Chapter 5

The Delta Method and Applications

5.1 Linear approximations of functions

In the simplest form of the central limit theorem, Theorem 4.18, we consider a sequence $X_1, X_2, \ldots$ of independent and identically distributed (univariate) random variables with finite variance $\sigma^2$. In this case, the central limit theorem states that

$$\sqrt{n}(X_n - \mu) \xrightarrow{d} \sigma Z,$$  \hspace{1cm} (5.1)

where $\mu = E[X_1]$ and $Z$ is a standard normal random variable.

In this chapter, we wish to consider the asymptotic distribution of, say, some function of $X_n$. In the simplest case, the answer depends on results already known: Consider a linear function $g(t) = at + b$ for some known constants $a$ and $b$. Since $E[X_n] = \mu$, clearly $E[g(X_n)] = a\mu + b = g(\mu)$ by the linearity of the expectation operator. Therefore, it is reasonable to ask whether $\sqrt{n}[g(X_n) - g(\mu)]$ tends to some distribution as $n \to \infty$. But the linearity of $g(t)$ allows one to write

$$\sqrt{n} [g(X_n) - g(\mu)] = a\sqrt{n}(X_n - \mu).$$

We conclude by Theorem 2.24 that

$$\sqrt{n} [g(X_n) - g(\mu)] \xrightarrow{d} a\sigma Z.$$

Of course, the distribution on the right hand side above is $N(0, a^2\sigma^2)$.

None of the preceding development is especially deep; one might even say that it is obvious that a linear transformation of the random variable $X_n$ alters its asymptotic distribution
by a constant multiple. Yet what if the function $g(t)$ is nonlinear? It is in this nonlinear case that a strong understanding of the argument above, as simple as it may be, pays real dividends. For if $X_n$ is consistent for $\mu$ (say), then we know that, roughly speaking, $X_n$ will be very close to $\mu$ for large $n$. Therefore, the only meaningful aspect of the behavior of $g(t)$ is its behavior in a small neighborhood of $\mu$. And in a small neighborhood of $\mu$, $g(\mu)$ may be considered to be roughly a linear function if we use a first-order Taylor expansion.

In particular, we may approximate

$$g(t) \approx g(\mu) + g'(\mu)(t - \mu)$$

for $t$ in a small neighborhood of $\mu$. We see that $g'(\mu)$ is the multiple of $t$, and so the logic of the linear case above suggests

$$\sqrt{n} \{ g(X_n) - g(\mu) \} \xrightarrow{d} g'(\mu)\sigma^2 Z.$$  \hspace{1cm} (5.2)

Indeed, expression (5.2) is a special case of the powerful theorem known as the delta method.

**Theorem 5.1**  \hspace{1cm} **Delta method:**  \hspace{1cm} If $g'(a)$ exists and $n^b(X_n - a) \xrightarrow{d} X$ for $b > 0$, then

$$n^b \{ g(X_n) - g(a) \} \xrightarrow{d} g'(a)X.$$  

The proof of the delta method uses Taylor’s theorem, Theorem 1.18: Since $X_n - a \xrightarrow{p} 0$,

$$n^b \{ g(X_n) - g(a) \} = n^b(X_n - a) \{ g'(a) + o_P(1) \},$$

and thus Slutsky’s theorem together with the fact that $n^b(X_n - a) \xrightarrow{d} X$ proves the result.

Expression (5.2) may be reexpressed as a corollary of Theorem 5.1:

**Corollary 5.2**  \hspace{1cm} The often-used special case of Theorem 5.1 in which $X$ is normally distributed states that if $g'(\mu)$ exists and $\sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$, then

$$\sqrt{n} \{ g(X_n) - g(\mu) \} \xrightarrow{d} N \{ 0, \sigma^2 g'(\mu)^2 \}.$$ 

Ultimately, we will extend Theorem 5.1 in two directions: Theorem 5.5 deals with the special case in which $g'(a) = 0$, and Theorem 5.6 is the multivariate version of the delta method. But we first apply the delta method to a couple of simple examples that illustrate a frequently understood but seldom stated principle: When we speak of the “asymptotic distribution” of a sequence of random variables, we generally refer to a nontrivial (i.e., nonconstant) distribution. Thus, in the case of an independent and identically distributed sequence $X_1, X_2, \ldots$ of random variables with finite variance, the phrase “asymptotic distribution of $X_n$” generally refers to the fact that

$$\sqrt{n} (X_n - E X_1) \xrightarrow{d} N(0, \text{Var } X_1),$$

not the fact that $X_n \xrightarrow{p} E X_1$. 

86
Example 5.3  Asymptotic distribution of $X_n^2$ Suppose $X_1, X_2, \ldots$ are iid with mean $\mu$ and finite variance $\sigma^2$. Then by the central limit theorem,

$$\sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2).$$

Therefore, the delta method gives

$$\sqrt{n}(X_n^2 - \mu^2) \xrightarrow{d} N(0, 4\mu^2\sigma^2). \quad (5.3)$$

However, this is not necessarily the end of the story. If $\mu = 0$, then the normal limit in (5.3) is degenerate—that is, expression (5.3) merely states that $\sqrt{n}X_n^2$ converges in probability to the constant 0. This is not what we mean by the asymptotic distribution! Thus, we must treat the case $\mu = 0$ separately, noting in that case that $\sqrt{n}X_n \xrightarrow{d} N(0, \sigma^2)$ by the central limit theorem, which implies that

$$nX_n \xrightarrow{d} \sigma^2 \chi^2_1.$$

Example 5.4  Estimating binomial variance: Suppose $X_n \sim \text{binomial}(n, p)$. Because $X_n/n$ is the maximum likelihood estimator for $p$, the maximum likelihood estimator for $p(1-p)$ is $\delta_n = X_n(n - X_n)/n^2$. The central limit theorem tells us that $\sqrt{n}(X_n/n - p) \xrightarrow{d} N\{0, p(1-p)\}$, so the delta method gives

$$\sqrt{n} \{\delta_n - p(1-p)\} \xrightarrow{d} N\{0, p(1-p)(1 - 2p)^2\}.$$

Note that in the case $p = 1/2$, this does not give the asymptotic distribution of $\delta_n$. Exercise 5.1 gives a hint about how to find the asymptotic distribution of $\delta_n$ in this case.

We have seen in the preceding examples that if $g'(a) = 0$, then the delta method gives something other than the asymptotic distribution we seek. However, by using more terms in the Taylor expansion, we obtain the following generalization of Theorem 5.1:

Theorem 5.5  If $g(t)$ has $r$ derivatives at the point $a$ and $g'(a) = g''(a) = \cdots = g^{(r-1)}(a) = 0$, then $n^b(X_n - a) \xrightarrow{d} X$ for $b > 0$ implies that

$$n^b \{g(X_n) - g(a)\} \xrightarrow{d} \frac{1}{r!} g^{(r)}(a)X^r.$$

It is straightforward using the multivariate notion of differentiability discussed in Definition 1.34 to prove the following theorem:
Theorem 5.6  Multivariate delta method: If \( g : R^k \to R^\ell \) has a derivative \( \nabla g(a) \) at \( a \in R^k \) and
\[
n^b (X_n - a) \overset{d}{\to} Y
\]
for some \( k \)-vector \( Y \) and some sequence \( X_1, X_2, \ldots \) of \( k \)-vectors, where \( b > 0 \), then
\[
n^b \{ g(X_n) - g(a) \} \overset{d}{\to} [\nabla g(a)]^T Y.
\]
The proof of Theorem 5.6 involves a simple application of the multivariate Taylor expansion of Equation (1.18).

Exercises for Section 5.1

Exercise 5.1  Let \( \delta_n \) be defined as in Example 5.4. Find the asymptotic distribution of \( \delta_n \) in the case \( p = 1/2 \). That is, find constant sequences \( a_n \) and \( b_n \) and a nontrivial random variable \( X \) such that \( a_n(\delta_n - b_n) \overset{d}{\to} X \).

Hint:  Let \( Y_n = X_n - (n/2) \). Apply the central limit theorem to \( Y_n \), then transform both sides of the resulting limit statement so that a statement involving \( \delta_n \) results.

Exercise 5.2  Prove Theorem 5.5.

5.2 Variance stabilizing transformations

Often, if \( E(X_i) = \mu \) is the parameter of interest, the central limit theorem gives
\[
\sqrt{n}(\bar{X}_n - \mu) \overset{d}{\to} N\{0, \sigma^2(\mu)\}.
\]
In other words, the variance of the limiting distribution is a function of \( \mu \). This is a problem if we wish to do inference for \( \mu \), because ideally the limiting distribution should not depend on the unknown \( \mu \). The delta method gives a possible solution: Since
\[
\sqrt{n} \left\{ g(\bar{X}_n) - g(\mu) \right\} \overset{d}{\to} N \left\{ 0, \sigma^2(\mu)g'(\mu)^2 \right\},
\]
we may search for a transformation \( g(x) \) such that \( g'(\mu)\sigma(\mu) \) is a constant. Such a transformation is called a variance stabilizing transformation.
Example 5.7  Suppose that $X_1, X_2, \ldots$ are independent normal random variables with mean 0 and variance $\sigma^2$. Let us define $\tau^2 = \text{Var} X_i^2$, which for the normal distribution may be seen to be $2\sigma^4$. (To verify this, try showing that $\mathbb{E} X_i^4 = 3\sigma^4$ by differentiating the normal characteristic function four times and evaluating at zero.) Thus, Example 4.11 shows that

$$\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \sigma^2 \right) \overset{d}{\to} N(0, 2\sigma^4).$$

To do inference for $\sigma^2$ when we believe that our data are truly independent and identically normally distributed, it would be helpful if the limiting distribution did not depend on the unknown $\sigma^2$. Therefore, it is sensible in light of Corollary 5.2 to search for a function $g(t)$ such that $2[g'(\sigma^2)]^2\sigma^4$ is not a function of $\sigma^2$. In other words, we want $g'(t)$ to be proportional to $\sqrt{t^{-2}} = \sqrt{|t|^{-1}}$. Clearly $g(t) = \log t$ is such a function. Therefore, we call the logarithm function a variance-stabilizing function in this example, and Corollary 5.2 shows that

$$\sqrt{n} \left\{ \log \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 \right) - \log (\sigma^2) \right\} \overset{d}{\to} N(0, 2).$$

Exercises for Section 5.2

Exercise 5.3  Suppose $X_n \sim \text{binomial}(n, p)$, where $0 < p < 1$.

(a) Find the asymptotic distribution of $g(X_n/n) - g(p)$, where $g(x) = \min\{x, 1-x\}$.

(b) Show that $h(x) = \sin^{-1}(\sqrt{x})$ is a variance-stabilizing transformation. 

Hint: $(d/du) \sin^{-1}(u) = 1/\sqrt{1-u^2}$.

Exercise 5.4  Let $X_1, X_2, \ldots$ be independent from $N(\mu, \sigma^2)$ where $\mu \neq 0$. Let

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

Find the asymptotic distribution of the coefficient of variation $S_n/\overline{X}_n$.

Exercise 5.5  Let $X_n \sim \text{binomial}(n, p)$, where $p \in (0, 1)$ is unknown. Obtain confidence intervals for $p$ in two different ways:
(a) Since $\sqrt{n}(X_n/n - p) \xrightarrow{d} N[0, p(1-p)]$, the variance of the limiting distribution depends only on $p$. Use the fact that $X_n/n \xrightarrow{P} p$ to find a consistent estimator of the variance and use it to derive a 95% confidence interval for $p$.

(b) Use the result of problem 5.3(b) to derive a 95% confidence interval for $p$.

(c) Evaluate the two confidence intervals in parts (a) and (b) numerically for all combinations of $n \in \{10, 100, 1000\}$ and $p \in \{.1, .3, .5\}$ as follows: For 1000 realizations of $X \sim \text{bin}(n, p)$, construct both 95% confidence intervals and keep track of how many times (out of 1000) that the confidence intervals contain $p$. Report the observed proportion of successes for each $(n, p)$ combination. Does your study reveal any differences in the performance of these two competing methods?

5.3 Sample Moments

The weak law of large numbers tells us that If $X_1, X_2, \ldots$ are independent and identically distributed with $E |X_1|^k < \infty$, then

$$\frac{1}{n} \sum_{i=1}^{n} X_i^k \xrightarrow{P} E X_i^k.$$ 

That is, sample moments are (weakly) consistent. For example, the sample variance, which we define as

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

is consistent for $\text{Var } X_i = E X_i^2 - (E X_i)^2$.

However, consistency is not the end of the story. The central limit theorem and the delta method will prove very useful in deriving asymptotic distribution results about sample moments. We consider two very important examples involving the sample variance of Equation (5.4).

Example 5.8 Distribution of sample $T$ statistic: Suppose $X_1, X_2, \ldots$ are iid with $E (X_i) = \mu$ and $\text{Var } (X_i) = \sigma^2 < \infty$. Define $s_n^2$ as in Equation (5.4), and let

$$T_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{s_n}.$$
Letting

\[ A_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \]

and \( B_n = \frac{\sigma}{s_n} \), we obtain \( T_n = A_nB_n \). Therefore, since \( A_n \xrightarrow{d} N(0, 1) \) by the central limit theorem and \( B_n \xrightarrow{P} 1 \) by the weak law of large numbers, Slutsky’s theorem implies that \( T_n \xrightarrow{d} N(0, 1) \). In other words, \( T \) statistics are asymptotically normal under the null hypothesis.

**Example 5.9** Let \( X_1, X_2, \ldots \) be independent and identically distributed with mean \( \mu \), variance \( \sigma^2 \), third central moment \( E((X_i - \mu)^3) = \gamma \), and \( \text{Var}(X_i - \mu)^2 = \tau^2 < \infty \). Define \( S_n^2 \) as in Equation (4.6). We have shown earlier that \( \sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{d} N(0, \tau^2) \). The same fact may be proven using Theorem 4.9 as follows.

First, let \( Y_i = X_i - \mu \) and \( Z_i = Y_i^2 \). We may use the multivariate central limit theorem to find the joint asymptotic distribution of \( \bar{Y}_n \) and \( \bar{Z}_n \), namely

\[
\sqrt{n} \left\{ \begin{pmatrix} \bar{Y}_n \\ \bar{Z}_n \end{pmatrix} - \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix} \right\} \xrightarrow{d} N_2 \left\{ 0, \begin{pmatrix} \sigma^2 & \gamma \\ \gamma & \tau^2 \end{pmatrix} \right\}.
\]

Note that the above result uses the fact that \( \text{Cov}(Y_1, Z_1) = \gamma \).

We may write \( S_n^2 = \bar{Z}_n - (\bar{Y}_n)^2 \). Therefore, define the function \( g(a, b) = b - a^2 \) and note that this gives \( \dot{g}(a, b) = (-2a, 1) \). To use the delta method, we should evaluate

\[
\dot{g}(0, \sigma^2) \begin{pmatrix} \sigma^2 & \gamma \\ \gamma & \tau^2 \end{pmatrix} \dot{g}(0, \sigma^2)^T = (0, 1) \begin{pmatrix} \sigma^2 & \gamma \\ \gamma & \tau^2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \dot{g}(0, \sigma^2)^T = \tau^2
\]

We conclude that

\[
\sqrt{n} \left\{ g\left( \begin{pmatrix} \bar{Y}_n \\ \bar{Z}_n \end{pmatrix} \right) - g\left( \begin{pmatrix} 0 \\ \sigma^2 \end{pmatrix} \right) \right\} = \sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{d} N(0, \tau^2)
\]

as we found earlier (using a different argument) in Example 4.11.

**Exercises for Section 5.3**

**Exercise 5.6** Suppose that \( X_1, X_2, \ldots \) are iid Normal \((0, \sigma^2)\) random variables.

(a) Based on the result of Example 5.7, Give an approximate test at \( \alpha = .05 \) for \( H_0 : \sigma^2 = \sigma_0^2 \) vs. \( H_a : \sigma^2 \neq \sigma_0^2 \).
For \( n = 25 \), estimate the true level of the test in part (a) for \( \sigma_0^2 = 1 \) by simulating 5000 samples of size \( n = 25 \) from the null distribution. Report the proportion of cases in which you reject the null hypothesis according to your test (ideally, this proportion will be about .05).

5.4 Sample Correlation

Suppose that \((X_1, Y_1), (X_2, Y_2), \ldots\) are iid vectors with \( E X_i^4 < \infty \) and \( E Y_i^4 < \infty \). For the sake of simplicity, we will assume without loss of generality that \( E X_i = E Y_i = 0 \) (alternatively, we could base all of the following derivations on the centered versions of the random variables).

We wish to find the asymptotic distribution of the sample correlation coefficient, \( r \). If we let

\[
\begin{pmatrix}
m_x \\
m_y \\
m_{xx} \\
m_{yy} \\
m_{xy}
\end{pmatrix} = \frac{1}{n} \begin{pmatrix}
\sum_{i=1}^{n} X_i \\
\sum_{i=1}^{n} Y_i \\
\sum_{i=1}^{n} X_i^2 \\
\sum_{i=1}^{n} Y_i^2 \\
\sum_{i=1}^{n} X_i Y_i
\end{pmatrix}
\]

and

\[
s_x^2 = m_{xx} - m_x^2, s_y^2 = m_{yy} - m_y^2, \text{ and } s_{xy} = m_{xy} - m_x m_y,
\]

then \( r = s_{xy} / (s_x s_y) \). According to the central limit theorem,

\[
\sqrt{n} \begin{pmatrix}
m_x \\
m_y \\
m_{xx} \\
m_{yy} \\
m_{xy}
\end{pmatrix} \rightarrow N_5 \begin{pmatrix}
0 \\
0 \\
\sigma_x^2 \\
\sigma_y^2 \\
\sigma_{xy}
\end{pmatrix}, \begin{pmatrix}
\text{Cov} (X_1, X_1) & \cdots & \text{Cov} (X_1, X_1 Y_1) \\
\text{Cov} (Y_1, X_1) & \cdots & \text{Cov} (Y_1, X_1 Y_1) \\
\vdots & \ddots & \vdots \\
\text{Cov} (X_1 Y_1, X_1) & \cdots & \text{Cov} (X_1 Y_1, X_1 Y_1)
\end{pmatrix}
\]

Let \( \Sigma \) denote the covariance matrix in expression (5.7). Define a function \( g : \mathbb{R}^5 \rightarrow \mathbb{R}^3 \) such that \( g \) applied to the vector of moments in Equation (5.5) yields the vector \((s_x^2, s_y^2, s_{xy})\) as defined in expression (5.6). Then

\[
\nabla g \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = \begin{pmatrix} -2a & 0 & -b \\ 0 & -2b & -a \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\]
Therefore, if we let

\[ \Sigma^* = \begin{bmatrix} \nabla g \left( \begin{array}{c} 0 \\ \sigma_x^2 \\ \sigma_y^2 \\ \sigma_{xy} \end{array} \right) \end{bmatrix} \Sigma \begin{bmatrix} \nabla g \left( \begin{array}{c} 0 \\ \sigma_x^2 \\ \sigma_y^2 \\ \sigma_{xy} \end{array} \right) \end{bmatrix}^T \]

\[ = \begin{pmatrix} \text{Cov}(X_1^2, X_1^2) & \text{Cov}(X_1^2, Y_1^2) & \text{Cov}(X_1^2, X_1 Y_1) \\ \text{Cov}(Y_1^2, X_1^2) & \text{Cov}(Y_1^2, Y_1^2) & \text{Cov}(Y_1^2, X_1 Y_1) \\ \text{Cov}(X_1 Y_1, X_1^2) & \text{Cov}(X_1 Y_1, Y_1^2) & \text{Cov}(X_1 Y_1, X_1 Y_1) \end{pmatrix}, \]

then by the delta method,

\[ \sqrt{n} \left\{ \begin{pmatrix} s_x^2 \\ s_y^2 \\ s_{xy} \end{pmatrix} - \begin{pmatrix} \sigma_x^2 \\ \sigma_y^2 \\ \sigma_{xy} \end{pmatrix} \right\} \xrightarrow{d} N_3(0, \Sigma^*). \] (5.8)

As an aside, note that expression (5.8) gives the same marginal asymptotic distribution for \( \sqrt{n} (s_x^2 - \sigma_x^2) \) as was derived using a different approach in Example 4.11, since Cov \((X_1^2, X_1^2)\) is the same as \(\tau^2\) in that example.

Next, define the function \( h(a, b, c) = c/\sqrt{ab} \), so that we have \( h(s_x^2, s_y^2, s_{xy}) = r \). Then

\[ [\nabla h(a, b, c)]^T = \frac{1}{2} \begin{pmatrix} -c/\sqrt{a^3b}, -c/\sqrt{ab^3}, 2/\sqrt{ab} \end{pmatrix}, \]

so that

\[ [\nabla h(\sigma_x^2, \sigma_y^2, \sigma_{xy})]^T = \begin{pmatrix} -\sigma_{xy}/2\sigma_x^4\sigma_y, -\sigma_{xy}/2\sigma_x^3\sigma_y^2, 1/\sigma_x\sigma_y \end{pmatrix} = \begin{pmatrix} -\rho/2\sigma_x^2, -\rho/2\sigma_y^2, 1/\sigma_x\sigma_y \end{pmatrix}. \] (5.9)

Therefore, if \( A \) denotes the \( 1 \times 3 \) matrix in Equation (5.9), using the delta method once again yields

\[ \sqrt{n}(r - \rho) \xrightarrow{d} N(0, A\Sigma^* A^T). \]

To recap, we have used the basic tools of the multivariate central limit theorem and the multivariate delta method to obtain a univariate result. This derivation of univariate facts via multivariate techniques is common practice in statistical large-sample theory.

**Example 5.10** Consider the special case of bivariate normal \((X_i, Y_i)\). In this case, we may derive

\[ \Sigma^* = \begin{pmatrix} 2\sigma_x^4 & 2\rho\sigma_x^2\sigma_y^2 & 2\rho\sigma_x^3\sigma_y \\ 2\rho^2\sigma_x^2\sigma_y^2 & 2\sigma_y^4 & 2\rho\sigma_y^3 \sigma_x \\ 2\rho^3\sigma_x^3\sigma_y & 2\rho\sigma_x\sigma_y^3 & (1 + \rho^2)^2 \sigma_x^2 \sigma_y^2 \end{pmatrix}. \] (5.10)

93
In this case, $A \Sigma^* A^T = (1 - \rho^2)^2$, which implies that
\[
\sqrt{n}(r - \rho) \xrightarrow{d} N\{0, (1 - \rho^2)^2\}. \quad (5.11)
\]

In the normal case, we may derive a variance-stabilizing transformation. According to Equation (5.11), we should find a function $f(x)$ satisfying $f''(x) = (1-x^2)^{-1}$. Since
\[
\frac{1}{1-x^2} = \frac{1}{2(1-x)} + \frac{1}{2(1+x)},
\]
we integrate to obtain
\[
f(x) = \frac{1}{2} \log \frac{1+x}{1-x}.
\]
This is called Fisher’s transformation; we conclude that
\[
\sqrt{n} \left( \frac{1}{2} \log \frac{1+r}{1-r} - \frac{1}{2} \log \frac{1+\rho}{1-\rho} \right) \xrightarrow{d} N(0,1).
\]

**Exercises for Section 5.4**

**Exercise 5.7** Verify expressions (5.10) and (5.11).

**Exercise 5.8** Assume $(X_1, Y_1), \ldots, (X_n, Y_n)$ are iid from some bivariate normal distribution. Let $\rho$ denote the population correlation coefficient and $r$ the sample correlation coefficient.

(a) Describe a test of $H_0 : \rho = 0$ against $H_1 : \rho \neq 0$ based on the fact that
\[
\sqrt{n}[f(r) - f(\rho)] \xrightarrow{d} N(0,1),
\]
where $f(x)$ is Fisher’s transformation $f(x) = (1/2)\log[(1 + x)/(1 - x)]$. Use $\alpha = .05$.

(b) Based on 5000 repetitions each, estimate the actual level for this test in the case when $E(X_i) = E(Y_i) = 0$, Var$(X_i) = \text{Var}(Y_i) = 1$, and $n \in \{3, 5, 10, 20\}$.

**Exercise 5.9** Suppose that $X$ and $Y$ are jointly distributed such that $X$ and $Y$ are Bernoulli $(1/2)$ random variables with $P(XY = 1) = \theta$ for $\theta \in (0, 1/2)$. Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be iid with $(X_i, Y_i)$ distributed as $(X, Y)$.
(a) Find the asymptotic distribution of $\sqrt{n} \left[ (X_n, Y_n) - (1/2, 1/2) \right]$.

(b) If $r_n$ is the sample correlation coefficient for a sample of size $n$, find the asymptotic distribution of $\sqrt{n}(r_n - \rho)$.

(c) Find a variance stabilizing transformation for $r_n$.

(d) Based on your answer to part (c), construct a 95% confidence interval for $\theta$.

(e) For each combination of $n \in \{5, 20\}$ and $\theta \in \{.05, .25, .45\}$, estimate the true coverage probability of the confidence interval in part (d) by simulating 5000 samples and the corresponding confidence intervals. One problem you will face is that in some samples, the sample correlation coefficient is undefined because with positive probability each of the $X_i$ or $Y_i$ will be the same. In such cases, consider the confidence interval to be undefined and the true parameter therefore not contained therein.

**Hint:** To generate a sample of $(X, Y)$, first simulate the $X$’s from their marginal distribution, then simulate the $Y$’s according to the conditional distribution of $Y$ given $X$. To obtain this conditional distribution, find $P(Y = 1 \mid X = 1)$ and $P(Y = 1 \mid X = 0)$. 