

CUANTAS SUBMUESTRAS SON NECESARIAS PARA OBTENER UNA BUENA MUESTRA DE SUELO PARA ANALISIS DE LABORATORIO?

F. Pimentel Gómes *

Introducción

Para realizar un análisis de suelo el agricultor obtiene, mediante un barreno u otra herramienta, muestras de suelo a una profundidad de 0-20 cm, en un cierto número (n) de sitios de toda el área en estudio. De cada sitio se toman alrededor de 500 ml de suelo (medio litro) y esto constituye lo que se denomina muestra simple (también conocida como sub-muestra). Al mezclarse cuidadosamente todas las (n) muestras simples se obtiene una muestra compuesta, de la cual se separa el material que se envía al laboratorio para ser analizado. Aun cuando el área muestreada (por ejemplo 10 ha) se considere razonablemente uniforme, las muestras simples siempre difieren entre si y su variación influye en los datos analíticos obtenidos de la muestra compuesta. Aquí cabe mencionar una ley estadística que dice que una muestra compuesta será más representativa cuanto mayor sea el número de muestras simples. En otras palabras, valores obtenidos de una muestra compuesta de 5 muestras simples son menos confiables o representativos de aquellos obtenidos de 20 o 30 muestras simples.

Como calcular la precisión del resultado

Si se considera, por ejemplo, que el pH de una muestra compuesta de n muestras simples es \hat{A} , entonces el error patrón de esa media \hat{A} , en porcentaje, será :

$$E\% = \frac{CV}{(n)^{1/2}}$$

CV = coeficiente de variación que se

calcula de la siguiente forma:

$$CV = \frac{s}{\hat{A}} \times 100$$

s = desviación estándar relativa a cada muestra simple. El valor de s es desconocido y no se calcula en los ejemplos que se dan a continuación, pero se puede obtener fácilmente por medio de métodos estadísticos simples (Pimentel Gomes, 1987).

Ejemplo 1

En su tesis de doctorado, Chitolina (1982) obtuvo en una área homogénea de 10 ha en un Oxisol arenoso una media de pH (\hat{A}) de 4.80 y una s de 0.13. Con estos datos se calcula el coeficiente de variación de la siguiente forma:

$$CV = \frac{0.13}{4.80} \times 100 = 2.70\%$$

Este coeficiente de variación es bastante bajo en relación a valores mayores al 20 % encontrados para potasio (K) intercambiable y el fósforo (P) soluble en ese mismo suelo y en otros suelos.

El error patrón de la media en porcentaje es en este caso:

$$E\% = \frac{2.70}{(n)^{1/2}}$$

Si n = 5 muestras simples (o sitios de toma de muestra) :

$$E\% = \frac{2.70}{(5)^{1/2}} = 1.21\%$$

Si se toma como 100% la media del

pH medido en las 5 muestras ($\hat{A} = 4.80$), se puede afirmar, con un 95% de probabilidad, que la verdadera media (desconocida) estará entre :

$$100 - (2 \times 1.21) = 97.6\%$$

y

$$100 + (2 \times 1.21) = 102.4\%$$

del valor $\hat{A} = 4.80$ estimado, esto es, entre :

$$\frac{97.6}{100} \times 4.80 = 4.68$$

y

$$\frac{102.4}{100} \times 4.80 = 4.92$$

El resultado obtenido es excelente y demuestra que para pH el uso apenas de 5 sitios (o muestras simples) es suficiente. En el caso del análisis de otros factores éste puede no ser suficiente.

Ejemplo 2

Consideremos el K, en el mismo Oxisol arenoso, estudiado por Chitolina (1982). Tenemos que n = 5 muestras simples, $\hat{A} = 15$ ug/ml, s = 5.93. Entonces el coeficiente de variación es :

$$CV = \frac{5.93}{15} \times 100 = 39.5\%$$

El error patrón de la media, en porcentaje, en el caso de la muestra compuesta formada por n = 5 muestras simples, es :

* Profesor Catedrático de ESALQ/USP y Director de la Revista de Agricultura. Casilla Postal 60, 13400-970. Piracicaba-SP.

$$E\% = \frac{39.5}{(5)^{1/2}} = 17.7\%$$

Tomando nuevamente como 100% el contenido de K obtenido $\hat{A} = 15$ ug/ml, la verdadera media (desconocida), con 95% de probabilidad debe estar entre :

$$100 - (2 \times 17.7) = 64.6\%$$

y

$$100 + (2 \times 17.7) = 135.4\%$$

Este intervalo es excesivo, demostrando claramente que 5 muestras simples son insuficientes.

En números absolutos, el contenido de K estaría entre :

$$\frac{64.6}{100} \times 15 = 9.7 \text{ ug/ml}$$

y

$$\frac{102.4}{100} \times 15 = 20.3 \text{ ug/ml}$$

Se obtendría un mejor resultado con 20 muestra simples ($n = 20$), como se demuestra a continuación :

$$E\% = \frac{CV}{n} = \frac{39.5}{(20)^{1/2}} = 8.83\%$$

$$100 - (2 \times 8.83) = 82.3\%$$

y

$$100 + (2 \times 8.83) = 117.7\%$$

La variación de la media, esto es, su intervalo de confianza, en porcentaje estaría entre 82,3% y 117,7% de la media $\hat{A} = 15$. Los límites absolutos del contenido de K serían :

$$\frac{82.3}{100} \times 15 = 12.3 \text{ ug/ml}$$

y

$$\frac{117.7}{100} \times 15 = 17.7 \text{ ug/ml}$$

Se concluye entonces que el intervalo de confianza, a 95% de probabilidad, estaría entre 12.3 y 17.7 mg K/ml. Esta precisión parece ser aceptable y se podría recomendar la obtención de 20 muestras simples para tener una buena precisión en la estimación del contenido de K en este suelo. Si se considera insuficiente la precisión obtenida se debe tomar un número mayor de muestras simples.

Otro camino a seguir

Al adoptarse el método antes expuesto, el intervalo de confianza, en porcentaje, se ubica entre los extremos $100 - 2E\%$ y $100 + 2E\%$. Se dice entonces que existe el 95% de probabilidad de que la verdadera media (desconocida) esté entre esos dos valores. Sin embargo, cuando el coeficiente de variación es muy grande el método obliga a usar un número (n) de muestras simples excesivamente alto (80 por ejemplo). En estos casos puede ser aconsejable considerar un nivel de probabilidad menos exigente, 90% por ejemplo. Si este fuera el caso, los extremos del intervalo de confianza en porcentaje, pasaría a ser, aproximadamente:

$$100 - 1.7 E\% \quad \text{y} \quad 100 + 1.7 E\%$$

Ejemplo 3

Utilizando los datos del ejemplo 2 (contenido de K) y reconstruyendo los cálculos con este nuevo nivel de probabilidad se tiene que el $CV = 39.5\%$ y $n = 5$. De esta forma, $E\% = 39.5/(5)^{1/2} = 17.7\%$, como en el caso anterior, pero los extremos del intervalo de confianza son menos discretos:

$$100 - 1.7 \times 17.7\% = 69.9\%$$

y

$$100 + 1.7 \times 17.7\% = 130.1\%$$

El tamaño del intervalo de confianza ahora es $130.1 - 69.9 = 60.2\%$, mientras tanto que en el caso de 95% de probabilidad era $135.4 - 64.6 = 70.8\%$

En el caso de adoptar el 90% de probabilidad y $n = 20$, el tamaño del intervalo de confianza descendería al 30% con extremos de 85.0 y 115.0%, lo que sería tal vez aceptable.

Comentario final

Como se acaba de ver, el error patrón de la media obtenida en análisis químico de una muestra compuesta de suelo depende del coeficiente de variación (CV) y del número de muestras simples (n). Sin embargo, el coeficiente de variación depende de la heterogeneidad del área muestreada, así como del tipo de análisis químico, como se puede observar en los datos de la Tabla I.

Los datos de la Tabla I, que presentan resultados de análisis en dos tipos de suelo, indican que existe un bajo coeficiente de variación para pH y materia orgánica, pero que este valor es

Tabla I. Valores de coeficiente de variación (CV) para diversos resultados analíticos, investigado por Chitolina (1982) teniendo como unidad de muestreo la muestra simple.

Suelo	pH	Materia Orgánica	Aluminio intercambiable	Potasio intercambiable	Fósforo soluble	CV %	
Oxisol arenoso	2.7	2.2	38.1	39.5	22.1		
Ultisol	5.2	3.9	17.5	30.3	36.3		

relativamente alto para aluminio intercambiable, potasio intercambiable y fósforo soluble. De lo anterior se deduce que el número de muestras simples suficientes para obtener buena precisión en los contenidos estimados de pH o materia orgánica, no es suficiente para aluminio intercambiable, potasio intercambiable o fósforo soluble. En consecuencia, se debe determinar el número necesario de muestras simples teniendo en cuenta los análisis de mayor coeficiente de variación. Este número de muestras simples dará más precisión de la necesaria a pH y materia orgánica, pero suficiente para las otras determinaciones.

Finalmente, los resultados de Chitolina (1982) expuestos en la Tabla I indican, que en el peor de los casos (controlados por fósforo soluble y el potasio intercambiable), es probable que con 30 muestras simples (o sitios) se obtenga un valor de E del orden del 10% y un intervalo de confianza de aproximadamente 30%, con un nivel de 90% de probabilidad. Esta precisión debe ser suficiente en la mayoría de los casos prácticos.

OBSERVACION IMPORTANTE

Lectores con conocimientos razonables de estadística habrán notado que en este trabajo se ha evitado un tratamiento más completo y sofisticado del asunto, especialmente en lo que se refiere a la consideración de los números de grados de libertad disponible. Esto se hizo para lograr un artículo más comprensible sin pérdida de su utilidad práctica.

Bibliografía

Chitolina, J.C. 1982. Contribuicao de algunos fatores nos resultados analiticos da análise química de terra e seus efeitos nas recomendações de adubacao e calagem. Piracicaba, 200p. (ESALQ/USP-Tese de Doutorado)

Pimentel-Gomes, F. A. 1987. Estadística moderna na pesquisa agropecuária. 3ed. Piracicaba, POTAFOS, 162p.♣