

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 כיס ודפי רשימות
(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 $\mathbf{A}(\boldsymbol{\theta}), \mathbf{B}(\boldsymbol{\theta})$ שתי מטריצות באשר \mathbf{A} הינה בגודל $k \times k$ ו- $\mathbf{B}(\boldsymbol{\theta})$ הינה בגודל $k \times m$.

א. הוכיחו כי

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

ב. הוכיחו כי

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

שאלה 2 (25 נקודות)

נתייחס 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

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

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

$\boldsymbol{\beta}_i$ 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:

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

בעת יש לנו פרט חדש שהערך $Y^*(2)$ שלו בזמן $t = 2$ 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 פלט מניתוח מרכיבים ראשיים (PCA). נניח כי החלטנו לשמור רק את שני המרכיבים הראשונים, דהיינו

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

לפי הסימון שהשתמשנו בו בשעורים. 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 וסטטיסטי המבחן מפולג $N(0, 1)$ וב- 2 מהן השערת האפס לא נכונה וסטטיסטי המבחן מפולג $N(1.5, 1)$. נניח עוד כי הסטטיסטיקאי החליט להשתמש בשיטת בופרוני עם $\alpha = 0.1$ חד צדדי. 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

A normal distribution curve is shown with a vertical line at a point labeled z on the horizontal axis. The area under the curve to the left of z is shaded gray, representing the cumulative probability $P(Z \leq z)$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	