2025/09/10
藪友良

不均一分散に頑健な標準誤差（ロバスト標準誤差）

均一分散は、ガウス=マルコフ定理の前提条件の 1 つであり、数式展開を簡単にしてくれる仮定でもあります。しかし、均一分散は、現実には成立しないことが多く、不均一分散が適切な仮定になります。本稿では、不均一分散を前提として、不均一分散に頑健な標準誤差(ロバスト標準誤差)を説明します。なお、本稿は、藪友良『入門 実践する計量経済学』(2023 年、東洋経済新報社)の補足資料になります。

1.　OLS 推定量の分散

ここでは単回帰モデルを考えます。

$$Y_i = \alpha + \beta X_i + u_i$$

横断面データであるため、各変数の右下添字はiとしています(サンプルサイズはn)。ここで、誤差項は不均一分散とします。

$$V(u_i) = E[u_i^2] = \sigma_i^2$$

誤差項の分散は、σ^2ではなく、σ_i^2としていることに注意してください。

このとき、OLS 推定量$\hat{\beta}$の分散$\sigma_{\hat{\beta}}^2$は、均一分散のもとで導出された

$$\frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

ではなく、不均一分散を考慮することで得られた次式です(証明は 9 章補足参照)。

$$\sigma_{\hat{\beta}}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sigma_i^2}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

次節では、分散$\sigma_{\hat{\beta}}^2$を推定する方法を説明します。

2.　$\sigma_{\hat{\beta}}^2$の推定方法

OLS 推定量$\hat{\beta}$の分散$\sigma_{\hat{\beta}}^2$は、次のようにも表現できます。$q_i = (X_i - \bar{X})^2 \sigma_i^2$と定義すると、

$$\sigma_{\hat{\beta}}^2 = \frac{\sum_{i=1}^{n} q_i}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

$$= \frac{n\bar{q}}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

ただし、$\bar{q}$は平均$\bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i$と定義しています。ここで、誤差項の分散$\sigma_i^2$は未知であるため、データから推定する必要があります。

誤差項の分散は$\sigma_i^2 = E[u_i^2]$ですから、σ_i^2の推定量として、誤差項の 2 乗u_i^2を用いることが自然でしょう。しかし、誤差項u_iは観察できないため、誤差項u_iの推定量である残差$\hat{u}_i$に置き換えることで、$q_i = (X_i - \bar{X})^2 \sigma_i^2$の不偏推定量$\hat{q}_i$が得られます。

$$\hat{q}_i = (X_i - \bar{X})^2 \hat{u}_i^2$$

しかし、$\hat{q}_i$は1つの観測値$\hat{u}_i^2$からσ_i^2を推定しており、その推定精度は低くなります。

H.ホワイト(Halbert White)は、個々のq_iではなく、その平均$\bar{q}$なら安定的に推定できるとしました[1]。つまり、$\hat{q}_i$は推定誤差が大きい一方、サンプルサイズnが十分大きいなら、その平均である

$$\hat{\bar{q}} = \frac{1}{n}\sum_{i=1}^{n}\hat{q}_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\hat{u}_i^2$$

は推定誤差が打ち消しあうことで正確に推定できるというわけです。

$\hat{\bar{q}}$を先の分散$\sigma_{\hat{\beta}}^2$の式に代入すると、**不均一分散に頑健な分散の推定量**が得られます(教科書9.3節を参照してください)[2]。

$$s_{\hat{\beta}}^2 = \frac{n\hat{\bar{q}}}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2\hat{u}_i^2}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

そして、分散の推定量$s_{\hat{\beta}}^2$の平方根である標準誤差は、**不均一分散に対して頑健な標準誤差**(heteroskedasticity-consistent standard error、**HC 標準誤差**)、**ロバスト標準誤差**(robust standard error)、または、開発者の名前にちなんで**ホワイト標準誤差**などと呼ばれます。

$$s_{\hat{\beta}} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2\hat{u}_i^2}{\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}}$$

ロバスト標準誤差は、通常の標準誤差より少し大きくなることが多いですが、小さくなることもあります。ロバスト標準誤差は、通常の標準誤差に大きな違いがあるようなら、何らかの問題が生じているかもしれません。

この点について、有名な計量経済学の教科書『Mostly Harmless Econometrics』(Joshua Angrist, Jorn-Steffen Pischke)で次のように書かれています。「もしロバスト標準誤差が、通常の標準誤差より 30%以上大きい、もしくは、著しく小さい場合、プログラミングエラーや他の問題を疑うべきだろう。たとえば、ロバスト標準誤差が、通常の標準誤差を下回る場合として、サンプルサイズが小さいことによるバイアスが考えられる」。

ロバスト標準誤差が、通常の標準誤差よりも小さくなる場合に関心がある方は、『入門実践する計量経済学』9章練習問題 10、11 を参照してください。

最後に、統計ソフトでは、OLS 推定を実行すると通常の標準誤差が算出されます。しか

---

[1] White, Halbert. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." Econometrica: journal of the Econometric Society (1980): 817-838.

[2] $\sqrt{n}\hat{\beta}$の分散は次のようになります。

$$\frac{\bar{q}}{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

これは次のように推定できます。

$$\frac{\hat{\bar{q}}}{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2}$$

サンプルサイズnが大きくなると、分子は$\bar{q} = \frac{1}{n}\sum_{i=1}^{n}q_i$に収束し、$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$は$X$の分散に収束します。

し、不均一分散が自然な前提であることから、統計ソフトのオプションを変更するなどし、ロバスト標準誤差を算出するようにしてください[3]。

H. ホワイトについて

ホワイトが不均一分散に頑健な標準誤差を提案した論文は、Google Scholar の引用回数が 36,218 回に達しています（2025 年 7 月 17 日時点）。経済分野では引用回数が 1000 回を超えたら有名論文と言ってよいでしょうから、この論文の影響力の強さがわかります。ホワイトは 2012 年に亡くなっていますが、存命であれば、ノーベル経済学賞を取得する可能性が高い計量経済学者といえます。

彼はコンサルティング会社 Bates White の共同創業者としても有名です。この会社では、経済分析を用いて、様々な訴訟案件に専門的知見を提供しています。訴訟案件は多岐にわたり、たとえば、契約紛争、損害賠償、エネルギー資産をめぐる紛争などがあります。余談ですが、私の知人の経済学者も PhD 取得後に、この会社で専門的知見を活かした仕事をしています。経済学者が起業することも多々見られるようになっていますが、ホワイトはその先駆けとなった 1 人でした。

---

[3] たとえば、Stata では、次のように最後に r を入れるとロバスト標準誤差が表示されます。

```
reg Y X, r
```