Περί μη-αρνητικότητας των BLUE και δύο αντιπαραδείγματα

(από τον Νίκο Παπαδάτο)

Ας υποθέσουμε ότι $\boldsymbol{Z}' = (Z_1, Z_2, \ldots, Z_n)$ $(n \geq 2)$ είναι ένα τυχαίο διάνυσμα με $Z_i \geq 0$, $i = 1, 2, \ldots, n$. Επίσης υποθέτουμε ότι

$$\mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu} \in \mathbb{R}^n, \quad \mathbb{D}[\boldsymbol{Z}] = \boldsymbol{\Sigma} \in \mathbb{R}^{n \times n},$$

όπου $\boldsymbol{\Sigma} > 0$ (positive definite). Εδώ τα $\boldsymbol{\mu}$ και $\boldsymbol{\Sigma}$ θεωρούνται γνωστά. Παρατηρούμε ότι, αναγκαστικά, $\mu_i = \mathbb{E}[Z_i] > 0$ (αλλιώς θα ήταν $\mathbb{P}[Z_i = 0] = 1$, το οποίο αντιβαίνει στην υπόθεση $\boldsymbol{\Sigma} > 0$).

Τώρα θεωρούμε ότι υπάρχει μία άγνωστη παράμετρος $\theta > 0$ (scale parameter), και ότι τα δεδομένα που παρατηρούμε περιγράφονται από το διάνυσμα

$$\boldsymbol{Y} = \theta \boldsymbol{Z},$$

από το οποίο προσπαθούμε να εκτιμήσουμε την θ.

Ορισμός. Η εκτιμήτρια L καλείται γραμμική αν είναι της μορφής $L = \boldsymbol{c}'\boldsymbol{Y} = \sum_{i=1}^n c_i Y_i$, όπου τα c_i είναι γνωστές σταθερές. Μία γραμμική εκτιμήτρια L καλείται BLUE αν είναι αμερόληπτη ($\mathbb{E}[L] = \theta$ για κάθε θ) και έχει ελάχιστη διασπορά (για κάθε θ), ενώ καλείται BLIE αν ελαχιστοποιεί το $\mathrm{MSE}[L] = \mathbb{E}[L - \theta]^2$ (για κάθε θ).

Θεώρημα 1. $\mathrm{BLUE} = \dfrac{1}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \stackrel{\mathrm{d}}{=} \dfrac{\theta}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}.$

Απόδειξη: Είναι $\mathbb{E}[L] = \mathbb{E}[\boldsymbol{c}'\boldsymbol{Y}] = \theta \mathbb{E}[\boldsymbol{c}'\boldsymbol{Z}] = \theta \cdot \boldsymbol{c}'\boldsymbol{\mu}$. Άρα $\mathbb{E}[L] = \theta$ (για κάθε θ) αν και μόνο αν $\boldsymbol{c}'\boldsymbol{\mu} = 1$. Τώρα $\mathrm{Var}[L] = \theta^2 \mathrm{Var}[\boldsymbol{c}'\boldsymbol{Z}] = \theta^2 \cdot \boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c}$, οπότε αρκεί να ελαχιστοποιηθεί η ποσότητα $\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c}$ υπό τον περιορισμό $\boldsymbol{c}'\boldsymbol{\mu} = 1$. Παρατηρούμε όμως ότι για τυχόν $\boldsymbol{c} \in \mathbb{R}^n$, με $\boldsymbol{c}'\boldsymbol{\mu} = 1$, ισχύει η ταυτότητα

$$\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c} = \frac{1}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} + \left(\boldsymbol{c} - \frac{1}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)' \boldsymbol{\Sigma} \left(\boldsymbol{c} - \frac{1}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right),$$

από την οποία προκύπτει άμεσα ότι το $\boldsymbol{c} = \dfrac{1}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ είναι το (μοναδικό) διάνυσμα που ελαχιστοποιεί την διασπορά της αμερόληπτης γραμμικής εκτιμήτριας L. (ο.ε.δ.)

Θεώρημα 2. Έστω $\boldsymbol{E} = \mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}'] = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'$ ο πίνακας δευτέρων ροπών του $\boldsymbol{Z}$ (που είναι γνωστός και positive definite). Τότε $\mathrm{BLIE} = \boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{Y} = \dfrac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{1 + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \mathrm{BLUE}.$

Απόδειξη: Είναι

$$\mathbb{E}[L - \theta]^2 = \mathbb{E}[\boldsymbol{c}'\boldsymbol{Y} - \theta]^2 = \theta^2 \mathbb{E}[\boldsymbol{c}'\boldsymbol{Z} - 1]^2 = \theta^2 \left\{ \mathbb{E}[\boldsymbol{c}'\boldsymbol{Z}]^2 - 2\boldsymbol{c}'\boldsymbol{\mu} + 1 \right\}.$$

Όμως $\mathbb{E}[\boldsymbol{c}'\boldsymbol{Z}]^2 = \mathbb{E}[\boldsymbol{c}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{c})'] = \mathbb{E}[\boldsymbol{c}'\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{c}] = \boldsymbol{c}'\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}']\boldsymbol{c} = \boldsymbol{c}'\boldsymbol{E}\boldsymbol{c}$, οπότε η προηγούμενη σχέση γίνεται

$$\mathbb{E}[L-\theta]^2 = \theta^2\left\{\boldsymbol{c}'\boldsymbol{E}\boldsymbol{c} - 2\boldsymbol{c}'\boldsymbol{\mu} + 1\right\}.$$

Παρατηρούμε ότι για τυχόν $\boldsymbol{c} \in \mathbb{R}^n$ ισχύει η ταυτότητα

$$\left(\boldsymbol{c} - \boldsymbol{E}^{-1}\boldsymbol{\mu}\right)' \boldsymbol{E}\left(\boldsymbol{c} - \boldsymbol{E}^{-1}\boldsymbol{\mu}\right) = \boldsymbol{c}'\boldsymbol{E}\boldsymbol{c} - 2\boldsymbol{c}'\boldsymbol{\mu} + \boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{\mu},$$

από την οποία προκύπτει άμεσα η εξής έκφραση για το MSE:

$$\mathbb{E}[L-\theta]^2 = \theta^2\left\{\boldsymbol{c}'\boldsymbol{E}\boldsymbol{c} - 2\boldsymbol{c}'\boldsymbol{\mu} + 1\right\} = \theta^2(1-\boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{\mu}) + \theta^2\left(\boldsymbol{c} - \boldsymbol{E}^{-1}\boldsymbol{\mu}\right)' \boldsymbol{E}\left(\boldsymbol{c} - \boldsymbol{E}^{-1}\boldsymbol{\mu}\right).$$

Επομένως το MSE ελαχιστοποιείται όταν και μόνο όταν $\boldsymbol{c} = \boldsymbol{E}^{-1}\boldsymbol{\mu}$, και συνεπώς, BLIE $= \boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{Y}$. Αφού $\boldsymbol{E} = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'$ έπεται ότι

$$\boldsymbol{E}^{-1} = \boldsymbol{\Sigma}^{-1} - \frac{1}{1+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)'.$$

[Για να αποδείξουμε αυτήν την σχέση είναι αρκετό να εκτελέσουμε τον πολλαπλασιασμό $(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')\boldsymbol{E}^{-1}$, με $\boldsymbol{E}^{-1}$ όπως παραπάνω, και θα διαπιστώσουμε ότι το αποτέλεσμα ισούται με τον $n \times n$ μοναδιαίο πίνακα $\boldsymbol{I}_n$.] Πολλαπλασιάζοντας την προηγούμενη σχέση από αριστερά με $\boldsymbol{\mu}'$ και από δεξιά με $\boldsymbol{Y}$, έχουμε

$$\begin{aligned}
\text{BLIE} &= \boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{Y} = \boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1} - \frac{1}{1+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{Y} \\
&= \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} - \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{1+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \\
&= \frac{1}{1+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \\
&= \frac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{1+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \cdot \text{BLUE},
\end{aligned}$$

λόγω του Θεωρήματος 1. (ο.ε.δ.)

Παρατηρώντας ότι $\theta > 0$ και $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} > 0$, έπεται από τις προηγούμενες εκφράσεις ότι BLUE ≥ 0 αν και μόνο αν BLIE ≥ 0 αν και μόνο αν $\boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{Z} \geq 0$ αν και μόνο αν $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} \geq 0$. Επομένως, για απλότητα, μπορούμε να χρησιμοποιούμε το τελευταίο κριτήριο, και έτσι, η μη-αρνητικότητα του BLUE ισοδυναμεί με το εξής:

Ερώτημα 1 (δικό μου – περί μη αρνητικότητας του BLUE του scale parameter στο γενικευμένο γραμμικό μοντέλο με μία άγνωστη παράμετρο, όταν τα δεδομένα προέχονται από μη-αρνητικό πληθυσμό). Αν $\boldsymbol{Z}' = (Z_1, Z_2, \ldots, Z_n)$ $(n \geq 2)$ είναι ένα τυχαίο διάνυσμα με $Z_i \geq 0$, $i = 1, 2, \ldots, n$, $\mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}$ και $\mathbb{D}[\boldsymbol{Z}] = \boldsymbol{\Sigma} > 0$ (positive definite), είναι αλήθεια ότι $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} \geq 0$;

Το Ερώτημα 1 γενικεύει το (μέχρι σήμερα ανοικτό) πρόβλημα που τέθηκε από τον Bala, και το οποίο ισοδυναμεί με το εξής:

Ερώτημα 2 – Εικασία του N. Balakrishnan (1992) περί μη-αρνητικότητας του BLUE του scale parameter στην γενική location-scale family. Έστω $\boldsymbol{Z}' = (Z_1, Z_2, \ldots, Z_n)$ $(n \geq 2)$ το τυχαίο διάνυσμα των spacings, δηλ. $Z_i = X_{(i+1)} - X_{(i)} \geq 0$, $i = 1, 2, \ldots, n$, όπου $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n+1)}$ είναι το διατεταγμένο δείγμα που αντιστοιχεί στο τυχαίο δείγμα $X_1, X_2, \ldots, X_{n+1}$, το οποίο προέρχεται από κατανομή με μη μηδενική, πεπερασμένη διασπορά. Αν θέσουμε $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}]$ και $\boldsymbol{\Sigma} = \mathbb{D}[\boldsymbol{Z}]$, είναι αλήθεια ότι $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} \geq 0$;

Αποδεικνύεται ότι το Ερώτημα 2 επιδέχεται θετική απάντηση αν και μόνον αν όλες οι συνιστώσες του n-διάστατου διανύσματος $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ είναι μη-αρνητικές. Επομένως, το Ερώτημα 2 διατυπώνεται ισοδύναμα ως εξής:

Ερώτημα 2′. Έστω $\boldsymbol{Z}' = (Z_1, Z_2, \ldots, Z_n)$ $(n \geq 2)$ το τυχαίο διάνυσμα των spacings από κατανομή με μη μηδενική, πεπερασμένη διασπορά. Αν θέσουμε $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}]$ και $\boldsymbol{\Sigma} = \mathbb{D}[\boldsymbol{Z}]$, είναι αλήθεια ότι το διάνυσμα $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ έχει όλες τις συνιστώσες του μη αρνητικές;

Είναι προφανές ότι τα spacings ικανοποιούν τις υποθέσεις του Ερωτήματος 1 χωρίς, φυσικά, να ισχύει το αντίστροφο. Για παράδειγμα, μπορεί να ισχύει ότι ο BLUE είναι μη-αρνητικός ακόμα και αν το διάνυσμα $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ έχει κάποια αρνητική συνιστώσα:

Παράδειγμα 0. Έστω $n = 2$, και I_1, I_2 δείκτριες Bernoulli με $I_1, I_2 \sim \mathrm{Be}(1/2)$ και $I_1 \cdot I_2 \sim \mathrm{Be}(2/5)$, δηλ. $\mathbb{P}[I_1 = I_2 = 1] = \mathbb{P}[I_1 = I_2 = 0] = 2/5$ και $\mathbb{P}[I_1 = 1, I_2 = 0] = \mathbb{P}[I_1 = 0, I_2 = 1] = 1/10$. Θέτουμε $Z_1 = I_1 \geq 0$, $Z_2 = I_2 + 1 \geq 0$, $\boldsymbol{Z}' = (Z_1, Z_2)$, οπότε τα $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}]$, $\boldsymbol{\Sigma} = \mathbb{D}[\boldsymbol{Z}]$ και $\boldsymbol{\Sigma}^{-1}$ δίδονται από τους τύπους

$$\boldsymbol{\mu} = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \frac{1}{20} \cdot \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}, \quad \boldsymbol{\Sigma}^{-1} = \frac{5}{4} \cdot \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}.$$

Μολονότι το διάνυσμα

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \frac{5}{2} \cdot \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

έχει μία αρνητική συνιστώσα, είναι προφανές ότι $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} = (5/2)(3Z_2 - Z_1) \geq 0$, αφού $3Z_2 = 3(I_2 + 1) \geq 3 > 1 \geq Z_1 = I_1$. Για την ιστορία, ο BLUE του μοντέλου αυτού είναι ο (βλ. Θεώρημα 1)

$$\mathrm{BLUE} = \frac{3Y_2 - Y_1}{4} \stackrel{\mathrm{d}}{=} \theta \cdot \frac{3Z_2 - Z_1}{4} \geq \frac{\theta}{2} > 0.$$

Σημειώνεται ότι **δεν υπάρχουν spacings με κατανομή όπως στο παράδειγμα αυτό**, διότι, όπως εύκολα διαπιστώνεται, αν τα Z_1 και Z_2 είναι spacings από οποιαδήποτε κατανομή και ισχύει ότι $\mathbb{P}[Z_1 = 0] > 0$, τότε θα πρέπει και $\mathbb{P}[Z_2 = 0] > 0$, πράγμα το οποίο αντιβαίνει στα δεδομένα του παραδείγματος. Φυσικά, δεν γνωρίζω αν υπάρχουν spacings με $\boldsymbol{\mu}$ και $\boldsymbol{\Sigma}$ όπως στο παράδειγμα (στοιχηματίζω ότι δεν υπάρχουν, αλλά αυτό δεν φαίνεται να είναι τόσο προφανές). Πάντως, αν υπήρχαν, τότε η απάντηση στην Εικασία του Bala θα ήταν αρνητική (βλ. Ερώτημα 2, σχόλιο που το ακολουθεί και Ερώτημα 2′).

Από την παραπάνω συζήτηση αρχίζει να διαφαίνεται κάποια συσχέτιση της Εικασίας του Bala με ένα γενικευμένο πρόβλημα ροπών (moment problem) των spacings – π.χ., μπορώ να διατυπώσω την εξής, όχι και τόσο αθώα, ερώτηση:

Για ποια ζεύγη $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ υπάρχουν spacings, $\boldsymbol{Z}$, με $\mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}$ και $\mathbb{D}[\boldsymbol{Z}] = \boldsymbol{\Sigma}$;

Χαρακτηρίζοντας αυτά τα ζεύγη ίσως γίνει δυνατό να αποδειχθεί η Εικασία του Bala, αφού θα μένει να δειχθεί ότι όλες οι συνιστώσες του διανύσματος $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ είναι μη-αρνητικές.

Όσο για το (δικό μου) γενικευμένο Ερώτημα 1 η απάντηση είναι, δυστυχώς, αρνητική. Και λέω «δυστυχώς» διότι ήλπιζα ότι το Ερώτημα 1 θα είχε απλούστερη απόδειξη από το Ερώτημα 2, αν φυσικά ίσχυε η θετική απάντηση. Για του λόγου το αληθές θα αναφέρω δύο αντιπαραδείγματα για το Ερώτημα 1 στην περίπτωση που $n = 2$ (το δεύτερο είναι του Μιχάλη Μπούτσικα – thanks Michael!).

Παράδειγμα 1. Έστω $n = 2$, και I_1, I_2 δείκτριες Bernoulli με $I_1, I_2 \sim \mathrm{Be}(1/4)$ και $I_1 \cdot I_2 \sim \mathrm{Be}(1/5)$, δηλ. $\mathbb{P}[I_1 = I_2 = 1] = 1/5$, $\mathbb{P}[I_1 = I_2 = 0] = 7/10$ και $\mathbb{P}[I_1 = 1, I_2 = 0] = \mathbb{P}[I_1 = 0, I_2 = 1] = 1/20$. Θέτουμε $Z_1 = I_1 \geq 0$, $Z_2 = 2I_2 + 1 \geq 0$, $\boldsymbol{Z}' = (Z_1, Z_2)$, οπότε τα $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}]$, $\boldsymbol{\Sigma} = \mathbb{D}[\boldsymbol{Z}]$ και $\boldsymbol{\Sigma}^{-1}$ δίδονται από τους τύπους

$$\boldsymbol{\mu} = \frac{1}{4} \cdot \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \frac{1}{80} \cdot \begin{bmatrix} 15 & 22 \\ 22 & 60 \end{bmatrix}, \quad \boldsymbol{\Sigma}^{-1} = \frac{5}{26} \cdot \begin{bmatrix} 60 & -22 \\ -22 & 15 \end{bmatrix}.$$

Εύκολα διαπιστώνουμε ότι

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \frac{5}{26} \cdot \begin{bmatrix} -18 \\ 17 \end{bmatrix}, \quad \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \frac{105}{26},$$

οπότε (βλ. Θεώρημα 1)

$$\mathrm{BLUE} = \frac{17Y_2 - 18Y_1}{21} \stackrel{\mathrm{d}}{=} \theta \cdot \frac{17Z_2 - 18Z_1}{21} = \frac{\theta}{21} \cdot \{17 + 34I_2 - 18I_1\}.$$

Είναι σαφές ότι

$$\mathbb{P}[\mathrm{BLUE} < 0] = \mathbb{P}[\mathrm{BLUE} = -\theta/21] = \mathbb{P}[I_1 = 1, I_2 = 0] = 1/20 > 0.$$

Φυσικά και ο BLIE παίρνει αρνητικές τιμές με θετική πιθανότητα στο παράδειγμα αυτό, αφού (βλ. Θεώρημα 2)

$$\text{BLIE} = \frac{105}{131} \cdot \text{BLUE}.$$

Σημειώνεται ότι **δεν υπάρχουν spacings με κατανομή όπως παραπάνω**, και ισχύουν τα ίδια σχόλια με αυτά του Παραδείγματος 0.

Παράδειγμα 2 (Μ. Μπούτσικας). Εδώ προκαθορίζουμε τα $\boldsymbol{\mu}$ και $\boldsymbol{\Sigma}$ ως εξής:

$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}, \quad \text{οπότε} \quad \boldsymbol{\Sigma}^{-1} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}.$$

Τα $\boldsymbol{\mu}$ και $\boldsymbol{\Sigma}$ επιλέχθηκαν έτσι ώστε οι πίνακες $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{\Sigma}^{1/2}$ και $\boldsymbol{\Sigma}^{-1/2}$ να είναι σχετικά απλοί και το διάνυσμα $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ να έχει μία αρνητική συνιστώσα. Πράγματι,

$$\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = 1, \quad \boldsymbol{\Sigma}^{1/2} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ και } \boldsymbol{\Sigma}^{-1/2} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix},$$

οπότε (βλ. Θεώρημα 1)

$$\text{BLUE} \stackrel{\text{d}}{=} \theta \cdot \{2Z_2 - Z_1\},$$

και μας ενδιαφέρει αν μπορεί να πάρει αρνητικές τιμές η τυχαία μεταβλητή $T = 2Z_2 - Z_1$, όταν $\boldsymbol{Z}' = (Z_1, Z_2)$ με $Z_1 \geq 0$, $Z_2 \geq 0$, $\mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}$ και $\mathbb{D}[\boldsymbol{Z}] = \boldsymbol{\Sigma}$, όπως παραπάνω.

Για να απλοποιηθούν οι λογαριασμοί θέτουμε $\boldsymbol{X}' = (X_1, X_2)$ με

$$\boldsymbol{X} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{Z} - \boldsymbol{\mu}) = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} Z_1 - 1 \\ Z_2 - 1 \end{bmatrix} = \begin{bmatrix} Z_1 - Z_2 \\ 2Z_2 - Z_1 - 1 \end{bmatrix},$$

και συνεπώς, $\mathbb{E}[\boldsymbol{X}] = \boldsymbol{0}$ και $\mathbb{D}[\boldsymbol{X}] = \boldsymbol{I}_2$, από κατασκευή. Επίσης,

$$\boldsymbol{Z} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{X} + \boldsymbol{\mu} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2X_1 + X_2 + 1 \\ X_1 + X_2 + 1 \end{bmatrix},$$

οπότε οι περιορισμοί μη-αρνητικότητας για τα Z_i αντιστοιχούν στις ανισότητες $2X_1 + X_2 \geq -1$ και $X_1 + X_2 \geq -1$. Τέλος, ο BLUE είναι θετικό πολλαπλάσιο της $T = 2Z_2 - Z_1 = 2(X_1 + X_2 + 1) - (2X_1 + X_2 + 1) = X_2 + 1$, επομένως θα μπορεί να πάρει αρνητικές τιμές αν και μόνο αν η X_2 μπορεί να πάρει τιμές μικρότερες του -1. Συνοψίζοντας, αναζητούμε $\boldsymbol{X}' = (X_1, X_2)$ τέτοιο ώστε να ισχύουν ταυτόχρονα οι εξής τρεις συνθήκες:

(1) $\mathbb{E}[\boldsymbol{X}] = \boldsymbol{0}$ και $\mathbb{D}[\boldsymbol{X}] = \boldsymbol{I}_2$,

(2) $2X_1 + X_2 \geq -1$ και $X_1 + X_2 \geq -1$ (με πιθ. 1), και

(3) $X_2 < -1$ με θετική πιθανότητα.

Από τις (2) και (3) προέκυψε το σχήμα που σας έδωσα. Προς στιγμήν πίστεψα ότι δεν είναι εφικτές οι παραπάνω συνθήκες, αλλά ο Μπούτσικας με διέψευσε με το εξής απλό παράδειγμα:

$$\mathbb{P}[X_1 = -1/2, X_2 = 0] = \frac{4}{5},\ \mathbb{P}[X_1 = 2, X_2 = -\sqrt{5}] = \mathbb{P}[X_1 = 2, X_2 = \sqrt{5}] = \frac{1}{10},$$

που ικανοποιεί τις (1)–(3) με $\mathbb{P}[X_2 < -1] = 1/10$. Αυτό το παράδειγμα αντιστοιχεί στις μη-αρνητικές Z_1, Z_2 με

$$\begin{aligned}
&\mathbb{P}[Z_1 = 5 - \sqrt{5}, Z_2 = 3 - \sqrt{5}] = \mathbb{P}[Z_1 = 5 + \sqrt{5}, Z_2 = 3 + \sqrt{5}] = \frac{1}{10},\\
&\mathbb{P}[Z_1 = 0, Z_2 = 1/2] = \frac{4}{5},
\end{aligned}$$

για τις οποίες ισχύει ότι $\mathbb{E}[\boldsymbol{Z}] = \boldsymbol{\mu}$, $\mathbb{D}[\boldsymbol{Z}] = \boldsymbol{\Sigma}$ και

$$\mathbb{P}[\text{BLUE} < 0] = \mathbb{P}[2Z_2 - Z_1 < 0] = \mathbb{P}[2Z_2 - Z_1 = -(\sqrt{5} - 1)] = \frac{1}{10}.$$

Αξίζει να σημειωθεί ότι η τυχαία μεταβλητή $T = \text{BLUE}/\theta = 2Z_2 - Z_1$ έχει συμμετρική κατανομή γύρω από την $\mathbb{E}[T] = 1$, αφού παίρνει τις τιμές $1 - \sqrt{5}$, 1 και $1 + \sqrt{5}$ με αντίστοιχες πιθανότητες $1/10$, $4/5$ και $1/10$. Πάντως, το συγκεκριμένο διάνυσμα $\boldsymbol{Z}$ δεν μπορεί να προέλθει από spacings, όπως και στα προηγούμενα παραδείγματα.

Σχετικές αναφορές

[1] Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. (1992). *A First Course in Order Statistics.* Wiley, N.Y. (see p. 174, Remark 4, where the question appeared for the first time).

[2] Bai, Z; Sarkar, S.K.; Wang, W. (1997). Positivity of the best linear unbiased L-estimator of the scale parameter with complete or selected order statistics from location–scale distribution. *Statist. Probab. Lett.* **32**, 181–188 (a partial answer).

[3] Balakrishnan, N.; Papadatos, N. (2002). The use of spacings in the estimation of a scale parameter. *Statist. Probab. Lett.* **57**, 193–204 (a partial answer).

[4] Burkschat, M. (2009). Multivariate dependence of spacings of generalized order statistics. *J. Multivariate Anal.* **100**, 1093–1106 (see Theorem 3.5 for a partial answer).

[5] Sarkadi, K. (1985). On an estimator of the scale parameter. *Statist. Decisions Suppl.* **2**, 231–236 (perhaps the first author who considers spacings; I do not have this paper).