

Modelos de Variable Dependiente Binaria

Relación de Ejercicios

1. El profesor chiflado desea conocer cuáles son los factores que influyen en que sus alumnos aprueben o no la asignatura que imparte. Con tal objetivo, se ha de tener en cuenta que la evaluación continua de la asignatura está formada por la calificación obtenida en el examen de ordenador (\mathbf{O} , de 0 a 1.5 puntos) y la calificación obtenida en los exámenes tipo test (\mathbf{T} , de 0 a 1.5 puntos). Al mismo tiempo, se tiene en cuenta otros aspectos de interés como el género del alumno (\mathbf{G} , 1 si es femenino y 0 si es masculino) y la elección del alumno en la votación del padrino de la promoción (\mathbf{P} , 1 si el alumno ha votado al profesor chiflado para ser padrino de su promoción y 0 en caso contrario). Si el modelo se estima mediante un Logit, se obtienen los siguientes resultados:

$$\hat{\mathbf{z}}_i = -3.07 + 3.39 \cdot \mathbf{O}_i + 1.61 \cdot \mathbf{T}_i - 0.49 \cdot \mathbf{G}_i + 2.25 \cdot \mathbf{P}_i,$$

mientras que si se estima mediante un probit:

$$\hat{\mathbf{z}}_i = -1.755 + 1.9 \cdot \mathbf{O}_i + 0.95 \cdot \mathbf{T}_i - 0.24 \cdot \mathbf{G}_i + 1.25 \cdot \mathbf{P}_i.$$

Se pide contestar de forma razonada a las siguientes cuestiones:

- A partir de la estimación del modelo Logit, ¿cuántas veces incrementa la probabilidad de aprobar respecto de suspender de un estudiante que pasa de no votar a votar como padrino de promoción al profesor chiflado suponiendo que el resto de variables permanecen constantes?
 - A partir de la estimación del modelo Logit y teniendo en cuenta que la verosimilitud del modelo es 0.0005 y la del modelo con sólo constante es 0.000001, calcular el coeficiente de McFadden y decidir si el modelo es conjuntamente significativo.
 - A partir de la estimación del modelo Probit, ¿cuál es la probabilidad de aprobar de un alumno que no ha votado al profesor chiflado como padrino de promoción y que ha obtenido 0.7 puntos en el examen de ordenador y en los exámenes tipo test?
 - A partir de la estimación del modelo Probit, calcular e interpretar el odd de un alumno que ha votado al profesor chiflado como padrino de promoción y ha obtenido 0.7 puntos en el examen de ordenador y en los exámenes tipo test.
2. Usando un modelo Logit se ha estudiado la posesión de ordenador portátil (\mathbf{O} , 1 si el individuo tiene un ordenador portátil) en 100 individuos en función de la renta (\mathbf{R} , medida en miles de euros al mes), la edad (\mathbf{E} , medida en años) y el género (\mathbf{G} , 1 si es hombre y 0 si es mujer) obteniéndose los siguientes resultados:

$$\hat{\mathbf{z}}_i = \begin{matrix} -12.061 & + & 3.021 \cdot \mathbf{R}_i & + & 0.1752 \cdot \mathbf{E}_i & + & 2.3822 \cdot \mathbf{G}_i \\ (2.986) & & (0.7392) & & (0.0591) & & (0.9139) \end{matrix}$$

Entre paréntesis se muestran las desviaciones típicas. Se pide contestar de forma razonada a las siguientes cuestiones:

- a) Obtenga el efecto marginal de la renta de un hombre de 30 años que gana 2000 euros al mes.
 - b) ¿Cuántas veces es más probable que ese hombre sea propietario de un ordenador portátil que una mujer en las mismas condiciones?
 - c) Estudie si la variable edad tiene un coeficiente significativamente distinto de cero.
 - d) Analice la bondad del modelo ajustado sabiendo que la log-verosimilitud con solo el término constante es -52.69 y la del modelo completo -20.676.
3. Se ha hecho un estudio con 150 personas y para cada una se ha observado si se ha ido o no de vacaciones (\mathbf{V} , 1 si se ha ido de vacaciones y 0 en caso contrario), la renta disponible (\mathbf{R} , medida en euros) y si trabaja (\mathbf{T} , 1 si trabaja y 0 si no lo hace). Para ver cómo influyen la renta y el trabajo en la probabilidad de ir de vacaciones, se estima un modelo Logit obteniéndose los siguientes resultados:

$$\hat{z}_i = \begin{array}{cccc} -1.54 & + & 0.0005 \cdot \mathbf{R}_i & + & 0.45 \cdot \mathbf{T}_i \\ (0.82) & & (0.0002) & & (0.01) \end{array}$$

Entre paréntesis se muestran las desviaciones típicas. Se pide contestar de forma razonada a las siguientes cuestiones:

- a) Interprete el signo de los coeficientes de las variables que son significativamente distintos de cero.
 - b) Obtenga e interprete el odd para una persona cuya renta es de 1800 euros y tiene trabajo.
 - c) Obtenga e interprete el odd ratio asociado a la variable trabajo.
 - d) Obtenga el efecto marginal de la variable trabajo teniendo en cuenta que la renta media es de 2200 euros.
 - e) Sabiendo que la log-verosimilitud del modelo con sólo constante es -9.59 y del modelo completo es -6.33, analice la bondad del modelo y si es conjuntamente significativo.
4. En la base de datos de cierta aseguradora se tiene información sobre si los propietarios de un vehículo tienen asegurado el mismo a todo riesgo (\mathbf{S} , 1 en caso afirmativo y 0 en caso contrario), el género del propietario (\mathbf{G} , 1 si es varón y 0 en caso contrario), el nivel de estudios (\mathbf{E} , 1 si son universitarios y 0 en caso contrario) y la renta anual del propietario (\mathbf{R} , medida en miles de euros). Para analizar los factores que influyen en la probabilidad de tener asegurado el vehículo a todo riesgo se ajusta un modelo Probit obteniéndose los siguientes resultados:

$$\hat{z}_i = \begin{array}{cccc} -5 & + & 1.05 \cdot \mathbf{G}_i & + & 1.3 \cdot \mathbf{E}_i & + & 0.06 \cdot \mathbf{R}_i \\ (0.3) & & (0.22) & & (0.3) & & (0.02) \end{array}$$

Entre paréntesis se muestran las desviaciones típicas. Se pide contestar de forma razonada a las siguientes cuestiones:

- a) Interprete el signo de los coeficientes de las variables que son significativamente distintos de cero.
- b) ¿Cuál es la probabilidad de que tenga asegurado a todo riesgo el vehículo un hombre con estudios universitarios con una renta anual de 40000 euros?
- c) Obtener e interpretar el odd asociado al individuo del apartado anterior.
- d) Obtenga el efecto marginal asociado a la variable renta para una mujer universitaria con una renta anual de 33000 euros.