# Section 6.3 Sums of Independent Random Variables

It is often important to be able to calculate the distribution of $X + Y$ from the distributions of $X$ and $Y$ when $X$ and $Y$ are independent. Suppose that $X$ and $Y$ are independent, continuous random variables having probability density functions $f_X$ and $f_Y$. The cumulative distribution function of $X + Y$ is obtained as follows:

$$
\begin{aligned}
F_{X+Y}(a) &= P\{X + Y \leq a\} \\
&= \iint_{x+y\leq a} f_X(x) f_Y(y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x) f_Y(y)\, dx\, dy \qquad (3.1) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)\, dx\, f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\, dy
\end{aligned}
$$

The cumulative distribution function $F_{X+Y}$ is called the *convolution* of the distributions $F_X$ and $F_Y$ (the cumulative distribution functions of $X$ and $Y$, respectively).

By differentiating Equation (3.1), we obtain that the probability density function $f_{X+Y}$ of $X + Y$ is given by

$$
\begin{aligned}
f_{X+Y}(a) &= \frac{d}{da} \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} \frac{d}{da} F_X(a - y) f_Y(y)\, dy \qquad (3.2) \\
&= \int_{-\infty}^{\infty} f_X(a - y) f_Y(y)\, dy
\end{aligned}
$$

**Example 3a.** ***Sum of two independent uniform random variables.*** If $X$ and $Y$ are independent random variables, both uniformly distributed on $(0, 1)$, calculate the probability density of $X + Y$.

**Solution** From Equation (3.2), since

$$
f_X(a) = f_Y(a) = \begin{cases} 1 & 0 < a < 1 \\ 0 & \text{otherwise} \end{cases}
$$

we obtain

$$
f_{X+Y}(a) = \int_0^1 f_X(a - y)\, dy
$$

For $0 \leq a \leq 1$, this yields

$$
f_{X+Y}(a) = \int_0^a dy = a
$$

For $1 < a < 2$, we get

$$
f_{X+Y}(a) = \int_{a-1}^1 dy = 2 - a
$$

Hence

$$f_{X+Y}(a) = \begin{cases} a & 0 \le a \le 1 \\ 2 - a & 1 < a < 2 \\ 0 & \text{otherwise} \end{cases}$$

Because of the shape of its density function (see Figure 6.3), the random variable $X + Y$ is said to have a *triangular* distribution. ∎

***Proposition 3.2***

If $X_i$, $i = 1, \ldots, n$, are independent random variables that are normally distributed with respective parameters $\mu_i$, $\sigma_i^2$, $i = 1, \ldots, n$, then $\sum_{i=1}^{n} X_i$ is normally distributed with parameters $\sum_{i=1}^{n} \mu_i$ and $\sum_{i=1}^{n} \sigma_i^2$.

**Proof of Proposition 3.2:** To begin, let $X$ and $Y$ be independent normal random variables, with $X$ having mean 0 and variance $\sigma^2$, and $Y$ having mean 0 and variance 1. We will determine the density function of $X + Y$ by utilizing Equation (3.2). Now, with

$$c = \frac{1}{2\sigma^2} + \frac{1}{2} = \frac{1 + \sigma^2}{2\sigma^2}$$

we have

$$\begin{aligned} f_X(a - y) f_Y(y) &= \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(a-y)^2}{2\sigma^2}\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} \\ &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{a^2}{2\sigma^2}\right\} \exp\left\{-c\left(y^2 - 2y\frac{a}{1+\sigma^2}\right)\right\} \end{aligned}$$

Hence, from Equation (3.2),

$$\begin{aligned} f_{X+Y}(a) &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{a^2}{2\sigma^2}\right\} \exp\left\{\frac{a^2}{2\sigma^2(1+\sigma^2)}\right\} \int_{-\infty}^{\infty} \exp\left\{-c\left(y - \frac{a}{1+\sigma^2}\right)^2\right\} dy \\ &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{a^2}{2(1+\sigma^2)}\right\} \int_{-\infty}^{\infty} \exp\{-cx^2\}\, dx \\ &= C \exp\left\{-\frac{a^2}{2(1+\sigma^2)}\right\} \end{aligned}$$

where $C$ doesn't depend on $a$. But this implies that $X + Y$ is normal with mean 0 and variance $1 + \sigma^2$.

Now, suppose that $X_1$ and $X_2$ are independent normal random variables, with $X_i$ having mean $\mu_i$ and variance $\sigma_i^2$, $i = 1, 2$. Then

$$X_1 + X_2 = \sigma_2\left(\frac{X_1 - \mu_1}{\sigma_2} + \frac{X_2 - \mu_2}{\sigma_2}\right) + \mu_1 + \mu_2$$

But since $(X_1 - \mu_1)/\sigma_2$ is normal with mean 0 and variance $\sigma_1^2/\sigma_2^2$, and $(X_2 - \mu_2)/\sigma_2$ is normal with mean 0 and variance 1, it follows from our previous result that $(X_1 - \mu_1)/\sigma_2 + (X_2 - \mu_2)/\sigma_2$ is normal with mean 0 and variance $1 + \sigma_1^2/\sigma_2^2$, implying that $X_1 + X_2$ is normal with mean $\mu_1 + \mu_2$ and variance $\sigma_2^2(1 + \sigma_1^2/\sigma_2^2) = \sigma_1^2 + \sigma_2^2$.

**Example 3c.** A club basketball team will play a 44-game season. Twenty-six of these games are against class A teams and 18 are against class B teams. Suppose that the team will win each game against a class A team with probability .4, and will win each game against a class B team with probability .7. Assume also that the results of the different games are independent. Approximate the probability that

(a) the team wins 25 games or more;

(b) the team wins more games against class A teams than it does against class B teams.

**Solution** (a) Let $X_A$ and $X_B$ denote, respectively, the number of games the team wins against class A and against class B teams. Note that $X_A$ and $X_B$ are independent binomial random variables, and

$$E[X_A] = 26(.4) = 10.4 \qquad \text{Var}(X_A) = 26(.4)(.6) = 6.24$$
$$E[X_B] = 18(.7) = 12.6 \qquad \text{Var}(X_B) = 18(.7)(.3) = 3.78$$

By the normal approximation to the binomial it follows that $X_A$ and $X_B$ will approximately have the same distribution as would independent normal random variables with expected values and variances as given in the preceding. Hence, by Proposition 3.2, $X_A + X_B$ will approximately have a normal distribution with mean 23 and variance 10.02. Therefore, letting $Z$ denote a standard normal random variable, we have

$$\begin{aligned} P\{X_A + X_B \geq 25\} &= P\{X_A + X_B \geq 24.5\} \\ &= P\left\{\frac{X_A + X_B - 23}{\sqrt{10.02}} \geq \frac{24.5 - 23}{\sqrt{10.02}}\right\} \\ &\approx P\left\{Z \geq \frac{1.5}{\sqrt{10.02}}\right\} \\ &\approx 1 - P\{Z < .4739\} \\ &\approx .3178 \end{aligned}$$

(b) We note that $X_A - X_B$ will approximately have a normal distribution with mean $-2.2$ and variance 10.02. Hence

$$\begin{aligned} P\{X_A - X_B \geq 1\} &= P\{X_A - X_B \geq .5\} \\ &= P\left\{\frac{X_A - X_B + 2.2}{\sqrt{10.02}} \geq \frac{.5 + 2.2}{\sqrt{10.02}}\right\} \\ &\approx P\left\{Z \geq \frac{2.7}{\sqrt{10.02}}\right\} \\ &\approx 1 - P\{Z < .8530\} \\ &\approx .1968 \end{aligned}$$

Therefore, there is approximately a 31.78 percent chance that the team will win at least 25 games, and approximately a 19.68 percent chance that it will win more games against class A teams than against class B teams. ∎