ACT-11302 Calculo Actuarial III

Sesion - Examen Parcial 1 - Soluciones

Juan Carlos Martinez-Ovando

Departamento Academico de Actuaria y Seguros

Pregunta 1.1

Bajo el enfoque de riesgo individual, el número potencial de siniestros nunca excederá al número de suscriptores.

R: Cierto, porque bajo el enfoque de `riesgo individual` *cada suscriptor de polizas, $j = 1, \ldots, J_t$ tiene asociada la posibilidad de sufrir a los mas un riesgo en un periodo de opreacion dado, i.e.*

$$\iota_j = \begin{cases} 1, & \textit{siniestro} \\ 0, & \textit{no siniestro} \end{cases}.$$

Por lo que el numero de siniestros a lo mas seria igual a J.

Pregunta 1.2

El número de parámetros a estimar con m datos, $\{n_1, \ldots, n_m\}$, del modelo para la frecuencia de siniestros más flexible es igual al de cualquier distribución paramétrica.

R: Falso. El modelo mas flexible es el que no presupone estructuras en las masas de probabiliades sobre $\mathcal{N} = \{0, 1, \ldots\}$ (puede ser finito o no), i.e.

$$\mathbb{P}(N = n) = p_n,$$

para $n \in \mathcal{N}$, sujetos a $p_n > 0$, para todo n, y $\sum_{n \in \mathcal{N}} p_n = 1$. Asi, el ***numero de parametos de este modelo es igual al numero de la coleccion*** $(p_n)_{n \in \mathcal{N}}$.

Por otro lado, los ***modelos parametricos*** *imponene restricciones sobe las p_ns, e.g. $p_n = \frac{e^{-\lambda}\lambda^n}{n!}$,* ***para todo*** *$n$ y una cierta $\lambda > 0$.*

Mas aun, con U datos distintos de $\{n_1, \ldots, n_m\}$ podremos estimar a lo mas U de las p_ns en el modelo mas flexible.

Pregunta 1.3

La modificación de distribuciones de probabilidad vía mezclas de probabilidades genera modelos que son típicamente subdispersos.

R: Falso. Los modelos de mezclas, tanto en el caso finito,

$$\mathbb{P}(N = n) = \sum_{k=1}^{K} \omega_k \mathbb{P}(N = n|\theta_k),$$

como en el caso no numerable,

$$\mathbb{P}(N = n) = \int_{\theta} \mathbb{P}(N = n|\theta)\pi(\theta)d\theta,$$

tienen mas varianza que en el modelo aislad (marginal/condicional)

$$\mathbb{P}(N = n|\theta),$$

para un θ fijo. Por lo que el resultados sera tipicamente mas disperso que el asilado, y por consecuencia podra ser (en la mayoria de los casos) sobredisperso.

Pregunta 1.4

El error de estimación de un modelo de probabilidad no tiene efecto sobre la cuantificación de riesgos.

R: Falso. Si tiene un efecto, derivado de un efecto de ***propagacion de incertidumbre****. Por ejemplo, si* $\mathbb{P}(S \leq s|\theta)$ *es un modelo especifico para un* θ *dado, tiene una varianza especifica* $var(S|\theta)$*, e.g.*

Ahora, si el se incorpora variabilidad en θ *asociada con el proceso de estimacion, grosso modo, promediariamos* $\mathbb{P}(S \leq s|\theta)$ *especto a ciertos valores de diferentes* θ*s, cuyo espacio es denotado por* $\hat{\Theta}$*, tendriamos que* `pseudo promediar` $\mathbb{P}(S \leq s|\theta)$ *entorno a* $\theta \in \hat{\Theta}$*.*

Esto ultimo hace que el `modelo presudo promediado` *sea mas disperso que el* `modelo especifico`*, por lo que la prevision de riesgos sera mas grande.*

Pregunta 1.5

Las masas de probabilidad para las diferentes frecuencias de siniestros de un portafolio de seguros siempre tienen un comportamiento monótono exponencial.

R: Falso. La clase de modelos $(\alpha, \beta, 0)$ *tiene ciertamente esta caracteristica. Pero, en casos mas generales, no necesariamente el comportamiento de las* p_n *sera monotono decreciente exponencial.*

Por ejemplo, si

$$\mathbb{P}(N = n) = p_n = \omega Po(n|\lambda) + (1 - \omega) Bin(n|R, \theta),$$

con $0 < \omega < 1$, *tendra un decaimiento exponencial, pero no monotono.*

Pregunta 2

Como vimos en clase, el procedimiento de mezcla de distribuciones permite flexibilizar la forma en cómo distribuciones de probabilidad describen diferentes características de un fenómeno aleatorio/desconocido. Considera que N denota la frecuencia de siniestros de un portafolio de seguros. Con probabilidad $0 < \theta < 1$ la variable N es descrita con la distribución $\text{Bin}(n|4, 0.5)$ y con probabilidad $(1 - \theta)$ se describe con $\text{Bin}(n|6, 0.75)$; considera θ conocido. Con base en esto:

2.a Calcula analíticamente $\mathbb{P}(N = 2)$

El modelo es de tipo mezcla, i.e.

$$\mathbb{P}(N = n) = \theta Bin(n|4, 0.5) + (1 - \theta)Bin(n|6, 0.75),$$

con soporte en $\mathcal{N} = \{0, 1, \ldots, 6\}$.

Asi, tenemos

$$\mathbb{P}(N = 2) = \theta Bin(2|4, 0.5) + (1 - \theta)Bin(2|6, 0.75).$$

2.b. Determina analíticamente la media y varianza de N.

Recordemos, media y varianza pueden verse como

$$\mathbb{E}_{\mathbb{P}}(\phi(N)) = \sum_{n \in \mathcal{N}} \phi(n) p_n.$$

Asi, sustituyendo $p_n = \theta Bin(n|4, 0.5) + (1-\theta) Bin(n|6, 0.75)$ *tenemos*

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}(N) &= \theta \mathbb{E}_{Bin(4,0.5)}(N) + (1-\theta)\mathbb{E}_{Bin(6,0.75)}(N) \\ &= \theta(2 \times 0.5) + (1-\theta)(6 \times 0.75). \end{aligned}$$

Se obtiene un resultado analogo para la varianza.

2.c. Calcula la función generadora de momentos de la distribución mezcla de N.

Siguiendo el mismo razonamiento al inciso (2.b) tenemos

$$M_{\mathbb{P}}(w) = \theta M_{Bin(4,0.5)}(w) + (1-\theta) M_{Bin(6,0.75)}(w).$$

2.d. Calcula $\mathbb{P}(N > n | N \geq 1)$, definiendo el soporte de esta probabilidad condicional.

La distribucion condicional tendra soporte en

$$\mathcal{N}^c = \mathcal{N} \cap \{n : n \geq 1\} = \{1, 2, \ldots\}.$$

La probabilidad de $N \in \mathcal{N}^c$ *es*

$$\mathbb{P}(N \geq 1) = 1 - \mathbb{P}(N = 0) = 1 - (\theta Bin(0|4, 0.5) + (1-\theta) Bin(0|6, 0.75))$$

Asi, la probabilidad condicional tendra la forma

$$\mathbb{P}(N > n | N \geq 1) = \sum_{m=n+1}^{6} \frac{\theta Bin(m|4, 0.5) + (1-\theta) Bin(m|6, 0.75)}{1 - (\theta Bin(0|4, 0.5) + (1-\theta) Bin(0|6, 0.75))} \mathbb{I}_{\mathcal{N}^c}(n).$$

2.e. Desarrolla una discusión acerca de la combinación de soportes que involucra esta mezcla.

R: El soporte del componente $Bin(n|4, 0.5)$ *esta definido en* $\{0, 1 \ldots, 4\}$, *mientras que el soporte del componente* $Bin(n|6, 0.75)$ *esta definido en* $\{0, 1 \ldots, 6\}$.

Ambos componentes se complementan, definiendo un soporte para la mezcla dado por

$$\mathcal{N} = \{0, 1 \ldots, 4\} \cup \{0, 1 \ldots, 6\} = \{0, 1 \ldots, 6\}.$$

Hasta los valores de $N \leq 4$, *ambos componentes contribuyen al calculo de* p_n *proporcionalmente a* θ *y* $(1 - \theta)$. *A partir de los casos* $N \geq 5$ *solo el componente* $Bin(n|6, 0.75)$ *contribuye, con proporcion* $(1 - \theta)$.

Pregunta 3

Considera que la frecuencia de siniestros de un portafolio de seguros exhibe históricamente que la varianza muestral es aproximadamente dos veces la media muestral. Examina y argumenta para los siguientes modelos la relevancia y adecuación para describir la frecuencia de siniestros de este portafolio.

3.a. Mezcla de dos distribuciones binomiales con diferentes medias.

R: Debemos empezar mencionando que la informacion proporcionanda es limitada (tenderiamos a pensar que la informacion es *incompleta* *para un proposito concluyente, pero la* *informacion es la informacion* ***como es y no necesariamente como deberia de ser o nos gustaria que fuera****).*

En el sentido anterior, sabemos que los datos son sobredispersos, por lo que podriamos pensar en una mezcla de la siguiente forma

$$\gamma Bin(n|J,\theta) + (1-\gamma) Bin(n|J',\theta'),$$

al rededor de la cual podemos definir escenarios:

a) $J' = 2J$ con $\theta' = \theta$

b) $J' = J$ con $\theta' = 2\theta$

c) $J' \neq J$ con $\theta' \neq \theta$

3.a. Mezcla de dos distribuciones binomiales con diferentes medias.

R: Debemos empezar mencionando que la informacion proporcionanda es limitada (tenderiamos a pensar que la informacion es incompleta para un proposito concluyente, pero la informacion es la informacion ***como es y no necesariamente como deberia de ser o nos gustaria que fuera****).*

En el sentido anterior, sabemos que los datos son sobredispersos, por lo que podriamos pensar en una mezcla de la siguiente forma

$$\gamma Bin(n|J,\theta) + (1-\gamma)Bin(n|J',\theta'),$$

al rededor de la cual podemos definir escenarios:

a) $J' = 2J$ con $\theta' = \theta$

b) $J' = J$ con $\theta' = 2\theta$

c) $J' \neq J$ con $\theta' \neq \theta$

3.a. Mezcla de dos distribuciones binomiales con diferentes medias.

R: Debemos empezar mencionando que la informacion proporcionanda es limitada (tenderiamos a pensar que la informacion es `incompleta` *para un proposito concluyente, pero la* `informacion es la informacion` ***como es y no necesariamente como deberia de ser o nos gustaria que fuera****).*

En el sentido anterior, sabemos que los datos son sobredispersos, por lo que podriamos pensar en una mezcla de la siguiente forma

$$\gamma Bin(n|J,\theta) + (1-\gamma) Bin(n|J',\theta'),$$

al rededor de la cual podemos definir escenarios:

a) $J' = 2J$ con $\theta' = \theta$

b) $J' = J$ con $\theta' = 2\theta$

c) $J' \neq J$ con $\theta' \neq \theta$

3.b. Mezcla de dos distribuciones Poisson con diferentes tasas de intensidad.

La relevancia del tipo de mezclas estara en funcion de como son especificadas.

3.c. Mezcla de dos distribuciones binomial negativa con diferente media.

Mexclas con dos medias distintas sera informativa para la varianza, dispersion, pero no necesariamente para la media.

La media puede ser `ambigua`*, en el sentido que no sea informativa (a.k.a. posiblemente resultado de distribuciones bimodales).*

3.d. Mezcla de tres distribuciones binomiales negativas con diferentes parámetos de probabilidades.

Lo mismo puede pasar con la varianza.

3.e. Mezcla de dos distribuciones donde una es Poisson y otra es binomial, ambas con la misma media.

Igualmente pasa con la combinacion de soportes.

En este caso, la informacion proporcionada no es concluyente en este sentido.

*En **grandes rasgos**, la informacion proporcionada es incompleta... lo que nos conduce a reflexionar acerca de la **utilidad/necesidad** de **recurperar la distribucion completa** de N.*

Pregunta 4

Demuestra que la distribución modificada en 0 de una variable aleatoria para el número de reclamos de un portafolio de seguros puede ser derivada usando una distribución tipo mezcla con dos componentes.

Supongamos que N *es la frecuencia de siniestros con soporte en* $\mathcal{N} = \{0, 1, 2, \ldots\}$ *para la cual tenemos la coleccion de masas de probabilidades* $(p_n)_{n\in\mathcal{N}}$.

La distribucion modificada en $\{0\}$ *incorpora un* q_0 *adicional.*

Podemos introducir una variable de asignacion Z, *dada por*

$$Z = \begin{cases} 1, & si\ N = 0, \\ 0, & si\ N \neq 0. \end{cases}$$

Pregunta 4

Con las siguientes probabilidades,

$$\begin{aligned}\mathbb{Q}(Z=1) &= q_0,\\ \mathbb{Q}(Z=0) &= (1-q_0).\end{aligned}$$

Asi, la distribucion modificada en $\{0\}$ *para* N *seria expresada como*

$$\mathbb{Q}(N=n) = q_0\mathbb{I}(n=0) + (1-q_0)\frac{p_n}{\sum_{m\in\mathcal{N}\setminus\{0\}} p_m}\mathbb{I}(n\neq 0).$$

Pregunta 4

De esta forma, el componente uno de la mezcla es

$$q_0 \mathbb{I}(n = 0),$$

mientras que el componente dos es

$$(1 - q_0) \frac{p_n}{\sum_{m \in \mathcal{N} \setminus \{0\}} p_m} \mathbb{I}(n \neq 0).$$

Pregunta 4

Los pesos q_0 y $(1 - p_0)$ denotan las dos porporciones de las mezclas.

Los `kernels/nucleos` de las mezclas estan expresados como

$$
\begin{aligned}
\mathbb{Q}(N = n | Z = 1) &= \delta_{\{0\}}(n), \\
\mathbb{Q}(N = n | Z = 0) &= \frac{p_n}{\sum_{m \in \mathcal{N} \setminus \{0\}} p_m}.
\end{aligned}
$$

Table of Contents

Pregunta 1

Pregunta 2

Pregunta 3

Pregunta 4