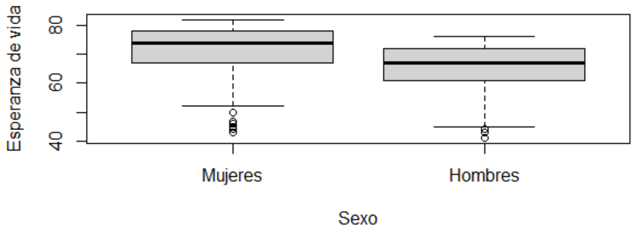
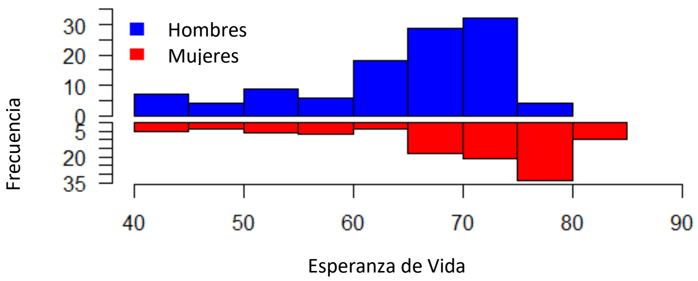
2 de Noviembre de 2022

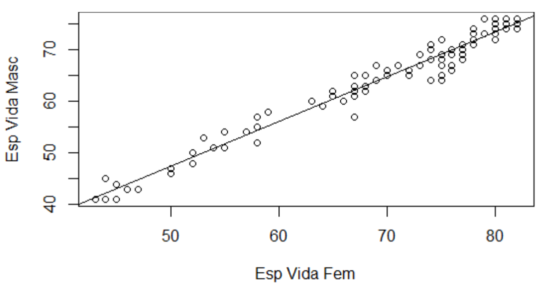
Tiempo: 75 minutos

- **No está pertimitido el uso de documentación excepto la hoja de fórmulas y la tabla de la normal Z(0,1).**
- **Utilice 4 decimales en todos los cálculos y resultados.**

1. A continuación se incluye una salida de R correspondiente al análisis descriptivo de las variables esperanza de vida masculina y esperanza de vida femenina con valores para 109 países del mundo.








```
Call:
lm(formula = Paises$"Esp vida Masc" ~ Paises$"Esp vida Fem")

Coefficients:
(Intercept)  Paises$"Esp vida Fem"
     4.4676               0.8616
```

```
Paises$`Esp vida Fem`
N: 109
                  Esp vida Fem
----------------- ----------------
           Mean         70.16
        Std.Dev         10.57
            Min         43.00
             Q1         67.00
         Median         74.00
             Q3         78.00
            Max         82.00
            MAD          7.41
            IQR         11.00
             CV          0.15

       Skewness         -1.08
    SE.Skewness          0.23
       Kurtosis          0.09
        N.Valid        109.00
      Pct.Valid        100.00
```

```
Paises$`Esp vida Masc`
N: 109
                  Esp vida Masc
----------------- -----------------
           Mean         64.92
        Std.Dev          9.27
            Min         41.00
             Q1         61.00
         Median         67.00
             Q3         72.00
            Max         76.00
            MAD          8.90
            IQR         11.00
             CV          0.14
       Skewness         -1.05
    SE.Skewness          0.23
       Kurtosis          0.21
        N.Valid        109.00
      Pct.Valid        100.00
```

a) (0,5 puntos) ¿Son simétricas las distribuciones? Justifique la respuesta mediante apreciación visual y confirmelo con las correspondientes medidas estadísticas.
No, ambas distribuciones presentan una cola en los valores menores, por lo que presentan asimetría negativa. En el caso de la esperanza de vida femenina el coeficiente de asímetria tiene un valor de -1,08, y en el de la esperanza de vida masculina -1,05.

b) (0,5 puntos) Explique qué género se espera que viva más cuantificando la respuesta.
Se espera que sean las mujeres quienes vivan más ya que el valor máximo de la esperanza de vida femenina es 82 años, mientras que el de la esperanza de vida masculina es 76 años.

c) (0,5 puntos)¿Hay valores extraños en ambas distribuciones? Si es así indique alguno en cada distribución.
Ambas distribuciones presentan valores extraños, por ejemplo, en el caso de la esperanza de vida femenina 43 años, y en el de la masculina 41 años.

d) (0,5 puntos)¿Qué distribución tiene una varianza mayor? Justifique la respuesta con al menos dos medidas estadísticas.
La varianza de la esperanza de vida femenina es mayor que la de la masculina, ya que su rango es mayor (82-43= 39años) frente a (76-41=35) de la esperanza de vida masculina. Además la desviación típica de la esperanza de vida femenina (10,57) es mayor que la desviación típica de la esperanza de vida masculina (9,27)

e) (0,5 puntos)¿Existe una relación lineal entre ambas variable? Justifique su respuesta, y en caso de ser positiva escriba un posible modelo de regresión para dicha relación.
Si existe una fuerte relación lineal entre ambas variables tal y como muestra el diagrama de dispersión que presenta una recta atravesando por la mitad la nube de puntos correspondientes a las observaciones. La ecuación del modelo es:
*Esperanza de vida masculina = 4,4676 + 0,8616*esperanza de vida femenina*

f) (0,5 puntos) Indique el valor de la covarianza entre la esperanza de vida masculina y la esperanza de vida femenina, y del coeficiente de correlación.

$Cov(x,y) = b * S^2_x$

$b = 0,8616$; $S^2_x = (10,57)^2 = 111,7249$

$Cov(x,y) = 96,2621$

$r_{xy} = Cov(x,y)/S_x S_y = 96,2621/10,57*9,27 = 0,9824$

2. De los datos obtenidos de las taquillas de una estación de esquí se sabe que la temporada pasada, de los 100.000 clientes que pasaron por caja, 20.000 clientes pagaron en metálico y el resto con tarjeta. Se sabe también que el 40% de los que pagaron en metálico gastaron en esa compra más de 100 euros, y que el 60% de los que pagaron con tarjeta gastaron más de 100 euros.

a) (0,5 puntos) Establezca los eventos considerados en el problema e indique sus probabilidades.
M – Pagar en metálico P(M)=0,2
T – Pagar con tarjeta P(T) = 0,8
C100 – Realizar una compra superior a 100€
Realizar una compra superior a 100€ cuando se paga en metálico P(C100/M) = 0,4
Realizar una compra superior a 100€ cuando se paga con tarjeta P(C100/T) = 0,6

b) (0,5 puntos) ¿Cuál es la probabilidad de que un cliente elegido al azar pague en metálico y gaste más de 100€?
*P(M∩C100) = P(C100/M) * P(M) = 0,4*0,2 = 0,08*

c) (1 punto) ¿Son independientes los sucesos "pagar en metálico" y "gastar más de 100€"?
Para que sean independientes se tiene que verificar que la P(C100) = P(C100/M)
Calculamos la probabilidad de C100 mediante el teorema de la probabilidad total:
*P(C100) = P(C100/M) * P(M) + P(C100/T) * P(T) = 0,4*0,2+0,6*0,8 = 0,56*
P(C100/M) =/ P(C100), por lo tanto, no son sucesos independientes

d) (0,5 puntos) ¿Cuál es la probabilidad de que pague en metálico si el importe de la compra es superior a 100€?
*Por el teorema de Bayes P(M/C100) = P(C100/M) * P(M) / P(C100)*
*P(M/C100) = 0,4*0,2/0,56 = 0,1428*

e) (0,5 puntos) Dos amigos llegan a la taquilla, ¿Cuál es la probabilidad de que los dos paguen en metálico? ¿Qué suposición hay que hacer para calcular esta probabilidad?
Suponemos que son independientes los sucesos de pagar cada amigo, por tanto, la probabilidad de que ambos paguen en metálico ser $P(M_1) * P(M_2) = 0,2*0,2 = 0,04$

3. El peso de unas piezas se distribuye normalmente. Se sabe que un 58% de las piezas pesan menos de 75 g, un 38% entre 75 y 80 g, y el 4% restante, más de 80 g.
   a) (2 puntos) Calcule la media y la desviación típica de la distribución
   X – Peso de las piezas y sigue una N(μ,σ)
   P(X<75) = 0,58; P(X<80)=0,96
   P(Z<(75-μ)/σ)=0,58 ---- (75-μ)/σ = 0,20
   P(Z<(80-μ)/σ)=0,96 ---- (80-μ)/σ = 1,75
   Resolviendo el sistema de estas dos ecuaciones obtenemos μ = 74,35 y σ = 3,226

   b) (1 punto) Las piezas se empaquetan en cajas de 10 unidades. Calcula la probabilidad de que una caja pese mas de 750 g. (Nota: si no ha obtenido los valores de μ y σ en el apartado anterior puede utilizar μ = 75 y σ = 3)
   La variable Y – Peso de la caja sigue una $N(\mu*10,RAIZ(10*\sigma^2))$ ---
   *N(74,35*10;RAIZ(10*10,407) = N(743,5; 10,2)*
   P(Y>750) = 1 – P(Y<=750) = 1 – P(Z<= (750-743,5)/10,2)= 1-P(Z<=0,637)=1-0,738 = 0,262

   c) (1 punto) Se toman tres cajas al azar. Calcule la probabilidad de que 2 de ellas pesen más de 750 g.

La variable W – Número de cajas entre 3 que pesan mas de 750 g, sigue una B(3;0,262)

Entonces la P(W=2) = $\binom{3}{2}\ (0{,}262)^2(0{,}738) = 0{,}152$