

POLS 602 Homework 4

Nicholas Ray

2023-11-07

Conduct a simulation to evaluate the effects of serial correlation (autocorrelation in the residuals) on a regression model. More specifically, simulate the model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ where $\beta_0 = 1$ and $\beta_1 = 2$. Use the function `stats::arima.sim(list(ar=p),N)` to generate ε_i with various levels of autocorrelation such that $\rho \in (0, 0.25, 0.5, 0.75, 0.99)$. For each level of ρ , conduct 100 simulations with 1000 observations for each model. Answer the following two questions using your results:

1. How does $\hat{\beta}_1$ vary across values of ρ ? Compare bias across different values of ρ [i.e., $mean(\hat{\beta}_1) - \beta_1$].
2. How does $se(\hat{\beta}_1)$ vary across values of ρ ? Compare the standard deviations of your estimates to their standard errors across different values of ρ [i.e., $\frac{sd(\hat{\beta}_1)}{mean(se(\hat{\beta}_1))}$]. If this ratio is bigger (smaller) than 1, the standard errors underestimate (overestimate) the true variability of the estimates.

The effect of serial correlation on standard errors.

500 simulated models each with 1000 observations.

Points represent the standard errors of the slope estimate from each model.

The line represents the standard deviation of the slope estimate for each value of $AR(p)$.

