Clase auxiliar 11

Aplicaciones de probabilidades y estadística en gestión

Departamento de Ingeniería Civil Industrial
Universidad de Chile

Ronald Leblebici Garo

19 de noviembre de 2017

Antes de imprimir esta presentación, piense si es realmente necesario.

Comente 1: efectos fijos y aleatorios

En un modelo de regresión lineal clásico con datos de panel, se estudia los rendimientos individuales de futbolistas profesionales (variable y) explicados por características exógenas (variables x). Si el campeonato es corto se prefiere efectos aleatorios y si el campeonato es largo se prefiere efectos fijos.

Comente 1: efectos fijos y aleatorios

En un modelo de regresión lineal clásico con datos de panel, se estudia los rendimientos individuales de futbolistas profesionales (variable y) explicados por características exógenas (variables x). Si el campeonato es corto se prefiere efectos aleatorios y si el campeonato es largo se prefiere efectos fijos.

Verdadero. Si T es grande el efecto fijo es preferido porque se puede estimar con un número mayor de observaciones. Con T pequeño es preferible estimar efectos aleatorios.

Comente 2: convergencia en distribución y en probabilidad

Convergencia en distribución implica convergencia en probabilidad.

Comente 2: convergencia en distribución y en probabilidad

Convergencia en distribución implica convergencia en probabilidad.

- **Convergencia en probabilidad:** $\hat{\theta}_N \xrightarrow{p} \theta$ si $\mathbb{P}(|\hat{\theta}_N - \theta| > \varepsilon)$
- **Convergencia en distribución:** $F_N \xrightarrow{p} F$ si $F_N(x) \to F(x), \forall x$

Comente 2: convergencia en distribución y en probabilidad

Convergencia en distribución implica convergencia en probabilidad.

- **Convergencia en probabilidad:** $\hat{\theta}_N \xrightarrow{p} \theta$ si $\mathbb{P}(|\hat{\theta}_N - \theta| > \varepsilon)$
- **Convergencia en distribución:** $F_N \xrightarrow{p} F$ si $F_N(x) \rightarrow F(x), \forall x$

Falso. La única excepción es la convergencia en distribución a una constante (distribución degenerada) en cuyo caso los conceptos son equivalentes.

Comente 3: test de endogeneidad

Para testear problemas de regresores endógenos, se debe regresionar los residuos $\hat{\varepsilon}$ en función de las variables explicativas x, es decir, $\hat{\varepsilon}_i = \alpha_i + x_i^{'}\gamma + u_i$. Si rechazo $\gamma = 0$, no hay problema de endogeneidad.

Comente 3: test de endogeneidad

Para testear problemas de regresores endógenos, se debe regresionar los residuos $\hat{\varepsilon}$ en función de las variables explicativas x, es decir, $\hat{\varepsilon}_i = \alpha_i + x_i^{'}\gamma + u_i$. Si rechazo $\gamma = 0$, no hay problema de endogeneidad.

Falso. Los coeficientes MCO estimados fueron generados para garantizar $X^{'}\hat{\varepsilon} = 0$. Aún en presencia de endogeneidad, siempre se obtendrá que $\gamma = 0$.

Comente 4: heterocedasticidad

En un modelo con heterocedasticidad no es posible estimar un modelo donde la varianza de cada residuo sea distinta.

Comente 4: heterocedasticidad

En un modelo con heterocedasticidad no es posible estimar un modelo donde la varianza de cada residuo sea distinta.

Falso. El modelo con N varianzas distintas se puede estimar y se denomina Mínimos Cuadrados Ponderados, donde la ponderación es el inverso de la varianza de cada observación.

$$\hat{\beta}_{MCP} = (X^T \hat{\Omega}^{-1} X)^{-1} X^T \hat{\Omega}^{-1} Y$$

$\hat{\Omega} = diag(\hat{\sigma}_i^2)$

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.
2. **Matriz X^TX invertible:** $rg(X^TX) = K$. Se puede verificar.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.
2. **Matriz $X^T X$ invertible:** $rg(X^T X) = K$. Se puede verificar.
3. **Errores nulos en esperanza:** $\mathbb{E}(U) = 0$. Supuesto aceptable.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.
2. **Matriz X^TX invertible:** $rg(X^TX) = K$. Se puede verificar.
3. **Errores nulos en esperanza:** $\mathbb{E}(U) = 0$. Supuesto aceptable.
4. **Independencia y homocedasticidad:** $\mathbb{V}(U) = \sigma^2 I_N$.
   Se puede testear.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.
2. **Matriz $X^T X$ invertible:** $rg(X^T X) = K$. Se puede verificar.
3. **Errores nulos en esperanza:** $\mathbb{E}(U) = 0$. Supuesto aceptable.
4. **Independencia y homocedasticidad:** $\mathbb{V}(U) = \sigma^2 I_N$.
   Se puede testear.
5. **Exogeneidad:** $\mathbb{E}(X^T U) = 0$. Supuesto fuerte y problemático.

Comente 5: ¿MCO es MELI?

MCO es el mejor estimador lineal insesgado si y solo si los errores ε del modelo distribuyen normales con esperanza cero y $\mathbb{E}(\varepsilon\varepsilon^{'}) = \sigma^2 Id_N$.

Recordemos los supuestos de MCO:

1. **Linealidad de parámetros:** $Y = X\beta + U$.
   Una vez estimado el modelo, podemos realizar tests de hipótesis y obtener métricas de bondad de ajuste.
2. **Matriz $X^T X$ invertible:** $rg(X^T X) = K$. Se puede verificar.
3. **Errores nulos en esperanza:** $\mathbb{E}(U) = 0$. Supuesto aceptable.
4. **Independencia y homocedasticidad:** $\mathbb{V}(U) = \sigma^2 I_N$.
   Se puede testear.
5. **Exogeneidad:** $\mathbb{E}(X^T U) = 0$. Supuesto fuerte y problemático.

Falso. El teorema de Gauss-Markov no requiere una distribución particular de los errores U, basta con que se cumplan los supuestos de MCO para garantizar que $\hat{\beta}_{MCO}$ es MELI.

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Supongamos que $\hat{\beta}$ es consistente ($\hat{\beta} \xrightarrow{p} \beta$) e insesgado ($\mathbb{E}(\hat{\beta}) = \beta$).

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Supongamos que $\hat{\beta}$ es consistente ($\hat{\beta} \xrightarrow{p} \beta$) e insesgado ($\mathbb{E}(\hat{\beta}) = \beta$).

Veamos que $\widetilde{\beta}$ es sesgado e inconsistente:

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Supongamos que $\hat{\beta}$ es consistente ($\hat{\beta} \xrightarrow{p} \beta$) e insesgado ($\mathbb{E}(\hat{\beta}) = \beta$).

Veamos que $\widetilde{\beta}$ es sesgado e inconsistente:

- **Sesgado:** $\mathbb{E}(\widetilde{\beta}) = \mathbb{E}\left(\frac{N}{N+1}\hat{\beta}\right) = \frac{N}{N+1}\mathbb{E}(\hat{\beta}) = \frac{N}{N+1}\beta \neq \beta$.

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Supongamos que $\hat{\beta}$ es consistente ($\hat{\beta} \xrightarrow{p} \beta$) e insesgado ($\mathbb{E}(\hat{\beta}) = \beta$).

Veamos que $\widetilde{\beta}$ es sesgado e inconsistente:

- **Sesgado:** $\mathbb{E}(\widetilde{\beta}) = \mathbb{E}\left(\frac{N}{N+1}\hat{\beta}\right) = \frac{N}{N+1}\mathbb{E}(\hat{\beta}) = \frac{N}{N+1}\beta \neq \beta$.
- **Inconsistente:** $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta} \xrightarrow{p} \beta$

Comente 6: sesgo y consistencia

Un estimador sesgado puede ser consistente.

Verdadero. El sesgo que es la desviación entre la esperanza y el parámetro desconocido puede desaparecer en una muestra infinita.

Ejemplo: $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta}$.

Donde $\widetilde{\beta}$ y $\hat{\beta}$ son estimadores de β.

Supongamos que $\hat{\beta}$ es consistente ($\hat{\beta} \xrightarrow{p} \beta$) e insesgado ($\mathbb{E}(\hat{\beta}) = \beta$).

Veamos que $\widetilde{\beta}$ es sesgado e inconsistente:

- **Sesgado:** $\mathbb{E}(\widetilde{\beta}) = \mathbb{E}\left(\frac{N}{N+1}\hat{\beta}\right) = \frac{N}{N+1}\mathbb{E}(\hat{\beta}) = \frac{N}{N+1}\beta \neq \beta$.
- **Inconsistente:** $\widetilde{\beta} = \frac{N}{N+1}\hat{\beta} \xrightarrow{p} \beta$

En el último cálculo hemos ocupado el teorema de Slutzky, pues $\frac{N}{N+1} \xrightarrow{p} \beta$ y $\hat{\beta} \xrightarrow{p} \beta$.

Comente 7: endogeneidad y efectos fijos

Utilizando datos de panel y en un modelo de regresión lineal, nunca el efecto fijo podrá solucionar el problema de endogeneidad por la variable no observable correlacionada con los regresores observables.

Comente 7: endogeneidad y efectos fijos

Utilizando datos de panel y en un modelo de regresión lineal, nunca el efecto fijo podrá solucionar el problema de endogeneidad por la variable no observable correlacionada con los regresores observables.

Falso. Si los no observables son constantes (invariantes) en el tiempo, entoces el incluir efectos fijos por individuo soluciona el problema de endogeneidad.

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Falso. La constante es primordial para testear la nula de homocedasticidad, $H_0 : \gamma_j = 0$, $\forall j$.

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Falso. La constante es primordial para testear la nula de homocedasticidad, $H_0 : \gamma_j = 0, \forall j$.

Si se propone el modelo $\hat{\varepsilon}_i{}^2 = \sigma^2 + \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$ y no se puede rechazar que $\gamma_j = 0$, entonces se supone que:

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Falso. La constante es primordial para testear la nula de homocedasticidad, $H_0 : \gamma_j = 0, \; \forall j$.

Si se propone el modelo $\hat{\varepsilon}_i{}^2 = \sigma^2 + \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$ y no se puede rechazar que $\gamma_j = 0$, entonces se supone que:

$$\hat{\varepsilon}_i{}^2 = \sigma^2 + u_i$$

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Falso. La constante es primordial para testear la nula de homocedasticidad, $H_0 : \gamma_j = 0, \forall j$.

Si se propone el modelo $\hat{\varepsilon}_i{}^2 = \sigma^2 + \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$ y no se puede rechazar que $\gamma_j = 0$, entonces se supone que:

$$\hat{\varepsilon}_i{}^2 = \sigma^2 + u_i$$

$$\Rightarrow \mathbb{E}(\hat{\varepsilon}_i{}^2) = \sigma^2$$

Comente 8: test de heterocedasticidad

Para testear problemas de heterocedasticidad una alternativa es regresionar los cuadrados de los residuos, $\hat{\varepsilon}^2$, en función de los cuadrados de las K variables explicativas , es decir, $\hat{\varepsilon}_i{}^2 = \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$, , donde se debe omitir la constante para todos los individuos $i = \{1, .., N\}$.

Falso. La constante es primordial para testear la nula de homocedasticidad, $H_0 : \gamma_j = 0, \forall j$.

Si se propone el modelo $\hat{\varepsilon}_i{}^2 = \sigma^2 + \sum_{j=1}^{K} \gamma_j x_{ji}^2 + u_i$ y no se puede rechazar que $\gamma_j = 0$, entonces se supone que:

$$\hat{\varepsilon}_i{}^2 = \sigma^2 + u_i$$

$$\Rightarrow \mathbb{E}(\hat{\varepsilon}_i{}^2) = \sigma^2$$

El decir, en valor esperado se cumple la homocedasticidad.

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

$$Y_{nt} = \alpha_n + \delta_t + X_{nt}\beta + U_{nt} \Rightarrow Y_n = \alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n$$

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

$$Y_{nt} = \alpha_n + \delta_t + X_{nt}\beta + U_{nt} \Rightarrow Y_n = \alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n$$

$$\Rightarrow y_{nt} = Y_{nt} - Y_n = (\alpha_n + \delta_t + X_{nt}\beta + U_{nt}) - (\alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n)$$

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

$$Y_{nt} = \alpha_n + \delta_t + X_{nt}\beta + U_{nt} \Rightarrow Y_n = \alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n$$

$$\Rightarrow y_{nt} = Y_{nt} - Y_n = (\alpha_n + \delta_t + X_{nt}\beta + U_{nt}) - (\alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n)$$

$$\Rightarrow y_{nt} = \underbrace{(\alpha_n - \alpha_n)}_{=0} + (\delta_t - \bar{\delta}) + \underbrace{(X_{nt} - \bar{X}_n)}_{x_{nt}}\beta + \underbrace{(U_{nt} - U_n)}_{u_n t}$$

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

$$Y_{nt} = \alpha_n + \delta_t + X_{nt}\beta + U_{nt} \Rightarrow Y_n = \alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n$$

$$\Rightarrow y_{nt} = Y_{nt} - Y_n = (\alpha_n + \delta_t + X_{nt}\beta + U_{nt}) - (\alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n)$$

$$\Rightarrow y_{nt} = \underbrace{(\alpha_n - \alpha_n)}_{=0} + (\delta_t - \bar{\delta}) + \underbrace{(X_{nt} - \bar{X}_n)}_{x_{nt}}\beta + \underbrace{(U_{nt} - U_n)}_{u_n t}$$

$$\Rightarrow y_{nt} = \delta_t - \bar{\delta} + x_{nt}\beta + u_{nt}$$

Comente 9: estimación de efectos fijos

Suponga usted dispone de datos de panel balanceado con 10 millones de individuos observados en $T \geq 2$ períodos de tiempo. Al estimar el modelo con efecto fijo, necesariamente se deben estimar los 10 millones de coeficientes para cada efecto fijo.

$$Y_{nt} = \alpha_n + \delta_t + X_{nt}\beta + U_{nt} \Rightarrow Y_n = \alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n$$

$$\Rightarrow y_{nt} = Y_{nt} - Y_n = (\alpha_n + \delta_t + X_{nt}\beta + U_{nt}) - (\alpha_n + \bar{\delta} + \bar{X}_n\beta + \bar{U}_n)$$

$$\Rightarrow y_{nt} = \underbrace{(\alpha_n - \alpha_n)}_{=0} + (\delta_t - \bar{\delta}) + \underbrace{(X_{nt} - \bar{X}_n)}_{x_{nt}}\beta + \underbrace{(U_{nt} - U_n)}_{u_n t}$$

$$\Rightarrow y_{nt} = \delta_t - \bar{\delta} + x_{nt}\beta + u_{nt}$$

Falso. Se puede restar la media de cada individuo y por lo tanto los efectos fijos se cancelan.

Comente 10: un único regresor endógeno

Un modelo de regresión lineal posee $K - 1$ regresores exógenos y solo un regresor endógeno. El estimador de MCO sera insesgado para los $K - 1$ coeficientes de regresores exógenos.

Comente 10: un único regresor endógeno

Un modelo de regresión lineal posee $K-1$ regresores exógenos y solo un regresor endógeno. El estimador de MCO sera insesgado para los $K-1$ coeficientes de regresores exógenos.

$$\mathbb{E}(\hat{\beta}|X) = \beta + \mathbb{E}((X^{'}X)^{-1}X^{'}U)$$

Comente 10: un único regresor endógeno

Un modelo de regresión lineal posee $K-1$ regresores exógenos y solo un regresor endógeno. El estimador de MCO sera insesgado para los $K-1$ coeficientes de regresores exógenos.

$$\mathbb{E}(\hat{\beta}|X) = \beta + \mathbb{E}((X'X)^{-1}X'U)$$

Sean $(X'X)^{-1} = \begin{pmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & & \vdots \\ a_{K1} & \cdots & a_{KK} \end{pmatrix}$ y $X'U = \begin{pmatrix} \sum_{i=1}^{K} X_{i1}U_i \\ \vdots \\ \sum_{i=1}^{K} X_{iK}U_i \end{pmatrix}$

Después de un poco de álgebra se llega a que:

Comente 10: un único regresor endógeno

Un modelo de regresión lineal posee $K-1$ regresores exógenos y solo un regresor endógeno. El estimador de MCO sera insesgado para los $K-1$ coeficientes de regresores exógenos.

$$\mathbb{E}(\hat{\beta}|X) = \beta + \mathbb{E}((X'X)^{-1}X'U)$$

Sean $(X'X)^{-1} = \begin{pmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & & \vdots \\ a_{K1} & \cdots & a_{KK} \end{pmatrix}$ y $X'U = \begin{pmatrix} \sum_{i=1}^{K} X_{i1}U_i \\ \vdots \\ \sum_{i=1}^{K} X_{iK}U_i \end{pmatrix}$

Después de un poco de álgebra se llega a que:

$$\mathbb{E}[(X'X)^{-1}X'U] = \begin{pmatrix} a_{1K}\sum_{i=1}^{N}\mathbb{E}(x_{iK}u_i) \\ \vdots \\ a_{KK}\sum_{i=1}^{N}\mathbb{E}(x_{iK}u_i) \end{pmatrix}$$

Comente 10: un único regresor endógeno

Un modelo de regresión lineal posee $K - 1$ regresores exógenos y solo un regresor endógeno. El estimador de MCO sera insesgado para los $K - 1$ coeficientes de regresores exógenos.

Falso. En general, la presencia de un regresor endógeno implica sesgo en todos los coeficientes.

Comente 11: ¿sesgo al agregar variables irrelevantes?

En el modelo clásico de mínimos cuadrados ordinarios, incluir variables irrelevantes exógenas jamás causará sesgo en las estimaciones de β.

Verdadero. En valor esperado, el coeficiente de una variable irrelevante será 0, por lo que al estimar el modelo, los coeficientes de las otras variables no serán afectados.

$$\hat{\beta} = (X^T X)^{-1} X^T Y = (X^T X)^{-1} X^T ([X_1 X_2] \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + U)$$

Comente 11: ¿sesgo al agregar variables irrelevantes?

En el modelo clásico de mínimos cuadrados ordinarios, incluir variables irrelevantes exógenas jamás causará sesgo en las estimaciones de β.

Verdadero. En valor esperado, el coeficiente de una variable irrelevante será 0, por lo que al estimar el modelo, los coeficientes de las otras variables no serán afectados.

$$\hat{\beta} = (X^T X)^{-1} X^T Y = (X^T X)^{-1} X^T ([X_1 X_2] \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + U)$$

$$= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + (X^T X)^{-1} X^T U$$

Comente 11: ¿sesgo al agregar variables irrelevantes?

En el modelo clásico de mínimos cuadrados ordinarios, incluir variables irrelevantes exógenas jamás causará sesgo en las estimaciones de β.

Verdadero. En valor esperado, el coeficiente de una variable irrelevante será 0, por lo que al estimar el modelo, los coeficientes de las otras variables no serán afectados.

$$\hat{\beta} = (X^T X)^{-1} X^T Y = (X^T X)^{-1} X^T ([X_1 X_2] \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + U)$$

$$= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + (X^T X)^{-1} X^T U$$

$$\Rightarrow \mathbb{E}(\hat{\beta}|X) = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ 0 \end{pmatrix}$$

Comente 11: ¿sesgo al agregar variables irrelevantes?

En el modelo clásico de mínimos cuadrados ordinarios, incluir variables irrelevantes exógenas jamás causará sesgo en las estimaciones de β.

Verdadero. En valor esperado, el coeficiente de una variable irrelevante será 0, por lo que al estimar el modelo, los coeficientes de las otras variables no serán afectados.

$$\hat{\beta} = (X^TX)^{-1}X^TY = (X^TX)^{-1}X^T([X_1X_2]\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + U)$$

$$= \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} + (X^TX)^{-1}X^TU$$

$$\Rightarrow \mathbb{E}(\hat{\beta}|X) = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} = \begin{pmatrix}\beta_1\\ 0\end{pmatrix}$$

$$\Rightarrow \mathbb{E}(\hat{\beta}_1) = \beta_1$$

Comente 12

Es deseable que el porcentaje de varianza del ECM atribuido al componente de covarianza sea lo más bajo posible.

Comente 12

Es deseable que el porcentaje de varianza del *ECM* atribuido al componente de covarianza sea lo más bajo posible.

Falso. El error cuadrático medio se puede descomponer como:

$$ECM = \underbrace{(\bar{Y} - \bar{\hat{Y}})^2}_{A} + \underbrace{(DS(Y) - DS(\hat{Y}))^2}_{B} + \underbrace{2[1 - Corr(Y, \hat{Y})]DS(Y)DS(}_{C}$$

Comente 12

Es deseable que el porcentaje de varianza del *ECM* atribuido al componente de covarianza sea lo más bajo posible.

Falso. El error cuadrático medio se puede descomponer como:

$$ECM = \underbrace{(\bar{Y} - \bar{\hat{Y}})^2}_{A} + \underbrace{(DS(Y) - DS(\hat{Y}))^2}_{B} + \underbrace{2[1 - Corr(Y, \hat{Y})]DS(Y)DS(}_{C}$$

Donde A es un sesgo sistemático de la predicción y B es una diferencia sistemática de la volatilidad de la predicción.

Comente 12

Es deseable que el porcentaje de varianza del ECM atribuido al componente de covarianza sea lo más bajo posible.

Falso. El error cuadrático medio se puede descomponer como:

$$ECM = \underbrace{(\bar{Y} - \bar{\hat{Y}})^2}_{A} + \underbrace{(DS(Y) - DS(\hat{Y}))^2}_{B} + \underbrace{2[1 - Corr(Y, \hat{Y})]DS(Y)DS(}_{C}$$

Donde A es un sesgo sistemático de la predicción y B es una diferencia sistemática de la volatilidad de la predicción.

- Si A/ECM es alto, el modelo tiene un problema serio (está sesgado).

Comente 12

Es deseable que el porcentaje de varianza del *ECM* atribuido al componente de covarianza sea lo más bajo posible.

Falso. El error cuadrático medio se puede descomponer como:

$$ECM = \underbrace{(\bar{Y} - \bar{\hat{Y}})^2}_{A} + \underbrace{(DS(Y) - DS(\hat{Y}))^2}_{B} + \underbrace{2[1 - Corr(Y, \hat{Y})]DS(Y)DS(}_{C}$$

Donde A es un sesgo sistemático de la predicción y B es una diferencia sistemática de la volatilidad de la predicción.

- Si A/ECM es alto, el modelo tiene un problema serio (está sesgado).
- Si B/ECM es alto, el modelo tiene un problema, aunque menos grave (existe poca variedad en las predicciones $\hat{Y}$).

Comente 12

Es deseable que el porcentaje de varianza del *ECM* atribuido al componente de covarianza sea lo más bajo posible.

Falso. El error cuadrático medio se puede descomponer como:

$$ECM = \underbrace{(\bar{Y} - \bar{\hat{Y}})^2}_{A} + \underbrace{(DS(Y) - DS(\hat{Y}))^2}_{B} + \underbrace{2[1 - Corr(Y, \hat{Y})]DS(Y)DS(}_{C}$$

Donde A es un sesgo sistemático de la predicción y B es una diferencia sistemática de la volatilidad de la predicción.

- Si A/ECM es alto, el modelo tiene un problema serio (está sesgado).
- Si B/ECM es alto, el modelo tiene un problema, aunque menos grave (existe poca variedad en las predicciones $\hat{Y}$).
- Por descarte, lo preferible sería que C/ECM tenga el mayor porcentaje posible (aunque nos gustaría que ECM en su conjunto sea lo más bajo posible).

Comente 13: ¿es siempre mejor agregar una variable?

Si existe una variable X_{K+1} que no estamos seguros de incluir en un modelo, es siempre mejor incluirla.

Comente 13: ¿es siempre mejor agregar una variable?

Si existe una variable X_{K+1} que no estamos seguros de incluir en un modelo, es siempre mejor incluirla.

Falso. El efecto de agregar una variable no es claro y dependería de su capacidad explicativa y de qué tan colineal es esta con otros regresores del modelo.

Comente 13: ¿es siempre mejor agregar una variable?

Si existe una variable X_{K+1} que no estamos seguros de incluir en un modelo, es siempre mejor incluirla.

Falso. El efecto de agregar una variable no es claro y dependería de su capacidad explicativa y de qué tan colineal es esta con otros regresores del modelo.

Si la variable es relevante y se omite, se tendrá sesgo e inconsistencia, aunque se los estimadores serán más eficientes $(\downarrow \widehat{\mathbb{V}}(\hat{\beta}_k))$.

Comente 13: ¿es siempre mejor agregar una variable?

Si existe una variable X_{K+1} que no estamos seguros de incluir en un modelo, es siempre mejor incluirla.

Falso. El efecto de agregar una variable no es claro y dependería de su capacidad explicativa y de qué tan colineal es esta con otros regresores del modelo.

Si la variable es relevante y se omite, se tendrá sesgo e inconsistencia, aunque se los estimadores serán más eficientes ($\downarrow \widehat{\mathbb{V}}(\hat{\beta}_k)$).

Si la variable es irrelevante y se incluye, $\hat{\beta}$ será insesgado, aunque se perderá eficiencia ($\uparrow \widehat{\mathbb{V}}(\hat{\beta}_k)$).

Comente 13: ¿es siempre mejor agregar una variable?

Si existe una variable X_{K+1} que no estamos seguros de incluir en un modelo, es siempre mejor incluirla.

Falso. El efecto de agregar una variable no es claro y dependería de su capacidad explicativa y de qué tan colineal es esta con otros regresores del modelo.

Si la variable es relevante y se omite, se tendrá sesgo e inconsistencia, aunque se los estimadores serán más eficientes $(\downarrow \widehat{\mathbb{V}}(\hat{\beta}_k))$.

Si la variable es irrelevante y se incluye, $\hat{\beta}$ será insesgado, aunque se perderá eficiencia $(\uparrow \widehat{\mathbb{V}}(\hat{\beta}_k))$.

Lo anterior se puede verificar a través de la siguiente fórmula:

$$\mathbb{V}(\hat{\beta}_k) = \frac{\mathbb{V}(Y)(1 - R^2)}{(N - K)\mathbb{V}(X_k)(1 - R^2_{k,-k})}$$

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Verdadero.

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Verdadero.

En un modelo estacionario, el mejor predictor sería el promedio de las observaciones ($y_t = \bar{y} + u_t$), por lo que el coeficiente R^2 queda descrito de la forma tradicional:

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Verdadero.

En un modelo estacionario, el mejor predictor sería el promedio de las observaciones ($y_t = \bar{y} + u_t$), por lo que el coeficiente R^2 queda descrito de la forma tradicional:

$$R^2 = 1 - \left[\sum_{i=1}^{N}(y_i - \hat{y}_i)^2\right] / \left[\sum_{i=1}^{N}(y_i - \bar{y}_i)^2\right]$$

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Verdadero.

En un modelo estacionario, el mejor predictor sería el promedio de las observaciones ($y_t = \bar{y} + u_t$), por lo que el coeficiente R^2 queda descrito de la forma tradicional:

$$R^2 = 1 - \left[\sum_{i=1}^{N}(y_i - \hat{y}_i)^2\right] / \left[\sum_{i=1}^{N}(y_i - \bar{y}_i)^2\right]$$

En una serie de tiempo, el promedio muestral no será un buen predictor. En este caso, un "camino aleatorio" ($y_t = y_{t-1} + u_t$) podría ser un mejor modelo alternativo para explicar y_t.

Comente 14: R^2 en series de tiempo

En un modelo de series de tiempo $R^2 = 1 - \frac{\sum_{t=1}^{T}(y-\hat{y})^2}{\sum_{t=1}^{T}(y-\bar{y})^2}$ resulta ser un indicador de ajuste engañoso.

Verdadero.

En un modelo estacionario, el mejor predictor sería el promedio de las observaciones ($y_t = \bar{y} + u_t$), por lo que el coeficiente R^2 queda descrito de la forma tradicional:

$$R^2 = 1 - \left[\sum_{i=1}^{N}(y_i - \hat{y}_i)^2\right] / \left[\sum_{i=1}^{N}(y_i - \bar{y}_i)^2\right]$$

En una serie de tiempo, el promedio muestral no será un buen predictor. En este caso, un "camino aleatorio" ($y_t = y_{t-1} + u_t$) podría ser un mejor modelo alternativo para explicar y_t.

$$R^2 = 1 - \left[\sum_{t=1}^{T}(y_t - \hat{y}_t)^2\right] / \left[\sum_{t=1}^{T}(y_t - y_{t-1})^2\right]$$

Comente 15: ¿Los errores siempre suman cero?

Los errores poblacionales de un modelo de regresión múltiple necesariamente deben sumar cero.

Falso. Los residuos o errores poblacionales son estocásticos y deben satisfacer que su valor esperado sea 0. En una muestra finita, pueden no sumar cero.

Comente 15: ¿Los errores siempre suman cero?

Los errores poblacionales de un modelo de regresión múltiple necesariamente deben sumar cero.

Falso. Los residuos o errores poblacionales son estocásticos y deben satisfacer que su valor esperado sea 0. En una muestra finita, pueden no sumar cero. Lo que si debe sumar cero son los errores muestrales de una regresión muestral por MCO. En efecto:

Comente 15: ¿Los errores siempre suman cero?

Los errores poblacionales de un modelo de regresión múltiple necesariamente deben sumar cero.

Falso. Los residuos o errores poblacionales son estocásticos y deben satisfacer que su valor esperado sea 0. En una muestra finita, pueden no sumar cero. Lo que si debe sumar cero son los errores muestrales de una regresión muestral por MCO. En efecto:

$$\sum_{i=1}^{N} \hat{u}_i^2 = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \sum_{k=1}^{K} \hat{\beta}_k x_{ik})^2$$

Comente 15: ¿Los errores siempre suman cero?

Los errores poblacionales de un modelo de regresión múltiple necesariamente deben sumar cero.

Falso. Los residuos o errores poblacionales son estocásticos y deben satisfacer que su valor esperado sea 0. En una muestra finita, pueden no sumar cero. Lo que si debe sumar cero son los errores muestrales de una regresión muestral por MCO. En efecto:

$$\sum_{i=1}^{N} \hat{u}_i^2 = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \sum_{k=1}^{K} \hat{\beta}_k x_{ik})^2$$

$$\Rightarrow \frac{\partial \hat{u}_i^2}{\partial \hat{\beta}_0} = -2 \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \sum_{k=1}^{K} \beta_k x_{ik}) \stackrel{!}{=} 0$$

Comente 15: ¿Los errores siempre suman cero?

Los errores poblacionales de un modelo de regresión múltiple necesariamente deben sumar cero.

Falso. Los residuos o errores poblacionales son estocásticos y deben satisfacer que su valor esperado sea 0. En una muestra finita, pueden no sumar cero. Lo que si debe sumar cero son los errores muestrales de una regresión muestral por MCO. En efecto:

$$\sum_{i=1}^{N} \hat{u}_i^2 = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \sum_{k=1}^{K} \hat{\beta}_k x_{ik})^2$$

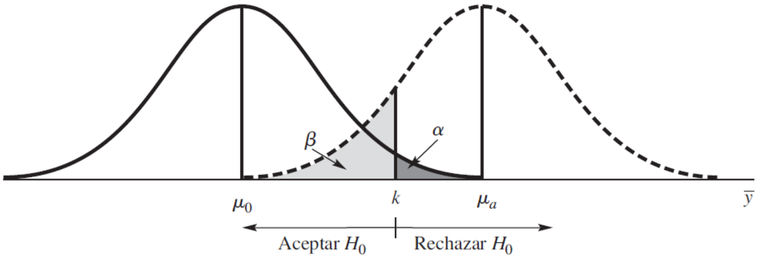
$$\Rightarrow \frac{\partial \hat{u}_i^2}{\partial \hat{\beta}_0} = -2 \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \sum_{k=1}^{K} \beta_k x_{ik}) \stackrel{!}{=} 0$$

$$\Rightarrow \sum_{i=1}^{N} \hat{u}_i = 0$$

Comente 16: confianza y potencia

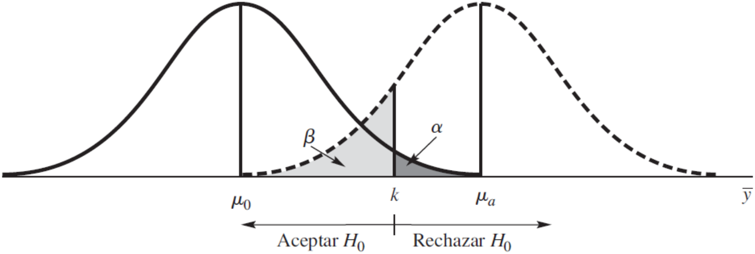
Es imposible mejorar la confianza y la potencia de un test de hipótesis simultáneamente.

Es imposible mejorar la confianza y la potencia de un test de hipótesis simultáneamente.



Comente 16: confianza y potencia

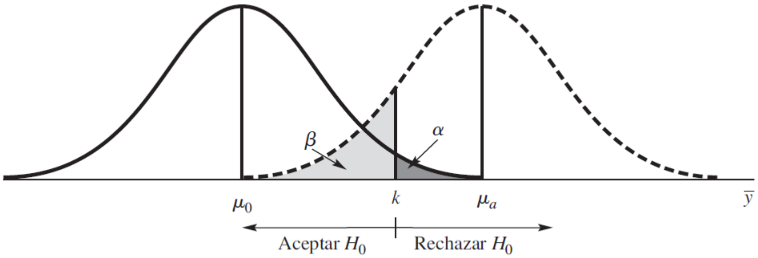
Es imposible mejorar la confianza y la potencia de un test de hipótesis simultáneamente.



$$\uparrow \underbrace{(1-\alpha)}_{\text{confianza}} \Leftrightarrow \downarrow \underbrace{\alpha}_{\mathbb{P}(\text{error tipo 1})} \Leftrightarrow \uparrow \underbrace{\beta}_{\mathbb{P}(\text{error tipo 1})} \Leftrightarrow \underbrace{(1-\beta)}_{\text{potencia}}$$

Comente 16: confianza y potencia

Es imposible mejorar la confianza y la potencia de un test de hipótesis simultáneamente.



$$\uparrow \underbrace{(1-\alpha)}_{\text{confianza}} \Leftrightarrow \downarrow \underbrace{\alpha}_{\mathbb{P}(\text{error tipo 1})} \Leftrightarrow \uparrow \underbrace{\beta}_{\mathbb{P}(\text{error tipo 1})} \Leftrightarrow \underbrace{(1-\beta)}_{\text{potencia}}$$

Verdadero. Si mejoramos la potencia, aumenta la probabilidad de rechazo por lo que la confianza disminuye.