Matemáticas Aplicadas a las CCSS for Dummies

Curso 2015-2016

Jesús García de Jalón de la Fuente

1. Programación lineal

1.1. Inecuaciones

Una inecuación lineal con dos incógnitas tiene la forma:

$$Ax + By + C < 0$$

En lugar del signo $<$ pueden aparecer $>$, $\leq$ o $\geq$.

Como es sabido, el conjunto de puntos que cumplen la ecuación $Ax + By + C = 0$ forman una línea recta. Esta recta divide el plano en dos semiplanos; en uno de ellos se cumple que $Ax + By + C < 0$ y en otro se cumple que $Ax + By + C > 0$. Por consiguiente, la solución de la inecuación está formada por los puntos de uno de los dos semiplanos (y la recta, en el caso de una inecuación con los signos $\leq$ o $\geq$).

Para determinar cuál de los dos semiplanos es la solución, basta probar si un punto particular de uno se los semiplanos (por ejemplo el origen de coordenadas en el caso de que no esté contenido en la recta es solución.

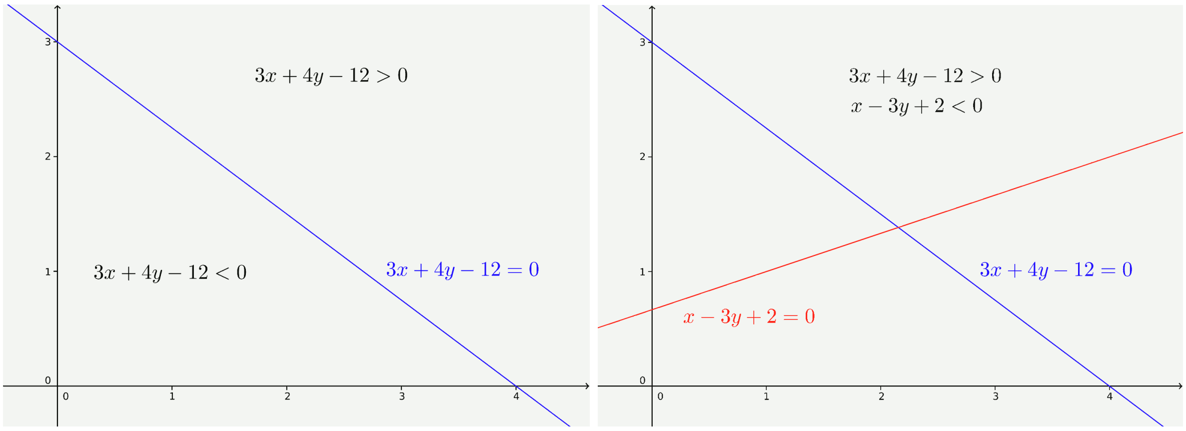


Figura 1: Solución de una inecuación y de un sistema de inecuaciones

La solución de un sistema de dos inecuaciones lineales será la intersección de los dos semiplanos solución, es decir, será una de las cuatro regiones en que las rectas que se cortan dividen el plano.

Las soluciones de un sistema formado por un número cualquiera de inecuaciones:

$$\begin{array}{c} A_1x + B_1y + C_1 < 0 \\ A_2x + B_2y + C_2 < 0 \\ \dots \\ A_nx + B_ny + C_n < 0 \end{array}$$

forman una región del plano limitada por una línea poligonal. Esta región puede estar acotada o no (un conjunto de puntos es acotado si existe un cuadrado que los contiene a todos) pero siempre es un conjunto convexo. Un conjunto de puntos es convexo, si dados dos puntos del conjunto P y Q, también pertenecen al conjunto todos los puntos del segmento PQ.

1.2. Programación lineal

Los problemas de programación lineal consisten en encontrar los valores de x e y que hacen que sea máxima o mínima una **función objetivo**:

$$F(x, y) = Mx + Ny$$

sujeta a un conjunto de **restricciones**:

$$A_1x + B_1y + C_1 \leq 0$$

$$A_2x + B_2y + C_2 \leq 0$$

$$\ldots\ldots\ldots$$

$$A_nx + B_ny + C_n \leq 0$$

El conjunto de restricciones, definen una región del plano llamada **región factible**. Como hemos visto, la región factible es **convexa** y puede ser o no ser acotada.

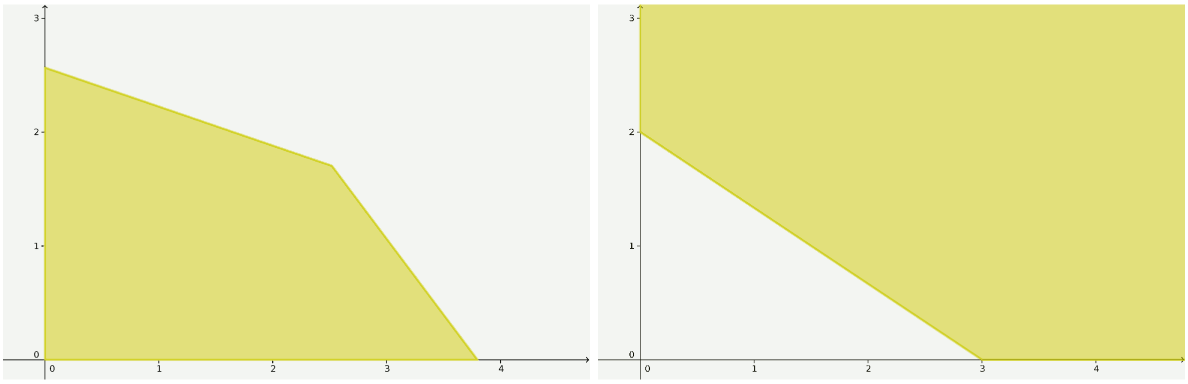


Figura 2: Región factible acotada y no acotada

Consideremos una región factible acotada como la que está representada en la figura (la superficie interior al polígono dibujado con trazo grueso). Los puntos para los que la función objetivo toma determinados valores C_1, C_2. C_3 y C_4, son rectas tales como las representadas en la segunda figura. Cuanto mayor es el valor de la función objetivo tanto más alejada está la recta del origen de coordenadas.

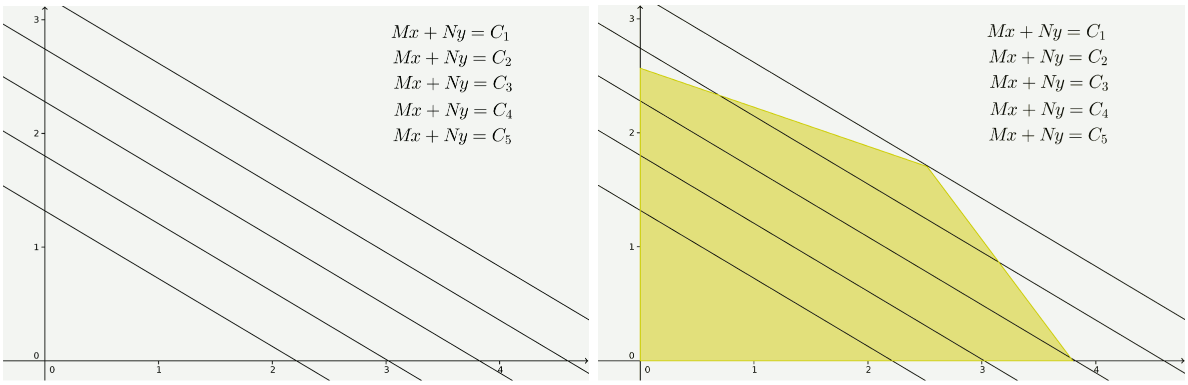


Figura 3: Solución de un problema de programación lineal

Queda claro que el valor máximo de la función objetivo se dará en un vértice de la región factible y entre los vértices deberemos buscar por tanto, la solución del problema. Es posible que el mismo valor máximo se dé en dos vértices. Entonces, como el polígono es convexo, el mismo valor se dará en el segmento que une los dos vértices. En este caso la solución no es un vértice sino un lado del polígono

1.3. Ejercicios resueltos

1. Calcular la solución del siguiente sistema de inecuaciones y las coordenadas de sus vértices:

$$\begin{cases} x + 3y \leq 15 \\ 2x + y \leq 10 \\ x - y \leq 2 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

Solución:
La región factible se puede ver en la figura 4.

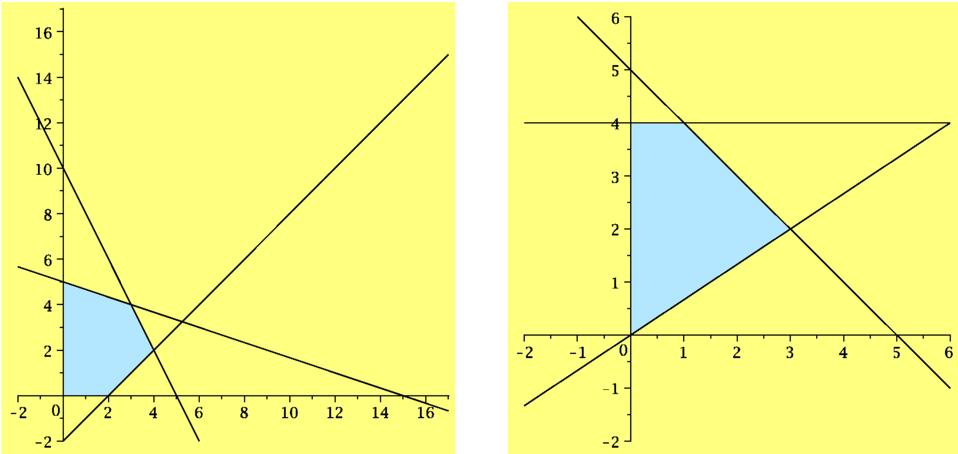


Figura 4: Ejercicios 1 y 2

Los vértices de la región son $O(0,0)$, $A(0,5)$, $B(3,4)$, $C(4,2)$ y $D(2,0)$.

2. Calcular el máximo y el mínimo de la función $F(x, y) = 3x + 8y$ con las siguientes restricciones:

$$\begin{cases} 2x - 3y \leq 0 \\ x + y \leq 5 \\ y \leq 4 \\ x \geq 0 \end{cases}$$

Solución:
Podemos ver la región factible en la figura 4. Los vértices de la región son los puntos $O(0,0)$, $A(0,4)$, $B(1,4)$ y $C(3,2)$. El mínimo de la función se da en el punto $O(0,0)$ y el máximo en $B(1,4)$.

3. Se desea invertir una cantidad de dinero menor o igual que 125000 euros, distribuidos entre acciones del tipo A y del tipo B. Las acciones del tipo A garantizan una ganancia del 10 % anual, siendo obligatorio invertir en ellas un mínimo de 30000 euros y un máximo de 81000 euros. Las acciones del tipo B garantizan una ganancia del 5 % anual, siendo obligatorio invertir en ellas un mínimo de 25000 euros. La cantidad invertida en acciones del tipo B no puede superar el triple de la cantidad invertida en acciones del tipo A. ¿Cuál debe ser la distribución de la inversión para maximizar la ganancia anual? Determínese dicha ganancia máxima.
Solución:
Llamemos:
x: cantidad invertida en acciones del tipo A
y: cantidad invertida en acciones del tipo B
La función objetivo es $F(x, y) = 0{,}10x + 0{,}05y$.
Las restricciones son las siguientes:

$$\begin{cases} x + y \leq 125 \\ x \geq 30 \\ x \leq 81 \\ y \geq 25 \\ y \leq 3x \\ x \geq 0\,,\ y \geq 0 \end{cases}$$

La región factible está representada en la figura 5.

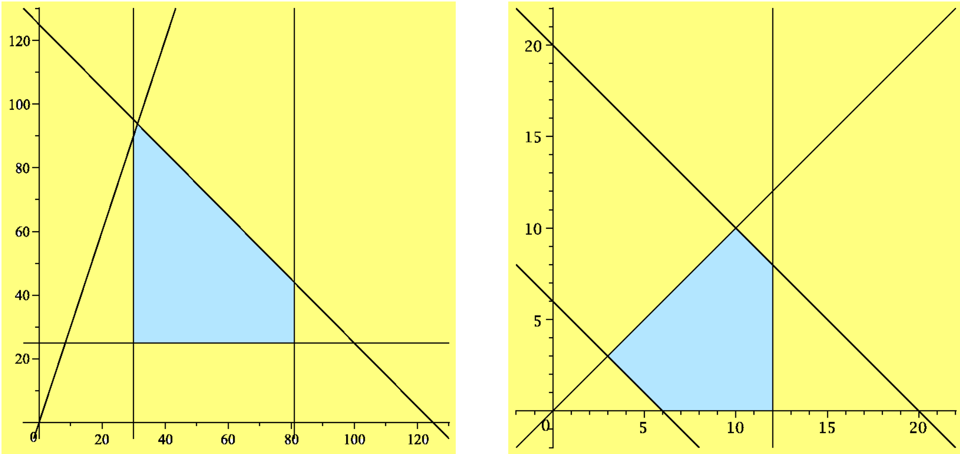


Figura 5: Ejercicios 3 y 4

El máximo puede darse en los vértices $A(31250, 93750)$ o en $B(81000, 44000)$. Calculamos la función objetivo en ambos:

$$F(31250, 93750) = 0{,}10 \cdot 31250 + 0{,}05 \cdot 93750 = 7812{,}50$$
$$F(81000, 44000) = 0{,}10 \cdot 81000 + 0{,}05 \cdot 44000 = 10300$$

El beneficio máximo de 10300 euros se da invirtiendo 81000 euros en acciones del tipo A y 44000 euros en acciones del tipo B.

4. Una compañía naviera dispone de dos barcos A y B para realizar un crucero. El barco A debe hacer tantos viajes o más que el barco B, pero no puede sobrepasar los 12 viajes. Entre los dos barcos deben hacer no menos de 6 viajes y no más de 20. La naviera obtiene un beneficio de 18000 euros por cada viaje del barco A y 12000 euros por cada viaje del B. Se desea que las ganancias sean máximas. Hallar el número de viajes que debe efectuar cada barco para obtener el máximo beneficio. Calcular dicho beneficio máximo.
Solución:
Llamamos:
x: número de viajes del barco A
y: número de viajes del barco B
La función objetivo es $F(x, y) = 18000x + 12000y$.
Las restricciones son:

$$\begin{cases} x \geq y \\ x \leq 12 \\ x + y \geq 6 \\ x + y \leq 20 \\ x \geq 0\,,\ \ y \geq 0 \end{cases}$$

Los vértices que pueden dar el beneficio máximo son $A(10, 10)$ o $B(12, 8)$. Claramente, el máximo beneficio se da en $B(12, 8)$ y vale:

$$F(12, 8) = 18000 \cdot 12 + 12000 \cdot 8 = 312000$$

1.4. Ejercicios propuestos

1. Calcular el mínimo de $F(x, y) = 9x + y$ con las restricciones:

$$\begin{cases} x + y \leq 5 \\ 3x + y \geq 2 \\ 3x - 2y \leq 5 \\ x \geq 0 \end{cases}$$

2. Calcular el mínimo de $F(x, y) = 4x + 5y$ con las restricciones:

$$\begin{cases} \dfrac{x}{10} + \dfrac{y}{8} \leq 1 \\ \dfrac{x}{5} + \dfrac{y}{8} \geq 1 \\ \dfrac{x}{10} + \dfrac{y}{4} \geq 1 \end{cases}$$

3. Una empresa cuenta con 3 empleados que trabajan durante 40 horas semanales para elaborar dos tipos de guitarras eléctricas, G_1 y G_2. Cada unidad de G_1 requiere tres horas de trabajo y cada unidad de G_2, cuatro. Independientemente del tipo que sea, cada guitarra proporciona un beneficio de 75 euros.

   Un estudio de mercado señala que no se deben producir en total más de 32 guitarras semanales. Determina la producción para que los beneficios sean máximos.

4. Se desea fabricar comida para gatos de dos clases diferentes: gama alta y gama media. La comida está formada por una mezcla de carne, cereales y grasa animal en diferentes proporciones según la gama. La mezcla de gama alta incluye 3 kg de carne, 2 kg de cereales y 1 kg de grasa animal por paquete, y produce un beneficio de 20 euros, mientras que la mezcla de gama media incluye 1 kg de carne, 2 kg de cereales y 2 g de grasa animal por paquete y produce un beneficio de 30 euros.

   Se cuenta con un total de 105 kg de carne, 110 de cereales y 85 de grasa animal para elaborar las mezclas.

   ¿Cuántos paquetes de cada gama se deberán fabricar para que el beneficio producido sea máximo?

5. En una empresa se editan revistas de dos tipos: de información deportiva y de cultura. Cada revista de información deportiva, precisa dos cartuchos de tinta negra y uno de color y se vende a 3 euros. Cada revista de cultura precisa dos cartuchos de tinta negra y dos de color y se vende a 5 euros. se dispone de 500 cartuchos de cada clase.
   ¿Cuántas revistas de cada tipo de deben editar para ingresar el máximo posible?

6. Para iluminar una sala de pintura es preciso colocar suficientes bombillas que sumen un total de 1440 vatios como mínimo. En el mercado se pueden adquirir bombillas incandescentes tradicionales de 90 vatios al precio de un euro y bombillas de bajo consumo de 9 vatios (equivalentes a 60 vatios) al precio de 5 euros la unidad.

   Debido a la estructura del espacio, el número total de bombillas no puede ser mayor de 20. Por otra parte, las normas del Ayuntamiento imponen que, para este tipo de salas, el número de bombillas de bajo consumo no puede ser inferior a la mitad del de bombillas tradicionales.

   Calcula el número de bombillas de cada clase que se deben colocar para que el coste sea mínimo.

7. Dos jóvenes empresarios se disponen a abrir un negocio de informática. Montarán y comercializarán dos tipos de ordenador: el tipo A llevará una unidad de memoria de pequeña capacidad y un disco duro; el tipo B llevará una unidad de memoria de alta capacidad y dos discos duros.

   En total se cuenta con 40 unidades de memoria de pequeña capacidad, 30 unidades de memoria de alta capacidad y 80 discos duros.

   Por cada ordenador de tipo A esperan obtener 150 euros de beneficios, y por cada ordenador de tipo B, 250 euros.
   - (a) Cuál es la mejor decisión sobre el número de ordenadores de cada tipo?
   - (b) Cuáles serían los beneficios en ese caso?
   - (c) Con esa producción, ¿habría algún excedente en el material mencionado?

8. En un taller de confección se van a elaborar trajes de cocinero y de camarero. Se dispone para ello de 30 m^2 de algodón, 10 m^2 de fibra sintética y 20 m^2 de lana.

   Para hacer cada traje de cocinero se precisan 1 m^2 de algodón, 2 m^2 de fibra sintética y 2 m^2 de lana. Cada unidad de este tipo deja 20 euros de beneficios.

   Para hacer cada traje de camarero se precisan 2 m^2 de algodón. 1 m^2 de fibra sintética y 1 m^2 de lana. Cada unidad de este tipo deja 30 euros de beneficios.

   Se deben confeccionar mayor o igual número de trajes de camarero que de cocinero y, como mínimo, se deben hacer un traje de cocinero y dos de camarero. El total no puede ser superior a 20.
   - (a) ¿Cuántos trajes de cada tipo se deberán confeccionar de forma que el beneficio sea máximo?
   - (b) ¿Sobrará algún tipo de material?
   - (c) ¿Hay alguna condición redundante?

9. Una empresaria desea invertir los beneficios de 7500 euros obtenidos en su negocio en dos tipos de acciones A y B.

   El tipo A produce un tipo de interés esperado del 6 % y el tipo B del 4 %. Como máximo desea invertir 5000 euros en A y, como mínimo, 1500 en B. Además, desea que la inversión en A sea superior a dos veces y media la inversión en B.

   ¿Cómo deberá realizar la inversión para que las ganancias sen máximas?

10. Una empresa de siderurgia cuenta con tres tipos de recursos productivos para producir dos tipos de aleaciones de hierro, A_1 y A_2: 1000 horas de trabajo de personal y 880 y 1161 toneladas, respectivamente, de dos materias primas M_1 y M_2, que se deben mezclar.

    Para fabricar una unidad de la aleación A_1 se precisan 10 horas de trabajo de personal, 20 toneladas de M_1 y 50 toneladas de M_2.

    Para fabricar una unidad de la aleación A_2 se precisan 40 horas de trabajo de personal, 50 toneladas de M_1 y 60 toneladas de M_2.

    Gracias a un estudio de mercado, se supone que, por cada unidad de A_1 se obtendrán unos beneficios de 125 unidades monetarias, y por cada unidad de A_2 se obtendrán 250 unidades monetarias.
    - (a) Halla la producción que maximiza los beneficios.
    - (b) Indica si se genera algún tipo de excedente en los recursos productivos.
    - (c) Si se produce una rebaja del 40 % en los beneficios obtenidos por cada unidad de A_2 y se mantienen los obtenidos por cada unidad de A_1, ¿cómo variará la producción óptima?

11. En la región determinada por:

$$\begin{cases} x + y \geq 2 \\ x \leq y \\ x \geq 0 \\ y \geq 0 \end{cases}$$

halla las coordenadas de los puntos en que la función $F(x, y) = 3x + 4y$ alcanza su máximo y su mínimo.

(*Solución: no existe máximo, hay un mínimo en* $(1, 1)$)

12. Con 80 kg de acero y 120 kg de aluminio se quieren fabricar bicicletas de montaña y de paseo que se venderán a 200 y 150 euros respectivamente. Para las de montaña son necesarios 1 kg de acero y 3 de aluminio y para las de paseo 2 kg de cada uno de los dos metales. ¿Cuántas bicicletas de paseo y cuántas de montaña se deben fabricar para obtener el máximo beneficio?

    (*Solución: 20 bicicletas de montaña y 30 de paseo*)

13. Una refinería de petróleo tiene dos fuentes de crudo: ligero y pesado.

    Cada barril de crudo ligero cuesta 70 dólares y con él la refinería produce 0,3 barriles de gasolina (G), 0,2 barriles de combustible de calefacción (C) y 0,3 barriles de combustible para turbinas (T). Cada barril de crudo pesado cuesta 60 dólares y produce 0,3 barriles de G, 0,4 barriles de C y 0,2 barriles de T.

    La refinería ha contratado el suministro de 900000 barriles de G, 800000 barriles de C y 500000 barriles de T. Hallar las cantidades de crudo ligero y pesado que debe comprar para poder cubrir sus necesidades con un coste mínimo.

    (*Solución:* 3 *millones de barriles de crudo pesado y ninguno de ligero*)

14. Un comerciante desea comprar dos tipos de frigoríficos F_1 y F_2. Los del tipo F_1 cuestan 300 euros y los del tipo F_2 500 euros. Solo dispone de sitio para 20 frigoríficos y de 7000 euros para hacer las compras.

    ¿Cuántos frigoríficos ha de comprar de cada tipo para obtener beneficios máximos con su venta posterior, sabiendo que en cada frigorífico gana el 30 % del precio de compra?

    (*Solución: puntos de coordenadas enteras del segmento de extremos* $(0, 14)$, $(15, 5)$.)

15. En la fabricación de piensos se utilizan tres ingredientes, P, Q y R. Se dispone de 90 toneladas de P, 90 de Q y 70 de R. Se desea fabricar dos tipos de pienso M_1 y M_2.

    Una vagoneta de pienso M_1 requiere 2 toneladas de P, 1 tonelada de Q y 1 tonelada de R y se vende a 1200 euros. Una vagoneta de M_2 requiere 1 tonelada de P, 2 de Q y 1 de R y se vende a 1000 euros.

    ¿Cuántas toneladas de cada pienso deben facturarse para obtener el máximo beneficio?

    (*Solución:* 30 *vagonetas de cada clase.*)

16. Una refinería utiliza dos tipos de petróleo, A y B, que compra a un precio de 350 euros y 400 euros por tonelada respectivamente. Por cada tonelada de petróleo de tipo A que refina, obtiene 0,10 toneladas de gasolina y 0,35 toneladas de fueloil. Por cada tonelada de petróleo de tipo B que refina, obtiene 0,05 toneladas de gasolina y 0,55 toneladas de fueloil. Para cubrir sus necesidades necesita obtener al menos 10 toneladas de gasolina y al menos 50 toneladas de fueloil. Por cuestiones de capacidad, no puede comprar más de 100 toneladas de cada tipo de petróleo. ¿Cuántas toneladas de petróleo de cada tipo debe comprar la refinería para cubrir sus necesidades a mínimo coste? Determinar dicho coste mínimo.

17. Un distribuidor de aceite de oliva compra la materia prima a dos almazaras, A y B. Las almazaras A y B venden el aceite a 2000 y 3000 euros por tonelada, respectivamente. Cada almazara le vende un mínimo de 2 toneladas y un máximo de 7 para atender a su demanda, el distribuidor debe comprar en total un mínimo de 6 toneladas. El distribuidor debe comprar como máximo a la almazara A el doble de aceite que a la almazara B. ¿Qué cantidad de aceite debe comprar el distribuidor a cada una de las almazaras para obtener el mínimo coste?. Determínese dicho coste mínimo.

18. Se desea invertir una cantidad de dinero menor o igual que 125000 euros, distribuidos entre acciones del tipo A y del tipo B. Las acciones del tipo A garantizan una ganancia del 10 % anual, siendo obligatorio invertir en ellas un mínimo de 30000 euros y un máximo de 81000 euros. Las acciones del tipo B garantizan una ganancia del 5 % anual, siendo obligatorio invertir en ellas un mínimo de 25000 euros. La cantidad invertida en acciones del tipo B no puede superar el triple de la cantidad invertida en acciones del tipo A. ¿Cuál debe ser la distribución de la inversión para maximizar la ganancia anual? Determínese dicha ganancia máxima.

1.5. Soluciones a los ejercicios

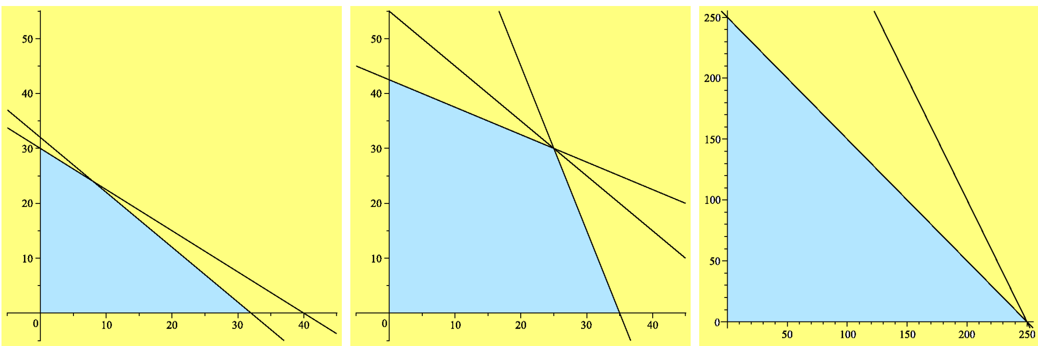


Figura 6: Problemas 3, 4 y 5

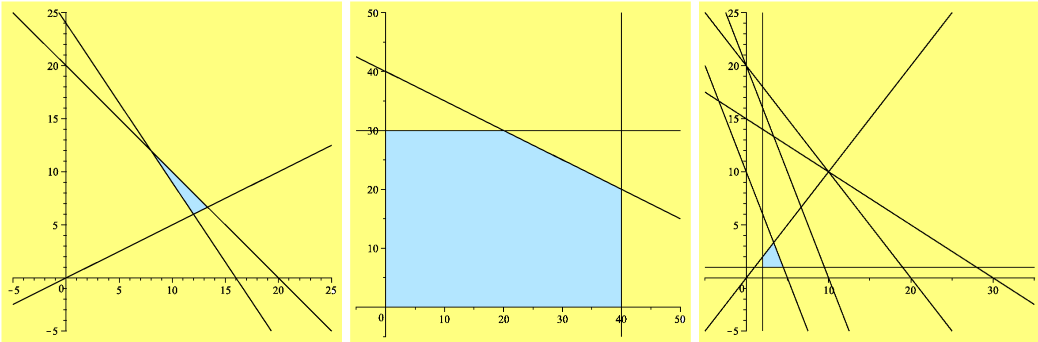


Figura 7: Problemas 6, 7 y 8

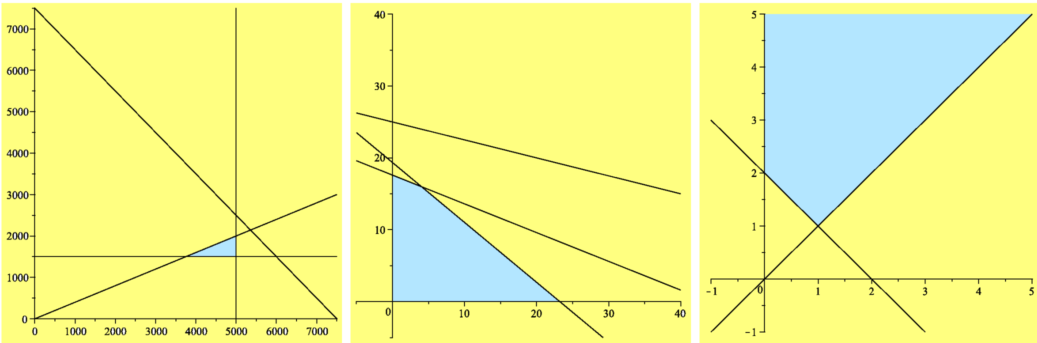


Figura 8: Problemas 9, 10 y 11

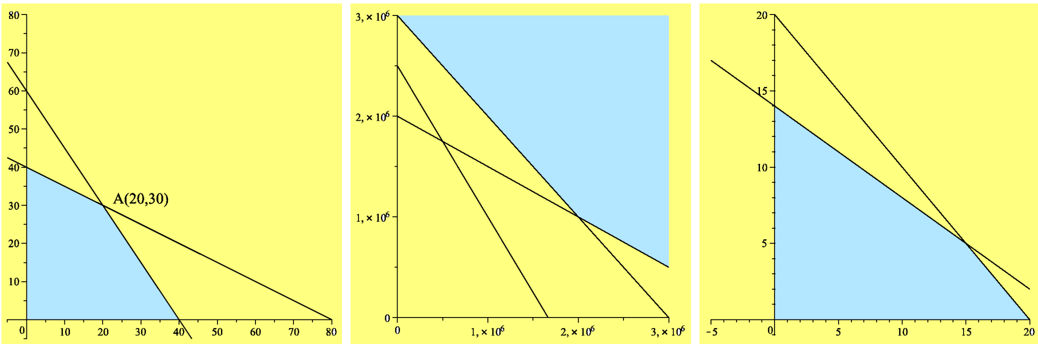


Figura 9: Problemas 12, 13 y 14

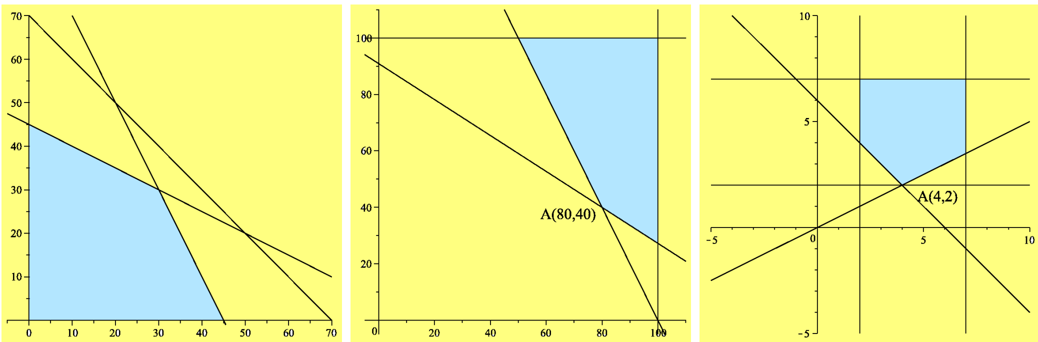


Figura 10: Problemas 15, 16 y 17

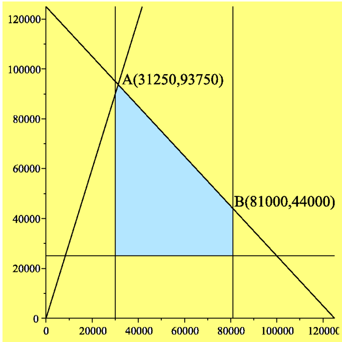


Figura 11: Problema 18

2. Estadística.

2.1. La distribución normal.

Una distribución de probabilidad es un modo de asignar probabilidades mediante una función. Sea X una variable estadística continua. Las probabilidades se definen mediante una función de densidad $g(x)$ de tal forma que la probabilidad de que la variable se encuentre comprendida entre dos valores x_1 y x_2 se define por (figura 12):

$$\mathrm{p}\left(x_1 < X < x_2\right) = \int_{x_1}^{x_2} g(x)\,\mathrm{d}x$$

Como la suma de todas las probabilidades debe ser igual a 1, la función de densidad debe cumplir que el

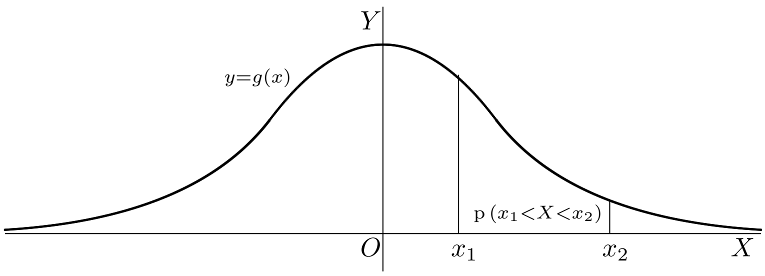


Figura 12: Distribución de probabilidad. Función de densidad.

área total bajo la curva debe ser 1, o lo que es lo mismo:

$$\int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x = 1$$

La media y la desviación típica de una distribución de probabilidad de variable continua se obtiene mediante fórmulas similares de las expuestas para variable discreta sustituyendo las sumas por integrales:

$$\mu = \int_{-\infty}^{\infty} xf(x)\,dx \qquad \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

La función de distribución $F(x)$ representa probabilidades acumuladas. Así, $F(x)$ es la probabilidad de obtener un resultado menor (o menor o igual cuya probabilidad es la misma) que x.

Conocida la función de distribución $F(x)$ puede calcularse fácilmente la probabilidad asociada a cualquier intervalo mediante diferencias:

$$\mathrm{p}\left(a < X < b\right) = F(b) - F(a)$$

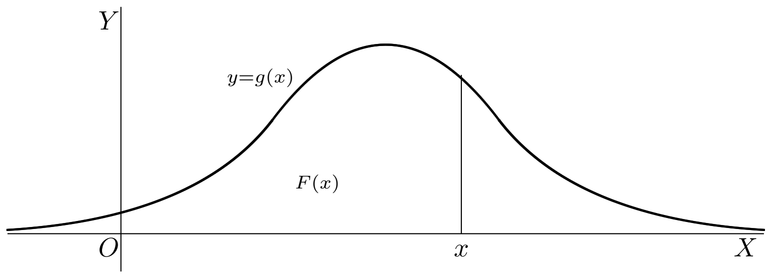


Figura 13: Función de distribución.

Una distribución de probabilidad de variable continua cuya importancia se comprenderá en el tema de inferencia estadística es la distribución normal. La distribución normal de media μ y desviación típica σ tiene la siguiente función de densidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

La distribución normal de media 0 y desviación típica 1 se indica mediante $N(0,1)$. Los valores de la función de distribución de $N(0,1)$ se encuentran en las tablas de la distribución normal (ver cuadro 1). En la tabla aparecen los valores $F(x)$ de la función de distribución solamente para valores positivos de x pero, por la simetría de la función de densidad, pueden hallarse también las probabilidades para valores negativos de la variable aleatoria.

Para obtener las probabilidades a partir de la tabla pueden aplicarse las siguientes reglas:

- ⋄ a y b positivos:
  - $\mathrm{p}\,(X < a) = F(a)$
  - $\mathrm{p}\,(X > a) = 1 - F(a)$
  - $\mathrm{p}\,(a < X < b) = F(b) - F(a)$
- ⋄ Para valores negativos:
  - $\mathrm{p}\,(X < -a) = \mathrm{p}\,(X > a) = 1 - F(a)$
  - $\mathrm{p}\,(-a < X < b) = F(a) + F(b) - 1$
  - $\mathrm{p}\,(-a < X < -b) = F(a) - F(b)$

A partir de estos valores puede obtenerse la función de distribución de $N(\mu,\sigma)$ tipificando la variable, esto es, si z es la variable de $N(0,1)$ y x la variable de $N(\mu,\sigma)$, se pasa de una a otra mediante el cambio:

$$z = \frac{x-\mu}{\sigma} \qquad \text{o bien} \qquad x = \mu + z\sigma$$

La función es simétrica respecto a $x = \mu$ y es tanto más aplanada cuanto mayor sea σ. La probabilidad correspondiente al intervalo $(\mu-\sigma,\mu+\sigma)$ (es decir al intervalo $(-1,1)$ para la variable tipificada z) es aproximadamente de $0,68$. Para el intervalo $(\mu-2\sigma,\mu+2\sigma)$ es aproximadamente de $0,95$ y para $(\mu-3\sigma,\mu+3\sigma)$ de $0,99$.

2.2. Muestras

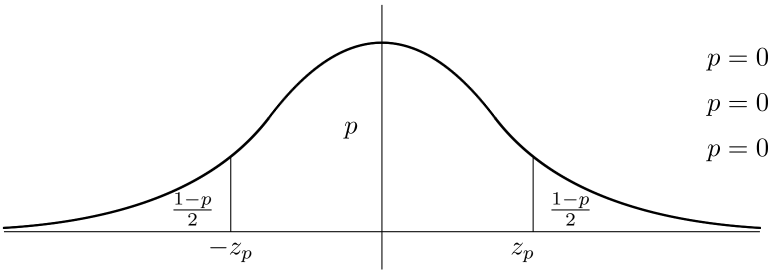
En ocasiones, resulta imposible efectuar medidas sobre todos los objetos de una población, bien debido a que su tamaño lo hace imposible, bien porque el proceso de medida es destructivo (por ejemplo cuando se mide la duración de una población de bombillas) o por otras razones. En este caso se toma una muestra y a partir de los resultados de las medidas efectuadas sobre la muestra se trata de sacar conclusiones sobre la población.

Estas conclusiones tienen necesariamente un carácter probabilístico y por tanto, junto al dato poblacional que se deduzca habrá que especificar cuál es la probabilidad de que sea erróneo.

Para poder sacar conclusiones es esencial que el **muestreo** es decir, el proceso de obtención de la muestra sea **aleatorio**. El muestreo es aleatorio si todos los elementos de la población tienen la misma probabilidad de ser elegidos.

2.3. Intervalos característicos

Consideremos una distribución normal de media cero y desviación típica 1, $N(0,1)$. Se llaman **intervalos característicos** a intervalos centrados en la media $(-z_p, z_p)$ a los que corresponde una probabilidad dada p.



$p = 0,9 \quad z_p = 1,645$

$p = 0,95 \quad z_p = 1,96$

$p = 0,99 \quad z_p = 2,575$

Figura 14: Intervalos característicos

Junto a la figura se han puesto los valores de los z_p para los valores de la probabilidad que se presentan más frecuentemente en la práctica.

El cálculo de intervalos característicos en una distribución normal cualquiera $N(\mu, \sigma)$ puede hacerse sin dificultad recordando las fórmulas de cambio a puntuaciones típicas:

$$z = \frac{x - \mu}{\sigma} \Longleftrightarrow x = \mu + z\sigma$$

por lo que el intervalo característico correspondienta a una probabilidad p será:

$$(\mu - z_p\sigma, \mu + z_p\sigma)$$

2.4. Distribución de medias muestrales. Intervalos de confianza

Supongamos que en una población se ha medido una magnitud y se ha obtenido una media μ y una desviación típica σ. Tomamos una muestra de tamaño N y nos preguntamos por los valores de la media de los valores de la muestra $\bar{x}$, es decir, nos preguntamos cuál es la probabilidad de que la media muestral $\bar{x}$ se encuentre en un cierto intervalo (a, b).

El **teorema central del límite** establece que si los valores de la variable en la población se distribuyen normalmente segun la distribución $N(\mu, \sigma)$, las medias muestrales se distribuyen normalmente con la misma media μ y una desviación típica $\sigma/\sqrt{N}$. Además, para muestras grandes ($N > 30$) se puede aplicar el teorema aunque la población de partida no sea normal.

En resumen, si la población es normal o si no siéndolo $N > 30$ podemos obtener probabilidades para los valores de la media muestral a partir de la distribución:

$$N\left(\mu, \frac{\sigma}{\sqrt{N}}\right)$$

En la práctica, el valor de la desviación típica de la población σ es desconocido por lo que se toma como valor aproximado, la desviación típica de la muestra s.

En una población, una variable estadística tiene una media μ y una desviación típica σ. Se toma una muestra de tamaño N y se mide la media muestral de la misma magnitud $\bar{x}$. Si las medias muestrales se distribuyen normalmente, podemos decir que con una probabilidad c, llamada **nivel de confianza**, $\bar{x}$ se encontrará en el intervalo $(\mu - z_c\sigma/\sqrt{N}\ ,\ \mu + z_c\sigma/\sqrt{N})$.

Con la misma probabilidad podemos decir que la media poblacional se encontrará en el siguiente intervalo:

$$\left(\bar{x} - z_c \frac{\sigma}{\sqrt{N}}\ ,\ \bar{x} + z_c \frac{\sigma}{\sqrt{N}}\right)$$

llamado **intervalo de confianza** para la media correspondiente a un nivel de confianza c.

2.5. Ejercicios resueltos.

1. Se supone que el nivel de glucosa en sangre de los individuos de una población (medido en miligramos por decilitro) se puede aproximar por una variable aleatoria con distribución normal de media μ desconocida y desviación típica igual a 35 mg/dl. ¿Cuál es el tamaño muestral mínimo que permite garantizar que el valor absoluto de la diferencia entre la media muestral y μ es menor que 20 mg/dl con una probabilidad mayor o igual que 0,98?

   Solución:

   Si el nivel de confianza es de 0,98 la tabla de la distribución normal nos da $z_c = 2{,}33$. El error en la estimación está dado por:

   $$E = \frac{z_c \sigma}{\sqrt{N}}$$

   donde N es el tamaño de la muestra. Puesto que el erro debe ser menor que 20, tenemos que:

   $$\frac{z_c \sigma}{\sqrt{N}} < 20 \quad \Longrightarrow \quad \frac{2{,}33 \cdot 35}{\sqrt{N}} < 20 \quad \Longrightarrow \quad \sqrt{N} > \frac{2{,}33 \cdot 35}{20}$$

   De aquí obtenemos:

   $$N > \left(\frac{2{,}33 \cdot 35}{20}\right)^2 = 16{,}626$$

   Como el tamaño de la muestra debe ser un número entero tiene que ocurrir que $N \geq 17$.

2. Se supone que el precio de un kilo de patatas en una cierta región se puede aproximar por una variable aleatoria con distribución normal de desviación típica igual a 10 céntimos de euro. una muestra aleatoria simple de tamaño 256 proporciona un precio medio del kilo de patatas igual a 19 céntimos de euro.

   (*a*) Determínese un intervalo de confianza del 95 % para el precio medio de un kilo de patatas en la región.

   (*b*) Se desea aumentar el nivel de confianza al 99 % sin aumentar el error en la estimación. ¿Cuál debe ser el tamaño muestral mínimo que ha de observarse?

   Solución:

   (*a*) El intervalo de confianza es:

   $$\left(\bar{x} - \frac{z_c \sigma}{\sqrt{N}}\ ,\ \bar{x} + \frac{z_c \sigma}{\sqrt{N}}\right)$$

   Con los datos del problema $\bar{x} = 19$, $\sigma = 10$, $N = 256$ y $z_c = 1{,}96$. Sustituyendo obtenemos el intervalo:

   $$(17{,}775\ ;\ 20,225)$$

   (*b*) El error en la estimación anterior es

   $$E = \frac{z_c \sigma}{\sqrt{N}} = \frac{1{,}96 \cdot 10}{\sqrt{2576}} = 1{,}225$$

   Si queremos mantener el mismo error aumentando el nivel de confianza al 99 % (es decir $z_c = 2{,}575$) tenemos que:

   $$1{,}225 = \frac{2{,}575 \cdot 10}{\sqrt{N}} \quad \Longrightarrow \quad N = \left(\frac{2{,}575 \cdot 10}{1{,}225}\right)^2 = 441,858$$

   Por consiguiente el tamaño mínimo debe ser como mínimo de 442.

3. (*a*) La edad de un determinado grupo de personas sigue una distribución $N(37, 9)$. Calcula la probabilidad de que una persona de ese grupo, elegido al azar, tenga más de 40 años.

   (*b*) La edad de los alumnos de segundo de bachillerato de cierto instituto sigue una distribución $N(17{,}8; 0{,}6)$. Los agrupamos al azar de 10 en 10 para una competición. Halla el intervalo característico del 95 % correspondiente a las edades medias de los grupos.

   Solución:

   (*a*) $\mathrm{p}\,(x \geq 40) = \mathrm{p}\left(z \geq \dfrac{40 - 35}{10}\right) = p(z \geq 0{,}50) = 0{,}3085$

   (*b*) Las medias muestrales se distribuyen de acuerdo con $N\left(17{,}6\ ;\ \frac{0{,}5}{\sqrt{10}}\right) = N(17{,}6\ ;\ 0{,}158)$: El intervalo característico es:

   $$(17{,}6 - 1{,}96 \cdot 0{,}158\ ;\ 17{,}6 + 1{,}96 \cdot 0{,}158) = (17{,}2903\ ;\ 17{,}9097)$$

4. En una determinada empresa, se seleccionó al azar una muestra de 100 empleados cuya media de ingresos mensuales resultó igual a 705 euros, con una desviación típica de 120 euros. Halla un intervalo de confianza al 99 % para la media de los ingresos mensuales de todos los empleados de la empresa.?

Solución:
Los datos son $\bar{x} = 705$, $c = 0{,}99$, $\sigma = 120$, $N = 100$. Con estos datos, el intervalo de confianza es:

$$\left(705 - 2{,}575\frac{120}{\sqrt{100}}\ ;\ 705 + 2{,}575\frac{120}{\sqrt{100}}\right) = (674{,}10\ ;\ 735{,}90)$$

5. La duración de un lavavajillas sigue una distribución normal con una desviación típica de 0,5 años. ¿Cuántos lavavajillas tenemos que seleccionar en la muestra si queremos que la media muestral no difiera en más de 0,25 años de la media de la población. con un nivel de confianza del 90 %?

Solución:
El error en la estimación debe ser menor que 0,25. Esto significa queda

$$E = \frac{z_c\sigma}{\sqrt{N}} \leq 0{,}25 \quad\Longrightarrow\quad N \geq \left(\frac{z_c\sigma}{0{,}25}\right)^2 = \left(\frac{1{,}645\cdot 0{,}5}{0{,}25}\right)^2 = 10{,}82$$

Por consiguiente, puesto que el tamaño de la muestra debe ser entero, $N \geq 11$.

6. Un fabricante de lámparas de bajo consumo sabe que el tiempo de duración, en horas, de las lámparas que fabrica sigue una distribución normal de media desconocida y desviación típica 180 horas. Con una muestra de dichas lámparas elegida a azar y con un nivel de confianza del 97 %, obtuvo para la media el intervalo de confianza $(10072,1; 10127,9)$. Si se quiere que el error de su estimación sea como máximo de 24 horas y se utiliza una muestra de tamaño 225, ¿cuál será entonces el nivel de confianza?

Solución:
Puesto que el error en la estimación debe ser menor que 24, debe cumplirse que:

$$E = \frac{z_c\sigma}{\sqrt{N}} = \frac{z_c\cdot 180}{15} \leq 24 \quad\Longrightarrow\quad z_c \leq \frac{24\cdot 15}{180} = 2$$

Por consiguiente:

$$c \leq \mathrm{p}\,(-2 \leq z \leq 2) = 0{,}9772 - 0{,}0228 = 0{,}9544$$

7. El número de días de ausencia en el trabajo de los empleados de cierta empresa para un período de seis meses, se puede aproximar mediante una distribución normal de desviación típica 1,5 días. Una muestra aleatoria de 10 empleados ha proporcionado los siguientes datos:

5 4 6 8 7 4 2 7 6 1

(*a*) Determinar el intervalo de confianza del 90 % para el número medio de días que los empleados de esa empresa han faltado durante los últimos seis meses.

(*b*) ¿Qué tamaño debe tener la muestra para que el error máximo de la estimación sea de 0,5 días con el mismo nivel de confianza?

Solución:

(*a*) La media de los datos es 5. Así pues, $\bar{x} = 5$, $c = 0{,}90$, $z_c = 1{,}645$ y $N = 10$. El intervalo de confianza es:

$$\left(5 - 1{,}645\frac{1{,}5}{\sqrt{10}}, 5 + 1{,}645\frac{1{,}5}{\sqrt{10}}\right) = (4{,}22; 5{,}78)$$

(*b*) El error en la estimación debe ser menor que 0,5. Entonces:

$$\frac{z_c\sigma}{\sqrt{N}} \leq 0{,}5 \quad\Longrightarrow\quad \frac{1{,}645\cdot 1{,}5}{\sqrt{N}} \leq= ,5 \quad\Longrightarrow\quad N \geq \left(\frac{1{,}645\cdot 1{,}5}{0{,}5}\right)^2 = 24{,}35$$

Por consiguiente, puesto que N es entero, debe verificarse que $N \geq 25$.

8. La temperatura corporal de una cierta especie animal es una variable aleatoria que tiene una distribución normal de media 36,7ºC y desviación típica 3,8ºC. se elige aleatoriamente una muestra de 100 ejemplares de esa especie. Hallar la probabilidad de que la temperatura corporal media de la muestra:

(*a*) Sea menor o igual que 36,9ºC.

(*b*) Esté comprendida entre 36,5ºC y 37,3ºC.

Solución:
Las medias muestrales siguen la distribución normal:

$$N\left(36{,}7, \frac{3{,}8}{\sqrt{100}}\right) = N(36{,}7; 0{,}38)$$

Entonces

$$\mathrm{p}\,(x \leq 36{,}9) = \mathrm{p}\left(z \leq \frac{36{,}9 - 36{,}7}{0{,}38}\right) = \mathrm{p}\,(z \leq 0{,}52) = 0{,}6985$$

$$\mathrm{p}\,(36{,}5 \leq x \leq 37{,}3) = \mathrm{p}\left(\frac{36{,}5 - 36{,}7}{0{,}38} \leq z \leq \frac{37{,}3 - 36{,}7}{0{,}38}\right) = \mathrm{p}\,(-0{,}53 \leq z \leq 1{,}58) = 0{,}6448$$

9. En una encuesta se pregunta a 10000 personas cuántos libros lee al año, obteniéndose una media de 5 libros. Se sabe que la población tiene una distribución normal con desviación típica 2.

(a) Hallar un intervalo de confianza al 80 % para la media poblacional.

(b) Para garantizar un error de estimación de la media poblacional no superior a 0,25 con un nivel de confianza del 95 %, ¿a cuántas personas como mínimo sería necesario entrevistar?

Solución:

Calculemos en primer lugar el valor de z_c para $c = 0{,}80$:

$$\mathrm{p}\left(-z_c \leq z \leq z_c\right) = 0{,}80 \quad \Longrightarrow \quad \mathrm{p}\left(z \leq z_c\right) = 0{,}90 \quad \Longrightarrow \quad z_c = 1{,}28$$

(a) Puesto que $\bar{x} = 5$, $\sigma = 2$, $z_c = 1{,}28$ y $N = 10000$, el intervalo es:

$$\left(5 - 1{,}28\frac{2}{\sqrt{1000}}; 5 + 1{,}28\frac{2}{\sqrt{1000}}\right) = (4{,}97; 5{,}03)$$

(b) Debe ocurrir queda

$$N \geq \left(\frac{1{,}96 \cdot 2}{0,25}\right)^2 = 245{,}86 \quad \Longrightarrow \quad N \geq 246$$

TABLA DE LA DISTRIBUCIÓN NORMAL

z	0.00	0.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Cuadro 1: Tabla de la distribución normal