

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 15

Konrad Menzel

7 Nisan 2009

1. Özel Dağılımlar (devam)

1.1 Poisson Dağılımı

Bazen, belli bir olayın belirli bir aralıkta hangi sıklıkla meydana geldiğini bilmek isteyebiliriz.

Örnek 1. *Hava yolu güvenliği söz konusu olunca, bir uçak modelinin ne kadar “güvenli” olduğu konusunda bir kanıya sahip olmak isteyebiliriz. Aşağıdaki veri www.airsafe.com veri sitesinden elde edildi ve toplam uçuş sayısı ile Aralık 2006’ya kadar belli bir uçak tipinin yer aldığı ölümcül kaza sayısını vermektedir.*

Model	Uçuşlar	Olaylar
Airbus A300	10.35M	9
Airbus A310	3.94M	6
Airbus A320/319/321	30.08M	7
Airbus A340	1.49M	0
Boeing 727	76.40M	48
Boeing 737	127.35M	64
Boeing 747	17.39M	28
Boeing 757	16.67M	7
Boeing 767	13.33M	6
Boeing 777	2.0M	0
Boeing DC9	61.69M	43
Boeing DC10	8.75M	15
Boeing MD11	1.69M	3
Boeing MD80/MD90	37.27M	14
Concorde	0.09M	1

Tablodan hemen görebileceğimiz gibi bazı uçak modelleri diğerlerine göre daha az kaza yapmış çünkü, basit bir ifadeyle, ya uzun süreli kullanılmamışlar ya da çok az

üretmişler. Daha anlamlı bir karşılaştırma yapabilmek için, ölümcül kaza sayılarının dağılımlarını açıklamak için daha iyi bir yönteme ihtiyacımız var.

Bu türden tesadüfi değişkenler genellikle “sayılabilir veri” olarak bilinir ve bu veriyi açıklamak için sıklıkla kullanılan dağılım Poisson Dağılımıdır.

Tanım 1. Eğer X aşağıdaki gibi bir p.d.f.ye sahip ise, o zaman λ parametresiyle bir Poisson Dağılımı olduğu söylenir,

$$f_X(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & \text{eğer } x \in \{0, 1, 2, \dots\} \text{ ise} \\ 0 & \text{diğer bütün durumlarda} \end{cases}$$

X 'in kesikli olduğuna özellikle dikkat ediniz.

Özelik 1. Bir Poisson rasgele değişkeni X için aşağıdaki ilişkiler yazılabilir:

$$\begin{aligned} \mathbb{E}[X] &= \lambda \\ \text{Var}(X) &= \lambda \end{aligned}$$

Sayılabilen bir değişkenin dağılımı için Poisson dağılımının neden uygun veya kabul edilebilir olduğunu görebilmek için aşağıdaki düşünsel deneyi yapalım: Varsayalım ki

- bir olayın bir $1/n$ zaman aralığında gerçekleşme olasılığı $p_n = \lambda/n$ 'dir
- zamanın herhangi bir anında gerçekleşen olayların zaman içerisinde bağımsız olduğunu da varsayıyoruz.

n 'nin sonsuza doğru gitmesine izine vererek alt-aralıkların bölüntüsünün giderek küçülmesini sağlarız. Eğer daralan iki alt aralığın olasılığı sıfıra doğru giderse, ve ondan sonra olayın en az bir kere gerçekleştiği aralıkları sayarsak, toplam gerçekleşmelerin sayısını elde etmiş oluruz. Bunun $p = \lambda/n$ ve n parametrelili bir binom rasgele değişken olduğuna dikkat ediniz.

Önerme 1. $X_n \sim B(n, \lambda/n)$ dağılımlı binom rasgele değişken için, $n \rightarrow \infty$ iken p.d.f. aşağıdakine yakınsar:

$$\lim_n f_{X_n}(x) = \lim_n \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!}$$

İSPAT: çarpımın limitini limitlerin çarpımı olarak alabiliriz ve her birini ayrı ayrı değerlendirebiliriz: Kalkülüs'te çok iyi bilinen bir sonuca göre (her iki tarafa Taylor serisi(açılımı) uygulanabilir) aşağıdaki ifadeyi yazılabilir,

$$\lim_n \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Böylece, x ve λ sabit ve bu nedenle n ile karşılaştırınca küçüldüğünden, elimizde aşağıdaki ifade kalır,

$$T_n = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x$$

Ve bunun için de şunu gösterebiliriz.

$$\lim_n T_n = \lim_n \frac{n(n-1)\dots(n-x+1)\lambda^x}{x!(n-x)!} = \frac{\lambda^x}{x!} \lim_n \frac{n}{n-\lambda} \cdot \frac{n-1}{n-\lambda} \cdots \frac{n-x+1}{n-\lambda} = \frac{\lambda^x}{x!} \cdot 1$$

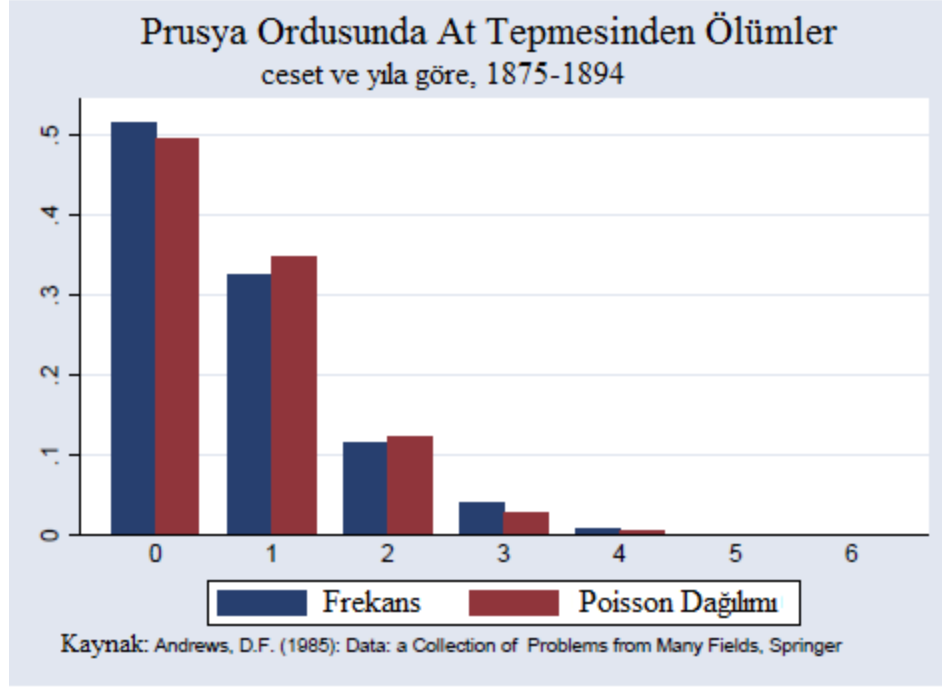
Bütün parçaları bir araya getirdiğimizde önermedeki ifadeyi elde ederiz.

Örnek 2. Sayılabilir veri için en klasik örnek (en azından İstatistiğin tarihinde bir klasiktir) 19ncu yüzyılda at tepmesi sonucunda ölen Prusya süvarileridir. 1898'de Rus istatistikçi Ladislaus Bortkiewicz'in keşfettiği gibi, Poisson dağılımı Prusya ordusunda bir tümende bir yıl içerisinde at tepmesinden ölenlerin gözlemlenmiş frekansını şaşırtıcı bir şekilde çok iyi tahmin etmiştir.

Gözlemlenmiş frekansları, en nihayetinde bilinmeyen λ 'ya bağlı olan Poisson dağılımının p.d.f.siyle nasıl karşılaştırabiliriz? Bu derste daha sonra tartışacağımız tahminin bir ön izlencesi olarak, λ 'nın kabul edilebilir bir değeri, örnekleme gözlemlendiğimiz at tepme ölümlerinin beklenen sayısının aynısını tahmin eden p.d.f.deki bir parametrenin değeri olabilir. O halde $X \sim P(\lambda)$ için, $\mathbb{E}[X]$ nedir?

Yukarıda tartıştığımız gibi, bir Poisson rasgele değişkeni Binom rasgele değişkeninin, $X_n \sim B(n, \lambda/n)$, limitidir. Burada n deneme sayısı sonsuza gitmektedir. Daha önceki Binom dağılım ile ilgili tartışmalarımıza göre, n 'den bağımsız olarak, $\mathbb{E}[X_n] = n(\lambda/n) = \lambda$ 'dir.

Böylece $\sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}$ sonsuz serilerini doğrudan kullanmadan, $\mathbb{E}[X] = \lambda$ olduğunu söyleyebiliriz.



At tepme veri setinde, örneklem ortalaması (yıl ve tümen itibariyle) $\hat{\lambda} = 0.7$ 'dir. Şimdi Şekil 2'de gösterildiği gibi örneklem frekansını $\lambda = 0.7$ parametre oranı için oluşturulan Poisson p.d.f.sinin teorik değerlerine karşı çizebiliriz. İki dağılım belirgin bir şekilde benzerdir ve bu durum sık sık "Küçük Sayılar Kanunu" olarak ifade edilir.

2. Asimptotik Teorisi

Şimdiye kadar, p.d.f.'yi, parametreleri (normal için μ ve σ^2 , üstel için λ gibi) bildiğimizi (ya da bulabileceğimizi) varsaydık ve sonra o bilgiye dayanarak olasılıklar ile ilgili ifadeler geliştirdik.

Dersin izleyen bölümünde, o bilgiye sahipmişiz gibi hareket etmeyeceğiz ama, olasılık üzerine olan bilgimizi kullanarak, söz konusu rasgele değişkenlerin dağılımı konusunda bize bir şeyler söyleyecek fonksiyonlar oluşturacağız. O fonksiyonlar *tahmin ediciler* olacaktır.

İstatistikte önemli rol oynayan bir tahmin edici *örneklem ortalaması*dır. Bu tahmin edici birazdan işleyeceğimiz gibi, rasgele bir değişkenin beklenen değerini tahmin eder.

Tanım 2. n büyüklüğündeki bir rasgele örneklem n sayıda *i.i.d*, olan, yani bütün X 'ler bağımsız ve aynı $f_X(x)$ p.d.f.sine sahip, X_1, \dots, X_n rasgele değişkenler dizisidir.

Biz sık sık *gerçekleşen* rasgele değişkenleri rasgele örneklem olarak ta ifade ederiz.

Eğer n büyüklüğünde rasgele bir örneklemimiz varsa, örneklemin ortalamasının dağılımını doğru bir şekilde açıklamak için X_i 'nin dağılımı konusunda çok şey bilmeye gerek duymayışımız (örneğin $f_X(x)$ 'i bilmeye gerek duymamak gibi) bu dersin ana mesajı olacaktır.

Buradaki ana fikir, örneklem büyüklüğü n 'i artırarak “gerçeğe” giderek yaklaşıp p.d.f.’yi tahmin etmektir. Burada oluşan iki sonuç şudur:

1. Büyük Sayılar Kanunu: Büyük n sayıları için, örneklemin ortalaması bütün olabilirliğiyle rasgele değişkenin beklenen değeri $\mathbb{E}[X]$ 'e “yaklaşacaktır”.
2. Merkezi Limit Teoremi: *Standardize* edilmiş örneklem ortalamasının (“standardize” son dersteki sıfır ortalama birim varyans manasında kullanılmaktadır) p.d.f.si standart normal rasgele değişkenin p.d.f.sine gayri ihtiyarı olarak yakın olacaktır.

Biçimsel olarak, asimptotik sonuçlar $n \rightarrow \infty$ iken neler olacağını ortaya koyar, ancak pratik uygulama açsısından (yani sonlu n için), o sonuçlar n 'in yeterince büyük olması durumunda tahminlerin oldukça doğru olacağını da ima eder.

2.1 Büyük Sayılar Kanunu

2.1.1 Chebyshev Eşitsizliği

Chebyshev Eşitsizliği bir rasgele değişkenin beklenen değerden “çok uzakta” bir yerde gerçekleşme olasılığının sınırlarını belirten biçimsel bir sonuçtur.

Önerme 2. X varyansı $\text{Var}(X) < \infty$ olan rasgele bir değişken olsun. O zaman herhangi bir $\varepsilon > 0$ için,

$$P(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}$$

İSPAT: X 'in p.d.f.si $f_X(x)$ ile verilmiş olsun. Bu durumda aşağıdaki ilişkiyi göstereceğiz

$$\text{Var}(X) \geq \varepsilon^2 P(|X - \mathbb{E}[X]| \geq \varepsilon)$$

Varyansın tanımını kullanarak şu ilişkiyi buluruz:

$$\begin{aligned}
\text{Var}(X) &= \int_{-\infty}^{\infty} (t - \mathbb{E}[X])^2 f_X(t) dt \\
&= \int_{\mathbb{E}[X]-\varepsilon}^{\mathbb{E}[X]+\varepsilon} (t - \mathbb{E}[X])^2 f_X(t) dt + \int_{-\infty}^{\mathbb{E}[X]-\varepsilon} (t - \mathbb{E}[X])^2 f_X(t) dt + \int_{\mathbb{E}[X]+\varepsilon}^{\infty} (t - \mathbb{E}[X])^2 f_X(t) dt
\end{aligned}$$

Üç integralin her birisi pozitifdir ve buna ilaveten, herhangi bir $t \leq \mathbb{E}[X] - \varepsilon$ veya $t \geq \mathbb{E}[X] + \varepsilon$ için

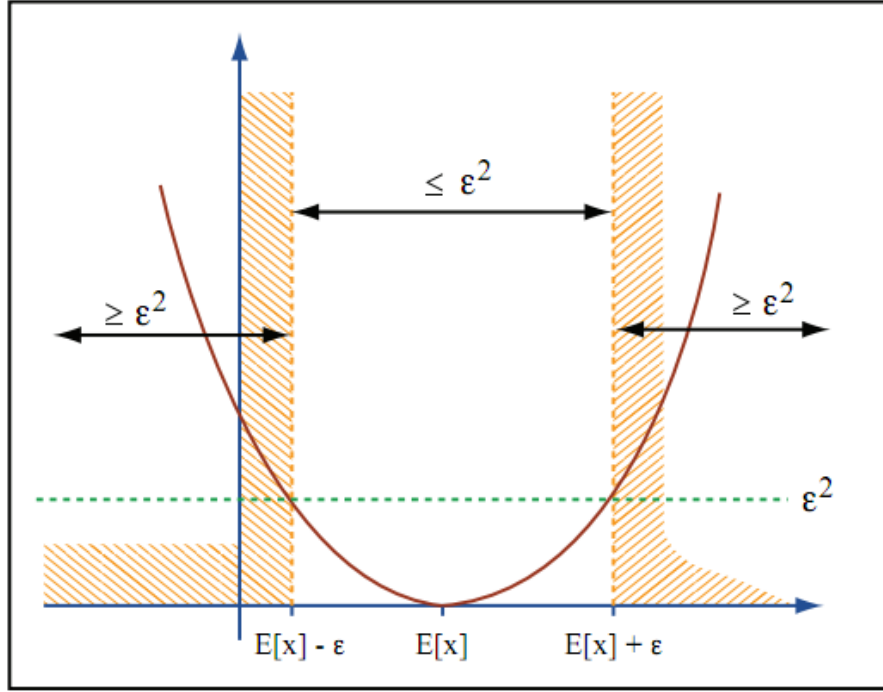
$$(t - \mathbb{E}[X])^2 \geq \varepsilon^2$$

Dolayısıyla, ilk integrali atabiliriz ve şu sonucuna ulaşabiliriz.

$$\begin{aligned}
\text{Var}(X) &\geq \int_{-\infty}^{\mathbb{E}[X]-\varepsilon} (t - \mathbb{E}[X])^2 f_X(t) dt + \int_{\mathbb{E}[X]+\varepsilon}^{\infty} (t - \mathbb{E}[X])^2 f_X(t) dt \\
&\geq \varepsilon^2 \int_{-\infty}^{\mathbb{E}[X]+\varepsilon} f_X(t) dt + \varepsilon^2 \int_{\mathbb{E}[X]+\varepsilon}^{\infty} f_X(t) dt \\
&= \varepsilon^2 P(|X - \mathbb{E}[X]| \geq \varepsilon)
\end{aligned}$$

Bu durumda her tarafı ε^2 'e bölerek sonucu elde ederiz.

Bu derste rasgele bir değişkenin varyansının değişkenin “yayılma”sının ölçüsü olduğunu daha önce bir ara söylediğimizi hatırlayınız. Chebyshev Eşitsizliği varyansı rasgele değişken X 'in “ekstrem” gerçekleşmelerini gözlemenin (yani ortalamadan çok uzakta olan değerlerin) olasılığı ile ilişkilendirerek bu ifadeyi doğrular.



Kaynak: MIT OpenCourseWare

2.1.2 Büyük Sayılar Kanun

Tanım 3. Örneklem ortalaması n büyüklüğündeki rasgele bir örneklemden elde edilen n rasgele değişkenin (gerçekleşmenin) aritmetik ortalamasıdır.

$$\bar{X}_n = \frac{1}{n} (X_1 + \dots + X_n) = \frac{1}{n} \sum_{i=1}^n X_i$$

X_i 'lerin rasgele değişken olmasından ötürü \bar{X}_n 'nin de rasgele değişken olduğunu not ediniz.

Örneklem ortalamasının beklenen değeri şöyledir:

$$\mathbb{E}[\bar{X}_n] = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] = \mathbb{E}[X_1]$$

Eğer X_1, \dots, X_n bağımsız ise, örneklem ortalamasının varyansı aşağıdaki gibi hesaplanır,

$$\text{Var}(\bar{X}_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n} \text{Var}(X_1)$$

Eğer X_i 'ler i.i.d. normal ise, $X_i \sim N(\mu, \sigma^2)$ ne olur? Normallerin doğrusal kombinasyonunun uygun varyans ve ortalama ile yine normal olduğunu biliyoruz, yani

$$\bar{X}_n \sim N\left(\mu, \frac{1}{n}\sigma^2\right)$$

n 'i artırdığımızda varyans azaldığı için, büyük bir olasılıkla ortalama en nihayetinde $\mathbb{E}[X]$ 'e çok yakın olacaktır. Esas itibarıyla Büyük Sayılar Kanunu'nun söylediği de budur.

Teorem 1 (Büyük Sayılar Kanunu). Varsayalım ki, bütün i 'ler için, X_1, \dots, X_n $\mathbb{E}[X] = \mu$ ve $\text{Var}(X_i) = \sigma^2$ ile bir i.i.d. çekiliş dizisi olsun. O zaman herhangi bir $\varepsilon > 0$ için (çok küçük bir sayı) örneklem ortalaması aşağıdaki ilişkiyi sağlar

$$\lim_n P(|\bar{X}_n - \mu| > \varepsilon) = 0$$

Bu durumda \bar{X}_n olasılıkla μ 'e yakınsar deriz.

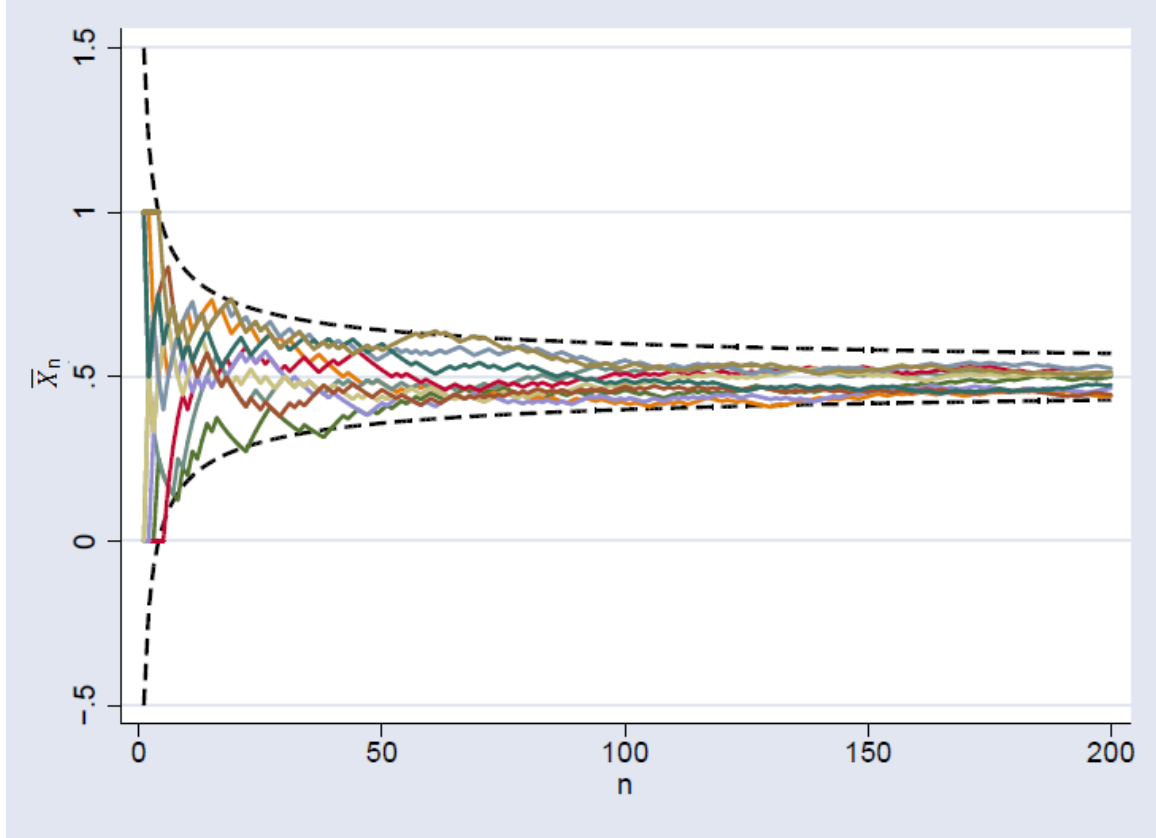
İSPAT: Önceki şu sonucumuzu kullanalım

$$\text{Var}(\bar{X}_n) = \frac{1}{n} \text{Var}(X_i) = \frac{\sigma^2}{n}$$

Chebyshev Eşitsizliği ile aşağıdaki elde edilir.

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

Bu ifade, büyük örneklem için, örnek ortalamasının büyük olasılıkla rasgele değişkenin beklenen değerinden çok uzakta olmayacağını söyler. n ve varyans σ^2 veri iken, Chebyshev Eşitsizliği'ni doğrudan kullanarak örnek ortalamasının verili mesafeden daha uzakta olma olasılığını sınırlarız.



Şekil 1. n tane madeni para atışının ortalama tura sayısının 10 dizisi- kesikli çizgiler $\pm \frac{1}{\sqrt{2n}}$ 'dir.

Örnek 3. Ölçüm Birimlerinin Standardizasyonu (Stigler'in kitabına bkz.): Ortaçağda genellikle her şehir ölçüm aracının uzunluğuna bağlı olarak "ayak", "inç" ve "yarda" vs için farklı ölçümler kullanırdı. Bu, ölçüm birimlerinde ticareti karmaşık hale getiren ve hukuki tartışmaların artmasına yol açan çok fazla farklılığın olduğu anlamına gelirdi: Örneğin verili bir kumaşın uzunluğunun gerçekten de 20 yarda uzunluğunda olup olmaması gibi yasal tartışmalar olurdu.

İnsanların bulduğu akıllıca bir çözüm şöyleydi: 16 ayak uzunluğundaki bir çubuğun uzunluğunu belirlemek için, rasgele 16 kişilik bir örneklem seçiyorsunuz (bu durumda bunlar Pazar günü Kiliseden çıkan ilk 16 kişidir), bunların ayak uzunluklarını toplayıp 16 ayak gibi bir ölçüyü belirliyorsunuz, sonra uzunluğu 16'ya bölüyorsunuz. Şekil 2'ye bakınız. 16 gözlemin ortalamasının varyansının formülüne göre, bu, yeni ölçüm biriminin varyansını $1/16$ kadar azaltmalı. Eğer bölgeler arasında ayak büyüklüklerinde (veya kiliseye giden insanlarda) sistematik bir farklılık yok ise, bu ölçü, farklı bölgelerdeki tacirlerin birbirleriyle daha kolay alış veriş yapmalarına yol açmalı.



Şekil 2. Köbel'in 16 kişinin 16 ayak uzunluğundaki bir sıriğın yasal tanımını belirlemesini betimleyen Ahşap Baskı eseri (1535)

2.1.3. Örnek: “Kalabalıkların Bilgeliği”

Varsayalım ki n büyüklüğündeki bir kitle, kamu idaresi için 2 aday arasından birini seçecektir. Burada basitçe en çok oyu alan aday kazanacaktır. Seçmen i 'nin A adayı için oy kullanması durumunda 1, diğer durumlarda 0 olan rasgele değişken X_i 'ye bakacağız. Eğer oy oranı aşağıdaki gibiyse Aday A kazanır:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n \geq \frac{1}{2}$$

Aday A şüphesiz en iyi tercihtir ancak bu kesin olarak A lehine oy kullanacakların sadece $2\varepsilon > 0$ oranı tarafından bilinmektedir, yani $i = 1, \dots, 2n\varepsilon$ için $P(X_i = 1) = 1$ 'dir. Seçmenin geri kalan $1 - 2\varepsilon$ 'u her iki aday hakkında temel bir bilgiye sahip değildir ve her iki aday için herhangi bir tercih belirtmeden oyunu kullanmaktadır, yani $i = 2n\varepsilon, \dots, n$ için, $P(X_i = 1) = P(X_i = 0) = 1/2$ 'dir. A adayının oy oranı aşağıdaki ile verilir:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i = \varepsilon + \frac{1}{n} \sum_{i=2n\varepsilon}^n X_i$$

Böylece, aday A'nın beklenen oy oranının değeri

$$\mu = 2\varepsilon + \frac{1}{2}(1 - 2\varepsilon) = \frac{1}{2} + \varepsilon$$

Binom dağılım sonuçlarına göre, varyansı ise

$$\sigma^2 = (1 - 2\varepsilon) \frac{\frac{1}{2} \left(1 - \frac{1}{2}\right)}{n} = \frac{1 - 2\varepsilon}{4n}$$

Chebyshev Eşitsizliği ile ilgili ispatın argümanlarına göre, olasılık sınırlarını çizmek için \bar{X}_n 'nin varyansından başlayabiliriz

$$\begin{aligned} \text{Var}(\bar{X}_n) &= \int_{-\infty}^{\mu-\varepsilon} (t - \mu)^2 dt + \int_{\mu-\varepsilon}^{\mu+\varepsilon} (t - \mu)^2 dt + \int_{\mu+\varepsilon}^{\infty} (t - \mu)^2 dt \\ &\geq \varepsilon^2 [P(\bar{X}_n - \mu > \varepsilon) + P(\bar{X}_n - \mu < -\varepsilon)] \end{aligned}$$

Kararsız seçmenler (gürültü) herhangi bir adayı özellikle tercih etmedikleri için, dağılım μ etrafında simetriktir, böylece

$$P(\bar{X}_n - \mu > \varepsilon) = P(\bar{X}_n - \mu < -\varepsilon)$$

$\mu = (1/2) + \varepsilon$ olduğu için, aday B'nin kaybetme olasılığı şudur:

$$\text{Var}(\bar{X}_n) \geq 2\varepsilon^2 P(\bar{X}_n - \mu < -\varepsilon) \Leftrightarrow P\left(\bar{X}_n < \frac{1}{2}\right) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{1 - 2\varepsilon}{4n\varepsilon^2}$$

Birkaç sayı deneyelim: Diyelim ki $2\varepsilon = \%5$ 'tir. Aday B'nin seçilme olasılığını $\%5$ 'in altında tutmak için n ne kadar büyük olmak zorundadır? Sınır şöyle olur,

$$P\left(\bar{X}_n < \frac{1}{2}\right) \leq \frac{90\%}{4(5\%)^2 n} = \frac{90}{n}$$

böylece $n \geq \frac{1800}{19} \approx 95$ 'dir. Eğer seçmenin $\%95$ 'i kararını rasgele verse bile, yanlış adayın seçilme olasılığını $\%5$ 'in altında tutmak için 95 yeterlidir. Bu durum “kalabalıkların bilgeliği” olarak bilinir: Seçimin sonucunda bilgilendirilmemiş seçmenin yarattığı stokastik

“gürültü” büyük örneklerde ortalamayı aşar, ancak sonuçta sadece bilgilendirilmiş seçmenlerden gelen sistematik “sinyaller” seçimin sonucunu belirler.

Büyük Sayılar Kanunu’nda olduğu gibi seçmenlerin bağımsız olduğunu varsaydığımıza dikkat ediniz. Bağımsızlık varsayımından vaz geçerse ne olur?

Varsayalım ki adaylar arasında televizyonda gösterilen bir tartışma sırasında TV stüdyosunda aday A veya B’nin yüzüne rasgele konabilecek (1/2 gibi eşit bir olasılıkla) bir sinek vardır ve bir süre etrafta uçuşan sinek adaylarda rahatsızlığa yol açmaktadır. Bilgilendirilmiş seçmenler düşüncelerini değiştirmezler, fakat bilgilendirilmemiş seçmenler 1/3 olasılıkla üzerine sinek konan, 2/3 olasılıkla da rakibi için oy kullanacaktır.

Yinelenen Olasılıklar Kanununa göre, \bar{X}_n ’in ortalaması

$$\begin{aligned}\mu &= \mathbb{E}[\bar{X}_n | \text{A'ya Sinek Konar}] \frac{1}{2} + \mathbb{E}[\bar{X}_n | \text{B'ye Sinek Konar}] \frac{1}{2} \\ &= \frac{1}{2} \left(2\varepsilon + (1 - 2\varepsilon) \frac{1}{3} \right) + \frac{1}{2} \left(2\varepsilon + (1 - 2\varepsilon) \frac{2}{3} \right) = \frac{1}{2} + \varepsilon\end{aligned}$$

dolayısıyla ortalama daha öncekinin aynısıdır. Ancak, varyans değişir: ANOVA özdeşliğine göre (koşulu varyans)

$$\text{Var}(\bar{X}_n) = \text{Var}(\mathbb{E}[\bar{X}_n | \text{fly}]) + \mathbb{E}[\text{Var}(\bar{X}_n | \text{fly})]$$

Buna göre aşağıdakileri hesaplayabiliriz

$$\begin{aligned}\text{Var}(\bar{X}_n | \text{sinek A'da}) &= \text{Var}(\bar{X}_n | \text{sinek B'de}) = \frac{2(1 - 2\varepsilon)}{9n} \\ \mathbb{E}[\bar{X}_n | \text{sinek A'da}] &= 2\varepsilon + (1 - 2\varepsilon) \frac{1}{3} = \frac{1}{3} + \frac{4}{3}\varepsilon \\ \mathbb{E}[\bar{X}_n | \text{sinek B'de}] &= 2\varepsilon + (1 - 2\varepsilon) \frac{2}{3} = \frac{2}{3} + \frac{2}{3}\varepsilon\end{aligned}$$

böylece

$$\text{Var}(\bar{X}_n) = \left(\frac{1}{6} - \frac{1}{3}\varepsilon \right)^2 + \frac{2(1 - 2\varepsilon)}{9n}$$

Adayların rolleri kararsız seçmen tarafında değiştirilebileceği için, dağılım ortalama etrafında simetriktir. Bu durumda yukarıda elde edilen olasılıklar için oluşturulan sınırları kullanabiliriz

$$P\left(\bar{X}_n < \frac{1}{2}\right) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2}$$

Ancak, ilk terim $\text{Var}(\mathbb{E}[\bar{X}_n|\text{sinek}])$ n'ye hiç bir şekilde bağlı olmadığı için, $n \rightarrow \infty$ iken $\text{Var}(\bar{X}_n)$ 'in sıfıra doğru gitmediğini artık görebiliyoruz.

Sayılar ile ifade edecek olursak, eğer $\varepsilon = \%15$ ise (önceki hesaplamaların altı katı), sınır aşağıdakine eşit olur

$$P\left(\bar{X}_n < \frac{1}{2}\right) \leq \frac{49}{81} + \frac{560}{81n}$$

böylece n ne kadar büyük olursa olsun sınır 1/2'nin üzerindedir. Bu sadece üst sınır olduğu için, bize olayın gerçekte ne kadar olası olduğunu söylemiyor, ancak varyans sıfıra düşmediği için, “gürültücü” seçmenlerin seçimin sonucu üzerinde çok güçlü bir etkisinin olacağı açıktır.

Burada Büyük Sayılar Kanunu başarısız olur çünkü sinek olayı bütün “gürültücü” seçmenleri aynı anda etkiler, böylece X_1, \dots, X_n artık bağımsız değildir. Bağımsızlık varsayımı çok önemlidir çünkü büyük sayılar kanununun genellikle çalışmasının nedeni “gürültü”nün birçok gözlemde ortalamayı aşmasıdır. Eğer “gürültü”nün bir bileşeni bütün gözlemler için ortak ise (veya en azında hepsiyle yüksek korelasyon içinde ise), bu bileşenin varyansa katkısı -örneğimizde olasılığının sınırındaki 49/81 terimi - örneklem çok büyük olsa da yok olmaz.