# CHAPTER 5
# SIMULATION STUDY

The main objective of this study is to derive a similarity measure, in particular $R_p^2$, to compare two digital images. The MULFR model

$$\boldsymbol{Y}_i = \boldsymbol{\alpha} + \beta \boldsymbol{X}_i,\ i = 1, 2, \ldots, n \tag{5.1}$$

$$\left.\begin{aligned} \boldsymbol{x}_i &= \boldsymbol{X}_i + \boldsymbol{\delta}_i \\ \boldsymbol{y}_i &= \boldsymbol{Y}_i + \boldsymbol{\varepsilon}_i \end{aligned}\right\} i = 1, 2, \ldots, n. \tag{5.2}$$

where $\boldsymbol{\delta}_i \sim IND\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ and $\boldsymbol{\varepsilon}_i \sim IND\left(\mathbf{0}, \tau^2\mathbf{I}\right)$ maybe be applied if the parameters $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_p\right]$, $\beta$, $p$, $n$, $\boldsymbol{X}$, $\lambda$, $\sigma$, together with specific error distribution are known. Since it is impossible to study the performance of $R_p^2$ by considering all possible 'real' images, a simulation study is initiated by firstly selecting frequently used images such as Lena, Airplane and Peppers images (USC-SIPI Image Database) from which observed $p$-summary statistics are calculated and labeled as $\boldsymbol{x} = \left[x_{1i}, x_{2i}, \ldots, x_{pi}\right]'$. The corresponding unobserved $p$-summary statistics are labeled as $\boldsymbol{X} = \left[X_{1i}, X_{2i}, \ldots, X_{pi}\right]'$ .

## 5.1 Selection of Parameters Values

Selection of parameters values should reflect real image applications.

### 5.1.1 Selection of $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_p\right]$

Since $R_p^2$ is invariant to translation of the data (see Section 4.4.4), the vector $\boldsymbol{\alpha}$ will be fixed at the value $\boldsymbol{\alpha} = \left[0, \ldots, 0\right]$.

### 5.1.2 Selection of $\beta$

The parameter $\beta$ should reflect the relationship with the 'original' image $\boldsymbol{X}$ and the 'transformed' image $\boldsymbol{Y}$. For example if $\beta = 1.0$, this corresponds to the case when $\boldsymbol{Y} = \boldsymbol{X}$, implying identical images even after the transformation stage. In general, $\beta$ is not equal to one. The choice of $\beta$ value was derived from subjecting four frequently studied images in an image compression problem, and the estimated value is ranged from 1.0001 to 38.5711. It is noted that if a different image problem was studied, a different set of $\beta$ values should be considered.

### 5.1.3 Selection of $p$ and $n$

The feature vectors $\boldsymbol{X} = \left[X_{1i}, X_{2i}, \ldots, X_{pi}\right]'$ and $\boldsymbol{Y} = \left[Y_{1i}, Y_{2i}, \ldots, Y_{pi}\right]'$ have the same summary statistics representing the images. For example, in the compression problem, $X_{1i}$ and $Y_{1i}$ are the image luminance, while $X_{2i}$ and $Y_{2i}$ may refer to image contrast. In general, it is sufficient to consider $p \leq 5$ since most of the imaging applications used only one image attribute such as luminance or contrast as their input, while MSSIM utilized two image attributes. For colour image, three-dimensional feature vector may be used.

The choice of $n$ is simply to investigate the sampling properties of the estimated parameters $\hat{\beta}$ and $R_p^2$. Sample sizes used in the simulation are $n$ = 10, 50, 100, 250, 1000 and 4000. These values are selected to represent the approximate number of data calculated from an image based on $8 \times 8$ window size. In JPEG compression problem, a standard test image can be dissected into sub-images of size of $8 \times 8$, to allow for suitable value of sample sizes ($n$).

### 5.1.4 Selection of $\sigma$ and $\lambda$

Various values of variance of the random error $\boldsymbol{\varepsilon}$ are also considered. In image applications, the larger value of error variance causes the higher level of noise in an image. Note that an error term with zero mean and 10 standard deviations can yield a completely distorted image. Therefore this simulation study will only consider standard deviation, $\sigma = 5,10$ to represent different degrees of image noise.

It was assumed that the MULFR model has the ratio of error variances equals to one. This assumption may not be true in some imaging applications. The choice of $\lambda$ value was derived from a pilot experiment in an image compression problem, and the estimated value is ranged from 1.0062 to 91.1108.

### 5.1.5 Selection of Distributions for $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\delta}_i$

Both $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\delta}_i$ are initially choose to be multivariate normal as defined in the MULFR model. It is of interest to see if violation of the normality assumption will affect the estimation of parameters, in particular $R_p^2$.

A total of 27 distributions of $\boldsymbol{\varepsilon}_i$ are being considered in this study involving one normal distribution and 26 non-normal distributions. The varying degrees of non-normality are achieved by introducing increasing levels of skewness and kurtosis into the generated data. Non-normality exists in many image processing applications since the transformed image is usually subjected to noise which has a non-normal distribution. For example, the Laplacian noise generated by JPEG compression process follows a Laplace distribution or Double exponential distribution (Cheng & Cheng, 2009), the shot noise has a Poisson distribution and Salt & Pepper noise caused by dead pixels, analog-to-digital converter and transmission processes has flat-tail error distribution. It is common in imaging applications to assume that the 'original' image is noise-free or this original image is subjected to Gaussian white noise.

## 5.2 Monte Carlo Simulation

### 5.2.1 Monte Carlo Simulation Procedure

Program subroutines (generate_error, Rm2normal) are written using MATLAB 7.1 environment to generate the simulation data needed for this study. For a particular choice of $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_p\right]$, $\beta$, $p$, $n$, $\lambda$, $\sigma$ and specified error distribution, the simulation processes can be explained by the following procedure.

Step 1: Select a set of frequently studied images which is represented by $\boldsymbol{X} = \left[X_{1i}, X_{2i}, \ldots, X_{pi}\right]'$, $i = 1, 2, \ldots, n$ of size $p \times n$.

Step 2: Select the parameters values of $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_p\right]$ and $\beta$. Obtain the $\boldsymbol{Y} = \left[Y_{1i}, Y_{2i}, \ldots, Y_{pi}\right]'$ by using the equation $\boldsymbol{Y} = \boldsymbol{\alpha} + \beta \boldsymbol{X}$ with fixed $\lambda$ and $\sigma$ values.

Step 3: Generate two sets of $p \times 1$ random errors, $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ using standard normal generator RANDN. This random errors can be transformed to a pre-defined distribution with mean zero and variance $\sigma^2$ using Equation (5.9) in the next section to obtain the *p*-dimensional random errors $\boldsymbol{\delta} = \left[\delta_{1i}, \delta_{2i}, \ldots, \delta_{pi}\right]'$ and $\boldsymbol{\varepsilon} = \left[\varepsilon_{1i}, \varepsilon_{2i}, \ldots, \varepsilon_{pi}\right]'$ respectively.

Step 4: Add the random errors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ into $\boldsymbol{Y}$ and $\boldsymbol{X}$ correspondingly to obtain the observed random values $\boldsymbol{y}$ and $\boldsymbol{x}$, respectively.

Step 5: Repeat Step 3 and Step 4 to derive $\left(\boldsymbol{x}_i, \boldsymbol{y}_i\right), i = 1, 2, \ldots, n$ using Equation (4.2).

Step 6: Calculate $\hat{\beta}$, followed by $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{X}}$ and $\hat{\sigma}^2$ using Result 3 in Section 4.1.2. Finally calculate $\hat{\lambda}$ as defined by constraint (iv) in Section 4.1.2 and

$R_p^2$ using Result 6 in Section 4.3. The confidence interval for $R_p^2$ is calculated by using Result 7, Section 4.4.6.

Step 7: Repeat Steps 3 to 6 to obtain $N$ replicates of the statistics in Step 5. $(N = 10000)$.

Step 8: Compute the sample statistics $\overline{R}_p^2$, $\overline{\hat{\beta}}$, $b\hat{i}as$, $\hat{S}D$, $M\hat{S}E$ and $\overline{\hat{\lambda}}$ as defined in Equations (5.3) to (5.8) over $N$ data sets of the estimators.

Step 9: Repeat for different choice of $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_p\right]$, $\beta$, $p$, $n$, $\lambda$, $\sigma$ and selected distribution of $\boldsymbol{\varepsilon}$.

The sample statistics such as sample mean, bias, standard deviation and mean square error for $\hat{\beta}$ (similarly for $R_p^2$) computed in Step 8 are used to measure the effects of changing parameter values and non-normality on the estimation of $\beta$ and $R_p^2$. These sample statistics are given below;

$$\overline{\hat{\beta}} = \frac{1}{N}\sum_{s=1}^{N}\hat{\beta}_s, \quad s = 1, 2, \ldots, S \text{ where } S \text{ is the simulation size} \tag{5.3}$$

$$b\hat{i}as = \overline{\hat{\beta}} - \beta \tag{5.4}$$

$$\hat{S}D = \sqrt{\frac{1}{(S-1)}\sum_{s=1}^{S}\left(\hat{\beta}_s - \overline{\hat{\beta}}\right)^2} \tag{5.5}$$

$$M\hat{S}E = \hat{S}D^2 + b\hat{i}as^2 \tag{5.6}$$

$$\overline{\hat{\lambda}} = \frac{1}{N}\sum_{s=1}^{N}\hat{\lambda}_s \text{ where } \hat{\lambda}_s \text{ is the ratio of error variances for the } s\text{-iteration} \tag{5.7}$$

$$\overline{R}_p^2 = \frac{1}{N}\sum_{s=1}^{N}R_{Ps}^2 \tag{5.8}$$

### 5.2.2 Transformation to a Non-Normal Distribution

There are many ways to perform a one-to-one transformation of a normal distributed variable to a unimodal non-normal distribution. Power transformation such as Box & Cox (1964) is a standard reference and several others are discussed in Johnson et al. (1994). A simple quadratic method used by Pooi (2003) and Ng (2006) has been chosen in this study because it is capable of generating a wide range of unimodal distributions which are skewed or having thin waist with known skewness and kurtosis values. Let $e$ be a random error with the standard normal distribution and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$ a vector of constant values. Consider the following non-linear function of $e$:

$$\varepsilon = \begin{cases} \gamma_1 e + \gamma_2 \left( e^2 - \dfrac{1+\gamma_3}{2} \right) & , e \geq 0 \\ \gamma_1 e + \gamma_2 \left( \gamma_3 e^2 - \dfrac{1+\gamma_3}{2} \right) & , e < 0 \end{cases} \tag{5.9}$$

in which the constants $\gamma_t$ ; $t = 1,2,3$ are such that for small value $q > 0$, $\varepsilon$ is a one-to-one function of $e$ when $|e| > Z_q$ when $\text{Prob}\{Z > Z_q\} = \dfrac{q}{2}$ where $Z$ has the standard normal distribution.

The mean of $\varepsilon$ is always zero. The shape of the distribution of $\varepsilon$ will vary as the value of $\boldsymbol{\gamma}$ varies. If $\gamma_2 = 0$, then $\varepsilon$ has the normal distribution with mean zero and variance $\gamma_1^2$. As the standard normal distribution has a third and fourth moments equal to zero and three respectively, the severity of departure from normality may be measured by the sample third and fourth moments (Pooi, 2003; Ng, 2006) of $\varepsilon_i$ ; $i = 1, 2, \ldots, n$ given by

$$\bar{m}_3 = \frac{m_3}{m_2^{3/2}} \text{ and } \bar{m}_4 = \frac{m_4}{m_2^2}$$

where $m_2 = \dfrac{1}{n-1} \sum_{i=1}^{n} \varepsilon_i^2$ , $m_3 = \dfrac{1}{n\left(1 - 3/n\right)} \sum_{i=1}^{n} \varepsilon_i^3$ and $m_4 = \dfrac{1}{n\left(1 - 4/n\right)} \sum_{i=1}^{n} \varepsilon_i^4$ .

In practice the error terms are non-normal, for example in JPEG compression the errors have a Laplacian distribution. Henceforth, the generated values of $\bar{m}_3$ and $\bar{m}_4$ will correspond to a particular Laplacian error term as determine by $\gamma = (0.015819, 0.568667, -1.000010)$.

The following Figures 5.1 to 5.3 show some examples of probability density function for various non-normal distributions transformed by Equation 5.9. Figure 5.1 for example, indicates the plots for a normal distribution with $(\bar{m}_3, \bar{m}_4) = (0.0, 3.0)$ derived by using $\gamma = (1, 0, -1)$, a Laplacian distribution with $(\bar{m}_3, \bar{m}_4) = (0.0, 12.0)$ derived by using $\gamma = (0.015819, 0.568667, -1.000010)$ and a symmetry distribution with $(\bar{m}_3, \bar{m}_4) = (0.0, 2.2)$ derived by using $\gamma = (1.273796, -0.174687, -0.999993)$. Figure 5.2 compares the normal distribution and two right-skewed distributions with $(\bar{m}_3, \bar{m}_4) = (1.5, 6.7)$ and $(\bar{m}_3, \bar{m}_4) = (3.0, 16.0)$. Whereas, Figure 5.3 compares the normal distribution and two left-skewed distributions with $(\bar{m}_3, \bar{m}_4) = (-1.5, 7.5)$ and $(\bar{m}_3, \bar{m}_4) = (-3.0, 16.7)$.

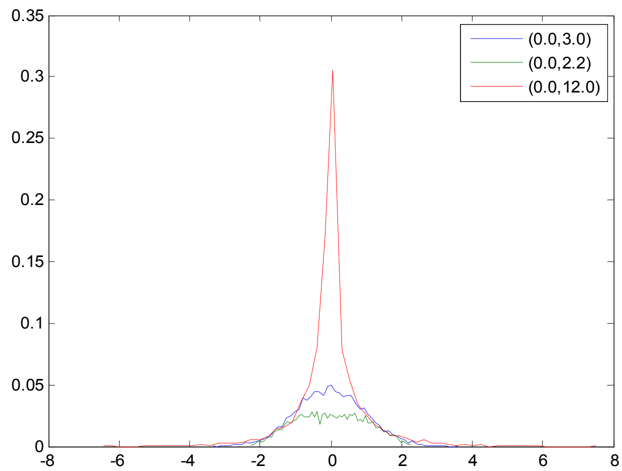



Figure 5.1: Probability density function of $\varepsilon$ with $(\bar{m}_3, \bar{m}_4) = (0.0, 3.0), (0.0, 2.2), (0.0, 12.0)$

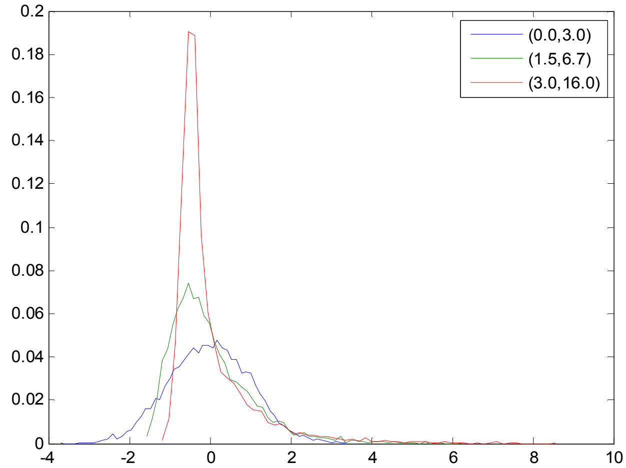



Figure 5.2: Probability density function of $\varepsilon$ with $(\bar{m}_3, \bar{m}_4) = (0.0, 3.0), (1.5, 6.7), (3.0, 16.0)$

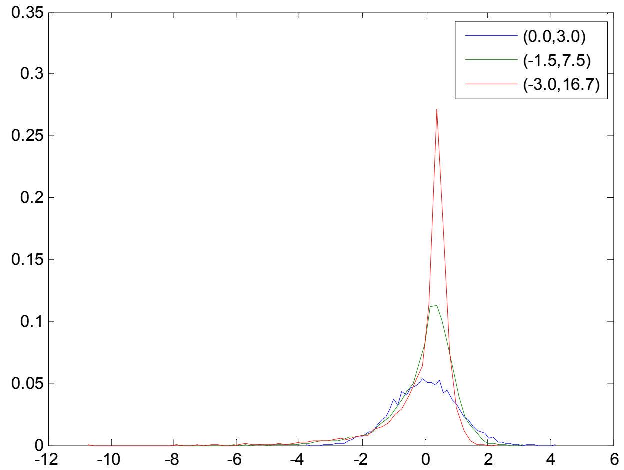



Figure 5.3: Probability density function of $\varepsilon$ with $(\bar{m}_3, \bar{m}_4) = (0.0, 3.0), (-1.5, 7.5), (-3.0, 16.7)$

## 5.3 Simulation A: Performance of the MULFR Model When the Basic Assumptions are Satisfied

The objective of this subsection is to investigate the effects of changing values of $\beta$, $n$, $p$ and $\sigma$ on the performance of MULFR model when their ratio of error variances, $\lambda = 1$ and both errors $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are normally distributed. In Simulation A, some parameters values are fixed where $\boldsymbol{\alpha} = [0, \ldots, 0]$, $p = 1, 2, 5$ and $n = 10, 50, 100, 250, 1000, 4000$. Other parameters values are changed in three cases: (i)

$\sigma = 1, \beta = 1$ , (ii) $\sigma = 1$ and $\beta = 1.5, 10, 40$ , (iii) $\beta = 10$ and $\sigma = 5, 10$ which should covered a broad range of image type.

### 5.3.1 Both $\delta$ and $\varepsilon$ are Normally Distributed, $\lambda = 1$ and $\sigma = 1$, $\beta = 1$

Under the given set of condition $(\lambda = 1, \sigma = 1, \beta = 1)$, Table 5.1 clearly show that when all the model assumptions are satisfied, the parameters $\beta$ and $R_P^2$ are well estimated, in the sense that their average values equal to the true values and also because of minimal mean square errors and a negligible length for confidence intervals.

The parameters are also well estimated for modest sample size $(n = 10)$ and moderately high dimension $(p = 5)$.

Table 5.1: All model assumptions satisfied with $\lambda = 1, \sigma = 1, \beta = 1$

| | $n$ | $\overline{\lambda}$ | $\overline{\hat{\beta}}$ | $M\hat{S}E_{\hat{\beta}}$ | $\overline{R}_P^2$ | $\hat{S}D_{R_P^2}$ | $M\hat{S}E_{R_P^2}$ | Length $CI\left(\overline{R}_P^2\right)$ |
|---|---|---|---|---|---|---|---|---|
| $p$=1 | 10 | 1.0000 | 1.0000 | 4.80E-30 | 1.0000 | 1.66E-15 | 2.77E-30 | 0.0000 |
| | 50 | 1.0000 | 1.0000 | 2.55E-29 | 1.0000 | 1.97E-15 | 3.87E-30 | 0.0000 |
| | 100 | 1.0000 | 1.0000 | 5.04E-29 | 1.0000 | 1.84E-15 | 3.38E-30 | 0.0000 |
| | 250 | 1.0000 | 1.0000 | 1.10E-28 | 1.0000 | 2.14E-15 | 4.58E-30 | 0.0000 |
| | 1000 | 1.0000 | 1.0000 | 1.31E-28 | 1.0000 | 2.73E-15 | 7.46E-30 | 0.0000 |
| | 4000 | 1.0000 | 1.0000 | 1.49E-28 | 1.0000 | 4.64E-15 | 2.16E-29 | 0.0000 |
| $p$=2 | 10 | 1.0000 | 1.0000 | 9.40E-30 | 1.0000 | 3.21E-15 | 1.03E-29 | 0.0000 |
| | 50 | 1.0000 | 1.0000 | 8.30E-30 | 1.0000 | 1.89E-15 | 3.58E-30 | 0.0000 |
| | 100 | 1.0000 | 1.0000 | 4.14E-29 | 1.0000 | 2.65E-15 | 7.05E-30 | 0.0000 |
| | 250 | 1.0000 | 1.0000 | 5.01E-29 | 1.0000 | 2.42E-15 | 5.84E-30 | 0.0000 |
| | 1000 | 1.0000 | 1.0000 | 1.63E-22 | 1.0000 | 3.43E-15 | 1.18E-29 | 0.0000 |
| | 4000 | 1.0000 | 1.0000 | 8.37E-29 | 1.0000 | 6.43E-15 | 4.13E-29 | 0.0000 |
| $p$=5 | 10 | 1.0000 | 1.0000 | 2.60E-30 | 1.0000 | 2.04E-15 | 4.17E-30 | 0.0000 |
| | 50 | 1.0000 | 1.0000 | 7.36E-30 | 1.0000 | 2.54E-15 | 6.43E-30 | 0.0000 |
| | 100 | 1.0000 | 1.0000 | 2.12E-29 | 1.0000 | 3.12E-15 | 9.73E-30 | 0.0000 |
| | 250 | 1.0000 | 1.0000 | 2.35E-29 | 1.0000 | 2.98E-15 | 8.87E-30 | 0.0000 |
| | 1000 | 1.0000 | 1.0000 | 3.17E-24 | 1.0000 | 4.07E-15 | 1.65E-29 | 0.0000 |
| | 4000 | 1.0000 | 1.0000 | 1.69E-25 | 1.0000 | 6.88E-15 | 4.73E-29 | 0.0000 |

### 5.3.2 Repeating Section 5.3.1 to Compare $\beta$ = 1.5, 10, 40

Table 5.2 shows simulation results using the given set of condition$(\lambda = 1, \sigma = 1, \beta = 1.5)$. The parameters $\beta$ and $R_P^2$ are still well estimated with small mean square errors and narrow confidence intervals at modest sample size $(n = 10)$ and moderately high dimension$(p = 5)$.

Figure 5.4 and Figure 5.5 (see Appendix A1) summarize the mean square errors for $\hat{\beta}$ and length of confidence interval for $R_P^2$, respectively when the true $\beta = 1.5, 10, 40$. The parameters $\beta$ and $R_P^2$ are well estimated for the selected true $\beta$ values, dimensions and sample sizes.

Table 5.2: All model assumptions satisfied with $\lambda = 1, \sigma = 1, \beta = 1.5$

| | $n$ | $\overline{\lambda}$ | $\overline{\hat{\beta}}$ | $M\hat{S}E_{\hat{\beta}}$ | $\overline{R}_P^2$ | $\hat{S}D_{R_P^2}$ | $M\hat{S}E_{R_P^2}$ | Length $CI\left(\overline{R}_P^2\right)$ |
|---|---|---|---|---|---|---|---|---|
| $p$=1 | 10 | 1.0000 | 1.4996 | 3.50E-05 | 1.0000 | 1.91E-05 | 1.85E-09 | 1.034E-03 |
| | 50 | 1.0000 | 1.4995 | 6.80E-06 | 1.0000 | 8.86E-06 | 1.95E-09 | 5.619E-04 |
| | 100 | 1.0000 | 1.4995 | 3.35E-06 | 1.0000 | 5.97E-06 | 1.80E-09 | 4.008E-04 |
| | 250 | 1.0000 | 1.4996 | 1.33E-06 | 1.0000 | 3.45E-06 | 1.51E-09 | 2.505E-04 |
| | 1000 | 1.0000 | 1.4995 | 5.16E-07 | 1.0000 | 1.83E-06 | 1.65E-09 | 1.285E-04 |
| | 4000 | 1.0000 | 1.4996 | 2.60E-07 | 1.0000 | 8.73E-07 | 1.51E-09 | 6.405E-05 |
| $p$=2 | 10 | 1.0000 | 1.4995 | 1.99E-05 | 1.0000 | 2.38E-05 | 3.03E-09 | 1.627E-03 |
| | 50 | 1.0000 | 1.4995 | 3.69E-06 | 1.0000 | 8.54E-06 | 1.84E-09 | 8.022E-04 |
| | 100 | 1.0000 | 1.4995 | 1.69E-06 | 1.0000 | 5.84E-06 | 1.72E-09 | 5.517E-04 |
| | 250 | 1.0000 | 1.4996 | 6.90E-07 | 1.0000 | 3.33E-06 | 1.40E-09 | 3.549E-04 |
| | 1000 | 1.0000 | 1.4996 | 3.15E-07 | 1.0000 | 1.66E-06 | 1.42E-09 | 1.781E-04 |
| | 4000 | 1.0000 | 1.4996 | 2.24E-07 | 1.0000 | 8.51E-07 | 1.46E-09 | 8.929E-05 |
| $p$=5 | 10 | 1.0000 | 1.4997 | 3.58E-06 | 1.0000 | 1.41E-05 | 1.07E-09 | 1.442E-03 |
| | 50 | 1.0000 | 1.4995 | 1.26E-06 | 1.0000 | 8.60E-06 | 1.87E-09 | 7.769E-04 |
| | 100 | 1.0000 | 1.4996 | 7.30E-07 | 1.0000 | 5.36E-06 | 1.46E-09 | 5.506E-04 |
| | 250 | 1.0000 | 1.4995 | 4.72E-07 | 1.0000 | 3.58E-06 | 1.62E-09 | 3.529E-04 |
| | 1000 | 1.0000 | 1.4996 | 2.30E-07 | 1.0000 | 1.68E-06 | 1.40E-09 | 1.783E-04 |
| | 4000 | 1.0000 | 1.4996 | 1.97E-07 | 1.0000 | 8.42E-07 | 1.43E-09 | 8.899E-05 |

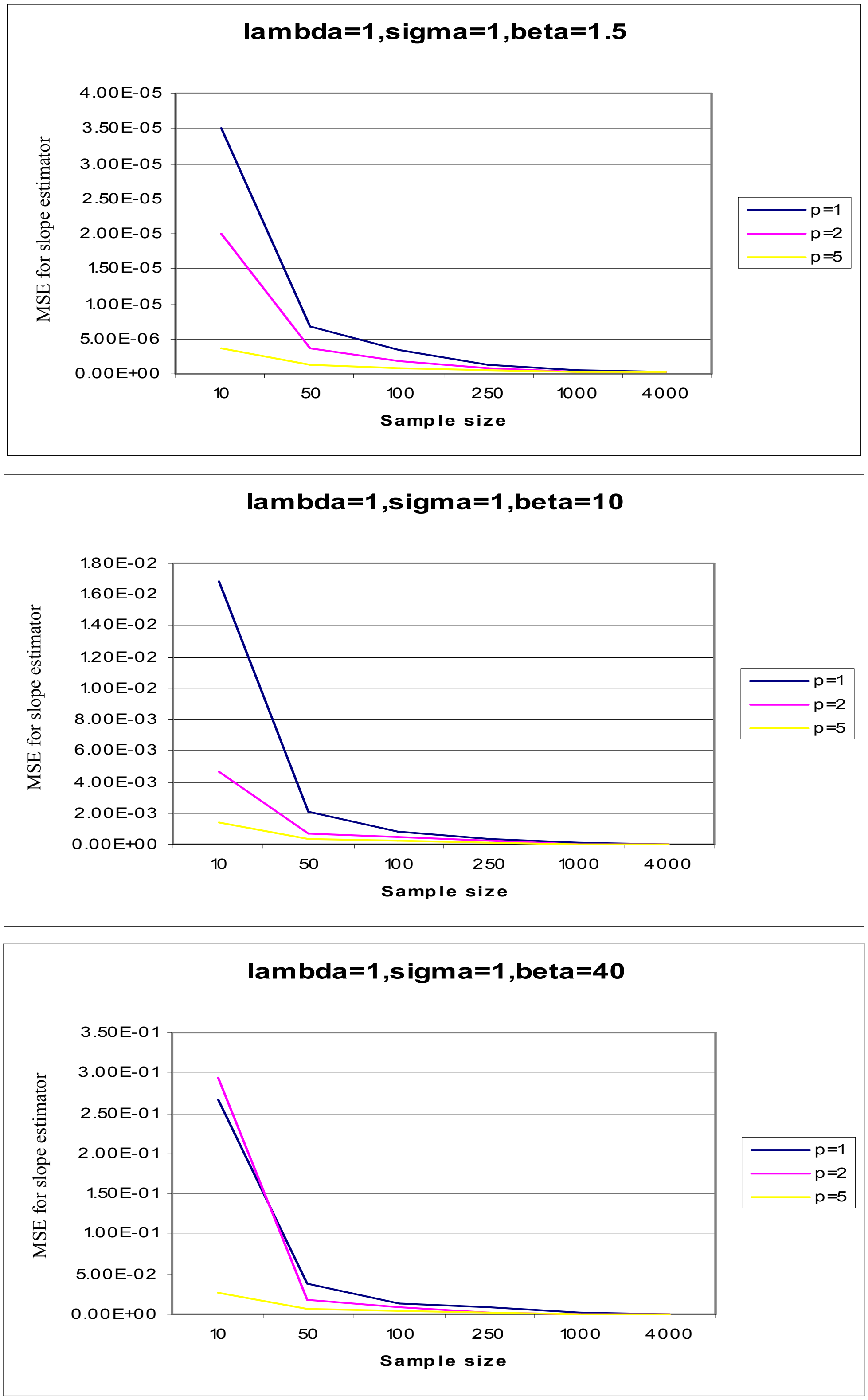



Figure 5.4: Mean square error for $\hat{\beta}$ when $\lambda = 1, \sigma = 1$ and $\beta = 1.5, 10, 40$

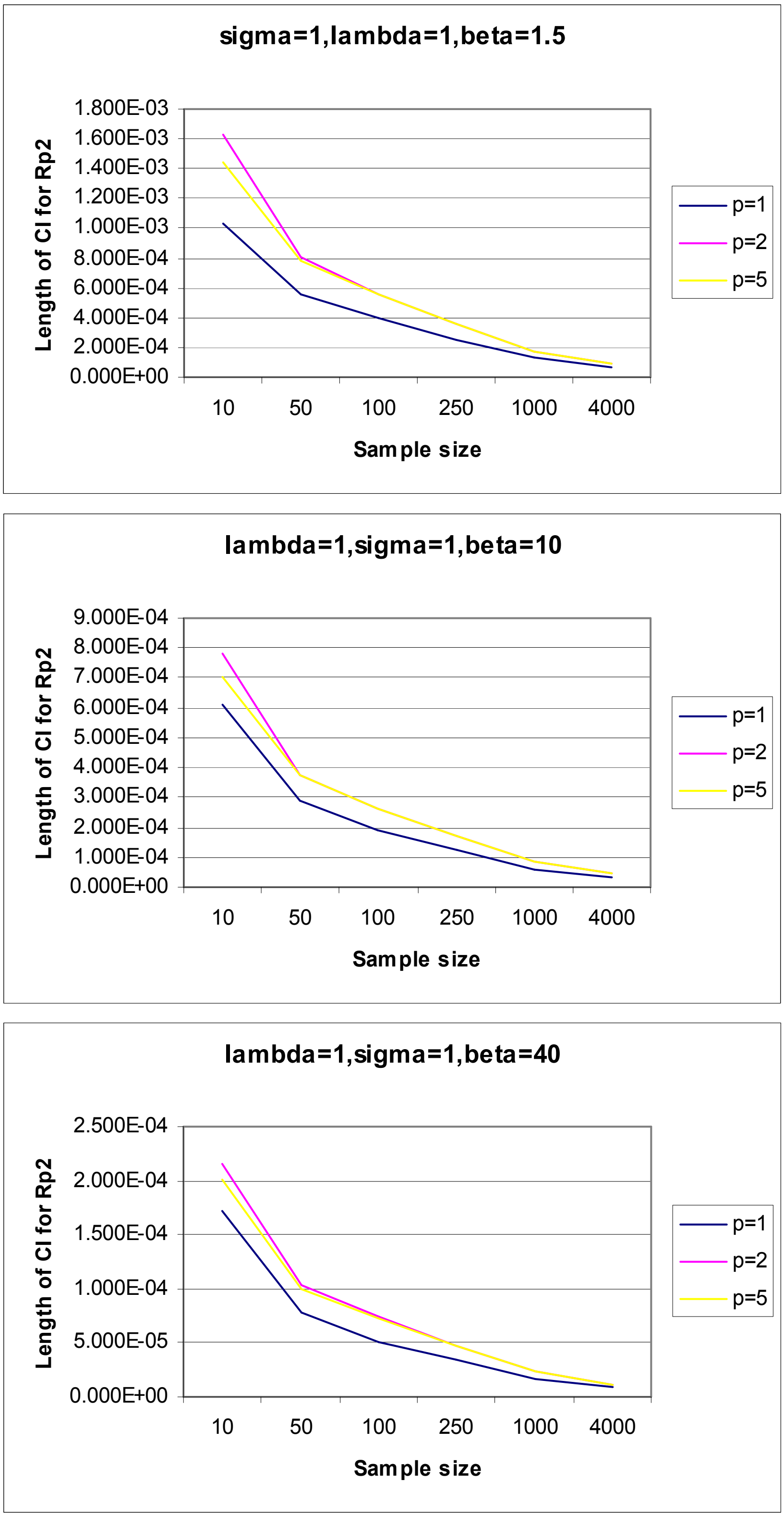
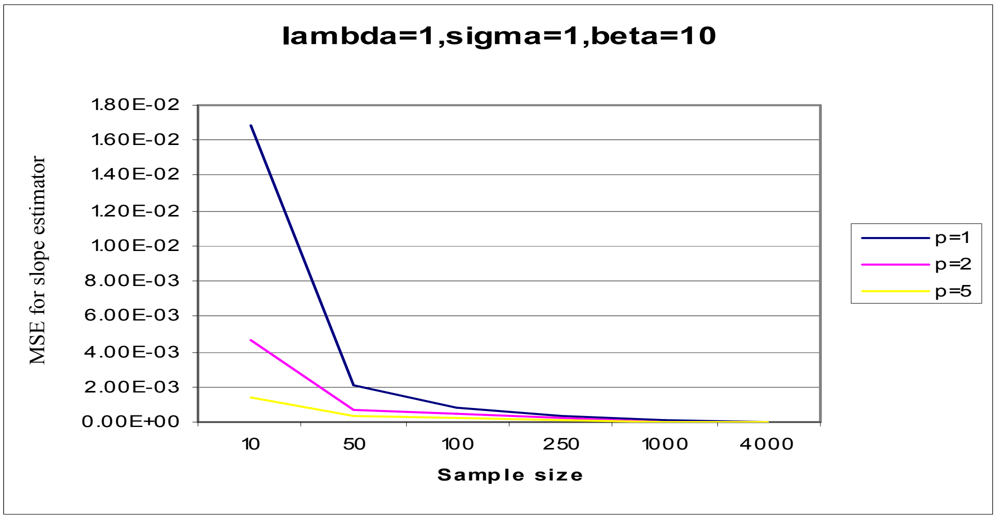
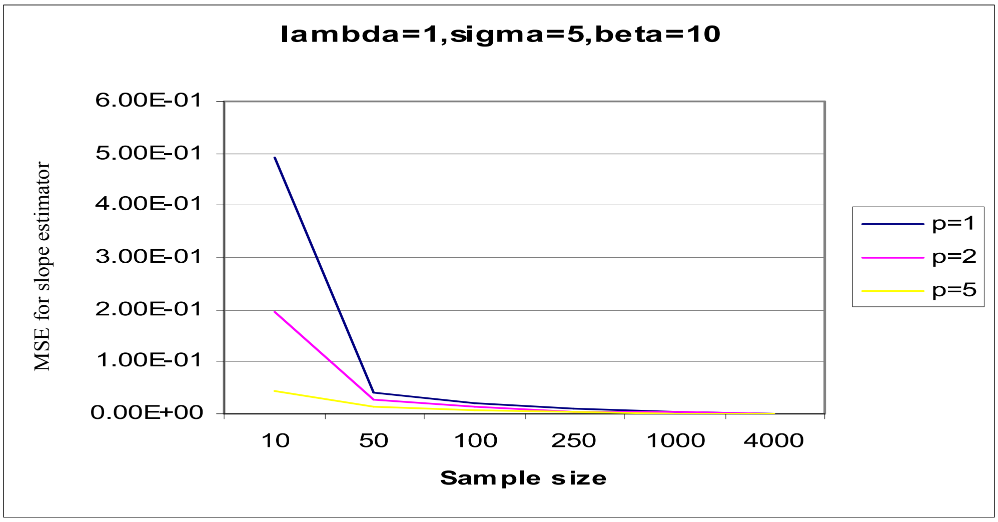



Figure 5.5: Length of confidence interval for $R_P^2$ when $\lambda=1, \sigma=1$ and $\beta=1.5, 10, 40$

### 5.3.3 Both $\delta$ and $\varepsilon$ are Normally Distributed, $\lambda = 1$, $\beta = 10$ and $\sigma = 1,5,10$

Figure 5.6 and Figure 5.7 (see Appendix A2) summarize the mean square errors for $\hat{\beta}$ and length of confidence interval for $R_P^2$, respectively when the true error standard deviation $\sigma = 1,5,10$. The parameters $\beta$ and $R_P^2$ are well estimated for the selected true $\beta$ values, dimensions and sample sizes.



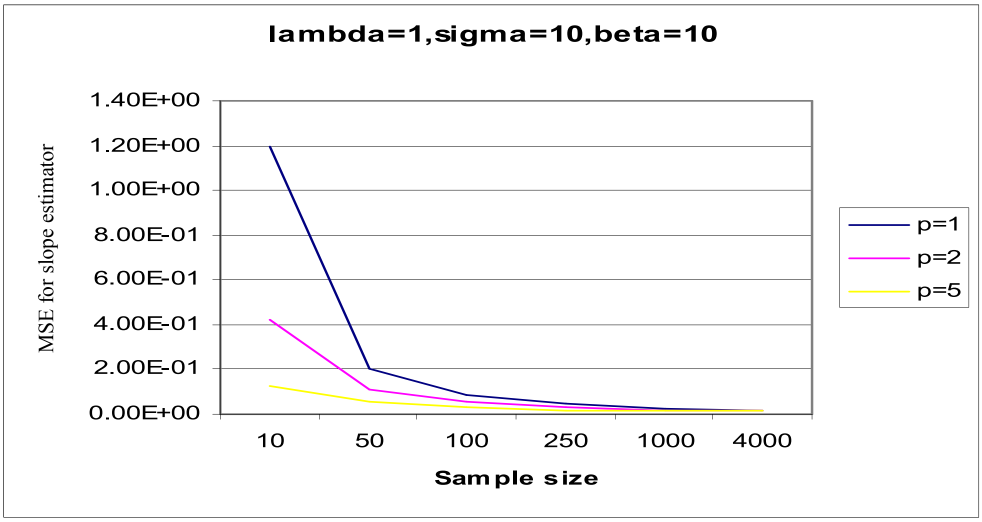
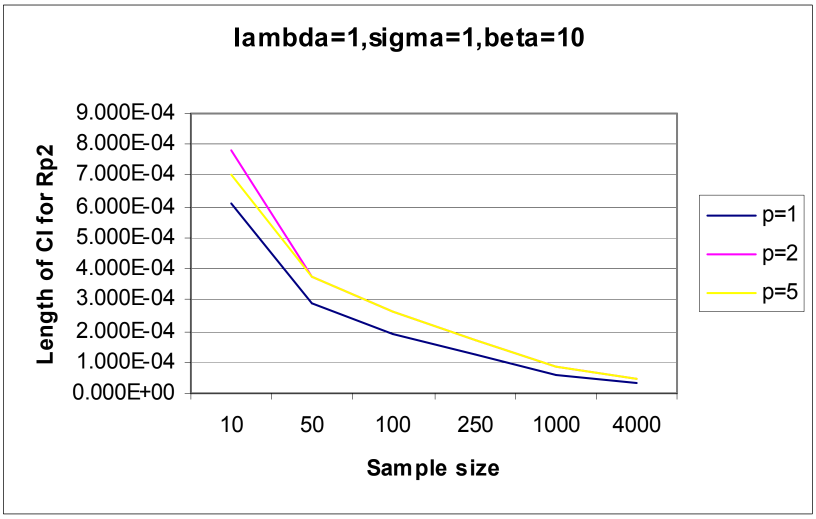
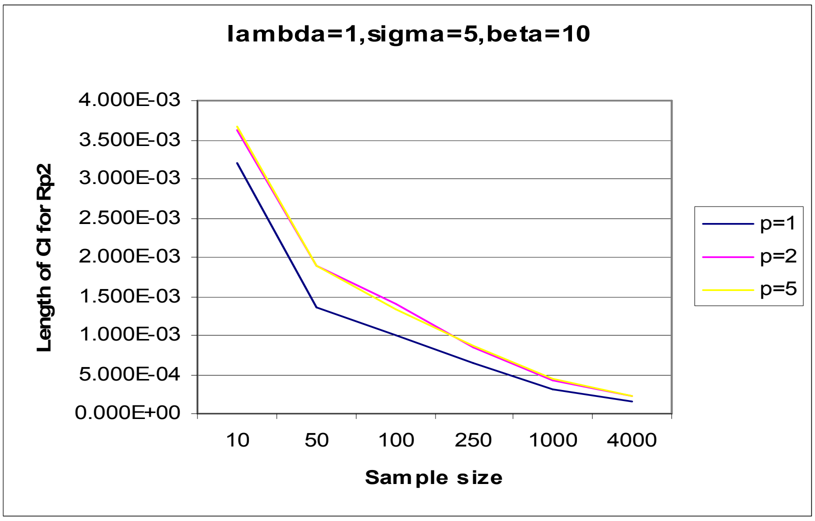



Figure 5.6: Mean square error for $\hat{\beta}$ when $\lambda=1, \beta=10, \sigma=1,5,10$



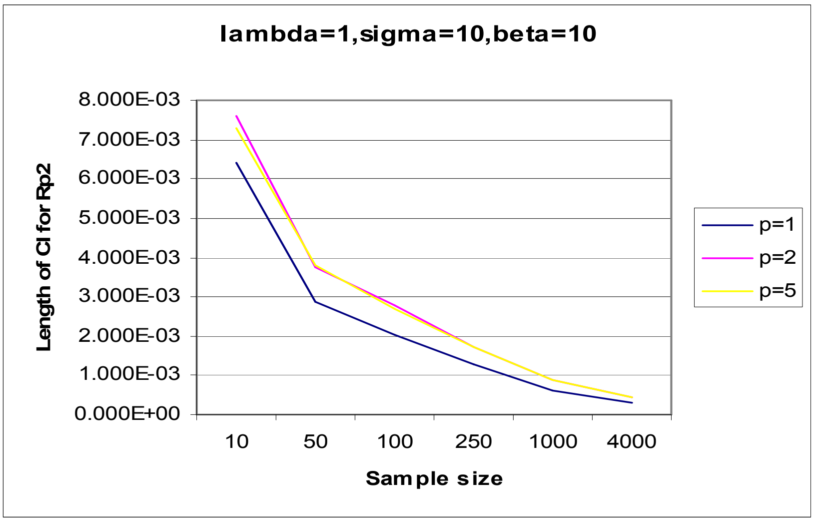



Figure 5.7: Length of confidence interval for $R_P^2$ when $\lambda = 1, \beta = 10, \sigma = 1,5,10$

## 5.4 Simulation B: Robustness of the MULFR Model When the Basic Assumptions are Violated

This subsection investigates the robustness of the estimators when one of the basic assumptions is violated. Two types of violation on the assumptions made are; (i) one of the random errors, say $\boldsymbol{\varepsilon}$, is either normal $\left(\bar{m}_3 = 0.0, \bar{m}_4 = 3.0\right)$ or non-normally distributed and (ii) the ratio of error variances, $\lambda \neq 1$. For the first case, 27 types of error distributions were obtained by using quadratic method illustrated in Section 5.2.2. These distributions are summarized by average skewness $\left(\bar{m}_3\right)$ and average kurtosis $\left(\bar{m}_4\right)$ as are given in Table 5.3. Other parameters values were set to $\boldsymbol{\alpha} = [0,\ldots,0]$, $\beta = 1$, $\sigma = 1$, $\lambda = 1$, $p = 1,2,5$ and $n = 10,100,1000$ due to heavily computation times. For the second case, the ratio of error variances change among values $\lambda = 1.5,10,30,100$ while other parameter values were set to $\boldsymbol{\alpha} = [0,\ldots,0]$, $\beta = 10$, $\sigma = 1$, $p = 1,2,5$, $n = 10,50,100,1000$, and $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ are normally distributed.

### 5.4.1 Robustness of $\hat{\beta}$ and $R_P^2$ to Non-Normality

Table 5.3 shows an example of simulation results when the error $\boldsymbol{\varepsilon}$ is deviated from normal assumption for small sample size $n=10$ and $p=1$. It is observed that the estimation of $\beta$ and $R_P^2$ are good for different distributions with small mean square errors and confidence intervals. Similar results on $M\hat{S}E_{\hat{\beta}}$, $M\hat{S}E_{R_P^2}$ and the length of confidence interval for $R_P^2$ using different sample sizes, $(n=10,100,1000)$ and dimensionalities $(p=1,2,5)$ are summarized in Figure 5.8 to Figure 5.9.

Table 5.3: Parameters estimation, standard deviation, mean square error and length of confidence interval for $\hat{\beta}$ and $R_P^2$ involving $\lambda=1, \sigma=1, \beta=1, p=1, n=10$ and varying distributions

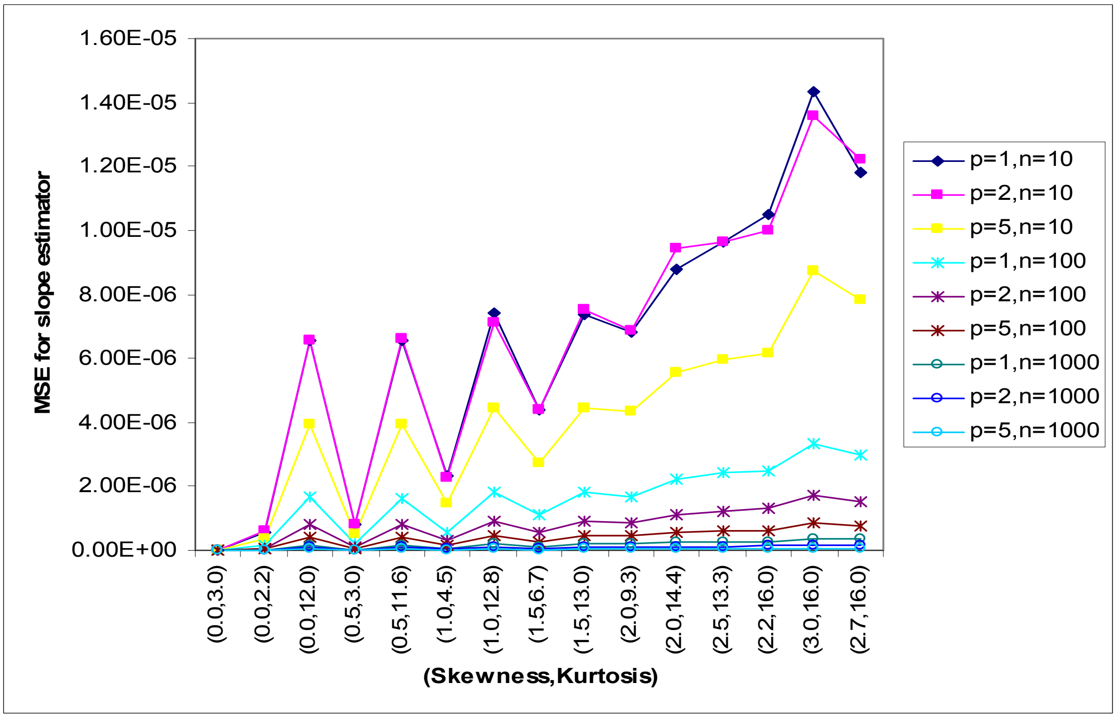
| $\bar{m}_3$ | $\bar{m}_4$ | $\bar{\hat{\beta}}$ | $M\hat{S}E_{\hat{\beta}}$ | $\bar{R}_P^2$ | $\hat{S}D_{R_P^2}$ | $M\hat{S}E_{R_P^2}$ | Length $CI\left(\bar{R}_P^2\right)$ |
|---|---|---|---|---|---|---|---|
| 0.0 | 3.0 | 1.0000 | 0.00E-00 | 1.0000 | 0.00E+00 | 0.00E+00 | 0.000E+00 |
| 0.0 | 2.2 | 1.0000 | 5.67E-07 | 1.0000 | 3.73E-06 | 1.93E-11 | 4.743E-04 |
| 0.0 | 12.0 | 1.0000 | 6.56E-06 | 1.0000 | 3.86E-05 | 2.20E-09 | 1.670E-03 |
| 0.5 | 3.0 | 1.0000 | 8.00E-07 | 1.0000 | 4.76E-06 | 3.28E-11 | 5.431E-04 |
| 0.5 | 11.6 | 1.0000 | 6.57E-06 | 1.0000 | 3.14E-05 | 1.64E-09 | 1.641E-03 |
| 1.0 | 4.5 | 1.0000 | 2.31E-06 | 1.0000 | 1.10E-05 | 2.09E-10 | 9.486E-04 |
| 1.0 | 12.8 | 1.0000 | 7.44E-06 | 1.0000 | 4.15E-05 | 2.59E-09 | 1.749E-03 |
| 1.5 | 6.7 | 1.0000 | 4.40E-06 | 1.0000 | 2.35E-05 | 8.81E-10 | 1.326E-03 |
| 1.5 | 13.0 | 1.0000 | 7.37E-06 | 1.0000 | 4.19E-05 | 2.57E-09 | 1.715E-03 |
| 2.0 | 9.3 | 1.0000 | 6.82E-06 | 1.0000 | 3.37E-05 | 1.92E-09 | 1.674E-03 |
| 2.0 | 14.4 | 1.0000 | 8.81E-06 | 1.0000 | 5.37E-05 | 4.10E-09 | 1.867E-03 |
| 2.5 | 13.3 | 1.0001 | 9.62E-06 | 1.0000 | 4.74E-05 | 3.73E-09 | 1.988E-03 |
| 2.2 | 16.0 | 1.0000 | 1.05E-05 | 1.0000 | 7.43E-05 | 7.21E-09 | 1.995E-03 |
| 3.0 | 16.0 | 1.0000 | 1.43E-05 | 0.9999 | 6.32E-05 | 6.91E-09 | 2.364E-03 |
| 2.7 | 16.0 | 1.0000 | 1.18E-05 | 1.0000 | 7.75E-05 | 8.26E-09 | 2.165E-03 |
| -0.5 | 4.1 | 1.0000 | 6.34E-07 | 1.0000 | 4.34E-06 | 2.51E-11 | 4.841E-04 |
| -0.5 | 11.9 | 1.0000 | 6.51E-06 | 1.0000 | 3.30E-05 | 1.74E-09 | 1.644E-03 |
| -1.0 | 5.7 | 1.0000 | 2.05E-06 | 1.0000 | 1.32E-05 | 2.46E-10 | 9.109E-04 |
| -1.0 | 11.6 | 1.0000 | 6.32E-06 | 1.0000 | 4.15E-05 | 2.43E-09 | 1.636E-03 |
| -1.5 | 7.5 | 1.0000 | 4.14E-06 | 1.0000 | 2.53E-05 | 9.13E-10 | 1.275E-03 |
| -1.5 | 12.1 | 1.0000 | 7.22E-06 | 1.0000 | 4.35E-05 | 2.73E-09 | 1.718E-03 |
| -2.0 | 10.2 | 1.0000 | 6.89E-06 | 1.0000 | 3.63E-05 | 2.06E-09 | 1.659E-03 |
| -2.0 | 14.7 | 1.0000 | 8.92E-06 | 1.0000 | 5.37E-05 | 4.18E-09 | 1.906E-03 |
| -2.5 | 13.5 | 1.0000 | 9.98E-06 | 1.0000 | 5.99E-05 | 5.04E-09 | 1.956E-03 |
| -2.2 | 16.7 | 1.0000 | 1.03E-05 | 1.0000 | 7.46E-05 | 7.36E-09 | 2.023E-03 |
| -3.0 | 16.7 | 1.0000 | 1.26E-05 | 0.9999 | 7.21E-05 | 7.81E-09 | 2.275E-03 |
| -2.7 | 16.7 | 1.0000 | 1.15E-05 | 1.0000 | 8.28E-05 | 9.03E-09 | 2.116E-03 |

Figure 5.8 (see Appendix A3) compares the mean square errors for $\hat{\beta}$ obtained from the normal and non-normal distributions. The *x*-axis refers to varying $\boldsymbol{\varepsilon}$ distributions labelled by (skewness, kurtosis) as given in the first two columns, Table 5.3. Figure 5.8(a) shows the effect of positive skewness and kurtosis, while Figure 5.8(b) displays the results for negative skewness and kurtosis. From the simulation study, it may be inferred that:

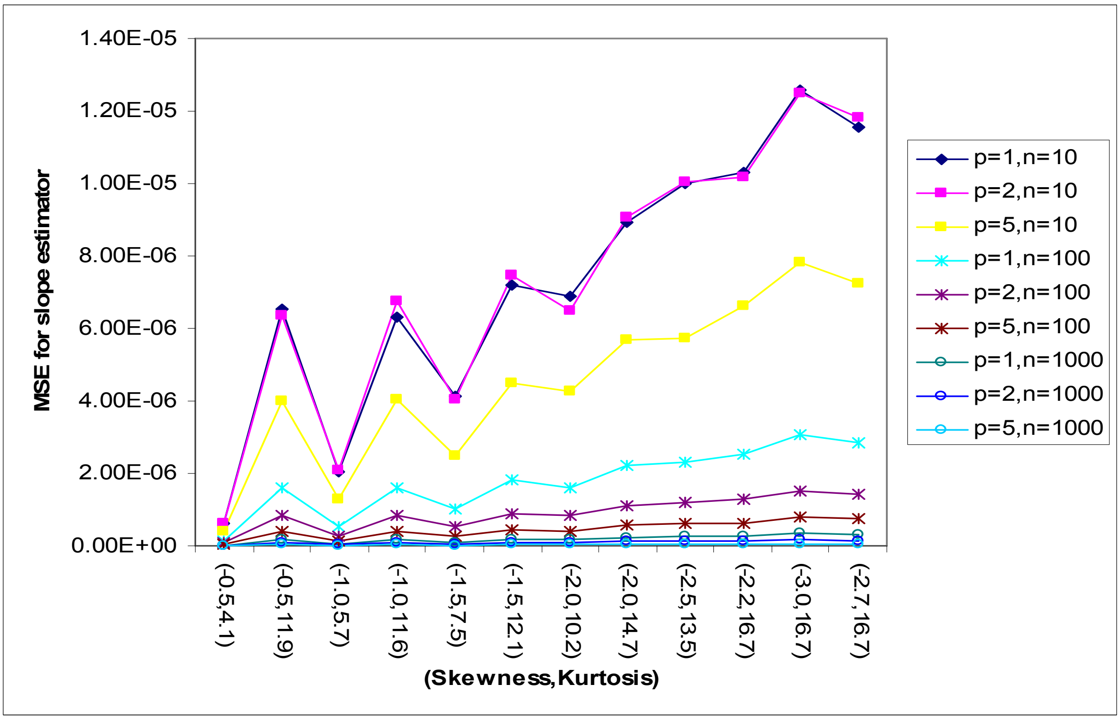
(1) The more the $\boldsymbol{\varepsilon}$ deviated (negative skewed or positive skewed) from normal distribution, the larger the $M\hat{S}E_{\hat{\beta}}$ value. For example the distributions $(\overline{m}_3 = 2.2, \overline{m}_4 = 16.0)$ , $(\overline{m}_3 = 2.7, \overline{m}_4 = 16.0)$ and $(\overline{m}_3 = 3.0, \overline{m}_4 = 16.0)$ in Figure 5.9(a), the $M\hat{S}E_{\hat{\beta}}$ increases as the skewness increases from $\overline{m}_3 = 2.2$ to $\overline{m}_3 = 3.0$ at the same kurtosis $\overline{m}_4 = 16.0$.

(2) The $M\hat{S}E_{\hat{\beta}}$ increases as the $\boldsymbol{\varepsilon}$ has larger kurtosis value given a particular level of skewness. For example at 0.5 skewness, the distribution $(\overline{m}_3 = 0.5, \overline{m}_4 = 11.6)$ has larger $M\hat{S}E_{\hat{\beta}}$ as compare to the distribution $(\overline{m}_3 = 0.5, \overline{m}_4 = 3.0)$.

(3) The $M\hat{S}E_{\hat{\beta}}$ increases as the sample size decreases when $\boldsymbol{\varepsilon}$ deviated from normal distribution.

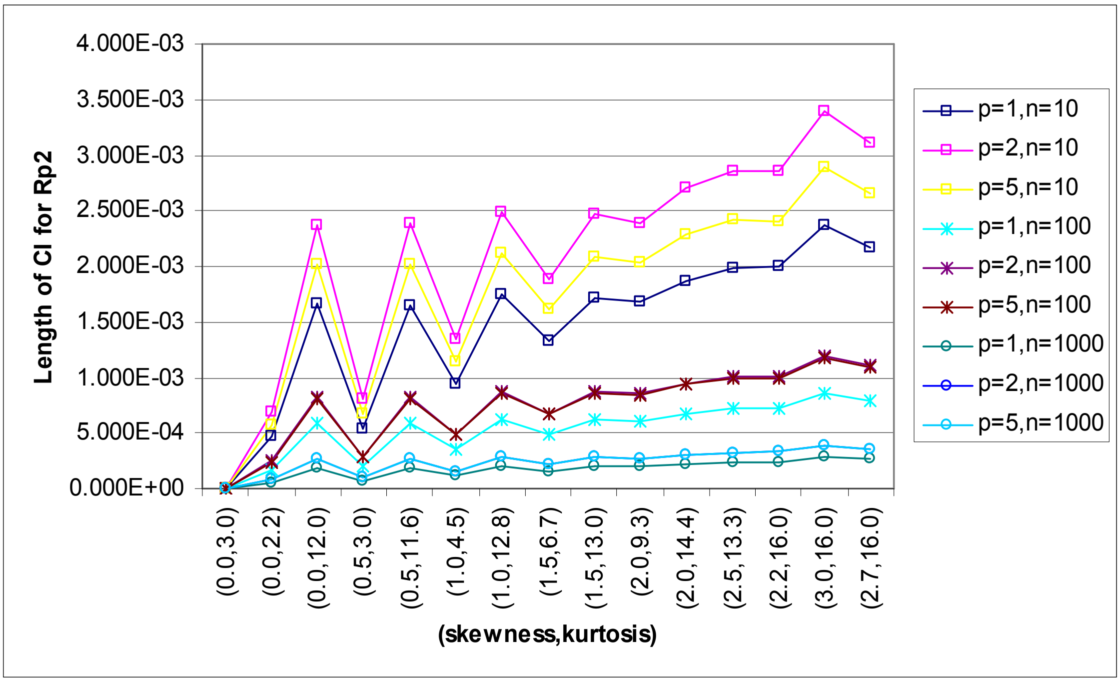



(a) Normal distribution and fourteen non-normal distributions with varying positive skewness and kurtosis values.

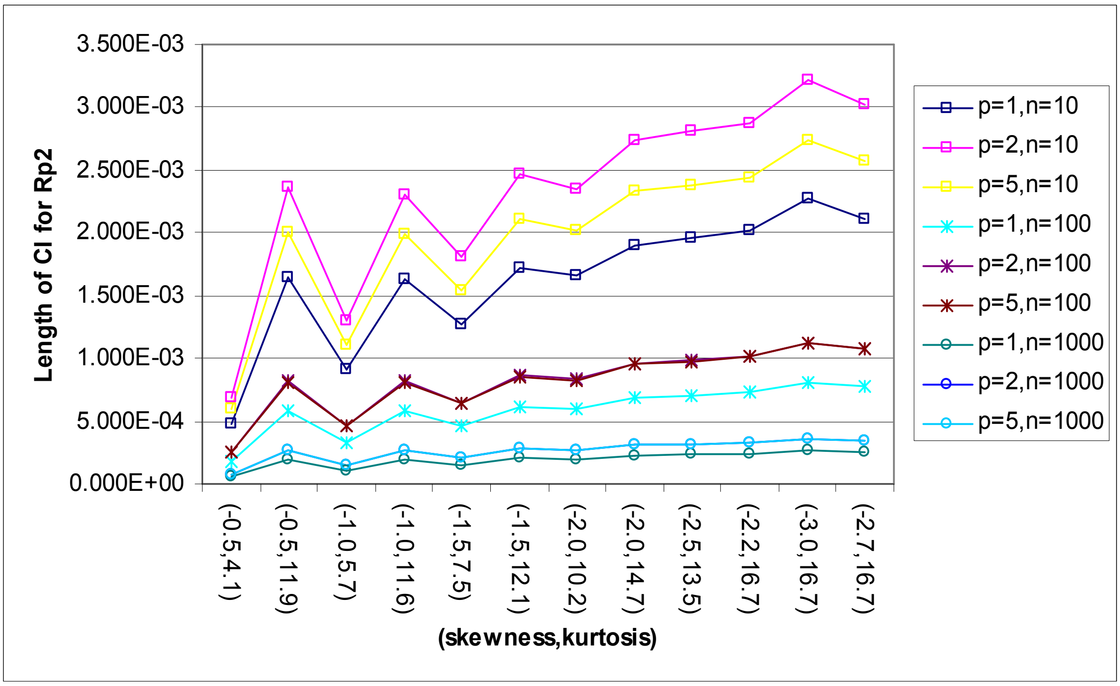



(b) Twelve non-normal distributions with varying negative skewness and kurtosis values.

Figure 5.8: Mean square error for $\hat{\beta}$ under varying error distributions when $p = 1,2,5$ and $n = 10,100,1000$.



(a) Normal distribution and fourteen non-normal distributions with varying positive skewness and kurtosis values.



(b) Twelve non-normal distributions with varying negative skewness and kurtosis values.

Figure 5.9: Length of confidence interval for $R_P^2$ under varying error distributions when $p = 1, 2, 5$ and $n = 10, 100, 1000$.

Figure 5.9 compares the length of confidence interval for $R_P^2$ obtained from the varying distributions. The *x*-axis was defined in the same manner as Figure 5.8. From the simulation study, it may be inferred that:

(1) The length of confidence interval increases as the $\boldsymbol{\varepsilon}$ deviated further (negative skewed or positive skewed) from normal distribution.

(2) The length of confidence interval increases as the $\boldsymbol{\varepsilon}$ has larger kurtosis values.

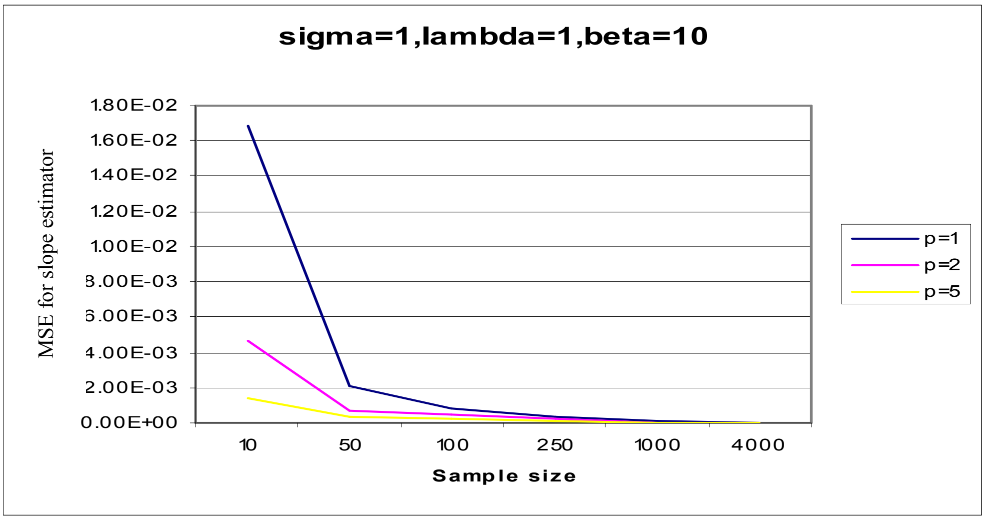
(3) The length of confidence interval increases as the sample size decreases when $\boldsymbol{\varepsilon}$ deviated from normal distribution.

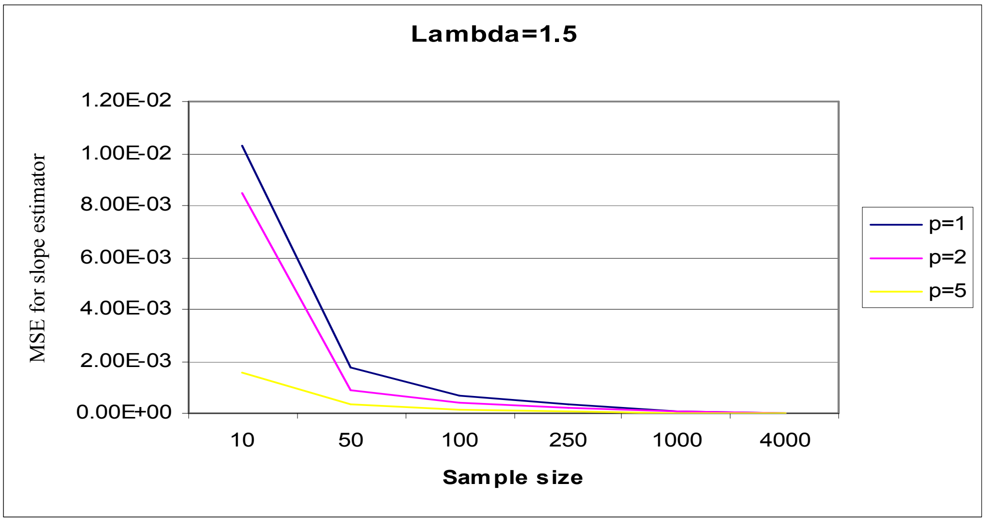
(4) Generally, the length of confidence interval increases as the dimension increases except for the case when sample size is small $(n=10)$.

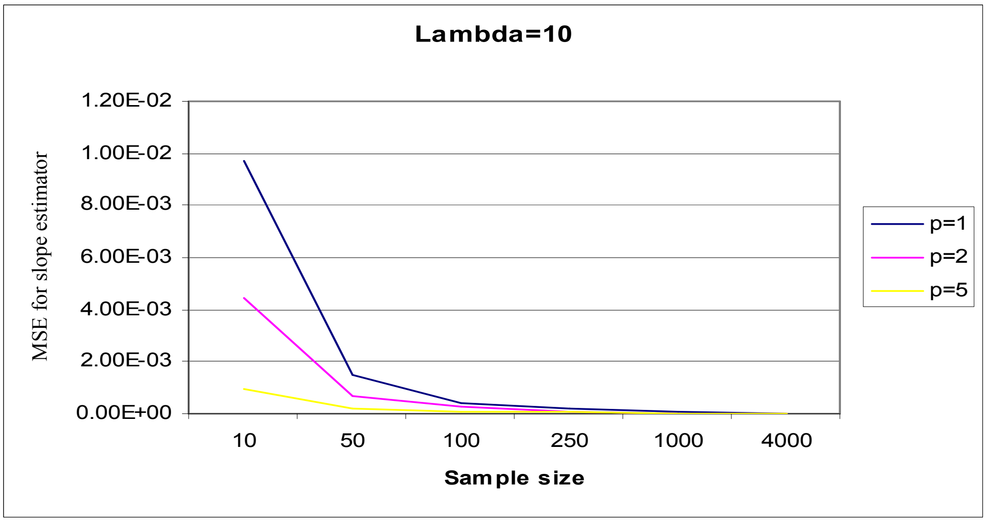
Generally, the increases in $M\hat{S}E_{\hat{\beta}}$, $M\hat{S}E_{R_P^2}$ and the length of confidence interval for $R_P^2$ is due to the sample size, error variance, ratio of the error variances and non-normality are negligible. For example, the largest $M\hat{S}E_{\hat{\beta}}$ value (see Figure 5.8) is $1.43\times10^{-5}$ and the largest length of confidence interval for $R_P^2$ (see Figure 5.9) is $3.397\times10^{-3}$ at error distribution with $\bar{m}_3=3.0$ and $\bar{m}_4=16.0$ (labelled with 14). These values are still considered very small and the estimated values were not significantly different from their true values. It is concluded that $\hat{\beta}$ and $R_P^2$ are robust estimators even when the normality assumption is not satisfied.

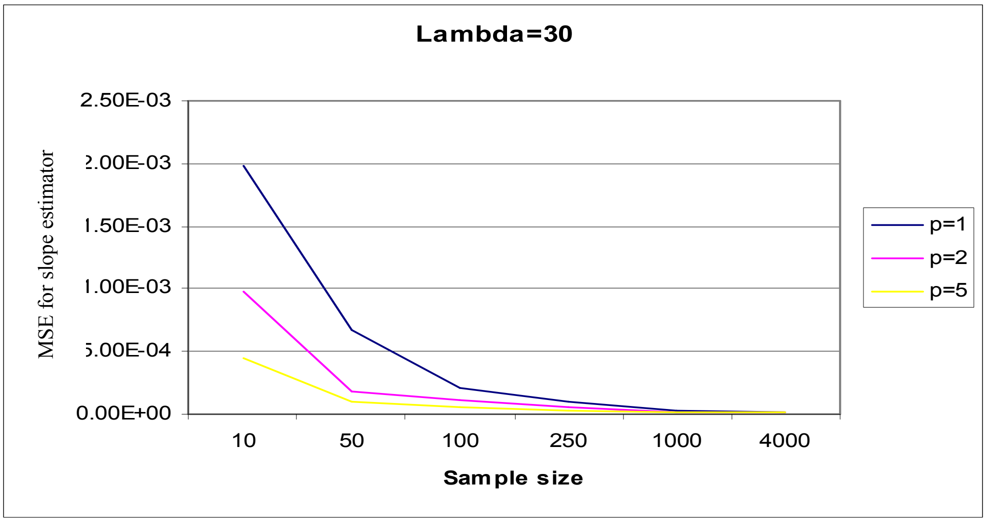
### 5.4.2 Robustness of $\hat{\beta}$ and $R_P^2$ to $\lambda\neq1$ When $\beta=10$ and $\sigma=1$

Figure 5.10 and Figure 5.11 (see Appendix A4) show the mean square error for $\hat{\beta}$ and the length of confidence interval for $R_P^2$, respectively when the true $\lambda$ =1.0, 1.5, 10, 30 and 100. The parameters $\beta$ and $R_P^2$ are well estimated for the selected true $\lambda$ values, dimensions and sample sizes $(n=10,50,100,250,1000,4000)$.


sigma=1,lambda=1,beta=10
MSE for slope estimator
1.80E-02
1.60E-02
1.40E-02
1.20E-02
1.00E-02
8.00E-03
6.00E-03
4.00E-03
2.00E-03
0.00E+00
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5


Lambda=1.5
MSE for slope estimator
1.20E-02
1.00E-02
8.00E-03
6.00E-03
4.00E-03
2.00E-03
0.00E+00
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5


Lambda=10
MSE for slope estimator
1.20E-02
1.00E-02
8.00E-03
6.00E-03
4.00E-03
2.00E-03
0.00E+00
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5



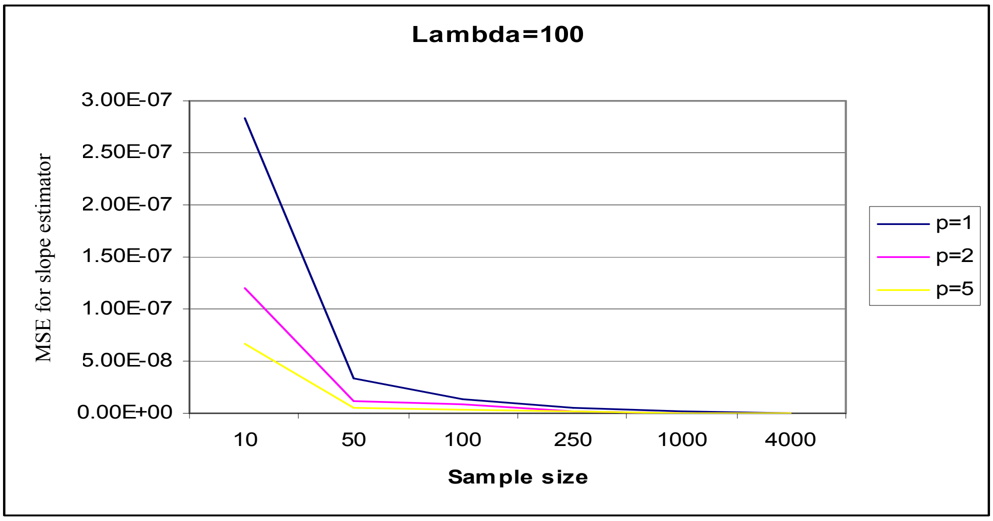
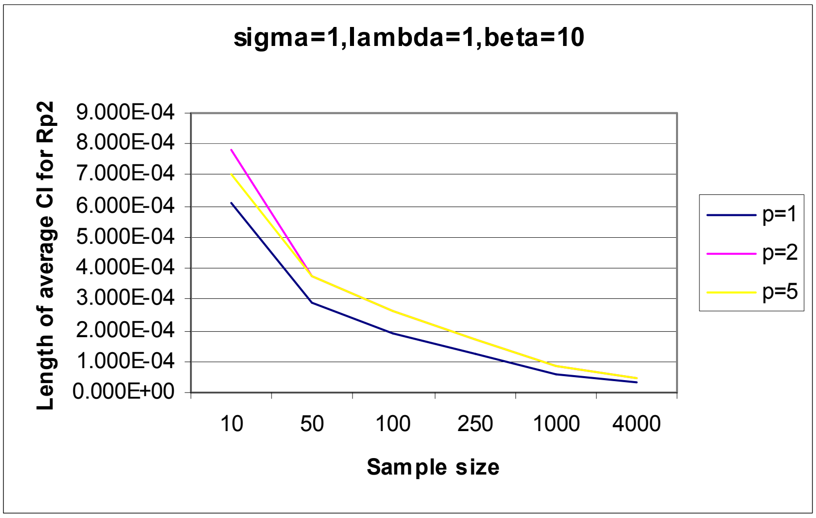
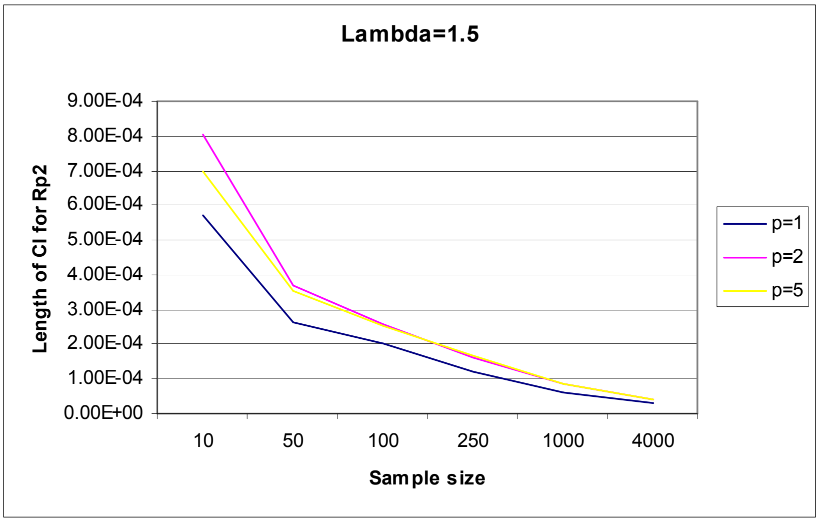
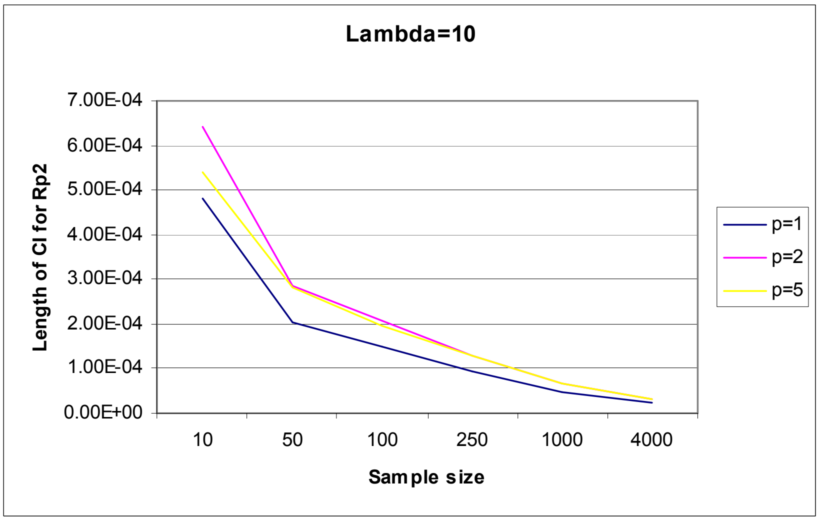




Figure 5.10: Mean square error for $\hat{\beta}$ when $\sigma = 1, \beta = 10, \lambda = 1.0, 1.5, 10, 30, 100$

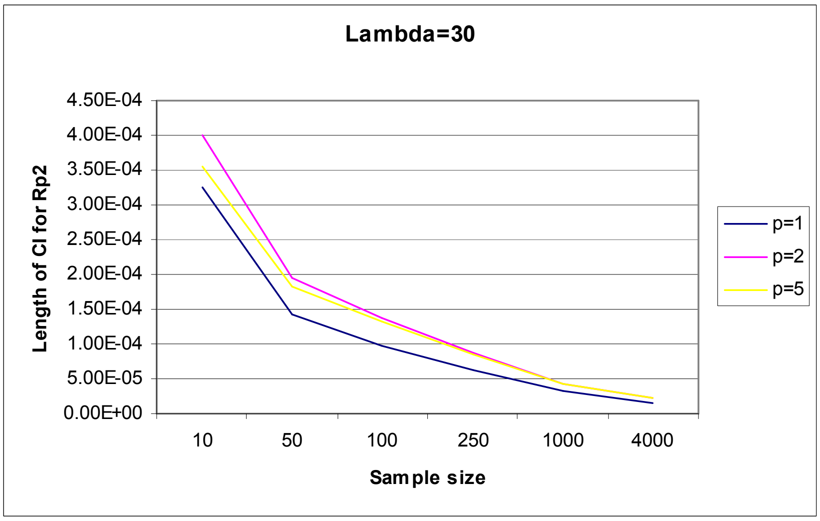




Lambda=1.5
9.00E-04
8.00E-04
7.00E-04
6.00E-04
5.00E-04
4.00E-04
3.00E-04
2.00E-04
1.00E-04
0.00E+00
Length of CI for Rp2
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5

Lambda=10
7.00E-04
6.00E-04
5.00E-04
4.00E-04
3.00E-04
2.00E-04
1.00E-04
0.00E+00
Length of CI for Rp2
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5

Lambda=30
4.50E-04
4.00E-04
3.50E-04
3.00E-04
2.50E-04
2.00E-04
1.50E-04
1.00E-04
5.00E-05
0.00E+00
Length of CI for Rp2
10
50
100
250
1000
4000
Sample size
p=1
p=2
p=5

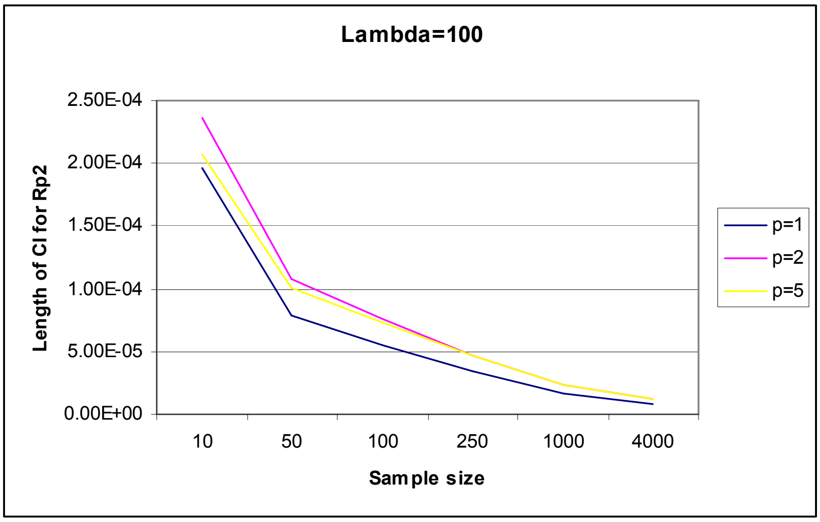



Figure 5.11: Length of confidence interval for $R_P^2$ when $\sigma=1, \beta=10, \lambda=1.0, 1.5, 10, 30, 100$

## 5.5 Properties of Maximum Likelihood Estimator, $\hat{\beta}$

Result 4 in Section 4.2.1 showed that $\hat{\beta}$ is an asymptotically an unbiased estimator of $\beta$. The unbiasedness of $\hat{\beta}$ (and hence of $R_P^2$ since it is only depends on $\hat{\beta}$) can be verified in Figure 5.4, Figure 5.6 and Figure 5.10 where there is a small mean square error for $\hat{\beta}$, for varying parameters conditions.

Furthermore, Figure 5.12 indicates that the estimated standard deviation of $\hat{\beta}$ approaches zero as the sample size increases. This trend verifies the result in Section 4.2.3 and showed that $\hat{\beta}$ (and $R_P^2$) is a consistent estimator of $\beta$.

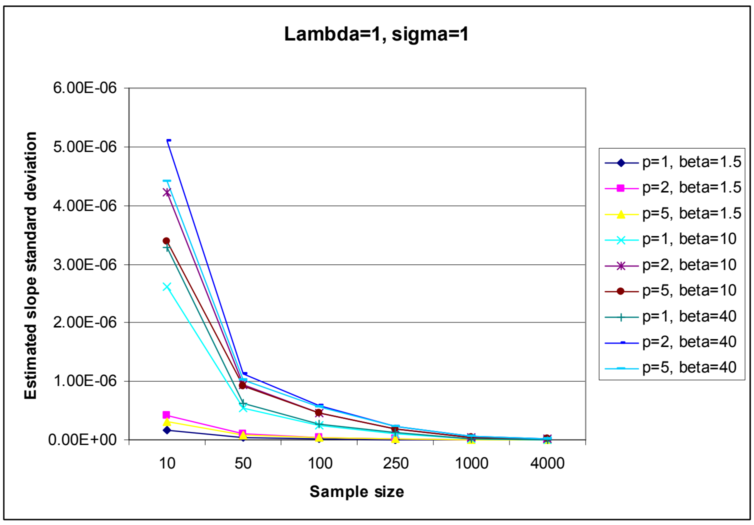



Figure 5.12: Consistency of $\hat{\beta}$

## 5.6 Empirical Distribution of $R_P^2$

In practice, there is one estimated value of $R_P^2$ from a given data set. Knowledge of the distribution of $R_P^2$ would be useful to investigate if $R_P^2$ is significantly far from the mean of $R_P^2$. The 10000 values of simulated $R_P^2$ from Section 5.1.2 were standardized using $\left(R_P^2 - \overline{R}_p^2\right)\Big/\sqrt{\operatorname{var}\left(R_P^2\right)}$. The empirical cumulative density function (CDF) for $R_P^2$ is studied and compared with CDF for standard normal distribution. The CDF plots are shown in Figure 5.13 to Figure 5.18.

Three levels of $R_P^2$ values are considered empirically, they are high $R_P^2$ values with $\overline{R}_p^2$ close to one (see Figure 5.13 and Figure 5.14), moderate $R_P^2$ values with $\overline{R}_p^2$ between 0.5 and 0.6 (see Figure 5.15 and Figure 5.16), and low $R_P^2$ values with $\overline{R}_p^2$ close to zero (see Figure 5.17 and Figure 5.18). These figures (after standardization) indicate that $R_P^2$ has non-symmetrical distribution for moderate sample sizes. A Kolmogorov-Smirnov one-sample goodness-of-fit test (Daniel, 1990) is applied to determine whether the sampled population of $R_P^2$ is normally distributed as stated in hypothesis below:

$$H_0 : F(x) = F_0(x) \text{ for all values of } x$$

$$H_1 : F(x) \neq F_0(x) \text{ for at least one value of } x$$

Let $F_0(x)$ be the hypothesized standard normal distribution function and $S(x)$ designates the empirical distribution function of standardized $R_P^2$. Specifically,

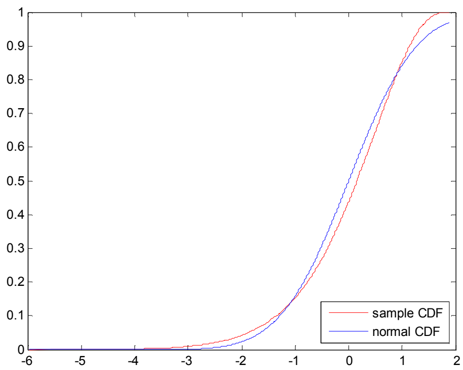
$$S(x) = \text{the proportion of sample observations less than or equal to } x.$$

The two-sided test statistics is

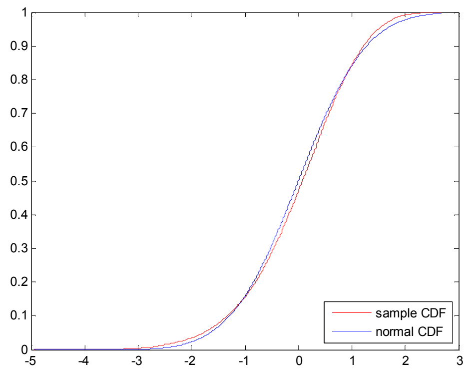
$$D = \sup_x \left| S(x) - F_0(x) \right|$$

Reject $H_0$ at the $\alpha$ specified level of significance if the test statistics exceeds the $1-\alpha$ quantiles of the Kolmogorov test statistic.

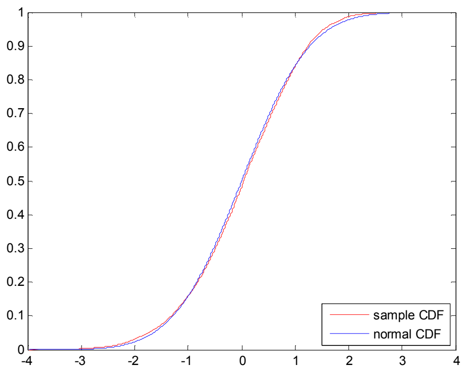
The hypothesis test results are summarized in Table 5.4. It indicates that the null hypothesis is not rejected for at least moderate sample sizes $(n \geq 250)$ when $R_P^2$ value is large and it is not significant at large sample sizes $(n = 1000, 4000)$ when $R_P^2$ value is moderate and small. This implies that $R_P^2$ is asymptotically (standard) normally distributed, which can also be derived from Section 4.2.4 where $\hat{\beta}$ has asymptotical normal distribution.



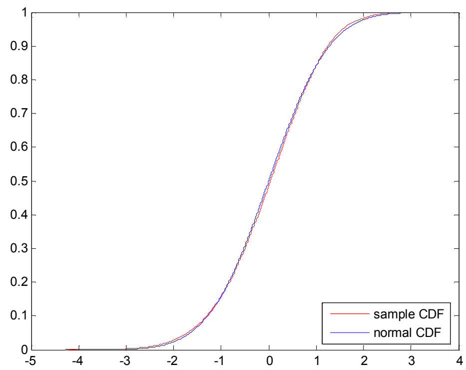

(a) $p = 1, n = 10$



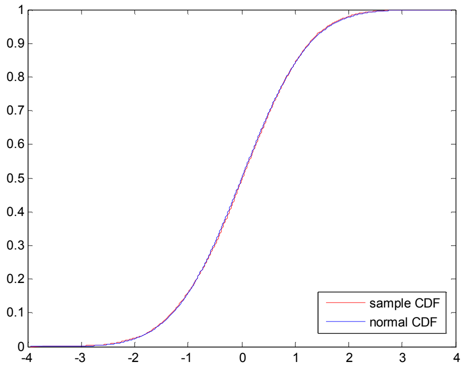

(b) $p = 1, n = 50$



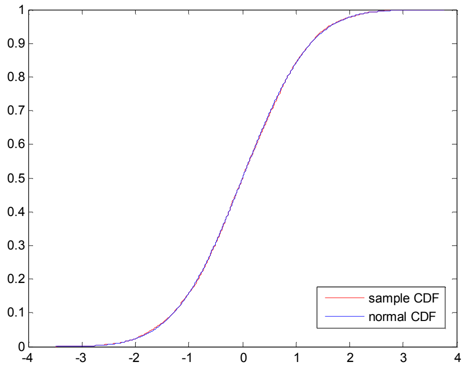

(c) $p = 1, n = 100$




(d) $p = 1, n = 250$

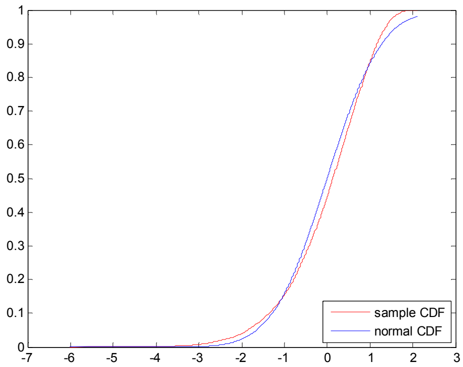



(e) $p=1, n=1000$

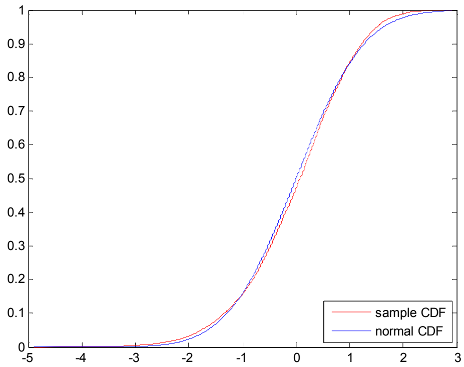



(f) $p=1, n=4000$

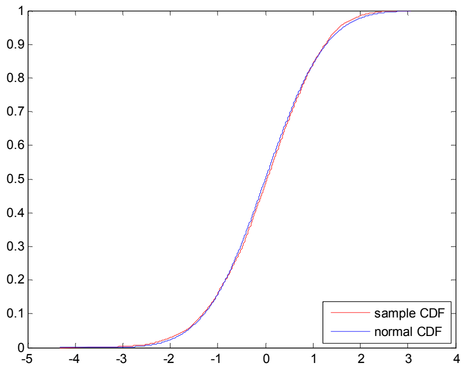
Figure 5.13: Empirical cumulative density functions of $R_P^2$ at high correlation value (average $R_P^2 \approx 0.99$ ) and CDF for standard normal distribution. Given $p=1$



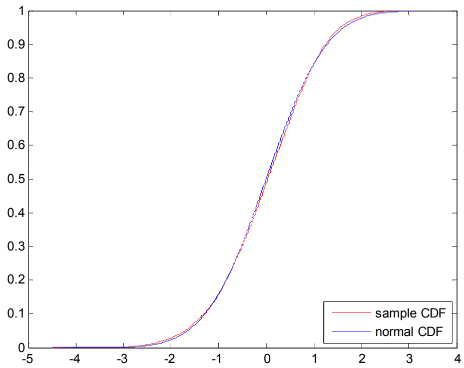

(a) $p=5, n=10$



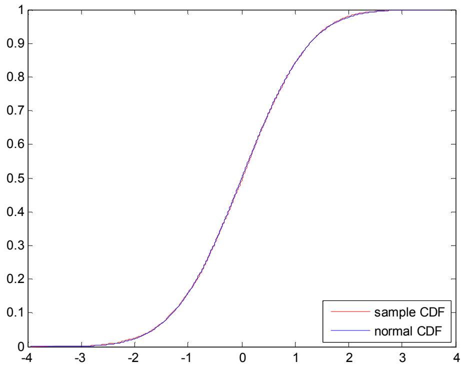

(b) $p=5, n=50$



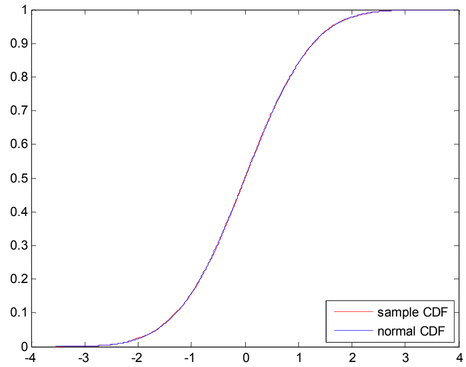

(c) $p=5, n=100$




(d) $p=5, n=250$

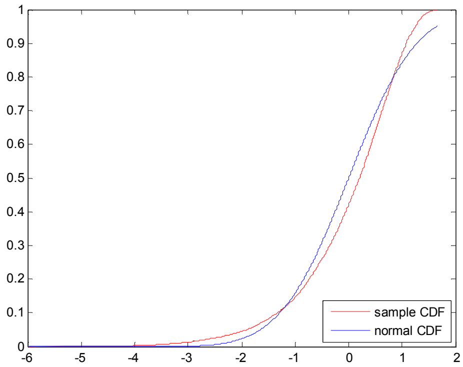



(e) $p = 5, n = 1000$

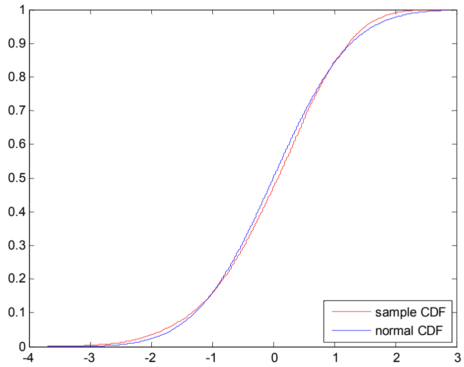



(f) $p = 5, n = 4000$

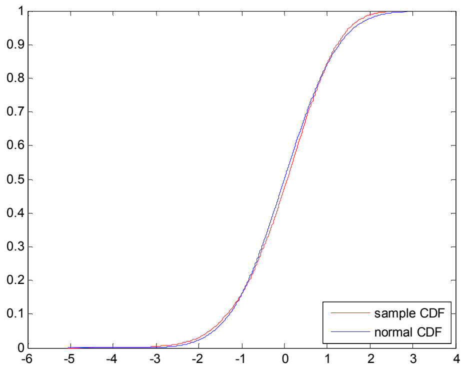
Figure 5.14: Empirical cumulative density functions of $R_P^2$ at high correlation value (average $R_P^2 \approx 0.99$ ) and CDF for standard normal distribution. Given $p = 5$.



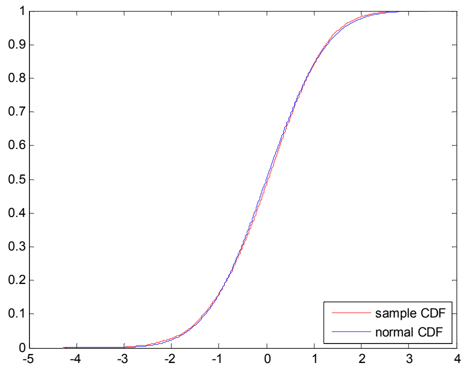

(a) $p = 1, n = 10$



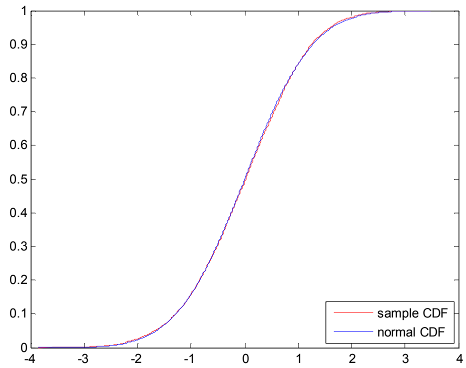

(b) $p = 1, n = 50$



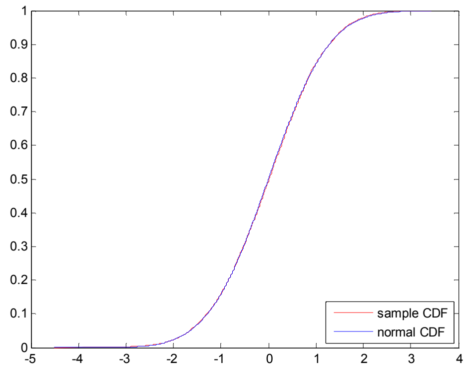

(c) $p = 1, n = 100$




(d) $p = 1, n = 250$

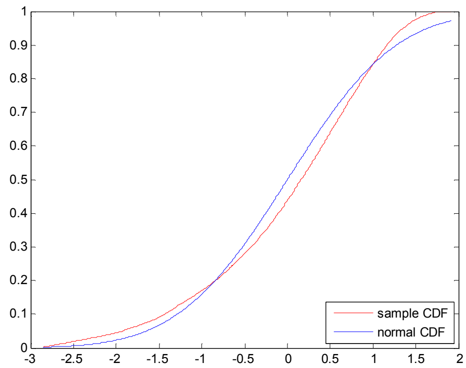



(e) $p = 1, n = 1000$

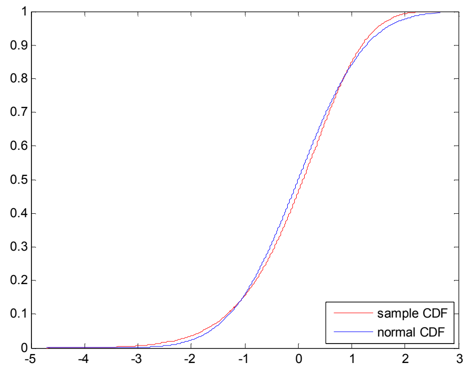



(f) $p = 1, n = 4000$

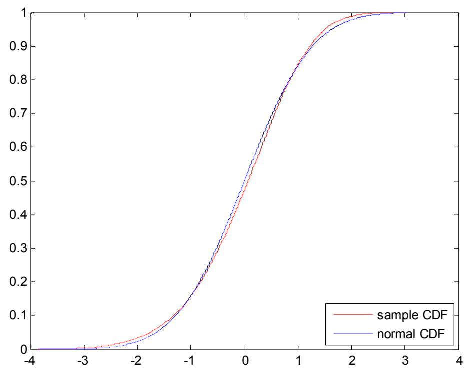
Figure 5.15: Empirical cumulative density functions of $R_P^2$ at moderate correlation value (average $R_P^2 \approx 0.56$ ) and CDF for standard normal distribution. Given $p = 1$ .



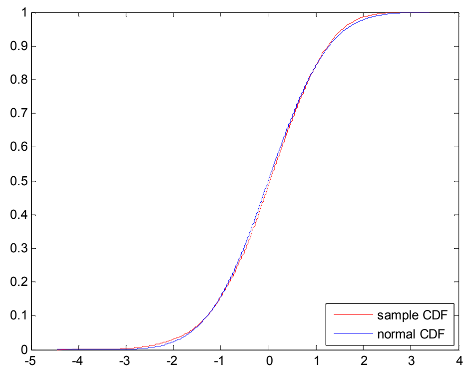

(a) $p = 5, n = 10$



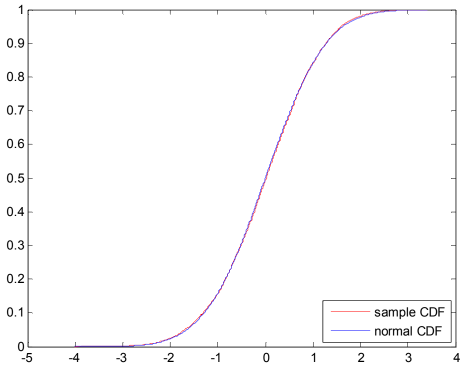

(b) $p = 5, n = 50$



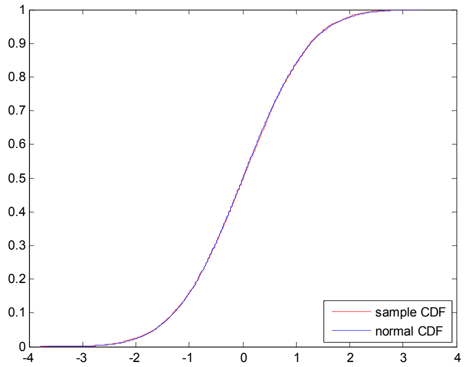

(c) $p = 5, n = 100$




(d) $p = 5, n = 250$

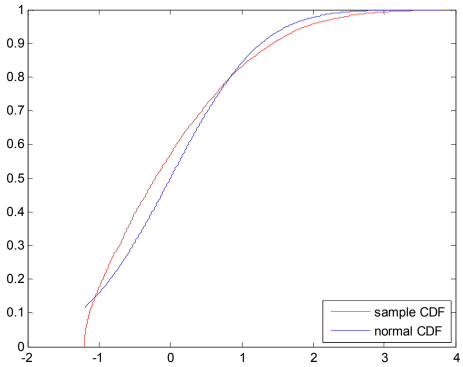



(e) $p = 5, n = 1000$

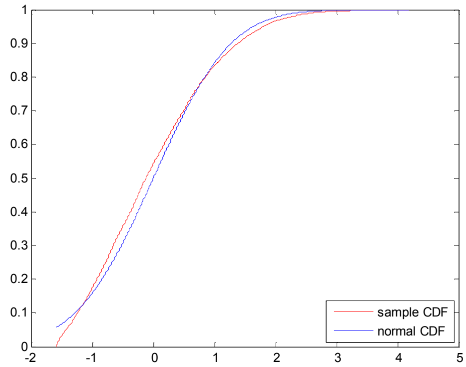



(f) $p = 5, n = 4000$

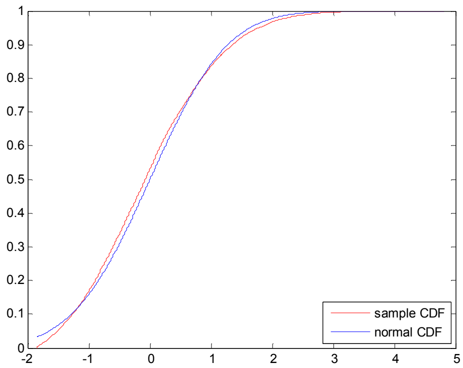
Figure 5.16: Empirical cumulative density functions of $R_P^2$ at moderate correlation value (average $R_P^2 \approx 0.58$ ) and CDF for standard normal distribution. Given $p = 5$ .



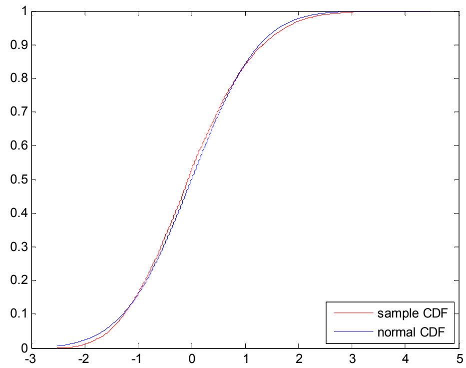

(a) $p = 1, n = 10$



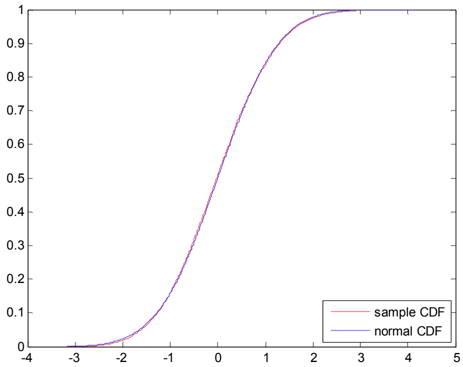

(b) $p = 1, n = 50$



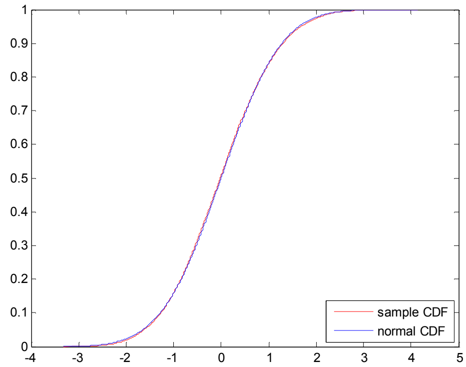

(c) $p = 1, n = 100$




(d) $p = 1, n = 250$

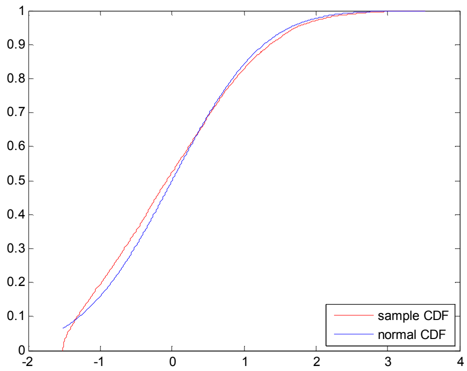



(e) $p = 1, n = 1000$

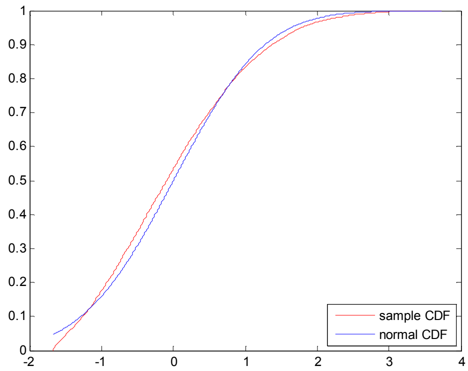



(f) $p = 1, n = 4000$

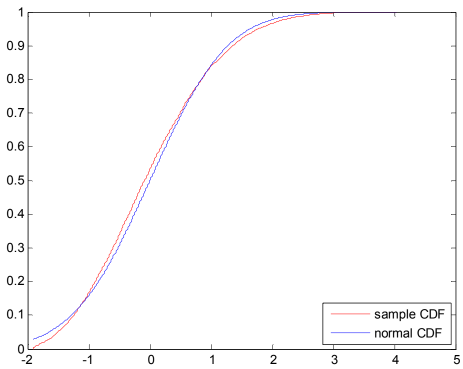
Figure 5.17: Empirical cumulative density functions of $R_P^2$ at low correlation value (average $R_P^2 \approx 0.1$) and CDF for standard normal distribution. Given $p = 1$.



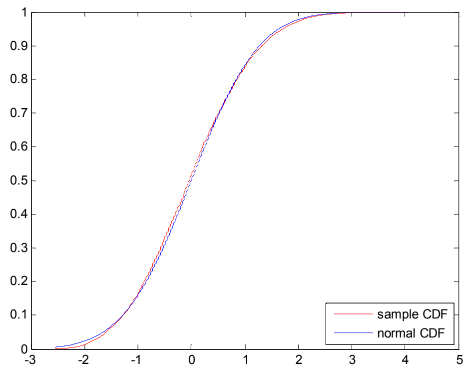

(a) $p = 5, n = 10$



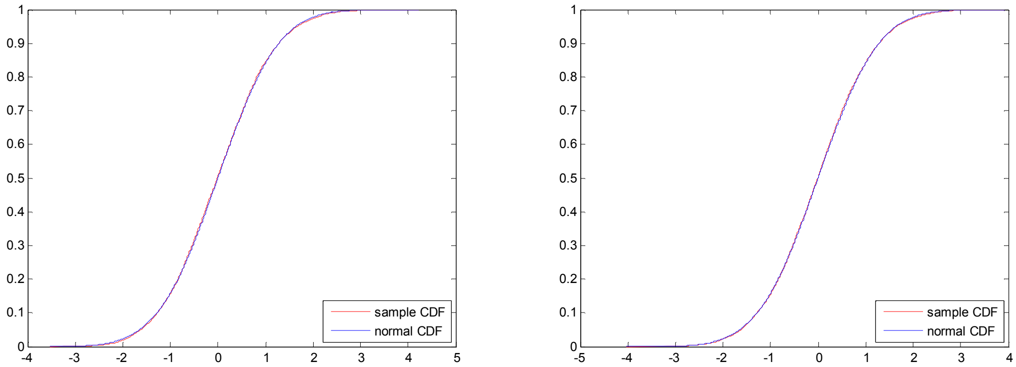

(b) $p = 5, n = 50$




(c) $p = 5, n = 100$




(d) $p = 5, n = 250$




(e) $p = 5, n = 1000$ (f) $p = 5, n = 4000$

Figure 5.18: Empirical cumulative density functions of $R_P^2$ at low correlation value (average $R_P^2 \approx 0.1$) and CDF for standard normal distribution. Given $p = 5$

Table 5.4: Summary for Kolmogorov-Smirnov test.

| | High $R_P^2$ value | | | | Moderate $R_P^2$ value | | | | Low $R_P^2$ value | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *p*=1 | | *p*=5 | | *p*=1 | | *p*=5 | | *p*=1 | | *p*=5 | |
| *n* | $H_0$ true? | *p*-value | $H_0$ true? | *p*-value | $H_0$ true? | *p*-value | $H_0$ true? | *p*-value | $H_0$ true? | *p*-value | $H_0$ true? | *p*-value |
| 10 | No | 1.29E-35 | No | 1.36E-27 | No | 1.37E-59 | No | 1.14E-37 | No | 2.12E-112 | No | 1.02E-36 |
| 50 | No | 8.71E-08 | No | 4.67E-09 | No | 2.00E-09 | No | 7.05E-13 | No | 1.74E-28 | No | 5.09E-20 |
| 100 | No | 9.40E-05 | No | 0.003 | No | 2.93E-08 | No | 7.61E-08 | No | 1.89E-10 | No | 7.96E-13 |
| 250 | Yes | 0.0551 | No | 0.0051 | No | 0.0019 | No | 0.0017 | No | 1.04E-08 | No | 2.60E-04 |
| 1000 | Yes | 0.4335 | Yes | 0.5038 | Yes | 0.1123 | Yes | 0.2752 | Yes | 0.112 | Yes | 0.2403 |
| 4000 | Yes | 0.677 | Yes | 0.9677 | Yes | 0.662 | Yes | 0.5907 | Yes | 0.213 | Yes | 0.36 |

## 5.7 Summary

Simulations were conducted to study the performance of $R_P^2$ (and $\hat{\beta}$) under a specific set of parameter values. These parameters are $\beta$, $\sigma$, $n$, $p$, $\lambda$ and $\varepsilon$'s distribution. The selection of parameters values were based on the real image applications.

Several results were observed from the simulation studies. When the assumptions of unit ratio of error variances and normality are satisfied, both $\hat{\beta}$ and $R_P^2$ are well estimated for various combinations of parameters. These parameters combinations are $\{(\beta,\sigma) \mid \beta = 1, 1.5, 10, 40; \sigma = 1, 5, 10\}$ for all $p = 1, 2, 5$ and $n = 10, 50, 100, 250, 1000, 4000$.

Both $\hat{\beta}$ and $R_P^2$ are robust estimators when there is mild violation of normality assumption of the error term, $\boldsymbol{\varepsilon}$. The skewness and kurtosis of these non-normal distributions considered a range between $m_3 \in [-3.0, 3.0]$ and $m_4 \in [2.2, 16.7]$, respectively. Similar results were observed when the assumption of $\lambda = 1$ is violated. Both $\hat{\beta}$ and $R_P^2$ still perform well for $\lambda \leq 100$.

It was showed that $\hat{\beta}$ and $R_P^2$ still maintained a good properties of consistent and unbiased estimators at small sample size of $n = 10$. Meanwhile, $R_P^2$ has asymptotically normal distribution for moderate sample sizes $(n \geq 250)$. This condition of moderate sample size is not difficult to fulfill as the number of data in an image are usually very large.

This simulation study considered a broad range of conditions of comparing two images by changing the parameters values. However, it is impossible to cover all possible types of images that may occur in the real applications. For example, this study only discussed the Full Reference case where the same information is available from both the 'transformed' image $\boldsymbol{Y}$ and the reference image $\boldsymbol{X}$. In JPEG compression application, the transformed $\boldsymbol{Y}$ image is referred to JPEG codec image that subjected to Laplacian noise. The image $\boldsymbol{X}$ is a perfect full reference when it is noise-free, while it is called an imperfect full reference when significant amount of noise $\boldsymbol{\delta}$ is presented. Another issue that was not addressed theoretically is the uncorrelated errors case. For example, the neighbour pixels within an image tend to be highly correlated. In fact, this issue was solved by an appropriate choice of window size of $8 \times 8$ in JPEG compression (Gonzalez et al., 2004) to minimize the possible correlation between windows.