Sampling Distributions

1 Introduction

Statistics concerns itself mainly with conclusions and predictions resulting from chance outcomes that occur in carefully planned experiments or investigations. Drawing such conclusions usually involves taking sample observations from a given population and using the results of the sample to make inferences about the population itself, its mean, its variance, and so forth. To do this requires that we first find the distributions of certain functions of the random variables whose values make up the sample, called **statistics**. (An example of such a statistic is the sample mean.) The properties of these distributions then allow us to make probability statements about the resulting inferences drawn from the sample about the population. The theory to be given in this chapter forms an important foundation for the theory of statistical inference.

Inasmuch as statistical inference can be loosely defined as a process of drawing conclusions from a sample about the population from which it is drawn, it is useful to have the following definition.

> **Definition 1. Population.** *A set of numbers from which a sample is drawn is referred to as a **population**. The distribution of the numbers constituting a population is called the **population distribution**.*

To illustrate, suppose a scientist must choose and then weigh 5 of 40 guinea pigs as part of an experiment, a layman might say that the ones she selects constitute the sample. This is how the term "sample" is used in everyday language. In statistics, it is preferable to look upon the weights of the 5 guinea pigs as a sample from the population, which consists of the weights of all 40 guinea pigs. In this way, the population as well as the sample consists of numbers. Also, suppose that, to estimate the average useful life of a certain kind of transistor, an engineer selects 10 of these transistors, tests them over a period of time, and records for each one the time to failure. If these times to failure are values of independent random variables having an exponential distribution with the parameter θ, we say that they constitute a sample from this exponential population.

As can well be imagined, not all samples lend themselves to valid generalizations about the populations from which they came. In fact, most of the methods of inference discussed in this chapter are based on the assumption that we are dealing with

From Chapter 8 of *John E. Freund's Mathematical Statistics with Applications*, Eighth Edition. Irwin Miller, Marylees Miller.

random samples. In practice, we often deal with random samples from populations that are finite, but large enough to be treated as if they were infinite. Thus, most statistical theory and most of the methods we shall discuss apply to samples from infinite populations, and we shall begin here with a definition of random samples from infinite populations. Random samples from finite populations will be treated later in Section 3.

DEFINITION 2. RANDOM SAMPLE. *If* $X_1, X_2, \ldots, X_n$ *are independent and identically distributed random variables, we say that they constitute a* ***random sample*** *from the infinite population given by their common distribution.*

If $f(x_1, x_2, \ldots, x_n)$ is the value of the joint distribution of such a set of random variables at $(x_1, x_2, \ldots, x_n)$, by virtue of independence we can write

$$f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f(x_i)$$

where $f(x_i)$ is the value of the population distribution at x_i. Observe that Definition 2 and the subsequent discussion apply also to sampling with replacement from finite populations; sampling without replacement from finite populations is discussed in section 3.

Statistical inferences are usually based on **statistics**, that is, on random variables that are functions of a set of random variables $X_1, X_2, \ldots, X_n$, constituting a random sample. Typical of what we mean by "statistic" are the **sample mean** and the **sample variance**.

DEFINITION 3. SAMPLE MEAN AND SAMPLE VARIANCE. *If* $X_1, X_2, \ldots, X_n$ *constitute a random sample, then the* ***sample mean*** *is given by*

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

and the ***sample variance*** *is given by*

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

As they are given here, these definitions apply only to random samples, but the sample mean and the sample variance can, similarly, be defined for any set of random variables $X_1, X_2, \ldots, X_n$.

It is common practice also to apply the terms "random sample," "statistic," "sample mean," and "sample variance" to the values of the random variables instead of the random variables themselves. Intuitively, this makes more sense and it conforms with colloquial usage. Thus, we might calculate

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$$

for observed sample data and refer to these statistics as the sample mean and the sample variance. Here, the x_i, $\overline{x}$, and s^2 are values of the corresponding random

†The note has been intentionally omitted for this edition.

variables X_i, $\overline{X}$, and S^2. Indeed, the formulas for $\bar{x}$ and s^2 are used even when we deal with any kind of data, not necessarily sample data, in which case we refer to $\bar{x}$ and s^2 simply as the mean and the variance.

These, and other statistics that will be introduced in this chapter, are those mainly used in statistical inference. Sample statistics such as the sample mean and sample variance play an important role in estimating the parameters of the population from which the corresponding random samples were drawn.

2 The Sampling Distribution of the Mean

Inasmuch as the values of sampling statistics can be expected to vary from sample to sample, it is necessary that we find the distribution of such statistics. We call these distributions **sampling distributions**, and we make important use of them in determining the properties of the inferences we draw from the sample about the parameters of the population from which it is drawn.

First let us study some theory about the **sampling distribution of the mean**, making only some very general assumptions about the nature of the populations sampled.

THEOREM 1. If $X_1, X_2, \ldots, X_n$ constitute a random sample from an infinite population with the mean μ and the variance σ^2, then

$$E(\overline{X}) = \mu \quad \text{and} \quad \text{var}(\overline{X}) = \frac{\sigma^2}{n}$$

Proof Letting $Y = \overline{X}$ and hence setting $a_i = \frac{1}{n}$, we get

$$E(\overline{X}) = \sum_{i=1}^{n} \frac{1}{n} \cdot \mu = n\left(\frac{1}{n} \cdot \mu\right) = \mu$$

since $E(X_i) = \mu$. Then, by the corollary of a theorem "If the random variables $X_1, X_2, \ldots, X_n$ are independent and $Y = \sum_{i=1}^{n} a_i X_i$, then $\text{var}(Y) = \sum_{i=1}^{n} a_i^2 \cdot \text{var}(X_i)$", we conclude that

$$\text{var}(\overline{X}) = \sum_{i=1}^{n} \frac{1}{n^2} \cdot \sigma^2 = n\left(\frac{1}{n^2} \cdot \sigma^2\right) = \frac{\sigma^2}{n}$$

It is customary to write $E(\overline{X})$ as $\mu_{\overline{X}}$ and $\text{var}(\overline{X})$ as $\sigma^2_{\overline{X}}$ and refer to $\sigma_{\overline{X}}$ as the **standard error of the mean**. The formula for the standard error of the mean, $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$, shows that the standard deviation of the distribution of $\overline{X}$ decreases when n, the **sample size**, is increased. This means that when n becomes larger and we actually have more information (the values of more random variables), we can expect values of $\overline{X}$ to be closer to μ, the quantity that they are intended to estimate.

THEOREM 2. For any positive constant c, the probability that $\overline{X}$ will take on a value between $\mu - c$ and $\mu + c$ is at least

$$1 - \frac{\sigma^2}{nc^2}$$

When $n \to \infty$, this probability approaches 1.

This result, called a **law of large numbers**, is primarily of theoretical interest. Of much more practical value is the **central limit theorem**, one of the most important theorems of statistics, which concerns the limiting distribution of the **standardized mean** of n random variables when $n \to \infty$. We shall prove this theorem here only for the case where the n random variables are a random sample from a population whose moment-generating function exists. More general conditions under which the theorem holds are given in Exercises 7 and 9, and the most general conditions under which it holds are referred to at the end of this chapter.

THEOREM 3. CENTRAL LIMIT THEOREM. If $X_1, X_2, \ldots, X_n$ constitute a random sample from an infinite population with the mean μ, the variance σ^2, and the moment-generating function $M_X(t)$, then the limiting distribution of

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

as $n \to \infty$ is the standard normal distribution.

Proof First using the third part and then the second of the given theorem "If a and b are constants, then **1.** $M_{X+a}(t) = E[e^{(X+a)t}] = e^{at} \cdot M_X(t)$; **2.** $M_{bX}(t) = E(e^{bXt}) = M_X(bt)$; **3.** $M_{\frac{X+a}{b}}(t) = E[e^{\left(\frac{X+a}{b}\right)t}] = e^{\frac{a}{b}t} \cdot M_X\left(\frac{t}{b}\right)$", we get

$$M_Z(t) = M_{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}(t) = e^{-\sqrt{n}\,\mu t/\sigma} \cdot M_{\overline{X}}\left(\frac{\sqrt{n}t}{\sigma}\right)$$

$$= e^{-\sqrt{n}\,\mu t/\sigma} \cdot M_{n\overline{X}}\left(\frac{t}{\sigma\sqrt{n}}\right)$$

Since $n\overline{X} = X_1 + X_2 + \cdots + X_n$,

$$M_Z(t) = e^{-\sqrt{n}\,\mu t/\sigma} \cdot \left[M_X\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

and hence that

$$\ln M_Z(t) = -\frac{\sqrt{n}\,\mu t}{\sigma} + n \cdot \ln M_X\left(\frac{t}{\sigma\sqrt{n}}\right)$$

Expanding $M_X\left(\frac{t}{\sigma\sqrt{n}}\right)$ as a power series in t, we obtain

$$\ln M_Z(t) = -\frac{\sqrt{n}\,\mu t}{\sigma} + n \cdot \ln\left[1 + \mu_1' \frac{t}{\sigma\sqrt{n}} + \mu_2' \frac{t^2}{2\sigma^2 n} + \mu_3' \frac{t^3}{6\sigma^3 n\sqrt{n}} + \cdots\right]$$

where μ_1', μ_2', and μ_3' are the moments about the origin of the population distribution, that is, those of the original random variables X_i.

If n is sufficiently large, we can use the expansion of $\ln(1+x)$ as a power series in x, getting

$$\ln M_Z(t) = -\frac{\sqrt{n}\,\mu t}{\sigma} + n\left\{\left[\mu_1'\frac{t}{\sigma\sqrt{n}} + \mu_2'\frac{t^2}{2\sigma^2 n} + \mu_3'\frac{t^3}{6\sigma^3 n\sqrt{n}} + \cdots\right]\right.$$

$$-\frac{1}{2}\left[\mu_1'\frac{t}{\sigma\sqrt{n}} + \mu_2'\frac{t^2}{2\sigma^2 n} + \mu_3'\frac{t^3}{6\sigma^3 n\sqrt{n}} + \cdots\right]^2$$

$$\left.+\frac{1}{3}\left[\mu_1'\frac{t}{\sigma\sqrt{n}} + \mu_2'\frac{t^2}{2\sigma^2 n} + \mu_3'\frac{t^3}{6\sigma^3 n\sqrt{n}} + \cdots\right]^3 - \cdots\right\}$$

Then, collecting powers of t, we obtain

$$\ln M_Z(t) = \left(-\frac{\sqrt{n}\,\mu}{\sigma} + \frac{\sqrt{n}\,\mu_1'}{\sigma}\right)t + \left(\frac{\mu_2'}{2\sigma^2} - \frac{\mu_1'^2}{2\sigma^2}\right)t^2$$

$$+\left(\frac{\mu_3'}{6\sigma^3\sqrt{n}} - \frac{\mu_1'\cdot\mu_2'}{2\sigma^3\sqrt{n}} + \frac{\mu_1'^3}{3\sigma^3\sqrt{n}}\right)t^3 + \cdots$$

and since $\mu_1' = \mu$ and $\mu_2' - (\mu_1')^2 = \sigma^2$, this reduces to

$$\ln M_Z(t) = \frac{1}{2}t^2 + \left(\frac{\mu_3'}{6} - \frac{\mu_1'\mu_2'}{2} + \frac{\mu_1'^3}{6}\right)\frac{t^3}{\sigma^3\sqrt{n}} + \cdots$$

Finally, observing that the coefficient of t^3 is a constant times $\frac{1}{\sqrt{n}}$ and in general, for $r \geqq 2$, the coefficient of t^r is a constant times $\frac{1}{\sqrt{n^{r-2}}}$, we get

$$\lim_{n\to\infty} \ln M_Z(t) = \frac{1}{2}t^2$$

and hence

$$\lim_{n\to\infty} M_Z(t) = e^{\frac{1}{2}t^2}$$

since the limit of a logarithm equals the logarithm of the limit (provided these limits exist). An illustration of this theorem is given in Exercise 13 and 14.

Sometimes, the central limit theorem is interpreted incorrectly as implying that the distribution of $\overline{X}$ approaches a normal distribution when $n \to \infty$. This is incorrect because $\operatorname{var}(\overline{X}) \to 0$ when $n \to \infty$; on the other hand, the central limit theorem does justify approximating the distribution of $\overline{X}$ with a normal distribution having the mean μ and the variance $\frac{\sigma^2}{n}$ when n is large. In practice, this approximation is used when $n \geqq 30$ regardless of the actual shape of the population sampled. For smaller values of n the approximation is questionable, but see Theorem 4.

EXAMPLE 1

A soft-drink vending machine is set so that the amount of drink dispensed is a random variable with a mean of 200 milliliters and a standard deviation of 15 milliliters. What is the probability that the average (mean) amount dispensed in a random sample of size 36 is at least 204 milliliters?

Solution

According to Theorem 1, the distribution of $\overline{X}$ has the mean $\mu_{\overline{X}} = 200$ and the standard deviation $\sigma_{\overline{X}} = \dfrac{15}{\sqrt{36}} = 2.5$, and according to the central limit theorem, this distribution is approximately normal. Since $z = \dfrac{204-200}{2.5} = 1.6$, it follows from Table III of "Statistical Tables" that $P(\overline{X} \geqq 204) = P(Z \geqq 1.6) = 0.5000 - 0.4452 = 0.0548$.

It is of interest to note that when the population we are sampling is normal, the distribution of $\overline{X}$ is a normal distribution regardless of the size of n.

THEOREM 4. If $\overline{X}$ is the mean of a random sample of size n from a normal population with the mean μ and the variance σ^2, its sampling distribution is a normal distribution with the mean μ and the variance σ^2/n.

Proof According to Theorems "If a and b are constants, then **1.** $M_{X+a}(t) = E[e^{(X+a)t}] = e^{at} \cdot M_X(t)$; **2.** $M_{bX}(t) = E(e^{bXt}) = M_X(bt)$; **3.** $M_{\frac{X+a}{b}}(t) = E[e^{\left(\frac{X+a}{b}\right)t}] = e^{\frac{a}{b}t} \cdot M_X\left(\dfrac{t}{b}\right)$. If $X_1, X_2, \ldots,$ and X_n are independent random variables and $Y = X_1 + X_2 + \cdots + X_n$, then $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$ where $M_{X_i}(t)$ is the value of the moment-generating function of X_i at t", we can write

$$M_{\overline{X}}(t) = \left[M_X\left(\frac{t}{n}\right)\right]^n$$

and since the moment-generating function of a normal distribution with the mean μ and the variance σ^2 is given by

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

according to the theorem $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$, we get

$$\begin{aligned} M_{\overline{X}}(t) &= \left[e^{\mu \cdot \frac{t}{n} + \frac{1}{2}\left(\frac{t}{n}\right)^2 \sigma^2}\right]^n \\ &= e^{\mu t + \frac{1}{2}t^2\left(\frac{\sigma^2}{n}\right)} \end{aligned}$$

This moment-generating function is readily seen to be that of a normal distribution with the mean μ and the variance σ^2/n.

3 The Sampling Distribution of the Mean: Finite Populations

If an experiment consists of selecting one or more values from a finite set of numbers $\{c_1, c_2, \ldots, c_N\}$, this set is referred to as a **finite population of size N**. In the definition that follows, it will be assumed that we are sampling *without replacement* from a finite population of size N.

DEFINITION 4. RANDOM SAMPLE—FINITE POPULATION. *If* X_1 *is the first value drawn from a finite population of size* N*,* X_2 *is the second value drawn, ...,* X_n *is the* n*th value drawn, and the joint probability distribution of these* n *random variables is given by*

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{N(N-1) \cdot \ldots \cdot (N-n+1)}$$

for each ordered n*-tuple of values of these random variables, then* $X_1, X_2, \ldots, X_n$ *are said to constitute a* ***random sample*** *from the given finite population.*

As in Definition 2, the random sample is a set of random variables, but here again it is common practice also to apply the term "random sample" to the values of the random variables, that is, to the actual numbers drawn.

From the joint probability distribution of Definition 4, it follows that the probability for each subset of n of the N elements of the finite population (regardless of the order in which the values are drawn) is

$$\frac{n!}{N(N-1) \cdot \ldots \cdot (N-n+1)} = \frac{1}{\binom{N}{n}}$$

This is often given as an alternative definition or as a criterion for the selection of a random sample of size n from a finite population of size N: *Each of the* $\binom{N}{n}$ *possible samples must have the same probability*.

It also follows from the joint probability distribution of Definition 4 that the marginal distribution of X_r is given by

$$f(x_r) = \frac{1}{N} \qquad \text{for } x_r = c_1, c_2, \ldots, c_N$$

for $r = 1, 2, \ldots, n$, and we refer to the mean and the variance of this discrete uniform distribution as the mean and the variance of the finite population. Therefore,

DEFINITION 5. SAMPLE MEAN AND VARIANCE—FINITE POPULATION. *The* ***sample mean*** *and the* ***sample variance*** *of the finite population* $\{c_1, c_2, \ldots, c_N\}$ *are*

$$\mu = \sum_{i=1}^{N} c_i \cdot \frac{1}{N} \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{N} (c_i - \mu)^2 \cdot \frac{1}{N}$$

Finally, it follows from the joint probability distribution of Definition 4 that the joint marginal distribution of any two of the random variables $X_1, X_2, \ldots, X_n$ is given by

$$g(x_r, x_s) = \frac{1}{N(N-1)}$$

for each ordered pair of elements of the finite population. Thus, we can prove the following theorem.

THEOREM 5. If X_r and X_s are the rth and sth random variables of a random sample of size n drawn from the finite population $\{c_1, c_2, \ldots, c_N\}$, then

$$\text{cov}(X_r, X_s) = -\frac{\sigma^2}{N-1}$$

Proof According to the definition given here "**COVARIANCE.** $\mu_{1,1}$ *is called the **covariance** of* X *and* Y*, and it is denoted by* σ_{XY}, *cov*(X, Y), or C(X, Y)",

$$\text{cov}(X_r, X_s) = \sum_{i=1}^{N}\sum_{\substack{j=1 \\ i \neq j}}^{N} \frac{1}{N(N-1)}(c_i - \mu)(c_j - \mu)$$

$$= \frac{1}{N(N-1)} \cdot \sum_{i=1}^{N}(c_i - \mu)\left[\sum_{\substack{j=1 \\ j \neq i}}^{N}(c_j - \mu)\right]$$

and since $\sum_{\substack{j=1 \\ j \neq i}}^{N}(c_j - \mu) = \sum_{j=1}^{N}(c_j - \mu) - (c_i - \mu) = -(c_i - \mu)$, we get

$$\text{cov}(X_r, X_s) = -\frac{1}{N(N-1)} \cdot \sum_{i=1}^{N}(c_i - \mu)^2$$

$$= -\frac{1}{N-1} \cdot \sigma^2$$

Making use of all these results, let us now prove the following theorem, which, for random samples from finite populations, corresponds to Theorem 1.

THEOREM 6. If $\overline{X}$ is the mean of a random sample of size n taken without replacement from a finite population of size N with the mean μ and the variance σ^2, then

$$E(\overline{X}) = \mu \quad \text{and} \quad \text{var}(\overline{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

Proof Substituting $a_i = \frac{1}{N}$, $\text{var}(X_i) = \sigma^2$, and $\text{cov}(X_i, X_j) = -\frac{\sigma^2}{N-1}$ into the formula $E(Y) = \sum_{i=1}^{n} a_i E(X_i)$, we get

$$E(\overline{X}) = \sum_{i=1}^{n} \frac{1}{n} \cdot \mu = \mu$$

and

$$\text{var}(\overline{X}) = \sum_{i=1}^{n} \frac{1}{n^2} \cdot \sigma^2 + 2 \cdot \sum\sum_{i<j} \frac{1}{n^2}\left(-\frac{\sigma^2}{N-1}\right)$$

$$= \frac{\sigma^2}{n} + 2 \cdot \frac{n(n-1)}{2} \cdot \frac{1}{n^2}\left(-\frac{\sigma^2}{N-1}\right)$$

$$= \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

It is of interest to note that the formulas we obtained for $\text{var}(\overline{X})$ in Theorems 1 and 6 differ only by the **finite population correction factor** $\frac{N-n}{N-1}$.† Indeed, when N is large compared to n, the difference between the two formulas for $\text{var}(\overline{X})$ is usually negligible, and the formula $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ is often used as an approximation when we are sampling from a large finite population. A general rule of thumb is to use this approximation when the sample does not constitute more than 5 percent of the population.

Exercises

1. This question has been intentionally omitted for this edition.

2. This question has been intentionally omitted for this edition.

3. With reference to Exercise 2, show that if the two samples come from normal populations, then $\overline{X}_1 - \overline{X}_2$ is a random variable having a normal distribution with the mean $\mu_1 - \mu_2$ and the variance $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. (*Hint*: Proceed as in the proof of Theorem 4.)

4. If $X_1, X_2, \ldots, X_n$ are independent random variables having identical Bernoulli distributions with the parameter θ, then $\overline{X}$ is the proportion of successes in n trials, which we denote by $\hat{\Theta}$. Verify that

(a) $E(\hat{\Theta}) = \theta$;

(b) $\text{var}(\hat{\Theta}) = \frac{\theta(1-\theta)}{n}$.

5. If the first n_1 random variables of Exercise 2 have Bernoulli distributions with the parameter θ_1 and the other n_2 random variables have Bernoulli distributions with the parameter θ_2, show that, in the notation of Exercise 4,

(a) $E(\hat{\Theta}_1 - \hat{\Theta}_2) = \theta_1 - \theta_2$;

(b) $\text{var}(\hat{\Theta}_1 - \hat{\Theta}_2) = \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}$.

6. This question has been intentionally omitted for this edition.

7. The following is a sufficient condition for the central limit theorem: If the random variables $X_1, X_2, \ldots, X_n$ are independent and uniformly bounded (that is, there exists a positive constant k such that the probability is zero that any one of the random variables X_i will take on a value greater than k or less than $-k$), then if the variance of

$$Y_n = X_1 + X_2 + \cdots + X_n$$

becomes infinite when $n \to \infty$, the distribution of the standardized mean of the X_i approaches the standard normal distribution. Show that this sufficient condition holds for a sequence of independent random variables X_i having the respective probability distributions

$$f_i(x_i) = \begin{cases} \frac{1}{2} & \text{for } x_i = 1 - (\frac{1}{2})^i \\ \frac{1}{2} & \text{for } x_i = (\frac{1}{2})^i - 1 \end{cases}$$

8. Consider the sequence of independent random variables $X_1, X_2, X_3, \ldots$ having the uniform densities

$$f_i(x_i) = \begin{cases} \frac{1}{2 - \frac{1}{i}} & \text{for } 0 < x_i < 2 - \frac{1}{i} \\ 0 & \text{elsewhere} \end{cases}$$

Use the sufficient condition of Exercise 7 to show that the central limit theorem holds.

9. The following is a sufficient condition, the *Laplace–Liapounoff condition*, for the central limit theorem: If $X_1, X_2, X_3, \ldots$ is a sequence of independent random variables, each having an absolute third moment

$$c_i = E(|X_i - \mu_i|^3)$$

and if

$$\lim_{n \to \infty} [\text{var}(Y_n)]^{-\frac{3}{2}} \cdot \sum_{i=1}^{n} c_i = 0$$

where $Y_n = X_1 + X_2 + \cdots + X_n$, then the distribution of the standardized mean of the X_i approaches the standard normal distribution when $n \to \infty$. Use this condition to show that the central limit theorem holds for the sequence of random variables of Exercise 7.

10. Use the condition of Exercise 9 to show that the central limit theorem holds for the sequence of random variables of Exercise 8.

†Since there are many problems in which we are interested in the standard deviation rather than the variance, the term "finite population correction factor" often refers to $\sqrt{\frac{N-n}{N-1}}$ instead of $\frac{N-n}{N-1}$. This does not matter, of course, as long as the usage is clearly understood.