2. Measurements were taken on three biometric characters X_1, X_2 and X_3. The means of the characters and the matrix of the pooled variances and covariances for the two groups A and B are given below:

Groups	Sample size	X_1	X_2	X_3
A	20	58.0	25.8	7.8
B	72	59.7	28.4	7.4

Pooled variance-covariance matrix

$$\begin{pmatrix} 13.86 & 06.97 & 01.08 \\ & 04.79 & 00.56 \\ & & 00.14 \end{pmatrix}.$$

(a) Examine whether the two groups differ significantly in respect of the means of the two groups.

(b) The mean vector based on 10 observations from a third group C was (59.0, 26.5, 07.6), and the var-covariance matrix as given below:

Variables	X_1	X_2	X_3
X_1	4.7350	0.5622	1.4685
X_2		0.1413	0.2174
X_3			0.5702

Test the hypothesis : $\boldsymbol{\mu}_C = \frac{1}{3}\boldsymbol{\mu}_A + \frac{2}{3}\boldsymbol{\mu}_B$.

Solution:

<u>Theoretical understanding of the problem:</u>

A two-sample problem: Consider 2 indep distributions $N_p(\boldsymbol{\mu}^{(1)}, \Sigma)$ and $N_p(\boldsymbol{\mu}^{(2)}, \Sigma)$, $\Sigma > 0$.

Let $(X_1^{(1)}, \dots, X_{N_1}^{(1)})$ and $(X_1^{(2)}, \dots, X_{N_2}^{(2)})$ be indep. random samples drawn from the above distributions. Further, let $(\overline{\boldsymbol{X}}^{(1)}, A_1)$ and $(\overline{\boldsymbol{X}}^{(2)}, A_2)$ denote the sample mean and corrected SS-SP matrix of the 2 samples.

Define $A = A_1 + A_2$.

To test H_0: $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$.

Case 1: Σ known.

We have $\overline{\boldsymbol{X}}^{(1)} - \overline{\boldsymbol{X}}^{(2)} \sim N_p\left(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}, \Sigma(\frac{1}{N_1} + \frac{1}{N_2})\right)$.

We, therefore, suggest the test criterion

$$\chi^2 = (\overline{\boldsymbol{X}}^{(1)} - \overline{\boldsymbol{X}}^{(2)})'[\Sigma(\frac{1}{N_1} + \frac{1}{N_2})]^{-1}(\overline{\boldsymbol{X}}^{(1)} - \overline{\boldsymbol{X}}^{(2)})$$

$$= \frac{N_1 N_2}{N_1 + N_2}(\overline{\boldsymbol{X}}^{(1)} - \overline{\boldsymbol{X}}^{(2)})'\Sigma^{-1}(\overline{\boldsymbol{X}}^{(1)} - \overline{\boldsymbol{X}}^{(2)}).$$

$\chi^2 \sim \chi^2_{p,\gamma}$, where $\gamma = \frac{N_1 N_2}{N_1 + N_2}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\Sigma^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$.

$\gamma = 0$ iff $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$ (since Σ is p.d.).

So, under H_0, $\chi^2 \sim \chi^2_p$.

We reject H_0 at α level of significance if

obs. $\chi^2 > \chi^2_p(\alpha)$.

R Code :

```
# Create a matrix of means for each group
group1_means <- c(58.0, 25.8, 7.8)
group2_means <- c(59.7, 28.4, 7.4)
means_matrix <- rbind(group1_means, group2_means)
# Create the pooled variance-covariance matrix
pooled_matrix <- matrix(c(13.86, 6.97, 1.08, 6.97, 4.79, 0.56, 1.08, 0.56, 0.14), nrow = 3, ncol = 3, byrow
= TRUE)
solve(pooled_matrix)
# Calculate the Hotelling's T-square statistic
n1 <- 20
n2 <- 72
n <- n1 + n2
t_sq <- (n1*n2/n)*t(group1_means - group2_means)%*%solve(pooled_matrix) %*%(group1_means -
group2_means)
t_sq
# Calculate the critical value and p-value
df1 <- 3
df2 <- n – 3
crit_val <- qf(0.95, df1, df2)
p_val <- pf(t_sq, df1, df2, lower.tail = FALSE)
# Print the results
cat("Hotelling's T-squared statistic:", round(t_sq, 2), "\n")
cat("Critical value at alpha = 0.05:", round(crit_val, 2), "\n")
cat("p-value:", round(p_val, 4), "\n")
if (t_sq > crit_val) {
  cat("Reject the null hypothesis of equal mean vectors.\n")
} else {
  cat("Fail to reject the null hypothesis of equal mean vectors.\n")
}
```

Output:

```
Hotelling's T-squared statistic: 127.16

Critical value at alpha = 0.05: 2.71

Reject the null hypothesis of equal mean vectors.
```