6. İSTATİSTİK DAĞILIM FONKSİYONLARI

Bu bölümde sürekli türden bir rastgele değişkenin kitle uzayındaki*(domain)* değişimini matematik olarak tasvir etmeye yarayan dağılım fonksiyonlarından bazıları ele alınmaktadır. Konu bir rastgele değişken olduğunda; o zaman ilgili dağılım da tek değişkenli*(univariate)* bir dağılım fonksiyonu, konu birden fazla değişken olduğunda, o zaman dağılım çok değişkenli*(multivariate)* bir dağılım fonksiyonu olmaktadır. Ancak, fiziksel ya da doğal olaylarla ilgili gözlemler; benzer özelliklere sahip jeodezik gözlemler de, normal dağılımda olduklarından burada; sadece *normal dağılım* ve onunla ilgili;

- *t-Student Dağılımı*
- $\chi^2 - Dağılımı$
- *F-Fisher Dağılımı*

biçiminde ele alınacaktır. Bunlarla ilgili matematiksel özellikler özet alarak verildikten sonra, gerekli hesaplamaların veya ilgili tabloların nasıl kullanılacağına ilişkin bazı sayısal örnekler verilerek konu daha da kolay anlaşılır hale getirilmeye ayrıca özen gösterilecektir.

6.1. NORMAL DAĞILIM

Normal dağılım fonksiyonu sözcüğü ilk defa 1889 yılında *Gaulton* tarafından kullanılmış bir kelimedir. Ancak, daha sonraları bazı özel kullanım amaçları için bunun yanında *Gauss dağılım fonksiyonu* veya kısaca *çan eğrisi* terimlerinin de kullanıldığı ilgili kaynaklardan görülmektedir. Böyle bir fonksiyon bağımsız değişkenlerin sayısına göre;

- *tek boyutlu (yada tek değişkenli) normal dağılım fonksiyonları,*
- *çok boyutlu (yada çok değişkenli) normal dağılım fonksiyon*ları

biçiminde ele alınabilirler. Tek boyutlu normal dağılım fonksiyonları düzlemde bir alan tanımlarken, çok boyutlu normal dağılım fonksiyonları n boyutlu uzayda daima bir yüzey veya hacim belirlemektedir. her şeyden önce, bir deny sonucunda elde edilmiş sonlu sayıdaki verilerin istatistik hipotez testleri kullanılarak analiz edilmeleri açısında öncelikle bu dağılımların genel özelliklerinin bilinmesi gerekir. Bu nedenle, burada her bir dağılım genel durumuyla özet de olsa ele alınarak incelenmektedir.

6.1.1. Tek Değişkenli Normal Dağılım Fonksiyonu

Tek değişkenli bir normal dağılım fonksiyonu için böyle bir amaca yönelik matematik ifadesi,

$$I = \int_{-\infty}^{+\infty} \exp(-y^2/2)\ dy \qquad 6\text{-}51$$

entegral bağıntısından faydalanılarak elde edilebilir. Böyle bir entegral değerinin hesabı için;

$$I^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp(-\frac{y^2+z^2}{2})\ dy\ dz \qquad 6\text{-}52$$

şeklinde kartezyen koordinatlara göre verilmiş olan bir entegral işleminde, kartezyen koordinat değerleri;

$$\begin{bmatrix} y \\ z \end{bmatrix} = r\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad 6\text{-}53$$

kutupsal koordinat dönüşümü yapılarak , (6-52) bağıntısının bu koordinatlar cinsinden bir ifadesi olan,

$$I^2 = \int_0^{2\pi}\int_0^{+\infty} \exp(\frac{-r^2}{2})\ r\,dr\ d\theta = \int_0^{2\pi} d\theta = 2\pi \qquad 6\text{-}54$$

entegral bağıntısında entegral alınarak sınır değerlerinin yerine yazılması ile,

$$I^2 = 2\pi$$

biçiminde elde edilir. Neticede; buradan görüleceği gibi, (6-51) fonksiyonunun entegral değeri de;

$$I = (2\pi)^{1/2} \qquad 6\text{-}55$$

kadar olur. (6-51)' deki entegral işlemi sonucunda; değeri (6-55) de verildiği gibi $I = (2\pi)^{1/2}$ değerine eşit olduğundan; o zaman (6-51) deki entegral bağıntısı da;

$$\int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{1/2}} \exp(-y^2/2)\ dy = 1 \qquad 6\text{-}56$$

şeklinde ifade edilebilir.

Bu şekildeki bir entegral alma işleminde entegral bağıntısının değişkeni için, a ve b herhangi bir sayı olmak üzere;

$$y = \frac{x-a}{b} \quad ; \quad b\rangle 0 \qquad 6\text{-}57$$

şeklinde bir değişken dönüşümü yapılırsa; sonuçta entegral bağıntısı,

$$\int_{-\infty}^{+\infty} \frac{1}{b(2\pi)^{1/2}} \exp(-\frac{(x-a)^2}{2b^2})\ dx = 1 \qquad 6\text{-}58$$

şeklini alır.

Bu (6-58) entegral bağıntısı incelendiğinde, bunun bir rastgele değişkenle ilgili dağılım *(kümülatif dağılım)* fonksiyonu özelliklerine sahip olduğu açıkça görülmektedir. Buradan(6-58) bağıntısının türevinin alınması ile, böyle bir dağılım fonksiyonuna karşılık gelen olasılık ya da yoğunluk *(frekans)* fonksiyonu; (6-58) entegralinin diferansiyeli olan,

$$f(x) = \frac{1}{b(2\pi)^{1/2}} \ \exp\left\{-\frac{(x-a)^2}{2b^2}\right\} \quad burada \ \ -\infty \langle \ x \ \langle \ +\infty \qquad \text{6-59a}$$

biçimindeki üstel bir fonksiyon olarak elde edilir. Burada; x değeri dağılım ve olasılık fonksiyonlarının değişken parametresini, a ve b değerleri de bu fonksiyonlarının dağılım parametrelerini temsil etmektedir.

Çoğu uygulamalarda; rastgele değişkenlerin dağılım ya da olasılık fonksiyonlarında genel bir gösterim olarak; *a için* μ değişkenin umut değeri ve b^2 *içi de* σ^2 varyans değeri kullanılmaktadır. O zaman (6-59a) fonksiyonu ile verilmiş olan olasılık ya da diğer adıyla yoğunluk fonksiyonu,

$$f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \ \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad burada \ \ -\infty \langle \ x \ \langle \ +\infty \qquad \text{6-59b}$$

şeklinde ifade edilir.

Uygulamada bu fonksiyonun temsil ettiği eğri de çan biçiminde olduğundan *Çan eğrisi, Gauss dağılım eğrisi* ya da *Normal dağılım* eğrisi olarak adlandırılır.
(6-59b) fonksiyonlarındaki x değişkeninin sürekli türden bir rastgele değişken olduğundan bu $f(x)$ olasılık fonksiyonu da sürekli türden bir rastgele değişkenin olasılık fonksiyonunu göstermektedir.

Aynı şekilde, (6-59b) olasılık fonksiyonunun, (6-58) ‘deki ifadeye benzer şekilde ifade edilmiş entegrali biçimindeki,

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma(2\pi)^{1/2}} \ \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx = \\ &= \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^{+\infty} \ \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \end{aligned} \qquad \text{6-59c}$$

(6-59c) dağılım fonksiyonu da sürekli bir değişkenin dağılım fonksiyonu olmaktadır. (6-58) deki özellikten dolayı (6-59c) entegralinin değeri de,

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^{+\infty} \ \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx = \\ &= F(+\infty) - F(-\infty) = 1 - 0 = 1 \end{aligned} \qquad \text{6-59c}$$

olmaktadır.

Burada, tekrar vurgulamak gerekirse; sürekli türden rastgele değişkenler yerine diskırıt*(ayrık)* türden rastgele değişkenlerin kullanılmasının söz konusu olduğu problemlerde ise, yukarıda sözü edilen entegral alma işlemi yerine toplam ifadesi kullanılarak aynı işlemler yapılır.

6.1.2 Moment Üreten Fonksiyonlar

Bir normal dağılımın moment üreten fonksiyonları, 4.1.1 paragrafında anlatılanlara benzer şekilde hareket edilerek, normal dağılım için $g(x) = e^{tx}$ alarak,

$$M(t) = \int_{-\infty}^{+\infty} e^{tx} f(x)\ dx = \int_{-\infty}^{+\infty} e^{tx} \frac{1}{b(2\pi)^{1/2}} \exp\left\{-\frac{(x-a)^2}{2b^2}\right\} dx \qquad \text{6-60}$$

şeklinde verilebilir. Burada; bir değişken dönüşümü için

$$y = \frac{x-a}{b} - bt$$

ifadesi yazılarak, bu ifadeden x elemanı çekildiğinde,

$$x = by + b^2 t + a$$

olarak elde edilir. Daha sonra bunun (6-60) da yerine yazılmasından,

$$M(t) = \int_{-\infty}^{+\infty} \exp\{t(by + b^2 t + a)\} \frac{1}{b(2\pi)^{1/2}} \exp\left\{-\frac{(by + b^2 t)^2}{2b^2}\right\} b\ dy$$

$$= \exp\left\{at + \frac{b^2 t^2}{2}\right\} \int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{1/2}}\ \exp(\frac{-y^2}{2})\ dy \qquad \text{6-61}$$

bağıntısı elde edilerek, burada gerekli işlemlerin yapılması neticesinde; bir normal dağılımın *M(t)* moment üreten fonksiyonları için nihai sonuç,

$$M(t) = \exp\left\{at + \frac{b^2 t^2}{2}\right\} \qquad \text{6-62}$$

olarak elde edilebilir.

Buradan, daha önceki konularda (4-10d) ifadesi ile verilmiş olan bir dağılımın μ ortalaması ya da umut değeri, onun moment üreten fonksiyonu ile

$$\mu = \frac{\partial M(t)}{\partial t} = M^{'}(0)$$

şeklinde ilişkili olduğu bilinmektedir. Normal dağılım için böyle bir ilişki;

$$M^{'}(t) = M(t)(a + b^2 t)$$

ve t elemanı için de $t = 0$ alınırsa bunun sonucunda, normal dağılımın umut değeri yada parametresi,

$$\mu = M^{'}(0) = a \qquad 6\text{-}63$$

olarak bulunur.
Benzer şekilde, (4-10e) denkleminde verilmiş olan, normal dağılımın ikinci merkezsel momenti olan dağılımın σ^2 varyansı da;

$$\sigma^2 = M^{''}(0) - \{M^{'}(0)\}^2$$

şeklinde hesaplanabilir. Bu bağıntının kullanılması ile bir normal dağılımın σ^2 değeri hesaplandığında;

$$M^{''}(t) = M(t)(b^2) + M(t)(a + b^2 t^2)^2 \quad ,$$

ve buradan da;

$$\sigma^2 = M^{''}(0) - \{M^{'}(0)\}^2 = (b^2 + a^2) - a^2 = b^2 \qquad 6\text{-}64$$

olarak elde edilir.

Sonuçta, bu açıklamaların ışığı altında (6-59) da verilmiş olan normal dağılımın olasılık fonksiyonu μ ve σ^2 parametrelerine göre;

$$f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \quad ; \quad -\infty \langle \ x \ \langle +\infty \qquad 6\text{-}65a$$

şeklinde yazılabilir.

Benzer şekilde, μ ve σ^2 parametreleri de kullanılarak, normal dağılım fonksiyonu da,

$$F(x) = \int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^{+\infty} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} dx = \qquad 6\text{-}65b$$
$$= F(+\infty) - F(-\infty) = 1 - 0 = 1$$

şeklinde ifade edilmiş olur.

Tekrar söylemek gerekirse; bazı uygulamalarda bu şekilde ifade edilmiş olan normal dağılımın olasılık fonksiyonu için,

$n(\mu, \sigma^2)$ ya da $x \to N(\mu, \sigma^2)$

şeklindeki bir kısa gösterim de kullanılmaktadır. Pratik anlamda, bu $x \to N(\mu, \sigma^2)$ ifadesi; x rastgele değişkeni μ ve σ^2 varyansına göre normal dağılımda olan bir rastgele değişkendir şeklinde okunur.

Özetle, yapılan bu işlemlerden normal dağılım için moment üreten fonksiyonun,

$$M(t) = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \qquad 6\text{-}66$$

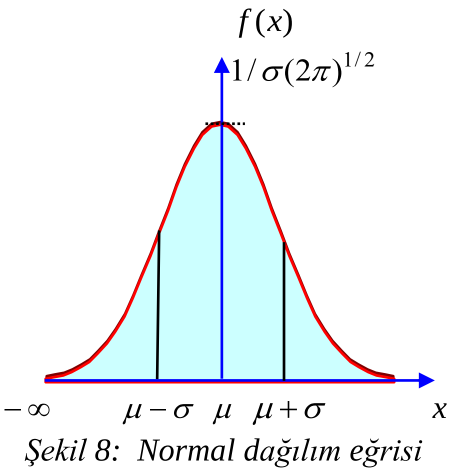
şeklindeki bir bağıntı olduğu verilebilir.

6.1.3 Normal Dağılımın Eğrisi ve Özellikleri

Bir normal dağılım ile ilgili,

$$f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad ; \quad -\infty \langle\ x\ \langle +\infty$$

yoğunluk fonksiyonu ya da daha pratik amaçlarına yönelik gösterimler için kullanılan $n(\mu, \sigma^2)$ yada $x \to N(\mu, \sigma^2)$ normal dağılım gösterimlerinde iki farklı parametre değeri bulunmaktadır. Bunlardan biri μ dağılım eğrisinin umut değeri, diğeri de onun yayılmasının bir ölçütü olan σ standart sapma değeridir. Bu iki parametrenin alacağı farklı değerlere göre tek boyutlu bir normal dağılım fonksiyonun bir düzlem dik kartezyen koordinat sisteminde grafiği çizildiğinde Şekil 8‘daki gibi bir eğrinin elde edilmiş olduğu görülür.



Şekil 8: Normal dağılım eğrisi

Şekil 8’dan görüleceği gibi; bu normal dağılım eğrisi, her zaman,

- μ *Parametresine göre simetriktir. Bu nedenle bütün tek sayıdaki ikinci merkezsel momentleri toplamı sıfırdır,*

$$\sigma^k = E\{(x-\mu)^k\} = 0 \quad ; \quad k = 1, 3, 5, 7 \ldots,$$

- $\mu = x$ *İçin* $1/\sigma(2\pi)^{1/2}$ *değerinde Maksimum olur. Bunun standart değeri 0.399 gibi bir miktar olmaktadır. Bu aynı zamanda normal dağılımın en olasılıklı yada tepe değeri olmaktadır.*
- *Yoğunluk fonksiyonu* $\pm\infty$ *değerleri için* x *eksenine asimtotiktir.*
- *x Ekseni ile f(x) yoğunluk fonksiyonunun sınırladığı alan* S *güven seviyesini veya* $1 - S = \alpha$ *yanılma olasılığını göstermektedir.*
- *f(x) Yoğunluk fonksiyonu* $x = \mu \pm \sigma$ *değerleri onun dönüm noktalarıdır.*
- *Bir normal dağılım için,* μ_x *gerçek değerinin* $x \pm \sigma_x$, $x \pm 2\sigma_x$ *ve* $x \pm 3\sigma_x$ *sınırları arasına düşme olasılığı,*

$$P(-\sigma_x < x - \mu_x < \sigma_x) = 0.6827$$
$$P(-2\sigma_x < x - \mu_x < 2\sigma_x) = 0.9545$$
$$P(-3\sigma_x < x - \mu_x < 3\sigma_x) = 0.9973$$

kadar bir değer olmaktadır.

- *Ayrıca bir normal dağılım için* $S = 0.90$, $S = 0.95$ *olasılıklarına göre güven alanı sınırları,*

$$P(-1.645\sigma_x < x - \mu_x < 1.645\sigma_x) = 0.90$$
$$P(-1.960\sigma_x < x - \mu_x < 1.960\sigma_x) = 0.95$$
$$P(-2.576\sigma_x < x - \mu_x < 2.576\sigma_x) = 0.99$$

değerinde olmaktadır.

- *Bir normal dağılımın modu medyana eşittir. Yani normal dağılım eğrisinin sınırladığı alanı,*

$$\int_{-\infty}^{m} f(x)dx = 0.5$$

şeklinde iki eşit parçaya bölen m sınır değeri aynı zamanda onu Maksimum yapan en olasılıklı değeri, diğer bir ifade ile hem mod hem de medyan değerleri olmaktadır.

- $x \to N(\mu, \sigma^2)$ *Rastgele değişkenin* $P(x \le \mu_x) = 0.50$ *olmaktadır.*
- *Bu eğrinin sınırladığı toplam alan,*

$$S = \int_{-\infty}^{+\infty} f(x)dx = 1$$

daima 1 değerinde olmaktadır. Bu durum bir olayın mutlak yani; %100 gerçekleşeceğini göstermektedir

özelliklerine sahip bir eğri olmaktadır.

Uygulamada çoğu zaman, bu normal dağılım eğrisi, bir çana benzediğinden dolayı buna *Çan eğrisi* yada bir diğer ifade ile *Gauss dağılım eğrisi* de denmektedir.

6.1.4 Normal Dağılıma Sahip Bir Rastgele Değişkenin Standartlaştırılması

Eğer bir x rastgele değişken $x \to N(\mu, \sigma^2)$ parametrelerine göre normal dağılıma sahipse, o zaman bunun standartlaştırılmış *(Normlandırılmış yada standart hale getirilmiş)* durumdaki z yeni rastgele değişkeni;

$$z = \frac{x-\mu}{\sigma}$$

şeklindeki bir lineer dönüşüm bağıntısından elde edilir. Neticede, z standartlaştırılmış z yeni rastgele değişkeni;

$$E\{z\} = E\left\{\frac{x-\mu}{\sigma}\right\} = \frac{1}{\sigma}(E\{x-\mu\}) = \frac{1}{\sigma}(E\{x\} - E\{\mu\}) = \frac{1}{\sigma}(\mu-\mu) = 0$$

umut değeri ve

$$E\{z^2\} = E\left\{(\frac{x-\mu}{\sigma})^2\right\} = \frac{1}{\sigma^2}(E\{(x-\mu)^2\}) =$$

$$= \frac{1}{\sigma^2}(E\{(x^2 - 2x\mu + \mu^2)\}) = \frac{1}{\sigma^2}(E\{x^2\} - E\{2\mu x\} + E\{\mu^2\}) =$$

$$= \frac{1}{\sigma^2}(E\{x^2\} - 2\mu E\{x\} + \mu^2) = \frac{1}{\sigma^2} E\{x^2\} - \mu^2 = \frac{\sigma^2}{\sigma^2} = 1$$

varyans değerlerine göre;

$$z \to N(0.1)$$

şeklinde ifade edilebilen,

- *Umut değeri :* $\mu = 0$ *sıfır*
- *Varyansı :* $\sigma^2 = 1$ *bir*

olan bir normal dağılıma sahip olur.

Matematik-istatistikte bir rastgele değişkenin bu şekilde eşdeğer özelliklere sahip bir diğer rastgele değişkene dönüştürülmesi işlemine; standartlaştırma yada rastgele değişkenlerin standartlaştırılması işlemi denir. Bunun yukarıda, umut değerlerine göre yapılmış açıklamasına göre de; sürekli türden bir rastgele değişken için matematik istatistik yasalarına göre bir diğer açıklaması aşağıdaki gibi yapılabilir.

Böyle bir standartlaştırılmış z rastgele değişkenin bilinen *Kümülatif dağılım* fonksiyonu;

$$F(z) = P_r(\frac{x-\mu}{\sigma} \le z) = P_r(x \le z\,\sigma + \mu)$$

ya da entegral ifadesi;

$$F(z) = \int_{-\infty}^{z\sigma+\mu} \frac{1}{\sigma(2\pi)^{1/2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$

olarak verilebilir. Burada; bir x rastgele değişkeni için

$$z = \frac{x-\mu}{\sigma}$$

rastgele değişken dönüşümü yapılarak standart hale getirilmiş z rastgele değişkene göre dağılım fonksiyonunun ifadesi;

$$F(z) = \int_{-\infty}^{z} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{z^2}{2}\right\} dz$$

ve buna karşılık gelen yoğunluk fonksiyonu da; dağılımın fonksiyonunu değişkene göre;

$$f(z) = \frac{\partial F(z)}{\partial z} = F^{'}(z)$$

şeklindeki bir türevi olduğundan;

$$f(z) = \frac{1}{(2\pi)^{1/2}} \exp(\frac{-z^2}{2}) \qquad 6\text{-}67$$

olarak verilebilir.

Bu şekilde (6-67) de elde edilmiş olan bir yoğunluk fonksiyonu ifadesinin; (6-65a) bağıntısında $\mu = 0$ ve $\sigma^2 = 1$ alınarak elde edilecek ifadenin aynısı olduğu açıkça görülür.

Bu durumda; $z \to N(0,1)$ standart normal dağılımın tanımladığı eğri, $x \to N(\mu,\sigma^2)$ normal dağılım fonksiyonda dağılımın parametreler için $\mu = 0$ ve $\sigma^2 = 1$ olduğu dikkate alınarak elde edilen fonksiyonun tanımladığı eğriye benzer olur.

Neticede, her iki değişkene göre ifade edilen normal dağılım bağıntıları birbirleriyle bağlantılı ve eşdeğer ifadeler olurlar. Aynı zamanda böyle bir eğri normal dağılım eğrisine, yukarıda sözü edilen durumlara benzer şekildeki,

- $x = 0$ *Parametresinin belirlediği düşey eksene göre simetriktir,*
- $x = 0$ *İçin* $1/\sigma(2\pi)^{1/2}$ *değerinde Maksimum olur,*
- $\pm\infty$*Değerleri için* x *eksenine asimtotiktir,*
- $x = \pm\sigma$ *Değerleri onun dönüm noktasıdır,*
- *Modu meydanına eşittir,*
- $\int_{-\infty}^{+\infty} f(x)dx = 1$ *Olmaktadır*

özelliklere de sahip olur.

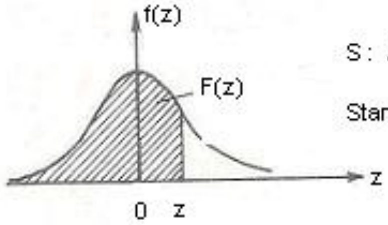
6.1.5 Normal Dağılımla İlgili Hesaplamalar

Bir boyutlu normal dağılımın fonksiyonu iki parametre ile tanımlanır. Bunlardan birisi; dağılımın μ ortalama ya da umut değeri, bir diğeri ise; dağılımın σ^2 varyans değeridir. Bu parametrelerin belirlediği dağılıma sahip bir rastgele değişkenin herhangi bir sınır değerinden küçük veya büyük olma, ya da belli sınır değerleri arasına düşme olasılığı, parametrelerin tanımladığı olasılık fonksiyonunun bu sınır değerleri arasındaki entegral değerine eşdeğer olur. Aynı zamanda, dağılım yada olasılık fonksiyonu da yoğunluk fonksiyonunun bu sınır değerleri arasındaki entegral değeri veya tanımladığı alan olmaktadır.

Bu alan, aynı zamanda, rastgele değişken için anlamlılık*(signifikant)* seviyesini gösteren *S-güven* alanını temsil eder. Arda kalan, $\alpha = 1 - S$ dağılımın uç kısımlarındaki alan da yanılma olasılığını gösteren güvensiz alan veya alanlar olur. İstatistik anlamda olayları açıklamak amacıyla, böyle bir değerin elde edilmesinde; her seferinde olasılık fonksiyonunun sınır değerleri arasındaki entegralinin alınması gerekmektedir. Olasılık fonksiyonlarının üstel yapıda bir fonksiyon olması nedeniyle, entegrallerinin her seferinde alınması oldukça karmaşık ve zaman alıcı matematiksel işlemler gerektirmektedir.

Uygulamada, günümüz teknolojisine paralel olarak hızlı bilgisayarların kullanım alanına girmiş olmasına rağmen, alışılagelen bir yol olarak, genellikle, bu gibi karmaşık işlemlerden kurtulmak için, değişkenlerin güven aralıklarının hesaplanmasında işlemler yönünden daha basit bir yol olan, parametreler için $\mu = 0$ ve $\sigma^2 = 1$ değerleri alınarak hesaplanmış, standart dağılım tablolarının kullanılarak istenen sonuçlara varılmasıdır(Tablo 2).

Tablo 2: Standart Normal Dağılım Tablosu



S : Anlamlılık seviyesine göre düzenlenmiş

Standart Normal Dağılımın Dağılım Tablosu

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0,1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0,2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0,3	.6179	.6217	.6256	.6293	.6331	.6368	.6406	.6413	.6480	.6517
0,4	.6554	.6594	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0,5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0,6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0,7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0,8	.7881	.7910	.7939	.7967	.7995	.8022	.8051	.8079	.8106	.8133
0,9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1,0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8539	.8621
1,1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1,2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1,3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1,5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1,6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9526	.9535	.9545
1,7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1,8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1,9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2,0	.9773	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2,1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2,2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2,3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2,4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2,5	.9938	.9940	.9943	.9944	.9945	.9946	.9948	.9949	.9951	.9952
2,6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2,7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2,8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2,9	.9981	.9982	.9983	.9983	.9984	.9984	.9984	.9985	.9986	.9986
	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
3,0	.9987	.9990	.9993	.9995	.9997	.9997	.9998	.9998	.9999	.9999

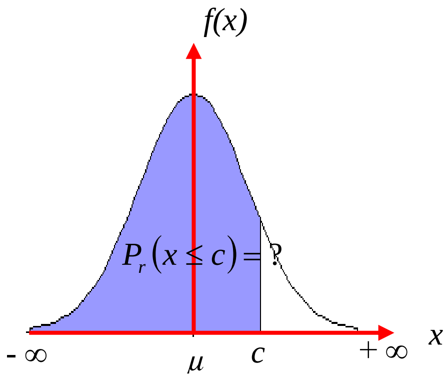
Çok amaçlı kullanıma yönelik önceden istatistikçiler tarafından hazır durumda sunulmuş olan bu tür standart normal dağılım tablolarının kolay kullanılabilir olmaları için kullanım parametreleri olarak; P_r olasılık değerleri ile bu olasılık değerlerine karşılık gelen z standart rastgele değişkeni apsis değerleri tabloların kullanım parametreleri olarak seçilmiştir.

Bu tablolardan doğrudan alınacak veya ara değerler için enterpolasyon yoluyla hesaplanacak z apsis değeri; verilen olasılık değerine karşılık gelen dağılım fonksiyonunun bağımsız değişkenine ilişkin herhangi bir özel yada sınır değerlerini göstermektedir.

Ayrıca, normal dağılımım simetrik yapılı bir dağılım olması nedeniyle, bu gibi tablolar $P_r(0 \leq z \leq +\infty)$ değerleri için düzenlenerek, $z = 0.00$ sınır değerine karşılık gelen $P_r(z = 0.00) = 0.5000$ değerinden başlamaktadır. z *standart rastgele* değişkeninin alacağı negatif değerler için olasılık değerleri normal dağılımın özelliklerinden faydalanarak tabloda verilmiş olan mevcut değerlerden doğrudan hesaplanır. Tersi yöndeki hesaplamalar için de durum aynı olmaktadır.

Uygulamada karşılaşılan bir diğer durum da; bazı normal dağılım tablolarının S güven alanı yerine $\alpha = 1 - S$ yanılma olasılığına göre düzenlenmiş olmalarıdır(Ek 1: standart normal dağılım tablosu). Ancak, böyle bir tablo düzenlemesi konunun ele alınış biçiminden başka hesaplamalar açısından hiçbir sorun teşkil etmemektedir.

Direkt problemlerde; tabloya z apsis değeri girilerek ona karşılık gelen olasılık değeri tablodan varsa hemen yoksa basit bir doğrusal enterpolasyonla hesaplanarak alınır. Buna karşılık, invers çözümlerde bunun tersi yol izlenerek verilen bir olasılık değerine karşılık gelen z apsis değeri bulunur. Böyle bir invers işlem için, önce verilen olasılık değeri tabloda bulunarak, ona karşılık gelen z apsis değeri elde edilir.



Şekil 9: Normal Dağılım Eğrisi

Normal dağılımla ilişkili problemlerin çözümünde temel esas; teorik değer P_r ve z ile tablo haline getirilmiş normal dağılım değeri arasındaki ilişkinin elde edilmesidir. Eğer bir rastgele değişken $x \rightarrow N(\mu, \sigma^2)$ dağılımda ise, o zaman $P_r(x \leq c) = ?$ şeklinde matematik ifadesi ile verilmiş olan; x 'in c gibi bir sınır değerine eşit veya küçük olma olasılığı hesaplanabilir(Şekil 9). Bu amaçla yapılacak bir $P_r(x \leq c) = ?$ olasılığını belirle işleminde, önce $x \rightarrow N(\mu, \sigma^2)$ parametrelerine göre normal dağılıma sahip olan x rastgele değişkeninin ve c sınır değerinin, $z \rightarrow N(0,1)$ parametrelerine göre standart normal dağılım tablosundaki karşılıkları olan z değişken ve z_s sınır değerleri hesaplanması gerekmektedir. Bunun için değişken ve sınır değerlerine ilişkin,

$$z = \frac{x - \mu}{\sigma} \quad ; \quad z_s = \frac{c - \mu}{\sigma}$$

değişken dönüşümleri yapılır. Yapılan dönüşüm sonucunda elde edilmiş olan z değişken ve z_s sınır değerlerine göre;

$$P_r(x \leq c) = P_r\left(\frac{x - \mu}{\sigma} \leq \frac{c - \mu}{\sigma}\right) = P_r(z \leq z_s)$$

olduğu dikkate alınarak,

$$P_r(x \le c) = P_r(z \le z_s) = \int_{-\infty}^{\frac{c-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2}) \; dz = \int_{-\infty}^{z_s} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2}) \; dz$$

$$= F(\frac{c-\mu}{\sigma}) - F(-\infty) = F(z_s) - F(-\infty)$$

olasılık ifadesi yazılabilir. Sonuçta; $F(-\infty) = 0$ olduğu dikkate alınarak, bu ifade;

$$F(\frac{c-\mu}{\sigma}) = F(z_s)$$

şeklinde ifade olacağı görülür. Bu sınır değerine ilişkin $F(z_s)$ olasılık değeri; $z \to N(0.1)$ parametrelerine göre önceden hazırlanmış mevcut standart normal dağılım tablosundan alınarak hesaplanmış olur(Tablo 2).

Özetle tekrar vurgulamak gerekirse, burada

$$z = \frac{x-\mu}{\sigma}$$

şeklindeki standartlaştırma işlemi, $x \to N(\mu, \sigma^2)$ gibi bir normal dağılıma sahip x değişkeninin $z \to N(0.1)$ gibi bir standart normal dağılma sahip z standartlaştırılmış rastgele değişkene dönüştürülerek olasılığın standart normal dağılım tablolarını kullanarak hesaplanabilmesi için yapılmaktadır. Başka bir deyişle; standart normal dağılım tablosu kullanılarak olasılık fonksiyonunun $(-\infty, c)$ aralığında entegrali alınmaktadır.

Sonuçta, böyle bir çözüm neticesinde bir rastgele değişkenin önceden öngörülmüş bir aralıktaki olasılık değerini gösteren güven alanı veya anlamlılık seviyesi hesaplanmış olmaktadır. Bir diğer yönüyle, tersi yöndeki bir çözüm için de bu düşünce aynen geçerliliğini korumaktadır.

Örnek 1: $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \le 4) = ?$ hesaplayınız.

Çözüm 1: $P_r(x \le 4) = ?$ olasılığını hesaplayabilmek için, $x \to N(2,\ 16)$ rastgele değişkeni $\mu = 2$ *ve* $\sigma^2 = 16$ parametre değerlerine göre normal dağılıma sahip olduğundan, öncelikle bu değişkenin hesaplamalarda kullanılacak standart normal dağılım tablosuna ilişkin parametre değerlerine dönüştürülmesi gerekir. Bunun için yukarıda sözü edilen işlem yolu izlenerek;

$$z = \frac{x-\mu}{\sigma}$$

normal dağılıma sahip bir $x \to N(\mu, \sigma^2)$ rastgele değişkenin $z \to N(0{,}1)$ gibi bir dağılıma sahip standart normal dağılımlı bir yeni rastgele değişkene dönüşümü bağıntısından faydalanarak,

$$z_s = \frac{c-\mu}{\sigma} = \frac{4-2}{4} = 0{,}5$$

standart normal dağılıma ilişkin rastgele değişken ve sınır değerleri hesaplanır. Bu değerlere göre problem;

$$P_r(\frac{x-\mu}{\sigma} \le \frac{4-\mu}{\sigma}) = P_r(z \le z_s) = P_r(z \le 0{,}5) = ?$$

olarak ifade edilebilir. Neticede bu ifade;

$$P_r(z \le 0{,}5) = \int_{-\infty}^{0{,}5} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2})\ dz$$
$$P_r(x \le 4) = P_r(z \le 0{,}5) = F(0{,}5) - F(-\infty) = F(0{,}5)$$

şekline dönüşmüş olur.

Buradan, standart normal dağılım tabloları kullanılarak; ilgili standart normal dağılım tablosundan $z = 0.5$ değerine karşılık gelen olasılık değeri alınarak;

$$P_r(x \le 4) = P_r(z \le 0{,}5) = F(0{,}5) = 0.6915$$

bulunur.

Sonuç: $P_r(x \le 4) = 0.6915$ olarak bulunur.
Örnek 2: $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \ge 4) = ?$ hesaplayınız.

Çözüm 2: $P_r(x \ge 4) = ?$ olasılığını hesaplayabilmek için, $x \to N(2,\ 16)$ rastgele değişkeni örnek 1 'de olduğu gibi önce $z \to N(0\ ,\ 1)$ standart normal dağılım tablosuna ilişkin parametre değerlerine dönüştürülür. Bu amaçla; z rastgele değişkeni ve z_s sınır değerleri;

$$z = \frac{x-\mu}{\sigma} = \frac{\text{x - 2}}{4} \qquad ; \qquad z_s = \frac{c-\mu}{\sigma} = \frac{4-2}{4} = 0{,}5$$

olarak hesaplanır. Bu değerlere göre problem;

$$P_r(z \geq 4) = P_r(\frac{x-\mu}{\sigma} \geq \frac{4-\mu}{\sigma}) = P_r(z \geq z_s) = P_r(z \geq 0,5) = ?$$

şekline dönüşmüş olur. Standart normal dağılım tablo değerleri, olasılık hesaplamalarına temel olan entegral işlemlerine uygun olarak $-\infty$ değeri ile z_s gibi bir sınır değeri arasındaki entegral değerlerine göre düzenlenmiş olduklarından, z_s sınır değeri ile $+\infty$ arasındaki olasılık değerleri bu aralıktaki değerlerden faydalanarak;

$$P_r(z \geq z_s) = 1 - \int_{-\infty}^{z_s} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2})\ dz$$

bağıntısına göre;

$$P_r(z \geq 0,5) = 1 - \int_{-\infty}^{0,5} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2})\ dz$$

$$P_r(x \geq 4) = P_r(z \geq 0,5) = 1 - F(0,5) =$$

$$= 1 - 0,6915 = 0,3085$$

hesaplanır.

Sonuç: $P_r(x \geq 4) = 0,3085$ olarak hesaplanır.

Örnek 3: $x \rightarrow N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \leq -4) = ?$ hesaplayınız.

Çözüm 3: Normal dağılım fonksiyonları ortalama değere göre simetrik olduklarından *mod* değeri *medyan* değerine eşittir. Bu özelliğin neticesinde standart normal dağılım tabloları normal dağılım fonksiyonunun yarı alanının pozitif tarafı için düzenlenmektedir. Bu nedenle; negatif yarı alanı içinde sorulan olasılık değeri öncelikle tablo değerleri mevcut olan, $P_r(x \leq -4) = 1 - P_r(x \leq 4)$ eşitliği kullanılarak, pozitif bölgeye çevrilir. Örnek 1 çözümünden $P_r(x \leq 4) = 0,6915$ olduğu bilindiğine göre;

$$P_r(x \leq -4) = 1 - P_r(x \leq 4) = 1 - 0,6915 = 0,3085$$

hesaplanır.

Sonuç: $P_r(x \leq -4) = 0,3085$ olarak bulunur.

Örnek 4: $x \rightarrow N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \geq -4) = ?$ hesaplayınız.

Çözüm 4: Yukarıdaki örneklere benzer şekilde düşünerek; Normal ve standart normal dağılım fonksiyonlarının ortalama değere göre simetrik olduklarından mod değeri medyan değerine eşittir. Bu özelliğinden dolayı normal dağılım aynı zamanda simetrik bir dağılım olmaktadır. Neticede; $P_r(x \geq -4) = P_r(x \leq 4)$ olduğu göz önüne alınarak,

$$P_r(x \geq -4) = P_r(x \leq 4) = 0{,}6915$$

bulunur.

Sonuç: $P_r(x \geq -4) = 0{,}6915$ olur.

Örnek 5: $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(0 \leq x \leq 4) = ?$ aralığına düşme olasılığını hesaplayınız.

Çözüm 5: $P_r(0 \leq x \leq 4) = ?$ için, $x \to N(2,\ 16)$ rastgele değişkeni $\mu = 2$ *ve* $\sigma^2 = 16$ parametre değerlerine göre normal dağılıma sahip olduğundan, önce bu değişken standart normal dağılımlı $z \to N(0,\ 1)$ standart dağılıma sahip rastgele değişkene dönüştürülerek sınır değerleri,

$$z = \frac{x - \mu}{\sigma} = \frac{\text{x - 2}}{4}$$

değişken dönüşümü bağıntısından faydalanarak,

$$z_{ü} = \frac{c - \mu}{\sigma} = \frac{4 - 2}{4} = 0{,}5 \quad ; \quad z_a = \frac{c - \mu}{\sigma} = \frac{0 - 2}{4} = -0{,}5$$

biçiminde hesaplanır. Bu sınır değerlerine göre $x \to N(2,\ 16)$ parametrelerine sahip normal dağılımlı rastgele değişkenin $P_r(0 \leq x \leq 4)$ aralığına düşme olasılığı için,

$$P_r(0 \leq x \leq 4) = P_r(-0{,}5 \leq z \leq 0{,}5) = \int_{-0{,}5}^{0{,}5} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z^2}{2})\ dz$$

entegral bağıntısı kurularak ve bu işlemlerde

$$F(-0{,}5) = 1 - F(0{,}5)$$

olduğunun da dikkate alınması ile;

$$= F(0{,}5) - F(-0{,}5) = F(0{,}5) - \{1 - F(0{,}5)\} =$$
$$= 2F(0{,}5) - 1 = 2(0{,}6915) - 1 = 0{,}3830$$

olarak hesaplanır.

Sonuç: $P_r(0 \leq x \leq 4) = 0{,}3830$ olarak bulunur.

Örnek 6: $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(1 \le x \le 4) = ?$ aralığına düşme olasılığını hesaplayınız.

Çözüm 6: Bu problemin çözümü örnek 5 'dekine benzer şekilde hareket edilerek,

$$z = \frac{x-\mu}{\sigma} = \frac{\mathrm{x - 2}}{4}$$

değişken dönüşümü bağıntısından, alt ve üst sınır değerleri için,

$$z_ü = \frac{c-\mu}{\sigma} = \frac{4-2}{4} = 0{,}5 \quad \text{ve} \quad z_a = \frac{c-\mu}{\sigma} = \frac{1-2}{4} = -0{,}25$$

hesaplanır. Böylece bu olasılık değeri,

$$P_r(1 \le x \le 4) = P_r(-0{,}25 \le z \le 0{,}5)$$

şeklinde standart normal dağılıma sahip z rastgele değişkene göre ifade edilmiş olunur. Sonra ilgili standart normal dağılım tablosundan 0,5 ve 0,25 sınır değerlerine karşılık gelen olasılık değerleri alınarak;

$$\begin{aligned} P_r(1 \le x \le 4) &= P_r(-0{,}25 \le z \le 0{,}5) = \\ &= F(0{,}5) - F(-0{,}25) = \\ &= F(0{,}5) - \{1 - F(0{,}25)\} = \\ &= 0{,}6915 + 0{,}5987 - 1 = 0{,}2902 \end{aligned}$$

hesaplanır.

Sonuç : $P_r(1 \le x \le 4) = 0{,}2902$ olarak bulunur.

Örnek 7: $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \le c) = 0{,}95$ olması için $c = ?$ sınır değerinin ne olması gerektiğini hesaplayınız.

Çözüm 7: Bu problemin çözümü için yukarıda yapılan çözümlerin tersi yönde ve sırada işlemler yapılır. Bunun için *S=0,95* olasılığa sahip standart normal dağılımına ilişkin;

$$z = \frac{x-\mu}{\sigma}$$

rastgele değişkeninin,

$$z_s = \frac{c-\mu}{\sigma} \to N(0\,,1)$$

sınır değeri ile ilgili normal dağılım tablosundan 1,645 olarak alınır. Sonra $z_s = 1{,}645$ olduğundan,

$$\frac{c-\mu}{\sigma} = 1{,}645$$

eşitliği kurularak çözülür. Sonuçta $P_r(x \le c) = 0{,}95$ için c sınır değeri

$$c = 1{,}645\sigma + \mu$$

olarak elde edilir. Bu bağıntıda $\mu = 2$ *ve* $\sigma = 4$ değerleri yerlerine yazılmakla; $c = 8{,}56$ olarak hesaplanır.

Sonuç : $c = 8{,}56$ olarak hesaplanmış olur.

Örnek 8: $x \to N(2,\, 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \ge c) = 0{,}2$ olması için $c = ?$ sınır değerini hesaplayınız.

Çözüm 8: Bu problemin çözümü için $P_r(x \ge c) = 1 - P_r(x \le c)$; şeklinde yazılabilir. Burada değişken dönüşümünden,

$$z = \frac{x-\mu}{\sigma}$$

rastgele değişkeninin,

$$z_s = \frac{c-\mu}{\sigma} \to N(0\,,1)$$

sınır değeri olur. Bu sınır değerine göre;

$$P_r(x \ge c) = 1 - P_r(x \le c) = 1 - F(z_s) = 0{,}2$$
$$F(z_s) = 1 - 0{,}2 = 0{,}80$$

Bu olasılığa karşılık gelen tablo değeri standart normal dağılım tablosundan $z_s = 0{,}842$ olarak alınır. Sonuçta;

$$z_s = \frac{c-\mu}{\sigma} = \frac{c-2}{4} = 0{,}842$$

eşitliği kurularak çözümünden, c sınır değeri,

$$c = 0{,}842\sigma + \mu$$

ve burada, $\mu = 2$ *ve* $\sigma = 4$ oldukları dikkate alınarak,

$$c = 0{,}842\sigma + \mu = (0{,}842)4 + 2 = 5{,}368$$

olarak bulunur.

Sonuç : $c = 5{,}368$ olarak hesaplanmış olur.

Örnek 9: $x \rightarrow N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(-c \langle x \leq 10) = 0{,}5$ olması için $c = ?$ sınır değerini hesaplayınız.

Çözüm 9: Bu problemin çözümü için, yukarıdaki işlemlere benzer şekilde hareket ederek,

$$z = \frac{x - \mu}{\sigma}$$

rastgele değişkeninin sınır değerleri,

$$z_a = \frac{x - \mu}{\sigma} = \frac{-c - 2}{4} \quad ; \quad z_ü = \frac{x - \mu}{\sigma} = \frac{10 - 2}{4} = 2$$

olarak hesaplanır. Bu sınır değerlerine göre;

$$P_r(-c \langle x \leq 10) = P_r(z_a \langle x \leq z_ü) = F(z_ü) - F(z_a) =$$
$$= F(2) - F(-\frac{c+2}{4}) = F(2) - \left\{1 - F(\frac{c+2}{4})\right\} = 0{,}5$$

ifadesi yazılır. Burada, gerekli işlemlerin yapılması neticesinde,

$$F(2) - 1 + F(\frac{c+2}{4}) = 0{,}5$$

elde edilmiş olur. Daha sonra, standart normal dağılım tablosundan alınan $F(2) = 0.9772$ değerinin yerine yazılması ile

$$F(\frac{c+2}{4}) = 0{,}5228$$

bulunur. Normal dağılım tablosundan 0,5228 karşılık gelen değişken değeri 0,057 alınarak, c sınır değeri; $c = 0.057\sigma + \mu = (0{,}057)4 + 2 = 2{,}228$ olarak hesaplanır.

Sonuç : Burada c sınır değeri : $c = 2{,}228$ olur.

Örnek 10: $x \rightarrow N(-2,\ 0{,}25)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = -2$ *ve* $\sigma^2 = 0{,}25$ *dır)* x rastgele değişkeninin $P_r(-2 - c \langle x \leq -2 + c) = 0{,}9$ olması için $c = ?$ sınır değerini hesaplayınız.

Çözüm 10: Yukarıdaki işlemlere benzer şekilde hareket ederek,

$$z = \frac{x-\mu}{\sigma}$$

rastgele değişkeninin sınır değerleri,

$$z_a = \frac{-2-c+2}{0,5} = -2c \qquad ; \qquad z_{ü} = \frac{-2+c+2}{0,5} = 2c$$

olarak hesaplanır. Bu sınır değerlerine göre;

$$\begin{aligned} P_r(-2-c \langle x \le -2+c) = P_r(z_a \langle x \le z_{ü}) &= \\ &= F(z_{ü}) - F(z_a) = \\ &= F(2c) - F(-2c) = \\ &= F(2c) - \{1 - F(2c)\} = \\ &= 2F(2c) - 1 = 0,9 \end{aligned}$$

ve buradan

$$F(2c) = \frac{1+0,9}{2} = 0,95$$

değerine göre; ilgili normal dağılım tablosundan $2c = 1,645$ alınarak buradan c sınır değeri,

$$c = \frac{1.645}{2} = 0.823$$

olarak hesaplanır.

Sonuç : $c = 0,823$ olarak bulunmuş olur.

Örnek 11: $x \to N(\mu,\ \sigma^2)$ normal dağılıma sahip x rastgele değişkeninin

$$P_r(|x-\mu| \le c-\mu) = P_r(-(c-\mu) \le x-\mu \le c-\mu) = 0,95$$

olması için $c =$? değerinin ne olması gerektiğinin hesaplanması sorulmaktadır.

Çözüm 11 : Bu problemde de benzer yollar izlenerek çözüm gerçekleştirilebilir. Şöyle ki : $P_r(|x-\mu| \le c-\mu) = P_r(-(c-\mu) \le x-\mu \le c-\mu) = 0,95$ şeklinde yazılır. Burada rastgele değişken ve alt, üst sınır değerleri için

$$z = \frac{x-\mu}{\sigma} \qquad ; \qquad z_a = -\frac{c-\mu}{\sigma} \qquad ; \qquad z_{ü} = \frac{c-\mu}{\sigma}$$

şeklinde hesaplanarak bu ifade;

$$P_r(z_a \le z \le z_{ü}) = 0,95$$

olarak yazılabilir. Sonuçta eşitlikleri yerlerine yazmakla,

$$F(z_{ü}) - F(z_a) = F(\frac{c-\mu}{\sigma}) - F(-\frac{c-\mu}{\sigma}) = 0{,}95$$
$$F(\frac{c-\mu}{\sigma}) - \left[1 - F(\frac{c-\mu}{\sigma})\right] = 0{,}95$$
$$F(\frac{c-\mu}{\sigma}) = \frac{1-0{,}95}{2} = 0{,}975$$

standart normal dağılım tablosundan 0,975 olasılık değerine karşılık gelen tablo değeri alınarak,

$$\frac{c-\mu}{\sigma} = 1{,}96$$

ve buradan gerekli işlemler yapılarak c sınır değeri,

$$c = 1{,}96\sigma + \mu$$

olarak elde edilir.

Sonuç : $c = 1{,}96\sigma + \mu$ olarak hesaplanmış olur.

Örnek 12: Herhangi bir okuldaki 500 kız öğrencinin boylarının ortalama uzunluğu 151cm. ve standart sapması da 15 cm. olarak verilmektedir. Bu öğrencilerin boylarının uzunlukları normal dağılımda oldukları kabul edilerek, bu öğrencilerden;

a) *kaç tanesinin boy uzunlukları 120cm. ile 155 cm arasındadır,*
b) *ve kaç tanesi 186 cm. den daha uzundur?* Hesaplayınız.

Çözüm 12:

a) Boyları 120cm. ile 155 cm arasında olanlarının sayısı x ile gösterilirse,

$$P_r(120 \le x \le 155) = P_r(119{,}5 \langle x \langle 155{,}5) = ?$$

olasılığı için, bu veriler normal dağılımda olduğundan, $z = \frac{x-\mu}{\sigma}$ rastgele değişken dönüşümü yapılarak;

$$z_a = \frac{119{,}5-\mu}{\sigma} = \frac{119{,}5-151}{15} = -2{,}10$$
$$z_{ü} = \frac{155{,}5-\mu}{\sigma} = \frac{155{,}5-151}{15} = 0{,}30$$

şeklinde sınır değerleri hesaplanır. Bu değerlere göre; $P_r(120 \le \mu \le 155) = ?$ olasılığı standart normal dağılıma sahip z rastgele değişkeni cinsinden,

$$P_r(z_a \langle z \langle z_{ü}) = ?$$

olarak ifade edilir. Buradan, standart normal dağılım tabloları kullanılarak,

$$\begin{aligned} P_r(z_a \langle z \langle z_{ü}) &= F(z_{ü}) - F(z_a) = F(0{,}30) - F(-2.10) \\ &= F(0{,}30) - (1 - F(2.10) = F(0{,}30) + F(2.10) - 1 = \\ &= 0{,}9821 + 0{,}6179 - 1. = 0{,}6000 \end{aligned}$$

olarak bulunur. Diğer bir ifade ile bu okuldaki 500 kız öğrenciden boyları 120cm. ile 155 cm. arasında olanlarının bu okuldaki tüm kız öğrencilerinin %60 kadar olmaktadır.

Sonuç 12:

a) Buradan, okuldaki tüm öğrencilerin sayısı 500 olduğuna göre boyları 120cm. ile 155 cm. arasında olanlarının kız öğrencilerin sayısı *500 x 0,60=300* olarak bulunur.

b) Benzer şekilde; 186 cm. den daha uzun olan öğrencilerin sayıları da,

$$P_r(x \geq 186) = P_r(x \rangle 185{,}5) = ?$$

olasılık ifadesinde

$$z = \frac{x - \mu}{\sigma}$$

şeklinde rastgele değişken dönüşümü yapılara,

$$z_s = \frac{185{,}5 - \mu}{\sigma} = \frac{185{,}5 - 151}{15} = 2{,}30$$

sınır değerlerine göre;

$$\begin{aligned} P_r(x \geq 186) = P_r(z \rangle z_s) = P_r(z \rangle 2{,}30) &= \\ &= 1 - F(2{,}30) = 1 - 0{,}9893 = \\ &= 0{,}0107 \end{aligned}$$

hesaplanan %0,107 yüzde değerine karşılık gelenlerinin sayısı *500(0,0107)=5* olarak bulunur. Bu öğrencilerden *5* tanesinin boyu 186 cm. ve daha uzundur.

Örnek 13: Bir fabrikada üretilen 200 adet çamaşır makinesi için ortalama iç yarıçapları 0,502 cm. ve standart sapması da 0.005 cm olarak verilmektedir. Bu çamaşır makinelerinin iç yarıçapları normal dağılıma sahip olduğu bilindiğine göre, iç yarıçapları 0,496cm. ile 0,508 cm. arasında olanları sağlam, diğerleri bozuk kabul edilmektedir. Bu fabrikada bozuk olarak üretilen çamaşır makinelerinin yüzde kaçı bozuktur. Ve sayılarını hesaplayınız.

Çözüm 13: Bu fabrikada sağlam üretilen çamaşır makinelerinin sayısı x ile gösterilirse, sağlam olarak üretilen çamaşır makinelerinin yüzdesi;

$$P_r(0{,}496 \langle x \langle 0{,}508) = ?$$

olasılığı için,

$$z_a = \frac{0{,}496-\mu}{\sigma} = \frac{0{,}496-0{,}502}{0{,}005} = -1{,}2$$

$$z_ü = \frac{0{,}508-\mu}{\sigma} = \frac{0{,}508-0{,}502}{0{,}005} = 1{,}2$$

sınır değerlerine göre,

$$P_r(0{,}496 \langle x \langle 0{,}508) = P_r(z_a \langle z \langle z_ü) = F(1{,}2) - F(-1{,}2) =$$
$$F(1{,}2) + F(1{,}2) - 1 = 2F(1{,}2) - 1 = 2(0{,}8849) - 1 = 0{,}7698$$

şeklinde %0,7698 olarak bulunur.

Buna göre, bozuk olan çamaşır makinelerinin yüzdesi de %100-%0,7698=%2302 $\cong$ %23 olarak hesaplanır. Buradan bozuk çamaşır makinelerinin sayısı de 200(0,23)=46 olarak elde edilir.

Sonuç : Bozuk çamaşır makinelerinin yüzdesi *%23* ve sayısı da *46* adet olur.

6.1.6 Çok Değişkenli Normal Dağılım

Buraya kadar tek rastgele değişkenle ilgili normal dağılım ele alınmaktadır. Ayni anda birden çok rastgele değişken birlikte ele alınarak tahmin edilmeleri söz konusu olduğundan, normal dağılım bu değişkenlerin hepsini birden karakterize edeceğinden böyle bir normal dağılım çok değişkenli normal dağılım olarak adlandırılmaktadır. Bu duruma bir örnek olarak, bütün noktalarının koordinatları birlikte tahmin edilen ve bunların çok değişkenli rastgele değişken oldukları düşünülen bir jeodezik kontrol ağı verilebilir. Neticede böyle bir durum için; m adet rastgele değişkeniçeren m-boyutlu çok değişkenli normal dağılımın olasılık fonksiyonu;

$$f(X) = C \ \exp\left[-\frac{(X-U)^T \Sigma_X^{-1}(X-U)}{2}\right] \qquad 6\text{-}68$$

olarak verilebilir. Burada; rastgele değişkenler vektörü;

$$X^T = [x_1 \quad x_2 \quad .... \quad x_m] \qquad 6\text{-}69$$

onunla ilgili tahmin parametreleri vektörü;

$$U^T = [\mu_1 \quad \mu_2 \quad .... \quad \mu_m] \qquad 6\text{-}70$$

ve kovaryans matrisi de;

$$\Sigma_X = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & & \sigma_{x_1 x_m} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & & \sigma_{x_2 x_m} \\ . & & & \\ . & & & \\ \sigma_{x_m x_1} & \sigma_{x_m x_2} & \ldots & \sigma_{x_m}^2 \end{bmatrix} \qquad 6\text{-}71$$

ve C sabiti de;

$$C = \frac{\{\det(\Sigma_X^{-1})\}^{1/2}}{(2\pi)^{m/2}} \qquad 6\text{-}72$$

göstermektedir.

Buradan, (6-61) tek değişkenli ve (6-68) de çok değişkenli normal dağılımların olasılık fonksiyonları arasındaki benzerlik ilişkileri;

Tablo 3: Tek ve çok değişkenli normal dağılımın olasılık fonksiyonları arasındaki benzerlik ilişkileri

Tek değişkenli normal dağılım olasılık fonksiyonu terimleri	*Çok değişkenli normal dağılım olasılık fonksiyonu terimleri*
$(\frac{1}{\sigma^2})^{1/2}$	$\{\det(\Sigma_X^{-1})\}^{1/2}$
$\frac{1}{(2\pi)^{1/2}}$	$\frac{1}{(2\pi)^{m/2}}$
$\frac{(x-\mu)^2}{2\sigma^2}$	$\frac{(X-U)^T \Sigma_X^{-1} (X-U)}{2}$

şeklinde bir tablo halinde özetlenebilir. Buradan görüldüğü gibi; sıfır ortalama *(U=0)* için çok değişkenli normal dağılım olasılık fonksiyonu;

$$F(X) = C\ \exp(-\frac{X^T \Sigma_X^{-1} X}{2}) \qquad 6\text{-}73$$

olarak ifade edilebilir.

6.1.6.1 İki Değişkenli Normal Dağılım

İki farklı rastgele değişken olan $x \to N(\mu_x, \sigma_x^2)$ ve $y \to N(\mu_y, \sigma_y^2)$ parametrelerine göre farklı birer normal dağılıma sahip iseler, bunların birbirinden bağımsız olmaları halinde,(4-18a) göre bunların ortak dağılım fonksiyonları,

$$f(x,y) = f(x)f(y) \qquad 6\text{-}74$$

olarak ifade edilebilir. Burada, $f(x)$ yoğunluk fonksiyonu, $x \to N(\mu_x, \sigma_x^2)$ rastgele değişkeninin tek boyutlu,

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu_x}{\sigma_x})^2} \qquad 6\text{-}75$$

$f(y)$ de $y \to N(\mu_y, \sigma_y^2)$ rastgele değişkeninin tek boyutlu,

$$f(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y-\mu_y}{\sigma_y})^2} \qquad 6\text{-}76$$

yoğunluk yada olasılık fonksiyonlarını göstermektedir. Bu durumda; her iki değişkenin $f(x,y)$ bileşik olasılık fonksiyonu, $f(x,y) = f(x)f(y)$ özelliğinden faydalanarak,

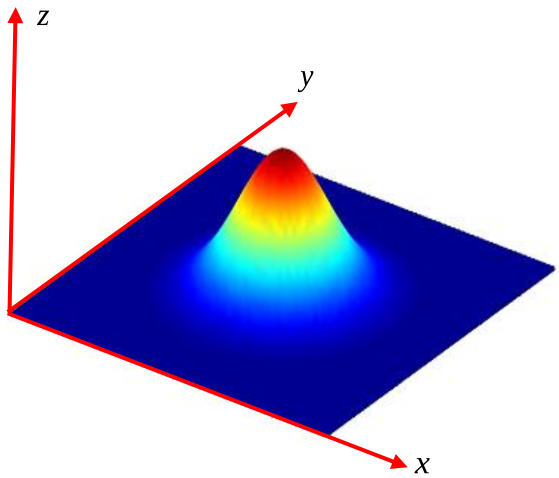
$$f(x.y) = f(x)f(y) = \left[\frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu_x}{\sigma_x})^2}\right]\left[\frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y-\mu_y}{\sigma_y})^2}\right] =$$

$$= \left[\frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left\{(\frac{x-\mu_x}{\sigma_x})^2 + (\frac{y-\mu_y}{\sigma_y})^2\right\}}\right] \qquad 6\text{-}78$$

şeklinde yazılabilir. Ancak, böyle bir işlem için her iki $x \to N(\mu_x, \sigma_x^2)$ ve $y \to N(\mu_y, \sigma_y^2)$ rastgele değişkenleri arasında ρ değerine sahip bir korelasyon var olduğu düşünülürse, sonuçta her iki değişkenle ilgili bileşik olasılık fonksiyonu,

$$f(x.y) = = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{(\frac{x-\mu_x}{\sigma_x})^2 + (\frac{y-\mu_y}{\sigma_y})^2 - 2\rho(\frac{x-\mu_x}{\sigma_x})(\frac{y-\mu_y}{\sigma_y})\right\}} \qquad 6\text{-}79$$

şeklindeki bir üstel fonksiyonla ifade edilebilir. (6-79) fonksiyonunun sergilediği geometrik şekil de,



Şekil 10: İki boyutlu dağılım fonksiyonu

olarak verilebilir(Şekil 10).

Burada, değişkenler arasındaki korelasyon değeri, varyans-kovaryans elemanlarından faydalanarak,

$$\rho = \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

bağıntısından hesaplanabilir. Bilindiği gibi, kuramsal anlamda bu şekilde tanımlanmış olan bir istatistik korelasyon değeri $-1 \le \rho \le +1$ sınır değerleri arasında çeşitli değerler alabilmektedir.

Uygulamada, böyle bir korelasyon değeri bilinemez. Bu gibi durumlarda, örnekleme verilerden elde edilmiş ve kuramsal korelasyon yerine kullanılabilecek en muhtemel, yani deneysel korelasyon değerleri alınır. Böyle bir değer, *n* adet iki farklı veri kümesinden,

$$\bar{x} = \frac{[x]}{n} \qquad \text{ve} \qquad \bar{y} = \frac{[y]}{n} \qquad \text{6-80a}$$

$$s_x = \pm\sqrt{\frac{\left[(x-\bar{x})^2\right]}{n-1}} \quad ; \quad s_y = \pm\sqrt{\frac{\left[(y-\bar{y})^2\right]}{n-1}} \quad \text{ve} \quad s_{xy} = \frac{\left[(x-\bar{x})(x-\bar{x})\right]}{n-1} \qquad \text{6-80b}$$

ortalama değerleri ile bunların deneysel standart sapma ve kovaryans değerlerinden faydalanılarak,

$$\rho_{xy} \approx r_{xy} = \frac{s_{xy}}{s_x s_y} \qquad \text{6-81}$$

şeklinde hesaplanabilir.

Eğer $\rho_{xy} \approx r_{xy} = 0$ ise; her biri farklı normal dağılıma sahip bu iki $x \to N(\mu_x, \sigma_x^2)$ ve $y \to N(\mu_y, \sigma_y^2)$ rastgele değişkenleri birbirinden bağımsızdır denir. Aksi halde korelasyonlu olurlar.

Her iki rastgele değişken arasındaki varyans-kovaryans ilişkileri, matris gösterimi kullanılarak,

$$\Sigma = Cov(x,y) = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \qquad \text{6-82a}$$

şeklinde ve bu matrisin standartlaştırılmasında da, buna karşılık gelen korelasyon matrisi,

$$R = \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix} \qquad \text{6-82b}$$

olarak ifade edilebilir. Ayrıca, böyle bir matrisin determinantı; varyans-kovaryans matrisi için,

$$\det|\Sigma| = \sigma_x^2\sigma_y^2 - \sigma_{xy}^2 = \sigma_x^2\sigma_y^2(1 - \frac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2}) = \sigma_x^2\sigma_y^2(1-\rho_{xy}^2) = \sigma_x^2\sigma_y^2 \det|R| \qquad \text{6-83a}$$

ve korelasyon matrisi için de

$$\det|R| = 1 - \rho_{xy}^2 = 1 - \rho^2 \qquad \text{6-83b}$$

olur. Diğer taraftan, değişkenlerde

$$\underline{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{ve} \quad \underline{\mu} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \qquad \text{6-84}$$

şeklinde tanımlanır ve işlemlerde $\det|\Sigma| = \sigma_x^2\sigma_y^2(1-\rho_{xy}^2)$ olduğu dikkate alınırsa,

$$(\underline{x}-\underline{\mu})^T \underline{\Sigma}^{-1}((\underline{x}-\underline{\mu}) = \begin{bmatrix} x-\mu_x & y-\mu_y \end{bmatrix} \frac{1}{\det|\underline{\Sigma}|} \begin{bmatrix} \sigma_y^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_x^2 \end{bmatrix} \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix} =$$

$$= \frac{1}{\det|\underline{\Sigma}|} \left\{ \sigma_x^2 (x-\mu_x)^2 + 2\sigma_{xy}(x-\mu_x)(y-\mu_y) + \sigma_y^2 (y-\mu_y)^2 \right\} =$$

$$= \frac{1}{\sigma_x^2\sigma_y^2(1-\rho_{xy}^2)} \left\{ (\frac{x-\mu_x}{\sigma_x})^2 + 2\rho(\frac{x-\mu_x}{\sigma_x})(\frac{y-\mu_y}{\sigma_y}) + (\frac{y-\mu_y}{\sigma_y})^2 \right\} \qquad \text{6-85}$$

bağıntısı elde edilmiş olur. Neticede, her iki rastgele değişkenin bileşik yoğunluk fonksiyonu için de,

$$f(x,y)=\frac{1}{(2\pi)^{2/2}|\underline{\Sigma}|^{1/2}}e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})} \qquad 6\text{-}86$$

yada bir diğer gösterim şekli olan,

$$f(x,y)=\frac{1}{(2\pi)|\underline{\Sigma}|^{1/2}}Exp-\frac{1}{2}(\underline{x}-E\{\underline{x}\})^T\underline{\Sigma}^{-1}(\underline{x}-E\{\underline{x}\}) \qquad 6\text{-}87$$

bağıntıları yazılabilir. Burada iki rastgele değişkenli bir rastgele vektörler için söylenenler; n rastgele değişkenli bir vektöre genişletilerek söylenirse; durum,

$$\underline{x}^T=[x_1 \quad x_2 \quad .... \quad x_n] \quad \text{ve} \quad \underline{\mu}^T=[\mu_1 \quad \mu_2 \quad .... \quad \mu_n] \qquad 6\text{-}88$$

olmak üzere, n değişkenli bir rastgele vektör için varyans-kovaryans ilişkileri de,

$$\underline{\Sigma}=E\{(\underline{x}-E\{\underline{x}\})(\underline{x}-E\{\underline{x}\})^T\} \qquad 6\text{-}89a$$

şeklinde yada bunun daha açık şekildeki bir ifadesi olarak,

$$\underline{\Sigma}=\begin{bmatrix}\sigma_1^2 & \sigma_{12} & .... & \sigma_{1n}\\ \sigma_{12} & \sigma_2^2 & .... & \sigma_{2n}\\ \vdots & \vdots & .... & \vdots\\ \sigma_{1n} & \sigma_{2n} & .... & \sigma_n^2\end{bmatrix} \qquad 6\text{-}89b$$

matrisi ile ifade edilebilir.

Özel durumda; her iki değişkenin bağımsız olmaları halinde, aralarındaki korelasyon $\rho_{xy}=0$ sıfır olacağından, bunlarla ilgili varyans kovaryans matris de,

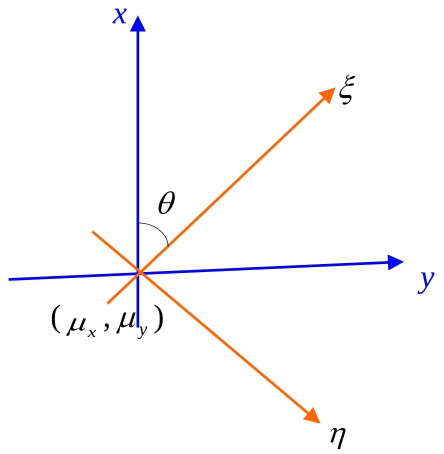
$$\underline{\Sigma}=\begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_y^2\end{bmatrix} \qquad \text{ve} \qquad \underline{\Sigma}^{-1}=\begin{bmatrix}1/\sigma_x^2 & 0\\ 0 & 1/\sigma_y^2\end{bmatrix} \qquad 6\text{-}89c$$

biçiminde köşegen bir matris olur. Bu durumda, bileşik olasılık fonksiyonu da

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}(x-\mu_x)^2+(y-\mu_y)^2} \qquad 6\text{-}90$$

olarak verilebilir.

Şimdi, orijin noktasının koordinatları μ_x ve μ_y olan bir x, y düzlem dik koordinat sistemi seçilerek, bunun eksenleri saat ibresi yönünde θ açısı kadar döndürülse,



Şekil 11: Düzlem dik koordinat sistemleri

eksenleri ξ ve η olan diğer düzlem dik koordinat sistemi elde edilmiş olur(Şekil 11). Bunların arasında,

$$\begin{aligned} x-\mu_x &= \xi\cos\theta-\eta\sin\theta \\ y-\mu_y &= \xi\sin\theta-\eta\cos\theta \end{aligned} \qquad \text{6-91a}$$

veya tersi çözümden,

$$\begin{aligned} \xi &= (x-\mu_x)\cos\theta+(y-\mu_y)\sin\theta \\ \eta &= (x-\mu_x)\sin\theta+(y-\mu_y)\cos\theta \end{aligned} \qquad \text{6-91b}$$

yazılabilir. Böyle bir koordinat dönüşümünün olasılık fonksiyonu da,

$$f(x,y)=\frac{1}{(2\pi)\left|\underline{\Sigma}\right|^{1/2}}\,Exp-\frac{1}{2}(\underline{x}-\underline{\mu}_x)^T\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_x) \qquad \text{6-92}$$

şeklindeki bir fonksiyon olur. Burada, yukarıdaki koordinat dönüşümü için,

$$A=\begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix} \qquad \text{6-93a}$$

olmak üzere,

$$\left|\underline{x}-\underline{\mu}\right| = A\,y \qquad \text{6-93b}$$

ve buradan da

$$y = A^{-1}\left|\underline{x}-\underline{\mu}\right| \qquad \text{6-93c}$$

matris bağıntıları yazılabilir. Daha sonra, bu bağıntılarda varyans-kovaryansların yayılması kuralı uygulandığında,

$$\underline{\Sigma}_x = A\underline{\Sigma}_y A^T \qquad \text{6-94a}$$

ve bu ifadenin inversinin alınmasından da,

$$\underline{\Sigma}_y = A^{-1}\underline{\Sigma}_x (A^{-1})^T \qquad \text{6-94b}$$

elde edilir. Burada A matrisi *ortagonal* bir matris olduğundan

$$A^{-1} = A^T \text{ ve } \det|A| = 1$$

olur. Ayrıca matrisin ortogonal bir matris olması neniyle de; satır veya sütun elemanlarının kareleri toplamı bire, karşılıklı satır ya da sütun elemanlarının çarpımının toplamı sıfır eşit olmaktadır. Bu özellik dikkate alınarak, *f(x,y)* bileşik olasılık fonksiyonunun üstel kısmı için,

$$\begin{aligned} &Exp - \frac{1}{2}\left\{(\underline{x} - \underline{\mu}_x)^T \underline{\Sigma}^{-1} (\underline{x} - \underline{\mu}_x)\right\} = \\ &Exp - \frac{1}{2}\left\{y^T A^T (A^T)^{-1} \underline{\Sigma}_y^{-1} A^{-1} A y\right\} \end{aligned} \qquad \text{6-95}$$

bağıntısı yazılabilir. Buradan, her iki değişkenle ilgili bileşik olasılık fonksiyonu,

$$f(x, y) = f(\xi, \eta) = \frac{1}{(2\pi)|\underline{\Sigma}|^{1/2}} Exp - \frac{1}{2}(y^T \underline{\Sigma}_y^{-1} y) \qquad \text{6-96}$$

şeklinde yazılabilir.

Neticede; burada verilmiş olan (6-96) bileşik olasılık fonksiyonunu aynı zamanda dönüştürülmüş koordinatların da bileşik olasılık fonksiyonuna eşit olduğu açıkça görülmektedir. Diğer tarafta, $A^{-1} = A^T$ olduğu dikkate alınarak,

$$\underline{\Sigma}_y = A^{-1}\underline{\Sigma}_x A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \sigma_\xi^2 & \sigma_{\xi\eta} \\ \sigma_{\xi\eta} & \sigma_\eta^2 \end{bmatrix} \qquad \text{6-97}$$

işlemler sonucunda,

$$\begin{aligned} \sigma_\xi^2 &= \sigma_x^2 \cos^2\theta + \sigma_y^2 \sin^2\theta \\ \sigma_\eta^2 &= \sigma_x^2 \sin^2\theta + \sigma_y^2 \cos^2\theta \end{aligned} \qquad \text{6-98a}$$

ve buradan da

$$\begin{aligned} \sigma_\xi^2 - \sigma_\eta^2 &= (\sigma_y^2 - \sigma_x^2)\cos 2\theta \\ \sigma_{\xi\eta} &= (\sigma_y^2 - \sigma_x^2)\sin\theta\cos\theta = \frac{1}{2}(\sigma_y^2 - \sigma_x^2)\sin 2\theta \end{aligned} \qquad \text{6-98b}$$

yazılarak, birbirlerine oranlamasından,

$$\tan 2\theta = \frac{2\sigma_{\xi\eta}}{\sigma_\xi^2 - \sigma_\eta^2} \qquad \text{6-99}$$

her iki koordinat sistemi arasındaki θ dönüklük açısı hesaplanabilir. Buradan görüldüğü gibi, korelasyonsuz koordinat çiftlerinden, korelasyonlu yeni bir koordinat çiftleri elde edilmiş olur. Bu durumun tersi yönde, korelasyonlu iki değişkenden bağımsız iki değişkene dönüşüm yapmak için, koordinat eksenleri θ dönüklük açısı kadar döndürülmesi gerekir. Aynı şekilde, $f(\xi,\eta)$ olasılık fonksiyonu bağımlı olarak verilmiş, buradan $f(x,y)$ bağımsız olasılık fonksiyonunu elde etmek için ξ , η koordinat sisteminin her iki eksenleri de θ dönüklük açısı kadar döndürmekle bunların bağımsız hale getirilebilmesi mümkün olmaktadır.

Aynı dönüşüm özellikleri, koordinat ekseni yönündeki gerçek hatalar içinde düşünülebilir. Bu amaçla, burada değişken olarak kullanılmış koordinat eksenleri yerine, her biriyle ilgili,

$$\varepsilon_x = x - \mu_x \qquad \text{6-100a}$$

$$\varepsilon_y = y - \mu_y \qquad \text{6-100b}$$

gerçek hataları ele alınarak benzer işlemler yapılırsa, bunlardan her biri

$$\varepsilon_x \rightarrow N(0,\sigma_x^2)$$

$$\varepsilon_y \rightarrow N(0,\sigma_y^2)$$

parametrelerine göre normal dağılımda olup, bileşik olasılık fonksiyonu,

$$f(\varepsilon_x,\varepsilon_y) = = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2-2\rho(\frac{\varepsilon_x}{\sigma_x})(\frac{\varepsilon_y}{\sigma_y})\right\}} \qquad \text{6-101}$$

şeklinde ifade edilebilir.

Daha önce de sözü edilmiş olduğu gibi burada, korelasyonlu gerçek hata değerleri için yazılmış olan bu bağıntıyı, korelasyon değerini sıfır yapmak, yani; korelasyonsuz hale dönüştürmek için, koordinat eksenler θ açısı kadar çevirmek yeterli olur.

Uygulamalarda, bu formüllerde geçen σ_x ve σ_y kuramsal değerleri bilinemez. Bunların yerine,

$$\sigma_x^2 \approx s_x^2 = \frac{[\varepsilon_x\varepsilon_x]}{n} \qquad yada \qquad s_x^2 = \frac{1}{n}\underline{\varepsilon}_x^T\underline{\varepsilon}_x \qquad \text{6-102a}$$

ve

$$\sigma_y^2 \approx s_y^2 = \frac{[\varepsilon_y\varepsilon_y]}{n} \qquad yada \qquad s_y^2 = \frac{1}{n}\underline{\varepsilon}_y^T\underline{\varepsilon}_y \qquad \text{6-102b}$$

şeklinde toplam veya matris gösterimi ile ifade edilmiş olan formüller yardımıyla hesaplanan yaklaşık değerleri kullanılarak ilgili işlemler yapılır.

Buraya kadar yapılan açıklamaların neticesinde; bağımsız ya da korelasyonsuz değişkenlerle ilgili bileşik olasılık fonksiyonu da benzer bir düşünce ile

$$f(\varepsilon_x,\varepsilon_y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2\right\}} \qquad \text{6-103}$$

olarak ifade edilebilir. Burada, bağımsız değişkenlerin bileşik olasılık fonksiyonun üst kısmı ele alınarak,

$$-\frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2\right\}=sabit=k^2 \qquad \text{6-104}$$

bir sabite eşitlendiğinde, bu bağıntı bir elips denklemi olduğu görülür. Bu elipslerden herhangi biri;

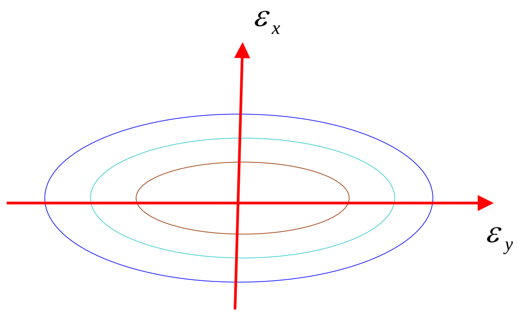
- *Elips eğrisi üzerindeki noktaların olasılıkları her zaman eşittir,*
- *Eksenleri; en büyük ve en küçük hata doğrultularını,*
- *Yapılan hatanın büyüklüğünü,*
- *İlgili noktaların presizyonu*

gösterirler şeklinde bazı özelliklere sahip olurlar.

Ayrıca; her bir elips A büyük yarı ve B küçük yarı eksenleri ile tanımlı olurlar. Böyle bir elipsin büyük ve küçük yarı eksenleri,

$$A=\sigma_x\, k\,\sqrt{2} \quad \text{ve} \quad B=\sigma_y\, k\,\sqrt{2}$$

değerinde olur. Bunlar, k parametresinin farklı şekildeki seçilecek değerlerine bağlı olarak her biri iç içe farklı büyüklükte elipsler olurlar.



Şekil 12: Hata elipsleri

Sonuçta; her birinin olasılılık fonksiyonu,

$$f(\varepsilon_x,\varepsilon_y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-k^2} \qquad \text{6-105}$$

gibi bir formülle ifade edilir. Diğer tarafta, n boyutlu bir dağılım fonksiyonunun diferansiyeli;

$$dF=dp=f(\underline{x})dx_1,dx_2,\ldots\ldots,dx_n \qquad \text{6-106}$$

olduğundan, benzer bir düşünce ile ε_x ve ε_y hataları bileşik dağılım fonksiyonu için,

$$dF = dp = \frac{1}{2\pi\sigma_x\sigma_y}\; e^{-k^2} d\varepsilon_x \;\; d\varepsilon_y \qquad \text{6-107}$$

ifadesi yazılabilir.

Bu bağıntı aynı zamanda olasılık fonksiyonu ile belirlenmiş bir elipsin alanı olmaktadır. Matematikten bilindiğine göre bir elipsin alanı,

$$F = A \; B \;\; \pi \qquad \text{6-108}$$

şeklinde ifade edilmektedir. Burada, A ve B eksenlerinin yukarda hesaplanan değerleri yerlerine yazıldığında, aynı elipsin eksenlere göre diğer bir ifadesi,

$$F = 2 \;\; \pi \;\; \sigma_{\mathrm{x}} \;\; \sigma_{\mathrm{y}} \;\;\; \mathrm{k}^2 \qquad \text{6-109}$$

olarak ifade edilebilir. Bu durumda,

$$dF = 4 \;\; \pi \;\; \sigma_{\mathrm{x}} \;\; \sigma_{\mathrm{y}} \;\;\; \mathrm{k} \;\; \mathrm{dk} \qquad \text{6-110}$$

olmaktadır.

Bu ifadenin yukarıda verilmiş olan diferansiyel anlamdaki alan bağıntısında yerine yazılması ile,

$$\begin{aligned} dF = dp &= \frac{1}{2\pi\sigma_x\sigma_y}\; e^{-k^2} d\varepsilon_x \;\; d\varepsilon_y = \frac{1}{2\pi\sigma_x\sigma_y}\; e^{-k^2} dF = \\ &= \frac{1}{2\pi\sigma_x\sigma_y}\; e^{-k^2}\, 4\pi\sigma_x\sigma_y k \; dk = 2 \;\; k \;\; e^{-k^2} dk \end{aligned} \qquad \text{6-111}$$

şeklinde bir formül elde edilmiş olur. Diferansiyel anlamdaki bu ifadenin entegralinden, herhangi bir elipsin alanı yada bu elipsin içine düşme olasılığı formülü,

$$P = 2 \int_0^{\mathrm{k}} k \;\; e^{-k^2} dk = 1 - e^{-k^2} \qquad \text{6-112}$$

olarak elde edilir. Bu elipsler, k 'nın farklı değerleri için hesaplanabilir. Bunlarla ilgili bazı hata elipsi örnekleri aşağıdaki tablodan görülmektedir.

Tablo 4: Çeşitli hata elipsleri

k	$P = 1 - e^{k^2}$ Olasılığı	Elipsin adı
0.70111	*0.39348*	*Helmert ortalama hata elipsi*
0.83225	*0.50000*	*Mühtemel hata elipsi*
0.92510	*0.57506*	*Mutlak ortalama hata elipsi*
1.00000	*0.63212*	*Andra ve Möhle ortalama hata elipsi*
1.07137	*0.68268*	*Wellisch ortalama hata elipsi*

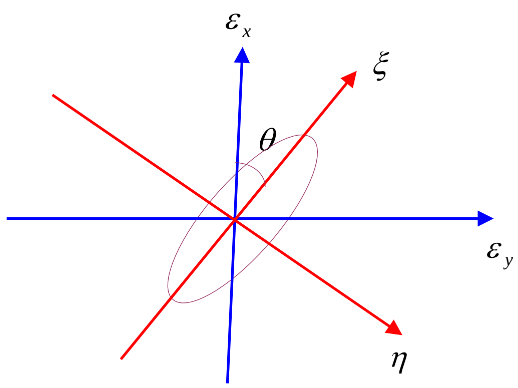
Benzer işlemler, korelasyonlu değişkenler için yapılırsa, her iki koordinat sisteminin eksenleri arasındaki dönüklük açısı,

$$\tan 2\theta = \frac{2\sigma_{xy}}{\sigma_y^2 - \sigma_x^2} \qquad \text{6-113}$$

kadar olur. Bu koordinat sistemleri arasındaki dönüşüm işlemi de,

$$\begin{aligned} \varepsilon_\xi &= \varepsilon_x \cos\theta + \varepsilon_y \sin\theta \\ \varepsilon_\eta &= -\varepsilon_x \sin\theta + \varepsilon_y \cos\theta \end{aligned} \qquad \text{6-114}$$

şeklinde ifade edilir.



Şekil 13: Hata elipsi

Burada, hata yayılması kuralı uygulanırsa;

$$\begin{aligned} \sigma_\xi^2 &= \sigma_x^2 \cos^2\theta + 2\sigma_{xy} \sin\theta\cos\theta + \sigma_y^2 \sin^2\theta \\ \sigma_\eta^2 &= \sigma_x^2 \sin^2\theta - 2\sigma_{xy} \sin\theta\cos\theta + \sigma_y^2 \cos^2\theta \end{aligned} \qquad \text{6-115}$$

olur. Sonuçta elipsin büyük ve küçük yarı eksenleri için,

$$\begin{aligned} A^2 &= 2\ \sigma_0^2\ Q_{\xi\xi}\ k^2 \\ B^2 &= 2\ \sigma_0^2\ Q_{\eta\eta}\ k^2 \end{aligned} \qquad \text{6-116}$$

bağıntılarından hesaplanır.

6.1.6.2 Üç yada Daha Fazla Değişkenli Normal Dağılımlar

Birbirinden farklı $x \rightarrow N(\mu_x, \sigma_x^2)$, $y \rightarrow N(\mu_y, \sigma_y^2)$, $z \rightarrow N(\mu_z, \sigma_z^2)$ gibi üç adet normal dağılıma sahip rastgele değişkenler bağımsız iseler aralarındaki korelasyon değeri de sıfır olmaktadır. O zaman bunların bileşik olasılık fonksiyonu, iki boyutlu normal dağılımınkine benzer şekilde,

$$f(x,y,z) = f(x)\,f(y)\,f(z) \qquad \text{6-117}$$

olarak yazılabilir. Burada, her bir değişkenle ilgili tek boyutlu olasılık fonksiyonları yerlerine yazılıp, gerekli işlemlerin sonucunda bileşik olasılık fonksiyonu,

$$f(x.y,z) = f(x)f(y)f(z) = \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} e^{-\frac{1}{2}\left\{(\frac{x-\mu_x}{\sigma_x})^2+(\frac{y-\mu_y}{\sigma_y})^2+(\frac{z-\mu_z}{\sigma_z})^2\right\}} \qquad \text{6-118}$$

şeklinde elde edilebilir. Burada,

$$\varepsilon_x = x - \mu_x \qquad \text{6-119a}$$

$$\varepsilon_y = y - \mu_y \qquad \text{6-119b}$$

$$\varepsilon_z = z - \mu_z \qquad \text{6-119c}$$

her bir değişken ile ilgili gerçek hatalar kullanılırsa, bunların bileşik olasılık fonksiyonu da

$$f(\varepsilon_x,\varepsilon_y,\varepsilon_z) = f(\varepsilon_x)f(\varepsilon_y)f(\varepsilon_z) = \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} e^{-\frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2+(\frac{\varepsilon_z}{\sigma_z})^2\right\}} \qquad \text{6-120}$$

şeklinde yazılabilir.

Yine, iki boyutlu normal dağılımlarda yapılanlara benzer şekildeki bir düşünce ile buradan,

$$-\frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2+(\frac{\varepsilon_z}{\sigma_z})^2\right\} = K^2 \qquad \text{6-121}$$

olduğu görülür. Bilindiği gibi, bu denklem bir elipsoid denklemi olmaktadır. Bu elipsoidin A, B, C yarı eksenleri;

$$A = \sqrt{2}\ K\ \sigma_x$$

$$B = \sqrt{2}\ K\ \sigma_y \qquad \text{6-122}$$

$$C = \sqrt{2}\ K\ \sigma_z$$

ve hacmi de,

$$V = \frac{4}{3}\pi\, A\,B\,C = \frac{8}{3}\sqrt{2}\,\pi\, K^3\sigma_x\sigma_y\sigma_z \qquad \text{6-123}$$

kadar olur. Aynı şekilde,

$$dp = \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} Exp - \frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2+(\frac{\varepsilon_z}{\sigma_z})^2\right\} d\varepsilon_x\, d\varepsilon_y\, d\varepsilon_z =$$

$$= \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} Exp - \frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2+(\frac{\varepsilon_y}{\sigma_y})^2+(\frac{\varepsilon_z}{\sigma_z})^2\right\} dv = \qquad \text{6-124}$$

yazılabilir.

Burada, dv hacim elemanını göstermektedir. Bu bağıntıda, verilmiş olan hacim bağıntısının diferansiyel değeri olan,

$$dv = 8\sqrt{2}\,\pi\,K^2\sigma_x\sigma_y\sigma_z dK \qquad \text{6-125}$$

hacim elemanı yerine yazılırsa,

$$P = \int_0^K \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} Exp - \frac{1}{2}\left\{(\frac{\varepsilon_x}{\sigma_x})^2 + (\frac{\varepsilon_y}{\sigma_y})^2 + (\frac{\varepsilon_z}{\sigma_z})^2\right\} 8\sqrt{2}\pi K^2\sigma_x\sigma_y\sigma_z dK =$$

$$= \int_0^K \frac{8\sqrt{2}\pi K^2}{(2\pi)^{3/2}} e^{-K^2} dK = \int \frac{4}{\sqrt{\pi}} K^2 e^{-K^2} dK = \frac{2}{\sqrt{\pi}}(\frac{2}{3}K^3 - \frac{4}{10}K^5 + \frac{6}{42}K^7 - ......) \qquad \text{6-126}$$

olasılık formülü elde edilmiş olur.

Böyle bir bağıntıyı kullanarak, K değerinin seçilmiş çeşitli değerlerine göre, bir noktanın elde edilen hata elipsoidi içine düşme olasılığı hesaplanabilir. Uygulamada, iki boyutlu dağılımlara karşılık burada da K ‘nın farklı değerleri için tanımlanmış hata elipsoitleri

Tablo 5: Çeşitli hata elipsleri

k	*P Olasılığı*	*Elipsin adı*
0.70111	*0.19874*	*Helmert ortalama hata elipsoidi*
1.08765	*0.50000*	*Mühtemel hata elipsoidi*
1.18137	*0.57506*	*Mutlak ortalama hata elipsoidi*
1.022474	*0.60837*	*Andra ve Möhle ortalama hata elipsoidi*
1.032791	*0.68268*	*Wellisch ortalama hata elipsoidi*

olarak elde edilir. Burada her biri ;

$$x \rightarrow N(\mu_x, \sigma_x^2) \quad , \quad y \rightarrow N(\mu_y, \sigma_y^2) \quad \text{ve} \quad z \rightarrow N(\mu_z, \sigma_z^2)$$

parametrelerine göre normal dağılıma sahip rastgele değişkenlerin korelasyonlu değişkenler olası halinde, bunların bileşik olasılık fonksiyonu,

$$f(x,y,z) = \frac{1}{(2\pi)^{3/2}|\underline{\Sigma}|^{1/2}} Exp - \frac{1}{2}(\underline{x} - \underline{\mu})^T \underline{\Sigma}^{-1}(\underline{x} - \underline{\mu}) \qquad \text{6-127}$$

olur. Burada kullanılan $\underline{\Sigma}$ varyans- kovaryans değerleri,

$$\underline{\Sigma} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{bmatrix} \qquad \text{6-128}$$

şeklindeki kare simetrik matrisin elemanları olarak verilmektedir. Bu durum, her biri farklı normal dağılıma sahip;

$$x_1 \rightarrow N(\mu_1, \sigma_1^2)$$
$$x_2 \rightarrow N(\mu_2, \sigma_2^2)$$
$$x_n \rightarrow N(\mu_n, \sigma_n^2)$$

gibi n adet rastgele değişkenli bileşik normal dağılıma genelleştirilirse;

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad \underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \quad \text{ve} \quad \underline{x} - \underline{\mu} = \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \\ \vdots \\ x_n - \mu_n \end{bmatrix} \qquad 6\text{-}129$$

olmak üzere, bileşik olasılık fonksiyonu,

$$f(\underline{x}) = \frac{1}{(2\pi)^{n/2} |\underline{\Sigma}|^{1/2}} Exp - \frac{1}{2} (\underline{x} - \underline{\mu})^T \underline{\Sigma}^{-1} (\underline{x} - \underline{\mu})$$

ve $\underline{\Sigma}$ varyans- kovaryans değerleri de,

$$\underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_{nn} \end{bmatrix}$$

biçiminde ifadeler olarak verilebilir.