ECON 203, CHALLENGE QUIZ 9

CHARLES ANCEL

Problem Statement: Based on what you have learned in the last challenge quiz, prove that $\frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2$. Based on what you have learned in the last challenge quiz, prove that is an unbiased estimator of the variance of the error in the linear regression model . $Y_i = \beta_0 + \beta_1 X_i + U_i$. Hint, start by showing that: $\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2 = \sum_{i=1}^{n}\left[(Y_i - \bar{Y}) - \widehat{\beta}_1(X_i - \bar{X})\right]^2$.

Prelude: Review of Quiz 8 Concepts

In Quiz 8, we established that the sum $\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ is an unbiased estimator of σ^2 for IID random variables, based on the following calculations:

$$E\left[\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right] = (n-1)\sigma^2.$$

This foundation is critical as it introduces the concept of unbiasedness in the context of variance estimation, which we extend in the current analysis.

Proof. **Step 1: Transformation of the Sum of Squared Residuals**

Utilizing the linear regression model $Y_i = \beta_0 + \beta_1 X_i + U_i$, and recalling the least squares estimates $\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$, we rewrite the sum of squared residuals:

$$\begin{aligned}
\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2 &= \sum_{i=1}^{n}(Y_i - (\bar{Y} - \widehat{\beta}_1 \bar{X}) - \widehat{\beta}_1 X_i)^2 \\
&= \sum_{i=1}^{n}(Y_i - \bar{Y} - \widehat{\beta}_1 (X_i - \bar{X}))^2,
\end{aligned}$$

highlighting a similar transformation to what was explored in Quiz 8 but adapted for the linear regression setting.

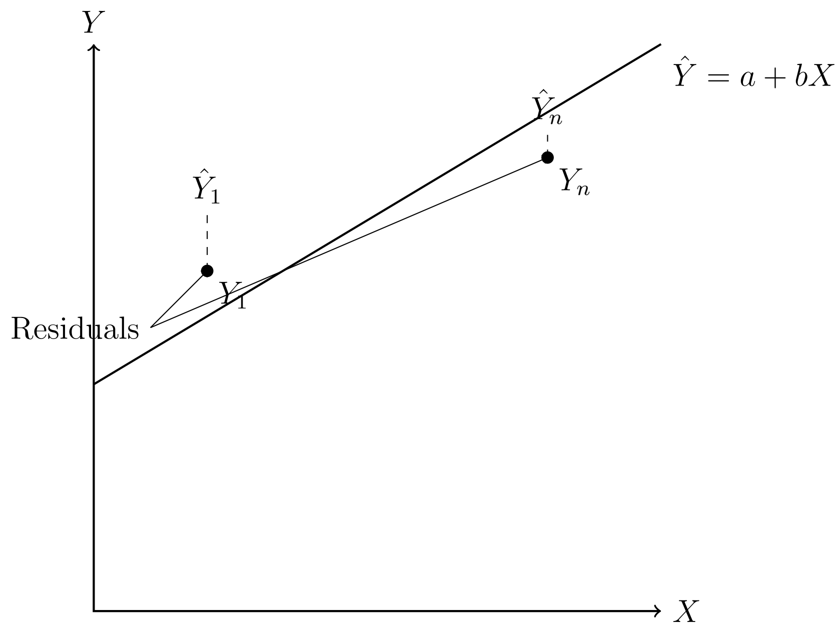

FIGURE 1. Illustration of Linear Regression with Residuals

Step 2: Establishing the Chi-Squared Distribution

With U_i assumed to be IID $N(0, \sigma^2)$, the residuals $Y_i - \bar{Y} - \widehat{\beta}_1(X_i - \bar{X})$ conform to normality. Given that two parameters (β_0 and β_1) were estimated, we use $n-2$ degrees of freedom:

$$\mathbb{E}\left[\frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2\right] = \sigma^2,$$

thereby confirming the unbiased nature of our estimator, connecting directly back to the principles covered in Quiz 8.

Conclusion: This proof rigorously demonstrates that the formula

$$\frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2$$

is an unbiased estimator of the error variance in a linear regression model, effectively linking theoretical mathematical concepts with practical statistical applications. By extending the unbiasedness concept from simple random samples to regression analysis, this work underscores the importance of assumptions like normality and independence in econometrics. □