Public Policy 529
Fall 2023: Problem Set #10

Due Wednesday, December 6

1. In the regression analysis below, the dependent variable is the number of minutes it takes for an ambulance to arrive when called to an emergency. The independent variable is the distance in miles from the location where the ambulance is waiting to the location of the emergency. You have data on n=52 cases in which an ambulance was called to an emergency.

$$\text{Minutes}_i = \beta_0 + \beta_1(\text{Miles}_i) + u_i$$

Variable	Coefficient	Standard Error
Intercept	1.2	0.7
Miles	1.8	0.4
		$R^2 = .71$

   (a) Interpret the estimated coefficient on Miles ($\hat{\beta}_1$). In other words, what does it predict about the relationship between distance and the time it takes for the ambulance to arrive?

   (b) Are these two coefficients statistically significant? In other words, for each estimated coefficient, can we reject the null hypothesis that the true coefficient is zero? Provide evidence for your answer by finding the t-statistic and p-value for each coefficient.

   (c) Suppose there is an emergency 6.2 miles from where the ambulance is waiting. What do we predict is the number of minutes it will take for the ambulance to arrive?

2. Use the dataset `StateData_2018` for this question. This dataset contains variables measured at the state level in the United States. You will estimate a bivariate linear regression using the following variables:

   - The dependent variable is `gun_deathrate`, which is the number of firearm-related deaths per 100,000 in population in the state. In the sample, this variable ranges from 3.3 to 22.9.
   - The independent variable is `lawtotal`, which is the number of laws restricting

gun ownership in the state. Specifically, it is a count of different types of possible gun restrictions: requiring a person be 21 years-old to possess a handgun, requiring a person be 21 years-old to purchase a handgun, a ban on "assault-style" weapons, a requirement for universal background checks for gun sales, etc. In the sample, this variable ranges from 1 to 109.

(a) Estimate the model and report the output. See lecture slides or help documents for appropriate commands.

(b) What is the substantive meaning of the estimated intercept ($\hat{\beta}_0$)?

(c) Interpret the substantive meaning of the coefficient on `lawtotal` ($\hat{\beta}_1$).

(d) Assess the statistical significance of each of these coefficients at $\alpha = .05$.

(e) Suppose a state has 20 laws that involve a restriction on gun ownership. What do we predict would be the level of firearm-related deaths in that state?

(f) Suppose this state were to pass five additional laws restricting gun ownership. What is the change in the predicted level of firearm-related deaths?

(g) Interpret the R^2 for this regression.

(h) Interpret the standard error of the estimate from this regression. In other words, what does this number mean substantively?

3. Continue with the `StateData_2018` for this question. You will estimate a bivariate regression using the following variables:

   - The dependent variable is `pctui`, which is the percentage of a state's population that does not have health insurance. In the sample, this variable ranges from 3.3 to 19.9.
   - The independent variable is `MedicaidExp`, which is a dichotomous (0,1) indicator of whether a state chose to expand Medicaid eligibility as part of the Affordable Care Act (ACA). The variable is coded as 1 if the state expanded Medicaid and 0 otherwise. As of 2018, 32 states, plus the District of Columbia, had expanded Medicaid eligibility, while 18 states did not.

   Since the independent variable is dichotomous, a bivariate regression is essentially a difference of means test.

   (a) Estimate the model and report the output.

   (b) What is the substantive meaning of the estimated intercept ($\hat{\beta}_0$)?

   (c) Interpret the substantive meaning of the coefficient on `MedicaidExp` ($\hat{\beta}_1$).

   (d) What do we predict is the mean level of `pctui` in a state that has expanded Medicaid as part of the ACA?

   (e) Using your software, perform a t-test for the difference of means. Identify on the output the items that match up with your regression output. Note: the regression

method for a difference of means test is equivalent to using the equal variance assumption in a two-sample t-test, so the relevant t-statistics will generally differ somewhat unless you apply the equal variance assumption in the t-test (which I would not recommend in this case).