Punto 1

$$i = 1, \dots, n \quad Y_i \quad \text{be} \quad \in \{1, 2, \dots, p+1\}$$

with probability $\pi_1, \dots, \pi_p, \pi_{p+1} > 0 \quad \sum_{j=1}^{p+1} \pi_j = 1$. The one-hot vector $Y_i = (Y_{i1}, \dots, Y_{ip+1})$ such that $Y_{ij} = 1 \quad \text{when} \quad Y_i = j \implies \sum_{i=1}^n Y_i$ has multinomial distribution ie $\sum_{i=1}^n Y_i \sim \text{Multinomial}(n, \pi_1, \dots, \pi_{p+1})$

$$Y_i = \begin{cases} 1 & \text{with probability } \pi_1 \\ 2 & \text{with probability } \pi_2 \\ & \quad \dots \\ p+1 & \quad \dots \qquad \pi_2 \end{cases}$$

$Y_i = (Y_{i1}, \dots, Y_{ip+1}) \to 1 \leq j \leq p+1 \quad \text{and} \quad Y_{ij} = 1$

So Y_i codified as one hot vectors is distributed as such $Y_i \sim \text{Multinomial}(1, \pi_1, \dots, \pi_{p+1})$.

By the CLT we can say

$$(X_1, \dots, X_{p+1}) \sim \mathcal{N}_{p+1}(0, \text{diag}(\pi) - \pi\pi^T) \implies \tag{1}$$

$$\Sigma = \begin{pmatrix} \pi_1 & & & \\ & \pi_2 & & \\ & & \ddots & \\ & & & \pi_{p+1} \end{pmatrix} - \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_{p+1} \end{pmatrix} \begin{pmatrix} \pi_1 & \dots & \pi_{p+1} \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} \pi_1 & & & \\ & \pi_2 & & \\ & & \ddots & \\ & & & \pi_{p+1} \end{pmatrix} - \begin{pmatrix} \pi_1^2 & \pi_1\pi_2 & \dots & \pi_1\pi_{p+1} \\ \pi_2\pi_1 & \pi_2^2 & \dots & \pi_2\pi_{p+1} \\ \vdots & & \ddots & \\ \pi_{p+1}\pi_1 & & & \pi_{p+1}^2 \end{pmatrix} = \tag{3}$$

$$\Sigma = \left(\begin{array}{ccc|c} \pi_1(1-\pi_1) & -\pi_1\pi_2 & \dots & -\pi_1\pi_{p+1} \\ -\pi_2\pi_1 & \pi_2(1-\pi_2) & \dots & -\pi_2\pi_{p+1} \\ \vdots & & & \\ \hline -\pi_{p+1}\pi_1 & -\pi_{p+1}\pi_2 & \dots & \pi_{p+1}(1-\pi_{p+1}) \end{array}\right) \tag{4}$$

It's like $X_1, \dots, X_p \sim \mathcal{N}_{p+1}(\underline{0}, \Sigma)$ so$X_1, \dots, X_p$ it's like the marginal of these component where $\underline{0}_p$ is the new vector of the mean with 0 component with dimension $p < p+1$ and $\Sigma_{p+1,p+1} = \underbrace{\pi_{p+1}}_{>0} \underbrace{(1 - \pi_{p+1})}_{>0(1)} > 0$ (1)

$1 - \pi_{p+1} > 0 \Leftrightarrow \pi_{p+1} < 1$ and this is true beacuse π_{p+1} is a probability so it is defined in (0,1) so the two condition are valid :

$$(X_1, \dots, X_p) \sim \mathcal{N}_p(\underline{0}, \Sigma_{p \times p})$$

where

$$\Sigma_{p \times p} = \begin{pmatrix} \pi_1(1-\pi_1) & -\pi_1\pi_2 & \dots & -\pi_1\pi_p \\ -\pi_2\pi_1 & \pi_2(1-\pi_2) & \dots & -\pi_2\pi_p \\ \vdots & & \ddots & \\ -\pi_p\pi_1 & -\pi_p\pi_2 & \dots & \pi_p(1-\pi_p) \end{pmatrix}$$

Punto 2

We want to find the inverse of the covariance matrix:

$$\Sigma = \begin{pmatrix} \pi_1^2 & -\pi_1\pi_2 & \dots & -\pi_1\pi_p \\ -\pi_2\pi_1 & \pi_2^2 & \dots & -\pi_2\pi_p \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_p\pi_1 & -\pi_p\pi_2 & \dots & \pi_p^2 \end{pmatrix}$$

First of all we need to check if the matrix is invertible, so we need to prove that $det(\Sigma) \neq 0$
Since evaluating the determinant is not quite easy, we find a Lemma that give us an equivalent condition for the invertibility:

Sherman Morrison Formula: Suppose $A \in \mathbb{R}^{n\times n}$ invertible and $u, v \in \mathbb{R}$ are column vectors. $A + uv^T$ is invertible $\iff 1 + v^T Au \neq 0$.In this case: $(A + uv^T) = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}(2)$.

Application: we know that

$$\Sigma = diag(\pi_1, \dots, \pi_p) - \pi\pi^T = A + uv^T$$

where: - $A = \text{diag}(\pi_1, \dots, \pi_p) \in \mathbb{R}^{p\times p}$ - $u = \begin{pmatrix} -\pi_1 \\ -\pi_2 \\ \vdots \\ -\pi_p \end{pmatrix}$ - $v^T = \begin{pmatrix} \pi_1 & \dots & \pi_p \end{pmatrix}$

Remark : in our case there is a minus so to be coherent with the notation of the formula (2) we take the minus inside the vector u.

Is A invertible?

$$|A| = \left| \begin{pmatrix} \pi_1 & & \\ & \ddots & \\ & & \pi_p \end{pmatrix} \right| = \pi_1\pi_2 \dots \pi_p > 0$$

so it is invertible beacuse the determinant is bigger than 0

To apply the theorem we check that $1 + v^T A^{-1}u \neq 0$ Since A is a diagonal matrix we know that the inverse is as such:

$$A^{-1} = \begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix}$$

now we calculate :

$$1 + v^T A^{-1}u = 1 + \begin{pmatrix} \pi_1 & \dots & \pi_p \end{pmatrix} \begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix} \begin{pmatrix} -\pi_1 \\ \vdots \\ -\pi_p \end{pmatrix} =$$

$$= 1 + \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} -\pi_1 \\ \vdots \\ -\pi_p \end{pmatrix} = \tag{5}$$

$$= 1 - (\pi_1 + \dots + \pi_p) \neq 0 \Leftrightarrow \pi_1 + \dots + \pi_p \neq 1 \tag{6}$$

this is always true by hypothesis $\sum_{i=1}^{p+1} \pi_i = 1$ and $\pi_i > 0$ so $\sum_{i=1}^{p} \pi_i < 1$ because π_{p+1} is missing The hypotesis of the theorem are verifed now we need to find

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{\underbrace{1 + v^tA^{-1}u}_{1-(\pi_1+\dots+\pi_p)=\pi_{p+1}}}$$

The numerator of the fraction is the following

$$\underbrace{\begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix}}_{p\times p} \underbrace{\begin{pmatrix} -\pi_1 \\ \vdots \\ -\pi_p \end{pmatrix}}_{p\times 1} \begin{pmatrix} \pi_1 & \dots & \pi_p \end{pmatrix} \begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix} = \tag{7}$$

$$\begin{pmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{pmatrix} \begin{pmatrix} \pi_1 & \dots & \pi_p \end{pmatrix} \begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix} = \tag{8}$$

$$\begin{pmatrix} -\pi_1 & \dots & -\pi_p \\ -\pi_1 & \dots & -\pi_p \\ \vdots & & \vdots \\ -\pi_1 & \dots & -\pi_p \end{pmatrix} \begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & \dots & -1 \\ -1 & \dots & -1 \\ \vdots & & \vdots \\ -1 & \dots & -1 \end{pmatrix}}_{p\times p} \tag{9}$$

$$\Sigma^{-1} = A^{-1} - \begin{pmatrix} -\frac{1}{\pi_{p+1}} & \dots & -\frac{1}{\pi_{p+1}} \\ \vdots & \ddots & \vdots \\ -\frac{1}{\pi_{p+1}} & \dots & -\frac{1}{\pi_{p+1}} \end{pmatrix} = \tag{10}$$

$$\begin{pmatrix} \frac{1}{\pi_1} & & \\ & \ddots & \\ & & \frac{1}{\pi_p} \end{pmatrix} + \begin{pmatrix} \frac{1}{\pi_{p+1}} & \dots & \frac{1}{\pi_{p+1}} \\ \vdots & \ddots & \vdots \\ \frac{1}{\pi_{p+1}} & \dots & \frac{1}{\pi_{p+1}} \end{pmatrix} = \tag{11}$$

$$\begin{pmatrix} \frac{1}{\pi_1} + \frac{1}{\pi_{p+1}} & \frac{1}{\pi_{p+1}} & \dots & \frac{1}{\pi_{p+1}} \\ \frac{1}{\pi_{p+1}} & \frac{1}{\pi_2} + \frac{1}{\pi_{p+1}} & \dots & \frac{1}{\pi_{p+1}} \\ \vdots & \vdots & & \vdots \\ \frac{1}{\pi_{p+1}} & \frac{1}{\pi_{p+1}} & \dots & \frac{1}{\pi_p} + \frac{1}{\pi_{p+1}} \end{pmatrix} \tag{12}$$

Punto 3

Punto 4

Punto 5

Given a vector a Gaussian random vector we know that conditioning on a Gaussian random variable the conditioned vector is still Gaussian with parameters :

$$\overline{\mu} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2)$$

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

We know that $(X_1, X_2, X_3) \sim \mathcal{N}_3(\underline{0}, \Sigma)$ where

$$\Sigma = \begin{pmatrix} \frac{3}{16} & -\frac{1}{16} & -\frac{1}{16} \\ -\frac{1}{16} & \frac{3}{16} & -\frac{1}{16} \\ -\frac{1}{16} & -\frac{1}{16} & \frac{3}{16} \end{pmatrix}$$

$$\left(\begin{array}{cc|c} \frac{3}{16} & -\frac{1}{16} & -\frac{1}{16} \\ -\frac{1}{16} & \frac{3}{16} & -\frac{1}{16} \\ \hline -\frac{1}{16} & -\frac{1}{16} & \frac{3}{16} \end{array}\right) = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right)$$

We need to find the distribution of $(X_1, X_2)|X_3 = x_3 \sim \mathcal{N}_2(\mu_{12|3}, \Sigma_{12|3})$.

So now we have to apply the formula above:

$$\mu_{12|3} := \mu_{12} + \Sigma_{12}\Sigma_{22}^{-1}(x_3 - \mu_3) = \tag{13}$$
$$\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} + \begin{pmatrix}-\frac{1}{16}\\ -\frac{1}{16}\end{pmatrix}\frac{16}{3}(x_3 - \mu_3) = \tag{14}$$
$$\begin{pmatrix}0\\ 0\end{pmatrix} + \begin{pmatrix}-\frac{1}{3}\\ -\frac{1}{3}\end{pmatrix} x_3 = \begin{pmatrix}-\frac{1}{3}x_3\\ -\frac{1}{3}x_3\end{pmatrix} \tag{15}$$

$$\Sigma_{12|3} := \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \tag{16}$$
$$= \begin{pmatrix}\frac{3}{16} & -\frac{1}{16}\\ -\frac{1}{16} & \frac{3}{16}\end{pmatrix} - \begin{pmatrix}-\frac{1}{16}\\ -\frac{1}{16}\end{pmatrix}\frac{16}{3}\begin{pmatrix}-\frac{1}{16} & -\frac{1}{16}\end{pmatrix} \tag{17}$$
$$= \begin{pmatrix}\frac{3}{16} & -\frac{1}{16}\\ -\frac{1}{16} & \frac{3}{16}\end{pmatrix} - \begin{pmatrix}\frac{1}{48} & \frac{1}{48}\\ \frac{1}{48} & \frac{1}{48}\end{pmatrix} \tag{18}$$
$$= \begin{pmatrix}\frac{1}{6} & -\frac{1}{12}\\ -\frac{1}{12} & \frac{1}{6}\end{pmatrix} \tag{19}$$

$$(X_1, X_2)|X_3 = x_3 \sim \mathcal{N}_2\left(\begin{bmatrix}-\frac{1}{3}x_3\\ -\frac{1}{3}x_3\end{bmatrix}, \begin{bmatrix}\frac{1}{6} & -\frac{1}{12}\\ -\frac{1}{12} & \frac{1}{6}\end{bmatrix}\right)$$

We now that

$$X_3|(X_1 = x_1, X_2 = x_2) \sim \mathcal{N}_1(\mu_{3|12}, \Sigma_{3|12})$$

now we define we get the mean vector is

$$\mu_{3|1,2} = \mu_3 + \Sigma_{12}\Sigma_{11}^{-1}\begin{pmatrix}x_1 - \mu_1\\ x_2 - \mu_2\end{pmatrix} \tag{20}$$
$$= 0 + \begin{pmatrix}-\frac{1}{16} & -\frac{1}{16}\end{pmatrix} 32 \begin{pmatrix}\frac{3}{16} & \frac{1}{16}\\ \frac{1}{16} & \frac{3}{16}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} \tag{21}$$
$$= \begin{pmatrix}-\frac{1}{16} & -\frac{1}{16}\end{pmatrix}\underbrace{\begin{pmatrix}6 & 2\\ 2 & 6\end{pmatrix}}_{\Sigma_{11}^{-1}}\begin{pmatrix}x_1\\ x_2\end{pmatrix} \tag{22}$$
$$= \begin{pmatrix}-\frac{1}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} \tag{23}$$
$$= -\frac{1}{2}x_1 - \frac{1}{2}x_2 \tag{24}$$

$$\Sigma_{3|12} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \tag{25}$$
$$= \frac{3}{16} - \begin{pmatrix}-\frac{1}{16} & -\frac{1}{16}\end{pmatrix}\begin{pmatrix}6 & 2\\ 2 & 6\end{pmatrix}\begin{pmatrix}-\frac{1}{16}\\ -\frac{1}{16}\end{pmatrix} \tag{26}$$
$$= \frac{3}{16} - \begin{pmatrix}-\frac{1}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}-\frac{1}{16}\\ -\frac{1}{16}\end{pmatrix} \tag{27}$$
$$= \frac{3}{16} - \frac{1}{16} = \frac{1}{8} \tag{28}$$

$$X_3|(X_1 = x_1, X_2 = x_2) \sim \mathcal{N}\left(-\frac{1}{2}x_1 - \frac{1}{2}x_2, \frac{1}{8}\right)$$