

Profesor: Héctor Bahamonde, PhD.

e: hector.bahamonde@uoh.cl

w: www.hectorbahamonde.com

Curso: MLE.

TA: Gonzalo Barria.

Scope

Este temario incluye las principales ideas que debes mencionar en tu ensayo. La manera en cómo las articulas es subjetiva. El número de ideas corresponde al mínimo absoluto, en consecuencia, deberás mencionarlas todas para obtener una buena nota.

Preguntas y aspectos mínimos a considerar en cada respuesta

1. Si la matemática de OLS y MLE (en su familia **gaussian**) dan los mismos resultados, por qué necesitamos el método MLE? Qué es *gaussian*?

Esto es clase #9.

- (a) Necesitamos el método MLE porque éste está capacitado para trabajar con distribuciones que generan residuos ϵ (o “distribuciones”) no normales.
- (b) *Gaussian* es “normal”. Es por esto que un modelo “normal” (*gaussian*) en MLE da lo mismos resultados que un modelo OLS. Sin embargo, cuando estimamos un modelo MLE con una distribución no normal (no **gaussian**), los resultados son diferentes en ambos métodos.
- (c) De hecho, si usamos un modelo OLS para una distribución no normal, violamos el concepto de normalidad de los residuos, donde $E(\epsilon_i) = 0$.

2. Qué cantidad hace explícita la relatividad de los parámetros estimados $\hat{\theta}$ en MLE? Explica qué limitaciones y posibilidades ofrece la relatividad de $\hat{\theta}$. Explica cómo el enfoque de probabilidad es absoluto.

Esto es clase #7.

- (a) Es la cantidad $k(y)$ en $L(\tilde{\theta}|y) = k(y)Pr(y|\tilde{\theta})$.
- (b) El estudiante debe explícitamente explicar que existe una relación de proporcionalidad (“ \propto ”) entre MLE y OLS. Es decir, $L(\tilde{\theta}|y) \propto Pr(y|\tilde{\theta})$.
- (c) Limitaciones: $k(y)$ hace que cada $\tilde{\theta}$ sea relativo a los datos. Perdemos comparabilidad.
- (d) El enfoque de probabilidad es “absoluto”: el p-value y R^2 de un modelo se puede comparar con otro (usando otros datos). En MLE no se puede.
- (e) Oportunidades: al mismo tiempo que perdemos comparabilidad, ganamos en flexibilidad: podemos estimar una gran cantidad de modelos flexibles (con distintos tipos de variables dependientes).

3. King (1998, p. 16) explica el hecho de que “*M is treated as given is a problem because uncertainty in inference lies with the model, not the data.*” Qué quiere decir?

Esto es clase # 7.

- (a) En OLS se dice que $Pr(y|M) = Pr(\text{datos}|\text{modelo})$. Explicar: $Pr(y|M)$ significa que nosotros obtenemos información de y (coeficientes, incertidumbre y probabilidades) condicional en un modelo M . Eso significa que el modelo está “dado” (“*given*”). El problema es que **el modelo no puede ser considerado como “dado” (porque es construido)**, y **los datos y no pueden ser considerados como *random* porque la naturaleza ya ha realizado ese evento**. Esto debe estar explícito.
 - (b) Entonces, es por esto que en probabilidad es $Pr(y|\tilde{\theta})$ y en “inference” (MLE) es al revés $L(\tilde{\theta}|y)$, es decir, en MLE la inferencia es del modelo respecto a los datos (no de los datos respecto al modelo). El estudiante debe ser explícito y claro respecto a esto.
4. Por qué la tangente en la punta de la curva en la Figura 2.3 de Long (1997, p. 26) es un concepto importante? Qué significa?

Esto es la clase #9 y #10.

- (a) Porque cuando la tangente es plana, el grado de cambio en el *log-likelihood* es cero, es decir, no podemos subir más, y hemos alcanzado el *global maxima*.
- (b) Ese punto significa que estamos en el *global maxima* y no el *local maxima* (clase #9) del *parameter space* (clase #10). Es decir, hemos encontrado el parámetro estimado θ que maximiza el *likelihood* de la distribución en cuestión.
- (c) Explicar el *LLR* y *Wald tests* (clase #10).

REFERENCES

- King, Gary (1998). *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Ann Arbor, MI: University of Michigan Press, pp. 1–274.
- Long, J. Scott (1997). *Regression Models for Categorical and Limited Dependent Variables*. 1st. SAGE Publications, Inc.