ACT-11302: Cálculo Actuarial III
Modelo de Probabilidad

Juan Carlos Martínez-Ovando
juan.martinez.ovando[AT]itam.mx

ITAM

Primavera 2018

Modelo de probabilidad: Definición

Definición

Notación

- X es una variable aleatoria observable (discreta o continua)

Modelo

Sin pérdida de generalidad, refirámonos al modelo de probabilidad como la función de distribución de probabilidades de X indizada por θ, i.e.

$$X \sim F(x|\theta). \tag{1}$$

El soporte de X, denotado por $\mathcal{X}$, se define como,

$$\mathcal{X} = \{x : F(x|\theta) > 0\}, \tag{2}$$

para todo θ, donde $\mathcal{X}$ forma un subconjunto de un espacio Euclidiano de dimensión finita.

El parámetro θ, toma valores en el espacio parametral Θ (generalmente de dimensión finita).

Definición

Notación

- X es una variable aleatoria observable (discreta o continua)

Modelo

Sin pérdida de generalidad, refirámonos al modelo de probabilidad como la función de distribución de probabilidades de X indizada por θ, i.e.

$$X \sim F(x|\theta). \tag{1}$$

El soporte de X, denotado por $\mathcal{X}$, se define como,

$$\mathcal{X} = \{x : F(x|\theta) > 0\}, \tag{2}$$

para todo θ, donde $\mathcal{X}$ forma un subconjunto de un espacio Euclidiano de dimensión finita.

El parámetro θ, toma valores en el espacio parametral Θ (generalmente de dimensión finita).

Definición

Densidades y masa de probabilidad

1. Cuando X es absolutamente continua, $F(x|\theta)$ admite una densidad, $f(x|\theta)$ y

$$\Pr(X = x) = 0,$$

para todo $x \in \mathcal{X}$.

2. Cuando X es discreta, $\Pr(X = x^*) > 0$ para algunos x^*s $\in \mathcal{X}$. Los valores de x que satisfacen lo anterior se llaman átomos,

$$\{x_1^*, \ldots, x_K^*\}.$$

3. Cuando X es del tipo mixta, $F(x|\theta)$ admite una parte absolutamente continua al mismo tiempo de admitir átomos, i.e.

$$\Pr(X \leq x) = F(x|\theta) = F_c(x|\theta_c) + \sum_{x_k^* \leq x} p(X = x_k|\theta_d). \tag{3}$$

El **soporte** de X en este caso es típicamente $\mathcal{X} = \Re_+ \cup \{x_1^*, \ldots, x_K^*\}$

Modelo de probabilidad: Verosimilitud

Verosimilitud

Consideremos un conjunto de datos $X_1 = x_1, \ldots, X_n = x_n$ en el caso continuo.

Función de verosimilitud

- Enfoque frecuentista (bajo **independencia estocástica**)

$$\Pr(X_1 = x_1, \ldots, X_n = x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta). \tag{4}$$

- Enfoque bayesiano (bajo **independencia condicional**)

$$\Pr(X_1 = x_1, \ldots, X_n = x_n) = \int \prod_{i=1}^{n} f(x_i|\theta)\pi(\theta)\mathrm{d}\theta. \tag{5}$$

Pregunta

¿Cómo será la expresión de la función de verosimilitud en el caso discreto y tipo mixta?

Verosimilitud

Consideremos un conjunto de datos $X_1 = x_1, \ldots, X_n = x_n$ en el caso continuo.

Función de verosimilitud

- Enfoque frecuentista (bajo **independencia estocástica**)

$$\Pr(X_1 = x_1, \ldots, X_n = x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta). \tag{4}$$

- Enfoque bayesiano (bajo **independencia condicional**)

$$\Pr(X_1 = x_1, \ldots, X_n = x_n) = \int \prod_{i=1}^{n} f(x_i|\theta)\pi(\theta)\mathrm{d}\theta. \tag{5}$$

Pregunta

¿Cómo será la expresión de la función de verosimilitud en el caso discreto y tipo mixta?

Verosimilitud

Tipos de datos

- Las expresiones anteriores son correctas cuando los datos son exactos.
- Cuando trabajamos con datos agrupados en $\Re_+$, modificamos el soporte $\mathcal{X}$ por una partición $\{c_j\}_{j=1}^{J}$ tal que,

$$c_1 < c_2 < \ldots < c_J, \tag{6}$$

sustituyendo $\mathcal{X}$ por el conjunto,

$$\mathcal{C} = \{(c_j, c_{j+1}] : c_j < c_{j+1}, j = 1, \ldots, J\}. \tag{7}$$

Pregunta

¿Cómo será la expresión de la función de verosimilitud para datos agrupados?

Verosimilitud

Tipos de datos

- Las expresiones anteriores son correctas cuando los datos son exactos.
- Cuando trabajamos con datos agrupados en $\Re_+$, modificamos el soporte $\mathcal{X}$ por una partición $\{c_j\}_{j=1}^{J}$ tal que,

$$c_1 < c_2 < \ldots < c_J, \tag{6}$$

sustituyendo $\mathcal{X}$ por el conjunto,

$$\mathcal{C} = \{(c_j, c_{j+1}] : c_j < c_{j+1}, j = 1, \ldots, J\}. \tag{7}$$

Pregunta

¿Cómo será la expresión de la función de verosimilitud para datos agrupados?

Modelo de probabilidad: Predicción

Predicción

Enfoque frecuentista

Bajo el enfoque frecuentista, la predicción de un valor futuro de X, X^f, se obtiene a través de la imputación del EMV de θ en el modelo, i.e.

$$X^f | x_1 \ldots, x_n \sim f(x^f | \widehat{\theta}_n), \tag{8}$$

donde $\widehat{\theta}_n = \widehat{\theta}_n(x_1 \ldots, x_n)$.

Enfoque bayesiano

Bajo el enfoque bayesiano, la predicción se obtiene usando argumentos probabilistas, como

$$p(x^f | x_1 \ldots, x_n) = \int_\Theta f(x^f | \theta) \pi(\theta | x_1 \ldots, x_n) \mathrm{d}\theta, \tag{9}$$

donde $\pi(\theta | x_1 \ldots, x_n) \propto f(x_1 \ldots, x_n | \theta) \pi(\theta)$ es la distribución de θ actualizada con la información contenida en $x_1 \ldots, x_n$.

Predicción

Enfoque frecuentista

Bajo el enfoque frecuentista, la predicción de un valor futuro de X, X^f, se obtiene a través de la imputación del EMV de θ en el modelo, i.e.

$$X^f|x_1 \ldots, x_n \sim f(x^f|\widehat{\theta}_n), \tag{8}$$

donde $\widehat{\theta}_n = \widehat{\theta}_n(x_1 \ldots, x_n)$.

Enfoque bayesiano

Bajo el enfoque bayesiano, la predicción se obtiene usando argumentos probabilistas, como

$$p(x^f|x_1 \ldots, x_n) = \int_\Theta f(x^f|\theta)\pi(\theta|x_1 \ldots, x_n)\mathrm{d}\theta, \tag{9}$$

donde $\pi(\theta|x_1 \ldots, x_n) \propto f(x_1 \ldots, x_n|\theta)\pi(\theta)$ es la distribución de θ actualizada con la información contenida en $x_1 \ldots, x_n$.

Modelo de probabilidad: Intercambiabilidad

Intercambiabilidad

Definición

Se dice que un conjunto (numerable) de variables aleatorias $\{X_j\}_{j=1}^{\infty}$ es intercambiabiable con respecto a Pr si para todo n finito,

$$\mathrm{Pr}(X_1, \ldots, X_n) = \mathrm{Pr}(X_{\sigma(1)}, \ldots, X_{\sigma(n)}), \tag{10}$$

donde $(\sigma(1), \ldots, \sigma(n))$ es cualquier permutación del vector $(1, \ldots, n)$.

Comentarios

- Cualquier sucesión de variables aleatorias iid es naturalmente intercambiable.
- La noción de intercambiabilidad, como la de independencia, se refiere a que el orden de la información es irrelevante (i.e. los resultados analíticos son invariantes ante permutaciones).

Intercambiabilidad

Definición

Se dice que un conjunto (numerable) de variables aleatorias $\{X_j\}_{j=1}^{\infty}$ es intercambiabiable con respecto a $\Pr$ si para todo n finito,

$$\Pr(X_1, \ldots, X_n) = \Pr(X_{\sigma(1)}, \ldots, X_{\sigma(n)}), \tag{10}$$

donde $(\sigma(1), \ldots, \sigma(n))$ es cualquier permutación del vector $(1, \ldots, n)$.

Comentarios

- Cualquier sucesión de variables aleatorias iid es naturalmente intercambiable.
- La noción de intercambiabilidad, como la de independencia, se refiere a que el orden de la información es irrelevante (i.e. los resultados analíticos son invariantes ante permutaciones).

Intercambiabilidad

Representación: de Finetti

Una consecuencia del supuesto de intercambiabilidad (numerable) es el teorema de representación en el que se admite que para toda sucesión de variables aleatorias intercambiables, para toda n finita, se tiene que existe un ente estocástico $\theta \in \Theta$ acompañado de una medida de probabilidad Π, tal que

$$\Pr(X_1, \ldots, X_n) = \int_\Theta \left\{ \prod_{j=1}^{n} \Pr(X_j|\theta) \right\} \Pi(\mathrm{d}\theta), \tag{11}$$

donde $(\sigma(1), \ldots, \sigma(n))$ es cualquier permutación del vector $(1, \ldots, n)$.

Comentarios

- El resultado anterior es de **existencia**, i.e. no nombra cómo se lleva a cabo tal representación.
- Para un conjunto de variables aleatorias intercambiables existen más de una posible representación como la anterior (en términos de diferentes especificaciones de θ y/o de Π).
- Este teorema de representación brinda una interpretación al paradigma bayesiano de inferencia.

Intercambiabilidad

Representación: de Finetti

Una consecuencia del supuesto de intercambiabilidad (numerable) es el teorema de representación en el que se admite que para toda sucesión de variables aleatorias intercambiables, para toda n finita, se tiene que existe un ente estocástico $\theta \in \Theta$ acompañado de una medida de probabilidad Π, tal que

$$\Pr(X_1, \ldots, X_n) = \int_\Theta \left\{ \prod_{j=1}^{n} \Pr(X_j|\theta) \right\} \Pi(\mathrm{d}\theta), \tag{11}$$

donde $(\sigma(1), \ldots, \sigma(n))$ es cualquier permutación del vector $(1, \ldots, n)$.

Comentarios

- El resultado anterior es de **existencia**, i.e. no nombra cómo se lleva a cabo tal representación.
- Para un conjunto de variables aleatorias intercambiables existen más de una posible representación como la anterior (en términos de diferentes especificaciones de θ y/o de Π).
- Este teorema de representación brinda una interpretación al paradigma bayesiano de inferencia.

Modelo de probabilidad: Suma de variables aleatorias

Independencia

Caso discreto

Si N_1 y N_2 denotan dos variables aleatorias independientes con el mismo soporte $\mathcal{N}$ y con la misma funcion de distribucion discreta, entonces la distribucion de la suma

$$N = N_1 + N_2,$$

se calcula como

$$\mathbb{P}(N = n) = \sum_{k \in \mathcal{N}} p_k p_{n-k},$$

donde $p_k = \mathbb{P}(N_j = k)$ para $j = 1, 2$.

Comentarios

- La definicion anterior depende funcamentalmente dle supuesto de independencia. En el caso general, los sumandos $p_k p_{n-k}$ deben reemplazarse por $\mathbb{P}(N_i = k, N_k = n - k)$.
- Pensemos que pasa en el caso **intercambiable**...

Independencia

Caso discreto

Si N_1 y N_2 denotan dos variables aleatorias independientes con el mismo soporte $\mathcal{N}$ y con la misma funcion de distribucion discreta, entonces la distribucion de la suma

$$N = N_1 + N_2,$$

se calcula como

$$\mathbb{P}(N = n) = \sum_{k \in \mathcal{N}} p_k p_{n-k},$$

donde $p_k = \mathbb{P}(N_j = k)$ para $j = 1, 2$.

Comentarios

- La definicion anterior depende funcamentalmente dle supuesto de independencia. En el caso general, los sumandos $p_k p_{n-k}$ deben reemplazarse por $\mathbb{P}(N_i = k, N_k = n - k)$.
- Pensemos que pasa en el caso **intercambiable**...

Independencia

Caso continuo

Si X_1 y X_2 denotan dos variables aleatorias independientes con el mismo soporte $\mathcal{X}$ y con la misma funcion de distribucion (absolutamente) continua, entonces la distribucion de la suma

$$X = X_1 + X_2,$$

se calcula como

$$\mathbb{P}(X \leq x) = \mathbb{P}(X_1 + X_2 \leq x) = \int_{x_1 + x_2 \leq x} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2.$$

Bajo independencia, y suponiendo que $\mathcal{X} = \Re$ se sigue,

$$\begin{aligned}
\mathbb{P}(X \leq x) &= \int_{\Re} \int_{(-\infty, x - x_1)} f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 && (12) \\
&= \int_{\Re} f_{X_1}(x_1) \left\{ \int_{(-\infty, x - x_1)} f_{X_2}(x_2) dx_2 \right\} dx_1 && (13) \\
&= \int_{\Re} f_{X_1}(x_1) F_{X_2}(x - x_1) dx_1. && (14)
\end{aligned}$$