Task #09 Bytewise Fellowship

Introduction to Statistics.

Mean: Mean is arithematic average of all data points.

Formula: $\bar{X} = \frac{\Sigma x}{n}$

Example:

Data Set $x = \{2, 4, 6, 8, 10\}$

Sum of all points $\Sigma x = 30$

Total number of data points $= 5$

$$\text{Mean} = \frac{30}{5}$$

$$= 6$$

Median: Median is middle most value when data is arranged in either ascending or descending order.

Formula: If n is Even

$$= \frac{\left(\left(\frac{n}{2}\right)^{th}\text{term} + \left(\frac{n}{2}+1\right)^{th}\text{term}\right)}{2}$$

If n is odd

$$= \frac{(n+1)}{2}$$

Example: Data Set $x = \{2, 4, 6, 8, 10, 12\}$

$$\text{Median} = \frac{(6+8)}{2} = 7$$

Mode : Mode is the value with highest frequency in data set.

Example : Data set = {2, 6, 8, 6, 3, 7, 6}

Mode = 6

---

1. **Binomial Probability Distribution:** When we have data that is quantitative discrete, we use binomial probability distribution.

Properties of binomial distribution.

1. Fixed number of trials
2. For each trial there's only two possible outcomes i.e success or failure.
3. Probability of success remains same for each trial
4. Successive trials are independent.

Formula: $P(X=x) = \binom{n}{x} p^x q^{n-x}$

Where n = number of trials (fixed)
x = Binomial Random Variable
p = probability of success in each trial
q = " " failure " "

Important note: ① $0 < x < p$
② $p + q = 1$
③ n & p are two parameters of binomial probability distribution.

Example: A multiple choice Quiz has 5 Question each with 4 options. If a student guesses on each question, what is the probability that student answers exactly 3 questions correctly?

Solution:

$n = 5$
$x = 3$
p = prob of success in each trial $= \frac{1}{4}$
$= 0.25$
$q = 1 - p = 0.75$

$$P(x=3) = \binom{5}{3} (0.25)^3 (0.75)^{5-3}$$

$$= \frac{5!}{2!\,3!} (0.015)(0.56)$$

$$= \frac{5 \times 4 \times \cancel{3!}}{2!\ \cancel{3!}} (0.008) = 0.087$$

So the prob that student answers exactly 3 questions correctly is 8.7%

2) Normal Distribution:

When we have quantitative continuous data, we use normal distribution.

Prob Density function

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

Note: μ & σ are two parameters of Normal Prob distribution.

- For computational ease we convert it into Z-Score.

$$Z = \frac{x-\mu}{\sigma} \qquad (-\infty < Z < +\infty)$$

Important properties of Z Score:

1. Mean is always 0; $\mu_z = 0$
2. Standard deviation = 1; $\sigma_z = 1$

Example.

Average temperature for a certain city in summers is 85°F with standard deviation 5°F. What is the probability that randomly chosen day in July will have average temperature greater than 90°F.

Solution

$\mu = 85°F$

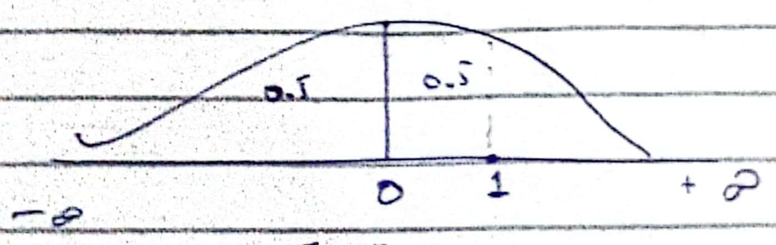
$\sigma = 5°F$, $x = 90°F$

$$Z = \frac{x - \mu}{\sigma}, \Rightarrow \frac{90-85}{5} = 1$$

$P(x > 90) = ?$

$P(x > 90) = P(Z > 1)$.

- We will be finding the prob $P(x>90)$ from Z score Area table.
- We are aware that total area under the curve is 1.



Z scores

From the above Curve:

$$P(Z>1) = \text{Area}(0 \to +\infty) - \text{Area}(0 \to +1)$$
$$\text{Area}(+\infty) - \text{Area}(+1)$$
$$= 0.5 - \text{Area}(1.00)$$

As Total area from $+\infty \to -\infty$ under the Curve is 1, Therefore half of it will be 0.5. (from $0 \to +\infty$).

- From the Z Score Table, the probability is 0.3413 When Z Score is 1.

Therefore

$$P(X>90) = P(Z>1)$$
$$= 0.5 - 0.3413$$
$$= 0.1587$$

Hence the probability that we will have ~~98~~ Temp greater Than 90F On any random day in July is 0.15.

3 Uniform Distribution

When all Outcomes are Equally likely to happen Over a

Specific range, we use uniform prob distribution.

Probability Density function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{Otherwise} \end{cases}$$

Example:

Suppose a random number is chosen uniformly along a line segment from 0 to 10 m

$$f(x) = \frac{1}{10-0} = \frac{1}{10}$$

→ Each random number has Equal Probability of Occurance in Outcome while Selecting.

---

(4) Poisson Distribution

When we have to Predict the Probability of given number of Events Occuring in a fixed interval of time, we use Poisson distribution

Probability Density - function

$$P(X = n) = \frac{e^{-\lambda} \lambda^{n}}{n!}$$

n = Total # of Events Occur in an interval.

λ = Average # of events per interval.

Example

Consider in a Store, 5 Customers per hour arrives on average. What is the probability that Exactly 3 Customers will arrive in next hour?

$$P(X = 3) = \frac{e^{-5} \cdot 5^{3}}{3!}$$

$$= 0.14$$

Prob that Exactly 3 Customers will arrive in next hour at the Store is approx 14 %.