

# Probabilidad y Estadística para Inteligencia Artificial

## Examen

16 de junio de 2022

Docentes: Magdalena Bouza, Pedro Cosatto

Por favor seguir los siguientes pasos para entregar las soluciones de los ejercicios:

- Enviar las soluciones por e-mail a mbouza@fi.uba.ar y a pcosatto@fi.uba.ar.
- En el Asunto del e-mail poner “Resolución de examen PEIA 2022 2b - *Nombre Apellido*”.
- Entregar todos los desarrollos en formato pdf.
- Entregar los resultados antes de las 17 hs del martes 14/06.
- El examen se aprueba con al menos 2 ejercicios bien, al menos uno de ellos de la parte de estadística.

1. La producción de gasolina mensual (en m3) en Neuquén sigue una distribución normal de media 95529 y desvío estándar 30127, mientras que la de Santa Cruz sigue una distribución también normal pero de media 8268 y desvío estándar 2481.
  - (a) Hallar la probabilidad de que la producción total entre Neuquén y Santa Cruz de un mes supere los 142925 m3.
  - (b) Hallar la probabilidad de que en un mes la producción de Neuquén sea 10 veces más grande que la de Santa Cruz.
  - (c) Si cada barril de gasolina tiene una capacidad de 159 l (0.159 m3), hallar la cantidad mínima de barriles necesarios para almacenar la producción de gasolina de Santa Cruz de un mes con 95 % de probabilidad.
  - (d) Simular el experimento y verificar los resultados obtenidos en los puntos a), b) y c).
2. Una técnica muy utilizada en el preprocesamiento de datos se conoce como *binning*. Este método consiste en dividir el soporte de la variable en distintos intervalos y reemplazar el valor original de la variable por un valor representativo del intervalo en el que cae.

Consideremos  $X$  una variable con distribución exponencial de parámetro  $\lambda = 5$ . Se desea aplicarle la técnica de *binning*, partiendo el soporte en 5 intervalos de igual probabilidad, representando los valores de cada intervalo por su punto medio.

Sea  $Y$  aleatoria resultante luego de aplicar el **binning**

  - a) hallar la función de probabilidad de  $Y$ ,
  - b) hallar la esperanza de  $Y$ ,
  - c) hallar la probabilidad de que  $Y$  sea mayor a 1.
3. En un juego de ruleta que cuenta con los números del 0 al 36, un jugador siempre apuesta a tercera docena, es decir que sólo lo benefician los números del 25 al 36 inclusive. El casino sospecha que un crupier intenta favorecer al jugador, y está dispuesto a despedirlo si encuentra evidencia suficiente de que lo favorece. Luego de 100 bolas tiradas por el crupier, salió la tercer docena 40 veces.
  - a) Hallar un test de hipótesis de nivel asintótico 0.05 adecuado a este problema y basándose en él decidir si el casino debe despedir al crupier.
  - b) Calcular el p-valor aproximado.
4. Distribución de los tamaños de los archivos del tráfico de Internet(en 1000KB) que utiliza el protocolo TCP (muchos archivos pequeños, pocos archivos grandes sigue una distribución de Pareto dada por:

$$f_X(x) = \theta(0.5)^\theta x^{-(\theta+1)} \mathbf{1}\{x > 0.5\}$$

- Hallar el EVM para  $\theta$
- Para siguiente muestra de tamaño 15, hallar el valor estimado de  $\theta$ :

$$\underline{x} = [10.8941, 1.3631, 24.1238, 85.8298, 1.2616, 681.7245, 20.9089, \\ 1673.6941, 3.7612, 37.5120, 243.4445, 10.6362, 773.1123, 12434.6597, 25.5089]$$