

**האוניברסיטה העברית בירושלים
המחלקה לסטטיסטיקה ומדע נתונים**

מודלים סטטיסטיים למתקדמים א' (52801)
פרופ' דוד צוקר
תשפ"ד סמסטר א', מועד ב'

תאריך : כ"ב בניסן תשפ"ד, 30.4.24

משך הבחינה : שלוש שעות

חומר עזר : מחשב CIS ודף רשיומות
(4 דפים בגודל A4 שני צדדים או 8 דפים בגודל A4 צד אחד)

יש בבחינה ארבע שאלות, כל אחת שווה 25 נקודות. עליכם לענות על כל השאלות.

חל איסור מוחלט להעתיק. תלמיד שייתפס יורחק לשנה מלימודיו.

בהצלחה!!

שאלה 1 (25 נקודות)

יהיו $A(\theta), B(\theta)$ שתי מטריצות באשר A הינה בגודל $k \times k$ ו- B הינה בגודל $m \times m$.

a. הוכיחו כי

$$\frac{\partial}{\partial \theta_r} A(\theta) B(\theta) = \left(\frac{\partial}{\partial \theta_r} A(\theta) \right) B(\theta) + A(\theta) \left(\frac{\partial}{\partial \theta_r} B(\theta) \right)$$

b. הוכיחו כי

$$\frac{\partial}{\partial \theta_r} A(\theta)^{-1} = -A(\theta)^{-1} \left(\frac{\partial}{\partial \theta_r} A(\theta) \right) A(\theta)^{-1}$$

שאלה 2 (25 נקודות)

נתוייחס ל- random line model

$$Y_{ij} = \beta_{1i} + \beta_{2i}t_j + \epsilon_{ij}, \quad i = 1, \dots, n \text{ and } j = 1, \dots, J$$

באשר

$$\beta_i = \begin{bmatrix} \beta_{1i} \\ \beta_{2i} \end{bmatrix} \sim N(\beta, A)$$

the ϵ_{ij} 's are i.i.d. $N(0, \sigma^2)$

β_i and the ϵ_{ij} 's are independent

נניח שהשתמשנו במודל הנ"ל לנתח נתונים עם זמני מדידה 2, 4, 6, 8, 10, 12 וקיבלנו את האומדנים הבאים:

$$\hat{\beta}_1 = 0.8, \hat{\beta}_2 = 0.15, \hat{A}_{11} = 10, \hat{A}_{12} = 0.3, \hat{A}_{22} = 0.05, \hat{\sigma}_\epsilon^2 = 0.6$$

כעת יש לנו פרט חדש שהערך $Y^*(2)$ שלו בזמן 2 שווה ל- 1.3 ואנחנו רוצים לנבא את הערך $Y^*(6)$ שלו בזמן 6. חשבו אומדן ל- $E[Y^*(6)|Y^*(2)]$.

שאלה 3 (25 נקודות)

בנספח מופיע פلت מניתוח מרכיבים ראשיים (PCA). נניח כי החלטנו לשמר רק את שני המרכיבים הראשונים, דהיינו

$$\tilde{X}_r = [U]_{r1}Y_1 + [U]_{r2}Y_2$$

לפי הסימון שהשתמשנו בו בשוררים. חשבו את מקדם המתאם (קורלציה) $\text{Corr}(X_3, \tilde{X}_3)$.

שאלה 4 (25 נקודות)

נניח כי יש לנו מחקר עם 4 השערות אפס בלתי תלויות אשר ב- 2 מהן השערת האפס נכונה וסטטיטיסטי המבחן מפולג $N(0, 1)$ וב- 2 מהן השערת האפס לא נכונה וסטטיטיסטי המבחן מפולג $N(1.5, 1)$. נניח עוד כי הסטטיטיסטי קאי הפעיל להשתמש בשיטת בופרוני עם $\alpha = 0.1$ חד צדי. מהו ה FDR של הפרוצדורה? לוח של התפלגות $N(0, 1)$ מצורף בסוף המבחן.

נספח

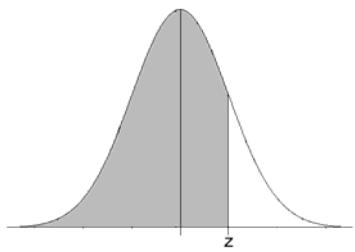
Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
Standard deviation	2.12132	2.0000000	1.8708287	1.7320508	1.4142136	1.00000000
Proportion of Variance	0.25000	0.2222222	0.1944444	0.1666667	0.1111111	0.05555556
Cumulative Proportion	0.25000	0.4722222	0.6666667	0.8333333	0.9444444	1.00000000

Loadings:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
[1,]	0.599	0.175	0.430	0.571	0.145	0.280
[2,]	0.128	-0.189	0.028	0.210	0.428	-0.848
[3,]	0.115	-0.604	0.593	-0.483	0.148	0.126
[4,]	0.012	0.550	0.006	-0.427	0.706	0.129
[5,]	-0.612	0.338	0.662	0.158	-0.131	-0.173
[6,]	-0.486	-0.391	-0.155	0.435	0.509	0.373

Standard Normal Cumulative Probability Table



Cumulative probabilities for POSITIVE z-values are shown in the following table: