Probability and Statistics

DR. AHMED TAYEL

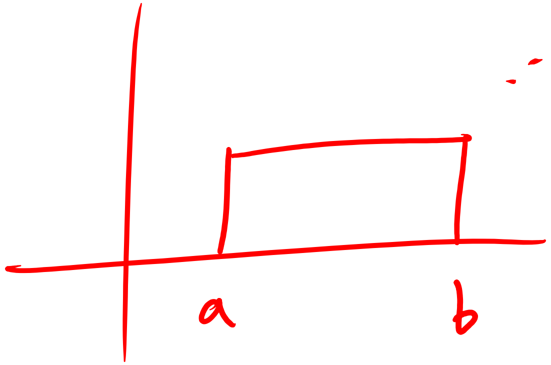
Department of Engineering Mathematics and Physics, Faculty of Engineering, Alexandria University

ahmed.tayel@alexu.edu.eg

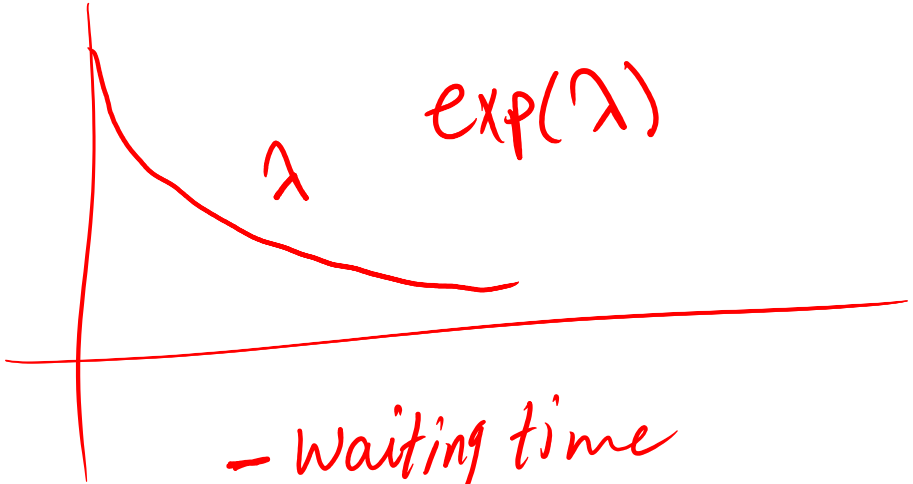
Outline

Some important continuous random variables

- Uniform PDF ✓ "Completely random process"
- Exponential PDF
- Normal PDF



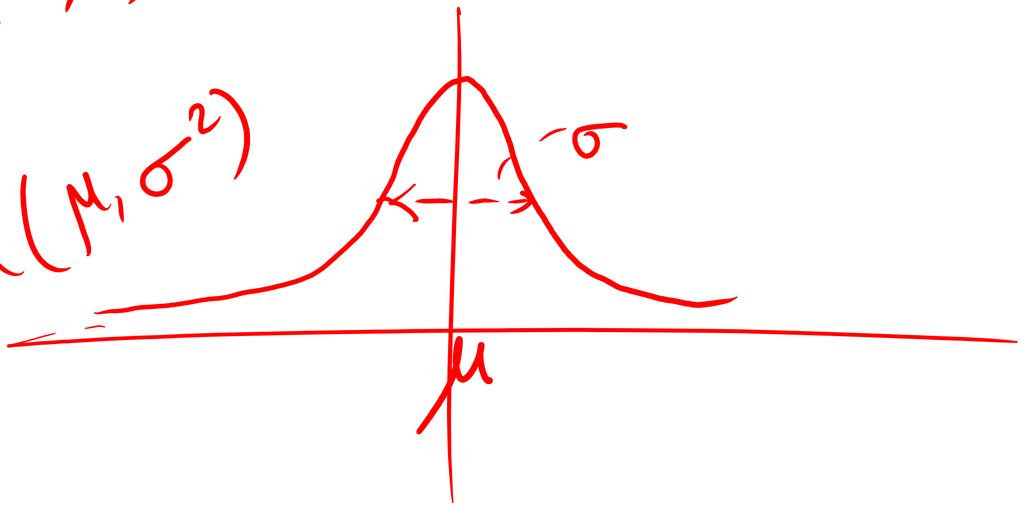
uniform(a,b)



$\exp(\lambda)$

- waiting time
- Life time of devices

Normal(μ,σ^2)



Uniform random variable

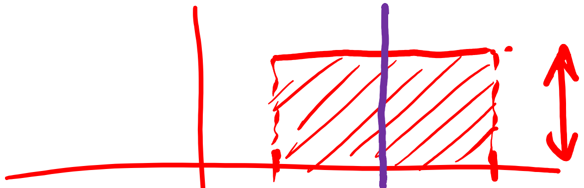
$P(X>5) = \int_5^{\infty} \text{pdf}\, dx = 1 - P(X<5) = 1 - CDF(5)$

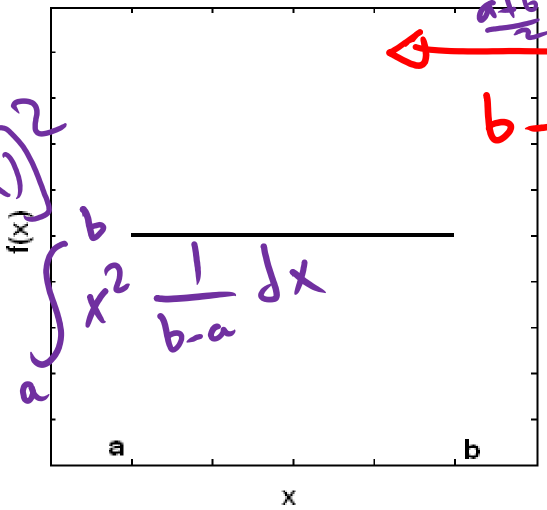
PDF

$$f_X(x) = \frac{1}{b-a}, \qquad a \le x \le b$$

constant

$$X \sim Uniform(a,b)$$





Mean & variance

$$E(X) = \frac{a+b}{2}, \qquad Var(X) = \frac{(b-a)^2}{12}$$

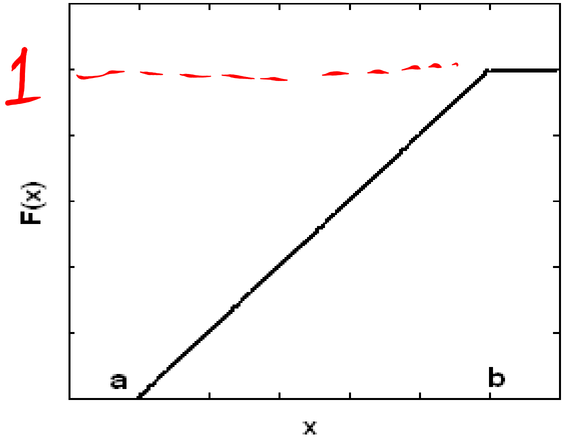
$E(x^2) - [E(x)]^2$

$\int_a^b x^2 \frac{1}{b-a}\, dx$

$F_X(x) = P(X \le x) = \int_{-\infty}^{x} \text{pdf}\, dx$

CDF

$$F_X(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$



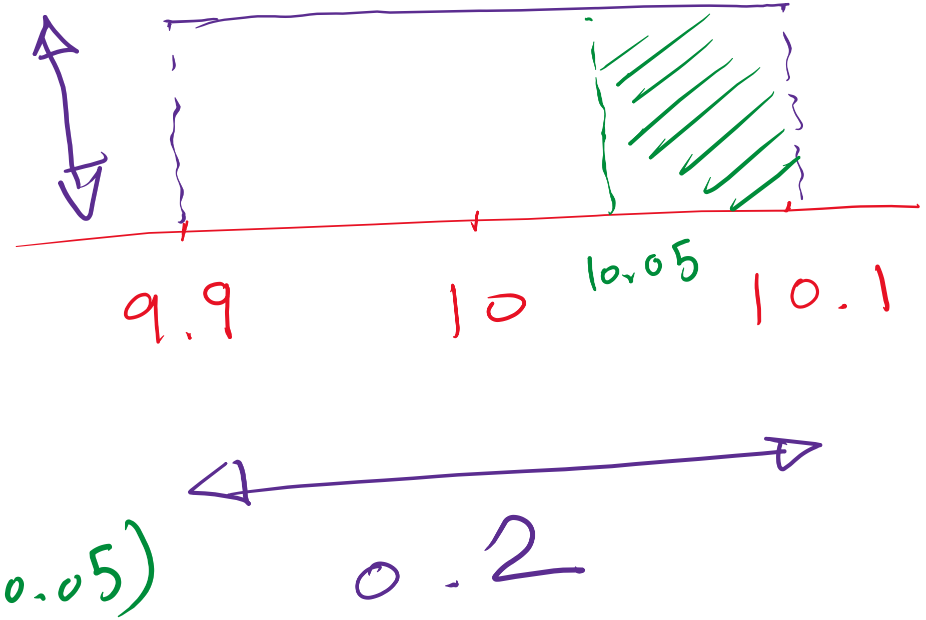
CDF non-decreasing fn.

Example

The voltage across a resistor is *uniformally* distributed over the interval (10 ± 0.1).

What the probability that the voltage exceeds 10.05?

$$\frac{1}{0.2} = 5$$



$$P(V > 10.05)$$

$= \text{shaded area} = (10.1 - 10.05) * 5$

$= 0.25$

or $\int_{10.05}^{10.1} 5\,dx = 5\,x\Big|_{10.05}^{10.1} = 5\,(10.1 - 10.05)$

$= 0.25$

or from shape symmetry = quarter of the total length

Example

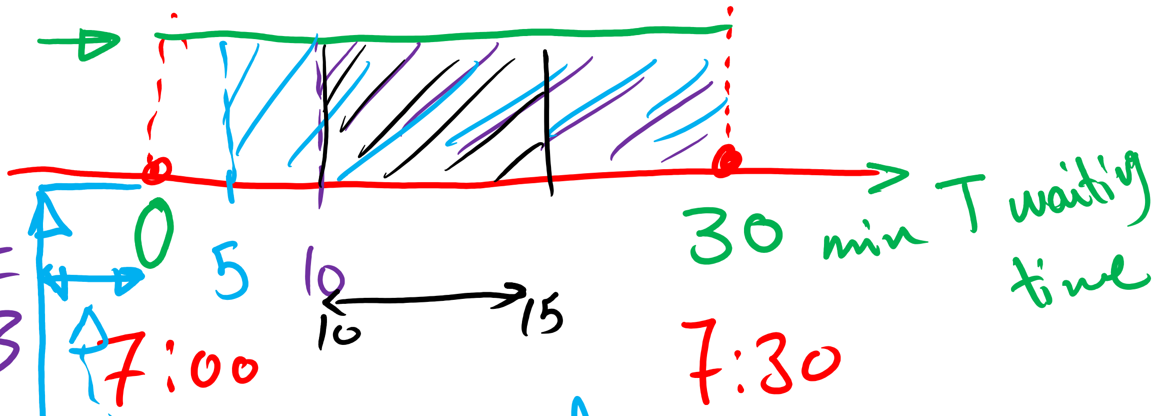
Mr. Ali arrives at a bus station every day at 7:00 A.M. If a bus arrives at a random time between 7:00 A.M. and 7:30 A.M. Most rand. ✓ uniform

(a) Find the probability that he waits more than 10 minutes. الاحصائي المطلوب → random

(b) Repeat (a) if Mr. Ali arrives at 6:55 A.M.

(c) If it is known that the waiting time exceeds 10 minutes, what is the probability that it does not exceed 15 mins.

أخر $\dfrac{1}{30-0} = \dfrac{1}{30}$



ⓐ $T \sim unif(0, 30)$

$P(T > 10) = (30-10) * \dfrac{1}{30} = \dfrac{20}{30} = \dfrac{2}{3}$

ⓑ $P(T > 5) = (30-5) * \dfrac{1}{30} = \dfrac{25}{30} = \dfrac{5}{6}$ 6:55

5 mins already waite

+5 already waited before 7:00 AM

ⓒ $P(T < 15 \mid T > 10) = \dfrac{P(10 < T < 15)}{P(T > 10)} = \dfrac{(15-10) * \frac{1}{30}}{2/3} =$

From(a)

اللي ورى يعلم الحالة continuous

Example

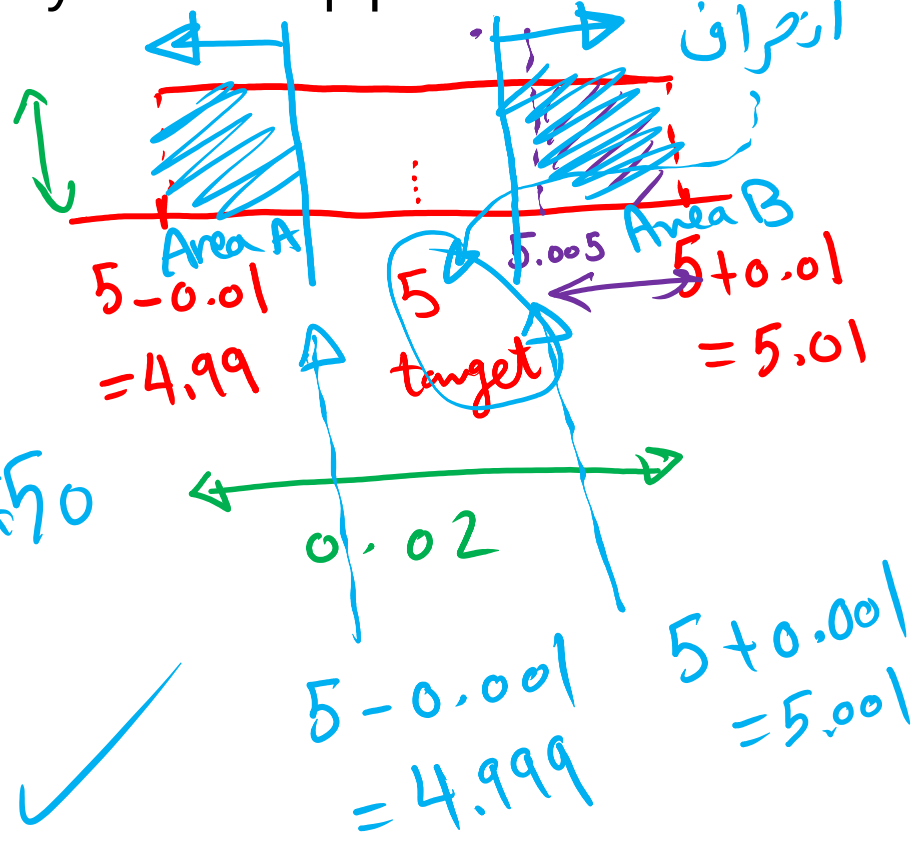
A production line is expected to produce pipes of diameter 5 mm. Actual product is assumed to follow a uniform distribution over the interval [5 ± 0.01].

(a) Compute the probability that the diameter of a randomly chosen pipe is greater than 5.005.

(b) Compute the probability that the diameter of a randomly chosen pipe has a deviation greater than 0.001.

(a) $P_a = (5.01 - 5.005) * 50 = \checkmark$

$\frac{1}{0.02} = 50$



(b) $P_b = \text{Area}_A + \text{Area}_B = (4.999 - 4.99) * 50$

$+ (5.01 - 5.001) * 50$

↳ or using complement

$= 1 - (5.001 - 4.999) * 50 = \checkmark$

Example

A production line is expected to produce pipes of diameter 5 mm. Actual product is assumed to follow a uniform distribution over the interval [5 ± 0.01].

(a) Compute the probability that the diameter of a randomly chosen pipe is greater than 5.005.

(b) Compute the probability that the diameter of a randomly chosen pipe has a deviation greater than 0.001. $= q$

(b) A pipe is not accepted if its diameter has a deviation greater than 0.001.
In a lot of 100 items, what is the probability of getting at most 2 unaccepted items?

$X \sim Bin(100, q)$ --- $P_X(x) = {}^{100}C_x \, q^x (1-q)^{100-x}$

no. of not accepted items

$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

Example

A production line is expected to produce pipes of diameter 5 mm. Actual product is assumed to follow a uniform distribution over the interval [5 ± 0.01].

(a) Compute the probability that the diameter of a randomly chosen pipe is greater than 5.005.

(b) Compute the probability that the diameter of a randomly chosen pipe has a deviation greater than 0.001.

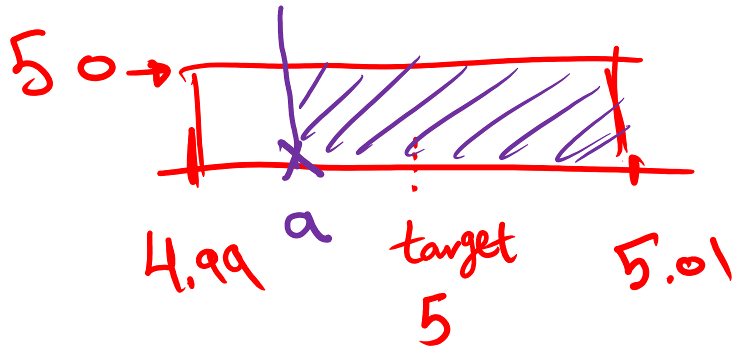
(b) A pipe is not accepted if its diameter has a deviation greater than 0.001.
In a lot of 100 items, what is the probability of getting at most 2 unaccepted items?

d) Determine the pipe diameter that is exceeded in 90% of the production.

$$P(D>a)=0.9$$

$$(5.01-a)*50=0.9$$

$\hookrightarrow$ Get a ✓

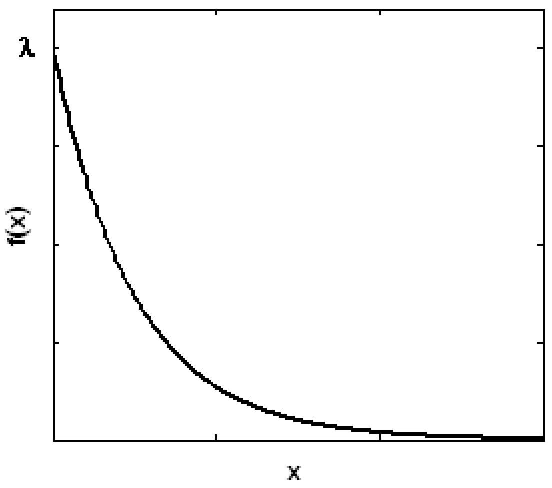


Exponential random variable

PDF

$$f_X(x) = \lambda e^{-\lambda x}, \qquad x \geq 0, \quad \lambda > 0$$

$$X \sim \text{Exponential}(\lambda)$$



Mean & variance

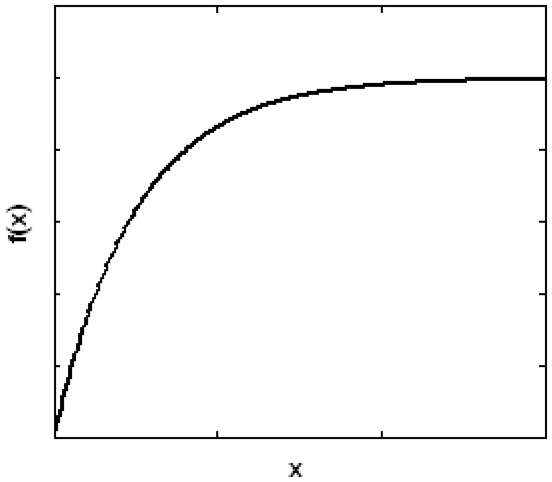
حفظ

$$E(X) = 1/\lambda, \; Var(X) = 1/\lambda^2$$

CDF

حفظ

$$F_X(x) = 1 - e^{-\lambda x}, \qquad x \geq 0,$$
$$= 0, \qquad x < 0$$



Example

The time to failure of a certain device is assumed to follow an exponential distribution of mean 3 years. $= \frac{1}{\lambda} \Rightarrow \lambda = 1/3$

What is the probability that it fails during the first year of operation?

$f_X(x) = \lambda e^{-\lambda x}$

T : time to failure $\sim \exp\left(\frac{1}{3}\right)$

$P(X \le x) = F_X(x) = 1 - e^{-\lambda x}$

$E(X) = \frac{1}{\lambda}$

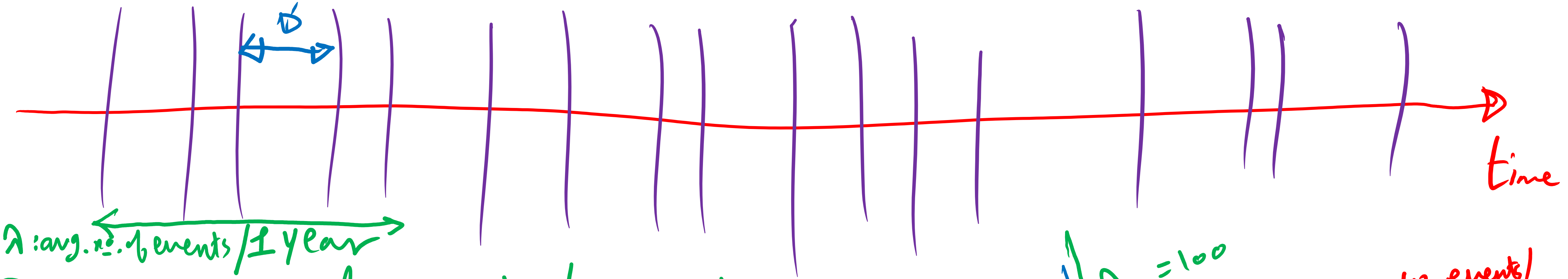
$P(T < 1) = F_X(1)$

$= 1 - e^{-\left(\frac{1}{3}\right)(1)}$

$V(X) = \frac{1}{\lambda^2}$

Number of events ~ Poisson (λ)	$\Longleftrightarrow$	*Time* between events ~ Exponential (λ)

Interarrival time $\sim \exp(\lambda)$



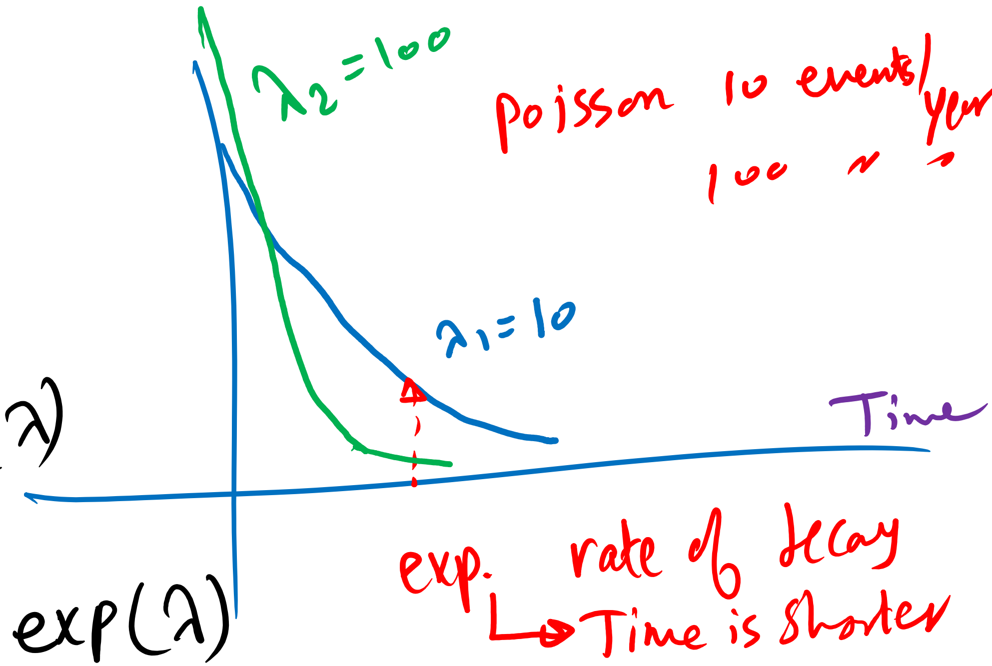
λ : avg. no. of events / 1 year

Poisson : no. of events / unit time

Q : $P(X=2)$ per year

Q : Describe as Poisson

- Q : no. of events $\longrightarrow$ Poisson(λ)
- Q : time betw. events $\longrightarrow$ exp(λ)



Poisson 10 events/year
100 " "

exp. rate of decay
$\rightarrow$ Time is shorter

Example

The number of received telephone calls per minute is assumed to follow a Poisson distribution with mean 2 calls. λ

There is only one person to answer and each call is assumed to have a *fixed* duration of 0.5 minute.

If that person receives a call, what is the probability that the next call will be blocked?

X: no. of calls $\sim$ Poisson(λ)

T: time betw. calls $\sim$ exp(λ)

– one person to answer

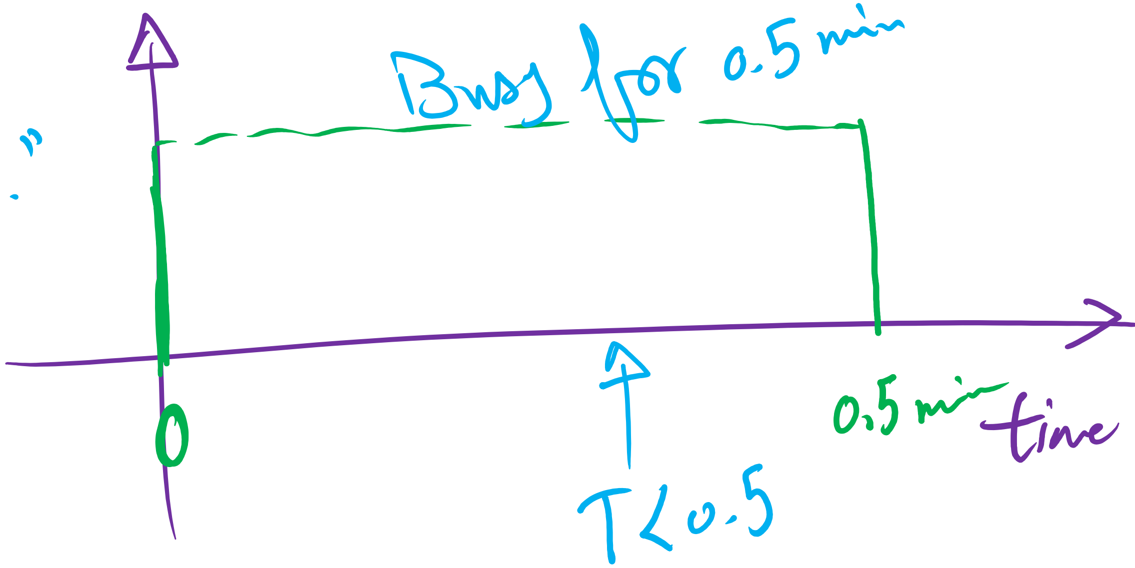
– Each call fixed 0.5 min

$P(T<0.5) = P_{blocked}$

"time → exp."

$= 1 - e^{-2(0.5)} = 1 - e^{-1}$

$= 0.63$



Example

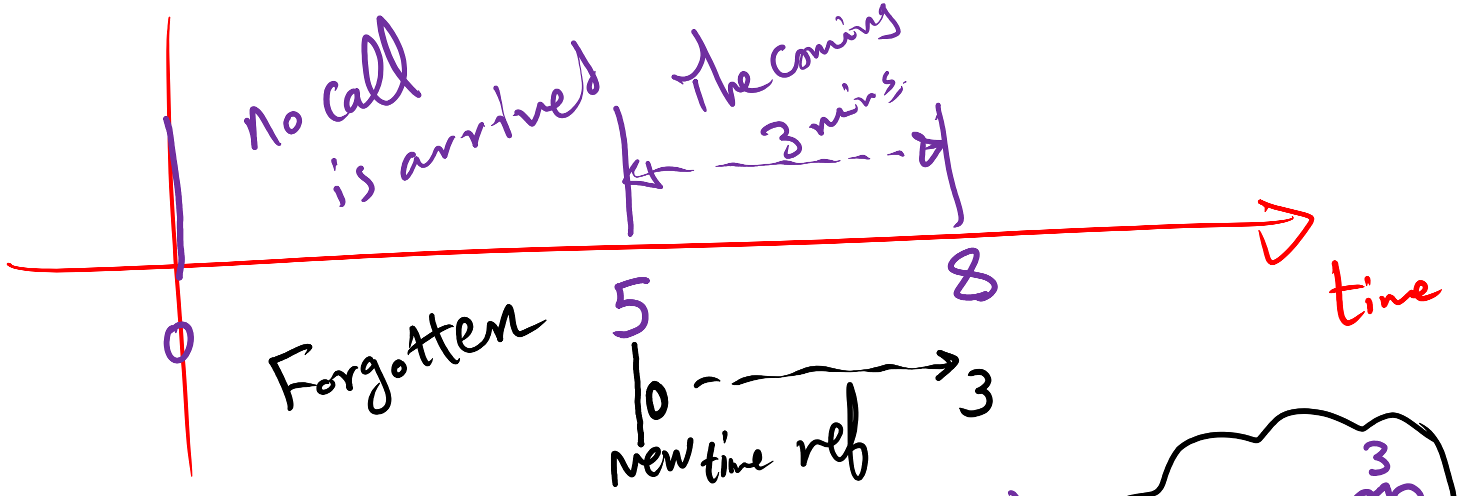
$$F_X(x) = 1 - e^{-\lambda x}$$

The number of received telephone calls per minute is assumed to follow a Poisson distribution with mean 2 calls.

There is only one person to answer and each call is assumed to have a *fixed* duration of 0.5 minute.

If that person receives a call, what is the probability that the next call will be blocked?

If No call arrived in the last 5 mins, what is the probability that the next call will arrive within the coming 3 mins?



$$P(T<8 \mid T>5)$$

$$= \frac{P(5<T<8)}{P(T>5)}$$

$$= \frac{F_T(8) - F_T(5)}{1 - F_T(5)} = \frac{\cancel{1} - e^{-(2)(8)} - (\cancel{1} - e^{-(2)(5)})}{\cancel{1} - [\cancel{1} - e^{-(2)(5)}]} = \frac{e^{-2(5)} - e^{-2(8)}}{e^{-2(5)}} = 1 - e^{-2\overbrace{(8-5)}^{3}}$$

memoryless property of the exp. distn.

Memoryless property of the exponential distribution

Assume that X follows an exponential distribution with rate λ.

$$\Pr(X \geq \overset{5}{x} + \overset{3}{y} | X \geq \overset{5}{x}) = P(X \geq y)$$

Exercise

Normal (Gaussian) random variable

$$\int e^{-x^2} dx$$

Numerical

PDF

ليس لها حل

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty$$

$$-\infty < \mu < \infty, \quad \sigma > 0$$

$$X \sim N(\mu, \sigma^2)$$

Mean & variance

$$E(X) = \mu,$$

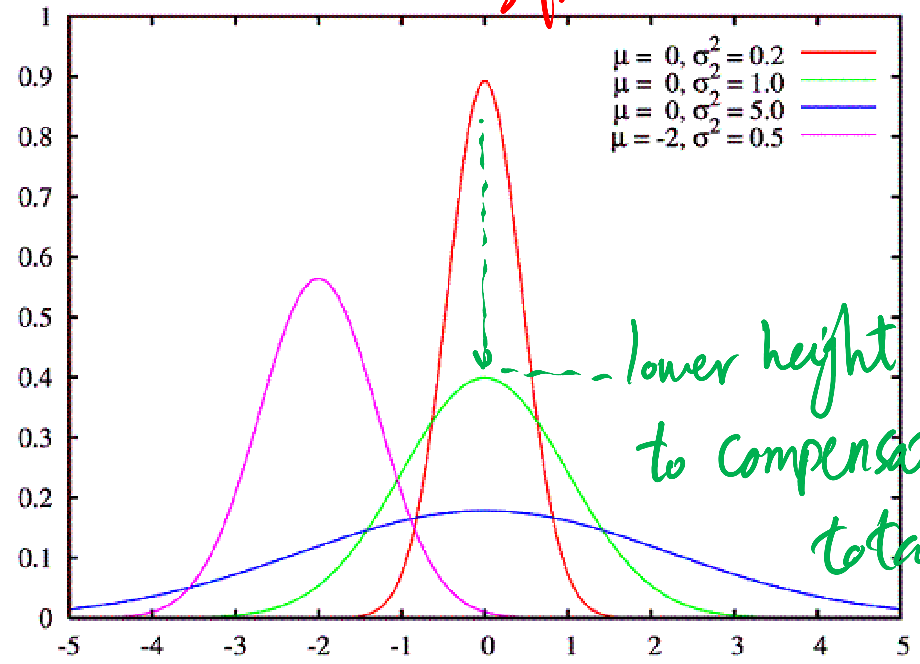
$$Var(X) = \sigma^2$$

CDF

ليس لها حل

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

symmetric



lower height to compensate for total area = 1

From http://en.wikipedia.org

Standard normal (Gaussian) RV

PDF $Z \sim N(\overset{\mu}{0}, \overset{\sigma^2}{1})$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, -\infty < z < \infty$$

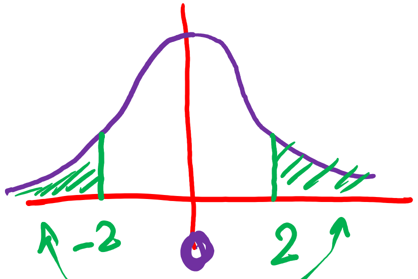
ليست حفظ

CDF $\phi(z) = \int_{-\infty}^{z} f_Z(t)\,dt$

$P(Z \le z)$

z	ϕ(z)
0.00	0.5000
0.01	0.5040
0.02	0.5080
. . .	. . .
2.99	0.9986

For the z values only



The same area from symmetry

$\boxed{\phi(0) =}$ 0.5

$\boxed{\phi(-\infty) =}$ 0

$\boxed{\phi(\infty) =}$ 1

$\phi(-ve) = 1 - \phi(+ve)$

$\boxed{\phi(-2) =}$ $P(Z > 2) = 1 - P(Z < 2) = 1 - \phi(2)$

Example

Compute

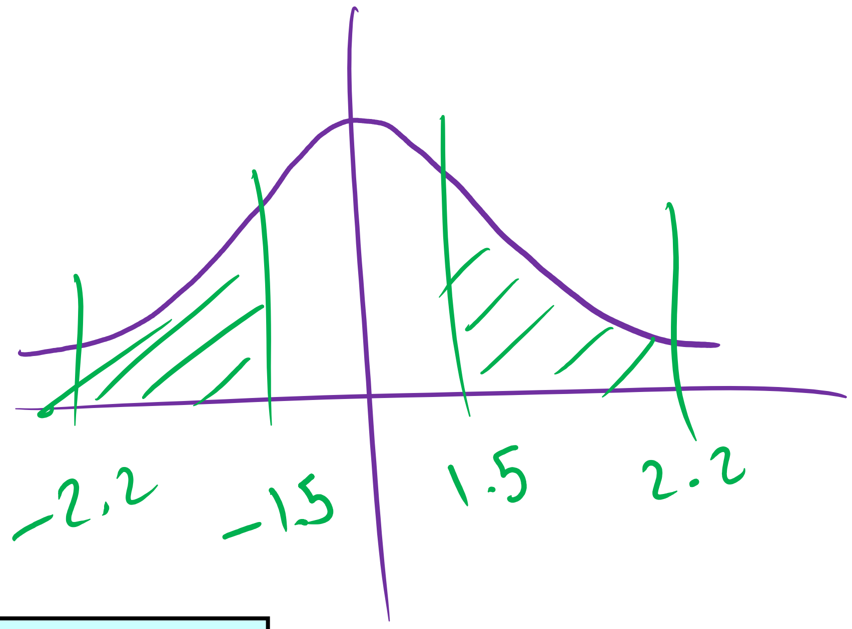
(a) $P(Z < 1) = \Phi(1)$

(b) $P(Z > 2) = 1 - P(Z \leq 2) = 1 - \Phi(2)$

(c) $P(1.5 < Z < 2.2) = \Phi(2.2) - \Phi(1.5)$

(d) $P(Z < -1) = \Phi(-1) = 1 - \Phi(1)$

(e) $P(-2.2 < Z < -1.5) = P(1.5 < Z < 2.2)$



$$X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

Example

Height of students in a certain college is assumed to follow a normal distribution with mean 170 cm and standard deviation 2 cm.

μ $\quad \sigma$ or $\sigma^2 \to 80$ $\quad \sigma$

(a) What is the probability that the height of a student chosen at random will be greater than 175 cm?

(b) In a group of 50 students, what is the probability of getting at least 2 students with height greater than 175 cm? [$\phi(2.5) = 0.99$] $\Phi(5) =$, $\Phi(\frac{1}{2})$, $\Phi(175)$

$H \sim Norm(170, \underbrace{4}_{\sigma^2})$

$P(H > 175) = 1 - \Phi(175)$ "wrong ... Not std. Normal"

std. normal

$P\left(\underbrace{\frac{H-170}{2}}_{Z} > \frac{175-170}{2}\right) = P(Z > 2.5) = 1 - \Phi(2.5) = 1 - 0.99$

$= 0.01$

Example

Height of students in a certain college is assumed to follow a normal distribution with mean 170 cm and standard deviation 2 cm.

(a) What is the probability that the height of a student chosen at random will be greater than 175 cm? 0.01

(b) In a group of 50 students, what is the probability of getting at least 2 students with height greater than 175 cm? [$\phi(2.5) = 0.99$]

X: no. of students of height > 175 $\sim Bin(50, 0.01)$

$$P_X(x) = {}^{50}C_x (0.01)^x (0.99)^{50-x}; \quad x = 0, 1, 2, \ldots, 50$$

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

$$= 1 - \left[{}^{50}C_0 (0.01)^0 (0.99)^{50-0} + {}^{50}C_1 (0.01)^1 (0.99)^{50-1}\right]$$

Example

Height of students in a certain college is assumed to follow a normal distribution with mean 170 cm and standard deviation 2 cm.

μ

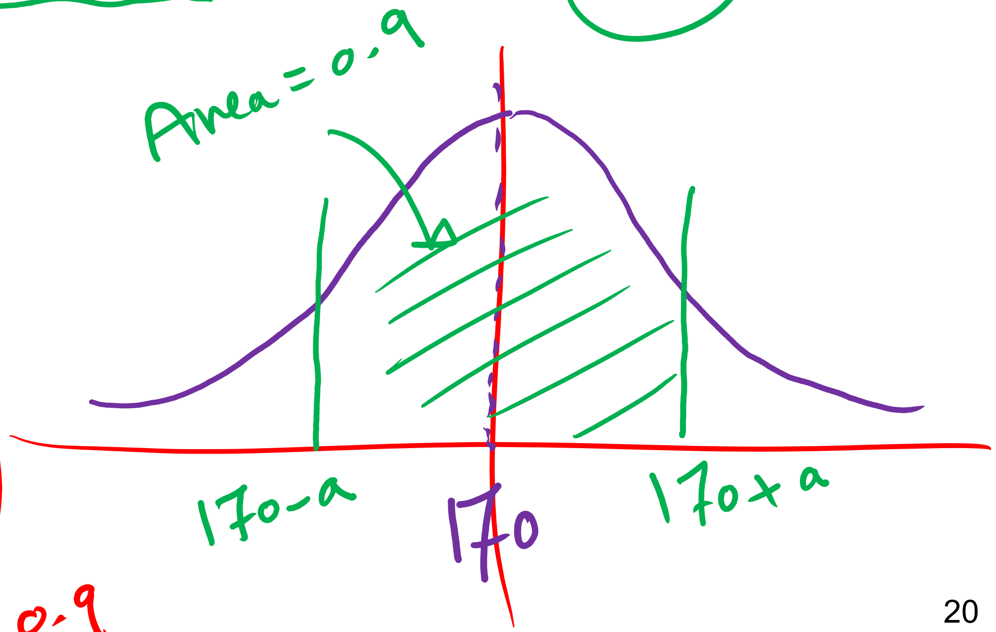
(a) What is the probability that the height of a student chosen at random will be greater than 175 cm?

(b) In a group of 50 students, what is the probability of getting at least 2 students with height greater than 175 cm?

(c) Determine the symmetric bounds around the mean that include 90% of all student lengths. $[\phi(1.645) = 0.95]$

$$P(170-a < H < 170+a) = 0.9$$

$$P\left(\frac{\cancel{170}-a-\cancel{170}}{2} < \overset{Z}{\frac{H-170}{2}} < \frac{\cancel{170}+a-\cancel{170}}{2}\right) = 0.9$$



Example

Height of students in a certain college is assumed to follow a normal distribution with mean 170 cm and standard deviation 2 cm.

(a) What is the probability that the height of a student chosen at random will be greater than 175 cm?

(b) In a group of 50 students, what is the probability of getting at least 2 students with height greater than 175 cm?

(c) Determine the symmetric bounds around the mean that include 90% of all student lengths. $[\phi(1.645) = 0.95]$

$P\left(-\frac{a}{2} < Z < \frac{a}{2}\right) = 0.9 \Rightarrow 2P\left(0 < Z < \frac{a}{2}\right) = 0.9$

$\Rightarrow P\left(0 < Z < \frac{a}{2}\right) = 0.45 \Rightarrow \Phi\left(\frac{a}{2}\right) - \underbrace{\cancel{\Phi(0)}}_{0.5} = 0.45$

$\Rightarrow \Phi\left(\frac{a}{2}\right) = 0.95 \Rightarrow \frac{a}{2} = 1.645 \Rightarrow a = 3.28$

$170 - 3.28$

$170 + 3.28$

Outline

Moment generating function

- Introduction
- Definition
- Proof
- Examples

Introduction

For a random variable X (discrete or continuous)

- The mean: E(X)
- The variance: $\mathrm{E}(\mathrm{X}^2) - \mathrm{E}(\mathrm{X})^2$

2^{nd} mom. $-$ $(1^{st}$ mom.$)^2$

- We need to calculate $\mathrm{E}(\mathrm{X}), \mathrm{E}(\mathrm{X}^2), \ldots\ldots, \mathrm{E}(\mathrm{X}^n), \ldots.$

1^{st} moment **2^{nd} moment** **n^{th} moment**

- Moment generating function (MGF): $\mathrm{M}_{\mathrm{X}}(t)$

Definition

$$E[g(X)] = \begin{cases} \text{Disc.} & \sum_{\forall x} g(x)\, P_X(x) \\ \text{Cont.} & \int_{\forall x} g(x)\, f_X(x) \end{cases}$$

Definition

$$M_X(t) = E(e^{tX})$$

Discrete

$$\sum_{\forall x} e^{tx} P_X(x)$$

Continuous

$$\int_{\forall x} e^{tx} f_X(x)\, dx$$

$$E(X) = \left[\frac{d}{dt} M_X(t)\right]_{t=0}$$

$$E(X^2) = \left[\frac{d^2}{dt^2} M_X(t)\right]_{t=0}$$

$$E(X^n) = \left[\frac{d^n}{dt^n} M_X(t)\right]_{t=0}$$

MGF for important RV's

Disc.	$Binomial(n,p)$	$(1-p+pe^t)^n$
	$Geometric(p)$	$\dfrac{pe^t}{1-(1-p)e^t}$
	$Poisson(\lambda)$	$e^{\lambda(e^t-1)}$
Cont.	$Uniform(a,b)$	$\dfrac{e^{bt}-e^{at}}{t(b-a)}$
	$Exponential(\lambda)$	$\dfrac{\lambda}{\lambda-t}$
	$N(\mu,\sigma^2)$	$e^{t\mu+\frac{1}{2}t^2\sigma^2}$

بس
حفظ

Proof

SKIP

Hint $e^{\square} = 1 + \square + \frac{\square^2}{2!} + \frac{\square^3}{3!} + \cdots$

$$E\left[e^{tX}\right] = E\left[1 + tX + \frac{t^2 X^2}{2!} + \frac{t^3 X^3}{3!} + \cdots\right]$$

$$M_X(t) = E(e^{tX}) = 1 + t\,E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots$$

$$\frac{dM_X}{dt} = E(X) + \cancel{t}\,E(X^2) + \cancel{\frac{3t^2}{3!}}E(X^3) \Big|_{t=0}$$

Example:

For a random variable with the following PMF

$$P_X(x) = \frac{1}{8}\binom{3}{x} \quad \text{for } x = 0\,,1\,,2\,,3$$

Discrete R.V.

Find the MGF and use it to get $E(X)$ and $V(X)$

$$M_X(t) = E\left[e^{tX}\right] = \sum_{\forall x} e^{tx} P_X(x)$$

0, 1, 2, 3

$3C_x$

1

1 2 1

1 3 3 1

1 4 6 4 1

$$= e^{t(0)}\frac{1}{8}\overset{1}{\cancel{{}^3C_0}} + e^{t(1)}\frac{1}{8}\overset{3}{\cancel{{}^3C_1}} + e^{t(2)}\frac{1}{8}\overset{3}{\cancel{{}^3C_2}} + e^{t(3)}\frac{1}{8}\overset{1}{\cancel{{}^3C_3}}$$

$$= \frac{1}{8} + \frac{3}{8}e^{t} + \frac{3}{8}e^{2t} + \frac{1}{8}e^{3t}$$

$$M_X(t) = \frac{1}{8} + \frac{3}{8}e^t + \frac{3}{8}e^{2t} + \frac{1}{8}e^{3t}$$

$$\frac{dM_X}{dt} = 0 + \frac{3}{8}e^t + \frac{3}{8}\cdot 2e^{2t} + \frac{1}{8}\cdot 3e^{3t}$$

$$E[X] = \left.\frac{dM_X}{dt}\right|_{t=0} = \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = \frac{3}{2} = 1.5$$

$$\frac{d^2M_X}{dt^2} = \frac{3}{8}e^t + \frac{6}{8}\cdot 2e^{2t} + \frac{3}{8}\cdot 3e^{3t} \rightarrow E(X^2) = \left.\frac{d^2M_X}{dt^2}\right|_{t=0} = \frac{3}{8} + \frac{12}{8} + \frac{9}{8} = \frac{24}{8} = 3$$

$$V(X) = E(X^2) - \mu_X^2 = 3 - (1.5)^2 = \checkmark$$

Example:

For a random variable with the following PDF $\quad \exp(1)$

$f_X(x) = e^{-x} \quad$ for $x > 0 \qquad$ Cont. RV.

$E(X) = \frac{1}{\lambda} = 1$

$V(X) = 1/\lambda^2 = 1$

Find the MGF and use it to get $E(X)$ and $V(X)$

$$M_X(t) = \int_{\forall x} e^{tx} f_X(x)\,dx = \int_0^\infty e^{tx} e^{-x}\,dx = \int_0^\infty e^{-(1-t)x}\,dx$$

عكس الديريف

$$= \frac{-1}{1-t} e^{-(1-t)x} \Big|_0^\infty = \frac{1}{1-t}(1-0)$$

$$\frac{1}{\square} \to \frac{-1}{\square^2}\,\square'$$

$$M_x(t) = \frac{1}{1-t}$$

$$\frac{dM_x}{dt} = \frac{-1}{(1-t)^2}(-1) = \frac{1}{(1-t)^2} \;;\; \frac{d^2M_x}{dt^2} = \frac{\cancel{-}1}{(1-t)^{\cancel{4}\,3}}\, 2\cancel{(1-t)}\cancel{(-1)}$$

$$= \frac{2}{(1-t)^3}$$

$t=0$ $\qquad$ $t=0$

$$E(X) = 1 \qquad\qquad E(X^2) = 2$$

$$V(X) = E(X^2) - \left(E(X)\right)^2 = 2 - (1)^2 = 1$$

Exercise:

Consider the discrete random variable X with probability mass function

x	0	1	2	3
$P(X = x)$	$1/8$	$2/8$	$2/8$	$3/8$

1. Find the moment generating function of X
2. Using the obtained formula, find $E[X]$ and $V[X]$

Exercise