Probability and Statistics

DR. AHMED TAYEL

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

ahmed.tayel@alexu.edu.eg

cont. { - Continuous RVs. (Pdf + Cdf + E(x) + V(x)) ⇐ 2 lecs.

- Special Cont. distributions (Uniform + Exponential + Normal)

cont. + discrete - Moment generating function

$V(x) = 2^{nd}$ moment $-(1^{st}$ mom.$)^2$

Statistics "method of moments"

- Introduction to Statistics 2 lecs.

Descriptive stat.

$\bar{x}$, s, histogram, ...

Inferential stat.

"Method of moments"

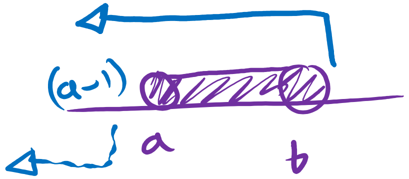
Outline

2.4 Continuous random variables

- Probability density function (PDF)
- Cumulative distribution function (CDF)
- Expectation and variance

PMF versus PDF

probability mass function



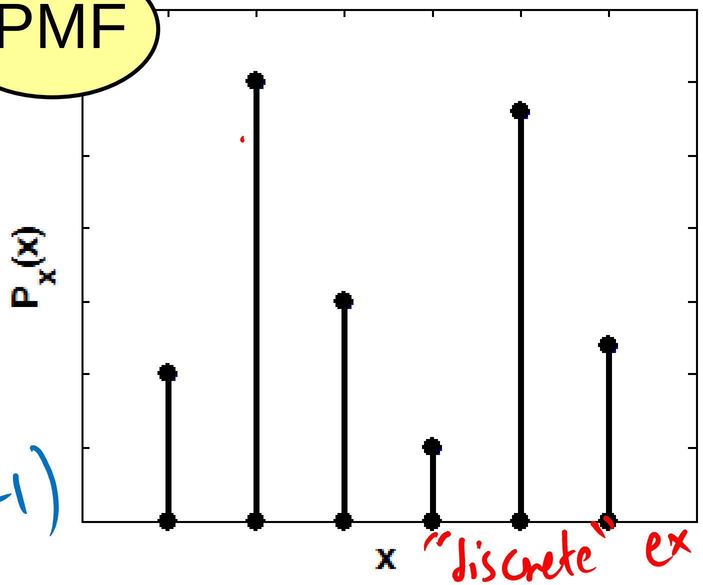
$P(a \le X \le b)$

$= F_X(b) - F_X(a-1)$

$P(a < X \le b)$

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

PMF



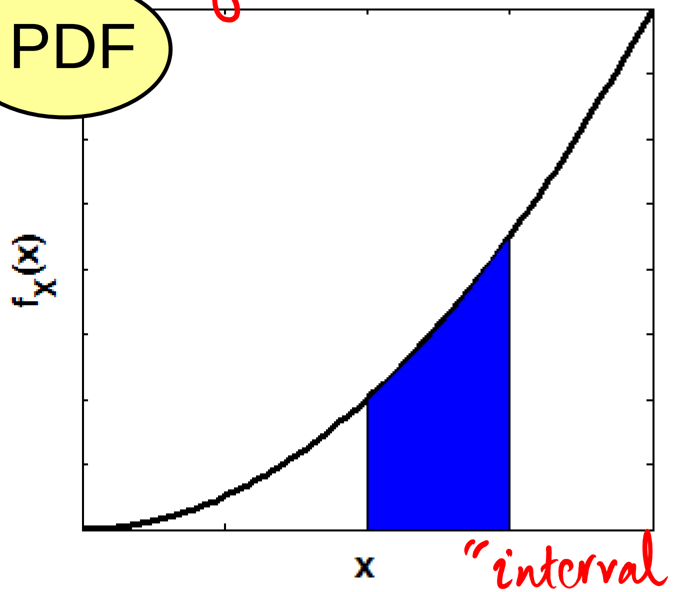
"discrete" ex $x = 1, 2, 3, \dots$

$$P(X = x) = P_X(x)$$

$$P(a \le X \le b) = \sum_{x=a}^{b} P_X(x)$$

Probability density function

PDF



"interval" $x \geq 0$

$$P(a \le X \le b) = \int_a^b f_X(x)dx$$

$$P(X = x) = 0$$

$P(a \le X \le b)$

$= P(a < X \le b)$

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

PDF properties

(1) $f_X(x) \geq 0$ for all x

(2) $\int_{-\infty}^{\infty} f_X(x)dx = 1$

Validation of pdf

Example

Let $\underbrace{\boxed{f_X(x) = 2x}}_{\text{pdf}}$, $\underbrace{0 \leq x \leq 1}$. Show that $f_X(x)$ is a valid PDF

$$\equiv \begin{cases} 0 & ; x < 0 \\ 2x & ; 0 \leq x \leq 1 \\ 0 & ; x > 1 \end{cases}$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 2x\,dx = x^2\Big|_0^1 = 1 - 0 = 1 \ \#$$

max x, min of x

Example

Let $f_X(x) = \cancel{k}\, e^{-3x}$, $\boxed{x \geq 0}$ "interval" $\equiv$ Cont. RV.

3

(a) Find k.

(b) Compute $P(1 \leq X < 2)$, $P(X \geq 3)$

$P(-3 \leq X < 1)$
$= P(0 \leq X < 1)$
$= \int_0^1 3e^{-3x}\,dx$

عكس ترتيب التعويض

(a) $\int_{0\ (\text{min})}^{\infty\ (\text{max})} k\, e^{-3x}\,dx = k \int_0^{\infty} e^{-3x}\,dx = k\, \frac{-1}{3} e^{-3x} \Big|_0^{\infty}$

$= \frac{k}{3}\left[1 - 0\right] = \boxed{\frac{k}{3} = 1} \Rightarrow k = 3$

(b) $P(1 \leq X < 2) = \int_1^2 3e^{-3x}\,dx = \cancel{3} \cdot \frac{-1}{\cancel{3}} e^{-3x}\Big|_1^2 = e^{-3(1)} - e^{-3(2)} = \checkmark$

$P(X \geq 3) = \int_3^{\infty} 3e^{-3x}\,dx = \cancel{3} \cdot \frac{-1}{\cancel{3}} e^{-3x}\Big|_3^{\infty} = e^{-3(3)} - 0 = \checkmark$

Example

The duration of a telephone call (in minutes) is assumed to have the PDF

$$f_X(x) = \cancel{k}^{\,3/32}\,(4x - x^2), \qquad 0 \le x \le 4.$$

(a) What is the probability that a call duration exceeds 3 minutes?

(b) If consecutive calls are independent, what is the probability that the duration of the next 10 calls exceed 3 minutes?

$$k = \frac{3}{32}$$

max, min

$$\int_0^4 k(4x - x^2)\,dx = k\left(2x^2 - \frac{1}{3}x^3\right)\Big|_0^4 = k\Big[\underbrace{32 - \frac{4^3}{3}}_{\frac{32}{3}} - 0\Big] = 1$$

$$\text{(a)}\quad P(X>3) = \int_3^{4\,\cancel{\infty}} \frac{3}{32}(4x - x^2)\,dx = \frac{3}{32}\left(2x^2 - \frac{1}{3}x^3\right)\Big|_3^4 = q$$

using calculator

$$\text{(b)}\quad P(1^{st} \cap 2^{nd} \cap \cdots \cap 10^{th}) = P(1^{st})P(2^{nd}) \cdots P(10^{th}) = q^{10}$$

indep.

Cumulative distribution function (CDF)

general pt.

Definition $F_X(x) = P(X \le x)$

Example

Find the CDF of the RV having a PDF $f_X(x) = 2x$, $0 \le x \le 1$.

pdf

$$\text{CDF: } F_X(x) \overset{\text{general}}{=} \int_{\text{min } 0}^{x} 2x \, dx$$

$$= x^2 \Big|_0^x = x^2$$

$$\frac{d}{dx}(x^2) = 2x \quad \text{"pdf"}$$

General properties

(1) Non-decreasing function

(2) $F(-\infty) = 0$

(3) $F(\infty) = 1$

CDF general pdf

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$$

min

pdf cdf

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Example

Find the CDF of the RV X where $f_X(x) = 3\,e^{-3x}$, $x \geq 0$.

Hence, compute

(a) $P(X \leq 2)$

(b) $P(X \geq 5)$

(c) $P(1 \leq X \leq 2)$

in discrete $P(X \geq 5) = 1 - P(X < 5) = 1 - F_X(4)$ assume step 1

$P(1 \leq X \leq 2) = F_X(2) - F_X(0)$

general

$$\text{CDF: } F_X(x) = \int_{0}^{x} 3e^{-3x}\,dx = \cancel{3}\cdot\frac{-1}{\cancel{3}}\,e^{-3x}\Big|_0^x = 1 - e^{-3x}$$

min

$$P(X \leq x) = 1 - e^{-3x}$$

$$P(X \leq 2) = F_X(2) = 1 - e^{-3(2)} = \checkmark$$

$$P(X \geq 5) = 1 - P(X < 5) = 1 - F_X(5) = 1 - \left[1 - e^{-3(5)}\right] = \checkmark$$

$$P(1 \leq X \leq 2) = F_X(2) - F_X(1) = \left[1 - e^{-3(2)}\right] - \left[1 - e^{-3(1)}\right]$$

Example

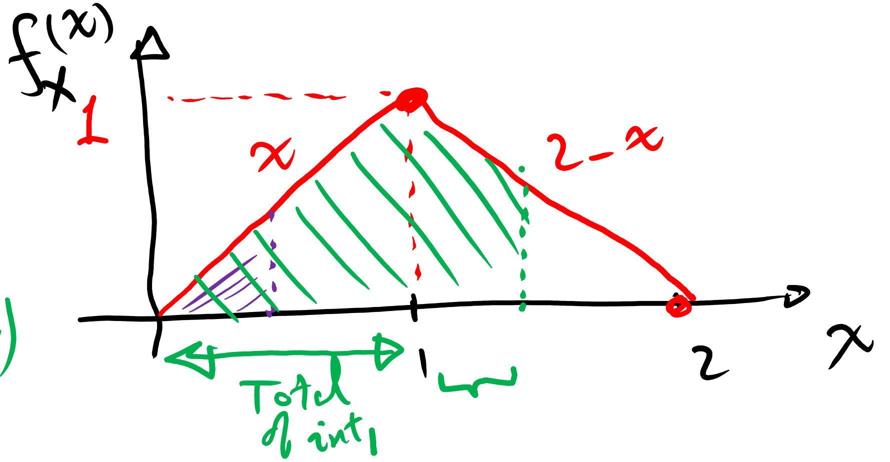
The total number of hours, measured in units of 100 hours, which a certain machine is running over a period of one year is a continuous random variable X that has the density function

$$f_X(x) = \begin{cases} x & ,0 < x < 1 \\ 2 - x & ,1 \le x < 2 \end{cases}$$

Find the associated CDF.

$$F_X(x) = \begin{cases} 0 & ; \ x<0 \\ \frac{1}{2}x^2 & ; \ 0<x<1 \\ 2x-\frac{1}{2}x^2-1 & ; \ 1\le x<2 \\ 1 & ; \ x\ge 2 \end{cases}$$

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



$0<x<1$

$$F_X(x) = \int_0^x x\,dx = \frac{1}{2}x^2\Big|_0^x = \frac{1}{2}x^2$$

Total of int1 $= F_X(1) = \frac{1}{2}(1)^2 = \frac{1}{2}$

$1\le x<2$

$$F_X(x) = \frac{1}{2} + \int_1^x (2-x)\,dx$$

$$= \frac{1}{2} + \left(2x-\frac{1}{2}x^2\right)\Big|_1^x$$

$$= \frac{1}{2} + \left[2x-\frac{1}{2}x^2-\left(2-\frac{1}{2}\right)\right]$$

$$= 2x-\frac{1}{2}x^2-1$$

Example

The target diameter of a hole drilled in a metal sheet is 12.5 mm. Most random disturbances result in larger diameters. Historical data show that the distribution of the diameter of the hole can be modelled by

$$f_X(x) = \begin{cases} 20\, e^{-20(x-12.5)} & , x \geq 12.5 \\ 0 & , \text{o.w} \end{cases}$$

Target

Error = | Current Value − Target |

| X − 12.5 |

(a) If a part with diameter exceeding 12.6 mm is scrapped, what proportion of parts is scrapped?

(b) Find the probability that the error in the diameter does not exceed 0.1 mm?

CDF: $F_X(x) = \int_{12.5}^{x} 20\, e^{-20(x-12.5)}\, dx = \cancel{20} \cdot \frac{-1}{\cancel{20}} e^{-20(x-12.5)} \Big|_{12.5}^{x}$

$= 1 - e^{-20(x-12.5)}$

(a) $P(X > 12.6) = 1 - P(X \leq 12.6) = 1 - F_X(12.6) = 1 - \left[1 - e^{-20(12.6-12.5)}\right] =$

(b) $P(\underbrace{|X-12.5|}_{\text{error}} < 0.1) = P(-0.1 < X - 12.5 < 0.1) = P(12.4 < X < 12.6)$

$= P(12.5 < X < 12.6) = F_X(12.6) = 1 - e^{-20(12.6-12.5)}$

$|x| < 1$

$\equiv \quad -1 < x < 1$

$|x| > 1$

$\equiv \quad x < -1$ or $x > 1$

$\equiv$ Complement of $|x| < 1$

Expectation

All values of x

Definition

max x

$$E(X) = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$

min x

General properties

(1) $E[a X + b] = a E(X) + b$

(2) $E\left[g(X)\right] = \int_{-\infty}^{\infty} g(x) \, f_X(x) \, dx$

Example

The duration T of a mobile call is a RV having a pdf

$$\boxed{f_T(t) = 0.5\, e^{-0.5t}}, \quad \boxed{t \geq 0.}$$

(a) What is the expected call duration?

(b) Plan A: Charge 0.25 EP per minute.

Plan B: Charge 0.15 EP per minute + fixed cost (0.15 EP).

Which plan has a lower average cost?

d	I
t	$e^{-\frac{1}{2}t}$
1	$-2e^{-\frac{1}{2}t}$
0	$4e^{-\frac{1}{2}t}$

$+$, $-$

(a) $$E(T) = \int_{0 \,(\text{min})}^{\infty \,(\text{max})} \underbrace{t}_{\text{All values}} \cdot \underbrace{0.5\, e^{-0.5t}}_{\text{pdf}}\, dt = \frac{1}{2}\int_0^\infty t\, e^{-\frac{1}{2}t}\, dt$$

[illegible]

$$= \frac{1}{2}\left[-2te^{-\frac{1}{2}t} - 4e^{-\frac{1}{2}t}\right]_0^\infty = \frac{1}{2}\left[4 - (0)\right] = 2$$

$$\lim_{t\to\infty} \frac{t}{e^{1/2 t}} = \frac{\infty}{\infty}$$

l'hopital $$\lim_{t\to\infty} \frac{1}{\frac{1}{2}e^{1/2 t}} = 0$$

By inspection $$\lim_{t\to\infty} t\, e^{-\frac{1}{2}t} = 0$$

Example

$E(aX+b) = aE(X) + b$

The duration T of a mobile call is a RV having a pdf

$$f_T(t) = 0.5\, e^{-0.5t}, \quad t \geq 0.$$

(a) What is the expected call duration? $E(T) = 2$ mins.

(b) Plan A: Charge 0.25 EP per minute.

Plan B: Charge 0.15 EP per minute + fixed cost (0.15 EP).

Which plan has a lower average cost?

Statistics

open call log.

1.4, 2.5, 3, 0.5,

method of moment

Expected cost of each plan

$E(A) = E[0.25\,T] = 0.25\,E(T) = 0.25 * 2 = 0.5$ EP

$E(B) = E[0.15 + 0.15\,T] = \underbrace{0.15 + 0.15\,\underbrace{E(T)}_{2}}_{0.3} = 0.45$ EP

plan B is better

Example

CDF

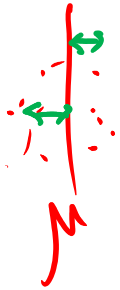
Given $\underbrace{F_X(x)} = \begin{cases} 0 & , x < 0 \\ 1 - e^{-0.01x} & , x \geq 0 \end{cases}$, find E(X)

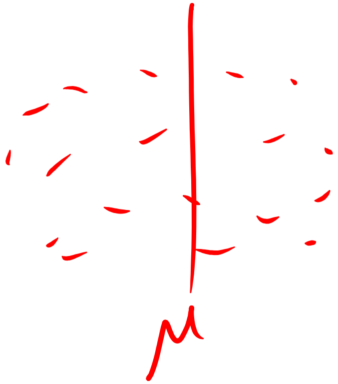
Pdf : $f_X(x) = \frac{d}{dx}\left(1 - e^{-0.01x}\right) = 0.01 e^{-0.01x}$

$E(X) = \int_0^\infty x \; 0.01 e^{-0.01x} \, dx$

$d \qquad I$

Variance and standard deviation





measure of dispersion

Variance

$$\mathrm{Var}(X) = E(X - \mu_X)^2$$

$$\int_{-\infty}^{\infty} (x-\mu_x)^2 f_X(x)\, dx$$

Computational formula

$$\mathrm{Var}(X) = E(X^2) - (\mu_X)^2$$

2nd mom. − (1st mom.)^2

General property

$$\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$$

Standard deviation

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

has the same units of X, μ_X

Example

The probability density function of the random variable X is given by

$V(x) = E(x^2) - (E(x))^2$

$$f_X(x) = \begin{cases} x & ,0 \le x \le 1 \\ 2 - x & ,1 \le x \le 2 \\ 0 & ,\text{o.w.} \end{cases}$$

(a) Find the expectation and variance of X?

(b) Find P(|X-μ|<σ)?

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_0^1 \underbrace{x \cdot x}_{x^2}\, dx + \int_1^2 \underbrace{x \cdot (2-x)}_{2x-x^2}\, dx$$

$$= \frac{1}{3}x^3\Big|_0^1 + \left(x^2 - \frac{1}{3}x^3\right)\Big|_1^2 = \frac{1}{3} + \left[4 - \frac{8}{3} - \left(1 - \frac{1}{3}\right)\right] = 1$$

$$E(X^2) = \int_0^1 \underbrace{x^2 \cdot x}_{x^3}\, dx + \int_1^2 \underbrace{x^2 \cdot (2-x)}_{2x^2-x^3}\, dx = \frac{1}{4}x^4\Big|_0^1 + \left(\frac{2}{3}x^3 - \frac{1}{4}x^4\right)\Big|_1^2 = \frac{7}{6}$$

$$V(X) = E(X^2) - (E(X))^2 = \frac{7}{6} - (1)^2 = \frac{1}{6}; \quad \sigma_X = \frac{1}{\sqrt{6}}$$

or using calculator

Example

The probability density function of the random variable X is given by

$$f_X(x) = \begin{cases} x & , 0 \le x \le 1 \\ 2 - x & , 1 \le x \le 2 \\ 0 & , \text{o.w.} \end{cases}$$

(a) Find the expectation and variance of X? $\rightarrow 1 = \mu \rightarrow \frac{1}{6}, \sigma = \frac{1}{\sqrt{6}}$

(b) Find P(|X-μ|<σ)?

$$P(|X-\mu| < \sigma) = P(-\sigma < X - \mu < \sigma)$$

$$= P(\mu - \sigma < X < \mu + \sigma) = P(0.6 < X < 1.4)$$

$$= \int_{0.6}^{1} x \, dx + \int_{1}^{1.4} 2 - x \, dx = \checkmark$$

Example

The probability density function of the random variable X is given by

$$f_X(x) = \begin{cases} x & , 0 \leq x \leq 1 \\ 2 - x & , 1 \leq x \leq 2 \\ 0 & , \text{o. w.} \end{cases}$$

(a) Find the expectation and variance of X?

(b) Find P(|X-μ|<σ)?

$$F_X(x) = \begin{cases} 0 & ; \quad x < 0 \\ \frac{1}{2} x^2 & ; \quad 0 < x < 1 \\ 2x - \frac{1}{2}x^2 - 1 & ; \quad 1 \leq x < 2 \\ 1 & ; \quad x \geq 2 \end{cases}$$

Given the CDF "From slide 10"

$$P(0.6 < X < 1.4) = F_X(1.4) - F_X(0.6)$$

$$= \left[2(1.4) - \frac{1}{2}(1.4)^2 - 1\right] - \left[\frac{1}{2}(0.6)^2\right]$$