רגרסיה ומודלים לינאריים 52320 תשע"ז 2016-17
בוחן 1 26.04.2017

בבוחן שאלות אמריקאיות ושאלות פתוחות.

משקל כל שאלה הוא 25 נקודות כך שמספר הנקודות הכולל הוא 125. בכל מקרה, **ציון הבוחן הוא 100 לכל היותר.**

שימו לב שהשאלות הן בדרגת קושי שונה כך שמומלץ לא להתעכב יתר על המידה על שאלה מסויימת.

אנא הקפידו על ההנחיות הבאות:

- כתבו את ת.ז. (**לא את השם!**) בראש כל עמוד של טופס הבחינה.
- אין לצרף לטופס דפים נוספים.
- אין לתלוש דפים מטופס הבחינה.

לתשומת לבכם לגבי השאלות הפתוחות:

- **תשובה סופית ללא דרך לא תזכה בניקוד כלשהו (ציון 0).**
- בשאלות הפתוחות יש לכתוב את הפתרון רק במקום המוקצה לכך, מעל לכל קו כתבו שורה אחת בלבד בכתב יד קריא. (השאלות נכתבו כך שניתן לכתוב פתרון תמציתי לכל סעיף).
- מגבלת המקום תאכף באופן קפדני. פתרונות אשר יחרגו מהמקום המותר, יהיו בכתב קטן מכדי שיהיה קריא, ו/או יכללו יותר משורת כתב אחת לכל קו לא ייבדקו.
- מומלץ מאוד לפתור תחילה את השאלה במחברת הטיוטה ולהעתיק את עיקר הפתרון אל הטופס רק לאחר בדיקה.

חומר עזר מותר: מחשבון, דף נוסחאות דו צדדי

משך הבוחן: שעה

בהצלחה!

סימונים: נכתוב משתנים בכתיב וקטורי, כאשר $x, y, ..$ הם וקטורי עמודה. x_i מסמן את האיבר ה־i של וקטור x ו־$\bar{x}$ מסמן את הממוצע של וקטור x. עבור שני וקטורים x, y באורך n המכפלה הסקלרית שלהם היא $x^T y = \sum_{i=1}^n x_i y_i$.

תזכורת: נגדיר מודל לרגרסיה פשוטה עם חותך: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ עבור נתונים $(x_1, y_1), ..., (x_n, y_n)$. עבור מודל זה שגיאת הריבועים הפחותים $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2$ ניתנת לכתיבה בצורות הבאות:

$SSE = \left(y - \hat{\beta}_0 - \hat{\beta}_1 x\right)^T \left(y - \hat{\beta}_0 - \hat{\beta}_1 x\right) = (y - \bar{y})^T (y - \bar{y}) - \hat{\beta}_1^2 (x - \bar{x})^T (x - \bar{x})$ כאשר האומדים למקדמים הם: $\hat{\beta}_1 = \frac{(x-\bar{x})^T (y-\bar{y})}{(x-\bar{x})^T (x-\bar{x})}, \ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

1. עבור מודל הרגרסיה הפשוטה עם חותך לעיל, יהיה $\hat{y}$ וקטור התחזיות ויהיה $e = y - \hat{y}$ וקטור השאריות. אילו מהגדלים הבאים תמיד שווים לאפס? $\sum_{i=1}^n e_i, \ \sum_{i=1}^n x_i e_i, \ \sum_{i=1}^n y_i e_i$

(i) $\sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = n(\bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x}) = n \cdot (\bar{y} - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 \bar{x}) = 0.$

(ii) $\sum_{i=1}^n x_i e_i = \sum_{i=1}^n x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = n(\overline{xy} - \hat{\beta}_0 \bar{x} - \hat{\beta}_1 \overline{x^2})$
$= n(\overline{xy} - (\bar{y}\bar{x} - \hat{\beta}_1 \bar{x}^2) - \hat{\beta}_1 \overline{x^2}) = S_{xy} - \frac{1}{n} S_x S_y - \hat{\beta}_1 (S_{xx} - \frac{1}{n} S_x^2)$
$= S_{xy} - \frac{1}{n} S_x S_y - \frac{S_{xy} - \frac{1}{n} S_x S_y}{S_{xx} - \frac{1}{n} S_x^2} (S_{xx} - \frac{1}{n} S_x^2) = 0.$

(iii) $n = 3$, $D = \{(0,0), (1,1), (-1,0)\}$
$\hat{\beta}_1 = \frac{1-0}{2-0} = \frac{1}{2}$
$\hat{\beta}_0 = \frac{1}{3} - 0 = \frac{1}{3}$
$\hat{y} = (\frac{1}{3}, \frac{5}{6}, -\frac{1}{6})$
$e = (-\frac{1}{3}, \frac{1}{6}, \frac{1}{6})$
$\sum_i y_i e_i = 0 + \frac{1}{6} + 0 = \frac{1}{6} \neq 0.$

(א) כולם

(ב) אף אחד

(ג) רק $\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i$

(ד) רק $\sum_{i=1}^n e_i$

(ה) רק $\sum_{i=1}^n x_i e_i, \sum_{i=1}^n e_i$

(ו) אף אחת מהתשובות לעיל אינה נכונה

2. נניח כעת כי נתונים כל הסכומים הבאים עבור נתונים $(x_1, y_1), ..., (x_n, y_n)$:

$$S_x = \sum_{i=1}^n x_i = -3.8, S_y = \sum_{i=1}^n y_i = 32.6, S_{xx} = x^T x = 14.45, \ \ S_{yy} = y^T y = 69.64, S_{xy} = x^T y = -15.45, n = 20$$

חשבו את $\hat{\beta}_0, \hat{\beta}_1$ ואת האומד לשונות s^2 עבור מודל הרגרסיה עם חותך והנתונים שלעיל.

$\boxed{\hat{\beta}_1} = \frac{S_{xy} - \frac{1}{n} S_x S_y}{S_{xx} - \frac{1}{n} S_x^2} = \frac{-15.45 - \frac{1}{20}(-3.8)32.6}{14.45 - \frac{1}{20}(-3.8)^2} = \frac{-9.256}{13.728} = \boxed{-0.674}$

$\boxed{\hat{\beta}_0} = \bar{y} - \bar{x}\hat{\beta}_1 = \frac{1}{20} 32.6 - \frac{1}{20}(-3.8)(-0.674) = \boxed{1.502}$

$SSE = SST - SSR = \sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 x_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{x})^2 = S_{yy} - \frac{1}{n} S_y^2 - \hat{\beta}_1^2 (S_{xx} - \frac{1}{n} S_x^2)$
$= 69.64 - \frac{1}{20}(32.6)^2 - (-0.674)^2 (14.45 - \frac{1}{20}(-3.8)^2) = 10.266$

$\boxed{s^2} = \frac{SSE}{n-2} = \frac{10.266}{18} = \boxed{0.57}$

3. נגדיר כעת מודל רגרסיה **ללא** חותך: $y_i = \gamma_1 x_i + \epsilon_i$ עבור נתונים $(x_1,y_1),...,(x_n,y_n)$. הוכיחו שבמודל זה אומד הרבועים הפחותים $\hat{\gamma}_1$ עבור γ_1 שווה ל־$\frac{x^T y}{x^T x}$.

נגדיר: $F(\gamma_1) \equiv \sum_{i=1}^{n} (y_i - \gamma_1 x_i)^2$ $\quad \hat{\gamma}_1 = \arg\min_{\gamma_1} \sum_{i=1}^{n} (y_i - \gamma_1 x_i)^2$

$$\frac{dF}{d\gamma_1} = \sum_{i=1}^{n} 2(y_i - \gamma_1 x_i)(-x_i) = 0 \Rightarrow \sum_{i=1}^{n} x_i y_i = \sum_i x_i^2 \gamma_1$$

$$\Rightarrow \hat{\gamma}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{x^T y}{x^T x}$$

$$\frac{d^2F}{d\gamma_1^2} = \sum_{i=1}^{n} 2x_i^2 > 0 \Rightarrow \text{מינימום}$$

4. עבור רגרסיה לינארית פשוטה, עם שגיאות נורמליות $\epsilon \sim N(0, \sigma^2 I_n)$ עבור $\sigma^2 > 0$ יהיו $\hat{\beta}_0, \hat{\beta}_1, \bar{y}$ אומדי הרבועים הפחותים הרגילים וממוצע ה־y_i בהתאמה עבור מדגם $D = (x_1,y_1),...,(x_n,y_n)$ עם $n > 1$. שלושת המשתנים המקריים $\hat{\beta}_0, \hat{\beta}_1, \bar{y}$ הם:

$(\hat{\beta}_0, \hat{\beta}_1, \bar{y})$ תלויים $\Leftarrow$ ערכי $\bar{y}, \hat{\beta}_1$ קובעים את $\hat{\beta}_0$ $\Leftarrow$ $\hat{\beta}_0 = \bar{y} - \bar{x}\hat{\beta}_1$

$\bar{y}, \hat{\beta}_1$ בת"ל (ראינו בכיתה) $\Leftarrow$ $cov(\bar{y}, \hat{\beta}_1) = 0$

$cov(\hat{\beta}_0, \bar{y}) = cov(\bar{y} - \bar{x}\hat{\beta}_1, \bar{y}) = Var(\bar{y}) - \bar{x}\, cov(\hat{\beta}_1, \bar{y}) = Var(\bar{y}) > 0$

כלומר $\hat{\beta}_0, \bar{y}$ מתואמים ולכן תלויים.

(א) תלויים אך בלתי תלויים בזוגות
(ב) בלתי תלויים
(ג) תלויים וכל הזוגות תלויים
(ד) תלויים, חלק מהזוגות תלויים וחלק מהזוגות בלתי תלויים

5. נתונה המטריצה $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix}$ והוקטור $v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. תהיה $P_A = A[A^T A]^{-1} A^T$ מטריצת הטלה על תת המרחב הנפרש על ידי עמודות A ותהי $P_{A^\perp}$ מטריצת הטלה על תת המרחב הניצב לתת המרחב הנפרש על ידי עמודות A. חשבו את הגדלים הבאים: (כתבו את התוצאות כוקטורי שורה. כתבו רק תוצאה סופית ־ אין צורך להראות דרך)

(א) $P_A v =$ $(\frac{1}{2}, \frac{1}{2}, 0)$

(ב) $P_{A^\perp} v =$ $(\frac{1}{2}, -\frac{1}{2}, 0)$

(ג) $P_{A^\perp} P_A v =$ $(0, 0, 0)$

$$A^T A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \Rightarrow [A^T A]^{-1} = \begin{bmatrix} 1/6 & 0 \\ 0 & 1/3 \end{bmatrix}$$

$$P_A v = A[A^T A]^{-1} A^T v = A[A^T A]^{-1} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = A[A^T A]^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= A \begin{bmatrix} 1/6 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = A \begin{bmatrix} 1/6 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1/6 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$P_{A^\perp} v = v - P_A v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}$$

$$P_A v \in Span(A) \Rightarrow P_A v \perp Span(A)^\perp$$

$$\Rightarrow P_{A^\perp} P_A v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

סימונים: נכתוב משתנים בכתיב וקטורי, כאשר $x, y, ..$ הם וקטורי עמודה. x_i מסמן את האיבר ה־i של וקטור x ו־$\bar{x}$ מסמן את הממוצע של וקטור x. עבור שני וקטורים x, y באורך n המכפלה הסקלרית שלהם היא $x^T y = \sum_{i=1}^n x_i y_i$.

תזכורת: נגדיר מודל לרגרסיה פשוטה עם חותך: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ עבור נתונים $(x_1,y_1), ..., (x_n, y_n)$. **עבור מודל זה** שגיאת הרבועים הפחותים $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2$ ניתנת לכתיבה בצורות הבאות:

$SSE = \left(y - \hat{\beta}_0 - \hat{\beta}_1 x\right)^T \left(y - \hat{\beta}_0 - \hat{\beta}_1 x\right) = (y - \bar{y})^T (y - \bar{y}) - \hat{\beta}_1^2 (x - \bar{x})^T (x - \bar{x})$ כאשר האומדים למקדמים הם: $\hat{\beta}_1 = \frac{(x-\bar{x})^T (y-\bar{y})}{(x-\bar{x})^T (x-\bar{x})},\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

1. עבור מודל הרגרסיה הפשוטה עם חותך לעיל, יהיה $\hat{y}$ וקטור התחזיות ויהיה $e = y - \hat{y}$ וקטור השאריות. אילו מהגדלים הבאים **תמיד** שווים לאפס? $\sum_{i=1}^n e_i,\ \sum_{i=1}^n x_i e_i,\ \sum_{i=1}^n y_i e_i$

   (א) כולם
   (ב) אף אחד
   (ג) רק $\sum_{i=1}^n e_i, \sum_{i=1}^n x_i e_i$
   (ד) רק $\sum_{i=1}^n e_i$
   (ה) רק $\sum_{i=1}^n y_i e_i, \sum_{i=1}^n x_i e_i$
   (ו) אף אחת מהתשובות לעיל אינה נכונה

2. נניח כעת כי נתונים כל הסכומים הבאים עבור נתונים $(x_1,y_1), ..., (x_n, y_n)$:

$$S_x = \sum_{i=1}^n x_i = -3.8, S_y = \sum_{i=1}^n y_i = 23.6, S_{xx} = x^T x = 41.45,\ \ S_{yy} = y^T y = 69.64, S_{xy} = x^T y = -15.45, n = 20$$

חשבו את $\hat{\beta}_0, \hat{\beta}_1$ ואת האומד לשונות s^2 עבור מודל הרגרסיה עם חותך והנתונים שלעיל.

$\boxed{\hat{\beta}_1} = \frac{-15.45 - \frac{1}{20}(-3.8)23.6}{41.45 - \frac{1}{20}(-3.8)^2} = \boxed{-0.269}$

$\boxed{\hat{\beta}_0} = \frac{1}{20}23.6 - \frac{1}{20}(-3.8)(-0.269) = \boxed{1.125}$

$SSE = 69.64 - \frac{1}{20}(23.6)^2 - (-0.269)^2\left(41.45 - \frac{1}{20}(-3.8)^2\right) = 38.845$

$\boxed{s^2} = \frac{SSE}{18} = \boxed{2.158}$

3. נגדיר כעת מודל רגרסיה **ללא** חותך: $y_i = \gamma_1 x_i + \epsilon_i$ עבור נתונים $(x_1,y_1),...,(x_n,y_n)$. הוכיחו שבמודל זה אומד הרבועים הפחותים $\hat{\gamma}_1$ עבור γ_1 שווה ל-$\frac{x^T y}{x^T x}$.

4. עבור רגרסיה לינארית פשוטה, עם שגיאות נורמליות $\epsilon \sim N(0, \sigma^2 I_n)$ עבור $\sigma^2 > 0$ יהיו $\hat{\beta}_0, \hat{\beta}_1, \bar{y}$ אומדי הרבועים הפחותים הרגילים וממוצע ה־y_i בהתאמה עבור מדגם $D = (x_1,y_1),...,(x_n,y_n)$ עם $n > 1$. שלושת המשתנים המקריים $\hat{\beta}_0, \hat{\beta}_1, \bar{y}$ הם:

(א) תלויים וכל הזוגות תלויים

(ב) בלתי תלויים

(ג) תלויים, חלק מהזוגות תלויים וחלק מהזוגות בלתי תלויים

(ד) תלויים, אך בלתי תלויים בזוגות

5. נתונה המטריצה $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 3 & -1 \end{bmatrix}$ והוקטור $v = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. תהיה $P_A = A[A^T A]^{-1} A^T$ מטריצת הטלה על תת המרחב הנפרש על ידי עמודות A ותהי $P_{A^\perp}$ מטריצת הטלה על תת המרחב הניצב לתת המרחב הנפרש על ידי עמודות A. חשבו את הגדלים הבאים: (כתבו את התוצאות כוקטורי שורה. כתבו רק תוצאה סופית ־ אין צורך להראות דרך)

$A^T A = \begin{bmatrix} 11 & 0 \\ 0 & 6 \end{bmatrix}$ $\quad A^T v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$P_A v = A[A^T A]^{-1} A^T v = A \begin{bmatrix} 1/11 \\ 1/3 \end{bmatrix} = \frac{1}{33}\begin{bmatrix} 14 \\ 25 \\ -2 \end{bmatrix} = \begin{bmatrix} 0.424 \\ 0.757 \\ -0.060 \end{bmatrix}$

$P_{A^\perp} v = v - P_A v$

(א) $P_A v = \frac{1}{33}(14, 25, -2)$

(ב) $P_{A^\perp} v = \frac{1}{33}(-14, 8, 2)$

(ג) $P_{A^\perp} P_A v = (0, 0, 0)$