

Assignment 1

Emre Beray Boztepe

19 03 2024

Problem 1

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

1-) The rule for a symmetric matrix is that its transpose is equal to the original matrix. Since $A^T = A$ holds true, A is symmetric.

2-) After solving $\det(A - \lambda I) = 0$, we find the eigenvalues to be $\lambda_1 = 2$ and $\lambda_2 = 4$, along with their corresponding eigenvectors $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Thus, the matrix can be decomposed as $P = \frac{1}{\sqrt{2}}[v_1 \ v_2]$, and $\Lambda = \text{diag}(\lambda_1, \lambda_2)$.

3-)

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

the first matrix corresponds to the eigenvector e_1 scaled by λ_1 and the second matrix corresponds to the eigenvector e_2 scaled by λ_2 .

$$4-) \sqrt{A} = \sqrt{\lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T}$$

The property: $(XY)^{1/2} = X^{1/2}Y^{1/2}$

$$\sqrt{\lambda_1 e_1 e_1^T} = \sqrt{\lambda_1} e_1 e_1^T$$

$$\sqrt{\lambda_2 e_2 e_2^T} = \sqrt{\lambda_2} e_2 e_2^T$$

$$\sqrt{A} = \frac{1}{2} \begin{bmatrix} \sqrt{2} + 2 & \sqrt{2} - 2 \\ \sqrt{2} - 2 & \sqrt{2} + 2 \end{bmatrix}$$

This matrix satisfies $\sqrt{A}\sqrt{A} = A$, verifying the result.

Problem 2

Consider the spectral decomposition of a positive definite matrix as given in Lecture 1:

$$A = P\Lambda P^T$$

The columns of P are made of eigenvectors $e_i, i = 1, \dots, n$ and they are orthonormalized, i.e. their lengths are one and they are orthogonal (perpendicular) one to another. The

diagonal matrix Λ has the corresponding (positive) eigenvalues on the diagonal. Provide argument for the following

$$1-) P^T = P^{-1}$$

Since, $e_i^T e_i = 1$ and $e_i^T e_j = 0$ for $i \neq j$

$$P^T P = \begin{bmatrix} e_1^T \\ - \\ \dots \\ - \\ e_n^T \end{bmatrix} [e_1 \mid \dots \mid e_n]$$

$$= \begin{bmatrix} e_1^T e_1 & \dots & e_1^T e_n \\ \dots & \dots & \dots \\ e_n^T e_1 & \dots & e_n^T e_n \end{bmatrix} = I$$

Each column of P represents an eigenvector, and the dot product of any two distinct eigenvectors is zero, while the dot product of an eigenvector with itself is one

2-) We can establish this by mathematical induction. Let's Assume that for an $n \times n$ matrix Λ_n , $\det(\Lambda_n) = \prod_{i=1}^n \lambda_i$. Now, we aim to demonstrate that this holds true for an $(n+1) \times (n+1)$ matrix as well. To do so, we'll utilize Laplace expansion on the $(n+1)$ -th row:

$$\det(\Lambda_{n+1}) = \det(\Lambda_n) * \lambda_{n+1} + 0 + \dots + 0$$

Thus, by our inductive hypothesis, $\det(\Lambda_{n+1}) = \prod_{i=1}^{n+1} \lambda_i$. This confirms the validity of the proposition for matrices of size $(n+1) \times (n+1)$

3-) Utilizing property 2 and the condition $\det(P) = \det(P^T) = \pm 1$ we can deduce:

$$\det(A) = \det(P \Lambda P^T) = \det(P) \det(\Lambda) \det(P^T) = \det(\Lambda)$$

the determinant of a positive definite matrix A equals the determinant of its diagonal matrix Λ . This implies that the determinant of A is the product of its eigenvalues.

4-) Let's assume that $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ then:

$$\Lambda^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1})$$

This process simplifies finding the inverse of diagonal matrices, requiring only the computation of reciprocals without complex matrix operations.

5-) Demonstrating that $A A^{-1} = I$ is sufficient:

$$A A^{-1} = (P \Lambda P^T) (P \Lambda^{-1} P^T) = P \Lambda \Lambda^{-1} P^T = P P^T = I$$

6-) Demonstration in R

```

# Print eigenvalues and eigenvectors of matrix A
A = matrix(c(3, -1, -1, 3), nrow = 2)
eigen_A = eigen(A)

# Calculate P * P^T
cat("P * P^T =\n")

## P * P^T =

print(round(eigen_A$vectors %*% t(eigen_A$vectors), 16))

##      [,1] [,2]
## [1,]    1    0
## [2,]    0    1

# Calculate determinant of A and product of eigenvalues
cat("\ndet(A) = ", det(A), "\n")

##
## det(A) = 8

cat("Π(λi) = ", prod(eigen_A$values), "\n\n")

## Π(λi) = 8

# Calculate determinant of Λ
cat("det(Λ) = ", det(diag(eigen_A$values)))

## det(Λ) = 8

# Test the inverse
cat("\nTest the inverse:\n")

##
## Test the inverse:

print(round(A %*% eigen_A$vectors %*% diag(1 / eigen_A$values) %*%
t(eigen_A$vectors), 16))

##      [,1] [,2]
## [1,]    1    0
## [2,]    0    1

```

Problem 3

In a medical study, length L and weight W of newborn children is considered. It was assumed that (L, W) will be modeled through a bivariate normal distribution. The following information has been known: the mean weight is 3343[g], with the standard deviation of 528[g], while the mean length is 49.8[cm], with the standard deviation of 2.5[cm]. Additionally the correlation between the length and the weight has been established and equal to 0.75. The joint distribution of (W, L) is bivariate normal, i.e. $(W, L) \sim N(\mu, \Sigma)$. Perform the following tasks and answer the questions:

1-) Covariance matrix (Σ) =

$$\begin{bmatrix} \sigma_W^2 & P\sigma_W\sigma_L \\ P\sigma_W\sigma_L & \sigma_L^2 \end{bmatrix}$$

Parameter $\mu = [3343, 49.8]$

Parameter

$$\begin{aligned} \Sigma &= \begin{bmatrix} 528^2 & 0.75 \cdot 528 & 2.5 \\ 0.75 \cdot 528 & (2.5)^2 & 2.5 \\ 2.5 & 2.5 & 2.5 \end{bmatrix} \\ &= \begin{bmatrix} 278794 & 990 \\ 990 & 6.25 \end{bmatrix} \end{aligned}$$

$$2-) f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

where:

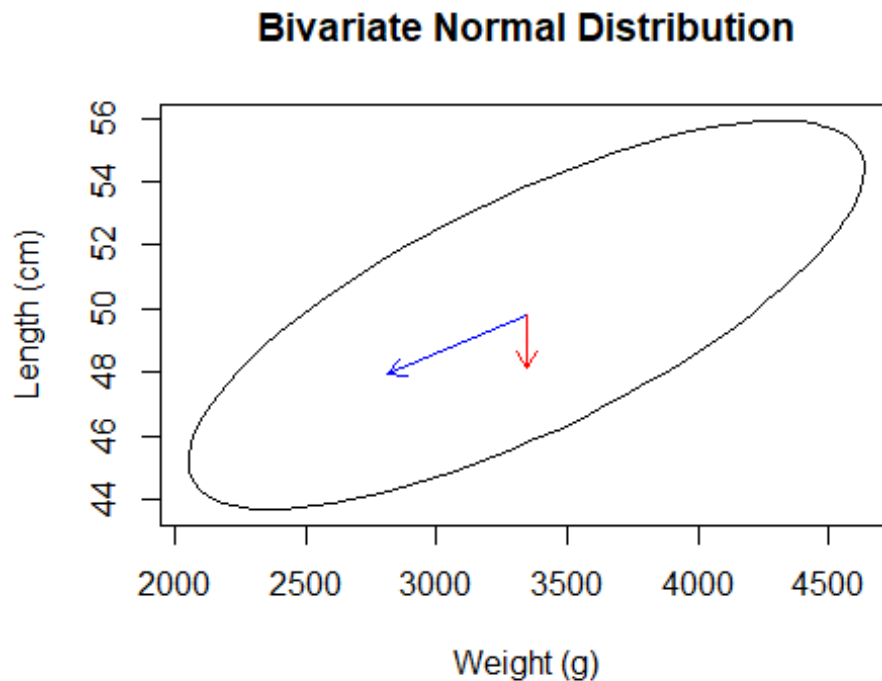
- $|\Sigma|$ = determinant of covariance matrix
- \mathbf{x} = vector of random variables
- $\boldsymbol{\mu}$ = mean vector

3-) Eigenvalues and eigenvectors:

```
## Eigen Values: 278787.5 2.734341
```

```
## Eigen Vectors: -0.9999937 -0.003551149 0.003551149 -0.9999937
```

Plotting the ellipse and adding eigenvectors



4-) For bivariate normal distribution, we have $\mu_w, \mu_l, \sigma_w, \sigma_l, \rho$; for p-dimensional normal distribution:

- p means: $\mu_1, \mu_2, \dots, \mu_p$
- p standard deviations: $\sigma_1, \sigma_2, \dots, \sigma_p$
- $\frac{p(p-1)}{2}$ correlation coefficients (covariance matrix which has $\frac{p(p+1)}{2}$ unique elements)

So, for a p-dimensional normal distribution, we'd have total: $2p + \frac{p(p-1)}{2}$ parameters.

5-) The distribution of L is a normal distribution. The parameters of the normal distribution are its mean $\mu_L = 49.8$ and its standard deviation $\sigma_L = 2.5$

Therefore, the distribution of L can be denoted as $L \sim N(49.8, 2.5)$ This information characterizes the distribution of L, providing its name (normal distribution) and its parameters (mean and standard deviation).

6-) The 3- σ rule states that approximately 99.7% of the data lies within three standard deviations of the mean in a normal distribution. The best guess for the length of the newborn child would be the mean of the distribution $\mu = 49.8$

To find lower and upper bounds = $\mu_L \quad (+ -) \quad 3 \times \sigma_L$

- The upper bound would be $49.8 + 3 \times 2.5 = 57.3$
- The lower bound would be $49.8 - 3 \times 2.5 = 42.3$

So:

- Approximately 68% of the data falls within one standard deviation (σ) of the mean.
- Approximately 95% of the data falls within two standard deviations (2σ) of the mean.
- Approximately 99.7% of the data falls within three standard deviations (3σ) of the mean.

Problem 4

In the setup of the previous problem, assume that it was reported by the mother of the child that weight was 4025[g].

1-) The conditional distribution of L given $W=w$ in a bivariate normal distribution can be represented as:

$$L|W = w \sim \mathcal{N}\left(\mu_L + \rho \frac{\sigma_L}{\sigma_W}(w - \mu_W), \sigma_L^2(1 - \rho^2)\right)$$

- μ_L and μ_W are the mean lengths and mean weights
- σ_L and σ_W are the standard deviations of lengths and weights
- w is the specific value of the weight (which is 4025g for this question)

when we apply the formula:

```
## mean = 52.22187 var = 2.734375
```

```
## standard dev = 1.653595
```

2-) $L \sim N(52.22187, 2.734375)$. The best guess for the length of the newborn child would be the mean of the distribution $\mu = 52.22187$

- The upper bound would be $52.22 + 3 \times 1.65 = 57.18044$
- The lower bound would be $52.22 - 3 \times 1.65 = 47.2633$

3-) With a 68% confidence level, we can expect the prediction error to fall within approximately 1.65 cm, representing one standard deviation. Notably, this standard deviation is smaller compared to our previous estimate, indicating a higher level of accuracy in our prediction.

Problem 5

Let X_1, X_2 , and X_3 be independent $N(\mu, \Sigma)$ random vectors of a dimension p .

1-) Given that X_1, X_2, X_3 are independent multivariate normal random vectors, and v_1 and v_2 are linear combinations of them:

$$v_1 = \frac{1}{4}x_1 - \frac{1}{2}x_2 + \frac{1}{4}x_3$$

$$\begin{aligned}
E(v_1) &= \frac{1}{4}E(x_1) - \frac{1}{2}E(x_2) + \frac{1}{4}E(x_3) \\
&= \frac{1}{4}\mu - \frac{1}{2}\mu + \frac{1}{4}\mu = 0 \\
\text{var}(v_1) &= \text{var}\left(\frac{1}{4}x_1 - \frac{1}{2}x_2 + \frac{1}{4}x_3\right) \\
&= \frac{1}{4^2}\text{var}(x_1) + \frac{1}{2^2}\text{var}(x_2) + \frac{1}{4^2}\text{var}(x_3) \\
&= \frac{3}{8}\Sigma \\
V_1 &\sim N\left(0, \frac{3}{8}\Sigma\right)
\end{aligned}$$

- Same for v_2 :

$$\begin{aligned}
v_2 &= \frac{1}{4}x_1 - \frac{1}{2}x_2 - \frac{1}{4}x_3 \\
E(v_2) &= \frac{1}{4}E(x_1) - \frac{1}{2}E(x_2) - \frac{1}{4}E(x_3) \\
&= -\frac{1}{2}\mu \\
\text{var}(v_2) &= \text{var}\left(\frac{1}{4}x_1 - \frac{1}{2}x_2 - \frac{1}{4}x_3\right) \\
&= \frac{1}{4^2}\text{var}(x_1) + \frac{1}{2^2}\text{var}(x_2) + \frac{1}{4^2}\text{var}(x_3) \\
&= \frac{3}{8}\Sigma \\
V_2 &\sim N\left(-\frac{1}{2}\mu, \frac{3}{8}\Sigma\right)
\end{aligned}$$

2-) Calculating the covariance between the linear combinations. $X \sim N(\mu, \Sigma)$ $AX \sim N(A\mu, A\Sigma A')$

$$V = AX$$

$$V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}I & -\frac{1}{2}I & \frac{1}{4}I \\ \frac{1}{4}I & -\frac{1}{2}I & -\frac{1}{4}I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$E(V) = A_{\mu} = \begin{bmatrix} \frac{1}{4}I & -\frac{1}{2}I & \frac{1}{4}I \\ \frac{1}{4}I & -\frac{1}{2}I & -\frac{1}{4}I \end{bmatrix} \begin{bmatrix} \mu \\ \mu \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2}\mu \end{bmatrix}$$

$$var(v) = A\Sigma A' = \begin{bmatrix} \frac{1}{4}I & -\frac{1}{2}I & \frac{1}{4}I \\ \frac{1}{4}I & -\frac{1}{2}I & -\frac{1}{4}I \end{bmatrix} \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix} \begin{bmatrix} \frac{1}{4}I & \frac{1}{4}I \\ -\frac{1}{2}I & -\frac{1}{2}I \\ \frac{1}{4}I & -\frac{1}{4}I \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{8}\Sigma & \frac{1}{4}\Sigma \\ \frac{1}{4}\Sigma & \frac{3}{8}\Sigma \end{bmatrix}$$

Joint Distribution of

$$v_1, v_2 \sim N\left(\begin{bmatrix} 0 \\ -\frac{1}{2}\mu \end{bmatrix}, \begin{bmatrix} \frac{3}{8}\Sigma & \frac{1}{4}\Sigma \\ \frac{1}{4}\Sigma & \frac{3}{8}\Sigma \end{bmatrix}\right)$$

Project 1: Weight and length of newborn children

Health services and health insurance companies are interested in determining what kind of medical examinations and diagnostic procedures should be administered to a newborn child. In one approach, there is a score system based on which it is determined when a child is healthy and does not require any special attention or when he/she is not in which case a series additional medical tests are performed.

Weight and length of a newborn child are most standard indicators of the health of a child. In order to decide on the score the following procedure is considered. If the weight and length fall outside 95% of the typical values for the population, the score of zero is given. If the measurements are falling in the category between 75% and 95% the score is one. In all other cases the score of two is assigned.

A random sample of records for 736 recently born children (singleton and not prematurely born) has been considered from hospital across a certain region. The records contain a large variety of information but extraction of weight and height data are given in the file WeightHeight.txt.

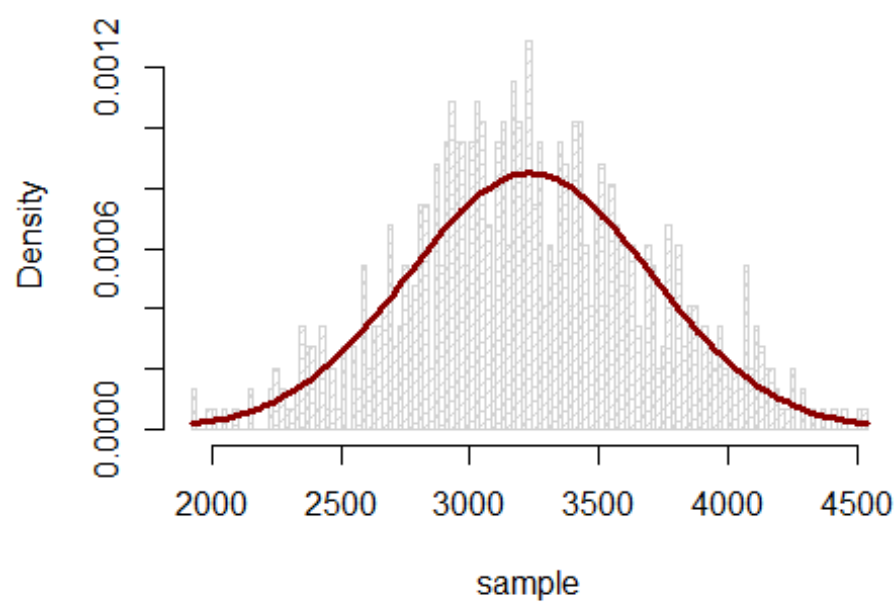
Part One

1-) Using the data estimate the mean and the covariance for the length and the weight of children.

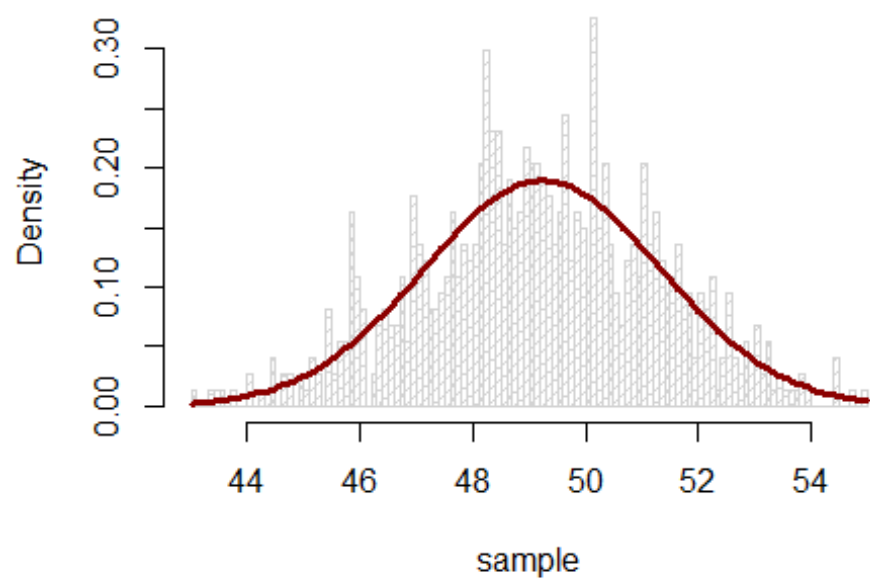

```
## mean_weight: 3233.545  
## mean_length: 49.23764  
## covariance: 915.2955
```

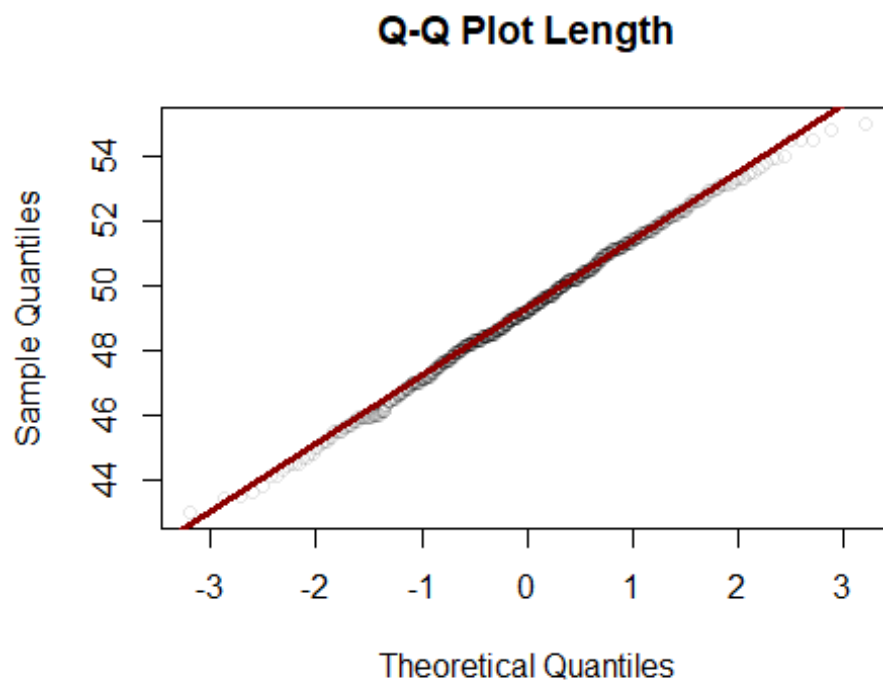
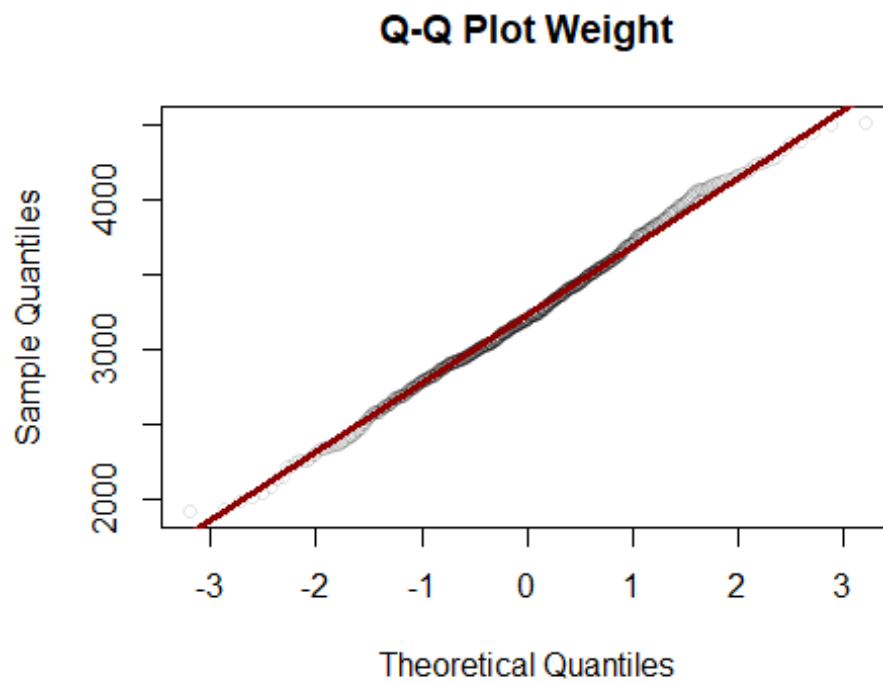
2-) Verify graphically normal distribution of the data. Use a scatterplot and qq-plots for the marginal distributions.

Histogram Weight

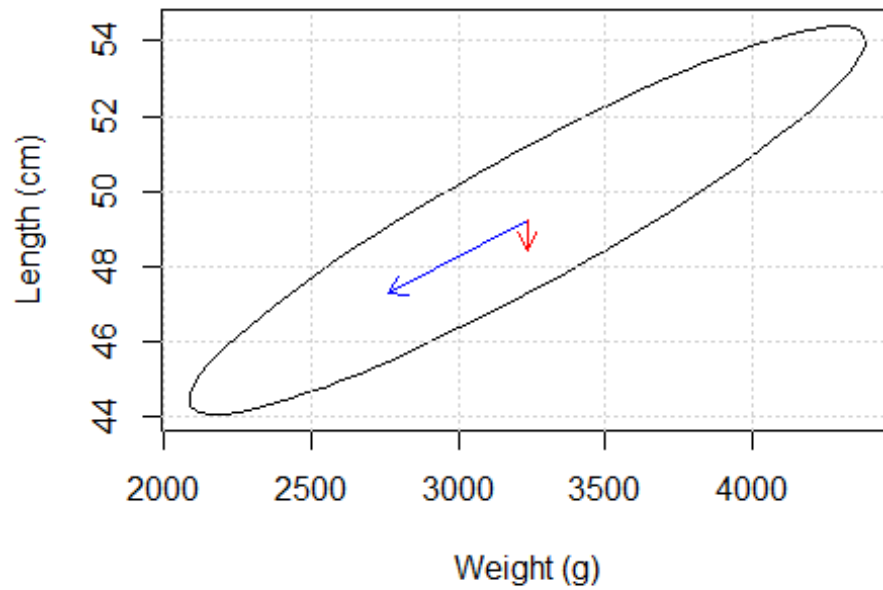


Histogram Length



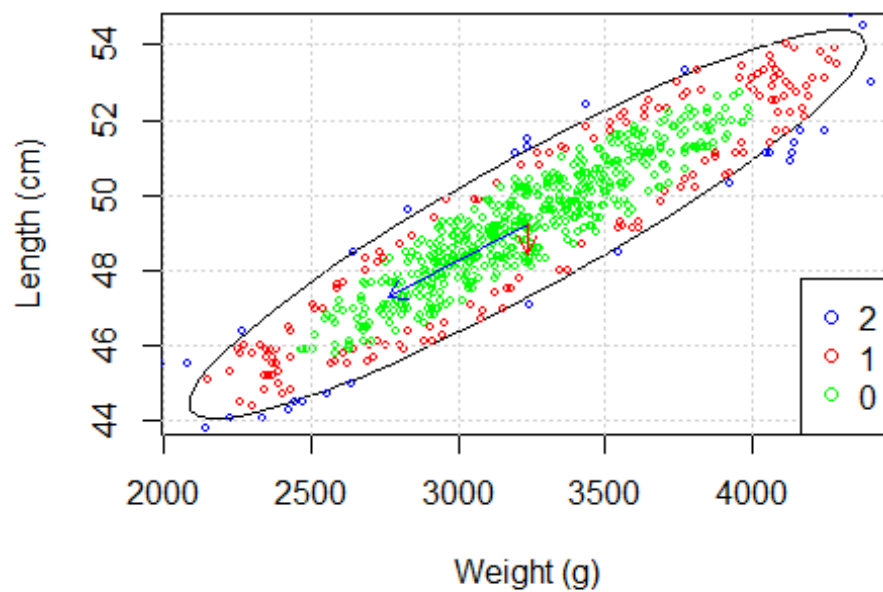
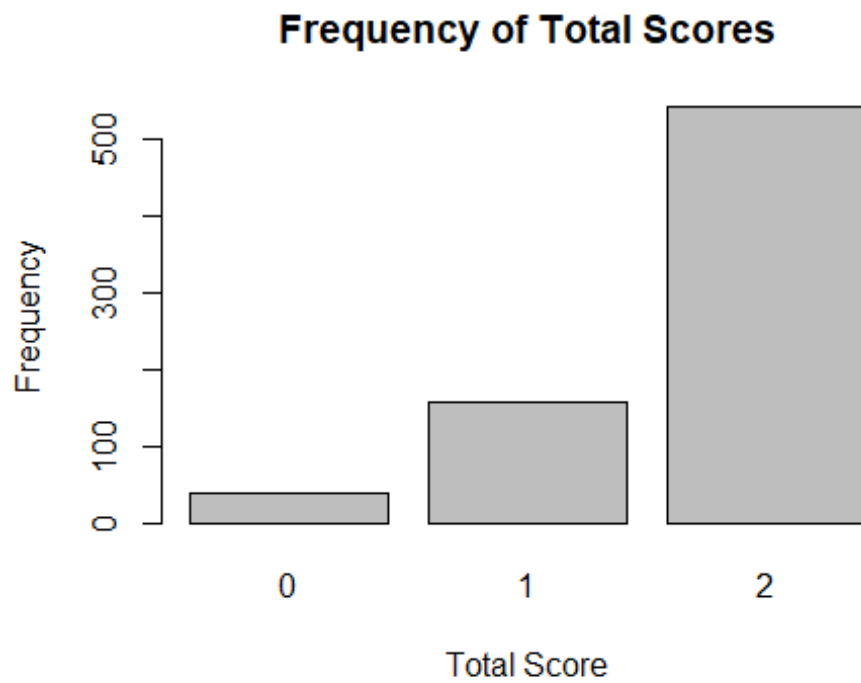


3-) Find the ellipsoids that would serve classification regions for scores as described above.



4-) How many children would score zero, one, and two, respectively? Illustrate this classification on the graphs.

##	Total_Score	Frequency
## 1	0	38
## 2	1	157
## 3	2	541

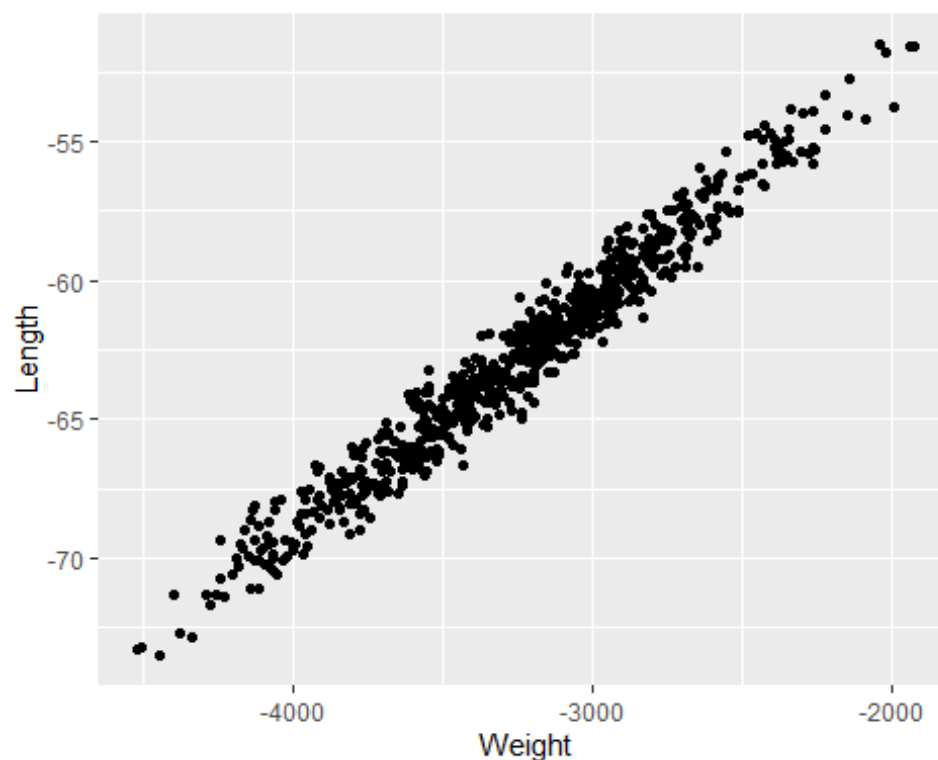


5-) Find the spectral decomposition of the estimated covariance matrix

```
## decomposition:
## P:  -0.9999914 -0.004155185 0.004155185 -0.9999914
## Λ:  220280.5 0 0 0.6400477
```

6-) Plot the data transformed according to $P^T X$, where P is the matrix made of the eigenvectors standing as the columns. Interpret the transformed data.

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```



Part Two

Additionally to weight and length of a child, also the height of parents is included in the records. In order to tune the procedure of scoring the height of parents can be also used. The ParentsWeightLength.txt file contains this information.

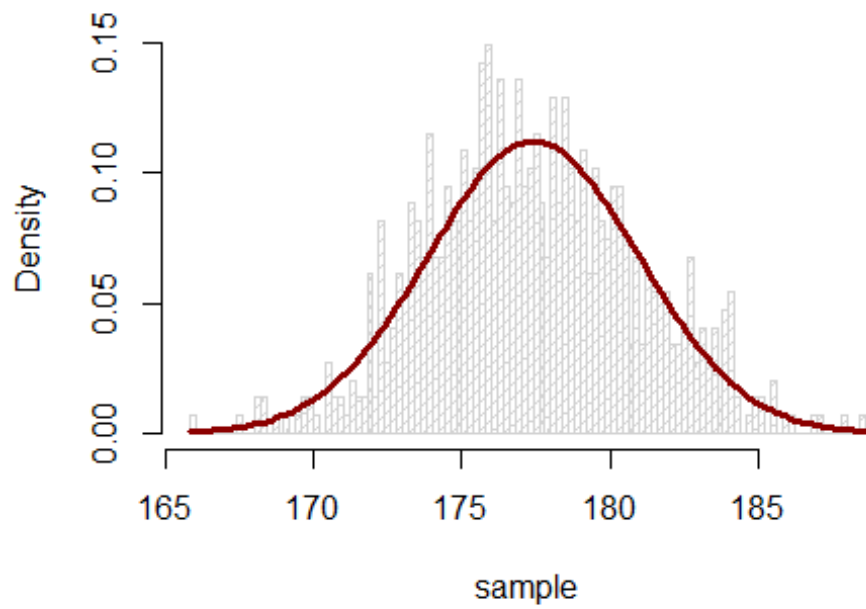
1-) Using the data estimate the mean and the covariance for all four variables.

```
##           FatherHeight MotherHeight      Weight      Length
## FatherHeight  12.6120963   0.6310056   931.8590   3.289545
## MotherHeight   0.6310056   9.7721065   827.2878   2.852147
## Weight        931.8589629  827.2877829 220276.6577  915.295511
## Length        3.2895454   2.8521470   915.2955   4.443303

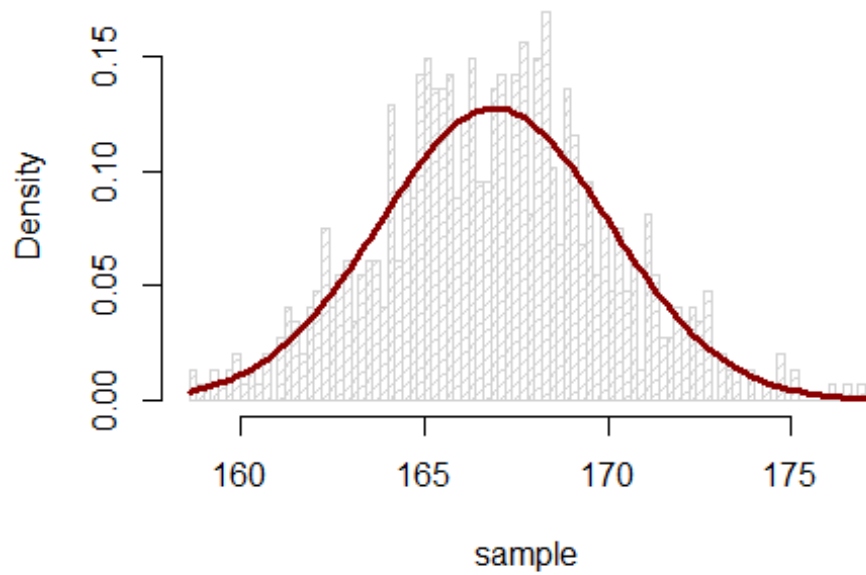
## Mean =  177.416 166.9196 3233.545 49.23764
```

2-) Verify graphically the normal distribution of the data. Use scatterplots and qq-plots for the marginal distributions.

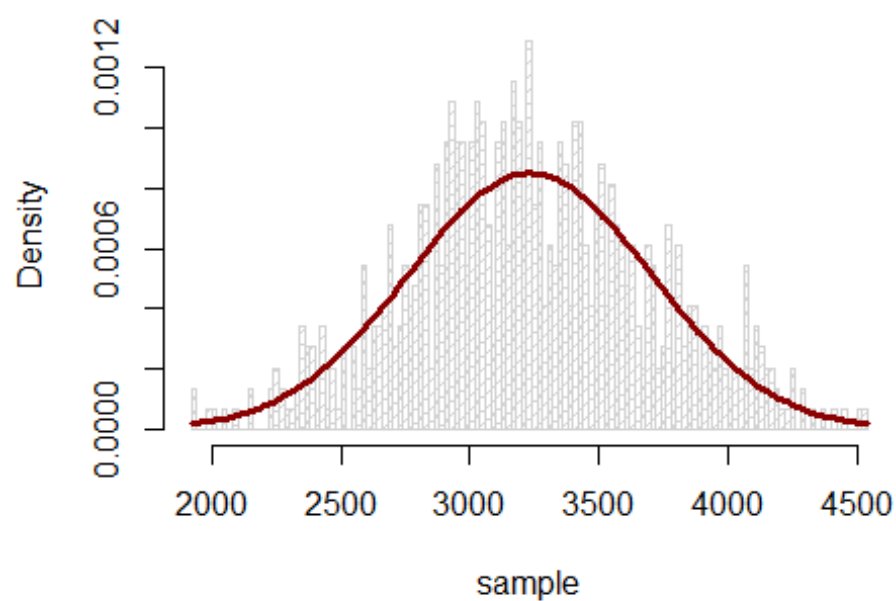
Histogram FatherHeight



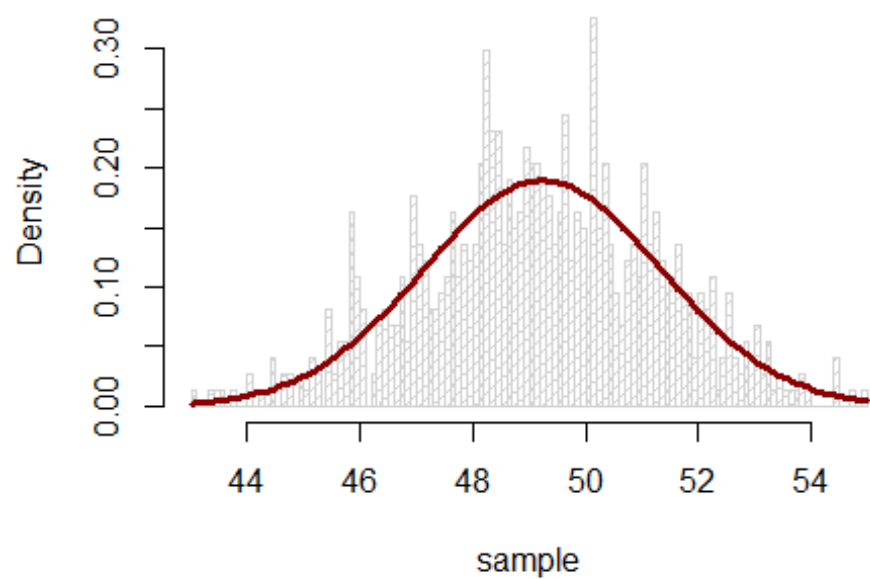
Histogram MotherHeight



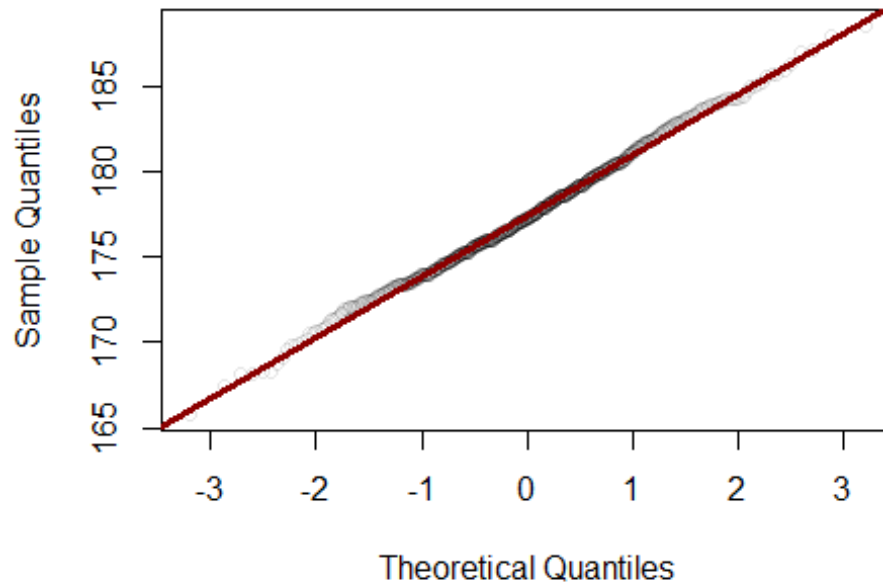
Histogram Weight



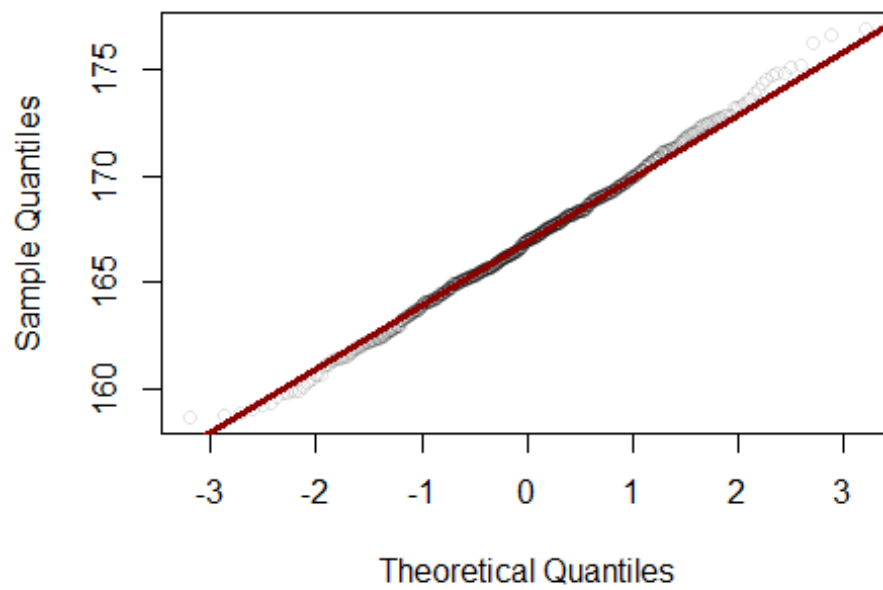
Histogram Length

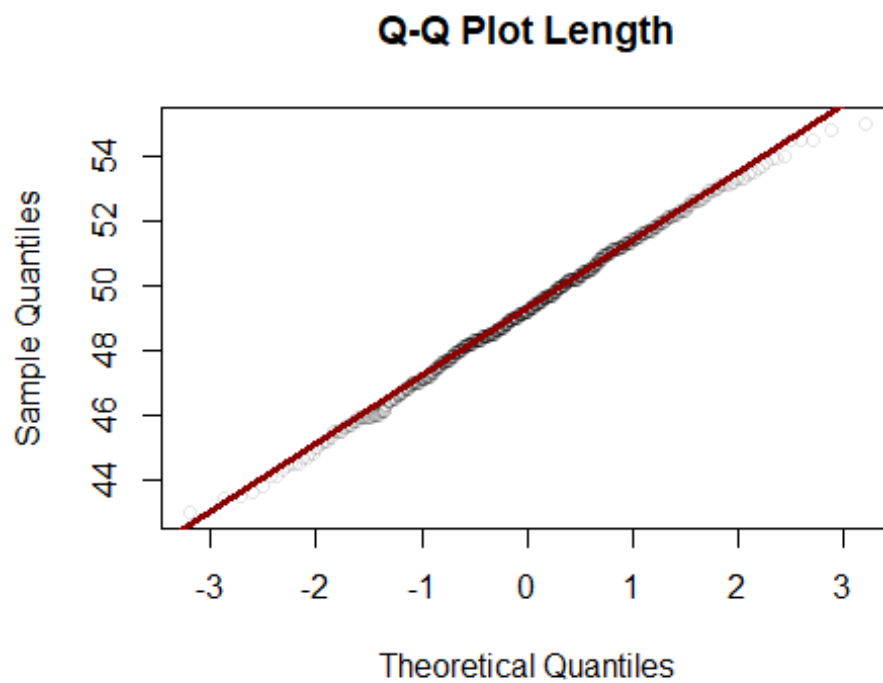
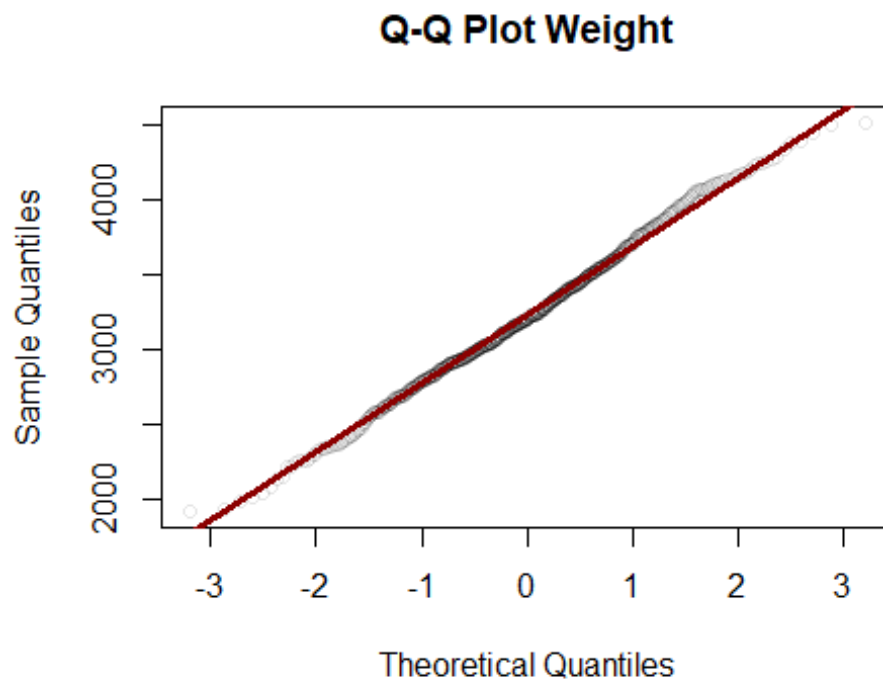


Q-Q Plot FatherHeight



Q-Q Plot MotherHeight





3-) Identify the conditional distribution of the weight and length of a child given the heights of parents. Find an estimate of the covariance matrix of the conditional distribution and compare it with the original unconditional covariance.

Conditional distribution of two bivariate normal distributions

$$L, W | F = f, M = m \sim \mathcal{N}(\mu_{WL} + \sum_{FMWL} \sum_{FM}^{-1} \left(\begin{bmatrix} f \\ m \end{bmatrix} - \mu_{FM} \right), \sum_{FMWL} \sum_{FM}^{-1} \sum_{WLFM})$$

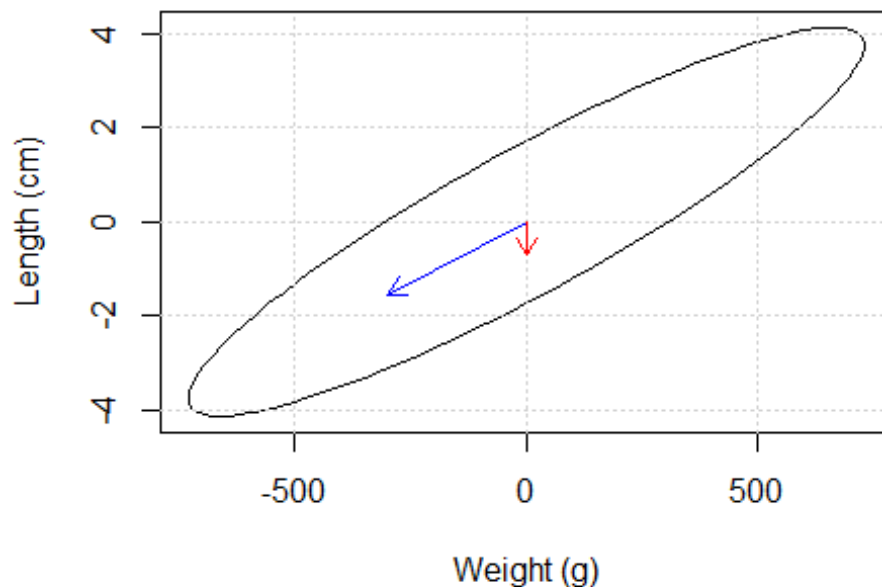
where:

- μ_{WL} mean vector of join distribution of weight and length
- μ_{FM} mean vector of join distribution of mother and father heights
- Σ_{FMWL} the covariance matrix of the joint distribution of all
- Σ_{FM} the covariance matrix of the joint distribution of father's and mother's heights

Sigma: 88858 456.8461 456.8461 2.843767

Variances in both Length and Weight has decreased significantly in conditional distribution. This indicates a reduction in uncertainty about both Weight and Length when conditioned on FatherHeight and MotherHeight and a weaker linear relationship between Weight and Length when conditioned on FatherHeight and MotherHeight.

4-) How the ellipsoids based on the conditional distribution will look like?

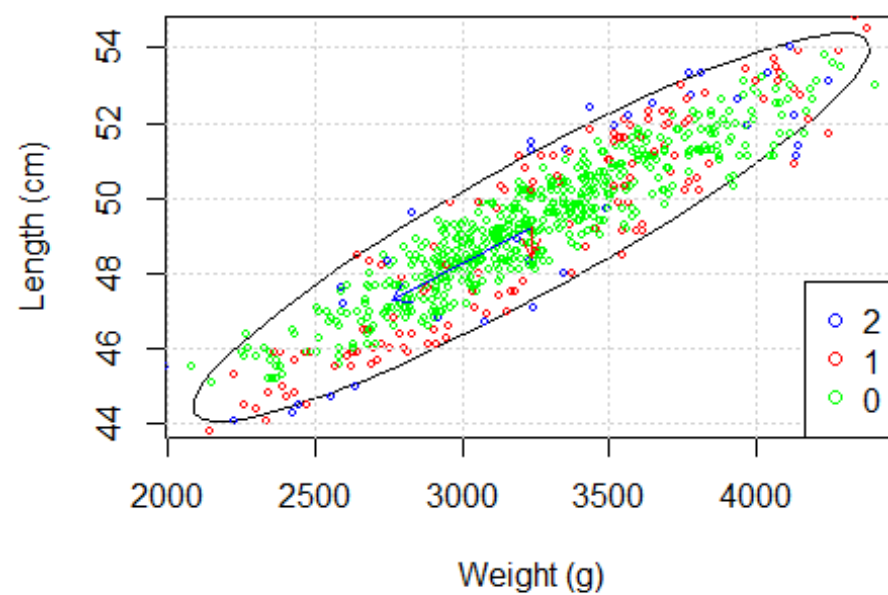
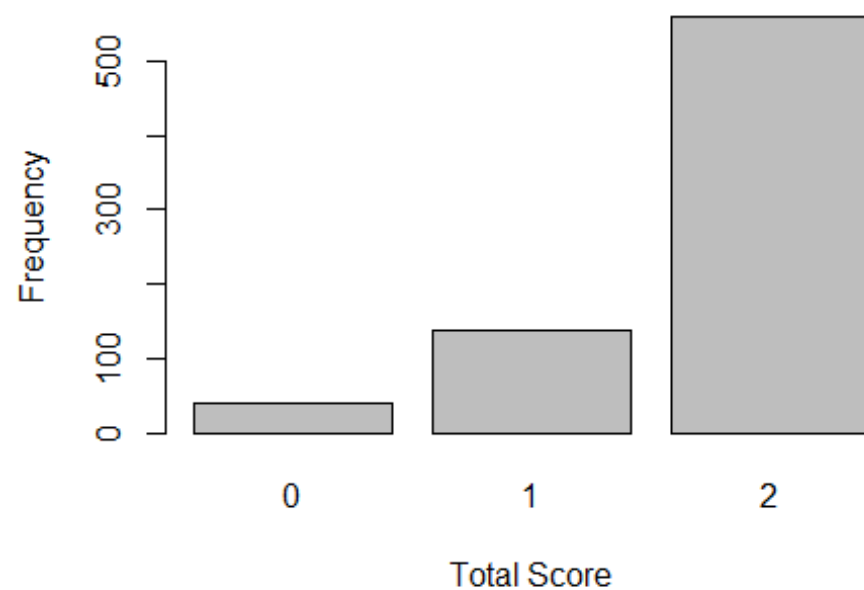


5-) How many children would score zero, one, and two, respectively? Illustrate this classification on the graph and compare with the one obtained without considering the heights of parents.

##	Total_Score	Frequency
## 1	0	40

##	2	1	137
##	3	2	559

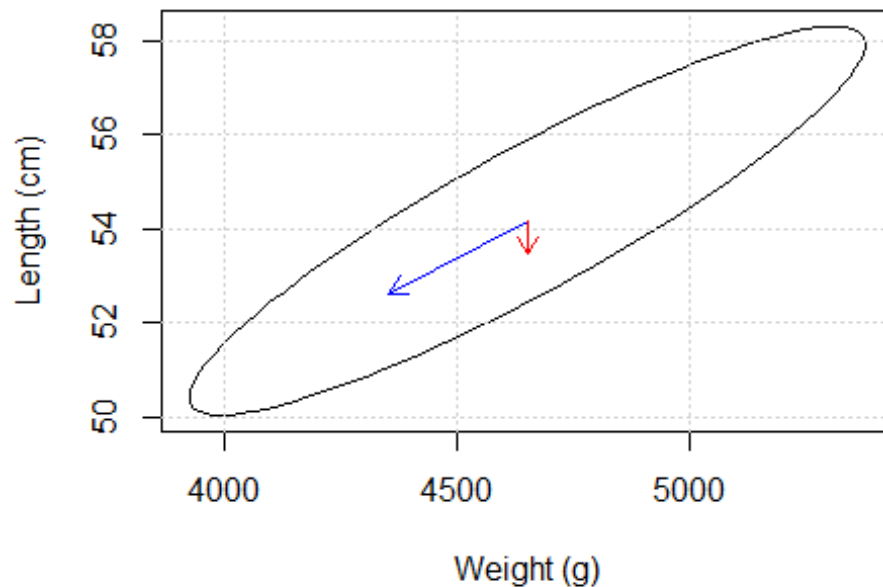
Frequency of Total Scores



When we compare the results, we can see that frequency of class 2 has significantly increased whilst class 1 frequency significantly decreased. We can also say that class 0 is increased a little bit.

6-) Suppose that the father of a child is 185[cm] tall and mother is 178[cm] tall. Plot the classification ellipsoids for their child.

```
## Example Mu: 4651.537 54.16829
```



7-) Find spectral decomposition of the estimated covariance matrix for the complete set of the data.

```
## decomposition:
## P: -0.9999868 -0.005141266 0.005141266 -0.9999868
## Λ: 88860.34 0 0 0.494969
```

8-) Transform the data the according to $P^T X$. Plot scatter plots of the transformed data.

