



Modelo bayesiano para la Reincidencia

Oslando Padilla
Rolando de la Cruz
Departamento de Salud Pública
PUC

Trabajo parcialmente financiado por Proyecto Anillo ACT87 “Métodos
Cuantitativos en Seguridad”



INDICE

- Dos aspectos de la Reincidencia
- Objetivos
- Estudios previos
- Dos modelos básicos.
 - Modelo de cura
 - Alternativas al modelo de Cura
 - Modelo de riesgos competitivos
- Modelo bayesiano de Riesgos competitivos + Cura
- Representación de algunos resultados
- Situación en Chile
- Aplicaciones
- Bibliografía



Reincidencia (Recidivism): Volver a ... cometer un delito.

- Probabilidad de Reincidir (En el mismo delito o en otro)
- Tiempo que demora en Reincidir (Desde que es condenado o liberado hasta que es apresado, formalizado o condenado)

Algunos Predictores:

1. Número de delitos previos
2. Edad
3. Trastornos de la personalidad
4. Entorno Familiar
5. Programa de Rehabilitación



Objetivos

- ¿Cómo influyen los predictores en la probabilidad y en el tiempo de reincidencia?
- ¿Qué características tienen los individuos que tienen una alta probabilidad de reincidir?
- ¿Qué diferencia a los individuos que reinciden rápidamente de aquellos que demoran más?
- ¿Sobre qué conjunto de individuos es más eficaz un programa de rehabilitación?
- ¿Es un programa de rehabilitación más efectivo que otro? ¿En qué sentido?

Estudios previos

- Loza, 2003: Predictores clínicos (psiquiatría) de la violencia.
- Beck & Shipley (prisioneros liberados en 1983, en 11 estados de USA).
 1. 5% de los exconvictos tenía 45 o más detenciones (pre y post) y eran responsables del 20% de todos los cargos.
 2. El 18,6% tenía 25 detenciones o más y eran responsables del 47.8% de todos los cargos.
 3. El 18.9% tenía menos de 5 detenciones y sólo le correspondían el 3,3% de todos los cargos.
 4. Los que habían cometido delito contra la propiedad fueron más rearrestados por delitos contra la propiedad que los que cometieron delitos violentos (49.8% vs. 32.1)

Modelos de Cura

Modelos estándar de sobrevida (Regresión de Cox)

$$\lim_{t \rightarrow \infty} S(t) = 0$$

Se desarrollaron modelos tales que:

$$\lim_{t \rightarrow \infty} S(t) = q > 0$$

Derivación del Modelo básico de Cura

Sean:

$F_i = 1$ si el i -ésimo individuo no se cura y

$F_i = 0$ si se cura

y sea T : tiempo hasta que el individuo “falla”
(p.ej. hace metástasis o reincide en el delito)

$$\begin{aligned} S(t) &= P(T > t) = P(T > t / F_i = 0)P(F_i = 0) + P(T > t / F_i = 1)P(F_i = 1) \\ &= S(t / F_i = 0) p + 1 \cdot q \end{aligned}$$

Alternativas

Promotion Time Cure Models

$$S(t / X_i, Z_i) = e^{-p(Z_i)F(t/X_i)}$$

Modelo general de cura

$$\frac{S(t / X_i, Z_i)^a - 1}{a} = -p(a, Z_i)F(t / X_i)$$

Si $a \rightarrow 0$ Promotion time cure Model

Si $a = 1$ Modelo estándar de cura

Modelo de Riesgos Competitivos

Existen K posibles resultados mutuamente excluyentes que ocurren en los tiempos T_1, \dots, T_K pero sólo observamos el que ocurre primero.

Sea R_{ik} el evento que para el individuo i se observa el resultado k (p.ej. el tumor regresa en el mismo lugar o en otro lugar o la próxima detención del individuo es por un delito de tipo sexual o de violencia no sexual o por otro delito como robo, estafa, etc)

p_{ik} probabilidad que para individuo i se observe el resultado k

$p_i = \sum_{k=1}^K p_{ik}$ p.ej. es la probabilidad de que en el individuo i retorne el cáncer o de que cometa algún delito

$q_i = 1 - p_i$ p. ej. es la probabilidad de que el individuo i se cure del cáncer o no cometa delito

$$F_i(t) = \sum_{k=1}^K P(T_i < t / R_{ik}) P(R_{ik}) = \sum_{k=1}^K p_{ik} F_{ik}(t)$$

$$L = \prod_{i=1}^n \left(\prod_{k=1}^K (p_{ik} f_{ik}(t_i))^{\delta_{ik}} \right) \left(1 - \sum_{k=1}^K p_{ik} F_{ik}(t_i) \right)^{1-\delta_i}$$

$$\delta_{ik} = \begin{cases} 1 & \text{si el } i\text{-ésimo individuo reincide en el delito } k \\ 0 & \text{e. o. c} \end{cases}$$

$$\delta_i = \sum_{k=1}^K \delta_{ik}$$

Distribuciones para el Tiempo

Exponencial

$$f_{ik}(t) = h_{ik} e^{-h_{ik}t} \quad F_{ik}(t) = 1 - e^{-h_{ik}t}$$

Weibull

$$f_{ik}(t) = r_k h_{ik} t^{r_k - 1} e^{-h_{ik}t^{r_k}}$$

$$F_{ik}(t) = 1 - e^{-h_{ik}t^{r_k}}$$

$$X \sim \Gamma(\alpha, \beta)$$

$$f(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}$$

$$F(x) = \int_0^{\frac{x}{\beta}} t^{\alpha-1} e^{-t} dt$$

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$$

$$Y \sim \text{LogNormal}(\mu, \sigma^2) \Leftrightarrow \ln(Y) \sim N(\mu, \sigma^2)$$

$$f(y) = \frac{1}{y\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} \quad F(y) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right)$$

Probabilidad (betas)

$$p(\mathbf{Z}_i) = \frac{e^{\beta^t \mathbf{Z}_i}}{1 + \sum_{k=1}^K e^{\beta^t \mathbf{Z}_i}}$$

Tiempo (deltas)

$$h_{ik} = e^{\delta_k^t \mathbf{X}_i}$$

Prioris

$$\beta_i \sim N(0; 0.05)$$

$$\delta_i \sim N(0; 0.05)$$

$$r_i \sim \exp(1.2)$$

Modelo para Inglaterra y Gales (Escarela, 2000)

- 3068 hombres condenados por delitos sexuales contra mujeres en 1973.
- Historia desde 1963 y seguimiento hasta 1994.
- Variables:
 1. NP: Número de delitos no sexuales previos en los 10 años anteriores.
 2. NPS: Número de delitos sexuales previos en los 10 años anteriores.
 3. Edad: Edad del convicto en 1973.
 4. AgeVict_16: 1 si la edad de la victima es 16 años o más , 0 si es menor.
 5. Type: Tipo de delito que comete (0: no reincide, 1: sexual, 2: violento no sexual, 3: Otro)
 6. TIME: tiempo en riesgo hasta que reincide o tiempo de seguimiento sin que reincida

Datos

NP	NPS	AGE	AGEVICT_16	type	SENLEN	SERVED	TIME
3	0	34,95	0	3	274	182	372
5	0	46,53	1	3	0	0	312
4	1	16,86	0	3	0	0	1649
9	1	20,02	0	3	731	304	36
11	3	39,82	0	1	122	81	16
2	1	19,35	1	1	1096	365	608
3	3	25,74	0	0	0	0	8024
2	1	21,18	0	1	1096	365	584

Estimaciones sin covariables

Categoría	Probabilidad	r (Weibull)
1	0.143 (0.006)	0.851 (0.787-0.918)
2	0.067 (0.004)	0.861 (0.761-0.967)
3	0.315 (0.008)	0.837 (0.792-0.885)

$$q = 1 - 0.141 - 0.066 - 0.313 = 0.48$$

Estimaciones de Probabilidad con covariables

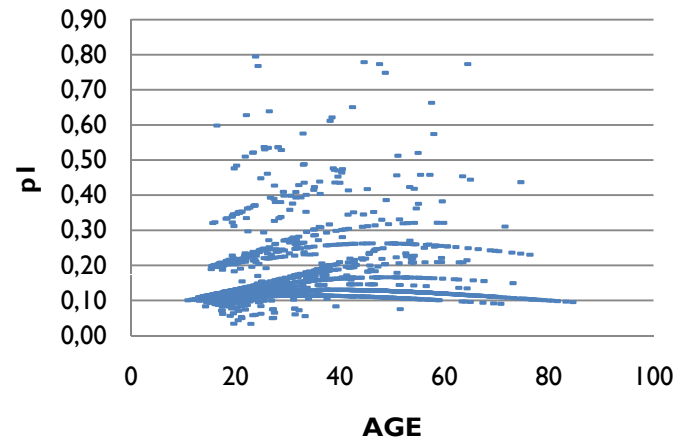
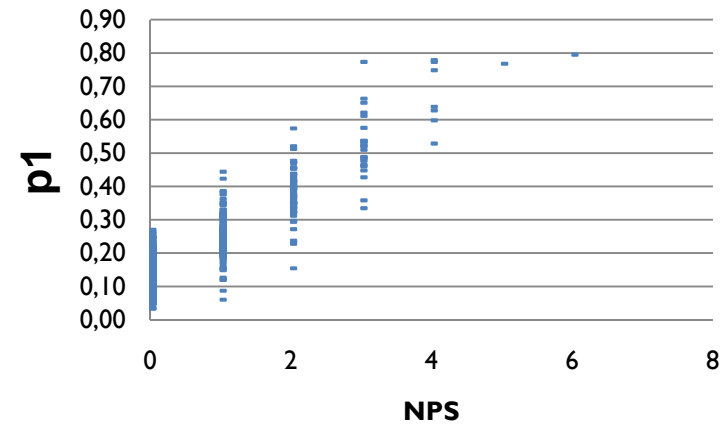
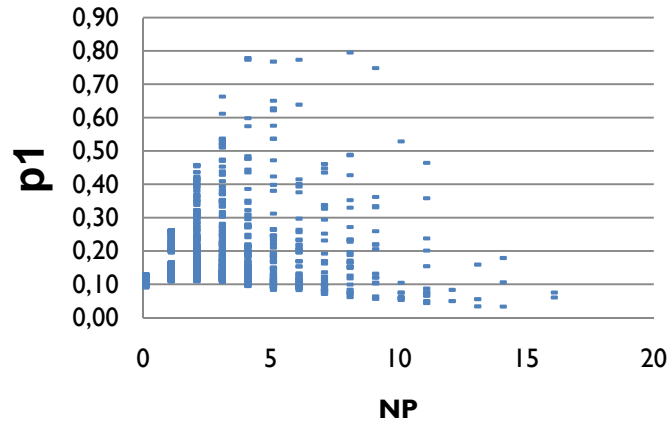
Parámetro	Estimación	SD	
Intercept (sex)	-1.385	0.171*	
Intercept (viol)	-1.046	0.23*	
Intercept (otro)	0.446	0.125*	
NP (sex)	0.412	0.043*	
NP (viol)	0.569	0.045*	
NP (otro)	0.583	0.039*	
NPS (sex)	0.558	0.129*	
NPS (viol)	0.135	0.165	ns
NPS (otro)	-0.173	0.135	ns
Edad (sex)	-0.013	0.004*	
Edad (viol)	-0.054	0.009*	
Edad (otro)	-0.059	0.004*	
AgeVict_16 (sex)	-0.22	0.132	ns
AgeVict_16 (viol)	-0.098	0.16	ns
AgeVict_16 (otro)	-0.37	0.098*	

Real vs. Pronosticado

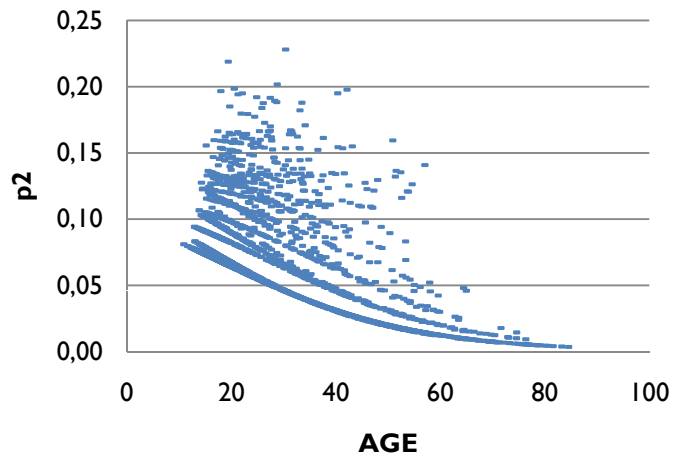
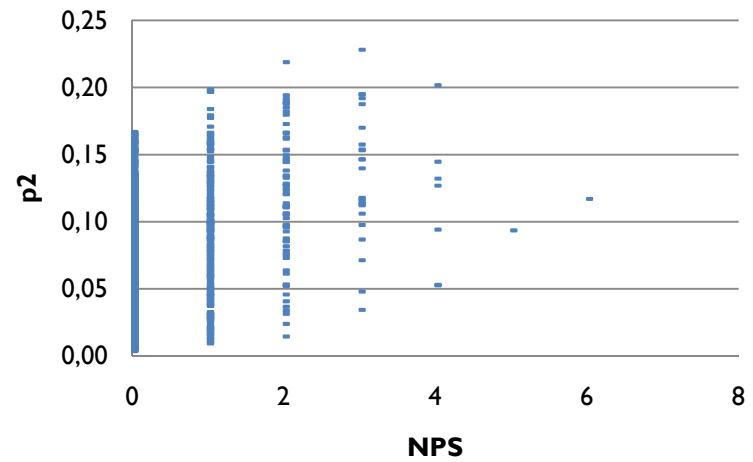
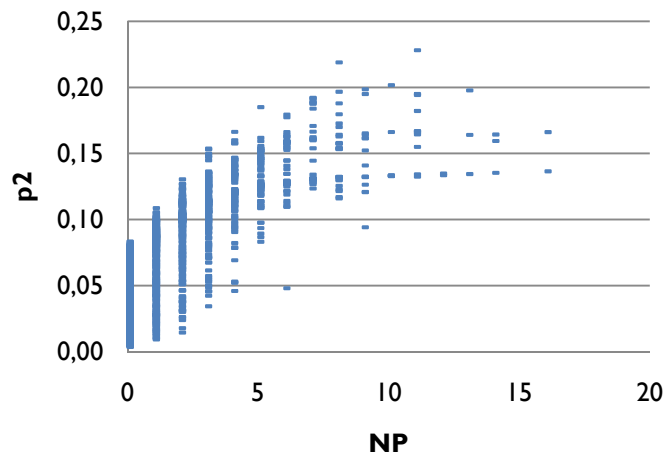
		type_pred			Total
		0	1	3	
type	0	1294	16	165	1475
	1	270	32	131	433
	2	102	8	92	202
	3	456	22	480	958
Total		2122	78	868	3068

1. El modelo falla al predecir que nadie cometerá un delito de tipo violento
2. El modelo predice bien al 58.8% de los sujetos $(1294+32+480)/3068$
3. El modelo predice bien al 87.7% de los que no reinciden $(1294/1475)$
4. De aquellos para los que el modelo predice que reinciden, sólo un 20% no lo hace $(16/78$ y $165/868)$

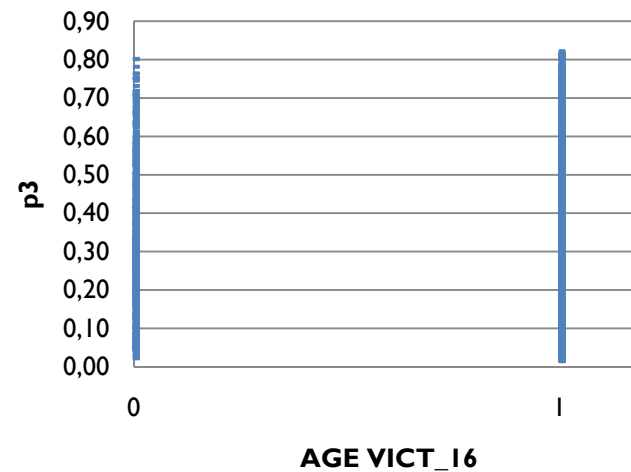
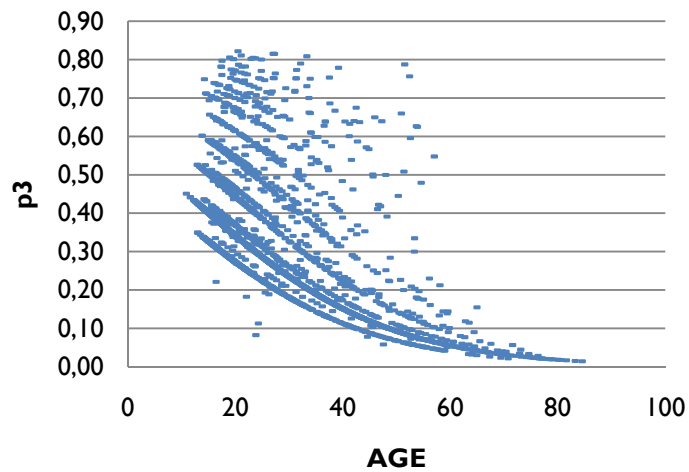
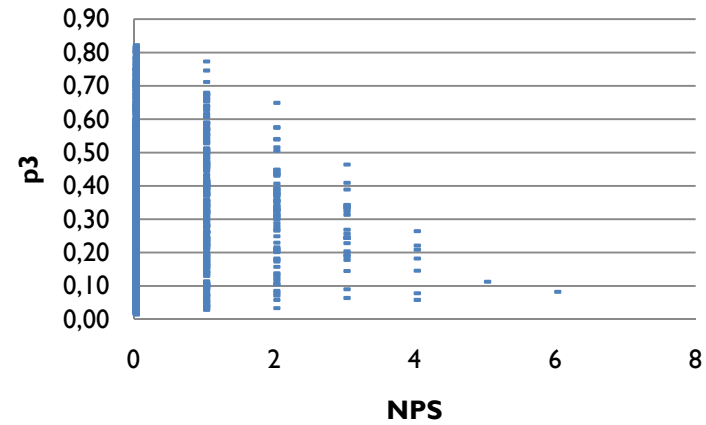
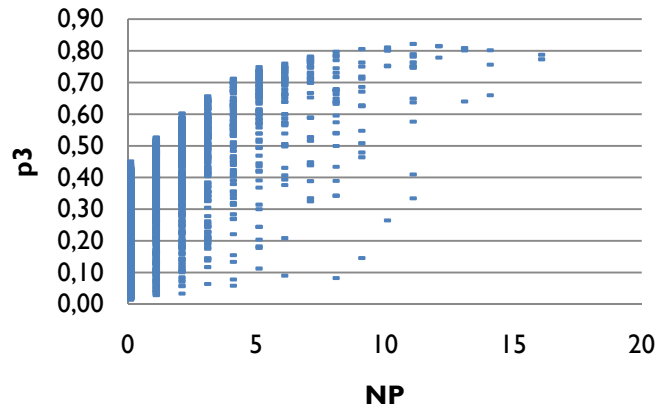
Probabilidad de reincidir por un delito sexual



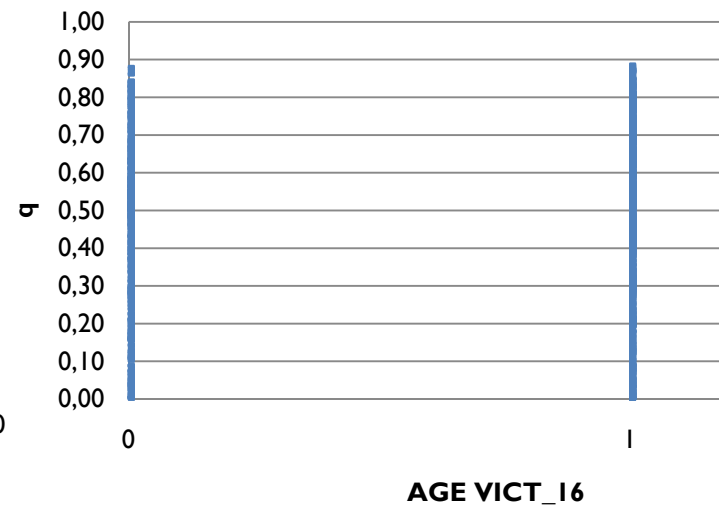
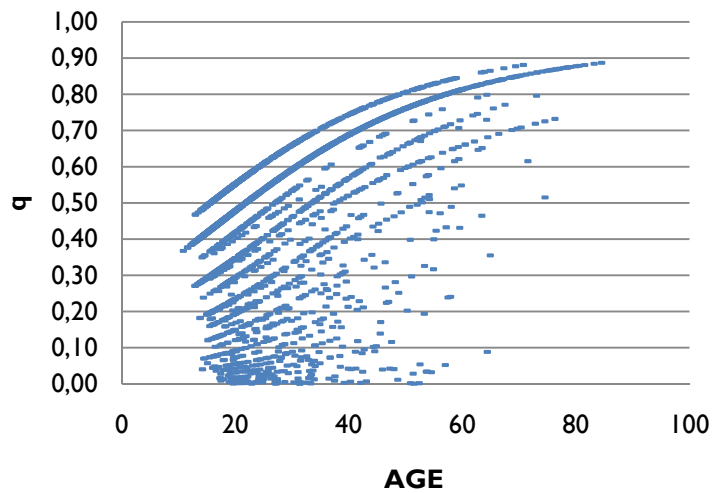
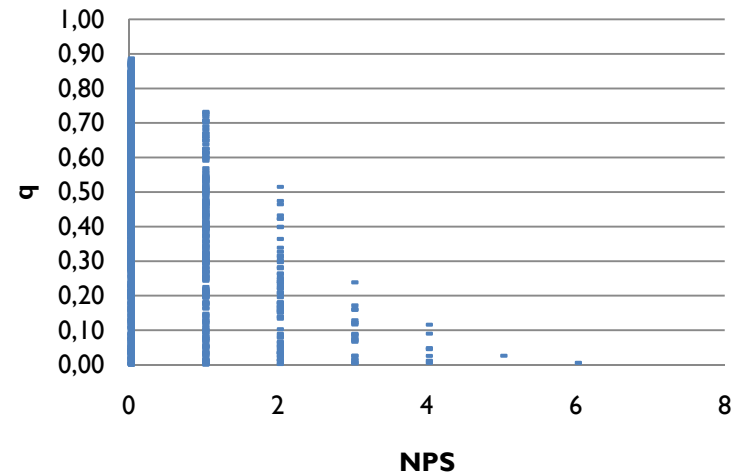
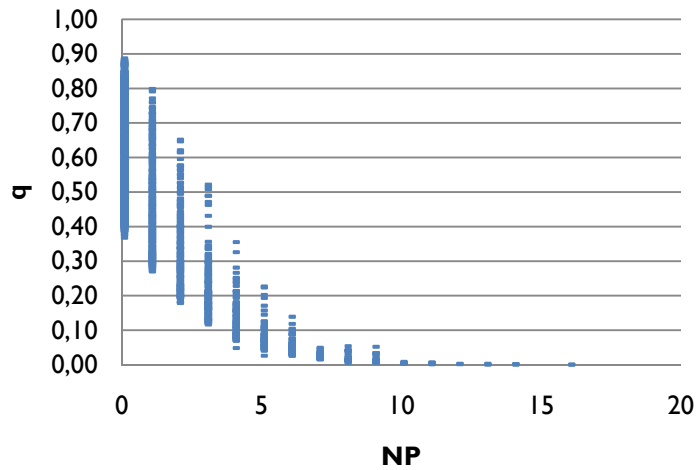
Probabilidad de reincidir por un delito violento



Probabilidad de reincidir por Otro delito



Probabilidad de no reincidir



Valores pronosticados

NP	NPS	AGE	AGE VICT 16	type	TIME	p1	p2	p3	q	l1	l2	l3	P10	P25	P50	P75	P90
6	2	33	1	3	500	0,40	0,14	0,41	0,05	1242	1019	761	80	219	527	1054	1751
9	1	54,1	1	3	422	0,36	0,13	0,48	0,03	1224	839	902	95	260	625	1251	2077
3	2	20,7	1	1	77	0,35	0,13	0,38	0,14	1130	911	669	70	192	464	927	1540
3	2	20,2	1	3	222	0,35	0,13	0,38	0,14	1141	931	675	71	194	468	936	1554
4	1	41	1	3	1178	0,34	0,09	0,36	0,20	912	515	642	68	185	445	890	1478
9	3	21,6	1	3	67	0,33	0,19	0,46	0,01	2007	2844	1069	113	308	741	1482	2462
9	1	50,9	1	3	1801	0,33	0,13	0,51	0,03	1299	956	953	100	274	661	1321	2194
8	1	47,5	1	1	514	0,33	0,13	0,50	0,04	1242	895	902	95	259	625	1250	2077
2	2	15,5	1	3	7075	0,32	0,12	0,37	0,18	1118	918	653	69	188	453	905	1504
4	2	18,9	1	3	868	0,32	0,15	0,45	0,08	1302	1205	773	81	222	536	1071	1779
2	2	15,1	1	2	280	0,32	0,12	0,38	0,18	1126	932	657	69	189	456	911	1514
3	1	37,4	1	3	2248	0,32	0,09	0,33	0,27	875	485	610	64	175	422	845	1403
3	1	35,8	1	0	7782	0,31	0,09	0,34	0,26	902	519	627	66	180	435	869	1443

Situación en Chile

1. La Información en Chile es rica (Informe Técnico de Gendarmería).
 2. Aún existe una falta de integración de diferentes fuentes de Información (Gendarmería, Registro Civil, Fiscalía).
 3. Reciente Firma de convenios
CEAMOS - Gendarmería
CEAMOS - Carabineros de Chile
- ⇒? Modelos de la Reincidencia



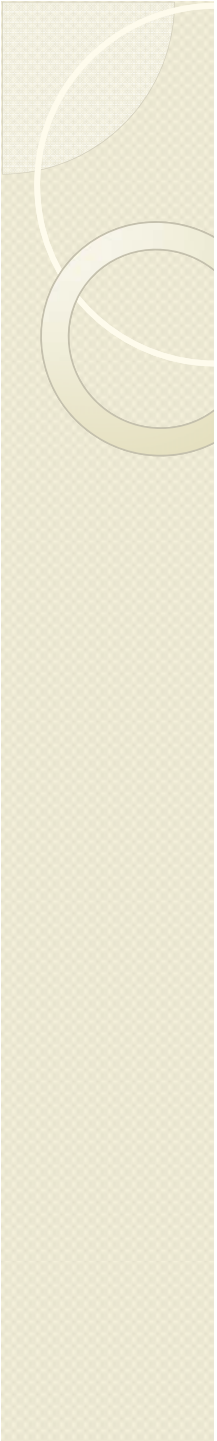
4. Metodología Estadística

Desarrollar algoritmos que permitan incluir otras distribuciones de probabilidad para el tiempo

Está implementado un análisis semiparamétrico que permite mayor flexibilidad sobre la distribución de probabilidades de los tiempos, basado en la distribución exponencial por tramos.

Aplicaciones

- Conocer las probabilidades y el tiempo de reincidencia \Rightarrow Mejor planificación y gestión.
- Permitiría incorporar la disminución de la probabilidad de reincidencia y/o el aumento del tiempo, entre los criterios para asignar a un reo a un programa de rehabilitación.

- 
- Permitiría incorporar estas probabilidades y tiempos en la búsqueda de sospechosos de un delito (Dificultad: Argumentos éticos).
 - El poder judicial pudiera usar esta información a la hora de asignar una condena de manera que esta cumpla dos objetivos claves: favorecer la rehabilitación y proteger a la sociedad de los delincuentes que más persisten en el delito.

Bibliografía

- Beck A. J., Shipley B.E. Recidivism of prisoners released in 1983. Bureau of Justice Statistics Special Report, 1989.
- Bierens H. J., Carvalho J. R. Semi-Nonparametric Competing Risks Analysis of Recidivism. Journal of Applied Econometrics. 2007.
- Escarela G. , Francis B. and Soothill K. Competing Risks, Persistence, and Desistance in Analyzing Recidivism. Journal of Quantitative Criminology. 2000.
- Loza W. Predicting violent and nonviolent recidivism of incarcerated male offenders. Aggression and Violent Behavior, 2003.
- Ross Maller and Xian Zhiou. Survival Analysis with Long –Term Survivors. John Wiley & Sons. 1996.
- Yin G ., Ibrahim J. A General Class of Bayesian Survival Models with Zero and Nonzero Cure Fraction. Biometrics. 2005.
- Yin G ., Ibrahim J. Cure Rate Models:A Unified Approach. The Canadian Journal of Statistics. 2005.
- Informe técnico “Tasas de Reincidencia Cohorte Egresados 2007 de los Programas evaluados por la DIPRES.” elaborado por la Unidad de Investigaciones Criminológicas Gendarmería de Chile, Noviembre, 2010.