



**TOWERS
PERRIN**

TILLINGHAST

Contribuciones recientes de la Estadística a la Actuaría en México

Tarifa y Segmentación de Riesgos usando Modelos Predictivos

Roberto Ambríz
Eduardo Esteva
Georgina Montiel

Noviembre 23, 2007

This document is incomplete without the accompanying discussion; it is confidential and intended solely for the information and benefit of the immediate recipient hereof.

Agenda

- Introducción
- ¿Qué son los modelos predictivos?
- Ciclo de la Tarifa
 - Estructura de la tarifa
 - Análisis Multivariado
 - Modelo de índice de siniestralidad
 - Tarifa teórica
 - Simulación de la tarifa
 - Prueba de mercado

Introducción

- La experiencia de siniestros para la mayoría de las líneas de negocio generalmente es analizada utilizando un análisis tradicional de frecuencia y severidad.
- Los modelos predictivos (GLM's, árboles de decisiones, redes neuronales, ...) otorgan mejoras significativas dentro del análisis.
- Las compañías han utilizado cada vez más y más análisis estadísticos tanto en el manejo de su información como en la toma de decisiones y elaboración y diseño de tarifas.
- El uso de esta herramienta no es común en todos los mercados, debido entre otras cosas a:
 - Falta o insuficiencia de información.
 - Pobre calidad de la información.
 - Falta de métodos y herramientas actuariales.

¿Qué son los modelos predictivos?

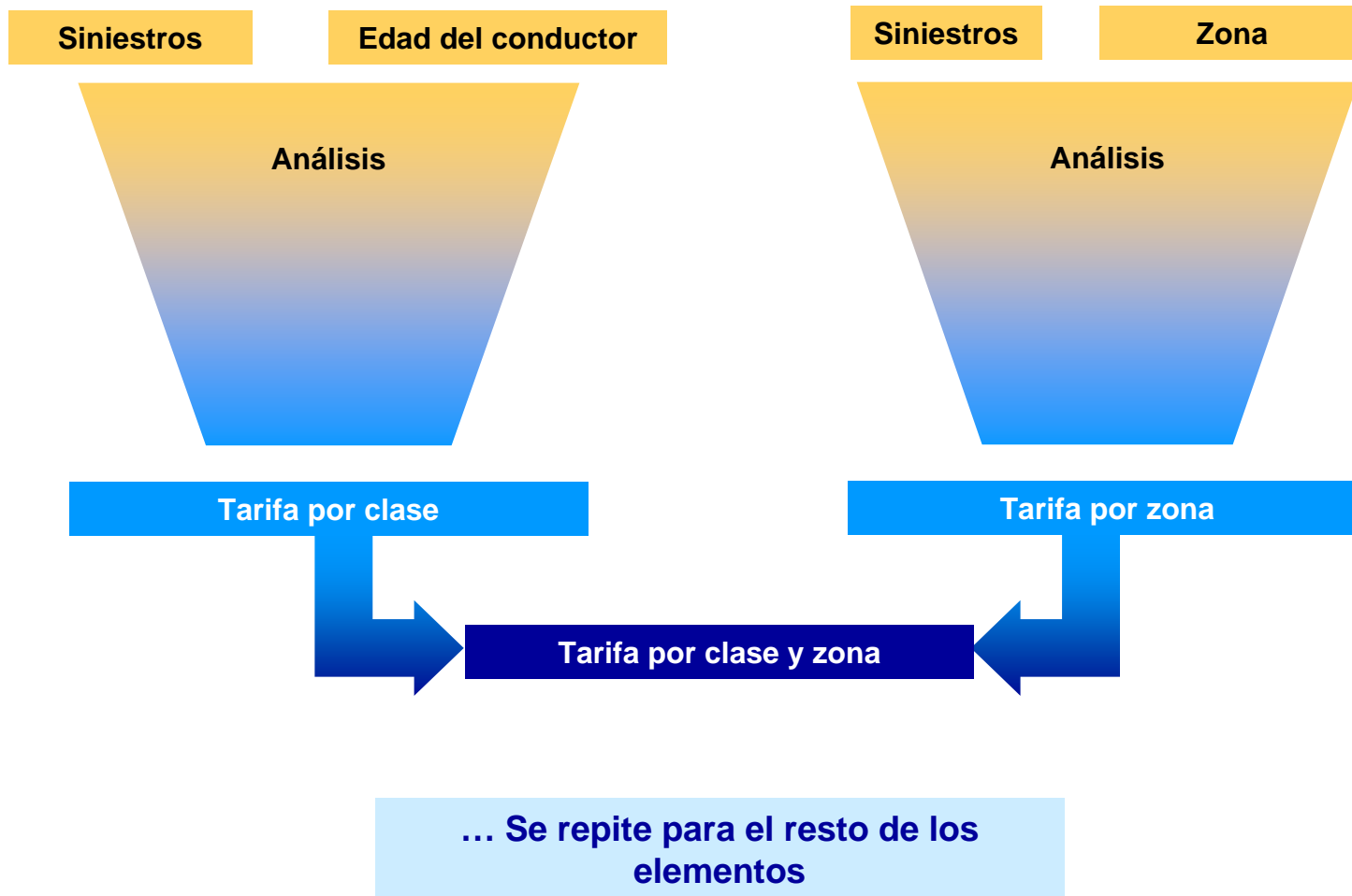
- Es el uso de técnicas estadísticas y computacionales que permiten:
 - Extraer cierta información de grandes bases de datos.
 - Identificar patrones, relaciones y correlaciones entre diversas variables.
- La definición estadística dice que un modelo predictivo predice el valor esperado de cierto análisis basado en el uso de variables simultáneas.

El objetivo de un modelo predictivo es ayudar a una organización a tomar mejores decisiones de negocio.

- Por qué utilizarlos
 - Exista información
 - Existen herramientas
 - Proporcionan información que se encuentra alineada con la estrategia y dirección de la compañía

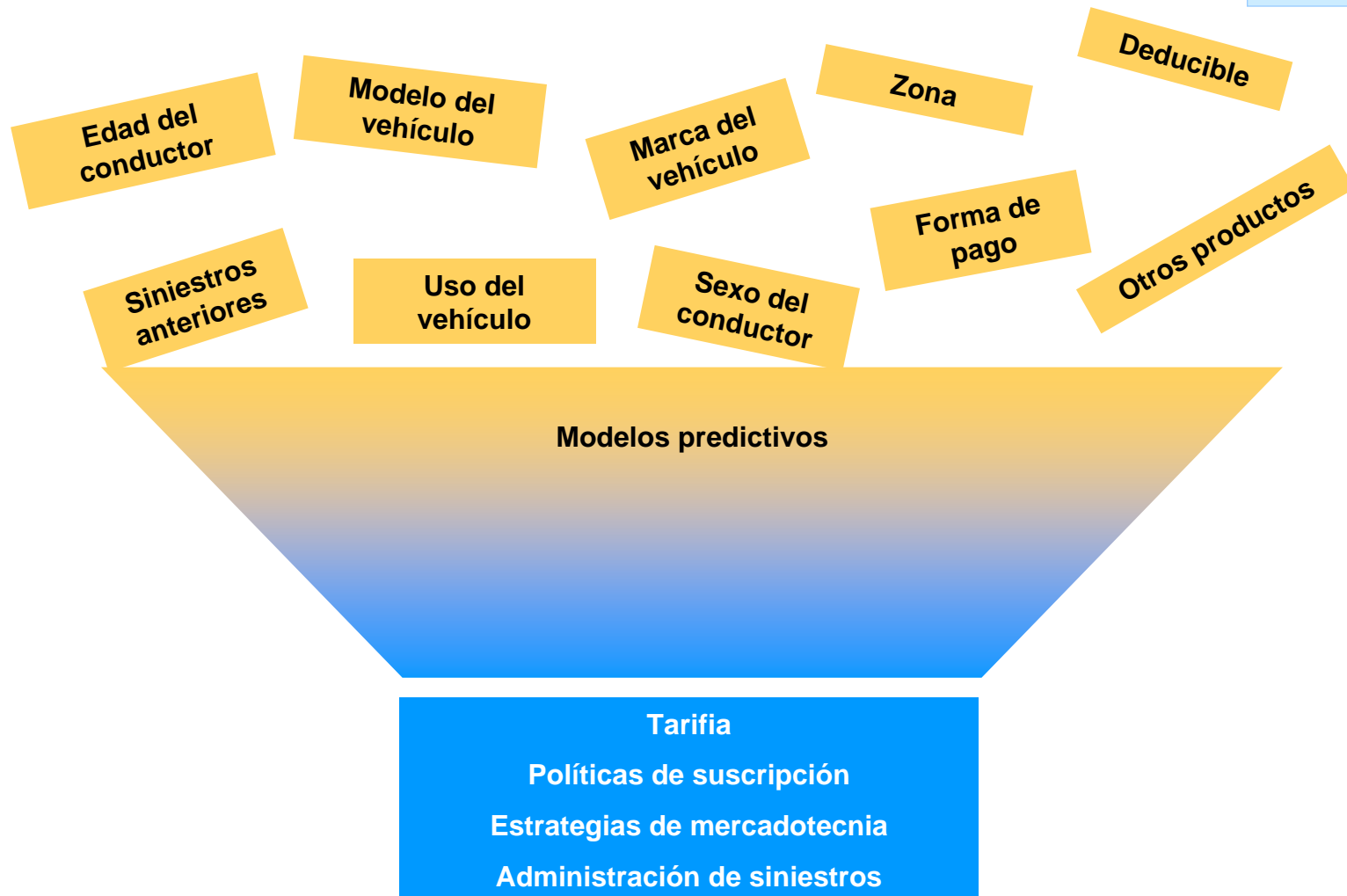
Métodos tradicionales

ILUSTRATIVO



Modelos predictivos

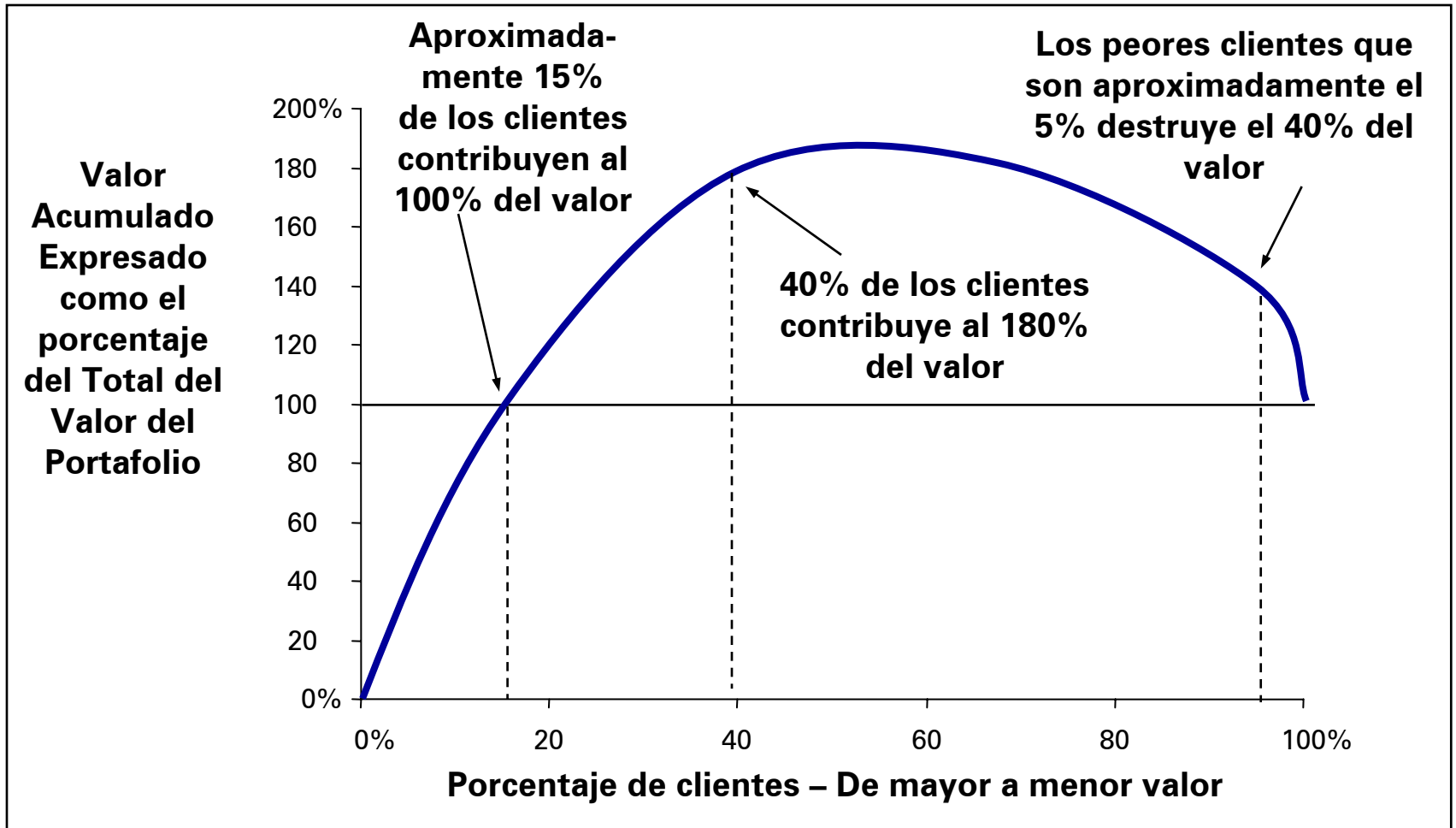
ILUSTRATIVO



Los modelos predictivos funcionan debido a las características básicas del mercado asegurador

Los atributos de los clientes pueden usarse para mostrar que no todos los clientes crean valor

ILUSTRATIVO



Visión global del ciclo de la tarifa

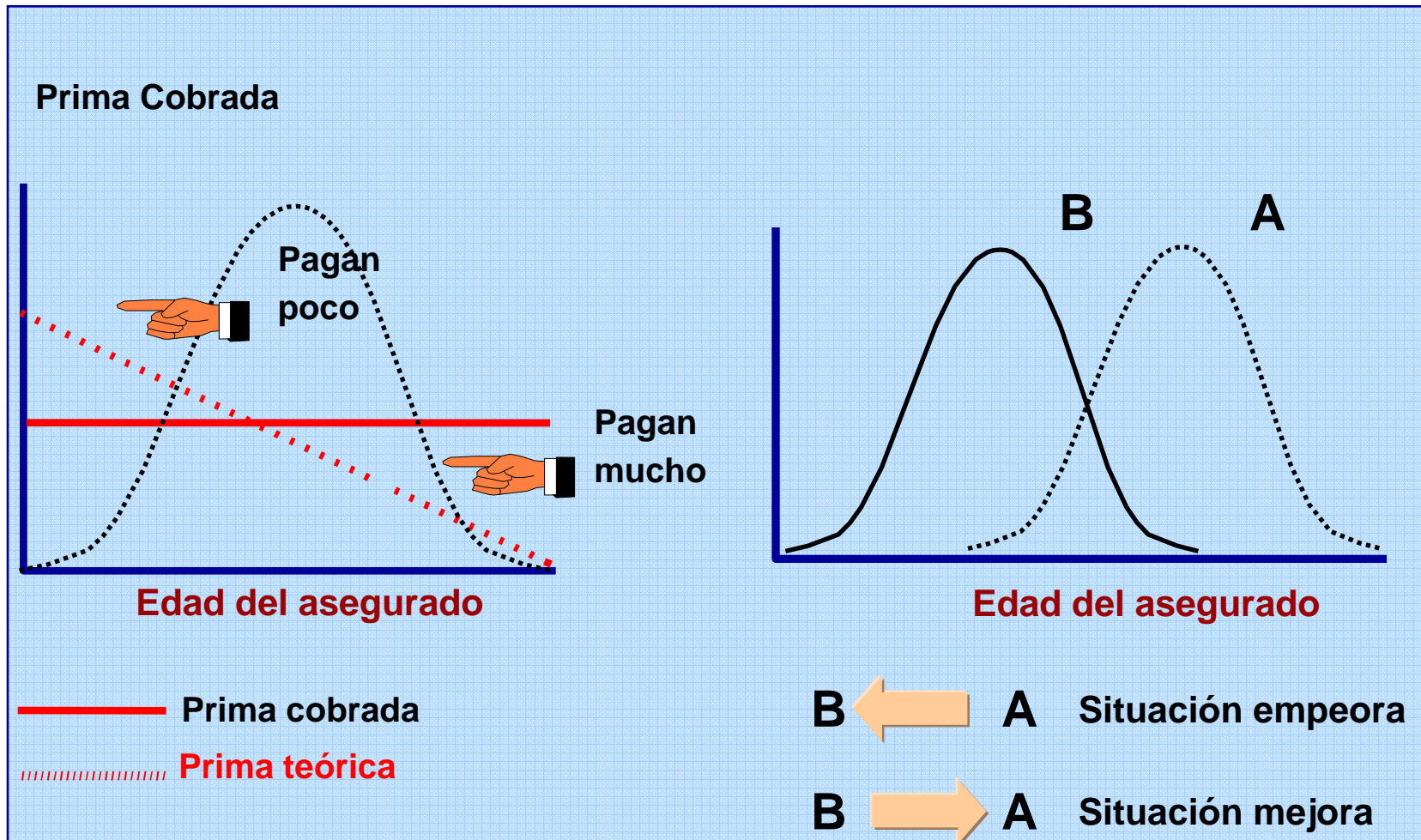


Estructura de la tarifa

- Existen varios tipos de análisis, incluyendo:
 - La determinación de recargos y descuentos que definen la estructura de la tarifa.
 - La determinación del volumen total de primas que se debe recaudar para alcanzar los objetivos de rentabilidad del producto.
- ¿Por qué es necesario contar con una adecuada estructura de la tarifa?
 - Alta competitividad en el mercado
 - Riesgo de anti-selección
 - Búsqueda de rentabilidad en todos los segmentos (jóvenes, provincias, todo tipo de vehículos, ...).
- La información debe ser analizada siniestro por siniestro y a nivel póliza.

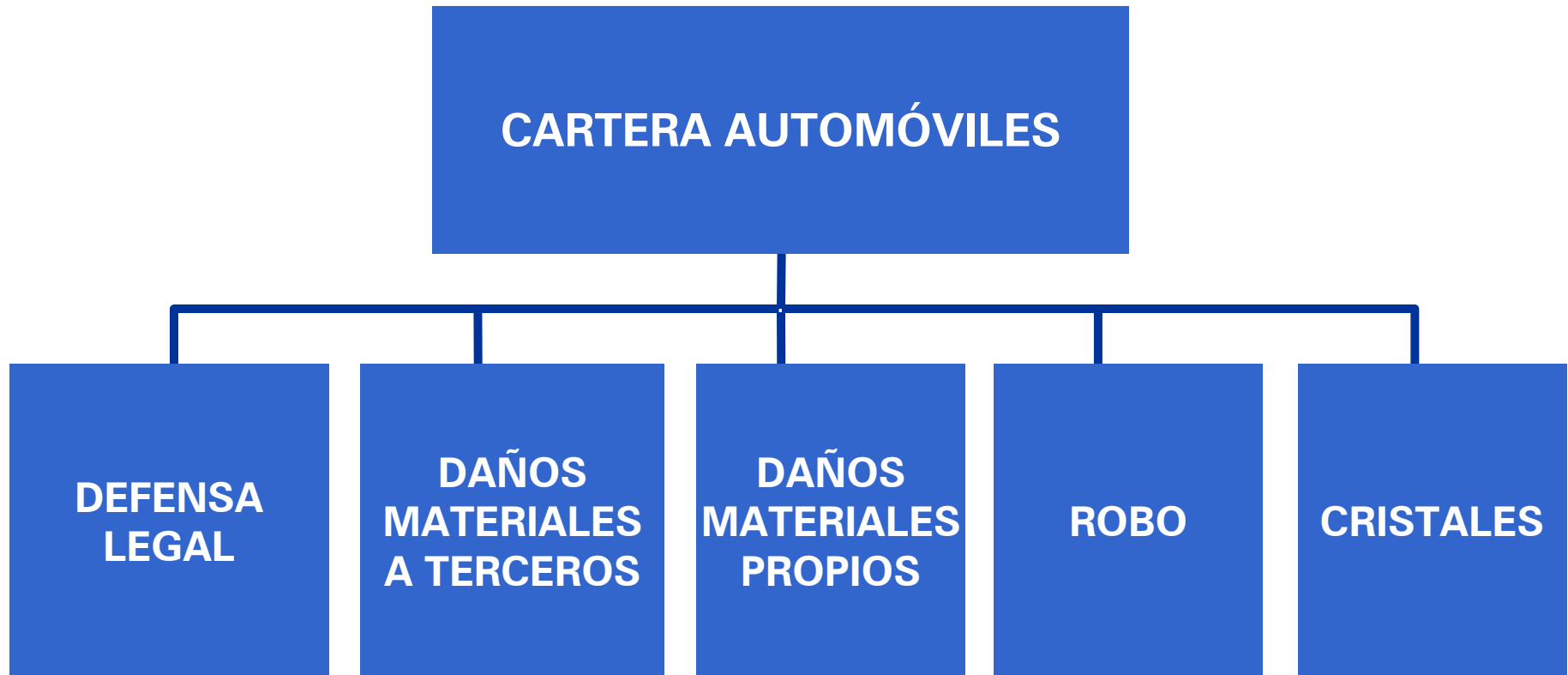
Inclusive, si una compañía alcanza sus metas, una incorrecta estructura de tarifa puede crear anti-selección y por lo tanto, deteriorar sus resultados.

Estructura de la tarifa:
Anti-Selección

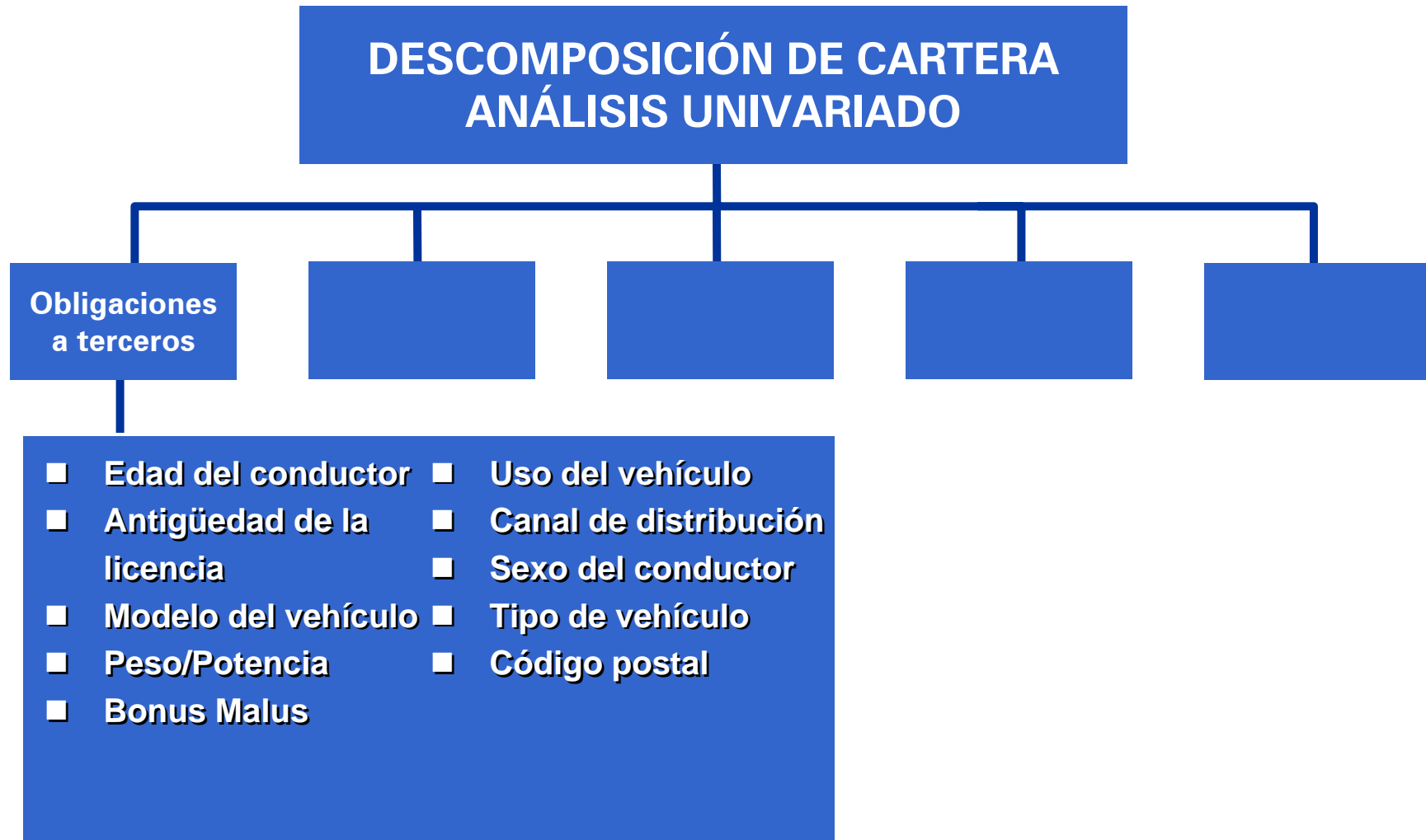


Estructura de la tarifa:

Base de datos (ejemplo autom3viles)



Estructura de la tarifa:
Base de datos



Principales factores utilizados en la tarifa

- Estándar:
 - Antigüedad de la póliza
 - Grupo del vehículo
 - Grupo de la zona
 - Tipo de vehículo
- Otros:
 - Sexo
 - Valor del vehículo
 - Modelo del vehículo
 - Uso del vehículo (personal, negocios)
 - Kilómetros recorridos
 - Lugar donde se estaciona
 - Estado civil del conductor
 - Ocupación
 - Número de años de propiedad del vehículo
 - Conductor joven adicional
 - Número de años asegurado con la compañía

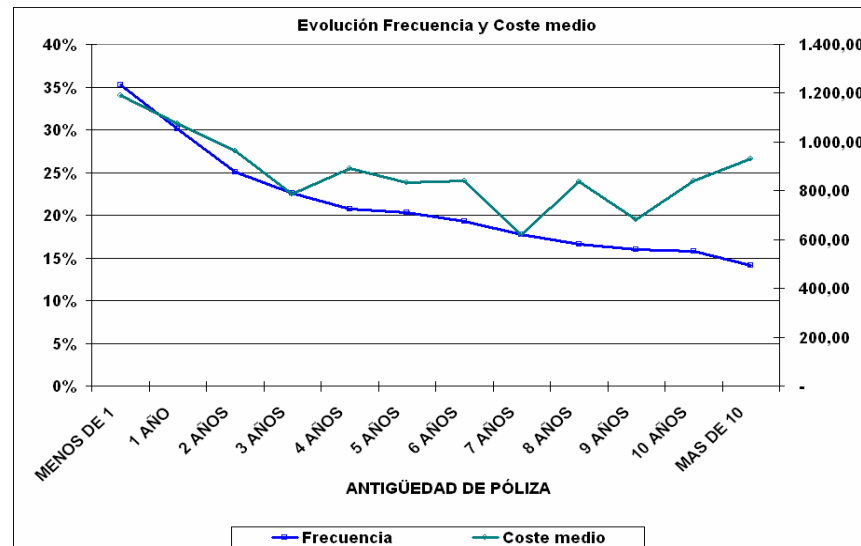
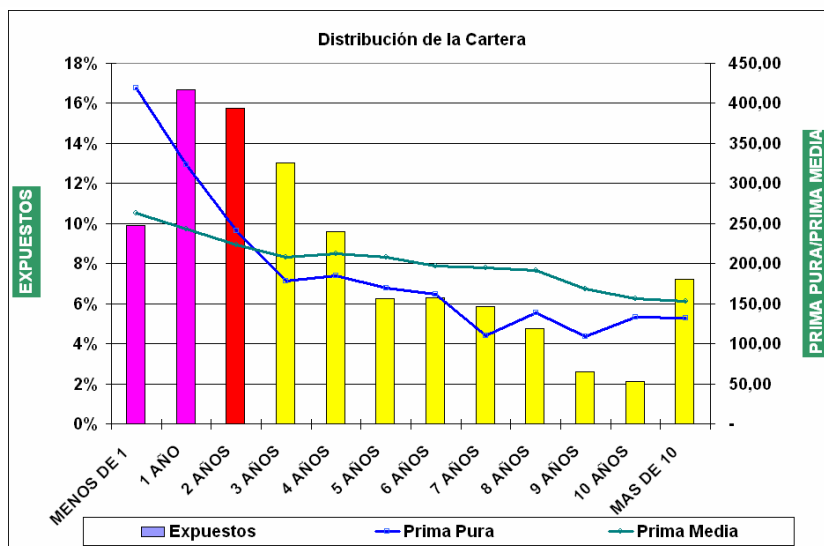
Estructura de la tarifa: Análisis Univariado

Ejemplo Ilustrativo

ANTIGÜEDAD DE PÓLIZA (1)	Expuestos / Año (2)	Distribución Expuestos (3)	Prima Adquirida (4)	Número Siniestros (5)	Incurrido < 60.000 (6)	Incurrido > 60.000 (7)	Incurrido Total (8)	Frecuencia (9)	Coste Medio (10)	Prima Media Adquirida (11)	Prima Pura		Ratio Siniestralidad	
											Total (12)	Truncada (13)	Total (14)	Truncado (15)
MENOS DE 1						7.999.906	28.479.785	35,2%	1.279	263	451	324	171,5%	123,3%
1 AÑO						9.308.006	36.696.887	30,1%	1.148	243	346	258	142,4%	106,3%
2 AÑOS						5.993.240	25.711.389	25,1%	1.023	224	256	197	114,6%	87,9%
3 AÑOS						2.539.342	15.362.485	22,6%	820	208	185	155	89,0%	74,3%
4 AÑOS						2.804.591	11.978.136	20,7%	946	212	196	150	92,4%	70,8%
5 AÑOS						1.518.850	7.125.700	20,3%	881	207	179	141	86,2%	67,8%
6 AÑOS						1.918.072	7.125.700	19,3%	901	197	174	126	88,3%	64,0%
7 AÑOS						1.418.072	7.125.700	17,7%	624	194	111	107	57,0%	55,0%
8 AÑOS						791.910	4.406.610	16,6%	877	191	145	119	76,0%	62,4%
9 AÑOS						317.357	1.893.308	16,0%	711	169	114	95	67,5%	56,2%
10 AÑOS						391.024	1.873.404	15,8%	886	156	140	111	89,6%	70,9%
MAS DE 10						2.253.534	6.641.283	14,2%	1.019	153	144	95	94,6%	62,5%

- Frecuencia
- Costo Medio
- Prima Media
- Prima Pura
- Índice Siniestralidad

Total	636.842	100,0%	136.102.218	149.478	115.265.028	35.976.285	151.241.313	23,5%	1.012	214	237	181	111,1%	84,7%
--------------	---------	--------	-------------	---------	-------------	------------	-------------	-------	-------	-----	-----	-----	--------	-------



Estructura de la tarifa:
Análisis Univariado

Problemas

- No identifica que factores de tarifa son los realmente importantes.
- No capta las correlaciones e interacciones entre los factores de tarifa (Ej. Sexo y edad del conductor).
- No ajusta los resultados por el hecho de que la distribución de la cartera es diferente para cada nivel de cada factor de tarifa.
- No permite separar lo relevante de lo redundante (correlaciones e interacciones).

Se necesita un método que pueda procesar muchos factores de tarifa para una mejor segmentación.

¿Qué son los GLM's?

- Los modelos generalizados lineales (GLM's) son una extensión de los modelos lineales tradicionales que permiten medir la relación entre múltiples variables predecibles y la variable dependiente, considerando la interdependencia entre los factores.
- El propósito de los GLM's es expresar la relación entre la variable de observada, "Y" y el número de covariables, "X".
- Las variables que se busca analizar en un proyecto para obtener una tarifa es:
 - Frecuencia (N° siniestros / Expuestos)
 - Severidad (Ocurridos / N° siniestros)
- Los GLM's consideran:
 - Efectos de multiplicación y sumas
 - El uso de distribuciones entre familias exponenciales
 - Cada observación es independiente

¿Qué son los GLM's?

- Una diferencia en que los GLM difiere de una regresión múltiple es en relación al número de variables dependientes que puede analizar.
- El vector \mathbf{Y} de n observaciones de una sola variable Y se puede reemplazar por una matriz \mathbf{Y} de n observaciones de m diferentes variables Y .
- Igualmente, el vector de coeficiente de regresión \mathbf{b} para una variable Y se puede reemplazar por una matriz de coeficientes de regresión \mathbf{b} , con un vector de coeficientes \mathbf{b} para cada una variable dependiente m .

¿Qué son los GLM's?

- Un modelo GLM puede ser expresado como:
- $Y\mathbf{M} = \mathbf{X}\mathbf{b} + \mathbf{e}$
- Donde Y , \mathbf{X} , \mathbf{b} , y \mathbf{e} son las mismas usadas en un modelo de regresión multivariable y \mathbf{M} es una matriz de coeficientes de $m \times s$ que definen la transformación lineal de las variables dependientes s . Las ecuaciones normales son:
- $\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{Y}\mathbf{M}$
- Y la solución a la ecuación normal está dada por:
- $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{M}$
- Donde la inversa de $\mathbf{X}'\mathbf{X}$ es la inversa generalizada si $\mathbf{X}'\mathbf{X}$ contiene columnas redundantes.

Análisis Multivariado: Definición de modelos

■ **Frecuencia:**

- Variable dependiente: *Número de siniestros*
- Distribución asociada: *Poisson*
- Variable vínculo: *Log*
- Offset (ponderador): *Expuestos*
- Escalar: *Desviación*

■ **Severidad:**

- Variable dependiente: *Costo promedio*
- Distribución asociada: *Gamma*
- Variable vínculo: *Log*
- Offset (ponderador): *N/A*
- Escalar: *Desviación*

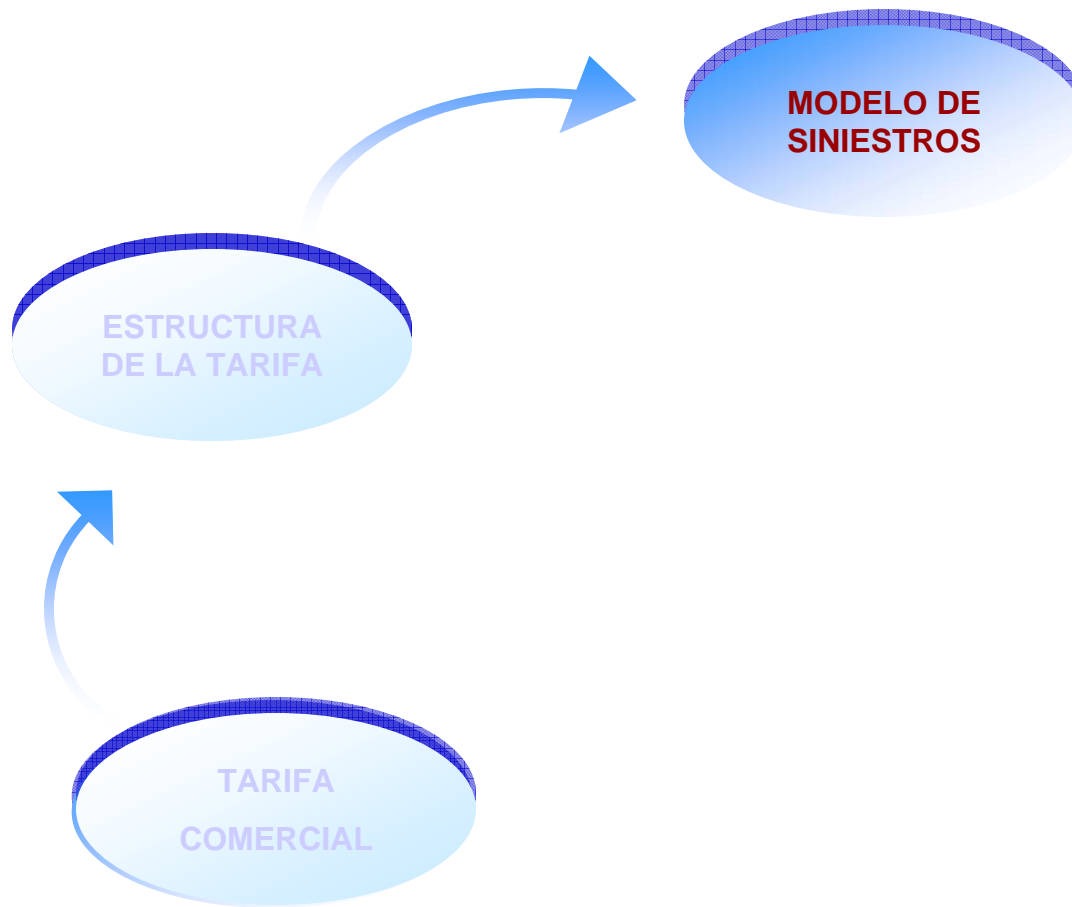
Estructura de la tarifa: Zonas geográficas

Problemas

- Cada compañía tiene una distribución geográfica diferente.
- Pocas compañías tienen suficiente experiencia a nivel nacional.
- Si las zonas no están bien agrupadas, no habrá diferenciación de riesgo.
- Es necesario vincular el análisis geográfico con los modelos multivariados.

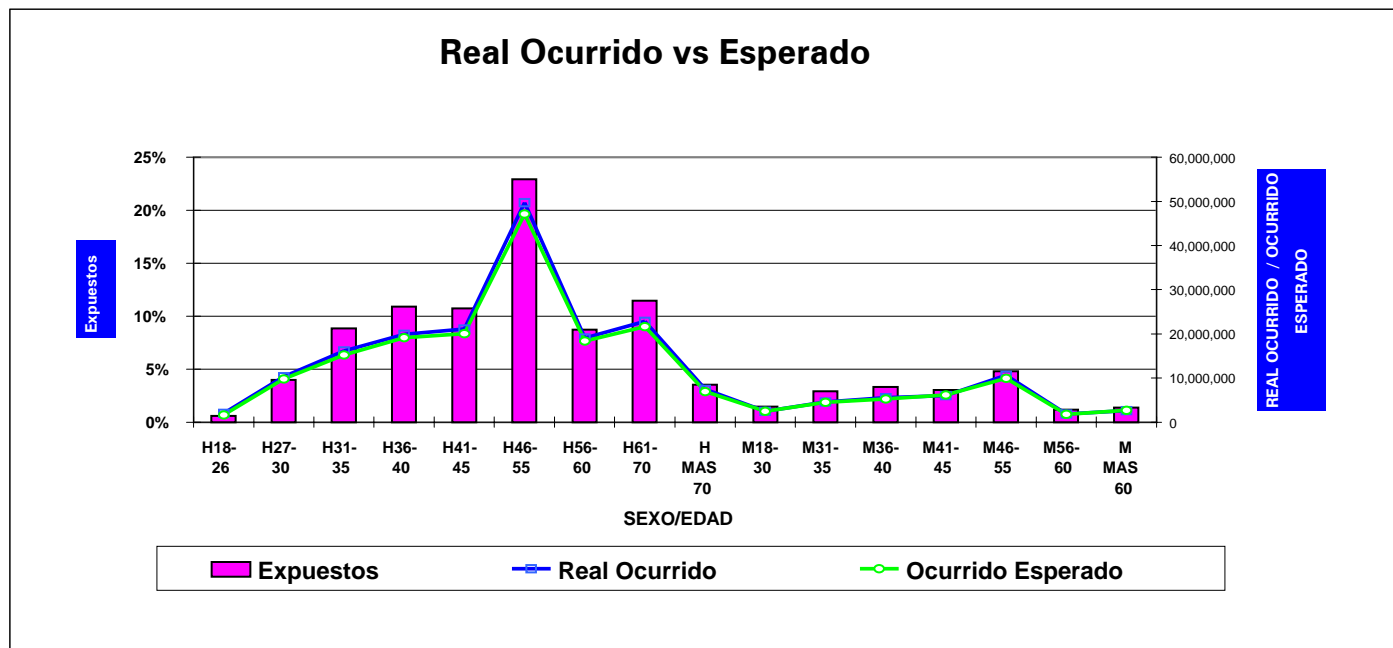


Visión global del ciclo de la tarifa

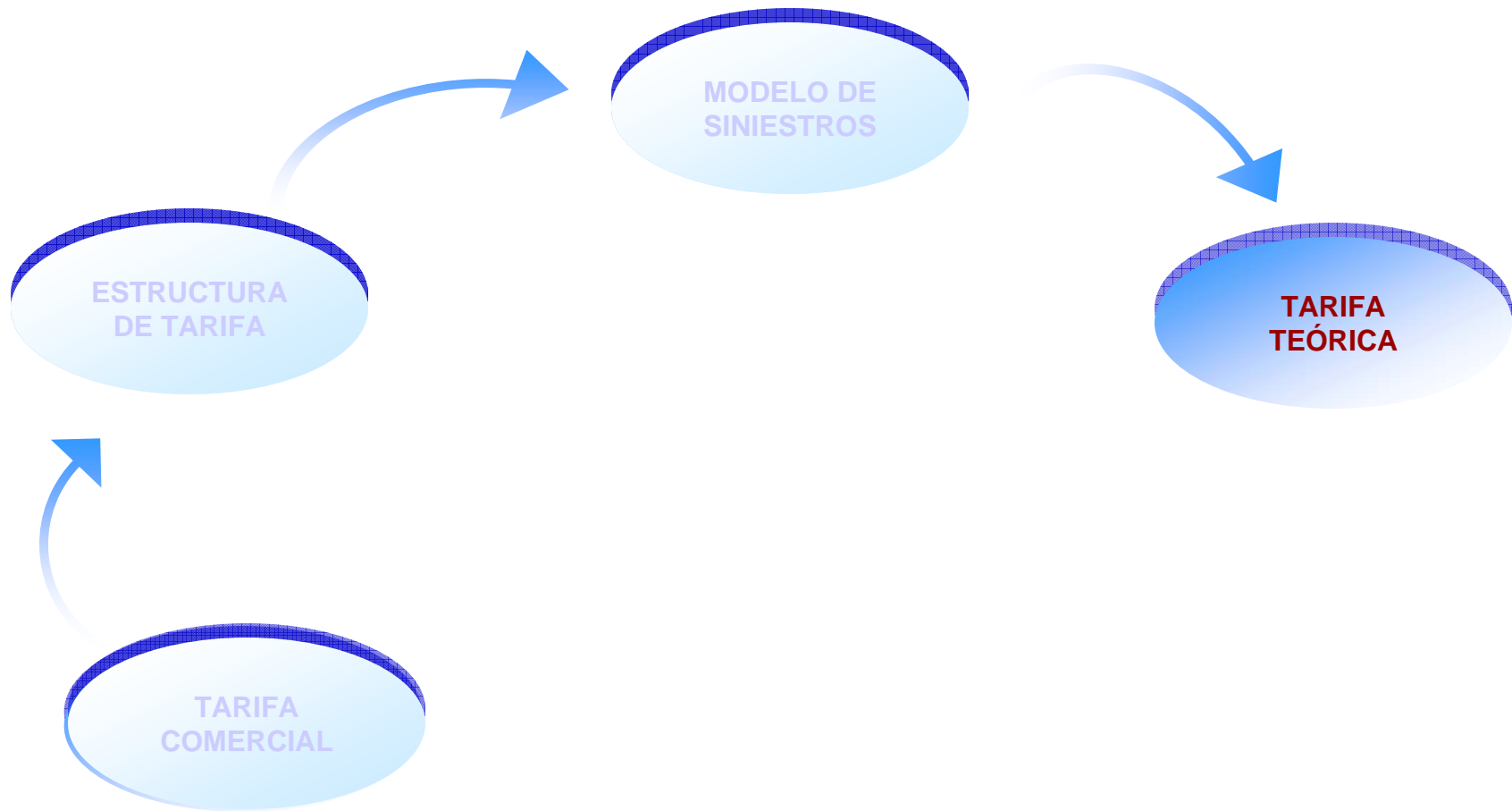


Modelo de índice de siniestralidad

- Es necesario reconstruir el costo de los siniestros generados por el modelo y compararlo con el costo real, a fin de obtener desviaciones entre el costo real y el reproducido por el modelo.
- En la selección del “mejor” modelo, la decisión usualmente se encuentra entre el número de factores de suscripción y la estabilidad del modelo.



Visión global del ciclo de la tarifa

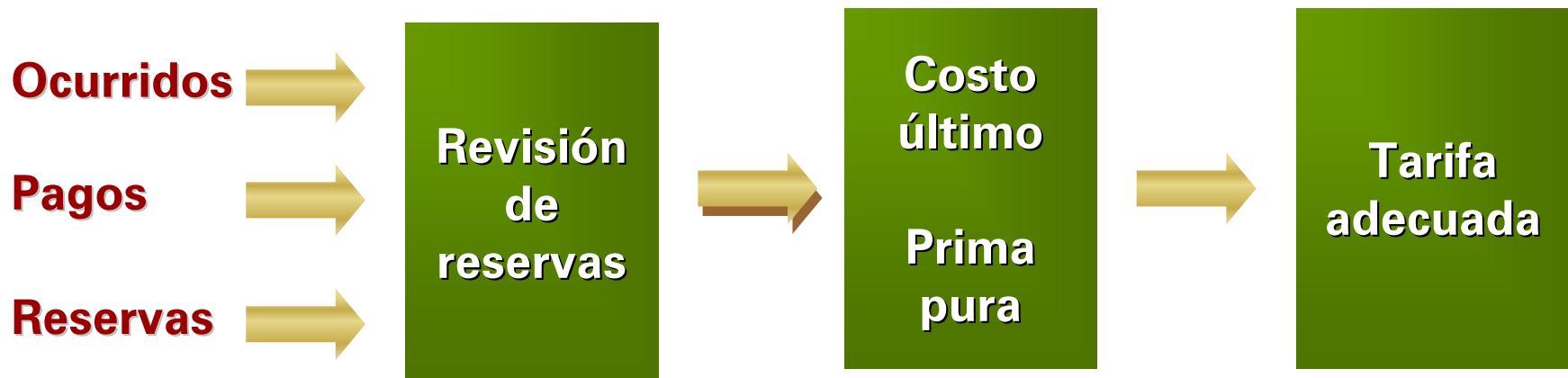


Riesgo técnico de la tarifa

- Investigación inicial de información.
- Análisis de los factores que conforman la tarifa.
- Modelo de frecuencia y severidad separadamente (GLM's).
- Ajuste del modelo a la información, es decir, diferente modelo para cada tipo de siniestro.
- Contar con factores de tarifa apropiados para cada grupo.
- Pruebas acorde con el ajuste del modelo y la información disponible.

Tarifa teórica: Prima pura

- Una vez que se ha establecido la estructura de la tarifa, necesitamos determinar el nivel de prima pura asociada con la estructura de la tarifa.
- La prima pura es generada por la experiencia de siniestros pasados de la compañía.
- El cálculo de la prima pura debe incluir el costo último, incluyendo las estimaciones de reserva SONR y SPV.
- También es necesario incluir hipótesis sobre inflación, costo del reaseguro y siniestros graves entre otros.



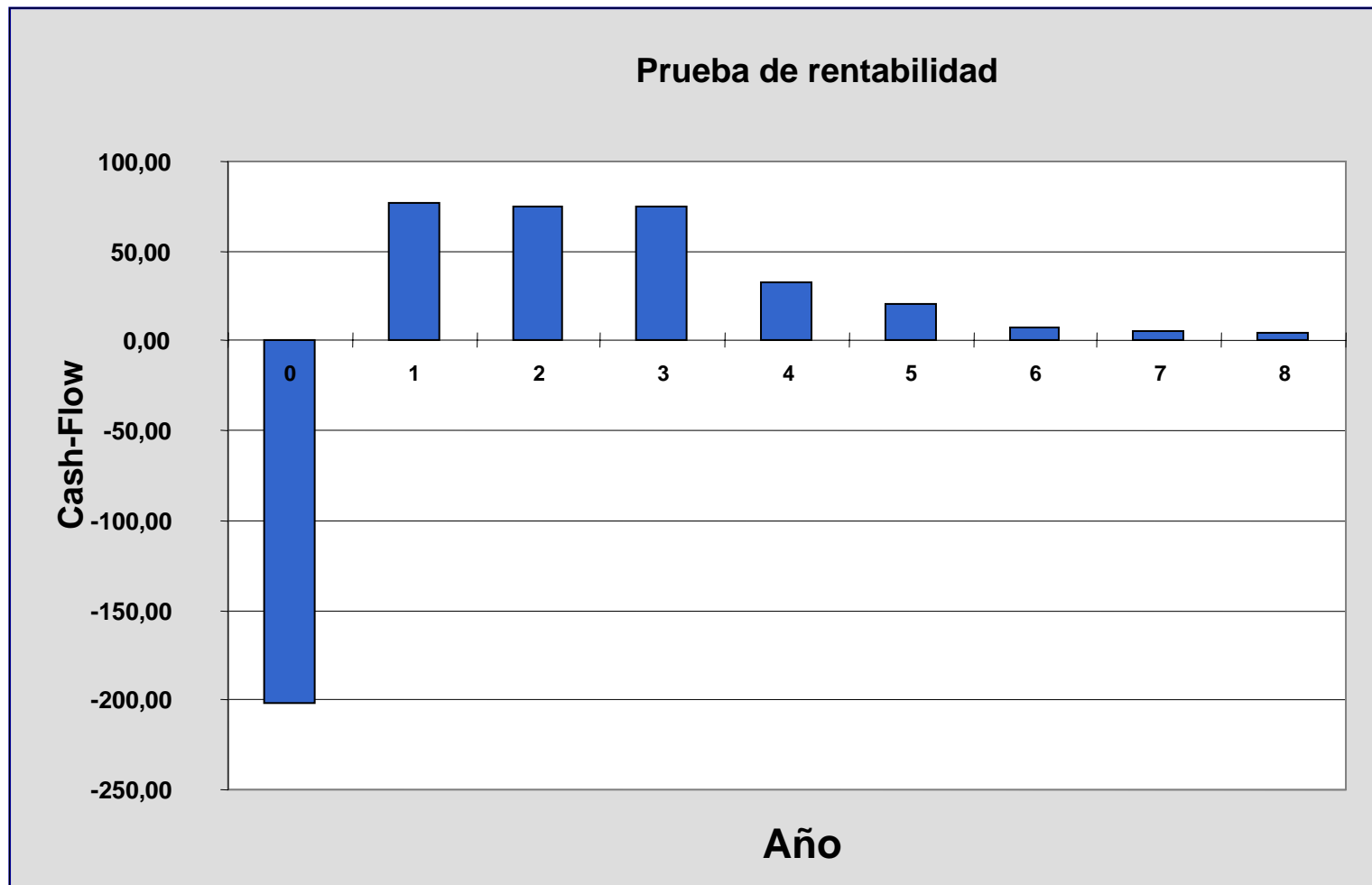
Tarifa teórica:

Prueba de rentabilidad

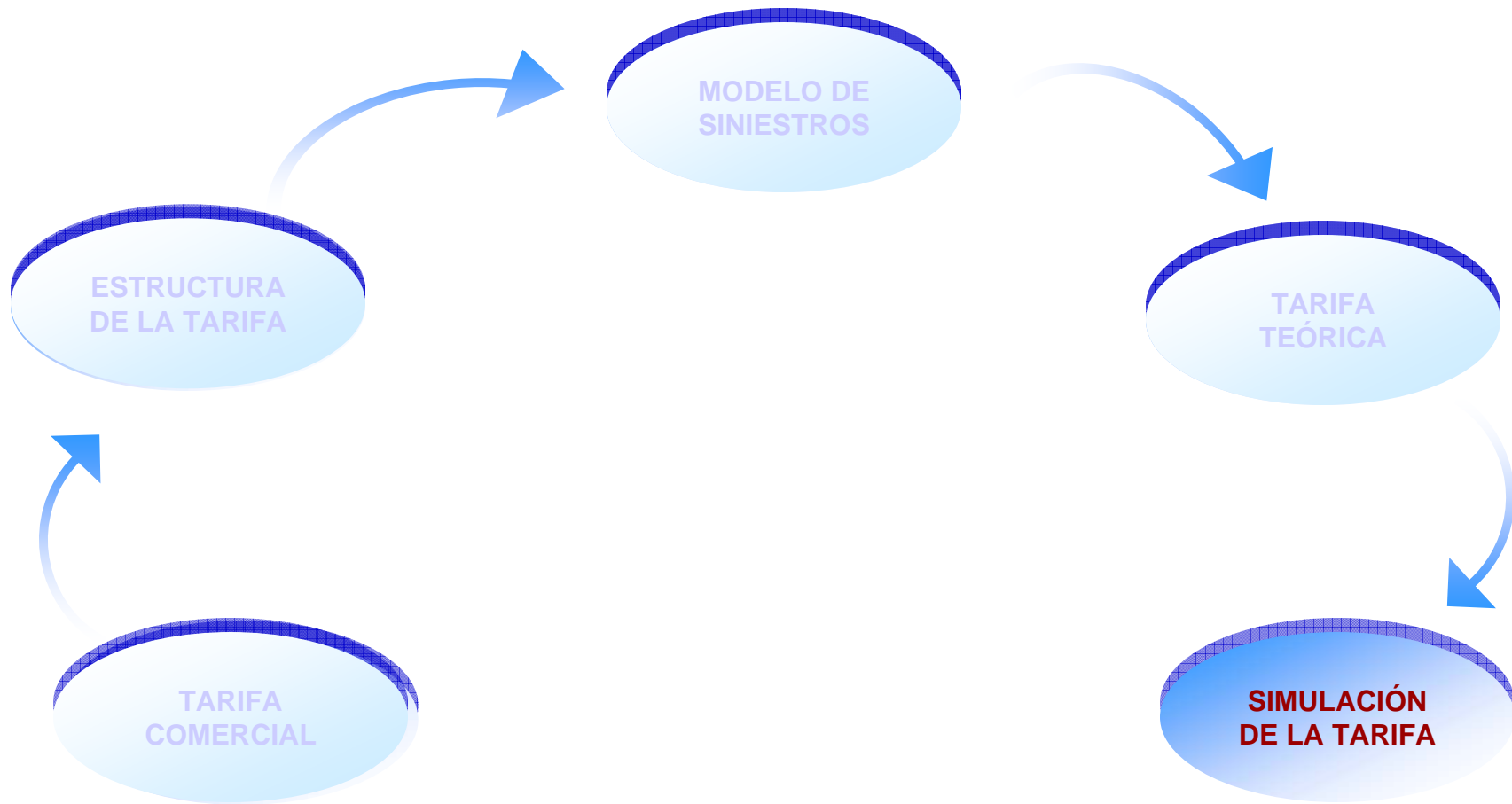
- ¿Qué es la prueba de rentabilidad?
La prueba de rentabilidad se refiere al uso de un modelo que proyecte los ingresos y los egresos de una póliza”.
- ¿Cómo se puede ajustar la prueba de rentabilidad a la tarifa?
La prueba de rentabilidad permite calcular los recargos y descuentos partiendo de la prima de riesgo para obtener la prima comercial.
- ¿Cómo funciona la prueba de rentabilidad?
Los principales componentes de la prueba de rentabilidad son el patrón de pagos, gastos, siniestros y costo de capital, todos estimados a valor presente.

Tarifa teórica:
Prueba de rentabilidad

Caso práctico

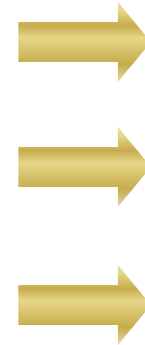
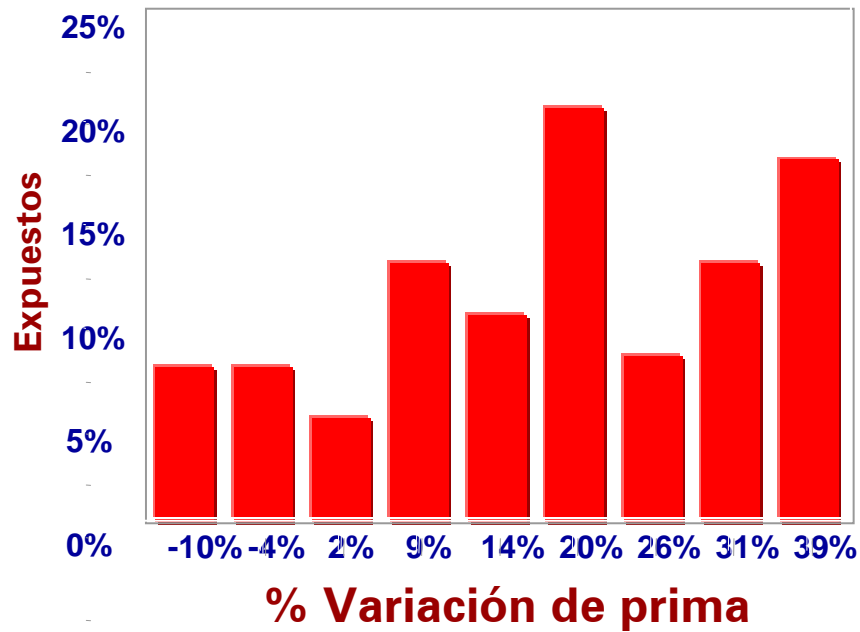


Visión global del ciclo de la tarifa

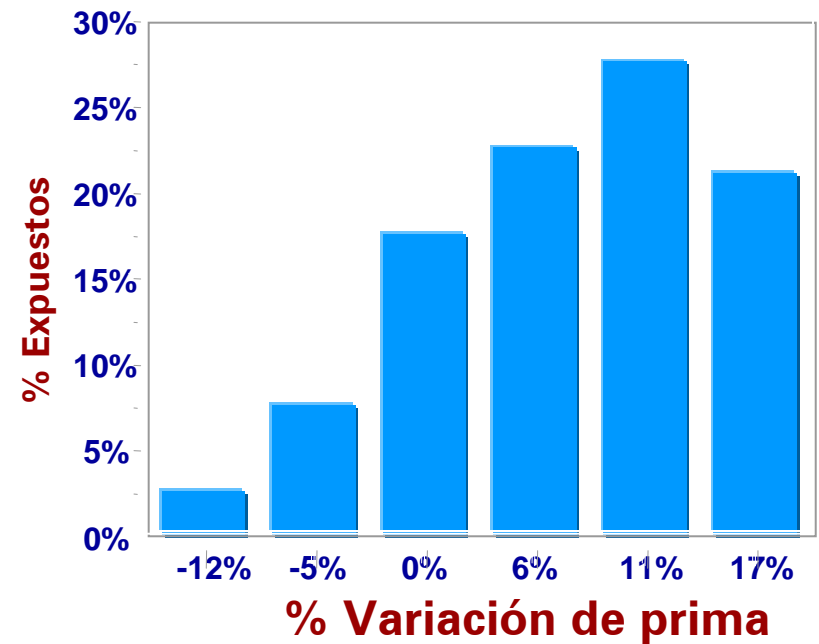


Simulación de la tarifa

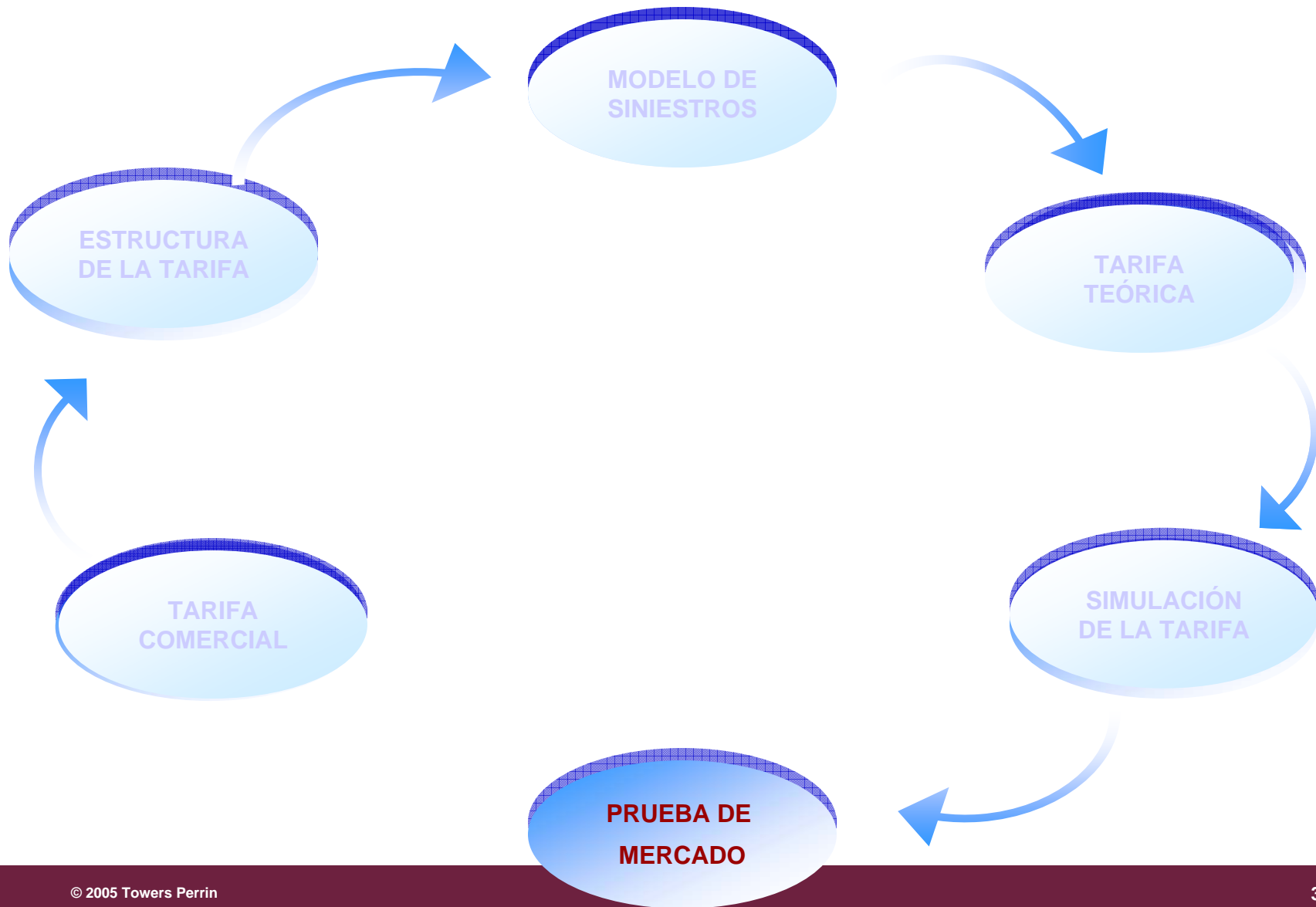
Tarifa teórica



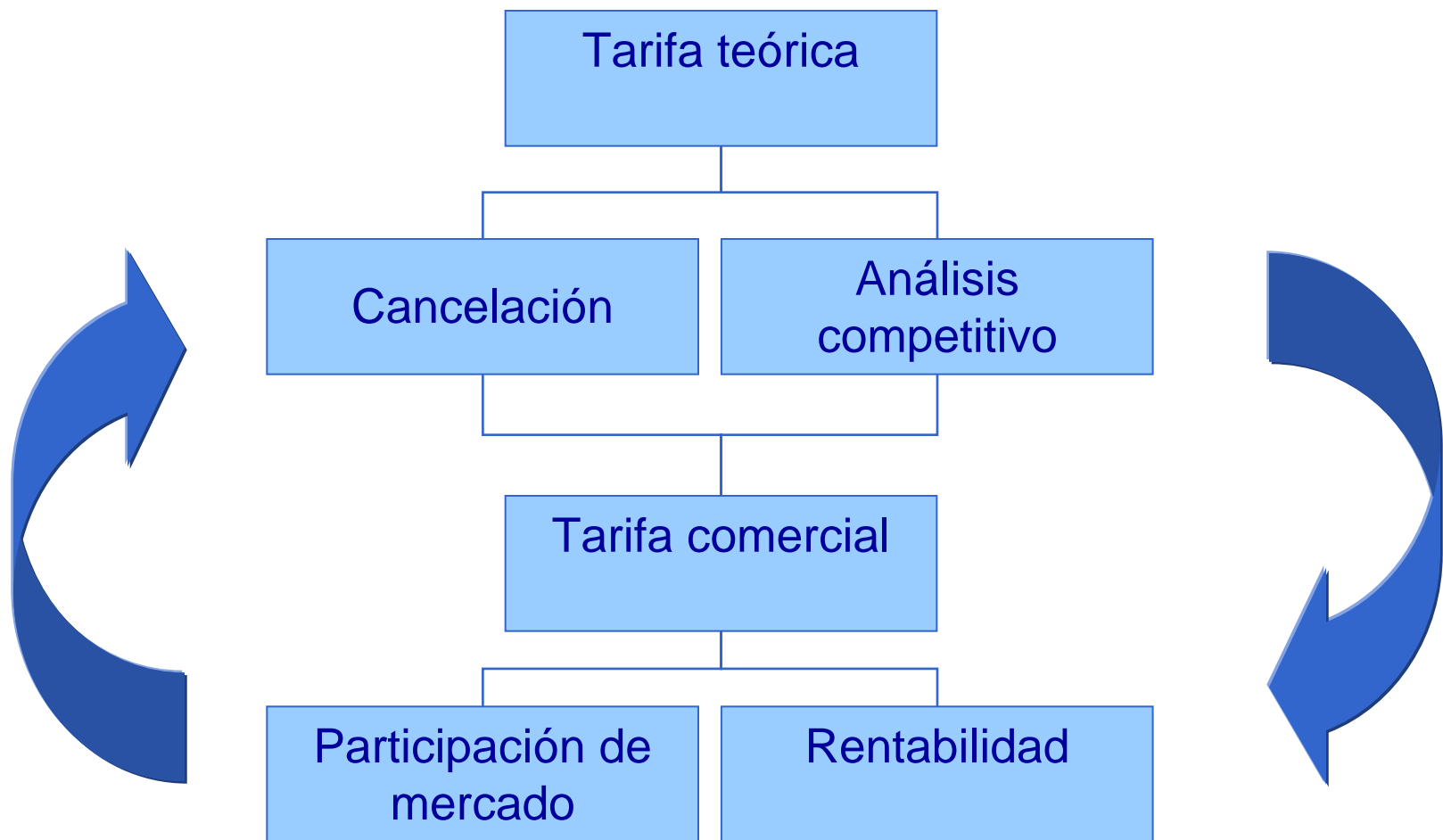
Tarifa comercial



Visión global del ciclo de tarifa



Simulación de mercado:
Prueba de mercado



Simulación de mercado:

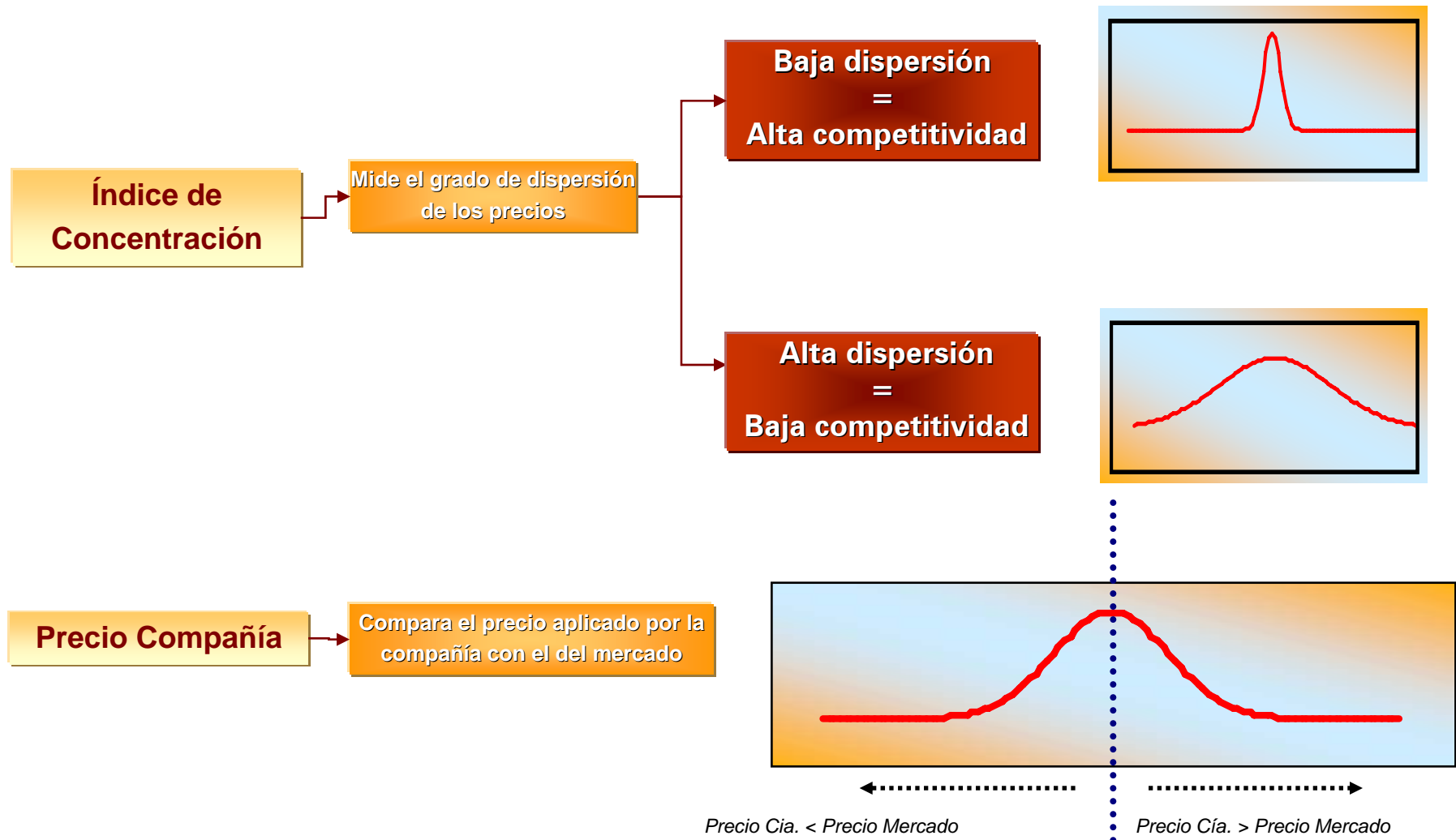
Análisis de cancelaciones

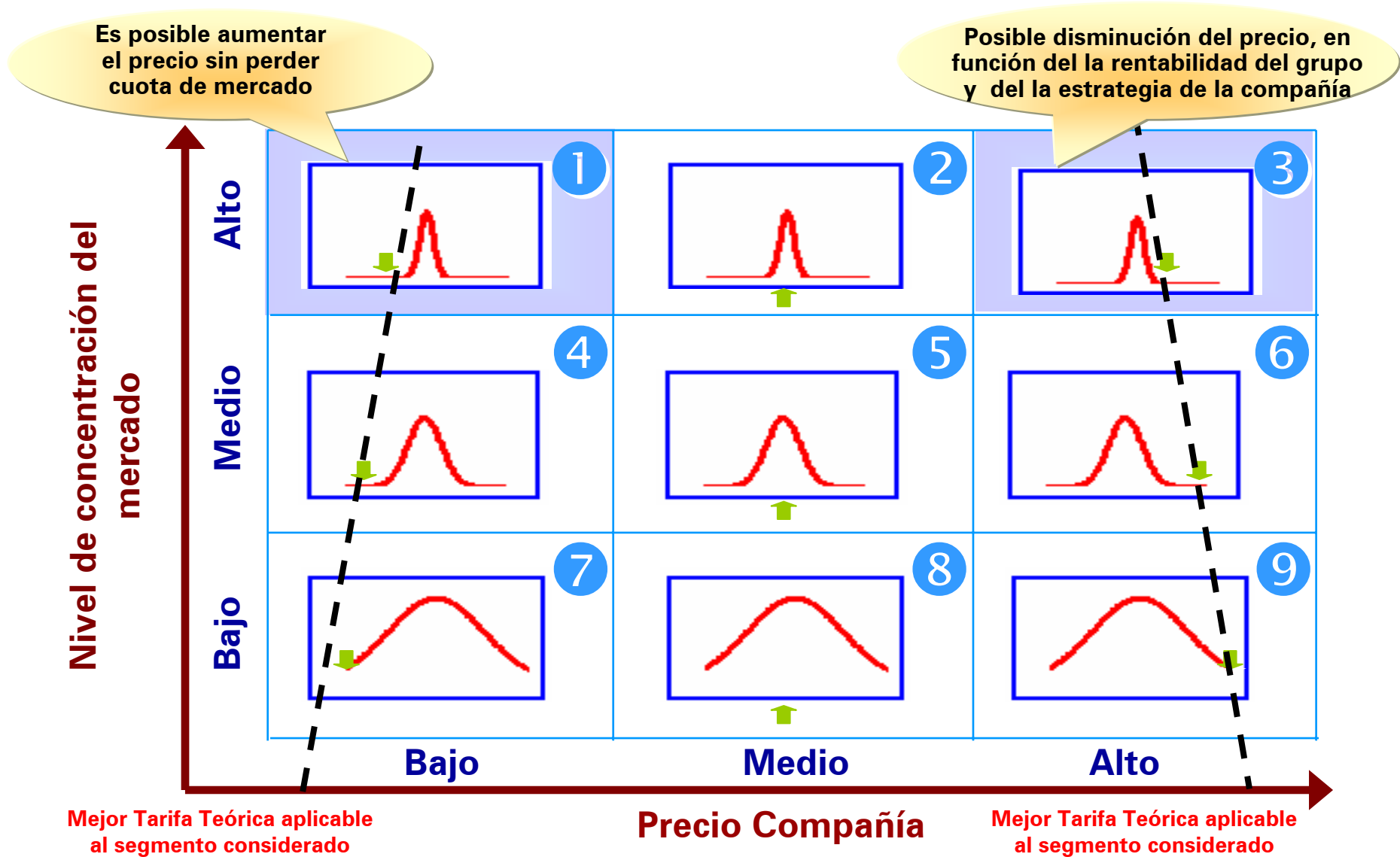
- El objetivo es predecir las cancelaciones dado un cambio de tarifa.
- Se modela la tasa de anulación utilizando GLM's.
- Requiere datos detallados (histórico de primas cobradas,...).

Análisis competitivo de mercado (CMA)

APLICACIONES POSIBLES:

- 1. Situación competitiva y posición estratégica.**
- 2. Fijación de objetivos comerciales:**
 - **Identificar segmentos alta/baja competitividad.**
- 3. Análisis de nueva producción.**
- 4. Análisis de cartera:**
 - **Identificar segmentos vulnerables.**
- 5. Seguimiento de cambios de precios en el tiempo.**







**TOWERS
PERRIN**

TILLINGHAST

Contribuciones recientes de la Estadística a la Actuaría en México

*Tarifa y Segmentación de Riesgos usando
Modelos Predictivos*

Roberto Ambríz
Eduardo Esteva
Georgina Montiel

Noviembre 23, 2007

This document is incomplete without the accompanying discussion; it is confidential and intended solely for the information and benefit of the immediate recipient hereof.