ECON 0150 | Spring 2025 | Homework 4.2

Due:

Homework is designed to both test your knowledge and challenge you to apply familiar concepts in new applications. Answer clearly and completely. You are welcomed and encouraged to work in groups so long as your work is your own. Use the provided datasets to answer the following questions. Then submit your figures and answers to Gradescope.

Q1. Working Hours and Weekly Income

An economist is studying the relationship between time spent working (hours per week) and weekly income (in dollars) for part-time employees. The data for 15 randomly selected employees is:

```
hours_worked = [15, 22, 12, 25, 18, 10, 20, 28, 14, 16, 24, 19, 21, 17, 23]
weekly_income = [320, 465, 255, 510, 385, 210, 420, 570, 305, 340,
                 490, 400, 445, 360, 475]
```

a) Create a scatter plot of weekly income versus hours worked. Does there appear to be a relationship?

b) Fit a linear regression model to predict weekly income based on hours worked. What is the interpretation of the intercept coefficient?

c) What is the interpretation of the slope coefficient?

d) What is the interpretation of the p-value of the slope coefficient?

Q2. County Unemployment and BMI

Read the article: Zhang, Q., Lamichhane, R., & Wang, Y. (2014). Associations between US adult obesity and state and county economic conditions in the recession. Journal of clinical medicine, 3(1), 153-166.
https://pmc.ncbi.nlm.nih.gov/articles/PMC4449673/

a) What is the primary research question the authors are trying to answer?

b) What dataset(s) do the authors use for their analysis? Be specific about the years and key variables.

c) What are the main outcome variables, and what are the key explanatory variables?

d) Summarize the main findings in 2–3 sentences. What is the estimated effect of county unemployment on BMI and physical activity? Is it statistically significant?

b)

You are provided with a sample of individual-level data from the 2011 Behavioral Risk Factor Surveillance System (BRFSS), which includes the following variables:

- `bmi` – individual Body Mass Index (kg/m²)
- `Mental Healt` – days of Poor Mental Health in the Past 30 Days
- `county_unemp` – county-level unemployment rate (in percentage points)

Using this data, estimate the following linear regression model:

$$\texttt{bmi}_i = \beta_0 + \beta_1 \cdot \texttt{county\_unemp}_i + \varepsilon_i \tag{1}$$

1. Report the estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$, along with their standard errors, t-statistics, and p-values. Also attach the code you run.

2. Interpret the coefficient $\hat{\beta}_1$. What does it suggest about the relationship between county unemployment and individual BMI?

3. Identify the corresponding table and column in the Chen et al. (2014) paper where a similar regression is reported. How does your estimated coefficient compare in magnitude, sign, and statistical significance?

c)

The 2011 BRFSS also includes information on mental health. The variable `menthlth` records the number of days in the past 30 days when the respondent's mental health was "not good."

1. Estimate the following regression model:

$$\texttt{menthlth}_i = \beta_0 + \beta_1 \cdot \texttt{county\_unemp}_i + \varepsilon_i \tag{2}$$

Report and interpret the estimated coefficient $\hat{\beta}_1$.