un modelo estadístico para evaluar cambios en un proceso de control de cantidad en productos alimenticios

OSCAR E. HERNANDEZ *

RESUMEN

En este artículo se desarrolla un modelo estadístico para describir la distribución de los pesos de paquetes de un producto alimenticio después de haber sido sometidos a un nuevo proceso de control de cantidad. Se reportan los resultados obtenidos al ajustar el modelo a un conjunto de datos obtenidos de una empresa de la ciudad de Birmingham, Inglaterra.

DESCRIPCION DE LOS DATOS

Se tienen datos sobre pesos de paquetes de un producto alimenticio producido en una fábrica mediante tres líneas de producción. En una cierta etapa los paquetes se llenan del producto alimenticio por medio de una máquina automática. Posteriormente aquellos que son encontrados faltos de peso, descoloridos o no satisfactorios por alguna otra razón son rechazados. Dos turnos de inspectores realizan el anterior procedimiento. Por otra parte pesadores entrenados registran el peso de los paquetes aceptados como desviaciones de un peso estándar, según se muestra en el Cuadro No. 1.

La fábrica consideró que la desviación promedio regalada era mayor de lo necesario por lo que decidió introducir un proceso de control de cantidad para aumentar la rentabilidad.

Se estudiaron tres períodos, dos de ellos correspondían al nuevo y viejo procesos, y el otro al período de transición durante el cual los inspectores se acostumbrarían al nuevo proceso.

DESCRIPCION DEL PROBLEMA ESTADISTICO

Para comparar los procesos de control de cantidad en diferentes períodos, líneas de producción y turnos es necesario conocer la distribución de los pesos de los paquetes aceptados.

Para obtener la distribución de los pesos de los paquetes se hacen dos supuestos:

i) una distribución normal $n(w\ ;\ \mu\ ,\ \sigma\ )$ para los pesos (w) de los paquetes antes de ser inspeccionados.

ii) P (rechazar un paquete de peso w) =

$$= \begin{cases} e^{-\gamma\,(w-\delta)} & , w > \delta \\ 1 & , w \leqslant \delta \end{cases}$$

donde $\gamma > 0$ y δ es igual al peso muestral mínimo.

CUADRO No. 1. Resumen de registro de peso del producto.

Pesador .... Hoja No. ....

Fecha .... Turno ....

Línea ....

Rango de pesos	Frecuencia	Punto medio	Peso total
7 a 7,9			
6 a 6,9	5	6,5	32,5
5 a 5,9	15	5,5	82,5
4 a 4,9	45	4,5	202,5
3 a 3,9	71	3,5	248,5
2 a 2,9	142	2,5	355,0
1 a 1,9	134	1,5	201,0
0 a 0,9	78	0,5	39,0
– 1,0 a – 0,1	5	– 0,5	– 2,5
– 2,0 a – 1,1	5	– 1,5	– 7,5
– 3,0 a – 2,1			
– 4,0 a – 3,1			
– 5,0 a – 4,1			
– 6,0 a – 5,1			
– 7,0 a – 6,1			
Total	500		1151,0
Promedio			2,3

* Profesor asociado de la Escuela de Estadística, Universidad de Costa Rica.

El efecto del segundo supuesto es la introducción de una modificación en la distribución de los pesos de los paquetes cuando salen de la máquina automática.

Combinando las dos anteriores densidades mediante el teorema de Bayes la distribución modificada es:

$$h(w;\mu,\sigma,\gamma) = \frac{\left(1-e^{-\gamma(w-\delta)}\right)\dfrac{1}{\sigma\sqrt{2\pi}}\,e^{-1/2\left(\frac{w-\mu}{\sigma}\right)^2}}{\displaystyle\int_{\delta}^{\infty}\left(1-e^{-\gamma(w-\delta)}\right)\dfrac{1}{\sigma\sqrt{2\pi}}\,e^{-1/2\left(\frac{w-\mu}{\sigma}\right)^2}}$$

$$= 0 \qquad , w \leqslant \delta$$

El denominador, $D(\mu,\sigma,\gamma)$ se puede expresar como:

$$D(\mu,\sigma,\gamma) = \Phi\left(\frac{\mu-\delta}{\sigma}\right) - e^{\gamma(\delta-\mu)+\frac{\gamma^2\sigma^2}{2}}\,\Phi\left(\frac{\mu-\gamma\sigma^2-\delta}{\sigma}\right)$$

donde ϕ (z) es la densidad normal estándar y ϕ (z) es la función de distribución acumulativa.

Tenemos así que:

$$h(w;\mu,\sigma,\gamma) = \frac{\left(1-2^{-\gamma(w-\delta)}\right)\Phi\left(\frac{w-\mu}{\sigma}\right)}{\sigma D} \qquad w > \delta$$

$$= 0 \qquad w \leqslant \delta$$

El modelo involucra entonces 4 parámetros. Usaremos el método de máxima verosimilitud para estimar μ, σ y γ, Para estimar δ usaremos un método ad hoc, i.e tomar el peso de la muestra.

Un caso especial e interesante del anterior modelo se obtiene haciendo tender γ a cero. La distribución límite es:

$$h(w;\mu;\sigma) = \frac{(w-\delta)\,\phi\left(\frac{w-\mu}{\sigma}\right)}{\sigma\left[\sigma\,\phi\left(\frac{\mu-\delta}{\sigma}\right) + (\mu-\delta)\,\Phi\left(\frac{\mu-\delta}{\sigma}\right)\right]} \qquad w > \delta$$

$$= 0 \qquad w \leqslant \delta$$

Pareciera que lo anterior no tiene sentido ya que hacer γ igual a cero, significaría rechazar todos los paquetes. Sin embargo, si consideramos una probabilidad de rechazo proporcional al peso del paquete:

$$P(\text{aceptar un paquete de peso } w) = \alpha(w-\delta) \qquad , w > \delta$$

$$= 0 \qquad , w \leqslant \delta$$

el teorema de Bayes daría la anterior distribución límite. Una interpretación para $\gamma = 0$ sería la de que los que inspeccionan los pesos no están realmente discriminando entre paquetes livianos o no sino solamente rechazando menos paquetes pesados que livianos de una manera aproximadamente lineal.

EL METODO DE MAXIMA VEROSIMILITUD

Para obtener las estimaciones máximo-verosímiles de $\theta = (\mu,\sigma,\gamma)$ hay que maximizar la función de verosimilitud L definida por:

$$e^L = c\prod_{i=1}^{k} P_i^{n_i} = c\prod_{i=1}^{k}\left[\int_{a_i}^{a_{i+1}} h(w;\mu,\sigma,\gamma)\,dw\right]^{n_i}$$

donde P_i es la probabilidad de que una observación caiga en la clase i, n_i es el número de observaciones en la muestra del tamaño n, c es una constante combinatoria y las a_i son los límites de clase.

Tenemos entonces que

$$e^L = c\prod_{i=1}^{k}\int_{a_i}^{a_i+1}\left[\frac{\left(1-e^{-\gamma(w-\delta)}\right)\Phi\left(\frac{w-\mu}{\sigma}\right)dw}{\sigma D(\mu,\sigma,\gamma)}\right]^{n_i}$$

o equivalentemente:

$$L = \log c + \sum_{i=1}^{k} n_i \log\int_{a_i}^{a_{i+1}}\left(1-e^{-\gamma(w-\sigma)}\right)\Phi\left(\frac{w-\mu}{\sigma}\right)dw - n\log\sigma - n\log D(\mu,\sigma)$$

Las ecuaciones de verosimilutd:

$$\frac{\partial L}{\partial \mu} = \frac{\partial L}{\partial \sigma} = \frac{\partial L}{\partial \gamma} = 0,$$

no son susceptibles de solución por métodos que no sean numéricos. Un método corriente es el método de Newton o una de sus conocidas variaciones. El método de Newton es iterativo y establece una sucesión θ_n que generalmente converge a θ, esto es:

$$\theta_n = \theta_{n-1} - \left[L''(\theta_{n-1})\right]^{-1} \nabla L(\theta_{n-1})$$

donde $L''(\theta_n)$ representa al Hessiano de la función de verosimilitud evaluado en θ_n, $\nabla L(\theta_n)$ es el gradiente evaluado en θ_n y $\theta = (\mu, \sigma, \gamma)$. Es posible a menudo obtener una aproximación inicial θ_o, considerando rasgos especiales del problema, con la cual iniciar el proceso iterativo. En nuestro problema μ y γ pueden aproximarse con $\overline{x}$ y s.

El método de Newton puede modificarse reemplazando el Hessiano de L por — I, donde I es la matriz de Información, asumiendo que $L''(\theta_n)$ está cercano a su valor esperado. Los elementos de la matriz de información son $E\left(\frac{\partial^2 L}{\partial\theta_i \partial\theta_j}\right)$, donde $\theta_1 = \mu$, $\theta_2 = \sigma$ y $\theta_3 = \gamma$.

De la matriz de información se puede obtener la matriz de varianza–covarianza y por lo tanto los errores de las estimaciones con los cuales se pueden comparar los procesos de control de cantidad para diferentes períodos, turnos y líneas de producción.

RESULTADOS

Se escogieron para estudio del turno A varias líneas de producción y períodos (A21, A23, A33 y A43 donde el primer número indica la línea y el segundo indica período). Para el turno B se escogieron B41, B23, B33 y B43. Estas escogencias se hicieron para propósitos comparativos con otro estudio que analizó los mismos datos pero con otros supuestos *(2)*.

En el ajuste de los datos surgieron problemas de convergencia en el modelo que involucraba a los parámetros μ, σ y γ, no así en el modelo límite obtenido haciendo γ tender a O. Para los detalles se puede consultar *(1)*.

Ambos modelos explicaron mejor los datos que otro modelo desarrollado por N. Intaravicha *(2)* y ajustado a los mismos datos. Sin embargo, el ajuste no fue tan bueno como se esperaba. El ajuste podría mejorarse si se consideran otros factores en la probabilidad de rechazo, la influencia de los distintos tiempos muestreados, y la racionalidad del supuesto de normalidad para la distribución de los pesos antes de ser inspeccionados.

El proceso de modelar situaciones nuevas es siempre difícil; el contar con modelos alternativos sin lugar a dudas facilita esta ardua labor, y esta es la contribución de este artículo al presentar una nueva experiencia en modelación.

LITERATURA CONSULTADA

1. Hernández Rodríguez, O.E. A new model in the investigation of weights of a food product. Universidad de Birmingham, Inglaterra. 1974. Tesis de MSc.

2. Intaravicha, N. Maximun likelihood estimation in the investigation of the weights of a food product. Universidad de Birmingham, Inglaterra. 1973. Tesis de MSc.

3. Silvey, S.D. Statistical Inference, Penguin Books Ltd, Library of University Mathematics.