Markov ve Chebychev Eşitsizlikleri

$$\Pr[X \geq a] = 1 - \Pr[X - a] = 1 - \int_{-\infty}^{a} f_X(x)dx = 1 - F_X(a)$$

bilinmiyor olabilir

- r.d. nin sadece ortalamasını ve varyansını bildiğimizi varsayalım.
- Ortalama ve varyans r.d. nin oyf si veya kdf si hakkında herzaman yeterli bilgi sağlamaz
- Fakat, bazı olasılıklar için sınırları elde etmekte yardımcı olurlar.

Markov Eşitsizliği

X sadece <u>negatif olmayan</u> değerler alan bir r.d. ise

$$\Pr\left[X \geq a\right] \leq \frac{E\left[X\right]}{a}$$

İspat:

$$E[X] = \int_0^{\infty} x\, f_X(x)dx = \int_0^{a} x\, f_X(x)dx + \int_a^{\infty} x\, f_X(x)dx$$

$$\geq \int_a^{\infty} x\, f_X(x)dx \geq a\int_a^{\infty} f_X(x)dx = a\Pr\left[X \geq a\right]$$

<u>Örnek</u>:

Bir radar dönüşünün ortalama kesit alanı 3 m^2; ve kesit alanının oyf/kdf si bilinmiyor.

$$\Pr\left[X \geq 5\ \mathrm{m}^2\right] \leq \frac{3}{5} = 0.6$$

Chebychev Eşitsizliği

$$\Pr\left[\left|X - m_X\right| \geq a\right] \leq \frac{\sigma_X^2}{a^2}$$

İspat: $Y = (X - m_X)^2$ olsun ve Y ye Markov eşitsilğini uygulayalım:

$$\Pr\left[\left(X - m_X\right)^2 \geq a^2\right] \leq \frac{E\left[\left(X - m_X\right)^2\right]}{a^2} = \frac{\sigma_X^2}{a^2}$$

$|X - m| \geq a$ ile aynı olay

<u>Örnek</u>:

RKA nın ortalaması 3 m^2 varyansı 4 olduğu bilinsin

$$\Pr\left[\left|X - 3\right| \geq 5\ \mathrm{m}^2\right] \leq \frac{4}{25} = 0.16$$

Tek Taraflı Chebychev Eşitsizliği

$$\Pr\left[X - m_X \geq a\right] \leq \frac{\sigma_X^2}{\sigma_X^2 + a^2}; \qquad \Pr\left[m_X - X \geq a\right] \leq \frac{\sigma_X^2}{\sigma_X^2 + a^2}$$

($X > m_X$ veya $X \leq m_X$ olmasına göre yukarıdaki eşitsizliklerden biri kullanılabilir)

<u>Örnek</u>:

Bir direncin uçlarında ölçülen gürültü voltajı [0, 5] volt aralığında değişen birbiçim rastgele değişkendir.

$$f_X(x) = \begin{cases} 0.2, & 0 \leq x \leq 5 \\ 0, & \text{diğer} \end{cases}$$

$\Pr\left[X \geq 3.5\right]$ için üst sınırları belirleyin. Doğru olasılık aşağıdaki gibidir

$$\Pr\left[X \geq 3.5\right] = \int_{3.5}^{5} f_X(x)dx = \int_{3.5}^{5} 0.2\; dx = 0.3$$

Örnek: (devam)

- Ortalama ve varyans:

$$E[X] = 0.2\int_0^5 x\,dx = \frac{5}{2} \qquad \sigma_X^2 = 0.2\int_0^5 (x-2.5)^2\,dx = \frac{25}{12}$$

Tek taraflı Chebychev eşitsizliğinden üst sınır şöyle çıkar

$$\Pr[X - 2.5 \geq 1] \leq \frac{\sigma_X^2}{\sigma_X^2 + 1} = \frac{25}{25+12} = 0.6757$$

Markov eşitsizliğinden bulunan sınır ise şöyledir

$$\Pr[X \geq 3.5] \leq \frac{E[X]}{3.5} = \frac{2.5}{3.5} = 0.7143$$

Tek taraflı Chebychev sınırı Markov sınırından daha az toleranslıdır yani, 0.30 gerçek olasılığına daha yakındır.

Yakınsama Modları

X_1, X_2, ... , X_n gibi rastgele değişkenler dizisi verilmiş olsun, her bir örnek uzayı çıktısı s rastgele değişkenlerin *gerçeklenmesi* adı verilen X_1(s), X_2(s), ... , X_n(s) dizisini oluşturur.

Rastgele değişkenler dizisi bir rastgele değişkene yakınsar mı?

1 olasılıkla yakınsama (hemen hemen her yerde yakınsama):

$$\Pr\left[\lim_{n\to\infty} X_n \to X\right] = 1, \quad \text{tüm } s \in S \text{ için}$$

1 olasılıkla yakınsama r.d. lerin güçlü yakınsaması olarak da adlandırılır.

Ortalama karede yakınsama:

$$\lim_{n\to\infty} E\left[\left|X_n - X\right|^2\right] = 0, \quad \text{tüm } s \in S \text{ için}$$

<u>Olasılık ile Yakınsama</u> (stokastik yakınsama):

$$\lim_{n\to\infty} \Pr\left[\left| X_n - X \right| > \varepsilon \right] = 0, \quad \text{tüm } s \in S \text{ için}$$

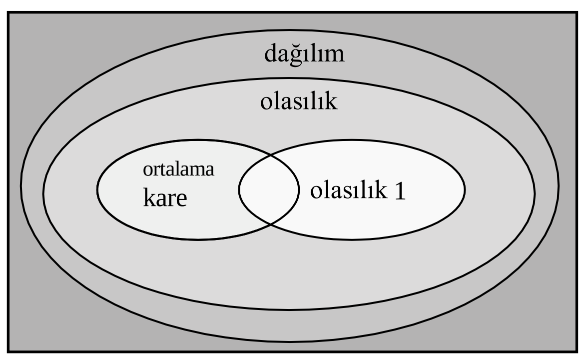
MS yakınsama olasılıkta yakınsama anlamına gelir, tersi doğru değildir.

Olasılık ile yakınsama r.d. lerin zayıf yakınsaması olarak da adlandırılır.

<u>Dağılımda yakınsama</u>:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

$F_X(x)$ in sürekli olduğu tüm x ler için. Dağılımda yakınsama S örnek uzayına bir göndermede bulunmaz.



Büyük Sayılar Kuralları

Büyük sayılar kuralları M_n örnek ortalamasının dağılımın ortalaması m_X e yakınsaması ile ilgilidir.

X_i, $i = 1, 2, \ldots, n$ gibi m_X dağılım ortalaması ve σ_X varyansına sahip IID rastgele değişken dizisi verildiğinde, örnek ortalaması aşağıdaki gibi verilir.

$$M_n = \frac{1}{n}\left[X_1 + X_2 + \cdots + X_n\right]$$

Zayıf Büyük Sayılar Kuralı

M_n örnek ortalaması, olasılıkta örnek ortalaması $m_X = E[X]$ e yakınsar, yani;

$$\lim_{n\to\infty} \Pr\left[|M_n - m_X| \geq \varepsilon \right] = 0$$

İspat: Chebychev eşitsizliğini M_n ye uygularsak

$$\Pr\left[|M_n - m_X| \geq \varepsilon \right] \leq \frac{\operatorname{var}(M_n)}{\varepsilon^2} = \frac{\sigma_X^2}{n\varepsilon^2}$$

Burada $\operatorname{var}(M_n) = \sigma_X^2 / n$ bağıntısı kullanılmıştır. $n \to \infty$ limiti eşitliğin her iki tarafına da uygulanırsa ispat tamamlanmış olur.

Bu sonucun yorumu; zayıf kural temelde belirlenmiş, ne kadar küçük olursa olsun, sıfır olmayan bir marjin için yeterince büyük sayıda örnek alındığında çok yüksek bir olasılıkla gözlemlerin ortalaması, bu marjin içinde, beklenen değere yakın olacaktır.

Güçlü Büyük Sayılar Kuralı

M_n, $m_X = E[X]$ e 1 olasılıkla yakınsar, yani;

$$\Pr\left[\lim_{n\to\infty} M_n = m_X\right] = 1$$

1 olasılıkla yakınsama olasılıkta yakınsamadan çok daha güçlü, dolayısıyla çok daha istenen bir durumdur.

Zayıf kural der ki; belirlenmiş büyük bir *n* sayısı için ortalama, $\overline{X}_n$ kuvvetle muhtemel gerçek ortalama μ ye yakın olacaktır. Yani, çok sık olmamakla beraber $|X_n - \mu| > \varepsilon$ durumu gerçekleşebilir.

Güçlü kural nerdeyse kesinlikle bunun olmyacağını gösterir. Özel olarak herhangi bir $\varepsilon > 0$ için 1 olasılıkla yeterince büyük her *n* için $|X_n - \mu| < \varepsilon$ eşitsizliği sağlanacağı sonucu çıkarılabilir.

Merkezi Limit Teoremi

IID X_i ler için $Y_n = X_1 + X_2 + \cdots X_n$ olsun: $E\left[X_i\right] = m_X$, $\operatorname{var}\left[X_i\right] = \sigma_X^2$

Buradan $\lim\limits_{n\to\infty} F_{Y_n}(y) = \Phi\left(\dfrac{y - nm_X}{\sigma_X\sqrt{n}}\right)$ (dağılımda yakınsama)

(Φ sıfır ortalamalı ve 1 varyanslı normalize edilmiş Gausyen kdf dir)

Hatırlatma:

$$z = \frac{y - nm_X}{\sigma_X\sqrt{n}} \quad \text{ve}$$

$$\Phi(z) = 1 - Q(z) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right)$$

Örnek:

Bir dükkanda ziyaretçiler tarafından günde harcanan TL, ortalaması m = 80 TL ve standart sapması σ = 20 TL IID rastgele değişkendir.

(a) Pr[100 ziyaretçinin harcadığı para > 8400 TL]= ?

X_k, k. ziyaretçi tarafından harcanan miktar. $Y_{100} = X_1 + X_2 + \cdots + X_{100}$

$$m_{Y_{100}} = nm_X = 100 \cdot 80TL = 8000TL, \qquad \sigma_{Y_{100}} = \sqrt{n}\sigma_X = 10 \cdot 20TL = 200TL$$

$$\Pr\left[Y_{100} > 8400TL\right] = 1 - \Pr\left[Y_{100} \leq 8400TL\right] = 1 - F_{Y_{100}}(8400)$$
$$= Q\left(\frac{8400-8000}{200}\right) = Q(2) = 2.28 \times 10^{-2}$$

burada normalize edilmiş değişken $z = \dfrac{y-8000}{200}$ kullanılmıştır.

Q Fonksiyonu Tablosu

x	$Q(x)$	x	$Q(x)$	x	$Q(x)$	x	$Q(x)$
0.00	5.0000E-01	1.00	1.5866E-01	2.00	2.2750E-02	5.00	2.8665E-07
0.02	4.9202E-01	1.02	1.5386E-01	2.05	2.0182E-02	5.10	1.6983E-07
0.04	4.8405E-01	1.04	1.4917E-01	2.10	1.7864E-02	5.20	9.9644E-08
0.06	4.7608E-01	1.06	1.4457E-01	2.15	1.5778E-02	5.30	5.7901E-08
0.08	4.6812E-01	1.08	1.4007E-01	2.20	1.3903E-02	5.40	3.3320E-08
0.10	4.6017E-01	1.10	1.3567E-01	2.25	1.2224E-02	5.50	1.8990E-08
0.12	4.5224E-01	1.12	1.3136E-01	2.30	1.0724E-02	5.60	1.0718E-08
0.14	4.4433E-01	1.14	1.2714E-01	2.35	9.3867E-03	5.70	5.9904E-09
0.16	4.3644E-01	1.16	1.2302E-01	2.40	8.1975E-03	5.80	3.3157E-09
0.18	4.2858E-01	1.18	1.1900E-01	2.45	7.1428E-03	5.90	1.8175E-09
0.20	4.2074E-01	1.20	1.1507E-01	2.50	6.2097E-03	6.00	9.8659E-10
0.22	4.1294E-01	1.22	1.1123E-01	2.55	5.3861E-03	6.10	5.3034E-10
0.24	4.0517E-01	1.24	1.0749E-01	2.60	4.6612E-03	6.20	2.8232E-10
0.26	3.9743E-01	1.26	1.0383E-01	2.65	4.0246E-03	6.30	1.4882E-10
0.28	3.8974E-01	1.28	1.0027E-01	2.70	3.4670E-03	6.40	7.7688E-11
0.30	3.8209E-01	1.30	9.6800E-02	2.75	2.9798E-03	6.50	4.0160E-11
0.32	3.7448E-01	1.32	9.3418E-02	2.80	2.5551E-03	6.60	2.0558E-11
0.34	3.6693E-01	1.34	9.0123E-02	2.85	2.1860E-03	6.70	1.0421E-11
0.36	3.5942E-01	1.36	8.6915E-02	2.90	1.8658E-03	6.80	5.2310E-12
0.38	3.5197E-01	1.38	8.3793E-02	2.95	1.5889E-03	6.90	2.6001E-12
0.40	3.4458E-01	1.40	8.0757E-02	3.00	1.3499E-03	7.00	1.2798E-12
0.42	3.3724E-01	1.42	7.7804E-02	3.05	1.1442E-03	7.10	6.2378E-13
0.44	3.2997E-01	1.44	7.4934E-02	3.10	9.6760E-04	7.20	3.0106E-13
0.46	3.2276E-01	1.46	7.2145E-02	3.15	8.1635E-04	7.30	1.4388E-13
0.48	3.1561E-01	1.48	6.9437E-02	3.20	6.8714E-04	7.40	6.8092E-14
0.50	3.0854E-01	1.50	6.6807E-02	3.25	5.7703E-04	7.50	3.1909E-14
0.52	3.0153E-01	1.52	6.4255E-02	3.30	4.8342E-04	7.60	1.4807E-14
0.54	2.9460E-01	1.54	6.1780E-02	3.35	4.0406E-04	7.70	6.8033E-15
0.56	2.8774E-01	1.56	5.9380E-02	3.40	3.3693E-04	7.80	3.0954E-15
0.58	2.8096E-01	1.58	5.7053E-02	3.45	2.8029E-04	7.90	1.3945E-15
0.60	2.7425E-01	1.60	5.4799E-02	3.50	2.3263E-04	8.00	6.2210E-16
0.62	2.6763E-01	1.62	5.2616E-02	3.55	1.9262E-04	8.10	2.7480E-16
0.64	2.6109E-01	1.64	5.0503E-02	3.60	1.5911E-04	8.20	1.2019E-16
0.66	2.5463E-01	1.66	4.8457E-02	3.65	1.3112E-04	8.30	5.2056E-17
0.68	2.4825E-01	1.68	4.6479E-02	3.70	1.0780E-04	8.40	2.2324E-17
0.70	2.4196E-01	1.70	4.4565E-02	3.75	8.8417E-05	8.50	9.4795E-18
0.72	2.3576E-01	1.72	4.2716E-02	3.80	7.2348E-05	8.60	3.9858E-18
0.74	2.2965E-01	1.74	4.0930E-02	3.85	5.9059E-05	8.70	1.6594E-18
0.76	2.2363E-01	1.76	3.9204E-02	3.90	4.8096E-05	8.80	6.8408E-19
0.78	2.1770E-01	1.78	3.7538E-02	3.95	3.9076E-05	8.90	2.7923E-19
0.80	2.1186E-01	1.80	3.5930E-02	4.00	3.1671E-05	9.00	1.1286E-19
0.82	2.0611E-01	1.82	3.4380E-02	4.10	2.0658E-05	9.10	4.5166E-20
0.84	2.0045E-01	1.84	3.2884E-02	4.20	1.3346E-05	9.20	1.7897E-20
0.86	1.9489E-01	1.86	3.1443E-02	4.30	8.5399E-06	9.30	7.0223E-21
0.88	1.8943E-01	1.88	3.0054E-02	4.40	5.4125E-06	9.40	2.7282E-21
0.90	1.8406E-01	1.90	2.8717E-02	4.50	3.3977E-06	9.50	1.0495E-21
0.92	1.7879E-01	1.92	2.7429E-02	4.60	2.1125E-06	9.60	3.9972E-22
0.94	1.7361E-01	1.94	2.6190E-02	4.70	1.3008E-06	9.70	1.5075E-22
0.96	1.6853E-01	1.96	2.4998E-02	4.80	7.9333E-07	9.80	5.6293E-23
0.98	1.6354E-01	1.98	2.3852E-02	4.90	4.7918E-07	9.90	2.0814E-23

Ters Q Fonksiyonu Tablosu

l	$x = Q^{-1}(10^{-l})$	l	$x = Q^{-1}(10^{-l})$	l	$x = Q^{-1}(10^{-l})$	l	$x = Q^{-1}(10^{-l})$
1	1.2816	7	5.1993	13	7.3488	19	9.0133
2	2.3263	8	5.6120	14	7.6506	20	9.2623
3	3.0902	9	5.9978	15	7.9413	21	9.5050
4	3.7190	10	6.3613	16	8.2221	22	9.7418
5	4.2649	11	6.7060	17	8.4938	23	9.9730
6	4.7534	12	7.0345	18	8.7573	24	10.1990

(b) 0.90 olasılıkla toplam 10,000 TL harcanmış olması için kaç ziyaretçinin dükkanı ziyaret etmiş olması gerekir.

$\Pr\left[Y_n > 10000TL\right] = 0.90$ isteniyor. $\quad m_{Y_n} = 80n \quad \sigma_{Y_n}^2 = 400n$

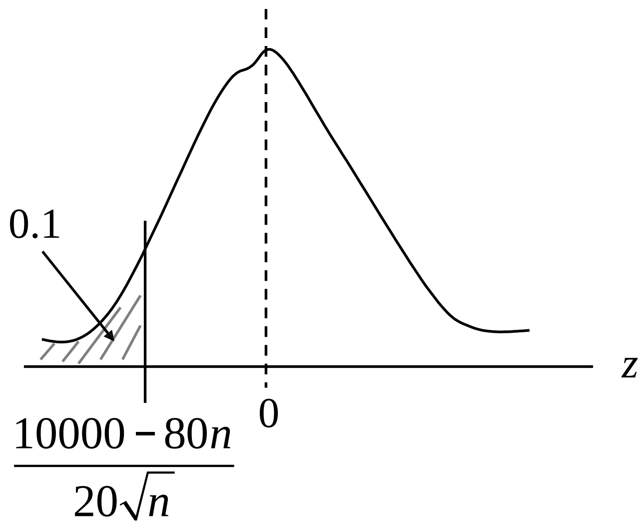
$$\Pr\left[Y_n > 10000TL\right] = \Pr\left[Z_n > \frac{10000-80n}{20\sqrt{n}}\right] = 0.9 \text{ veya } \Pr\left[Z_n \le \frac{10000-80n}{20\sqrt{n}}\right] = 0.1$$

Tablo B den:

$$\frac{10000-80n}{20\sqrt{n}} = -1.2815$$

$$80n - 25.63\sqrt{n} - 10000 = 0$$

(pozitif kök)



$$\sqrt{n} = \frac{25.63 \pm \sqrt{25.63^2 + 3200000}}{160} = 11.34 \qquad \therefore\ n = 11.34^2 \cong 129$$

Örnek Ortalaması ve Varyansı

Örnek Ortalaması:

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

X_i ler m_X ortalamalı σ_X^2 varyanslı IID dir.

M_n in ortalaması:

$$E\left[M_n\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n}\sum_{i=1}^{n} m_X = m_X$$

M_n in varyansı:

$$\mathrm{var}\left[M_n\right] = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{var}\left[X_i\right] = \frac{\sigma_X^2}{n}$$

$n \to \infty$ için sıfır olur.

Örnek varyansı:

$$\sum\nolimits_n^2 \frac{\left(X_1 - M_n\right)^2 + \cdots + \left(X_n - M_n\right)^2}{n} = \frac{1}{n}\sum_{i=1}^{n} \left(X_i - M_n\right)^2$$

Örnek Varyansın Ortalaması:

$$E\left[\Sigma_n^2\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[\left(X_i - M_n\right)^2\right]$$

$X_i - M_n = \left(X - m_X\right) - \left(M_n - m_X\right)$ koyarsak

$$E\left[\Sigma_n^2\right] = \frac{1}{n}\sum_{i=1}^{n} E\left[\left(X_i - m_X\right)^2\right] + \frac{1}{n}\sum_{i=1}^{n} E\left[\left(M_n - m_X\right)^2\right]$$

$$-\frac{2}{n} E\left[\left(M_n - m_X\right)\sum_{i=1}^{n}\left(X_i - m_X\right)\right] \text{ (not a bakın)}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left[\left(X_i - m_X\right)^2\right] - E\left[\left(M_n - m_X\right)^2\right] = \left(1 - \frac{1}{n}\right)\sigma_X^2 .$$

Not: $\frac{1}{n}\sum_{i=1}^{n}\left(X_i - m_X\right) = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - m_X = M_n - m_X$