

Demanda de Transporte

Estudio de Caso 6: Especificación y Pruebas Estadísticas de Modelos Multivariados de Valor Extremo

Entrega: Martes 29 de Noviembre de 2022

I. Estudio de Caso:

- **Trabajar en parejas, indicar nombres de ambas personas en el informe y en la resolución de los problemas adicionales.**
- Revisar y replicar los ejemplos descritos en los **anexos 1 y 2**. Los resultados presentados en los anexos fueron estimados en Biogeme, que para los modelos Logit Anidados estima la escala (μ). Sin embargo, Apollo estima su inverso ($\lambda=1/\mu$). Discutir al respecto.
- Formular hipótesis respecto a la relación de las variables explicativas con las probabilidades de uso. Validar con test estadísticos.
- Estimar los modelos propuestos e interpretar los resultados y verificar si acaso es plausible. Utilizar diferentes tipos de pruebas de hipótesis (anidadas o no anidadas). **Como mínimo hacer al menos un test IIA de McFadden.**
- Dar una interpretación a los resultados obtenidos recordando que la especificación de modelos es un proceso iterativo.

El reporte final debe contener:

- a) Presentación de diferentes especificaciones y la motivación y razonamiento tras ellas.
- b) Reporte de la mejor especificación lograda.
- c) Por favor incluya los resultados finales de Apollo de los principales modelos estimados

El informe auto-contenido debe reportar cada tarea y resumir las principales conclusiones. Longitud de referencia es 10 páginas. Estructura: Resumen (300 palabras), Introducción, Tareas y Conclusiones. Referencias (formato APA) no contarán en las 10 páginas.

Pauta de evaluación del informe

- Cumplir con lo solicitado (30%).
- Trabajo creativo (30%). El alumno es capaz de utilizar/adaptar/completar la información disponible para cumplir con el objetivo propuesto.
- Análisis (30%). El alumno demuestra comprender la teoría y analiza los resultados de manera asertiva y creativa. Demuestra comprensión del problema (por ejemplo, las relaciones de causalidad); demuestra comprensión de análisis de regresión lineal; Utiliza software de regresión (pruebas de que aprendió a estimar un modelo de regresión con un software apropiado); Demuestra comprensión de estadística de modelos de regresión y pruebas de hipótesis (explicar lo que significan los estadísticos). Recuerde que siempre debe examinar y comentar sobre sus resultados. Salidas de software sin explicaciones no son aceptables (notas escritas a mano están bien).

- Estructura y redacción (10%). Sugerencias: a) ordenamiento lógico de ideas; b) una idea un párrafo; c) primera frase del párrafo presenta la idea y última frase la concluye; d) no más de tres líneas por frase.

La revisión del Informe se realizará con metodología “doble ciego”. **El 20% de la evaluación personal corresponderá al desempeño en la revisión crítica del trabajo de sus pares.**

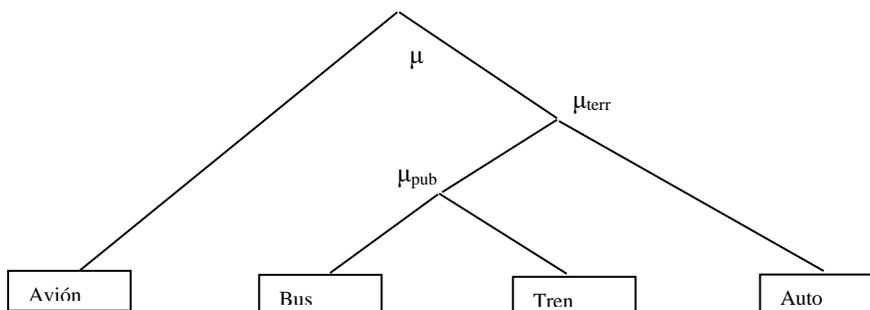
II. Problemas Adicionales

Pregunta 1: Para cada afirmación indique si acaso ella es verdadera o falsa y dé una breve explicación. (Si no hay explicación, o la explicación es errónea en algún sentido, no se otorgarán puntos)

- 1) Aplicando un test IIA de variables omitidas de McFadden, ocurre que la hipótesis nula de que los coeficientes de las variables auxiliares son cero es rechazada sólo con un 90% de confianza. Se debiera entonces recomendar usar un modelo Logit para representar el fenómeno en estudio.
- 2) La propiedad IIA se puede evitar en modelos Logit Binario usando muestreo de alternativas.
- 3) El estimador secuencial de Logit Jerárquico es menos eficiente que estimador simultáneo.
- 4) El propósito de hacer muestreo estratificado en regresión lineal es incrementar la varianza de las variables independientes del modelo. Con esto se busca aumentar la varianza de los estimadores.
- 5) Para estimar un modelo lineal con una muestra estratificada debiera usarse el método de mínimos cuadrados ponderados (weighed least squares), usando el inverso de la probabilidad de muestreo de cada observación como ponderador. De esta manera se logrará maximizar la eficiencia de los estimadores.
- 6) Para medir el impacto de un cambio porcentual de una variable independiente en un modelo lineal, bastará con evaluar el impacto de dicho ese cambio sobre el promedio de las variables independientes, multiplicadas por el respectivo estimador mínimo cuadrado (OLS). El equivalente no es cierto para el modelo Logit.

Pregunta 2

La siguiente figura muestra un Logit Jerárquico (NL) de tres niveles. En el modelo NL hay primero un nivel superior en el cual se elige entre avión y modos terrestres. Luego hay una elección de nivel medio entre transporte público y privado (auto), condicional en la elección de modos terrestres, y finalmente en el nivel inferior está la elección entre tren y bus. Los resultados de la estimación de este modelo NL, así como también de un modelo Logit simple (MNL) se presentan en la tabla siguiente.



	Modo	Logit (MNL)		Logit Jerárquico (NL)	
		parámetro	t-test	parámetro	t-test
Costo en el vehículo	Todos	-0.0060	(-0.85)	-0.0137	(-0.98)
Tiempo de terminal	Todos	-0.0214	(-1.97)	-0.0114	(-0.66)
Ingreso Familiar	Avión	0.0023	-1.97	0.0041	-2.97
Tamaño del Grupo de Viaje	Avión	-0.4952	(-2.10)	-0.8811	(-2.97)
Tiempo de Viaje	Todos	-0.1920	(-2.51)	-0.1169	-0.81
ASC_Avión	Avión	-0.1382	(-0.09)	0.0566	-0.03
ASC_Tren	Tren	2.2693	-2.88	1.9073	-0.75
ASC_Bus	Bus	0.1454	-0.25	0.9724	-0.71
$\mu_{Pub.}$				1.956	-6.39
$\mu_{Terr.}$				1.268	-1.86
Log-likelihood		-125.55		-117.41	
Log-likelihood (0)		-163.58		-163.58	
ρ^2		0.23		0.26	

- Escriba la probabilidad de elección de Tren para el modelo Logit Jerárquico (NL).
- Interprete las variables “Ingreso Familiar” y “Tamaño del Grupo de Viaje” para el modelo Logit (MNL)
- Examine los dos modelos presentados anteriormente. Suponga que usted tiene que recomendar el uso de uno de estos modelos a su cliente. ¿Cuál elegiría y por qué? Suponga que no puede realizar más pruebas.
- Los resultados de las estimaciones presentadas anteriormente se calcularon suponiendo que las observaciones se obtuvieron mediante discado digital al azar de números de teléfono. Sin embargo, indagando más detalles, se encuentra con que 2/3 de las observaciones de avión se obtuvieron de una base de datos de pasajeros aéreos. ¿Necesita re-estimar el modelo Logit (MNL) o el modelo NL? Si es así, ¿cómo y por qué?