a. What fraction of bottles end up in the glass furnace?

$P(\text{plastic}) = \%.20$ $P(R \mid \text{plastic}) = 0.98$

$P(\text{cam}) = \%.80$

$P(R \mid \text{cam}) = 0.03 \rightarrow P(R' \mid \text{cam}) = 0.97$

$P(A \mid X) + P(A' \mid X) = 1$

remove

$P(R) = P(R \mid \text{plastic}) \cdot P(\text{plastic}) + P(R \mid \text{cam}) \cdot P(\text{cam})$

$= 0.98 \cdot 0.2 + 0.03 \cdot 0.8 = 0.22$

$P(R') = 1 - P(R) = 1 - 0.22 = 0.78$

b. Given that a bottle ends up in the glass furnace is the prob. that it is a glass bottle?

$$P(\text{cam} \mid R') = \frac{P(R' \mid \text{cam}) \cdot P(\text{cam})}{P(R')} = \frac{0.97 \cdot 0.8}{0.78} \approx 0.995$$

%0.5 plastik şişe birliği

EXAMPLE: Carlton's disease affects %0.1 of the population. A medical test is developed to detect the disease by a simple breath test at home. If the test indicates that a person has the disease, then this test result is said to be "positive". If a person has the disease, then there is %99 chance that the medical test will be positive. If a person doesn't have the disease, then there is just %2 chance that the test will be positive (false positive).

a. What is the prob. that a randomly-selected person will be tested positive?

D: Person has disease

D': Person doesn't have disease

+: Medical test positive

−: " " not positive

$P(+ \mid D) = 0.99$

$P(+ \mid D') = 0.02$

$P(D) = 0.001$

$P(D') = 0.999$

$P(+) = P(+|D) \cdot P(D) + P(+|D') \cdot P(D')$

$= 0.99 \cdot 0.001 + 0.02 \cdot 0.999 = 0.02097 \approx \%2.1$

b. Out of a population of 60 million people, how many will be tested positive?

Burada önceki soruya ek olarak popülasyon sayısı vardır. Yani $P(D) = 0.001$ değil de $0.001 \times 60 \times 10^6$ olur, aynı şekilde $P(D')$ de $0.999 \times 60 \times 10^6$ olur. Yani aslında sonucu 60×10^6 ile çarpmak yeterli olacaktır.

$P(+) = 0.02097 \times 60 \times 10^6 = 1258200$ kişi positive test edilir!

c. How many people in the population actually have the disease?

Popülasyonun $\%0.1$'i hastadır.

$P(D) = 60 \times 10^6 \times 0.001 = 60000$ kişi gerçekten hastadır.

d. Given a positive, what is the prob. that the person has Carlton's disease?

$$P(D|+) = \frac{P(+|D) \cdot P(D)}{P(+)} = \frac{0.99 \cdot 0.001}{0.02097} \approx 0.047 \ (\%4.7)$$

Lecture 5 - The Random Variable and Expectation Values

The Random Variable; Bir random variable X, S örnek uzayın-daki her öğeye rastgele bir sürecin sonucunu temsil eden bir sayı atayan bir fonksiyondur

- Yani random variable X, bir fonksiyondur
  Burada zaten var olan bir sample space (set) vardır.

$$\underset{\text{domain}}{X : S} \rightarrow \underset{\text{range}}{\mathbb{R}}$$

- S → prob. space
  Bu yüzden X için de olasılık yapısını kullanırız

Deney: Toss a coin 2 times

Sample Set: $S = \{HH, HT, TH, TT\}$

$P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}$

Hepsi aynı (equally likely)

A random variable:

X: the number of heads → head hiç gelmeyebilir
1 tane gelebilir
2 " "

$P(X=0) = P(\{TT\}) = \frac{1}{4}$

$P(X=1) = P(\{HT, TH\}) = \frac{1}{2}$

$P(X=2) = P(\{HH\}) = \frac{1}{4}$

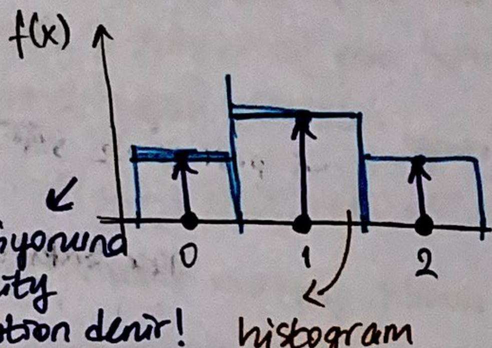
$f(x) = P(X=x) \Rightarrow f(0) = P(X=0) = \frac{1}{4}$

$f(1) = P(X=1) = \frac{1}{2}$

$f(2) = P(X=2) = \frac{1}{4}$

Probability Mass Function

↳ Fonksiyon grafiği çizilmesinde çok işe yarar.



f fonksiyonuna probability distribution denir!

Bu grafiği sadece noktalarla belirtmektense oklar kullanmak daha iyidir. Hatta histogram kullanmak daha da iyidir.

! HISTOGRAM? : Gruplandırılmış bir veri dağılımının sütun grafiğiyle gösterimidir.

Sample set: The people in a country
(The students in a class)

Experiment: Select a person from the population at random

Burada kişilerin boylarına bakalım:

A random variable: X seçilen kişinin boyu olsun.

Fonksiyonumuzu oluşturalım:

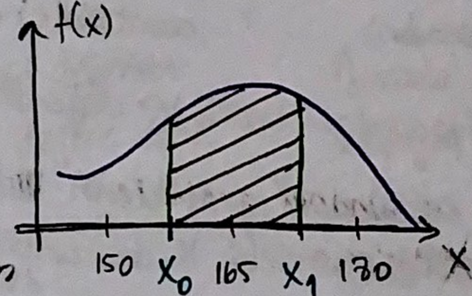
X = 165 olsun. Peki olasılık fonksiyonumuzu nasıl bulacağız? → $P(X=165) = ?$

Burada bir sorun oluşur. Tamamen 165 olan kişileri bulamayız. Mutlaka küsüratlı olur. Bu yüzden $P(X=165)=0$ olur. Burada sonsuz durum vardır. Mesela Ali'yi seçmiş olalım. Boyu 165.237856 olsun. Normalde onun da 165 olarak alınması gerekirken bu fonksiyona göre alınamaz. Fakat şöyle düşünülebilir:

$$P(164.5 < X < 165.5)$$

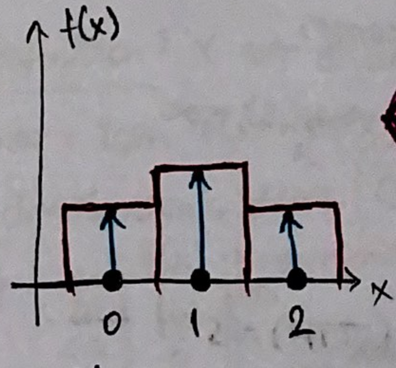
Bu şekilde olan prob. fonksiyonumuzun grafiği de şöyledir:

- f, süreklidir.
- Buradaki f de probability distr.'a sahiptir.



Bu fonksiyona da <u>probability density function</u> denir.

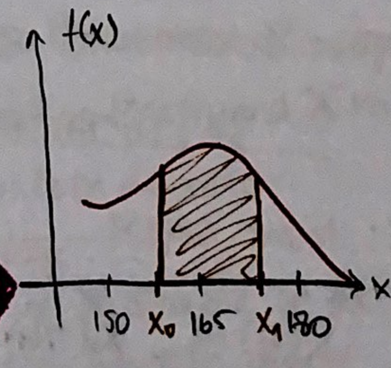
$$P(X_0 < X < X_1) = \int_{X_0}^{X_1} f(x) \cdot dx$$



Histogram

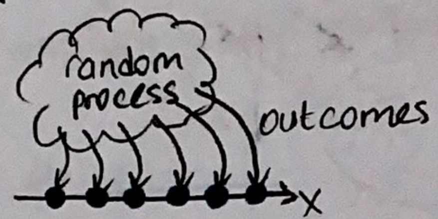
~~Prob. Mass Func.~~

Prob. Mass Func.

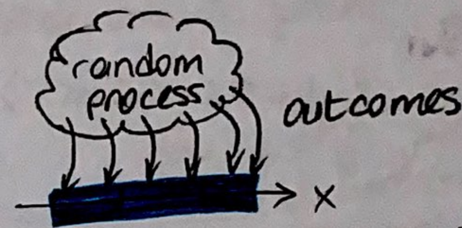


Prob. Density Func.

✓ X'in değer aralıkları ve dasılıkları ile ilgileniyoruz. Burada, sonuçları discrete space'te olan rastgele bir süreç ile continuous space'te olan rastgele bir süreç arasındaki ayrımı yapmak önemlidir.



Discrete space; uzaydaki outcomeların countable (sayılabilir) olmasıdır.



Continuous space; uzaydaki outcomeların sayılamaz (uncountable) olmasıdır.

NOT: x, X koleksiyonundaki belirli bir değerdir.

✓ Önceki sayfada en sondaki grafiklerin ilkindeki X discrete random variable, ikincisindeki X de continuous random variable'dır.

✓ Range aralık ise X continuous random variable, belirli bir sayı ise discrete random variable'dır.

Deney: Toss a coin until a Head comes

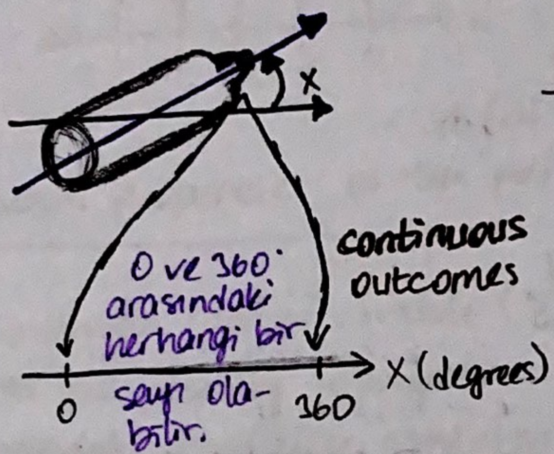
$$S = \{H, TH, TTH, \dots\} \rightarrow n(S) = \infty$$

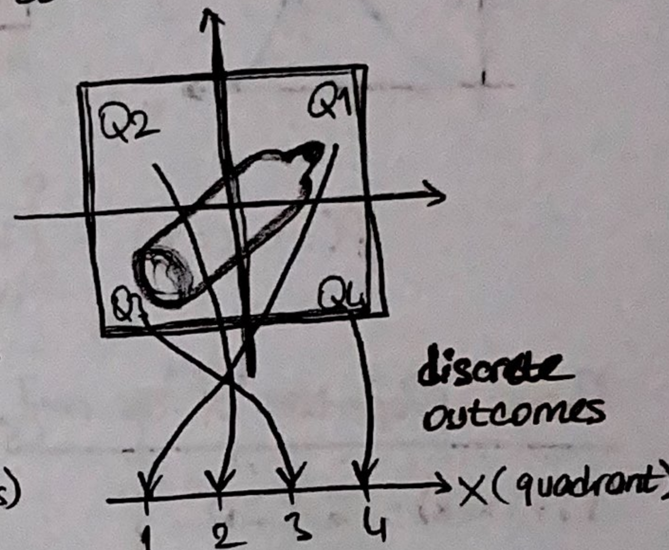
X: deneme sayısı

$$X(H) = 1 \,,\; X(TH) = 2 \,,\; X(TTH) = 3$$

X burada 1, 2, 3, ... olabilir. Bunlar sayılabildiği için de discrete olur.

EXAMPLE: Dönen bir şişenin düz bir yüzey üzerinde durduğu pozisyon rastgele olarak kabul edilsin





Burada x bir açı değeridir ve herhangi bir değere denk gelebilir.

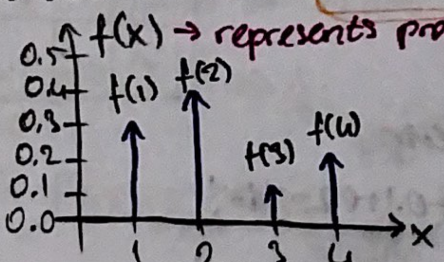
↳ Continuous

Burada her bir çeyrek bir sayıya karşılık gelir. Bu yüzden tek bir değeri olabilir.

↳ Discrete

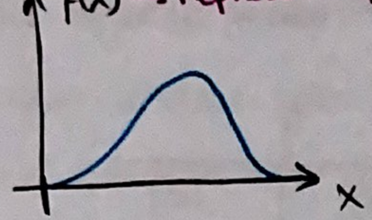
The Probability Distribution

Definition: X bir discrete random variable ise, X range'indeki her x için $f(x) = P(X=x)$ ile verilen fonksiyona X'in prob. distribution'ı denir. Continuous için de vardır



Eğer X discrete ise f(x) probab. mass function olarak adlandırılır.

pmf $P(X=x) = f(x)$

* 

$f(x) \rightarrow$ represents prob. density

Eğer X continuous ise f(x) prob. density function olarak adlandırılır $\rightarrow$ pdf

$$P(a < X < b) = \int_a^b f(x)\,dx$$

Basic Properties of the pmf

1. $P(X=x) = f(x) \geq 0$

   f(x), her x outcome'ı için olasılığı tanımlar, bu nedenle pozitif olmalıdır.

2. $\sum_x f(x) = 1 \rightarrow P(S)$ $\rightarrow$ Random variable (X) tüm olası sonuçları kapsar ve bu nedenle toplam olasılık 1'dir.

3. $P(x_j \leq X \leq x_k) = \sum_{i=j}^{k} f(x_i)$ $\rightarrow$ Tek tek olasılıkları basitçe toplayarak bir dizi sonucun olasılığını çıkarabiliriz (her sonuç disjoint'tir).

✓ Pmf'yi ifade etmenin 3 yolu vardır:

1- $X = \{1, 2, 3, 4\}$

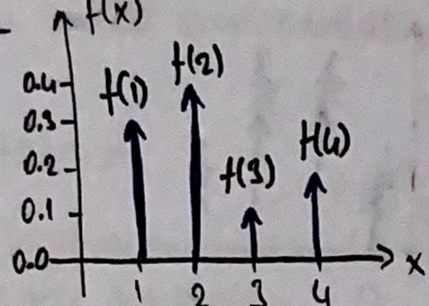
$f(x) = \{0.3, 0.4, 0.1, 0.2\}$

2 kurala göre bu değerlerin toplamı 1 olmalı! $0.3+0.4+0.1+0.2=1$'dir!

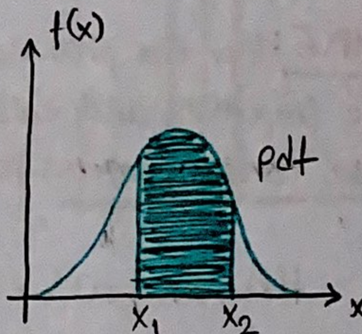
2-

x	1	2	3	4
f(x)	0.3	0.4	0.1	0.2

3-



Basic properties of the pdf



Continuous random variable (X) için prob. density fonksiyonu aşağıdaki özelliklere sahiptir:

1. $f(x) \geq 0$, tüm x'ler için
   Olasılık negatif olamaz.
2. $\int_x f(x)dx = 1$
   Random variable (X) tüm mümkün çıktılara sahiptir ve bu yüzden toplam olasılık 1'dir.
3. $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)\,dx$

↳ Basitçe prob. density üzerinden integral alarak bir dizi outcome'ın olasılığını çıkartabiliriz.

EXAMPLE: For the experiment of tossing a coin 3 times, find the prob. distribution for the random variable X representing the number of heads in the outcome.

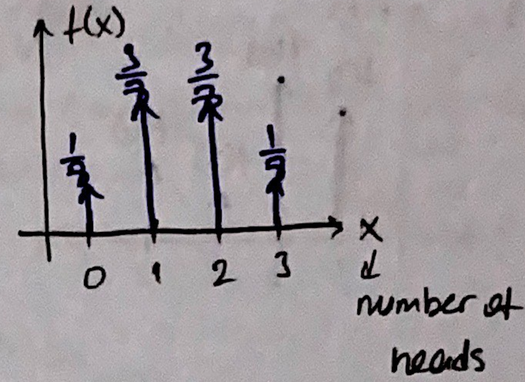
X : number of heads → {0,1,2,3}

$f(0) \leftarrow P(X=0) \Rightarrow \frac{1}{8}$, $f(1) \leftarrow P(X=1) \Rightarrow \frac{3}{8}$

$f(2) \leftarrow P(X=2) \Rightarrow \frac{3}{8}$, $f(3) \leftarrow P(X=3) \Rightarrow \frac{1}{8}$

> 2. koşul da sağlanır. Toplam = 1'dir.

x	0	1	2	3
f(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

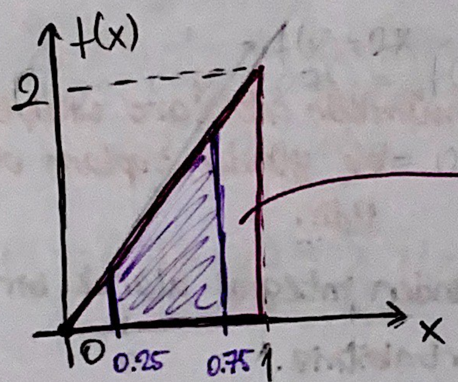


EXAMPLE: For the prob. density function below, investigate its basic properties, and calculate the prob. that the outcome of the random process is between 0.25 and 0.75.

$f(x) = 2x \; ; \; 0 \leq x < 1$

$f(x) \geq 0 \; \boxed{\checkmark}$

$\int_x f(x) \cdot dx = 1$ olmak zorundadır



$\frac{2 \times 1}{2} = 1 \; \boxed{\checkmark}$

$P(0.25 < X < 0.75) = \int_{0.25}^{0.75} 2x \cdot dx$

$\Rightarrow x^2 \Big|_{0.25}^{0.75} = 0.5$

Expectation Values

Expectation değerleri, bir dağılımı (distribution) birkaç tanımlayıcı değer (tanımlayıcı istatistikler) biçiminde özetleyen matematiksel nesnelerdir.

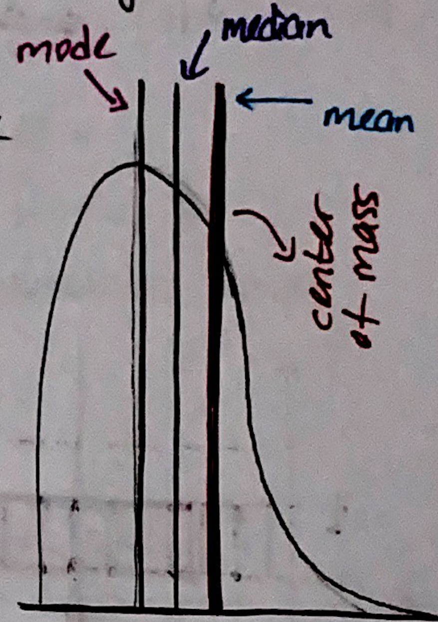
Bu kursta bunlardan 2'sine odaklanacağız:

1. Mean, μ_X

Dağılımın ortasıdır / ortalama. → grafik simetrik ise ortadadır!
Kütle merkezidir

2. Standard Deviation, σ_X

Dağılımın ortalamaya ne kadar yayıldığının bir ölçüsüdür



! Dağılımı popülasyon gibi düşün.

Mode → En çok tekrar eden öğe

Medyan → Ortanca değer.

! $f(x) = P(X=x)$ olsun ve X discrete olsun.

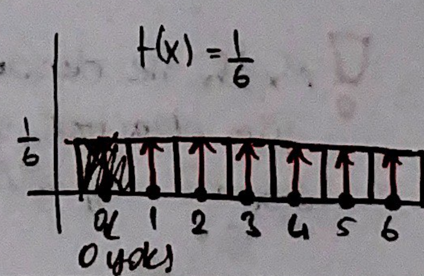
Buradan: Mean $(\mu_X) = \sum_x x \cdot f(x)$ olur.

$$\text{Standard Dev. } (\sigma_X) = \sqrt{\sum_x (x-\mu_X)^2 \cdot f(x)}$$

Deney: Throw a die

Sample set: $S = \{1,2,3,4,5,6\}$

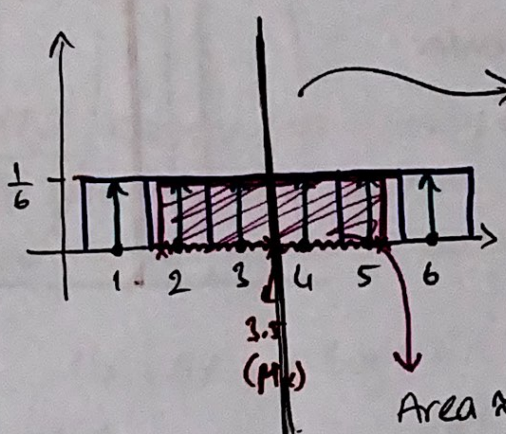
A random variable: $X(w) = w$



$$\mu_X = \sum_x x \cdot f(x) \longrightarrow 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}$$

$$= \frac{21}{6} \doteq 3.5$$

Direkt ortalamayı bulabiliriz!



Hem x eksenine diktir, hem de lacivert alanın merkezidir. Mean value budur!

Area $\approx \frac{2}{3}$ — Tüm alanın $\frac{2}{3}$'üdür.

Area(Lacivert) = 1

$$\sigma_X = \sqrt{\underbrace{\sum_x (x-\mu_X)^2 \cdot f(x)}_{A}} \Rightarrow A = (1-3.5)^2\cdot\frac{1}{6} + (2-3.5)^2\cdot\frac{1}{6} + \dots + (6-3.5)^2\cdot\frac{1}{6}$$

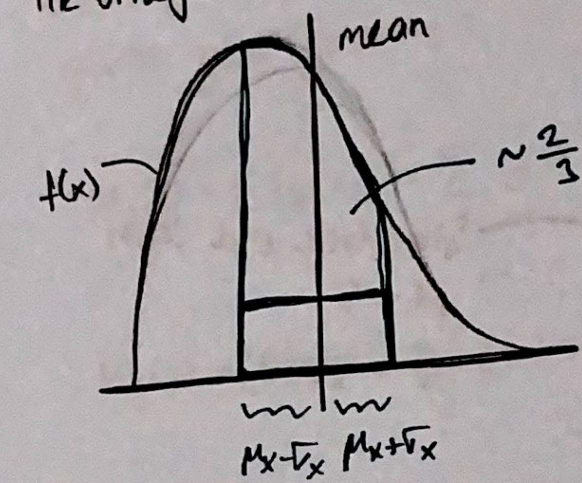
$$= 2.916667$$

$$\sigma_X = \sqrt{A} \approx 1.7078$$

$$\mu_X \pm \sigma_X = 5.2078$$
$$1.7922$$

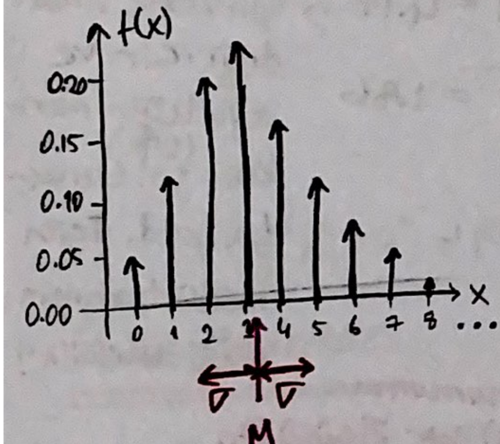
! $\mu_X \pm \sigma_X$ ile oluşan şekil her zaman tüm şeklin $\frac{2}{3}$'üdür. Tüm alan her zaman 1 olacağı için de direkt $\frac{2}{3}$'tür diyebiliriz.

İlk örneğe tekrar bakarsak :



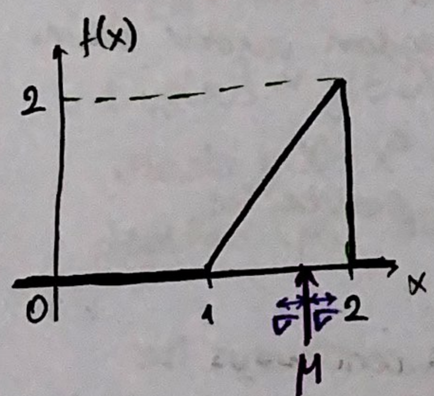
$$\int_{\mu_x-\sigma_x}^{\mu_x+\sigma_x} f(x)\,dx \approx \frac{2}{3}$$

Examples of the mean μ and standard deviation σ



$\mu = 3.17$ → Eğer simetrikse tam ortasıdır değilse formülden hesaplama-usın!

$\sigma = 1.78$



$\mu = 1.67$

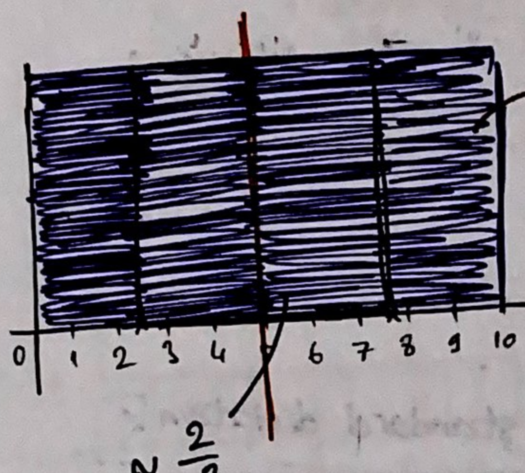
$\sigma = 0.24$

! μ = center of mass

$\mu \pm \sigma$ covers $\approx \frac{2}{3}$ of the distribution

Estimation

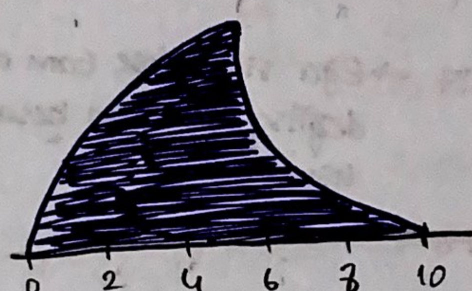
Examples (Estimate μ and σ for the following pdfs)



Simetrik old için

$\mu = 5$

$\sigma = 2.89$



$\mu = 4.17 \rightarrow$ simetrik olmadığı için ve

$\sigma = 1.86$ ağırlığın merkezi (5'in) sol tarafta old için belki tahmin yürütülebilir!

Definition : mean, μ ; distribution'ın konumunun (center of mass) bir ölçüsüdür

X, $f(x)$ prob. distr. fonksiyonu ile bir random variable olsun.

X'in mean'i :

$$\mu_X = E[X] = \sum_x x \cdot f(x) \rightarrow \text{Eğer } X \text{ discrete ise}$$

$$\mu_X = E[X] = \int_x x \cdot f(x) \cdot dx \rightarrow \text{Eğer } X \text{ continuous ise}$$

g(x) fonksiyonunun expectation'ına bakacak olursak x yerine g(x) yazarız:

$$\mu_g(x) = E[g(X)] = \sum_x g(x) \cdot f(x) \rightarrow X \text{ discrete ise}$$

$$\mu_g(x) = E[g(X)] = \int_x g(x) \cdot f(x) \cdot dx \rightarrow X \text{ continuous ise}$$

Deney: Toss a coin 2 times

X: number of head

$g(x) = x^3$ olsun. g(x)

x^3	0	1	8
x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$P(g(X) = 8) = P(X^3 = 8) = P(X = 2) = f(2)$$

Some Properties

✓ $E[aX^n + b] = a \cdot E[X^n] + b$; a, b, n sabitlerdir.

✓ $E[g(x)] \sim g(E[X])$

Mesela $g(X) = X^2$ olsun. Burada $E[g(x)]$ in yaklaşık değerinin $(\mu_x)^2$ old söyleyebiliriz. Çünkü $E[X] = \mu_x$'tir.

EXAMPLE: What is the mean number of heads obtained when three coins are tossed?

$$\mu_X = \sum_x x.f(x) = E[X]$$

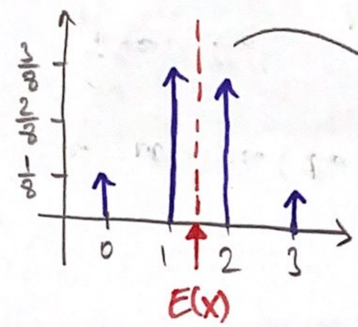
$$\boxed{f(x) = P(X=x)}$$

X: number of heads = $\{0,1,2,3\}$

X	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\mu_X = 0.\frac{1}{8} + 1.\frac{3}{8} + 2.\frac{3}{8} + 3.\frac{1}{8}$$

$$= \frac{12}{8} = \frac{3}{2} = 1.5$$



Grafik simetrik old. için direkt $\overset{\mu}{\boxed{1.5}}$ diyebiliriz. Bu aynı zamanda ağırlığın merkezidir

EXAMPLE: A factory produces circular metal discs with various diameters X (cm) with the probs. P given below:

x(cm)	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

a. Calculate the average diameter of the disc.

çap

$\mu = ?$

$$\mu_X = E[X] = \sum_x x.f(x) \Rightarrow 4.\frac{1}{12} + 5.\frac{1}{12} + 6.\underset{(3)}{\frac{1}{4}} + 7.\underset{(3)}{\frac{1}{4}} + 8.\underset{(2)}{\frac{1}{6}} + 9.\underset{(2)}{\frac{1}{6}}$$

$$\Rightarrow 82/12 \approx 6.83$$

b. Calculate the average area of the discs.

Burada farklı bir fonksiyonun yaklaşık değerini bulmamız isteniyor.

$$g(X) = A = \pi r^2 = \pi\left(\frac{X}{2}\right)^2 = \frac{\pi}{4}\cdot X^2$$

I. yol)

X	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$
$g(X)$	$\frac{\pi}{4}4^2$	$\frac{\pi}{4}5^2$	$\frac{\pi}{4}6^2$	$\frac{\pi}{4}7^2$	$\frac{\pi}{4}8^2$	$\frac{\pi}{4}9^2$

$$g(x) = E[A] = \sum_x g(x)\cdot f(x) \Rightarrow \frac{\pi}{4}\left(4^2\cdot\frac{1}{12} + 5^2\cdot\frac{1}{12} + 6^2\cdot\frac{1}{4} + 7^2\cdot\frac{1}{4} + 8^2\cdot\frac{1}{6} + 9^2\cdot\frac{1}{6}\right)$$

$$\Rightarrow 38.3536 \approx 38.4\,cm^2$$

II. yol) → Yaklaşık Değer / Tahmin

$$E[A] = E\left[\frac{\pi}{4}\cdot X^2\right] = \frac{\pi}{4}\cdot E[X^2] \sim \frac{\pi}{4}\cdot \underbrace{E[X]^2}_{\mu^2}$$

$$= \frac{\pi}{4}\left(\frac{41}{6}\right)^2 = 36.7\,cm^2$$

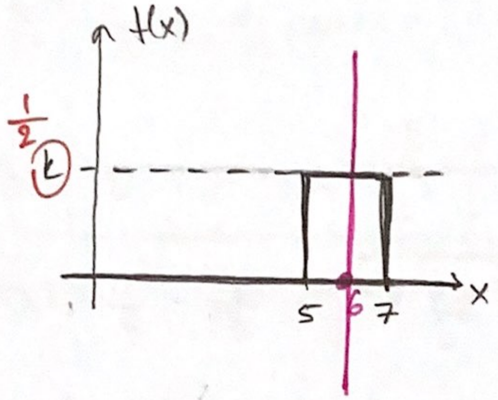
Benzer

! $E[X^2] \neq E[X]^2$

Soruda ilk yolda kullandığımız bize gerçek değeri verirken 2. yolda kullandığımız yaklaşık değeri vermiştir. Buna dikkat etmelisin!

EXAMPLE: The diameter, x, in cm of a sample of steel spheres can be considered to be a uniform continuous random variable of the form: $f(x)=k$ for $5<X<7$ cm.

a. Sketch the pdf and determine the value of k for this pdf.



$k=?$

Tüm alanın 1 olması gerektiğini biliyoruz.

$$\int_5^7 \underset{f(x)}{k}\cdot dx = 1 \rightarrow kx\Big|_5^7 = 1$$

$$\rightarrow 7k-5k=1$$

$$2k=1$$

$$\boxed{k=\frac{1}{2}}$$

b. Determine the mean diameter of the steel spheres.

$$M_x = \int_5^7 x\cdot f(x)\cdot dx = \frac{1}{2}\cdot\int_5^7 x\cdot dx \rightarrow \frac{1}{2}\left(\frac{x^2}{2}\Big|_5^7\right) = \frac{1}{2}\cdot\left(\frac{49-25}{2}\right)$$

$$\Rightarrow 6$$

Zaten şekil simetrik old. için direkt 6 denebilir.

c. Calculate the mean volume of the steel spheres.

$$V=\frac{4}{3}\pi r^3 \rightarrow E[V]=?\ ,\ r=\frac{x}{2}$$

$$E[V]=\frac{4}{3}\cdot\pi\cdot E\left[\left(\frac{x}{2}\right)^3\right] = \frac{\cancel{4}}{3}\cdot\pi\cdot\frac{1}{\cancel{8}_2}\cdot E[x^3] = \frac{\pi}{6}\cdot E[\underbrace{x^3}_{g(x)}]$$

$$=\frac{\pi}{6}\cdot\int_5^7 x^3\cdot \overset{\frac{1}{2}}{f(x)}\cdot dx = \frac{\pi}{12}\cdot\int_5^7 x^3\cdot dx = \frac{\pi}{12}\cdot\left(\frac{x^4}{4}\Big|_5^7\right)$$

$$= \frac{\pi}{48} \cdot x^4 \Big|_5^7 = \frac{\pi}{48} \cdot (7^4 - 5^4) = 116.298928 \text{ cm}^3 \approx 116 \text{ cm}^3$$

Yaklaşıma da bakalım: $E[X^3] \sim \underbrace{E[X]}_{\mu_X = 6}{}^3$

Formülde yerine yazarsak

$$\frac{\pi}{6} \cdot E[X]^3 = \frac{\pi}{6} \cdot 6^3 = 113 \text{ cm}^3$$

Yaklaşık!

Variance & Standard Deviation

X random variable'ının variance'ı (σ^2) şöyle tanımlanır:

$$\sigma^2 = E[(X - \underset{\text{mean}}{\mu})^2] = \sum_x (x-\mu)^2 \cdot f(x) \rightarrow X \text{ discrete ise}$$

$$\sigma^2 = E[(X-\mu)^2] = \int_x (x-\mu)^2 \cdot f(x) \cdot dx \rightarrow X \text{ continuous ise}$$

Bu, distribution'ın mean değere göre ne kadar değiştiğinin bir ölçüsüdür.

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[X^2 - 2X\mu_X + \mu_X^2]$$
$$= E[X^2] - 2\mu_X \underbrace{E[X]}_{\mu_X} + \mu_X^2$$
$$= E[X^2] - \mu_X^2$$

Bunu ← kullanacağız!

$$\boxed{\sigma_X^2 = E[X^2] - E[X]^2}$$ → Genelde bunu hesaplamak daha kolaydır

Sonra $\sigma_X = \sqrt{\sigma_X^2}$ 'den standard deviation bulunur.

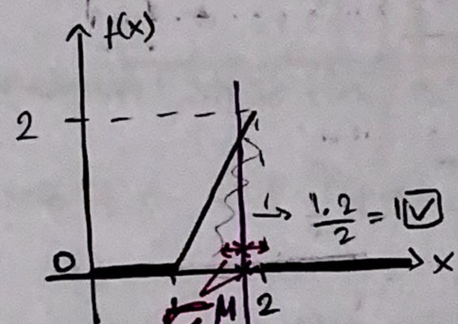
Property: $\sigma_{(aX+b)} = a \cdot \sigma_X$

$$\boxed{y = 100X \text{ ise } \sigma_y = 100 \cdot \sigma_X}$$

EXAMPLE: The power generated (GWatts) by a tidal power station is a continuous random variable X having the prob. density shown right.

$$f(x) = \begin{cases} 2(x-1), & 1<x<2 \\ 0, & \text{elsewhere} \end{cases}$$

Find the mean and standard deviation of X.



- $\mu_X = E[X] = \int_x x \cdot f(x) \cdot dx$

$$\to \int_1^2 \underbrace{x \cdot 2(x-1)}_{2x^2-2x} dx$$

$$\to \left(\frac{2}{3}x^3 - \frac{\cancel{2}x^2}{\cancel{2}}\right)\Big|_1^2 = \left(\frac{16}{3} - 4\right) - \left(\frac{2}{3} - 1\right)$$

$$= \frac{14}{3} - 3 = \frac{5}{3}$$

$\boxed{\approx 1.67}$

- $\sigma_X^2 = E[X^2] - E[X]^2$

$$E[X^2] = \int_x x^2 \cdot f(x) \cdot dx \to \int_x x^2 \cdot (2x-2) dx = \int_x (2x^3 - 2x^2) dx$$

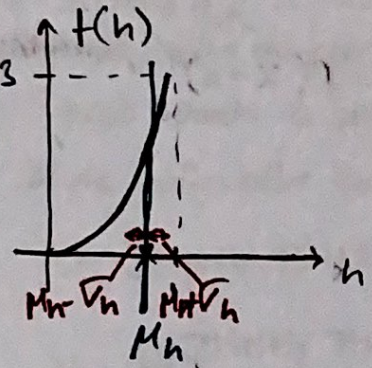
$$\to \int_1^2 (2x^3 - 2x^2) dx = \left(\frac{x^4}{2} - \frac{2}{3}x^3\right)\Big|_1^2 = \frac{17}{6}$$

$$\sigma_X^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \underset{(3)}{\frac{17}{6}} - \underset{(2)}{\frac{25}{9}} = \frac{1}{18}$$

$\sigma_X = \sqrt{\frac{1}{18}} = 0.2357 \to$ Yani dağılım 1.67 ± 0.24 GW'dır.

EXAMPLE: In a computer game, a virtual first aid kit provides a random healing value, H, to a player represented by the following prob. density function: $f(h) = 3h^2;\ 0 \le H < 1$

a. Show that this is a valid pdf.



1. $f(x) \ge 0$ $\boxed{✓}$

2. $\int_x f(x)dx = 1$

$\int_0^1 3h^2 dh = h^3 \Big|_0^1 = 1$ $\boxed{✓}$

b. Calculate the average healing value.

$$\mu_h = \int_0^1 h \cdot f(h) \cdot dh \rightarrow \int_0^1 h \cdot 3h^2 \cdot dh = \frac{3}{4} h^4 \Big|_0^1 = \frac{3}{4}$$

c. Calculate the standard deviation of the healing value.

$$\sigma_h^2 = E[H^2] - \underbrace{E[H]}_{\mu_h}{}^2$$

$$E[H^2] = \int_0^1 h^2 \cdot 3h^2 \cdot dh = \frac{3}{5} h^5 \Big|_0^1 = \frac{3}{5}$$

$$\sigma_h^2 = \frac{3}{5} - \frac{9}{16} = \frac{3}{80} \rightarrow \sigma_h = \sqrt{\frac{3}{80}} \approx 0.194$$

d. What is the average and standard deviation of 100H?

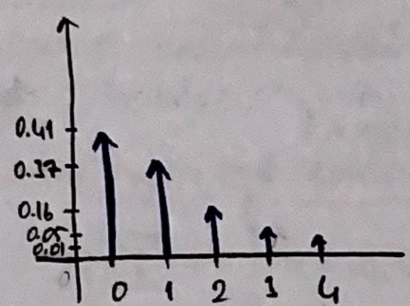
! $E[aX^n + b] = a \cdot E[X^n] + b \rightarrow E[100H] = 100 \cdot \underbrace{E[H]}_{\frac{3}{4}} = 75$

$\sigma_{aX+b} = a \cdot \sigma_X \rightarrow \sigma_{100H} = 100 \cdot \underbrace{\sigma_H}_{0.194} = 19.4$

EXAMPLE:

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

a. Plot the prob. function.



$$f(x) = P(X = x)$$

discrete random variable

b. Find the expected number of imperfections, $E(X) = \mu$.

$$E[X] = \mu = \sum_x x \cdot f(x) \Rightarrow \cancel{0 \cdot 0.41} + 1 \cdot 0.37 + 2 \cdot 0.16 + 3 \cdot 0.05 + 4 \cdot 0.01$$

$$= 0.88$$

c. Find $E[X^2]$.

$$E[X^2] = \sum_x x^2 \cdot f(x) \Rightarrow 0 \cdot 0.41 + 1 \cdot 0.37 + 4 \cdot 0.16 + 9 \cdot 0.05 + 16 \cdot 0.01$$

$$= 1.62$$

d. Find σ.

$$\sigma^2 = E[X^2] - E[X]^2 = 1.62 - 0.88^2 = 0.8456$$

$$\sigma = \sqrt{0.8456} \approx 0.92$$

EXAMPLE: A coin is flipped until 3 heads in succession occur. List only those elements of the sample space that require 6 or less tosses. Is this a discrete sample space? Explain.

arka arkaya 3 head

Sample space 6 atış koşulu olmasaydı sonsuz öğeden oluşurdu. $n(S) = \infty$ olurdu.

Fakat sonsda 6 atış veya daha az için bakılması sebebiyle mümkün çıktılar:

C = {HHH, THHH, TTHHH, TTTHHH, HTHHH, HTTHHH, HHTHHH, THTHHH}

! Gibi sayılabilir old. için discrete'dir.

$\boxed{X: \text{number of tosses}}$

(X'in range'i) Gibi sonluysa mutlaka discrete olur!

Lecture 6 – Special Discrete Prob. Distributions (I)

The Geometric Distribution

✓ Geometrik dağılım, ilk başarıyı elde etmek için gereken deneme sayısının (x) olasılık dağılımını (prob. distribution) tanımlar.

✓ Her denemenin değişmeyen bir başarı olasılığı p vardır ve bu nedenle bir başarısızlık olasılığı $q = 1-p$'dir.

success prob.

failure prob.

✓ İlk başarının ortaya çıkması, deneme yapma sürecini sona erdirir.

! Buradaki varsayım, başarı olasılığının değişmediğidir (her denemeden sonra giderde daha iyi olmayız).

Deney: Toss a coin until you get Head!

Sample space: $S = \{H, TH, TTH, TTTH, \dots\}$, $n(S) = \infty$

success = getting a head

success olasılığı p ile gösterilir.

Böylece failure prob da $q = 1-p$ olur.

X: The number of tosses

$f(1) = P(X=1) = P(\{H\}) = \frac{1}{2}$

$f(2) = P(X=2) = P(\{TH\}) = \frac{1}{4} = \frac{1}{2^2}$

$f(3) = P(X=3) = P(\{TTH\}) = \frac{1}{8} = \frac{1}{2^3}$

$\vdots$

$f(n) = P(X=n) = P(\{\underbrace{TTT\dots}_{(n-1)\text{ tane } T}H\}) = \frac{1}{2^n}$

$g(n) = p \cdot q^{n-1}$ olur

prob. mass func in the geometric distribution

Derivation of the Geometric Prob. Distribution

Her denemenin bir başarı olasılığı p ve bir başarısızlık olasılığı $q = 1-p$ vardır.

İlk başarının ortaya çıkması deneme sürecini sona erdirir. Çünkü istenene ulaşılmıştır.