Olasılık Teorisi ve Stokastik Süreçler:

Bölüm 1: Elementer Olasılık Teorisi

§1.1. Neden bir teoriye ihtiyaç duyuyoruz?

Olasılığın ilk bakışta karmaşık gözüken problemleri, doğru şekilde sunulduğu zaman teorisinin anlaşılmasını kolaylaştırabilir. Bunu bir örnekle açıklayalım:

"Bir deste oyun kağıdından sadece as ve papazları alalım (toplamda 8 kağıt). Bunlardan iki tanesini, kapalı şekilde, bir arkadaşınıza verelim. Arkadaşınız kağıtlara bakıp elinde bir as olduğunu söylüyorsa her ikisinin de as olma olasılığı nedir? Eğer elindeki kağıdın maça as olduğunu söyleseydi bu olasılık ne olurdu, değişir miydi?

Şimdi bu soruları sistematik şekilde çözelim:

Soru 1) $\binom{8}{2} = 8!/(2!\times 6!) = 28$ olduğu için arkadaşınıza birbirinden farklı 28 değişik el verebilirsiniz. Bunlardan sadece $\binom{4}{2} = 4!/(2!\times 2!) = 6$ tanesinden hiç as bulunmaz. O halde $28-6=22$ elde en az bir as bulunur. Ve yine bunlardan sadece $\binom{4}{2}=6$ tanesi sadece aslardan oluşur. O halde, arkadaşınızın elindeki her iki kağıdın da as olma ihtimali $6/22 = 3/11$'dir.

Soru 2) Arkadaşınızın elindeki kağıtlardan bir maça as ise ikinci kağıt için $8-1=7$ olasılık kalmıştır ve bunlardan sadece 3 tanesi yine bir astır. O halde, cevap $3/7$'dir.

§1.2. Örneklem Uzayı:

Tanım: Rasgele bir deneyin tüm olası sonuçlarının kümesine deneyin örneklem uzayı denir ve genelde S ile gösterilir.

Deneyin doğasına bağlı olarak S örneklem uzayı sonlu, sayılabilir sonsuz veya sayılamaz (sonsuz) olabilir. Örneğin bir yazı-tura deneyi için $S=\{Y,T\}$ ve dolayısıyla

$|S| = 2$ olur. Diğer yandan yukarıdaki kart oyununda $|S| = 28$'dir.

__Başka Örnekler:__ 1) Zar atma oyunu: $S = \{1,2,3,4,5,6\}$, $|S| = 6$.

2) Tura gelene kadar parayı atma oyunu: $S = \{1,2,3,\dots\}$, $|S| = \infty$, sayılabilir sonsuz.

3) "0" ile "1" arasında rasgele sayı seçme deneyi: $S = [0,1]$, sayılamaz sonsuz bir küme.

4) Yarıçapı bir birim olan dairesel bir hedef tahtasına ok atma deneyi:
$S = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=1\}$, sayılamaz sonsuz küme.

§1.3. __Olaylar:__ Rasgele bir deneyin S örneklem uzayının alt kümelerine olay denir. S alt kümesine kesin olay, ϕ boş alt kümesine ise boş olay denir.

§1.4. __Verilen olaylardan Yeni olaylar elde etme:__

Temelde üç çeşit yeni olay elde etme yöntemi vardır:

i) __Tümleyen:__ Verilen bir $E \subseteq S$ olayının $E^c = S \setminus E$ tümleyeni de bir olaydır.

ii) __Arakesit:__ Verilen bir $\{E_\alpha\}$ olay ailesinin $\bigcap_\alpha E_\alpha$ arakesiti de bir olaydır.

iii) __Birleşim:__ Benzer şekilde $\{E_\alpha\}$ olay ailesinin $\bigcup_\alpha E_\alpha$ birleşimi de bir olaydır.

Genellikle bir olay, cevabı istenen bir soruya "Evet" yanıtı anlamına gelen deney sonuçlarından oluşur. Deneyin sonucu $A \subseteq S$ olayının içinde demek "A olayı doğrudur" anlamına gelir. Örneğin oyun kartları deneyinde doğru cevap olarak görülen olay her iki kartında as olmasıdır.

Alıştırma 1. E, F ve G rasgele bir deneyin üç olayı olsun. Aşağıdaki olayları formülüze edelim:
i) Sadece E doğrudur, ii) E ve F doğrudur ama G doğru değildir, iii) Bu olayların en az ikisi doğrudur.

Çözüm: i) Sadece $E = E \cap F^c \cap G^c$.
ii) $E \cap F \cap G^c$
iii) $(E \cap F) \cup (E \cap G) \cup (F \cap G)$.

§1.5. Olasılık Atama İşlemi:

S bir deneyin örneklem uzayı ise aşağıdaki koşulları sağlayan bir $P: \mathcal{P}(S) \longrightarrow [0,1]$ fonksiyonuna örneklem uzayı üzerinde bir olasılık dağılımı denir.

1) $P(E^c) = 1 - P(E)$, $\forall E \in S$
2) $P(S) = 1$
3) $E \subseteq F$, $E, F \in S$ ise $P(E) \leq P(F)$.
4) Eğer, $E, F \in S$ için $E \cap F = \emptyset$ ise $P(E \cup F) = P(E) + P(F)$.
5) $E, F \in S$ ise $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.
6) $\{E_i\}$, $E_i \in S$ ve her $i \neq j$ için $E_i \cap E_j = \emptyset$ ise

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i)$$

Alıştırma 3. (1), (3) ve (5) koşullarının (2) ve (4)'den elde edildiğini gösteriniz.

Uyarı: (6) diğerlerinin bir sonucu değildir.

Alıştırma 4. E ve F iki olay olmak üzere her ikisinin de yanlış olma olasılığı 0.6 ve her ikisinin de doğru olma olasılığı 0.2 ise bu iki olaydan sadece birinin doğru olma olasılığını hesaplayınız.

Çözüm: $0.6 = P((E \cup F)^c) = 1 - P(E \cup F)$.

$$\Rightarrow 0.6 = 1 - (P(E) + P(F) - P(E \cap F))$$
$$= 1 - P(E) - P(F) + 0.2$$

$\Rightarrow P(E) + P(F) = 0.6$. O halde, aranan olasılık
$(P(E) - P(E \cap F)) + (P(F) - P(E \cap F)) = 0.6 - 2 \times 0.2 = 0.2$ olur.

S örneklem uzayının sonlu veya sayılabilir sonsuz olduğu elementer olasılık teorisinde her $s \in S$ elemanının $p(s)$ olasılık değeri verilir. Dolayısıyla, her $A \in S$ için

$$\mathbb{P}(A) = \sum_{s \in A} p(s) \quad \text{olur (koşul (6))}.$$

Burada, $A = \bigcup_{s \in A} \{s\}$ ayrık birleşimdir ve $\mathbb{P}(S) = 1$ olduğu için yukarıdaki toplam yakınsaktır.

Alıştırma 5: $A_n \subseteq S, \forall n$, olmak üzere

$$\mathbb{P}(A) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n) \quad \text{olduğunu gösteriniz.}$$

§1.7. Bağıl (Koşullu) Olasılık:

$A, B \subseteq S$ olmak üzere $\mathbb{P}(B|A)$ ile "B'nin A'nın doğru olma durumunda doğru olma" koşullu olasılığını göstereceğiz.

Bunu bir örnek üzerinde görelim: Yazı tura atılan bir oyunda, bir deste kağıttan, tura geldiğinde bir kağıt, yazı geldiğinde iki kağıt seçtiğimizi düşünelim. A ile yazı turada tura gelme olasılığını, B ile de çekilen oyun kağıtların-dan en az birinin as olması olasılığını gösterelim. Bu durumda, $P(B|A) = 4/52$ olurken $\mathbb{P}(A \cap B) = \frac{1}{2} \times \frac{1}{3} = 1/6$ olur.

Dolayısıyla, bu örnek için $\mathbb{P}(A \cap B) = 1/2 \times 1/3 = \mathbb{P}(A).\mathbb{P}(B|A)$ olur.

Tanım: $\mathbb{P}(B|A) = \mathbb{P}(B \cap A) / \mathbb{P}(A)$ $\quad (\mathbb{P}(A) \neq 0$ olmamak üzere.)

Alıştırma 6. $\mathbb{P}(B|A) > \mathbb{P}(B)$ olsun. Bu durumda $\mathbb{P}(A|B)$ ve $\mathbb{P}(A)$ hakkında ne söyleyebilirsiniz.

Çözüm: $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(A)\, \mathbb{P}(B|A)}{\mathbb{P}(B)}$ olduğundan

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)}{\mathbb{P}(B)} \mathbb{P}(A) \geq 1.\mathbb{P}(A) = \mathbb{P}(A) \text{ elde ederiz.}$$

Alıştırma 7: $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\cdots\mathbb{P}(A_n|A_1\cap\cdots\cap A_{n-1})$ olduğunu tümevarım metodu ile kanıtlayınız.

Alıştırma 8: 52'lik bir oyun kağıdı destesi rasgele şekilde 13'er kağıtlık dört ele ayrılıyor. Her bir elde en az bir as olma olasılığını hesaplayınız.

İpucu: $k=1,2,3,4$ için $A_k = k$. elin sadece bir ası olması olayı olsun. Şimdi Alıştırma 7'yi kullanın.

Toplam Olasılık: $A_1,\ldots,A_n$ bir örneklem uzayının birbirinden ayrık olayları olsun: $A_i \cap A_j = \emptyset$, $\forall i \neq j$. Ayrıca $S = A_1 \cup \cdots \cup A_n$ verilsin. Bu durumda, her $B \subseteq S$ için $B = \bigcup_{i=1}^{n}(A_i \cap B)$ ayrık bir birleşimdir. Dolayısıyla

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}\Big(\bigcup_{i=1}^{n}(A_i\cap B)\Big) \\ &= \mathbb{P}(B\cap A_1) + \cdots + \mathbb{P}(B\cap A_n) \\ &= \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \cdots + \mathbb{P}(B|A_n)\mathbb{P}(A_n) \end{aligned}$$

olur.

Bayes Formülü: Yukarıdaki formülden, her $i = 1,\ldots,n$, için

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{i=1}^{n}\mathbb{P}(B|A_i)\mathbb{P}(A_i)}$$

elde ederiz.

Buna Bayes Formülü denir.

Alıştırma 10. Ateş ve üşüme hissi şikayeti olan bir hasta doktora gider. Doktor gördüğü belirtiler altında hastada tüberküloz (TB) olma olasılığını merak eder. A ile hastanın TB olma olayını, B ile de hastada ateş ve üşüme hissi olma olayını gösterelim. Diyelim ki toplumdaki TB hastası oranı 0.01 ve yine toplumda ateş ve üşüme belirtileri gösterenlerin oranı da %3 olsun. $\mathbb{P}(B|A) = 0.5$ olduğunu kabul edelim. Başka bir deyişle, TB hastalarının yarısı ateş ve üşüme şikayeti göstersin. Bu durumda, doktorun sorusunun cevabı nedir?

$\mathbb{P}(A\cap B) = \mathbb{P}(B|A)\mathbb{P}(A) = 0.5\times 10^{-4} = 5\times 10^{-5}$ olur.

Buradan, $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \dfrac{5\times 10^{-5}}{3\times 10^{-2}} = \dfrac{5}{3000} = \dfrac{1}{600}$

Alıştırma 11. En başta vermiş olduğumuz oyun kâğıdı problemini koşullu olasılığı kullanarak çözünüz.

§1.8. Bağımsız Olaylar: A ve B gibi iki olay için $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$ veya eşdeğer şekilde $\mathbb{P}(A|B) = \mathbb{P}(A)$ veya $\mathbb{P}(B|A) = \mathbb{P}(B)$ oluyorsa bu iki olaya bağımsız olaylar denir.

Benzer şekilde bir $A_1, A_2, \ldots, A_n, \ldots$ olaylar dizisinin her sonlu $A_{i_1}, \ldots, A_{i_n}$ elemanı için
$\mathbb{P}(A_{i_1}\cap \cdots \cap A_{i_n}) = \mathbb{P}(A_{i_1})\cdots \mathbb{P}(A_{i_n})$ oluyorsa bu diziye bağımsız olaylar dizisi denir.

Alıştırma 12. Ard arda zar atma olayları bağımsızdır. Hileli bir paranın tura gelme olasılığı p olsun. On kere atılan bu paranın en az iki kere tura gelme olasılığı nedir?

Çözüm: A = 10 kere atılan paranın her seferinde yazı gelme olasılığı, ve
B = 10 kere atılan paranın sadece bir defa tura gelme olasılığı, olsun.

Bu durumda sorumuzun cevabı $1-\mathbb{P}(A)-\mathbb{P}(B)$ olur. Diğer yandan, $q = 1-p$ olmak üzere $\mathbb{P}(A) = q^{10}$ ve $\mathbb{P}(B) = 10pq^9$ olur. O halde cevap

$$1 - q^{10} - 10pq^9 \text{ olur.}$$

Alıştırma 13. İki zar defalarca atılıyor ve her seferinde zarların üzerindeki sayılar toplanıyor. Bu sayının 7 gelmeden önce 8 gelme olasılığı kaçtır?

Çözüm: $S = \{(a,b) \mid 1 \le a, b \le 6\}$ olsun. Bu durumda $|S| = 36$ olur. A ve B ile 7 ve 8 gelme olaylarını gösterelim. Bu durumda $\mathbb{P}(A) = 6/36 = 1/6$ ve $\mathbb{P}(B) = 5/36$ olur. C ise 7 ve 8'in gelmeme olasılığı olsun: $\mathbb{P}(C) = 1 - 11/36 = \dfrac{25}{36}$.

Şimdi de D_n ile n'inci denemeden önce 7 veya 8'in hiç gelmeme ve n'inci denemede 8 gelme olayı olsun.

O halde, $\mathbb{P}(D_n) = (\mathbb{P}(C))^{n-1}\mathbb{P}(B) = \left(\frac{25}{36}\right)^{n-1} \times \frac{5}{36} = \frac{1}{5}\left(\frac{25}{36}\right)^n.$

Bu durumda sorumuzun cevabı

$$\sum_{n=1}^{\infty} \mathbb{P}(D_n) = \frac{1}{5}\sum_{n=1}^{\infty}\left(\frac{25}{36}\right)^n = \frac{1}{5}\cdot\frac{25}{36}\,\frac{1}{1-25/36} = \frac{5}{36}\cdot\frac{36}{11} = \frac{5}{11}.$$

§1.9. Rasgele Değişkenler: Verilen bir S örneklem uzayı üzerinde tanımlı $X: S \to \mathbb{R}$, $s \mapsto X(s)$, fonksiyonuna rasgele değişken denir. Rasgele değişkeni deneyin sonuçlarının bir ölçüsü olarak görebiliriz.

Verilen bir $x \in \mathbb{R}$ değeri için $\{X = x\} \doteq \{s \in S \mid X(s) = x\} = X^{-1}(x)$ olarak tanımlanır. $\{X=x\} \subseteq S$ olayının olasılığı $\mathbb{P}(X=x)$ ile ifade edilir. Elementer olasılık teorisinde S sonlu veya sayılabilir sonsuz olduğu için $X: S \to \mathbb{R}$ fonksiyonun değerlerinin kümesi de sonlu veya sayılabilir sonsuzdur. $\mathrm{Ran}(X) = \{x_1, x_2, \ldots, x_k, \ldots\}$ ve $p_k = \mathbb{P}(X = x_k)$ ise $(p_1, x_1), (p_2, x_2), \ldots, (p_k, x_k), \ldots$ dizisine rasgele değişkenin olasılık kütle fonksiyonu (pmf) denir.

Alıştırma 14. İki zar atalım ve Y gelen sayıların maksimumu olsun. Y rasgele değişkeninin olasılık kütle fonksiyonunu bulalım.

$S = \{(m,n) \mid 1 \le m, n \le 6\}$, $\mathrm{Ran}(Y) = \{1, 2, \ldots, 6\}$.

$\{Y=1\} = \{(1,1)\}$, $\{Y=2\} = \{(1,2),(2,1),(2,2)\} \ldots$, $\{Y=6\} = \{(1,6), \ldots, (6,1)\}$.

Bu durumda $p_1 = \mathbb{P}(\{Y=1\}) = 1/36$, $p_2 = \mathbb{P}(\{Y=2\}) = 3/36 = 1/12$, $p_3 = \mathbb{P}(\{Y=3\}) = 5/36$, $p_4 = \mathbb{P}(\{Y=4\}) = 7/36$, $p_5 = \mathbb{P}(\{Y=5\}) = 9/36 = 1/4$ ve $p_6 = \mathbb{P}(\{Y=6\}) = 11/36$ olur.

Soru: Bir X olasılık kütle fonksiyonu verilirse bu fonksiyonu olasılık kütle dağılımı olarak gören bir örneklem uzayı var mıdır? Bu sorunun cevabı evettir ve çok sayıda örneklem uzayı bulunabilir. Fakat bunlardan birisi en küçüktür! Gerçekten de eğer $\{(x_i, p_i)\}$ verilirse, $X: S \to S$, $s \mapsto s$, ve $\mathbb{P}(\{x_i\}) = p_i$, $i \in I$, olarak tanımlanan örneklem uzayı aranan özelliğe sahiptir.

Örnek: İki tane hilesiz zar atma deneyi için aşağıdaki rasgele değişkeni tanımlayalım:

$$X = \begin{cases} 0 & \text{eğer ikisi farklı ise,} \\ 1 & \text{eğer ikisi de aynı ise.} \end{cases}$$

Bu değişken için doğal örneklem uzayı 36 elemanlı $S' = \{(m,n) \mid 1 \leq m,n \leq 6\}$ olur. Örneğin, $X(1,3) = 0$ ve $X(4,4) = 1$'dir. Ayrıca, $P(X=0) = 6/36 = 1/6$ ve $P(X=1) = 30/36 = 5/6$ dır.

Diğer yandan aynı değişken için örneklem uzayını $S' = \{0,1\}$ ve $P(X=0) = 1/6$, $P(X=1) = 5/6$ olarak da seçebiliriz.

Önemli Ayrık Rasgele Değişken Örnekleri:

Bernoulli: $Ran(X) = \{0,1\}$, $p = P(X=1)$, $1-p = P(X=0)$.

Binom: $Ran(X) = \{0,1,2,\ldots,n\}$, $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

Örneğin, X'i n defa atılan hileli bir paranın k defa tura gelme olasılığı olarak görebiliriz.

Poisson: $Ran(X) = \{0,1,2,\ldots\}$, $P(X=k) = e^{-\lambda} \frac{\lambda^k}{k}$.

Binom ve Poisson rasgele değişkenleri için şunu biliyoruz:

Lemma: $\lambda = np$ alalım ve λ bir sabit olsun. Bu durumda,

$$\lim_{\substack{n \to \infty \\ (p \to 0)}} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \lambda^k / k!.$$

Kanıt:
$$\binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \frac{1}{\left(1-\frac{\lambda}{n}\right)^k} \left(1 - \lambda/n\right)^n$$
$$= \binom{n}{k} \left(\frac{\lambda}{n-\lambda}\right)^k \left(1 - \lambda/n\right)^n$$
$$= \frac{n!}{k!\,(n-k)!} \frac{\lambda^k}{(n-\lambda)^k} \left(1 - \lambda/n\right)^n$$

$$= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{(n-\lambda)^k} \left(1-\lambda/n\right)^n.$$

O halde, $\lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1-\lambda/n\right)^{n-k} = \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda} = \frac{\lambda^k}{k!} e^{-\lambda}$.

elde edilir. ∎

Alıştırma 15. Hileli bir paranın tura gelme olasılığı p olsun $(0<p\leq 1)$. Tura gelene kadar parayı atalım. N sayısı ilk tura geldiği atışın numarası olsun. N rasgele değişkeninin pmf'sini bulalım. (Buna geometrik dağılım adını vereceğiz.)

§1.10. Birleşik Dağılımlar:

$X_1, \ldots, X_n$ ayrık rasgele dağılımlar ve $R_i = Ran(X_i)$, $i=1,\ldots,n$, olsun. $X = (X_1, \ldots, X_n)$ fonksiyonunun görüntüsü $R_1\times\cdots\times R_n$ çarpım kümesinin içinde kalacaktır. Bu değişkenlerin birleşik olasılık kütle fonksiyonu (p.m.f.) $P(X_1=x_1,\ldots,X_n=x_n)$, $(x_1,\ldots,x_n)\in R_1\times\cdots\times R_n$, ile tanımlanır.

Alıştırma 16. X ve Y peş peşe atılan bir çift zarın en küçük ve en büyük değerlerini gösteren rasgele değişkenler olsun.

Çözüm: $x=y$ ise $P(X=x, Y=y) = 2/36 = 1/18$ ve $x\neq y$ ise

$P(X=x, Y=y) = 1/36$ olur.

Birleşik dağılımdan her bir değişkenin dağılımını elde edebiliriz. Örneğin,
$$P(X=x_1) = \sum_{x_2,\ldots,x_n} P(X_1=x_1, X_2=x_2, \ldots, X_n=x_n).$$

Buna X_1 değişkenin marjinal pmf'sini bulma işlemi denir.

Örnek: Alıştırma 16'yı tekrar ele alalım. Bu durumda

$$11/36 = P(X=1) = \sum_{y=1}^{6} P(X=1, Y=y) = P(X=1, Y=1) + \cdots + P(X=1, Y=6) = 1/36 + \frac{5\times 2}{36}.$$

$X_1, X_2, \ldots, X_n$ rasgele değişkenleri verilsin. Eğer her $(x_1, \ldots, x_n) \in R_1 \times \cdots \times R_n$ için $\{X_1 = x_1\}, \ldots, \{X_n = x_n\}$ olayları bağımsız ise bu değişkenlere de bağımsız denir. Başka bir değişle

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n), \ \forall (x_1, \ldots, x_n) \in R_1 \times \cdots \times R_n$$

oluyorsa bu değişkenlere bağımsızdır denir.

Alıştırma 17: Birisi hilesiz iki paramız olsun. Hileli olan $2/3$ olasılıkla tura geliyor. Her iki parayı ikişer defa atalım. X hilesiz paranın tura sayısını, Y ise hileli paranın tura sayısını göstersin. $P(X>Y)$ olasılığını hesaplayınız.

Çözüm: $$P(X>Y) = \sum_{2 \geq i > j \geq 0} P(X=i, Y=j)$$

(X ve Y bağımsız!)
$$\begin{aligned}
&= P(X=2, Y=1) + P(X=2, Y=0) + P(X=1, Y=0) \\
&= P(X=2)P(Y=1) + P(X=2)P(Y=0) + P(X=1)P(Y=0) \\
&= \frac{1}{2}\cdot\frac{1}{2}\left(\frac{2}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{2}{3}\right) + \frac{1}{2}\cdot\frac{1}{2}\left(\frac{1}{3}\cdot\frac{1}{3}\right) + \left(\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\right)\cdot\left(\frac{1}{3}\cdot\frac{1}{3}\right) \\
&= \frac{1}{4}\cdot\frac{4}{9} + \frac{1}{4}\cdot\frac{1}{9} + \frac{2}{4}\cdot\frac{1}{9} \\
&= 7/36
\end{aligned}$$

Alıştırma 18: Tura gelme olasılığı p olan iki hileli paramız var. İlk para tura gelene kadar atılıyor. Bu atış sayısı N rasgele değişkeni olsun. İkinci para ise N defa atılıyor. İkinci paranın tura gelme sayısını da X ile gösterelim. $P(X=k)$ olasılığını hesaplayınız.

Çözüm: $P(N=n) = p(1-p)^{n-1}$ olduğu açıktır. İkinci paranın n defa atıldığında tam olarak k kere tura gelme olasılığı ise $\binom{n}{k} p^k (1-p)^{n-k}$'dır.

Hence,
$$\begin{aligned}
P(X=k) &= \sum_{n=k}^{\infty} P(X=k, N=n) \\
&= \sum_{n=k}^{\infty} P(X=k / N=n) P(N=n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \cdot p(1-p)^{n-1} \\
&= \sum_{n=k}^{\infty} \binom{n}{k} (1-p)^{2n-k-1} p^{k+1}
\end{aligned}$$

elde edilir.

§1.11 Beklenen Değer (Beklenti):

$X(S)$ görüntü kümesi sonlu olan bir rasgele değişkenin beklenen değeri aşağıdaki gibi tanımlanır:

$$\mathbb{E}X = \sum_n p_n x_n, \quad p_n = \mathbb{P}(X = x_n).$$

Örneğin bir zarın beklenen değeri $\sum_{n=1}^{\infty} \frac{n}{6} = \frac{6\times 7}{2\times 6} = 7/2$'dir.

Eğer $X(S)$ kümesi sayılabilir sonsuz ise daha dikkatli olunması gerekir. İlk önce X rasgele değişkenini pozitif ve negatif kısımlarına ayırmamız gerekir:

$$X^+ = \max\{X, 0\} \text{ ve } X^- = -\min\{X, 0\}.$$

Bu durumda, hem X^+ hem de X^- pozitif değerlidir ve $X = X^+ - X^-$ olur. $\mathbb{E}X^+$ ve $\mathbb{E}X^-$ aşağıdaki toplamlar yardımıyla tanımlanır:

$$\mathbb{E}X^+ = \sum_{n;\, x_n \geq 0} x_n p_n \quad \text{ve} \quad \mathbb{E}X^- = \sum_{n,\, x_n < 0} x_n |p_n|.$$

Bu toplamlar $[0, +\infty]$ aralığından değer alır. İkisi birden $+\infty$ olmadığı durumda $\mathbb{E}X = \mathbb{E}X^+ - \mathbb{E}X^-$ olarak tanımlanır.

$|X| = X^+ + X^-$ olur ve $\mathbb{E}X$ iyi tanımlı ve sonludur ancak ve ancak $\mathbb{E}|X|$ iyi tanımlı ve sonludur.

<u>Uyarı</u>: n defa tekrarlanan bir deneyin ortalama sonucu $Or(X;n)$ ile gösterilsin. Bu durumda aksiyomlardan

$$\mathbb{E}X = \lim_{n\to\infty} Or(X;n)$$ olduğu gösterilebilir.

<u>Alıştırma 19</u>: Aşağıdaki durumlar için $\mathbb{E}X$'i hesaplayınız:

a) X atılan iki zarın en büyük değeridir.

b) X tura gelme olasılığı p olan hileli bir paranın tura gelene kadar atılma sayısıdır.

Çözüm: a) $\mathbb{E}X = \sum_{n=1}^{\infty} x_n p_n = 1 \times \frac{1}{36} + 2 \times \frac{3}{36} + 3 \times \frac{5}{36} + 4 \times \frac{7}{36} + 5 \times \frac{9}{36} + 6 \times \frac{11}{36}$

$$\Rightarrow \mathbb{E}X = \frac{161}{36}.$$

b) Bir önceki alıştırmadan $P(N=n) = p(1-p)^{n-1}$ olduğunu biliyoruz. O halde cevap

$$\mathbb{E}X = \sum_{n=1}^{\infty} p_n x_n = \sum_{n=1}^{\infty} n p_n = \sum_{n=1}^{\infty} n p (1-p)^{n-1} \text{ olur.}$$

Toplamı hesaplamak için $f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, $|x|<1$, fonksiyonunu ele alalım.

$f'(x) = \sum_{n=1}^{\infty} n x^{n-1}$ olduğu için

$$\mathbb{E}X = \sum_{n=1}^{\infty} np(1-p)^{n-1} = p \sum_{n=1}^{\infty} n(1-p)^{n-1} = p f'(1-p) = \left.\frac{p}{(1-x)^2}\right|_{x=1-p}$$

$\Rightarrow \mathbb{E}X = p\left(1/(1-(1-p))^2\right) = p/(1/p^2) = p^3$ elde edilir.

§1.13. Rasgele Değişkenlerin Fonksiyonları:

$X: S \to \mathbb{R}$ ayrık rasgele bir değişken ve $g: \mathbb{R} \to \mathbb{R}$ bir fonksiyon ise $Y = g \circ X : S \to \mathbb{R}$ bir başka ayrık rasgele değişken olur. Bu durumda

$$\mathrm{Ran}(Y) = g(\mathrm{Ran}(X)) = \{g(x_k) \mid x_k \in \mathrm{Ran}(X)\} \text{ ve}$$

$$P(Y=y) = P(\{s \in X \mid g(X(s)) = y\}) = \sum_{s : g(X(s)) = y} p(s) = \sum_{k : g(x_k) = y} P(X = x_k) \text{ olur.}$$

Ayrıca, $Y = g(X)$ için

$$\mathbb{E}Y^+ = \sum_{y_j \geq 0} y_j P(Y = y_j) = \sum_{y_j \geq 0} y_k \sum_{k : g(x_k) = y_j} P(X = x_k) = \sum_{y_j \geq 0} \sum_{k : g(x_k) = y_j} g(x_k) P(X = x_k) = \sum_{y_j \geq 0} \sum_{k} \mathbb{1}_{\{g(x_k) = y_j\}} g(x_k) P(X = x_k)$$

olur.

Burada $\mathbb{1}_A$, A kümesi üzerinde 1, A'nın dışında ise 0 değerini alan gösterge fonksiyonudur.

Toplamdaki tüm terimler pozitif olduğu için terimlerin yerini istediğimiz gibi değiştirebiliriz.

$$\mathbb{E}Y^+ = \sum_{k:g(x_k)\geq 0} g(x_k)\, P(X_k = x_k) \text{ olduğu gibi}$$

$$\mathbb{E}Y^- = \sum_{k:g(x_k)<0} -g(x_k)\, P(X_k = x_k) \text{ olur.}$$

$\mathbb{E}Y$ tanımlandığı sürece $\mathbb{E}Y = \sum_k g(x_k)\, P(X = x_k)$ veya kısaca $\mathbb{E}Y = \sum_x g(x)\, P(X=x)$ yazacağız.

<u>Alıştırma 20.</u> $a_{ij} \geq 0$, $\forall i,j$ olduğunu kabul edelim.

O halde, $\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{ij}\right) = \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} a_{ij}\right)$ olur. Burada her iki toplamın da $+\infty$ olmasına izin veriyoruz.

<u>Çözüm:</u> Her i için $b_i = \sum_{j=1}^{\infty} a_{ij}$ ve $\sum_{i=1}^{\infty} b_i$ toplamlarının yakınsak olduğunu kabul edelim.

$E = \{1/n \mid n \in \mathbb{Z}^+ \cup \{0\}\}$ $\mathbb{R}$'nin tıkız bir alt kümesidir. Bu tıkız küme üzerinde aşağıdaki fonksiyon dizisini tanımlayalım:

$i \in \mathbb{Z}^+$, $f_i : E \to \mathbb{R}$, $f_i(0) = b_i = \sum_j a_{ij}$ ve $f_i(1/n) = \sum_{j=1}^{n} a_{ij}$.

Ayrıca, $g : E \to \mathbb{R}$, $g(x) = \sum_{i=1}^{\infty} f_i(x)$ ile tanımlansın.

<u>İddia 1:</u> Her f_i fonksiyonu süreklidir.

<u>Kanıt:</u> Sadece 0 noktasında sürekli olduğunu göstermek yeterlidir:

$$f_i(0) = \sum_{j=1}^{\infty} a_{ij} = \lim_{n\to\infty} \sum_{j=1}^{n} a_{ij} = \lim_{n\to\infty} f_i(1/n).$$

Bu kanıtı tamamlar. ∎

<u>İddia 2:</u> $g = \sum_{i=1}^{\infty} f_i$ E üzerinde düzgün yakınsaktır.

Kanıt: Weierstrass M-testi kullanacağız.

$|f_i(x)| \le \sum_{j=1}^{\infty} a_{ij} = b_i$ ve $\sum_{i=1}^{\infty} b_i = L$ yakınsak olduğu

için $g = \sum f_i$ toplamı düzgün yakınsaktır. ∎

Her bir f_i sürekli olduğu için g de süreklidir.

O halde,

$$\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{ij}\right) = \sum_{i=1}^{\infty} f_i(0) = g(0) = g\left(\lim_n \tfrac{1}{n}\right)$$

$$= \lim_n g(1/n)$$

$$= \lim_n \sum_{i=1}^{\infty} f_i(1/n)$$

$$= \lim_n \sum_{i=1}^{\infty}\left(\sum_{j=1}^{n} a_{ij}\right)$$

$$= \lim_n \sum_{j=1}^{n}\left(\sum_{i=1}^{\infty} a_{ij}\right)$$

$$= \sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} a_{ij}\right).$$

Bu kanıtı bitirir. ∎

Alıştırma 21: $\mathbb{E}[\cdot]$ fonksiyonunun doğrusal olduğunu gösteriniz.

Çözüm: $a, b \in \mathbb{R}$ ve X, Y rasgele değişkenler ise

$$\mathbb{E}(aX+bY) = \sum_{s\in S} (aX+bY)(s)\, \mathbb{P}(\{s\})$$

$$= a\sum_{s\in S} X(s)\, \mathbb{P}(\{s\}) + b\sum_{s\in S} Y(s)\, \mathbb{P}(\{s\})$$ olur.

(Burada hem $\mathbb{E}X$ hem de $\mathbb{E}Y$'nin tanımlı olduğunu kabul ediyoruz.)

Alıştırma 22: Eğer $Ran(N) = \{1, 2, 3, \dots\}$ ise

$$\mathbb{E}N = \sum_{n=1}^{\infty} \mathbb{P}(N \ge n)$$ olduğunu gösteriniz.

Çözüm: $$\sum_{n=1}^{\infty} \mathbb{P}(N\ge n) = \sum_{n=1}^{\infty}\sum_{k=n} P(N=k)$$

$$= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mathbb{1}_{\{k \geq n\}} \mathbb{P}(N=k)$$

(Alıştırma 20) $$= \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mathbb{1}_{\{k \geq n\}} \mathbb{P}(N=k)$$

$(n \Leftrightarrow k)$ $$= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mathbb{1}_{\{n \geq k\}} \mathbb{P}(N=n)$$

$$= \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} \mathbb{1}_{\{n \geq k\}} \right) \mathbb{P}(N=n)$$

$$= \sum_{n=1}^{\infty} n\, \mathbb{P}(N=n) = \mathbb{E} N.$$

Alıştırma 23: N yazı tura deneyinde tura gelene kadar yapılan atış sayısını gösteren rasgele değişken olsun. Alıştırma 19'dan $\mathbb{P}(N=n) = p(1-p)^{n-1}$ olduğunu biliyoruz.

O halde, $$\mathbb{E} N^2 = \sum_{n=1}^{\infty} n^2\, \mathbb{P}(N=n)$$

$$= \sum_{n=1}^{\infty} n^2 p (1-p)^{n-1}$$

$$= (2-p)/p^2$$ elde edilir.

Burada göre $f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ fonksiyonunu kullandık!

§1.14 Momentumlar: X rasgele değişkeninin k'inci momentumu $\mathbb{E} X^k$ olarak tanımlanır (eğer yakınsak ise). Dolayısıyla birinci momentum beklenen değer yani ortalama değerdir.

Alıştırma 24. Eğer $\mathbb{E} X^k$ varsa tüm $1 \leq j \leq k$ için $\mathbb{E} X^j$'de vardır.

Çözüm: $$\mathbb{E} |X|^j = \sum_x |x^j| \mathbb{P}(X=x) \leq \sum_x (1+|x|^k) \mathbb{P}(X=x)$$

$$\Rightarrow \mathbb{E} |X|^j \leq \mathbb{E} |X|^k + 1.$$

Tanım: Bir rasgele değişkenin Varyansı aşağıdaki gibi tanımlanır:

$$VAR(X) = \mathbb{E}(X - \mathbb{E}X)^2.$$

Hatırlatma: 1) $\mathbb{E}(X-\mathbb{E}X) = \mathbb{E}X - \mathbb{E}X \cdot 1 = 0.$

2) Eğer X'in 2. momentumu varsa, o zaman

$$VAR(X) = \mathbb{E}(X^2 - 2\mathbb{E}X\, X + (\mathbb{E}X)^2) = \mathbb{E}X^2 - 2\mathbb{E}X\,\mathbb{E}X + (\mathbb{E}X)^2$$

$$\Rightarrow VAR(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 \text{ olur.}$$

Tanım: X rasgele değişkeninin standard sapması

$$STD(X) = \sqrt{VAR(X)} \text{ olarak tanımlanır.}$$

Alıştırma 25: Bir rasgele X değişkeni için $Ran(X) = \{1, 2, \ldots\}$ ve $\mathbb{P}(X=k) = c\,k^{-t}$, $t>1$ ve $c>0$ sabit olmak üzere, olsun. X'in hangi momentumları vardır?

Tanım: Bir X rasgele değişkeninin momentum üretme fonksiyonu (mgf) şu şekilde tanımlanır: $t \in \mathbb{R}$ olmak üzere

$$M_X(t) = \mathbb{E}\, e^{tX} = \sum_x e^{tx}\, \mathbb{P}(X=x).$$

Uyarı: $M_X(0) = 1$ olduğu açıktır. Ayrıca, $e^{tX} > 0$ olduğu için mgf ya sonludur ya da $+\infty$'dur.

§1.15. Rasgele Değişken Vektörlerin Fonksiyonları:

$X_1, \ldots, X_n$ rasgele değişkenler, $p(x_1, \ldots, x_n)$ bunların birleşik olasılık kütle fonksiyonu ve $g: \mathbb{R}^n \to \mathbb{R}$ bir fonksiyon olsun. Bu durumda

$$\mathbb{E}\, g(X_1, \ldots, X_n) = \sum_{x_1, \ldots, x_n} g(x_1, \ldots, x_n)\, p(x_1, \ldots, x_n) \text{ olur.}$$

Eğer $g(x_1, \ldots, x_n) = x_i$, izdüşüm fonksiyonu ise

$$\mathbb{E}\, g(X_1, \ldots, X_n) = \mathbb{E}X_k = \sum_{x_1, \ldots, x_n} x_k\, p(x_1, \ldots, x_n) = \sum_{x_k} x_k\, p(x_k),$$

$p(x_k) = \mathbb{P}(X_k = x_k)$ olur.

Bu gözlemin bir uygulaması olarak $\mathbb{E}(aX+bY) = a\mathbb{E}X + b\mathbb{E}Y$ ve $COV(X,Y) \doteq \mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$ elde ederiz.

COV(X,Y), X ve Y'nin kovaryansı olarak adlandırılır ve bu değişkenlerin arasındaki ilişkinin bir ölçüsüdür. Aşağıdaki alıştırma buna bir destek sunmaktadır.

Alıştırma 26: X ve Y bağımsız değişkenler ise $COV(X,Y)=0$ olur.

Çözüm: $Ran(X)=\{x_i\}$ ve $Ran(Y)=\{y_j\}$ olsun. Bu durumda $\mathbb{E}(XY)=\sum_{i,j} x_i y_j \, \mathbb{P}(X=x_i, Y=y_j)$

$$
\begin{aligned}
(X,Y \text{ bağımsız!}) &= \sum_{i,j} x_i y_j \, \mathbb{P}(X=x_i)\mathbb{P}(Y=y_j) \\
&= \Big(\sum_i x_i \, \mathbb{P}(X=x_i)\Big)\Big(\sum_j y_j \mathbb{P}(Y=y_j)\Big) \\
&= (\mathbb{E}X)(\mathbb{E}Y) \text{ elde edilir.}
\end{aligned}
$$

O halde, $COV(XY)=\mathbb{E}(XY)-(\mathbb{E}X)(\mathbb{E}Y)=0$ olur.

Üçüncü eşitlik için biraz daha detay verebiliriz:

$A=\sum a_n$, $B=\sum b_n$ yakınsak seriler olsun. Bu durumda

$$
\begin{aligned}
AB &= A\Big(\lim_M \sum_{m=1}^{M} b_m\Big) \\
&= \lim_M \Big(A\sum_{m=1}^{M} b_m\Big) \\
&= \lim_M \sum_{m=1}^{M} (b_m A) \\
&= \lim_M \sum_{m=1}^{M} b_m \Big(\sum_{n=1}^{\infty} a_n\Big) \\
&= \lim_M \sum_{m=1}^{M} \Big(\sum_{n=1}^{\infty} b_m a_n\Big) \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} b_m a_n \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_n b_m
\end{aligned}
$$

elde edilir. Bu kanıtı bitirir.

Alıştırma 27: X ve Y atılan bir çift zarın en büyük ve en küçük değerlerini gösteren rasgele değişkenler olmak üzere COV(X,Y) kovaryansını hesaplayınız.

Çözüm: $\mathbb{E}X = 161/36$ olduğunu göstermiştik. Benzer şekilde $\mathbb{E}Y$ ve $\mathbb{E}(XY)$ de hesaplanabilir.

X ve Y bağımsız ise $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$ olduğunu biliyoruz. Benzer şekilde $X_1, \dots, X_n$ değişkenleri ikili olarak bağımsızsa

$$
\begin{aligned}
VAR[X_1 + X_2 + \dots + X_n] &= \mathbb{E}(X_1 + \dots + X_n - \mathbb{E}(X_1 + \dots + X_n))^2 \\
&= \mathbb{E}(X_1 - \mathbb{E}X_1 + \dots + X_n - \mathbb{E}X_n)^2 \\
&= \sum_i \mathbb{E}(X_i - \mathbb{E}X_i)^2 + \sum_{i \neq j} \mathbb{E}(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j) \\
&= \sum_i VAR[X_i] + \sum_{i \neq j} COV(X_i, X_j) \\
&= \sum_i VAR[X_i] + 0 \\
&= \sum_i VAR[X_i].
\end{aligned}
$$

Alıştırma 28: Binomial rasgele değişkenin ortalama değerini ve varyansını hesaplayın.

Çözüm: $Ran(X) = \{0, 1, 2, \dots, n\}$ ve $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ olduğunu biliyoruz. O halde,

$$
\begin{aligned}
\mathbb{E}X &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\
(l = k-1) \quad &= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-l)!} p^l (1-p)^{n-1-l} \\
&= np\,(p + (1-p))^{n-1} \\
&= np.
\end{aligned}
$$

Bir diğer çözüm şu olabilir: X_k ile k'ıncı atışta tura gelirse 1 yoksa sıfır değerini alan değişkeni gösterelim.

O halde, $X = X_1 + \cdots + X_n$ olur. $\mathbb{E}X_i = 1\cdot p + 0\cdot(1-p) = p$ olduğu açıktır. O halde, $\mathbb{E}X = \mathbb{E}X_1 + \cdots + \mathbb{E}X_n = np$ olur.

Bunu ve bir önceki paragraftaki çıkarımı kullanırsak

$$VAR(X) = \sum_{k=1}^{n} VAR(X_i) = n\,VAR(X_1) = n\,\mathbb{E}(X_1 - \mathbb{E}X_1)^2$$
$$= n\left(\mathbb{E}X_1^2 - (\mathbb{E}X_1)^2\right) = n\left(\mathbb{E}X_1 - (\mathbb{E}X_1)^2\right) \quad (X_1^2 = X_1!)$$
$$= n(p - p^2) = np(1-p) \text{ elde ederiz.}$$

(Burada X_i'lerin bağımsız olduğunu kullandık!)

<u>Alıştırma 29</u>: $VAR[X_1 + \cdots + X_n] = \sum_{k=1}^{n} VAR[X_k] + 2\sum_{i<j} COV(X_i, X_j)$ olduğunu kanıtlayınız.

§1.16. <u>Koşullu Dağılım ve Beklenti</u>:

X ve Y iki rasgele değişken olmak üzere, X'in $\{Y=y\}$ olayına koşullu dağılımı

$$P(X=x \mid Y=y) \doteq \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)} \quad (P(Y=y) \neq 0 \text{ olmamak üzere})$$

ile tanımlanır.

Benzer şekilde, X değişkeninin verilen bir A olayına koşullu beklenen değeri

$$\mathbb{E}[X \mid A] = \sum_x x\,\mathbb{P}(X=x \mid A) \text{ ile tanımlanır.}$$

Daha önce yaptığımız gibi pozitif bir değişken için bu toplam ya sonludur yada $+\infty$'dur. Genel bir X değişkeni için ise yine $X = X^+ - X^-$ yazarız ve koşullu beklenen değeri

$$\mathbb{E}[X \mid A] \doteq \mathbb{E}[X^+ \mid A] - \mathbb{E}[X^- \mid A] \text{ ile tanımlarız.}$$

Eğer $A = \{Y=s\}$, $s \in S$, ise $\mathbb{E}[X \mid Y]$ ifadesini S üzerinde bir fonksiyon olarak görebiliriz:

$$\mathbb{E}[X \mid Y] : S \longrightarrow \mathbb{R}, \quad s \longmapsto \mathbb{E}[X \mid Y = Y(s)], \quad s \in S.$$

Bu fonksiyonun kendisini rasgele değişken olarak alıp onun beklenen değerini hesaplayabiliriz:

$$\mathbb{E}[\mathbb{E}[X|Y]] = \sum_y \mathbb{E}[X|Y=y]\,\mathbb{P}(Y=y).$$

Alıştırma 30: Eğer $\mathbb{E}X$ tanımlı ise $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}X$ olur.

Çözüm:
$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|Y]] &= \sum_y \mathbb{E}[X|Y=y]\,\mathbb{P}(Y=y)\\
&= \sum_y \sum_x x\,\mathbb{P}(X=x|Y=y)\,\mathbb{P}(Y=y)\\
&= \sum_x \sum_y x\,\frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)}\,\mathbb{P}(Y=y)\\
&= \sum_x \sum_y x\,\mathbb{P}(X=x, Y=y)\\
&= \sum_x x\,\mathbb{P}(X=x, \bigcup_y \{Y=y\})\\
&= \sum_x x\,\mathbb{P}(X=x)\\
&= \mathbb{E}[X].\end{aligned}$$

Alıştırma 31. $N, X_1, X_2, \dots$ S ayrık örneklem uzayı üzerinde bağımsız rasgele değişkenler olsun. X_k'lar ortalamaları $\mathbb{E}X_k = \mu$ olan sonlu denk (identically distributed) değişkenler olsun. Ayrıca $Ran(N) = \{1, 2, \dots\} = \mathbb{N}$ ve $\mathbb{E}N < +\infty$ olduğunu kabul edelim.

$$Y(s) = \sum_{n=1}^{N(s)} X_n(s)$$ ile tanımlansın.

$\mathbb{E}Y = \mu\,\mathbb{E}N$ olduğunu kanıtlayınız.

Çözüm: $\mathbb{E}[Y|N] : S \to \mathbb{R}$, $\mathbb{E}[Y|N](s) = \mathbb{E}[Y|N=N(s)]$ olarak tanımlanmıştı.

O halde,
$$\begin{aligned}\mathbb{E}[Y|N](s) &= \mathbb{E}\Big[\sum_{n=1}^{N(s)} X_n\Big]\\
&= \sum_{n=1}^{N(s)} \mathbb{E}[X_n]\\
&= \mu\,N(s).\end{aligned}$$

O halde, bir önceki alıştırmadan

$$\mathbb{E}Y = \mathbb{E}[\mathbb{E}[Y|N]] = \mathbb{E}[\mu N] = \mu \mathbb{E}N$$ elde ederiz.

Alıştırma 32. Bir önceki alıştırmayı kullanarak şunu kanıtlayın. $VAR[X_k] < +\infty$ ve $VAR[N] < +\infty$ olduğunu kabul edelim.

$$VAR[Y] = (\mathbb{E}N)\, VAR[X_1] + \mu^2 VAR[N].$$

Çözüm: $VAR[Y] = \mathbb{E}Y^2 - (\mathbb{E}Y)^2$, $Y = \sum_{n=1}^{N} X_n$ olduğunu biliyoruz.

Ayrıca, $\mathbb{E}Y = \mu \mathbb{E}N$ olduğunu gösterdiğimiz için $(\mathbb{E}Y)^2 = \mu^2 (\mathbb{E}N)^2$ olur.

O halde, $Y^2 = \sum_{n=1}^{N} X_n^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{N} X_i X_j$ ve buradan da

$$\mathbb{E}Y^2 = \mathbb{E}\Big(\sum_{n=1}^{N} X_n^2\Big) + \mathbb{E}\Big(\sum_{\substack{i,j=1 \\ i \neq j}}^{N} X_i X_j\Big)$$

$$= \mathbb{E}X_1^2\, \mathbb{E}N + \mathbb{E}(X_i X_j)\, \mathbb{E}(N^2 - N) \quad \big(X_i \neq X_j \text{ için } \mathbb{E}(X_i X_j) = (\mathbb{E}X_1)^2\big)$$

$$= \mathbb{E}X_1^2\, \mathbb{E}N + (\mathbb{E}X_1)^2\, \mathbb{E}(N^2 - N)$$

$$= \mathbb{E}X_1^2\, \mathbb{E}N + \mu^2 (\mathbb{E}N^2 - \mathbb{E}N)$$ olur.

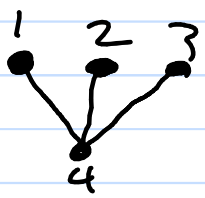
O halde, $VAR[Y] = \mathbb{E}Y^2 - (\mathbb{E}Y)^2$

$$= \mathbb{E}X_1^2\, \mathbb{E}N + \mu^2(\mathbb{E}N^2 - \mathbb{E}N) - \mu^2(\mathbb{E}N)^2$$

$$= (\mathbb{E}X_1^2 - \mu^2)\, \mathbb{E}N + \mu^2(\mathbb{E}N^2 - (\mathbb{E}N)^2)$$

$$= VAR[X_1]\, \mathbb{E}N + \mu^2 VAR[N]$$ elde edilir.

Alıştırma 33. Labirentteki bir farenin önünde üç kapı vardır. 1. kapıdan çıkarsa 1 dakika içinde özgürlüğüne kavuşuyor. 2. kapıdan çıkarsa 3 dakika sonra tekrar aynı odaya dönüyor. Son olarak, 3. kapıdan çıkarsa 5 dakika sonra aynı odaya dönüyor. Farenin üç kapıyı da seçme ihtimali aynı ise ortalama kaç dakika sonra özgürlüğüne kavuşur.

Çözüm: 

E_i, $i = 1,2,3,4$, ile i. köşeden başladığında özgürlüğüne kavuşmasının ortalama zamanı olsun.

O halde, $E_1 = 0$, $E_2 = E_4 + 3$, $E_3 = E_4 + 5$ ve son olarak

$$E_4 = \frac{1}{3}\cdot E_1 + \frac{1}{3}E_2 + \frac{1}{3}E_3 = \frac{1}{3}\cdot 0 + \frac{1}{3}(E_4+3) + \frac{1}{3}(E_4+5)$$

$\Rightarrow \frac{1}{3}E_4 = 1 + \frac{5}{3} \Rightarrow E_4 = 3+5 = 8$ dakika bulunur.

Bölüm 2. Ayrık zaman sonlu durum Markov zincileri.

Markov zincirleri hem bilim hemde mühendislik alanında çokça başvurulan bir matematiksel yapıdır.

§2.1. Tanımlar:

$S = \{1, 2, \ldots, N\}$ olsun. S değerli rasgele (X_n) değişkenleri, her $n \geq 1$ ve $j, i_0, \ldots, i_{n-1} \in S$ için

$$P(X_n = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = P(X_n = j \mid X_{n-1} = i_{n-1})$$

Markov Zincir Koşulu'nu sağlıyorsa bu değişken ailesine bir Markov Zinciri denir.

Eğer $\{X_n\}$ Markov zinciri her n için

$$P(X_n = j \mid X_{n-1} = i) = P(X_1 = j \mid X_0 = i) = p_{ij} \text{ koşulunu}$$

sağlıyorsa bu zincire homogen zincir denir. Bu durumda $P = (p_{ij})_{N\times N}$ matrisine zincirin geçişme matrisi denir. $P = (p_{ij})$ matrisi şu özelliklere sahiptir.

$P1)$ $p_{ij} \geq 0 \quad \forall i, j \in S$

$P2)$ $\sum_{j\in S} p_{ij} = 1, \quad \forall i \in S.$

Bu özelliklere sahip matrise stokastik olmayan matris denir.

$\alpha_i = P(X_0 = i)$ olarak tanımlanırsa

$P(X_0 = i_0, X_1 = i_1, \ldots, X_m = i_m) = \alpha_{i_0}\, p_{i_0 i_1} \cdot p_{i_1 i_2} \cdots p_{i_{m-1} i_m}$ elde edilir.

Bu formül $P(A_0 \cap \cdots \cap A_n) = P(A_0) P(A_1 \mid A_0) P(A_2 \mid A_1 \cap A_0) \cdots P(A_n \mid A_0 \cap \cdots \cap A_{n-1})$ eşitliğinin bir sonucudur. Gerçekten de $A_k = \{X_k = i_k\}$ ile

$$\mathbb{P}(X_0=i_0,\dots,X_m=i_m)=\mathbb{P}(A_0\cap\dots\cap A_m)$$
$$=\mathbb{P}(A_0)\,\mathbb{P}(A_1|A_0)\cdots\mathbb{P}(A_m|A_0\cap\dots\cap A_{m-1})$$
$$=\mathbb{P}(X_0=i_0)\,\mathbb{P}(X_1=i_1|X_0=i_0)\,\mathbb{P}(X_2=i_2|X_1=i_1,X_0=i_0)\cdots\mathbb{P}(X_m=i_m|X_{m-1}=i_{m-1},\dots,X_0=i_0)$$
$$=\mathbb{P}(X_0=i_0)\,\mathbb{P}(X_1=i_1|X_0=i_0)\,\mathbb{P}(X_2=i_2|X_1=i_1)\cdots\mathbb{P}(X_m=i_m|X_{m-1}=i_{m-1})$$
$$=\alpha_{i_0}P_{i_0i_1}P_{i_1i_2}\cdots P_{i_{m-1}i_m}$$ elde edilir.

Geçişme matrisi zincirin nasıl değişeceğini tamamen belirler. Örneğin,

$$\mathbb{P}(X_2=j|X_0=i)=\sum_k \mathbb{P}(X_2=j,X_1=k|X_0=i)$$
$$=\sum_k \mathbb{P}(X_2=j,X_k=k,X_0=i)/\mathbb{P}(X_0=i)$$
$$=\sum_k \frac{\mathbb{P}(X_2=j,X_k=k,X_0=i)}{\mathbb{P}(X_0=i,X_1=k)}\,\frac{\mathbb{P}(X_0=i,X_1=k)}{\mathbb{P}(X_0=i)}$$
$$=\sum_k \mathbb{P}(X_2=j|X_1=k,X_0=i)\,\mathbb{P}(X_1=k|X_0=i)$$
$$=\sum_k \mathbb{P}(X_2=j|X_1=k)\,\mathbb{P}(X_1=k|X_0=i)$$
$$=\sum_k P_{kj}P_{ik}=(P^2)_{ij}$$

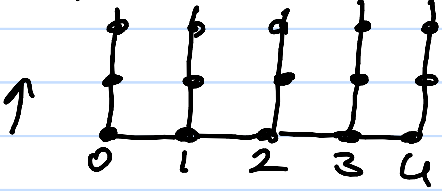
Benzer şekilde, $\mathbb{P}(X_k=j|X_0=i)=(P^k)_{ij}$ olur. Başka bir deyişle P^k k-adım geçiş fonksiyonudur.

$P_{ij}(n)$ ile $(P^n)_{ij}$ elemanını göstereceğiz. $P_{ij}(n)>0$ olması n adımda i'inci durumdan j'inci duruma geçmenin mümkün olduğu anlamına gelir.

Örnekler: 1) Şu modeli düşünelim: İki kutuda ikisi siyah, ikisi beyaz toplam 4 top vardır. İlk başta ikisinde de ikişer top olsun. Her seferinde iki kutudan bir top seçiliyor ve diğer kutuya konuluyor. n adım sonra ilk kutudaki beyaz top sayısını X_n değişkeni göstersin. O halde, $\mathrm{Ran}(X_n)=\{0,1,2\}$ dir. X_{n+1} sadece X_n'e bağlı olduğu için $\{X_n\}$ bir Markov zinciridir. P geçişme matrisi ise

$$P=\begin{pmatrix}0&1&0\\1/4&1/2&1/4\\0&1&0\end{pmatrix}$$ olur.

2) <u>Sarhoşun Yürüyüşü</u>: Bir sarhoş her seferinde bir adım atarak en alttan yürümeye başlıyor. Her adımda $1/2$ olasılıkla ya sağa yada sola bir adım kayıyor. $\{0,1,2,3,4\}$ durum uzayında değer alan X_n rasgele değişkeni sarhoşun n-adım sonra hangi çizgi üzerinde olduğunu göstersin. Bu modelin geçişim matrisini bulalım.



$$P=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

<u>Alıştırma 34.</u> Aşağıdaki modellerin birer Markov zinciri olup olmadığına karar veriniz. Bir deste oyun kağıdı karılıyor. 1) En üstteki kağıt alınıyor, değeri kaydediliyor ve tekrar desteye rasgele şekilde geri konuluyor. Bu şekilde devam ediliyor. 2) Üsteki modelin aynısı, fakat bu sefer kağıt desteye geri konmuyor.

<u>Alıştırma 35.</u> $\{X_i\}$'ler aynı olasılık dağılımına sahip rasgele değişkenler ailesi olsun. $S_n=\sum_{i=1}^{n} X_i$ olmak üzere aşağıdaki dizilerin Markov zinciri olup olmadığına karar veriniz.

i) $\{S_n\}$, ii) $\{S_{\gamma_n}\}$, $\gamma_n=\min\{k<n \mid S_k=\max\{S_1,\dots,S_n\}\}$

iii) (S_n, S_{γ_n}) sıralı ikilisi.

(İpucu: (ii) için $Ran(X_n)=\{0,-1,1\}$ alın!)

§2.2. Yutan Zincirler:

<u>Tanım.</u> Bir $i\in S$ durumu için $P_{ii}=1$ ise bu duruma yutan durum denir. $P_{ii}=1$ olduğu için $P_{ij}=0$, $\forall j\neq i$ olur. Dolayısıyla, i durumundan başka bir duruma geçiş mümkün değildir. $P_{ii}\neq 0$ değilse $i\in S$ durumuna geçiş durumu denir.

Bir zincirde r tane yutan durum, t tane de geçiş durumu olsun. Bu durumda durumları yeniden sıralayarak geçiş matrisinin şu şekilde olmasını sağlayabiliriz:

$$P=\begin{pmatrix} Q & R \\ 0 & I_r \end{pmatrix}, \quad I_r=\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}_{r\times r}.$$

Alıştırma 36. Sarhoşun yürüyüşü için Q, R ve P matrislerini belirleyiniz.

Alıştırma 36'. $P=\begin{pmatrix} Q & R \\ 0 & I_r \end{pmatrix}$ ise $P^n=\begin{pmatrix} Q^n & R_n \\ 0 & I_r \end{pmatrix}$ olduğunu gösteriniz.

Yardımcı Teorem 4. $n\to+\infty$ sonsuza giderken her i,j için $(Q^n)_{ij}\to 0$ olduğunu gösteriniz.

Kanıt: Her $i\in T$ geçiş durumu için bir $k\in S$ yutan durumu, bir $n_i\in\mathbb{Z}$ tamsayısı ve bir $\delta_i>0$ gerçel sayısı vardır ki $p_{ik}(n_i)=\delta_i>0$ olur. $n=\max\{n_i\}$ ve $\delta=\min\{\delta_i\}>0$ olsun. O halde, her $i\in T$ geçiş durumu için bir $k\in S$ yutan durumu vardır ve $p_{ik}(n)\geq\delta$ olur. Buradan,

$$\sum_{j\in T} Q^n_{ij}=\sum_{j\in T} P^n_{ij}=1-\sum_{k\in R} P^n_{ik}=1-\sum_{k\in R} p_{ik}(n)\leq 1-\delta.$$

Özel olarak, her $i\in T$ için $\sum_{j\in T} Q^n_{ij}\leq 1-\delta$ elde ederiz.

O halde, her $i\in T$ için,

$$\sum_{j\in T} Q^{2n}_{ij}=\sum_{k\in T} Q^n_{ik}Q^n_{kj}\leq\sum_{k\in T} Q^n_{ik}\Big(\sum_{j\in T} Q^n_{kj}\Big)\leq(1-\delta)\sum_{k\in T} Q^n_{ik}\leq(1-\delta)^2$$

olur. Tümevarım yöntemiyle $\sum_{j\in T} Q^{kn}_{ij}\leq(1-\delta)^k$ ve dolayısıyla

$k\to\infty$ giderken $\sum_{j\in T} Q^{kn}_{ij}\to 0$ elde edilir ($\forall i\in T$ için).

Son olarak, $\sum_{j\in T} Q^{m+1}_{ij}=\sum_{j\in T}\sum_{k\in T} Q^m_{ik}Q_{kj}=\sum_{k\in T} Q^m_{ik}\underbrace{\Big(\sum_{j\in T} Q_{kj}\Big)}_{\leq 1}\leq\sum_{k\in T} Q^m_{ik}$ olur ve

$\Big(\sum_{k\in T} Q^m_{ik}\Big)_{m=1}^{\infty}$ dizisinin monotone azalan olduğunu görünüz.

Bir alt dizisi 0'a yakınsadığı için kendisi de 0'a yakınsar. Böylece kanıt tamamlanır. ■

Dolayısıyla, geçiş durumlarının olasılıkları, zamanla yutan durumlara kayarak, sıfıra yaklaşır.

Son olarak $x = Qx$ denklemini ele alalım. Her iki tarafı da Q ile çarpalım: $x = Qx = Q^2x$ ve buradan $x = Q^n x$ elde ederiz. $Q^n \to 0$ olduğu için $x = 0$ elde ederiz. Dolayısıyla, $I-Q$ matrisi tekil değildir. Başka bir deyişle tersi vardır. $N = (I-Q)^{-1}$ olsun. $(I+Q+\cdots+Q^n)(I-Q) = I - Q^{n+1}$ olduğu için

$$N = (I-Q)^{-1} = \sum_{n=0}^{\infty} Q^n \text{ elde ederiz.}$$

<u>Teorem 5.</u> $i, j \in T \subseteq S$ geçiş durumları olsun. O halde,

(1) N_{ij} i'den başlayan turların j'ye uğrama sayısının ortalama değeridir.

(2) $\sum_j N_{ij}$ i'den başlayan ve yutan bir duruma ulaşılana kadar (yutulana kadar) geçen ortalama adım sayısıdır.

(3) $B = NR$, $t \times n$-matrisi olsun. Bu durumda B_{ik} i'den başlayıp k yutan durumuna ulaşmanın olasılığını gösterir.

<u>Kanıt:</u> Zincirin ilk adımı $X_0 = i$ olsun. Verilen bir $j \in T$ geçiş durumu ve $k \geq 0$ tam sayısı için

$$Y^{(k)} = \begin{cases} 1 & , X_k = j \\ 0 & , X_k \neq j \end{cases} \quad \text{rasgele değişkenini tanımlayalım.}$$

O halde, $\mathbb{E}Y^{(k)} = \mathbb{P}(Y^{(k)} = 1) = \mathbb{P}(X_k = j) = P_{ij}(k) = (Q^k)_{ij}$, $k \geq 1$ ve $\mathbb{E}Y^{(0)} = \delta_{ij}$ olur.

Dolayısıyla ilk n adımda j durumuna ulaşma sayısı $Y^{(0)} + Y^{(1)} + \cdots + Y^{(n)}$ olur. Bu toplamın beklenen (ortalama) değeri ise

$$\delta_{ij} + Q_{ij} + (Q^2)_{ij} + \cdots + (Q^n)_{ij} = (I + Q + \cdots + Q^n)_{ij}$$

olur. Bu ifade $n \to \infty$ durumunda N_{ij} sayısına yakınsar ve böylece (1)'in kanıtı tamamlanır.

(2) ise (1)'in sonucudur.

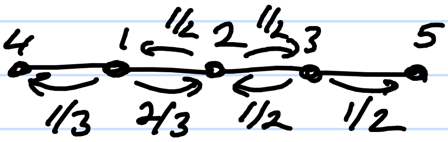
(3) için ise şu şekilde düşünelim: $i \in T$, $k \in S \setminus T$.

$$(NR)_{ik} = \sum_{j \in T} N_{ij} R_{jk} = \sum_{j \in T} \sum_{n=0}^{\infty} (Q^n)_{ij} R_{jk} = \sum_{n=0}^{\infty} \sum_{j \in T} (Q^n)_{ij} R_{jk},$$

elde edilir. Burada $\sum_{j\in T}(Q^n)_{ij}R_{jk}$ zincirin n-adım sonunda $k\in R$ yutan durumuna ulaşma olasılığını göstermektedir. Bu kanıtı tamamlar ■

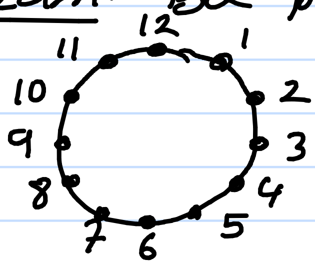
Alıştırma 37. Sarhoşun yürüyüşü modelinde $Q=\begin{pmatrix}0 & 1/2 & 0\\ 1/2 & 0 & 1/2\\ 0 & 1/2 & 0\end{pmatrix}$, $R=\begin{pmatrix}1/2 & 0\\ 0 & 0\\ 0 & 1/2\end{pmatrix}$ ve $P=I_2=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ dir. O halde, $Q^{2n+1}=2^{-n}Q$, $Q^{2n+2}=2^{-n}Q^2$, $N=\begin{pmatrix}3/2 & 1 & 1/2\\ 1 & 2 & 1\\ 1/2 & 1 & 3/2\end{pmatrix}$ ve $B=\begin{pmatrix}3/4 & 1/4\\ 1/2 & 1/2\\ 1/4 & 3/4\end{pmatrix}$ olduğunu gösteriniz.

Alıştırma 38. Yukarıdaki örneği olasılıkları aşağıdaki gibi alarak tekrar yapınız.



Örnek: Ali'nin 7 ve Berna'nın 5 bilyesi vardır. İkisi şu oyunu oynuyor. Yazı tura atıyorlar ve kazanan diğerinden bir bilye alıyor. 12 bilyeyi de toplayan oyunu kazanıyor. Bu oyun ortalama kaç yazı-tura atışı sonunda biter.

Çözüm. Bu problemi şu şekilde modelleyelim:



Her saniye 5 numaralı delikten bir parçacık çıkıyor. Her bir parçacık eşit olasılıkla ya bir adım ileri yada bir adım geri gidiyor. 12'ye ulaşan parçacık yok oluyor.

Yukarıdaki problemin cevabı uzun vadede (zaman sonsuza giderken sistemde kalan ortalama parçacık sayısıdır.

Her $k\in S=\{1,2,\dots,12\}$ ve $n\geq 0$ tam sayısı için $E_k(n)$ değişkenini tanımlayalım:

$E_k(n)$ = "5 nolu delikten çıkan bir parçacığın n saniye sonra k numaralı noktada olma olasılığı".

O halde, bu değişkenler için şu bağıntıları yazabiliriz:

$$E_1(n) = \tfrac{1}{2} E_2(n-1),$$

$$E_2(n) = \tfrac{1}{2} E_1(n-1) + \tfrac{1}{2} E_3(n-1),$$

$$\vdots$$

$$E_{10}(n) = \tfrac{1}{2} E_9(n-1) + \tfrac{1}{2} E_{11}(n-1),$$

$$E_{11}(n) = \tfrac{1}{2} E_9(n-1)$$

$$E_{12}(n) = \tfrac{1}{2} E_1(n-1) + \tfrac{1}{2} E_{11}(n-1).$$

Ayrıca $E(n) = [E_1(n), \dots, E_{12}(n)]^T$ ve $E(0) = [0\ 0\ 0\ 0\ 1\ 0 \cdots 0]^T$ olsun.

Bu durumda $E_n = A E_{n-1}$ denklemini yazabiliriz.

Burada A matrisi şudur:

$$A = \begin{bmatrix} 0 & 1/2 & 0 & \cdots & 0 & 0 \\ 1/2 & 0 & 1/2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & & \ddots & 1/2 & \vdots \\ \vdots & & & & 0 & \vdots \\ 1/2 & 0 & 0 & \cdots & 1/2 & 0 \end{bmatrix}$$

$E(n) = A E(n-1)$ denkleminden $E(n) = A^n E(0)$ elde edilir.

Dolayısıyla, $t=0$'da 5. delikten çıkan bir parçacığın n saniye sonra k'ıncı noktada olma olasılığı $E_k(n)$, $E(n)$ vektörünün k'ıncı bileşenidir. O halde, n saniye sonra k'ıncı noktadaki ortalama parçacık sayısı $E_k(0) + \dots + E_k(n)$ olur.

Sonuç olarak zaman sonsuza giderken k'ıncı noktadaki parçacık sayısı aşağıdaki sonsuz toplamdır:

$$\sum_{i=0}^{\infty} E_k(i) = \left(\sum_{i=0}^{\infty} A^i E(0)\right)_k = \left(\left(\sum_{i=0}^{\infty} A^i\right) E(0)\right)_k$$

A matrisinin normu $\|A\| = \frac{1}{2} < 1$ olduğu için $\sum_{i=0}^{\infty} A^i$ yakınsaktır ve $(I_{12} - A)^{-1}$ matrisine eşittir.

Hesaplamalardan $(I_{12} - A)^{-1}$ matrisinin 5'inci sütunu

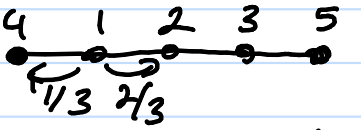
$[7/6\ \ 7/3\ \ 7/2\ \ 14/3\ \ 35/6\ \ 5\ \ 25/6\ \ 10/3\ \ 5/12\ \ 5/3\ \ 5/6\ \ 1]^T$ olarak bulunur.

$$E_T(\infty) \doteq \left(\sum_{i=0}^{\infty} A^i\right) E(0) = (I_{12} - A)^{-1} E(0) = (I_{12} - A)^{-1}$$ matrisinin 5'inci sütunu.

Bu sayılar, zaman sonsuza giderken tabla üzerinde her bir noktadaki parçacık sayısını gösterir. Son bileşenin 1 olması her saniye bir parçacığın yok olduğunu gösteriyor.

Bu durum beklenendir çünkü her saniye bir parçacık doğuyor ve tabla üzerindeki parçacık sayısı bir dengeye ulaşıyor.
Bizim başlangıçtaki sorumuzun cevabı ise bu sayıların toplamıdır: $7/6+7/3+\cdots+5/6+1=35\ (=5\cdot 7)$.

<u>Gözlem:</u> $P=E^T+\left[\begin{array}{c|c}0&0\\ \hline 0&I\end{array}\right]$. Ayrıca $E^n\to 0$ olduğu için $Q^n\to 0$ olduğu da görülür.

<u>Örnek:</u> 

Diğer bütün olasılıklar $1/2$ olsun.

1 ve 5 yutan durumlardır. Bu durumda

$$E=\begin{bmatrix}0&1/2&0&0&0\\2/3&0&1/2&0&0\\0&1/2&0&0&0\\1/3&0&0&0&0\\0&0&1/2&0&0\end{bmatrix}$$ olur.

<u>Alıştırma 39.</u> Üç bölgeye ayrılmış bir şehrin bölgeleri günlük olarak, sırasıyla, u_1, u_2 ve u_3 miktarında kirlilik üretmektedir. q_{ij} ile bir gün içinde i'inci bölgeden j'inci bölgeye taşınan kirlilik miktarının oranını gösterelim. $q_i=1-\sum_j q_{ij}>0$ ise bu bölgeden atmosfere karışan kirlilik miktarının oranı olsun. $w_i^{(n)}$ ise i'inci bölgede n gün içinde biriken kirlilik miktarını göstersin. Aşağıdaki ifadeleri kanıtlayınız:

a) $w^{(n)}=[w_1\ w_2\ w_3]=u+uQ+\cdots+uQ^{n-1}$
b) $w^{(n)}$ bir w değerine yakınsaktır: $w^{(n)}\to w$.
c) w verildiğinde u'nun nasıl belirlendiğini açıklayınız.

<u>Alıştırma 40.</u> (Kumarbazın Felaket Sonu)
Bir kumar makinesinde, her seferinde, oyuncu p olasılıkla kazanırken $1-p$ olasılıkla kaybetmektedir. Her oyun diğerlerinden bağımsızdır. Kumarbazın tüm parasını kaybetmeden önce tüm parasının verilen bir N değerine ulaşma olasılığını hesaplayınız.

<u>Çözüm:</u> Bu problem için aşağıdaki modeli düşünelim:

$p+q=1$.

Geçiş matrisi $P=\begin{pmatrix}1&0&\cdots&\cdots&0\\q&0&p&0\cdots&0\\&\ddots&\ddots&\ddots&\vdots\\&&q&0&p\\0&0&\cdots&0&1\end{pmatrix}$ olsun. $P_\infty=\lim\limits_{n\to\infty}P^n$ matrisi

$$P_\infty = \begin{bmatrix} 1-P_0 & 0 & \cdots & 0 & P_0 \\ 1-P_1 & \vdots & & \vdots & P_1 \\ \vdots & & & & \vdots \\ 1-P_N & 0 & \cdots & 0 & P_n \end{bmatrix}$$ olur. Burada $P_0=0$, $P_N=1$ dir. Ayrıca $P_1 = pP_2$, $P_2 = qP_1 + pP_3$, $P_i = qP_{i-1} + pP_{i+1}, \ldots, P_{N-2} = qP_{n-3} + pP_{N-1}$ ve $P_{N-1} = qP_{N-2} + pP_N$ olur.

Hesabı kolaylaştırmak için $p=q=1/2$ alalım. Başka bir deyişle makine hilesiz olsun. Yine $N=5$ olsun.

Buradan $P_2 = 2P_1$, $P_2 = \frac{1}{2}P_1 + \frac{1}{2}P_3 \Rightarrow P_3 = 2P_2 - P_1 = 3P_1$, $P_4 = 4P_1$ ve $P_5 = 5P_1$ çıkar.

Genel durumda ($p=q=1/2$), $P_i = i/N$ olur. Dolayısıyla, kumarbazın bütçesi 1 birim ve kasa N birim ise kumarbazın tüm parayı kazanma olasılığı $1/N$ olur. Başka bir deyişle kumarbazın felaketi kaçınılmazdır.

Bu alıştırmayı genel p için tekrarlayınız.

§2.3. Ergodic Markov Chains

Yutan (durumlu) zincirlerin aksine ergodik zincirler belli bir duruma yakınlaşmazlar, bunun yerine sürekli bir durumdan diğerine geçerler.

Bir T matrisi için her $T_{ij} \geq 0$ ise $T \geq 0$ ve her $T_{ij} > 0$ ise $T > 0$ yazacağız.

Tanım: P bir Markov zincirinin geçişme matrisi olsun.
1) Eğer bir $n \in \mathbb{Z}$ için $P^n > 0$ ise zincire ilkel zincir denir.
2) Her i,j durumları için $p_{ij}(n(i,j)) > 0$ olacak şekilde bir $n(i,j) \in \mathbb{Z}$ sayısı varsa zincire indirgenemez zincir denir.

Açık bir şekilde her ilkel zincir indirgenemezdir.

Alıştırma 41. $P = \begin{pmatrix} 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{pmatrix}$ olsun (Kutulardaki toplar modeli).
$P^2 > 0$ olur. Dolayısıyla, bu zincir ilkeldir.

Alıştırma 42. $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ olsun. P ilkel değildir ama indirgenemezdir.

Teorem (Perron-Frobenius)

P indirgenemez $n\times n$-geçişme matrisi olsun. O halde, bir ve sadece bir $w \in \mathbb{R}^n$, $w>0$ vektörü vardır öyle ki, $w^T P = w^T$ ve $e=(1,1,\ldots,1)^T$ olmak üzere $k\to\infty$ sonsuza giderken $P^k \to e w^T$ olur.

Kanıt: İlk önce her $y\in\mathbb{R}^n$ vektörü için $P^k y \to e w^T y$ olduğunu göstereceğiz. Kolaylık olması açısından önce $P>0$ olduğunu kabul edelim. O halde, $d=\min p_{ij}>0$ olur. Bu durumda, $d\leq 1/2$'dir. Her $y\in\mathbb{R}^n$ vektörü için $m_0=\min\limits_j y_j$, $M_0=\max\limits_j y_j$, $m_1=\min\limits_j (Py)_j$ ve

$M_1=\max\limits_j (Py)_j$ olsun. $(Py)_i=\sum\limits_j p_{ij}y_j$ elemanını ele alalım.

O halde, $M_1=\max\limits_i (Py)_i = \max\limits_i \sum\limits_j p_{ij}y_j \leq (1-d)M_0+dm_0$.

(Bu eşitsizliği görmek için şunu gözlemleyin: y vektörünün en küçük bileşeni p_{ij}'lerin en küçüğü "d" ile, ve en büyük bileşeni M_0 da $(1-d)$ ile çarpılıyor.)

Benzer bir sebeble $m_1=\min\limits_i (Py)_i \geq (1-d)m_0+dM_0$ olur. O halde

$$M_1-m_1 \leq (1-d)M_0+dm_0-(1-d)m_0-dM_0=(1-2d)M_0+(2d-1)m_0$$

$$\Rightarrow M_1-m_1\leq(1-2d)(M_0-m_0) \text{ elde edilir.}$$

y vektörünü tekrar P ile çarpıp bu işlemi k defa tekrarlarsak

$$M_k-m_k\leq(1-2d)^k(M_0-m_0), \text{ öyle ki}$$

$M_k=\max\limits_i (P^k y)_i$ ve $m_k=\min\limits_i (P^k y)_i$ olarak tanımlanmıştır.

(M_k) dizisi azalandır:

$$M_{k+1}=\max_i (PP^k y)_i=\max_i \sum_j p_{ij}(P^k y)_j \leq \max_i \sum_j p_{ij} M_k = \max_i M_k = M_k.$$

Benzer şekilde (m_k) dizisi de artandır. Dolayısıyla, ikisi de yakınsaktır. $0\leq(1-2d)<1$ olduğu için $\lim\limits_k M_k=\lim\limits_k m_k = m$ olmalıdır. Dolayısıyla, $P^k y\to me$ $(k\to\infty)$ olur.

Şimdi m sayısının pozitif olduğunu görelim: İlk önce $m(e^Te) = e^Tme = e^T \lim_k P^k y = \lim_k e^T P^k y$ olduğunu görelim.

Ayrıca, $e^T P^k y \geq e^T(m_k e) = m_k(e^Te) \geq m_1(e^Te)$ olur. Buradaki ilk eşitsizlik $m_k = \min_i (P^k y)_i$ tanımının, ikincisi ise (m_k) dizisinin artan olmasının sonucudur.

O halde, $e^T P^k y \geq m_1(e^Te) = \min_i (Py)_i (e^Te)$ olur.

Şimdi $P>0$ olarak kabul edildiği için, eğer her $y_i > 0$ ise $(Py)_i > 0$, $\forall i$ olur ve dolayısıyla $m>0$ olmalıdır.

Son olarak $w_j = \lim_k P^k e_j / (e^Te)$, $e_j = [0 \cdots 1 \cdots 0]^T$, ile w_j sayısını tanımlayalım. Yukarıda yaptıklarımızdan dolayı $w_j > 0$ ve $P^k \to ew^T$, $w^T = [w_1 \; w_2 \cdots w_n]$, olur. Buradan

$$\lim_{k\to\infty} P^k P = ew^T P$$ ve dolayısıyla $ew^T = ew^T P \Rightarrow w^T = w^T P$ elde edilir.

Böylece $P>0$ durumda kanıt tamamlanmış oldu.

Şimdi P'nin ilkel olduğunu kabul edelim ve bir $N>0$ tam sayısı için $P^N > 0$ olsun. O halde, yukarıda yaptığımızdan dolayı $P^{kN} \to ew^T$ olacak şekilde tek bir $w \in \mathbb{R}^n$, $w>0$ vektörü vardır. $P^N>0$ olduğu için $P^{N+1}>0$ olur. O halde bir $v>0$ vektörü için $P^{k(N+1)} \to ev^T$ olur. O halde, $P^{kN(N+1)}$ hem ew^T hem de ev^T vektörüne yakınsar ve dolayısıyla, $w=v$ elde edilir. O halde, $w^T P^{N+1} = v^T P^{N+1} = v^T = w^T = w^T P^N$ ve buradan da $(w^T P) P^N = w^T P^N$ elde edilir. Bu ise $w^T P = w^T$ sonucunu verir.

Şu an her $y \in \mathbb{R}^n$ için $P^{kN} y \to ew^T y$ olduğunu biliyoruz ve $P^m y \to ew^T y$ olduğunu göstermek istiyoruz. O halde, verilen herhangi bir $\forall \epsilon > 0$ için öyle bir $K \in \mathbb{N}$ sayısı vardır ki her $k \geq K$ ve $y \in \mathbb{R}^n$ için

$$\|(P^{kN} - ew^T)y\| < \epsilon$$ sağlanır. $m = kN + J$, $J < N$ olsun.

O zaman,
$$\|(P^m - ew^T)y\| = \|(P^{kN+J} - ew^T)y\| = \|(P^{kN} - ew^T)P^J y\| \leq \epsilon$$ olur.

Burada, $w^T P = w^T$ eşitliğini J-kere kullandık. Böylece kanıt biter. ■

Hatırlatma: Yukarıdaki teoremin bir sonucu olarak w vektörü $w^T P = w^T$ denkleminin tek çözümüdür.

Ayrıca her olasılık vektörü v için $(v = (v_i), \sum v_i = 1)$

$$\lim_n v^T P^n = v^T \lim_n P^n = v^T e w^T = (v^T e) w = (\sum v_i) w = w$$

olur.

Başka bir deyişle her başlangıç (olasılık) vektörü zaman ilerledikçe w dağılımına yakınsar.

Alıştırma 43. Tekrar kutulardaki toplar modeline bakalım. P matrisinin

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{pmatrix}$$

olduğunu biliyoruz. P matrisinin ilkel olduğunu görmüştük. $w^T P = w^T$ denklemini çözerek $w^T = (1/6, 2/3, 1/6)$ olduğunu görürüz.

Bu teoremin sonucu olarak, $P^n \to e w^T$ sonucu P'nin tüm özdeğerlerinin mutlak değerce ≤ 1 ve 1 özdeğerinin tek öz vektörünün de w olduğunu görmüş olduk.

Teoremin Yorumu: Diyelim ki X_0 başlangıç durumunun olasılık dağılımı $P(X_0 = i) = \alpha_i$, $i \in S$ olsun. O zaman

$$P(X_k = j) = \sum_i P(X_k = j \mid X_0 = i)\, P(X_0 = i)$$

$$= \sum_i (P^k)_{ij} \alpha_i = (\alpha^T P^k)_j \text{ olur.}$$

Teoreme göre $P(X_j = j) = (\alpha^T P^k)_j \to w_j$, $k \to \infty$, olduğu için şunu söyleyebiliriz. Başlangıç olasılık dağılımı ne olursa olsun zincir hızla $w^T = (w_1, \ldots, w_n)$ denge veya kalıcı olasılık dağılımına yakınsar.

Bu teorem indirgenemez zincirler için aşağıdaki zayıf versiyonuna dönüşür:

Teorem 8. P indirgenemez bir Markov zincirinin geçişme matrisi olsun. Bu durumda, $w^T P = w^T$ eşitliğini sağlayan tek bir pozitif $w > 0$ olasılık dağılım vektörü vardır.

Ayrıca, $\frac{1}{n+1}(I+P+\cdots+P^n) \to ew^T$, $n\to\infty$, yakınsar.

<u>Kanıt:</u> $Q \doteq \frac{1}{2}(I+P)$ olarak tanımlansın. Q matrisi de bir geçişme matrisidir. Ayrıca,

$$2^n Q^n = (I+P)^n = \sum_{k=0}^{n} \binom{n}{k} P^k$$ olur. Zincir indirgenemez olduğu için verilen her i,j için bir $n(i,j)\in\mathbb{N}$ vardır öyle ki $(P^{n(i,j)})_{ij} > 0$. Her $n = \max_{i,j} n(i,j)$ olsun.

O zaman, $$2^n (Q^n)_{ij} = \sum_{k=0}^{n} \binom{n}{k} (P^k)_{ij}$$
$$\geq \binom{n}{n(i,j)} (P^{n(i,j)})_{ij} > 0$$ olur.

O halde, Q ilkel bir matristir. w ile Q matrisinin yegane denge olasılık dağılım vektörünü gösterelim. Bu durumda,

$$w^T Q = w^T \Leftrightarrow w^T \frac{1}{2}(I+P) = w^T \Leftrightarrow w^T P = w^T$$
olur.

Dolayısıyla, P matrisi için w vektörünün varlığı ve tekliği gösterilmiş oldu.

Şimdi $W = ew^T$ matrisini tanımlayalım. O zaman

$$(I-P+W)(I+P+\cdots+P^{n-1}) = I - P^n + \sum_{k=0}^{n-1} W P^k$$
$$= I - P^n + \sum_{k=0}^{n-1} e w^T P^k$$
$$= I - P^n + \sum_{k=0}^{n-1} e w^T$$
$$= I - P^n + nW,$$ elde edilir.

<u>İddia:</u> $I-P+W$ matrisinin tersi vardır.

<u>Kanıt:</u> $y^T(I-P+W) = 0$ olduğunu kabul edelim. Buradan

$y^T - y^T P + (y^T e) w^T = 0$ elde edilir. Bu denklemi e ile çarparsak ve $Pe = e$ olduğunu kullanırsak

$$y^T e - y^T \underbrace{Pe}_{e} + (y^T e) w^T e = 0 \Rightarrow (y^T e) w^T e = 0$$ elde ederiz.

$w^Te = 1 > 0$ olduğu için $y^Te = 0$ ve buradan da $y^T - y^TP = 0$ buluruz. w vektörünün tekliğinden dolayı $y = \lambda w$ elde ederiz. Son olarak

$$0 = y^Te = \lambda w^Te = \lambda \Rightarrow \lambda = 0 \Rightarrow y = 0 \cdot w = 0 \text{ olur.}$$

O halde, $I-P+W$ matrisinin tersi vardır. ▪

$I-P+W$ matrisinin tersi olduğuna göre

$$(I-P+W)(I+P+\cdots+P^{n-1}) = I-P^n+nW \text{ eşitliğinden}$$

$$I+P+\cdots+P^{n-1} = (I-P+W)^{-1}(I-P^n+nW) \text{ buluruz.}$$

Diğer taraftan, $WP = ew^TP = ew^T = W$ ve

$W^2 = ew^Tew^T = e \cdot 1 \cdot w^T = ew^T = W$'dur. O halde,

$$W(I-P+W) = W - WP + W^2 = W - W + W = W \text{ olur.}$$

Buradan da

$$W = W(I-P+W) \Rightarrow (I-P+W)^{-1}W = W \text{ elde edilir.}$$

O halde,
$$\begin{aligned} I+P+\cdots+P^{n-1} &= (I-P+W)^{-1}(I-P^n+nW) \\ &= (I-P+W)^{-1}(I-P^n) + n(I-P+W)^{-1}W \\ &= (I-P+W)^{-1}(I-P^n) + nW \end{aligned}$$

$$\Rightarrow \frac{1}{n}(I+P+\cdots+P^{n-1}) = W + \frac{1}{n}(I-P+W)^{-1}(I-P^n)$$

elde edilir. Son olarak,

$$\|P^nz\|_1 = \sum_{i,j}(P^n)_{ij}|z_j| = \sum_j \underbrace{\left(\sum_i (P^n)_{ij}\right)}_{\leq 1}|z_j|$$

$$\leq \sum_j |z_j| \text{ elde edilir.}$$

Dolayısıyla P^n ve $I-P^n$ matrisleri uniform olarak sınırlıdır. O halde, $\frac{1}{n}(I-P^n) \to 0$, $n\to\infty$.

Bu kanıtı bitirir. ▪

Alıştırma 44. Bir P geçişme matrisinin her sütununun toplamı da 1 oluyorsa P'ye ikili periyodik denir. Böyle bir matris aynı zamanda indirgenemez ise bu matrisin denge dağılım vektörünü hesaplayınız.

Çözüm: $w^T = \frac{1}{n}[1\ 1 \cdots 1]$ olsun. Bu durumda

$$(w^T P)_i = \sum_{j=1}^{n} w_j^T P_{ji} = \sum_{j=1}^{n} \frac{1}{n} P_{ji} = \frac{1}{n}$$ ve dolayısıyla

$w^T P = w^T$ elde edilir. Matris indirgenemez olduğu için bunu sağlayan w^T tektir.

Alıştırma 45. Bir deneyde peş peşe deneme yapılıyor. Eğer son iki deneme başarılı ise bir sonrakinin başarılı olma ihtimali 0.8, aksi halde 0.5 dir. Bu deneyde bir denemenin başarılı olma olasılığı kaçtır?

Çözüm: Üç farklı durum tanımlayalım:

$S_1 = \{$En son deneme başarısızdır$\}$

$S_2 = \{$En son deneme başarılı, fakat bir önceki başarısızdır.$\}$

$S_3 = \{$Son iki deneme de başarısızdır.$\}$

X_n ile n'inci denemenin durumunu gösterelim.

$P = (P_{ij})$ geçişme matrisi aşağıdaki gibidir.

$P_{11} = 0.5$, $P_{12} = 0.5$, $P_{13} = 0$
$P_{21} = 0.5$, $P_{22} = 0$, $P_{23} = 0.5$
$P_{31} = 0.2$, $P_{32} = 0$, $P_{33} = 0.8$

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0.2 & 0 & 0.8 \end{bmatrix}$$

Bir önceki teoremden w_j sayısının değişkenin j durumuna gelme olasılığı olduğunu biliyoruz. O halde $w^T = w^T P$ denklemini çözersek $w^T = [4/11, 2/11, 5/11]$ elde ederiz. O halde, bizim aradığımız cevap $1 - 4/11 = 7/11$ olur.

Alıştırma 46. (X_n) geçişme matrisi P ve kalıcı olasılık dağılımı w olan sonlu durumlu ilkel bir Markov zinciri olsun. (Y_n), $Y_n=(X_n,X_{n-1})$ şeklinde yeni bir zincir tanımlayalım. Bunun da bir Markov zinciri olduğunu gösterin ve $\lim\limits_{n\to\infty} P(Y_n=(i,j))$ limitini hesaplayınız.

Çözüm: $P(Y_n=(i,j)\mid Y_1=(i_1,j_1),\dots,Y_{n-1}=(i_{n-1},j_{n-1}))$

$$=P(X_n=j,X_{n-1}=i\mid X_0=i_1, X_1=j_1,\dots,X_{n-2}=i_{n-1},X_{n-1}=j_{n-1})$$

olur.

Şimdi, $P(A\cap B\mid C)=P(A\mid B\cap C)P(B\mid C)$ eşitliğini kullanarak

$$=P(X_n=j\mid X_{n-1}=i, X_0=i_1, X_1=j_1, X_2=i_2, X_3=j_2,\dots, X_{n-2}=i_{n-1}, X_{n-1}=j_{n-1})$$
$$\cdot\, P(X_{n-1}=i\mid X_0=i, X_1=j, X_1=i_2, X_2=j_2,\dots, X_{n-2}=i_{n-1}, X_{n-1}=j_{n-1})$$

$$= P(X_n=j\mid X_{n-1}=i=j_{n-1})\, P(X_{n-1}=i\mid X_{n-1}=j_{n-1}, X_{n-2}=i_{n-1})$$
$$= P(X_n=j\mid X_{n-1}=i)\cdot\delta_{ij_{n-1}}\cdot P(X_{n-1}=i\mid X_{n-2}=i_{n-1})$$
$$= P(X_n=j, X_{n-1}=i\mid X_{n-1}=i, X_{n-2}=i_{n-1})\,\delta_{ij_{n-1}}$$
$$= P(Y_n=(i,j)\mid Y_{n-1}=(i_{n-1},j_{n-1}))\,\delta_{ij_{n-1}}.$$

Dolayısıyla, (Y_n) bir Markov zinciridir. Ayrıca bu hesaplamalardan

$$P_{(i_1,j_1),(i_2,j_2)}=\delta_{j_1i_2}P_{i_2j_2}P_{i_1j_1}=\begin{cases}0, & j_1\neq i_2\\ P_{i_1j_1}P_{i_2j_2}, & j_1=i_2,\end{cases}$$

elde edilir.

Diğer taraftan, $\lim\limits_{n\to\infty}P(Y_n=(i,j))=\lim\limits_{n}P(X_n=j, X_{n-1}=i)$

$$=\lim_n P(X_n=j\mid X_{n-1}=i)\,P(X_{n-1}=i)$$
$$=\lim_n P_{ij}\,P(X_{n-1}=i)$$
$$=P_{ij}\lim_n P(X_{n-1}=i)$$
$$=P_{ij}\,w_i,$$

olur.

Tanım: (X_n) indirgenemez bir Markov zinciri olsun.

1) i'nin durumundan başlayarak j'inci duruma ilk defa ulaşmak için beklenen adım sayısını m_{ij} ile gösterelim ($m_{ii}=0$ olarak tanımlanır).

2) r_i ise i'inci durumdan başlayıp tekrar bu duruma dönmek için geçen adım sayısının beklenen değeri olsun.

3) $Z=(I-P+W)^{-1}$ matrisine zincirin temel matrisi denir.

<u>Teorem:</u> w verilen indirgenemez bir Markov zincirinin kalıcı olasılık dağılımı olsun. Bu durumda her $i,j \in S$ durumları için $r_i = 1/w_i$, $m_{ij} = (z_{jj}-z_{ij})/w_j$, $Z=(z_{ij})$ olur.

<u>Kanıt:</u> $M=(m_{ij})$, $E=(E_{ij})$, $E_{ij}=1, \forall i,j$ ve $D=(D_{ij})$, $D_{ij}=\delta_{ij} r_i$, $\forall i$, olsun. Bu durumda her $i \neq j$ için

$$m_{ij} = 1 \cdot p_{ij} + \sum_{k \neq j} p_{ik}(m_{kj}+1) = \sum_{k=1}^{n} p_{ik} + \sum_{k \neq j} p_{ik} m_{kj}, \text{ olur.}$$

Benzer şekilde, $r_i = \sum_k p_{ik}(m_{ik}+1) = 1 + \sum_k p_{ik} m_{ki}$, olur.

O halde, her i,j için, $m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{ki} - D_{ij}$ ve dolayısıyla,

$M=(m_{ij}) = E + PM - D$ elde edilir. Şimdi bu eşitliği w^T ile çarpalım:

$w^T M = w^T E + w^T P M - w^T D$. $w^T P = w^T$ olduğu için $0 = w^T E - w^T D$ ve buradan da $0 = (\sum_i w_i) - w_i r_i$
$\Rightarrow w_i r_i = \sum w_i = 1$ elde edilir. O halde, $r_i = 1/w_i$ olmalıdır.

$e_i = (1,\ldots,1)^T$ olsun. $Pe = e = We$ olur çünkü $W = ew^T$ olarak tanımlanmıştı ve $w^T e = 1$'dir. O halde,
$(I-P+W)e = e - Pe + We = e$ ve dolayısıyla $Ze = e$ olur. Buradan, $ZE = E = ee^T$ elde edilir. Ayrıca,
$w^T P = w^T = w^T W$ ve benzer şekilde, $w^T(I-P+W) = w^T$ olduğu açıktır ve dolayısıyla, $w^T Z = w^T$ olur. O halde,
$ZW = Zew^T$ ve $ZW = Zew^T = ew^T = W$ olur. Şimdi $M = E + PM - D$ olduğu için,
$Z(I-P)M = ZM - ZPM = ZM - Z(M-E+D) = ZE - ZD$ olur.

Ayrıca, $Z(I-P) = Z(I-P+W) - ZW$
$= I - ZW \quad (Z = (I-P+W)^{-1}!)$
$= I - W$ ve dolayısıyla,

$E - ZD = Z(I-P)M = (I-W)M = M - WM$ ve sonuç olarak $M = E - ZD + WM$ olur. O halde, ij- bileşini

$$m_{ij} = 1 - z_{ij} r_j + (ew^T M)_{ij}$$
$$= 1 - z_{ij} r_j + (w^T M)_j \text{, olur.}$$

$i=j$ alırsak, $0 = m_{jj} = 1 - z_{jj} r_j + (w^T M)_j$ ve dolayısıyla,

$(w^T M)_j = z_{jj} r_j - 1$ elde edilir. Son olarak,

$$m_{ij} = 1 - z_{ij} r_j + z_{jj} r_j - 1 = (z_{jj} - z_{ij}) r_j = \frac{z_{jj} - z_{ij}}{w_j} \text{ olur.}$$

Böylece kanıt biter.

§ 2.4. Classification of Finite-State Markov Chains

Durum uzayının i ve j gibi iki elemanı için $p_{ij}(m) > 0$ ve $p_{ji}(n) > 0$ olacak şekilde m ve n (≥ 0) tam sayıları varsa bu iki duruma birbiriyle konuşuyorlar denir. Bir $C \subseteq S$ alt kümesinde her $i, j \in C$ eleman çifti birbiriyle konuşuyorsa C kümesine indirgenemez denir. O halde indirgenemez bir zincirde S durum uzayının kendisi indirgenemezdir.

Teorem: Bir Markov zincirinin S durum uzayı sonlu ise S kümesi tek bir şekilde $S = T \cup C_1 \cup \cdots \cup C_n$ şeklinde ayrık kümelerin birleşimi olarak yazılabilir, öyle ki, T tüm geçiş durumlarının kümesidir ve her bir C_i kapalı (C_i'in kendisi de bir zincirdir) ve indirgenemez dir. Bu bileşenlere zincirin sınıfları denir.

Kanıt sonra verilecek.

Alıştırma 47.
$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/4 & 3/4 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

Bu zincirin tüm sınıflarını bulunuz.

Bölüm 3. Markov Zincirlerinin Varlığı.

Bu bölümde $\lim_{n\to\infty} P(X_n = j)$ limiti ne anlama geliyor gibi soruları cevaplamaya çalışacağız.

§ 3.1. Bir Markov Zincirinin Örneklem Uzayı:

S bir Markov zincirinin sayılabilir durum uzayı olsun. X_0 başlangıç olasılık dağılımı olmak üzere $\alpha_i = P(X_0 = i)$ ve $P(X_{n+1} = j \mid X_n = i) = p_{ij}$ olsun $(i, j \in S)$, $n \geq 1$.

Bu durumda

$$P(X_2 = k, X_1 = j, X_0 = i) = P(X_2 = k \mid X_1 = j, X_0 = i)\, P(X_1 = j \mid X_0 = i)\, P(X_0 = i)$$
$$= p_{jk}\, p_{ij}\, \alpha_i \text{ olur.}$$

Daha genel halde ise,

$$P(X_0 = i, X_1 = i_1, \ldots, X_n = i_n) = \alpha_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \text{ olur.}$$

Amacımız yukarıdaki zincir yapısını veren bir Ω örneklem uzayı ve P olasılık fonksiyonunu oluşturmak. Bunu şöyle oluşturabiliriz:

a) $\Omega = [0,1]$ ve P uzunluk ölçümü olsun. $[0,1]$ aralığını $|S|$ alt aralığa bölelim $\{I_i^{(0)}\}$ öyle ki $P(I_i^{(0)}) = \alpha_i$, $\forall i$, olsun. Bu durumda $\sum \alpha_i = 1$ olacaktır.

b) Daha sonra her bir $I_i^{(0)}$ aralığını $I_{ij}^{(1)}$, $j \in S$, aralıklarına bölelim öyle ki $P(I_{ij}^{(1)}) = \alpha_i p_{ij}$ olsun. $\sum_j p_{ij} = 1$ olduğu için bu mümkündür.

c) Bu işleme devam ederek her bir $I_{i_0 i_1 \ldots i_{n-1}}^{(n-1)}$ aralığını $I_{i_0 i_1 \ldots i_n}^{(n)}$ aralıklarına bölelim öyle ki $P(I_{i_0 \ldots i_n}^{(n)}) = \alpha_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$ olsun.

d) X_n değişkeni, $x \in I_{i_0 i_1 \ldots i_n}^{(n)}$ ise $P(x) = i_n$ olarak tanımlansın.

<u>İddia:</u> X_n istenilen birleşik dağılım fonksiyonuna sahiptir.

<u>Kanıt:</u> $P(X_0 = i) = P(x \in I_i^{(0)}) = \alpha_i$. Diğer yandan,

$P(X_1 = j, X_0 = i) = P(x \in I_{ij}^{(1)}) = p_{ij}\alpha_i$. Genel durumda ise

$P(X_0=i_0,\dots,X_n=i_n) = P(x \in I^{(n)}_{i_0 i_1 \dots i_n}) = \alpha_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$ olur.

Ayrıca, $x \in I^{(n)}_{i_0 i_1 \dots i_n}$ olduğu için $X_0(x)=i_0,\ X_1(x)=i_1,\dots,$ $X_{n-1}(x)=i_{n-1}$'dir.

Böylece kanıt tamamlanır. ■

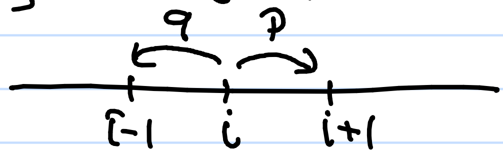
§3.2. Lebesgue Ölçümü: Zincirlerin varlığı ile daha sonra 5. Bölüm'de ele alınacak. Bunu yapabilmek için aralıkların boyunu veren ölçümün $[0,1]$ aralığının daha genel altkümelerine genişletilmesi gerekiyor. Söz konusu bir ölçüm Lebesgue Ölçümü diyeceğiz.

Bölüm 4. Ayrık Zaman Sayılabilir Sonsuz Durumlu Markov Zincirleri:

S durum uzayı sayılabilir $S=\{x_1,x_2,\dots,x_n,\dots\}$ kümesi olsun. $p_{ij}=P(X_n=j \mid X_{n-1}=i)$ olmak üzere yine $\sum_j p_{ij}=1$ koşulunu sağlayacaktır.

§4.1. Örnekler 1) Bir boyutlu rasgele yürüyüş. Durum uzayı $S=\mathbb{Z}=\{\dots,-1,0,1,2,\dots\}$ ve geçiş olasılıkları

$$p_{ij}=\begin{cases} p, & j=i+1 \\ q, & j=i-1 \\ 0, & \text{aksi halde} \end{cases} \qquad (p+q=1)$$



2) Bu bir nüfus artışı modeli olacak. X_n ile n'inci neslin nüfusunu gösterelim. B_i ise i'inci nesildeki bireylerin bir sonraki nesle aktardıkları bireylerin sayısı olsun (kendisi dahil). O halde,

$$X_{n+1}=\sum_{i=1}^{X_n} B_i$$ olur.

Burada anlamlı soru bu nüfusun ne kadar süre var olacağıdır.

§4.2. Durumların Sınıflandırılması: Herhangi i,j durumu için

$f_{ij}(n) = P(X_1 \neq j, \ldots, X_{n-1} \neq j, X_n = j \mid X_0 = i)$ olasılığı olsun. Bu i'inci durumda başlayan bir zincirin n adım sonra ilk defa j'inci duruma ulaşma olasılığını göstermektedir. O halde, $f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$ ise i'inci durumdan başlayan bir zincirin j'inci duruma ulaşma olasılığıdır.

Tanım. Bir j durumu için $f_{jj} = 1$ ise bu duruma kalıcıdır denir. $f_{jj} < 1$ ise j'ye geçici durum denir.

Tanım. j'inci duruma ortalama dönüş zamanı

$$\mu_j = \begin{cases} \sum_{n=1}^{\infty} n f_{jj}(n) & , j \text{ kalıcıysa,} \\ 0 & , j \text{ geçiciyse.} \end{cases}$$

μ_j zamanı sonsuz çıkabilir. Daha önceki bölümlerde μ_j ifadesini r_j ile göstermiştik.

Tanım. Kalıcı bir j durumu için $\mu_j = \infty$ ise ona görünmez kalıcı durum, aksi halde ise $(\mu_j < \infty)$ görünür kalıcı durum denir.

O halde sayılabilir sonsuz durumlu Markov zincirlerinde üç çeşit durum vardır: Geçici durum, görünür kalıcı durum ve görünmez kalıcı durum.

Alıştırma 48. Aşağıdaki üreteç fonksiyonlarını tanımlayalım:

$$P_{ij}(s) = \sum_{n=0}^{\infty} s^n P_{ij}(n) \quad \text{ve} \quad F_{ij}(s) = \sum_{n=0}^{\infty} s^n f_{ij}(n), \text{ öyle ki}$$

$p_{ij}(0) = \delta_{ij}$ ve $f_{ij}(0) = 0$ olarak tanımlansın.

Bu durumda $P_{ii}(s) = 1 + F_{ii}(s) P_{ii}(s)$ olduğunu gösteriniz.

Çözüm: İlk önce Abel'in şu teoremini hatırlayalım: $a_n \geq 0$, $\forall n$, ve her $0 \leq s < 1$ için $\sum a_n s^n < +\infty$ olsun. Bu durumda

$$\lim_{s \to 1^+} \sum_{n=0}^{\infty} a_n s^n = \sum_{n=0}^{\infty} a_n$$ olur. Başka bir deyişle

$s \mapsto \sum_{n=0}^{\infty} a_n s^n$ fonksiyonu $[-1, 1]$ aralığında süreklidir.

İlk önce $F_{ii}(s)\,P_{ii}(s)$ çarpımını hesaplayalım:

$$F_{ii}(s)\,P_{ii}(s) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} s^k p_{ii}(k)\, s^{n-k} f_{ii}(n-k)$$
$$= \sum_{n=0}^{\infty} s^n \Big(\sum_{k=0}^{n} p_{ii}(k)\, f_{ii}(n-k)\Big).$$

O halde, $p_{ii}(n) = \sum_{k=0}^{n} p_{ii}(k)\, f_{ii}(n-k)$ olduğunu göstermeliyiz.

Bunu şu şekilde görebiliriz: i'den i'ye giden her n adımlık yolculuk için tek bir $n-k$'ıncı adım vardır öyle ki bu başlangıçtan sonra i'ye ilk defa gelme adım sayısıdır. Ve bunun olasılığı $f_{ii}(n-k)$'dır. Ayrıca bu adımdan k adım sonra tekrar i'ye gelme olasılığı $p_{ii}(k)$'dır (toplamda, $(n-k)+k=n$ adım). Bu kanıtı bitirir.

<u>Alıştırma 48'</u>: Bir i durumu kalıcıdır ancak ve ancak $\sum_n p_{ii}(n) < \infty$ toplamı sonludur.

<u>Çözüm</u>: İlk önce i durumunun kalıcı olduğunu, başka bir deyişle $f_{ii}=1$ olduğunu kabul edelim.

Her $f_{ij}(n) \le 1$ ve $p_{ij}(n) \le 1$ olduğu için her iki fonksiyon serisi de $(-1,1)$ aralığında yakınsaktır. Bizim $\sum p_{ii}(n) = +\infty$ olduğunu göstermeliyiz. Bunun yerine $\sum p_{ii}(n)$ toplamının sonlu olduğunu kabul edelim. O halde, $P_{ii}(s)$ fonksiyonu $[-1,1]$ aralığında süreklidir (Abel Teoremi).

$P_{ii}(s) = 1 + F_{ii}(s)\,P_{ii}(s)$, $s\in(-1,1)$, eşitliği $P_{ii}(s)\ge 1$ $\forall s\in(-1,1)$ ve $P_{ii}(s)$ sürekli olduğu için de $P_{ii}(s)\ge 1, \forall s\in[-1,1]$, elde edilir. O halde,

$$F_{ii}(s) = \frac{P_{ii}(s)-1}{P_{ii}(s)}, \quad s\in[-1,1], \text{ yazabiliriz.}$$

Ayrıca, $F_{ii}(s)$ fonksiyonunun da $[-1,1]$ aralığında sürekli olduğunu görmüş olduk. O halde, $f_{ii}=1$ olduğunu kullanarak şu çelişkiye varabiliriz:

$$1 = f_{ii} = \sum_{n=0}^{\infty} f_{ii}(n) = F_{ii}(1) = \frac{P_{ii}(1)-1}{P_{ii}(1)} \Rightarrow 0 = -1 \quad \rightarrow\leftarrow.$$

O halde, $\sum_{n=0}^{\infty} p_{ii}(n) = +\infty$ olmalıdır.

Şimdi de $\sum_n p_{ii}(n) = +\infty$ olduğunu kabul edelim. O halde, $\lim_{s\to 1^+} P_{ii}(s) = +\infty$ olur. Şimdi

$$F_{ii}(s) = \frac{P_{ii}(s)-1}{P_{ii}(s)}, \quad s\in(-1,1),$$ eşitliğini kullanırsak $|F_{ii}(s)| \leq 1$, $\forall s\in(-1,1)$ elde ederiz. $F_{ii}(s) = \sum_{n=0}^{\infty} s^n f_{ii}(n)$ toplamının tüm terimleri pozitif olduğu için bu toplamın $[-1,1]$ aralığında da yakınsak olduğunu görürüz. Yine Abel Teoremi'nden dolayı $\lim_{s\to 1^+} F_{ii}(s)$ limiti vardır ve ≤ 1 dir.

Özel olarak F_{ii} fonksiyonu $[-1,1]$ aralığında süreklidir. O halde,

$$F_{ii}(1) = \lim_{s\to 1^+} F_{ii}(s) = \lim_{s\to 1^+} \frac{P_{ii}(s)-1}{P_{ii}(s)} = 1 \text{ olur}$$

($P_{ii}(s) \to +\infty$ olduğu için). Sonuç olarak

$f_{ii} = F_{ii}(1) = 1$ olur ve dolayısıyla i durumu kalıcı olur ve kanıt tamamlanır. ∎

§ 4.3. Sayılabilir Sonsuz Durumlu Ayrık Markov Zincirlerin Sınıflandırması:

Teorem: Eğer iki durum birbiriyle komunikasyonda ($p_{ij}(n)>0$ ve $p_{ji}(m)>0$ olacak şekilde m ve n varsa) bu iki durumun ikisi de geçiş, görünür kalıcı veya görünmez kalıcıdır.

Kanıt: $h = p_{ij}(m)\, p_{ji}(n) > 0$ olsun. O zaman her $r \geq 0$ tam sayısı için

$$p_{ii}(n+m+r) \geq p_{ij}(m)\, p_{jj}(r)\, p_{ji}(n) = h\, p_{jj}(r)$$ olur. Buradan

$$\sum_k p_{ii}(k) \geq \sum_r p_{ii}(n+m+r) \geq h \sum_r p_{jj}(r)$$ elde edilir. Sağ ve sol uçtaki toplamların ya ikisi birden sonludur ya da sonsuzdur. Dolayısıyla, ikisi birden geçiş durumudur ya da ikisi birden kalıcı durumdur.

Kanıtın geri kalanı alıştırma olarak okuyucuya bırakılmıştır. ∎

Alıştırma 49. $i \in S$ geçiş durumu bir $j \in S$ durumundan erişilebilir olsun. Bu durumda $p_{ij}(n) \to 0$ olduğunu gösteriniz.

Çözüm: Verilenlerden dolayı $p_{ji} > 0$ ve $\sum_{n=0}^{\infty} p_{ii}(n) < +\infty$.

Aynı zamanda, her $n \geq 1$ için $p_{ii}(n) \geq p_{ij}(n-1) p_{ji}(1) = p_{ji}\, p_{ij}(n-1)$ elde edilir.

O halde, $\dfrac{p_{ii}(n)}{p_{ji}} \geq p_{ij}(n-1) \geq 0$ olur.

Son olarak $\sum_{n=0}^{\infty} p_{ii}(n) < +\infty$ yakınsak olduğu için $p_{ii}(n) \to 0$ olduğu için $p_{ij}(n-1) \to 0$ elde edilir.

Tanımlar. Bir $C \subseteq S$ alt kümesinde her $i \in C$ ve $j \in S \setminus C$ için $p_{ij} = 0$ ise C kümesine kapalıdır denir çünkü C'deki hiç bir durumdan C'nin dışına çıkma olasılığı sıfırdır.

Sonlu durumlarda verdiğimiz sınıflandırma sonucu burada da geçerlidir.

Teorem. Verilen her sayılabilir durumlu Markov zinciri için S kümesi tek bir şekilde

$$S = T \cup C_1 \cup \cdots \cup C_i \cup \cdots$$

yazılabilir.

Burada T geçiş durumlarının kümesi ve her bir C_i kapalı ve indirgenemez kalıcı durumların kümesidir. Ayrıca, her bir C_i'deki durumlar ya hep görünen kalıcı yada görünmez kalıcıdır.

Kanıt: C_i'lerin kapalı olmaları dışındaki diğer iddialar açıktır. Bunlardan birinin kapalı olmadığını kabul edelim. O halde, bir $i \in C$ ve $j \notin C$ için $p_{ij} > 0$ olsun.

Buradan

$$1 - f_{ii} = P(X_n \neq i, \forall n \geq 1 \mid X_0 = i) \geq P(X_1 = j \mid X_0 = i) = p_{ij} > 0$$

ve dolayısıyla $f_{ii} < 1$ elde edilir. Bu ise i'nin kalıcı durum olmasıyla çelişir. Böylece kanıt tamamlanır.

Hatırlatma: Eğer (X_n) zinciri $x_0 \in C_i$ ile başlarsa hep C_i içinde kalır. $X_0 \in T$ olsa böyle er ya da geç sonunda bir C_i içine düşecek ve sonrasında da orada kalacaktır. Bundan sonra sadece indirgenemez zincirleri ele alacağız. $\pi = (\pi_1, \pi_2, \ldots)$, $\pi_i \geq 0, \forall i$, $\sum \pi_i = 1$, bir zincirin kalıcı olasılık dağılımı olsun. $\pi_j = \sum_i \pi_i p_{ij}$ olduğunu hatırlayalım.

Teorem: S geçiş olasılıkları p_{ij} olan bir indirgenemez zincir olsun.

1) Zincirin geçişken olması için $\sum_n p_{jj}(n) < \infty$ olacak şekilde bir $j \in S$ (ve dolayısıyla, $\forall j \in S$ için) olması gerek ve yeter koşuldur.

2) Benzer şekilde, zincirin kararlı olması için $\sum_n p_{jj}(n) = +\infty$ olacak şekilde bir $j \in S$ (ve dolayısıyla, $\forall j \in S$ için) olması gerek ve yeter koşuldur.

3) $x^T = x^T P$ ve $x_j = \sum_i x_i p_{ij}$, $\forall j$, koşulunu sağlayan bir $x = (x_i \geq 0)$ vektörü vardır. Zincir görünür kalıcıdır ancak ve ancak $\sum_i x_i < \infty$ sağlanır.

4) Eğer zincirin kalıcı olasılık dağılımı varsa, zincir görünür kalıcıdır.

Uyarı: x vektörünü $\sum x_i$ ile bölerek π kalıcı olasılık dağılımını elde ederiz. Dolayısıyla, bu teoremin bir sonucu olarak şunu ifade edebiliriz: Bir zincirin kalıcı olasılık dağılımı olması için gerek ve yeter şart zincirin görünür kalıcı olmasıdır.

Kanıt: (1) ve (2) Alıştırma 48 ve sonrasındaki teoremde kanıtlanmıştı.

(3) Bir k durumu için T_k tekrar k durumuna gelmek için geçen ortalama adım sayısı olsun. $N_k(i)$ ise bu süre zarfında i'inci duruma uğrama sayısı olsun:

$$N_k(i) = \sum_{n=1}^{\infty} 1_{\{X_n = i\} \cap \{T_k \geq n\}}.$$

Kolayca, $N_k(k)=1$ ve $T_k=\sum_{i\in S} N_i(k)$ olur. Ayrıca, tanımından dolayı

$$\mu_k=\mathbb{E}[T_k \mid X_0=k] \text{ olur.}$$

$\rho_i(k)=\mathbb{E}[N_i(k)\mid X_0=k]$ olarak tanımlansın. O halde,

$$\mu_k=\sum_{i\in S}\rho_i(k) \text{ eşitliği doğrudur.}$$

<u>İddia:</u> $\rho_i(k)<\infty$.

Bunu görmek için $L_{ki}(n) \doteq \mathbb{E}[1_{\{X_n=i\}\cap\{T_k\geq n\}}]=\mathbb{P}(\{X_n=i\}\cap\{T_k\geq n\})$ olsun, öyle ki

$$\mathbb{E}[N_i(k)] = \sum_{n=1}^{\infty} L_{ki}(n) \text{ olur.}$$

$f_{kk}(m+n)\geq L_{ki}(n)f_{ik}(m)$ olduğu kolayca görülür.

Zincir indirgenemez olduğu için $f_{ik}(m)>0$ olacak şekilde bir m seçebiliriz. O halde, $L_{ki}(n)\leq f_{kk}(m+n)/f_{ik}(m)$ olur. Buradan,

$$\rho_i(k)=\mathbb{E}[N_i(k)\mid X_0=k]=\sum_{n=1}^{\infty}L_{ki}(n)\leq\sum_{n=1}^{\infty}\frac{f_{kk}(m+n)}{f_{ik}(m)}\leq\frac{1}{f_{ik}(m)}<\infty.$$

Böylece iddianın kanıtı tamamlanır.

Şimdi de ρ_i dağılımının kalıcı olduğunu gösterelim: $n\geq 2$ olmak üzere $L_{ki}(n)=\sum_{j\neq k}L_{kj}(n-1)P_{ji}$ olur ve buradan

$$\begin{aligned}\rho_i(k)&=L_{ki}(1)+\sum_{j\neq k}\sum_{n=2}^{\infty}L_{kj}(n-1)P_{ji}\\&=P_{ki}+\sum_{j\neq k}\sum_{n=2}^{\infty}L_{kj}(n-1)P_{ji}\\&=P_{ki}+\sum_{j\neq k}\rho_j(k)P_{ji}=\sum_{j\in S}\rho_j(k)P_{ji}\end{aligned}$$

($\rho_k(k)=1$ olduğunu kullandık!) Dolayısıyla, $\rho_i(k)$ kalıcıdır.

O halde, her $k\in S$ için bir kalıcı vektörümüz var: $\rho_i(k)$. Eğer bu görünür kalıcı ise, $\sum\rho_i(k)=\mu_k<\infty$, vektörü μ_k ile bölerek görünür kalıcı olasılık dağılım vektörünü buluruz.

Kanıtı tamamlamak için bu dağılımın biricik ve pozitif olarak göstermeliyiz.

İlk önce bir j için $\pi_j = 0$ olduğunu kabul edelim. O halde, $0 = \pi_j = \sum \pi_i p_{ij}(n) \geq \pi_i p_{ij}(n)$, $\forall i, n$, olur. Eğer i ve j haberleşiyorsa, $p_{ij}(n) > 0$ bazı n için, $\pi_i = 0$ olmalı. Fakat zincir indirgenemez olduğu için her i için $\pi_i = 0$ olur, ve bu çelişki verir. Biricik olması bir sonraki teoremin sonucudur.

4) Alıştırma olarak okuyucuya bırakılmıştır. ■

<u>Tanım:</u> Bir $i \in S$ durumu için $\text{ebob}\{n \mid p_{ii}(n) > 0\} = 1$ ise i'ye döngüsel değildir denir.

<u>Uyarı:</u> İndirgenemez bir zincir için ya tüm durumlar döngüsel değildir yada hiçbiri döngüsel değildir. Bunu şöyle görebiliriz: Diyelim ki i döngüsel olmasın. Eğer j bir başka durum ise $h = p_{ji}(n) p_{ij}(m) > 0$ olacak şekilde m ve n seçelim. O halde,

$$p_{jj}(m+n+r) \geq p_{ji}(n) p_{ii}(r) p_{ij}(m) = h\, p_{ii}(r). \quad \xi = 1, \ldots, k, \text{ için}$$

$p_{ii}(r_\xi) > 0$ olsun. O halde, $\text{obeb}\{r_1, \ldots, r_k\} = 1$ olur. Bu durumda,

$$p_{jj}(l(m+n) + r_\xi) \geq h\, p_{ii}(r_\xi) > 0 \text{ ve dolayısıyla}$$

$$\text{obeb}\{l(m+n) + r_\xi \mid \xi = 1, \ldots, k,\ l \geq 1\} = 1 \text{ olur.}$$

<u>Alıştırma 50.</u> X ve Y aynı geçişme matrisine sahip olan bağımsız değişkenler olsun. Eğer X ve Y indirgenemez ve döngüsel değillerse $Z = (X, Y)$ değişkeni de indirgenemez ve döngüsel değildir.

<u>İpucu:</u> Eğer tam sayılar kümesinin toplama altında kapalı bir kümesi varsa bu küme elemanlarının en büyük ortak böleninin sonlu tanesi dışında tüm katlarını içerir.

<u>Çözüm:</u> $Z = (X, Y)$ ve $(i, j) \in Ran(X) \times Ran(Y)$ için

$$\mathbb{P}(Z_n = (i, j)) = \mathbb{P}(X_n = i, Y_n = j) = \mathbb{P}(X_n = i)\, \mathbb{P}(Y_n = j) \text{ olur.}$$

Ayrıca, $\mathbb{P}(Z_n=(i,j) \mid Z_0=(i_0,j_0),\dots,Z_{n-1}=(i_{n-1},j_{n-1}))$

$$= \mathbb{P}(X_n=i, Y_n=j \mid X_0=i_0,\dots,X_{n-1}=i_{n-1}, Y_0=j_0,\dots,Y_{n-1}=j_{n-1})$$

$$= \mathbb{P}(X_n=i \mid Y_n=j, X_0=i_0,\dots,Y_{n-1}=j_{n-1})\ \mathbb{P}(Y_n=j \mid X_0=i_0,\dots,Y_{n-1}=j_{n-1})$$

$$= \mathbb{P}(X_n=i \mid X_{n-1}=i_{n-1})\ \mathbb{P}(Y_n=j \mid Y_{n-1}=j_{n-1})$$

Dolayısıyla, Z bir Markov zinciridir ve

$\mathbb{P}_{(i_1,j_1)(i_2,j_2)} = P_{i_1 i_2} P_{j_1 j_2}$. Başka bir deyişle, Z zincirinin geçişme matrisi $\mathbb{P} \otimes \mathbb{P}'$ dir.

<u>Z indirgenemezdir</u>: Her $(i_1,j_1),(i_2,j_2)$ ikilileri için öyle m ve n seçelim ki $P_{i_1 i_2}(m)>0$ ve $P_{j_1 j_2}(n)>0$ olsun.

Diğer yandan, her A,B,C,D matrisi için

$(A\otimes B)\cdot(C\otimes D) = AC \otimes BD$ olduğu için

$(A\otimes B)^n = A^n \otimes B^n$ ve $(\mathbb{P}\otimes\mathbb{P})^n = \mathbb{P}^n \otimes \mathbb{P}^n$ elde edilir.

Buradan, $(\mathbb{P}\otimes\mathbb{P})^{(k)}_{(i_1,j_1)(i_2,j_2)} = \mathbb{P}_{i_1 i_2}(k)\,\mathbb{P}_{j_1 j_2}(k)$ olur.

Zincir döngüsel olmadığı için (ipucunu kullanarak) öyle bir K vardır ki, her $k \geq K$ için $\mathbb{P}_{i_1 i_2}(k)>0$ ve $\mathbb{P}_{j_1 j_2}(k)>0$ olur.

O halde, $(\mathbb{P}\otimes\mathbb{P})_{(i_1,j_1)(i_2,j_2)}(m+n+k) = \mathbb{P}_{i_1 i_2}(m+n+k)\,\mathbb{P}_{j_1 j_2}(m+n+k)$

$$\geq \mathbb{P}_{i_1 i_1}(n+k)\,\mathbb{P}_{i_1 i_2}(m)\,\mathbb{P}_{j_1 j_1}(m+k)\,\mathbb{P}_{j_1 j_2}(n) > 0$$ olur.

Dolayısıyla, $Z=(X,Y)$ indirgenemezdir.

<u>Z döngüsel değildir</u>: X ve Y döngüsel olmadığı için, ipucundan dolayı, öyle bir N vardır ki her $n \geq N$ için $\mathbb{P}_{ii}(n)>0$ ve $\mathbb{P}_{jj}(n)>0$ olur. Buradan $(\mathbb{P}\otimes\mathbb{P})_{(i,j)(i,j)}(n) = \mathbb{P}_{ii}(n)\,\mathbb{P}_{jj}(n) > 0$, $\forall n \geq N$, elde edilir

ve dolayısıyla $Z=(X,Y)$ döngüsel değildir.

<u>Tanım:</u> İndirgenemez, döngüsel olmayan ve görünür bir Markov zincirine Ergodik zincir denir.

<u>Teorem:</u> Bir ergodik zincir için, $n \to \infty$ iken her $i, j \in S$ için $p_{ij}(n) \to \pi_j = 1/\mu_j$ olur.

<u>Kanıt:</u> Bir önceki alıştırmada (Alıştırma 50) incelediğimiz ikili zincirleri kullanacağız. X, Y bir olasılık dağılım matrisine sahip iki bağımsız değişken ve $X_0 = i$, $Y_0 = j$, $Z = (X, Y)$ ve dolayısıyla $Z_0 = (i, j)$ olsun.

Verilen herhangi bir $s \in S$ durumu için T şöyle tanımlansın:

$$T = \min\{n \geq 1 \mid Z_n = (s, s)\},$$ zincirin (s,s)'den ilk geçiş zamanı.

Şimdi,
$$
\begin{aligned}
P_{ik}(n) &= P(X_n = k) \\
&= P(X_n = k, T \leq n) + P(X_n = k, T > n) \\
&= P(Y_n = k, T \leq n) + P(X_n = k, T > n) \\
&\leq P(Y_n = k) + P(T > n) \\
&= p_{jk}(n) + P(T > n) \text{ olur.}
\end{aligned}
$$

3. eşitlik için $T \leq n$ durumdan X_n ve Y_n'nin dağılımlarının denk olduğunu kullandık (T anında X ve Y eşitlerdir!).

Şimdi i ve j'yi yer değiştirirsek, bu defa $P_{jk}(n) \leq P_{ik}(n) + P(T > n)$ elde ederiz. O halde,

$$|P_{ik}(n) - P_{jk}(n)| \leq P(T > n) \text{ elde edilir.}$$

$1 = \sum_{n=1}^{\infty} P(T = n)$ olduğu için, n sonsuza giderken, $n \to \infty$, $P(T > n) \to 0$, sıfıra gider.

O halde, $n \to \infty$ iken $|P_{ik}(n) - P_{jk}(n)| \to 0$.

Şimdi, π X değişkeni için kalıcı dağılım olsun. Bu durumda

$$
\begin{aligned}
\pi_k - P_{jk}(n) &= \sum_i \pi_i P_{ik}(n) - \Big(\underbrace{\sum_i \pi_i}_{=1}\Big) P_{jk}(n) \\
&= \sum_i \pi_i \left(P_{ik}(n) - P_{jk}(n)\right) \to 0, \quad n \to \infty \text{ iken.}
\end{aligned}
$$

⌈Yukarıdaki limiti daha açık hesaplayalım:
$F \subseteq S$ sonlu bir alt küme olsun. O zaman

$$\sum_i \pi_i \left|P_{ik}(n) - P_{jk}(n)\right| \le \sum_{i\in F} \pi_i \left|P_{ik}(n)-P_{jk}(n)\right| + 2\sum_{i\notin F}\pi_i$$

olur, çünkü her n için $|P_{ik}(n)-P_{jk}(n)|\le 1$'dir
(sonlu toplam $n\to\infty$ için sıfıra yakınsar!)

Son olarak S kümesine yakınsayan ve artan sonlu bir $F_a \subseteq S$ alt küme dizisi alalım. $\sum_{i\in S}\pi_i=1$ olduğu için $\sum_{i\notin F_a}\pi_i \to 0$ olur ve kanıtı tamamlanır.⌋

Şimdi teoremin kanıtını tamamlayalım:

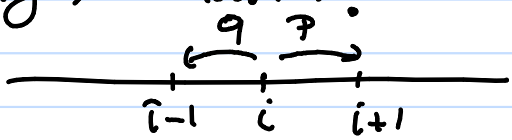
Hem $|P_{ik}(n)-P_{jk}(n)|\to 0$ ve $\pi_k - P_{jk}(n)\to 0$ olduğunu gördük. Ayrıca, $P_{jk}(n)\to\pi_k$, j durumundan bağımsızdır.

Dolayısıyla, X için tek bir kalıcı dağılım vardır. Ayrıca bir önceki teoremin kanıtından $\rho_k(k)=1$ ve böylece $\pi_k = \rho_k(k)/\sum_i \rho_i(k) = 1/\mu_k$ olur. (????)

<u>Alıştırma 51.</u> Bir boyutlu rasgele yürüyüşün $p\neq 1/2$ olduğu zaman geçici ve $p=1/2$ olduğu zaman görünmez kalıcı olduğunu gösteriniz.

<u>Çözüm:</u> Stirling Formülü: $n! \sim n^n e^{-n}\sqrt{2\pi n}$ as $n\to\infty$.

$X_0 = i$



$p+q=1.$

$$P_{ij} = \begin{cases} p & j=i+1 \\ q & j=i-1 \\ 0 & |j-i|\neq 1 \end{cases}$$

$$(px+qx^{-1})^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} x^{2k-n}$$

$$2k-n = j-i \Rightarrow k = \frac{n+j-i}{2}.$$

$$P_{ij}(n) = P_{0\,j-i}(n) = \binom{n}{k} p^k q^{n-k} = \binom{n}{\frac{n+j-i}{2}} p^{\frac{n+j-i}{2}} q^{\frac{n-j+i}{2}}$$

O halde, $p_{ii}(n) = \binom{n}{n/2}(pq)^{n/2} = \dfrac{n!}{((n/2)!)^2}(pq)^{n/2}$

Bu arada, eğer $n \geq 1$ tek sayı ise $p_{ii}(n) = 0$'dır.

Şimdi Stirling Formülü'nü kullanalım:

$$p_{ii}(n) \sim \frac{n^n e^{-n}\sqrt{2\pi n}}{\left((n/2)^{n/2} e^{-n/2}\sqrt{\pi n}\right)^2}(pq)^{n/2}$$

$$= \frac{n^n\sqrt{2\pi n}}{(n/2)^n\,\pi n} = 2^n\sqrt{\frac{2}{\pi n}}\,(pq)^{n/2}$$

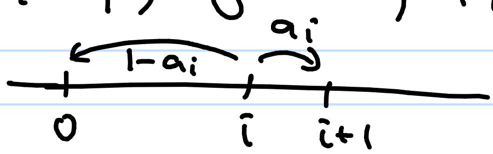
Eğer $p = q = 1/2$ ise $p_{ii}(n) \sim \sqrt{\frac{2}{\pi n}}$ ve buradan $\sum_n p_{ii}(n) = +\infty$ olur.

Diğer yandan $p \neq 1/2$ ise $pq < 1/4$ olur.

Buradan, $p_{ii}(n) \sim 2^n\sqrt{\frac{2}{n\pi}}(pq)^{n/2} \leq \sqrt{\frac{2}{n\pi}}\, r^n$, $r = \sqrt{4pq} < 1$, ve dolayısıyla $\sum_{n=1}^{\infty} p_{ii}(n) < +\infty$ olur. Bu kanıtı bitirir.

<u>Alıştırma 52.</u> Durum uzayı $S = \{0,1,2,\dots\}$ ve geçişme olasılıkları

$p_{i\,i+1} = a_i$, $p_{i,0} = 1 - a_i$, öyle ki $\{a_i \mid i \geq 0\}$, $0 < a_i < 1$, $\forall i$, bir dizi olsun.



$b_0 = 1$ ve $i \geq 1$ için $b_i = a_0 a_1 \cdots a_{i-1}$ olsun.

a) Zincir kalıcıdır ancak ve ancak $b_i \to 0$, $i \to \infty$.

b) Zincir görünür kalıcıdır ancak ve ancak $\sum_i b_i < +\infty$.

(b) durumunda kalıcı dağılımı yazınız.

Çözüm: a) $f_{ii}(n) = P(X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i, X_n = i \mid X_0 = i)$

So, $f_{00}(n) = P(X_1 \neq 0, \dots, X_{n-1} \neq 0, X_n = 0 \mid X_0 = 0)$

$$= P_{01} P_{12} \cdots P_{n-1\,n} P_{n0}$$
$$= a_0 a_1 \cdots a_{n-1} (1 - a_n)$$
$$= b_n - b_{n+1}.$$

Buradan $f_{00} = \lim_{N\to\infty} \sum_{n=0}^{N} f_{00}(n)$

$$= \lim_{N\to\infty} \sum_{n=0}^{N} (b_n - b_{n+1}) = b_0 - \lim_{N\to\infty} b_{N+1}$$

$$\Rightarrow f_{00} = 1 - \lim_{N\to\infty} b_{N+1}.$$

Zincirin kalıcı olması $f_{00} = 1$ olmasına denk olduğuna göre zincirin kalıcı olması $\lim_{n\to\infty} b_n = 0$ olmasına denktir.

b) $\mu_0 = \sum_{n=0}^{\infty} n f_{00}(n) = \sum_{n=0}^{\infty} n (b_n - b_{n+1})$

$$= \lim_{N\to\infty} 1\cdot(b_0 - b_1) + 2(b_1 - b_2) + \dots + N(b_N - b_{N+1})$$

$$= \lim_{N\to\infty} \left(\left(\sum_{n=1}^{N} b_n\right) - N b_{N+1} \right)$$

Şimdi ilk önce $\sum_{n=1}^{\infty} b_n < +\infty$ olduğunu kabul edelim.

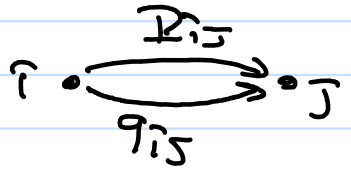
Eğer $\lim_{n\to\infty} n b_{n+1} \neq 0$ ise sonsuz tane $n b_{n+1}$ terimi için $n b_{n+1} > \epsilon$ olacak şekilde bir ϵ vardır.

Alıştırma 53. Bir dallanma işleminde n adım sonra oluşan dalların sayısını
$$X_n = \sum_{i=0}^{X_{n-1}} Z_i$$
olarak yazabiliriz. Burada $X_0 = 1$ ve Z_i de i'inci dalın oluşturduğu yeni fülizlerin sayısını göstermektedir. Z_i'lerin hepsi aynı olasılık dağılımına sahip bağımsız değişkenler olsun.
$$\pi_0 = \lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 1)$$
tüm topluluğun ortadan kalkma olasılığı olsun. π_0'ın $\pi_0 = \sum_{j=0}^{\infty} \pi_0^j \, \mathbb{P}(Z=j)$ denklemini sağlayan en küçük pozitif sayı olduğunu gösteriniz.

İpucu: $\phi(s) = \mathbb{E}\, s^Z$ ve $\phi_n(s) = \mathbb{E}\, s^{X_n}$, $s > 0$, fonksiyonlarını tanımlayalım. $\phi_{n+1}(s) = \phi(\phi_n(s))$ olduğunu gösterin. Bu durumda π_0 ϕ fonksiyonunun sabit noktası olacaktır.

§ 44. Zamanda Geri Çevrilebilen Markov Zincirleri.

Geçişme matrisi p_{ij}, kalıcı olasılık dağılımı π olan bir (X_n) Markov zincirini alalım. Bu zincirin tersi $q_{ij} = P(X_{n-1} = j \mid X_n = i)$ ile tanımlanır.



Kolayca,

$$\begin{aligned} q_{ij} &= P(X_{n-1}=j \mid X_n = i) \\ &= P(X_{n-1}=j, X_n = i) / P(X_n = i) \\ &= P(X_n = i \mid X_{n-1} = j) P(X_{n-1} = j) / P(X_n = i) \\ &= p_{ij} \pi_j / \pi_i \text{, olur.} \end{aligned}$$

$\pi_i = P(X_n = i)$

O halde, $q_{ij} \pi_i = p_{ij} \pi_j$ olur.

Tanım: Bir Markov zinciri için $p_{ij} = q_{ij}$, $\forall i, j$, oluyorsa zincire (zamanda) geri çevrilebilir zincir denir.

Yardımcı Teorem: (X_n) geçişme matrisi p_{ij} olan gözümlü kalıcı bir Markov zinciri olsun. Her $i, j \in S$ için $x_i > 0$, $\sum x_i < +\infty$ ve $x_i p_{ij} = x_j p_{ji}$ koşullarını sağlayan bir x_i kalıcı dağılımı verilsin.

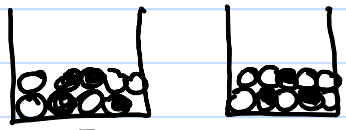
Bu durumda zincir (zamanda) geri çevrilebilir ve x_j vektörü de bu zincirin kalıcı olasılık dağılımının bir katıdır.

Bu sonucun kanıtı da okuyucuya bırakılmıştır.

Alıştırma 54. m adet beyaz ve m adet siyah top iki kutuya, her birinde m tane olacak şekilde, rasgele yerleştiriliyor. Her adımda iki kutudan rasgele seçilen iki top değiştirilerek tekrar kutulara konuluyor. Bu zincirin ters çevrilebilir olduğunu gösteriniz ve kalıcı olasılık dağılımını bulunuz.

X_n n-adım sonra birinci kutudaki beyaz topların sayısı olsun.

Çözüm:



i B $\quad$ i S

$m-i$ S $\quad$ $m-i$ B

$$p_{i\,i+1} = \left(\frac{m-i}{m}\right)^2, \ \forall i \leq m-1.$$

$$p_{i,i-1} = \left(i/m\right)^2, \ \forall i \leq m$$

$$p_{ii} = \frac{2i\,(m-i)}{m^2}, \ \forall i \leq m.$$

Tüm diğer $p_{ij} = 0$. Şimdi $x_i > 0$ ve $x_i p_{ij} = x_j p_{ji}$ koşulunu sağlayan (x_i) vektörünü bulmaya çalışalım.

$x_1 = 1$ olsun. $x_2 p_{21} = x_1 p_{12} \Rightarrow x_2 = p_{12}/p_{21} = (m-1)^2/2^2$ olur. Benzer şekilde,

$$x_3 = x_2 \cdot \frac{p_{23}}{p_{32}} = \frac{(m-1)^2}{2^2} \cdot \frac{(m-2)^2}{3^2}$$

ve genelde $x_i = \dfrac{(m-1)^2 \cdots (m-i+1)^2}{1^2 \cdot 2^2 \cdots i^2}$ olur, $1 \leq i \leq m$.

Kalıcı olasılık dağılımı ise $w_i = x_i / \sum x_i$ olur.

Alıştırma 55. $\mathbb{Z}$ üzerinde değerler alan ve sadece komşu durumlara sıçrayan bir zincirin geri çevrilebilir olduğunu gösteriniz.

Alıştırma 56. Yukarıdaki örnekte $\mathbb{Z}$ kümesini sonlu bir graf ile değiştirelim. Her bir köşedeki olasılıklar eşit şekilde dağıtılmış olsun. Bu zincirin de geri çevrilebilir olduğunu gösteriniz.

Alıştırma Bir topluluğun %10'nun belirli bir virüsü taşıdığını kabul edelim. Bu virüsü tespit etmesi için geliştirilen bir testin %90 ihtimalle doğru sonuç verdiği biliniyor. Test sonucu pozitif çıkan bir kişinin gerçekten bu virüsü taşıma ihtimali nedir?

Çözüm: Topluluğun 100 kişi olduğunu kabul edelim. Test sonucu pozitif çıkacak kişi sayısı 10 virüslünün %90'ı 9 kişi ve 90 virüs taşımayanların %10'u 9 kişi olur. O halde, cevap $9/(9+9) = \%50$ dur.

Bölüm 5. Olasılık Uzayları:

Bu bölümde sayılamaz büyüklükte örneklem uzaylarını çalışmaya başlayacağız. Bunu yaparken Lebesgue ölçümünü ve sürekli rasgele değişkenleri tanımlayıp kullanacağız.

§ 5.3. σ-Cebirleri.

Genellikle herhangi bir cebirsel yapıda sonlu toplamlar veya çarpımlar ele alınır. Bizim ihtiyaç duyduğumuz yapıda ise sayılabilir sonsuz toplamları hesaba katmamız gerekecek. O yüzden "cebir" sözcüğü yerine "σ-cebiri" tabirini kullanacağız.

Tanım: S boş olmayan bir küme olsun. S kümesinin aşağıdaki koşulları sağlayan bir $\mathcal{A}$ alt kümeler ailesine S üzerinde bir σ-cebiri denir:

1) $S \in \mathcal{A}$,
2) $A \in \mathcal{A} \Rightarrow A^c = S \setminus A \in \mathcal{A}$,
3) $A_n \in \mathcal{A},\ n=1,2,3,\ldots \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Alıştırma 57. $\mathcal{A}$ ailesinin sayılabilir sonsuz kesişim işlemi altında kapalı olduğunu gösteriniz: $A_n \in \mathcal{A},\ n=1,2,3,\ldots \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.

Alıştırma 58. 1) $\mathcal{P}(S)$, S'nin kuvvet kümesi de bir σ-cebiridir.

2) $\mathcal{A}$ S'nin sayılabilir ya da tümleyeni sayılabilir alt kümelerinin ailesi olsun. $\mathcal{A}$ bir σ-cebirdir.

Tanım: S boş olmayan bir küme, $\mathcal{A}$ S üzerinde bir σ-cebiri ve P aşağıdaki koşulları sağlayan bir fonksiyon ile $(S, \mathcal{A}, P)$ üçlüsüne bir olasılık üçlüsü denir:

1) $P : \mathcal{A} \to [0,1]$,
2) $P(S) = 1$
3) $A_1, A_2, \ldots, A_n, \ldots, A_n \in \mathcal{A}$ ve $A_i \cap A_j = \emptyset$ $(i \neq j)$ ile

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Hatırlatma: Yukarıdaki aksiyomların bazı sonuçları aşağıdaki gibidir.

a) Her $A \in \mathcal{A}$ için $S = A \cup A^c$, $A^c = S \setminus A$, ayrık birleşim olduğu için $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$ ve dolayısıyla $P(A^c) = 1 - P(A)$ olur.

b) Eğer $A \subseteq B$ ve $A, B \in \mathcal{A}$ ise $B = A \cup (B \setminus A)$ yazarak

$P(B) = P(A) + P(B \setminus A) \geq P(A)$ elde ederiz.

Alıştırma 59. Aşağıdaki kapsama-dışlama prensibinin kanıtını yap:

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_i A_i - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$
$$- \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n).$$

Hatırlatma. Eğer S sayılabilir sonsuz ve $\mathcal{A} = 2^S = \mathcal{P}(S)$ ise her alt küme bir olaydır ve $\sum_{s \in S} P(s) = 1$ olur.

§ 5.4. Süreklilik.

$\mathcal{A}$ bir σ-cebiri ve $A_1, A_2, \ldots, A_n, \ldots, A_n \in \mathcal{A}$ bir dizi olsun. Eğer her $n \geq 1$ için $A_n \subseteq A_{n+1}$ ise bu diziye artan, $A_n \supseteq A_{n+1}$ ise azalan dizi denir. Artan veya azalan bir diziye ayrıca monoton dizi denir.

Eğer (A_n) artan bir dizi ise $\lim A_n \doteq \bigcup A_n$ ve azalan bir dizi ise $\lim A_n \doteq \bigcap A_n$ olarak tanımlanır.

Yardımcı Teorem. (A_n) monoton bir dizi ise

$P(\lim A_n) = \lim P(A_n)$ olur. Başka bir deyişle P süreklidir.

Kanıt: Sadece (A_n)'nin artan olduğu durumu ele alacağız. Bu durumda, $B_{n+1} = A_{n+1} \setminus A_n$ olarak tanımlansın, $B_1 = A_1$ olmak üzere. Bu durumda, B_n dizisi ayrık bir dizidir ve $\bigcup_{n=1}^{N} A_n = \bigcup_{n=1}^{N} B_n$ ve $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$ olur.

O halde, $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} B_n)$

$$= \sum_{n=1}^{\infty} \mathbb{P}(B_n)$$
$$= \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{P}(B_n)$$
$$= \lim_{N\to\infty} \mathbb{P}(\bigcup_{n=1}^{N} B_n)$$
$$= \lim_{N\to\infty} \mathbb{P}(\bigcup_{n=1}^{N} A_n)$$
$$= \lim_{N\to\infty} \mathbb{P}(A_N)$$

$\Rightarrow \mathbb{P}(\lim A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$ olur ve kanıt tamamlanır. ∎

<u>Alıştırma 60.</u> Bu kanıtı azalan diziler için yapınız.

<u>Hatırlatma.</u> (A_n) herhangi bir dizi olsun ve $B_n = \bigcup_{k=1}^{n} A_k$ ile tanımlansın. Bu durumda (B_n) artan bir dizidir ve $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$ olur. O halde,

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} B_n) = \lim_{n} \mathbb{P}(B_n)$$
$$= \lim_{n} \mathbb{P}(\bigcup_{k=1}^{n} A_k)$$
$$\leq \lim_{n} \sum_{k=1}^{n} \mathbb{P}(A_k)$$
$$= \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

<u>Yardımcı Teorem.</u> (A_n) ve (B_n) $\bigcup_n A_n = \bigcup_n B_n$ koşulunu sağlayan ama ayrık terimlerden oluşan iki dizi olsun. $(A_i \cap A_j = \emptyset = B_i \cap B_j,\ i \neq j)$

Bu durumda, $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \mathbb{P}(B_n)$ olur.

<u>Kanıt:</u> $E = \bigcup A_n = \bigcup B_n$ olsun. $C_{n,m} \doteq A_n \cap B_m$ olarak tanımlansın. Bu durumda, $C_{n,m}$ ayrık kümelerden oluşan bir dizidir ve

$$\bigcup_{n=1}^{\infty} C_{n,m} = E \cap B_m = B_m \text{ ve } \bigcup_{m=1}^{\infty} C_{n,m} = E \cap A_n = A_n$$

olur. O halde,

$$\begin{aligned}\sum_{n=1}^{\infty} P(A_n) &= \sum_{n=1}^{\infty} P\Big(\bigcup_{m=1}^{\infty} C_{n,m}\Big) \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} P(C_{n,m}) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} P(C_{n,m}) \\ &= \sum_{m=1}^{\infty} P\Big(\bigcup_{n=1}^{\infty} C_{n,m}\Big) \\ &= \sum_{m=1}^{\infty} P(B_m).\end{aligned}$$

Alıştırma 61. $\mathcal{A}$ ve $\mathcal{B}$ bir S kümesi üzerinde σ-cebirleri ise $\mathcal{A} \cap \mathcal{B}$ ailesi de bir σ-cebirdir.

§ 5.6. Bir Sınıf tarafından Üretilen σ-Cebiri.

$\mathcal{C}$ verilen bir S kümesinin alt kümelerinin bir ailesi olsun. $\mathcal{C}$ ailesini içeren tüm σ-cebirlerinin arakesitide bir σ-cebirdir. Bu σ-cebiri $\sigma(\mathcal{C})$ ile gösterilecektir. Aşağıdaki yardımcı teoremin kanıtını okuyucuya bırakıyoruz.

Yardımcı Teorem: $\mathcal{T}$ bir S kümesi üzerinde verilen σ-cebirlerinin bir ailesi olsun. Bu durumda
$\mathcal{B} = \bigcap_{\mathcal{A} \in \mathcal{T}} \mathcal{A}$ arakesiti de bir σ-cebirdir.

Alıştırma 62. $\mathcal{C}$ ile gerçel sayıların tüm $(a, b]$ şeklindeki aralıklarının ailesini düşünelim. $\sigma(\mathcal{C})$ σ-cebirinin tüm aralıkları kapsadığını gösteriniz.

§ 5.7. Borel Kümeleri.

Yukarıdaki alıştırmada tanımlanan σ-cebirine $\mathbb{R}$ üzerindeki Borel cebiri denir. $\mathcal{B}$ ile gösterilen bu cebirin elemanlarına Borel kümeler denir.

Alıştırma 63. $\mathcal{B}$ cebiri $\mathbb{R}$'nin tüm açık ve kapalı kümelerini içerir. $\mathcal{B}$ sadece açık aralıklar ile de üretilebilir.

Alıştırma 64. $f:\mathbb{R}\to\mathbb{R}$ sürekli bir fonksiyon olsun. $\mathcal{C}\doteq\{E\subseteq\mathbb{R} \mid f^{-1}(E)\in\mathcal{B}\}$ kümesini tanımlayalım. Bu durumda $\mathcal{C}$ tüm açık aralıkları içeren bir σ-cebirdir. Dolayısıyla, $\mathcal{C}$ Borel cebiri de içerir.

§ 5.8. Lebesgue Ölçümü.

$l(a,b]=l(a,b)=|a-b|$ ile verilen aralık ölçümünü tüm Borel kümelere genişletmek istiyoruz. Bunu yapabilmek için ilk önce "dış ölçüm" olarak adlandırılan ölçümü tanımlayacağız:

$A\in\mathcal{B}$ bir Borel küme olmak üzere

$$m^*(A)=\inf\{\sum_{n=1}^{\infty} l(I_n) \mid A\subseteq\bigcup_{n=1}^{\infty} I_n\},$$ $I_n\subseteq\mathbb{R}$ aralıkları üzerinden infimum olarak

tanımlanır. Bu tanımın aralıklar üzerinde uzunluk ölçümü ile aynı olduğunu göstermeliyiz.

İddia: $I\subseteq\mathbb{R}$ bir aralık ise $m^*(I)=l(I)$ olur.

Kanıt: İlk önce $I=[a,b]$ alalım. O halde, $l(I)=l[a,b]=b-a$ olur. O halde, $m^*(I)=b-a$ olduğunu göstermeliyiz. $\epsilon>0$ herhangi bir pozitif gerçel sayı olsun. $I\subseteq\bigcup_{n=1}^{\infty} I_n$ olacak şekilde bir aralık koleksiyonu alalım. Her $I_n=(a_n,b_n),[a_n,b_n),(a_n,b_n]$ veya $[a_n,b_n]$ için $a_n'=a_n-\frac{\epsilon}{2^{n+1}}$ ve $b_n'=b_n+\frac{\epsilon}{2^{n+1}}$ olsun. O halde, $I_n\subseteq(a_n',b_n')\doteq I_n'$ olur.

$I\subseteq\bigcup_{n=1}^{\infty} I_n'$ ve I tıkız olduğu için $I\subseteq I_{n_1}'\cup\cdots\cup I_{n_k}'$ olacak şekilde $n_1,\dots,n_k$ vardır. O halde,

$$|a-b|=l(I)\leq l(I_{n_1}')+\cdots+l(I_{n_k}')\leq\sum_{n=1}^{\infty} l(I_n')=\epsilon+\sum_{n=1}^{\infty} l(I_n)$$

elde edilir. İnfimum alarak, $|a-b|=l(I)\leq\epsilon+m^*(I)$ buluruz. $\epsilon>0$ rasgele olduğu için $m^*(I)\geq|a-b|$ olur. Diğer yandan, $I=I_1=[a,b]$ olduğu için $|a-b|\geq m^*(I)$ olur. Dolayısıyla, $I=[a,b]$ için $m^*(I)=l(I)$ olmalıdır.

Diğer aralık tipleri için $\{I_n\}$ koleksiyonuna $[a,a+\epsilon/2)$ veya $(b-\epsilon/2,b]$ aralıklarını da ekleyerek kanıtı tamamlarız. ■

Tanım: Eğer bir $E \subseteq \mathbb{R}$ alt kümesi her $A \subseteq \mathbb{R}$ alt kümesi için
$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c), \quad E^c = \mathbb{R} \setminus E, \text{ eşitliği}$$
sağlanıyorsa E kümesine ölçülebilir küme denir.

$\mathcal{M}$ tüm ölçülebilir kümelerin koleksiyonu olsun.

Yardımcı Teorem: $\mathcal{M}$ bir σ-cebirdir ve $\mathcal{B}$ Borel cebirini kapsar. Ayrıca $E_n \in \mathcal{M}$, $n \in \mathbb{N}$, kümeleri ayrık ise
$$m^*\left(\bigcup_n E_n\right) = \sum_n m^*(E_n).$$

Dolayısıyla, m^* $\mathcal{M}$ σ-cebiri üzerinde bir ölçümdür.

Uyarı: Bu ölçüm ötelemeler altında değişmezdir. Ayrıca, $\mathcal{M}$ cebiri $\mathcal{B}$ Borel cebirinden büyüktür.

Bu yardımcı teorem İntegrasyon Teorisinin ana sonuçlarından biridir. Aslında, Carathéodory–Hahn genişleme teoreminin bir sonucudur. Robert G. Bartle'in "The Elements of Integration and Lebesgue Measure" kitabının 9. kısmına bakabilirsiniz.

Alıştırma 65. Her sayılabilir küme ölçülebilirdir ve ölçümü sıfırdır. Özel durumda, $\mathbb{Q}$ kümesinin ölçümü sıfırdır.

Yardımcı Teorem. $B \in \mathcal{B}$ bir Borel küme olsun. Bu durumda, her $\epsilon > 0$ için bir F kapalı ve G açık kümeleri vardır, öyle ki $F \subseteq B \subseteq G$, $\lambda(B \setminus F) < \epsilon$ ve $\lambda(G \setminus B) < \epsilon$ olur.

Kanıt: İlk önce G kümesini kuralım. Dış ölçümün tanımı gereği her $\epsilon > 0$ için sayılabilir $\{I_n\}$ açık aralık kümesi vardır öyle ki $B \subseteq \bigcup I_n$ ve
$$\lambda(B) = m^*(B) > \sum_{n=1}^{\infty} l(I_n) - \epsilon \quad \text{olur.}$$

$G = \bigcup_n I_n$ olarak tanımlansın. Öyleyse, G açık bir kümedir, $B \subseteq G$ koşulunu sağlar ve
$$\lambda(G \setminus B) = \lambda(G) - \lambda(B) \leq \sum_n l(I_n) - \lambda(B) < \epsilon \quad \text{olur.}$$

F'yi bulmak için B^c kümesi için bulacağımız G kümesinin tümleyeni alalım. Bu küme aranan şartları sağlar. ∎

Alıştırma 66. m^* dış ölçümünün ötelemeler altında değişmez olduğunu gösteriniz: $m^*(A+\{x\})=m^*(A)$, $\forall A\subseteq \mathbb{R}$, $x\in\mathbb{R}$.

Alıştırma 67. Lebesgue ölçümü tamdır. Başka bir deyişle eğer B ölçülebilir ve $\lambda(B)=0$ ise her $A\subseteq B$ alt kümesi de ölçülebilirdir ve $\lambda(A)=0$'dır.

Çözüm. Dış ölçümün tanımı gereği $m^*(B)=0$ ve $A\subseteq B$ olduğundan $m^*(A)=0$ olur. O halde, A'nın ölçülebilir olduğunu gösterirsek $\lambda(A)=m^*(A)=0$ olur. Şimdi A'nın ölçülebilir olduğunu gösterelim.

B ölçülebilir olduğu için her $C\subseteq\mathbb{R}$ alt kümesi için $m^*(C)=m^*(B\cap C)+m^*(B^c\cap C)$ eşitliği sağlanır. $m^*(B)=\lambda(B)=0$ olduğu için $m^*(B\cap C)=0$ olur ve dolayısıyla $m^*(C)=m^*(B^c\cap C)$ olur. Ayrıca, $m^*(A)=0$ olduğundan $m^*(A\cap C)=0$ dır. O halde, sadece $m^*(A^c\cap C)=m^*(C)$ kanıtlamak yeterlidir. $m^*(A^c\cap C)\leq m^*(C)$ olduğu açıktır, çünkü $A^c\cap C\subseteq C$'dir.

Ters eşitsizliği göstermek için rasgele bir $\epsilon>0$ alalım. $\{I_n\}$, $A\subseteq\{I_n\}$ ve $\sum_n l(I_n)<\epsilon$ olan bir açık aralık ailesi (sayılabilir) olsun.

Şimdi, $C\setminus(A^c\cap C)=C\cap(A^c\cap C)^c=C\cap A\subseteq A$ ve dolayısıyla $C\subseteq A\cup(A^c\cap C)$ olur.

$\{J_n\}$ sayılabilir açık aralıklar dizisi olsun öyle ki $A^c\cap C\subseteq\bigcup J_n$ sağlansın. O zaman $\{I_n, J_n\}$ C'yi içeren bir sayılabilir açık aralıklar dizisi olur. O halde,

$$m^*(C)\leq\sum l(I_n)+\sum l(J_n)<\epsilon+\sum_n l(J_n).$$

$\{J_n\}$ diziler üzerinden infimum alarak

$$m^*(C)\leq\epsilon+\inf\Big\{\sum_n l(J_n)\;\Big|\;A^c\cap C\subseteq\bigcup J_n\Big\}=\epsilon+m^*(A^c\cap C)$$

elde ederiz.

$\epsilon>0$ rasgele olduğu için $m^*(A^c\cap C)\geq m^*(C)$ olmalıdır ve böylece $m^*(C)=m^*(A^c\cap C)$ elde edilir. Bu kanıtı tamamlar. ■

§5.9. Lebesgue-Stieltjes Ölçümü.

Sürekli rasgele değişkenleri çalışmamız sırasında Lebesgue ölçümünün aşağıdaki genelleştirmesine ihtiyaç duyacağız:

$F:\mathbb{R}\to\mathbb{R}$ artan ve sağdan sürekli bir fonksiyon olsun: $\forall x<y$, $f(x)\leq f(y)$ ve $F(x)=\lim_{h\to 0^+} F(x+h)$.

Bu durumda $(a,b]$ yarı açık aralıklarını şu şekilde ölçebiliriz:

$$\mu(a,b]=F(b)-F(a).$$

Sonrasında l'den Lebesgue ölçümü elde ettiğimiz gibi bu yeni ölçümden de μ_F ölçümünü elde ederiz (yine $\mathcal{B}$ üzerinde).

<u>Lemma.</u> $F:\mathbb{R}\to\mathbb{R}$ sağdan sürekli artan fonksiyon olsun, öyle ki $\lim_{x\to-\infty} F(x)=0$. Bu durumda $\mathcal{B}$ Borel cebiri üzerinde

$\mu_F(a,b]=F(b)-F(a)$ koşulunu sağlayan tek bir μ_F ölçümü vardır.

<u>Alıştırma 68.</u> $F(x)=\begin{cases} 0, & x<0 \\ x^2, & 0<x<1/2 \\ 1/2, & 1/2\leq x<1 \\ 1, & x\geq 1 \end{cases}$ ile tanımlansın.

$\mu_F(0,1/2]$, $\mu_F(0,1/2)$, $\mu_F(\{1/2\})$, $\mu_F[1/2,1)$, $\mu_F[1/2,1]$ değerlerini hesaplayınız.

<u>§ 5.10. $\mathbb{R}^n$ üzerinde Lebesgue-Stieltjes Ölçümü.</u>

$\mathbb{R}^n$ üzerindeki Borel cebiri, $\mathcal{B}(\mathbb{R}^n)$ veya $\mathcal{B}^n$, $\mathbb{R}^n$ içindeki açık kümeler ya da $(a_1,b_1]\times\cdots\times(a_n,b_n]$ şeklindeki dikdörtgen prizmalar ile üretilen σ-cebiridir. $F:\mathbb{R}^n\to\mathbb{R}$ her değişkene göre artan ve sağdan sürekli bir fonksiyon olsun. Bu durumda $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ üzerinde aşağıdaki koşulu sağlayan tek bir μ_F ölçümü vardır:

$$F(x_1,\ldots,x_n)=\mu_F\big((-\infty,x_1]\times\cdots\times(-\infty,x_n]\big),\ \forall(x_1,\ldots,x_n)\in\mathbb{R}^n.$$

Buna F tarafından tanımlanan Lebesgue-Stieltjes ölçümü denir.

Eğer $F(x_1,\ldots,x_n)=F_1(x_1)\cdots F_n(x_n)$ şeklinde bir fonksiyon ise $\mu_F=\mu_{F_1}\times\cdots\times\mu_{F_n}$ olur.

§5.11. Rasgele Değişkenler.

$\mathcal{A}$ ve $\mathcal{C}$ sırasıyla S ve T kümeleri üzerinde σ-cebirleri olsun. Eğer bir $f: S \to T$ fonksiyonu her $C \in \mathcal{C}$ için $f^{-1}(C) \in \mathcal{A}$ koşulunu sağlıyorsa f fonksiyonuna ölçülebilir fonksiyon denir.
$T = \mathbb{R}$ üzerinde $\mathcal{B}$ Borel cebir olmak üzere ölçülebilir $f: S \to \mathbb{R}$ fonksiyonlarına rasgele değişken diyeceğiz.

Alıştırma 69. $A \in \mathcal{A}$ olmak üzere $\mathbb{1}_A$ işaret fonksiyonu rasgele değişkendir.

Alıştırma 70. $f, g: \mathbb{R} \to \mathbb{R}$ rasgele değişkenler ise $f+g$ ve $f \cdot g$ fonksiyonları da rasgele değişkenlerdir.

Ayrık uzaylarda kullandığımız "pmf" bu durumda tamamen anlamsızlaşır. Onun yerine, verilen bir $X: S \to \mathbb{R}$ rasgele değişkeni için $F_X: \mathbb{R} \to [0,1]$, $F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X^{-1}(-\infty, x])$ fonksiyonunu tanımlarız. Bazen gösterimden rasgele değişkeni atarak sadece F de yazarız. $F = F_X$'in aşağıdaki özellikleri sağladığı açıktır:

a) $0 \leq F(x) \leq 1 \quad \forall x \in \mathbb{R}$
b) $x < y \Rightarrow F(x) \leq F(y)$
c) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$

d) F sağdan süreklidir: $\lim_{x \to x_0^+} F(x) = F(x_0)$.

Alıştırma 71. (a) – (d) özelliklerini kanıtlayınız.

Alıştırma 72. $\mathbb{P}(X = x) = F(x) - \lim_{h \to 0^+} F(x-h)$

F_X fonksiyonuna yığınsal üreteç fonksiyonu denir (cumulative generating function) ve cdf. F_X fonksiyonu artan ve sağdan süreklidir. Dolayısıyla da, tek bir Lebesgue-Stieltjes ölçümünü belirler: $\mu_F: \mathcal{B} \to \mathbb{R}$ öyle ki, $\mu_F(a,b] = F(b) - F(a)$, $\forall a,b \in \mathbb{R}$.
$\mathbb{P}$ olasılık ölçümü, $\mathbb{P}(a < x \leq b) = \mu_F(a,b]$ eşitliğini sağlar.

Bunu, $X: (S, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B})$ ölçülebilir fonksiyonu bir ölçüm belirler (pushes forward) diye ifade edebiliriz: $\mu_F(B) = \mathbb{P}(x \mid X(x) \in B)$.

Alıştırma 73. $f: S \to T$ $(S, \mathcal{A}, \mu)$ ölçüm uzayından $(T, \mathcal{C})$'ye ölçülebilir fonksiyon olsun. Her $C \in \mathcal{C}$ için $\nu(C) = \mu(f^{-1}(C))$ olarak tanımlansın. ν'nin $(T, \mathcal{C})$ üzerinde bir ölçüm tanımladığını gösteriniz.

§5.12. Sürekli Rasgele Değişkenler

İlk önce sürekli olmanın daha kuvvetli bir halini tanımlayacağız.
Bir $g: \mathbb{R} \to \mathbb{R}$ fonksiyonu aşağıdaki koşulu sağlarsa bu fonksiyona kesin-sürekli fonsiyon denir:
Verilen her $\epsilon > 0$ için öyle bir $\delta > 0$ sayısı vardır ki, uzunlukları toplamı $\sum_{i=1}^{n} |y_i - x_i| < \delta$ olan her ayrık $\{(x_i, y_i)\}_{i=1}^{n}$ aralık kümesi için

$$\sum_{i=1}^{n} |g(y_i) - g(x_i)| < \epsilon \text{ olur.}$$

Teorem. Bir $F: [a,b] \to \mathbb{R}$ fonksiyonunun kesin-sürekli olması için gerek ve yeter şart,

$$F(x) = \int_c^x f(t)\,dt + F(c), \quad \forall\, c, x \in [a,b],$$

eşitliğini sağlayacak bir integrallenebilir f fonksiyonun var olmasıdır.

Bu teorem Radon-Nikodym teoremini özel bir halidir ve f fonksiyonuna F'nin Radon-Nikodym türevi denir.

Definition: Verilen bir X rasgele değişkeni için F_X kesin sürekli ise X rasgele değişkenine süreklidir denir.

Eğer X sürekli bir rasgele değişken ise F_X cdf fonksiyonu f_X pdf (probability distribution function) tarafından belirlenir ve aşağıdaki koşulları sağlar:

1) $f_X \geq 0$ ve ölçülebilirdir.

2) Her $a \in \mathbb{R}$ için $\mathbb{P}(X \leq a) = F_X(a) = \int_{-\infty}^{a} f_X(x)\,dx$ ve

dolayısıyla, $\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx$ olur.

Buradan, $\mathbb{P}(X = x) = 0$ ve $\mathbb{P}(X < a) = \mathbb{P}(X \leq a)$ elde edilir.
Ayrıca $\int_{-\infty}^{\infty} f_X(x)\,dx = \mathbb{P}(\mathbb{R}) = 1$ normalizasyon koşulu sağlanır.

Bazı Önemli Örnekler.

Düzgün Dağılım: $a<b$ olmak üzere $f(x)=\begin{cases} \frac{1}{b-a}, & x\in[a,b] \\ 0, & x\notin[a,b]. \end{cases}$

Bu durumda, X'e "$[a,b]$ aralığında eşit olasılıklı değişken" denir.

Üssel Dağılım: $f(x)=\begin{cases} ke^{-kx}, & x\geq 0 \\ 0, & x<0 \end{cases}$, ($k>0$ sabit).

Normal Dağılım: $f(x)=\dfrac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $\mu\in\mathbb{R}$ sayısına dağılımın ortalaması, $\sigma^2>0$'na ise varyansı denir. Bu en yaygın bilinen ve kullanılan rasgele değişkendir.

Alıştırma 74. Düzgün ve üssel değişkenler için cdf fonksiyonlarını hesaplayınız.

Alıştırma 75. Üssel dağılım "hafızasızlık özelliği" olarak bilinen

$$P(X>s+t \mid X>s) = P(X>t), \ \forall s,t>0,$$

eşitliğini sağladığını gösteriniz.

Alıştırma 76. $\displaystyle\int_{-\infty}^{+\infty} e^{-ax^2+bx}\,dx = \sqrt{\pi/a}\; e^{b^2/4a}$ eşitliğini kanıtlayarak normal dağılım için $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = 1$ olduğunu gösteriniz.

Alıştırma 77. Boyutu 1×1-birim^2 olan karoların oluşturduğu bir yüzeye yarıçapı $r<1/2$ olan bir bozuk para atılıyor. R paranın merkezinin en yakın karo merkezine olan uzaklık olsun. Bu değişkenini pdf fonksiyonunu bulunuz.

§5.13. Rasgele Vektör Fonksiyonları (Değişkenleri)

$X_1,\dots,X_n$ bir $(S,\mathcal{A},P)$ ölçüm uzayı üzerinde tanımlı rasgele değişkenler olsun. $X=(X_1,\dots,X_n): S\to\mathbb{R}^n$ fonksiyonu ölçülebilir bir fonksiyondur.

X değişkeninin birleşik cdf fonksiyonu

$F_X(x_1,\dots,x_n) = \mathbb{P}(X_1 \leq x_1,\dots,X_n \leq x_n)$ ile verilir. Birleşik

f_X pdf fonksiyonu da $\mathbb{P}(x\in B) = \int_B f_X(x_1,\dots,x_n)\,dx_1\cdots dx_n$, $\forall B \in \mathcal{B}(\mathbb{R}^n)$ eşitliğini sağlar.

§ 5.14. Bağımsızlık.

Ayrık teoride olduğu gibi verilen $A_1,\dots,A_n$ olayları için $\mathbb{P}(A_1\cap\dots\cap A_n) = \mathbb{P}(A_1)\cdots\mathbb{P}(A_n)$ oluyorsa bu olaylara bağımsız olaylar denir.

Benzer şekilde verilen $\mathcal{A}_1,\dots,\mathcal{A}_n$ olay aileleri için her $A_i \in \mathcal{A}_i$, $i=1,\dots,n$, için $A_1,\dots,A_n$ bağımsız ise ailelere de bağımsız denir.

Bir rasgele $X: S \to \mathbb{R}$ değişkeni için $\sigma(X)$ ile X'i ölçülebilir yapan S üzerindeki en küçük σ-cebirini gösterelim. Dolayısıyla, $\sigma(X)$ $\{X^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R})\}$ kümesini içerir.

Özel olarak X değişkeninin görüntüsü sayılabilir ayrık bir küme ise, $\mathrm{Ran}(X) = \{x_1, x_2,\dots,x_n,\dots\}$, $\sigma(X)$ $A_i = \sigma^{-1}(x_i)$ kümeleri tarafından üretilir.

Tanım. $X_1,\dots,X_n$ değişkenleri için $\sigma(X_1),\dots,\sigma(X_n)$ cebirleri bağımsız ise bu değişkenlere de bağımsız denir.

Başka bir deyişle, her $B_i \in \mathcal{B}$, $i=1,\dots,n$, için

$\mathbb{P}(X_1\in B_1,\dots X_n \in B_n) = \mathbb{P}(X_1 \in B_1)\cdots\mathbb{P}(X_n, B_n)$ oluyorsa değişkenlere bağımsız denir.

Aslında $\mathcal{B}$ Borel cebiri $(-\infty, x]$ şeklindeki aralıklar tarafından üretildiği için bağımsızlık koşulu şöyle de ifade edilebilir: Her $x_1,\dots,x_n \in \mathbb{R}$ için

$F_{X_1\dots X_n}(x_1,\dots,x_n) = F_{X_1}(x_1)\cdots F_{X_n}(x_n)$, yada benzer şekilde

$$f_{X_1\dots X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n).$$

Alıştırma 78. Bir dart birim kare şeklindeki hedef tahtasına atılıyor. Saplandığı noktanın koordinatları (X,Y) olsun. $|X-Y| \leq 1/4$ olma olasılığını hesaplayınız.

§ 5.15. Beklenti Fonksiyonları:

X rasgele bir değişken, F ve μ_F'de karşılık gelen cdf ve Lebesgue–Stieltjes ölçümü olsun. X'in beklenen değeri bu ölçüme göre hesaplanır: $B \in \mathcal{B}$ bir Borel kümesi ise gösterim fonksiyonun μ_F ölçümüne göre integrali $\int 1_B \, dF \doteq \mu_F(B) = P(X \in B)$ ile tanımlanır.

Eğer $\phi = \sum_{i=1}^{n} c_i 1_{B_i}$ ise $\int \phi = \sum_{n=1}^{n} c_i \int 1_{B_i} \, dF = \sum_{i=1}^{n} c_i \mu_F(B_i)$ olur.

Alıştırma 79. Eğer $\phi = \sum c_i 1_{A_i} = \sum d_j 1_{B_j}$ ise

$$\sum c_i \int 1_{A_i} \, dF = \sum d_j \int 1_{B_j} \, dF \text{ olur.}$$

O halde, sınırlı ölçülebilir bir $g \geq 0$ fonksiyonun integrali

$$\int g \, dF = \sup_{\phi \leq g} \int \phi \, dF \text{ ile tanımlanır.}$$

Yardımcı Teorem. $\int \cdot \, dF$ integrali tüm $f \geq 0$ ölçülebilir $f: \mathbb{R} \to \mathbb{R}$ fonksiyonları için iyi tanımlıdır ve aşağıdaki koşulları sağlar

i) $\int c g \, dF = c \int g \, dF, \ \forall c \in \mathbb{R}$

ii) $\int (g + h) \, dF = \int g \, dF + \int h \, dF.$

iii) $g \geq 0 \implies \int g \, dF \geq 0.$

iv) Eğer $g_n \uparrow g$ ise $\int g_n \, dF = \int g \, dF$ olur.

Uyarı. Dördüncü özellik Monoton Yakınsaklık Teoremi olarak bilinir.

Alıştırma 80. g sınırlı ve ölçülebilir fonksiyon olsun. $|g(x)| \leq M, \forall x$ ise her $n \geq 1$ için aşağıdaki kümeleri tanımlayalım:

$$E_k = \{x \mid \tfrac{kM}{n} \geq g(x) \geq \tfrac{(k-1)M}{n}\}, \quad -n \leq k \leq n.$$

$\psi_n(x) \doteq \frac{M}{n} \sum_{k=-n}^{n} k \, 1\!\!1_{E_k}(x)$ ve $\phi_n(x) = \frac{M}{n} \sum_{k=-n}^{n} (k-1) \, 1\!\!1_{E_k}(x)$ fonksiyonlarını tanımlayalım. O halde, $\psi_n(x) \leq g(x) \leq \phi_n(x), \forall x, n$ olduğunu gösteriniz.

Dolayısıyla,
$$\inf_{g \leq \phi} \int \psi \, dF = \sup_{g \geq \phi} \int \phi \, dF \text{ olur.}$$
(ϕ basit fonksiyon olmak üzere.)

~o~

g ölçülebilir fonksiyon ve $g^+, g^- \geq 0$ olmak üzere $g = g^+ - g^-$ şeklinde yazalım. Eğer her $\int g^{\mp} dF$ iki integral de sonluysa
$$\int g \, dF = \int g^+ dF - \int g^- dF \text{ olarak tanımlanır.}$$

<u>Yardımcı Teorem</u> (Sınırlandırılmış Yakınsaklık Teoremi)
(Dominated Convergence Teorem)

$\{g_n\}$ ölçülebilir fonksiyonlar dizisi olsun öyle ki $g_n \to g$ fonksiyonuna yakınsasın. Ayrıca integrallenebilir bir h fonksiyonu için $|g_n| \leq h \ \forall n$, olsun. O halde,
$$\lim_n \int g_n \, dF = \int g \, dF \text{ olur.}$$

<u>Alıştırma 81.</u> Yukarıdaki teoremi kullanarak aşağıdaki limitin varlığını gösteriniz ve limiti hesaplayınız:
$$\lim_{n \to \infty} \int_0^1 n \sin\left(\frac{1}{n\sqrt{x}}\right) dx.$$

<u>Tanım.</u> X bir rasgele değişken ve F onun cdf'si olsun. Herhangi bir ölçülebilir $g: \mathbb{R} \to \mathbb{R}$ fonksiyonu için
$$\mathbb{E}[g(x)] \doteq \int_{-\infty}^{\infty} g(x) \, dF \text{ olarak tanımlanır.}$$

Eğer X ayrık ve $Ran(X) = \{x_1, x_2, \ldots, x_n, \ldots\}$ ise
$$\mathbb{E}[g(x)] = \sum_i g(x_i) \, \mathbb{P}(X = x_i)$$ ve eğer X sürekli

ve f onun pdf fonksiyonu ile $\mathbb{E}[g(X)]=\int_{-\infty}^{+\infty} f(x)g(x)\,dx$ olur.

§5.16. (Kesin) Sürekli Rasgele Değişkenlerle Hesaplamalar.

Değişken Değişimi. X sürekli bir değişken ve g ölçülebilir fonksiyon olsun. $Y=g(X)$ değişkeninin cdf fonksiyonu

$$\mathbb{P}(Y\leq y)=\mathbb{P}(g(X)\leq y)=\int_{x:\,g(x)\leq y} f_X(x)\,dx$$ ile verilir, çünkü

$S \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$ bileşke fonksiyonu için

$$\mathbb{P}(a<X\leq b)=\int_a^b f_X(x)\,dx \quad \text{ve} \quad F_X(a)=\int_{-\infty}^a f_X(x)\,dx$$ olur.

Alıştırma 82. Z standart normal dağılım ise $Y=Z^2$'nin pdf fonksiyonu

$$f_Y(y)=\begin{cases} 0, & y\leq 0 \\ 1/\sqrt{2\pi}\; y^{-1/2} e^{-y/2}, & y>0 \end{cases}$$ ile verilir.

Çözüm. $f_Z(z)=\frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ ile verilir. O halde,

$$\mathbb{P}(Z\leq a)=F_Z(a)=\int_{-\infty}^a f_Z(z)\,dz=\int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$ olur.

Buradan

$$\begin{aligned}\mathbb{P}(Y\leq a) &= \mathbb{P}(Z^2\leq a)\\ &=\begin{cases} 0, & a<0 \\ \mathbb{P}(-\sqrt{a}\leq Z\leq \sqrt{a}), & a\geq 0 \end{cases}\\ &=\begin{cases} 0, & a<0 \\ \int_{-\sqrt{a}}^{\sqrt{a}} f_Z(z)\,dz \end{cases} =\begin{cases} 0, & a<0 \\ \int_{-\sqrt{a}}^{\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz, & a\geq 0. \end{cases}\end{aligned}$$

$$\Rightarrow \int_{-\infty}^a f_Y(y)\,dy=\begin{cases} 0, & a<0 \\ \int_{-\sqrt{a}}^{\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz, & a\geq 0, \end{cases}$$ elde edilir.

O halde, eğer $a<0$ ise $f_Y(y)=0$ ve $a>0$ ise

$$\int_{-\infty}^{a} f_Y(y)\,dy = \int_{-\sqrt{a}}^{\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$

$$= \int_{-\infty}^{\sqrt{a}} \cdots - \int_{-\infty}^{-\sqrt{a}} .$$ Burada her iki tarafın d/da'sını alırsak

$$f_Y(y) = \left(\frac{1}{2\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-a/2}\right) - \left(-\frac{1}{2\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-a/2}\right)$$

$$= \frac{1}{\sqrt{a}} \frac{1}{\sqrt{2\pi}} e^{-a/2}$$ buluruz.

~ o ~ o ~ o ~ o ~

Eğer $g: \mathbb{R}^n \to \mathbb{R}^n$ türevlenebilir ve türevlenebilir tersi olan bir fonksiyon ise (hem $g, T = g^{-1} \in C^1$) $Y = g \circ X$ olmak üzere

$f_Y(y) = f_X(T(y))\,|J(y)|$ olarak yazabiliriz (J Jakobyen matrix)

Alıştırma 83. X_1, X_2 normal dağılım değişkenleri olsun öyle ki birleşik pdf fonksiyonları

$$f_{X_1,X_2}(x_1,x_2) = \frac{3}{2\pi} e^{-\frac{1}{2}(2x_1^2 + 2x_1x_2 + 5x_2^2)}$$ ile verilsin.

$U = X_1 - X_2$, $V = X_1 + 2X_2$ alalım. $f_{U,V}(u,v) = \frac{1}{2\pi} e^{-\frac{1}{2}(u^2+v^2)}$ olduğunu gösteriniz.

Çözüm: $\left|\frac{\partial(u,v)}{\partial(x_1,x_2)}\right| = \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 3. \Rightarrow J = \frac{1}{3}.$

$X_1 = \frac{2U+V}{3}$, $X_2 = \frac{V-U}{3}$. O halde,

$$-\frac{1}{2}(2x_1^2 + 2x_1x_2 + 5x_2^2) = -\frac{1}{2}\left(\frac{8u^2+2v^2+8uv}{9} + \frac{2v^2-4u^2+2uv}{9} + \frac{5u^2+5v^2-10uv}{9}\right)$$

$$= -\frac{1}{2}\left(\frac{9u^2+9v^2}{9}\right) = -\frac{1}{2}(u^2+v^2)$$

$$\Rightarrow f_{U,V}(u,v) = \frac{3}{2\pi}\,\frac{1}{3}\cdot e^{-\frac{1}{2}(u^2+v^2)}.$$

∎

Bağımsız Rasgele Değişken İçeren olaylar:

Birden fazla bağımsız rasgele değişken içeren olayların olasılığını katlı integral kullanarak hesaplayabiliriz. Örneğin X ve Y iki bağımsız değişken ise bir $B \in \mathcal{B}$ olayısının olasılığı

$$P((X,Y)\in B)=\int_{-\infty}^{\infty} P((x,Y)\in B)\,dF_X=\int_{-\infty}^{+\infty}\Big(\int_{y:(x,y)\in B} dF_Y\Big)\,dF_X$$

olarak hesaplanabilir.

Alıştırma 84. X $[0,1]$ aralığı üzerinde üssel, Y ise düzgün bağımsız dağılımlar olsun. $P(X+Y\leq 1)=e^{-1}$ olduğunu gösteriniz.

Çözüm. $f_X(x)=\begin{cases} ke^{-kx}, & x\geq 0 \\ 0, & x<0 \end{cases}$ ve $f_Y(y)=\begin{cases} 1, & 0\leq x\leq 1 \\ 0, & \text{aksi halde} \end{cases}$

O halde, $F_X(x)=\int_{-\infty}^{x} f_X(t)\,dt=\int_0^x ke^{-kt}\,dt=-e^{-kt}\Big|_0^x=1-e^{-kx}$ ve

dolayısıyla, $F_X(x)=\begin{cases} 0, & x<0 \\ 1-e^{kx}, & x\geq 0 \end{cases}$ elde edilir.

Benzer şekilde, $F_Y(y)=\int_{-\infty}^{y} f_Y(t)\,dt=\int_0^y dt=y$ ve böylece

$F_Y(y)=\begin{cases} 0, & y<0 \\ y, & y\geq 0 \end{cases}$ olur.

Şimdi, $P(X+Y\leq 1)=\int_{-\infty}^{+\infty}\Big(\int_{y:x+y\leq 1} dF_Y\Big)\,dF_X$

$$=\int_0^{\infty}\int_{y:\,x+y\leq 1} dy\, ke^{-kx}\,dx$$

$$=\int_0^{1}\int_0^{1-x} dy\, ke^{-kx}\,dx$$

$$=(1-x)(-e^{-kx})\Big|_0^1-\int_0^1(-1)(-e^{-kx})\,dx$$

$$=1+\frac{1}{ke^k}-\frac{1}{k}=\frac{1}{e}, \quad (k=1 \text{ olursa}).$$

Y değişkeninin bir başka X değişkenine göre koşullu olasılığı $\mathbb{P}((x,Y)\in B) = \mathbb{P}((X,Y)\in B \mid X=x)$ olarak yazılabilir.

Ayrıca, $\mathbb{P}((X,Y)\in B) = \int_{-\infty}^{+\infty} \mathbb{P}((x,Y)\in B)\, dF_X$

$$= \int_{-\infty}^{\infty} \mathbb{P}((X,Y)\in B \mid X=x)\, dF_X$$

$$= \mathbb{E}\left[\mathbb{P}((X,Y)\in B \mid X=x)\right] \text{ olur.}$$

<u>Örnek</u>. X ve Y ortalamaları 1 olan bağımsız üssel rasgele değişkenler olsun. $z>0$ olmak üzere $\mathbb{P}(X+Y\geq z)$ olasılığını hesaplayalım.

$$\mathbb{P}(X+Y\geq z \mid X=x) = \mathbb{P}(Y\geq z-x \mid X=x)$$
$$= \mathbb{P}(Y\geq z-x,\, X=x) / \mathbb{P}(X=x)$$
$$= \mathbb{P}(Y\geq z-x)\,\mathbb{P}(X=x) / \mathbb{P}(X=x) \quad (X,Y \text{ bağımsız!})$$
$$= \mathbb{P}(Y\geq z-x) \text{ elde edilir.}$$

Şimdi, $\mathbb{P}(Y\geq z-x) = \begin{cases} e^{-(z-x)} & , z-x\geq 0 \\ 1 & , z-x<0 \end{cases}$ $\quad (k=1!)$

olduğundan

$$\mathbb{P}(X+Y\geq z) = \int_{-\infty}^{\infty} \mathbb{P}(X+Y\geq z \mid X=x)\, f_X(x)\, dx$$

$z-x\geq 0 \Updownarrow z\geq x$

$$= \int_{0}^{\infty} \mathbb{P}(X+Y\geq z \mid X=x)\, e^{-x}\, dx$$
$$= \int_{0}^{\infty} \mathbb{P}(Y\geq z-x)\, e^{-x}\, dx$$
$$= \int_{0}^{z} e^{-(z-x)} e^{-x}\, dx + \int_{z}^{\infty} e^{-x}\, dx$$
$$= \int_{0}^{z} e^{-z}\, dx - (e^{-x})\Big|_{z}^{\infty}$$
$$= z e^{-z} + e^{-z} \quad \text{olur.} \quad \blacksquare$$

<u>Alıştırma 85.</u> X $[0,1]$ aralığında, Y ise $[0,X]$ aralığında düzgün dağılımlar olsunlar. $\mathbb{E}Y$ ortalama değerini hesaplayınız.

§ <u>5.17. Stokastik Süreçler.</u>

Aşağıdaki teorem sürekli bağımsız bir $\{X_k\}$ rasgele değişken dizisinin bölgesi bu dizinin $\{F_k\}$ cdf dizisinden elde edilebileceğini gösterir.

<u>Teorem.</u> $\{F_k\}$ $\mathbb{R}$ üzerinde bir cdf dizisi olsun. Bu durumda bir $(S, \mathcal{A}, \mathbb{P})$ ölçüm uzayı ve bu uzay üzerinde tanımlı bir $\{X_k\}$ rasgele değişken dizisi vardır öyle ki X_k değişkeninin cdf'si F_k olur.

<u>Kanıtın Taslağı:</u> $\mathbb{R}^\infty = \mathbb{R}\times\mathbb{R}\times\cdots$ uzayının bir A alt kümesi aşağıdaki koşulu sağlıyorsa A'ya $\mathbb{R}^\infty$ içinde bir silindir denir: Bir $(i_1, \ldots, i_k)$ tam sayı dizisi ve $B_{i_1}, \ldots, B_{i_k}$ Borel kümeleri için

$$A = \{ s = (s_i) \mid s_{i_1} \in B_{i_1}, \ldots, s_{i_k} \in B_{i_k} \}.$$

A kümesinin ölçümünü $\mathbb{P}(A) = \mu_{F_{i_1}}(B_{i_1}) \cdots \mu_{F_{i_k}}(B_{i_k})$ olarak tanımlayalım.

$\mathbb{P}$'nin silindirik kümeler üzerinde sonlu toplamsal olduğu kolayca görülür. $\mathbb{P}$'nin silindirlerin ürettiği σ-cebirine genişletilebileceğini aşağıdaki iki alıştırmada göreceğiz.

<u>Alıştırma 86.</u> A, B ve $A\cup B$ kümelerinin üçü de silindir olsun. Eğer $A\cap B = \emptyset$ ise $\mathbb{P}(A\cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ olur.

<u>Alıştırma 87.</u> $\mathcal{J}$ bir kümeler koleksiyonu ve $\mathbb{P}$ bir olasılık ölçümü olsun. Eğer $T_1, \ldots, T_n \in \mathcal{J}$ ikili olarak ayrık kümeler ve $\bigcup_{i=1}^{n} T_i \in \mathcal{J}$ ise $\mathbb{P}(\bigcup_{i=1}^{n} T_i) = \sum_{i=1}^{n} \mathbb{P}(T_i)$ olsun.

Ayrıca verilen her $T_1, T_2, \ldots \in \mathcal{J}$ dizisi için $T_{n+1} \subseteq T_n$ ve $\bigcap_n T_n = \emptyset$ ise $\mathbb{P}(T_n) \to 0$ olduğunu kabul edelim.

Bu durumda, $\mathbb{P}$'nin $\mathcal{J}$ üzerinde ölçüm verdiğini gösteriniz: $T_i \in \mathcal{J}$, $T_i \cap T_j = \emptyset$, $\forall i \neq j$ ve $\bigcup_i T_i \in \mathcal{J}$ ise $\mathbb{P}(\bigcup_i T_i) = \sum_i \mathbb{P}(T_i)$ olur.

Sürekli olasılık yerine ayrık olasılık teorisinde çalışsaydık S'yi ayrık bir küme ve $\mathbb{R}^\infty$ yerine de S^∞ alabilirdik. ∎

Bölüm 6. Stokastik Diziler İçin Limit Teoremleri.

Bu bölümün amacı Büyük Sayılar Kanunu'nu (LLN, Law of Large Numbers) ve Merkezi Limit Teoremi'ni (CLT, Central Limit Teorem) kanıtlamak.

§6.1. Temel Tanımlar, Ortalamalar ve Varyanslar.

$X_1, \ldots, X_n$ rasgele değişkenlerin örneklem ortalaması

$$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n) \text{ olarak tanımlanır.}$$

Beklenen değer doğrusal olduğu için $\mathbb{E}(\overline{X}) = \frac{1}{n}(\mathbb{E}X_1 + \cdots + \mathbb{E}X_n)$ olur.

Varyans lineer olmasa da bağımsız değişkenler için

$$VAR[\overline{X}] = \frac{1}{n^2}(VAR[X_1] + \cdots + VAR[X_n]) \text{ olur.}$$

§6.2. Sayı ve Olay Dizileri.

İlk önce sayı dizilerinin yakınsaklığını hatırlayalım:
Bir (a_n) gerçel sayı dizisinin, bir $a \in \mathbb{R}$ sayısına yakınsak olması şöyle tanımlanır: Verilen $\epsilon > 0$ sayısı için bir N tam sayısı varsa, öyle ki, $n \geq N$ için $|a_n - a| < \epsilon$ eşitsizliği sağlansın. Bu durumda (a_n) dizisi a'ya yakınsıyor deriz ve bunu $\lim_{n\to\infty} a_n = a$ olarak yazarız.

Alıştırma 88. $\lim n \sin \frac{x}{n} = x$ olduğunu kanıtlayınız.

Ayrıca limsup ve liminf şöyle tanımlanır:

$$\limsup a_n = \overline{\lim}\, a_n = \inf_{n \geq 1} \sup_{k \geq n} a_k, \text{ ve}$$

$$\liminf a_n = \underline{\lim}\, a_n = \sup_{n \geq 1} \inf_{k \geq n} a_k.$$

Bu durumda, (a_n) dizisi yakınsaktır ancak ve ancak $\liminf a_n = \lim a_n = \limsup a_n$.

Alıştırma 89. Yukarıdaki ifadeleri kanıtlayınız.

Alıştırma 90. $(a_n)=\left(\frac{n\cos n\pi}{n+1}\right)$ dizisi için $\underline{\lim}\, a_n$ ve $\overline{\lim}\, a_n$ limitlerini hesaplayınız.

Olay Dizileri. Bir S örneklem uzayından alınan $A_1, A_2, \ldots, A_n, \ldots$ olaylar dizisi için aşağıdaki ifadeleri tanımlayalım:

$$\overline{\lim}\, A_n \doteq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \text{ve} \quad \underline{\lim}\, A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k .$$

Bunları ayrıca şöyle de yazabiliriz:

$\overline{\lim}\, A_n = \{\omega \in S \mid \omega \in A_n$ i.o. (infinitely often)$\}$ ve

$\underline{\lim}\, A_n = \{\omega \in S \mid \omega \in A_n$ eninde-sonunda (eventually)$\}$.

Definition: Bir $\{A_n\}$ olay dizisi için $\underline{\lim}\, A_n = \overline{\lim}\, A_n$ ise bu diziye yakınsaktır denir ve bu ortak limit değeri $\lim A_n$ ile gösterilir.

Hatırlatma. Eğer her A_n ölçülebilir ise $\overline{\lim}\, A_n$ ve $\underline{\lim}\, A_n$ kümeleri de ölçülebilir kümelerdir.

Alıştırma 91. (X_n) bir $(S, \mathcal{A}, P)$ olasılık uzayında tanımlı ölçülebilir rasgele değişkenler dizisi olsun. Aşağıdaki eşitlikleri kanıtlayınız:

$$\{s \mid \sup X_n(s) \leq x\} = \bigcap_{n=1}^{\infty} \{s \mid X_n(s) \leq x\} \quad \text{ve}$$

$$\{s \mid \inf X_n(s) \geq x\} = \bigcap_{n=1}^{\infty} \{s \mid X_n(s) \geq x\}.$$

Ayrıca hem $\sup X_n$ hem de $\inf X_n$ ölçülebilir rasgele değişkenlerdir. Dolayısıyla, $\overline{\lim}\, X_n$ ve $\underline{\lim}\, X_n$ de ölçülebilir değişkenlerdir.

Tanım: $\{X_n\}$ dizisi X'e a.s. ("almost surely", neredeyse kesin şekilde) eğer $P(\{s \in S \mid \lim X_n(s) = X(s)\}) = 1$ eşitliğini sağlarsa.

<u>Yardımcı Teorem</u>. $\{X_n\}$ dizisinin X'e a.s. yakınsaması için gerek ve yeter şart, verilen her $\epsilon>0$ sayı için

$P(|X_n-X|\geq\epsilon \text{ i.o.})=0$ (1) olmasıdır.

<u>Kanıt:</u> $s\in S$ alalım. Eğer $\lim X_n(s)\neq X(s)$ olması için gerek ve yeter şart şudur: Öyle bir $\epsilon>0$ vardır ki, $|X_n(s)-X(s)|>\epsilon$ koşulunu sağlayan sonsuz tane $n\geq 1$ tam sayısı vardır. Dolayısıyla,

(2) $\{s \mid \lim_n X_n(s)\neq X(s)\}=\bigcup_{\epsilon>0}\{s \mid |X_n(s)-X(s)|>\epsilon \text{ i.o.}\}$ olur.

O halde, her $\epsilon>0$ için

(3) $P(\lim_n X_n\neq X)\geqslant P(|X_n-X|>\epsilon \text{ i.o.})$.

Şimdi, eğer $\{X_n\}$ dizisi X'e yakınsıyorsa $P(\lim_n X_n\neq X)=0$ olur ve (3)'den dolayı (1) elde edilir.

Diğer yön için şu şekilde ilerleyeceğiz. İlk önce (2) nolu eşitliğin sağ tarafındaki birleşimi $\epsilon>0$ rasyonel sayılar üzerinden de alabiliriz. Bu eşitliği bozmaz. Şimdi, eğer (1) her $\epsilon>0$ (gerçel) için geçerli ise sayılabilir toplam alarak (rasyonel $\epsilon>0$ sayıları üzerinden) (2)'nin sağ tarafı ve dolayısıyla sol tarafının ölçüsü sıfırdır. Başka bir deyişle, $\{X_n\}$ dizisi X'e a.s. yakınsar.

§6.3. The Borel Cantelli Yardımcı Teoremleri ve 0-1 Kuralı

<u>Yardımcı Teorem</u> (Birinci Borel-Cantelli Teoremi)

$\{A_n\}$, $\sum_n P(A_n)<+\infty$ olan bir olay dizisi olsun. O halde,

$$P(\limsup A_n)=P(A_n \text{ i.o.})=0 \text{ olur.}$$

<u>Kanıt:</u> $\limsup A_n=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k\subseteq\bigcup_{k=n}^{\infty}A_k$, $\forall n$, olduğu için

$P(\limsup A_n)\leqslant P(\bigcup_{k=n}^{\infty}A_k)\leqslant\sum_{k=n}^{\infty}P(A_k)$ olur. Fakat, bu seri yakınsak olduğu için $\lim_{n\to\infty}\sum_{k=n}^{\infty}P(A_k)=0$ olur.

Böylece kanıt tamamlanır. ■

<u>Yardımcı Teorem</u> (İkinci Borel-Cantelli Teoremi)

$\{A_n\}$, $\sum_n P(A_n) = +\infty$ olan bağımsız olaylar dizisi olsun.

Bu durumda, $P(\limsup A_n) = P(A_n \text{ i.o.}) = 1$ olur.

<u>Kanıt:</u> $P((\limsup A_n)^c) = 0$ olduğunu kanıtlamak yeterli olacaktır.

Şimdi, $(\limsup A_n)^c = \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n\right)^c$

$$= \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_n^c \quad \text{olur.}$$

O halde, her $n \geq 1$ için $P(\bigcap_{k=n}^{\infty} A_n^c) = 0$ olduğunu göstermeliyiz.

Bunu yapmak için olayların bağımsız olduğu bilgisini kullanacağız. Analiz dersinden, $1 - x \leq e^{-x}$, $\forall x \in \mathbb{R}$, olduğunu biliyoruz. O halde,

$$P\left(\bigcap_{k=n}^{\infty} A_n^c\right) \leq P\left(\bigcap_{k=n}^{n+m} A_k^c\right) = \prod_{k=n}^{n+m} P(A_k^c)$$

$$= \prod_{k=n}^{n+m} (1 - P(A_k))$$

$$\leq \prod_{k=n}^{n+m} \exp(-P(A_k))$$

$$= \exp\left(-\sum_{k=n}^{n+m} P(A_k)\right) \text{ elde ederiz.}$$

Kabulden dolayı $m \to \infty$ giderken $\sum_{k=n}^{n+m} P(A_k) \to +\infty$ olduğunu biliyoruz.

O halde, $\exp(-\sum_{k=n}^{n+m} P(A_k)) \to 0$ olur kanıt tamamlanır. ■

<u>Hatırlatma.</u> A_n'ler bağımsız olmak kaydıyla, yukarıdaki iki sonuçtan dolayı $\limsup A_n$ ya 0 ya da 1 değeri alacaktır. İkisinin arasında başka değer alma şansı yoktur. Buna Borel'in 0-1 Kanunu denir.

Bunun oldukça ilginç sonuçları vardır. Örneğin hilesiz bir parayı sürekli attığımızı ve sonuçları yan yana

yazdığımızı düşünelim. Elde edilen Y, T dizisinin içinde TYYTY dizisinin var olup olmadığını soralım. Borel'in Kanunu'na göre bu 5 harfli kelime bu dizi içinde ya hiç yoktur ya da sonsuz defa yer alır.
Aslında ikinci ihtimal doğrudur, çünkü hiç yer almasaydı, bu diziyi 5 harfli parçalara ayırarak,

$$\underset{1.}{-----},\underset{2.}{-----},\underset{3.}{-----},\ \cdots,\underset{n.\ 5\text{-bit}}{-----},\cdots$$

hiç olmama ihtimalinin $\lim_{n\to\infty}\left(\frac{31}{32}\right)^n=0$ olduğunu görürüz.

0-1 kanunu şu şekilde genelleştirilebilir. Verilen bir A_n dizisinin kuyruğu şöyle tanımlanır:

$$\tau=\bigcap_{n=1}^{\infty}\sigma(A_n,A_{n+1},\ldots).$$

<u>Yardımcı Teorem</u> (Kolmogorov'un 0-1 Kanunu)

Eğer $\{A_n\}$ bir bağımsız olaylar dizisi ise kuyruğun içinde kalan bir $A\in\tau$ olayı için ya $P(A)=1$ ya da $P(A)=0$ olur.

<u>Uyarı</u>: Yukarıdaki Yazı-Tura örneğinde A_n olayını dizinin içindeki n'inci 5-bitlik kısmın TYYTY olmaması olayıdır.

<u>Sonuç</u>. Eğer bir X değişkeni τ göre sürekli ise (τ'nun kendisi bir σ-cebiridir) X değişkeni neredeyse kesin sabittir (almost surely constant).

Bu sonucun anlamı şudur: Kuyruk içinde kalan her olayın ihtimali ya sıfır ya da birdir. Ayrıca, kuyruk üzerinde tanımlı her ölçüm (sürekli değişken) sabittir.

Bir örnek olarak sonsuz ortalamayı ele alalım:

$$\overline{X}=\lim_{n\to\infty}\frac{1}{n}(X_1+\cdots+X_n).$$

Henüz bu limit varlığını bilmiyoruz. Fakat limit varsa $\overline{X}$ sadece kuyruğa bağlıdır, çünkü ilk n_0 terim asla $\overline{X}$'i etkilemez. O halde, yukarıdaki sonuçtan dolayı sabit olmalıdır. Böylece, her X_i'nin ortalaması μ ise $\bar{X}$'nin ortalaması da μ olmak durumunda kalır.

İlerki bölümlerde $\bar{X}$'nin varlığını kanıtlayacağız.

Sonucun Kanıtı: $X:(S,\mathcal{T},P)\to\mathbb{R}$ ölçülebilir rasgele değişken olsun. $\mathbb{R}=\bigcup_n[n,n+1]$ olduğundan en az bir n için $\mathbb{P}(X^{-1}([n,n+1]))\neq 0$ olmalı. Dolayısıyla, $\mathbb{P}(X^{-1}([n,n+1]))=1$ olur. Genellikten hiç bir şey kaybetmeden bu aralığın $[0,1]$ olduğunun kabul edelim. Sonrasında bu aralığı her seferinde ikiye bölerek $I_0=[0,1]\supseteq I_1\supset I_2\supset\cdots\supset I_n\supset\cdots$ aralıkları bulalım, öyle ki $\ell(I_n)=1/2^n$ ve $\mathbb{P}(X^{-1}(I_n))=1$ olsun. I_n'ler tıkız, $I_{n+1}\subseteq I_n$ ve $I_n\neq 0$ $\forall n$, olduğu için öyle bir $x_0\in\mathbb{R}$ vardır ki $\{x_0\}=\bigcap I_n=\lim I_n=\lim\sup I_n=\lim\inf I_n$ olur. $A_n=X^{-1}(I_n)$ dizisi de monotone olur, $\mathbb{P}$ sürekli olduğu için $\mathbb{P}(\bigcap_n A_n)=\mathbb{P}(\lim A_n)=\lim\mathbb{P}(A_n)=\lim 1=1$ elde edilir. $A\doteq\bigcap_n A_n$ olsun. Son olarak, $X(A)=x_0$ ve $\mathbb{P}(A)=1$ olduğu için X neredeyse kesin sabittir. Böylece kanıt tamamlanır. $\blacksquare$

§ 6.4. Bazı Eşitsizlikler.

Markov Eşitsizliği. X rasgele değişken ve $a>0$, $k>0$ gerçel sayılar olmak üzere
$$\mathbb{P}(|X|>a)\leqslant\frac{1}{a^k}\mathbb{E}[|X|^k]$$
eşitsizliği sağlanır.

Proof: $\mathbb{E}[|X|^k]=\int|X|^k dF\geq\int_{|X|\geq a}|X|^k dF$
$$\Rightarrow\mathbb{E}[|X|^k]\geq a^k\int_{|X|\geq a}dF=a^k\,\mathbb{P}(|X|\geq a)$$
ve böylece kanıt tamamlanır. $\blacksquare$

Özel durumda, $X=Y-\mathbb{E}Y$ ve $k=2$ alırsak
$$\mathbb{P}(|Y-\mathbb{E}Y|\geq a)\leqslant\frac{1}{a^2}\mathbb{E}[|Y-\mathbb{E}Y|^2]=\frac{1}{a^2}VAR[Y].$$

Bu eşitsizlik Chebyshev Eşitsizliği olarak bilinir.

Alıştırma 9.2. Y üssel dağılım olsun. $a>0$ sayısını yeterince büyük seçelim öyle ki sağ taraf, $\frac{1}{a^2}VAR[Y]<0.1$ olsun.

§6.5. Yakınsama Seçitleri.

Bu bölümde 4 değişik dizi yakınsaklığı ve bunların arasındaki ilişkileri göreceğiz. $\{X_n\}$ bir rasgele değişken dizisi olsun:

1) Eğer $\mathbb{P}(\lim_n X_n = X) = \mathbb{P}(\omega \mid \lim_n X_n(\omega) = X(\omega))$ ise $\{X_n\}$ dizisi X'e neredeyse kesin yakınsıyor denir (almost surely convergent).

2) Eğer $\mathbb{E}[X_n - X]^2 \longrightarrow 0$ ise $\{X_n\}$ dizisi X'e L^2'de yakınsıyor denir.

3) Eğer her $\epsilon > 0$ ise $P(|X_n - X| > \epsilon) \longrightarrow 0$ $(n \to \infty)$ ise $\{X_n\}$ dizisi X'e olasılıksal yakınsıyor denir.

4) Eğer $\lim_n \mathbb{P}(X_n \leq t) = \mathbb{P}(X \leq t)$, $\forall t \in \mathbb{R}$, oluyorsa $\{X_n\}$ dizisi X'e zayıf yakınsar denir. Burada cdf fonksiyonu $F_X(t) = \mathbb{P}(X \leq t)$'in sürekli olduğunu kabul ediyoruz.

Yardımcı Teorem. (1), (2) $\Rightarrow$ (3) $\Rightarrow$ (4).

Kanıt: (2) $\Rightarrow$ (3). Bu Chebyshev eşitsizliğinden elde edilir: Şöyle ki, her $\epsilon > 0$ için Chebyshev eşitsizliği

$$\mathbb{P}(|X_n - X| \geq \epsilon) \leq \frac{\mathbb{E}|X_n - X|^2}{\epsilon^2}$$ olduğunu söyler ve (2)'

den dolayı $n \to \infty$ giderken sağ taraf sıfıra gider. O halde sol taraf da sıfıra gider. Bu ise tam olarak (3)'tür.

(1) $\Rightarrow$ (3) Her $\epsilon > 0$ için neredeyse kesin yakınsama bize

$$\mathbb{P}\Big(\bigcup_{k \geq n} \{\omega \mid |X_k - X| \geq \epsilon\}\Big) \longrightarrow 0 \quad (n \to \infty)$$ olduğunu garanti

eder. Dolayısıyla, $\mathbb{P}(\{\omega \mid |X_k - X| \geq \epsilon\}) \leq \mathbb{P}\Big(\bigcup_{k \geq n} \{\omega \mid |X_k - X| \geq \epsilon\}\Big) \to 0$

olduğu için $\mathbb{P}(\{\omega \mid |X_k - X| \geq \epsilon\}) \to 0$ $(k \to \infty)$ elde edilir.

(3) $\Rightarrow$ (4) $F(t) \doteq \mathbb{P}(X \leq t)$ ve $F_n(t) \doteq \mathbb{P}(X_n \leq t)$ olarak tanımlansın.

O zaman, $F(t-\epsilon) = \mathbb{P}(X \leq t-\epsilon)$
$$= \mathbb{P}(X \leq t-\epsilon, X_n \leq t) + \mathbb{P}(X \leq t-\epsilon, X_n > t)$$
$$\leq \mathbb{P}(X_n \leq t) + \mathbb{P}(X_n - X \geq \epsilon)$$

[Burada, $X \leq t-\epsilon$ ve $X_n > t$ ise $X_n - X \geq t-(t-\epsilon) = \epsilon$ olmasını kullandık.]

(3) kabulünden dolayı $n \to \infty$ durumunda sağ taraftaki ikinci terim sıfıra gider. Dolayısıyla, $F(t-\epsilon) \leq \liminf F_n(t)$ elde edilir. Benzer şekilde, $F(t+\epsilon) \geq \limsup F_n(t)$ olur.

O halde, her $\epsilon > 0$ için

$$F(t-\epsilon) \leq \liminf F_n(t) \leq \limsup F_n(t) \leq F(t+\epsilon) \text{ olur.}$$

X_n ve dolayısıyla $F_n(t)$ sürekli olduğu için, $\epsilon \to 0$ giderken limit olarak $F_n(t) \to F(t)$ elde ederiz. ∎

__Alıştırma 93.__ $X_n \to X$ olasılıksal yaklaşsın ve f fonksiyonu da düzgün sürekli olsun. O zaman, $f(X_n) \to f(X)$ olasılıksal yakınsar.

__Alıştırma 94.__ X_n $\{1, 2, \ldots, n\}$ kümesi üzerinde ayrık ve düzgün dağılım olsun ($X_n(k) = 1/n$, $k = 1, \ldots, n$, ve $X_n(k) = 0$ aksi halde $X_n(k) = 0$ olsun).
Bu durumda, her $0 \leq x \leq 1$ için $\lim_n \mathbb{P}(n^{-1} X_n \leq x) = x$ olur.

§ 6.6. Büyük Sayıların Zayıf Kuralı.

$X_1, X_2, \ldots, X_n, \ldots$ bir IID rasgele değişken dizisi olsun. Örneğin bir deneyin sürekli tekrarlarının sonuçları olsunlar (Yazı-Tura gibi). μ ve σ^2 sayıları ortak ortalamaları ve varyansları olsun. Her ikisinin de sonlu sayılar olduğunu kabul edelim. Değişkenler ayrıca bağımsız olduğu için $\mathrm{COV}(X_i, X_j) = 0$ olur (yada $\mathbb{E}[XY] = (\mathbb{E}X)(\mathbb{E}Y)$).

__Teorem.__ Yukarıdaki dizi için $S_n = X_1 + \cdots + X_n$ değişkenini tanımla alım. Bu durumda S_n/n dizisi hem L^2 hem de olasılıksal olarak $S_n/n \to \mu$ sayısına yakınsar.

Kanıt: Bir önceki yardımcı teoremden dolayı sadece L^2'de yakınsadığını göstermek yeterlidir.

$EX_n = \mu$ olduğu için $\mathbb{E}S_n = n\mu$ olur. Buradan da

$$\mathbb{E}\left[\frac{S_n}{n}-\mu\right]^2 = VAR\,\frac{S_n}{n} = \frac{1}{n^2}VAR[X_1+\cdots+X_n]$$
$$= \frac{1}{n^2}\left(VAR[X_1]+\cdots+VAR[X_n]\right)$$
$$= \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} \text{ elde edilir.}$$

$\sigma^2/n \to 0$ olduğundan $\frac{S_n}{n} \to \mu$ $(n\to\infty)$ L^2-de yakınsar. ∎

Alıştırma 95. (Shannon's Teorem)

X_i'ler $\{1,\ldots,r\}$ değerlerini sırasıyla $p_1,\ldots,p_r$ olasılıklarıyla alan IID rasgele değişkenler olsun. $i_j \in \{1,\ldots,r\}$ olmak üzere $p_n(i_1,\ldots,i_n) = p_{i_1}\cdots p_{i_n}$ olsun. $Y_n \doteq p_n(X_1,\ldots,X_n)$ olarak tanımlansın. B

$$-\frac{1}{n}\log Y_n \to -\sum_{i=1}^{n} p_i \log p_i \text{ olasılıksal olarak yakınsadığını gösteriniz.}$$

Kanıt: $\frac{1}{n}\log Y_n(\omega) = \frac{1}{n}\log p_n(X_1(\omega),\ldots,X_n(\omega))$

$$= \frac{1}{n}\log p_{i_{X_1(\omega)}}\cdots p_{i_{X_n(\omega)}}$$
$$= \frac{1}{n}\sum_{k=1}^{n}\log p_{i_{X_k(\omega)}} \to \mu \text{ olur, öyle ki}$$

μ $\log p_{X_k}$ değişkenlerinin ortak ortalamasıdır. ∎

Yukarıdaki teoremin daha genel ama zayıf bir halini verelim.

Teorem. X_i, $\mu = \mathbb{E}[X_i] < +\infty$, $\forall i$, olan IID rasgele değişkenler olsun. Bu durumda, $S_n/n \to \mu$ olasılıksal olarak yakınsar.

§6.7. Büyük Sayıların Kuvvetli Kanunu (Strong Law of Large numbers, SLLN)

Teorem. X_n, $\mu = \mathbb{E}[X_n]$, $\forall n$ olan IID rasgele değişkenler olsun. Her $p=1,2,3,4$ ve $n=1,2,\ldots$ için $\mathbb{E}[|X_n|^p] < +\infty$ olduğunu kabul edelim. $S_n = X_1+\cdots+X_n$ olsun. O zaman $S_n/n \to \mu$ a.s. $(n\to\infty)$ olur.

Kanıt: X_n değişkenini $X_n - \mu$ ile değiştirerek $\mu = 0$ olduğunu kabul edebiliriz. Dolayısıyla, $\frac{1}{n} S_n \to 0$ a.s. olduğunu göstermeliyiz. Daha doğrusu

$\{\omega \mid \lim_n n^{-1} S_n(\omega) = 0\}$ kümesinin ölçümünün 1 olduğunu göstermek yeterlidir. §6.2'nin sonundaki yardımcı teoreminden dolayı, her $\epsilon > 0$ için $\mathbb{P}(|\frac{1}{n} S_n(\omega)| > \epsilon \text{ i.o.}) = 0$ olduğunu göstermeliyiz.

$A_n = \{\omega \mid |n^{-1} S_n(\omega)| \geq \epsilon\}$ olarak tanımlansın. O zaman, $\{\omega \mid |n^{-1} S_n(\omega)| \geq \epsilon \text{ i.o.}\}$ olayı $\limsup A_n$'dir. O halde, $\mathbb{P}(\limsup A_n) = 0$ olduğunu göstermeliyiz. Bunun için ise, Borel-Cantelli Teoremi'nden dolayı, $\sum_k \mathbb{P}(A_k)$ serisinin yakınsak olduğunu göstermek yeterlidir.

Kuvvet 4 olan Markov eşitsizliğinden

$\mathbb{P}(A_k) = \mathbb{P}(|S_k| \geq k\epsilon) \leq \frac{\mathbb{E} S_k^4}{k^4 \epsilon^4}$ elde ederiz. Diğer yandan,

$\mathbb{E} S_k^4 = \sum_{a,b,c,d} \mathbb{E}[X_a X_b X_c X_d]$ (her endeks $1, \dots, k$ arasındaki tüm değerleri alır) vardır.

Değişkenler bağımsız ve $\mathbb{E}[X_i] = 0$ $\forall i$, olduğu için a, b, c, d tamamen farklılar ise $\mathbb{E}[X_a X_b X_c X_d] = 0$ olur. Diğer yandan $a = b \neq c \neq d$ ise yine $\mathbb{E}[X_a X_b X_c X_d] = \mathbb{E} X_a^2 \, \mathbb{E} X_c \, \mathbb{E} X_d = 0$ olur. Buradan

$$\mathbb{E} S_k^4 = \sum_a \mathbb{E}[X_a^4] + 3 \sum_{a \neq b} \mathbb{E}[X_a^2] \, \mathbb{E}[X_b^2]$$

$$= k \, \mathbb{E}[X^4] + 3k(k-1)(\mathbb{E}[X^2])^2 \text{ olur.}$$

k değişkeninin fonksiyonu olarak görüldüğünde öyle bir $C > 0$ vardır ki $\mathbb{E}[S_k^4] \leq C k^2$ olur.

Dolayısıyla, $\mathbb{P}(A_k) \leq \frac{\mathbb{E}[S_k^4]}{k^4 \epsilon^4} \leq \frac{C k^2}{k^4 \epsilon^4} = \frac{C}{k^2 \epsilon^4}$ elde ederiz.

Son olarak $\sum \frac{1}{k^2} < +\infty$ olduğu için $\sum_k \mathbb{P}(A_k) < +\infty$ olur ve kanıt tamamlanır. ■

Aslında bu sonuç daha genel durumda da geçerlidir.

Teorem (SLLN) $\mathbb{E}[X_i] < +\infty$ ise SSLN geçerlidir.

Alıştırma 96. X_n dizisi şöyle tanımlansın: $X_0 = 1$ ve X_{n+1} $(n \geq 0)$ $[0, X_n]$ aralığındaki düzgün dağılım ile rasgele seçilsin. Bu durumda $\frac{1}{n}\log X_n$ dizisinin neredeyse kesin şekilde bir sabite yakınsadığını gösteriniz.

Çözüm: Y_n $[0,1]$ aralığındaki düzgün rasgele değişken olsun.

Bu durumda, $X_n = Y_1 \cdots Y_n$ olur. $Z_n = \log X_n = \sum_{k=1}^{n} \log Y_k$ olarak tanımlansın.

Dolayısıyla, Z_n/n dizisi a.s. şekilde bir μ sayısına yakınsar, öyle ki $\mu = \mathbb{E}\log Y_n = \log 1/2$. Başka bir deyişle $\frac{1}{n}\log X_n \to \log 1/2$ a.s. ∎

§6.8. SSLN'nin Uygulamaları.

X_i pozitif değerli IID rasgele değişken dizisi ve $T_k = X_1 + \cdots + X_k$ olsun. Örneğin, X_i bir elektrik ampulünün ömrü olabilir öyle ki, bozulan bozulmaz ilk fırsatta değiştirilsin. Bu durumda T_k ampulün k'ıncı defa bozulana kadar ki süre olur.
Ayrıca, $N_t = \sup\{n \mid T_n \leq t\}$ t anına kadar bozulan ampul sayısı olsun.

Yardımcı Teorem. $\mathbb{E}X_i = \mu < +\infty$ olsun. Bu durumda
$$N_t/t \to 1/\mu \quad \text{a.s. yakınsar.}$$

Hatırlatma. Bu sonuç şu beklentiyi doğrular: Ampullerin bozulma hızı ampulün ortalama ömrünün tersidir.

Kanıt: SLLN'den dolayı $T_n/n \to \mu$ a.s. yakınsar. Ayrıca tanımından dolayı
$$T_{N_t} \leq t < T_{N_t+1} \text{ olur.}$$

O halde,
$$T_{N_t}/N_t \leq t/N_t < T_{N_t+1}/N_t = \frac{T_{N_t+1}}{N_t+1} \cdot \frac{N_t+1}{N_t} \quad (*)$$
olur.

Her $T_n < +\infty$ olduğu için $t \to \infty$ giderken $N_t \to \infty$ gider.

SLLN'den dolayı $P(B)=1$ olacak şekilde bir B olayı vardır öyle ki $t\to\infty$ giderken $N_t\to\infty$ gider ve $n\to\infty$ giderken $T_n/n \to \mu$ gider.

O halde, bu B olayında, $T_{N_t}/N_t \to \mu$ ve $N_{t+1}/N_t \to 1$ $(t\to+\infty)$ olur.

Son olarak (*) satırından dolayı

$$\mu \le \lim_{t\to\infty} t/N_t \le \mu.1 \Rightarrow t/N_t \to \mu \text{ a.s. } (t\to\infty) \text{ elde edilir}$$

BÖLÜM 7. Moment Üreteç Fonksiyonları.

§7.1. Rasgele Değişkenin Momenti.

X rasgele değişkeninin k'ıncı momenti $\mathbb{E}X^k=\int x^k dF(x)$ integrali olarak tanımlanır.

Alıştırma 97. Düzgün dağılımın ve üssel dağılımın momentlerini hesaplayınız.

§7.2. Moment Üreteç Fonksiyonları.

Bir X rasgele değişkeninin Moment Üreteç Fonksiyonu (mgf - Moment Generating function) şu şekilde tanımlanır:

$$M(t)=\mathbb{E}[e^{tX}]=\int_{\mathbb{R}} e^{tX} dF(x).$$

Eğer $M(t)$ tanımlıysa, teknik ayrıntıları şimdilik göz ardı ederek, şu şekilde türevini alalım:

$$\frac{d}{dt}M(t)=\frac{d}{dt}\int_{\mathbb{R}} e^{tX} dF(x)=\int_{\mathbb{R}}\frac{d}{dt}(e^{tX})\,dF_X=\int_{\mathbb{R}} xe^{tx} dF(x)$$

$$=\mathbb{E}[xe^{tX}].$$

Dolayısıyla, $\frac{d}{dt}M(t)\Big|_{t=0}=\mathbb{E}[X]$ olur. Benzer şekilde $\mathbb{E}X^k = d^k/dt^k\, M(t)\Big|_{t=0}$ olur, $\forall k>0$.

Örneğin, $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ ise mgf $M(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx$ olur. Üssel ifadeyi tam kareye tamamlayarak

$$M(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t^2/2} e^{-(x-t)^2/2} dx = e^{t^2/2} \text{ olur.}$$

Alıştırma 98. pdf'si $f(x) = ke^{-kx}$ olan üssel dağılımı için $M(t)$ ancak $t<k$ olduğunda tanımlıdır. Bu dağılım için $M(t)$'yi ve bulunuz ve n'inci momentin $n!/k^n$ olduğunu gösteriniz.

Teknik Ayrıntılar:

$M(t) = \mathbb{E}[e^{tX}] = \int_{\mathbb{R}} e^{tx} dF(x)$ fonksiyonunun bir $(-t_0, t_0)$ aralığında tanımlı olduğunu kabul edelim.

Diğer yandan Taylor Teoremi bize şunu verir:

$$|e^{tx}| = \left|\sum_k (tx)^k/k!\right| \leq \sum_k \frac{|tx|^k}{k!} = e^{|tx|}.$$

Dolayısıyla, $e^{|tx|}$ bir $(-t_0, t_0)$ aralığında integrallenebilir ise Sınırlandırılmış Yakınsaklık Teoremi (Dominated Convergence Theorem) sayesinde şu eşitlikleri elde ederiz:

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \int_{\mathbb{R}} x^k dF = \sum_{k=0}^{\infty} \int_{\mathbb{R}} \frac{(tx)^k}{k!} dF = \int_{\mathbb{R}} \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} dF = \int_{\mathbb{R}} e^{tx} dF = M(t).$$

Aslında bu $M(t)$'nin Taylor açılımından başka birşey değildir ve Taylor katsayıları ise $\int_{\mathbb{R}} x^k dF = \mathbb{E}[X^k]$ olur. Açılım $(-t_0, t_0)$ aralığında geçerli olduğu için $\mathbb{E}[X^k] = \frac{d^k}{dt^k} M(t)\Big|_{t=0}$ olur.

Dolayısıyla, gerek duyulan esas nokta sıfırı içeren bir aralıkta $M(t)$'nin varlığıdır.

Alıştırma 99. f'in sürekli olduğunu ve bazı $s>1$, $c, K \in \mathbb{R}$ için $f(x) \geq c|x|^{-s}$, $\forall |x| \geq K$, eşitsizliğinin sağlandığını düşünelim. Bu durumda $M(t)$'nin sadece $t=0$ var olduğunu gösteriniz. Peki momentler var mıdır?

Alıştırma 100. Bazı $t_1 < t_2$ için $M(t_1) < +\infty$ ve $M(t_2) < +\infty$ olduğunu kabul edelim. Bu durumda, her $t \in [t_1, t_2]$ için $M(t) < +\infty$ olduğunu gösteriniz.

Şimdi de bağımsız $X_1, \ldots, X_n$ değişkenlerinin $M_1, \ldots, M_n$ moment üreteç fonksiyonlarının hepsinin bir $(-t_0, t_0)$ aralığında varolduğunu kabul edelim. Bu durumda $X_1 + \ldots + X_n$'nin mgf'nu da vardır ve

$$\mathbb{E}[e^{t(X_1+\ldots+X_n)}] = \mathbb{E}[e^{tX_1} \cdots e^{tX_n}]$$
$$= \mathbb{E}[e^{tX_1}] \cdots \mathbb{E}[e^{tX_n}]$$
$$= M_1(t) \cdots M_n(t_n) \text{ olur.}$$

Alıştırma 101. $[0, L]$ aralığında tanımlı n tane IID düzgün rasgele değişkenin toplamının mgf'sini hesaplayınız.

Bölüm 8. Merkezi Limit Teoremi (Central Limit Theorem)

Merkezi Limit Teoremi zayıf ve kuvvetli büyük sayılar teoremlerindeki yakınsamaların türü ve hızı hakkında daha kesin bilgiye ulaşmamızı sağlar.

Bir Z rasgele değişkenin $\mathbb{P}(a < Z < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx, \forall a, b,$ eşitliğini sağlıyorsa bu değişkene standart normal (dağılım) değişken denir. Standart normal dağılımın mgf'sinin $M_Z(t) = e^{t^2/2}$ olduğunu hatırlayalım.

Bir $\{X_n\}$ rasgele değişken dizisi ve X rasgele değişkeni için $\mathbb{P}(a \leq X_n \leq b) \longrightarrow \mathbb{P}(a \leq X \leq b)$, $n \to \infty$, oluyorsa $\{X_n\}$ dizisi X'e dağılımsal olarak veya "zayıf" olarak yakınsıyor denir.

Teorem (CLT) $X_1, X_2, \ldots$ $\mathbb{E}[X_i] = \mu < +\infty$ ve $VAR[X_i] = \sigma^2 < +\infty$ olan bir IID rasgele değişken dizisi olsun. $S_n = X_1 + \ldots + X_n$ ise $(S_n - n\mu)/(\sigma\sqrt{n})$ dizisi standart normal dağılıma dağılımsal olarak yakınsar.

Kanıt: $Y_n = \frac{X_n - \mu}{\sigma}$ olarak tanımlansın. Bu durumda Y_n'nin ortalaması 0 ve varyansı 1 olur.

Ayrıca, $T_n = Y_1 + \dots + Y_n = \frac{S_n - n\mu}{\sigma}$ olsun. O halde, $T_n/\sqrt{n}$'nin standart normal dağılıma zayıf şekilde yakınsadığını göstermeliyiz.

$M(t) = \mathbb{E}[e^{tY}]$ ve $M_n(t) = \mathbb{E}[e^{tT_n/\sqrt{n}}]$ olarak tanımlansın. T_n bağımsız değişkenlerin toplamı olduğu için

$$M_n(t) = \mathbb{E}[e^{t(Y_1+\dots+Y_n)/\sqrt{n}}] = \mathbb{E}[e^{tY/\sqrt{n}}]^n = M(t/\sqrt{n})^n$$ olur.

Sabit bir $t \in \mathbb{R}$ değeri için $t/\sqrt{n} \to 0$ $(n \to \infty)$ olduğundan e^{tY}'nin Taylor açılımından

$$M(t/\sqrt{n}) = \mathbb{E}\left[1 + \frac{tY}{\sqrt{n}} + \frac{t^2Y^2}{2n} + \dots\right] = 1 + \frac{t^2}{2n} + R_n(t)$$ elde ederiz.

Bu hesapta $\mathbb{E}[Y]=0$, $\mathbb{E}[Y^2]=1$ ve $R_n(t)$ serinin kalan terimidir.

O halde, $M_n(t) = \left(1 + \frac{t^2}{2n} + R_n(t)\right)^n = \left(1 + \frac{t^2}{2n}\right)^n \left(1 + R_n(t)/\left(1+\frac{t^2}{2n}\right)\right)^n$

olur. Diğer yandan, $\lim_{n\to\infty}\left(1+\frac{t^2}{2n}\right)^n = e^{t^2/2}$, standart normal dağılımın mgf'si olur.

O halde, kanıtı tamamlamak için şunları göstermeliyiz:

Adım 1. Her t için $\lim_{n\to\infty}\left(1 + R_n(t)/(1+t^2/2n)\right)^n = 1$.

Adım 2. $M_n(t)$'lerin bir $M_\infty(t)$'a noktasal yakınsaması $F_n(x)$'lerin bir $F_\infty(x)$'e noktasal yakınsamasını sağlar.

Adım 3. mgf fonksiyonu $e^{t^2/2}$ olan tek bir cdf fonksiyonu vardır. Dolayısıyla, $F_n(x)$'lerin cdf fonksiyonları standart normal dağılımın cdf'sine noktasal olarak yakınsar.

Bu adımlarda birçok sorunlu nokta bulunur. Örneğin, Y değişkeninin mdf fonksiyonu var olmayabilir. Tüm bu teknik sorunları "karmaşık integraller" ile aşabiliriz:

$$\phi(t) = \mathbb{E}[e^{itY}] = \int e^{itY}\, dF.$$

§8.1. Merkezi Limit Teoreminin Uygulamaları.

CLT büyük n değerleri için olasılıkların yaklaşık değerlerini verir. Aşağıda bunun bazı örneklerini göreceğiz.

Binomsal Dağılım. Bir Yazı-Tura deneyinin sonuçları X_k dizisi ile verilsin. Hilesiz olduğu kabul edilen para yazı gelirse $X=0$, tura gelirse $X=1$ değerini alsın. O halde, $S_n = X_1 + \cdots + X_n$ n atış sonra turaların toplam sayısını gösterir. $\mathbb{E}X = 1/2$ ve $VAR[X] = 1/4$ olduğu için CLT bize aşağıdaki ifadenin

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \sqrt{\frac{4}{n}}\left(S_n - \frac{n}{2}\right) \longrightarrow Z$$ yakınsadığını söyler.

Burada Z standart normal dağılımdır. O halde

$$\mathbb{P}\left(a \le \sqrt{\frac{4}{n}}\left(S_n - \frac{n}{2}\right) < b\right) \longrightarrow \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}\,dt, \ \forall a<b,$$

yakınsar.

Alıştırma 102. X_i $[0,1]$ aralığındaki düzgün dağılım olsun, $i=1,\ldots,10$. CLT'yi kullanarak $X_1 + \cdots + X_{10} \ge 7$ olma olasılığını hesaplayınız.

Çözüm. $\mu = \mathbb{E}[X_i] = 1/2$, $VAR[X_i] = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2$

$$= \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

O halde, $\mathbb{P}\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} < b\right) \longrightarrow \int_a^b e^{-t^2/2}\,dt$ sonucundan

$$\mathbb{P}\left(a \le \frac{S_{10} - 10(1/2)}{1/12\,\sqrt{10}} < \infty\right) \longrightarrow \int_a^\infty e^{-t^2/2}\,dt$$

$$\mathbb{P}\left(5 + \frac{\sqrt{10}\,a}{12} \le S_{10} < \infty\right) \longrightarrow \int_a^\infty e^{-t^2/2}\,dt.$$

$5 + \frac{\sqrt{10}a}{12} = 7 \Rightarrow a = 24/\sqrt{10}$ olur ve dolayısıyla

$\mathbb{P}(7 \le S_{10} < \infty) \sim \int_{24/\sqrt{10}}^\infty e^{-t^2/2}\,dt$ elde edilir.

Alıştırma 103. X ve Y ortalamaları sırasıyla λ ve μ olan bağımsız Poisson rasgele değişkenler olsun. $X+Y$'nin ortalamasının $\lambda+\mu$ olduğunu gösteriniz. Daha sonra bu sonucu ve CLT'yi kullanarak aşağıdaki limiti hesaplayınız:

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

Çözüm. X'in Poisson olması demek $Ran(X)=\{0,1,2,\ldots\}$ ve $P(X=k)=e^{-\lambda}\lambda^k/k!$ demektir (λ X'in ortalama değeri).

Gerçekten de, $E[X]=\sum_{k=0}^{\infty} k\,P(X=k)=\sum_{k=0}^{\infty} k\,e^{-\lambda}\lambda^k/k!$

$$= e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda \text{ olur.}$$

Diğer yandan $VAR[X]=E[X^2]-(E[X])^2$

$= E[X^2]-\lambda^2$ olur. Burada

$$E[X^2]=\sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=0}^{\infty} k^2 e^{-\lambda}\lambda^k/k! = \lambda^2+\lambda \text{ elde edilir.}$$

$\Rightarrow VAR[X]=\lambda^2+\lambda-\lambda^2=\lambda$ olur.

Şimdi $X+Y$'ye dönelim. $Ran(X+Y)=\{0,1,2,\ldots\}$ olur.

Ayrıca, $P(X+Y=k)=\sum_{i=0}^{k} P(X=i, Y=k-i)$

$$= \sum_{i=0}^{k} P(X=i)\,P(Y=k-i) \quad (X, Y \text{ bağımsız!})$$

$$= \sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^i}{i!}\,\frac{e^{-\mu}\mu^{k-i}}{(k-i)!}$$

$$= e^{-(\lambda+\mu)}\sum_{i=0}^{k}\frac{\lambda^i \mu^{k-i}}{i!\,(k-i)!}$$

$$= e^{-(\lambda+\mu)}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!\,(k-i)!}\lambda^i\mu^{k-i}$$

$$= e^{-(\lambda+\mu)}\frac{1}{k!}(\lambda+\mu)^k.$$

Dolayısıyla, $X+Y$ ortalaması $\lambda+\mu$ olan bir Poisson rasgele değişkendir.

Şimdi her X_i'nin ortalaması $\lambda=1$ olan bağımsız Poisson değişkenler olduğunu kabul edelim. O halde, her birinin varyansı da 1'dir ve $S_n = X_1 + \cdots + X_n$ de ortalaması n olan bir Poisson değişkendir. Bu durumda CLT bize şunu verir: $\frac{S_n - n}{\sqrt{n}} \to Z$ veya

$$P\left(a \leq \frac{S_n-n}{\sqrt{n}} < b\right) \sim \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\,dx.$$

$$\Rightarrow P\left(0 \leq \frac{S_n-n}{\sqrt{n}} < \infty\right) \sim \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/2}\,dx = 1 \text{ olur.}$$

Diğer yandan, $P\left(0 \leq \frac{S_n-n}{\sqrt{n}} < \infty\right) = P(n \leq S_n < \infty)$

$$\Rightarrow \sum_{k=n}^{\infty} P(S_n = k) \sim 1, \quad \forall n >> 1.$$

$$\Rightarrow \sum_{k=n}^{\infty} e^{-n} \frac{n^k}{k!} \sim 1, \quad \forall n >> 1, \text{ elde edilir.}$$

Ayrıca, $\sum_{k=0}^{\infty} e^{-n} \frac{n^k}{k!} = e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} = e^{-n} e^{n} = 1$'dir.

O halde, $$e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \sim 1 - \sum_{k=n+1}^{\infty} e^{-n} \frac{n^k}{k!}$$
$$\sim 1 - \Bigg(\underbrace{\sum_{k=n}^{\infty} e^{-n} \frac{n^k}{k!}}_{\sim 1} - \frac{e^{-n} n^n}{n!} \Bigg)$$
$$\sim 1 - 1 + \frac{e^{-n} n^n}{n!} = \frac{e^{-n} n^n}{n!} \sim \frac{1}{\sqrt{2\pi n}} \text{ olur.}$$

Buradaki en son limit değeri Stirling Formülü olarak bilinir.

Bu bölümü bir başka örnek ile bitirelim.

Örnek: Matematik Bölümü Kantini'ne hergün yaklaşık 100 müşteri gelir. Herhangi bir müşteri 1/2 olasılıkla hiç elmalı kurabiye almazken, 1/3 olasılıkla 1 ve 1/6 olasılıkla da 2 kurabiye alır. Buna göre kantin sahibi gün boyunca kurabiye miktarının %95 olasılıkla yetması için kaç kurabiye getirmelidir.

Çözüm. X_i i'inci müşterinin satın aldığı elmalı kurabiye sayısı olsun. Ayrıca $Y = X_1 + \dots + X_{100}$ olarak tanımlansın.

O halde, $\mathbb{E}[X_i] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} = 2/3$ ve dolayısıyla

$$\mathbb{E}[Y] = 100\, \mathbb{E}[X] = 200/3 \text{ olur.}$$

Yine

$$\begin{aligned} VAR[X_i] &= \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 \\ &= 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{3} + 2^2 \frac{1}{6} - (2/3)^2 \\ &= \frac{1}{3} + \frac{2}{3} - \frac{4}{9} = \frac{5}{9} \text{ olur.} \end{aligned}$$

$$\Rightarrow VAR[Y] = 100\, VAR[X] = 500/9.$$

$Y = S_{100}$, $\mu = 2/3$, $\sigma = \sqrt{5}/3$ olmak üzere CLT bize

$$\mathbb{P}(Y < a) = \mathbb{P}\left(\frac{S_{100} - 100\mu}{\sigma\sqrt{100}} < \frac{a - 100\mu}{\sigma\sqrt{100}} \right)$$

$$\sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(a-\mu)/10\sigma} e^{-x^2/2}\, dx$$

Bizim $\mathbb{P}(Y < a) \simeq 0.95$ olacak şekilde a sayısını bulmamız gerekiyor.

$$\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(a-100\mu)/10\sigma} e^{-x^2/2}\, dx \simeq 0.95$$

$$\Rightarrow \frac{a - 100\mu}{10\sigma} \simeq 1.6449 \Rightarrow a \sim 79 \text{ elde edilir.}$$

Örnek (Ortalama Polinom)

Bu örnekte iki farklı ölçüm kullanarak derecesi sabit polinomların ortalamalarını hesaplayacağız.

A) $(1+x+y)^d = \sum_{j_1+j_2\le d} \binom{d}{j_1, j_2} x^{j_1} y^{j_2} = \sum_{j_1+j_2\le d} \frac{d!\, x^{j_1} y^{j_2}}{j_1!\, j_2!\, (d-j_1-j_2)!}$

açılımını göz önüne alarak d'inci dereceden polinomun genel halini

$$f(x,y) = \sum_{j_1+j_2\le d} a_j \sqrt{\binom{d}{j}}\; x^{j_1} y^{j_2}, \quad j=(j_1,j_2)$$

olarak alalım. Burada polinomu belirleyen a_j katsayılarını tipi $N(0,1)$ olan normal dağılım olarak alalım (ortalaması 0, varyansı 1).

O halde, $\mathbb{E}[a_j]=0$, $\forall j$ olur. Dolayısıyla, $\mathbb{E}[f(x,y)]=0$ olur.

Peki $\mathbb{E}[f^2(x,y)]$ nedir?

$$f^2(x,y) = \sum_{\substack{j_1+j_2\le d \\ j_1'+j_2'\le d}} a_j a_{j'} \sqrt{\binom{d}{j}} \sqrt{\binom{d}{j'}}\; x^{j_1+j_1'} y^{j_2+j_2'}$$

Tüm değişkenler bağımsız olduğu için $j\neq j'$ ise

$\mathbb{E}[a_j a_{j'}] = \mathbb{E}[a_j]\,\mathbb{E}[a_{j'}] = 0$ ve $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2}\, dt = 1$

olduğu için $\mathbb{E}[a_j^2]=1$ olur. [Burada, $M(t)=e^{t^2/2}$ olduğu için $\mathbb{E}[a_j]=M'(0)=0$ ve $\mathbb{E}[a_j^2]=M''(0)=1$ olur.]

Son olarak $\mathbb{E}$ doğrusal olduğu için

$$\mathbb{E}[f^2(x,y)] = \sum_{j} \mathbb{E}[a_j^2] \binom{d}{j} x^{2j_1} y^{2j_2}$$

$$= \sum_{j_1+j_2\le d} \binom{d}{j} x^{2j_1} y^{2j_2}$$

$$= (1+x^2+y^2)^d \text{ olur.}$$

Başka bir değişle derecesi $\le d$ olan polinomların karelerinin ortalaması $(1+x^2+y^2)^d$ olur.

B) $f(x,y) = \sum_{i,j\le d} a_{ij}\, x^i y^j$ genel polinomunu ele alalım.

a_{ij}'ler $N(0,1)$ tipinde normal dağılımlı rasgele değişken-olsunlar. Yine $\mathbb{E}[f^2(x,y)]$ ortalama değerini hesaplayalım.

<u>Çözüm.</u> Yine $\mathbb{E}[a_{ij}]=0$ olduğundan $\mathbb{E}[f(x,y)]=0$ elde edilir. Ayrıca $\mathbb{E}[a_{ij}^2]=1$ olur (önceki durumunda olduğu gibi).

Şimdi $f^2(x,y)=\sum_{i,j}\sum_{k,l} a_{ij}a_{kl}\, x^{i+k}\, y^{j+l}$ olduğu için

$$\mathbb{E}[f^2(x,y)] = \sum_{0\le i,j\le d} \mathbb{E}[a_{ij}^2]\, x^{2i}\, y^{2j}$$

$$= \sum_{0\le i,j\le d} x^{2i}y^{2j} = \frac{(1-x^{2d+2})(1-y^{2d+2})}{(1-x^2)(1-y^2)} \text{ olur.}$$

$$\partial/\partial x\ \mathbb{E}[f^2(x,y)] = \partial/\partial x \Big(\qquad \Big) \text{ yazarsak}$$

$$\mathbb{E}[f f_x] = \frac{1}{2}\frac{\partial}{\partial x}\left(\frac{(1-x^{2d+2})(1-y^{2d+2})}{(1-x^2)(1-y^2)}\right) \text{ olur.}$$

§8.2. LNN'de Yakınsamanın Hızı.

Bu bölümde S_n/n'nin limitine (ortalama değere) üssel bir hızla yakınsadığını göreceğiz.

<u>Yardımcı Teorem.</u> $X_1, X_2, \dots$ ortalama değeri μ olan bir IID rasgele değişken dizisi olsun. Bazı $a,b>0$ ve her $t\in(-a,b)$ için $M_{X_i}(t)<+\infty$ olsun. O zaman

$$P\left(\frac{S_n}{n}\ge \mu+\epsilon\right)\le e^{-n\lambda} \text{ olur, öyle ki}$$

$$e^{-\lambda}=\inf_{0<s<b}\left\{e^{-s(\mu+\epsilon)}M_{X_i}(s)<1\right\}\text{'dir.}$$

<u>Kanıt:</u> Her $0<s<b$ ve Y rasgele değişkeni için

$$P(Y\ge 0)=P(e^{sY}\ge 1)\le \mathbb{E}[e^{sY}]/1 = M_Y(s) \quad \text{(Markov eşitsizliği.)}$$

elde edilir. Bu eşitsizlik her $0<s<b$ için geçerli olduğundan $P(Y\ge 0)\le \inf_{0<s<b} M_Y(s)$ olur.

Şimdi $Y_i = X_i - \mu - \epsilon$ alalım. O zaman

$$\begin{aligned} M_{Y_i}(s) &= \mathbb{E}\left[e^{sY_i}\right] \\ &= \mathbb{E}\left[e^{s(X_i-\mu-\epsilon)}\right] \\ &= \mathbb{E}\left[e^{sX_i}\, e^{-s(\mu+\epsilon)}\right] \\ &= e^{-s(\mu+\epsilon)}\, \mathbb{E}\left[e^{sX_i}\right] \\ &= e^{-s(\mu+\epsilon)}\, M_{X_i}(s) \text{ olur.} \end{aligned}$$

Dolayısıyla, $\mathbb{P}\left(\frac{S_n}{n} \geq \mu+\epsilon\right) = \mathbb{P}(Y_1+\cdots+Y_n \geq 0)$
$$\leq \inf_{0<s<b} M_{Y_1+\cdots+Y_n}(s)$$

$$\Rightarrow \mathbb{P}\left(\frac{S_n}{n} \geq \mu+\epsilon\right) \leq \inf_{0<s<b} \left(M_Y(s)\right)^n$$
$$= \inf_{0<s<b} e^{-sn(\mu+\epsilon)} \left(M_{X_i}(s)\right)^n \text{ olur.}$$

<u>Alıştırma 104.</u> X_i'ler $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = 1/2$ olan IID rasgele değişken dizisi olsun. $\mathbb{P}(S_n/n \geq 0.1)$ için üssel azalan bir üst sınır bulunuz.