Definition 12 Standard deviation If X is a random variable, the *standard deviation* of X, denoted by σ_X, is defined as $+\sqrt{\text{var}[X]}$. ////

The standard deviation of a random variable, like the variance, is a measure of the spread or dispersion of the values of the random variable. In many applications it is preferable to the variance as such a measure since it will have the same measurement units as the random variable itself.

EXAMPLE 14 Let X be the total of the two dice in the experiment of tossing two dice.

$$\begin{aligned}\text{var}[X] &= \sum (x_j - \mu_X)^2 f_X(x_j)\\ &= (2-7)^2\tfrac{1}{36} + (3-7)^2\tfrac{2}{36} + (4-7)^2\tfrac{3}{36} + (5-7)^2\tfrac{4}{36}\\ &\quad + (6-7)^2\tfrac{5}{36} + (7-7)^2\tfrac{6}{36} + (8-7)^2\tfrac{5}{36} + (9-7)^2\tfrac{4}{36}\\ &\quad + (10-7)^2\tfrac{3}{36} + (11-7)^2\tfrac{2}{36} + (12-7)^2\tfrac{1}{36} = \tfrac{210}{36}.\end{aligned}$$ ////

EXAMPLE 15 Let X be a random variable with probability density given by $f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$; then

$$\begin{aligned}\text{Var}[X] &= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx\\ &= \int_0^{\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x}\,dx\\ &\quad \cdot \quad \cdot \quad \cdot\\ &= \frac{1}{\lambda^2}.\end{aligned}$$ ////

EXAMPLE 16 Let X be a random variable with cumulative distribution given by $F_X(x) = (1 - pe^{-\lambda x}) I_{[0,\infty)}(x)$; then

$$\begin{aligned}\text{Var}[X] &= \int_0^{\infty} 2x[1 - F(x) + F(-x)]\,dx - \mu_X^2\\ &= \int_0^{\infty} 2xpe^{-\lambda x}\,dx - \left(\frac{p}{\lambda}\right)^2\\ &= 2\frac{p}{\lambda^2} - \left(\frac{p}{\lambda}\right)^2 = \frac{p(2-p)}{\lambda^2}.\end{aligned}$$ ////

4.3 Expected Value of a Function of a Random Variable

We defined the expectation of an arbitrary random variable X, called the mean of X, in Subsec. 4.1. In this subsection, we will define the expectation of a function of a random variable for discrete or continuous random variables.

Definition 13 Expectation Let X be a random variable and $g(\cdot)$ be a function with both domain and counterdomain the real line. The *expectation* or *expected value* of the function $g(\cdot)$ of the random variable X, denoted by $\mathcal{E}[g(X)]$, is defined by:

(i)
$$\mathcal{E}[g(X)] = \sum_j g(x_j) f_X(x_j) \tag{10}$$

if X is discrete with mass points $x_1, x_2, \ldots, x_j, \ldots$ (provided this series is absolutely convergent).

(ii)
$$\mathcal{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx \tag{11}$$

if X is continuous with probability density function $f_X(x)$ (provided $\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx < \infty$).* ////

Remark If $g(x) = x$, then $\mathcal{E}[g(X)] = \mathcal{E}[X]$ is the mean of X. If $g(x) = (x - \mu_X)^2$, then $\mathcal{E}[g(X)] = \mathcal{E}[(X - \mu_X)^2] = \text{var}[X]$. ////

Theorem 3 Below are properties of expected value:

(i) $\mathcal{E}[c] = c$ for a constant c.

(ii) $\mathcal{E}[cg(X)] = c\mathcal{E}[g(X)]$ for a constant c.

(iii) $\mathcal{E}[c_1 g_1(X) + c_2 g_2(X)] = c_1 \mathcal{E}[g_1(X)] + c_2 \mathcal{E}[g_2(X)]$.

(iv) $\mathcal{E}[g_1(X)] \le \mathcal{E}[g_2(X)]$ if $g_1(x) \le g_2(x)$ for all x.

Theorem 4 If X is a random variable, $\text{var}[X] = \mathcal{E}[(X - \mathcal{E}[X])^2] = \mathcal{E}[X^2] - (\mathcal{E}[X])^2$ provided $\mathcal{E}[X^2]$ exists.

Theorem 5 Let X be a random variable and $g(\cdot)$ a nonnegative function with domain the real line; then

$$P[g(X) \geq k] \leq \frac{\mathscr{E}[g(X)]}{k} \qquad \text{for every } k > 0. \tag{12}$$

PROOF Assume that X is a continuous random variable with probability density function $f_X(\cdot)$; then

$$\begin{aligned} \mathscr{E}[g(X)] &= \int_{-\infty}^{\infty} g(x) f_X(x)\,dx = \int_{\{x:\, g(x) \geq k\}} g(x) f_X(x)\,dx \\ &+ \int_{\{x:\, g(x) < k\}} g(x) f_X(x)\,dx \geq \int_{\{x:\, g(x) \geq k\}} g(x) f_X(x)\,dx \\ &\geq \int_{\{x:\, g(x) \geq k\}} k f_X(x)\,dx = kP[g(X) \geq k]. \end{aligned}$$

Divide by k, and the result follows. A similar proof holds for X discrete. ////

Corollary Chebyshev inequality If X is a random variable with finite variance,

$$P[|X - \mu_X| \geq r\sigma_X] = P[(X - \mu_X)^2 \geq r^2\sigma_X^2] \leq \frac{1}{r^2} \qquad \text{for every } r > 0. \tag{13}$$

PROOF Take $g(x) = (x - \mu_X)^2$ and $k = r^2\sigma_X^2$ in Eq. (12) of Theorem 5. ////