Estadística Bayesiana y Riesgos

Manuel Mendoza Ramírez
Instituto Tecnológico Autónomo de México

Seminario Aleatorio. Contribuciones Recientes de la Estadística a la Actuaría en México.
ITAM. México, D.F. Noviembre 23, 2007.

Estadística Bayesiana y Solvencia

Manuel Mendoza Ramírez
Instituto Tecnológico Autónomo de México

Seminario Aleatorio. Contribuciones Recientes de la Estadística a la Actuaría en México.
ITAM. México, D.F. Noviembre 23, 2007.

Contenido

- Introducción
- Condición de solvencia
- Formulación estadística
- Alternativa Bayesiana
- Ejemplos
- Consideraciones finales

Introducción

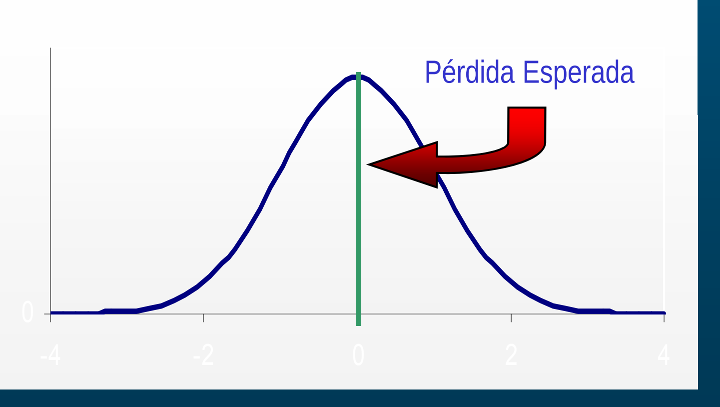
Actuaría y Riesgos Financieros

- La incertidumbre está implícita y es inevitable en el origen de las Ciencias Actuariales
- Una parte de la Actuaría se ocupa del estudio de los fenómenos (aleatorios) que pueden producir las llamadas pérdidas contingentes
- La información disponible se utiliza para pronosticar el comportamiento futuro de las pérdidas

Introducción

Modelos

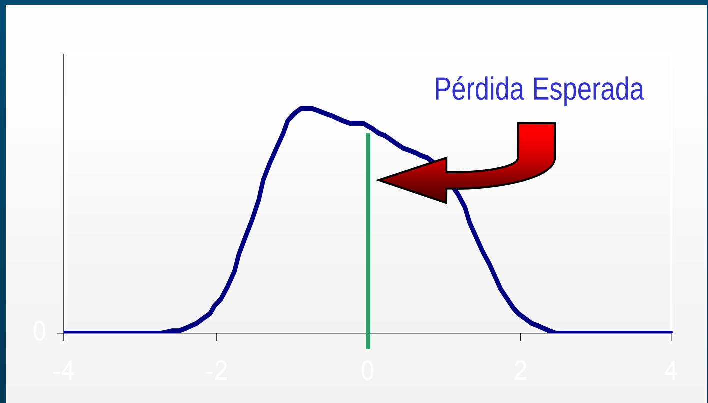
- Una componente fundamental es la distribución de pérdidas



Introducción

Modelos

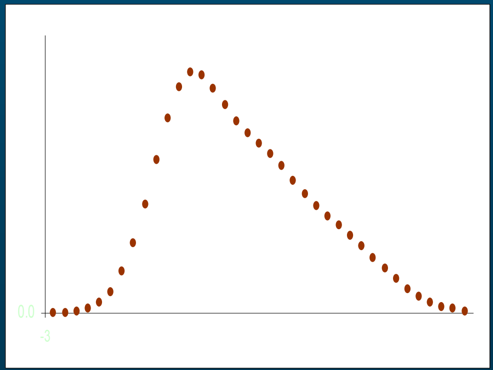
- Una componente fundamental es la distribución de pérdidas



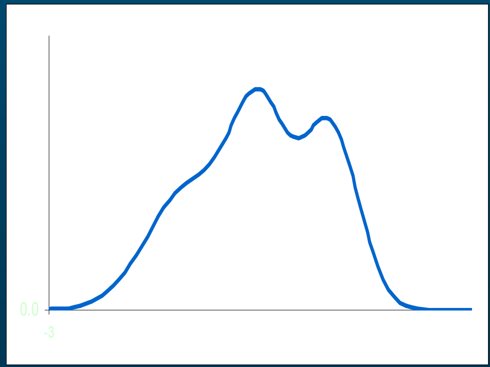
Introducción

Supuestos

- En forma natural se incorporan supuestos en el proceso de modelado. Por ejemplo, la hipótesis de independencia entre frecuencia y severidad de las pérdidas



Frecuencia



Severidad

Introducción

Supuestos

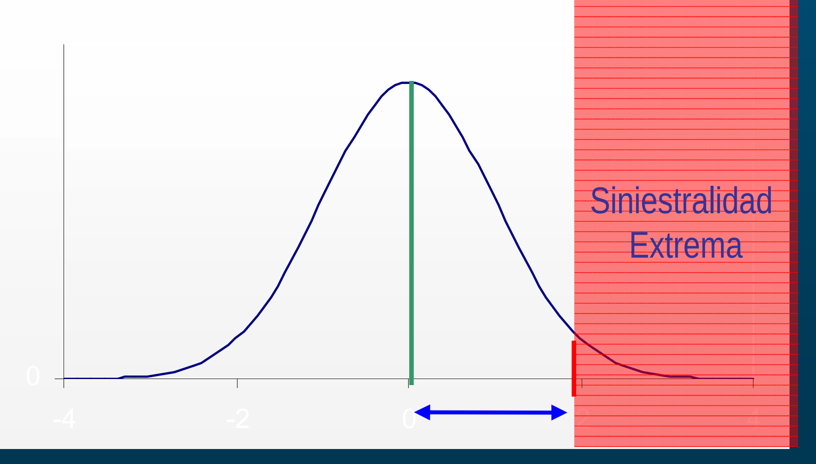
- En forma natural se incorporan supuestos en el proceso de modelado. Por ejemplo, la hipótesis de independencia entre frecuencia y severidad de las pérdidas

$$L = S \times F \quad \Rightarrow \quad E(L) = E(S) \times E(F)$$

Solvencia

Pérdidas Extremas

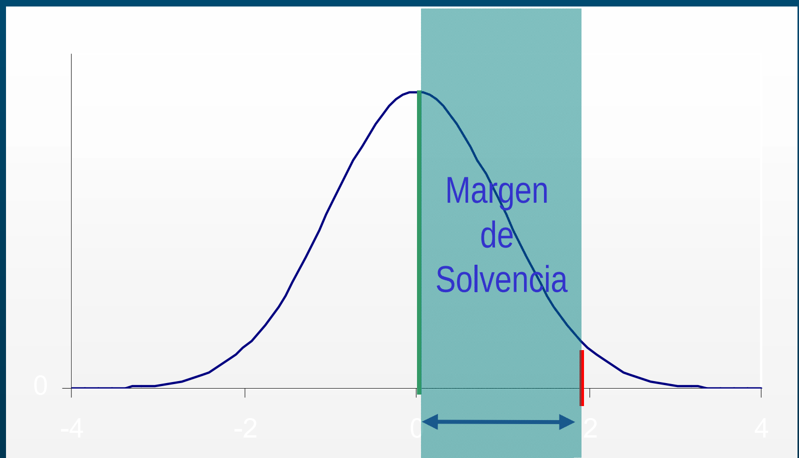
- Un análisis estadístico debe incluir una descripción completa de la distribución de interés. En Actuaría, en particular, es conveniente describir las pérdidas extremas



Solvencia

Pérdidas Extremas

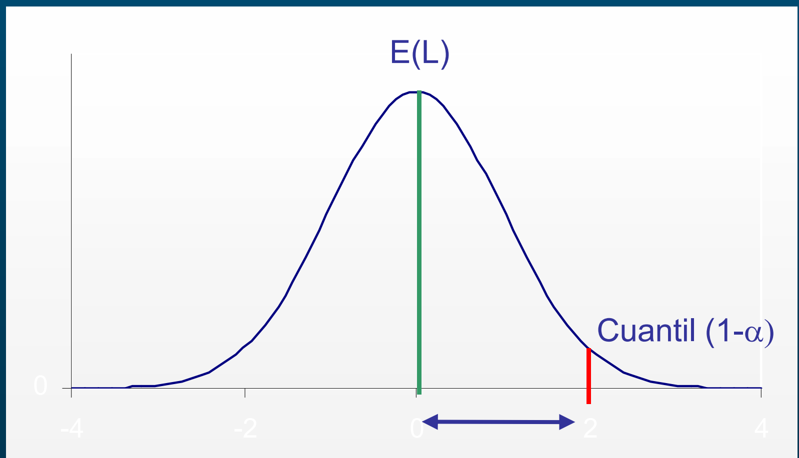
- Para las agencias reguladoras es importante establecer un margen de solvencia.



Solvencia

Pérdidas Extremas

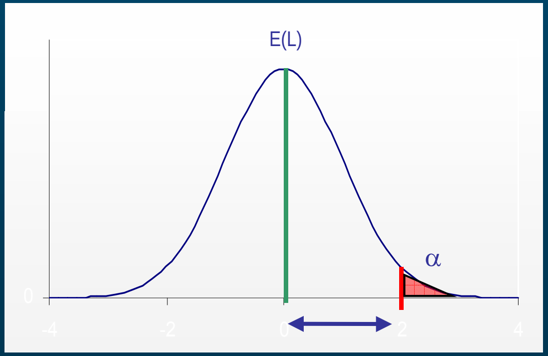
- Para las agencias reguladoras es importante establecer un margen de solvencia.



Solvencia

Condición de Solvencia

- Un sistema de seguros es solvente si cuenta con recursos suficientes para hacer frente a sus obligaciones.
- Un sistema es (1 – α)-solvente si cuenta con recursos para hacer frente a sus obligaciones con probabilidad 1 - α.



Formulación Estadística

- Dos series históricas:
  - $X_1, X_2, \ldots, X_T$ (primas)
  - $Y_1, Y_2, \ldots, Y_T$ (reclamaciones)
- Objetivo:
  - Determinar el valor $q^{1-\alpha}_{Y,T+1}$ tal que $P(Y_{T+1} > q^{1-\alpha}_{Y,T+1}) = \alpha$

Formulación Estadística

- La serie de siniestralidad relativa
  - $W_1, W_2, \ldots, W_T$ donde $W_i = \dfrac{Y_i}{X_i}$ (i = 1,...,T)
  - $W_1, W_2, \ldots, W_T$ se suelen considerar i.i.d. (estabilidad)
  - El interés se concentra en el valor $q_W^{1-\alpha}$ tal que

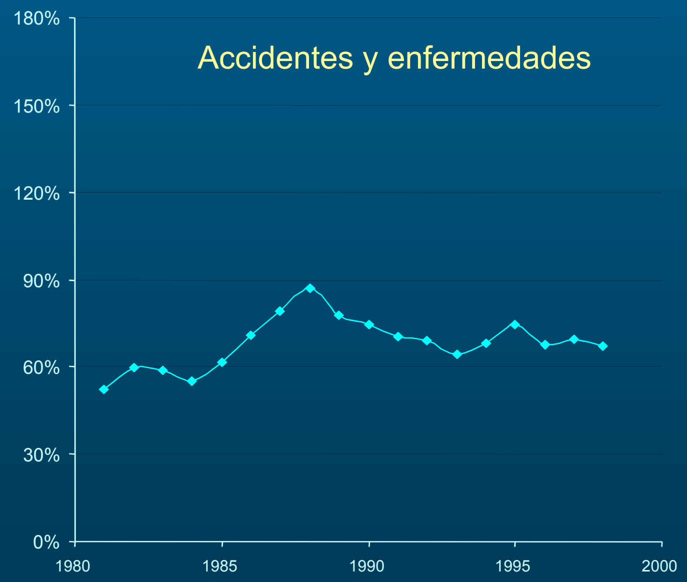
$$P(W > q_W^{1-\alpha}) = \alpha$$

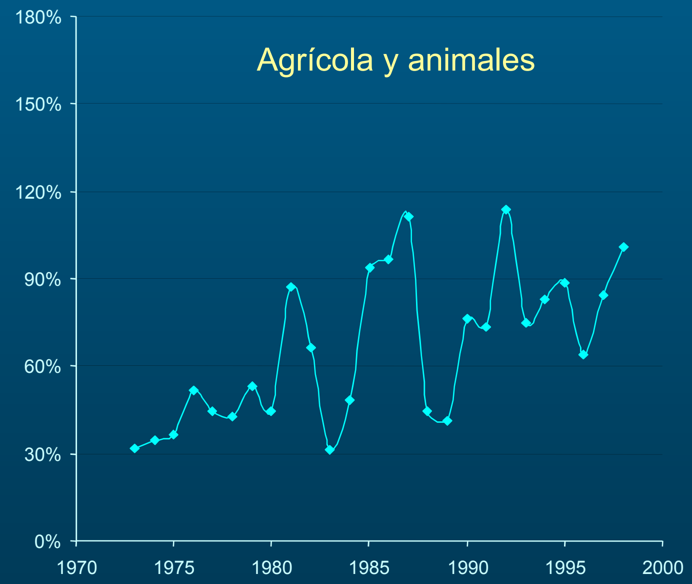
Formulación Estadística

- Entonces, para la siguiente observación,
  - $P(W_{T+1} > q_W^{1-\alpha}) = \alpha \quad \Rightarrow \quad P\left(\frac{Y_{T+1}}{X_{T+1}} > q_W^{1-\alpha}\right) = \alpha$
- De manera que
  - $P(Y_{T+1} > (q_W^{1-\alpha})(X_{T+1})) = \alpha \quad \Rightarrow \quad \boxed{q_{Y,T+1}^{1-\alpha} = (q_W^{1-\alpha})(X_{T+1})}$

Formulación Estadística

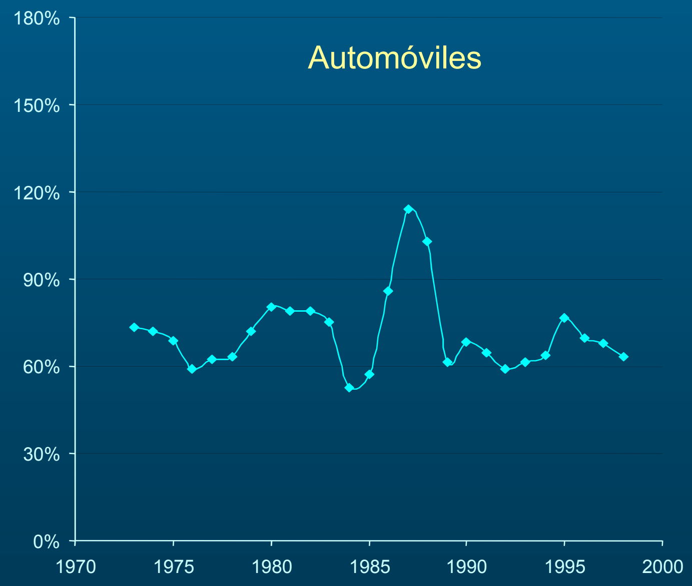
Datos históricos





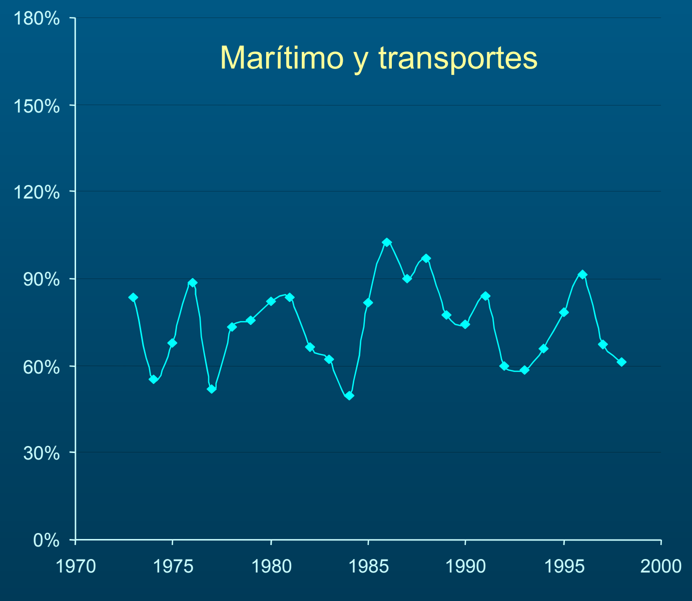
Formulación Estadística

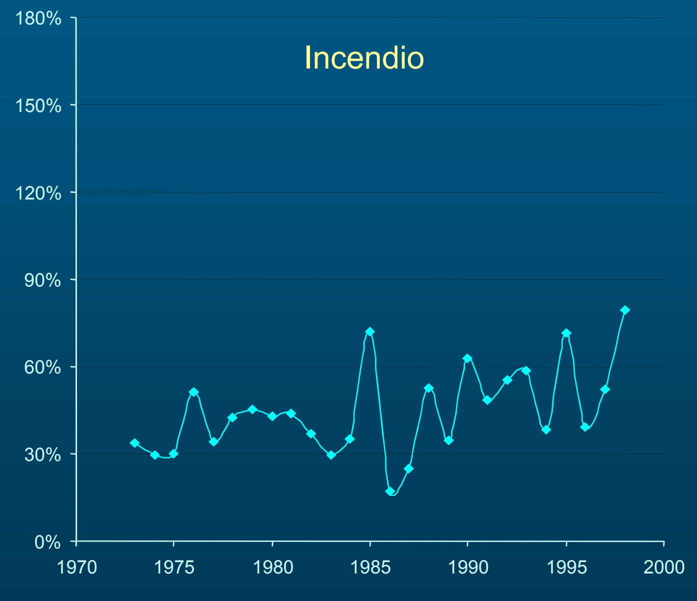
Datos históricos



Formulación Estadística

Datos históricos





Formulación Estadística

Distribución para W

- Problema estadístico
  - Estimar los cuantiles de la distribución de W
- Normalidad como primera aproximación
  - El cálculo es extremadamente simple $\left(q_W^{1-\alpha} = \hat{\mu}_W + z_{(1-\alpha)}\hat{\sigma}_W \right)$
  - Los datos históricos no son compatibles con el supuesto

Formulación Estadística

Distribución para W

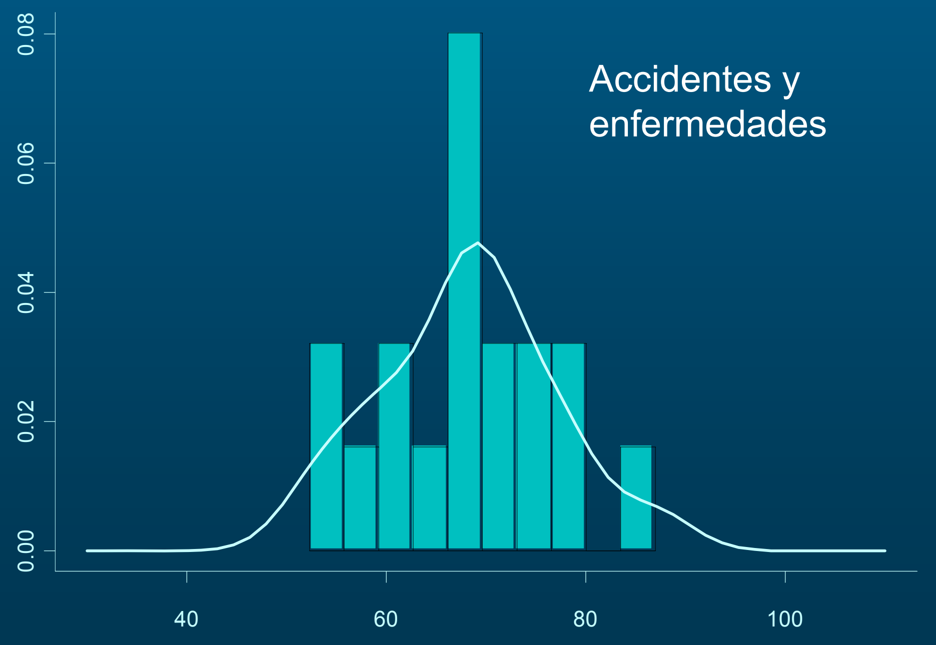
- Aproximar la densidad de W con mezclas de normales

$$\boxed{f(w) = \tfrac{1}{T}\sum_{i=1}^{T} N(w \mid w_i, \sigma^2)}$$

- $w_1,\ldots,w_T$ valores observados; σ^2 se determina con un criterio de ajuste.

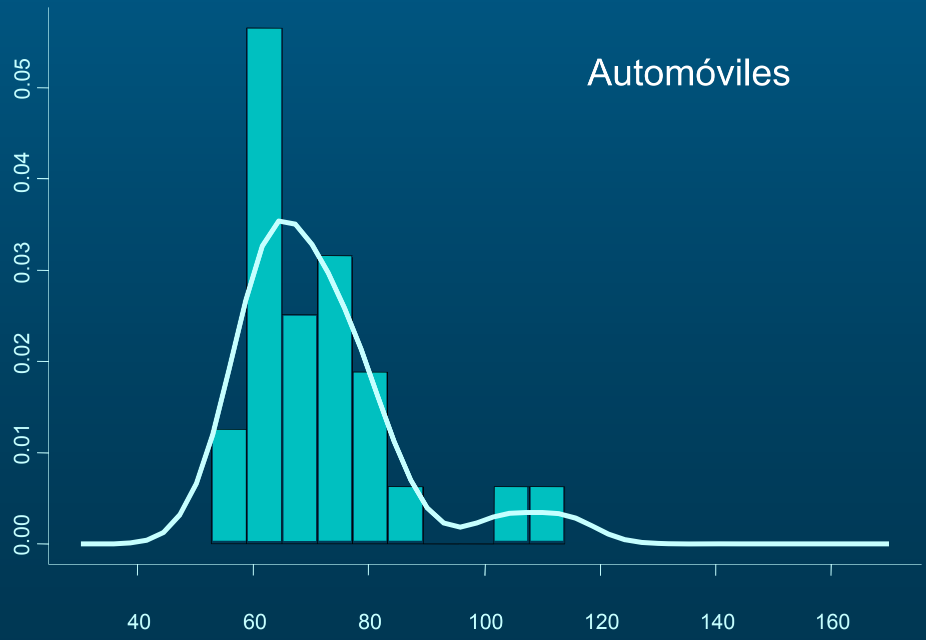
Formulación Estadística

Distribución para W



Formulación Estadística

Distribución para W



Formulación Estadística

Distribución para W

- La mezcla captura el patrón de asimetría
- No incorpora el efecto de estimación de los parámetros

Formulación Estadística

Modelo Condicional

- En cada periodo de tiempo: $Y_i = W_i \times X_i \qquad (i = 1, \ldots)$
  - X_i = primas emitidas
  - W_i = tasa de siniestralidad relativa
  - Y_i = siniestralidad

Formulación Estadística

Distribución para W

- $\{ X_1, X_2, \ldots \}$ una serie de observaciones correlacionadas
- $\{ W_1, W_2, \ldots \}$ realizaciones i.i.d. de un modelo $P(W \mid \phi)$
- $\{ Y_1, Y_2, \ldots \}$ debe su aleatoriedad a $\{ X_i \}$, $\{ W_i \}$ y son correlacionadas como resultado de la correlación en $\{ X_i \}$

Formulación Estadística

Distribución para W

- La distribución de Y_{T+1}, dado $X_{T+1} = x$, está totalmente determinada por la de W_{T+1}.

$$Y_{T+1} = W_{T+1} \times x$$

- Basta entonces
  - Asignar una distribución $P(W_{T+1} \mid \phi)$
  - Estimar el cuantil de interés para W
  - Estimar del cuantil condicional de Y_{T+1}

Alternativa Bayesiana

Características Generales

- Fundado sobre una base axiomática

Inferencia $\Leftrightarrow$ problema de decisión

- Proceso de aprendizaje basado en la fórmula de Bayes

Final $\propto$ Verosimilitud $\times$ Inicial

Alternativa Bayesiana

Características Generales

- Fórmula de Bayes

$$p(\theta \mid \text{datos}) \propto p(\text{datos} \mid \theta) \times p(\theta)$$

- Mecanismo general para la producción de pronósticos

Distribución Predictiva

$$p(X \mid \text{datos}) = \int p(X \mid \theta)\, p(\theta \mid \text{datos})\, d\theta$$

Alternativa Bayesiana

Análisis Predictivo

- Dado $X_{T+1} = x$, el comportamiento de Y_{T+1}, está determinado por W_{T+1} :

$$Y_{T+1} = W_{T+1} \times X_{T+1}$$

- Algoritmo, si se adopta un modelo $P(W_{T+1} \mid \phi)$
  - Asignar una inicial para ϕ y combinarla con la información histórica para obtener la final para ϕ,
  - Determinar la distribución predictiva para W_{T+1}
  - Calcular el cuantil predictivo condicional para Y_{T+1}, dado X_{T+1}

Alternativa Bayesiana

Distribución para W

- La selección del modelo P(W | ϕ) constituye un reto interesante.
- La Normal no es, en general, una buena elección.

  (Asimetría, valores positivos)
- Una mezcla de Normales también presenta inconvenientes.

Alternativa Bayesiana

Distribución para W

- Una posibilidad es suponer que una transformación de W es Normal.

$$U = \frac{W^{\lambda} - 1}{\lambda} \qquad \text{(Transformación de Box-Cox)}$$

- La introducción de λ da lugar a otro problema de decisión (la selección de un valor concreto para λ).

Alternativa Bayesiana

Distribución para W

- Con propósitos de ilustración

$$\lambda \to 0 \qquad \Rightarrow \qquad U = \ln(W)$$

$$Y_i = W_i \times X_i \qquad \Rightarrow \qquad \ln(Y_i) = \ln(W_i) + \ln(X_i)$$

$$V_i = U_i + c_i$$

$$V_i = \ln(Y_i)\,, \quad U_i = \ln(W_i)\,, \quad c_i = \ln(X_i)$$

Alternativa Bayesiana

Distribución para W

- En particular, $V_{T+1} = U_{T+1} + c_{T+1}$

- Dado un valor fijo de X_{T+1} (c_{T+1} constante),

predictiva para U_{T+1} $\Leftrightarrow$ predictiva para V_{T+1}

$\Leftrightarrow$ predictiva para Y_{T+1}

Alternativa Bayesiana

- El problema en estos términos es muy simple
  - Datos: $D = \{ U_1, U_2, ..., U_T \}$ i.i.d. $N(\mu,\tau)$ $(\tau^{-1} = \sigma^2)$
  - $\phi = (\mu, \tau)$; Distribución inicial de referencia: $P(\mu, \tau) \propto \tau^{-1}$

$\Rightarrow$

Distribución Final: $P(\mu, \tau \mid D)$ = Normal - Gamma

Distribución Predictiva Final: $P(U_{T+1} \mid D)$ = Student

Alternativa Bayesiana

Análisis Predictivo

- Predictiva Final: $P(\ U_{T+1} \mid D\ ) = \text{Student}\ ,\qquad U_{T+1} = \ln(W_{T+1})$

$$V_{T+1} = U_{T+1} + c_{T+1},\qquad c_{T+1} = \ln(X_{T+1})$$

$\Rightarrow$ (dado X_{T+1})

- Predictiva Final: $P(\ V_{T+1} \mid D\ ) = \text{Student}\ ,\qquad V_{T+1} = \ln(Y_{T+1})$
- Predictiva Final: $P(\ Y_{T+1} \mid D\ ) = \log - \text{Student}$

Alternativa Bayesiana

Análisis Predictivo

- El cuantil de orden (1-α) de la distribución predictiva de Y_{T+1}, dado X_{T+1}, resulta:

$$Y_{T+1}^{(1-\alpha)} = X_{T+1} \left(\prod_{i=1}^{T} w_i^{1/T} \right) \exp\left(t_{T-1}^{(1-\alpha)} [(1+1/T)\ \tilde{S}_U^2\ ]^{1/2} \right)$$

donde

$t_{T-1}^{(1-\alpha)}$ es el cuantil de una Student con T-1 g. de l. y $\tilde{S}_U^2 = \frac{1}{T-1} \sum (u_i - \bar{u})^2$

Alternativa Bayesiana

El supuesto de independencia

- Una alternativa al supuesto de independencia es un modelo auto regresivo estacionario de primer orden:

Para i = 1:

$$U_1 \sim N(\mu, \tau)$$

mientras que para i = 2,..., T+1:

$$U_i \mid U_{i-1} \sim N(\mu + \rho(U_{i-1} - \mu), \tau / (1-\rho^2))$$

Alternativa Bayesiana

El supuesto de independencia

- El modelo es estacionario

$$U_i \sim N(\mu, \tau) \quad \text{para} \quad i = 1, 2, ..., T+1$$

- La distribución conjunta de U_1, U_2, ..., U_{T+1} está determinada por los parámetros μ, τ y ρ. El modelo de independencia se recupera con $\rho = 0$.

$$U_i \mid U_{i-1} \sim N(\mu + \rho(U_{i-1} - \mu), \tau / (1 - \rho^2))$$

Alternativa Bayesiana

Modelo correlación común

- El modelo requiere la especificación de una distribución inicial para los tres parámetros μ, τ y ρ.

- A partir del algoritmo propuesto por Berger & Bernardo (1992) se determinó la distribución inicial de referencia para la parametrización ordenada $\{\rho\}$, $\{\tau\}$, $\{\mu\}$ (invariante ante permutaciones).

$$P(\mu, \tau, \rho) \propto \tau^{-1} \frac{(1+\rho^2)^{1/2}}{(1-\rho^2)}$$

Alternativa Bayesiana

Modelo correlación común

- La distribución final de μ, τ y ρ no tiene una expresión analítica completa.

- El efecto se reproduce para la distribución predictiva de U_{T+1}.

$\Rightarrow$

Simulación (Gibbs Sampler y Metropolis -Hastings)

Alternativa Bayesiana

Modelo correlaciones diferentes

- El modelo se puede generalizar

Para $i = 1$:

$$U_1 \sim N(\mu, \tau)$$

mientras que para $i = 2, \ldots, T+1$:

$$U_i \mid U_{i-1} \sim N(\mu + \rho_{i-1}(U_{i-1} - \mu), \tau / (1 - \rho_{i-1}^2))$$

Alternativa Bayesiana

Modelo correlaciones diferentes

- El modelo, para las T+1 observaciones, involucra a los parámetros μ, τ, ρ_1, ρ_2, ..., ρ_T. Para eliminar problemas de estimabilidad se introduce una estructura parcialmente jerárquica

$$P(\mu, \tau, \rho_1, \ldots, \rho_T \mid \theta) \propto \tau^{-1} \prod_{i=1}^{T} P(\rho_i \mid \theta)$$

en donde

$$P(\rho_i \mid \theta) = \frac{\theta}{2^{\theta}} (1+\rho_i)^{\theta-1} \; ; \quad -1 < \rho_i < 1$$

y, por simplicidad, $P(\theta \mid \gamma) = \gamma \exp(-\gamma\theta); \quad \theta > 0$

Alternativa Bayesiana

Modelo correlaciones diferentes

- De nuevo, la distribución final para μ, τ, ρ_1, ..., ρ_T no tiene una expresión analítica completa.
- Asimismo, el efecto se reproduce para la distribución predictiva de U_{T+1}.

$\Rightarrow$

Simulación (Gibbs Sampler y Metropolis -Hastings)

Mendoza, M. y Nieto-Barajas, L. E. (2006). Bayesian Solvency analysis with auto correlated observations. *Applied Stochastic Models in Business and Industry*, **22**, 169-180.

Alternativa Bayesiana

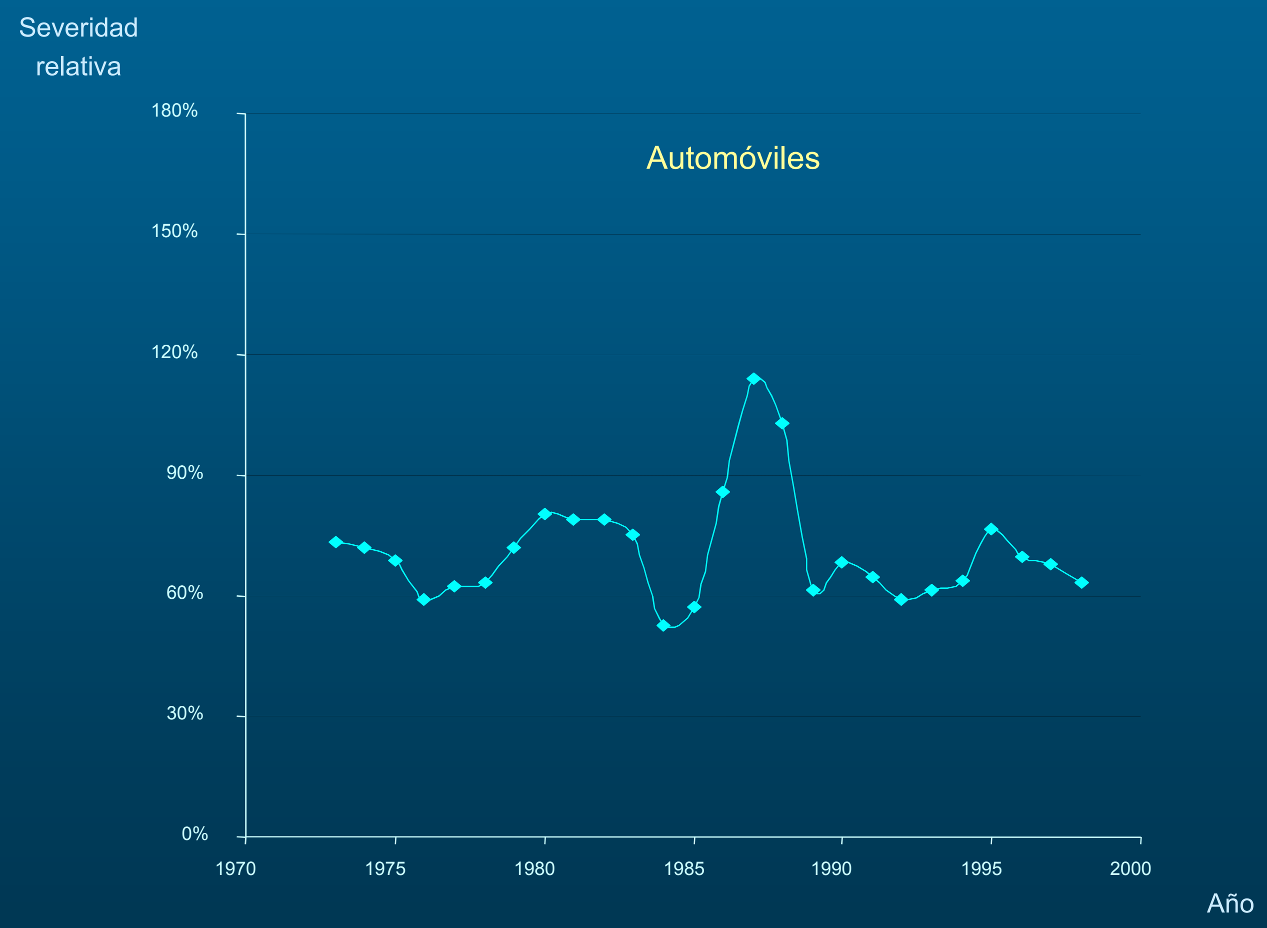
Propósito del modelado

- Una distribución más flexible y general que la Normal
- Posibilidad de incorporar patrones de dependencia
- Empleo de una herramienta de pronóstico general

$\Longrightarrow$

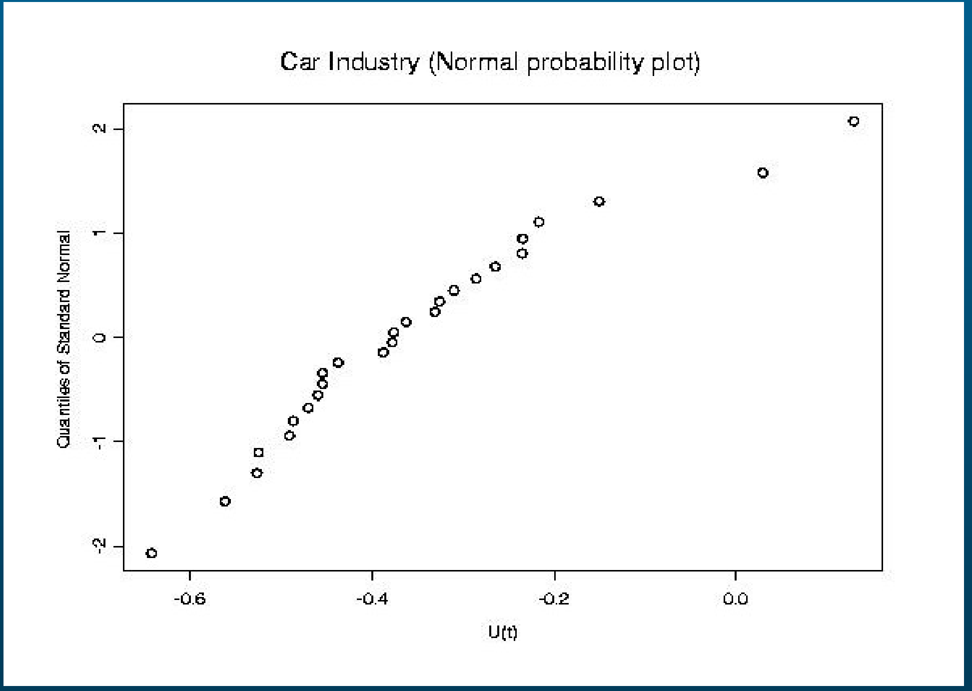
Procedimientos más robustos
para el cálculo de los factores

Ejemplos



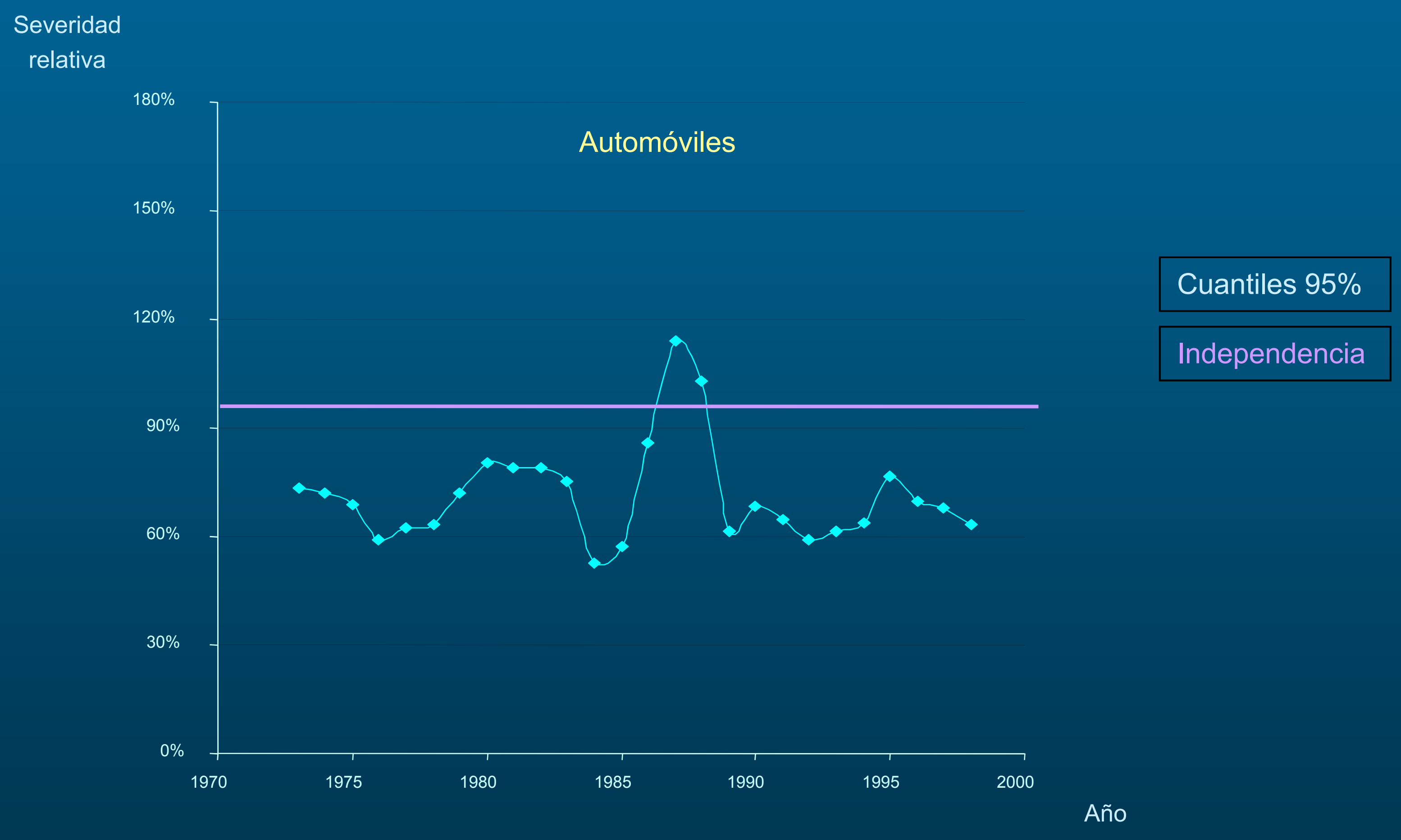
Severidad relativa
180%
150%
120%
90%
60%
30%
0%
Automóviles
1970
1975
1980
1985
1990
1995
2000
Año

Ejemplos



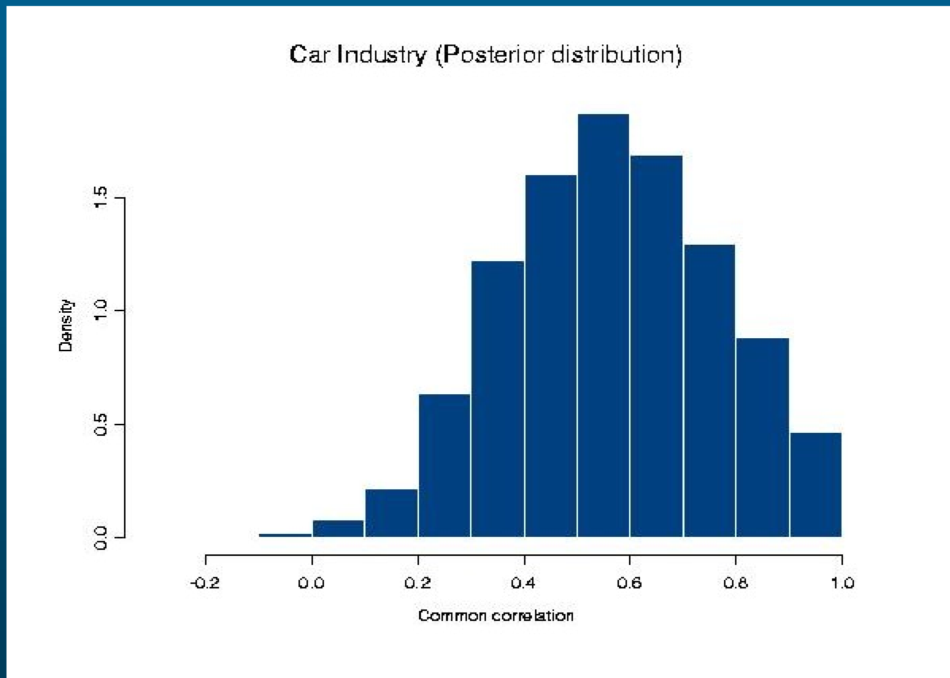
Car Industry (Normal probability plot)
Quantiles of Standard Normal
2
1
0
-1
-2
-0.6
-0.4
-0.2
0.0
U(t)

Ejemplos

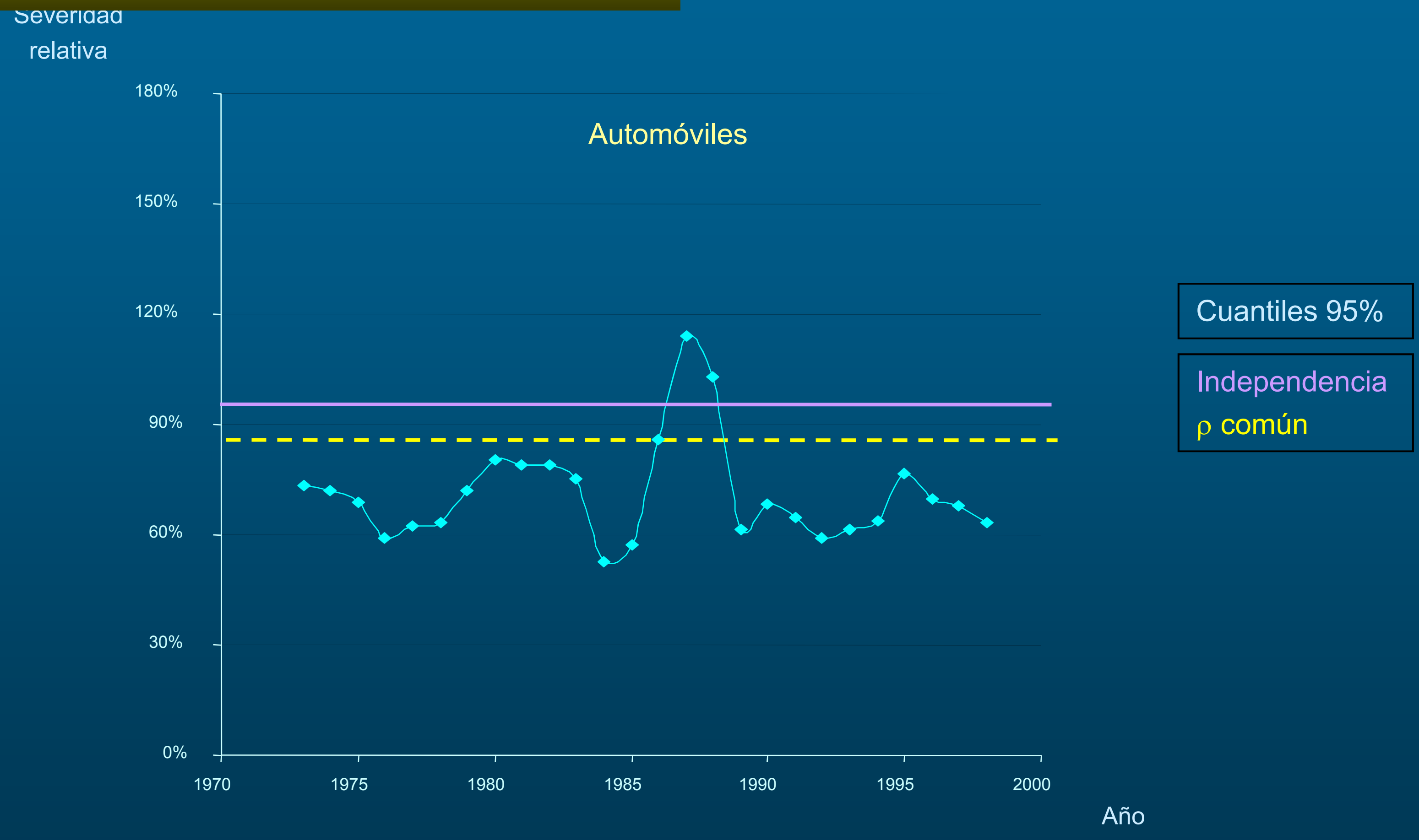


Severidad
relativa
180%
150%
120%
90%
60%
30%
0%
Automóviles
1970
1975
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia

Ejemplos

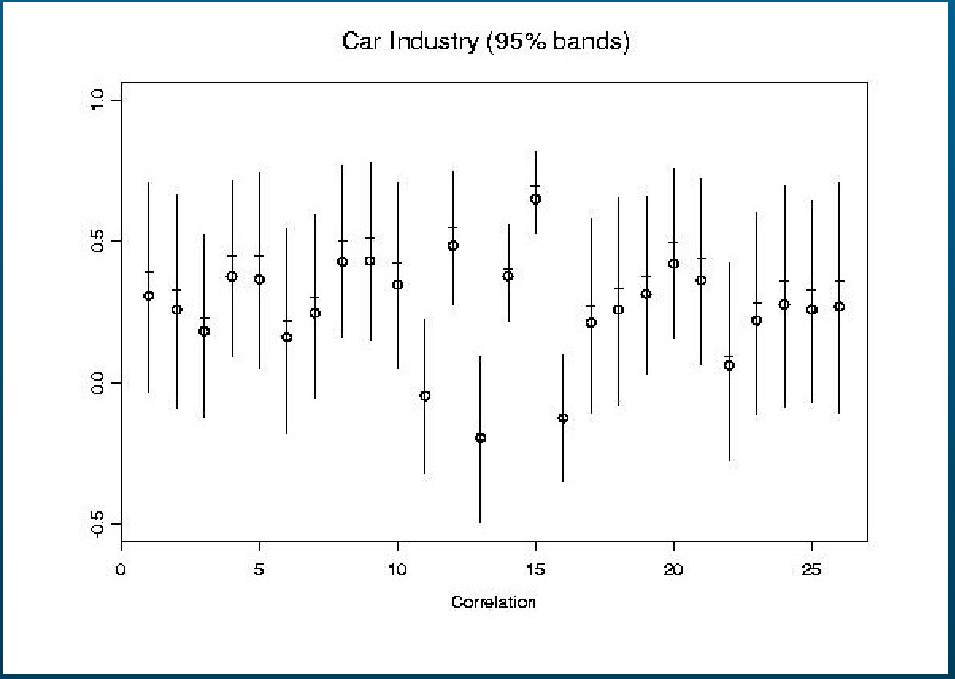

Car Industry (Posterior distribution)
Density
0.0
0.5
1.0
1.5
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
Common correlation

Ejemplos

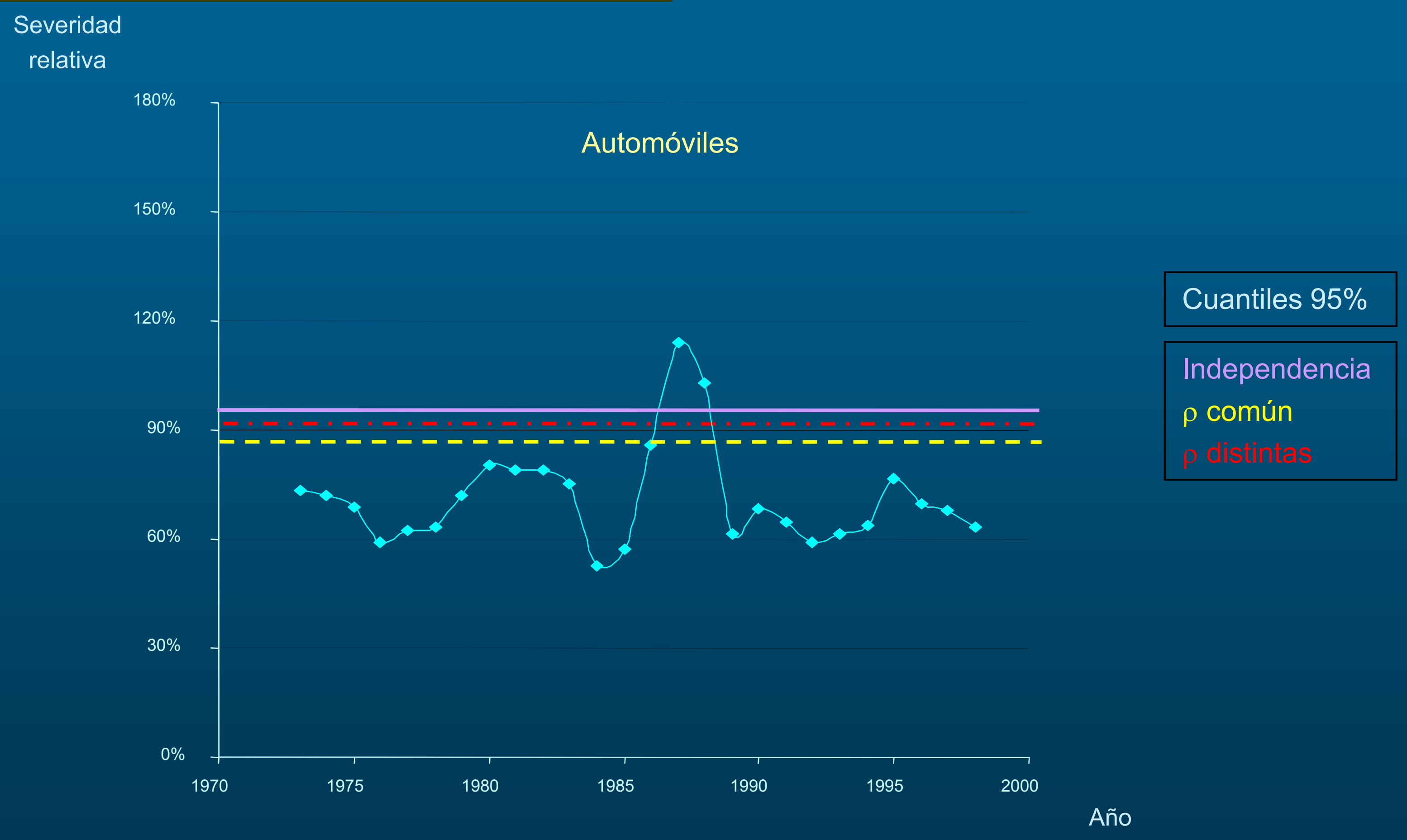


Severidad
relativa
180%
150%
120%
90%
60%
30%
0%
Automóviles
1970
1975
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia
ρ común

Ejemplos

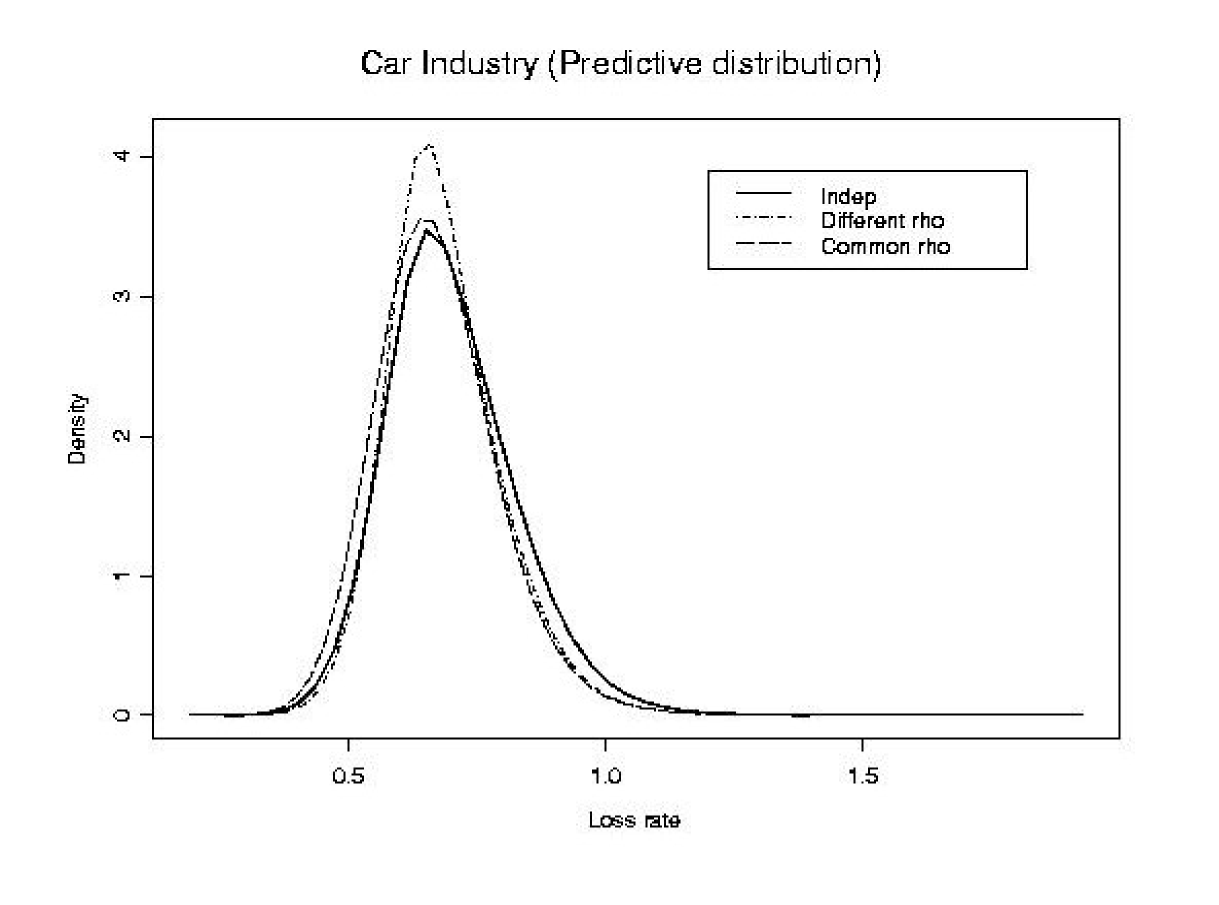


Ejemplos



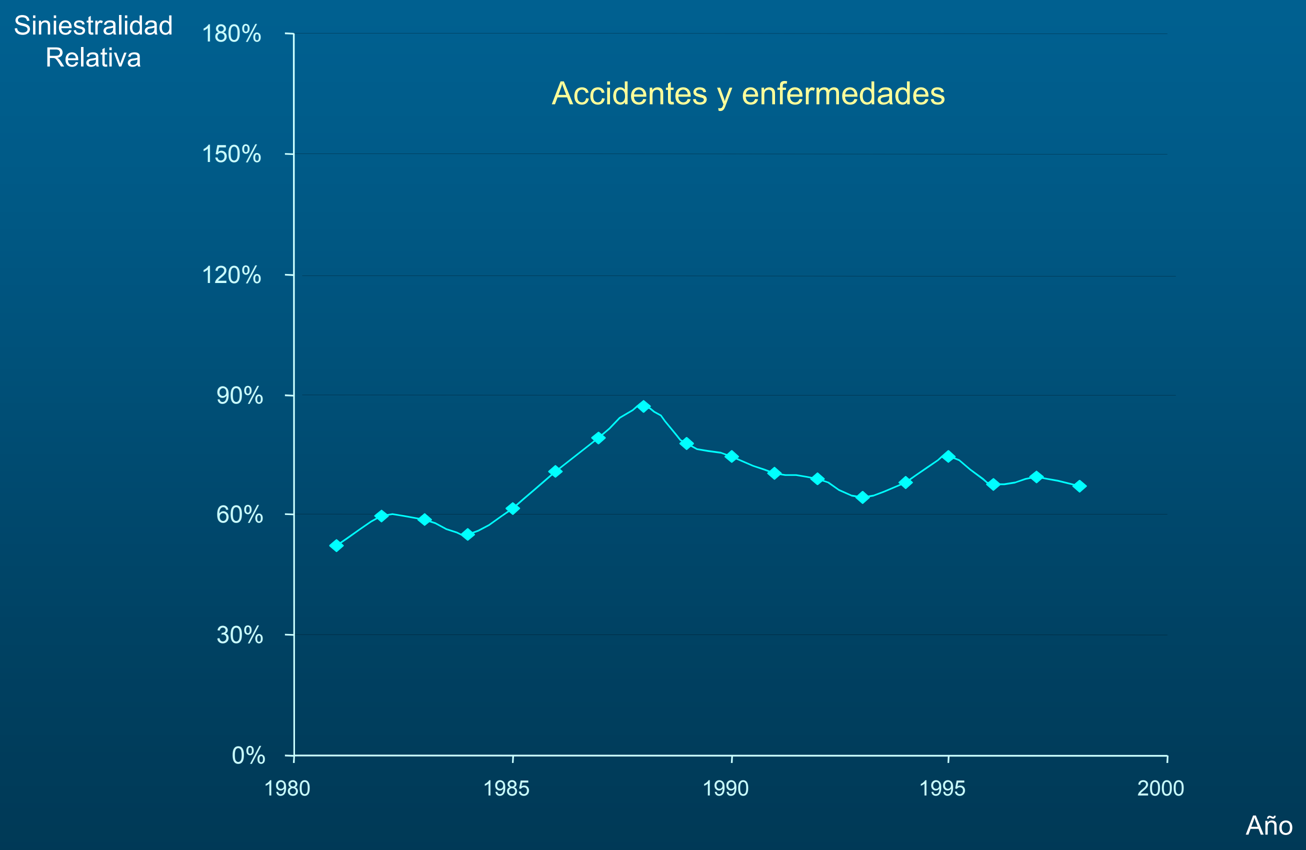
Severidad relativa
Automóviles
180%
150%
120%
90%
60%
30%
0%
1970
1975
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia
ρ común
ρ distintas

Ejemplos



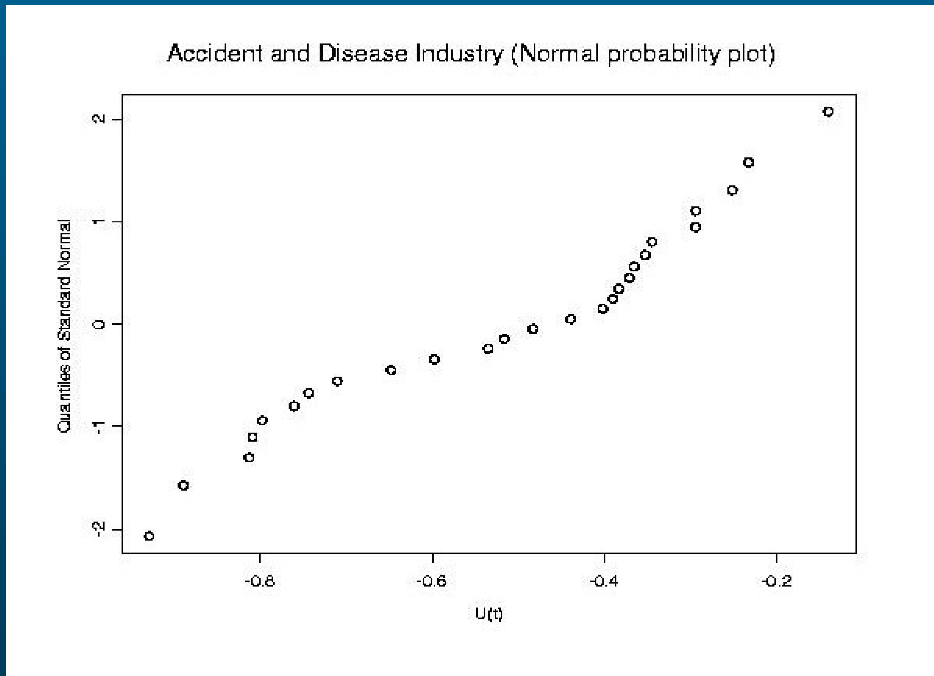
Car Industry (Predictive distribution)
Indep
Different rho
Common rho
Density
0
1
2
3
4
0.5
1.0
1.5
Loss rate

Ejemplos



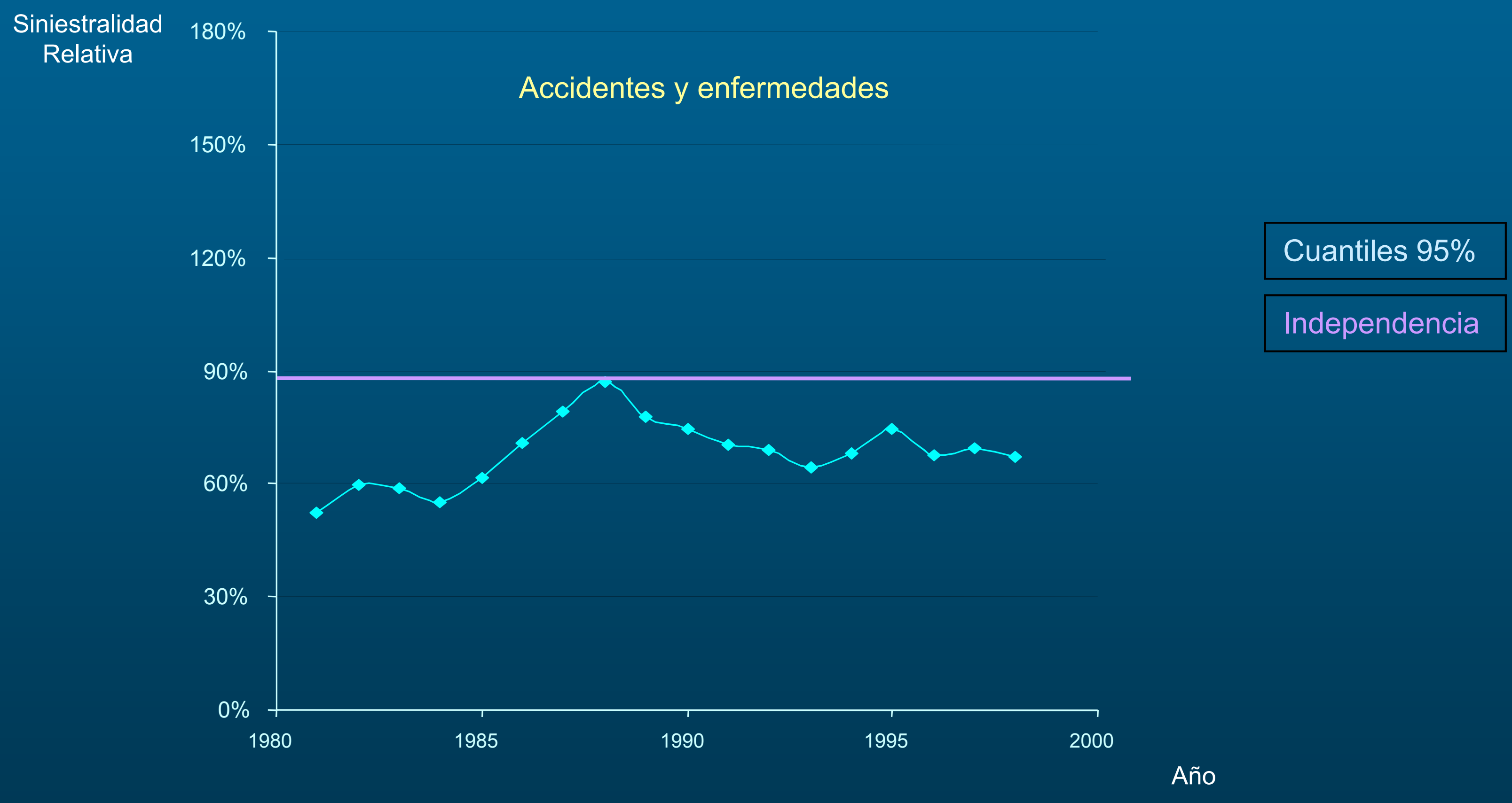
Siniestralidad
Relativa
Accidentes y enfermedades
180%
150%
120%
90%
60%
30%
0%
1980
1985
1990
1995
2000
Año

Ejemplos



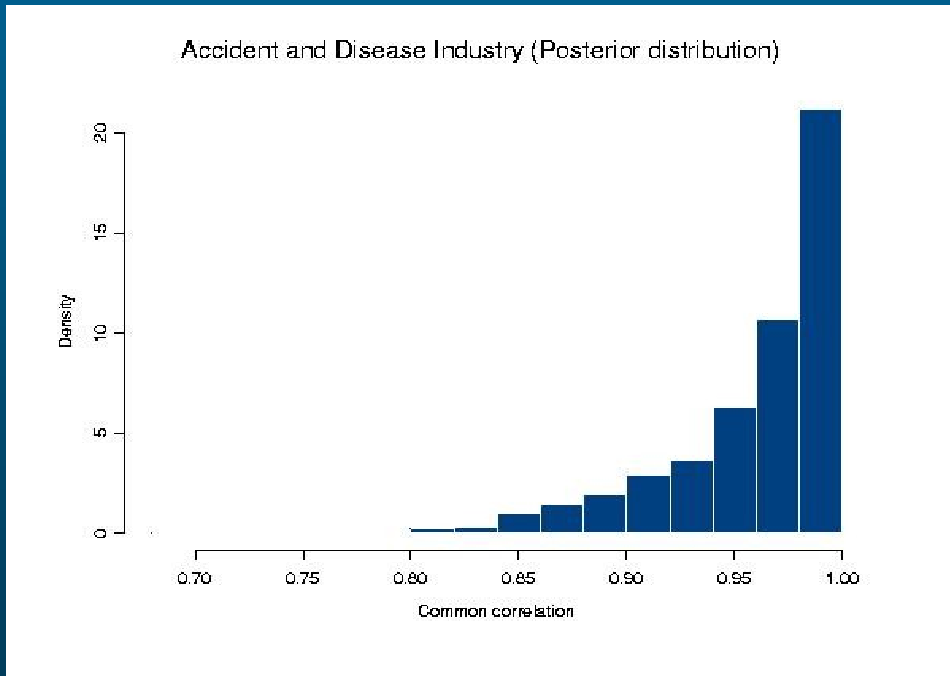
Accident and Disease Industry (Normal probability plot)
Quantiles of Standard Normal
-2
-1
0
1
2
-0.8
-0.6
-0.4
-0.2
U(t)

Ejemplos



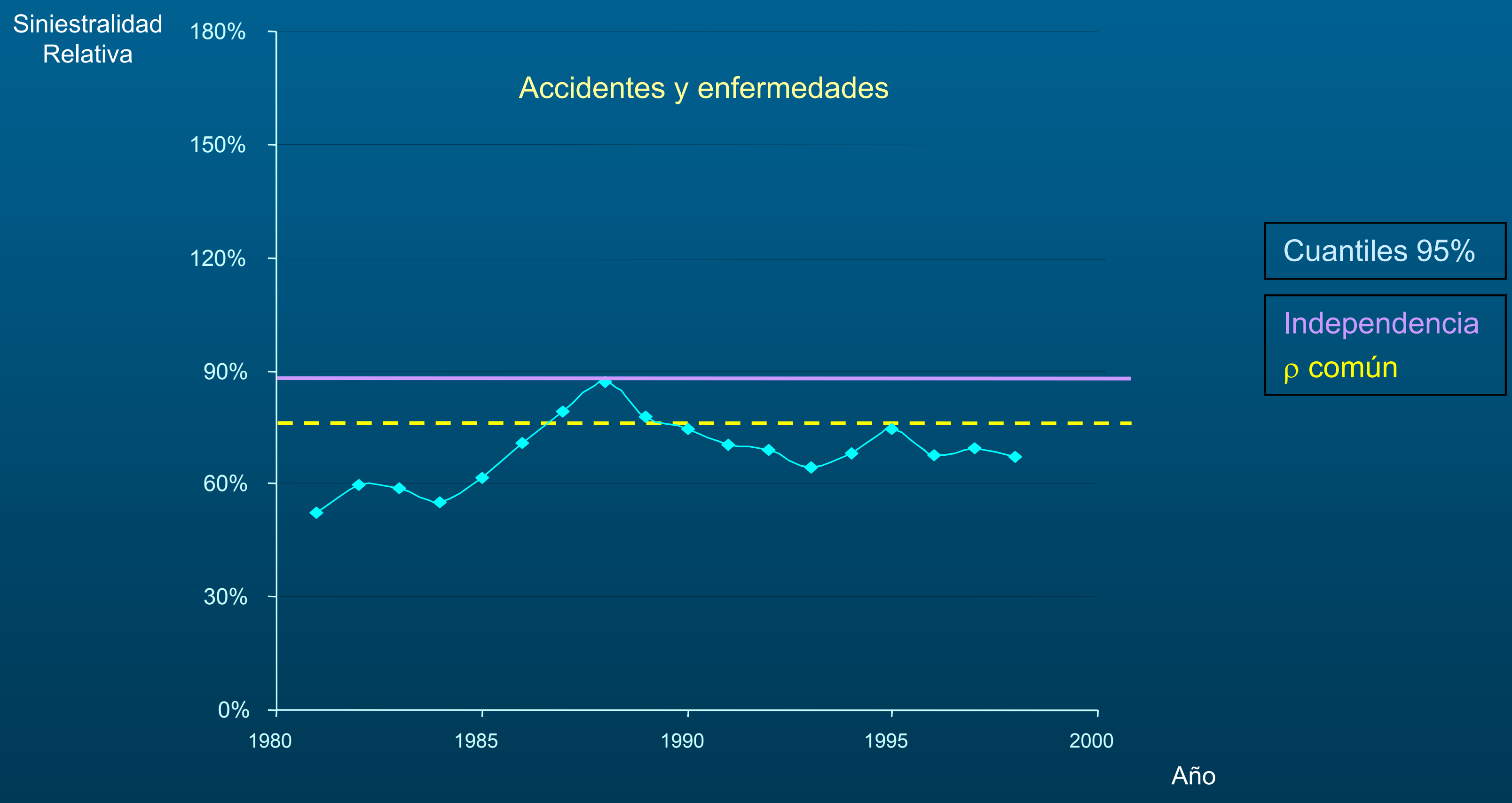
Siniestralidad Relativa
180%
150%
120%
90%
60%
30%
0%
Accidentes y enfermedades
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia

Ejemplos

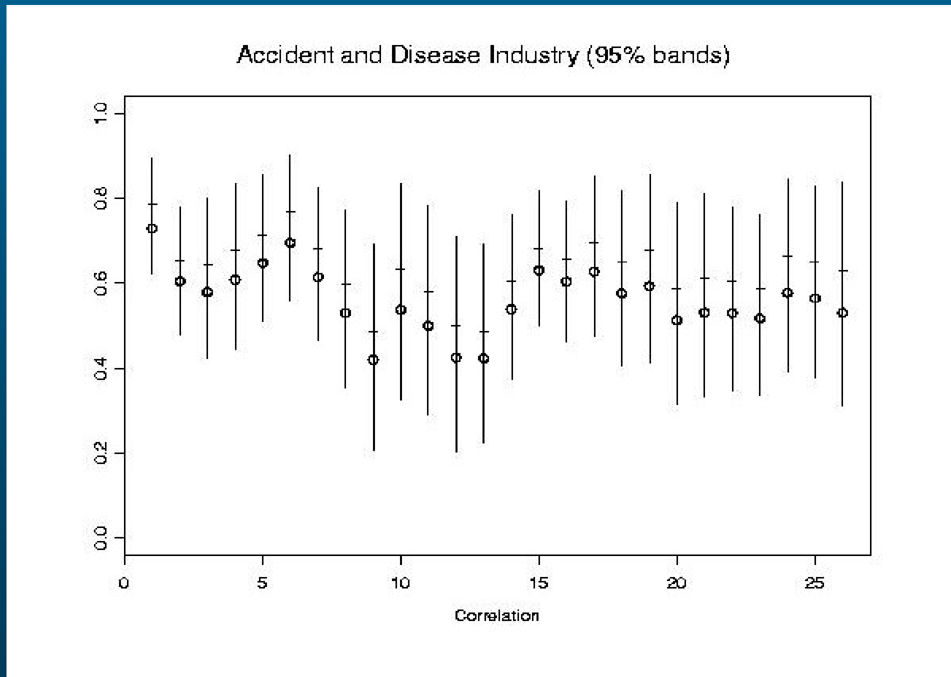


Accident and Disease Industry (Posterior distribution)
Density
0
5
10
15
20
0.70
0.75
0.80
0.85
0.90
0.95
1.00
Common correlation

Ejemplos

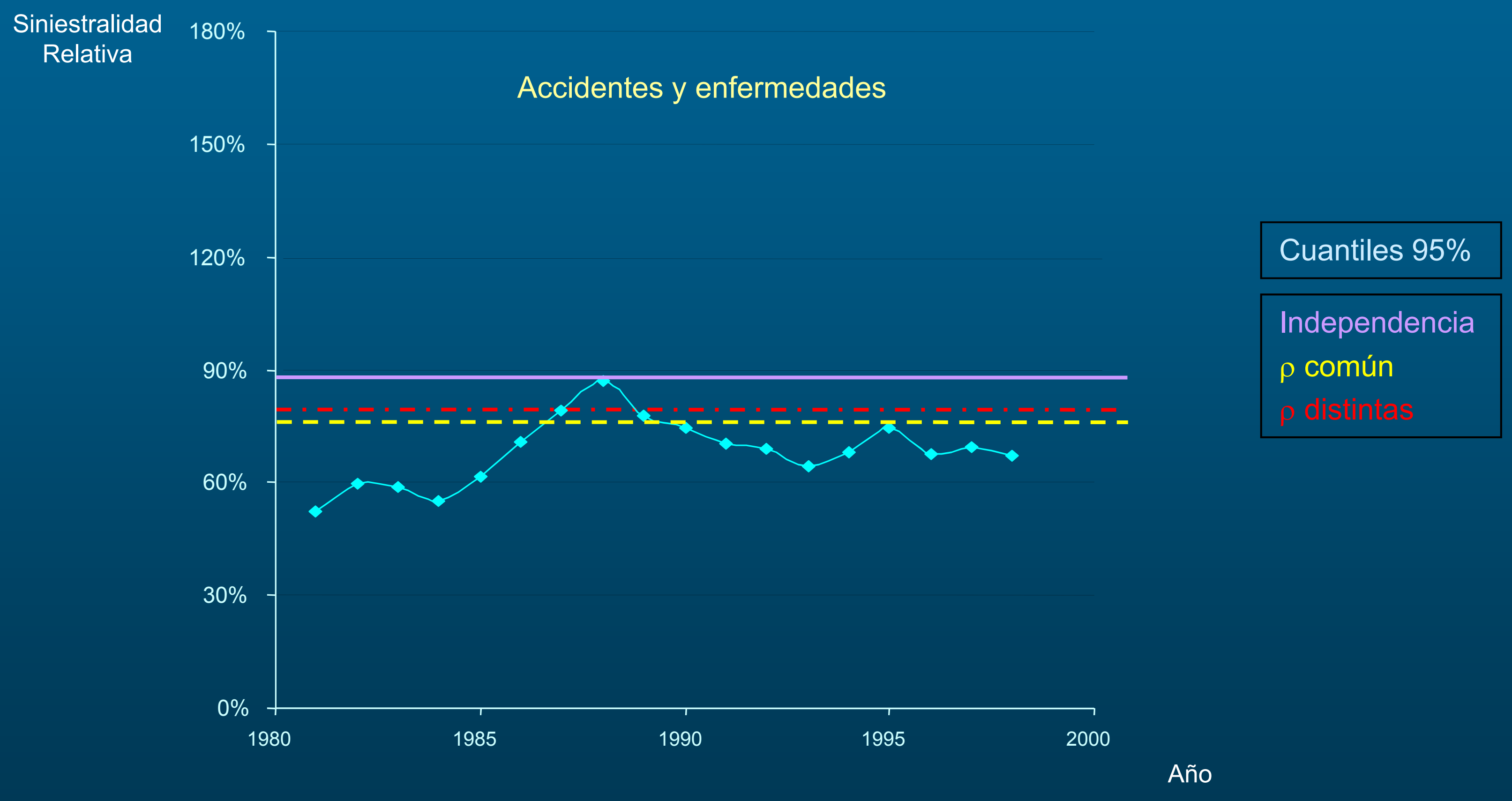

Siniestralidad Relativa
180%
150%
120%
90%
60%
30%
0%
Accidentes y enfermedades
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia
ρ común

Ejemplos



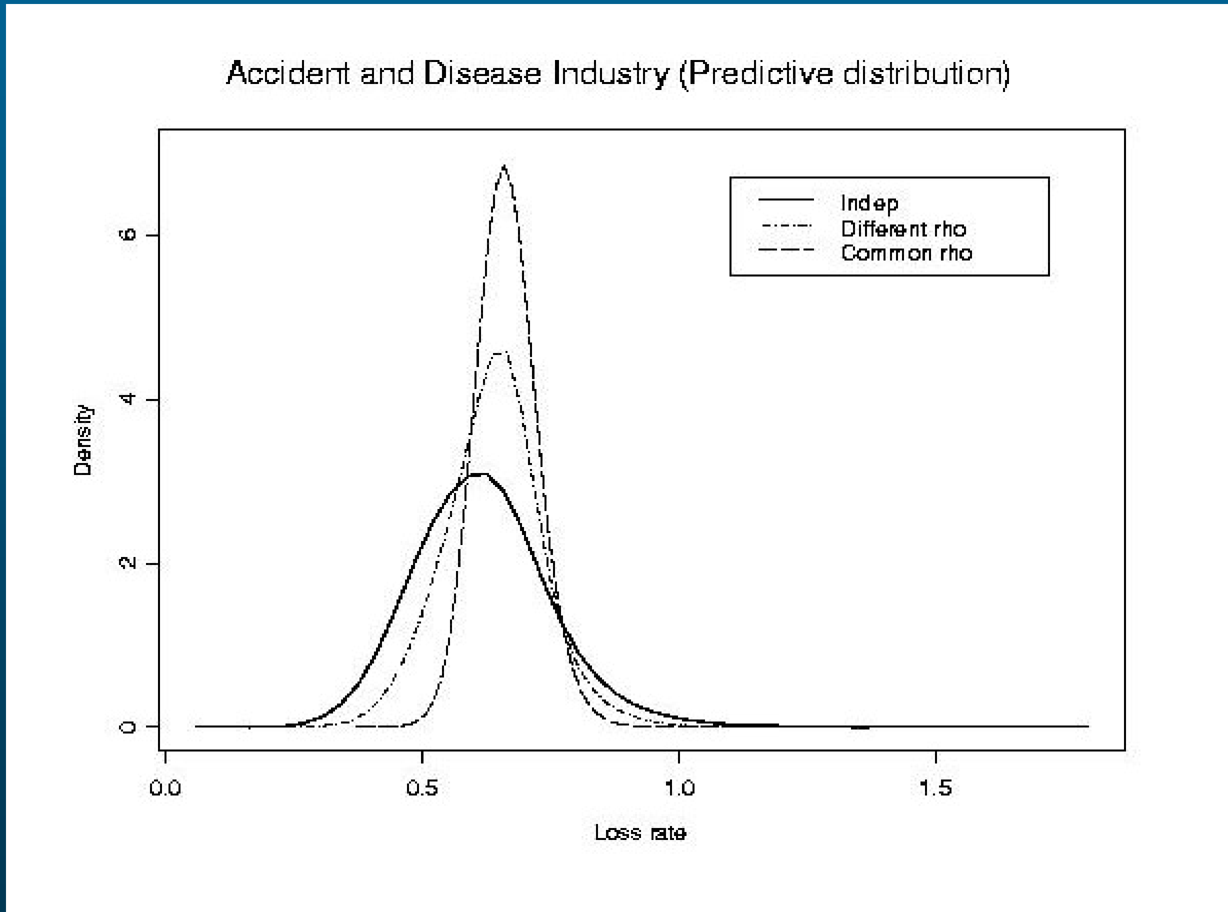
Accident and Disease Industry (95% bands)
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
Correlation

Ejemplos



Siniestralidad
Relativa
Accidentes y enfermedades
180%
150%
120%
90%
60%
30%
0%
1980
1985
1990
1995
2000
Año
Cuantiles 95%
Independencia
ρ común
ρ distintas

Ejemplos



Accident and Disease Industry (Predictive distribution)
Indep
Different rho
Common rho
Density
0
2
4
6
0.0
0.5
1.0
1.5
Loss rate

Consideraciones finales

- El valor del factor modifica si se cambia el modelo
- El modelo se debe seleccionar por su capacidad predictiva
- El análisis hace patente el riesgo de modelo
- Los modelos se deben evaluar periódicamente
- Es conveniente considerar modelos alternativos