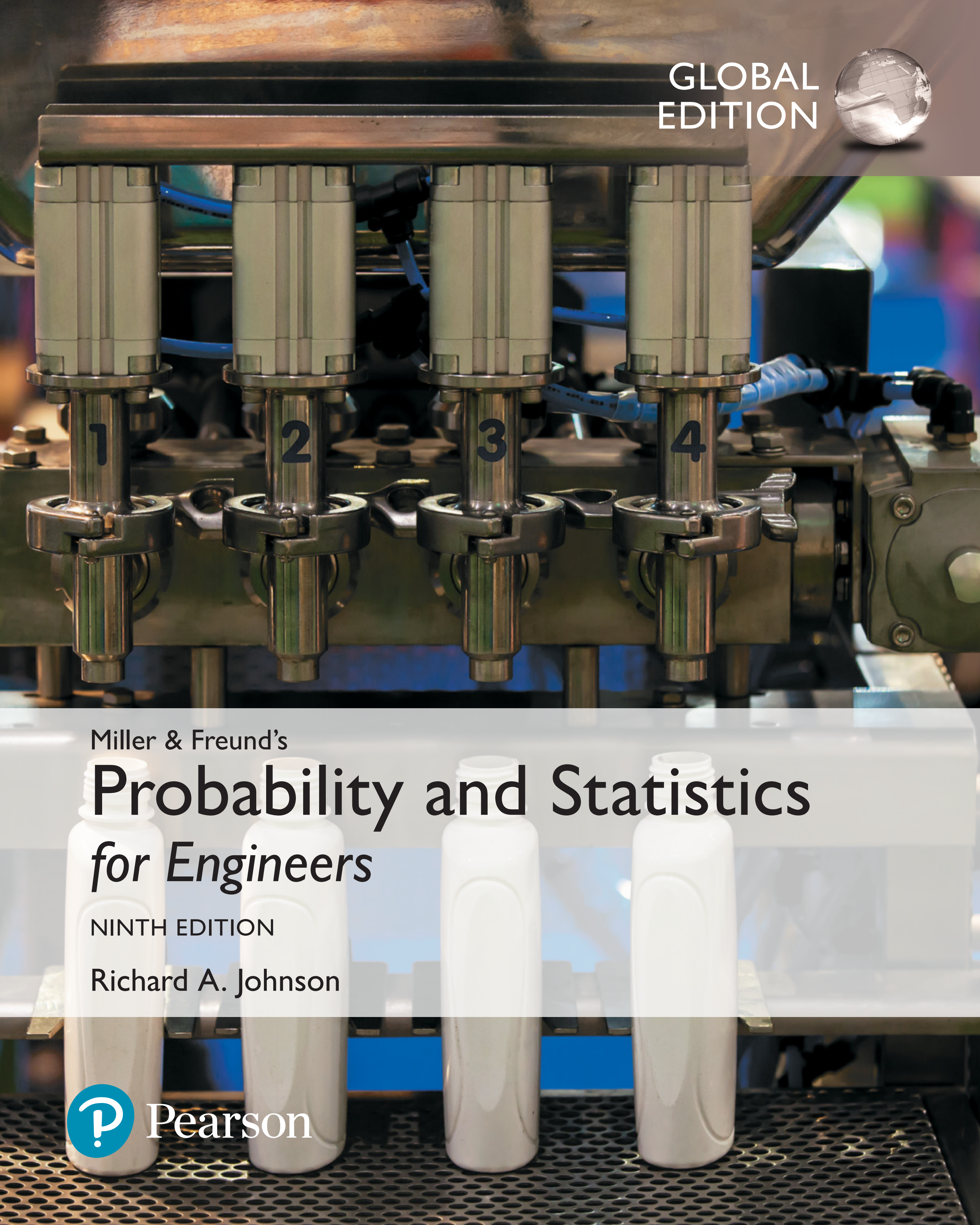

GLOBAL EDITION

Miller & Freund's

Probability and Statistics *for Engineers*

NINTH EDITION

Richard A. Johnson

Pearson

MILLER & FREUND'S

PROBABILITY AND STATISTICS FOR ENGINEERS

NINTH EDITION
Global Edition

Richard A. Johnson

University of Wisconsin–Madison

Pearson

Boston Columbus Indianapolis New York San Francisco Amsterdam
Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

EXAMPLE 23 **Determining a joint cumulative distribution function**

Find the joint cumulative distribution function of the two random variables of the preceding exercise, and use it to find the probability that both random variables will take on values less than 1.

Solution By definition,

$$F(x_1, x_2) = \begin{cases} \int_0^{x_2} \int_0^{x_1} 6\, e^{-2u-3v}\, du\, dv & \text{for } x_1 > 0,\ x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

so that

$$F(x_1, x_2) = \begin{cases} (1 - e^{-2x_1})(1 - e^{-3x_2}) & \text{for } x_1 > 0,\ x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and, hence,

$$\begin{aligned} F(1, 1) &= (1 - e^{-2})(1 - e^{-3}) \\ &= 0.8216 \end{aligned}$$

■

Given the joint probability density of k random variables, the probability density of the ith random variable can be obtained by integrating out the other variables; symbolically,

Marginal density

$$f_i(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_k)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_k$$

and, in this context, the function f_i is called the **marginal density** of the ith random variable. Integrating out only some of the k random variables, we can similarly define **joint marginal densities** of any two, three, or more of the k random variables.

EXAMPLE 24 **Determining a marginal density from a joint density**

With reference to Example 22. find the marginal density of the first random variable.

Solution Integrating out x_2, we get

$$f_1(x_1) = \begin{cases} \int_0^{\infty} 6\, e^{-2x_1-3x_2}\, dx_2 & \text{for } x_1 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

or

$$f_1(x_1) = \begin{cases} 2\, e^{-2x_1} & \text{for } x_1 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

■

To explain what we mean by the **independence** of continuous random variables, we could proceed as with discrete random variables and define *conditional probability densities* first; however, it will be easier to say that

Independent random variables

> k random variables $X_1, \ldots, X_k$ are **independent** if and only if
>
> $$F(x_1, x_2, \ldots, x_k) = F_1(x_1) \cdot F_2(x_2) \cdots F_k(x_k)$$
>
> for all values $x_1, x_2, \ldots, x_k$ of these random variables.

In this notation $F(x_1, x_2, \ldots, x_k)$ is, as before, the joint distribution function of the k random variables, while $F_i(x_i)$ for $i = 1, 2, \ldots, k$ are the corresponding individual distribution function of the respective random variables. The same condition applies for discrete random variables.

EXAMPLE 25 **Checking independence via the joint cumulative distribution**

With reference to Example 23, check whether the two random variables are independent.

Solution As we already saw in Example 23, the joint distribution function of the two random variables is given by

$$F(x_1, x_2) = \begin{cases} (1 - e^{-2x_1})(1 - e^{-3x_2}) & \text{for } x_1 > 0 \text{ and } x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Now, since $F_1(x_1) = F(x_1, \infty)$ and $F_2(x_2) = F(\infty, x_2)$, it follows that

$$F_1(x_1) = \begin{cases} 1 - e^{-2x_1} & \text{for } x_1 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$F_2(x_2) = \begin{cases} 1 - e^{-3x_2} & \text{for } x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Thus, $F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$ for all (x_1, x_2) and the two random variables are independent. ■

When k random variables have a joint probability density, the k **random variables are independent** if and only if their joint probability density equals the product of the corresponding values of the marginal densities of the k random variables; symbolically,

$$f(x_1, x_2, \ldots, x_k) = f_1(x_1) \cdot f_2(x_2) \cdots f_k(x_k) \qquad \text{for all } (x_1, \ldots, x_k).$$

EXAMPLE 26 **Establishing independence by factoring the joint probability density**

With reference to Example 22, verify that

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

Solution Example 24 shows that

$$f_1(x_1) = \begin{cases} 2\,e^{-2x_1} & \text{for } x_1 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and in the same way,

$$f_2(x_2) = \begin{cases} 3\,e^{-3x_2} & \text{for } x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Thus,

$$f_1(x_1) \cdot f_2(x_2) = \begin{cases} 6\,e^{-2x_1-3x_2} & \text{for } x_1 > 0 \text{ and } x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and it can be seen that $f_1(x_1) \cdot f_2(x_2) = f(x_1, x_2)$ for all (x_1, x_2). ■

Given two continuous random variables X_1 and X_2, we define the **conditional probability density** of the first given that the second takes on the value x_2 as

Conditional probability density

$$f_1(\,x_1\,|\,x_2\,) = \frac{f(\,x_1, x_2\,)}{f_2(\,x_2\,)} \quad \text{provided } f_2(\,x_2\,) \neq 0$$

where $f(x_1, x_2)$ and $f_2(x_2)$ are, as before, the joint density of the two random variables and the marginal density of the second. Note that this definition parallels that of the conditional probability distribution on page 163. Also, the joint probability density is the product

$$f\,(\,x_1, x_2\,) \;=\; f_1\,(\,x_1\,|\;x_2\,)\,f_2\,(\,x_2\,).$$

EXAMPLE 27

Determining a conditional probability density

If two random variables have the joint probability density

$$f(x_1, x_2) = \begin{cases} \dfrac{2}{3}\,(\,x_1 + 2x_2\,) & \text{for } 0 < x_1 < 1,\, 0 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

find the conditional density of the first given that the second takes on the value x_2.

Solution First we find the marginal density of the second random variable by integrating out x_1, and we get

$$f_2(x_2) = \int_0^1 \frac{2}{3}\,(\,x_1 + 2\,x_2\,)\,dx_1 \;=\; \frac{1}{3}\,(\,1 + 4\,x_2\,) \qquad \text{for } 0 < x_2 < 1$$

and $f_2(x_2) = 0$ elsewhere. Hence, by definition, the conditional density of the first random variable given that the second takes on the value x_2 is given by

$$f_1(x_1\,|\,x_2) = \frac{\dfrac{2}{3}\,(\,x_1 + 2\,x_2\,)}{\dfrac{1}{3}\,(\,1 + 4\,x_2\,)} = \frac{2\,x_1 + 4\,x_2}{1 + 4\,x_2} \qquad \text{for } 0 < x_1 < 1,\, 0 < x_2 < 1$$

and $f_1(x_1\,|\,x_2) = 0$ for $x_1 \leq 0$ or $x_1 \geq 1$ and $0 < x_2 < 1$. ■

Properties of Expectation

Consider a function $g(X)$ of a single random variable X. For instance, if X is an oven temperature in degrees centigrade, then

$$g(X) = \frac{9}{5}X + 32$$

is the same temperature in degrees Fahrenheit.

The **expectation** of the function $g(X)$ is again the *sum* of the products *value* × *probability*.

Expected value of g(X)

> In the discrete case, where X has probability distribution $f(x)$
>
> $$E[g(X)] = \sum_{x_i} g(x_i) f(x_i)$$
>
> In the continuous case, where X has probability density function $f(x)$
>
> $$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

If X has mean $\mu = E(X)$, then taking $g(x) = (x - \mu)^2$, we have $E[g(X)] = E(X - \mu)^2$, which is just the variance σ^2 of X.

For any random variable Y, let $E(Y)$ denote its expectation, which is also its mean μ_Y. Its variance is $Var(Y)$ which is also written as σ_Y^2.

When $g(x) = ax + b$, for given constants a and b, then random variable $g(X)$ has expectation

$$\begin{aligned} E(aX + b) &= \int_{-\infty}^{\infty} (ax + b) f(x)\, dx = a \int_{-\infty}^{\infty} x f(x)\, dx + b \int_{-\infty}^{\infty} f(x)\, dx \\ &= aE(X) + b \end{aligned}$$

and variance

$$\begin{aligned} Var(aX + b) &= \int_{-\infty}^{\infty} (ax + b - a\mu_X - b)^2 f(x)\, dx \\ &= a^2 \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)\, dx = a^2 Var(X) \end{aligned}$$

To summarize,

> For given constants a and b
>
> $$E(aX + b) = aE(X) + b \quad \text{and} \quad Var(aX + b) = a^2 Var(X)$$

EXAMPLE 28 **The mean and standard deviation of a standardized random variable**

Let X have mean μ and standard deviation σ. Use the properties of expectation to show that the standardized random variable

$$Z = \frac{X - \mu}{\sigma}$$

has mean 0 and standard deviation 1.

Solution Since Z is of the form

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} = aX + b$$

where $a = 1/\sigma$ and $b = -\mu/\sigma$,

$$E(Z) = \frac{1}{\sigma} E(X) - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

and the variance of Z is

$$\left(\frac{1}{\sigma}\right)^2 Var(X) = \frac{\sigma^2}{\sigma^2} = 1$$

because $E(X) = \mu$ and $Var(X) = \sigma^2$. ■

EXAMPLE 29

Determining the mean and variance of 20 *X*

Suppose the daily amount of electricity X required for a plating process has mean 10 and standard deviation 3 kilowatt-hours. If the cost of electricity is 20 dollars per kilowatt hour, find the mean, variance, and standard deviation of the daily cost of electricity.

Solution The daily cost of electricity, $g(X) = 20X$, has mean $20\,E(X) = 20 \times 10 = 200$ dollars and variance $(20)^2\,Var(X) = (20)^2\,3^2 = 3{,}600$. Its standard deviation is $\sqrt{3{,}600} = 60$ dollars. ■

Given any collection of k random variables, the function $Y = g(X_1, X_2, \ldots, X_k)$ is also a random variable. Examples include $Y = X_1 - X_2$ when $g(x_1, x_2) = x_1 - x_2$ and $Y = 2X_1 + 3X_2$ when $g(x_1, x_2) = 2x_1 + 3x_2$. The random variable $g(X_1, X_2, \ldots, X_k)$ has expected value, or mean, which is the sum of the products *value* × *probability*.

Expected value of $g(X_1, X_2, \ldots, X_k)$

In the discrete case,

$$E[g(X_1, X_2, \ldots, X_k)] = \sum_{x_1}\sum_{x_2}\cdots\sum_{x_k} g(x_1, x_2, \ldots, x_k) f(x_1, x_2, \ldots, x_k)$$

In the continuous case,

$$E[g(X_1, X_2, \ldots, X_k)]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_k) f(x_1, x_2, \ldots, x_k)\, dx_1\, dx_2 \cdots dx_k$$

Several important properties of expectation can be deduced from this definition. Taking $g(x_1, x_2) = (x_1 - \mu_1)(x_2 - \mu_2)$, we see that the product $(x_1 - \mu_1)(x_2 - \mu_2)$ will be positive if both values x_1 and x_2 are above their respective means or both are below their respective means. Otherwise it will be negative. The expected value $E[(X_1 - \mu_1)(X_2 - \mu_2)]$ will tend to be positive when large X_1 and X_2 tend to occur together and small X_1 and X_2 tend to occur together, with high probability. This measure $E[(X_1 - \mu_1)(X_2 - \mu_2)]$ of joint variation is called the population **covariance** of X_1 and X_2.

If X_1 and X_2 are independent so $f(x_1, x_2) = f_1(x_1) f_2(x_2)$,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - \mu_1)(x_2 - \mu_2) f(x_1, x_2)\, dx_1\, dx_2$$
$$= \int_{-\infty}^{\infty} (x_1 - \mu_1) f_1(x_1)\, dx_1 \cdot \int_{-\infty}^{\infty} (x_2 - \mu_2) f_2(x_2)\, dx_2 = 0$$

This result concerning zero covariance can be stated as

Independence implies that the covariance is zero

When X_1 and X_2 are independent, their covariance

$$E[(X_1 - \mu_1)(X_2 - \mu_2)] = 0$$

Further, the expectation of a linear combination of two independent random variables $Y = a_1X_1 + a_2X_2$ is

$$\begin{aligned}
\mu_Y &= E(Y) = E(a_1X_1 + a_2X_2) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a_1x_1 + a_2x_2)f_1(x)f_2(x_2)\,dx_1\,dx_2 \\
&= a_1\int_{-\infty}^{\infty}x_1f_1(x_1)\,dx_1\int_{-\infty}^{\infty}f_2(x_2)\,dx_2 \\
&\quad + a_2\int_{-\infty}^{\infty}f_1(x_1)\,dx_1\int_{-\infty}^{\infty}x_2f_2(x_2)\,dx_2 \\
&= a_1E(X_1) + a_2E(X_2)
\end{aligned}$$

This result holds even if the two random variables are not independent. Also,

$$\begin{aligned}
Var(Y) &= E(Y - \mu_Y)^2 = E[(a_1X_1 + a_2X_2 - a_1\mu_1 - a_2\mu_2)^2] \\
&= E[(a_1(X_1 - \mu_1) + a_2(X_2 - \mu_2))^2] \\
&= E[a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)] \\
&= a_1^2E[(X_1 - \mu_1)^2] + a_2^2E[(X_2 - \mu_2)^2] + 2a_1a_2E[(X_1 - \mu_1)(X_2 - \mu_2)] \\
&= a_1^2\,Var(X_1) + a_2^2\,Var(X_2)
\end{aligned}$$

since the third term is zero because we assumed X_1 and X_2 are independent.

These properties hold for any number of random variables whether they are continuous or discrete.

The mean and variance of linear combinations

Let X_i have mean μ_i and variance σ_i^2 for $i = 1, 2, \ldots, k$. The linear combination $Y = a_1X_1 + a_2X_2 + \cdots + a_kX_k$ has

$$E(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = a_1E(X_1) + a_2E(X_2) + \cdots + a_kE(X_k)$$

or

$$\mu_Y = \sum_{i=1}^{k} a_i\,\mu_i$$

When the random variables are independent,

$$\begin{aligned}
Var(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = {} & a_1^2\,Var(X_1) \\
& + a_2^2\,Var(X_2) + \cdots + a_k^2\,Var(X_k)
\end{aligned}$$

or

$$\sigma_Y^2 = \sum_{i=1}^{k} a_i^2\,\sigma_i^2$$