1. We can express $\boldsymbol{A}$ and $\boldsymbol{B}$ in terms of vectors:

$$\underset{3\times 2}{\boldsymbol{A}} = \begin{bmatrix} 1 & 5 \\ 4 & -1 \\ 0 & 1 \end{bmatrix} = \left[\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \\ 1 \end{pmatrix} \right] = [\boldsymbol{a}_1,\ \boldsymbol{a}_2],$$

$$\underset{2\times 4}{\boldsymbol{B}} = \begin{bmatrix} 3 & 3 & -2 & 4 \\ 5 & -1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} (3 & 3 & -2 & 4) \\ (5 & -1 & 2 & 3) \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}'_{(1)} \\ \boldsymbol{b}'_{(2)} \end{bmatrix}.$$

Then $\boldsymbol{AB}$ is a sum of two matrices:

$$\begin{aligned} \underset{3\times 4}{\boldsymbol{AB}} &= [\boldsymbol{a}_1, \boldsymbol{a}_2] \begin{bmatrix} \boldsymbol{b}'_{(1)} \\ \boldsymbol{b}'_{(2)} \end{bmatrix} = \sum_{i=1}^{2} \boldsymbol{a}_i \boldsymbol{b}'_{(i)} \\ &= \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 & -1 & 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3 & -2 & 4 \\ 12 & 12 & -8 & 16 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 25 & -5 & 10 & 15 \\ -5 & 1 & -2 & -3 \\ 5 & -1 & 2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 28 & -2 & 8 & 19 \\ 7 & 13 & -10 & 13 \\ 5 & -1 & 2 & 3 \end{bmatrix}. \end{aligned}$$

2. Given the matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 5 & 2 \\ 3 & -1 & 7 \\ 5 & -7 & 12 \end{bmatrix}$$

(a) By silde 15 of set 1,

1. Note that the matrix $\boldsymbol{A}$ has rank 2 because

$$\boldsymbol{A} = \begin{bmatrix} 1 & 5 & 2 \\ 3 & -1 & 7 \\ 5 & -7 & 12 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & 2 \\ 0 & -16 & 1 \\ 0 & -32 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & 2 \\ 0 & -16 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Find a $r \times r$ nonsingular submatrix of $\boldsymbol{A}$ where $r = rank(\boldsymbol{A}) = 2$. Call this matrix $\boldsymbol{W}$.

$$\boldsymbol{W} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$$

2. Compute $(\boldsymbol{W}^{-1})'$.

$$(\boldsymbol{W}^{-1})' = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$$

3. Replace each element of $\boldsymbol{W}$ in $\boldsymbol{A}$ with the corresponding element of $(\boldsymbol{W}^{-1})'$. Then the corresponding matrix is

$$\begin{bmatrix} 7 & 5 & -3 \\ -2 & -1 & 1 \\ 5 & -7 & 12 \end{bmatrix}.$$

4. Replace all other elements in $\boldsymbol{A}$ with zeros. The resulting matrix is

$$\begin{bmatrix} 7 & 0 & -3 \\ -2 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

5. Transpose the resulting matrix to obtain $\boldsymbol{G}$, a generalized inverse for $\boldsymbol{A}$.

$$\boldsymbol{G} = \begin{bmatrix} 7 & -2 & 0 \\ 0 & 0 & 0 \\ -3 & 1 & 0 \end{bmatrix}$$

(b) Use the `R` function `ginv` in the `MASS` package .

```
> A=matrix(c(1,5,2,3,-1,7,5,-7,12),nrow=3, byrow=T)
> MASS::ginv(A)
```

We can obtain a generalized inverse matrix

$$\boldsymbol{G}^* = \begin{bmatrix} 0.02777778 & 0.02070521 & 0.01363264 \\ 0.13888889 & 0.04448544 & -0.04991800 \\ 0.05555556 & 0.04510045 & 0.03464535 \end{bmatrix}.$$

Notice that the generalized inverses for a singular matrix (such as $\boldsymbol{A}$ in this problem) are not unique.

3. Suppose $u \sim t_m(\delta)$. Let y and w be independent random variables with $y \sim \mathcal{N}(\delta, 1)$ and $w \sim \chi^2_m$. By slide 33 of set 1,

$$\frac{y}{\sqrt{w/m}} \sim t_m(\delta).$$

Since both u and the random variable defined by $y/\sqrt{w/m}$ have $t_m(\delta)$ distributions, they must have the same cdf, and thus they are equal in distribution:

$$u \overset{\mathrm{d}}{=} \frac{y}{\sqrt{w/m}}.$$

Squaring, we obtain

$$u^2 \overset{\mathrm{d}}{=} \frac{y^2}{w/m}.$$

Since $y \sim \mathcal{N}(\delta, 1)$, we have $y^2 \sim \chi^2_1(\delta^2/2)$ by slide 29 of set 1. Further, since y and w are independent, y^2 and w are also independent (see Theorem 4.3.5 in Casella and Berger[1]). By slide 35 of set 1, it follows that

$$u^2 \overset{\mathrm{d}}{=} \frac{y^2}{w/m} = \frac{y^2/1}{w/m} \sim F_{1,m}(\delta^2/2).$$

[1] Casella, G. and Berger, R.L. (2002) *Statistical Inference*, 2nd edition, Duxubry Press.

Hence, $u^2 \sim F_{1,m}(\delta^2/2)$.

Comments:

- Many of you forgot about the independence condition! This is a very important condition, highlighted by this counterexample. Suppose $y \sim \mathcal{N}(0,1)$. Let $w = y^2$ and $u = y/\sqrt{w/1}$. By slide 30 of set 1, $w \sim \chi_1^2$. Then we have

$$u = \frac{y}{\sqrt{w/1}}, \qquad \text{where } y \sim \mathcal{N}(0,1) \text{ and } w \sim \chi_1^2. \tag{1}$$

  Since $w = y^2$,

$$u = \frac{y}{\sqrt{w/1}} = \frac{y}{\sqrt{y^2/1}} = \frac{y}{|y|} = \begin{cases} -1 & \text{with probability } 0.5 \\ 1 & \text{with probability } 0.5 \end{cases}$$

  Obviously u does not have a t_1 distribution; in fact, u isn't even a continuous random variable. Thus, we've found that $u \not\sim t_1$ despite having (1) hold.

4. Since $x \sim \mathcal{N}(2,1)$ independent of $y \sim \mathcal{N}(1,1)$, together they are multivariate normal:

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right),$$

where $\mathrm{Cov}(x,y) = \mathrm{Cov}(y,x) = 0$ since $x \perp y$.

Put $\boldsymbol{\mu} = [2,1]'$ so that we can write $[x,y]' \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I}_{2\times 2})$. Using the results on non-central chi-square distributions (slide 29 of set 1),

$$x^2 + y^2 = [x,y]\begin{bmatrix} x \\ y \end{bmatrix} \sim \chi_2^2(\boldsymbol{\mu}'\boldsymbol{\mu}/2),$$

where the non-centrality parameter reduces to

$$\boldsymbol{\mu}'\boldsymbol{\mu}/2 = \frac{1}{2}[2,1]\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{5}{2}.$$

This gives us $x^2 + y^2 \sim \chi_2^2(2.5)$, so that

$$\mathrm{P}\left(\sqrt{x^2+y^2} > 6\right) = \mathrm{P}\left(x^2+y^2 > 36\right) = \mathrm{P}\left(\chi_2^2(2.5) > 36\right).$$

We can compute the above probability in `R` using the `pchisq()` function. The documentation brought up by `?pchisq` says

> "The non-central chi-squared distribution with df $= n$ degrees of freedom and non-centrality parameter ncp $= \lambda$ [...] this is the distribution of the sum of squares of n normals each with variance one, λ being the sum of squares of the normal means."

This implies `R` parameterizes the non-centrality parameter using $\boldsymbol{\mu}'\boldsymbol{\mu} = \mu_1^2 + \cdots + \mu_n^2$ (this is the sum of squares of normal means) rather than $\boldsymbol{\mu}'\boldsymbol{\mu}/2$. So, we need to double the non-centrality parameter we obtained above when inputing the function arguments. The code

```
> pchisq(36, df = 2, ncp = 5, lower.tail = FALSE)
```

gives the result

$$\mathrm{P}\left(\sqrt{x^2 + y^2} > 6\right) = 0.00014096.$$

Comments: Some students did a Monte Carlo experiment to approximate this probability; others did a change of variable transformation or used a moment generating function to establish the distribution of x^2+y^2. Noticing that we can write x^2+y^2 in terms of a multivariate normal vector and applying the results in the slide set, as done above, is not an approximation (as Monte Carlo methods are), and may be easier than a change of variable transformation or using a moment generating function.

5. We have $z_1, z_2 \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$. By slide 25 of set 1,

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim \mathcal{N}(\mathbf{0}_{2\times 2}, \boldsymbol{I}_{2\times 2}).$$

   (a) Let $\boldsymbol{A} = \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right]$. By slide 26 of set 1,

$$\frac{1}{\sqrt{2}}(z_1 - z_2) = \frac{1}{\sqrt{2}}\left[1, -1\right]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \left[\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \boldsymbol{A}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim \mathcal{N}(0, \boldsymbol{A}\boldsymbol{A}').$$

   Since $\boldsymbol{A}\boldsymbol{A}' = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}}\right) \cdot \left(-\frac{1}{\sqrt{2}}\right) = 1$, we have

$$\frac{1}{\sqrt{2}}(z_1 - z_2) \sim \mathcal{N}(0, 1).$$

   Squaring, by slide 30 of set 1, we obtain a central chi-square random variable:

$$\frac{1}{2}(z_1 - z_2)^2 \sim \chi_1^2.$$

   (b) Notice that

$$\frac{z_1 + z_2}{|z_1 - z_2|} = \frac{\frac{1}{\sqrt{2}}(z_1 + z_2)}{\sqrt{\frac{1}{2}(z_1 - z_2)^2/1}}. \tag{2}$$

   Similar to the first step in part (a), the numerator in (2) is standard normal:

$$\frac{1}{\sqrt{2}}(z_1 + z_2) \sim \mathcal{N}(0, 1).$$

   From the result in part (a), the part of the denominator in (2) under the square root is a central chi-square on one degree of freedom.

To have a t_1 distribution, we need to show that the random variables in the numerator and under the square root in the denominator in (2) are independent. We can do this by using slide 38 of slide set 1. Let $\boldsymbol{z} = [z_1, z_2]'$. Then $\frac{1}{\sqrt{2}}(z_1 + z_2) = \boldsymbol{A}_1 \boldsymbol{z}$ where $\boldsymbol{A}_1 = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$. Also,

$$\frac{1}{2}(z_1 - z_2)^2 = \frac{1}{2} \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \boldsymbol{z}' \boldsymbol{A}_2 \boldsymbol{z}$$

where

$$\boldsymbol{A}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Now note that $\boldsymbol{A}_1 \boldsymbol{I}_{2\times 2} \boldsymbol{A}_2 = \boldsymbol{0}_{1\times 2}$. Consequently, $\frac{1}{\sqrt{2}}(z_1 + z_2)$ is independent of $\frac{1}{2}(z_1 - z_2)^2$ by slide 38 of slide set 1. By the result on slide 34 of set 1,

$$\frac{z_1 + z_2}{|z_1 - z_2|} = \frac{\frac{1}{\sqrt{2}}(z_1 + z_2)}{\sqrt{\frac{1}{2}(z_1 - z_2)^2/1}} \sim t_1.$$

Comments:

- Many students failed to check the independence condition needed to establish a t_1 distribution in (b). See the comment section in question 1 for the potential perils of doing so.
- Some students did a change of variable transformation or used a moment generating function. Others showed that $\frac{z_1+z_2}{|z_1-z_2|}$ is a ratio of independent standard normal random variables and hence Cauchy(0,1), which is the same distribution as t_1.

6. We are given $y_1, \ldots, y_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$.

(a) Using properties of matrix transposition,

$$
\begin{aligned}
s^2 &= \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}_.)^2 \\
&= \frac{1}{n-1} (\boldsymbol{y} - \bar{y}_. \mathbf{1}_{n\times 1})'(\boldsymbol{y} - \bar{y}_. \mathbf{1}_{n\times 1}) \\
&= \frac{1}{n-1} \left(\boldsymbol{y} - \frac{1}{n} \sum_{i=1}^{n} y_i \mathbf{1}_{n\times 1} \right)' \left(\boldsymbol{y} - \frac{1}{n} \sum_{i=1}^{n} y_i \mathbf{1}_{n\times 1} \right) \\
&= \frac{1}{n-1} \left(\boldsymbol{y} - \frac{1}{n} \mathbf{1}_{n\times 1} \sum_{i=1}^{n} y_i \right)' \left(\boldsymbol{y} - \frac{1}{n} \mathbf{1}_{n\times 1} \sum_{i=1}^{n} y_i \right) \\
&= \frac{1}{n-1} \left(\boldsymbol{y} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \boldsymbol{y} \right)' \left(\boldsymbol{y} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \boldsymbol{y} \right) \\
&= \frac{1}{n-1} \left[\left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right) \boldsymbol{y} \right]' \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right) \boldsymbol{y} \\
&= \frac{1}{n-1} \boldsymbol{y}' \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right)' \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right) \boldsymbol{y} \\
&= \boldsymbol{y}' \left[\frac{1}{n-1} \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right)' \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right) \right] \boldsymbol{y} \\
&= \boldsymbol{y}' \boldsymbol{B} \boldsymbol{y},
\end{aligned}
$$

where

$$
\boldsymbol{B} = \frac{1}{n-1} \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right)' \left(\boldsymbol{I}_{n\times n} - \frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} \right).
$$

Notice that $\boldsymbol{B} = \frac{1}{n-1} \left(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}} \right)$, because

$$
\begin{aligned}
\frac{1}{n} \mathbf{1}_{n\times 1} \mathbf{1}'_{n\times 1} &= \mathbf{1}_{n\times 1} (n^{-1}) \mathbf{1}'_{n\times 1} \\
&= \mathbf{1}_{n\times 1} (\mathbf{1}'_{n\times 1} \mathbf{1}_{n\times 1})^{-1} \mathbf{1}'_{n\times 1} \\
&= \boldsymbol{P}_{\mathbf{1}_{n\times 1}}.
\end{aligned}
$$

Comments: As seen above, we do not need to assume any model for $\boldsymbol{y}$ (e.g., we do not need the GMMNE on slide 16 of set 2) to show the existence of $\boldsymbol{B}$ such that $s^2 = \boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$.

(b) From part (a), we have $\boldsymbol{B} = \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\mathbf{1}_{n\times 1}})$. Since $y_i \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, it follows that

$$
[y_1, \ldots, y_n]' = \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),
$$

where $\boldsymbol{\mu} = \mu \mathbf{1}_{n\times 1}$ and $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_{n\times n}$. Clearly $\boldsymbol{\Sigma}$ is positive definite, assuming that $\sigma^2 > 0$:

for any non-zero $\boldsymbol{t} = [t_1, \dots, t_n]' \in \mathbb{R}^n$ (i.e., $\boldsymbol{t} \neq \boldsymbol{0}_{n\times 1}$), we have

$$\begin{aligned}
\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} &= \boldsymbol{t}'[\sigma^2 \boldsymbol{I}_{n\times n}]\boldsymbol{t} \\
&= \sigma^2 \boldsymbol{t}'\boldsymbol{I}_{n\times n}\boldsymbol{t} \\
&= \sigma^2 \boldsymbol{t}'\boldsymbol{t} \\
&= \sigma^2 (t_1^2 + \cdots + t_n^2) \\
&> 0
\end{aligned}$$

since $\boldsymbol{t} \neq \boldsymbol{0}_{n\times 1}$ implies $t_i \neq 0$ for at least one $i \in \{1, \dots, n\}$.

Set $\boldsymbol{A} = \frac{n-1}{\sigma^2}\boldsymbol{B}$. By properties of rank and the results on slide 5 of set 2 that pertain to the rank of an orthogonal projection matrix,

$$\begin{aligned}
\operatorname{rank}(\boldsymbol{A}) &= \operatorname{rank}\left(\frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\right) \\
&= \operatorname{rank}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \operatorname{tr}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \operatorname{tr}(\boldsymbol{I}_{n\times n}) - \operatorname{tr}(\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= n - \operatorname{rank}(\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= n - \operatorname{rank}(\boldsymbol{1}_{n\times 1}) \\
&= n - 1.
\end{aligned}$$

As an orthogonal projection matrix, $\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is symmetric, and clearly $\boldsymbol{I}_{n\times n}$ is symmetric. Hence $\boldsymbol{A}$ is also symmetric. We also have that $\boldsymbol{A}\boldsymbol{\Sigma}$ is idempotent:

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma} &= \frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2 \boldsymbol{I}_{n\times n}\right]\frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2 \boldsymbol{I}_{n\times n}\right] \\
&= (\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= (\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}) \\
&= \frac{n-1}{\sigma^2} \cdot \frac{1}{n-1}(\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}})\left[\sigma^2 \boldsymbol{I}_{n\times n}\right] \\
&= \boldsymbol{A}\boldsymbol{\Sigma},
\end{aligned}$$

since $\boldsymbol{I}_{n\times n} - \boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is idempotent (this is easy to show using the fact that $\boldsymbol{P}_{\boldsymbol{1}_{n\times 1}}$ is idempotent).

We have now established all the ingredients that we need to apply the result on slide 31 of set 1. Hence,

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \frac{n-1}{\sigma^2}s^2 \sim \chi^2_{n-1}(\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2).$$

The non-centrality parameter reduces to zero:

$$\begin{aligned}
\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2 &= [\mu\mathbf{1}_{n\times1}]'\frac{n-1}{\sigma^2}\frac{1}{n-1}(\boldsymbol{I}_{n\times n}-\boldsymbol{P}_{\mathbf{1}_{n\times1}})\,[\mu\mathbf{1}_{n\times1}]\cdot\frac{1}{2} \\
&= \frac{\mu^2}{2\sigma^2}\mathbf{1}'_{n\times1}(\boldsymbol{I}_{n\times n}-\boldsymbol{P}_{\mathbf{1}_{n\times1}})\mathbf{1}_{n\times1} \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times1}\boldsymbol{I}_{n\times n}\mathbf{1}_{n\times1}-\mathbf{1}'_{n\times1}\boldsymbol{P}_{\mathbf{1}_{n\times1}}\mathbf{1}_{n\times1}) \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times1}\mathbf{1}_{n\times1}-\mathbf{1}'_{n\times1}\left[\mathbf{1}_{n\times1}(\mathbf{1}'_{n\times1}\mathbf{1}_{n\times1})^{-1}\mathbf{1}'_{n\times1}\right]\mathbf{1}_{n\times1}) \\
&= \frac{\mu^2}{2\sigma^2}(\mathbf{1}'_{n\times1}\mathbf{1}_{n\times1}-\mathbf{1}'_{n\times1}\mathbf{1}_{n\times1}) \\
&= \frac{\mu^2}{2\sigma^2}(n-n) \\
&= 0,
\end{aligned}$$

proving the desired result that

$$\frac{n-1}{\sigma^2}s^2\sim\chi^2_{n-1}.$$

Comments: Many students didn't establish all the conditions listed on slide 31, such as $\boldsymbol{A}$ is symmetric or that $\boldsymbol{y}$ is multivariate normal. A few students assumed the non-centrality parameter was zero without any mention of $\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2$, and others only showed minimal work in establishing that $\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2=0$.

7. (a) Consider a matrix

$$\boldsymbol{A}_{m\times n}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix},$$

so that

$$\begin{aligned}
\boldsymbol{A}'\boldsymbol{A} &= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm}\end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^m a_{i1}^2 & \sum_{i=1}^m a_{i1}a_{i2} & \cdots & \sum_{i=1}^m a_{i1}a_{in} \\ \sum_{i=1}^m a_{i1}a_{i2} & \sum_{i=1}^m a_{i2}^2 & \cdots & \sum_{i=1}^m a_{i2}a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{i1}a_{in} & \sum_{i=1}^m a_{i2}a_{in} & \cdots & \sum_{i=1}^m a_{in}^2\end{bmatrix}.
\end{aligned}$$

$\Rightarrow$: ["only if" part] Suppose that $\boldsymbol{A}=\mathbf{0}_{m\times n}$. Clearly then $\boldsymbol{A}'_{n\times m}=\mathbf{0}_{n\times m}$, which implies $\boldsymbol{A}'\boldsymbol{A}=\mathbf{0}_{n\times m}\mathbf{0}_{m\times n}=\mathbf{0}_{n\times n}$.

$\Leftarrow$: ["if" part] Suppose that $\boldsymbol{A}'\boldsymbol{A}=\mathbf{0}_{n\times n}$. This requires that the diagonal elements of $\boldsymbol{A}'\boldsymbol{A}$ are only zeros, that is, $\sum_{i=1}^m a_{ij}^2=0$ for $j=1,\ldots n$. Consequently, $a_{ij}=0$ for

$j = 1, \dots, n,\ i = 1, \dots, m$, which implies $\boldsymbol{A} = \boldsymbol{0}_{m \times n}$.

Thus, $\boldsymbol{A} = \boldsymbol{0} \iff \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{0}$.

Comments: This question, as well as the next one, requires a proof of an "if and only if" statement. This means you need to consider both directions (both the "if" and "only if" parts) to prove the desired conclusion. While it is trivial to many that $\boldsymbol{A} = \boldsymbol{0} \implies \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{0}$, if you don't include this in your proof, you can't claim to have proven the desired result. Additionally, your proof needs to hold for a matrix $\boldsymbol{A}$ with any dimension, not just 2×2.

(b) $\Leftarrow$: ["if" part] Suppose $\boldsymbol{XA} = \boldsymbol{XB}$. This part of the proof is trivial: multiplying both sides by $\boldsymbol{X}'$ gives $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB}$.

$\Rightarrow$: ["only if" part] Conversely, suppose $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB}$. By properties of matrix algebra and transpose, as well as the result of part (a), we have

$$\begin{aligned}
\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB} &\implies \boldsymbol{X}'\boldsymbol{XA} - \boldsymbol{X}'\boldsymbol{XB} = \boldsymbol{0} \\
&\implies \boldsymbol{X}'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \\
&\implies (\boldsymbol{A} - \boldsymbol{B})'\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \\
&\implies \big(\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B})\big)'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} \\
&\implies \boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0} && \text{by part (a)} \\
&\implies \boldsymbol{XA} - \boldsymbol{XB} = \boldsymbol{0} \\
&\implies \boldsymbol{XA} = \boldsymbol{XB}.
\end{aligned}$$

Therefore, $\boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB} \iff \boldsymbol{XA} = \boldsymbol{XB}$.

Comments: As in part (a), a proof of an "if and only if" statement requires both directions to be complete. Additionally, note that depending on the matrix dimensions, we may have

$$\big(\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B})\big)'\boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{0}_{n \times n} \neq \boldsymbol{0}_{m \times n} = \boldsymbol{X}(\boldsymbol{A} - \boldsymbol{B}),$$

where $m, n \in \{1, 2, \dots\}$.

(c) Let $(\boldsymbol{X}'\boldsymbol{X})^-$ be any generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$, which by definition implies

$$\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{X},$$

where $\boldsymbol{X}$ has dimension $m \times n$, say. Put $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}$ and $\boldsymbol{B} = \boldsymbol{I}_{n \times n}$, so that

$$\boldsymbol{X}'\boldsymbol{X}\underbrace{(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}}_{\boldsymbol{A}} = \boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{X}\underbrace{\boldsymbol{I}_{n \times n}}_{\boldsymbol{B}} \implies \boldsymbol{X}'\boldsymbol{XA} = \boldsymbol{X}'\boldsymbol{XB}.$$

By the result of part (b), it follows that $\boldsymbol{XA} = \boldsymbol{XB}$, and hence

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}.$$

(d) Let $\boldsymbol{A}$ be any symmetric matrix and $\boldsymbol{G}$ be any generalized inverse of $\boldsymbol{A}$. By definition,

$$\boldsymbol{AGA} = \boldsymbol{A}.$$

Now, transpose both sides and use the fact that $\boldsymbol{A'} = \boldsymbol{A}$ by symmetry:

$$\begin{aligned}(\boldsymbol{AGA})' = \boldsymbol{A}' &\implies \boldsymbol{A}'\boldsymbol{G}'\boldsymbol{A}' = \boldsymbol{A}' \\ &\implies \boldsymbol{A}\boldsymbol{G}'\boldsymbol{A} = \boldsymbol{A}.\end{aligned}$$

Hence, $\boldsymbol{G'}$ is a generalized inverse of $\boldsymbol{A}$.

(e) Let $\boldsymbol{G}$ be any generalized inverse of $\boldsymbol{X'X}$. Notice that $\boldsymbol{X'X}$ is symmetric, so by part (d), $\boldsymbol{G'}$ is also a generalized inverse of $\boldsymbol{X'X}$. The result of part (c) holds for any generalized inverse of $\boldsymbol{X'X}$, and hence holds using $\boldsymbol{G}'$. Using the result of part (c) with $\boldsymbol{G'}$ and then taking transposes gives

$$\begin{aligned}\boldsymbol{X}\boldsymbol{G}'\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X} &\implies (\boldsymbol{X}\boldsymbol{G}'\boldsymbol{X}'\boldsymbol{X})' = \boldsymbol{X}' \\ &\implies \boldsymbol{X}'[\boldsymbol{X}']'[\boldsymbol{G}']'\boldsymbol{X}' = \boldsymbol{X}' \\ &\implies \boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}' = \boldsymbol{X}'.\end{aligned}$$

Because we chose $\boldsymbol{G}$ to be any generalized inverse of $\boldsymbol{X'X}$,

$$\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' = \boldsymbol{X}'.$$

Comments: We could have

$$(\boldsymbol{X'X})^{-} \neq [(\boldsymbol{X'X})^{-}]',$$

so it is important that (c) and (d) are used at the right steps in your proof so it is clear that you aren't trying to say $(\boldsymbol{X'X})^{-} = [(\boldsymbol{X'X})^{-}]'$. On a related note, we may also have $[(\boldsymbol{X'X})^{-}]' \neq [(\boldsymbol{X'X})']^{-}$.

(f) This part requires *two* proofs that $\boldsymbol{P_X}$ is idempotent for full credit.

1. By part (c),

$$\begin{aligned}\boldsymbol{P_X}\boldsymbol{P_X} &= \underbrace{\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}}_{\boldsymbol{X}}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' \\ &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' \\ &= \boldsymbol{P_X}.\end{aligned}$$

2. By part (e),

$$\begin{aligned}\boldsymbol{P_X}\boldsymbol{P_X} &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\underbrace{\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'}_{\boldsymbol{X}'} \\ &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' \\ &= \boldsymbol{P_X}.\end{aligned}$$

(g) Let $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ be any generalized inverses of $\boldsymbol{X}'\boldsymbol{X}$. By parts (c) and (e), we have

$$\begin{aligned}
\boldsymbol{X}\boldsymbol{G}_1\boldsymbol{X}' &= \boldsymbol{X}\boldsymbol{G}_1\underbrace{\boldsymbol{X}'\boldsymbol{X}\boldsymbol{G}_2\boldsymbol{X}'}_{\boldsymbol{X}'} && \text{part (e) holds for any generalized inverse of } \boldsymbol{X}'\boldsymbol{X} \\
&= \underbrace{\boldsymbol{X}\boldsymbol{G}_1\boldsymbol{X}'\boldsymbol{X}}_{\boldsymbol{X}}\boldsymbol{G}_2\boldsymbol{X}' && \text{part (c) holds for any generalized inverse of } \boldsymbol{X}'\boldsymbol{X} \\
&= \boldsymbol{X}\boldsymbol{G}_2\boldsymbol{X}'.
\end{aligned}$$

Comments: A few students tried to use the fact that $\boldsymbol{P}_{\boldsymbol{X}}$ is the same matrix regardless of which generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$ is used, but this is what we are trying to show.

(h) Let $(\boldsymbol{X}'\boldsymbol{X})^-$ be any generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$. We know that $\boldsymbol{X}'\boldsymbol{X}$ is a symmetric matrix, so the result of part (d) says that if $(\boldsymbol{X}'\boldsymbol{X})^-$ is a generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$, then $[(\boldsymbol{X}'\boldsymbol{X})^-]'$ is a generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$. The result of part (g) then establishes that $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' = \boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{X}'$. Hence, these results and properties of matrix transpose give

$$\begin{aligned}
\boldsymbol{P}_{\boldsymbol{X}}' &= \left(\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\right)' \\
&= [\boldsymbol{X}']'[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{X}' \\
&= \boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^-]'\boldsymbol{X}' \\
&= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}' && \text{by parts (d,g) as explained above} \\
&= \boldsymbol{P}_{\boldsymbol{X}}.
\end{aligned}$$

Comments: It is important to use parts (d) and (g) at the right steps in your proof so it is clear that you aren't trying to say $(\boldsymbol{X}'\boldsymbol{X})^- = [(\boldsymbol{X}'\boldsymbol{X})^-]'$.