5.5 Modelo de regresión

El modelo de regresión lineal que se aplicó para explicar la relación entre producción antigüedad, edad, capital humano, apiarios y camioneta de los apicultores nayaritas se especificó en los términos siguientes:

$P_i = \alpha + \beta ANT_i + \beta Ed_i + \beta CH_i + \beta A_i + \beta C_i + e_i$

(6.3.1) $_{i=}$ 1,2,3,…..N

Donde:

P = Producción; ANT = Antigüedad; Ed = edad; CH = Capital Humano; A = Apiarios C = Camioneta; β = Coeficientes de la regresión; e_i = variable que recoge los errores i = Subíndice porque la muestra contiene datos de sección cruzada, y N = el número de observaciones de la muestra. Una vez que se corrió el modelo se obtuvieron lo resultados siguientes:

Tabla 5.2 Variable dependiente: Producción

	C	Antigüedad	Edad	Capital Humano	Apiarios	Camioneta
ŷ =	8253.01	74.58	-57.57	-727.50	81.12	-1104.51
Estadístico T=	(3.84)	(2.32)	(-2.02)	(-2.34)	(2.46)	(-1.78)
Probabilidad=	0.00	0.02	0.05	0.02	0.02	0.08

Observaciones = 112 g de l=106. F-Estadístico experimental = 5.85.
F. Estadístico 0.00.

Para conocer la significación global del modelo, se realizo la prueba de hipótesi de igualdad de variables por medio del estadístico F, que se obtiene del cociente d dividir la varianza explicada entre la varianza no explicada. Su fórmula es:

$$F_{k-1,n-k} = \frac{\frac{\sum \hat{y}_i^2}{(k-1)}}{\frac{\sum e_i^2}{(n-k)}} = \frac{\frac{R^2}{k-1}}{\frac{(1-R^2)}{(n-k)}} \qquad (5.7)$$

Donde: $(k-1)$ son los grados de libertad de la varianza explicada; $(n-k)$ los grados de libertad de la varianza no explicada.

En las tablas, Greene (2001) referidas a la distribución F teórica, significancia de 5%, y nivel de confianza del 95% con 6 grados de libertad para el numerador (varianza explicada) y 106 grados de libertad para el denominador, vemos que $F_{\alpha} = 2.19$. En cambio la distribución F empírica que calculamos enseguida:

$$F_{6,106} = \frac{\frac{0.21}{6}}{\frac{1-0.21}{106}} = \frac{0.035}{0.00745} = 4.69$$

Como $F_{6,106}$ =*4.69*> F_{α} =*2.19*, se rechaza la hipótesis nula. Por ello el modelo econométrico es globalmente significativo. En consecuencia se dice que la producción de miel, es explicada por la Antigüedad en el trabajo de los apicultores, su edad, su capital humano, la forma en que distribuyen sus apiarios y el uso de camionetas.

Por otra parte se obtuvo la prueba de Wald para obtener información del estadístico F y de χ^2, así como la probabilidad que dejan a la derecha de la distribución del valor muestral de los mismos, lo que permite obtener información directamente. El resultado de la prueba arrojó un valor de 8.63 para el estadístico F y probabilidad cercana a 0 en tanto que para el estadístico χ^2 fue de 25.9 y con una probabilidad cercana a 0 (véase tabla 5.3).

En ambos casos la probabilidad es menor que el nivel de significancia elegido (se utiliza normalmente 5%) se rechaza la hipótesis nula, por lo que se rechaza la igualdad entre los parámetros y de acuerdo a los datos que se utilizaron el modelo explica a la variable dependiente.

Tabla 5.3 Prueba de Wald[2]. Hipótesis nula = 0

Estadístico F	8.63		Probabilidad	0.000001
χ^2	25.90		Probabilidad	0.000001

Fuente: Elaboración propia con datos de la regresión.

Por otra parte, en lo que se refiere al resultado de la regresión se tiene que e estadístico *"t"* para la variable Antigüedad reporta, un valor de 2.32; para la variabl edad de -2.02; para Capital Humano -2.34; para apiarios 2.46 y Camioneta -1.78 co un valor de probabilidad de 0.0001, por lo cual se afirma que en conjunto las variable son significativas a un 95% de confianza.

Pero el coeficiente de correlación (R^2) nos está indicando que las variables tomada en cuenta explican en un 21% el modelo, en general en los modelos de regresión d corte transversal ese estadístico es bajo. En tanto que el estadístico F nos da cuent que tiene un valor de 5.85 de acuerdo al valor teórico y al calculado se puede afirma que en su conjunto las variables explican el modelo.

La probabilidad de F- estadístico con valor de 0.0001, cantidad cercana a cer muestra que el modelo es confiable. En general este modelo econométrico de cort transversal, dio cuenta de la situación en que se relacionan las variables.

Notas:

1 Se simboliza con la letra I, se puede calcular a partir de la ecuación:

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y},$$

Donde: σ_{XY} es la covarianza de (X,Y) ; σ_X es la desviación típica de la variable X ; σ_Y es la desviación típica de la variable Y (SPSS 15.0) también se usa la forma semejante para calcular ese coeficiente sobre la muestra estadística, denotándolo como r_{xy} a:

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\ \sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

2 Prueba de Wald: toma en consideración la correlación existente dentro de conglomerados, lo que hace que sea un procedimiento válido para probar la bondad del ajuste en algunos diseños muestrales complejos. El estadístico de

Wald para probar la bondad de ajuste es: $X_w^2 = (\hat{p} - p_0)' \hat{V}^{-1} (\hat{p} - p_0)$. Donde $\hat{p}$ es el vector de proporcione estimadas $(\hat{p}_1, \ldots \hat{p}_{k-1})$; $\hat{V}$ es un estimador de la matriz de covarianza de $\hat{p}$; y p_0 es el vector de proporciones teórica $(\hat{p}_{01}, \ldots \hat{p}_{0k-1})$. El estadístico X_w^2 se distribuye aproximadamente como una chi-cuadrado con $k-1$ grados de libertad.