

Inferencia estadística

En este capítulo se verá lo que es una distribución tipo mezcla en general, así como algunas de sus utilidades, para después enfocarnos en la mezcla que nos es de interés, es decir, la mezcla compuesta. Luego se ahondará en cómo estimar los parámetros de las mezclas compuestas a través del método de máxima verosimilitud, y la necesidad de extender la verosimilitud. Por último, se propondrá seleccionar el muestreador de Gibbs para encontrar los parámetros de interés de nuestra mezcla.

Distribuciones tipo mezcla contable

Este tipo de distribución consiste en ponderar un conjunto numerable de vectores aleatorios discretos con el mismo soporte y la misma dimensión con un conjunto de pesos $0 < u(i) < 1$ cuya suma es 1, $u(i)$ puede ser generado por una distribución de probabilidad discreta, por lo que a los pesos también se les conoce como probabilidades, entonces la función de densidad es de la siguiente forma:

$f_x(x) = \sum_{i=1}^n u(i) f_{i_x}(x)$, tal que $0 < u(i) < 1$, $\sum_{i=1}^n u(i) = 1$, y $f_{i_x}(x)$ un vector aleatorio discreto de dimensión $p \forall i$.

Es común encontrar este tipo de distribuciones en un conjunto de datos que provienen de dos o más distribuciones diferentes, y se caracterizan por ser multimodales. Si el conjunto de vectores aleatorios provienen de la misma distribución, entonces se dice que es una mezcla de familias paramétricas, mientras que si el vector aleatorio tienen una distribución continua, entonces se le conoce como densidad tipo mezcla.

Densidad tipo mezcla compuesta

Una densidad tipo mezcla compuesta resulta de considerar que los pesos $u(i)$ están dados por una densidad continua u con soporte en R^+ , a esta densidad se le conoce como variable de mezcla. El vector aleatorio X está parametrizado por (θ, u) , por lo que ahora la familia de vectores aleatorios queda dada por la familia no numerable $X|_u$, tal que u se distribuye $f_u(u)$, de aquí que la densidad de X esté dada por:

$$f_x(x) = \int_u f(x|u) f_u(u) du$$

Si u es una variable aleatoria parametrizada, es común que la distribución de X dependa de los parámetros de u , y además ciertas características como

la dispersión o la cola de la distribución de X también dependerán de la variable de mezcla u .

Este tipo de distribución tiene gran aplicación en estadística bayesiana, pues se considera a u como distribución a-priori, mientras que $f|u = \int_X f(x|u)f_u(u)dx$ como distribución a-posteriori. También en ocasiones resulta mas sencillo obtener características numéricas, o hacer inferencia sobre la parametrización de algún vector aleatorio X de dimensión p , si a este se le considera como una densidad tipo mezcla compuesta, como es el caso de las distribuciones hiperbólicas generalizadas, pues se obtienen a partir de una distribución normal, con una distribución gaussiana inversa como variable de mezcla. Otra de las principales características de este tipo de distribuciones es que son muy flexibles al ajustar datos, siempre y cuando se mezclen las densidades correctas, por lo que en las siguientes secciones se verá el problema de estimación.

De aquí en adelante solamente se trabajará con densidades de tipo mezcla compuesta, refiriéndose a ellas como distribuciones tipo mezcla.

Estimación puntual a través del método de máxima verosimilitud

Entre los distintos métodos de estimación parametral se usará el de MV debido a las propiedades de estos estimadores, y a que actualmente existen métodos computacionales que permiten obtener dichos estimadores sin necesidad de que la distribución del vector aleatorio X tenga una forma cerrada.

Supongamos que tenemos una muestra aleatoria iid de tamaño n del vector aleatorio $X \in R^p$ y que además la distribución de X puede ser estudiada a partir de una distribución tipo mezcla, con variable de mezcla u , entonces como los parámetros de X dependerán de la variable de mezcla u , la propiedad de invarianza de los EMV nos permite hacer la inferencia sobre los parámetros de la distribución de mezcla y la variable de mezcla, por lo que la función de verosimilitud vendría dada por la siguiente expresión:

$$P(X_1 = x_1, \dots, X_n) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n \int_u f(X_i = x_i | \theta, u) f_u(u | \alpha) du$$

Suponiendo que se cumplen las condiciones de regularidad, $Lik(\theta, \alpha) = \int_u \prod_{i=1}^n f(X_i = x_i | \theta, u) f_u(u | \alpha) du$
(Aquí introducir la noción de extender la verosimilitud)

Muestreador de Gibbs

El muestreador de Gibbs tiene un enfoque bayesiano, y consiste en construir una cadena de markov, para así conseguir una muestra de una distribución dada. La construcción de la cadena de Markov es necesaria para asegurar que nuestra muestra proviene efectivamente de la distribución que nos interesaría muestrear.

Supongamos que tenemos una distribución de la forma $f(X, \Theta)$ donde X es un vector aleatorio y Θ algún vector de parámetros del cual se desearía tener una muestra mediante una cadena de markov, es decir, obtener una muestra de $f(\theta^k|X)$ a partir de una cadena θ_1^∞ , tal que su probabilidad de transición converga a la distribución a muestrea, esto es, $k \xrightarrow{Lim} \infty$ $P(\theta^{k+1}|\theta^k, X) = P(\theta|X)$. Dada la descripción del modelo anterior, la cadena debe ser: estacionaria, irreducible, aperiódica y recurrente positiva. La propiedad de estacionariedad asegurar que $P(\theta|X)$ no cambie cuando $k \rightarrow \infty$, irreducible para no descartar ningún valor posible de θ^k en cada iteración, mientras que aperiodicidad para que la cadena no tenga ciclos, y así permanecer en el estado estacionario. Por último, recurrente positiva, para que se pueda visitar cualquier estado un número infinito de veces para así asegurar que la distribución límite tomó toda la información disponible en consideración.

Ahora bien, para encontrar una cadena que nos permita obtener las características anteriores German & German (citar bien esta parte) demuestra que $\lim_{k \rightarrow \infty} P(\theta^k|X) = P(\theta|X)$, donde $P(\theta^k|X) = \prod_{j=1}^t P(\theta_j|X, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_n)$, con $\theta_j|X, \theta_{\theta_j}$ independiente $\forall j$, para alguna medida de probabilidad arbitraria $P(\cdot)$. Entonces si θ_j^k se distribulle $f(\theta_j^k|X, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_n)$, de aquí sería fácil obtener una muestra de θ_j^k para toda j , si su densidad tiene una forma cerrada, y así se construiría una cadena de markov con las características deseadas, $[\theta^k]_1^\infty$, con probabilidad de transición $P(\theta^k|X)$.

Entonces los pasos para implementar el Muestreador de Gibbs son los siguientes:

Dar un conjunto de valores iniciales Θ_0 , no importa que valores se den, la propiedad de estacionariedad, asegurará la convergencia al estado estacionario. El vector θ puede ser particionado en d subconjuntos excluyentes, tal que $d \leq n$, con la finalidad de simplificar los cálculos. Estratégicamente se pueden seleccionar los subconjuntos de modo que los que tienen alta correlación queden juntos.

Para cada θ_j obtener su distribución $f(\theta_j|X, \dots, \theta_{0,j-1}, \theta_{0,j+1}, \dots, \theta_{0,n})$,

y realizar una simulación de esta. El valor de θ_j obtenido reemplazará a los valores iniciales. En caso de que no se pueda tener la distribución de $\theta|_{X, \theta_j}$ se realiza un procedimiento de aceptación rechazo para obtener una realización de θ_j .

Una vez obtenidos todos los valores de $\theta|_{X, \theta_j}$ se reemplazan por los valores iniciales y se repite el procedimiento hasta que la cadena se estabilice. computacionalmente el algoritmo se vería de la siguiente forma:

```

 $\Theta_k < -(\theta_{1,0}, \dots, \theta_{d,0})$ 

dowhile  $||\theta^{k+1} - varaux < \epsilon||$ 

 $varaux < -\theta_k$ 

for (j in 1:d)

 $\theta_j^{k+1} < - \text{random} (f(\theta|_{X, \theta^k_j}))$ 

end do while

```