We can use Theorem A of Section 4.4.1 to find $E(\hat{\boldsymbol{\beta}})$ under the new model:

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= E(E(\hat{\boldsymbol{\beta}}|\boldsymbol{\Xi})) \\ &= E(\boldsymbol{\beta}) \\ &= \boldsymbol{\beta} \end{aligned}$$

where the outer expectation is with respect to the distribution of $\boldsymbol{\Xi}$. The least squares estimate is thus unbiased under the new model as well.

Next we consider the variance of the least squares estimate. From Theorem B of Section 14.2.2, $\mathrm{Var}(\hat{\beta}_i|\boldsymbol{\Xi} = \mathbf{X}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}_{ii}$. This is the conditional variance. To find the unconditional variance we can use Theorem B of Section 4.4.1, according to which

$$\begin{aligned} \mathrm{Var}(\hat{\beta}_i) &= \mathrm{Var}\left(E(\hat{\beta}_i|\boldsymbol{\Xi})\right) + E\left(\mathrm{Var}(\hat{\beta}_i|\boldsymbol{\Xi})\right) \\ &= \mathrm{Var}(\beta_i) + E\left(\sigma^2(\boldsymbol{\Xi}^T\boldsymbol{\Xi})^{-1}_{ii}\right) \\ &= \sigma^2 E(\boldsymbol{\Xi}^T\boldsymbol{\Xi})^{-1}_{ii} \end{aligned}$$

This is a highly nonlinear function of the random vectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n$ and would generally be difficult to evaluate analytically.

Thus for the new, unconditional model, the least squares estimates are still unbiased, but their variances (and covariances) are different. Surprisingly, it turns out that the confidence intervals we have developed still hold at their nominal levels of coverage. Let $C(\mathbf{X})$ denote the $100(1-\alpha)\%$ confidence interval for β_j that we developed under the old model. Using I_A to denote the indicator variable of the event A, we can express the fact that this is a $100(1-\alpha)\%$ confidence interval as

$$E\left(I_{\{\beta_j \in C(\mathbf{X})\}}|\boldsymbol{\Xi} = \mathbf{X}\right) = 1 - \alpha$$

that is, the conditional probability of coverage is $1-\alpha$. Because the conditional probability of coverage is the same for every value of $\boldsymbol{\Xi}$, the unconditional probability of coverage is also $1-\alpha$:

$$\begin{aligned} EI_{\{\beta_j \in C(\Xi)\}} &= E\left(E(I_{\{\beta_j \in C(\Xi)\}}|\boldsymbol{\Xi})\right) \\ &= E(1-\alpha) \\ &= 1-\alpha \end{aligned}$$

This very useful result says that for forming confidence intervals we can use the old fixed-**X** model and that the intervals we thus form have the correct coverage in the new random-**X** model as well.

We complete this section by discussing how the bootstrap can be used to estimate the variability of a parameter estimate under the new model according to which the parameter estimate, say $\hat{\theta}$, is based on n i.i.d. random vectors $(Y_1, \xi_1), (Y_2, \xi_2), \ldots, (Y_n, \xi_n)$. Depending on the context, there are a variety of parameters θ that might be of interest. For example, θ could be one of the regression coefficients, β_i; θ could be $E(Y|\xi = \mathbf{x}_0)$, the expected response at a fixed level $\mathbf{x}_0$ of the independent variables (see Problem 13); in simple linear regression, θ could be that value x_0 such that $E(Y|\xi = x_0) = \mu_0$ for some fixed μ_0; in simple linear regression, θ could be the correlation coefficient of Y and ξ. Now if we knew the probability distribution of the random vector (Y, ξ), we could simulate the sampling distribution of the parameter