

Hipotézisvizsgálat és intervallumbecslés lineáris modellben

Ferenci Tamás
tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 12.

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter
- 3 Modell egésze
- 4 Tetszőleges számú paraméter
- 5 Lineáris megkötés(ek)

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter
- 3 Modell egésze
- 4 Tetszőleges számú paraméter
- 5 Lineáris megkötés(ek)

Emlékeztetőül

- A most következő eredmények csak akkor egzaktak, ha a hibanormalitás is fennáll
- Ám aszimptotikusak, így közelítőleg akkor is fennállnak, ha elég nagy a mintanagyság (minél nagyobb, annál inkább)

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter**
- 3 Modell egésze
- 4 Tetszőleges számú paraméter
- 5 Lineáris megkötés(ek)

Becsült regressziós koefficiensek mintavételi eloszlása

- A $\hat{\beta}_i$ becsült regressziós koefficiens mintavételi ingadozását tehát a következő összefüggés írja le:

$$\frac{\hat{\beta}_i - \beta_i}{\text{se}(\hat{\beta}_i)} \sim \mathcal{N}(0, 1),$$

$$\text{ahol } \text{se}(\hat{\beta}_i) = \sqrt{\sigma^2 \left[\left(\underline{\underline{X}}^T \underline{\underline{X}} \right)^{-1} \right]_{kk}}$$

- Sajnos ezzel a gyakorlatban nem sokra megyünk, mert σ^2 -et általában nem ismerjük
- Helyettesítsük a jó tulajdonságú becslőjével, $\hat{\sigma}^2$ -tel!
- Így persze már más lesz az eloszlás, de szerencsére meghatározható, hogy mi, és nem bonyolult: $n - (k + 1)$ szabadságfokú t -eloszlás

Becsült regressziós koefficiensek mintavételi eloszlása

- A $\hat{\beta}_i$ becsült regressziós koefficiens mintavételi ingadozását tehát a következő összefüggés írja le:

$$\frac{\hat{\beta}_i - \beta_i}{\text{se}(\hat{\beta}_i)} \sim \mathcal{N}(0, 1),$$

$$\text{ahol } \text{se}(\hat{\beta}_i) = \sqrt{\sigma^2 \left[\left(\underline{\underline{\mathbf{X}^T \mathbf{X}}} \right)^{-1} \right]_{kk}}$$

- Sajnos ezzel a gyakorlatban nem sokra megyünk, mert σ^2 -et általában nem ismerjük
- Helyettesítsük a jó tulajdonságú becslőjével, $\hat{\sigma}^2$ -tel!
- Így persze már más lesz az eloszlás, de szerencsére meghatározható, hogy mi, és nem bonyolult: $n - (k + 1)$ szabadságfokú t -eloszlás

Becsült regressziós koefficiensek mintavételi eloszlása

- A $\hat{\beta}_i$ becsült regressziós koefficiens mintavételi ingadozását tehát a következő összefüggés írja le:

$$\frac{\hat{\beta}_i - \beta_i}{\text{se}(\hat{\beta}_i)} \sim \mathcal{N}(0, 1),$$

$$\text{ahol } \text{se}(\hat{\beta}_i) = \sqrt{\sigma^2 \left[\left(\underline{\underline{\mathbf{X}^T \mathbf{X}}} \right)^{-1} \right]_{kk}}$$

- Sajnos ezzel a gyakorlatban nem sokra megyünk, mert σ^2 -et általában nem ismerjük
- Helyettesítsük a jó tulajdonságú becslőjével, $\hat{\sigma}^2$ -tel!
- Így persze már más lesz az eloszlás, de szerencsére meghatározható, hogy mi, és nem bonyolult: $n - (k + 1)$ szabadságfokú t -eloszlás

Becsült regressziós koefficiensek mintavételi eloszlása

- A $\hat{\beta}_i$ becsült regressziós koefficiens mintavételi ingadozását tehát a következő összefüggés írja le:

$$\frac{\hat{\beta}_i - \beta_i}{\text{se}(\hat{\beta}_i)} \sim \mathcal{N}(0, 1),$$

$$\text{ahol } \text{se}(\hat{\beta}_i) = \sqrt{\sigma^2 \left[(\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \right]_{kk}}$$

- Sajnos ezzel a gyakorlatban nem sokra megyünk, mert σ^2 -et általában nem ismerjük
- Helyettesítsük a jó tulajdonságú becslőjével, $\hat{\sigma}^2$ -tel!
- Így persze már más lesz az eloszlás, de szerencsére meghatározható, hogy mi, és nem bonyolult: $n - (k + 1)$ szabadságfokú t -eloszlás

Változó relevanciája

Egy változót relevánsnak nevezünk, ha a sokasági paramétere nem nulla: $\beta_i \neq 0$.

Hipotézisvizsgálat változó relevanciájára

Ez alapján már konstruálhatunk próbát változó relevanciájának vizsgálatára:

- 1 $H_0 : \beta_i = 0$
- 2 Ekkor (azaz *ha* ez fennáll!) a $t_{\text{emp},i} = \frac{\hat{\beta}_i}{\text{se}(\hat{\beta}_i)}$ kifejezés $n - (k + 1)$ szabadságfokú t -eloszlást követ (nulleloszlás)
- 3 Számítsuk ki a konkrét $t_{\text{emp},i}$ -t a mintánkból és döntsük el, hogy hihető-e, hogy $t_{n-(k+1)}$ -ből származik

Hipotézisvizsgálat változó relevanciájára

A hipotézisvizsgálat elvégzéséhez szükséges minden tudnivalót – a nullhipotézisen kívül – összefoglal tehát a következő kifejezés (a későbbiekben is ezt a sémát fogjuk használni hipotézisvizsgálatok megadására):

$$t_{\text{emp},i} = \frac{\widehat{\beta}_i}{\text{se}(\widehat{\beta}_i)} \stackrel{H_0}{\sim} t_{n-(k+1)}.$$

E próba precíz neve: változó relevanciájára irányuló (parciális) t -próba

Hipotézisvizsgálat változó relevanciájára

A hipotézisvizsgálat elvégzéséhez szükséges minden tudnivalót – a nullhipotézisen kívül – összefoglal tehát a következő kifejezés (a későbbiekben is ezt a sémát fogjuk használni hipotézisvizsgálatok megadására):

$$t_{\text{emp},i} = \frac{\widehat{\beta}_i}{\text{se}(\widehat{\beta}_i)} \stackrel{H_0}{\sim} t_{n-(k+1)}.$$

E próba precíz neve: változó relevanciájára irányuló (parciális) t -próba

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter
- 3 Modell egésze**
- 4 Tetszőleges számú paraméter
- 5 Lineáris megkötés(ek)

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (semmilyen más eset jelölésére *ne* használjuk az egyenlőségjelcímet!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (semmilyen más eset jelölésére *ne* használjuk az egyenlőségjelölést!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (*semmilyen más* eset jelölésére *ne* használjuk az egyenlőségjelölést!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (*semmilyen más* eset jelölésére *ne* használjuk az egyenlőségjelölést!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (*semmilyen más* eset jelölésére *ne* használjuk az egyenlőségjelét!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A korábban látott t -próba azért volt „parciális”, mert egy változó irrelevanciáját vizsgálta
- Felmerül a kérdés, hogy definiálható-e a modell *egészének* irrelevanciája
- Igen, mégpedig úgy, hogy *valamennyi* magyarázó változó paramétere *együttesen is* irreleváns:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

- (Természetesen a β_0 nincs felsorolva!)
- Rövid jelölés arra, hogy $\beta_1 = 0$ és $\beta_2 = 0$ stb. és $\beta_k = 0$ (*semmilyen más* eset jelölésére *ne* használjuk az egyenlőségjelölést!)
- Figyelem: az „egyszerre nulla mindegyik” *több* mint, hogy „külön-külön nulla mindegyik”!

Modell egészének relevanciája

- A modell egészének irrelevanciájára magyarul azt jelenti, hogy a modell nem tér el lényegesen a nullmodelltől
- Implikálja, hogy minden magyarázó változó külön-külön is irreleváns (tartalmazza ezeket a hipotéziseket) → előbb teszteljük a modell egészének irrelevanciáját, és csak ennek elvetése utána teszteljük a változókat parciálisan
- A próba konkrét alakja:

$$F_{\text{emp}} = \frac{RSS/k}{ESS/[n - (k + 1)]} \stackrel{H_0}{\sim} \mathcal{F}_{k, n-(k+1)}$$

Modell egészének relevanciája

- A modell egészének irrelevanciájára magyarul azt jelenti, hogy a modell nem tér el lényegesen a nullmodelltől
- Implikálja, hogy minden magyarázó változó külön-külön is irreleváns (tartalmazza ezeket a hipotéziseket) → előbb teszteljük a modell egészének irrelevanciáját, és csak ennek elvetése után teszteljük a változókat parciálisan
- A próba konkrét alakja:

$$F_{\text{emp}} = \frac{RSS/k}{ESS/[n - (k + 1)]} \stackrel{H_0}{\sim} \mathcal{F}_{k, n-(k+1)}$$

Modell egészének relevanciája

- A modell egészének irrelevanciájára magyarul azt jelenti, hogy a modell nem tér el lényegesen a nullmodelltől
- Implikálja, hogy minden magyarázó változó külön-külön is irreleváns (tartalmazza ezeket a hipotéziseket) → előbb tesztljük a modell egészének irrelevanciáját, és csak ennek elvetése utána tesztljük a változókat parciálisan
- A próba konkrét alakja:

$$F_{\text{emp}} = \frac{RSS/k}{ESS/[n - (k + 1)]} \stackrel{H_0}{\sim} \mathcal{F}_{k, n-(k+1)}$$

Modell egészének relevanciája

- A tesztstatisztika átírható mint

$$\frac{RSS/k}{ESS/[n - (k + 1)]} = \frac{R^2/k}{(1 - R^2) / [n - (k + 1)]}$$

- Persze: a „nem tér el lényegesen a nullmodelltől” úgy is megfogalmazható, hogy az „ R^2 nem tér el lényegesen a nullától” ($H_0 : R^2 = 0$ is mondható lett volna)

Modell egészének relevanciája

- A tesztstatisztika átírható mint

$$\frac{RSS/k}{ESS/[n - (k + 1)]} = \frac{R^2/k}{(1 - R^2) / [n - (k + 1)]}$$

- Persze: a „nem tér el lényegesen a nullmodelltől” úgy is megfogalmazható, hogy az „ R^2 nem tér el lényegesen a nullától” ($H_0 : R^2 = 0$ is mondható lett volna)

Modell egészének relevanciája

- A próba neve: a modell egészének relevanciájára irányuló (globális) F -próba
- Szokás ANOVA-próbának is nevezni (a $TSS = ESS + RSS$ variancia-felbontáson alapszik; számlálóban és nevezőben a fokszámmal normált szórásnégyzetek vannak)
- Tipikus eredményközlés az ún. ANOVA-táblában

Modell egészének relevanciája

- A próba neve: a modell egészének relevanciájára irányuló (globális) F -próba
- Szokás ANOVA-próbának is nevezni (a $TSS = ESS + RSS$ variancia-felbontáson alapszik; számlálóban és nevezőben a fokszámmal normált szórásnégyzetek vannak)
- Tipikus eredményközlés az ún. ANOVA-táblában

Modell egészének relevanciája

- A próba neve: a modell egészének relevanciájára irányuló (globális) F -próba
- Szokás ANOVA-próbának is nevezni (a $TSS = ESS + RSS$ variancia-felbontáson alapszik; számlálóban és nevezőben a fokszámmal normált szórásnégyzetek vannak)
- Tipikus eredményközlés az ún. ANOVA-táblában

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter
- 3 Modell egésze
- 4 Tetszőleges számú paraméter**
- 5 Lineáris megkötés(ek)

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
- Wald-elv, LM-elv, LR-elv
- Az eddigi két próba Wald-elven is kihozható

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
- Wald-elv, LM-elv, LR-elv
- Az eddigi két próba Wald-elven is kihozható

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
- Wald-elv, LM-elv, LR-elv
- Az eddigi két próba Wald-elven is kihozható

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
 - Wald-elv, LM-elv, LR-elv
 - Az eddigi két próba Wald-elven is kihozható

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
- Wald-elv, LM-elv, LR-elv
- Az eddigi két próba Wald-elven is kihozható

Felvezető gondolatok

- Valamennyi eddigi próba felírható úgy, hogy van egy modellünk, a nullhipotézis pedig egy megkötést jelent arra a modellre
- Azaz lényegében két modellünk van, egy megkötés nélküli és egy megkötött
- Mellesleg a megkötött modell szükségképp rosszabb, de legalábbis nem jobb (szűkebb tartományon vett optimum nem lehet jobb, mint egy bővebben vett), emiatt úgy is megfogalmazható a kérdés, hogy a különbség lényeges-e
- Az ilyen helyzetre – mint bármilyen helyzetre – többféle elven lehet tesztet konstruálni
- Wald-elv, LM-elv, LR-elv
- Az eddigi két próba Wald-elven is kihozható

Tetszőleges számú paraméter tesztelése Wald-elven

- Most felírjuk a két modellt explicite is, mert a nullhipotézis alakja szebb lesz (ez pusztán formai kérdés):
- Az egyik modell a bővebb (U – unrestricted), a másik a szűkebb (R – restricted):

$$U: Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R: Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

- $H_0: \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$, tehát megadott m darab változó még összességében sem bír lényeges magyarázó erővel

Tetszőleges számú paraméter tesztelése Wald-elven

- Most felírjuk a két modellt explicite is, mert a nullhipotézis alakja szebb lesz (ez pusztán formai kérdés):
- Az egyik modell a bővebb (U – unrestricted), a másik a szűkebb (R – restricted):

$$U : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

- $H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$, tehát megadott m darab változó még összességében sem bír lényeges magyarázó erővel

Tetszőleges számú paraméter tesztelése Wald-elven

- Most felírjuk a két modellt explicite is, mert a nullhipotézis alakja szebb lesz (ez pusztán formai kérdés):
- Az egyik modell a bővebb (U – unrestricted), a másik a szűkebb (R – restricted):

$$U : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

- $H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$, tehát megadott m darab változó még összességében sem bír lényeges magyarázó erővel

Tetszőleges számú paraméter tesztelése Wald-elven

A próba:

$$\begin{aligned} F_{\text{emp}} &= \frac{(ESS_R - ESS_U) / m}{ESS_U / (n - q - m)} = \\ &= \frac{(R_U^2 - R_R^2) / m}{(1 - R_U^2) / (n - q - m)} \stackrel{H_0}{\sim} \mathcal{F}_{m, n-q-m}. \end{aligned}$$

Speciális esetek

- Vegyük észre, hogy ez az általános megközelítés a két, eddig látott tesztet is tartalmazza speciális esetként!
- Ha $m = 1$, akkor $F = t_j^2$: visszakaptuk a t -tesztet
 - *Ami figyelem: a Wald-teszt nem ekvivalens a t -próba m -szeri elvégzésével (külön-külön az egyes változókra)!*
- Ha $m = k$, akkor $F_{\text{Wald}} = F_{\text{ANOVA}}$: visszakaptuk a függetlenségvizsgálatot
- Logikusak, hiszen a nullhipotézisek is azonos alakúak lettek

Speciális esetek

- Vegyük észre, hogy ez az általános megközelítés a két, eddig látott tesztet is tartalmazza speciális esetként!
- Ha $m = 1$, akkor $F = t_j^2$: visszakaptuk a t -tesztet
 - Ám figyelem: a Wald-teszt *nem* ekvivalens a t -próba m -szeri elvégzésével (külön-külön az egyes változókra)!
- Ha $m = k$, akkor $F_{\text{Wald}} = F_{\text{ANOVA}}$: visszakaptuk a függetlenségvizsgálatot
- Logikusak, hiszen a nullhipotézisek is azonos alakúak lettek

Speciális esetek

- Vegyük észre, hogy ez az általános megközelítés a két, eddig látott tesztet is tartalmazza speciális esetként!
- Ha $m = 1$, akkor $F = t_j^2$: visszakaptuk a t -tesztet
 - Ám figyelem: a Wald-teszt *nem* ekvivalens a t -próba m -szeri elvégzésével (külön-külön az egyes változókra)!
- Ha $m = k$, akkor $F_{\text{Wald}} = F_{\text{ANOVA}}$: visszakaptuk a függetlenségvizsgálatot
- Logikusak, hiszen a nullhipotézisek is azonos alakúak lettek

Speciális esetek

- Vegyük észre, hogy ez az általános megközelítés a két, eddig látott tesztet is tartalmazza speciális esetként!
- Ha $m = 1$, akkor $F = t_j^2$: visszakaptuk a t -tesztet
 - Ám figyelem: a Wald-teszt *nem* ekvivalens a t -próba m -szeri elvégzésével (külön-külön az egyes változókra)!
- Ha $m = k$, akkor $F_{\text{Wald}} = F_{\text{ANOVA}}$: visszakaptuk a függetlenségvizsgálatot
- Logikusak, hiszen a nullhipotézisek is azonos alakúak lettek

Speciális esetek

- Vegyük észre, hogy ez az általános megközelítés a két, eddig látott tesztet is tartalmazza speciális esetként!
- Ha $m = 1$, akkor $F = t_j^2$: visszakaptuk a t -tesztet
 - Ám figyelem: a Wald-teszt *nem* ekvivalens a t -próba m -szeri elvégzésével (külön-külön az egyes változókra)!
- Ha $m = k$, akkor $F_{\text{Wald}} = F_{\text{ANOVA}}$: visszakaptuk a függetlenségvizsgálatot
- Logikusak, hiszen a nullhipotézisek is azonos alakúak lettek

Kitérő: a Lagrange Multiplikátor (LM)-elv

- Az LM (Lagrange Multiplikátor) próba hipotézispárja *teljesen* azonos alakú a Wald- F -teszttel:

$$U : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

és $H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$

- A különbség a modellezés filozófiájában van (ld. később), a teszt tulajdonságai, alkalmazhatósága is eltérő
- Alapötlet: becsüljük meg a szűkebb modellt, és számítsuk ki ez alapján a becsült reziduumokat. Ha fennáll H_0 , akkor ezek a reziduumok nem magyarázhatóak lényegesen sem a szűkebb modell változóival (OLS következménye), sem a vizsgált változókkal (H_0 következménye). Azaz: ha a becsült reziduumokat kiregresszáljuk az összes változóval, akkor sem tudjuk azt lényegesen magyarázni, ha fennáll a H_0 .

Kitérő: a Lagrange Multiplikátor (LM)-elv

- Az LM (Lagrange Multiplikátor) próba hipotézispárja *teljesen* azonos alakú a Wald- F -teszttel:

$$U : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

és $H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$

- A különbség a modellezés filozófiájában van (ld. később), a teszt tulajdonságai, alkalmazhatósága is eltérő
- Alapötlet: becsüljük meg a szűkebb modellt, és számítsuk ki ez alapján a becsült reziduumokat. Ha fennáll H_0 , akkor ezek a reziduumok nem magyarázhatóak lényegesen sem a szűkebb modell változóival (OLS következménye), sem a vizsgált változókkal (H_0 következménye). Azaz: ha a becsült reziduumokat kiregresszáljuk az összes változóval, akkor sem tudjuk azt lényegesen magyarázni, *ha* fennáll a H_0 .

Kitérő: a Lagrange Multiplikátor (LM)-elv

- Az LM (Lagrange Multiplikátor) próba hipotézispárja *teljesen* azonos alakú a Wald- F -teszttel:

$$U : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_{q+m} X_{q+m} + \varepsilon_U$$

$$R : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{q-1} X_{q-1} + \beta_q X_q + \varepsilon_R$$

és $H_0 : \beta_{q+1} = \beta_{q+2} = \dots = \beta_{q+m} = 0$

- A különbség a modellezés filozófiájában van (ld. később), a teszt tulajdonságai, alkalmazhatósága is eltérő
- Alapötlet: becsüljük meg a szűkebb modellt, és számítsuk ki ez alapján a becsült reziduumokat. Ha fennáll H_0 , akkor ezek a reziduumok nem magyarázhatóak lényegesen sem a szűkebb modell változóival (OLS következménye), sem a vizsgált változókkal (H_0 következménye). Azaz: ha a becsült reziduumokat kiregresszáljuk az összes változóval, akkor sem tudjuk azt lényegesen magyarázni, ha fennáll a H_0 .

Az LM-próba próbafüggvénye

- Ezen intuitív indoklás után a próbafüggvény:

$$n \cdot R_{\hat{\epsilon}_R | X_1, X_2, \dots, X_{q+m}}^2 \stackrel{H_0}{\sim} \chi_m^2$$

- Itt $\hat{\epsilon}_R$ jelölés arra utal, hogy a szűkebb (R) modellből kapott reziduumokról van szó

Az LM-próba próbafüggvénye

- Ezen intuitív indoklás után a próbafüggvény:

$$n \cdot R_{\hat{\epsilon}_R | X_1, X_2, \dots, X_{q+m}}^2 \stackrel{H_0}{\sim} \chi_m^2$$

- Itt $\hat{\epsilon}_R$ jelölés arra utal, hogy a szűkebb (R) modellből kapott reziduumokról van szó

Tartalom

- 1 Alkalmazási feltételek
- 2 Egy paraméter
- 3 Modell egésze
- 4 Tetszőleges számú paraméter
- 5 **Lineáris megkötés(ek)**

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficienst is érinthet, de csak egy egyenletet tartalmazhat
- Például:

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficienst is érinthet, de csak egy egyenletet tartalmazhat
- Például:

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficiens is érinthet, de csak egy egyenletet tartalmazhat
- Például:

• Két koefficiens egyezik, $\beta_i = \beta_w$ (ekkor $r_i = +1$, $r_w = -1$, a többi r nulla és $r = 0$)

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficienst is érinthet, de csak egy egyenletet tartalmazhat
- Például:
 - Két koefficiens egyezik, $\beta_l = \beta_m$ (ekkor $r_l = +1$, $r_m = -1$, a többi r_i nulla és $r = 0$)
 - Egyik koefficiens c -szerese a másiknak, $\beta_l = c\beta_m$ (ekkor $r_l = +1$, $r_m = -c$, a többi r_i nulla és $r = 0$)
 - Az összes koefficiens összege épp nulla (ekkor mindegyik r_i 1 és $r = 0$)

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficiens is érinthet, de csak egy egyenletet tartalmazhat
- Például:
 - Két koefficiens egyezik, $\beta_l = \beta_m$ (ekkor $r_l = +1$, $r_m = -1$, a többi r_i nulla és $r = 0$)
 - Egyik koefficiens c -szerese a másiknak, $\beta_l = c\beta_m$ (ekkor $r_l = +1$, $r_m = -c$, a többi r_i nulla és $r = 0$)
 - Az összes koefficiens összege épp nulla (ekkor mindegyik r_i 1 és $r = 0$)

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficiens is érinthet, de csak egy egyenletet tartalmazhat
- Például:
 - Két koefficiens egyezik, $\beta_l = \beta_m$ (ekkor $r_l = +1$, $r_m = -1$, a többi r_i nulla és $r = 0$)
 - Egyik koefficiens c -szerese a másiknak, $\beta_l = c\beta_m$ (ekkor $r_l = +1$, $r_m = -c$, a többi r_i nulla és $r = 0$)
 - Az összes koefficiens összege épp nulla (ekkor mindegyik r_i 1 és $r = 0$)

Lineáris kombináció tesztelése

- A séma:

$$r_1\beta_1 + r_2\beta_2 + \dots + r_k\beta_k = r$$

- Avagy röviden: $\mathbf{r}^T\boldsymbol{\beta} = r$
- Több koefficiens is érinthet, de csak egy egyenletet tartalmazhat
- Például:
 - Két koefficiens egyezik, $\beta_l = \beta_m$ (ekkor $r_l = +1$, $r_m = -1$, a többi r_i nulla és $r = 0$)
 - Egyik koefficiens c -szerese a másiknak, $\beta_l = c\beta_m$ (ekkor $r_l = +1$, $r_m = -c$, a többi r_i nulla és $r = 0$)
 - Az összes koefficiens összege épp nulla (ekkor mindegyik r_i 1 és $r = 0$)

Lineáris kombináció tesztelése

- A normális lineáris modellben erre teszt szerkeszthető
- Megvalósítás: egyik lehetőség, hogy a t -próbához hasonló alakra vezetjük vissza
- Legyen $r_1\hat{\beta}_1 + r_2\hat{\beta}_2 + \dots + r_k\hat{\beta}_k = \hat{r}$, ekkor

$$\frac{\hat{r} - r}{\text{se}(\hat{r})} \stackrel{H_0}{\sim} t_{n-(k+1)}$$

- Ez az ún. *közvetlen t -próba*
- Vizsgálható Wald-jellegű próbával is

Lineáris kombináció tesztelése

- A normális lineáris modellben erre teszt szerkeszthető
- Megvalósítás: egyik lehetőség, hogy a t -próbához hasonló alakra vezetjük vissza
- Legyen $r_1\hat{\beta}_1 + r_2\hat{\beta}_2 + \dots + r_k\hat{\beta}_k = \hat{r}$, ekkor

$$\frac{\hat{r} - r}{\text{se}(\hat{r})} \stackrel{H_0}{\sim} t_{n-(k+1)}$$

- Ez az ún. *közvetlen t -próba*
- Vizsgálható Wald-jellegű próbával is

Lineáris kombináció tesztelése

- A normális lineáris modellben erre teszt szerkeszthető
- Megvalósítás: egyik lehetőség, hogy a t -próbához hasonló alakra vezetjük vissza
- Legyen $r_1\hat{\beta}_1 + r_2\hat{\beta}_2 + \dots + r_k\hat{\beta}_k = \hat{r}$, ekkor

$$\frac{\hat{r} - r}{\text{se}(\hat{r})} \stackrel{H_0}{\sim} t_{n-(k+1)}$$

- Ez az ún. *közvetlen t -próba*
- Vizsgálható Wald-jellegű próbával is

Lineáris kombináció tesztelése

- A normális lineáris modellben erre teszt szerkeszthető
- Megvalósítás: egyik lehetőség, hogy a t -próbához hasonló alakra vezetjük vissza
- Legyen $r_1\hat{\beta}_1 + r_2\hat{\beta}_2 + \dots + r_k\hat{\beta}_k = \hat{r}$, ekkor

$$\frac{\hat{r} - r}{\text{se}(\hat{r})} \stackrel{H_0}{\sim} t_{n-(k+1)}$$

- Ez az ún. *közvetlen t -próba*
- Vizsgálható Wald-jellegű próbával is

Lineáris kombináció tesztelése

- A normális lineáris modellben erre teszt szerkeszthető
- Megvalósítás: egyik lehetőség, hogy a t -próbához hasonló alakra vezetjük vissza
- Legyen $r_1\hat{\beta}_1 + r_2\hat{\beta}_2 + \dots + r_k\hat{\beta}_k = \hat{r}$, ekkor

$$\frac{\hat{r} - r}{\text{se}(\hat{r})} \stackrel{H_0}{\sim} t_{n-(k+1)}$$

- Ez az ún. *közvetlen t -próba*
- Vizsgálható Wald-jellegű próbával is

Speciális esetek

- Ez tartalmazza speciális esetként a parciális t -próbát
- De mást nem: kettő vagy több paraméter *egyidejű* nulla mivolta több megkötést jelent
- Szerencsére az előbbi kiterjeszhető több megkötés tesztelésére is:

$$\mathbf{r}_1^T \boldsymbol{\beta} = r_1$$

$$\mathbf{r}_2^T \boldsymbol{\beta} = r_2$$

⋮

$$\mathbf{r}_m^T \boldsymbol{\beta} = r_m$$

- Az \mathbf{r}_i^T sorvektorokat rakjuk össze egy \mathbf{R} mátrixba, az r_i skalárokat egy r oszlopvektorba

Több megkötés egyidejű tesztelése

- Célszerű felírás:

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r},$$

ahol \mathbf{R} $m \times k$ típusú (tehát m a megszorítások száma)

- Az erre adható teszt:

$$F_{\text{emp}} = \frac{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})^T [\mathbf{R}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{R}^T]^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}) / m}{\text{ESS} / [n - (k + 1)]} \stackrel{H_0}{\sim} \mathcal{F}[m, n - (k + 1)]$$

Több megkötés egyidejű tesztelése

- Célszerű felírás:

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r},$$

ahol \mathbf{R} $m \times k$ típusú (tehát m a megszorítások száma)

- Az erre adható teszt:

$$F_{\text{emp}} = \frac{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})^T [\mathbf{R}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{R}^T]^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}) / m}{\text{ESS} / [n - (k + 1)]} \stackrel{H_0}{\sim} \mathcal{F} [m, n - (k + 1)]$$

Konkrét példák a fenti sémára

- Ellenőrizhető, hogy ha például...

• $\dots R = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 \dots 0 \end{pmatrix}$ és $r = 0$, akkor a t -tesztet ...

• $\dots R = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ és $r = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$ akkor az ANOVA-t...

• $\dots R = \begin{pmatrix} \lambda_{\beta_1} & \lambda_{\beta_2} & \dots & \lambda_{\beta_k} \end{pmatrix}$ és $r = \Lambda$, akkor a lineáris kombináció tesztelését...

- ...kapjuk vissza.

Konkrét példák a fenti sémára

- Ellenőrizhető, hogy ha például...

- ... $\mathbf{R} = (0 \ 0 \ \dots \ 0 \ 1 \ 0 \dots 0)$ és $r = 0$, akkor a t -tesztet ...

- ... $\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ és $r = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$ akkor az ANOVA-t...

- ... $\mathbf{R} = (\lambda_{\beta_1} \ \lambda_{\beta_2} \ \dots \ \lambda_{\beta_k})$ és $r = \Lambda$, akkor a lineáris kombináció tesztelését...

- ...kapjuk vissza.

Konkrét példák a fenti sémára

- Ellenőrizhető, hogy ha például...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 \dots 0 \end{pmatrix}$ és $r = 0$, akkor a t -tesztet ...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ és $\mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$ akkor az ANOVA-t...

- $\dots \mathbf{R} = \begin{pmatrix} \lambda_{\beta_1} & \lambda_{\beta_2} & \dots & \lambda_{\beta_k} \end{pmatrix}$ és $r = \Lambda$, akkor a lineáris kombináció tesztelését...

- ...kapjuk vissza.

Konkrét példák a fenti sémára

- Ellenőrizhető, hogy ha például...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 \dots 0 \end{pmatrix}$ és $r = 0$, akkor a t -tesztet ...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ és $\mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$ akkor az ANOVA-t...

- $\dots \mathbf{R} = \begin{pmatrix} \lambda_{\beta_1} & \lambda_{\beta_2} & \dots & \lambda_{\beta_k} \end{pmatrix}$ és $r = \Lambda$, akkor a lineáris kombináció tesztelését...

- ...kapjuk vissza.

Konkrét példák a fenti sémára

- Ellenőrizhető, hogy ha például...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 \dots 0 \end{pmatrix}$ és $r = 0$, akkor a t -tesztet ...

- $\dots \mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$ és $\mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$ akkor az ANOVA-t...

- $\dots \mathbf{R} = \begin{pmatrix} \lambda_{\beta_1} & \lambda_{\beta_2} & \dots & \lambda_{\beta_k} \end{pmatrix}$ és $r = \Lambda$, akkor a lineáris kombináció tesztelését...

- ...kapjuk vissza.

Speciális esetek

- Ez a képlet viszont *minden* eddig látott dolgot tartalmaz speciális esetként!
- Wald-elven