# Probability

## Discrete Random Variable [ Mean and Variance ]

Prof. Dr. Mostafa Elhosseini

https://youtube.com/drmelhosseini

## Sample Mean

defects

- The number of flaws in a 1-inch length of copper wire manufactured by a certain process varies from wire to wire. Overall, **48%** of the wires produced have **no flaws**, **39%** have **one flaw**, **12%** have **two flaws**, and **1%** have **three flaws**. Let $X$ be the **number of flaws** in a randomly selected piece of wire. Now **imagine** that we have a **sample of 100 wires**, and that the sample follows this distribution perfectly
- What is the **sample mean**?

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | 0.48 | 0.39 | 0.12 | 0.01 |
| # | 48 | 39 | 12 | 1 |

$\downarrow$ 100

$$\bar{X} = \frac{48(0) + 39(1) + 12(2) + 1(3)}{100}$$

$$= 0.48(0) + 0.39(1) + 0.12(2) + 0.01(3)$$

$$\text{Sample mean} = P(X=x)\,x$$

## Population Mean

- Let $X$ be a discrete random variable with probability mass function $p(x) = P(X = x)$.

**The mean of $X$ is** given by $\boldsymbol{\mu_X = \sum_x xP(X = x)}$

Expected value $\Rightarrow$ $E(x)$

Expectation $\Rightarrow$

population mean $\Rightarrow$ $\mu_x$ $\Rightarrow$ $\mu$

mean

$$\mu_x = \sum_{x} x P(X=x)$$

$\Downarrow$

Expected Value

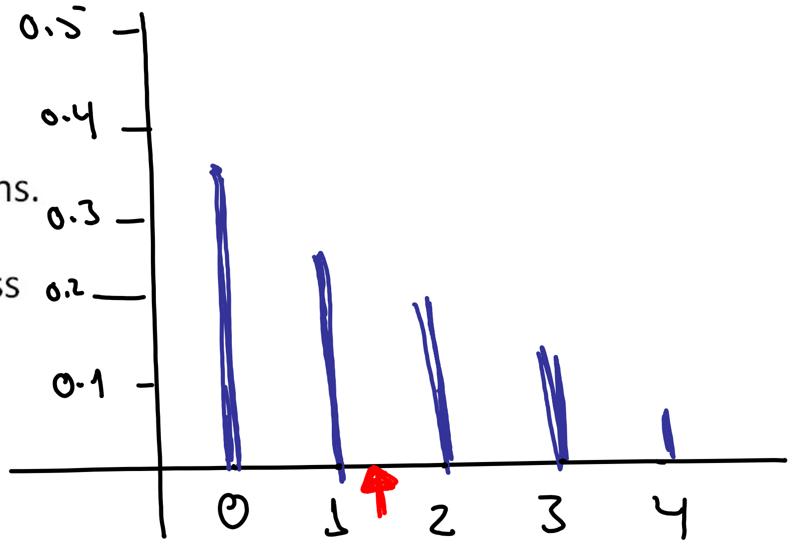
# Example - 01

اعادة معايرة

- A certain industrial process is brought down for recalibration whenever the quality of the items produced falls below specifications. Let $X$ **represent the number of times the process is recalibrated during a week and** assume that $X$ has the following probability mass function.

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p(x)$ | 0.35 | 0.25 | 0.20 | 0.15 | 0.05 |

Find the mean of $X$.




$$\mu_X \Rightarrow E(X) = \sum_x x\,P(X=x)$$

$$= 0*0.35 + 1*0.25 + 2*0.2 + 3*0.15 + 4*0.05$$

$$\mu_X = \boxed{1.3}$$

# Recall: Sample variance

- Recall that for a sample $X_1, \ldots, X_n$, the **sample variance** is given by

$$\sum (X_i - \bar{X})^2/(n-1)$$

$\mu_X$

$(x - \mu_X)^2$

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 \, P(X = x)$$

## Population Variance

- By analogy, the **population variance of a discrete random** variable $X$ is a weighted average of the squared differences $(x - \mu_X)^2$ where $x$ ranges through all the possible values of the random variable X. This weighted average is computed by **multiplying each squared difference $(x - \mu_X)^2$ by the probability $P(X = x)$ and summing the results**

# Population Variance

$$\boldsymbol{\sigma_X^2} = \sum_{x} (x - \mu_X)^2 P(X \;=\; x)$$

$$\boldsymbol{\sigma_X^2} = \sum_{x} x^2 P(X \;=\; x)\, \mu_X^2$$

$\mu_X = 1.3$

## Example - 02

- Find the **variance and standard deviation** for the random variable X described in Example - 01, representing the number of times a process is recalibrated

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p(x)$ | 0.35 | 0.25 | 0.20 | 0.15 | 0.05 |

$$\sigma_X^2 = (0-1.3)^2(0.35) + (1-1.3)^2(0.25) + (2-1.3)^2(0.2) + (3-1.3)^2(0.15) + (4-1.3)^2(0.05)$$

$$= 1.51$$

$$\sigma_X = \sqrt{1.51} = 1.23$$

$$\sigma_X^2 = 0^2(0.35) + 1^2(0.25) + 2^2(0.2) + 3^2(0.15) + 4^2(0.05) - (1.3)^2$$

$$= 1.51$$

# References


- Statistics for Engineers and Scientists, Fourth Edition, William Navidi
  Colorado School of Mines
- Khan Academy