

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinealidad

Caracterización Diagnóstico

Perturbaciones esféricas

Heteroscedaticidad

Endogeneida

Anexos

References

Econometría Básica

Tópico 3: Violación de supuestos

Luis Chávez

Escuela Profesional de Economía USMP

Lima, 2025



Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterizació: Diagnóstico

Perturbacione

Heteroscedaticidad

Endogeneida

Anexos

References

Contenido

- Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA
- 2 Multicolinealidad Caracterización Diagnóstico
- 3 Perturbaciones esféricas Heteroscedaticidad Autocorrelación
- 4 Endogeneidad
- 6 Anexos



Econometría

Luis Chávez

Modelos ANOVA

Modelos ANCOVA

Diagnóstico

Heteroscedaticidad

Contenido

- Modelos con variables dicótomas Modelos ANOVA
- Multicolinealidad



Generalidades

Econometría

Luis Chávez

Modelos cor variables dicótomas

Modelos ANCOVA

Multicolinealidae

Caracterizació Diagnóstico

Perturbaciones esféricas Heteroscedaticidad

Endogeneida

Anexo

- A veces los modelos econométricos involucran variables cualitativas, a quienes se conocen como variables dicótomas.
- Las que admite más de dos atributos se denominan variables politómicas, pero requieren de cierto tratamiento para su implementación dentro de los modelos.



Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Multicolinealidad

Caracterizació Diagnóstico

Perturbaciones

Heteroscedaticidae

Endogeneid

A

Reference

Definición 1 (variable dummy)

Una variable **dummy** es aquella variable dicotómica, binaria o cero-uno si admite únicamente dos categorías



Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANOVA Modelos ANCOVA

Multicolinealida Caracterización

Diagnóstico

Perturbacione esféricas

Endononald

Lindogeneral

Doforonce

Sea el caso de los salarios de maestros de escuelas públicas de Gujarati and Porter (2010). El modelo base es

$$salar_i = \beta_1 + \beta_2 D_{i2} + \beta_3 D_{i3} + u_i \tag{1}$$

donde $salar_i$ es el salario promedio de los maestros de escuelas públicas del estado i, $D_{i2}\{1=\text{estado noreste o norte-centro},0\}$ y $D_{i3}\{1=\text{estado sur},0\}$. La categoría de comparación es el **oeste**.



Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Multicolinealidad

Caracterizacio Diagnóstico

Perturbacione esféricas

Heteroscedaticidad Autocorrelación

Endogeneida

D (

Esperanzas:

$$E(salar_i|D_{i2}=1, D_{i3}=0) = \beta_1 + \beta_2$$
 (2)

$$E(salar_i|D_{i2}=0, D_{i3}=1) = \beta_1 + \beta_3$$
 (3)

$$E(salar_i|D_{i2} = 0, D_{i3} = 0) = \beta_1$$
 (4)



Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Multicolinealidad

Caracterizació Diagnóstico

Perturbaciones

Heteroscedaticidad

Endogeneida

Anexo

References

Nota 1

Para evitar la trampa de la variable dicótoma, si una variable cualitativa tiene n categorías, sólo se debe añadir n-1 categorías tipo dummy.



Contenido

Econometría

Luis Chávez

Modelos cor variables dicótomas

Modelos ANCOVA

Multicolineal

Caracterizació Diagnóstico

Perturbaciones

Heteroscedaticidad Autocorrelación

Endogeneida

Anexos

References

 Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA

- 2 Multicolinealidad Caracterización Diagnóstico
 - 3 Perturbaciones esféricas Heteroscedaticidad Autocorrelación
 - 4 Endogeneidad
 - 5 Anexos



Econometría

Luis Chávez

variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Multicolinealida

Caracterización Diagnóstico

Perturbacione esféricas

Heteroscedaticidad

Endogeneida

Anexo

Reference

Ejemplo 1

Sea el modelo:

$$ing_i = \beta_1 + \beta_2 educ_{i2} + \beta_3 raza_{i3} + \beta_4 sexo_{i3} + \beta_5 latin_{i4} + u_i$$
 (5)

Calcular la brecha remunerativa de los varones blancos, latinos y no latinos. Reformule el modelo anterior si se añade el tipo de universidad donde estudió el pregrado i (pub_i) y su status de pobreza (pob_i) de su hogar según SISFOH.



Quiebre estructural

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Modelos ANCOVA

Caracterización Diagnóstico

Perturbacione esféricas

Heteroscedaticida Autocorrelación

Endogeneid

Anexos

Reference

Una vez que se identificó quiebre estructural en un modelo estimado, se puede solucionar utilizando variables *dummy*.



Quiebre estructural

Econometría

Luis Chávez

Modelos o variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Multicolinealidae

Caracterizaci Diagnóstico

Perturbaciones esféricas

Heteroscedaticidad Autocorrelación

E-4-----

Anexo

References

Ejemplo 2

Revisita al modelo ahorro-ingreso de Gujarati and Porter (2010).



Integración

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA

Modelos ANCOVA

Caracterización Diagnóstico

Perturbacion

esféricas Heteroscedaticidad

Autocorrelación

Endogeneida

Anexos

References

Actividad 1. Replicar la sección 9.12 de Gujarati and Porter (2010).



Contenido

Econometría

Luis Chávez

variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolineal

Caracterización Diagnóstico

Perturbaciones

Heteroscedaticidad Autocorrelación

Endogeneidad

Anexos

- Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA
- 2 Multicolinealidad
 Caracterización
- 3 Perturbaciones esféricas Heteroscedaticidad Autocorrelación
- 4 Endogeneidad
- 5 Anexos



Naturaleza

Econometría

Luis Chávez

Modelos con variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinealidad

Caracterización

Perturbacione esféricas

Heteroscedaticidad Autocorrelación

Endogeneidad

Anexo

Reference

Conforme a Ragnar Frisch, la **multicolinealidad** puede generar problemas en las estimaciones econométricas.

Definición 2 (multicolinealidad perfecta)

Un modelo presenta multicolinealidad perfecta si las variables independientes puede expresarse como una combinación lineal exacta.

$$\alpha_1 x_{i1} + \dots + \alpha_k x_{ik} = 0 \tag{6}$$

donde $\exists \alpha_j \neq 0$.



Naturaleza

Econometría

Luis Chávez

dicótomas

Modelos ANOVA

Marking the collision

Caracterización

Diagnóstico

Perturbacione esféricas

Heteroscedaticida Autocorrelación

Endogeneidad

Anexo

Reference

Definición 3 (multicolinealidad imperfecta)

Un modelo presenta multicolinealidad imperfecta si las variables independientes puede expresarse como una combinación lineal no exacta.

$$\alpha_1 x_{i1} + \dots + \alpha_k x_{ik} + v_i = 0 (7)$$

donde v_i es un elemento estocástico.



Naturaleza

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealida

Caracterización Diagnóstico

Perturbacion esféricas

Heteroscedaticidae

Heteroscedaticida Autocorrelación

Lildogeneida

- La colinealidad perfecta involucra coeficientes indeterminados y ee infinitos.
 Problema de indeterminación.
- La colinealidad imperfecta involucra coeficientes determinados pero ee explosivos. Problema de precisión.



Mecanismos causales

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterización

Perturbaciones esféricas

esféricas Heteroscedaticidad

Endogeneidad

Anexo

Reference

Mecanismos causales de Montgomery et al. (2021):

- Recopilación de datos (escasa varianza).
- Restricciones en la población de análisis.
- Especificación del modelo.
- Sobredeterminación del modelo.

En time series suele evidenciarse tendencia común.



Estimación

Econometría

Luis Chávez

variables dicótomas

Modelos ANCOVA

Caracterización

Diagnóstico

esféricas Heteroscedaticidad

Endogeneidad

Anexo

- En el caso perfecto, la estimación no es factible. ¿Porqué? Los cambios marginales de una variable es cíclica, por lo que el supuesto cetaris paribus deja de funcionar. No hay una solución única.
- En el caso imperfecto, la estimación es factible pero a veces puede ser ruidosa.



Consecuencias teóricas

Econometría

Luis Chávez

variables dicótomas

Modelos ANCOVA

Caracterización

Diagnóstico

Perturbaciones esféricas

Heteroscedaticida Autocorrelación

Endogeneid

Reference

Dado que se trata de un problema de la muestra,

- Los LSE siguen siendo MELI.
- La varianza de los LSE no necesariamente son pequeños a pesar de ser eficientes.



Consecuencias prácticas

Econometría

Luis Chávez

variables dicótomas

KA IST DE DEL

Caracterización

Diagnóstico

esféricas Heteroscedaticidae

Autocorrelación

Endogeneidad

Anexos

- Varianzas y covarianzas grandes en los LSE.
- Intervalos de confianza amplios.
- Sesgo hacia la no significación de variables.
- R² alto.
- Las estimaciones son sensibles a cambios en los datos.



Consecuencias prácticas

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Multicolinealidad

Caracterización Diagnóstico

Perturbaciones esféricas

Heteroscedaticidad Autocorrelación

Endogeneid:

Anexos

References

Ejemplo 3

Véase el ejemplo 10.1 de Gujarati and Porter (2010).



Contenido

Econometría

Luis Chávez

variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinea

Diagnóstico

Perturbaciones

Heteroscedaticidad

Anexos

References

 Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA

2 Multicolinealidad

Caracterización

Diagnóstico

- ③ Perturbaciones esféricas Heteroscedaticidad Autocorrelación
- 4 Endogeneidad
- 5 Anexos



Identificación

Econometría

Luis Chávez

Modelos con variables dicótomas Modelos ANOVA

Multicolinealida

Caracterizaci

Diagnóstico

esféricas

Heteroscedaticidad

Endamentia d

Anexo

- R² alto pero pocas variables significativas.
- Alta correlación bivariada entre variables independientes.
- Valores propios e índices de condición.
- Tolerancia y VIF.
- Diagramas de dispersión pareadas.



Corrección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealida

Diagnóstico

Diagnostico

esféricas Heteroscedaticida

Endononoidad

^

- No hacer algo.
- Eliminación de variables (overfitting).
- Transformación de variables.
- Mejorar la muestra (tamaño).
- Análisis factorial y ACP.



Corrección

Econometría

Luis Chávez

dicótomas

Modelos ANOVA

Modelos ANCOVA

Caracterizació

Diagnóstico

Perturbaciones

Heteroscedaticidad Autocorrelación

Endogeneida

Anexos

References

Ejemplo 4

Véase el ejemplo de la sección 10.10 de Gujarati and Porter (2010).



Contenido

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterizació Diagnóstico

Perturbacio esféricas

Heteroscedaticidad

Endogeneidad

Anexos

- Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA
- 2 Multicolinealidad Caracterización Diagnóstico
- 3 Perturbaciones esféricas Heteroscedaticidad
- 4 Endogeneidad
- 6 Anexos



Conceptos

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANOVA Modelos ANCOVA

Multicolinealidad

Caracterizació Diagnóstico

Perturbacione esféricas

Heteroscedaticidad

Autocorrelación

Reference

El supuesto de homoscedasticidad asumía que

$$E(u_i^2|x) = \sigma^2, \forall i = 1, ..., n$$
 (8)

Cuando no se conserva, la heteroscedasticidad implica:

$$E(u_i^2|x) = \sigma_i^2, \forall i = 1, ..., n$$
 (9)

¿Gráficamente?



Causas

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealidae Caracterización

Diagnóstico

esféricas

Heteroscedaticidad

Autocorrelación

Endogeneid

Anexos

- Los patrones de comportamiento se van modulando con el paso del tiempo, lo cual reduce volatilidad.
- Las mejoras en los procesos de recopilación de datos está permitiendo reducir errores.
- Presencia de outliers.
- Incorrecta especificación del modelo de regresión (omisión, redundancia o forma funcional).
- Variables alejadas de una distribución normal (asimetría).



Problemas

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinealidad

Caracterización Diagnóstico

Perturbacione esféricas

Heteroscedaticidad Autocorrelación

Endogeneidad

Anexo

- El LSE sigue siendo insesgado en presencia de heteroscedasticidad. Por tanto, asintóticamente seguirá siendo consistente y normalmente distribuido. Sin embargo, ya no es eficiente dentro de la familia de estimadores lineales: la varianza ya no es mínima.
- ¿Qué sigue? MCG



MCG

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANCOVA

Multicolinealidad

Caracterización

Diagnóstico

Perturbacion esféricas

Heteroscedaticidad

Endogeneida

Anexos

References

Dado el modelo estándar,

$$y_i = \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i$$
 (10)

Si se conociera σ_i^2 , la ecuación (10) se puede escribir en forma ponderada:

$$\frac{y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{x_{i2}}{\sigma_i} + \dots + \beta_k \frac{x_{ik}}{\sigma_i} + \frac{u_i}{\sigma_i}$$
(11)

0

$$y_i^* = \beta_1^* + \beta_2^* x_{i2}^* + \dots + \beta_k^* x_{ik}^* + u_i^*$$
 (12)

¿Que pasa con u_i^* ?



Heteroscedasticidad

Econometría

Luis Chávez

Modelos co variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinealidad

Caracterización

Caracterización Diagnóstico

Perturbaciones esféricas

Heteroscedaticidad Autocorrelación

Endogeneid

Anexos

References

De (11), se tiene:

$$\min \sum w_i(u_i)^2 = \sum w_i(y_i - \beta_1 - \beta_2 x_{i2} - \dots - \beta_k x_{ik})^2$$
 (13)

donde $w_i = 1/\sigma_i^2$. El WLSE se puede escribir como:

$$\hat{\beta}_{WLS} = (X'WX)^{-1}X'Wy \tag{14}$$

donde

$$W = egin{bmatrix} w_1 & 0 & \cdots & 0 \ 0 & w_2 & \cdots & 0 \ dots & dots & \ddots & dots \ 0 & 0 & \cdots & w_n \end{bmatrix}$$



Consecuencias

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterización Diagnóstico

Perturbación esféricas

Heteroscedaticidad

- IC grandes (WLS).
- Sesgo hacia la no significancia de variables (WLS).
- La estimación por OLS en presencia de heteroscedasticidad generará conclusiones erróneas (grave).



Detección

Econometría

Luis Chávez

dicótomas

Modelos ANOVA

Multicolinealidae

Caracterizació Diagnóstico

esféricas Heteroscedaticidad

Autocorrelación

Endogeneidad

Anexo

- Graficar los residuos al cuadrado y las predicciones.
- Prueba de Breusch-Pagan.
- Test de White.
- Prueba de rango de Szroeter.
- Test de Cameron y Trivedi.
- ..



Corrección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

withticonnealida

Diagnóstico

Perturbacione esféricas

Heteroscedaticidad

Autocorrelación

.

References

Ejemplo 4

Véase el ejemplo 11.7 de Gujarati and Porter (2010).



Corrección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealidae

Caracterizació Diagnóstico

Perturbaciones esféricas

esféricas Heteroscedaticidad

Autocorrelación

Endogeneida

Anexos

- WLS (si se conoce σ_i^2).
- Heteroscedasticidad de White¹ (si no se conoce σ_i^2).
- Imponer otros supuestos razonables sobre los errores.

¹En stata de implementa con *vce*(*robust*).



Contenido

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterizació Diagnóstico

Perturbaciones esféricas

Autocorrelación

Endogeneidad

Anexos

- Modelos con variables dicótomas Modelos ANOVA Modelos ANCOVA
- MulticolinealidadCaracterizaciónDiagnóstico
- 3 Perturbaciones esféricas Heteroscedaticidad Autocorrelación
- 4 Endogeneidad
- 6 Anexos



Conceptos

Econometría Luis Chávez

. . . .

variables dicótomas

Modelos ANOVA Modelos ANCOVA

Multicolinealida

Caracterización

Diagnóstico

esféricas Heteroscedaticida

Autocorrelación

Endogeneida

Anexos

References

Otros supuesto gaussiano fue:

$$cov(u_i, u_j | x_i, x_j) = E(u_i, u_j) = 0, \ \forall i \neq j$$
 (15)

Si no se preserva, se tiene autocorrelación²:

$$E(u_i, u_j) \neq 0 \ \forall i \neq j \tag{16}$$

²Típicamente es un problema en series temporales.



Causas

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

KALINDA - P. . . P. . .

Caracterizació Diagnóstico

esféricas

Heteroscedaticida Autocorrelación

Endogeneidad

Anexo

- Inercia (naturalidad).
- Sesgo de especificación.
- Rezagos (en ts).
- Corrección al efectuar AED.
- No estacionariedad (en ts).



Consecuencias

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealida

Caracterizacio Diagnóstico

esféricas Heteroscedaticida

Heteroscedaticida Autocorrelación

Endogeneidad

Anexo

- Inercia (naturalidad).
- Varianza residual implosiva.
- Pruebas t y F inválidas: sesgo hacia la significación de variables.
- R^2 alto.
- No estacionariedad (en ts).



Detección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Multicolinealida

Caracterización Diagnóstico

Perturbacione esféricas

Autocorrelación

Anavos

- Gráficos de los residuos.
- Durbin-Watson (en ts). Valores cercanos a 2 reflejan autocorrelación, mientras que cercanos a 0 o 4 rechazan esa conjetura³.
- Breusch-Godfrey (en ts).
- ..

³Hay zonas de indecisión



Corrección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Multicolinealid

Caracterización Diagnóstico

Perturbaciones esféricas

Heteroscedaticida Autocorrelación

Endogeneida

Anexos

- GLS (si la matriz V-C es conocida se conoce como FGLS y, en caso contrario, NFGLS).
- Método Newey-West (asintóticamente).



GLS

Econometría

Luis Chávez

Modelos con variables dicótomas Modelos ANOVA

Caracterización

Diagnóstico

esféricas

Heteroscedaticidae

Autocorrelación

References

A partir del modelo base,

$$y_i = X\beta + u \tag{17}$$

$$E(uu'|X) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = \Omega$$

$$\Omega^{1/2} y_i = \Omega^{1/2} X \beta + \Omega^{1/2} u \tag{18}$$

Luego, se tiene el GLSE:

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$
(19)



Conceptos

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterización

Diagnóstico

esféricas

Autocorrelación

 ${\bf Endogeneidad}$

Anexos

Reference

¿Qué pasa si las variables independientes están correlacionadas con el error?

$$plim\hat{eta}
eqeta$$

(20)

Se trata de **endogeneidad**.



Corrección

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA Modelos ANCOVA

Multicolinealidae

Caracterizació Diagnóstico

esféricas

Heteroscedatició Autocorrelación

Endogeneidad

Anexos

References

 En el caso de identificación, se puede usar IV. Todo instrumento debe garantizar:

$$cov(z_i, u_i) = 0$$
 (ortogonalidad)
 $cov(z_i, x_i^k) \neq 0$ (relevancia)

• Usar multiepáticos: 2SLS o 3SLS.



Referencias

Econometría

Luis Chávez

variables dicótomas Modelos ANOVA

Modelos ANCOVA

Caracterizació Diagnóstico

Perturbaciones

Heteroscedaticida Autocorrelación

Anexo

References

Gujarati, P. and Porter, D. (2010). *Econometría*. McGraw-Hill, 5th edition. Montgomery, D., Peck, E. A., and Vining, G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.