

תרגיל בית 5

הנחיות להגשת התרגיל

יש להגיש שני קבצים נפרדים, אחד עבור הקוד ואחד לפתרון החלק היבש. בתוך קובץ פתרון החלק היבש הוסיפו גם את הפלט של הקוד.

שאלה 1

יהיו $X_1, \dots, X_n \sim F$. ותהי $\hat{F}_n(x)$ פונקציית ההתפלגות האמפירית. הוכיחו כי עבור x נתון מתקיים:

- א. $\mathbb{E}[\hat{F}_n(x)] = F(x)$
- ב. $\text{Var}(\hat{F}_n(x)) = \frac{F(x)(1-F(x))}{n}$
- ג. $\hat{F}_n(x)$ הוא אומד עקיב ל- $F(x)$

הדרכה: שימו לב שעבור x נתון, המשתנה $I(X \leq x)$ הוא משתנה ברנולי.

שאלה 2

יהיו $X_1, \dots, X_n \sim F$. ותהי $\hat{F}_n(x)$ פונקציית ההתפלגות האמפירית.

- א. עבור x נתון השתמשו במשפט הגבול המרכזי כדי למצא את ההתפלגות הגבולית של $\hat{F}_n(x)$.
- ב. יהיו $a, b \in \mathbb{R}$ כך ש- $a < b$ ונגדיר $\theta = T(F) = P(a < x \leq b)$. הראו כי זהו פונקציונל לינארי, הציעו אומד plug-in ל- θ ומצאו רווח סמך ברמת סמך מקורבת $1 - \alpha$ עבור θ .

שאלה 3

יהיו $X_1, \dots, X_n \sim F$ ותהי $\hat{F}_n(x)$ פונקציית ההתפלגות האמפירית. עבור x, y נתונים ושונים מצאו את $\text{Cov}(\hat{F}_n(x), \hat{F}_n(y))$

שאלה 4

נגדיר את השברון ה- α של פונקציית התפלגות אמפירית באופן הבא:

$$\hat{F}_n^{-1}(\alpha) = \inf\{x : \hat{F}_n(x) \geq \alpha\}$$

יהי $\hat{\theta}_\alpha^*$ השברון ה- α של ערכי הבוטסטראפ שהתקבלו. הראו שרווח הסמך המבוסס על פיווט ניתן לכתיבה כ-

$$\left[2\hat{\theta}_n - \hat{\theta}_{1-\frac{\alpha}{2}}^*, 2\hat{\theta}_n - \hat{\theta}_{\frac{\alpha}{2}}^* \right]$$

שאלה 5

חזרו על שקף 12 במצגת 3.4 (מבחן וולד). עבור הנתונים המתוארים בשקף זה, כתבו פסאודו-קוד לחישוב ה-standard error של הפרש החציונים באמצעות בוטסטראפ.

ללא חובת הגשה: הצדיקו את התהליך שתיארתם בסעיף א' באמצעות שלבי הבוטסטראפ שלמדנו.

שאלה 6

עבור 15 סטודנטים בפקולטה כלשהי למשפטים נאספו נתונים על ממוצע הציונים (GPA) וציוני מבחן הכניסה לפקולטה (LSAT). הנתונים מובאים בקובץ הנתונים המצורף לתרגיל. העזרו בפייתון וענו על הסעיפים הבאים.

א. חשבו את אומדן ה-plug-in למתאם (קורלציה) בין ציוני מבחן הכניסה (LSAT) לבין ממוצע הציונים (GPA).

ב. אמדו את ה-standard error של המתאם באמצעות $B = 1000$ דגימות בוטסטראפ

ג. מצאו רווחי סמך ברמת סמך 95% למתאם באמצעות שלוש השיטות שלמדנו: רווח סמך מבוסס קירוב נורמלי, רווח סמך פיווטלי ורווח סמך מבוסס שברונים.