

MIT OpenCourseWare
<http://ocw.mit.edu>

14.30 Ekonomide İstatistiksel Yöntemlere Giriş
Bahar 2009

Bu materyale atıfta bulunmak ve kullanım koşulları için <http://ocw.mit.edu/terms> sayfasını ziyaret ediniz.

14.30 Ekonomide İstatistiksel Yöntemlere Giriş Ders Notları 11

Konrad Menzel

17 Mart 2009

1. Sıra İstatistikleri

X_1, \dots, X_n p.d.f.leri $f_{X_1}(x) = \dots = f_{X_n}(x)$ benzer olan bağımsız rasgele değişkenler olsun – genellikle böyle bir sıralı ifadeyi “bağımsız ve aynı(benzer) dağılımlı” olarak adlandırırız ve *i.i.d.* olarak kısaltırız. Aşağıdaki fonksiyon ile ilgileniyoruz.

$$Y_n = \max\{X_1, \dots, X_n\}$$

yani Y_n örneklemin en büyük değeridir.

Bağımsızlığı kullanarak Y_n 'nin c.d.f.sini türetebiliriz.

$$\begin{aligned} F_{Y_n}(y) &= P(Y_n \leq y) = P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y) \\ &= P(X_1 \leq y)P(X_2 \leq y) \dots P(X_n \leq y) \\ &= F_{X_1}(y)F_{X_2}(y) \dots F_{X_n}(y) \\ &= [F_X(y)]^n \end{aligned}$$

Zincir kuralını kullanarak, maksimumun p.d.f.sini elde edebiliriz.

$$f_{Y_n}(y) = \frac{d}{dy} F_{Y_n}(y) = n [F_X(y)]^{n-1} f_X(y)$$

Örnek 1. Eski bir resim bir açık arttırmada satılır. n kişi açık arttırmada bağımsız olarak B_1, \dots, B_n tekliflerini sunarlar ve tekliflerin marjinal c.d.f.si $F_B(b)$ 'dir. En yüksek teklifi veren potansiyel alıcı resmi alacak olan kişidir ve teklif miktarını ödemek zorundadır (bu tür açık artırmalar Dutch, ya da birinci fiyat açık artırması olarak bilinir). Bu durumda resim satıcısının hasılasının p.d.f.si aşağıdaki ile verilir:

$$f_Y(y) = f_{\max\{B_1, \dots, B_n\}}(y) = n [F_B(y)]^{n-1} f_B(y)$$

Şimdi bunu örneklemdaki diğer sıralamalara genelleştirebiliriz. Örneğin,

$$Y_{n-1} = "X_1, \dots, X_n\text{'nin en yüksek ikinci değeri}"$$

Bu rasgele değişken X_1, \dots, X_n 'nin $(n-1)$ 'nci sıralı istatistiği olarak adlandırılır ve onun p.d.f.sini belirleyebiliriz.

Önerme 1. X_1, \dots, X_n p.d.f.si $f_X(x)$ ve c.d.f.si $F_X(x)$ bir i.i.d. rasgele değişkenler silsilesi olsun. Bu durumda k 'nci sıralı istatistik Y_k 'nin p.d.f.si şöyledir:

$$f_{Y_k}(y) = k \binom{n}{k} [F_X(y)]^{k-1} [1 - F_X(y)]^{n-k} f_X(y)$$

İSPAT: Deneyi iki bölüme ayırabiliriz, (a) X 'lerden biri yoğunluk $f_X(y)$ 'e göre y değerini almak zorunda olsun, ve (b) y değeri veriyken, diğer çekilişler dizisi y 'nin etrafında y örneklemin en küçük k 'nci değer olacak şekilde gruplandırılabilir.

Bölüm (b) n deneyin X_1, \dots, X_n 'in n çekilişine karşılık gelen bir binom deneydir ve i 'nci turdaki "başarı" olay ($X_i \leq y$) olarak tanımlanır. Çekilişler bağımsız ve aynı p.d.f. ile ilintili oldukları için, binom dağılımdaki p parametresi $F_X(y)$ 'e eşittir. y 'nin daha küçük olması veya en küçük k 'nci değere eşit olması binom dağılımdaki en az k kadar "başarı" ile ilgilidir ve bu nedenle ilgili c.d.f aşağıdaki gibidir.

$$F_{Y_k}(y) = \sum_{l=k}^n \binom{n}{l} [F_X(y)]^l [1 - F_X(y)]^{n-l}$$

Şimdi c.d.f.'nin y 'ye göre türevini çarpım ve zincir kuralı ile alarak p.d.f.yi elde edebiliriz.

$$\begin{aligned} f_{Y_k}(y) &= \frac{d}{dy} F_{Y_k}(y) \\ &= \sum_{l=k}^n \binom{n}{l} l [F_X(y)]^{l-1} [1 - F_X(y)]^{n-l} f_X(y) - \sum_{l=k}^n \binom{n}{l} (n-l) [F_X(y)]^l [1 - F_X(y)]^{n-l-1} f_X(y) = T_1 - T_2 \end{aligned}$$

Bu ifade karmaşık görünüyor, fakat bunun esasında teleskopik bir toplam olduğu anlaşılabilir, bunda ötürü toplamı terimlerin (summand) çoğu düşecektir. İkinci terimdeki $l = n$ ile ilgili toplam değerin sıfır olduğuna dikkat ediniz. Onu yeniden yazabiliriz

$$T_2 = \sum_{l=k}^{n-1} \binom{n}{l} (n-l) [F_X(y)]^l [1 - F_X(y)]^{n-l-1} f_X(y) = \sum_{l=k+1}^n \binom{n}{l-1} (n-l+1) [F_X(y)]^{l-1} [1 - F_X(y)]^{n-l} f_X(y)$$

burada mevcut l endeksi $l-1$ ile yer değiştirmiştir. Birinci terim için aşağıdaki söz konusudur:

$$l \cdot \binom{n}{l} = \frac{n!}{l!(n-l)!} = \frac{n!}{(l-1)!(n-l+1)!} (n-l+1) = \binom{n}{l-1} (n-l+1)$$

Bu durumda birinci terim aşağıdaki gibi olur

$$T_1 = \sum_{l=k}^n \binom{n}{l-1} (n-l+1) [F_X(y)]^{l-1} [1-F_X(y)]^{n-l} f_X(y)$$

Böylece yoğunluk T_1 'i tanımlayan $l = k$ terimin toplamına eşittir çünkü T_2 'yi çıkardığımızda yok olmayan tek terimdir. Bu nedenle,

$$\begin{aligned} f_{Y_k}(y) &= T_1 - T_2 = \sum_{l=k}^n \binom{n}{l-1} (n-l+1) [F_X(y)]^{l-1} [1-F_X(y)]^{n-l} f_X(y) \\ &= \frac{n!}{k!(n-k+1)!} (n-k+1) [F_X(y)]^{k-1} [1-F_X(y)]^{n-k} f_X(y) \\ &= k \binom{n}{k} [F_X(y)]^{k-1} [1-F_X(y)]^{n-k} f_X(y) \end{aligned}$$

bu ispatı yapılan sonuçtur.

Örnek 2. *Alicının en yüksek fiyat teklifini verip resmi aldığı birinci açık artırmadan farklı bir açık artırmayı şimdi düşünebiliriz. Buna göre en yüksek fiyatı teklif eden yine resmi alır ancak bu durumda ikinci en büyük fiyat teklifi kadar ödemek yapmak zorundadır (bu açık artırma şekli ilkine göre daha yaygındır ve İngiliz yada ikinci-fiyat açık artırması olarak bilinir). Eğer teklif edilen fiyat rasgele değişkenler, C_1, \dots, C_n , ise, satıcının geliri Y şimdi aşağıdaki p.d.f.'ye sahiptir.*

$$f_Y(y) = 2 \binom{n}{2} F_C(y) [1-F_C(y)]^{n-2} f_C(y)$$

Aynı fiyat teklifini veren kişinin iki farklı açık artırma formatına farklı fiyat teklifi vermesi gerektiğini ekonomi teorisinden bildiğimiz için, fiyat teklifleri için farklı harf kullandığımıza dikkat ediniz.

2. Bir parantez açmak: “Herhangi bir sayının ilk basamağı”nın dağılımı (derste işlenmedi).

Burada, kendisi için daha önce gördüğümüz yöntemleri kullanmayacağımız, hoş ama standart olmayan bir problem var. Bu kesinlikle problem seti veya sınava hazırlık için üzerinde durmayacağınız bir şey, ama ben yine de değinmek istiyorum.

Hakkında hiçbir şey bilmediğimiz sayıların birinci basamaklarının dağılımı nedir? Daha açık olmak gerekirse, X 'in neyi temsil ettiğini veya ne tür birim ($1/Y$) kullandığını bilmediğimiz herhangi bir şeyin ölçüsü olsun. Örneğin bir gazeteyi inceleyip herhangi bir şeyi (gelir, borsa endeksleri, nüfus vs.) ölçen rakamları toplayabiliriz. Nereden geldiği hakkında başka hiçbir şey bilmediğimiz bu rakamların birinci basamaklarının p.d.f.'si nedir? Yani X ve Y 'nin pozitif olma dışında herhangi bir değer olabilecekleri bildiğimiz tek şey ise, $Z = X.Y$ rasgele değişkenin ilk ondalığının p.d.f.sini nasıl türetebiliriz?

Sezgisel olarak, uniform dağılım sanki rakamlar ve birimleri hakkında çok fazla “bilgi” içermediği için, ilk tahminimiz, ilk basamağın uniform (kesikli) dağılımlı olması olabilir. Ancak, eğer uniform dağılımı alırsak ve birimleri değiştirirsek (örneğin varsayılan dağılımda bütün rakamları ikiye ve dörde katlarsak), birinci basamakların dağılımı uniform olarak kalmaz. Örneğin, eğer gerçek rakamlar $X \sim U[1, 10]$, $4X \sim U[0, 40]$ ise, $4X$ 'in birinci basamağı Y aşağıdaki p.d.f.ye sahiptir.

$$f_Y(y) = \begin{cases} \frac{1}{40} + \frac{1}{4} & \text{eğer } y \in \{1, 2, 3\} \text{ ise} \\ \frac{1}{40} & \text{eğer } y \in \{4, 5, 6, 7, 8, 9\} \text{ ise} \\ 0 & \text{diğer bütün değerlerde} \end{cases}$$

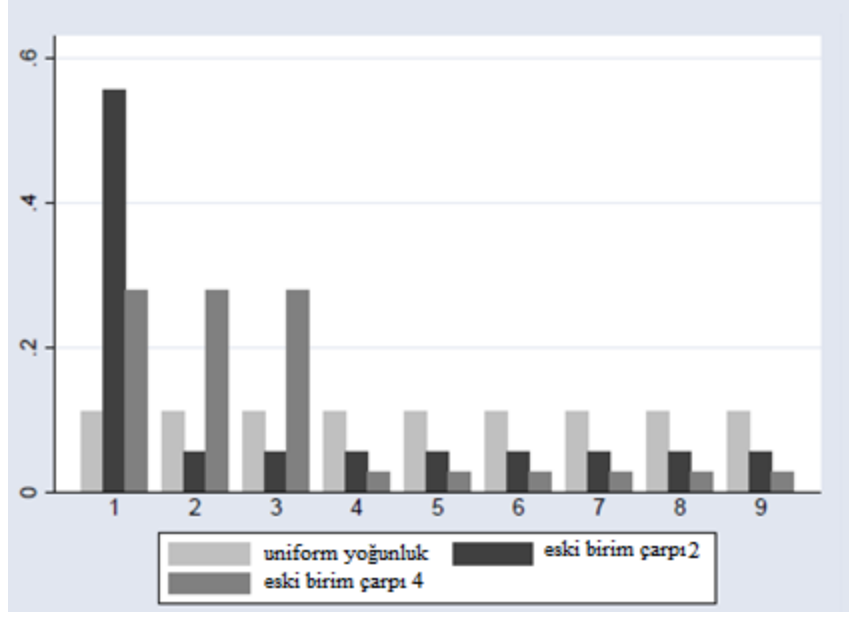
Ya da, görüldüğü gibi, birinci basamakların dağılımı hakkında minimal düzeyde gerekli olan bilgi, ölçüm birimini değiştirdiğimizde dağılımın değişmeyeceğidir.

Gerçekte aradığımız, ölçek değişimine bağlı olarak dağılımı değişmeyen bir rasgele değişken X 'tir, yani $a > X$ için aX . Eğer $Z = \log(X) \sim U[\log(1), \log(10)]$ varsayarsak bu doğrudur, çünkü bir ölçek kayması için aşağıdaki elde edilir:

$$\begin{aligned} P(z \leq aX \leq z+1) &= P\left(\frac{z}{a} \leq X \leq \frac{z+1}{a}\right) = \frac{\log\left(\frac{z}{a}\right) - \log\left(\frac{z+1}{a}\right)}{\log(10a) - \log(a)} \\ &= \frac{\log(z+1) - \log(z)}{\log(10) - \log(1)} = P(z \leq X \leq z+1) \end{aligned}$$

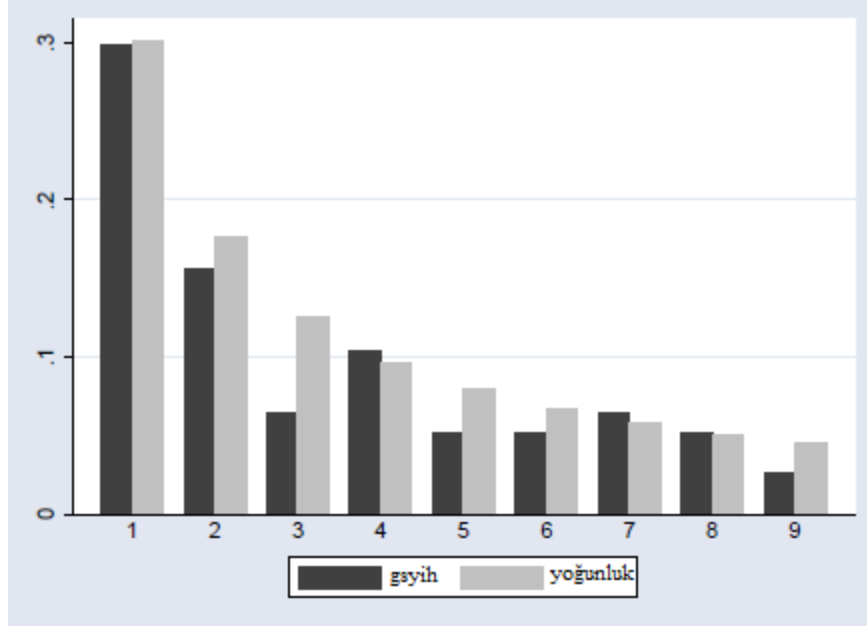
O zaman Z 'nin ilk basamağı Y aşağıdaki p.d.f.ye sahiptir.

$$f_Y(y) = \begin{cases} \frac{\log(z+1) - \log(z)}{\log(10)} & \text{eğer } y \in \{1, 2, 3, \dots, 9\} \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$



Şekil 1: Uniform Üzerinde Ölçüm Birimi Değişikliğinin Etkisi

Bu değişmeyen fikir bir dağılımı elde etmenin çok yapay bir yolu gibi görünebilir, çünkü dağılımın X ölçümü veya ölçme birimi ile çok belirgin bir bağı yoktur. Ancak, ortaya çıkan p.d.f. kategoriye düşen “gerçek-dünya verisi” hakkında çok iyi bir tahmin veriyor gibi görünüyor. Örneğin New York Times’ta görünen rakamların bir şeyi ölçmesini temsil etmeleri gibi. Aşağıdaki şekil, Economist’in “Rakamlarla Dünya 2007” cep kitabında yer alan 77 ülkenin ulusal para birimi cinsinden (Japonya için Yen, Kanada için C\$ gibi) GSYİH’ların birinci basamaklarının histogramı ile beraber “teorik” yoğunluklarını göstermektedir. Özetleyecek olursak, bu örnek verilen iki rasgele değişkenin dağılımını belirlemede değişik radikal bir yaklaşım ortaya koymaktadır: burada X ile Y ’nin p.d.f.lerini bilmeden başladık, fakat her nasıl bir dağılım ortaya çıkacaksa, birim değişikliklerinden etkilenmemek zorunda olduğunu belirttik, yani Y ’nin gerçekleşmesi gibi. “Değişmeyen (invariance)” illeri istatistikte çok önemli yeri olan bir kavramdır, fakat bu dersin amacı nedeniyle, bu örnekten öteye gitmeyeceğiz.



Şekil 2: Yerel Para Cinsinden GSYİH'ların ilk Rakamlarının Dağılımı ve Teorik Yoğunlukları (rakamlar The Economist, Pocket World in Figure 2007'den alınmadır)

3. Beklenen Değer ve Medyan

P.d.f.si $f_X(x)$ olan bir rasgele değişken X verilmişken, tüm yoğunluk dağılımını vermek zorunda kalmadan tüm dağılımın en önemli özelliklerini özetlemek istiyoruz. Beklenen değer esas itibariyle bize X 'in dağılımının nerede merkezlendiğini söyler.

3.1 Tanımlar

Tanım 1. Eğer X kesikli rasgele bir değişken ise, toplamı sonluysa, X 'in $E[X]$ ile belirtilen beklenen değeri aşağıdaki gibidir.

$$E[X] := \sum_x x f_X(x)$$

Eğer X sürekli ise, integrali sonluysa, beklenen değer aşağıdaki gibi tanımlanır

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

Örnek 3. Binom rasgele değişkenin, $X \sim B(n, p)$, beklenen değeri nedir?

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^n \frac{n!x}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^n np \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!} p^{x-1} (1-p)^{(n-1)-(x-1)} \\
&= np \sum_{x=0}^{n-1} \binom{n-1}{x} p^x (1-p)^{n-1-x} \\
&= np \cdot 1
\end{aligned}$$

burada ikinci sırada, $x = 0$ ile ilintili toplam değeri görmezlikten gelebiliriz çünkü sıfıra eşittir. Üçüncü sırada n 'yi binom katsayısından attık ve izleyen adımda, toplam endeksini x 'ten $x-1$ 'e dönüştürdük. Sonuç olarak, eğer np 'yi çekersek, toplamalar $\hat{X} \sim B(n-1, p)$ 'in binom olasılıkları olur ve bu nedenle toplamaları birdir.

Sonsuz sayıda değer alabilen bir rasgele değişkenin sonlu bir beklenen değere sahip olamayabileceğine dikkat ediniz. Bu durumda beklenen değer tanımlanmamıştır. Her ne kadar beklenen değer dağılımın “konumu” hakkında bilgi verse de, genel olarak onun rasgele bir değişken için “tipik bir değer” olmadığına da dikkat ediniz: örneğin zar atmanın beklenen değeri $\frac{1}{6}(1+2+3+4+5+6) = 3\frac{1}{2}$ 'dir ki bu mümkün olan bir sonuç değildir.

Dağılımın konumunu hesaplamamanın diğer bir seçeneği de *medyandır*.

Tanım 2. *Rasgele bir X değişkenin medyan $m(X)$ 'i reel bir sayıdır yani*

$$P(X < m) = 1/2$$

Eğer rasgele değişken X 'in dağılımı $m(X)$ etrafında simetrik ise X 'in medyan ve beklenen değeri çakışır, yani $f_x(m(X) - x) = f_x(m(X) + x)$ 'dir, fakat bu durum genelde aynı değildir.

Örnek 4. *Diyelim ki X 'in p.d.f.si şöyledir:*

$$f_X(x) = \begin{cases} \frac{1}{9}x^2 & \text{eğer } 0 \leq x \leq 3 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

Beklenen değeri ise,

$$\int_0^3 t \cdot \frac{1}{9} t^2 dt = \frac{1}{9} \int_0^3 t^3 dt = \left[\frac{1}{36} t^4 \right]_0^3 = \frac{81}{36} = \frac{9}{4} = 2.25$$

Medyanı elde etmek için, önce X 'in c.d.f.sini hesaplayalım

$$F_X(x) = \int_0^x \frac{1}{9} t^2 dt = \left[\frac{1}{27} t^3 \right]_0^x = \frac{x^3}{27}$$

Dolayısıyla, m için $F_X(m) = 1/2$ 'i çözmek aşağıdakini verir

$$m^3 = \frac{27}{2} \Leftrightarrow m = \frac{3}{\sqrt[3]{2}} \approx 2.38 > 2.25$$

Bundan ötürü, bu dağılımın medyanı ortalamasından büyüktür.

Medyanın tek olmayabileceğine dikkat ediniz.

Örnek 5. X adil bir zarın atışlarının sonucu olsun. Herhangi bir sayı için $m \in (3, 4]$, $P(X < m) = P(X \leq 3) = 1/2$ 'dir. Dolayısıyla, o aralıktaki herhangi bir sayı medyandır.

3.2 Beklenen Değerin Özellikleri

Özelik 1. Eğer $X = c$ ve c sabit bir değer ise, o zaman

$$E[X] = c$$

Özelik 2. Eğer $Y = aX + b$, ise, o zaman

$$E[Y] = aE[X] + b$$

İSPAT: Sadece sürekli duruma bakalım: Eğer X p.d.f.si $f_X(x)$ olan bir sürekli rasgele değişken ise, o zaman Y 'nin beklenen değeri şöyledir:

$$E[Y] = \int_{-\infty}^{\infty} (ax + b) f_X(x) dx = a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} f_X(x) dx = aE[X] + b \cdot 1$$

Böylece görüldüğü gibi, integralin doğrusallığı doğrudan beklenen değer in doğrusallığına dönüşür.

Özelik 3.

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n + b, \text{ için}$$

$$\mathbb{E}[Y] = a_1\mathbb{E}[X_1] + a_2\mathbb{E}[X_2] + \dots + a_n\mathbb{E}[X_n] + b$$

Bu doğrusal beklentilerin en genel durumudur ve bu özelliği bundan sonraki derslerde tekrar tekrar kullanacağız.

Örnek 6. Yukarıda, $X \sim B(n, p)$ 'nin beklenen değerini, $\mathbb{E}[X] = np$, X 'in olası bütün sonuçlarının üzerinden toplam yaparak hesapladık. Fakat son sonuçtan, aynı sonucu elde etmenin başka daha kolay bir yolu olduğunu görebiliriz: X ardışık n denemenin başarı sonucu olduğu için, her bir deneyin sonucunu Z_1, Z_2, \dots, Z_n olarak kodlayabiliriz. Burada eğer i 'nci deney başarı ise $Z_i = 1$ 'dir, diğer durumlarda $Z_i = 0$ 'dir.

$$\mathbb{E}[Z_i] = 1 \cdot p + 0 \cdot (1 - p) = p$$

ve dolayısıyla

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n Z_i\right] = \sum_{i=1}^n \mathbb{E}[Z_i] = \sum_{i=1}^n p = np$$

Özelik 4. Eğer X ile Y bağımsız ise, o zaman

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Eğer X ile Y bağımsız değil ise, bu genel olarak doğru değildir.

3.3 Rasgele Değişkenlerin Fonksiyonlarının Beklenen Değeri

$Y = r(X)$ olsun. Geçen hafta, eğer $f_X(x)$ 'i biliyorsak, Y 'nin p.d.f.sini nasıl türetebileceğimizi görmüştük. Beklenen değer için bu sorun daha kolaydır çünkü biz dağılımın sadece bir tek özeliğine bakıyoruz.

$Y = r(X)$ beklenen değeri aşağıdaki gibidir

$$\mathbb{E}[Y] = \mathbb{E}[r(X)] = \begin{cases} \sum_x r(x) f_X(x) & \text{eğer } X \text{ kesikli ise} \\ \int_{-\infty}^{\infty} r(t) f_X(t) dt & \text{eğer } X \text{ sürekli ise} \end{cases}$$

Örnek 7. Varsayalım ki, $Y = X^{1/2}$ ve X 'in p.d.f.si aşağıdaki gibidir

$$f_X(x) = \begin{cases} 2x & \text{eğer } 0 < x < 1 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

$$\mathbb{E}[Y] = \int_0^1 t^{1/2} f_X(t) dt = 2 \int_0^1 t^{3/2} dt = 2 \left[\frac{2}{5} t^{5/2} \right]_0^1 = \frac{4}{5}$$

Aynı kurallar 2 veya daha fazla rasgele değişken içinde işe yarar.

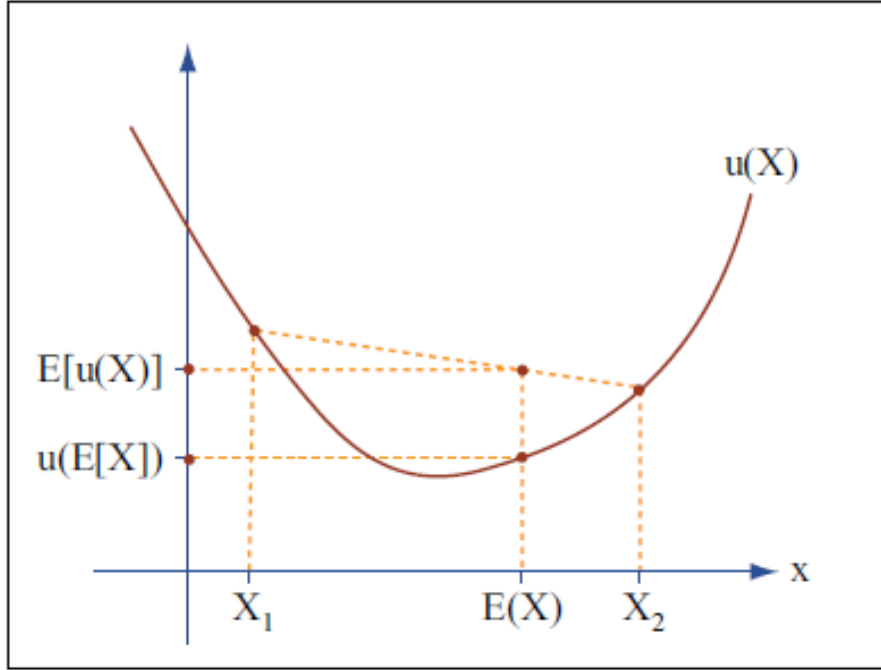
Örnek 8. Varsayalım ki birleşik p.d.f.si, aşağıda verilen X ve Y gibi iki rasgele değişkenin $Z = X^2 + Y^2$ fonksiyonu ile ilgileniyoruz.

$$f_{XY}(x, y) = \begin{cases} 1 & \text{eğer } 0 \leq x, y \leq 1 \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

O zaman

$$\begin{aligned} \mathbb{E}[Z] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2 + y^2) f_{XY}(x, y) dx dy \\ &= \int_0^1 \int_0^1 (x^2 + y^2) dx dy \\ &= \int_0^1 \left[\frac{1}{3} x^3 + y^2 x \right]_0^1 dy \\ &= \int_0^1 \left(\frac{1}{3} + y^2 \right) dy \\ &= \left[\frac{1}{3} y + \frac{1}{3} y^3 \right]_0^1 = \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \end{aligned}$$

Rasgele değişken X 'in $Y = aX + b$ türünden doğrusal fonksiyonları için, yukarıda $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$ olduğunu görmüştük. Bu doğrusal olmayan rasgele değişkenlerin fonksiyonları için çalışmaz. Bunun en tipik sonucu Jensen'in Eşitsizliği'dir:



Kaynak: MIT OpenCourseWare

Şekil 3. $\{x_1, x_2\}$ için Bir Kesikli Dağılım Örneği

Önerme 2 (Jensen'in Eşitsizliği). X rasgele bir değişken ve $u(x)$ konveks bir fonksiyon olsun. O zaman,

$$\mathbb{E}[u(X)] \geq u(\mathbb{E}[X])$$

Eğer $u(\cdot)$ kesin konveks ve X pozitif olasılıkla en az iki farklı değer alır ise eşitsizlik kesindir(strict).

İSPAT: $(\mathbb{E}[X], u(\mathbb{E}[X]))$ noktasından geçen ve $u(x)$ 'e teğet bir doğrusal fonksiyon tanımlayabiliriz:

$$r(x) = u(\mathbb{E}[X]) + u'(\mathbb{E}[X])(x - \mathbb{E}[X])$$

$u(\cdot)$ konveks olduğu için, bütün x 'ler için aşağıdaki ilişkiyi elde ederiz:

$$u(x) \geq r(x)$$

Özellikle,

$$\mathbb{E}[u(X)] = \int_{-\infty}^{\infty} u(t) f_X(t) dt \geq \int_{-\infty}^{\infty} r(t) f_X(t) dt = \mathbb{E}[r(X)]$$

$r(x)$ doğrusal olarak oluşturulduğu için, $a = u'(\mathbb{E}[X])$ ve $b = u(\mathbb{E}[X]) - u'(\mathbb{E}[X])\mathbb{E}[X]$ li doğrusal fonksiyonun beklenen değeri ile ilgili Özellik 2'yi, aşağıdakini elde etmek için kullanabiliriz.

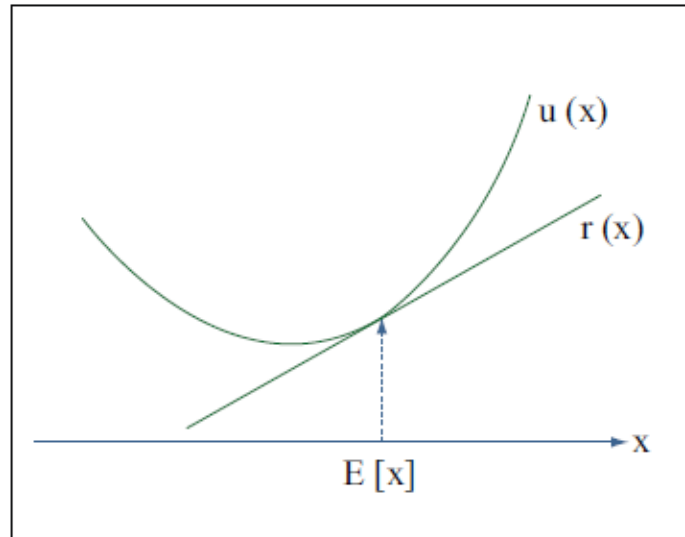
$$\mathbb{E}[r(X)] = \mathbb{E}[aX + b] = a\mathbb{E}[X] + b = u'(\mathbb{E}[X])\mathbb{E}[X] + u(\mathbb{E}[X]) - u'(\mathbb{E}[X])\mathbb{E}[X] = u(\mathbb{E}[X])$$

Bunu daha önce türetilen eşitsizlikle bir araya getirecek olursak ispat tamamlanmış olur:

$$\mathbb{E}[u(X)] \geq \mathbb{E}[r(X)] = u(\mathbb{E}[X])$$

Bir konkav fonksiyon $v(x)$ 'in negatifi $-v(x)$ konveks olduğundan, Jensen'in Eşitsizliği de bir konkav $v(\cdot)$ için aşağıdakini sağlar:

$$\mathbb{E}[v(X)] \leq v(\mathbb{E}[X])$$



Kaynak: MIT OpenCourseWare

Şekil 4. $r(x)$ her zaman $u(x)$ 'ten küçüktür.

Örnek 9 (riskten kaçınma): Varsayalım ki ilk senesi için sınırlı garantiyle gelen 1200 dolarlık bir dizüstü bilgisayar aldınız. O ilk yıl süresince, $p = \%10$ olasılıkla bir bardak kahveyi dizüstü bilgisayarın üzerine dökme (ya da sizin hatanız olan başka bir kaza) ve 1100 dolara mal olacak anakartı değiştirme ihtimaliniz var. Bu tamirat sınırlı garanti kapsamına girmez ama siz 115 dolara uzatılmış bir garanti (servis) alabilirsiniz. Bu ilave “sigorta”yı almalı mısınız?

İlave sigorta olmadan, $1-p$ olasılıkla, dizüstü bilgisayarın toplam maliyetini rasgele bir değişken olarak $X = 120$ dolar, p olasılıkla, $X = 1200 + 1100 = 2300$ dolar olacak gibi

düşünebiliriz (bu problemi farklı şekillerde oluşturmak mümkündür, ancak şimdilik her şeyi basit tutalım). Uzatılmış servis planıyla dizüstü bilgisayarınız size $X = 1200 + 115 = 1315$ dolara mal olacaktır.

Eğer siz sadece dizüstü bilgisayarın beklenen değeriyle ilgileniyorsanız, o zaman $\mathbb{E}[X] = 2300p + 1200(1 - p) = 1200 + 1100p$. Bu, eğer $p \geq \%10.45$ ise, $\mathbb{E}[Y] = 1315$ 'ten daha büyüktür. Fakat $p = \%10$ dediğimiz için, uzatılmış servis planını satın almak hala iyi bir fikir midir? - İktisatçılar, insanların belirsizlik durumunda karar aldıkları zaman, beklenen harcama miktarı W 'yla pek ilgilenmediklerini varsayarlar, ama harcadıkları dolardan elde edecekleri fayda (toplam harcama miktarında ilave bir dolarlık artışın ilave değeri olduğu için) $u(W)$ miktarıyla ilgilenirler. Bu $U(.)$ 'nin maliyette konkav olduğunu varsaydığımız anlamına gelir, diyelim ki

$$u(c) = \sqrt{4,800 - c}$$

burada başlangıçtaki varlığımızın 4800 olduğunu ve $4800 - C$ harcayabileceğimizi varsayıyoruz. C dizüstü bilgisayarın toplam maliyetidir. Bu durumda, ilave servis planına sahip olmamanın beklenen faydası

$$\mathbb{E}[u(C_1)] = 0.9\sqrt{4,800 - 1,200} + 0.1\sqrt{4,800 - 2,300} = 0.9 \cdot 60 + 0.1 \cdot 50 = 59$$

Ancak sigorta planıyla da şunu buluruz:

$$\mathbb{E}[u(C_2)] = \sqrt{4,800 - 1,315} > \sqrt{3,481} = 59$$

Gerçekten de, siz $4800 - 3481 = 119$ doları sigorta için harcamak isteyeceksiniz, hâlbuki beklenen ilave maliyet sadece $1100p = 110$ dolardır. Sigorta için ödemek istediğimiz bu 9 dolarlık farka $u(.)$ 'nin konkav olmasında gelen risk-primi denilir. Jensen'in Eşitsizliği 'ne göre, eğer $u(.)$ konkav ise bu risk-primi pozitifdir ve bu tür tercihlerin riskten kaçınma göstergesi olduğunu söyleyebiliriz.

Örnek 10. İzleyen örnek St. Petersburg Paradoksu olarak bilinir ve sonlu beklenen değeri olmayan rasgele bir değişken örneğini verir.

Bize aşağıdaki bir kumar önerilir: Varsayalım ki adil bir madeni para tura gelinceye kadar tekrar tekrar atılır. İlk atışta tura gelirse 2 dolar, 2'ncide gelirse 2^2 dolar ve genel olarak x 'nci seferde görünürse 2^x dolar kazanacaksınız.

Bu oyunu oynamak için ne kadar öderdiniz? Prensip olarak, beklenen kazancınız kadar vermek niyetinde olursunuz, bu nedenle şimdi hesaplamaları yapalım: gerek duyulan tam x atışın olasılığı aşağıdakine eşittir:

$$f_X(x) = P(x-1 \text{ yazı, } 1 \text{ tura}) = \frac{1}{2^{x-1}} \frac{1}{2} = 2^{-x}$$

Bu nedenle, beklenen kazanımlar, Y , aşağıdaki gibi hesaplanır:

$$\mathbb{E}[Y] = \sum_{x=1}^{\infty} 2^x f_X(x) = \sum_{x=1}^{\infty} \left(\frac{2}{2}\right)^x = \sum_{x=1}^{\infty} 1 = \infty$$

Bundan ötürü, beklenen kazanımlar ile ilgili üst sınır yoktur.

Bu, bu tür bahisler için insanların sonsuz miktarda ödeme yapacağını göreceğimiz anlamına mı geliyor? Kesinlikle hayır: genellikle insanlar oyunu oynamak için aşağı yukarı en fazla 25 dolar verir. Bu paradoks farklı yollardan çözülebilir:

- insanlar genellikle beklenen para miktarıyla ilgilenmez, ancak sahip oldukları toplam miktar içinde değer verdikleri para miktarı azalır, yani önceki örnekte olduğu gibi insanlar bir çeşit konkav $u(\cdot)$ fonksiyonunu maksimize ederler.
- çok küçük olasılıklarla, kazanacağınız miktar çok yüksektir- yani trilyon, katrilyon dolarlar gibi, ve ödemeyi yapacak olan karşı tarafın verdiği söze bağlı kalacağına inanmayız, bu durumda gerçekte böyle bir bahisten en iyi umutla ne kadar kazanabileceğimize dair bir çeşit üst sınır vardır.

Jensen'in Eşitsizliğiyle bağlantıyı tekrar kurmak için, $p = 1/a$ olasılıkla tura gelen bir madeni para ile oynanan oyunun beklenen atış sayısını hesaplayalım:

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} x a^{-x} = \frac{1}{a} \sum_{x=1}^{\infty} x \left(\frac{1}{a}\right)^{x-1} =: \frac{1}{a} G' \left(\frac{1}{a}\right)$$

burada kolaylıkla kontrol edebileceğiniz gibi $G'(a)$ aşağıdaki ilişkinin $1/a$ 'ya göre birinci türevidir

$$G\left(\frac{1}{a}\right) = \sum_{x=1}^{\infty} \left(\frac{1}{a}\right)^x = \frac{1}{1 - \frac{1}{a}}$$

Bu nedenle, yeni ifadenin $G(1/a)$ için türevini alınca, aşağıdakini elde ederiz

$$\mathbb{E}[X] = \frac{1}{a} G' \left(\frac{1}{a}\right) = \frac{1}{a} \frac{1}{\left(1 - \frac{1}{a}\right)^2}$$

ve örneğimizde $1/a = 1/2$ olduğu için, beklenen atış sayısı 2'dir.

Böylece, insanların hala bahis için kullanmak istedikleri 25 dolar ortalama atışların kazancına $2^{\mathbb{E}[X]} = 2^2 = 4$ dolar kadar uzaktır. Bunun açıklaması bir kere daha Jensen'in Eşitsizliği'dir, ve gerçekte $u(x) = 2^x$ x'in (ekstrem) bir konveks fonksiyonudur.

Örnek 11. Varsayalım ki iki değerli kağıt arasından seçim yapmak durumundasınız: Birincisi insanlara internet sayfalarını bedava arattıran meçhul yeni başlayan bir internet firmasının hisse senetleridir. %90 olasılıkla kar payları $e^{0t} = 1$ 'de sabit kalmasından ötürü çok risklidir ve %10 olasılıkla firmanın adı Google'dır. t zamanında her hangi bir anda ödeyeceği kar payları $e^{0.1t}$ kadar büyür, yani sırasıyla %90 olasılıkla %0, %10 olasılıkla %10 değerini alan rasgele bir büyüme oranı G_1 vardır. Diğer seçenek ise, gelecekte herhangi bir t zamanda $e^{0.02t}$ faiz ödeyecek olan devlet tahvilini tutmak olabilir, yani kesin olarak $G_2 = \%2$ 'dir.

t zamanda alacağınız bir dolara şimdi sahip olduğunuz doların $e^{-0.15t}$ 'i kadar değer verirsiniz, fakat yine de ikisinden birine yatırırsınız. Yani genel olarak, getirisi g oranında büyüyen değerli bir kağıda kesin olarak aşağıdaki kadar değeri biçersiniz:

$$V(g) = \int_0^{\infty} e^{(g-r)t} dt = \frac{0-1}{g-r} = \frac{1}{r-g}$$

Riskli hisse senedinin beklenen kar payı büyüme oranı ise,

$$\mathbb{E}[G_1] = 0.9 \cdot 0\% + 0.1 \cdot 10\% = 1\% < 2\% = \mathbb{E}[G_2]$$

Ancak, hisse senedine biçtiğiniz değer şudur:

$$\mathbb{E}[V(G_1)] = 0.9 \frac{1}{0.15-0} + 0.1 \frac{1}{0.15-0.1} = \frac{90}{15} + \frac{10}{5} = \frac{40}{5} = 8$$

Halbuki tahvil için biçtiğiniz değer aşağıdaki kadardır:

$$\mathbb{E}[V(G_2)] = \frac{1}{0.15-0.02} = \frac{100}{13} = \frac{200}{26} < \frac{200}{25} = \mathbb{E}[V(G_1)]$$

Sezgisel olarak, büyüme oranı üzerindeki belirsizlik, her ne kadar yeni faaliyete başlayanların %90'ını büyümese de (hata iflasta edebilirler), %10'u inanılmaz bir şekilde kötü giden yatırımları telafi ettikleri anlamına gelir. Biçimsel olarak, büyüme oranlarında $V(g)$ fonksiyonu konvekstir, böylece Jensen'in Eşitsizliğine göre, yatırımcılar büyüme oranındaki riske değer biçmelidirler – miktara (düzeye) değil.