VCEMATHSMETHODS.COM

MATHS METHODS

PODCAST

Discrete probability distributions

- Probability distributions
- Discrete random variables
- Expected values (mean)
- Variance
- Standard deviation
- Linear functions - mean & standard deviation

Probability distributions

- A **probability distribution** is the set of probabilities related to outcomes.
- For example, a driver passes through five sets of traffic lights every day.
- The probability of stopping at a number of red lights (X) might look something like this probability distribution.

x	0	1	2	3	4	5
Pr (X=x)	6%	11%	18%	25%	21%	p

- Pr(X=5) must be p = 100% - 81% =19%.

Discrete random variables

- The number of traffic light stops is an example of a **discrete random variable**.
- Only a finite set of numbers appears in the event space (the set of outcomes).
- $x=0, x=1, x=2, x=3, x=4, x=5$
- The sum of all probabilities is equal to 1.

$6\% + 11\% + 18\% + 25\% + 21\% + 19\% = 100\%$

Expected values (mean)

- A discrete random variable has an **expected (mean) value**.
- The mean is found by multiplying all values of X by the probabilities Pr(X=x) & adding these together.

Expected value (mean) → $$\mu = E(X) = \sum_{x=1}^{n} x\Pr(X = x)$$ ← Values of x multiplied by probabilities.

↑ Sum for all values of x (starting at 1)

- For the traffic light example, the expected number of red lights is:

$$= (0 \times 6\%) + (1 \times 11\%) + (2 \times 18\%) + (3 \times 25\%) + (4 \times 21\%) + (5 \times 19\%)$$

$$= 301\% \quad = 3.01$$

- While it is impossible to actually get 3.01 red lights as a result, this would be the longer term average.

Variance

- A discrete random variable has a **variance**, a measure of spread of data.
- The variance is found by finding differences between X values & the mean.

Variance ⟶ $$Var(X)=\sum_{x=1}^{n}(x-\mu)^2 Pr(X=x) = E(x-\mu)^2 = E(x^2)-E(X)^2$$

Sum for all values of x (starting at 1)

Expected value of differences from the mean.

- Find differences between x and the mean.
- Square these differences.
- Multiply these squares by the probability of associated x values.
- Find the sum of the squares.

Calculating variance

- Variance can be calculated by finding the expected value of the difference between the values of x & the mean.

$$Var(X)=\sum_{x=1}^{n}(x-\mu)^2 Pr(X=x) \ = E(x-\mu)^2$$

x	Pr (X=x)	x.Pr (X=x)
0	6%	0
1	11%	0.11
2	18%	0.36
3	25%	0.75
4	21%	0.84
5	19%	0.95
	E(x)=μ	3.01

(x-μ)	$(x-\mu)^2$.Pr (X=x)
-3.01	0.54
-2.01	0.44
-1.01	0.18
-0.01	0.00
0.99	0.21
1.99	0.75
Var (x)	2.13

Calculating variance

- Variance can also be calculated by finding the difference between the expected value of the squares & the square of the expected value of x.

$$Var(X) = \sum_{i=1}^{n} (x-\mu)^2 Pr(X=x) = E(X^2) - E(X)^2$$

x	Pr (X=x)	x^2.Pr (X=x)
0	6%	0
1	11%	0.11
2	18%	0.72
3	25%	2.25
4	21%	3.36
5	19%	4.75
	$E(X^2)$	11.19

$E(X^2)$	11.19
$E(X)^2$	9.06
Var(X) = $E(X^2) - E(X)^2$	2.13

Linear functions - mean & standard deviation

- If a linear function is applied to the values of X, then expected (mean) value and variation will be affected.

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

(The expected value is changed according to the rule - multiplying & adding.)

$$Var(aX+b)=a^2Var(X)$$

(Adding an amount to all values does not change the variance.)

$$E(X+Y)=E(X)+E(Y)$$

(Expected value of a sum or difference is equal to the sum or difference of expected means.)

Standard deviation

- The **standard deviation**, is the more commonly used measure of spread of data.
- It is the measure of average difference of values from the mean.
- The standard deviation is the square root of the variance.
- For any variable that follows normal distribution, there will always be the proportions within the same variation from the mean.

Standard deviation → $\sigma = SD(X) = \sqrt{Var(X)}$

↑ Variance of X

$$SD(aX + b) = a\sigma$$

Standard deviation

- For the example of the traffic lights:

$$\sigma = \sqrt{2.13} \approx 1.5$$

- For normally distributed data, 68% of values fall within one standard deviation of the mean.
- For normally distributed data, 95% of values fall within two standard deviation of the mean.
- One standard deviation: While 1.5 or 4.5 red lights is impossible, the closest approximations are 1 and 5.
- This technique of modelling works better when there are more values of X in the distribution.

VCEMATHSMETHODS.COM

MATHS METHODS

PODCAST