Reclamos intercambiables en un proceso tipo Poisson Compuesto

Luis E. Nieto-Barajas

(conjunto con Ramses Mena)

Departmento de Estadística, ITAM, Mexico

Seminario Aleatorio, abril 2009

Contenido

- Introducción
- Proceso de reclamos intercambiables
- Sucesiones de variables intercambiables
  - Método paramétrico
  - Método no paramétrico
- Inferencia Bayesiana
- Ejemplo

Introducción

- En la teoría de riesgo colectivo uno de los principales objetos de estudio es la determinación de reservas para hacerle frente a los posibles reclamos futuros y prevenir la insolvencia de la compañía.

Introducción

- En la teoría de riesgo colectivo uno de los principales objetos de estudio es la determinación de reservas para hacerle frente a los posibles reclamos futuros y prevenir la insolvencia de la compañía.
- La herramienta básica ha sido modelar los reclamos agregados a través de un Proceso Poisson Compuesto (CPP).

Introducción

- En la teoría de riesgo colectivo uno de los principales objetos de estudio es la determinación de reservas para hacerle frente a los posibles reclamos futuros y prevenir la insolvencia de la compañía.
- La herramienta básica ha sido modelar los reclamos agregados a través de un Proceso Poisson Compuesto (CPP).
- Este proceso se define como:

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

donde $\{N_t; t \geq 0\}$ es un proceso Poisson homogéneo con intensidad $\lambda > 0$ y $Y_1, Y_2, \ldots$ una sucesión de v.a.i.i.d. (no negativas) con distribución F, independientes de $\{N_t\}$.

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.
- Propiedades del PPC:

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.
- Propiedades del PPC:
  1. $\mathrm{E}(X_t) = \lambda t \, \mathrm{E}(Y_i)$

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.
- Propiedades del PPC:
  1. $\mathrm{E}(X_t) = \lambda t \, \mathrm{E}(Y_i)$
  2. $\mathrm{Var}(X_t) = \lambda t \, \mathrm{E}(Y_i^2)$

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.
- Propiedades del PPC:
  1. $\mathrm{E}(X_t) = \lambda t\, \mathrm{E}(Y_i)$
  2. $\mathrm{Var}(X_t) = \lambda t\, \mathrm{E}(Y_i^2)$
  3. $\mathrm{Cov}(X_t, X_s) = \mathrm{Var}(X_{t \wedge s})$
     donde $t \wedge s := \min(s, t)$

Introducción

- En aplicaciones de seguros:
  - $\{N_t\}$ es el número de reclamos hechos a la compañía durante el intervalo de tiempo $(0, t]$,
  - Y_i es la magnitud del i-ésimo reclamo.
  - X_t son los reclamos agregados en el intervalo de tiempo $(0, t]$.
- Propiedades del PPC:
  1. $\mathrm{E}(X_t) = \lambda t \, \mathrm{E}(Y_i)$
  2. $\mathrm{Var}(X_t) = \lambda t \, \mathrm{E}(Y_i^2)$
  3. $\mathrm{Cov}(X_t, X_s) = \mathrm{Var}(X_{t \wedge s})$
     donde $t \wedge s := \min(s, t)$
  4. $\mathrm{Corr}(X_t, X_s) = (t \wedge s)/\sqrt{ts}$
     $\Rightarrow$ si $t < s$ entonces $\mathrm{Corr}(X_t, X_s) = \sqrt{t/s}$

Introducción

- El proceso de riesgo de un portafolio (Lundberg, 1926) consiste en compensar el PPC mediante:

$$Z_t = r\,t - X_t,$$

donde $r > 0$ denota la tasa de prima de riesgo.

Introducción

- El proceso de riesgo de un portafolio (Lundberg, 1926) consiste en compensar el PPC mediante:

$$Z_t = r\,t - X_t,$$

donde $r > 0$ denota la tasa de prima de riesgo.
  - En seguros Z_t es la ganacia del portafolio en el intervalo $(0, t]$

Introducción

- El proceso de riesgo de un portafolio (Lundberg, 1926) consiste en compensar el PPC mediante:

$$Z_t = r\,t - X_t,$$

  donde $r > 0$ denota la tasa de prima de riesgo.
  - En seguros Z_t es la ganacia del portafolio en el intervalo $(0, t]$
- Otro proceso de interés es el proceso de reserva definido como:

$$R_t^u := u + Z_t$$

  donde u es el capital inicial

Introducción

- Considerando la riqueza del portafolio, es de interés determinar si el proceso R_t^u cae por debajo de cero $\Rightarrow$ Probabilidad de ruina

$$\Psi(u) := \mathrm{P}\left(\inf_{t \geq 0}\{t; R_t^u < 0\} < \infty\right)$$

Introducción

- Considerando la riqueza del portafolio, es de interés determinar si el proceso R_t^u cae por debajo de cero $\Rightarrow$ Probabilidad de ruina

$$\Psi(u) := \mathrm{P}\left(\inf_{t\geq 0}\{t; R_t^u < 0\} < \infty\right)$$

- Alternativamente se puede considerar la probabilidad de ruina en horizote finito $\Psi(u, T) := \mathrm{P}\left(\inf_{0\leq t<T}\{t; R_t^u < 0\} < T\right)$.

Introducción

- Considerando la riqueza del portafolio, es de interés determinar si el proceso R_t^u cae por debajo de cero $\Rightarrow$ Probabilidad de ruina

$$\Psi(u) := \mathrm{P}\left(\inf_{t\geq 0}\{t; R_t^u < 0\} < \infty\right)$$

- Alternativamente se puede considerar la probabilidad de ruina en horizote finito $\Psi(u, T) := \mathrm{P}\left(\inf_{0\leq t<T}\{t; R_t^u < 0\} < T\right)$.
- Otra cantidad de interés es el factor de seguridad relativo ρ,

$$\rho := \frac{\mathrm{E}(Z_t)}{\mathrm{E}(X_t)}$$

Introducción

- Considerando la riqueza del portafolio, es de interés determinar si el proceso R_t^u cae por debajo de cero $\Rightarrow$ Probabilidad de ruina

$$\Psi(u) := \mathrm{P}\left(\inf_{t \geq 0}\{t; R_t^u < 0\} < \infty\right)$$

- Alternativamente se puede considerar la probabilidad de ruina en horizote finito $\Psi(u, T) := \mathrm{P}\left(\inf_{0 \leq t < T}\{t; R_t^u < 0\} < T\right)$.
- Otra cantidad de interés es el factor de seguridad relativo ρ,

$$\rho := \frac{\mathrm{E}(Z_t)}{\mathrm{E}(X_t)}$$

- $\Rightarrow$ Un portafolio es redituable si $\rho > 0$ y $\Psi(u)$ es pequeña.

Generalizaciones

- Daboni (1974):

$$N_t \sim \text{Proceso Poisson Mixto o de Cox}$$

$\Rightarrow X_t$ tiene incrementos intercambiables

Generalizaciones

- Daboni (1974):

$$N_t \sim \text{Proceso Poisson Mixto o de Cox}$$

$\Rightarrow X_t$ tiene incrementos intercambiables

- Morales y Shoutens (2003):

$$X_t \sim \text{Proceso de Lévy}$$

$\Rightarrow$ Incrementos y frecuencias dependientes del tiempo

Generalizaciones

- Daboni (1974):

$$N_t \sim \text{Proceso Poisson Mixto o de Cox}$$

$\Rightarrow X_t$ tiene incrementos intercambiables

- Morales y Shoutens (2003):

$$X_t \sim \text{Proceso de Lévy}$$

$\Rightarrow$ Incrementos y frecuencias dependientes del tiempo

- En otra línea, Rolski et al. (2001): ingreso por primas no lineal, i.e.,

$$r\,t \longrightarrow r(t).$$

Motivación

- Escenario común: la aseguradora considera la información agregada de todas las pólizas de un portafolios (sin importar si las reclamaciones fueron hechas por el mismo o distinto asegurado.

Motivación

- Escenario común: la aseguradora considera la información agregada de todas las pólizas de un portafolios (sin importar si las reclamaciones fueron hechas por el mismo o distinto asegurado.
  $\Rightarrow$ Independencia entre reclamos

Motivación

- Escenario común: la aseguradora considera la información agregada de todas las pólizas de un portafolios (sin importar si las reclamaciones fueron hechas por el mismo o distinto asegurado.
  $\Rightarrow$ Independencia entre reclamos
- En la actualidad la colectividad es más especializada a grupos más homogeneos y es posible que un reclamo hecho por un asegurado esté relacionado con un reclamo futuro del mismo asegurado o de otro asegurado del mismo portafolio.

Motivación

- Escenario común: la aseguradora considera la información agregada de todas las pólizas de un portafolios (sin importar si las reclamaciones fueron hechas por el mismo o distinto asegurado.
  $\Rightarrow$ Independencia entre reclamos
- En la actualidad la colectividad es más especializada a grupos más homogeneos y es posible que un reclamo hecho por un asegurado esté relacionado con un reclamo futuro del mismo asegurado o de otro asegurado del mismo portafolio.
  $\Rightarrow$ Dependencia entre reclamos

Motivación

- Supongamos que podemos identificar los reclamos hechos por el mismo individuo j, digamos $Y_{1j}, Y_{2j}, \ldots$

Motivación

- Supongamos que podemos identificar los reclamos hechos por el mismo individuo j, digamos $Y_{1j}, Y_{2j}, \ldots$
- El objetivo es modelar los patrones de reclamo del individuo j reconociendo una posible dependencia entre los reclamos de la misma póliza.

Motivación

- Supongamos que podemos identificar los reclamos hechos por el mismo individuo j, digamos $Y_{1j}, Y_{2j}, \ldots$
- El objetivo es modelar los patrones de reclamo del individuo j reconociendo una posible dependencia entre los reclamos de la misma póliza.
- ¿cómo?

Motivación

- Supongamos que podemos identificar los reclamos hechos por el mismo individuo j, digamos $Y_{1j}, Y_{2j}, \ldots$
- El objetivo es modelar los patrones de reclamo del individuo j reconociendo una posible dependencia entre los reclamos de la misma póliza.
- ¿cómo? consideraremos un escenario similar a los modelos mixtos endonde un efecto aleatorio modela la heterogeneidad de los individuous e introduce una dependencia en las reclamaciones.

Motivación

- Supongamos que podemos identificar los reclamos hechos por el mismo individuo j, digamos $Y_{1j}, Y_{2j}, \ldots$
- El objetivo es modelar los patrones de reclamo del individuo j reconociendo una posible dependencia entre los reclamos de la misma póliza.
- ¿cómo? consideraremos un escenario similar a los modelos mixtos endonde un efecto aleatorio modela la heterogeneidad de los individuous e introduce una dependencia en las reclamaciones.
- Efecto aleatorio $\Rightarrow$ Reclamos intercambiables

Motivación

- Requerimos información por póliza o individuo (datos tipo panel) que nos da lugar a distintas trayectorias o realizaciones

$$X_{t1}, X_{t2}, \ldots, X_{tm}$$

Motivación

- Requerimos información por póliza o individuo (datos tipo panel) que nos da lugar a distintas trayectorias o realizaciones

$$X_{t1}, X_{t2}, \ldots, X_{tm}$$

- Es posible determinar la reserva promedio por individuo, así como la la reserva agregada del portafolio

$$X_t^a := \sum_{j=1}^{m} X_{tj}.$$

Modelo

- Proceso con incrementos intercambiables (ECP):

$$X_t := \sum_{i=1}^{N_t} Y_i \text{ donde}$$

  - $\{N_t; t \geq 0\}$ es un p. Poisson homogéneo con intensidad $\lambda > 0$
  - $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables no negativas con distribución marginal común F, indep. de $\{N_t\}$.

Modelo

- Proceso con incrementos intercambiables (ECP):

$$X_t := \sum_{i=1}^{N_t} Y_i \text{ donde}$$

  - $\{N_t; t \geq 0\}$ es un p. Poisson homogéneo con intensidad $\lambda > 0$
  - $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables no negativas con distribución marginal común F, indep. de $\{N_t\}$.

- Intercambiabilidad (de Finetti, 1937): $Y_1, Y_2, \ldots$ son intercambiables si

$$(Y_1, \ldots, Y_n) \stackrel{d}{=} (Y_{\sigma(1)}, \ldots, Y_{\sigma(n)}),$$

  $\forall$ permutación σ de $\{1, \ldots, n\}$ y para $n \geq 1$. $\Rightarrow \exists$ una variable/medida G t.q. $\{Y_i\}$, $i = 1, 2, \ldots$ son cond. indep. dado G, y G es una variable/medida con ley descrita por $\mathcal{P}$.

Modelo

- Propiedades:

Modelo

- Propiedades:
  1. X_t es un proceso de Markov con incrementos intercambiables

Modelo

- Propiedades:
  1. X_t es un proceso de Markov con incrementos intercambiables
  2. $\mathrm{E}(X_t) = \lambda t \mathrm{E}(Y_i)$

Modelo

- Propiedades:
  1. X_t es un proceso de Markov con incrementos intercambiables
  2. $\mathrm{E}(X_t) = \lambda t \mathrm{E}(Y_i)$
  3. $\mathrm{Var}(X_t) = \lambda t \mathrm{E}(Y_i^2) + \lambda^2 t^2 \mathrm{Cov}(Y_i, Y_j)$

Modelo

- Propiedades:
  1. X_t es un proceso de Markov con incrementos intercambiables
  2. $\text{E}(X_t) = \lambda t \text{E}(Y_i)$
  3. $\text{Var}(X_t) = \lambda t \text{E}(Y_i^2) + \lambda^2 t^2 \text{Cov}(Y_i, Y_j)$
  4. $\text{Cov}(X_t, X_s) = \lambda\,(t \wedge s)\,\text{E}[Y_i^2] + \lambda^2\, ts\,\text{Cov}(Y_i, Y_j)$

Modelo

- Propiedades:
  1. X_t es un proceso de Markov con incrementos intercambiables
  2. $\mathrm{E}(X_t) = \lambda t \mathrm{E}(Y_i)$
  3. $\mathrm{Var}(X_t) = \lambda t \mathrm{E}(Y_i^2) + \lambda^2 t^2 \mathrm{Cov}(Y_i, Y_j)$
  4. $\mathrm{Cov}(X_t, X_s) = \lambda\,(t \wedge s)\,\mathrm{E}[Y_i^2] + \lambda^2\, ts\,\mathrm{Cov}(Y_i, Y_j)$
- Pregunta: ¿Cómo obtener una sucesión de v.a. intercambiables $Y_1, Y_2, \ldots$ tal que cada Y_i tenga la misma distribución marginal $F(y)$ (ó $f(y)$)?

Sucesiones intercambiables

- Caso paramétrico:
  Sea Z una v.a. latente con distribución arbitraria $f(z \mid y)$
  $\Rightarrow$ definimos $f(y \mid z)$ a través del Teo. de Bayes

$$f(y \mid z) = \frac{f(z \mid y)f(y)}{f(z)}$$

  con

$$f(z) = \int f(z \mid y)f(y)\mathrm{d}\mu_1(y),$$

  donde $\mu_1(y)$ representa una medida de referencia de conteo si Y es discreta o la medida de Lebesgue si Y is continua.
  $\Rightarrow$ si marginalizamos sobre Z,

$$\int f(y \mid z)f(z)\mathrm{d}\mu_2(z) = f(y),$$

  donde $\mu_2(\cdot)$ es otra medida de referencia que actua sobre Z.

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid Z \sim f(y \mid z)$, como arriba, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado $Z = z$, con distribución marginal para Z, como arriba, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con densidad marginal $f(y)$.

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid Z \sim f(y \mid z)$, como arriba, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado $Z = z$, con distribución marginal para Z, como arriba, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con densidad marginal $f(y)$.
- Existen muchas posibilidades para $f(z \mid y)$: discreta, continua, univariada o multivariada

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid Z \sim f(y \mid z)$, como arriba, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado $Z = z$, con distribución marginal para Z, como arriba, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con densidad marginal $f(y)$.
- Existen muchas posibilidades para $f(z \mid y)$: discreta, continua, univariada o multivariada
- Distintas características en Z darán lugar a distintas formas de dependencia

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid Z \sim f(y \mid z)$, como arriba, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado $Z = z$, con distribución marginal para Z, como arriba, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con densidad marginal $f(y)$.
- Existen muchas posibilidades para $f(z \mid y)$: discreta, continua, univariada o multivariada
- Distintas características en Z darán lugar a distintas formas de dependencia
- Expresiones cerradas para $f(y \mid z)$ y $f(z)$ se obtienen cuando (y, z) son un par con distribuciones conjugadas en estadística Bayesiana

Ejemplo 1

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$

Ejemplo 1

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \mathrm{Ga}(a, b)$
- Proponemos $f(z \mid y) = \mathrm{Ga}(z \mid c, y)$, $c > 0$

Ejemplo 1

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $f(z \mid y) = \text{Ga}(z \mid c, y)$, $c > 0$
- $\Rightarrow f(y \mid z) = \text{Ga}(y \mid a + c, b + z)$ y $f(z) = \text{Gga}(z \mid a, b, c)$.

Ejemplo 1

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $f(z \mid y) = \text{Ga}(z \mid c, y)$, $c > 0$
- $\Rightarrow f(y \mid z) = \text{Ga}(y \mid a + c, b + z)$ y $f(z) = \text{Gga}(z \mid a, b, c)$.
- $\therefore\ Y_i \mid Z \sim \text{Ga}(a + c, b + z)$, $i = 1, 2, \ldots$ cond. indep. dado Z y $Z \sim \text{Gga}(a, b, c) \Rightarrow Y_i \sim \text{Ga}(a, b)$ con $\text{Corr}(Y_i, Y_j) = c/(a + c + 1)$, $\forall\ i \neq j$.

Ejemplo 1

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $f(z \mid y) = \text{Ga}(z \mid c, y)$, $c > 0$
- $\Rightarrow f(y \mid z) = \text{Ga}(y \mid a + c, b + z)$ y $f(z) = \text{Gga}(z \mid a, b, c)$.
- $\therefore Y_i \mid Z \sim \text{Ga}(a + c, b + z)$, $i = 1, 2, \ldots$ cond. indep. dado Z y $Z \sim \text{Gga}(a, b, c) \Rightarrow Y_i \sim \text{Ga}(a, b)$ con $\text{Corr}(Y_i, Y_j) = c/(a + c + 1)$, $\forall\ i \neq j$.
- Consideramos dos procesos: X_t^I y X_t^E y definimos el proceso de reserva R_t^u, correspondiente, con $r = \lambda = a = c = 1$ y $b = 1.1 \Rightarrow$ factor de seguridad relativo $\rho = (r\,b)/(\lambda\,a) - 1 = 0.1$.

Ejemplo 1

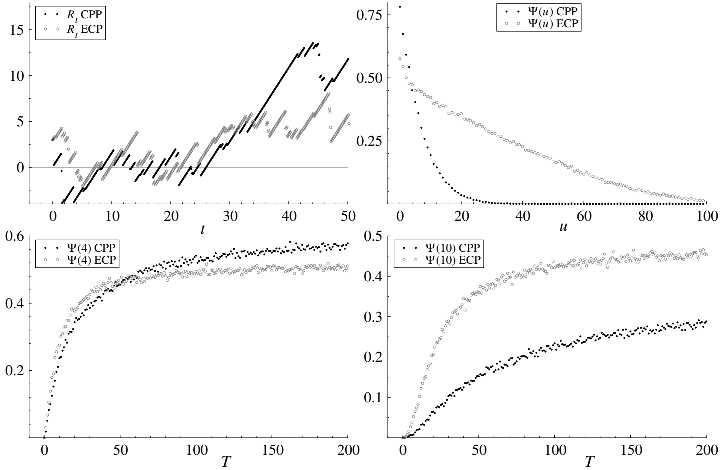

Figure: Procesos de reserva y estimadores Monte Carlo de la prob. de ruina para los modelos CPP y $\mathrm{ECP}_{\mathrm{p}}$.

Sucesiones intercambiables

- Caso no paramétrico: v. latente $Z \rightarrow$ medida latente G
  Sea G una medida a. latente con distribución arbitraria
  $G \mid Y \sim \mathcal{P}(\cdot \mid y)$
  $\Rightarrow$ En este caso
  $$Y \mid G \sim G$$
  con
  $$G \sim \mathcal{P}$$
  $\Rightarrow$ $\mathcal{P}$ debe de satisfacer que
  $$\mathrm{E}_{\mathcal{P}}\{G(y)\} = \int G(y)\mathcal{P}(\mathrm{d}G) = F(y),$$
  i.e., $F(y)$ es la media inicial de una medida aleatoria G

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid G \sim G$, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado G, con $G \sim \mathcal{P}$, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con distribución marginal $F(y)$.

Sucesiones intercambiables

- Entonces, si tomamos $Y_i \mid G \sim G$, para $i = 1, 2, \ldots$ una sucesión de v.a. cond. indep. dado G, con $G \sim \mathcal{P}$, entonces $Y_1, Y_2, \ldots$ es una sucesión de v.a. intercambiables con distribución marginal $F(y)$.
- La propiedad $\mathrm{E}(G) = F$ es una característica de muchas de las distribuciones iniciales no paramétricas $\mathcal{P}$, e.g., proceso Dirichlet, medidas aleatorias normalizadas con incrementos independientes, etc.

Ejemplo 2

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$

Ejemplo 2

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $G \sim \mathcal{DP}(F/c)$ como la distribución $\mathcal{P}$ de G, donde
  $1/c$ es parámetro de precisión
  F es una f.d.a. paramétrica que coincide con la media de G

Ejemplo 2

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $G \sim \mathcal{DP}(F/c)$ como la distribución $\mathcal{P}$ de G, donde
  $1/c$ es parámetro de precisión
  F es una f.d.a. paramétrica que coincide con la media de G
- $\Rightarrow$ Queremos que $\text{E}(G) = F$ con F la f.d.a. de una $\text{Ga}(a, b)$

Ejemplo 2

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $G \sim \mathcal{DP}(F/c)$ como la distribución $\mathcal{P}$ de G, donde
  $1/c$ es parámetro de precisión
  F es una f.d.a. paramétrica que coincide con la media de G
- $\Rightarrow$ Queremos que $\text{E}(G) = F$ con F la f.d.a. de una $\text{Ga}(a, b)$
- $\therefore$ $Y_i \mid G \sim G$, $i = 1, 2, \ldots$ cond. indep. dado G y $G \sim \mathcal{DP}(F/c) \Rightarrow Y_i \sim \text{Ga}(a, b)$ con $\text{Corr}(Y_i, Y_j) = c/(c+1)$, $\forall\ i \neq j$.

Ejemplo 2

- Queremos una sucesión de v.a. intercambiables t.q. $Y_i \sim \text{Ga}(a, b)$
- Proponemos $G \sim \mathcal{DP}(F/c)$ como la distribución $\mathcal{P}$ de G, donde
  $1/c$ es parámetro de precisión
  F es una f.d.a. paramétrica que coincide con la media de G
- $\Rightarrow$ Queremos que $\text{E}(G) = F$ con F la f.d.a. de una $\text{Ga}(a, b)$
- $\therefore\ Y_i \mid G \sim G$, $i = 1, 2, \ldots$ cond. indep. dado G y
  $G \sim \mathcal{DP}(F/c) \Rightarrow Y_i \sim \text{Ga}(a, b)$ con
  $\text{Corr}(Y_i, Y_j) = c/(c + 1)$, $\forall\ i \neq j$.
- Nota: La correlación es no paramétrica !

Supuestos

- Supongamos que para cada individuo (póliza) j tenemos información sobre sus reclamos $\{Y_{ij}\}$, para $i = 1, \ldots, N_{tj}$, donde

Supuestos

- Supongamos que para cada individuo (póliza) j tenemos información sobre sus reclamos $\{Y_{ij}\}$, para $i = 1, \ldots, N_{tj}$, donde
  - $Y_{ij} \sim f(y \mid \theta)$ independientes $\forall j$

Supuestos

- Supongamos que para cada individuo (póliza) j tenemos información sobre sus reclamos $\{Y_{ij}\}$, para $i = 1, \ldots, N_{tj}$, donde
  - $Y_{ij} \sim f(y \mid \theta)$ independientes $\forall j$
  - $N_{tj} \sim \mathrm{Po}(\lambda t)$ independientes $\forall j$ e indep. de $\{Y_{ij}\}$.

Supuestos

- Supongamos que para cada individuo (póliza) j tenemos información sobre sus reclamos $\{Y_{ij}\}$, para $i = 1, \ldots, N_{tj}$, donde
  - $Y_{ij} \sim f(y \mid \theta)$ independientes $\forall j$
  - $N_{tj} \sim \text{Po}(\lambda t)$ independientes $\forall j$ e indep. de $\{Y_{ij}\}$.
- Los reclamos agregados para cada individuo j son

$$X_{tj} = \sum_{i=1}^{N_{tj}} Y_{ij}$$

Verosimilitud

- En el caso de que $Y_{ij} \sim \text{Ga}(a, b)$, y dado que en $(0, T]$, $N_{Tj} = n_j \Rightarrow$

Verosimilitud

- En el caso de que $Y_{ij} \sim \text{Ga}(a, b)$, y dado que en $(0, T]$, $N_{Tj} = n_j \Rightarrow$
- Para CPP:

$$f(\mathbf{y}_1, \ldots, \mathbf{y}_m \mid \theta) = \prod_{j=1}^{m} \prod_{i=1}^{n_j} \text{Ga}(y_{ij} \mid a, b)$$

Verosimilitud

- En el caso de que $Y_{ij} \sim \mathrm{Ga}(a, b)$, y dado que en $(0, T]$, $N_{Tj} = n_j \Rightarrow$
- Para CPP:

$$f(\mathbf{y}_1, \ldots, \mathbf{y}_m \mid \theta) = \prod_{j=1}^{m} \prod_{i=1}^{n_j} \mathrm{Ga}(y_{ij} \mid a, b)$$

- Para $\mathrm{ECP}_{\mathrm{p}}$:

$$\begin{aligned} f(\mathbf{y}_1, \ldots, \mathbf{y}_m \mid a, b, c) &= \prod_{j=1}^{m} \int \left\{ \prod_{i=1}^{n_j} \mathrm{Ga}(y_{ij} \mid a + c, b + z_j) \right\} \\ &\times \mathrm{Gga}(z_j \mid a, b, c) \mathrm{d}z_j \end{aligned}$$

Verosimilitud

- Para $\mathrm{ECP}_{\mathrm{np}}$:

$$
\begin{aligned}
f(\mathbf{y}_1,\ldots,\mathbf{y}_m \mid a,b,c) &= \prod_{j=1}^{m} \mathsf{E}_{\mathcal{DP}}\left\{\prod_{i=1}^{n_j} G(\mathrm{d}y_{ij})\right\} \\
&= \prod_{j=1}^{m} \frac{(1/c)^{k_j}\Gamma(1/c)}{\Gamma(1/c+n_j)} \prod_{i=1}^{k_j} \mathrm{Ga}(y_{ij}^{*} \mid a,b),
\end{aligned}
$$

donde $y_{1j}^{*},\ldots,y_{kj}^{*}$ los distintos y_{ij}'s para $i=1,\ldots,n_j$, con $k_j \leq n_j$

Verosimilitud

- Para ECP$_{\rm np}$:

$$
\begin{aligned}
f(\mathbf{y}_1,\ldots,\mathbf{y}_m \mid a,b,c) &= \prod_{j=1}^{m} \mathrm{E}_{\mathcal{DP}}\left\{\prod_{i=1}^{n_j} G(\mathrm{d}y_{ij})\right\} \\
&= \prod_{j=1}^{m} \frac{(1/c)^{k_j}\Gamma(1/c)}{\Gamma(1/c+n_j)} \prod_{i=1}^{k_j} \mathrm{Ga}(y_{ij}^{*} \mid a,b),
\end{aligned}
$$

donde $y_{1j}^{*},\ldots,y_{kj}^{*}$ los distintos y_{ij}'s para $i=1,\ldots,n_j$, con $k_j \leq n_j$

- Iniciales: $\pi(a,b,c) = \mathrm{Ga}(a|\alpha_a,\beta_a)\mathrm{Ga}(b|\alpha_b,\beta_b)\mathrm{Ga}(c|\alpha_c,\beta_c)$

Distribución final

- Para λ:

$$\pi(\lambda \mid \mathbf{w}) = \mathrm{Ga}\left(\lambda \,\middle|\, \alpha_\lambda + \sum_{j=1}^{m} n_j,\ \beta_\lambda + \sum_{j=1}^{m}\sum_{i=1}^{n_j} w_{ij}\right),$$

donde $W_{ij} = J_{ij} - J_{i-1,j} \sim \mathrm{Ga}(1, \lambda)$ y $J_{1j}, J_{2j}, \ldots, J_{n_j j}$ son los tiempos de salto del proceso $\mathrm{Po}(\lambda t)$ para el individuo j

MEPS

- MEPS= Encuesta panel de gastos médicos

MEPS

- MEPS= Encuesta panel de gastos médicos
- MEPS dá información sobre el uso del sistema de salud, gastos, formas de pago, coberture de seguro médico, etc. en E.U. del 2005

MEPS

- MEPS= Encuesta panel de gastos médicos
- MEPS dá información sobre el uso del sistema de salud, gastos, formas de pago, coberture de seguro médico, etc. en E.U. del 2005
- Hay 3341 eventos (estancias en hospital) reportados

MEPS

- MEPS= Encuesta panel de gastos médicos
- MEPS dá información sobre el uso del sistema de salud, gastos, formas de pago, coberture de seguro médico, etc. en E.U. del 2005
- Hay 3341 eventos (estancias en hospital) reportados
- Depués de quitar los eventos incompletos y de sumar los reclamos del mismo día $\Rightarrow$ 1729 eventos

MEPS

- MEPS= Encuesta panel de gastos médicos
- MEPS dá información sobre el uso del sistema de salud, gastos, formas de pago, coberture de seguro médico, etc. en E.U. del 2005
- Hay 3341 eventos (estancias en hospital) reportados
- Depués de quitar los eventos incompletos y de sumar los reclamos del mismo día $\Rightarrow$ 1729 eventos
- Los 1729 eventos corresponden a 66 individuos $\Rightarrow$ promedio de 26 eventos por individuo al año, que van de 1 a 122 eventos por persona.

Iniciales

- Se transformaron los datos con una potencia de $1/4$

Iniciales

- Se transformaron los datos con una potencia de $1/4$
- Las fechas se transformaron en días transcurridos del año 2005

Iniciales

- Se transformaron los datos con una potencia de $1/4$
- Las fechas se transformaron en días transcurridos del año 2005
- Se consideraron iniciales vagas para todos los parámetros de los modelos: $(\alpha_a, \beta_a) = (0.01, 0.01)$, $(\alpha_b, \beta_b) = (0.01, 0.01)$, $(\alpha_c, \beta_c) = (0.01, 0.01)$ y $(\alpha_\lambda, \beta_\lambda) = (0.01, 0.01)$.

Iniciales

- Se transformaron los datos con una potencia de $1/4$
- Las fechas se transformaron en días transcurridos del año 2005
- Se consideraron iniciales vagas para todos los parámetros de los modelos: $(\alpha_a, \beta_a) = (0.01, 0.01)$, $(\alpha_b, \beta_b) = (0.01, 0.01)$, $(\alpha_c, \beta_c) = (0.01, 0.01)$ y $(\alpha_\lambda, \beta_\lambda) = (0.01, 0.01)$.
- Se implementó un muestreador de Gibbs con 600,000 iteraciones con un período de calentamiento de 100,000 manteniedo una observación cada 50 para reducir la autocorrelación de la cadena.

Dist. finales para c

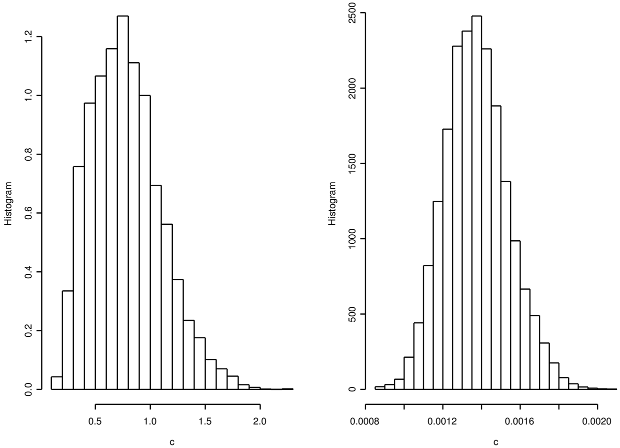

Figure: Distribución posterior del parametro c: (*izq.*) construcción paramétrica y (*der.*) construcción no paramétrica.

Resultados

Quantity	Model			
	CPP		ECP_p	
	Mean	95% CI	Mean	95% CI
LPML	-4124.6	–	-4117.9	–
a	11.99	(11.25, 12.72)	11.45	(10.61, 12.24)
b	1.28	(1.20, 1.36)	1.25	(1.17, 1.34)
c	–	–	0.78	(0.27, 1.50)

Table: Comparación de modelos y resúmenes posteriores de (a, b, c) para los modelos CCP y ECP_p. Media posterior e IC al 95%

Resultados

Quantity	Model			
	CPP		$\mathrm{ECP_p}$	
	Mean	95% CI	Mean	95% CI
LPML	-4124.6	–	-4117.9	–
a	11.99	(11.25, 12.72)	11.45	(10.61, 12.24)
b	1.28	(1.20, 1.36)	1.25	(1.17, 1.34)
c	–	–	0.78	(0.27, 1.50)

Table: Comparación de modelos y resúmenes posteriores de (a, b, c) para los modelos CCP y $\mathrm{ECP_p}$. Media posterior e IC al 95%

- Correlación con $\mathrm{ECP_p}$: $\rho \in (0.021, 0.112)$

Resultados

Quantity	Model			
	CPP		$\mathrm{ECP_p}$	
	Mean	95% CI	Mean	95% CI
LPML	-4124.6	–	-4117.9	–
a	11.99	(11.25, 12.72)	11.45	(10.61, 12.24)
b	1.28	(1.20, 1.36)	1.25	(1.17, 1.34)
c	–	–	0.78	(0.27, 1.50)

Table: Comparación de modelos y resúmenes posteriores de (a, b, c) para los modelos CCP y $\mathrm{ECP_p}$. Media posterior e IC al 95%

- Correlación con $\mathrm{ECP_p}$: $\rho \in (0.021, 0.112)$
- $\lambda \mid \text{data} \sim \text{Ga}(1729.01, 52.74) \Rightarrow \hat{\lambda} = 32.78$

Probabilidades de ruina

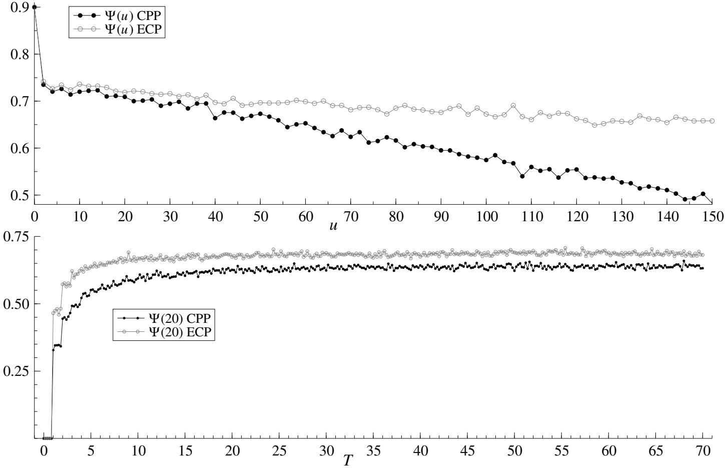

Figure: Estimadores MC de la prob. de ruina para CPP y ECP_p ajustados a los datos MEPS.

Predictiva final para X_t de un individuo

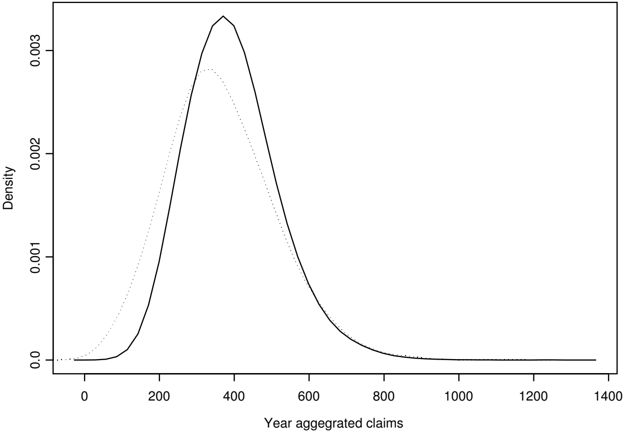

Figure: Predictiva final para los reclamos agregados en un año para un individuo (en miles de dolares). (—) CPP, $(\cdots)$ $\mathrm{ECP}_{\mathrm{p}}$.

Resultados

- Determinación de la reserva por individuo:

Resultados

- Determinación de la reserva por individuo:
- Cuantil 95%: 606,438 usd para CPP y 618,370 usd para $\mathrm{ECP}_{\mathrm{p}}$

Resultados

- Determinación de la reserva por individuo:
- Cuantil 95%: 606,438 usd para CPP y 618,370 usd para $\mathrm{ECP_p}$
- Media posterior: 394,796 usd para CPP y 364,120 usd para $\mathrm{ECP_p}$
  $\Rightarrow$ Ahorro de 30,676 usd por individuo por año.