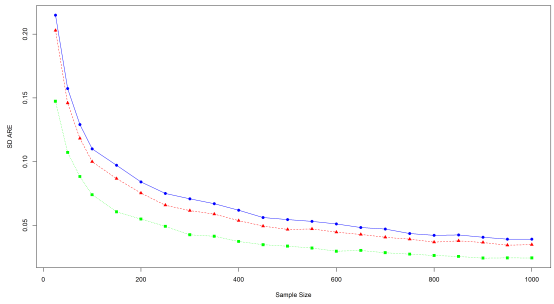
Supplementary Material for: Estimating Standard Deviation via Sample Mean Extended Quantile Estimation

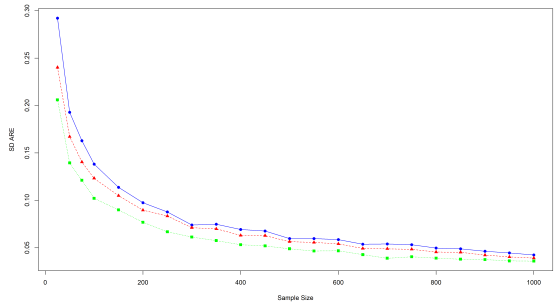
Contents

1 ARE Comparisons (With Absolute Value)

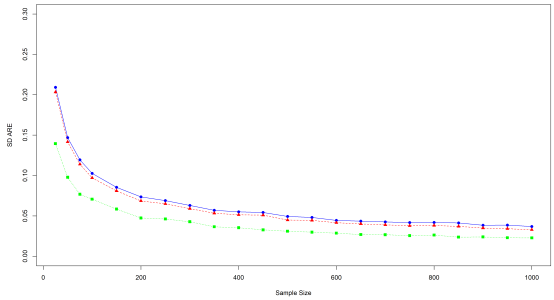
1.1 Against theoretical SD under Scenario S2



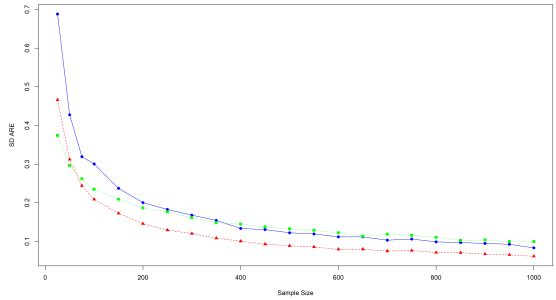
Log-normal (4, 0.3)



Log-normal (5, 0.5)

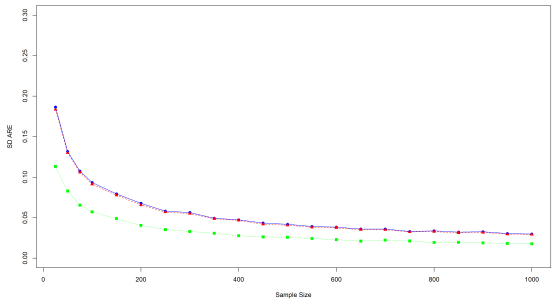


Log-normal (5, 0.25)

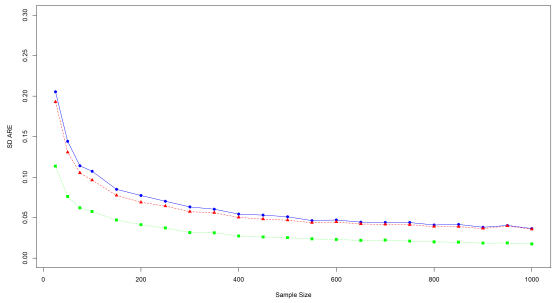


Log-normal (5, 1)

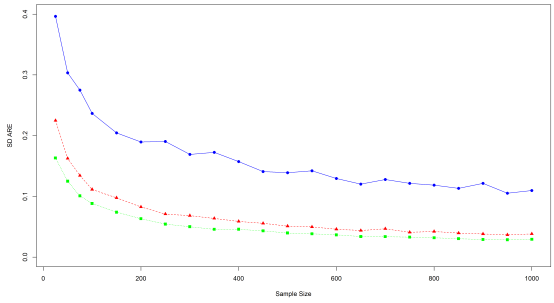
Figure 1: Comparisons against theoretical standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1) in scenario S2 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



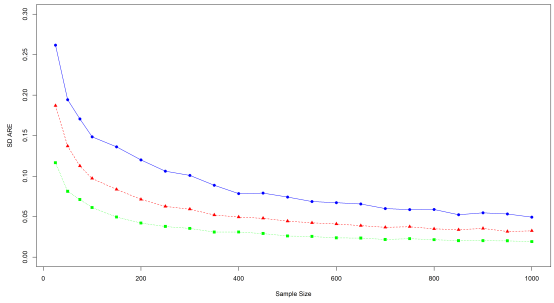
Normal (5,1)



Normal (50, 17)



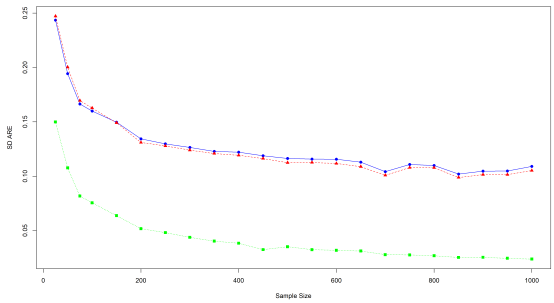
Gamma (2, 5)



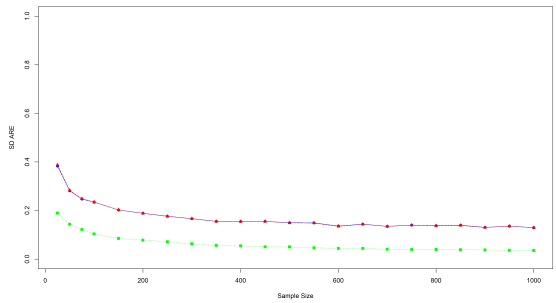
Weibull (2,35)

Figure 2: Comparisons against theoretical standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S2 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

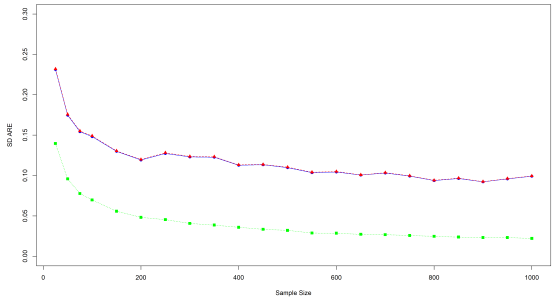
1.2 Against theoretical SD under Scenario S3



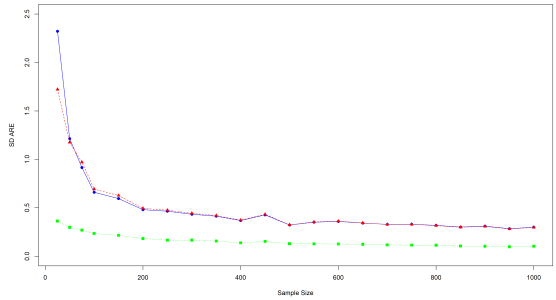
Log-normal (4, 0.3)



Log-normal (5, 0.5)

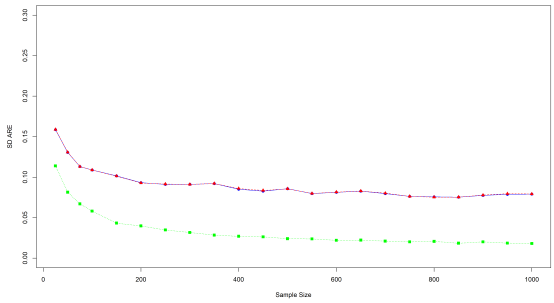


Log-normal (5, 0.25)

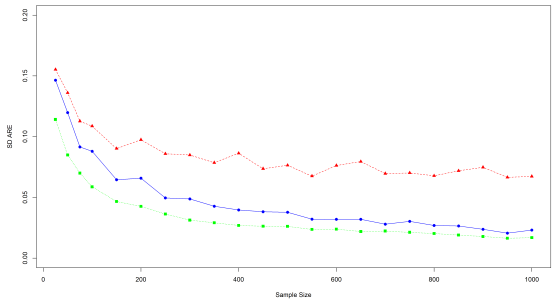


Log-normal(5, 1)

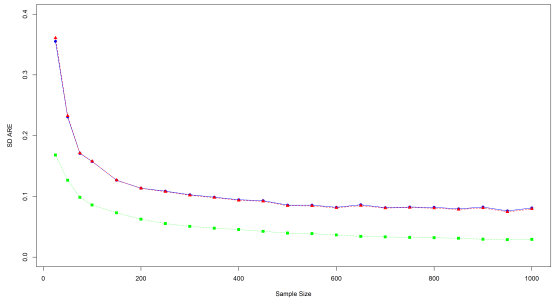
Figure 3: Comparisons against theoretical standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1) in scenario S3 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



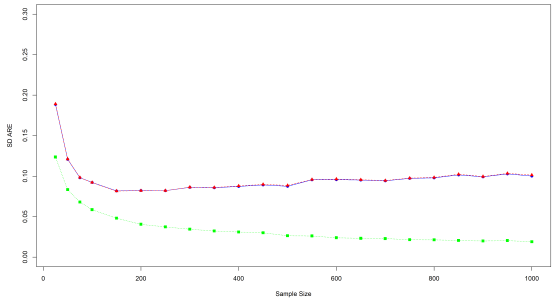
Normal (5,1)



Normal (50, 17)



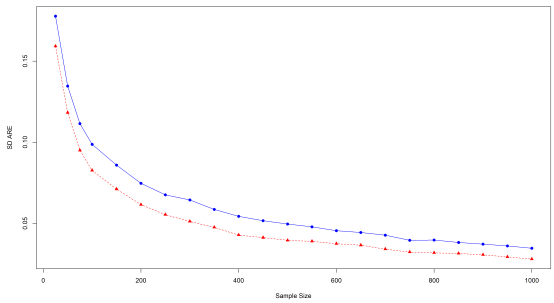
Gamma (2, 5)



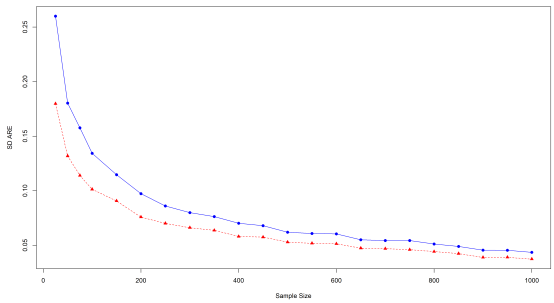
Weibull (2, 35)

Figure 4: Comparisons against theoretical standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S3 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

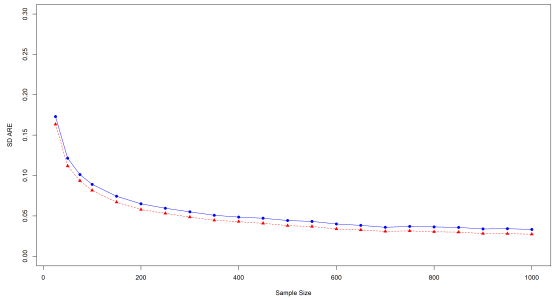
1.3 Against Sample SD under Scenario S2



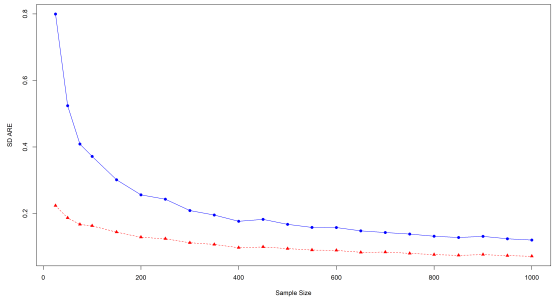
Log-normal (4, 0.3)



Log-normal (5, 0.5)

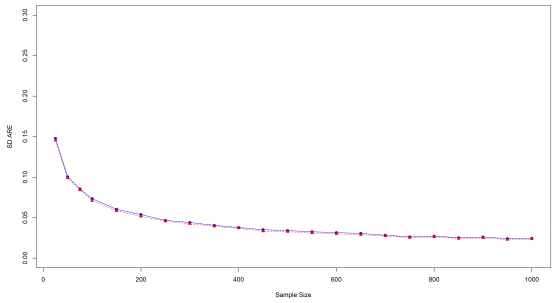


Log-normal (5, 0.25)

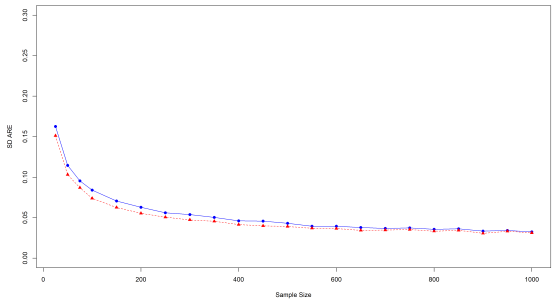


Log-normal (5, 1)

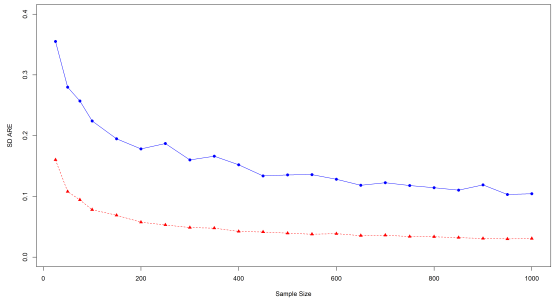
Figure 5: Comparisons against sample standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1), in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



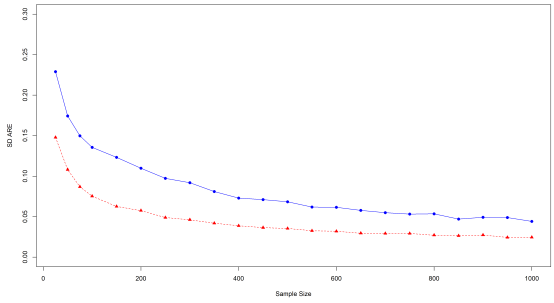
Normal (5, 1)



Normal (50, 17)



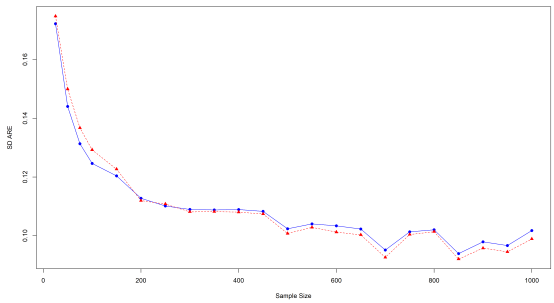
Gamma (2,5)



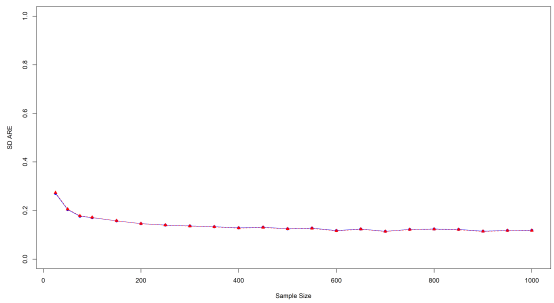
Weibull (2, 35)

Figure 6: Comparisons against sample standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

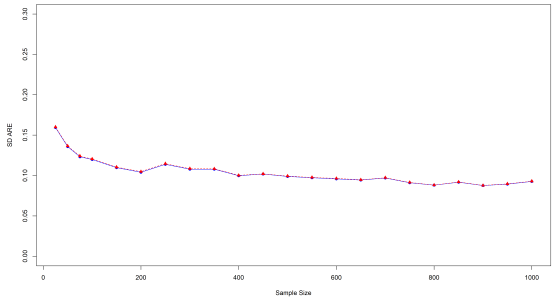
1.4 Against Sample SD under Scenario S3



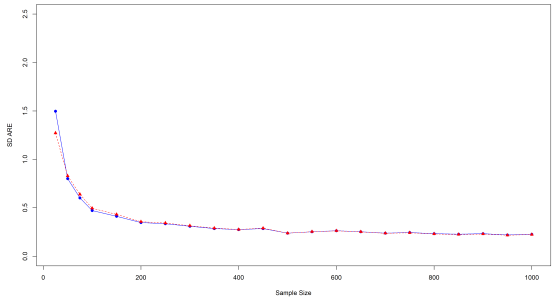
Log-normal (4, 0.3)



Log-normal (5, 0.5)

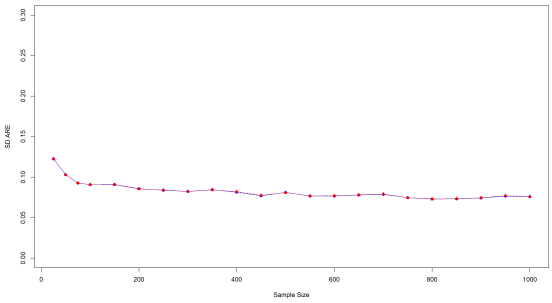


Log-normal (5, 0.25)

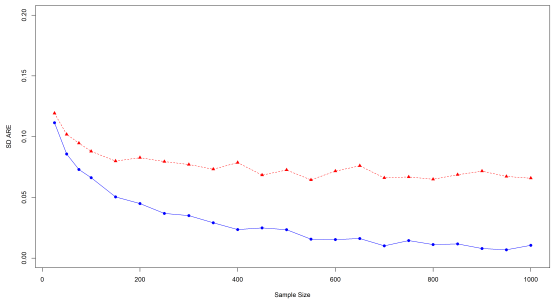


Log-normal (5, 1)

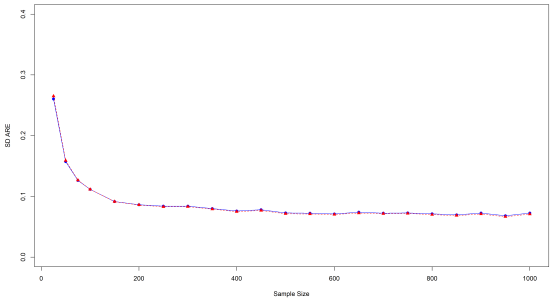
Figure 7: Comparisons against sample standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1), in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



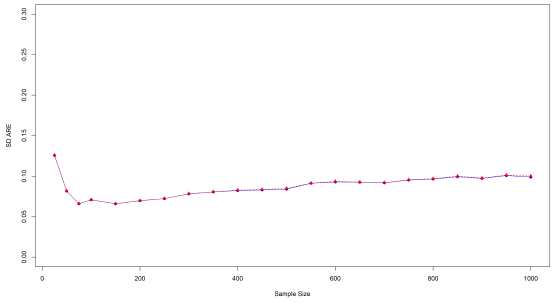
Normal (5, 1)



Normal (50, 17)



Gamma (2,5)

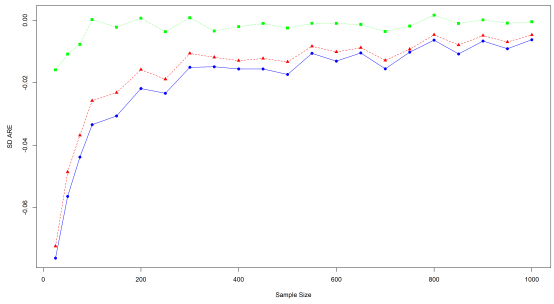


Weibull (2, 35)

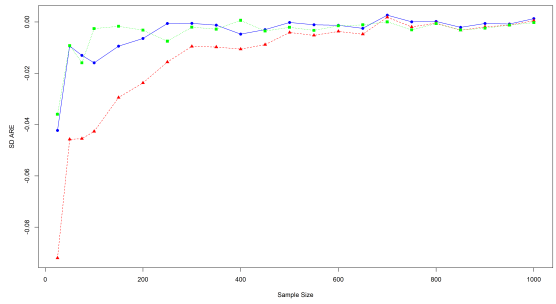
Figure 8: Comparisons against sample standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

2 ARE Comparisons (Without Absolute Value)

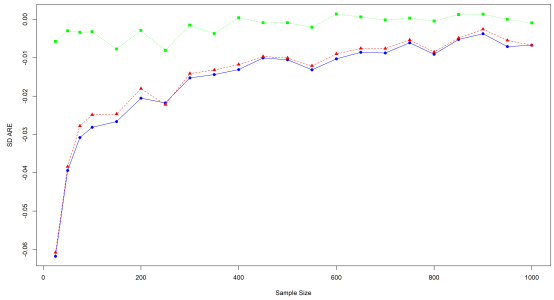
2.1 Against theoretical SD under Scenario S2



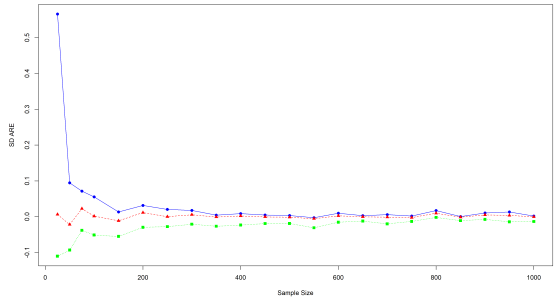
Log-normal (4, 0.3)



Log-normal (5, 0.5)

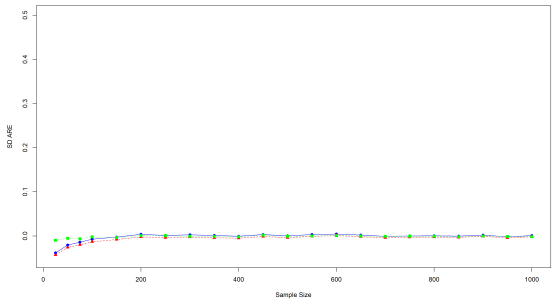


Log-normal (5, 0.25)

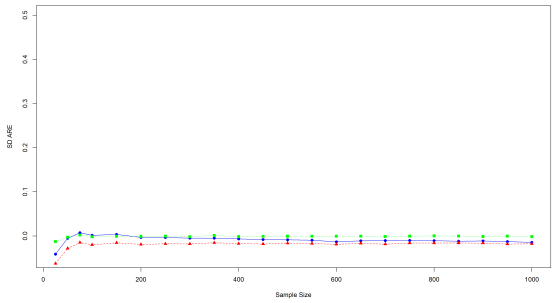


Log-normal (5, 1)

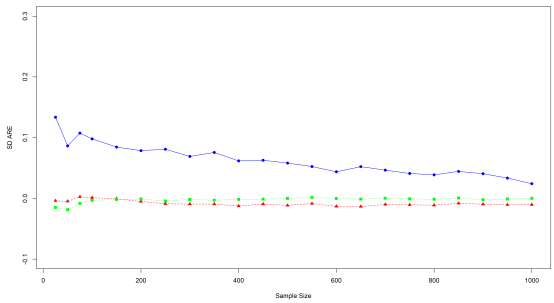
Figure 9: Comparisons against theoretical standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1) in scenario S2 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



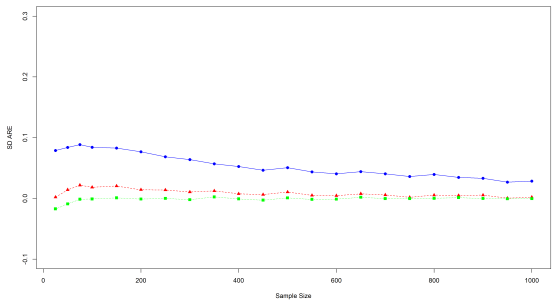
Normal (5,1)



Normal (50, 17)



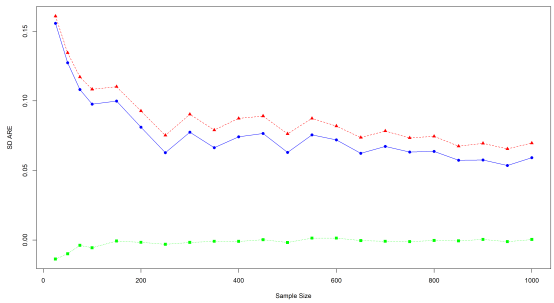
Gamma (2, 5)



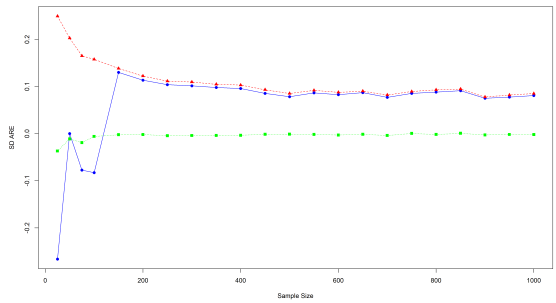
Weibull (2, 35)

Figure 10: Comparisons against theoretical standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S2 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

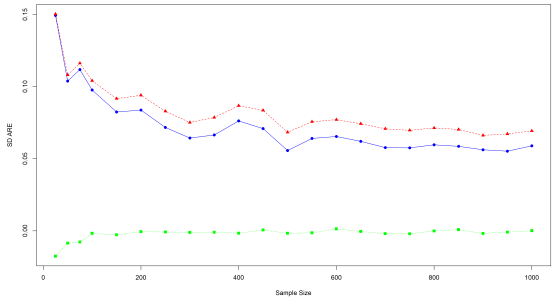
2.2 Against theoretical SD under Scenario S3



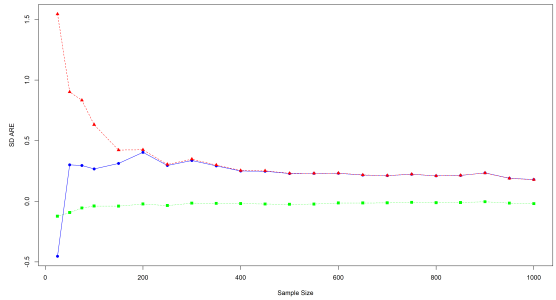
Log-normal (4, 0.3)



Log-normal (5, 0.5)

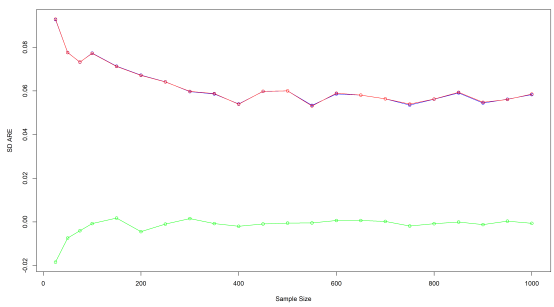


Log-normal (5, 0.25)

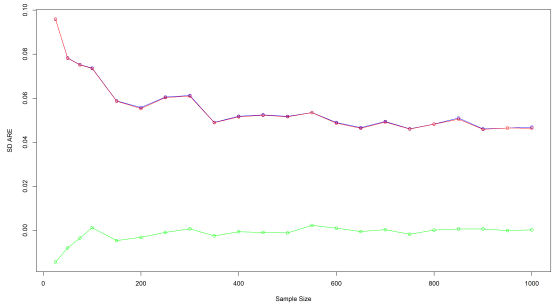


Log-normal (5, 1)

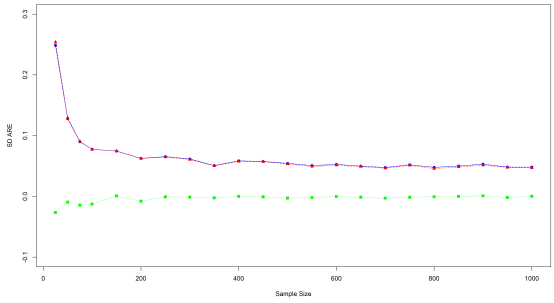
Figure 11: Comparisons against theoretical standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1) in scenario S2 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



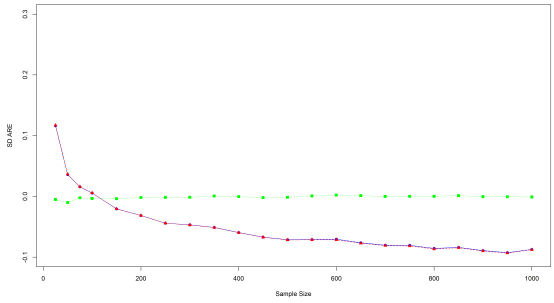
Normal (5,1)



Normal (50, 17)



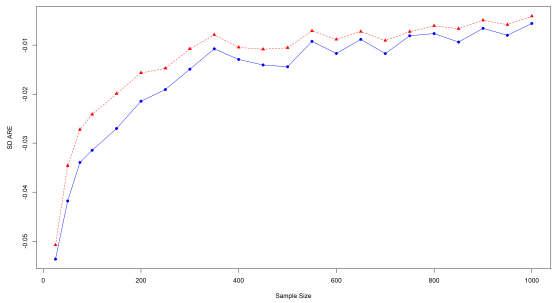
Gamma (2, 5)



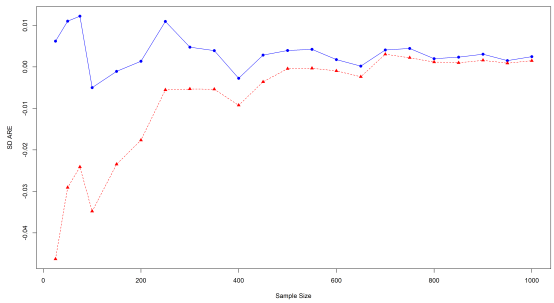
Weibull (2, 35)

Figure 12: Comparisons against theoretical standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S3 (MEQE - Red line, solid triangle, sample SD - green line, solid square, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

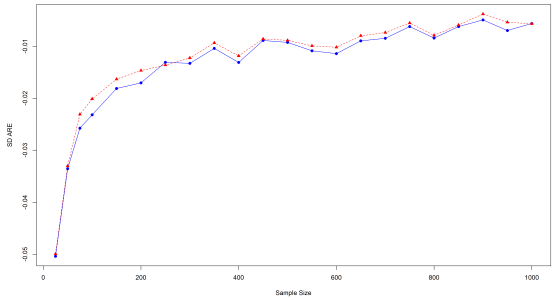
2.3 Against sample SD under Scenario S2



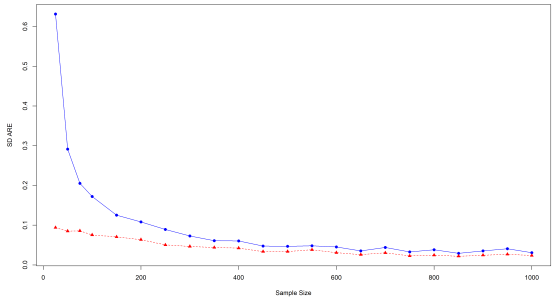
Log-normal (4, 0.3)



Log-normal (5, 0.5)

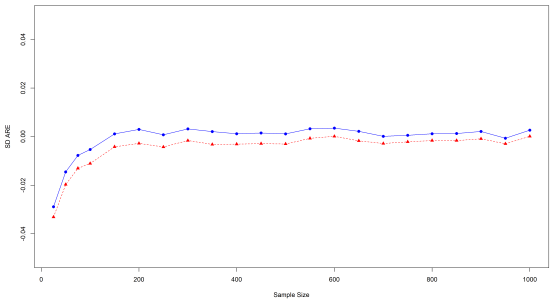


Log-normal (5, 0.25)

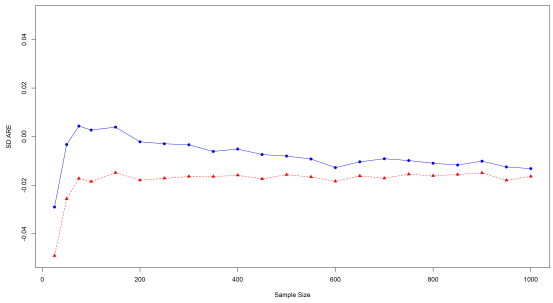


Log-normal (5, 1)

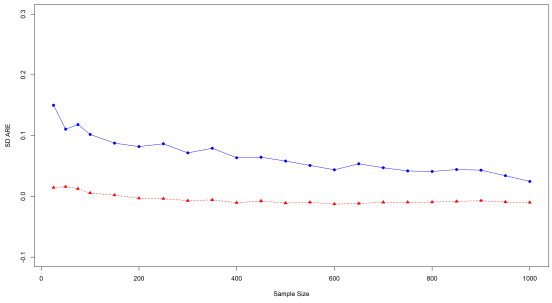
Figure 13: Comparisons against sample standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1) in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



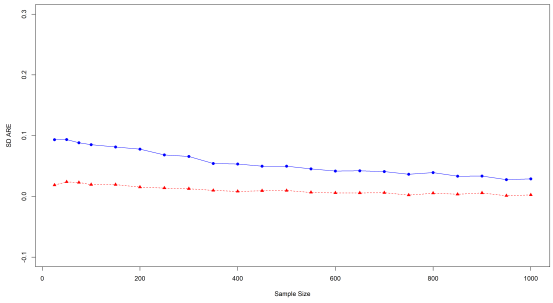
Normal (5, 1)



Normal (50, 17)



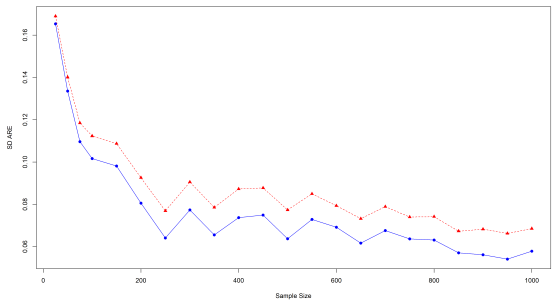
Gamma (2,5)



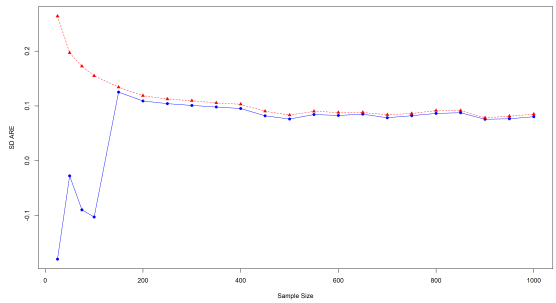
Weibull (2, 35)

Figure 14: Comparisons against sample standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the without SD ARE.

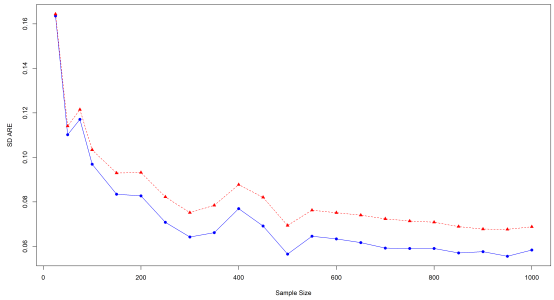
2.4 Against sample SD under Scenario S3



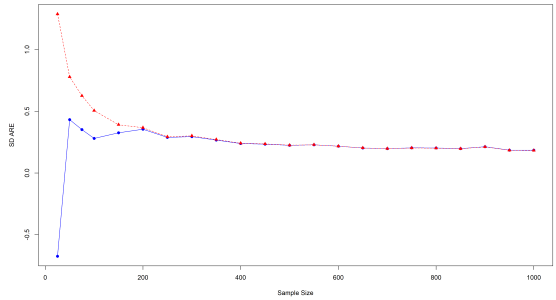
Log-normal (4, 0.3)



Log-normal (5, 0.5)

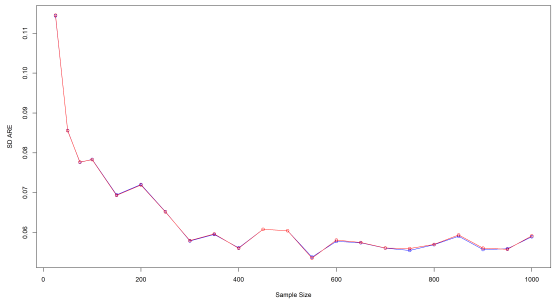


Log-normal (5, 0.25)

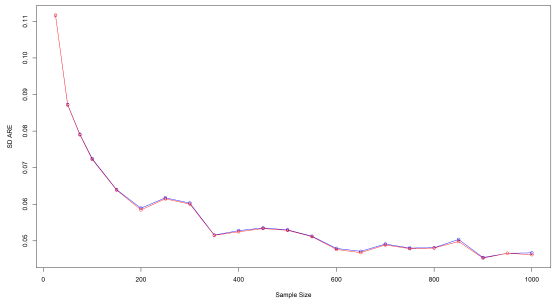


Log-normal (5, 1)

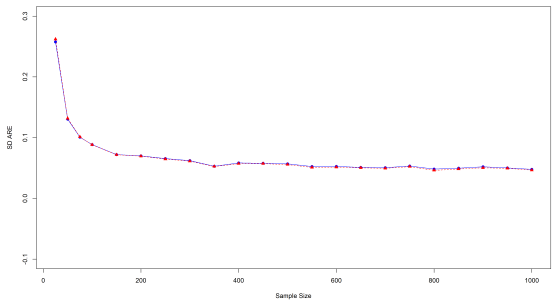
Figure 15: Comparisons against sample standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1), in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.



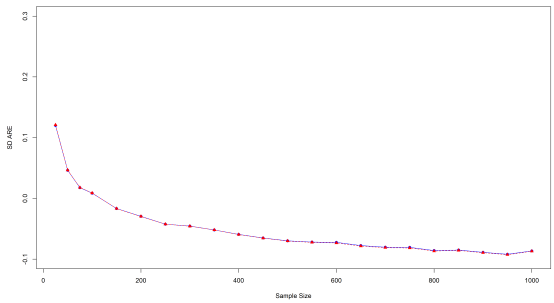
Normal (5, 1)



Normal (50, 17)



Gamma (2,5)



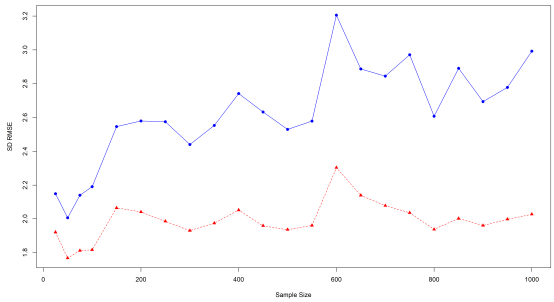
Weibull (2, 35)

Figure 16: Comparisons against sample standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD ARE.

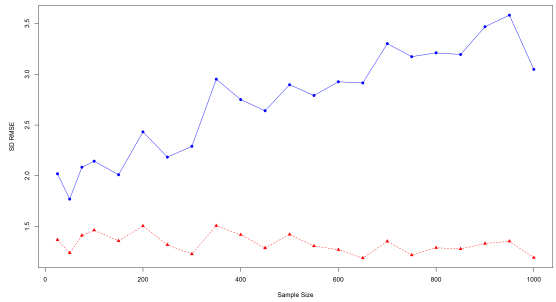
3 RMSE Comparisons

The results from both figures strongly indicate that the MEQE (red) method is the superior choice for reducing RMSE SD, providing more stable and reliable error estimates across various probability distributions. In contrast, the QE (blue) method exhibits high variability, making it less dependable, especially for distributions with greater variance.

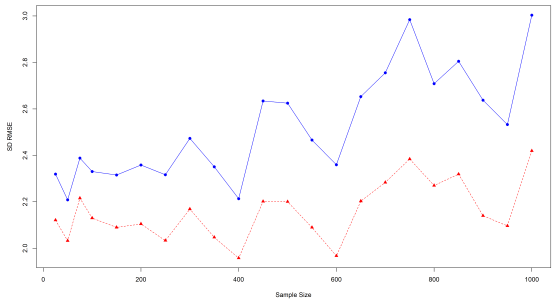
3.1 RMSE on S2



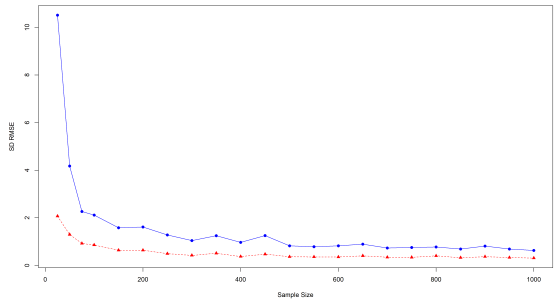
Log-normal (4, 0.3)



Log-normal (5, 0.5)

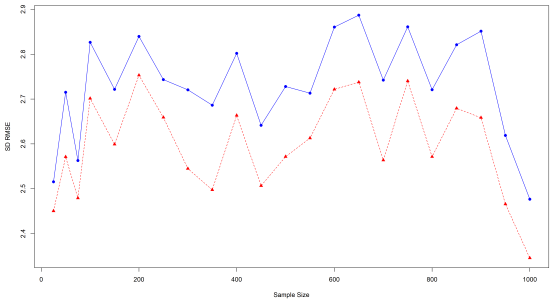


Log-normal (5, 0.25)

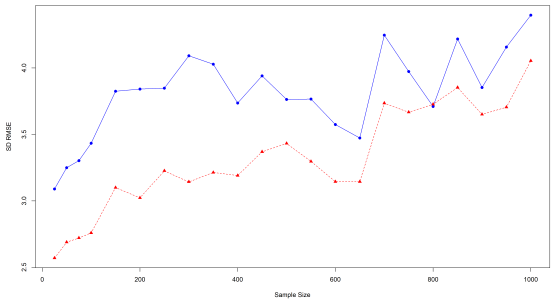


Log-normal (5, 1)

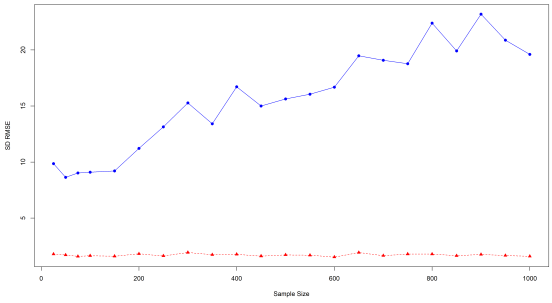
Figure 17: Comparisons RMSE standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1), in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD RMSE.



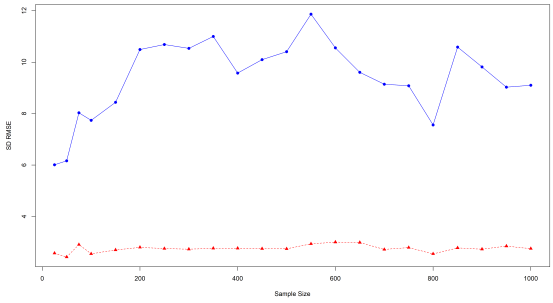
Normal (5, 1)



Normal (50, 17)



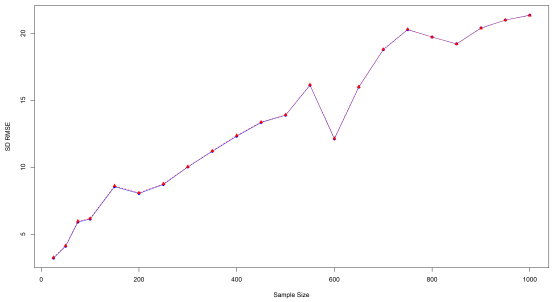
Gamma (2,5)



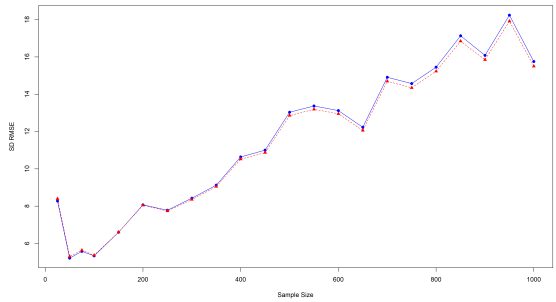
Weibull (2, 35)

Figure 18: Comparisons RMSE standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S2 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD RMSE.

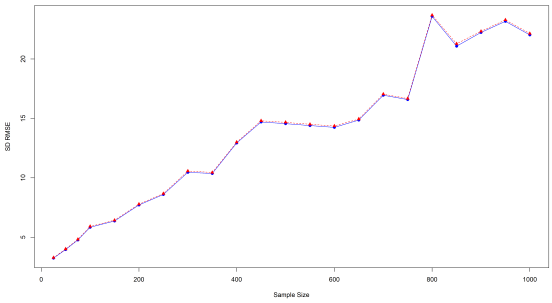
3.2 RMSE on S3



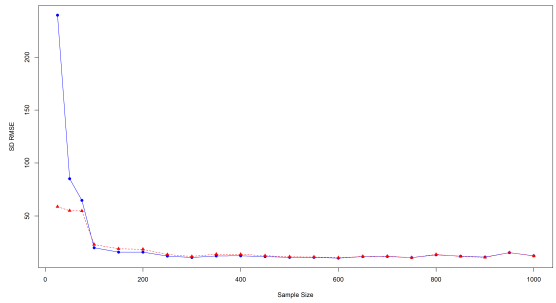
Log-normal (4, 0.3)



Log-normal (5, 0.5)

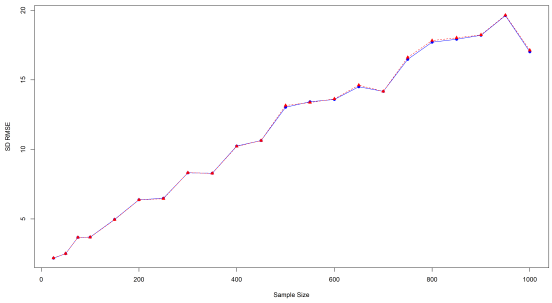


Log-normal (5, 0.25)

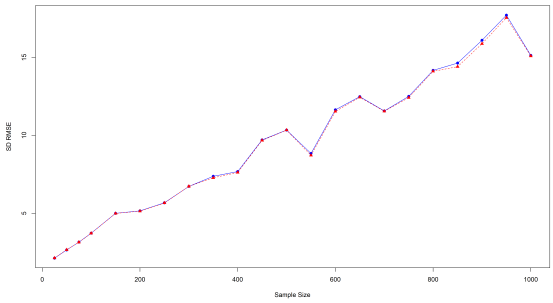


Log-normal (5, 1)

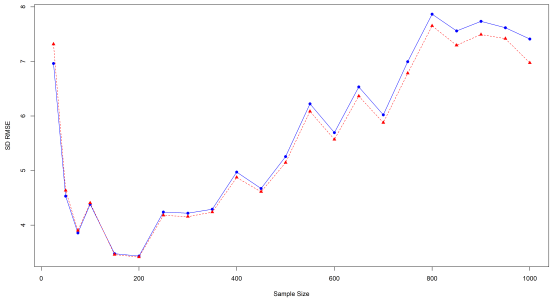
Figure 19: Comparisons RMSE standard deviation for log-normal distribution with parameters (4,0.3), (5,0.5), (5,0.25) and (5,1), in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD RMSE.



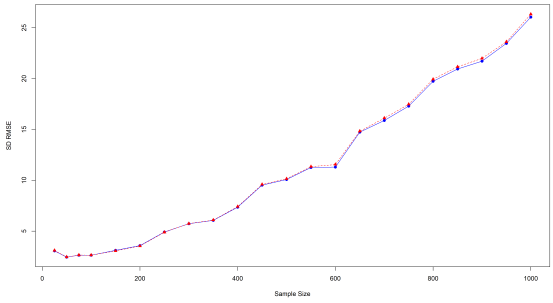
Normal (5, 1)



Normal (50, 17)



Gamma (2,5)



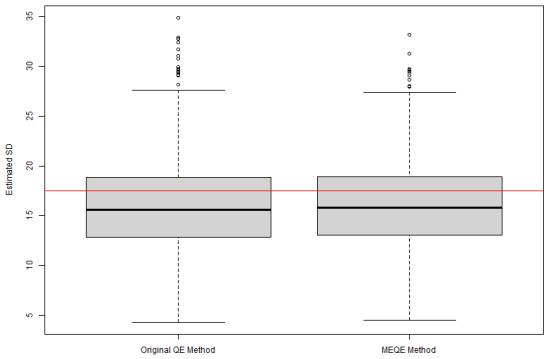
Weibull (2, 35)

Figure 20: Comparisons RMSE standard deviation for normal distribution with parameters (5,1), (50,17), gamma and weibull distributions with parameters (2,5) and (2,35), respectively, in scenario S3 (MEQE - Red line, solid triangle, QE - blue line, solid circle). The x-axis represents sample size, while the y-axis represents the SD RMSE.

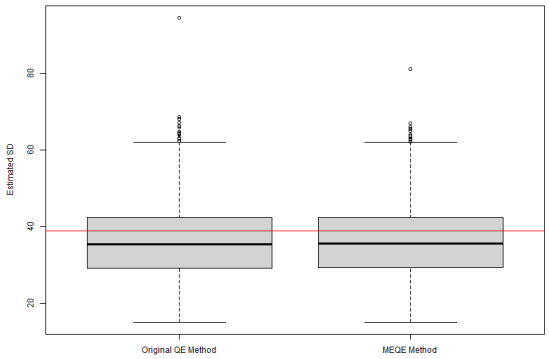
4 Box-Plot Comparisons

In this section, a (Red line, solid triangle) displaying the theoretical background value follows the fixed n, box plot of the estimation for different sizes.

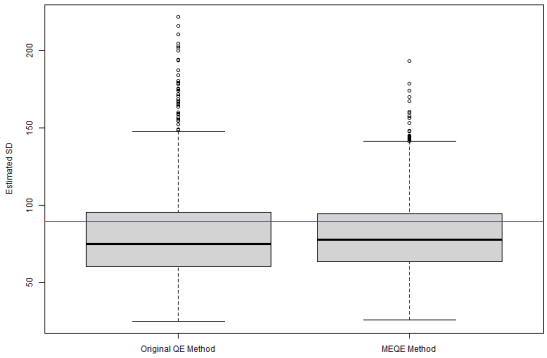
4.1 Box Plot comparison for Scenario S2



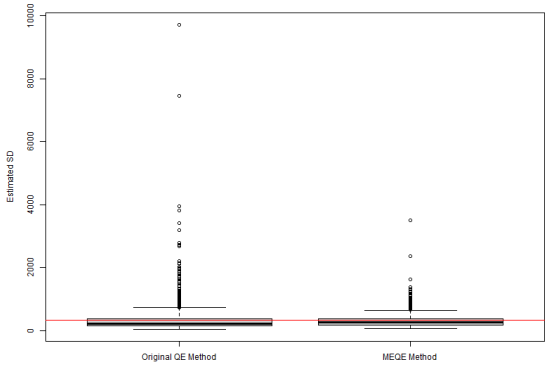
Log-normal (4, 0.3)



Log-normal (5, 0.5)

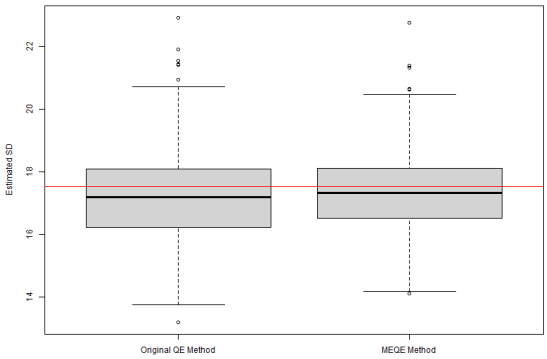


Log-normal (5, 0.25)

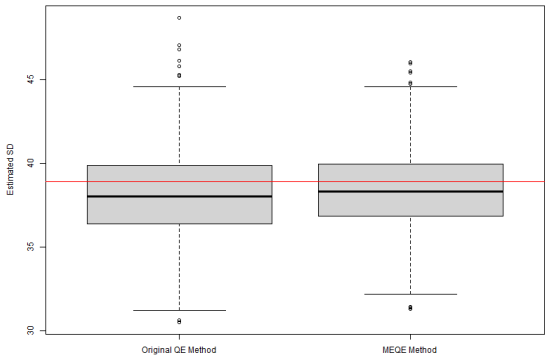


Log-normal (5, 1)

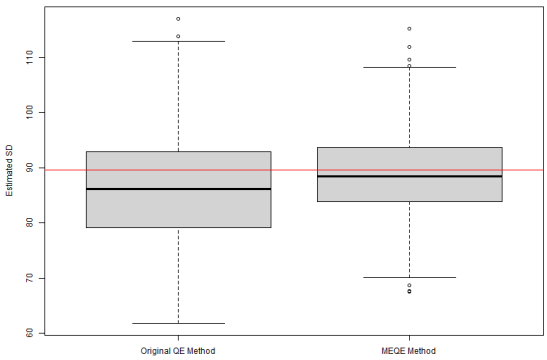
Figure 21: Comparison of SD estimation methods for sample size 25 in Scenario S2 across log-normal distribution.



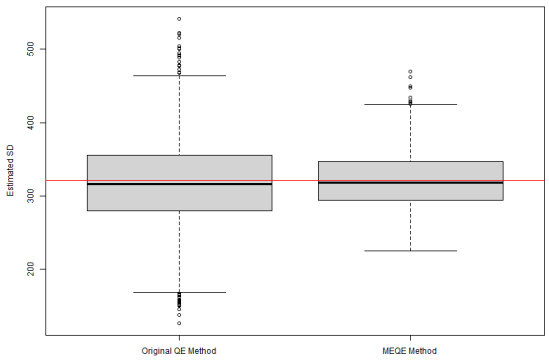
Log-normal (4, 0.3)



Log-normal (5, 0.5)

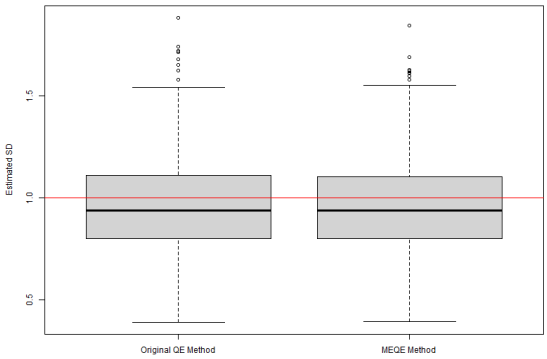


Log-normal (5, 0.25)

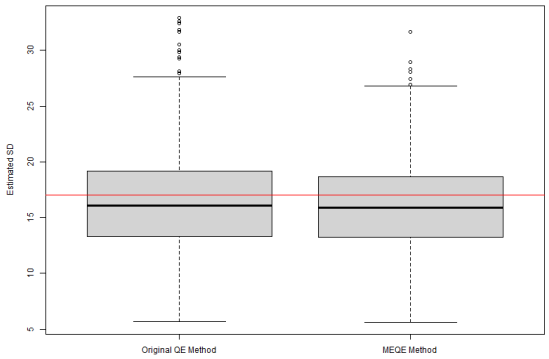


Log-normal (5, 1)

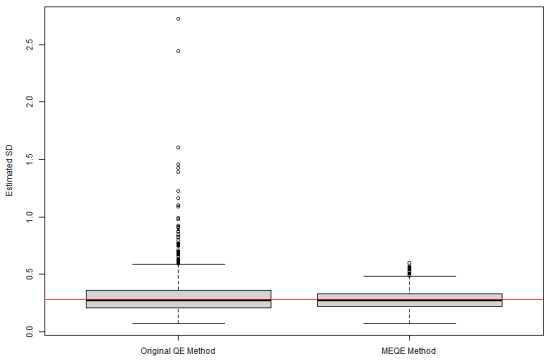
Figure 22: Comparison of SD estimation methods for sample size 400 in Scenario S2 across log-normal distribution.



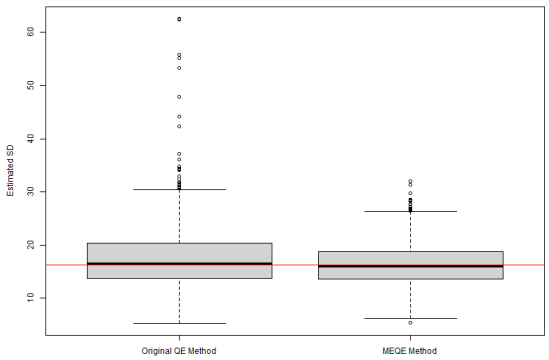
Normal (5, 1)



Normal (50, 17)

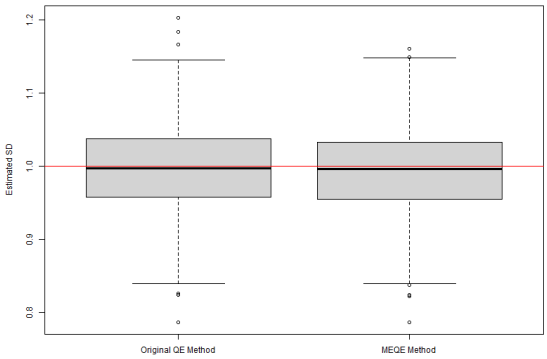


Gamma (2, 5)

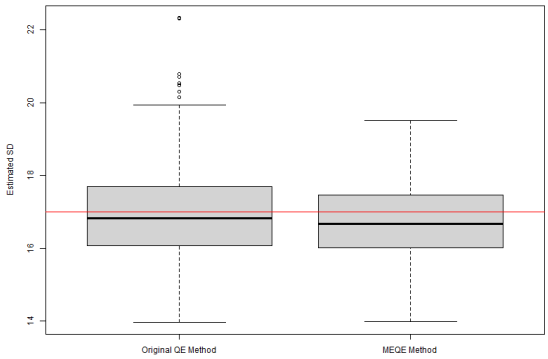


Weibull (2, 35)

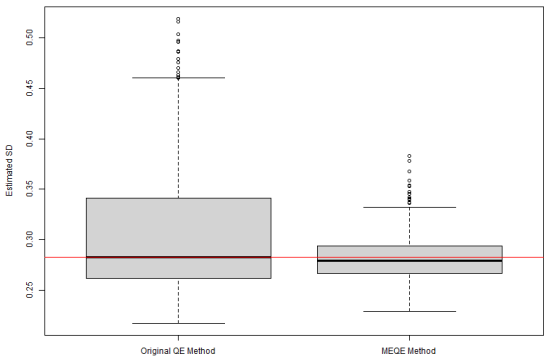
Figure 23: Comparison of SD estimation methods for sample size 25 in Scenario S2 across normal, gamma and weibull distributions.



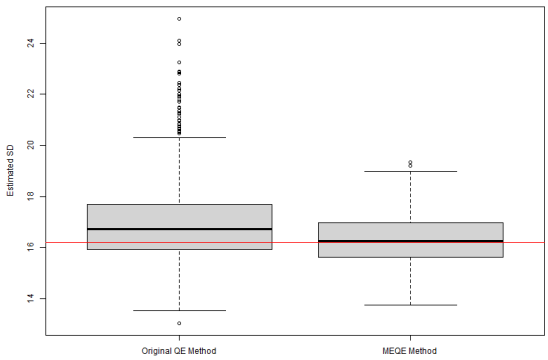
Normal (5, 1)



Normal (50, 17)



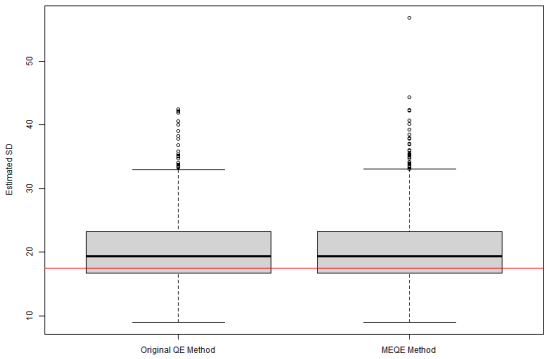
Gamma (2, 5)



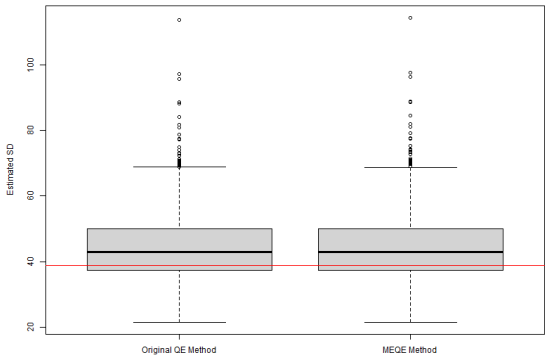
Weibull (2, 35)

Figure 24: Comparison of SD estimation methods for sample size 400 in Scenario S2 across normal, gamma and weibull distributions.

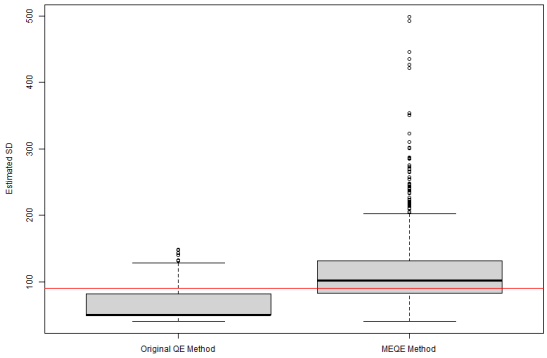
4.2 Box Plot comparison for Scenario S3



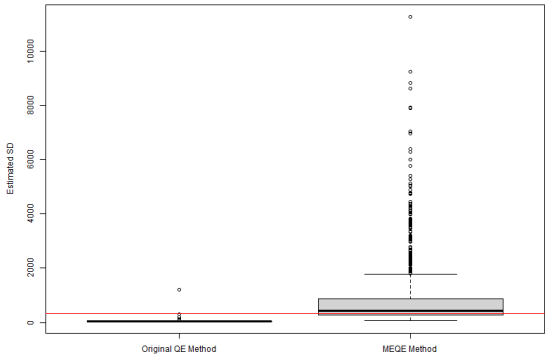
Log-normal (4, 0.3)



Log-normal (5, 0.5)

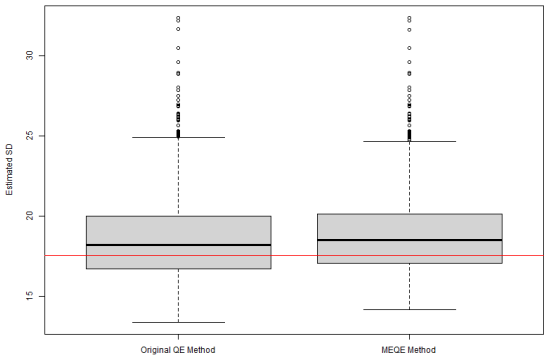


Log-normal (5, 0.25)

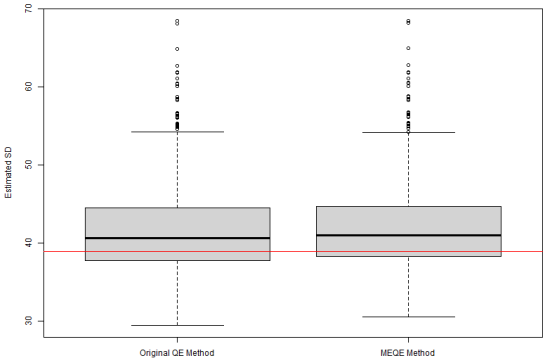


Log-normal (5, 1)

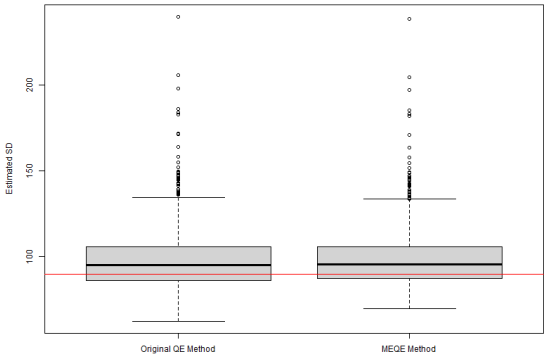
Figure 25: Comparison of SD estimation methods for sample size 25 in Scenario S3 across log-normal distribution.



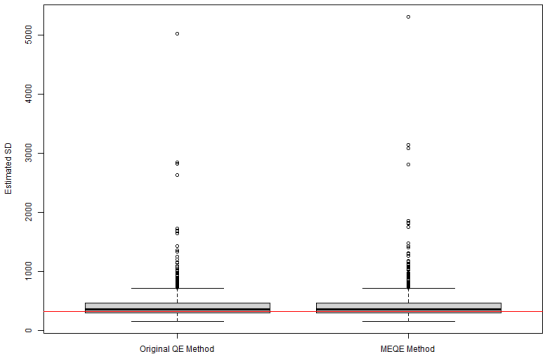
Log-normal (4, 0.3)



Log-normal (5, 0.5)

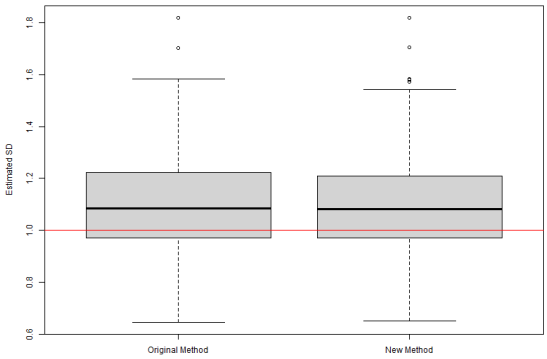


Log-normal (5, 0.25)

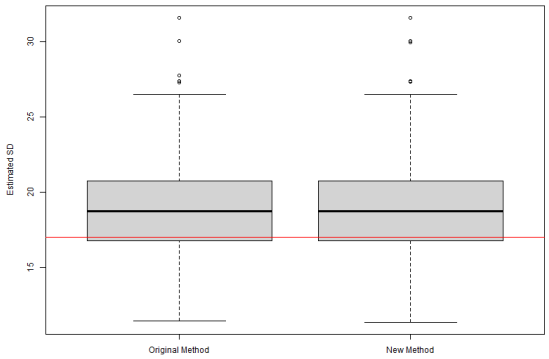


Log-normal (5, 1)

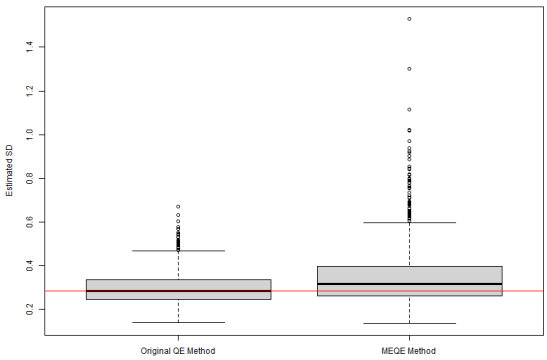
Figure 26: Comparison of SD estimation methods for sample size 400 in Scenario S3 across log-normal distribution.



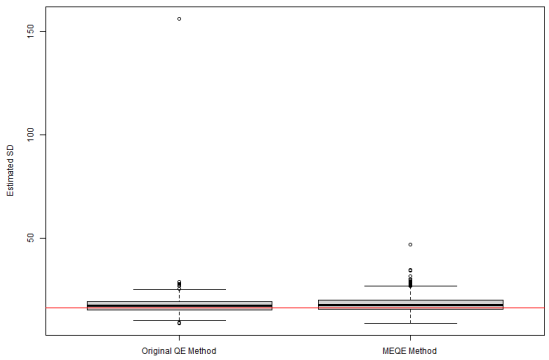
Normal (5, 1)



Normal (50, 17)

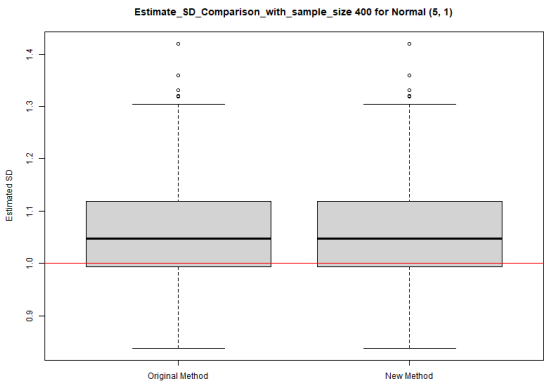


Gamma (2, 5)

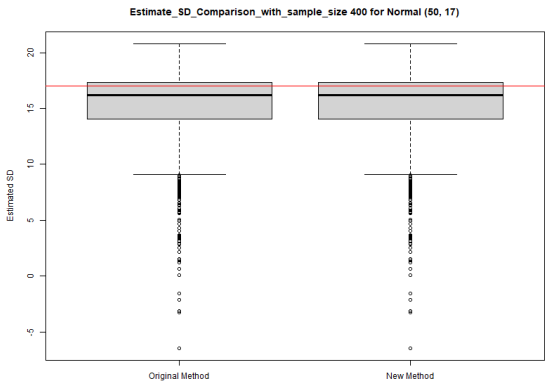


Weibull (2, 35)

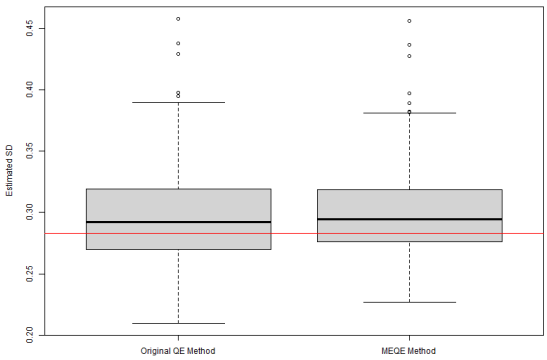
Figure 27: Comparison of SD estimation methods for sample size 25 in Scenario S3 across normal, gamma and weibull distributions.



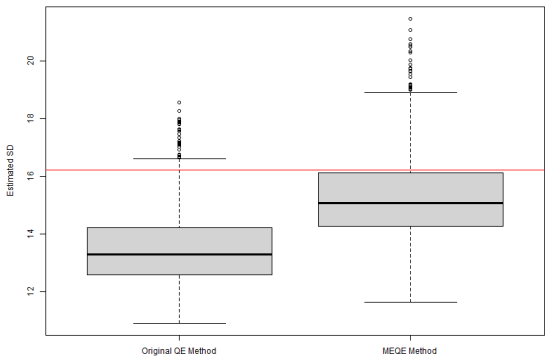
Normal (5, 1)



Normal (50, 17)



Gamma (2, 5)

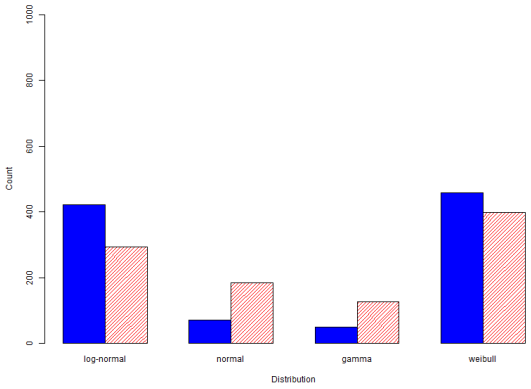


Weibull (2, 35)

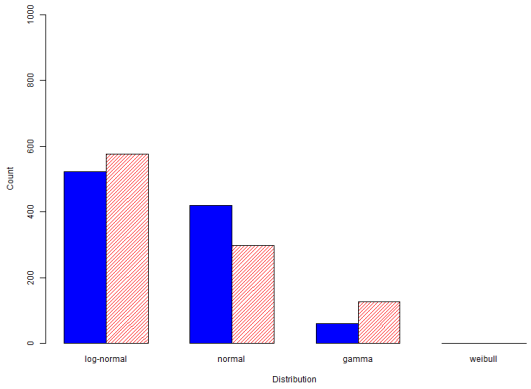
Figure 28: Comparison of SD estimation methods for sample size 400 in Scenario S3 across normal, gamma and weibull distributions.

5 Comparing Distribution Detection Abilities

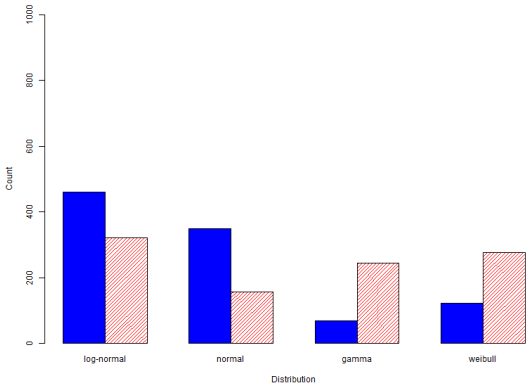
In this section, shows compression the MEQE and Original QE techniques for a different sample sizes 25 and 400, for selecting the best-fit distribution on S2.



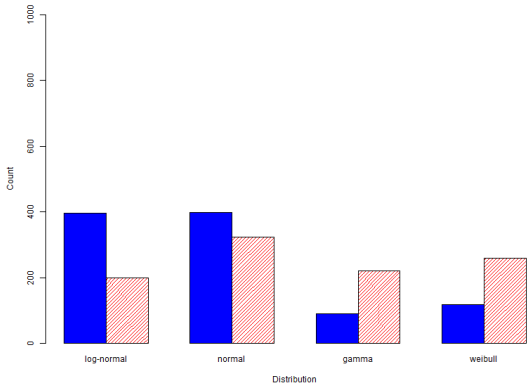
Normal (5, 1)



Log-normal (5, 0.25)

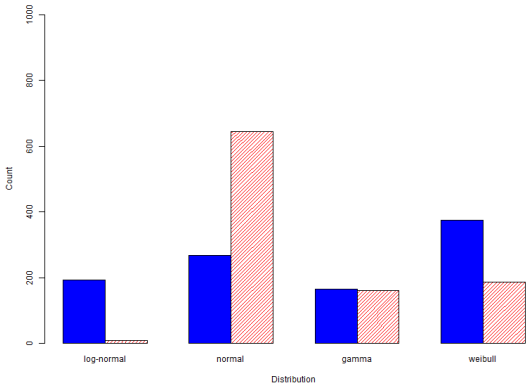


Gamma (2, 5)

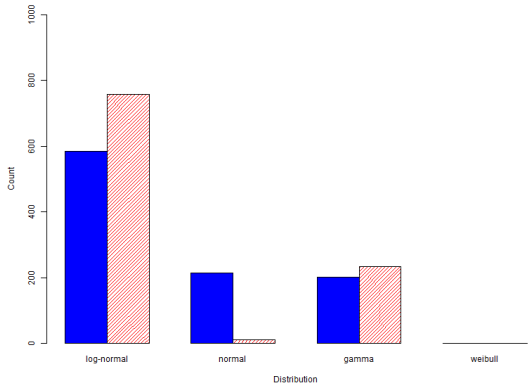


Weibull (2, 35)

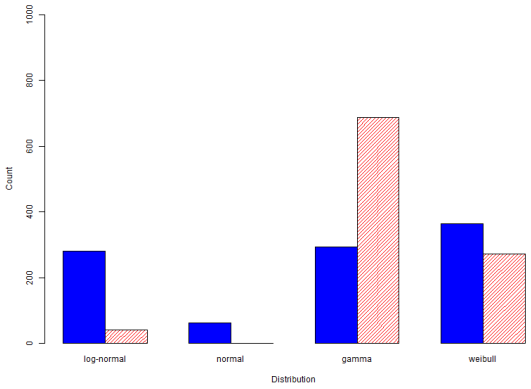
Figure 29: Comparisons of distribution selection counts between the MEQE method and the original QE method for sample size 25 across different data distributions. (MEQE - red, dashed line, QE - blue, solid line).



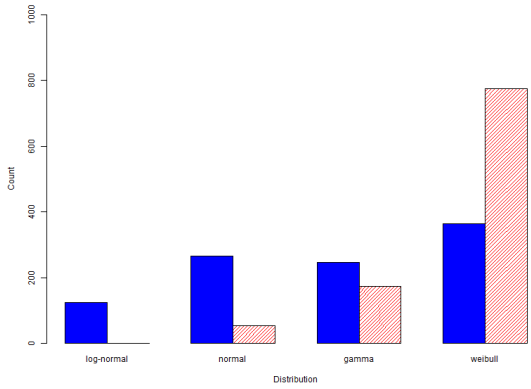
Normal (5, 1)



Log-normal (5, 0.25)



Gamma (2, 5)



Weibull (2, 35)

Figure 30: Comparisons of distribution selection counts between the MEQE method and the original QE method for sample size 400 across different data distributions. (MEQE - red, dashed line, QE - blue, solid line).

6 Scenarios between S2 and S3

In this section, between S2 and S3 we define scenarios S2_min_extended and S2_max_extended:

$$S2\_min\_extended = \{Q_{\min}, Q_1, Q_2, Q_3, n\} \tag{1}$$

$$S2\_max\_extended = \{Q_1, Q_2, Q_3, Q_{\max}, n\} \tag{2}$$

In the simulation, we used the log-normal distribution with parameters (5, 0.25) and compared the performance of the MEQE method across all the four scenarios.

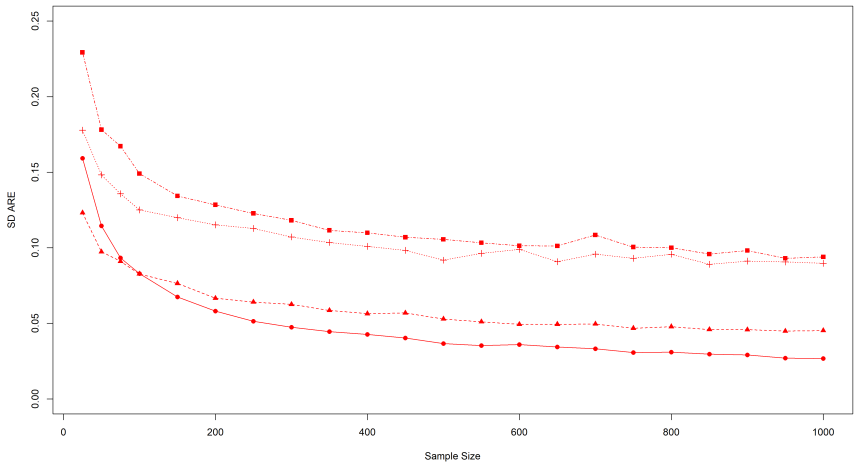


Figure 31: Absolute ARE comparisons of SD estimates of MEQE method across different scenarios (S2 - solid line, circles, S2_min_extended - dashed line, filled triangles, S2_max_extended - dotted line, plus signs, S3 - dot-dash line, filled squares). The x-axis represents sample size, while the y-axis represents the SD ARE.

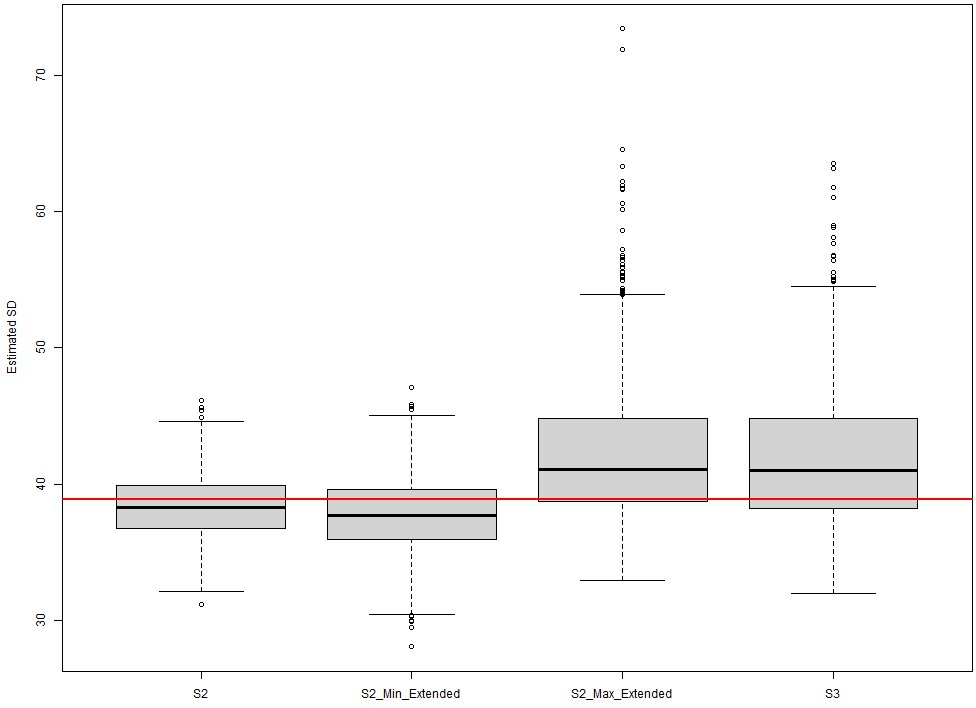


Figure 32: Comparison of the MEQE estimates under different scenarios when the sample size is 400.