Tarea 1 (1)

1 La siguiente tabla muestra los resultados parciales de dos encuestas que forman parte de un estudio para evaluar el desempeño del Primer Ministro del Canadá. Se tomó una muestra aleatoria de 1600 ciudadanos canadienses mayores de edad y en los renglones se observa que 944 ciudadanos aprobaban el desempeño del funcionario, mientras que las columnas muestran que, seis meses después de la primera encuesta, sólo 880 aprueban su desempeño.

Primera encuesta	Segunda encuesta $Y=1$, Aprueba	$Y=0$, Desaprueba	Total
$x=1$, Aprueba	794	150	944
$x=0$, Desaprueba	86	570	656
Total	880	720	1600

Tarea 1 (2)

1 ··· Cont.

a Considere el modelo de regresión logística

$$\log \frac{P(Y_i = 1|x_i)}{1 - P(Y_i = 1|x_i)} = \beta_0 + \beta_1 x_i$$

Escriba la logverosimilitud correspondiente. Muestre explícitamente (i.e. maximizando la logverosimilitud), que el estimador máximo verosimilitud para β_1 es el logaritmo de la tasa de momios de la tabla dada (En general, en regresión logística los estimadores de máxima verosimilitud no tienen una forma explícita, sin embargo, en el presente caso si).

b Sea p_1 la proporción de ciudadanos que aprueban el desempeño del ministro al tiempo inicial y sea p_2 la proporción correspondiente seis meses después. Considere la hipótesis $H_0 : p_1 = p_2$, ¿Cómo puede hacerse esta prueba?.

Tarea 1 (3)

2 Suponga $(x_1, y_1), \ldots, (x_n, y_n)$ observaciones independientes de variables aleatorias definidas como sigue:

$$Y_i \sim \text{Bernoulli}(p), \quad i = 1, \cdots, n$$
$$X_i \mid \{Y_i = 1\} \sim N(\mu_1, \sigma^2)$$
$$X_i \mid \{Y_i = 0\} \sim N(\mu_0, \sigma^2)$$

Usando el Teorema de Bayes, muestre que $\text{P}(Y_i = 1|X_i)$ satisface el modelo de regresión logística, esto es

$$\text{logit}(\ \text{P}(Y_i = 1|X_i)\ ) = \alpha + \beta X_i$$

con $\beta = (\mu_1 - \mu_0)/\sigma^2$.

Entregar: Miércoles 19 de agosto.

Tarea 2

1 La siguiente tabla muestra conteos de células T_4 por mm^3 en muestras de sangre de 20 pacientes (en remisión) con enfermedad de Hodgkin, así como conteos en 20 pacientes en remisión de otras enfermedades. Una cuestión de interés es si existen diferencias en las distribuciones de conteos en ambos grupos.

H	396	568	1212	171	554	1104	257	435	295	397
No-H	375	375	752	208	151	116	736	192	315	1252
H	288	1004	431	795	1621	1378	902	958	1283	2415
No-H	675	700	440	771	688	426	410	979	377	503

a. Haga una comparación gráfica exploratoria de estos datos.
b. Ajuste un modelo de Poisson apropiado.
c. Usando la normalidad asintótica de los estimadores de máxima verosimilitud, dé un intervalo del 90% de confianza para la diferencia en medias. ¿Hay evidencia de diferencias en los dos grupos en cuanto a las medias de los conteos?.

(Problema adaptado de Wakefield (2013), p.300)