Expectation, Variance and Standard Deviation

Measure of central tendency: gives an idea about the central value of the probability distribution around which values of the random variable are clustered. Mean, median and mode are three commonly used measures of central tendency.

Measure of dispersion: Measures of central tendency give us the idea about the location of only central part of the distribution. Other measures are often needed to describe a probability distribution. A probability distribution (or the corresponding random variable) is said to have a high dispersion if its support contains many values that are significantly higher or lower than the mean or median value. Some of the commonly used measures of dispersion are standard deviation, quartile deviation (or semi-inter-quartile range) and coefficient of variation.

Definition 1. (1) *Let X be a discrete random variable with p.m.f. f_X and support E_X. We say that the expected value of X or mean of X or expectation of X (denoted by $E(X)$) is finite and equals*

$$E(X) = \sum_{x \in E_X} x f_X(x)$$

provided the series $\sum_{x \in E_X} |x| f_X(x)$ is convergent, i.e., $\sum_{x \in E_X} |x| f_X(x) < \infty$.

(2) *Let X be a continuous random variable with p.d.f. f_X. We say that the expected value of X or mean of X or expectation of X (denoted by $E(X)$) is finite and equals*

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

provided the integral $\int_{-\infty}^{\infty} |x| f_X(x) dx$ is convergent, i.e., $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$.

Remark 2. (1) *Let X be a discrete random variable with finite support E_X and the p.m.f. f_X. Then $\sum_{x \in E_X} |x| f_X(x) < \infty$. Hence, $E(X)$ is finite.*

(2) *Let X be a continuous random variable with support $E_X \subseteq [-a, a]$ and p.d.f. f_X, for some $0 < a < \infty$. Then $\int_{-\infty}^{\infty} |x| f_X(x) dx = \int_{-a}^{a} |x| f_X(x) dx \leq a \int_{-a}^{a} f_X(x) dx = a$. Hence, $E(X)$ is finite.*

(3) *Let X be a continuous random variable with the d.f. F_X and p.d.f. f_X. Then*

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f_X(x)dx \\
&= \int_{-\infty}^{0} x f_X(x)dx + \int_{0}^{\infty} x f_X(x)dx \\
&= -\int_{-\infty}^{0}\int_{x}^{0} f_X(x)dtdx + \int_{0}^{\infty}\int_{0}^{x} f_X(x)dtdx \\
&= -\int_{-\infty}^{0}\int_{-\infty}^{t} f_X(x)dxdt + \int_{0}^{\infty}\int_{t}^{\infty} f_X(x)dxdt \ (\textit{by the change of order of integration}) \\
&= -\int_{-\infty}^{0} P(\{X < t\})dt + \int_{0}^{\infty} P(\{X > t\})dt \\
&= \int_{0}^{\infty} (1 - F_X(t))dt - \int_{-\infty}^{0} F_X(t-)dt
\end{aligned}
$$

Hence $E(X)$ does not depend on the version of p.d.f. although p.d.f. may not be unique.

Example 3. *Let X be a random variable with p.m.f.*

$$
f_X(x) = \begin{cases} \frac{1}{2^x}, & \textit{if } x \in \{1, 2, 3, \cdots\} \\ 0, & \textit{otherwise} \end{cases}
$$

Show that the expected value of X is finite and find its value.

Solution: The support $E_X = \{1, 2, 3, \cdots\}$. By the ratio test, we can see that the infinite series $\sum_{x \in E_X} |x| f_X(x) = \sum_{n=1}^{\infty} \frac{n}{2^n}$ is convergent. Hence, the expected value of X is finite.

Now $E(X) = \sum_{x \in E_X} x f_X(x) = \sum_{n=1}^{\infty} \frac{n}{2^n} = \lim_{n \to \infty} s_n$, where the partial sum $s_n = 2[1 - \frac{1}{2^n} - \frac{n}{2^{n+1}}]$. This implies that $E(X) = 2$.

Example 4. *Let X be a random variable with p.d.f. $f_X(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < \infty$ Show that the expected value of X is not finite.*

Solution: We have $\int_{-\infty}^{\infty} |x| f_X(x)dx = \int_{-\infty}^{\infty} |x| \frac{1}{\pi(1+x^2)}dx = \int_{0}^{\infty} \frac{2x}{\pi(1+x^2)}dx$. Since the improper integral $\int_{0}^{\infty} \frac{2x}{\pi(1+x^2)}dx$ is not convergent, the expected value of X is not finite.

Theorem 5. (1) *Let X be a random variable of discrete type with the support E_X and the p.m.f. f_X. Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ be a function such that $h^{-1}(A) \in \mathbb{B}_{\mathbb{R}}$, for every $A \in \mathbb{B}_{\mathbb{R}}$. Then*

$$
E(h(X)) = \sum_{x \in E_X} h(x) f_X(x);
$$

(2) *Let X be a random variable of continuous type with p.d.f. f_X. Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ be a function such that $h^{-1}(A) \in \mathbb{B}_{\mathbb{R}}$, for every $A \in \mathbb{B}_{\mathbb{R}}$. Then*

$$E(h(X)) = \int_{-\infty}^{\infty} h(x) f_X(x);$$

(3) *If, for real constants a and b with $a \leq b$, $P(\{a \leq X \leq b\}) = 1$, then $a \leq E(X) \leq b$;*

(4) *Let $h_i : \mathbb{R} \longrightarrow \mathbb{R}$ be a function such that $h_i^{-1}(A) \in \mathbb{B}_{\mathbb{R}}$, for every $A \in \mathbb{B}_{\mathbb{R}}$, for $i = 1, 2$. If $P(\{h_1(X) \leq h_2(X)\}) = 1$, then $E(h_1(X)) \leq E(h_2(X))$, provided the involved expectations are finite;*

(5) *Let $h_i : \mathbb{R} \longrightarrow \mathbb{R}$ be a function such that $h_i^{-1}(A) \in \mathbb{B}_{\mathbb{R}}$, for every $A \in \mathbb{B}_{\mathbb{R}}$, for $i = 1, 2, \cdots, m$. Then*

$$E(\sum_{i=1}^{m} h_i(X)) = \sum_{i=1}^{m} E(h_i(X)),$$

provided the involved expectations are finite;

(6) *If $P(\{X \geq 0\}) = 1$ and $E(X) = 0$, then $P(\{X = 0\}) = 1$.*

Definition 6. *Let X be a random variable.*

(1) *For $r \in \{1, 2, \cdots\}$, $\mu_r' = E(X^r)$, provided it is finite, is called the r-th moment of the random variable X;*

(2) *For $r \in \{1, 2, \cdots\}$, $E(|X|^r)$, provided it is finite, is called the r-th absolute moment of the random variable X;*

(3) *For $r \in \{1, 2, \cdots\}$, $\mu_r = E((X - \mu_1')^r)$, provided it is finite, is called the r-th central moment of the random variable X;*

(4) *$\mu_2 = E((X - \mu_1')^2) = E((X - E(X))^2)$, provided it is finite, is called the variance of the random variable X. The variance of the random variable X is denoted by $Var(X)$. The quantity $\sigma = \sqrt{Var(X)} = \sqrt{E((X - E(X))^2)}$ is called the standard deviation of the random variable X.*

Proposition 7. *Let X be a random variable with finite first two moments and $\mu = E(X)$. Then*

(1) *For real constants a and b, $E(aX + b) = aE(X) + b$;*

(2) *$Var(X) = E(X^2) - (E(X))^2$;*

(3) *$Var(X) \geq 0$. Moreover, $Var(X) = 0$ if and only if, $P(\{X = \mu\}) = 1$;*

(4) *$E(X^2) \geq (E(X))^2$ (Cauchy- Schwarz inequality);*

(5) *For real constants a and b, $Var(aX + b) = a^2 Var(X)$.*

Proof. (1) Consider the function $h : \mathbb{R} \longrightarrow \mathbb{R}$, defined by $h(x) = ax + b$, for all $x \in \mathbb{R}$. Since h is a continuous function, $h(X) = aX + b$ is random variable.

Suppose X is a discrete type random variable with the p.m.f. f_X and the support E_X. Then by Theorem 5(1),

$$\begin{aligned}
E(h(X)) &= E(aX+b)\\
&= \sum_{x\in E_X} h(x)f_X(x)\\
&= \sum_{x\in E_X} (ax+b)f_X(x)\\
&= a\sum_{x\in E_X} xf_X(x) + b\sum_{x\in E_X} f_X(x)\\
&= aE(X)+b \text{ (since } \sum_{x\in E_X} f_X(x)=1)
\end{aligned}$$

Similarly, we can prove it for continuous case.

(2)

$$\begin{aligned}
Var(X) &= E((X-\mu)^2)\\
&= E(X^2-2\mu X+\mu^2)\\
&= E(X^2)-2\mu E(X)+\mu^2 \text{ (by using (1) and Theorem 5(4))}\\
&= E(X^2)-(E(X))^2
\end{aligned}$$

(3) Since $P(\{(X-\mu)^2\geq 0\}) = P(\mathcal{S}) = 1$, using Theorem 5(3), $E(0)\leq E((X-\mu)^2)$. So, $Var(X)\geq 0$.

Also, using Theorem 5(5), if $Var(X) = E((X-\mu)^2) = 0$, then $P(\{(X-\mu)^2 = 0\}) = 1$, i.e., $P(\{X=\mu\}) = 1$. Conversely, if $P(\{X=\mu\}) = 1$, then $E(X)=\mu$ and $E(X^2)=\mu^2$ (as support of X is $\{\mu\}$). Now, $Var(X) = E(X^2)-(E(X))^2 = \mu^2-\mu^2 = 0$.

(4) Since $Var(X)\geq 0$ and $Var(X) = E(X^2)-(E(X))^2$, $E(X^2)\geq (E(X))^2$.

(5) Let $Y = aX+b$. Then, (1), we have $E(Y) = aE(X)+b$. So, $Y-E(Y) = a(X-E(X))$.

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

$\square$

Example 8. *Let X be a random variable with p.m.f.*

$$f_X(x) = \begin{cases} \frac{1}{6}, & \text{if } x\in\{1,2,3,4,5,6\}\\ 0, & \text{otherwise}\end{cases}$$

Find the expectation and variance of X.

Solution: $E(X) = \sum_{x\in E_X} xf_X(x) = \frac{7}{2}$.

$E(X^2) = \sum\limits_{x \in E_X} x^2 f_X(x) = \frac{91}{6}$.

Hence, $Var(X) = E(X^2) - (E(X))^2 = \frac{25}{12}$.

Example 9. *Let X be a random variable with p.m.f.*

$$f_X(x) = \begin{cases} 0.2, & \textit{if } x = 0 \\ 0.5, & \textit{if } x = 1 \\ 0.3, & \textit{if } x = 2 \\ 0, & \textit{otherwise} \end{cases}$$

(1) *Find the p.m.f. of $Y = X^2$ and hence find expectation of Y.*
(2) *Find expectation of Y directly.*

Solution:

(1) The support of Y is $E_Y = \{x^2 \mid x \in E_X\} = \{0, 1, 4\}$. Now, the p.m.f. of $Y = X^2$ is

$$f_Y(y) = \begin{cases} 0.2, & \text{if } y = 0 \\ 0.5, & \text{if } y = 1 \\ 0.3, & \text{if } y = 4 \\ 0, & \text{otherwise} \end{cases}$$

Hence, $E(Y) = \sum\limits_{y \in E_Y} y f_Y(y) = 1.7$

(2) $E(X^2) = \sum\limits_{x \in E_X} x^2 f_X(x) = 1.7$

Example 10. *Let X be a random variable with p.d.f.*

$$f_X(x) = \begin{cases} 1, & \textit{if } 0 < x < 1 \\ 0, & \textit{otherwise} \end{cases}$$

(1) *Find the p.d.f. of $Y = X^3$ and hence find expectation of Y.*
(2) *Find expectation of Y directly.*

Solution:

(1) The c.d.f. of Y is

$$F_Y(y) = P(Y \leq y) = \begin{cases} 0, & \text{if } y \leq 0 \\ \int\limits_0^{y^{1/3}} dx, & \text{if } 0 < y < 1 \\ \int\limits_0^1 dx, & \text{if } y \geq 1 \end{cases}$$

Thus

$$F_Y(y) = \begin{cases} 0, & \text{if } y \leq 0 \\ y^{1/3}, & \text{if } 0 < y < 1 \\ 1, & \text{if } y \geq 1 \end{cases}$$

Hence the p.d.f. of Y is

$$f_Y(y) = \begin{cases} \frac{1}{3} y^{-2/3}, & \text{if } 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Now, $E(Y) = \int\limits_{-\infty}^{\infty} y f_Y(y)dy = \int\limits_0^1 y f_Y(y)dy = \int\limits_0^1 \frac{1}{3} y^{1/3} dy = \frac{1}{4}$.

(2) $E(Y) = E(X^3) = \int\limits_{-\infty}^{\infty} x^3 f_X(x)dx = \int\limits_0^1 x^3 dx = \frac{1}{4}$.

Example 11. *Let X be a random variable with p.d.f.*

$$f_X(x) = \begin{cases} \frac{1}{2}, & \textit{if } -2 < x < -1 \\ \frac{x}{9}, & \textit{if } 0 < x < 3 \\ 0, & \textit{otherwise} \end{cases}$$

(1) *If $Y_1 = \max(X, 0)$, find the mean and variance of Y_1.*
(2) *If $Y_2 = 2X + 3e^{-\max(X,0)} + 4$, find the mean Y_2.*

Solution:

(1) Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ be a function defined by $h(x) = \max(x, 0)$, for all $x \in \mathbb{R}$. Then $Y_1 = h(X)$. For $r > 0$, $E(Y_1^r) = \int\limits_{-\infty}^{\infty} (\max(x, 0))^r f_X(x)dx = \int\limits_0^3 \frac{x^{r+1}}{9} dx = \frac{3^r}{r+2}$. Hence $E(Y_1) = 1$ and $Var(Y_1) = E(Y_1^2) - (E(Y_1))^2 = \frac{9}{4} - 1 = \frac{5}{4}$.
(2) We have

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x)dx \\ &= \int_{-2}^{-1} \frac{x}{2} dx + \int_0^3 \frac{x^2}{9} dx \\ &= \frac{1}{4} \end{aligned}$$

and

$$\begin{aligned} E(e^{-\max(X,0)}) &= \int_{-\infty}^{\infty} e^{-\max(x,0)} f_X(x)dx \\ &= \int_{-2}^{-1} \frac{1}{2} dx + \int_0^3 \frac{xe^{-x}}{9} dx \\ &= \frac{11 - 8e^{-3}}{18}. \end{aligned}$$

Therefore, $E(Y_2) = 2E(X) + 3E(e^{-\max(X,0)}) + 4 = \frac{19-4e^{-3}}{3}$

Definition 12. *Let X be a random variable with the d.f. F_X. A real number x satisfying*

$$F_X(x-) \leq \frac{1}{2} \leq F_X(x), \;\; i.e., \;\; P(\{X < x\}) \leq \frac{1}{2} \leq P(\{X \leq x\})$$

is called a median of X. A median is not necessarily unique.