

Economics 250
Introductory Statistics
Exercise 1

Due Wednesday 3 February 2021 by noon via onQ

Instructions: No late submissions are possible. Show intermediate steps (such as formulas) in your work both for ease of review and part marks in case of error. Do the exercise on your own and submit only your own work. Plots and graphs must be drawn with software (such as Excel) not by hand.

1. Imagine that a portfolio's investment returns in eight successive years take these values (in percentage points): -5.0, 2.8, 3.0, -1.5, 2.1, 2.8, -1.0, 3.3

- (a) Find the mean, median and mode of these returns.
- (b) Is the distribution symmetric, skewed right, or skewed left? What indicator did you use to answer this question?
- (c) Find the interquartile range.
- (d) Find the sample variance and the sample standard deviation.

2. A researcher is studying percentage scores of 20 people who have written an aptitude test. She sees 4 scores of 69%, 10 scores of 76%, and 6 scores of 84%.

- (a) Find the median and mode of these scores.
- (b) Find the sample mean and the sample standard deviation of these scores.

3. Suppose that waiting times in a dental office are continuously and uniformly distributed between 5 minutes and 15 minutes.

- (a) What is the probability of waiting less than 9 minutes?
- (b) What is the probability of waiting exactly 7 minutes?
- (c) What is the standard deviation of the waiting time?

4. Suppose that $z \sim N(0, 1)$.

- (a) Find the probability that $z > 1$.
- (b) Find the probability that $z \leq -1.96$.
- (c) Find the probability that $0 < z < 1.645$.
- (d) Find the value z_c such that there is a 34% probability that $-z_c < z < z_c$.

5. Suppose that a researcher approximates the density of the CAD/USD exchange rate in 2020 by $x \sim N(0.74, 0.01)$ (so the mean is 0.74 and the standard deviation is 0.01, measured in US dollars).

- (a) Find the probability of observing an exchange rate less than 0.75.
- (b) Find the probability of observing an exchange rate between 0.74 and 0.76.
- (c) Find a value m such that $0.74 + m > x > 0.74 - m$ 95% of the time.
- (d) Find a value m such that $0.74 + m > x > 0.74 - m$ 99% of the time.

6. Suppose that investing in the bond market has a return with a mean of 1% and a standard deviation of 0.1%. Investing in the stock market has a return with a mean of 4% and a standard deviation of 2.1%.

- (a) What is the coefficient of variation of each return?
- (b) If the stock return is normally distributed then what is the probability that it is greater than the average bond return?

7. A survey of 1200 households classifies them according to:

- (a) whether they contain a minimum-wage earner or not and
- (b) whether the household is below Statistics Canada's low income cutoff (LICO) or not.

Here are the results:

	Minimum Wage	Not Minimum Wage
Below LICO	75	260
Above LICO	235	630

- (a) Find both marginal distributions.
- (b) Find the distribution of minimum-wage status first conditional on being below the LICO and then separately conditional on being above the LICO.
- (c) Is there a strong relationship between the two categorical variables? Can you think of an omitted variable/category that might be useful in understanding the relationship?

8. From FRED, <https://fred.stlouisfed.org/>, search for ‘Canada unemployment’ and download the unemployment rate for all persons 15 and over, monthly, seasonally adjusted from 2000M1 to 2020M10. Then search for ‘Canada CPI’, find the all-items index, and download the CPI growth rate, from the same period previous year, monthly, for the same time period. (Find that from the orange ‘edit graph’ button then ‘percent change from a year ago’.)

- (a) Find the mean unemployment rate and the mean inflation rate for this sample period.
- (b) Find the correlation between the unemployment rate and the inflation rate.
- (c) Draw and label a time series plot of the unemployment rate. Make sure your chart has a title, axis labels, and a source note. Format the scale of the vertical axis appropriately.
- (d) Draw and label a scatterplot of the two variables, with the unemployment rate on the horizontal axis and the inflation rate on the vertical axis. Make sure your chart has a title, axis labels, and a source note. Format the scale of the axes appropriately.
- (e) Use your findings to briefly discuss whether there is a relationship between the two variables (*e.g.* comment on the sign, strength, and on the role of outliers). What is the traditional name for this scatter plot in macroeconomics?

Economics 250 Exercise 1 Answer Guide

February 2021

1. (a: 3 marks) Let us order the data: -50, -1.5, -1, 2.1, 2.8, 2.8, 3, 3.3 Then you can see that the mode is 2.8. The median is $(2.1+2.8)/2 = 2.45$. The mean is 0.8125.

(b: 2 marks) It is skewed left. You can see this from the fact that the mean is less than the median.

(c: 2 marks) The IQR is (-1.25, 2.9) for a range of 4.15.

(d: 3 marks) The sample variance is $s^2 = 8.9925$ and the standard deviation is 2.998 (with some rounding).

2. (a: 2 marks) The mode is 76 and the median also is 76.

(b: 6 marks) The sample mean is

$$\bar{x} = 0.4(69) + 0.5(76) + 0.3(84) = 77$$

The sample variance is:

$$\frac{1}{19}[4(69 - 77)^2 + 10(76 - 77)^2 + 6(84 - 77)^2] = 29.97$$

so the sample standard deviation is $s = 5.43$.

3. (a: 2 marks) $4/10$ or 0.40 or 40%

(b: 2 marks) This is zero because the distribution is continuous.

(c: 2 marks) From *uniform.pdf* the variance is $(15-5)^2/12 = 8.33$ so the standard deviation is 2.89 (depending on how you round) minutes. (Notice by the way that the standard deviation is in the same units as the waiting time itself.)

4. (a: 2 marks) The probability is $1-0.8413 = 0.1587$ or 15.87%.

(b: 2 marks) This probability is 0.025 or 2.5%. (The equality sign makes no difference.)

(c: 2 marks) This probability is $0.95-0.5 = 0.45$ or 45%.

(d: 2 marks) This value is 0.44.

5. (a: 2 marks) This gives $z = 1.0$ so the probability is 0.8413 or 84.13%.
- (b: 2 marks) This gives z between 0 and 2 so the probability is $0.9772 - 0.5 = 0.4772$.
- (c: 2 marks) Here $m = 1.96 \times 0.01 = 0.0196$.
- (d: 2 marks) Here $m = 2.575 \times 0.01 = 0.02575$.

6. (a: 2 marks) For bonds $CV = (0.1/1)(100) = 10\%$. For stocks $CV = (2.1/4)100 = 52.5\%$. Thus the stock return is more variable (even relative to its higher mean) than the bond return.

(b: 3 marks) To answer this question we standardize:

$$z = \frac{1 - 4}{2.1} = -1.42857.$$

I'll use -1.43 from Table A so the probability of a value above this level is $1 - 0.0764 = 0.9236$. (You will get a slightly smaller value if you use your calculator or excel to find the exact value.)

7. (a: 3 marks) The LICO marginal distribution is 0.279 below LICO and 0.721 above it. The minimum wage status marginal distribution is 0.258 with minimum wage and 0.742 without it.

(b: 3 marks) Conditional on being below LICO the distribution is: 0.224 at minimum wage and 0.776 not at minimum wage (with some rounding). Conditional on being above LICO the corresponding values are: 0.263 (rounding) and 0.737.

(c: 4 marks) Thus there is not a strong relationship between these two categorical variables, because the two conditional distributions are similar. An omitted variable or category that might affect the relationship is whether or not there are workers age 15–24 in the household. Often, those workers receive minimum wage even if they are not in LICO households. (You may have other good ideas for a lurking variable too.)

8. (a: 5 marks) The average unemployment rate 7.0% while the average inflation rate was 1.73%.

(b: 3 marks) The correlation coefficient is -0.276.

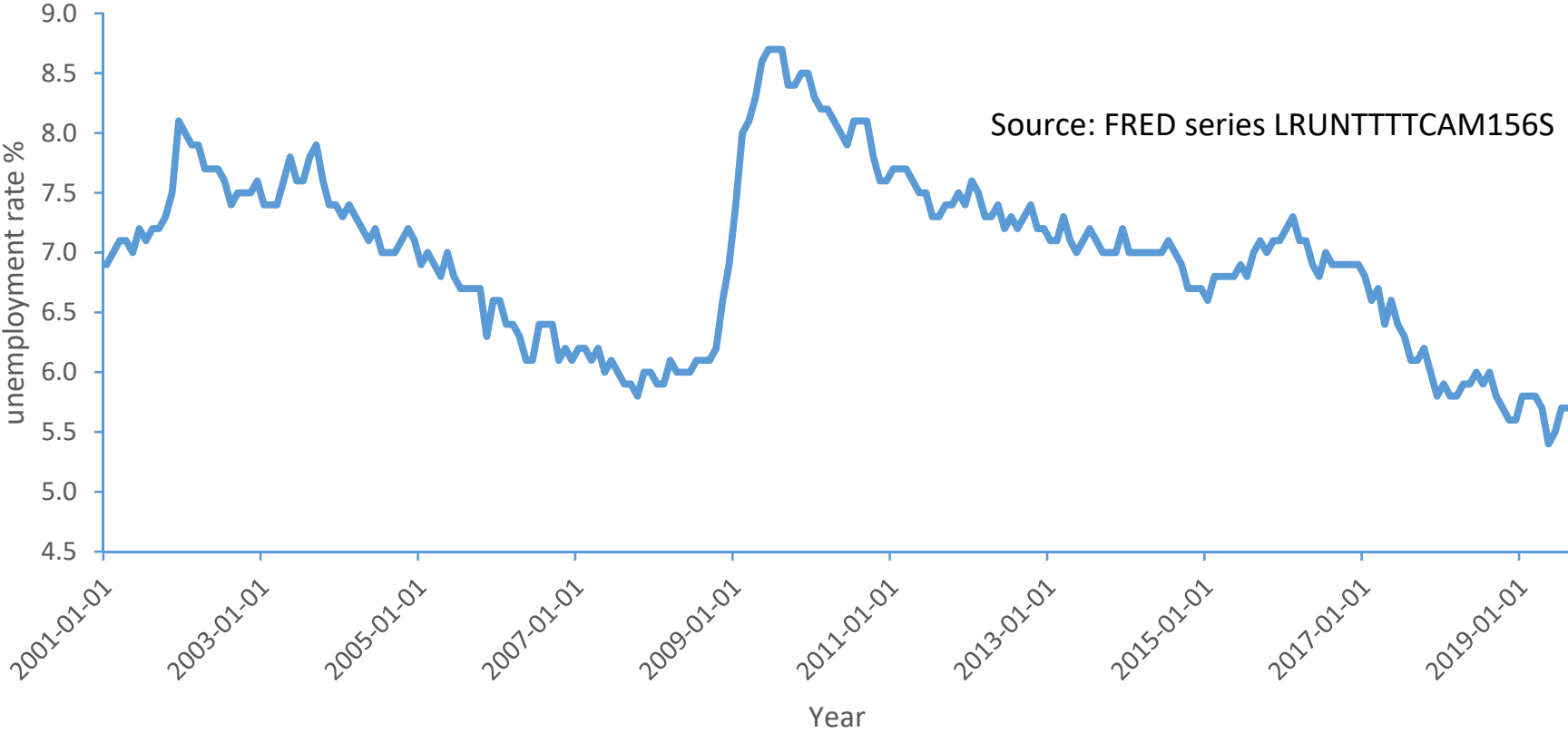
(c: 6 marks) [Please see the figure attached.]

(d: 6 marks) [Please see the figure attached.]

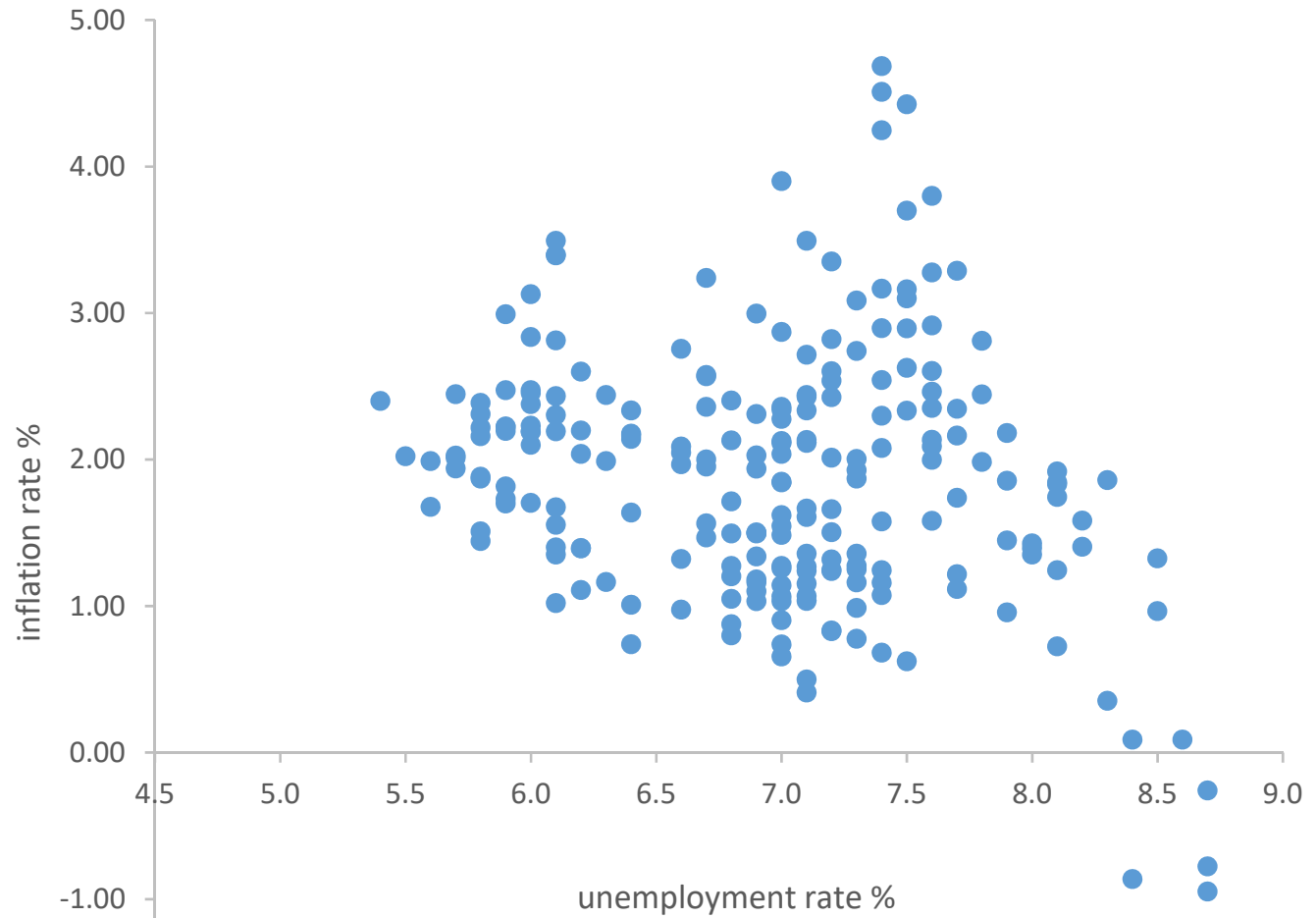
(e: 3 marks) The relationship is negative but weak, as shown by the correlation coefficient (and evident in the scatter plot). The result does not seem sensitive to outliers. This relationship traditionally is called a Phillips curve.

[Note: We discovered that two places in FRED give Canadian inflation rates that mysteriously differ. So you may have found a slightly different scatter plot and an even weaker negative correlation. Answers correctly using either series received full marks.]

Canadian Unemployment Rate



Canadian Unemployment and Inflation Rates



Source: FRED series LRUNTTTTCAM156S and CPALTT01CAM659N