

Taller 5 Estadística II

Autores

Giselle Tatiana Fernández López
José Luis Sánchez Escobar
Santiago Rendón Giraldo

Docente

Raúl Alberto Pérez Agamez

Asignatura

Estadística II



Sede Medellín
Noviembre de 2021

Índice

1. Ejercicio 1	2
1.1. Estimación del modelo, significancia e interpretación de coeficientes	2
1.2. Significancia de la regresión	2
1.3. Significancia parámetros individuales	3

Índice de figuras

Índice de cuadros

Warning: package 'leaps' was built under R version 4.0.4

1. Ejercicio 1

Se tiene del texto que se desea estimar un modelo de regresión múltiple que explique el riesgo de infección en términos de todas las variables predictoras, este modelo tendrá la forma

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i, \varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2); 1 \leq i \leq 76$$

1.1. Estimación del modelo, significancia e interpretación de coeficientes

Estimación del modelo

##	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	0.6025077830	0.745565728	0.8081216	0.4216108776
## X1	0.3102389507	0.089702236	3.4585420	0.0009042477
## X2	0.0420308437	0.011665795	3.6029130	0.0005671690
## X3	-0.0003404187	0.003116501	-0.1092311	0.9133148963
## X4	-0.0025909567	0.003982866	-0.6505257	0.5173674098
## X5	0.0042372532	0.002343412	1.8081553	0.0746453707

De esta tabla se obtienen los estimadores para los diferentes parámetros que utilizamos en el modelo.

$$\hat{y}_i = 0.6025 + 0.3102x_{i1} + 0.0420x_{i2} - 0.0003x_{i3} - 0.0026x_{i4} + 0.0042x_{i5}$$

Significancia e interpretación de coeficientes Para esta prueba utilizaremos el siguiente juego de hipótesis

$$H_0 : \beta_j = 0 \text{ vs } H_1 : \beta_j \neq 0; j = 0, \dots, 5$$

De la tabla obtenemos los valores p para cada uno de los parámetros, y usando un $\alpha = 0.05$ llegamos a estas conclusiones:

β_0 tiene un valor-p muy por encima de 0.05, por tanto no se rechaza la hipótesis nula, dejando así que no es significativo; por otro lado no se puede interpretar

β_1 tiene un valor-p menor que 0.05, entonces es significativo, y se puede interpretar como el aumento en Y en promedio 0.310 unidades cuando en X_1 se aumenta en una unidad, siempre que las otras variables de predicción se tengan constantes

1.2. Significancia de la regresión

Para probar la significancia de la regresión estableceremos las siguientes hipótesis

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0, \text{ vs}$$

$$H_1 : \text{algún } \beta_j \neq 0, j = 1, 2, \dots, 5$$

y utilizaremos la siguiente tabla ANOVA

##	Sum_of_Squares	DF	Mean_Square	F_Value	P_value
## Model	63.2061	5	12.641215	13.3516	2.82658e-09
## Error	70.0628	74	0.946795		

Haciendo el análisis de la tabla ANOVA se concluye que el modelo de regresión sí es significativo, puesto que su valor-p es menor a 0.05(2.82658e-09). Rechazando así H_0 , concluyendo que el riesgo de infección depende de al menos una de las variables predictoras del modelo.

1.3. Significancia parámetros individuales

Para probar la significancia utilizaremos