The Pre-Test Estimator

James J. Heckman and Rodrigo Pinto
University of Chicago

Econ 312

Spring 2022

1 The Model

The True Model

$$Model: Y_i = X_i \left[\begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right] + \varepsilon_i; \;\; i = 1, ..., N; \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.};$$

$$Y_i = X_{1,i}\beta_1 + X_{2,i}\beta_2 + \varepsilon_i$$

$$X_i \equiv \left[\begin{array}{c} X_{1i} \\ X_{2i} \end{array} \right] \sim \mathcal{N}\left[0, \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right] \quad \text{i.i.d.}; \quad \left[\begin{array}{c} \beta_1 \\ \beta_2 \end{array} \right] = \left[\begin{array}{c} 1 \\ 1 \end{array} \right]$$

Parameters estimation

Procedure 1 $\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} \equiv \widehat{\beta}$ are OLS estimation of β using data from the model above.

Procedure 2 $\overline{\beta}_1$is OLS estimation of $Y_i = \beta_1 \; X_{1i} + \xi_i$ using data from the true model above

Procedure 3 $\widetilde{\beta} = \begin{cases} \widetilde{\beta} = \widehat{\beta} \text{ if } |t|_{\widehat{\beta}_2} > 2 \\ \widetilde{\beta} = \overline{\beta}_1 \text{ otherwise} \end{cases}$,

where $|t|_{\widehat{\beta}_2}$ means the absolute value of the random variable t-statistic for the estimator $\widehat{\beta}_2$.

The first procedure estimates the OLS parameters. The second procedure estimates a wrong model in which the parameter β_2 is not computed.

The third procedure uses two estimations and the parameter β_2 belongs to the final one only if it is statistically significant. This is the common method of estimating a OLS with a number of variables in order to observe which of them are statistically significant. The second and final estimation uses only the variables which had significant estimated parameters in the first estimation.

Observe that:

$$\widetilde{\beta}_1 = \begin{cases} \widehat{\beta}_1 \text{ with Prob.} = \Pr\left(|t|_{\widehat{\beta}_2} > 2\right) \\ \overline{\beta}_1 \ \text{ with Prob.} = 1 - \Pr\left(|t|_{\widehat{\beta}_2} > 2\right) \end{cases}$$

2 The Variance Analysis in the True Model

The OLS Estimation

$$\begin{aligned}\widehat{\beta} &= \left(X'X\right)^{-1} X'Y = \beta + \left(X'X\right)^{-1} X'\varepsilon \\ &= \beta + \left(\frac{X'X}{N}\right)^{-1} \frac{X'\varepsilon}{N} \to_p \beta\end{aligned}$$

$$\begin{aligned}\sqrt{N} \cdot \left(\widehat{\beta} - \beta\right) &\to {}_d\mathcal{N}\left(\mathbf{0}, Q_{XX} \cdot \sigma^2\right) \\ Q_{XX} &= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\end{aligned}$$

For N big enough:

$$\widehat{\beta} \sim \mathcal{N}\left(\beta, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\right)$$

$$\rho \to 0 \Rightarrow var(\widehat{\beta}) \to \frac{\sigma^2}{N} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\rho \to \pm 1 \Rightarrow var(\widehat{\beta}) \to \infty \text{ (perfect colinearity)}$$

$$\begin{aligned} \widehat{\beta}_1 &\sim \mathcal{N}\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right) \\ \widehat{\beta}_2 &\sim \mathcal{N}\left(\beta_2, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right) \end{aligned}$$

3 The Bias and Variance Analysis in the second Procedure

The Exclusion of β_2 Estimation

$$\overline{\beta}_1 = \left(X_1' X_1\right)^{-1} X_1' Y = \frac{\sum_{i=1}^{N} X_{1,i} Y_i}{\sum_{i=1}^{N} X_{1,i}^2}$$

$$\begin{aligned}
&= \frac{\sum_{i=1}^{N} \left(X_{1,i}\beta_1 + X_{2,i}\beta_2 + \varepsilon_i\right) X_{1,i}}{\sum_{i=1}^{N} X_{1,i}^2} \\
&= \beta_1 + \frac{\beta_2 \sum_{i=1}^{N} X_{1,i} X_{2,i} + \sum_{i=1}^{N} \varepsilon_i X_{1,i}}{\sum_{i=1}^{N} X_{1,i}^2} \\
&= \beta_1 + \frac{\beta_2 \sum_{i=1}^{N} \frac{X_{1,i} X_{2,i}}{N} + \sum_{i=1}^{N} \frac{\varepsilon_i X_{1,i}}{N}}{\sum_{i=1}^{N} \frac{X_{1,i}^2}{N}}
\end{aligned}$$

But:

$$\sum_{i=1}^{N} \frac{\varepsilon_i X_{1,i}}{N} \rightarrow_p E\left(\varepsilon_i X_{1,i}\right) = 0 \ ;$$

$$\sum_{i=1}^{N} \frac{X_{1,i} X_{2,i}}{N} \rightarrow_p E\left(X_{1,i} X_{2,i}\right) = \rho \quad \text{(WLLN)}$$

$$\sum_{i=1}^{N} X_{1,i}^2 = \mathcal{X}_N^2 \Rightarrow \sum_{i=1}^{N} \frac{X_{1,i}^2}{N} \rightarrow_p 1$$

$$\Rightarrow \overline{\beta}_1 \rightarrow_p \beta_1 + \rho \beta_2$$

The absolute bias in the second procedure gets bigger as

$$|\rho| \rightarrow_p 1.$$

The variance of the second procedure:

The $\overline{\beta}_1$ Variance: Exclusion of β_2

$$\begin{aligned} assvar(\overline{\beta}_1) &= assvar\left((X_1'X_1)^{-1} X_1'Y\right) \\ &= assvar\left((X_1'X_1)^{-1} X_1' (X_1\beta_1 + X_2\beta_2 + \varepsilon)\right) \\ &= assvar\left(\beta_1 + \beta_2 (X_1'X_1)^{-1} X_1'X_2 + (X_1'X_1)^{-1} X_1'\varepsilon\right) \end{aligned}$$

$$
\begin{aligned}
&= p\lim\left(\beta_2^2\left[\left(\left(\frac{X_1'X_1}{N}\right)^{-1}\frac{X_1'X_2}{\sqrt{N}}\right)-\rho\right]^2\right) \\
&\quad +p\lim\left(\left[\left(\frac{X_1'X_1}{N}\right)^{-1}\frac{X_1'\varepsilon}{\sqrt{N}}\right]^2\right) \\
&= E\left(X_{i,1}\right)^{-1}\cdot\left(\beta_2^2\cdot var\left[X_{i,1}X_{i,2}\right]+\sigma^2\right) \\
&= \beta_2^2\cdot var\left[X_{i,1}X_{i,2}\right]+\sigma^2 \\
&= \beta_2^2\cdot\left(E\left[X_{i,1}^2X_{i,2}^2\right]-\rho^2\right)+\sigma^2
\end{aligned}
$$

Observe that:

$$
\begin{aligned}
X_1 &= \mathcal{P}_{X_2}\left(X_1\right)+\mathcal{P}_{\perp X_2}\left(X_1\right) \\
X_1 &= \rho X_2+v;\;\; v\perp X_2
\end{aligned}
$$

But X_2, X_1 are normals $\Rightarrow v$ is normal

$$var[X_1] = 1 = \rho^2 + var[v]$$
$$\Rightarrow var[v] = 1 - \rho^2 \text{ and } E[v] = 0.$$

(The Jacobian method of change of variables can also provide the same result: $v = X_1 - \rho X_2$.)

$$\Rightarrow E(X_1^2 X_2^2) = E((\rho X_2 + v)^2 \cdot X_2^2)$$

but:

$$E\left(X_2^j\right) = \begin{cases} \frac{j}{(j/2)!} \cdot \frac{\sigma_{\beta_2}^j}{2^{j/2}} \text{ for } j \text{ even} \\ 0 \text{ for } j \text{ odd} \end{cases}$$

$$\Rightarrow \quad E\left(X_2^2\right) = 3\sigma_{\beta_2}^4 = 3$$
$$\Rightarrow \quad assvar(\overline{\beta}_1) = \beta_2^2 \cdot \left(1 - \rho^2\right) + \sigma^2$$

For N big enough:

$$\overline{\beta}_1 \approx \mathcal{N}\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot (1-\rho^2)}{N}\right)$$

The Pre-Test Estimator has the distribution:

$$\widetilde{\beta}_1 \sim \begin{cases} \mathcal{N}\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right) \text{ with Prob.= } \Pr\left(|t|_{\widehat{\beta}_2} > 2\right) \\ \mathcal{N}\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot \left(1-\rho^2\right)}{N}\right) \text{ with Prob.= } 1 - \Pr\left(|t|_{\widehat{\beta}_2} > 2\right) \end{cases}$$

for N large enough

The bias (due to $\rho\beta_2 \neq 0$) depends on

- the level of the correlation ρ;
- the value of β_2;
- the size of $\rho\beta_2$;
- the probability $\Pr\left(|t|_{\widehat{\beta}_2} < 2\right)$

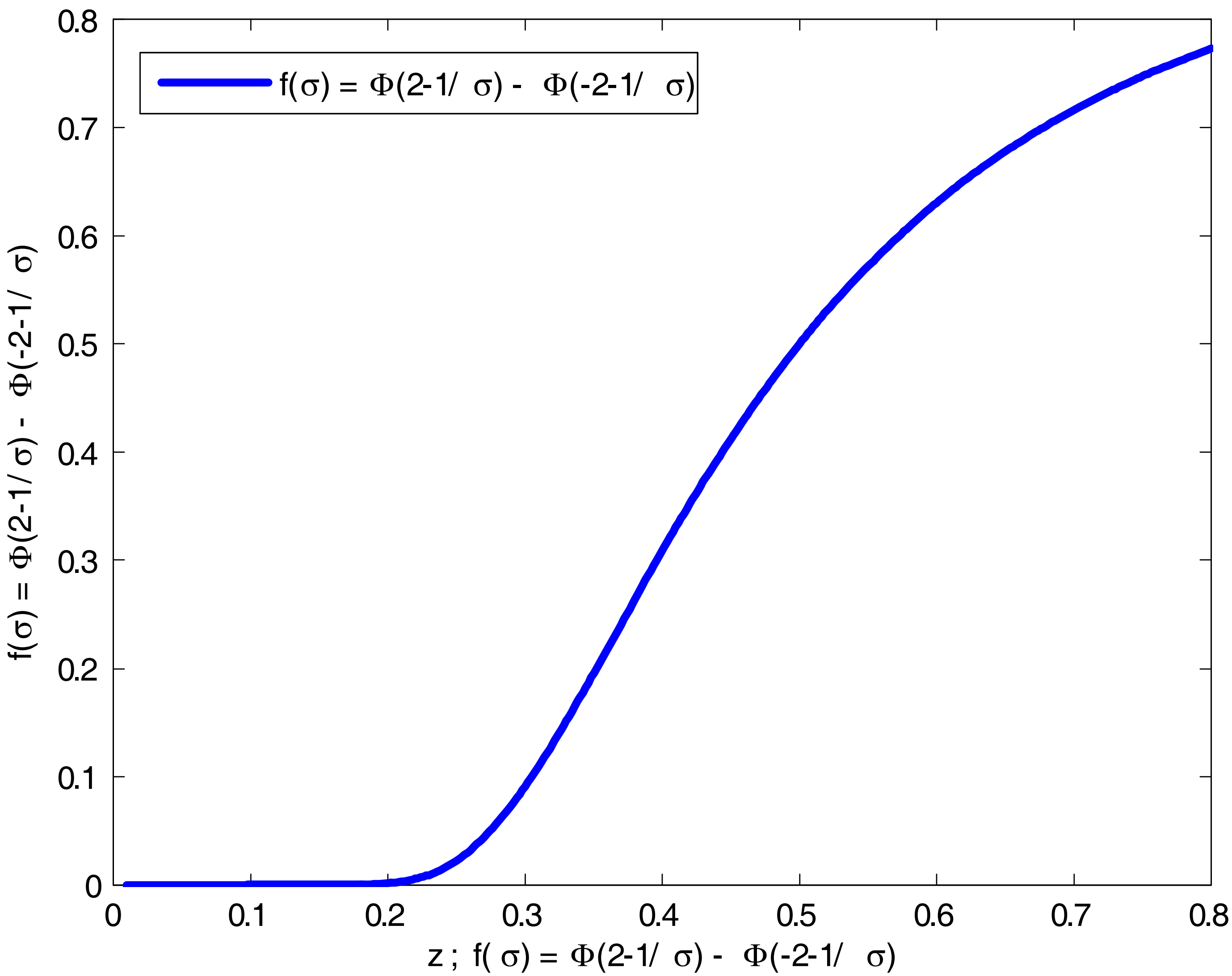
Observe that:

$$\widehat{\beta}_2 \sim \mathcal{N}\left(\beta_2, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right) \Rightarrow$$

$$\Pr\left(|t|_{\widehat{\beta}_2} < 2\right) \approx \Phi\left(2 - \frac{\beta_2}{\sqrt{\frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}}}\right) - \Phi\left(-2 - \frac{\beta_2}{\sqrt{\frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}}}\right)$$

The variance of ε_i which affects $|t|_{\widehat{\beta}_2}$; $\sigma^2 \uparrow \Rightarrow \Pr\left(|t|_{\widehat{\beta}_2} < 2\right) \uparrow$
The Expectation of $E(X_i'X_i)$ which affects $|t|_{\widehat{\beta}_2}$.
An increase in $|\rho|$ increases the variance of $\widehat{\beta}_2$ which increases $\Pr\left(|t|_{\widehat{\beta}_2} < 2\right)$

Example of Function



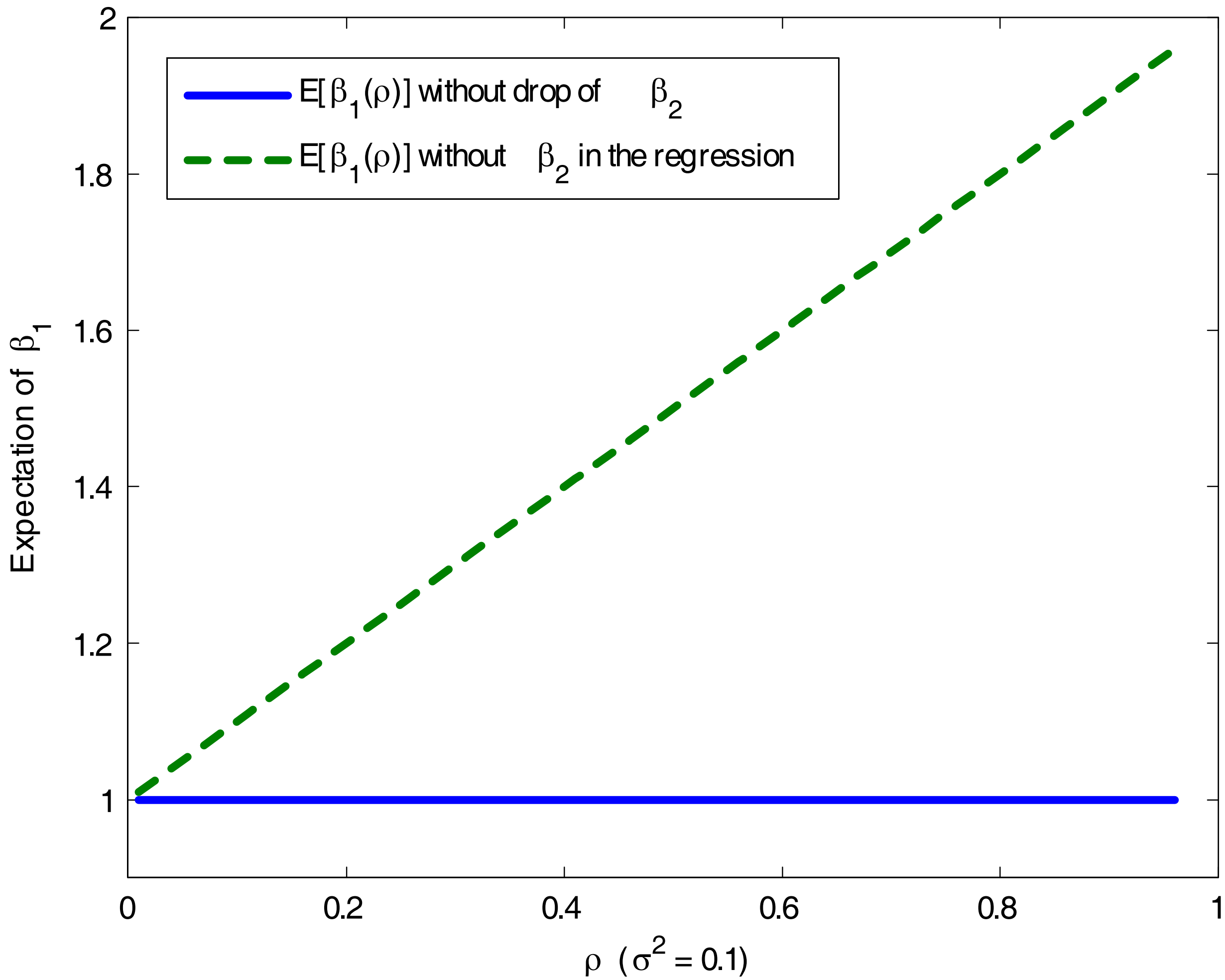
$$f(\sigma) = \Phi\left(2 - \frac{1}{\sigma}\right) - \Phi\left(-2 - \frac{1}{\sigma}\right)$$

Suppose:

$$
\begin{aligned}
Y_i &= X_i \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon_i;\ i = 1, ..., N; \\
\varepsilon_i &\sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}; \quad \sigma^2 = 0.1 \\
X_i &\equiv \begin{bmatrix} X_{1i} \\ X_{2i} \end{bmatrix} \sim \mathcal{N}\begin{bmatrix} 0, & \begin{matrix} 1 & \rho \\ \rho & 1 \end{matrix} \end{bmatrix} \quad \text{i.i.d.}; \ \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}
$$

We obtain the results:

Expectation of $\widehat{\beta}_1$, $\overline{\beta}_1$,



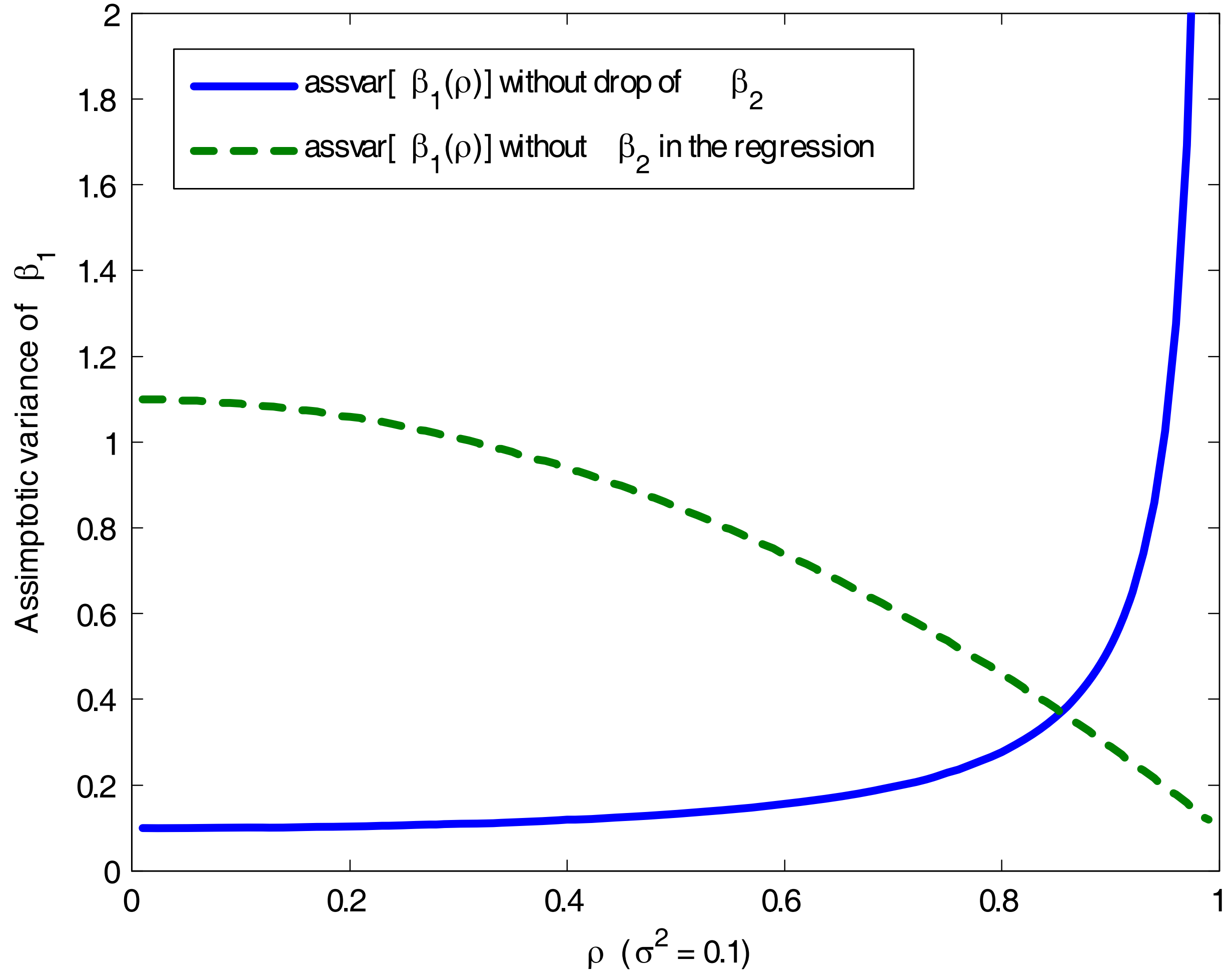
Expectation of β_1

$\rho = 0.5; \quad \sigma^2 = 0.1$

$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right); \beta_1 = \beta_2 = 1$$

$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot \left(1-\rho^2\right)}{N}\right)$$

Asymptotic Variance of $\widehat{\beta}_1$, $\overline{\beta}_1$,



Expectation of β_1

$\rho = 0.5; \quad \sigma^2 = 0.1$

$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right); \beta_1 = \beta_2 = 1$$

$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot (1-\rho^2)}{N}\right)$$

4 The Parameter Analysis of Procedures 1 and 3

This section provides the following facts for $\rho > 0$, $\beta_2 > 0$.

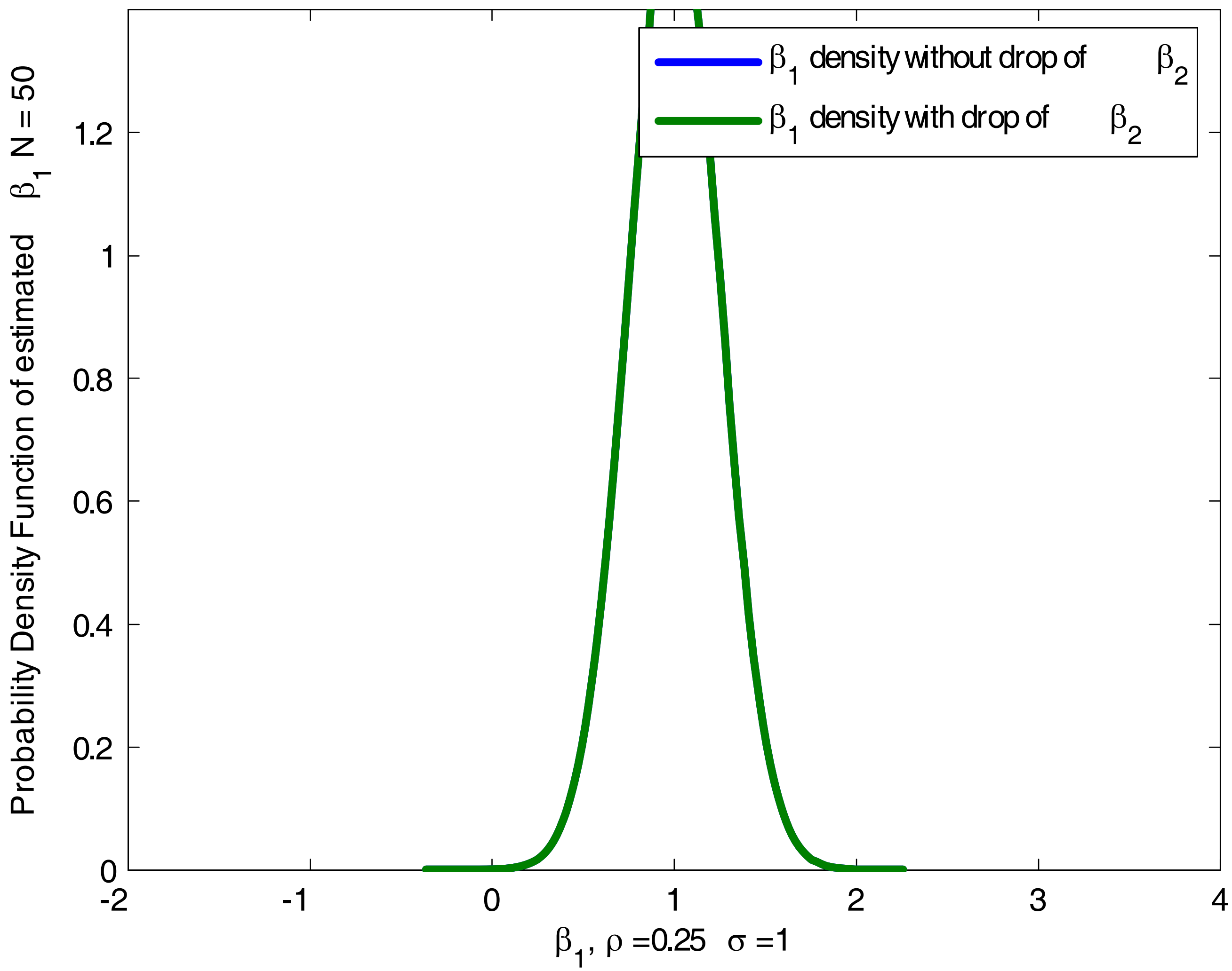
1. The increase in the error variance makes the OLS and Pre-Test Estimators less precise.

2. The increase in the error variance makes the Pre-Test more biased.

3. The increase in the correlation among variables X makes the OLS and Pre-Test Estimators less precise.

4. The increase in the correlation among variables X makes Pre-Test Estimators more biased.

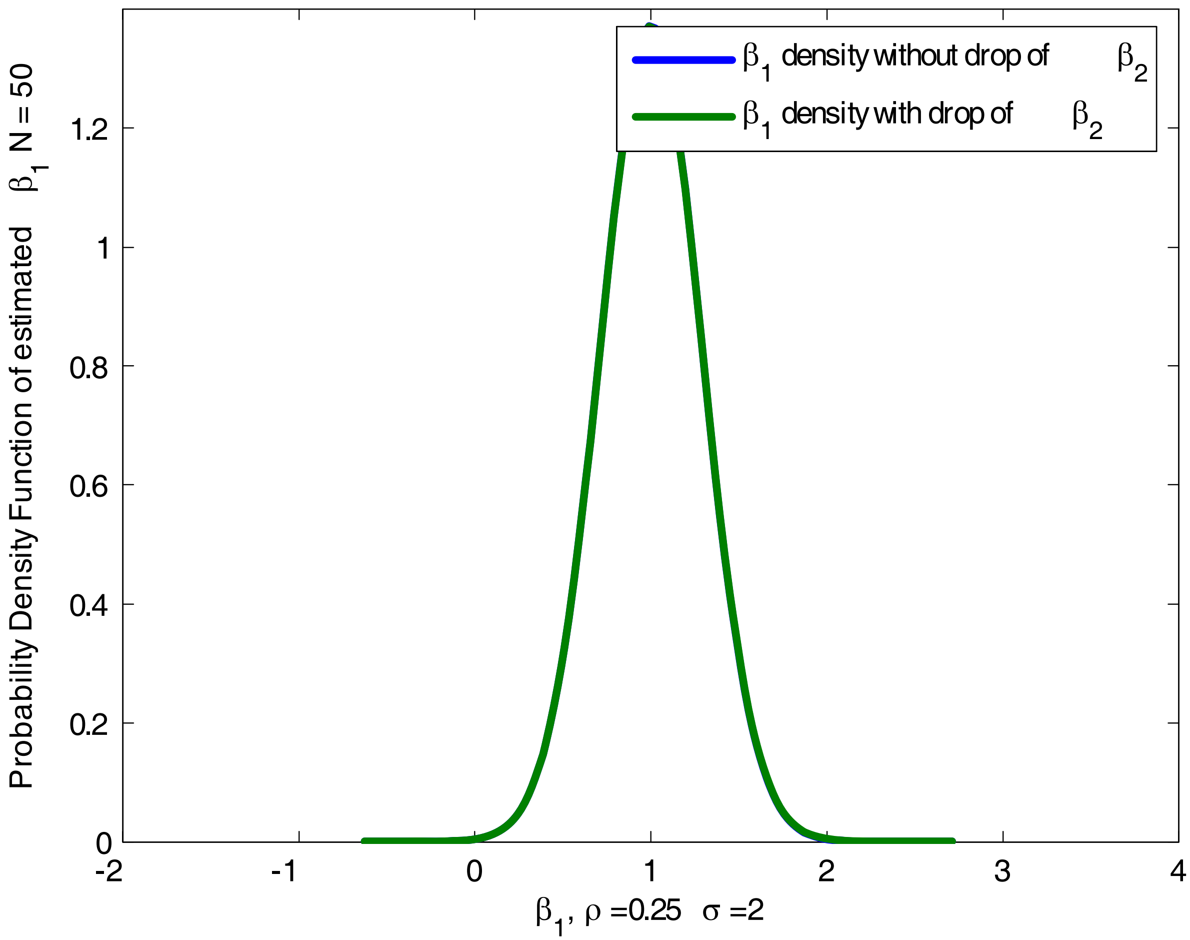
Pre-Test Estimator Density Function



Distribution of β_1

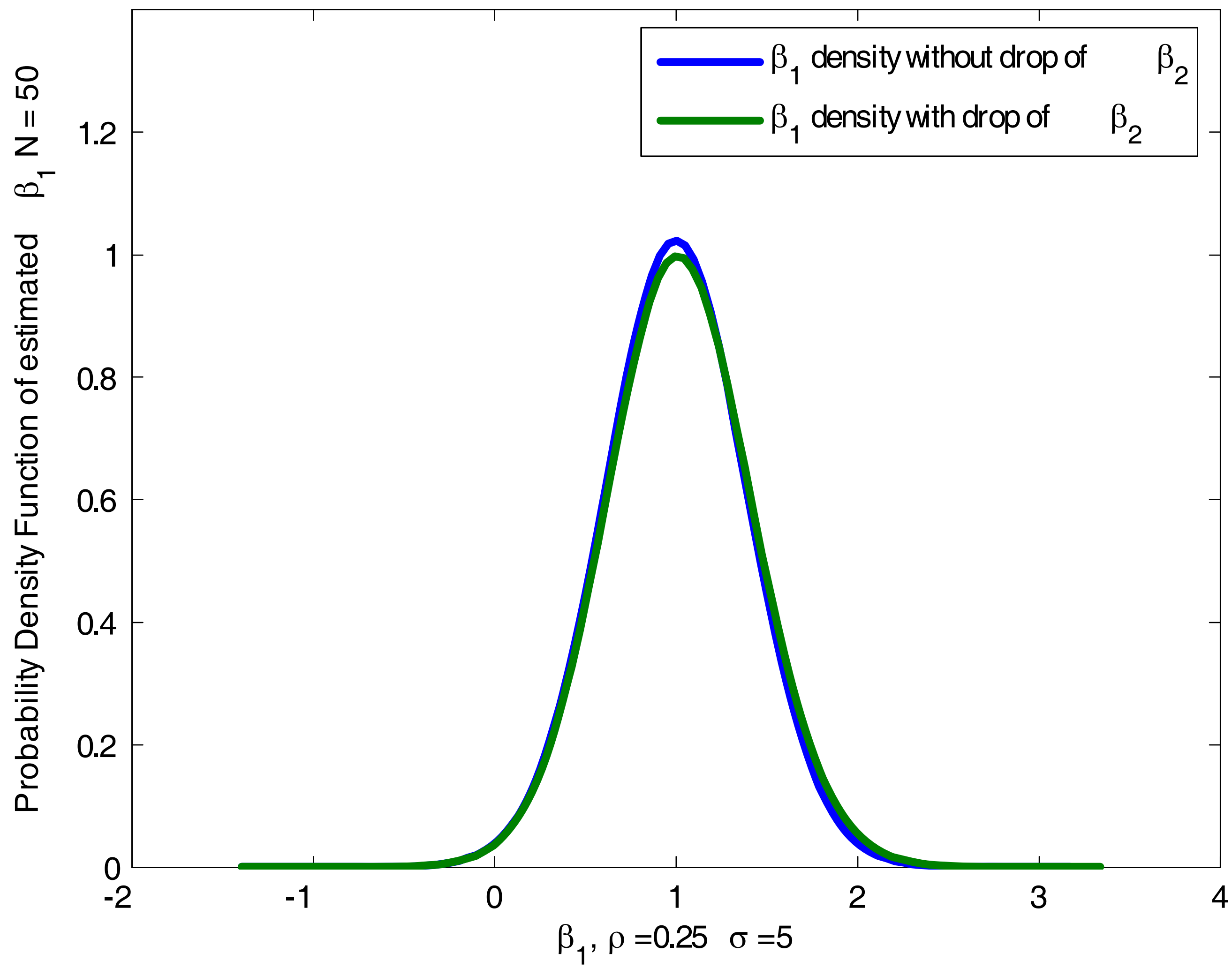
$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 1$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 2$$

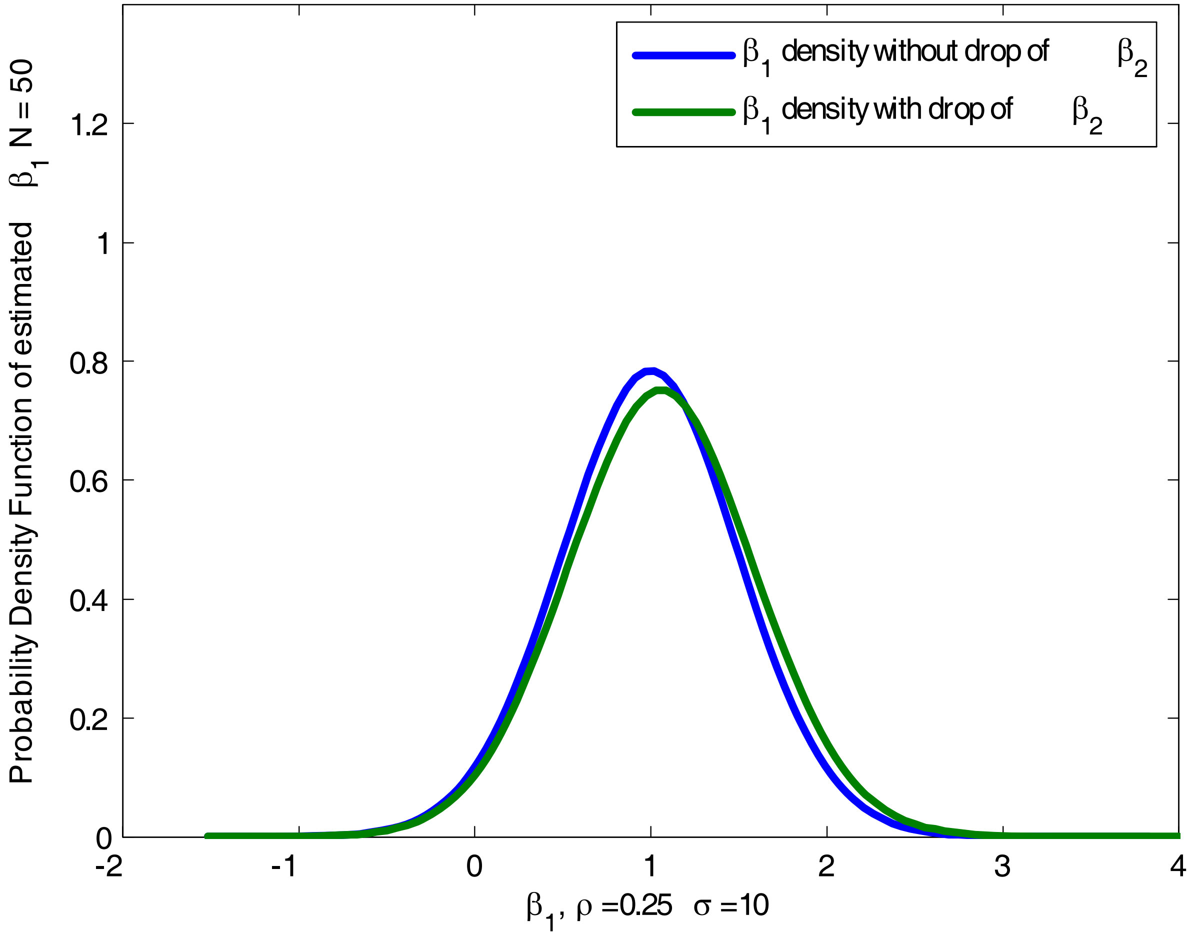
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 5$

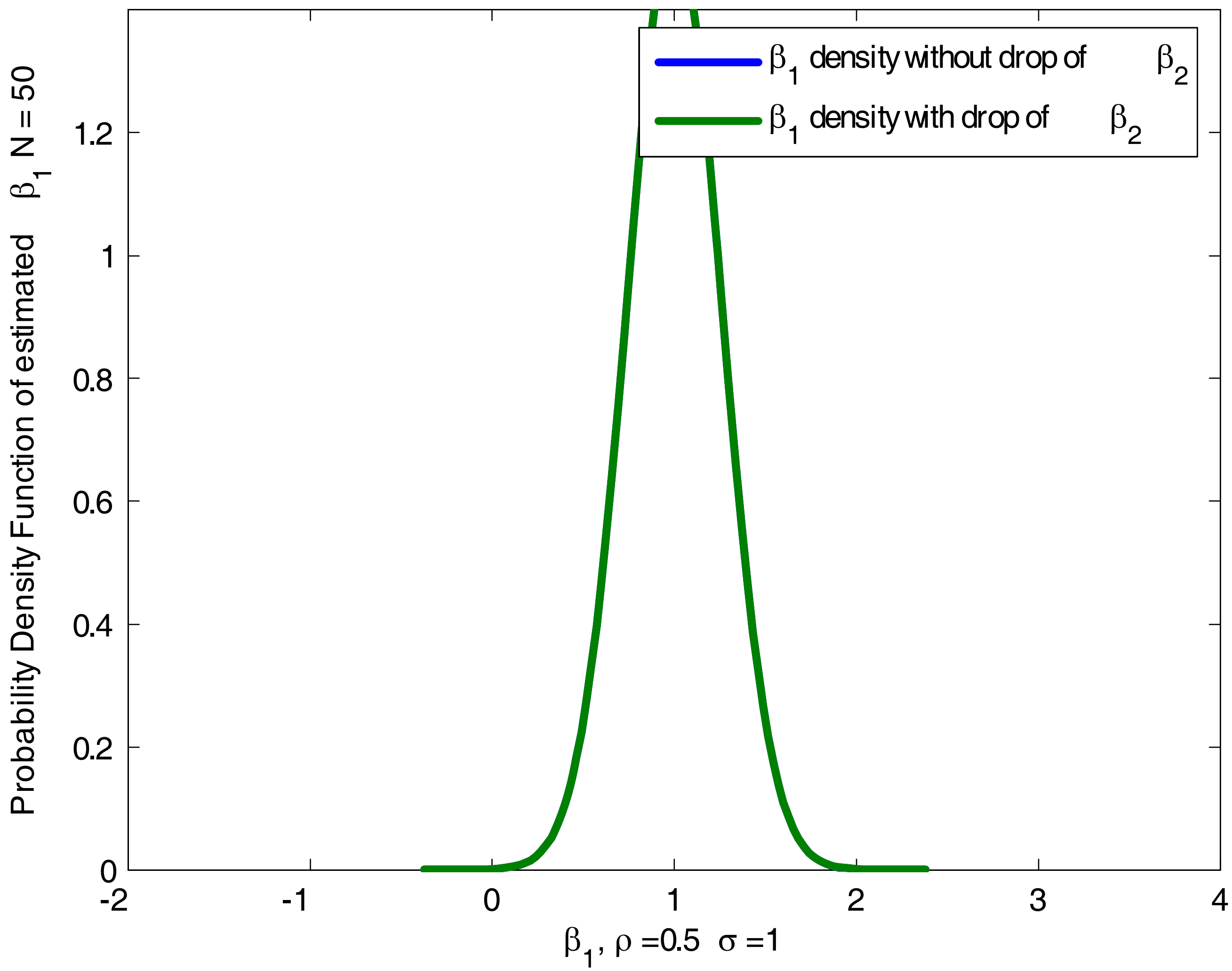
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 10$

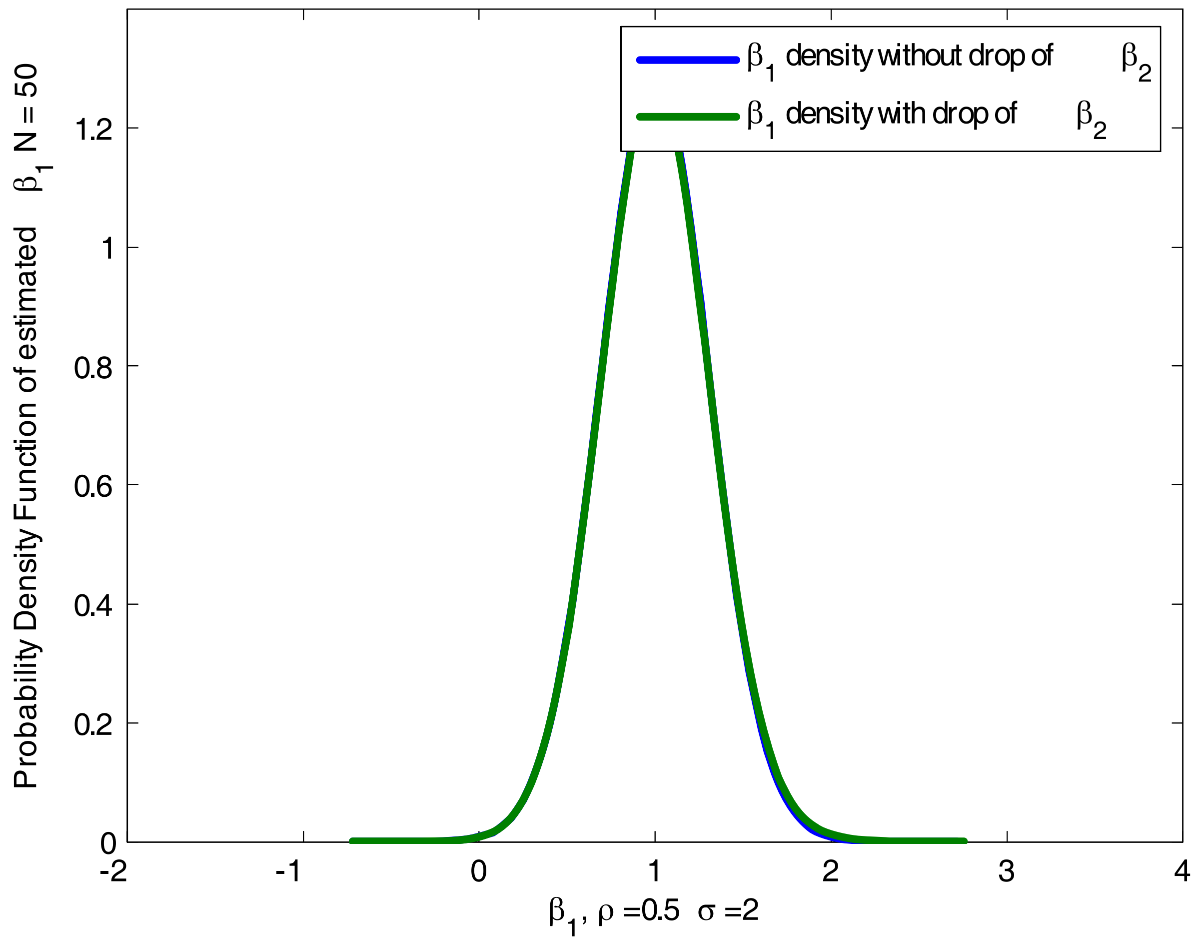
Pre-Test Estimator Density Function



Distribution of β_1

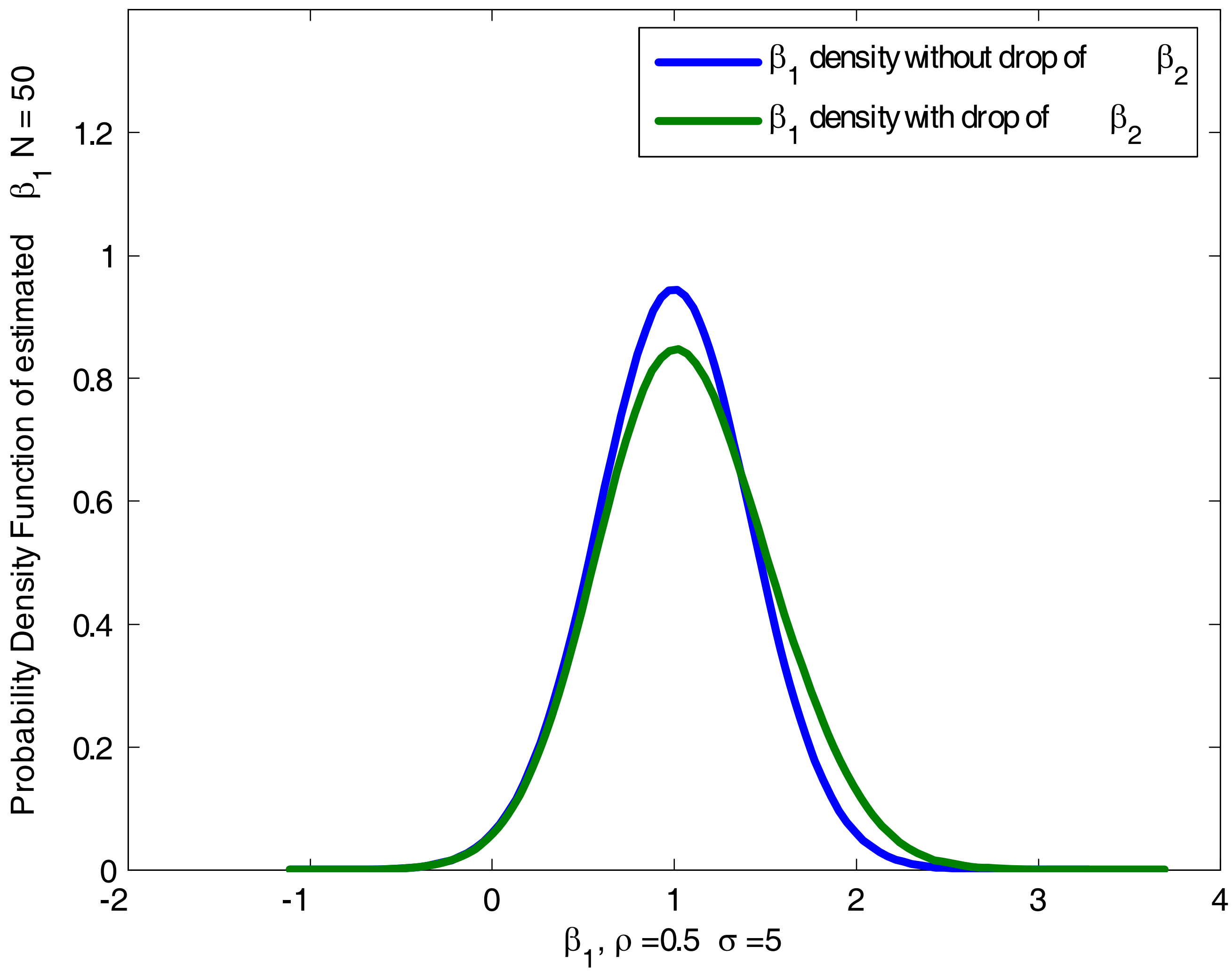
$$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 1$$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 2$$

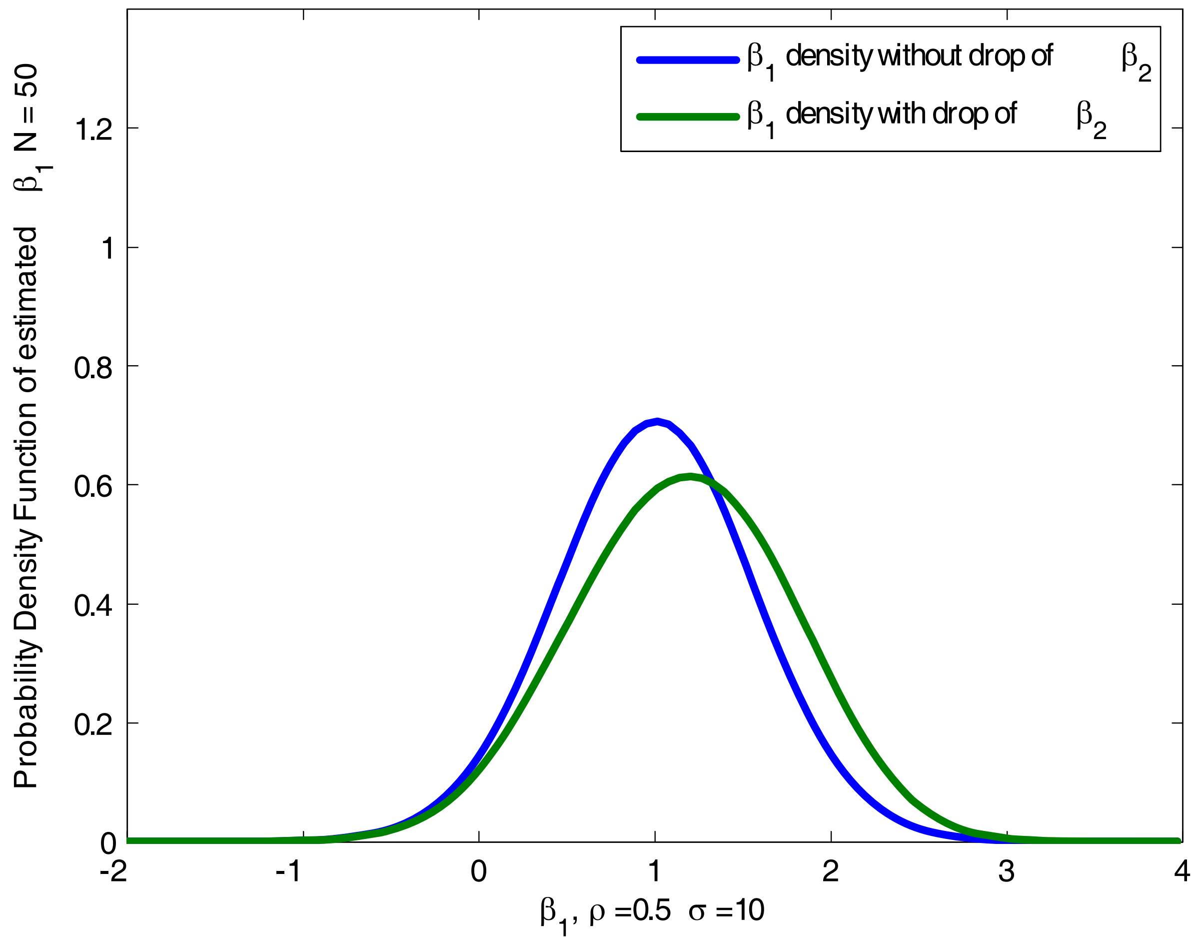
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 5$

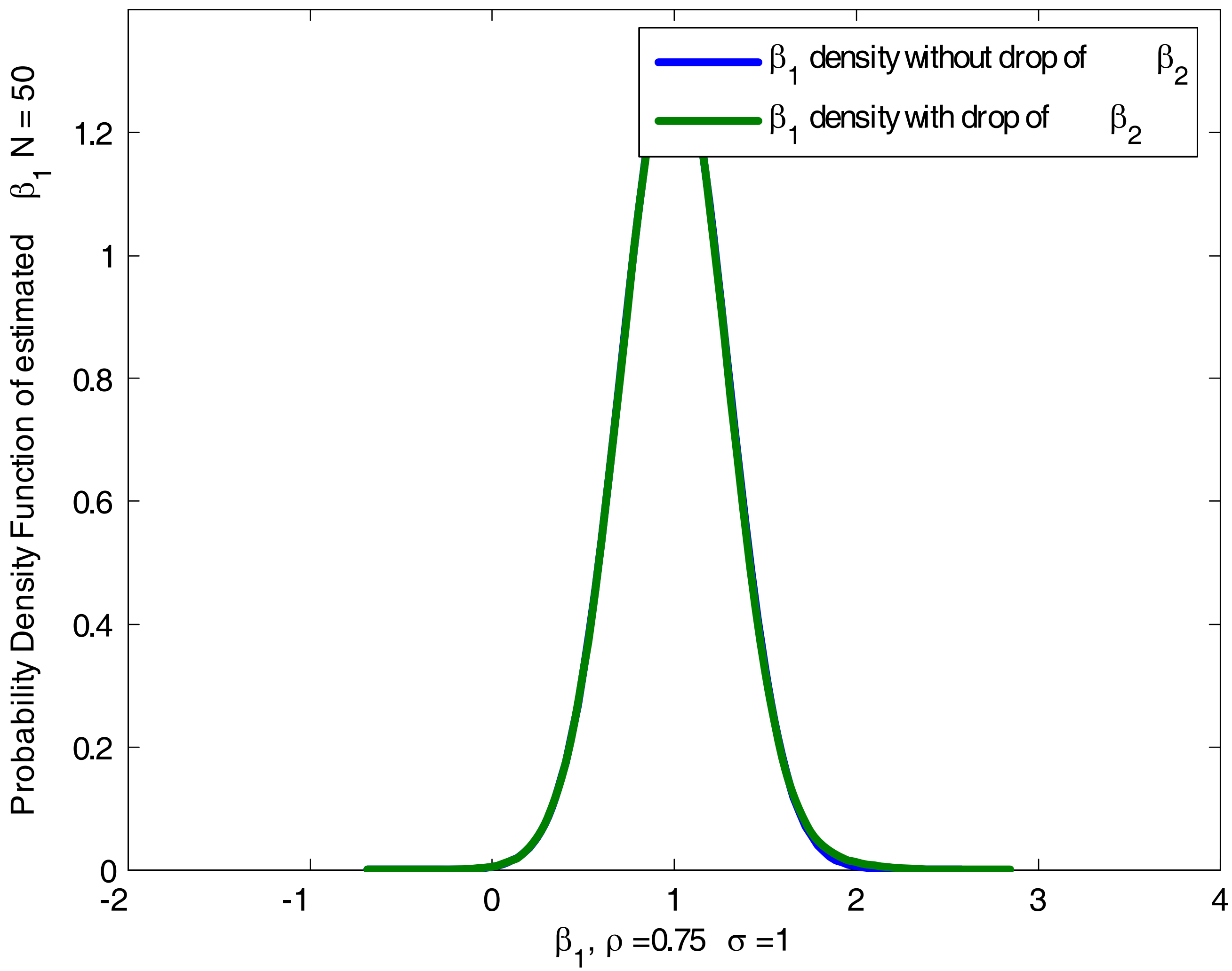
Pre-Test Estimator Density Function



Distribution of β_1

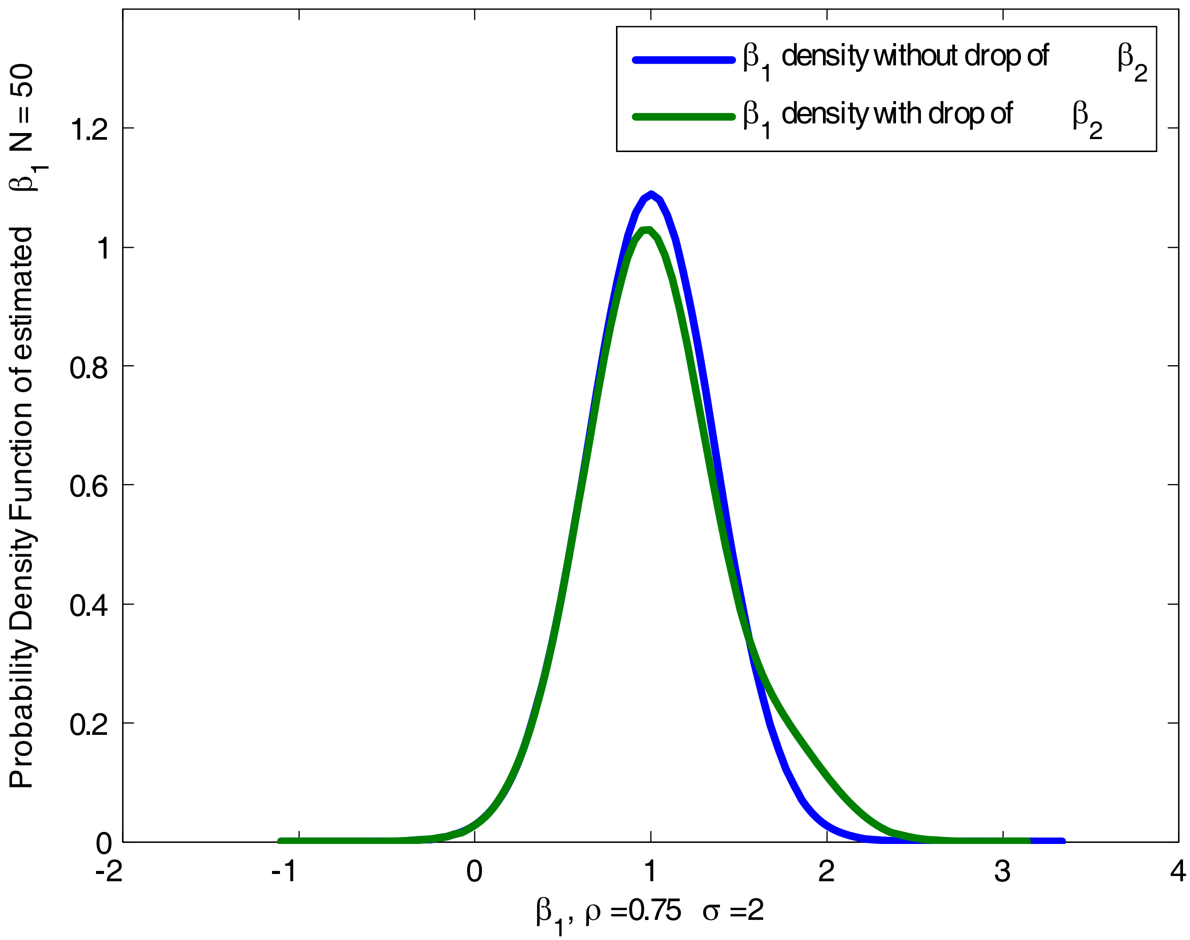
$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 10$

Pre-Test Estimator Density Function



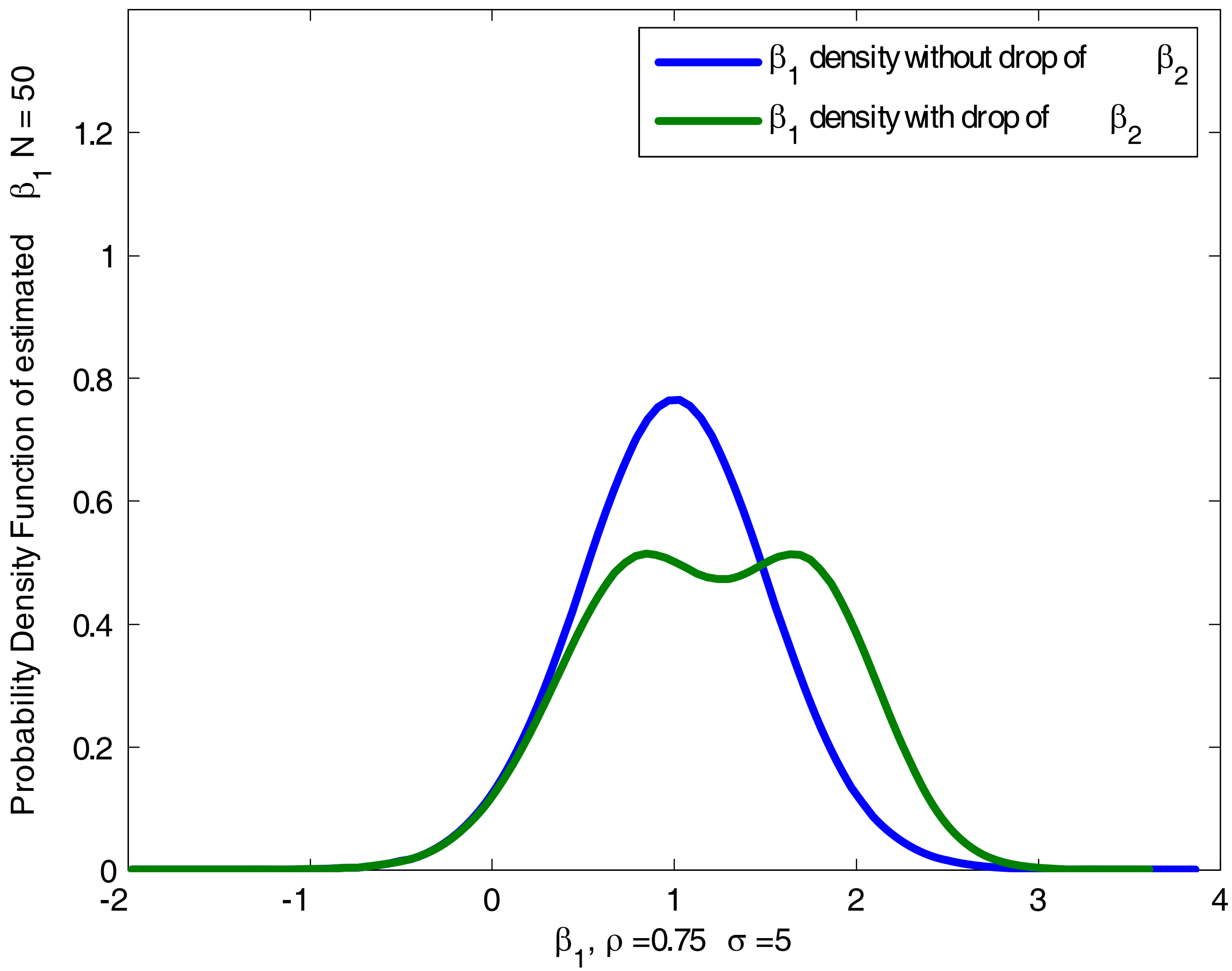
$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 1$$

Pre-Test Estimator Density Function



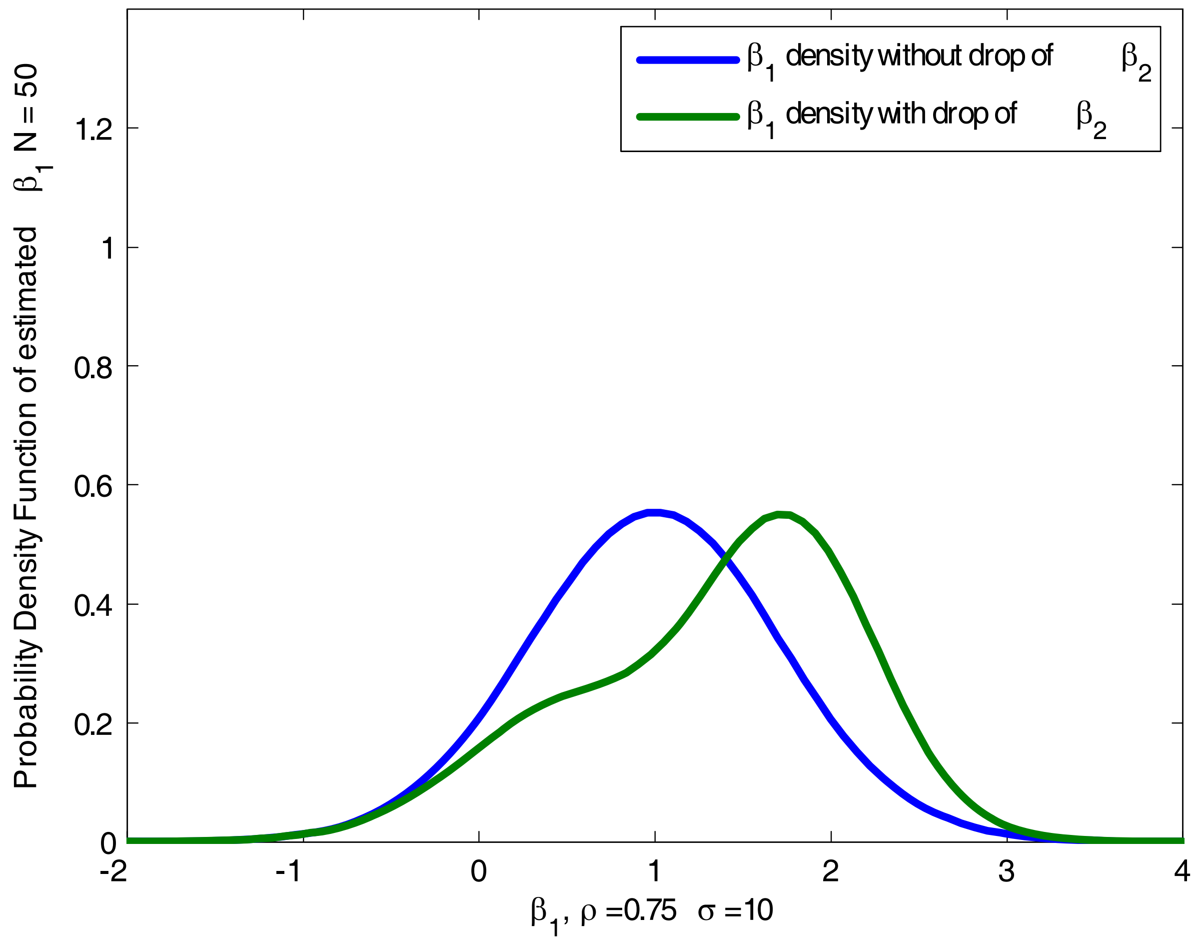
$$\textit{Distribution of } \beta_1$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 2$$

Pre-Test Estimator Density Function



$$\textit{Distribution of } \beta_1$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 5$$

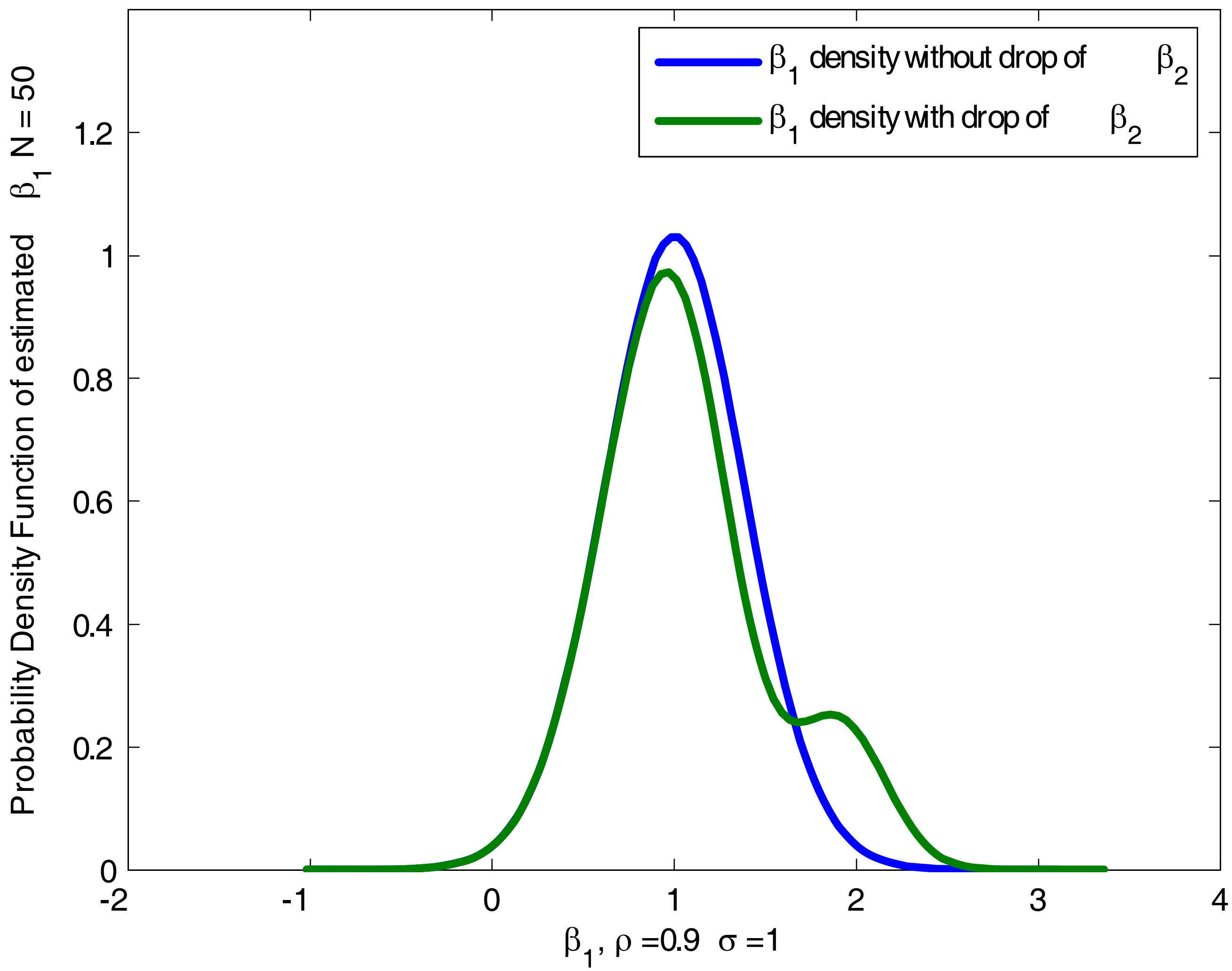
Pre-Test Estimator Density Function



Distribution of β_1

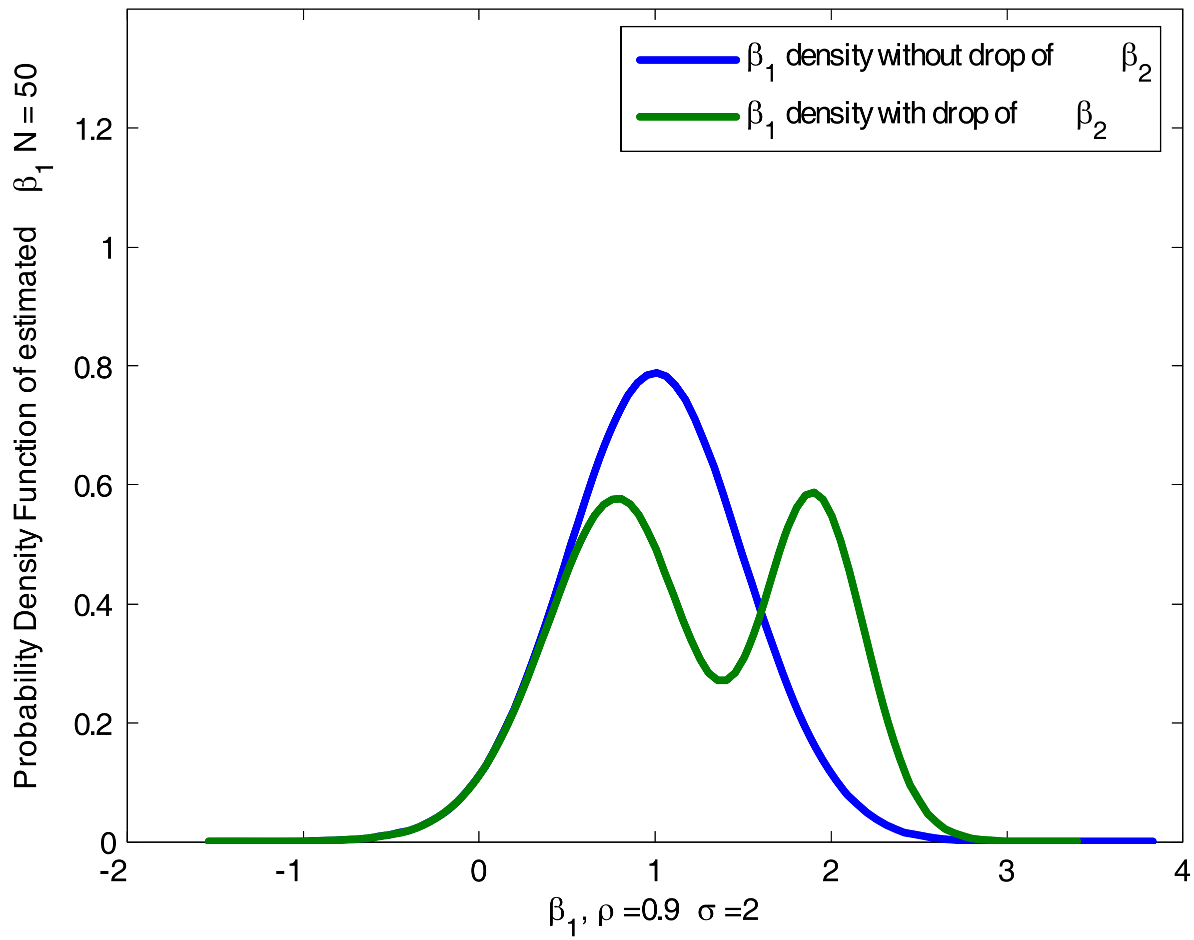
$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 10$

Pre-Test Estimator Density Function



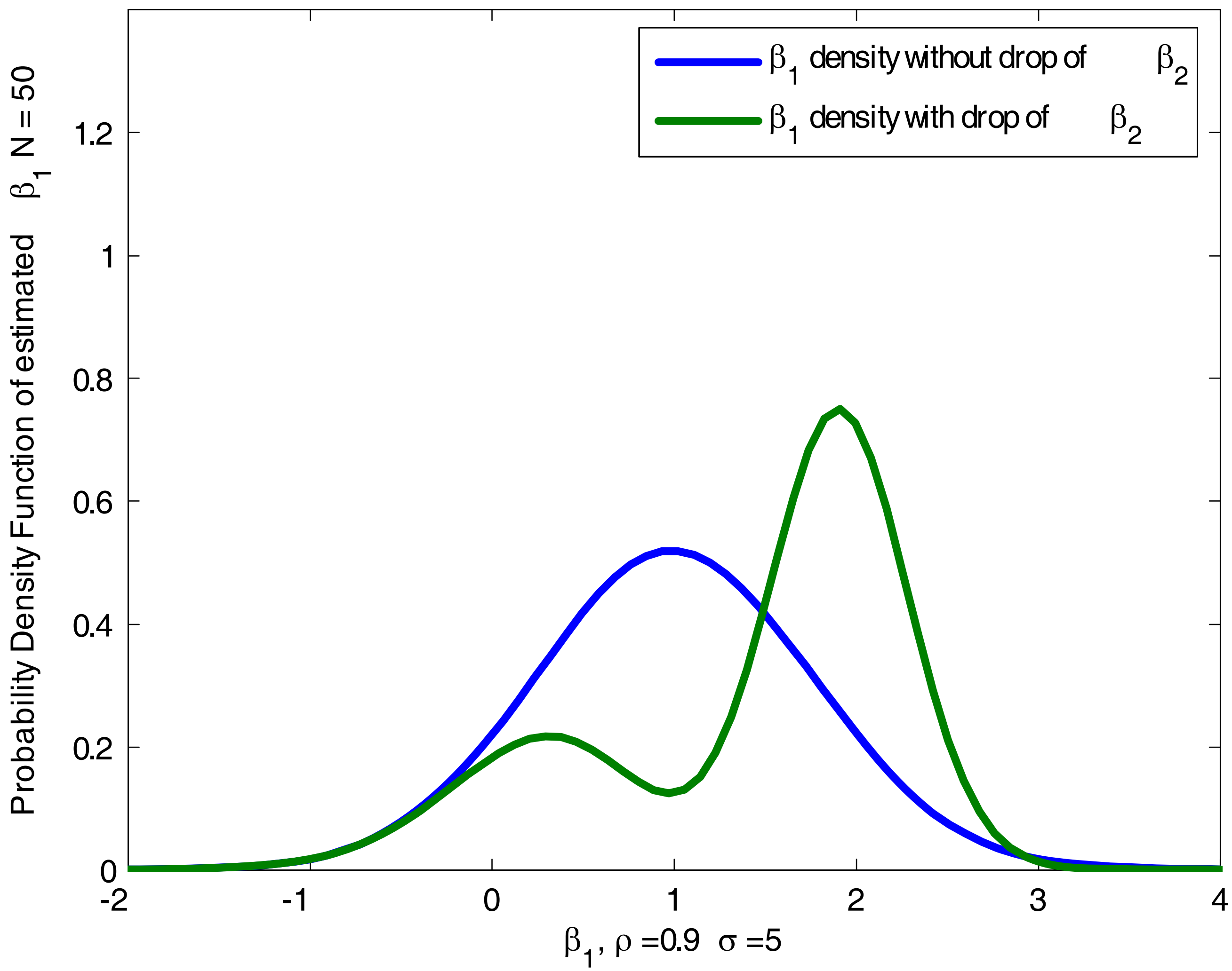
$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 1$$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 2$$

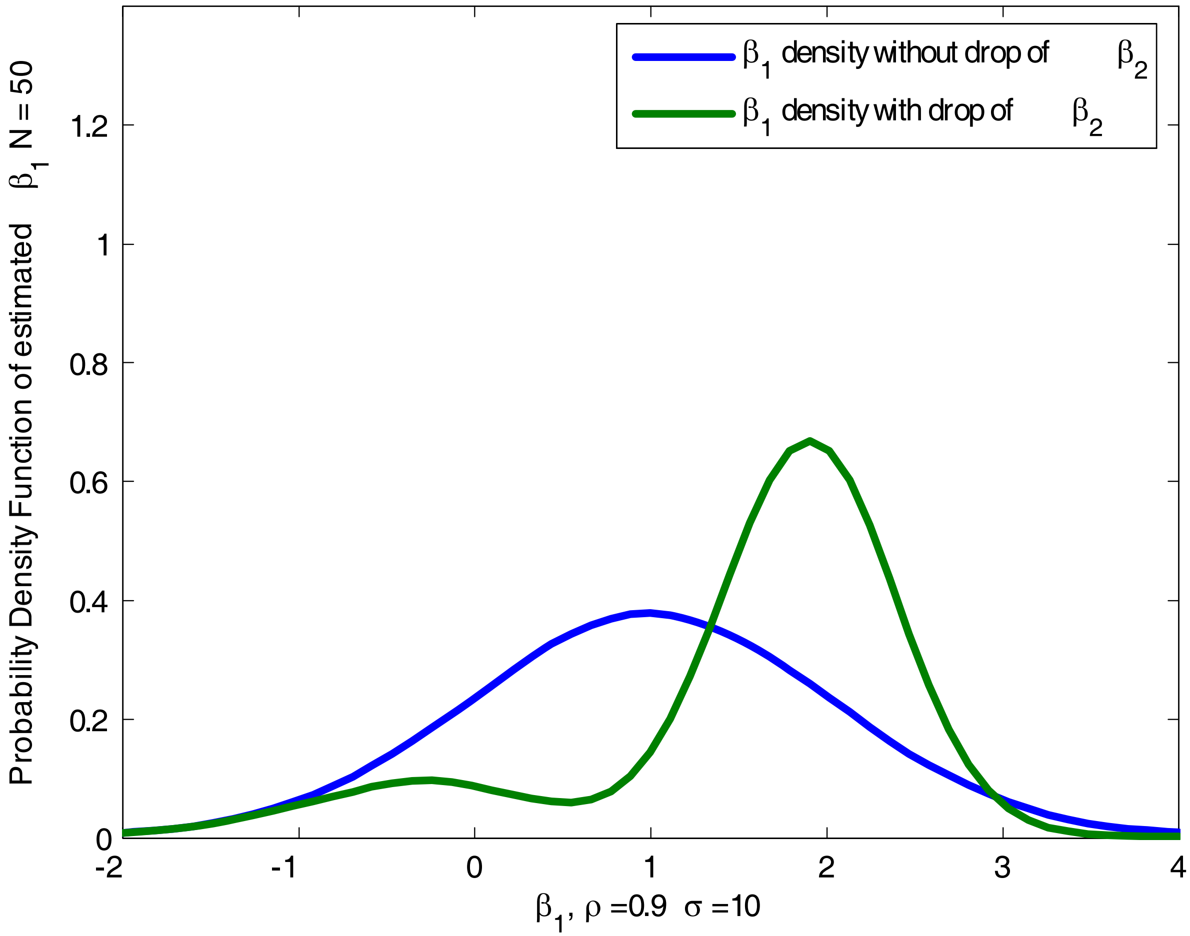
Pre-Test Estimator Density Function



$Distribution\ of\ \beta_1$

$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 5$

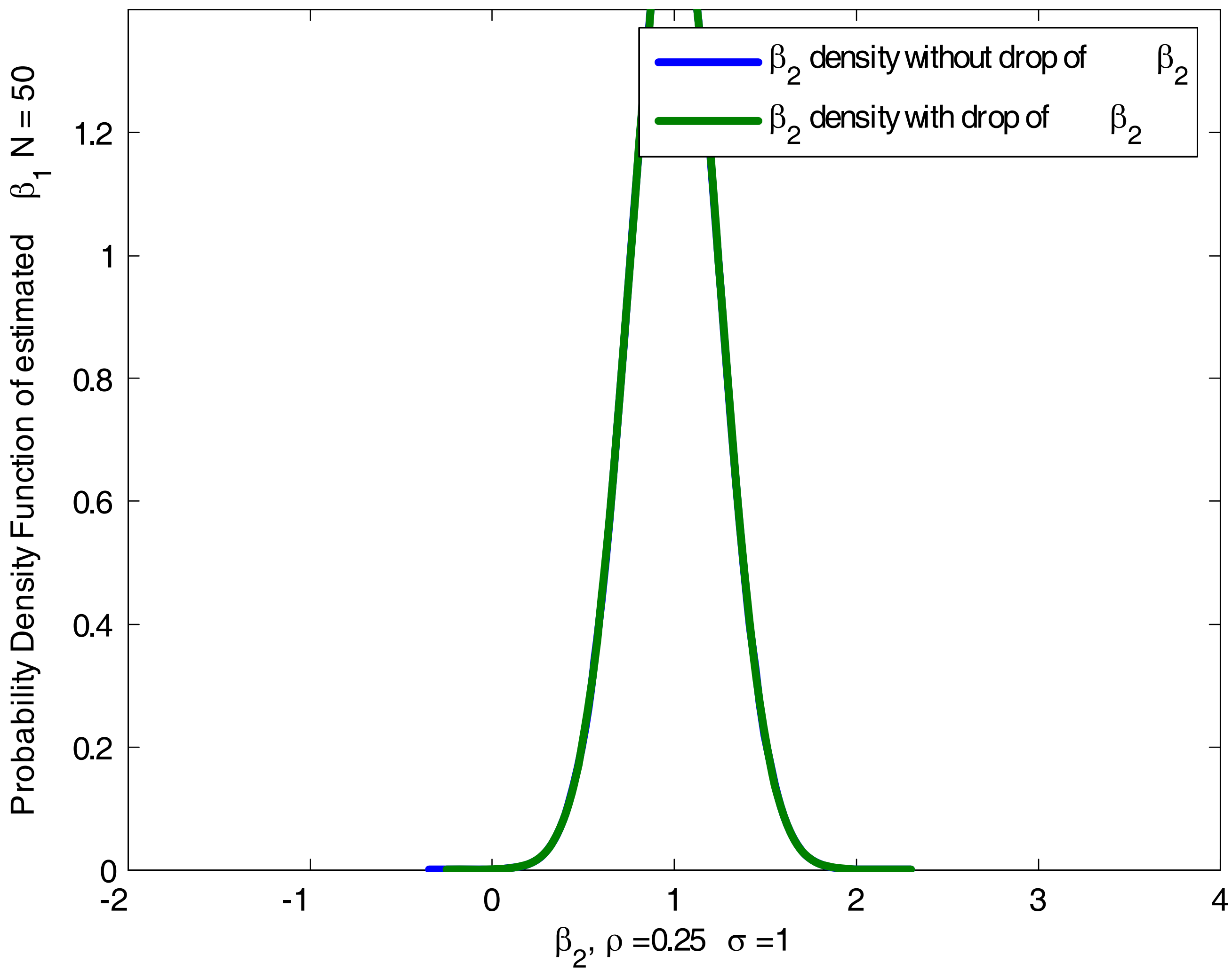
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 10$

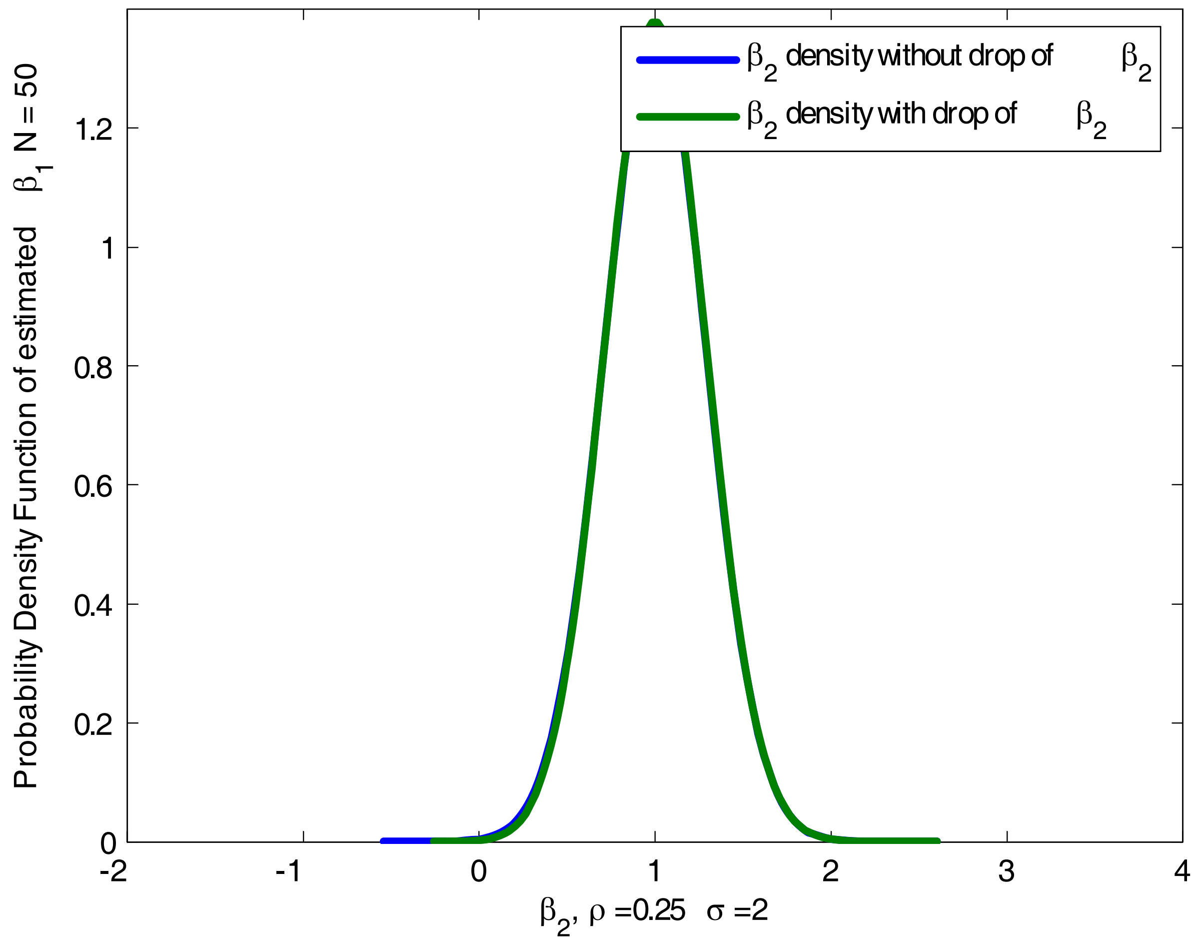
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 1$

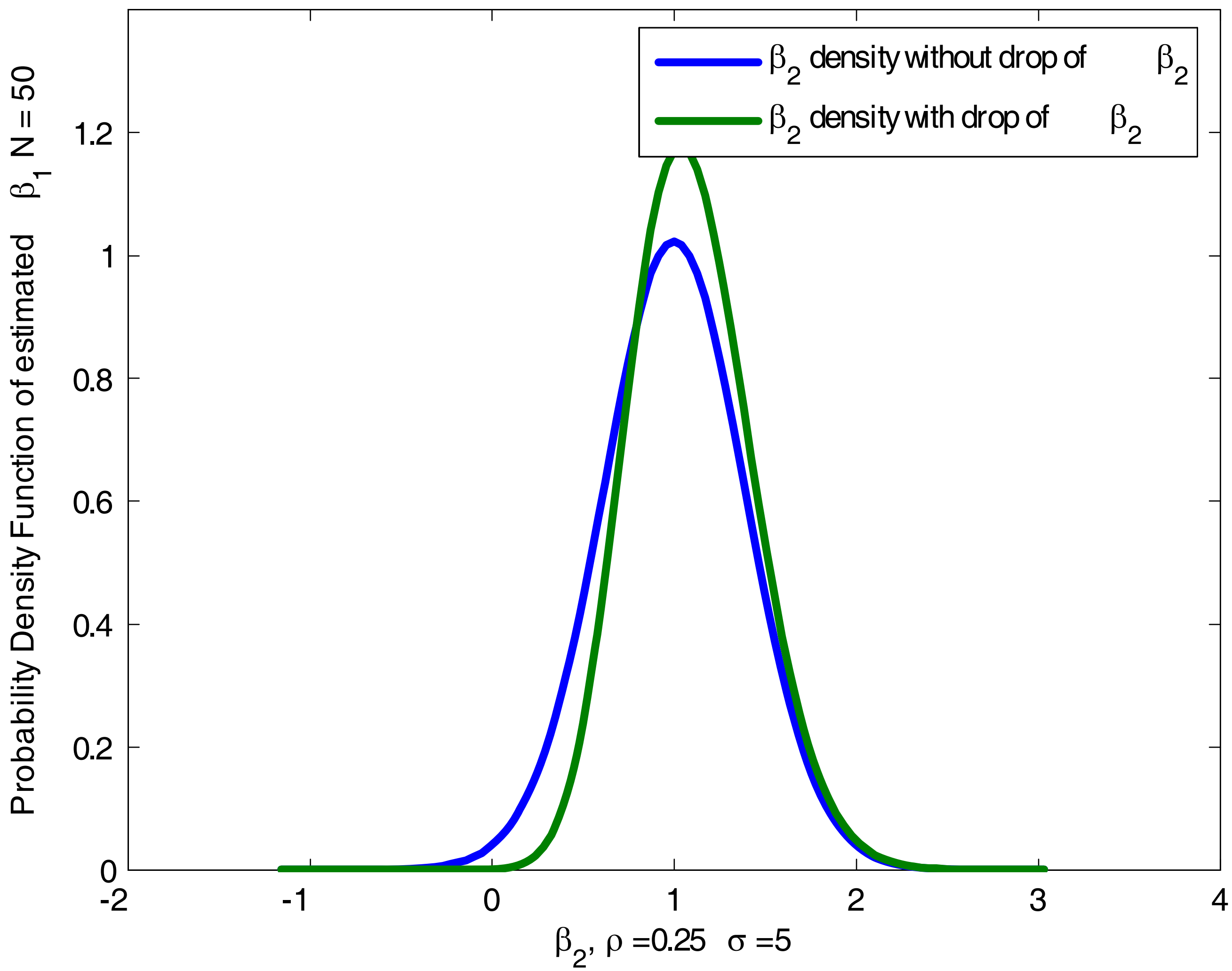
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 2$

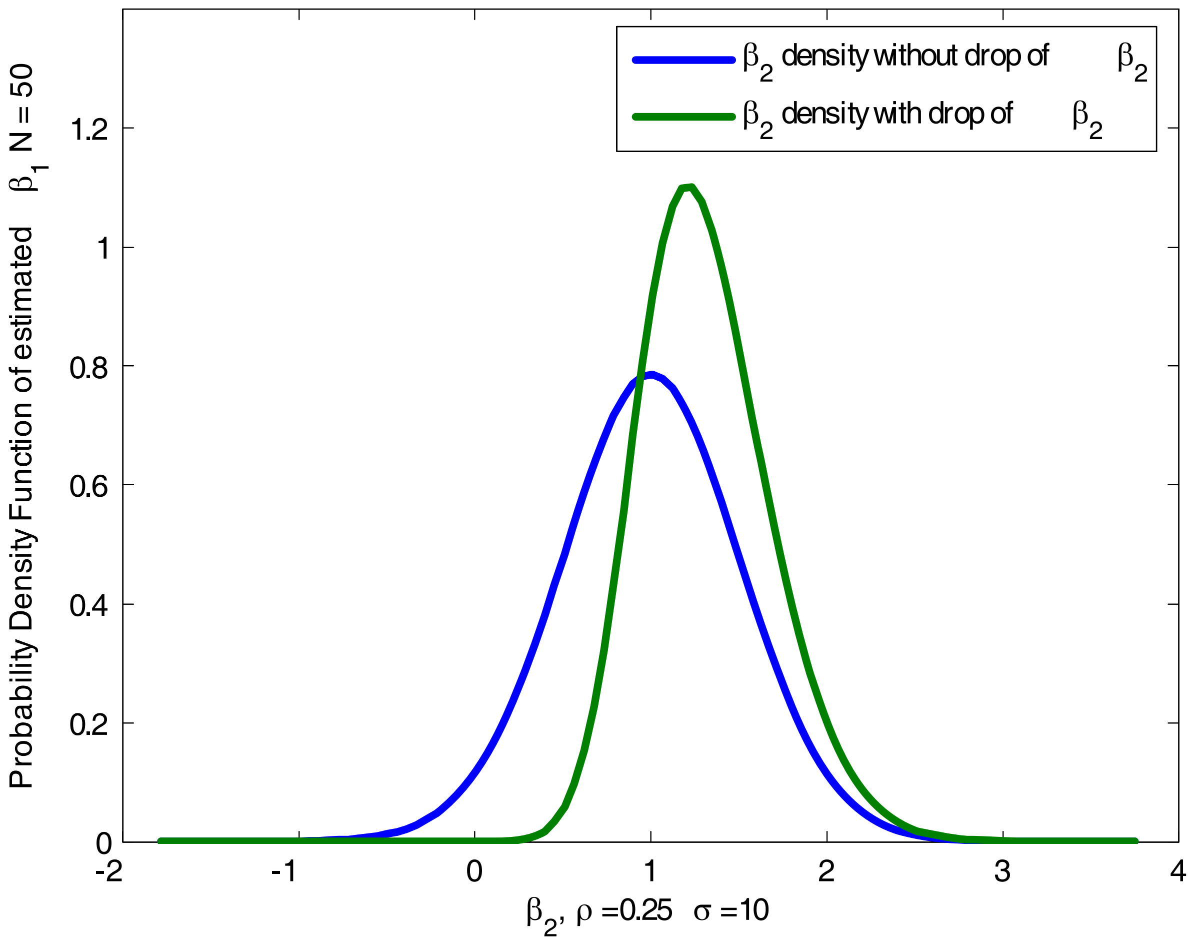
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 5$

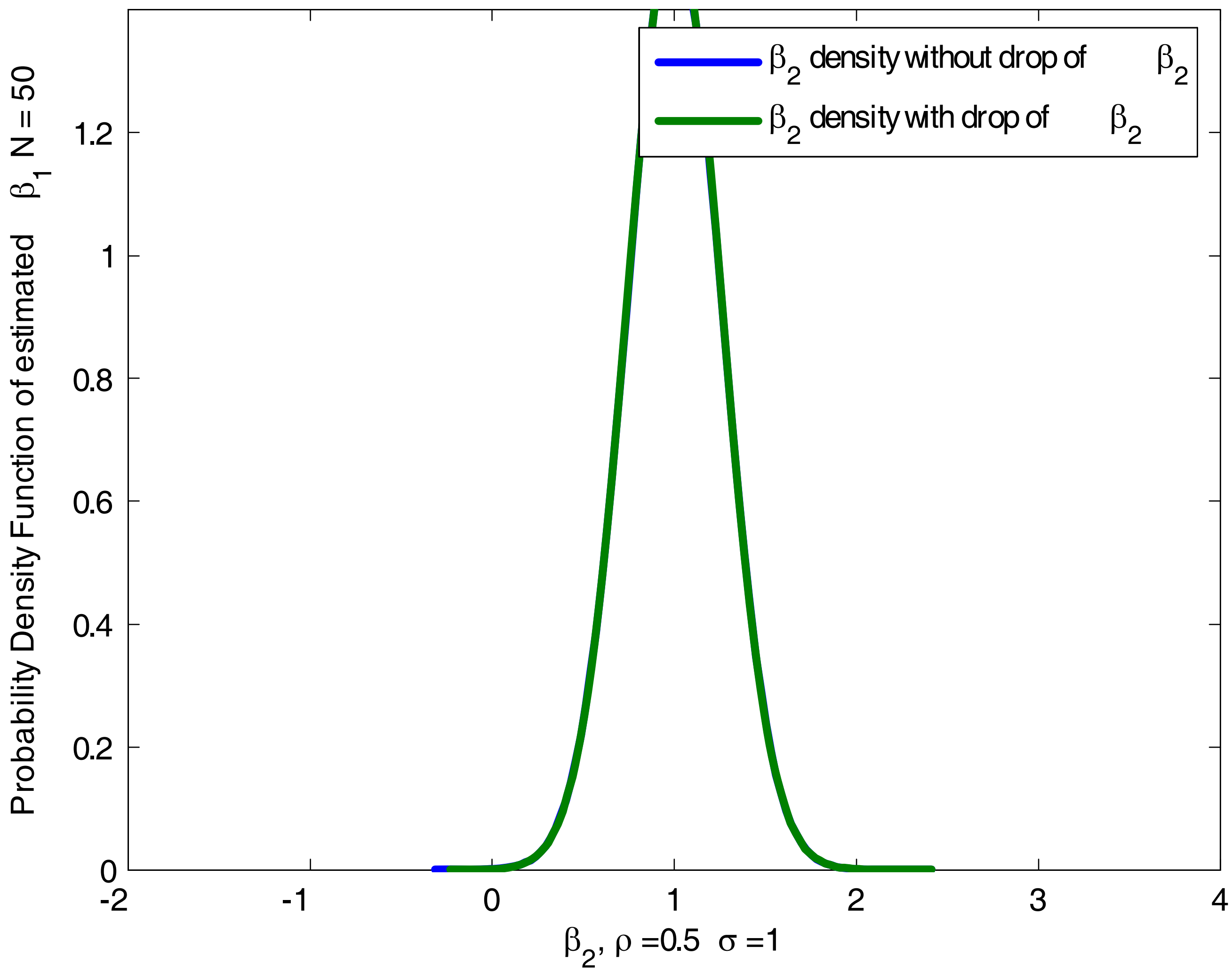
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 10$

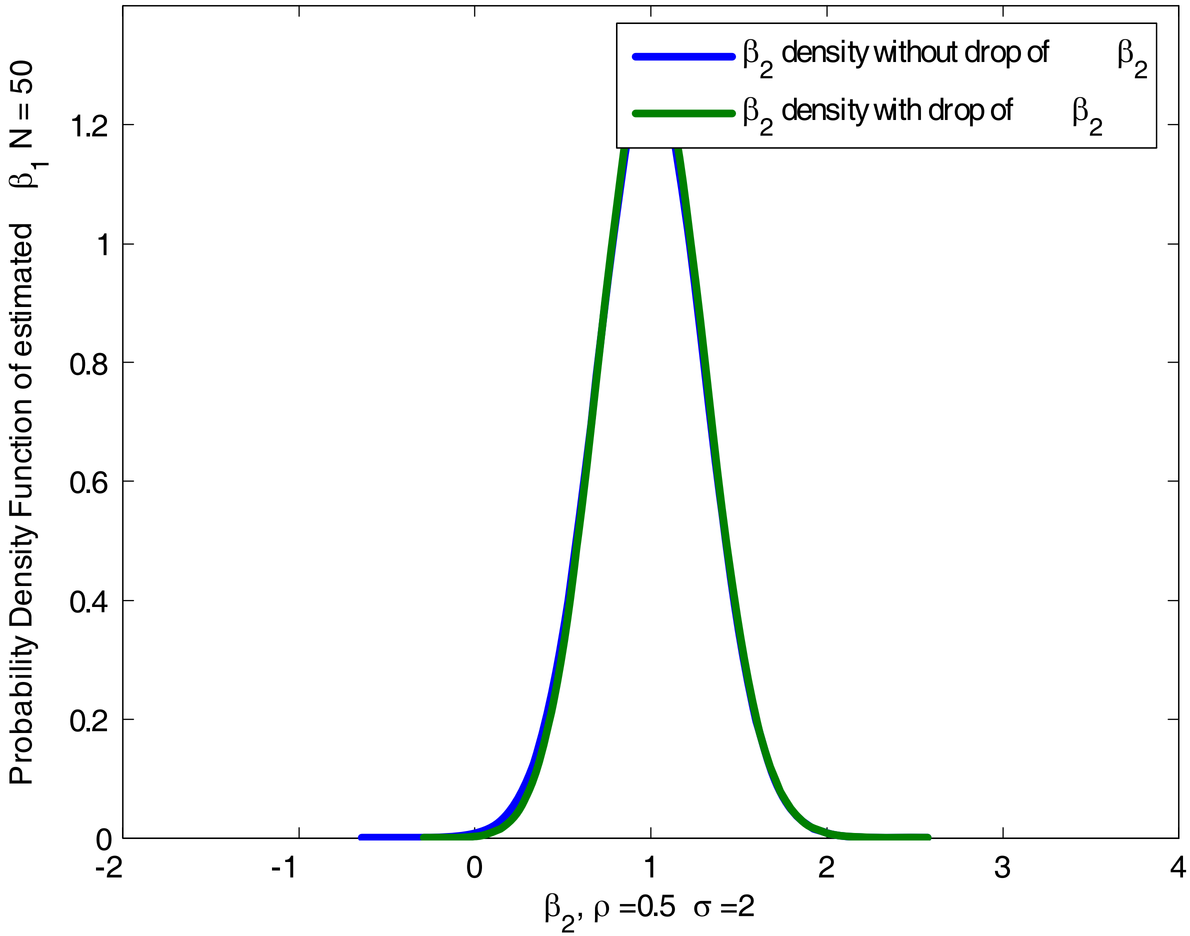
Pre-Test Estimator Density Function



Distribution of β_2

$$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 1$$

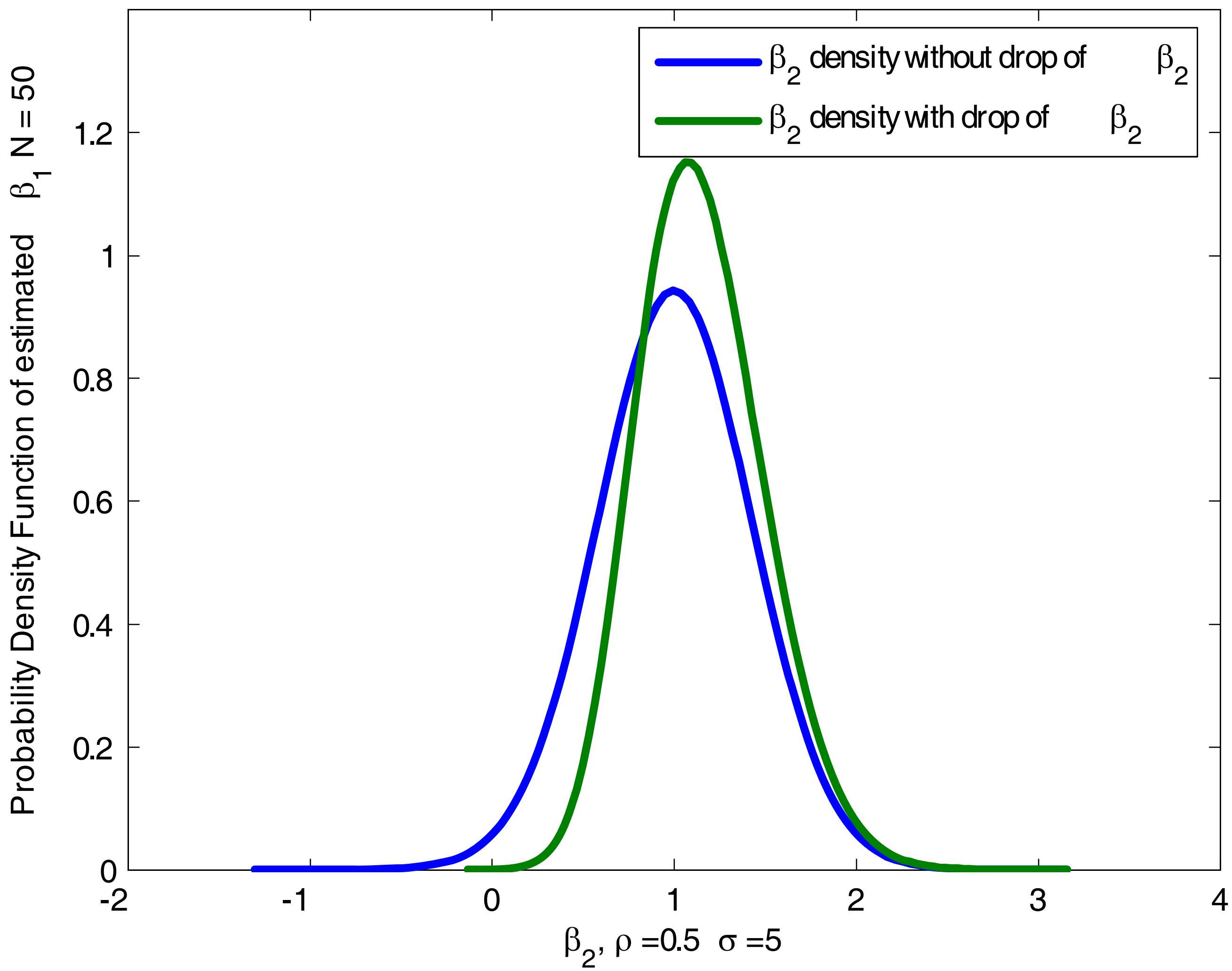
Pre-Test Estimator Density Function



Distribution of β_2

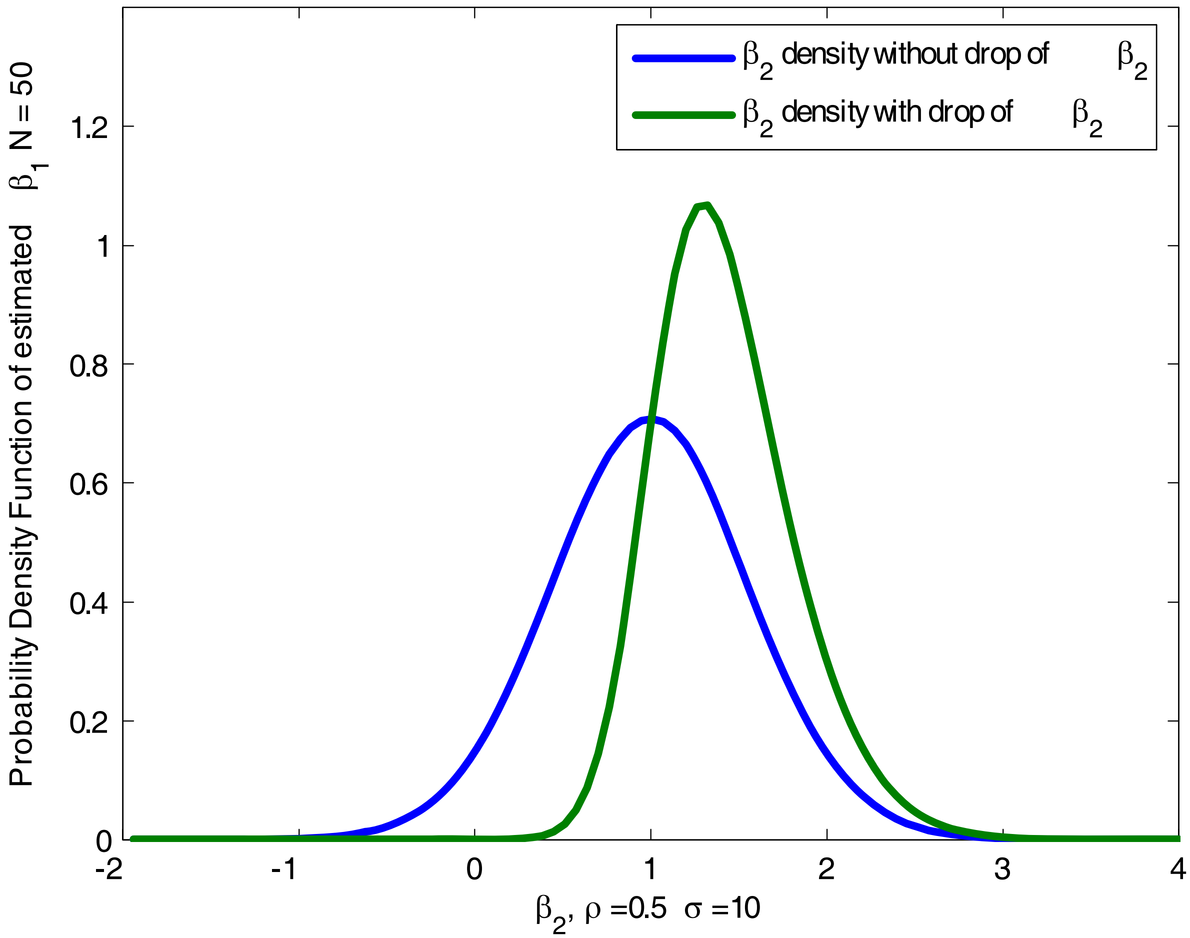
$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 2$

Pre-Test Estimator Density Function



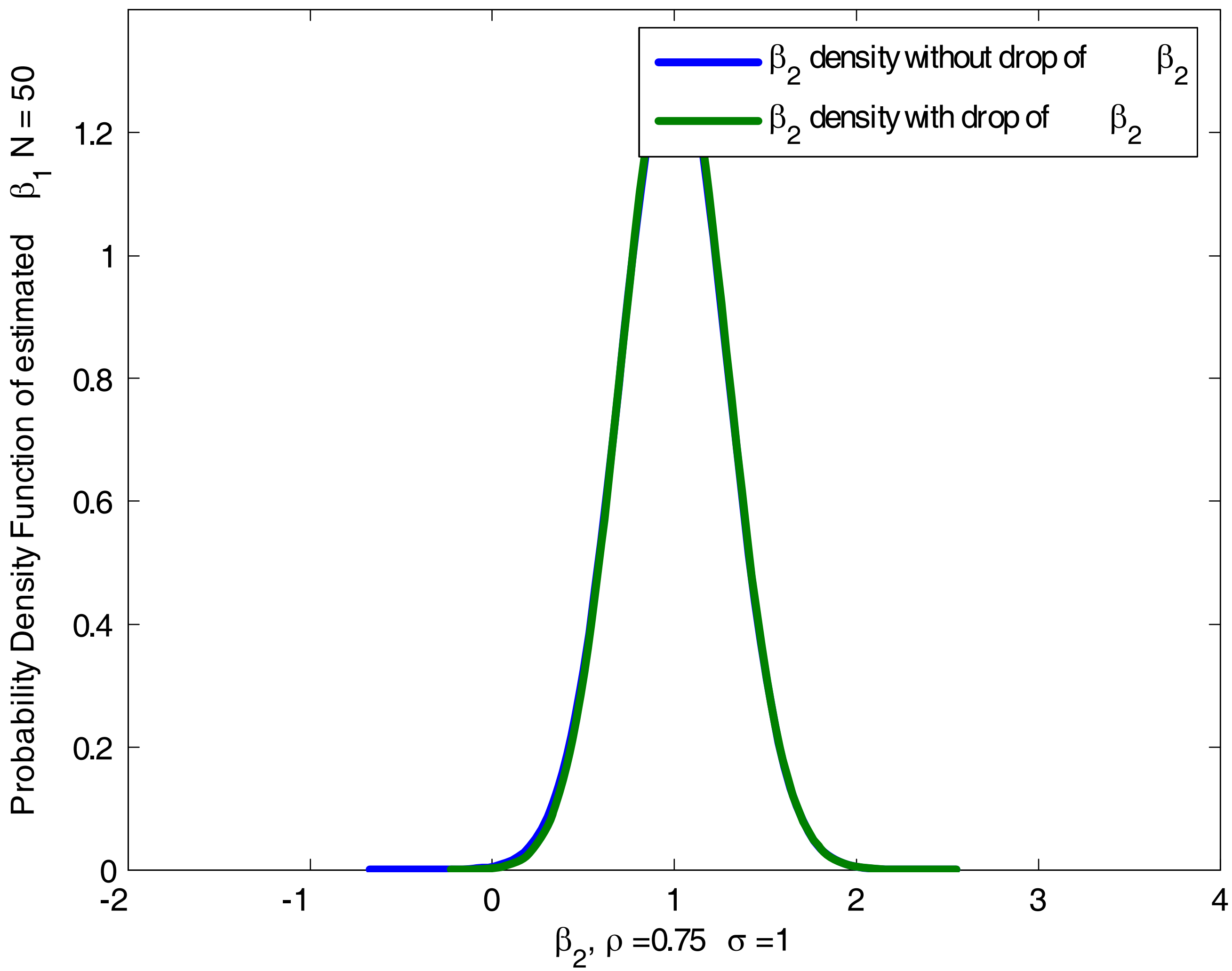
$$Distribution\ of\ \beta_2$$
$$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 5$$

Pre-Test Estimator Density Function



Distribution of β_2
$N = 50, \quad \rho = 0.5; \quad \sigma^2 = 10$

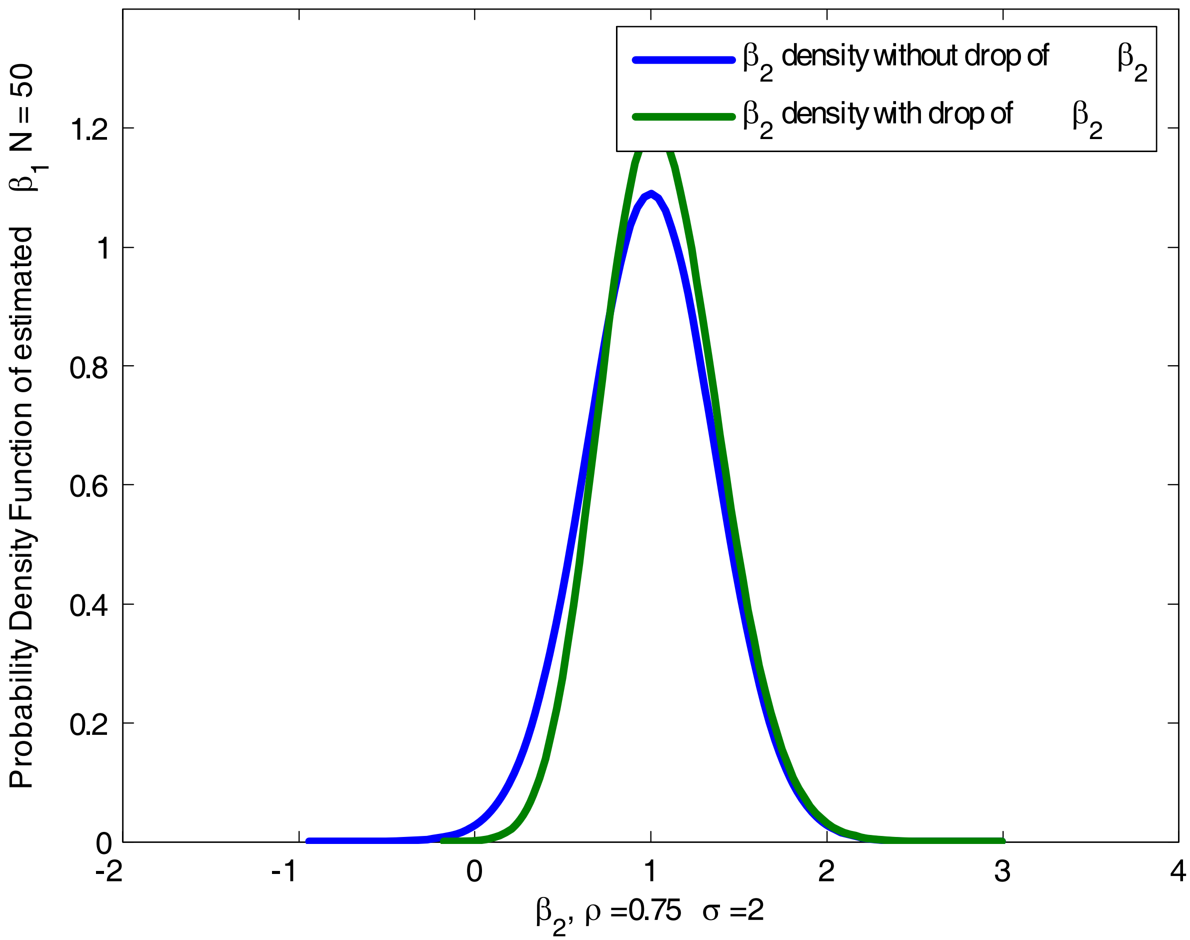
Pre-Test Estimator Density Function



Distribution of β_2

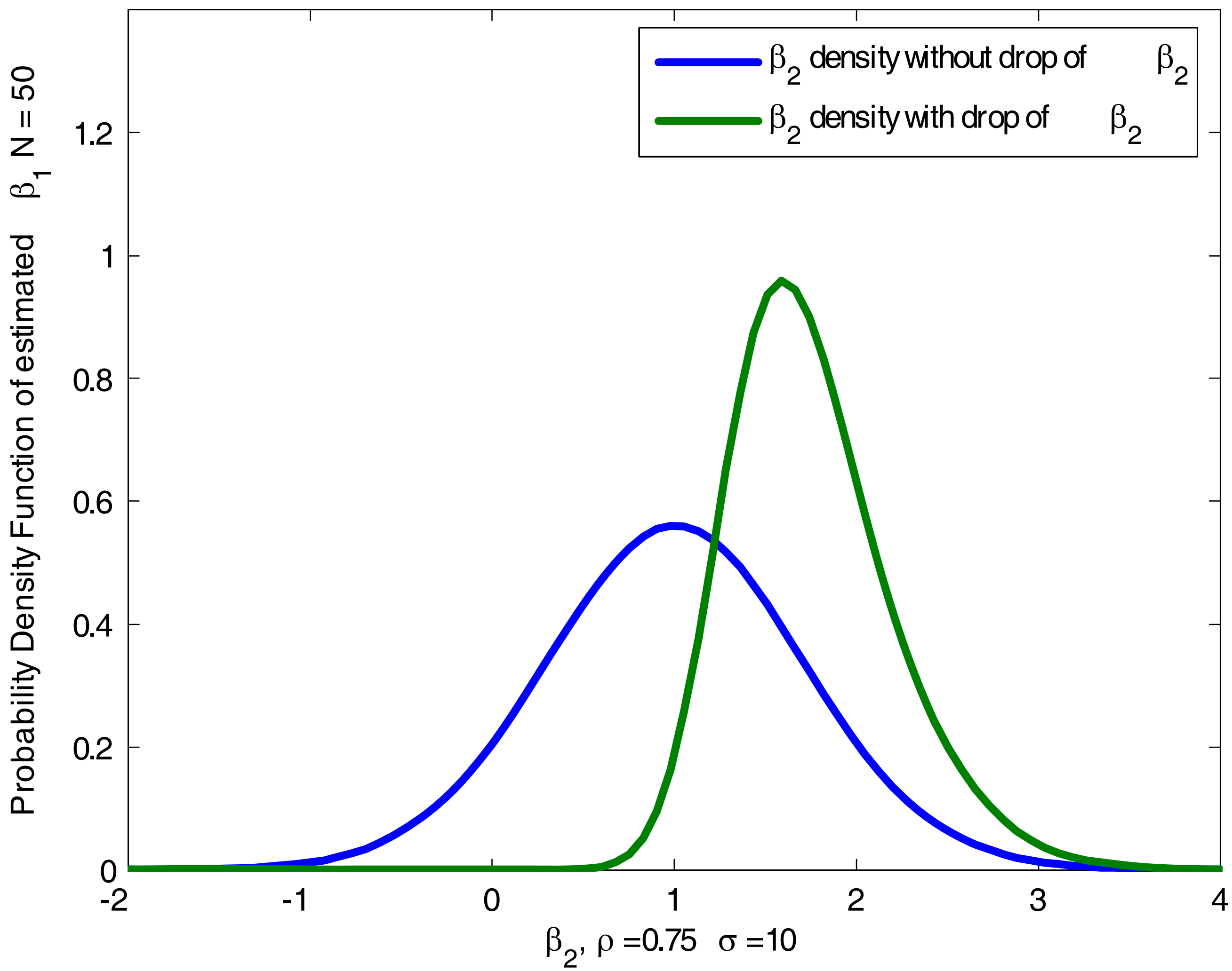
$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 1$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_2$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 2$$

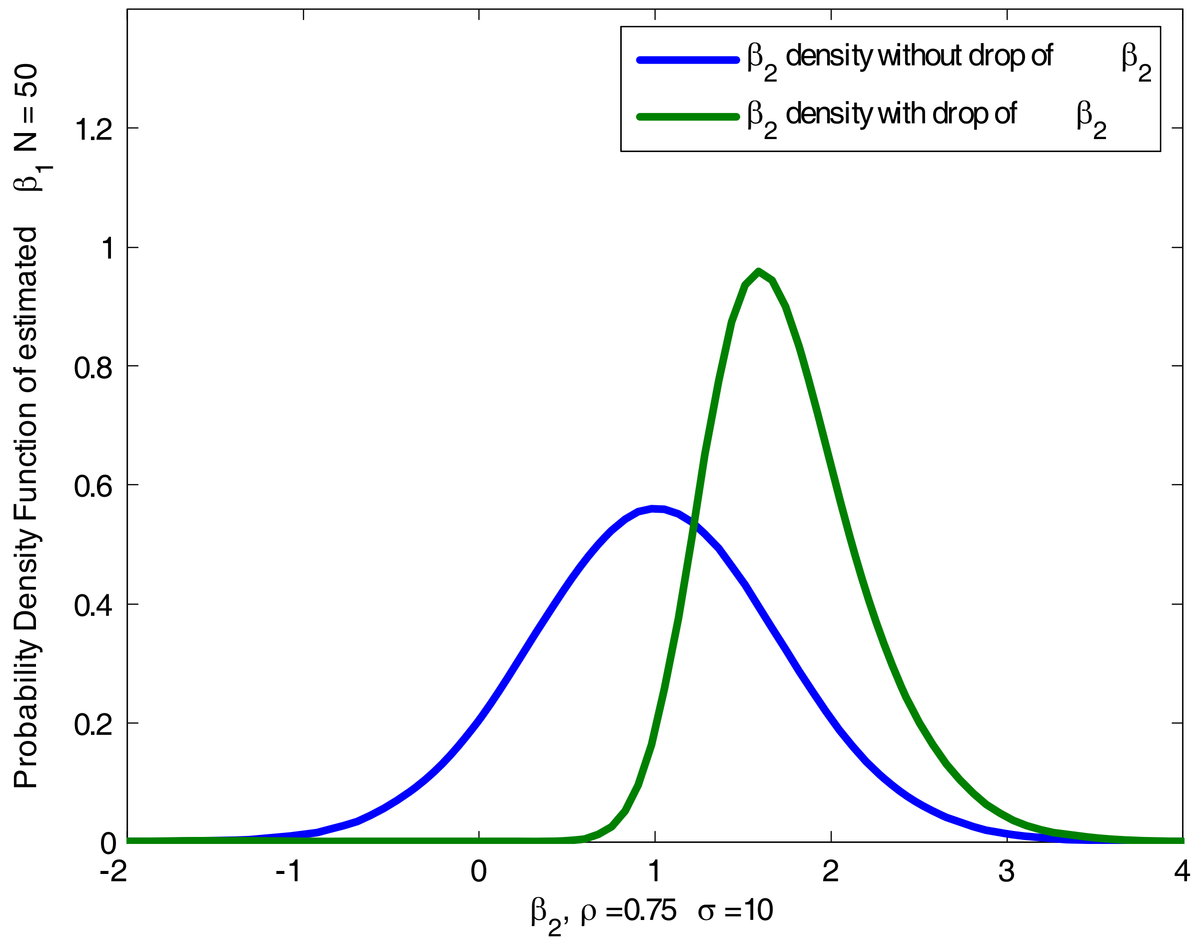
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 5$

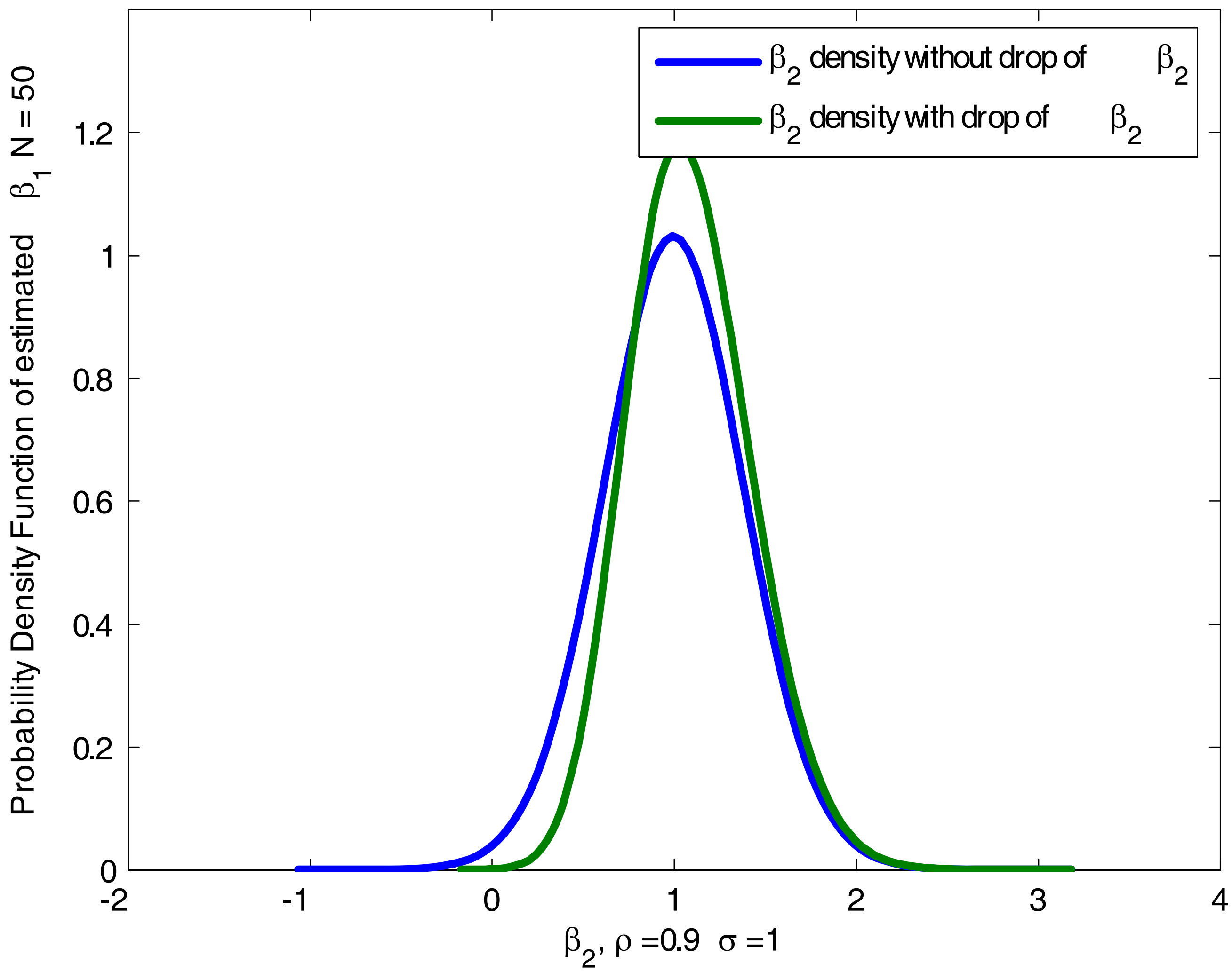
Pre-Test Estimator Density Function



Distribution of β_2

$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 10$

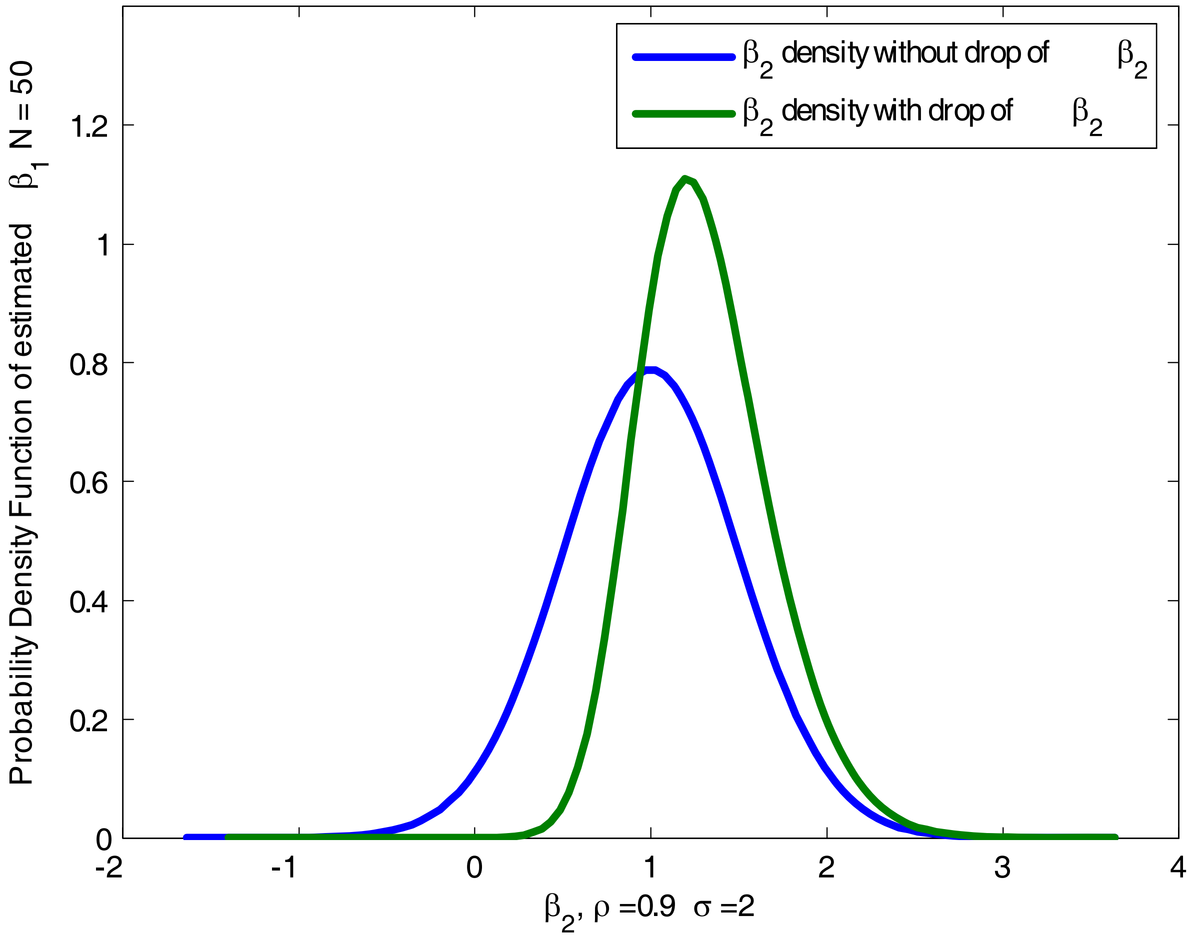
Pre-Test Estimator Density Function



Distribution of β_2

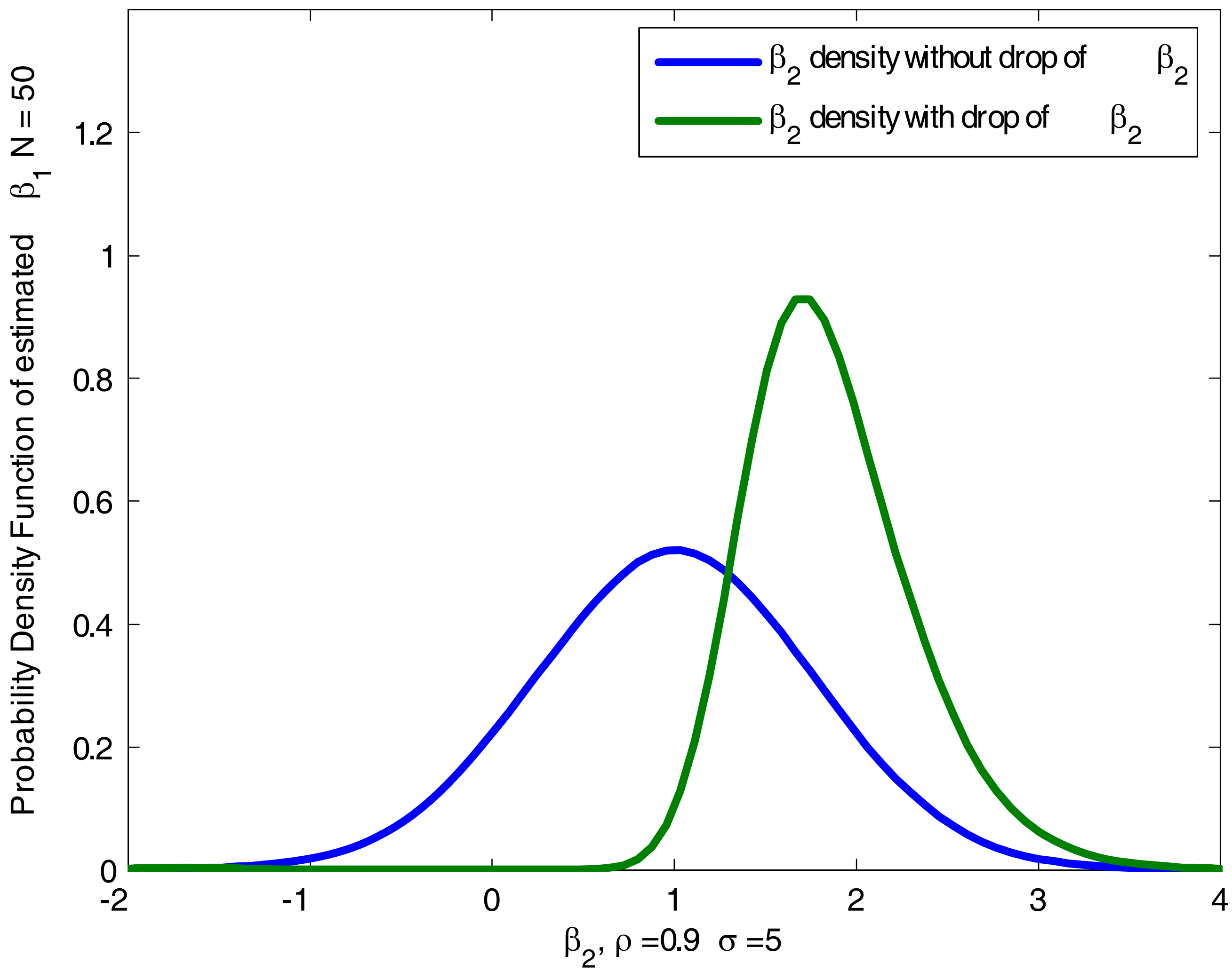
$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 1$

Pre-Test Estimator Density Function



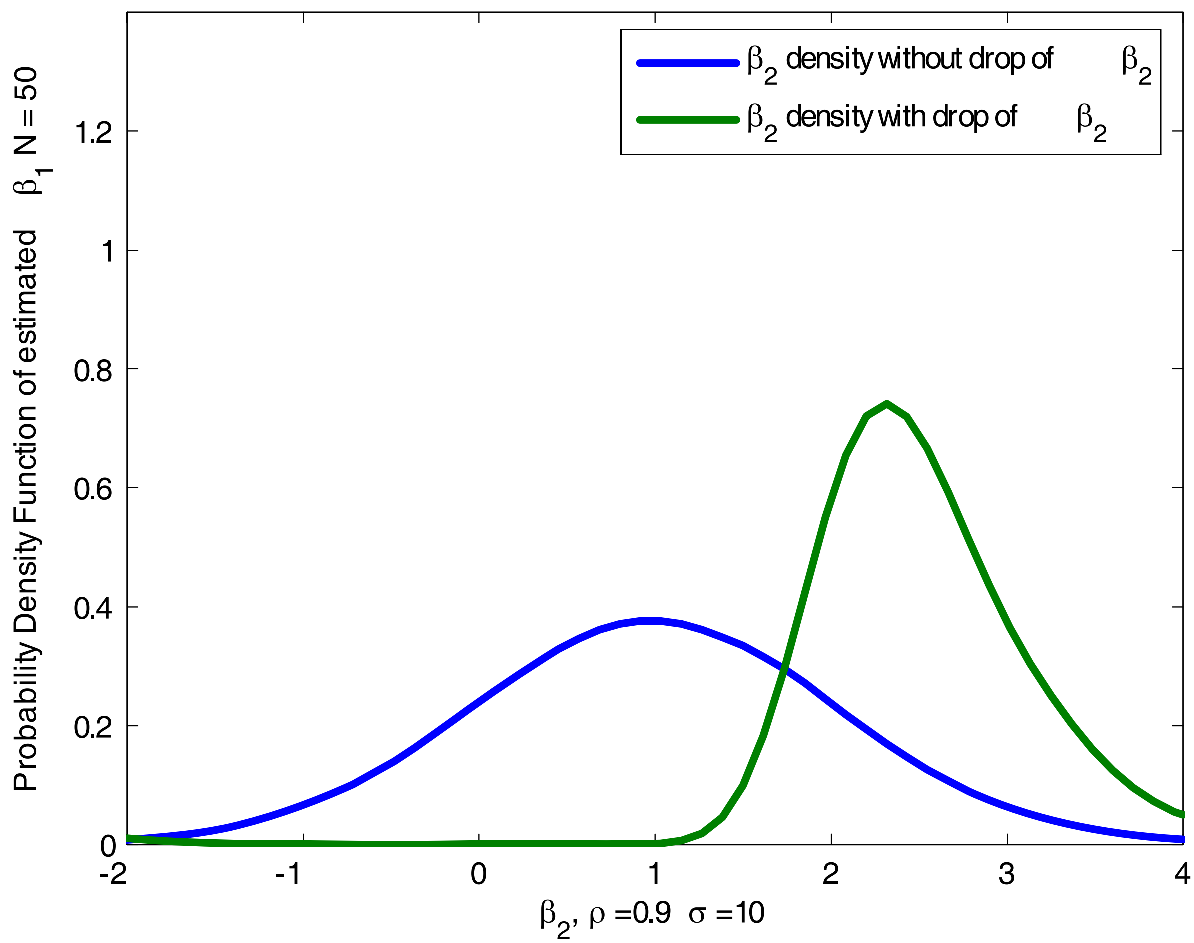
Distribution of β_2
$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 2$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_2$$
$$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 5$$

Pre-Test Estimator Density Function



Distribution of β_2
$N = 50, \quad \rho = 0.9; \quad \sigma^2 = 10$

5 The β_1 Analysis of Procedures 1, 2 and 3 for a specific Model

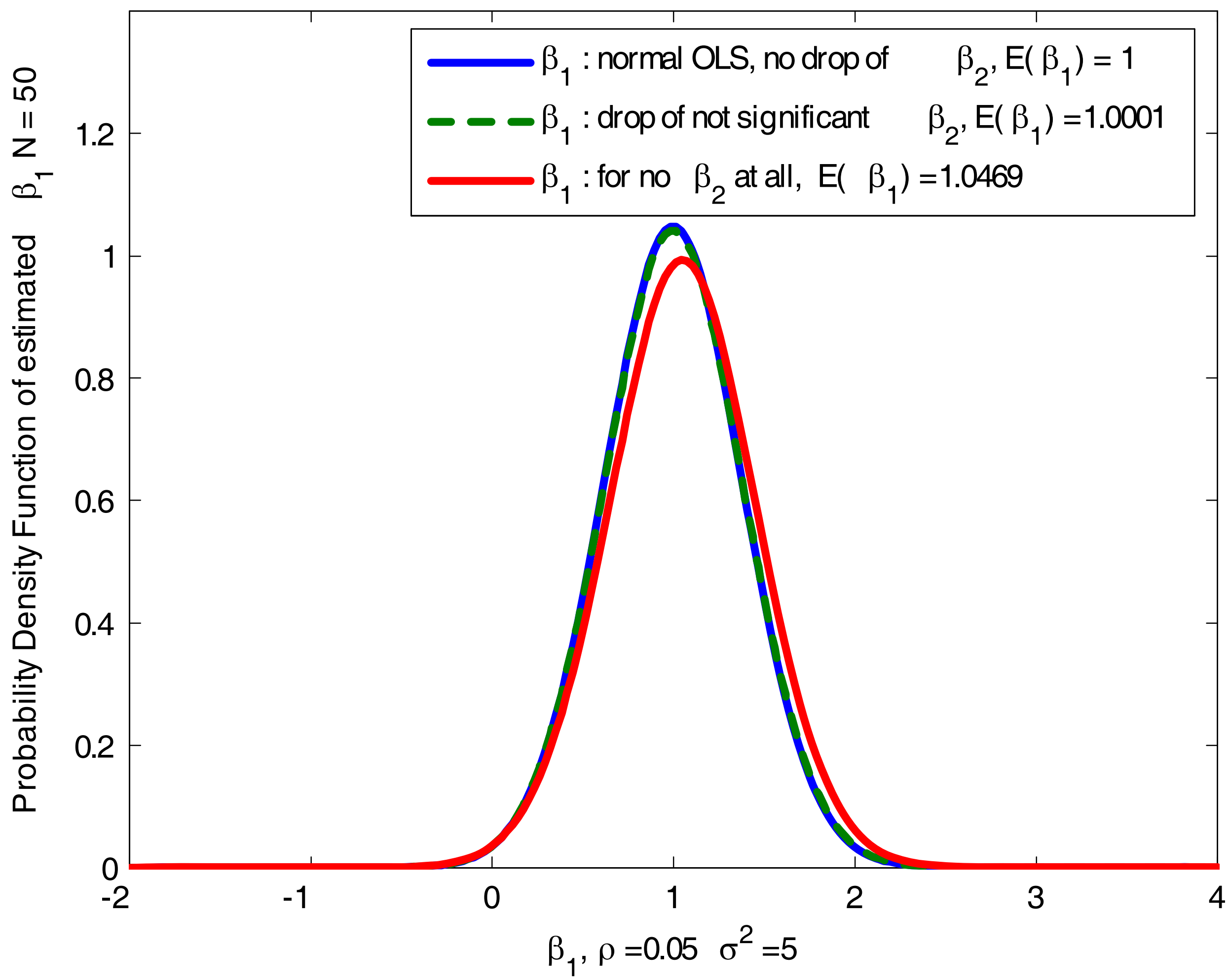
Let the model be:
The True Model:

$$Y_i = X_i \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \varepsilon_i; \; i = 1, ..., 50; \quad \varepsilon_i \sim \mathcal{N}(0, 5) \quad \text{i.i.d.};$$

$$Y_i = X_{1,i} + X_{2,i} + \varepsilon_i; \quad \sigma^2 = 5; \quad N = 50;$$

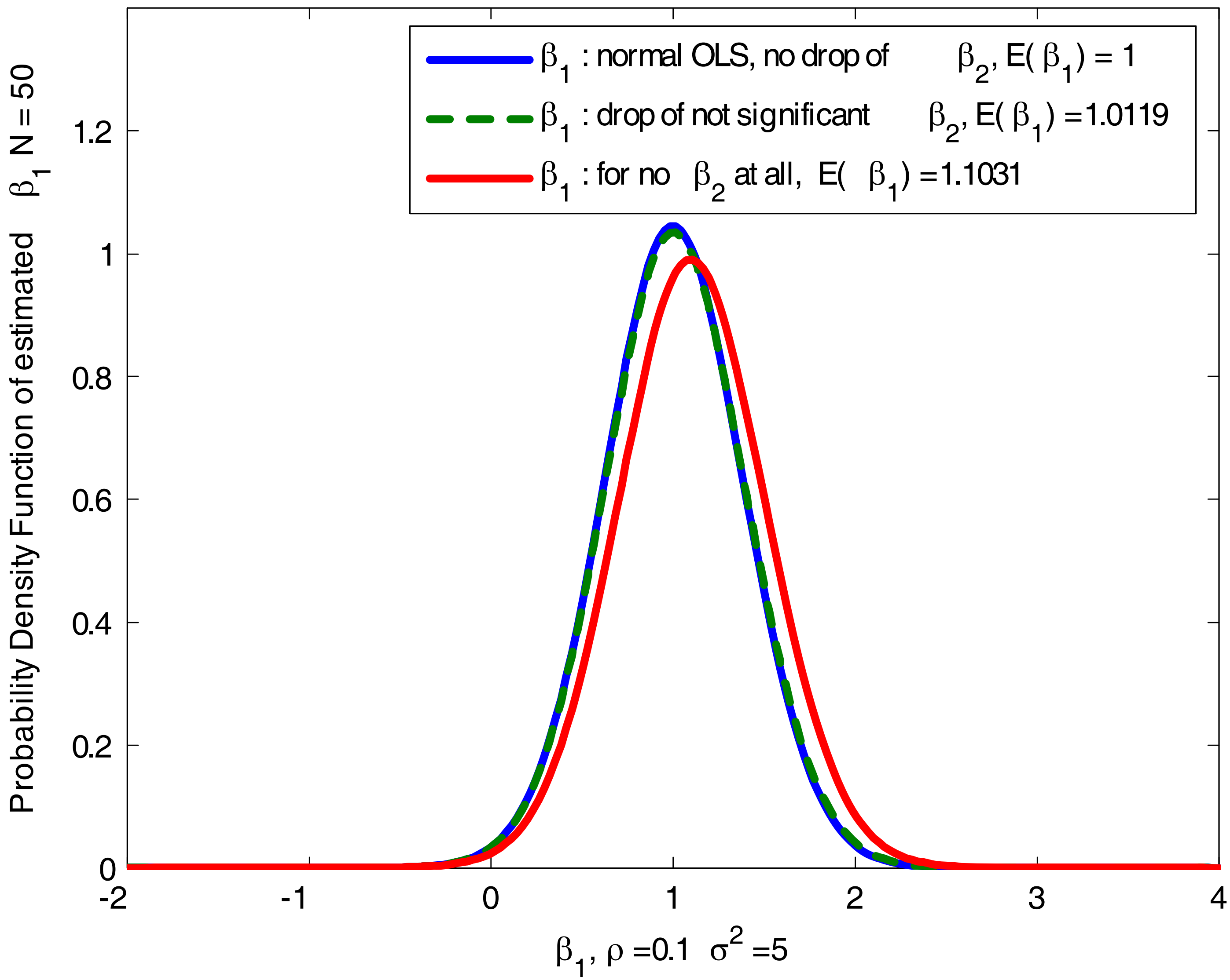
$$X_i \equiv \begin{bmatrix} X_{1i} \\ X_{2i} \end{bmatrix} \sim \mathcal{N} \begin{bmatrix} 0, & \begin{matrix} 1 & \rho \\ \rho & 1 \end{matrix} \end{bmatrix} \quad \text{i.i.d.}; \qquad \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Pre-Test Estimator Density Function



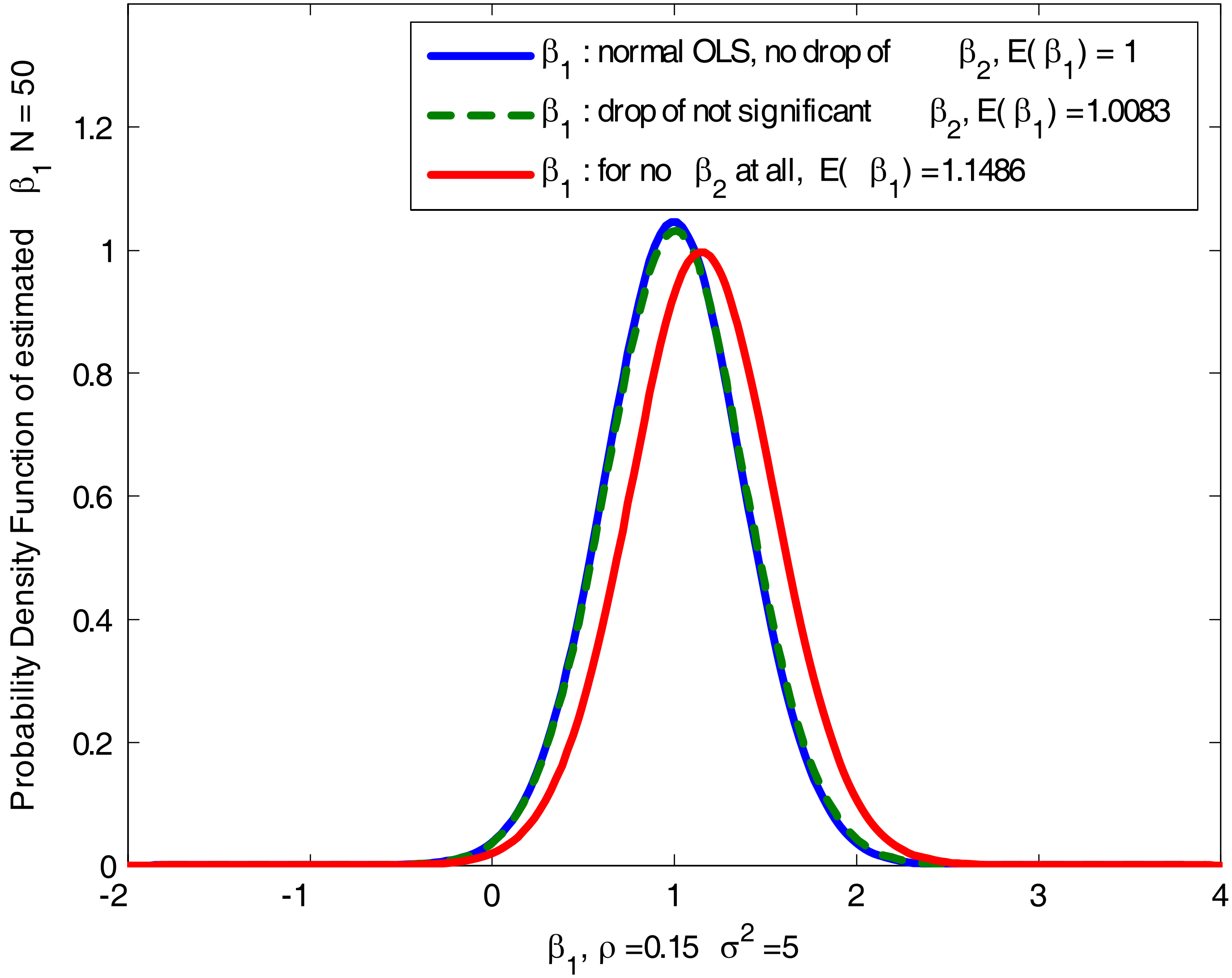
$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.05; \quad \sigma^2 = 5$$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.10; \quad \sigma^2 = 5$$

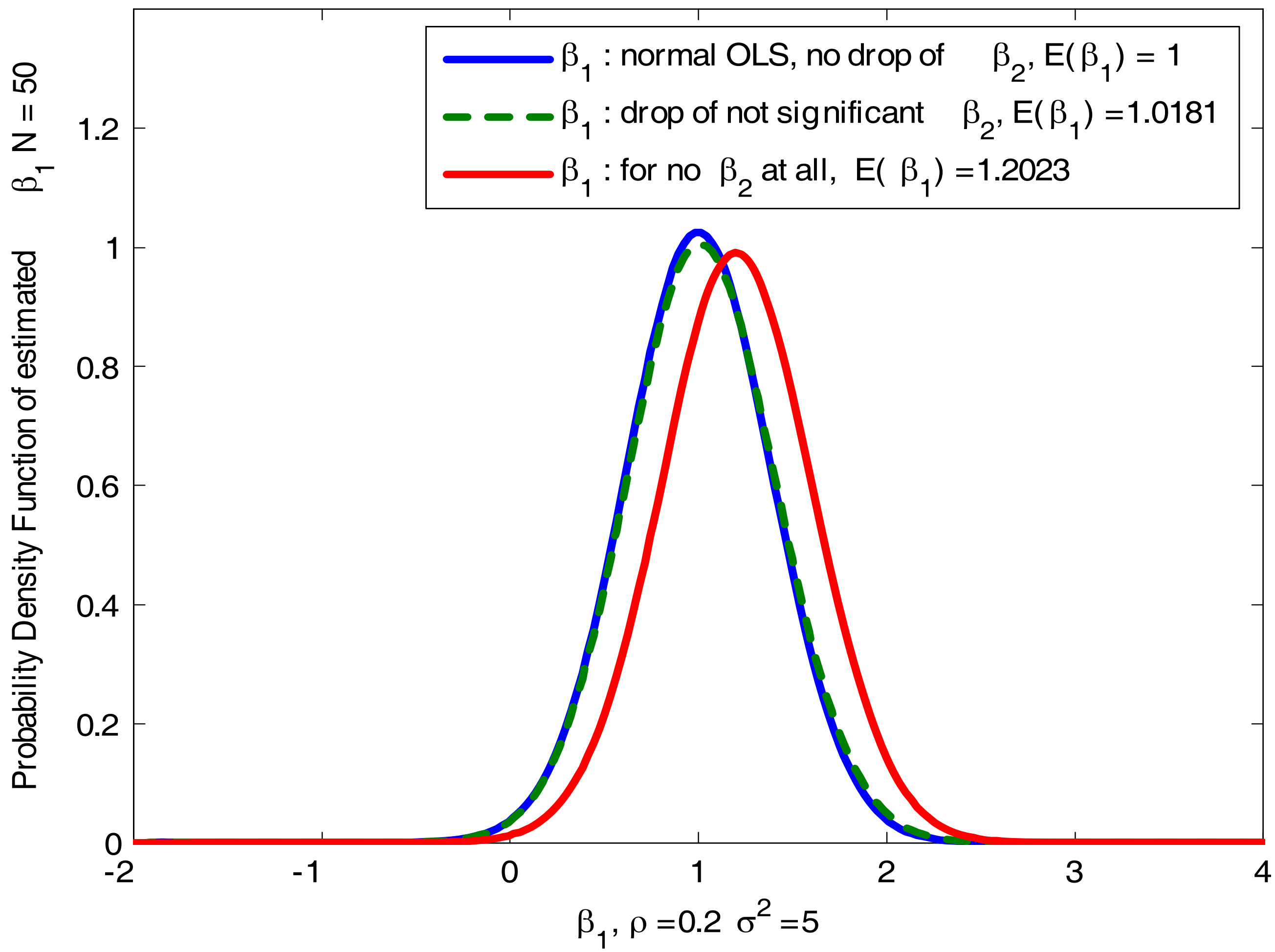
Pre-Test Estimator Density Function



Distribution of β_1

$$N = 50, \quad \rho = 0.15; \quad \sigma^2 = 5$$

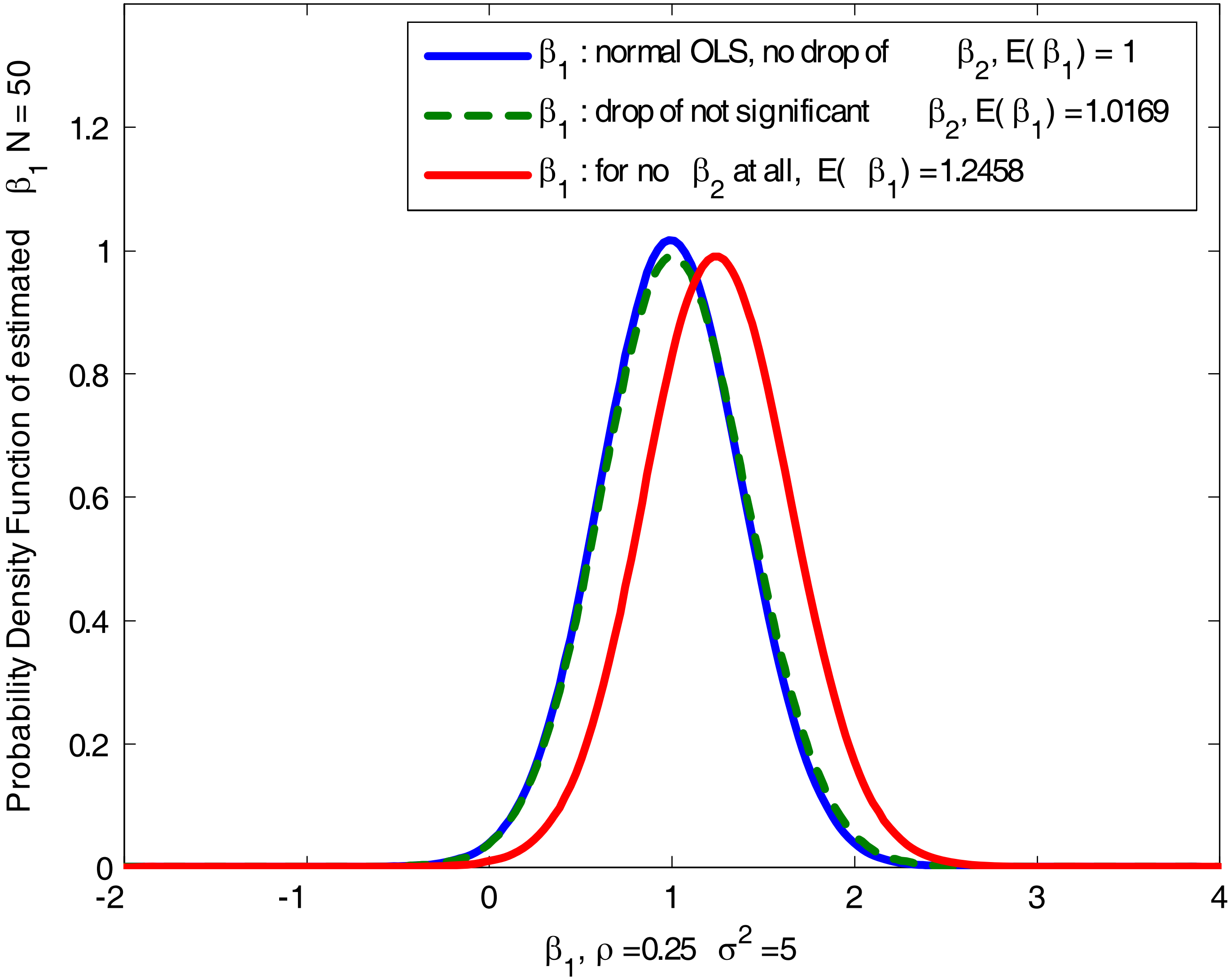
Pre-Test Estimator Density Function



Distribution of β_1

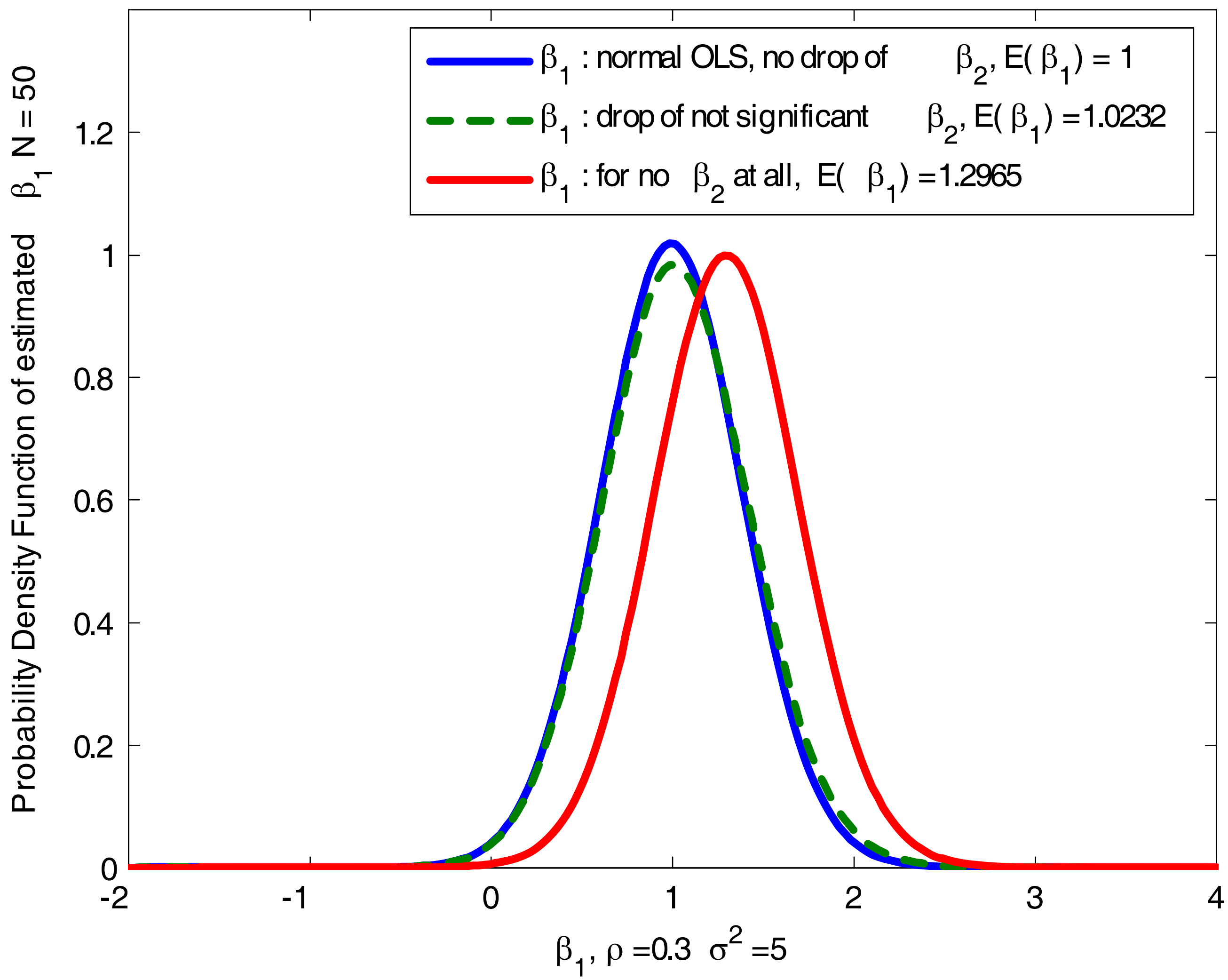
$N = 50, \quad \rho = 0.20; \quad \sigma^2 = 5$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 5$$

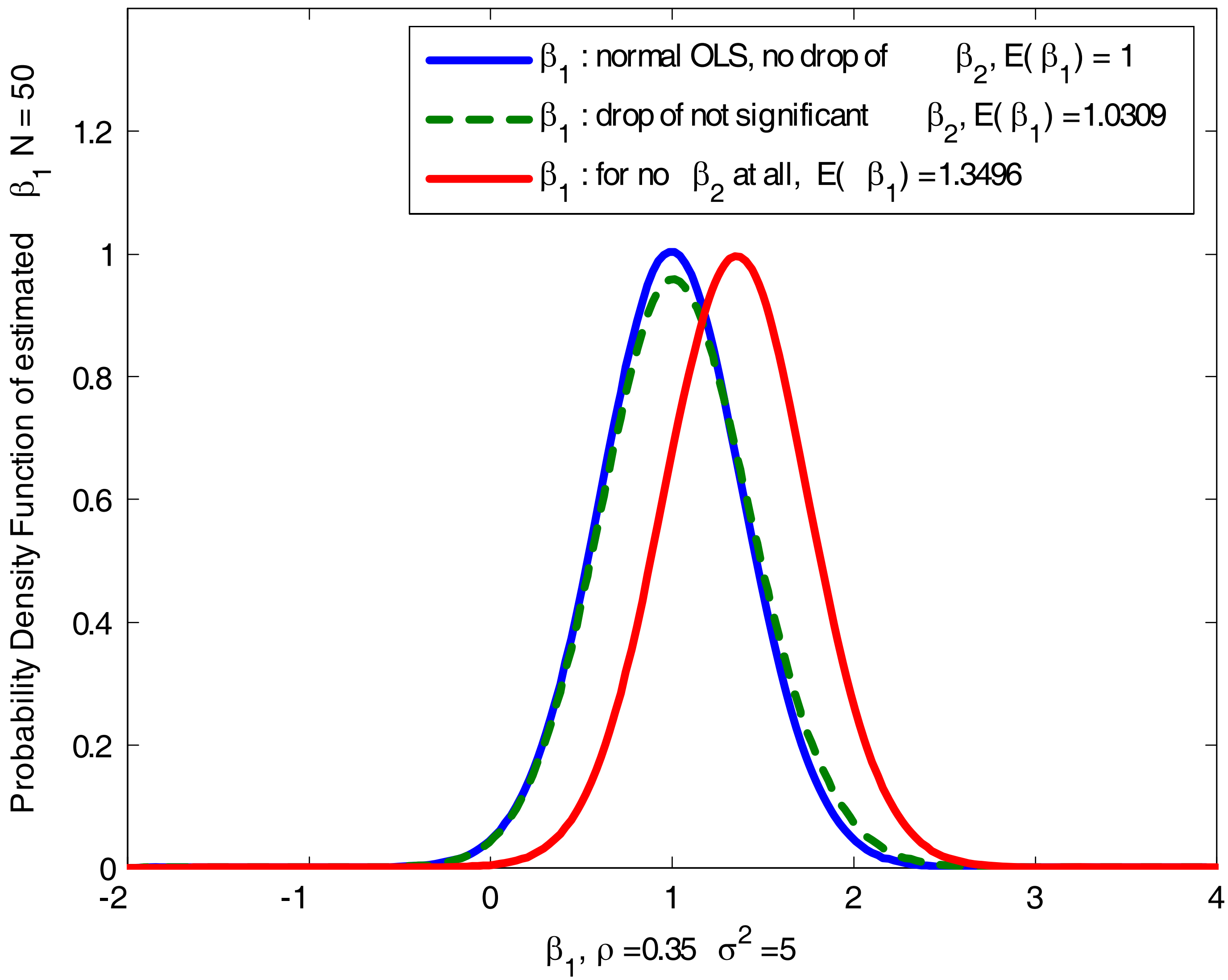
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.30; \quad \sigma^2 = 5$

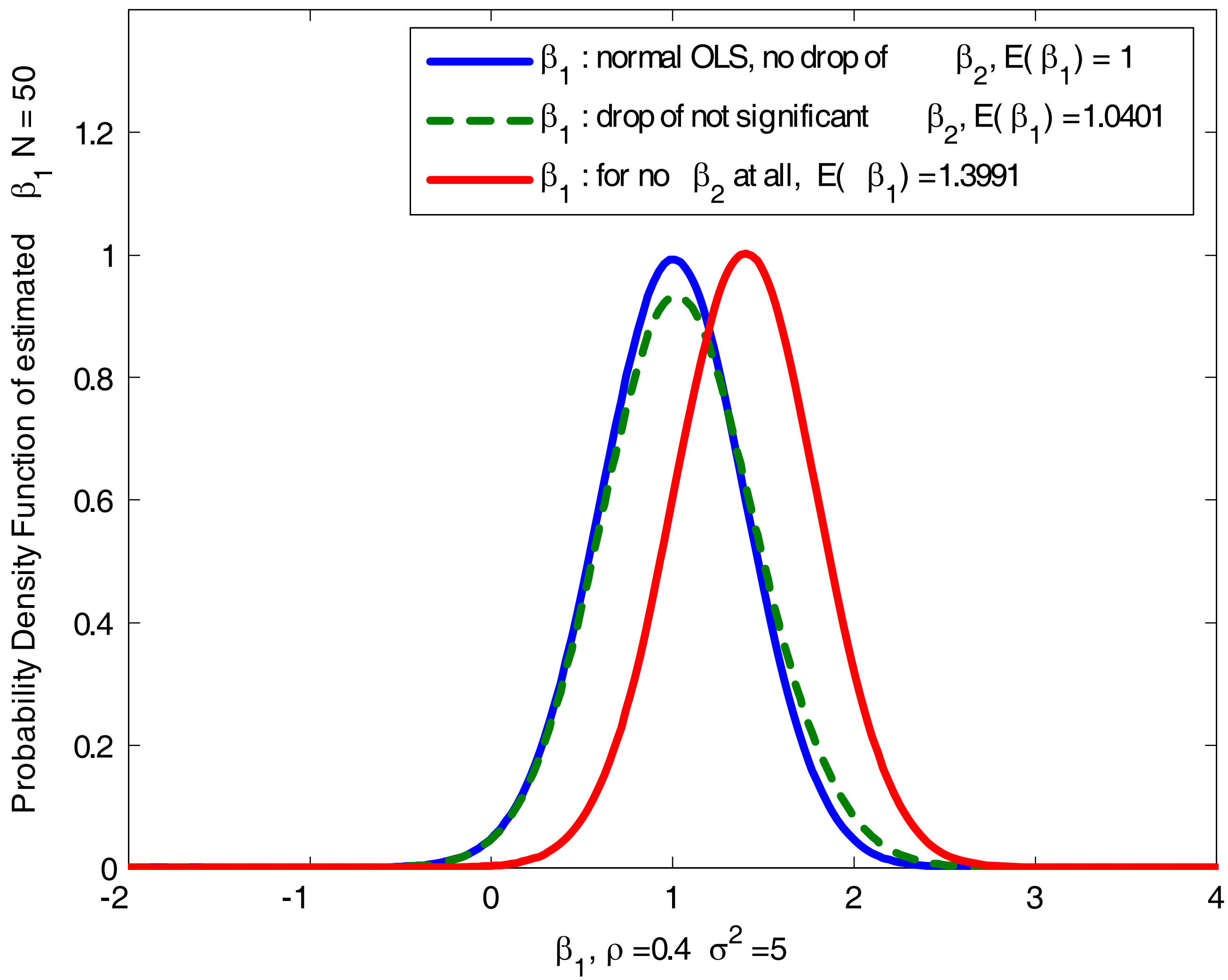
Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$

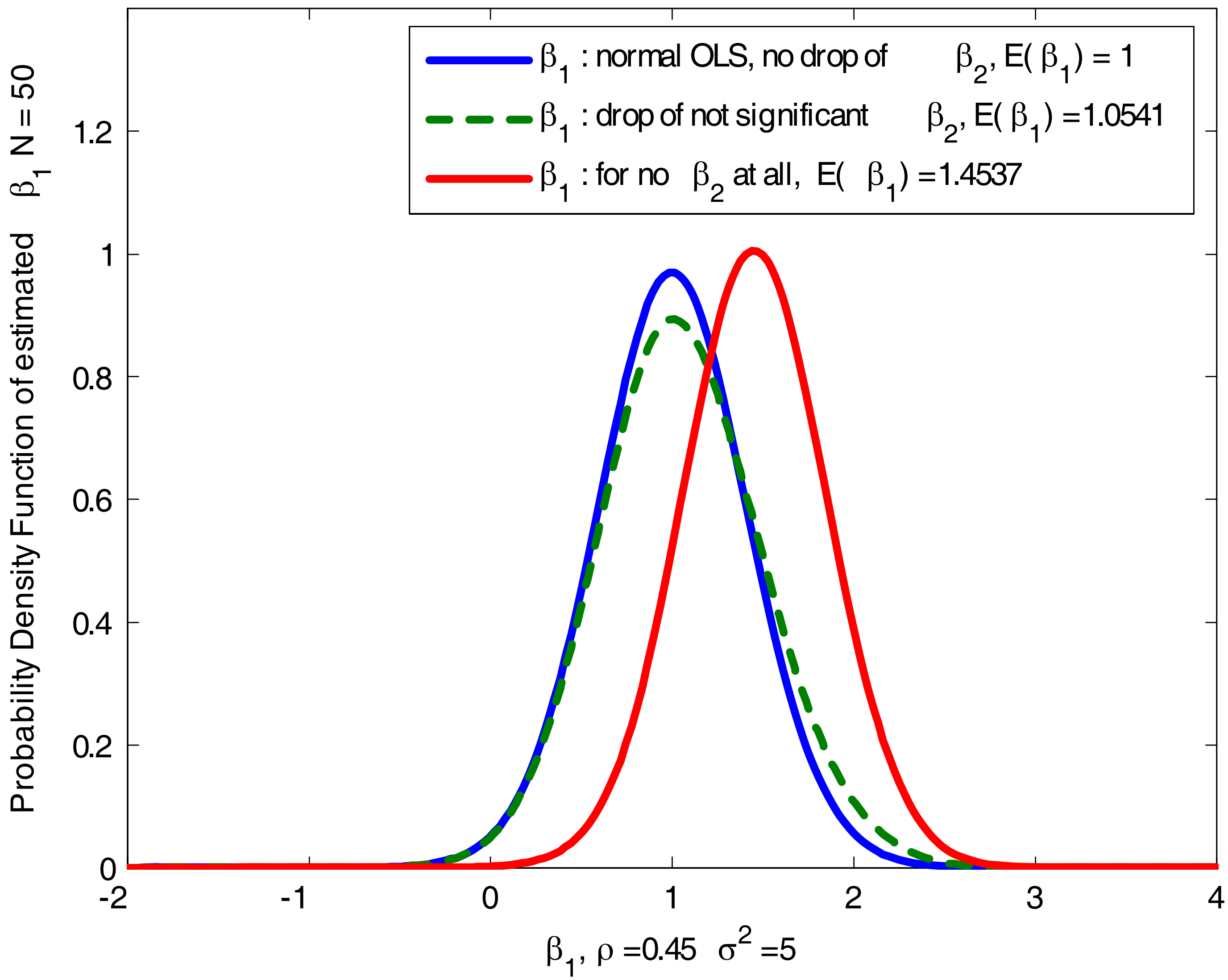
$$N = 50, \quad \rho = 0.35; \quad \sigma^2 = 5$$

Pre-Test Estimator Density Function



Distribution of β_1
$N = 50, \quad \rho = 0.40; \quad \sigma^2 = 5$

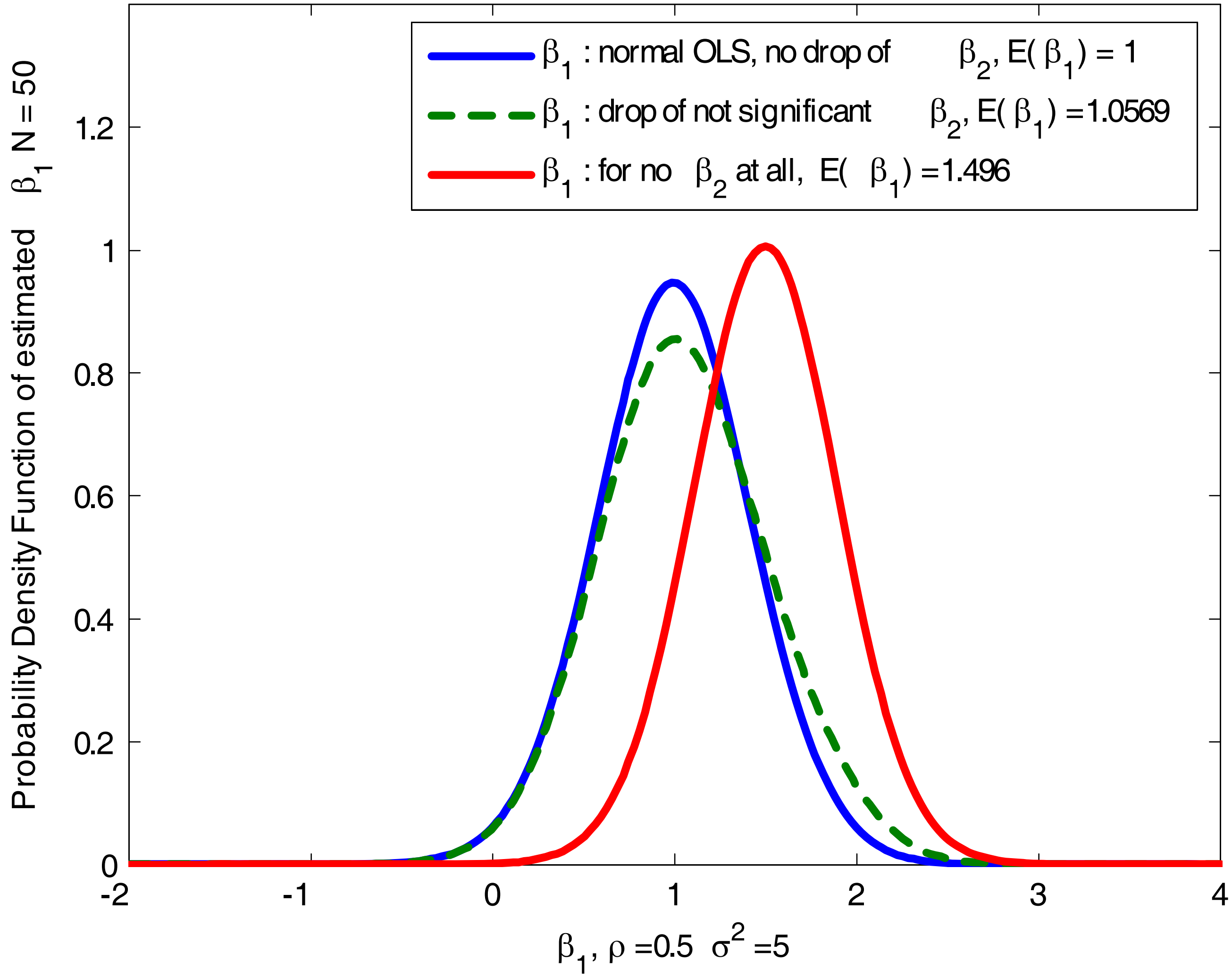
Pre-Test Estimator Density Function



Distribution of β_1

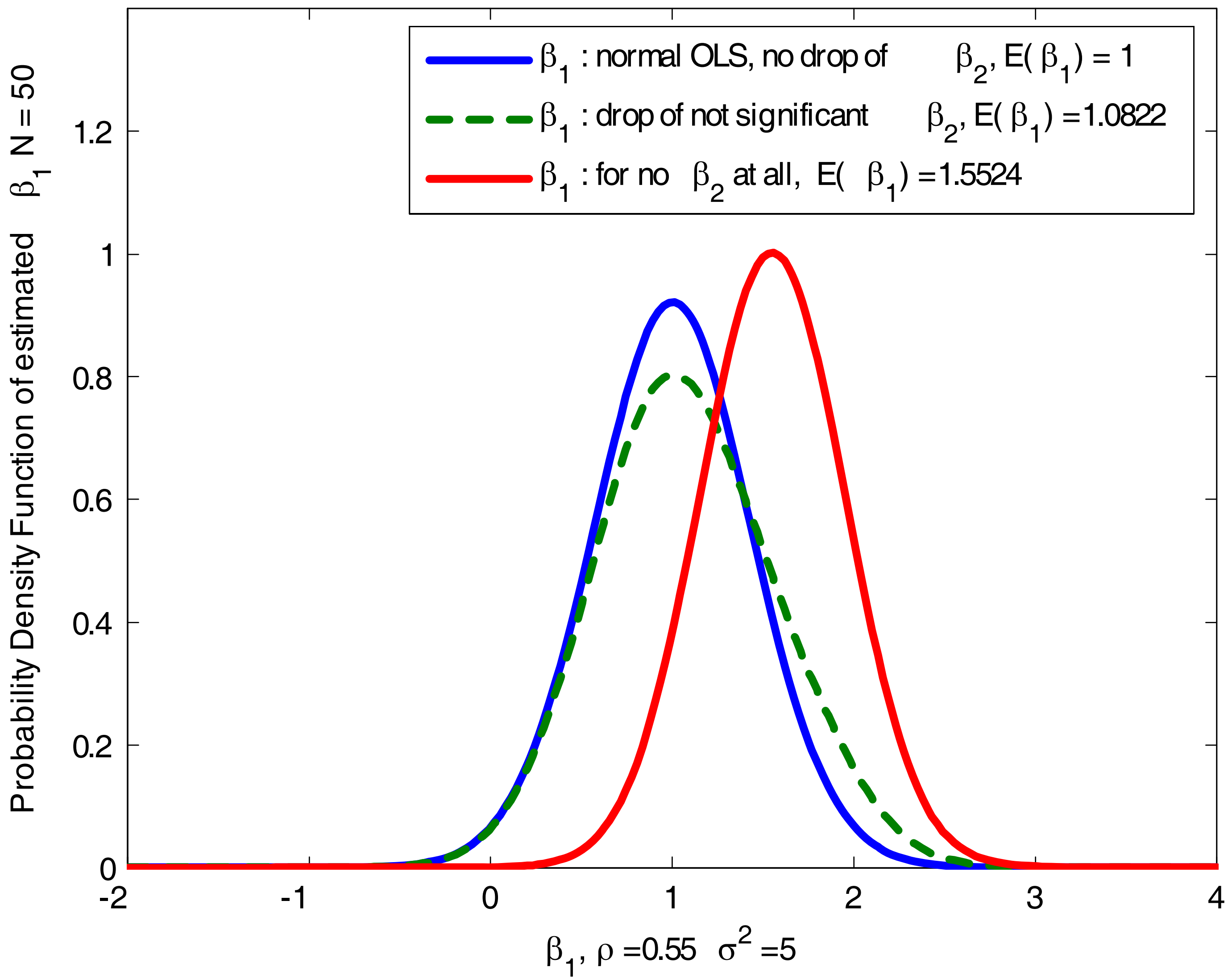
$N = 50, \quad \rho = 0.45; \quad \sigma^2 = 5$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.50; \quad \sigma^2 = 5$$

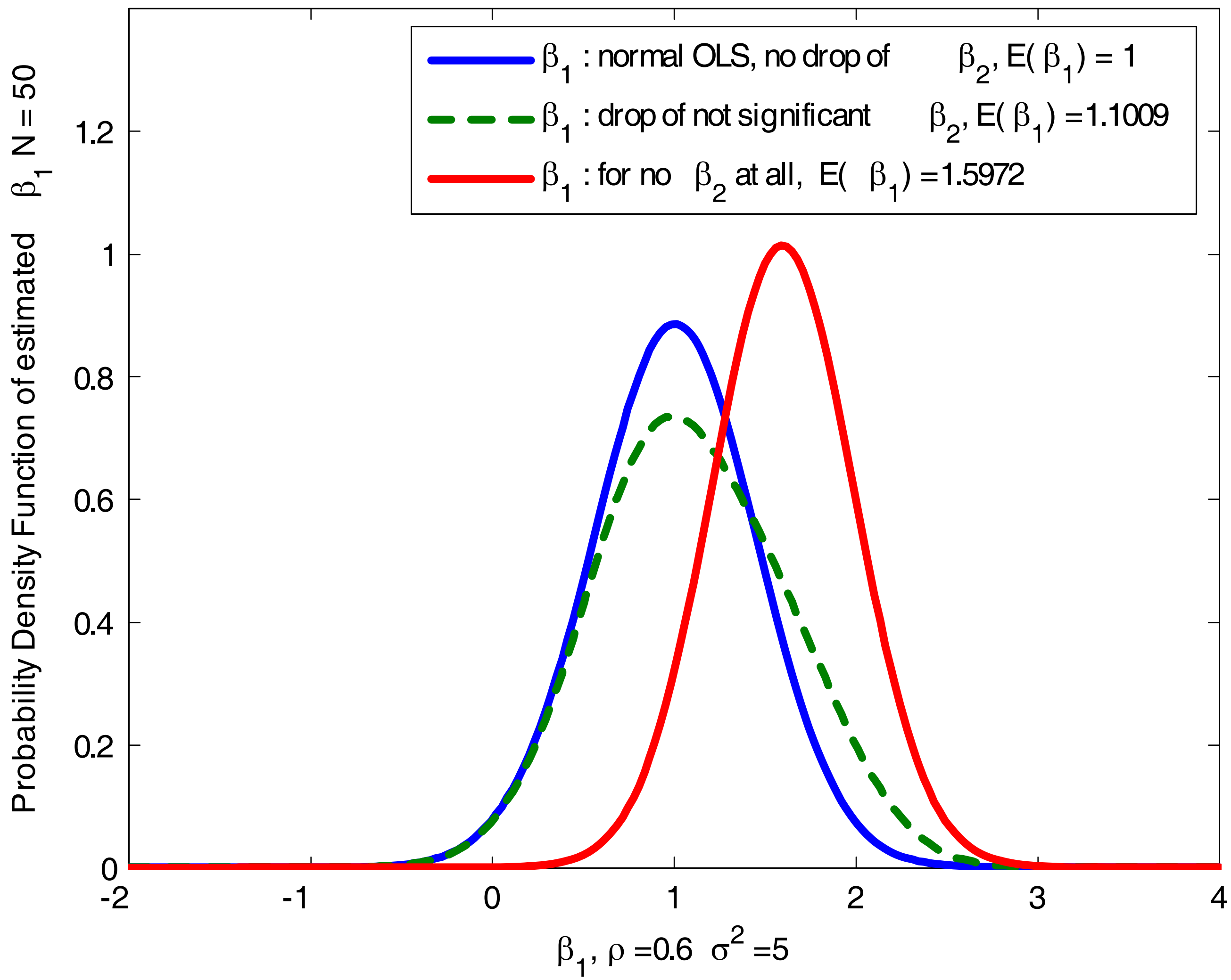
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.55; \quad \sigma^2 = 5$

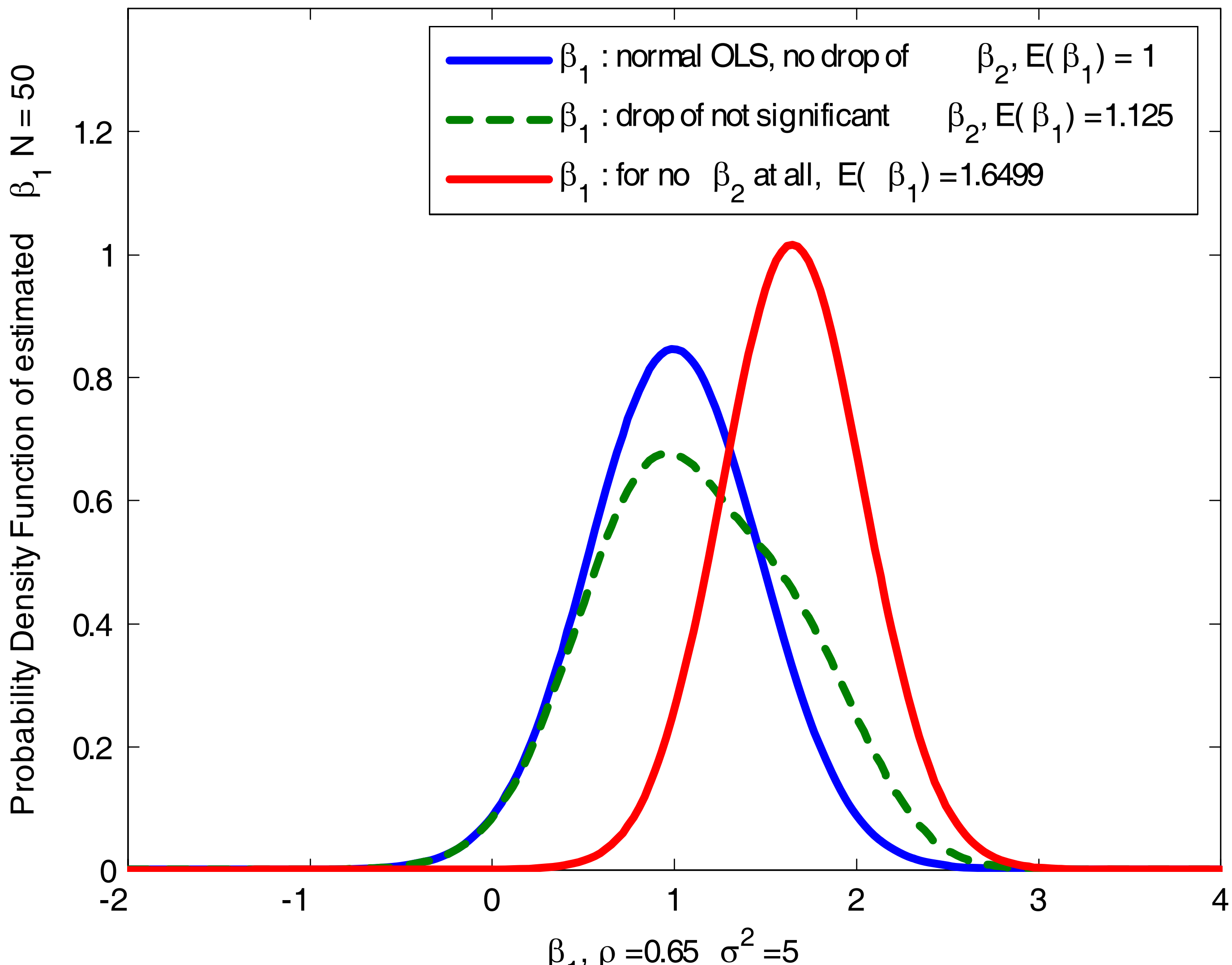
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.60; \quad \sigma^2 = 5$

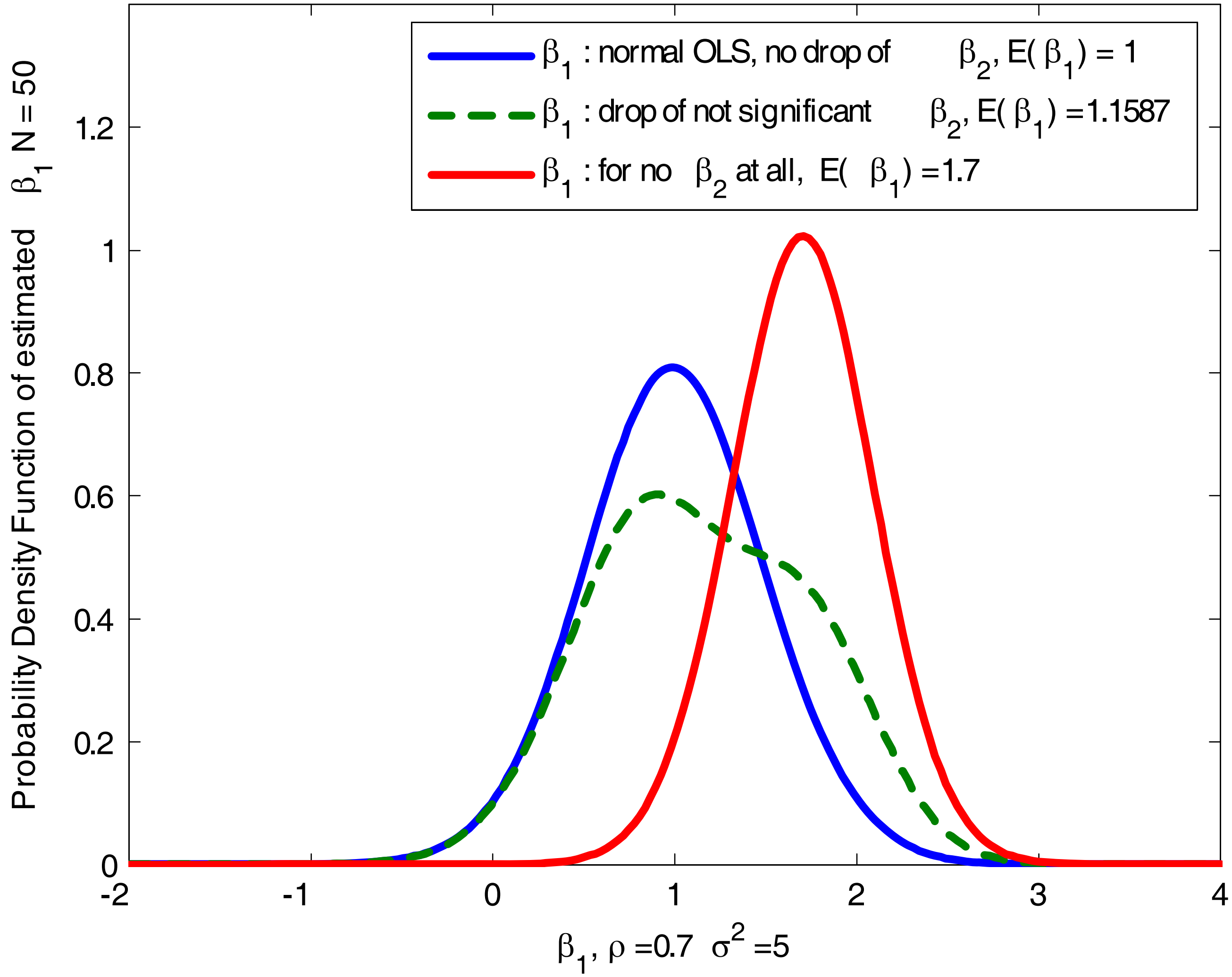
Pre-Test Estimator Density Function



Distribution of β_1

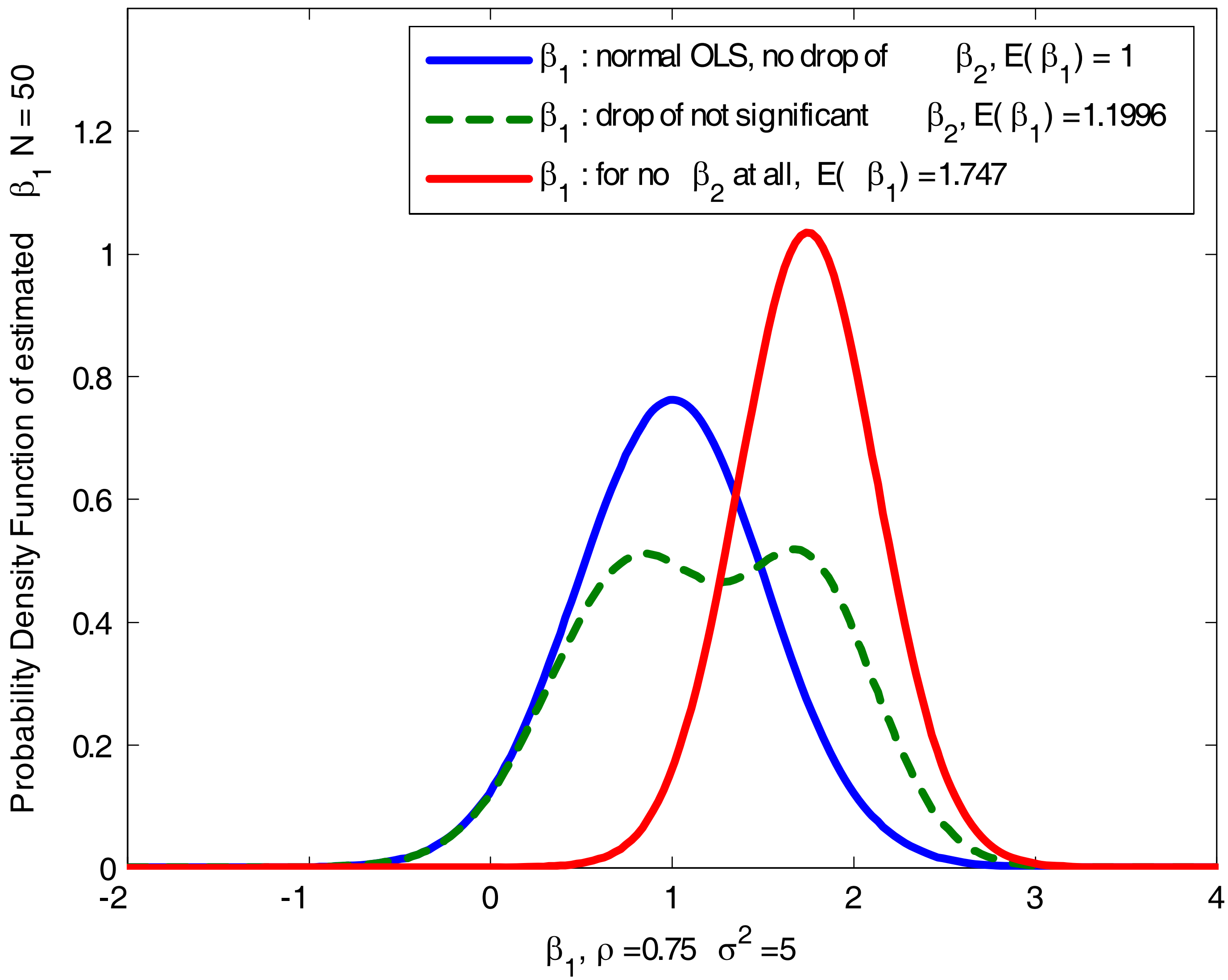
$N = 50, \quad \rho = 0.65; \quad \sigma^2 = 5$

Pre-Test Estimator Density Function



Distribution of β_1
$N = 50, \quad \rho = 0.70; \quad \sigma^2 = 5$

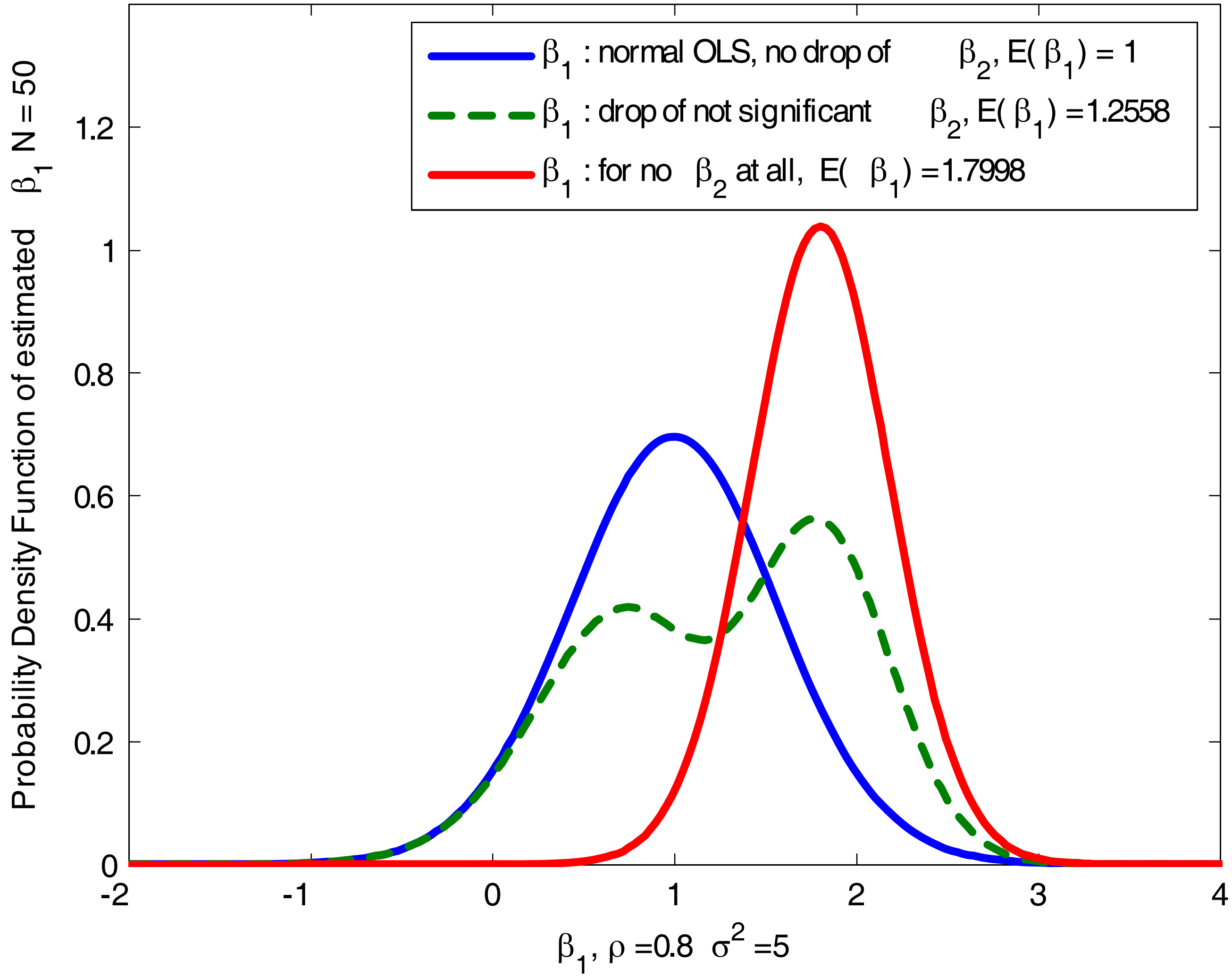
Pre-Test Estimator Density Function



Distribution of β_1

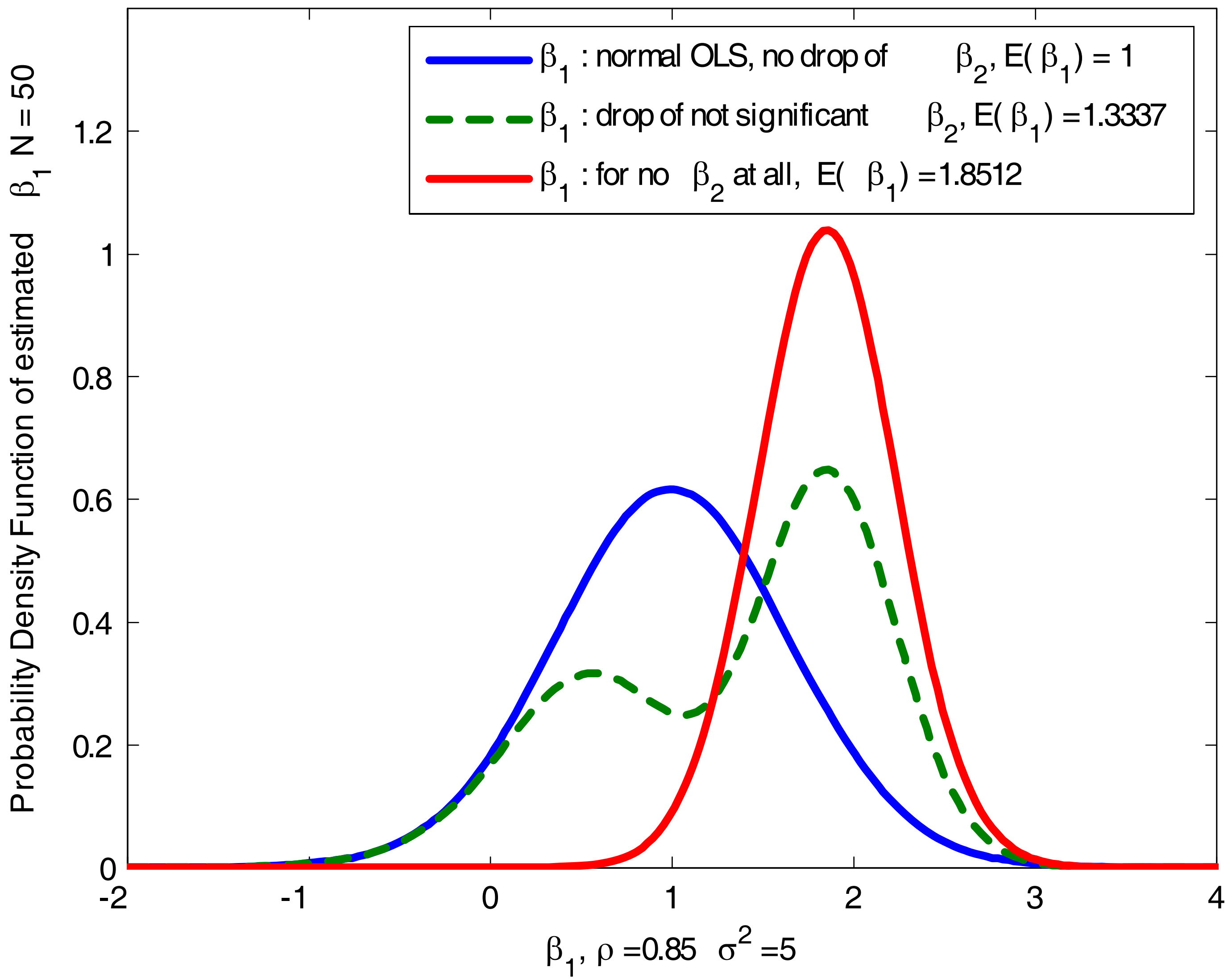
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 5$$

Pre-Test Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \rho = 0.80; \quad \sigma^2 = 5$$

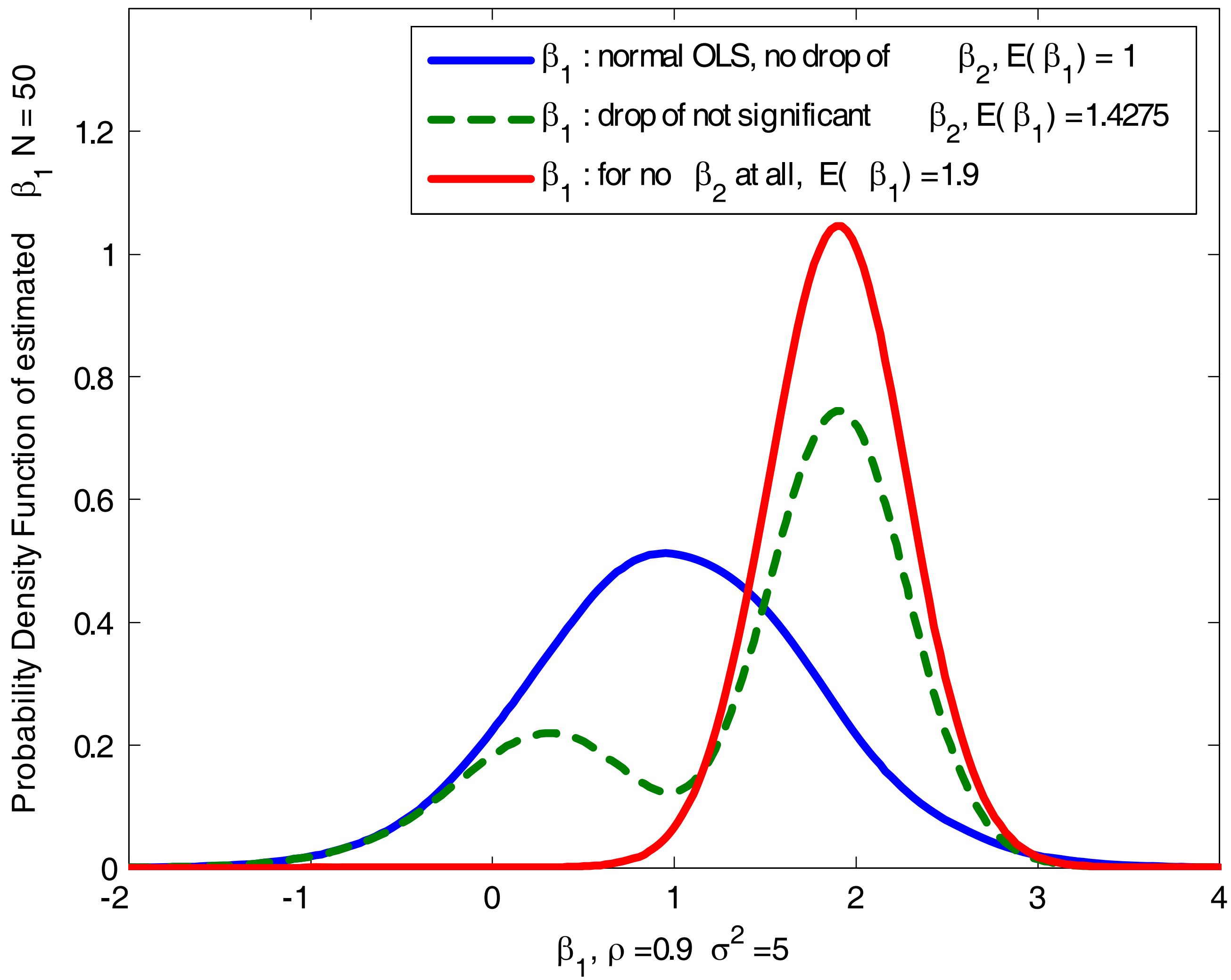
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.85; \quad \sigma^2 = 5$

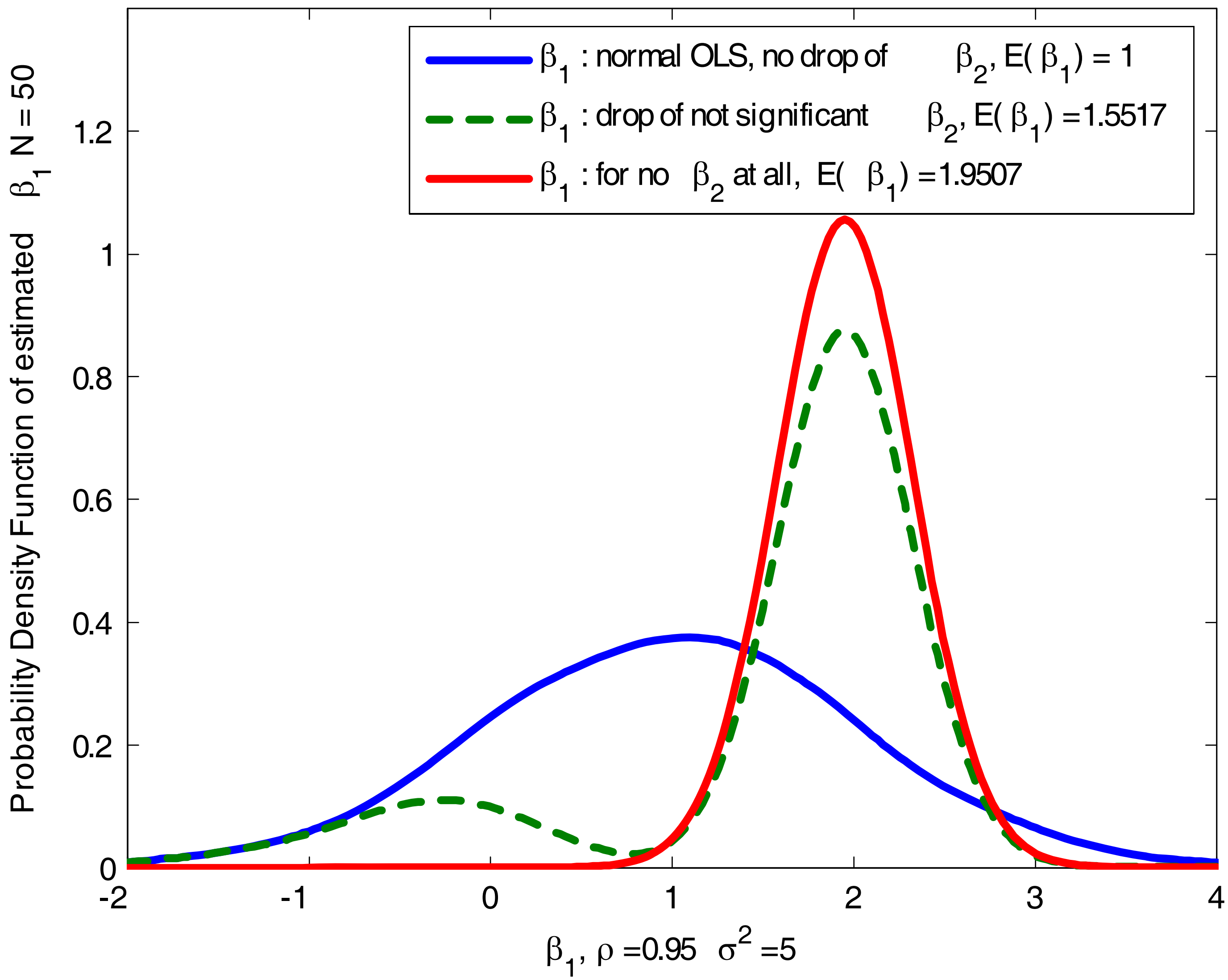
Pre-Test Estimator Density Function



Distribution of β_1

$N = 50, \quad \rho = 0.90; \quad \sigma^2 = 5$

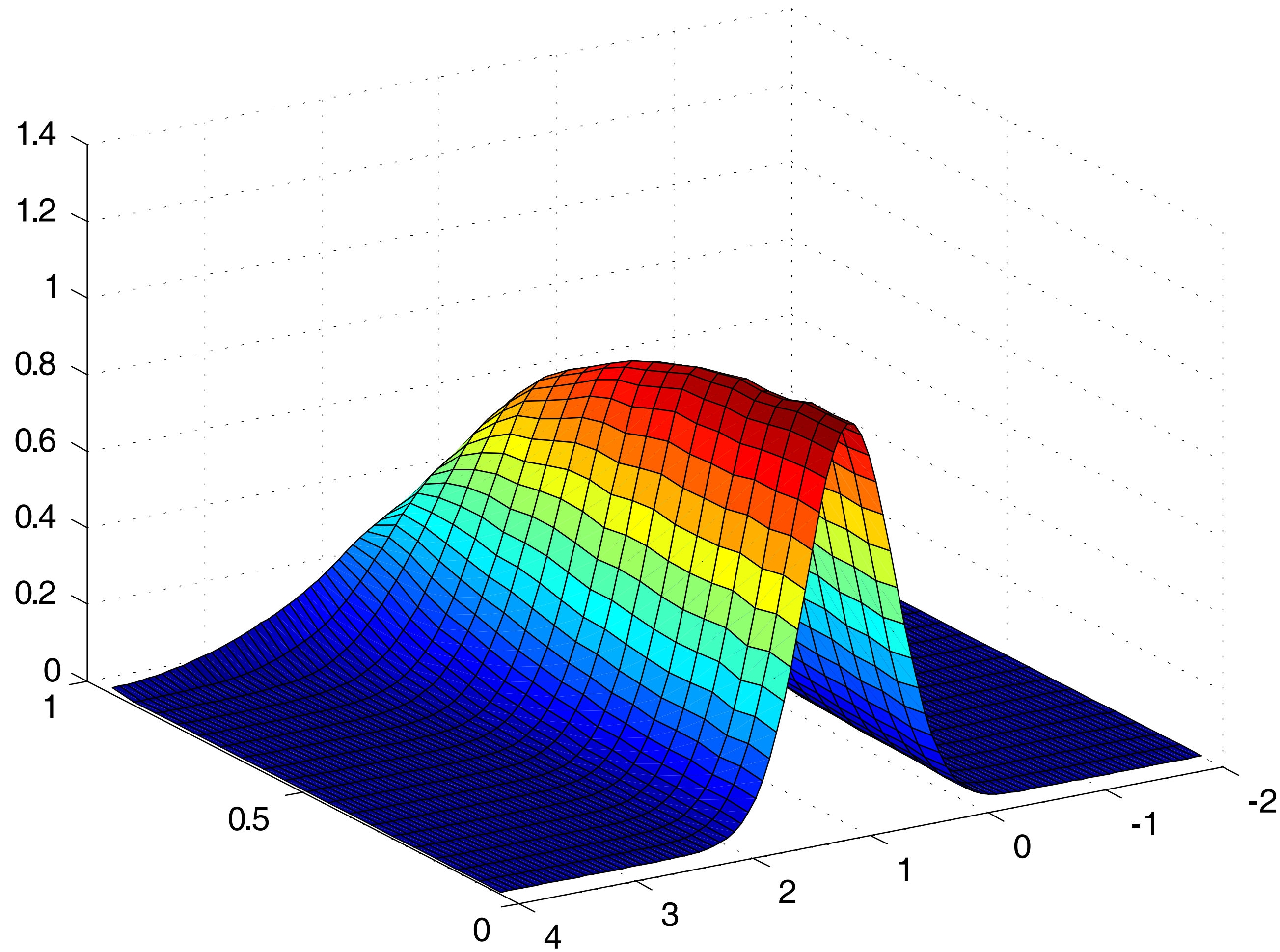
Pre-Test Estimator Density Function



Distribution of β_1

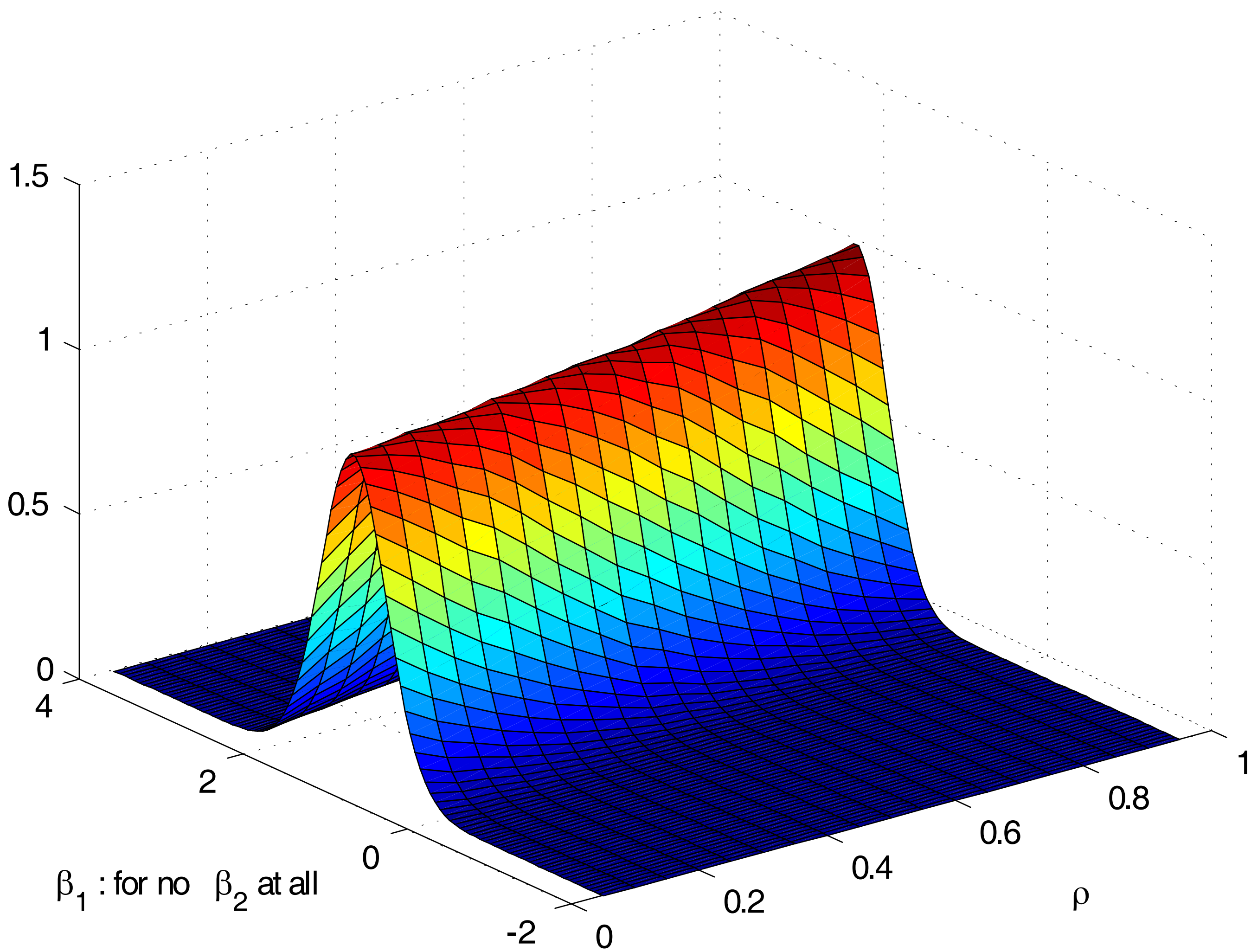
$N = 50, \quad \rho = 0.95; \quad \sigma^2 = 5$

OLS $\widehat{\beta}_1$ Estimator Density Function



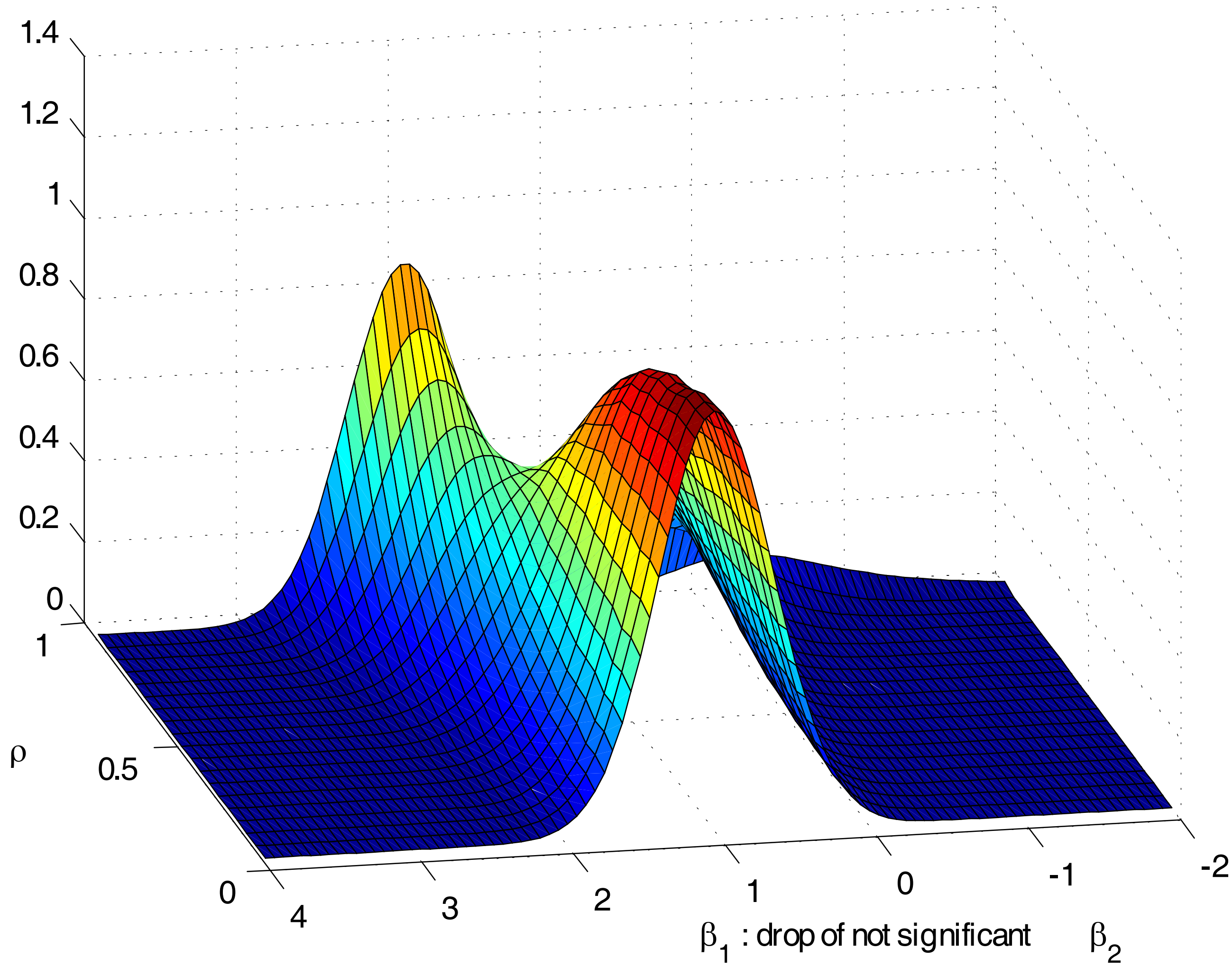
$Distribution\ of\ \beta_1$
$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$

Procedure 2 $\overline{\beta}_1$ Estimator Density Function



$$Distribution\ of\ \beta_1$$
$$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$$

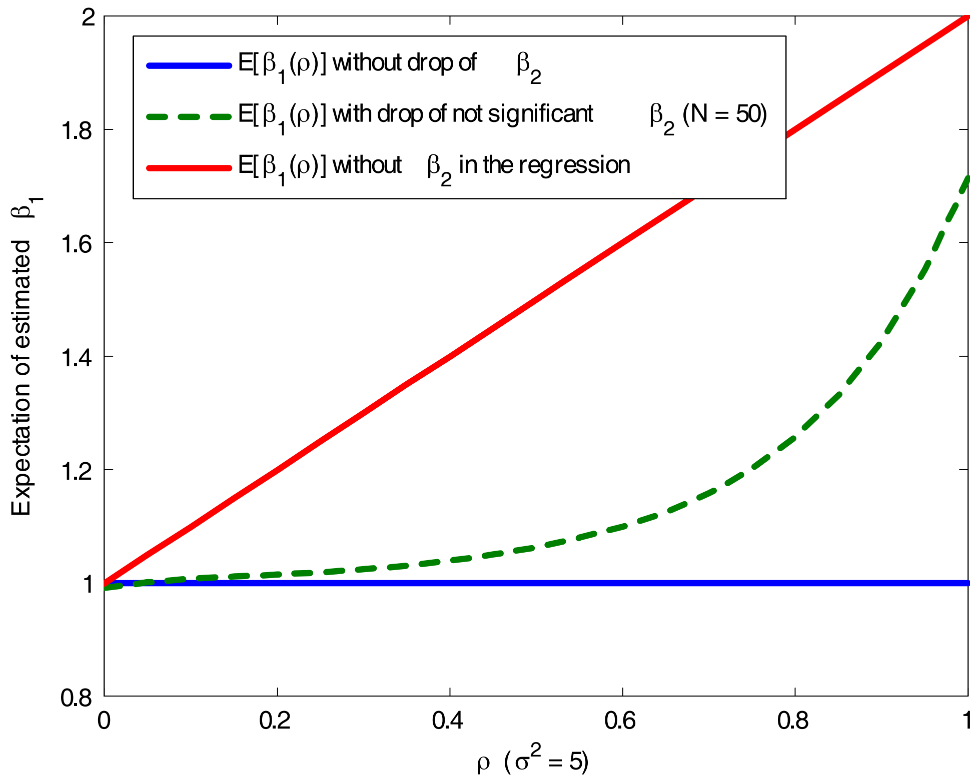
Pre-test $\widetilde{\beta}_1$ Estimator Density Function



$Distribution\ of\ \beta_1$

$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$

Pre-Test Estimator Expectation



$$\textit{Expectation of } \beta_1$$

$$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$$

$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right)$$

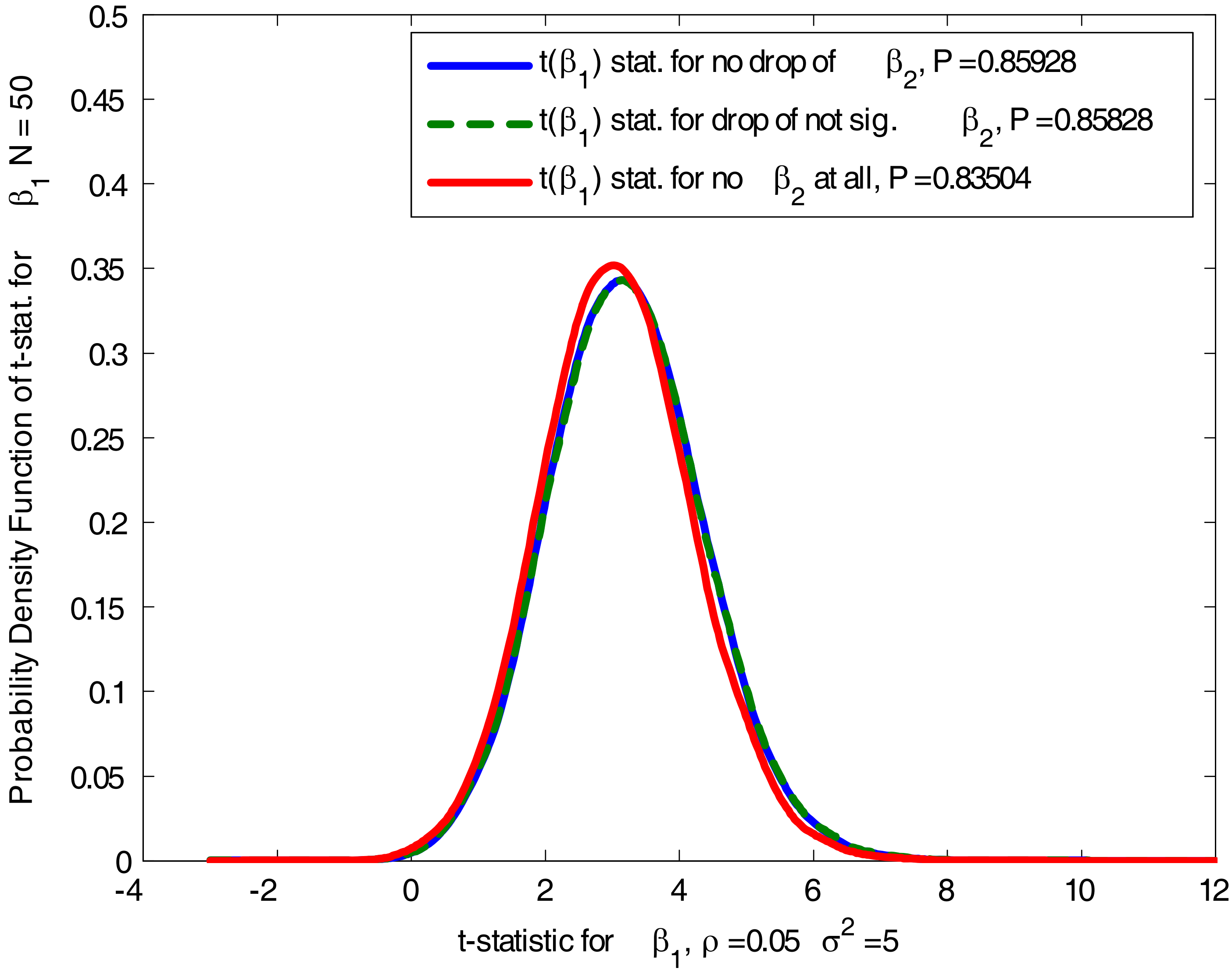
$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot (1-\rho^2)}{N}\right)$$

$$(\text{Pre-test}) \ \widetilde{\beta}_1 = \begin{cases} \widehat{\beta}_1 \text{ with Prob.} = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \\ \overline{\beta}_1 \text{ with Prob.} = 1 - \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \end{cases}$$

$$E\left(\widetilde{\beta}_1\right) = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \cdot \beta_1 + \left[1 - \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right)\right] \cdot (\beta_1 + \rho\beta_2)$$

6 The β_1 T-statistic Analysis of Procedures 1, 2 and 3 for a specific Model

Pre-Test T-statistic

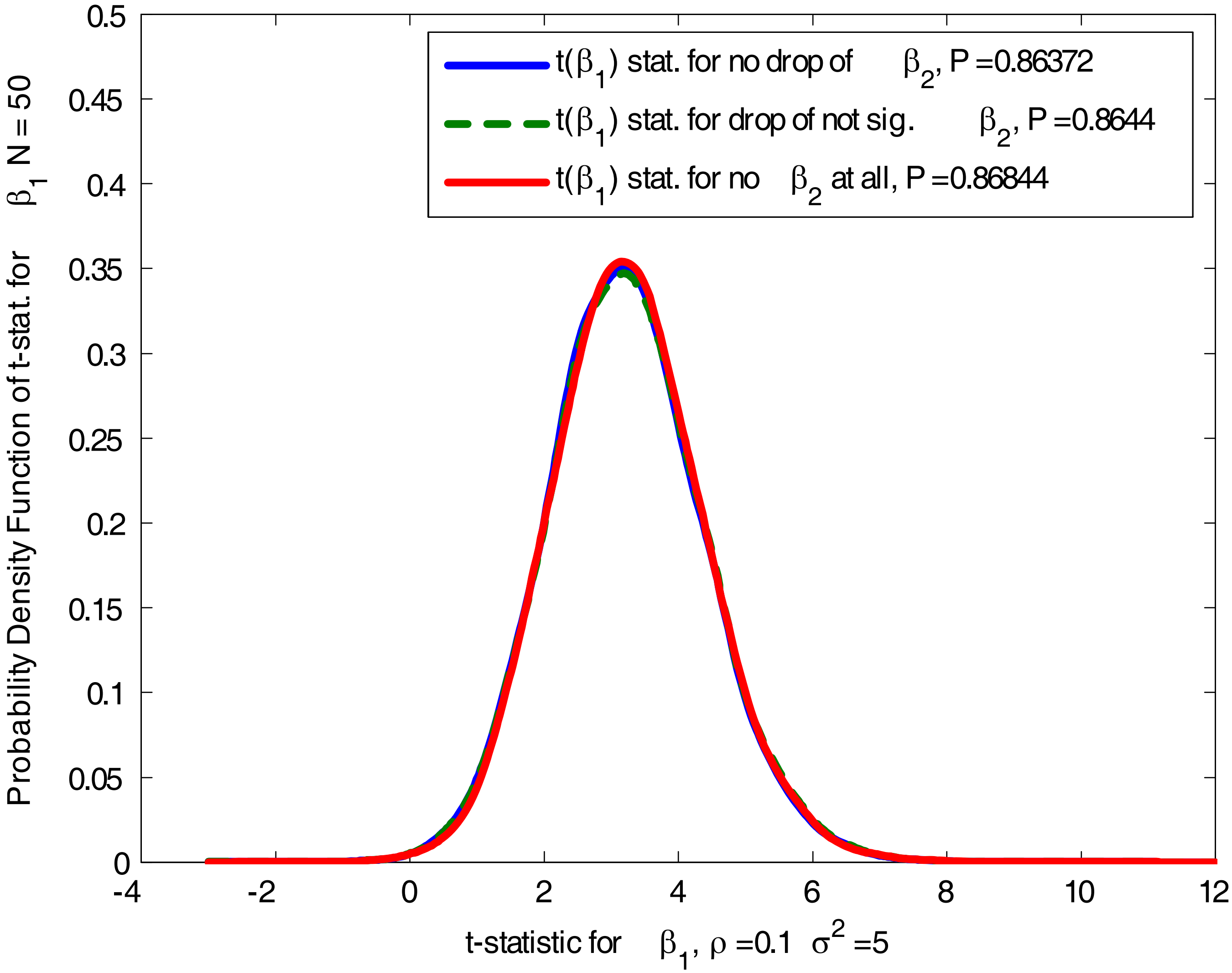


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

$$N = 50, \quad \rho = 0.05; \quad \sigma^2 = 5$$

Pre-Test T-statistic

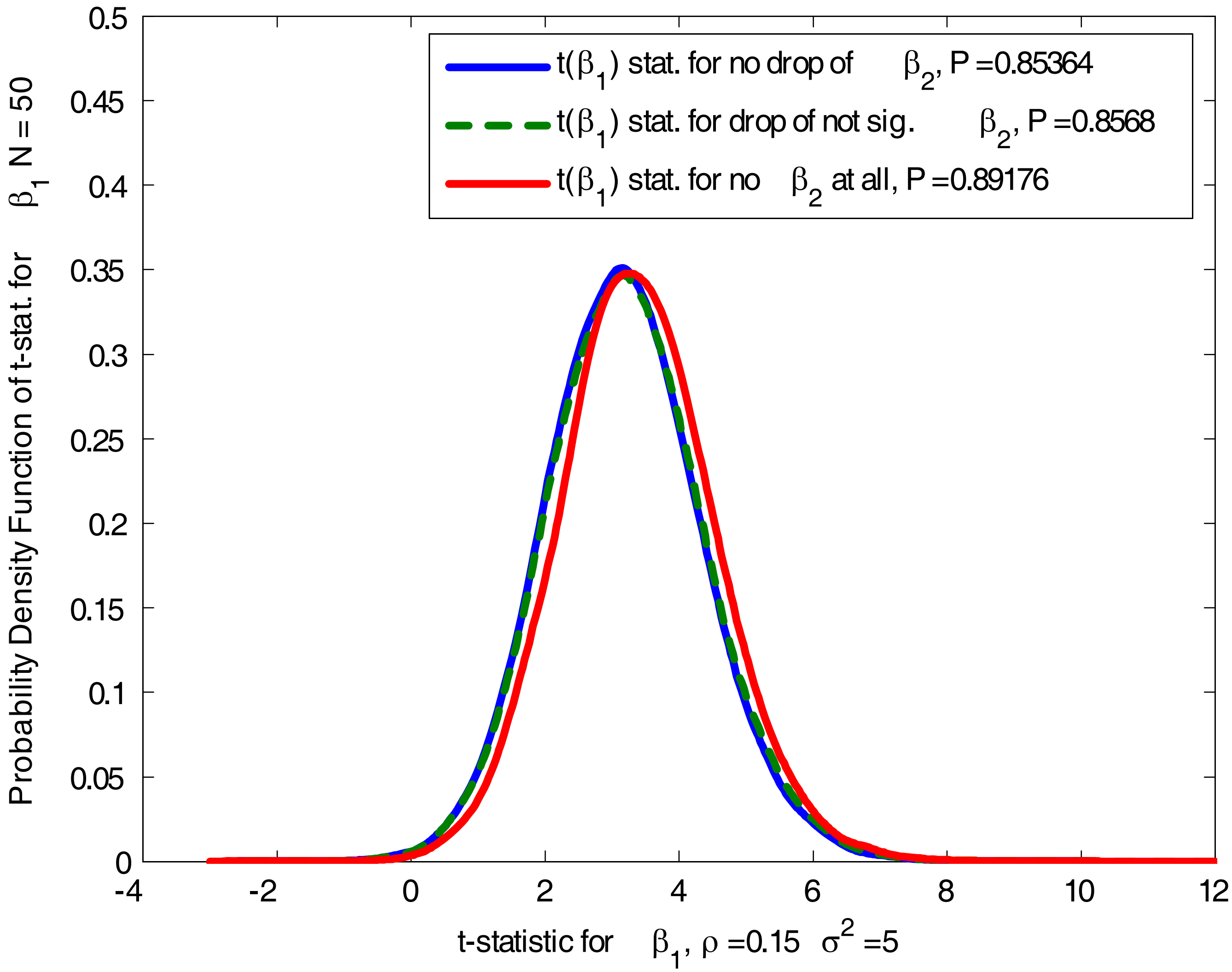


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

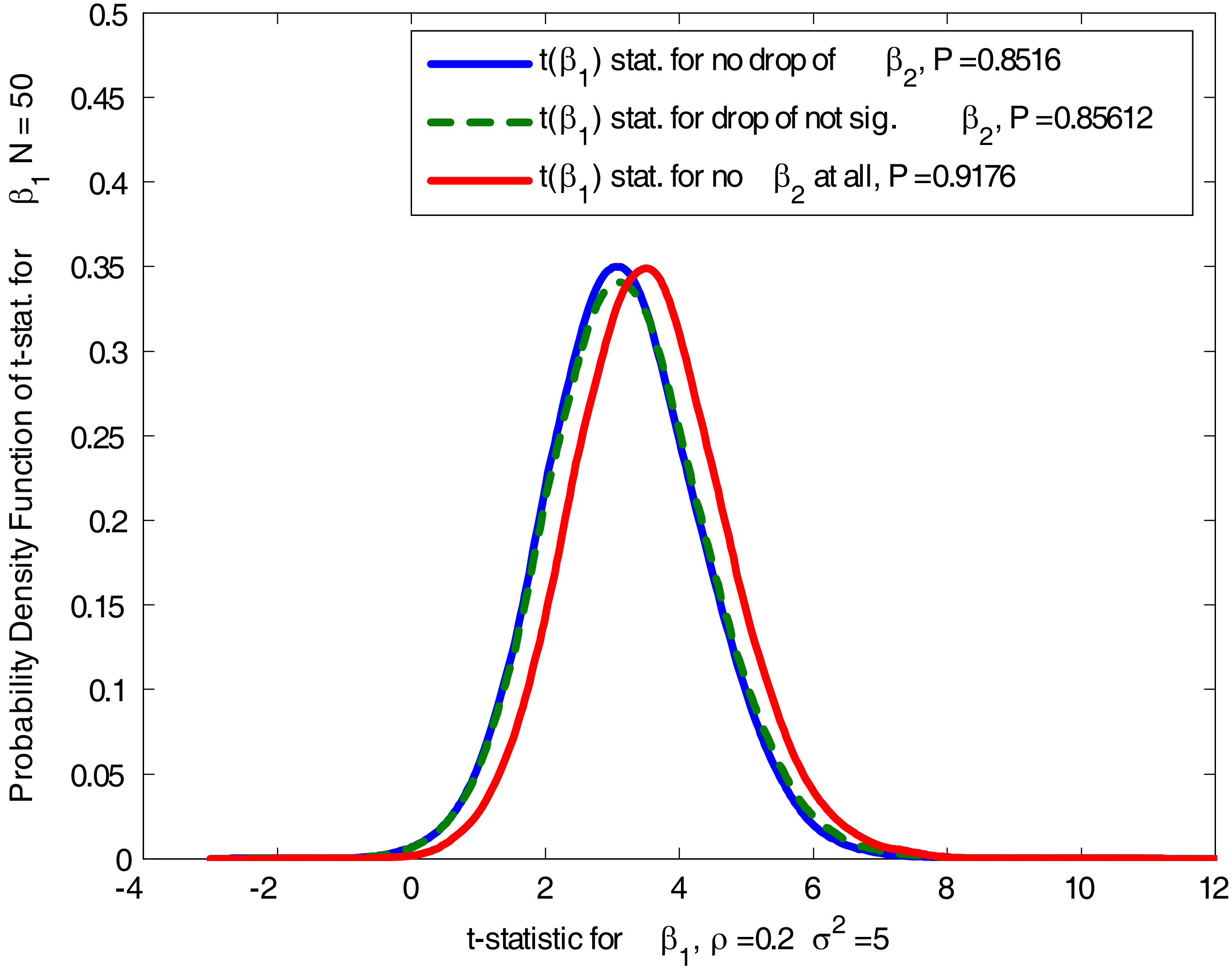
$$N = 50, \quad \rho = 0.10; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.15; \quad \sigma^2 = 5$$

Pre-Test T-statistic

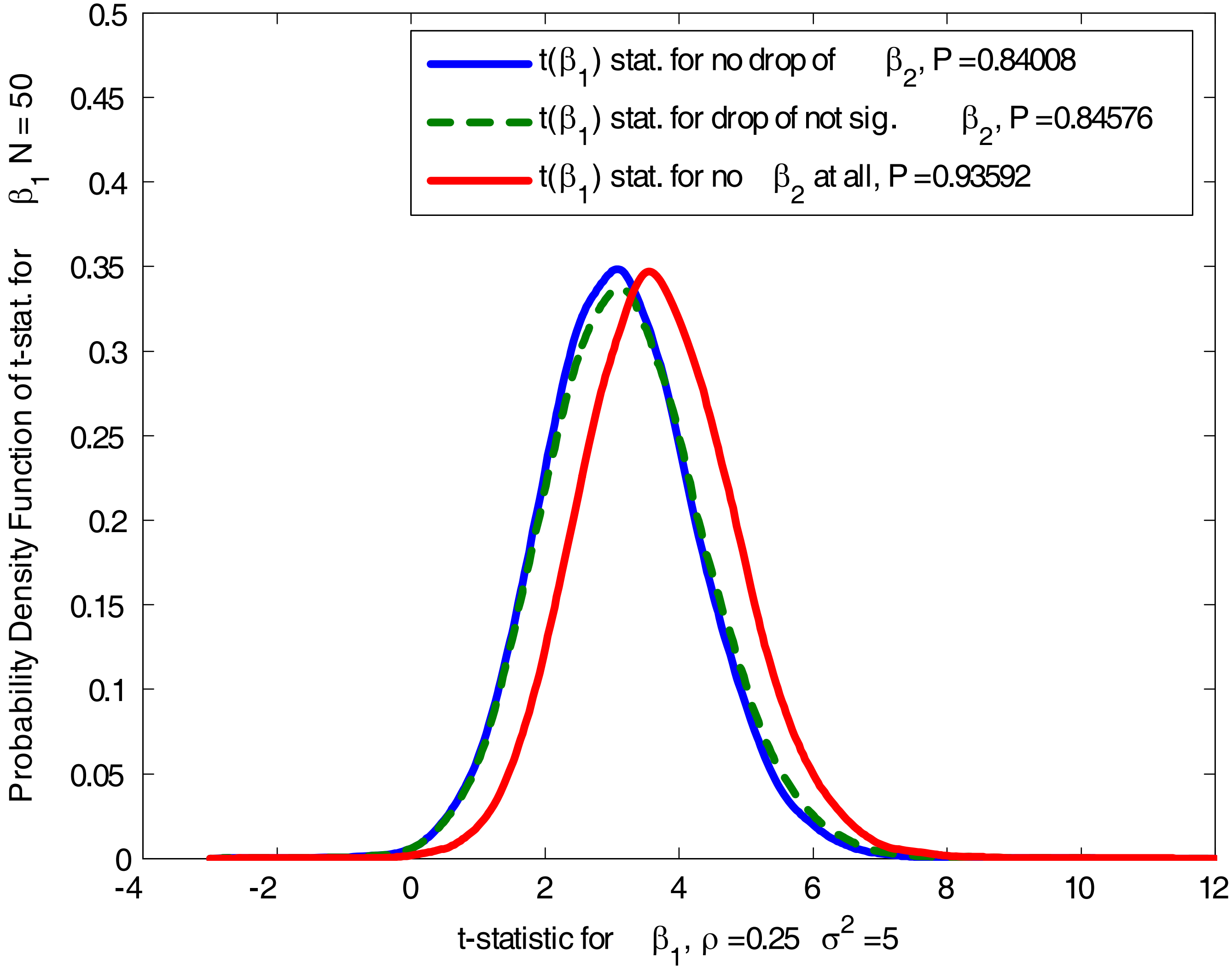


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

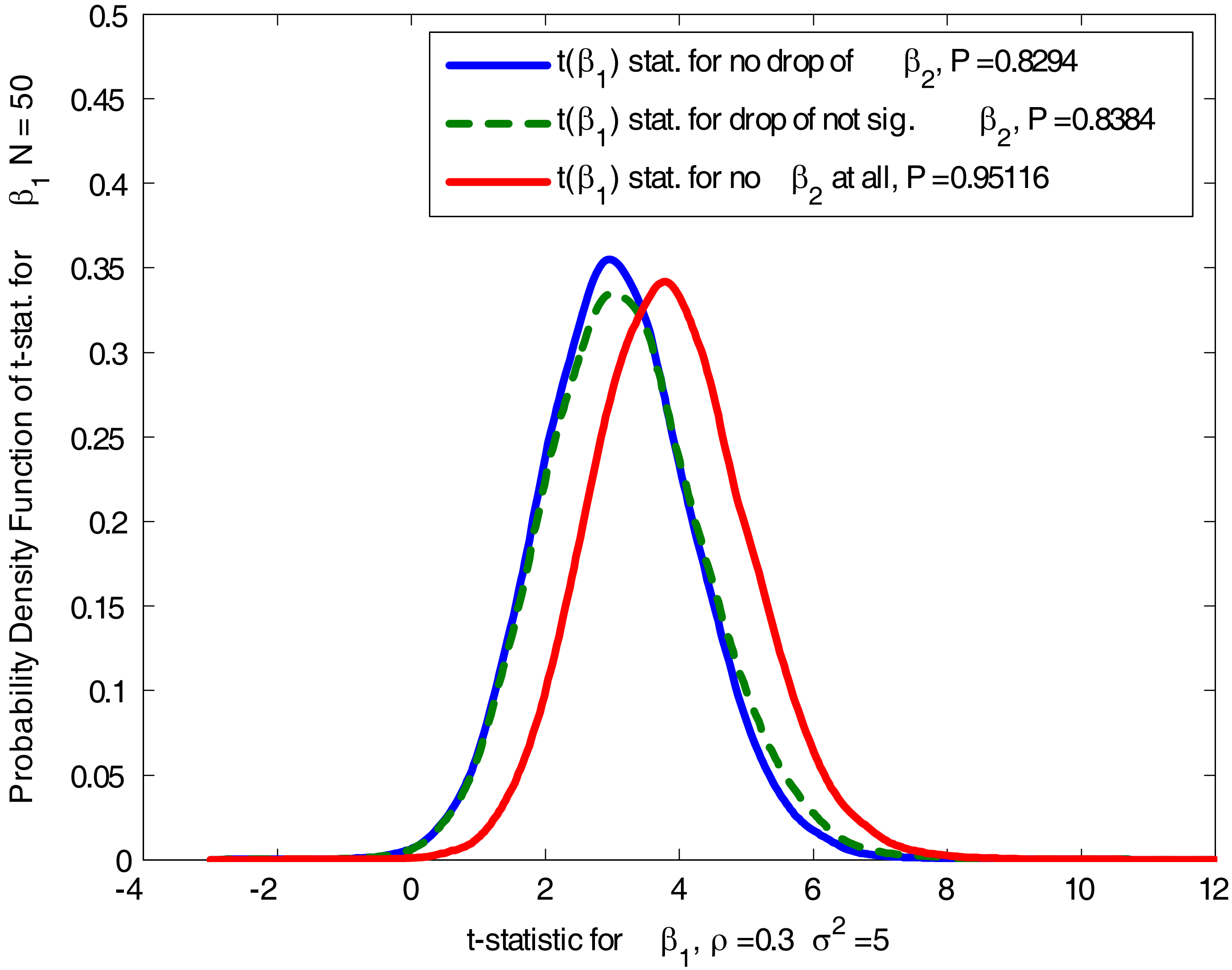
$$N = 50, \quad \rho = 0.20; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 5$$

Pre-Test T-statistic

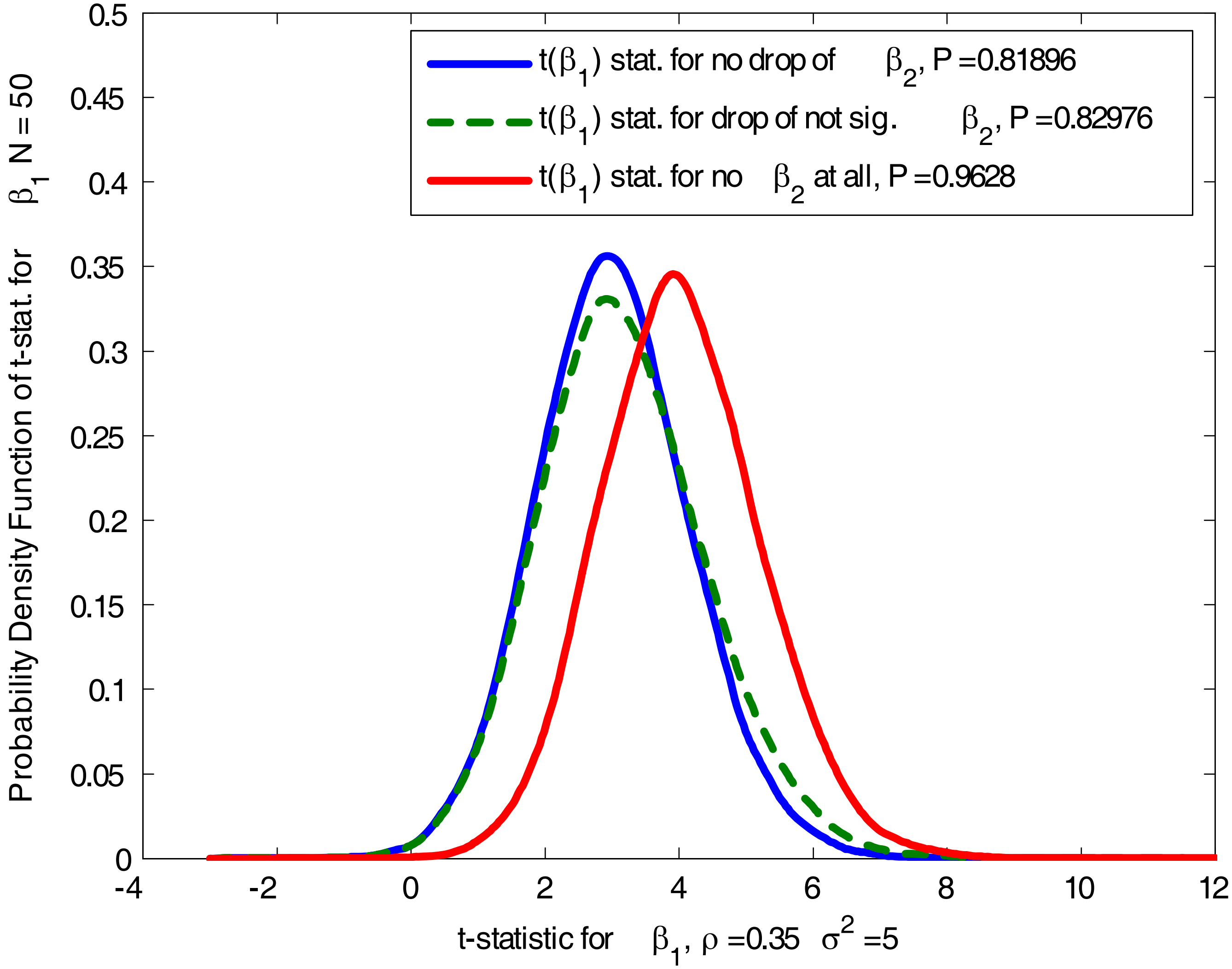


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

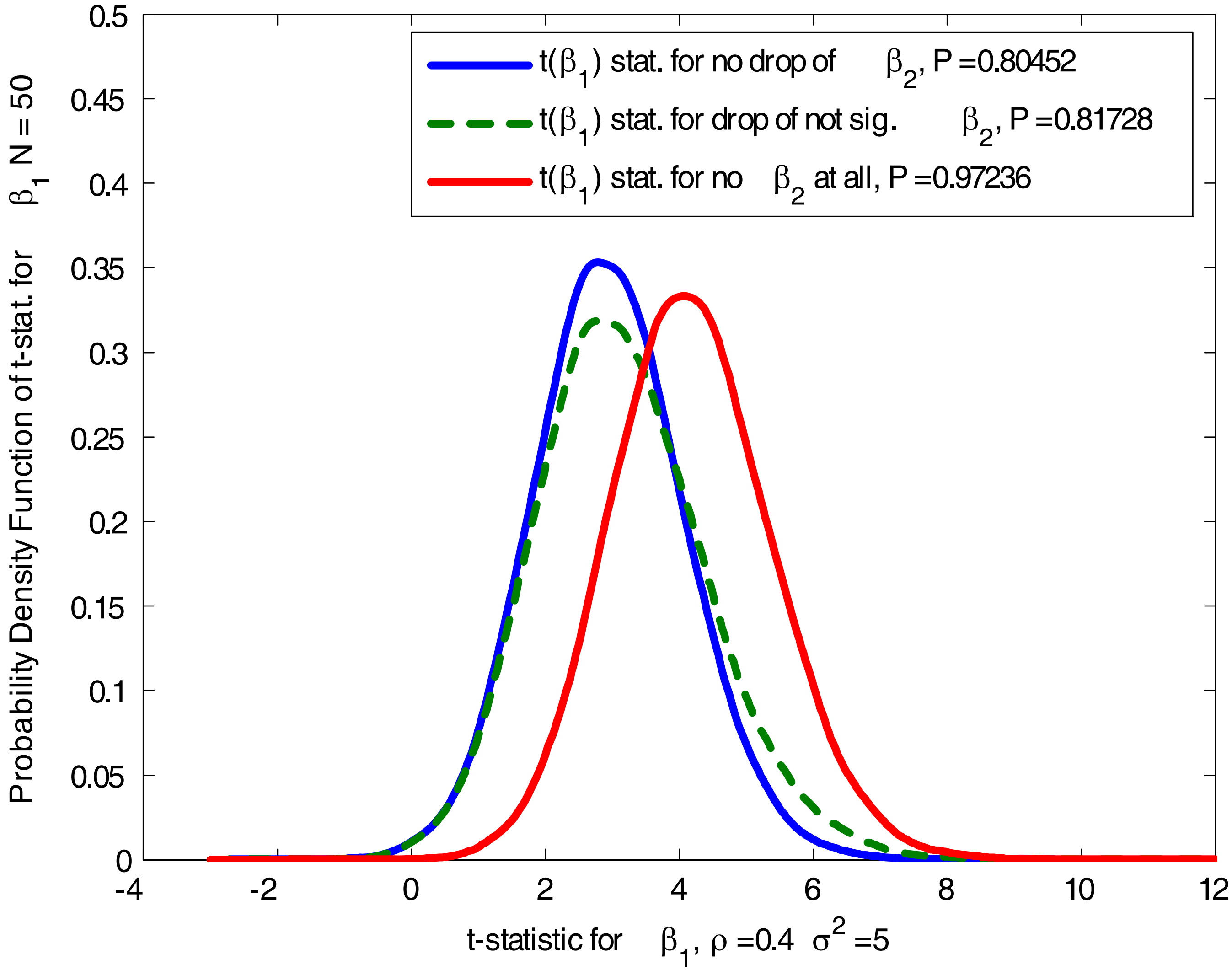
$$N = 50, \quad \rho = 0.30; \quad \sigma^2 = 5$$

Pre-Test T-statistic



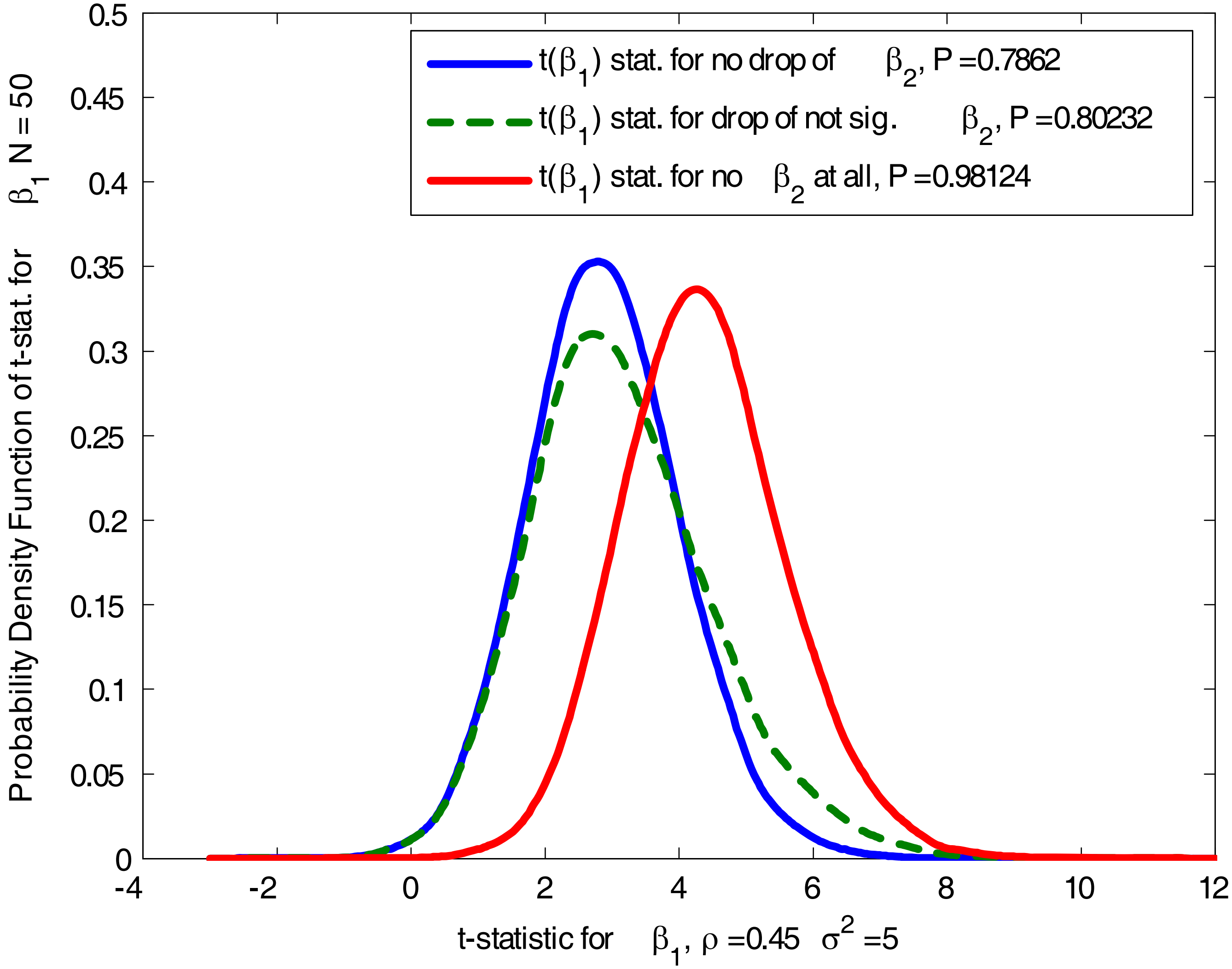
$$Distribution\ of\ T - statistics\quad for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.35; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.40; \quad \sigma^2 = 5$$

Pre-Test T-statistic

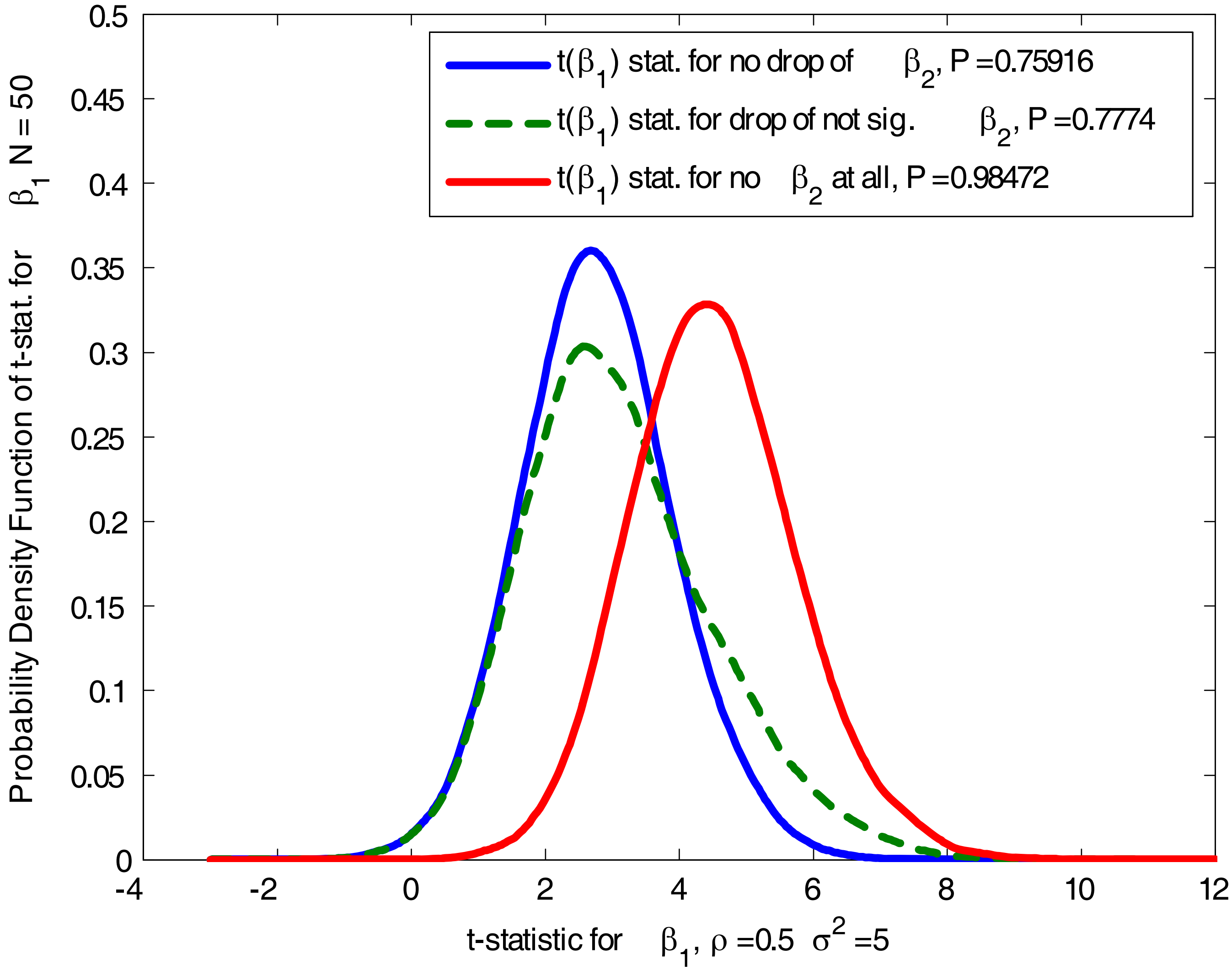


$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

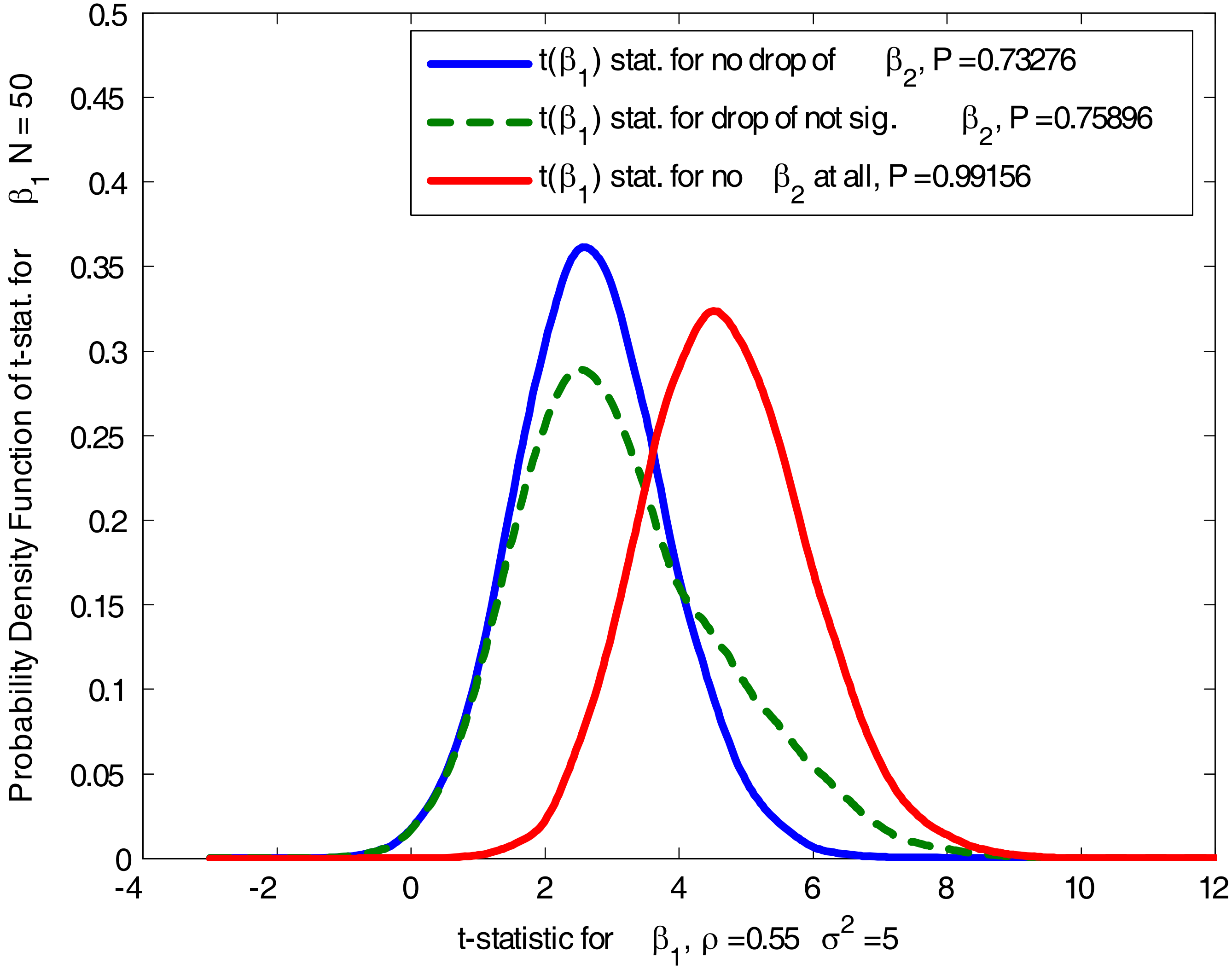
$$N = 50, \quad \rho = 0.45; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.50; \quad \sigma^2 = 5$$

Pre-Test T-statistic

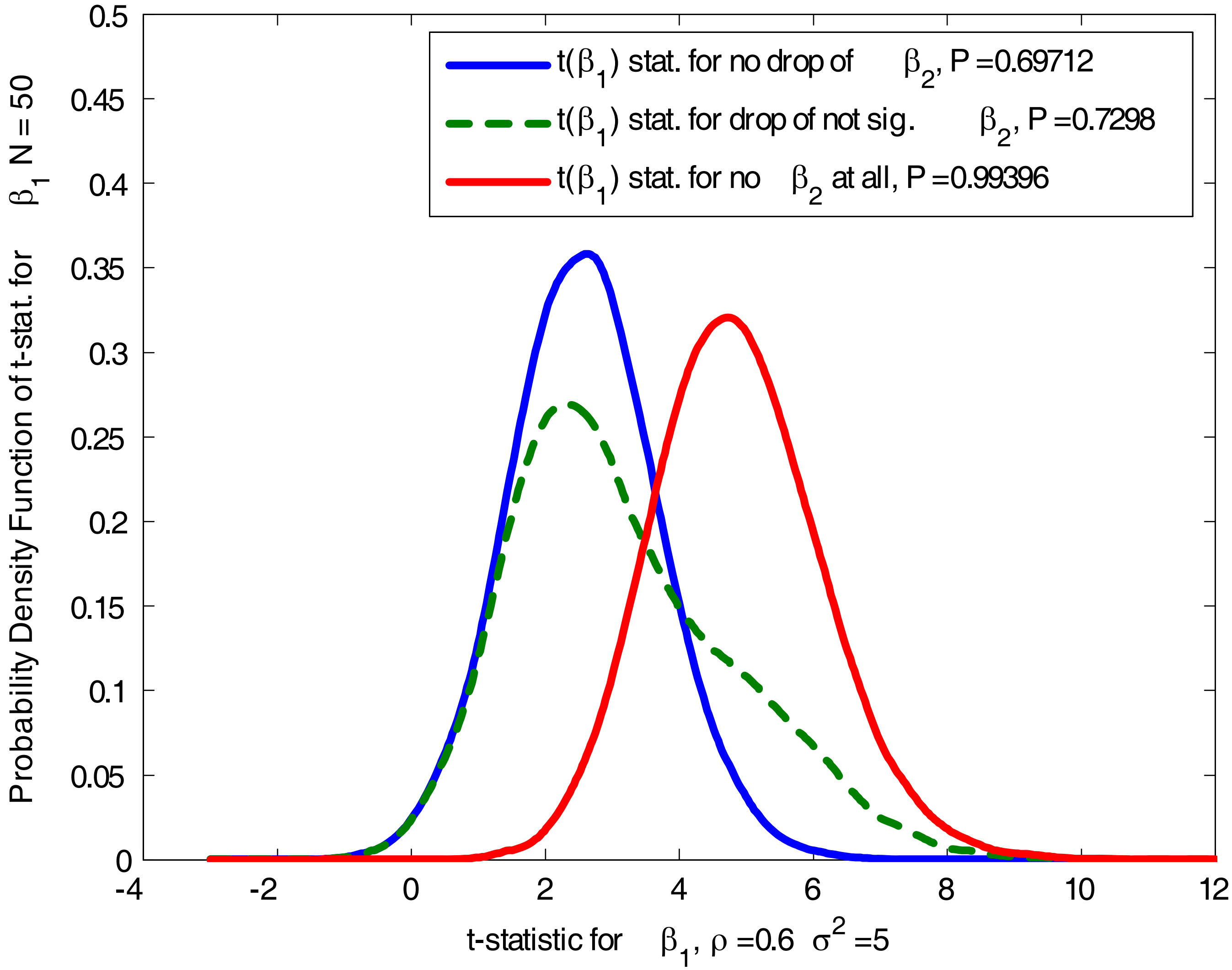


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

$$N = 50, \quad \rho = 0.55; \quad \sigma^2 = 5$$

Pre-Test T-statistic

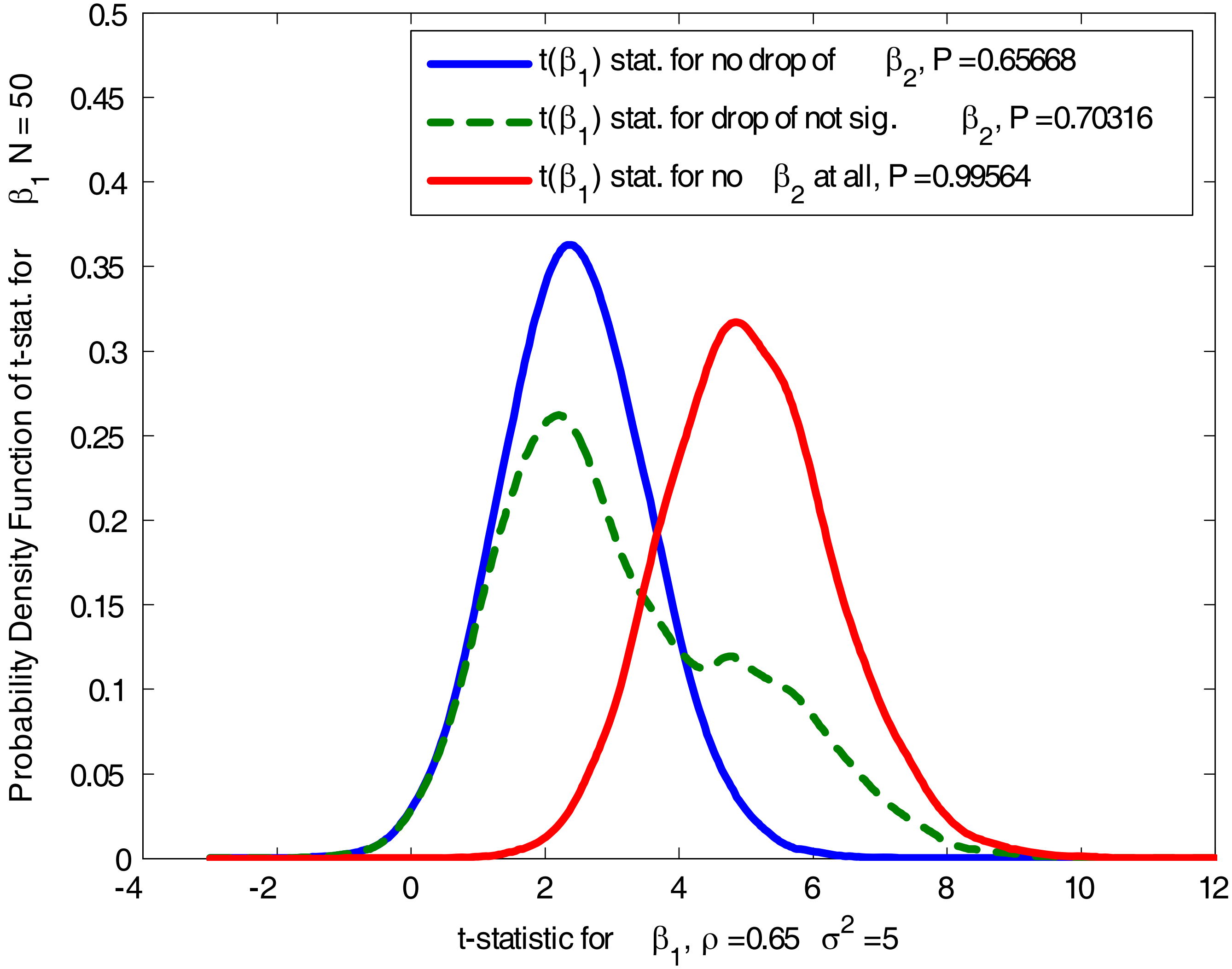


$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

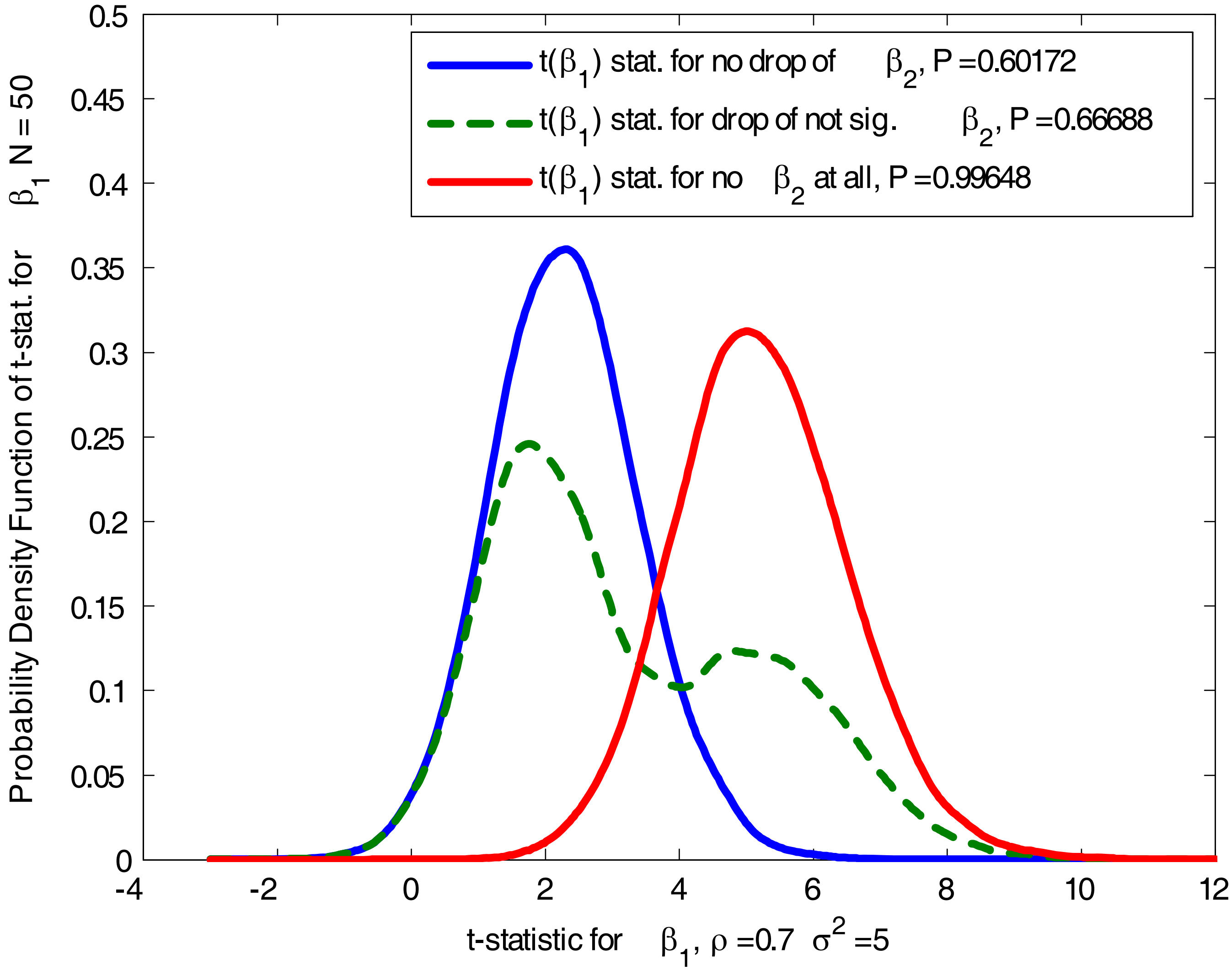
$$N = 50, \quad \rho = 0.60; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.65; \quad \sigma^2 = 5$$

Pre-Test T-statistic

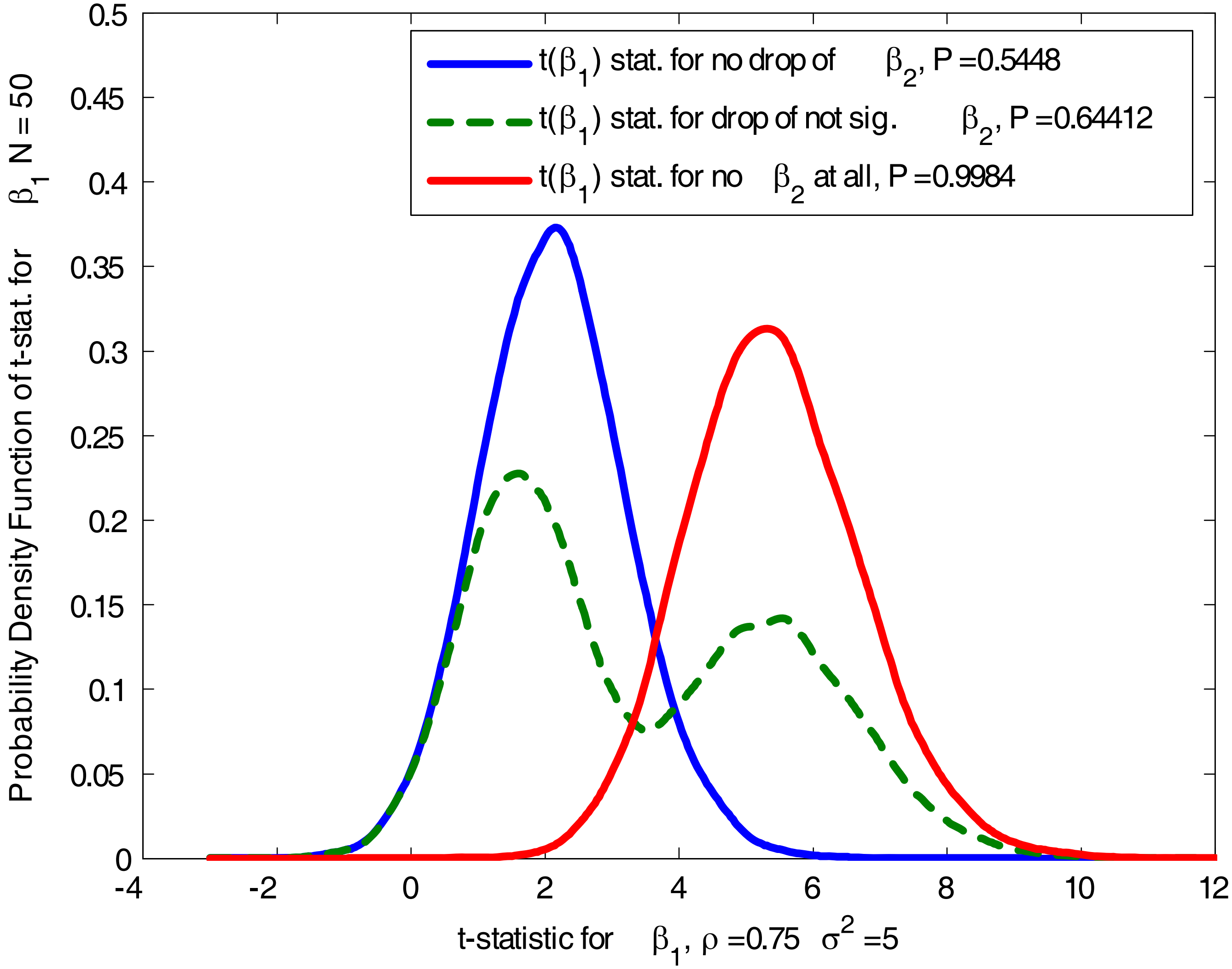


$$\textit{Distribution of } T-\textit{statistics for } \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

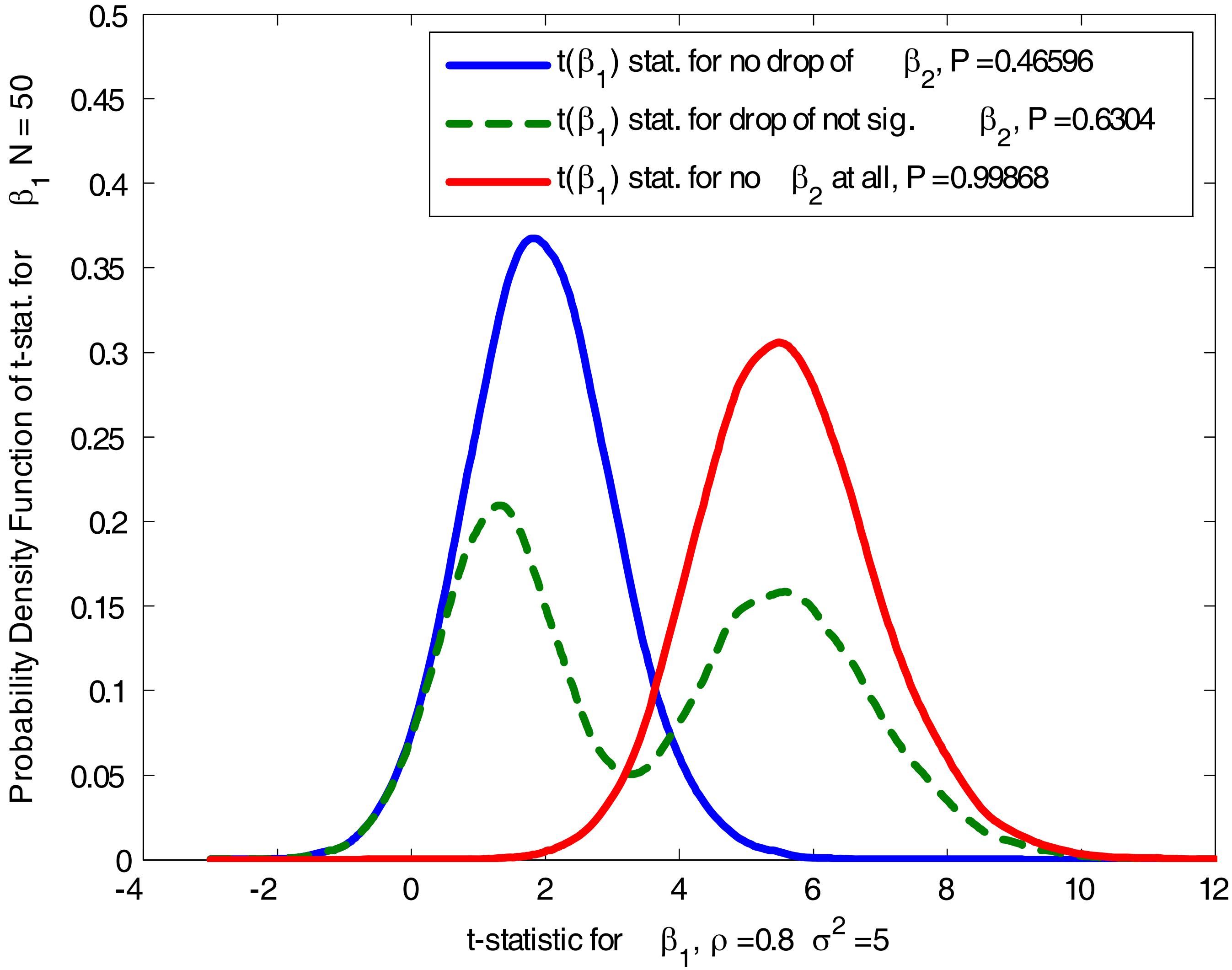
$$N = 50, \quad \rho = 0.70; \quad \sigma^2 = 5$$

Pre-Test T-statistic



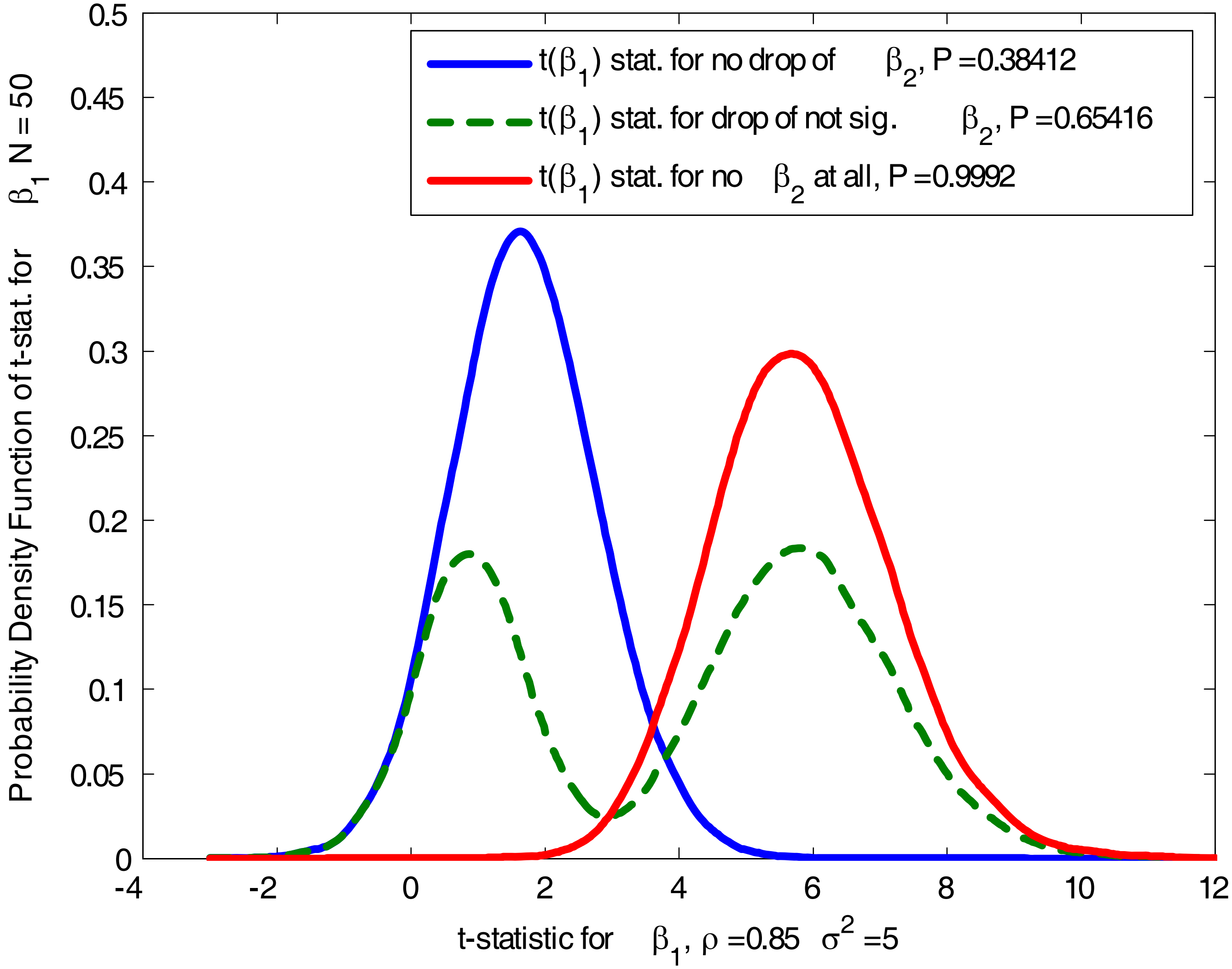
$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 5$$

Pre-Test T-statistic



$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$
$$N = 50, \quad \rho = 0.80; \quad \sigma^2 = 5$$

Pre-Test T-statistic

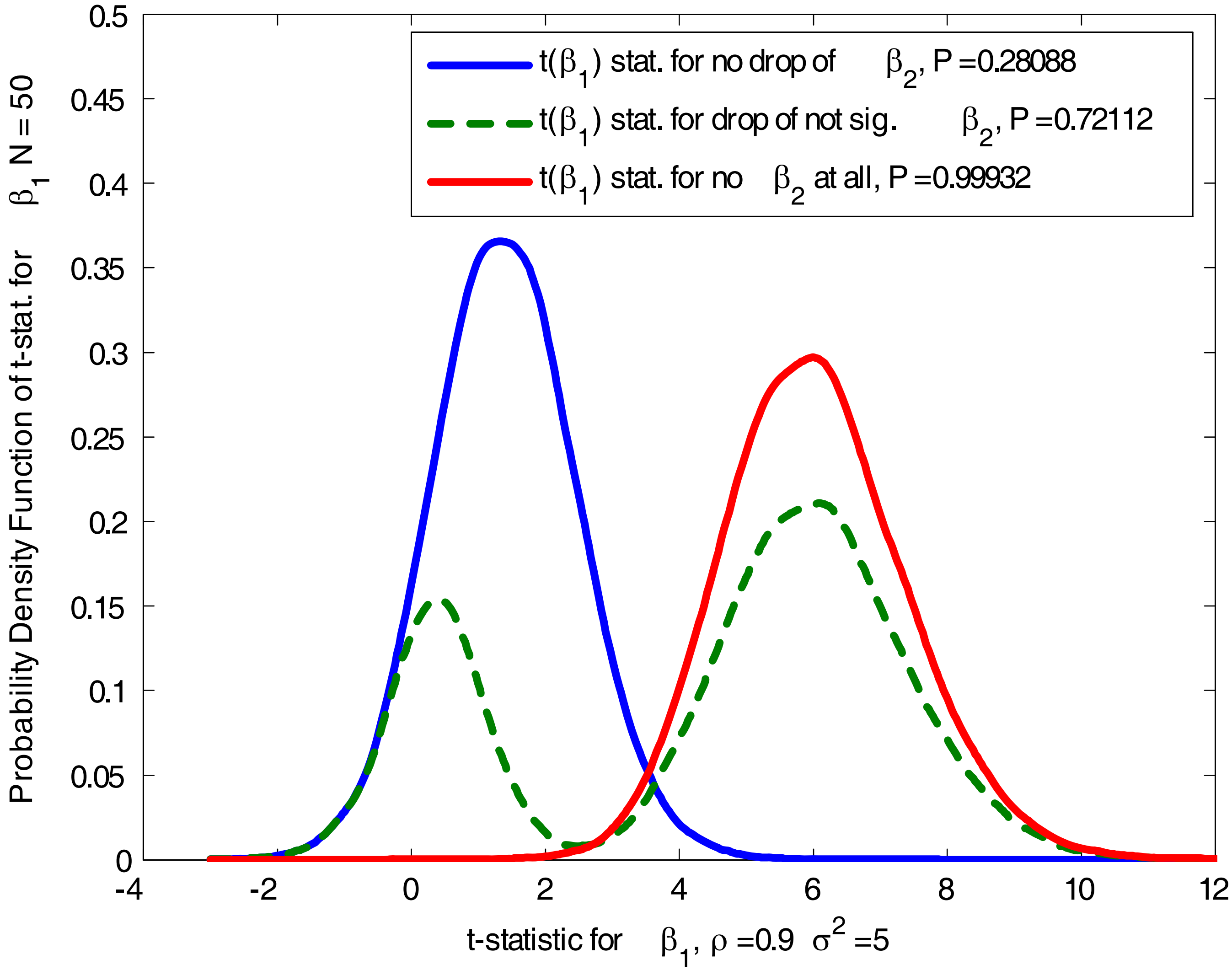


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

$$N = 50, \quad \rho = 0.85; \quad \sigma^2 = 5$$

Pre-Test T-statistic

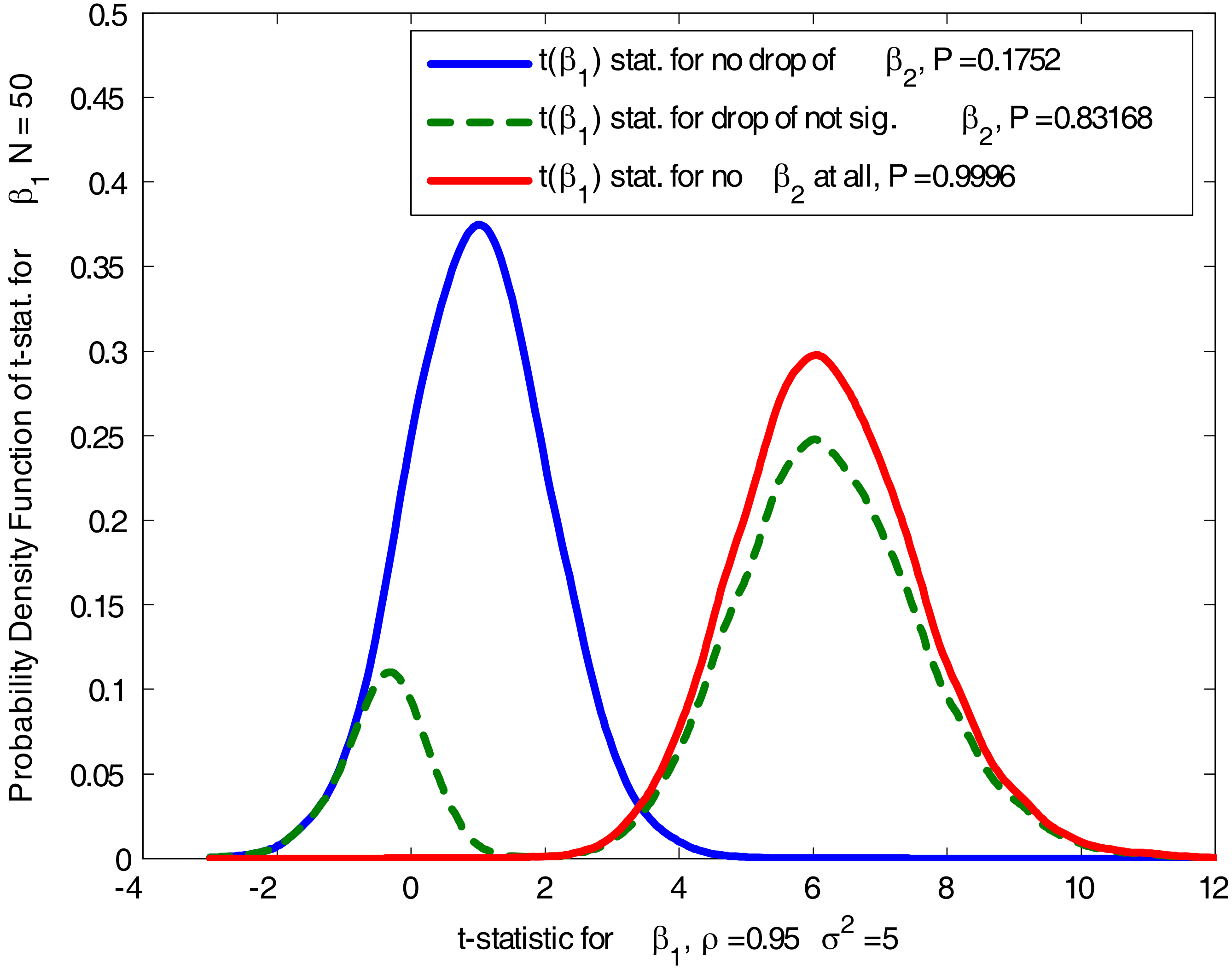


$$Distribution\ of\ T-statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

$$N = 50, \quad \rho = 0.90; \quad \sigma^2 = 5$$

Pre-Test T-statistic

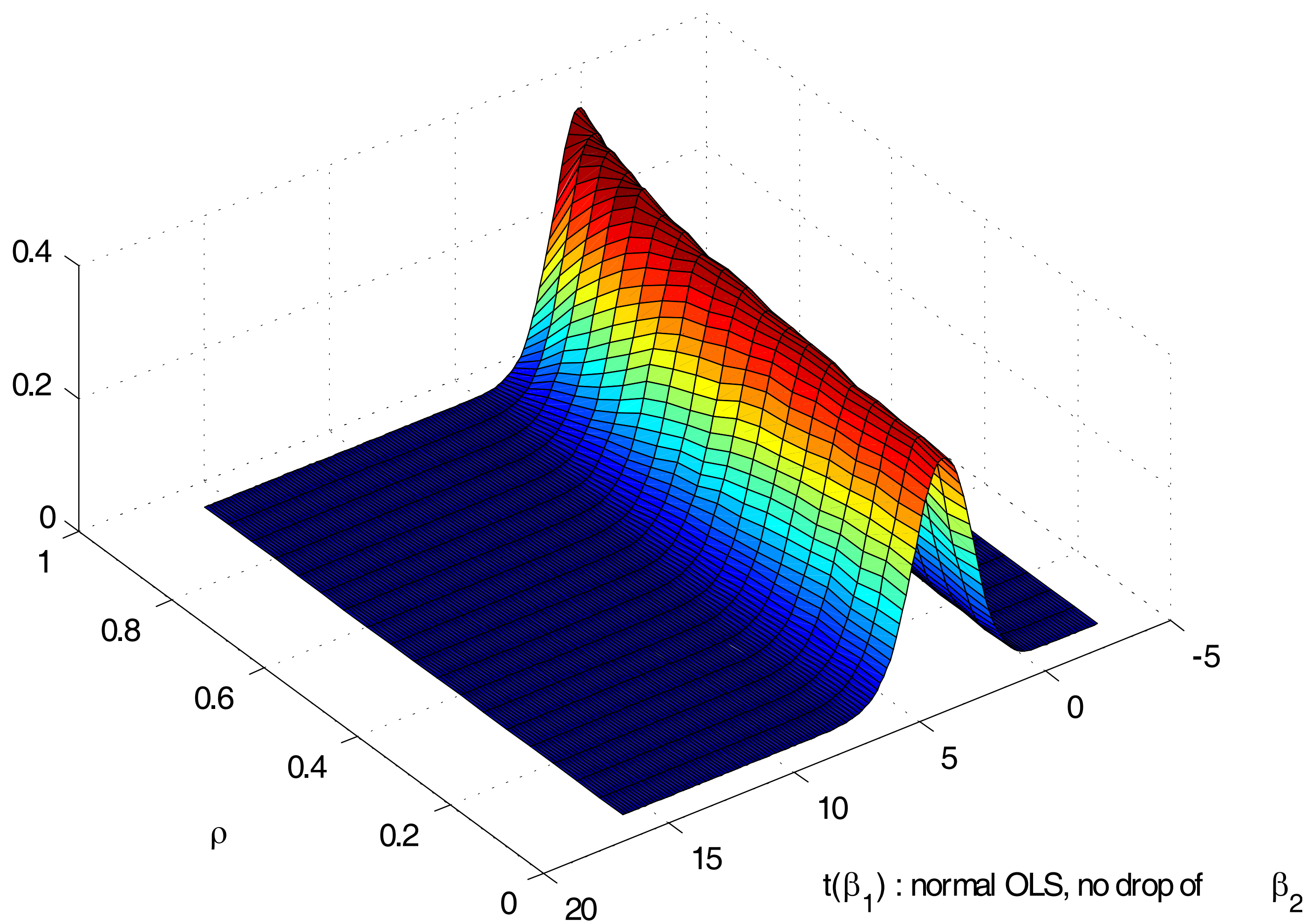


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$

$$P \text{ is the test power} \equiv \Pr(|t|_{\widehat{\beta}_2} > 2)$$

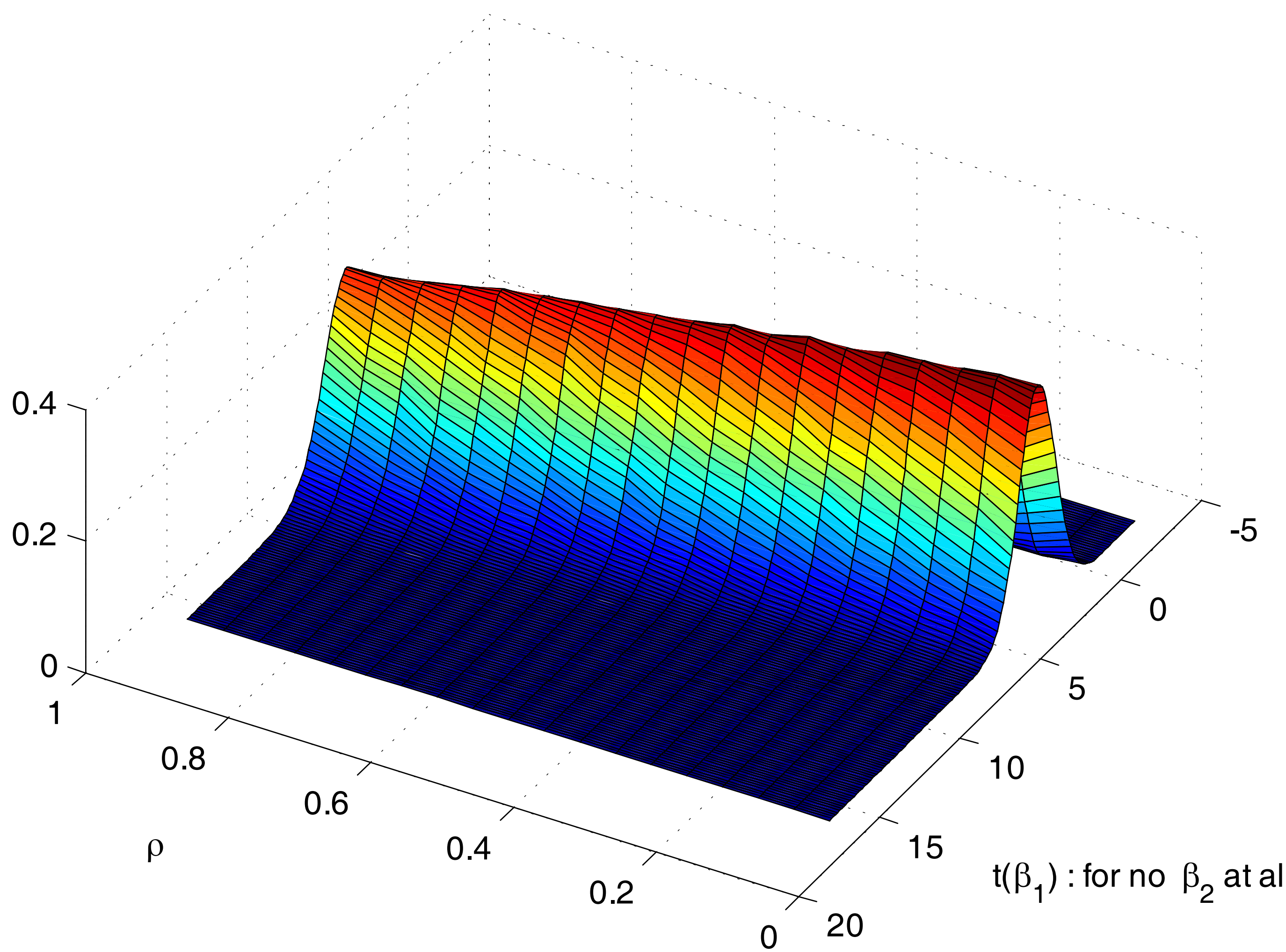
$$N = 50, \quad \rho = 0.95; \quad \sigma^2 = 5$$

T-statistic Density Function for OLS $\widehat{\beta}_1$ Estimator



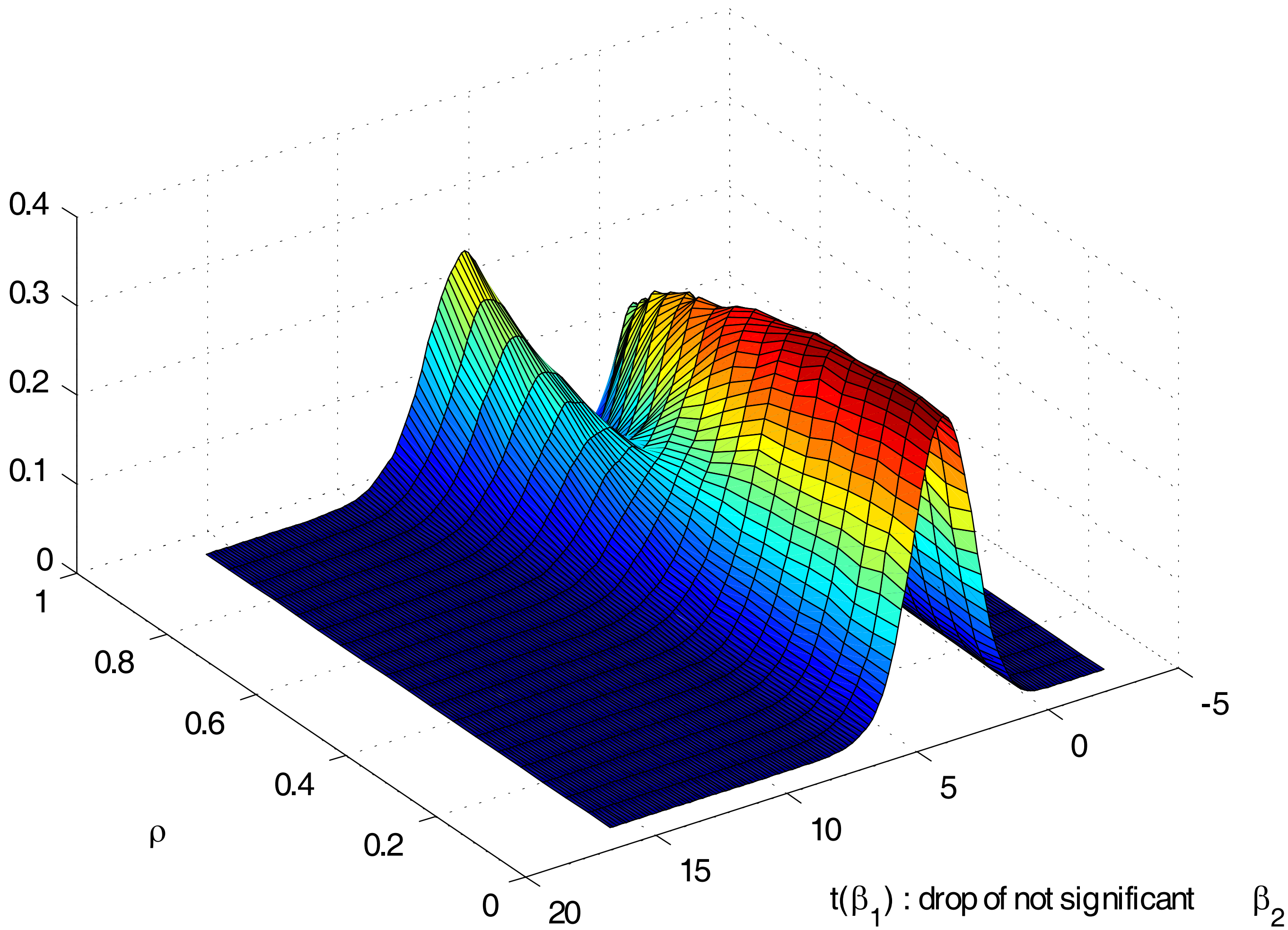
$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$$

T-statistic Density Function for Procedure 2 $\overline{\beta}_1$ Estimator

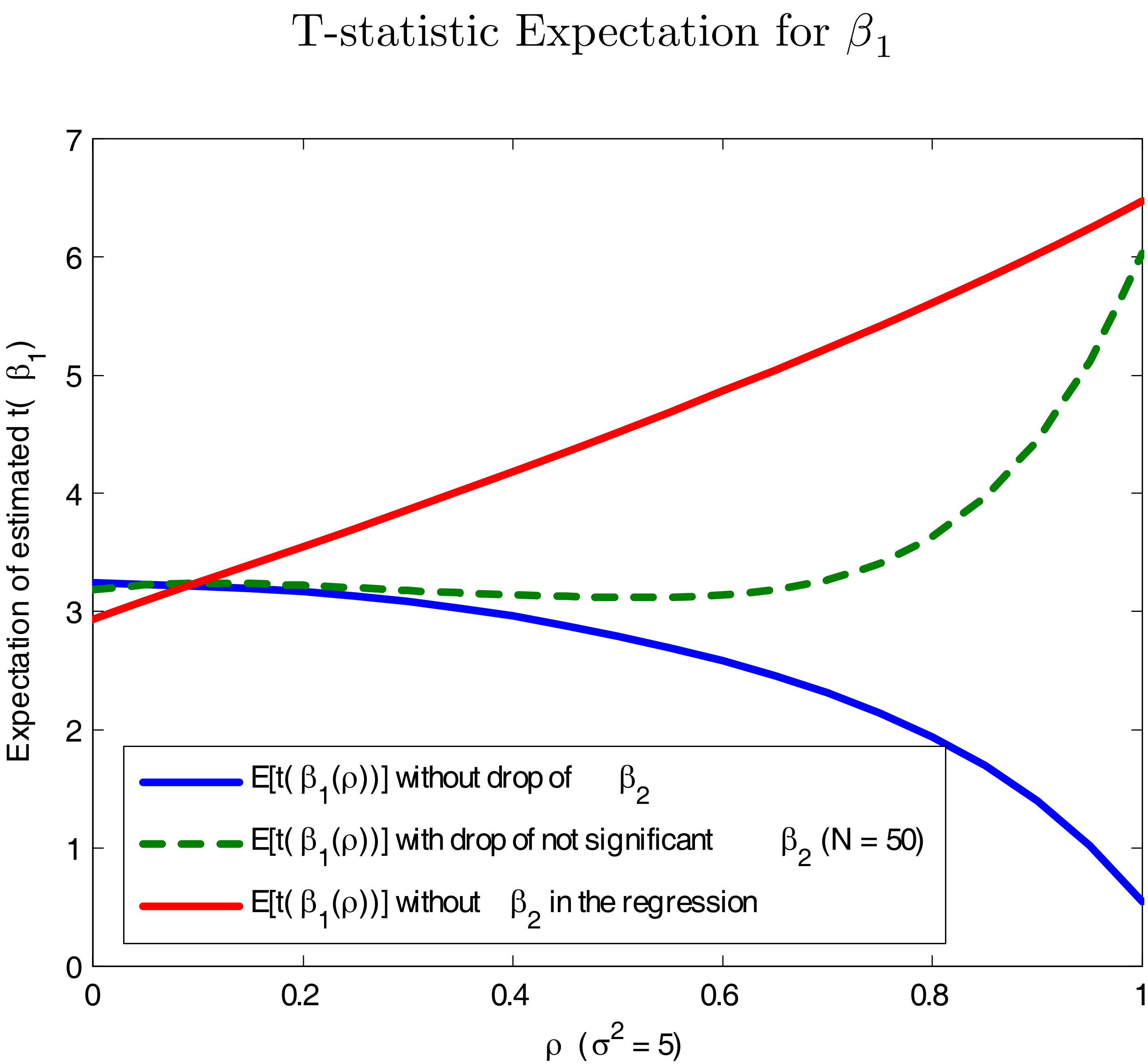


$Distribution\ of\ T-statistics\ \ for\ \beta_1$
$N=50,\quad \sigma^2=5; \rho\in[0,1]$

T-statistic Density Function for Pre-test $\widetilde{\beta}_1$ Estimator

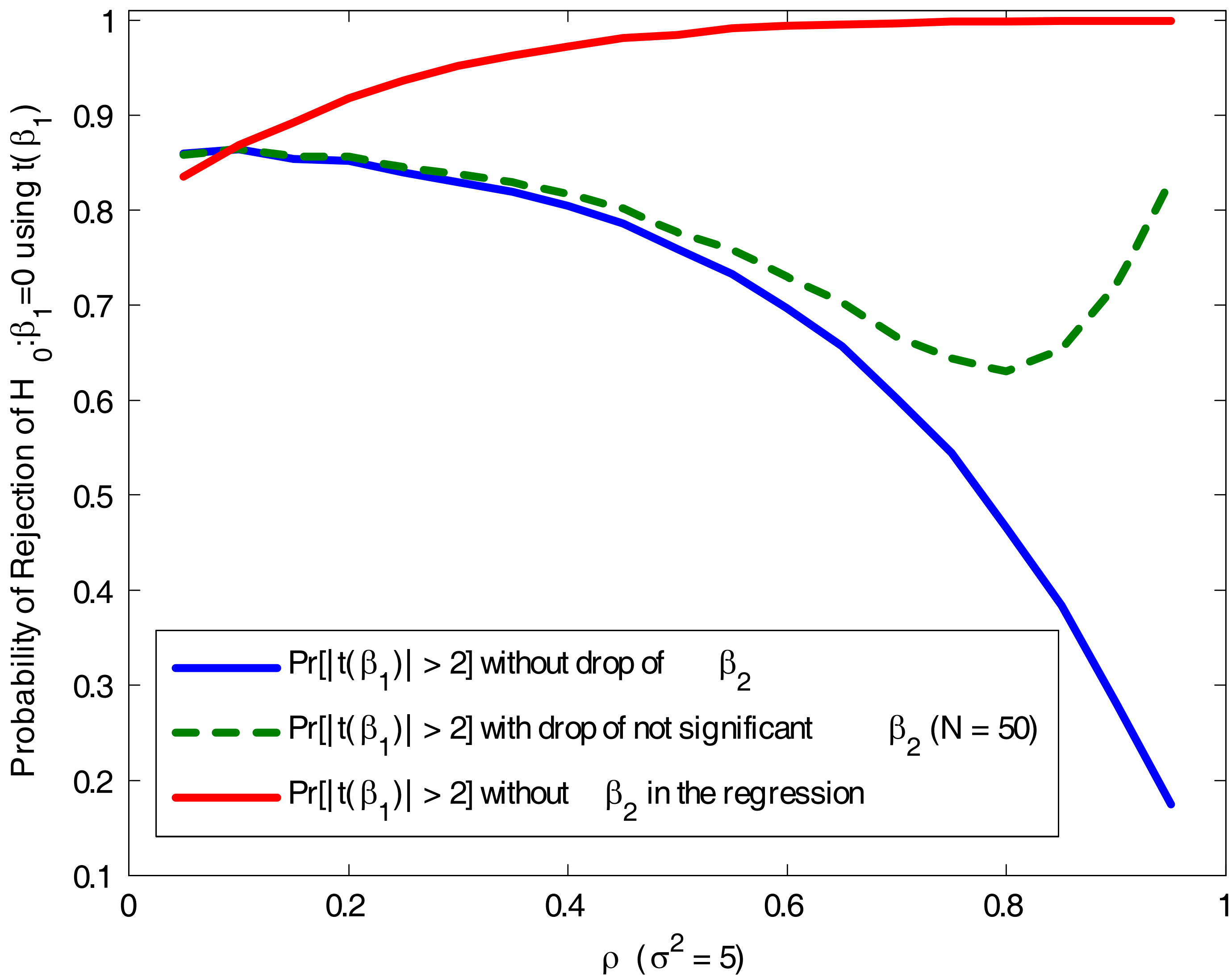


$$Distribution\ of\ T - statistics\ \ for\ \beta_1$$
$$N = 50,\ \ \sigma^2 = 5; \rho \in [0, 1]$$



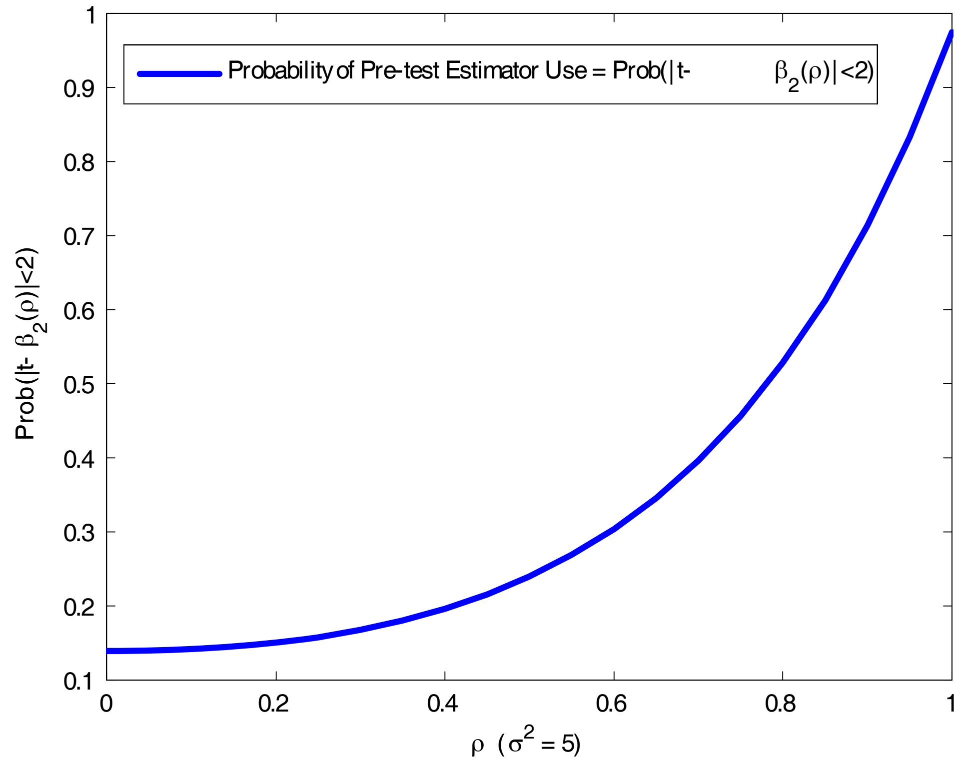
$$Expectation\ of\ T-statistics\ \ for\ \beta_1$$
$$N = 50,\ \ \sigma^2 = 5; \rho \in [0,1]$$

Probability of Rejection of β_1 Using the t-test rule



$$\text{Probability of Rejection } H_0 : \beta_1 = 0 \therefore \Pr(|t|_{\widehat{\beta}_1} > 2)$$
$$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$$

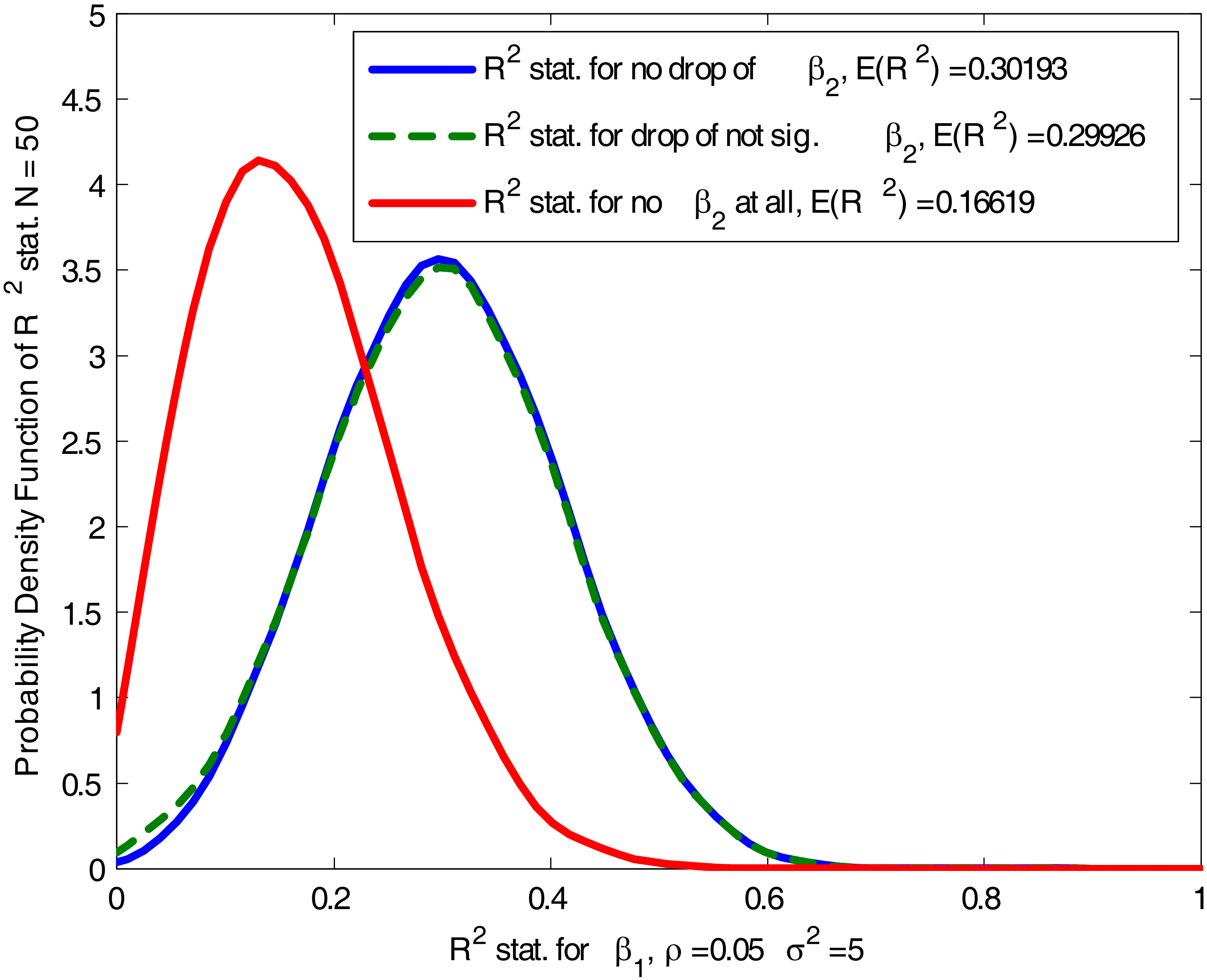
Probability of Use of the Pre-test Estimator $\Pr(|t|_{\widehat{\beta}_2} < 2)$



$$\text{Probability of Non Rejection of } H_0 : \widehat{\beta}_2 = 0$$
$$\Pr(|t|_{\widehat{\beta}_2} < 2)$$
$$N = 50, \quad \sigma^2 = 5; \rho \in [0,1]$$
$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right)$$
$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot (1-\rho^2)}{N}\right)$$
$$(\text{Pre-test})\ \widetilde{\beta}_1 = \begin{cases} \widehat{\beta}_1 \text{ with Prob.} = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \\ \overline{\beta}_1 \text{ with Prob.} = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} < 2\right) \end{cases}$$
$$E\left(\widetilde{\beta}_1\right) = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \cdot \beta_1 + \left[\Pr\left(|\widehat{t}|_{\widehat{\beta}_2} < 2\right)\right] \cdot (\beta_1 + \rho\beta_2)$$

7 The R^2 Analysis

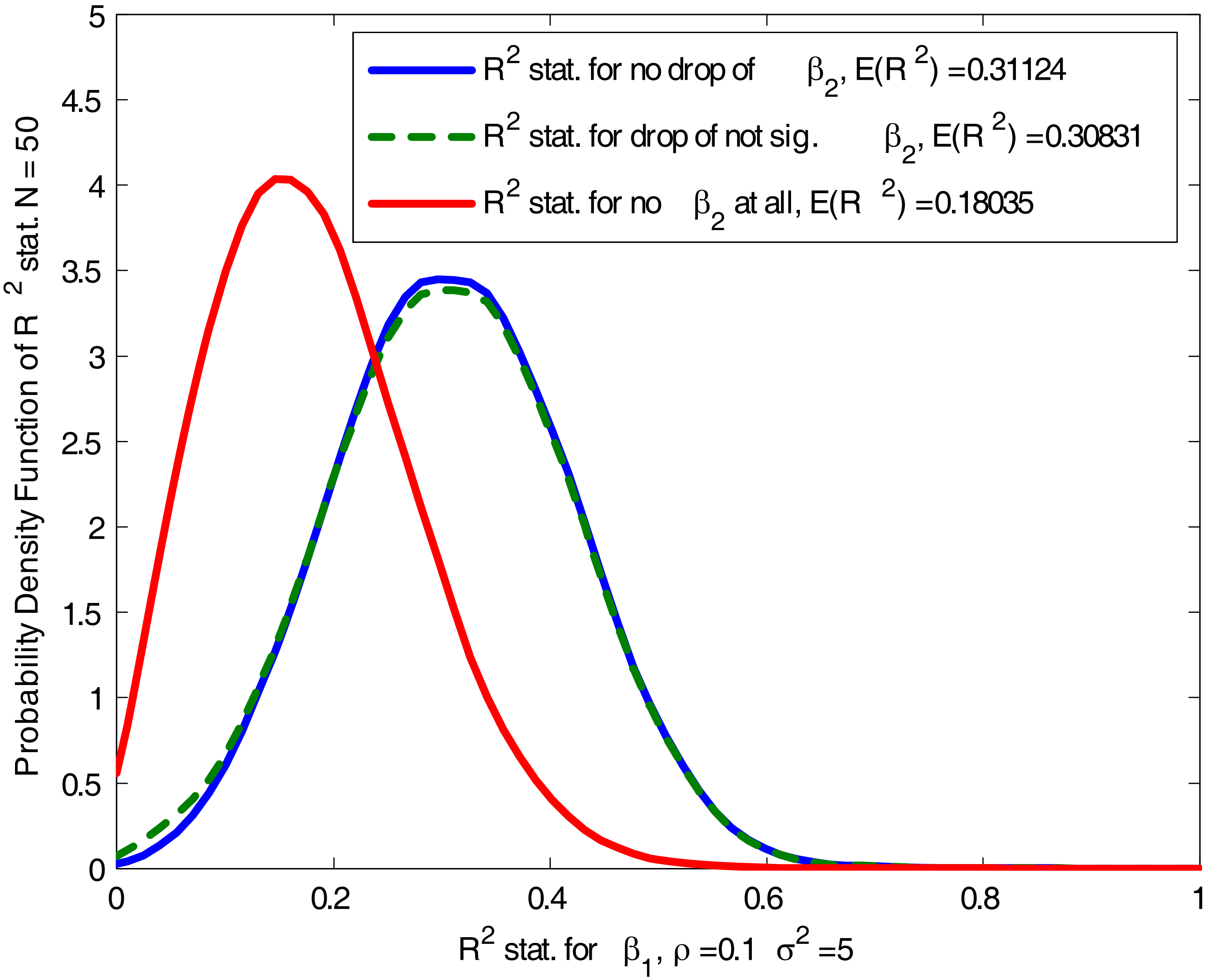
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.05; \quad \sigma^2 = 5$

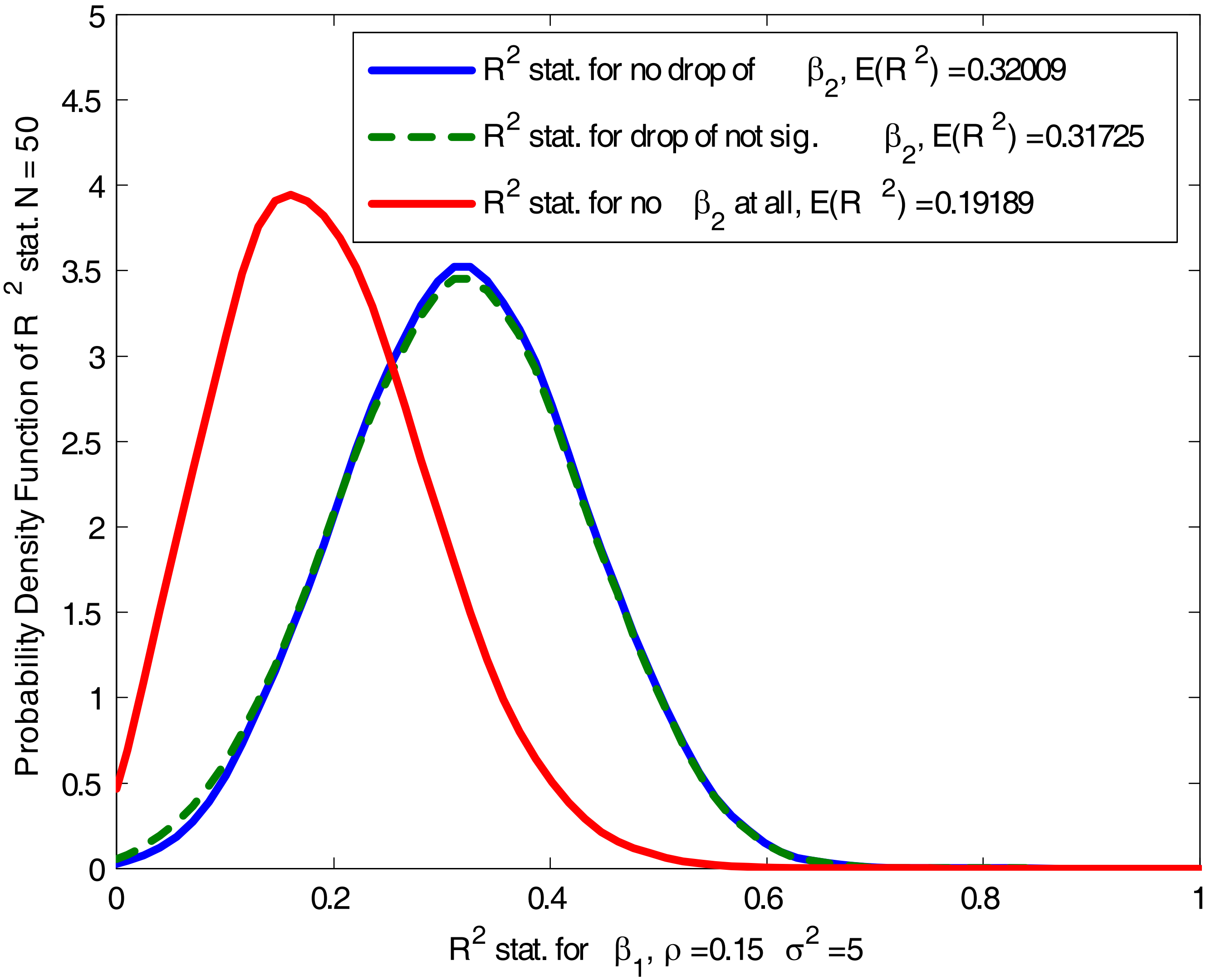
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.10; \quad \sigma^2 = 5$

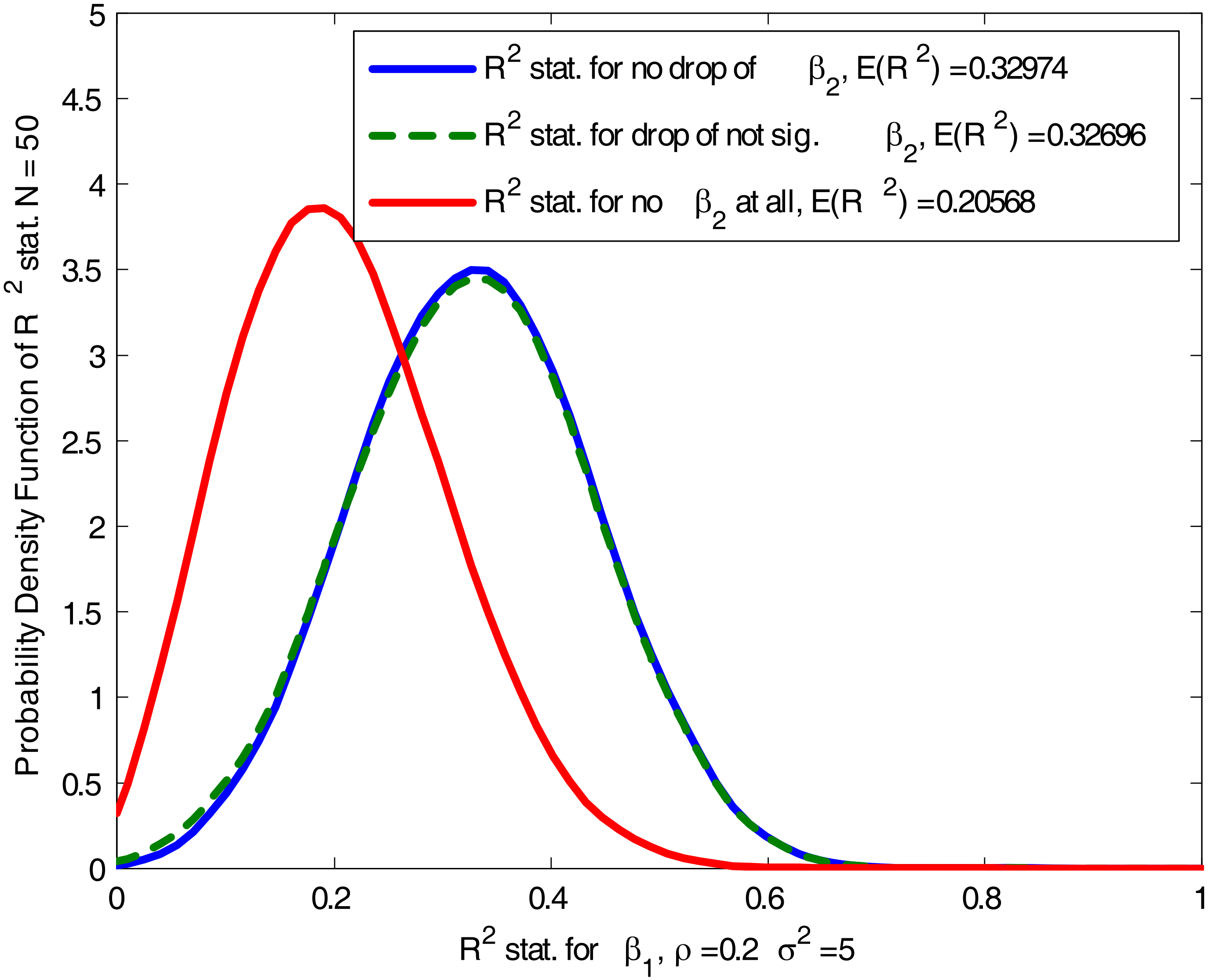
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.15; \quad \sigma^2 = 5$

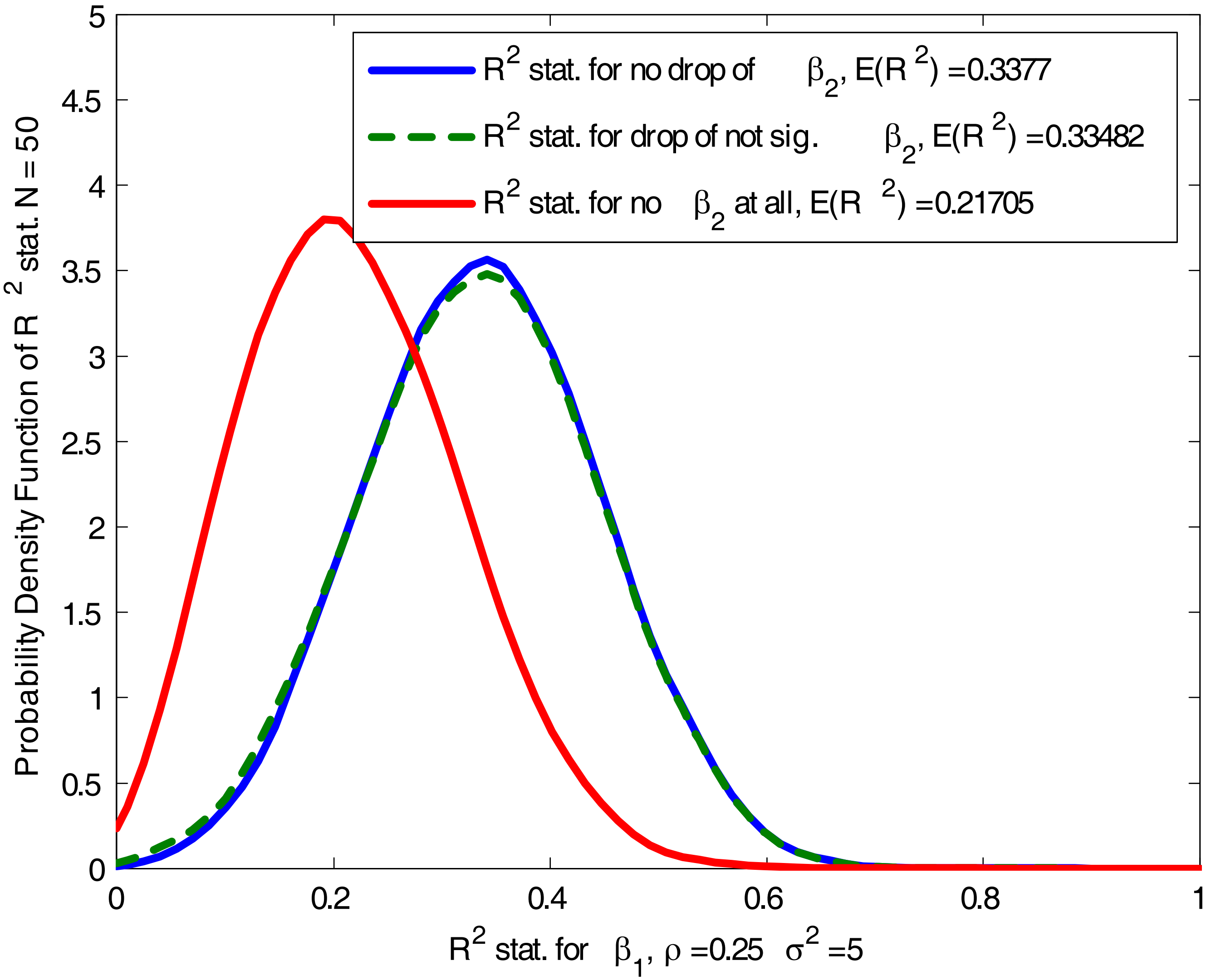
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.20; \quad \sigma^2 = 5$

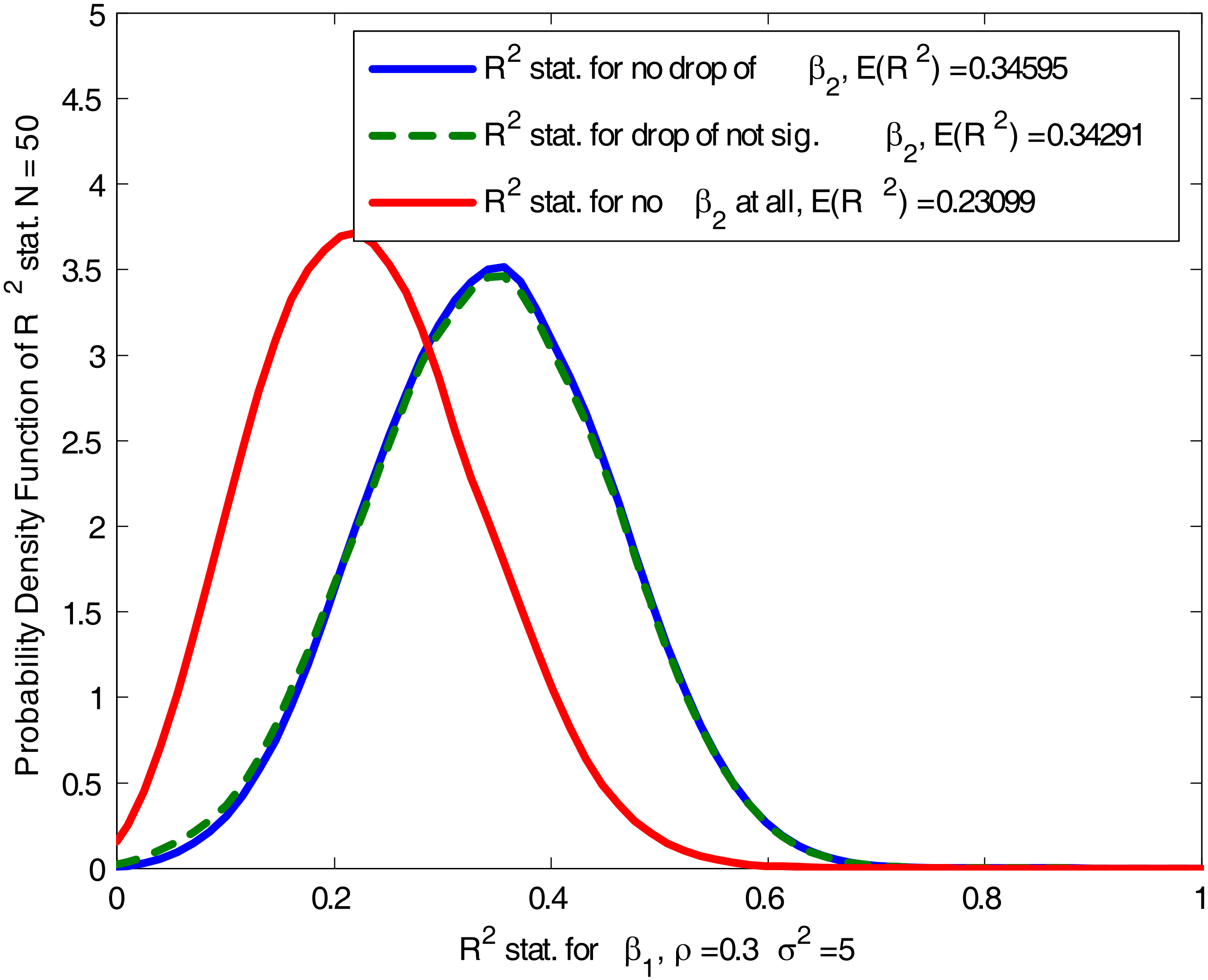
R^2 density Function



Distribution of R^2

$$N = 50, \quad \rho = 0.25; \quad \sigma^2 = 5$$

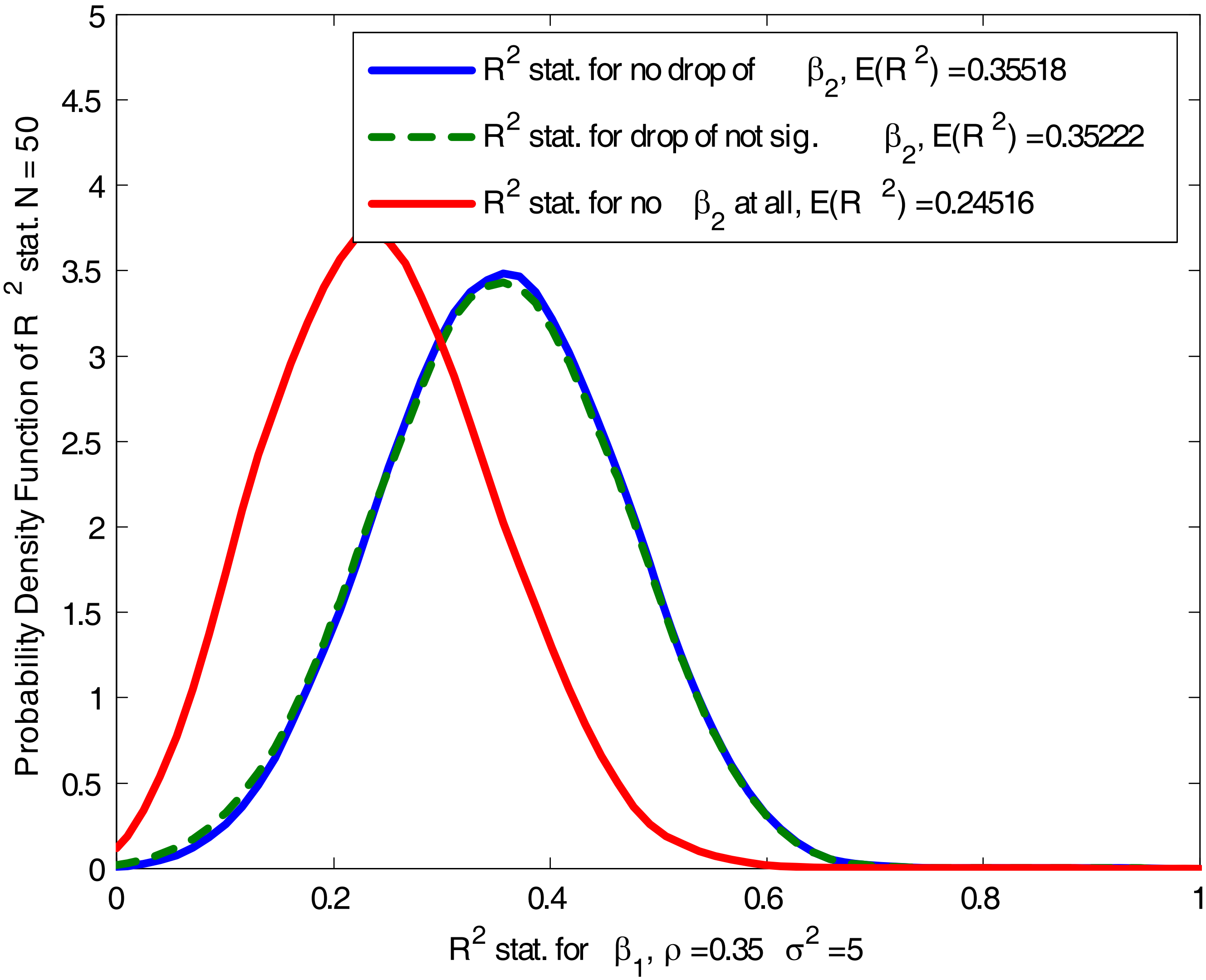
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.30; \quad \sigma^2 = 5$

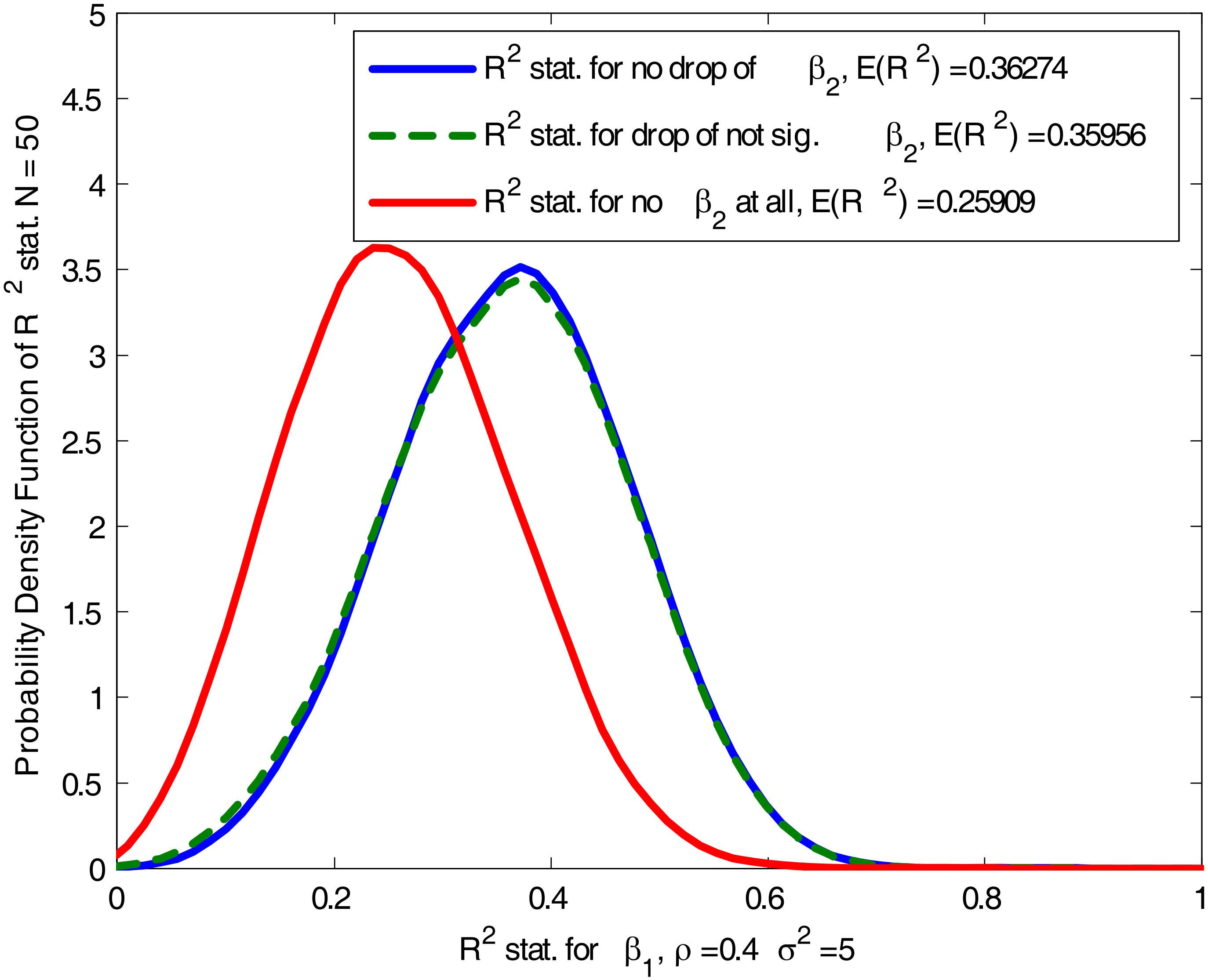
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.35; \quad \sigma^2 = 5$

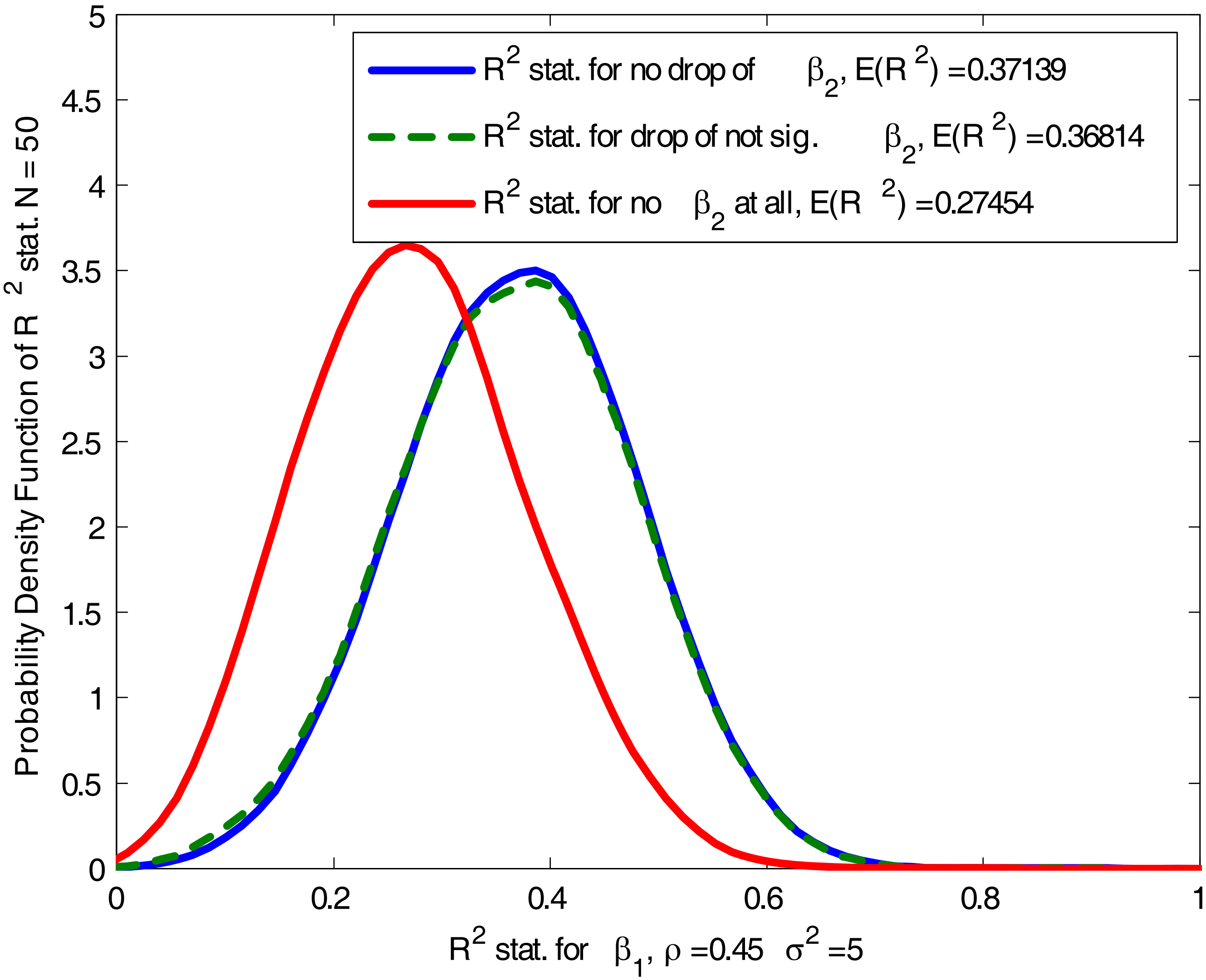
R^2 density Function



Distribution of R^2

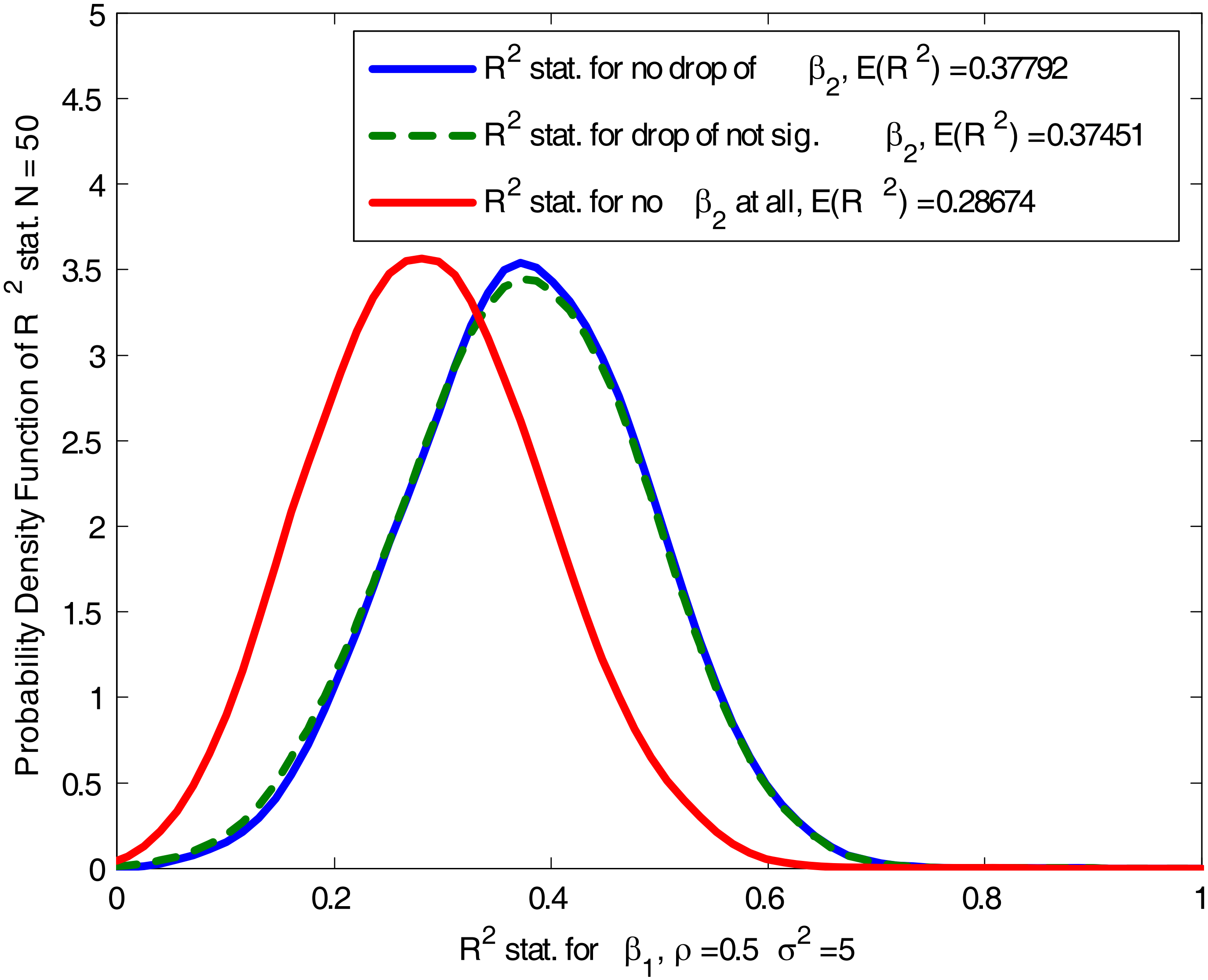
$N = 50, \quad \rho = 0.40; \quad \sigma^2 = 5$

R^2 density Function



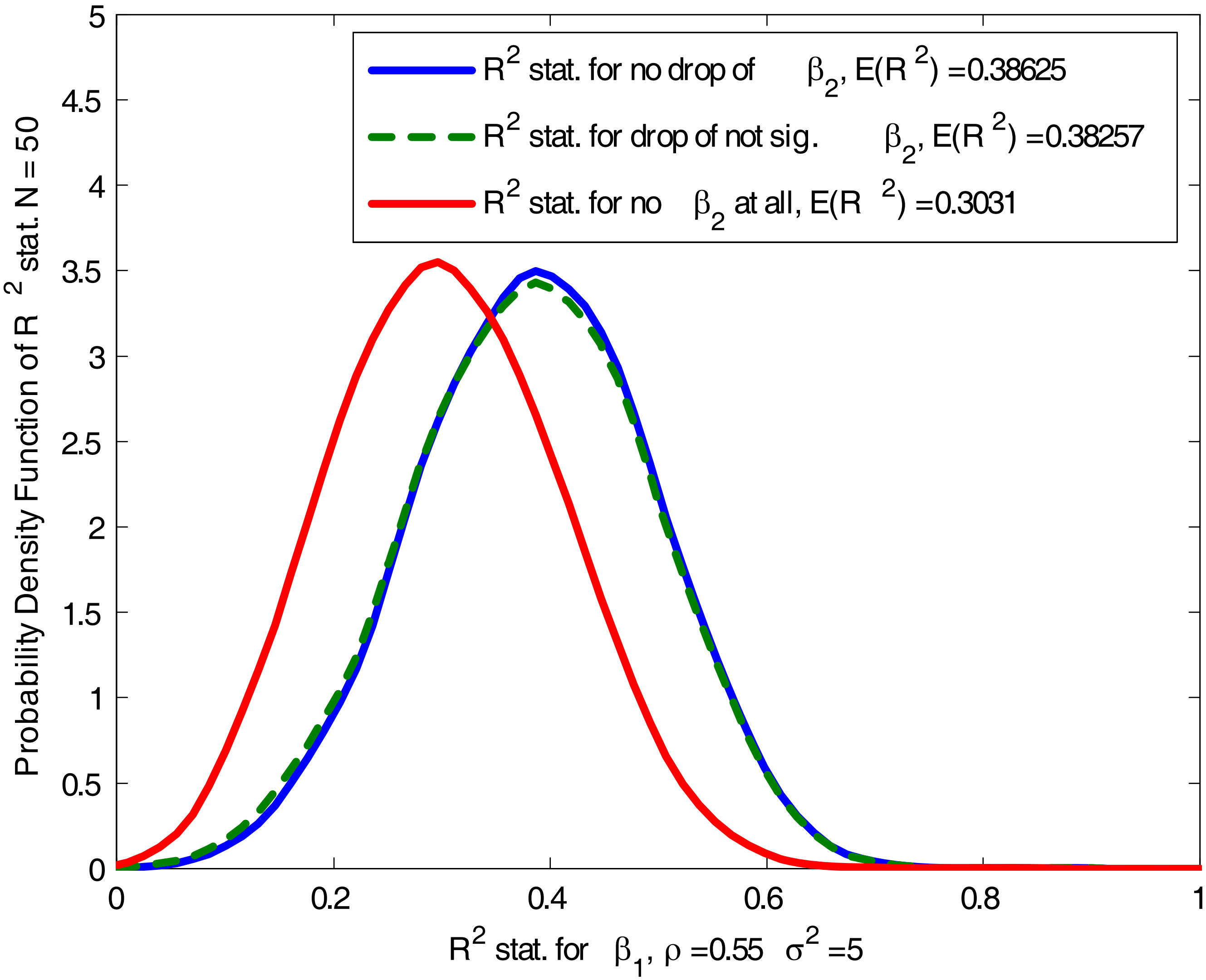
$$Distribution\ of\ R^2$$
$$N = 50, \quad \rho = 0.45; \quad \sigma^2 = 5$$

R^2 density Function



$$Distribution\ of\ R^2$$
$$N = 50, \quad \rho = 0.50; \quad \sigma^2 = 5$$

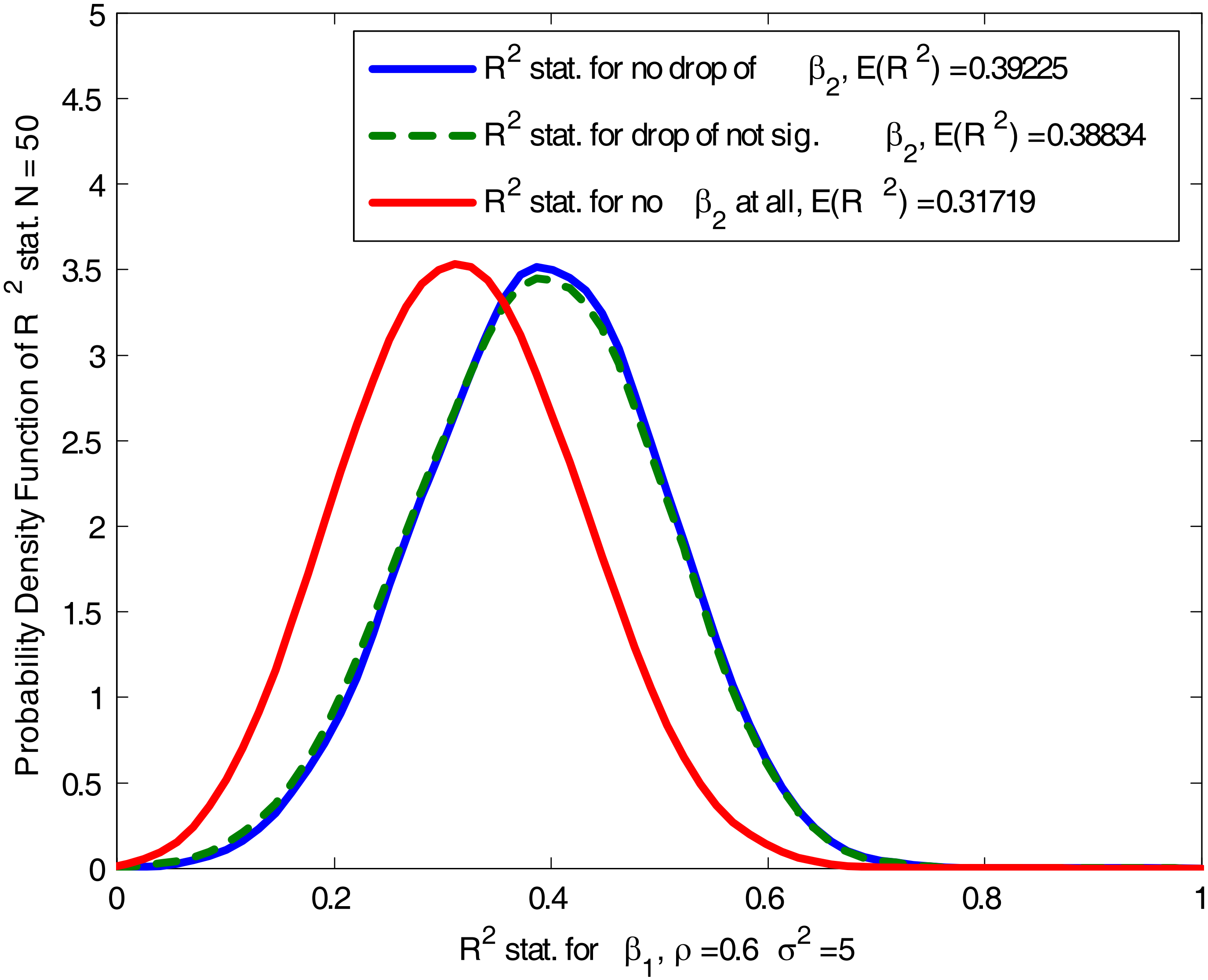
R^2 density Function



Distribution of R^2

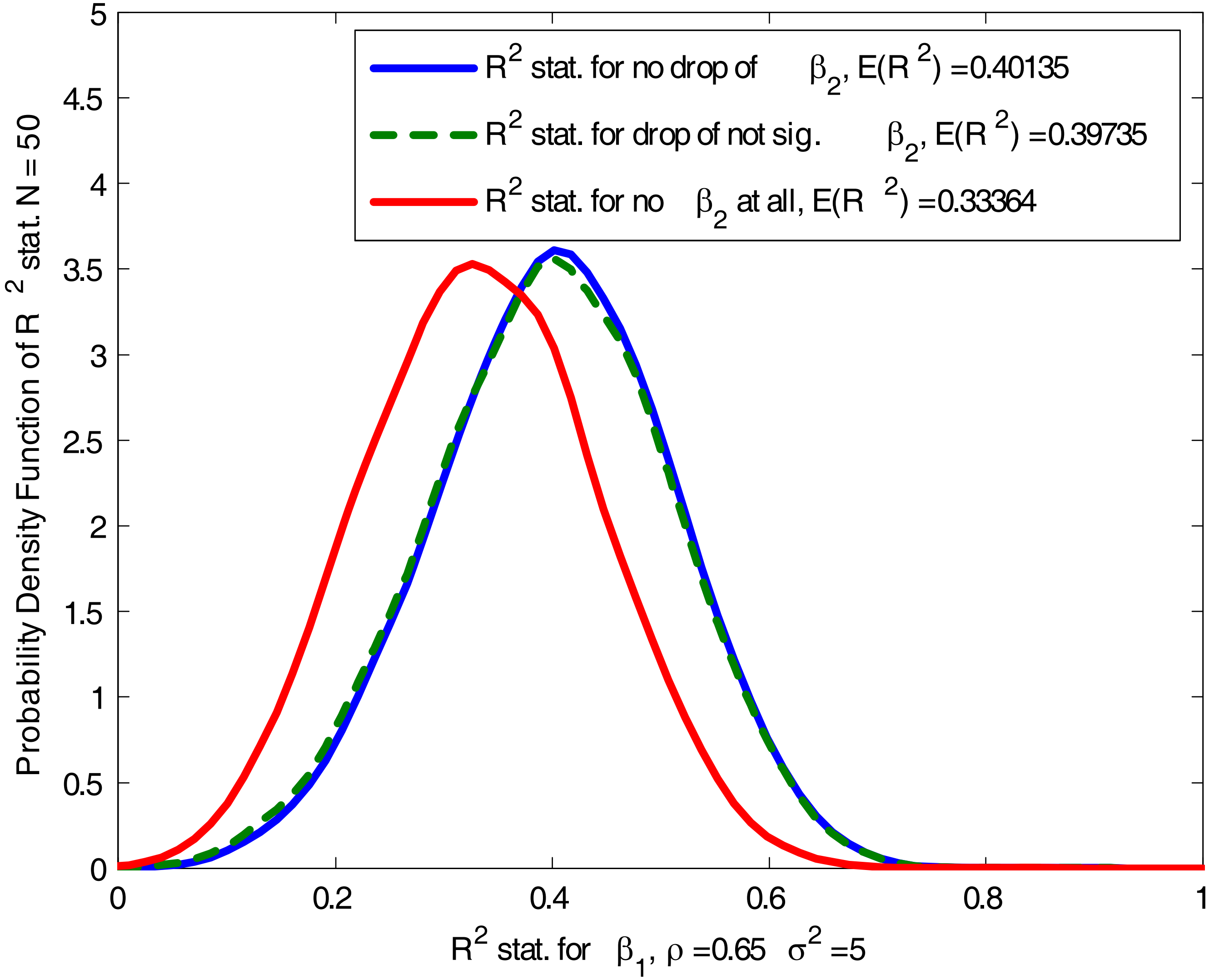
$N = 50, \quad \rho = 0.55; \quad \sigma^2 = 5$

R^2 density Function



$$Distribution\ of\ R^2$$
$$N = 50, \quad \rho = 0.60; \quad \sigma^2 = 5$$

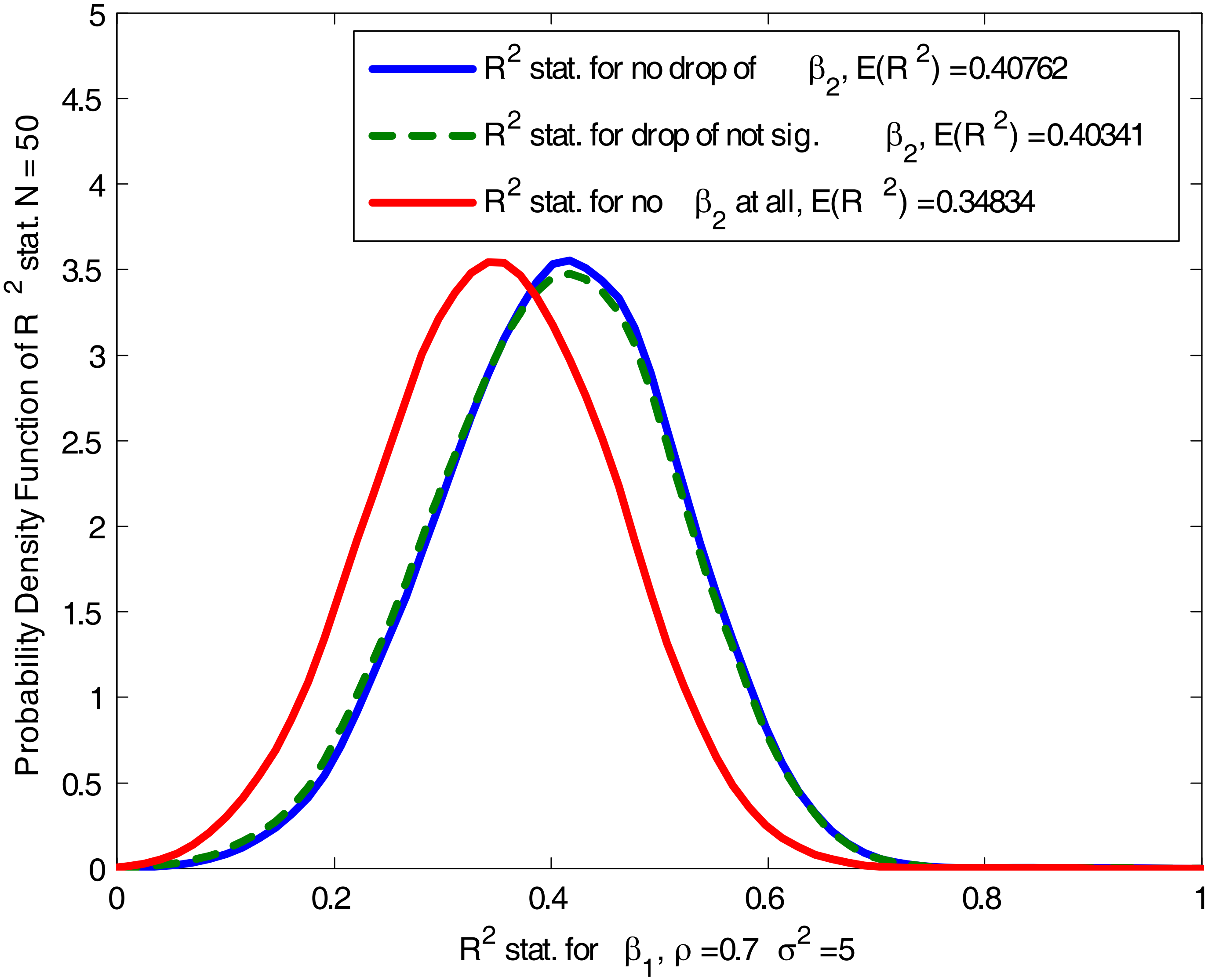
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.65; \quad \sigma^2 = 5$

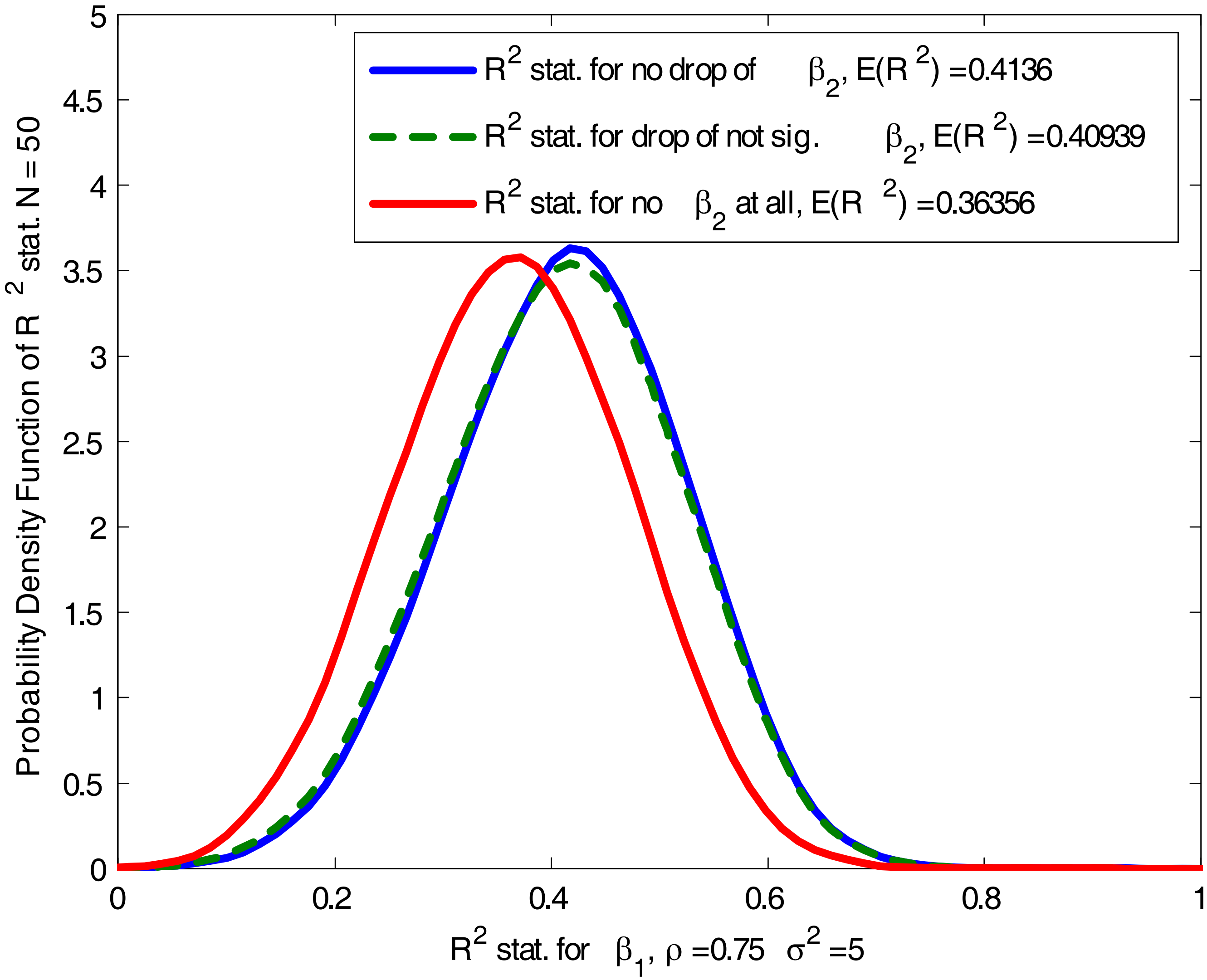
R^2 density Function



Distribution of R^2

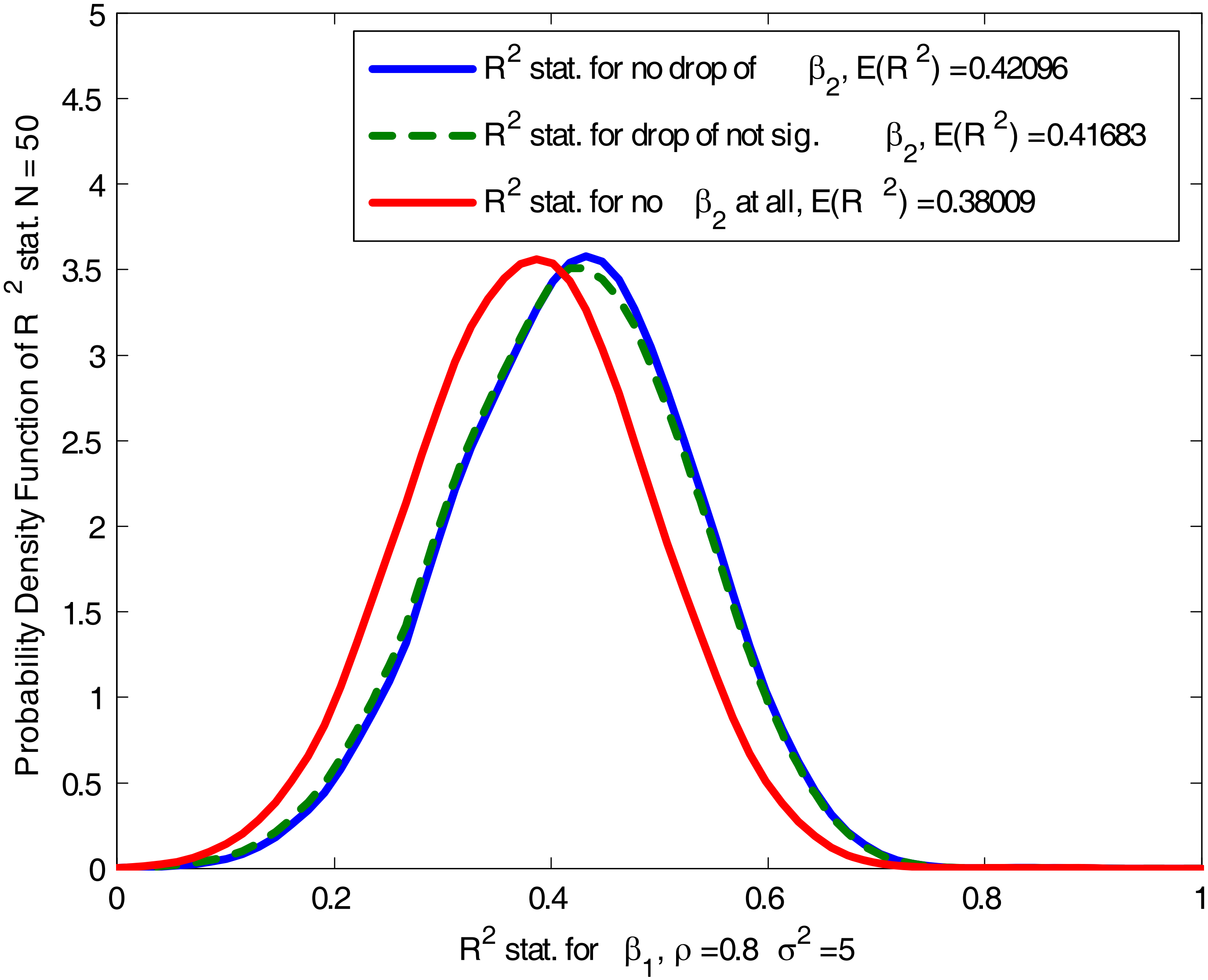
$N = 50, \quad \rho = 0.70; \quad \sigma^2 = 5$

R^2 density Function



$$Distribution\ of\ R^2$$
$$N = 50, \quad \rho = 0.75; \quad \sigma^2 = 5$$

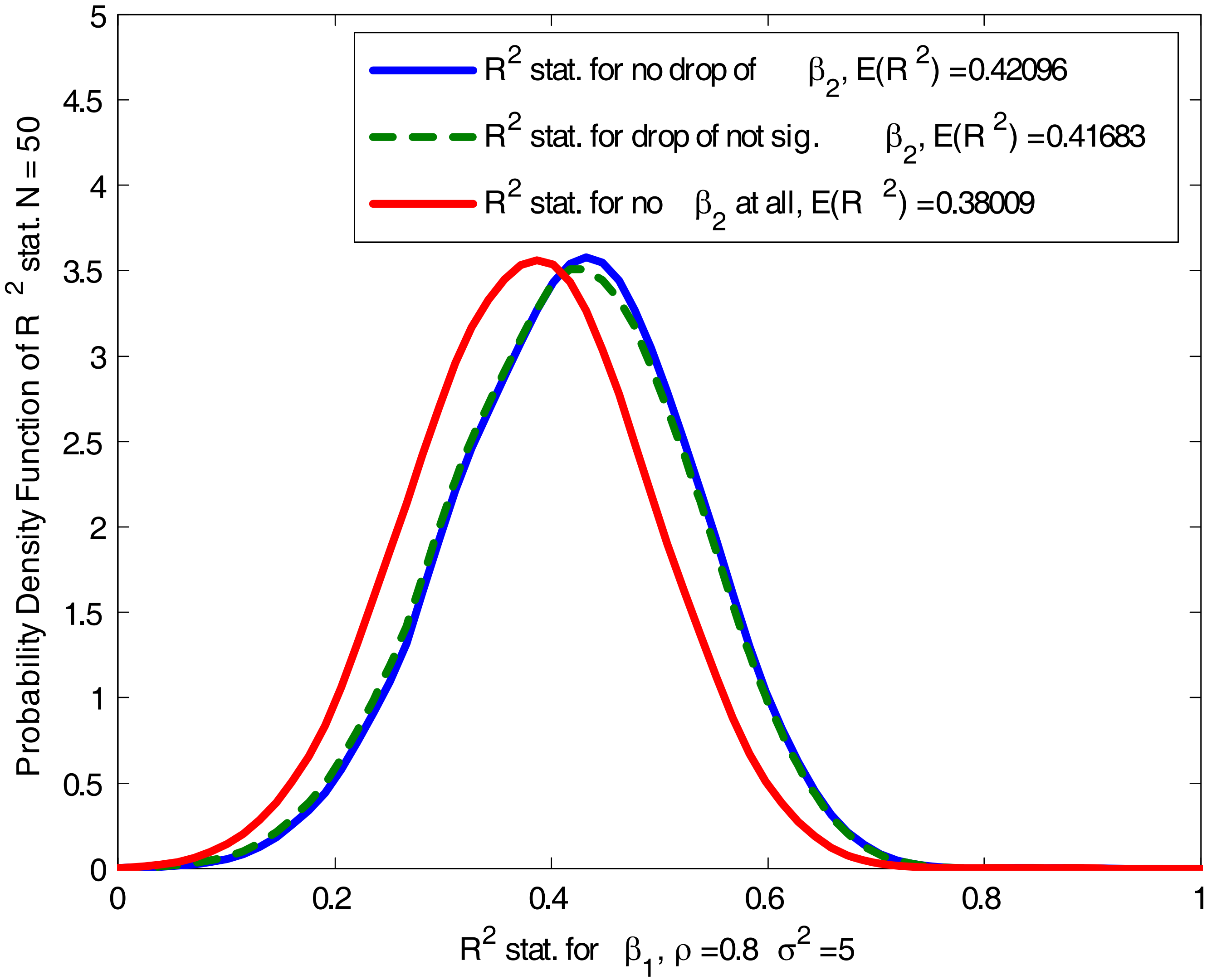
R^2 density Function



Distribution of R^2

$N = 50, \quad \rho = 0.80; \quad \sigma^2 = 5$

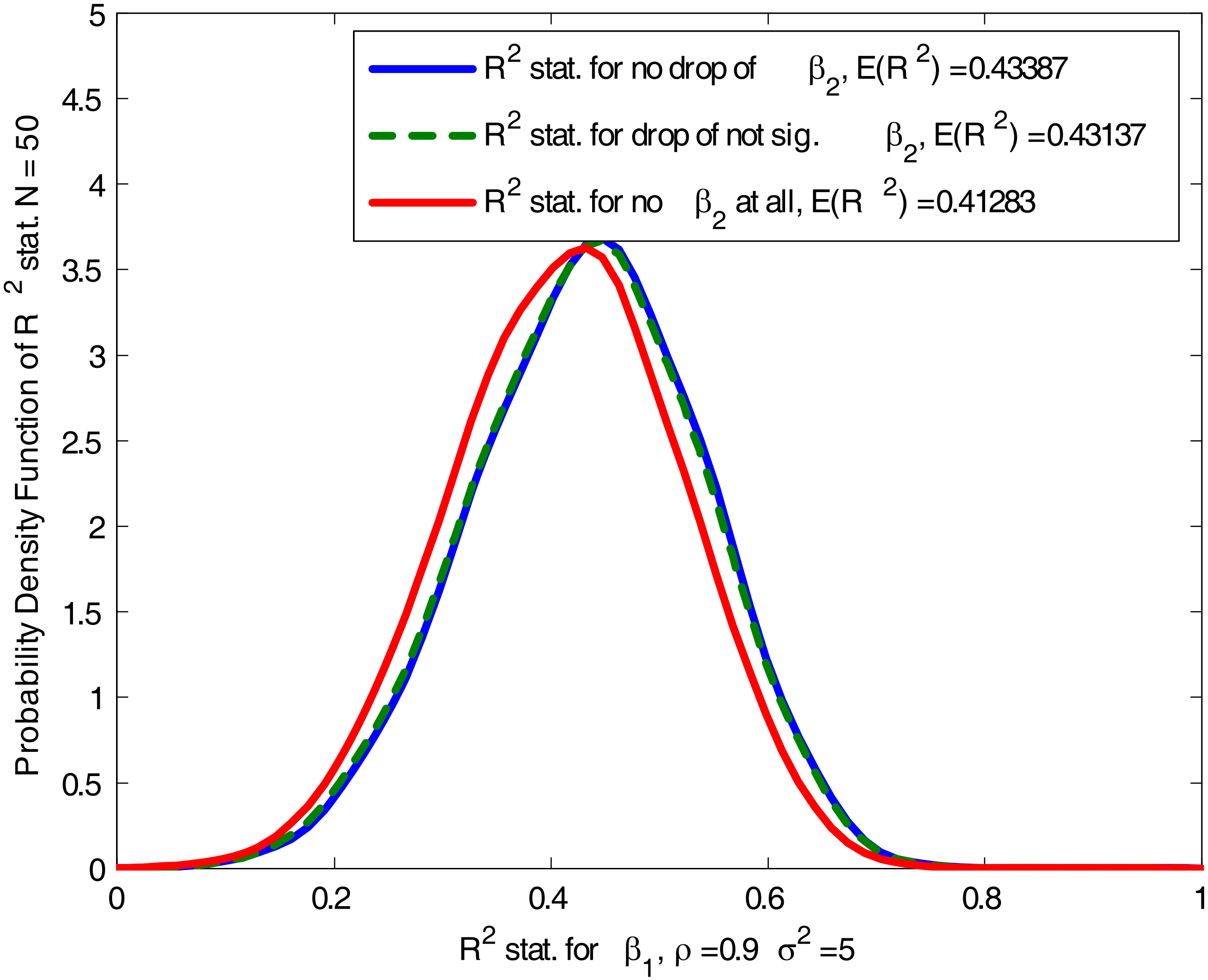
R^2 density Function



Distribution of R^2

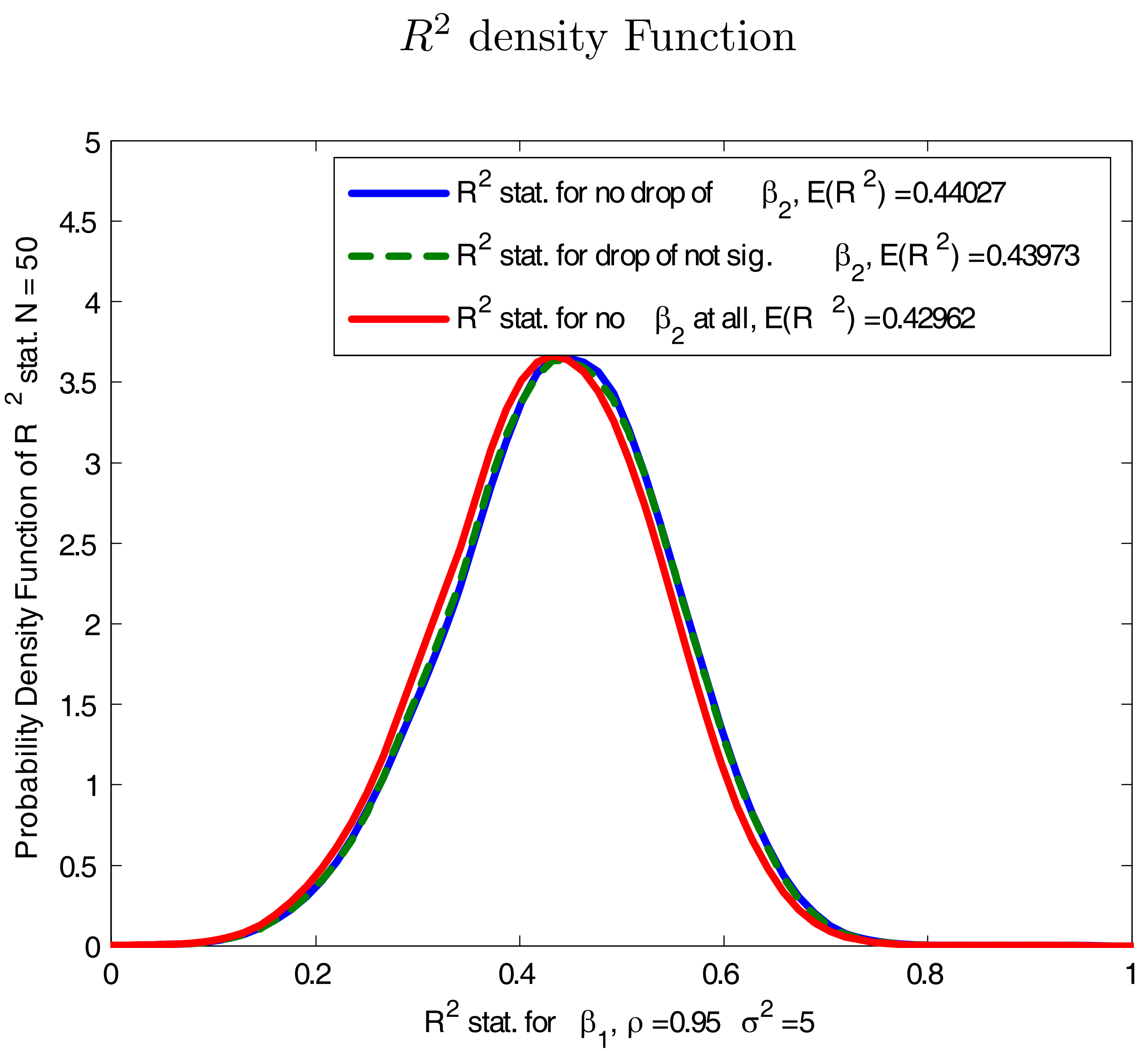
$N = 50, \quad \rho = 0.85; \quad \sigma^2 = 5$

R^2 density Function



$Distribution\ of\ R^2$

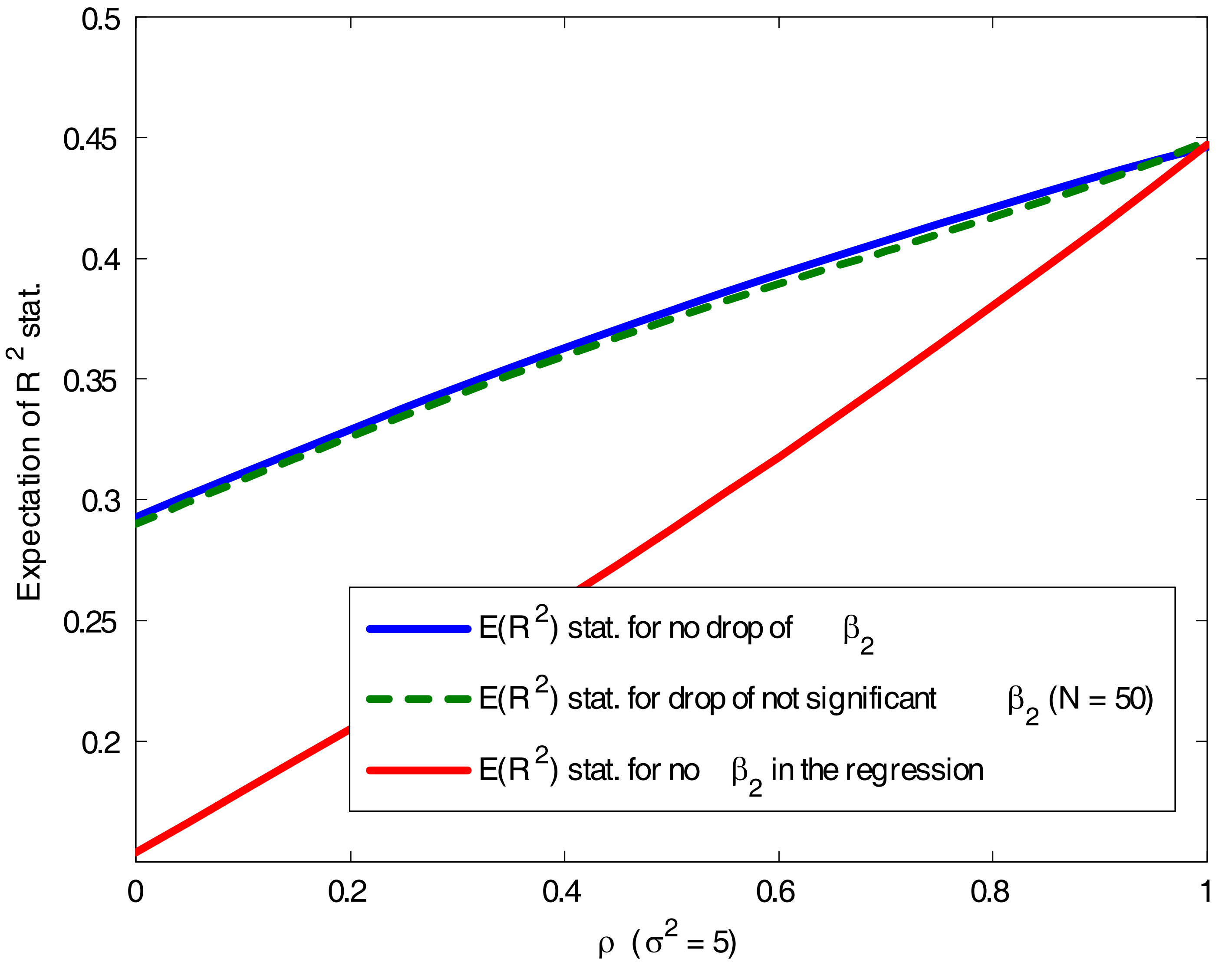
$$N = 50, \quad \rho = 0.90; \quad \sigma^2 = 5$$



Distribution of R^2

$N = 50, \quad \rho = 0.95; \quad \sigma^2 = 5$

R^2 Expectation



Expectation of R^2

$N = 50, \quad \rho \in [0, 1]\,; \quad \sigma^2 = 5$

Observe that for the case of $\rho = 1$, we obtain the model:

$$Y_i = X_i \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \varepsilon_i; \; i = 1, ..., N; \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.};$$

$$\begin{aligned} Y_i &= X_{1,i}\beta_1 + X_{2,i}\beta_2 + \varepsilon_i \\ &= X_{1,i}\beta_1 + (X_{1,i}\xi)\,\beta_2 + \varepsilon_i \\ Y_i &= X_{1,i}\,(\beta_1 + \xi) + \varepsilon_i \end{aligned}$$

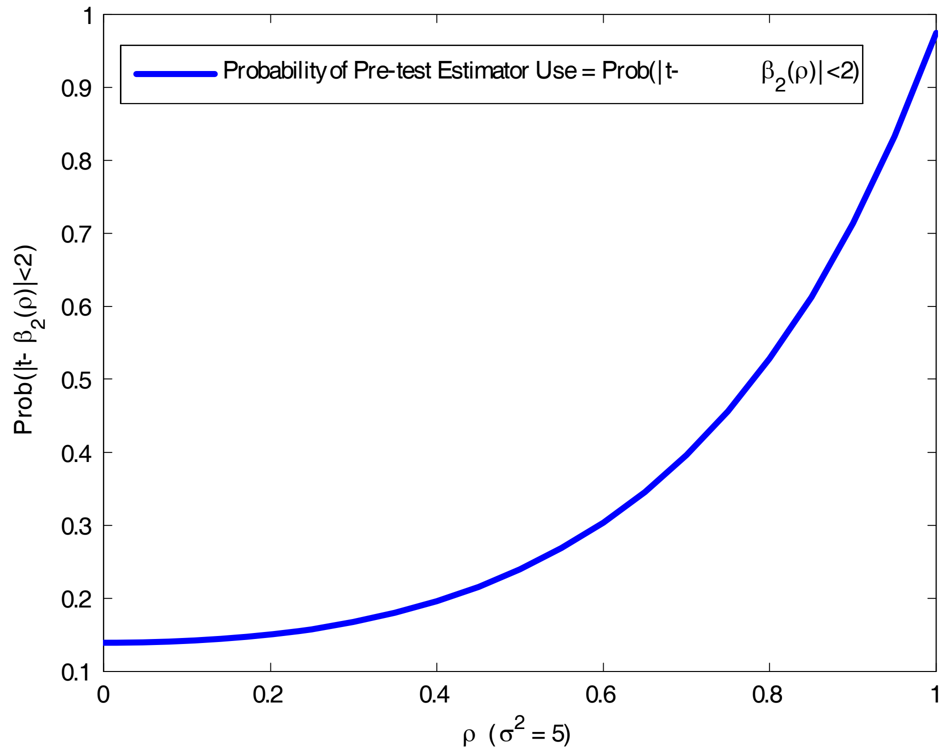
No loss of information. R^2 remains the same, but the estimator is highly biased.

8 Conclusion

Suppose the variables have a reasonable correlation.

$\Rightarrow$ There is a high Probability to Procedure the second Estimation:

Probability of Use of the Pre-test Estimator $\Pr(|t|_{\widehat{\beta}_2} < 2)$



Probability of Non Rejection of $H_0 : \widehat{\beta}_2 = 0$

$$\Pr(|t|_{\widehat{\beta}_2} < 2)$$

$$N = 50, \quad \sigma^2 = 5; \rho \in [0,1]$$

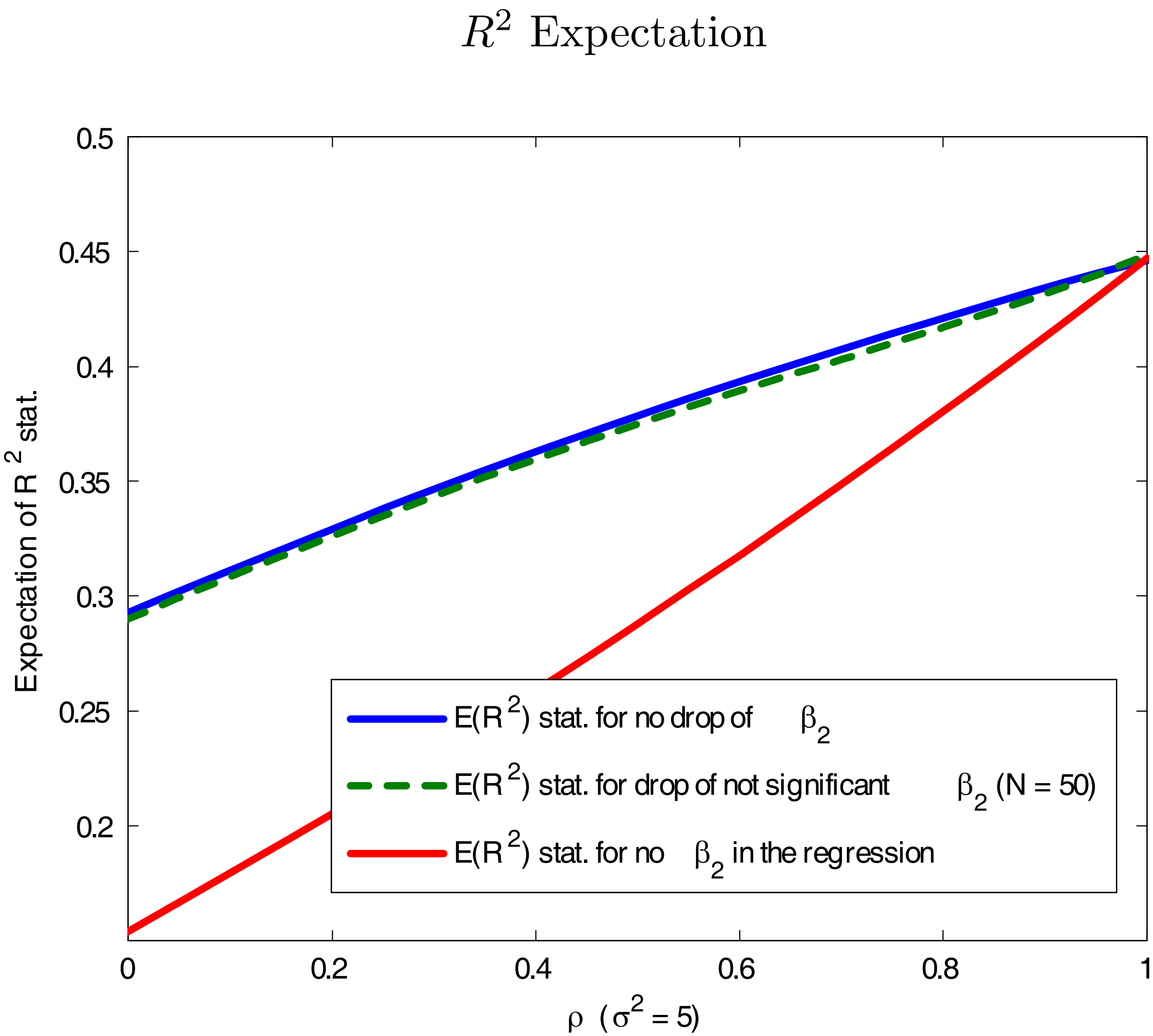
$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right)$$

$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot (1-\rho^2)}{N}\right)$$

$$(\text{Pre-test})\ \widetilde{\beta}_1 = \begin{cases} \widehat{\beta}_1 \text{ with Prob.} = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \\ \overline{\beta}_1 \text{ with Prob.} = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} < 2\right) \end{cases}$$

$$E\left(\widetilde{\beta}_1\right) = \Pr\left(|\widehat{t}|_{\widehat{\beta}_2} > 2\right) \cdot \beta_1 + \left[\Pr\left(|\widehat{t}|_{\widehat{\beta}_2} < 2\right)\right] \cdot (\beta_1 + \rho\beta_2)$$

R^2 won't change much:

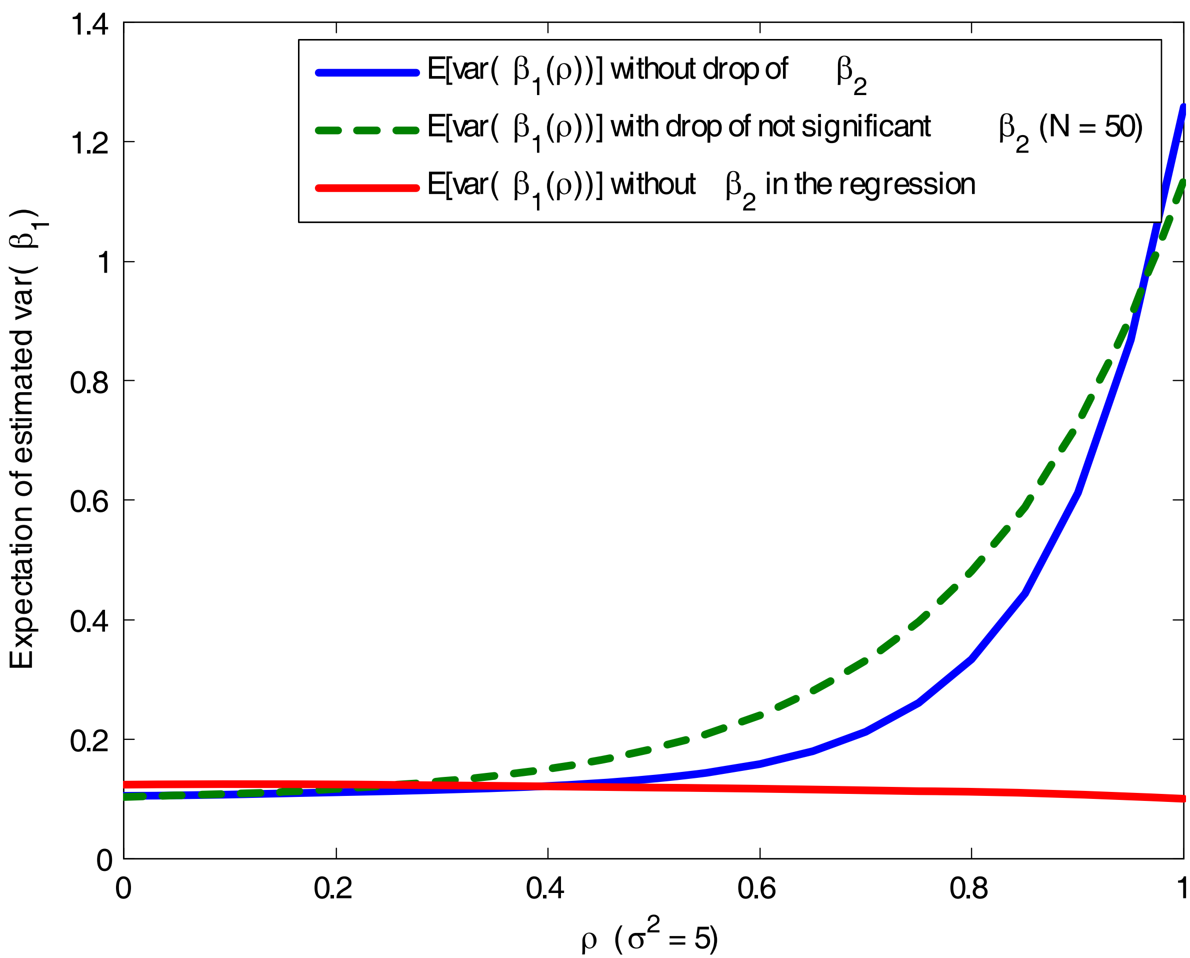


Expectation of R^2

$N = 50, \quad \rho \in [0, 1]; \quad \sigma^2 = 5$

The Variance Might Decrease:

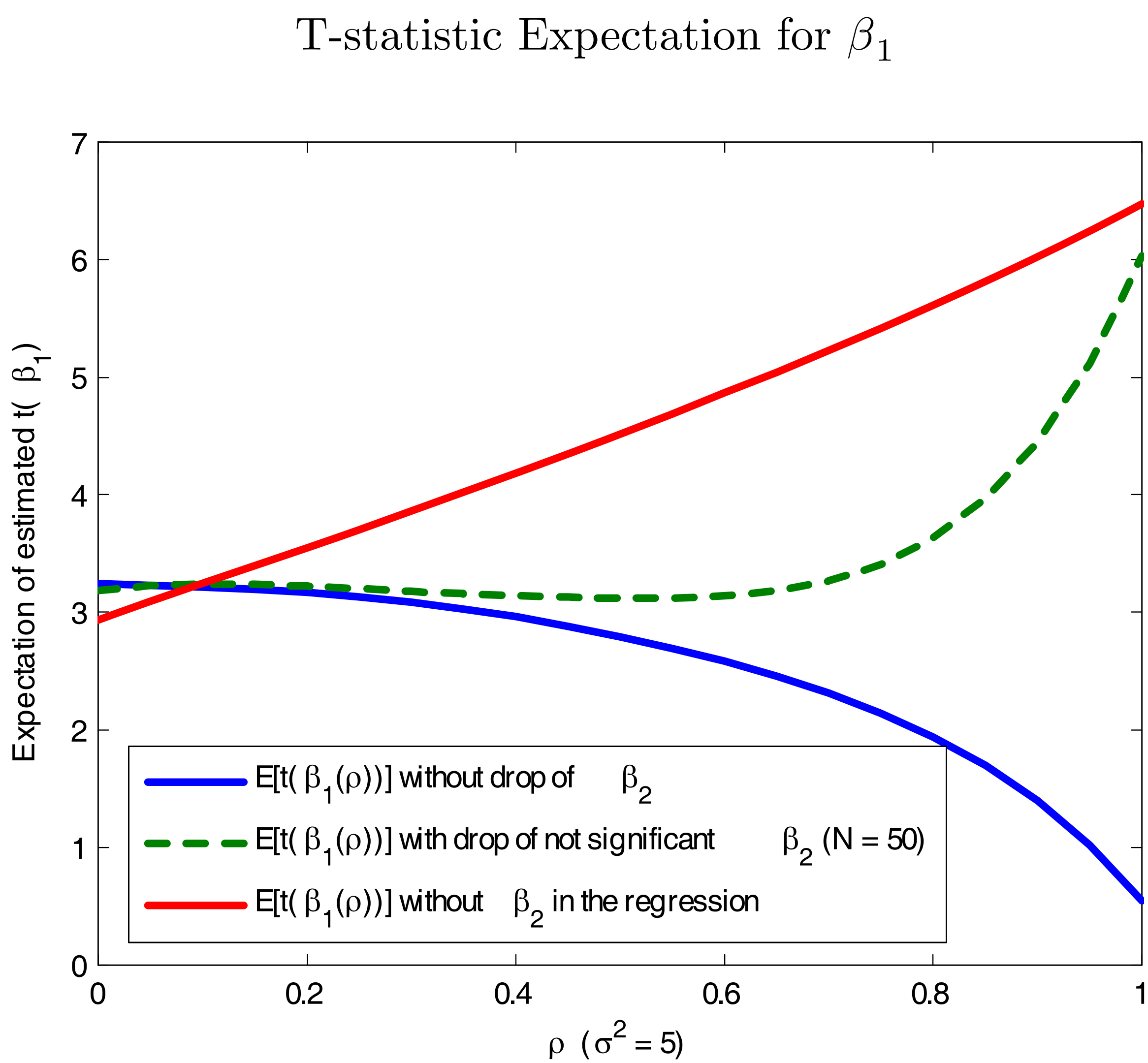
Expectation of The Variance of β_1



Expectation of R^2

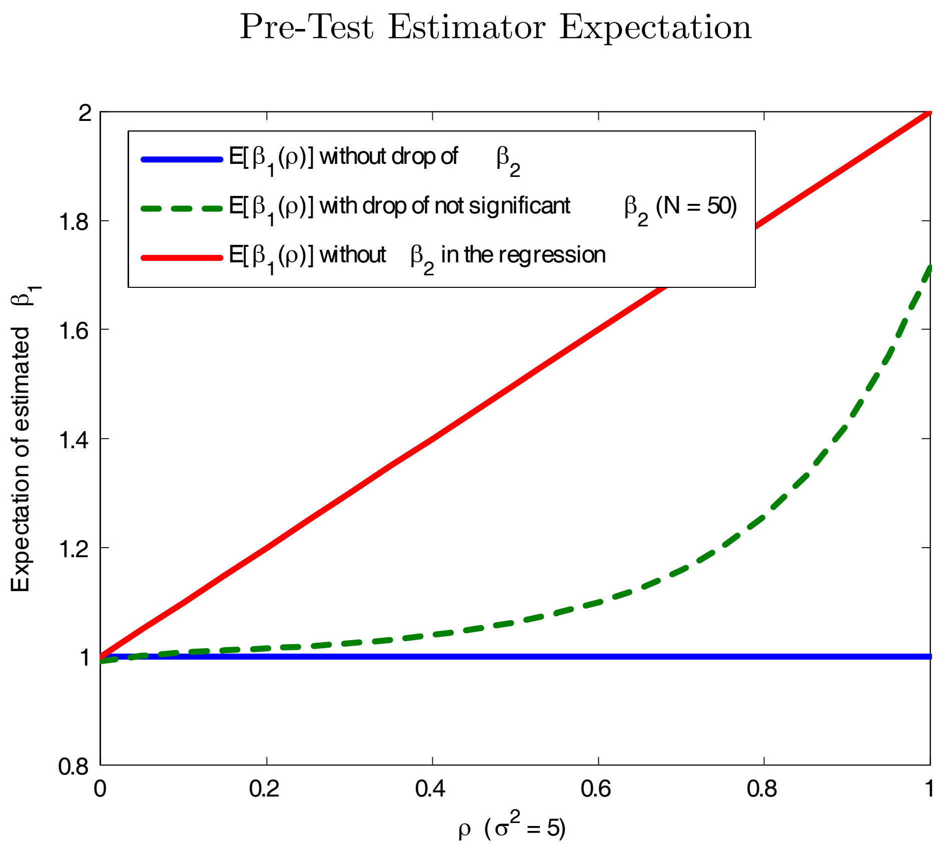
$N = 50, \quad \rho \in [0,1]; \quad \sigma^2 = 5$

The Estimator $\overline{\beta}_1$ is very significant:



$$Expectation\ of\ T-statistics\ \ for\ \beta_1$$
$$N = 50,\quad \sigma^2 = 5; \rho \in [0, 1]$$

And Extreme Biased:



$$\textit{Expectation of } \beta_1$$

$$N = 50, \quad \sigma^2 = 5; \rho \in [0, 1]$$

$$\widehat{\beta}_1 \approx N\left(\beta_1, \frac{\sigma^2}{N} \cdot \frac{1}{1-\rho^2}\right)$$

$$\overline{\beta}_1 \approx N\left(\beta_1 + \rho\beta_2, \frac{\sigma^2}{N} + \frac{\beta_2^2 \cdot \left(1-\rho^2\right)}{N}\right)$$

$$\text{(Pre-test) } \widetilde{\beta}_1 = \begin{cases} \widehat{\beta}_1 \text{ with Prob.= } \Pr\left(\left|\widehat{t}\right|_{\widehat{\beta}_2} > 2\right) \\ \overline{\beta}_1 \text{ with Prob.= } 1 - \Pr\left(\left|\widehat{t}\right|_{\widehat{\beta}_2} > 2\right) \end{cases}$$

$$E\left(\widetilde{\beta}_1\right) = \Pr\left(\left|\widehat{t}\right|_{\widehat{\beta}_2} > 2\right) \cdot \beta_1 + \left[1 - \Pr\left(\left|\widehat{t}\right|_{\widehat{\beta}_2} > 2\right)\right] \cdot (\beta_1 + \rho\beta_2)$$