Definition 1. *The general rule of computing expectation (or easier word, mean) of some real values $g(x)$, is defined by*

$$E(g(X)) = \sum_x g(x)p_X(x)$$

Expectation has a very nice property, it is a linear operator.

Property 1.

$$E\left(\sum_{i=1}^{n} a_i g_i(X)\right) = \sum_{i=1}^{n} a_i E(g_i(x))$$

Where $a_i, g_i(x) \in \mathbb{R}, \forall i = 1, 2, ..., n$.

Proof.

$$\begin{aligned}
E\left(\sum_{i=1}^{n} a_i g_i(X)\right) &= \sum_x \left(\sum_{i=1}^{n} a_i g_i(X)\right) p_X(x) \\
&= \sum_x (a_1 g_1(x) + a_2 g_2(x) + \cdots + a_n g_n(x)) p_X(x) \\
&= a_1 \sum_x g_1(x)p_X(x) + a_2 \sum_x g_2(x)p_X(x) + \cdots + a_n \sum_x g_n(x)p_X(x) \\
&= \sum_{i=1}^{n} a_i E(g(X))
\end{aligned}$$

□

Definition 2. *The expectation value of X is defined by*

$$E(X) = \sum_x x p_X(x)$$

Since the expectation value is linear, we easily see that

Property 2.

$$E(aX + b) = aE(X) + b, \quad a, b \in \mathbb{R}$$

This property states that if a random variable X is transformed linearly to $aX + b$, then the expectation value of this new random variable is just the transformed version of the original one, namely, $aE(X) + b$.

Geometrically, if all data on the histogram is transformed to some where else, then the measure of location(in this case, mean) will be also transformed in the same manner.

Definition 3. *The variance of the X is defined by*

$$\begin{aligned}
Var(X) &= E\left((X - E(X))^2\right) \\
&= E(X^2) - (E(X))^2 \\
&= \sum_x x^2 p_X(x) - \left(\sum_x x p_X(x)\right)^2
\end{aligned}$$

Variance also has a meaningful property due to the linearity of expectation.

Property 3.

$$Var(aX + b) = a^2 Var(X), \quad a, b \in \mathbb{R}$$

Proof.

$$\begin{aligned}
Var(aX + b) &= E((aX + b)^2) - (E(aX + b))^2 \\
&= E((aX)^2) + E(2abX) + E(b^2) - (aE(X) + b)^2 \\
&= a^2E(X^2) + 2abE(X) + b^2 - a^2(E(X))^2 - 2abE(X) - b^2 \\
&= a^2(E(X^2) - (E(X))^2) \\
&= a^2Var(X)
\end{aligned}$$

□

This property states that if data is being shifted left or right, this spread of the data would not changed, thus the value b does not have effect on variance; however, if the scale of the data is shrink or stretched, then the spread of the data would also be changed, thus the value a has effect on variance.