Se quiere averiguar si tres fertilizantes, A, B y C presentan diferencias significativas en cuanto a sus efectos sobre el aumento de la cosecha.

Con este propósito se eligieron al azar 15 parcelas a las que se fertilizó aleatoriamente con cada uno de los fertilizantes en cuestión. Los aumentos de cosecha obtenidos fueron los siguientes

Fertilizante	*Aumento de cosecha*				
A	39	33	39	35	32
B	36	40	35	30	29
C	33	33	36	26	35

A la vista de estos datos, ¿puede inferirse que existen diferencias significativas entre los tres fertilizantes a nivel $\alpha = 0'05$?

Se trata de un *Análisis de la Varianza para un factor en un diseño completamente aleatorizado*, cuyos fundamentos y desarrollos teóricos aparecen en CB-sección 8.2, con el que se quiere contrastar la hipótesis nula de igualdad de los efectos medios de los tres fertilizantes, $H_0 : \mu_A = \mu_B = \mu_C$, frente a la alternativa de no ser los tres iguales.

Como en todos los contrastes de este tipo, lo primero que debemos determinar es la tabla de Análisis de la Varianza, la cual es

F. de variación	Suma de cuadrados	g.l.	c. medios	Estadístico
Fertilizantes	$SST_i = \sum_{i=1}^{r} \frac{T_i^2}{n_i} - \frac{T^2}{n}$	$r-1$	$\frac{SST_i}{r-1}$	$\frac{SST_i/(r-1)}{SSE/(n-r)}$
Residual	$SSE = SST - SST_i$	$n-r$	$\frac{SSE}{n-r}$	
Total	$SST = \sum_{i=1}^{r} \sum_{j=1}^{n_i} x_{ij}^2 - \frac{T^2}{n}$	$n-1$		

Para calcular la suma de cuadrados SST_i, partiendo de la tabla de datos del enunciado, calculamos, en una última columna, los totales de cada tratamiento (totales por filas)

$$T_i = \sum_{j=1}^{n_i} x_{ij}$$

así como la suma de todos los totales,

$$T = \sum_{i=1}^{r} T_i$$

Fertilizante						T_i
A	39	33	39	35	32	178
B	36	40	35	30	29	170
C	33	33	36	26	35	163
						$T = 511$

El número de observaciones realizadas de cada tratamiento es $n_i = 5$ $i = 1, 2, 3$, y el número total de observaciones es

$$n = \sum_{i=1}^{r} n_i = 5 + 5 + 5 = 15.$$

La *suma de cuadrados debida a los fertilizantes* será, por tanto,

$$\begin{aligned} SST_i &= \sum_{i=1}^{r} \frac{T_i^2}{n_i} - \frac{T^2}{n} \\ &= \frac{178^2}{5} + \frac{170^2}{5} + \frac{163^2}{5} - \frac{511^2}{15} \\ &= 17430'6 - 17408'067 = 22'533. \end{aligned}$$

Sus grados de libertad son igual al número de tratamientos menos uno, $r - 1 = 3 - 1 = 2$.

La *suma total de cuadrados* es igual a la suma de los cuadrados de las observaciones menos el valor antes calculado T^2/n,

$$SST = \sum_{i=1}^{r} \sum_{j=1}^{n_i} x_{ij}^2 - \frac{T^2}{n}.$$

En nuestro ejercicio es igual a

$$\begin{aligned} SST &= \sum_{i=1}^{r}\sum_{j=1}^{n_i} x_{ij}^2 - \frac{T^2}{n} \\ &= 39^2 + 33^3 + \cdots + 26^2 + 35^2 - 17408'067 \\ &= 17617 - 17408'067 = 208'933. \end{aligned}$$

Sus grados de libertad son el número de observaciones menos uno, $n - 1 = 15 - 1 = 14$.

Por último, la *suma residual de cuadrados* se calcula como diferencia de las otras dos sumas de cuadrados antes calculadas:

$$SSE = SST - SST_i = 208'933 - 22'533 = 186'4.$$

Sus grados de libertad se calculan también como diferencia de los grados de libertad de las otras dos sumas de cuadrados que sirvieron para obtener SSE; es decir, $(n-1)-(r-1) = 14-2 = 12$.

De esta forma, determinamos las *sumas de cuadrados* y los *grados de libertad* de las tres fuentes de variación que forman la tabla de Análisis de la Varianza. Los *cuadrados medios* correspondientes a cada fuente de variación, se determinan ahora, simplemente, dividiendo cada suma de cuadrados por sus grados de libertad:
Cuadrado medio correspondiente a los fertilizantes:

$$\frac{SST_i}{r-1} = \frac{22'533}{2} = 11'266.$$

Cuadrado medio correspondiente a la suma residual de cuadrados:

$$\frac{SSE}{n-r} = \frac{186'4}{12} = 15'533.$$

Finalmente, el estadístico a utilizar en el contraste, se calcula dividiendo los cuadrados medios antes determinados:

$$F = \frac{SST_i/(r-1)}{SSE/(n-r)} = \frac{11'266}{15'533} = 0'73.$$

Todos estos cálculos se recogen en la tabla ANOVA siguiente:

F. de variación	Suma de cuadrados	g.l.	c. medios	Estadístico
Fertilizantes	$SST_i = 22'533$	2	$11'266$	$F = 0'73$
Residual	$SSE = 186'4$	12	$15'533$	
Total	$SST = 208'933$	14		

El estadístico F tiene, si es cierta la hipótesis nula de igualdad de los efectos medios de los tres fertilizantes, una distribución F de Snedecor con grados de libertad igual al par formado por los grados de libertad correspondientes a las fuentes de variación *Fertilizantes* y *Residual*, antes determinados, $(r-1, n-r) = (2, 12)$, por lo que para determinar el punto crítico, a un nivel de significación $\alpha = 0'05$, buscaremos en la tabla de la F de Snedecor (tabla 6) el valor $F_{(2,12);0'05} = 3'8853$. Al ser $F = 0'73$ menor que dicho punto crítico, se acepta H_0, concluyendo con la no existencia de diferencias significativas entre los tres fertilizantes.

De dicha tabla también se obtiene una acotación del p-valor:

$$\text{p-valor} = P\{F_{(2,12)} > 0'73\} > P\{F_{(2,12)} > 2'8068\} = 0'1$$

es decir, p-valor$> 0'1$, acotación suficiente como para poder concluir con la confirmación de la decisión tomada.