

[ראשי](#) / [הקורסים שלי](#) / [57322 מבוא לאקונומטריקה](#) / [יחידות הוראה](#) / [מבחן 2021](#) / [מבחן 2021](#)

התחיל ב: 13:59, 15/02/2021

מצב הסתיים

הושלם ב- 12:56, 19/02/2021

הזמן שלקח 3 ימים 22 שעות

ציון 76.00 מתוך הציון המירבי 90.00 (84%)

מבחן בית - פברואר 2021

מועד העלאת הבחינה למודל: יום ב' - 15.02.2021 (בשעה 13:00)
מועד אחרון להגשת הבחינה למודל: יום ו' - 19.02.2021 (בשעה 13:00)

חומר עזר: מותר להיעזר בכל חומרי השיעור והתרגולים. אסור לפנות לאף אדם אחר לצורך עזרה, סיעור מוחות או התייעצות על פתרון השאלות, גם לא אם מדובר בהתייעצות כללית. אם יש לכם שאלות, אתם מוזמנים להפנות אותן למתרגלים.

זו עבודה אישית. אנחנו נותנים בכס אמון ומבקשים מאוד שתכבדו זאת, בהתאם להצהרה עליה חתמתם כעת. אסור לתקשר בשום דרך עם אנשים אחרים בנושאי הבחינה. אין לבקש מאנשים אחרים עזרה עם R או STATA.

שאלות והבהרות במהלך הבחינה: שאלות שמתעוררות בעת הבחינה ניתן לשלוח לצוות הקורס דרך פורום השאלות המופיע בעמוד המודל הראשי, מתחת למבחן. הסגל יראה את כל ההודעות שלכם אבל לא יענה על כל שאלה בנפרד באופן מרוכז בקובץ נפרד. תיקונים והערות יפורסמו בקובץ נפרד שיתעדכן בכל כמה שעות (במידת הצורך) ועל עדכונים נשלח מייל בתפוצת כל הקורס. קובץ התיקונים יהיה זמין בכל עת בדף המודל. אם שאלתכם לא נענתה בקובץ העדכונים האחרון זה אומר שראינו אותה אבל לא נדרשים תיקונים או הערות עבודה.

קראו בעיון את ההוראות:

- רמת המובהקות בכל המבחנים הסטטיסטיים הינה 5%.
- בחישובים ובהקלדת ערכים מתוך פלט הרגרסיה, יש להקליד לפחות **5 ספרות מימין לנקודה** על מנת להגיע לדיוק הרצוי; אם לאחר הנקודה מופיעים 2 אפסים או יותר, יש לקחת מספר ספרות כך שתהייה 3 ספרות שאינן אפס.
- במקרים בהם התוכנה מדווחת פחות מ-5 ספרות, אין צורך בדיוק של 5 ספרות, אך יש להקליד את המספר המדווח בפלט כפי שהוא, בשלמותו. כלומר אם התוכנה מדווחת ברמת דיוק של שני מקומות עשרוניים בלבד, זה מה שעליכם לדווח.
- אפשר לדייק יותר מההנחה.
- דווחו אחוזים כשבר עשרוני (למשל 0.5). בכל מקרה אין להוסיף סימן %.
- עבור מספרים גדולים, כתבו ללא פסיק מפריד. לדוגמה כתבו 1000 ולא 1,000.
- כאשר אתם מבצעים מבחן סטטיסטי (למשל מבחן t) הניקוד המלא יתקבל גם עבור הערך הנכון ללא תלות בסימן (למשל, 1 או -1).
- בראש כל שאלה פתוחה מופיע מספר המילים המקסימלי המותר לתשובה. על אף שהדבר אפשרי מבחינה טכנית - אין לחרוג ממגבלה זו. תשובות ארוכות יותר לא ייבדקו.
- בשאלות הפתוחות: אם אינכם בטוחים כיצד לפרש את השאלה, אפשר להסביר את התשובה לשתי הפרשנויות הסבירות בעינכם. עם זאת, גם במקרה זה אין לחרוג ממגבלת המילים.
- אלא אם הוגדר אחרת, אין לתקן עבור הטרוסקדסטיות.
- אין צורך להשתמש ב-cluster.
- על שאלות פתוחות ניתן לענות באנגלית.
- אם אתם נדרשים לנתח משהו בעזרת התוכנה כדי לענות על שאלה פתוחה יש לכלול את הניתוח בקובץ הלוג ולדון בתוצאות בתשובה הפתוחה שלכם ב moodle.
- בשאלה האחרונה במבחן אתם נדרשים להגיש קובץ לוג. וודאו שהקובץ מכיל את כל העבודה עליה מסתמכות התשובות שלכם, שהוא ניתן לקריאה ואינו מכיל "רעש" מיותר כגון ניסיונות קודמים או תשובות לא רלוונטיות. התשובות שהוקלדו חייבות להתאים לתוצאות בקובץ הלוג המצורף על מנת לקבל את הנקודות בשאלה.
- הנקודות במבחן מסתכמות ל 90 וזהו הציון המקסימלי האפשרי.

בהצלחה!

חלק א' (60 נקודות)

הורידו את קובץ הנתונים הבא:



קובץ הנתונים לשאלה זו מגיע מסקר שנערך בבריטניה ונקרא "Understanding Society". קובץ הנתונים (ukdata.dta) כולל נתוני פאנל של פרטים שונים בין השנים 2009-2020. לכל משתנה בו חסר מידע למשיב בסקר, המשתנה מקודד בתור הערך 999- לכל משתנה בו אתם משתמשים בנייתו שלכם, הקפידו לדאוג לערכים שיש להפוך לערכים חסרים. אם תתבקשו לבדוק השערה, עליכם לענות תחת ההנחה כי כל ההנחות הקלאסיות (MLR1-MLR6) מתקיימות במודל המדובר.

המשתנים בקובץ הנתונים מוגדרים כדלקמן:

- $year_i$ = שנת הסקר עבור פרט i
- $wage_{it}$ = השכר של פרט i בשנה t
- age_{it} = הגיל של פרט i בשנה t
- $female_i$ = משתנה דמי השווה ל-1 אם פרט i הינה אישה; ושווה 0 אחרת
- $married_{it}$ = משתנה דמי השווה ל-1 אם פרט i נשוי/אה בשנה t ; ושווה 0 אחרת
- $bomuk_{it}$ = משתנה דמי השווה ל-1 אם פרט i נולדה בבריטניה; ושווה 0 אחרת
- $educ_{it}$ = משתנה המייצג אחת משלוש קטגוריות השכלה של פרט i בשנה t
- $firmsize_{it}$ = משתנה המייצג אחת משלוש קטגוריות גודל החברה בה פרט i עובד/ת בשנה t
- id_i = מספר תעודת זהות פיקטיבי לפרט i

השתמשו בפקודות המתאימות כדי לראות כיצד כל משתנה מוגדר ומהן התוויות.

בסטטה: בפקודה `tab` עם ובלי האופציה `no label`,R: בפקודה `unique`

שאלה 1

תקין

2.00 נקודות מתוך 2.00

(2 נק')

עבור הנשאלים בסקר בגיל 40 בדיוק, לאיזה אחוז מהנשים יש לפחות תואר ראשון?

תשובה: ✓

התשובה הנכונה היא: 0.31894

שאלה 2

תקין

2.00 נקודות מתוך 2.00

(2 נק')

עבור הנשאלים בסקר בגיל 40 בדיוק, לאיזה אחוז מהגברים יש לפחות תואר ראשון?



0.258785

תשובה:

התשובה הנכונה היא: 0.32213439

שאלה 3

תקין

4.00 נקודות מתוך 4.00

(4 נק')

מהו ה-t-סטטיסטי להשערה ששני המספרים בשאלות הקודמות שווים?



0.80267

תשובה:

התשובה הנכונה היא: 0.10985548

מנקודה זו ועד סוף חלק א', יש להגביל את המדגם לנשים בגילאי 20 עד 65 (כולל שתי הקצוות).

הריצו את המודלים הבאים והציגו את הפלט מכל רגרסיה בקובץ log.

בסעיפים הבאים ענו לפי התוצאות שקיבלתם.

$$\text{model (1): } lwage = \beta_0 + \beta_1 age + \beta_2 educ + u_1$$

$$\text{model (2): } lwage = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 educ + u_2$$

$$\text{model (3): } lwage = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 educ + \beta_4 hsg + u_3$$

(2 נק')

מלאו את האומדים וטעויות התקן עבור המשתנים age ו-educ מכל אחד מהמודלים.

Model 1

age: Coef.

✓ SE

✓

educ: Coef.

✓ SE

✓

Model 2

age: Coef.

✓ SE

✓

educ: Coef.

✓ SE

✓

Model 3

age: Coef.

✓ SE

0.0031403



educ: Coef.

0.3089981

✓ SE

0.0054571



שימו לב - אל תלחצו על כפתור Submit\הגשת תשובה - המערכת שומרת את התשובות אוטומטית

שאלה 5

הושלם

0.00 נקודות מתוך 4.00

(4 נק') (אורך מקסימלי: 75 מילים)

בהנחה שכל ההנחות MLR1-MLR6 מתקיימות במודל 2- האם התוצאות ממודלים 1-2 מצביעות על כך שההשפעה של השכלה על לוג השכר תלויה בגיל הפרט? הסבירו.

ניתן לראות שיש הבדל במקדם (בטא) של educ ממודל 1 למודל 2. לפיכך, אנו מסיקים שאכן השכלה מתואמת עם גיל בריבוע. גיל בריבוע מאפשר לנו לתפוס יחס לא ליניארי במודל, ומכיוון שההכנסה של גיל בריבוע לרגרסיה שינתה את הערכים, אנו יכולים להסיק שההשפעה של השכלה על לוג השכר תלויה בגיל והיא לא קבועה.

הערה:

שאלה 6

תקין

4.00 נקודות מתוך 4.00

(4 נק')

מהתוצאות של מודל (3) וביחס למי שנשרה מהתיכון, מה ההשפעה של סיום תיכון על לוג השכר?

תשובה:

התשובה הנכונה היא: 0.0968958

שאלה 7

תקין

4.00 נקודות מתוך 4.00

(4 נק')

מהתוצאות של מודל (3) וביחס למי שנשרה מהתיכון, מה ההשפעה של תואר ראשון על לוג השכר?

תשובה:

התשובה הנכונה היא: 0.61799611

שאלה 8

תקין

4.00 נקודות מתוך 4.00

(4 נק')

מהתוצאות של מודל (3), מה ערך הסטטיסטי לבדיקת ההשערה (H_0) שהשכלה לא משפיעה על לוג השכר?

תשובה:

התשובה הנכונה היא: 1764.64

שאלה 9

תקין

2.00 נקודות מתוך 2.00

(2 נק')

הוסיפו למודל (3) את *married* כמשתנה מסביר. מודל זה ייקרא מודל (4).

מה ההשפעה הנאמדת של נישואין על לוג השכר?



-0.0110127

תשובה:

התשובה הנכונה היא: -0.0110127

שאלה 10

הושלם

4.00 נקודות מתוך 4.00

(4 נק') (אורך מקסימלי: 75 מילים)

הוסיפו למודל (4) את *id* כמשתנה מסביר.

לפי התוצאות, איך הוספת המשתנה הזה משפיעה על האומד של *married*?

הסבירו עם סיבות ספציפיות

ניתן לראות שהאומד השתנה מ -0.0110127 ל -0.0110105, כלומר השינוי הוא **מזערי**.

זה נובע מכך שלמספר של תעודת הזהות אין השפעה על הנישואים עצמם. כלומר, העבודה שאדם נולד עם מספר ת.ז גבוה או נמוך לא משפיעה על הסיכוי שלו להתחתן.

כמובן שיייתכן שבמדגם מקרי יהיה מצב שבו עבור מי שמספר ת.ז שלו הוא גבוה- שיעור הנישואים יהיה גבוה יותר מאשר מי שמספר ת.ז שלו הוא נמוך. לכן בכל זאת השינוי באומד הוא לא **בדיוק** 0.

הערה:

(4 נק') (אורך מקסימלי: 100 מילים)

הוסיפו למודל (4) משתני דמי לכל ערך של id כמשתנים מסבירים. מודל זה ייקרא מודל (5).

לפי התוצאות, איך הוספת המשתנים הללו משפיעה על האומד של $married$?

הסבירו עם סיבות ספציפיות.

אנו רואים כי המקדם של $married$ ירד. כאשר אנו מסתכלים על נתוני פאנל אנו מנסים לנטרל את ההשפעה של כל הגורמים שב- u עם נישואים ובכך לקבל אומד מדויק יותר עבור נישואים עצמם. במקרה שלנו אנו רואים שההטיה הייתה כלפי מעלה. דוגמא אפשרית להטיה יכולה להיות שלנשואים יש IQ גבוה יותר מלרווקים כלומר $cov(married, IQ) > 0$. כאשר אנו מסתכלים על אותו אדם לאורך זמן ($fixed$ -effect), שאר הפרמטרים נשארים קבועים ואנו רואים את ההשפעה נטו של הנישואים. מודל OLS רגיל משווה בין כל האנשים אפילו שיש התאמה בין u למשתנים המסבירים. הפתרון של משתני דמי עבור כל אדם מאפשר חותך שונה עבור כל אדם.

הערה:

(4 נק') (אורך מקסימלי: 100 מילים)

הוסיפו למודל (5) משתני דמי לכל קטגוריה של *firm size* כמשתנים מסבירים. מודל זה ייקרא מודל (6).
האם תוצאות אלו תומכות בפרשנות שהאומד של *married* מהתוצאות של מודל (5) מייצג את ההשפעה הסיבתית של נישואין על לוג השכר?
הסבירו.

המקדם של נישואין עלה (במעט) לאחר הוספת משתנה הדאמי *firm size*. מאחר והוא בעל מקדם חיובי, זה מרמז לנו שיש קשר שלילי בין גודל החברה שבה הפרט עובד לנישואין.

מההבדל בין מודל 5 למודל 6 אנו רואים שנישואים מתואמים עם גורמים שנמצאים בט. כלומר, באותו אופן שהוספת משתנה *firm size* שינתה את מודל 5, כך ייתכן שישנם משתנים נוספים שמתואמים עם *married*, מפרים את ההנחות של המודל ולכן גורמים לאומד להיות מוטעה. לפיכך, האומד של *married* אינו מייצג בצורה מדויקת את ההשפעה של נישואים על השכר אלא מכיל גם את ההשפעה של גורמים אחרים (שנמצאים בט) על השכר.

הערה:

(4 נק')

מתוצאות מודל (6), מה ההשפעה הנאמדת של עבודה בחברה עם 30 עובדים שמפטרת 10 מהם?

תשובה: ❌

התשובה הנכונה היא: -0.10044807

שאלה 14

תקין

4.00 נקודות מתוך 4.00

(4 נק')

מתוצאות מודל (6), מה ההשפעה הנאמדת של עבודה בחברה עם 20 עובדים שמפטרות 10 מהם?

תשובה:



0

התשובה הנכונה היא: 0

שאלה 15

תקין

4.00 נקודות מתוך 4.00

(4 נק')

ניתן להסיק מתוצאות מודל (6) שלפחות אישה אחת במדגם סיימה תואר ראשון בתקופת המדגם

יש לבחור תשובה אחת:

☒ נכון

☐ לא-נכון

התשובה הנכונה היא: "נכון"

(4 נק') (אורך מקסימלי: 100 מילים)

בהקשר החומר שנדון בקורס, האם לדעתכם האומד על *married* במודל (5) מייצג את ההשפעה הסיבתית של נישואין על לוג השכר? הסבירו עם סיבות ספציפיות

לדעתי האומד על *married* לא מייצג את ההשפעה הסיבתית של נישואין על לוג השכר. אני סובר כך מכיוון שברגרסיה התייחסתי לנשים בלבד, אך ייתכנו הבדלים בהשפעה של נישואים בין גברים לנשים כך שהתוצאה שקיבלתי אינה מדויקת לנישואים באופן כללי. בנוסף ייתכנו דברים נוספים שגורמים לכך שהשכר בקרב נשואים נמוך יותר. דוגמא אפשרית לכך היא שמחוץ לעיר (שם המשכורות נמוכות יותר מבעיר) שיעור הנשואים גבוה יותר וכד'. כלומר, ייתכנו גורמים נוספים (שמתואמים עם נישואים) שהם הגורמים האמיתיים להשפעה על השכר אך לא הכנסנו אותם למודל. לכן אני חושב שנישואים **לא גורמים** לשכר להיות נמוך יותר כפי שניתן היה להסיק מהאומד שקיבלנו.

הערה:

(4 נק')

חוקרת רוצה להוסיף משתנים מסבירים למודל (5) על מנת לבדוק את ההשערה (H_0) כי ההשפעה של להיוולד בבריטניה קבועה לאורך שנות הסקר.



☒ דוחים את H_0

☐ לא דוחים את H_0

☐ לא ניתן לבדוק את ההשערה

תשובתך נכונה.

התשובה הנכונה:

דוחים את H_0

(אורך מקסימלי: 50 מילים)

הסבירו את תשובתכם

יצרתי משתנה שמקבל את המכפלה בין משתנה הדמי bornuk לבין השנה של המדגם. משתנה זה יקבל ערך שונה עבור כל שנה (במידה והפרט אכן נולד בבריטניה), לכן באמצעות הכנסה שלו למודל ניתן לראות האם ההשפעה של להיוולד בבריטניה קבועה לאורך שנות הסקר. ה-p-value שמתקבל נמוך (0.001) ולכן דוחים H_0 .

חלק ב' (30 נקודות)

בשאלה זו נבחן את ההשפעה של איכות סגל הפקולטה על התוצאות של תלמידי הדוקטורט במחלקה. נשתמש בניסוי טבעי שקרה בגרמניה הנאצית.

ב-1933, עם עליית כוחה של המפלגה הנאצית, החליטה הממשלה לפטר את כל החוקרים היהודים ואלו "שאינם מהימנים פוליטית" מהאוניברסיטאות הגרמניות. חלק מהאוניברסיטאות היו צריכות לפטר יותר מחצי מחברי הסגל שלהן בעוד שאוניברסיטאות אחרות כלל לא נפגעו, בהתאם לשיעור של חוקרים אלו בכל מחלקה. הגרפים הבאים ממחישים את ההשפעה של אירוע זה על אוניברסיטאות עם שיעור גבוה של פיטורים (קו מקווקו) ועל אוניברסיטאות בלי פיטורים כלל (קו רציף).

המשתנים:

- $PhDCitations_{uni,y}$ = מדד לפרודוקטיביות של בוגרי הדוקטורט מהפקולטה - מספר הציטוטים לאורך החיים של תלמידי שנתון (y) מאוניברסיטה (uni).
- $FQual_{uni,y}$ = מדד לאיכות סגל ההוראה - מספר הפרסומים (ממושקל לפי ציטוטים) לחבר סגל באוניברסיטה (uni) שהיה מועסק בשנה (y). ככל שהמדד מקבל ערך גבוה יותר - חברי הסגל הם יותר "איכותיים".



$$PhDCitations_{uni,y} = \beta_0 + \beta_1 FQual_{uni,y} + u_{uni,y}$$

(2 נק') (אנרך מקסימלי: 50 מילים)

נניח ויש לכם נתוני פאנל על פרודוקטיביות בוגרי הדוקטורט ועל איכות סגל ההוראה בפקולטה ברמת האוניברסיטה עבור כל שנה. מלבד טעות מדידה, תארו סיבה אחת לכך שהמשוואה לעיל לא תמדוד כראוי את הקשר הסיבתי בין איכות סגל ההוראה לפרודוקטיביות תלמידי הדוקטורט. התייחסו באופן ספציפי לאופן בו הניסוי הטבעי המוצג לעיל פותר את הבעיה.

סיבה אפשרית היא שFQual של המורים באוניברסיטה מסוימת מתואם עם גורמים נוספים באוניברסיטה שמשפיעים על phDCitations (דבר שגורם להטיה במודל). דוגמא לכך יכולה להיות המיקום של האוניברסיטה שיכול למשוך מרצים מאיכויות שונות. הפתרון הוא להסתכל על אותה האוניברסיטה לפני ואחרי הפיטורים ובכך לראות את הקפיצה שנגרמת בעקבות שינוי סגל ההוראה.

הערה:

6 נק') (אורך מקסימלי: 125 מילים)

בהתבסס על התיאור המופיע מעלה, הציעו מודל הפרש-ההפרשים שבאמצעותו תוכל לאמוד את ההשפעה הסיבתית של איכות סגל ההוראה על פרודוקטיביות בוגרי הדוקטורט. הגדירו באופן מדויק כל משתנה הנדרש לצורך האמידה ומהו האומד המעיד על סיבתיות.

נשתמש בשני מאגרי נתוני חתך ביחס למצב לפני ואחרי גל הפיטורים ב1933. בכל מאגר יש שני קבוצות: ישנה קבוצת הטיפול-האוניברסיטאות שהושפעו מגל הפיטורים, וישנה קבוצת ביקורת-האוניברסיטאות שלא הושפעו מגל הפיטורים. אנו נשווה בין השינויים בתוך כל קבוצה כדי לקבל את ההשפעה של הפיטורים.

אגדיר, $Tit=1$ קבוצת טיפול (0 אחרת)

$Y_t=1$ תצפית אחרי 1933 (0 אחרת)

נשים לב ש $Tit*Y_t=1$ רק אם שני האיברים שווים 1.

$$Y_{it} = \alpha + \beta_1 Y_{it} + \beta_2 Y_t + \beta_3 Tit * Y_t + u_{it}$$

כלומר, מחשבים את ההפרש בין $phDCitations$ (לפני הפיטורים) של קבוצת הביקורת לבין $phDCitations$ (לפני הפיטורים) של קבוצת הטיפול.

בנוסף מחשבים את ההפרש בין $phDCitations$ (אחרי הפיטורים) של קבוצת הביקורת לבין $phDCitations$ (אחרי הפיטורים) של קבוצת הטיפול.

לבסוף מחשבים את ההפרש בין שני ההפרשים שמצאנו וזהו האומד שמצביע על הסיבתיות בין איכות הסגל לפרודוקטיביות של בוגרי הדוקטורט של האוניברסיטה.

הערה:

(2 נק') (אורך מקסימלי: 50 מילים)

אם נדרשות הנחות כלשהן למודל זה כדי שהוא יעיד על קשר סיבתי, ציינו מה הן.

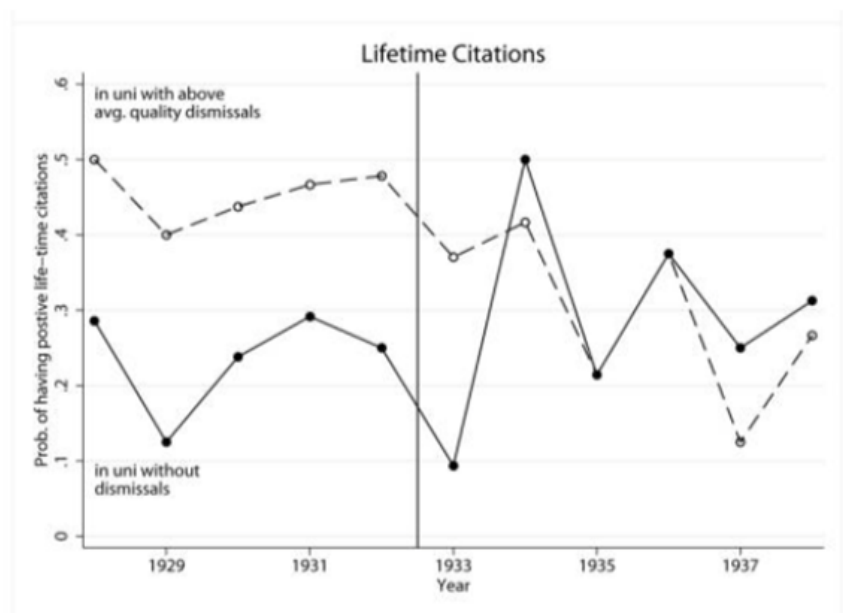
כדי שהמודל יעיד על קשר סיבתי אנו מניחים שבהיעדר גל הפיטורים, השינוי בתוצאות של קבוצת הביקורת יהיה זהה לשינוי בתוצאות של קבוצת הטיפול.

בנוסף לכך אנו מניחים שאין שינויים בלתי נצפים שנגרמו בעקבות גל הפיטורים (דבר זה יגרום לחוסר עיקביות של המודל).

הערה:

(6 נק') (אורך מקסימלי: 150 מילים)

1. הגרף המופיע בהמשך מציג את הרמה הממוצעת של פרודוקטיביות תלמידי הדוקטורט (PHDCitations) בכל שנה בחלוקה לפי אוניברסיטאות עם שיעור גבוה של פיטורים ושל אוניברסיטאות בלי פיטורים. בהתבסס על גרף זה, האם איכות חברי הסגל משפיעה על פרודוקטיביות בוגרי הדוקטורט? הסבירו וציינו באופן ברור את ההנחות הנדרשות.



לפי הגרף שלעיל לא ניתן לקבוע האם ישנה השפעה של איכות חברי הסגל על הפרודוקטיביות של בוגרי הדוקטורט.

אנו רואים כי לפני הפיטורים ב-1933 המגמות של הרמה הממוצעת של הפרודוקטיביות של תלמידי הדוקטורט משתנות בין קבוצת הביקורת לקבוצת הטיפול. בין השנים 1931-1932 הייתה עליה בפרודוקטיביות הממוצעת של קבוצת הטיפול ואילו במהלך השנים הללו הייתה ירידה בפרודוקטיביות הממוצעת של קבוצת הביקורת.

בנוסף, בין 1929-1931 העלייה בפרודוקטיביות הממוצעת בקבוצת הטיפול מתונה ולעומת זאת באותן השנים העלייה בקבוצת הביקורת הייתה חדה.

לפיכך, לא ניתן לקבוע שהגורם שהשפיע על הפרודוקטיביות של בוגרי הדוקטורט הוא דווקא איכות הסגל שהרי המגמות של הפרודוקטיביות היו שונות עוד לפני אירוע הפיטורים.

אציין שלא העובדה שהגרפים אינם נחתכים במדויק לפני הפיטורים היא שגרמה לי לקבוע שלא ניתן להכריע בשאלת ההשפעה של איכות חברי הסגל על הפרודוקטיביות של בוגרי הדוקטורט אלא שינוי המגמות שהיה בהם עוד לפני הפיטורים עצמם.

הערה:

(4 נק')

בכדי לוודא באופן סטטיסטי כי משתנה עזר מקיים את תנאי החיצוניות/השמטה (exclusion criteria), ניתן לבדוק את ...:

☐ הקורלציה במדגם בין המשתנה התלוי ומשתנה התוצאה

☐ יותר מאחת התשובות נכונה

☐ המובהקות הסטטיסטית ברגרסיה של משתנה העזר על המשתנה התלוי.



☒ אף תשובה אינה נכונה

☐ המובהקות הסטטיסטית ברגרסיה של משתנה העזר על המשתנה הבלתי-תלוי האנדוגני

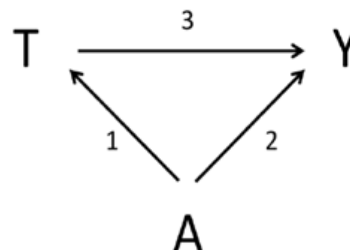
תשובתך נכונה.

התשובה הנכונה:

אף תשובה אינה נכונה

(4 נק')

התייחסו לגרף המופיע מטה כאשר T מציין את המשתנה הבלתי תלוי, Y את משתנה התוצאה ו- A הוא משתנה אחר. שימו לב לכך שהחיצים ממספרים וכיוון החץ מעיד על כיוון ההשפעה. הניחו כי אין עוד משתנים משפיעים.



על כל אחד מהמשפטים הבאים סמנו אם הוא נכון \ לא נכון:

- ☒ לא נכון בהיעדר חץ 1, אם מעניינת אותנו ההשפעה של T על Y צריך להשתמש ב- A כמשתנה בקרה
- ☒ לא נכון בהיעדר חץ 3 לא תהיה קורלציה בין T ובין Y .
- ☒ נכון בהיעדר חץ 2 היינו יכולים להשתמש ב- A כמשתנה עזר בעבור T
- ☒ לא נכון אם הייתה סיבתיות הפוכה: Y משפיע על T , מספיק לשלוט ב- A כדי לאמוד את ההשפעה של T על Y .

תשובתך נכונה.

התשובה הנכונה היא:

- בהיעדר חץ 1, אם מעניינת אותנו ההשפעה של T על Y צריך להשתמש ב- A כמשתנה בקרה – לא נכון,
- בהיעדר חץ 3 לא תהיה קורלציה בין T ובין Y . – לא נכון,
- בהיעדר חץ 2 היינו יכולים להשתמש ב- A כמשתנה עזר בעבור T – נכון,
- אם הייתה סיבתיות הפוכה: Y משפיע על T , מספיק לשלוט ב- A כדי לאמוד את ההשפעה של T על Y . – לא נכון

(6 נק')

להלן פלט קובץ נתונים על נשים לאחר לידה ומאפיינים שלהן. אתם מעוניינים לבדוק את ההשפעה של עישון בזמן ההיריון על משקל התינוק.

המשתנה התלוי *low_weight* הוא משתנה דמי השווה ל 1 אם התינוקת נולדת במשקל נמוך; ושווה 0 אחרת.

```
. des low_weight male white packs motheduc
```

variable name	storage type	display format	value label	variable label
low_weight	float	%9.0g		low birth weight (<107 oz)
male	byte	%8.0g		=1 if male child
white	byte	%8.0g		=1 if white
packs	float	%9.0g		packs smked per day while preg
motheduc	byte	%8.0g		mother's yrs of educ

```
. su low_weight male white packs motheduc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
low_weight	1,388	.2629683	.440404	0	1
male	1,388	.5208934	.4997433	0	1
white	1,388	.7845821	.4112601	0	1
packs	1,388	.1043588	.2986344	0	2.5
motheduc	1,387	12.93583	2.376728	2	18

```
. probit low_weight male white packs motheduc , robust
```

```
Iteration 0:  log pseudolikelihood = -798.34396
Iteration 1:  log pseudolikelihood = -764.23117
Iteration 2:  log pseudolikelihood = -764.18336
Iteration 3:  log pseudolikelihood = -764.18336
```

```
Probit regression
```

Number of obs	=	1,387
Wald chi2(4)	=	62.30
Prob > chi2	=	0.0000
Pseudo R2	=	0.0428

```
Log pseudolikelihood = -764.18336
```

low_weight	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
male	-.2926817	.0740056	-3.95	0.000	-.4377299 -.1476335
white	-.4430008	.0869408	-5.10	0.000	-.6134016 -.2726
packs	.5835045	.1265521	4.61	0.000	.335467 .831542
motheduc	-.0074411	.0162725	-0.46	0.647	-.0393346 .0244525
_cons	-.1253025	.2210088	-0.57	0.571	-.5584717 .3078668

$$p_{xb} = \Phi(xb)$$

	x _b	p _{x_b}
1.	-2	.0227501
2.	-1.9	.0287166
3.	-1.8	.0359303
4.	-1.7	.0445655
5.	-1.6	.0547993
6.	-1.5	.0668072
7.	-1.4	.0807567
8.	-1.3	.0968005
9.	-1.2	.1150697
10.	-1	.1586553
11.	-.9	.1840601
12.	-.8	.2118554
13.	-.7	.2419637
14.	-.6	.2742531
15.	-.5	.3085375
16.	-.4	.3445783
17.	-.3	.3820886
18.	-.2	.4207403
19.	-.1	.4601722
20.	0	.5
21.	.1	.5398278
22.	.2	.5792597
23.	.3	.6179114
24.	.4	.6554217
25.	.5	.6914625

באמצעות המידע הניתן בצילומי המסך וכאשר המשתנים *motheduc*, *white* ו-*male* שווים לערכים הממוצעים שלהם (בהתאמה), במעבר מעישון חבילה אחת של סיגריות ביום לעישון שתי חבילות של סיגריות ליום מהו השינוי בהסתברות שהתינוק יהיה במשקל נמוך? אם לא ניתן לחשב תשובה מדויקת חשבו בקירוב הטוב ביותר והסבירו בקצרה בתוך הקוד.


תשובה:

התשובה הנכונה היא: 0.2098515

אנא חיזרו וודאו כי הקבצים שהגשתם ניתנים לקריאה ומכילים את כל התשובות שלכם לפני שאתם מגישים את הבחינה. קובץ חסר או שגוי יהיה מבחינתנו כמו להגיש את המחברת הלא נכונה למשיחות: לא תהיה לנו דרך. לוודא שאכן ביצעתם את הבחינה כנדרש, הדבר עלול לפגוע בציון שלכם באופן משמעותי - וחבל

בהצלחה,
צוות הקורס.

[Exam log file.do](#) 

[Final Exam.log](#) 

► טופס ערעור - המחלקה לכלכלה

מעבר ל...

הנחיות למבחן מקוון 2021 ◀