Γενικό Γραμμικό Μοντέλο

Μοντέλο #1: $\underline{y} = \underline{\mu} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim N_n(0, \sigma^2 I)$, με $V \subseteq \Re^n$, με dim$V = p < n$, σ^2>0.
Ορίζω $\hat{\mu} = \mathrm{P}_V \underline{y}$, $\hat{\sigma}^2 = \| \mathrm{P}_{V^\perp} \underline{y} \|^2 / (n-p)$.

Μοντέλο #2: $\underline{y} = \mathrm{X}\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim N_n(0, \sigma^2 I)$, $\underline{\beta} \in \Re^p$, rank$(\mathrm{X}) = p < n$, σ^2>0.

$$\hat{\underline{\beta}} = (\mathrm{X}^T \mathrm{X})^{-1} \mathrm{X}^T \underline{y}, \quad \hat{\sigma}^2 = \frac{\underline{y}^T (I - \mathrm{P}_V) \underline{y}}{n-p}$$

Παρατήρηση: 1) Για το μοντέλο #1 το εκτιμώμενο μοντέλο είναι $\hat{\underline{y}} = \hat{\underline{\mu}} = \mathrm{P}_V \underline{y}$
Τα υπόλοιπα: $\underline{e} = \underline{y} - \hat{\underline{y}} = \underline{y} - \underline{\mu} = \underline{y} - \mathrm{P}_V \underline{y}$
Το άθροισμα τετραγώνων } $SS_{res} = \| \underline{e} \|^2 = \| \underline{y} - \hat{\underline{y}} \|^2 = \| \underline{y} - \mathrm{P}_V \underline{y} \|^2$
των υπολοίπων } $= \| (I - \mathrm{P}_V)\, \underline{y} \|^2 = \| \mathrm{P}_{V^\perp} \underline{y} \|^2$.
Άρα $\hat{\sigma}^2 = SS_{res} / (n-p)$.
2) Για το μοντέλο #2 έχω αντίστοιχα ότι $\hat{\underline{y}} = \mathrm{X}\hat{\underline{\beta}}$, $\underline{e} = \underline{y} - \hat{\underline{y}} = \underline{y} - \mathrm{X}\hat{\underline{\beta}}$.
Το άθροισμα των τετραγώνων των υπολοίπων: $SS_{res} = \| \underline{e} \|^2 = \| \underline{y} - \mathrm{X}\hat{\underline{\beta}} \|^2$.

Ιδιότητες των εκτιμητριών όταν δεν ισχύει η υπόθεση της κανονικότηττας αλλά ισχύει:
Ε($\underline{\varepsilon}$)=0 και Var($\underline{\varepsilon}$)=$\sigma^2 I$.
Για το μοντέλο #1:
1) Η εκτιμήτρια $\hat{\underline{\mu}}$ είναι αμερόληπτη εκτιμήτρια του $\underline{\mu}$.
2) Η εκτιμήτρια $\hat{\sigma}^2$ είναι αμερόληπτη της σ^2.
3) Η εκτιμήτρια $\hat{\underline{\mu}}$ είναι εκτιμήτρια ελαχίστων τετραγώνων.
Απόδειξη:
1) Αφού $\hat{\mu} = \mathrm{P}_V \underline{y}$ έχω ότι Ε($\hat{\mu}$)=P_V Ε($\underline{y}$)= $\mathrm{P}_V \underline{\mu} = \underline{\mu}$, αφού $\underline{\mu} \in V$.

2) $$E(\hat{\sigma}^2) = E\frac{[\underline{y}^T (I - \mathrm{P}_V)\underline{y}]}{n-p} = \frac{\underline{\mu}^T (I - \mathrm{P}_V)\underline{\mu} + tr[(I - \mathrm{P}_V)\sigma^2 I]}{n-p}$$
$$= \frac{(\underline{\mu}^T \underline{\mu} - \underline{\mu}^T \underline{\mu}) + tr(I - \mathrm{P}_V)}{n-p} = \frac{n-p}{n-p} = 1,$$
αφού $tr(\mathrm{P}_V)$=rank(V) εφόσον $\mathrm{P}_V{}^2 = \mathrm{P}_V$.

3) Αναζητούμε $\underline{\mu} \in V$: $\| \underline{e} \|^2 = \sum_{i=1}^{n} e_i{}^2 = \| \underline{y} - \underline{\mu} \|$ να γίνεται ελάχιστο.
Αλλά $\min_{\underline{\mu} \in V} \| \underline{y} - \underline{\mu} \|^2 = \| \underline{y} - \mathrm{P}_V \underline{y} \|^2$, από τις ιδιότητες της προβολής.

Για το μοντέλο #2 ισχύει η παρακάτω πρόταση:

Πρόταση:
1) Η εκτιμήτρια $\hat{\underline{\beta}}$ είναι αμερόληπτη για το $\underline{\beta}$.
2) Η εκτιμήτρια $\hat{\sigma}^2$ είναι αμερόληπτη για το σ^2.

3) Η εκτιμήτρια $\hat{\underline{\beta}}$ είναι εκτιμήτρια ελαχίστων τετραγώνων.

<u>Απόδειξη:</u>

1) και 2) όπως προηγουμένως.

3) Αναζητούμε $\underline{\beta}$ έτσι ώστε $\|\underline{e}\|^2 = \|\underline{y} - \mathrm{X}\underline{\beta}\|^2$ να γίνεται ελάχιστο.

Αλλά $\min\limits_{\underline{\beta}\in\Re^p} \|\underline{y} - \mathrm{X}\underline{\beta}\|^2 = \min\limits_{\underline{\mu}\in V} \|\underline{y} - \underline{\mu}\|^2 = \|\underline{y} - \mathrm{P}_V\underline{y}\|^2 \, \|\underline{y} - \mathrm{X}\hat{\underline{\beta}}\|^2$,

όπου $\hat{\underline{\beta}} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\underline{y}$ □ $\hat{\underline{\beta}}$ είναι Ε.Ε.Τ. των $\underline{\beta}$.

Έστω τώρα $\underline{c}\in\Re^n$ δεδομένο. Ζητάμε να εκτιμήσουμε την ποσότητα $\underline{c}^T\underline{\mu}$.

Θα περιοριστώ στην κλάση των γραμμικών και αμερόληπτων εκτιμητριών $\mathrm{T}(y_1,\ldots,y_n)$:

α) $\mathrm{T}(y_1,\ldots,y_n) = \underline{b}^T\underline{y} = b_1y_1 + \ldots + b_ny_n = \sum_{i=1}^{n} b_iy_i$

β) $\mathrm{E}[\mathrm{T}(\underline{y})] = \underline{c}^T\underline{\mu}$, $\forall\underline{\mu}\in V$.

Τότε ισχύει το παρακάτω Θεώρημα.

<u>Θεώρημα Gauss-Markov:</u>

Η εκτιμήτρια $\underline{c}^T\hat{\underline{\mu}}$ είναι <u>**η**</u> μοναδική γραμμική αμερόληπτη εκτιμήτρια ελάχιστης διασποράς.

<u>Απόδειξη:</u>

1) Αποδεικνύουμε ότι $\underline{c}^T\hat{\underline{\mu}}$ ανήκει στην κλάση των γραμμικών αμερόληπτων εκτιμητριών:

$\underline{c}^T\hat{\underline{\mu}} = \underline{c}^T\,\mathrm{P}_V\underline{y} = (\mathrm{P}_V\underline{c})^T\underline{y} = \underline{b}\cdot\underline{y}$ (γραμμικότητα).

$\mathrm{E}[\underline{c}^T\hat{\underline{\mu}}] = \underline{c}^T\mathrm{E}(\hat{\underline{\mu}}) = \underline{c}^T\underline{\mu}$, $\forall\underline{\mu}\in V$.

2) Έστω $\mathrm{T}(\underline{y}) = \underline{a}^T\underline{y}$ αμερόληπτη και γραμμική εκτιμήτρια του $\underline{c}^T\underline{\mu}$.

Τότε, λόγω της αμεροληψίας:

$\mathrm{E}(\underline{a}^T\underline{y}) = \underline{c}^T\underline{\mu}$ □ $\underline{a}^T\underline{\mu} = \underline{c}^T\underline{\mu}$ □ $(\underline{a}-\underline{c})^T\cdot\underline{\mu} = 0, \forall\underline{\mu}$ □ $(\underline{a}-\underline{c})^T \perp V$.

Επομένως, $\mathrm{P}_V(\underline{a}-\underline{c}) = 0$, ή $\mathrm{P}_V\underline{a} = \mathrm{P}_V\underline{c}$.

$\mathrm{Var}(\underline{a}^T\underline{y}) = \underline{a}^T\sigma^2\underline{a} = \sigma^2\|\underline{a}\|^2 = \sigma^2\|\mathrm{P}_V\underline{a} + \mathrm{P}_{V^\perp}\underline{a}\|^2 = \sigma^2(\|\mathrm{P}_V\underline{a}\|^2 + \|\mathrm{P}_{V^\perp}\underline{a}\|^2)$

□ $\sigma^2\|\mathrm{P}_V\underline{a}\|^2 = \sigma^2\|\mathrm{P}_V\underline{c}\|^2 = \sigma^2\,\underline{c}^T\,\mathrm{P}_V\underline{c} = \mathrm{Var}(\underline{c}^T\hat{\underline{\mu}})$.

Αφού $\mathrm{Var}(\underline{c}^T\hat{\underline{\mu}}) = \underline{c}^T\mathrm{Var}(\hat{\underline{\mu}})\underline{c} = \underline{c}^T\,\sigma^2\,\mathrm{P}_V\underline{c}$.

Άρα, $\mathrm{Var}(\underline{a}^T\underline{y})$ □ $\mathrm{Var}(\underline{c}^T\hat{\underline{\mu}})$ □ ελάχιστη διακύμανση.

<u>Μοναδικότητα:</u> Αν υπήρχε άλλη εκτιμήτρια $\underline{a}^T\underline{y}$, τότε $\mathrm{Var}(\underline{a}^T\underline{y}) = \mathrm{Var}(\underline{c}^T\hat{\underline{\mu}})$

□ $\mathrm{P}_{V^\perp}\underline{a} = 0$, από την παραπάνω ανάλυση. □ $\underline{a}\in V$ □ $\underline{a} = \mathrm{P}_V\underline{a}$.

Συνεπώς $\underline{a}^T\underline{y} = (\mathrm{P}_V\underline{a})^T\underline{y} = (\mathrm{P}_V\underline{c})^T\underline{y} = \underline{c}\,\mathrm{P}_V\underline{y} = \underline{c}^T\hat{\underline{\mu}}$.

<u>Πόρισμα:</u>

Έστω $\underline{\mu} = \begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix}$, $\hat{\underline{\mu}} = \begin{pmatrix}\hat{\mu}_1\\ \vdots\\ \hat{\mu}_n\end{pmatrix}$. Τότε $\hat{\mu}_i$ είναι Α.Ο.Ε.Δ. των μ_i, κατά συνιστώσα.

Θέτω $\underline{c} = (0,\ldots,0,\underset{\uparrow\, i \text{ θέση}}{1},0,\ldots,0)^T$.