EC4301 MACROECONOMETRÍA
Asist: Juan Andrés Montero Zúñiga

Laboratorio #1:

Máxima Verosimilitud

Creado: Agosto 2024

Cobb-Douglas

Se tiene una función de producción (Q_i) dependiente de trabajo (L_i) y capital (K_i) de la forma Cobb-Douglas asumiendo un error multiplicativo log-normal y homocedástico.

$$Q_i = \beta_0 L_i^{\beta_1} K_i^{\beta_2} \epsilon_i$$

Repaso distribución log-normal

Se tiene una media (m) y una varianza (v):

$$m = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$v = \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right)$$

Puede calcular los parámetros de distribución lognormal μ y σ a partir de la media m y la varianza v:

$$\mu = \log\left(\frac{m^2}{\sqrt{v + m^2}}\right)$$

$$\sigma = \sqrt{\log\left(\frac{v}{m^2} + 1\right)}$$

Función de densidad de probabilidad

$$y = f(x \mid \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\}, \quad \text{for } x > 0.$$

Fuente: MATLAB

Máxima Verosimilitud en STATA

En nuestro caso $\epsilon_i \sim \mathrm{LN}(0, \sigma^2)$

En STATA crearemos el programa "**cobbdouglas_mv**", definimos su versión, y los "**args**" a usar, en este caso:

- "**lnf**" indica que es una función de logverosimilitud $\ln(\beta_0) + \beta_1 \ln(L_i) + \beta_2 \ln(K_i)$
- "**mu**" puede tener otro nombre, en este momento señala que la combinación especificada es la media.
- "**sigma**" es la desviación estándar del error.

Finalmente, se plantea la relación que van a tener los argumentos. Ahora se calcula el logaritmo de las variables y se plantea el modelo con **ml**.

Estimar el modelo

En **ml model** se especifica el modelo, **lf** especifica que el contexto es de forma lineal, luego se definen las variables y un paréntesis vacío indicando error homocedástico. Ahora, con **ml max** se estima el modelo.

En los resultados se obtienen los coeficientes (Betas) en la sección **eq1**. En **eq2** se obtiene la varianza (σ^2) del error ϵ_i. Note que los coeficientes de trabajo y capital no suman 1 y podemos plantear la hipótesis nula de que la función de producción presenta rendimientos constantes a escala con la prueba de *razón de verosimilitud*.

Para esto se impone una restricción con el comando **constraint** señalando que la suma de los coeficientes es igual a 1. En ambos modelos se guardan los resultados bajo los nombres **m1** y **m2**.

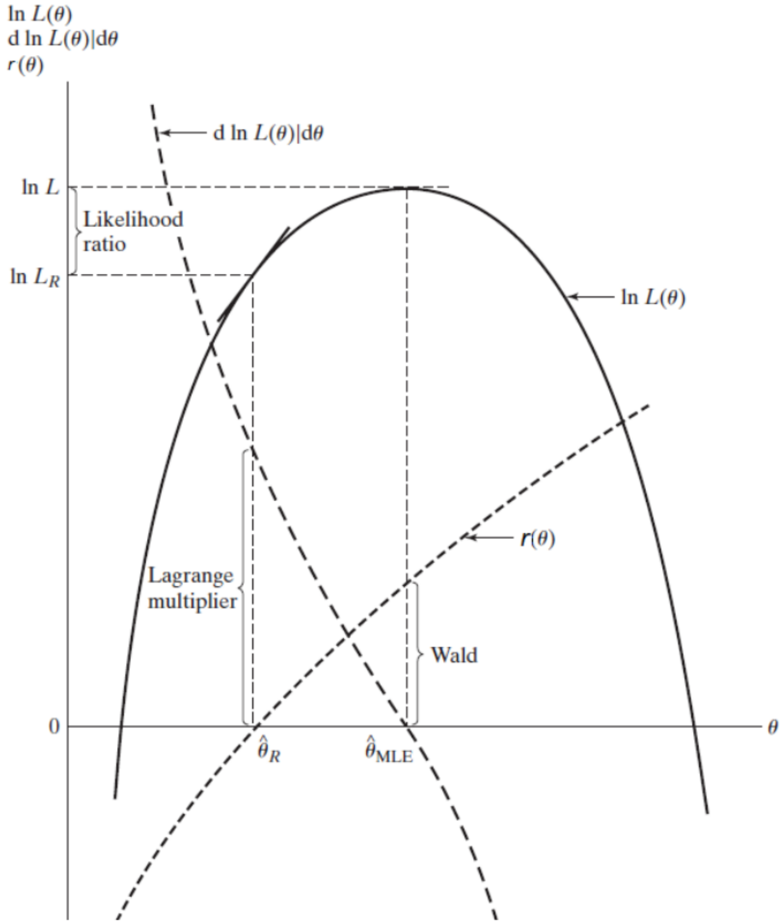
Testeando los resultados

Razón de verosimilitud

Simplemente debemos correr **lrtest m1 m2**. Con un p-value muy bajo rechazamos la hipótesis nula de rendimientos constantes a escala a todos los niveles de significancia.

Wald

Se llega al mismo resultado con un simple **test** que compare los resultados del modelo sin restricción.



Especificación del Modelo con Error Aditivo

En el modelo anterior el término error multiplica a la regresión, esto puede no ser del todo apropiado porque asume que la variabilidad o el "ruido" en los datos es proporcional al valor de la regresión. Ahora se plantea lo siguiente:

$$Q_i = \beta_0 L_i^{\beta_1} K_i^{\beta_2} + \epsilon_i \quad \text{con} \quad \epsilon_i \sim N(0, \sigma^2)$$

Despejando el error se tiene:

$$\epsilon_i = Q_i - \beta_0 L_i^{\beta_1} K_i^{\beta_2} = Q_i - \beta_0 \exp\{\beta_1 \ln(L_i) + \beta_2 \ln(K_i)\}$$

El proceso es muy similar, se crea un programa nuevo, indica versión y sus **args**, lo diferente es que ahora se agrega el factor tecnología **b0** y los demas factores ponderados **b1x**, note que los residuos (**res**) son iguales a la diferencia de la regresión y estos residuos se distribuyen de manera normal.

Repaso distribución normal

Los estimadores de máxima verosimilitud de μ y σ2 para la distribución normal, respectivamente, son:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$s^2_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad \text{(insesgado)}$$

Función de densidad de probabilidad

$$y = f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{para } x \in \mathbb{R}.$$

Log-Verosimilitud

La log-verosimilitud, $\ln L$, es el logaritmo natural de la función de verosimilitud. Para n observaciones independientes, la log-verosimilitud de los datos bajo este modelo normal es la suma de los logaritmos de las densidades de probabilidad:

$$\ln L(\mu, \sigma) = \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)\right)$$

Desglosando esta expresión:

$$\sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n}\left(-\frac{(x_i-\mu_i)^2}{2\sigma^2}\right)$$

$$\Rightarrow \sum_{i=1}^{n}(\ln(1) - \ln(\sqrt{2\pi\sigma^2})) = -\frac{n}{2}\ln(2\pi) - n\ln(\sigma)$$

$$\therefore \ln L(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Caso un único error

$$\epsilon \sim N(0, \sigma^2)$$

$$f(e) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)$$

$$\ln L = \ln\left(\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)\right)$$

$$\ln L = \ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{\epsilon^2}{2\sigma^2}$$

Por lo tanto, la función de log-verosimilitud completa se convierte en:

$$\ln L = -\frac{1}{2}\ln(2\pi) - \ln(\sigma) - \frac{\epsilon^2}{2\sigma^2}$$

Esta función permitirá comparar diferentes funciones **log-likelihood** en STATA, por lo tanto las incluimos en el **program** de la forma:

```
-0.5*ln(2*_pi)-ln(`sigma')-0.5*`res'^2/`sigma^2'
```

Finalmente ejecutamos el modelo "maximum likelihood" **ml model lf cobbdouglas2 (b1: Q=ln_L ln_K, nocons) (b0:) (sigma:)**

Finalmente realizamos los mismos tests que en modelo anterior.

Ejemplo de estimación de máxima verosimilitud a mano

FDP exponencial:

$$y = f(x \mid \theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

Paso 1: Multiplicatoria

La función de verosimilitud $L(\theta)$ para una muestra de tamaño n es el producto de las funciones de densidad individuales:

$$L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}}$$

Paso 2: Log-Verosimilitud

Tomamos el logaritmo natural de la función de verosimilitud para obtener la función de log-verosimilitud (\ell(\theta)):

$$\ell(\theta) = \ln L(\theta) = \ln\left(\prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}}\right)$$

$$\ell(\theta) = -n \ln \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i$$

Paso 3: Derivada de la Log-Verosimilitud

Para encontrar el MLE, derivamos $\ell(\theta)$ con respecto a θ e igualamos a cero:

$$\frac{\partial \ell(\theta)}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0 \implies -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i = 0$$

Paso 4: Resolución para el Estimador de Máxima Verosimilitud

Resolviendo la ecuación anterior para θ:

$$-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i=0$$

Multiplicamos ambos lados por θ^2:

$$-n\theta+\sum_{i=1}^{n}x_i=0$$

Por lo tanto, el estimador de máxima verosimilitud para θ es:

$$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$$