

REDUCCION DEL TAMAÑO DE LA MUESTRA EN LA MEDIDA DEL GRANO DE CAFE

Germán Moreno-Ruiz *
Jaime Castillo-Zapata **

INTRODUCCION

El actual programa de mejoramiento genético en el Centro Nacional de Investigaciones de Café incluye el estudio de un número de árboles superior a 30.000, que tiende a incrementarse al avanzar la obtención de nuevas generaciones de los cruzamientos. Para la evaluación del tamaño del grano generalmente se toman dos muestras a cada árbol en cada una de tres cosechas, labor que implica el manejo de un volumen considerable de café.

La utilidad de emplear muestras pequeñas es evidente. En primer lugar, el sistema de beneficio húmedo es laborioso y requiere mucho espacio, mano de obra, energía y equipo. El menor tamaño de la muestra implica, desde luego, ahorro en cada uno de estos aspectos.

En segundo lugar, no siempre es viable la alternativa de reducir al mínimo el número de muestras, ya que algunos estudios requieren información sobre las variaciones más notables entre cosechas diferentes o aún dentro de una misma cosecha. Este hecho conduce a la multiplicación del número de muestras por factores numéricos que dependen de la variabilidad observada y de la precisión requerida en los estudios.

* Asistente de la Sección de Fitomejoramiento del Centro Nacional de Investigaciones de Café, Cenicafé, Chinchiná, Caldas, Colombia.

** Jefe de la Sección de Fitomejoramiento del Centro Nacional de Investigaciones de Café, Cenicafé, Chinchiná, Caldas, Colombia.

El trabajo aquí descrito tiene por objeto explorar la posibilidad de disminuir la magnitud de la muestra empleada para medir el tamaño del grano en café. La meta es seleccionar una muestra suficientemente pequeña que resulte práctica en su manipulación física, y al mismo tiempo permita medir las diferencias entre tratamientos experimentales con suficiente exactitud. Eso se lograría si los resultados obtenidos con la muestra tradicionalmente usada y con las más pequeñas fueran esencialmente los mismos, es decir, si las discrepancias entre ellas fueran exclusivamente de origen aleatorio.

MATERIALES Y METODOS

La muestra empleada tradicionalmente en Cenicafé para el estudio del tamaño de las semillas es de 250 gramos de café almendra (café verde), con una humedad aproximada de 110/0. En los ensayos reseñados se evalúa la posibilidad de reducir el tamaño de esta muestra.

El estudio se hizo en semillas de 35 plantas de tercera generación, (F3), de cruzamientos de *C. arabica* variedad Caturra x Híbrido de Timor, material en que se presenta una amplia variación de tamaño de grano. Los resultados obtenidos en estos materiales se comprobaron posteriormente por medio de un ensayo adicional, en un conjunto de variedades comerciales, dentro de las cuales las dimensiones de las semillas son menos variables.

En cada uno de los 35 árboles mencionados, se obtuvieron cantidades de 250, 100 y 50 gramos de café almendra. La cosecha de un árbol en una recolección (frutos maduros) se dividió en tres porciones que fueron procesadas separadamente hasta completar las cantidades indicadas. Posteriormente, cada muestra se pasó dos veces a través de un conjunto de tres zarandas cuyos orificios tenían 17/64, 16/64 y 14/64 de pulgada respectivamente. Se empleó el promedio de ambas medidas para la evaluación estadística.

Como índice del tamaño de grano se utilizó el café "supremo", que es el porcentaje retenido por la zaranda con orificio de 17/64 de pulgada. Se analizaron además los porcentajes de café de las zarandas menores, ya que éstos pueden emplearse en el cálculo de otros índices o medidas de dicho tamaño.

Análisis estadístico

La información proveniente de los árboles F3 fue analizada estadísticamente por medio de dos pruebas no paramétricas. Estas pruebas son de fácil ejecución y no requieren que los datos tengan determinada distribución. Además, permiten utilizar los registros en forma de rangos, aspecto que en la práctica es de uso común para la selección de los mejores individuos.

El análisis de varianza de la clasificación de rangos, diseñado por Kruskall y Wallis (1, 2), permite decidir si las discrepancias entre las muestras evaluadas son genuinas diferencias entre poblaciones distintas o, si por el contrario, son simples variaciones aleatorias, como las esperadas entre muestras de la misma población. Se parte de este último supuesto como hipótesis nula.

Cada valor de las N observaciones, que comprende el conjunto de k muestras, es reemplazado por el rango que le corresponde, partiendo del valor más bajo, que lleva el número 1, hasta el más alto, que lleva el número N. Se efectúa la suma de los rangos, (R_j), en cada una de las k muestras y se calcula el valor H según la fórmula siguiente:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1)$$

Esta fórmula se ajusta al procedimiento común para el análisis de varianza y está calculada a partir de los valores algebraicos de la suma de los N primeros números naturales y de la suma de sus cuadrados. El valor H está distribuido como Ji cuadrado, con k-1 grados de libertad.

La correlación de rangos de Spearman (1,2), otra prueba no paramétrica, se empleó como un complemento de la anterior, puesto que permite conocer el grado de asociación entre dos clasificaciones o variables medidas en los mismos individuos. La hipótesis nula supone que no existe asociación entre las clasificaciones.

Se determinan los rangos de cada una de las variables y se calcula su coeficiente de correlación (r_s), según la fórmula descrita a continuación, la cual se expresa en función de las diferencias observadas entre los rangos de un mismo individuo (d_i) en las dos series de datos.

$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N}$$

La comprobación de los resultados obtenidos en los árboles F3 se efectuó por medio de un ensayo adicional en ocho variedades, en las cuales está representada la variación del grano en cultivares comerciales. En este caso, se emplearon solamente dos tamaños de mues-

tra, 250 y 100 gramos. La medida de la clasificación tuvo tres repeticiones y la evaluación estadística se hizo por medio de análisis de varianza y correlación.

RESULTADOS

En la tabla 1, aparecen los porcentajes de café supremo (índice del tamaño del grano) para cada árbol, determinados en muestras de 250, 100 y 50 gramos. En la parte derecha de la misma tabla se detallan los rangos ocupados por cada valor en la serie de 105 datos.

Se destaca la amplia variación de los árboles en cuanto a la cantidad de café supremo, que va de 6.6% hasta 88.9%.

Es notoria la coincidencia de los rangos correspondientes a cada árbol. En los casos en que la diferencia entre ellos es grande, árboles Nº 19 y 21, por ejemplo, las diferencias en café supremo solo son del orden de 5% y 7%, aproximadamente.

Los valores de H en la prueba de Kruskal y Wallis, que se presentan en la tabla 2, indican una alta probabilidad de que los datos obtenidos en las tres muestras sean fundamentalmente equivalentes. Esta probabilidad ocurre tanto en el café supremo, tomado como índice del tamaño, como en los tres tamaños de grano separados por las otras zarandas que se emplearon para el tamizado. El valor de $H = 0,083$, por ejemplo, correspondiente a la zaranda de 17/64, está muy distante de 5,99 y de 9,21, que son los niveles críticos calculados con probabilidades de 5% y 1%, las más empleadas en experimentación. En consecuencia, puede afirmarse, con una gran probabilidad de acierto, que las series de rangos, establecidas con base en las muestras de 250, 100 y 50 gramos, tienen origen en una misma población.

El segundo aspecto analizado fue el grado de asociación que tienen los datos del tamaño del grano de un mismo árbol, cuando proceden de muestras de diferente magnitud. Los coeficientes de correlación de rango de Spearman, que se presentan en la tabla 3, indican que estos datos están estrechamente asociados en forma positiva, pues los altos valores de r_s encontrados tienen una probabilidad muy baja (menor de 0,0005) de ser producto del azar.

Los resultados referentes a las variedades comerciales se muestran en las tablas 4 y 5. Las cifras de la tabla 4, indican que las diferencias entre los porcentajes de café, separados para las diferentes variedades, en cada tamaño de muestra, son muy pequeñas, por lo general menores de 1%, y que estos porcentajes están estrechamente correlacionados, con coeficientes r que varían entre 0.96 y 0.99. Se aprecia también que los coeficientes de variación

TABLA 1.- PORCENTAJES DE CAFE SUPREMO (MAYOR DE 17/64 DE PULGADA) DE 35 ARBOLES, DETERMINADOS EN MUESTRAS DE TRES TAMAÑOS Y RANGOS OCUPADOS POR CADA ARBOL EN LA SERIE DE 105 DATOS.

Arbol Nº	PORCENTAJES				RANGOS		
	Tamaño de muestra (gramos)				Tamaño de muestra (gramos)		
	Prom.	250	100	50	250	100	50
1	6.6	5.4	8.7	5.6	1	3	2
2	12.9	19.4	10.0	9.2	7	5	4
3	19.3	20.4	20.3	17.3	9	8	6
4	28.0	21.3	32.8	30.0	10	16	14
5	29.7	29.8	31.6	27.6	12	15	11
6	33.8	35.9	30.0	35.5	18	13	17
7	38.5	40.0	38.2	37.2	22	20	19
8	43.8	45.3	46.8	39.3	26	28	21
9	45.5	46.6	45.2	44.8	27	25	23
10	48.2	52.5	46.9	45.1	33	29	24
11	50.3	52.7	48.3	49.8	34	30	31
12	54.8	54.5	58.2	51.6	35	38	32
13	58.3	56.8	58.0	60.2	36	37	40
14	61.5	58.4	61.8	64.2	39	43	52
15	63.7	65.2	63.9	61.9	56	50	44
16	63.9	61.3	64.9	65.5	42	53	59
17	64.0	63.1	63.7	65.1	47	49	55
18	64.2	63.9	66.7	62.0	51	62	45
19	64.3	65.3	61.3	66.2	58	41	61
20	64.5	65.0	65.2	63.2	54	57	48
21	66.0	65.9	69.9	62.3	60	67	46
22	68.8	68.9	68.6	68.9	66	64	65
23	72.2	76.2	73.5	67.0	76	71	63
24	72.9	74.4	74.2	70.1	74	73	68
25	74.2	72.2	76.4	74.0	70	77	72
26	75.0	78.6	75.3	71.0	80	75	69
27	79.2	76.8	82.5	78.2	78	91	79
28	80.9	82.1	79.7	80.8	90	82	85
29	81.2	81.8	81.4	80.5	89	87	83
30	81.6	79.6	84.3	81.0	81	96	86
31	82.9	83.6	80.5	84.5	95	84	97
32	83.0	84.6	81.6	82.9	98	88	92
33	84.0	83.6	85.4	83.0	94	99	93
34	87.3	87.2	85.6	89.0	101	100	104
35	88.9	88.0	90.0	88.8	102	105	103

(c. v.) calculados para las muestras de 250 y 100 gramos en cada tamaño de semilla, solamente difieren en 1.50/0, lo cual indica que la precisión en los experimentos no se afecta sustancialmente al reducir el tamaño de las muestras.

El análisis de varianza combinado, que aparece en la tabla 5, denota que no existen diferencias significativas entre "tamaños de muestra" y que tampoco existe interacción "tamaño de muestra x variedad". Estos resultados confirman los anteriormente expuestos, en el sentido que el comportamiento de las variedades es el mismo al medir el tamaño de sus semillas con cualquiera de las muestras evaluadas.

TABLA 2.- PORCENTAJE MEDIO DE SEMILLAS DE DIFERENTE TAMAÑO SEPARADAS POR TRES ZARANDAS, EMPLEANDO MUESTRAS DE TRES TAMAÑOS; VALORES DE "H" EN LA PRUEBA DE KRUSKALL Y WALLIS.

Tamaño de la muestra (gramos)	Porcentaje			
	Tamaño de las semillas (n/64")			
	17	16	14	Menor de 14
250	60.17	18.15	16.52	5.16
100	60.30	19.12	16.41	4.17
50	58.95	18.70	17.73	4.62
Valor de H *	0.083	0.184	0.667	3.770

* Los niveles críticos de H para 2 g. l., con $p \geq 0.05$ y $p \geq 0.01$, son 5.99 y 9.21, respectivamente.

TABLA 3.- COEFICIENTES DE CORRELACION DE RANGO DE SPEARMAN (r_s), PARA LA CLASIFICACION DE 35 ARBOLES SEGUN LOS TAMAÑOS DE LAS SEMILLAS SEPARADOS POR TRES ZARANDAS, EMPLEANDO TRES TAMAÑOS DE MUESTRA.

Clasificaciones comparadas	Valores de r_s			
	Tamaño de semilla (n/64")			
	17	16	14	Menor de 14
250 gramos vs 100 gramos	0.97*	0.96*	0.88*	0.75*
250 gramos vs 50 gramos	0.97*	0.95*	0.94*	0.72*
100 gramos vs 50 gramos	0.96*	0.94*	0.88*	0.75*

* La significación de r_s se mide por el criterio de t, que se calcula por la fórmula:

$$t = r_s \sqrt{\frac{N - 2}{1 - r_s^2}}$$

Para $P \geq 0.0005$, con 33 g. l., el valor de t es 3.6.

TABLA 4.- PORCENTAJE MEDIO DE SEMILLAS DE DIFERENTE TAMAÑO SEPARADAS POR TRES ZARANDAS EN OCHO VARIEDADES DE CAFE, EMPLEANDO DOS TAMAÑOS DE MUESTRA; CORRELACION (r) ENTRE ESTOS PORCENTAJES Y COEFICIENTES DE VARIACION (C. V.) PARA CADA TAMAÑO DE MUESTRA Y DE SEMILLA.

Variedad	Tamaños de semillas (n/64'')							
	17		16		14		Menor de 14	
	Tamaño de muestra (gramos)							
	250	100	250	100	250	100	250	100
A	77.1	76.4	10.7	11.9	9.7	9.7	2.3	1.9
B	72.0	70.8	14.3	15.0	11.5	12.3	2.1	1.7
C	68.8	70.9	16.4	15.6	12.5	11.6	2.2	1.7
D	62.2	61.9	22.7	22.4	12.8	13.9	2.2	1.7
E	44.9	47.2	34.1	34.4	16.7	15.1	4.1	3.4
F	36.4	37.1	33.2	32.9	26.0	25.7	4.2	4.3
G	34.4	36.4	35.9	36.8	24.7	22.5	4.9	4.3
H	25.6	27.3	38.6	37.8	29.7	28.8	5.9	6.0
r*	0.99**		0.99**		0.97**		0.96**	
c. v. (°/o)	12.7	14.2	10.6	11.6	14.6	16.1	16.6	18.1

* Correlación calculada para 24 pares de datos, originados por ocho variedades y tres repeticiones.

** Significativo con $P \leq 1\%$.

TABLA 5.- CUADRADOS MEDIOS DETERMINADOS EN LOS ANALISIS DE VARIANZA DEL CAFE RETENIDO POR TRES ZARANDAS EN OCHO VARIEDADES, CON MUESTRAS DE 250 Y 100 GRAMOS DE CAFE VERDE.

Fuente de variación	G. L.	Tamaño de semilla (n/64'')			
		17	16	14	Menor de 14
Variedades	7	798.5**	370.7**	177.1**	35.8**
Tamaño de muestra	1	6.3	0.1	1.4	5.2
Variedad x tamaño	7	2.7	0.6	1.1	0.3
Error global	28	40.2	11.2	14.1	3.1

** F significativo al nivel de 1% .

DISCUSION

Los resultados expuestos indican que la muestra de 250 gramos de café verde, empleada actualmente para medir el tamaño de la semilla, puede reducirse a 100 ó 50 gramos, sin causar alteraciones en los resultados experimentales. En primer lugar, en el caso de árboles individuales, las diferencias detectadas entre las muestras son estrictamente de carácter aleatorio y de pequeña magnitud, y por lo tanto, cualquiera de ellas es válida para evaluar la población original. Además, existe una estrecha asociación positiva entre las series de datos obtenidos para los mismos árboles con los diferentes tamaños de muestra, lo cual indica que éstos quedan clasificados en la misma forma, independientemente de los tamaños de muestra probados.

En segundo lugar, en el caso de poblaciones más homogéneas, como las variedades comerciales, la reducción del tamaño de la muestra tampoco produce modificaciones en la medida del grano. Esta afirmación está soportada en el análisis de varianza por la carencia de diferencias significativas entre "tamaño de muestra" y por la ausencia de interacción "tamaño de muestra x variedad". Como podía esperarse, las cantidades de café separadas por las zarandas, en muestras de diferente tamaño, están correlacionadas y la variación experimental permanece prácticamente inalterada.

Los bajos valores de H y las correlaciones significativas observadas para las zarandas con orificios menores, indican que el tamaño de grano, apreciado por medio del café supremo, o por otro índice basado en estas zarandas, será el mismo, independientemente del tamaño de muestra empleado.

Aparte de las ventajas prácticas ya mencionadas, que se refieren al ahorro de material, espacio, mano de obra y energía para el secado de las muestras, la reducción de la magnitud de las muestras facilita los estudios de tamaño de grano. Por una parte, el tamaño reducido permite que una mayor proporción de árboles sea muestreada en cada recolección. Por otra parte, se hace posible estudiar un mayor número de variables en las exploraciones de la variación debida a diferentes factores experimentales, y éstos podrán analizarse en forma más completa.

RESUMEN

Con el propósito de disminuir la cantidad de café verde empleada como muestra para determinar el tamaño de las semillas, se analizó el efecto de emplear 100 y 50 gramos de semilla en lugar de la muestra tradicional de 250 gramos. Para ello, se compararon los por-

centajes de café retenidos por zarandas con orificios de 17/64, 16/64 y 14/64 de pulgada en cada tamaño de muestra, en un grupo de 35 árboles F3 del cruzamiento Caturra x Híbrido de Timor y en ocho variedades comerciales.

Se encontró que la clasificación de los árboles, según el porcentaje de café retenido en cada zaranda, no está influida por el tamaño de muestra empleado. Además, las tres series de medidas del tamaño del grano, obtenidas con muestras de diferente magnitud, están estrechamente correlacionadas.

La disminución del tamaño de la muestra tampoco afectó la medida del grano en las variedades comerciales, y solamente produjo aumentos menores del 2% en la variabilidad del experimento.

Se concluye que la muestra de 250 gramos, usada tradicionalmente para los estudios de tamaño de grano, puede reducirse a 100 ó 50 gramos sin causar alteraciones en los resultados experimentales.

SUMMARY

One experiment was designed with the purpose of decreasing the sample size of green coffee. Fifty, one hundred, and two hundred and fifty gram samples were compared on the basis of beans retained by sieves measuring 14/64", 16/64" and 17/64", for each one of the above sample sizes, which in turn were taken from 35 coffee trees belonging to the F3 generation of the cross Caturra x Timor Hybrid and eight commercial varieties.

The results showed that the classification of the trees, according to the percentage of coffee withheld in each sieve, is not influenced by the sample size. Furthermore, the three series of measurements taken from the different sample sizes were closely correlated.

The reduction of the sample size did not affect the measurements of the coffee beans originated from different commercial varieties and only slightly increased the variability of the experiment in about two percent.

It is concluded that the traditional 250 gram sample used for estimating bean size could be reduced to either fifty or 100 grams without altering the experimental results.

BIBLIOGRAFIA

- 1.- SIEGEL, S. Estadística no paramétrica aplicada a las ciencias de la conducta. México, Trillas, 1974 346 p.
- 2.- STEEL, R. G. and TORRIE, J. H. Principles and procedures of statistics with special reference to the biological sciences. New York, McGraw-Hill, 1960. 481 p.