



Maestría en Ciencias de los Datos y Analítica
Examen de admisión en Matemáticas, Estadística y
Computación

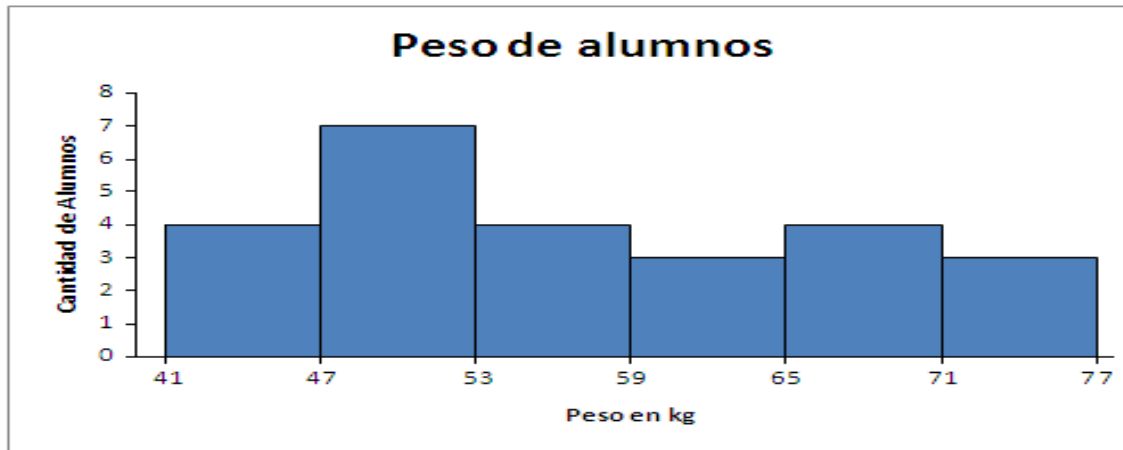
NOTA: _____

NOMBRE: _____ CÓDIGO: _____

GRUPO: _____ PROFESOR: Henry Laniado; Santiago Ortiz; Juan G. Lalinde FECHA: _____

1. [13 %] Juan debe trabajar al menos 15 horas a la semana para complementar sus ingresos. Tiene la oportunidad de trabajar en dos tiendas de menudeo. En la tienda A puede trabajar entre 6 y 10 horas a la semana, y en la tienda B le permiten trabajar entre 5 y 12 horas. Ambas tiendas pagan el mismo salario por hora. Para decidir cuántas horas trabajar en cada tienda, Juan desea tomar su decisión según el estrés del trabajo. Basado en entrevistas con otros empleados, Juan estima que, en una escala del 1 al 10, los factores de estrés son 6 y 8 en las tiendas A y B , respectivamente. El estrés aumenta proporcionalmente a las horas que trabaja en cada tienda.
- a) [5 %] Formule la función objetivo y las restricciones.
 - b) [5 %] ¿Cuántas horas debería trabajar Juan en cada tienda?
 - c) [3 %]Cuál es el estrés acumulado al final de la jornada?

2. [25 %] El peso de un conjunto de alumnos de un curso de analítica de datos se resume en el siguiente histograma



- a) [5 %] Estime el peso superado por el 90 % de los alumnos
- b) [10 %] Si el peso mediano es mayor al peso medio, hay indicios que hay estudiantes que tienen bajo peso con relación a los demás. Qué se puede concluir de este grupo?.
- c) [10 %] Se promociona una campaña de hábitos y complementos alimenticios para los estudiantes cuyo peso esté a mas de dos desviaciones típicas a la izquierda del peso medio del grupo. El programa cuesta por estudiante 100 mil pesos. Estime aproximadamente cuánto costaría la campaña en este grupo.

3. [50 %] El fichero coches.xls contiene características especiales de un conjunto de autos nuevos. Suponga que T_{100} es la variable que representa el tiempo en segundos hasta que el coche alcanza los $100km/h$ (Columna M del fichero)
- a) [5 %] Calcule un intervalo al 95 % confianza para el tiempo medio en segundos hasta alcanzar los $100km/h$
 - b) [10 %] Formule un modelo de regresión lineal múltiple que explique la variable T_{100} en términos de las otras. (que sólo tenga variables significativas)
 - c) [5 %] La variable binaria *Ensamble* es significativa para T_{100} ? Justifique su respuesta. Identifique una variable significativa e interprete su coeficiente.
 - d) [5 %] Considera que hay un buen ajuste?. Con qué valor lo relaciona? Interprete ese valor.
 - e) [10 %] Calcule un intervalo de confianza para la proporción de coches que se ensamblan localmente. Pruebe la hipótesis de que la producción local supera la extranjera.
 - f) [5 %] Construya una sucesión, que sea convergente a la potencia media de los coches.
 - g) [10 %] Qué coche, distinto del primero, es el mas parecido al primer coche, según la métrica inducida por la norma infinito?

4. [4 %] A continuación se presenta el pseudocódigo¹ de dos formas diferentes para buscar un elemento en un arreglo que está ordenado. Una primera forma de buscar es con el algoritmo 1.

Algoritmo 1 Buscar un elemento en un arreglo ordenado

```
1: procedure  $B(x, n, a)$ 
                                     ▷  $a$  es el arreglo ordenado.
                                     ▷  $a[i]$  se refiere al  $i$ -ésimo elemento del arreglo
                                     ▷  $x$  es el elemento a buscar
                                     ▷  $n$  es el tamaño del arreglo
                                     ▷ Retorna  $i < n$  si  $x$  está en la posición  $i$  o  $-1$  si no está.
2:    $i \leftarrow 0$ 
3:   while  $i < n$  do
4:     if  $a[i] == x$  then
5:       return  $i$ 
6:     else  $i \leftarrow i + 1$ 
7:     end if
8:   end while
9:   return  $-1$ 
10: end procedure
```

Otra alternativa es el algoritmo 2.

- a) [4 %] Compare el algoritmo 1 con el algoritmo 2 ¿Cuál de ellos es más eficiente desde el punto de vista computacional? Justifique su respuesta.

¹En el pseudocódigo:

- El símbolo ▷ significa que esa línea es un comentario, es decir, una explicación que no forma parte del algoritmo.
- El símbolo \leftarrow representa la operación de asignación. El símbolo $=$ se reserva para comparaciones.

Algoritmo 2 Buscar un elemento en un arreglo ordenado

```
1: procedure  $B(x, n, a)$   
  
2:    $l \leftarrow 0$   
3:    $r \leftarrow n - 1$   
4:   while  $l \leq r$  do  
5:      $m \leftarrow \frac{l+r}{2}$  ▷ Al ser enteros, es el cociente  
6:     if  $a[m] < x$  then  
7:        $r \leftarrow m - 1$   
8:     else  
9:       if  $a[m] > x$  then  
10:         $l \leftarrow m + 1$   
11:      else  
12:        return  $m$   
13:      end if  
14:    end if  
15:  end while  
16:  return -1  
17: end procedure
```

5. [8 %] Sean dos $\mathcal{F}(n) = 10^5 n$ y $\mathcal{G}(n) = 10^{-1} n^2$ las funciones que relacionan el tamaño del problema (n) con el tiempo de ejecución del mismo ($\mathcal{F}(n)$ y $\mathcal{G}(n)$).

- [4 %] ¿Para qué valores de n es más eficiente el algoritmo representado por $\mathcal{F}(n)$ que el representado por $\mathcal{G}(n)$?
- [4 %] ¿Cuál algoritmo escala mejor?