



## ALGORITMO REGRESIVO DE INTEGRACIÓN NUMÉRICA.

**Ing. Reinier Fernández López<sup>1</sup>**

Universidad de Pinar del Río  
[rflopez@upr.edu.cu](mailto:rflopez@upr.edu.cu)

**MSc. Ledy Raúl Díaz González<sup>2</sup>**

Universidad de Pinar del Río  
[ledy@upr.edu.cu](mailto:ledy@upr.edu.cu)

Para citar este artículo puede utilizar el siguiente formato:

Reinier Fernández López y Ledy Raúl Díaz González (2018): "Algoritmo regresivo de integración numérica", Revista Caribeña de Ciencias Sociales (noviembre 2018). En línea

[//www.eumed.net/rev/caribe/2018/11/algoritmo-regresivo.html](http://www.eumed.net/rev/caribe/2018/11/algoritmo-regresivo.html)

### RESUMEN

En el presente trabajo se expone la aplicación de un algoritmo regresivo de integración numérica en la Empresa de Producción Agroindustrial "Enrique Troncoso" de Pinar del Río. Se crea como una herramienta propia orientada al área de la Ingeniería de la Calidad. Se utiliza como Entorno de Desarrollo Integrado (IDE) R Studio y el software libre R, lenguaje de programación y asistente matemático de actualidad, novedoso para el trabajo estadístico como ciencia, en combinación con elementos de la Matemática Numérica. El objetivo es mostrar la funcionalidad práctica de este nuevo procedimiento, resultado de la combinación de cualquier método de integración y optimización numéricas con el funcional.

**PALABRAS CLAVE:** calidad-algoritmo-integración numérica-probabilidad.

In the present work, the application of a regressive numerical integration algorithm in the "Enrique Troncoso" Agroindustrial Production Company of Pinar del Río is exposed. It is created as a proprietary tool oriented to the area of Quality Engineering. It is used as an Integrated Development Environment (IDE) R Studio and free software R, programming language and current mathematical assistant, novel for statistical work as a science, in combination with elements of Numerical Mathematics. The objective is to show the practical functionality of this new procedure, resulting from the combination of any numerical integration and optimization method with the functional one.

### ABSTRACT

In the present work, the application of a regressive numerical integration algorithm in the "Enrique Troncoso" Agroindustrial Production Company of Pinar del Río is exposed. It is

<sup>1</sup>Ingeniero Industrial. Profesor Dpto. de Matemáticas. Facultad de Ciencias Técnicas. Universidad de Pinar del Río.

<sup>2</sup> Master en Ciencias Pedagógicas. Ingeniero en Telecomunicaciones y Electrónica. Profesor Dpto. de Matemáticas. Facultad de Ciencias Técnicas. Universidad de Pinar del Río.

created as a proprietary tool oriented to the area of Quality Engineering. It is used as an Integrated Development Environment (IDE) R Studio and free software R, programming language and current mathematical assistant, novel for statistical work as a science, in combination with elements of

KEYWORDS: quality-algorithm-numerical-integration-probability.

## 1. INTRODUCCIÓN

El algoritmo regresivo de integración numérica se utilizó en Empresa de Producción Agroindustrial “Enrique Troncoso” de Pinar del Río, la empresa constituye una de las fuentes económicas fundamentales en el desarrollo alimenticio de esta provincia. Su función principal es la producción de alimentos agroindustriales para el consumo humano, entre estos productos se encuentran la barra de guayaba, encurtidos, mermeladas y siropes de distintos sabores. Por estas características se seleccionó para la aplicación de un algoritmo matemático creado con el propósito de introducirse como una nueva herramienta para el control y análisis de las diferentes características de calidad, modelando mediante funciones matemáticas el comportamiento productivo.

Pues, desde el punto de vista de la ingeniería y de la gestión de la calidad no existe en esta ciencia un algoritmo matemático para el análisis funcional de las características de calidad (López, 2017, p.32).

El proceso seleccionado fue la fabricación de sirope, por ser actualmente el de mayor impacto en la economía de la empresa según el departamento de contabilidad de la institución. La característica seleccionada fue los grados *Brix*<sup>3</sup> por ser la variable de mayor incidencia en la calidad del producto sirope. Como apoyo a dicho trabajo se utilizó el lenguaje R y el IDE R Studio.

R, es un software o más bien un lenguaje de comandos de manipulación y análisis estadístico basado en el lenguaje estadístico S desarrollado por AT&T, con la diferencia de que R es un programa de código abierto y gratis, muy popular en los ámbitos académicos (Bande et al., 2008, p. 11). Dadas sus características, R tiene gran potencial para ser usado en la educación pública, pues al no pagarse el software, puede ser instalado en diversos sistemas operativos: IOS de MAC, Linux o Windows, entre otros.

El lenguaje R como herramienta de análisis estadístico es uno de los paquetes informáticos cuyo uso ha crecido de manera más acelerada en los últimos años, tiene como una de las principales ventajas que cuenta con una amplia comunidad de expertos que continuamente introducen innovaciones, así como proveen de útiles materiales de consulta disponible gratuitamente en internet (Sulmont, 2015, p. 5).

Quizás el mayor inconveniente que ven en un principio los nuevos usuarios de R es su interfaz gráfica, ‘poco amigable’, en el sentido de que hay que programar las funciones, pues a diferencia de Windows no hay botones o ventanas que despliegan menús donde el usuario puede elegir opciones (Araya et al., 2012, p. 1).

R, es un sistema para el análisis estadísticos y gráficos creado por Ross Ihaka y Robert Gentleman. Tiene una naturaleza doble de programa y lenguaje de programación y es considerado como un dialecto del lenguaje S (Paradis, 2003, p.4).

En la investigación se utilizó el Método de Simpson de integración numérica y el Método de Bisección para la optimización numérica. Conformando un algoritmo que permite la programación de la función que describe los datos a utilizar. Todos estos algoritmos son programados en R, utilizando como soporte de programación al software R Studio 1.0 como IDE.

## 2. ELEMENTOS DEL ALGORITMO REGRESIVO DE INTEGRACIÓN NUMÉRICA

---

<sup>3</sup> Los grados *Brix* describen la cantidad de sólidos solubles en agua, expresando niveles de azúcares en el producto terminado.

Se hace imprescindible comentar breve y teóricamente el algoritmo matemático para su posterior comprensión y aplicación. Se descompone en tres partes fundamentales: método de integración numérica, método de optimización numérica y método funcional de los datos con su respectiva función de densidad probabilística. Es válido aclarar que cualquier método de integración y de optimización numérica se puede utilizar para conformar el algoritmo en cuestión.

### 2.1 El método de integración numérica

En esta investigación se implementa el método de integración numérica de Simpson descrito por Álvarez et al., (2004), el cual plantea que la obtención de la fórmula de Simpson requiere conocer el valor de la integral de una función cuadrática determinada por tres puntos  $(x_0, y_0)$ ,  $(x_1, y_1)$  y  $(x_2, y_2)$  donde  $x_1 - x_0 = h$  y  $x_2 - x_1 = h$  como muestra figura 1 (p. 316). El resultado, naturalmente, no cambia si toda la figura se traslada de modo que  $x_1 = 0$ , de esta manera el trabajo algebraico se simplifica (p. 314).

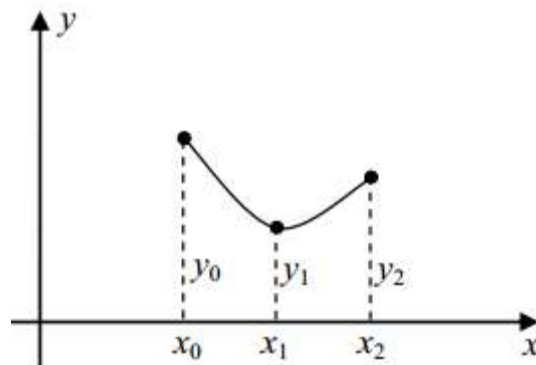


Figura 1: Función cuadrática determinada por tres puntos.  
Fuente: Álvarez et al., (2004),

Sea entonces  $p(x)$  el polinomio de segundo grado cuya gráfica aparece en la figura 2 (p. 316), determinado por los tres puntos  $(-h, y_0)$ ,  $(0, y_1)$  y  $(h, y_2)$ , cuya ecuación será de la forma:  
 $p(x) = ax^2 + bx + c$  (1)

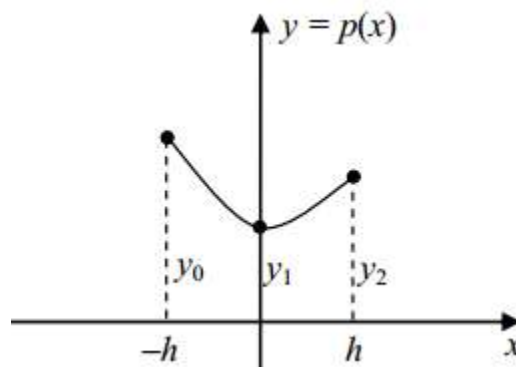


Figura 2: Polinomio de segundo grado  
Fuente: Álvarez et al., (2004),

Aplicando ahora la integral de  $f(x)$  en  $[a, b]$  y la propiedad aditiva respecto al intervalo de integración resulta:

$$\int_a^b f(x)dx = \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \dots + \int_{x_{n-2}}^{x_n} f(x)dx$$

Sustituyendo en cada integral la función en  $[a, b]$  por el polinomio interpolador de segundo grado que corresponde a esos nodos, la integral resulta:

$$\int_a^b f(x)dx \approx \frac{1}{3}h(y_0 + 4y_1 + y_2) + \frac{1}{3}h(y_2 + 4y_3 + y_4) + \dots + \frac{1}{3}h(y_{n-2} + 4y_{n-1} + y_n) \quad (1)$$

Que se simplifica como:

$$\int_a^b f(x)dx \approx \frac{1}{3}h(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n) \quad (2)$$

O, más brevemente:

$$\int_a^b f(x)dx \approx \frac{1}{3}h(E + 4I + 2P) \quad (3)$$

Donde:

$$E = y_0 + y_n \text{ (Suma de las ordenadas en los extremos)} \quad (4)$$

$$I = y_1 + y_3 + \dots + y_{n-1} \text{ (Suma de las ordenadas de índices impares)} \quad (5)$$

$$P = y_2 + y_4 + \dots + y_{n-2} \text{ (Suma de las ordenadas de índices par)} \quad (6)$$

El algoritmo que sigue calcula aproximadamente la integral definida de la función continua  $f(x)$  en el intervalo  $[a,b]$  mediante el Método de Simpson, tomando un número  $n$  de subintervalos. Se supone conocida la función  $f(x)$ , en el intervalo  $[a,b]$  y el número de enteros  $n$ . El código en R es:

*# Método de Simpson*

```
> simpson=function(b) {
+   a=0
+   n=1000
+   h=(b-a)/n
+   E=f(a)+f(b)
+   l=0
+   x=0
+   i=1
+   while(i<n) {
+     x=a+i*h
+     l=l+f(x)
+     i=i+2
+   }
+   P=0
+   i=2
+   while(i<n) {
+     x=a+i*h
+     P=P+f(x)
+     i=i+2
+   }
+   integral=h*(E+4*l+2*P)/3
+   integral
+ }
```

## 2.2 El método de optimización numérica

Según Álvarez et al., (2004) describe que si en un problema de optimización unidimensional se conoce un intervalo donde se encuentra el punto máximo de la función unimodal, el problema puede resolverse de modo más efectivo que una búsqueda secuencial. El intervalo puede estar dado desde un inicio a partir de consideraciones físicas, económicas, etc...O puede llegarse a

él a partir de una etapa previa de búsqueda secuencial (a partir de un conjunto de datos) (p. 376).

Sea  $f(x)$  una función unimodal definida en  $[a, b]$  y con un punto máximo en  $x^*$  en ese intervalo. Tómesese dos puntos experimentales  $x_1$  y  $x_2$  muy próximos entre sí y a ambos lados del centro del intervalo  $[a, b]$ . Para concretar, sea  $\delta$  la distancia entre  $x_1$  y  $x_2$ , entonces, como se aprecia en la figura 3 (p. 377).

$$x_1 = \frac{a+b}{2} - \frac{\delta}{2} \tag{7}$$

$$x_2 = \frac{a+b}{2} + \frac{\delta}{2} \tag{8}$$

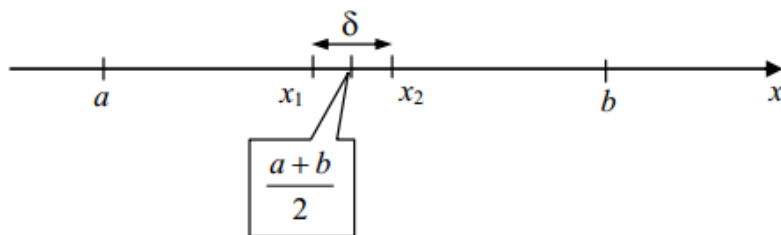


Figura 3: distancia entre  $x_1$  y  $x_2$   
Fuente: Álvarez et al., (2004),

El procedimiento se ilustra en la figura 4 (p. 377):

$$y_1 < y_2 \Rightarrow x_1 \leq x^* \leq b) \quad y_1 > y_2 \Rightarrow a \leq x^* \leq x_2) \quad y_1 = y_2 \Rightarrow x_1 \leq x^* \leq x_2)$$

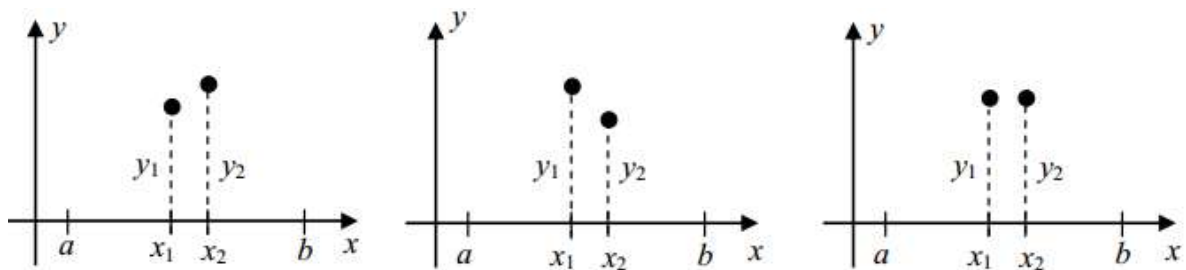


Figura 4: Propiedad básica de la optimización unidimensional  
Fuente: Álvarez et al., (2004),

Si consideramos que  $\delta$  es un número muy pequeño en relación con el amplitud del intervalo, puede suponerse que en cada iteración (requiere una evaluación de  $f(x)$ ) la longitud del intervalo de búsqueda se reduce a la mitad. Más formalmente, llamando:

$L_n$  : Amplitud del intervalo de búsqueda de  $n$  evaluaciones (  $n$  par)

$$L_0 = b - a \tag{9}$$

$$\text{Entonces: } L_n = \frac{L_0}{2^{n/2}} \tag{10}$$

Lo cual prueba que  $L_n$  converge hacia 0 cuando  $n$  tiende hacia el infinito. En la práctica,  $L_n$  no puede reducirse a valores menores que  $\delta$ .

Se supone que la función  $f(x)$  es unimodal con máximo en el intervalo  $[a, b]$  y está definida en todos los puntos del intervalo. El algoritmo da como resultado un intervalo cerrado de amplitud menor que un número especificado  $\varepsilon$  y que contiene el punto de máximo. Los datos que se requieren son: la función, la distancia  $\delta$  que separa los puntos experimentales y la tolerancia  $\varepsilon$ , que determina la amplitud del intervalo final.

*#Método de Bisección*

```
> bisección=function(i) {
+   a=0
+   b=72
+   e=0.0000001
+   x=(a+b)/2
+   while (b-a>e) {
+     if(simpson(x)>i) {
+       b=x
+     } else {
+       a=x
+     }
+     x=(a+b)/2
+   }
}
```

### 2.3 El método de optimización numérica

En este método solo se describe la función de densidad  $f(x)$  computacional o numéricamente o sea mediante un algoritmo, especificando todo el comportamiento de la misma en el intervalo  $[a, b]$  como sus restricciones, etc. A continuación, se muestra dicho código en R.

```
> f=function(x){
# Aquí se introduce la función a utilizar
# En el caso de estudio se utiliza una función cúbica
}
```

### 3. DEFINICIÓN TEÓRICA DEL MÉTODO REGRESIVO DE INTEGRACIÓN NUMÉRICA.

El algoritmo regresivo de integración numérica es un conjunto de métodos modificados, con el objetivo de encontrar un extremo de una integral, para poder inferir un análisis funcional probable en cuanto al comportamiento de la variable seleccionada.

Sea una integral  $\int_a^b f(x)dx = I$  donde  $f(x)$  es una función de densidad cualquiera, definida en el intervalo  $[a, b]$  sin primitiva en términos de funciones elementales. El problema radica en cómo conociendo el extremo  $a$  y el parámetro  $I$ , se puede hallar  $b$ .

Por tanto, planteando que  $P(x)$  es una función de probabilidad, figura 5, conformada por los puntos  $I_n$ , generada por cualquier método de integración mediante la combinación del método de bisección, dado el parámetro  $I_0$  y conociendo un extremo, entonces es posible hallar  $x_0$  con un error épsilon ( $\varepsilon$ ) como condición de parada. Siendo esta la descripción teórica para dar paso al algoritmo.

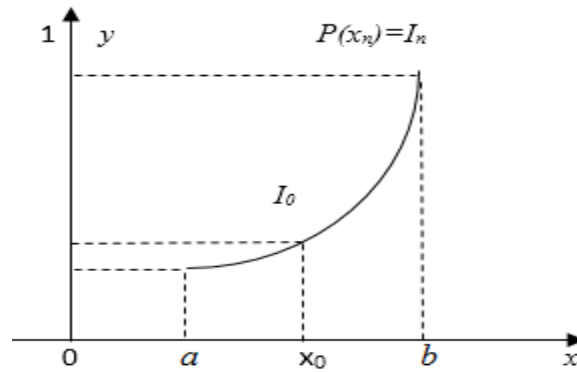


Figura 5: Gráfica de probabilidad.  
Fuente: Elaboración propia,

En el siguiente código se describe el algoritmo regresivo de integración numérica, el cual se aporta como una nueva herramienta matemática, numérica y computacional a la Ingeniería de la Calidad.

```
# Método funcional
> f=function(x) {
# Aquí se introduce la función a utilizar
}
# Método de integración numérica
> simpson=function(b) {
+ a=0
+ n=1000
+ h=(b-a)/n
+ E=f(a)+f(b)
+ l=0
+ x=0
+ i=1
+ while(i<n) {
+ x=a+i*h
+ l=l+f(x)
+ i=i+2
+ }
+ P=0
+ i=2
+ while(i<n){
+ x=a+i*h
+ P=P+f(x)
+ i=i+2
+ }
+ integral=h*(E+4*l+2*P)/3
+ integral
+ }
```

```
# Método de optimización numérica
```

```
> bisección=function(i) {  
+ a=0  
+ b=72  
+ e=0.0000001  
+ x=(a+b)/2  
+ while (b-a>e) {  
+   if(simpson(x)>i) {  
+     b=x  
+   } else{  
+     a=x  
+   }  
+   x=(a+b)/2  
+ }
```

#### 4. APLICACIÓN DEL ALGORITMO REGRESIVO DE INTEGRACIÓN NUMÉRICA A LA CARACTERÍSTICA DE CALIDAD GRADOS BRUX MEDIANTE EL SOFTWARE R STUDIO.

Para la aplicación del algoritmo de integración numérica a la característica de calidad grados Brix se seleccionó una muestra aleatoria de 72 datos, para formular una expresión matemática que describe el comportamiento mediante el método de regresión lineal múltiple utilizando el comando `lm()` según Cayuela (2015) la salida de R Studio fue la siguiente:

```
# x: variable que contiene los grados Brix
```

```
# xx: variable x cuadrática
```

```
# xxx: variable xxx cúbica
```

```
# lm: función que llama al modelo lineal
```

```
> lm(y~x+xx+xxx)
```

```
Call:
```

```
lm(formula = y ~ x + xx + xxx)
```

```
Coefficients:
```

(Intercept)	x
5.688e+01	-7.442e-02
xx	xxx
1.486e-03	-1.016e-05

Con esta salida, se confeccionó la función cúbica que describe el comportamiento de la variable Brix:

$$brix(y) = -0.00001x^3 + 0.001486x^2 - 0.07442x + 56.88 \quad (11)$$

Como es lógico se comprueban los supuestos teóricos para verificar la calidad de la curva por lo que se le realiza la prueba de normalidad y de homogeneidad de varianzas como sigue a continuación.

```
> lillie.test(x)
```

```
Lilliefors (Kolmogorov-Smirnov) normality test
```

```
data: x
```



$D = 0.059$ ,  $p\text{-value} = 0.15$

La homogeneidad de varianzas se puede comprobar en la figura 6 la cual muestra los residuos del modelo contra los valores predichos.

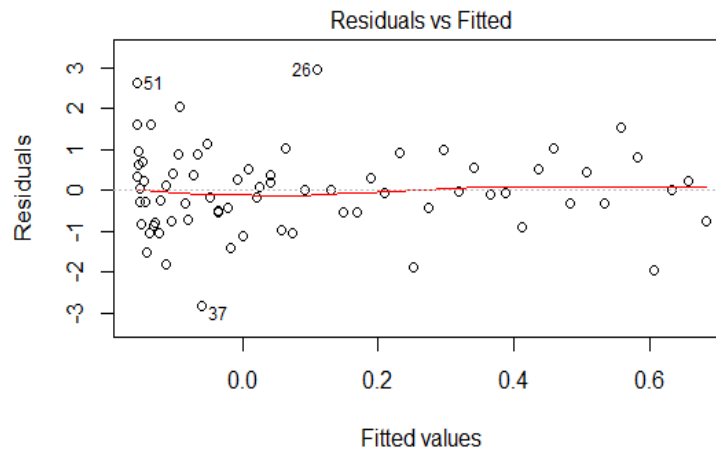


Figura 6: Gráfica de residuos contra valores predichos.  
Fuente: Elaboración propia.

La figura 7 muestra el gráfico de dispersión con su respectiva curva ajustada, la cual se obtuvo mediante los siguientes comandos:

```
> plot(y,xlab = "x",ylab = "brix",ylim = c(53,59))  
> par(new=T)  
> matplot(56.88-0.07442*x+0.001486*xx-0.00001*xxx,type = "l",ylab = "brix",ylim = c(53,59))
```

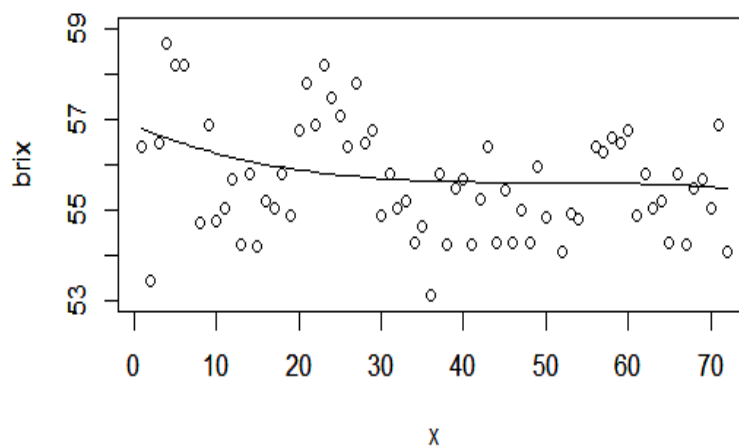


Figura 7: Gráfica de dispersión con modelo cúbico ajustado.  
Fuente: Elaboración propia

Como se analizó, es necesario buscar un coeficiente para obtener una función de densidad probabilística que nos permita realizar el estudio de los datos de manera probable. Utilizando el método de Simpson para integrar la función cúbica (1) se obtuvo la siguiente salida de R Studio:

```
> simpson(72)  
[1] 4020.161
```

Ahora se puede ejecutar el algoritmo regresivo de integración numérica con el objetivo de determinar ¿cuál es valor de calidad Brix con el cual podemos sustentar una determinada probabilidad?

La empresa de Cítricos de Pinar del Río requiere saber cuál sería el valor de Brix que se puede lograr si ellos garantizan una probabilidad de cumplir con ese valor más del 90% de las veces. Las salidas de R Studio fueron las siguientes:

```
# llamada del método de bisección ya programado
```

```
> bisección(0.90)
```

```
[1] 64.76164
```

```
# llamada de la función
```

```
> f(0)
```

```
[1] 56.88
```

Este resultado significa que dicha empresa puede garantizar que el 90% de las veces la característica de calidad logrará valores entre 56.88 y 64.76 grados Brix, saliéndose de los parámetros de calidad establecidos por la industria productiva que son 56.4 más menos 0.2 grados Brix. Se recomienda observar detalladamente el proceso productivo para ver cuáles operaciones son las causantes de este problema de calidad.

## 5. CONCLUSIONES

El algoritmo regresivo de integración numérica comprende o combina cualquier método de integración y de optimización numérica para formar un nuevo procedimiento, herramienta o método con el objetivo de analizar datos o conjuntos de datos a partir de la descripción de los mismo mediante una función matemática de probabilidad, la cual permite tomar decisiones respecto a las variables que se valoran. Esta herramienta de probada utilidad muestra los resultados en la calidad y puede ser empleada para el estudio y mejores resultados en los procesos de esta índole en la empresa elegida y similares.

## 6. REFERENCIAS

- Álvarez, M., Guerra, A., & Lau, R. (2004). *Matemática Numérica. 2da Edición*.
- Araya, J. A. Z., Brenes, R. T. V., & Acuña, M. F. (2012). *Aprendiendo estadística con R. VIII FESTIVAL INTERNACIONAL DE MATEMÁTICA, 7 al 9 de junio de 2012. Sede Chorotega, Universidad Nacional, Liberia, Costa Rica*.
- Bande, M. F., Miguel, P. G. S., Díaz, J. G., & López, B. P. (2008). *Prácticas de Estadística en R*.
- Cayuela, L. (2015). *Modelos lineales: Regresión, ANOVA y ANCOVA*.
- López, R. F. (2017). *Aplicación del Algoritmo Regresivo de Integración Numérica a la Característica de Calidad Grados Brix en la Empresa de Cítricos de Pinar del Río.*, Universidad de Pinar del Río, Pinar del Río.
- Paradis, E. (2003). *R para Principiantes. Institut des Sciences de l'Evolution Universit Montpellier II, F-34095 Montpellier cedex 05, France*.
- Sulmont, D. (2015). *Estadística para las Ciencias Sociales con R. Pontificia Universidad Católica del Perú*.