## Instituto Tecnológico de Costa Rica Escuela de Computación

Programa de Especialidad en Ciencias de los Datos Curso: Estadistica

Profesor: Ph. D. Saúl Calderón Ramírez

QUIZ 0

Entrega: Domingo 17 de setiembre del 2023

Debe subir un *pdf* con la respuesta.

A través del TEC digital

Valor: 100 pts.
Puntos Obtenidos:

Nota: \_\_\_\_\_

Nombre del (la) estudiante:	

Carné:

1. Suponga que, usted desea ajustar una función de densidad de distribución a un conjunto de datos que se refiere a la probabilidad de que un paciente sea atendido o no en menos de 5 minutos. Su objetivo es predicir la probabilidad con la que el paciente será atendido o no. Para ello, su equipo de científicos de datos observa que el histograma de frecuencias puede ajustarse muy bien a un modelo paramétrico binomial.

$$p(t|\rho) = \rho^t (1-\rho)^{(t-1)}$$

Este modelo también se conoce como un modelo de Bernoulli. Como se observa, el modelo cuenta con un solo parámetro a estimar, en este caso  $\rho$ .

(a) (50 pts) Suponga que se utilizará el algoritmo del descenso del gradiente, para obtener el  $\rho$  óptimo, maximizando la verosimlitud de tal modelo, para la serie de valores (conjunto de datos)  $\overrightarrow{t}$ :

$$p\left(\overrightarrow{t}|\rho\right) = \ln\left(\prod_{i=1}^{n} \rho^{t} \left(1-\rho\right)^{(t-1)}\right)$$

En este caso, el algoritmo del descenso del gradiente implementa la actualización del parámetro  $\lambda (\tau + 1)$  para una iteración  $\tau + 1$ , como sigue:

$$\rho(\tau + 1) = \rho(\tau) + \Delta\rho(\tau).$$

Calcule el cambio en el parámetro  $\rho$ ,  $\Delta\rho$  (au). Muestre todos los pasos intermedios.

(b) **(50 pts)** Un miembro de su equipo sospecha que puede usarse una forma más eficiente para optimizar el parámetro  $\rho$ : derivando respecto a tal parámetro, igualando la ecuación a 0 y despejandolo. Demuestre si es posible hacerlo, y en caso de ser posible, muestre todos los pasos para llegar a la expresión del  $\rho$  óptimo. Finalmente responda, porqué el esquema sugerido por su colega es más efectivo que el implementado en el apartado anterior?