Cumulative Distribution function

PMF $P_X(x_k) = 1/4$

PMF discrete random variable

CDF: can be used for any kind of random variable

Definition:

A CDF of random variable X is defined as

$$F_X(x) = P(X \leq x) \text{ for all } x \in R$$

Example: I have tossed a coin twice. RV X is the number of heads. find its PMF and CDF

HH — 2
HT — 1
TH — 1
TT — 0

$R_X = \{0, 1, 2\}$

$$\underbrace{P_X(0) = 1/4 \quad P_X(1) = 2/4 \quad P_X(2) = 1/4}_{\text{PMF}}$$

CDF: $F_X(x) = P(X \leq x)$

$\text{CDF}: F_X(x) = P(X \le x)$ (with α written under x)

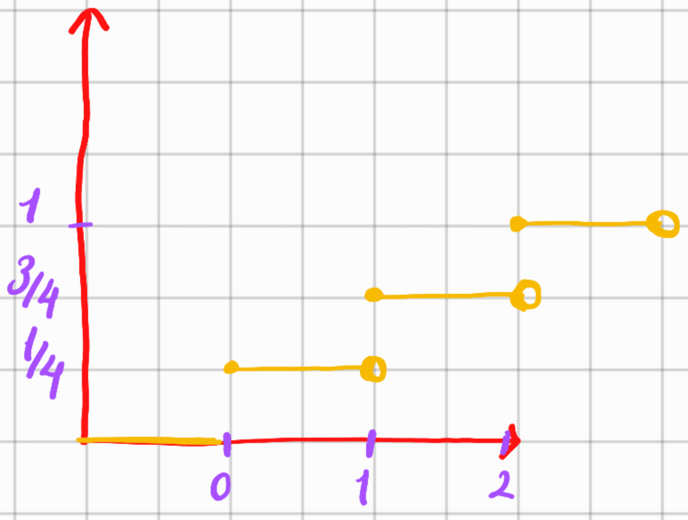
for $\alpha < 0 \quad F_X(\alpha) = 0$

$\alpha \geq 2 \quad F_X(\alpha) = 1 \longrightarrow P_X(0) + P_X(1) + P_X(2) = \underbrace{\frac{1}{4} + \frac{2}{4} + \frac{1}{4}}_{1}$

For $0 \leq \alpha < 1 \quad F_X(\alpha) = P_X(0) = \frac{1}{4}$

For $1 \leq \alpha < 2 \quad F_X(\alpha) = P(X \leq \alpha) = P_X(0) + P_X(1) = \underbrace{\frac{1}{4} + \frac{2}{4}}_{3/4}$

$$F_X(\alpha) = \begin{cases} 0 & \alpha < 0 \\ 1/4 & 0 \leq \alpha < 1 \\ 3/4 & 1 \leq \alpha < 2 \\ 1 & x \geq 2 \end{cases}$$



Properties of Cumulative Dist. Function

1. $F_X(-\infty) = 0 \qquad F(\infty) = 1$

CDF'nin değeri her zaman artar

2. CDF is a non-decreasing function

$$\alpha \leq \beta \Longrightarrow F_X(\alpha) \leq F_X(\beta)$$

3. For any $x_k \in R_X$

$$F_X(x_k) - F_X(x_k - \epsilon) = P(x = x_k)$$

Very small value

4. $P(a < x \leq b) = P(x \leq b) - P(X \leq a)$

$$= F_X(b) - F_X(a)$$

Expectation

$a_1, a_2, a_3 \ldots a_n$

average $\underbrace{\sum_{i=0}^{n} a_i}_{\text{mean}}$

average

Consider a R.V X, how we find its average?

Let X be a R.V with $R_X = \{x_1, x_2, x_3 \dots\}$

the expected value of X, EX

$$EX = \sum_{x_k \in R_k} x_k \underbrace{P(X = x_k)}_{P_X(x_k)}$$

$$\underbrace{E[X] = E(X) = \mu_X}_{\text{expected value}}$$

yukarıdaki ifade bazı kitaplarda böyle de gösteriliyor

example: X'in beklenen değerini bulalım

X	P(x)
2	0,75
3	0,25

$$EX = \sum x_k P_X(x_k)$$

$$= 2 \cdot 0.75 + 3 \cdot 0.25$$

$$= 1,5 + 0.75$$

$$\boxed{= 2,25} \; \mu_X$$

Functions of random variables

X is a random variable

$Y = g(x)$

Y → random variable

g → function

x → R.V

yani bir rastgele degiskenin fonksiyonları bunlar girdi olarak rastgele bir degisken alır ve bunlar yine rastgele bir degisken olan Y üretir

$R_Y = \{ g(x) \mid x \in R_X \}$

R_Y → range

We already know the PMF(x),

biz PMF(x)'i biliyoruz o zaman PMF(Y) nasıl elde ederiz

$$P_Y(y) = P(Y=y) = P(g(x)=y) = \sum_{x : g(x)=y} P_X(x)$$

Ex: Let X be a discrete RV with $P_X(k) = 1/5$

for $k = -1, 0, 1, 2, 3$

Let $Y = 2|X|$ find the range and PMF of Y

$R_X = \{-1, 0, 1, 2, 3\}$

$R_Y = \{2, 0, 2, 4, 6\}$

x'in range'indaki her elemanı $2|X|$'de yerine koyduk

$R_Y = \{0, 2, 4, 6\}$

$P_X(0) = P_X(-1) = P_X(1) = P_X(2) = P_X(3) = 1/5$

$P_Y(0) = P(Y=0) = (2|X|=0) = P_X(X=0) = 1/5$

$P_Y(2) = P_Y(2|\underset{+1 \;\; -1}{X}| = 2) = P_X(1) + P_X(-1) = \underbrace{1/5 + 1/5}_{2/5}$

$P_Y(4) = P_Y(2|\underset{2}{X}| = 4) = P_X(2) = 1/5$

$P_Y(6) = 1/5$

Law of unconscious statistician (Lotus)
for discrete random variable

$$\underbrace{E[g(x)]}_{Y} = \sum_{x_i \in R_X} g(x_i)\, P_X(x_i)$$

(for proof see your book)

Ex: Prev question find $E[Y]$ where $Y = 2|x|$

$R_X = \{-1, 0, 1, 2, 3\}$

Lotus

$$E[Y] = 2|-1| \cdot P_X(-1) + 2|0| \cdot P_X(0) + 2|1| \cdot P_X(1) + 2|2| \cdot P_X(2) + 2|3| \cdot P_X(3)$$

$$\rightarrow 2 \cdot 1/5 + 0 + 2 \cdot 1/5 + 4 \cdot 1/5 + 6 \cdot 1/5$$

$= \mu_{E(Y)}$

$\frac{2+2+4+6}{\;} = 14$

$P(X \leq x)$
Cumulative distri-

$\frac{\quad}{5} - \frac{\quad}{5}$ //

bution function

$\mu \rightarrow$ average of random variable

Variance and standart deviation

I am offered to invest \$800 2 investment accounts:

1- will give me \$1000 in one year

2- will give me either 500\$ or \$1500 (equally likely) in one year

which one should I choose?

$P(X=1000)=1 \; - \; Ex = 1000$

$\rightarrow$ expected value

$Var(x) = 1000 - 1000 = 0$

Same

Y: $\frac{1}{2}$ 500, $\frac{1}{2}$ 1500 $\; - \; Ex = 500 \cdot \frac{1}{2} + 1500 \cdot \frac{1}{2} = 1000$

$\downarrow$ Variance

Only expectation is not sufficient

Bu ikisinin expectationları aynı çıktı yani bunlar yeterli değil

O yüzden bizim variance and standart deviation'a ihtiyacımız var

$Var(Y) = (500-1000)^2 + (1500-1000)$

$\downarrow$ something different then zero and standart deviation'a ihtiyacımız var

The variance of a RV X with $EX = \mu_X$

Burda çıkan değer expectation ya da average-dan farklı değer gösteriyor

$$\underbrace{Var(x)}_{\sigma^2} = E[(X - \underbrace{\mu_X}_{mean})^2]$$

$\underbrace{\sigma^2}_{variance}$

Variance bazı kitaplarda böyle gösterilir

σ^2 variance

σ standart deviation

Var = 0

↳ yukarıda variance'ım 0 olucak

ama 500, 1500 değerlerini koyduğumuzda variance > 0 olur

Definition → standart deviation is a square root of variance

$$\sigma = Std(X) = \sqrt{Var(x)}$$

↳ standart deviation varyansın karekökü'dür

standart deviation

Covariance and Correlation

Ex, var. std. dev → distribution of a single R.V

Two R.V → covariance, correlation (iki tane random variables varsa bunları kullan)

→ Covariance measure the association of two R.V

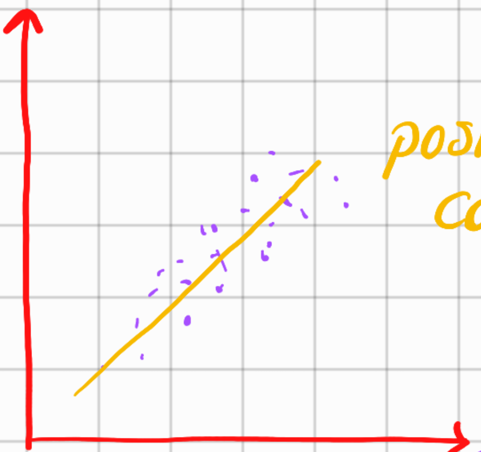
Covariance: Two R.V X and Y

Cov(X,Y) is defined as =

$= E(X-EX).E(Y-EY)$

$E(XY)-E(X).E(Y)$ → proof see book

→ covariance formulu

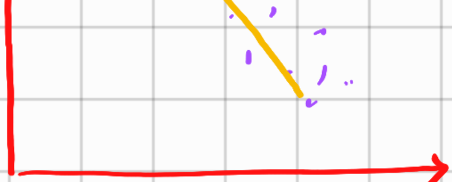


positive covariance

Covaryans 0'dan büyükse bu ikisinin de pozitif olduğu anlamına gelir

$Cov(X,Y) > 0$ } Both increasing

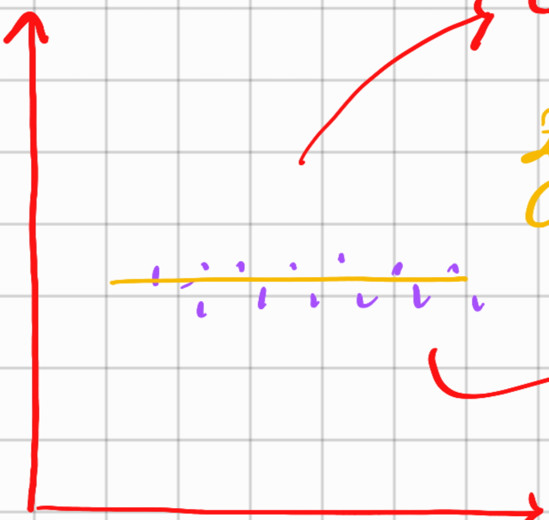
negative covariance → Burda bunlardan biri artıyor biri azalıyor



$Cor(X,Y) < 0$ → one of X,Y is increasing and the other one is decrease

Covaryans 0 ise x ve y arasında bir ilişki yok-tur

zero Covariance



Bu x ve y arasında herhangi bir ilişki olmadığı anlamına gelir

$Cor(X,Y) = 0$ no relation

Correlation

Definition

Like covariance, the values of correlation is between [-1,1]

$$\rho = \frac{Cor(x,y)}{(std\ X)(std\ Y)} = \frac{Cor(x,y)}{\sigma_x \sigma_y}$$

Bu ro olarak okunuyor p harfi değil bu

Sigma

eğer covariance

1 or -1 $\longrightarrow$ perfect correlation $\longrightarrow$ 1 veya -1 ise biz buna perfect correlation deriz

$Cor(X,Y) < 0 \longrightarrow$ small X, large Y

$cor(X,Y) = 0 \longrightarrow$ X and Y are uncorrelated

Ex: A program consists of two modules. the number of errors in the first module X and the number of errors in the second module Y have joint distribution

$P(X,Y)$

$P(0,0) = P(0,1) = P(1,0) = 0.2$ $\longrightarrow$ joint dist

$P(1,1) = P(1,2) = P(1,3) = 0.1$

$P(0,2) = P(0,3) = 0.05$

a.) marginal prob X and Y $P(x)$
$P(Y)$

b.) X and Y are independent ?

c.) $P(X,Y) = ?$

a-) $Var(X) = ?$ $Var(Y) = ?$

	0	1	2	3
0	0.2	0.2	0.05	0.05
1	0.2	0.1	0.1	0.1

$P_Y(0)$ $P_Y(3)$

$$P_X(0) = \sum_Y P(x,y) = 0.2 + 0.2 + 0.05 + 0.05 = 0.5$$

$$P_X(1) = 0.5$$

$$P_Y(0) = \sum_X P(x,0) = P(0,0) + P(1,0) = 0.4$$

$$P_Y(1) = 0.3$$

$$P_Y(2) = 0.15$$

$$P_Y(3) = 0.15$$

b) Are X and Y independent?

$P(X,Y) = P(X).P(Y) \longrightarrow$ X and Y are independent

Proof with contradictio

assume X and Y are independent

$X = 0$ and $Y = 1$

$$P_{X,Y}(0,1) = P_X(0)\,P_Y(1)$$

$$0.2 = 0.5 \cdot 0.5$$

No they are not equal

X and Y are not independent with
profley contradiction

C) $Var(x) = ?$ $Var(Y) = ?$

$$Var(x) = E(X - \mu)$$

$\mu \rightarrow EX$

$$E(x) = \mu_X = \sum_{X=0}^{1} x\,P(x) = 0.05 + 1.05 = 0.5$$

$$E(Y) = \mu_y = \sum_{y=0}^{3} y\,P(y) = \underbrace{0.0.4}_{0} + \underbrace{1.0.3}_{0.3}$$

$$+\underbrace{2.0.15}_{0.3} + \underbrace{3.0.15}_{0.45}$$

$\rightarrow 1.05$

$$Var(x) = E\left[(X - \mu_X)^2\right] = \sum^{1} (0 - 0.5)^2 + (1 - 0.5)^2$$

$X=0$

$\rightarrow 0.5$

$$Var(Y) = E\left[(Y-\mu_Y)^2\right] = (0-1.0.5)^2 + (1-1.05)^2$$

$\underbrace{\qquad\qquad}_{\text{4 values}}$

homework

solve it

d-) $\rho_{X,y} = Cov(X,y)$

$$Cov(X,y) = E((X-\mu_X)(y-\mu_y))$$

$$= E(X.y) - E(x).E(y)$$

Covariance

$a, b, c \rightarrow$ constant

properties of variances and covariances

$$Var(aX + bY + c) = a^2 Var(x) + b^2 Var(Y) + 2ab\, Cov(X,Y)$$

X ve y random variable olur

$$Cov(aX+bY, cZ+dW) = ac\,Cov(X,Z) + ad\,Cov(X,W) + bc\,Cov(Y,Z) + bd\,Cov(Y,W)$$

$$Cov(X,Y) = Cov(Y,X)$$

$$\rho(X,Y) = \rho(Y,X)$$

in particular

$$Var(aX+b) = a^2 Var(X)$$

$$Cov(aX+b, cY+d) = ac\,Cov(X,Y)$$

$$\rho(aX+b, cY+d) = \rho(X,Y)$$

for independent X and Y

$$Cov(X,Y) = 0$$

$$Var(X+Y) = Var(X) + Var(Y)$$

4. partı kaldı