



## **Análisis de múltiples puntos de cambio temporales con el factor de Bayes**

**Bulmaro Juárez Hernández<sup>a</sup>, Lucila Muñoz Merino<sup>b</sup>**

<sup>a,b</sup> *Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias Físico Matemáticas, Puebla, Puebla, México.*

<sup>a</sup>[bjuaraz@cfm.buap.mx](mailto:bjuaraz@cfm.buap.mx), <sup>b</sup>[lucymerino74@hotmail.com](mailto:lucymerino74@hotmail.com).

### **Resumen**

En este trabajo se hace un análisis de puntos de cambio con el factor de Bayes. Para el análisis de puntos de cambio con tal factor se simularon datos Poisson y se utilizaron para hacer divisiones en el método de bisección.

**Palabra claves:** Factor de Bayes, Poisson y método de bisección.

---

### **1. Introducción**

Cuando se escucha decir punto de cambio la primera pregunta que surge es ¿qué es un punto de cambio? Chen y Gupta (2012) lo definen como el sitio o punto en el tiempo  $t$ , en una sucesión de datos  $\{x_{t_i}\}$ ,  $i = 1, \dots, n$  observados y ordenados respecto al tiempo tal que dichas observaciones siguen una distribución  $F_1$ , antes de un punto, y en otro posterior la distribución a seguir es  $F_2$ . Es decir, desde el punto de vista estadístico, la sucesión de observaciones muestra un comportamiento no homogéneo. El problema de punto de cambio es considerado como uno de los problemas centrales de inferencia estadística, pues relaciona a la teoría de control estadístico, a las pruebas de hipótesis (al detectar si existe algún cambio en la sucesión de variables aleatorias observadas), y a la teoría de estimación (al estimar el número de cambios y sus correspondientes localizaciones). Esto bajo

los enfoques clásico y Bayesiano, según Chen y Gupta (2012).

Chen y Gupta (2012) indican que los problemas de puntos de cambio originalmente surgieron en control de calidad y en general pueden ser encontrados en la modelación matemática de diversas disciplinas tales como Medio Ambiente, Epidemiología, Procesos de señal sísmica, Economía, Finanzas, Geología, Medicina, Biología, Física, etc.

En general el problema de puntos de cambio según Chen y Gupta (2012) se visualiza de la forma siguiente:

Sean  $X_1, X_2, \dots, X_n$  una colección de vectores (variables) aleatorios independientes con funciones de distribución de probabilidad  $F_1, F_2, \dots, F_n$ , respectivamente. Entonces el problema de puntos de cambio consiste en probar la hipótesis nula  $H_0$  de la no existencia de

cambio contra la alternativa  $H_a$  de que existe al menos un punto de cambio lo cual se expresa de la siguiente manera:

$$H_0: F_1 = F_2 = \dots = F_n$$

vs

$$H_a: F_1 = \dots = F_{k_1} \neq F_{k_1+1} = \dots = F_{k_2} \neq F_{k_2+1} \\ \dots = F_{k_q} \neq F_{k_q+1} = \dots = F_n$$

donde  $1 < k_1 < k_2 < \dots < k_q < n$ ,  $q$  es el número desconocido de puntos de cambio y  $k_1, k_2, \dots, k_q$  son las posiciones desconocidas respectivas que tienen que ser estimadas. Si las distribuciones  $F_1, F_2, \dots, F_n$  pertenecen a una familia paramétrica común  $F(\theta)$ , donde  $\theta \in \mathbb{R}^p$ , entonces el problema de puntos de cambio consiste en probar la hipótesis nula  $H_0$  sobre la no existencia de cambio en los parámetros  $\theta_i, i = 1, \dots, n$  de la población contra la alternativa  $H_a$  de que existe al menos un punto de cambio; lo cual se expresa de la siguiente forma:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_n = \theta \text{ (desconocido)}$$

vs

$$H_a: \theta_1 = \dots = \theta_{k_1} \neq \theta_{k_1+1} = \dots = \theta_{k_2} \neq \theta_{k_2+1} \\ \dots = \theta_{k_q} \neq \theta_{k_q+1} = \dots = \theta_n$$

donde  $q$  y  $k_1, k_2, \dots, k_n$  tienen que ser estimados. Estas hipótesis revelan los aspectos de inferencia de puntos de cambio para determinar si existe algún punto de cambio en el proceso, estimar el número de ellos y sus respectivas posiciones.

El objetivo de este trabajo es desarrollar un programa para aplicar algunos procedimientos en la detección de los puntos de cambio, en particular con factor de Bayes. También se utiliza el proceso Poisson homogéneo, ya que se simulan datos con este proceso.

Se detectaron con el factor de Bayes múltiples puntos de cambio temporales. Para este método

se tomó como base los conceptos presentados en Altieri (2015).

En este trabajo se elaboró un programa, el cual se hizo con el objetivo de detectar los puntos de cambio, se analizan y comparan sus resultados. Fue creado con algunas instrucciones del paquete INLA de R presentada en los libros Gómez (2020) y Blangiardo y Cameletti (2015) para ello se usaron las librerías INLA y Poisson.

Se utilizó una base de datos simulada de un proceso Poisson, con la cual se detectaron los cambios temporales con el factor de Bayes.

## 2. Materiales y métodos

### 2.1 Factor de Bayes

Si se presenta un problema de selección de modelos en el que se debe elegir entre dos posibles modelos, en base a un conjunto de datos observados  $D$ , la plausibilidad de los dos diferentes modelos  $M_1$  y  $M_2$ , parametrizados por vectores de parámetros  $\theta_1$  y  $\theta_2$  se puede medir mediante el factor Bayes.

El factor de Bayes se define como:

$$B = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\int_{\theta_1} P(D|\theta_1, M_1)\Pi(\theta_1|M_1)d\theta_1}{\int_{\theta_2} P(D|\theta_2, M_2)\Pi(\theta_2|M_2)d\theta_2}$$

donde  $P(D|M_1)$  se denomina verosimilitud marginal o verosimilitud integrada. Esto es similar a lo que se hace en las pruebas de la razón de verosimilitudes, pero ahora, en lugar de maximizar la verosimilitud, el factor Bayes realiza un promedio ponderado mediante la distribución de los parámetros.

Un valor de  $B > 1$  significa que  $M_1$  es apoyado por los datos más que  $M_2$ . En el caso del factor de Bayes, Jeffreys estableció una escala de interpretación de  $B$ , la cual se muestra en la Tabla 1.

**Tabla 1.-** Escala de interpretación de B, según Jeffreys.

B	Fuerza de la evidencia a favor de $M_1$
$B \leq 1$	Negativa apoya $M_2$
$1 < B \leq 3$	Muy escasa
$3 < B \leq 10$	Sustancial
$10 < B \leq 30$	Fuerte
$30 < B \leq 100$	Muy fuerte
$> 100$	Decisiva

Otra forma de considerar el factor de Bayes es la siguiente:

Supóngase dos hipótesis  $H_0$  y  $H_a$  tales que las probabilidades de las distribuciones a priori son  $f_0 = P(H_0)$  y  $f_1 = P(H_1)$ . Después de observar una muestra  $\mathbf{x}$ , las probabilidades a posteriori de ambas hipótesis son  $\alpha_0 = P(H_0|\mathbf{x})$  y  $\alpha_1 = P(H_1|\mathbf{x})$ . Se define el factor de Bayes a favor de  $H_0$  como

$$B = \frac{\frac{f_0}{\alpha_0}}{\frac{f_1}{\alpha_1}} = \frac{\alpha_1 f_0}{\alpha_0 f_1}.$$

Así, el factor de Bayes representa la plausibilidad a posteriori dividida entre la plausibilidad a priori. Nos informa de los cambios en nuestras creencias introducidas por los datos. Tiene la propiedad de que es casi objetivo y elimina parcialmente la influencia de la distribución a priori.

Como ejemplo supóngase el contraste simple:

$$H_0: \theta = \theta_0 \text{ vs } H_1: \theta = \theta_1.$$

Tenemos que las distribuciones a posteriori son:

$$\alpha_0 = P(H_0|\mathbf{x}) = \frac{f_0 L(\theta_0|\mathbf{x})}{f_0 L(\theta_0|\mathbf{x}) + f_1 L(\theta_1|\mathbf{x})}.$$

Y

$$\alpha_1 = P(H_1|\mathbf{x}) = \frac{f_1 L(\theta_1|\mathbf{x})}{f_0 L(\theta_0|\mathbf{x}) + f_1 L(\theta_1|\mathbf{x})}.$$

Entonces el factor de Bayes es:

$$B = \frac{\alpha_1 f_0}{\alpha_0 f_1} = \frac{f_1 L(\theta_1|\mathbf{x}) f_0}{f_0 L(\theta_0|\mathbf{x}) f_1} = \frac{L(\theta_1|\mathbf{x})}{L(\theta_0|\mathbf{x})},$$

que coincide con la razón de verosimilitudes, de modo que la distribución a priori no influiría en este caso en el factor de Bayes.

Así el factor de Bayes para el punto de cambio cuando se divide en dos segmentos está dado por la razón de verosimilitudes:

$$\frac{L_1}{L_0} = \frac{Q_1 Q_2}{L_0},$$

donde  $Q_1$  es la verosimilitud del segmento 1 y  $Q_2$  es la verosimilitud del segmento 2 bajo la hipótesis alternativa y  $L_0$  es la verosimilitud bajo la hipótesis nula.

Aplicando el logaritmo se tiene

$$\log\{B\} = \log\{Q_1\} + \log\{Q_2\} - \log\{L_0\}.$$

### 2.1.1. Proceso de Poisson homogéneo

Para encontrar los puntos de cambio primero se simula un proceso Poisson homogéneo y se trabaja los datos obtenidos, éste está definido de la siguiente forma:

**Definición 1:** Una colección de variables, aleatorias  $\{N(t): t \geq 0\}$  definidas en un espacio de probabilidad  $(\omega, F, P)$  se dice ser un proceso de Poisson (homogéneo) con intensidad  $\lambda > 0$  si satisface las siguientes propiedades:

- i)  $P(N(0) = 0) = 1$
- ii) Para todo  $0 < s < t$ ,  $N(t) - N(s)$  tiene distribución de Poisson de parámetro  $\lambda(t - s)$ .
- iii) Para todo  $0 \leq t_1 < \dots < t_n$ ,  $n \geq 1$  (es decir para todo conjunto finito de

tiempos), las variables aleatorias  $N(t_n) - N(t_{n-1}), \dots, N(t_2) - N(t_1), N(t_1) - N(0)$ , son independientes. Esta propiedad se conoce como propiedad de incrementos independientes.

### 2.1.2. Método de bisección

Para determinar los puntos de cambio con el factor de Bayes se utilizó el método de bisección. Éste consiste en dividir los datos a la mitad y se busca un punto de cambio con el factor de Bayes y las verosimilitudes correspondientes a la hipótesis nula y la alternativa, enseguida se va a la izquierda y se divide también a la mitad, posteriormente se divide el lado derecho y así se sigue sucesivamente yendo a la izquierda y a la derecha dividiendo a la mitad y buscando puntos de cambio con el factor de Bayes.

## 3. Resultados y discusión

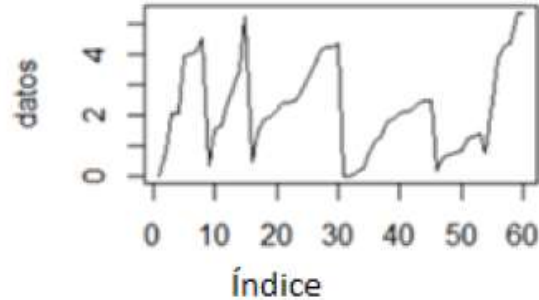
### 3.1 Análisis de múltiples puntos de cambio temporales con el factor de Bayes

Con el objetivo de detectar, analizar y comparar resultados de puntos de cambio se hizo un programa para detectar 5 puntos de cambio, los datos utilizados fueron simulaciones de un proceso Poisson obteniéndose un conjunto de 60 datos con 6 valores diferentes para el parámetro  $\lambda$ . Se pretende en este problema detectar los cambios que se generan con los distintos valores de  $\lambda$ , además, se hizo un comparativo de los resultados al utilizar las a priori uniforme, log-gamma y Gaussiana.

Se corrió un programa en R, para detectar múltiples puntos de cambio temporales, en total para 5, se utilizaron los 60 datos simulados y el segmento más pequeño fue de 4 puntos. Así, se simuló un proceso Poisson uniforme con diferentes parámetros  $\lambda$  y se aproximó a la distribución a posteriori con el paquete de R INLA, el cual aproxima con series de Taylor.

Para detectar los puntos de cambio se utilizó el factor de Bayes y el método de bisección.

Se utilizó una distribución a priori uniforme, log-gamma y una Gaussiana. La Figura 1, muestra 5 puntos de cambio.



**Figura 1.-** Gráfica de 5 puntos de cambio.

Para el siguiente problema se usan los 60 datos y se llega a la división más pequeña que fue aquella que sólo contiene cuatro datos, es decir se tienen quince divisiones (Tabla 5).

**Tabla 5.-** Puntos de cambio.

Num	uniforme	Log gamma	Gaussiana
1	<b>49.565011</b>	<b>43.644575</b>	<b>42.972975</b>
2	<b>24.392282</b>	<b>25.829232</b>	<b>26.592785</b>
3	<b>13.940830</b>	<b>26.041694</b>	<b>28.329433</b>
4	<b>16.000633</b>	<b>9.990938</b>	<b>11.641769</b>
5	12.478724	2.387813	4.477951
6	12.478724	2.387813	4.477951
7	<b>12.572475</b>	<b>4.498905</b>	<b>5.523642</b>
8	5.701123	1.835601	5.801758
9	4.732525	1.384310	4.912561
10	5.701123	1.835601	5.801758
11	4.732525	1.384310	4.912561
12	5.701123	1.835601	5.801758
13	4.732525	1.384310	4.912561
14	5.701123	1.835601	5.801758
15	4.732525	1.384310	4.912561

Los puntos de cambio para una distribución a priori uniforme con  $\lambda = 2, 1, 4, 7, 6, 1$  y en las divisiones de datos o segmentos 8, 7, 15, 15, 8, 7, esto debido a que en los segmentos dividiendo de dos en dos para 30 datos, quedo en un segmento 15 y en el otro 15, al dividir 15 quedo

en un segmento 8 datos y en el otro 7. Los resultados para la detección de los puntos de cambio fueron como se muestra en la Tabla 5. Los puntos de cambio son los primeros cuatro y el séptimo.

Con los mismos puntos de cambio y con la distribución a priori log-gamma con parámetros 0.01 y 0.01, el resultado se muestra en la Tabla 5. Se puede observar que en comparación con la distribución uniforme las log de verosimilitudes de la log-gamma son más pequeñas, sin embargo, también detectan los puntos de cambio.

Con la distribución a priori Gaussiana con media cero y parámetro de precisión 0.001 para el mismo número de puntos de cambio como en la distribución uniforme y la log-gamma se detectaron los cinco puntos de cambio, se puede observar en los datos dados a continuación en la Tabla 5, que son los primeros cuatro y el siete, aunque el valor de la log-verosimilitud del séptimo punto de cambio es menor en comparación con las siguientes posiciones como se puede observar.

Los 5 puntos de cambio que se observan en la Tabla 5, fueron detectados con el factor de Bayes B mencionado anteriormente, los valores para B indican la fuerza de la evidencia a favor de la hipótesis del punto de cambio, valores menores a 1 indican que no existe punto de cambio. Los datos en negrita que son los puntos de cambio se detectaron de la siguiente forma: el valor de 1 fue detectado a la mitad de los datos, el valor 2 en el dato número 15, el valor 3 en el dato 45, el número 4 en el dato número 8 y el 7 en el dato 53, como se puede observar los valores en negrita de los puntos de cambio no son iguales entre sí ni con el resto, en cambio en donde no sucedió un punto de cambio los valores son iguales, por ejemplo los valores 5 y 6 son iguales; el 8,10 y 14 también son iguales; el 9, 11, 13 y 15 son iguales, por lo que todos estos valores están indicando que no existen puntos de cambio. La fuerza de la evidencia a favor de la hipótesis alternativa indica según los puntos de

cambio que al aplicar el método de bisección e ir detectando los puntos de cambio, la evidencia es fuerte en el primer punto de cambio y va disminuyendo en los siguientes, obsérvese esto en la tabla 5.

Los puntos de cambio se muestran en la Figura 1 y son en 8, 15, 30, 45 y 53.

### **3.1.1. Seis puntos de cambio con el factor de Bayes**

Con el mismo modelo uniforme se detectaron 6 puntos de cambio para  $\lambda = 2, 1, 4, 7, 6, 1, 2$  y en las divisiones de datos o segmentos de 8, 7, 15, 15, 8, 4, 3.

Los puntos de cambio que se detectaron son los primeros cuatro, el séptimo y el último. Se muestran estos en las Tablas 6 y 7.

Con la distribución a priori el logaritmo de gamma, para los 6 puntos de cambio que son los 4 primeros, el siete y el último, también se detectaron los 6 puntos de cambio, la diferencia con la uniforme es que el log de la verosimilitud es más pequeño, como se puede observar en la Tabla 6.

Para la distribución a priori Gaussiana también se detectaron los 6 puntos de cambio, se muestran enseguida en valores de log verosimilitud, es un poco parecida a la uniforme. Estos son los puntos de cambio, los cuatro primeros, el siete y el último, se muestran enseguida en la Tabla 6. La diferencia entre la Gaussiana y la uniforme es que el último punto de cambio es menor a los valores de log verosimilitud anteriores, en donde no hay punto de cambio.

Con la distribución logit-beta con parámetros 2 y 2 se detectaron 6 puntos de cambio, y son los mismos que para la distribución uniforme y la distribución log-gamma, los puntos de cambio se detectan bien, son los cuatro primeros, el

séptimo y último. No se visualizan grandes diferencias con respecto a la uniforme.

Para la distribución a priori normal truncada con parámetros media 0 y precisión 0.001, la verosimilitud es más pequeña comparada con la uniforme, esta es muy parecida a la log-gamma, sin embargo, también se detectaron 6 puntos de cambio, los cuatro primeros, el siete y el último, el resultado está en las Tablas 6 y 7.

**Tabla 6.-** 6 puntos de cambio.

<i>Num</i>	uniforme	Log gamma	Gaussiana
1	<b>37.382628</b>	<b>44.717137</b>	<b>38.258157</b>
2	<b>23.107912</b>	<b>28.913661</b>	<b>26.506732</b>
3	<b>22.987741</b>	<b>27.851420</b>	<b>26.529322</b>
4	<b>16.466681</b>	<b>5.045601</b>	<b>11.798824</b>
5	12.478724	2.387813	4.477951
6	12.478724	2.387813	4.477951
7	<b>16.555598</b>	<b>3.062915</b>	<b>12.572375</b>
8	5.701123	1.835601	5.801758
9	4.732525	1.384310	4.912561
10	5.701123	1.835601	5.801758
11	4.732525	1.384310	4.912561
12	5.701123	1.835601	5.801758
13	4.732525	1.384310	4.912561
14	5.701123	1.835601	5.801758
15	<b>4.733385</b>	<b>1.720949</b>	<b>4.280097</b>

De forma similar como en el problema anterior se presentan los puntos de cambio que se observan en las tablas 6 y 7, estos valores no se repiten entre ellos ni entre el resto. En cambio, el resto de los valores si se repiten, se puede observar por ejemplo que los valores 5 y 6 son iguales; el 8, 10, 12 y 14 también son iguales; el 9, 11 y 13 son iguales, por lo tanto, todos estos valores aun cuando sean mayores que 1 no indican punto de cambio, pues se están repitiendo.

Los puntos de cambio se muestran en la Figura 2, los cuales se detectan en 8, 15, 30, 45, 53 y 57.

El programa se hizo en R, es etiquetado como Programa 1 y se presenta al final en el Apéndice A.

**Tabla 7.-** 6 puntos de cambio.

<i>Num</i>	Logit-beta	Normal truncada
1	<b>49.664317</b>	<b>51.841263</b>
2	<b>30.567246</b>	<b>16.016904</b>
3	<b>18.174484</b>	<b>33.658036</b>
4	<b>9.944739</b>	<b>7.969972</b>
5	4.633706	2.186421
6	4.633706	2.186421
7	<b>11.248933</b>	<b>1.984581</b>
8	4.076750	1.881876
9	3.794214	1.783849
10	4.076750	1.881876
11	3.794214	1.783849
12	4.076750	1.881876
13	3.794214	1.783849
14	4.076750	1.881876
15	<b>4.405551</b>	<b>2.260447</b>



**Figura 2.-** Gráfica de 6 puntos de cambio.

#### 4. Conclusiones

Se elaboro un programa para el método de factor de Bayes, la finalidad fue detectar numéricamente los puntos de cambio y comparar los resultados de las distribuciones a priori utilizadas.

El método del factor de Bayes detectó bien los puntos de cambio, la debilidad de este es que detecta cambios hasta una división que contiene cuatro datos, para menos datos en la división ya no se detectan los puntos de cambio.

## Referencias

Altieri Linda. (2015) A Bayesian changepoint analysis on spatio-temporal point processes, Tesis.

Blangiardo Marta y Cameletti Michela. (2012) Spatial and Spatio-temporal Bayesian Models with R-INLA, John Wiley and Sons.

Chen, J., y Gupta, A. (2012). Parametric Statistical Changepoint Analysis. Bogota: Birkhauser.

Gómez Rubio Virgilio. (2020). Bayesian inference with INLA, Champman and Hall.

## Apéndice A

### Programa 1

```
rm(list=ls())
library(poisson)
library(INLA)
vector<-hpp.event.times(2, 15, num.sims = 1, t0 = 0)
plot(vector)
vector1<-hpp.event.times(1, 15, num.sims = 1, t0 = 0)
vector2<-hpp.event.times(4, 15, num.sims = 1, t0 = 0)
vector4<-hpp.event.times(6, 15, num.sims = 1, t0 = 0)
datos<-c(vector,vector1,vector2,vector4)
plot(datos,type="l")
datosA=datos[1:30]
datosB=datos[31:60]
datos
datosA
datosB
vec_res=rep(0,20)
factor_b<-function(datosA,datosB,datos){
p1<-
data.frame("num"=seq(1,length(datosA),1),"datosA"=datosA)
p2<-
data.frame("num"=seq(1,length(datosB),1),"datosB"=datosB)
mp1<-inla(num~f(datosA,
model="ar1"),data=p1,family="poisson")
mp2<-inla(num~f(datosB,
model="ar1"),data=p2,family="poisson")
midf <- data.frame( "num" = seq(1,length(datos),1), "datos" =
datos)
mp<-inla(num~f(datos,
model="ar1"),data=midf,family="poisson")
respuesta<- as.numeric(mp1$mlik[1,1])
+as.numeric(mp2$mlik[1,1])-as.numeric(mp$mlik[1,1])
return(respuesta) }
```

```
vec_res[1]=factor_b(datosA,datosB,datos)
vec_res
datosC=datosA[1:15]
datosD=datosA[16:30]
vec_res[2]=factor_b(datosC,datosD,datosA)
vec_res
datosE=datosB[1:15]
datosF=datosB[16:30]
vec_res[3]=factor_b(datosE,datosF,datosB)
vec_res
datosG=datosC[1:8]
datosH=datosC[9:15]
vec_res[4]=factor_b(datosG,datosH,datosC)
vec_res
datosI=datosD[1:8]
datosJ=datosD[9:15]
vec_res[5]=factor_b(datosI,datosJ,datosD)
vec_res
datosK=datosE[1:8]
datosL=datosE[9:15]
vec_res[6]=factor_b(datosK,datosL,datosE)
vec_res
datosM=datosF[1:8]
datosN=datosF[9:15]
vec_res[7]=factor_b(datosM,datosN,datosF)
vec_res
datosO=datosG[1:4]
datosP=datosG[5:8]
vec_res[8]=factor_b(datosO,datosP,datosG)
vec_res
datosQ=datosH[1:4]
datosR=datosH[5:7]
vec_res[9]=factor_b(datosQ,datosR,datosH)
vec_res
datosS=datosI[1:4]
datosT=datosI[5:8]
vec_res[10]=factor_b(datosS,datosT,datosI)
vec_res
datosU=datosJ[1:4]
datosV=datosJ[5:7]
vec_res[11]=factor_b(datosU,datosV,datosJ)
vec_res
datosX=datosK[1:4]
datosY=datosK[5:8]
vec_res[12]=factor_b(datosX,datosY,datosK)
vec_res
datosW=datosL[1:4]
datosZ=datosL[5:7]
vec_res[13]=factor_b(datosW,datosZ,datosL)
vec_res
datosAA=datosM[1:4]
datosBB=datosM[5:8]
vec_res[14]=factor_b(datosAA,datosBB,datosM)
vec_res
datosCC=datosN[1:4]
datosDD=datosN[5:7]
vec_res[15]=factor_b(datosCC,datosDD,datosN)
vec_res
```