1. שלום, שמי יבגני קולונסקי.  
    היום אני בכובע של מרצה בכיר במתמטיקה יישומית.  
    יוליה הזמנה אותי לתת לכם כמה טיפים טובים בניתוח תוצאות מדידה כדי שתוכלו להשתמש בהם גם במחקר וגם בהוראה. הטיפ הראשון - קצת טריוויאלי אבל טוב - בבקשה תשתמשו בממוצע משוקלל.
2. תזכורת קטנה על ממוצע אריתמטי.  
    אם יצא כך שמדדנו N מדידות, אנחנו כמעט אוטומטית נוטים לחשב ממוצע אריתמטי וגם כך אנחנו מלמדים את הסטודנטים.  
    זה בסדר גמור! רק צריך לזכור שזה תיקף בשתי הנחות:  
    א') המדידות היו בלתי תלויות  
    ב') כל מדידה שייכת לאותה ההתפלגות האקראית גאוסיאנית.  
    כאשר זה נכון, חוץ מזה שאנחנו יכולים לחשב את הממוצע, ולתאר בעזרתו את המדידה שלנו, אנחנו גם יכולים לחשב סטיית תקן וגם סטיית התקן של הממוצע הזה. מעולה!  
    אבל מה לעשות אם מדידות לא שייכות לאותה ההתפלגות?
3. לדוגמה ניקח מדידה XA ששייכת להתפלגות A ומדידה XB ששייכת להתפלגות B, שתי המדידות קרובות אחת לשנייה בטווח השגיאה.  
    נרצה למקסם את פונקציית ההסתברות P שהיא מכפלה של הסתברויות A והסתברות B.  
    כדי למקסם את P נצטרך לפתור את המשווה הזאת.  
    הפתרון של המשוואה נותן לנו הגדרה של ממוצע משוקלל וסטית התקן החדשה.  
    נוח להגדיר אחד חלקי סיגמא בריבוע של מדידת A כ"משקל" של מדידה A וסיגמא בריבוע של מדידת B כ"משקל" של מדידה B.
4. עכשיו אתן לכם דוגמה פשוטה וברורה.  
    בניסוי של מעבדה 2 "קבל לוחות" אנחנו מודדים קיבול בכמה דרכים.  
    אני מציג לכם רק שתי שיטות: השיטה המבוססת פריקה והשיטה הגאומטרית.  
    בשיטת הפריקה קיבלנו ערך של הקיבול 445- פיקו-פאראד עם שגיאה 10 פיקו-פאראד.  
    בשיטה הגאומטרית קיבלנו 430 פיקו-פאראד עם שגיאה של 20 פיקו-פאראד.  
    שאלת קיטבג: מה יותר נכון לעשות – לחשב ממוצע אריתמטי או משוקלל? נכון!
5. אני רואה שאתם בפוקוס מלא, זאת אומרת הגיע זמן לקוד קצר נחמד.  
    אני רושם קיבול עם שגיאה שלו לכול שיטה בצורה קריאה ומשׂמחת עין.  
    אחרי זה מחשב משקל של כל מדידה, ממוצע משוקלל וסטיית תקן חדשה.  
    יצא לי ארבע מאות ארבעים ושניים פיקו-פאראד.  
    זה קצת לא צפוי, נכון?  
    גם פחות צפוי ששגיאה למדידה ששקללנו – קָטְנה!  
    מעניין, נכון? אני ממליץ בחום!
6. עכשיו אני רוצה להזכיר לכם איך להשתמש במבחן χ² ואולי אצליח לשכנע אותכם להשתמש בו בהוראה.

כמה מכם כבר משתמשים בו בהוראה? מעולה! כל הכבוד

מי שלא כול כך משתמש עדיין – אל דאגו – אספר בצורה קלילה עם דוגמאות.

1. נתחיל מהגדרה מספיק מדויקת למטרות שלנו.  
    נעשה ניסוי מחשבתי אידאלי, ונקבל ***n*** מדידות.  
    כל מדידה שלנו עמוסה בשגיאה אקראית כלשהי שמפולגת גאוסיאנית עם סטיית תקן.  
    נניח יש לנו גם מודל מתמטי שיכול לתת ערך צפוי לכל מדידה.  
    בתנאים האלה נוכל לחשב χ² לפי הנוסחה (). מעולה.  
    זה כבר ערך מעניין אבל יותר נוח להשתמש בχ² מצומצם.
2. הגדרת χ² מצומצם מאוד פשוטה – זה χ² חלקי מספר דרגות חופש.  
    קל להגדיר מבחן χ² על ידי המספר הזה.  
    באופן גס: אם χ² מצומצם יותר גדול משמעותית מ1 – המדידות שלנו לא תואמות מודל.  
    או במילים אחרות הסתברות לחפיפה בין מדידות וערך הצפוי קנטה עד זניחה.  
    כאשר χ² מצומצם קטן מאחד כבר יש על מה לדבר.  
    אני מיד מביא לכם דוגמה פשוטה.  
    רק חשוב לי להזכיר לכם מה כוונה מאחורי מושג "דרגות חופש"
3. נניח מדדתי 10 מדידות או יותר מדויק להגיד 10 זוגות - (x,y)  
    .המודל שבחרתי מניח שיש קשר פולינומיאלי בחזקה 2 בין x לy.  
    לבחירתי 3 אופציות התאמה: איבר ריבוע בלבד, איבר ריבועי פלוס אופ-סט, או איבר ריבועי פלוס איבר לינארי פלוס אופ-סט.  
    במקרה הראשון יהיה 9 דרגות חופש, במקרה השני – שמונה דרגו חופש ו במקרה השלישי – שבע דרגות חופש.  
    כדי לבדוק שלא נרדמתם כבר עכשיו אני אשאל: מה יותר טוב – יותר דרגות חופש או פחות? מעולה.  
    מתקדמים הלאה ונבדוק על הדוגמה.
4. ניקח דוגמה של אלקטרון חופשי נע בשדה חשמלי אחיד ניצב למהירות של האלקטרון.  
    דוגמה מוכרת ופשוטה.  
    אלקטרון מואץ בעזרת אלקטרודות האצה ונכנס לשדה אחיד של קבל לוחות.  
    תחת השפעה של השדה החשמלי של הלוחות אלקטרון מוסט לכיוון אחד מהלוחות - תלוי בסימן המטען של הלוח.  
    המסלול שאלקטרון יעשה בין הלוחות יהיה פרבולי לפי המודל הכי פשוט.  
    האם יש צורך להזכיר איך מקבלים את הביטוי? מעולה, לא צריך.  
    כמו ש אתם רואים בתמונה – אנחנו הולכים למדוד ולבדוק את זה.
5. כמו שרואים בעין – המסלול בוודאות לא לינארי.  
    זאת מסקנה איכותנית.  
    אנחנו מעוניינים לקבל מסקנה כמותית.  
    בעזרת קנהֶ מידה שיש בתמונה (כל ריבוע הוא סנטימטר לסנטימטר), נוכל להפוך את המידע ויזואלי לזוגות מדידות (x,y).  
    מעולה. אני באופן ידני סימנתי 7 מדידות וחישבתי χ².
6. בוא נקרא ביחד קוד קצר שכתבתי – הוא קל וקריא.  
    נעשה יבוא של הספרייה הנדרשת.  
    נרשום זוגות מדידות שלנו. אני אוהב לעשות ככה.  
    כאן – נקודה חשובה!  
    כל מדידה צריכה להיות מלווה בערכת שגיאה שלה.  
    לפעמים הערכה נוּבעַת מהרבה מדידות ואנחנו מסוגלים לחשב סטיית תקן, ולפעמים יש רק מדידה בדידה ואנחנו עושים עכרת שגיאה משיקולים אחרים.  
    במקרה הזה אני מעריך שגיאה כחצי מילימטר.  
    מה אתם חושבים על הערכה שלי?  
    פסימית? אופטימית? אופטימלית? נראה ביחד.  
    אחרי זה אני מגדיר מודל.  
    מיד אחרי זה עושה התאמה ומחשב χ².  
    כמו שאתם רואים χ² יצא 0.64.  
    זה טוב לדעתכם?  
    אתם זוכרים את המבחן שהצגתי לכם לפני 5 דקות?  
    נכון, זה טוב, כי פחות מ1.  
    אבל כבר יש לנו קריטריון כזה המום וקוראים לו R².  
    למה χ² יותר טוב?
7. מי שיודע תְכוּנוֹת של התפלגות חי, הוא יכול לתת מסקנה עמוקה במבוססת.  
    מה זה התפלגות χ²?  
    זאת אחת מההתפלגויות סטטיסטיות ידועות כמו שהתפלגות פואסונית או גאוסיאנית.  
    בשקף הזה הצגתי אותה לכמה פרמטרים רלוונטיים לדיון שלנו.  
    ננסה להבין מה נותנת לנו את ההתפלגות הזאת.  
    קיבלנו χ² מצומצם 0.64. מעולה.  
    בציר אופקי אני אמצא את 0.64 בערך.  
    אני זוכר שמספר דרגות חופש היה 6.  
    הקו הכחול נותן לי הסתברות 70%. אוקי.  
    בואו נחבר הכול ביחד למסקנה:  
    מדדתי 7 מדידות עם שגיאת מדידה חצי מילימטר,  
    הנחתי מודל פרבולי עם פרמטר התאמה אחד,  
    חישבתי χ² מצומצם שאומר -  
    הסתברות שהמודל המתמטי הזה תקף לסט של המדידות היא – 70%. וואו!  
    זאת תשובה מבוססת.  
    בלי אפילו לראות איך נתונים יושבים על הקו.
8. בכל זאת נראה התאמה גם בצורה גרפית.   
   האם טוב לי 70%? איך אפשר לשפר? נכון!  
    גם מהתפלגות אנחנו רואים שכדאי לנו למדוד יותר זוגות (x,y).
9. לקחתי 20 מדידות עם שגיאת מדידה חצי מילימטר,  
    הנחתי מודל פרבולי עם פרמטר התאמה אחד,  
    חישבתי χ² מצומצם שאומר -  
    הסתברות שהמודל המתמטי הזה תָקַף לסט של 20 מדידות היא – 93%.  
   אני רוצה להדגיש כאן שני היתרונות של χ² מצומצם:  
    א') הוא נותן הערכה סטטיסטית לסיכוי התאמה  
    ב') הוא מתייחס לדיוק של המדידות!  
    ולא רק מתייחס הוא גם מאוד רגיש להערכת שגיאה.  
    אני מיד נותן דוגמה...
10. החלטתי להקטין הערכת שגיאה לחצי.  
     במקום חצי מילימטר – אניח שהשגיאה היא רבע מילימטר.   
    לאותו סט הנתונים קיבלתי הסתברות 1%!  
     המודל כבר לא מסוגל לתאר את הנותנים למרות שוויזואלית עדיין יש רוֹשֶם של התאמה.
11. האם הצלחתי לשכנע אותם שχ² זה דבר טוב יותר מאר בריבוע?
12. נשאר לנו עוד פרק מאוד מעניין של אנליזת נתונים.  
     אתחיל אותו מתזכורת קצרה על רגרסיה לינארית.
13. זה כלי קלאסי אמין ומפורסם.  
     כולכם משתמשים ברגרסיה לינארית להוראה ואני רק רוצה להזכיר את ההנחות בה רגרסיה מוגדרת היטב.  
     ההנחה הראשונה ששגיאה בציר X משמעותית קטנה (זניחה) משגיאה בציר Y.  
     שנית - מאוד חשוב שכל נקודה תהיה עמוסה בשגיאה זהה בציר Y.
14. מתִגרוּל פייתון של מעבדה 2 אנחנו יודעים שאפשר לחשב רגרסיה על ידי קוד פשוט בעזרת ספריה linregress . הזנתי תוצאות מדידה לשני וקטורים. אני מגדיר רגרסיה ומחשב שיפוע ונקודת חיתוך עם ציר Y.  
     אלה תוצאות החישוב. מצוין.
15. כאשר לפי המודל התאורטי אני צריך לקַבֶּעַ איבר חופשי כאפס אני בבעיה עם הקוד שכתבתי לפני.  
     הספריה לא מאפשר לאפס איבר חופשי בכוח.  
     אני מציע לכם פתרון.  
     בבקשה תגידו מה אתם חושבים על זה.  
     אני מציע לחשב שיפוע לכל זוג נתונים לכוד.  
     בהנחה ששגיאה בציר X זניחה אני יכול להגיד ששגיאה יחסית בשיפוע שווה של שגיאה יחסית בY. ו...
16. אפשר ליישׂם ממוצע משוקלל שהסברתי לכם קודם! מגניב!  
     תראו אני מגדיר את פונקציית הממוצע המשוקלל ומחשב את השיפוע והשגיאה שלו.  
     לרוב המקרים או מודל לא לינאריים או מדידה עקיפה. כמו שידוע לכם אנחנו משתדלים לבחור לינאריזציה מתאימה למודל כדי בכל זאת להשתמש ברגרסיה לינארית – כול כך אנחנו אוהבים אותה.  
     אראה לכם כמה דוגמאות.
17. במעבדה 4 יש ניסוי יחסית פשוט - בליעת קרינת **γ** על ידי פלטות של עופרת בעוביים שונים.   
    אנחנו ממקמים פלטה מעופרת בין מקור הקרינה לגלאי וסופרים קריאות של הגלאי.  
     תאורטית אנחנו מצפים קשר אקספוננציאלי בין קצב הקרינה ועובי של בולע מעופרת.  
     אנחנו מהנדסים ניסוי כך ששגיאה בכל מדידה תהיה כ3%.  
     נרצה לעשות לינאריזציה למודל. מצוין!  
     למה? כי שמרנו על הנחות לרגרסיה לינארית: מדדנו את העוביים בדיוק גבוה שמאפשר להזניח שגיאה בציר X, ושגיאות בציר Y כמעט זהות.  
     זאת דוגמה מצוינת איך אפשר לעשות לינאריזציה כדי להשתמש ברגרסיה לינארית בגבולות הגְדַרַתֶיִהַ.
18. עוד דוגמה. עכשיו ממעבדה 2.  
     אסביר בקצרה.  
     אנחנו מחבְּרים קבל לוחות למֵחוֹלֵל אות חשמלי מרובע.  
     זה מאפשר לטעוֹן ולִפְרוֹק את הקבל העגול שאתם רואים בתמונה.  
     המודל מתאר את הפריקה הוא אקספוננציאלי.  
     אני מעוניין למדוד זמן הפריקה של הקבל ולחלץ ממנו קיבל.  
     כמו שבמקרה הקודם עשיתי לינאריזציה למודל.  
     מה אני רואה בנתונים?  
     בזמנים ארוּכּים אני מגלה קִיוּם של ארטִפקט.  
     אתם רואים מדרגות?  
     כן, זה קשור לאופן הפעולה של האוסצלוסקופ.  
     אבל מה שיותר חשוב שהסיגנל המקורי שלנו היה מדוד עם אותו ערך השגיאה לכל נקודה.  
     אם אנחנו נצייר קוי שגיאה בציר LOG הם יהיו שונים בסדרי גודל.  
     זה מונע מאיתנו להשתמש ברגרסיה לינארית בגדול.  
     אבל מה לעשות? בכל זאת נשתמש. רק קצת יותר חכם.  
     אקח רק מדידות מהחתלה עד לבערך אמצע ואעביר רגרסיה.  
     אני רואה שקודם כל זה יותר פיזיקאלי ושנית זה יותר מתאים לתוצאות של השיטות האחרות להערכה של קיבול.  
     בסדר גמור
19. ועוד הדוגמה האחרונה בנושא.  
     זאת דוגמה ממעבדה 3.  
     בניסוי של גלי מיקרו סטודנטים מתבקשים לבדוק שחוק מַאלוּס מתקיים.  
     החוק האמפירי הזה אומר שעוצמת הגל האלקטרו-מגנטי המקוּטַב של המֵשַדֵר לאחר עברה דרך המקטב תהיה פרופורציונלית לריבוע של קוסינוס הזווית של המקטב.  
     כדי להוכיח את החוק מבקשים לעשות לינארזציה המתאימה לחוק מאלוס ולהראות שהקשר הוא לינארי.  
     כמו שאתם רואים בגרף אין טעם לעשות רגרסיה לינארית כי לפי העין רואים שאין התאמה.  
     חוץ מזה אנחנו לא עומדים אפילו קרוב לדרישה ששגיאה בציר X זניחה ביחס לשגיאה בציר Y.  
     בניסוי הזה – הפוך.  
     מה רציתי לוֹמָר – שלא תמיד אפשר וכדאי לעשות לינארזציה.   
    מה לעשות??? נכון מאוד! התאמה לא לינארית.
20. יאללה. ניקח דוגמה חמודה של התאמה לפונקציה מחזורית בעלת 4 פרמטרי התאמה.  
     דרך אגב זה מהניסוי מתנד הרמוני לא מרוסן.  
     נגדיר פונקציה מחזורית.  
     אתם איתי?  
     לפעמים מומלץ גם להגדיר פרמטרי התחלה.  
     אתם רואים שיצאה התאמה איכותית. מעולה!
21. כאשר יש ריסון אפילו קטן אנחנו צריכים להוסיף גם דעיכה למודל.  
     אנחנו מגלים שהתאמה לא מתכַּנֶסֶת.  
     באופן כללי כאשר יש במודל יותר פרמטרים לספרייה הזאת יותר "קשה" להתכנס.  
     כדי "לעזור" לאלגוריתם להתכנס אנחנו מגדירים גם גבולות השתנות של פרמטרי ההתאמה.  
     סוג של פיגומים לאלגוריתם.
22. חחח. כן, עכשיו שישה פרמטרים.  
     מה שרציתי להראות לכם בדוגמה החישובית הזאת שכמעט שום דבר יותר לא עוזר לאלגוריתם להתכנס - חוץ מלדעת מראש את הערכים שאנחנו מחפשים.  
     קצת מצחיק, לא? לא נורא.  
     עכשיו אני אתן לכם טיפ מאה דולר. תרשמו.
23. פשוט צריך להשתמש בספריות יודעות לחפש המינימום הגלובלי. כתבתי לנוחותכם קוד שמיישם ספרייה בשם LMFIT. מכירים?  
     מה שחשוב לציין. אני ממליץ לכם להגדיר שיטה **חיִשׁוּל כַפוּל**.  
     מי שרוצה לדעת על זה יותר מוזמן לדבר איתי!
24. תודה רבה על ריכוז ושיתוף הפעולה שלכם!