**Theorem 1.1** If $X_1, X_2, \ldots, X_n$ is a random sample from a $N\left(\mu, \sigma^2\right)$ population, then

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

**Proof** Consider mutually independent random variables $X_1, X_2, \ldots, X_n$, where $X_i \sim N(\mu_i, \sigma_i^2)$, for $i = 1, 2, \ldots, n$, and nonzero real constants $a_1, a_2, \ldots, a_n$. From probability theory, the linear combination $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ is normally distributed with population mean

$$\mu_Y = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n$$

and population variance

$$\sigma_Y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2.$$

The present result is a special case of this result from probability theory in which

$$\mu_1 = \mu_2 = \cdots = \mu_n = \mu,$$

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2,$$

and

$$a_1 = a_2 = \cdots = a_n = \frac{1}{n}.$$ □

Sometimes the parameters for the population are written as $\mu_X$ and $\sigma_X^2$. Likewise, sometimes the parameters associated with the sample mean can be written as $\mu_{\bar{X}} = \mu_X$ and $\sigma_{\bar{X}}^2 = \sigma_X^2/n$. The conclusion to Theorem 1.1 can also be written in normalized form, that is

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

In order to illustrate the application of this theorem, we return to the elevator problem posed earlier.

**Example 1.30** Ten people crowd into a small elevator. Their weights are mutually independent and identically distributed $N(140, 900)$ random variables. What is the probability that the elevator capacity of 1500 pounds is exceeded?

Let $X_1, X_2, \ldots, X_{10}$ denote the weights of the ten people. The probability that the elevator capacity is exceeded is

$$P\left(X_1 + X_2 + \cdots + X_{10} > 1500\right).$$

Since Theorem 1.1 concerns the sample mean, this probability can be rewritten in terms of $\bar{X}$ by dividing both sides of the inequality by $n = 10$, which gives

$$P\left(\bar{X} > 150\right).$$

Since the population mean is $\mu = 140$ and the population standard deviation is $\sigma = 30$, Theorem 1.1 can now be invoked by normalizing the left-hand side of the inequality:

$$\begin{aligned} P(\bar{X} > 150) &= P\left(\frac{\bar{X} - 140}{30/\sqrt{10}} > \frac{150 - 140}{30/\sqrt{10}}\right) \\ &= P\left(Z > \frac{\sqrt{10}}{3}\right) \\ &\cong P(Z > 1.0541) \\ &\cong 0.1459. \end{aligned}$$

This probability can be calculated in R with the statement

```
1 - pnorm(1.0541)
```

or you can leave the normalizing to R and use

```
1 - pnorm(150, 140, 30 / sqrt(10))
```

Figure 1.42 shows the population distribution, which is $N(140, 900)$, and the sampling distribution of the sample mean $\bar{X}$, which is $N(140, 90)$. As in the previous examples, the effect of computing a sample mean is to maintain the same mean value, but decrease the variance. The area under the probability density function of $\bar{X}$ to the right of 150, which is $P(\bar{X} > 150) \cong 0.1459$, is shaded. The R commands

```
x  = 80:200
y1 = dnorm(x, 140, 30)
y2 = dnorm(x, 140, 30 / sqrt(10))
matplot(x, cbind(y1, y2), type = "l")
```

generate the graphs of the two probability density functions. The `matplot` function plots two curves on the same set of axes. The column bind (`cbind`) function pastes the two vectors together, treating each as a column.

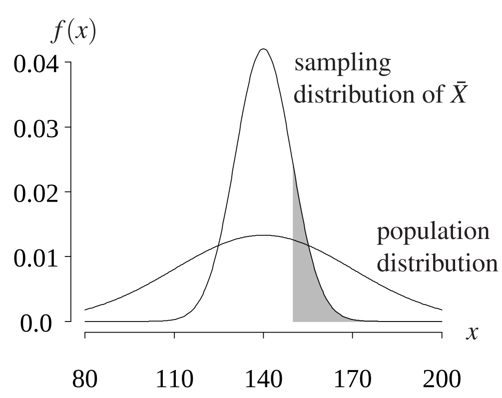


Figure 1.42: Population distribution and sampling distribution of $\bar{X}$.

**Theorem 1.2** If $X_1, X_2, \ldots, X_n$ is a random sample from a $N\left(\mu, \sigma^2\right)$ population, then

$$\sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n).$$

**Proof** First, standardize the normal random variable $X_i$:

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1),$$

for $i = 1, 2, \ldots, n$. Since the square of a standard normal random variable has the chi-square distribution with one degree of freedom,

$$\left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(1),$$

for $i = 1, 2, \ldots, n$. Finally, since the sum of $n$ mutually independent chi-square random variables is also chi-square (with degrees of freedom equal to the sum of the degrees of freedom of the component chi-square random variables),

$$\sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n),$$

which proves the result. □

**Example 1.31** Ten people crowd into a small elevator. Denote their mutually independent and identically distributed weights, which are $N(140, 900)$ random variables, by $X_1, X_2, \ldots, X_{10}$. Find a constant $c$ such that

$$P\left( \sum_{i=1}^{10} \left( \frac{X_i - 140}{30} \right)^2 < c \right) = 0.99.$$

Invoking Theorem 1.2, the random variable

$$\sum_{i=1}^{10} \left( \frac{X_i - 140}{30} \right)^2 \sim \chi^2(10).$$

Therefore,

$$P\left( \sum_{i=1}^{10} \left( \frac{X_i - 140}{30} \right)^2 < c \right) = P\left(\chi^2(10) < c\right) = 0.99.$$

The value of the constant $c$ can be calculated with the R statement

```
qchisq(0.99, 10)
```

which returns $c \cong 23.2093$.