

İ.T.O.
KUTUPHANESİ - KÜTÜPMANTASYON
LİSANS BAŞKANLIĞI

T. C.
İSTANBUL TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

102
mm

**VERİ İŞLEME PROSEDÜRLERİ
VE
HATALAR**



YÜKSEK LİSANS TEZİ

ANA BİLİM DALI : ELEKTRİK
PROGRAMI : DEVRELER VE SİSTEMLER

ELK. MÜH. HASAN ÇAM

TEZ DANIŞMANI : PROF. DR. ENİSE ERİMEZ

HAZİRAN, 1984

Ö N S Ö Z

Bu tez çalışmamızda, bilhassa bir adet random veri fonksiyonunun istatistiksel özelliklerini analiz etmek için gerekli olan prosedürlerden birisini inceledik. Olasılık yoğunluk fonksiyonları, korelasyon fonksiyonları ve güç spektral yoğunluk fonksiyonlarını tahmin ederken yaptığımız istatistiksel hataları belirledik. Ayrıca, random veriyle ilgili temel kavramların önemli özelliklerini ve bu özellikleri ölçmek için kullanılan bazı analog ve sayısal metodları anlattık.

Tez çalışması boyunca bana yardımlarını esirgemeyen ve önemli katkılarda bulunan çok değerli hocam Sayın Prof. Dr. Enise Erimez'e teşekkür etmeyi bir borç bilirim.

Hasan ÇAM

Kayıt No. 4378
Şön 1985-4.
UDC

TLK
U29
ÇAM

İÇİNDEKİLER

	<u>Sahife No:</u>
Giriş	1
BÖLÜM 1. Örnek Fonksiyonları Analiz Etme Metodu	4
1.1 Veriyi Elde Etme	6
1.2 Veriyi Örnekleme	13
1.2.1 Random Örnekleme	14
1.2.2 Random Veriyle İlgili Örnek Alma Teoremleri	16
1.3 Depolama ve Transmisyon	18
1.4 Seri Bakış Analizi	19
1.5 Stasyonierlik Testi	19
1.6 Verinin Random Olup Olmadığını Anlamak İçin Yapılan Test	24
1.7 Normal Dağılıma Uygunluğu (Normaliteyi) Anlamak İçin Yapılan Testler	30
1.8 Ortalama ve Karesel Ortalama Seviye Analizi	39
1.8.1 Ortalama ve Karesel Ortalama Değer Ölçümleri	42
1.9 Güç Spektrumu Yoğunluk Analizi	47
1.10 Otokorelasyon Analizi	55
1.11 Olasılık Yoğunluk Fonksiyonunun Analizi	61
1.12 Özel Amaçlı Analizler	65
1.13 Nonstasyonier Veri Analizi	66
1.14 Sinüzoidal Elemanların Ayırımı	68
BÖLÜM 2. Parametre Tahminlerinde Yapılan İstatistiksel Hatalar	70
2.1 Ortalama Değer Tahminleri	73
2.2 Karesel Ortalama Değer Tahminleri	76
2.3 Olasılık Yoğunluk Fonksiyonu Tahminleri	77

2.3.1 Tahmin Edilen Fonksiyonun Varyansı	79
2.3.2 Biased Tahminin Hatası	81
2.3.3 Ortak Olasılık Yoğunluk Tahminleri	82
2.4 Korelasyon Fonksiyonunun Tahminleri	84
2.5 Ortalama ve Karasıl Ortalama Değerler	86
FAYDALANILAN KAYNAKLAR	89

G İ R İ Ş

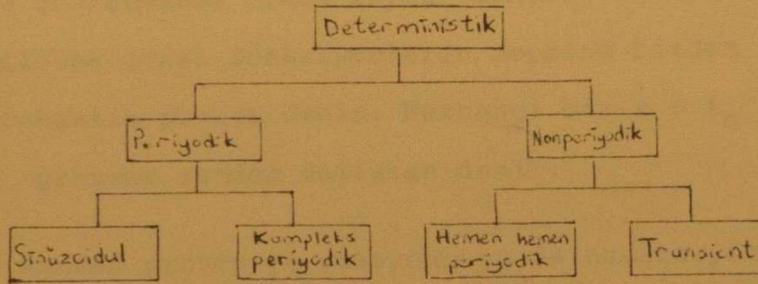
Random verinin özelliklerini belirleyen prosedürleri iki sınıfa ayırırız. Birinci sınıf, bir örnek kaydın istatistiksel özelliklerini analiz eden prosedürdür. Diğeri ise, bir grup örnek kaydın her birinin istatistiksel özelliklerini belirleyen prosedürdür. Bu prosedürler, tek ve grup halindeki örnek kayıtları incelemek için gerekli olan parametreleri elde etmenin, temel özelliklerini belirlemenin ve analiz etmenin sistematik metodlarını araştırırlar.

İlgilendiğimiz veri random olduğundan parametre tekniklerinde yapılan istatistiksel hataları doğru şekilde hesaplamamız gerekir. Bundan dolayı, ortalama değerleri, karesel ortalama değerleri, olasılık yoğunluk fonksiyonlarını, korelasyon fonksiyonlarını ve spektral yoğunluk fonksiyonlarını belirlemeye çalışırken yapılan tahminlerin sapma hatalarını hassas şekilde belirlemeye çalışacağız.

Fiziksel olaylarla ilgili herhangi bir gözlenen veri ya deterministik ya da nondeterministik (random)'dır. Deterministik veri belli matematik fonksiyonlarla tarif edilebilir. Gözlenen olayın gelecekteki herhangi bir zamanda ne değeri alabileceği denklem ve fonksiyonlarla tesbit edilir. Ancak bazı olayların ilerideki herhangi bir zamanda alacağı değerler hassas şekilde belirlenemez. İşte bu şekildeki verilere random veri denir. Bu verilerin olasılık fonksiyonları, istatistiksel

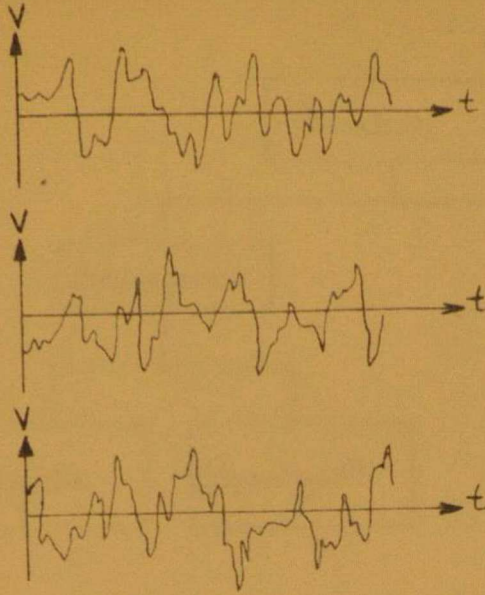
ortalamalar vasıtasıyla belirlenir. Örneğin, gürültü jeneratörünün elektriksel çıkışı, kabaran denizdeki dalgaların yüksekliği bu tip verilerdendir.

Deterministik veriler ya periyodiktir ya da nonperiyodiktir. Periyodik veri kendi arasında "sinüzoidal" ve "nonsinüzoidal (kompleks periyodik)" olmak üzere ikiye ayrılır. Nonperiyodik veri de kendi arasında "hemen hemen periyodik" ve "transient" olmak üzere ikiye ayrılır. Bu sınıflandırma aşağıdaki şemada gösterilmiştir.



Şema 1: Deterministik verinin Sınıflandırılması

Random fiziksel olayda her bir gözlem değerinden bağımsız olduğundan dolayı random veri belli matematik fonksiyonlarla tanımlanamaz. Gürültü jeneratörü çıkış geriliminin, zamanın fonksiyonu olduğunu kabul edersek aşağıdaki şekilde gösterildiği gibi değişik örnek fonksiyonları elde ederiz.

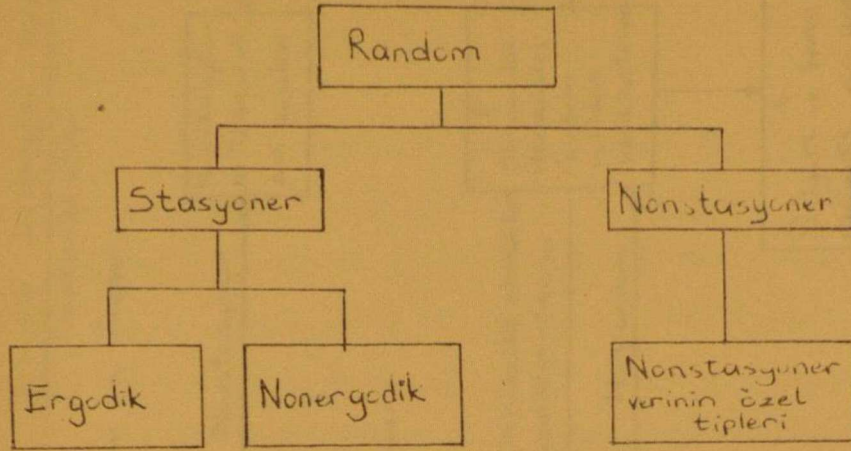


Şekil 1: Gürültü Jenaratörünün çıkışından elde edilen örnek kayıtlar.

Random olayın zamana göre durumunu gösteren fonksiyonlardan her birine örnek fonksiyon (bu olay sonlu bir zaman aralığında gözlenmişse örnek kayıd) denir. Random olaydan elde edilebilecek örnek fonksiyonların hepsine birden random proses veya stokastik proses denir. Herhangi bir $t = t_0 = st$ anındaki $\{x(t)\}$ prosese random değişken denir.

Random proses ya stasyoner ya da nonstasyonerdir. Stasyoner random proses ise kendi arasında ergodik ve nonergodik olmak üzere ikiye ayrılır. Nonstasyoner random proses de nonstasyoner özelliklerin özel tiplerine ayrılır. Bu sınıflandırma aşağıdaki şemada görülmektedir.

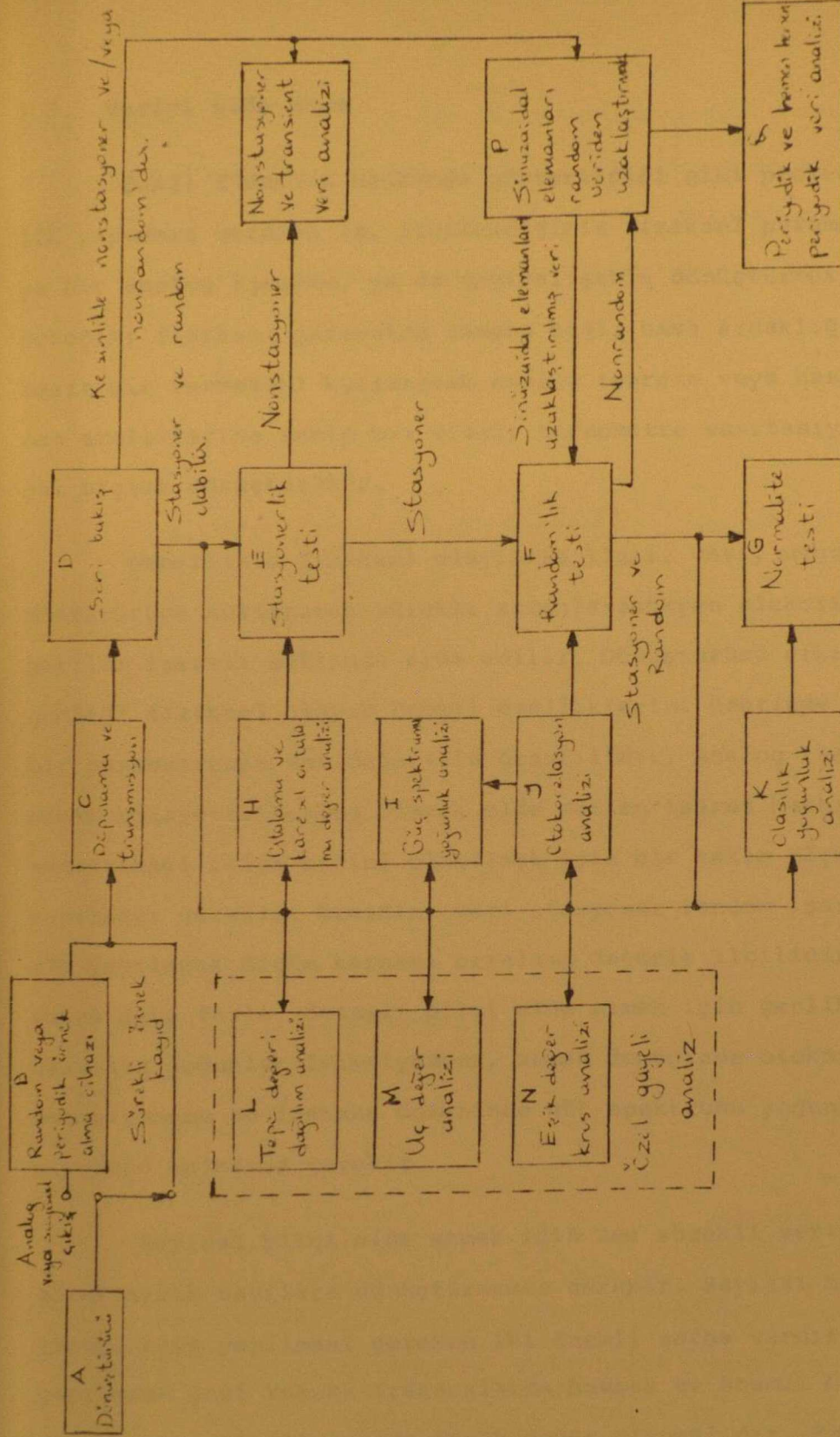
$\{x(t)\}$ random değişkenler kümesi (random proses) ve bu kümenin olasılık dağılımlarının hepsine birden stokastik proses denir.



Şema 2: Random verinin sınıflandırılması.

1. ÖRNEK FONKSİYONLARI ANALİZ ETME METODU

Bir örnek kaydın uygun istatistiksel özelliklerini analiz etmek için kullanılan prosedür aşağıdaki şekilde gösterilmiştir (Şekil 2). Şekildeki her bir blok sayısal ya da analog tekniklerle gerçekleştirilebilir. Şekilden de görülebileceği gibi random verinin detaylı analizine başlamadan önce, random veriyle ilgili istatistiksel tahminlerin hassas şekilde yapılması gerekir. Önce, örneklenmiş verinin a) random, b) stasyonær ve c) normal olup olmadığının belirlenmesi gerekir. Ancak, bu belirleme işlemlerini yapabilmek için bir takım cihaz ve aletlerin kullanılması şart değildir. Tecrübeli bir analizci eldeki mevcut verileri gözleyerek periyodikliği, stasyonærliği veya normalitedeki standart sapmayı tesbit edebilir. Demek ki, random veriyi analiz etmek için 3 temel karakteristik olan stasyonærlik, randomness ve normalite'nin bilinmesi gerekmektedir.



Şekil 2: Örnek Kayıtları tek tek analiz eden prosedür.

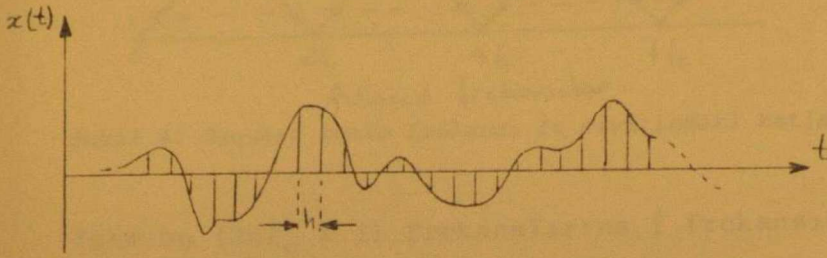
1.1. Veriyi Elde Etme

Şekil 2'deki A blokunda gösterildiği gibi prosedürde ilk yapmamız gereken iş, ilgilendiğimiz fiziksel parametreyi ya bir analog işarete, ya da sayısal şekle dönüştürmektir. Örneğin; fiziksel parametre zamana bağlı hava sıcaklığı ise basit bir termokupl kullanarak analog işarete veya hassas zaman aralıklarına sahip bir civalı termometre vasıtasıyla sayısal biçime dönüştürülür.

Genellikle fiziksel olaylarla ilgili veri, uygun bir dönüştürücü kullanarak sürekli zamanla-değişen elektriksel gerilim işareti şeklinde elde edilir. Dönüştürücü çıkışının genliği fiziksel olayın önemli özelliklerini üzerinde toplayan bir parametrenin ani değeriyle orantılıdır. Analog işaret elde etmek kolaydır. Analog olarak elde edilen işaret hakkında istatistiksel bilgi sahibi olabilmek için bir takım ölçmelerin yapılması gerekir. Örneğin, veri stasyoner random işareti ise ilk yapılacak ölçüm karesel ortalama değerle ilgilidir. Bundan sonra daha fazla detaylı bilgi elde etmek için genlik domeninde olasılık yoğunluk fonksiyonunu, zaman domeninde otokorelasyon fonksiyonunu ve frekans domeninde güç spektrumu yoğunluk fonksiyonunu bulmamız gerekir.

Sayısal bilgi elde etmek için ise sürekli veriyi (zamana göre) ayırık sayılara dönüştürmemiz gerekir. Sayısal hale dönüştürmek için yapılması gereken iki önemli safha vardır. Birinci yapılacak şey: Yüksek frekanslarda hassas ve önemli bilgi elde etmek için yeterli sayıda örnek değer alınmalıdır. Fakat birbirine çok yakın olan noktalardan örnek almak hem hesaplamalarda

çok uğraşmayı gerektirir hem de hesaplamaların, ölçmelerin maliyetini arttırır. Bundan dolayı, örneklerin sayısı azaltılmalıdır, yani örnekleme hızının aliasing hatalarına sebep olmayan en düşük hıza indirilmesi gerekmektedir.



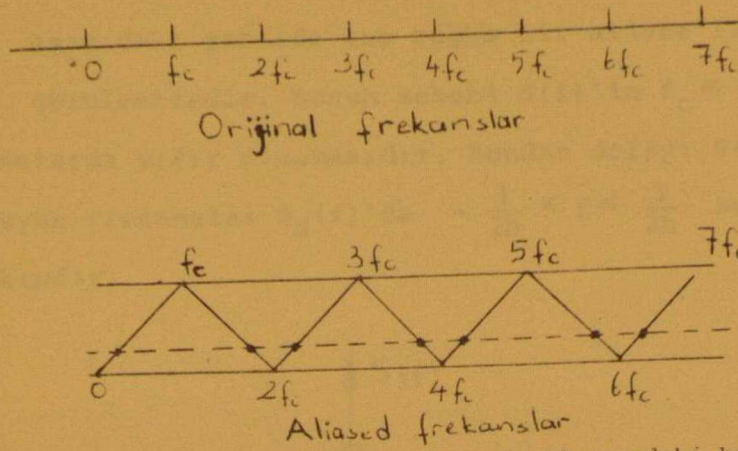
Şekil 3: Sürekli örnek kaydın örneklenmesi

Şekil 3'de örnek alma vasıtasıyla sürekli verinin ayırık veriye dönüştürülüşü gösterilmektedir. Örnekler arasındaki Δt zaman aralığı h sn ise örnekleme hızı saniyede $1/h$ 'dir. Eğer $h = 0.1$ sn ise veride 5 Hz'den daha büyük frekansların olmaması gerekir. Çünkü 5Hz'den büyük frekanslar (0-5) Hz aralığındaki frekanslara katlanır ve bu aralıktaki verinin anlaşılmasının zorlaştırırlar. Kesim frekansı

$$f_c = \frac{1}{2h} \quad (1.1)$$

Nyquist frekansı olarak bilinir. Frekans ekseninin katlanması aşağıdaki şekilde gösterilmiştir.

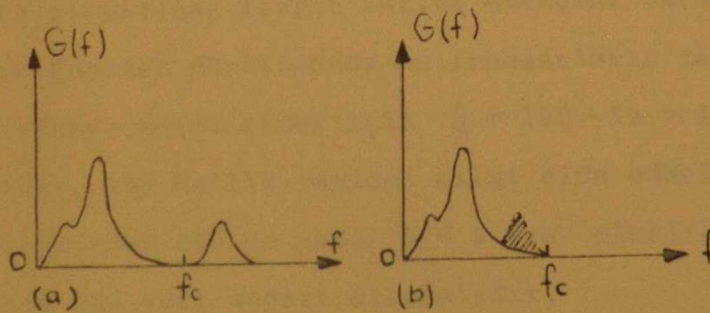
Eğer veride f_c 'den daha büyük frekanslar varsa $0 \leq f \leq f_c$ aralığındaki herhangi bir frekansın ayırık transformu, bu frekans ve $(2f_c \pm f)$, $(4f_c \pm f)$, ... $(2nf_c \pm f)$, frekanslarının birleşiminden meydana gelir.



Şekil 4: Nyquist kesim frekansı f_c civarındaki katlanma.

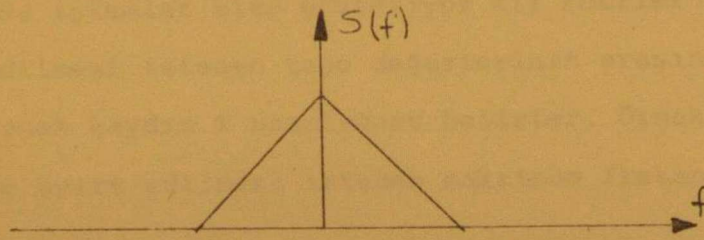
İşte bu $(2nf_c \pm f)$ frekanslarına f frekansının aliases'leri ve Fourier dönüşümündeki etkilerine ise aliasing denir.

Örnek: $f_c = 100$ Hz ise 30 Hz'deki veri 170 Hz, 370 Hz, 430 Hz, ... gibi frekanslardaki verilerin toplamıyla birleşecektir. Yani 100 Hz'den fazla frekanslar veride varsa, örnek alma işleminde aliasing hataları olacaktır. Benzer şekilde, bu istenmeyen yüksek frekanslardaki güç daha düşük frekanslardaki güçle üst üste gelecektir. f_c frekansı aşağıdaki şekilde gösterildiği gibi ise a şıkında gösterilen gerçek güç spektral yoğunluk fonksiyonu b şıkında gösterilen aliased güç spektral yoğunluk fonksiyonuna dönüşür.

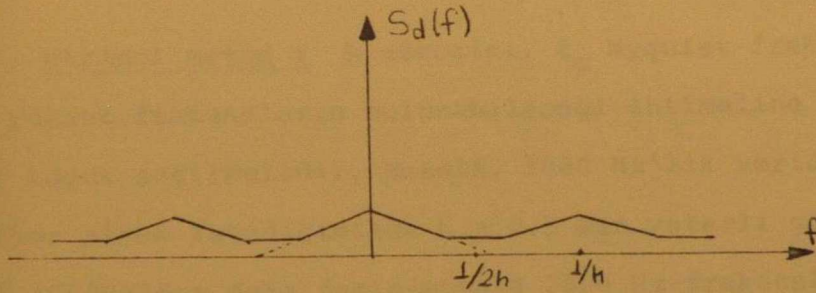


Şekil 5: Katlanmaya sebep olan aliased güç spektrumu a) Gerçek Spektrum
b) Aliased spektrumu.

Aşağıdaki şekilde ise başka bir aliase'li ayırık dönüşüm $S_d(f)$ görülmektedir. Bunun sebebi $S(f)$ 'in $f_c = \frac{1}{2h}$ dan yüksek frekanslarda sıfır olmamasıdır. Bundan dolayı $S(f)$ 'daki $1/2h$ dan büyük frekanslar $S_d(f)$ 'de $-\frac{1}{2h} \leq f \leq \frac{1}{2h}$ aralığında görülmektedir.



Şekil 6.a: İşaretin Fourier dönüşümü



Şekil 6.b: $f_c = \frac{1}{2h}$ olan örneklenmiş işaretin Fourier dönüşümü.

Örnek : Uzunluğu T ve örnekleme aralığı olan bir örnek kayıt düşünelim. İşaretimizin 100 Hz ve 99 Hz olmak üzere iki sinüzoidal işaret ihtiva ettiğini varsayalım. Örnek kaydın Fourier dönüşümünde bu frekansların tepe değerlerini birbirinden ayırt etmek için $\frac{1}{T} = 100 - 99 = 1$ Hz, $T = 1$ sn olmalıdır. 100 Hz'lik veriden bilgi elde edebilmek için $\frac{1}{2\Delta}$ aralığı en azından 100 Hz yani $\Delta \leq 5$ msn olmalıdır. Böylece 1 saniyede 200 veri değeri alınmalıdır.

Eğer 999 Hz ve 1000 Hz frekans bileşenlerinin tepe değerlerini birbirinden ayırmak istiyorsak, yine 1 sn'lik örnek kayıt süresi gerekli olur. Fakat örnekleme aralığı $\Delta \leq 0.5$ msn olmalıdır. Bu şekilde 1sn'de 2000 veri değeri elde alınmalıdır.

Bu işlemler bize gösteriyor ki; Fourier dönüşümünde ayırt edilmesi istenen tepe değerlerinin arasındaki frekans farkı örnek kaydın T uzunluğunu belirler. Örnekleme aralığı Δ 'yı ise ayırt edilmesi istenen maksimum frekans tesbit eder.

Aliasing problemini önlemek için 2 tane pratik metod vardır.

Birinci metod : h süresini, f_c Nyquist frekansından daha yüksek frekansların bulunabileceği ihtimaline göre yeteri kadar küçük seçilmelidir. Meselâ, 1000 Hz'lik veriden örnek değerler almak istediğimizde $h = 0.5$ msn yeterli görülebilir. Ancak 1000Hz'den daha fazla-meselâ 2000 Hz-frekansların bulunabileceği ihtimali varsa Nyquist frekansı $f_c = 2000$ Hz ve $h = 0.25$ msn olmalıdır. Genellikle f_c seçilirken, verideki maksimum frekansın 1.5 veya 2 katı alınır.

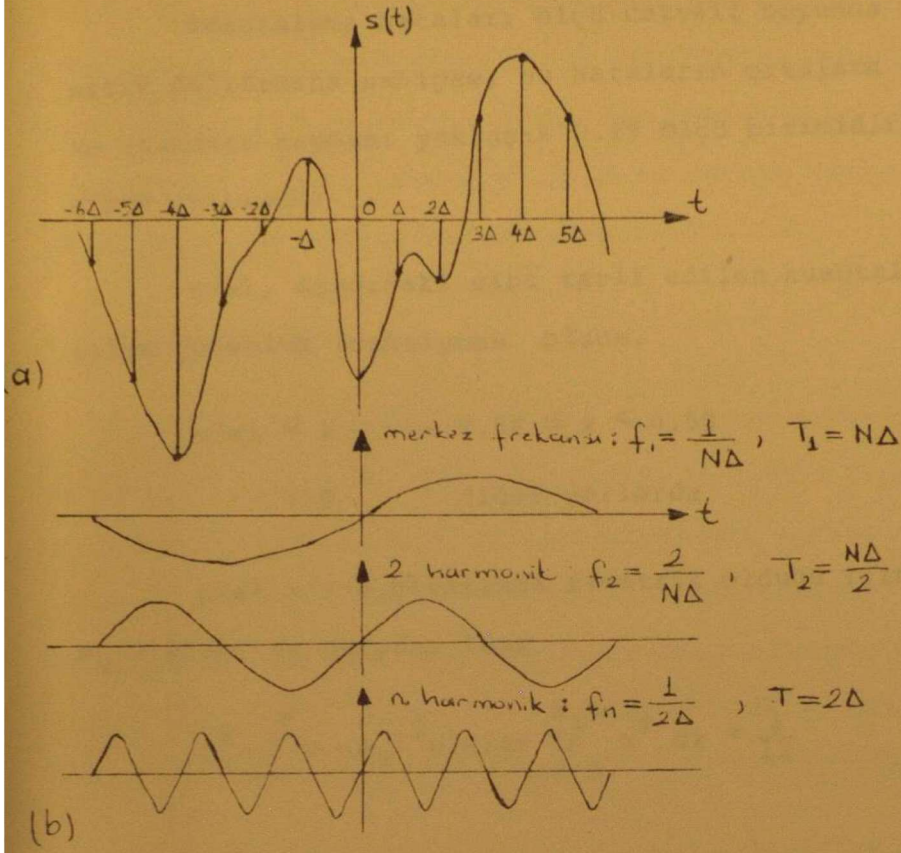
İkinci metod : Örnek alınması istenen maksimum frekans-tan daha büyük frekansın veride olmaması için, veri örneklenmeden önce bir filtreden geçirilir. Filtreden geçirdikten sonra f_c 'yi istenen maksimum frekansta seçeriz ve doğru sonuçlar elde ederiz. Bu metod birinci metoda göre tercih edilir.

Periyodik fonksiyonları kullanarak non-periyodik fonksiyonları göstermek mümkündür. Eğer faz açıları önemsenmezse periyodik veri gibi hemen hemen periyodik verilerin de ayrık frekans spektrumu elde edilir. T uzunluğundaki sürekli bir işaretin Δ aralıklarla örneklendiğini ve periyodik sinüs ve cosinüs fonksiyonlar cinsinden ifade edilmek istendiğini farzedelim. $N = \frac{T}{\Delta}$ kadar örnek değer elde edilir. N çiftse $2n$ 'e, tekse $2n-1$ 'e eşittir (n tam sayıdır). Aşağıdaki şekilde gösterildiği gibi, ayrık işaret $\tilde{s}(t)$ fundemantal sin dalgası ve harmoniklerinin (temel frekansların katları) toplamından oluşan ve sürekli işaretin bütün özelliklerini üzerinde yaklaşık olarak taşıyan bir işaret elde edilir.

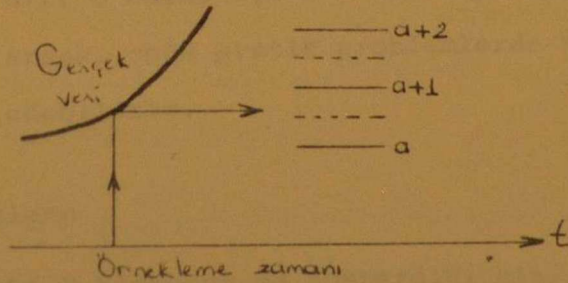
$$\tilde{s}(t) = A_0 + 2 \sum_{m=1}^{n-1} \{ A_m \cos 2\pi m f_1 t + B_m \sin 2\pi m f_1 t \} + A_n \cos 2\pi n f_1 t \quad (1-2)$$

N tek olursa $A_n = 0$ 'dır. $f_1 = \frac{1}{N\Delta}$ frekansına işaretin fundemantal frekansı denir ve şekilde gösterildiği gibi örnek kaydın uzunluğuna eşit periyodu vardır. $\tilde{s}(t)$ fonksiyonunda mevcut olan en yüksek frekans $\frac{n}{N\Delta} = \frac{1}{2\Delta}$ Hz' dir ve periyodu 2 örnek aralığının uzunluğuna eşittir.

Sayısal hale getirme (digitization) metodunun ikinci kısmı kuantalama (quantization) problemidir. Yani gözlenen veriye en uygun nümerik değeri karşılık getirirken ortaya çıkan problemi inceliyeceğiz. Bazen, iki ardışıl nümerik değerden birisini seçmek mecburiyetinde kalırız. Bu problemi gösteren aşağıdaki şekilde, en yakın nümerik değer olarak (a+1) seçilir (Şekil 8).



Şekil 7: a) Sürekli bir işareti örnekleyerek elde edilen bir ayrık işaret.
b) Fundamental sin dalgası ve harmonikleri.



Şekil 8: Kuantalama hatası

Kuantalama hataları ölçü cetveli boyunca uniform olasılık dağılımına sahipse, bu hataların ortalama değeri sıfır ve standart sapması yaklaşık 0.29 ölçü birimidir. Bunu ispat etmek istersek:

$p(x)$, aşağıdaki gibi tarif edilen kuantalama hata olasılık yoğunluk fonksiyonu olsun.

$$\begin{aligned} p(x) &= 1 & -0.50 \leq x \leq 0.50 \\ &= 0 & \text{diğer yerlerde} \end{aligned}$$

$p(x)$ $x = 0$ civarında simetrik olduğu için ortalama değer $\mu_x = 0$ 'dır ve varyans ise;

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x) dx = \int_{-0.5}^{0.5} x^2 dx = \frac{1}{12}$$

$$\sigma_x = \sqrt{\frac{1}{12}} \approx 0.29 \text{ ölçü brimi (Standart sapma)}$$

Standart sapma kuantalama hatasının rms değeridir ve rms gürültüsü olarak da düşünülebilir. İşaret 256 tane ölçü değerinde kuantalanıyorsa rms gürültüsünün işarete oranı $0.29/256 \approx 0.001$ 'dir. O halde işareti 256 veya daha fazla ölçü biriminde kuantalarsak, çoğu pratik problemlerde kuantalama hatalarını ihmal edebiliriz.

1.2. Veriyi Örnekleme

Şekil 2'deki B blokunun da gösterdiği gibi, dönüştürücüden çıkan veriden herhangi bir örnekleme metodu yardımıyla örnekler alınmalıdır. Her bir zamana bağlı örnek kaydın uzunluğu, analizin istatistiksel hassasiyeti bakımından çok önemlidir.

Ayrıca, elde edilen örneklerin sayısı ve seçilen örnek alma metodu grup halindeki örnek kayıtların istatistiksel özelliklerini belirlemede çok önemlidir. Genellikle örnekler ya uzun süreli sürekli kayıt ya da kısa süreli random veya periyodik kayıtlar kümesi olarak elde edilmelidir.

1.2.1. Random Örneklem

Veri aralığında en önemli nokta çok az veriden çok fazla bilgi elde etmektir. Uygun örneklem metodunu (random veya periyodik) seçenek istenilen hassasiyetdeki parametreleri tanımlamak için gerekli olan veri hacmini azaltabiliriz. Her bir örnek kaydın gerekli uzunluğu istatistiksel hata formülleriyle tesbit edilir. Bundan dolayı random örneklem metodunda, sadece örnekler arasındaki zaman aralığı random olur.

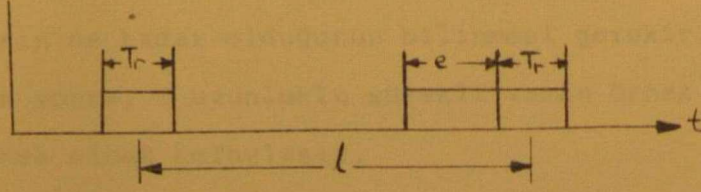
Random örneklem metodunda en zor yapılan şey, örnek kayıtlar arasında meydana gelen beklenmeyen önemli olayların kaybolma olasılığını belirlemektir. Bu olasılığı belirleyebilmemiz için aşağıdaki özellikleri bilmemiz gerekir.

e = saniye cinsinden beklenmeyen olayın uzunluğu

T_r = saniye cinsinden her bir örneğin uzunluğu

ℓ = saniye cinsinden örnekler arasındaki ortalama zaman (merkezden merkeze ölçülüdür).

e 'nin ℓ boyunca bir kez meydana geldiğini ve $T_r + e \leq \ell$ olduğunu farzederek aşağıdaki şekli çizebiliriz.



Şekil 9: Örnek değerlerin alınışı

Random örnekleme olasılık yoğunluk fonksiyonunun uniform olduğunu kabul edersek, e 'nin bütün parçalarının ortalama kaybolma olasılığı:

$$\bar{p}(\text{kaybolan } e) = \frac{l - (T_r + e)}{l} = 1 - \frac{T_r + e}{l} \quad (1.3)$$

Fakat önceden tesbit edilmiş örnekleme zamanları olduğunda yukarıdaki denklem uygulanmaz ve $\bar{p}$ aşağıdaki gibi belirlenir.

$$\bar{p}(\text{kaybolan } e) = \begin{cases} 0 & e, \text{ örnek kayıt uzunluğunun içinde oluşuyorsa} \\ 1 & e, \text{ örnek kayıt uzunluğunun dışında oluşuyorsa} \end{cases}$$

e 'nin herhangi bir parçasının ortalama kaydedilme olasılığını bulabiliriz.

Örnek kaydın random şekilde örneklendiğini varsayarak 100 sn, $T_r = 5$ sn ve $e = 10$ sn alırsak;

$$\bar{p}(\text{kaydedilen } e) = 1 - \left(1 - \frac{T_r + e}{l}\right)^{10} = 0.84$$

T_r uzunluklu örnek değerlerin merkezleri arasındaki uzunluğu uygun şekilde seçebilmek için, önce beklenmeyen olayın

ne kadar uzunlukta olduğu ve müsaade edilen kaybolma olasılığının ne kadar olduğunun bilinmesi gerekir. ℓ tesbit edildikten sonra, T uzunluklu sürekli zaman örnek kaydından N tane örnek olmak kolaylaşır.

$$N = \frac{T}{\ell} \quad (1.4)$$

e uzunluklu beklenmeyen olayların istenen deteksiyon (detection) olasılığını bulabilmek için her birinin uzunluğu T_r olan N tane örnek değer gereklidir.

Örnekler arasındaki uzaklık ℓ sn olduğunda, örnekleme periyodu saniyede $1/\ell$ örnektir. Verideki lineer eğilimleri görmek periyodik örnekleme ile daha kolay olduğu için random örnekleme tercih edilir. Veri hem random hem stasyoner ise keyfi random zamanlarda örnekler alınır ve hata yapılmaz, fakat veri nonstasyoner ise sonuçlar gözlemin belirli zamanlarına bağlıdır ve özel metodlarla analiz edilir.

1.2.2. Random veriyle ilgili Örnek Alma Teoremleri

Sadece 0 ve T saniyeleri arasında değer alan ve diğer bütün zamanlarda sıfır olan $x(t)$ örnek fonksiyonun Fourier dönüşümü

$$X(f) = \int_0^T x(t) e^{-j2\pi f t} dt \quad (1.5)$$

dir. Eğer $x(t)$, periyodu T olan bir periyodik zaman fonksiyonu ise temel frekans artışı (fundamental frequency increment) $f = 1/T$ 'dir ve Fourier serisinin açılımı,

$$x(t) = \sum_{-\infty}^{\infty} A_n e^{j2\pi nt/T} \quad (1.6)$$

$$A_n = \frac{1}{T} \int_0^T x(t) e^{-j2\pi nt/T} dt \quad (1.7)$$

$$X\left(\frac{n}{T}\right) = \int_0^T x(t) e^{-j2\pi nt/T} dt = T \cdot A_n \quad (1.8)$$

Böylece $X\left(\frac{n}{T}\right)$ bütün f 'ler için $X(f)$ 'i belirler. Bu işleme frekans domeninde örnek alma teoremi denir. Temel frekans artışı $1/T$ 'ye Nyquist aralığı (cointerval) denir.

Ters Fourier dönüşümü vasıtasıyla da $x(t)$ 'yi bulabiliriz. Örneğin, $x(t)$ 'nin sadece $(-B)$ Hz ile $(+B)$ Hz frekansları arasında mevcut, diğer frekanslarda sıfır olduğu farzedilirse:

$$x(t) = \int_{-B}^B X(f) e^{j2\pi ft} df \quad (1.9)$$

$X(f)$, periyodu $2B$ Hz olan bir periyodik frekans fonksiyonu ise temel zaman artışı (fundamental time increment) $t = \frac{1}{2B}$ 'dir.

$$X(f) = \sum_{-\infty}^{\infty} C_n e^{-j\pi nf/B} \quad (1.10)$$

$$C_n = \frac{1}{2B} \int_{-B}^B X(f) e^{j\pi nf/B} df \quad (1.11)$$

$$X\left(\frac{n}{2B}\right) = \int_{-B}^B X(f) e^{j\pi nf/B} df = 2B C_n \quad (1.12)$$

Böylece bütün t 'lerde $x(t)$ belirlenmiş olur. Bu işleme, zaman domeninde örnek alma teoremi denir. Temel zaman artışı $\frac{1}{2B}$ 'ye Nyquist aralığı ve $n = 2BT$ 'ye serbestlik derecesi miktarı

denir. Çünkü, n istatistiksel bakımından bağımsız olan ayrık örneklerin sayısını gösterir.

$X(f)$ 'i $-B$ ve B frekansları arasında Nyquist aralığı noktaları $1/T$ 'de örnekleyerek $x(t)$ 'yi belirlemek için gerekli olan ayrık örneklerin sayısını buluruz.

$$n = \frac{2B}{1/T} = 2BT \quad (1.13)$$

$x(t)$ 'yi 0 ve T seviyeleri arasında Nyquist aralığı noktaları $1/2B$ 'de örnekleyerek n bulunur.

$$n = \frac{T}{1/2B} = 2BT \quad (1.14)$$

Böylece frekans skalasındaki Nyquist aralıkları ile zaman skalasında Nyquist aralıkları örneklendiğinde ayrık örneklerin sayısı aynı olmalıdır.

1.3. Depolama ve Transmisyon

Örnek kayıtlar elde edildikten sonra, örnekler ya hemen analiz edilmeli ya da daha sonraki analizler için uygun şekilde kaydedilmeli ve depo edilmelidir. Çünkü analiz teknikleri her bir örnek kaydın incelenmesini tekrarlattırabilir. Analog veriyi depolamak için manyetik teyde kullanılabilir. Sayısal veriyi depolamak için delgi kartları, manyetik teyp gibi pek çok teknik kullanılabilir.

Örnek kayıtlar işaret kaynağındaki manyetik teyp kaydedilebilir. Ancak bazı durumlarda, veriyi analiz etmek veya

kaydetmek için daha uygun olan başka bir yere işaret kaynağın-
dan transfer etmek gerekebilir. Örneğin, uçaklarda çevreyle
ilgili olarak toplanan veri sık sık telemetri aleti kullanan
yer'e nakledilir.

1.4. Seri Bakış (Quick Look) Analizi

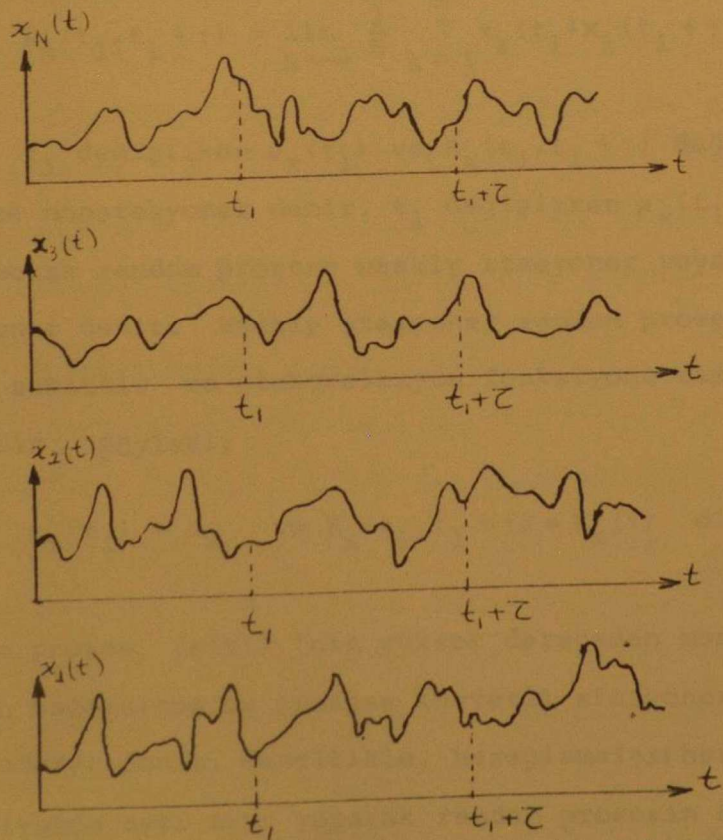
Bu kademedede, prosedür örnek kayıtlarının iki sınıfa
böler. Birinci sınıf, stasyoner ve random olan kayıtları; ikin-
ci sınıf ise nonstasyoner ve/veya nonrandom olan kayıtları ih-
tiva eder. E ve F bloklarında ise örnek kaydın stasyoner veya
random olup olmadığını belirleyen istatistiksel prosedürler
tartışılacaktır. Bununla beraber detaylı bir analize başlama-
dan önce, D blokunda gösterildiği gibi örnek kaydın nonstasyoner
veya nonrandom olup olmadığını görmek için seri bakış analizi
yapılır. Analog veri kayıtlarının seri bakış analizini yapa-
bilmek için kayıtlar katot ışıklı tüpte veya galvanometre osi-
lografide gösterilebilir. Zaman eğilimleri (time trends) veya
kuvvetli sinüzoidal elemanların varlığı hemen görülebiliyorsa,
daha teferruatlı analizlerde vakit harcamadan örnek kaydın non-
stasyoner veya nonrandom olduğuna hemen karar verilebilir.
Veri nonstasyonerse Q blokuna, nonrandom ise P blokuna trans-
fer edilir.

1.5. Stasyonerlik Testi

Seri bakış analizi yapıldığında stasyoner olması mümkün
kabul edilen örnek kayıtlar E blokunda daha dikkatli bir şe-
kilde araştırılır.

Örnek kayıd stasyonerse, bazı analiz prosedürlerini yapmak kolaylaşır. Deterministik verinin transient biçimi genellikle nonstasyoner gözükür. Bundan dolayı bu şekilde sınıflandırılmalıdır. Deterministik transient veriyle nonstasyoner random veri arasındaki farkı örnek kayıtları tek tek inceleyerek göremeyiz; grup halindeki örnek kayıtları incelememiz gerekir. Nonstasyoner (veya transient) olan örnek kayıtlar P blokuna transfer edilir. Hemen hemen periyodik verinin dışındaki nonperiyodik verilere transient veri denir.)

Random prosesi oluşturan bir grup örnek fonksiyon aşağıdaki şekilde gösterilmiştir.



Şekil 10: Random prosesi oluşturan örnek fonksiyonlar

Random prosesin t_1 anındaki ortalama değeri (birinci momenti), gruptaki her bir örnek fonksiyonun t_1 anındaki değerleri toplamının, örnek fonksiyonlarının toplam sayısına bölünmesiyle elde edilir. Benzer şekilde, random prosesin iki farklı zamandaki değerlerinin arasındaki otokorelasyon fonksiyonu da t_1 ve $(t_1 + \tau)$ zamanlarındaki ani değerler çarpımının ortalamasını alarak hesaplanır. Random process $\{x(t)\}$ (örnek fonksiyonlar kümesi)'nin ortalama değeri $\mu_x(t_1)$ ve otokorelasyon fonksiyonu $R_x(t_1, t_1 + \tau)$ aşağıdaki gibi formüllerle ifade edilebilir.

$$\mu_x(t_1) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x_k(t_1) \quad (1.15)$$

$$R_x(t_1, t_1 + \tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x_k(t_1) x_k(t_1 + \tau) \quad (1.16)$$

t_1 değişirken $\mu_x(t_1)$ ve $R_x(t_1, t_1 + \tau)$ değişirse random prosese nonstasyonier denir. t_1 değişirken $\mu_x(t_1)$ ve $R_x(t_1, t_1 + \tau)$ değişmezse random prosese weakly stasyonier veya geniş anlamda stasyonier denir. Weakly stasyonier random proses'de ortalama değer sabittir ve otokorelasyon fonksiyonu zaman aralığı τ 'ya bağlıdır. Şöyleki;

$$\mu_x(t_1) = \mu_x \quad \text{ve} \quad R_x(t_1, t_1 + \tau) = R_x(\tau) \quad \text{dir.} \quad (1.17)$$

Random proses $\{x(t)\}$ 'nin yüksek dereceden momentleri de zamandan bağımsızsa bu prosese kuvvetli stasyonierdir (strongly stationary) denir. Genellikle, hesaplamaları her bir örnek fonksiyonda ayrı ayrı yaparak random prosesin özelliklerini

bulmaya çalışırız. Örneğin, k . örnek fonksiyonun ortalama değerini (μ_x) ve otokorelasyon fonksiyonunu [$R_x(\tau, k)$] aşağıdaki gibi hesaplayabiliriz.

$$\mu_x(k) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) dt \quad (1.18)$$

$$R_x(\tau, k) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) x_k(t + \tau) dt \quad (1.19)$$

Random proses $\{x(t)\}$ stasyonerse ve $\mu_x(k)$, $R_x(\tau, k)$ farklı örnek fonksiyonlarda değişmiyorsa, random prosese ergodiktir denir. Bu durumda $\mu_x(k) = \mu_x$ ve $R_x(\tau, k) = R_x(\tau)$ 'dir.

Teorik olarak stasyonerliğin isbatlanması için, ilgili random proses 'in bütün istatistiksel özelliklerinin zaman değişimlerinden bağımsız olması gerekmektedir. Ancak sonsuz sayıda istatistiksel özellikler olabileceğinden pratikte bu özelliklerin hepsinin incelenmesi uygun olmadığı gibi mümkün de değildir. Bundan dolayı random verinin çoğunda geçerli olan belirli önemli tahminlerin incelenmesi gerekir.

Birinci önemli tahmin: Küçük zaman aralıklarında hesaplanan istatistiksel özelliklerin aralıklara göre, önemli bir şekilde (istatistiksel örnekleme dönüşümlerine tesir edecek şekilde) değişip değişmediğidir. Eğer istatistiksel özellikler aralıklara göre önemli ölçüde değişmiyorsa random veri stasyonerdir, aksi halde nonstasyonerdir. Bu tahmini yaparken bir grup örnek kaydın davranışlarını incelemek yerine tek tek örnek kayıtları incelemek stasyonerliği inceleme bakımından daha faydalıdır.

İkinci önemli tahmin ise zayıf stasyonerliğin (ortalama değer ve otokorelasyon fonksiyonunun zamana göre değişmez) olup olmadığıdır. Bu tahmin geçerli olarak kabul edilirse stasyonerliğin isbatı için verinin sadece ortalama değer ve otokorelasyon fonksiyonlarının bilinmesi gerekir.

Üçüncü önemli tahmin: Örnek fonksiyonun, uzun süreli eğilimlerinin (trends) belirlenmesine imkan verecek derecede uzun olması gerekir.

Bir değişkenin birbirini takip eden zaman birimlerinde aldığı değerlerin sıralanışına zaman serisi denir. Zaman serisinin uzun dönemde hâkim eğilimine trend denir. Ekonomik konularla ilgileniyorsak şu unsurlara dikkat etmemiz gerekir. Nüfus artışı, teknolojik değişme, sermaye stokunun büyümesi, zevk ve tercihlerde ve tüketim kalıplarındaki değişiklikler gibi faktörlerin etkisi ile seri uzun dönemde artma veya azalma eğilimi gösterebilir. Bu eğilimin ortaya çıkarılabilmesi için zaman serisinin oldukça uzun bir dönemde izlenmesi gerekir. Örneğin, stok market fiyatlarında nonstasyonerliğin olup olmadığını araştıralım. Sadece bir haftalık fiyatlarla ilgili fonksiyonlara bakarak uzun süreli trendler hakkında karar verilemez. Her bir problem için özel durumlu fonksiyonlar varsa bu tahmin yapılır. Böylece, kısa süreli ortalamaların verinin özelliklerine doğru şekilde yansımaya gayret gösterilir. Eğer bu tahmin yapılmıyorsa burada bahsedilen stasyonerlik testleri yapılamaz ve verinin topluluk ortalamaları (ensemble averages) cinsinden incelenmesi gerekir.

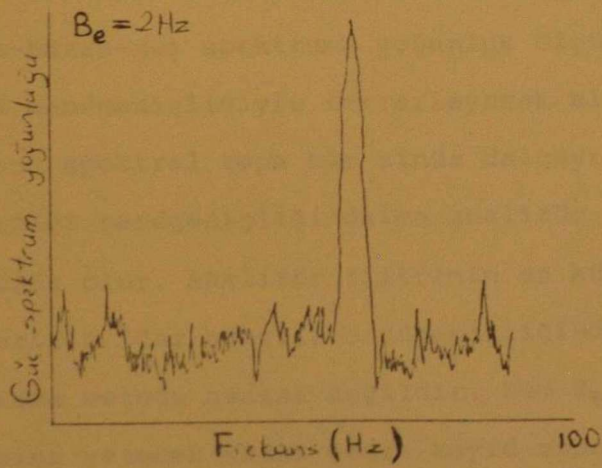
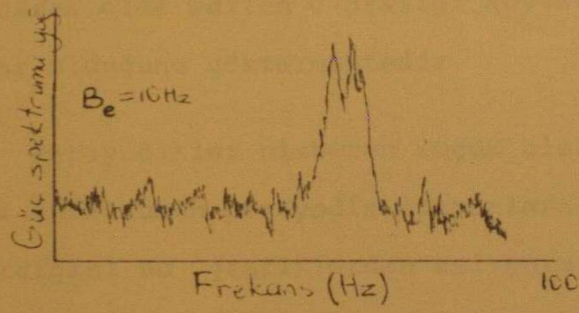
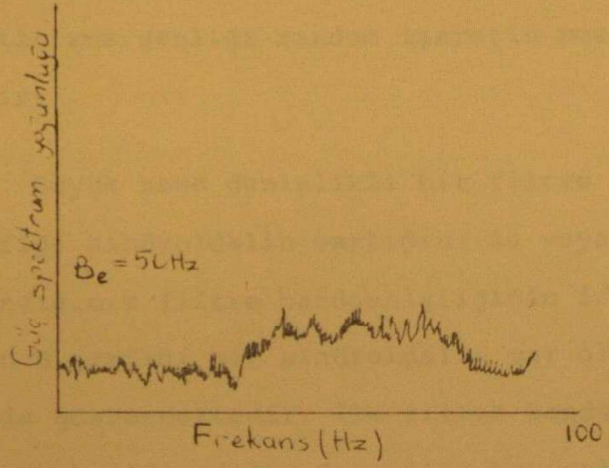
Dördüncü önemli tahmin: Verinin karesel ortalama değeri stasyonerse, otokorelasyon fonksiyonu (veya otokovaryans fonksiyonu) da stasyonerdir. Bu tahmin yukarıdaki üç tahmin kadar güvenilir değildir. Karesel ortalama değer $\tau = 0$ 'da otokorelasyona eşit olduğundan (Varyans $\tau = 0$ 'da otokovaryansa eşittir), genellikle karesel ortalama değere (varyansa) daha çok dikkat edilir.

Stasyonerliği bulmak için verinin frekans band genişliğini bilmek, gözlem zamanı boyunca ortalama ve karesel ortalama değerlerini ölçmek gerekli değildir.

1.6. Verinin Random Olup Olmadığını Anlamak için Yapılan Test

Eğer veri stasyonerlik testinden geçmişse, bu yapacağımız testde sadece veride periyodik veya hemen hemen periyodik elemanların bulunmasına sebep olan sinüzoidallerin olup olmadığına bakılır. Veride sinüzoidal elemanlar çoksa periyodik veya hemen hemen periyodik elemanların olacağı aşıkardır. Ancak random veride düşük yoğunlukta sinüzoidal eleman varsa periyodik elemanların varlığı şüphelidir. Pratikte genellikle stasyoner veriyle ilgili güç spektrumu yoğunluk fonksiyonu, genlik olasılık yoğunluk fonksiyonu ve otokorelasyon fonksiyonu gözle incelenerek verinin random olup olmadığı araştırılır.

Güç spektrumunun random verideki sinüzoidal elemanların varlığını nasıl gösterdiğini aşağıdaki şekle bakarak anlayabiliriz.



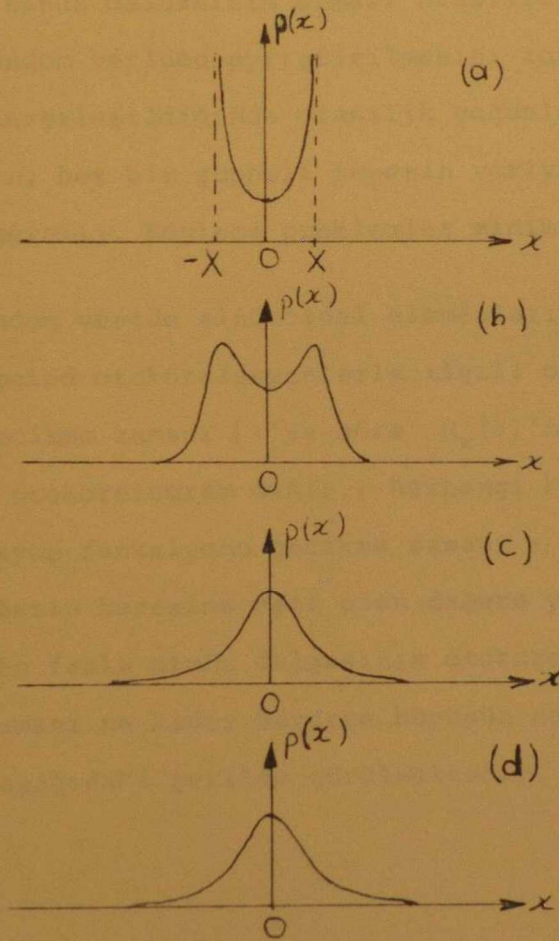
Şekil 11: Gerçek güç spektrumu grafikleri

Şekil 11'deki örnek fonksiyonda random gürültü jeneratörünün çıkışı sinüzoidal işaretle karıştırılır. Sinüzoidal işaretin rms genliği random işaretin rms genliğinin $1/20$ 'ine eşittir.

Büyük band genişlikli bir filtre kullanarak elde edilen A grafiği sinüzoidalın varlığını az veya hiç göstermez. Burada kullandığımız filtre bandgenişliğinin $1/5$ 'ini kullanarak elde edilen B grafiği bir sinüzoidalın var olabileceğini açık bir şekilde göstermektedir. İlk filtre bandgenişliğinin $1/25$ 'ini kullanarak elde edilen C grafiği kuvvetli bir şekilde sinüzoidalın var olduğunu göstermektedir.

Periyodikler nisbeten küçük olsalar bile, yüksek derecedeki spektrumun periyodik elemanları keskin tepeler şeklinde gösterdiğini bu grafiklerden anlıyoruz. Bununla beraber, bir keskin tepe dar bantlı bir random veriyi gösterebilir. Bu iki durumu-bazen-güç spektrumu yoğunluk ölçümünü en dar analizör filtre bandgenişliğiyle tekrarlayarak birbirinden ayırt ederiz. Ölçülen spektral tepe bir sinüs dalgayı gösterirse, tepenin gösterdiği bandgenişliği daima analizör filtrenin bandgenişliğine eşit olur. Analizör filtrenin en küçük bandgenişliği random verinin (dar bantlı) bandgenişliğinden daha küçük olmadıkça, bu ayırma metodu hassas değildir. Hem $B_e Tr \geq 1$ şartının sağlanmasına yetecek kadar örnek kayıd uzun olmalıdır. Aksi takdirde, verinin spektral ölçmelerinde sahte tepeler vardır ve bu tepeler sinüs dalgasının meydana getirdiği tepeleri gözden saklar.

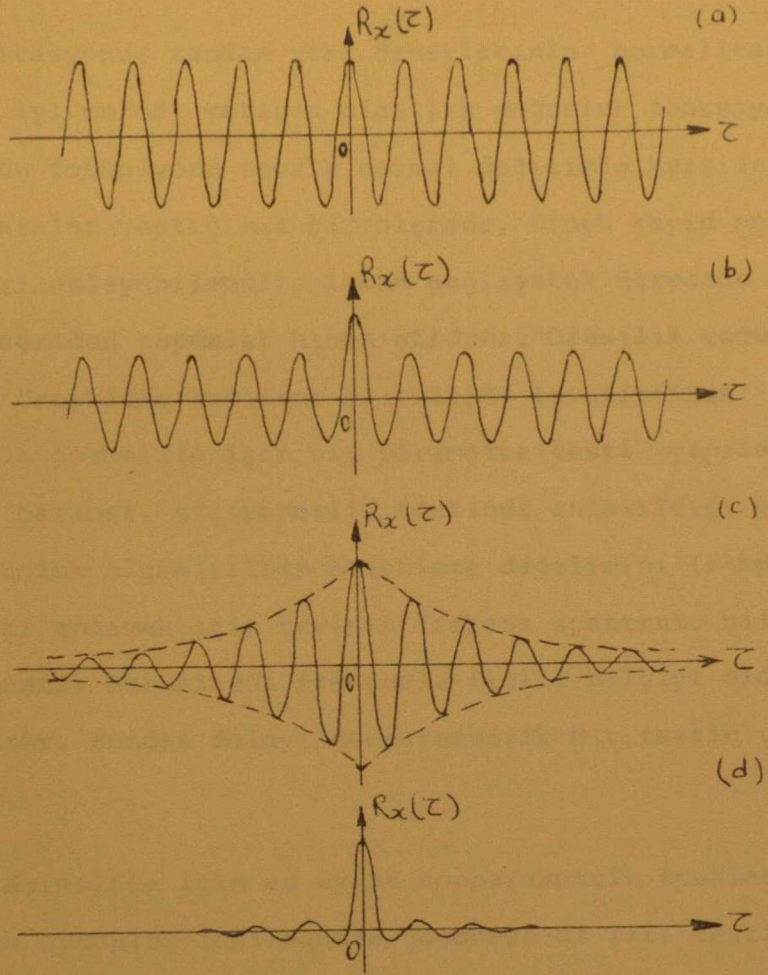
Random verideki sinüzoidal elemanların varlığı verinin genlik olasılık yoğunluk fonksiyonlarıyla da anlaşılabilir. Çünkü aşağıdaki şekilde gösterildiği gibi sinüzoidal ve random verinin olasılık yoğunluk fonksiyonlarının grafikleri birbirinden farklıdır.



Şekil 12: Olasılık yoğunluk fonksiyonları a) Sinüs dalgası, b) Random gürültülü sinüs dalgası, c) Dar-bandlı random gürültü, d) Geniş-bandlı random gürültü.

Genellikle random işaretin çan şeklindeki Gauss karakteristiğine benzeyen olasılık yoğunluk fonksiyonu vardır. Ancak, sinüs dalgasının tabak şeklinde olasılık yoğunluk fonksiyonu da vardır. Sinüzoidal ve random verinin karışımları önemli karakteristiklerde görülür. Eğer sinüs dalgasının rms değeri, verinin random kısmının rms değerinin 2 katından daha az olursa sinüs dalgasının varlığını tesbit etmek çok güçtür. Ayrıca birden fazla sinüs dalgasının olması olasılık yoğunluk fonksiyonlarının random veriden ayrıştırılmasını zorlaştırır. Bu problemlerle karşılaşıldığında olasılık yoğunluk analizini yapabilmek için, her bir şüpheli tepenin veriyi filtreleyerek izole edilmesi gerekir. Böylece problemler minimuma indirilmiş olur.

Random veride sinüzoidal elemanların varlığını araştıran en güçlü metod otokorelasyonlarla ilgili olanıdır. (Otokorelasyonunun gecikme zamanı τ 'ye göre $R_x(\tau)$ 'nin τ 'ya göre] olan grafiğine otokorelogram denir.) Herhangi bir saf random verinin otokorelasyon fonksiyonu gecikme zamanı τ büyürken daima ortalama değerin karesine eşit olan değere yükselir; fakat bir veya birden fazla sinüs dalgasının otokorelasyon fonksiyonu gecikme zamanı ne kadar büyürse büyüsün salınımına devam eder. Bu durum aşağıdaki şekilde görülmektedir.



Şekil 13: Otokorelogramlar, a) Sinüs dalgası, b) Random gürültülü sinüs dalgası, c) Dar-bandlı random gürültü, d) Geniş-bandlı random gürültü.

1.7. Normal Dağılıma Uygunluğu (Normaliteyi) Anlamak için Yapılan Testler

Stasyonær random veri örneklerinin normalitesini bulmak için en iyi metod, verinin olasılık yoğunluk fonksiyonunu ölçmek ve bu fonksiyonu teorik normal dağılımla karşılaştırmaktır. Küçük hatalar yaptığımız bir ölçmede, örnek kayıd normalitedeki sapmaları anlayabilmemize imkan sağlayacak derecede uzun olursa, normal eğriden sapmalar hemen görünür. Olasılık yoğunluk tahmininin örnekleme dağılımı bilinirse, istatistiksel hata büyük olsa bile normalite için bir parametre testi yapılabilir. Bununla beraber, stasyonærlik testinde anlatıldığı gibi olasılık yoğunluk ölçmelerinin örnekleme dağılımını (sampling distribution) anlamak için verinin frekans spektrumu bilinmelidir. Ancak pratik durumlarda frekansla ilgili bilgiyi elde etmek çok güçtür. Bundan dolayı nonparametrik bir testin yapılması gerekir.

Normalite için en uygun nonparametrik testlerden biri chi-kare uygunluk (chi-square goodness-of fit) testidir. Bu test random verinin ayrık (discrete) örnek değerlerine uygulanmalıdır. Veri sayısal tekniklerle analiz edildiğinde hiç bir problem yoktur. Çünkü veri halen ayrık örnek değerlerden oluşur. Fakat analog veri analizinde verinin testde kullanılması için sayısal şekle dönüştürülmesi gerekir.

Örneklenmiş verinin olasılık yoğunluk fonksiyonunun teorik yoğunluk fonksiyonuyla eşdeğer olup olmadığını anlamak için kullanılan özel tip hipotez tezine chi-kare uygunluk testi

denir. Genel prosedür, gözlenen olasılık yoğunluk fonksiyonu ile teorik yoğunluk fonksiyonu arasındaki farkı ölçen ve yaklaşık chi-kare dağılımı olan bir istatistiğin kullanılmasını gerektirir. Bu istatistiğin örneklem dağılımı incelenerek eş-değerlilik testini gerçekleştiririz.

Yapılan işlemi daha iyi anlayabilmek için, olasılık yoğunluk fonksiyonu $p(x)$ olan $x(k)$ random değişkeninin N bağımsız gözlemi olan bir örnek kaydı göz önüne alalım. Gözlemleri sınıf aralıkları denen K tane aralığa ayıralım. Bu aralıklar tekrarlanma histogramını (frequency histogram) oluştururlar. i . sınıf aralığındaki gözlem sayısına i . sınıfın gözlenen tekrarlanma sayısı (observed frequency) denir ve f_i ile gösterilir. $x(k)$ 'nin gerçek olasılık yoğunluk fonksiyonu $p_0(x)$ 'e göre i sınıfta olması beklenen gözlem sayısına ise i . sınıf aralığının teorik tekrarlanma sayısı (expected frequency) denir ve F_i ile gösterilir. Her bir sınıf aralığının gözlenen tekrarlanma sayısı ile teorik tekrarlanma sayısı arasındaki fark $(f_i - F_i)$ ile hesaplanır. Örnek kayıdla ilgili istatistiği elde edebilmek için bütün sınıf aralıklarındaki toplam farkın ölçülmesi gerekir. Bundan dolayı her bir aralıktaki farkların karesi aralığın teorik tekrarlanma sayısına bölünür ve hepsi toplanır.

$$\chi^2 = \sum_{i=1}^K \frac{(f_i - F_i)^2}{F_i} \quad (1.20)$$

(Bu formülden de anlayabileceğimiz gibi; "chi-kare", bir miktarın teorik değeri ile gözlenen değeri arasındaki farkın karesini teorik değere bölerek elde edilen miktarların toplamına denir.)

$z_1(k), z_2(k), z_3(k), \dots, z_N(k)$ adet bağımsız random değişken ve bu değişkenlerin normal dağılımının ortalama değeri sıfır ve varyansı bir ($\mu = 0, \sigma^2 = 1$) olsun. Şimdi yeni bir random değişken olan x_n^2 'yi tarif edebiliriz.

$$x_n^2 = z_1^2(k) + z_2^2(k) + z_3^2(k) + \dots + z_n^2(k) \quad (1.21)$$

x_n^2 random değişkenine, serbestlik derecesi n olan chi-kare değişkeni denir. n serbestlik derecesi, denkleme giren serbest karelerin bağımsız elemanlarının sayısını gösterir. x_n^2 'in olasılık yoğunluk fonksiyonu ve dağılım fonksiyonları aşağıdaki denklemlerle tarif edilir. $x^2 \geq 0$ ise:

$$p(x^2) = [2^{n/2} \Gamma(n/2)]^{-1} (x^2)^{(n/2)-1} e^{-x^2/2} \quad (1.22)$$

$$P(x^2) = \int_{-\infty}^{x^2} p(\xi) d\xi \quad (1.23)$$

x^2 dağılımının 100α yüzdeler noktası (The 100α percentage point of the x^2 distribution) $x_{n;\alpha}^2$ ile gösterilir.

$$\int_{x_{n;\alpha}^2}^{\infty} p(x^2) dx^2 = \text{Olasılık } [x_n^2 > x_{n;\alpha}^2] = \alpha \quad (1.24)$$

x_n^2 değişkeninin ortalama değeri ve varyansı ise şu şekilde bulunabilir.

$$E[x_n^2] = \mu_{x^2} = n \quad (1.25)$$

$$E[(x_n^2 - \mu_{x^2})^2] = \sigma_{x^2}^2 = 2n \quad (1.26)$$

Chi-kare uygunluk testinde bulduğumuz χ^2 ile χ_n^2 birbirine çok yaklaşıktır. Chi-kare uygunluk testinde n serbestlik derecesi, K ve gözlemlere mecburen yüklenen farklı bağımsız lineer sınırlamaların sayısı arasındaki farka eşittir. Sınırlamalardan birisi, baştan itibaren (K-1) sınıf aralıklarındaki frekanslar bilindiğinde son sınıf aralığındaki frekansın bir kez belirlenmesi içindir. Bundan başka en az bir tane daha lineer sınırlamanın olması gerekir. Çünkü beklenen teorik yoğunluk fonksiyonu gözlenen verinin tekrarlanma histogramına uygun olmalıdır. Genellikle, beklenen teorik yoğunluk fonksiyonu normal yoğunluk fonksiyonudur. Bundan dolayı ortalama değer ve varyansın normal yoğunluk fonksiyonuna uygun olabilmesi için 2 lineer sınırlama daha yapılır. Böylece normalite testi için kullanılan chi-kare uygunluk testinde χ^2 'nin serbestlik derecesi $n = K-3$ 'dür.

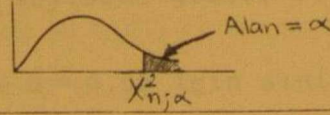
χ^2 'nin uygun serbestlik derecesini tesbit ettikten sonra, random değişken $x(k)$ 'nin olasılık yoğunluk fonksiyonu $p(x) = p_0(x)$ olduğu kabul edilir. Örneklenmiş gözlemleri K adet sınıf aralığına böldükten ve her bir aralığın teorik tekrarlanma sayısını ($p(x) = p_0(x)$ farzederek) hesapladıktan sonra, χ^2 yukarıdaki formül yardımı ile hesaplanır. $p(x)$ 'in $p_0(x)$ 'den sapması χ^2 'nin artmasına sebep olabileceği için bir kenarlı test kullanılır. Yani kabul edilen bölge

$$\chi^2 \leq \chi_{n;\alpha}^2 \quad (1.27)$$

dir. $\chi_{n;\alpha}^2$ Tablodan bulunur. χ^2 'nin örnek değeri $\chi_{n;\alpha}^2$ 'dan büyükse $p(x) = p_0(x)$ hipotezi α seviyesinde red edilir. χ^2 'nin değeri $\chi_{n;\alpha}^2$ 'dan küçük veya eşitse hipotez kabul edilir.

Tablo 1: Chi-kare dağılımının yüzdelik noktaları

Olasılık $[X_n^2 > X_{n;\alpha}^2] = \alpha$ eşitliğinin sağlanabilmesi için gerekli $X_{n;\alpha}^2$ değeri.



α	0.995	0.990	0.975	0.950	0.900	0.10	0.05	0.025	0.010	0.005
1	0.000039	0.00016	0.00098	0.0039	0.0138	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	4.61	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.06	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.75
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.67	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.08	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.67	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.88	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.94	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.58	95.70	100.62	140.23	146.57	152.21	158.95	163.65

$n > 120$ olduğunda $X_{n;\alpha}^2 \approx n \left[1 - \frac{2}{9n} + Z_{\alpha} \sqrt{\frac{2}{9n}} \right]^3$ 'dir

Formüldeki Z_{α} , standartlaştırılmış normal dağılımın istenen yüzdelik noktasıdır.

Chi-kare uygunluk testinin gücü sınıf aralıklarının seçimine bağlıdır. Bu seçimle ilgili değişik pratik ve teorik metodlar mevcuttur. Bunlardan birisi $\alpha = 0.05$ seviyesinde sınıf aralıklarının minimum sayısını gösterir.

Örnek hacmi N ve $\alpha = 0.05$ için sınıf aralıklarının minimum optimum sayısı K ile gösterilir.

N	200	400	600	800	1000	1500	2000
K	16	20	24	27	30	35	39

Şema 3: Örnek hacim N ve $\alpha = 0.05$ için sınıf aralıklarının minimum optimum sayısı.

Testde uygulanabilecek en iyi metod eşit genişlikte sınıf aralıklarını seçmektir. Bu metod, bitişik aralıklarda farklı teorik tekrarlanma sayılarının olmasına sebep olur. Sınıf aralıklarında eşit sayıda teorik tekrarlanma sayısının olması istenir, ancak bunu elde etmek zordur. Fakat genellikle en güçlü testdir. Herhangi bir durumda, her bir sınıf aralığında en azından tekrarlanma sayısının 5 olması istenir. Bununla beraber, sadece uç aralıklarda tekrarlanma sayısının 2 olması kabul edilebilir. Örnek: $x(k)$ random değişkenini chi-kare uygunluk testini kullanarak $N = 200$ bağımsız gözlemi olan bir örnekte $\alpha = 0.05$ seviyesinde inceleyelim. Gözlemlerin nümerik değerlerinin artan derecede düzenlendiğini (aşağıdaki tabloda gösterildiği gibi) varsayalım.

Tablodaki gibi $K = 16$ sınıf aralığı yapalım. Basitlik olması için eşit genişlikli sınıf aralıkları seçmek istersek

tablodan görüldüğü gibi birim genişlikli 16 sınıf aralığı bütün sayıları içerecektir. -6.5; -5.5; ...; 6.5; 7.5; sınıf aralıklarının sınırlarıdır.

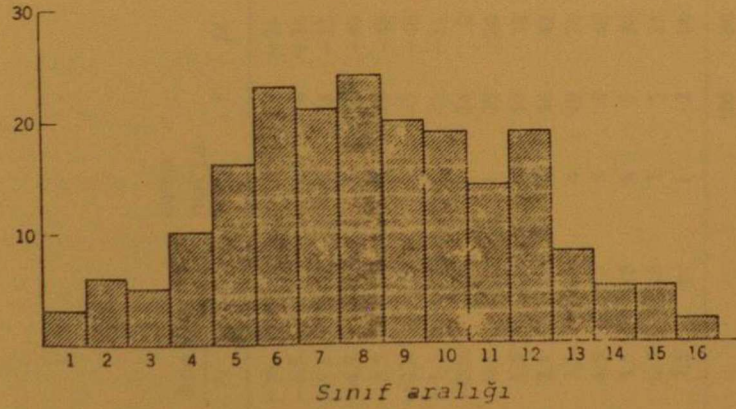
-6.5'dan daha az değerleri 1. sınıf aralığına 7.5'dan daha büyük değerleri son aralığa ve sınırdaki değerleri gelecek aralığa dahil ederek elde edilen tekrarlanma sayısı histogram aşağıdaki şekilde görülmektedir.

Aşağıdaki tabloda testde yaptığımız hesaplamalar görülmektedir. Tablo 1'den $n = 13$ ve $\alpha = 0.05$ için $\chi^2_{n;\alpha}$ 'nin değeri 22.36 ($\chi^2_{13;0.05} = 22.36$) olduğu görülür. Hesaplanan örnek değer χ^2 ise 8.8 'dir ($\chi^2 = 8.8$). Bundan dolayı random değişken $x(k)$ nin normalite hipotezi kabul edilir.

Şekil 2'de G bloku, normalite testinin yapıldığı yeri gösterilmektedir. Genellikle, herhangi bir delil olmamasına rağmen random verinin normal (Gaussian) olasılık yoğunluk fonksiyonuna sahip olduğu farzedilir. Merkezi Limit Teoremi, random etkilerin tesirinde kalan herhangi bir fiziksel olayın normal dağılıma meyilli olduğunu söyler. Olayı belirleyen fiziksel işlemler lineer olduğu sürece bu kabul pratik hayatta geçerlidir. Ancak nonlinear işlemler varsa, ideal Gauss şekliinden sapmalar meydana gelecektir. Bundan dolayı normalite testinin yapılması gerekir.

Tablo 2: Artan derecede düzenlenmiş örnek gözlemler

-7.6	-3.8	-2.5	-1.6	-0.7	0.2	1.1	2.0	3.4	4.6
-6.9	-3.8	-2.5	-1.6	-0.7	0.2	1.1	2.1	3.5	4.8
-6.6	-3.7	-2.4	-1.6	-0.6	0.2	1.2	2.3	3.5	4.8
-6.4	-3.6	-2.3	-1.5	-0.6	0.3	1.2	2.3	3.6	4.9
-6.4	-3.5	-2.3	-1.5	-0.5	0.3	1.3	2.3	3.6	5.0
-6.1	-3.4	-2.3	-1.4	-0.5	0.3	1.3	2.4	3.6	5.2
-6.0	-3.4	-2.2	-1.4	-0.4	0.4	1.3	2.4	3.7	5.3
-5.7	-3.4	-2.2	-1.2	-0.4	0.4	1.4	2.5	3.7	5.4
-5.6	-3.3	-2.1	-1.2	-0.4	0.5	1.5	2.5	3.7	5.6
-5.5	-3.2	-2.1	-1.2	-0.3	0.5	1.5	2.6	3.7	5.9
-5.1	-3.2	-2.0	-1.1	-0.3	0.6	1.6	2.6	3.8	6.1
-4.8	-3.1	-2.0	-1.1	-0.2	0.6	1.6	2.6	3.8	6.3
-4.8	-3.0	-1.9	-1.0	-0.2	0.7	1.6	2.7	3.9	6.3
-4.6	-3.0	-1.9	-1.0	-0.2	0.8	1.7	2.8	4.0	6.5
-4.4	-2.9	-1.8	-1.0	-0.1	0.9	1.8	2.8	4.2	6.9
-4.4	-2.9	-1.8	-0.9	-0.0	0.9	1.8	2.9	4.2	7.1
-4.3	-2.9	-1.8	-0.9	0.0	1.0	1.8	3.1	4.3	7.2
-4.1	-2.7	-1.7	-0.8	0.1	1.0	1.9	3.2	4.3	7.4
-4.0	-2.6	-1.7	-0.8	0.1	1.1	1.9	3.2	4.4	7.9
-3.8	-2.6	-1.6	-0.7	0.2	1.1	2.0	3.3	4.4	9.0



Örnek Gözlemlerin Frekans Histogramı

Tablo 3: Chi-kare uygunluk testini yapabilmek için gerekli hesaplamalar.

Sınıfın numarası	Aralık x_i	Orta nokta x	f	$f x$	x^2	$f x^2$	$z_{i+1} = \frac{x_{i+1} - \bar{x}}{s}$	A^*	$F = N A$	$ F - f $	$\frac{(F - f)^2}{F}$
1	$-\infty$	-7	3	-21	49	147	-2.06	0.0197	3.9	0.9	0.2
2	-6.5	-6	6	-36	36	216	-1.76	0.0195	3.9	2.1	1.1
3	-5.5	-5	5	-25	25	125	-1.45	0.0343	6.9	1.9	0.5
4	-4.5	-4	10	-40	16	160	-1.15	0.0516	10.3	0.3	0.0
5	-3.5	-3	16	-48	9	144	-0.85	0.0726	14.5	1.5	0.2
6	-2.5	-2	23	-46	4	92	-0.54	0.0969	19.4	3.6	0.7
7	-1.5	-1	21	-21	1	21	-0.24	0.1106	22.1	1.1	0.1
8	-0.5	0	24	0	0	0	0.06	0.1187	23.7	0.3	0.0
9	0.5	1	20	20	1	20	0.36	0.1167	23.3	3.3	0.5
10	1.5	2	19	38	4	76	0.67	0.1080	21.6	2.6	0.3
11	2.5	3	14	42	9	126	0.97	0.0854	17.1	3.1	0.6
12	3.5	4	19	76	16	304	1.27	0.0640	12.8	6.2	3.0
13	4.5	5	8	40	25	200	1.58	0.0449	9.0	1.0	0.1
14	5.5	6	5	30	36	180	1.88	0.0270	5.4	0.4	0.0
15	6.5	7	5	35	49	245	2.18	0.0155	3.1	1.9	1.2
16	7.5	8	2	16	64	128	∞	0.0146	2.9	0.9	0.3
			200	60		2184					8.8

$$N = 200 \quad \bar{x} = \frac{1}{N} \sum f x = 0.30 \quad s = \left(\frac{\sum f x^2 - N(\bar{x})^2}{N-1} \right)^{1/2} = 3.30 \quad n = K - 3 = 13 \quad \chi^2 = 8.8$$

* A_1 Standartlaştırılmış normal uygunluk fonksiyonu $p(z)$ 'nin z_i ve z_{i+1} sınırları arasındaki alanına eşittir.

1.8. Ortalama ve Karesel Ortalama Seviye Analizi

H blokuyla gösterilen bu basamakta ortalama değer ve rms değerinin ölçümü yapılır. Aslında, bu basamakda diğer analizlerden ayrı olarak yeni ve önemli bir bilgi elde edemeyiz. Çünkü örnek kaydın ortalama ve rms değerlerinin tahmini, olasılık yoğunluk fonksiyonundan, güç spektral yoğunluk fonksiyonundan, güç spektral yoğunluk fonksiyonundan veya otokorelasyon fonksiyonundan hemen elde edilebilir.

Fiziksel random verinin statik (zamanla değişmeyen) ve dinamik (zamanla değişen) elemanlardan oluştuğu varsayılır. Statik eleman, bütün değerlerin ortalaması olan ortalama değer; dinamik elemanı ise varyans temsil eder.

$$\mu_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt \quad (1.28)$$

$x(k)$ random değişkenin $(-\infty)$ ile $(+\infty)$ arasında değerler aldığını farzedelim. $x(k)$ 'nin ortalama (μ_x) , diğer adıyla beklenen değeri $E[x(k)]$, $x(k)$ 'nin olasılık yoğunluk fonksiyonuyla çarpımının uygun sınırlarla integralini almaya eşittir. Kesikli değişkenler olduğunda $E(X)$, X 'in olası değerlerinin ağırlıklı ortalamasıdır.

$$E[x(k)] = \int_{-\infty}^{\infty} x p(x) dx = \mu_x \quad (1.29)$$

Random değişken, $x(k)$ 'nin sürekli fonksiyonu $g(x)$ ise

$$E[g(x(k))] = \int_{-\infty}^{\infty} g(x) p(x) dx \quad (1.30)$$

Bilhassa, $g(x) = x^2$ ise $x(k)$ 'nin karesel ortalama değeri şu şekildedir.

$$E [x^2(k)] = \int_{-\infty}^{\infty} x^2 p(x) dx = \psi_x^2 \quad (1.31)$$

ψ_x^2 bir random verinin genel yoğunluğunu gösterir.

Dinamik eleman olan varyans, ortalama değer civarındaki karesel ortalama değere eşittir.

$$\sigma_x^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [x(t) - \mu(x)]^2 dt \quad (1.32)$$

Aynı zamanda varyans, $x(k)$ 'nin ortalama değer civarındaki karesel ortalama değerinin beklenen değeri olarak da tarif edilebilir. Burada $g(x) = (x - \mu_x)^2$

$$E[(x(k) - \mu_x)^2] = \int_{-\infty}^{\infty} (x - \mu_x)^2 p(x) dx = \sigma_x^2 \quad (1.33)$$

Bir random değişkenin sadece beklenen değerinin bilinmesi çoğu kere yeterli değildir. Buna ek olarak bu merkezi eğilimden ne ölçüde sapmalar olabileceğini de bilmek gerekir. X random değişkeninin olabilecek değerlerinin beklenen değerden sapmalarının cebrik toplamı sıfır olduğundan, terimlerin beklenen değerden sapmalarının karelerini almak ve bunları x değerlerinin olasılıkları ile çarpıp toplamak dağılımın iyi bir ölçüsünü verir. Yani, varyans, random değişkenin değerleri ile beklenen değer arasındaki farkın karesinin beklenen değeridir.

$$\text{Var}(X) = E [X - E(X)]^2 \quad (1.34)$$

X kesikli random değişken ise

$$\text{Var}(X) = \sum_{i=1}^N [x_i - E(X)]^2 \cdot p_i \quad (1.35)$$

X sürekli random değişken ise

$$\text{Var}(X) = \int_{-\infty}^{\infty} [x - E(x)]^2 f(x) dx \quad (1.36)$$

Varyansın (+) kare köküne standart sapma denir.

Eğer X ve Y random değişkenleri arasındaki ilişkiyi ölçü birimlerine bağlı olarak bilmek istiyorsak kovaryans fonksiyonunu $[(\text{Cov}(X,Y))]$ bilmemiz gerekir.

$$\text{Cov}(X,Y) = E [X-E(X)] [Y-E(Y)] \quad (1.37)$$

X,Y stokastik olarak bağımsız iseler $\text{Cov}(X,Y) = 0$ olur ve aşağıdaki bağıntıyı yazabiliriz.

$$\text{Var} (X + Y) = \text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) \quad (1.38)$$

$$\text{Var}(X_1 + X_2 + X_3 + \dots + X_n) = \sum_{i=1}^n \text{Var} (X_i)$$

Çünkü, X ve Y random değişkenlerinin beklenen değerleri $E(X)$ ve $E(Y)$ ise;

$$E(X + Y) = E(X) + E(Y) \quad (1.39)$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y) \quad (1.40)$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X,Y) \quad (1.41)$$

1.8.1. Ortalama ve Karesel Ortalama Değer Ölçümleri

Stasyonær random zamana baēlı bir gerilim iřaretinin ortalama deēeri μ_x ařaēıdaki formül yardımıyla tahmin edilebilir.

$$\hat{\mu}_x = \frac{1}{T} \int_0^T x(t) dt \quad (1.42)$$

Bu iřlem doēru akım voltmetresiyle yapılabilir. Genellikle dc voltmetresinde basit bir ortalama alma devresi vardır. Ortalama iki řekilde alınabilir. Birincisi; iřareti integral alan devreye, yani geri besleme kondansatörü olan iřlemsel kuvvetlendiriciye göndermektir. İkincisi; iřareti alçak geçiren bir filtreden (genellikle RC tipi) geçirmektir. Bu metoda RC ortalama denir.

iřaretin karesel ortalama deēeri ψ_x^2 ařaēıdaki formül yardımıyla ölçülebilir.

$$\hat{\psi}_x^2 = \frac{1}{T} \int_0^T x^2(t) dt \quad (1.43)$$

ψ_x^2 iki iřlem yardımıyla tahmin edilir. 1) iřaretin ani deēerinin karesini almak, 2) Örnekleme zamanı boyunca karesi alınan ani deēerin ortalaması alınır. rms deēerinin bu tahmini ise ölçülen karesel ortalama deēerin kare köküdür.

Gerilim iřaretinin ortalama deēirini bulmakla iřaretin statik kısmını tesbit etmiř oluyoruz. Varyansı ölçmekle ise iřaretin dinamik kısmını bulmuř oluyoruz. Karesel ortalama

değer ise aşağıdaki formülden de görülebileceği gibi işaretin hem statik hem de dinamik kısımlarıyla ilgili bir ölçümdür.

$$\psi_x^2 = \sigma_x^2 + \mu_x^2 \quad (1.44)$$

Çoğu zaman işaretin statik (ortalama değer) kısmı ve dinamik (varyans) kısmı ayrı ayrı ölçülür.

Örnek kayıd sayısal olduğunda aşağıdaki gibi istatistiksel hesaplamalar yapılabilir. $\{U_n\}$ kümesinin ($n=1,2,\dots,N$) $u(t)$ örnek kaydını $t_n = t_0 + nh$ noktalarında bulunan değerlerini içerdiğini varsayalım. t_0 noktası keyfidir. Noktalar arasındaki mesafe h ile ölçülür ve h sn yardımıyla Nyquist frekansı $f_c = 1/2h$ belirlenir. Orijinal veri değerlerini veren formül ise aşağıdadır.

$$U_n = U(t_0 + nh), \quad n = 1, 2, \dots, N$$

Sayısal olarak verilen örnek kaydın ortalama değeri ise;

$$\bar{u} = \frac{1}{N} \sum_{n=1}^N u_n \quad (1.45)$$

N örnek değerlerinin sayısına, u_n ise örnek değerlerine eşittir. $\bar{u}$, gerçek ortalama değer μ_n 'nin unbiased tahminidir. Bağımsız veri örnekleri için normalize edilmiş standart hata $1/\sqrt{N}$ 'dir.

Formüllerin ve hesaplamaların basitleştirilebilmesi için veri ortalama değerinin sıfır olmasını isteriz. Bundan dolayı

Yeni bir $x(t)$ örnek kaydı tarif edilir $(x(t) = u(t) - \bar{u}) \cdot x(t)$ örnek kaydının aşağıdaki değerleri $\{x_n\}$ prosesini meydana getirirler.

$$x_n = x(t_0 + nh) = u_n - \bar{u} \quad n = 1, 2, \dots, N \quad (1.46)$$

$$\bar{x} = 0 \text{ (ortalama değer)}$$

Örneklerin karesel ortalama değeri aşağıdaki formülle hesaplanır.

$$\bar{x}^2 = \frac{1}{N} \sum_{n=1}^N (x_n)^2 \quad (1.47)$$

Bu formüldeki N , veri örneklerinin sayısına eşittir ve x_n , ortalama değeri sıfır ($\bar{x} = 0$) olan dönüştürülmüş veri değerleridir. Hesapladığımız $\bar{x}^2$ ise gerçek karesel ortalama değer ψ_x^2 'nin biased tahminidir. Bağımsız örnekler için normalize edilmiş standart hata $\sqrt{2/N}$ 'dir.

Örnek kaydın standart sapmasını

$$s = \left[\sum_{n=1}^N \frac{(x_n)^2}{N-1} \right]^{1/2} \quad (1.48)$$

formülüyle hesaplarız. x_n ortalaması sıfır olan ($\bar{x} = 0$) dönüştürülmüş veri değerleridir.

s ve s^2 ise sırasıyla gerçek standart sapma (σ_x) ve gerçek varyansın (σ_x^2) unbiased tahminleridir.

Gerçek karesel ortalama değer ψ_x^2 'nin biased tahmini karesel ortalama değer $\bar{x}^2$, s^2 cinsinden şu şekilde hesaplanır.

$$\overline{x^2} = \left(\frac{N-1}{N}\right) s^2 \quad (1.49)$$

1.8.2. Örnek Değerlerin Hacmi

Örneklem aralığı ($h = \Delta t$) aliasing problemin olmamasını sağlayacak kadar küçük olmalıdır.

$$h = \frac{1}{2f_c} \quad (1.50)$$

$1/f_c$, örnek kayıtdaki en küçük periyottur. Pratikte, daha iyi sonuç elde edebilmek için daha fazla noktadan örnek değer alınmasını isteriz. Meselâ, korelasyon fonksiyonunun f_c 'ye yakın frekansları varsa, hassas korelasyon fonksiyonu ölçümlerinde $h = 1/4f_c$ seçilmelidir. Güç spektrumu ölçümlerinde $h = 2/5f_c$ yeterli olmaktadır. h 'ı küçük seçmenin hesaplanma ve ölçmelerde her ne kadar faydalı tarafları varsa da, ekonomik (maliyet) sebeplerden dolayı $h = 1/2f_c$ 'ye yakın seçilmelidir.

Korelasyon gecikme değerlerinin maksimum sayısı m aşağıdaki formülle bulunur.

$$m = \frac{1}{B_e h} \quad (1.51)$$

Formüldeki B_e , güç spektrumu hesaplamalarında istenen eşdeğer bandgenişliği miktarıdır.

$$B_e = \frac{1}{mh} \quad (1.52)$$

ϵ , spektral hesaplamalarda istenen normalize edilmiş standart hata ve T_r örnek kayıd uzunluğu ise,

$$N = \frac{m}{\epsilon^2} \quad T_r = Nh \quad (1.53)$$

dir. Güç spektrumu hesaplamalarında serbestlik derecesi n ve normalize edilmiş standart hata aşağıdaki formüllerle bulunur.

$$n = 2B_e T_r = \frac{2N}{m} \quad \epsilon = \sqrt{m/N} \quad (1.54)$$

Formülden de görüldüğü gibi m küçüldüğünde ϵ 'da küçülecektir.

Örnek 1: $f_c = 2000$ Hz ve $m = 50$ olduğunda $\epsilon = 0.10$ olması için h, B_e, N ve T_r 'yi hesaplayalım.

$$h = \frac{1}{2f_c} = 0.25 \text{ msn} \quad B_e = \frac{1}{mh} = 80 \text{ Hz} \quad N = \frac{m}{\epsilon^2} = 5000$$

$$T_r = Nh = 1.25 \text{ sn.}$$

Örnek 2 : $f_c = 1000$ Hz ve $\epsilon = 0.10$ olduğunda $B_e = 20$ Hz olması için h, m, N ve T_r 'nin alacağı değerleri bulalım.

$$h = \frac{1}{2f_c} = 0.50 \text{ msn} \quad m = \frac{1}{B_e h} = 100 \quad N = \frac{m}{\epsilon^2} = 10000$$

$$T_r = Nh = 5 \text{ sn.}$$

Örnek 3 : $f_c = 500$ Hz ve $T_r = 20$ sn olduğunda $B_e = 25$ Hz olması için h, m, N ve ϵ 'un alması gereken değerleri hesaplayalım.

$$h = \frac{1}{2f_c} = 1 \text{ msn} \quad m = \frac{1}{B_e h} = 40 \quad N = \frac{T_r}{h} = 2000$$

$$\frac{m}{N} = 0.14$$

1.9. Güç Spektrumu Yoğunluk Analizi

Stasyoner random verinin en önemli karakteristiği, verinin frekans bileşimini gösteren güç spektrumu yoğunluk fonksiyonudur diyebiliriz. Lineer fiziksel sistemlerde, çıkış güç spektrumu sistemin kazanç faktörünün karesi ile giriş güç spektrumunun çarpımına eşittir. Güç spektrumunun altındaki toplam alan karesel ortalama değer ψ^2 'ye eşittir. Yani, herhangi bir frekans bölgesindeki verinin karesel ortalama değeri frekans bölgesinin sınırları içerisinde kalan güç spektrumunun alanına eşittir. Blok I'da gösterilen ve sistemin dinamik karakteristikleriyle ilgili olan güç spektrumu çoğu analizlerde önemli bilgi verir.

Spektral yoğunluk fonksiyonları Fourier dönüşümleri ile tarif edilir. Otokorelasyon fonksiyonlarının Fourier dönüşümleri varsa, $(-\infty, +\infty)$ aralığında frekansları olan iki taraflı (two-sided) spektral yoğunluk fonksiyonları vardır. Bir taraflı spektral yoğunluk fonksiyonunu (fiziksel olarak gerçek olan fonksiyon) bulmak için veri süzülür, sonra karesi ve ortalaması alınır.

Stasyoner random proses $\{x_k(t)\}$ 'nin her bir $x_k(t)$ örnek fonksiyonunun aşağıda gösterildiği gibi bir $X_k(f)$ Fourier dönüşümü vardır.

$$X_k(f) = \int_{-\infty}^{\infty} x_k(t) \cdot e^{-j2\pi ft} \cdot dt \quad (1.55)$$

$$x_k(t) = \int_{-\infty}^{\infty} X_k(f) \cdot e^{j2\pi ft} \cdot df \quad (1.56)$$

Benzer şekilde, $\{y_k(t)\}$ stasyoner random prosessin her bir örnek fonksiyonu $y_k(t)$ 'nin de aşağıdaki gibi kompleks Fourier dönüşümü $Y_k(f)$ vardır.

$$Y_k(f) = \int_{-\infty}^{\infty} y_k(t) e^{-j2\pi ft} dt \quad (1.57)$$

$$y_k(t) = \int_{-\infty}^{\infty} Y_k(f) e^{j2\pi ft} df \quad (1.58)$$

Örnek 1: $x_k(t) = 1$ ise Fourier dönüşümü:

$$1 = \int_{-\infty}^{\infty} x_k(f) e^{j2\pi ft} df ; x_k(f) = \delta(f) = \int_{-\infty}^{\infty} e^{-j2\pi ft} dt$$

$\delta(f)$ 'e delta fonksiyonu denir ve

$$f \neq 0 \quad \text{ise} \quad \delta(f) = 0$$

$$f = 0 \quad \text{ise} \quad \delta(0) = \infty$$

Herhangi bir $\epsilon > 0$ için $\int_{-\epsilon}^{\epsilon} \delta(f) df = 1$ ve $\delta(-f) = \delta(f)$

$F(f)$, herhangi bir fonksiyon ise $\int_{-\epsilon}^{\epsilon} F(f) \delta(f) df = F(0)$

Örnek 2: $x_k(t) = e^{j2\pi f_0 t}$ ise $x_k(f) = ?$

$$e^{j2\pi f_0 t} = \int_{-\infty}^{\infty} x_k(f) e^{j2\pi ft} df$$

$$1 = \int_{-\infty}^{\infty} x_k(f) e^{j2\pi (f-f_0)t} df = \int_{-\infty}^{\infty} x_k(f+f_0) e^{j2\pi ft} df$$

$$x_k(f+f_0) = \delta(f) \Rightarrow x_k(f) = \delta(f-f_0)$$

Örnek 3 : $x_k(t) = x_c(t) = \cos 2\pi f_0 t = \frac{1}{2} (e^{j2\pi f_0 t} + e^{-j2\pi f_0 t})$

$$x_c(f) = \frac{1}{2} [\delta(f - f_0) + \delta(f + f_0)]$$

Örnek 4: $x_k(t) = x_s(t) = \sin 2\pi f_0 t = \frac{1}{2} (e^{j2\pi f_0 t} - e^{-j2\pi f_0 t})$

$$x_s(f) = \frac{1}{2} [\delta(f - f_0) - \delta(f + f_0)]$$

$\{x_k(t)\}$ ve $\{y_k(t)\}$ 'nin ortalama değerleri her t için sıfırsa, Fourier dönüşümleri olan $\{x_k(f)\}$ ve $\{y_k(f)\}$ 'nin de ortalama değerleri her f için sıfırdır.

$$E[X_k(f)] = E[Y_k(f)] = 0 \quad (1.59)$$

$x_k(t)$ gerçel ise $x_k(f)$ nin eşleniği olan $x_k^*(f)$ ile de tarif edilebilir.

$$x_k(t) = \int_{-\infty}^{\infty} x_k^*(f) e^{-j2\pi ft} df ; x_k(t) = \int_{-\infty}^{\infty} x_k(g) e^{j2\pi gt} dg \quad (1.60)$$

$$\begin{aligned} x_k^2(t) &= \int_{-\infty}^{\infty} x_k^*(f) e^{-j2\pi ft} df \cdot \int_{-\infty}^{\infty} x_k(g) e^{j2\pi gt} dg \\ &= \iint_{-\infty}^{\infty} x_k^*(f) x_k(g) e^{j2\pi(g-f)t} dgdf \end{aligned} \quad (1.61)$$

$x_k^2(t)$ 'nin beklenen değeri:

$$E[x_k^2(t)] = \iint_{-\infty}^{\infty} E[x_k^*(f) x_k(g)] e^{j2\pi(g-f)t} dgdf = R_x(0) \quad (1.62)$$

(Bu son denklemde $\{x_k(t)\}$ 'nin stasyonier random proses t 'den bağımsız olduğu kabul edildi).

Güç spektrumu yoğunluk fonksiyonu $S_x(f)$ 'nin beklenen değeri:

$$R_x(0) = E [x_k^2(t)] = \int_{-\infty}^{\infty} S_x(f) df \quad (1.63)$$

$R_x(0)$, $\{x_k(t)\}$ nin karesel ortalama değerinin $(-\infty, \infty)$ frekans aralığındaki dağılım şeklini gösterir. Son iki denklemi karşılaştırırsak $S_x(f)$ 'i elde ederiz.

$$S_x(f) = \int_{-\infty}^{\infty} E [x_k^*(f) x_k(g)] e^{j2\pi(g-f)t} dg \quad (1.64)$$

$$E [x_k^*(f) x_k(g)] = S_x(f) \cdot \delta(f-g) \quad (1.65)$$

$S_x(f)$, random proses $\{x_k(t)\}$ 'nin iki-taraflı güç spektral yoğunluk fonksiyonudur.

$X(f)$, $X(g)$, $Y(f)$ ve $Y(g)$ kompleks random değişkenlerinin stasyonier random prosesdeki korelasyon fonksiyonları aşağıdaki gibi tarif edilir.

$$E [x_k^*(f) x_k(g)] = S_x(f) \delta(f-g) \quad (1.66)$$

$$E [y_k^*(f) y_k(g)] = S_y(f) \cdot \delta(f-g) \quad (1.67)$$

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f\tau} df \quad (1.68)$$

$$R_y(\tau) = \int_{-\infty}^{\infty} S_y(f) e^{j2\pi f\tau} df \quad (1.69)$$

$R_x(\tau)$ ve $R_y(\tau)$ stasyonier korelasyon fonksiyonlardır.

Bu fonksiyonları aşağıdaki gibi basitleştirebiliriz.

$$R_X(\tau) = 2 \int_0^{\infty} S_X(f) \cos 2\pi f \tau df \quad (1.70)$$

$$R_Y(\tau) = 2 \int_0^{\infty} S_Y(f) \cos 2\pi f \tau df \quad (1.71)$$

$R_X(\tau)$ ve $R_Y(\tau)$ korelasyon fonksiyonlarının Fourier dönüşümlerinin olabilmesi için yeterli şart: Hem $|R_X(\tau)|$ hem de $|R_Y(\tau)|$ 'lerinin $(-\infty, \infty)$ aralığında integrallerinin var olmasıdır.

Yani, $R(\tau)$ mevcut ve

$$\int_{-\infty}^{\infty} |R(\tau)| d\tau < \infty \quad (1.72)$$

ise

$$S(f) = \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f \tau} d\tau \quad (1.73)$$

Fiziksel olarak gerçek olan bir-terafli güç spektral yoğunluk fonksiyonları (f sadece $(0, \infty)$ aralığında değişir). $G_X(f)$ ve $G_Y(f)$ aşağıdaki gibi tarif edilir.

$$G_X(f) = 2S_X(f) \leftarrow 0 \leq f < \infty$$

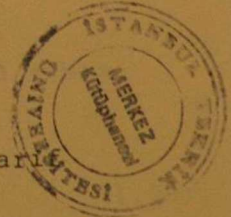
$$G_X(f) = 0 \quad \leftarrow \text{diğer frekanslarda} \quad (1.74)$$

$$G_Y(f) = 2S_Y(f) \leftarrow 0 \leq f < \infty$$

$$G_Y(f) = 0 \quad \leftarrow \text{diğer frekanslarda}$$

$G_X(f)$, $R_X(\tau)$ cinsinden ve $G_Y(f)$, $R_Y(\tau)$ cinsinden tanımlanabilir.

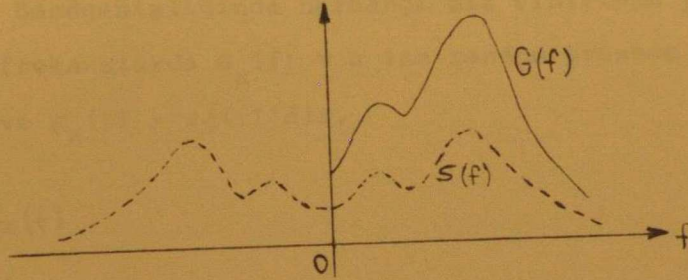
$$G_X(f) = 4 \int_0^{\infty} R_X(\tau) \cos 2\pi f \tau d\tau \quad 0 \leq f < \infty \quad (1.75)$$



$$G_Y(f) = 4 \int_0^{\infty} R_Y(\tau) \cos 2\pi f \tau \, d\tau \quad 0 \leq f < \infty \quad (1.76)$$

$R_X(\tau)$ ve $R_Y(\tau)$ ise sırasıyla $G_X(f)$ ve $G_Y(f)$ cinsinden aşağıdaki gibi tarif edilir.

$$\begin{aligned} R_X(\tau) &= \int_0^{\infty} G_X(f) \cos 2\pi f \tau \, df \\ R_Y(\tau) &= \int_0^{\infty} G_Y(f) \cos 2\pi f \tau \, df \end{aligned} \quad (1.77)$$



Şekil 14: Bir-taraflı ve iki-taraflı spektral yoğunluk fonksiyonu.

1.9.1. Bandgenişliği Sınırlı Beyaz Gürültü

Bandgenişliği sınırlı beyaz gürültü güç spektral yoğunluk fonksiyonu sabit olan bir random prosesdir.

$$\begin{aligned} G_X(f) &= a \quad 0 \leq f_0 - (B/2) \leq f_0 + (B/2) \\ &= 0 \quad \text{diğer frekanslarda} \end{aligned} \quad (1.78)$$

f_0 merkez frekansıdır, B bandgenişliğidir ve a bir sabittir.

$$R_X(\tau) = \int_0^{\infty} G_X(f) \cos 2\pi f \tau \, df \quad (1.79)$$

formülünü kullanarak bandgenişliği sınırlı beyaz gürültünün otokorelasyon fonksiyonunu bulabiliriz.

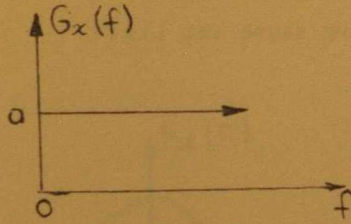
$$R_x(\tau) = \int_{f_0 - (B/2)}^{f_0 + (B/2)} a \cos 2\pi f \tau df = aB \left(\frac{\sin \pi B \tau}{\pi B \tau} \right) \cos 2\pi f_0 \tau \quad (1.80)$$

$$f_0 = B/2 \text{ ise } G_x(f) = a \quad 0 \leq f \leq B \quad (1.81)$$

$$= 0 \quad \text{diğer frekanslarda}$$

$$R_x(\tau) = aB \left(\frac{\sin 2\pi B \tau}{2\pi B \tau} \right) \quad (1.82)$$

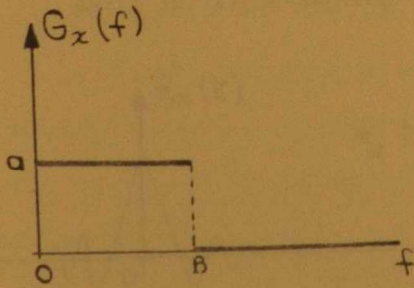
Bandgenişliğinde herhangi bir sınırlama yoksa, yani bütün frekanslarda $G_x(f) = a$ ise random prosese beyaz gürültü denir ve $R_x(\tau) = a\delta(\tau)$ 'dir.



$$f \geq 0 \text{ için } G_x(f) = a$$

$$\text{diğer frekanslarda } G_x(f) = 0$$

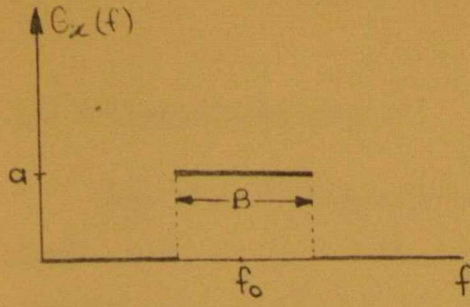
Şekil 15: Beyaz gürültünün bir-terafli güç spektral yoğunluk fonksiyonu



$$0 \leq f < B \text{ için } G_x(f) = a$$

$$\text{diğer frekanslarda } G_x(f) = 0$$

Şekil 16: Alçak-geçiren beyaz gürültünün güç spektral yoğunluk fonksiyonu.

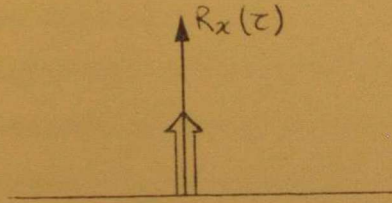


$0 < f_0 - (B/2) \leq f \leq f_0 + (B/2)$ için

$G_x(f) = a$; diğer frekanslarda

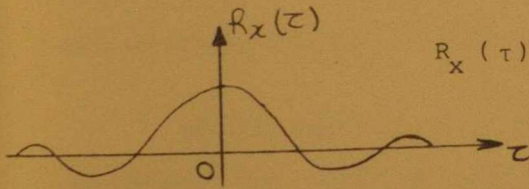
$G_x(f) = 0$ 'dır.

Şekil 17: Band-geçiren beyaz gürültünün bir taraflı güç spektral yoğunluk fonksiyonu.



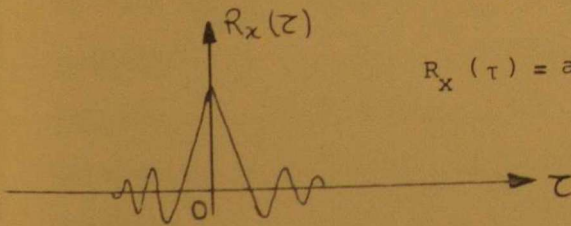
$$R_x(\tau) = a\delta(\tau)$$

Şekil 18: Beyaz gürültünün otokorelasyon fonksiyonu.



$$R_x(\tau) = aB \left(\frac{\sin 2\pi B\tau}{2\pi B\tau} \right) \quad (1.83)$$

Şekil 19: Alçak-geçiren beyaz gürültünün otokorelasyon fonksiyonu.



$$R_x(\tau) = aB \left(\frac{\sin \pi B\tau}{\pi B\tau} \right) \cos 2\pi f_0 \tau \quad (1.84)$$

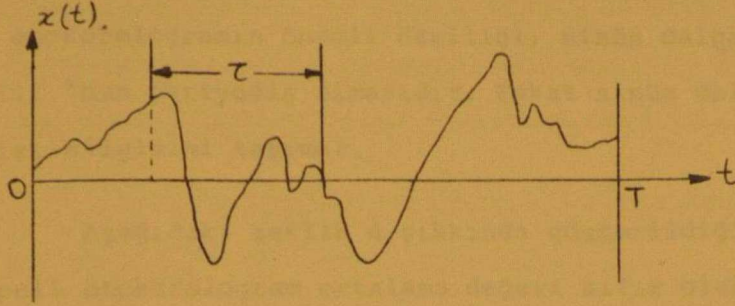
Şekil 20: Band-geçiren beyaz gürültünün otokorelasyon fonksiyonu.

1.10. Otokorelasyon Analizi

Blok 1'de otokorelasyon analizi yapılır. Stasyon verinin otokorelasyon fonksiyonu güç spektral yoğunluk fonksiyonunun ters Fourier dönüşümüdür. Otokorelasyon fonksiyonunun tesbiti güç spektrumuyla ilgili - teknik bakımdan - yeni bir bilgi vermez. Ancak bazı durumlarda, otokorelasyon fonksiyonu güç spektral yoğunluk fonksiyonuna göre daha uygun biçimde istenen bilgiyi saklar. Meselâ, random veride sinüzoidal elemanların varlığını araştırmak istediğimizde otokorelasyonun fonksiyonu daha faydalıdır. Çünkü, güç spektrumunda yorumlayamadığımız sinüzoidal elemanlar otokorelasyon analizi vasıtasıyla daha hızlı teşhis ederiz. Otokorelasyon analizi F ve P bloklarında yapılan işlemler bakımından da çok önemlidir. Otokorelogramın tepe değeri $|R_x(0)|$ karesel ortalama değer ψ^2 'ye eşittir.

Random verinin otokorelasyon fonksiyonu, iki ayrı zamandaki veri değerlerinin birbirlerine olan bağımlılığını inceler. Aşağıdaki şekilde gösterilen örnek kaydın t ve $(t + \tau)$ zamanlarındaki değerleri arasındaki otokorelasyonun tahmini, bu değerleri çarparak $(0, T)$ zaman aralığında ortalamasını almakla elde edilir. T sonsuza yaklaşırken, otokorelasyon tahmini de gerçek otokorelasyon fonksiyonuna yaklaşır.

$$R_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) \cdot x(t+\tau) dt \quad (1.85)$$



Şekil 21: Otokorelasyon ölçümü.

$R_x(\tau)$ fonksiyonu daima $\tau = 0$ 'da maksimumu olan gerçek değerli çift fonksiyondur ve pozitif veya negatif olabilir.

$$R_x(-\tau) = R_x(\tau) \quad (1.86)$$

Bütün τ değerleri için $R_x(0) \geq |R_x(\tau)|$ bağıntısı geçerlidir. τ değeri çok büyükse, $x(t)$ 'nin ortalama değeri (μ_x) otokorelasyonun pozitif kare köküne eşittir.

$$\mu_x = +\sqrt{R_x(\infty)} \quad (1.87)$$

$x(t)$ 'nin karesel ortalama değeri ψ_x^2 ise $R_x(0)$ 'a eşittir

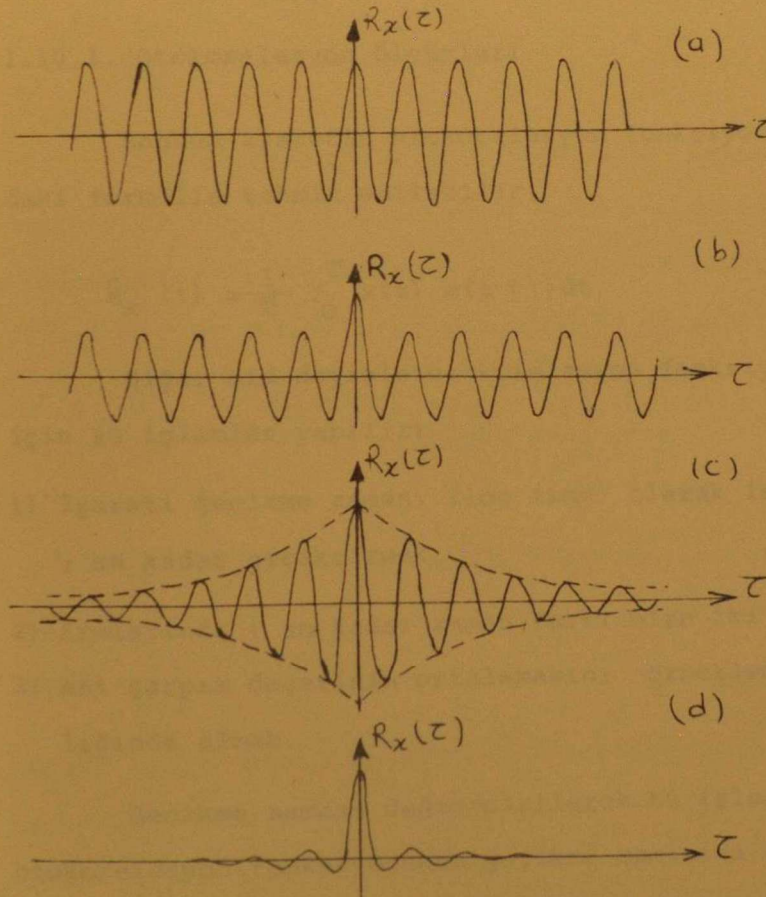
$$\psi_x^2 = R_x(0) \quad (1.88)$$

Otokorelasyonun gecikme süresi (τ)'ya göre olan grafiğine otokorelogram denir. Sinüs dalgasının otokorelasyon fonksiyonu aşağıdaki formülle hesaplanır.

$$R_x(\tau) = \frac{x^2}{2} \cos 2\pi f_0 \tau \quad (1.89)$$

Bu otokorelogramın önemli özelliği, sinüs dalgası gibi $R_x(\tau)$ 'nın periyodik olmasıdır. Fakat sinüs dalgasının faz açısı bilgisini taşımaz.

Aşağıdaki şeklin d şikkında gösterildiği gibi, keskin tepeli otokorelogram ortalama değeri sıfır olan geniş-band random verinin tipik bir örneğidir (Eğer ortalama değeri sıfır değilse, otokorelogram μ_x^2 değerine yaklaşır). Beyaz gürültünün, yani her frekansta ayrı enerjiye sahip olan random verinin otokorelogramı $\tau = 0$ 'da bir Dirac delta fonksiyonudur.



Şekil 22: Otokorelasyon fonksiyon grafikleri (otokorelogramlar)
 (a) Sinüs dalgası. (b) Sinüs dalgası + random gürültü
 (c) Dar-band random gürültü (d) Geniş-band random gürültü.

(b) şıkında gösterilen random gürültülü sinüs dalgasının otokorelogramı, sinüs dalgası otokorelogramı ile random gürültü otokorelogramlarının toplamına eşittir.

$x(t)$ 'nin $R(t_1, t_2)$ otokorelasyonu random değişken $x(t_1)$ ve $x(t_2)$ 'nin ortak momentidir.

$$R_{xx}(t_1, t_2) = R(t_1, t_2) = E\{x(t_1) x(t_2)\} =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \cdot f(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (1.90)$$

1.10.1. Otokorelasyon Ölçümleri

Random işaretin otokorelasyon fonksiyonu $R_x(\tau)$ aşağıdaki formülle tahmin edilebilir.

$$\hat{R}_x(\tau) = \frac{1}{T} \int_0^T x(t) x(t+\tau) dt \quad (1.91)$$

Diğer bir deyişle otokorelasyon fonksiyonunu bulmak için şu işlemler yapılır:

- 1) İşareti gecikme zamanı (lag time) olarak isimlendirilen τ sn kadar geciktirmek.
- 2) Aralarında τ sn kadar zaman farkı olan iki işareti çarpmak.
- 3) Ani çarpım değerinin ortalamasını örnekleme zaman aralığında almak.

Gecikme zamanı değiştirilerek bu işlemler yapılırsa, otokorelasyon fonksiyonunun gecikme zamanına göre (otokorelogram) grafiği elde edilir.

$$T = \frac{1}{2 B_e^2} \frac{R_x^2(0)/R_x^2(\tau)}{2(1000)(0,01)} = \frac{101}{2(1000)(0,01)} = 5,05 \text{ sn.}$$

Demek ki, en küçük korelasyon değerleri maksimum korelasyon değerinin %10'nu olduğunda sonuçlar %10 hatadan daha az hatalı olabilmesi için örnek kaydın uzunluğu en azından 5 sn olmalıdır.

Random veriyi sayısal değerlerle ifade etmek istediğimizde otokorelasyon fonksiyonunu şu şekilde bulabiliriz. Ortalama değeri $\tilde{x} = 0$ olan dönüştürülmüş stasyonere $x(t)$ örnek kaydının N tane veri değeri $\{x_n\}$ prosesini oluşturur ($n = 1, 2, \dots, N$). rh aralıklı tahmin edilen otokorelasyon fonksiyonu aşağıdaki formülle hesaplanır.

$$\hat{R}_r = \hat{R}_x(rh) = \frac{1}{N-r} \sum_{n=1}^{N-r} x_n x_{n+r} \quad r = 0, 1, 2, \dots, m. \quad (1.93)$$

r gecikme sayısıdır, m ise maksimum gecikme sayısıdır. $\hat{R}_r$ ise gecikme sayısı r olduğunda gerçek R_r değerinin tahminidir. Maksimum gecikme sayısı m , $(0, f_c)$ frekans aralığındaki güç spektral yoğunluk fonksiyonunun en küçük eşdeğer frekans bandgenişliği miktarını belirler. Eşdeğer bandgenişliği aşağıdaki gibi hesaplanır.

$$B_e = \frac{2f_c}{m} = \frac{1}{mh} = \frac{1}{\tau_{\max}} \quad (1.94)$$

B_e , teorik frekans aralığı $(-f_c, f_c)$ 'yi m adet eşit parçaya böler. Böylece f_c ve B_e belli ise m 'i bulabiliriz. Maksimum

yer deęiřtirme aralıęı (the maximum displacement) $\tau_{\max} = mh'$ dir. Genellikle, maksimum gecikme sayısı m örnek kaydın hacmi N' nin $0,1$ 'inden daha küçük olması gerekir. Aksi halde, otokorelasyon fonksiyonu tahminlerinde kararsızlıklar meydana gelir.

Otokorelasyon fonksiyonu pozitif deęerler aldıęı gibi negatif deęerler de alabilir. Otokorelasyon fonksiyonunun normalize edilmiř deęeri $\hat{R}_r/\hat{R}_0$ 'dir.

$$\hat{R}_0 = \hat{R}_x(0) = \frac{1}{N} \sum_{n=1}^N (x_n)^2 = \sigma^2 \quad (1.95)$$

$\hat{R}_0$, verinin gerek karesel ortalama deęerinin tahminidir. Varyansı s^2 ile gosterirsek;

$$\hat{R}_0 = \left(\frac{N-1}{N} \right) s^2 \quad (1.96)$$

dir. Buyk N deęerlerinde s^2 ile $\hat{R}_0$ arasındaki fark ihmal edilir. Otokorelasyon fonksiyonunun normalize deęeri ařaęıdaki gibi ifade edilir.

$$-1 \leq \frac{\hat{R}_r}{\hat{R}_0} \leq 1 \quad (1.97)$$

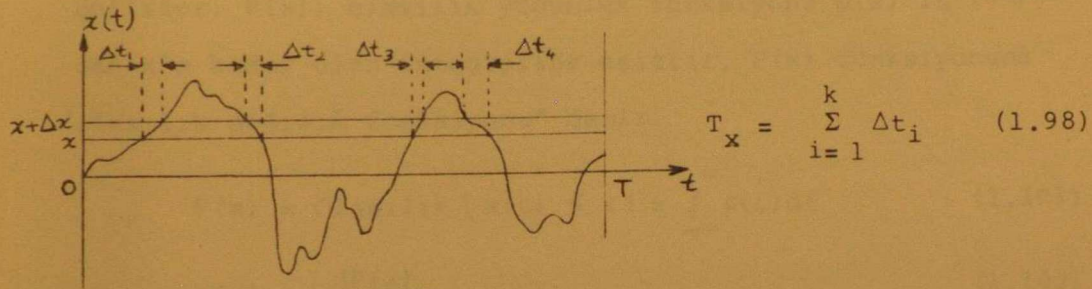
1.11. Olasılık Yoęunluk Fonksiyonunun Analizi

Prosedre dahil edilen son nemli analiz K blokunda gosterildięi gibi olasılık yoęunluk fonksiyonu analizidir. Ancak, btn random olayların normal daęılımı olduęu farz edilirse, olasılık yoęunluk analizi veri analizi prosedrnden ıkarılır. Aslında, bazı durumlarda random veri normali-

teden sapabilir. Bu gibi sapmaları normalite testiyle belirledikten sonra, verinin gerçek olasılık karakteristiklerini belirleyebilmemiz için olasılık yoğunluk fonksiyonunu ölçmemiz gerekir. Verinin karesel ortalama değeri

$\psi^2 = \sigma^2 + \mu^2$ 'dir. Bu formüldeki σ , olasılık yoğunluk fonksiyonunun standart sapmasıdır, μ ise ortalama değerdir.

Olasılık yoğunluk fonksiyonu, random veri $x(t)$ 'nin herhangi bir zamanda belirli bir aralıkta değer alma olasılığını tesbit eder.



Şekil 24: Olasılık ölçümü.

Yukarıdaki şekilde gösterilen örnek kayıd $x(t)$ 'nin, x ve $(x + \Delta x)$ arasında bir değer alma olasılığı (T_x/T) oranıyla bulunur.

$$\text{Olasılık } [x < x(t) \leq x + \Delta x] = \lim_{T \rightarrow \infty} \frac{T_x}{T} \quad (1.99)$$

Formüldeki T_x , $x(t)$ 'nin $(x, x + \Delta x)$ aralığına düştüğü toplam zaman miktarıdır. T ise toplam gözleme süresidir. T sonsuza yaklaşırken bu oran mükemmel olasılık kavramına yaklaşır.

Küçük Δx 'ler için birinci dereceden olasılık yoğunluk fonksiyonu $p(x)$ aşağıdaki gibi ifade edilir.

$$\text{Olasılık } [x < x(t) \leq x + \Delta x] \approx p(x)\Delta x$$

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{\text{olasılık } [x < x(t) \leq x + \Delta x]}{\Delta x} = \lim_{\Delta x \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{T} \left(\frac{T}{\Delta x} \right) \quad (1.100)$$

Olasılık yoğunluk fonksiyonu daima gerçek değerli ve pozitif fonksiyondur. $x(t)$ 'nin ani değerinin herhangi bir x değerine eşit veyadaha küçük olma olasılığını $P(x)$ belirler. $P(x)$, olasılık yoğunluk fonksiyonu $p(x)$ 'in $(-\infty)$ ' dan x 'e kadar olan integraline eşittir. $P(x)$ fonksiyonuna "olasılık dağılım fonksiyonu" denir.

$$P(x) = \text{Olasılık } [x(t) \leq x] = \int_{-\infty}^{\infty} p(\xi) d\xi \quad (1.101)$$

$$p(x) = \frac{dP(x)}{dx} \quad (1.102)$$

$P(x)$ dağılım fonksiyonu 0 ve 1 arasında değerler alır. $x(t)$ nin herhangi bir (x_1, x_2) oralığında değer alma olasılığı:

$$P(x_2) - P(x_1) = \text{Olasılık } [x_1 < x(t) \leq x_2] = \int_{x_1}^{x_2} p(x) dx \quad (1.103)$$

$x(t)$ 'nin ortalama değerini olasılık yoğunluk fonksiyonu cinsinden yazmak istersek:

$$\mu_x = \int_{-\infty}^{\infty} xp(x) dx \quad (1.104)$$

Benzer şekilde karesel ortalama değeri de olasılık yoğunluk fonksiyonu $p(x)$ cinsinden yazabiliriz.

$$\psi_x^2 = \int_{-\infty}^{\infty} x^2 p(x) dx \quad (1.105)$$

1.11.1. Olasılık Yoğunluk Ölçümleri

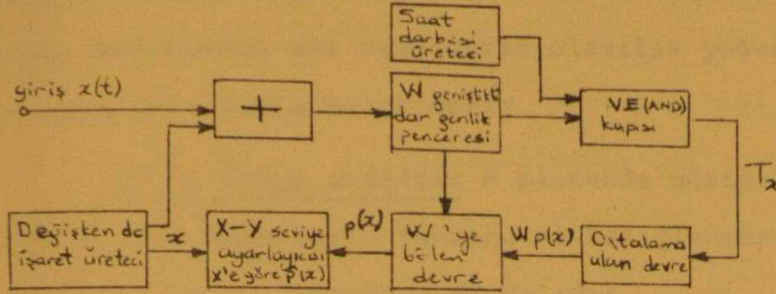
Random işaretin olasılık yoğunluk fonksiyonu $p(x)$ aşağıdaki denklem vasıtasıyla tahmin edilir:

$$\hat{p}(x) = \frac{T_x}{TW} \quad (1.106)$$

Stasyonær random prosesden bir örnek kayıd olan random gerilim işaretimizin olasılık yoğunluk fonksiyonunu tahmin edebilmek için aşağıdaki işlemlerin yapılması gerekir.

- 1) Pencere genişliği W volt olan dar bir genlik penceresiyle işaretin genliğini süzmek.
- 2) Pencere içine giren işaretin harcadığı toplam zamanı ölçmek.
- 3) Pencere içine giren işaretin harcadığı zamanın ortalama oranını bulabilmek için örnekleme zamanı boyunca pencerede harcanan zamanın ortalamasını almak.
- 4) Pencere içinde harcanan zamanın ortalama oranını W genişlikli pencereye bölmek.

Bu işlemler APD analizörü olarak isimlendirilen analog genlik olasılık yoğunluk analizörü ile gerçekleştirilir.



Şekil 25: APD analizörü.

Pencerenin merkez gerilimi hareket ettirilirse, olasılık yoğunluk fonksiyonunun gerilim seviyesine göre grafiği elde edilir. APD analizörü genlik penceresi kullanarak küçük bir gerilim aralığında işaret tarafından harcanan zamanı ölçer. Giriş işaretinin seviyesi pencereye girerse AND kapısı çalışır ve saat darbeleri pencereden geçer. İşaretin seviyesi pencereye girmezse AND kapısı kapanır. Bundan sonra, pencereden geçen işaretin harcadığı zamanın ortalama oranını bulabilmek için, AND kapısından geçen darbelerin toplam örnekleme zamanına göre ortalaması alınır ve dağılım fonksiyonu ($W.p(x)$) elde edilir. Dağılım fonksiyonu da pencere genliği olan W 'ye bölünerek olasılık yoğunluk fonksiyonu $p(x)$ elde edilir.

1.12. Özel Amaçlı Analizler

Özel gayeler için özel amaçlı analizler yapmak çok faydalıdır. Özel amaçlı analizlerden en önemli 3 tanesi Şekil 2'de gösterildi.

1) Tepe değeri dağılımının analizi: L blokunda gösterilen tepe değeri dağılımının analizi, tepe değerlerinin

özelliklêriyle ilgili parametreler hakkında G ve K bloklarında belirlenen ani deęerlerin olasılık yoğunluk fonksiyonuna göre daha fazla bilgi verir.

2) Uç deęer analizi: M blokunda gösterilen uç deęer analizi, uç deęerlerin tahmini ve belirlenmesiyle ilgili problemlerin çözülmesinde faydalıdır.

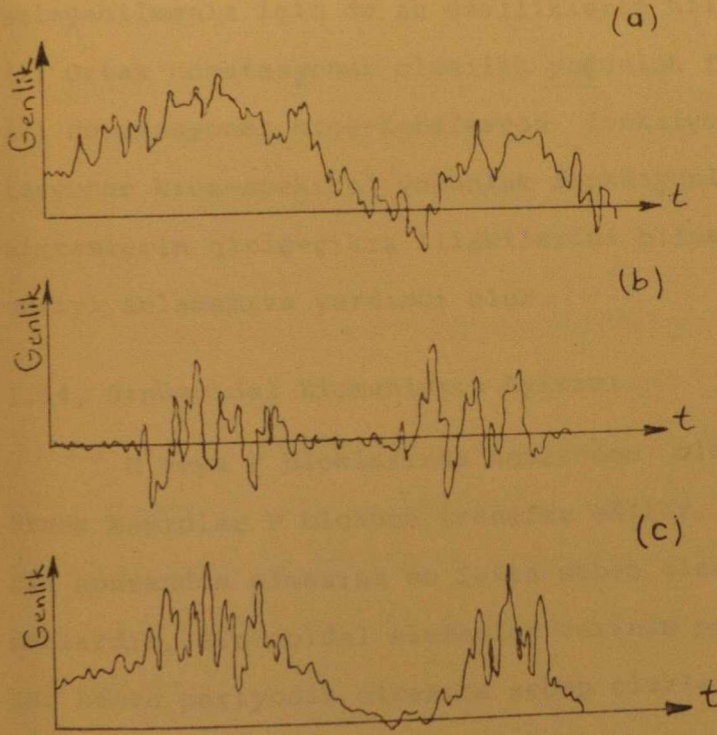
3) Eşik kros analizi (Threshold Crossing Analysis): Çapraz seviyelerin arasındaki ortalama zamanla ilgili problemlerde eşik çaprazların (sıfır çaprazların) analizi faydalı olur.

1.13. Nonstasyonêr Veri Analizi

D ve E bloklarında nonstasyonêr olduęu tesbit edilen örnek kayıtlar özel inceleme için O blokuna gönderilir.

Bundan önce gördüğümüz teorik fikirlerin, proses tekniklerin ve istatistiksel hata formüllerinin çoęu veri nonstasyonêr olduęunda uygulanmaz. Nonstasyonêr veriyi incelemek için özel prosedürler ve teknikler gereklidir. Daha öncede gördüğümüz gibi, nonstasyonêr veride istatistiksel özellikler zamana göre deęişir. Bundan dolayı fiziksel verilerin çoęu nonstasyonêr veridir. Ancak besitlik olması ve yaklaşık ifadeler elde edebilmek için çoęu verilerin stasyonêr olduęu farzedilir. Örneęin; uzun süreli işlemlerde verilen girişin deęişik çıkışlar üretebilmesi için sistem özellikleri deęiştirildiğinde ve kısa süreli işlemlerde çevre şartları aniden deęiştiiğinde nonstasyonêr veri elde edilir.

Nonstasyon verinin bütün tiplerini inceleyecek bir prosedür henüz mevcut değildir. Bundan dolayı nonstasyon verinin özelliklerini kesin ve hassas olarak bilemeyiz. Dolayısıyla nonstasyon veriyi inceleyebilmemiz için, önce nonstasyonluğun en önemli tiplerini bilmemiz gerekir. Aşağıdaki şekilde 3 tane önemli ve temel tip görülmektedir.



Şekil 26: 3 tane nonstasyon verisi, a) Zamanla değişen ortalama değer. (b) Zamanla değişen karesel ortalama değer. (c) Zamanla değişen ortalama ve karesel ortalama değerlerin toplamı (a ve b şiklerinin kombinasyonu)

Bir nonstasyonier prosesin farklı zamanlardaki random verilerin ortak istatistiksel özelliklerini belirleyebilme-miz için prosesin şu özelliklerinin bilinmesi gerekir:

1) Nonstasyonier olasılık yoğunluk fonksiyonu, (2) Nonstasyonier otokorelasyon fonksiyonu (3) Nonstasyonier güç spektral yoğunluk fonksiyonu.

İki adet nonstasyonier prosesin arasındaki ilişkileri anlayabilmemiz için de şu özelliklerin bilinmesi gerekir:

(1) Ortak nonstasyonier olasılık yoğunluk fonksiyonları, (2) Nonstasyonier kros-korelasyon fonksiyonları. (3) Nonstasyonier kros-spektral yoğunluk fonksiyonları. Fiziksel sistemlerin giriş-çıkış ilişkilerini bilmemiz nonstasyonier veriyi anlamamıza yardımcı olur.

1.14. Sinüzoidal Elemanların Ayırımı

D veya F bloklarında nonrandom olduğu tesbit edilen örnek kayıtlar P blokuna transfer edilir. Stasyonier verilerin nonrandom olmasına en fazla sebep olan sinüzoidal elemanlardır. Sinüzoidal elemanlar verinin periyodik ya da hemen hemen periyodik olmasına sebep olurlar. Random kısımlardan sinüzoidal elemanları ayırdıktan sonra tekrar random veriyle ilgili prosedüre döneriz.

P blokunda yapacağımız ilk iş sinüzoidal elemanları tesbit etmektir. Random veride bulunan yüksek yoğunluktaki sinüzoidal elemanların varlığını gözümüzle anlayabiliriz. Fakat, daha az yoğunluktaki sinüzoidal elemanları güç spektrumu analizi veya olasılık yoğunluk analizi ile ilgili

grafiklere bakarak anlayabiliriz. En güçlü deteksiyon prosedürü ise J blokunda gösterilen otokorelasyon analizidir.

Deteksiyon işleminden sonra yapılacak iş sinüzoidal elemanları random veriden ayırmaktır. Bu ayırma işlemi darbantlı reddetme filtreleriyle (narrow-band rejection filters) gerçekleştirilebilir. Reddetme filtresiyle sinüs dalgasını süzerken random elemanın dar frekans bandı da süzülür. Ancak, reddetme filtresi çok dar ise uzaklaştırılan random veri toplam random veriye göre çok küçük ve önemsiz kalır. Verinin etkin şekilde normalite testini yapabilmemiz için verinin random kısmından sinüzoidal elemanları uzaklaştırmak gerekir.

Sinüzoidal elemanları uzaklaştırdıktan sonra örnek kayıd tekrar F blokuna gönderilir. Eğer örnek kaydın tekrar nonrandom olduğuna karar verilirse, periyodik deterministik (transient) veri incelemesini yapmak için örnek kayıd Q blokuna transfer edilir. Q blokunda, random kayıddan uzaklaştırılan periyodik ve hemen hemen periyodik elemanlar analiz edilir.

2. PARAMETRE TAHMİNLERİNDE YAPILAN İSTATİSTİKSEL HATALAR

Random veri olan örnek kayıtların analizinde yapılan hatalar üç ana gruba ayrılır.

a) Tahmin hataları (bazen "istatistiksel hatalar" isimlendirilir).

b) Alet hataları (bazen "kalibrasyon hatalar" isimlendirilir).

c) Kullanım hataları (bazen "şahıs hataları" isimlendirilir).

Önce, değişik örnek değerlerinde yapılan tahmin hatalarını inceleyelim. Tahmin hataları, kullanım ve klasik alet hatalarından daha fazladır. Her bir analiz prosedürünün alet ve kalibrasyon hatası kullanılan aletin bir fonksiyonudur. Benzer şekilde, kullanma hatası da aleti kullanan insanların fonksiyonudur.

$\{x(t)\}$ nin stasyonær random proses olduğunu kabul edelim. Parantezdeki t zamandır veya zamanın yerini alabilecek diğær bir değışkendir. ϕ 'nin random proses $\{x(t)\}$ nin bilinmeyen bir parametrenin gerçæk değıeri olduğunu farzedelim. Örneğın, ϕ belirli bir frekansta ortalama değıer veya güç spektrumu yoğunluğu olabilir. T uzunluklu örnek kayıddá yaptığımız ölçmeden elde edilen ϕ nin tahmini değıeri de $\hat{\phi}$ olsun.

$x(t)$ örnek kaydı, aynısının tekrar edilmesi mümkün olmayan olaylar kümesinden bir eleman temsil ettiğinden dolayı $\hat{\phi}$ de bir random değişkendir. Yani, $\{x(t)\}$ random prosesindeki farklı örnek kayıtlar için hesaplanan $\hat{\phi}$ 'nin değerleri de değişiktir. Burada şu soruya ne derece kesin cevap verdiğimiz önemlidir. Tahmini $\hat{\phi}$ değeri gerçek ϕ değerini ne kadar yaklaşıklıkla temsil ediyor?

Önce, örnekleme zamanı T 'ye bağlı olan $\hat{\phi}$ tahmininin beklenen değerini bulalım. Şüphesiz, $\hat{\phi}$ nin beklenen değerinin gerçek değer ϕ yi vermesini isteriz. Diğer bir deyişle, $\{x(t)\}$ nin farklı örnek kayıtlarından ölçülen $\{\hat{\phi}\}$ tahminler kümesi beklenen değerinin hatasız olarak ϕ yi vermesini isteriz. Bu özelliğe sahip olan parametre tahminine unbiased denir. Yani

$$E[\hat{\phi}] = \phi \quad (2.1)$$

ise bu tahmin unbiased'dir. Eğer $E[\hat{\phi}] \neq \phi$ ise, etkileşim hatası (bias error) aşağıdaki gibi ifade edilir.

$$b[\hat{\phi}] = E[\hat{\phi}] - \phi = E[\hat{\phi}] - E[\phi] = E[\hat{\phi} - \phi] \quad (2.2)$$

Bu formüldeki $b[\hat{\phi}]$ notasyonu, parantezlerdeki tahminin etkileşimi anlamına gelir. Unbiased tahminlerde $b[\hat{\phi}] = 0$ 'dır. Sabit T aralığında $\hat{\phi}$ nin unbiased olması $\hat{\phi}$ nin gerçek değer ϕ 'ye yakın olacağını ima etmez. Gerçek değerden önemli bir sapma meydana gelebilir. Bu sapma T artarken azalabilir. Bu durumdan analiz etmek için sapmanın karesinin beklenen değeri olarak tarif edeceğimiz karesel ortalama hatayı hesaplarız. Şöyleki;

$$\text{Karesel ortalama hata} = E[(\hat{\phi} - \phi)^2] \quad (2.3)$$

T'yi büyütürken karesel ortalama hatanın da sifıra yaklaşmasını isteriz. Böylece herhangi bir $\hat{\phi}$ tahmininin büyük T değerlerinde gerçek değer ϕ 'ye yaklaşma eğiliminin kuvvetli olması gerekir. Bu istenen özelliğe sahip tahminlere consistent'dir denir. Yani,

$$\lim_{T \rightarrow \infty} E[(\hat{\phi} - \phi)^2] = 0 \quad (2.4)$$

ise bu tahmin consistent'dir. Karesel ortalama hatayı aşağıdaki gibi yazabiliriz.

$$\begin{aligned} E[(\hat{\phi} - \phi)^2] &= E[(\hat{\phi} - E[\hat{\phi}] + E[\hat{\phi}] - \phi)^2] \\ &= E[(\hat{\phi} - E[\hat{\phi}])^2] + 2E[(\hat{\phi} - E[\hat{\phi}])(E[\hat{\phi}] - \phi)] \\ &\quad + E[(E[\hat{\phi}] - \phi)^2] \\ E[\hat{\phi} - E[\hat{\phi}]] &= E[\hat{\phi}] - E[\hat{\phi}] = 0 \end{aligned} \quad (2.5)$$

olduğundan $E[(\hat{\phi} - \phi)^2]$ 'yi aşağıdaki gibi ifade edebiliriz.

$$E[(\hat{\phi} - \phi)^2] = E[(\hat{\phi} - E[\hat{\phi}])^2] + E[(E[\hat{\phi}] - \phi)^2] \quad (2.6)$$

(2.6) denklemi, beklenen değer civarındaki karesel ortalama değer ile gerçek değer civarındaki beklenen değer karesinin toplamının gerçek değer civarındaki karesel ortalama değere eşit olduğunu gösterir. Demek ki, karesel ortalama hata iki parçanın toplamıdır. Birinci parça aşağıdaki denklemle gösterildiği gibi tahminin varyansıdır.

$$\text{Var} [\hat{\phi}] = \sigma_{\hat{\phi}}^2 = E[(\hat{\phi} - E[\hat{\phi}])^2] = E[\hat{\phi}^2] - E^2[\hat{\phi}] \quad (2.7)$$

İkinci parça, aşağıdaki denklemde görüldüğü gibi tahminin etkileşim hatasının karesidir. (Bir sabitin beklenen değeri de sabittir).

$$b^2[\hat{\phi}] = E[b^2[\hat{\phi}]] = E[(E[\hat{\phi}] - \phi)^2] \quad (2.8)$$

Genellikle, T büyürken hem varyans hem de etkileşim hatasını sifıra yaklaştırmaya çalışırız. Karesel ortalama hatayı varyans ve etkileşim hatası cinsinden ifade etmek istersek:

$$E[(\hat{\phi} - \phi)^2] = \text{Var}[\hat{\phi}] + b^2[\hat{\phi}] \quad (2.9)$$

2.1. ORTALAMA DEĞER TAHMİNLERİ

Stasyoner (ergodik) random proses $\{x(t)\}$ 'nin örnek kayıdı olan $x(t)$ 'yi inceleyelim. Bu örnek kaydın ortalama değeri aşağıdaki denklemle tahmin edilir.

$$\hat{\mu}_x = \frac{1}{T} \int_0^T x(t) dt \quad (2.10)$$

Gerçek ortalama değer ise

$$\mu_x = E[x(t)] \quad (2.11)$$

dir. ve $\{x(t)\}$ stasyoner olduğu için t'den bağımsızdır. $\hat{\mu}_x$ tahmininin beklenen değeri ise;

$$E[\hat{\mu}_x] = E\left[\frac{1}{T} \int_0^T x(t) dt\right] = \frac{1}{T} \int_0^T E[x(t)] dt = \frac{1}{T} \int_0^T \mu_x dt = \mu_x \quad (2.12)$$

$\hat{\mu}_x$ T'den bağımsızdır ve μ_x 'in unbiased tahminidir.

$\hat{\mu}_x$ unbiased olduğundan, μ_x tahmininin karesel ortalama hatası varyansa eşittir.

$$\text{Var}[\hat{\mu}_x] = E[(\hat{\mu}_x - \mu_x)^2] = E[\hat{\mu}_x^2] - \mu_x^2 \quad (2.13)$$

(2.10) denklemini kullanarak $E[\hat{\mu}_x^2]$ yi hesaplayabiliriz.

$$E[\hat{\mu}_x^2] = \frac{1}{T^2} \int_0^T \int_0^T E[x(\xi) x(\eta)] d\eta d\xi \quad (2.14)$$

Stasyoner random proses $\{x(t)\}$ 'nin otokorelasyon fonksiyonu $R_x(\tau)$ aşağıdaki şekilde ifade edilir.

$$R_x(\tau) = E[x(t) x(t + \tau)] \quad (2.15)$$

$\{x(t)\}$ stasyoner olduğundan $R_x(\tau)$ t'den bağımsızdır.

$R_x(\tau)$ 'nın sürekli, τ 'nin bütün değerleri için sonlu ve girişte $R_x(\tau)$ 'daki bütün periyodik elemanların uzaklaştırılmış olduğu farzedilir. Otokovaryans fonksiyon $C_x(\tau)$ aşağıdaki gibi ifade edilir.

$$C_x(\tau) = R_x(\tau) - \mu_x^2 \quad (2.16)$$

Büyük T değerlerinde

$$\text{Var}[\hat{\mu}_x] \approx \frac{1}{T} \int_{-\infty}^{\infty} C_x(\tau) d\tau \quad (2.17)$$

denklemini yazılabilir, burada $|\tau| \ll T$ 'dir.

Örnek 2.1: Band genişliği sınırlı beyaz gürültü ortalama değer tahminlerinin varyansı.

$\{x(t)\}$, ortalama değeri $\mu_x \neq 0$ olan bandgenişliği sınırlı beyaz gürültü prosesi olsun. Güç spektrum yoğunluk fonksiyonu aşağıdaki gibi ifade edilir.

$$G_x(f) = \frac{1}{B} + \mu_x^2 \delta(f) \quad 0 \leq f \leq B \quad (2.18)$$

$$= 0 \quad f > B$$

Formüldeki B bandgenişliğidir. Otokovaryans fonksiyonu $C_x(\tau)$ ise:

$$C_x(\tau) = \int_0^{\infty} G_x(f) \cos 2\pi f \tau df - \mu_x^2$$

$$= \int_0^B \frac{1}{B} \cos 2\pi f \tau df = \frac{\sin 2\pi B \tau}{2\pi B \tau} \quad (2.19)$$

$C_x(0) = 1$ ve $C_x(1/2B) = 0$ dir. Bu durum gösteriyor ki, aralarındaki mesafe $\frac{1}{2B}$ olan noktalar ilişkisizdir. $C_x(\tau)$ 'yu (2.17) denkleminde yerine koyarsak;

$$\text{Var}[\hat{\mu}_x] \approx \frac{1}{T} \int_{-\infty}^{\infty} \left(\frac{\sin 2\pi B \tau}{2\pi B \tau} \right) d\tau = \frac{1}{2BT} \quad (2.20)$$

elde edilir. Bu denklem, B bandgenişliği ve T örnekleme zamanı olan bandgenişliği sınırlı beyaz gürültünün ortalama değer tahminlerinin varyansıdır. Ancak (2.20) denkleminde T'nin yeterince büyük olması gerekir. Pratikteki tatminkar şartlar $T \geq 10|\tau|$ ve $BT \geq 5$ dir.

$C_x(0) \neq 1$ ise (2.20) denklemi aşağıdaki gibi yazılır.

$$\text{Var}[\hat{\mu}_x] \approx \frac{C_x(0)}{2BT} = \frac{\sigma_x^2}{2BT}$$

Bundan dolayı, $\mu_x \neq 0$ olduğunda normalize edilmiş karesel ortalama hata aşağıdaki denklemle verilir.

$$\epsilon^2 = \frac{\text{Var}[\hat{\mu}_x]}{\mu_x^2} \approx \frac{1}{2BT} \left(\frac{\sigma_x}{\mu_x}\right)^2 \quad (2.21)$$

ϵ 'a tahminin normalize edilmiş standart hatası denir.

2.2. KARESEL ORTALAMA DEĞER TAHMİNLERİ

$x(t)$, stasyoner (ergodik) random proses $\{x(t)\}$ nin zamana bağlı örnek kaydı olsun. $\{x(t)\}$ nin karesel ortalama değeri aşağıdaki gibi T aralığında zaman ortalaması alınarak tahmin edilebilir.

$$\hat{\psi}_x^2 = \frac{1}{T} \int_0^T x^2(t) dt \quad (2.22)$$

Gerçek karesel ortalama değer ise:

$$\psi_x^2 = E[x^2(t)] \quad (2.23)$$

$\{x(t)\}$ stasyoner olduğundan ψ_x^2 t 'den bağımsızdır.
 $\hat{\psi}_x^2$ tahmininin beklenen değeri

$$E[\hat{\psi}_x^2] = \frac{1}{T} \int_0^T E[x^2(t)] dt = \frac{1}{T} \int_0^T \psi_x^2 dt = \psi_x^2 \quad (2.24)$$

dir. $\hat{\psi}_x^2$, T 'den bağımsız olan ψ_x^2 nin unbiased tahminidir.

O halde karesel ortalama varyansa eşittir. Ortalama karesel hata varyansla gösterilir.

$$\begin{aligned} \text{Var} [\hat{\psi}_x^2] &= E [(\hat{\psi}_x^2 - \psi_x^2)^2] = E [\hat{\psi}_x^4] - \psi_x^4 \\ &= \frac{1}{T^2} \int_0^T \int_0^T (E [x^2(\xi) x^2(\eta)] - \psi_x^4) d\eta d\xi \quad (2.25) \end{aligned}$$

2.3. OLASILIK YOĞUNLUK FONKSİYONU TAHMİNLERİ

Burada stasyoner (ergodik) random proses $\{x(t)\}$ nin bir örnek kaydının olasılık yoğunluk ölçümünü inceleyeceğiz. $x(t)$ 'nin T zaman aralığında $(x - \frac{w}{2})$ ve $(x + \frac{w}{2})$ arasındaki genlik değerlerini alma olasılığı aşağıdaki gibi tahmin edilebilir.

$$\hat{P}[x, w] = \text{olasılık} [(x - \frac{w}{2}) \leq x(t) \leq (x + \frac{w}{2})] = \frac{1}{T} \sum_i \Delta t_i = \frac{T_x}{T} \quad (2.26)$$

Formüldeki Δt_i , i . girişin w genlik aralığında $x(t)$ tarafından harcanan zamandır ve $T_x = \sum_i \Delta t_i$ 'dir. T_x/T oranı, $[x - (w/2), x + (w/2)]$ aralığında $x(t)$ tarafından harcanan zamanın gözlem zamanına (T) oranıdır. T_x genellikle x genliğinin bir fonksiyonudur. T sonsuza yaklaşırken tahmin edilen olasılık $\hat{P}[x, w]$ de gerçek olasılık $P[x, w]$ ye yaklaşacaktır. Ayrıca, bu tahmin edilen olasılık gerçek olasılığın unbiased tahminidir. Bundan dolayı,

$$P[x, w] = E[\hat{P}[x, w]] = \lim_{T \rightarrow \infty} \hat{P}[x, w] = \lim_{T \rightarrow \infty} \frac{T_x}{T} \quad (2.27)$$

Birinci dereceden olasılık yoğunluk fonksiyonu $p(x)$ aşağıdaki gibi ifade edilir.

$$p(x) = \lim_{W \rightarrow 0} \frac{P[x, W]}{W} = \lim_{\substack{T \rightarrow \infty \\ W \rightarrow 0}} \frac{\hat{P}[x, W]}{W} = \lim_{\substack{T \rightarrow \infty \\ W \rightarrow 0}} \hat{p}(x) \quad (2.28)$$

$$\hat{p}(x) = \frac{\hat{P}[x, W]}{W} = \frac{T_x}{TW} \quad (2.29)$$

$\hat{p}(x)$ tahmin edilen değerdir.

$x(t)$ örnek kaydının herhangi iki genlik değeri (x_1 ve x_2) arasında değer alma olasılığı, gerçek olasılık yoğunluk fonksiyonu $p(x)$ cinsinden aşağıdaki gibi ifade edilir.

$$\text{Olasılık } [x_1 \leq x \leq x_2] = \int_{x_1}^{x_2} p(x) dx \quad (2.30)$$

$$P[x, W] = \text{Olasılık } \left[x - \frac{W}{2} \leq x(t) \leq x + \frac{W}{2} \right] = \int_{x - (W/2)}^{x + (W/2)} p(\xi) d\xi \quad (2.31)$$

(2.29) denkleminde faydalananarak aşağıdaki denklemi yazabiliriz.

$$E[\hat{p}(x)] = \frac{E[\hat{P}[x, W]]}{W} = \frac{P[x, W]}{W} = \frac{1}{W} \int_{x - (W/2)}^{x + (W/2)} p(\xi) d\xi \quad (2.32)$$

Genellikle $\hat{p}(x)$, $p(x)$ 'in biased tahminidir. Yani,

$$E[\hat{p}(x)] \neq p(x) \quad (2.33)$$

dir.

$\hat{p}(x)$ tahmininin karesel ortalama hatasını (2.9) denklemi yardımıyla hesaplayabiliriz.

$$E[(\hat{p}(x) - p(x))^2] = \text{Var}[\hat{p}(x)] + b^2[\hat{p}(x)] \quad (2.34)$$

$$\text{Var}[\hat{p}(x)] = E[(\hat{p}(x) - E[\hat{p}(x)])^2] \quad (2.35)$$

$b[\hat{p}(x)]$, biased tahminin hatasıdır ve aşağıdaki gibi ifade edilir.

$$b[\hat{p}(x)] = E[\hat{p}(x) - p(x)] \quad (2.36)$$

2.3.1. Tahmin Edilen Fonksiyonun Varyansı

$\hat{p}(x)$ tahmininin varyansını hesaplayabileceğimiz için T_x 'i oluşturan Δt_i zaman aralıklarının istatistiksel özelliklerini bilmemiz gereklidir. Ancak random prosesin zamanla ilgili istatistiksel özelliklerini öğrenmek çok zordur. Örnek değerlerin ortalamasının varyansı örnek hacmine bölünen örneklenmiş random değişkeninin varyansına eşittir.

$$\sigma_x^2 = \text{Var} \left[\frac{1}{N} \sum_{i=1}^N x_i \right] = \frac{1}{N^2} \sum_{i=1}^N \sigma_x^2 = \frac{\sigma_x^2}{N} ; \text{Var}[\hat{p}(x)] = \frac{\text{Var}[p_i(x)]}{N} \quad (2.37)$$

Formüldeki $\text{Var}[p_i(x)]$, serbestlik derecesi bir olan olasılık yoğunluk ölçmesinin etkin varyansıdır. Bir varyansın özel ifadesi bilinmez, fakat en genel ifadesi muhtemelen şu olacaktır.

$$\text{Var}[p_i(x)] = 2A_1^2 p^2(x) \quad (2.38)$$

Buradaki A_1 bir sabittir. Diğer bir deyişle $A_1, p_i(x)$ ' in standart sapmasının olasılık yoğunluk fonksiyonu $p(x)$ ile ne kadar orantılı olduğuyla ilgilidir.

Örneğin, $p_i(x)$ 'in olasılık yoğunluk fonksiyonunun uniform olduğunu farzedelim. $p_i(x)$ asla negatif olmadığından, uniform olasılık fonksiyonunun sınırları sıfır ve $2p(x)$, yoğunluk da $1/2p(x)$ olacaktır.

Bu özel durum için;

$$\text{Var}[p_i(x)] = \int_0^{2p(x)} [\xi - p(x)] \left[\frac{1}{2p(x)} \right] d\xi = \frac{1}{3} p^2(x) \quad (2.39)$$

olur. Bu özel durum için $A \approx 0.41$ 'dir.

(2.37) denklemindeki $N, \hat{p}(x)$ 'i hesaplamak için kullanılan bağımsız gözlemlerin sayısını gösterir. N için uygun bir ifade zaman domenindeki örnekleme teoremini kullanarak elde edilir. B bandgenişliği ve T uzunluğu olan bir örnek kayıd $n = 2BT$ ayrık değerleriyle tamamen yeniden üretilebilir. Zaman $x - (W/2)$ ve $x + (W/2)$ arasındaki aralığa girdiğinde, olasılık yoğunluk ölçümleri için veri gözlenebilir ve analizi yapılabilir. (2.29) denkleminden görülebileceği gibi bu zaman aralığı $T_x = TW\hat{p}(x)$ 'e eşittir. Bundan dolayı veriyi yeniden üretecek olan ayrık değerlerin sayısı $n = 2BTW\hat{p}(x)$ 'dir. $x(t)$ 'nin band genişliği sınırlı beyaz gürültü olduğu özel durumda bu n değerleri bağımsızdır. Bu özel durumda

$$N = 2 BTW \hat{p}(x) \quad (2.39)$$

olması için $n = N'$ dir.

(2.38) ve (2.39) denklemlerini (2.37) de yerine koyarsak

$$\text{Var} [\hat{p}(x)] = \frac{A_1^2 p^2(x)}{BTW \hat{p}(x)} \quad (2.40)$$

denklemini elde edilir.

2.3.2. Biased Tahminin Hatası

Gerçek olasılık yoğunluk fonksiyonları cinsinden (2.32) denklemini yeniden yazalım.

$$E [\hat{p}(x)] = \frac{1}{W} \int_{x-(W/2)}^{x+(W/2)} p(\xi) d\xi$$

$p(\xi)$, $\xi = x$ noktası civarında Taylor serisine açılır ve sadece 3 terimi yazılırsa;

$$p(\xi) \approx p(x) + (\xi - x)p'(x) + \frac{(\xi - x)^2}{2} p''(x) \quad (2.41)$$

Yukarıdaki 2 denklem vasıtasıyla aşağıdaki denklemler elde edilebilir.

$$\int_{x-(W/2)}^{x+(W/2)} (\xi - x) d\xi = 0 ; \int_{x-(W/2)}^{x+(W/2)} \frac{(\xi - x)^2}{2} d\xi = \frac{W^2}{24} \quad (2.42)$$

$$E [\hat{p}(x)] \approx p(x) + \frac{W^2}{24} p''(x) \quad (2.43)$$

Böylece, biased tahmininin hatasını bulmuş olduk.

$$b[\hat{p}(x)] \approx \frac{W^2}{24} p''(x) \quad (2.44)$$

Buradaki $p''(x)$, x 'e göre $p(x)$ 'in ikinci türevidir.

KARESEL ORTALAMA HATA: (2.34) denklemiyle verilen tahmini olasılık yoğunluk fonksiyonu $\hat{p}(x)$ 'in toplam karesel ortalama hatası, (2.40) ve (2.44) denklemleri yardımıyla hesaplanır.

$$E[(\hat{p}(x) - p(x))^2] \approx \frac{A_1^2 p^2(x)}{BTW\hat{p}(x)} + \left(\frac{W^2 p''(x)}{24} \right) \quad (2.45)$$

$W \rightarrow 0$ ve $WT \rightarrow \infty$ 'a gitmesi için W sınırlanırsa $T \rightarrow \infty$ 'a giderken karesel ortalama hata da sifıra yaklaşır.

2.3.3. ORTAK OLASILIK YOĞUNLUK TAHMİNLERİ: Stasyoner (ergodik) random prosesler $\{x(t)\}$ ve $\{y(t)\}$ nin örnek kayıtları olan $x(t)$ ve $y(t)$ nin ortak olasılık yoğunluk tahminlerini aşağıdaki denklemlerle tarif edebiliriz.

$$\hat{P}[x, W_x; y, W_y] = \frac{T_{x,y}}{T} \quad (2.46)$$

(2.46) denklemiyle ortak olasılık tahminini bulabiliriz. $y(t)$ W_y genlik aralığına girdiğinde $x(t)$ 'nin de W_x genlik aralığına girme olasılığını bu denklemle bulabiliriz. Bu tahmin $\frac{T_{x,y}}{T}$ oranıyla ölçülür. Formüldeki $T_{x,y}$, $x(t)$ ve $y(t)$ nin T zamanı boyunca aynı anda sırasıyla W_x ve W_y aralıklarında bulundukları zamanların toplamıdır.

Bu tahmin edilen ortak olasılık T sonsuza yaklaşırken gerçek olasılık $P[x, W_x; y, W_y]$ ya yaklaşır. Şöyle ki,

$$P[x, W_x; y, W_y] = \lim_{T \rightarrow \infty} P[x, W_x; y, W_y] = \lim_{T \rightarrow \infty} \frac{T_{x,y}}{T} \quad (2.47)$$

Ortak olasılık yoğunluk fonksiyonu $p(x, y)$ de şu şekilde izah edilebilir:

$$p(x, y) = \lim_{\substack{W_x \rightarrow 0 \\ W_y \rightarrow 0}} \frac{P[x, W_x; y, W_y]}{W_x W_y} = \lim_{\substack{T \rightarrow \infty \\ W_x \rightarrow 0 \\ W_y \rightarrow 0}} \frac{\hat{P}[x, W_x; y, W_y]}{W_x W_y} = \lim_{\substack{T \rightarrow \infty \\ W_x \rightarrow 0 \\ W_y \rightarrow 0}} \hat{p}(x, y) \quad (2.48)$$

$$\hat{p}(x, y) = \frac{\hat{P}[x, W_x; y, W_y]}{W_x W_y} = \frac{T_{x,y}}{T W_x W_y} \quad (2.49)$$

Etkileşim hataların (bias errors) ihmal edebilmek için W_x ve W_y 'nin yeterince küçük olduğunu farzedelim. Bu durumda, $\hat{p}(x, y)$ tahmininin karesel ortalama hatası varyansına eşittir. $x(t)$ ve $y(t)$ 'nin aynı band genişliklerine sahip band genişliği sınırlı beyaz gürültü olduğunu kabul edersek,

$$\text{Var}[\hat{p}(x, y)] \approx \frac{A_2^2 p^2(x, y)}{B T W_x W_y \hat{p}(x, y)} \quad (2.50)$$

elde edilir. Formüldeki A_2 bilinmeyen bir sabittir.

Normalize edilmiş hata : $\hat{p}(x)$ tahmininin normalize edilmiş karesel ortalama hatası aşağıdaki denklemle tesbit edilir.

$$\epsilon^2 = \frac{E[(\hat{p}(x) - p(x))^2]}{p^2(x)} \approx \frac{A_1^2}{BT W \hat{p}(x)} + \frac{W^4}{576} \left(\frac{p''(x)}{p(x)} \right)^2 \quad (2.51)$$

ϵ normalize edilmiş standart hatadır. $W^4 [p''/x]^2$ küçük olursa denklemin sağ tarafındaki bias hatası ihmal edilir.

2.4. KORELASYON FONKSİYONUNUN TAHMİNLERİ

Stasyoner (ergodik) random prosesler $\{x(t)\}$ ve $\{y(t)\}$ 'deki $x(t)$ ve $y(t)$ örnek kayıtları değerlerinin birbirine olan bağımlılığını incelemek için kros korelasyon fonksiyonunu bulmamız gerekir. Hesapları basitleştirmek için μ_x ve μ_y nin sıfır olduğunu kabul edeceğiz. T gözlem zamanında mevcut olan $x(t)$ ve $y(t)$ örnek kayıtlarının tahmini kros-korelasyon fonksiyonu:

$$\begin{aligned} \hat{R}_{xy}(\tau) &= \frac{1}{T-\tau} \int_0^{T-\tau} x(t)y(t+\tau)dt \quad 0 \leq \tau < T \\ &= \frac{1}{T-|\tau|} \int_{|\tau|}^T x(t)y(t+\tau)dt \quad -T < \tau \leq 0 \end{aligned} \quad (2.52)$$

Mutlak değer işaretlerini kullanmamak için τ 'nin bundan sonra pozitif olduğunu kabul edeceğiz. $x(t)$ ve $y(t)$ örnek kayıtlarının aynı zaman periyodunda meydana geldiğini kabul ettiğimizde otokorelasyon fonksiyon tahminleri $\hat{R}_x(\tau)$ ve $\hat{R}_y(\tau)$:

$$\hat{R}_x(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} x(t)x(t+\tau) dt \quad 0 \leq \tau < T$$

$$\hat{R}_y(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} y(t)y(t+\tau) dt \quad 0 \leq \tau < T$$
(2.53)

Ortalama değeri $\mu_x = 0$, bandgenişliği B ve gözlem süresi T olan bandgenişliği sınırlı beyaz gürültünün bütün τ değerleri için varyansı yaklaşık olarak tesbit edilebilir.

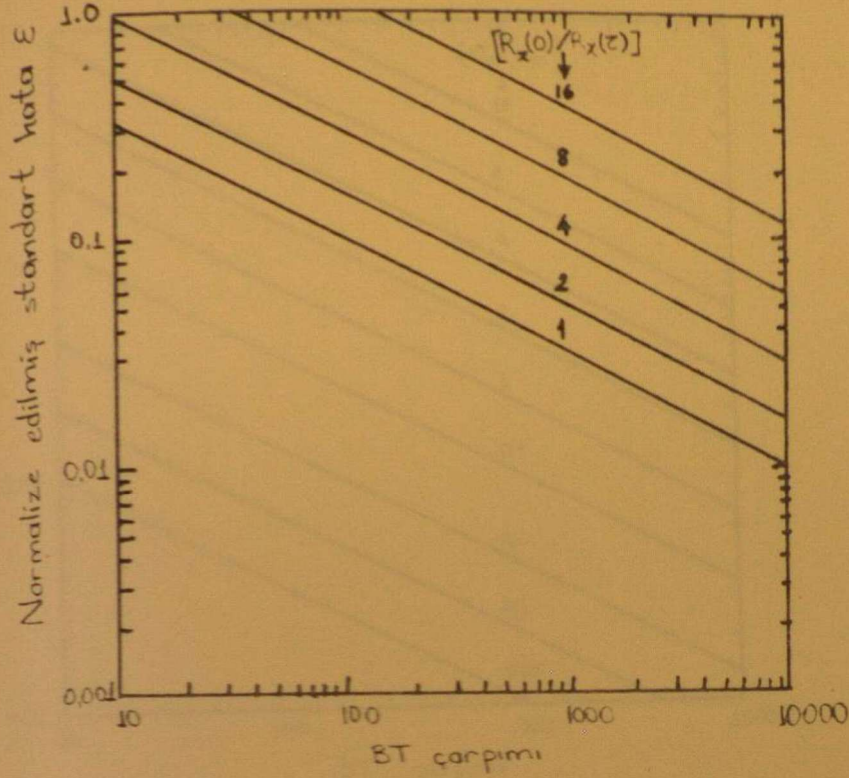
$$\text{Var}[\hat{R}_x(\tau)] \approx \frac{1}{2BT} (R_x^2(0) + R_x^2(\tau))$$
(2.54)

Tahmini otokorelasyon fonksiyonu $\hat{R}_x(\tau)$ 'nin normalize edilmiş standart hatası:

$$\epsilon^2 = \frac{\text{Var}[\hat{R}_x(\tau)]}{R_x^2(\tau)} \Rightarrow \epsilon = \frac{\sqrt{\text{Var}[\hat{R}_x(\tau)]}}{R_x(\tau)} \approx \frac{1}{\sqrt{2BT}} \left(1 + \frac{R_x^2(0)}{R_x^2(\tau)}\right)^{1/2}$$
(2.55)

$x(t)$ örnek kaydının B bandgenişliğinde uniform güç spektrumu gösterdiğini varsayarak arzu edilen ϵ hatayı gerçekleştirebilmek için gerekli minimum örnek kayıt uzunluğu T_r 'yi hesaplayabiliriz.

$$T_r = \frac{1}{2B\epsilon^2} \left[1 + \frac{R_x^2(0)}{R_x^2(\tau)}\right]$$
(2.56)

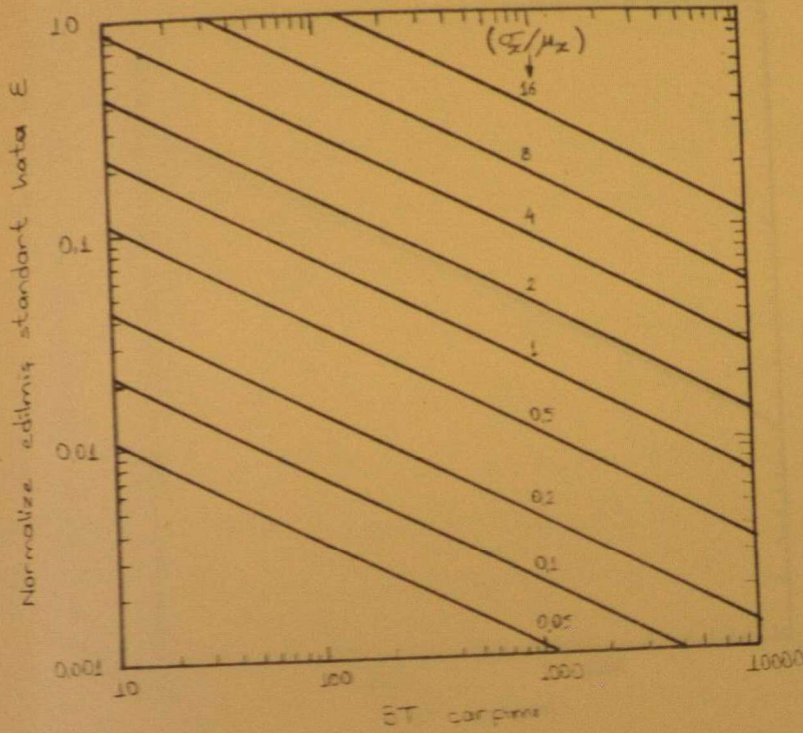


řekil 2.1. Otokorelasyon fonksiyonlarında yapılan istatistiksel hata

2.5. Ortalama ve Karesel Ortalama Deęerler: $x(t)$  rnek kaydının ger ek ortalama deęeri $\mu_x \neq 0$ ve standart sapmasının σ_x olduęunu g z n nde bulundurarak tahmini ortalama deęer $\hat{\mu}_x$ in normalize edilmiř standart hatasını ve minumum  rnek kayıd uzunluęu T_r 'yi ařaęıdaki form ller yardımıyla hesaplayabiliriz.

$$\epsilon^2 = \frac{\text{Var} [\hat{\mu}_x]}{\mu_x^2} \Rightarrow \epsilon = \frac{\sqrt{\text{Var} [\hat{\mu}_x]}}{\mu_x} \approx \frac{1}{\sqrt{2BT}} \left(\frac{\sigma_x}{\mu_x} \right) \quad (2.57)$$

$$T_r = \frac{1}{2B\epsilon^2} \left(\frac{\sigma_x}{\mu_x} \right)^2 \quad (2.58)$$



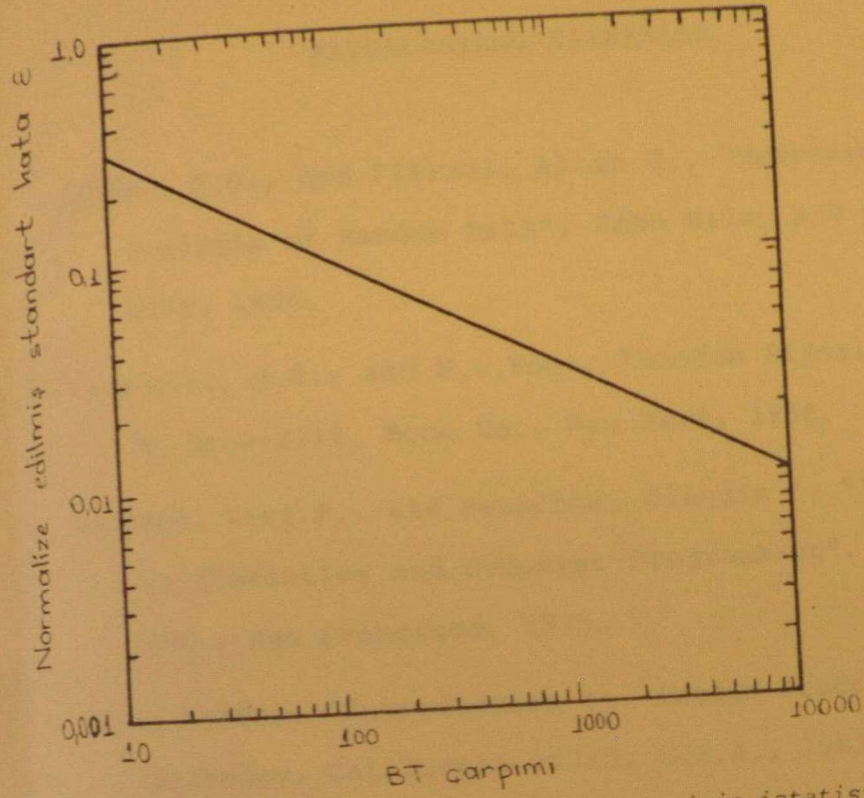
Şekil 2.2.: Ortalama değer tahminlerinin istatistiksel hatası

Gerçek ortalama değeri $\mu_x = 0$ ve gerçek rms değeri ψ_x olan $x(t)$ örnek kaydının karesel ortalama değer tahmini $\hat{\psi}_x^2$ nin normalize edilmiş standart hatası:

$$\epsilon^2 = \frac{\text{Var}[\hat{\psi}_x^2]}{\psi_x^4} \Rightarrow \epsilon = \frac{\sqrt{\text{Var}[\hat{\psi}_x^2]}}{\psi_x^2} \approx \frac{1}{\sqrt{BT}} \quad (2.59)$$

$x(t)$ örnek kaydının B bandgenişliğinde uniform güç spektrumunun olduğunu farzederek, herhangi bir ϵ hatasındaki minimum örnek kayıt uzunluğunu aşağıdaki formül yardımı ile bulabiliriz.

$$T_r = \frac{1}{B\epsilon^2}$$



Şekil 2.3: Karesel ortalama değer tahminlerinin istatistiksel hatası.

Not : Tez çalışmamız esnasında, uygun türkçe karşılıklarını bulamadığımız ingilizce kelimeleri aynen kullandık. Bazen de türkçe karşılığının yanına parantez içinde İngilizcesini yazdık.

FAYDALANILAN KAYNAKLAR

1. Bendat, J.S., and Piersol, Allan G., "Measurement and Analysis of Random Data", John Wiley and Sons, New York, 1966.
2. Davenport, W.B., and W.L.Root, "Random Signals and Noise", Mc Graw-Hill Book Co., New York, 1958.
3. Kossack, Carl F., and Henschke, Claudia I, "Introduction to Statistics and Computer Programming", Holden- Day, INC., San Francisco, 1975.
4. Wong, Eugene, "Introduction to Random Processes", Berkeley, California 94720, U.S.A., 1983.
5. Brillinger, R.David, "Time Series Data Analysis and Theory", Holt, Rinehart and Winston, 1975.
6. Mosteller, Frederick and Tukey, W. John, "Data Analysis and Regression", Addison - Wesley Publishing Company, Inc., 1977.
7. Taylor, R. John, "An Introduction to Error Analysis", Colarodo, University Science books, U.S.A., 1982.