School of Statistics, University of the Philippines (Diliman)
Linangan ng Estadistika, Unibersidad ng Pilipinas (Diliman)

31 of October 2015

Al-Ahmadgaid B. Asaad
PS 2 | **Stat 233**

1. Let $\mathbf{Y} \sim \mathcal{N}_n(\boldsymbol{\mu}, \sigma^2\mathbf{I}), \sigma^2 > 0$. If $V \perp W$, then $P_V\mathbf{Y}$ and $P_W\mathbf{Y}$ are independent.

   Proof. From Theorem 2.14 (a), $P_V\mathbf{Y} \sim \mathcal{N}_n(P_V\boldsymbol{\mu}, \sigma^2 P_V)$ and $P_W\mathbf{Y} \sim \mathcal{N}_n(P_W\boldsymbol{\mu}, \sigma^2 P_W)$. If $\mathbf{A} = \begin{bmatrix} P_V \\ P_W \end{bmatrix}$, then $\mathbf{A}$ has $p \times n$ where $p = 2n$ dimension. From Theorem 2.7 (c),

$$\mathbf{AY} \sim \mathcal{N}_p(\mathbf{A}\boldsymbol{\mu}, \mathbf{A\Sigma A}'),$$

   where $\mathbf{A}\boldsymbol{\mu} = \begin{bmatrix} P_V \\ P_W \end{bmatrix}\boldsymbol{\mu} = \begin{bmatrix} P_V\boldsymbol{\mu} \\ P_W\boldsymbol{\mu} \end{bmatrix}$ and

$$\begin{aligned}
\mathbf{A\Sigma A}^{\mathrm{T}} &= \begin{bmatrix} P_V \\ P_W \end{bmatrix}\sigma^2\mathrm{I}\begin{bmatrix} P_V & P_W \end{bmatrix}\\
&= \sigma^2\begin{bmatrix} P_V \\ P_W \end{bmatrix}\begin{bmatrix} P_V & P_W \end{bmatrix}\\
&= \sigma^2\begin{bmatrix} P_V^2 & P_VP_W \\ P_WP_V & P_W^2 \end{bmatrix}\\
&= \sigma^2\begin{bmatrix} P_V & \mathbf{0} \\ \mathbf{0} & P_W \end{bmatrix}, \text{ by Theorem 1.4 (g) and since } P_V \text{ and } P_W \\
&\qquad \text{are idempotent.}
\end{aligned}$$

   And because $\mathrm{cov}(P_V, P_W) = \mathbf{0}$, then $P_V\mathbf{Y}$ and $P_W\mathbf{Y}$ are independent. □

2. Let $\mathbf{Y} \sim \mathcal{N}_n(\boldsymbol{\mu}, \sigma^2\mathbf{I}), \sigma^2 > 0$. Let $\mathbf{C}$ and $\mathbf{D}$ be $n \times n$ nonnegative definite matrices. If $\mathbf{CD} = \mathbf{0}$, then $\mathbf{Y}'\mathbf{CY}$ and $\mathbf{Y}'\mathbf{DY}$ are independent.

   Proof. Let $\mathbf{C}$ and $\mathbf{D}$ have rank r, and define $\mathbf{E}$ and $\mathbf{F}$ be matrices of dimension $r \times n$, such that $\mathbf{C} = \mathbf{E}'\mathbf{E}$ and $\mathbf{D} = \mathbf{F}'\mathbf{F}$. If $\mathbf{CD} = \mathbf{0}$, then

$$\begin{aligned}
(\mathbf{E}'\mathbf{E})(\mathbf{F}'\mathbf{F}) &= \mathbf{0} && \text{(dimension: } n \times n)\\
\mathbf{E}(\mathbf{E}'\mathbf{E})(\mathbf{F}'\mathbf{F}) &= \mathbf{E0} = \mathbf{0} && \text{(dimension: } r \times n)\\
(\mathbf{EE}')\mathbf{E}(\mathbf{F}'\mathbf{F})\mathbf{F}' &= \mathbf{0F}' = \mathbf{0} && \text{(dimension: } r \times r)\\
(\mathbf{EE}')^{-1}(\mathbf{EE}')\mathbf{EF}'(\mathbf{FF}') &= (\mathbf{EE}')^{-1}\mathbf{0} = \mathbf{0} && (\mathbf{EE}' \text{ is invertible.})\\
\mathbf{EF}'(\mathbf{FF}')(\mathbf{FF}')^{-1} &= \mathbf{0}(\mathbf{FF}')^{-1} = \mathbf{0} && (\mathbf{FF}' \text{ is invertible.})\\
\mathbf{EF}' &= \mathbf{0}.
\end{aligned}$$

   Then by Theorem 2.16 (a), $\mathbf{EY}$ and $\mathbf{FY}$ are independent. So that,

$$(\mathbf{EY})'(\mathbf{EY}) = \mathbf{Y}'\mathbf{E}'\mathbf{EY} = \mathbf{Y}'\mathbf{CY},$$

   and

$$(\mathbf{FY})'(\mathbf{FY}) = \mathbf{Y}'\mathbf{F}'\mathbf{FY} = \mathbf{Y}'\mathbf{DY},$$

   are independent. □

3. $(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2)$ is a complete sufficient statistics for the coordinatized linear model. $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are independent and

$$\hat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}), (n-p)\hat{\sigma}^2 \sim \sigma^2\chi^2_{n-p}(0).$$

   Proof. Let $\mathbf{X}$ be a known $n \times p$ basis matrix of the subspace V. The distribution of the response variable $\mathbf{Y}$, of the coordinatized version of the general linear model is given by:

$$\begin{aligned}
f(\mathbf{y}|\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}) &= \frac{1}{(2\pi)^{n/2}|\sigma^2\mathbf{I}|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\left(\sigma^2\mathbf{I}\right)^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right]\\
&= \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2\right]
\end{aligned} \tag{1}$$

   By Pythagorean theorem,

$$\begin{aligned}
\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2 &= \|\mathbf{y}-P_V\mathbf{y}\|^2 + \|P_V\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2\\
&= \|\mathbf{y}-\hat{\boldsymbol{\mu}}\|^2 + \|\hat{\boldsymbol{\mu}}-\mathbf{X}\boldsymbol{\beta}\|^2.
\end{aligned}$$

   Define $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, by Theorem 1.3 (c)

$$\hat{\boldsymbol{\mu}} = P_V\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}},$$

and

$$\begin{aligned} \mathbf{X}'\hat{\boldsymbol{\mu}} &= \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\mu}} \end{aligned} \tag{2}$$

Also $(n-p)\sigma^2 = \|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2$ (follows from Eq. 5.1 of Arnold 1981), thus

$$\begin{aligned} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 &= \|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2 + \|\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta}\|^2 \\ &= (n-p)\sigma^2 + \|\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta}\|^2 \\ &= (n-p)\sigma^2 + (\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}\boldsymbol{\beta}) \\ &= (n-p)\sigma^2 + \|\mathbf{X}\hat{\boldsymbol{\beta}}\|^2 - 2(\mathbf{X}\boldsymbol{\beta})'(\mathbf{X}\hat{\boldsymbol{\beta}}) + \|\mathbf{X}\boldsymbol{\beta}\|^2. \end{aligned}$$

By exponential criterion, the pdf (1) can be factorized into the following components:

$$h(\mathbf{y}) = 1, \qquad k(\mathbf{X}\boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{\|\mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}\right],$$

$$Q(\mathbf{X}\boldsymbol{\beta}, \sigma^2) = \begin{bmatrix} -\dfrac{1}{2\sigma^2} \\ \dfrac{\mathbf{X}\boldsymbol{\beta}}{\sigma^2} \end{bmatrix}, \qquad S(\mathbf{y}) = \begin{bmatrix} (n-p)\hat{\sigma}^2 + \|\mathbf{X}\hat{\boldsymbol{\beta}}\|^2 \\ \mathbf{X}\hat{\boldsymbol{\beta}} \end{bmatrix},$$

and since $\mathbf{X}$ spans $V \subset \mathbb{R}^p$ then $\mathbf{X}\boldsymbol{\beta}$ lies in the subspace V which is a hyperplane, and so it has an (open) p-ball of radius $\varepsilon > 0$ centered at point x – the interior point in $\mathbb{R}^p$. It follows that, $Q(\mathbf{X}\boldsymbol{\beta}, \sigma^2)$ lies in the space spanned by $\mathbb{R}^{p+1}$, that is the image of the parameter space under Q is the set of all vectors in $\mathbb{R}^{p+1}$ with (open) $(p+1)$-ball, so $\exists$ an interior point x with ε-neighbourhood for every unique point $Q(\mathbf{X}\boldsymbol{\beta}, \sigma^2)$. And if

$$\begin{aligned} (n-p)\hat{\sigma}^2 + \|\mathbf{X}\hat{\boldsymbol{\beta}}\|^2 &= T_1(\mathbf{y}) \\ \mathbf{X}\hat{\boldsymbol{\beta}} &= T_2(\mathbf{y}) \end{aligned}$$

then

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'T_2(\mathbf{y}) \quad \text{(from Equation (2))},$$

so that

$$\hat{\sigma}^2 = \frac{T_1(\mathbf{y}) - \|\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'T_2(\mathbf{y})\|^2}{n-p}.$$

$\begin{pmatrix} \hat{\sigma}^2 \\ \hat{\boldsymbol{\beta}} \end{pmatrix}$ is an invertible function of $S(\mathbf{y})$, and so $\begin{pmatrix} \hat{\sigma}^2 \\ \hat{\boldsymbol{\beta}} \end{pmatrix}$ is a complete sufficient statistics.

Now $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}$ are independent because $\hat{\boldsymbol{\mu}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ lies in V and $\hat{\sigma}$ lies in $V^\perp$, which by Theorem 2.14 (b), $P_V\mathbf{Y}$ and $P_{V^\perp}\mathbf{Y}$ are independent since $V \perp V^\perp$. So that $\hat{\boldsymbol{\beta}} \perp\!\!\!\perp \hat{\sigma}$. And by Theorem 2.14 (a)

$$\hat{\boldsymbol{\mu}} = P_V\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} \sim \mathcal{N}_n(P_V\mathbf{X}\boldsymbol{\beta}, \sigma^2 P_V)$$

where

$$\begin{aligned} P_V\mathbf{X}\boldsymbol{\beta} &= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{X}\boldsymbol{\beta} \end{aligned}$$

and

$$\sigma^2 P_V = \sigma^2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

thus,

$$\hat{\boldsymbol{\mu}} = \mathbf{X}\hat{\boldsymbol{\beta}} \sim \mathcal{N}_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}').$$

Now $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\mu}} = \mathbf{A}\hat{\boldsymbol{\mu}}$, where $\mathbf{A}$ is a $p \times n$ matrix. By Theorem 2.7 (c)

$$\hat{\boldsymbol{\beta}} = \mathbf{A}\hat{\boldsymbol{\mu}} \sim \mathcal{N}_p(\mathbf{A}\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{A}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}')$$

where

$$\mathbf{A}\mathbf{X}\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta},$$

and

$$\begin{aligned} \sigma^2\mathbf{A}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}' &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}]' \\ &= \sigma^2[(\mathbf{X}'\mathbf{X})']^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \end{aligned}$$

Therefore,

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}).$$

On the other hand by Theorem 3.12,

$$\|P_{V^\perp}\mathbf{Y}\|^2 \sim \sigma^2\chi^2_{\dim(V^\perp)}\left(\frac{\|P_{V^\perp}\mathbf{X}\boldsymbol{\beta}\|^2}{\sigma^2}\right),$$

and since $\dim(V^\perp) = n-p$, and $\mathbf{X}\boldsymbol{\beta} \in V$, then $P_{V^\perp}\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$. Suggesting

$$(n-p)\hat{\sigma}^2 = \|P_{V^\perp}\mathbf{Y}\|^2 \sim \chi^2_{n-p}(\mathbf{0}).$$

$\square$

4. $\mathbf{C}\hat{\boldsymbol{\beta}}$ is MVUE and MLE of $\mathbf{C}\boldsymbol{\beta}$.

Proof. By Theorem 4.4, $S(\mathbf{X}) = (\hat{\boldsymbol{\beta}}, \hat{\sigma}^2)$ is a complete sufficient statistics for the coordinatized linear model. By the same theorem, let $\mathrm{E}\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$, i.e. $\hat{\boldsymbol{\beta}}$ is unbiased for $\boldsymbol{\beta}$. By Lehmann-Scheffe Theorem, $\hat{\boldsymbol{\beta}}$ is the MVUE for $\boldsymbol{\beta}$. By invariance property of MLE, $g(\hat{\boldsymbol{\beta}})$ is an unbiased estimator of $\mathbf{C}\boldsymbol{\beta}$ where $g(\hat{\boldsymbol{\beta}}) = \mathbf{C}\hat{\boldsymbol{\beta}}$. Therefore if $T(S) = \mathbf{C}\hat{\boldsymbol{\beta}}$ then $\mathrm{E}T(S) = \mathrm{E}\mathbf{C}\hat{\boldsymbol{\beta}} = \mathbf{C}\boldsymbol{\beta}$. Then again by Lehmann-Scheffe, $T(S)$ is the MVUE of $\tau(\boldsymbol{\beta}) = \mathbf{C}\boldsymbol{\beta}$. □

5. Let $\mathbf{c} \in \mathbb{R}^n, \mathbf{c} \notin V^{\perp}$, and $\mathbf{d} \in \mathbb{R}^p, \mathbf{d} \neq 0$. Then

$$\frac{\mathbf{c}'\hat{\boldsymbol{\mu}} - \mathbf{c}'\boldsymbol{\mu}}{\hat{\sigma}\sqrt{\mathbf{c}'P_V\mathbf{c}}} \sim t_{n-p}(0), \quad \frac{\mathbf{d}'\hat{\boldsymbol{\beta}} - \mathbf{d}'\boldsymbol{\beta}}{\hat{\sigma}\sqrt{\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d}}} \sim t_{n-p}(0).$$

Proof. From Theorem 4.1 $\hat{\boldsymbol{\mu}} \sim \mathcal{N}_n\left(\boldsymbol{\mu}, \sigma^2 P_V\right)$. Let

$$W = \frac{\mathbf{c}'\hat{\boldsymbol{\mu}} - \mathbf{c}'\boldsymbol{\mu}}{\sqrt{\sigma^2\mathbf{c}'P_V\mathbf{c}}}$$

since $\mathbf{c}'\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\mathbf{c}'\boldsymbol{\mu}, \sigma^2\mathbf{c}'P_V\mathbf{c})$, then by location and scale transformation $W \sim \mathcal{N}(0,1)$. Also, from the same Theorem let

$$S = \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}(0),$$

then W and S are independent random variables. Therefore,

$$T = \frac{W}{\sqrt{S/(n-p)}} = \frac{\dfrac{\mathbf{c}'\hat{\boldsymbol{\mu}} - \mathbf{c}'\boldsymbol{\mu}}{\sqrt{\sigma^2\mathbf{c}'P_V\mathbf{c}}}}{\sqrt{\dfrac{\frac{(n-p)\hat{\sigma}^2}{\sigma^2}}{(n-p)}}} = \frac{\mathbf{c}'\hat{\boldsymbol{\mu}} - \mathbf{c}'\boldsymbol{\mu}}{\hat{\sigma}\sqrt{\mathbf{c}'P_V\mathbf{c}}} \sim t_{n-p}(0).$$

Similarly from Theorem 4.4, $\hat{\boldsymbol{\beta}} \sim \mathcal{N}_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$, let

$$G = \frac{\mathbf{d}'\hat{\boldsymbol{\beta}} - \mathbf{d}'\beta}{\sqrt{\sigma^2\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d}}},$$

since $\mathbf{d}'\hat{\boldsymbol{\beta}} \sim \mathcal{N}(\mathbf{d}\boldsymbol{\beta}, \sigma^2\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d})$, and by location-scale transformation $G \sim \mathcal{N}(0,1)$ is a univariate standard normal distribution. From the same Theorem let

$$H = \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}(0),$$

then G and H are independent, so that

$$K = \frac{G}{\sqrt{H/(n-p)}} = \frac{\dfrac{\mathbf{d}'\hat{\boldsymbol{\beta}} - \mathbf{d}'\beta}{\sqrt{\sigma^2\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d}}}}{\sqrt{\dfrac{\frac{(n-p)\hat{\sigma}^2}{\sigma^2}}{(n-p)}}} = \frac{\mathbf{d}'\hat{\boldsymbol{\beta}} - \mathbf{d}'\beta}{\hat{\sigma}\sqrt{\mathbf{d}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d}}} \sim t_{n-p}(0).$$

□