Estadística – Licenciatura en Nutrición - UCEL

Capítulo 5: Modelizar para decidir

RESOLUCIÓN DE EJERCICIOS PROPUESTOS EN LAS DIAPOSITIVAS

Manos a la Obra N° 1

VARIABLES ALEATORIAS

1.
   a) Seleccionar al azar una mujer y registrar el número de frutas que consumen en un día.
   b) X: número de frutas que se consume en un día. Es una variable aleatoria discreta.
   c) Se sabe que la suma de todos los p(*x*) es 1, entonces P(X=6)=0,03

X=x	**0**	**1**	**2**	**3**	**4**	**5**	**6**
P(X=x)	0,15	0,22	0,10	0,15	0,35	0	0,03

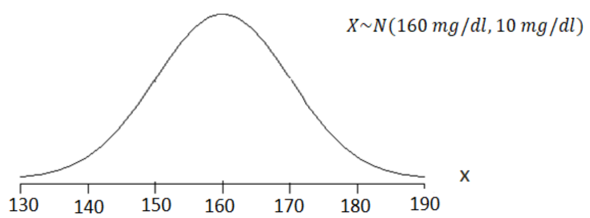
   d) La probabilidad de seleccionar una mujer al azar y que no haya consumido frutas es 0,15.

2.
   a) X: tiempo que demora (en minutos) un medicamento en calmar el dolor de cabeza. Variable aleatoria continua.
   b) No, ya que el comportamiento de una variable aleatoria continua viene dada por una función de densidad de probabilidad.
   c) La probabilidad de un valor puntual, cuando estamos analizando variables aleatorias continuas, es igual a cero. Por lo tanto, la probabilidad de que una mujer seleccionada al azar se le pase el dolor de cabeza a los 30 minutos es igual a cero. Sin embargo, se asignan probabilidades a intervalos de resultados, por lo tanto, la probabilidad de que el tiempo en calmar el dolor de cabeza sea más de 30 minutos pero menos de 60 minutos es posible calcularla obteniendo el área debajo de la curva.

Manos a la Obra N° 2

DISTRIBUCIÓN NORMAL

1. Sea X: el nivel de colesterol demujeres sanas

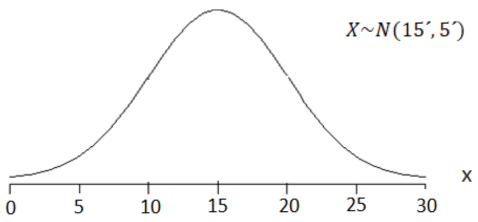


a) Si aplicamos la regla empírica, el 95% de las mujeres analizadas está entre $\mu - 2\sigma$ y $\mu + 2\sigma$, es decir, que el 95% de las mujeres analizadas está entre 140 mg/dl y 180 mg/dl.

b) La curva de la distribución del colesterol de hombres estaría más a la derecha y más aplanada con respecto a la curva de la distribución del colesterol de mujeres porque la media poblacional y desvío estándar poblacional son mayores.

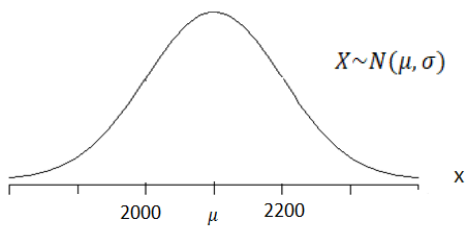
2. Sea X: la demora (en minutos) de un paciente hasta ser atendido

a)



b) Si aplicamos la regla empírica, el 99,7% de los pacientes analizados está entre $\mu - 3\sigma$ y $\mu + 3\sigma$, es decir, que el 99,7% de los pacientes tiene de demora entre 0 minutos y 30 minutos.

3. Sea X: el consumo de calorías de adolescentes por día



a) Si aplicamos la regla empírica, el 68% de los adolescentes está entre $\mu - \sigma$ y $\mu + \sigma$, es decir, que el 68% de los pacientes consumen entre 2000 calorías y 2200 calorías.

Obtengo un sistema de ecuación lineal de 2x2:

$$\begin{cases} \mu - \sigma = 2000 \\ \mu + \sigma = 2200 \end{cases} \Rightarrow \begin{cases} \mu - \sigma = 2000 \Rightarrow \mu = 2000 + \sigma \\ \qquad\qquad \mu + \sigma = 2200 \end{cases}$$

(Despejo una incógnita en la primera ecuación y la reemplazo en la segunda ecuación)

$$\mu + \sigma = 2200 \Rightarrow 2000 + \sigma + \sigma = 2200 \Rightarrow 2\sigma = 2200 - 2000 \Rightarrow \sigma = 100$$

(Una vez que obtengo una incógnita la reemplazo en cualquiera de las dos ecuaciones)

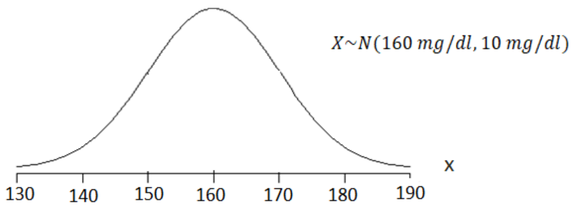
$$\mu - \sigma = 2000 \Rightarrow \mu - 100 = 2000 \Rightarrow \mu = 2100$$

Por lo tanto, el valor medio del consumo de calorías por día en adolescentes es de 2100 calorías y el desvío estándar es de 100 calorías.

Manos a la Obra N° 3

DISTRIBUCIÓN NORMAL ESTÁNDAR

Sea X: el nivel de colesterol de mujeres sanas



a) $z = \frac{X-\mu}{\sigma} = \frac{137-160}{10} = -2,3$
El valor 137 mg/dl está a 2,3 desvíos estándares por debajo del nivel medio de colesterol de mujeres sanas.

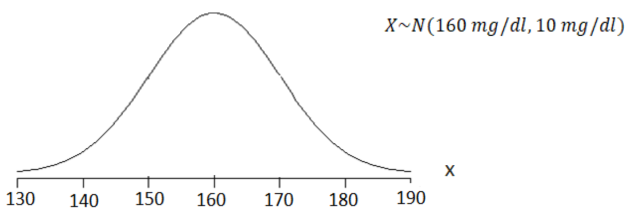
b) $z = \frac{X-\mu}{\sigma} = \frac{195-160}{10} = 3,5$
El valor 195 mg/dl está a 3,5 desvíos estándares por encima del nivel medio de colesterol de mujeres sanas.

c) $z = \frac{X-\mu}{\sigma} \Longrightarrow 1,2 = \frac{X-160}{10} \Longrightarrow 1,2.10 + 160 = X \Longrightarrow 172 = X$
El nivel de colesterol es de 172 mg/dl para un valor de z=1,2.

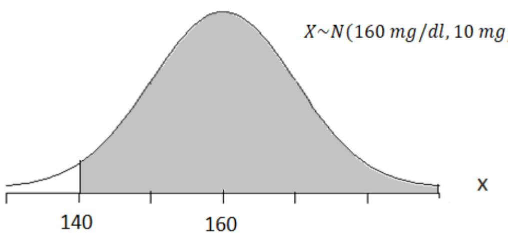
Manos a la Obra N° 4

CALCULANDO PROPORCIONES NORMALES

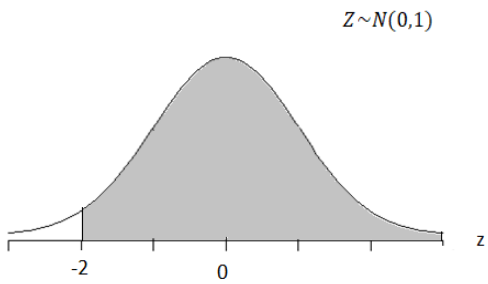
1. Sea X: el nivel de colesterol para mujeres sanas



a) P(X>140)



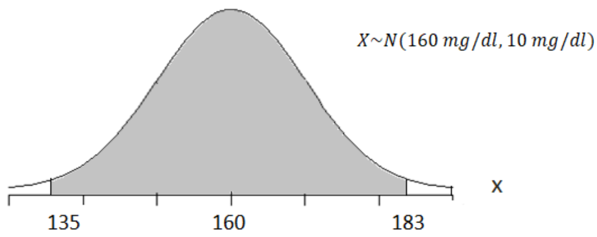
Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{140-160}{10} = -2$



P(X>140)= P(Z>-2)=1-P(Z<-2)=1-0.0228=0,9772

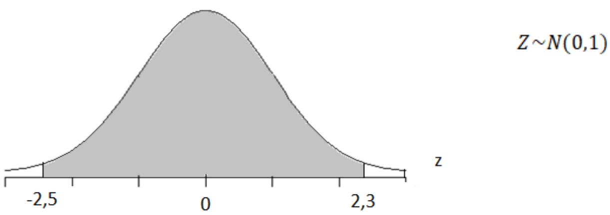
El 97,72% de las mujeres sanas tienen al menos 140 mg/dl de colesterol.

b) P(135<X<183)



Se estandariza: $z_1 = \frac{X-\mu}{\sigma} = \frac{135-160}{10} = -2{,}5$

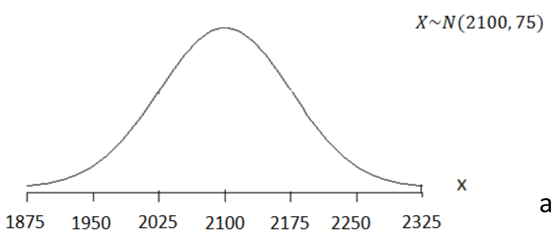
$$z_2 = \frac{X-\mu}{\sigma} = \frac{183-160}{10} = 2{,}3$$



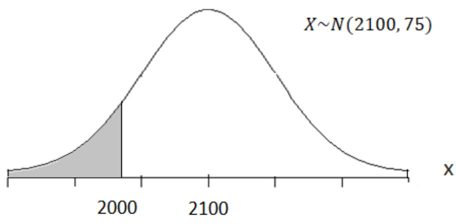
P(135<X<183)= P(-2,5<Z<2,3)=P(Z<2,3)- P(Z<-2,5)=0,9893-0,0062=0,9831

El 98,31% de las mujeres sanas tienen entre 135 mg/dl y 183 mg/dl de colesterol.

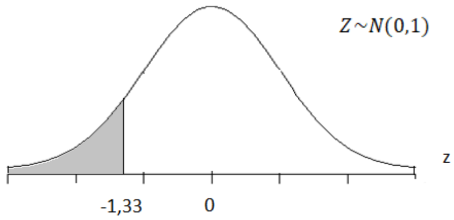
2. Sea X: el consumo de calorías de adolescentes por día



a) P(X<2000)



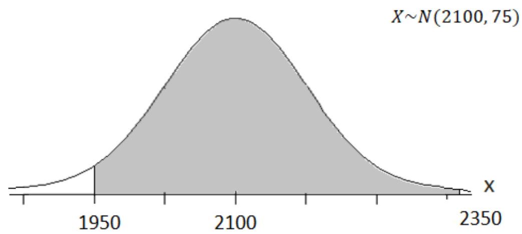
Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{2000-2100}{75} = -1{,}33$



P(X<2000)=P(Z<-1,33)=0,0918

La proporción de adolescentes que consumen a lo sumo 2000 calorías al día es 0,0918.

b) P(1950<X<2350)=



Se estandariza: $z_1 = \frac{X-\mu}{\sigma} = \frac{1950-2100}{75} = -2$

$z_2 = \frac{X-\mu}{\sigma} = \frac{2350-2100}{75} = 3{,}33$

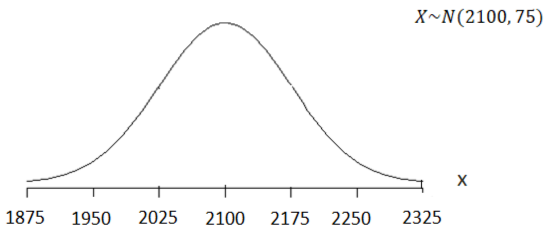
P(1950<X<2350)= P(-2<Z<3,33)=P(Z<3,33)-P(Z<-2)=0,996-0,0228=0,9732

El 97,32% de adolescentes consumen entre 1950 y 2350 calorías al día.

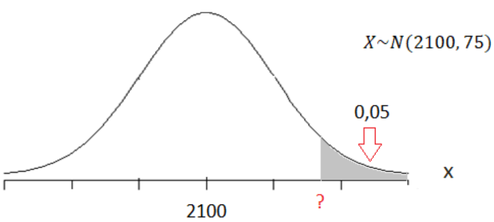
<u>Manos a la Obra N° 5</u>

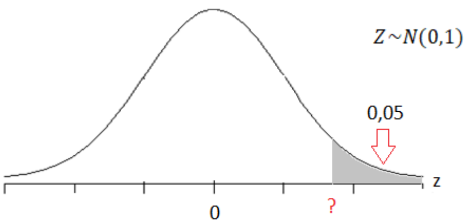
PERCENTILES

Sea X: el consumo de calorías de adolescentes por día



a) P(X>x)=0,05





P(Z>z)=0,05⟹ P(Z>z)=1-P(Z<z) ⟹0,05=1-P(Z<z) ⟹ P(Z<z)=0,95⟹z=1,645

$$z = \frac{X-\mu}{\sigma} \Longrightarrow 1{,}645 = \frac{X-2100}{75} \Longrightarrow 1{,}645.75 + 2100 = X \Longrightarrow 2223{,}375 = X$$

La caloría que representa el 5% superior de la distribución es 2223,375.

b) Si se aplica la regla empírica se sabe que el 95% de las observaciones está entre $\mu - 2\sigma$ y $\mu + 2\sigma$, por lo tanto debido a que el área debajo de la curva es 1 y por la simetría de la distribución, las calorías que representan el 2,5% inferior de la distribución y el 2,5% superior de la distribución son 1950 y 2250 calorías.

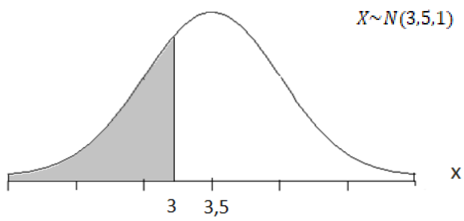
Manos a la Obra N° 6

PRUEBA DE HIPÓTESIS

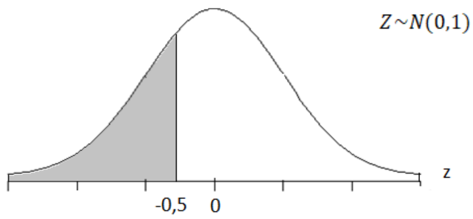
1. Sea X: ingesta diaria de frutas y hortalizas de los operarios.

   a) $H_0)\ \mu = 3{,}5\ porciones\ vs.\ H_1)\ \mu > 3{,}5\ porciones$

   b) P(X<3)



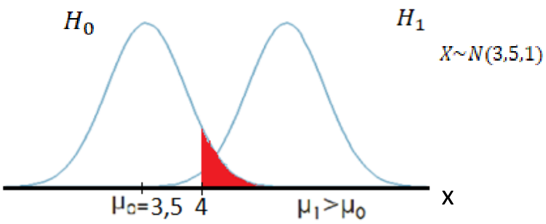
Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{3-3,5}{1} = -0,5$



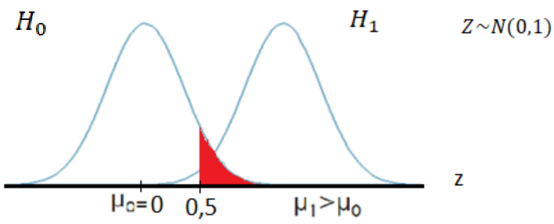
P(X<3)= P(Z<-0,5)=0,3085

La probabilidad de que un operario al azar consuma como máximo 3 porciones es igual a 0,3085

c) p-value=P(X>4)



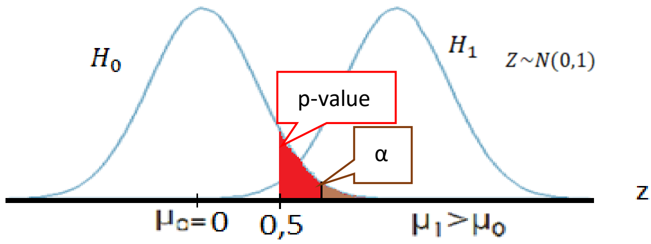
Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{4-3,5}{1} = 0,5$



p-value=P(X>4)= P(Z>0,5)=1-P(Z<0,5)=1-0,6915=0,3085

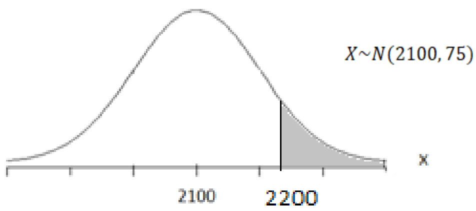
La probabilidad de observar que un operario consuma 4 porciones de frutas y hortalizas o más es 0,3085.

d) p–value>0,05⇒ No Rechazo H_0

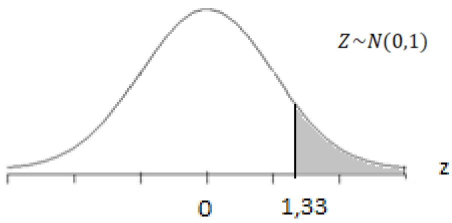


Con un nivel de significación del 5% y en base a la evidencia muestral es de esperar que la ingesta diaria promedio de frutas y hortalizas de los operarios no haya incrementado.

2. Sea X el consumo de calorías de adolescentes por día.
   a) $H_0)\ \mu = 2100\ calorías\ vs.H_1)\ \mu < 2100\ calorías$
   b) P(X>2200)

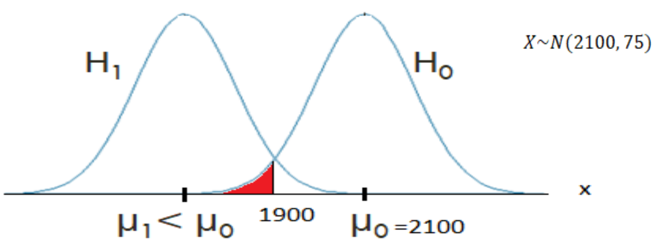


Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{2200-2100}{75} = 1{,}33$

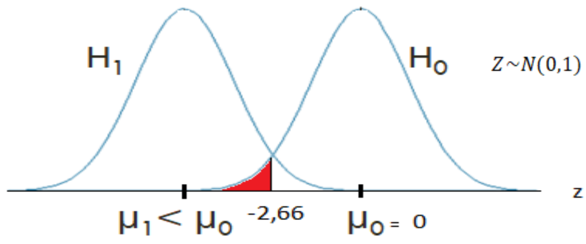


P(X>2200)= P(Z>1,33)=1- P(Z<1,33)=1-0,9082=0,0918

   c) p-value=P(X<1900)



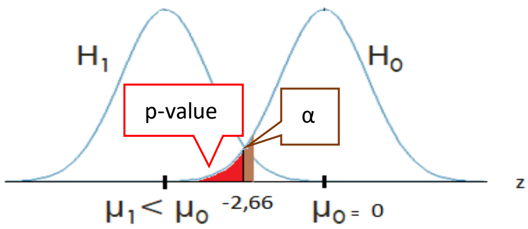
Se estandariza: $z = \frac{X-\mu}{\sigma} = \frac{1900-2100}{75} = -2{,}66$



p-value=P(X<1900)= P(Z<-2,66)=0,0039

La probabilidad de observar que un adolescente consuma 1900 calorías o menos es 0,0039.

e) p-value<0,01⇒ Rechazo H_0



Con un nivel de significación del 1% y en base a la evidencia muestral es de esperar que el consumo diario promedio de calorías de los adolescentes disminuyó.