Economics 250
Introductory Statistics
Exercise 1
Due Tuesday 28 January 2025

Instructions: 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. Write in pen (or type) and not in pencil. Plots and graphs must be drawn with software (such as Excel) not by hand. Work can be submitted on paper in class or via OnQ as a single file.

1. Imagine that a portfolio's investment returns in eight successive years take these values (in percentage points): -10.0, -2.0, 2.0, -3.0, 6.7, 4.8, 4.8, 6.2.

(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 6 scores of 60%, 10 scores of 72%, and 4 scores of 80%.

(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 medical office are continuously and uniformly distributed between 4 minutes and 16 minutes.

(a) What is the probability of waiting less than 8 minutes?

(b) What is the probability of waiting exactly 6 minutes?

(c) What is the standard deviation of the waiting time?

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

(a) Find the probability that $z > 2$.

(b) Find the probability that $z \leq -1.645$.

(c) Find the probability that $0 > z > -1$.

(d) Find the value z_c such that there is a 95% probability that $-z_c < z < z_c$.

5. Suppose that a researcher approximates the density of the CAD/USD exchange rate in January 2018 by $x \sim N(0.80, 0.02)$ (so the mean is 0.80 and the standard deviation is 0.02, measured in US dollars).

(a) Find the probability of observing an exchange rate less than 0.81.

(b) Find the probability of observing an exchange rate between 0.80 and 0.826.

(c) Find a value m such that $0.80 + m > x > 0.80 - m$ 95% of the time.

(d) Find a value m such that $0.80 + m > x > 0.80 - m$ 99% of the time.

6. Suppose that investing in the bond market has a return with a mean of 0.8% and a standard deviation of 0.2%. Investing in the stock market has a return with a mean of 3.0% 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 400 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	16	90
Above LICO	42	252

(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' (series LRUNTTTTCAM156S) and download (as an excel file) the unemployment rate for all persons 15 and over, monthly, seasonally adjusted from 2001:1 to 2024:10. Then search for the Canadian CPI growth rate (series CPALTT01CAM659N) (it is from the same period previous year, monthly) for the same time span.

(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
January 2025

1. (a: 3 marks) Let us order the data: -10.0, -3.0, -2.0, 2.0, 4.8, 4.8, 6.2, 6.7 Then you can see that the mode is 4.8. The median is (2.0+4.8)/2 = 3.4. The mean is 1.1875.

(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 (-2.5, 5.5).

(d: 3 marks) The sample variance is $s^2 = 33.59$ and the standard deviation is 5.796.

2. (a: 2 marks) The mode is 72 and the median also is 72.

(b: 6 marks) The sample mean is

$$\overline{x} = 0.3(60) + 0.5(72) + 0.2(80) = 70$$

The sample variance is:

$$\frac{1}{19}[6(60-70)^2 + 10(72-70)^2 + 4(80-70)^2)] = 54.73$$

so the sample standard deviation is $s = 7.39$.

3. (a: 2 marks) 4/12 or 0.33 or 33% (rounding to the nearest percentage point)

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

(c: 2 marks) From *uniform.pdf* the variance is 12 so the standard deviation is 3.46 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 1-0.9772= 0.0228 or 2.28%.

(b: 2 marks) This probability is 5.0% (the equality sign makes no difference). (I averaged the values in Table A. You may get a slightly different value if you use your calculator or an on-line version: 4.9985%).

(c: 2 marks) This probability is 0.5 - .1587 = 0.3413 or 34.13%.

(d: 2 marks) This value is 1.96.

5. (a: 2 marks) This gives $z = 0.5$ so the probability is 0.6915 or 69.15%.

(b: 2 marks) This gives z between 0 and 1.3 so the probability is 0.9032 - 0.5 = 0.4032.

(c: 2 marks) Here $m = 1.96 \times 0.02 = 0.0392$.

(d: 2 marks) Here $m = 2.58 \times 0.02 = 0.0516$.

6. (a: 2 marks) For bonds CV = (0.2/0.8)(100) = 25%. For stocks CV = (2.1/3.0)100 = 70%. Thus the stock return is much more variable, relative to its mean, than the bond return.

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

$$z = \frac{0.8 - 3.0}{2.1} = -1.048.$$

I round this to -1.05 and use Table A, so the probability of a value above this level is 1-0.1469 = 0.8531.

7. (a: 3 marks) The LICO marginal distribution is 0.265 below LICO and 0.735 above it. The minimum wage status marginal distribution is 0.145 with minimum wage and 0.855 without it.

(b: 3 marks) Conditional on being below LICO the distribution is: 0.151 at minimum wage and 0.849 not at minimum wage. Conditional on being above LICO the corresponding values are: 0.143 (rounding) and 0.857.

(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 2.21%.

(b: 3 marks) The correlation coefficient is -0.53.

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

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

(e: 3 marks) The relationship is negative and of intermediate strength, as shown by the correlation coefficient (and evident in the scatter plot). There are five outliers: the unemployment rates greater than 10% from April to August 2020. Notice that these are outliers in the scatter plot because they are not associated with deflation (negative inflation rates). The negative correlation would be stronger without those observations. This relationship traditionally is called a Phillips curve.