HAFTA 8

2.5. Beklenen Değer

Bu kısımda, kitlenin önemli özelliklerinden rasgele değişkenlerin momentleri üzerinde durulacaktır. X bir rasgele değişken g de tanım kümesi reel sayılar olan herhangi bir fonksiyon olmak üzere $g(X)$ de bir rasgele değişkendir. X in olasılık veya olasılık yoğunluk fonksiyonu verildiğinde $g(X)$ in de olasılık veya olasılık yoğunluk fonksiyonu da bulunabilir. Rasgele değişkenlerin dönüşümlerinin dağılımlarının bulunması bir sonraki bölümde incelenecektir. Bu kısımda, X in olasılık veya olasılık yoğunluk fonksiyonu verildiğinde rasgele değişkenin momentleri üzerinde durulacaktır. Ayrıca, momentler ile ilgili bazı kavramlar da kısaca ele alınacaktır.

Tanım 2.5.1 Değer kümesi D_X olan X rasgele değişkenin olasılık veya olasılık yoğunluk fonksiyonu $f(x)$ olsun. $g(X)$ rasgele değişkeninin *beklenen değeri*

$$\begin{cases} \int\limits_{D_X} |g(x)|\, f(x)\, dx < \infty & , \ X \text{ sürekli} \\ \sum\limits_{D_X} |g(x)|\, f(x) < \infty & , \ X \text{ kesikli} \end{cases} \quad \text{ise} \quad E(g(X)) = \begin{cases} \int\limits_{D_X} g(x)\, f(x)\, dx < \infty & , \ X \text{ sürekli} \\ \sum\limits_{D_X} g(x)\, f(x) < \infty & , \ X \text{ kesikli} \end{cases}$$

dir$\otimes$

Rasgele değişkenin beklenen değeri her zaman olmayabilir. Yani, $\int\limits_{D_X} g(x)\, f(x)\, dx$ veya $\sum\limits_{D_X} g(x)\, f(x)$ değerleri sonlu olsa bile, $\int\limits_{D_X} |g(x)|\, f(x)\, dx$ (veya $\sum\limits_{D_X} |g(x)|\, f(x)$) integrali (veya toplamı) sonlu olmayabilir. Böyle durumlarda rasgele değişkenin beklenen değeri yoktur diyeceğiz. Şimdi, bunu bir örnek üzerinde görelim.

Örnek 2.5.1 X rasgele değişkeninin olasılık fonksiyonu,

$$P(X = x) = 1/2^x \ , \ x = 1, 2, 3, \ldots$$

olarak verilmiş olsun (Öztürk, 1993, sayfa 241). $g(x) = (-1)^x 2^x / x$ alındığında

$$\sum_{D_X} g(x)\, f(x) = \sum_{x=1}^{\infty} \left((-1)^x \frac{2^x}{x} \right) \frac{1}{2^x} = \sum_{x=1}^{\infty} \frac{(-1)^x}{x} = -\ln(2)$$

olmasına rağmen,

$$\sum_{D_X} |g(x)|\, f(x) = \sum_{x=1}^{\infty} \left| (-1)^x \frac{2^x}{x} \right| \frac{1}{2^x} = \sum_{x=1}^{\infty} \frac{1}{x} = \infty$$

olduğundan $g(X)$ rasgele değişkeninin beklenen değeri yoktur$\oplus$

X herhangi bir rasgele değişken ve $x_0 \in \mathbb{R}$ olsun. Eğer her $x \in \mathbb{R}$ için $P(X \geq x_0 + x) = P(X \leq x_0 - x)$ ise X rasgele değişkeni x_0 noktasına göre *simetriktir* denir. Herhangi bir X rasgele değişkeni c noktasına göre simetrik ise, $E(X) = c$ dir (Biswal, sayfa 64).

Tanım 2.5.2 X rasgele değişkeninin olasılık veya olasılık yoğunluk fonksiyonu $f(x)$ ve her $k \in \mathbb{N}$ için $E(X^k)$ var olsun. $E(X^k)$ değerine X rasgele değişkeninin $k.$*momenti* denir ve μ_k ile gösterilir. Ayrıca,

a) $k = 1$ için $E(X)$ değerine X in *beklenen değeri* denir ve μ ile gösterilir.

b) $E(X-\mu)^k$ değerine X in μ ye göre $k.$ *merkezi momenti* denir ve m_k ile gösterilir.

c) $m_2 = E(X-\mu)^2$ değerine X in *varyansı* denir ve $Var(X)$ veya σ^2 ile gösterilir.

d) $E(X(X-1)(X-2)...(X-(k+1)))$ değerine X in $k.$ *çarpımsal momenti* denir$\otimes$

$Var(X) = \mu_2 - \mu_1^2$ olup X rasgele değişkeninin varyansının pozitif kareköküne X in *standart sapması* denir. X in varyansı σ^2 ise standart sapması σ $(\sigma > 0)$ dır.

Örnek 2.5.2 X rasgele değişkeninin olasılık fonksiyonu $\lambda > 0$ için,

$P(X = x) = e^{-\lambda}\lambda^x / x!$, $x = 0, 1, 2, 3, ...$

şeklinde verilmiş olsun. Bu rasgele değişkenin beklenen değeri

$$\mu = E(X) = \sum_{x=0}^{\infty} x P(X = x) = \sum_{x=1}^{\infty} x P(X = x) = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1+1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

dir. Yani, $E(X) = \lambda$ dir. İkinci moment ise ($y = x - 2$ denirse),

$$\mu_2 = E(X^2) = \sum_{x=0}^{\infty} x^2 P(X = x) = \sum_{x=0}^{\infty} (x + x(x-1)) \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} = E(X) + \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda + e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2+2}}{(x-2)!} = \lambda + \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = \lambda + \lambda^2 e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!}$$

$$= \lambda + \lambda^2 e^{-\lambda} e^{\lambda} = \lambda + \lambda^2$$

şeklinde bulunmuştur. Rasgele değişkenin bu iki momentinden X in varyansı da

$$Var(X) = \mu_2 - \mu_1^2 = (\lambda + \lambda^2) - \lambda^2 = \lambda$$

olarak bulunmuştur $\oplus$

Tanım 2.5.3 X rasgele değişkeninin olasılık veya olasılık yoğunluk fonksiyonu $f(x)$ ve $E(g(X))$ var olsun.

a) $t \in \mathbb{R}$ ve $g(x) = e^{tx}$ için $E(e^{tX})$ fonksiyonuna, X in *moment çıkaran fonksiyonu* denir ve $M_X(t)$ ile gösterilir.

b) $i = \sqrt{-1}$ olmak üzere, $t \in \mathbb{R}$ ve $g(x) = e^{itx}$ için $E(e^{itX})$ fonksiyonuna, X in *karekteristik fonksiyonu* denir ve $\phi_X(t)$ ile gösterilir.

c) $t \in \mathbb{R}$ ve $g(x) = t^x$ için $E(t^X)$ fonksiyonuna, X in *çarpımsal moment üreten fonksiyonu* denir ve $N_X(t)$ ile gösterilir $\otimes$

Bu fonksiyonlar kullanılarak da rasgele değişkenlerin momentleri hesaplanabilir. Örneğin, X in moment çıkaran fonksiyonu $M_X(t)$ varsa, X in momentleri

$$\mu_k = E(X^k) = \frac{d^k}{dt^k} M_X(t) \Big|_{t=0}$$

şekilde moment çıkaran fonksiyonunun türevinden hesaplanır. Türev ile beklenen değer operatörlerinin yer değiştirebildiği varsayımı altında, moment çıkaran fonksiyonunun $k.$ türevi,

$$\frac{d^k}{dt^k} M_X(t) = \frac{d^k}{dt^k} E\left(e^{tX}\right) = E\left(\frac{d^k}{dt^k} e^{tX}\right) = E\left(X^k e^{tX}\right)$$

olup, türevde $t = 0$ yazıldığında rasgele değişkenin $k.$ momentinin

$$\frac{d^k}{dt^k} M_X(t) \Big|_{t=0} = E\left(X^k e^{tX}\right)\Big|_{t=0} = E(X^k) = \mu_k$$

olduğu görülür.

Örnek 2.5.3 a) X rasgele değişkeninin olasılık fonksiyonu $\lambda > 0$ için

$$P(X = x) = e^{-\lambda} \lambda^x / x!\,,\ x = 0,1,2,3,\ldots$$

olarak verilmiş olsun. Bu rasgele değişkenin ilk iki momenti Örnek (2.5.2) de $E(X)=\lambda$ ve $E(X^2)=\lambda+\lambda^2$ olarak hesaplanmıştı. Bu rasgele değişkenin moment çıkaran fonksiyonu (e^x fonksiyonunun Taylor serisi açılımı kullanılarak),

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^t\lambda)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

olup rasgele değişkenin birinci momenti,

$$\mu_1 = E(X) = \left.\frac{dM_X(t)}{dt}\right|_{t=0} = \left.\frac{d}{dt}\left(e^{\lambda(e^t-1)}\right)\right|_{t=0} = \left.\left(\lambda e^t e^{\lambda(e^t-1)}\right)\right|_{t=0} = \lambda$$

ve ikinci momenti,

$$\mu_2 = E(X^2) = \left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \left.\frac{d^2}{dt^2}\left(e^{\lambda(e^t-1)}\right)\right|_{t=0} = \left.\frac{d}{dt}\left(\lambda\ e^t\ e^{\lambda(e^t-1)}\right)\right|_{t=0}$$

$$= \left.\left[\lambda\ e^t\ e^{\lambda(e^t-1)} + \lambda^2\ e^{2t}\ e^{\lambda(e^t-1)}\right]\right|_{t=0} = \lambda + \lambda^2$$

şeklinde hesaplanabilir. Bu moment çıkaran fonksiyonu kullanılarak, rasgele değişkenin diğer tüm momentleri, $k\in\mathbb{N}$ ler için

$$\mu_{k+1} = \lambda\left(\mu_k + \frac{d\mu_k}{d\lambda}\right)$$

şeklinde ardışık olarak bulunabilir. Şimdi, bu bağıntının doğru olduğunu görelim. Bunun için matematiksel tümevarım yöntemi kullanılabilir. Önce, $k=1$ için eşitliğin sol tarafı $\mu_2 = E(X^2) = \lambda+\lambda^2$ dir. Eşitliğin sağ tarafı ise,

$$\lambda\left(\mu_1 + \frac{d\mu_1}{d\lambda}\right) = \lambda\left(\lambda + \frac{d\lambda}{d\lambda}\right) = \lambda(\lambda+1) = \lambda+\lambda^2$$

olup bağıntı $k=1$ için geçerlidir. Matematiksel tümevarım gereği, eşitliğin k için doğru olduğunu varsayarak $k+1$ için doğru olduğunu gösterelim. Eşitlik k için doğru ise, moment çıkaran fonksiyonunun türevinde $t=0$ yazılarak momentlerin elde edildiğini de göz önüne alarak

$$\mu_{k+1}(t) = \lambda\left(\mu_k(t) + \frac{d\mu_k(t)}{d\lambda}\right)$$

yazalım. Buradan, $k+1$ için eşitliğin sol tarafı

$$\mu_{k+2}(t) = \frac{d^{k+2}}{dt^{k+2}}M_X(t) = \frac{d}{dt}\left(\frac{d^{k+1}}{dt^{k+1}}M_X(t)\right) = \frac{d}{dt}\left(\mu_{k+1}(t)\right) = \frac{d}{dt}\left[\lambda\left(\mu_k(t) + \frac{d\mu_k(t)}{d\lambda}\right)\right]$$

$$=\lambda\left(\frac{d}{dt}\mu_k(t)+\frac{d}{dt}\frac{d\mu_k(t)}{d\lambda}\right)=\lambda\left(\mu_{k+1}(t)+\frac{d}{d\lambda}\frac{d\mu_k(t)}{dt}\right)=\lambda\left(\mu_{k+1}(t)+\frac{d\,\mu_k(t)}{d\lambda}\right)$$

şeklinde yazılır. Burada $t=0$ yazıldığında eşitlik $k+1$ için,

$$\mu_{k+2}=\lambda\left(\mu_{k+1}+\frac{d\mu_{k+1}}{d\lambda}\right)$$

olup iddia $k+1$ için de doğrudur. Yani, iddia her k için doğrudur.

b) Negatif değerler almayan kesikli bir X rasgele değişkeninin beklenen değeri varsa, bu beklenen değerin

$$E(X)=\sum_{n=0}^{\infty}P(X>n)$$

şeklinde hesaplanabileceğini görelim. Sağ taraftaki serinin açık olarak yazılması ile

$$\begin{aligned}\sum_{n=0}^{\infty}P(X>n)&=P(X>0)+P(X>1)+P(X>2)+P(X>3)+\ldots\\&=P(X=1)+P(X=2)+P(X=3)+P(X=4)+\ldots\\&\qquad+P(X=2)+P(X=3)+P(X=4)+P(X=5)+\ldots\\&\qquad\qquad+P(X=3)+P(X=4)+P(X=5)+\ldots\\&\qquad\qquad\qquad\ldots\ldots\ldots\\&=1P(X=1)+2P(X=2)+3P(X=3)+4P(X=4)+\ldots\\&=\sum_{x=1}^{\infty}xP(X=x)=\sum_{x=0}^{\infty}xP(X=x)=E(X)\end{aligned}$$

şeklinde aranan eşitlik elde edilir.

Negatif olmayan değerler alan sürekli bir X rasgele değişkeninin dağılım fonksiyonu $F(x)$ olsun. X rasgele değişkenin beklenen değeri varsa,

$$E(X)=\int_0^{\infty}(1-F(x))\,dx$$

dir (Öztürk, 1993) $\oplus$

Bir rasgele değişkenin varyansı, ortalamadan (beklenen değer) sapmasının karesinin beklenen değeridir. Tanım (2.5.2) de bir rasgele değişkenin varyansı $Var(X)=E(X-\mu)^2$ olarak tanımlandı.Bu varyans genellikle $Var(X)=E(X^2)-(E(X))^2$ hesaplanır. Yani,

$$Var(X)=\mu_2-\mu_1^2=E(X^2)-\left(E(X)\right)^2=E(X-\mu)^2$$

dir. $a,b\in\mathbb{R}$ olmak üzere $g(x)=ax+b$ fonksiyonunu tanımlayalım. Buradan kesikli $g(X)$ rasgele değişkeninin beklenen değeri (sürekli durumda toplam sembolleri yerine integraller gelir),

$$E(g(X)) = E(aX+b) = \sum_{x \in D_X} (ax+b) P(X=x)$$
$$= a \sum_{x \in D_X} x P(X=x) + b \sum_{x \in D_X} P(X=x) = a E(X) + b$$

ve

$$Var(g(X)) = Var(aX+b) = E\left((aX+b) - E(aX+b)\right)^2 = E\left(aX + b - aE(X) - b\right)^2$$
$$= E\left(a(X-\mu)\right)^2 = a^2 E(X-\mu)^2 = a^2 Var(X)$$

dir. Yani bir rasgele değişkenin lineer birleşiminin beklenen değer ve varyansı $a,b \in \mathbb{R}$ olmak üzere, $E(aX+b) = aE(X)+b$ ve $Var(aX+b) = a^2 Var(X)$ dir.

Buraya kadar bir boyutlu rasgele değişkenlerin beklenen değeri incelendi. Şimdi, bileşenleri X ve Y olan iki boyutlu rasgele vektörün ortak olasılık veya olasılık yoğunluk fonksiyonu $f(x,y)$ olsun. Marjinal olasılık fonksiyonlarından $E(X)$, $E(Y)$, $Var(X)$ ve $Var(Y)$ değerleri hesaplanabilir. Ayrıca, X ve Y aynı örnek uzay üzerinde tanımlı rasgele değişkenler, g de $\mathbb{R}^2$ den $\mathbb{R}$ ye tanımlı bir fonksiyon olmak üzere $g(X,Y)$ tek boyutlu bir rasgele değişkendir. Buna göre $E(g(X,Y))$ beklenen değeri varsa, bu beklenen değer

$$E(g(X,Y)) = \begin{cases} \int\limits_{x \in D_X} \int\limits_{y \in D_Y} g(x,y) f(x,y)\, dy\, dx & , \quad X \text{ sürekli} \\ \sum\limits_{x \in D_X} \sum\limits_{y \in D_Y} g(x,y) P(X=x, Y=y) & , \quad X \text{ kesikli} \end{cases}$$

şeklinde hesaplanır.

Tanım 2.5.4 Bileşenleri X ve Y olan iki boyutlu bir rasgele vektörün ortak olasılık veya olasılık yoğunluk fonksiyonu $f(x,y)$ olsun.

a) X ile Y arasındaki *kovaryans* $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$,

b) X ile Y arasındaki *korelasyon* ise,

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X) Var(Y)}},$$

c) Bileşenleri X ve Y olan iki boyutlu bir rasgele vektör $\underset{\sim}{X}$ olmak üzere, $\underset{\sim}{X}$ rasgele vektörün beklenen değer vektörü ile varyans kovaryans matrisi sırası ile,

$$\underset{\sim}{\mu} = \begin{bmatrix} E(X) \\ E(Y) \end{bmatrix} \quad , \quad V = \begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{bmatrix}$$

dir$\otimes$

X ile Y arasındaki kovaryans genellikle $Cov(X,Y)=E(XY)-E(X)E(Y)$ şeklinde hesaplanır.

Örnek 2.5.4 a) Bileşenleri X ve Y olan iki boyutlu bir rasgele vektörün ortak olasılık yoğunluk fonksiyonu

$$f(x,y)=\begin{cases} e^{-y} & , \quad 0<x<y<\infty \\ 0 & , \quad d.y. \end{cases}$$

olsun. X ve Y nin marjinal olasılık yoğunluk fonksiyonları,

$$f_X(x)=\begin{cases} e^{-x} & , \quad x>0 \\ 0 & , \quad d.y. \end{cases} \qquad \text{ve} \qquad f_Y(y)=\begin{cases} y e^{-y} & , \quad y>0 \\ 0 & , \quad d.y. \end{cases}$$

dir (Örnek (2.4.1b)). Marjinal olasılık yoğunluk fonksiyonlarından $E(X)=1$, $Var(X)=1$, $E(Y)=2$, $Var(Y)=2$ elde edilir. Ayrıca,

$$\begin{aligned} E(XY) &= \int_{y=0}^{\infty}\int_{x=0}^{y} x\,y\,e^{-y}\,dx\,dy = \int_{y=0}^{\infty} y\,e^{-y}\left(\int_{x=0}^{y} dx\right)dy \\ &= \int_{y=0}^{\infty} y\,e^{-y}\left(\frac{y^2}{2}\right)dy = \frac{1}{2}\int_{y=0}^{\infty} y^3 e^{-y}dy = \frac{\Gamma(4)}{2}=\frac{3!}{2}=3 \end{aligned}$$

olup X ile Y arasındaki korelasyon

$$\rho_{X,Y}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{E(XY)-E(X)E(Y)}{\sqrt{Var(X)Var(Y)}}=\frac{3-(1)(2)}{\sqrt{(1)(2)}}=\frac{1}{\sqrt{2}}$$

olarak hesaplanmıştır.

b) X ile Y nin ortak olasılık yoğunluk fonksiyonu,

$$f(x,y)=\begin{cases} 1 & , \quad 0<x<1,\ x<y<x+1 \\ 0 & , \quad d.y. \end{cases}$$

olarak verilmiş olsun (Casella ve Berger, 2002, sayfa 170). $D_X=[0,1]$ ve $D_Y=[0,2]$ olmak üzere marjinal olasılık yoğunluk fonksiyonları,

$$f_X(x)=\begin{cases} 1 & , \quad 0<x<1 \\ 0 & , \quad d.y. \end{cases} \qquad \text{ve} \qquad f_Y(y)=\begin{cases} y & , \quad 0<y<1 \\ 2-y & , \quad 1<y<2 \\ 0 & , \quad d.y. \end{cases}$$

dir. Burada, X rasgele değişkeninin marjinal olasılık yoğunluk fonksiyonu,

$$f_X(x)=\int_{y=x}^{x+1} dy = 1$$

ve Y nin marjinal olasılık yoğunluk fonksiyonu da,

$0 < y < 1$ için $f_Y(y) = \int\limits_{x=0}^{y} dx = y$ ve $1 < y < 2$ için $f_Y(y) = \int\limits_{x=y-1}^{1} dx = 2 - y$

olarak hesaplanmıştır. Marjinal olasılık yoğunluk fonksiyonlarından, X ve Y nin beklenen değer ve varyansı

$$E(X) = 1/2, \quad Var(X) = 1/12, \quad E(Y) = 1 \text{ ve } Var(Y) = 1/6$$

dır. Diğer taraftan,

$$E(XY) = \int\limits_{x=0}^{1} \int\limits_{y=x}^{x+1} x\,y\;dy\,dx = \frac{1}{2} \int\limits_{x=0}^{1} x\left((x+1)^2 - x^2\right) dx = \frac{7}{12}$$

olup, X ile Y arasındaki korelasyon

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\frac{7}{12} - \frac{1}{2}(1)}{\sqrt{\frac{1}{12}\frac{1}{6}}} = \frac{\frac{7}{12} - \frac{6}{12}}{\sqrt{\frac{1}{12}\frac{1}{6}}} = \frac{1}{\sqrt{2}}$$

dir.

c) $(\Omega, \mathcal{U}, P)$ bir olasılık uzayı olsun. $A, B \in \mathcal{U}$ için $P(A) = 0.5$, $P(B|A) = 0.4$ ve $P(A \cup B) = 0.8$ olmak üzere, X ve Y rasgele değişkenleri

$$X : \Omega \to \mathbb{R} \qquad\qquad Y : \Omega \to \mathbb{R}$$

$$w \to X(w) = \begin{cases} 1 & , \quad w \in A \\ 0 & , \quad w \notin A \end{cases} \qquad\qquad w \to Y(w) = \begin{cases} 1 & , \quad w \in B \\ 0 & , \quad w \notin B \end{cases}$$

şeklinde tanımlansın. Her iki rasgele değişkenin değer kümesi aynıdır ($D_X = D_Y = \{0,1\}$). Buna göre, X in olasılık fonksiyonu,

$$P(X = 1) = P(A) = 0.5 \text{ ve } P(X = 0) = P(A^c) = 1 - 0.5 = 0.5$$

den,

$$P(X = x) = (0.5)^x (0.5)^{1-x} \;, \; x = 0,1$$

şeklinde yazılabilir. Y nin olasılık fonksiyonu için

$$P(A \cap B) = P(A)P(B|A) = (0.5)(0.4) = 0.2$$

olduğundan B olayının olasılığı,

$$P(B) = P(A \cup B) - P(A) + P(A \cap B) = 0.8 - 0.5 + 0.2 = 0.5$$

şeklinde olup, Y nin olasılık fonksiyonu da aynıdır. Yani,

$$P(Y = 1) = P(B) = 0.5 \text{ ve } P(Y = 0) = P(B^c) = 1 - 0.5 = 0.5$$

olasılıklarının hesabından, Y nin olasılık fonksiyonu

$$P(Y=y)=(0.5)^y(0.5)^{1-y}\ ,\ y=0,1$$

şeklinde bulunur. Yani, X ve Y nin olasılık fonksiyonları aynıdır. Ayrıca,

$$P(X=0,Y=0)=P(A^c\cap B^c)=1-P(A\cup B)=1-0.8=0.2$$

$$P(X=0,Y=1)=P(A^c\cap B)=P(B)-P(A\cap B)=0.5-0.2=0.3$$

$$P(X=1,Y=0)=P(A\cap B^c)=P(A)-P(A\cap B)=0.5-0.2=0.3$$

$$P(X=1,Y=1)=P(A\cap B)=0.2$$

olasılıklarından ortak olasılık fonksiyonu da,

Y/X	0	1
0	0.2	0.3
1	0.3	0.2

şeklinde bulunmuştur. Buradan,

$$E(XY)=\sum_{x=0}^{1}\sum_{y=0}^{1}x\,y\,P(X=x,Y=y)=P(X=1,Y=1)=0.2$$

olup X ile Y arasındaki korelasyon

$$\rho_{X,Y}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{E(XY)-E(X)E(Y)}{\sqrt{Var(X)Var(Y)}}=\frac{0.2-(0.5)(0.5)}{\sqrt{(0.25)(0.25)}}=\frac{-0.05}{0.25}=-\frac{1}{5}$$

dir. X ve Y aynı olasılık fonksiyonlarına sahiptir ancak bağımsız değildir ⊕

Tanım 2.5.5 Herhangi bir X rasgele değişkeninin beklenen değeri μ varyansı σ^2, dağılım fonksiyonu da $F(x)$ olsun. Buna göre,

a) (Medyan) X in medyanı

$$P(X\geq M)\geq\frac{1}{2}\quad \textbf{ve}\quad P(X\leq M)\geq\frac{1}{2}$$

özelliğini sağlayan M sayısı,

b) *Varyasyon katsayısı* (Coefficient of Variation): $V=\sigma/\mu$,

c) *Çarpıklık (skewness) katsayısı* : $\gamma=E(X-\mu)^3/\sigma^3$,

d) *Basıklık (kürtosis) katsayısı* : $\eta=(E(X-\mu)^4/\sigma^4)-3$

dür ⊗

Örnek 2.5.5 X rasgele değişkeninin olasılık fonksiyonu $\lambda>0$ için

$P(X=x)=e^{-\lambda}\lambda^{x}/x!\,,\ x=0,1,2,3,...$

olarak verilsin. Moment çıkaran fonksiyonu yardımı ile ilk üç moment ve merkezi momentler,

$$E(X)=\lambda\,,\ \ E(X^2)=\lambda+\lambda^2\ ,\ E(X-\mu)^2=\lambda\,,\quad E(X^3)=\lambda^3+3\lambda^2+\lambda$$

$$E(X-\mu)^3=\lambda\,,\quad E(X-\mu)^4=3\lambda^2+\lambda\quad \text{ve}\quad Var(X)=\lambda$$

olarak hesaplanmıştır. Buradan da çarpıklık ve basıklık katsayıları sırası ile,

$$\gamma=\frac{E(X-\mu)^3}{\sigma^3}=\frac{\lambda}{\lambda^{3/2}}=\frac{1}{\sqrt{\lambda}}$$

$$\eta=\frac{E(X-\mu)^4}{\sigma^4}-3\ =\ \frac{\lambda+3\lambda^2}{\lambda^2}-3\ =\ \frac{1}{\lambda}$$

dır$\oplus$

Çok değişkenli dağılım fonksiyonlarından (Kısım (2.4)) koşullu olasılık fonksiyonları ve buradan da koşullu beklenen değer bulunabilir. Örneğin, $X_1,\ldots,X_k$ ve Y rasgele değişkenlerinin ortak olasılık veya olasılık yoğunluk fonksiyonu $f(x_1,\ldots,x_k,y)$ olsun. $X_1=x_1,\ldots,X_k=x_k$ verildiğinde Y rasgele değişkeninin koşullu olasılık veya olasılık yoğunluk fonksiyonu $f(y\,|\,x_1,\ldots,x_k)$ olup $X_1=x_1,X_2=x_2,\ldots,X_k=x_k$ verildiğinde Y nin koşullu beklenen değeri

$$E(Y\,|\,X_1=x_1,X_2=x_2,\ldots,X_k=x_k)=\int_{y\in D_Y} y\,f(y\,|x_1,\ldots,x_k)\,d\,y$$

dir (Y kesikli ise formülde integral yerine toplam sembolü gelir). Bu koşullu beklenen değer $x_1,\ldots,x_k$ nin bir fonksiyonudur. Yani,

$$E(Y\,|\,X_1=x_1,\ldots,X_k=x_k)=h(x_1,\ldots,x_k)$$

dir. Bir değişkenin diğer değişkenler üzerine regresyonu (burada Y nin $X_1,\ldots,X_k$ ler üzerine regresyonu) bu koşullu beklenen değerdir. Regresyon konusu kitabın son bölümünde ayrıntılı olarak incelenecektir. Buradaki $h(x_1,\ldots,x_k)$ fonksiyonu $x_1,\ldots,x_k$ lerin lineer birleşimi ise regresyon, linner regresyon denklemi, değilse lineer olmayan regresyon denklemi adını alır. Koşullu beklenen değer $E(Y\,|\,X_1=x_1)=h(x_1)$ ve $h(x_1)=a+b\,x_1$ şeklinde ise regresyona basit doğrusal regresyon denir.

Teorem 2.5.1 X ve Y rasgele değişkenlerinin ortak olasılık veya olasılık yoğunluk fonksiyonu $f(x,y)$ olmak üzere,

$$E(E(X\,|Y)) = E(X) \quad \text{ve} \quad Var(X) = E(Var(X\,|Y)) + Var(E(X\,|Y))$$

dir.

İspat: İspatı sürekli durum için yapalım (kesikli durumda integraller yerine toplamlar gelir). X ve Y nin ortak olasılık yoğunluk fonksiyonu $f(x,y)$ ise, $Y = y$ verildiğinde X in koşullu olasılık yoğunluk fonksiyonu $f(x\,|\,y)$ olsun. Buradan koşullu beklenen değer,

$$\begin{aligned}
E\left(E(X\,|Y)\right) &= \int_{y\in D_Y} E\left(X\,|Y=y\right) f_Y(y)\,dy = \int_{y\in D_Y}\left(\int_{x\in D_X} x f(x\,|\,y)dx\right) f_Y(y)\,dy \\
&= \int_{x\in D_X} x\left(\int_{y\in D_Y} f(x\,|\,y) f_Y(y)dy\right)dx = \int_{x\in D_X} x\left(\int_{y\in D_Y} \frac{f(x,y)}{f_Y(y)} f_Y(y)dy\right)dx \\
&= \int_{x\in D_X} x\left(\int_{y\in D_Y} f(x,y)dy\right)dx = \int_{x\in D_X} x f_X(x)dx = E(X)
\end{aligned}$$

şeklinde bulunur. Varyansın tanımından, X in varyansı

$$\begin{aligned}
Var(X) = E([X - E(X)]^2) &= E([X - E(X\,|Y) + E(X\,|Y) - E(X)]^2) \\
&= E([X - E(X\,|Y)]^2) + E([E(X\,|Y) - E(X)]^2) \\
&\quad + 2E\left(([X - E(X\,|Y)])[E(X\,|Y) - E(X)]\right)
\end{aligned}$$

olarak yazılabilir. Bu ifadedeki son terim sıfırdır. Bunu görebilmek için,

$$E\left(\left[X - E(X\,|Y)\right]\left[E(X\,|Y) - E(X)\right]\right) = E\left(E\left\{\left[X - E(X\,|Y)\right]\left[E(X\,|Y) - E(X)\right]\right\}\right)$$

eşitliğini yazalım. Koşullu dağılımdaki $X\,|\,Y$ ve X rasgele değişkenler olup $E(X\,|Y)$ ve $E(X)$ sabittir. Dolayısı ile,

$$\begin{aligned}
&E\left(E\left\{\left[X - E(X\,|Y)\right]\left[E(X\,|Y) - E(X)\right]\right\}\right) \\
&\qquad = \left(E(X\,|Y) - E(X)\right)\left(E\left\{\left[X - E(X\,|Y)\right]|\,Y\right\}\right) \\
&\qquad = \left(E(X\,|Y) - E(X)\right)\left(E(X\,|Y) - E(X\,|Y)\right) = \left(E(X\,|Y) - E(X)\right).0 = 0
\end{aligned}$$

elde edilir. Buradan, $E\left((X - E(X\,|Y))[E(X\,|Y) - E(X)]\right) = E(0) = 0$ olup

$$E([X - E(X\,|Y)]^2) = E(E\{[X - E(X\,|Y)]^2\,|Y\}) = E(Var(X\,|Y))$$

ve

$$E([E(X\,|Y) - E(X)]^2) = Var(E(X\,|Y))$$

olduğundan,

$$Var(X) = E([X - E(X\,|\,Y)]^2) + E([E(X\,|\,Y) - E(X)]^2)$$
$$= E(Var(X\,|\,Y)) + Var(E(X\,|\,Y))$$

şeklinde aranan eşitlik elde edilir ◊

Teorem 2.5.2 X ve Y herhangi iki rasgele değişken, g de tanım kümesi reel sayılar olan herhangi bir fonksiyon olmak üzere,

$$\min_{g(x)} E([Y - g(X)]^2) = E([Y - E(Y\,|\,X)]^2)$$

dir.

İspat: Kolayca görüleceği gibi,

$$E([Y - g(X)]^2) = E([Y - E(Y\,|\,X) + E(Y\,|\,X) - g(X)]^2)$$
$$= E([Y - E(Y\,|\,X)]^2) + E([E(Y\,|\,X) - g(X)]^2)$$
$$+ 2E\big([Y - E(Y\,|\,X)]\ [E(Y\,|\,X) - g(X)]\big)$$

olup son terim sıfırdır. Dolayısı ile,

$$E([Y - g(X)]^2) = E([Y - E(Y\,|\,X)]^2) + E([E(Y\,|\,X) - g(X)]^2)$$
$$\geq E([Y - E(Y\,|\,X)]^2)$$

olup $g(X) = E(Y\,|\,X)$ olduğunda eşitlik sağlanır ◊

Örnek 2.5.6 a) X ve Y rasgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu,

$$f(x,y) = \begin{cases} e^{-y} & , \quad 0 < x < y < \infty \\ 0 & , \quad d.y. \end{cases}$$

olsun. $X = x$ verildiğinde, Y nin koşullu olasılık yoğunluk fonksiyonu,

$$f(y\,|\,x) = \begin{cases} e^{-(y-x)} & , \quad y > x \\ 0 & , \quad d.y. \end{cases}$$

şeklinde olup Y nin koşullu beklenen değeri,

$$E(Y\,|\,X = x) = \int_{y=x}^{\infty} y\, f(y\,|\,x)\, dy = \int_{y=x}^{\infty} y\, e^{-(y-x)}\, dy = e^{x} \int_{y=x}^{\infty} y\, e^{-y}\, dy$$
$$= e^{x}\left(-e^{-y} - y e^{-y} \Big|_{y=x}^{\infty} \right) = e^{x}\left(e^{-x} + x e^{-x}\right) = x + 1$$

dir.

b) X ve Y nin ortak olasılık fonksiyonu $x = 0,1,2,\ldots,n$, $y = 0,1,2,\ldots,n$ ve $x + y \leq n$ için,

$$f(x,y) = \binom{n}{x, y, n-x-y} \theta_1^{x}\, \theta_2^{y} (1 - \theta_1 - \theta_2)^{n-x-y}$$

olarak verilmiş olsun. Burada,

$$\binom{n}{x, y, n-x-y} = \frac{n!}{x!\,y!\,(n-x-y)!}$$

dir. X rasgele değişkeninin marjinal olasılık fonksiyonu $x = 0, 1, 2, \ldots, n$ için,

$$f_X(x) = \sum_{y=0}^{n-x} f(x,y) = \binom{n}{x}\theta_1^x (1-\theta_1)^{n-x} = \frac{n!}{x!(n-x)!}\theta_1^x (1-\theta_1)^{n-x}$$

olup $X = x$ verildiğinde Y nin koşullu olasılık fonksiyonu da

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)} = \binom{n-x}{y}\left(\frac{\theta_2}{1-\theta_1}\right)^y \left(\frac{1-\theta_1-\theta_2}{1-\theta_1}\right)^{n-x-y} \quad , \; y = 0,1,2,\ldots,n-x$$

olarak bulunur. Buradan, $X = x$ verildiğinde Y nin koşullu beklenen değeri

$$E(Y \mid X = x) = \sum_{y=0}^{n-x} y\, f(y \mid x) = \sum_{y=0}^{n-x} y\, P(Y = y \mid X = x)$$

$$= \sum_{y=0}^{n-x} y \binom{n-x}{y}\left(\frac{\theta_2}{1-\theta_1}\right)^y \left(\frac{1-\theta_1-\theta_2}{1-\theta_1}\right)^{n-x-y} = \frac{(n-x)\theta_2}{1-\theta_1} = \frac{n\theta_2}{1-\theta_1} - \frac{\theta_2}{1-\theta_1}x = \alpha + \beta x$$

olur. Burada, $\alpha = n\theta_2/(1-\theta_1)$ ve $\beta = -\theta_2/(1-\theta_1)$ dir.

c) X ve Y kesikli iki rasgele değişken ve ortak olasılık fonksiyonları $h(x,y)$ olsun. $E(Y \mid X = x) = a + bx$ ise a ve b değerlerini bulmak isteyelim. X ve Y kesikli rasgele değişkenlerinin marjinal olasılık fonksiyonları $f(x)$ ve $g(y)$ olmak üzere, $X = x$ verildiğinde Y nin koşullu beklenen değeri

$$E(Y \mid X = x) = \sum_{y \in D_Y} y\, h(y \mid x) = \sum_{y \in D_Y} y \left(\frac{h(x,y)}{f(x)}\right) = a + bx$$

dir. Başka bir ifade ile,

$$\sum_{y \in D_Y} y\, h(x,y) = (a + bx)\, f(x)$$

dir. Diğer taraftan,

$$E(Y) = \sum_{y \in D_Y} y\, g(y) = \sum_{y \in D_Y} y \left(\sum_{x \in D_X} h(x,y)\right) = \sum_{x \in D_X} \left(\sum_{y \in D_Y} y\, h(x,y)\right)$$

$$= \sum_{x \in D_X} (a + bx)\, f(x) = a + b\, E(X)$$

dir. Benzer şekilde X ile Y arasındaki korelasyon ρ olmak üzere,

$$\rho\sqrt{Var(X)Var(Y)}+E(X)E(Y)=E(XY)=\sum_{x\in D_X}\sum_{y\in D_Y}(x\,y)\,h(x,y)=\sum_{x\in D_X}x\left(\sum_{y\in D_Y}y\,h(x,y)\right)$$

$$=\sum_{x\in D_X}x\left(a+b\,x\right)f(x)=\sum_{x\in D_X}(ax+b\,x^2)\,f(x)=a\,E(X)+b\,E(X^2)$$

$$=a\,E(X)+b(Var(X)+[E(X)]^2)$$

eşitliği yazılabilir. Bu iki eşitlik, $E(Y)=a+b\,E(X)$ ve

$$\rho\sqrt{Var(X)Var(Y)}+E(X)E(Y)=a\,E(X)+b\left(Var(X)+\left(E(X)^2\right)\right)$$

şeklinde yazıldığında iki bilinmeyenli iki denklemin çözümünden,

$$a=E(Y)-\rho\left(\sqrt{\frac{Var(Y)}{Var(X}}\right)E(X) \quad \text{ve} \quad b=\rho\left(\sqrt{\frac{Var(Y)}{Var(X}}\right)$$

bulunur$\oplus$

Teorem 2.5.3 X ve Y herhangi iki rasgele değişken ve $a,b\in\mathbb{R}$ olsun. Buna göre,

a) $E(aX+bY)=a\,E(X)+b\,E(Y)$

b) $Var(aX+bY)=a^2Var(X)+b^2Var(Y)+2\,a\,b\,Cov(X,Y)$

dir.

*İspat:*Kesikli durumu göz önüne alalım (sürekli durumda toplam yerine integral gelir).

a) Beklenen değerin tanımından aranan eşitlik

$$E(aX+bY)=\sum_{x\in D_X}\sum_{y\in D_Y}(a\,x+b\,y)\,P(X=x,Y=y)$$

$$=a\sum_{x\in D_X}x\sum_{y\in D_Y}P(X=x,Y=y)+b\sum_{x\in D_X}\sum_{y\in D_Y}y\,P(X=x,Y=y)$$

$$=a\sum_{x\in D_X}xP(X=x)+b\sum_{y\in D_Y}y\,P(Y=y)$$

$$=a\,E(X)+b\,E(Y)$$

şeklinde gösterilmiş olur. Bu da Teoremin (a) kısmını tamamlar.

b) Varyansın tanımında $E(aX+bY)$ in yukarıdaki değeri yerine konursa,

$$Var(aX+bY)=E\left([(aX+bY)-E(aX+bY)]^2\right)=E\left([a(X-E(X))+b(Y-E(Y))]^2\right)$$

$$=a^2E([X-E(X)]^2)+b^2E([Y-E(Y)]^2)+2\,a\,b\,E([X-E(X)][Y-E(Y)])$$

$$=a^2\,Var(X)+b^2\,Var(Y)+2\,a\,b\,Cov(X,Y)$$

şeklinde aranan sonuç elde edilir$\Diamond$

2.6. Üretici Fonksiyonlar

Bu kısımda, rasgele değişkenlerin beklenen değerine bağlı bazı üretici fonksiyonlar ve bunların kullanıldığı yerler üzerinde durulacaktır. Bunlardan, moment çıkaran fonksiyonu ile karekteristik fonksiyon bir önceki kısımda tanımlanmıştı.

1. Moment Çıkaran Fonksiyonu:

Bu fonksiyonun tanımı bir önceki kısımda (Tanım(2.5.3a)) verildi. X in moment çıkaran fonksiyonu, $t \in \mathbb{R}$ için $M_X(t) = E(e^{tX})$ dir. Moment çıkaran fonksiyonu yardımı ile var olması halinde herhangi bir rasgele değişkenin bütün momentlerinin hesaplanabileceğini biliyoruz. Ayrıca, rasgele değişkenin bütün momentleri biliniyorsa, rasgele değişkenin olasılık veya olasılık yoğunluk fonksiyonuna ihtiyaç duyulmadan moment çıkaran fonksiyonu da bulunabilir. $M_X(t) = E(e^{tX})$ olduğundan beklenen değer ile sonsuz toplamın yer değiştirebildiği varsayımı altında, e^{tX} fonksiyonunun Taylor serisi açılımından moment çıkaran fonksiyonu,

$$M_X(t) = E\left(e^{tX}\right) = E\left(\sum_{k=0}^{\infty} \frac{(X\,t)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E(X^k)$$

şeklinde de ifade edilebilir. Yani, rasgele değişkenin bütün momentleri ile moment çıkaran fonksiyonu arasında bir ilişki de vardır. Ancak, buradaki geçişin yapılabilmesi için sonsuz toplam ile beklenen değer operatörünün yer değiştirebilir olması gerekir. Benzer durum, moment çıkaran fonksiyonundan momentlere geçiş için de vardır. Orada da beklenen değer operatörü ile türev operatörlerinin yer değiştirebilmesi varsayımı yapılır. Bu yer değiştirebilme varsayımları momentlerin var olmasından dolayı genellikle sağlanır.

Örnek 2.6.1 a) X rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$f(x) = \begin{cases} e^{-x} & , \quad x > 0 \\ 0 & , \quad d.y. \end{cases}$$

olarak verilmiş olsun. X rasgele değişkeninin moment çıkaran fonksiyonu, $t < 1$ için

$$M_X(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} f(x)\,dx = \int_0^{\infty} e^{tx} e^{-x}\,dx = \int_0^{\infty} e^{-x(t-1)} dx = \frac{1}{1-t} e^{-x(t-1)} \Big|_{x=0}^{\infty} = \frac{1}{1-t}$$

dir. Rasgele değişkenin momentlerinin

$$\frac{d^k}{dx^k} M_X(t) \Big|_{t=0} = E(X^k e^{tX}) \Big|_{t=0} = E(X^k) = \mu_k$$

şeklinde hesaplanabildiğini bir önceki kısımdan biliyoruz. Moment çıkaran fonksiyonunu kullanarak, bu rasgele değişkenin beklenen değer ve varyansını bulalım. Birinci moment, yani rasgele değişkenin beklenen değeri,

$$M_X(t)=\frac{1}{1-t}\Rightarrow\frac{d}{dt}M_X(t)=\frac{1}{(1-t)^2}\Rightarrow E(X)=\left.\frac{d}{dt}M_X(t)\right|_{t=0}=\left.\frac{1}{(1-t)^2}\right|_{t=0}=1$$

olup ikinci momenti de

$$M_X(t)=\frac{1}{1-t}\Rightarrow\frac{d^2}{dt^2}M_X(t)=\frac{2}{(1-t)^3}\Rightarrow E(X^2)=\left.\frac{d^2}{dt^2}M_X(t)\right|_{t=0}=\left.\frac{2}{(1-t)^3}\right|_{t=0}=2$$

dir. Dolayısı ile, rasgele değişkenin varyansı $Var(X)=E(X^2)-(E(X))^2=2-1=1$ dir.

b) Herhangi bir X rasgele değişkeninin bütün $k\in\mathbb{N}$ için $E(X^k)=2^k\,k!$ olarak verilsin. Bu rasgele değişkenin moment çıkaran fonksiyonu için yukarıdaki formülde $E(X^k)=2^k\,k!$ yazıldığında moment çıkaran fonksiyonu, $t<(1/2)$ olmak üzere,

$$M_X(t)=E(e^{tX})=E\left(\sum_{k=0}^{\infty}\frac{(X\,t)^k}{k!}\right)=\sum_{k=0}^{\infty}\frac{t^k}{k!}E(X^k)=\sum_{k=0}^{\infty}\frac{t^k}{k!}(2^k\,k!)=\sum_{k=0}^{\infty}(2t)^k=\frac{1}{1-2t}$$

olur. Aslında, bir X rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x)=\begin{cases}\frac{1}{2}e^{-x/2} & ,\quad x>0\\ 0 & ,\quad d.d.\end{cases}$$

şeklinde verilmiş ise, bu rasgele değişkenin $k.$ nci momenti

$$E(X^k)=\int_0^{\infty}x^k f(x)dx=\frac{1}{2}\int_0^{\infty}x^k e^{-x/2}dx=\frac{1}{2}\Gamma(k+1)\,2^{k+1}=2^k\,k!$$

ve moment çıkaran fonksiyonu da $t<(1/2)$ olmak üzere

$$M_X(t)=E(e^{tX})=\int_0^{\infty}e^{tx}f(x)dx=\frac{1}{2}\int_0^{\infty}e^{tx}e^{-x/2}dx=\frac{1}{2}\int_0^{\infty}e^{-x(0.5-t)}dx=\frac{1}{2\big((1/2)-t\big)}=\frac{1}{1-2t}$$

dir$\oplus$

2. Kümülant Üreten Fonksiyonu

Tanım 2.6.1 X rasgele değişkeninin moment çıkaran fonksiyonu var ve $M_X(t)$ olsun. Buna göre, X in *kümülant üreten fonksiyonu,*

$$K_n(t)=\frac{d^n}{dt^n}\ln\big(M_X(t)\big)$$

dir$\otimes$

X rasgele değişkeninin moment çıkaran fonksiyonu (dolayısı ile, kümülant üreten fonksiyonu) bilindiğinde X in n. kümülantı

$$K_n = \frac{d^n}{d\,t^n} \ln\left(M_X(t)\right)\Big|_{t=0}$$

formülü ile bulunur.

Örnek 2.6.2 X rasgele değişkeninin kümülant üreten fonksiyonu

$$K_1(t) = \frac{d}{d\,t} \ln\left(M_X(t)\right) = \frac{M'_X(t)}{\left(M_X(t)\right)}$$

olup birinci kümülant değeri

$$K_1 = \frac{d}{d\,t} \ln\left(M_X(t)\right)\Big|_{t=0} = \frac{M'_X(t)}{M_X(t)}\Big|_{t=0} = E(X)$$

dir. İkinci kümülant değeri de,

$$K_2 = \frac{d^2}{d\,t^2} \ln\left(M_X(t)\right)\Big|_{t=0} = \frac{M'_X(t) M_X(t) - \left(M'_X(t)\right)^2}{\left(M_X(t)\right)^2}\Bigg|_{t=0} = E(X^2) - \left(E(X)\right)^2 = Var(X)$$

dir. Türevler ardışık olarak devam ettirildiğinde üçüncü ve dördüncü kümülantların,

$$K_3 = E(X^3) - 3E(X^2)E(X) + 2(E(X))^3$$
$$K_4 = E(X^4) - 4E(X^3)E(X) - 3(E(X^2))^2 + 12E(X^2)(E(X))^2 - 6(E(X))^4$$

şeklinde olduğu görülür. X rasgele değişkeninin olasılık fonksiyonu $\lambda > 0$ için

$$P(X = x) = e^{-\lambda} \lambda^x / x!\,,\ x = 0, 1, 2, 3, \ldots$$

şeklinde verildiğinde, X in moment çıkaran fonksiyonu $M_X(t) = e^{\lambda(e^t - 1)}$ olup kümülant üreten fonksiyonu,

$$K_n(t) = \frac{d^n}{d\,t^n} \ln\left(M_X(t)\right) = \frac{d^n}{d\,t^n} \ln\left(e^{\lambda(e^t - 1)}\right) = \lambda\, e^t$$

dir. Buradan kümülantlar bütün $n \in \mathbb{N}$ ler için,

$$K_n = \frac{d^n}{d\,t^n} \ln\left(M_X(t)\right)\Big|_{t=0} = \lambda\, e^t\Big|_{t=0} = \lambda$$

dir. Yani, bu rasgele değişkenin bütün kümülantları aynıdır $\oplus$

3. Çarpımsal Moment Üreten Fonksiyonu

Bu fonksiyonun tanımı da bir önceki kısımda (Tanım (2.5.3c)) verildi. X rasgele değişkeninin olasılık fonksiyonu, $f(x)$ olmak üzere, X in çarpımsal moment üreten fonksiyonu $t \in \mathbb{R}$ için

$N_X(t) = E(t^X)$ dir. Rasgele değişkenin çarpımsal üreten fonksiyonu verildiğinde çarpımsal momentler,

$$E\big(X(X-1)(X-2)...(X-(k+1))\big) = \left.\frac{d^k N_X(t)}{d\,t^k}\right|_{t=1}$$

şeklinde hesaplanır.

Örnek 2.6.3 X rasgele değişkeninin olasılık fonksiyonu $\lambda > 0$ için

$P(X = x) = e^{-\lambda}\lambda^x / x!\,,\; x = 0,1,2,3,...$

olarak verilmiş olsun. Bu rasgele değişkenin çarpımsal moment üreten fonksiyonu,

$$N_X(t) = E(t^X) = \sum_{x=0}^{\infty} t^x P(X = x) = e^{-\lambda}\sum_{x=0}^{\infty} t^x \frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(t\lambda)^x}{x!} = e^{-\lambda}e^{\lambda t} = e^{\lambda(t-1)}$$

olup çarpımsal momentler,

$$E(X) = \left.\frac{d}{dt}N_X(t)\right|_{t=1} = \lambda e^{\lambda(t-1)}\Big|_{t=1} = \lambda$$

$$E(X(X-1)) = \left.\frac{d^2}{dt^2}N_X(t)\right|_{t=1} = \lambda^2 e^{\lambda(t-1)}\Big|_{t=1} = \lambda^2$$

$$E(X(X-1)(X-2)...(X-(k+1))) = \left.\frac{d^k}{dt^k}N_X(t)\right|_{t=1} = \lambda^k e^{\lambda(t-1)}\Big|_{t=1} = \lambda^k$$

şeklindedir$\oplus$

4. Karekteristik Fonksiyon

Karekteristik fonksiyonun tanımı da daha önce (Tanım (2.5.3b)) verildi. Karekteristik fonksiyon, üretici fonksiyonlar içinde önemli bir yer tutar. Bir rasgele değişkenin moment çıkaran fonksiyonu bazen olmayabilir. Ancak, karekteristik fonksiyonu her zaman vardır. Ayrıca, karekteristik fonksiyon biliniyorsa, olasılık veya olasılık yoğunluk fonksiyonu bulunabilir. $i = \sqrt{-1}$ olmak üzere, X rasgele değişkeninin karekteristik fonksiyonu $t \in \mathbb{R}$ için $\phi_X(t) = E(e^{itX})$ dir. Daha açık olarak, X in karekteristik fonksiyonu

$$\phi_X(t) = E(e^{itX}) = \begin{cases} \displaystyle\sum_{x\in D_X} e^{itx}P(X = x) & , \quad X \text{ kesikli} \\ \displaystyle\int_{x\in D_X} e^{itx} f(x)\,dx & , \quad X \text{ sürekli} \end{cases}$$

dir. X rasgele değişkeninin karekteristik fonksiyonu $\phi_X(t)$ ise momentler (var olması halinde),

$$E(X^k)=\frac{d^k}{i^k\,dt^k}\phi_X(t)\bigg|_{t=0}$$

formülü ile hesaplanır.

Örnek 2.6.4 X rasgele değişkeninin olasılık yoğunluk fonksiyonu $\beta>0$ için

$$f(x)=\begin{cases}(1/\beta)\,e^{-x/\beta} & , \quad x>0\\ 0 & , \quad d.y.\end{cases}$$

şeklinde verilmiş olsun. X in karekteristik fonksiyonu,

$$\phi_X(t)=E(e^{itX})=\frac{1}{\beta}\int_0^\infty e^{itx}e^{-x/\beta}dx=\frac{1}{\beta}\int_0^\infty e^{-x(1-it\beta)/\beta}dx=(1-it\beta)^{-1}$$

olup ilk iki moment,

$$E(X)=\frac{d}{i\,dt}\phi_X(t)\bigg|_{t=0}=\frac{1}{i}\left(\frac{i\beta}{(1-it\beta)^2}\right)\bigg|_{t=0}=\beta$$

$$E(X^2)=\frac{d^2}{i\,dt^2}\phi_X(t)\bigg|_{t=0}=\frac{1}{i^2}\left(\frac{2i^2\beta^2}{(1-it\beta)^3}\right)\bigg|_{t=0}=2\beta^2$$

dir. Rasgele değişkenin varyansı da $Var(X)=E(X^2)-\left(E(X)\right)^2=2\beta^2-\beta^2=\beta^2$ dir$\oplus$

Örnek 2.6.5 X rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x)=\frac{1}{\pi}\frac{1}{1+x^2} \quad , \quad x\in\mathbb{R}$$

olarak verilmiş olsun. Bu rasgele değişkenin moment çıkaran fonksiyonu tanımlı değildir. Oysa, X in karekteristik fonksiyonu Euler formülü olarak bilinen $e^{itx}=\cos(tx)+i\sin(tx)$ eşitliği kullanılarak bulunabilir. X in karekteristik fonksiyonunu,

$$\phi_X(t)=E(e^{itX})=\frac{1}{\pi}\int_{-\infty}^{\infty}e^{itx}\left(\frac{1}{1+x^2}\right)dx=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\cos(tx)+i\sin(tx)}{1+x^2}dx$$

$$=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\cos(tx)}{1+x^2}dx+\frac{i}{\pi}\int_{-\infty}^{\infty}\frac{\sin(tx)}{1+x^2}dx=\frac{1}{\pi}\int_{-\infty}^{\infty}h_1(x)\,dx+\frac{i}{\pi}\int_{-\infty}^{\infty}h_2(x)\,dx=I_1+I_2$$

şeklinde yazılalım ve I_1 ve I_2 değerlerini hesaplayalım. Önce,

$$h_2(-x)=\frac{\sin(-tx)}{1+(-x)^2}=-\frac{\sin(tx)}{1+x^2}=-h_2(x)$$

olduğundan $h_2(x)$ fonksiyonu tektir. Tek bir fonksiyonun simetrik bir bölge üzerinden integrali de sıfırdır. Yani,

$$I_2 = \frac{i}{\pi}\int_{-\infty}^{\infty} \frac{\sin(t\,x)}{1+x^2}\,dx = 0$$

dır. Benzer şekilde,

$$h_1(-x) = \frac{\cos(-t\,x)}{1+(-x)^2} = \frac{\cos(t\,x)}{1+x^2} = h_1(x)$$

olduğundan $h_1(x)$ fonksiyonu çifttir. Buna göre, karekteristik fonksiyon,

$$\phi_X(t) = E(e^{itX}) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\cos(t\,x)}{1+x^2}\,dx = \frac{2}{\pi}\int_{0}^{\infty} \frac{\cos(t\,x)}{1+x^2}\,dx = e^{-|t|}$$

olarak bulunur (Billingsley, 1986) $\oplus$

Karekteristik fonksiyonunun olasılık teorisinde önemli olduğunu belirtmiştik. Şimdi, bu fonksiyonun özelliklerinden bazılarını kısaca hatırlayalım.

a) herhangi bir X rasgele değişkeni, örnek uzaydan reel sayılara giden reel değerli bir fonksiyondur. Reel değerli bir X rasgele değişkeninin karekteristik fonksiyonu, $e^{itX} = \cos(t\,X) + i\sin(t\,X)$ gibi kompleks değerli bir rasgele değişkenin beklenen değeridir. Ayrıca,

$$|e^{itX}| = |\cos(t\,X) + i\sin(t\,X)| = 1 \quad \text{ve} \quad \int_{-\infty}^{\infty} |e^{itx}|\, f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

dir.

b) $\phi_X(0) = 1$ olduğu açıktır.

c) $|\phi_X(t)| = |E(e^{itX})| \le E(|e^{itX}|) = 1$

olup her $t \in \mathbb{R}$ için $|\phi_X(t)| \le \phi_X(0)$ dır.

d) z kompleks bir sayı olsun (yani $a, b \in \mathbb{R}$ için $z = a + i\,b$ ve z nin kompleks konjugesi $\bar{z} = a - i\,b$ dir). Herhangi bir X rasgele değişkeninin karekteristik fonksiyonu $\phi_X(t)$ olmak üzere,

$$\phi_X(-t) = E(e^{-itX}) = E(\cos(t\,X) - i\sin(t\,X)) = E(\cos(t\,X)) - i\,E(\sin(t\,X)) = \overline{\phi}_X(t)$$

dir. Yani, $\phi_X(t)$ karekteristik fonksiyonu *Hermitian özelliğine* sahiptir.

e) X rasgele değişkeninin karekteristik fonksiyonu $\phi_X(t)$ ise $a, b \in \mathbb{R}$ olmak üzere $Y = b + a\,X$ nin karekteristik fonksiyonu da

$$\phi_Y(t) = E(e^{itY}) = E(e^{it(b+aX)}) = e^{itb}\phi_X(a\,t)$$

şeklindedir.

f) Bir rasgele değişkenin karekteristik fonksiyonu $\phi_X(t)$ biliniyorsa, dağılım fonksiyonu da bulunabilir. Aşağıdaki teorem bunu ifade etmektedir.

Teorem 2.6.1 Herhangi bir X rasgele değişkeninin dağılım fonksiyonu $F(x)$, karekteristik fonksiyonu da $\phi_X(t)$ olsun. h pozitif bir sayı olmak üzere $F(x)$ nin sürekli olduğu yerlerde,

$$\frac{F(x+h)-F(x)}{h}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itx}-e^{-it(x+h)}}{ith}\phi_X(t)\,dt$$

ve

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\phi(t)\,dt$$

dir (Billingsley, 1986, sayfa 357) ◊

Örnek 2.6.6 a) Karekteristik fonksiyonu, $\phi_X(t)=e^{-|t|}$ olan X rasgele değişkeninin olasılık yoğunluk fonksiyonunu bulalım. Bunun için, $\cos(x)=(e^{ix}+e^{-ix})/2$ eşitliğinden X in olasılık yoğunluk fonksiyonu (Teorem (2.6.1))

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\phi_X(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}e^{-|t|}\,dt=\frac{1}{2\pi}\left[\int_{-\infty}^{0}e^{-itx}e^{-t}\,dt+\int_{0}^{\infty}e^{-itx}e^{-t}\,dt\right]$$

$$=\frac{1}{2\pi}\left[\int_{0}^{\infty}e^{itx}e^{t}\,dt+\int_{0}^{\infty}e^{-itx}e^{-t}\,dt\right]=\frac{1}{2\pi}\left[\int_{0}^{\infty}e^{-t}\left(e^{itx}+e^{-itx}\right)dt\right]$$

$$=\frac{1}{\pi}\int_{0}^{\infty}e^{-t}\cos(tx)\,dx=\frac{1}{\pi}\left(-e^{-t}\cos(tx)\Big|_{t=0}^{\infty}\right)-\frac{x}{\pi}\int_{0}^{\infty}e^{-t}\sin(tx)\,dx$$

$$=\frac{1}{\pi}(0+1)-\frac{x}{\pi}\int_{0}^{\infty}e^{-t}\sin(tx)\,dx=\frac{1}{\pi}-\frac{x^2}{\pi}\left[\frac{e^{-t}}{1+x^2}\left(x\sin(tx)-\cos(tx)\right)\right]_{t=0}^{\infty}$$

$$=\frac{1}{\pi}-x^2\left(\frac{1}{\pi}\frac{1}{1+x^2}\right)=\frac{1}{\pi}\left[1-\frac{x^2}{1+x^2}\right]=\frac{1}{\pi}\frac{1}{1+x^2}$$

şeklinde bulunur. Yani, X rasgele değişkeninin olasılık yoğunluk fonksiyonu $x\in\mathbb{R}$ için,

$$f(x)=\frac{1}{\pi}\frac{1}{1+x^2}$$

dir.

b) X rasgele değişkeninin karekteristik fonksiyonu $\mu\in\mathbb{R}$ ve $\sigma>0$ için,

$$\phi_X(t)=e^{i\mu t-t^2\sigma^2/2}=\exp(i\mu t-t^2\sigma^2/2)$$

olarak verildiğinde yine Teorem (2.6.1) den olasılık yoğunluk fonksiyonu için

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\phi_X(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\exp\left(i\mu t-\frac{t^2\sigma^2}{2}\right)dt$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-\frac{t^2\sigma^2}{2}-\frac{t^2\sigma^2}{2\sigma^2}(ix-i\mu)\right)dt$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-\frac{\sigma^2}{2}\left[t^2+2t\left(\frac{ix-i\mu}{\sigma^2}\right)+\left(\frac{ix-i\mu}{\sigma^2}\right)^2\right]+\frac{\sigma^2}{2}\left(\frac{ix-i\mu}{\sigma^2}\right)^2\right)dt$$
$$=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{\sigma^2}{2}\left(\frac{ix-i\mu}{\sigma^2}\right)^2\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{\sigma^2}{2}\left[t+\left(\frac{ix-i\mu}{\sigma^2}\right)\right]^2\right)dt$$

ifadesi elde edilir. Bu son ifade,

$$u=\sigma\left(t+\left(\frac{ix-i\mu}{\sigma^2}\right)\right)\Rightarrow dt=\frac{du}{\sigma}$$

dönüşümü ile,

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-u^2/2}du=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

haline gelir. Yani, X in olasılık yoğunluk fonksiyonu, $x\in\mathbb{R}$ için

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

dir.