# STATISTICAL ANALYSIS

## MEAN AND VARIANCE FOR DISCRETE RANDOM VARIABLES (II)

Contents include: Calculating mean, variance and standard deviation for discrete random variables

www.mathifyhsc.com.au

- Mean of a Discrete Distribution

Another name for the mean of a discrete distribution is the "**Expected Value**", with the symbols being **either $\boldsymbol{\mu}$ or $\boldsymbol{E(x)}$** and used interchangeably.

For the discrete random variable X, let $p(x) = P(X = x)$. The formula for the expected value is given by:

$$\mu = E(x) = \sum x\, p(x)$$

Where:

$x \text{ is the value of } X$
$p(x) \text{ is the probability of } x \text{ occuring}$
$\text{The } \Sigma \text{ symbol denotes that we are summing over all values of the distribution}$

**Example 1:** Alex conducts an experiment in school where he throws four normal coins and records the number of heads he observes. Calculate the expected number of heads that he will see.

Solution:

$$\boldsymbol{Step\ 1: Identify\ the\ discrete\ random\ variable\ in\ the\ question}$$

In this case, let $X$ equal the number of heads recorded, meaning that this is the discrete random variable.

$$\boldsymbol{Step\ 2: Draw\ a\ table\ for\ all\ the\ values\ that\ X\ can\ take}$$

Now the possible observations that can be made by Alex are:

$$P(X = 0), P(X = 1), P(X = 2), P(X = 3), P(X = 4)$$

Therefore, the table including the probability $p(x)$ of each case occurring can be drawn:

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p(x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

$$\boldsymbol{Step\ 3: Using\ our\ formula, calculate\ E(x)}$$

Using the formula $E(x) = \mu = \sum x\, p(x)$:

$$\begin{aligned} E(x) = \mu &= 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16} \\ &= 0 + \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16} \\ &= 2 \end{aligned}$$

$\therefore$ The expected number of heads would be 2

- Variance and standard deviation of a Discrete Distribution

The variance and standard deviation are essentially measures of how spread out the data is, each respectively being represented by $Var(x)$ and $\sigma$.

For the discrete random variable X, let $p(x) = P(X = x)$. The formula for the variance may be given by:

$$Var(x) = E\big((X - \mu^2)\big) = \sum (x - \mu)^2 p(x)$$

An alternative form that is easier to use is:

$$Var(x) = E(X^2) - \mu^2 = \sum x^2 p(x) - \mu^2$$

Where:

$\mu \text{ is the mean}$
$p(x) \text{ is the probability of } x \text{ occurring}$

The standard deviation is found by rooting the variance:

$$\sigma = \sqrt{Var(x)}$$

Where:

$\sigma$ is the standard deviation

$Var(x)$ is the variance

**Example 2:** Taking the same question as example 1, calculate variance and the standard deviation:

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $p(x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

And it is known that $E(x) = \mu = 2$

Solution:

$$Step\ 1: Add\ an\ extra\ row\ in\ the\ table\ for\ x^2$$

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $x^2$ | 0 | 1 | 4 | 9 | 16 |
| $p(x)$ | $\frac{1}{16}$ | $\frac{4}{16}$ | $\frac{6}{16}$ | $\frac{4}{16}$ | $\frac{1}{16}$ |

$$Step\ 2: Calculate\ Var(x)$$

We will be using the formula: $Var(x) = E(X^2) - \mu^2 = \sum x^2 p(x) - \mu^2$

$$\sum x^2 p(x) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 4 \times \frac{6}{16} + 9 \times \frac{4}{16} + 16 \times \frac{1}{16}$$
$$= 5$$

$\mu = 2\ (from\ Example\ 1)$
$so\ \mu^2 = 4$

$$\therefore \sum x^2 p(x) - \mu^2 = 5 - 4$$
$$= 1$$

$$\therefore Variance\ is\ 1$$

$$Step\ 3: Calculate\ standard\ deviation$$

Since $Var(x) = 1$,

$$\sigma = \sqrt{1} = 1$$

$$\therefore standard\ deviation\ is\ 1$$

## Discrete Probability Exercises

1. The discrete random variable X has this probability distribution:

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.1 | 0.2 | 0.4 | 0.2 | 0.1 |

a) Find $P(1 < X \leq 3)$
b) Find the expected value of X, showing all working
c) Find the standard deviation of X, showing all working

2. Consider the data from 100 trials of a weighted spinner which has five outcomes: 1, 2, 3, 4, 5 that each have a separate probability of occurring:

| $Score\ x$ | 1 | 2 | 3 | 4 | 5 | $Total$ |
|---|---|---|---|---|---|---|
| $Frequency$ | 10 | 20 | 45 | 15 | 10 | 100 |
| $Relative\ Frequency$ | $\frac{1}{10}$ | $\frac{2}{10}$ | $\frac{9}{20}$ | $\frac{3}{20}$ | $x$ | 1 |

i) What is the value of $x$?
ii) Calculate the expected value
iii) Calculate the variance and standard deviation

3. A discrete random variable X has the following probability distribution:

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $2q$ | $6q$ | $3q$ | $4q$ |

Calculate the mean of $X$

4. The probability distribution of a discrete random variable X is given by the table below:

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.2 | $0.6p^2$ | 0.1 | $1 - p$ | 0.1 |

a) Show that $p = \frac{2}{3}$ or $p = 1$

b) Let $p = \frac{2}{3}$. Calculate, to two decimal places:

   a. $P(X \geq E(X))$
   b. $E(X)$
   c. The variance of X

5. The discrete random variable X has the following probability distribution:

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X = x)$ | $a$ | $b$ | 0.3 |

Given that $E(X) = 0.8$, then find the values of $a$ and $b$

## Discrete Probability Exercise Answers

1.

a) Notice how for a discrete probability distribution, 2 and 3 are the only integers that can fit within the interval $1 < X \leq 3$

$$\begin{aligned} P(1 < X \leq 3) &= P(X = 2) + P(X = 3) \\ &= 0.4 + 0.2 \\ &= 0.6 \end{aligned}$$

b) To find the expected value:

$$\begin{aligned} E(X) &= \sum x\, p(x) \\ &= 0 \times 0.1 + 1 \times 0.2 + 2 \times 0.4 + 3 \times 0.2 + 4 \times 0.1 \\ &= 2 \end{aligned}$$

c) To find the variance:

$$Var(x) = E(X^2) - \mu^2$$

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $X^2$ | 0 | 1 | 4 | 9 | 16 |
| $P(X=x)$ | 0.1 | 0.2 | 0.4 | 0.2 | 0.1 |

$$E(X^2) = 0 \times 0.1 + 1 \times 0.2 + 4 \times 0.4 + 9 \times 0.2 + 16 \times 0.1$$
$$= 5.2$$

$$\mu^2 = 2^2 = 4$$

$$\therefore Var(X) = 5.2 - 4 = 1.2$$

Since the standard deviation, $\sigma = \sqrt{Var(x)}$

$$\therefore \sigma = \sqrt{1.2} \approx 1.095$$

2.

i) Since the sum of all probabilities in a discrete distribution is equal to 1:

$$\frac{1}{10} + \frac{2}{10} + \frac{9}{20} + \frac{3}{20} + x = 1$$
$$\therefore x = 0.1$$

ii) To calculate the expected value: $E(X) = \sum x\, p(x)$

$$E(x) = 1 \times \frac{1}{10} + 2 \times \frac{2}{10} + 3 \times \frac{9}{20} + 4 \times \frac{3}{20} + 5 \times \frac{1}{10}$$
$$= 2.95$$

iii) To calculate variance:

$$Var(X) = E(X^2) - \mu^2$$

| $Score\ x$ | 1 | 2 | 3 | 4 | 5 | $Total$ |
|---|---|---|---|---|---|---|
| $x^2$ | 1 | 4 | 9 | 16 | 25 | |
| $Frequency$ | 10 | 20 | 45 | 15 | 10 | 100 |
| $Relative\ Frequency$ | $\frac{1}{10}$ | $\frac{2}{10}$ | $\frac{9}{20}$ | $\frac{3}{20}$ | $x$ | 1 |

$$E(X^2) = 1 \times \frac{1}{10} + 4 \times \frac{2}{10} + 9 \times \frac{9}{20} + 16 \times \frac{3}{20} + 25 \times \frac{1}{10}$$
$$= 9.85$$

$$\therefore Var(X) = 9.85 - 2.95^2$$
$$= 1.1475$$

$$\sigma = \sqrt{Var(X)} \approx 1.07$$

3. $Step\ 1: Evalute\ the\ value\ of\ q$

Since the sum of all probabilities equals to 1:

$$2q + 6q + 3q + 4q = 1$$
$$\therefore q = \frac{1}{15}$$

$$Step\ 2: Substitute\ in\ the\ value\ of\ q\ and\ evaluate\ E(X)$$

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $2q = \frac{2}{15}$ | $6q = \frac{6}{15}$ | $3q = \frac{3}{15}$ | $4q = \frac{4}{15}$ |

$$E(X) = 0 \times \frac{2}{15} + 1 \times \frac{6}{15} + 2 \times \frac{3}{15} + 3 \times \frac{4}{15}$$
$$= \frac{8}{5}$$

4. A

a) Since the sum of probabilities equals to 1:

$$0.2 + 0.6p^2 + 0.1 + 1 - p + 0.1 = 1$$
$$0.6p^2 - p + 0.4 = 0$$

Using the quadratic equation to evaluate p:

$$p = \frac{- - 1 \pm \sqrt{(-1)^2 - 4 \times 0.4 \times 0.6}}{2 \times 0.6}$$
$$= \frac{1 \pm \sqrt{0.04}}{1.2}$$
$$= 1\ OR\ \frac{2}{3}$$

b)

a) Before finding $E(X) = \sum x\, p(x)$, we should first refill our table now that the value of p is known

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.2 | $\frac{4}{15}$ | 0.1 | $\frac{1}{3}$ | 0.1 |

$$E(X) = 0 \times 0.2 + 1 \times \frac{4}{15} + 2 \times 0.1 + 3 \times \frac{1}{3} + 4 \times 0.1$$
$$= \frac{4}{15} + 0.2 + 1 + 0.4$$
$$= \frac{28}{15}$$

b) Consider $P(X \geq E(X))$, where $E(X) = \frac{28}{15}$

$$P\left(X \geq \frac{28}{15}\right) = P(X = 2) + P(X = 3) + P(X = 4)$$

$$= 0.1 + \frac{1}{3} + 0.1$$
$$= \frac{8}{15}$$

c) The variance is given by: $Var(X) = E(X^2) - \mu^2$

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $X^2$ | 0 | 1 | 4 | 9 | 16 |
| $P(X = x)$ | 0.2 | $\frac{4}{15}$ | 0.1 | $\frac{1}{3}$ | 0.1 |

$$E(X^2) = 0 \times 0.2 + 1 \times \frac{4}{15} + 4 \times 0.1 + 9 \times \frac{1}{3} + 16 \times 0.1$$
$$= \frac{79}{15}$$

$$\therefore E(X^2) - \mu^2 = \frac{79}{15} - \left(\frac{28}{15}\right)^2$$
$$= \frac{401}{225}$$
$$Var(X) = \frac{401}{225}$$

5. To evaluate a and b, we have to create a pair of simultaneous equations

Since the sum of probabilities equals to 1:

$$a + b + 0.3 = 1$$
$$a + b = 0.7$$

Since $E(X) = \sum x\, p(x) = 0.8$ :

$$0 \times a + 1 \times b + 2 \times 0.3 = 0.8$$
$$b + 0.6 = 0.8$$
$$\therefore b = 0.2$$

Sub $b = 0.2$ into the previous equation:

$$a + 0.2 = 0.7$$
$$\therefore a = 0.5$$