# Section 2.3 Mathematical Expectation and Variance

School of Sciences, BUPT

# Contents

# Expectation of a Random Variable

The distribution of a random variable $X$ contains all of the probabilistic information about $X$.

But it is difficult to get the d.f. of the random variable.

Without trying to describe the entire distribution, How could we obtain useful information of r.v.?

In this section, we mainly introduce two measures: the expectation and the variance.

# Expectation of a Random Variable

| Winning | $n = 100$ | | $n = 10000$ | |
|---|---|---|---|---|
| | Frequency | Relative Frequency | Frequency | Relative Frequency |
| 1 | 17 | 0.17 | 1681 | 0.1681 |
| -2 | 17 | 0.17 | 1678 | 0.1678 |
| 3 | 16 | 0.16 | 1626 | 0.1626 |
| -4 | 18 | 0.18 | 1696 | 0.1696 |
| 5 | 16 | 0.16 | 1686 | 0.1686 |
| -6 | 16 | 0.16 | 1633 | 0.1633 |

Table: Frequencies for dice game

# Expectation of a Random Variable

In the first run, we have played the game 100 times.

$$\begin{aligned}\text{average gain} &= 1 \cdot F_{100}(1) + (-2) \cdot F_{100}(-2) + 3 \cdot F_{100}(3) \\ &\quad + (-4) \cdot F_{100}(-4) + 5 \cdot F_{100}(5) + (-6) \cdot F_{100}(-6) \\ &= -0.57\end{aligned}$$

To get a better idea, we have played the game 10,000 times.

$$\begin{aligned}\text{average gain} &= 1 \cdot F_{10000}(1) + (-2) \cdot F_{10000}(-2) + 3 \cdot F_{10000}(3) \\ &\quad + (-4) \cdot F_{10000}(-4) + 5 \cdot F_{10000}(5) + (-6) \cdot F_{10000}(-6) \\ &= 1 \cdot 0.1681 - 2 \cdot 0.1678 + 3 \cdot 0.1626 \\ &\quad - 4 \cdot 0.1696 + 5 \cdot 0.1686 - 6 \cdot 0.1633 \\ &= -0.4949\end{aligned}$$

$$1 \cdot p(1) - 2 \cdot p(-2) + 3 \cdot p(3) - 4 \cdot p(-4) + 5 \cdot p(5) - 6 \cdot p(-6) = -0.5.$$

# Expectation of a Random Variable

加权平均，期望公式

**Definition**

*The **expectation** (the **mean** or the **expected value**) of a random variable $X$ is given by*

求 到积、再和

$$\mu := E(X) = \begin{cases} \displaystyle\sum_{i=1}^{\infty} x_i p(x_i) & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{\infty} x f(x) dx & \text{if } X \text{ is continuous.} \end{cases} \tag{1}$$

In words, the expected value of $X$ is a weighted average of the possible values that $X$ can take on, each value being weighted by the probability that $X$ assumes it.

# Expectation of a Random Variable

However, the idea of adding the value times the probability to find the expectation does not quite work for continuous random variables.

Why?

By returning to the Riemann sum approach, we will find the continuous case of equation (1) in Definition is reasonable.

# Expectation of a Random Variable

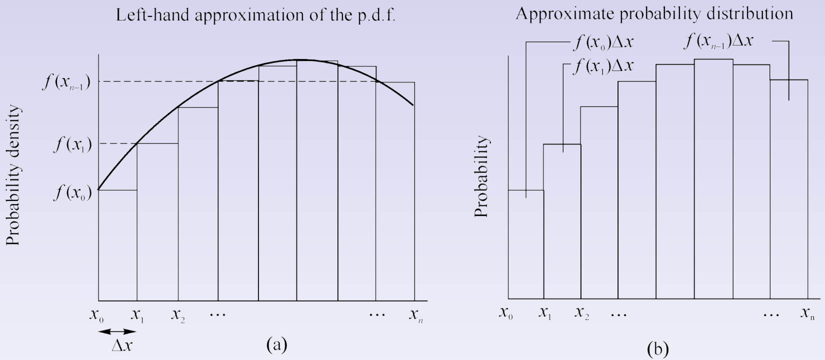


Figure: From discrete to continuous

$$\sum_{i=0}^{n-1} x_i f(x_i) \triangle x \to \int_{-\infty}^{\infty} x f(x) dx, \quad n \to \infty \;\; and \; \triangle x \to 0$$

# Expectation of a Random Variable

## Example

Find $E(X)$, where the p.f. of $X$ is $p(0) = 0.1$ and $p(1) = 0.9$.

**Solution**. Since $p(0) = 0.1$ and $p(1) = 0.9$, we obtain

$$E(X) = 0 \cdot 0.1 + 1 \cdot 0.9 = 0.9.$$

□

## Remark

*The probability concept of expectation is analogous to the physical concept of the center of gravity of a distribution of a mass.*

# Expectation of a Random Variable

## Example

Suppose that you plan on starting a new business and would like to know the expected return for the business. You develop a subjective probability distribution of the returns and their associated probabilities. Note: $X$=return in \$1,000/year (e.g., $-10$ means a loss of \$10,000).

| $x$ | $-10$ | 0 | 10 | 20 |
|---|---|---|---|---|
| $p(x)$ | 0.20 | 0.25 | 0.40 | 0.15 |

What is the expected return for this business?

# Expectation of a Random Variable

**Solution.**

$$\begin{aligned} E(X) &= \sum x \cdot p(x) \\ &= x_1 P(x_1) + x_2 P(x_2) + x_3 P(x_3) + x_4 P(x_4) \\ &= -10 \cdot (0.2) + 0 \cdot (0.25) + 10 \cdot (0.4) + 20 \cdot (0.15) \\ &= -2 + 0 + 4 + 3 = 5. \end{aligned}$$

The expected return for this business is \$5,000. □

# Expectation of a Random Variable

## Example

Suppose that we flip a fair coin until a head first appears, and if the number of tosses equals $n$, then we are paid $2^n$ dollars. What is the expected value of the payment?

**Solution.** We let $Y$ represent the payment. Then,

$$P(Y = 2^n) = \frac{1}{2^n}$$

for $n \geqslant 1$. Thus,

$$E(Y) = \sum_{n=1}^{\infty} 2^n \cdot \frac{1}{2^n}$$

which is a divergent sum. Thus, $Y$ has no expectation. □

***St. Petersburg Paradox.***

# Expectation of a Random Variable

The expectation $E(X)$ exists if and only if the summation in equation (1) is absolutely convergent, that is,

$$\sum_{all\ \ x} |x|p(x) < \infty \quad \text{or} \quad \int_x |x|f(x)dx < \infty. \tag{2}$$

Thus, we know if relation (2) is satisfied, then $E(X)$ exists and its value is given by equation (1).

If relation (2) is not satisfied, then $E(X)$ does not exists.

# Expectation of a Random Variable

Now let us see some examples of continuous case.

## Example

Suppose that the p.d.f. of a random variable $X$ is

$$f(x) = \begin{cases} cx & \text{for } 0 < x < 4, \\ 0 & \text{otherwise,} \end{cases}$$

where $c$ is a given constant. Determine the value of $c$ and the expected value of $X$.

**Solution**. For every p.d.f., it must be true that $\int_{-\infty}^{+\infty} f(x)dx = 1$. Therefore, in this example,

$$\int_0^4 cxdx = 8c = 1.$$

Hence, $c = 1/8$. It follows that

$$E(X) = \int_0^4 x \cdot \frac{1}{8} xdx = \frac{8}{3}.$$

$\square$

# Expectation of a Random Variable

## Example

Suppose that the d.f. of a random variable $X$ is as follows:

$$F(x) = \begin{cases} e^{x-3} & \text{for } x \leqslant 3, \\ 1 & \text{for } x > 3. \end{cases}$$

Determine the expected value of $X$.

**Solution.** Of course we can find the p.d.f. of $X$ from d.f., then $E(X)$ can be obtained by using the formula in Definition. But another way is using d.f. directly to get $E(X)$.

$$E(X) = \int_{-\infty}^{+\infty} x dF(x) = \int_{-\infty}^{3} x de^{x-3}$$

$$= xe^{x-3}\Big|_{-\infty}^{3} - e^{x-3}\Big|_{-\infty}^{3} = 3 - 1 = 2.$$

□

# Expectation of a Random Variable

## Example

Suppose that $X$ is a continuous random variable and the p.d.f of $X$ is

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

This distribution is called the ***Cauchy distribution***. $E(X)$ does not exist since

$$\int_{-\infty}^{\infty} |x| f(x) dx = 2\int_0^{\infty} \frac{1}{\pi}\frac{x}{1+x^2} dx = \frac{1}{\pi}\ln(1+x^2)\Big|_0^{\infty} \text{ diverges.}$$

□

# Contents

# Expectation of Functions of a Random Variable

How can we compute the expected value of $g(X)$,

some function of $X$?

# Expectation of Functions of a Random Variable

In fact, we can compute $E[g(X)]$ by the following easier way.

**Proposition**

*The mathematical expectation of a function $g(X)$ of a random variable $X$ can be calculated by*

$$E[g(X)] = \begin{cases} \displaystyle\sum_{i=1}^{\infty} g(x_i)p(x_i) & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{\infty} g(x)f(x)dx & \text{if } X \text{ is continuous.} \end{cases} \tag{3}$$

The expectation $E[g(X)]$ will exist if and only if

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty.$$

# Expectation of Functions of a Random Variable

**Example**

Let $X$ denote a random variable that takes on any of the values $-1, 0$, and $1$ with respective probabilities

$$P(X = -1) = 0.2 \qquad P(X = 0) = 0.5 \qquad P(X = 1) = 0.3$$

Compute $E(X^2)$.

**Solution**. Let $Y = X^2$. Then the probability mass function of $Y$ is given by

$$P(Y = 1) = P(X = -1) + P(X = 1) = 0.5$$
$$P(Y = 0) = P(X = 0) = 0.5.$$

Hence, $E(Y) = E(X^2) = 1 \cdot 0.5 + 0 \cdot 0.5 = 0.5.$

$E(Y) = E(X^2) = (-1)^2 \cdot 0.2 + 0 \cdot 0.5 + 1^2 \cdot 0.3 = 0.5.$

# Expectation of Functions of a Random Variable

**Proof of Proposition** As in the preceding verification, we will group together all the terms in $\sum_i g(x_i)p(x_i)$ having the same value of $g(x_i)$. Specifically, suppose that $y_j, j \geqslant 1$, represent the different values of $g(x_i), i \geqslant 1$. Then, grouping all the $g(x_i)$ having the same value gives

$$\begin{aligned}
\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\
&= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \\
&= \sum_j y_j P(g(X) = y_j) \\
&= E(g(X)).
\end{aligned}$$

$\square$

# Expectation of Functions of a Random Variable

## Example

Let $X$ be a continuous nonnegative random variable with density function

$$f_X(x) = \begin{cases} \frac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise.} \end{cases}$$

and let $Y = 2X + 1$. Compute $E(Y)$.

**Solution.** We have known the p.d.f. of $Y$. So

$$E(Y) = \int_{-\infty}^{+\infty} y f_Y(y) dy = \int_1^9 y \cdot \frac{y-1}{32} dy = \frac{19}{3}.$$

$$E(Y) = E(2X+1) = \int_{-\infty}^{+\infty} (2x+1) f(x) dx = \int_0^4 (2x+1) \cdot \frac{x}{8} dx = \frac{19}{3}.$$

# Contents

# Properties of Expectation of a Random Variable

## Proposition

(i) *$E(c) = c$, for any constant $c$.*
(ii) *If there exists a constant such that $P(X \geqslant a) = 1$, then $E(X) \geqslant a$. If there exists a constant such that $P(X \leqslant b) = 1$, then $E(X) \leqslant b$.*

From Proposition, we can get

## Corralary

(i) *If $P(a \leqslant X \leqslant b) = 1$, then $a \leqslant E(X) \leqslant b$.*

(ii) *If $P(X \geqslant a) = 1$ and $E(X) = a$, then it must be true that $P(X > a) = 0$ and $P(X = a) = 1$.*

# Properties of Expectation of a Random Variable

## Proposition

*For all constants a and b, we have*

$$E(aX + b) = aE(X) + b. \tag{4}$$

linearity

# Properties of Expectation of a Random Variable

## Example

Suppose that the p.d.f. of $V$ is

$$f(v) = \begin{cases} \dfrac{1}{a} & \text{for } 0 < v < a, \\ 0 & \text{otherwise,} \end{cases}$$

(1)Determine the value of $E(W) = E[g(V)] = E(kV^2)$, where $k$ is a constant; (2) Calculate the value of $E(3W + 5)$.

**Solution**.

$$(1) E(W) = E(g(V)) = \int_{-\infty}^{+\infty} kv^2 \cdot f(v) dv = \int_0^a kv^2 \cdot \frac{1}{a} dv = \frac{1}{3} ka^2.$$

$$(2) E(3W + 5) = 3E(W) + 5 = ka^2 + 5.$$

# Contents

# Variance of a Random Variable

In this section we shall introduce a measure, the variance, which is an indication of the variability of a random variable.

## Definition

*The **variance** of a random variable $X$ is defined by*

Definition

$$\sigma^2 \equiv E\left[\left(X - E(X)\right)^2\right] = Var(X) = \begin{cases} \sum_{i=1}^{\infty} (x_i - \mu)^2 p(x_i) \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \end{cases}$$

*where* $\mu = E(X)$.

# Variance of a Random Variable

## Remark

$$Var(X) \geq 0.$$

Finally,
the larger the variance is,
the more spread out is the distribution of the random variable around its mean.

# Variance of a Random Variable

**Question**

What does the variance mean?

The quantity is rather hard to interpret.

So we take square root to change it.

**Definition**

*The* ***standard deviation*** *of a random variable $X$ is given by*

$$\sigma = \sqrt{Var(X)}.$$

# Variance of a Random Variable

**Proposition**

$$Var(X) = E(X^2) - [E(X)]^2. \tag{5}$$

**Proof.** Assume $X$ is discrete. Suppose that $\mu = E(X)$.

$$\begin{aligned} Var(X) &= E\left[\left(X - E(X)\right)^2\right] = E[(X-\mu)^2] \\ &= \sum_x (x-\mu)^2 p(x) = \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\ &= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \sum_x \mu^2 p(x) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2. \end{aligned}$$

The continuous case is similar to prove. $\square$

# Variance of a Random Variable

**Example**

Suppose that the p.d.f. of a random variable $X$ is

$$f(x) = \begin{cases} \frac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise,} \end{cases}$$

Find $Var(X)$.

**Solution.**

$$Var(X) = \int_0^4 \left(x - \frac{8}{3}\right)^2 \cdot \frac{1}{8} x dx = \int_0^4 x^2 \cdot \frac{1}{8} x dx - \left(\frac{8}{3}\right)^2 = \frac{8}{9}.$$

By using the equation (5), we also have

$$Var(X) = E(X^2) - E^2(X) = \int_0^4 x^2 \frac{1}{8} x dx - \left(\frac{8}{3}\right)^2 = \frac{8}{9}.$$

$\square$

# Variance of a Random Variable

## Theorem

*$Var(X) = 0$ if and only if there exists a constant $c$ such that $P(X = c) = 1$.*

$c = E(X)$

**Proof.** Suppose first that there exists a constant $c$ such that $P(X = c) = 1$. Then $E(X) = c$, and $P[(X - c)^2 = 0] = 1$. Therefore,

$$Var(X) = E[(X - c)^2] = 0.$$

Conversely, suppose that $Var(X) = 0$. Then $P\big[\big(X - E(X)\big)^2 \geqslant 0\big] = 1$ but $E\big[\big(X - E(X)\big)^2\big] = 0$. Therefore, it can be seen that

$$P\big[\big(X - E(X)\big)^2 = 0\big] = 1.$$

Hence, $P(X = E(X)) = 1$. $\square$

# Variance of a Random Variable

## Proposition

*For any constants $a$ and $b$,*

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

**Proof**. For any constants $a$ and $b$,

$$\begin{aligned} Var(aX+b) &= E\left\{\left[(aX+b) - E(aX+b)\right]^2\right\} \\ &= E\left\{a^2\left[X - E(X)\right]^2\right\} = a^2 Var(X). \end{aligned}$$

$\square$

Thus, the operator $Var$ is not linear.

# Variance of a Random Variable

**Example**

Let $X^* = \frac{X - E(X)}{\sqrt{Var(X)}}$. Find $E(X^*)$ and $Var(X^*)$.

**Solution**.

$$E(X^*) = E\left(\frac{X - E(X)}{\sqrt{Var(X)}}\right) = \frac{1}{\sqrt{Var(X)}}[E(X) - E(X)] = 0.$$

$$Var(X^*) = Var\left(\frac{X - E(X)}{\sqrt{Var(X)}}\right) = \frac{Var(X)}{Var(X)} = 1.$$

$\square$

Since $E(X^*) = 0$ and $Var(X^*) = 1$, $X^*$ is usually said to be *standard random variable*.

# Moments of a Random Variable

The mean and the variance of a random variable are particular cases of the quantities known as the ***moments*** of this variable.

**Definition**

*The $k$**th moment about the origin**, or **noncentral moment** of a discrete random variable $X$ is defined by*

$$E(X^k) \equiv \mu_k' = \sum_{i=1}^{\infty} x_i^k p(x_i)$$

**Definition**

*The $k$**th moment about the center**, or **central moment** of a discrete random variable $X$ is defined by*

$$E[(X-\mu)^k] \equiv \mu_k = \sum_{i=1}^{\infty} (x_i-\mu)^k p(x_i).$$

# Moments of a Random Variable

**Theorem**

*If $E(|X|^k) < \infty$ for some positive integer $k$, then $E(|X|^j) < \infty$ for every integer $j$ such that $j < k$.*

**Proof.** For convenience, we assume that the distribution of $X$ is continuous and the p.d.f. is $f$. Then

$$\begin{aligned} E(|X^j|) &= \int_{-\infty}^{\infty} |x^j| f(x) dx \\ &= \int_{|x| \leqslant 1} |x|^j f(x) dx + \int_{|x|>1} |x|^j f(x) dx \\ &\leqslant \int_{|x| \leqslant 1} 1 \cdot f(x) dx + \int_{|x|>1} |x|^j f(x) dx \\ &\leqslant P(|X| \leqslant 1) + E(|X|^k). \end{aligned}$$

By hypothesis, $E(|X|^k) < \infty$. It therefore follows that $E(|X|^j) < \infty$. $\square$

# Moments of a Random Variable

## Example

Suppose that $X$ has a continuous distribution for which the p.d.f. has the following form

$$f(x) = ce^{-(x-3)^2}, \qquad -\infty < x < \infty.$$

Determine the mean of $X$ and all the odd central moments.

**Solution**. It can be shown that for every positive integer $k$,

$$\int_{-\infty}^{\infty} |x|^k e^{-(x-3)^2} dx < \infty.$$

# Moments of a Random Variable

**Solution**. Hence, all the moments of $X$ exist.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} cxe^{-(x-3)^2}dx \\ &= \int_{-\infty}^{\infty} c(x-3)e^{-(x-3)^2}dx + 3\int_{-\infty}^{\infty} ce^{-(x-3)^2}dx \quad \text{let } t = x-3, \\ &= \int_{-\infty}^{\infty} cte^{-t^2}dt + 3 = 0 + 3 = 3. \end{aligned}$$

For every odd positive integer $k$, it follows that

$$E[(X-3)^k] = \int_{-\infty}^{\infty} c(x-3)^k e^{-(x-3)^2}dx = 0.$$

□

# Contents

### Proposition

(***Markov's inequality***) *If $X$ is a random variable that takes only nonnegative values, then for any value $\varepsilon > 0$,*

$$P(X \geqslant \varepsilon) \leqslant \frac{E(X)}{\varepsilon}.$$

**Proof.** We give a proof for the case where $X$ is continuous with density $f$.

$$\begin{aligned} E(X) &= \int_0^\infty xf(x)dx = \int_0^\varepsilon xf(x)dx + \int_\varepsilon^\infty xf(x)dx \\ &\geqslant \int_\varepsilon^\infty xf(x)dx \geqslant \int_\varepsilon^\infty \varepsilon f(x)dx = \varepsilon \int_\varepsilon^\infty f(x)dx = \varepsilon P(X \geqslant \varepsilon), \end{aligned}$$

and the result is proven. $\square$

# The Application of Expectation and Variation

## Proposition

(***Chebyshev's inequality***) *If $X$ is a random variable with mean $\mu$ and variance $\sigma^2$, then, for any value $\varepsilon > 0$,*

$$P(|X-\mu| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{\varepsilon^2}.$$

**Proof.** Since $(X-\mu)^2$ is a nonnegative random variable, we can apply Markov's inequality to obtain

$$P((X-\mu)^2 \geqslant \varepsilon^2) \leqslant \frac{E[(X-\mu)^2]}{\varepsilon^2}.$$

But since $(X-\mu)^2 \geqslant \varepsilon^2$ if and only if $|X-\mu| \geqslant \varepsilon$, the preceding is equivalent to

$$P(|X-\mu| \geqslant \varepsilon) \leqslant \frac{E[(X-\mu)^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

and the proof is complete. □

# The Application of Expectation and Variation

From Markov's inequality, we can see if $E(X)$ is small, then it is not too likely that $X$ is large.

Chebyshev's inequality tells us that if $\sigma^2 = Var(X)$ is small, then it is not too likely that $X$ is far from its mean.

The importance of Markov's and Chebyshev's inequalities is that they enable us to derive bounds on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known.

# The Application of Expectation and Variation

**Example**

Suppose that we know that the number of items produced in a factory during a week is a random variable with mean 500.

(a) What can be said about the probability that this week's production will be at least 1000?

(b) If the variance of a week's production is known to equal 100, then what can be said about the probability that this week's production will be between 400 and 600?

**Solution**. Let $X$ be the number of items that will be produced in a week.

(a) By Markov's inequality,

$$P(X \geqslant 1000) \leqslant \frac{E(X)}{1000} = \frac{500}{1000} = \frac{1}{2}.$$

(b) By Chebyshev's inequality,

$$P(|X - 500| \geqslant 100) \leqslant \frac{\sigma^2}{(100)^2} = \frac{1}{100}.$$

Hence,

$$P(|X - 500| \leqslant 100) \geqslant 1 - \frac{1}{100} = \frac{99}{100},$$

and so the probability that this week's production will be between 400 and 600 is at least 0.99. $\square$

**Thank you for your patience !**