6. BÖLÜM

PARAMETRİK DAĞILIM FONKSİYONLARI VE GENEL ÖZELLİKLERİ

6. İstatistikte Parametrik Dağılım Fonksiyonları ve Özellikleri

Bu bölümde nicel özellikli 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 dolaylı da olsa onunla ilgili diğer;

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

sırası ile ele alınacaktır. Bu amaçla, burada her bir dağılımla ilgili matematiksel özellikler özet alarak verildikten sonra, gerekli hesaplamaların veya ilgili tabloların nasıl kullanılacağına ilişkin bazı sayısal örnekler verilerek konu daha da kolay anlaşılır ve uygulanabilir olmasına ayrıca özen gösterilecektir.

6.1. Normal Dağılım

Literatürde yer aldığı şekliyle, matematik-istatistikte, ilk defa birbirinden habersiz olarak, *Charles S. Peirce, Francis Gaulton* ve *Wilhelm Lexis* tarafından *1872* yıllarında kullanıldığı söylenen *normal dağılım* sözcüğü, özetle gözlemlerin bir ortalama etrafındaki simetrik veya homojen yayılmalarının ölçüsü olarak tanımlanmaktadır. Ancak, daha sonraları bazı özel kullanım amaçları için bunun yanında *Gauss dağılım fonksiyonu* veya kısaca *çan eğrisi* terimlerinin de kullanıldığı ilgili kaynaklardan görülmektedir.

Pratikte, normal dağılıma sahip rastgele değişkenler için istatistik anlamı olan böyle bir kuramsal bilgi; sonlu sayıda elemanı içeren örnekleme veri kümesindeki gözlem değerlerinin dağılımının normal dağılımda olmasına ilişkin uygulanacak istatistik hipotez testleri açısından da oldukça fazla önem taşımaktadır. Böyle bir kuramsal bilgi, aynı zamanda normal dağılımla ilgili hipotez testlerinin de temelini oluşturmaktadır. Pratikte her zaman, bir olayla ilgili sonlu sayıda yapılmış gözlemelerden elde edilmiş örnekleme verilere ilişkin dağılımın normal dağılımda olduğunun bilinmiş olması, istatistikte güçlü testlerden *parametrik* ve *non-parametrik* testlerden herhangi birinin seçimi için de gerekli olan en önemli koşulun yerine getirilmesine yardımcı olmaktadır. Böyle bir fonksiyon bağımsız değişkenlerin sayısına göre;

- *Tek boyutlu (ya da tek değişkenli) normal dağılım fonksiyonları,*
- *Çok boyutlu (ya da çok değişkenli) normal dağılım fonksiyon*ları

biçiminde ele alınabilir. Tek boyutlu normal dağılım fonksiyonları düzlemsel bir alan tanımlarken, çok boyutlu normal dağılım fonksiyonları n boyutlu uzayda daima bir yüzey veya hacim belirlemektedir. her şeyden önce, bir deney 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 incelenmiştir.

6.1.1. Tek Değişkenli Normal Dağılım Fonksiyonu

Tek değişkenli bir normal dağılım fonksiyonu için böyle bir amaca yönelik analitik ifade,

$$I = \int_{-\infty}^{+\infty} \exp\,(-y^2/2)\; dy \qquad 6\text{-}1$$

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

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

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

$$\begin{bmatrix} y \\ z \end{bmatrix} = r\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad 6\text{-}3$$

kutupsal koordinat dönüşümü yapılarak, (6-2) bağıntısının bu koordinatlar cinsinden bir ifadesi olan,

$$I^2 = \int_0^{2\pi} \int_0^{+\infty} \exp(\frac{-r^2}{2}) \; r\,dr \; d\theta = \int_0^{2\pi} d\theta = 2\pi \qquad 6\text{-}4$$

integral bağıntısında, integral alınarak sınır değerlerinin yerine yazılması ile,

$$I^2 = 2\pi$$

biçiminde elde edilir. Neticede; buradan görüleceği gibi, (6-1) fonksiyonunun integral değeri de;

$$I = (2\pi)^{1/2} \qquad 6\text{-}5$$

kadar olur. (6-1)'deki integral işlemi sonucunda; değeri (6-5) de verildiği gibi $I = (2\pi)^{1/2}$ değerine eşit olduğundan; o zaman (6-1) deki integral bağıntısı da;

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

şeklinde ifade edilebilir. Bu şekildeki bir integral alma işleminde integral bağıntısının değişkeni için, *a* ve *b* herhangi bir sayı olmak üzere;

$$y = \frac{x-a}{b} \quad ; \quad b \rangle 0 \qquad 6\text{-}7$$

şeklinde bir değişken dönüşümü yapılırsa; sonuçta integral bağıntısı,

$$\int_{-\infty}^{+\infty} \frac{1}{b(2\pi)^{1/2}} \exp\,(-\frac{(x-a)^2}{2b^2})\; dx = 1 \qquad 6\text{-}8$$

şeklini alır.

Burada verilmiş olan (6-8) integral bağıntısı incelendiğinde, bunun bir rastgele değişkenle ilgili dağılım *(kümulatif dağılım)* fonksiyonu özelliklerine sahip olduğu açıkça görülmektedir. Buradan (6-8) bağıntısının türevinin alınması ile, böyle bir dağılım fonksiyonuna karşılık gelen rastgele değişkenlerin bir parametre etrafındaki yayılmasını temsil eden olasılık ya da yoğunluk *(frekans)* fonksiyonu; (6-8) integralinin diferansiyeli olan,

$$f(x)=\frac{1}{b(2\pi)^{1/2}}\;\exp\left\{-\frac{(x-a)^2}{2b^2}\right\}\quad \textit{burada}\quad -\infty \langle\; x \;\langle\; +\infty \qquad\qquad 6\text{-}9a$$

biçimindeki üstel bir fonksiyon olarak elde edilir. Burada; x değeri dağılım ve olasılık fonksiyonlarının değişken parametresini, a ve b değerleri de bu fonksiyonlarının istatistik 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 (6-9a) fonksiyonu ile verilmiş olan olasılık ya da diğer adıyla yoğunluk fonksiyonu,

$$f(x)=\frac{1}{\sigma(2\pi)^{1/2}}\;\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\quad \textit{burada}\quad -\infty \langle\; x \;\langle\; +\infty \qquad\qquad 6\text{-}9b$$

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. (6-9b) fonksiyonlarındaki x değişkeninin sürekli türden bir rastgele değişken olduğundan bu $f(x)$ olasılık fonksiyonu da benzer şekilde sürekli türden bir rastgele değişkenin olasılık ya da frekans fonksiyonunu göstermektedir.

Aynı şekilde, (6-9b) olasılık fonksiyonunun, (6-8)'deki ifadeye benzer şekilde ifade edilmiş integrali biçimindeki,

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma(2\pi)^{1/2}}\;\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = \\ &= \frac{1}{\sigma(2\pi)^{1/2}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx \end{aligned} \qquad\qquad 6\text{-}9c$$

(6-9c) dağılım fonksiyonu da sürekli bir değişkenin dağılım fonksiyonu olmaktadır. (6-8) deki özellikten dolayı (6-9c) integralinin değeri de,

$$\begin{aligned} F(x) &= \int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sigma(2\pi)^{1/2}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx = \\ &= F(+\infty)-F(-\infty)=1-0=1 \end{aligned} \qquad\qquad 6\text{-}9d$$

olmaktadır.

Burada, tekrar vurgulamak gerekirse; sürekli türden rastgele değişkenler yerine diskırıt*(ayrık)* türden rastgele değişkenlerin kullanılmasının söz konusu olduğu problemlerde ise, yukarıda sözü edilen integral alma işlemi yerine toplam ifadesi kullanılarak aynı işlemler n sonlu sayıdan oluşan örnekleme veri kümeleri için de benzer bir yaklaşımla olasılık kuramına göre yapılabilir.

6.1.2. Moment Üreten Fonksiyonları

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

$$M(t) = \int_{-\infty}^{+\infty} e^{tx} f(x)\ dx = \int_{-\infty}^{+\infty} e^{tx} \frac{1}{b(2\pi)^{1/2}} \exp\left\{-\frac{(x-a)^2}{2b^2}\right\} dx \qquad \text{6-10}$$

şeklinde verilebilir. Burada; bir değişken dönüşümü için

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

ifadesi yazılır ve bu ifadeden x elemanı çekildiğinde,

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

olarak elde edilir. Daha sonra bunun (6-10) formülünde yerine yazılmasından,

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

$$= \exp\left\{at + \frac{b^2 t^2}{2}\right\} \int_{-\infty}^{+\infty} \frac{1}{(2\pi)^{1/2}}\ \exp\left(\frac{-y^2}{2}\right)\ dy \qquad \text{6-11}$$

bağıntısı elde edilir. Burada, (6-11)'de, gerekli ara işlemlerin yapılması neticesinde bir normal dağılımın *M(t)* moment üreten fonksiyonları için nihai sonuç,

$$M(t) = \exp\left\{at + \frac{b^2 t^2}{2}\right\} \qquad \text{6-12}$$

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 ya da parametresi,

$$\mu = M^{'}(0) = a \qquad 6\text{-}13$$

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

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

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

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

ve buradan da;

$$\sigma^2 = M^{''}(0) - \left\{M^{'}(0)\right\}^2 = (b^2 + a^2) - a^2 = b^2 \qquad 6\text{-}14$$

olarak elde edilir.

Sonuçta, bu açıklamaların ışığı altında (6-15a) da verilmiş olan normal dağılımın $f(x)$ olasılık fonksiyonu μ ve σ^2 parametrelerine göre

$$f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \quad ; \quad -\infty \langle \; x \; \langle +\infty \qquad 6\text{-}15a$$

şeklinde yazılabilir. Benzer şekilde, bu dağılım için μ ve σ^2 parametreleri de kullanılarak, normal dağılımın *F(x)* dağılım fonksiyonu da,

$$F(x) = \int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx = F(+\infty) - F(-\infty) = 1 - 0 = 1 \qquad \text{6-15b}$$

şeklinde ifade edilmiş olur.

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

$$n(\mu, \sigma^2) \qquad \text{ya da} \qquad x \to N(\mu, \sigma^2)$$

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

Sonuçta, yapılan bu işlemlerden sonra normal dağılım için moment üreten fonksiyonun,

$$M(t) = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \qquad \text{6-16}$$

şeklindeki bir bağıntı olduğu özet olarak verilebilir.

6.1.3. Normal Dağılımın Eğrisi ve Özellikleri

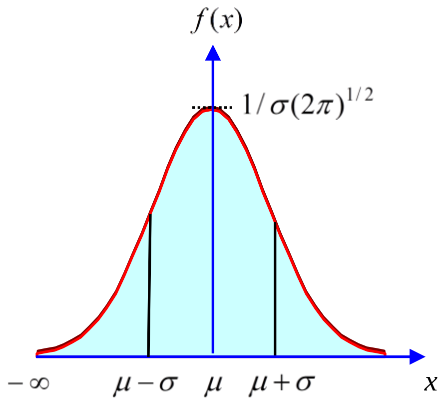
Bir normal dağılım ile ilgili,

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

yoğunluk fonksiyonu ya da daha pratik amaçlara yönelik gösterimler için kullanılan

$$n(\mu, \sigma^2) \quad \text{ya da} \quad x \to N(\mu, \sigma^2)$$

normal dağılım gösterimlerinde iki farklı parametre değeri bulunmaktadır. Bunlardan biri; μ dağılım eğrisinin umut değeri, diğeri de onun yayılmasının bir ölçütü olan σ^2 varyansı ya da onun karekökü olan σ standart sapma değeridir. Bu iki parametrenin alacağı farklı değerlere göre tek boyutlu bir normal dağılım fonksiyonunun bir düzlem kartezyen dik koordinat sisteminde grafiği çizildiğinde Şekil 13‘deki gibi bir eğrinin elde edilmiş olacağı görülür.



Şekil 13: Normal dağılım eğrisi

Şekil 13 ’den görüleceği gibi; bir normal dağılım eğrisi her zaman aşağıda sözü edilen özelliklere sahip bir eğri fonksiyon olmaktadır.

- μ *Parametresine göre simetriktir. Bu nedenle bütün tek dereceden 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 değerli olur. Bunun standart değeri 0.399 gibi bir miktar olmaktadır. Bu aynı zamanda normal dağılımın en olasılıklı ya da tepe noktası değeri olmaktadır.*

- *Yoğunluk fonksiyonu* $\pm\infty$ *değerleri için* x *eksenine asimtotiktir.*
- *x Ekseni ile f(x) yoğunluk fonksiyonunun sınırladığı alan* $S = 1-\alpha$ *güven seviyesini veya* $\alpha = 1-S$ *yanılma olasılığını göstermektedir. Çünkü bir dağılım için daima* $\alpha + S = 1$ *olmaktadır.*

- *f(x) Yoğunluk fonksiyonunun* $x=\mu \pm \sigma$ *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 medyanına eşittir. Yani normal dağılım eğrisinin sınırladığı alanı,*

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

şeklinde iki eşit parçaya bölen b sınır değeri aynı zamanda onu maksimum yapan en olasılıklı değeri, diğer bir ifade ile hem mod hem de medyan değerleri olmaktadır. Yani ortalamaya eşit bir değer olmaktadır.

- $x \rightarrow N(\mu,\sigma^2)$ *Rastgele değişken için* $P(x \le \mu_x) = 0.50$ *olmaktadır.*

- *Bu eğrinin sınırladığı toplam alan daima,*

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

olmaktadır. Pratikte bu durum bir olayın mutlak anlamda yani; %100 gerçekleşeceğini göstermektedir

Burada tekrar vurgulamak gerekirse, bir normal dağılım eğrisi için yukarıda belirtilmiş olan özellikler aynı zamanda bir örnekleme veri kümesinin normal dağılımda olup olmadığının irdelenmesinde kullanılmakta olan *“Normal dağılıma uyum iyiliği hipotez testlerinin”* geliştirilmesindeki temel dayanak konularını oluşturmaktadır.

6.1.4. Normal Dağılıma Sahip Bir Rastgele Değişkenin Standartlaştırılması

Eğer bir x rastgele değişkeni $x \to N(\mu, \sigma^2)$ parametrelerine göre normal dağılıma sahipse, o zaman bunun standartlaştırılmış *(Normlandırılmış ya da* $\mu = 0$ *ortalama ve* $\sigma^2 = 1$ *varyansına göre merkezi normal dağılıma göre ifade edilmiş şekli)* durumdaki z yeni rastgele değişkeni;

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

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

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

aritmetik ortalama ya da umut değeri ve

$$\begin{aligned} E\{z^2\} &= E\left\{(\frac{x-\mu}{\sigma})^2\right\} = \frac{1}{\sigma^2}(E\{(x-\mu)^2\}) = \frac{1}{\sigma^2}(E\{(x^2 - 2x\mu + \mu^2)\}) = \\ &= \frac{1}{\sigma^2}(E\{x^2\} - E\{2\mu x\} + E\{\mu^2\}) = \frac{1}{\sigma^2}(E\{x^2\} - 2\mu E\{x\} + \mu^2) = \\ &= \frac{1}{\sigma^2}E\{x^2\} - \mu^2 = \frac{\sigma^2}{\sigma^2} = 1 \end{aligned}$$

varyans değerlerine göre; $z \to N(0.1)$ şeklinde ifade edilebilen,

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

olan bir merkezi normal dağılıma sahip olur. Matematik-istatistikte bir rastgele değişkenin bu şekilde eşdeğer özelliklere sahip bir diğer merkezi normal dağılıma sahip rastgele değişken biçimine dönüştürülmesi işlemine normlandırma ya da rastgele değişkenlerin standartlaştırılması işlemi denir. Bunun daha önceki konularda umut değerlerine göre yapılmış kısa açıklamasına göre de, sürekli türden bir rastgele değişken için matematik istatistik yasalarına göre benzer bir diğer açıklaması aşağıdaki gibi yapılabilir.

Böyle bir açıklamaya göre yukarıda sözü edildiği şekilde standartlaştırılma işlemi yapılmış bir z rastgele değişkeninin bilinen *Kümülatif dağılım* fonksiyonu;

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

ya da integral 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 fonksiyonunun değişkene göre;

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

şeklindeki bir türevi olduğundan;

$$f(z) = \frac{1}{(2\pi)^{1/2}} \exp(\frac{-z^2}{2}) \qquad 6\text{-}17$$

olarak verilebilir. Bu şekilde (6-17) de elde edilmiş olan bir yoğunluk fonksiyonu ifadesinin; (6-15a) bağıntısında

$$\mu = 0 \qquad \text{ve} \qquad \sigma^2 = 1$$

parametre değerleri kullanılarak elde edilecek ifadenin aynısı olduğu açıkça görülür. Bu durumda;

$$z \rightarrow N(0,1)$$

standart normal dağılımın tanımladığı eğri

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

normal dağılım fonksiyonunda dağılımın parametreler için $\mu = 0$ ve $\sigma^2 = 1$ olduğu dikkate alınmak suretiyle elde edilen merkezi bir olasılık fonksiyonun tanımladığı eğrinin benzeri olur.

Neticede, bunların incelenmesinden her iki değişkene göre ifade edilen normal dağılım bağıntıları birbirleriyle bağlantılı ve eşdeğer ifadeler oldukları görülür. Aralarındaki fark sadece gerek umut değeri gerekse varyans değeri için sabit bir öteleme değerinden başka bir şey olmamaktadır.

Sonuçta böyle bir merkezi normal dağılıma sahip rastgele değişken için de daha önce sözü edilmiş olduğu gibi normal dağılıma sahip herhangi bir x rastgele değişkeni için söylenenlere benzer özelliklerden standart normal dağılım eğrisi, genel durumdaki kartezyen koordinat sisteminin x ekseni yerine z eksen gösterimi kullanılarak aşağıdaki sonuçlara varılabilir.

- $z = 0$ *Parametresinin belirlediği düşey eksene göre simetriktir,*
- $z = 0$ *İçin* $1/(2\pi)^{1/2} = 0.399$ *değerinde Maksimum olur,*
- $\pm\infty$ *Değerleri için* z *eksenine asimtotiktir,*
- $z = \pm 1$ *Değerleri onun dönüm noktasıdır,*
- *Modu meydanına eşittir,*
- $\int_{-\infty}^{+\infty} f(z)dz = 1$ *dir.*

Bu gibi özelliklere sahip bir merkezi normal dağılım eğrisi, Şekil 13 deki normal dağılım eğrisi grafiğinde $\mu = 0$ ve olduğu kabul edilerek elde edilecek bir görünüm sergilediği söylenebilir.

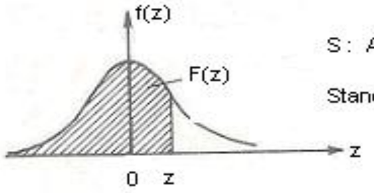
6.1.5. Normal Dağılımla İlgili Hesaplamalar

Bilindiği gibi, tek boyutlu bir normal dağılımın fonksiyonu iki parametre ile tanımlanır. Bunlardan birisi; dağılımın μ ortalama ya da umut değeri, bir diğeri de; dağılımın dispersiyonunu temsil eden σ^2 varyans ya da karekökü olan σ değeridir. Bu parametrelerin belirlediği dağılıma sahip bir rastgele değişkenin daha önceden tanımlanmış herhangi bir sınır *(kritik)* değerinden küçük veya büyük olma, ya da bunlardan üretilmiş belli sınır değerleri arasına düşme olasılığı, parametrelerin tanımladığı olasılık fonksiyonunun bu sınır değerleri arasındaki integral değerine eşdeğer olur. Aynı zamanda bu integral değeri, onun dağılım ya da olasılık miktarı yoğunluk fonksiyonunun bu sınır değerleri arasındaki integral veya tanımladığı alan değeri kadar bir miktar olur. Bu alan, aynı zamanda, rastgele değişken için anlamlılık *(Signifikant)* seviyesini gösteren *S-güven* alanını temsil eder. Arda kalan alan $\alpha = 1 - S$ ise; dağılımın her iki ya da uç kısımlarındaki alan miktarı kadar olur ve yanılma olasılığını gösteren güvensiz alan veya çift taraflı hipotezler için tanımlanan yanılma alanları olur.

İstatistik anlamda olayları açıklamak amacıyla böyle bir değerin elde edilmesinde; her seferinde olasılık fonksiyonunun sınır değerleri arasındaki integralinin alınması gerekmektedir. Olasılık fonksiyonlarının üstel yapıda bir fonksiyon olması nedeniyle, integrallerinin her seferinde yeniden alınması oldukça karmaşık ve zaman alıcı zahmetli matematiksel işlemler gerektirmektedir. Uygulamada bu gibi güçlüklerden kurtulmak için; günümüz teknolojisine paralel olarak hızlı bilgisayarların kullanım alanına girmiş olmasına rağmen, alışılagelen bir yol olarak, genellikle, bu gibi karmaşık işlemlerden kurtulmak amacıyla değişkenlerin $S = 1 - \alpha$ güven aralıklarının hesaplanmasındaki işlemler yönünden daha basit bir yol olan ve parametreler için $\mu = 0$ ve $\sigma^2 = 1$ değerleri alınarak merkezi normal dağılım fonksiyonuna göre hesaplanmış, neticede daha önceden hazırlanmış vaziyetteki standart dağılım tablolarının kullanılması suretiyle istenen sonuçlara varılır (*Tablo 2* veya *Ek.1* standart normal dağılım tablosu).

Pratik hayatta, hazır vaziyette ve kullanımdaki mevcut durumlarına göre her iki şekilde düzenlenmiş olan bu standart normal dağılım tablolarının incelenmesinden, bir kısmının $S = 1 - \alpha$ güven ya da anlamlılık seviyesine göre hazırlanmış olmalarına rağmen (Tablo 2), bir diğer kısmının $\alpha = 1 - S$ yanılma olasılığına göre düzenlenmiş olduklar açıkça görülmektedir (*Ek 1*).

Tablo 2: Standart Normal Dağılım Tablosu

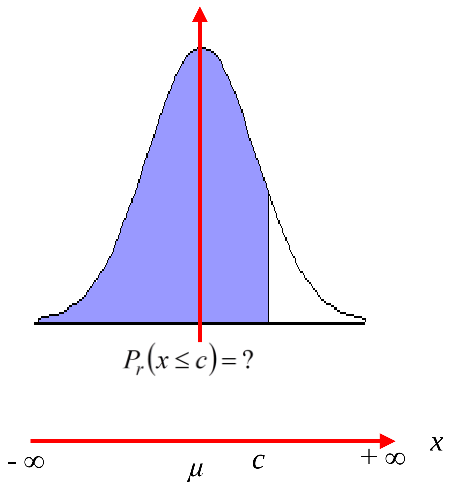


S : Anlamlılık seviyesine göre düzenlenmiş

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

Ancak, bunların pratik kullanımları yönünden aralarında önemli bir fark bulunmamakla birlikte, her iki düzende tasarlanmış ve düzenlenmiş mevcut standart normal dağılım tabloları birbirinin oldukça fazla benzeri tablolar oldukları rahatlıkla söylenebilir. Çok amaçlı kullanıma yönelik istatistikçiler tarafından önceden hazır durumda sunulmuş olan bu tür standart normal dağılım tablolarının kolay kullanılabilir olmaları için her biriyle ilgili kullanım parametreleri; P_r olasılık değeri ile ilgili $S = 1 - \alpha$ *güven* ya da $\alpha = 1 - S$ *yanılma* olasılık değerlerine karşılık gelen z standart rastgele değişkeninin apsis sınır değerleri seçilmiştir. Bu tablolardan doğrudan veya ara değerler için doğrusal enterpolasyon yoluyla hesaplanabilecek z sınır değerleri; önceden öngörülmüş $S = 1 - \alpha$ güven ya da $\alpha = 1 - S$ yanılma olasılık değerlerine karşılık gelen, *F(x)* dağılım fonksiyonunun bağımsız değişkenine ilişkin herhangi bir özel ya da sınır değerlerini göstermektedir (Şekil 14).



Şekil 14: Normal dağılım eğrisi

İstatistik problemlerin çözümünde böyle bir değerin kullanılmış olması ilgili problemlerin irdelenmesindeki *kritik* ya da *kuramsal sınır* değerini oluşturmaktadır.

Ayrıca, normal dağılımın simetrik yapılı bir dağılım olması nedeniyle, bu gibi tablolar $P_r(0 \le z \le +\infty)$ değerleri için düzenlenerek, $z = \mu = 0$ ortalama değerine karşılık gelen $P_r(z = 0) = 0.5000$ olasılık değerinden başlamaktadır. z *standart rastgele* değişkeninin alacağı negatif değerlere karşılık gelen P_r olasılık değerleri z rastgele değişken için tablolarda verilmiş mevcut pozitif normal dağılım değerlerinden, normal dağılımın simetriklik özelliğine göre doğrudan hesaplanır. Tersi yöndeki hesaplamalar için de durum benzer 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, pratikte böyle bir tablonun düzenlenmiş olması, konunun ele alınış biçiminden başka hesaplamalar açısından hiçbir farklı sorun teşkil etmemektedir. Bu tablolarda da, aynen, $S = 1 - \alpha$ güven alanına göre düzenlemiş standart normal dağılım tablolarında olduğu gibi direkt problemlerde; z tablosuna apsis sınır değeri girilerek ona karşılık gelen olasılık değeri tabloda varsa doğrudan, yoksa basit bir doğrusal enterpolasyonla hesaplanarak alınır. Buna karşılık, invers çözümlerde bunun tersi yol izlenerek verilen bir olasılık değerine karşılık gelen z apsis sınır değeri doğrudan ya da enterpolasyonla bulunur. Uygulamada böyle bir invers işlem için, önce verilen olasılık değeri tabloda bulunur ve ona karşılık gelen z apsis sınır değeri de ilgili ilk sütun ya da satırdan elde edilir. Normal dağılımla ilişkili problemlerin çözümünde tablo değerlerinin kullanılmasındaki temel esas; teorik değer P_r olasılığı ve

z sınır değerlerine göre tablo haline getirilmiş normal dağılım değeri arasındaki ilişkinin öncelikle elde edilmesidir. Eğer bir rastgele değişken $x \to N(\mu, \sigma^2)$ dağılımda ise, o zaman $P_r(x \le c) = ?$ şeklinde matematik ifadesi ile verilmiş olan; x'in c gibi bir sınır değerine eşit veya küçük olma olasılığı bu tablolar yardımı ile kolayca hesaplanabilir (Şekil 14). Bu amaçla yapılacak bir $P_r(x \le c) = ?$ olasılığını belirle işleminde, önce $x \to N(\mu, \sigma^2)$ parametrelerine göre normal dağılıma sahip olan x rastgele değişkeninin ve c sınır değerinin, $z \to N(0,1)$ parametrelerine göre standart normal dağılım tablosundaki karşılıkları olan z değişken ve z_s sınır değerleri hesaplanması gerekmektedir. Bunun için öncelikle rastgele değişken ve sınır değerlerine ilişkin,

$$z = \frac{x-\mu}{\sigma} \qquad : \qquad \textit{Rastgele değişken dönüşümü,}$$

ve

$$z_s = \frac{c-\mu}{\sigma} \qquad : \qquad \textit{Sınır değeri dönüşümü,}$$

biçiminde değişken dönüşümü işlemleri yapılır. Bu şekilde gerçekleştirilen bir dönüşüm işlemi sonucunda elde edilmiş olan z standart normal dağılımlı rastgele değişkeni ve z_s onun P_r olasılıktaki üst sınır değeri olmak üzere;

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

olduğu dikkate alınarak,

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

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

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

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

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

Burada, konuyla ilgili bazı kavram ve işlemleri özetle tekrar ele almak gerekirse, $x \rightarrow N(\mu,\sigma^2)$ gibi bir normal dağılıma sahip x rastgele değişkeninin $z \rightarrow N(0.1)$ gibi bir standart normal dağılma sahip z standartlaştırılmış rastgele değişkene dönüştürülerek olasılığın standart normal dağılım tablolarından faydalanılarak hesaplanabilmesi için öncelikle bu değişkenler ve sınır değerleri arasında

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

şeklinde bir dönüşüm uygulanır. Daha sonra, problemin yapısına göre, ya standart normal dağılım sınır değerleri kullanılarak olasılık değeri, ya da olasılık değeri kullanılarak standart rastgele değişken sınır değeri ve ters dönüşümle de $x \rightarrow N(\mu,\sigma^2)$ parametreli rastgele değişkenin sınır değeri bulunur. Sonuçta, standart normal dağılım tablosu kullanılarak olasılık fonksiyonunun $(-\infty\,, c)$ aralığında integrali alınmaktadır. Böyle bir çözüm neticesinde de bir rastgele değişkenin önceden öngörülmüş bir aralıktaki olasılık değerini gösteren güven alanı veya anlamlılık seviyesi hesaplanmış olmaktadır.

Diğer bir yönüyle burada vurgulamak gerekirse, bir rastgele değişkenin belirli sınır değerleri arasına düşme olasılığı olan tersi yöndeki bir problem çözümü için de bu düşünce aynen geçerliliğini korumaktadır.

Örnek 1: $x \rightarrow N(\,2,\;16\,)$ parametrelerine göre normal dağılıma sahip *(burada normal dağılımın parametreleri için* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \le 4) = ?$ olma olasılığının hesaplanması istenmektedir.

Çözüm 1: Böyle bir soruda $P_r(x \le 4) = ?$ olma olasılığını hesaplayabilmek için, $x \rightarrow N(\,2,\;16\,)$ rastgele değişkeni $\mu = 2$ *ve* $\sigma^2 = 16$ parametre değerlerine göre normal dağılıma sahip olduğundan, öncelikle x değişkenin hesaplamalarda kullanılacak standart normal dağılım tablosuna ilişkin parametre değerlerine dönüştürülmesi gerekir. Bunun için yukarıda sözü edilen işlem yolu izlenerek normal dağılıma sahip bir $x \rightarrow N(\mu,\sigma^2)$ rastgele değişkenin $z \rightarrow N(0,1)$ standart normal dağılımlı bir diğer rastgele değişken haline dönüştürülmüş şekli için,

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

dönüşüm bağıntısından faydalanılarak aşağıdaki hesaplamalar yapılır. Bu amaçla $z \rightarrow N(0,1)$ standart rastgele değişkeninin güven aralığının üst sınır değeri olan z_s sınır değeri de, sınır değeri için $x = c = 4$ alınarak

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

olarak elde edilir. Sonuçta, ilk şekliyle $P_r(x \le 4) = ?$ biçiminde verilmiş olan problem,

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

rastgele değişken dönüşümü neticesinde $P_r(z \le 0{,}5) = ?$ şeklinde ifade edilmiş olur. Matematiğin bir diğer gösterim şekli kullanılarak aynı ifade yoğunluk fonksiyonlarına göre entegral biçiminde

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

ya da dağılım fonksiyonları kullanılarak da

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

şekline ifade edilebilir. Buradan, ilgili standart normal dağılım tablosundan $z_s = 0.5$ sınır değerine karşılık gelen olasılık değeri $F(0{,}5) = 0.6915$ alınarak, $-\infty$ alt sınır değeri için olasılık $F(-\infty) = 0$ olduğu dikkate alınmak üzere;

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

sonucu bulunmuş olur (Tablo 2).

Sonuç: $P_r(x \le 4) = 0.6915$ veya (%69) olarak bulunmuş olur.

Örnek 2: Bu örnek problemde $x \to N(2,\ 16)$ normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \ge 4) = ?$ olma olasılığının hesaplanması istenmektedir.

Çözüm 2: Böyle bir problemde,

$$P_r(x \ge 4) = ?$$

olma olasılığını hesaplayabilmek için öncelikle; hem $x \rightarrow N(2,\ 16)$ rastgele değişkeni hem de sınır değeri, örnek 1'de olduğu gibi $z \rightarrow N(0\ ,\ 1)$ standart normal dağılıma ilişkin rastgele değişken değerlerine dönüştürülür. Bu amaçla; z rastgele değişkeni ve z_s sınır değerleri;

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

olarak hesaplanır. Bu değerlere göre mevcut $P_r(x \geq 4) = ?$ ilk şekline göre verilmiş olan 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, bir rastgele değişkenin olasılığının hesaplanmasına temel olan yoğunluk fonksiyonunun integralinin alınması işlemine uygun olarak $-\infty$ alt sınır değeri ile z_s üst sınır değeri gibi iki sınır değeri arasındaki integral değeri bulma (alanı hesaplama) konusuna göre düzenlenmiş olduklarından, z_s sınır değeri ile $+\infty$ arasındaki olasılık değeri, $-\infty$ ile z_s sınır değerlerden faydalanılarak;

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

bağıntısından hesaplanabilir. Buna göre aranan sonuç, gerekli ara işlemlerin yapılması neticesinde,

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

$$P_r(x \geq 4) = P_r(z \geq 0{,}5) = 1 - F(0{,}5) = 1 - 0{,}6915 = 0{,}3085$$

olarak hesaplanır.

Sonuç: Burada gerçekleştirilmiş çözüm için $P_r(x \geq 4) = 0{,}3085$ veya (%31) olasılığı hesaplanmış olur.

Örnek 3: Konuyla ilgili bir diğer problem; $x \rightarrow N(2,\ 16)$ parametrelerine göre normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* bir x rastgele değişkeninin,

$$P_r(x \leq -4) = ?$$

olma olasılığının hesaplanması istenmektedir.

Çözüm 3: Böyle bir problemin çözümü için öncelikle normal dağılımın bazı geometrik özelliklerinden faydalanılır. Hatırlanacağı gibi, normal dağılımla ilgili böyle bir özellik "Normal dağılım fonksiyonları ortalama değere göre simetriktir. Yani normal dağılımda mod değeri medyan değerine eşittir" şeklinde verilmişti. Bu özelliğin neticesinde, bütün standart normal dağılım tabloları normal dağılım fonksiyonunun pozitif apsis değerli yarı alan değerleri için düzenlenmektedir. Bu nedenle; negatif yarı alanı içinde sorulmuş olasılık değerini hesaplayabilmek için, problem öncelikle tablo değerleri mevcut olan pozitif apsis değerli olasılık bölgesine dönüştürülür. Bu amaçla, negatif apsis değerli bir sınır için sorulmuş olasılık değeri,

$$z = \frac{x-\mu}{\sigma} = \frac{x\text{-}2}{4} \qquad ; \qquad z_s = \frac{c-\mu}{\sigma} = \frac{-4-2}{4} = -1{,}5$$

$$P_r(x \leq -4) = P_r(z \leq -1.5) = 1 - P_r(z \leq 1.5)$$

formülü kullanılarak normal dağılım eğrisinin pozitif bölgesiyle ilgili apsis değerlerine çevrilir. Bu haliyle problem; Örnek 1 'de yapılan çözüme benzer şekilde ve Örnek 2 'deki düşüncenin de göz önüne alınması suretiyle,

$$P_r(x \leq -4) = P_r(z \leq -1.5) = 1 - P_r(z \leq 1.5) = 1 - 0.9332 = 0.0668$$

olduğu hesaplanır.

Sonuç: $P_r(x \leq -4) = 0{,}0668$ veya $(\approx \%07)$ olarak bulunmuş olur.

Örnek 4: Konuyla ilgili benzer bir diğer örnek problem de, $x \rightarrow N(2,\ 16)$ parametrelerine göre normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(x \geq -4) = ?$ olma olasılığının hesaplanmasıdır.

Çözüm 4: Böyle bir problemin çözümü; yukarıdaki örnek problem çözümlerinde izlenen yollara benzer şekildeki düşüncelerden hareket edilerek, istene çözüm gerçekleştirilebilir. Bu amaçla, bir normal dağılım veya standart normal dağılım fonksiyonlarının tanımladığı grafiklerde, her birinin ortalama değere göre simetrik bir dağılım olması, bunların ayrıca mod *(tepe)* değerinin medyan *(ortanca)* değerine eşit bir değer olması gibi özellikler büyük önem taşımaktadır. Buna göre, örnek 3 ‘deki gibi bir değişken dönüşümü sonucundan

$$P_r(x \geq -4) = P_r(z \geq -1{,}5) = 1 - P_r(z \leq -1{,}5)$$

Neticede bu özelliğe göre; standart normal dağılım eğrisinin her iki kuyruk tarafındaki $P_r(x \geq -1{,}5) = P_r(x \leq 1{,}5)$ alanları daima birbirine eşit oldukları göz önüne alınarak problemin çözümü için,

$$P_r(x \geq -4) = P_r(z \geq -1{,}5) = 1 - P_r(z \leq -1{,}5) = P_r(z \leq 1{,}5) = 0{,}9332$$

değeri basit bir düşünce ile elde edilmiş olur.

Sonuç: Buradan $P_r(x \geq -4) = 0{,}9332$ veya $(\approx \%93)$ olur.

Örnek 5: $x \rightarrow N(2,\ 16)$ parametrelerine göre normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* bir x rastgele değişkeninin $P_r(0 \leq x \leq 4) = ?$ aralığına düşme olasılığının hesaplanması istenmektedir.

Çözüm 5: Böyle bir problemin $P_r(0 \leq x \leq 4) = ?$ çözümü için, $x \rightarrow N(2,\ 16)$ rastgele değişkeni $\mu = 2$ *ve* $\sigma^2 = 16$ parametre değerlerine göre normal dağılıma sahip olduğu bilindiğine göre, önce bu değişken standart normal dağılımlı bir $z \rightarrow N(0,\ 1)$ standart dağılıma sahip rastgele değişkene dönüştürülerek alt ve üst sınır değerleri,

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

değişken dönüşümü bağıntısında x yerine alt ve üst sınır değerleri kullanılarak her birine karşılık gelen standart normal dağılıma ilişkin sınır değerleri için,

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

değerleri hesaplanır. Daha sonra, bu sınır değerleri kullanılarak, $x \rightarrow N(2,\ 16)$ parametrelerine sahip normal dağılımlı bir rastgele değişkenin $P_r(0 \leq x \leq 4)$ aralığına düşme olasılığı problemi için,

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

integral bağıntısı yazılabilir. Bu entegral biçimde ifade edilmiş olan olasılık bağıntısında,

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

olduğunun da dikkate alınması neticesinde $P_r(0 \le x \le 4)$ olasılığı için

$$P_r(0 \le x \le 4) = F(0{,}5) - F(-0{,}5) = F(0{,}5) - \{1 - F(0{,}5)\} =$$
$$= 2F(0{,}5) - 1 = 2(0{,}6915) - 1 = 0{,}3830$$

değeri hesaplanmış olur.

Sonuç: Problemle ilgili çözüm için $P_r(0 \le x \le 4) = 0{,}3830$ ya da (%38) bulunmuş olur.

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

Çözüm 6: Böyle bir problemin çözümü için örnek 5'dekine benzer bir yol izlenerek,

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

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. Buna göre de $P_r(1 \le x \le 4)$ şeklinde verilmiş olan 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 olur.

Sonra ilgili standart normal dağılım tablosundan *-0,25* ve *0,5* sınır değerlerine karşılık gelen olasılık değerleri alınarak problemin çözümü

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

olarak hesaplanır.

Sonuç: $P_r(1 \le x \le 4) = 0,2902$ olarak hesaplanmış olur.

Örnek 7: $x \to N(2,\ 16)$ parametrelerine göre 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 gerektiğinin hesaplanması istenmektedir.

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 tablosundaki z sütununda 1,645 olarak alınır (Tablo 2).

Bu değerin *S=0,95* anlamlılık seviyesine veya olasılıkla standart rastgele değişken için $z_s = 1,645$ sınır değeri olarak alınması neticesinde,

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

eşitliği kurularak c sınır değeri için, $c = 1,645\sigma + \mu$ bağıntısı elde edilmiş olur. Bu bağıntıda $\mu = 2$ *ve* $\sigma = 4$ değerlerinin yerlerine yazılması ile $c = 8,58$ olarak hesaplanmış olur.

Sonuç: Bu problem için $c = 8,58$ olarak hesaplanmış olur.

Örnek 8: $x \to N(2,\ 16)$ parametreli 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 değerinin hesaplanması istenmektedir.

Çözüm 8: Bu problemin çözümü için daha önceki örneklerde sözü edilen nedenlerden dolayı $P_r(x \geq c) = 1 - P_r(x \leq c)$ şeklinde yazılabilir. Burada değişken dönüşümünden elde edilecek olan standart normal dağılımlı bir

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

rastgele değişkeninin sınır değeri için,

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

eşitliği yazılabilir. Bu sınır değeri eşitliğine göre;

$$P_r(x \geq c) = 1 - P_r(x \leq c) = 1 - F(z_s) = 0{,}2$$

denkliği yazılarak ve buradan $P_r(z \leq z_s)$ olasılığı için

$$F(z_s) = 1 - 0{,}2 = 0{,}80$$

değeri elde edilir. Bu olasılığa karşılık gelen sınır değeri standart normal dağılım tablosundan $z_s = 0{,}842$ olarak alınır (Tablo 2).

Sonuçta;

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

eşitliği kurularak, bu eşitliğin çözümünden, c sınır değeri için,

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

formülü elde edilir. Daha sonra bu formülde $\mu = 2$ *ve* $\sigma = 4$ oldukları dikkate alınarak c sınır değeri,

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

olarak hesaplanır.

Sonuç: Böylece aranan sınır değeri $c = 5{,}368$ olarak hesaplanmış olur.

Örnek 9: $x \to N(2,\ 16)$ parametrelerine göre normal dağılıma sahip *(burada normal dağılımın parametreleri* $\mu = 2$ *ve* $\sigma^2 = 16$ *dır)* x rastgele değişkeninin $P_r(-c \langle x \le 10) = 0{,}5$ olması için c=? sınır değerinin hesaplanması istenmektedir.

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

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

rastgele değişkeninin,

$$z_a = \frac{x-\mu}{\sigma} = \frac{-c-2}{4} \qquad ; \quad \textit{Alt sınır,}$$

ve

$$z_ü = \frac{x-\mu}{\sigma} = \frac{10-2}{4} = 2 \qquad ; \quad \textit{Üst sınır değerleri}$$

değerleri hesaplanır. Bu sınır değerlerine göre ilgili olasılık için,

$$\begin{aligned} P_r(-c\langle x \le 10) &= P_r(z_a \langle z \le z_ü) = F(z_ü) - F(z_a) = \\ &= F(2) - F(-\frac{c+2}{4}) = \\ &= F(2) - \left\{1 - F(\frac{c+2}{4})\right\} = 0{,}5 \end{aligned}$$

ifadesi yazılır. Buradan, gerekli ara işlemlerin yapılması neticesinde,

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

elde edilmiş olur.

Daha sonra, Tablo 2'de verilmiş olan standart normal dağılım tablosundan alınan $F(2) = 0.9773$ değerinin yerine yazılması ile

$$F(\frac{c+2}{4}) = 0{,}5227$$

bulunur. Normal dağılım tablosundan 0,5227 olasılık değerine karşılık gelen değişken değeri $(c+2)/4 = 0{,}5227$ alınarak, buradan c sınır değeri;

$$c = 0.057\sigma + \mu = (0{,}057)4 + 2 = 2{,}228$$

olarak hesaplanır.

Sonuç: Böylece problemin çözümünden sınır değeri $c = 2{,}228$ olarak bulunmuş olur.

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

Çözüm 10: Yukarıdaki verilmiş örnek problemlerin çözümüne benzer işlem yolları bu problemin çözümü içinde uygulanarak,

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

rastgele değişkeninin her iki alt ve üst 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;

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

bulunur ve buradan da

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

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

$$c = \frac{1.645}{2} = 0.8225$$

olarak hesaplanır.

Sonuç: $c = 0{,}8225$ 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 istenmektedir.

Çözüm 11: Bu problemde de benzer yollar izlenerek problemin aranan çözümü 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 ile alt ve üst sınır değerleri için

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

şeklinde hesaplanarak, bu ifade için

$$P_r(z_a \le z \le z_ü) = 0{,}95$$

bağıntısı yazılır. Sonuçta bu bağıntıdaki her bir terimin yerine karşılıkları yazılarak el edilen

$$F(z_ü) - F(z_a) = F(\frac{c-\mu}{\sigma}) - F(-\frac{c-\mu}{\sigma}) = 0{,}95$$

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

$$F(\frac{c-\mu}{\sigma}) = \frac{1-0{,}95}{2} = 0{,}975$$

son bağıntıda standart normal dağılım tablosundan 0,975 olasılık değerine karşılık gelen sınır değeri 1,96 alınarak her iki değer arasında,

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

biçiminde yazılan eşitlikte gerekli işlemlerin yapılması neticesinde c sınır değeri,

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

olarak elde edilir.

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

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

a) kaç tanesinin boy uzunlukları 120cm. ile 155 cm arasındadır?

b) *kaç tanesi 186 cm. den daha uzundur?*

hesaplanmak istenmektedir.

Çözüm 12:

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

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

olasılığı için, bu verilerin normal dağılımda olması özelliğinden faydalanılarak bir $z = \dfrac{x-\mu}{\sigma}$ rastgele değişken dönüşümü yapılır. Bu dönüşüm neticesinde, z standart rastgele değişkenin z_a alt ve $z_ü$ üst sınır değerleri,

$$z_a = \frac{119{,}5-\mu}{\sigma} = \frac{119{,}5-151}{15} = -2{,}10$$

$$z_ü = \frac{155{,}5-\mu}{\sigma} = \frac{155{,}5-151}{15} = 0{,}30$$

olarak hesaplanır.

Bu değerlere göre; $P_r(120 \leq \mu \leq 155) = ?$ bulunma 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,

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

olarak bulunur. Diğer bir ifade ile bu okuldaki 500 kız öğrenciden boyları 120*cm*. 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 hesaplanmış olur.

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,

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

hesaplanan %0,107 yüzde değerine karşılık gelenlerinin sayısı *500(0,0107)=5* olarak elde edilir. Sonuçta; bu öğrencilerden ancak *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 0,496 *cm.* ile 0,508 *cm.* arasında olanları sağlam, diğerleri bozuk kabul edilmektedir. Bu fabrikada bozuk olarak üretilen çamaşır makinelerinin yüzde kaçı bozuktur. Sayılarının hesaplanması istenmektedir.

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 dağılımın alt ve üst sınırları 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 değerleri hesaplanarak bu değerler, göre,

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

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

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