# ECON STAT-1 LECTURE NOTES / CHAPTERS 5&6

## Random Variables

**Discrete CH:5**

X takes Numbers

Ex: # of people who suppors Galatasaray

# of shots scored by a Basketball Player...etc.

**Continuous CH:6**

X takes values over an interval

Ex: weight of a newborn baby

weekly food expenditure ...etc.

5.3 For each of the following indicate if a discrete or a continuous random variable provides the best definition:

a. The number of cars that arrive each day for repair in a two-person repair shop
b. The number of cars produced annually by General Motors
c. Total daily e-commerce sales in dollars
d. The number of passengers that are bumped from a specific airline flight 3 days before Christmas

a) Discrete

b) Discrete

c) Continuous

d) Discrete

**Probability mass function (pmf) of a Discrete random variable;**

Let $f(x)$ be given. If $f(x)$ is a pmf;

(i) $f(x) \geq 0$ → Nonnegative probabilities

(ii) $\sum_x f(x) = 1$ → Total probability is 1.

Also, for discrete case; we have

$P(X = x) = f(x)$ ⇒ X'in bir sayıya eşit olması olasılığını fonksiyonda o sayıyı yerine koyarak buluyoruz.

Random Variable ← X; x → Numerical Value



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Cummulative Probability Distribution;

$$\boxed{F(x) = P(X \leq x) = \sum_{k \leq x} f(k)}$$

5.13 The number of computers sold per day at Dan's Computer Works is defined by the following probability distribution:

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $P(x)$ | 0.05 | 0.10 | 0.20 | 0.20 | 0.20 | 0.15 | 0.10 |

a. $P(3 \leq x < 6) = ?$
b. $P(x > 3) = ?$
c. $P(x \leq 4) = ?$
d. $P(2 < x \leq 5) = ?$

5.13)

a) $P(3 \leq x < 6)$
$= f(3) + f(4) + f(5)$
$= 0{,}20 + 0{,}20 + 0{,}15 = 0{,}55$

b) $P(X > 3) = f(4) + f(5) + f(6)$
$= 0{,}20 + 0{,}15 + 0{,}10 = 0{,}45$

c) $P(X \leq 4) = 1 - P(X > 4) = 1 - [f(5) + f(6)] = 1 - [0{,}15 + 0{,}10]$
$= 0{,}75$

d) $P(2 < x \leq 5) = f(3) + f(4) + f(5) = 0{,}55$

Note; pmf'deki olasılıklardan birisi verilmezse, toplam 1 olacak şekilde verilmeyen olasılığı bul.

* we have; $\boxed{P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)}$

Ex; (5.13 cont.) Cummulative Distribution function is;

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x)$ | 0,05 | 0,10 | 0,20 | 0,20 | 0,20 | 0,15 | 0,10 |
| $F(x)$ | 0,05 | 0,15 | 0,35 | 0,55 | 0,75 | 0,90 | 1,00 |

d) $P(2 < x \leq 5) = F(5) - F(2) = 0{,}90 - 0{,}35 = 0{,}55$

c) $P(X \leq 4) = F(4) = 0{,}75$   b) $P(X > 3) = 1 - F(3) = 0{,}45$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* Also note the following;

WPT; X is more than 3 → $P(X>3) = 1 - F(3)$

X is at least 3 → $P(X \geq 3) = 1 - F(2)$

X is less than 3 → $P(X<3) = F(2)$

X is at most 3 → $P(X \leq 3) = F(3)$

Expectation, Variance & Standard Deviation;

Remember; Expected Value, Mean & Average have the same meaning. Expected Value of a Random Variable X is its long run average.

We have;

$$\boxed{\mu = E(X) = \sum_x x . f(x)}$$

* Expectation of any function g(x) is;

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

* To find Variance, we need $E(X^2) = \sum x^2 . f(x)$

$$\boxed{\sigma^2 = Var(X) = E(X^2) - \mu^2} = \underbrace{E[(X-\mu)^2]}_{\text{Definition Formula}}$$

* Remember; $\sigma$ : Standard Deviation.

$$\boxed{\sigma = \sqrt{\sigma^2}}$$

$\sigma^2$ is mean squared deviation around the mean.



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

## Linear Functions of a Random Variable;

Let $W = a + bX$. We have;

$$E(W) = E(a+bX) = a + bE(X)$$

$$Var(W) = Var(a+bX) = Var(bX) = b^2 \cdot Var(X)$$

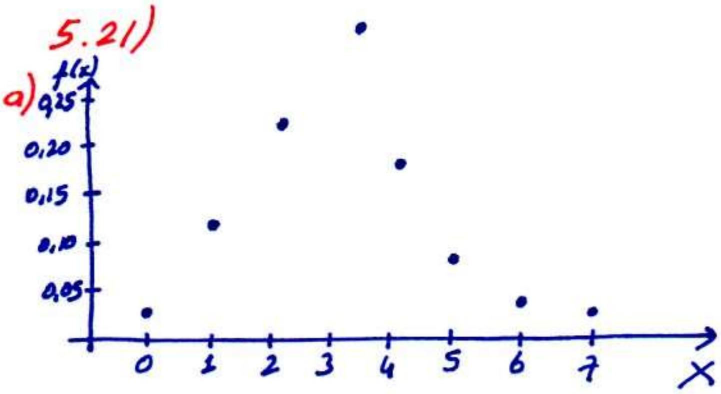
5.21 A municipal bus company has started operations in a new subdivision. Records were kept on the numbers of riders from this subdivision during the early-morning service. The accompanying table shows proportions over all weekdays.

| Number of riders | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Proportion | 0.02 | 0.12 | 0.23 | 0.31 | 0.19 | 0.08 | 0.03 | 0.02 |
| $F(x)$ | 0.02 | 0.14 | 0.37 | 0.68 | 0.87 | 0.95 | 0.98 | 1 |

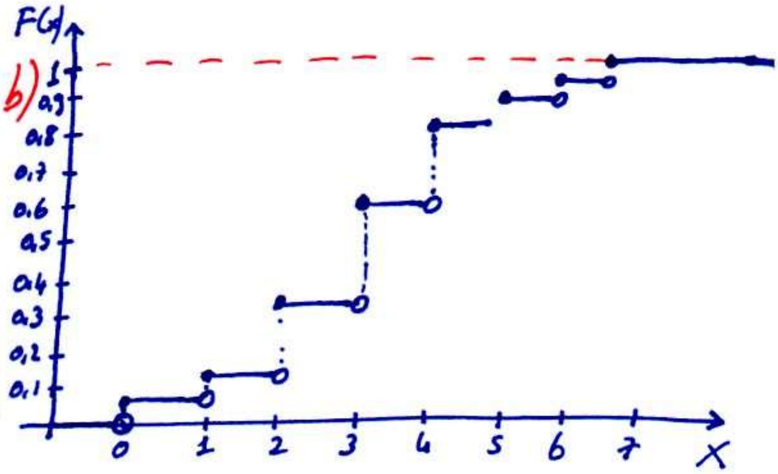
a. Draw the probability function.
b. Calculate and draw the cumulative probability function.
c. What is the probability that on a randomly chosen weekday there will be at least four riders from the subdivision on this service?
d. Two weekdays are chosen at random. What is the probability that on both of these days there will be fewer than three riders from the subdivision on this service?
e. Find the mean and standard deviation of the number of riders from this subdivision on this service on a weekday.
f. If the cost of a ride is 50 cents, find the mean and standard deviation of the total payments of riders from this subdivision on this service on a weekday.

5.21)

a)




b)




c) $P(X \geq 4) = 1 - P(X \leq 3) = 1 - F(3) = 1 - 0{,}68 = 0{,}32$

d) $P(X \leq 3) = F(3) = 0{,}68$

f-cont) $\text{Std-dev}(50X) = \sqrt{5000} = 70{,}71$

$P(\text{Both are less than } 3) = F(3) \cdot F(3) = 0{,}68^2 = 0{,}4624$

e) $\mu = E(X) = \sum x \cdot f(x) = 0 \cdot 0{,}02 + 1 \cdot 0{,}12 + \dots + 7 \cdot 0{,}02 = 2{,}99$

$E(X^2) = \sum x^2 \cdot f(x) = 0^2 \cdot 0{,}02 + 1^2 \cdot 0{,}12 + \dots + 7^2 \cdot 0{,}02 = 10{,}93$

$\sigma^2 = Var(X) = 10{,}93 - 2{,}99^2 \cong 2 \Rightarrow \sigma = \sqrt{2} = 1{,}41$

f) $E(50X) = 50 \cdot E(X) = 50 \cdot 2{,}99 = 149{,}5$ cents ; $Var(50X) = 2500 \cdot Var(X) = 5000$

5.22 a. A very large shipment of parts contains 10% defectives. Two parts are chosen at random from the shipment and checked. Let the random variable $X$ denote the number of defectives found. Find the probability function of this random variable.

b. A shipment of 20 parts contains 2 defectives. Two parts are chosen at random from the shipment and checked. Let the random variable $Y$ denote the number of defectives found. Find the probability function of this random variable. Explain why your answer is different from that for part (a).

c. Find the mean and variance of the random variable $X$ in part (a).

d. Find the mean and variance of the random variable $Y$ in part (b).

5.22) a) $p = 0{,}10$ for each part

$n = 2$ ; $X$: # of defective parts

$P(X=2) = 0{,}10 \cdot 0{,}10 = 0{,}01$

$P(X=1) = 0{,}10 \cdot 0{,}90 + 0{,}90 \cdot 0{,}10 = 0{,}18$

$P(X=1) = 0{,}9 \cdot 0{,}9 = 0{,}81$

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $f_x(x)$ | 0,81 | 0,18 | 0,01 |

b) 2 def. / 18 Nondef. — $T = 20$ parts → Select $n=2$ parts.

$$P(Y=0) = \frac{\binom{2}{0}\binom{18}{2}}{\binom{20}{2}} = 0{,}805$$

$$P(Y=1) = \frac{\binom{2}{1}\binom{18}{1}}{\binom{20}{2}} = 0{,}189 \qquad P(Y=2) = \frac{\binom{2}{2}\binom{18}{0}}{\binom{20}{2}} = 0{,}001$$

| $y$ | 0 | 1 | 2 |
|---|---|---|---|
| $f_y(y)$ | 0,805 | 0,189 | 0,006 |

c) $\mu_x = E(X) = 0 \cdot 0{,}81 + 1 \cdot 0{,}18 + 2 \cdot 0{,}01 = 0{,}2$

$E(X^2) = 0^2 \cdot 0{,}81 + 1^2 \cdot 0{,}18 + 2^2 \cdot 0{,}01 = 0{,}22$

$\sigma_x^2 = Var(X) = 0{,}22 - 0{,}02^2 = 0{,}18$ ; $\sigma = \sqrt{0{,}18} = 0{,}424$

d) $\mu_y = E(Y) = 0 \cdot 0{,}805 + 1 \cdot 0{,}189 + 2 \cdot 0{,}00\cancel{6} = 0{,}2$

$E(Y^2) = 0^2 \cdot 0{,}805 + 1^2 \cdot 0{,}189 + 2^2 \cdot 0{,}001 = 0{,}201$

$\sigma_y^2 = Var(Y) = 0{,}201 - 0{,}2^2 = 0{,}161$ ; $\sigma = \sqrt{0{,}161} = 0{,}401$

5.23 A student needs to know details of a class assignment that is due the next day and decides to call fellow class members for this information. She believes that for any particular call the probability of obtaining the necessary information is 0.40. She decides to continue calling class members until the information is obtained. Let the random variable $X$ denote the number of calls needed to obtain the information.

a. Find the probability function of $X$.
b. Find the cumulative probability function of $X$.
c. Find the probability that at least three calls are required.

5.23 a) $p = 0{,}40$

$1 - p = 0{,}60$

$X$: # of calls to obtain necessary information

$$P(X=x) = f(x) = \underbrace{0{,}60^{x-1}}_{x-1 \text{ failures}} \cdot \underbrace{0{,}40}_{x^{th} \text{ call is success}}$$

b) $$F(x) = P(X \le x) = \sum_{k=1}^{x} f(k) = \sum_{k=1}^{x} 0{,}60^{k-1} \cdot 0{,}40$$

$$= 0{,}40 \cdot \frac{1 - 0{,}60^x}{1 - 0{,}60} = 1 - 0{,}60^x$$

c) $$P(X \le 2) = F(2) = 1 - 0{,}60^2 = 0{,}64$$

$$P(X \ge 3) = 1 - F(2) = 1 - 0{,}64 = 0{,}36$$

5.29 An investor is considering three strategies for a \$1,000 investment. The probable returns are estimated as follows:

- *Strategy* 1: A profit of \$10,000 with probability 0.15 and a loss of \$1,000 with probability 0.85
- *Strategy* 2: A profit of \$1,000 with probability 0.50, a profit of \$500 with probability 0.30, and a loss of \$500 with probability 0.20
- *Strategy* 3: A certain profit of \$400

Which strategy has the highest expected profit? Would you necessarily advise the investor to adopt this strategy?

Strategy 1: $E(Profit) = 10000 \cdot 0{,}15 - 1000 \cdot 0{,}85 = 650$

Strategy 2: $E(Profit) = 1000 \cdot 0{,}50 + 500 \cdot 0{,}30 - 500 \cdot 0{,}2 = 550$

Strategy 3: $E(Profit) = 400 \cdot 1 = 400$

Strategy 1 is most profitable. However, it is too risky because all money can be lost with a high probability. I think strategy 2 is better.

# Binomial Distribution;

$X \sim Binomial(n;p)$

where $n$: # of independent trials
$p$: Fixed probability of success
$X$: # of success.

Then; $P(X=x)=f(x)=\binom{n}{x}\cdot p^{x}\cdot(1-p)^{n-x}$

$\mu=E(X)=n.p \qquad \sigma^{2}=Var(X)=n.p.(1-p)$

* Remember; $\binom{n}{x}=\frac{n!}{x!(n-x)!}=\frac{n.(n-1).\ \dots\ .(n-x+1)}{x!}$

5.39 A company installs new central heating furnaces and has found that for 15% of all installations a return visit is needed to make some modifications. Six installations were made in a particular week. Assume independence of outcomes for these installations.

a. What is the probability that a return visit will be needed in all of these cases?
b. What is the probability that a return visit will be needed in none of these cases?
c. What is the probability that a return visit will be needed in more than one of these cases?

5.39) $X$: # of installations that need a return visit.

$X \sim Binomial(n=6;p=0{,}15)$

a) $P(X=6)=\binom{6}{6}.0{,}15^{6}.0{,}85^{0}=0{,}000011$

b) $P(X=0)=\binom{6}{0}.0{,}15^{0}.0{,}85^{6}=0{,}377$

c) $P(X>1)=1-P(X\leq 1)=1-[f(0)+f(1)]$

$P(X=1)=f(1)=\binom{6}{1}.0{,}15^{1}.0{,}85^{4}=0{,}470$

$P(X>1)=1-[0{,}377+0{,}470]=0{,}153$

d) Find expected # of returns and std. dev.

$\mu=E(X)=6.0{,}15=0{,}9 \qquad \sigma^{2}=Var(X)=6.0{,}15.0{,}85=0{,}765$

$\sigma=\sqrt{0{,}765}=0{,}875$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

5.46 A campus finance officer finds that, for all parking tickets issued, fines of 78% are paid. The fine is $2. In the most recent week 620 parking tickets have been issued.

a. Find the mean and standard deviation of the number of these tickets for which the fines will be paid.
b. Find the mean and standard deviation of the amount of money that will be obtained from the payment of these fines.

5.46) X: # of parking tickets paid.

$X \sim Binomial\,(n=620;\ p=0{,}78)$

a) $E(X) = 620 \cdot 0{,}78 = 483{,}6$

$Var(X) = 620 \cdot 0{,}78 \cdot 0{,}22 = 106{,}4$

$\sigma = \sqrt{106{,}4} = 10{,}31$

b) $E(2X) = 2E(X) = 2 \cdot 483{,}6 = 967{,}2$

$Var(2X) = 4\,Var(X) = 4 \cdot 106{,}4 = 425{,}6$

$Std.Dev(2X) = \sqrt{425{,}6} = 20{,}63$

5.48 The following two acceptance rules are being considered for determining whether to take delivery of a large shipment of components:

- A random sample of 10 components is checked, and the shipment is accepted only if none of them is defective.
- A random sample of 20 components is checked, and the shipment is accepted only if no more than 1 of them is defective.

Which of these acceptance rules has the smaller probability of accepting a shipment containing 20% defectives?

5.48) X: # of defective items.

$X \sim Binomial\,(n;\ p=0{,}20)$

if $n=10$ and $P(Accept) = P(X=0)$

$P(Accept) = \binom{10}{0} 0{,}20^0 \cdot 0{,}80^{10} = 0{,}107$

if $n=20$ and $P(Accept) = P(X \leq 1)$

$P(Accept) = P(X \leq 1) = f(0) + f(1) = \binom{20}{0} 0{,}20^0 \cdot 0{,}80^{20} + \binom{20}{1} 0{,}20^1 \cdot 0{,}80^{19}$

$= 0{,}0692$

First case $n=10$ and $P(Accept) = P(X=0)$ has higher probability of acceptence, second has smaller.

5.49 A company receives large shipments of parts from two sources. Seventy percent of the shipments come from a supplier whose shipments typically contain 10% defectives, while the remainder are from a supplier whose shipments typically contain 20% defectives. A manager receives a shipment but does not know the source. A random sample of 20 items from this shipment is tested, and 1 of the parts is found to be defective. What is the probability that this shipment came from the more reliable supplier? [*Hint:* Use Bayes' theorem.]



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

5.49) Let, Source A has 10% defective and $P(A)=0{,}70$

Source B has 20% defective $P(B)=0{,}30$

$X \sim \text{Binomial}(n=20;\, p)$

For source A, we have;

$$P(X=1|A)=\binom{20}{1}\cdot 0{,}10^{1}\cdot 0{,}90^{19}=0{,}270$$

$$P(X=1|B)=\binom{20}{1}\,0{,}20^{1}\cdot 0{,}80^{19}=0{,}058$$

$$P(A|X=1)=\frac{P(A)\cdot P(X=1|A)}{P(A)\cdot P(X=1|A)+P(B)\cdot P(X=1|B)}$$

$$=\frac{0{,}70\cdot 0{,}270}{0{,}70\cdot 0{,}270+0{,}30\cdot 0{,}058}=\underline{0{,}916}$$

## Hypergeometric Distribution

S: Type I
N−S: Type II
Total: N objects

→ Select n objects randomly
X: # of Type-I objects.

Then; $X \sim \text{Hypergeometric}(N;\, S;\, n)$

$$P(X=x)=f(x)=\frac{\binom{S}{x}\binom{N-S}{n-x}}{\binom{N}{n}}$$

$$\mu=E(X)=n\cdot\frac{S}{N}\qquad \sigma^2=Var(X)=n\cdot\frac{S}{N}\cdot\frac{N-S}{N}\cdot\frac{N-1}{N-n}$$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

5.57 A bond analyst was given a list of 12 corporate bonds. From that list she selected 3 whose ratings she felt were in danger of being downgraded in the next year. In actuality, a total of 4 of the 12 bonds on the list had their ratings downgraded in the next year. Suppose that the analyst had simply chosen 3 bonds randomly from this list. What is the probability that at least 2 of the chosen bonds would be among those whose ratings were to be downgraded in the next year?

$\boxed{\begin{array}{l} 4 \text{ in danger} \\ 8 \text{ NOT in danger} \end{array}}$ → Select $n=3$ bonds

$X$: # of bonds really in danger.

$N=12$ bonds

$X \sim \text{Hypergeometric}(N=12; S=4; n=3)$

$$P(X \geq 2) = f(2) + f(3) = \frac{\binom{4}{2}\binom{8}{1}}{\binom{12}{3}} + \frac{\binom{4}{3}\binom{8}{0}}{\binom{12}{3}} = 0{,}236$$

## Poisson Distribution

$X \sim \text{Poisson}(\lambda)$ $\quad \lambda$: Rate or Mean

$$P(X=x) = f(x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

$$\mu = E(X) = \lambda \; ; \; \sigma^2 = Var(X) = \lambda$$

* Two types of questions

(i) Poisson distribution is mentioned or a discrete random variable, only mean is known ⇒ Use Poisson dist.

(ii) Distribution is Binomial but $n$ is high, $p$ is small (preferably $np \leq 7$) then set $\lambda = n.p$ and use Poisson Approximation to Binomial.

Note that; If $\lambda$ = # of events / unit time, then unit times must be the same. If, for example, 3 events/week and month is asked, use 12 events/month.

**Ex:** On the average, 3 accidents occur in a week at E5 highway. WPT

a) At least 2 accidents will occur next week?

b) 5 accidents will occur next month?

**Ans:** X: # of accidents in a week

$X \sim \text{Poisson}(\lambda = 3)$

a) $P(X \geq 2) = 1 - P(X < 2) = 1 - [f(0) + f(1)]$

$= 1 - \left[ \frac{e^{-3} \cdot 3^0}{0!} + \frac{e^{-3} \cdot 3^1}{1!} \right] = 0{,}801$

b) y: # of accidents in a month.

$y \sim \text{Poisson}(\lambda = 3.4 = 12)$

$P(y = 5) = \frac{e^{-12} \cdot 12^5}{5!} = 0{,}013$

**Ex:** An insurance company holds 6000 policies. Any single policy will result in a claim with probability 0,001. WPT at least 3 claims are made in a given year?

**Ans:** X: # of claims in a year $\sim$ Binomial

$n = 6000$; $p = 0{,}001$; $\lambda = np = 6$

$X \sim \text{Poisson}(\lambda = 6)$

$P(X \geq 3) = 1 - \left[ \frac{e^{-6} \cdot 6^0}{0!} + \frac{e^{-6} \cdot 6^1}{1!} + \frac{e^{-6} \cdot 6^2}{2!} \right] = 0{,}893$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

# Jointly Distributed Discrete R.V's.

Remember; 3 types of probabilities;

(i) Marginal (ii) Joint (iii) Conditional.

* Given the Joint pmf of X and Y;

↳ We sum over X to find marginal of Y and sum over Y to find Marginal of X

↳ To find $P(X \mid Y=y)$, we use $P_{X|Y}(X \mid Y=y) = \frac{P(X=x, Y=y)}{P_Y(Y=y)}$ for every choice of x.

↳ X and Y are independent if $P(x,y) = P_X(x) \cdot P_Y(y)$

↳ To find $Corr(X,Y)$;

(i) $E(X) = \sum_x x \cdot P_x(x)$ ; $E(Y) = \sum_y y \cdot P_y(y)$ ;

$E(X^2) = \sum_x x^2 \cdot P_x(x)$ ; $E(Y^2) = \sum_y y^2 \cdot P_y(y)$ ;

$$E(XY) = \sum_x \sum_y x \cdot y \cdot P(x,y)$$

(ii) $\mu_x = E(X)$ ; $\sigma_x^2 = E(X^2) - \mu_x^2$

$\mu_y = E(Y)$ ; $\sigma_y^2 = E(Y^2) - \mu_y^2$

$Cov(X,Y) = E(XY) - \mu_x \mu_y$

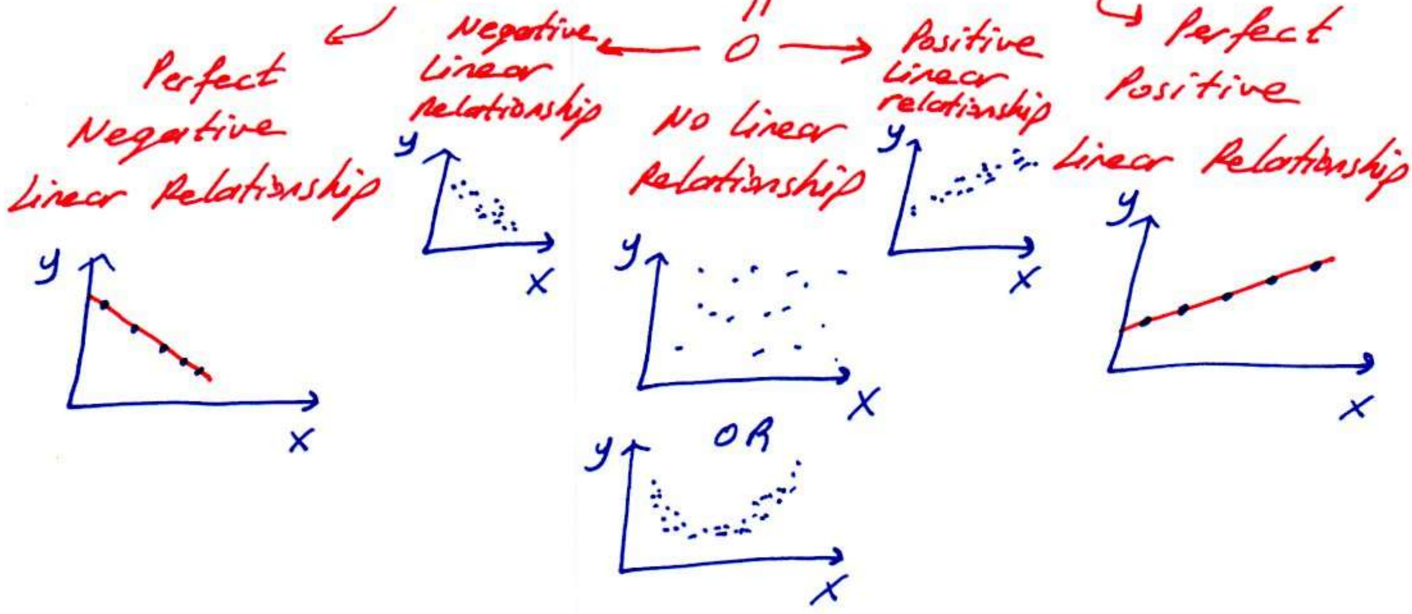
(iii) $\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{\sigma_x^2 \cdot \sigma_y^2}}$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* Remember that, $\rho$ shows the LINEAR relationship between $X$ and $Y$.

$$-1 \leq \rho \leq 1$$




Then; $X, Y$ independent $\Rightarrow$ $Corr(X,Y) = 0$
$(Cov(X,Y) = 0)$

But converse is not true. Although $Cov(X,Y)=0$, there may be some relation between $X$ and $Y$ and $X$ and $Y$ may not be independent. $Cov(X,Y) \neq 0$ implies linear dependence between $X$ and $Y$.

$Cov(X,Y) \neq 0 \Rightarrow X, Y$ NOT independent
$(Corr(X,Y) \neq 0)$



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

# Linear Combinations of Random Variables;

$$E(aX+bY+c)=aE(X)+bE(Y)+c$$

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

5.77 Consider the joint probability distribution

| | | X | |
|---|---|---|---|
| | | 1 | 2 |
| Y | 0 | 0.70 | 0.0 |
| | 1 | 0.0 | 0.30 |

a. Compute the marginal probability distributions for X and Y.
b. Compute the covariance and correlation for X and Y.
c. Compute the mean and variance for the linear function $W=3X+4Y$.

5.77)

| | | X | | |
|---|---|---|---|---|
| | | 1 | 2 | $P(Y=y)$ |
| y | 0 | 0,70 | 0,0 | 0,70 |
| | 1 | 0,0 | 0,30 | 0,30 |
| | $P(X=x)$ | 0,70 | 0,30 | 1,00 |

a)

| X | 1 | 2 |
|---|---|---|
| $P_x(X=x)$ | 0,70 | 0,30 |

| y | 0 | 1 |
|---|---|---|
| $P_y(Y=y)$ | 0,70 | 0,30 |

b) $\mu_x=E(X)=1.0{,}70+2.0{,}30=1{,}3$ $\quad \mu_y=E(Y)=0{,}3$

$E(X^2)=1^2.0{,}70+2^2.0{,}30=1{,}9$ $\quad E(Y^2)=0{,}3$

$\sigma_x^2=1{,}9-1{,}3^2=0{,}21$ $\quad \sigma_y^2=0{,}3-0{,}3^2=0{,}21$

$E(XY)=1.0.0{,}70+2.0.0{,}0+1.1.0{,}0+2.1.0{,}30=0{,}60$

$Cov(X,Y)=0{,}60-1{,}3.0{,}3=0{,}21$

$\rho=\dfrac{0{,}21}{\sqrt{0{,}21.0{,}21}}=1$ ⤳ Perfect Positive Linear Relationship

c) $Cov(X,Y)\neq 0$ so they are NOT independent.



Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

5.83 A real estate agent is interested in the relationship between the number of lines in a newspaper advertisement for an apartment and the volume of inquiries from potential renters. Let volume of inquiries be denoted by the random variable X, with the value 0 for little interest, 1 for moderate interest, and 2 for strong interest. The real estate agent estimated the joint probability function shown in the accompanying table.

| Number of Lines (Y) | Number of Enquiries (X) 0 | 1 | 2 |
|---|---|---|---|
| 3 | 0.09 | 0.14 | 0.07 |
| 4 | 0.07 | 0.23 | 0.16 |
| 5 | 0.03 | 0.10 | 0.11 |

a. Find the joint cumulative probability function at $X = 1$, $Y = 4$, and interpret your result.
b. Find and interpret the conditional probability function for Y, given $X = 0$.
c. Find and interpret the conditional probability function for X, given $Y = 5$.
d. Find and interpret the covariance between X and Y. and correlation
e. Are number of lines in the advertisement and volume of inquiries independent of one another?

a) $F(X=1, y=4)$

$= P(X \leq 1, y \leq 4)$

$= 0{,}09 + 0{,}14 + 0{,}23 + 0{,}07 = 0{,}53$

b) $P(X=0) = 0{,}09 + 0{,}07 + 0{,}03 = 0{,}19$

| $y$ | 3 | 4 | 5 |
|---|---|---|---|
| $P(y \mid X=0)$ | $\frac{0{,}09}{0{,}19} = 0{,}47$ | $\frac{0{,}07}{0{,}19} = 0{,}37$ | $\frac{0{,}03}{0{,}19} = 0{,}16$ |

c)

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(x \mid y=5)$ | $\frac{0{,}03}{0{,}24} = 0{,}125$ | $\frac{0{,}10}{0{,}24} = 0{,}417$ | $\frac{0{,}11}{0{,}24} = 0{,}458$ |

$P(y=5) = 0{,}03 + 0{,}10 + 0{,}11 = 0{,}24$

d) $\mu_x = E(X) = 0 \cdot 0{,}19 + 1 \cdot 0{,}47 + 2 \cdot 0{,}34 = 1{,}15$

$E(X^2) = 0^2 \cdot 0{,}19 + 1^2 \cdot 0{,}47 + 2^2 \cdot 0{,}34 = 1{,}83$

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $f_x(x)$ | 0,19 | 0,47 | 0,34 |

$\sigma_x^2 = 1{,}83 - 1{,}15^2 = 0{,}51$

$\mu_y = E(y) = 3 \cdot 0{,}30 + 4 \cdot 0{,}46 + 5 \cdot 0{,}24 = 3{,}94$

$E(y^2) = 3^2 \cdot 0{,}30 + 4^2 \cdot 0{,}46 + 5^2 \cdot 0{,}24 = 16{,}06$

| $y$ | 3 | 4 | 5 |
|---|---|---|---|
| $f_y(y)$ | 0,30 | 0,46 | 0,24 |

$\sigma_y^2 = 16{,}06 - 3{,}94^2 = 0{,}54$

$E(xy) = 0 \cdot 3 \cdot 0{,}09 + 0 \cdot 4 \cdot 0{,}07 + \text{-----} + 2 \cdot 5 \cdot 0{,}11 = 4{,}64$

$Cov(x, y) = 4{,}64 - 1{,}15 \cdot 3{,}94 = 0{,}109$

$\rho = Corr(x, y) = \frac{0{,}109}{\sqrt{0{,}51 \cdot 0{,}54}} = 0{,}208 \neq 0$

e) $x, y$ are NOT independent.