



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ESTUDIO DEL FENÓMENO DE NON-ATTENDANCE EN EL PROCESO DE TOMA DE  
DECISIONES DE LAS PERSONAS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

MACARENA SOFÍA GANDULFO KLEIN

PROFESOR GUÍA:  
RICARDO MONTOYA MOREIRA

MIEMBROS DE LA COMISIÓN:  
ANDRÉS MUSALEM SAID  
ALEJANDRA PUENTE CHANDIA

SANTIAGO DE CHILE  
2018

RESUMEN DE MEMORIA PARA OPTAR AL  
TÍTULO DE INGENIERA CIVIL INDUSTRIAL

POR: Macarena Sofía Gandulfo Klein

FECHA: 05 de Octubre de 2018

PROFESOR GUÍA: Ricardo Montoya Moreira

ESTUDIO DEL FENÓMENO DE NON-ATTENDANCE EN EL PROCESO DE TOMA DE  
DECISIONES DE LAS PERSONAS

Los experimentos de elección discreta, tales como análisis conjunto, son ampliamente utilizados para estimar las preferencias de los consumidores sobre ciertos productos o servicios y a través de los años han sido capaces de entregar valiosa información a los investigadores. Sin embargo, durante los últimos años ha crecido de forma considerable el interés sobre el impacto que genera el fenómeno de *non-attendance* o la ignorancia de atributos en la toma de decisiones de las personas. Se ha encontrado evidencia que no considerar este fenómeno genera sesgos en las estimaciones y/o bajo desempeño de los modelos. Investigadores han buscado incorporar este fenómeno, ya sea a través de preguntas directas o mediante inferencia a partir de sus respuestas.

En base a revisión bibliográfica, hasta el momento no se ha visto que los estudios realicen validaciones sobre los modelos que presentan como solución al problema y, por ende, no es posible afirmar que estos enfoques logren capturar el fenómeno de *non-attendance*. Debido a esto, la investigación se enfoca en estudiar el fenómeno de *non-attendance* mediante inferencia en los datos, proponiendo un Modelo Logit con Clase Latente y validando su desempeño contra el Modelo Logit utilizado tradicionalmente. Mediante escenarios simulados se mide la capacidad de predicción que posee el modelo sobre las probabilidades de atender los atributos. Se calculan métricas de ajuste y analizan las diferencias entre ambos modelos. Por último, se aplican ambos modelos en un set de datos reales.

Los resultados muestran que el Modelo Logit con Clase Latente es capaz de recuperar los parámetros de las simulaciones con alta precisión, así también la probabilidad de *non-attendance*, obteniéndose para el 92% de los casos un error bajo el 1%. Destaca un BIC y AIC bastante inferior que el Modelo Logit para casos con alta probabilidad de ignorar atributos. Resultados para datos *in-sample* y *out-of-sample* son consistentes. Además, se aplica el Modelo Logit con Clase Latente sobre datos reales obteniéndose un desempeño, según *hit rate*, y ajuste, según BIC y AIC, superior que el Modelo Logit.

Se concluye que siempre que se posean datos de una muestra representativa, el Modelo Logit con Clase Latente entrega ventajas sobre el Modelo Logit en las estimaciones de las preferencias de los consumidores, presentando un mejor ajuste y desempeño, tanto para casos con presencia de *non-attendance* como para casos donde los participantes atienden todos los atributos.

# Tabla de Contenido

Tabla de Contenido.....	ii
Índice de Tablas.....	v
Índice de Ilustraciones.....	x
1. Introducción.....	1
2. Especificaciones del trabajo.....	3
2.1. Objetivos.....	3
2.1.1. Objetivo General.....	3
2.1.2. Objetivos Específicos.....	3
2.2. Metodología.....	3
2.2.1. Generación de datos.....	3
2.2.2. Evaluación del modelo.....	4
2.2.3. Aplicación del modelo.....	5
2.3. Resultados esperados.....	6
2.4. Alcances.....	6
3. Marco conceptual.....	7
3.1. Análisis Conjunto.....	7
3.2. Modelos de Elección Discreta.....	8
3.3. Reglas de Decisión.....	9
3.4. Non-Attendance.....	9
3.5. Modelo Logit con Clase Latente.....	11
3.5.1. Especificaciones del Modelo.....	11
3.5.2. Asignación de Segmentos.....	13
4. Desarrollo del trabajo.....	15
4.1. Generación de Datos.....	15
4.2. Evaluación del Modelo.....	17

4.2.1.	Hit Rate .....	17
4.2.2.	Criterio de Información Bayesiano (BIC) .....	20
4.2.3.	Criterio de Información de Akaike (AIC) .....	21
4.2.4.	Probabilidad de Observar Atributo .....	23
4.2.5.	Matriz de Confusión .....	24
4.2.6.	Hit Rate por Atributo .....	28
4.2.7.	Probabilidad Promedio de Observar Atributo .....	32
4.2.3.	Conclusiones Evaluación del Modelo .....	36
4.3.	Aplicación del Modelo Logit con Clase Latente .....	38
4.3.1.	Datos .....	38
4.3.2.	Aplicación.....	40
5.	Conclusiones .....	45
6.	Bibliografía.....	46
7.	Anexos.....	48
7.1.	Coeficientes utilizados para la simulación.....	48
7.2.	Listado de Clases .....	49
7.3.	Resultados de las estimaciones para los distintos escenarios y resultados de las evaluaciones para datos de calibración y evaluación.....	50
7.3.1.	Escenario 1 .....	50
7.3.2.	Escenario 2 .....	54
7.3.3.	Escenario 3 .....	58
7.3.4.	Escenario 4 .....	62
7.3.5.	Escenario 5 .....	66
7.3.6.	Escenario 6 .....	70
7.3.7.	Escenario 7 .....	74
7.4.	Matriz de confusión .....	78
7.4.1.	Escenario 1 .....	78

7.4.2. Escenario 2 .....	80
7.4.3. Escenario 3 .....	82
7.4.4. Escenario 4 .....	84
7.4.5. Escenario 5 .....	86
7.4.6. Escenario 6 .....	88
7.4.7. Escenario 7 .....	90

## Índice de Tablas

Tabla 1: Escenarios Simulados en el Experimento .....	16
Tabla 2: Error sobre Probabilidad de Observar los Atributos para los Escenarios Simulados Calculado por el Modelo Logit con Clase Latente .....	23
Tabla 3: Matriz de Confusión Resumida del Escenario 1 Datos Calibración .....	24
Tabla 4: Probabilidad de Observar Atributo Promedio para las Metodologías de Asignación Regla de Bayes y Máximo Alpha, además de Atención Atributo Individual versus Probabilidad Muestral Real Datos Calibración, Escenario 3 .....	33
Tabla 5: Error sobre Probabilidad Muestral Real para Asignación mediante Regla de Bayes y Máximo Alpha, además de Atención Atributo Individual Datos Calibración, Escenario 3.....	34
Tabla 6: Error sobre Probabilidad Promedio para Regla de Bayes, Máximo Alpha y Atención Atributo Individual Datos Calibración .....	34
Tabla 7: Atributos versus Nivel de Satisfacción .....	38
Tabla 8: Hit Rate Aplicado en Datos Reales .....	40
Tabla 9: Métricas de Ajuste Aplicadas en Datos Reales .....	41
Tabla 10: Coeficientes y Errores Estándar Estimados Modelo Logit, Modelo Mixed Logit y Modelo Logit con Clase Latente, Datos Reales. * = Coeficiente Significativo al 95% .....	42
Tabla 11: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Datos Reales .....	42
Tabla 12: Probabilidad Estimada de Atención Atributos, Datos Reales .....	43
Tabla 13: Coeficientes Base para Simulaciones.....	48
Tabla 14: Listado de Clases basado en Vector Alpha .....	49
Tabla 15: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 1 .....	50
Tabla 16: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 1 .....	50
Tabla 17: Probabilidad Estimada de Atención Atributos, Escenario 1 .....	51
Tabla 18: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 1 .....	52
Tabla 19: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 1 .....	52

Tabla 20: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 1 .....	52
Tabla 21: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 1 .....	52
Tabla 22: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 1 .....	53
Tabla 23: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 1 .....	53
Tabla 24: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 1 .....	53
Tabla 25: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 2 .....	54
Tabla 26: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 2 .....	54
Tabla 27: Probabilidad Estimada de Atención Atributos, Escenario 2 .....	55
Tabla 28: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 2 .....	56
Tabla 29: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 2 .....	56
Tabla 30: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 2 .....	56
Tabla 31: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 2 .....	56
Tabla 32: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 2 .....	57
Tabla 33: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 2 .....	57
Tabla 34: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 2 .....	57
Tabla 35: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 3 .....	58
Tabla 36: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 3 .....	58
Tabla 37: Probabilidad Estimada de Atención Atributos, Escenario 3 .....	59

Tabla 38: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 3.....	60
Tabla 39: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 3 .....	60
Tabla 40: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 3 .....	60
Tabla 41: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 3 .....	60
Tabla 42: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 3 .....	61
Tabla 43: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 3 .....	61
Tabla 44: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 3.....	61
Tabla 45: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 4 .....	62
Tabla 46: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 4 .....	62
Tabla 47: Probabilidad Estimada de Atención Atributos, Escenario 4 .....	63
Tabla 48: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 4.....	64
Tabla 49: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 4 .....	64
Tabla 50: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 4 .....	64
Tabla 51: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 4 .....	64
Tabla 52: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 4 .....	65
Tabla 53: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 4 .....	65
Tabla 54: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 4.....	65
Tabla 55: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 5 .....	66



Tabla 56: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 5 .....	66
Tabla 57: Probabilidad Estimada de Atención Atributos, Escenario 5 .....	67
Tabla 58: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 5 .....	68
Tabla 59: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 5 .....	68
Tabla 60: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 5 .....	68
Tabla 61: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 5 .....	68
Tabla 62: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 5 .....	69
Tabla 63: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 5 .....	69
Tabla 64: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 5 .....	69
Tabla 65: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 6 .....	70
Tabla 66: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 6 .....	70
Tabla 67: Probabilidad Estimada de Atención Atributos, Escenario 6 .....	71
Tabla 68: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 6 .....	72
Tabla 69: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 6 .....	72
Tabla 70: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 6 .....	72
Tabla 71: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 6 .....	72
Tabla 72: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 6 .....	73
Tabla 73: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 6 .....	73

Tabla 74: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 6.....	73
Tabla 75: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 7 .....	74
Tabla 76: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 7 .....	74
Tabla 77: Probabilidad Estimada de Atención Atributos, Escenario 7 .....	75
Tabla 78: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 7.....	76
Tabla 79: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 7 .....	76
Tabla 80: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 7 .....	76
Tabla 81: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 7 .....	76
Tabla 82: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 7 .....	77
Tabla 83: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 7 .....	77
Tabla 84: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 7.....	77

## Índice de Ilustraciones

Gráfico 1: Hit Rate Datos Calibración .....	18
Gráfico 2: Hit Rate Datos Evaluación .....	19
Gráfico 3: Criterio de Información Bayesiano (BIC) Datos Calibración.....	20
Gráfico 4: Criterio de Información de Akaike (AIC) Datos Calibración.....	22
Gráfico 5: Hit Rate Asignación Segmentos mediante Regla de Bayes y Máximo Alpha Datos Calibración .....	25
Gráfico 6: Hit Rate Asignación Segmentos mediante Regla de Bayes y Máximo Alpha Datos Evaluación .....	26
Gráfico 7: Hit Rate por Atributo para la Metodología de Asignación Regla de Bayes Datos Calibración .....	28
Gráfico 8: Hit Rate por Atributo para la Metodología de Asignación Máximo Alpha Datos Calibración .....	28
Gráfico 9: Hit Rate por Atributo del Escenario 7 para Metodología de Asignación Regla de Bayes y Máximo Alpha Datos Calibración.....	29
Gráfico 10: Hit Rate por Atributo Promedio para los Escenarios Simulados Datos Calibración .....	30
Gráfico 11: Hit Rate por Atributo para la Metodología de Asignación Regla de Bayes Datos Evaluación .....	31
Gráfico 12: Hit Rate por Atributo para la Metodología de Asignación Máximo Alpha Datos Evaluación .....	31
Gráfico 13: Hit Rate por Atributo Promedio para los seis Escenarios Base Simulados Datos Evaluación .....	32
Gráfico 14: Frecuencia Non-Attendance sobre Preguntas Experimento Elección Discreta .....	39
Ilustración 1: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 1 .....	78
Ilustración 2: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 1 .....	78
Ilustración 3: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 1 .....	79

Ilustración 4: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 1	79
Ilustración 5: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 2	80
Ilustración 6: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 2	80
Ilustración 7: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 2	81
Ilustración 8: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 2	81
Ilustración 9: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 3	82
Ilustración 10: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 3	82
Ilustración 11: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 3	83
Ilustración 12: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 3	83
Ilustración 13: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 4	84
Ilustración 14: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 4	84
Ilustración 15: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 4	85
Ilustración 16: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 4	85
Ilustración 17: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 5	86
Ilustración 18: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 5	86
Ilustración 19: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 5	87
Ilustración 20: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 5	87

Ilustración 21: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 6	88
Ilustración 22: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 6	88
Ilustración 23: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 6	89
Ilustración 24: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 6	89
Ilustración 25: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 7	90
Ilustración 26: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 7	90
Ilustración 27: Matriz de Confusión Asignación Máximo Alpha, Datos Calibración, Escenario 7	91
Ilustración 28: Matriz de Confusión Asignación Máximo Alpha, Datos Evaluación, Escenario 7	91

# 1. Introducción

Los consumidores son el pilar fundamental del éxito de cualquier negocio, por lo que el objetivo principal debiese ser siempre identificarlos, conocerlos y determinar sus necesidades, con el fin último de satisfacerlas. Es por esto, que el marketing tiene como objetivo crear valor para el consumidor, maximizando su utilidad al entregar el producto o servicio óptimo a cada uno.

En esta línea, el objetivo de la investigación es medir las preferencias que poseen los consumidores en situaciones no consideradas bajo los modelos de estimación tradicionales, con el fin de obtener *insights* diferenciadores que satisfagan sus necesidades de forma más eficiente. Para esto, se conceptualiza un producto o servicio como un conjunto de atributos y se utiliza la técnica de análisis conjunto como método de estimación de preferencias.

Típicamente cuando se realizan experimentos de elección discreta se asume que los individuos atienden todos los atributos que definen un producto y que realizan *trade-offs* entre todos estos en su proceso de toma de decisiones. Sin embargo, diversos estudios concluyen que no todos los individuos poseen este comportamiento compensatorio y que sus preferencias son de carácter discontinuo, lo que implica que no atienden o ignoran atributos al tomar sus decisiones. A este fenómeno, de no considerar o ignorar ciertos atributos en el proceso de toma de decisiones, se le conoce comúnmente como *non-attendance* (Campbell et al., 2008).

En la literatura actual es posible observar que, durante años los investigadores han puesto sus esfuerzos en determinar el impacto que el fenómeno de *non-attendance* genera en las estimaciones de las preferencias de los consumidores. Hensher (2006) destaca que existe una posibilidad real que este fenómeno pueda tener un impacto sobre las estimaciones de la disposición a pagar en el caso de ahorro en los tiempos de viajes. Más adelante, Hensher, Scarpa, Campbell y otros investigadores (2009), siguen desarrollando análisis sobre este tema, reforzando la idea que no incorporar esta información genera sesgos en las estimaciones y/o bajo desempeño del modelo. Esto implica que, a través de los años se ha ido reforzando la idea que el fenómeno de *non-attendance* genera impactos significativos sobre las estimaciones de las preferencias de las personas.

Los métodos que se utilizan para estudiar *non-attendance* provienen de dos corrientes distintas: *Stated ANA* e *Inferred ANA*, donde la sigla ANA corresponde a *Attribute Non-Attendance*. El primero, *Stated ANA o follow-up questions* (Hensher et al., 2005), utiliza preguntas directas a los encuestados sobre si consideran o no los atributos, lo que trae el beneficio de no calcular e inferir su comportamiento. Sin embargo, también trae problemas puesto que no es posible determinar cuánto se puede confiar en que las personas entiendan realmente lo que significa no considerar un atributo y, por lo tanto, en las respuestas entregadas por los participantes. Pueden existir los casos en que el participante observó el atributo y no lo consideró significativo, pero si fue considerado en su proceso de toma de decisiones. La literatura muestra también que estos métodos no generan resultados tan confiables y sus resultados difieren bastante de los obtenidos mediante el segundo método.

Por otro lado, el segundo método, *Inferred ANA o analytical* (Scarpa et al., 2009), procura inferir los atributos que son considerados por los encuestados en su proceso de toma de decisiones, mediante modelos probabilísticos en el contexto de análisis conjunto.

En este campo es en el cual se desarrolla el estudio del fenómeno de *non-attendance*, ya que hasta el momento no existe un modelo que pruebe realmente que es capaz de predecir *non-attendance* en el proceso de toma de decisiones de los consumidores. La gran mayoría de las investigaciones se concentran en la utilización de datos reales y proponen modelos que podrían capturar este fenómeno, estimando el impacto que esto podría generar en los resultados finales. Ninguna de estas investigaciones ha comprobado que sus modelos efectivamente capturan el impacto que genera este fenómeno sobre las preferencias.

Esta investigación busca no sólo proponer un modelo capaz de capturar este fenómeno, sino también comprobar que el modelo es capaz efectivamente de capturar la atención de los consumidores a los atributos y estudiar las implicancias que esto genera en las estimaciones de las preferencias. Una forma adecuada de verificar que un modelo cumple con sus objetivos es a través de la simulación de datos y posterior recuperación de parámetros, lo que pretende el trabajo desarrollado.

## **2. Especificaciones del trabajo**

### **2.1. Objetivos**

#### *2.1.1. Objetivo General*

Estudiar el fenómeno de *non-attendance* sobre las preferencias que poseen las personas respecto a los atributos que describen un producto o servicio.

#### *2.1.2. Objetivos Específicos*

- Determinar una metodología adecuada para incorporar el fenómeno de *non-attendance* efectivamente en las estimaciones de preferencias de los consumidores.
- Determinar cómo el fenómeno de *non-attendance* afecta los resultados obtenidos mediante el modelo tradicional aplicado en análisis conjunto.
- Analizar y determinar la existencia de *non-attendance* en datos reales.

### **2.2. Metodología**

La metodología establecida para el desarrollo de la investigación se divide en tres partes: generación de datos, evaluación del modelo y aplicación del modelo.

#### *2.2.1. Generación de datos*

Con el fin de replicar el proceso de toma de decisiones que siguen los consumidores al enfrentarse a un producto o servicio, se genera un set de datos el cual abarca posibles comportamientos que pudiesen adoptar las personas. Estos comportamientos comprenden desde la atención completa a los atributos hasta la ignorancia de estos.

Es posible incorporar la atención a los atributos mediante vectores, sobre los cuales se simulan las decisiones.



Para esto, es necesario seguir los siguientes pasos:

- Se comienza con un set de datos con estructura de experimento de elección discreta.
- Se establece la probabilidad con la que cada individuo va a atender cada atributo, en base al escenario a simular correspondiente.
- Se genera el vector de atención de cada individuo a los atributos, en base a las probabilidades establecidas con anterioridad. Es necesario destacar que, para todas las preguntas del experimento la simulación establece que el individuo mantiene el mismo comportamiento, es decir, el mismo grado de atención a los atributos.
- En base al vector de atención de los atributos generada anteriormente, se simulan las elecciones de los individuos.

### 2.2.2. Evaluación del modelo

La evaluación del modelo corresponde a la parte fundamental de la investigación, con lo cual se es posible determinar la capacidad de predicción del fenómeno de *non-attendance* del modelo. Además, permite desarrollar una comparación fiable con respecto al modelo tradicional.

Para la evaluación del modelo se utilizan las siguientes métricas:

- *Hit rate*: porcentaje total de aciertos en la predicción de las elecciones de los individuos.
- *Criterio de información bayesiano (BIC)*: criterio que penaliza según el número de parámetros a estimar, con el fin de reducir la probabilidad de sobreajuste del modelo. Se aplica mediante la fórmula  $BIC = -2LL(\hat{\beta}) + k \ln(n)$ , donde  $LL(\hat{\beta})$  corresponde a la log-verosimilitud de la muestra,  $k$  corresponde al número de parámetros a estimar y  $n$  el tamaño de la muestra.
- *Criterio de información de Akaike (AIC)*: criterio que penaliza según el número de parámetros a estimar, aunque en menor medida que el criterio BIC. Se aplica mediante la fórmula  $AIC = -2LL(\hat{\beta}) + 2k$ , donde  $LL(\hat{\beta})$  corresponde a la log-verosimilitud de la muestra y  $k$  corresponde al número de parámetros a estimar.
- *Probabilidad de observar el atributo*: se calcula la probabilidad de observar el atributo en la muestra y se compara con la probabilidad con la cual fueron simulados los escenarios, identificando las diferencias.

Mientras que para evaluar la asignación a las clases correspondientes se utilizan las siguientes métricas:

- *Matriz de confusión*: matriz que permite verificar visualmente el desempeño del modelo en la predicción de los segmentos. Las predicciones correspondientes a cada segmento realizadas por el modelo se muestran en las columnas de la matriz, mientras que los individuos clasificados realmente en los segmentos se muestran en las filas. Así, es posible identificar posibles confusiones provocadas por el modelo entre segmentos.
- *Hit rate por atributo*: porcentaje total de aciertos en la predicción por parte del modelo con respecto a si el atributo es atendido para el total de los individuos.
- *Probabilidad Promedio de observar el atributo*: se calcula la probabilidad de observar cada atributo para cada metodología de asignación y se obtiene el promedio de la muestra.

Estas métricas sirven además como base de comparación para medir la efectividad entre el modelo propuesto y el modelo tradicional, en relación con los datos simulados. Se procede luego a tomar un porcentaje de los datos para calibración (*in-sample*) y un porcentaje para análisis de precisión de ambos modelos (*out-of-sample*).

### 2.2.3. Aplicación del modelo

Se aplica el modelo propuesto sobre los datos base reales con los cuales se generó la simulación. Se analizan los resultados mediante el uso de datos *in-sample* y *out-of-sample*, comparando los resultados obtenidos con ambos modelos mediante las siguientes métricas, con el fin de determinar su desempeño y ajuste:

- *Hit rate*: porcentaje total de aciertos en la predicción de las elecciones de la muestra simulada.
- Criterio de información bayesiano (BIC)
- Criterio de información de Akaike (AIC)

### 2.3. Resultados esperados

Se esperan los siguientes resultados del desarrollo de la investigación:

- Descripción del problema y la situación actual, en base a la literatura disponible.
- Simulación de escenarios que incorporen tanto preferencias continuas como discontinuas en las decisiones de los consumidores.
- Desarrollar un modelo que sea capaz de capturar probabilidad de ignorancia de atributos en la toma de decisiones de los consumidores.
- Determinar métricas de evaluación del modelo.
- Determinar métricas de comparación del modelo propuesto versus la teoría tradicional.
- Recuperación de probabilidades de atender atributos.
- Resultados de la aplicación del modelo propuesto y el enfoque tradicional en datos reales.

### 2.4. Alcances

Los alcances de la memoria son los siguientes:

- Dada la naturaleza de investigación del estudio, sólo serán utilizados datos recolectados experimentalmente para el desarrollo de la memoria.
- No serán realizados nuevos experimentos de elección discreta, se trabajará con datos ya disponibles.
- El estudio no variará la dimensión del experimento de elección discreta, por lo que no será estudiado el efecto entre la dimensionalidad del experimento y el fenómeno *non-attendance*.
- Sólo será evaluado un modelo de predicción de *non-attendance*.
- La predicción de *non-attendance* se realizará por atributo completo, sin entrar en los distintos niveles que posee cada uno.
- No serán analizadas las implicancias directas en las organizaciones u aplicaciones la industria, ya que el fin de la memoria es investigación.

### 3. Marco conceptual

#### 3.1. Análisis Conjunto

El análisis conjunto es una técnica estadística que ha sido ampliamente utilizada para modelar la estructura de preferencias de las personas sobre un producto o servicio. Esta técnica posee un enfoque descomposicional basado en el supuesto que los productos son definidos como un conjunto finito de atributos, los cuales son evaluados de forma global por los individuos al tomar sus decisiones. De esta forma, es posible estimar sus preferencias mediante modelos que intentan reconstruir las relaciones entre los distintos atributos, o reglas de composición, que llevaron finalmente al individuo a escoger una alternativa sobre otra (Green y Srinivasan, 1978).

Esta técnica permite ser aplicada a nivel individual, permitiendo analizar las posibles combinaciones entre los atributos que definen el producto que maximizan la utilidad de los consumidores. Cabe destacar que, el análisis conjunto utilizado tradicionalmente asume que los individuos consideran todos los atributos en su proceso de toma de decisiones por lo que cada uno de estos aporta en cierta medida en su utilidad final. Así, los consumidores realizan *trade-offs* entre los distintos niveles de los atributos y son capaces de compensar la disminución de categoría de alguno con el aumento en la categoría de otro. El análisis conjunto permite capturar la forma en que los consumidores realizan estos *trade-offs* y medir la importancia que tienen los atributos en ellos.

Uno de los conceptos fundamentales en el Marketing es que los consumidores son diferentes unos de otros y en la actualidad se utilizan tres modelos básicos de comportamiento que intentan capturar esta heterogeneidad natural presente: los modelos de duración, conteo y elección. Cada uno de estos modelos intenta capturar alguna conducta específica presente en las personas. A su vez, estos modelos se dividen en discretos, donde se agrupan los clientes bajo características homogéneas, y continuos, donde se asume que cada cliente es diferente a los demás.

En la investigación se trabaja bajo el modelo de elección ya que, se pretende estimar la probabilidad que los individuos elijan un determinado comportamiento y determinar sus preferencias.

### 3.2. Modelos de Elección Discreta

Los modelos de elección discreta son apropiados para estimar la probabilidad bajo la cual un individuo va a adquirir un determinado comportamiento en situaciones en que las alternativas de decisión son discretas, se pueden agrupar comportamientos y se posee un conjunto finito de alternativas mutuamente excluyentes y exhaustivas (Train, 2004).

Estos modelos permiten analizar de forma desagregada los factores que llevan a la persona a tomar la decisión de una alternativa sobre otra, bajo el supuesto que sus decisiones son tomadas maximizando su utilidad. La utilidad de cada consumidor sólo es conocida por éste, por lo que es trabajo del investigador intentar determinarla. Así, se descompone la utilidad  $u$  en una componente determinística  $v$  que es observable para el investigador y que incorpora las características del producto o servicio, y otra estocástica  $\varepsilon$  que muestra la distribución de los errores que no son observables, es decir, características innatas que sólo conoce la persona. Finalmente, lo que importa son las diferencias en utilidad y no el valor exacto de los parámetros en los modelos.

$$u_{in} = v_{in} + \varepsilon_{in}$$

Donde se asume que el consumidor  $n$  escoge la alternativa  $i$  si y sólo si se cumple que:

$$u_{in} > u_{jn} \quad \forall j \neq i$$

Usualmente se utilizan Modelos Logit que asumen que los errores distribuyen valor extremo tipo I o distribución Gumbel, ya que resultan más fáciles de estimar. Sin embargo, genera ciertas restricciones en datos con dependencia en el tiempo o problemas con patrones de sustitución entre atributos de mayor complejidad. Agregar heterogeneidad discreta o continua permite al modelo capturar patrones más complejos.

En base a la distribución Gumbel de los errores obtenemos la probabilidad que el individuo  $n$  escoja la alternativa  $i$  dentro del conjunto de alternativas  $I$ :

$$Pr_{in} = \frac{\exp(v_{in})}{\sum_{j=1}^I \exp(v_{jn})} \quad \text{con} \quad \sum_{j=1}^I Pr_{jn} = 1$$

O bien

$$Pr_{in} = \frac{\exp(\sum_{k=1}^K \beta_k x_{ikn})}{\sum_{j=1}^I \exp(\sum_{k=1}^K \beta_k x_{jkn})} \quad \text{con} \quad \sum_{j=1}^I Pr_{jn} = 1$$

Donde  $x_{ikn}$  corresponde al atributo  $k$  observable de la alternativa  $i$  para la persona  $n$  y  $\beta_k$  es su coeficiente asociado.

### 3.3. Reglas de Decisión

Los modelos presentados anteriormente asumen que los consumidores maximizan su utilidad al momento de tomar una decisión, en base a las alternativas que tienen disponibles. Sin embargo, existen casos en los cuales las personas necesitan utilizar heurísticas que les permiten minimizar el esfuerzo necesario para tomar una decisión.

Las reglas de decisión son estrategias de procesamiento de información que utilizan las personas para guiar su proceso de toma de decisiones y se definen como el conjunto de condiciones que permiten aceptar o rechazar una alternativa dentro del conjunto de alternativas. Estas estrategias pueden ser de dos tipos (Riedl et al., 2008):

- *Compensatorias*: existe compensación entre las características de los distintos atributos, es decir, la disminución en la utilidad generada por el nivel poco satisfactorio en un atributo puede ser compensada con el aumento en el nivel de otro, dentro de la misma alternativa. Esto implica que existe sustitución ilimitada entre los distintos atributos y sus niveles.
- *No compensatorias*: no permiten la compensación entre los atributos dentro de la misma alternativa.

Estas reglas de decisión no compensatorias dan paso a preferencias discontinuas, lo que genera discrepancias con el supuesto base de la teoría tradicional utilizada en el método de análisis conjunto, el axioma de continuidad. Sin continuidad, no existen *trade-offs* entre los distintos atributos lo que ocasiona problemas al calcular la tasa marginal de sustitución entre atributos (Lockwood, 1996).

### 3.4. Non-Attendance

El fenómeno *non-attendance* o *attribute non-attendance* se refiere al comportamiento no compensatorio donde las personas toman como regla de decisión la ignorancia de atributos específicos (Hensher et al., 2005). Esto implica que no importa cuánto incremente o mejore el nivel de cierto atributo, si la persona lo ignora al tomar decisiones, no compensará la disminución de nivel en otro.

Este fenómeno ha sido ampliamente estudiado durante los últimos años, por autores como Hensher (2005), Campbell (2008), Scarpa (2009), Erdem y Hole (2015), entre otros, llegando a la conclusión que considerar este comportamiento que adoptan los participantes en las estimaciones de las preferencias genera un impacto significativo en éstas. Usualmente, se han encontrados mejoras significativas en el desempeño de los modelos al simplemente dar cuenta de la existencia de *non-attendance* en los modelos. Estas mejoras implican que se llega a expresiones de utilidad más fieles a la realidad de los consumidores.

A lo largo de estos años, se han utilizado dos corrientes distintas para el estudio de este fenómeno. La primera es “*Stated Attribute Non-Attendance*” o “*Stated Ana*”, la cual busca identificar las preferencias nulas sobre los atributos a través de preguntas directas a los sujetos sometidos al experimento. Estas preguntas suelen realizarse luego de cada respuesta entregada por el participante o al finalizar el experimento. A través de las investigaciones, se ha llegado a la conclusión que este método no resulta ser óptimo para incorporar *non-attendance* a las estimaciones. Esto, debido a que las personas no necesariamente comprenden la diferencia entre ignorar realmente un atributo en su decisión o tener preferencias muy bajas sobre éste. Además, los resultados muestran inconsistencias con respecto al método analítico. Por otro lado, se obtiene como ventaja que no es necesario realizar cálculos de inferencia adicionales y es posible utilizar los modelos tradicionales en las estimaciones.

La segunda corriente corresponde a “*Inferred Attribute Non-Attendance*” o “*Inferred Ana*”, la cual utiliza modelos probabilísticos con el fin de inferir desde los datos puros si los consumidores ignoran o no los atributos. Posee como gran ventaja que no existen sesgos generados por las personas que participan en los experimentos ya que, sus comportamientos son determinados por los datos y no por sus creencias. Como se mencionó anteriormente, existen diferencias entre los resultados obtenidos mediante ambos métodos, superando en ajuste “*Inferred ANA*” sobre “*Stated Ana*”.

Modelos probabilísticos conocidos que se han utilizado hasta el momento incluyen restringir los coeficientes a cero en el contexto de clases latentes (Scarpa et al., 2009), estimaciones bayesianas (Scarpa et al., 2009), logit multinomial (Hensher et al., 2011), entre otros.

## 3.5. Modelo Logit con Clase Latente

### 3.5.1. Especificaciones del Modelo

Con el fin de determinar si es posible capturar el fenómeno de *non-attendance*, se plantea un Modelo Logit con Clase Latente (a partir de ahora MLCL) en base al trabajo de Erdem, S., Campbell, D. y Hole, A.R. (2013) “*Attribute-level non-attendance in a choice experiment investigating preferences for health service innovations*”. Su investigación incorpora la atención a los atributos mediante un Modelo Logit con Clase Latente y concluye que incorporar *non-attendance* en los niveles de los atributos genera una mejora substancial en el ajuste del modelo e impacta en las estimaciones del *willingness to pay*.

El modelo propuesto para la investigación es una variación del modelo expuesto anteriormente ya que, no incorpora *non-attendance* en los niveles de los atributos particularmente como lo realizan ellos, en cambio incorpora el fenómeno sobre el atributo como un todo independiente del grado o valor de sus niveles. Esto, con el fin de asemejarse más al común de los modelos propuestos en la literatura. Se estudia entonces si cada atributo en particular es atendido o no en el proceso de toma de decisiones de las personas.

La investigación realizada por los autores propone que los individuos no sólo ignoran los atributos al momento de tomar una decisión, sino que además pueden ignorar un subconjunto de niveles de los atributos. Esto implica que el atributo sólo puede ser ignorado completamente si todos los niveles de ese atributo son ignorados. Así, se incorpora la atención a los atributos de forma más específica, directamente en los niveles.

Se propone entonces un MLCL al cual se incorpora un vector de variables dummy  $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_K\}$  compuesto por tantos elementos como número de atributos posea el producto o servicio, que hace referencia a la atención de la persona sobre cada atributo. Los elementos  $\alpha_k$  toman valor igual a uno cuando el atributo  $k$  es atendido e igual a cero cuando es ignorado. Por ejemplo, si la persona atiende todos los atributos al tomar sus decisiones, su vector correspondiente sería  $\alpha = \{1, 1, 1, 1, 1, 1\}$ .

Así, el número de posibles combinaciones para el vector  $\alpha$  viene dado por sus elementos, obteniéndose un número finito  $S = 2^K$  de segmentos heterogéneos, lo cual resulta considerablemente menor que el número posible de combinaciones de probabilidades de atención para cada  $\beta_k$  de los atributos. Así, la probabilidad de escoger la alternativa  $i$  dentro del conjunto de alternativas  $I$  por la persona  $n$  está dada por:

$$Pr_{in} = \frac{\exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{ikn})}{\sum_{j=1}^I \exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{jkn})}$$



Dado que el vector  $\alpha$  solo es conocido por la persona, se busca estimar la probabilidad de cada individuo de considerar cada atributo. Se define entonces la probabilidad de obtener  $\alpha_k = 1$ , es decir, la probabilidad de atender el atributo  $k$ , como  $\pi_k^1$  y la probabilidad de ignorar el atributo y obtener  $\alpha_k = 0$ , como  $\pi_k^0$ .

Las probabilidades se encuentran sujetas a tres restricciones, con el fin de imponer que se constituya una ley de probabilidad:

$$0 \leq \pi_k^0 \leq 1 \quad \forall k \in K$$

$$0 \leq \pi_k^1 \leq 1 \quad \forall k \in K$$

$$\pi_k^0 + \pi_k^1 = 1 \quad \forall k \in K$$

Así, las probabilidades se calculan en base a la siguiente fórmula, donde los coeficientes son obtenidos en la optimización.

$$\pi_k^1 = \frac{1}{1 + \exp(-\gamma_k)} \quad \forall k \in K$$

De esta forma es posible estimar la probabilidad de pertenecer a cada segmento, en base a la probabilidad de observar cada posible combinación de  $\alpha_K$  que define cada segmento, definida como  $\alpha_k^s$  con  $s = \{1, 2, 3, \dots, S\} \wedge k \in K$ . Por ejemplo, se definió que la combinación de atención completa a los atributos corresponde a la Clase 64, estableciéndose  $\alpha^{64} = \{1, 1, 1, 1, 1, 1\}$ .

$$\phi_s = \pi_1^{\alpha_1^s} * \pi_2^{\alpha_2^s} * \pi_3^{\alpha_3^s} * \dots * \pi_K^{\alpha_K^s}$$

Lo que genera finalmente que la probabilidad de escoger la alternativa  $i$  dentro del conjunto de alternativas  $I$  por la persona  $n$  dada la segmentación:

$$Pr_{in} = \sum_{s=1}^S \phi_s \frac{\exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{ikn})}{\sum_{j=1}^I \exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{jkn})}$$

Generalizando para incorporar la secuencia de decisiones que toma cada persona  $y_n = \{y_{n1}, y_{n2}, \dots, y_{nT_n}\}$  en las  $T_n$  ocasiones de decisión:

$$Pr(y_n | x_n, \beta, \pi) = \sum_{s=1}^S \phi_s \prod_{t=1}^{T_n} \frac{\exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{y_{ntkn}})}{\sum_{j=1}^I \exp(\sum_{k=1}^K \alpha_k^s \beta_k x_{jkn})}$$

El nivel de referencia utilizado en los niveles de los atributos es el nivel bajo. Esto implica que los niveles de satisfacción altos y medios quedan especificados sobre el bajo, el cual toma valor cero en sus coeficientes.

### 3.5.2. Asignación de Segmentos

La asignación a los segmentos se realiza utilizando dos metodologías distintas. La primera corresponde a la utilizada tradicionalmente en modelos con clases latente, donde se asigna mediante Regla de Bayes y la segunda es mediante la máxima probabilidad de observar el atributo, a lo que llamamos “Máximo Alpha”.

#### 3.5.2.1. Regla de Bayes

La asignación a los diferentes segmentos se realizó en base a la que presenta mayor probabilidad de elección para cada persona, utilizando la regla de Bayes. Así, la probabilidad de pertenecer a cada segmento para cada individuo viene dada por:

$$Pr_{sn} = \frac{LL_{ns} * \phi_s}{\sum_{r=1}^S LL_{nr} * \phi_r} \quad \forall s = \{1,2,3, \dots, S\} \text{ y } \forall n \in N$$

Una vez calculadas las probabilidades de pertenecer a todos los segmentos, se escoge simplemente la mayor, asumiendo que el individuo posee ese comportamiento.

#### 3.5.2.2. Máximo Alpha

Se establece una asignación alternativa basada en la regla de Bayes, diferenciándose en que no se escoge simplemente la mayor probabilidad, en cambio se consideran todos los segmentos y se realiza una ponderación de las probabilidades sobre los valores del vector  $\alpha_k^s$  para cada atributo. Si el valor de la ponderación es menor a 50% entonces se asume que el atributo no es considerado, obteniendo  $\alpha_k = 0$ . En cambio, si el valor es mayor o igual a 50%, se obtiene  $\alpha_k = 1$ .

$$Pr_{kn} = \sum_{s=1}^S \frac{LL_{ns} * \phi_s}{\sum_{r=1}^S LL_{nr} * \phi_r} * \alpha_k^s \quad \forall k \in K \text{ y } \forall n \in N$$

$$\widehat{\alpha}_k = \begin{cases} 0 & \text{si } Pr_{kn} < 0,5 \\ 1 & \text{si } Pr_{kn} \geq 0,5 \end{cases} \quad \forall k \in K$$

Este procedimiento se realiza para todos los atributos, obteniéndose así el vector  $\alpha^s$  y por consiguiente la clase asignada.

### 3.5.2.3. Atención Atributo Individual

El modelo propuesto entrega la probabilidad de observar *non-attendance* en la muestra completa, además de asignar con cierta probabilidad a las personas a los distintos segmentos de atención a los atributos. Se propone como alternativa, y con el fin de entregar la mayor cantidad posible de información, una probabilidad de atender a cada atributo de forma individual y no agregada como en la asignación tradicional de clases latentes mediante la regla de Bayes.

La propuesta corresponde a una variación de la asignación mediante “*máximo alpha*”, deteniéndose un paso antes en la probabilidad ponderada de los segmentos por atributo. Esto implica que, “*máximo alpha*” corresponde a un caso particular de “*atención atributo individual*”, donde las probabilidades  $Pr_{kn}$  son aproximadas. Se construye así el vector  $A_n$  que muestra la probabilidad de atender cada atributo por individuo:

$$Pr_{kn} = \sum_{s=1}^S \frac{LL_{ns} * \phi_s}{\sum_{r=1}^S LL_{nr} * \phi_r} * \alpha_k^s \quad \forall k \in K \text{ y } \forall n \in N$$

$$A_n = (Pr_{1n}, Pr_{2n}, \dots, Pr_{Kn}) \quad \forall n \in N$$

La probabilidad de observar el fenómeno *non-attendance* es complementaria a la probabilidad de atender los atributos, por lo que el vector  $NA_n$  que muestra la probabilidad de no atender cada atributo se define como:

$$NA_n = (1 - Pr_{1n}, 1 - Pr_{2n}, \dots, 1 - Pr_{Kn}) \quad \forall n \in N$$

## 4. Desarrollo del trabajo

### 4.1. Generación de Datos

Los datos utilizados para el desarrollo de la investigación provienen de distintos escenarios simulados, los cuales representan la presencia del fenómeno de *non-attendance* en diferente escala. Para todos los escenarios de *non-attendance* establecidos, la generación de datos sigue el mismo procedimiento, variando únicamente la probabilidad de atención a los atributos. Se establece que esta probabilidad es transversal a todos los atributos, es decir, es igual para todos los atributos en un mismo escenario. Esto, con el fin de identificar cualquier posible falla o diferencia generada por el modelo y determinar su estabilidad. Además, se asume que los individuos mantienen su comportamiento durante todo el experimento, por lo que todas las preguntas que responden mantienen el mismo vector  $\alpha$ .

La atención a los atributos fue simulada mediante la creación del vector  $\alpha$  para cada individuo, siendo los valores de sus elementos generados en base a una Distribución Bernoulli con probabilidad igual a la establecida para cada escenario. Así, los elementos de  $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_K\}$  toman valor de 1 en base a la probabilidad de atención  $\pi_k^1$  y 0 en base a la probabilidad complementaria  $\pi_k^0$  de observar *non-attendance*. Como se ha mencionado anteriormente, este vector permanece constante para las 12 preguntas que componen el experimento, manteniendo así la consistencia en el comportamiento de los individuos en su proceso de toma de decisiones.

Se simularon entonces 800 personas con 12 preguntas cada una, dando un total de 9.600 preguntas. Los valores de los coeficientes utilizados para cada atributo se mantienen fijos para todos los individuos, con el fin de mantener consistencia. Incorporando estos coeficientes, los vectores  $\alpha$  creados anteriormente y las alternativas seleccionadas en los datos originales, con sus respectivos valores, se obtienen las probabilidades de decisiones simuladas por cada persona en el experimento y se generan entonces los distintos escenarios.

En base a la literatura disponible (Scarpa et al., 2008), se establecen seis escenarios para el estudio del fenómeno *non-attendance*, que corresponden a distintos niveles de atención a los atributos. Sin embargo, a lo largo del estudio se observó que el sexto escenario generaba resultados bastante extremos en las distintas métricas de evaluación de asignación, por lo que se tomó la decisión de incluir un séptimo escenario correspondiente a un caso con alta probabilidad de observar *non-attendance*. Así, de ser necesario es posible identificar si los resultados obtenidos corresponden efectivamente al aumento de la probabilidad de ignorar atributos o simplemente al caso extremo del sexto escenario, donde el modelo podría confundir las probabilidades.

<b>Simulación</b>	<b>Probabilidad de Atención</b>	<b>Probabilidad Non-Attendance</b>
<b>Escenario 1</b>	100%	0%
<b>Escenario 2</b>	99%	1%
<b>Escenario 3</b>	95%	5%
<b>Escenario 4</b>	90%	10%
<b>Escenario 5</b>	75%	25%
<b>Escenario 6</b>	50%	50%
<b>Escenario 7</b>	25%	75%

*Tabla 1: Escenarios Simulados en el Experimento*

Todas las simulaciones, cálculos y demás fueron realizados en el software estadístico R del cual es posible obtener mayor información en <https://www.r-project.org/>.

## 4.2. Evaluación del Modelo

La evaluación del modelo propuesto se realiza contra el enfoque tradicional utilizado en análisis conjunto. En este caso, las comparaciones se realizaron contra un Modelo Logit (a partir de ahora ML), el cual define la verosimilitud en base a la probabilidad de escoger la alternativa  $i$  dentro del conjunto de alternativas  $I$  por la persona  $n$ :

$$L(\beta|x) = \prod_i \frac{\exp(\sum_{k=1}^K \beta_k x_{ikn})}{\sum_{j=1}^I \exp(\sum_{k=1}^K \beta_k x_{jkn})}$$

Para ambos modelos, la alternativa predicha por el individuo corresponde a la con mayor probabilidad y se optimizan los parámetros para maximizar la log-verosimilitud.

Los modelos son sometidos a los escenarios establecidos, aplicándose ambos sobre la misma base de datos simulada. Se estudiaron los casos con separación *in-sample* y *out-of-sample*. Las primeras 10 preguntas de cada persona son utilizadas para calibrar el modelo, siendo en total 8.000 preguntas de calibración. Las últimas dos preguntas son utilizadas para la evaluación, lo que da un total de 1.600 preguntas. Así, los datos se dividen en 80% *in-sample* y 20% *out-of-sample*.

Las primeras tres métricas corresponden a la evaluación directa del desempeño de los modelos, mientras que las siguientes tres evalúan la asignación a las clases mediante las distintas metodologías.

### 4.2.1. Hit Rate

*Hit rate* mide el porcentaje total de aciertos en la predicción de las elecciones de los individuos, mediante la siguiente fórmula:

$$\text{Hit Rate} = \frac{\text{Total aciertos}}{\text{Total preguntas}}$$

En este caso, para calcular el *hit rate in-sample* se utilizan 8.000 preguntas como total, las cuales corresponden a 10 preguntas para cada una de las 800 personas. Para el caso *out-of-sample* se utilizan 1.600, que implican 2 preguntas para la misma cantidad de personas.

El Gráfico 1 muestra los resultados obtenidos para esta métrica sobre los datos de calibración. El eje vertical izquierdo corresponde al gráfico de barras el cual muestra el porcentaje de aciertos obtenidos para cada escenario por el modelo tradicional y el modelo propuesto. El eje vertical derecho corresponde al gráfico de línea el cual representa la diferencia entre los resultados de ambos modelos, lo que para esta métrica corresponde a la fórmula  $ML - MLCL = \text{Hit rate}_{ML} - \text{Hit rate}_{MLCL}$ , en sentido inverso.

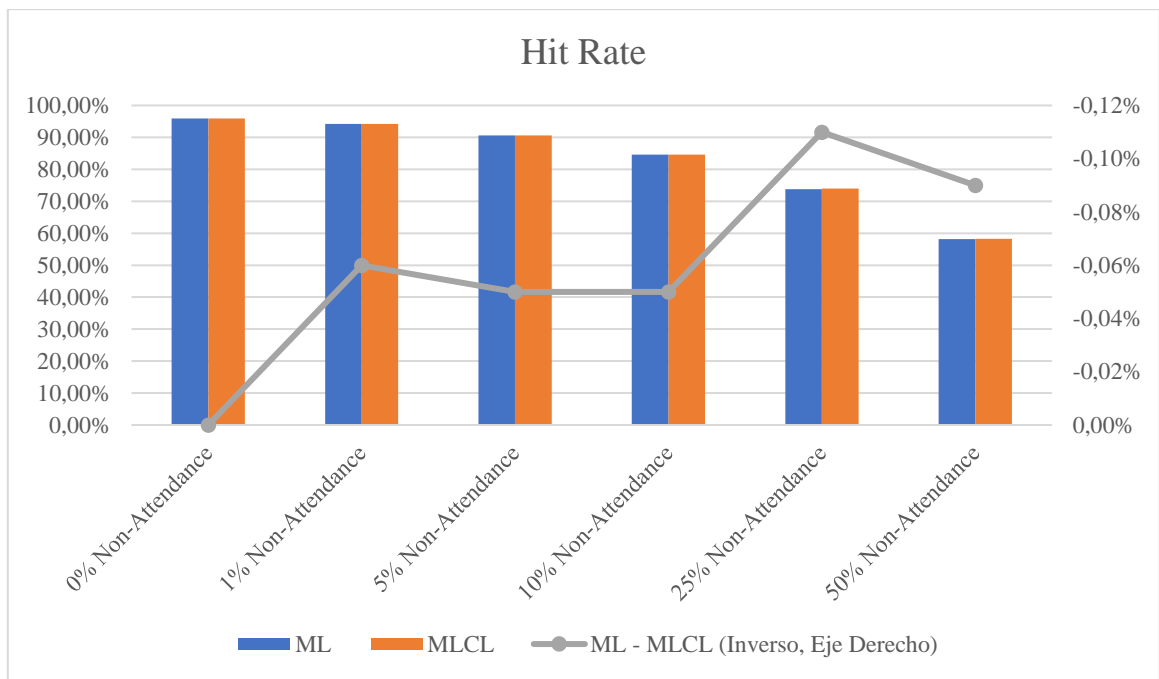


Gráfico 1: Hit Rate Datos Calibración

El primer escenario, con 0% de probabilidad de observar *non-attendance*, corresponde al escenario ideal para el ML, ya que se cumple el supuesto base de preferencias continuas en las decisiones de los consumidores. Se esperaría entonces que el ML fuese superior o igual en capacidad de predicción al MLCL. En el Gráfico 1 es posible observar que ambos modelos poseen igual capacidad de predicción en este escenario, ya que obtienen igual valor de *hit rate*. La diferencia entre ambos modelos se muestra en el gráfico de línea, correspondiente al eje vertical derecho. Para este escenario, la diferencia entre ambos modelos es nula. De esto se puede deducir que el modelo propuesto se comporta de forma similar, si no idéntica, al modelo tradicional en casos de atención completa a los atributos.

Los distintos escenarios muestran que ambos modelos pierden capacidad de predicción al aumentar la probabilidad de observar *non-attendance* en la muestra. Como se mencionó anteriormente, el eje vertical derecho muestra la diferencia entre el *hit rate* del ML y el *hit rate* del MLCL, en sentido inverso. Valores negativos implican que el modelo propuesto posee un mayor *hit rate* que el modelo tradicional, mientras que valores positivos indican lo contrario. Se muestra claramente que a mayor probabilidad de observar *non-attendance* el MLCL posee un mayor poder de predicción que el ML.

A pesar de que la diferencia entre ambos modelos es pequeña, esta se mantiene consistentemente en todos los escenarios simulados. En ningún caso se observa que el modelo tradicional posea una mayor cantidad de aciertos por sobre el modelo propuesto, sino que todo lo contrario. Además, se observa que la capacidad de predicción de ambos modelos se ve condicionada a la atención que los consumidores les dan a los atributos al momento de tomar una decisión, es decir, a mayor probabilidad de no atender los atributos, menor capacidad de predicción poseen ambos modelos. Tendencialmente, el MLCL tiende a superar en aciertos al ML al aumentar la probabilidad de observar *non-attendance* en la muestra.

El Escenario 6, el cual describe la probabilidad de atender los atributos en 50% muestra un escenario bastante significativo para los modelos. Dado que, la probabilidad de observar o no el atributo es la misma, es posible que los modelos tiendan a confundirse entre ambas posibilidades. Se destaca que a pesar de esto, ambos modelos mantienen el *hit-rate* sobre el 50%.

Los datos de evaluación muestran resultados similares a los obtenidos en calibración, donde el modelo propuesto presenta un mayor *hit rate* al aumentar la probabilidad de *non-attendance* que el modelo tradicional. A pesar que las diferencias entre los resultados obtenidos en ambos modelos sea baja, se observa que el MLCL posee consistentemente una mejor precisión en la predicción de las decisiones que toman los individuos que el ML, superándolo en aciertos.

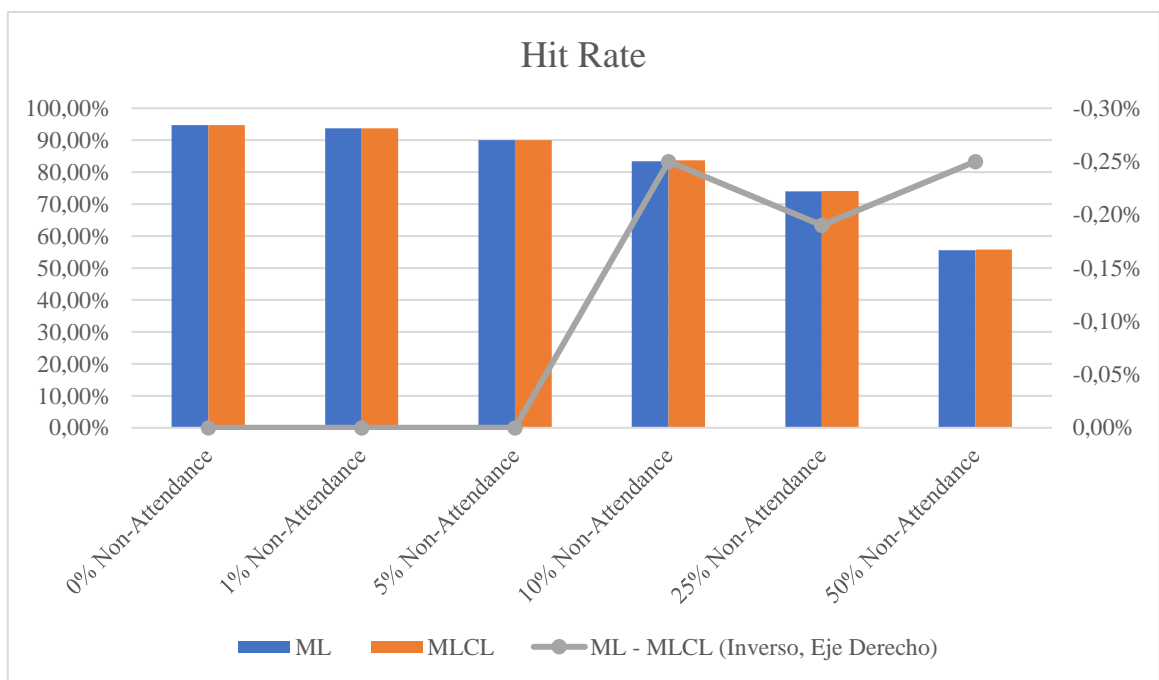


Gráfico 2: Hit Rate Datos Evaluación

El Gráfico 2 muestra que a mayor probabilidad de observar *non-attendance*, menor capacidad de predicción poseen ambos modelos. Sin embargo, el modelo propuesto comienza a superar en porcentaje de aciertos al modelo tradicional al aumentar la probabilidad de ignorar atributos. Estos resultados demuestran consistencia con los obtenidos *in-sample*.

Cabe destacar que las diferencias en *hit-rate* de ambos modelos son bastante pequeñas, no llegando ni siquiera al 1% por lo que probablemente no sean significativas.



#### 4.2.2. Criterio de Información Bayesiano (BIC)

El criterio de información bayesiano, BIC por sus siglas en inglés, penaliza a los modelos según el número de parámetros a estimar, a mayor número de parámetros, mayor penalización. Esto, con el fin de evitar sobreajustes del modelo ya que, incorporar más variables implica una mayor verosimilitud, pero disminuye la capacidad predictiva del modelo ante distintos escenarios. Su fórmula viene dada por:

$$BIC = -2LL(\hat{\beta}) + k \ln(n)$$

Donde  $LL(\hat{\beta})$  corresponde a la log-verosimilitud de la muestra,  $k$  corresponde al número de parámetros a estimar y  $n$  el tamaño de la muestra.

El MLCL agrega tantas variables como atributos se están evaluando en el experimento, por lo que en este caso se agregan seis variables extras. Así, el ML posee doce variables a estimar, mientras que el MLCL posee dieciocho.

Se considera que el modelo propuesto posee un mejor ajuste que el modelo tradicional si, a pesar de que el número de variables a estimar sea mayor, posea un menor criterio de información bayesiano.

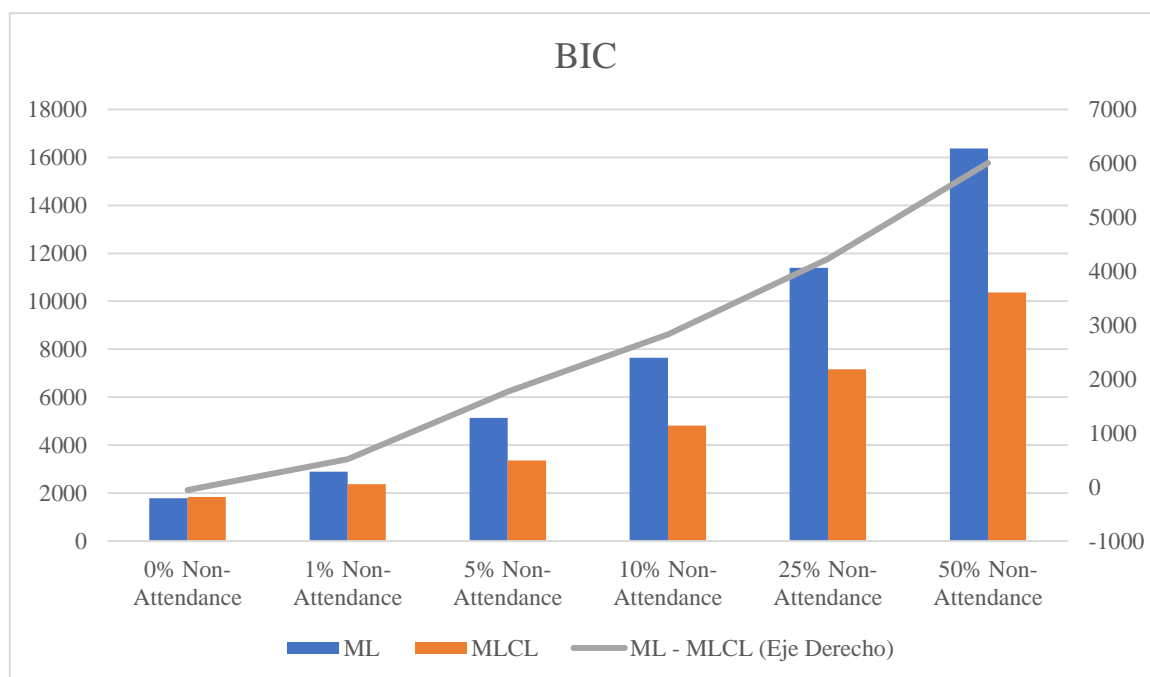


Gráfico 3: Criterio de Información Bayesiano (BIC) Datos Calibración

El primer escenario, con atención completa a los atributos, corresponde al único escenario donde el ML posee un menor BIC que el modelo propuesto. Los resultados muestran que ambos modelos obtienen igual valor de log-verosimilitud por lo que el modelo propuesto es penalizado por poseer un mayor número de variables a estimar. Sin embargo, la diferencia entre ambos modelos es muy pequeña en relación a las observadas en los cinco escenarios restantes por lo que no se considera relevante.

Para todos los casos con probabilidad mayor a cero de observar *non-attendance*, el modelo propuesto presentó un mejor ajuste que el modelo tradicional. La diferencia entre el ML y el MLCL corresponde al eje vertical derecho, donde se observa claramente que a mayor probabilidad de ignorar atributos el MLCL posee un ajuste bastante superior que el ML.

La diferencia entre los ajustes de los modelos incrementa a mayor probabilidad de ignorar los atributos. La tendencia al alza es claramente observable en el Gráfico 3, donde el valor del criterio de información bayesiano correspondiente al modelo tradicional alcanza 1,58 veces la del modelo propuesto. En base a esto, se considera que el MLCL posee un ajuste superior al ML bajo la presencia o no de *non-attendance*.

Se deduce entonces que el ajuste del MLCL es superior que el ML utilizado tradicionalmente, a pesar de la penalización por el aumento en el número de variables. Además, para los casos donde los atributos son considerados en su totalidad por las personas, el modelo propuesto compensa la leve disminución en el desempeño con el aumento en información que entrega ya que permite saber a los investigadores la probabilidad con la que los participantes están atendiendo los atributos, en relación al modelo tradicional.

#### 4.2.3. Criterio de Información de Akaike (AIC)

El criterio de información de Akaike, AIC por sus siglas en inglés, representa una métrica similar de ajuste al criterio de información bayesiano, penalizando en menor medida por el número de variables, como se puede observar en la siguiente fórmula:

$$AIC = -2LL(\hat{\beta}) + 2k$$

Donde  $LL(\hat{\beta})$  corresponde a la log-verosimilitud de la muestra y  $k$  corresponde al número de parámetros a estimar.

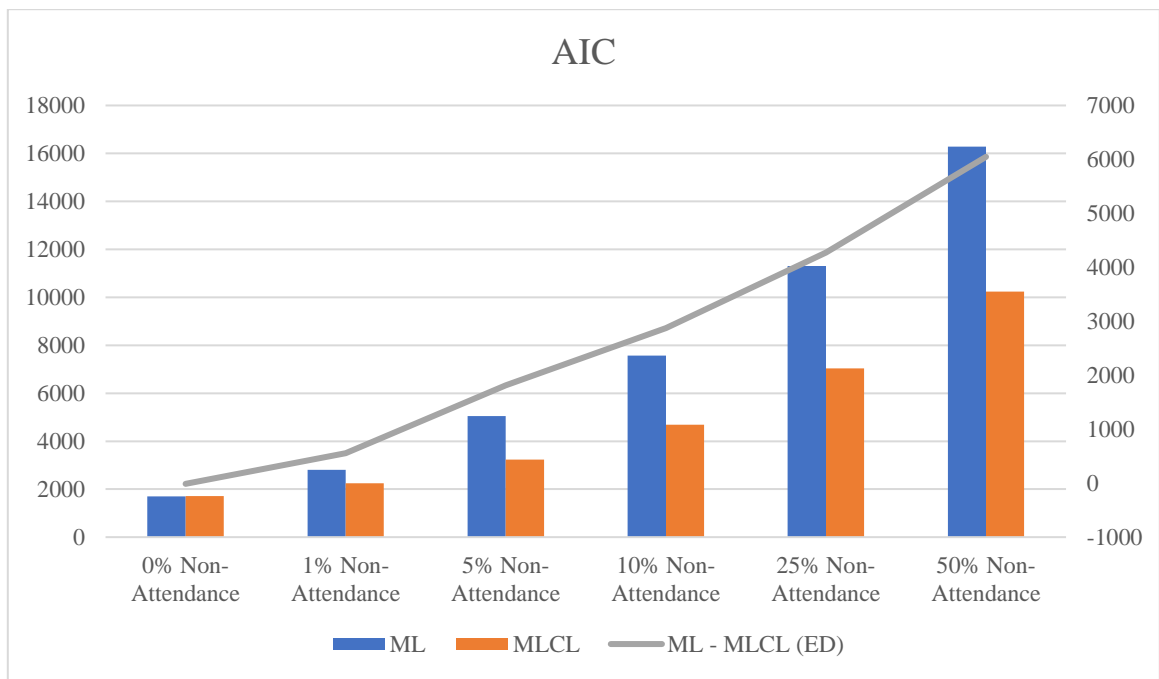


Gráfico 4: Criterio de Información de Akaike (AIC) Datos Calibración

Al igual que en el criterio anterior, lo que se busca es encontrar el modelo que realiza la mejor predicción, encontrando equilibrio entre complejidad y capacidad de ajuste. En el Gráfico 4 es posible observar los resultados obtenidos para los distintos escenarios.

Demostrando consistencia con el criterio de información bayesiano, es posible observar una clara tendencia incremental de las diferencias de ajuste del modelo propuesto sobre el modelo tradicional en relación al aumento de la probabilidad de observar *non-attendance*. En el primer escenario el ML supera ligeramente en ajuste al MLCL. Como se explicó anteriormente, dado que ambos poseen igual valor de log-verosimilitud el modelo propuesto es penalizado por poseer un mayor número de variables a estimar. Sin embargo, al aumentar en 1% la probabilidad de observar *non-attendance* ya es posible identificar una diferencia considerable entre ambos AIC.

Para casos donde existe probabilidad no nula de ignorar algún atributo por parte de los individuos, el modelo propuesto genera mejores predicciones. Se observa además que las diferencias en ajuste entre ambos modelos tienden a aumentar junto con el incremento en la probabilidad de observar *non-attendance*.

En base a ambos criterios de información es posible concluir que de existir probabilidad mayor a cero de observar ignorancia de atributos en la muestra el modelo propuesto tendrá un ajuste superior al modelo tradicional. Para los casos con total atención a los atributos, la diferencia entre ambos modelos es marginal, por lo que podrían ser utilizados indistintamente en base a estos criterios. La ventaja que se obtiene al utilizar el MLCL sobre el ML, es que entrega además información del comportamiento de los consumidores sobre si atienden o no los atributos del producto.

#### 4.2.4. Probabilidad de Observar Atributo

La probabilidad de observar los atributos se obtiene directamente de la optimización, donde se calculan los coeficientes respectivos, como se menciona en las especificaciones del modelo.

$$\pi_k^1 = \frac{1}{1 + \exp(-\gamma_k)} \quad \forall k \in K$$

Para todos los escenarios se calcula la diferencia entre la probabilidad de observar los atributos con la cual fue realizada la simulación y la determinada por el modelo. El error para todos los atributos se calcula mediante la siguiente fórmula:

$$error_k = |\pi_k^1 - Pr_k| \quad \forall k \in K$$

Un error de 0% implica que la probabilidad de observar los atributos con la cual fue realizada la simulación y la probabilidad determinada por el modelo son exactamente iguales. Es decir, el modelo fue capaz de recuperar la probabilidad perfectamente. Mientras menor sea el error, mejor el resultado obtenido por el modelo. Los resultados obtenidos para todos los escenarios se muestran a continuación:

	<b>Probabilidad Non- Attendance</b>	<b>Atributo 1</b>	<b>Atributo 2</b>	<b>Atributo 3</b>	<b>Atributo 4</b>	<b>Atributo 5</b>	<b>Atributo 6</b>
<b>Escenario 1</b>	<b>0%</b>	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
<b>Escenario 2</b>	<b>1%</b>	0,61%	0,59%	0,10%	0,21%	0,02%	0,00%
<b>Escenario 3</b>	<b>5%</b>	0,37%	0,00%	0,12%	0,50%	0,03%	0,45%
<b>Escenario 4</b>	<b>10%</b>	0,06%	0,62%	0,34%	0,99%	0,08%	0,02%
<b>Escenario 5</b>	<b>25%</b>	1,11%	0,05%	0,40%	1,14%	0,21%	0,88%
<b>Escenario 6</b>	<b>50%</b>	1,66%	0,97%	0,51%	0,53%	0,49%	0,56%

Tabla 2: Error sobre Probabilidad de Observar los Atributos para los Escenarios Simulados Calculado por el Modelo Logit con Clase Latente

La Tabla 2 muestra que para el primer escenario con atención completa a los atributos el modelo es capaz de predecir con total precisión el porcentaje de atención, obteniendo  $\pi_k^1 = 100\%$  en todos los atributos, igual a la probabilidad con la cual fue realizada la simulación. Esto implica que para todos los atributos el error porcentual sea nulo.

Además, es posible observar que el 92% de las veces el error es menor al 1%, lo que implica que el modelo entrega un porcentaje de probabilidad de atención por cada atributo bastante acertado. El error máximo alcanzado por el modelo es de 1,66% lo que implica que la probabilidad predicha por el modelo para ese atributo fue cercana a la probabilidad con la cual fue realizada la simulación. Cabe destacar que este valor fue alcanzado además en el escenario con 50% de probabilidad de observar *non-attendance*, un escenario que ha mostrado ser más difícil de estimar que los demás, posiblemente debido a que existe igual probabilidad de observar o no los atributos.

Este cálculo tiene como fin identificar la capacidad de predicción que tiene el modelo sobre la probabilidad de la muestra de observar cada atributo. La Tabla 2 nos muestra que el error oscila entre 0% y 1,66% por lo que el modelo es capaz de identificar de forma adecuada las probabilidades. Esto implica que entrega información confiable a los investigadores sobre el comportamiento que adoptan las personas al tomar decisiones, por lo menos a nivel agregado de la muestra.

#### 4.2.5. Matriz de Confusión

Las tres métricas presentadas anteriormente tenían como propósito evaluar el modelo tradicional y el modelo propuesto, con el fin de identificar cual entrega un mejor ajuste. Por otro lado, las siguientes métricas tienen como objetivo evaluar la asignación a los distintos segmentos mediante las metodologías de asignación Regla de Bayes y Máximo Alpha.

La matriz de confusión es una herramienta útil para identificar rápidamente si el modelo es capaz de predecir correctamente las clases a las cuales pertenece cada individuo y si está confundiendo clases. El listado de clases se muestra en el Anexo 7.2. y el orden de éstas se mantiene a lo largo de todo el experimento. En una matriz de confusión las columnas representan el número de predicciones por el modelo, mientras que las filas corresponden a las clases reales de cada individuo. Así, cualquier elemento que difiera de la diagonal en una matriz de confusión correspondería a un error en la clasificación.

	<b>Clase 1</b>	<b>Clase 2</b>	<b>Clase 3</b>	<b>...</b>	<b>Clase 64</b>
<b>Clase 1</b>	0	0	0	...	0
<b>Clase 2</b>	0	0	0	...	0
<b>Clase 3</b>	0	0	0	...	0
<b>...</b>	...	...	...	...	...
<b>Clase 64</b>	0	0	0	...	800

Tabla 3: Matriz de Confusión Resumida del Escenario 1 Datos Calibración

En la Tabla 3 se muestra a modo de ejemplo la matriz de confusión para la metodología regla de Bayes en el primer escenario con atención completa a los atributos, donde todos los individuos pertenecen a la Clase 64: (1 1 1 1 1). Para este caso, el modelo ha clasificado correctamente el 100% de los individuos, ya que coinciden todas las personas en la clasificación. La metodología Máximo Alpha posee una matriz de confusión idéntica para el mismo escenario, cumpliendo con la misma precisión de asignación. Todas las matrices de confusión pueden ser encontradas en el Anexo 7.4.

Las matrices de confusión muestran que la asignación mediante Máximo Alpha posee una mayor precisión en la clasificación de los individuos en las distintas clases. Además, se observa que cuando los modelos se equivocan en la asignación ambas metodologías tienden a predecir mayor atención a los atributos de la real, aunque Máximo Alpha en mayor medida. En casos de baja probabilidad de observar *non-attendance*, los modelos muestran una mayor precisión en la asignación a las clases.

Las matrices de confusión permiten además calcular el *hit rate* de asignación a las distintas clases. El Gráfico 5 muestra el *hit rate* de asignación de ambas metodologías para los distintos escenarios simulados.

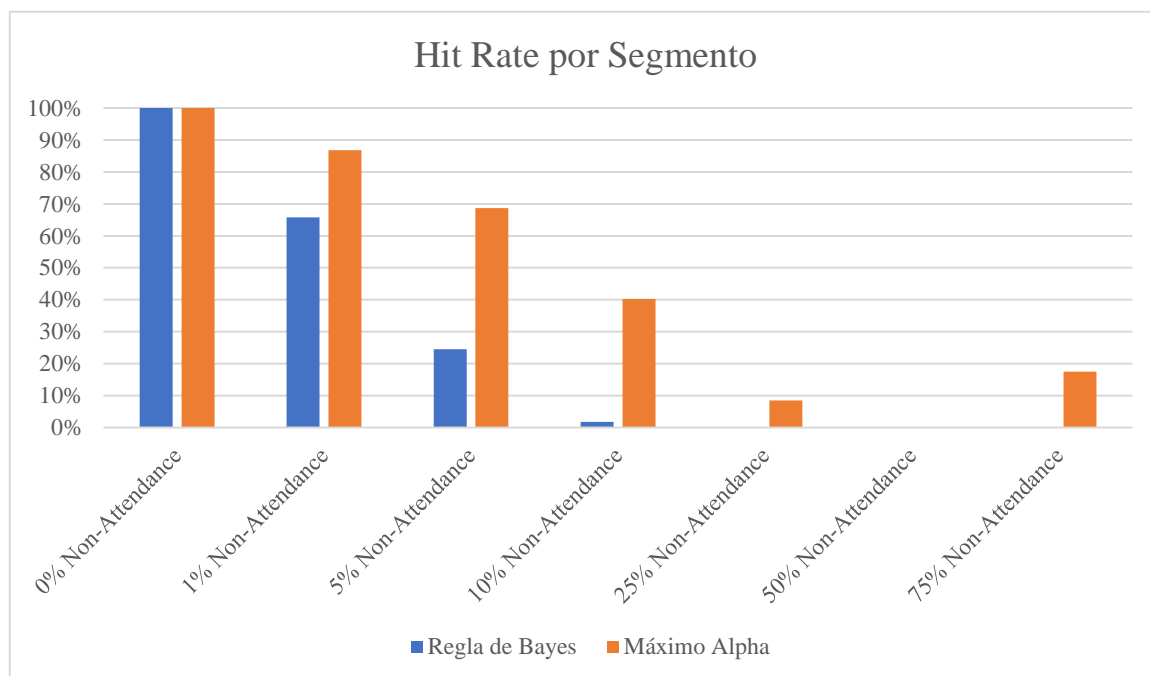


Gráfico 5: Hit Rate Asignación Segmentos mediante Regla de Bayes y Máximo Alpha Datos Calibración

Como se mencionó anteriormente, para los casos con atención completa a los atributos ambas metodologías asignaron a las personas en los distintos segmentos con 100% de precisión. Al aumentar la probabilidad de ignorar atributos, se observa claramente que Máximo Alpha supera considerablemente en *hit rate* a Regla de Bayes.

El sexto escenario, con 50% de probabilidad de atender los atributos, presenta un *hit rate* de 0% para ambas asignaciones. Dado este resultado, se temió que el modelo pudiese no asignar de forma correcta las personas a los distintos segmentos al existir una probabilidad mayor al 51% de atender los atributos. En base a esto, se incluyó el séptimo escenario con 25% de probabilidad de observar los atributos, con el fin de identificar si efectivamente no es posible asignar las personas correctamente a los segmentos con probabilidades altas de *non-attendance*.

Los resultados muestran que Regla de Bayes no es capaz de asignar ningún individuo a su clase correctamente al existir una probabilidad superior al 10% de *non-attendance*. Por otro lado, Máximo Alpha puede asignar en todos los escenarios a excepción del caso extremo de 50% para todos los atributos.

Dado que la probabilidad de observar o no *non-attendance* es la misma, se sospecha que el modelo tiende a confundir las clases y asigna de forma errónea los consumidores a los distintos segmentos. Las matrices de confusión para este escenario muestran una gran dispersión en cuanto a la asignación en comparación con los demás, tanto para los datos *in-sample* como *out-of-sample*. Esta dispersión vuelve a desaparecer al pasar al Escenario 7, donde se vuelve a observar que las asignaciones tienden hacia las clases correctas.

En cuanto a los resultados obtenidos *out-of-sample* se observan resultados consistentes con los *in-sample*. Para situaciones donde no existe probabilidad de observra *non-attendance*, ambas metodologías son capaces de asignar de forma correcta a los consumidores a sus segmentos. Sin embargo, al aumentar en 1% la probabilidad de ignorar atributos se observa una disminución considerable en el *hit rate* por parte de Regla de Bayes y Máximo Alpha, bajando del 50%.

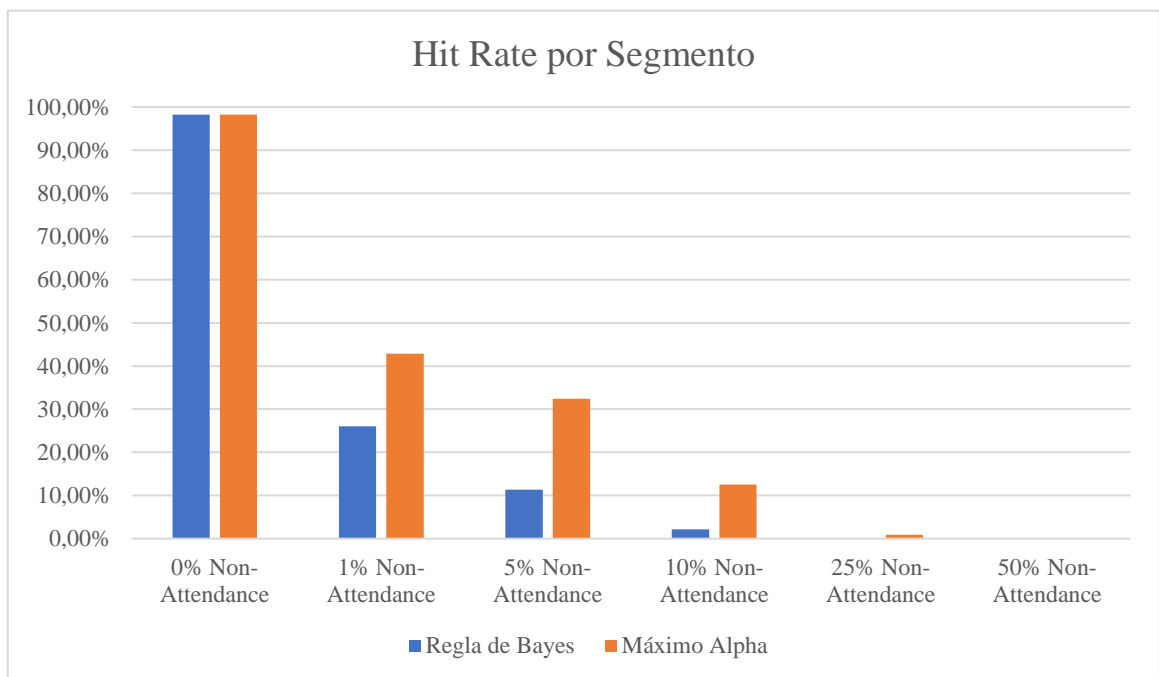


Gráfico 6: Hit Rate Asignación Segmentos mediante Regla de Bayes y Máximo Alpha Datos Evaluación

El Gráfico 6 muestra que la asignación mediante la metodología Máximo Alpha entrega resultados más certeros que Regla de Bayes, al igual que en los datos de calibración. Es por esto que, hasta encontrar una metodología de asignación más certera se recomienda su uso para asignar a los individuos a los distintos segmentos. Cabe mencionar que, a pesar que esta metodología demuestra no ser óptima para casos donde exista probabilidad mayor a cero de *non-attendance*, es la que entrega mejores resultados dentro de las opciones.

En escenarios donde exista igual probabilidad de atender o no los atributos, como en el sexto escenario, no es posible asignar correctamente a ningún individuo en los distintos segmentos. Sin embargo, las estimaciones del modelo si entregan información precisa sobre la probabilidad de *non-attendance* de la muestra por lo que aunque no sea posible asignar a los individuos, es posible saber que la muestra atenderá los atributos con determinada probabilidad.

Las matrices de confusión muestran que las dos metodologías de asignación poseen una alta capacidad de predicción de segmentos en escenarios con probabilidades altas de atención a los atributos. El *hit rate* de asignación decae fuertemente a medida que incrementa la probabilidad de *non-attendance*, llegando a ser nulo en el sexto escenario.

Así, la información entregada por el modelo es un complemento de la información entregada en la asignación, no recomendándose su utilización de forma individual.



#### 4.2.6. Hit Rate por Atributo

La asignación de las clases permite evaluar cuantas veces las distintas metodologías predijeron correctamente la atención a los atributos. El *hit rate* que posee cada metodología sobre cada atributo *in-sample* se muestra en los Gráficos 7 y 8 a continuación.

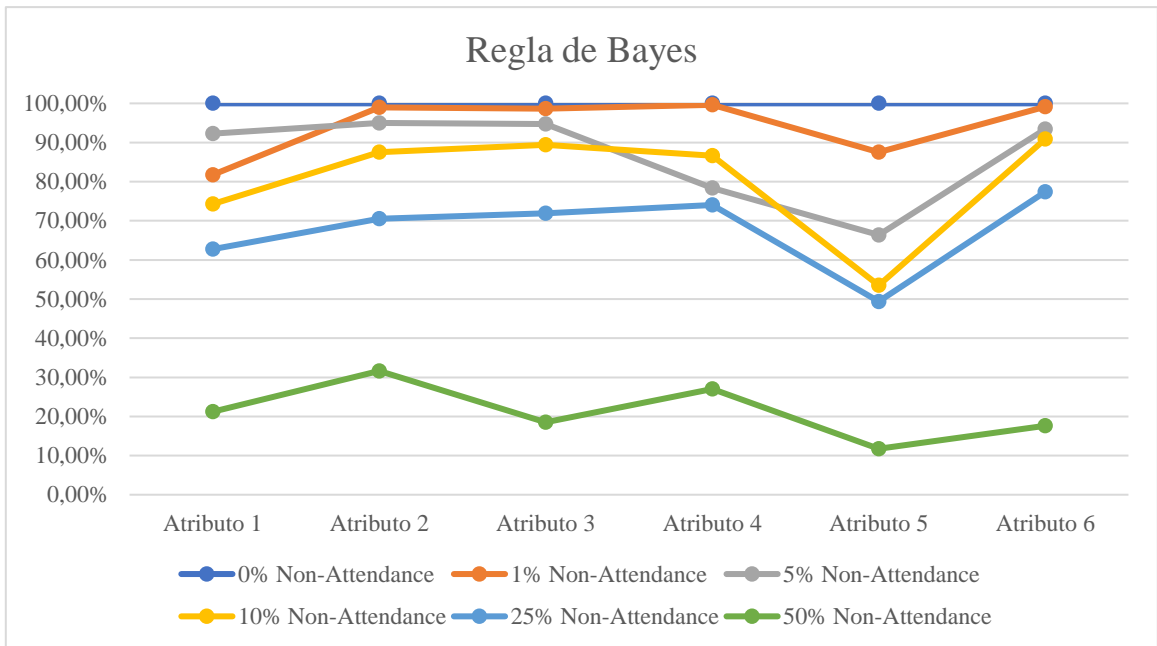


Gráfico 7: Hit Rate por Atributo para la Metodología de Asignación Regla de Bayes Datos Calibración

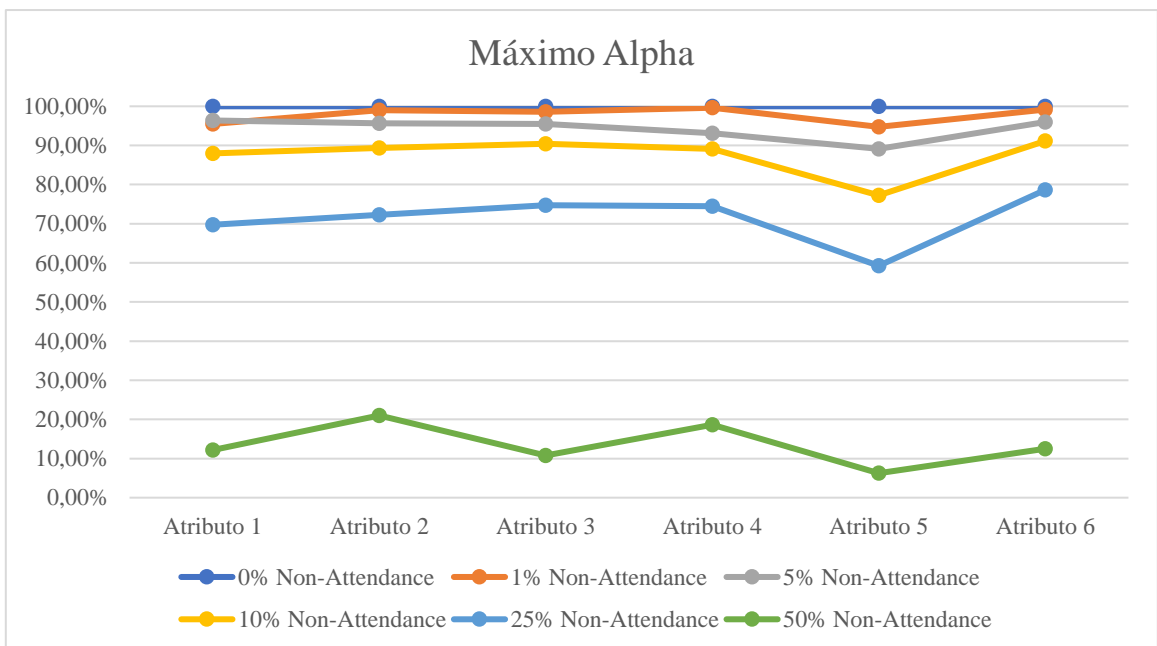


Gráfico 8: Hit Rate por Atributo para la Metodología de Asignación Máximo Alpha Datos Calibración

Como es posible observar, Máximo Alpha presenta mayor estabilidad en las estimaciones a través de todos los atributos que Regla de Bayes. Así mismo, el *hit rate* obtenido mediante esta metodología resulta ser mayor en todos los escenarios a excepción del sexto que posee 50% de probabilidad de observar *non-attendance*.

Estos resultados pueden parecer extraños al ser comparados con los obtenidos con las matrices de confusión, donde se observa que Regla de Bayes no es capaz de asignar a ninguna persona a su segmento correcto. Sin embargo, es necesario recordar que en la sección anterior se está buscando predecir la clase completa de forma correcta y en esta sección se busca predecir si el atributo es atendido o no correctamente de forma individual.

Debido a la disminución considerable en el Escenario 6, se busca determinar si esto es provocado por el aumento en la probabilidad de *non-attendance* o si corresponde a un resultado provocado por las características de éste. Así, se utiliza el séptimo escenario representado en el Gráfico 9.

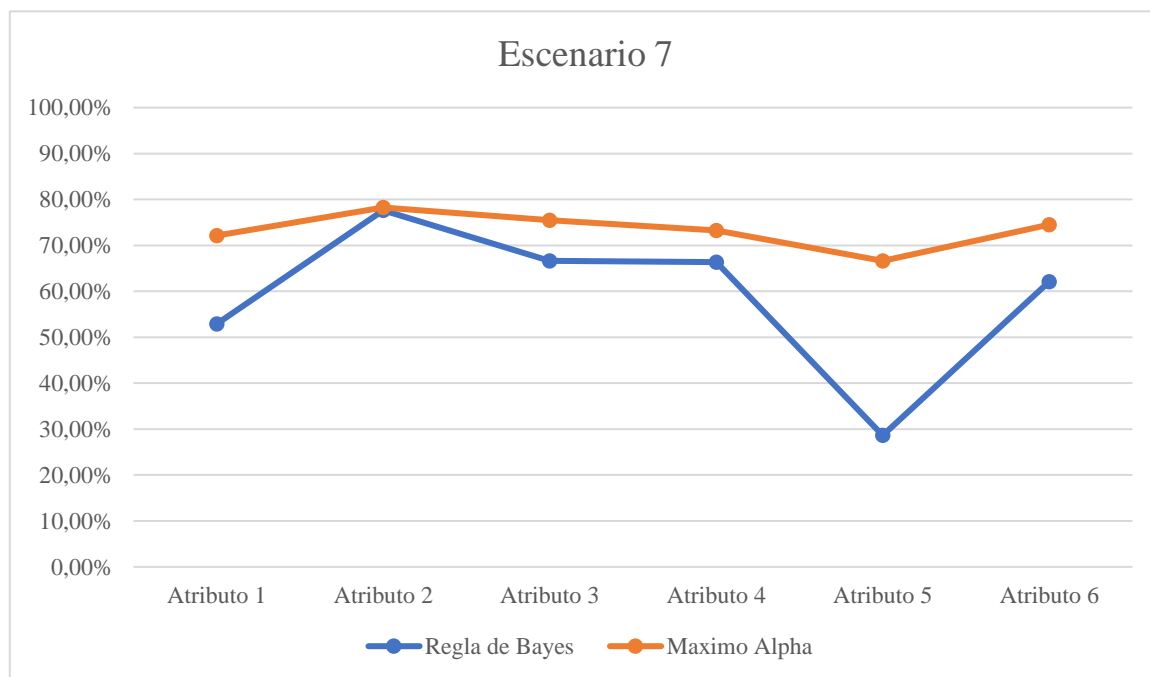


Gráfico 9: Hit Rate por Atributo del Escenario 7 para Metodología de Asignación Regla de Bayes y Máximo Alpha  
Datos Calibración

Nuevamente observamos una mayor consistencia por parte de Máximo Alpha en comparación con la asignación mediante Regla de Bayes. Cabe notar que, el Escenario 7 muestra un aumento en el *hit rate* de todos los atributos, volviendo a la zona del 70%. Esto implica que la baja en precisión demostrada en los gráficos anteriores se da simplemente por las características del sexto escenario y no por el aumento en la probabilidad de ignorar atributos.

El Gráfico 10 muestra el *hit rate* promedio de ambas metodologías para cada escenario. Se observa que a medida que aumenta la probabilidad de observar el fenómeno *non-attendance*, es más difícil predecir si los atributos serán atendidos. Sin embargo, a excepción del Escenario 6, los *hit rates* para la asignación mediante Regla de Bayes se mantienen sobre el 60% y los de Máximo Alpha, sobre 70%, lo que se considera un buen resultado.

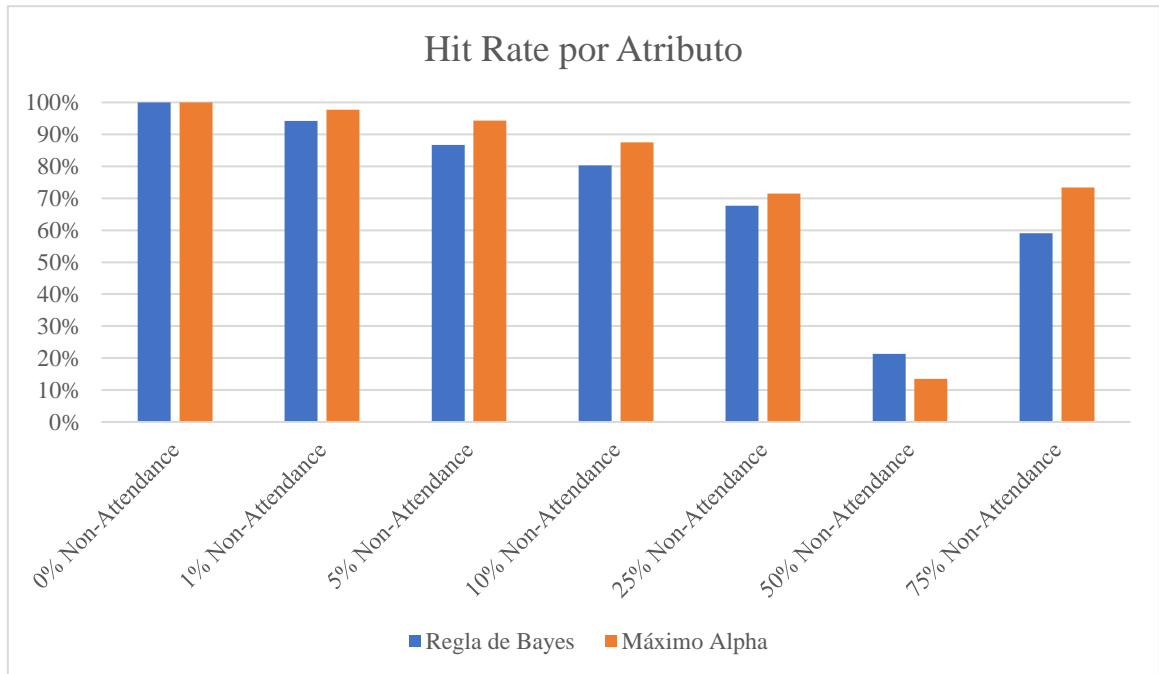


Gráfico 10: Hit Rate por Atributo Promedio para los Escenarios Simulados Datos Calibración

Los resultados obtenidos *out-of-sample* son bastante similares para ambas metodologías de asignación. En los Gráficos 11 y 12 es posible observar que Máximo Alpha presenta mayor estabilidad en las estimaciones a través de todos los atributos en los distintos escenarios que Regla de Bayes. Esto implica que, los resultados obtenidos en un atributo pueden ser comparados con los obtenidos en otro, al mismo tiempo que aumenta la confianza en estos.

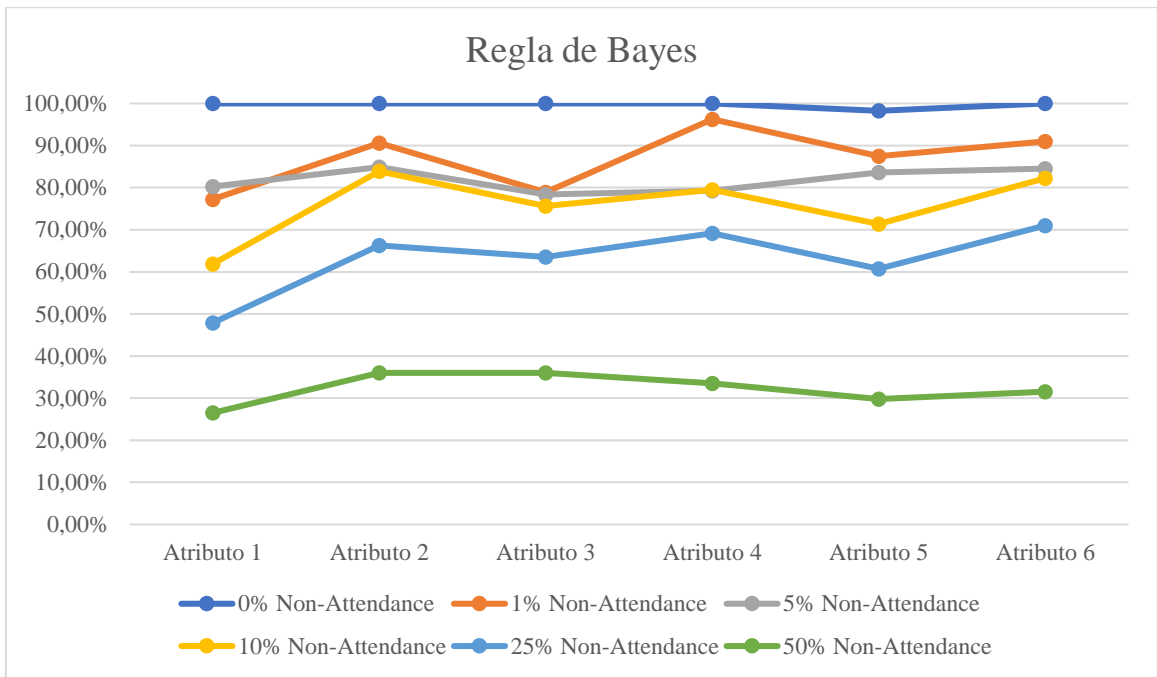


Gráfico 11: Hit Rate por Atributo para la Metodología de Asignación Regla de Bayes Datos Evaluación

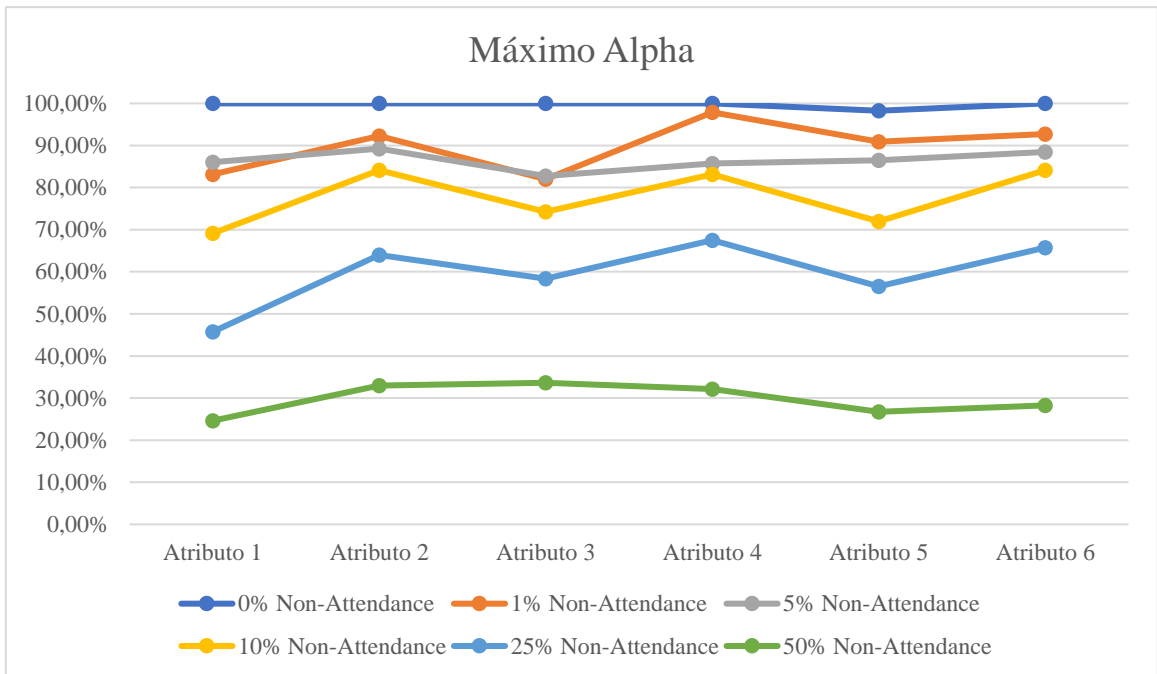


Gráfico 12: Hit Rate por Atributo para la Metodología de Asignación Máximo Alpha Datos Evaluación

Al analizar el *hit rate* por atributo en promedio para cada escenario se observan pequeñas diferencias entre las dos metodologías en comparación con los datos de calibración. En este caso, en tres de seis escenarios Máximo Alpha obtiene un *hit rate* superior a Regla de Bayes, mientras que en un escenario obtienen un resultado idéntico y en dos Regla de Bayes supera a Máximo Alpha.

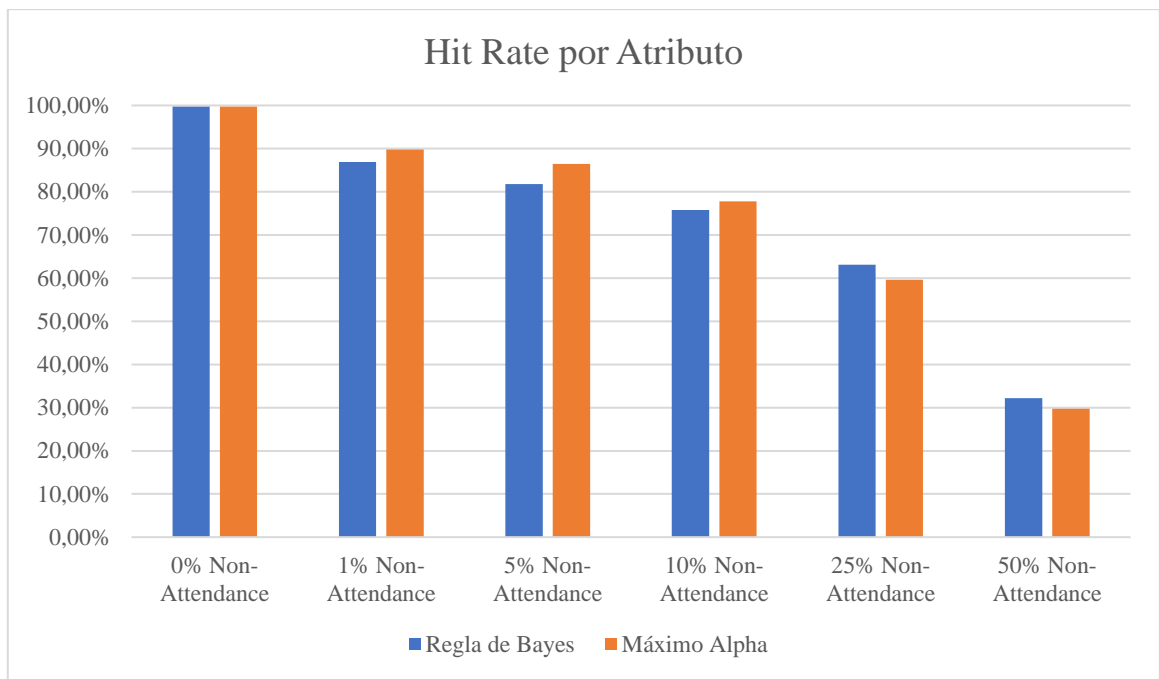


Gráfico 13: Hit Rate por Atributo Promedio para los seis Escenarios Base Simulados Datos Evaluación

En base a una mayor consistencia en todos los escenarios a través de todos los atributos y un mayor *hit rate*, se considera que la metodología Máximo Alpha es más adecuada para la asignación de clases que Regla de Bayes.

Complementando los resultados con los obtenidos en los datos de calibración y en las métricas anteriores, se fortalece la recomendación de utilizar Máximo Alpha como metodología de asignación. Máximo Alpha demuestra ser una metodología superior en todas las métricas presentadas y representa una fuente aceptable de información.

#### 4.2.7. Probabilidad Promedio de Observar Atributo

La probabilidad promedio de observar los atributos se calcula como el promedio simple de los valores de atención de los atributos  $\alpha_k$  obtenidos de las asignaciones que realizan las distintas metodologías sobre los individuos a las clases. Esto implica que, para cada metodología de asignación se realizará este cálculo.

Por ejemplo, si se poseen tres personas con las siguientes clases asignadas mediante Regla de Bayes:  $\{1, 1, 1, 1, 1, 1\}$ ,  $\{0, 1, 1, 0, 1, 0\}$  y  $\{0, 1, 0, 0, 1, 0\}$ . La probabilidad promedio de observar el primer atributo es  $(1 + 0 + 0)/3 = 33,33\%$ . Así, se obtiene una aproximación a la probabilidad de observar el atributo de la muestra en base a las asignaciones a las distintas clases.

La fórmula para calcular esta probabilidad para cada metodología es la siguiente:

$$Pr_k^{avg} = \frac{1}{n} * \sum_{n=1}^N \alpha_{k_n} \quad \forall k \in K$$

Se determina entonces el porcentaje de error entre la probabilidad con la que fue realizada la simulación y la probabilidad promedio de observar el atributo, de la siguiente forma:

$$error_k = |Pr_k^{avg} - Pr_k| \quad \forall k \in K$$

Por ejemplo, para el tercer escenario con 95% de probabilidad real simulada de observar los atributos el resultado sería el siguiente:

	Atributo 1	Atributo 2	Atributo 3	Atributo 4	Atributo 5	Atributo 6	Promedio
<b>Regla de Bayes</b>	95,88%	99,38%	99,25%	84,13%	70,88%	97,38%	91,15%
<b>Máximo Alpha</b>	100,00%	100,00%	100,00%	98,88%	93,63%	100,00%	98,75%
<b>Atención Atributo Individual</b>	87,19%	90,49%	88,07%	82,48%	75,62%	86,84%	82,15%
<b>Probabilidad Muestra</b>	96,38%	95,63%	95,50%	94,25%	95,50%	96,00%	95,54%

Tabla 4: Probabilidad de Observar Atributo Promedio para las Metodologías de Asignación Regla de Bayes y Máximo Alpha, además de Atención Atributo Individual versus Probabilidad Muestral Real Datos Calibración, Escenario 3

Donde la última fila muestra la probabilidad con la cual fue realizada la simulación. Así, el error para el Atributo 1 de Regla de Bayes sobre la probabilidad de la muestra tendría el siguiente cálculo:

$$error = |95,88\% - 96,38\%|$$

La Tabla 5 muestra los errores para Regla de Bayes, Máximo Alpha y Atención Atributo Individual para el ejemplo anterior del Escenario 3:

	<b>Atributo 1</b>	<b>Atributo 2</b>	<b>Atributo 3</b>	<b>Atributo 4</b>	<b>Atributo 5</b>	<b>Atributo 6</b>	<b>Promedio</b>
<b>Regla de Bayes</b>	0,50%	3,75%	3,75%	10,13%	24,63%	1,38%	7,35%
<b>Máximo Alpha</b>	3,63%	4,38%	4,50%	4,63%	1,87%	4,00%	3,83%
<b>Atención Atributo Individual</b>	9,18%	5,13%	7,43%	11,77%	19,88%	9,16%	10,43%

*Tabla 5: Error sobre Probabilidad Muestral Real para Asignación mediante Regla de Bayes y Máximo Alpha, además de Atención Atributo Individual Datos Calibración, Escenario 3*

Para este escenario, Máximo Alpha posee un menor porcentaje de error en promedio de todos los atributos.

Al realizar este mismo análisis para todos los escenarios se observa que en promedio Máximo Alpha posee un error más bajo que Regla de Bayes. Así mismo, el porcentaje de error de Atención Atributo Individual es en promedio menor que ambas metodologías tal como se muestra en la Tabla 6.

	<b>Escenario 1</b>	<b>Escenario 2</b>	<b>Escenario 3</b>	<b>Escenario 4</b>	<b>Escenario 5</b>	<b>Escenario 6</b>	<b>Promedio</b>
<b>Regla de Bayes</b>	0,00%	4,98%	7,35%	10,02%	17,06%	6,63%	7,67%
<b>Máximo Alpha</b>	0,00%	1,48%	3,83%	8,21%	20,15%	3,00%	6,11%
<b>Atención Atributo Individual</b>	0,00%	9,02%	10,43%	9,24%	5,49%	0,72%	5,82%

*Tabla 6: Error sobre Probabilidad Promedio para Regla de Bayes, Máximo Alpha y Atención Atributo Individual Datos Calibración*

Cabe destacar que todas las metodologías poseen una probabilidad promedio de observar el atributo igual a la probabilidad muestral para el escenario con atención absoluta a los atributos.

Esto conversa con las métricas anteriores, donde se observa que las metodologías son capaces de asignar de forma correcta en este escenario.

Al aumentar la probabilidad de observar *non-attendance* es posible observar errores mas altos en todas las metodologías. Destaca en estos escenarios que Atención Atributo Individual posee los errores más bajos, seguido por Máximo Alpha y quedando último Regla de Bayes.



### 4.2.3. Conclusiones Evaluación del Modelo

Los resultados obtenidos al analizar los datos de calibración muestran que el modelo propuesto posee un mejor desempeño y ajuste en todas las métricas a las cuales fue sometido que el modelo tradicional. Métricas como *hit rate*, AIC y BIC revelan que a mayor probabilidad de *non-attendance*, mayor diferencia existe entre los resultados obtenidos por ambos modelos. Esto implica que, a mayor probabilidad de *non-attendance* mayor diferencia podemos ver entre el desempeño del Modelo Logit con Clase Latente y el Modelo Logit. Cabe destacar que bajo un escenario con probabilidad de 1% de *non-attendance* ya es posible identificar estas diferencias en ajuste y desempeño a favor del modelo propuesto, además de este entregar mayor información.

El Modelo Logit con Clase Latente es capaz de identificar las probabilidades de atención para cada atributo con alta fiabilidad, con lo cual es posible identificar los atributos relevantes del producto. Esto, combinado con la importancia de los atributos entrega información muy valiosa a los investigadores. Del total de estimaciones realizadas el 22% de las veces el modelo propuesto estimó la probabilidad de atención a los atributos de forma exacta y el 39% de las veces el modelo erró por menos de un 0,1%. Los rangos de error varían entre 0% y 2%, donde el 92% de las estimaciones se sitúan bajo el 1%, teniendo un promedio de 0,38% y mediana de 0,27%. Estos resultados demuestran que el modelo es capaz de estimar la probabilidad de *non-attendance* en la muestra con bastante exactitud.

Por otro lado, con el fin de determinar la capacidad de predicción de los modelos sobre las decisiones de las personas, se utilizó la métrica *hit rate* obteniendo resultados favorables para el Modelo Logit con Clase Latente sobre el Modelo Logit. Para todos los escenarios el modelo propuesto obtuvo un *hit rate* mayor o igual al modelo tradicional, manteniéndose ambos sobre 55%. En los escenarios con menor probabilidad de *non-attendance*, se observan valores de *hit rate* sobre el 90% para ambos modelos.

Es necesario mencionar que, a mayor probabilidad de observar *non-attendance*, menor *hit rate* alcanzan ambos modelos. Los resultados muestran una tendencia clara a decrecer la capacidad de predicción al aumentar la probabilidad de ignorar atributos. Sin embargo, como se mencionó anteriormente los valores mínimos alcanzados en esta métrica se encuentran sobre el 55% por lo que sigue generando un aporte a los investigadores.

Para todas las métricas a las cuales fueron sometidos los modelos, los resultados obtenidos *out-of-sample* muestran tener la misma estructura que *in-sample*. Además, las métricas de ajuste de modelo muestran que a mayor probabilidad de ignorar atributos, mayor diferencia existe en el desempeño y ajuste a favor del modelo propuesto.

En cuanto a la ejecución, se observa que las diferencias en tiempos de procesamiento son considerables entre ambos modelos. Mientras que el Modelo Logit tarda menos de diez segundos, el Modelo Logit con Clase Latente demora entre una y dos horas en procesar. Hasta el momento, se ha trabajado con una base de datos de 9.600 preguntas, las cuales corresponden a 800 personas, por lo que un aumento importante en la cantidad de datos a procesar puede generar un aumento considerable en estos tiempos. Este debe ser un punto a considerar en las futuras aplicaciones, analizando el costo en tiempo versus el beneficio que conlleva poseer mayor información y estimaciones más precisas de las preferencias de los consumidores.

Con respecto a las metodologías de asignación, Máximo Alpha demuestra ser superior en todos los casos donde exista probabilidad no nula de observar *non-attendance* e igual de efectivo en los casos de atención completa a los atributos que Regla de Bayes. Además, Máximo Alpha demuestra poseer una estabilidad mayor en las métricas de evaluación a través de todos los atributos, lo que no se observa en Regla de Bayes. En base a esto, se considera adecuado utilizar esta metodología de asignación por sobre la segunda, ya que entrega resultados más confiables y consistentes a través de todos los atributos.

Un punto importante a considerar es que los resultados del Escenario 6 muestran que las metodologías de asignación utilizadas no son adecuadas ya que sus *hit rates* son nulos. Dado que estamos bajo una probabilidad de 50% de observar *non-attendance* a través de todos los atributos, se considera este escenario como extremo. Las matrices de confusión generadas para este escenario muestran una dispersión muy alta con respecto a los demás. En todos los demás escenarios, se observa una clara tendencia hacia la derecha o izquierda dependiendo del nivel de *non-attendance*, lo que no ocurre en el Escenario 6.

En base a los resultados de los demás escenarios, se espera que cuando al menos un atributo posea una probabilidad distinta a 50% de ser atendido, esto ayude a las metodologías a asignar de forma más precisa y se obtenga un mejor *hit rate*.

### 4.3. Aplicación del Modelo Logit con Clase Latente

#### 4.3.1. Datos

Los datos utilizados para el desarrollo de la investigación provienen del experimento lexicográfico realizado por Paula Urrutia y Delphine Faller, el cual consistía en un experimento de elección discreta donde los participantes debían escoger el mejor hotel en base a seis atributos: calidad de la comida, distancia al centro de la ciudad, categoría de la habitación, precio, recomendación de los clientes y vista al mar; cada uno con sus respectivos niveles de satisfacción.

Este experimento incluye además el uso de la herramienta *eye tracker*, la cual permite identificar las fijaciones que realizan los participantes sobre los atributos al momento de tomar sus decisiones y, por ende, muestra cuando un atributo es o no considerado.

	<b>Nivel de Satisfacción</b>		
<b>Atributo</b>	Bajo	Medio	Alto
<b>Calidad de comida</b>	Deficiente	Buena	Excelente
<b>Distancia al centro de la ciudad</b>	3 km	2 km	1 km
<b>Categoría de habitación</b>	Estándar	Superior	Deluxe
<b>Precio</b>	\$699.000	\$599.000	\$499.000
<b>Recomendación de clientes</b>	50%	70%	90%
<b>Vista al mar</b>	Sin vista al mar	Vista lateral al mar	Vista completa al mar

Tabla 7: Atributos versus Nivel de Satisfacción

Durante el experimento los participantes debían responder doce preguntas, donde se les solicitaba escoger entre cuatro alternativas un hotel para un amigo hipotético siguiendo una de seis reglas de decisión lexicográfica. Cada uno tenía información disponible sobre el funcionamiento de la regla asignada y los niveles mínimos a considerar para cada atributo. Esto quiere decir que, si un atributo no cumplía con los niveles mínimos establecidos no podía ser considerado en la toma de decisión. Así, se permite forzar en los participantes la ignorancia de atributos. El experimento fue diseñado forzando una única alternativa correcta luego de aplicada la regla.

Los encuestados debían seguir la misma regla en las doce preguntas, donde se variaban los niveles de cada atributo, con el fin de reproducir la elección de un mismo producto en reiteradas ocasiones. Para el experimento se obtuvo información de sesenta participantes, los que respondieron las doce preguntas designadas, dando un total de 720 preguntas a analizar.

Como se mencionó antes, este experimento fue realizado además con la herramienta *eye tracker*, la cual permite obtener información sobre las fijaciones que realizan los encuestados en la pantalla durante el experimento. Esto, ya que sigue los movimientos oculares que las personas realizan al mirar la pantalla mediante la utilización de lentes especiales. Así, se obtiene información sobre qué áreas se observaron en la pantalla, cuánto tiempo duró la observación y el orden en que fueron observadas.

Del total de participantes, los resultados muestran que un 42% no consideró algún atributo importante en su proceso de toma de decisiones, dada su regla de decisión asignada, durante por lo menos una de las doce preguntas que se les realizó. En otras palabras, cuando era requerido observar un atributo dada la regla lexicográfica asignada a cada participante, el 42% de las personas ignoraron en al menos una pregunta aquel atributo. Esto demuestra que a pesar de existir *non-attendance* forzado, se observa el fenómeno naturalmente. A continuación, se muestra la frecuencia con la que los participantes ignoraron los atributos, es decir, es posible observar *non-attendance*.

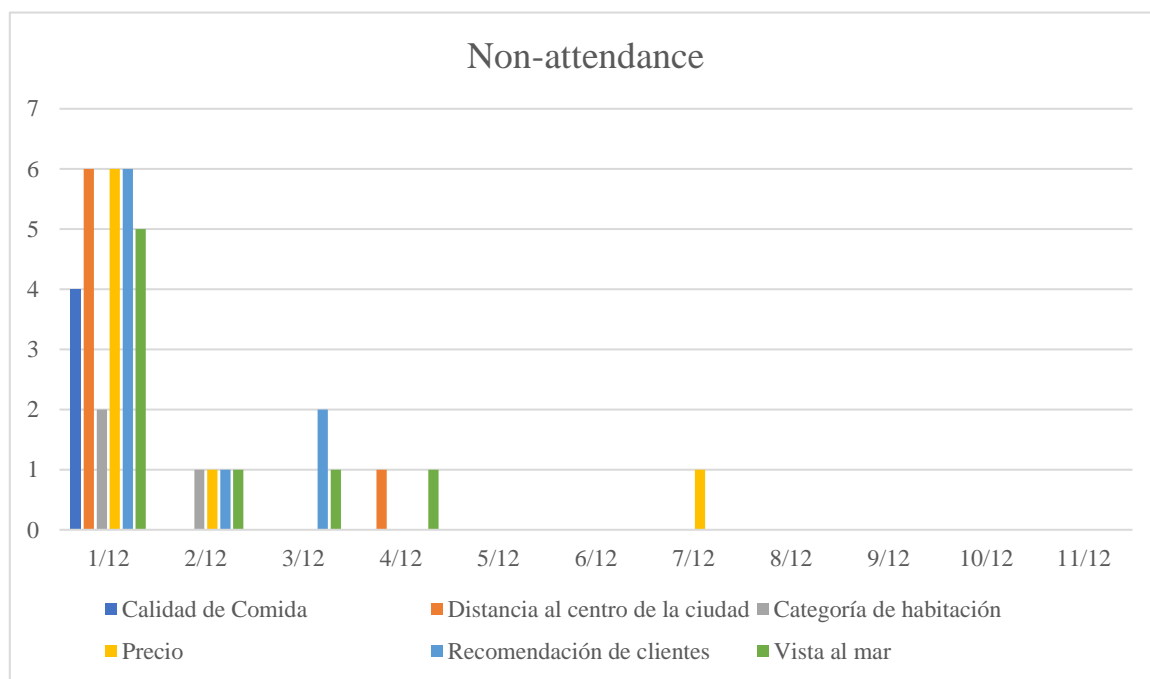


Gráfico 14: Frecuencia Non-Attendance sobre Preguntas Experimento Elección Discreta

Como se mencionó anteriormente, estos resultados no incorporan *non-attendance* forzado por las reglas de decisión lexicográficas, sino que representan casos en los cuales existió real ignorancia de atributos por parte de los participantes. Atributos que debiesen haberse considerado dada la regla lexicográfica asignada, fueron ignorados en al menos una pregunta por casi la mitad de los participantes.

Esto, dado que eran atributos que estaban forzados a considerar y aún así no fueron atendidos cuando los participantes tomaron sus decisiones.

### 4.3.2. Aplicación

La aplicación del MLCL en los datos originales es la etapa final de la investigación, donde se puede apreciar finalmente las implicancias que genera agregar un vector capaz de capturar el fenómeno *non-attendance*. Ya fue comprobado que el modelo propuesto es capaz de capturar las probabilidades de observar los atributos de forma precisa en la muestra, dependiendo del nivel de *non-attendance*, además de estimar las preferencias de los consumidores. Al igual que en la evaluación del modelo con datos simulados, para datos reales se separa la muestra en 10 preguntas de calibración y 2 preguntas de evaluación. El experimento cuenta con respuestas válidas de 60 personas, quedando 600 preguntas *in-sample* y 120 preguntas *out-of-sample*.

Para esta etapa se calculó, además, un Modelo Mixed Logit, en adelante MML, el cual toma en cuenta la heterogeneidad de los participantes asumiendo que los coeficientes varían de persona a persona, siguiendo una distribución con densidad  $f(\beta)$  en la población. Cabe destacar que el Modelo Logit es un caso particular del Modelo Mixed Logit, donde los coeficientes son fijos para la población.

Los modelos fueron sometidos a las métricas de evaluación *hit-rate*, criterio de información bayesiano y criterio de información de Akaike. La Tabla 8 muestra los resultados obtenidos por los tres modelos con respecto al *hit-rate* alcanzado tanto en datos de calibración como de evaluación.

	<b>Calibración</b>	<b>Evaluación</b>
<b>ML</b>	62,33%	57,50%
<b>MML</b>	60,50%	60,00%
<b>MLCL</b>	57,83%	59,17%

Tabla 8: Hit Rate Aplicado en Datos Reales

Con respecto a los datos de calibración, se observa que el modelo tradicional supera en aciertos tanto al Modelo Mixed Logit como al Modelo Logit con Clase Latente. Sin embargo, al observar los datos de evaluación podemos evidenciar que el Modelo Mixed Logit presenta el *hit-rate* más alto de los tres, seguido por el modelo propuesto y por último el tradicional. Esto sigue la lógica que el ML es un caso particular del MLCL, el cual a su vez lo es para el MML, por lo que los resultados son consistentes. A lo largo del experimento se observó que el *hit-rate* del modelo propuesto fue superior en todos los casos con presencia de *non-attendance* que el modelo tradicional, contrario a lo que vemos en los datos de calibración. Esto puede ser provocado porque la muestra no sea representativa debido a su tamaño. Los resultados obtenidos se consideran positivos para el modelo propuesto y aún más para el MML.

En cuanto a las métricas de ajuste, criterio de información bayesiano y criterio de información de Akaike, los resultados para los tres modelos se muestran en la Tabla 9.

	<b>Modelo</b>	<b>Log Verosimilitud</b>	<b>Muestra (n)</b>	<b>Variabes (k)</b>	<b>BIC</b>	<b>AIC</b>
<b>Calibración</b>	<b>ML</b>	-574,578	600	12	1.225,92	1.173,16
	<b>MML</b>	-541,000	600	12	1.158,76	1.106,00
	<b>MLCL</b>	-476,955	600	18	1.069,06	989,91

Tabla 9: Métricas de Ajuste Aplicadas en Datos Reales

En el análisis de los datos de calibración, es posible observar que el modelo que maximiza la Log Verosimilitud es el Modelo Logit con Clase Latente, el cual posee además un menor BIC y AIC, demostrando mejor ajuste a pesar de estar penalizado por el aumento de variables a estimar. Se destaca que, el ajuste del modelo propuesto supera tanto al ML como al MML, obteniendo mejores resultados en las métricas de ajuste, a pesar de ser segundo en los resultados de *hit rate*.

En tiempos de ejecución, se observa que el ML es 717 veces más rápido que el del MLCL en la calibración. Sin embargo, dado que los tiempos de ambos modelos son considerablemente bajos no se ve como un inconveniente para esta cantidad de datos ya que, la diferencia fue entre 0,675 segundos versus 8,061 minutos. De ser necesario analizar una muestra de datos de mayor magnitud, debe considerarse el costo entre obtener más información del modelo propuesto y el tiempo que demora en el procesamiento.

Las métricas asociadas a las asignaciones de los segmentos no pueden ser aplicadas en los datos reales, ya que no se posee información sobre la clase a la cual pertenece cada individuo. Se utilizan entonces *hit rate*, criterio de información bayesiano y criterio de información de Akaike para la comparación entre el modelo propuesto y el modelo tradicional.

Los coeficientes obtenidos en las estimaciones y sus respectivos errores estándar se observan en la Tabla 10.

		ML		MML		MLCL	
Atributo	Nivel	Coefficiente	Error Estándar	Coefficiente	Error Estándar	Coefficiente	Error Estándar
Calidad de Comida	$\beta_{Excelente}$	1,670*	0,164	56,66*	5,1	15,155*	1,877
	$\beta_{Buena}$	0,627*	0,158	19,17*	5,12	10,228*	1,600
Distancia al Centro de la Ciudad	$\beta_{2 km}$	0,384*	0,150	17,17*	4,18	0,895*	0,231
	$\beta_{1 km}$	0,932*	0,148	39,61*	4,48	2,029*	0,253
Categoría de Habitación	$\beta_{Deluxe}$	1,361*	0,157	54,59*	5,15	17,223*	2,436
	$\beta_{Superior}$	0,620*	0,163	23,63*	4,92	1,154*	1,996
Precio	$\beta_{Bajo}$	1,182*	0,154	39,1*	4,34	2,897*	0,437
	$\beta_{Medio}$	0,562*	0,148	19,85*	3,24	1,513*	0,327
Recomendación de Clientes	$\beta_{70\%}$	0,572*	0,162	22,65*	5,66	1,898*	0,368
	$\beta_{90\%}$	1,957*	0,167	87,3*	6,2	5,302*	0,533
Vista al mar	$\beta_{Completa}$	1,234*	0,148	36,66*	4,46	12,503*	1,598
	$\beta_{Parcial}$	0,546*	0,154	16,89*	4,45	4,125*	0,922

Tabla 10: Coeficientes y Errores Estándar Estimados Modelo Logit, Modelo Mixed Logit y Modelo Logit con Clase Latente, Datos Reales. \* = Coeficiente Significativo al 95%

Sus respectivas importancias se encuentran en la Tabla 11, las cuales permiten realizar comparaciones entre los tres modelos.

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	20,04%	11,18%	16,33%	14,18%	23,47%	14,80%
MML	18,05%	12,62%	17,39%	12,46%	27,81%	11,68%
MLCL	27,50%	3,68%	31,25%	5,26%	9,62%	22,69%

Tabla 11: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Datos Reales

Por último, las probabilidades de observar los atributos calculadas por el modelo propuesto son las siguientes:

<b>Atributo</b>	<b>Probabilidad de Atender Atributo</b>
<b>Calidad de Comida</b>	31,82%
<b>Distancia al Centro de la Ciudad</b>	99,98%
<b>Categoría de Habitación</b>	27,68%
<b>Precio</b>	82,18%
<b>Recomendación de Clientes</b>	80,58%
<b>Vista al Mar</b>	30,59%

Tabla 12: Probabilidad Estimada de Atención Atributos, Datos Reales

Los resultados demuestran la presencia del fenómeno *non-attendance* en el proceso de toma de decisiones de las personas. En tres de seis atributos se observan probabilidades cercanas al 70% de ignorar atributos, lo que implica que los consumidores son más proclives a no considerar estos atributos al tomar sus decisiones y, por ende, no importa cuánto mejoren estos atributos, no compensarán la disminución en el nivel de los otros. Los atributos con mayor probabilidad a ser ignorados son calidad de la comida, categoría de habitación y vista al mar, mientras que los atributos con mayor probabilidad a ser atendidos son distancia al centro de la ciudad, precio y recomendación de clientes.

Por otro lado, los valores de los coeficientes difieren con gran magnitud entre el MLCL y los otros dos modelos, notándose que los atributos con mayor probabilidad de ser ignorados son aquellos que poseen mayor importancia según el modelo propuesto. A medida que aumenta la probabilidad de *non-attendance*, los coeficientes estimados por el modelo tradicional tienden a disminuir en magnitud mientras que, los estimados por el modelo propuesto mantienen una magnitud cercana a la real. Esto, dado que el efecto provocado por este fenómeno se absorbe a través del vector  $\alpha$  y no a través de los coeficientes.

La importancia relativa de los atributos muestra que los atributos más importantes para los consumidores son efectivamente los con mayor probabilidad de *non-attendance*. Esto podría significar que a pesar de que no muchas personas atienden estos atributos en su proceso de toma de decisiones, las personas que si los atienden los consideran los más importantes en comparación con los demás.



El modelo tradicional, sin embargo, no entrega este tipo de información pudiendo provocar conclusiones sesgadas a los investigadores. Por ejemplo, es posible notar que el atributo recomendación de clientes es el considerado más importante por el ML, seguido por calidad de comida. No obstante, el modelo no incorpora la probabilidad de ignorar estos atributos, el cual demostraría que recomendación de clientes es uno de los atributos menos importante pero altamente considerado en las decisiones y que de forma opuesta, calidad de comida posee una importancia muy alta, pero es ignorado con una probabilidad cercana al 70%.

Finalmente, el Modelo Logit con Clase Latente propuesto en la investigación demuestra mejorar de forma importante las estimaciones y el ajuste del modelo. Además, entrega información valiosa para el investigador con respecto al comportamiento que poseen las personas en su proceso de toma de decisiones, al proveer probabilidades estimadas bastante precisas de observar *non-attendance*.

La clasificación de los consumidores en los distintos segmentos se considera adecuada, sin embargo, no óptima por lo que se propone una mayor profundización en investigaciones futuras.

A través de las simulaciones fue posible comprobar que el modelo propuesto es capaz de capturar el fenómeno *non-attendance* de forma efectiva en sus estimaciones, de forma agregada en la muestra, mientras que la aplicación en datos reales demuestra que el fenómeno se encuentra presente y el no considerarlo conlleva a estimaciones y conclusiones sesgadas y/o erróneas.

## 5. Conclusiones

El propósito de la investigación es estudiar el fenómeno de *non-attendance* mediante inferencia en los datos, utilizando un Modelo Logit con Clase. A través de simulaciones se generaron escenarios con distintas probabilidades de ignorar atributos. Estos escenarios permitieron evaluar y validar ambos modelos, antes de ser sometidos a datos reales. Los resultados son contundentes al demostrar superioridad del modelo propuesto sobre el modelo tradicional, en todas las métricas a las cuales fueron sometidos.

Las métricas de comparación de modelos, *hit rate*, criterio de información bayesiano y criterio de información de Akaike, demuestran tanto *in-sample* como *out-of-sample* que el modelo propuesto resulta ser mejor para realizar estimaciones sobre las preferencias de los consumidores que el modelo tradicional bajo situaciones donde exista probabilidad no nula de *non-attendance*, teniendo un mejor ajuste y desempeño. Esto, a pesar de que el Modelo Logit con Clase Latente incluye la estimación de seis parámetros adicionales, uno por cada probabilidad de observar el atributo. Bajo situaciones donde no exista probabilidad de ignorar atributos, ambos modelos resultan igualmente adecuados. En base a esto, se concluye que es recomendable utilizar en toda ocasión el modelo propuesto ya que, se ha visto en la teoría este fenómeno muy presente en los experimentos de elección discreta.

El modelo propuesto permite incorporar la heterogeneidad presente en el proceso de toma de decisiones de las personas directamente en las estimaciones de sus preferencias, identificando diferencias en sus utilidades. La ventaja que entrega la definición de las clases mediante el vector de elementos binarios  $\alpha$  es la reducción en la complejidad del modelo y, por ende, menores tiempos de procesamiento y capacidad computacional necesaria. Además, el Modelo Logit con Clase Latente resulta ser un modelo simple y fácil de interpretar por los investigadores que además les aporta mayor información.

La aplicación en datos reales demuestra la importancia de la información adicional que entrega el Modelo Logit con Clase Latente, además de estimar con mayor precisión los parámetros. Incorporar el fenómeno de *non-attendance* al modelo adecuadamente implica un mejor ajuste y menor sesgo en los parámetros del modelo, lo que conlleva finalmente a mejores estimaciones. Siempre que estemos bajo la presencia de una muestra representativa, el Modelo Logit con Clase Latente entrega ventajas sobre el Modelo Logit en las estimaciones de las preferencias de los consumidores, presentando un mejor ajuste y desempeño.

La investigación permitió comprobar que no es necesario utilizar preguntas complementarias para incorporar el fenómeno de *non-attendance* en las estimaciones. Esto, debido a que es posible identificar las probabilidades mediante el método de inferencia propuesto. Como se mencionó en los antecedentes, esto supone una ventaja en la investigación, eliminando los posibles errores humanos en la identificación de los atributos ignorados.

Se propone como trabajo futuro ampliar el modelo con tal de identificar *non-attendance* en los niveles de los atributos directamente. Se cree que así será más eficiente identificar en la muestra las preferencias de las personas y estimar de forma más efectiva las decisiones de los futuros consumidores. Esto tal vez conlleve a una mejor asignación de los segmentos en los consumidores, otro punto en el cual se propone profundizar en futuras investigaciones.

## 6. Bibliografía

Green, P. & Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. DOI: 10.2307/1251756

Lockwood, M. (1996). Non-compensatory preference structures in non-market valuation of natural area policy. DOI: 10.1111/j.1467-8489.1996.tb00557.x

Train, K. (2004). Discrete choice methods with simulation. DOI: 10.1017/CBO9780511805271. URL: <https://eml.berkeley.edu/books/choice2.html> [Consulta: 30 Septiembre 2018]

Hensher, D. A., Rose, J. & Greene, W. (2005). The implications on willingness to pay of respondents ignoring specific attributes. DOI: 10.1007/s11116-004-7613-8

Campbell, D., Hutchinson, G. & Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments. DOI: 10.1007/s10640-008-9198-8

Riedl, R., Brandstätter, E. & Roithmayr, F. (2008). Identifying decision strategies: a process- and outcome-based classification method. DOI: 10.3758/BRM.40.3.795

Scarpa, R., Gilbride, T.J., Campbell, D. & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. DOI: 10.1093/erae/jbp012

Campbell, D., Hensher, D. A. & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. DOI: 10.1080/09640568.2010.549367

Hensher, D. A., Rose, J. & Greene, W. (2011). Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design. DOI: 10.1007/s11116-011-9347-8

Alemu, M. H., Mørkbak, M. R., Olsen, S. B. & Jensen, C. L. (2012). Attending to the reasons for attribute non-attendance in choice experiments. DOI: 10.1007/s10640-012-9597-8

Erdem, S., Campbell, D. & Hole, A. R. (2013). Attribute-level non-attendance in a choice experiment investigating preferences for health service innovations. DOI: Sin información. URL: <http://www.icmconference.org.uk/index.php/icmc/icmc2013/paper/viewfile/717/351> [Consulta: 03 Julio 2018]

Weller, P., Oehlmann, M., Mariel, P. & Meyerhoff, J. (2014). Stated and inferred attribute non-attendance in a design of designs approach. DOI: 10.1016/j.jocm.2014.04.002

Erdem, S., Campbell, D. & Hole, A. R. (2015). Accounting for attribute-level non-attendance in a health choice experiment: Does it matter?. DOI: 10.1002/hec.3059

Dominique-Ferreira, S. (2011). Análisis conjunto: teoría, campos de aplicación y conceptos inherentes. DOI: Sin información. URL: [https://www.researchgate.net/publication/262752147\\_Analisis\\_conjunto\\_Teoria\\_campos\\_de\\_aplicacion\\_y\\_conceptos\\_inherentes/references](https://www.researchgate.net/publication/262752147_Analisis_conjunto_Teoria_campos_de_aplicacion_y_conceptos_inherentes/references) [Consulta: 03 Julio 2018]

Díaz, Fernanda Nicole, 2017, Evaluación de eficiencia y precisión en el contexto de decisiones repetidas basado en datos de eye-tracking. Memoria para optar al título de Ingeniera Civil Industrial, Santiago, Universidad de Chile.

Montoya, R. (2016). Curso Marketing II, Sección 1. Código: IN5602-1

## 7. Anexos

### 7.1. Coeficientes utilizados para la simulación

Atributo	Nivel	Coeficiente
<b>Atributo 1</b>	$\beta_{Atributo\ 1,Nivel\ 1}$	16,3910
	$\beta_{Atributo\ 1,Nivel\ 2}$	7,6663
<b>Atributo 2</b>	$\beta_{Atributo\ 2,Nivel\ 1}$	4,9001
	$\beta_{Atributo\ 2,Nivel\ 2}$	10,2472
<b>Atributo 3</b>	$\beta_{Atributo\ 3,Nivel\ 1}$	13,7469
	$\beta_{Atributo\ 3,Nivel\ 2}$	7,1236
<b>Atributo 4</b>	$\beta_{Atributo\ 4,Nivel\ 1}$	12,0345
	$\beta_{Atributo\ 4,Nivel\ 2}$	6,1505
<b>Atributo 5</b>	$\beta_{Atributo\ 5,Nivel\ 1}$	6,6865
	$\beta_{Atributo\ 5,Nivel\ 2}$	20,2488
<b>Atributo 6</b>	$\beta_{Atributo\ 6,Nivel\ 1}$	12,8012
	$\beta_{Atributo\ 6,Nivel\ 2}$	6,0152

Tabla 13: Coeficientes Base para Simulaciones

## 7.2. Listado de Clases

Clase	Vector Alpha	Clase	Vector Alpha	Clase	Vector Alpha
1	{0,0,0,0,0,0}	23	{0,1,0,1,1,0}	45	{1,0,1,1,0,0}
2	{0,0,0,0,0,1}	24	{0,1,0,1,1,1}	46	{1,0,1,1,0,1}
3	{0,0,0,0,1,0}	25	{0,1,1,0,0,0}	47	{1,0,1,1,1,0}
4	{0,0,0,0,1,1}	26	{0,1,1,0,0,1}	48	{1,0,1,1,1,1}
5	{0,0,0,1,0,0}	27	{0,1,1,0,1,0}	49	{1,1,0,0,0,0}
6	{0,0,0,1,0,1}	28	{0,1,1,0,1,1}	50	{1,1,0,0,0,1}
7	{0,0,0,1,1,0}	29	{0,1,1,1,0,0}	51	{1,1,0,0,1,0}
8	{0,0,0,1,1,1}	30	{0,1,1,1,0,1}	52	{1,1,0,0,1,1}
9	{0,0,1,0,0,0}	31	{0,1,1,1,1,0}	53	{1,1,0,1,0,0}
10	{0,0,1,0,0,1}	32	{0,1,1,1,1,1}	54	{1,1,0,1,0,1}
11	{0,0,1,0,1,0}	33	{1,0,0,0,0,0}	55	{1,1,0,1,1,0}
12	{0,0,1,0,1,1}	34	{1,0,0,0,0,1}	56	{1,1,0,1,1,1}
13	{0,0,1,1,0,0}	35	{1,0,0,0,1,0}	57	{1,1,1,0,0,0}
14	{0,0,1,1,0,1}	36	{1,0,0,0,1,1}	58	{1,1,1,0,0,1}
15	{0,0,1,1,1,0}	37	{1,0,0,1,0,0}	59	{1,1,1,0,1,0}
16	{0,0,1,1,1,1}	38	{1,0,0,1,0,1}	60	{1,1,1,0,1,1}
17	{0,1,0,0,0,0}	39	{1,0,0,1,1,0}	61	{1,1,1,1,0,0}
18	{0,1,0,0,0,1}	40	{1,0,0,1,1,1}	62	{1,1,1,1,0,1}
19	{0,1,0,0,1,0}	41	{1,0,1,0,0,0}	63	{1,1,1,1,1,0}
20	{0,1,0,0,1,1}	42	{1,0,1,0,0,1}	64	{1,1,1,1,1,1}
21	{0,1,0,1,0,0}	43	{1,0,1,0,1,0}		
22	{0,1,0,1,0,1}	44	{1,0,1,0,1,1}		

Tabla 14: Listado de Clases basado en Vector Alpha

### 7.3. Resultados de las estimaciones para los distintos escenarios y resultados de las evaluaciones para datos de calibración y evaluación

#### 7.3.1. Escenario 1

##### 7.3.1.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	16,376	0,588	16,377	0,588
$\beta_{\text{Comida Buena}}$	7,742	0,319	7,742	0,319
$\beta_{\text{Distancia 2 km}}$	4,926	0,206	4,926	0,206
$\beta_{\text{Distancia 1 km}}$	10,275	0,369	10,275	0,369
$\beta_{\text{Habitación Deluxe}}$	13,678	0,520	13,679	0,520
$\beta_{\text{Habitación Superior}}$	7,183	0,288	7,184	0,288
$\beta_{\text{Precio Bajo}}$	11,848	0,414	11,848	0,414
$\beta_{\text{Precio Medio}}$	6,115	0,222	6,116	0,222
$\beta_{\text{Recomendación 70}}$	6,729	0,287	6,730	0,287
$\beta_{\text{Recomendación 90}}$	20,018	0,718	20,019	0,718
$\beta_{\text{Vista Completa al Mar}}$	12,422	0,466	12,422	0,466
$\beta_{\text{Vista Parcial al Mar}}$	5,755	0,238	5,755	0,238

Tabla 15: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 1

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,35%	12,14%	16,16%	14,00%	23,66%	14,68%
MLCL	19,35%	12,14%	16,16%	14,00%	23,66%	14,68%

Tabla 16: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 1

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	16,437	100,00%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	12,328	100,00%
$\beta_{\text{Categoría de Habitación}}$	15,045	100,00%
$\beta_{\text{Precio}}$	15,208	100,00%
$\beta_{\text{Recomendación de Clientes}}$	15,913	100,00%
$\beta_{\text{Vista al Mar}}$	15,579	100,00%

Tabla 17: Probabilidad Estimada de Atención Atributos, Escenario 1



### 7.3.1.2. Calibración

Modelo	Hit rate	Tiempo
ML	95,95%	7,088 segundos
MLCL	95,95%	1,147 horas

Tabla 18: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 1

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-836,91	8.000	12	1.781,67	1.697,82
MLCL	-836,91	8.000	18	1.835,59	1.709,82

Tabla 19: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 1

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
Máximo Alpha	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tabla 20: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 1

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
Máximo Alpha	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%
Atención Atributo Individual	100,00%	99,98%	100,00%	100,00%	100,00%	100,00%
Probabilidad Simulada	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tabla 21: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 1

### 7.3.1.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	94,69%	N/A
<b>MLCL</b>	94,69%	N/A

Tabla 22: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 1

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	100,00%	100,00%	100,00%	100,00%	98,25%	100,00%
<b>Máximo Alpha</b>	100,00%	100,00%	100,00%	100,00%	98,25%	100,00%

Tabla 23: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 1

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	100,00%	100,00%	100,00%	100,00%	98,25%	100,00%
<b>Máximo Alpha</b>	100,00%	100,00%	100,00%	100,00%	98,25%	100,00%
<b>Atención Atributo Individual</b>	99,98%	99,28%	99,72%	99,95%	98,80%	99,93%
<b>Probabilidad Simulada</b>	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tabla 24: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 1

### 7.3.2. Escenario 2

#### 7.3.2.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	10,178	0,272	16,250	0,625
$\beta_{\text{Comida Buena}}$	4,634	0,160	7,521	0,325
$\beta_{\text{Distancia 2 km}}$	3,085	0,122	4,896	0,221
$\beta_{\text{Distancia 1 km}}$	6,316	0,179	9,927	0,392
$\beta_{\text{Habitación Deluxe}}$	8,582	0,248	13,751	0,551
$\beta_{\text{Habitación Superior}}$	4,499	0,159	7,185	0,310
$\beta_{\text{Precio Bajo}}$	7,693	0,201	11,817	0,440
$\beta_{\text{Precio Medio}}$	4,095	0,128	6,204	0,240
$\beta_{\text{Recomedación 70}}$	4,042	0,150	6,406	0,292
$\beta_{\text{Recomendación 90}}$	12,529	0,329	19,882	0,763
$\beta_{\text{Vista Completa al Mar}}$	7,944	0,218	12,707	0,503
$\beta_{\text{Vista Parcial al Mar}}$	3,779	0,132	6,001	0,260

Tabla 25: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 2

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,12%	11,86%	16,12%	14,45%	23,53%	14,92%
MLCL	19,27%	11,77%	16,31%	14,01%	23,58%	15,07%

Tabla 26: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 2

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	3,967	98,14%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	4,127	98,41%
$\beta_{\text{Categoría de Habitación}}$	4,198	98,52%
$\beta_{\text{Precio}}$	5,143	99,42%
$\beta_{\text{Recomendación de Clientes}}$	4,577	98,98%
$\beta_{\text{Vista al Mar}}$	4,732	99,13%

Tabla 27: Probabilidad Estimada de Atención Atributos, Escenario 2

### 7.3.2.2. Calibración

Modelo	Hit rate	Tiempo
ML	94,20%	6,250 segundos
MLCL	94,26%	1,496 horas

Tabla 28: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 2

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-1.391,62	8.000	12	2.891,09	2.807,25
MLCL	-1.107,37	8.000	18	2.376,50	2.250,73

Tabla 29: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 2

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	81,75%	99,00%	98,63%	99,63%	87,50%	99,13%
Máximo Alpha	95,50%	99,00%	98,63%	99,63%	94,75%	99,13%

Tabla 30: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 2

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	83,00%	100,00%	100,00%	100,00%	88,50%	100,00%
Máximo Alpha	96,75%	100,00%	100,00%	100,00%	95,75%	100,00%
Atención Atributo Individual	82,33%	93,11%	90,34%	96,06%	84,23%	93,93%
Probabilidad Simulada	98,75%	99,00%	98,63%	99,63%	99,00%	99,13%

Tabla 31: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 2

### 7.3.2.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	93,75%	N/A
<b>MLCL</b>	93,75%	N/A

Tabla 32: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 2

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	77,25%	90,63%	78,88%	96,25%	87,50%	91,00%
<b>Máximo Alpha</b>	83,13%	92,25%	82,00%	97,88%	90,88%	92,75%

Tabla 33: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 2

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	78,50%	91,63%	80,25%	96,63%	88,50%	91,88%
<b>Máximo Alpha</b>	84,38%	93,25%	83,38%	98,25%	91,88%	93,63%
<b>Atención Atributo Individual</b>	76,80%	88,81%	78,76%	94,35%	86,72%	87,96%
<b>Probabilidad Simulada</b>	98,75%	99,00%	98,63%	99,63%	99,00%	99,13%

Tabla 34: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 2

### 7.3.3. Escenario 3

#### 7.3.3.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	6,254	0,135	16,441	0,648
$\beta_{\text{Comida Buena}}$	2,796	0,092	7,541	0,337
$\beta_{\text{Distancia 2 km}}$	1,864	0,079	4,926	0,228
$\beta_{\text{Distancia 1 km}}$	3,906	0,100	10,294	0,418
$\beta_{\text{Habitación Deluxe}}$	4,936	0,119	13,752	0,564
$\beta_{\text{Habitación Superior}}$	2,587	0,095	7,125	0,317
$\beta_{\text{Precio Bajo}}$	4,731	0,107	12,351	0,480
$\beta_{\text{Precio Medio}}$	2,451	0,080	6,193	0,253
$\beta_{\text{Recomedación 70}}$	2,240	0,090	6,529	0,302
$\beta_{\text{Recomendación 90}}$	7,279	0,154	20,042	0,781
$\beta_{\text{Vista Completa al Mar}}$	4,777	0,109	12,973	0,532
$\beta_{\text{Vista Parcial al Mar}}$	2,337	0,083	6,193	0,281

Tabla 35: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 3

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,62%	12,25%	15,48%	14,84%	22,83%	14,98%
MLCL	19,15%	11,99%	16,02%	14,39%	23,34%	15,11%

Tabla 36: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 3

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	3,393	96,75%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	3,085	95,63%
$\beta_{\text{Categoría de Habitación}}$	3,084	95,62%
$\beta_{\text{Precio}}$	2,707	93,75%
$\beta_{\text{Recomendación de Clientes}}$	3,049	95,47%
$\beta_{\text{Vista al Mar}}$	3,066	95,55%

Tabla 37: Probabilidad Estimada de Atención Atributos, Escenario 3



### 7.3.3.2. Calibración

Modelo	Hit rate	Tiempo
ML	90,58%	5,93 segundos
MLCL	90,63%	2,575 horas

Tabla 38: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 3

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-2.512,34	8.000	12	5.132,53	5.048,69
MLCL	-1.600,09	8.000	18	3.361,94	3.236,17

Tabla 39: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 3

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	92,25%	95,00%	94,75%	78,38%	66,38%	93,38%
Máximo Alpha	96,38%	95,63%	95,50%	93,13%	89,13%	96,00%

Tabla 40: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 3

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	95,88%	99,38%	99,25%	84,13%	70,88%	97,38%
Máximo Alpha	100,00%	100,00%	100,00%	98,88%	93,63%	100,00%
Atención Atributo Individual	87,19%	90,49%	88,07%	82,48%	75,62%	86,84%
Probabilidad Simulada	96,38%	95,63%	95,50%	94,25%	95,50%	96,00%

Tabla 41: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 3

### 7.3.3.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	90,00%	N/A
<b>MLCL</b>	90,00%	N/A

Tabla 42: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 3

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	80,25%	84,88%	78,38%	79,25%	83,63%	84,50%
<b>Máximo Alpha</b>	86,00%	89,25%	82,75%	85,75%	86,50%	88,50%

Tabla 43: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 3

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	83,88%	89,00%	81,63%	84,75%	88,13%	88,25%
<b>Máximo Alpha</b>	89,63%	93,63%	86,00%	91,25%	91,00%	92,25%
<b>Atención Atributo Individual</b>	81,13%	87,57%	77,33%	82,15%	81,20%	79,19%
<b>Probabilidad Simulada</b>	96,38%	95,63%	95,50%	94,25%	95,50%	96,00%

Tabla 44: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 3

### 7.3.4. Escenario 4

#### 7.3.4.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	4,102	0,086	17,202	0,695
$\beta_{\text{Comida Buena}}$	1,886	0,068	7,994	0,355
$\beta_{\text{Distancia 2 km}}$	1,220	0,061	5,090	0,247
$\beta_{\text{Distancia 1 km}}$	2,575	0,070	10,705	0,458
$\beta_{\text{Habitación Deluxe}}$	3,466	0,081	14,263	0,592
$\beta_{\text{Habitación Superior}}$	1,788	0,072	7,440	0,332
$\beta_{\text{Precio Bajo}}$	3,188	0,073	12,642	0,514
$\beta_{\text{Precio Medio}}$	1,698	0,062	6,428	0,272
$\beta_{\text{Recomendación 70}}$	1,550	0,069	6,982	0,321
$\beta_{\text{Recomendación 90}}$	4,898	0,097	21,101	0,854
$\beta_{\text{Vista Completa al Mar}}$	3,274	0,073	13,402	0,566
$\beta_{\text{Vista Parcial al Mar}}$	1,519	0,063	6,204	0,288

Tabla 45: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 4

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,08%	11,97%	16,12%	14,82%	22,78%	15,23%
MLCL	19,26%	11,99%	15,97%	14,15%	23,63%	15,01%

Tabla 46: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 4

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta$ <i>Calidad de Comida</i>	2,071	88,81%
$\beta$ <i>Distancia al Centro de la Ciudad</i>	2,066	88,75%
$\beta$ <i>Categoría de Habitación</i>	2,201	90,04%
$\beta$ <i>Precio</i>	2,066	88,76%
$\beta$ <i>Recomendación de Clientes</i>	2,061	88,70%
$\beta$ <i>Vista al Mar</i>	2,331	91,14%

Tabla 47: Probabilidad Estimada de Atención Atributos, Escenario 4

### 7.3.4.2. Calibración

Modelo	Hit rate	Tiempo
ML	84,59%	4,997 segundos
MLCL	84,64%	55,734 minutos

Tabla 48: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 4

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-3.770,50	8.000	12	7.648,85	7.565,00
MLCL	-2.327,10	8.000	18	4.815,97	4.690,20

Tabla 49: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 4

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	74,25%	87,50%	89,38%	86,63%	53,50%	90,88%
Máximo Alpha	88,00%	89,38%	90,38%	89,13%	77,25%	91,13%

Tabla 50: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 4

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	85,50%	98,13%	99,00%	96,88%	64,88%	99,75%
Máximo Alpha	99,25%	100,00%	100,00%	99,38%	88,63%	100,00%
Atención Atributo Individual	77,82%	84,08%	83,75%	81,60%	70,66%	84,66%
Probabilidad Simulada	88,75%	89,38%	90,38%	89,75%	88,63%	91,13%

Tabla 51: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 4

### 7.3.4.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	83,44%	N/A
<b>MLCL</b>	83,69%	N/A

Tabla 52: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 4

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	61,88%	83,88%	75,63%	79,50%	71,38%	82,25%
<b>Máximo Alpha</b>	69,13%	84,13%	74,25%	83,13%	72,00%	84,13%

Tabla 53: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 4

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	70,63%	93,25%	84,25%	88,75%	81,75%	91,13%
<b>Máximo Alpha</b>	78,88%	93,50%	82,63%	92,38%	82,38%	93,00%
<b>Atención Atributo Individual</b>	71,57%	80,84%	74,67%	80,27%	75,21%	77,58%
<b>Probabilidad Simulada</b>	88,75%	89,38%	90,38%	89,75%	88,63%	91,13%

Tabla 54: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 4

### 7.3.5. Escenario 5

#### 7.3.5.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	2,623	0,059	16,281	0,641
$\beta_{\text{Comida Buena}}$	1,209	0,053	7,637	0,327
$\beta_{\text{Distancia 2 km}}$	0,785	0,049	4,799	0,244
$\beta_{\text{Distancia 1 km}}$	1,607	0,051	10,101	0,433
$\beta_{\text{Habitación Deluxe}}$	2,228	0,056	14,025	0,569
$\beta_{\text{Habitación Superior}}$	1,199	0,055	7,278	0,319
$\beta_{\text{Precio Bajo}}$	2,005	0,053	11,941	0,489
$\beta_{\text{Precio Medio}}$	1,061	0,049	6,341	0,276
$\beta_{\text{Recomedación 70}}$	0,887	0,053	6,457	0,306
$\beta_{\text{Recomendación 90}}$	3,016	0,061	19,911	0,785
$\beta_{\text{Vista Completa al Mar}}$	2,039	0,051	12,476	0,514
$\beta_{\text{Vista Parcial al Mar}}$	0,904	0,050	5,710	0,267

Tabla 55: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 5

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,41%	11,89%	16,48%	14,83%	22,31%	15,08%
MLCL	19,21%	11,92%	16,55%	14,09%	23,50%	14,72%

Tabla 56: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 5

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	1,079	74,64%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	0,986	72,83%
$\beta_{\text{Categoría de Habitación}}$	1,117	75,35%
$\beta_{\text{Precio}}$	1,174	76,39%
$\beta_{\text{Recomendación de Clientes}}$	1,157	76,09%
$\beta_{\text{Vista al Mar}}$	1,251	77,75%

Tabla 57: Probabilidad Estimada de Atención Atributos, Escenario 5



### 7.3.5.2. Calibración

Modelo	Hit rate	Tiempo
ML	73,84%	5,916 segundos
MLCL	73,95%	1,598 horas

Tabla 58: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 5

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-5.643,26	8.000	12	11.394,36	11.310,52
MLCL	-3.500,13	8.000	18	7.162,02	7.036,25

Tabla 59: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 5

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	62,75%	70,50%	71,88%	74,00%	49,38%	77,38%
Máximo Alpha	69,75%	72,25%	74,75%	74,50%	59,25%	78,63%

Tabla 60: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 5

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	87,00%	97,63%	96,13%	98,75%	73,50%	98,75%
Máximo Alpha	94,00%	99,38%	99,00%	99,25%	83,38%	100,00%
Atención Atributo Individual	68,15%	70,56%	70,97%	72,88%	65,04%	73,59%
Probabilidad Simulada	75,75%	72,88%	75,75%	75,25%	75,88%	78,63%

Tabla 61: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 5

### 7.3.5.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	73,94%	N/A
<b>MLCL</b>	74,13%	N/A

Tabla 62: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 5

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	47,88%	66,25%	63,50%	69,13%	60,75%	71,00%
<b>Máximo Alpha</b>	45,75%	64,00%	58,38%	67,50%	56,50%	65,75%

Tabla 63: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 5

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	66,38%	90,88%	85,50%	92,88%	82,88%	90,13%
<b>Máximo Alpha</b>	65,25%	88,38%	80,38%	91,25%	77,63%	82,13%
<b>Atención Atributo Individual</b>	62,40%	68,51%	66,32%	70,76%	66,85%	67,73%
<b>Probabilidad Simulada</b>	75,75%	72,88%	75,75%	75,25%	75,88%	78,63%

Tabla 64: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 5

### 7.3.6. Escenario 6

#### 7.3.6.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{Comida\ Excelente}$	1,539	0,043	16,296	0,687
$\beta_{Comida\ Buena}$	0,700	0,042	7,592	0,355
$\beta_{Distancia\ 2\ km}$	0,374	0,040	4,733	0,286
$\beta_{Distancia\ 1\ km}$	0,891	0,039	10,163	0,495
$\beta_{Habitación\ Deluxe}$	1,210	0,041	14,041	0,609
$\beta_{Habitación\ Superior}$	0,571	0,043	7,194	0,348
$\beta_{Precio\ Bajo}$	1,076	0,041	12,192	0,572
$\beta_{Precio\ Medio}$	0,545	0,039	6,426	0,334
$\beta_{Recomedación\ 70}$	0,495	0,042	6,473	0,349
$\beta_{Recomendación\ 90}$	1,617	0,043	20,289	0,891
$\beta_{Vista\ Completa\ al\ Mar}$	1,161	0,039	13,352	0,602
$\beta_{Vista\ Parcial\ al\ Mar}$	0,495	0,041	6,505	0,341

Tabla 65: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 6

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	20,54%	11,89%	16,14%	14,36%	21,57%	15,50%
MLCL	18,88%	11,77%	16,26%	14,12%	23,50%	15,47%

Tabla 66: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 6

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	0,137	53,41%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	-0,084	47,90%
$\beta_{\text{Categoría de Habitación}}$	0,030	50,76%
$\beta_{\text{Precio}}$	-0,041	48,97%
$\beta_{\text{Recomendación de Clientes}}$	0,100	52,49%
$\beta_{\text{Vista al Mar}}$	0,018	50,44%

Tabla 67: Probabilidad Estimada de Atención Atributos, Escenario 6

### 7.3.6.2. Calibración

Modelo	Hit rate	Tiempo
ML	58,16%	5,62 segundos
MLCL	58,25%	1,081 horas

Tabla 68: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 6

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-8.128,94	8.000	12	16.365,72	16.281,88
MLCL	-5.098,75	8.000	18	10.359,27	10.233,50

Tabla 69: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 6

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	21,25%	31,63%	18,50%	27,00%	11,75%	17,63%
Máximo Alpha	12,13%	21,00%	10,75%	18,63%	6,25%	12,50%

Tabla 70: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 6

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	60,50%	33,50%	55,00%	44,50%	55,25%	52,50%
Máximo Alpha	55,13%	42,13%	52,25%	51,13%	51,50%	53,63%
Atención Atributo Individual	53,12%	48,83%	51,25%	49,96%	52,13%	51,19%
Probabilidad Simulada	51,75%	48,88%	50,25%	49,50%	52,00%	49,88%

Tabla 71: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 6

### 7.3.6.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	55,56%	N/A
<b>MLCL</b>	55,81%	N/A

Tabla 72: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 6

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	26,50%	36,00%	36,00%	33,50%	29,75%	31,50%
<b>Máximo Alpha</b>	24,63%	33,00%	33,63%	32,13%	26,75%	28,25%

Tabla 73: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 6

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	56,00%	39,38%	56,50%	46,25%	56,00%	53,13%
<b>Máximo Alpha</b>	54,38%	43,63%	52,38%	48,63%	50,25%	46,38%
<b>Atención Atributo Individual</b>	52,51%	47,64%	49,31%	49,25%	50,62%	48,40%
<b>Probabilidad Simulada</b>	51,75%	48,88%	50,25%	49,50%	52,00%	49,88%

Tabla 74: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 6

### 7.3.7. Escenario 7

#### 7.3.7.1. Estimaciones

	ML		MLCL	
	Coefficiente	Error Estándar	Coefficiente	Error Estándar
$\beta_{\text{Comida Excelente}}$	0,701	0,036	16,456	0,891
$\beta_{\text{Comida Buena}}$	0,260	0,036	7,573	0,500
$\beta_{\text{Distancia 2 km}}$	0,153	0,035	4,475	0,375
$\beta_{\text{Distancia 1 km}}$	0,396	0,034	9,909	0,609
$\beta_{\text{Habitación Deluxe}}$	0,595	0,035	13,730	0,786
$\beta_{\text{Habitación Superior}}$	0,288	0,037	7,616	0,515
$\beta_{\text{Precio Bajo}}$	0,607	0,035	11,826	0,684
$\beta_{\text{Precio Medio}}$	0,254	0,035	6,041	0,408
$\beta_{\text{Recomendación 70}}$	0,182	0,036	6,574	0,516
$\beta_{\text{Recomendación 90}}$	0,675	0,035	20,248	1,157
$\beta_{\text{Vista Completa al Mar}}$	0,648	0,034	12,756	0,731
$\beta_{\text{Vista Parcial al Mar}}$	0,306	0,036	6,120	0,432

Tabla 75: Estimaciones Coeficientes Modelo Logit y Modelo Logit con Clase Latente, Escenario 7

Modelo	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
ML	19,37%	10,93%	16,42%	16,76%	18,65%	17,88%
MLCL	19,38%	11,67%	16,17%	13,93%	23,84%	15,02%

Tabla 76: Importancia Atributos Modelo Logit y Modelo Logit con Clase Latente, Escenario 7

	<b>Coefficiente</b>	<b>Probabilidad de Atención</b>
$\beta_{\text{Calidad de Comida}}$	-1,073	25,48%
$\beta_{\text{Distancia al Centro de la Ciudad}}$	-1,273	21,87%
$\beta_{\text{Categoría de Habitación}}$	-1,096	25,04%
$\beta_{\text{Precio}}$	-0,996	26,97%
$\beta_{\text{Recomendación de Clientes}}$	-1,092	25,12%
$\beta_{\text{Vista al Mar}}$	-1,018	26,53%

Tabla 77: Probabilidad Estimada de Atención Atributos, Escenario 7



### 7.3.7.2. Calibración

Modelo	Hit rate	Tiempo
ML	42,74%	4,663 segundos
MLCL	42,98%	1,019 horas

Tabla 78: Hit Rate y Tiempo de Cálculo Modelo Logit y Modelo Logit con Clase Latente, Datos Calibración, Escenario 7

Modelo	Log Verosimilitud	Muestra (n)	Variabes (k)	BIC	AIC
ML	-10.122,18	8.000	12	20.352,21	20.268,36
MLCL	-7.185,88	8.000	18	14.533,54	14.407,77

Tabla 79: Máxima Verosimilitud, BIC y AIC, Datos Calibración, Escenario 7

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	52,88%	77,63%	66,63%	66,38%	28,63%	62,13%
Máximo Alpha	72,13%	78,25%	75,50%	73,25%	66,63%	74,50%

Tabla 80: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Calibración, Escenario 7

Asignación	Calidad de Comida	Distancia al Centro de la Ciudad	Categoría de Habitación	Precio	Recomendación de Clientes	Vista al Mar
Regla de Bayes	22,00%	0,63%	8,88%	8,00%	46,75%	13,50%
Máximo Alpha	2,25%	0,00%	0,00%	0,13%	8,00%	0,38%
Atención Atributo Individual	32,53%	25,86%	30,49%	31,58%	35,20%	31,99%
Probabilidad Simulada	25,63%	21,75%	24,50%	26,63%	25,38%	25,13%

Tabla 81: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Calibración, Escenario 7

### 7.3.7.3. Evaluación

<b>Modelo</b>	<b>Hit rate</b>	<b>Tiempo</b>
<b>ML</b>	44,13%	N/A
<b>MLCL</b>	44,06%	N/A

Tabla 82: Hit Rate Modelo Logit y Modelo Logit con Clase Latente, Datos Evaluación, Escenario 7

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	48,38%	74,50%	63,63%	61,88%	46,50%	65,13%
<b>Máximo Alpha</b>	56,50%	77,38%	68,75%	67,75%	57,88%	66,75%

Tabla 83: Hit Rate por Atributo Regla de Bayes y Máximo Alpha, Datos Evaluación, Escenario 7

<b>Asignación</b>	<b>Calidad de Comida</b>	<b>Distancia al Centro de la Ciudad</b>	<b>Categoría de Habitación</b>	<b>Precio</b>	<b>Recomendación de Clientes</b>	<b>Vista al Mar</b>
<b>Regla de Bayes</b>	29,25%	4,50%	12,63%	12,50%	30,63%	10,50%
<b>Máximo Alpha</b>	18,88%	0,88%	7,25%	5,63%	17,75%	8,63%
<b>Atención Atributo Individual</b>	32,98%	25,37%	30,43%	31,28%	33,01%	30,43%
<b>Probabilidad Simulada</b>	25,63%	21,75%	24,50%	26,63%	25,38%	25,13%

Tabla 84: Probabilidad Promedio de Observar Atributo Regla de Bayes, Máximo Alpha, Atención Atributo Individual y Muestra Simulada, Datos Evaluación, Escenario 7























## 7.4.6. Escenario 6

### 7.4.6.1. Regla de Bayes

Ilustración 21: Matriz de Confusión Asignación Regla de Bayes, Datos Calibración, Escenario 6

Ilustración 22: Matriz de Confusión Asignación Regla de Bayes, Datos Evaluación, Escenario 6





