

# Trabajo Multivariado

## Parcial 3

Ana María López - Pedro Pablo Villegas - Esteban Tabares

Noviembre, 2017

### PUNTO 5.4

Use los datos del sudor de la tabla 5.1:

| <b>Table 5.1</b> Sweat Data             |                       |                   |                      |
|---|-----------------------|-------------------|----------------------|
| Individual                              | $X_1$<br>(Sweat rate) | $X_2$<br>(Sodium) | $X_3$<br>(Potassium) |
| 1                                       | 3.7                   | 48.5              | 9.3                  |
| 2                                       | 5.7                   | 65.1              | 8.0                  |
| 3                                       | 3.8                   | 47.2              | 10.9                 |
| 4                                       | 3.2                   | 53.2              | 12.0                 |
| 5                                       | 3.1                   | 55.5              | 9.7                  |
| 6                                       | 4.6                   | 36.1              | 7.9                  |
| 7                                       | 2.4                   | 24.8              | 14.0                 |
| 8                                       | 7.2                   | 33.1              | 7.6                  |
| 9                                       | 6.7                   | 47.4              | 8.5                  |
| 10                                      | 5.4                   | 54.1              | 11.3                 |
| 11                                      | 3.9                   | 36.9              | 12.7                 |
| 12                                      | 4.5                   | 58.8              | 12.3                 |
| 13                                      | 3.5                   | 27.8              | 9.8                  |
| 14                                      | 4.5                   | 40.2              | 8.4                  |
| 15                                      | 1.5                   | 13.5              | 10.1                 |
| 16                                      | 8.5                   | 56.4              | 7.1                  |
| 17                                      | 4.5                   | 71.6              | 8.2                  |
| 18                                      | 6.5                   | 52.8              | 10.9                 |
| 19                                      | 4.1                   | 44.1              | 11.2                 |
| 20                                      | 5.5                   | 40.9              | 9.4                  |
| Source: Courtesy of Dr. Gerald Bargman. |                       |                   |                      |

Definimos entonces las variables de la siguiente manera:

X1: Sweat Rate (Tasa de Sudor)

X2: Sodium (Contenido de Sodio)

X3: Potassium (Contenido de Potasio)

En total son 20 observaciones por cada una de las tres variables, adicionalmente no nos indican que la muestra proviene de una distribución Normal 3-variada por lo que usamos el estadístico  $T^2$  de Hotelling el cual se distribuye como una F

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

- Determine los ejes de la elipsoide del 90% de confianza para  $\mu$ . Determine la longitud de los ejes.

Para el caso de  $p = 3$ , los ejes de la región de confianza o Elipse de confianza y sus respectivas longitudes relativas, son determinados a partir de los eigen-valores y eigen-vectores de S. Para los datos de la tabla 5.1 tenemos  $\bar{X}$  y S definida de la siguiente manera:

## [ ,1]

```
## [1,] 4.640
## [2,] 45.400
## [3,] 9.965

##          X1          X2          X3
## X1  2.879368  10.0100 -1.809053
## X2  10.010000 199.7884 -5.640000
## X3  -1.809053  -5.6400  3.627658
```

Para los eigen-valores y eigen-vectores se tiene:

```
## eigen() decomposition
## $values
## [1] 200.462464  4.531591  1.301392
##
## $vectors
##          [,1]          [,2]          [,3]
## [1,] -0.05084144 -0.57370364  0.81748351
## [2,] -0.99828352  0.05302042 -0.02487655
## [3,]  0.02907156  0.81734508  0.57541452
```

Iniciando en el centro  $\bar{X}$ , los ejes del elipsoide de confianza son:

$$\pm \sqrt{\lambda_i} \sqrt{\frac{(n-1)p}{(n-p)n}} F_{\alpha; p, n-p} e_i$$

con  $Se_i = \lambda_i e_i$  para  $i = 1, 2, 3$ .

Para el calculo de las semi-longitudes tenemos:  $\pm \sqrt{\lambda_i} \sqrt{\frac{(n-1)p}{(n-p)n}} F_{\alpha; p, n-p}$

Tenemos entonces para los datos que las semi-longitudes son las siguientes:

```
##          [,1]
## [1,] 9.0506741
## [2,] 1.3607857
## [3,] 0.7292367
```

Para el calculo de los ejes, usaremos los eigen-vectores, teniendo como resultado lo siguiente:

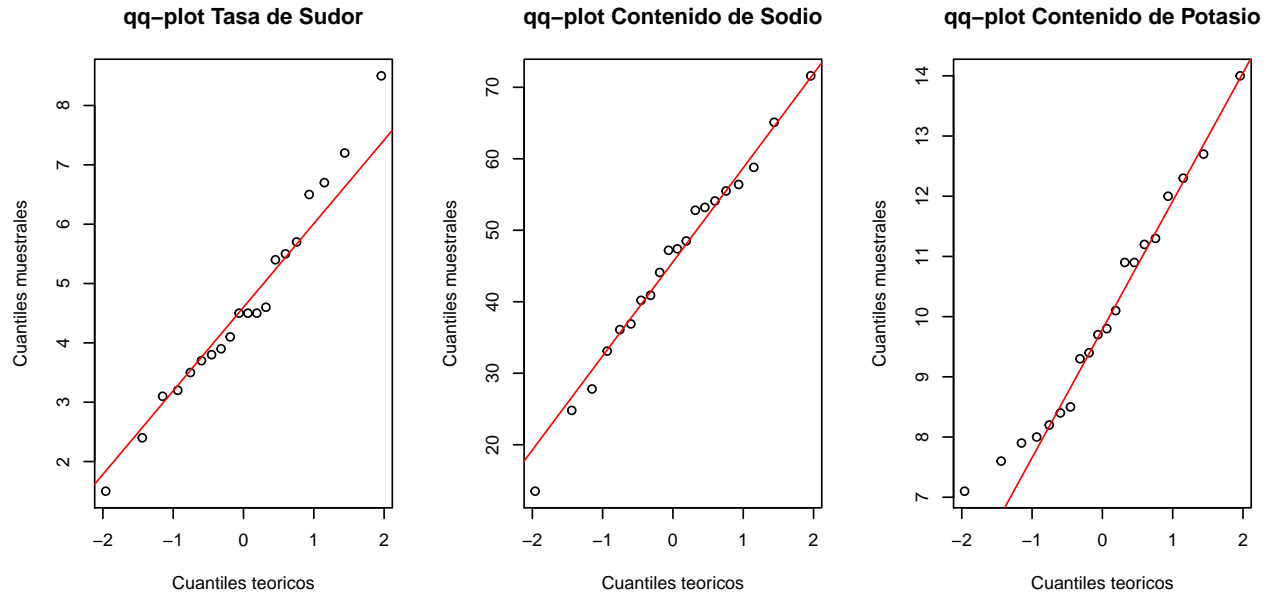
$$\begin{aligned} &\pm(9.0506741) \begin{bmatrix} -0.05084144 \\ 0.99828352 \\ 0.02907156 \end{bmatrix} \\ &\pm(1.3607857) \begin{bmatrix} -0.57370364 \\ 0.05302042 \\ 0.81734508 \end{bmatrix} \\ &\pm(0.7292367) \begin{bmatrix} 0.81748351 \\ -0.02487655 \\ 0.57541452 \end{bmatrix} \end{aligned}$$

Construimos entonces los intervalos de confianza con base en el estadístico  $T^2$  con un nivel de confianza del 90% seria el siguiente:

```
##      N  Media Desv_Estandar   T2_Low   T2_Up
## X1 20  4.640      2.879368  3.555292  5.724708
## X2 20 45.400     199.788421 36.364555 54.435445
## X3 20  9.965      3.627658  8.747476 11.182524
```

- Construya un grafico QQ para las observaciones en tasa de sudor, contenido de sodio y contenido de potasio respectivamente. Construir las tres posibles graficas de dispersión para las parejas de observaciones. Es la suposición de normal multivariada aceptada en este caso?

Se realiza el grafico QQ para cada una de las variables:

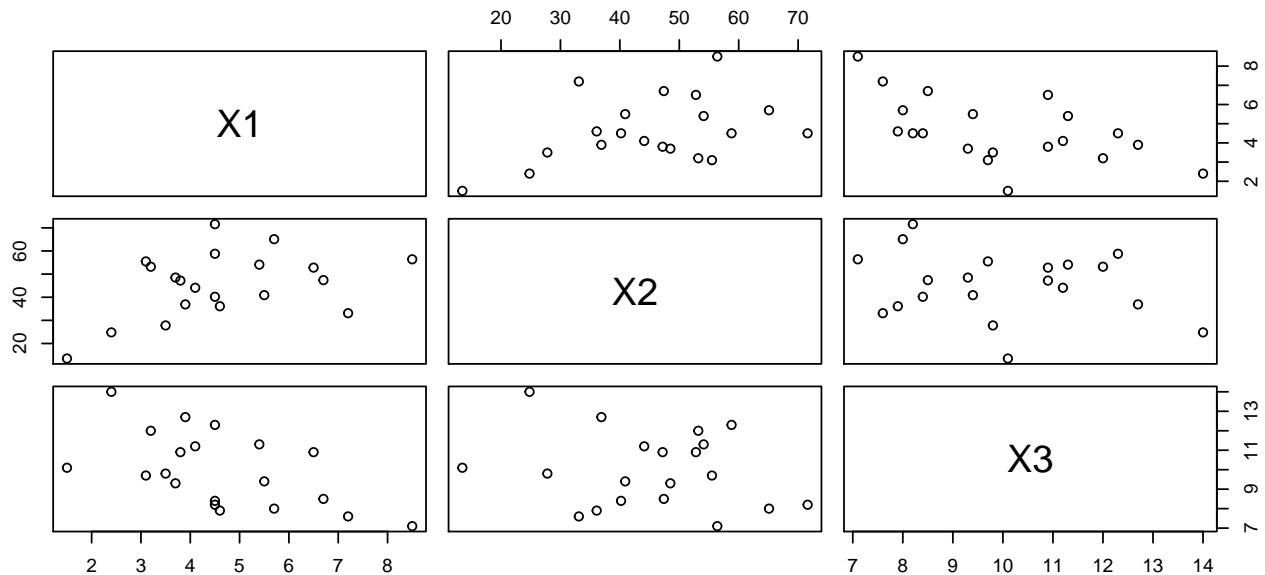


Del gráfico qqplot se observa que X1 y X2 son claramente normales, para X3 observamos que al principio hay un desvío de los quantiles teóricos, sin embargo podemos concluir que son normales pues solo tres observaciones se desvían de los quantiles teóricos.

```
##                               p-valor
## Tasa de Sudor                 0.8689242
## Contenido de Sodio           0.9861883
## Contenido de Potasio         0.6232620
```

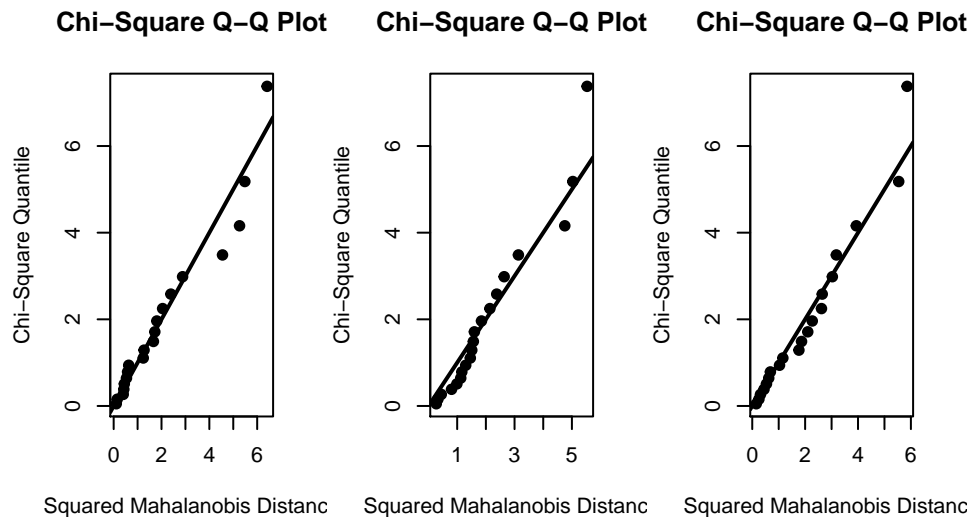
Revisando el resultado de las pruebas formales, se obtiene que los datos son normales ya que tenemos un valor de mayor de 0.5.

Graficos de dispersión:



En los gráficos de dispersión bi-variados se observa que X1-X2 y X1-X3 tienen una clara dispersión elipsoidal lo que nos puede indicar una normalidad bi-variada. Para X2-X3 no es tan claro en la gráfica, pero da la impresión de normalidad, esto lo comprobaremos con los gráficos chi-cuadrado y pruebas de hipótesis de Mardia.

Ahora se realiza el análisis bivariado por las diferentes combinaciones entre las 3 variables:



Se observa que los datos para X1-X2, X1-X3 y X2-X3 siguen claramente la línea de los cuantiles teóricos de una normal bi-variada.

```
##   bivariados p.value.kurt p.value.skew p.value.small
## 1   X1 - X2   0.8133465   0.3263974   0.2070319
## 2   X1 - X3   0.3224551   0.7077270   0.6029315
## 3   X2 - X3   0.4622371   0.8863096   0.8334699
```

De la prueba de Mardia, y los valores p de la kurtosis y la asimetría y con un nivel de confianza del 95%, tenemos que ninguno rechaza la hipótesis nula de que los datos provienen de una distribución normal bi-variada.

Dado que tanto las distribuciones marginales como las distribuciones bi-variadas son normales, podemos concluir que los datos distribuyen normal tres-variado.

## PUNTO 5.9

Harry Roberts, un naturalista para el departamento de Alaska Fish and Game, estudio los osos pardos con la meta de mantener una población saludable. Mediciones en  $n = 61$  de osos proveen el siguiente resumen de estadísticas:

| Variable            | Peso (Kg) | Longitud Cuerpo (cm) | Cuello (cm) | Cintura (cm) | Longitud Cabeza (cm) | Ancho Cabeza (cm) |
|---------------------|-----------|----------------------|-------------|--------------|----------------------|-------------------|
| Media ( $\bar{x}$ ) | 95.52     | 164.38               | 55.69       | 93.39        | 17.98                | 31.13             |

Matriz de Varianzas-Covarianzas S:

|    | X1      | X2      | X3     | X4      | X5     | X6     |
|----|---------|---------|--------|---------|--------|--------|
| X1 | 3266.46 | 1343.97 | 731.54 | 1175.50 | 162.68 | 238.37 |
| X2 | 1343.97 | 721.91  | 324.25 | 537.35  | 80.17  | 117.73 |
| X3 | 731.54  | 324.25  | 179.28 | 281.17  | 39.15  | 56.80  |
| X4 | 1175.50 | 537.35  | 281.17 | 474.98  | 63.73  | 94.85  |
| X5 | 162.68  | 80.17   | 39.15  | 63.73   | 9.95   | 13.88  |
| X6 | 238.37  | 117.73  | 56.80  | 94.85   | 13.88  | 21.26  |

Se pide:

a) **Obtenga intervalos simultáneos del 95% de confianza para muestras grandes para las seis poblaciones de las medias de los cuerpos de los osos.**

El análisis se realiza con base al vector de medias muestrales y a la matriz de varianzas y covarianzas muestrales de las siguientes variables:

- Peso (kg)
- longitud corporal (cm)
- cuello (cm)
- circunferencia (cm)
- longitud de la cabeza (cm)
- ancho de la cabeza (cm).

La siguiente función será utilizada para obtener los intervalos de confianza con una  $\alpha = 0.05$ , como el enunciado indica que es una muestra grande se trabajará con la  $\chi^2$ :

$$\bar{x}_i \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{S_{ii}}{n}}$$

Recordemos que la formula para hallar los intervalos  $T^2$  es la siguiente:

$$\bar{x}_i \pm \sqrt{\frac{(n-1)p}{(n-1)} F_{p, n-p}(\alpha)} \sqrt{\frac{S_{ii}}{n}}$$

Para  $i = 1, 2, \dots, 6$

Donde:

- $\bar{x}_i$ : Media muestral.
- $\mu$ : Media poblacional.
- $S_{ii}$ : Componente de la diagonal de la matriz de varianza y covarianza.

- $n$ : Número de osos.
- $p$ : Número de variables.
- $\alpha = 0.05$ .

Nota: Para hallar el intervalo de confianza simultáneo de las variables, ya se tienen todos los datos y procedemos a calcular los límites superiores e inferiores y se obtiene:

| ##                      | Media  | Varianza | Lim inferior | Lim superior | Longitud  |
|-------------------------|--------|----------|--------------|--------------|-----------|
| ## Peso (Kg)            | 95.52  | 3266.46  | 86.16199     | 104.87801    | 18.716019 |
| ## Longitud Cuerpo (cm) | 164.38 | 721.91   | 159.98067    | 168.77933    | 8.798656  |
| ## Cuello (cm)          | 55.69  | 179.28   | 53.49765     | 57.88235     | 4.384708  |
| ## Cintura (cm)         | 93.39  | 474.98   | 89.82153     | 96.95847     | 7.136947  |
| ## Longitud Cabeza (cm) | 17.98  | 9.95     | 17.46352     | 18.49648     | 1.032967  |
| ## Ancho Cabeza (cm)    | 31.13  | 21.26    | 30.37504     | 31.88496     | 1.509929  |

- Peso (kg): El valor real de la media,  $\mu$ , de la variable estará entre  $86.16kg$  y  $104.88kg$  con una probabilidad del 95%, siendo  $86.16kg$  y  $104.88kg$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 95.52$  y a la varianza  $S_{ii} = 3266.46$ .
- Longitud Cuerpo (cm): El valor real de la media,  $\mu$ , de la variable estará entre  $159.98cm$  y  $168.78cm$  con una probabilidad del 95%, siendo  $159.98cm$  y  $168.78cm$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 164.38$  y a la varianza  $S_{ii} = 721.91$ .
- Cuello (cm): El valor real de la media,  $\mu$ , de la variable estará entre  $53.49cm$  y  $57.88cm$  con una probabilidad del 95%, siendo  $53.49cm$  y  $57.88cm$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 55.69$  y a la varianza  $S_{ii} = 179.28$ .
- Cintura (cm): El valor real de la media,  $\mu$ , de la variable estará entre  $89.82cm$  y  $96.96cm$  con una probabilidad del 95%, siendo  $89.82cm$  y  $96.96cm$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 93.39$  y a la varianza  $S_{ii} = 474.98$ .
- Longitud Cabeza (cm): El valor real de la media,  $\mu$ , de la variable estará entre  $17.46cm$  y  $18.49cm$  con una probabilidad del 95%, siendo  $17.46cm$  y  $18.49cm$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 17.98$  y a la varianza  $S_{ii} = 9.95$ .
- Ancho Cabeza (cm): El valor real de la media,  $\mu$ , de la variable estará entre  $30.38cm$  y  $31.88cm$  con una probabilidad del 95%, siendo  $30.38cm$  y  $31.88cm$  el intervalo de confianza arrojado después de calcular el límite inferior y superior con base a  $\bar{x}_i = 31.13$  y a la varianza  $S_{ii} = 21.26$ .

Comparando el resultado del intervalo con el calculo de la  $T^2$  para muestras pequeñas, el resultado obtenido es el siguiente:

| ##                      | Media  | Varianza | Lim inferior | Lim superior | Longitud  |
|-------------------------|--------|----------|--------------|--------------|-----------|
| ## Peso (Kg)            | 95.52  | 3266.46  | 67.32099     | 123.71901    | 56.398021 |
| ## Longitud Cuerpo (cm) | 164.38 | 721.91   | 151.12326    | 177.63674    | 26.513479 |
| ## Cuello (cm)          | 55.69  | 179.28   | 49.08366     | 62.29634     | 13.212686 |
| ## Cintura (cm)         | 93.39  | 474.98   | 82.63692     | 104.14308    | 21.506159 |
| ## Longitud Cabeza (cm) | 17.98  | 9.95     | 16.42365     | 19.53635     | 3.112696  |
| ## Ancho Cabeza (cm)    | 31.13  | 21.26    | 28.85502     | 33.40498     | 4.549952  |

Las longitudes de los rangos de la  $T^2$  son muy grandes comparadas con las intervalos calculados con la  $\chi^2$

#### b) Obtenga una elipse de confianza simultanea del 95% para el peso medio y la cintura media para muestras grandes.

La siguiente ecuación es la fórmula general de una elipse de confianza simultanea para muestras grandes con un nivel del 95% para el peso medio  $\mu_1$  y la cintura media  $\mu_4$  tomada de una muestra de  $n = 61$  osos pardos.

$$n \begin{bmatrix} \bar{x}_1 - \mu_1 \\ \bar{x}_4 - \mu_4 \end{bmatrix}^t \begin{bmatrix} s_{11} & s_{14} \\ s_{14} & s_{44} \end{bmatrix}^{-1} \begin{bmatrix} \bar{x}_1 - \mu_1 \\ \bar{x}_4 - \mu_4 \end{bmatrix} \leq \chi_p^2(\alpha)$$

Se tiene entonces:

$$61 \begin{bmatrix} 95.52 - \mu_1 \\ 93.39 - \mu_4 \end{bmatrix}^t \begin{bmatrix} 3266.46 & 1175.50 \\ 1175.50 & 474.98 \end{bmatrix}^{-1} \begin{bmatrix} 95.52 - \mu_1 \\ 93.39 - \mu_4 \end{bmatrix} \leq \chi_6^2(0.05) = 1.635383$$

Donde la matriz de varianza y covarianza para las variables peso y cintura tomadas de los 61 datos de los osos pardos está dada por:

$$S^{(1)} = \begin{bmatrix} 3266.46 & 1175.50 \\ 1175.50 & 474.98 \end{bmatrix} \setminus$$

Teniendo la  $S^{(1)}$  se hallan los eigen-valores y eigen-vectores para poder realizar el calculo de la elipse, se obtiene entonces:

```
## eigen() decomposition
## $values
## [1] 3695.51873    45.92127
##
## $vectors
##           [,1]      [,2]
## [1,] -0.9393810  0.3428751
## [2,] -0.3428751 -0.9393810
```

Así los valores propios son:

$\lambda_1 = 3695.52$  y  $\lambda_2 = 45.92$ , y tenemos como vectores propios por:

$$\underline{e}_1 = \begin{bmatrix} -0.94 \\ -0.34 \end{bmatrix}$$

y

$$\underline{e}_2 = \begin{bmatrix} 0.34 \\ -0.94 \end{bmatrix}$$

Las semi-longitudes estan dadas por:

$$\underline{x}_i \pm \sqrt{\lambda_i} \sqrt{\chi_p^2(\alpha)}$$

Teniendo como resultado de las semi-longitudes:

```
##           [,1]
## [1,] 77.74052
## [2,]  8.66596
```

Ahora una indicación de la elongación de la elipse de confianza está dada por el cociente entre la longitud del eje mayor y longitud del eje menor, en este caso es:

```
## [1] 8.970791
```

Ahora los ejes del elipsoide de confianza del 95% son:

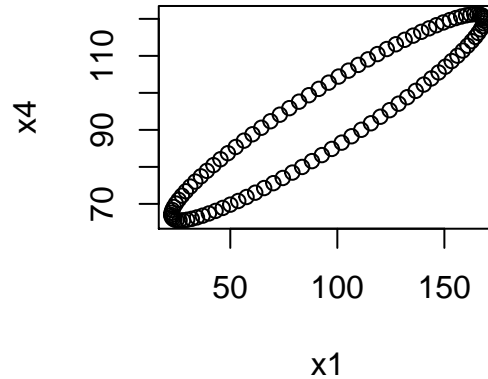
$$\pm \sqrt{3695.52} \sqrt{\chi_6^2(0.05)} \times \begin{bmatrix} -0.939 \\ -0.343 \end{bmatrix}$$

y

$$\pm \sqrt{45.92} \sqrt{\chi_6^2(0.05)} \times \begin{bmatrix} 0.342 \\ -0.939 \end{bmatrix} \setminus$$

donde  $\chi_6^2(0.05) = 1.635383$ .

Utilizando librería “ellipse” de *R* se procedió a calcular la elipse de confianza del 95% para las variables peso y cintura de los osos pardos.



La figura representa una elipse de confianza simultánea del 95% para las variables peso (kg) y cintura (cm) media de una muestra de 61 osos pardos.

**c) Obtenga un intervalo de confianza Bonferroni del 95% para las seis medias de la parte (a)**

En este punto a pesar de que se indica que es una muestra grande, se realizara el analisis de la Bonferroni con la distribución *t* ya que hacerlo con la  $\chi$  nos daría los mismos valores del punto a), se hallarán los intervalos de confianza simultáneos de Bonferroni de 95% para los valores de de las  $x_i$ , los valores de los límites inferiores y superiores de cada variable, deben ser mayores y menores, respectivamente, en Bonferroni respecto a los hallados en punto a).

La ecuación para calcular los intervalos de confianza de Bonferroni es:

$$\bar{x}_i \pm t_{\frac{\alpha}{2p}, n-1} \sqrt{\frac{S_{ii}}{n}}$$

Para  $i = 1, 2, \dots, 6$

A continuación se observa que los intervalos de confianza de Bonferroni hallados, efectivamente se encuentran entre los valores de los intervalos de confianza del punto a):

| Variable             | Límite Inferior | Límite Superior |
|----------------------|-----------------|-----------------|
| Peso (Kg)            | 75.55           | 115.49          |
| Longitud Cuerpo (cm) | 154.99          | 173.77          |
| Cuello (cm)          | 51.01           | 60.37           |
| Cintura (cm)         | 85.78           | 101.00          |
| Longitud Cabeza (cm) | 16.88           | 19.08           |
| Ancho Cabeza (cm)    | 29.52           | 32.74           |

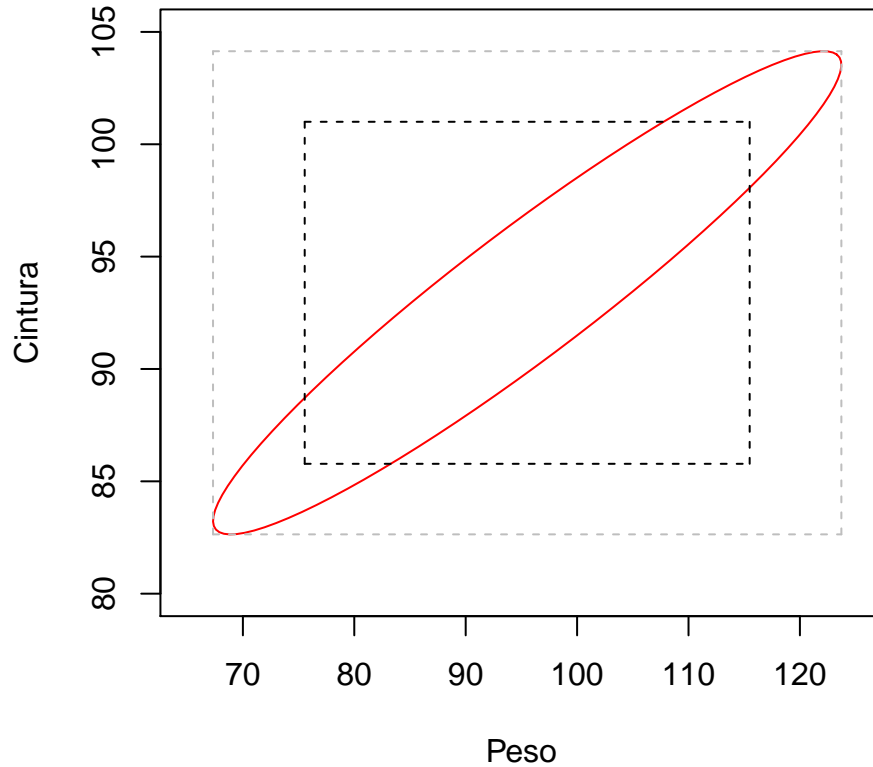
**d) Consulte la parte (b) construya un rectángulo de confianza del 95% para el peso medio y la cintura media usando  $m = 6$ , compare esté de confianza con la elipse de confianza en la parte (b)**

Para construir un rectángulo de Bonferroni del 95% simultaneo para las variables peso y cintura de los osos pardos y compararlo con la elipse de confianza simultánea del 95% de la parte (b) se utiliza un código en el programa estadístico R.

Para este código se utilizan los intervalos de confianzas de Bonferroni del inciso (c) tomando  $t_{\frac{\alpha}{2m}}$  con  $m = 6$



### Intervalo de confianza del 95% para tamaño de muestra de n



La matriz de varianza y covarianza para las variables peso y cintura de los osos pardos tiene una correlación muy grande de 1175.68, esta correlación es positiva, por eso observamos que la inclinación de la gráfica para elipse de confianza esta inclinada hacia la derecha.

Notemos que los intervalos de confianza individuales de Bonferroni calculados para cada una de las seis variables en la parte (c) están contenidos en los intervalos de confianza individuales calculados en la parte (b) para cada una de las seis variables es decir el intervalo de confianza para la variable peso para un intervalo de Bonferroni es (75.55, 115.49) y está contenido en el intervalo (67.32, 123.72) por lo cual los de Bonferroni individuales tienen menor confianza, pero más precisos de igual forma el resto de intervalos para cada variable.

**e) Obtenga un intervalo de confianza del 95% de Bonferroni para Ancho medio de la cabeza - longitud media de la cabeza. El uso  $m = 6 + 1 = 7$  permite esta afirmación, así como la declaración sobre cada media individual.**

Para hallar el intervalo de confianza Bonferroni simultaneo del 95% para la diferencia entre el ancho medio de la cabeza ( $\mu_5$ ) y la longitud media de la cabeza ( $\mu_6$ ), la diferencia de medias  $\mu_6 - \mu_5$  está dentro del siguiente intervalo con  $m = 7$

$$\left( \bar{x}_6 - \bar{x}_5 - t_{\frac{\alpha}{2m}, n-1} \sqrt{\frac{S_{66} - 2S_{65} + S_{55}}{61}}; \bar{x}_6 - \bar{x}_5 + t_{\frac{\alpha}{2m}, n-1} \sqrt{\frac{S_{66} - 2S_{65} + S_{55}}{61}} \right)$$

Sustituyendo los valores en la fórmula de arriba tenemos:

$$31.13 - 17.98 \pm t_{\frac{0.05}{14}}(60) \times \sqrt{\frac{21.26 - 2(13.88) + 9.95}{61}}$$

Utilizando en programa R se calcula  $t_{\frac{0.05}{14}}(60) = t_{0.0036}(60) = 2.783$

Haciendo las operaciones pertinentes se llega a que

$$12.49 \leq \mu_6 - \mu_5 \leq 13.81$$

De donde se puede concluir que el intervalo de confianza simultaneo de Bonferroni del 95% para las dos variables de la diferencia media entre el ancho de la cabeza y la longitud de la cabeza se encuentra entre 15.49cm y 13.81cm

## PUNTO 5.21

Usando los datos del contenido mineral de los huesos de la Tabla 1.8, construya el intervalo de Bonferroni al 95% para las medias individuales. Tambien encuentre el  $T^2$ -Intervalo simultaneo, compare los dos intervalos hallados.

| <b>Table 1.8 Mineral Content in Bones</b> |                 |        |                  |         |               |      |
|---|-----------------|--------|------------------|---------|---------------|------|
| Subject number                            | Dominant radius | Radius | Dominant humerus | Humerus | Dominant ulna | Ulna |
| 1   | 1.103           | 1.052  | 2.139            | 2.238   | .873          | .872 |
| 2   | .842            | .859   | 1.873            | 1.741   | .590          | .744 |
| 3   | .925            | .873   | 1.887            | 1.809   | .767          | .713 |
| 4   | .857            | .744   | 1.739            | 1.547   | .706          | .674 |
| 5   | .795            | .809   | 1.734            | 1.715   | .549          | .654 |
| 6   | .787            | .779   | 1.509            | 1.474   | .782          | .571 |
| 7   | .933            | .880   | 1.695            | 1.656   | .737          | .803 |
| 8   | .799            | .851   | 1.740            | 1.777   | .618          | .682 |
| 9   | .945            | .876   | 1.811            | 1.759   | .853          | .777 |
| 10  | .921            | .906   | 1.954            | 2.009   | .823          | .765 |
| 11  | .792            | .825   | 1.624            | 1.657   | .686          | .668 |
| 12  | .815            | .751   | 2.204            | 1.846   | .678          | .546 |
| 13  | .755            | .724   | 1.508            | 1.458   | .662          | .595 |
| 14  | .880            | .866   | 1.786            | 1.811   | .810          | .819 |
| 15  | .900            | .838   | 1.902            | 1.606   | .723          | .677 |
| 16  | .764            | .757   | 1.743            | 1.794   | .586          | .541 |
| 17  | .733            | .748   | 1.863            | 1.869   | .672          | .752 |
| 18  | .932            | .898   | 2.028            | 2.032   | .836          | .805 |
| 19  | .856            | .786   | 1.390            | 1.324   | .578          | .610 |
| 20  | .890            | .950   | 2.187            | 2.087   | .758          | .718 |
| 21  | .688            | .532   | 1.650            | 1.378   | .533          | .482 |
| 22  | .940            | .850   | 2.334            | 2.225   | .757          | .731 |
| 23  | .493            | .616   | 1.037            | 1.268   | .546          | .615 |
| 24  | .835            | .752   | 1.509            | 1.422   | .618          | .664 |
| 25  | .915            | .936   | 1.971            | 1.869   | .869          | .868 |

Source: Data courtesy of Everett Smith.

Definimos entonces las variables de las siguiente manera:

- X1: Dominant Radius
- X2: Radius
- X3: Dominant Humerus
- X4: Humerus
- X5: Dominant ulna
- X6: Ulna

Para hallar los intervalos de Bonferroni y los  $T^2$  se calcula  $\bar{X}$  y S, para Bonferroni los intervalos estan dados por:

$$\bar{x}_i \pm t_{\frac{\alpha}{2p}, n-1} \sqrt{\frac{S_{ii}}{n}}$$

Para  $T^2$  los intervalos estan dados por:

$$\bar{x}_i \pm \sqrt{\frac{(n-1)p}{(n-1)} F_{p,n-p}(\alpha)} \sqrt{\frac{S_{ii}}{n}}$$

Así que para los datos del ejercicio los intervalos de Bonferroni con un nivel de confianza del 95% serian:

```
## [1] "Vector de Medias"

##      Media estimada
## X1      0.84380
## X2      0.81832
## X3      1.79268
## X4      1.73484
## X5      0.70440
## X6      0.69384

## [1] "Matriz de Varianzas-Covarianzas"

##           X1           X2           X3           X4           X5           X6
## X1 0.013001583 0.010378442 0.02234997 0.02008568 0.009120708 0.007957842
## X2 0.010378442 0.011417893 0.01853519 0.02109951 0.008529783 0.008908512
## X3 0.022349975 0.018535190 0.08035723 0.06677620 0.016836925 0.012847030
## X4 0.020085675 0.021099512 0.06677620 0.06948447 0.017735483 0.016793598
## X5 0.009120708 0.008529783 0.01683692 0.01773548 0.011568417 0.008071150
## X6 0.007957842 0.008908512 0.01284703 0.01679360 0.008071150 0.010599140

## [1] "Elementos de la diagonal de la inversa de la Matriz de Varianzas-Covarianzas"

##           X1           X2           X3           X4           X5           X6
## 0.01300158 0.01141789 0.08035723 0.06948447 0.01156842 0.01059914

## [1] "Estadístico de prueba de la distribución T"

## [1] 2.875094

## [1] "Intervalos Bonferroni al 95% de confianza"

##      Media estimada Media estimada
## X1      0.7782338      0.9093662
## X2      0.7568766      0.8797634
## X3      1.6296774      1.9556826
## X4      1.5832656      1.8864144
## X5      0.6425529      0.7662471
## X6      0.6346406      0.7530394
```

A continuación se calculan los mismos intervalos pero con el método de  $T^2$  fijando el  $\alpha$  al 95%.

```
## [1] 4.46317

## [1] "Intervalos de T^2 al 95% de confianza"

##      Media estimada Media estimada
## X1      0.7420179      0.9455821
## X2      0.7229380      0.9137020
## X3      1.5396419      2.0457181
## X4      1.4995425      1.9701375
## X5      0.6083914      0.8004086
## X6      0.6019414      0.7857386
```

Tenemos entonces como resumen de los resultados hallados:

```
## [1] "Resultados consolidados"
```

| ##    | Media.estimada | Desv_Estandar | T2_Low    | T2_Up     | Longitud_T2 | B_Low     |
|-------|----------------|---------------|-----------|-----------|-------------|-----------|
| ## X1 | 0.84380        | 0.01300158    | 0.7420179 | 0.9455821 | 0.2035643   | 0.7782338 |
| ## X2 | 0.81832        | 0.01141789    | 0.7229380 | 0.9137020 | 0.1907640   | 0.7568766 |
| ## X3 | 1.79268        | 0.08035723    | 1.5396419 | 2.0457181 | 0.5060761   | 1.6296774 |
| ## X4 | 1.73484        | 0.06948447    | 1.4995425 | 1.9701375 | 0.4705950   | 1.5832656 |
| ## X5 | 0.70440        | 0.01156842    | 0.6083914 | 0.8004086 | 0.1920173   | 0.6425529 |
| ## X6 | 0.69384        | 0.01059914    | 0.6019414 | 0.7857386 | 0.1837971   | 0.6346406 |

| ##    | B_Up      | Longitud_B |
|-------|-----------|------------|
| ## X1 | 0.9093662 | 0.1311325  |
| ## X2 | 0.8797634 | 0.1228868  |
| ## X3 | 1.9556826 | 0.3260052  |
| ## X4 | 1.8864144 | 0.3031489  |
| ## X5 | 0.7662471 | 0.1236941  |
| ## X6 | 0.7530394 | 0.1183988  |

Se concluye que con un nivel de confianza las medias de las poblaciones se encuentren en los intervalos calculados, siendo los de Bonferroni más angostos (precisos).

Se observa que los intervalos de Bonferroni siempre son más precisos que los intervalos producidos por el método  $T^2$ , siendo los de *Bonferroni* más angostos lo que es una característica mejor.