HAFTA 10

EŞİTSİZLİKLER

Olasılık ve istatistikte eşitsizlikler önemli bir yer tutar. Bazen, olasılıkların ve rasgele değişkenin momentlerinin hesaplanması güç olabilir. Böyle durumlarda, olasılık ve momentler için bir alt veya üst sınır verilir. Bu bölümde, olasılık ve momentlere ilişkin bazı eşitsizliklerden bahsedilecektir.

4.1. Markov ve Chebyshev Eşitsizlikleri

4.1.1. Markov Eşitsizliği

Teorem 4.1.1 (*Markov Eşitsizliği*) X bir rasgele değişken g de $g:\mathbb{R}\to\mathbb{R}^+$ şeklinde tanımlı negatif değerler almayan bir fonksiyon olsun. Bu durumda, $r\in\mathbb{R}^+$ için

$$P\big(g(X)>r\big) \;\le\; \frac{E\big(g(X)\big)}{r}$$

dir.

İspat: X sürekli olsun (kesikli durumda integral yerine toplam gelir). X in olasılık yoğunluk fonksiyonu $f(x)$ olmak üzere, $g(X)$ in beklenen değeri doğrudan yazıldığında

$$\begin{aligned} E\big(g(X)\big) &= \int_{-\infty}^{\infty} g(x)\,f(x)\,dx = \int_{\{g(x)>r\}} g(x)\,f(x)\,dx + \int_{\{g(x)\le r\}} g(x)\,f(x)\,dx \\ &\ge \int_{\{g(x)>r\}} g(x)\,f(x)\,dx \ge r\int_{\{g(x)>r\}} f(x)\,dx = r\,P(g(X)>r) \end{aligned}$$

şeklinde bir eşitsizlik elde edilir. Burada, $g(x)$ negatif olmayan değerler aldığından,

$$g(x)>r \Rightarrow \int_{\{g(x)>r\}} g(x)\,f(x)\,dx \ge r\int_{\{g(x)>r\}} f(x)\,dx$$

dir. Buradan da, $r\in\mathbb{R}^+$ için $E(g(X)) \ge r\,P(g(X)>r)$ şeklinde aranan eşitsizlik ispat edilmiş olur ◊

4.1.2. Chebyshev Eşitsizliği

Markov eşitsizliğinin bir sonucu olan *Chebyshev eşitsizliği* olasılık ve istatistikte en çok kullanılan eşitsizliklerden biridir. Herhangi bir X rasgele değişkeninin beklenen değeri μ, varyansı σ^2 olsun. g fonksiyonunu $g(x)=(x-\mu)^2/\sigma^2$ şeklinde tanımlayalım. Buna göre, $E(g(X))=1$ olup, Markov eşitsizliğinden

$$P(|X-\mu|>k\sigma)=P(|X-\mu|^2>k^2\sigma^2)=P(|X-\mu|^2/\sigma^2>k^2)$$
$$=P((X-\mu)^2/\sigma^2>k^2)=P(g(X)>k^2)\leq\frac{E(g(X))}{k^2}=\frac{1}{k^2}$$

elde edilir. Yani, beklenen değeri μ, varyansı σ^2 olan bir X rasgele değişkeni için

$$P(|X-\mu|>k\sigma)\leq 1/k^2$$

eşitsizliği yazılır. Bu eşitsizlik bazen, $P(|X-\mu|\leq k\sigma)\geq 1-(1/k^2)$ olarak da ifade edilir ve literatürde *Chebyshev eşitsizliği* olarak bilinir. Chebyshev eşitsizliği genellikle olasılıklar için bir alt veya üst sınır belirlemek için kullanılır.

Örnek 4.1.1 Düzgün bir paranın 10 defa atılması deneyi için X rasgele değişkeni gelen turaların sayısı olsun. X in olasılık fonksiyonu,

$$P(X=x)=\binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}\quad, x=0,1,2,...,10$$

olup $E(X)=\mu=5$, $Var(X)=\sigma^2=2.5$ dır. $P(|X-\mu|\leq k\sqrt{2.5})$ olasılığı ise

$$P(|X-\mu|\leq k\sqrt{2.5})=P(|X-\mu|\leq k\sigma)$$
$$=P(5-k\sqrt{2.5}\leq X\leq 5+k\sqrt{2.5})=P(k_1\leq X\leq k_2)$$

şeklinde yazılılabilir. Chebyshev sınırı $1-1/k^2$ olmak üzere, k nin bazı değerleri için bu olasılıklar hesaplanarak aşağıda verilmiştir.

k	k_1	k_2	$P(\|X-\mu\|\leq k\sqrt{2.5})$	Chebyshev Sınırı
1	3.42	6.58	0.656250	0.000000
1.5	2.63	7.37	0.890625	0.555556
2	1.84	8.16	0.978516	0.750000
2.5	1.05	8.95	0.978516	0.840000
3	0.26	9.74	0.998047	0.888889
3.5	–0.53	10.53	1.000000	0.918367

Tablodaki olasılık değerlerinden $k=2$ durumu için $P(|X-\mu|\leq k\sqrt{2.5})$ olasılığı,

$$p=P\left(|X-\mu|\leq k\sqrt{2.5}\right)=P\left(5-2\sqrt{2.5}\leq X\leq 5+\sqrt{2.5}\right)=P(1.84\leq X\leq 8.16)$$
$$=\sum_{x=2}^{8}P(X=x)=\sum_{x=2}^{8}\binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}=0.978515625\cong 0.978516$$

dır. Diğer olasılıklar benzer şekilde hesaplanır $\oplus$

Markov eşitsizliğinde rasgele değişkenin olasılık veya olasılık yoğunluk fonksiyonu ve g fonksiyonunun özel seçimi ile birçok eşitsizlik üretilebilir. Örneğin, negatif değerler almayan X in olasılık yoğunluk fonksiyonu $f(x)$ olsun. Ayrıca, $E(X^k)$ varsa,

$$E(X^k)=\int_0^{\infty} x^k\, f(x)\,dx \geq \int_a^{\infty} x^k\, f(x)\,dx \geq a^k \int_a^{\infty} f(x)\,dx = a^k P(X>a)$$

olup $P(X>a)\ \leq E(X^k)/a^k$ eşitsizliği yazılır. Örneğin, X in olasılık yoğunluk fonksiyonu,

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

olarak verildiğinde, $P(X>a)$ olasılığı ile $E(X^k)$ değeri

$$P(X>a)=e^{-a} \quad \text{ve} \quad E(X^k)=\int_0^{\infty} x^k\, f(x)\,d\,x=\int_0^{\infty} x^k\, e^{-x}\,d\,x \geq \ \Gamma(k+1)=k!$$

şeklindedir. Bu eşitsizlikte $k=n=a$ alınırsa, $e^{-n} \leq n!\,/\,n^n$ elde edilir. Bu eşitsizlik biraz düzenlendiğinde, $n! \geq\ n^n\ e^{-n}$ veya $n!\ \geq\ (n/e)^n$ şeklinde bir eşitsizlik de elde edilebilir.

$(\Omega,\mathcal{U},P)$ bir olasılık uzayı ve Ω üzerinde tanımlı bir X rasgele değişkeninin her $w\in\Omega$ için $|X(w)|<M<\infty$ özelliğini (sınırlılık) sağladığını varsayalım. X in ikinci momenti kullanılarak

$$\begin{aligned}E(X^2)&=\int_{\mathbb{R}} x^2 f(x)\,dx=\int_{\{|x|\leq c\}} x^2 f(x)\,dx+\int_{\{|x|>c\}} x^2 f(x)\,dx \leq\ c^2\int_{\{|x|\leq c\}} f(x)\,dx+\int_{\{|x|>c\}} x^2 f(x)\,dx\\ &\leq\ c^2+M^2\int_{\{|x|>c\}} f(x)\,dx, \qquad \int_{\mathbb{R}} f(x)\,dx=1\\ &=c^2+M^2P(X>c)\end{aligned}$$

şeklinde bir eşitsizlik yazılabilir. Buradan da,

$$P(X>c)\geq\frac{E(X^2)-c^2}{M^2}$$

eşitsizliği elde edilmiş olur.

4.2. Hölder, Minkowsky ve Jensen Eşitsizlikleri

Bu eşitsizliklere geçmeden önce, rasgele değişkenler için norm kavramını hatırlayalım.

Tanım 4.2.1. X bir rasgele değişken ve $p\in\mathbb{R}^+$ için $E(X^p)$ varsa, X in *p. normu,*

$$\|X\|_p=\left[\int_{\mathbb{R}}|x|^p\ f(x)\,dx\right]^{1/p}$$

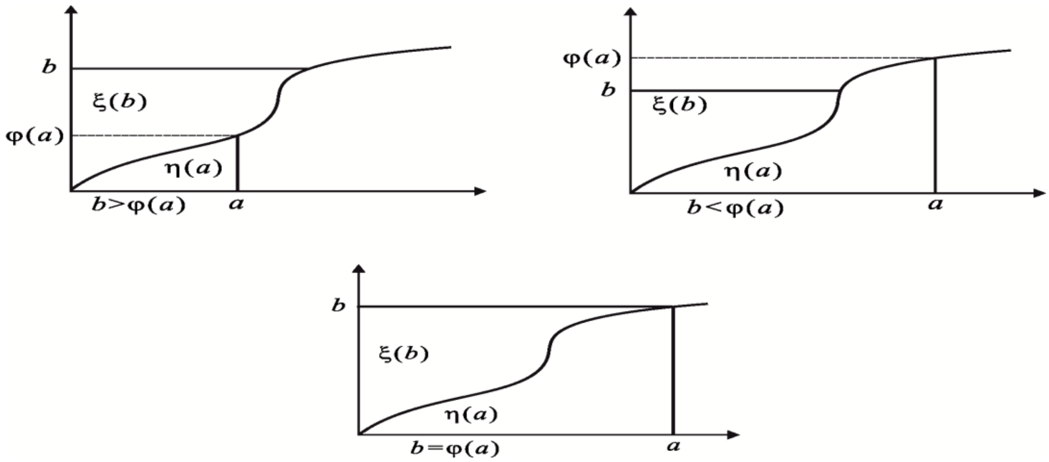
dir$\otimes$

X ve Y herhangi iki rasgele değişken ve $p\geq 1$ için $E(|X|^p)<\infty$ ve $E(|Y|^p)<\infty$ olsun. Kolayca görüleceği gibi (basit aritmetik işlemlerden sonra),

$$| X+Y |^p \le (| X |^p + | Y |^p) \le (2\max\{| X |,| Y |\})^p \le 2^p(| X |^p + | Y |^p)$$

dir (Balcı, 2000, sayfa 93). $E(| X |^p) < \infty$ ve $E(| Y |^p) < \infty$ ise $E(| X+Y |^p) < \infty$ dir.

$\varphi : \mathbb{R}^+ \to \mathbb{R}$ sürekli ve artan φ fonksiyonu $\varphi(0)=0$ ve $\lim\limits_{x\to\infty} \varphi(x)=\infty$ özelliklerine sahip olsun. $\varphi^{-1}=\psi$ olmak üzere, $\eta(x)=\int_0^x \varphi(u)\,du$ ve $\xi(x)=\int_0^x \psi(u)\,du$ diyelim.



Şekil 4.2.1 $ab \le \eta(a)+\xi(b)$ *eşitsizliğinin görsel ifadesi*

Grafiklerden de görüldüğü gibi, $ab \le \eta(a)+\xi(b)$ dir (Balcı, 2000). $b=\varphi(a)$ için eşitlik sağlanır. Grafiksel olarak da desteklenen bu bağıntı kullanılarak aşağıdaki eşitsizliği yazabiliriz.

4.2.1. Yardımcı Eşitsizlik

Teorem 4.2.1 $a,b \in \mathbb{R}^+$ ve $p,q \in \mathbb{R}^+$ ve $(1/p)+(1/q)=1$ olmak üzere,

$$\frac{a^p}{p}+\frac{b^q}{q} \ge ab$$

dir.

İspat Değişik ispat teknikleri olmakla birlikte, bu eşitsizliğin ispatını iki farklı şekilde yapacağız. $\varphi(u)=u^{p-1}$ denirse $\psi(u)=u^{1/(p-1)}$ olur. Ayrıca, $(1/p)+(1/q)=1$ olduğundan $q=p/(p-1)$ olup fonksiyondaki integral değerleri

$$\eta(a)=\int_0^a \varphi(u)\,du=\int_0^a u^{p-1}\,du=\frac{a^p}{p} \quad \text{ve} \quad \xi(b)=\int_0^b u^{1/(p-1)}\,du=\left.\frac{u^{p/(p-1)}}{p/(p-1)}\right|_{u=0}^{b}=\frac{b^q}{q}$$

şeklinde hesaplanır. Bu değerler $ab \le \eta(a)+\xi(b)$ eşitsizliğinde yerine konursa,

$$\frac{a^p}{p}+\frac{b^q}{q} \ge ab$$

eşitsizliği elde edilir. $b=\varphi(a)$ için (yani $b=u^{p-1}$) eşitlik sağlanır. İkinci ispat için,

$$g(a)=\frac{a^p}{p}+\frac{b^q}{q}-ab$$

fonksiyonunu tanımlayalım. $g(a)\geq 0$ ise ispat tamamlanmış olur. Fonksiyonun birinci türevinin sıfıra eşitlenmesi ile

$$\frac{d\,g(a)}{d\,a}=\frac{p\,a^{p-1}}{p}-b=a^{p-1}-b=0\Rightarrow a=b^{1/(p-1)}$$

bulunur. İkinci türev bu noktada pozitiftir. Yani, fonksiyonu minimum yapan değer $a=b^{1/(p-1)}$ dir. $p=q(p-1)$ olmak üzere fonksiyonun minimum değeri,

$$\frac{a^p}{p}+\frac{a^{q(p-1)}}{q}-a\,a^{p-1}=\frac{a^p}{p}+\frac{a^p}{q}-a^p=a^p\left(\frac{1}{p}+\frac{1}{q}-1\right)=0$$

dır. Yani, $g(a)\geq 0$ olup ispat tamamlanır◊

$0<p<\infty$ için rasgele değişkenlerin bir sınıfı $\mathcal{L}_p=\{X:E(|X|^p)<\infty,\ 0<p<\infty\}$ ve $p,q\in\mathbb{R}^+$ için $X\in\mathcal{L}_p$ *ve* $Y\in\mathcal{L}_q$ olsun. $(1/p)+(1/q)=1$ ise $X\,Y\in\mathcal{L}_1$ dir.

Şimdi bunun doğru olduğunu gösterelim. Önce, $|Y|\leq|X|^{p-1}$ ise $|XY|\leq|X|^p$ olup $XY\in\mathcal{L}_1$ olduğu açıktır. Diğer taraftan, $|Y|>|X|^{p-1}$ ise

$$|Y|^{1/(p-1)}>|X|\Rightarrow|Y|^{1/(p-1)}|Y|>|X||Y|\Rightarrow|Y|^q>|XY|\Rightarrow X\,Y\in\mathcal{L}_1$$

dir. Yani, her iki durumda da $X\in\mathcal{L}_p$ ve $Y\in\mathcal{L}_q$ ise $X\,Y\in\mathcal{L}_1$ dir.

4.2.2. Hölder Eşitsizliği

Teorem 4.2.2 $p,q\in\mathbb{R}^+$ için $(1/p)+(1/q)=1$ olsun. $X\in\mathcal{L}_p$ *ve* $Y\in\mathcal{L}_q$ ise

$$\|X\,Y\|_1\leq\|X\|_p\|Y\|_q$$

dir.

İspat: Önce, $\|X\|_p=0$ veya $\|Y\|_q=0$ ise ifadenin her iki tarafı da sıfır olacağından eşitsizlik sağlanır. $\|X\|_p\neq 0$ ve $\|Y\|_q\neq 0$ olsun. Yardımcı eşitsizlikte, $a=|X|/\|X\|_p$ ve $b=|Y|/\|Y\|_q$ alınırsa, X ve Y rasgele değişkenleri için

$$\frac{|X|}{\|X\|_p}\frac{|Y|}{\|Y\|_q}=\frac{|X\,Y|}{\|X\|_p\|Y\|_q}\leq\frac{|X|^p}{p\ \|X\|_p^p}+\frac{|Y|^q}{q\ \|Y\|_q^q}$$

eşitsizliği elde edilir. Her iki tarafın beklenen değeri alınıp aşağıdaki şekilde düzenlenirse,

$$E\left(\frac{|XY|}{\|X\|_p\|Y\|_q}\right) \le E\left(\frac{|X|^p}{p\|X\|_p^p}\right) + E\left(\frac{|Y|^q}{q\|Y\|_q^q}\right)$$

$$\frac{E(|XY|)}{\|X\|_p\|Y\|_q} \le \frac{E(|X|^p)}{p\|X\|_p^p} + \frac{E(|Y|^q)}{q\|Y\|_q^q} = \frac{\|X\|_p^p}{p\|X\|_p^p} + \frac{\|Y\|_q^q}{q\|Y\|_q^q} = \frac{1}{p}+\frac{1}{q} = 1$$

den

$$E\left(\frac{|XY|}{\|X\|_p\|Y\|_q}\right) = \frac{E(|XY|)}{\|X\|_p\|Y\|_q} = \frac{\|X\ Y\|_1}{\|X\|_p\|Y\|_q} \le 1$$

elde edilir. Buradan da, $\|X\ Y\|_1 \le \|X\|_p\|Y\|_q$ şeklinde aranan eşitsizlik elde edilmiş olur ◊

Hölder eşitsizliğinde rasgele değişkenlerin özel seçimleri ile olasılık ve istatistikte kullanılan bazı eşitsizlikler türetilebilir. Bunlardan bazıları aşağıdadır.

***a) Cauchy Schwartzd Eşitsizliği*:** Hölder eşistsiliği $p=q=2$ için,

$$\|XY\|_1 \le \|X\|_2\|Y\|_2$$

şeklinde yazılır. X ve Y rasgele değişkenleri için beklenen değerleri $E(X)=\mu_x$ ve $E(Y)=\mu_y$ olsun. Hölder eşitsizliği X yerine $X-\mu_x$ ve Y yerine de $Y-\mu_y$ yazılarak uygulandığında,

$$|Cov(X,Y)| = |E((X-\mu_x)(Y-\mu_y))| \le E(|(X-\mu_x)(Y-\mu_y)|)$$
$$\le \sqrt{E(X-\mu_x)^2E(Y-\mu_y)^2} = \sqrt{Var(X)Var(Y)}$$

elde edilir. Buradan, Cauchy-Schwartzd eşitsizliği,

$$|Cov(X,Y)| \le \sqrt{Var(X)Var(Y)}$$

veya

$$\left[Cov(X,Y)\right]^2 \le Var(X)Var(Y)$$

şeklinde yazılır. Bu eşitsizlikten, korelasyon katsayısı ile ilgili

$$\rho_{X,Y}^2 = \frac{\left[Cov(X,Y)\right]^2}{Var(X)Var(Y)} \le 1 \quad \text{veya} \quad -1 \le \rho_{X,Y} \le 1$$

eşitsizliği de elde edilir.

b) Hölder eşitsizliğinde $Y=1$ alalım. Bu durumda, $p\in\mathbb{R}^+$ için $\|X\|_1 \le \|X\|_p$ olup,

$$\|X\|_1 \le \|X\|_p \Rightarrow E(|X|) \le (E(|X|^p))^{1/p}$$

eşitsizliği elde edilmiş olur.

c) Hölder eşitsizliğinde $Y=1$ ve X yerine de X^r yazalım. Bu durumda, Hölder eşitsizliği $E(|X^r|) \le (E(|X^r|^p))^{1/p}$ veya $E(|X|^r) \le (E(|X|^{rp}))^{1/p}$ şekline dönüşür. Her iki tarafın $1/r$ kuvveti alındığında Hölder eşitsizliği,

$$[E(|X|^r)]^{1/r} \le [E(|X|^{rp}))^{1/(rp)}$$

olur. $p>1$ olduğundan, $s=pr$ denirse, $s>r$ dir. Dolayısı ile $s>r$ için,

$$[E(|X|^r)]^{1/r} \le [E(|X|^{rp}))^{1/(rp)} \text{ veya } \| X \|_r \le \| X \|_s$$

eşitsizliği elde edilir. Bu eşitsizlik literatürde *Liaponov eşistsizliği* olarak bilinir.

Örnek 4.2.1 Cauchy-Schwartzd eşitsizliği, vektörler için benzer şekilde yazılır. $\underset{\sim}{a}$ ve $\underset{\sim}{b}$ vektörleri, $\underset{\sim}{a}=(a_1,a_2,...,a_n)'$ ve $\underset{\sim}{b}=(b_1,b_2,...,b_n)'$ olarak verildiğinde Cauchy-Schwartzd eşitsizliği, $(\underset{\sim}{a}'\underset{\sim}{b})^2 \le (\underset{\sim}{a}'\underset{\sim}{a})(\underset{\sim}{b}'\underset{\sim}{b})$ şeklinde ifade edilir (herhangi bir lineer cebir kitabına bakılabilir). Son eşitsizlikte, $\underset{\sim}{a}=(1,1,...,1)'$ ve $\underset{\sim}{b}=(x_1,x_2,...,x_n)'$ alındığında, vektör çarpımları

$$(\underset{\sim}{a}'\underset{\sim}{b}) = \sum_{i=1}^{n} x_i\,,\quad (\underset{\sim}{a}'\underset{\sim}{a}) = n \text{ ve } (\underset{\sim}{b}'\underset{\sim}{b}) = \sum_{i=1}^{n} x_i^2$$

şeklinde olur. Cauchy-Schwartzd eşitsizliğinden,

$$\left(\sum_{i=1}^{n} x_i\right)^2 \le n\sum_{i=1}^{n} x_i^2 \text{ veya } \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \le n\left(\sum_{i=1}^{n} x_i\right)^{-2}$$

eşitsizliği elde edilir. Her i için $x_i \ne 0$, $\underset{\sim}{a}$ ve $\underset{\sim}{b}$ vektörleri de

$$\underset{\sim}{a}=(x_1,x_2,...,x_n)',\ \underset{\sim}{b}=(x_1^{-1},x_2^{-1},...,x_n^{-1})'$$

olarak seçildiğinde ise vektör çarpımları,

$$(\underset{\sim}{a}'\underset{\sim}{a}) = \sum_{i=1}^{n} x_i^2\ ,\quad (\underset{\sim}{b}'\underset{\sim}{b}) = \sum_{i=1}^{n} \frac{1}{x_i^2} \text{ ve } (\underset{\sim}{a}'\underset{\sim}{b}) = n$$

olur. $(\underset{\sim}{a}'\underset{\sim}{b})^2 \le (\underset{\sim}{a}'\underset{\sim}{a})(\underset{\sim}{b}'\underset{\sim}{b})$ eşitsizliğinde bu değerler yerine yazıldığında ise

$$n^2 \le \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} \frac{1}{x_i^2}\right) \text{ veya } \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \le \frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{x_i^2}$$

eşitsizliği elde edilir$\oplus$

4.2.3. Minkowsky Eşitsizliği

Aşağıda ifadesi verilen bu eşitsizlik, Hölder eşitsizliğine benzer. Literatürde, değişik ispat teknikleri vardır (Chow ve Teicher (1988), sayfa 111, Bauer (1981), sayfa 65).

Teorem 4.2.3 $X, Y \in \mathcal{L}_p$ olsun. Bu durumda,

$$\|X+Y\|_p \leq \| X \|_p + \| Y \|_p$$

dir.

İspat: Önce, $a, b \in \mathbb{R}$ ve $p \in \mathbb{R}^+$ için,

$$| a+b |^p \leq (| a |^p + | b |^p) \leq (2\max\{| a |,| b |\})^p \leq 2^p (| a |^p + | b |^p)$$

olduğunu hatırlayalım. $E(| X |^p) < \infty$ ve $E(| Y |^p) < \infty$ ise, $E(| X+Y |^p) < \infty$ dir. Yani, $X, Y \in \mathcal{L}_p$ ise $X+Y \in \mathcal{L}_p$ dir. Ayrıca

$$\|X+Y\|_1 = E\left(|X+Y|\right) \leq E\left(|X|\right) + E\left(|Y|\right) = \| X \|_1 + \| Y \|_1$$

olduğundan $p=1$ için eşitsizlik sağlanır. Şimdi, $p>1$ olsun. Buradan,

$$|X+Y|^p = |X+Y|\,|X+Y|^{p-1} \leq |X|\,|X+Y|^{p-1} + |Y|\,|X+Y|^{p-1}$$

olup eşitsizlikte her iki tarafın da beklenen değeri alınırsa,

$$\|X+Y\|_p^p = E(| X+Y |^p) \leq E(| X ||X+Y |^{p-1}) + E(| Y || X+Y |^{p-1})$$

eşitsizliği elde edilir. $(1/p)+(1/q)=1$ olacak şekilde q sayısı seçildiğinde, $p=q(p-1)$ olup, $X+Y \in \mathcal{L}_q$ olur. Hölder eşitsizliğinden,

$$E(| X || X+Y |^{p-1}) \leq [E(| X |^p)]^{1/p} [E(| X+Y |^{q(p-1)})]^{1/q} = \|X\|_p [E(| X+Y |^{q(p-1)})]^{1/q}$$

$$= \|X\|_p [E(| X+Y |^p)]^{\frac{1}{p}\frac{p}{q}} = \|X\|_p \|X+Y\|_p^{p/q}$$

yazılır. Yani,

$$E\left(|Y|\,|X+Y|^{p-1}\right) \leq \|X\|_p \|X+Y\|_p^{p/q}$$

dir. Buradan da,

$$\|X+Y\|_p^p = E\left(|X+Y|^p\right) \leq \|X\|_p \|X+Y\|_p^{p/q} + \|Y\|_p \|X+Y\|_p^{p/q}$$

$$= \left(\|X\|_p + \|Y\|_p\right)\|X+Y\|_p^{p/q}$$

eşitsizliği elde edilir. $\|X+Y\|_p = 0$ ise eşitsizlik sağlanır. $\|X+Y\|_p \neq 0$ ise, son eşitsizliğin her iki tarafı $\|X+Y\|_p^{p/q}$ ile bölündüğünde, $p-(p/q)=1$ olduğundan,

$$\|X+Y\|_p^p \leq \left(\|X\|_p + \|Y\|_p\right)\|X+Y\|_p^{p/q}$$

$$\Rightarrow \frac{\|X+Y\|_p^p}{\|X+Y\|_p^{p/q}} \leq \frac{\left(\|X\|_p + \|Y\|_p\right)\|X+Y\|_p^{p/q}}{\|X+Y\|_p^{p/q}}$$

$$\Rightarrow \|X+Y\|_p^{p-p/q} \leq \|X\|_p + \|Y\|_p$$

$$\Rightarrow \|X+Y\|_p \leq \|X\|_p + \|Y\|_p$$

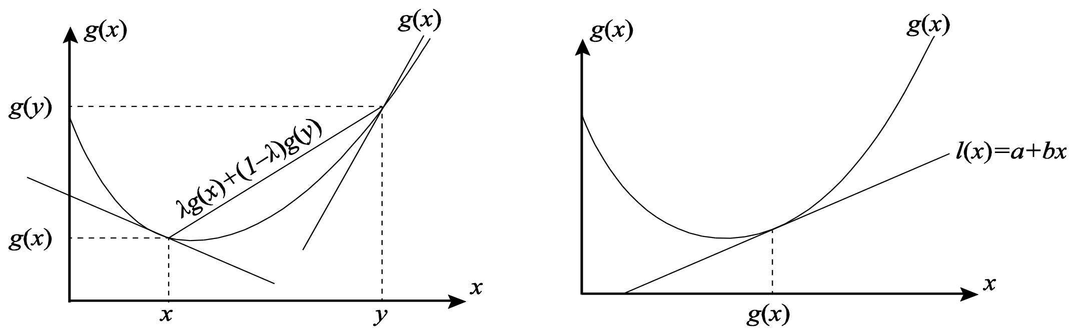
elde edilir ◊

4.2.4. Jensen Eşitsizliği

Jensen eşitsizliğini yazmadan önce konveks fonksiyonları hatırlayalım. Bir $g:\mathbb{R}\rightarrow\mathbb{R}$ fonksiyonu, $0<\lambda<1$ ve her $x,y\in\mathbb{R}$ için

$$g(\lambda x+(1-\lambda)y) \leq \lambda g(x)+(1-\lambda)g(y)$$

özelliğini sağlıyorsa, *konvekstir,* aksi halde *konkavdır* denir. Konveks $g(x)$ fonksiyonun grafiği aşağıda Şekil (4.2.2) de verildiği gibidir.



Şekil 4.2.2 Konveks fonksiyon

Türevlenebilir bir fonksiyonun ikinci türevi pozitif ise konveks, negatif ise konkavdır. Örneğin, $g(x)=x^2$ fonksiyonu $g''(x)=2>0$ olduğundan konvekstir. Benzer şekilde, $g(x)=1/x$ fonksiyonunu için $g''(x)=2/x^3$ olup, $x>0$ için fonksiyon konveks, $x<0$ için de konkavdır. Şimdi, Jensen eşitsizliğini yazabiliriz.

Teorem 4.2.4 *(Jensen Eşitsizliği):* X herhangi bir rasgele değişken, g de $\mathbb{R}$ den $\mathbb{R}$ ye herhangi bir konveks fonksiyon olmak üzere,

$$E\left(g(X)\right) \geq g\left(E(X)\right)$$

dir. Fonksiyon konkav ise eşitsizlik yön değiştirir. $P(g(X)=a+bX)=1$ ise eşitlik vardır.

İspat: g fonksiyonu sürekli ve türevlenebilir olsun. Bu durumda herhangi bir I aralığında, $x_0 \in I$ için x_0 değerine bağlı en az bir λ ($\lambda = \lambda(x_0)$) vardır ve her $x \in I$ için $g(x) - g(x_0) \geq \lambda(x - x_0)$ dir. g nin x_0 noktası komşuluğundaki birinci dereceden Taylor serisi açılımı

$$g(x) = g(x_0) + g'(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^2 g''(\xi)$$

olup g konveks olduğundan $g''(\xi) \geq 0$ ve $0.5\ (x - x_0)^2 g''(\xi) \geq 0$ dır. Böylece $\lambda = \lambda(x_0) = g'(x_0)$ için $g(x) - g(x_0) \geq \lambda(x - x_0)$ dır. Eşitsizlikte $x_0 = E(X)$ alınırsa, $g(X) - g(E(X)) \geq \lambda(X - E(X))$ elde edilir. Her iki tarafın da beklenen değeri alınırsa ($E(X - E(X)) = 0$),

$$E[g(X) - g(E(X))] = E(g(X)) - g(E(X)) \geq 0$$

bulunur. Yani, $E(g(X)) \geq g(E(X))$ şeklinde Jensen eşitsizliği elde edilmiş olur.

g fonksiyonu yukarıda verildiği gibi sürekli ve türevlenebilir olmayabilir. Bu durumda teoremin ispatı farklıdır. $g(x)$ konveks olduğundan, fonksiyonun $E(X)$ noktasındaki tanjant doğrusu Şekil (4.2.2) de görüldüğü gibidir. Bu doğruya (şekilde görüldüğü gibi) $\ell(x)$ diyelim. Bazı $a, b \in \mathbb{R}$ için bu tanjant doğrusu $\ell(x) = a + bx$ şeklinde yazılabilir. $g(x)$ fonksiyonu konveks olduğundan, $g(x) \geq \ell(x)$ dir. Buradan,

$$g(X) \geq \ell(X) \Rightarrow E(g(X)) \geq E(\ell(X)) = E(a + bX) = a + bE(X) = \ell(E(X)$$

yazılır. $\ell(x) = a + bx$ doğrusu $g(x)$ fonksiyonunun $E(X)$ noktasındaki tanjant değeri olduğundan, $\ell(E(X)) = g(E(X))$ dir. O halde,

$$E(g(X)) \geq \ell(E(X)) = g(E(X)) \text{ veya } E(g(X)) \geq g(E(X))$$

olur. $g(x)$ fonksiyonu lineer ise eşitlik sağlanır ◊

Örnek 4.2.2 a) $g(x) = x^2$ olsun. $g''(x) = 2 > 0$ olduğundan $g(x) = x^2$ fonksiyonu konvekstir. Jensen eşitsizliğine göre,

$$E(g(X)) = E(X^2) \geq g(E(X)) = (E(X))^2$$

dir. Yani, $E(X^2) \geq \left(E(X)\right)^2$ veya $Var(X) = E(X^2) - (E(X))^2 \geq 0$ dır.

b) $x > 0$ için $g(x) = 1/x$ konveks ($g''(x) = 2/x^3 > 0$) olup Jensen eşitsizliğinden,

$$E(g(X)) = E(1/X) \geq g(E(X)) = 1/E(X)$$

yazılır. Yani, $E(1/X) \geq 1/E(X)$ dir.

c) $g(x)=|x|$ fonksiyonu, $0<\lambda<1$ ve $x,y\in\mathbb{R}$ için,

$$g(\lambda x+(1-\lambda)y)=|\lambda x+(1-\lambda)y| \le \lambda|x|+(1-\lambda)|y|=\lambda g(x)+(1-\lambda)\, g(y)$$

olduğundan konvekstir. Jensen eşitsizliğine göre,

$E(g(X))=E(|X|)\ge g(E(X))=|E(X)|$ dir. Yani, $E(|X|)\ge|(E(X)|$ dir$\oplus$

Örnek 4.2.3 Hölder eşitsizliği Jensen eşitsizliği yardımı ile ispat edilebilir. Önce, $g(x)=\ln(x)$ fonksiyonu $g''(x)<0$ ($g''(x)=-1/x^2$) olduğundan konkavdır. $x,y\in\mathbb{R}^+$ ve $\lambda\in[0,1]$ olsun. Jensen eşitsizliğine göre ($E(g(X))\le g(E(X))$)

$$\lambda\ln(x)+(1-\lambda)\ln(y)\le\ln(\lambda x+(1-\lambda)y)$$

eşitsizliği yazılabilir. Bununla birlikte, $h(x)=\mathrm{e}^x$ fonksiyonu kesin artan olduğundan,

$$e^{\lambda\ln(x)+(1-\lambda)\ln(y)}\le e^{\ln(\lambda x+(1-\lambda)y)}$$

olup logaritma ve üstel fonksiyonun özelliklerinden ($\lambda\ln(x)=\ln(x^\lambda)$ ve $e^{\ln(a)}=a$)

$$e^{\ln(x^\lambda)+\ln(y^{(1-\lambda)})}\le e^{\ln(\lambda x+(1-\lambda)y)}\Rightarrow e^{\ln(x^\lambda)}e^{\ln(y^{(1-\lambda)})}\le e^{\ln(\lambda x+(1-\lambda)y)}$$
$$\Rightarrow x^\lambda y^{(1-\lambda)}\le\lambda x+(1-\lambda)y$$

eşitsizliğini yazabiliriz. Son eşitsizlik yardımı ile Hölder eşitsizliğinin ispatını yapabiliriz. Teorem (4.2.2) de $\lambda=1/p$ ve $1-\lambda=1/q$ alındığında $(1/p)+(1/q)=1$ dir. Eşitsizlikte, X ve Y rasgele değişkenler olmak üzere $\lambda=1/p$, $1-\lambda=1/q$, $x=|X|^p/E(|X|^p)$ ve $y=|Y|^q/E(|Y|^q)$ diyelim. Buna göre,

$$x^\lambda y^{(1-\lambda)}\le\lambda x+(1-\lambda)y\Rightarrow\frac{|X|}{(E(|X|^p)^{1/p}}\frac{|Y|}{(E(|Y|^q)^{1/q}}\le\frac{1}{p}\frac{|X|^p}{(E(|X|^p)}+\frac{1}{q}\frac{|Y|^q}{(E(|Y|^q)}$$

dir. Her iki tarafın beklenen değeri alındığında ($(1/p)+(1/q)=1$ dir)

$$\frac{E(|X|)}{(E(|X|^p)^{1/p}}\frac{E(|Y|)}{(E(|Y|^q)^{1/q}}\le\frac{1}{p}\frac{E(|X|^p)}{(E(|X|^p)}+\frac{1}{q}\frac{E(|Y|^q)}{(E(|Y|^q)}=\frac{1}{p}+\frac{1}{q}=1$$

eşitsizliği elde edilir. Buradan da,

$$\frac{E(|X|)}{(E(|X|^p)^{1/p}}\frac{E(|Y|)}{(E(|Y|^q)^{1/q}}\le1\Rightarrow\frac{E(|XY|)}{(E(|X|^p)^{1/p}(E(|Y|^q)^{1/q}}\le1$$
$$\Rightarrow E(|XY|)\le(E(|X|^p)^{1/p}(E(|Y|^q)^{1/q}$$
$$\Rightarrow\|XY\|_1\le\|X\|_p\|Y\|_q$$

şeklinde Hölder eşitsizliği elde edilir. Burada, Hölder eşitsizliği doğrudan Jensen eşitsizliğinin bir sonucu değildir. Ancak, yukarıdaki logaritma fonksiyonunun konkav olmasından dolayı elde edilen eşitsizliğin bir sonucudur $\oplus$

X ve Y beklenen değerleri mevcut herhangi iki rasgele değişken olsun. Genellikle $E(X/Y) \neq E(X)/E(Y)$ dir. Ancak, $Cov(X/Y, Y) = 0$ ise $E(X/Y) = E(X)/E(Y)$ eşitliği yazılabilir. Kovaryansın tanımından,

$$0 = Cov\left(\frac{X}{Y}, Y\right) = E\left(\frac{X}{Y}Y\right) - E\left(\frac{X}{Y}\right)E(Y) = E\left(X\right) - E\left(\frac{X}{Y}\right)E(Y)$$

eşitliğinden

$$0 = E\left(X\right) - E\left(\frac{X}{Y}\right)E(Y) \Rightarrow E\left(\frac{X}{Y}\right)E(Y) = E(X) \Rightarrow E\left(\frac{X}{Y}\right) = \frac{E(X)}{E(Y)}$$

elde edilir.