# Final - Fall 2024

*ECON 4753 — University of Arkansas*

1. (10 pts) In this course, we have described a set of methods that are very 'flexible' at modeling $y = f(X)$. We have also learned about a simple linear regression model ($y = \alpha + X\beta$). Describe one advantage of each method (a flexible model and a linear model).

    **Answer:** A more flexible model is able to better describe the relationship between $X$ and $y$ (e.g. if the relationship is non-linear in $X$). The advantage of a more simple linear regression model is that you can describe the result simply because it is a line (intercept and slope).

2. The following regression uses the "College Scorecard" which describes all U.S. colleges/universities. The outcome variable is the average annual earnings ($) of students 10 years after they enroll. The explanatory varaible is the median SAT Math score of the student body. I include both the variable itself and its square (quadratic in SAT math):

    ```
    OLS estimation, Dep. Var.: mean_earnings_10yr_after

    Observations: 935

    Standard-errors: Heteroskedasticity-robust

                             Estimate   Std. Error  t value    Pr(>|t|)
    (Intercept)          108369.503822 19595.243683  5.53040 4.1500e-08 ***

    sat_math_median        -337.941666    68.781472 -4.91327 1.0577e-06 ***

    I(sat_math_median^2)      0.411678     0.059815  6.88258 1.0783e-11 ***

    ---

    Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    ```

    i. (10 pts) What is the predicted earnings for a school with an average SAT math score of 500 (round to the nearest dollar)?

       **Answer:** $\hat{y} = 108369.50 - 337.941666 * 500 + 0.411678 * 500^2 = \$42318$

    ii. (10 pts) Say you take a school with an average SAT math score of 500. What is the predicted marginal change in $Y$ for a school with a 1 unit increase in average SAT math score?

**Answer:** Our estimated marginal change is given by $\hat{\beta}_1 + 2 * \hat{\beta}_2 * 500 = -337.941666 + 2 * 0.411678 * 500 = 73.73$. In words, a school with one point higher in SAT math is expected to have \$73.73 higher average annual earnings.

3. Say you have a sample of stores where you observe the average daily revenue and the number of employees on the sales floor. You regress the $\log$ of average daily revenue on the number of employees and estimate a coefficient of $\hat{\beta}_1 = 0.03$ and a standard error of $\text{SE}(\hat{\beta}_1) = 0.005$.

   i. (10 pts) Interpret this coefficient estimate in words.

   **Answer:** For every one additional employee on the sales floor, we predict sales to be $3\%$ higher.

   ii. (5 pts) The company does not want to increase the number of staff if these results are not statistically significant. Perform a test of the null that $\beta_1 = 0$. The company is risk adverse and want you to use a level of significance of $\alpha = 0.01$ (the z-score associated with this is 2.58).

   **Answer:** The confidence interval can be calculated as: $(0.03 - 2.58 * 0.005, 0.03 + 2.58 * 0.005) = (0.0171, 0.0429)$. Since this confidence interval does not contain 0, we can reject the null that $\beta_1 = 0$ with a $\alpha = 0.01$ level of significance.

   Alternatively, we could calculate the $t$-statistic as

   $$\frac{\hat{\beta}_1 - 0}{\text{SE}(\hat{\beta}_1)} = \frac{0.03 - 0}{0.005} = 6$$

   Since 6 is larger than the critical value of 2.58, we reject the null.

4. For the following questions, we will look at daily bitcoin price data (see Figure 1).

   i. (15 pts) Consider conducing inference on this time-series using a 30-day one-sided moving average. How do you think this method would perform on the bitcoin price data? Please explain why.

   **Answer:** This method will do a bad job at predicting the bitcoin because (i) it will oversmooth and miss short-run significant fluctuations (e.g. the bubble in 2021) and (ii) will also do a bad job at following the general upwards trend in the recent two years

   ii. (10 pts) Say, instead, you were to use a flexible piecewise linear function (say 10+ breaks)

to forecast future bitcoin prices. Why might you be concerned about extrapolating your regression estimate into the future to predict bitcoin prices.

**Answer:** The bitcoin forecast will have a line with a very large slope in recent years (because of the recent jump post Trump's election). If this line is extrapolated into the future, it would suggest that bitcoin's price will shoot off very quickly. There is a good chance this recent 'burst' will plateau or go away (as in a bubble).
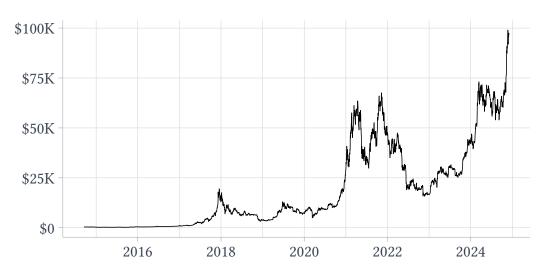
**Figure 1 — Daily Closing Prices of Bitcoin**



5. (10 pts) Say you take a sample of size 64 and estimate a sample mean of $\bar{X} = 40.6$. Additionally, you know the variance of $X$, $\sigma^2 = 24$. Construct a 95% confidence interval for this sample mean.

**Answer:** The standard deviation of the sample mean is $\sqrt{24}/\sqrt{64} = 0.612$. Therefore the 95% confidence interval can be formed as

$$40.6 \pm 1.96 * 0.612 \approx (39.4, 41.8)$$

6. The following regression uses the "College Scorecard" which describes all U.S. colleges/universities. The outcome variable is the average earnings of students 10 years after they enroll. There

are two covariates including `hbcu` which is an indicator if the college is a historically-Black college or univeristy and `share_low_income` which is the share of students considered 'low income' and takes values between 0 and 1.

```
OLS estimation, Dep. Var.: mean_earnings_10yr_after
Observations: 2,078
                 Estimate Std. Error   t value  Pr(>|t|)
(Intercept)      65281.86    809.581  80.63656 < 2.2e-16 ***
hbcu::1          -6752.49    701.296  -9.62859 < 2.2e-16 ***
share_low_income -40604.40   1840.844 -22.05749 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

i. (10 pts) Interpret in words the estimated coefficient on the `hbcu` indicator.

**Answer:** An HBCU college is on average expected to have \$6752 less in average earnings than a non-HBCU college.

ii. (10 pts) Interpret in words the estimated coefficient on the `share_low_income` variable. Is a 'one unit' increase a meaningful quantity in this case?

**Answer:** A school with a 1 unit increase in share (i.e. going from a school with a 0% low-income share to a school with 100%), we predict earnings to be \$40,000 lower. This is an extreme quantity, so instead, we can interpret a 10% increase in low-income share (change of 0.1) would predict a decrease in earnings of about \$4,000.