CAPITULO 19

Muestreo del trabajo y estimación estructurada

El muestreo del trabajo es una técnica para determinar, mediante muestreo estadístico y observaciones aleatorias, el porcentaje de aparición de determinada actividad.

1. Necesidad del muestreo del trabajo

El muestreo del trabajo (conocido también por «muestreo de actividades», «método de observaciones instantáneas», «método de observaciones aleatorias» y «control estadístico de actividades») es una técnica que, como su nombre indica, se basa en el muestreo. Veamos ante todo por qué resulta necesaria.

Para obtener una visión completa y exacta del tiempo productivo y del tiempo inactivo de todas las máquinas en una zona dada de producción, sería necesario observar continuamente cada una de las máquinas de dicha zona y registrar el momento y la causa de cada interrupción. Pero es algo evidentemente imposible de realizar, a menos que una multitud de trabajadores se dedicaran exclusivamente a esa tarea, lo que sería absurdo en la práctica.

Sin embargo, si fuera posible observar de una ojeada qué hace cada máquina de una fábrica en determinado momento, quizá se descubriera que, por ejemplo, 80 por ciento de las máquinas están funcionando y 20 por ciento están paradas. Si se hiciera lo mismo veinte veces más a distintas horas del día, y si cada vez la proporción de máquinas que estuviera funcionando fuera de 80 por ciento, podría decirse con cierta seguridad que en todo momento hay 80 por ciento de las máquinas en funcionamiento.

Como generalmente tampoco es posible aplicar esta técnica, hay que optar por la que le sigue en orden de preferencia: se hace una serie de recorridos del taller a intervalos aleatorios observando las máquinas que funcionan, las que están paradas y la causa de cada inmovilización. He aquí la base de la técnica de **muestreo del trabajo**. Si el tamaño de la muestra es suficientemente grande y las observaciones se efectúan realmente al azar, existe una buena probabilidad de que dichas observaciones reflejen la situación real, con un margen determinado de error por exceso o por defecto.

2. Algunas palabras sobre el muestreo

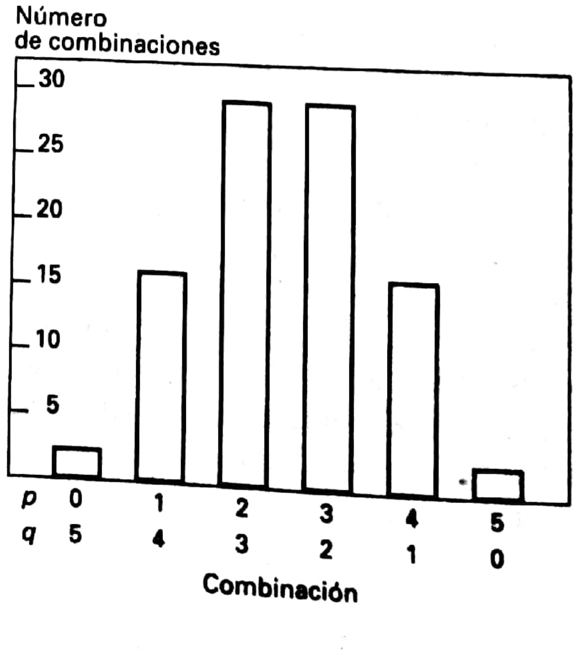
A diferencia del costoso y poco práctico método de observación continua, el muestreo del trabajo se basa principalmente en la **ley de probabilidades**. La probabilidad se ha definido como «el grado de posibilidad de que se produzca un acontecimiento». El ejemplo más sencillo, y frecuentemente mencionado para ilustrar esta idea, es el juego de cara y cruz con una moneda. Cuando lanzamos una moneda al aire pueden suceder dos cosas: que salga «cara» o que salga «cruz». La ley de probabilidades dice que de cada 100 veces que la lancemos, es probable que 50 veces salga cara y 50 cruz. Obsérvese la expresión «es probable que»; en realidad puede suceder que el resultado sea, por ejemplo, 55-45, 48-52 o cualquier otra proporción. Sin embargo, está demostrado que al aumentar el número de lanzamientos aumenta la exactitud de la ley de probabilidades. En otras palabras, cuanto mayor sea el número de lanzamientos de la moneda, tanto mayores serán las posibilidades de llegar a una proporción de 50 caras y 50 cruces. De ello se desprende que cuanto mayor sea la muestra, más exactamente representará la «población» o «universo» inicial, es decir, el grupo de factores que se están estudiando.

Ahora podemos imaginar una escala en la cual uno de los extremos corresponda a la precisión absoluta lograda por observación continua y el otro a resultados muy inciertos obtenidos mediante unas pocas observaciones aisladas. El tamaño de la muestra tiene, pues, su importancia, y podemos indicar si creemos o no en la representatividad de la muestra utilizando cierto **nivel de confianza**.

3. Cómo establecer niveles de confianza

Volvamos ahora al ejemplo citado y lancemos al aire cinco monedas simultáneamente, anotando el número de caras y cruces que salgan. Repitamos luego esta operación 99 veces más. Los resultados de estos lanzamientos podrían representarse como en el cuadro 11 o, gráficamente, como en la figura 87.

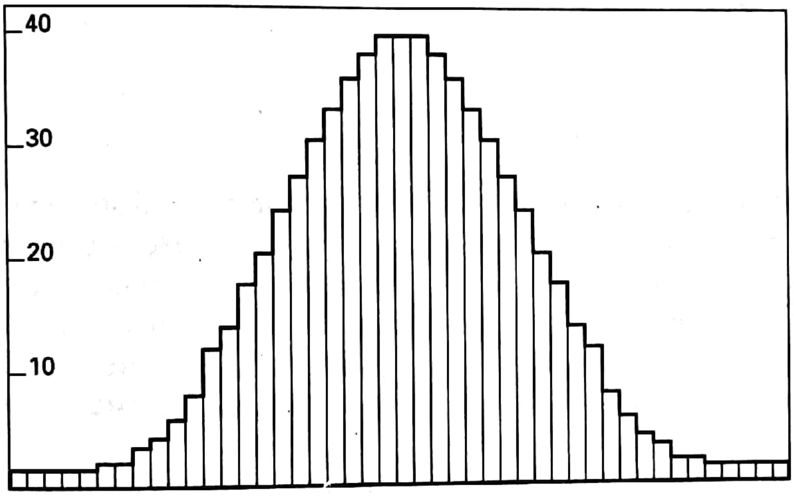
Figura 87. Distribución proporcional de «caras» y «cruces» (cien lanzamientos de cinco monedas a la vez)



Cuadro 11. **Distribución proporcional de «caras» y «cruces» (cien lanzamientos de cinco monedas a la vez)**

Combinación		Número de combinaciones
Caras *(p)*	Cruces *(q)*	
5	0	3
4	1	17
3	2	30
2	3	30
1	4	17
0	5	3
		100

Figura 88. **Curva de distribución que indica las probabilidades de combinaciones al utilizar grandes muestras**

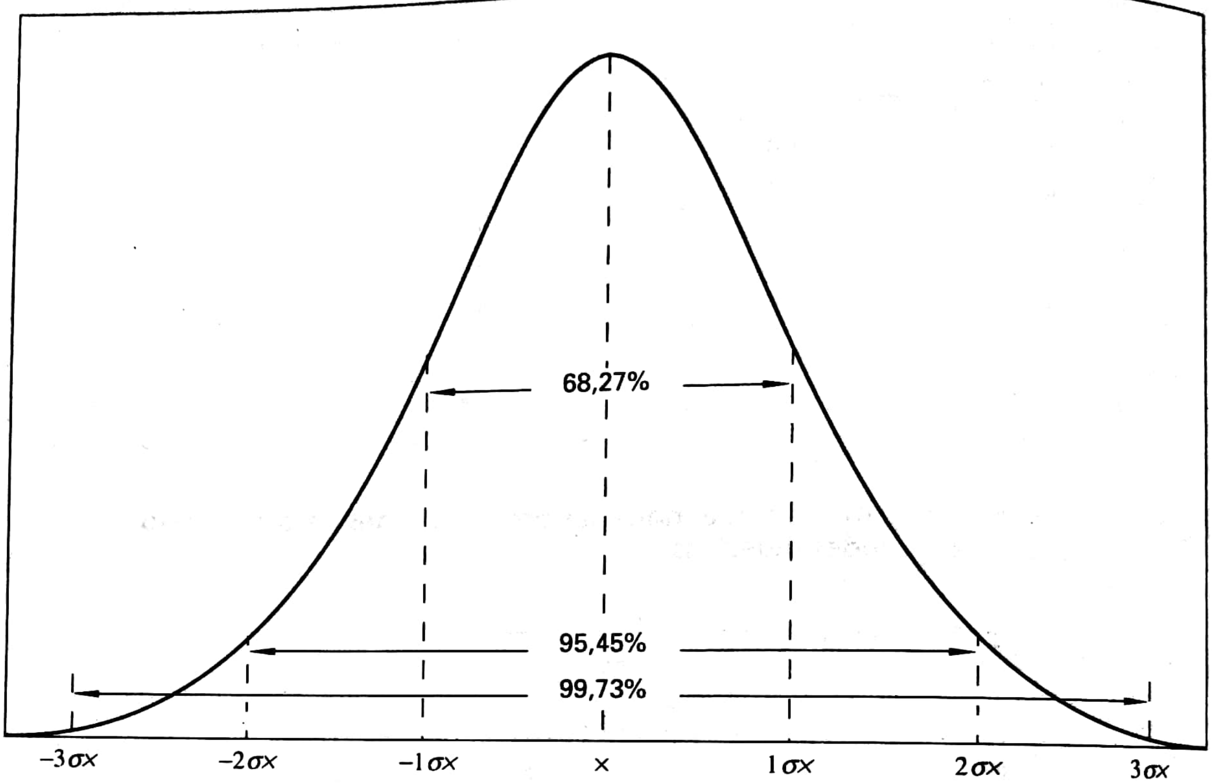


Si se aumenta considerablemente el número de lanzamientos, utilizando cada vez un gran número de monedas, podrá obtenerse una curva más progresiva, como la ilustrada en la figura 88.

Esta curva, llamada **curva de distribución normal**, también puede representarse como en la figura 89. Básicamente, esta curva indica que en la mayoría de los casos el número de caras tiende a igualar al de cruces en cualquier serie de lanzamientos (cuando $p = q$, el número de lanzamientos es un máximo). En pocos casos, sin embargo, p es muy diferente de q por mera casualidad.

Las curvas de distribución normal pueden tener numerosas configuraciones; según el caso, pueden ser más achatadas o más redondeadas. Para describir estas curvas se utilizan dos parámetros: $\bar{x}$, que es la media *o* la medida de la dispersión, y σ, que es la desviación de la media, denominada desviación típica o estándar. Dado que aquí se trata de una proporción, para indicar el error típico o estándar de la proporción se utilizará la expresión σp.

Figura 89. Curva de distribución normal



El área delimitada por la curva de distribución normal se puede calcular. En la figura 89, un σp a ambos lados de $\bar{x}$ da un área de 68,27 por ciento del área total; dos σp a ambos lados de $\bar{x}$ dan un área de 95,45 por ciento, y tres σp a ambos lados de $\bar{x}$ dan un área de 99,73 por ciento. En otros términos, si el muestreo realizado ha sido realmente aleatorio, 95,45 por ciento de las observaciones estarán comprendidas entre $\bar{x} \pm 2\ \sigma p$ y 99,73 por ciento estarán comprendidas entre $\bar{x} + 3\ \sigma p$.

Este es, de hecho, el grado de confianza que inspiran las observaciones. Sin embargo, para facilitar las cosas más vale evitar el uso de porcentajes decimales, pues es más sencillo hablar de un nivel de confianza de 95 por ciento que de 95,45 por ciento. Con ese fin pueden cambiarse los cálculos, obteniéndose:

- nivel de confianza de 95 por ciento, o sea 95 por ciento del área comprendida por la curva = 1,96 σp;
- nivel de confianza de 99 por ciento, o sea 99 por ciento del área comprendida por la curva = 2,58 σp;
- nivel de confianza de 99,9 por ciento, o sea 99,9 por ciento del área comprendida por la curva = 3,3 σp.

En este caso podemos decir que si tomamos una muestra **aleatoria** de gran tamaño, podemos confiar en que en 95 por ciento de los casos las observaciones estarán comprendidas entre ± 1,96 σp, y así sucesivamente para los demás valores.

En el muestreo del trabajo, el nivel de confianza más generalmente utilizado es el de 95 por ciento.

4. Cómo determinar el tamaño de la muestra

Además de definir el nivel de confianza de nuestras observaciones, también debemos decidir el margen de error que admitiremos. Debemos poder decir que «tenemos confianza en que 95 por ciento de las veces la observación que hagamos tendrá una exactitud de ± 5 por ciento», o 10 por ciento, o cualquier otro margen de exactitud que adoptemos.

Volvamos ahora a nuestro ejemplo del tiempo productivo y del tiempo inactivo de las máquinas de una fábrica. Para determinar el tamaño de la muestra que se necesita con este ejemplo existen dos métodos: el método estadístico y el método nomográfico.

Método estadístico

La fórmula utilizada en este método es la siguiente:

$$\sigma p = \sqrt{\frac{pq}{n}}$$

en la que:

σp = error estándar de la proporción;
p = porcentaje de tiempo inactivo;
q = porcentaje de tiempo en marcha;
n = número de observaciones o tamaño de la muestra que determinar.

Sin embargo, antes de poder aplicar esta fórmula debemos tener por lo menos una idea de los valores de p y q. Así, pues, el primer paso consiste en efectuar cierto número de observaciones aleatorias en el lugar de trabajo. Supongamos que, como estudio preliminar y aleatorio, se efectuaron 100 observaciones, de las que se dedujo que las máquinas estaban paradas 25 por ciento del tiempo ($p = 25$) y en marcha el restante 75 por ciento ($q = 75$). Ahora ya disponemos de los valores aproximados de p y q; para poder determinar el valor de n debemos calcular antes el valor de σp.

Tomemos, por ejemplo, un nivel de confianza de 95 por ciento con un margen de error de 10 por ciento (es decir, que tenemos confianza en que en nuestros cálculos el 95 por ciento de los casos corresponderán a ± 10 por ciento del valor real).

Al nivel de confianza de 95 por ciento,

$$1{,}96\ \sigma p = 10$$
$$\sigma p = 5 \text{ (aproximadamente).}$$

Ahora podemos volver a nuestra ecuación inicial para derivar n:

$$\sigma p = \sqrt{\frac{pq}{n}}$$

$$5 = \sqrt{\frac{25 \times 75}{n}}$$

$$= 75 \text{ observaciones.}$$

Si reducimos a ± 5 por ciento el margen de error, tendremos:

$$1{,}96\ \sigma p = 5$$
$$\sigma p = 2{,}5 \quad \text{(aproximadamente)}$$
$$2{,}5 = \sqrt{\frac{25 \times 75}{n}}$$
$$n = \sqrt{\frac{25 \times 75}{(2{,}5)^2}}$$
$$= 300 \text{ observaciones.}$$

En otras palabras, para reducir el margen de error a la mitad habrá que cuadruplicar el tamaño de la muestra.

Método nomográfico

El tamaño de la muestra puede determinarse con mayor facilidad leyendo directamente el número de observaciones requeridas en un nomograma como el presentado en la figura 90. Tomando nuevamente el ejemplo precedente, tracemos una línea recta que partiendo de la ordenada p «porcentaje de aparición» (en este caso, 25-75) corte la ordenada «error (precisión requerida)» (digamos 5 por ciento) y se prolongue hasta encontrar la ordenada n «número de observaciones»; se ve que la corta a 300 para un nivel de confianza de 95 por ciento. Este sistema para determinar el tamaño de la muestra es rapidísimo.

5. Cómo efectuar observaciones aleatorias

Las conclusiones a que hemos llegado son válidas siempre que podamos efectuar el número de observaciones necesarias para lograr el nivel de confianza y la precisión requeridos, y a condición de que las observaciones se hagan **al azar**.

Para asegurarnos de que las observaciones son efectivamente aleatorias podemos utilizar una tabla de números aleatorios como la del cuadro 12. Existen varios tipos de tablas de ese género, que pueden utilizarse de diferentes maneras. En nuestro caso, supongamos que nuestras observaciones se llevarán a cabo durante un turno de trabajo de ocho horas, de las 7,00 a las 15,00 horas. Una jornada de trabajo de ocho horas tiene 480 minutos, que pueden dividirse en 48 períodos de diez minutos.

Podemos empezar escogiendo en la tabla un número al azar, por ejemplo, cerrando los ojos y colocando la punta de un lápiz en algún lugar de la tabla. Supongamos que en este caso hemos caído por pura casualidad en el número 11, que se encuentra en el segundo bloque de la primera hilera vertical, cuarta columna, cuarta línea (cuadro 12). Seguidamente escogemos un número cualquiera de 1 a 10. Supongamos nuevamente que hemos elegido el número 2; bajando ahora por la columna, seleccionamos una cifra de cada dos y la anotamos, como se indica a continuación (si hubiéramos escogido el número 3, deberíamos seleccionar una cifra de cada tres, y así sucesivamente):

11 38 45 87 68 20 11 26 49 05