassal değişkenlerde tanımlanan kavramların sürekli durumdaki karşılıkları mevcuttur.

xy düzleminde tanımlanmış, f(x,y) değerli iki değişkenli bir fonksiyona, ancak aşağıdaki koşullar sağlanırsa, X ile Y sürekli rassal değişkenlerinin ortak olasılık yoğunluk fonksiyonu denir:

$$P[(X,Y) \in A] = \int_A \int f(x,y)\,dxdy \qquad 3.42$$

Bu durum xy düzlemindeki herhangi bir A bölgesi için tanımlanmıştır.

İki değişkenli bir fonksiyon, kendi f(x,y) değerleri şu koşullan sağlıyorsa $-\infty < x < \infty$, $-\infty < y < \infty$ iken, X ile Y sürekli rassal değişken çiftinin ortak olasılık yoğunluk fonksiyonu işlevini görebilir:

1. $f(x,y) \ge 0$ $\quad -\infty < x < \infty$ ve $-\infty < y < \infty$ için

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dxdy = 1$

X ile Y sürekli rassal değişkenler ise, $-\infty < x < \infty$, $-\infty < y < \infty$ iken,

$$F(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(s,t)\,ds\,dt \qquad 3.43$$

fonksiyonuna, X ile Y' nin ortak dağılım fonksiyonu denir. Burada f(s,t), X ile Y ortak olasılık yoğunluğunun (s,t) noktasındaki değeridir.

Sürekli iki rassal değişkenin ortak dağılım fonksiyonu, ortak yoğunluğun sürekli olduğu bütün (x,y) noktalarında ortak yoğunluğu belirler.

3.11. Marjinal Dağılımlar

Marjinal dağılımlar, değişkenlerden biri sabit tutularak sadece diğeri üzerinden hesaplanan olasılık dağılımlarıdır. X ile Y kesikli rassal değişkenler, f(x,y) de bunların (x,y) noktasındaki ortak olasılık dağılımı ise, X' in aralığındaki her x için,

$$g(x) = \sum_{y} f(x,y) \qquad 3.44$$

ile gösterilen fonksiyona X' in marjinal dağılımı denir.

Benzer biçimde, Y' nin aralığındaki her y için,

$$h(y) = \sum_{x} f(x,y) \qquad 3.45$$

ile gösterilen fonksiyona Y' nin marjinal dağılımı denir.

X ve Y sürekli rassal değişkenlerken olasılık dağılımları yerine olasılık yoğunlukları, toplamalar yerine integraller kullanılır.

Bu durumda $-\infty < x < \infty$ iken,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \qquad 3.46$$

X'in marjinal yoğunluğu ve aynı biçimde $-\infty < y < \infty$ iken,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \qquad 3.47$$

Y'nin marjinal yoğunluğu olarak gösterilir.

İkiden çok rassal değişkenle ilgilenildiğinde tekil rassal değişkenlerin marjinal dağılımları hesaplanabildiği gibi çeşitli rassal değişkenlerin ortak marjinal dağılımları da bulunabilir.

$X_1, X_2, ..., X_n$ rassal değişkenlerinin ortak olasılık dağılımı $f(x_1, x_2, ..., x_n)$ ise tek başına x_1'in marjinal dağılımı, x_1' in aralığındaki tüm değerler için; kesikli durumda,

$$g(x_1) = \sum_{x_2} ... \sum_{x_n} f(x_1, x_2, ..., x_n) \qquad 3.48$$

sürekli durumda,

$$h(x_1) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, x_2, ... x_n) dx_2 dx_3 ... dx_n \qquad 3.49$$

şeklinde marjinal yoğunluk fonksiyonları hesaplanabilir.

Aynı koşullarda, X_1, X_2 ve X_3 ün ortak marjinal dağılımı ise; kesikli durumda,

$$m(x_1, x_2, x_3) = \sum_{x_4} ... \sum_{x_n} f(x_1, x_2, ..., x_n)$$

ve sürekli durumda,

$$\phi(x_1, x_2, x_3) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) dx_4 dx_5 ... dx_n$$

şeklinde denklemler kullanılır.Burada $m(x_1, x_2, x_3)$ ve $\Phi(x_1, x_2, x_3)$ X_1, X_2, X_3 ün birleşik olasılık fonksiyonlarıdır.

3.12. Koşullu Dağılımlar

X ile Y, ortak dağılımlı kesikli bir rassal değişken çifti olsun. Bu değişkenlerin ortak olasılık dağılımının (x,y)' deki değeri f(x,y), Y' nin marjinal dağılımının y'deki değeri de h(y) ise, X aralığındaki her bir x için,

$$f(x/y) = \frac{f(x, y)}{h(y)} \qquad h(y) \neq 0$$

ile gösterilen fonksiyona, Y=y verilmişken X' in koşullu dağılımı denir.

Buna karşılık olarak X' in marjinal dağılımının x' teki değeri g(x) ise, Y aralığındaki her bir y için,

$$w(y/x) = \frac{f(x, y)}{g(x)} \qquad (x) \neq 0$$

ile gösterilen fonksiyona X=x verilmişken Y'nin koşullu dağılımı denir.

X ile Y sürekli rassal değişkenler olduğunda, olasılık dağılımları yerine olasılık yoğunlukları geçer. Hesaplamalar toplama operatörleri, integral operatörleri ile değiştirilerek gerçekleştirilir.

İki ya da daha çok rassal değişken için istatistiksel bağımsızlık konusu genellikle büyük önem taşır. Mesela Y = y verilmişken X'in koşullu dağılım değerleri y'ye bağlı değilse **f(x\y)=g(x)** olur, böylece

f(x\y)=f(x\y).h(y)=g(x).h(y)

Burada ortak dağılım değerleri iki marjinal dağılımın ilgili değerlerinin çarpımına eşit olmaktadır.Daha genel bir tanımla; $\mathbf{f(x_1,x_2,...,x_n)}$, n tane kesikli $\mathbf{X_1,X_2,...,X_n}$ rassal değişkenin ortak olasılık dağılımının $\mathbf{(x_1,x_2,...,x_n)}$1 deki değeriyse ve $\mathbf{f_i(X_i)}$ de, i =1,2,...,n için X_i'nin marjinal dağılımının X_i'deki değeriyse, bunların aralarındaki bütün $\mathbf{(x_1,x_2,...,x_n)}$ için aşağıdaki koşul sağlanırsa n tane rassal değişken bağımsız olur.(Hsu ,H.,1997)

$\mathbf{f(x_1,x_2,...,x_n) = f_1(x_1).f_2(x_2)\ \ f_n(x_n)}$

Beklenen Değerin Bazı Özellikleri

İster sürekli isterse kesikli olsun tek değişkenli rassal değişkenlerin beklenen değerleri için aşağıdaki bazı özelliklerden söz etmek uygun olacaktır:

a ve b sabit sayılar olmak üzere;

i) E(a.X)= a.E(X)

ii) E(b)=b

iii) E(a.X+b)= a.E(X)+b

iv) $g_1(X)$ ve $g_2(X)$ Aynı örnek uzayı S'den reel sayılar kümesiyle tanımlanmış iki fonksiyon ise X'in birer fonksiyonu olmak üzere

$E[g_1(X)+g_2(X)]=E[g_1(X)]+E[g_2(X)]$

v) Tüm x değerleri için $g_1(x)<g_2(x)$ ise, $E[g_1(X)] < E[g_2(X)]$ olur.

Çok Değişkenli Durumlarda Beklenen Değer

Beklenen değer kavramı birden çok rassal değişken içeren çok değişkenli durumda da uygulanır.

X ile Y kesikli rassal değişkenler, f(x,y) bunların ortak olasılık dağılımının (x,y)' deki değeriyse g(X,Y)' nin beklenen değeri: $E[g(X,Y)] = \sum_x \sum_y g(x,y).f(x,y)$ olur.Aynı şekilde, X ile Y sürekli rassal değişkenler, f(x,y) de bunların ortak olasılık yoğunluğunun (x,y)' deki değeriyse g(X,Y)' nin beklenen değeri:

BÖLÜM 4: TAHMİN

4.1 Parametreler

Anakütle hakkında bilgi veren karekteristik değere parametre denir . Örneğin, üretilen bir cihazın boyunun λ parametreli üstel bir dağılıma sahip olduğu ama bu parametrenin gerçek değerinin bilinmediği varsayılsın.. Bu türden birkaç cihazın boyu gözlemlenebilirse, bu gözlem değerlerinden yola çıkılarak λ'nın değeri hakkında bir tahminde bulunmak mümkün olabilir.

Belirli bir topluluktaki bireylerin boylarının dağılımı; μ ortalamasıyla ve σ^2 varyansıyla normal dağılsın. (μ ve σ^2'nın tam değeri bilinmemektedir.) Bu topluluktan seçilmiş rassal bir örnekteki bireylerin boylarını gözlemlenirse μ'nün ve σ^2'nin değerleri hakkında bir sonuç çıkarabileceği düşünülebilir.

Bir θ parametresinin alması olası tüm ($\theta_1 \ldots \theta_n$) değerlerini içeren kümeye parametre uzayı denir ve Ω ile gösterilir. Sunulan ilk örnekte , üstel dağılımının parametresi λ pozitif olmalıdır. Bu nedenle, parametre uzayı Ω tüm pozitif reel sayılar kümesinden oluşur. İkinci örnekte μ'nün değeri herhangi bir reel sayı olabilir. σ^2 de pozitiftir.

4.2 Tahmin Yöntemleri

İki temel tip tahmin probleminden söz edilebilir: Birinci tipte, bir veya daha fazla rassal değişken parametrelerinin tahmin edilmesiyle ilgilenilir. İkinci tipte ise, bağımlı bir rassal değişkenin değerinin; bağımsız rassal değişkenlerin gözlem değerleri yardımı ile tahmini ile ilgilenilir.

4.2.1. Parametrenin Tahmini

İstatistik

X' in olasılık fonksiyonu f(x) olsun ve ($X_1,\ldots\ldots X_n$), ise, sözkonusu olasılık dağılımına tabi bir anakütleden çekilen bir dizi gözlem olsun. ($X_1,\ldots\ldots X_n$) 'in reel değerli bir fonksiyonu $s(X_1,\ldots\ldots X_n)$ olsun. Bu durumda $s(X_1,\ldots\ldots X_n)$ bir istatistik olarak adlandırılır. İstatistik'de anakütle tamamen gözlenmez. Dolayısıyla anakütle dağılımını niteleyen parametrelerin tahmini değerleri sözkonusu istatistikler yardımı ile gerçekleştirilmeye çalışılır.

Ancak Bayesgil istatistiğe göre sözkonusu parametreler birer sabit olarak ele alınmazlar. Onların da birer olasılık fonksiyonları vardır. Dolayısıyla θ parametreli bir dağılımdan çekilen rastlantı değişkeni X'in olasılık dağılımı f(X/ θ) biçiminde koşullu olarak ifade edilir.

Ya da X ile birlikte θ da bir rastlantı değişkeni olarak kabul edileceği için onların ortak olasılık dağılımından sözedilir. (X_1,.......X_n), X in rastgele bir örneği olsun. Bu durumda (X_1,.......X_n)'in ortak olasılık yoğunluk fonksiyonu aşağıdaki eşitlikle verilmektedir:

$$f(x;\theta)=f(x_1,\ldots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)$$

Burada θ uygun istatistikler aracılığı ile tahmin edilmektedir. Örnekten gelen rassal gözlemlerin bir fonksiyonu olacağı için, θ'nın tahmin edicisi de bir rastlantı değişkeni olmaktadır. Tahmin edicinin özel bir değeri de de onun gerçekleştirilmiş hali olmaktadır.

Nokta Tahmini ve Aralık Tahmini

Bir parametre tahmininin tek bir değer olması gerekli değildir. Bunun yerine belirli bir olasılıkla tahmin değerleri dizisi öngörülebilir. Tek bir değeri tanımlayan tahminler nokta tahminleri olarak adlandırılır. Belirli bir olasılıkla bir değer aralığını tanımlayan tahminler de aralık tahminleri olarak adlandırılmaktadır.[4]

4.3.2. Nokta Tahmin Edicilerin Özellikleri

A. Tarafsız Tahmin Ediciler (Unbiased Estimators)

Θ = s(X_1,.......X_n), tahmin edicisi ya da istatistiği, E(Θ) = θ koşulunu sağlıyorsa θ parametresinin tarafsız tahmin edicisi olarak tanımlanır.

Θ nın tarafsız bir tahmin edici olması durumunda, hata kareleri ortalaması aşağıdaki gibi tanımlanır:

$$E\left[(\Theta-\theta)^2\right]=E\left\{[\Theta-E(\Theta)]^2\right\}=Var(\Theta)$$

Θ , θ 'nın yansız bir tahmin edicisi ise , kareler ortalaması hatası, varyansına eşittir.

B. Etkili Tahmin Ediciler

Θ_1 tahmin edicisi, aşağıdaki durumda, Θ_2 tahmin edicisine göre, θ parametresinin daha etkili bir tahmin edicisi olarak adlandırılır:

1. Θ_1 ve Θ_2, θ nın tarafsız tahmin edicileri ise.

[4] Literatürde parametrelere ilişkin aralık tahminleri güven aralıkları (confidence intervals) olarak adlandırılmaktadır.

2. $Var(\Theta_1) < Var(\Theta_2)$

Tahmin edici $\Theta_{MV} = s(X_1, \ldots\ldots X_n)$,, aşağıdaki durumda θ parametresinin en etkili (veya minimum varyanslı) tarafsız tahmin edicisi olarak adlandırılır

1. θ nın tarafsız bir tahmin edicisidir

2. Tüm Θ lar için var (Θ_{MV}) ≤ Var(Θ) dır.

Bu durumda Θ_{MV} ; θ 'nın minimum varyanslı (ya da en etkin) tahmincisi olarak da adlandırılabilir.

C. Tutarlı Tahmin Ediciler

Θ ; θ 'nın yansız bir tahmincisi olmayabilir. Başka bir deyişle $E(\Theta) \neq \theta$ olur. Bununla birlikte örnek büyüklüğü arttırıldığında Θ ile θ arasındaki fark sistematik olarak azalabilir.

n büyüklüğündeki rastgele bir örneğe dayalı olarak θ nın Θ_n tahmin edicisi, her küçük ε>0 için

$$\lim_{n\to\infty} P(|\Theta_n - \theta| < \varepsilon) = 1$$

veya eşlenik olarak,

$$\lim_{n\to\infty} P(|\Theta_n - \theta| \geq \varepsilon) = 0$$

oluyorsa Θ_n , θ'nın tutarlı bir tahmin edicisi denir.

Aşağıdaki eşitlikler tutarlılığı tanımlamak için yeterlidir

1. $\lim_{n\to\infty} E(\Theta_n) = \theta$

2. $\lim_{n\to\infty} Var(\Theta_n) = \theta$

Tutarlılık asimptotik bir özelliktir.

4.3 En Küçük Kareler Yöntemi

En Küçük Kareler Yöntemi, doğrusal (ya da doğrusal olmayan) , çoklu regresyon modellerinin çözümlenmesinde kullanıldığı gibi, çok denklemli ekonometrik modellerin çözümünde de kullanılır. Kurulan regresyon modellerinde gözlemlerin , anakütle gözlem değerlerinden herhangi şekilde alınmış gözlemler olduğu düşünülür. Kurulan regresyon modeli eldeki örnekten hareketle oluşturulmaya çalışılır. Bu nedenle kurulan modeldeki değerler tahmini değerler olacaktır. Tahmin edilmeye çalışılan açıklanan (ya da bağımlı) değişkenin alacağı değer;

açıklayıcı (ya da bağımsız) değişken (ya da değişkenlerin) gözlem değerleri yardımı ve genellikle, doğrusal bir fonksiyon aracılığı ile tahmin edilir. Bu regresyon denkleminde bulunan sabitlerin (gözlem değerleri yardımı ile) tahmin edilmesi ile uygun bir matematiksel tahmin modeli oluşturulmaya çalışılır. Ancak tahmini değerler ile gözlem değerleri nadiren birbirleri ile çakışacağından ayrı bir notasyona ihtiyaç vardır. Genellikle tahmini değerler "şapka" notasyonu ile ifade edilirler. Sözgelimi θ parametresinin eldeki örnekten hareketle hesaplanmış olan bir tahmincisi $\hat{\theta}$ ile gösterilir. Sözgelimi $\hat{\theta}=3$ olduğunda bu özel 3 değeri ; θ'nın bir tahmini değeri (ya da kestirimidir.) Özetlemek gerekirse θ bir parametredir. $\hat{\theta}$ ise bir formül ya da istatistik. Bu yüzden $\hat{\theta}$; θ'nın değerinin bir tahmincisi ya da kestirimcisidir (estimator). Bu tahmincinin farklı gözlem değerlerine göre farklı θ tahmini değerleri vereceği açıktır. Bu tahmini değerlerin her birine de θ'nın birer tahmini (estimate) denir.

Tek açıklayıcı değişkenli doğrusal regresyon modeli ele alınsın. Kurulan ana kütle regresyon modeli,

$$Y_i=\alpha+\beta X_i+u_i$$

Anakütle regresyon tahmini modeli ise

$\hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i$ denklemleri ile ifade edilmektedir.

Burada α *ve* β anakütle regresyon modelinin parametreleridir. $\hat{\alpha}$ *ve* $\hat{\beta}$ ise sözkonusu parametrelerin (eldeki örneğe dayanılarak oluşturulacak olan) tahmincileridirler.

Tahmin modelinin güvenilirliğinin sınanması da (daha sonra değinilecek bazı varsayımlar altında oluşturulacak olan) parametrelerin aralık tahminlerine ve benzeri istatistiklere dayanarak gerçekleştirilmektedir.

Regresyon analizi için kurulan modelde, bağımlı ve bağımsız değişkenin (ya da değişkenlerin) yanısıra hata terimi olarak isimlendirilen ve anakütle regresyon denkleminde u_i şeklinde ifade edilen rassal değişken yer almaktadır. Sözkonusu hata terimi (ya da distürbans) modele rassal olma özelliğini katan değişkendir. Çünkü bilindiği gibi

α *ve* β parametreleri sabittir. Yine açıklayıcı değişken ya da değişkenlerin değerleri sabit kabul edilmektedir. Dolayısıyla bağımlı değişkenin rassal değişken olabilmesi için geriye bir tek hata teriminin rassal değişken olması koşulu kalmaktadır.

Örnek regresyon modelinde ise bağımlı değişkenin gözlenen değerleri ile tahmin edilen değerleri arasında bir fark (genellikle) bulunur. Bu farklar hata terimleri ya da kalıntılar (residuals) olarak adlandırılırlar. Kalıntılar e_i notasyonu ile gösterilirler.

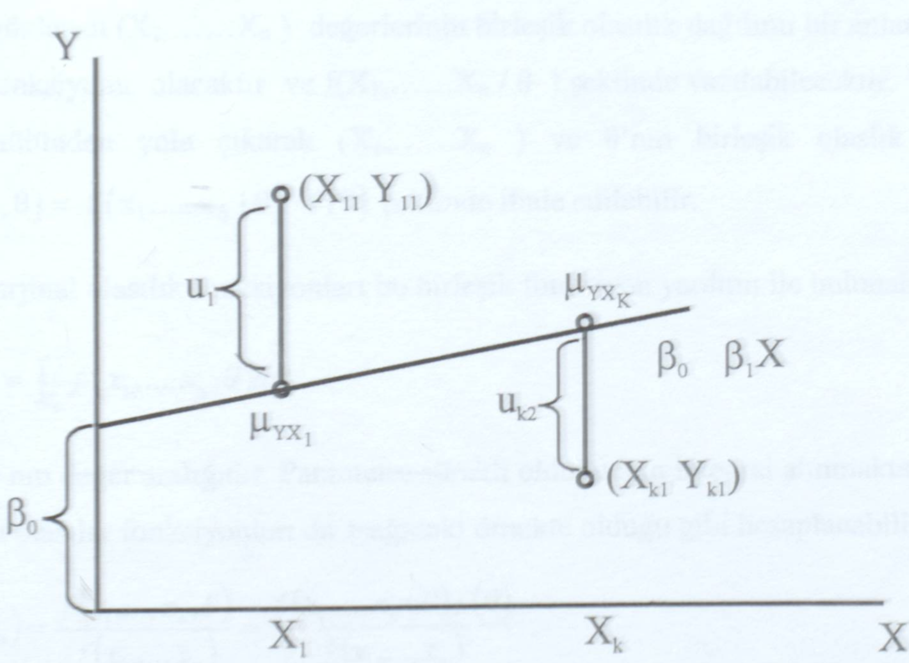
$e_i = Y_i - \hat{Y}_i$ olur. Başka bir değişle i. kalıntı, bağımlı değişkenin i. fiili ya da gözlenen değeri ile tahmin edilen değeri arasındaki farka eşittir.

Hata terimlerinin karelerinin toplamını minimum yapan yöntemler arasında en çok kulanılanlardan bir tanesi en küçük kareler yöntemidir. Bu yöntem kısaca kalıntı kareleri toplamı adı da verilebilecek aşağıdaki fonksiyonu minimum kılmayı sağlayacak olan parametre tahminlerini verir:

$$Q = \sum_{i=1}^{n} e_i^2$$

Burada Q fonksiyonu en küçük kareler fonksiyonudur ve parametrelerin en küçük kareler tahminleri bu fonksiyonu eşzamanlı olarak minimize eden parametre tahminlerinin bulunması ile gerçekleştirilir. Sözgelimi, yukarıdaki tek açıklayıcı değişkenli doğrusal regresyon modelinde

$\frac{\partial Q}{\partial \hat{\alpha}} = 0$ ve $\frac{\partial Q}{\partial \hat{\beta}} = 0$ denklemleri eşzamanlı olarak çözülür.



Şekil 4.1. Klasik Regresyon Örneği

4.4 En Çok Benzerlik Yöntemi

R.A. Fisher'e göre geliştirilen bu yönteme göre parametre tahminleri, elde bulunan gözlemleri (eşzamanlı ya da birlikte) elde etme olasığını maksimum kılacak şekilde gerçekleştirilir. Bu da olabilirlik fonksiyonu adı verilen bir fonksiyon yardımı ile yapılır. $X_1, X_2, ..., X_n$; θ parametreli bir dağılımdan gözlenmiş n birimlik bir örneği oluştursun. Başka bir deyişle X_1 bu örneğin birinci gözlem değeri, ... , X_n örneğin n. gözlem değeri olsun. Olabilirlik fonksiyonu $f(X_1, ..., X_n, \theta)$ şeklindedir. Maksimum olabilirlik yöntemine göre olabilirlik fonksiyonunu maksimize edecek $\hat{\theta}$ değeri bulunmaya çalışılır. Bu konuya ileride bazı olasılık dağılımlarının parametrelerinin maksimum olabilirlik yöntemi ile bulunması sırasında yeniden dönülecektir.

4.4.1. Bayesçi Tahmin

Klasik yaklaşıma göre bir anakütleyi ya da olasılık dağılımını şeklendiren parametreler sabittir. Bayesçi yaklaşıma göre ise bu parametrelerin bizzat kendileri de birer olasılık dağılımına uymaktadılar. Dolayısıyla parametrelerin kendileri de birer rastlantı değişkenidirler.

Bilinmeyen θ parametresinin olasılık fonksiyonu f(θ) olsun. Bu durumda θ parametreli bir dağılımdan gözlenen (X_1,……X_n) değerlerinin birleşik olasılık dağılımı bir anlamda koşullu bir olasılık fonksiyonu olacaktır ve f(X_1,……X_n / θ) şeklinde yazılabilecektir. Yine koşullu olaslık formülünden yola çıkarak (X_1,……X_n) ve θ'nın birleşik olaslık fonksiyonu $f(x_1,..., x_n,\theta) = f(x_1,...,x_n \mid \theta)\ f(\theta)$ şeklinde ifade edilebilir.

Yine bazı marjinal olasılık fonksiyonları bu birleşik fonksiyon yardımı ile bulunabilir:

$$f(x_1,...,x_n) = \int_{R_0} f(x_1,...,x_n,\theta) d\theta$$

Burada, R, θ nın değer aralığıdır. Parametre sürekli olduğu için integral alınmaktadır. Buradan diğer koşullu olasılık fonksiyonları da aşağıdaki örnekte olduğu gibi hesaplanabilirler:

$$f(\theta \mid x_1,...,x_n) = \frac{f(x_1,...,x_n,\theta)}{f(x_1,...,x_n)} = \frac{f(x_1,...,x_n \mid \theta) f(\theta)}{f(x_1,...,x_n)}$$

Yukarıdaki olasılık fonksiyonu θ' nın posterior olasılık fonksiyonu olarak tanımlanır. Gözlem öncesi ya da prior f(θ), (X_1,……X_n) sonuçlarının gözlenmesinden önce θ konusundaki bilgileri ifade eder. Posterior olasılık dağılım fonksiyonu f(θ/x_1,…..x_n), ise örnekteki bilgiler ile prior f(θ) ‘dan yararlanılarak bulunabilmektedir. Başka bir deyişle posterior dağılım, prior dağılımdan gelen bilgiler ile örnekten gelen bilgilerin bir sentezini oluşturmaktadır.

θ'nın koşullu ortalaması aşağıdaki eşitlikle ifade edilir:

$$\theta_B = E(\theta \mid x_1,\dots,x_n) = \int_{R_0} \theta f(\theta \mid x_1,\dots,x_n)d\theta$$

Başka bir deyişle bu beklenen değer posterior dağılımın beklenen değerine eşit olmaktadır.

Bu beklenen değer ; θ ‘nın Bayes tahmini olarak da adlandırılmaktadır.

$$\Theta_B = E(\theta \mid X_1,\dots,X_n)$$

Bayes Yaklaşımının Temeli

Bayes teoremi, bu yaklaşımın temelidir.

A ve B aynı örnek uzayı S içerisinde iki olay olsun. Bu durumda Bayes Teoremine göre B olayı verildiğinde A'nın koşullu gerçekleşme olasılığı

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad P(B) \neq 0 \quad için$$

şeklinde hesaplanır. Bayes Teoremi az önceki tartışmaya uyarlanacak olursa Y verildiğinde (veya gözlendiğinde) θ'nın posterior dağılımı aşağıdaki şekilde hesaplanabilir:

$$P(\theta|Y) = \frac{P(\theta|Y)P(\theta)}{P(Y)} \quad P(Y) \neq 0 \quad için \tag{4.9}$$

Bu ifade literatürde ters olasılık ilkesi (principle of inverse probability) olarak da geçmektedir.

(4.1) no'lu ifadeden hareketle,

$$P(Y,\theta) = P(Y/\theta)P(\theta) \tag{4.10}$$

$$P(Y,\theta) = P(\theta/Y)P(Y) \tag{4.11}$$

$$P(\theta/Y) = \frac{P(\theta)P(Y/\theta)}{P(Y)}$$

. P(Y) terimi normalleştirme sabiti olarak adlandırılır ve aşağıdaki gibi yazılabilir

$$P(Y) = \int P(\theta)P(Y/\theta)d\theta \tag{4.12}$$

(4.11) eşitliğinden P(Y) terimi çıkarıldığında,

$$P(\theta / Y) \propto P(\theta)P(Y / \theta) \tag{4.13}$$

elde edilir, $\propto$ simgesi orantısallığı belirtmektedir. P(Y/θ), örneklem bilgisini gösteren olabilirlik fonksiyonu, P(θ) ise θ parametresinin prior olasılık yoğunluk fonksiyonunu ya da ön bilgiyi vermektedir. P(θ/Y) ise, Y verildiğinde parametre θ için posterior olasılık yoğunluk fonksiyonunu göstermektedir. Ayrıca içinde bütün ön bilgiyi örneklem bilgisi ile birlikte taşımaktadır. Son bilgiler, Bayesci analizde çıkarımları elde etmede doğrudan kullanılırlar.

Ön bilgi, posterior olasılık yoğunluk fonksiyonuna ön olasılık yoğunluk fonksiyonu aracılığıyla girer. Örneklem bilgisi ise benzerlik fonksiyonu yardımıyla analizde yer alır.

Bayesci analiz sonucunda bulunan son dağılım, bir sonraki analizde ön dağılım olarak kullanılabilir. Burada ikinci analiz için alınacak olabilirlik fonksiyonunun, ilk olabilirlik fonksiyonundan bağımsız olması koşulu vardır. (Kahyaoğlu ,1999)

$P(\theta/X) \propto P(\theta)L(\theta/X)$ iken X'ten bağımsız Y örneklem için aşağıdaki ifadeleri oluşturmak mümkündür.

$$P(\theta/X,Y) \propto P(\theta)L(\theta/Y) \tag{4.14}$$

$$L(\theta/X,Y) \propto L(\theta/X)L(\theta/Y) \tag{4.15}$$

Buradan yola çıkarak,

$$P(\theta/X,Y) \propto P(\theta)L(\theta/X)l(\theta/Y) \tag{4.16}$$

$$P(\theta/X,Y) \propto P(\theta/X)l(\theta/Y) \tag{4.17}$$

elde edilir.

4.5. Klasik İstatistikle Bayesgil İstatistiğin Karşılaştırılması

Genelde klasik istatistik yanlıları, Bayesgil istatistiği aşırı sübjektif davranmakla suçlarlar. Ancak klasik hipotez testlerinde bilindiği gibi, bu testlerin anlamlılık düzeyi bir anlamda sübjektif olarak belirlenmektedir. Klasik istatistikte gerek tahmincilerin elde edilmesi, gerekse aralık tahminleri ve hipotez testleri gibi konuların temeli, örnek çekmenin tekrarlanmasına dayanır. Diğer bir deyişle, kuramsal olarak klasik istatistik sürekli olarak örneklem çekmenin sürdürüldüğü ifadesine dayanır. Bunun nedeni ise, şüphesiz doğru sonucu daha yüksek bir olasılıkla bulma isteğinde yatar. Ancak klasik istatistiğin bu özelliğine karşılık Bayesgil istatistik, böyle bir temele dayanmaz. Bayesgil istatistik'de bireylerin inanç derecelerine dayanan sübjektif olasılıklardan da yararlanılır ve klasik istatistiğe göre karar kavramı daha sık ve ağırlıklı bir şekilde vurgulanarak, örneklem dışı bilgilerin de biçimsel olarak analize katılması sağlanır.

BÖLÜM 5 : PRİOR VE POSTERİOR DAĞILIMLAR

5.1 Prior Dağılımın Belirlenmesi

Buradaki problem , kısaca olasılık fonksiyonu f(x/θ) olarak verilen bir anakütleden çekilen gözlemleri temel alarak, θ'nin asıl değerinin parametreler uzayı Ω'nin neresinde var olabileceğini saptamaya çalışmaktan ibarettir.

Birçok problemde, f(x/θ) ile ilgili herhangi bir gözlem yapılmadan önce,deney yapan bir kişi, önceki bilgilerini gözden geçirerek, Ω uzayında bir olasılık dağılımı oluşturmak suretiyle, θ'in değerinin nerede bulunabileceğini tahmin edebilir. Diğer bir deyişle, herhangi bir deneysel veri elde edilmeden ya da gözlemlenmeden önce, deneyi yapan kişinin geçmiş tecrübeleri ve bilgi birikimi onu, θ'nın Ω'nın diğer bölgelerine oranla, belli başlı bazı bölgelerinde bulunduğuna inanmaya sevk edecektir. θ'nın Ω uzayında belirli bölgelerde bulunma olasılıklarından hareketle bir ön dağılım düşünülebilir. Bu dağılıma θ'nın prior dağılımı denir.

İstatistikte, prior dağılım kavramı çok fazla tartışma yaratmaktadır. Bir araştırmacının θ'nın gerçek değerinin nerede olabileceğine dair öznel inançları bulunması bağlamında , bu dağılımın bir öznel olasılık dağılımı olduğu savı bazı istatistikçiler tarafından ileri sürülür. Öte yandan , bir prior dağılımın istatistik alanında kullanılan diğer herhangi bir olasılık dağılımından farksız olduğuna ve olasılık teorisinin tüm kurallarının bir prior dağılım için de geçerli olacağına inanılır.

Diğer bazı istatistikçiler birçok problemde θ'nın olasılık dağılımından söz etmenin doğru olmadığı yorumunda bulunur. Çünkü θ'nın gerçek değeri daha çok deney yapan kişi tarafından belirlenen sabit bir sayı olarak düşünülür. Bu istatistikçiler; ancak θ'nın geçmişteki olası değerleri ile ilgili yoğun bir ön bilgi sahibi olunduğunda, θ parametresinde bir prior dağılımın kullanılabileceği düşünmektedirler. Ancak o zaman, iki farklı istatistik ekolünün kullanılacak uygun bir prior dağılım üzerinde uzlaşması mümkün olacaktır.

5.1.2 Kesikli ve Sürekli Prior Dağılımlar

Bazı problemlerde, θ parametresi sadece sonlu sayıda değişik değer alabilir. Diğer problemlerde, θ parametresi reel sayılar kümesi üzerinde ya da reel sayılar kümesinin bir aralığında sonsuz sayıda değer alabilir . Bu durumda bu dağılımın olasılık yoğunluk fonksiyonuna, θ'nın prior olasılık yoğunluk fonksiyonu denir.

5.1.3 Posterior Dağılım

$X_1 \ldots X_n$ nin; olasılık fonksiyonunun f(x/θ) olduğu bir anakütleden; rassal bir örnek oluşturduğu varsayılsın. Ayrıca parametre θ'nın değerinin bilinmediği ve θ'nın prior olasılık olasılık fonksiyonunun da ξ(θ) olduğu düşünülsün.

Bu şartlara bağlı $X_1 \ldots X_n$ in olasılık fonksiyonu şekilde yazılabilir:

$$g_n(x) = \int_\Omega f_n(x|\theta)\xi(\theta)d(\theta) \tag{5.2}$$

Benzer bir şekilde

$X_1=x_1$ $X_n=x_n$ verildiğinde*, θ'nın koşullu olasılık yoğunluk fonksiyonu, $X_1 \ldots X_n$'nin bileşik olasılık yoğunluk fonksiyonuna ve θ'nın $x_1 \ldots x_n$ marjinal bileşik olasılık yoğunluk fonksiyonuna bölünerek elde edilir.

$$\xi(\theta|x) = \frac{f_n(x|\theta)\xi(\theta)}{g_n(x)} \tag{5.3}$$

Buradaki koşullu olasılık yoğunluk fonksiyonuna posterior dağılım denir. Çünkü bu, $X_1 \ldots X_n$'in değerleri gözlemlendikten sonra elde edilen θ'nın dağılımıdır. . Posterior olasılık yoğunluk fonksiyonu ξ (θ/x); $X_1=x_1$, $X_n=x_n$ gözlemledikten sonra θ 'nın bu göreli olabilirliği temsil eder

5.1.4 Eşlenik Dağılımlar

Bazı durumlarda θ parametresinin prior dağılımının bilinmesi, posterior dağılımının tahmin edilmesine de yaramaktadır. Örneğin θ'nın prioru beta ise posterioru da farklı parametreli bir diğer beta dağılımına uymaktadır. Bu durumda prior ve posterior dağılımlar eşlenik (conjugate) dağılımlar olarak nitelendirilmektedir. Bu durumu aşağıdaki çizelge yardımı ile de incelemek mümkündür.

---

* (θ/x) şeklinde belirtelecek

Çizelge 5.1: Bilgi Verici Örneklem Dağılımları ile Bu dağılımlara Ait Ön ve Son Dağılımlar

Örneklem dağılımı	Ön dağılım	Son dağılım
Bernoulli	Beta	Beta
Poisson	Gama	Gama
Negatif Binomial	Beta	Beta
Üstel	Gama	Gama
Normal	Normal	Normal

5.1.4.1 Bir Bernoulli Dağılımından Örnekleme

θ parametresinin değerinin bilinmediği bir Bernoulli dağılımından rassal bir örnek alınsın. Eğer θ'nın prior dağılımı bir beta dağılımıysa, o zaman herhangi bir rassal örnek yardımı ile , θ'nın posterior dağılımı da yine bir beta dağılımı olacaktır.

$X_1 \ldots X_n$ 'in, θ parametresinin değerinin bilinmediği bir Bernoulli dağılımından rassal bir örnek olsun ($0< \theta <1$). Ayrıca θ'un prior dağılımının, verilen α ve β($\alpha>0$ ve $\beta>0$) parametreleriyle, bir beta dağılımı olduğu varsayılsın. $y = \sum_{n=1}^{n} x_i$ olsun.

Prior olasılık yoğunluk fonksiyonu için şu orantısallıktan söz edilebilir.

$$\xi(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \quad 0 \langle \theta \langle 1 \tag{5.5}$$

Posterior olasılık yoğunluk fonksiyonu f_n $(x/\theta)\xi(\theta)$ ξ (θ/x)'ya oranlı olduğundan, şu sonuç çıkmaktadır:

$$\xi(\theta|x) \propto \theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-1} \quad 0 \langle \theta \langle 1 \tag{5.6}$$

Sabit bir faktör dışında, bu bağıntının sağ tarafı, α+y ve β+n-y parametrelerine sahip bir beta dağılımının olasılık yoğunluk fonksiyonuna eşittir.

5.1.4.3 Bir Poisson Dağılımından Örnekleme

$X_1 \ldots X_n$ Poisson dağılımında θ'nın ortalama değerinin bilinmediği rassal bir örnek oluştursun ($\theta>0$). θ'nın prior dağılımı, α ve β'nın ($\alpha>0$ ve $\beta>0$) belli parametrelerle gama dağılımı olduğu varsayılsın. Sonra θ'nın posterior dağılımı, $X_i=x_i$ (i=1,…,n) verildiğinde, parametreler $\alpha + \sum_{i=1}^{n} x_i$ ve β+n ile gama dağılımıdır.

Yine $y = \sum_{n=1}^{n} x_i$ olsun.

$$f_n(x|\theta) \propto e^{-n\theta}\theta^{y}$$
(5.7)

Bu bağıntıda x'i içeren ancak θ'ya bağlı olmayan bir faktör sağ taraftan atılır. Ayrıca θ'nın prior olasılık yoğunluk fonksiyonu

$$\xi(\theta) \propto \theta^{\alpha-1}\, e^{-\beta\theta} \quad \theta \rangle 0$$

$$\xi(\theta|x) \propto \theta^{\alpha+y-1}\, e^{-(\beta+n)\theta}$$
(5.8)

şekline sahiptir.

Olasılık dağılım fonksiyonu (θ /x) posteriorı, f_n (x/ θ)ξ (θ)'ya oransal olduğu için

$$\xi(\theta|x) \propto \theta^{\alpha+y-1}\, e^{-(\beta+n)\theta}$$

ya uyar (5.9)

Bu bağıntının sağ tarafı, α+y ve β+n parametreleriyle sabit bir faktör dışında bir gama dağılımının olasılık yoğunluk fonksiyonu olarak görülebilir.

5.1.4.4 Bir Normal Dağılımdan Örnekleme

$X_1 \ldots X_n$, ortalama θ'nın değerinin bilinmediği(.-∞<ө<∞) ve σ^2 varyansının değerinin bilindiği(σ^2>0) bir normal dağılımdan rassal bir örnek oluşturduğu varsayılsın. θ'nın prior dağılımının, ortalama μ'nın ve σ^2 varyansının belli değerleriyle bir normal dağılım olduğu düşünülsün. θ'nın posterior dağılımı,

$$\mu_1 = \frac{\sigma^2\mu + n\upsilon^2 \overline{x}_n}{\sigma^2 + n\upsilon^2}$$
(5.10)

$$\upsilon_1^2 = \frac{\sigma^2\upsilon^2}{\sigma^2 + n\upsilon^2}$$
(5.11)

Parametreli normal bir dağılıma uymaktadır.

Olasılık fonksiyonu f_n(x/ θ)

$$f_n(x|\theta) \propto \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\right]$$
(5.12)

şekline sahiptir.

Burada sabit faktör sağ taraftan atılır. Bu ifadeyi dönüştürmek için şu özdeşlik kullanılacaksa,

$$\sum_{i=1}^{n}(x_i-\theta)^2 = n(\theta-\bar{x}_n)^2+\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 \tag{5.13}$$

ve $x_1 \ldots x_n$ içeren ancak $f_n(x/\theta)$ 'ya bağlı olmayan bir faktör olarak ortaya çıkar. Sonuç olarak $f_n(x/\theta)$' yı aşağıdaki şekillerle yeniden yazabilir:

$$f_n(x|\theta) \propto \exp\left[-\frac{1}{2\sigma^2}(\theta-\bar{x}_n)^2\right] \tag{5.14}$$

Prior olasılık yoğunluk fonksiyonu $\xi(\theta)$ şekline sahip olduğu için, onu izleyen posterior olasılık yoğunluk fonksiyonu aşağıdaki formdadır:

$$\xi(\theta) \propto \exp\left[-\frac{1}{2\upsilon^2}(\theta-\mu)^2\right] \tag{5.15}$$

ve daha önceden belirtildiği gibi

$$\xi(\theta|x) \propto \exp\left\{-\frac{1}{2}\left[\frac{n}{\sigma^2}(\theta-\bar{x}_n)^2\ \frac{1}{\upsilon^2}(\theta-\mu)^2\right]\right\} \tag{5.16}$$

elde edilir. Bu denklemin sağ tarafındaki son terim θ içermediği için, oransallık faktörü içersinde düşünülebilir, ve şu bağıntı elde edilir:

$$\frac{n}{\sigma^2}(\theta-\bar{x}_n)^2+\frac{1}{\upsilon^2}(\theta-\mu)^2 = \frac{1}{\upsilon_1^2}(\theta-\mu_1)^2+\frac{n}{\sigma^2+n\upsilon^2}(\bar{x}_n-\mu)^2 \tag{5.17}$$

Bu bağıntının sağ tarafı, sabit faktör dışında, ortalaması μ ve varyansı σ^2 olan normal bir olasılık yoğunluk fonksiyonu olarak değerlendirilebilir. Bu nedenle, θ'nın posterior dağılımı teoremde belirtildiği gibidir. θ'nın posterior dağılımının ortalaması, aşağıdaki gibi yeniden yazılabilir:

$$\xi(\theta|x) \propto \exp\left[-\frac{1}{2\upsilon^2}(\theta-\mu_1)^2\right] \tag{5.18}$$

Denklem (5.18)'den, örnek ortalama x ve prior dağılımın ortalama μ'sı bir ağırlıklı ortalamadır sonucuna varılabilir. X'e verilen değişim ağırlığının sıradaki üç niteliği doğruladığı görülebilir.

v ve σ^2'nın değişmez değerleri için, daha geniş örnek büyüklüğü n, daha büyük x'e verilen değişim ağırlığı olacaktır.

V ve n'nin değişmez değerleri için, örnekteki her gözlem σ^2'nin varyansından ne kadar geniş olursa, x'e verilen değişim ağırlığı o kadar küçük olacaktır.

σ^2 ve n'nin değişmez değerleri için, prior dağılımın σ^2 varyansı ne kadar geniş olursa, x'e verilen değişim ağırlığı o kadar geniş olacaktır.

$$\mu_1 = \frac{\sigma^2}{\sigma^2 + n\upsilon^2}\mu + \frac{n\upsilon^2}{\sigma^2 + n\upsilon^2}\bar{x}_n \tag{5.19}$$

5.1.4.5 Bir Üstel Dağılımdan Örnekleme

$X_1....X_n$, θ parametresinin (θ >0), bir üstel dağılıma uyduğu varsayılsın. (θ'nın prior dağılımının, α ve β'nın (α>0 ve β>0) belli olduğu parametrelerle,bir gama dağılımı olduğu varsayılsın.) θ'nın posterior dağılımı, X_1=x_1 (i=1,...,n) verildiğinde, α+n ve β+y parametreleriyle bir gama dağılımıdır.

$y = \sum_{n=1}^{n} x_i$ olsun.

Prior olasılık yoğunluk fonksiyonu

$$f_n(x|\theta) = \propto \theta^n e^{-\theta y}$$

$$\xi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta} \quad \theta \rangle 0 \tag{5.20}$$

şeklindedir. Bu nedenle

$$\xi(\theta|x) \propto \theta^{\alpha+n-1} e^{-(\beta+y)\theta} \quad \theta \rangle 0 \tag{5.21}$$

şeklindeki orantısallık ilişkisi elde edilir.

Bu bağıntının sağ tarafı, sabit faktör dışında, α+n ve β +y parametreleriyle bir gama dağılımının olasılık yoğunluk fonksiyonuna eşittir

5.3 Kayıp (Loss) Fonksiyonları

İyi bir tahminci , bir parametreyi gerçek değerine yakın olarak tahmin etmelidir. Diğer bir deyişle, iyi bir tahmin edici için , tahminin hatası ortalama seviyede 0'a yakın olmalıdır . Herhangi bir θ parametresinin tahmini a olsun. $(a \neq \theta)$ Bu durumda böyle bir tahmin ve dolayısıyla tahmin hatası yapmaktan dolayı ortaya çıkan kaybı ölçmek için bir fonksiyona ihtiyaç vardır. Bu fonksiyon L(θ, a) şeklinde ifade edilsin ve kayıp fonksiyonu olarak adlandırılsın. Doğal olarak, θ ve a arasındaki uzaklık ne kadar büyükse, L(θ, a)'nın değeri o kadar büyük olacaktır. Ancak yine de bir tahmincinin iyi olup olmadığını test etmek için daha istatistiksel bir ölçüye gereksinim vardır. Eldeki her farklı x örneğinden hareketle elde edilecek kayıp fonksiyonlarının ortalaması (ya da beklenen değeri) böyle bir ihtiyaca cevap verebilir.

Risk fonksiyonu kayıp fonksiyonunun beklenen değeri olarak tanımlanır ve aşağıdaki gibi hesaplanır:

$$E[L(\theta,a)|\mathrm{x}] = \int_{\Omega} L(\theta,a)\xi(\theta|\mathrm{x})d\theta \qquad (5.23)$$

Bir istatistikçinin yukarıdaki beklenen değeri minimum kılacak şekilde bir tahminde bulunması gerektiği de açıktır. Bu şekilde elde edilecek olan tahmin edici, literatürde, θ'nın bir Bayes tahmin edicisi olarak ifade edilmektedir. (5.25)

Bununla birlikte farklı kayıp fonksiyonları düşünmek her zaman mümkündür. Dolayısıyla a'nın seçimi aynı zamanda kayıp fonksiyonunun tipine de bağlı olmaktadır.

5.3.1 Farklı Kayıp Fonksiyonları

Tahmin problemlerinde yaygın bir şekilde kullanılan kayıp fonksiyonu kareli kayıp fonksiyonudur. Bu fonksiyon şu şekilde tanımlanır:

$$L(\theta,a) = (\theta - a)^2$$

Kareli kayıp fonksiyonu kullanıldığında Bayes tahmin edicisi posterior dağılımın beklenen değeri olan E(θ/X) olmaktadır.

5.3.1.2 Mutlak Hata Kayıp Fonksiyonu

Tahmin problemlerinde sıkça kullanılan diğer bir fonksiyon, mutlak hata kayıp fonksiyonudur. Bu fonksiyon şu şekilde belirtilir:

$$\mathrm{L}(\theta,\mathrm{a}) = |\theta - \mathrm{a}| \quad .$$

Mutlak hata kayıp fonksiyonu kullanıldığında, Bayes tahmin edicisi θ'nın posterior dağılımının medyanına eşit olmaktadır.

Örnek 1: Bir Bernoulli Dağılımının Parametresinin Tahmin Edilmesi. θ'nın değerinin bilinmediği, tahmin edilmek zorunda olduğu ve θ'nın prior dağılımı, α ve β (α>0 ve β <0) belli dağılımlarıyla bir beta dağılımı olduğu bir Bernoulli dağılımından alınan bir rassal örnek $X_1 \ldots X_n$'i varsayılmaktadır. Karesel hata kayıp fonksiyonlarının, Denklem(5.26)'de belirtildiği gibi, $0< \theta <1$ ve $0<\alpha<1$, kullanıldığı varsayılır.

$x_1 \ldots x_n$'in her gözlemlenmiş değerleri için, $y = \sum_{n=1}^{n} x_i$ olsun. Teorem 5.1'inde görüldüğü üzere; θ'nın posterior dağılımı $\alpha+y$ ve $\beta+n-y$ parametreleriyle bir beta dağılımı olacaktır. Parametreler α ve β ile bir beta dağılımının anlamı $\alpha(\alpha+\beta)$ olacağı için, θ'nın bu posterior dağılımının ortalaması $(\alpha + y) / (\alpha + \beta + n)$ olacaktır. Bayes $\delta^*(X)$ tahmini her gözlemlenmiş vektör x'e eşit olacaktır. Bu nedenle Bayes tahmin edicisi $\delta^*(X)$ şöyle gösterilir:

$$\delta^*(X) = \frac{\alpha + \sum_{i=1}^{n} X_i}{\alpha+\beta+n}. \tag{5.27}$$

Örnek 2: Bir Normal Dağılımın Ortalamasını Tahmin Edilmesi.

Bir rassal örnek $X_1 \ldots X_n$, θ'nın değerinin bilinmediği ve σ^2 varyansının bilindiği bir normal dağılımdan alındığı varsayılmaktadır. θ'nın prior dağılımının, μ ortalamasının ve σ^2 varyansının değerlerinin belli olduğu bir normal dağılım olduğu varsayılmaktadır..

$x_1 \ldots x_n$ 'in her gözlemlenmiş değerleri için, θ'nın posterior dağılımı, ortalama μ'nın Denklem (5.23)'de belirtildiği, bir normal dağılım olacaktır. Bu nedenle Bayes tahmin edicisi $\delta^*(X)$

$$\delta^*(X) = \frac{\sigma^2\mu + n\upsilon^2\overline{X}_n}{\sigma^2 + n\upsilon^2}. \tag{5.27}$$

şeklinde belirtilir.

(5.28)

Örnek 3: Bir Bernoulli Dağılımının Parametresinin Tahmin Edilmesi. Örnek 1'in koşullarını tekrar varsayılmaktadır. Şimdi mutlak kayıp fonksiyonunun kullanıldığı, Denklem (5.28)'de belirtildiği gibi, varsayılmaktadır. $x_1 \ldots . x_n$ 'in her gözlemlenmiş değeri için, Bayes tahmini $\delta^*(X)$, $\alpha + y$ ve $\beta + n - y$ parametreleriyle bir beta dağılımı olan θ'nın posterior dağılımının bir ortancasına eşit olacaktır. Bu ortancanın basit ifadesi yoktur. Ortanca gözlemlenmiş değerlerin her belli kümesi için sayısal ortalama ile belirtilmelidir.

$$\delta^*(X) = \frac{\sigma^2\mu + n\upsilon^2\overline{X}_n}{\sigma^2 + n\upsilon^2}. \tag{5.29}$$

Örnek 4: Bir Normal Dağılımın Ortalamasının Tahmin Edilmesi. Örnek 2'deki koşullar tekrarlansın.Bununla beraber Denklem (5.29)'de belirtildiği gibi, mutlak hata kayıp fonksiyonunun kullanılsın. $X_1 \ldots X_n$ 'in her gözlemlenmiş değeri için, Bayes tahmini $\delta^*(X)$, θ'nın posterior normal dağılımının ortancasına eşit olacaktır. Ancak, her normal dağılımın ortancası ve ortalaması eşit olduğunda, $\delta^*(X)$ da, posterior dağılımın ortalamasına eşit olur. Bu nedenle, mutlak hata kayıp fonksiyonuyla ilgili Bayes tahmin edicisi, karesel hata kayıp fonksiyonuyla ilgili Bayes tahmin edicisi aynıdır. Bu Denklem (5.30)'da tekrar gösterilir.

$$\delta^*(X) = \frac{\sigma^2\mu + n\upsilon^2\overline{X}_n}{\sigma^2 + n\upsilon^2}. \tag{5.30}$$

5.3.1 Bayes Tahmin Edicisinin Tutarlılığı

Literatürde Bayes tahmin edicilerinin tutarlı oldukları belirtilmektedir. Bunun anlamı gözlem sayısı arttırıldığında Bayes tahmin edicilerinin θ'nın bilinmeyen değerine oldukça yakın değerler verdikleridir.

5.4.1 Bayes Tahmin Edicilerin Sınırlamaları

Bayes yaklaşımının istatistik problemlerinde uygulanabilirliğinde belli sınırlar bulunmaktadır. Bu yaklaşımı benimseyebilmek için kareli hata veya mutlak hata gibi, bir kayıp fonksiyonunun ve parametre (ya da parametreler) için bir prior dağılımı belirlenmesi gerekmektedir. Bütün bu işlemler çok zor ve uzun zaman alıcı olabilir. Diğer olası güçlük, θ'nın tek bir parametreye değil de bir parametreler vektörüne karşılık geldiği durumda karşımıza çıkmaktadır. Böyle bir durumda , çok boyutlu parametre uzayı Ω içinde anlamlı bir birleşik prior dağılım belirlemek çok zorlaşmaktadır.

Bu sorunları çözmek için basit bir yolun olmadığı vurgulanmalıdır. Prior dağılımlara ve kayıp fonksiyonu üzerine kurulmamış tahminin diğer metotları sadece teorik yapıda çok eksiklikler bulundurmamaktadırlar. Aynı zamanda pratikte kesin sınırlamalar barındırmaktadırlar.

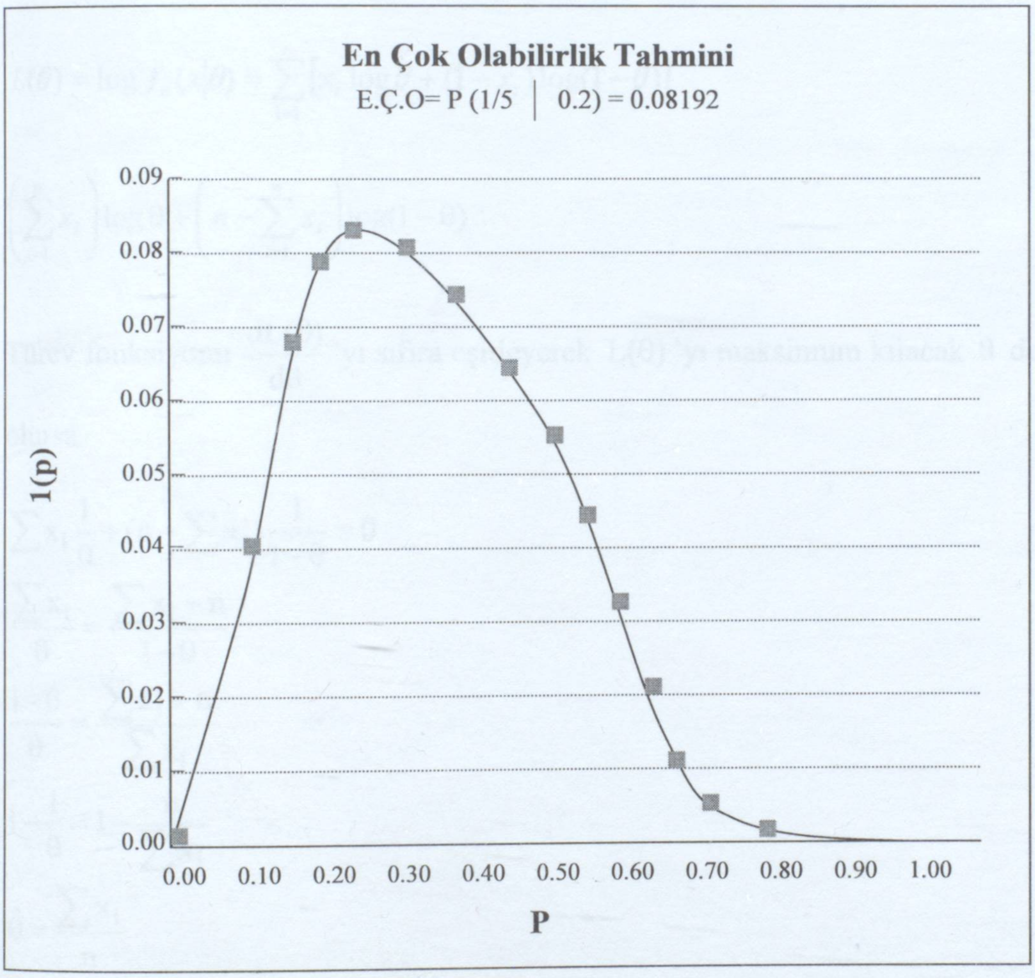
5.4.2 En Çok Olabilirlik Tahmin Edicisinin Tanımı

Az önce de değinildiği gibi bir kayıp fonksiyonu ve bir prior dağılım belirlemeden Bayesgil tahmincileri bulmak olanaksızdır. Yine de bu türden zorluklarla karşılaşmamak için başka bazı yöntemlere başvurmak yararlı olacaktır. Bu yöntemlerden bir tanesinin adı en çok olabilirlik yöntemidir. R.A.Fisher tarafından 1912'de geliştirilen bu yöntem, birçok probleme uygulanabilir. Güçlü sezgisel bir yapıya sahiptir. θ'nın mantıklı bir tahmin edicisini verir. Üstelik, örnek büyükse, θ'nın tahminindeki isabet artmaktadır.

Herhangi bir ana kütledeki X rastlantısal değişkeninin θ parametresine bağlı olasılık dağılımını f(x/ θ) ve bu ana kütleden çekilen n birimlik bir gözlemler $x_1, x_2, x_3, \ldots x_n$ olsun. Örnekteki birimlerinin bağımsız olmaları halinde birlikte meydana gelen olasılıkları

$$L(\theta):f(x_1, x_2, x_3, \ldots x_n; \theta) = \sum_{n=1}^{n} f(x;\theta) \text{ dır.}$$

Söz konusu L(θ) fonksiyonu çok olabilirlik fonksiyonu olarak tanımlanır.Burada, daha önceden de değinildiği gibi θ reel değerli bir parametre veya parametreler vektörü olabilir. Önceden olduğu gibi, f(x/θ), x vektörü için, θ'nın bir olasılık fonksiyonu olarak tanımlanır. En çok olabilirlik yöntemi bu fonksiyonu (ya da bu fonksiyonun logaritmasını) maksimize etmemizi sağlayacak parametre tahminlerini bulmamızı sağlar.



Şekil 5.1 E.Ç.O= P (1/5 | 0.2) = 0.08192

Yukarıdaki şekilden de anlaşılacağı gibi L(p) fonksiyonunu maksimum kılan p değeri 0,2 dir ve bu noktadaki L(p) değeri 0,08 civarındadır.

5.4.3 Bazı En Çok Olabilirlik Tahmin Edicilerinin Bulunuşu

Bazen L(θ)'yı maksimum kılan birden çok θ'nın tahmini değeri bulunabilir. Böyle bir durumda tek bir en çok olabilirlik tahmin edicisinden söz etmek mümkün değildir. Bulunan bu noktalardan herhangi bir tanesi θ'nın tahmini olarak seçilebilir. Bazen de , θ'nın bir E.Ç.O. tahmincisi mevcut olmayabilir. Bu durumda da yapacak pek bir şey yoktur.

Bir Bernoulli Dağılımından Örnekleme. $X_1 \ldots X_n$'in θ parametreli $(0 \le \theta \le 1)$ bir Bernoulli dağılımından çekildiği varsayılsın . Her $X_1 \ldots X_n$ için, olasılık yoğunluk fonksiyonu f(x/ θ)= (1)'dır.

f(x/ θ) fonksiyonunun olasılığını en yüksek dereceye çıkaran θ'nın değeri, f(x/ θ) logaritmasını en yüksek dereceye çıkaran θ'nın değerine eşittir. Bu nedenle en çok olabilirlik tahmin edicisi

$$f_n(x|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} \tag{5.32}$$

$$L(\theta) = \log f_n(x|\theta) = \sum_{i=1}^{n}\left[x_i \log\theta + (1-x_i)\log(1-\theta)\right] \tag{5.33}$$

$$\left(\sum_{i=1}^{n} x_i\right)\log\theta + \left(n - \sum_{i=1}^{n} x_i\right)\log(1-\theta) \tag{5.34}$$

Türev fonksiyonu $\frac{dL(\theta)}{d\theta}$'yı sıfıra eşitleyerek $L(\theta)$'yı maksimum kılacak θ değeri bulunacak olursa

$$\sum x_i \frac{1}{\theta} + (n - \sum x_i)\frac{1}{1-\theta} = 0$$

$$\frac{\sum x_i}{\theta} = \frac{\sum x_i - n}{1-\theta}$$

$$\frac{1-\theta}{\theta} = \frac{\sum x_i - n}{\sum x_i}$$

$$1 - \frac{1}{\theta} = 1 - \frac{n}{\sum x_i}$$

$$\hat{\theta} = \frac{\sum x_i}{n}$$

Buradan çıkan sonucua göre θ'nın en çok olabilirlik tahmin edicisi örnek ortalamasına eşittir: $\hat{\theta} = \bar{X}$ olur.

Bir Normal Dağılımdan Örnekleme. $X_1 \ldots X_n$, μ'nın ortalamasının bilinmediği ve varyans σ^2'ın bilindiği bir normal dağılımdan bir rassal örnek oluşturduğu varsayılsın. Her gözlemlenmiş $X_1 \ldots X_n$ değeri için, olasılık fonksiyonu f(x/μ),

$$f(x/\mu) = f_n(x|\mu) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right] \tag{5.35}$$

Denklem (5.35)'ten f(x/μ)'nün,

$$Q(\mu) = \sum_{i=1}^{n} (x_i - \mu)^2 = \sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \tag{5.36}$$

en aza indirgeyen μ'ün değeri ile en yüksek dereceye çıkaracağı görülebilir.

Türev fonksiyonu dQ(μ)/dμ 0'a eşitlenecek olursak $\hat{\mu} = \bar{X}$ bulunur. Dolayısıyla μ'nün en çok olabilirlik tahmincisi örnek ortalamasına eşit olmaktadır.

Not: μ'nün tahmincisinin bilindiği kabul edilen σ^2 varyansından etkilenmediği görebilir. Bilinmeyen ortalama μ'ün en çok olabilirlik tahmin edicisi, σ^2'in değerine bakmaksızın, örnek ortalama $\bar{X}$'ya eşit olmaktadır.

Bilinmeyen Varyanslı Bir Dağılımdan Örnek Alma. Yine $X_1 \ldots. X_n$ 'in normal bir dağılımdan rassal bir örnek oluşturduğu ama ortalama μ ve varyans σ^2'in bilinmediği varsayılsın. Bu durumda olabilirlik fonksiyonunun logaritması

$$L(\mu, \sigma^2) = \log f_n\left(x|\mu, \sigma^2\right)$$

$$-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{5.37}$$

Şeklindedir. Bu fonksiyonu maksimum kılmak için gerekli kısmi türevler alınıp sıfıra eşitlenecek olursa

$$\frac{\partial L(\mu, \sigma^2)}{\partial \mu} = 0 \tag{5.38}$$

$$\frac{\partial L(\mu, \sigma^2)}{\partial \sigma^2} = 0 \tag{5.39}$$

Bu denklemlerden yola çıkılarak şu bağıntı elde edilir:

$$\frac{\partial L(\mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) \tag{5.40}$$

Böylece $\hat{\mu} = \bar{X}$ bulunur.

Bunun yanı sıra

$$\frac{\partial L(\mu,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{5.41}$$

μ'ya, yeni elde edilen $\hat{\mu} = \overline{X}$ değeri verildiğinde

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n} \quad \text{bulunur.} \tag{5.42}$$

Bu denklemin sağ tarafında yer alan istatistik örnek varyansı

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1} \text{ ile ilişkilidir.}$$

Şöyle ki

$$\frac{n-1}{n}S^2 = \hat{\sigma}^2 \quad \text{olmaktadır.} \tag{5.44}$$

Tekdüze Bir Dağılımdan Örnek Alma. $X_1 \ldots\ldots X_n$ 'in, θ parametresinin değerinin (θ >0) bilinmediği, (0, θ) aralığındaki bir tekdüze dağılımdan çekildiğini varsayalım.. Her bir gözlemin olasılık dağılım fonksiyonu f(x/θ) aşağıdaki şekle sahiptir:

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} & \text{için } 0 \le x \le \theta \\ 0 & \text{değilse} \end{cases} \tag{5.45}$$

Böylece $X_1 \ldots\ldots X_n$'in bileşik olasılık fonksiyonu olan f(x/θ) şu şekli alır:

$$f_n(x|\theta) = \begin{cases} \frac{1}{\theta^n} & \text{için } 0 \le x_i \le \theta \ (i = 1, \ldots, n) \\ 0 & \text{değilse} \end{cases}. \tag{5.46}$$

Yukarıdaki denklemden de görüldüğü gibi θ'nin E.Ç.O, 1/θ en yüksek değerini veren bir değer olmalıdır. Bu değer θ=Maks($x_1 \ldots\ldots x_n$) olacağından:

θ'nın en çok olabilirlik tahmin edicisi : θ =Maks($X_1 \ldots\ldots X_n$) olur.

Burada en çok olabilirlik tahmin edicisinin uygun bir tahmin edici olmadığı belirtilmelidir. e Maks($X_1 \ldots\ldots X_n$)< θ olduğu için, en çok olabilirlik tahmin edicisini, θ'nın değerini olduğundan daha az tahmin etmekte olduğu görülmektedir.

Örnek 5: Bir en çok olabilir tahmin edicisi'nin tek olmaması. $X_1 \ldots X_n$ 'in (θ, θ +1) aralığındaki bir tekdüze dağılımdan rassal bir örnek oluşturduğu varsayılsın. Birleşik olasılık yoğunluk fonksiyonu $\theta < x < \theta + 1 (i=1 \ldots n)$ için

$$f_n(x|\theta) = \begin{cases} 1 & \text{için } \theta \leq x_i \leq \theta + 1 \; (i = 1, \ldots, n) \\ 0 & \text{değilse} \end{cases} \tag{5.48}$$

şeklindedir.

$i=1 \ldots n$ için $\theta < x$ koşulu $\theta < Min(x_1 \ldots x_n)$ koşuluna denktir. Benzer şekilde, $i=1 \ldots n$ için $x < \theta + 1$ koşulu da, $\theta > Maks(x_1 \ldots x_n)$ koşuluna denk olmaktadır. Bu yüzden

$Maks(x \ldots x) - 1 < \theta < Min(x_1 \ldots x_n)$ için ve

$$f_n(x|\theta) = \begin{cases} 1 & \text{için } Maks(x_1, \ldots, x_n) - 1 \leq \theta \leq Min(x_1, \ldots, x_n) \\ 0 & \text{degilse} \end{cases} \tag{5.49}$$

Böylelikle θ'ın herhangi bir değeri $Maks(x_1 \ldots x_n) - 1 < \theta < Min(x \ldots x)$ aralığında bir E.Ç.O tahmini değeri olarak seçilebilir.

BÖLÜM 6 : EN YÜKSEK OLABİLİRLİK TAHMİN EDİCİLERİNİN ÖZELLİKLERİ

6.1 Değişmezlik

X_1……X_n değişkenlerinin, olasılık fonksiyonu ya da olasılık yoğunluk fonksiyonu $f(x/\theta)$ olsun. Başka bir deyişle (X_1……X_n)'in θ parametresinin bilinmediği bir dağılımdan rassal bir örnek oluşturduğu varsayılsın. Θ de, θ'nin E.Ç.O. tahmincisi olsun. Şimdi dağılımdaki θ parametresini bir fonksiyonu olan g(θ) dikkate alınsın. Eğer Θ, θ'nin en yüksek olabilirlik tahmin edicisiyse, o zaman g(Θ) de, g(θ)'nin en yüksek olabilirlik tahmin edicisi olmaktadır. Sözgelimi tahmincisi en yüksek olabilirlik yöntemine göre bulunmuş olan θ parametresi sözkonusu olasılık dağılımının varyansına $\left(\theta=\sigma^2\right)$ eşit olsun. Bu durumda standart sapmanın (σ) en çok olabilirlik tahmincisi de hiç şüphesiz θ'nın en çok olabilirlik tahmincisinin kareköküne eşit olmaktadır. Bu özelliğe değişmezlik özelliği (invariance property) denmektedir. En çok olabilirlik tahmincileri değişmezlik özelliğine sahiptirler.

6.3 Tutarlılık

Örnek birim mevcudu n sonsuza giderken $n\to\infty$ Θ tahmin edicisinin örnekleme dağılımının ortalaması θ etrafında yoğunlaşması halinde tutarlıktan söz edilir. Diğer bir yönden Θ' nın odak olduğu bir yoğunlaşma için asimptotik sistematik hatasızlık gereklidir. Küçük örnekten büyük örneğe geçildiğinde sistematik hata azaldığı için Θ tahmin edicilerinin örnekleme dağılımı da θ etrafında toplanacaktır.

6.4 Yeterli İstatistikler

Hatırlanacağı gibi (X_1…X_n) gözlem değerlerinden oluşan rassal bir örnekteki gözlemlerin T(X_1…X_n) gibi reel değerli fonksiyonu bir istatistik olarak adlandırılır. $\theta\in\Omega$'nın herhangi bir sabit değeri için, her belli istatistik T'nin dağılımı, (X_1….X_n)'in bileşik olasılık yoğunluk fonksiyonundan alınabilir. Genelde, bu dağılım θ'nın değerine dayanır. Bu nedenle,

$\theta\in\Omega$'nın farklı olası değerlerine karşılık gelen T'nin olasılık dağılımlarının oluşacağını söyleyebiliriz. T verildiğinde (X_1….X_n)'nin koşullu olasılığının θ'ya bağlı olmaması halinde ise T istatistiğine θ'nın yeterli bir istatistiği denmektedir.

T'nin yeterli bir istatistik olması halinde (başka bir deyişle $f\left(X_1, X_2, \ldots, X_n / T\right)$ koşullu olasılığının θ'ya bağlı olmaması, sözkonusu örnek yerine onun bir istatistiği olan T'nin geçirilebileceği (ikame edilebileceği) anlamına gelir. Çünkü madem bu koşullu olasılık θ'ya bağlı değildir o halde böyle bir örnekten yola çıkılarak θ hakkında (yeni bir) bilgi edinilemez. Öte yandan bilindiği gibi istatistik oluşturmanın bir amacı örnekten gelen verinin özetlenmesidir. Örneğin, belirli koşullar altında örnekten gelen ($X_1 \ldots . X_n$) gözlemleri yerine örnek ortalaması, örnek varyansı vb. istatistikler kullanılarak boyut indirgenebilir. n tane gözlem değeri yerine onların bazı istatistiklerini kullanmak hiç şüphesiz daha pratik olacaktır. T'nin yeterli bir istatistik olması halinde artık ($X_1 \ldots . X_n$) gözlemleri yerine bu T istatistiği kullanılabilir.[5] Çünkü bu gözlemler kümesi T'nin varlığında θ hakkında (yeni bir) bilgi vermeyecektir.

Öte yandan yeterli istatistiği az önceki tartışma çerçevesinde rassal örnekteki tüm bilgiyi kapsayan istatistik olarak da tanımlamak mümkündür.

[5] Mood, A.M.; Graybill; F.A.; Boes, D. C.; "Introduction to the Theory of Statistics" McGraw-Hill International Editions, Statistics Series, Third Edition, s300-301

BÖLÜM 7: ÖN BİLGİNİN BİR DAĞILIM BİÇİMİNDE BELİRLENMESİ

.parametre θ ön yoğunluk fonksiyonu subjektif

1:kesikli

2:sürekli

kesikli halinde tanımını yapılır. Sürekli olması halinde ise öncelikle olasılığı bulunur. Bunun içinde bilgiler dağılıma dönüştürülecek ondan sonra tanıtılması yapılma aşamasına gelecektir.(Kahyaoğlu,1999)

7.1.1 Histogram Yaklaşımı

θ'nın dağılımını bulunmasındaki yaklaşımlardan histogram yaklaşımından bahsedilmek istenirse buradaki metot parametre uzay olarak ele alınır ve eşit şekilde bölmelere ayrılır. Her bölmenin kendisine ait bir subjektif olasılığı olur. Bu olasılıklardan bir histogram meydana gelir. Olasılık yoğunluk fonksiyonu histogramdan faydalanılarak tahmin edilebilir..

7.1.2 Göreceli (Nisbi) Benzerlik Yaklaşımı

Buradaki ön yoğunluk fonksiyonun tanımlanmasındaki yöntem daha çok her noktaysa ait kişisel yakınlıklarının birbirleri ile ilişkilendirilip tartıştırılması sonucu ortaya çıkar.

Çerçeve θ'nın sınırlanmamış şeklindeki haliyle sınırlanmış olan haline göre daha zordur. Benzerlik yapmak için sadece sınırlı bölge kullanılmaktadır.

Burada iki teknikten bahsedilmelidir: İlk teknik sınırlı bölgenin dışında kalan alanın olasılık yoğunluk fonksiyonunun tanımlanmasını sağlamaktır.Daha sonraki teknik ise belki yoğunluğun normalleştirilmesini sağlamak olacaktır.

7.1.3 Verilen Fonksiyonel Biçim İle Eşleştirme

Ön olasılığını tanımlanmasında kullanımına en sık rastlanılan önceden verilmiş bir fonksiyon biçimiyle eşleştirme yöntemidir. Varsayım olarak P(θ)'nın verilmiş bir fonsiyonel bir form olduğunu ve formun olasılığını daha yakın olacak bir şekilde seçilmesini sağlamaktır. Bunun için yapılması gereken en önemli iş, tahmin edilmiş momentler yöntemini kullanmak olacaktır. Varsayım olarak parametrelerin normal dağıldığı varsayılmaktadır. Yoğunluk formülünün ortaya çıkartılması için ortalama ve varyans bulmamız gerek olacaktır. Karşılaşacak en önemli sorun fonksiyon kuyruklarının momentlerin belirlenmesi üzerindeki etkisidir.

Eğer, [0,∞] arasındaki bir yoğunluğun kuyruğu K. θ_2 biçiminde olursa;

$$\int_b^\infty \theta(K\theta^{-2})d\theta = K\int_b^\infty \theta^{-1}d\theta = \infty \qquad \forall b > 0 \tag{7.1}$$

şeklinde elde edilecektir.

Moment değerine sahip olan bu yoğunluk k küçüldüğü takdirde etki anlamsız kalacaktır. Bu yoğunluk, bir moment değerine sahip olamaz. Ancak K küçükse bu kuyruk çok önemsiz ve anlamsız bir olasılık değerine sahip bulunacaktır. Bu subjektif olarak oldukça belirli olasılıkları ifade ettiği için küçük olasılığın kuyruğu kesin olarak bilinemez. Bu durumun momentler üzerindeki etkisi, sınırsız parametre uzayı ile ilgilenildiği zaman büyük olacaktır.

Eğer θ, [0,1] gibi sınırlı bir parametre uzayında ise, ön momentlerin sübjektif olarak belirlenmesi daha mantıklıdır. Bunun sebebi kuyruğun momentler üzerindeki etkisinin azlığıdır.

Parametrelerin belirlenmesinde daha iyi bir yöntem olarak, yığına ait dağılımın çeşitli kısımlarını, sübjektif olarak tahmin etmek ve daha sonra bu kısımlara mümkün olduğu kadar yakın bir yoğunluk elde etmek için verilen fonksiyonel şeklin parametrelerini seçmek önerilmektedir Sürekli bir dağılımın α kısmı, bu dağılıma sahip rassal bir değişkenin Z(α)'ya eşit veya küçük olması olasılığını a olarak veren bir Z(α) noktasıdır. Bu, sübjektif olarak yapılması en kolay olan, bölgelerin olasılıklarının tahminini kesin bir şekilde ifade eden yöntemdir. (Kahyaoğlu,1999)

7.1.4 Kümülatif Dağılım Fonksiyonunun Saptanması

Ön yoğunluk fonksiyonun belirlemek için kümülatif dağılım fonksiyonunun subjektif olarak meydana getirilmesi gerekmektedir. Burada eğri metodu kullanılmaktadır. Bu noktalar bir araya geldikleri zaman anlaşılır bir fonksiyonel eğri ortaya çıkmaktadır.

7.2 Ön Yoğunluk Fonksiyonu Belirlenmesinde Marjinal Dağılımın Kullanılması

X in marjinal dağılımından bahsetmek istersek x in f(x) olasılık dağılımına sahip olduğu ve parametre olarak ta q nın p(q) olasılık yoğunluk fonksiyonu bilindiği varsayımı altında x ve q bileşik yoğunluğu

$$h(X/\theta) = f(X/\theta).P(\theta) \tag{7.2}$$

marjinal yoğunluk

$$m(X|P) = \int_{\theta} f(X|\theta) df^{p}(\theta) = \int_{\theta} f(X|\theta) Pf(\theta) df \quad \text{sürekli} \tag{7.3}$$

$$m(X|P) = \int_{\theta} f(X|\theta) df^{p}(\theta) = \sum_{\theta} f(X|\theta) P(\theta) \quad \text{kesikli} \tag{7.4}$$

marjinal dağılımı için kullanılır.(Kahyaoğlu,1999)

7.3 Hiyerarşik Dağılımlar

Çok aşamalı dağılımların en önemli olan özelliği yapısal ve ön bilgiye sahip olup aynı zamanda kullanmalarıdır. Tanımlamak gerekirse;

$$P_i(\theta) = \prod_{i=1}^{k} P_0(\theta_i) \tag{7.5}$$

Burada yapısal ve ön bilgi mevcuttur.

Hiyerarşik yaklaşımda sübjektif olasılık ikinci aşamaya yerleştirilerek yığına ait bir dağılım oluşturulmaya çalışılır. İlk aşama içinde sübjektif olan P_1, böylece ikinci aşama için kullanılacaktır.

İlk aşama önlerinin meydana getirdiği bir Γ sınıfına sebep oluyorsa, ikinci aşama Γ'daki ön dağılımı yerine koyarak elde edilecektir Kısacası Γ üzerine bir ön dağılım oluşturulacaktır. Belirli bir fonksiyonel şeklin örneklerinden meydana gelen Γ, ilk aşama için çok daha önemli bir ön sınıftır. Γ'yı şu biçimde tanımlamak mümkündür.

$$\Gamma = P_1(\theta / \lambda) : P_1 \qquad \lambda \in \wedge \tag{7.6}$$

İkinci aşama, bir ön dağılımın yerine konması yolu ile elde edildiğine göre, λ yerine $P_2(\lambda)$ yerleştirildiğinde bu aşama tamamlanacaktır. $P_2(\lambda)$ ifadesinde hiper-parametre olarak isimlendirilen λ'dır. Böylece iki aşamalı önden bir hiper-ön elde edilmiş olur. P_1'in bağımsızlığının yapısal tahmininin normal dağıldığı varsayımı altında;

$$\Gamma\{P_1(\theta / \lambda) : P_1(\theta / \lambda)\} = \prod_{i=1}^{k} P_0(\theta_i) \tag{7.7}$$

elde edilir.

P_0 normal dağılıyor ise, $N(\mu_p, \sigma_p^2)$, $-\infty < \mu_p < +\infty$, $\sigma_p^2 > 0$ olacaktır. İkinci aşama ön dağılımı olan $P_2(\lambda)$ sübjektif kararlara karşılık gelen hiper-parametreler için seçilebilir. Bazen ikinci aşamadaki ön dağılımların sübjektif olarak belirlenmelerinin kesinliği, tek bir ön dağılım belirlenmesinden daha etkindir. Dolayısıyla P_2'nin yanlış tanımlanmasının oluşturacağı tehlike daha azdır denilebilir.

Hiyerarşik dağılımların iki aşamayla sınırlanmasının teorik bir nedeni yoktur. Fakat ikiden fazla uygulamanın da kullanımda pek fazla bir yararı olmayacaktır. Hiyerarşik yaklaşımın hesaplamada avantajlı olması nedeniyle Bayes yaklaşımı için de önemli bir yöntem olduğunu belirtmek gerekir. Böylece hiyerarşik bir yapıdan elde edilen dağılımı standart bir dağılım olarak kullanmanın bir sakıncası olmayacaktır. Bayeste hiyearşik yaklaşımın yaygın olarak kullanılmasının en önemli nedenleri arasında buradan elde edilecek dağılımın standart halde olmasıdır.Ve bu şekilde kullanılabilir olması da büyük bir avantaj yaratmaktadır. (Kahyaoğlu,1999)

7.4 En Büyük Entropi Dağılımları

Entropi belirsizlik ölçütüdür. Dağılmanın büyüklüğü ile vermiş olduğu bilginin miktarı ters orantılıdır.

Entropi, θ'nın kesikli olduğu durumlarda kullanılabilir. $\varepsilon_n(P)$ ile gösterilen P'nin entropisi, kesikli durumda

$$\varepsilon_n(P) = -\sum_{\varsigma} P(\theta_i)\log P(\theta_i) \tag{7.8}$$

olarak tanımlanır.

Eğer $P(\theta_i) = 0$ ise; $P(\theta_j)\log P(\theta_j)$ sıfır olarak elde edilecektir. $\theta = \{\theta_1, \theta_2, \theta_n\}$ ise, $P(\theta_k) = 1$ olduğunda, $P(\theta_1) = 1$ iken $i \neq k$ için parametre noktasını tamamen ifade eden olasılık dağılımı ortaya çıkacaktır. Böylece belirsizlik sıfırdır. Bu durumda,

$$\varepsilon_n(P) = -\sum_{i=1}^{n} P(\theta_i)\log P(\theta_i) = 0 \tag{7.9}$$

ifadesi yazılabilir.

Belirsiz ya da en yüksek entropi olasılık dağılımı, bütün i ler için

$P(\theta_1) = 1/n$

biçiminde ifade edilir. Bu, P için şöyle yazılır.

$$\varepsilon_n(P) = -\sum_{i=1}^{n} \frac{1}{n}\log\left(\frac{1}{n}\right) = \log n \tag{7.10}$$

Bütün uygun P'ler için $\varepsilon_n(P) \leq \log n$ durumu söz konusudur. Kesikli bir θ için, bu en büyük entropi dağılımı bilgi verici olmayan ön dağılım ile aynıdır. (Kahyaoğlu,1999)

BÖLÜM 8: REGRESYON MODELLERİ

8.1. 1 Modelin Ana Hatları

İktisadi olaylarda bağımlı değişken çok ender olarak tek bir bağımsız değişkene bağlıdır. Genellikle bir iktisadi olayda bağımlı değişken birden çok bağımsız değişkenin etkisi altındadır. Bağımlı değişken ile k bağımsız değişken X_1, X_2, ..., X_k arasındaki bağıntıyı doğrusal model içinde ele alırsak modeli

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + u_i \tag{8.1}$$

$$Y_k = \beta_0 + \sum_{j=1}^{k} \beta j X_{ji} + u_i \tag{8.2}$$

veya

$$Y_i = \sum_{j=1}^{k} \beta j X_{ji} + u_i \qquad X_{0i} = 1 \text{ her i için} \tag{8.3}$$

olarak ifade edebiliriz.

Modelde Y_i niceldir. Buna karşın bağımsız değişkenlerin tümünün nicel, tümünün nitel, kiminin nicel kiminin nitel olması gibi seçenekler söz konusu olabilir.

Bağımsız değişkenlerin tümünün nicel olması halinde Regresyon Modeli, tümünün nitel olması durumunda da varyans analizi karşımıza çıkmaktadır. Kovaryans analizi ise hem nicel hem de nitel bağımsız değişkenlerden oluşmaktadır. Bu bağlamda regresyona gölge değişkenlerin eklenmesi ile bu model de hem nicel hem de nitel bağımsız değişkenlerden oluşacaktır.

O halde regresyon ile kovaryans modelleri arasındaki farkın vurgulanması gerekmektedir. Kovaryans modellerinde nitellik ön planda olup ilgi alanını oluşturan başlıca bağımsız değişkenler niteldir. Nicel bağımsız değişkenler ise sadece hata payları varyansını düşürmek için model kapsamına alınmıştır. Halbuki Regresyon'da nicellik odak olup nitel değişkenler verilerdeki farklılığı yansıtmaktadır.

8.1.2 Çoklu Doğrusal Regresyon Modeli

İki değişkenli doğrusal regresyon modelinin bir uzantısı olan bu model birden fazla bağımsız değişkene sahip regresyon modellerin açıklanmasında sık sık uygulanan bir tekniktir.

Çoklu doğrusal regresyon modelinin en basit biçimi bir bağımlı ve iki bağımsız değişkenden oluşan birinci mertebeden bir modeldir:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + u_i \tag{8.4}$$

Bu doğrusal modeli ikinci mertebeye yükseltebilmek için

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_{11}.X_{21i} + u_i \tag{8.5}$$

yazabilir.

Basit doğrusal regresyonda bağımlı değişkendeki yani etkilenen değişkendeki değişmeler bir tek bağımsız değişken, yani etkileyen değişken ile açıklanıyordu. Genelde bir bağımlı değişkendeki değişimleri başka değişkenle açıklamanın yeterli olmayacağı açıktır. Sosyal ve fen bilimlerinde ele alınan olaylarda çok sayıda faktörün etkisi söz konusu olacağından, modelde daha fazla bağımsız değişkene ihtiyaç duyulacaktır. (Genceli,M.,1989)

8.3.3. Çoklu Doğrusal Regresyon Modelinin Matrislerle Gösterilmesi

Örnek n gözlemden oluştuğundan her bir gözlem Y, X_2, X_3, ..., X_k değişkenleri için birer değerden oluşacaktır. Model,

$$Y_i = \beta_1 + \beta_2 \xi_{12} + \beta_{13} + ... + \beta_k X_{ik} + u_i$$

olarak tanımlandığından herbir gözlem için,

$$Y_1 = \beta_1 + \beta_2 X_{22} + \beta_3 X_{23} + \beta_k X_{2\bar{k}} + u_1$$

$$Y_2 = \beta_1 + \beta_2 X_{32} + \beta_3 X_{33} + ... + \beta_k X_{3k} + u_2$$

$$Y_{n-1} = \beta_1 + \beta_2 X_{(n-1)2} + \beta_3 X_{n3} + ... + \beta_k X_{nk} + u_n$$

$$Y_n = \beta_1 + \beta_2 X_{n2} + \beta_3 X_{n3} + ... + \beta_k X_{nk} + u_n$$

eşitlikleri yazılabilir.

Bu eşitlikleri matrislerle ifade edersek,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_{n-1} \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{12} & X_{13} & \cdots & X_{1k} \\ 1 & X_{22} & X_{23} & \cdots & X_{2k} \\ 1 & X_{32} & X_{33} & \cdots & X_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{(n-1)2} & X_{(n-1)3} & \cdots & X_{(n-1)k} \\ 1 & X_{n2} & X_{n3} & \cdots & X_{nk} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{k-1} \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_{n-1} \\ \varepsilon_n \end{bmatrix}$$

olacaktır. Görüldüğü gibi β_1 modelde yer aldığından ve değişkeni olmadığından, X' ler matrisi yazılırken değerleri 1 olan bir X_1 değişkeni varmış gibi birinci sütunda 1 değerleri yer almıştır. (Neil.H. .,1975)

Yukarıda yer alan matris ve vektörleri,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{nx1} \qquad X = \begin{bmatrix} 1 & X_{12} & X_{13} & \cdots & X_{1k} \\ 1 & X_{22} & X_{23} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n2} & X_{n3} & \cdots & X_{nk} \end{bmatrix}_{nxk}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}_{kx1} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{nx1}$$

olarak tanımlarsak model,

$Y = \beta X + \varepsilon$

olarak ifade edilecektir.

Oluşturulan X matrisi vektörlere bölünebilir. Bu durumda vektörler,

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} X_2 = \begin{bmatrix} X_{12} \\ X_{22} \\ \vdots \\ X_{n2} \end{bmatrix} X_3 = \begin{bmatrix} X_{13} \\ X_{23} \\ \vdots \\ X_{n3} \end{bmatrix} \cdots X_k = \begin{bmatrix} X_{1k} \\ X_{2k} \\ \vdots \\ X_{nk} \end{bmatrix}$$

olacaktır. modelin bu vektörlerle,

$Y = X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + ... + X_k\beta_k + \varepsilon$

şeklinde gösterilmesi de mümkündür.

8.3 4 Modelin Varsayımları

1) Hata payları birer rastlantı değişkeni olup

$$E(u_i) = 0 \tag{8.6}$$

2) u normal olarak dağılmaktadır:

$$u_i \sim N(0, \sigma^2 u) \tag{8.7}$$

3) Homoskedasite söz konusudur:

$$E(u_i^2) = \sigma^2 u \tag{8.8}$$

4) Otokorelasyon bulunmamaktadır

$$Kov(u_i u_j) = 0 \quad i \neq j \tag{8.9}$$

5) u_i hata payları bağımsız değişkenlerden bağımsızdır.

$$Kov(u_i, X_{1i}) = Kov(u_i, X_{2i}) = \ldots = Kov(u_i, X_{ki}) = 0 \tag{8.10}$$

6) Bağımsız değişkenler arasında kesin bir doğrusal bağlantı bulunmamaktadır. Bağımsız değişkenlerden hiçbirinin diğerlerinin doğrusal bağlantısı yoktur. Örneğin $X_{2i} = 4X_{3i}$ veya $X_{3j} = 5\text{-}4\ X_{2j} + 9X_{3j}$ çoklu doğrusal bağlantıyı yansıtmaktadır. Buna karşın $X_{2i} = X^2_{3i}$ veya $X_{4i} = X_{5i}.X_{7j}$ gibi doğrusal olmayan ilişkiler çoklu doğrusal bağlantı olmadığı varsayımını değiştirmektedir.

Bu bağlamda 6. varsayımın örneğe ilişkin bir sorun olduğuna ve dolayısıyla hipotez testine konu olamayacağına değinilmelidir.

7) Bağımsız değişkenlerde ölçme hatası bulunmamaktadır.

8) Gözlem sayısı, n, parametre sayısından, p, fazladır: $n > p$

Bu varsayımları gerçekleştiren model yerine "Klasik Normal Çoklu Regresyon" modeli denilmektedir. İki bağımsız değişkenden oluşan en basit çoklu doğrusal regresyonun herhangi bir birimi;

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2j} + u_j \tag{8.11}$$

dir. Beklenen değer alındığı takdirde de ana kütle regresyon denklemi elde edilir:

$$E(Y_i) = E(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i) \tag{8.12}$$

β_0 parametresi regresyon sabiti veya kesişimidir. Regresyon yüzeyi ile Y'nin kesişimini vermektedir, diğer bir deyişle $X_1 = X_2 = 0$ olduğu zaman E(Y)'nin orijinden yüksekliğini göstermektedir. Ekonometrik açıdan ise, β_0, modele alınmayan diğer değişkenlerin Y'ye yaptıkları ortalama etkidir. $X_1 = 0$, $X_2 = 0$'ın modelin veri kapsamında bulunması halinde E $(Y/X_1 = 0, X_2 = 0)$ değerini vermektedir. Aksi takdirde özel bir anlamı bulunmamaktadır.

Genel olarak β_1, ... β_k'ya kısmi regresyon parametreleri denilmesinin sebebi bunların $X_1, X_2, ..., X_k$'nın kısmi etkilerini göstermeleridir.

β_k yorumlanırken bu parametreye ilişkin değişken, X_k dışındaki diğer tüm bağımsız değişkenlerin etkilerinin sabit olduğu varsayılmaktadır. Bu varsayım altında X_k'daki bir birim değişmenin bağımlı değişkenin ortalamasında ne kadar değişmeye yol açtığını β_k vermektedir.

Matematik olarak da burada;

$$\frac{\partial E(Y / X_1,...,X_k)}{\partial X_k} = \beta_k \tag{8.13}$$

kısmi türevi alınarak aynı sonuca ulaşılır.

β_1, ... β_k'ya işaretleri ilişkin oldukları bağımsız değişken ile bağımlı değişken arasındaki bağıntının yönünü vermektedir. Buna karşın hangi bağımsız değişkenin etkisinin daha fazla olduğu hususunda β'nin büyüklüğü gerek olmakla birlikte yeter değildir. Parametrelerin karşılaştırmalarda kullanılabilmesi için tüm bağımsız değişkenlerin aynı ölçü birimine tabi olmaları gerekir. İktisadi olaylarda çoğu kez rastlanıldığı gibi X_1, X_2,..., X_k bütün bağımsız değişkenler aynı ölçüye, örneğin TL'ye bağlı iseler parametrelerin büyüklüğü etkili olmada bir ölçüdür. Aksi takdirde, yani aynı ölçü birimleri söz konusu değilse böyle bir karşılaştırma (Genceli,M.,1989)

8.3.5 Parametrelerin Tahmini

n birimlik bir örnekten $p = k + 1$ parametre tahmini elde etmek istendiği takdirde gene E.K.K. yöntemine başvurulacaktır.

Aranılan örnek regresyon denklemi

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 \text{ dir.} \tag{8.14}$$

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + e_i \tag{8.15}$$

$$e_i = Y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}\right) \tag{8.16}$$

$$\Sigma e_i^{\,2} = S = \sum_{t=t}^{n}(Y_i - \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}) \rightarrow \min \tag{8.17}$$

Minimum olma gerekli koşulu

$$\frac{\partial S}{\partial \hat{\beta}_0} = 0 \quad ; \quad \frac{\partial S}{\partial \hat{\beta}_1} = 0 \quad ; \quad \frac{\partial S}{\partial \hat{\beta}_2} = 0 \quad \text{dir.} \tag{8.18}$$

Genel olarak, k parametreden oluşan doğrusal bir regresyonun çözümü için, 0(0,0,...) orijini kullanıldığı, yani asli değişkenler tercih edildiği takdirde p = (k+1) kısmi türev alınmaktadır.

$$\frac{\partial S}{\partial \hat{\beta}_0} = -2\Sigma(Y_i - \beta_0 - \beta_1 X_{1i} - \hat{\beta}_2 X_{2i} - ... - \beta_k X_{ki}) = 0$$

$$\frac{\partial S}{\partial \hat{\beta}_1} = -2\Sigma(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - ... - \hat{\beta}_k X_{ki})(X_{1i}) = 0$$

$$\frac{\partial S}{\partial \hat{\beta}_k} = -2\Sigma(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - ... - \hat{\beta}_k X_{ki})(X_{ki}) = 0 \tag{8.19}$$

Buradan da şu normal denklemler bulunur:

$$\Sigma Y_i = n.\beta_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 \Sigma X_{2i} + ... + \hat{\beta}_k \Sigma X_{ki} \tag{8.20}$$

$$\Sigma X_{1i} Y_i = \hat{\beta}_0 \Sigma X_{1i} + \hat{\beta}_1 X^2{}_{1i} + \hat{\beta}_2 \Sigma X_{1i} \Sigma X_{2i} + ... + \hat{\beta}_k \Sigma X_{1i} \Sigma X_{ki} \tag{8.21}$$

$$\Sigma X_{2i} Y_i = \hat{\beta}_0 \Sigma X_{2i} + \hat{\beta}_1 \Sigma X_{1i} \Sigma X_{2i} + \hat{\beta}_2 \Sigma X^2{}_{2i} + ... + \hat{\beta}_k \Sigma X_{2i} \Sigma X_{ki} \tag{8.22}$$

$$\Sigma X_{ki} Y_i = \hat{\beta}_0 \Sigma X_{ki} + \hat{\beta}_1 \Sigma X_{1i} \Sigma X_{ki} + \hat{\beta}_2 \Sigma X_{2i} \Sigma X_{ki} + ... + \hat{\beta}_k \Sigma X^2{}_{ki} \tag{8.23}$$

8.3.6 Genel Doğrusal Regresyon Modelinin Matris Yaklaşımı

Genel doğrusal model,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_{p-1} X_{i.p-1} + u_i$$

olarak ifade edilirse, matris formunda aşağıdaki matrisler tanımlanmalıdır:

$$\underset{nx1}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{bmatrix} \qquad \underset{nxp}{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1.p-1} \\ 1 & X_{21} & X_{22} & \dots & X_{2.p-1} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & X_{n1} & X_{n2} & \dots & X_{n.p-1} \end{bmatrix}$$

$$\underset{px1}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_{p-1} \end{bmatrix} \qquad \underset{nx1}{u} = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix}$$

Y ve u vektörlerinin basit doğrusal regresyon modeli için benzer olduklarına dikkat edilmelidir. β vektörü ilave regresyon parametrelerinden oluşmaktadır.

Matris terimleriyle genel doğrusal regresyon modeli:

$$\underset{nx1}{Y} = \underset{nxp}{X}\ \underset{px1}{\beta} + \underset{nx1}{u}$$

Y bağımlı değişken vektörü

β parametre vektörü

X bağımsız değişken matrisi

u bağımsız normal rastlantısal değişken vektörü

$E\{u\} = 0$ ve varyans- kovaryans matrisi:

$$\underset{nxn}{\sigma^2}\{u\} = \begin{bmatrix} \sigma^2 & 0 & 0 & . & 0 \\ 0 & \sigma^2 & 0 & . & 0 \\ 0 & 0 & \sigma^2 & . & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & . & \sigma^2 \end{bmatrix} = \sigma^2 I$$

$$\underset{nx1}{E\{Y\}} = X\beta \tag{8.24}$$

ve Y'nin varyans kovaryans matrisi u'ya eşittir:

$$\underset{nxn}{\sigma^2}\{Y\} = \sigma^2 I \tag{8.25}$$

8.4 REGRESYON PARAMETRELERİNİN TAHMİNİ

Genel doğrusal regresyon modeli için en küçük kareler kriteri :

$$Q = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - .... - \beta_{p-1} Xi.p - 1)^2 \tag{8.26}$$

$\beta_0, \beta_1, ... \beta_{p-1}$ parametrelerinin en küçük kareler yöntemi tahmincileri Q'yu minimize eden tahmincilerdir. Tahmin edilen regresyon katsayılarını $b_0, b_1, ..., b_{p-1}$ vektörü göstermektedir:

$$\underset{px1}{b} = \begin{bmatrix} b_0 \\ b_1 \\ . \\ . \\ b_{p-1} \end{bmatrix} \tag{8.27}$$

genel doğrusal model için en küçük karelerin normal denklemleri aşağıdaki gibidir:

$$X^{'}Xb = X^{'}Y \tag{8.28}$$

ve en küçük kareler tahmincileri:

$$\underset{px1}{b} = \underset{pxp}{(X'X)^{-1}} \underset{px1}{(X'Y)} \tag{8.29}$$

$$\mathrm{Var(b)} = (\mathrm{X'X})^{-1}\hat{\sigma}^2$$

$$\hat{\sigma}^2 = \frac{\sum \mathrm{e_i^2}}{\mathrm{n-p}} \tag{8.30}$$

Seklinde tahmin edilir.

8.4.1Öngörü Değerleri ve Kalıntılar

$\hat{Y}_i$ vektörü yerine $\hat{Y}$ ve $e_i = Y_i - \hat{Y}_i$ vektörü yerine e notasyonu kullanılacaktır:

$$\underset{nx1}{\hat{Y}} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ . \\ . \\ \hat{Y}_n \end{bmatrix} \qquad \underset{nx1}{e} = \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ e_n \end{bmatrix}$$

öngörü değerleri,

$$\underset{nx1}{\hat{Y}} = Xb = X\left(X'X\right)^{-1} X'Y = HY$$

$$(H = X(X'X)^{-1}X')$$

sonucuna ulaşılacaktır. H ile gösterilen nx1 boyutlu matris Y gözlem değerlerini $\hat{Y}$, öngörü değerlerine dönüştürmektedir. Bu nedenle de H, dönüşüm matrisi olarak adlandırılmaktadır.

Kalıntılar,

$$\underset{nx1}{e} = Y - \hat{Y} = Y - HY = \left(I - H\right)Y \tag{8.31}$$

kalıntıların varyans- kovaryans matrisi,

$$\underset{nxn}{\sigma^2\{e\}} = \sigma^2\left(I - H\right) \tag{8.32}$$

tahmincisi,

$$\underset{nxn}{s^2\{e\}} = MSE\left(I - H\right) \tag{8.33}$$

şeklinde bulunur.

8.5 Orijinden Geçen Çoklu Doğrusal Regresyon

Y = (y_1,y_2, ... y_m) gözlemleri, m tane değişken üzerine n×m gözlem matrisi boyutuyla, Y = X_p+U biçiminde modele alındığında, burada X, k bağımsız değişken üzerine verilen gözlemlerin k'nıncı sırasındaki bir n×k matrisini gösterir.

β=(β_1,β_2,....β_m) ise, regresyon parametreleri üzerine k×m boyutlu matristir. U = (u_1,u_2,…u_m) hata terimlerinin nxm boyutlu matrisidir. U'ların birbirinden bağımsız dağıldığını, her birinin sıfır vektör ortalamalı ve pozitif tanımlı mxm boyutlu kovaryans matrisi Σ ile normal dağılıma sahip olduğu varsayılmaktadır.

Bu varsayımlar ışığında verilen X, β, Σ için Y'nin ön olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$P(Y/X,\beta,\Sigma) \propto |\Sigma|^{-n/2} \exp\left[-\frac{1}{2}tr(Y-X\beta)'(Y-X\beta)\Sigma^{-1}\right]$$

$$(Y-X\beta)'(Y-X\beta) = (Y-X\hat{\beta})'(Y-X\hat{\beta}) + (\beta-\hat{\beta})'X'X(\beta-\hat{\beta})$$

$$(Y-X\beta)'(Y-X\beta) = S + (\beta-\hat{\beta})'X'X(\beta-\hat{\beta})$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

$$S = (Y-X\hat{\beta})'(Y-X\hat{\beta}) \tag{8.34}$$

Buradan β ve Σ için olabilirlik fonksiyonu çıkarımı,

$$\ell(\beta,\Sigma/Y,X) \propto |\Sigma|^{-n/2} \exp\left[-\frac{1}{2}trS\Sigma^{-1} - \frac{1}{2}tr(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\Sigma^{-1}\right] \tag{8.35}$$

olacaktır. (Zellner, 1985)

8.6 Bayesgil Tekli Doğrusal Regresyon Modeli

$$Y_i = \beta_1 + \beta_2 X_i + U_i \qquad i = 1,2......,n \tag{8.36}$$

Y_i = bağımlı değişkenin i'inci gözlemi

X_i := bağımsız değişkenin i'inci gözlemi

U_i = hata payının i'inci gözlenmemiş değeri

doğrusal klasik regresyonun varsayımlarına uygun doğrusal bir modelle çalışılmaktadır.

$Y = (y_1, y_2, ..., y_n)$ bağımlı değişkene ait gözlem değerlerine , $X = (X_1, X_2, ..., X_n)$ bağımsız değişkene ait gözlem değerleri olmasından ötürü birleşik olasılık yoğunluk fonksiyonu şu şekilde yazabilir.

$$P(Y,X/\beta_1,\beta_2,\sigma^2,\theta) = P(Y/X,\beta_1,\beta_2,\sigma^2)g(X/0) \tag{8.37}$$

Burada θ, X için marjinal olasılık yoğunluk fonksiyonunun parametrelerini göstermektedir. X_i nin değerlerinin dağılımı β_1, β_2 ve σ ile aynı olmadığından β_1, β_2 ve σ için benzerlik fonksiyonu (8.34)'nin sağ tarafında verilen $P(Y,X/\beta_1,\beta_2,\sigma^2)$ den bulunabilir. Y_i, verilen X_i, β_1, β_2, σ^2 değerlerinden bağımsız bir dağılıma sahip olacaktır.

$$E(Y_i/X_i,\beta_1,\beta_2,\sigma^2) = \beta_1 + \beta_2 X_i$$

$$Var(Y_i/X_i,\beta_1,\beta_2,\sigma^2) = \sigma^2 \qquad i = 1,2,......,n \tag{8.38}$$

Olabilirlik fonksiyonu,

$$P(Y/X,\beta_1\beta,\sigma)\propto\frac{1}{\sigma^n}\exp\left[-\frac{1}{2\sigma^2}\sum(Y_i-\beta_1-\beta_2X_i)^2\right] \tag{8.39}$$

gösterilir.

Son olasılık yoğunluk fonksiyonun meydan gelmesi için 2 tane varsayıma ihtiyaç duyulmaktadır. Bunlardan ilki yetersiz ön bilgiye sahip olunacak ikincisi ise belirsiz bir ön dağılıma sahip olması gerekmektedir.

$$P(\beta_1,\beta_2\sigma)\propto 1/\sigma \qquad -\infty<\beta_1,\beta_2<+\infty \qquad 0<\sigma<+\infty \tag{8.40}$$

β_1,β_2 ve σ ya ait olan ön olasılık yoğunluk fonksiyonu β_1, β_2, log σ'nın tekdüze ve bağımsız dağıldığı varsayımı altında,) son olasılık dağılım fonksiyonu ile birleştiğinde, β_1, β_2 ve σ için birleşik son olasılık dağılımı gösterimi şu şekildedir.

$$P(\beta_1,\beta_2,\sigma Y,X)\propto\frac{1}{\sigma^{n+1}}\exp\left[-\frac{1}{2\sigma^2}\sum(Y_i-\beta_1-\beta_2X_1)^2\right] \tag{8.41}$$

Bu ifade şu şekilde açılırsa, son olasılık yoğunluk fonksiyonunu ifadesi aşağıda gösterilmiş olduğu gibidir.

$$\sum(Y_i-\beta_1-\beta_2X_i)^2=vs^2+n(\beta_1-\hat{\beta}_1)^2+(\beta_2-\hat{\beta}_2)^2\sum X_i^2+2(\beta_1-\hat{\beta}_1)(\beta_2-\hat{\beta}_2)\sum X_i$$

$$\hat{\beta}_i=\overline{Y}-\hat{\beta}_2\overline{X} \qquad \hat{\beta}_2=\frac{\sum(X_i-\overline{X}_i)(Y_i-\overline{Y})}{\sum(X_i-\overline{X})^2} \qquad s^2=v^{-1}\sum(Y_i-\hat{\beta}_1-\hat{\beta}_2X_i)^2 \tag{8.42}$$

(8.39) nolu ifadeyi aynı zamanda şu biçimde de yazmak mümkündür.

$$\sum(Y_i-\beta_1-\beta_2X_i)^2=\sum\left\{\left(Y_i-\hat{\beta}_1-\hat{\beta}_2X_i\right)-\left[\left(\beta-\hat{\beta}_1\right)+\left(\beta_2+\hat{\beta}_2\right)X_i\right]\right\}^2 \tag{8.43}$$

Eğer (8.39)'deki ifade (8.38)'da yerine yazılırsa,

$$P(\beta_1,\beta_2,\sigma/Y,X)\propto\frac{1}{\sigma^{n+1}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[vs^2+n\left(\beta_1-\hat{\beta}_1\right)^2\left(\beta_2-\hat{\beta}_2\right)^2\sum X_i^2+2\left(\beta_1-\hat{\beta}_1\right)\left(\beta_2-\hat{\beta}_2\right)\sum X_i\right]\right\} \tag{8.44}$$

elde edilir. β_1 ve β_2 için bilinen σ ile koşullu son olasılık yoğunluk fonksiyonu, iki değişkenli normal dağılıma uymaktadır ve,

$$\sigma^2\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}^{-1} = \sigma^2 \begin{bmatrix} \dfrac{\sum X_i^2}{n\sum\left(X_i-\overline{X}\right)^2} & \dfrac{-\overline{X}}{\sum\left(X_i-\overline{X}\right)^2} \\ \dfrac{-\overline{X}}{\sum\left(X_i-\overline{X}\right)^2} & \dfrac{1}{\sum\left(X_i-\overline{X}\right)^2} \end{bmatrix} \tag{8.45}$$

biçimindeki varyans-kovaryans matrisine sahiptir.

Ancak uygulamada σ^2 çok nadiren bilindiğinden bu sonuç çok kullanışlı değildir. Parametrelere ilişkin marjinal son dağılımların kullanımı daha kolaydır. β_1 ve β_2 için marjinal olasılık yoğunluk fonksiyonunu bulmak için (8.41) nolu denklemde integral alınırsa,

$$P(\beta_1, \beta_2 / Y, X) =$$

$$\int_0^\infty P(\beta_1, \beta_2 \sigma / Y, X) d\sigma \propto \left[vs^2 + n\left(\beta_1 - \hat{\beta}_1\right)^2 \left(\beta_2 - \hat{\beta}_2\right)^2 \sum X_i^2 + 2\left(\beta_1 - \hat{\beta}_1\right)\left(\beta_2 - \hat{\beta}_2\right) \sum X_i \right]^{-n/2} \tag{8.46}$$

İki değişkenli Student-t dağılımı,

$$P(\beta_1 / Y, X) \propto \left[v + \frac{\sum\left(X_i - \overline{X}\right)^2}{s^2 \sum X_i^2 / n} (\beta_1 - \hat{\beta}_1)^2 \right]^{-(v+1)/2} \qquad -\infty < \beta_1 < +\infty \tag{8.47}$$

$$P(\beta_2 / Y, X) \propto \left[v + \frac{\sum\left(X_i - \overline{X}\right)^2}{s^2} (\beta_2 - \hat{\beta}_2)^2 \right]^{-(v+1)/2} \qquad -\infty < \beta_2 < +\infty \tag{8.48}$$

$$\left[\frac{\sum\left(X_i - \overline{X}\right)^2}{s^2 \sum X_i^2 / n} \right]^{-1/2} (\beta_1 - \hat{\beta}_1) = t_v \qquad \frac{\beta_2 - \hat{\beta}_2}{s / \left[\sum\left(X_i - \overline{X}\right)^2\right]^{1/2}} = t_v \tag{8.49}$$

dönüşümleri yapılabilir. Burada t_v, v serbestlik derecesindeki Student t olasılık yoğunluk fonksiyonuna sahip rastlantı değişkenidir. Bu sonuçlar t çizelge kullanılarak β_1 ve β_2 hakkında çıkarsamalar yapılmasını sağlar. Son (posterior) olasılık yoğunluk fonksiyonu (8.41)'dan β_1 ve β_2 parametre integrali alınarak,

$$P(\sigma / Y, X) \propto \frac{1}{\sigma^{v+1}} \exp\left(-\frac{vs^2}{2\sigma^2} \right)$$

$$0 < \sigma < \infty \tag{8.50}$$

sonucuna ulaşılır. Bu ifade, dönüşük(inverted) gama dağılımı yapısındadır. Bu dağılımın ortalaması ve varyansı sırasıyla,

$$E_{(v)} = s\sqrt{\sqrt{v/2}}\,\frac{\Gamma[(v-1)/2]}{\Gamma[v/2]} \qquad Var(\sigma) = \frac{s^2 v}{v-2} - (E_\sigma)^2 \tag{8.51}$$

biçimindedir. Varyansın posterior olasılık yoğunluk fonksiyonu aşağıdaki formdadır:

$$P(\sigma^2 / Y, X) \propto \left[(\sigma^2)^{n/2}\right]^{-1} \exp\left(-\frac{vs^2}{2\sigma^2}\right) \qquad 0 < \sigma^2 < \infty$$

8.6.1 Belirsiz Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı

Ön bilginin dağınık veya belirsiz olduğu durumda β ve log σ elemanları:

P(β, σ) ∝ 1 / σ, -∞ < βi < ∞, 0 < σ < ∞ i = 1,2:...k olarak ifade edilir.

(8.51) ve yukarıdaki ifadeden β ve σ parametreleri için bileşik posterior olasılık yoğunluk fonksiyonu,

$$P(\beta, Y, X) \propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{1}{2\sigma^2}\left[vs^2 + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta})\right]\right\} \tag{8.52}$$

olarak elde edilir. Bu dağılım, çok değişkenli normal olasılık yoğunluk fonksiyonu biçimindedir, σ^2 nadiren bilindiği için $(X'X)^{-1}\sigma^2$ genellikle değerlendirilemez. β elemanları için marjinal posterior olasılık yoğunluk fonksiyonu dağılımını oluşturmak için σ 'ya göre integral alındığında,

$$P(\beta, \sigma / Y, X) = \int_0^\infty P(\beta, \sigma / Y, X) d\sigma \propto \left[vs^2 + (\beta - \hat{\beta}) X' X (\beta - \hat{\beta})\right]^{-n/2} \tag{8.53}$$

sonucuna ulaşılır. Bu ise çok değişkenli Student t olasılık yoğunluk fonksiyonu biçimindedir ve p hakkında çıkarsamalar yapmak için kullanılır.

σ için marjinal posterior olasılık yoğunluk fonksiyonunun β'ya göre (8.52)'den integral yoluyla elde edilişi şöyledir.

$$P(\sigma / Y, X) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P(\beta, \sigma / Y, X) d\beta \propto \frac{1}{\sigma^{v+1}} \exp\left(-\frac{vs}{2\sigma^2}\right) \tag{8.54}$$

Bu dağılım dönüşük (inverted) gama olasılık yoğunluk fonksiyonu biçimindedir. (Kahyaoğlu,1996)

8.6.2 Bilgi Verici Ön Olasılık Yoğunluk Fonksiyonu ile Son Olasılık Yoğunluk Fonksiyonu Çıkarımı

Dağınık ön olasılık yoğunluk fonksiyonu altında β ve σ parametreleri için elde edilen bileşik posterior olasılık yoğunluk fonksiyonu olan

$$P(\beta,\sigma / Y, X) \propto \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{1}{2\sigma^2}\left[vs^2 + (\beta-\hat{\beta})' X' X(\beta-\hat{\beta})\right]\right\} \tag{8.55}$$

ifadesi prior olasılık yoğunluk fonksiyonu olarak kullanıldığında ve örneklem n_1 ve n_2 biçiminde iki bölümde tanımlanırsa, elde edilecek yeni ifade şu biçimdedir.

$$P(\beta, \sigma / Y_1, X_1) \propto \frac{1}{\sigma^{n+1}} \exp\left[-\frac{1}{2\sigma^2}(Y_1 - X_1\beta)'(Y_1 - X_1\beta\right] \propto \frac{1}{\sigma^{n+1}}$$
$$\exp\left[\frac{1}{2\sigma^2}\left[v_1 s_1^2 + (\beta-\hat{\beta}_1)' X_1' X_1 (\beta - \hat{\beta}_1)\right]\right] \tag{8.56}$$

Burada,

$v_1 = n_1 - k$ $\quad \hat{\beta}_1 (X_1' X_1)^{-1} X_1 Y$ ve $\quad v_1 s_1^2 = (Y_1 - X_1\hat{\beta}_1)'(Y_1 - X_1\hat{\beta}_1)$

olarak ele alınmıştır.

Bu ifade, ön olasılık yoğunluk fonksiyonu kabul edildiğinde birinci örnek için β normal dağılıyor ve bunun yanı sıra σ biliniyorsa, ortalaması $\hat{\beta}_1$ kovaryans matrisi $(X_1'X_1)^{-1} . \sigma$ olacaktır, σ için marjinal bir olasılık yoğunluk fonksiyonunu ise, v_1 ve s_1^2 parametreleriyle dönüşük gama formundadır. (8.56)'ün ikinci satırından hareketle,

$$P(\beta / \sigma, \hat{\beta}_1, s_1^2) \propto \frac{1}{\sigma^k} \exp\left[-\frac{1}{2\sigma^2}(\beta - \hat{\beta}_1)' X_1' X_1 (\beta - \hat{\beta}_1)\right] \tag{8.57}$$

yazılabilir.

İkinci örnek için olabilirlik fonksiyonu ise,

$$\ell(\beta / \sigma, Y_2, X_2) \propto \frac{1}{\sigma^{n_2}} \exp\left[-\frac{1}{2\sigma^2}(Y_2 - X_2\beta)'(Y_2 - X_2\beta)\right] \tag{8.58}$$

olarak elde edilir.

Bu örnek için β ve σ'nın, birinci örnekteki değerlerle aynı parametre olduğu varsayımı altında (8.56)'deki prior olasılık yoğunluk fonksiyonu (8.58)'daki olabilirlik fonksiyonu ile birleştirilirse, posterior olasılık yoğunluk fonksiyonu olan,

$$P(\beta,\sigma / Y_1, X_1, X_2) \propto \frac{1}{\sigma^{n_1+n_2+1}} \exp\left[-\frac{1}{2\sigma^2}(Y_1 - X_1\beta)'(Y_1 - X_1\beta) + (Y_2 - X_2\beta)'(Y_2 - X_2\beta)\right] \tag{8.59}$$

elde edilir.

Böylece ilk örneğin prior olasılık yoğunluk fonksiyonu ve ikinci örneğin olabilirlik fonksiyonu yardımıyla posterior olasılık yoğunluk fonksiyonuna ulaşılmıştır. Bu (8.52)'de verilen ifade ile aynı dağılıma sahiptir.(Zellner ,1985)

8.7 Bayesgil Çoklu Doğrusal Regresyon

Y = (y_1,y_2, ... y_m) gözlemleri, m tane değişken üzerine n×m gözlem matrisi boyutuyla, Y = X_p+U biçiminde modele alındığında, burada X, k bağımsız değişken üzerine verilen gözlemlerin k'nıncı sırasındaki bir n×k matrisini gösterir.

β=(β_1,β_2,....β_m) ise, regresyon parametreleri üzerine k×m boyutlu matristir. U = (u_1,u_2,...u_m) hata terimlerinin nxm boyutlu matrisidir. U'ların birbirinden bağımsız dağıldığını, her birinin sıfır vektör ortalamalı ve pozitif tanımlı mxm boyutlu kovaryans matrisi Σ ile normal dağılıma sahip olduğu varsayılmaktadır.

Bu varsayımlar ışığında verilen X, β, Σ için Y'nin ön olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$P(Y / X,\beta,\Sigma) \propto |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} tr(Y - X\beta)'(Y - X\beta)\Sigma^{-1}\right]$$

$$(Y - X\beta)'(Y - X\beta) = (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\beta - \hat{\beta})' X' X(\beta - \hat{\beta})$$

$$(Y - X\beta)'(Y - X\beta) = S + (\beta - \hat{\beta})' X' X(\beta - \hat{\beta})$$

$$\hat{\beta} = (X' X)^{-1} X' Y$$

$$S = (Y - X\hat{\beta})'(Y - X\hat{\beta}) \tag{8.60}$$

Buradan β ve Σ için olabilirlik fonksiyonu çıkarımı,

$$\ell(\beta,\Sigma / Y, X) \propto |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} trS\Sigma^{-1} - \frac{1}{2} tr(\beta - \hat{\beta})' X' X(\beta - \hat{\beta})\Sigma^{-1}\right] \tag{8.61}$$

olacaktır. (Zellner, 1985)

8.7.1 Belirsiz Ön Bilgi Durumunda Çoklu Doğrusal Bayesgili Regresyon

Σ'nın, m(m+1)/2 ayırıcı elemanları ve p'nın parametre değerleri hakkında ön olarak az bir bilgi veya belirsizlik söz konusu olduğunda, β ve Σ'nın elemanlarının bağımsız dağıldığı varsayımı altında prior olasılık yoğunluk fonksiyonu çıkarımı,

$$P(\beta,\Sigma) = P(\beta).P(\Sigma) \tag{8.62}$$

şeklindedir.

Eğer, (8.30)'a Jeffreys'in Değişmezlik Teoremi uyarınca,

$$P(\beta) = \text{sabit} \tag{8.63}$$

$$P(\Sigma) \propto |\Sigma|^{-(m+1)/2} \tag{8.64}$$

yazılabilir. Buradan hareketle,

$$P(\Sigma^{-1}) \propto |\Sigma|^{-(m+1)/2} \tag{8.65}$$

ifadesini yazmak mümkündür.

Bayes kuralına göre posterior dağılım aşağıdaki gibidir:

$$p(B_0,\ldots,B_k, \sigma^2 | Y, X) \propto p(B_0,\ldots,B_k,\sigma^2) \times \prod_i p(y_i | \mu_i, \sigma^2) \tag{8.66}$$

Regresyon katsayılarıyla ilgili bir çıkarım yapmak için B, σ^2için bir önsel dağılımın seçilmesi gerektiği açıktır. Standart bilgi vermeyen prior (B, log σ^2) üzerinde sabittir. Bu da

$p(B, \log \sigma^2) \propto \sigma^{-2}$ 'a eşittir.

Birçok veri noktası ve çok az parametre varsa, bu prior istatistiksel modeller için iyi bir seçimdir. Çünkü olabilirlik fonksiyonu örnekten büyük ölçüde etkilenecektir. Σ matrisine küçük değerler vererek prior bilgiye daha çok ağırlık veriliyorsa, varyans tahminleri daha büyük olacaktır. Standart EKK tahminleri prior değerlerden farklılık gösteriyorsa, özellikle det(X^TX) büyükse, regresyon katsayıları için varyans tahminleri daha büyük olacaktır. Küçük örneklerle ve birçok parametreyle önsel dağılımlar veya hiyerarşik modeller, analiz için daha önemli hale gelmektedir. *

---

* http://209.85.135.104/search?q=cache:FfxMo6M_XpEJ:home.uchicago.edu.

$Y|B, \sigma^2, X \sim N(XB, \sigma^2)$ ve $p(B, \log \sigma^2) \propto \sigma^2$ ise, sonsal koşullu dağılım $p(B|\sigma^2;\text{veri})$ şu şekilde yazılabilir:

$B' = (X^TX)^{-1}X^TY$ olduğunda, $p(B|\sigma^2;\text{veri}) \sim N$ çokludeğişken $(B', \sigma^2 (X^TX)^{-1})$ (8.67)

σ^2'nin sonsal dağılımı şu şekilde yazılabilir:

$s^2 = (Y-XB)^T(Y-XB)/(n-k)$ olduğunda, $p(\sigma^2|\text{veri}) \sim$ Scaled Inv- $\chi 2(n-k,s2)$ (8.68)

B'nin marjinal prior ile ilgili çıkarımda bulunmak için:

1) σ^2'nin posteriordan tekrarlanan örnekler alınır. Bu durumda bilinmeyen ortalamalı normal modelde görüldüğü gibi $B|\sigma^2$olur.

2) σ^2'yi koşullu posteriordan B için ayırılır ve B'nin marjinal dağılımının çok değişkenli bir t dağılımı olduğu bulunur:

$p(B \mid \text{veri}) \sim$ çoklu değişken $t_{n-k}(B', s^2(X^TX)^{-1})$ (8.69)

Bulunan sonuçlar klasik regresyon yöntemi ile bulunan sonuçlarla karşılaştırıldığında standart hataların yorumu farklı olacaktır.

(8.65)'de verilen bu belirsiz olasılık yoğunluk fonksiyonu Savage tarafından önerilmiştir.

(8.63), (8.64), (8.65)'deki belirsiz prior olasılık yoğunluk fonksiyonları (8.61)'daki olabilirlik fonksiyonu ile birleştirildiğinde parametreler için bileşik posterior olasılık yoğunluk fonksiyonu,

$$P(\beta,\Sigma/Y,X) \propto |\Sigma|^{-1/2(n+m+1)} \exp\left[-\frac{1}{2}trS\,\Sigma^{-1} - \frac{1}{2}tr\left[S + (\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\right]\dot{\Sigma}^{-1}\right] \tag{8.70}$$

yukarıdaki gösterime dayanarak aşağıdaki denklemlere ulaşabilir.

$$P(\beta,\Sigma/Y,X) = P(\beta/\Sigma,Y,X).P(\Sigma/Y,X)$$

$$P(\beta/\Sigma,Y,X) \propto |\Sigma|^{-k/2} \exp\left[-\frac{1}{2}tr\left[(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\Sigma^{-1}\right]\right]$$

$$P(\beta/\Sigma,Y,X) \propto |\Sigma|^{-k/2} \exp\left[-\frac{1}{2}tr\left[(\beta-\hat{\beta})'X'X \otimes X'X(\beta-\hat{\beta})\right]\right] \tag{8.71}$$

Katsayılar Σ^{-1} koşullu prior kesinlikle X^TX verisiyle bulunur. Buna göre B için önsel ağırlığı arttırdıkça (prior varyans azaltılarak), veriyle ilgili daha önceki bilgilere daha büyük katsayılar (ya da önem) verilmektedir. *

Burada $\beta' = (\beta_1', \beta_2'........, \beta_m')$ ve $\hat{\beta}' = (\beta_1', \beta_2'........, \beta_m')$ biçimindedir. Verilen Σ ile β için koşullu posterior olasılık yoğunluk fonksiyonu $\hat{\beta}$ ortalama ve $\Sigma \otimes (X'X)^{-1}$ kovaryans matrisi ile çok değişkenli normal dağılıma uymaktadır.

Parametrelerin koşullu posterior olasılık yoğunluk fonksiyonları aşağıdaki gibi gösterilebilir.

$$P(\beta_1 / \Sigma, Y, X) \propto \frac{1}{\sigma_{11}^{k/2}} \exp\left[-\frac{1}{2\sigma_{11}} (\beta_1 - \hat{\beta}_1)' X' X (\beta_1 - \hat{\beta}_1) \right] \tag{8.72}$$

$\hat{\beta}_1$ ortalama ve $(X'X)^{-1}\,\sigma_{11}$ varyans kovaryans matrisli çok değişkenli normal dağılımı Σ için marjinal posterior olasılık yoğunluk fonksiyonu,

$$P(\Sigma / Y, X) \propto |\Sigma|^{-v/2} \exp\left(-\frac{1}{2} tr\, \Sigma^{-1} S \right) \tag{8.73}$$

Burada v = n-k+m+1'dir.

v açıklayıcı değişken sayısıdır.

Aynı şekilde, a parametresi için marjinal posterior dağılımdan hareket edilirse,

$$P(\sigma_{11} / Y, X) \propto \frac{1}{\sigma_{11}^{[v-2((m-1)]/2}} \exp\left[-\frac{S_{11}}{2\sigma_{11}} \right] \tag{8.74}$$

dönüşük (inverted) gama olasılık yoğunluk fonksiyonu.

β_1 marjinal posterior olasılık yoğunluk fonksiyonu

$$P(\beta_1 / \Sigma, Y, X) \propto \frac{1}{\sigma_{11}^{k/2}} \exp\left[-\frac{1}{2\sigma_{11}} (\beta_1 - \hat{\beta}_1)' X' X (\beta_1 - \hat{\beta}_1) \right] \tag{8.75}$$

$$P(\sigma_{11} / Y, X) \propto \frac{1}{\sigma^{[v-2(m-1)]/2}} \exp\left(-\frac{S_{11}}{2\sigma_{11}} \right) \tag{8.76}$$

---

* Not: Zellner (1971) ve önsel varyans Σ'yi aşağıdaki gibi alır: iki ver dizininizin olduğunu varsayılmaktadır - Y_1, X_1 ve Y_2, X_2. Bileşik önsel $1/\sigma^2$ 'li X_1, Y_1'in regresyon analizi için $B_{önsel}$'i sonsal ortalamaya ve Σ'yi $X_1^T X_1$'e eşit sayar.

$$P(\beta_1 / Y_1 X) \propto \int P(r_{11} / Y_1 X) P(\beta_1 / \Sigma, Y_1 X) d\sigma_{11}$$

$$P(\beta_1 / Y_1 X) \propto \left[S_{11} + (\beta_1 + \hat{\beta}_1)' X' X (\beta_1 + \hat{\beta}_1) \right]^{-[n-(m-1)]/2} \tag{8.77}$$

8.7.2 Bilgi Verici Prior Dağılım Fonksiyonu ile Çoklu Regresyon Modeli

Burada en önemli sorun ön bilgiyi göstermede kullanılacak olasılık yoğunluk fonksiyonunun bütün durumlar için uygun olup olmadığıdır.

Eğer çok değişkenli regresyon modeli için basit bir bileşik prior dağılım kullanılırsa,varyans-kovaryans matrisi bu dağılımdan etkilenecek dolayısıyla parametre varyanslarına yansıyacaktır. Bu problemden kaçınmanın çözümü olarak modelin tüm katsayıları için genel çok değişkenli normal olasılık yoğunluk fonksiyonu kullanımı önerilmektedir. Fakat buradaki sorun ise, böyle genel çok değişkenli normal olasılık yoğunluk fonksiyonunun, basit bileşik olasılık yoğunluk fonksiyonu kullanımında olduğu gibi bir işlemsel uygunluk sağlayamayışıdır. Bu matematiksel zorluk bir çok durumda yine de göz ardı edilebilir. Çünkü basit bir olasılık yoğunluk fonksiyonu kullanılmasını gerektiren durumlarla da karşılaşılmaktadır. Bu bilgiler ışığında ön olasılık yoğunluk fonksiyonu şu şekilde yazılabilir.

$$P(\beta / \Sigma^{-1}) \propto \left|\Sigma^{-1}\right|^{-(m+1/2} \exp\left[-\frac{1}{2} (\beta - \overline{\beta})' C^{-1} X (\beta - \overline{\beta}) \right] \tag{8.78}$$

$\hat{\beta}$, kx1 boyutunda bir vektördür. Aynı zamanda prior olasılık yoğunluk fonksiyonunun ortalamasıdır ve araştırıcı tarafından belirlenir. C ise kxk boyutunda prior varyans-kovaryans matrisidir. E elemanları hakkında prior bir bilgi olduğu varsayılmaktadır ve daha önce kullanılan belirsiz prior olasılık yoğunluk fonksiyonu ele alınırsa, (8.61)'da verilen olabilirlik fonksiyonuyla birleştirildiğinde bileşik posterior olasılık yoğunluk, fonksiyonu,

$$P(\beta, \Sigma^{-1} / Y) \propto \left|\Sigma^{-1}\right|^{(n-m-1)/2} \exp\left\{ -\frac{1}{2} tr \Sigma^{-1} \left[S + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \right\}$$

$$P(\beta, \Sigma^{-1} / Y) \propto \left|\Sigma^{-1}\right|^{(n-m-1)/2} \exp\left[-\frac{1}{2} (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \right] \tag{8.79}$$

olacaktır. Burada, $\beta=(\beta_1,\beta_2,....,\beta_m)$, olarak $\beta' = (\beta_1', \beta_2', ...\beta_m')$ ve $\hat{\beta} = (X'X)^{-1}X'Y$ olarak, belirlenmiştir. (Kahyaoğlu,F.,1996)

Bu ifade Σ^{-1} için integre edildiğinde,

$$P(\beta,/Y) \propto \left|S+(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\right|^{-n/2} \exp\left[-\frac{1}{2}(\beta-\overline{\beta})'C^{-1}(\beta-\overline{\beta})\right] \tag{8.80}$$

β için posterior olasılık yoğunluk fonksiyonu elde edilir ve bu ifade, genelleştirilmiş çoklu Student t dağılımına uyar (Zellner, 1985, s.239).

BÖLÜM 9: UYGULAMA

4C : Kesim, Karat, Berraklık, Renk

Tüm dünyada uzmanlar pırlantaları 4C nitelikleri açısından değerlendirir: Kesim, karat ağırlığı, renk ve berraklık. Bu sınıflandırma, perakendecilerin ve tüketicilerin taşları kıyaslamasına ve değer biçmesine olanak tanır.

4C Kuralları Nasıl İşler? En Önemli Kural Hangisidir? Göz önüne alınması gereken en önemli husus, hiçbir kuralın bir diğerinden daha önemli olmadığıdır. Kendi bütçesine, satın alma nedenine, takının tasarımı ya da mücevherin çeşidine uyum sağlamak üzere 4C kurallarının herhangi bir karışımını seçlebilir. Aşağıda 4C kurallarının her biri ayrıca anlatılacak; böylece siz de değişik unsurların nasıl bir araya geldiğini ve bir elmasın değerinin nasıl belirlendiğini belirlenir.

Her ne kadar pırlanta, özünde bulunan 4C kurallarının birleşimine bağlı olarak tanımlansa da; taşın kendine özgü güzelliği onun bir özellikler listesi içermesinden daha da öte bir anlam taşır.

9.1 KESİM

Bir elmasın özelliklerinden birçoğunu doğa belirler, ancak taşın gerçek pırıltısını, ateşini ve nihai güzelliğini ortaya çıkartmak deneyimli bir kesim ustasının becerisine bağlıdır.

Boyutu ve şekli ne olursa olsun, iyi kesimli bir pırlanta ışığı kendi içinden geçirerek yansıtır. Işık, ayna gibi fasetalara çarpar ve taşın tepesinden dışarıya dağılarak ateşli ışıltısını ortaya çıkarır. Bir elmas çok derin ya da çok yassı bir şekilde kesilirse, o zaman yanlardan veya dipten ışık kaybedilir. Bu yüzden pırlantanın parlaklığı ve bu nedenle de değeri azalır.

Bir pırlantanın kesimi görsel boyutunu da etkileyebilir. Aynı karat ağırlığındaki iki pırlanta, kesimlerinin derinliği ya da şekline bağlı olarak değişik boyutlardaymış gibi gözükebilir.

Kesim sözcüğü ile aslında pırlantanın şekli de ifade edilir. Yuvarlak, kare, damla ya da kalp şekilleri en çok bilinenler arasındadır ama ilerleyen teknoloji sayesinde çok daha farklı şekiller ve kesimler de ortaya çıkmaktadır.

9.2 KARAT

Karat sözcüğü bazen yanlışlıkla pırlantanın boyutlarını ifade etmek için kullanılır ancak karat aslında bir ağırlık birimidir. Bir karat(200 miligram eşdeğeri) 100 'puana' bölünebilir.

0.75 karatlık bir pırlanta aynı zamanda '75 puan' ya da 3 / 4 karatlık bir pırlanta olarak tanımlanabilir.

Daha büyük elmaslar doğada daha az bulunur, bu yüzden daha pahalıdırlar. Örneğin, bir karatlık tek bir pırlanta, aynı renk, berraklık ve kesimdeki yarım karat ağırlığındaki iki pırlantadan daha maliyetli olacaktır.

Bu arada pırlantanıza sadece karat ağırlığı ile değer biçilmez. Kesim, berraklık ve renklerine bağlı olarak, eşit karat ağırlığı olan iki pırlantanın değerleri çok değişik olabilir. Bunun ötesinde, daha küçük ama mükemmel kesimli, daha beyaz renkli ve lekesiz bir pırlanta, daha az beyaz ve daha fazla lekeli olan daha büyük bir pırlantadan daha değerli olabilir.

9.3 RENK

Elmas, gökkuşağının neredeyse her renginde görülür ama en popüler olanlar hala daha beyaz renkli olanlarıdır. Elmasta, aralarında çok ince farklılıklar olan 20'den fazla renk bulunmaktadır. Bunlar alfabetik sırayla D-Z arasında dizilir.

Aradaki renk farklılıkları o kadar belirsizdir ki, renk sınıflandırması bir uzman tarafından kontrollü ışıklar altında yapılmalı ve doğruluk için standart bir renk seti ile karşılaştırmalıdır. Renk skalasının üst ucunda olanlar daha nadide olduklarından daha pahalıdır, ancak bunların rengini çıplak gözle ayırt etmek zordur.

Diğer 4C kurallarında olduğu gibi bir pırlantanın değeri rengine göre farklılık gösterebilir. Aynı berraklıkta ve karattaki iki pırlanta, renklerine bağlı olarak farklı fiyatlara sahip olabilir.

D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Renksiz			Nadir Beyaz				Beyaz			Renkli Beyaz					Renkli							

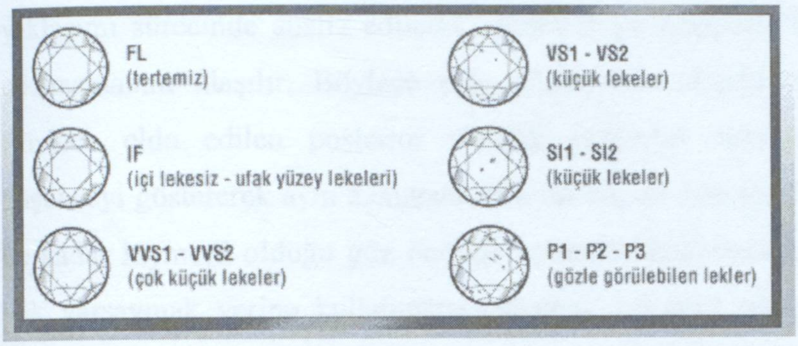
Şekil 9.5 Pırlanta Çeşitleri

Doğada daha koyu renklerde mavi, yeşil, sarı, turuncu, pembe ve en nadiri olan kırmızı renkli elmaslar vardır. Bu elmaslara "renkli fantezi" adı verilir ve çok az olan bu taşlara yüksek değerler biçilmektedir.

9.4 BERRAKLIK

Her elmas tektir. Doğa, her pırlantanın onu takan kişi kadar eşsiz olmasını sağlamıştır. Pırlantanın içinde var olan ve 'leke' olarak adlandırılan doğal özellikler, taşın içerisinde özgün bir parmak izi yaratır. Genellikle çıplak gözle görülemeyen bu küçük izler, elmas yerkabuğunun altında oluşurken içinde hapsolan mineraller ya da meydana gelen çatlaklardır.

Bu uyarin sayısı, türü, rengi, boyutu ve konumu da pırlantanın değerini etkileyebilir ve çoğu görüntüyü 10 kez büyüten bir mercek ile ancak uzmanlar tarafından görülebilir. Hatta bu mercekle bile en küçük kusurları bulmak çok zor olabilir.



Şekil 9. 6 Berraklık Türleri

İçleri kusursuz elmaslar daha az bulunur, bu yüzden de daha pahalıdır; ancak küçük lekeler pırlantanın güzelliğini ya da parlaklığını etkilemez. Ağırlık ve rengi eşit olan pırlantaların fiyatı, lekelerine bağlı olarak büyük ölçüde değişebilir.

9.5 Uygulama:

Uygulamaya öncelikle deneysel verilerimizin analizi ile başlanacak. Deneysel verileri MINITAB programını kullanarak analiz edildi. Çoklu regresyon modeli şeklinde yapmış olunan çalışma sonucunda elde etmiş olduğum değerleri istatiksel olarak analizini elde ettim. Bu sonuçları teorik olarak göstermiş olmak üzere diğer subjektif analize katılır. Çıkan değerler yorumlanır. Kişisel deneyimleri ve daha önceki benzer araştırmalarda elde edilen sonuçlar analizin yapıldığı alanların kendilerine olan kısıtlamaları ve örneklem bilgisi,bayes yaklaşımı sürecinde analiz edilerek parametreler hakkında bütün bilgilerin kullanılabileceği çıkarsamalara ulaşılır. Böylece elde edilen son olasılık yoğunluk fonksiyonuna ulaşılır. Böylece elde edilen posterior olasılık yoğunluk fonksiyonu parametreleri hakkındaki düşünceyi göstererek aynı zamanda hem ön bilgiyi hem de örnek bilgisini birleştirir. Bilginin ne kadar kıymetli olduğu göz önünde bulundurulduğunda kişilerden elde edilen ön bilgileri yok varsaymak yerine kullanmaya çalışmak bilimsel yaklaşımın gereği olacaktır. Normal dağılıma sahip örneklem bilgisi ön bilgi ile birleştirildiğinde posterior olasılık yoğunluk fonksiyonu da normal dağılıma uyacaktır. Çalışmada normal dağılıma uyan örneklem bilgisi kullanılarak pırlanta fiyatını etkileyen parametrelerin analiz yapılmıştır.

9.6 Çalışma

Çalışmada 67 adet taş üzerinden yapılmıştır.Bir taşa sertifika verilmesi için o taş ayrı odalarda bulunan 3 kişiye sırasıyla verilir ve taşın özellikleri hakkında bilgi istenir. Bu şahıslar konularında uzman kişilerden seçilir.aynı sürede taşa bakılmaları istenir. Uzmanlardan alınan bilgiler aynı olduğu takdirde taşa sertifika verilir. Aksi takdirde tekrar süreç tekrarlanır ama bu süreç esnasında 3 şahıs birbiri ile görüştürülmez. Çünkü genel olarak taşın özellikleri subjektiflik gösterir. Herkes kendine göre yorum yapabilir. Tabi taşların özellikleri son zamanlarda yaşanan teknolojik gelişmeler ışığı altında bilgisayar ya da diğer teknolojik aletlerde kullanılmaya başlandı. Ağırlık hesaplanırken elektronik terazi renge bakarken colormetre, berraklığa bakılırken de elektronik mikroskop kullanılarak bakılır. Buradan anlaşılacağı gibi taşın kalitesi hem deneysel veri aynı zamanda subjektif verilere dayalı olarak da ortaya konabilir. Tezde ilgilenmiş olduğum Bayes teorisiyle alt yapı bakımıyla birbiri ile örtüşmektedir. Pırlantanın fiyatını berraklığı mı, yoksa renginin mi daha çok etkilediğini ilk klasik istatistiksel metotlarla araştırıp daha sonra kişisel verilere dayalı olmak üzere yapmış olunan analizleri de ortaya koyup Bayesgil yöntemlerle analiz edilecek

$$B^{**}=[\textstyle\sum_p{}^{*} + X'X/\sigma^2]^{-1}[\textstyle\sum_p{}^{-1}{}^{*}\, B^{*}+(X'X/\sigma^2)b]$$

Ve kovaryans matrisi de şu şekilde bulunur.

$$\textstyle\sum_p{}^{**}=[\textstyle\sum^{-1}{}_p{}^{*}+ X'X/\sigma^2]^{-1}$$

$\sum_p{}^{*}$ **Anket Verisi**

0,08	0,051	0,116	-0,12
0,051	0,04	-0,0007	-0,165
0,116	-0,0007	1,218	-0,24
-0,12	-0,165	-0,24	1,61

Çizelge9.1 Ön Bilgi Var.Cov Matrisi

$\sum^{-1*}{}_p$ **Anket Verisi**

-91,6212	165,7851	11,15091	11,82374
165,7851	-255,634	-19,2282	-16,70813
11,15091	-19,2282	-0,49097	-1,212658
11,82374	-16,7081	-1,21266	-0,390703

Çizelge 9.2 Ön Bilgi Var.Cov Matrisinin Tersi

X'X/ σ^2 **Anket Verisi**

469986,2	44067,41	137203,4	175785,5
44067,41	4377,2	12704	15362
137203,4	12704	47700	49700
175785,5	15362	49700	77100

Çizelge 9.3

ÖN BİLGİ ORTALAMA DEĞERİ

Anket Verisi

6,7	1,43	0,11	0,08

Çizelge 9.4

- **Ekk Yöntemi İle Elde Edilen Parametre Tahminleri (Deneysel Veri)**

6,92	1,29	0,1	0,05

Çizelge 9.5

İlk önce **Çizelge 1** ve **Çizelge 3** toplanır ve tersi alınır. **Çizelge 9.6** elde edilir.

0,000206	-0,001204	-9,85161E-05	-0,000166
-0,001204	0,008268	0,000357019	0,000867
-9,85E-05	0,000357	0,000150266	5,66E-05
-0,000166	0,000867	5,66103E-05	0,000182

Çizelge 9.6

Çizelge 2 Çizelge 4 ile çarpılır.**Çizelge 3 Çizelge 5** ile çarpılır.çıkan sonuçlar toplanır.ve çıkan sonuç **Çizelge 6** ile çarpılarak son bilgi ortalama değerine ulaşılır.

- **Son bilgi Ortalama değeri**

5,61831	8,738979	0,453432	0,86933

Çizelge 9.7

Son bilgi varyans-kovaryans matrisine ulaşmak için **Çizelge 2** ile **Çizelge 3** toplanır.ve tersi alınarak istenilen varyans-kovaryans matrisi elde edilir:

9,07945E-05	0,000795	-3,21E-06	5,09163E-05
0,00079472	0,0052	-0,000285	-0,00059337
-3,20737E-06	-0,000285	0,00012	-1,30164E-05
5,09163E-05	-0,000593	-1,3E-05	2,33644E-05

Çizelge 9.8

- **Güven aralıkları karşılaştırması:**

Bayesgil regresyonla bulunan güven aralıkları

8,73+- 1,96*0,005=**8,72 ve 8,74**
0,453+-1,96*0,00012=**0,453005 ve 0,452995**
0,89633+-1,96*0,0002=**0,89638 ve 0,89628**

Klasik regresyonla bulunan güven aralıkları

1,29+-1,96*0,093=**(1,44 ve -1,14)**
0,1+-1,96*0,016=**(0,12 ve -0,08)**
0,051+-1,96*0,016=**(0,053 ve -0,049)**

- **DENEYSEL VERİ SETİ VE MINITAB UYGULAMASI**

Fiyat	Carat	Renk	Berraklık
7,97247	0,28	3	6
7,90101	0,29	4	5
7,97247	0,34	4	4
7,90101	0,37	4	5
7,64969	0,48	3	3
8,00637	0,52	2	4
8,00637	0,56	1	6
8,10168	0,56	3	3
8,18869	0,57	2	5
8,00637	0,54	2	5
7,97247	0,58	1	3
8,10168	0,62	3	4
8,24276	0,58	2	4
8,24276	0,57	2	3
8,00637	0,59	2	4
8,03916	0,63	3	3
8,10168	0,65	2	2
8,29405	0,68	3	4
8,21609	0,72	2	2
8,41183	0,73	2	5
8,45532	0,74	3	4
8,18869	0,71	1	3
8,18869	0,70	1	3
8,45532	0,70	3	3
8,38936	0,72	3	4
8,29405	0,68	2	4
8,29405	0,70	2	4
8,41183	0,73	3	3
8,41183	0,73	2	4
8,61250	0,73	4	5

8,29405	0,73	3	2
8,34284	0,73	2	4
8,45532	0,69	3	5
8,24276	0,71	1	4
8,13153	0,75	3	1
8,18869	0,75	2	3
8,29405	0,75	3	3
7,90101	0,80	1	5
8,13153	0,90	2	1
8,43381	0,90	3	1
8,69951	0,90	4	4
8,64822	0,90	2	3
8,16052	0,90	2	4
8,68271	0,90	3	3
8,61250	0,90	4	2
8,51719	1,02	1	2
8,51719	1,01	2	1
8,94898	1,03	3	3
8,57546	1,03	2	3
8,68271	1,03	3	1
8,64822	1,03	3	3
8,57546	1,03	3	1
8,94898	1,03	5	3
9,03599	1,05	5	3
8,88184	1,05	2	3
8,57546	1,05	3	1
8,73230	1,05	3	3
8,64822	1,05	3	2
8,64822	1,05	2	3
8,64822	1,05	2	2
8,43381	1,05	1	3
8,43381	1,05	1	3
8,79482	1,05	5	3
8,53700	1,05	3	2
8,68271	1,06	3	2
8,64822	1,08	2	2
8,43381	1,10	1	3

Çizelge9.9 Deneysel Veri

Regression Analysis: C1 versus C2; C3; C4

```
The regression equation is
C1 = 6,92 + 1,30 C2 + 0,100 C3 + 0,0518 C4

Predictor        Coef    SE Coef       T       P    VIF
Constant       6,9200     0,1245   55,58   0,000
C2            1,29793    0,09326   13,92   0,000    1,6
C3            0,10011    0,01611    6,22   0,000    1,0
C4            0,05180    0,01656    3,13   0,003    1,6

S = 0,1313         R-Sq = 81,3%          R-Sq(adj) = 80,4%
PRESS = 1,24137    R-Sq(pred) = 78,64%

Analysis of Variance

Source            DF       SS       MS       F       P
Regression         3   4,7254   1,5751   91,38   0,000
Residual Error    63   1,0860   0,0172
  Lack of Fit     56   0,9701   0,0173    1,05   0,528
  Pure Error       7   0,1159   0,0166
Total             66   5,8114

53 rows with no replicates

Durbin-Watson statistic = 1,62

Predicted Values for New Observations

New Obs     Fit   SE Fit        95,0% CI              95,0%
PI
1        9,8196   0,1029   (  9,6140; 10,0253)   (  9,4863;
10,1530) XX
X denotes a row with X values away from the center
XX denotes a row with very extreme X values

Values of Predictors for New Observations

New Obs     C2       C3       C4
1         2,00     2,00     2,00
```

MODELİN YORUMLANMASI

Yukarıda görülen deneysel veri üstünden yapılan çalışmanın normal dağılım varsayımlarına uyup uymadığı konusunda analiz yapmak istersek ilk önce otokorelayon testi yapmak gerekir. Otokorelasyon testinde Durbin Watson test istatistiği 1,67 bulunmuştur ve bunu çizelge değerleri ile karşılaştırdığımızda otokorelasyonun olmadığına karar verilir.

N=67 için k=2 d_L =1,41 d_U =1,47 otokorelasyon olmayan bölgeye denk gelmektedir.

Şimdi de değişen varyans olup olmadığına karar verilir. Bunun için white testine karar verilmiştir.

0,11*67=7,9 < 24,882 çizelge değeri hesaplanan değer çizelge değerinden küçük olduğundan dolayı sabit varyans varsayımı geçerlidir. Ayrıca çoklu doğrusal bağlılık olup olmadığına karar vermek için VİF değerlerine bakılması gereklidir.

1/1-0,11=1.1 Bu değer 5 ten küçük olduğundan dolayı çoklu doğrusal bağlılık yoktur ifadesini kabul ederiz.

Normal dağılıp dağılmadığını Jarque-Bera testi ile kontrol ederiz.

$JB=(N/6)\ B_1\ +(N/24)(\ B_2\ -3)^2$

Ki kare çizelge göre 2 serbestlik derecesine göre karşılık gelen değer esas alınır.Hesaplanan kritik değeri aşarsa hipotez red edilir.

Ele alınan veri setinde Jarque -Bera test istatistiği 1,631 olarak elde edilmiştir.0,05 anlamlılık düzeyinde 2 serbestlik dereceli ki kare değeri ise 9,882 dir JB<24,882 olduğu için H_0 kabul edilir.Normal dağılım söz konusudur.

Şimdi alınan bu 67 veri üzerinden yapılacak olan bayes analizinin son dağılımı da normal olacaktır. Bu veri setini tecrübe sahibi olan insanların düşüncelerine ve daha önceki deneyimlerine dayanılarak yapılmıştır.Taşlar kendilerine hiçbir etki altında kalmamaları için kendilerine ait odalar gösterilip taşların özelliklerini yazmaları istenmiştir.

ANKET VERİSİ VE BİLGİSAYAR UYGULAMASI

.Fiyat	Carat	Renk	Berraklık
2900	0,3	4	5
2700	0,31	4	4
2900	0,33	4	5
2700	0,36	3	5
2100	0,49	2	2
3000	0,5	1	5
3000	0,52	1	5
3300	0,52	2	4
3600	0,53	2	5
3000	0,55	1	5
2900	0,55	2	3
3300	0,55	2	4
3800	0,56	3	4
3800	0,56	3	4
3000	0,58	1	5
3100	0,6	4	2
3300	0,61	3	3
4000	0,7	2	3
3700	0,7	2	2

4500	0,7	2	5
4700	0,7	3	4
3600	0,7	1	3
3600	0,7	1	3
4700	0,7	4	3
4400	0,7	3	3
4000	0,7	2	3
4000	0,71	1	5
4500	0,71	2	5
4500	0,71	2	5
5500	0,71	4	5
4000	0,71	3	2
4200	0,71	2	4
4700	0,72	3	4
3800	0,72	1	4
3400	0,74	2	1
3600	0,75	1	3
4000	0,75	3	2
2700	0,81	1	4
3400	0,89	2	1
4600	0,9	3	1
6000	0,9	4	3
5700	0,9	3	3
3500	0,91	1	4
5900	0,91	3	4
5500	0,92	4	2
5000	1	1	2
5000	1	2	1
7700	1	4	3
5300	1	1	3
5900	1	3	2
5700	1	2	2
5300	1,01	4	1
7700	1,01	4	3
8400	1,01	4	4
7200	1,01	3	3
5300	1,01	4	1
6200	1,01	2	3
5700	1,01	2	2
5700	1,01	2	2
5700	1,01	2	2
4600	1,02	1	3
4600	1,02	1	3
6600	1,03	5	2
5100	1,04	3	1
5900	1,05	3	2
5700	1,05	2	2
4600	1,14	1	3

Çizelge9.10 Anket Verisi

Fiyat	Carat	Renk	Berraklık
7,97247	0,30	4	5
7,90101	0,31	4	4
7,97247	0,33	4	5
7,90101	0,36	3	5
7,64969	0,49	2	2
8,00637	0,50	1	5
8,00637	0,52	1	5
8,10168	0,52	2	4
8,18869	0,53	2	5
8,00637	0,55	1	5
7,97247	0,55	2	3
8,10168	0,55	2	4
8,24276	0,56	3	4
8,24276	0,56	3	4
8,00637	0,58	1	5
8,03916	0,60	4	2
8,10168	0,61	3	3
8,29405	0,70	2	3
8,21609	0,70	2	2
8,41183	0,70	2	5
8,45532	0,70	3	4
8,18869	0,70	1	3
8,18869	0,70	1	3
8,45532	0,70	4	3
8,38936	0,70	3	3
8,29405	0,70	2	3
8,29405	0,71	1	5
8,41183	0,71	2	5
8,41183	0,71	2	5
8,61250	0,71	4	5
8,29405	0,71	3	2
8,34284	0,71	2	4
8,45532	0,72	3	4
8,24276	0,72	1	4
8,13153	0,74	2	1
8,18869	0,75	1	3
8,29405	0,75	3	2
7,90101	0,81	1	4
8,13153	0,89	2	1
8,43381	0,90	3	1
8,69951	0,90	4	3
8,64822	0,90	3	3
8,16052	0,91	1	4
8,68271	0,91	3	4
8,61250	0,92	4	2
8,51719	1,00	1	2
8,51719	1,00	2	1

8,94898	1,00	4	3
8,57546	1,00	1	3
8,68271	1,00	3	2
8,64822	1,00	2	2
8,57546	1,01	4	1
8,94898	1,01	4	3
9,03599	1,01	4	4
8,88184	1,01	3	3
8,57546	1,01	4	1
8,73230	1,01	2	3
8,64822	1,01	2	2
8,64822	1,01	2	2
8,64822	1,01	2	2
8,43381	1,02	1	3
8,43381	1,02	1	3
8,79482	1,03	5	2
8,53700	1,04	3	1
8,68271	1,05	3	2
8,64822	1,05	2	2
8,43381	1,14	1	3

Çizelge 9.11 Logaritmik şekli

```
Welcome to Minitab, press F1 for help.
```

Results for: Worksheet 1

Regression Analysis: C1 versus C2; C3; C4

```
The regression equation is
C1 = 6,70 + 1,44 C2 + 0,114 C3 + 0,0868 C4

Predictor        Coef    SE Coef         T        P      VIF
Constant       6,7000     0,1056     63,42    0,000
C2            1,43542    0,07940     18,08    0,000      1,6
C3            0,11412    0,01248      9,14    0,000      1,1
C4            0,08681    0,01363      6,37    0,000      1,7

S = 0,1091     R-Sq = 87,1%     R-Sq(adj) = 86,5%

Analysis of Variance

Source           DF        SS        MS        F        P
Regression        3    5,0613    1,6871   141,69    0,000
Residual Error   63    0,7501    0,0119
Total            66    5,8114

Sum of squares for pure error is (nearly) zero.
Cannot do pure error test.
```

```
Durbin-Watson statistic = 1,37

Predicted Values for New Observations

New Obs     Fit      SE Fit           95,0% CI               95,0%
PI
1       9,9727     0,0887    (  9,7955; 10,1499)   (  9,6917;
10,2537) XX
X  denotes a row with X values away from the center
XX denotes a row with very extreme X values

Values of Predictors for New Observations

New Obs          C2          C3          C4
1              2,00        2,00        2,00
```

SONUÇ

Bayes analizi ile elde edilen parametreler anlamlı bulunmuştur. Bayes analizi ile elde edilen sonuçlar EKK yöntemiyle elde edilen sonuçlarla karşılaştırılmıştır. Bayes analizi ile elde edilen parametrelerin EKK ile elde edilen parametrelere göre daha düşük standart sapmaya ve daha dar bir güven aralığına sahip oldukları görülmüştür. Sonuç olarak Bayes analizi EKK yöntemine göre daha tutarlı ve daha güvenilir sonuçlar vermektedir.

Buradan anlaşılacağı üzere pırlanta taşının fiyatında etkili olan parametreleri, EKK yöntemine dayalı olarak analiz edip ona göre yorumlama yapmak,daha geniş bir fiyat aralığı bulmamıza neden olur. Ama pırlanta taşında istenen daha kesin bir sonuçtur. Tam olmasa bile fiyat aralığının daha dar olması istenir. Bu hem alıcı hem de satıcı için istenen bir olgudur. Müşteriye fiyat söylenirken öncelikle taşın özelliklerinden konuşulur. Parametreler hakkında bilgi verilir. Bunlara dayanılarak bir fiyatlandırmaya gidilir. Müşteri de kendisine bilgisi verilen pırlanta fiyatı hakkında , pırlantanın özelliklerine dayanarak bir fikir oluşturmaya başlar. Bu fiyatı diğer yerlerden almış olduğu fiyatlarla kıyaslar. Diğer yandan bu kıyaslamada sadece satıcının konuşmalarını belirleyici bir faktör olarak almaz aynı zamanda taşın sertifikasını da dikkate alır. Sonuç olarak pırlanta taşının kalitesini ve diğer parametrelerini anlamlı bir şekilde değerlendirmek istersek tecrübeli insanların birikimlerinden elde edilen ön bilginin de değerlendirmeye dahil edilmesi gerektiği ortaya çıkmaktadır.

KAYNAKLAR

Birkes, D. Ve Dodge, Y., (1993), "Alternative Methods of Regression", Wiley

Box, G.E.P and G.C. Tiao,1973,"Bayesian Inference in Statistical Analysis";Addison Wesley

Bradley P.Carlin and Thomas A.Louis ,2000,"Bayes and Empirical Bayes Methods for Data Analysis"; Chapman –Hall/

Blue Diamonds Academy Yayınları

Dodge, Y. Ve Arthanari, T.S., (1993), "Mathematical Programming in Statistics", Wiley,

New York:

De Groot,M.H,1970,"Optimal Statistical Decisions";McGraw-Hill Inc.,Newyork

De Groot,M.H,1985,"Probability and Statistics"; Addison Wesley

Evren.A.,2006,"Olasılık Ve Çok Değişkeni Olasılık Dağılımlarına Giriş",YTÜ Yayınları

Genceli,M.,1989,"Ekonometride İstatistik İlkeler", Filiz Kitabevi,İstanbul

Güriş,S.; 2000"Ekonometri", Der Yayınları;

Gürsakal,N., 1982 "Bayesgil İstatistik", B.Ü.Basımevi Bursa,

Hsu ,H.,1997,"Schaum's Outline of Probability" , McGraw-Hill Inc.,Newyork,

Random Variables and Random Processes"1995,Mc Graw- Hill Inc.,Newyork

Kahyaoğlu,F.,1996,"Bayes Analizinde Öngörüsel Olasılık Yoğunluk Fonksiyonu Çıkarımı":MÜ.İstatistik ve Ekonometri Araştırma ve Uygulama Merkezi Dergisi ,

Sayı 1,sf 171-178

Kahyaoğlu,F.,1999, "Regresyonda BayesYaklaşımı ve Cobb-Douglass Üretim Fonksiyonu Üzerine Bir Uygulama," Doktora Tezi

Lee,P.M. 1989,"Bayesian Statistics;An Introduction";Oxford.

Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W., (1993) "Applied Linear Statistical Methods", Wiley

Neil.H.Timm .,1975, "Multivariate Analysis with Applications", Brooks/Cole

Mood, A.M.; Graybill; F.A.; Boes, D. C. 1974; "Introduction to the Theory of Statistics" McGraw-Hill International Editions, Statistics Series, Third Edition,

Press, S.J. et al; 2003 "Subjective and Objective Bayesian Statistics : Principles, Models, and Applications" ; Wiley-Interscience,

Zellner,A, 1971"An Introduction to Bayesian Inference in Econometrics"; John Wiley ,

INTERNET KAYNAKLARI

http://en.wikipedia.org/wiki/Bayesian_inference

http://www.abelard.org/briefings/bayes.htm

http://www.fon.hum.uva.nl/.../CDROM/Literature/LOTwinterschool2006/bidug.pnl.gov/presentations/PEP/

http://www.bayesian.org/

http:// courses.ncssm.edu/math/TALKS/PDFS/BullardNCTM2001.pdf

http://amath.colorado.edu/seminars/2003fall/ProbStatSlides/linreg.pdf

http://209.85.135.104/search?q=cache:FfxMo6M_XpEJ:home.uchicago.edu

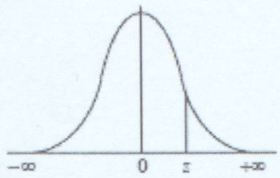
http://limnology.wisc.edu/regime/appendix_14jul03.pdf

http://www.aris.com

EKLER

EK1

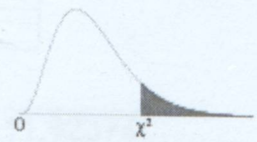
NORMAL DAĞILIM TABLOSU



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

EK 2

KİKARE DAĞILIM TABLOSU



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

EK3

EXCEL ÇALIŞMA KISIMLARI

0,08	0,051	0,116	-0,12
0,051	0,04	-0,0007	-0,165
0,116	-0,0007	1,218	-0,24
-0,12	-0,165	-0,24	1,61

-91,6212	165,7851	11,15091	11,82374
165,7851	-255,634	-19,2282	- 16,70813
11,15091	-19,2282	-0,49097	- 1,212658
11,82374	-16,7081	-1,21266	- 0,390703

469986,2	44067,41	137203,4	175785,5
44067,41	4377,2	12704	15362
137203,4	12704	47700	49700
175785,5	15362	49700	77100

6,7	1,43	0,11	0,08

6,92	1,29	0,1	0,05

0,000206	- 0,001204	-9,85161E- 05	- 0,000166
- 0,001204	0,008268	0,000357019	0,000867
-9,85E- 05	0,000357	0,000150266	5,66E-05
- 0,000166	0,000867	5,66103E-05	0,000182

5,61831	8,738979	0,453432	0,86933

9,07945E- 05	0,000795	-3,21E- 06	5,09163E- 05
0,00079472	0,0052	- 0,000285	- 0,00059337
-3,20737E- 06	-0,0002	0,00012	-1,30164E- 05
5,09163E- 05	- 0,000593	-1,3E-05	2,33644E- 05

-0,0007	-0,165	44067,46	4377,24	12703,84	15361,84	-0,001204	0,008268	0,000357019	[illegible]
1,218	-0,24	137203,5	12704	47699,76	49699,76	-9,85E-05	0,000357	0,000150266	[illegible]
-0,24	1,61	175785,4	15361,84	49701,61	77101,61	-0,000166	0,000867	5,66103E-05	[illegible]

621	165,7851	11,15091	11,82374	6,92	1,29	0,1	0,05	-418,45	814,7074	52,25016	60
78	-255,634	-19,2282	-16,70813								
5091	-19,2282	-0,49097	-1,212658								
2374	-16,7081	-1,21266	-0,390703								
986,2	44067,41	137203,4	175785,5	6,7	1,43	0,11	0,08	3241079	304137,4	946652,5	12
67,41	4377,2	12704	15362					3240661	304952,2	946704,8	12
203,4	12704	47700	49700								
785,5	15362	49700	77100								

ÖZGEÇMİŞ

Doğum tarihi	20.11.1980	
Doğum yeri	Mardin	
Lise	1994-1997	Beşiktaş Lisesi
Lisans	1998-2003	Yıldız Üniversitesi Fen Edebiyat Fak. İstatistik Bölümü

Çalıştığı kurumlar

	1989-Devam ediyor	Murat Kuyumculuk

00103111