

CAPITULO IX

LA TEORIA DEL MUESTREO

Una muestra es una pequeña porción de un lote de material, el cual contiene todos los componentes, *en la misma proporción que existe en el original*. El objeto de cualquier procedimiento de muestreo es obtener una porción representativa del todo. Las propiedades analizadas de la muestra pueden ser usadas para predecir o pronosticar las propiedades de la población de la cual la muestra fue extraída.

El control de calidad de cualquier industria se basa en procedimientos de muestreos constantes y regulares. Además, es necesario un conocimiento de estadística aplicada para el caso que se desee pronosticar las características de la población a partir de los resultados de muestreo. El rol del análisis estadístico es grande en cualquier tecnología donde se desee establecer la validez de un conjunto de datos de pruebas experimentales y donde se requiera tomar decisiones.

El muestreo y la inferencia estadística son utilizados en circunstancias en las cuales no es factible obtener información de todos los integrantes de una población, tales como en el análisis químico y biológico, control de calidad industrial ó investigaciones sociales. El método más elemental de muestreo es el muestreo aleatorio, basado en la teoría de probabilidades. Según esto, cada elemento de la población que es muestreada tiene la misma probabilidad de ser seleccionado. Por ejemplo, en una clase de 50 estudiantes, para obtener una muestra aleatoria de ese grupo, cada estudiante tiene exactamente la misma probabilidad, $1/50$, de ser seleccionado.

Las prácticas deficientes de muestreo son una causa primordial de inexactitudes en los valores calculados. Desafortunadamente, el muestreo preciso requiere de grandes cantidades de materiales, lo que hace del muestreo una tarea difícil y costosa. El costo de las instalaciones y personal para el muestreo es directo y tangible, pero los beneficios de un muestreo preciso, aunque indirectos, menos tangibles y de corto alcance, pueden ser más significativos. Como los análisis numéricos pueden obtenerse con tanta facilidad sobre muestras deficientes que sobre muestras buenas, fácilmente se genera un falso sentido de seguridad en la mente del ingeniero en la oficina, cuando solamente ve los resultados numéricos y no observa como fueron tomadas las muestras en el campo.

La composición del material que se mueve en un proceso es característicamente irregular, excepto en casos muy especiales. Idealmente un proceso continuo operando en estado uniforme deberá dar un flujo de materiales uniforme, pero no se puede suponer tal estado; son necesarios el muestreo y el análisis para conocer las características reales de la población que se analiza.

9.1. Razones para el muestreo

Al desarrollarse una investigación estadística se deben seguir ciertos pasos o etapas después de definir el problema en términos estadísticos. Antes de recolectar los datos, se debe decidir de que forma se van a coleccionar dichos datos; es decir los procedimientos a utilizar para obtener la muestra, de manera que podamos aprender algo acerca de la población sobre la base de datos extraídos de una parte de ella.

Las razones para el muestreo se pueden definir como las siguientes: primero, las poblaciones que se investigan pueden ser infinitas, y en tales casos el muestreo es el único procedimiento posible; segundo, aún en casos de poblaciones finitas, muy frecuentemente el muestreo es el único procedimiento práctico, pues la población finita puede constar de millares o millones de elementos y su recolección completa es prácticamente imposible; en tercer lugar, la medición de propiedades de una población a veces requiere de la destrucción de sus elementos, para determinar determinadas propiedades físicas por ejemplo.

Los resultados obtenidos por el estudio de una muestra pueden ser iguales o aún más precisos que los hallados de una cuenta completa de la población, puesto que, ante la posibilidad de evaluar a toda una población, el tiempo que requeriría ello podría conducir a que en el transcurso de la evaluación se produjesen cambios importantes que al finalizar la evaluación, el resultado sería diferente.

9.2 Bases del muestreo

Aunque la diversidad es una cualidad universal de los datos, no hay virtualmente ninguna población estadística en la práctica cuyos elementos varíen entre si sin límite. Los hechos de que cualquier población tiene propiedades características y que las variaciones de sus elementos son claramente limitadas, hace posible que se escoja una muestra relativamente pequeña y al azar que pueda reflejar bastante bien las características de la población.

Otra propiedad interesante e importante de los datos es la regularidad o uniformidad. Las fuerzas relacionadas, pero independientes que producen la variabilidad de una población están a menudo tan equilibradas y concentradas que tienden a generar iguales valores arriba y abajo de cierto valor central alrededor del cual tienden a agruparse la mayor parte de los valores.

Debido a la uniformidad estadística, si se escoge una muestra grande al azar, las características de esta muestra difieren muy poco de las características de la población. La uniformidad o regularidad es la tendencia de las características mensurables al agruparse alrededor de cierto “centro de gravedad”. La medida de tal tendencia central es un promedio del que divergen las observaciones individuales en cierta forma definida. Los promedios son más estables que los valores individuales. Además, los promedios resultan más estables cuantas más observaciones haya en la muestra.

La razón básica de que cantidades mayores de datos tienden a exhibir menos fluctuación es una tendencia de las unidades de datos a “compensarse unas con otras”. El mismo efecto de compensación puede esperarse cuando aumenta el tamaño de la muestra.

En la práctica, para hacer inferencias generalmente solo tomamos una muestra. Sea esta grande o pequeña, estamos casi seguros de que sus características no son exactamente las de la población. ¿Cómo podemos estar seguros entonces sobre el grado de confianza de nuestras conclusiones?. La respuesta a la pregunta es *aleatoriedad*. La medición objetiva de los errores por muestreo requiere que la muestra sea al azar. Esto es así porque este conocimiento solo puede ser obtenido aplicando las leyes matemáticas de probabilidad, las que a su vez, solo pueden ser aplicadas a muestras al azar.

9.3 El problema del muestreo

Los procedimientos para diseñar un procedimiento de muestreo son: a) definir el objetivo del muestreo; b) especificar los materiales a muestrear en términos de cantidad de flujo y los rangos de los parámetros de calidad del material; c) definir la cantidad de muestra a extraerse para lograr un determinado nivel de precisión; y, d) definir el equipo a utilizar para el muestreo primario que permita obtener una muestra representativa y que no introduzca errores según las definiciones estadísticas.

El problema cualitativo en un mineral triturado se describe en la Figura 9.1. En esta Figura los elementos negros representan las partículas de interés y el resto materiales de la ganga.

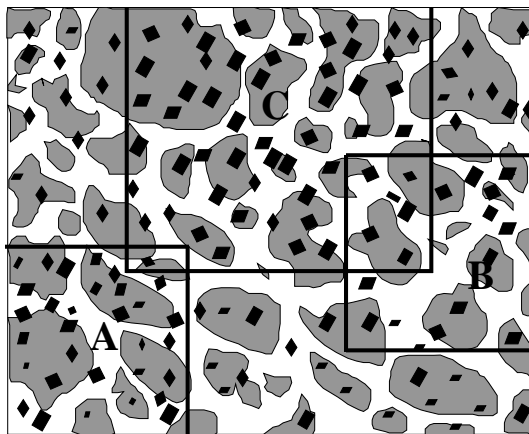


Figura 9.1: Sección de mineral.

Una muestra del tamaño apropiado deberá contener una representación estadísticamente válida del todo. En esta Figura se puede contabilizar un total de 100 partículas visibles en la sección de mineral que representa. Naturalmente, se encontrarán diferente número de partículas en otras secciones no visibles, pero se puede asumir que las partículas ocultas se encuentran distribuidas completamente aleatorizadas de modo que cada sección contenga un número proporcional de partículas de acuerdo al tamaño de la sección.

El concepto de representación de muestra se ilustra tomando tres secciones del material de la Figura 9.1. Las partes A y B poseen la misma área cada una y la sección C es del doble del área que las otras. Obsérvese la superposición de áreas. Los granos de mineral contabilizados son: $A = 32$; $B = 19$; y $C = 56$. Si establecemos una evaluación de granos por área observada, se notará que la sección completa representada en la Figura 3.1 contiene 2.35 granos/cm^2 ; la sección $A = 3.01 \text{ granos/cm}^2$; la sección $B = 1.79 \text{ granos/cm}^2$; y la sección $C = 2.64 \text{ granos/cm}^2$. Entonces para cada sección el error en relación al todo son de $+28\%$, -24% y $+12\%$ respectivamente para cada área con respecto al total. De esto resulta que las muestras de menor tamaño difieren en mayor proporción del verdadero valor que las muestra de mayor tamaño. Esto ilustra el efecto del tamaño de muestra (o peso) al caracterizar un material por muestreo, y también proporciona un ejemplo cualitativo del problema de la representación estadística de la técnica de muestreo.

Las muestras trituradas de mineral comprende una colección *discreta* de piezas de roca, cada una con sus propias características en tamaño y contenido de mineral. Porciones de rocas seleccionadas aleatoriamente

de una pila de mineral o de una corriente en movimiento de mineral mostrarán variaciones entre trozos al efectuar los ensayos respectivos, esto debido a la distribución no uniforme del material, observándose variaciones de un fragmento de roca a otro. Al existir variaciones entre los fragmentos de roca se condiciona a un mayor nivel de trituración y así un mayor número de partículas minerales se liberan de la ganga estéril. Con una muestra consistente de partículas de diferente tamaño, las variaciones entre muestra disminuirán al incrementarse el tamaño de muestra (mayor cantidad de partículas se incluyen en la muestra), ya que las partículas con diverso contenido de mineral y diverso tamaño se incluirán en la muestra final. Las variaciones entre muestra se podrán disminuir al nivel deseado con el incremento del tamaño de muestra. Naturalmente, esto acarrea mayores costos en la operación de muestreo.

Diversos aspectos emergen de esas observaciones. Primero, el tamaño de muestra está relacionado al nivel de variación entre muestra. Segundo, para observar variaciones en un caso específico es necesario observar muestras del mismo peso; a mayor número de muestras, mejor será el estimado estadístico de las variaciones entre muestra. Tercero, para obtener un valor de variación entre muestra, fijar, ya sea el tamaño de muestra y variar el número de muestras, o fijar el número de muestras y variar el tamaño de muestra para obtener la precisión de muestreo. Cuarto, el tamaño de muestra es influenciado por la abundancia del mineral útil; un mineral de alta ley se podrá caracterizar adecuadamente con una muestra pequeña, manteniendo todos los demás factores iguales; lo contrario sucederá con un mineral de baja ley. Un quinto punto refiere la relación entre el tamaño del mineral valioso y el tamaño de la roca; si la relación es pequeña, la muestra deberá ser necesariamente mayor para obtener una caracterización más adecuada. Este punto refiere por lo tanto al grado de dispersión del mineral. Finalmente; para minimizar el tamaño de muestra, es preferible muestrear pequeñas partículas en vez de grandes trozos.

9.4. Probabilidad y muestreo

Para ilustrar estadísticamente el problema del muestreo, consideremos el siguiente ejemplo. Se tiene una población de 100 frijoles, la cual consiste de 10 frijoles negros y 90 frijoles blancos (los frijoles negros representan al elemento de interés y los blancos a los demás constituyentes de una población). Si se sacara una muestra de 10 frijoles de la población, ¿cuáles son las probabilidades de conseguir una muestra que contenga: 0, 1, 2, 3 y 4 frijoles negros respectivamente?

Solución:

La probabilidad de obtener un número de frijoles negros de una muestra total se calcula mediante la siguiente expresión:

$$P(x = \text{negro}) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

donde:

p = probabilidad de que haya un negro en la población.

q = probabilidad de que haya un blanco en la población

x = número de negros en la muestra extraída

n = tamaño de la muestra.

$$p = \frac{10}{100} = 0.1; \quad q = 1 - 0.1 = 0.9; \quad n = 10$$

Para x = 0, es decir, cero negros y diez blancos:

$$P(0 \text{ negros}) = \frac{10!}{0! 10!} (0.1)^0 (0.9)^{10} = 0.349$$

$$P(1 \text{ negro}) = \frac{10!}{1! 9!} (0.1)^1 (0.9)^9 = 0.387$$

Similarmente:

$$P(2 \text{ negros}) = 0.194$$

$$P(3 \text{ negros}) = 0.057$$

$$P(4 \text{ negros}) = 0.011$$

Obsérvese que en una muestra de 10 frijoles, obtener un frijol negro y nueve blancos, sería una muestra representativa del total, y esta combinación tiene una probabilidad de 0.387, o sea, se tiene aproximadamente 39 % de posibilidades de obtener una muestra totalmente representativa del todo.

Obsérvese ahora el problema del muestreo de otra manera. Se tiene un recipiente con 250 frijoles negros y 750 frijoles blancos. Los frijoles negros representan al elemento de interés y los blancos a los demás constituyentes de una población. Se sacan del recipiente muestras de 100 frijoles y se contabiliza el número de frijoles negros de cada muestra extraída. Se repite el procedimiento por 1,000 veces. Los resultados reales del experimento se dan en la Tabla 9.1. y se comparan con los resultados teóricos que se obtienen utilizando la formula:

$$P(x = \text{negro}) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

De estos resultados tabulados se puede observar que si se sacan 1,000 muestras, solamente 92 de las 1,000 representarían correctamente el valor de la población; es decir, 92 muestra contienen 25 frijoles negros y 75 blancos cada una. Sin embargo, las otras 908 muestras también serían valores verdaderos, causales o al azar y no pueden ser descartados.

La Figura 9.2 describe la distribución normal de la extracción de frijoles, según los valores calculados y los valores reales observados. Nótese la gran aproximación existente entre ambos gráficos. Obsérvese también que la gran mayoría de extracciones caen en el intervalo [19, 31] frijoles negros en una muestra de 100. Asociado este concepto de probabilidad a una operación de muestreo, el obtener una muestra totalmente representativa del total tiene muy bajas probabilidades; mas bien, la mayoría de muestras obtenidas adecuadamente se distribuirán en un rango determinado alrededor del verdadero valor.

9.5. Prueba de normalidad

La distribución normal es un modelo teórico que describe la variabilidad en la variable X , y se asume que los parámetros de población μ y σ son conocidos. Supóngase que estos parámetros son desconocidos y se desea estimar estos de una muestra de tamaño n , los naturales estimadores son el promedio y varianza de muestra respectivamente, \bar{x} y s^2 . Ya que son sólo estimados de los parámetros de población μ y σ , no se sabe si la distribución normal proporciona el modelo apropiado para el análisis de la variabilidad de las observaciones en consideración.

Tal vez la manera más simple de verificar si las observaciones provienen de una distribución normal es elaborando el histograma de las observaciones. Ya que la frecuencia relativa de un histograma es un estimado de la densidad de probabilidad, se debe verificar si el gráfico se asemeja a la curva de la "campana" correspondiente a la distribución normal. Sin embargo, es difícil verificar si un pequeño número de datos proviene de una distribución normal.

Los percentiles, q_i , normales dependen de los parámetros μ y σ . En la práctica se puede reemplazar estos parámetros por sus estimados \bar{x} y s^2 . Alternativamente, se puede graficar los valores estandarizados z_i versus los respectivos percentiles x_i . Ya que $q_i = \bar{x} + s z_i$ y que bajo normalidad $x_i \cong q_i$ se encuentra que los puntos de un diagrama de dispersión de z_i vs. x_i deben caer aproximadamente en una línea recta con pendiente $1/s$ y que pasa por el punto $(\bar{x}, 0)$. Tal gráfico se denomina *gráfico de probabilidad normal*. Si se observa desviaciones del patrón lineal, es suficiente evidencia de que los datos no provienen de una distribución normal.

Tabla 9.1: Valores reales y calculados del experimento de extracción de frijoles.

n= 100		negros= 250 total= 1000		p= 0.25 q= 0.75			
					Acum	Acum	
x	n-x	P(x)	P(x) *1000	REAL	P(x)	P(x)*1000	
9	91	0.000	0		0.000	0	
10	90	0.000	0		0.000	0	
11	89	0.000	0		0.000	0	
12	88	0.001	1		0.001	1	
13	87	0.001	1		0.002	2	
14	86	0.003	3		0.005	5	
15	85	0.006	6	6	0.011	11	
16	84	0.010	10	10	0.021	21	
17	83	0.017	17	12	0.038	38	
18	82	0.025	25	22	0.063	63	
19	81	0.037	37	35	0.100	100	
20	80	0.049	49	48	0.149	149	
21	79	0.063	63	61	0.211	211	
22	78	0.075	75	78	0.286	286	
23	77	0.085	85	89	0.371	371	
24	76	0.091	91	100	0.462	462	
25	75	0.092	92	100	0.553	553	
26	74	0.088	88	94	0.642	642	
27	73	0.081	81	84	0.722	722	
28	72	0.070	70	72	0.792	792	
29	71	0.058	58	57	0.850	850	
30	70	0.046	46	44	0.896	896	
31	69	0.034	34	31	0.931	931	
32	68	0.025	25	24	0.955	955	
33	67	0.017	17	14	0.972	972	
34	66	0.011	11	9	0.984	984	
35	65	0.007	7	5	0.991	991	
36	64	0.004	4	3	0.995	995	
37	63	0.002	2	2	0.997	997	
38	62	0.001	1		0.999	999	
39	61	0.001	1		0.999	999	
40	60	0.000	0		1.000	1000	
41	59	0.000	0		1.000	1000	
42	58	0.000	0		1.000	1000	

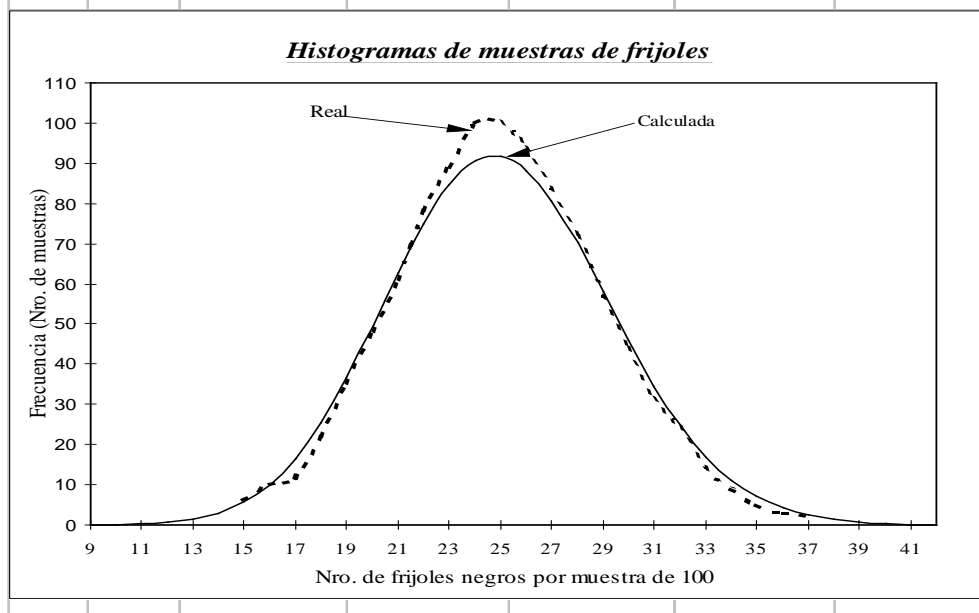


Figura 9.2: Histograma del experimento de extracción de muestras de 100 frijoles.

Ejemplo: Para ilustrar el procedimiento, se tomara los 10 primeros datos de la Tabla 1.2 de Resistencia a la Compresión de Bloques de Concreto. Estos datos ordenados según su posición ordinal son como sigue:

Orden	1	2	3	4	5	6	7	8	9	10
Valor	31.3	32.3	42.2	42.3	44.5	47.5	49.2	50.0	53.9	60.9

Si diversos valores se repiten, se les asigna la posición ordinal promedio.

Se calcula $P_i = (i - 0.5)/n$. Los valores estandarizados normales satisfacen $P_i = P(Z \leq z_i) = \Phi(z_i)$ para $i = 1, 2, \dots, n$ y se determinan de la Tabla Normal. Por ejemplo, z_1 satisface $P_1 = (1 - 0.5)/10 = 0.05 = P(Z \leq z_1)$ y esta dado por $z_1 = -1.645$. El segundo satisface $P_2 = (2 - 0.5)/10 = 0.15 = P(Z \leq z_2)$ y esta dado por $z_2 = -1.04$; y así sucesivamente. Estos valores se anotan en la última columna de la Tabla 9.2. El gráfico de los valores normales vs. las observaciones, se da en la Figura 9.2.

Tabla 9.2. Primeras 10 observaciones de la Tabla 1.2 de Resistencia a la compresión de bloques de concreto y sus valores normales.

Observación	Posición	$P_i = \frac{i-0.5}{n}$	Valor Normal
49.2	7	0.65	0.39
53.9	9	0.85	1.04
50.0	8	0.75	0.67
44.5	5	0.45	-0.13
42.2	3	0.25	-0.67
42.3	4	0.35	-0.39
32.3	2	0.15	-1.04
31.3	1	0.05	-1.64
60.9	10	0.95	1.64
47.5	6	0.55	0.13

Ya que los datos exhiben una relación lineal bastante buena, se puede aceptar con seguridad que los datos provienen de una distribución normal.

9.6. Eliminación de datos.

Normalmente se acepta que cuanto más grande sea el número de medidas de una muestra dada, mayor será la confianza de que su media sea el valor verdadero; por tanto es conveniente obtener tantas medidas como sea posible; sin embargo, hay un límite práctico, sea por el método de medida utilizado o la cantidad de muestra disponible. El problema es decidir cuales deben ser aceptadas y cuales rechazadas. La prueba de la *desviación estándar* y el *Criterio Q* son métodos para decidir el rechazo de algunos de los datos obtenidos.

Probabilidad Normal

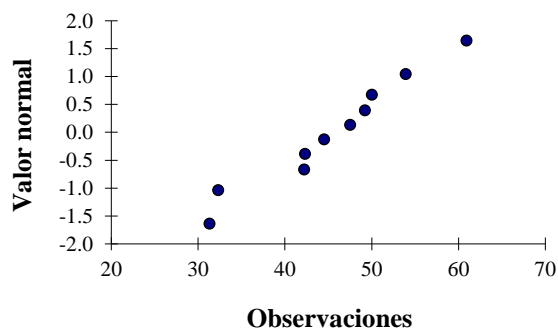


Figura 9.2: Gráfico de probabilidad normal.

a). Prueba de la Desviación Estándar.

Si se han hecho muchas medidas a una muestra, las observaciones se aproximarán a una curva de distribución normal. Esta distribución indica que cualquier dato tiene un 99/7% de probabilidad de caer en el intervalo $\bar{x} \pm 3\sigma$. Normalmente, cuando se utiliza la desviación estándar como criterio de rechazo, se toma el valor de 3σ . Se rechaza si el valor de una medida particular esta fuera de este intervalo.

b). El Criterio Q.

Puede utilizarse para rechazar una medida sobre un población de 3 a 10 medidas de una cantidad a cualquier nivel de confianza, aunque en circunstancias normales se utiliza el 90%; es decir, si se rechaza una medida según este criterio, se hace con una confianza del 90%. Para el aplicar el criterio Q se ordenan las medidas en orden creciente, tomándose los valores más divergentes, superior e inferior, calculándose el rango ($M_n - M_1$) y con el los cocientes Q para el mayor y menor valores, según:

$$Q_1 = \frac{M_2 - M_1}{M_n - M_1} \quad Q_2 = \frac{M_n - M_{n-1}}{M_n - M_1}$$

Como los valores que podrían eliminarse son M_1 ó M_n , el criterio Q se aplica a ambos. Estos valores se comparan con los que se anotan en la Tabla Nro. 9.3. Si Q_1 y Q_2 son mayores que el valor dado en la Tabla, el dato puede ser rechazado con cierto límite de confianza; por ejemplo, si Q_1 tiene un valor de 0.60 y el número de datos, $n = 7$, el dato puede ser eliminado con un límite de confianza del 90%.

Tabla Nro. 9.3: Valores de Q para rechazar datos con límite de confianza del 90%

Número de observaciones	$Q_{90\%}$
3	0.94
4	0.76
5	0.64
6	0.56
7	0.51
8	0.47
9	0.44
10	0.41

Ejemplo: Se efectúan 6 determinaciones de Fe en una misma muestra mineral, siendo los resultados los siguientes:

% Fe	14.28	14.20	13.99	13.18	13.92	14.30
------	-------	-------	-------	-------	-------	-------

De donde : $\bar{x} = 13.98\%$ y $s = 0.420$

El valor de s considera los 6 datos. Si se descarta la peor medida, 13.18 %, los nuevos valores calculados son: $\bar{x} = 14.14\%$ y $s = 0.170$

De modo que $3s = 3 * 0.170 = 0.510$.

El límite de aceptación es: $\%Fe = 14.14 \pm 0.510 = [13.63; 14.65]$

El valor 13.18 % no se encuentra dentro de estos límites, por lo que se puede desechar. El procedimiento se repite ahora con el peor valor de los restantes cinco datos. Hecho ello, se observa que 13.92 cae dentro de los nuevos límites, por lo que este dato no se desecha.

El criterio Q . Considera los datos ordenados

% Fe	13.18	13.92	13.99	14.20	14.28	14.30
------	-------	-------	-------	-------	-------	-------

El rango $\{M_6 - M_1\} = (14.30 - 13.18) = 1.12$

$$Q_1 = \frac{M_2 - M_1}{M_n - M_1} = \frac{13.92 - 13.18}{1.12} = 0.66 \quad Q_2 = \frac{M_n - M_{n-1}}{M_n - M_1} = \frac{14.30 - 14.28}{1.12} = 0.01$$

El $Q_{90\%} = 0.56$; con lo que se determina que el dato de 13.18 % se elimina. Repitiendo el procedimiento se encuentra que el resto de valores es aceptable. Por lo tanto, el valor aceptables de %Fe = 14.14.

9.7 Determinación del tamaño de muestra.

La determinación del tamaño de muestra es muy importante, puesto que si se toma una muestra muy pequeña, no será representativa; y si la muestra es muy grande, se estaría desperdiciando recursos. Por ello el tamaño de muestra adecuado deberá ser considerado correctamente.

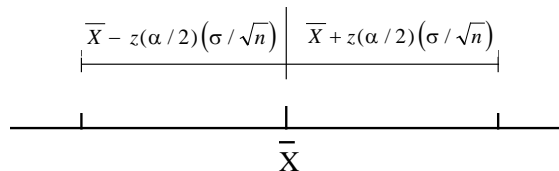
9.7.1. Tamaño de muestra para definir el intervalo de confianza del promedio

Se está planificando un experimento y se quiere estar $100(1-\alpha)$ por ciento seguro de que el estimado esta a h unidades del parámetro desconocido μ . Por ejemplo, queremos estar con el 95 % de confianza que nuestro estimado se encuentre a $h=2$ unidades de μ . ¿Cuán grande debemos tomar la muestra n para estimar μ con tal precisión?.

Ya que:

$$P\left[\bar{X} - z(\alpha/2) \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z(\alpha/2) \frac{\sigma}{\sqrt{n}}\right] = 1 - \alpha$$

lo que se puede graficar según:



De modo que la longitud del intervalo alrededor del promedio, h , es:

$$h = z(\alpha/2) \left(\sigma / \sqrt{n} \right)$$

De ésta última expresión se tiene:

$$n = \frac{\sigma^2 [z(\alpha/2)]^2}{h^2}$$

expresión que permite calcular el número mínimo de muestras, n , a considerarse en un muestreo. Esto requiere conocer σ^2 . En muchas circunstancias, se puede tener una aproximación razonable a σ^2 de anteriores pruebas o datos.

La expresión anterior denota que el tamaño de muestra, n , se incrementa al incrementarse el coeficiente de confianza $(1 - \alpha)$ y disminuye al incrementarse h . Algunas veces se requerirá bastante precisión, por lo que n debe ser muy grande. Pero si no es posible recolectar grandes muestras, se deberá reducir el coeficiente de confianza o incrementar el ancho del rango h .

9.7.2. Tamaño de muestra para evaluar la diferencia de dos promedios

Si se conocen (al menos aproximadamente) las desviaciones estándar σ_1 y σ_2 y los tamaños de las dos muestras se consideran iguales, $n_1 = n_2 = n$, entonces puede determinarse el tamaño requerido de la muestra, de modo que se tenga confianza del $100(1-\alpha)$ por ciento en que el error en la estimación de $\mu_1 - \mu_2$ sea menor que h . El tamaño requerido para la muestra de cada población es:

$$n = \frac{(\sigma_1^2 + \sigma_2^2) [z(\alpha/2)]^2}{h^2}$$

9.8. Muestreo mediante incrementos

El muestreo por incrementos es el término aplicado a la forma de muestreo, en el que un número de incrementos, específicamente medidos, se sacan de la población y se compositan para formar la muestra apropiada. En número real de incrementos requeridos y el tamaño de cada incremento, son funciones de los

patrones se segregación y dispersión de los valores de la población. Cuando el muestreo mediante incrementos se efectúa correctamente, los resultados deben ser representados por una curva de distribución normal, donde la frecuencia se plotea contra el grado. Entonces es posible aplicar la teoría de la probabilidad a los resultados iniciales de muestreo y predecir el número de incrementos que sería requerido para dar un error específico probable de estimación a cualquier nivel de confianza deseado.

Ya que :

$$Z = \frac{(\bar{x} - \mu)}{\frac{S}{\sqrt{n}}}$$

$$ZS = (\bar{x} - \mu)\sqrt{n}$$

sea $e = (\bar{x} - \mu)$ = error de estimación = número de unidades de diferencia
con respecto al valor promedio

$$n = \left(\frac{ZS}{e} \right)^2 = \text{número de muestras}$$

Ejemplo de aplicación

Doce muestras de 5 libras son tomadas de un lote de mineral que contiene cinc, los cuales ensayaron lo siguiente en % se Zn:

7.04; 8.10; 8.85; 6.25; 7.69; 8.76; 8.35; 7.57; 7.96; 9.30; 8.50 y 7.49.

¿Cuántos incrementos deben ser compositados con el objeto de que la muestra compósito ensaye dentro del 2 % del ensaye verdadero a un nivel de confianza del 95 %?

Solución:

Valor promedio, $\bar{x} = 7.99$

Varianza, $S^2 = 0.718$

Desviación estándar, $S = 0.847$

Error permisible de estimación $e = \left(\frac{2}{100} \right) * 7.99 = 0.1598$

$Z = 1.96$ para 95% de nivel de confianza

$$n = \left(\frac{1.96 * 0.847}{0.1598} \right)^2 = 108$$

Se debe compositar 108 muestras de mineral para que el resultado final se encuentre en un rango 2% alrededor del verdadero promedio.

9.2 Peso mínimo de muestra de un lote homogéneo.

Un procedimiento para calcular la cantidad de muestra de una mezcla uniforme (homogénea) de un lote de mineral fue desarrollado por Gy*. El peso de muestra según el procedimiento descrito es una cantidad ideal y que puede ser considerada como un estimado de peso mínimo y toma en cuenta los errores normales de la operación de muestreo y ausencia de otras fuentes de error. Según esta definición, el error de muestreo es aquel que surge de la dispersión “no ideal” del mineral, y se le denomina *varianza del error fundamental*, V_F . Los errores que se cometen durante el muestreo pueden ser disminuidos usando las técnicas apropiadas, pero el más importante de ellos, el Error Fundamental, resulta de la heterogeneidad del material muestreado. El adjetivo “fundamental” está justificado por el hecho de que no puede ser eliminado; es el que perdura así sea perfecta la operación de muestreo. Sin embargo, el error fundamental es el único que puede ser estimado de antemano.

Factores empíricos se proveen para evaluar la incidencia del tamaño de grano del mineral útil, dispersión del mineral, tamaño de trozos de roca y promedio de contenido metálico. El error fundamental de muestreo puede ser calculado para una muestra de peso W_s tomada de un lote uniforme de peso W_L mediante la siguiente ecuación:

* Gy, P. M.; “The Sampling of Broken Ores: A Review of Principle and Practice”, Proceedings, Institution of Mining Metallurgy Symposium, London, 1963, pp.26-27.

$$V_F = \left[\frac{1}{W_S} - \frac{1}{W_L} \right] \left[\frac{1-A_L}{A_L} (1-A_L) a_m + A_L a_g \right] f g b d^3$$

donde el error estándar se relaciona a la varianza del error fundamental por:

$$S_F = \sqrt{V_F}$$

A_L es la fracción en peso del mineral en el lote;

a_m es la densidad del mineral;

a_g es la densidad de la ganga;

f es un factor adimensional que relaciona la forma de la partícula;

g es un factor adimensional que relaciona la distribución de tamaño;

b es un factor adimensional que relaciona la liberación del mineral; y

d es el tamaño de partícula relacionada a la malla por la que pasa el 95% de una muestra.

Se escogen la unidades apropiadas para W , A y d de tal modo que el lado derecho de la ecuación sea adimensional. Los términos empíricos propuestos por Gy derivados de una serie de experimentaciones, son como sigue:

f de 0 a 1, con un valor promedio de 0.5 para minerales típicos; varía hasta 0.2 para minerales con metales preciosos y minerales con partículas metálicas.

g de 0 a 1 con valores de 0.25 para minerales extraídos y no clasificados. Minerales clasificados según tamaño toman valor de g de 0.5 o más.

b de 0 a 1, valor seleccionado de la Tabla 3.2 de acuerdo a la proporción de d para estimar el tamaño de liberación del mineral, d_b

El parámetro b es una medida de dispersión. Para minerales completamente dispersos, b toma valores que se aproximan al límite más bajo y se aproxima al límite superior para material heterogéneo con granos de mineral en estado libre, tales como el oro metálico.

Tabla 3.2: Factor de liberación para minerales.

Factor de liberación, b	0.8	0.4	0.2	0.1	0.05	0.02
$\frac{d}{d_b}$	1	4	10	40	100	
	heterogéneo		intermedio		homogéneo	

Ejemplo de aplicación

Calcular el error de ensayo para una muestra de 10 Kg. de mineral de Zinc con contenido de 6.6% de Zn como esfalerita, ZnS, tomado de una pila uniformemente mezclada de 20,000 Kg. 95 % del mineral atraviesa una malla de 2 cm. El tamaño práctico de liberación, d_b , es de 0.2 mm. La densidad de la ganga es, a_g , es de 2.5 y de la esfalerita, a_m es 5.3 g/cc.

Peso molecular del ZnS = 97.43 g/mol; Peso atómico: Zn = 65.37; S = 32.064 g/mol

Solución:

$$\frac{ZnS}{Zn} = \frac{97.43}{65.37} = 1.49$$

$$\% ZnS = 6.6\% \cdot 1.49 = 9.83 \%$$

$$\text{Contenido de ZnS en 10 Kg. de mineral: } 10 \text{ Kg.} \cdot 0.0983 = 0.983 \text{ Kg.}$$

La fracción de ZnS en el mineral, A_L , es:

$$A_L = \frac{0.983 \text{ Kg}}{10 \text{ Kg}} = 0.0983$$

$$\frac{d}{d_b} = \frac{2.0}{0.02} = 100$$

De la Tabla 3.2, el valor de $b = 0.05$

Usando: $f = 0.5$; $g = 0.25$, y aplicando la ecuación descrita anteriormente, se tiene:

$$V_F = \left[\frac{1}{10} - \frac{1}{20,000} \right] * 10^{-3} \left[\frac{1-0.098}{0.098} (1-0.098) * 5.3 + (0.098 * 2.5) \right] * 0.5 * 0.25 * 0.05 * 2^3$$

$$V_F = 0.000216 = 2.16 * 10^{-4}$$

$$S_F = \sqrt{V_F} = 0.0147$$

El error fundamental de muestreo es de aproximadamente 1.5 % para el ZnS o de 1.0 % para el Zn. En base a las probabilidades estadísticas, dos veces el valor de S_F proporciona un 95% de confianza de tener los valores de ensayo para el Zn en el intervalo $6.6\% \pm 2.0$.

Ejemplo de aplicación:

En el ejemplo de aplicación anterior, si el valor $2 S_F$ para el Zn es de 0.2%, ¿qué peso de muestra se debe tomar?

Solución:

El valor $2 S_F$ para el ZnS es de 0.3 % o equivalente a 0.003. De aquí $S_F = 0.0015$. Aplicando la ecuación:

$$V_F = \left[\frac{1}{W_s} - \frac{1}{W_L} \right] \left[\frac{1-A_L}{A_L} (1-A_L) a_m + A_L a_g \right] f g b d^3$$

y teniendo en cuenta que el valor $1/W_L = 1/20,000 = 5 \times 10^{-5}$ se puede despreciar por pequeño, se tiene:

$$(0.0015)^2 = \left[\frac{1}{W_s} \right] * 10^{-3} \left[\frac{1-0.098}{0.098} (1-0.098) * 5.3 + (0.098 * 2.5) \right] * 0.5 * 0.25 * 0.05 * 2^3$$

$$2.25 * 10^{-6} = \left[\frac{1}{W_s} \right] * 10^{-3} [9.2 * 4.78 + (0.245)] * 0.05$$

$$2.25 * 10^{-6} = \left[\frac{1}{W_s} \right] * 2.211 * 10^{-3}$$

$$W_s = 982.7 \text{ Kg}$$

Es decir, de un lote de 20,000 Kg. Se debe tomar aproximadamente 1,000 Kg. de mineral, lo que representa aproximadamente un 5 % del total. Téngase en cuenta de que el intervalo de certeza para los ensayos de Zn se desea que estén en $6.6\% \pm 0.2$, lo cual de por si es una exigencia de alta precisión para el muestreo, por ello la necesidad de esa elevada cantidad de muestra.

Ejemplo de aplicación:

El intervalo de confianza para los ensayos de Zn puede ser establecido arbitrariamente, de acuerdo a criterios prácticos de operación. Supóngase para el ejemplo anterior que se desea establecer el siguiente intervalo: $6.6\% \pm 0.5$, ó $[6.1, 7.1]\%$ en el cual con 95% de certeza se obtenga el valor real de ensayo de Zn, ¿cuál deberá ser el peso de muestra a tomarse del lote de 20,000 Kg.?

Solución:

Para el ZnS, 9.9 ± 0.75

$2 S_F = 0.75\%$; $S_F = 0.375\%$ ó $S_F = 0.00375$

$$(0.00375)^2 = \left[\frac{1}{W_s} \right] * 10^{-3} \left[\frac{1-0.098}{0.098} (1-0.098) * 5.3 + (0.098 * 2.5) \right] * 0.5 * 0.25 * 0.05 * 2^3$$

$$1.406 * 10^{-5} = \left[\frac{1}{W_s} \right] * 10^{-3} [9.2 * 4.78 + (0.245)] * 0.05$$

$$1.406 * 10^{-5} = \left[\frac{1}{W_s} \right] * 2.211 * 10^{-3}$$

$$W_s = 157.23 \text{ Kg}$$

De estos dos últimos ejemplos se concluye que mientras más precisión se requiera en el muestreo, mayor deberá ser la cantidad de muestra a tomar.

Este método de cálculo teórico de muestreo de minerales tiene en consideración los siguientes detalles:

1. Se debe tener alto nivel de confianza del valor promedio de ensaye del elemento de interés. En el caso de los ejemplos, el valor de 6.6 % de Zn se obtuvo como valor promedio de una gran cantidad de ensayos del cuerpo mineral del que fue extraído el lote
2. El método de cálculo descrito solo considera la existencia de una sola especie mineralógica de interés, ZnS. Para el caso en que dos ó más especies mineralógicas contengan el elemento de interés, se deberá considerar un promedio ponderado de todas las propiedades que se necesiten en la ecuación.
3. El valor W_s calculado es el mínimo teórico requerido desde el punto de vista estadístico. Como quedó mencionado anteriormente, se considera el error fundamental como el único susceptible de ser cuantificado. Los otros errores, tales como método de muestreo, heterogeneidad variable del mineral, instrumentos de muestreo, preparación de muestra, etc., al no poder ser cuantificados, contribuirán a hacer la operación de muestreo más imprecisa, lo que obligará a tomar mayor cantidad de muestra.

9.9. Planes de muestreo.

En la práctica muchos planes de muestreo son el resultado de una combinación de sentido común, reglas empíricas y factores prácticos, todos los cuales son difíciles de cuantificar. Consecuentemente, es difícil evaluar los errores inherentes a los planes de muestreo basados en tales criterios. Existe tres patrones básicos de muestreo a los cuales se hace frecuente referencia: a) Un bloque en el que cada una de las muestras es ubicada aleatoriamente; b) Un bloque dividido en n pequeños bloques con una sola y aleatoriamente ubicada muestra en cada uno de los bloques; y c) Un bloque dividido en pequeños bloques con una sola muestra cada bloque ubicada al centro de cada bloque; Figura 9.3.

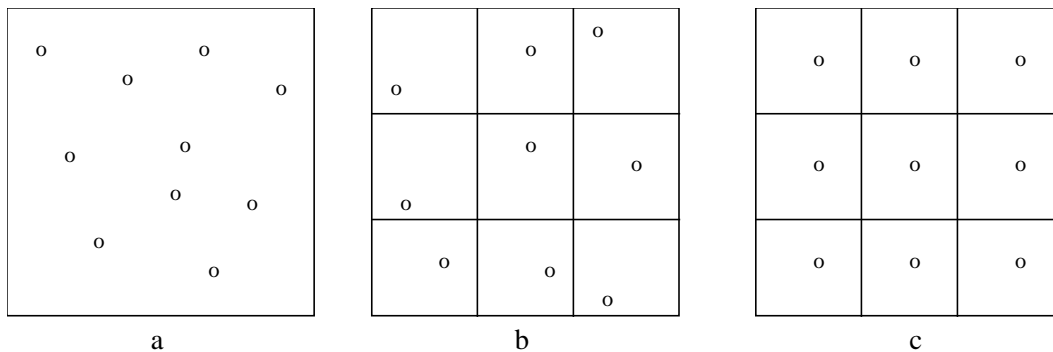


Figura 9.3: Patrones de muestreo: a) muestreo aleatoria en un bloque; b) muestreo aleatorio en pequeños bloques; c) muestreo ordenado en el centro de pequeños bloques

Con alguna base estadística, una regularidad en el plan de muestreo es preferible. Para variables que están fuertemente correlacionadas, tales como ley, distribución, segregación, etc., resulta mucho más conveniente efectuar muestreo con distribuciones regulares, tal como la mostrada en la Figura 9.3 c) .

Adicionalmente, se proponen una serie de patrones empíricos basados en planes regulares de muestreo, tal como los mostrados en la Figura 9.4.

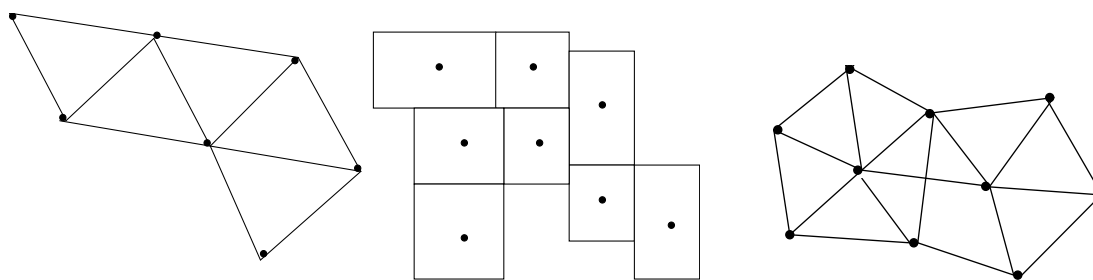


Figura 9.4: Modelos empíricos para obtener muestras de lotes.

El plan de muestreo depende de elecciones arbitrarias, tales como área y volumen de influencia, necesidades de precisión, cantidades de muestra requerida, etc. Los procedimientos que se establezcan se pueden ir modificando conforme se vaya ganando experiencia[•].

[•] Mular, A. L; Roshan Bhappu; “Mineral Processing Plant Design” 2nd Edition; Society of Mining Engineers of the AIME; New York, 1980.