Resumen Marketing II

Gustavo Avilés Segovia

16 de diciembre de 2013

Capítulo 1

Recuerdo Marketing I

1.0.1. Entender el Mercado

Las tres **C**: **Cliente** $\leftrightarrow$**Compañía** $\leftrightarrow$**Competencia**

1.0.2. Diseñar una estrategia de Marketing

STP: **Segmentación** $\leftrightarrow$**Targeting** $\leftrightarrow$**Posicionamiento**

1.0.3. Plan de Marketing

Las 4 **P:** **Producto** $\leftrightarrow$ **Precio** $\leftrightarrow$ **Promoción** $\leftrightarrow$ **Plaza**

1.1. Tipos de modelos

Depende de el supuesto de comportamiento de usuario.

Probabilísticos Los usuarios actúan de forma aleatoria.

Estructurales Los usuarios actúan de forma racional.

Agregados Usan datos agregados

Desagregados Usan datos desagregados.

Adicionalmente se puede dividir según la naturaleza de sus variables, que puede ser *Continua*, *Discreta* o *Mixta*.

Capítulo 2

Modelos de Comportamiento

Estos son modelos que buscan predecir el comportamiento de los usuarios, considerando que estos toman sus decisiones de forma aleatoria.

Los modelos de comportamiento se clasifican básicamente en:

1. *Tiempo*, resuelven preguntas del tipo ¿Cuándo?
   a) Tiempo Continuo.
      1) Proyecciones de adopción de nuevos productos.
   b) Tiempo Discreto
      1) Proyección de tasas de retención de clientes.
2. *Conteo*, resuelve preguntas del tipo ¿Cuántos?
   a) Estimaciones de exposición publicitaria.
3. Elección, resuelve preguntas del tipo ¿Cuál?
   a) Respuestas a una campaña de marketing directo.

Estos modelos pueden ser utilizados de forma conjunta, para modelar situaciones más complejas.

2.1. Proyección de Retención de Clientes

En este modelo se pretende entender el comportamiento de los clientes en cuanto a el tiempo de permanencia en un servicio.

2.1.1. Modelo sencillo: Distribución Geométrica Desplazada

2.1.1.1. Homogéneo.

$$Pr\left(T=t|\theta\right)=\theta\left(1-\theta\right)^{t-1}\quad Pr\left(T>t|\theta\right)=\left(1-\theta\right)^{t}$$

2.1.1.2. Heterogéneo Discreto: Con dos grupos de Clientes.

Para esto se estima que existen dos grupos de clientes los π y $1-\pi$. Cada uno de los segmentos posee un $\theta_i$diferente, agregando heterogeneidad al modelo.

$$Pr\left(T=t|\theta_1,\theta_2,\pi\right)=\theta_1\left(1-\theta_1\right)^{t-1}\pi+\theta_2\left(1-\theta_2\right)^{t-1}\left(1-\pi\right)$$

$$Pr\left(T>t|\theta_1,\theta_2,\pi\right)=\left(1-\theta_1\right)^{t}\pi+\left(1-\theta_2\right)^{t}\left(1-\pi\right)$$

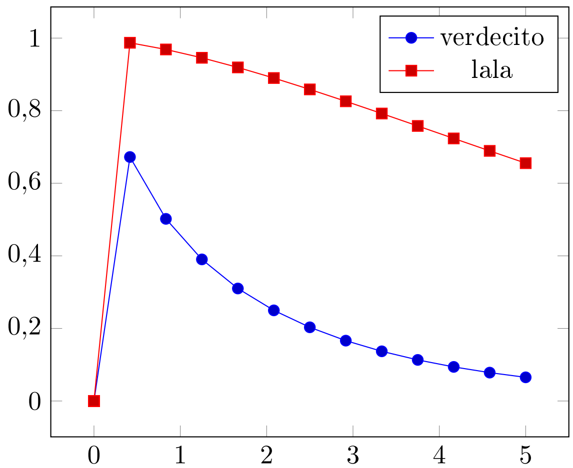
2.1.1.3. Heterogéneo Continuo: Distribución Beta.

Al igual que el modelo visto anteriormente, si se calculase una cantidad infinita de segmentos se puede utilizar una distribución continua, llamada mezcla continua de distribuciones geométricas, para modelar el comportamiento de los usuarios.

$$Pr\left(T=t|\alpha,\beta\right)=\begin{cases}\frac{\alpha}{\alpha+\beta} & t=1\\ \frac{\beta+t-2}{\alpha+\beta+t-1}Pr\left(T=t-1|\alpha,\beta\right) & t=2,3,4,...\end{cases}$$

Como se utiliza una distribución *Beta* para describir la heterogeneidad, el modelo se llama *Beta Geométrica desplazada.*

2.1.1.4. Distribución Beta.



Función de densidad

$$f\left(x|\alpha,\beta\right)=\frac{x^{\alpha-1}\left(1-x\right)^{\beta-1}}{\int_0^1 u^{\alpha-1}\left(1-u\right)^{\beta-1}du}=\frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}x^{\alpha-1}\left(1-x\right)^{\beta-1}$$

$$=\frac{1}{B\left(\alpha,\beta\right)}x^{\alpha-1}\left(1-x\right)^{\beta-1}$$

Media

$$\mu=\int_0^1 xf\left(x|\alpha,\beta\right)dx=\int_0^1 x\frac{x^{\alpha-1}\left(1-x\right)^{\beta-1}}{B\left(\alpha,\beta\right)}dx=\frac{\alpha}{\alpha+\beta}$$

Función Beta

$$B\left(a,\beta\right)=\int_0^1 t^{\alpha-1}\left(1-t\right)^{\beta-1}dt$$

$$B\left(\alpha,\beta\right)=B\left(\beta,\alpha\right)$$

$$B\left(\alpha,\beta\right)=\frac{\left(\alpha-1\right)!\left(\beta-1\right)!}{\left(\alpha+\beta-1\right)!}$$

$$B\left(\alpha,\beta\right)=\frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)}$$

2.1.1.5. Distribución Beta Geométrica Desplazada.

$$Pr\left(T = t|\alpha, \beta\right) = \int_0^1 Pr\left(T = t|\theta\right) Beta\left(\theta|\alpha, \beta\right) d\theta$$

$$Pr\left(T > t|\alpha, \beta\right) = \int_0^1 Pr\left(T > t|\theta\right) Beta\left(\theta|\alpha, \beta\right) d\theta$$

$$Beta\left(\theta|\alpha, \beta\right) = \frac{\theta^{a-1}\left(1-\theta\right)^{\beta-1}}{B(\alpha, \beta)} \quad B\left(\alpha, \beta\right) = \int_0^1 t^{\alpha-1}\left(1-t\right)^{\beta-1} dt$$

2.1.1.6. Distribución Gamma

$$g\left(\lambda\right) = \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma\left(r\right)}$$

$r > 0$: Parámetro de forma.
$\alpha > 0$:Parámetro de escala.

2.1.1.7. Distribución Gamma-Exponencial

2.2. Modelos Probabilísticos.

Estos modelos tratan de describir o predecir el comportamiento de los usuarios, usando variables observables, de esta forma se pretende que los modelos puedan describir el comportamiento de los usuarios y tomar decisiones en base a las proyecciones que el modelo muestre.

Dentro de sus usos se encuentra: *Predicciones*, *Comparaciones* (benchmarks), *Patrones*, entre otros.

Los modelos se pueden ver con datos agregados y des agregados, las principal ventaja de utiliza un modelo con datos desagregados, es la facilidad del uso de modelos estructurales. En cambio con datos agregados, se es más fácil la identificación de un patrón de comportamiento.

2.2.0.8. Metodología.

1. Determinar cuál es el problema o decisión de marketing que se resolver.
2. Identificar la información requerida para analizar dicho problema.
3. Verificar cuál es el comportamiento individual que se quiere analizar (debe ser observable).
4. Una vez verificado el comportamiento, seleccionar una distribución de probabilidad que se ajuste al comportamiento individual.
5. Luego de seleccionar una distribución que caracteriza a nivel individual, encontrar una que se ajuste a nivel agregado.
6. Derivar la distribución agregada, $f\left(x\right) = \int f\left(x|\theta\right) g\left(\theta\right) d\theta$
7. Estimar los parámetros ajustando los datos observados a el modelo propuesto.
8. Utilizar los resultados del modelo para resolver el problema de marketing inicial.

2.3. Modelo de duración en tiempo Discreto

Los modelos de duración, tal como su descripción dice, tratan de predecir el intervalo de tiempo (Puede ser medido en períodos) que una persona o cliente permanecerá dentro.

Es importante destacar que todos los modelos deben contener una componente discreta, esto no es dado por el comportamiento del fenómeno, el que si puede ser del tipo continuo, si no por que las muestras o evaluaciones del proceso no pueden ser continuas. De esta forma siempre se va a tener un intervalo de tiempo entre muestra y muestra, el que puede ser visto como un modelo discreto.

2.4. Modelos de duración en tiempo Continuo.

A pesar de lo descrito anteriormente en los modelos de duración en tiempo discreto, existen ciertos casos en donde es recomendable utilizar modelos de duración en tiempo continuo como:

1. Tiempo de respuesta a una campaña promocional de marketing directo.
2. Tiempo entre visitas a un sitio web.
3. Tiempos entre llamas recibidas por un call center.
4. Tiempo de operación para una industria de servicios.

2.4.1. Modelo Homogéneo.

Sea T el tiempo en se va a probar un producto, luego el tiempo T se puede decir que este tiempo posee una distribución, en particular se puede escoger la exponencial. Como la exponencial considera que a largo plazo todos adoptan, se debe considerar al utilizar esta distribución que una fracción de la población nunca va a probarlo.

$$F\left(t\right) = Pr\left(T \leq t\right) = 1 - e^{-\lambda t}$$

Luego como los datos estan en tiempo discreto, se ajusta la verosimilitud a los datos muestrales.

$$Pr\left(t_i \leq T \leq t_{i+1}\right) = F\left(t_{i+1}\right) - F\left(t_i\right)$$

Al agregar una fracción π que nunca probar el producto entonces.

$$F\left(t\right) = \left(1-\pi\right) Pr\left(T \leq t\right) + \pi Pr\left(\bigcirc\right) = \left(1-\pi\right)\left(1 - e^{-\lambda t}\right)$$

2.4.2. Modelo Heterogéneo.

Al modelo anteriormente descrito, se puede agregar heterogeneidad, considerando que la población distribuye de forma gamma. De esta forma la función de densidad esta dada por:

$$Pr\left(T \leq t\right) = \int_0^\infty Pr\left(T \leq t|\lambda\right) g\left(\lambda\right) d\lambda = \int_0^\infty \frac{\left(1 - e^{-\lambda t}\right)\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma\left(r\right)} d\lambda$$

$$Pr\left(T \leq t\right) = 1 - \left(\frac{\alpha}{\alpha + t}\right)^r$$

Este modelo es llamado Gamma-Exponencial, pero este modelo esta limitado a que el tiempo de adopción debe estar exponenecialmente distribuido, adicionalmente se sabe que la distribución exponencial caracteriza a un proceso sin memoria, lo que no siempre representa el comportamiento de las personas.

2.4.3. Modelo con Dependencia en la Duración.

2.4.4. Tasa de Riesgo.

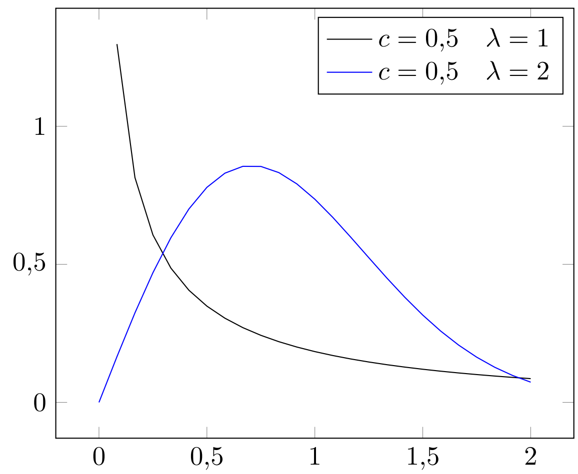
También conocida como "Hazard Ratio" se define como la probabilidad de que un evento ocurra ahora dado que hasta el momento no a ocurrido.

$$h\left(t\right) = \frac{f\left(t\right)}{1 - F\left(t\right)}$$

Es importante destacar que la tasa de riesgo se define como no negativa, dado que es un tiempo de ocurrencia. La metodología a utilizar es

2.4.5. Distribución Weibull

La distribución weibul es utilizada habitualmente para describir procesos de duración, como por ejemplo la esperanza de vida de las personas. Esta posee dos parámetros λ llamado factor de escala y c es el factor que define la forma de la distribución. Cuando $c = 1$ la distribución es equivalente a la exponencial.



Tasa de fallos de la distribución weibull:

$$h(t) = c\lambda t^{c-1}$$

Función de fiabilidad:

$$F(t) = e^{-\lambda t^c}$$

Luego la función de densidad de la distribución weibull.

$$f(x; \lambda, \kappa) = \begin{cases} c\lambda t^{c-1} e^{-\lambda t^c} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

La esperanza de la distribución se obtiene integrando $h(t)$

$$\mu = \int_0^\infty tc\lambda t^{c-1} e^{-\lambda t^c} dt$$

Luego haciendo un cambio de variable: $u = \lambda t^c$, $dt = \frac{du}{\lambda c t^{c-1}}$

$$\mu = \int_0^\infty \frac{cue^{-u}}{\frac{u}{t}} du = \int_0^\infty e^{-u} t du = \int_0^\infty \frac{u^{\frac{1}{c}}}{\lambda^{\frac{1}{c}}} e^{-u} du = \lambda^{-\frac{1}{c}} \int_0^\infty u^{\frac{1}{c}} e^{-u} du$$

Luego como la integral corresponde a una función gamma $\Gamma\left(1 + \frac{1}{c}\right)$

$$\mu = \lambda^{-\frac{1}{c}} \Gamma\left(1 + \frac{1}{c}\right)$$

2.4.6. Distribución Gamma-Weibull

Se puede introducir heterogeneidad al modelo weibull anteriormente descrito suponiendo que el parámetro λ distribuye en base una función $Gamma(\alpha, \beta)$.

$$Pr(T \leq t) = \int_0^\infty \frac{\left(1 - e^{-\lambda t^c}\right) \alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma(r)} d\lambda = 1 - \left(\frac{\alpha}{\alpha + t^c}\right)^r$$

2.5. Modelos de Conteo.

Los modelos de conteo son utilizados para describir la "cantidad de veces" que los consumidores realizarán una determinada acción. Ejemplos de esto: Cuantas veces visitarán una tienda, llamarán a un call center, usaran un producto o servicio, etc.

2.5.1. La Exposición Publicitaria.

Cuando se realizan campañas publicitarias, el gestor de la campaña esta interesado en estimar la efectividad de la campaña realiza, midiendo el alcance, la frecuencia y el rating.

Existen una medida para determinar la efectividad de una campaña llamada el *GRP (Gross Rating Points).*

$$GRP = \frac{\text{Cobertura de la campaña (\%)}}{\text{Frecuencia Media}}$$

Alcance Es la proporción de la población que fue expuesta a la campaña publicitaria al menos una vez durante la duración de la misma.

Frecuencia Promedio Es el número promedio de veces que la población que fue expuesta (los que la vieron) vieron la campaña.

Puntos de Rating Brutos (GRP) Es el número Promedio de exposiciones a una campaña por cada 100 personas.

2.5.1.1. Modelo de nivel individual

Para un modelo a nivel individual se puede considerar que la cantidad de veces que una persona es expuesta a una campaña publicitaria distribuye como una *Poissson* de tasa λ durante un tiempo de t de una semana.

$$Pr\left(X = x|\lambda\right) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Esta corresponde a una distribución Poisson-Exponencial. Ahora is consideramos un pe´riodo de tiempo de mayor duración (t semanas).

$$Pr\left(X\left(t\right) = x|\lambda\right) = \frac{\left(\lambda t\right)^x e^{-\lambda t}}{x!}$$

2.5.1.2. Modelo Heterogéneo.

El modelo anteriormente descrito considera homogeneidad en el parámetro λ. Se puede adicionar heterogeneidad al parámetro asumiendo que éste distribuye como una $Gamma\left(\alpha, r\right)$.

$$g\left(\lambda|\alpha, r\right) = \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma\left(r\right)}$$

$$Pr\left(X = x|r, \alpha\right) = \int_0^\infty Pr\left(X = x|\lambda\right) g\left(\lambda|\alpha, r\right) d\lambda$$

$$= \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma\left(r\right)} d\lambda = \frac{\alpha^r}{\Gamma\left(r\right) x!} \int_0^\infty \lambda^{r+x-1} e^{-\lambda(\alpha+1)} d\lambda$$

$$= \frac{\Gamma\left(r + r\right)}{\Gamma\left(r\right) x!} \left(\frac{\alpha}{\alpha + 1}\right)^r \left(\frac{1}{\alpha + 1}\right)^x$$

Este modelo también es conocido como Binomial Negativo (NBD).

La media y la varianza de la NBD son:

$$\mu = \frac{r}{\alpha} \quad \sigma^2 = \frac{r}{\alpha} + \frac{r}{\alpha^2}$$

Para la utilización de la NBD con una duración diferente de 1 el modelo queda:

$$Pr\left(X = x\right) = \frac{\Gamma\left(r + r\right)}{\Gamma\left(r\right) x!} \left(\frac{\alpha}{\alpha + t}\right)^r \left(\frac{1}{\alpha + t}\right)^x$$

Alcance $1 - Pr\left(X\left(t\right)\right) = 0$

Frecuencia $\frac{E(X(t))}{1-Pr(X(t)=0)}$

Rating $100 \cdot E\left(X\left(t\right)\right)$

Fórumula Recursiva NBD.

$$\frac{Pr\left(X = x\right)}{Pr\left(X = x-1\right)} = \frac{r+x-1}{x\left(\alpha+1\right)}$$

$$Pr\left(X = x\right) = \begin{cases} \left(\frac{\alpha}{\alpha+1}\right)^{r} & x = 0 \\ \frac{r+x-1}{x(\alpha+1)} Pr\left(X = x-1\right) & x \geq 1 \end{cases}$$

2.6. Modelos de Elección.

Los modelos de elección son utilizados para describir la "probabilidad de elección" que los consumidores poseen de realizar una determinada elección. Ejemplos de esto: Probabilidad de comprar al visitar una tienda, responder a una campaña de marketing directo, usar un producto o servicio en un período de tiempo, etc.

Para esto usando una base de datos, con descripciones de los clientes, se puede realizar una segmentación, con el objetivo de utilizar los recursos de marketing directo con los clientes que tengan una mayor tasa de respuesta. De esta forma, si consideramos que $X_s$es el número de respuestas del segmento s, asumiendo adicionalmente que X_s distribuye como una Binomial.

2.6.1. Modelo Homogéneo.

$$Pr\left(X_s = x_s | m_s, p_s\right) = \left(\begin{array}{c} m_s \\ x_s \end{array}\right) p_s^{x_s}\left(1-p_s\right)^{m_s - x_s}$$

$$E\left(X_s | m_s, p_s\right) = m_s p_s$$

2.6.2. Modelo Heterogéneo.

Para agregar heterogeneidad, se considera que p_s distribuye como una $Beta\left(\alpha, \beta\right)$.

$$g\left(p_s | \alpha, \beta\right) = \frac{p_s^{\alpha-1}\left(1-p_s\right)^{\beta-1}}{B\left(\alpha,\beta\right)} \quad \text{Con } \alpha > 0,\ \beta > 0,\ p_s \epsilon [0,1]$$

$$E\left(p_s | \alpha, \beta\right) = \frac{\alpha}{\alpha+\beta} \quad B\left(\alpha,\beta\right) = \frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)}$$

Luego la distribución de respuestas queda:

$$Pr\left(X_s = x_s\right) = \int_0^1 Pr\left(X_s = x_s | m_s, p_s\right) g\left(p_s | \alpha, \beta\right) dp_s$$

$$= \int_0^1 \left(\begin{array}{c} m_s \\ x_s \end{array}\right) p_s^{x_s}\left(1-p_s\right)^{m_s-x_s} \frac{p_s^{\alpha-1}\left(1-p_s\right)^{\beta-1}}{B\left(\alpha,\beta\right)} dp_s$$

$$= \left(\begin{array}{c} m_s \\ x_s \end{array}\right) \frac{B\left(\alpha + x_s, \beta + m_s - x_s\right)}{B\left(\alpha,\beta\right)}$$

Capítulo 3

Esperanzas condicionales

Las esperanzas condicionales son utilizadas para tomar decisiones desagregadas sobre los usuarios, de esta forma se pueden realizar promociones a ciertos segmentos de usuarios.

3.0.2.1. Teorema de Bayes

$$Pr(A_i|B) = \frac{Pr\left(B|A_i\right) Pr\left(A_i\right)}{Pr\left(B\right)}$$

3.1. Caso Distreto "Modelo de Conteo"

$$Pr\left(Segmento_1|N_x = x\right) = \frac{Pr\left(Segmento_1 \cap N_x = x\right)}{Pr\left(N_x = x\right)}$$

$$Pr\left(Segmento_2|N_x = x\right) = 1 - Pr\left(Segmento_1|N_x = x\right)$$

3.2. Caso Continuo "Modelo de Elección"

Suponiendo que todos los miembros de un mismo segmento s tienen la misma probabilidad de responder p_s y tamaño m_s, el número de respuestas en el segmento x_s tiene distribución binomial. La heterogeneidad del parámetro λ va a estar distribuida $Beta\left(\alpha, \beta\right)$.

$$\begin{aligned} Pr\left(X_s = x_s\right) &= \int_0^1 Pr\left(X_s = x_s|m_s, p_s\right) g\left(p_s|\alpha, \beta\right) dp_s \\ &= \begin{pmatrix} m_s \\ x_s \end{pmatrix} \frac{B\left(\alpha + x_s, \beta + m_s - x_s\right)}{B\left(\alpha, \beta\right)} \end{aligned}$$

3.2.1. Estimación Número de respuestas θ_s

$g\left(\theta|x\right)$ Es la distribución posterior, dada la información sobre el segmento s conocida $\left(x_s\right)$

$$g\left(\theta|x\right) = \frac{Pr\left(X_s = x_s|\theta_s, m_s\right) g\left(\theta_s\right)}{\int_0^1 Pr\left(X_s = x_s|\theta_s, m_s\right) g\left(\theta_s\right) d\theta_s}$$

$$= \frac{1}{B\left(\alpha + x_s, \beta + m_s - x_s\right)} \theta_s^{\alpha + x_s - 1} \left(1 - \theta_s\right)^{\beta + m_s - x_s - 1} = Beta\left(\alpha + x_s, \beta + m_s - x_s\right)$$

$$E\left(\theta_s|X_s = x_s, m_s\right) = \frac{\alpha + x_s}{\alpha + \beta + m_s}$$

3.3. Variables Explicativas

La idea principal es agregar datos que puedan explicar el comportamiento observado, como estación del año, lugar geográfico, temperatura, etc.

Para eso se utilizaron múltiples distribuciones, con el objetivo de "buscar" cuál es la que más se ajustaba a la data. Las distribuciones testeadas fueron: Poisson, Binomial Negativa y Gamma.

Con el fin de evaluar si la variable explicativa es significativa, se deben calcular los errores estándares de los diferentes modelos y comparar. Si se hizo el ajuste de cada modelo maximizando la verosimilitud, entonces se puede estimar los errores estándar a partir del Hessiano.

$$\sqrt{n}\left(\theta_{MaximaVerosimilitud} - \theta_0\right) \rightsquigarrow N\left(0, -H^{-1}\left(\theta_0\right)\right)$$

Capítulo 4

Evaluación de Modelos

¿Qué se le debe pedir a un Modelo?

Cualitativamente Que sea capaz de describir el comportamiento de los clientes en algún aspecto.

Cuantitavamente Que el modelo se ajuste a los datos observados (calibración) y que sea capaz de predecir en alguna medida u horizonte de tiempo futuro.

4.1. Métricas típicas de Ajuste

R^2 Coeficiente de determinación, nos indica cuanta variabilidad de la variable dependiente es explicada por el modelo.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \overline{y_i})^2}$$

MAE Mean absolute error, mide cuánto se equivoca el modelo.

$$MAE = \frac{1}{n} \sum |y_i - \overline{y_i}|$$

MAPE Mean absolute percentage error, mide cuánto se equivoca porcentualmente el modelo.

$$MAPE = \frac{1}{n} \sum \left| \frac{y_i - \overline{y_i}}{y_i} \right|$$

4.2. Test Bondad de Ajuste

Uno de los test para probar la bondad del ajuste es el test χ^2 de Pearson.

H_0 El modelo describe bien la data observada.

H_1 El modelo no describe bien la data observada.

Capítulo 5

Modelos Integrados

Cuándo? Modelo de Tiempo de Ocurrencia.

Cuántos? Modelo de Conteo.

Cuál? Modelo de Elección.

Esto modelos son sencillos y permiten integrar heterogeneidad en los comportamientos. Estos modelos pueden ser combinados entre sí. El problema de la mezcla de modelos es el aumento en la complejidad algebraica y de implementación.

5.0.1. Modelamiento de Conteo de Poisson no Reportados

EL problema en los modelos de Conteo es que al encuestar a los consumidores, estos premeditadamente reporten más o menos, dependiendiendo de la connotación de la encuesta aplicada.

5.0.1.1. Función Hypergeometrica Gaussiana

$$F(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)\, z^j}{\Gamma(c+j)\, j!}$$

$$F(a,b,c,z) = \sum_{j=0}^{\infty} u_j \approx \sum_{j=0}^{M} u_j$$

$$\text{Donde} u_0 = 1 \qquad \frac{u_j}{u_{j-1}} = \frac{(a+j-1)(b+j-1)}{(c+j-1)\, j} z \qquad j = 1,2,3,...$$

Las aplicaciones de este tipo de modelos tiene que ver con la búsqueda que un parámetro como visitas a una tienda se puede desprender de las compras realizadas por una persona.

Capítulo 6

Modelos Estructurales

Permiten probar y estimar relaciones causales, a partir de los datos secundarios. Este mezcla el comportamiento de los consumidores y que estos son agentes racionales. También existen modelos de *forma reducida* que poseen un enfoque estadístico (Foco en ajustar los datos al modelo). Es importante destacar que finalmente se debe llegar a un modelo que equilibrie lo *estructural* y su *forma reducida*.

6.1. Crítica de Lucas

Robert Lucas, critica el intento por predecir el comportamiento de los usuarios en base a datos agregados. Estos muchas veces no logran predecir el comportamiento de los usuarios si es que alguno de los factores que define el comportamiento actual es cambiado *¿ Al subir el precio las personas van a seguir consumiendo al mismo régimen ?*.

6.2. Modelo de Elección Continua

LES Linear Expenditure System

Suponiendo que x_{kt}como la cantidad a comprar en la categoría k, en la oportunidad de compra t, con un presupuesto μ_t. Pero este modelo posee muchos parámetros y no es consistente entre la compra de los productos y el uso del presupuesto.

$$f\left(x_{kt}\right)=\sum_{j\neq k}\alpha_j f_j\left(x_{jt}\right)+\sum_{j\neq k}\beta_j g_j\left(p_{jt}\right)+\gamma h\left(\mu_t\right)$$

Asumiendo que las demandas observadas son el producto de la racionalidad de los consumidores, entonces estos maximizan la utilidad, restringido a el presupuesto disponible.

$$\begin{array}{ll} Max & \left\{u\left(x_t\right)=\sum_k \alpha_k Ln\left(x_{kt}-\beta_k\right)\right\} \\ s.a & \sum_k p_{kt}x_{kt}\leq \mu_t \end{array}$$

La solución es de la forma: $x_{kt}=\beta_k-\sum_j \alpha_k\beta_j\frac{p_{jt}}{p_{kt}}+\alpha_k\frac{\mu_t}{p_{kt}}$ que puede ser vista como:

$$x_{kt}=\lambda_0-\sum_j \lambda_j\frac{p_{jt}}{p_{kt}}+\lambda_{k+1}\frac{\mu_t}{p_{kt}}$$

Considerando la componente estocástica como aditiva: $x_{kt}=\lambda_0-\sum_j\lambda_j\frac{p_{jt}}{p_{kt}}+\lambda_{k+1}\frac{\mu_t}{p_{kt}}+\epsilon_{kt}$
La verosimilitud del modelo se escribe:

$$LL_{kt}=Log\left(\phi\frac{x_{kt}-\left(\lambda_0-\sum_j\lambda_j\frac{p_{jt}}{p_{kt}}+\lambda_{k+1}\frac{\mu_t}{p_{kt}}\right)}{\sigma}\right)$$

Capítulo 7

Logit

La regresión logística, busca predecir el resultado de una o más variables categóricas en función de las variables independientes. Aplicando esto en marketing, se intenta predecir las elecciones de los usuarios en aspectos como marca, color, tamaño, compra o no un producto, etc.

7.0.1. Función Logit

$$Logit\left(p\right) = Log\left(\frac{p}{1-p}\right) = Log\left(p\right) - Log\left(1-p\right)$$

7.1. Maximización de Utilidades

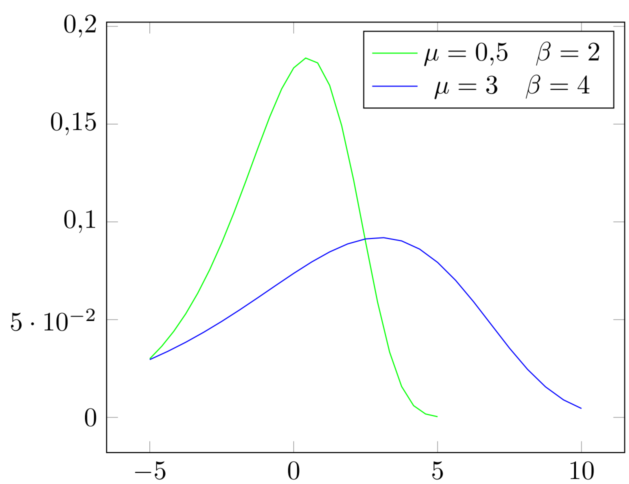
Sea P_{ni} la probabilidad de que un consumidor n elija al alternativa i, maximizando su utilidad u_{ni}. Estas utilidades no son conocidas cuando se realiza el análisis, pero si son conocidas por los usuarios, es por esto que se agrega la variable determinista v_{ni} y ϵ_{ni} estocástica.

$$P_{ni} = Pr\left(u_{ni} > u_{nj}, \forall j \neq i\right) = Pr\left(v_{ni} + \epsilon_{ni} > v_{nj} + \epsilon_{nj}, \forall j \neq i\right)$$

$$= \int_{\epsilon_n} 1_{\left(\epsilon_{nj} - \epsilon_{ni} < v_{ni} - v_{nj}, \forall j \neq i\right)} f\left(\epsilon_n\right) d\epsilon_n$$

Para proseguir con el análisis se debe conocer la utilidad de los consumidores, las que pueden ser estimadas al observar cambios en su comportamiento de consumo v/s cambios en el precio de los productos. Recordar que estas utilidades son relativas al usuario y típicamente necesitan que se normalice alguna de sus componentes en el modelo.

7.2. Distribución Valor Extremo (Gumbel)



$$f(x,\mu,\beta)=\frac{e^{\left(\frac{x-\mu}{\beta}\right)}e^{\left(-e^{\left(\frac{x-\mu}{\beta}\right)}\right)}}{\beta}$$

$$F(x,\mu,\beta)=1-e^{\left(-e^{\left(\frac{x-\mu}{\beta}\right)}\right)}$$

Media: $\mu+\gamma\beta$
Desviación: $\frac{\beta\pi}{\sqrt{6}}$

Con γ la constante de Euler-Mascheroni.

7.3. Modelo Logit

Este resulta de la asumción de que cada ϵ_{ni} se distribuye valor extremo.

$$\begin{aligned}P_{ni} &= Pr\left(e_{ni}<\epsilon_{ni}+v_{ni}-v_{nj},\forall j\neq i\right)\\ &= \int_{\epsilon_{ni}}\left(\prod_{j\neq i}e^{\left(-e^{-(\epsilon_{ni}+v_{ni}-v_{nj})}\right)}\right)e^{-\epsilon_{ni}}e^{\left(-e^{(-\epsilon_{ni})}\right)}d\epsilon_{ni}\\ &= \frac{e^{v_{ni}}}{\sum_j e^{v_{nj}}}\end{aligned}$$

7.3.1. Propiedades

1. Puede explicar variaciones sistemáticas pero no aleatorias.
2. Puede capturar comportamiento con los datos medidos, pero no se puede predecir con los datos no observados. Esto se debe a la asumción de independencia en en la componente aleatoria de la utilidad.
3. Logit retorna valores de substitución proporcionales entre los componentes medidos.

7.3.2. Patrones de substitución

Ratio de Independencia: Es aplicado sobre las alternativas irrelevantes $\frac{P_{ni}}{P_{nj}}=e^{v_{ni}-v_{nj}}$, (escojo iphone negro o blanco)

Substitución Proporcional: Al variar el valor de la probailidad de adopción en un $x\,\%$ el resto de las alternativas decrecerán un $\beta x\,\%$ $E_{(i,Z_{nj})}=-\beta_Z Z_{nj}P_{nj}$

7.3.3. Verosimilitud

Considerando un caso lineal $v_{ni}=\beta_z' Z_{ni}$.

$$\begin{aligned}LL(\beta) &= \sum_n\sum_i y_{ni}Log\left(P_{ni}\right)\\ &= \sum_n\sum_i y_{ni}Log\left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}\right)\\ &= \sum_n\sum_i y_{ni}\left(e^{\beta' x_{ni}}\right)-\sum_n\sum_i y_{ni}Log\left(\sum_j e^{\beta' x_{nj}}\right)\\ \frac{dLL}{d\beta} &= \sum_n\sum_i\left(y_{ni}-P_{ni}\right)x_{ni}\end{aligned}$$

7.3.4. Bondad de Ajuste

Para un modelo con k parámetros estimados sobre n observaciones.

1. Ratios de Verosimilitud:

$$\rho=1-\frac{LL\left(\hat{\beta}\right)}{LL(0)}$$

2. Criterio de Información Akaike (AIC):

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

3. Criterio de Información Bayesiano (BIC):

$$BIC = -2LL\left(\hat{\beta}\right) + kLog\left(n\right)$$

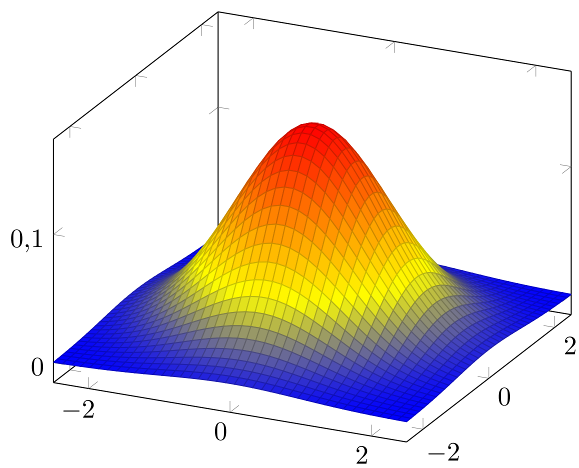
Capítulo 8

Probit

Este modelo nace al asumir que el vector de los errores ϵ_{ni} distribuye normal multivariada centrada en cero. Tiene la capacidad de acomodar variaciones sistemáticas de los consumidores en factores *no* observables, pero no tiene la propiedad de *"substitución proporcional"* es decir que las alternativas irrelevantes sean independientes, tampoco impone restricciones en dichos patrones.

$$\phi\left(\epsilon_n\right) = \frac{1}{\left(2\pi\right)^{I/2}\left|\sum\right|^{1/2}} e^{\left(-\frac{1}{2}\epsilon_n'\,\Sigma^{-1}\,\epsilon_n\right)}$$

8.1. Distribución Multivariada



8.1.1. Caso de dos dimensiones.

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-\frac{2\rho xy}{(\sigma_x\sigma_y)}\right)\right)$$

Donde ρ es el coeficiente de correlación entre x e y.

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

8.1.2. Caso general

Donde $|\sum|$ es el determinante de la matriz $\sum$ que es semidefinida positiva (matriz varianza covarianza), x es el vector compuesto por los x_i y μes el vector con los μ_i.

$$f_X(x_1,\ldots,x_n) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top\Sigma^{-1}(x-\mu)\right)$$

La función característica de x es:

$$\phi_x\left(u;\mu,\Sigma\right)=\exp\left(i\mu^{\top}\mu-\frac{1}{2}\mu^{\top}\Sigma\mu\right)$$

8.2. Identificación

Para ajustar el modelo Provit se deben realizar ciertas imposiciones, tales como que el valor "numérico" de las utilidades no es relevante, si no que la diferencia de utilidad.

$$\hat{\sum}=Cov\begin{pmatrix}\epsilon_{n2}-\epsilon_{n1}\\ \epsilon_{n3}-\epsilon_{n1}\\ \epsilon_{n4}-\epsilon_{n1}\end{pmatrix}=\begin{pmatrix}\hat{\theta_{22}} & \hat{\theta_{23}} & \hat{\theta_{24}}\\ & \hat{\theta_{33}} & \hat{\theta_{34}}\\ & & \hat{\theta_{44}}\end{pmatrix}$$

Luego fijando la escala, es decir $\hat{\theta_{22}}=1$

$$\tilde{\sum}=\begin{pmatrix}1 & \tilde{\theta_{23}} & \tilde{\theta_{24}}\\ & \tilde{\theta_{33}} & \tilde{\theta_{34}}\\ & & \tilde{\theta_{44}}\end{pmatrix}$$

En esta última matriz, típicamente se impone una estructura que agrupe las componentes por sus características, ejemplo marcas normales y premmium.

8.3. Dependencia en el tiempo

Es esencialmente utilizado para analizar datos de panel en dónde sus elecciones pueden estar correlacionadas en el tiempo.

$$u_{ni}=v_{ni}+\epsilon_{ni}\qquad\left(\begin{bmatrix}\epsilon_{n11}\\ \epsilon_{n21}\\ \vdots\\ \epsilon_{nI1}\end{bmatrix},\begin{bmatrix}\epsilon_{n12}\\ \epsilon_{n22}\\ \vdots\\ \epsilon_{nI2}\end{bmatrix},\cdots,\begin{bmatrix}\epsilon_{n1T}\\ \epsilon_{n2T}\\ \vdots\\ \epsilon_{nIT}\end{bmatrix}\right)\rightsquigarrow N\left(0,\sum\right)$$

8.4. Estimación

Como P_{ni} no tiene forma cerrada para este modelo, no se puede calcular la $LL\left(\right)$, es por esto que se debe aproximar numéricamente.

$$P_{ni}\approx\frac{1}{R}\sum_{r=1}^{r}1_{\left[\epsilon_{nj}^{r}-\epsilon_{ni}^{r}<v_{ni}-v_{nj},\forall j\neq i\right]}\qquad\left\{\epsilon_{n}^{r}\right\}_{r=1}^{R}\,iid,\ N\left(0,\sum\right)$$

De esta forma la *"verosimilitud simulada"* es de la forma.

$$\hat{\theta}=argMax\left\{\sum_{n}\sum_{i}y_{ni}Log\left(\hat{P}_{ni}\right)\right\}$$

Éste método realiza algo así como una simulación de Montecarlo, en dónde cada vez se generan R valores normales multivariados y se calcula finalmente el número de veces que se escogió cada alternativa.

Capítulo 9

Tobit

El modelo Tobit relaciona una variable independiente y_i^* también llamada variable latente, que no es observable, y un vector independiente x_i. β es un factor que transforma las variables independientes en la variable latente que no es observada.

$$y_i = \begin{cases} y_i^* & y_i^* > 0 \\ 0 & y_i^* \leq 0 \end{cases} \qquad y_i^* = \beta x_i + \epsilon_i \qquad \text{Con: } \epsilon_i \sim N(0, \sigma^2)$$

9.1. Consistencia

Para que el estimador β sea consistente debe ser calculado mediante el método de la máxima verosimilitud, si este es calculado utilizando mínimos cuadrados, este es inconsistente.

El ponderador β no corresponde a el efecto que x_i tiene sobre y_i^*, como podría deducirse al utilizar una estimación lineal, si no es la combinación de la variación se encuentra por encima de un cierto límite, ponderado por la probabilidad de estar por encima de ese límite y el cambio en la probabilidad de estar por encima del límite, ponderado por la esperanza de y_i^*dado que se sabe que tiene un valor superior al límite.

9.2. Tobit tipo 1

$$y_i = \begin{cases} y_i^* & y_i^* > 0 \\ 0 & y_i^* \leq 0 \end{cases} \qquad y_i^* = \beta x_i + \epsilon_i \qquad \text{Con: } \epsilon_i \sim N(0, \sigma^2)$$

9.3. Tobit tipo 2

$$y_i = \begin{cases} y_i^* & w_i^* > 0 \\ 0 & w_i^* \leq 0 \end{cases} \qquad \begin{matrix} y_i^* = \beta x_i + \epsilon_i \\ w_i^* = \gamma Z_i + e_i \end{matrix} \qquad \begin{pmatrix} \epsilon_i \\ e_i \end{pmatrix} \rightsquigarrow N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right)$$

9.4. Verosimilitud

La contribución depende si el dato es observado (medido) o no. Como $\epsilon_i \rightsquigarrow N\left(0, \sigma^2\right)$:

$$L_i = \begin{cases} f\left(y_i^*\right) & y_i > 0 \\ Pr\left(y_i^* \leq 0\right) & y_i = 0 \end{cases} \quad \rightarrow \quad L_i = \begin{cases} \frac{1}{\sigma}\phi\left(\frac{y_i - \beta x_i}{\sigma}\right) & y_i > 0 \\ 1 - \Phi\left(\frac{\beta x_i}{\sigma}\right) & y_i = 0 \end{cases}$$

9.4.1. Tobit tipo 1

$$L = \prod_{y_i=0} \left(1 - \Phi\left(\frac{\beta x_i}{\sigma}\right)\right) \prod_{y_i>0} \phi\left(\frac{y_i - \beta x_i}{\sigma}\right) \frac{1}{\sigma}$$

9.4.2. Tobit tipo 2

$$L = \prod_{y_i=0} \left(1 - \Phi\left(\frac{\gamma Z_i}{\sigma_2}\right)\right) \prod_{y_i>0} \phi\left(\frac{y_i \beta x_i}{\sigma_1}\right) \frac{1}{\sigma} \Phi\left(\left[\gamma Z_i + \frac{\sigma_{12}}{\sigma_1^2}\left(y_i - \beta x_i\right)\right] \sqrt{\sigma_2^2 + \frac{\sigma_{12}^2}{\sigma_1^2}}\right)$$

Capítulo 10

Teorías de Comportamiento

10.1. Lealtad en la elección de Marca

EL principio es que las personas suelen comprar productos de marcas que ya han comprado antes.

$$u_{nkt} = \beta_0 + \beta_1 P_{nkt} + \beta_2 x_{nkt} \qquad x_{nkt} = \lambda x_{nkt-1} + y_{nkt-1}(1-\lambda)$$

Dónde:

- n es el identificador de cliente.
- k es el identificador de la opción o marca que compró.
- t es la oportunidad de compra.
- $Y_{nkt} = 1$ si el cliente n elige la alternativa k en la oportunidad t de compra.
- P_{nkt} es el precio de la alternativa.
- x_{nkt} lealtad de la alternativa.

10.2. Efectos de Referencia de Precios y Promociones

Los consumidores poseen una percepción de el valor de las cosas y con ello su disposición a pagar. Esta percepción puede estar dada por el recuerdo de el valor en visitas pasadas al supermercado *"precio de referencia"*. Si no se usan precios anteriores (se desconoce) se compara mediante los precios de otras marcas o los anuncios en la publicidad.

$$u_{nkt} = \beta_0 + \beta_1 P_{nkt} + \beta_2 (P_{nkt} RP_{nkt}) \qquad RP_{nkt} = \lambda RP_{nkt} + P_{nkt-1}(1-\lambda)$$

10.3. Teoría de Prospectos

Los consumidores basan sus elecciones en el precio de referencia, más una componente personal de aversión al riesgo. La función de valor es cóncava y convexa para las perdidas generadas por el riesgo tomado. Las pendientes son diferentes, compensando negativamente las pérdidas.

$$u = \beta_0 + \beta_1 (PG + \lambda PL)$$

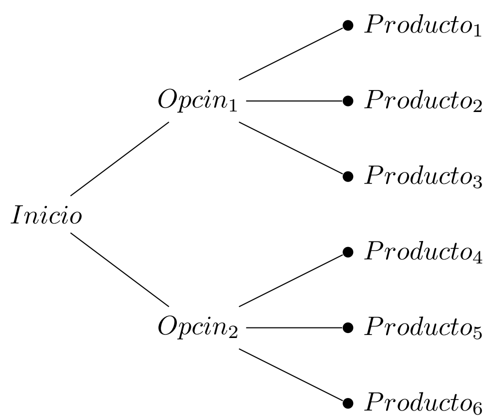
Dónde:

- PG : es la diferencia entre el precio de referencia el observado (diferencia positiva)
- PL : Es la diferencia entre el precio de referencia y el observado (diferencia negativa)

Capítulo 11

Decisiones Anidadas

Considera que los usuarios realizan múltiples decisiones y de forma en serie para llegar a la compra un producto final. La utilización de un *"Modelo Logit Simple"* no se ajusta a que los usuarios toman las decisiones de forma anidadas. De esta forma todas las opciones final poseen un mismo peso y son evaluadas de forma simultanea. En cambio un *"Modelo Logit Anidado"* si ajusta a el echo que las decisiones entre categorías no son comparables.



Luego de esto las probabilidades de que un usuario escoja cada uno de los productos posee la siguiente forma:

$$Pr\left(\text{Compre } Producto_1 | Opción_1\right) = \frac{\exp\left(Utilidad_1\right)}{\exp\left(Utilidad_1\right) + \exp\left(Utilidad_2\right) + \exp\left(Utilidad_3\right)}$$

A su vez la probabilidades de escoger la *Opción* o $Opción_1$esta dada por:

$$Pr\left(Opción_1\right) = \frac{\exp\left(Utilidad_{Opción1}\right)}{\exp\left(Utilidad_{Opción1}\right) + \exp\left(Utilidad_{Opción2}\right)}$$

Sin embargo dado que cada una de las opciones son diferentes, por son siguiente poseen diferentes factores que definen su utilidad. Si se asume racionalidad de los consumidores, entonces la utilidad de cada opción debería ser el valor esperado de elegir dicha opción.

$$I_{Opción1} = E\left[\text{máx}\left\{Utilidad_1 + \epsilon_1; Utilidad_2 + \epsilon_2; Utilidad_3 + \epsilon_3\right\}\right]$$

ϵ distribuyen valor extremo.

$$I_{Opción1} = \ln\left(\exp\left(Utilidad_1\right) + \exp\left(Utilidad_2\right) + \exp\left(Utilidad_3\right)\right)$$

$$Pr\left(Opción_1\right) = \frac{\exp\left(I_{Opción1}\right)}{\exp\left(I_{Opción1}\right) + \exp\left(I_{Opción2}\right)}$$

Capítulo 12

Heterogeneidad en modelos Logit

Heterogeidad Observable La sensibilidad de los consumidores con mayores ingresos es menor al precio, los clientes que compran una marca, es más probable que la siguiente vez la vuelvan a comprar.

Logit Mezcla finita: Utiliza clase latente.

Logit Mezcla Continua: Mixed Logit.

12.1. Mezcla Finita

12.1.1. Verosimilitud

El número de segmentos debe ser estimado mediante los estimadores AIC y BIC por ejemplo.

$$LL(\beta_m, s_m) = \sum_n \sum_i y_{ni} \ln \left(\sum_m s_m P_{nim}(\beta_m) \right)$$

$$s_m = \frac{\exp(\lambda_m)}{\sum_m \exp(\lambda_m)} \qquad \lambda_m = 0$$

12.1.2. Probabilidad de que un cliente pertenezca a un segmento:

$$Pr(n \in m) = \frac{L_m(y_n|x_n, \beta_m) \cdot s_m}{\sum_k L_k(y_n|x_n, \beta_k) \cdot s_k}$$

12.1.3. Elasticidad

$$E_{i,x_i} = \frac{\delta P_{mi}}{\delta x_i} \frac{x_i}{P_{mi}} = \beta_m (1 - P_{mi}) x_i$$

$$E_{i,x_j} = \frac{\delta P_{mi}}{\delta x_j} \frac{x_j}{P_{mi}} = -\beta_m P_{mj} x_j$$

Las elasticidades sirven para hacer análisis de influencia competitiva (IC) y de vulnerabilidad.

12.2. Mezcla Continua

Los agentes maximizan su utilidad. $u_{ni} = \alpha x_{ni} + \epsilon_{ni}$

12.2.1. Elasticidad

Cambio en la participación de la marca i, si es que cambia un atributo m de la alternativa j.

$$E_{i,x_j^m} = -\frac{x_j^m}{P_{nj}} \int \beta^m P_{ni}(\beta) P_{nj}(\beta) f(\beta) \, d\beta$$

Dónde:

$$P_{ni}(\beta) = \frac{\exp\left(v_{ni(\beta)}\right)}{\sum_k \exp\left(v_{nk(\beta)}\right)} \qquad P_{ni} = \int \left(\frac{\exp\left(v_{ni(\beta)}\right)}{\sum_k \exp\left(v_{nk(\beta)}\right)} \right) f(\beta) \, d\beta$$

En la mayoría de los casos no hay una fórmula cerrada para la probabilidad de elección P_{ni} de esta forma se utilizan herramientas como el método de la máxima verosimilitud simulada.

Capítulo 13

Enfoque Bayesiano

$$Pr\left(A|B\right) = \frac{Pr\left(A \cap B\right)}{Pr\left(B\right)} = \frac{Pr\left(B|A\right) Pr\left(A\right)}{Pr\left(B\right)}$$

Prior $Pr\left(B\right)$

Verosimilitud $Pr\left(B|A\right)$

Posterior $Pr\left(A|B\right)$

Se busca la posterior. La distribución posterior de los parámetros es proporcional a la verosimilitud y una distribución prior. Para casos conjugados, la posterior tiene fórmula analítica, en otros se debe realizar samplers desde la posterior.

Una distribución a priori es conjugada a una verosimilitud si la posterior esta en la misma familia de distribuciones que el prior.

Beneficios de la Inferencia Bayesiana:

- No recurre a argumentos asintóticos para inferir (teorema central del límite).
- Integra de forma coherente múltiples fuentes de información.
- Los métodos de simulación entregan flexibilidad para estimar modelos complejos.

Costos:

- Computacional
- Se debe definir una distribución a prior.

13.1. Montecarlo Markov Chains

Explotan una estructura particular del problema, para luego formular una cadena de Markov con la distribución estacionaria que coincida con la posterior. Como el método funciona en base a iteraciones, los primeros resultados son de baja calidad, por consiguiente son descartados.

13.1.1. Gibbs

Asume que las secuencias de distribuciones condicionales son conocidas, de esta forma su descomposición en ocasiones no es sencilla.

13.1.2. Metropolis

Es conocida como random walk *"Metropolis-Hastings"*, se propone un candidato y luego se acepta con una probabilidad, de las opciones disponibles, se vuelve a iterar con la que posea un mayor probabilidad.

13.2. Modelo Jerárquico Bayesiano

El objetivo de estos modelos es el mantener la individualidad de los consumidores, adicionalmente se usa información agregada de todos los consumidores para predecir el comportamiento de clientes que se observan parcialmente.

Capítulo 14

Sistema de Apoyo a Las Decisiones de Marketing

1. Escaneo, codificación limpieza e interpretación de los datos. Para recolectar la información necesaria para el estudio.
2. DATA: Administrar la base de datos
3. INFORMACIÓN: Modelos de desicion y mentales.
4. APRENDIZAJE: Juicios sobre los modelos utilizados.
5. DECISIONES: Recursos múltiples, como presupuesto, humanos, de datos, etc.
6. IMPLEMENTACIÓN.