

**Modelos de respuesta discreta. Máster oficial en Estadística Aplicada**  
**Relación de ejercicios del Tema 3**

Para explicar las relaciones existentes entre la ocurrencia de problemas coronarios y algunas de las variables relacionadas se dispone de una muestra de 200 hombres para los que se han observado las siguientes siete variables explicativas:

- Edad en años
- Presión arterial, categorizada en cuatro niveles: óptima, normal, alta y descompensada:
- Colesterol en miligramos por DL
- Índice de masa corporal (IMC) categorizado en tres niveles: normal, sobrepeso y obesidad
- Peso en libras
- Problemas coronarios (1 si se tiene alguna incidencia en los 10 años previos y 0 en otro caso)

**Ejercicio 1.** Ajuste un modelo de regresión logística simple para explicar la ocurrencia de problemas coronarios a partir del IMC utilizando codificación parcial. Tras el ajuste conteste a las siguientes cuestiones:

1. Indique la estimación puntual del parámetro independiente e interprete su resultado. (Solución: -2.5552874)
2. Indique el error estándar de estimación del parámetro independiente. (Solución: 0.3670264)
3. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro independiente (Solución: -6.9621349). Concluya sobre el test a una significación  $\alpha = 0.1$
4. Indique el p-valor del test de significación (Wald) del parámetro independiente (Solución:  $3.3515421 \times 10^{-12}$ ). Concluya sobre el test a una significación  $\alpha = 0.05$ .
5. Indique la estimación puntual del parámetro asociado a la categoría Sobrepeso e interprete su resultado. (Solución: 1.1044546)
6. Indique el error estándar de estimación del parámetro asociado a la categoría Sobrepeso. (Solución: 0.4658326)
7. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Sobrepeso (Solución: 2.3709256). Concluya sobre el test a una significación  $\alpha = 0.1$
8. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Sobrepeso (Solución: 0.0177436). Concluya sobre el test a una significación  $\alpha = 0.05$ .
9. Indique la estimación puntual del parámetro asociado a la categoría Obesidad e interprete su resultado. (Solución: 1.7079896)
10. Indique el error estándar de estimación del parámetro asociado a la categoría Obesidad. (Solución: 0.7816002)
11. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Obesidad (Solución: 2.1852471). Concluya sobre el test a una significación  $\alpha = 0.1$

- 
12. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Obesidad (Solución: 0.0288707). Concluya sobre el test a una significación  $\alpha = 0.05$ .
  13. Reflexione cuánto cambiaría multiplicativamente (por cuánto se multiplicaría) la ventaja de padecer enfermedad coronaria (frente a no padecerla) si el IMC pasa de Sobrepeso a Obesidad (Solución: 1.8285714)
  14. Indique un intervalo de confianza al 95% para el parámetro independiente (Solución:  $\beta_0 \in (-3.274646, -1.8359289)$ )
  15. Indique un intervalo de confianza al 90% para el parámetro asociado a la categoría Sobrepeso (Solución:  $\tau_2 \in (0.338228, 1.8706811)$ )
  16. Indique un intervalo de confianza al 99% para el parámetro asociado a la categoría Obesidad (Solución:  $\tau_1 \in (-0.3052791, 3.7212583)$ )
  17. Indique el valor experimental y el p-valor del test de bondad de ajuste de Hosmer y Lemeshow (tomando los deciles de riesgo para su obtención). Concluya sobre el test.  
(Solución:  $4.1392464 \times 10^{-23}$  y 1)
  18. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con sobrepeso según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados  
(Solución:  $\hat{p}(\text{Sobrepeso}) = 0.1898734$ ,  $1 - \hat{p}(\text{Sobrepeso}) = 0.8101266$ ,  $\hat{p}(\text{Sobrepeso})/(1 - \hat{p}(\text{Sobrepeso})) = 0.234375$ ,  $(1 - \hat{p}(\text{Sobrepeso}))/\hat{p}(\text{Sobrepeso}) = 4.2666667$ )
  19. Indique un intervalo de confianza al 95% para la probabilidad de padecer problemas coronarios de una persona con sobrepeso según el modelo (utilice aproximación normal) (Solución:  $p(\text{Sobrepeso}) \in (0.103388, 0.2763589)$ )
  20. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con obesidad según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(\text{Obesidad}) = 0.3$ ,  $1 - \hat{p}(\text{Obesidad}) = 0.7$ ,  $\hat{p}(\text{Obesidad})/(1 - \hat{p}(\text{Obesidad})) = 0.4285714$ ,  $(1 - \hat{p}(\text{Obesidad}))/\hat{p}(\text{Obesidad}) = 2.3333333$ )
  21. Indique un intervalo de confianza al 90% para la probabilidad de padecer problemas coronarios de una persona con obesidad según el modelo (utilice aproximación normal) (Solución:  $p(\text{Obesidad}) \in (0.0616381, 0.5383619)$ )
  22. Indique cuántos residuos de Pearson estandarizados son significativos y el porcentaje de residuos de la deviance no significativos (Solución: 23, 96)
  23. Calcule CCR, Sensibilidad y Especificidad (en %) del test diagnóstico asociado a este modelo, utilizando como punto de corte 0.13. Interprete las tres cantidades  
(Solución: CC= 60.5%, Sen=69.2307692%, Espec=59.1954023%)
  24. Calcule el área bajo la curva ROC  
(Solución: 0.6517462)

**Ejercicio 2.** La relación entre la ocurrencia de problemas coronarios, la presión arterial y el IMC se muestran a continuación:

Presión	IMC	Coronarios=0	Coronarios = 1
Optima			
	Normal	38	3
	Obesidad	0	0
	Sobrepeso	12	3
Alta			
	Normal	7	1
	Obesidad	0	1
	Sobrepeso	5	3
Descompensada			
	Normal	42	4
	Obesidad	7	1
	Sobrepeso	35	9
Normal			
	Normal	16	0
	Obesidad	0	1
	Sobrepeso	12	0

Ajuste un modelo de regresión logística múltiple sin interacción para explicar la ocurrencia de problemas coronarios a partir del IMC y de la Presión.

1. Indique la estimación puntual del parámetro independiente e interprete su resultado. (Solución: -2.5197882)
2. Indique el error estándar de estimación del parámetro independiente. (Solución: 0.4908522)
3. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro independiente (Solución: -5.1334969). Concluya sobre el test a una significación  $\alpha = 0.1$
4. Indique el p-valor del test de significación (Wald) del parámetro independiente (Solución:  $2.8440767 \times 10^{-7}$ ). Concluya sobre el test a una significación  $\alpha = 0.05$ .
5. Indique la estimación puntual del parámetro asociado a la categoría Sobrepeso de IMC e interprete su resultado. (Solución: 1.1109311)
6. Indique el error estándar de estimación del parámetro asociado a la categoría Sobrepeso de IMC. (Solución: 0.4826383)
7. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Sobrepeso de IMC (Solución: 2.3017878). Concluya sobre el test a una significación  $\alpha = 0.1$
8. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Sobrepeso de IMC (Solución: 0.0213471). Concluya sobre el test a una significación  $\alpha = 0.05$ .
9. Indique la estimación puntual del parámetro asociado a la categoría Obesidad de IMC e interprete su resultado. (Solución: 1.7034312)
10. Indique el error estándar de estimación del parámetro asociado a la categoría Obesidad de IMC. (Solución: 0.8275292)
11. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Obesidad de IMC (Solución: 2.0584545). Concluya sobre el test a una significación  $\alpha = 0.1$

- 
12. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Obesidad de IMC (Solución: 0.0395465). Concluya sobre el test a una significación  $\alpha = 0.05$ .
  13. Reflexione cuánto cambiaría multiplicativamente (por cuánto se multiplicaría) la ventaja de padecer enfermedad coronaria (frente a no padecerla) si el IMC pasa de Sobrepeso a Obesidad permaneciendo la presión inalterada (Solución: 1.8085043)
  14. Indique la estimación puntual del parámetro asociado a la categoría Normal de Presion e interprete su resultado. (Solución: -1.4922458)
  15. Indique el error estándar de estimación del parámetro asociado a la categoría Normal de Presion. (Solución: 1.1209333)
  16. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Normal de Presion (Solución: -1.331253). Concluya sobre el test a una significación  $\alpha = 0.1$
  17. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Normal de Presion (Solución: 0.1831058). Concluya sobre el test a una significación  $\alpha = 0.05$ .
  18. Indique la estimación puntual del parámetro asociado a la categoría Alta de Presion e interprete su resultado. (Solución: 0.9471427)
  19. Indique el error estándar de estimación del parámetro asociado a la categoría Alta de Presion. (Solución: 0.7122531)
  20. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Alta de Presion (Solución: 1.3297839). Concluya sobre el test a una significación  $\alpha = 0.1$
  21. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Alta de Presion (Solución: 0.1835895). Concluya sobre el test a una significación  $\alpha = 0.05$ .
  22. Indique la estimación puntual del parámetro asociado a la categoría Descompensada de Presion e interprete su resultado. (Solución: -0.0447068)
  23. Indique el error estándar de estimación del parámetro asociado a la categoría Descompensada de Presion. (Solución: 0.5517025)
  24. Indique el valor experimental (valor estimado del estadístico) del test de significación (Wald) del parámetro asociado a la categoría Descompensada de Presion (Solución: -0.0810342). Concluya sobre el test a una significación  $\alpha = 0.1$
  25. Indique el p-valor del test de significación (Wald) del parámetro asociado a la categoría Descompensada de Presion (Solución: 0.9354147). Concluya sobre el test a una significación  $\alpha = 0.05$ .
  26. Reflexione cuánto cambiaría multiplicativamente (por cuánto se multiplicaría) la ventaja de padecer enfermedad coronaria (frente a no padecerla) si la presión pasa de Normal a Alta permaneciendo el IMC inalterado (Solución: 11.4660264)
  27. Reflexione cuánto cambiaría multiplicativamente (por cuánto se multiplicaría) la ventaja de padecer enfermedad coronaria (frente a no padecerla) si la presión pasa de Normal a Descompensada permaneciendo el IMC inalterado (Solución: 4.2526358)

- 
28. Reflexione cuánto cambiaría multiplicativamente (por cuánto se multiplicaría) la ventaja de padecer enfermedad coronaria (frente a no padecerla) si la presión pasa de Alta a Descompensada permaneciendo el IMC inalterado (Solución: 0.3708901)
29. Indique un intervalo de confianza al 95% para el parámetro independiente (Solución:  $\beta_0 \in (-3.4818408, -1.5577356)$ )
30. Indique un intervalo de confianza al 90% para el parámetro asociado a la categoría Sobrepeso de IMC (Solución:  $\tau_2^{IMC} \in (0.3170616, 1.9048005)$ )
31. Indique un intervalo de confianza al 99% para el parámetro asociado a la categoría Obesidad de IMC (Solución:  $\tau_1^{IMC} \in (-0.4281428, 3.8350052)$ )
32. Indique un intervalo de confianza al 90% para el parámetro asociado a la categoría Normal de Presión (Solución:  $\tau_3^{Presion} \in (-3.3360169, 0.3515254)$ )
33. Indique un intervalo de confianza al 95% para el parámetro asociado a la categoría Alta de Presión (Solución:  $\tau_1^{Presion} \in (-0.4488477, 2.343133)$ )
34. Indique un intervalo de confianza al 99% para el parámetro asociado a la categoría Descompensada de Presión (Solución:  $\tau_2^{presion} \in (-1.4657983, 1.3763847)$ )
35. Indique un intervalo de confianza al 90% para el cociente de ventajas de respuesta 1 (frente a respuesta 0) cuando la presión pasa de Ótima a Normal (Solución:  $\theta \in (0.0355784, 1.4212339)$ )
36. Indique un intervalo de confianza al 95% para el cociente de ventajas de respuesta 1 (frente a respuesta 0) cuando la presión pasa de Ótima a Alta (Solución:  $\theta \in (0.6383633, 10.4138122)$ )
37. Indique un intervalo de confianza al 99% para el cociente de ventajas de respuesta 1 (frente a respuesta 0) cuando la presión pasa de Ótima a Descompensada (Solución:  $\theta \in (0.2308936, 3.9605572)$ )
38. Indique un intervalo de confianza al 95% para el cociente de ventajas de respuesta 1 (frente a respuesta 0) cuando el IMC pasa de Normal a Sobrepeso (Solución:  $\theta \in (1.1793663, 7.8215664)$ )
39. Indique un intervalo de confianza al 90% para el cociente de ventajas de respuesta 1 (frente a respuesta 0) cuando el IMC pasa de Normal a Obesidad (Solución:  $\theta \in (1.4081359, 21.4257962)$ )
40. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presión Ótima e IMC Normal según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados  
(Solución:  $\hat{p}(Optima, Normal) = 0.0744825$ ,  $1 - \hat{p}(Optima, Normal) = 0.9255175$ ,  $\hat{p}(Optima, Normal)/(1 - \hat{p}(Optima, Normal)) = 0.0804767$ ,  $(1 - \hat{p}(Optima, Normal))/\hat{p}(Optima, Normal) = 12.4259645$ )
41. Indique un intervalo de confianza al 95% para la probabilidad de padecer problemas coronarios de una persona con Presión Ótima e IMC Normal según el modelo (utilice aproximación normal) (Solución:  $p(Optima, Normal) \in (0.0081635, 0.1408015)$ )
42. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presión Normal e IMC de Sobrepeso según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(Normal, Sobrepeso) = 0.0520991$ ,  $1 - \hat{p}(Normal, Sobrepeso) = 0.9479009$ ,  $\hat{p}(Normal, Sobrepeso)/(1 - \hat{p}(Normal, Sobrepeso)) = 0.0549626$ ,  $(1 - \hat{p}(Normal, Sobrepeso))/\hat{p}(Normal, Sobrepeso) = 18.1942006$ )

- 
43. Indique un intervalo de confianza al 90% para la probabilidad de padecer problemas coronarios de una persona con Presion Normal e IMC de Sobre peso según el modelo (utilice aproximación normal) (Solución:  $p(\text{Normal}, \text{Sobre peso}) \in (-0.0321693, 0.1363674)$ )
  44. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presion Alta e IMC de Obesidad según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(\text{Alta}, \text{Obesidad}) = 0.5326499$ ,  $1 - \hat{p}(\text{Alta}, \text{Obesidad}) = 0.4673501$ ,  $\hat{p}(\text{Alta}, \text{Obesidad}) / (1 - \hat{p}(\text{Alta}, \text{Obesidad})) = 1.1397235$ ,  $(1 - \hat{p}(\text{Alta}, \text{Obesidad})) / \hat{p}(\text{Alta}, \text{Obesidad}) = 0.8774058$ )
  45. Indique un intervalo de confianza al 99% para la probabilidad de padecer problemas coronarios de una persona con Presion Alta e IMC de Obesidad según el modelo (utilice aproximación normal) (Solución:  $p(\text{Alta}, \text{Obesidad}) \in (-0.0401521, 1.1054519)$ )
  46. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presion Descompensada e IMC Normal según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(\text{Descompensada}, \text{Normal}) = 0.0714587$ ,  $1 - \hat{p}(\text{Descompensada}, \text{Normal}) = 0.9285413$ ,  $\hat{p}(\text{Descompensada}, \text{Normal}) / (1 - \hat{p}(\text{Descompensada}, \text{Normal})) = 0.076958$ ,  $(1 - \hat{p}(\text{Descompensada}, \text{Normal})) / \hat{p}(\text{Descompensada}, \text{Normal}) = 12.9940944$ )
  47. Indique un intervalo de confianza al 90% para la probabilidad de padecer problemas coronarios de una persona con Presion Descompensada e IMC Normal según el modelo (utilice aproximación normal) (Solución:  $p(\text{Alta}, \text{Obesidad}) \in (0.0215103, 0.1214071)$ )
  48. Indique cuántos residuos de Pearson estandarizados son significativos y el porcentaje de residuos de la deviance no significativos (Solución: 62, 69)

**Ejercicio 3.** Ajuste un modelo de regresión logística múltiple con interacción para explicar la ocurrencia de problemas coronarios a partir de la presión y el nivel de colesterol. Tras el ajuste conteste a las siguientes cuestiones:

1. Indique la estimación puntual del parámetro independiente e interprete su resultado. (Solución: -4.3730086)
2. Indique la estimación puntual del parámetro asociado al nivel de colesterol e interprete su resultado. (Solución: 0.0080993)
3. Indique la estimación puntual del parámetro asociado a la presión normal e interprete su resultado. (Solución: -10.0643314)
4. Indique la estimación puntual del parámetro asociado a la presión alta e interprete su resultado. (Solución: 7.7673904)
5. Indique la estimación puntual del parámetro asociado a la presión descompensada e interprete su resultado. (Solución: -0.9395891)
6. Indique la estimación puntual del parámetro asociado a la interacción entre el nivel de colesterol y la presión normal e interprete su resultado. (Solución: 0.0270485)
7. Indique la estimación puntual del parámetro asociado a la interacción entre el nivel de colesterol y la presión alta e interprete su resultado. (Solución: -0.0226108)

- 
8. Indique la estimación puntual del parámetro asociado a la interacción entre el nivel de colesterol y la presión descompensada e interprete su resultado. (Solución: 0.0031133)
  9. Indique el valor experimental y el p-valor del test de bondad de ajuste de Hosmer y Lemeshow (tomando los deciles de riesgo para su obtención). Concluya sobre el test.  
(Solución: 5.2718817 y 0.728161)
  10. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presion Normal y nivel de colesterol de 300 según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados  
(Solución:  $\hat{p}(Normal, 300) = 0.0199769$ ,  $1 - \hat{p}(Normal, 300) = 0.9800231$ ,  $\hat{p}(Normal, 300)/(1 - \hat{p}(Normal, 300)) = 0.0203841$ ,  $(1 - \hat{p}(Normal, 300))/\hat{p}(Normal, 300) = 49.0577843$ )
  11. Indique un intervalo de confianza al 95% para la probabilidad de padecer problemas coronarios de una persona de Presion Normal y nivel de colesterol de 300 según el modelo (utilice aproximación normal)  
(Solución:  $p(Normal, 300) \in (-0.0530756, 0.0930294)$ )
  12. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presion Alta y nivel de colesterol de 250 según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(Alta, 250) = 0.4418917$ ,  $1 - \hat{p}(Alta, 250) = 0.5581083$ ,  $\hat{p}(Alta, 250)/(1 - \hat{p}(Alta, 250)) = 0.7917669$ ,  $(1 - \hat{p}(Alta, 250))/\hat{p}(Alta, 250) = 1.2629979$ )
  13. Indique un intervalo de confianza al 90% para la probabilidad de padecer problemas coronarios de una persona de Presion Alta y nivel de colesterol de 250 según el modelo (utilice aproximación normal)  
(Solución:  $p(Alta, 250) \in (0.1525838, 0.7311996)$ )
  14. Indique cuál sería la probabilidad de padecer problemas coronarios de una persona con Presion Optima y nivel de colesterol de 275 según el modelo. Indique también cuál sería la probabilidad de no padecer problemas coronarios, la ventaja de padecer problemas coronarios y la ventaja de no padecer problemas coronarios. Interprete estos dos últimos resultados (Solución:  $\hat{p}(Optima, 275) = 0.1047348$ ,  $1 - \hat{p}(Optima, 275) = 0.8952652$ ,  $\hat{p}(Optima, 275)/(1 - \hat{p}(Optima, 275)) = 0.1169875$ ,  $(1 - \hat{p}(Optima, 275))/\hat{p}(Optima, 275) = 8.5479223$ )
  15. Indique un intervalo de confianza al 99% para la probabilidad de padecer problemas coronarios de una persona de Presion Optima y nivel de colesterol de 275 según el modelo (utilice aproximación normal)  
(Solución:  $p(Optima, 275) \in (-0.0032201, 0.2126898)$ )
  16. Indique cuántos residuos de Pearson estandarizados son no significativos y el porcentaje de residuos de la deviance no significativos (Solución: 185, 94)
  17. Calcule CCR, Sensibilidad y Especificidad (en %) del test diagnóstico asociado a este modelo, utilizando como punto de corte 0.13. Interprete las tres cantidades  
(Solución: CC= 70%, Sen=61.5384615%, Espec=71.2643678%)
  18. Calcule el área bajo la curva ROC  
(Solución: 0.7358532)

**Ejercicio 4.** Repita el ejercicio anterior sin incluir la interacción e interprete los parámetros. Observe las diferencias.