

# STAT 665 Midterm

*Maggie Buffum*

*February 21, 2019*

## Problem 1

Suppose that  $\mathbf{X} = (X_1, X_2, X_3)'$  is a trivariate normal random vector with the moment generating function

$$m(\mathbf{t}) = \exp\left(t_1 - t_2 + 2t_3 + t_1^2 + \frac{1}{2}t_2^2 + 2t_3^2 - \frac{1}{2}t_1t_2 - t_1t_3\right), \quad \mathbf{t} \in \mathcal{R}^3$$

Let  $\bar{X} = (X_1 + X_2 + X_3)/3$ . Find the distribution of  $X_1$ , conditional on  $\bar{X} = 6$ .

---

From the mgf we find  $\mu$  and  $\Sigma$  by recalling that the mgf of a multivariate normal random variable takes the form

$$M_X(\mathbf{t}) = \exp\{\mathbf{t}'\mu + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\}$$

Clearly,

$$\mu = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

and

$$\Sigma = \begin{bmatrix} 2 & -\frac{1}{2} & -1 \\ -\frac{1}{2} & 1 & 0 \\ -1 & 0 & 4 \end{bmatrix}$$

We need to find the conditional probability of  $X_1$  given  $\bar{X} = 6$ . From Result 5.2.6 we know that  $[X_1, \bar{X}]' \sim N(\mathbf{B}\mu, \mathbf{B}\Sigma\mathbf{B}')$ . We can rewrite  $\bar{X}$  as  $\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$  such that

$$\begin{bmatrix} X_1 \\ \bar{X} \end{bmatrix} = \begin{bmatrix} X_1 \\ \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

Solving for  $\mathbf{B}\mu$  we get

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$$

and

$$\mathbf{B}\Sigma\mathbf{B}' = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & -\frac{1}{2} & -1 \\ -\frac{1}{2} & 1 & 0 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 2 & \frac{1}{6} \\ \frac{1}{6} & \frac{4}{9} \end{bmatrix}$$

From Result 5.2.10 we know that if  $\mathbf{x} \sim N_k(\mu, \Sigma)$  with rank of  $\Sigma$   $k$ , then the conditional distribution of  $x_1$  given that  $x_2 = c_2$  is multivariate normal with mean vector

$$\mu_{1.2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{c}_2 - \mu_2)$$

and

$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

We can partition  $\mathbf{x}$ ,  $\mathbf{B}\mu$ , and  $\mathbf{B}'\Sigma\mathbf{B}$  as

$$\begin{bmatrix} X_1 \\ X \end{bmatrix}, \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} 2 & | & \frac{1}{6} \\ \frac{1}{6} & | & \frac{4}{9} \end{bmatrix}.$$

Plugging into Result 5.2.10 we have that the distribution of  $X_1$  given  $\bar{X} = 6$  is normal with mean

$$\mu_{1.2} = 1 + \left(\frac{1}{6}\right)\left(\frac{4}{9}\right)^{-1}\left(6 - \frac{2}{3}\right) = 1 + \frac{9}{24} \frac{16}{3} = 1 + 2 = 3$$

and variance

$$\Sigma_{11.2} = 2 - \left(\frac{1}{6}\right)\left(\frac{4}{9}\right)^{-1}\left(\frac{1}{6}\right) = 2 - \frac{9}{24} \frac{1}{6} = 2 - \frac{1}{16} = \frac{31}{16}$$

## Problem 2

Suppose that  $\mathbf{Y} = X\beta + \epsilon$ , where  $\epsilon$  are  $N(\mathbf{0}, \sigma^2 \mathbf{I})$ . The design matrix is given as follows:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

(a) Compute the rank of  $P = X(X'X)^{-1}X'$ .

Solving for  $P_X$ , we get

```
X <- matrix(c(
  rep(1,6),rep(1,3),rep(0,3),rep(c(1,0),3)
),nrow = 6,byrow = F)

P_X <- X %*% solve(t(X) %*% X) %*% t(X)
print(P_X)
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,]  0.4166667  0.1666667  0.4166667 -0.08333333  0.1666667 -0.08333333
## [2,]  0.1666667  0.6666667  0.1666667  0.1666667 -0.3333333  0.1666667
## [3,]  0.4166667  0.1666667  0.4166667 -0.08333333  0.1666667 -0.08333333
## [4,] -0.08333333  0.1666667 -0.08333333  0.4166667  0.1666667  0.4166667
## [5,]  0.1666667 -0.3333333  0.1666667  0.1666667  0.6666667  0.1666667
## [6,] -0.08333333  0.1666667 -0.08333333  0.4166667  0.1666667  0.4166667
```

After row reduction, we have the following matrix, which has 3 linearly independent columns, and thus a rank of 3.

```
rref(P_X)

##           [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]      1    0    1    0    1    0
## [2,]      0    1    0    0   -1    0
## [3,]      0    0    0    1    1    1
## [4,]      0    0    0    0    0    0
## [5,]      0    0    0    0    0    0
## [6,]      0    0    0    0    0    0
```

(b) Compute the rank of  $Q = I - P$ .

---

We know that  $P_X$  is idempotent, so  $I_n - P_X$  is also idempotent with rank equal to  $n - \text{tr}(P_X) = n - \text{rank}(P_X) = 6 - 3 = 3$  (by Result 2.3.9). So the rank of  $Q$  is 3.

(c) If  $\rho(Q)$  is the rank of  $Q$ , prove that  $\mathbf{Y}'Q\mathbf{Y}/\rho(Q)$  is an unbiased estimator of  $\sigma^2$ , where  $\mathbf{Y} = (Y_1, \dots, Y_6)'$ .

---

As a result of Exercise 5.23, we have that  $E(\mathbf{x}'\mathbf{A}\mathbf{x}) = \text{tr}(\mathbf{A}\Sigma) + \mu'\mathbf{A}\mu$ . Therefore,

$$E[\mathbf{Y}'Q\mathbf{Y}/\rho(Q)] = \frac{1}{\rho(Q)}E[\mathbf{Y}'Q\mathbf{Y}] = \frac{1}{\rho(Q)}[\text{tr}(Q\Sigma) + \mu'Q\mu]$$

Note that  $\text{Var}(\mathbf{Y}) = \text{Var}(\mathbf{X}\beta + \epsilon) = \text{Var}(\epsilon) = \mathbf{I}\sigma^2$  and  $\mu = E[\mathbf{Y}] = E[\mathbf{X}\beta + \epsilon] = E[\mathbf{X}\beta] + E[\epsilon] = \mathbf{X}\beta$ .

Now we have

$$\begin{aligned} E[\mathbf{Y}'Q\mathbf{Y}/\rho(Q)] &= \frac{1}{\rho(Q)}[\text{tr}(Q\Sigma) + \mu'Q\mu] \\ &= \frac{1}{3}[\text{tr}(Q\mathbf{I}\sigma^2) + (X\beta)'(I - P)X\beta] \\ &= \frac{1}{3}[\text{tr}(Q\sigma^2) + \beta'X'X\beta - \beta'X'PX\beta] \\ \text{Recalling that } PX &= X, \\ &= \frac{1}{3}[\sigma^2\text{tr}(Q) + \beta'X'X\beta - \beta'X'X\beta] \\ &= \frac{1}{3}[3\sigma^2] \\ &= \sigma^2 \end{aligned}$$

Since  $E\left[\frac{\mathbf{Y}'Q\mathbf{Y}}{\rho(Q)}\right] = \sigma^2$ ,  $\frac{\mathbf{Y}'Q\mathbf{Y}}{\rho(Q)}$  is an unbiased estimator for  $\sigma^2$ .

### Problem 3

Suppose that  $X_1, X_2$ , and  $X_3$  are iid  $N(0, 1)$ .

(a) Find the distribution of  $Q$ , where

$$Q = \frac{1}{6}(5x_1^2 + 2x_2^2 + 5x_3^2 + 4x_1x_2 - 2x_1x_3 + 4x_2x_3)$$


---

We can write  $Q$  as

$$\begin{aligned} Q = \mathbf{x}'\mathbf{A}\mathbf{x} &= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} x_1a_{11} + x_2a_{21} + x_3a_{31} & x_1a_{12} + x_2a_{22} + x_3a_{32} & x_1a_{13} + x_2a_{23} + x_3a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= x_1^2a_{11} + x_2^2a_{22} + x_3^2a_{33} + x_1x_2(a_{21} + a_{12}) + x_1x_3(a_{31} + a_{13}) + x_2x_3(a_{32} + a_{23}) \end{aligned}$$

which implies that

$$A = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

If  $A$  is idempotent, we can use Result 5.4.1. Let's check:

```
A <- matrix(c(
  5/6, 2/6, -1/6,
  2/6, 2/6, 2/6,
  -1/6, 2/6, 5/6
), nrow = 3)
A

##           [,1]      [,2]      [,3]
## [1,]  0.8333333 0.3333333 -0.1666667
## [2,]  0.3333333 0.3333333  0.3333333
## [3,] -0.1666667 0.3333333  0.8333333
```

```
A%*%A

##           [,1]      [,2]      [,3]
## [1,]  0.8333333 0.3333333 -0.1666667
## [2,]  0.3333333 0.3333333  0.3333333
## [3,] -0.1666667 0.3333333  0.8333333
```

Since  $A$  is idempotent, we can use Result 5.4.1 to show that  $Q \sim \chi_m^2$ , where  $m = \text{rank}(A)$  which, in this case, is 2 (see  $A$  in row-reduced Echelon form below).

```
rref(A)

##           [,1] [,2] [,3]
## [1,]      1   0  -1
## [2,]      0   1   2
## [3,]      0   0   0
```

Therefore,  $Q \sim \chi_2^2$ .

(b) Find the distribution of  $Y$  and  $\text{Var}(Y)$ , where

$$Y = 4x_1^2 + 3x_2^2 + 4x_3^2 + 2\sqrt{2}x_2x_3.$$

(Hint: Express  $Y$  as a quadratic form with a matrix being nonidempotent. Try to use a famous theorem to represent the matrix in the quadratic form as a linear combination of some matrices.)

---

Similar to part (a), we can express  $Y$  as

$$Y = \mathbf{x}'\mathbf{A}\mathbf{x}$$

where

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & \sqrt{2} \\ 0 & \sqrt{2} & 4 \end{bmatrix}$$

```
A <- matrix(c(
  4, 0, 0,
  0, 3, sqrt(2),
  0, sqrt(2), 4
), nrow = 3)
A%*%A

##           [,1]      [,2]      [,3]
## [1,]      16  0.000000  0.000000
```

```
## [2,]    0 11.000000  9.899495
## [3,]    0  9.899495 18.000000
```

Clearly,  $A$  is not idempotent. Instead, let's consider the spectral decomposition theorem, since  $A$  is symmetric. Using this theorem we can rewrite  $\mathbf{Y}$  as

$$\begin{aligned} Y &= x'Ax \\ &= x'(\lambda_1 p_1 p_1' + \lambda_2 p_2 p_2' + \lambda_3 p_3 p_3')x \\ &= x'(\lambda_1 p_1 p_1')x + x'(\lambda_2 p_2 p_2')x + x'(\lambda_3 p_3 p_3')x \\ &= \lambda_1 x'(p_1 p_1')x + \lambda_2 x'(p_2 p_2')x + \lambda_3 x'(p_3 p_3')x \end{aligned}$$

where  $p_k$  are the eigenvectors for corresponding  $\lambda_k$ . Note that  $p_i p_i'$  are  $3 \times 3$  orthogonal, idempotent matrices of rank 1. Using Result 5.4.1 we know that since  $p_i p_i'$  are idempotent and  $X \sim iid N(0, 1)$ ,  $x_i'(p_i p_i')x_i$  are  $\chi_1^2$  independent random variables.

Let's determine the distribution of  $\lambda_1(x_i' p_i p_i' x) = \lambda(x' p p' x)$ . Let  $y = \lambda x$ . Then  $x = y/\lambda$ , and  $\frac{dx}{dy} \frac{y}{\lambda} = \frac{1}{\lambda}$ . The distribution of  $y$  is

$$f_Y(y) = \frac{\left(\frac{y}{\lambda}\right)^{1/2-1} e^{-(1/2)(y/\lambda)} \frac{1}{\lambda}}{\Gamma\left(\frac{1}{2}\right) 2^{1/2}} = \frac{y^{1/2-1} e^{-(y/2\lambda)}}{\Gamma\left(\frac{1}{2}\right) (2\lambda)^{1/2}},$$

that is,  $Y \sim \text{Gamma}(\frac{1}{2}, 2\lambda)$ . Now we have the sum of three independent  $\text{Gamma}(\frac{1}{2}, 2\lambda_i)$ , with  $\lambda_1 = 5, \lambda_2 = 4, \lambda_3 = 2$ . Since they do not have the same scale parameter (since  $\lambda_i$ s are unique), they don't form a known distribution. But they are independent, making the variance of  $\mathbf{Y}$  easy to calculate:

$$\begin{aligned} \text{Var}(\mathbf{Y}) &= \text{Var}(Y_1) + \text{Var}(Y_2) + \text{Var}(Y_3) \\ &= \frac{1}{2}(2\lambda_1)^2 + \frac{1}{2}(2\lambda_2)^2 + \frac{1}{2}(2\lambda_3)^2 \\ &= 2 \sum_{i=1}^3 \lambda_i^2 \end{aligned}$$

`eig(A)`

```
## [1] 5 4 2
```

The eigenvalues of  $\mathbf{A}$  are 5, 4, and 2, giving us

$$\text{Var}(\mathbf{Y}) = 2 \sum_{i=1}^3 \lambda_i^2 = 2(5^2 + 4^2 + 2^2) = 90.$$

## Problem 4

Consider the two-way crossed classification model without interaction with exactly one observation per cell, i.e.,

$$y_{ij} = \mu + \alpha_i + \beta_j + d_{ij}, \quad (i = 1, \dots, p; j = 1, \dots, q).$$

Suppose further that  $\mathbf{d} = \{d_{ij}\}$  has a multivariate normal distribution with mean vector  $\mathbf{0}$  and variances and covariances given by

$$\begin{aligned} \text{cov}(d_{ij}, d_{i',j'}) &= \lambda + 2\gamma_i, & \text{if } i' = i \text{ and } j' = j \\ &= \gamma_i + \gamma_{i'}, & \text{if } i' \neq i \text{ and } j' = j \\ &= 0, & \text{if } j' \neq j. \end{aligned}$$

Here,  $\lambda$  and  $\gamma_i$  are known quantities satisfying  $\lambda > 0$  and  $\lambda + 2\gamma_i > 0$ , for all  $i$ .

- (a) Express  $q \sum_{i=1}^p (\bar{y}_{i.} - \bar{y}_{..})^2$  as a quadratic form in the vector  $\mathbf{z}$ , where  $\mathbf{z}' = (\bar{y}_1, \dots, \bar{y}_p)$ .

---

Let's expand  $q \sum_{i=1}^p (\bar{y}_{i.} - \bar{y}_{..})^2$ :

$$\begin{aligned}
q \sum_{i=1}^p (\bar{y}_{i.} - \bar{y}_{..})^2 &= q \sum_{i=1}^p (\bar{y}_{i.}^2 + \bar{y}_{..}^2 - 2\bar{y}_{i.}\bar{y}_{..}) \\
&= q \left[ \sum_{i=1}^p \bar{y}_{i.}^2 + \sum_{i=1}^p \bar{y}_{..}^2 - \sum_{i=1}^p 2\bar{y}_{i.}\bar{y}_{..} \right] \\
&= q \left[ \sum_{i=1}^p \bar{y}_{i.}^2 + p\bar{y}_{..}^2 - 2p\bar{y}_{..}^2 \right] \\
&= q \left[ \sum_{i=1}^p \bar{y}_{i.}^2 - p\bar{y}_{..}^2 \right]
\end{aligned}$$

We are looking for some matrix  $\mathbf{A}$  such that  $\mathbf{z}'\mathbf{A}\mathbf{z} = q [\sum_{i=1}^p \bar{y}_{i.}^2 - p\bar{y}_{..}^2]$ . Consider  $\mathbf{A} = \mathbf{B} - \mathbf{C}$  such that  $\mathbf{z}'\mathbf{A}\mathbf{z} = \mathbf{z}'q(\mathbf{B} - \mathbf{C})\mathbf{z}$ . Note that

$$\mathbf{z}'\mathbf{I}_p\mathbf{z} = \begin{bmatrix} \bar{y}_{1.} & \bar{y}_{2.} & \dots & \bar{y}_{p.} \end{bmatrix} I_p \begin{bmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix} = \bar{y}_{1.}^2 + \bar{y}_{2.}^2 + \dots + \bar{y}_{p.}^2 = \sum_{i=1}^p \bar{y}_{i.}^2$$

and

$$\begin{aligned}
\mathbf{z}'\frac{1}{p}\mathbf{J}_p\mathbf{z} &= \begin{bmatrix} \bar{y}_{1.} & \bar{y}_{2.} & \dots & \bar{y}_{p.} \end{bmatrix} \frac{1}{p} J_p \begin{bmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix} \\
&= \frac{1}{p} \begin{bmatrix} \sum_{i=1}^p \bar{y}_{i.} & \sum_{i=1}^p \bar{y}_{i.} & \dots & \sum_{i=1}^p \bar{y}_{i.} \end{bmatrix} \begin{bmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix} \\
&= \frac{1}{p} \sum_{i=1}^p \bar{y}_{i.} \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \vdots \\ \bar{y}_{p.} \end{bmatrix} \\
&= \frac{1}{p} \sum_{i=1}^p \bar{y}_{i.} \times \sum_{i=1}^p \bar{y}_{i.} \\
&= p\bar{y}_{..}^2
\end{aligned}$$

Therefore,  $q \sum_{i=1}^p (\bar{y}_{i.} - \bar{y}_{..})^2 = \mathbf{z}'q(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p)\mathbf{z}$ .

(b) Obtain the covariance matrix of  $\mathbf{z}$ .

---

To begin, let's look at the full model variance-covariance matrix, based on the description of  $cov(d_{ij}, d_{i'j'})$  given to us. We know that within each  $q \times q$  submatrix where  $i = i'$ , that is, the submatrices along the diagonal, the covariance is  $\lambda + 2\gamma_i$  whenever  $j = j'$ , giving us  $(\lambda + 2\gamma_i)I_q$  submatrices along the diagonal. The off-diagonal  $q \times q$  partitions include all cases where  $i \neq i'$ . Within these off-diagonal submatrices,

the covariance is  $(\gamma_i + \gamma_{i'})I_q$ . For all submatrices, the covariance is 0 whenever  $j \neq j'$ , which is why each submatrix is a scalar multiplied by the identity matrix.

$$\begin{bmatrix} (\lambda + 2\gamma_1)I_q & (\gamma_1 + \gamma_2)I_q & \dots & (\gamma_1 + \gamma_p)I_q \\ (\gamma_2 + \gamma_1)I_q & (\lambda + 2\gamma_2)I_q & \dots & (\gamma_2 + \gamma_p)I_q \\ \vdots & \vdots & \ddots & \vdots \\ (\gamma_p + \gamma_1)I_q & (\gamma_p + \gamma_2)I_q & \dots & (\lambda + 2\gamma_p)I_q \end{bmatrix}_{pq \times pq}$$

But we want to find the covariance of  $\mathbf{z}' = [\bar{y}_1, \bar{y}_2, \dots, \bar{y}_p]'$ . Because  $\bar{y}_i$  is the sum across all  $q$  within group  $i$ , and  $\mathbf{y}$  is normally distributed (since its variation comes from the  $d_{ij}$ s),  $\bar{y}_i$ s will be normally distributed (the reproductive property) with the following variance-covariance structure:

$$Cov(\bar{Y}_i, \bar{Y}_{i'}) = \frac{1}{q} \begin{bmatrix} \lambda + 2\gamma_1 & \gamma_1 + \gamma_2 & \dots & \gamma_1 + \gamma_p \\ \gamma_2 + \gamma_1 & \lambda + 2\gamma_2 & \dots & \gamma_2 + \gamma_p \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_p + \gamma_1 & \gamma_p + \gamma_2 & \dots & \lambda + 2\gamma_p \end{bmatrix}_{p \times p}$$

- (c) Show that  $q \sum_{i=1}^p (\bar{y}_i - \bar{y}_{..})^2$  is distributed as a scalar of a noncentral chi-square random variable, and determine the noncentrality parameter. You may assume that the covariance matrix of  $\mathbf{z}$  is positive definite.

---

By Result 5.4.5, if we can show that  $q(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p)\Sigma_z$  is idempotent, then  $q \sum_{i=1}^p (\bar{y}_i - \bar{y}_{..})^2$  has a noncentral chi-square distribution. Note that  $\Sigma_z$  can be rewritten as  $\lambda\mathbf{I} + \mathbf{1}\gamma' + \gamma\mathbf{1}'$ .

Let's check for idempotency by first simplifying  $A\Sigma_z$ :

$$\begin{aligned} A\Sigma_z &= q(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p)\frac{1}{q}(\lambda\mathbf{I}_p + \mathbf{1}\gamma' + \gamma\mathbf{1}') \\ &= \lambda\mathbf{I}_p + \mathbf{1}\gamma' + \gamma\mathbf{1}' - \frac{1}{p}\mathbf{J}_p\lambda\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p\mathbf{1}\gamma' - \frac{1}{p}\mathbf{J}_p\gamma\mathbf{1}' \\ &= \lambda\mathbf{I}_p + \mathbf{1}\gamma' + \gamma\mathbf{1}' - \frac{1}{p}\mathbf{J}_p\lambda\mathbf{I}_p - \frac{1}{p}\mathbf{1}\mathbf{1}'\mathbf{1}\gamma' - \frac{1}{p}\mathbf{J}_p\gamma\mathbf{1}' \\ &= \lambda\mathbf{I}_p + \mathbf{1}\gamma' + \gamma\mathbf{1}' - \frac{1}{p}\mathbf{J}_p\lambda\mathbf{I}_p - \frac{1}{p}\mathbf{1}p\gamma' - \frac{1}{p}\mathbf{J}_p\gamma\mathbf{1}' \\ &= \lambda\mathbf{I}_p + \gamma\mathbf{1}' - \frac{1}{p}\mathbf{J}_p\lambda\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p\gamma\mathbf{1}' \\ &= \lambda(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p) + \gamma\mathbf{1}'(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p) \\ &= (\lambda + \gamma\mathbf{1}')(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p) \end{aligned}$$

Now we can check for idempotency of  $A\Sigma_z$ .

$$\begin{aligned}
\mathbf{A}\Sigma_z\mathbf{A}\Sigma_z &= (\lambda + \gamma\mathbf{1}')(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p)(\lambda + \gamma\mathbf{1}')(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p) \\
&= (\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma\mathbf{1}'\mathbf{J}_p)(\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma\mathbf{1}'\mathbf{J}_p) \\
&= (\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma\mathbf{1}'\mathbf{1}\mathbf{1}')(\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma\mathbf{1}'\mathbf{1}\mathbf{1}') \\
&= (\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma p\mathbf{1}')(\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \frac{1}{p}\gamma p\mathbf{1}') \\
&= (\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \gamma\mathbf{1}')(\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p + \gamma\mathbf{1}' - \gamma\mathbf{1}') \\
&= (\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p)(\lambda\mathbf{I}_p - \frac{1}{p}\lambda\mathbf{J}_p) \\
&= \lambda^2(\mathbf{I}_p - \frac{1}{p}\mathbf{J}_p)
\end{aligned}$$

$\lambda$  is obviously a scalar. If we define  $\mathbf{A}^* = \frac{1}{\lambda}\mathbf{A}$ , we quickly see that  $\mathbf{z}'\mathbf{A}^*\mathbf{z}$  has a noncentral chi-square distribution with  $rank(A^*) = rank(\frac{1}{\lambda}(I_p - \frac{1}{p}J_p)) = rank(I_p - \frac{1}{p}J_p) = p - rank(J_p) = p - 1$  and noncentrality parameter



$$\begin{aligned}
\mu' \mathbf{A}^* \mu &= \frac{q}{\lambda} \mu' (\mathbf{I}_p - \frac{1}{p} \mathbf{J}_p) \mu \\
&= \frac{q}{\lambda} [\mu_{1.} \quad \mu_{2.} \quad \dots \quad \mu_{p.}] (\mathbf{I}_p - \frac{1}{p} \mathbf{J}_p) \begin{bmatrix} \mu_{1.} \\ \mu_{2.} \\ \dots \\ \mu_{p.} \end{bmatrix} \\
&= \frac{q}{\lambda} [\mu + \alpha_1 \quad \mu + \alpha_2 \quad \dots \quad \mu + \alpha_p] \begin{bmatrix} 1 - \frac{1}{p} & -\frac{1}{p} & \dots & -\frac{1}{p} \\ -\frac{1}{p} & 1 - \frac{1}{p} & \dots & -\frac{1}{p} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{p} & -\frac{1}{p} & \dots & 1 - \frac{1}{p} \end{bmatrix} \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_2 \\ \dots \\ \mu + \alpha_p \end{bmatrix} \\
&= \frac{q}{\lambda} [\mu + \alpha_1 - \frac{1}{p}(\mu + \alpha_1) - \frac{1}{p}(\mu + \alpha_2) - \frac{1}{p}(\mu + \alpha_s) \quad \dots \quad \dots \quad \dots] \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_2 \\ \dots \\ \mu + \alpha_p \end{bmatrix} \\
&= \frac{q}{\lambda} [\mu + \alpha_1 - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i) \quad \mu + \alpha_2 - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i) \quad \dots \quad \mu + \alpha_p - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i)] \begin{bmatrix} \mu + \alpha_1 \\ \mu + \alpha_2 \\ \dots \\ \mu + \alpha_p \end{bmatrix} \\
&= \frac{q}{\lambda} (\mu + \alpha_1 - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i))(\mu + \alpha_1) \\
&\quad + (\mu + \alpha_2 - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i))(\mu + \alpha_2) + \dots + (\mu + \alpha_p - \frac{1}{p}(p\mu + \sum_{i=1}^p \alpha_i))(\mu + \alpha_p) \\
&= \frac{q}{\lambda} (\alpha_1 + \mu)(\alpha_1 - \bar{\alpha}_{.}) + (\alpha_2 + \mu)(\alpha_2 - \bar{\alpha}_{.}) + \dots + (\alpha_p + \mu)(\alpha_p - \bar{\alpha}_{.}) \\
&= \frac{q}{\lambda} \sum_{i=1}^p (\alpha_i + \mu)(\alpha_i - \bar{\alpha}_{.}) \\
&= \frac{q}{\lambda} \sum_{i=1}^p (\alpha_i^2 - \alpha_i \bar{\alpha}_{.} + \mu \alpha_i - \mu \bar{\alpha}_{.}) \\
&= \frac{q}{\lambda} \left( \sum_{i=1}^p \alpha_i^2 - p \bar{\alpha}_{.}^2 + p \mu \bar{\alpha}_{.} - p \mu \bar{\alpha}_{.} \right) \\
&= \frac{q}{\lambda} \left( \sum_{i=1}^p \alpha_i^2 - p \bar{\alpha}_{.}^2 \right)
\end{aligned}$$

As a result, we see that  $q \sum_{i=1}^p (\bar{y}_{i.} - \bar{y}_{..})^2 = \mathbf{z}' \mathbf{A} \mathbf{z} = \mathbf{z}' \lambda \mathbf{A}^* \mathbf{z}$  is distributed as a scalar of a noncentral chi-square distribution with parameters  $p - 1$  and  $\frac{q}{\lambda} (\sum_{i=1}^p \alpha_i^2 - p \bar{\alpha}_{.}^2)$ .