

Método para la estimación del número de muestras para la cuantificación de un principio industrial

Ing. Mariano A. Rivara – LADIAC S.A.

Como laboratorio de control de calidad industrial nuestra tarea consiste no sólo en la cuantificación de un principio industrial sino también en el análisis estadístico de la cantidad de replicados imprescindibles para asegurar un error máximo determinado. Es por este motivo que en diferentes oportunidades recibimos requerimientos de clientes quienes nos solicitan ayuda a fin de establecer el número de muestras que es necesario analizar en un lote de productos a fin de garantizar un nivel de error en el resultado estadísticamente sustentable.

El planteo de una estrategia que, al mismo tiempo, brinde una solución estadísticamente válida y minimice los costos asociados requiere, primariamente, de la identificación y valoración de las variables involucradas.

Para cumplimentar el paso anterior, se requiere contar con una batería mínima de datos estadísticamente independientes, obtenidos a partir de un estudio piloto o bien de datos históricos, a fin de poder calcular los indicadores básicos de caracterización poblacional, a saber: Promedio (\bar{X}), Desvío estándar (SD) y Coeficiente de Variación (CV).

Cabe aclarar, en este punto, que dichos estimadores estadísticos, comprenden tanto, dado su origen, a la variación debida al desempeño del método de análisis como a la debida a la variabilidad poblacional; presentando entonces una aproximación totalmente realista al momento de la estimación del número de muestras necesarias para conocer el valor de la variable en estudio con un nivel de error dado.

Las variables intervinientes en el modelo matemático son las siguientes:

CV: Coeficiente de Variación (obtenido a partir de los resultados del estudio piloto o de datos históricos, expresado como porcentaje).

t: Estadístico de Student para un test a doble cola y $p = 0,05$ con $(R-1)$ grados de libertad; adonde R es el número de muestras consideradas para el cálculo de los estimadores estadísticos (estudio piloto o datos históricos).

E: Error Total porcentual que se requiere para la medición.

N: Total de unidades muestrales que conforman la población en estudio

n: Número de unidades muestrales a analizar

Así, el desarrollo del cálculo correspondiente al modelo planteado es el siguiente:

En el siguiente ejemplo, podremos apreciar la aplicación del modelo propuesto.

Una empresa fabrica un producto sobre el cual debe evaluar la cantidad de un principio activo.

Dicha empresa produce 85 lotes por mes, y cada uno de ellos se compone de 5500 envases.

Se ha analizado, con anterioridad, la concentración del principio activo en 7 muestras de lotes distintos, por el mismo método y en el mismo laboratorio.

El cliente, entonces, quiere saber cuantos envases de cada lote debe analizar para determinar el contenido de activo con un error menor al 10% y cuantos lotes por mes debe analizar para mantener el mismo criterio de error en toda la producción del mes

Dado éste requerimiento, haremos algunas distinciones:

- Cada envase (de cada lote), será la unidad muestral.
- 5500 será el total poblacional para el primer caso en cuestión.
- 85 será el total poblacional para el segundo caso en cuestión

Bajo este planteo, la resolución comienza con el cálculo de los estimadores estadísticos de la población a partir de los datos históricos:

Muestra	Principio Activo [%]
1	15,4
2	13,5
3	14,0
4	14,0
5	13,4
6	15,3
7	11,7
Promedio	13,9
Desvío Estándar	1,16
CV	8,37

Hecho esto, procedemos a la aplicación del modelo propuesto a fin de calcular el número de muestras (n) que deben analizarse de cada lote de (N) envases.

Parámetro	CV	t	E	N	n
Concentrado	8,37	2,447	10	5500	4,2

Concluyendo entonces que se necesitan analizar 5 unidades muestrales (primer número entero mayor a n) obtenidas al azar de cada lote a fin de poder evaluar la concentración del principio activo con un error menor al 10%.

Ahora procedemos a la aplicación del modelo propuesto a fin de calcular el número de lotes (n) que deben analizarse del total de la producción del mes de (N) lotes.

Parámetro	CV	t	E	N	n
Concentrado	8,37	2,447	10	85	4,0

Concluyendo entonces que se necesitan analizar 4 unidades muestrales (lotes) seleccionados al

azar del total de lotes producidos durante un mes, a fin de poder evaluar la concentración del principio activo con un error menor al 10%.

Vinculando ambas conclusiones, a fin de obtener un error menor al 10% en la evaluación de la concentración del principio activo en toda la producción del mes (467.500 envases producidos en 85 lotes de 5500 envases cada uno), el cliente deberá analizar un total de 20 envases obtenidos de manera aleatoria en grupos de 5 unidades, cada uno de 4 lotes seleccionados al azar.

Veamos ahora la influencia del nivel de error aceptable en el método propuesto, si bajo las mismas condiciones, se desea reducir el error al 5%, el planteo sería el siguiente:

Número de envases (n) a analizar de cada lote de (N) envases

Parámetro	CV	t	E	N	n
Concentrado	8,37	2,447	5	5500	16,7

Numero de lotes (n) a analizar del total de (N) lotes mensuales

Parámetro	CV	t	E	N	n
Concentrado	8,37	2,447	5	85	14,0

Como se puede apreciar al vincular ambas conclusiones, de manera análoga al caso anterior, el total de muestras a analizar es mucho mayor (238 muestras tomadas aleatoriamente en grupos de 17 de 14 lotes seleccionados aleatoriamente) dado el aumento del nivel de ajuste requerido.

Como aclaración final, podemos citar que puede realizarse un ajuste de la cantidad de datos mínimos necesarios para el cálculo de los estimadores poblacionales (X y DE) aplicando el método del "promedio corrido" (o acumulado) hasta obtener un Coeficiente de Variación estable.

La realización de este ajuste permite tomar un número de datos previos tal que el CV resultante sea el menor posible (y el mas representativo de la población), reduciendo entonces al mínimo el número de muestras necesarias. Este ajuste puede ser significativo en poblaciones con alto niveles de dispersión de la variable en estudio.

Bibliografía:

YAMANE, Taro. 1997, Elementary Sampling Theory. 3 Ed. Harla Editores.

Mostacedo, Bonifacio; Fredericksen, Todd S. 2000. Manual de Métodos Básicos de Muestreo y Análisis en Ecología Vegetal. 20-23. BOLFOR.

Gy, Pierre. 1998. Sampling for Analytical Purposes. Wiley:Ed.

Smith, P. L. 2001. A Primer for Sampling Solids, Liquids, and Gases. SIAM.