Az OLS becslő modellfeltevései és az OLS szolgáltatta becslések statisztikai tulajdonságai

Ferenci Tamás
tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 12.

Tartalom

1. Mintavételi helyzet
2. A mintavétel tulajdonságok szemléltetése szimulációval
3. A mintavétel tulajdonságok matematikai levezetése
4. Az OLS modellfeltevései

Deskriptív és következtető statisztika

- Az előbbi tárgyalás pusztán deskriptív volt: egy darab mintát tekintett, amire meghatározott egy darab regressziós függvényt és kész
- Mintha a feladat csak annyi lenne, hogy pontokra húzzunk egy rájuk jól illeszkedő görbét
- Ez a „görbeillesztési" szemlélet első ránézésre könnyen megérthető, és látszólag egyszerűsíti a helyzetet, valójában azonban rendkívül hátráltató a valódi megértésre nézve
- Nem teszi lehetővé ugyanis annak megértését, hogy a háttérben van egy sokaság, és a görbe nem univerzálisan jellemzi azt, hanem csak az adott, konkrét mintára illeszkedik legjobban, és *másik mintából másik görbét kaptunk volna*
- Azaz: figyelmen kívül hagyja a mintavételi helyzetet

A mintavételi helyzet hatásai

- Van egy elméleti regresszió a **sokaságban** ($\boldsymbol{\beta}$)
- Az adatbázisunk alapján megkaptuk a regressziós egyenest ($\widehat{\boldsymbol{\beta}}$)
- Az adatbázis azonban csak egy **minta** a sokaságából, így a $\widehat{\beta}_i$ paraméterek annak hatását *is* tükrözik, hogy konkrétan milyen mintát választottunk
- *Mintavételi ingadozás* lép fel (még akkor is, ha tökéletesen véletlen a mintavétel, ennek tehát semmi köze pl. a reprezentativitáshoz)
- Tehát: az egyes $\widehat{\beta}_i$ paraméterek „mintáról-mintára ingadoznak”: minden mintából más paramétereket kapnánk
- (Természetesen reméljük, hogy az ingadozás „kellemes” tulajdonságokkal bír, például a valós érték körül történik, szorosan körülötte stb., erről később)
- Ez tehát egy becslési feladat; az OLS-nek, mint becslőfüggvénynek vizsgálhatóak a tulajdonságai

Még egy fontos megjegyzés

- Nem elég annyit mondani, hogy „jó, hát akkor a háttérben van egy sokaság is”, mintha ezzel el lenne intézve ez a kérdés
- Azt is világosan látni kell, hogy az egész tárgyalás *kiindulópontja*, hogy erre *feltételezünk* egy modellt (pl. azt, hogy $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon$)
- Ez egy feltételezés, mivel a sokaságot nem ismerjük, így biztosan nem tudhatjuk, hogy igaz-e (csak következtethetünk rá)
- De minden további levezetés mögött ott lesz, hogy mi mit gondoltunk, hogyan viselkedik a sokaság, mi a *sokasági modell*

Előkészület a mintavétel vizsgálatához

- Ahhoz, hogy a mintavétel hatását matematikailag tudjuk vizsgálni, az OLS-becslőt val. változókra kell ráereszteni (szemben az eddigi képlettel – $\widehat{\beta_{\mathrm{OLS}}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$ – ahol konkrét számokra futtattuk)
- Pontosan ugyanúgy, ahogy az $\frac{1}{n}\sum_{i=1}^n x_i$-t sem tudjuk következtető statisztikailag vizsgálni (az egy szám), hanem az $\frac{1}{n}\sum_{i=1}^n X_i$-t nézzük
- Minket tehát $\widehat{\beta_{\mathrm{OLS}}} = \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{Y}$ fog érdekelni!
- Ahogy az előbbi átlagos példában, így itt is igaz lesz, hogy ekkor a $\widehat{\underline{\beta}_{\mathrm{OLS}}}$ már nem egy konkrét érték (vektor), hanem egy – vektor értékű – val. változó, tehát eloszlása van!
- Ez a mintavételi eloszlás, mi erre, ennek tulajdonságaira, a jó tulajdonságok feltételeire stb. leszünk kíváncsiak
- Előbb szimulációval nyerünk képet, aztán matematikailag is levezetjük

Monte Carlo szimuláció használata

- Számos konkrét véletlen mintát veszünk egy előre specifikált populációból (véletlenszám-generátort használunk)
- Lényegében: empirikusan vizsgálunk egy elméleti kérdést
- Valszámos embert játszunk: ugye azt mondtuk, hogy a valszámosok úgy dolgoznak mintha valahonnan ismernék a sokasági eloszlást – hát most tényleg ismerjük!
- Példának okáért, legyen a valódi sokasági eloszlás

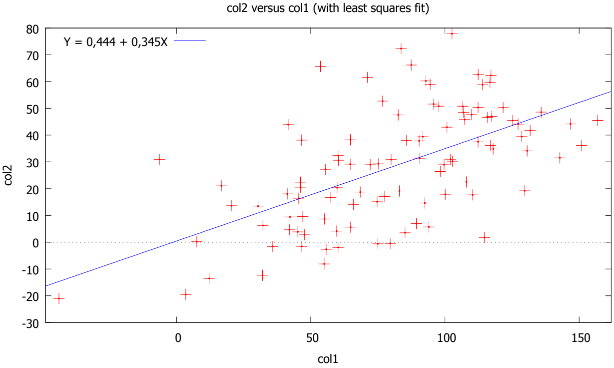
$$(X, Y) \sim \mathcal{N}\left(\begin{pmatrix} 77 \\ 26 \end{pmatrix}, \begin{pmatrix} 42^2 & 0{,}6 \cdot 20 \cdot 42 \\ 0{,}6 \cdot 20 \cdot 42 & 20^2 \end{pmatrix}\right)$$

- Ezért a *valódi* regressziós egyenes, a már látottak szerint:

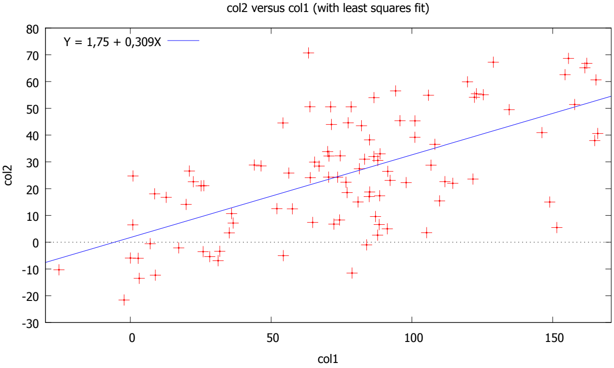
$$\mathbb{E}\left(Y \mid X\right) = 4 + \frac{12}{42}X \approx 4 + 0{,}2857X$$

- Szimulációs paraméterek: $n = 100$ elemű minta a fenti sokaságból, 1000 ismétlés

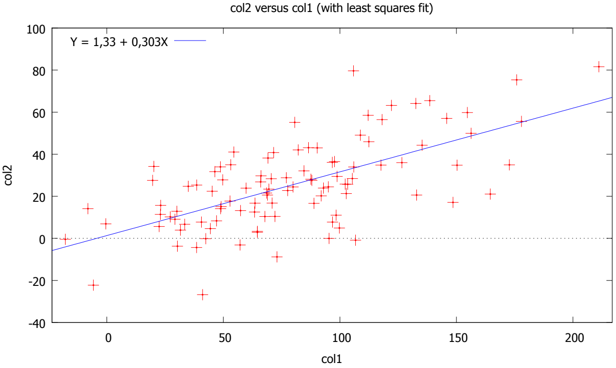
A szimuláció eredményei: 1. futtatás



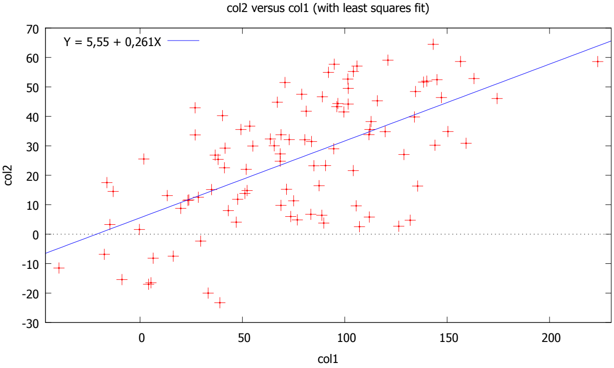
A szimuláció eredményei: 2. futtatás



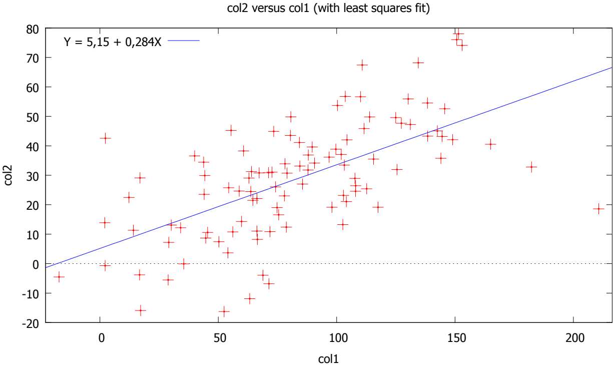
A szimuláció eredményei: 3. futtatás



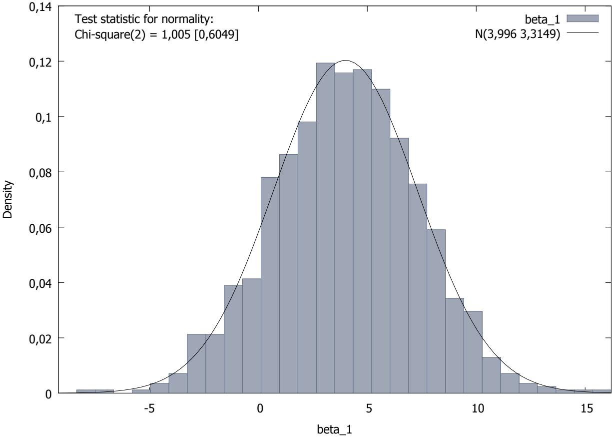
A szimuláció eredményei: 4. futtatás



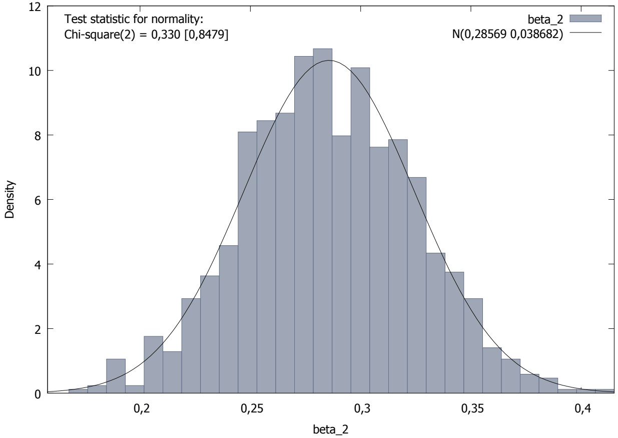
A szimuláció eredményei: 5. futtatás



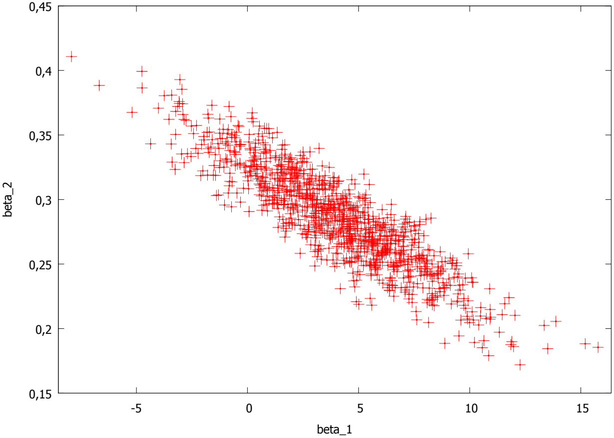
A szimuláció eredményei: konstans



A szimuláció eredményei: meredekség



A szimuláció eredményei: mindkét paraméter együtt



Az OLS-becslő mintavételi eloszlása

- Tudjuk, hogy $\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} = \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{Y}$
- Valamint elfogadtuk feltételezésként, hogy a sokasági modell $Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon = \underline{X}^T \boldsymbol{\beta} + \varepsilon$
  - És ez van mindegyik megfigyelési egység mögött is, tehát a mintavétel elemzéséhez ezt is írhatjuk:
    $$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_i$$
  - Röviden, értelemszerű vektorokba/mátrixokba fogással: $Y_i = \underline{X}_i^T \boldsymbol{\beta} + \varepsilon$ avagy az egész adatbázisra: $\underline{Y} = \underline{\underline{X}} \boldsymbol{\beta} + \underline{\varepsilon}$
- Na de rakjuk csak össze a kettőt:

$$\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} = \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{Y} = \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \left(\underline{\underline{X}} \boldsymbol{\beta} + \underline{\varepsilon}\right) =$$
$$= \left(X^T X\right)^{-1} \underline{\underline{X}}^T \underline{\underline{X}} \boldsymbol{\beta} + \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon} = \boldsymbol{\beta} + \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon}$$

Az OLS standard modellfeltevési

Ahhoz, hogy az OLS-nek fennálljanak bizonyos előnyös tulajdonságai, meghatározott feltevéseknek teljesülniük kell. Az ún. standard lineáris modell feltevései:

1. Linearitás
2. Nincs egzakt multikollinearitás
3. Erős (vagy szigorú) exogenitás
4. Homoszkedaszticitás
5. Autokorrelálatlanság

Linearitás

A sokaságot *valójában* leíró modell tényleg a feltételezett, azaz fennáll, hogy

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$$

és ez igaz mindegyik megfigyelési egységre, és így az egész mintára is:

$$\underline{Y} = \underline{\underline{X}}\boldsymbol{\beta} + \underline{\varepsilon}$$

Nincs egzakt multikollinearitás

- Egzakt multikollinearitásnak nevezzük, ha az adatmátrix nem teljes oszloprangú
- Tehát: az oszlopok között lineáris kapcsolat van
- Azaz valamelyik változó előállítható a többi lineáris kombinációjaként
- Érezhető, hogy nem túl szerencsés: minek használjuk egyáltalán azt a változót...? (Úgyis lineáris kombinációt képezünk a többiből is!) $\rightarrow$ a hatások nem lesznek szétválaszthatóak
- Sőt: az OLS becslőfüggvényéből az is látszik, hogy ilyenkor teljesen elakadunk: $\mathbf{X}^T\mathbf{X}$ szinguláris ($\underline{\underline{X}}^T\underline{\underline{X}}$ 1 valószínűséggel szinguláris)
- Ennek feltétele: $\mathbf{X}$ ($\underline{\underline{X}}$) nem teljes oszloprangú

Nincs egzakt multikollinearitás

- A feltétel tehát: az adatmátrix 1 valószínűséggel legyen teljes oszloprangú:

$$\mathbb{P}\left(\operatorname{rank}\underline{\underline{X}} = k+1\right) = 1$$

- Ez implikálja, hogy $n \geq k+1$ (kevesebb mint $k+1$-dimenziós vektorból nincs $k+1$ független)

Erős exogenitás

- Minden $i = 1, 2, \ldots, n$-re
$$\mathbb{E}\left(\varepsilon_i \mid \underline{X}_i\right) = 0$$
- Tartalma: a hibák – az ún. várható érték függetlenség értelemben – függetlenek a magyarázó változóktól

Az erős exogenitás következményei

- Toronyszabály miatt a feltétel *nélküli* várható érték is nulla:

$$\mathbb{E}\left[\mathbb{E}\left(\varepsilon_i \mid \underline{X}_i\right)\right] = \mathbb{E}\varepsilon_i = \mathbb{E}\left(0\right) = 0$$

- A várható érték függetlenség implikálja a korrelálatlanságot: $\operatorname{cov}\left(X_{ik}, \varepsilon_j\right) = 0$ avagy – ezzel egyenértékűen, hiszen $\mathbb{E}\varepsilon_i = 0$ – $\mathbb{E}\left(X_{ik}\varepsilon_j\right) = 0$
- Szokás a korrelálatlanság helyett azt is mondani, hogy a hibák ortogonálisak a magyarázóváltozókra

Az erős exogenitás sérülésének tipikus esetei

- Van olyan változó, ami lényeges magyarázó változó lenne (tehát valódi – sokasági – β-ja nem nulla), de mégsem szerepel a modellben, miközben legalább egy magyarázó változóval korrelál (kihagyott változó esete, „omitted variable bias") – ez épp a confounding!
- Mérési hiba magyarázó változónál (tehát a mérési változók valódi értékét nem, csak valamilyen zajjal terhelve tudjuk mérni)
- Szimultaneitás (többegyenletes modelleknél)

Az erős exogenitás sérülésének kezelése

- A problémát orvosolhatjuk a megfelelő(bb) modellspecifikációval, függően attól, hogy pontosan mi a baj oka...
- ...illetve bizonyos statisztikai eszközök is a rendelkezésünkre állnak, ilyen az instrumentális változós (IV) becslés, a kétfázisú legkisebb négyzetek módszere (TSLS) stb.

Homoszkedaszticitás

- A feltétel azt köti ki, hogy $\sigma_i^2 := \mathbb{D}^2\left(\varepsilon_i \mid \underline{\underline{X}}\right) = \sigma^2$ i-től függetlenül minden $i = 1, 2, \ldots, n$-re
- Tartalma: a hibák különböző megfigyelésekhez tartozó szórása állandó (nem függ attól, hogy melyik megfigyelésről van szó) avagy – másként megfogalmazva ugyanez – a becsült értékek szóródása a tényleges körül állandó
- Jellemzően keresztmetszeti adatoknál felmerülő kérdés (hamarosan foglalkozunk is vele bővebben)

Autokorrelálatlanság

- Tartalma: a különböző megfigyelésekhez tartozó hibák korrelálatlanok egymással
- Fae mintavételezésnél ez tehát *automatikusan* teljesül!
- Nem fae esetben a feltétel azt köti ki, hogy $\mathrm{cov}\left(\varepsilon_i, \varepsilon_j \mid \underline{\underline{X}}\right) = 0$ minden $i, j = 1, 2, \ldots, n$, $i \neq j$-re
- Ezzel egyenértékű $\mathbb{E}\left(\varepsilon_i \varepsilon_j \mid \underline{\underline{X}}\right) = 0$ (hiszen $\mathbb{E}\varepsilon_i = 0$, így a kovariancia a két változó szorzatának várható értéke)
- Elsősorban idősoros adatok kérdésköre, most nem is foglalkozunk vele bővebben

A homoszkedaszticitás és az autokorrelálatlanság együtt

- Mindkettő felfogható úgy, mint az ε_i hibák (feltételes) kovarianciamátrixára vonatkozó megkötés
  - Homoszkedaszticitás: a kovarianciamátrix főátlójában ugyanazok az elemek (σ^2) vannak (ugye itt vannak a szórásnégyzetek)
  - Autokorrelálatlanság: a kovarianciamátrix főátlóján kívüli elemek nullák (a mátrix diagonális)
- A kettő *együtt*: a kovarianciamátrix $\sigma^2\mathbf{I}$ alakú (szokás az ilyet skalármátrixnak is nevezni)

σ^2 becslése

Nem részletezzük, de belátható, hogy ez esetben a σ^2-re adható OLS-becslés:

$$\widehat{\sigma^2} = \frac{ESS}{n-(k+1)} = \frac{\widehat{\mathbf{e}}^T\widehat{\mathbf{e}}}{n-(k+1)}$$

Mintavételileg rögzített magyarázó változók

- Egyszerűbb tárgyalások azt feltételezik, hogy a magyarázó változók mintavételileg rögzítettek (mintha determinisztikusan megszabhatnánk az értéküket: $\underline{X}_i$ igazából $\mathbf{x}_i$)
- Ennek sok baja van:
  1. Nem annyira szép és elegáns (nyilván ez speciális esete a mi tárgyalásunknak!)
  2. Nem teszi lehetővé egy sor kérdés mélyebb tárgyalását
  3. Alapjában megkérdőjelezhető az alkalmazása nem-experimentális tudományokban (mint a közgazdaságtan...)
- Az előnye, hogy egyszerűsít: ekkor a hiba feltételes és feltétel nélküli eloszlása ugyanaz lesz, a '$\mid \underline{X}_i$' jellegű feltételek elhagyhatóak...
- ...emiatt a modellfeltevések a következőkre egyszerűsödnek:
  - Erős exogenitás: $\mathbb{E}\varepsilon_i = 0$ minden $i = 1, 2, \ldots, n$-re
  - Homoszkedaszticitás: $\mathbb{D}^2\varepsilon_i = \sigma^2$ minden $i = 1, 2, \ldots, n$-re
  - Autokorrelálatlanság: $\mathbb{E}\left(\varepsilon_i\varepsilon_j\right) = 0$ minden $i \neq j = 1, 2, \ldots, n$

A mintavételi tulajdonságok

- Ezek lesznek a standard modellfeltevések...
- ...most nekiállunk megvizsgálni, hogy a teljesülésük esetén milyen tulajdonságokkal bír az OLS-becslő

Várható érték

- Tudjuk, hogy

$$\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} = \boldsymbol{\beta} + \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon}$$

- Ez alapján mi $\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}}$ várható értéke (várható érték-vektora)?

$$\begin{aligned}\mathbb{E}\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} &= \boldsymbol{\beta} + \mathbb{E}\left[\left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon}\right] = \\ &= \boldsymbol{\beta} + \mathbb{E}\left\{\mathbb{E}\left[\left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon} \mid \underline{\underline{X}}\right]\right\} = \\ &= \boldsymbol{\beta} + \mathbb{E}\left\{\left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \mathbb{E}\left[\underline{\varepsilon} \mid \underline{\underline{X}}\right]\right\} = \boldsymbol{\beta}\end{aligned}$$

- Az erős exogenitás fennállása esetén tehát az OLS szolgáltatta becslések *torzítatlanok*
- Nem bizonyítjuk, de az is igaz, hogy *konzisztensek*

Kovarianciamátrix

Az előbbi ismeretében:

$$\begin{aligned}
\mathbb{D}^2\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} &= \mathbb{E}\left[\left(\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} - \mathbb{E}\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}}\right)\cdot\left(\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} - \mathbb{E}\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}}\right)^T\right] = \\
&= \mathbb{E}\left[\left(\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} - \boldsymbol{\beta}\right)\cdot\left(\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} - \boldsymbol{\beta}\right)^T\right] = \\
&= \mathbb{E}\left\{\left[\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}\right]\cdot\left[\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}\right]^T\right\} = \\
&= \mathbb{E}\left[\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}\underline{\varepsilon}^T\underline{\underline{X}}\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\right] = \\
&= \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\mathbb{E}\left(\underline{\varepsilon}\underline{\varepsilon}^T\right)\underline{\underline{X}}\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1} = \\
&= \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\cdot\sigma^2\mathbf{I}\cdot\underline{\underline{X}}\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1} = \sigma^2\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}
\end{aligned}$$

A Gauss–Markov tétel

- Ha mindegyik feltevés teljesül, akkor lineáris torzítatlan becslők körében az OLS-becslő minimális varianciájú (azaz hatásos)
- Tehát: $\mathbb{D}^2\left(\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}}\right) \leq \mathbb{D}^2\left(\widehat{\boldsymbol{\beta'}}\right)$ bármely más $\widehat{\boldsymbol{\beta'}}$ lineáris becslőre, amire $\mathbb{E}\left(\widehat{\boldsymbol{\beta'}}\right) = \boldsymbol{\beta}$ (azaz torzítatlan)

Összefoglalva

- Amennyiben a standard modellfeltevések közül teljesül a:
  - Linearitás
  - Nincs egzakt multikollinearitás
  - Erős exogenitás

  akkor az OLS szolgáltatatta becslések *torzítatlanok* és *konzisztensek*
- Ha ezen felül teljesül a:
  - Homoszkedaszticitás
  - Autokorrelálatlanság

  akkor az OLS szolgáltatta becslések *hatásosak* (minimális varianciájuk) is

BLUE-tulajdonság

Ezt röviden úgy szokták megfogalmazni, hogy ha valamennyi standard modellfeltétel teljesül, akkor az OLS szolgáltatta becslések BLUE-k:

- Best (minimális varianciájú)
- Linear (lineáris a mintaelemekben)
- Unbiased (torzítatlan)

A σ^2 és a koefficiensek kovarianciamátrixának becslői

- A σ^2-nek a $\widehat{\sigma^2} = \frac{ESS}{n-(k+1)}$ becslője torzítatlan, ha mindegyik feltétel fennáll
- A β_i koefficiensek kovarianciamátrixának $\widehat{\sigma^2} \left(\underline{\underline{X}}^T \underline{\underline{X}} \right)^{-1}$ becslője szintén
- Tehát vigyázat: itt *már* a torzítatlansághoz *is* kell mindegyik feltétel (a homoszkedaszticitás és az autokorrelálatlanság is)!

A $\widehat{\beta}_i$ koefficiensek eloszlása

- Az eddigi eredmények ugyan nagyon biztatóak, de még mindig nem mondanak semmit arról, hogy konkrétan mi a becsült koefficiensek (mintavételi) eloszlása
- A $\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} = \boldsymbol{\beta} + \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{\varepsilon}$ nem sok jót sejtet: ebből úgy tűnik, hogy ez $\underline{\underline{X}}$-től és ε-tól is függ, ráadásul egy elég komplexnek kinéző módon...
- Szerencsére nem ennyire rossz a helyzet!
- Van egy nevezetes speciális eset, amikor a becsült koefficiensek eloszlása egyszerű alakú, és *nem is függ* $\underline{\underline{X}}$ eloszlásától, ez pedig az, ha a hibák feltételes eloszlása normális
- Vigyázat: a hibák normalitása *nem* része a standard modellfeltevéseknek, azaz a BLUE-ság akkor is megvalósul, ha a hibák eloszlása nem normális!
- Ráadásul, még ha nem is tudjuk, hogy a normalitás teljesül, de nagy a mintánk, akkor a centrális határeloszlás-tétel miatt aszimptotikus közelítésként akkor is használhatjuk az így nyert eredményeket

Hibák normalitása

- $\underline{\varepsilon}$ feltételes eloszlása feltéve $\underline{\underline{X}}$-et többváltozós normális
- A standard modellfeltevéseket is felhasználva ez azt jelenti, hogy

$$\underline{\varepsilon} \mid \underline{\underline{X}} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$$

- Ez láthatóan nem függ $\underline{\underline{X}}$-től, így persze a hibák feltétel nélküli eloszlása is $\mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$

Hibanormalitás és a becsült koefficiensek eloszlása

- Ha $\underline{\varepsilon}$ eloszlása normális, akkor $\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}$-é is az
- Ez azért nagyon jó hír, mert a normális eloszláshoz csak két dolgot kell tudnunk: várható érték-vektort és kovarianciamátrixot!
- Az viszont könnyen meghatározható (az egyszerűség kedvéért a $\mid \underline{\underline{X}}$ feltételt nem írjuk ki a következőkben)
- Várható érték: $\mathbb{E}\left[\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}\right] = \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\mathbb{E}\underline{\varepsilon} = \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\mathbf{0} = \mathbf{0}$
- Kovarianciamátrix: $\mathbb{D}^2\left[\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\underline{\varepsilon}\right] = \left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\cdot\mathbb{D}^2\underline{\varepsilon}\cdot\underline{\underline{X}}\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1} =$ $\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\underline{\underline{X}}^T\cdot\sigma^2\mathbf{I}\cdot\underline{\underline{X}}\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1} = \sigma^2\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}$
- Összefoglalva: $\widehat{\boldsymbol{\beta}_{\mathrm{OLS}}} \sim \mathcal{N}\left(\boldsymbol{\beta}, \sigma^2\left(\underline{\underline{X}}^T\underline{\underline{X}}\right)^{-1}\right)$

Konfidenciaintervallum a paraméterekre

Hibanormalitás esetén, vagy aszimptotikusan könnyen szerkeszthető konfidenciaintervallum is, $1-\alpha$ megbízhatósági szinten:

$$\widehat{\beta}_i \pm t_{n-(k+1)}^{(1-\alpha/2)} \cdot \mathrm{se}\left(\widehat{\beta}_i\right)$$