Economics 250
Introductory Statistics
Exercise 2
Due Wednesday 3 March 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. Note: Please submit text in a single .pdf file and not a .doc or multiple photo files. Marks will be deducted otherwise.

1. Suppose at a track-and-field competition 10% of athletes are cheating (C) by taking banned substances. Everyone is tested, and 8% of the athletes test positive (Pos) for those substances. The test sometimes yields 'false positive' results though: Given that someone is not cheating, the test gives a positive result 5% of the time.

(a) What is $P(Not\, C \cap Pos)$?

(b) What is $P(Pos|C)$?

(c) What is $P(C|Pos)$?

(d) Are the events testing positive (Pos) and cheating (C) independent?

2. Suppose that we describe the return on the stock market as a discrete random variable, labelled r_s. Suppose that the possible outcomes are $\{-3, 6, 12\}$ with probabilities $\{0.2, 0.5, 0.3\}$.

(a) Find the expected value of r_s.

(b) Find the standard deviation of r_s.

(c) An investor considers combining this investment with a bond return labelled r_b that has mean 2 and variance 0. The new portfolio has equal amounts in each investment so its return is:

$$r_p = 0.5r_s + 0.5r_b.$$

Find the mean and standard deviation of r_p.

3. Statistics are central to managing an investment portfolio. Imagine that an investor is choosing how to divide each dollar between two investments, labelled A and B. Investment A has a return with mean 4 and standard deviation 1. Investment B has a return with mean 6 and standard deviation 1. The correlation between the two returns is 0.4. The portfolio invests a fraction ω in investment A and a fraction $1-\omega$ in investment B so that the portfolio's return is:

$$r_p = \omega r_A + (1-\omega) r_B.$$

Suppose that ω lies in the range $[0,1]$.

(a) Find formulas for the mean of the portfolio return, labelled μ_p, and the variance of the portfolio return, labelled σ_p^2, as functions of ω. What value of ω maximizes the mean return on the portfolio?

(b) In finance, risk somestimes is measured using the variance. Find the value of ω that minimizes the variance of r_p. (Hint: Use your mathematics skills to minimize the function you derived.)

4. Suppose that a financial asset pays a dividend of \$2 sometimes and \$4 at other times. Let us study the dividends in two successive years, denoted d_1 and d_2. These two discrete random variables are jointly distributed like this:

		d_1	
		2	4
d_2	2	0.4	0.1
	4	0.1	0.4

(a) Find the marginal distribution for d_2. Then find the mean and standard deviation using that distribution. (Sometimes these are called the unconditional mean and standard deviation.)

(b) Suppose that you observe that $d_1 = 4$. Find the conditional distribution for d_2. Then find the conditional mean and standard deviation.

(c) Are the two random variables independent? Find the covariance and the correlation.

5. Suppose that medical researchers know that a treatment has a 40 percent chance of being successful for any patient.

(a) If they treat 8 patients what is the probability that the sample success rate is 50%? What is the probability that the sample success rate is 50% or more?

(b) If they treat 100 patients what is the probability that the sample success rate is 50% or more?

6. Suppose that an after-school tutoring program succeeds in raising any student's mathematics test score with a probability of 0.7.

(a) If 10 students participate in the program then what is the probability that at least 6 of them (*i.e.* 6 or more) experience a test score increase?

(b) What is the probability that only 1, 2 or 3 of them experience a test score increase?

(c) If 100 students participate then what is the probability that at least 60 of them experience a test score increase? (Hint: Think of this sample size as large and do not use the binomial formula.)

(d) For large n what is the distribution of the success rate or proportion, defined as $\hat{p} = x/n$?

7. Use the data set *Fisher2021* from the Projects web page for this course. For years from 1962 to 2020 it shows Canadian 10-year bond interest rates and CPI inflation rates, both in percentage points for each month.

(a) Use Excel to produce and properly label a scatter plot of the 10-year interest rate (on the vertical axis) plotted against the inflation rate (on the horizontal axis). Also add a linear trendline to your graph.

(b) Run a linear regression like this:

$$y = a + bx,$$

In Excel, go to the 'data' tab then to 'data analysis" and 'regression'. Report the coefficient on the x-variable (*i.e.* the inflation rate) and its standard error. Also report the R-square (R^2) statistic.

(c) Sections 2.3–2.4 of the textbook tell us how to interpret these statistics. Based on these statistics is there much evidence of a relationship between this interest rate rate and the inflation rate, in these time series data?

(d) If the inflation rate is 0.5% then what will the predicted interest rate rate be, given this historical relationship?

Economics 250 Exercise 2 Answer Guide
March 2021

1. (a: 2 marks) Clearly $P(NotC) = 0.90$. The multiplication rule gives:

$$P(NotC \cap Pos) = P(Pos|NotC)P(NotC) = 0.05 \times 0.9 = 0.045$$

(b: 2 marks) Completing the two-way table using the multiplication rule:

$$P(Pos|C) = \frac{P(Pos \cap C)}{P(C)} = 0.035/0.10 = 0.35.$$

(c: 2 marks) From the table or from Bayes rule:

$$P(C|Pos) = \frac{P(Pos|C)P(C)}{P(Pos)} = \frac{0.35 \times 0.10}{0.08} = 0.4375.$$

(d: 2 marks) No they are not independent. $P(C|Pos) = 0.4375 > P(C) = 0.10$.

2. (a: 2 marks)

$$E(s) = 0.2 \times (-3) + 0.5 \times 6 + 0.3 \times 12 = -0.6 + 3.0 + 3.6 = 6.$$

(b: 2 marks) The variance is:

$$\sigma_s^2 = 0.2(-9)^2 + 0.5(0)^2 + 0.3(6)^2 = 16.2 + 0 + 10.8 = 27$$

so the standard deviation is:

$$\sigma_s = 5.196$$

(c: 2 marks) Here r_b is just a constant, so

$$r_p = 1 + 0.5r_b.$$

Thus the mean is $1 + 0.5(6) = 4$. The standard deviation is $0.5(5.196) = 2.598$.

3. (a: 3 marks) The mean is:

$$\mu_p = \omega(4) + (1-\omega)6 = 4\omega + 6 - 6\omega = 6 - 2\omega.$$

The variance is:

$$\sigma_p^2 = \omega^2(1^2) + (1-\omega)^2(1^2) + 2\omega(1-\omega)(0.4)(1)(1) = 1.2\omega^2 + 1 - 1.2\omega.$$

You can see that setting $\omega = 0$—investing only in asset B—maximizes the mean return on the portfolio.

(b: 3 marks) Now

$$\sigma_p^2 = 1.2\omega^2 + 1 - 1.2\omega$$

Differentiate with respect to ω and set the derivative to zero:

$$\frac{d\sigma_p^2}{d\omega} = 1.2(2\omega - 1) = 0,$$

which gives

$$\omega = 0.5.$$

Even though the two investments have the same variance, the portfolio variance can be reduced by exactly diversifying as long as the correlation is not 1. (Incidentally, the second derivative is positive, so this is indeed a minimum.)

4. (a: 2 marks) The marginal distribution for d_2 has probability 0.5 on the value 2 and 0.5 on the value 4. Thus the mean is 3. The variance is 1. The standard deviation is 1.

(b: 3 marks) The conditional distribution of d_2, given $d_1 = 4$, has weight 0.2 on the value 2 and 0.8 on the value 4. Thus the conditional mean is 3.6. The conditional variance uses the conditional mean and the conditional probabilities, which gives 0.64 so the standard deviation is 0.8. (Notice, by the way, that the new information leads the researcher to revise up the mean and revise down the standard deviation.)

(c: 4 marks) The two random variables are not independent. Our work in parts (a) and (b) shows that the conditional probabilities differ from the marginal ones. Another way to show they are not independent is to find the correlation. The covariance is 0.6 so the correlation is also 0.6.

5. (a: 2 marks) Here $n = 8$, $p = 0.4$ and so from Table C or the binomial formula for $k = 4$ the probability is 0.2322 or 23.22%. The probability that it succeeds for 4 or more is found by adding up the values for $k = 4, ..., 8$ which gives 0.406 or 40.6%.

(b: 2 marks) With $n = 100$ we use the normal approximation that $X \sim N(np, \sqrt{np(1-p)})$ so $X \sim N(40, 4.8989)$. Thus $\hat{p} \sim N(p, \sqrt{p(1-p)/n})$ so $\hat{p} \sim N(0.40, 0.0489)$. Standardizing gives $z = (0.50 - 0.40)/0.0489 = 2.044$ so from table A the probability is 1-0.9793 = 0.0207 or 2.07%. (In table A I used the value 2.04 but you also could use your calculator to find the exact value or you could the average of the values for 2.04 and 2.05.)

6. (a: 3 marks) This is a binomial problem with $n = 10$ and $p = 0.7$. (We can use Table C remembering that the probability of failure is thus 0.3 and we want to count 4 or fewer failures.) From Table C the answer is 0.8497.

(b: 3 marks) Reading up the column in Table C gives 0.0105. (The wording might be ambiguous, so some people might write the individual values but not sum them.)

(c: 3 marks) We use the normal approximation, so that

$$x \sim N(70, 4.58).$$

That means that:

$$z = \frac{60 - 70}{4.58} = -2.18.$$

So from Table A the probability above this value is 1-0.0146=0.9854 or 98.54%.

(d: 3 marks) The distribution for large n is:

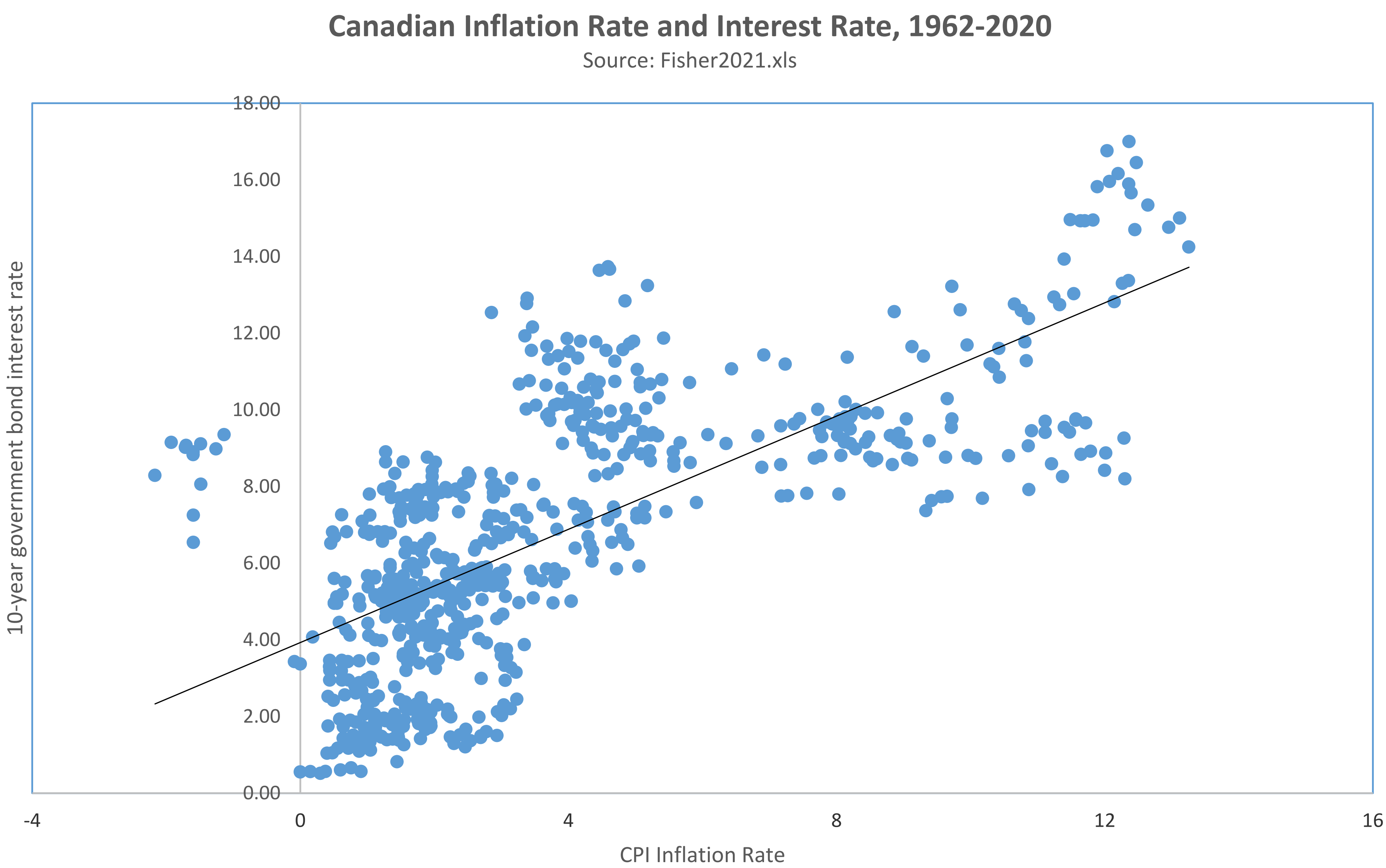
$$\hat{p} \sim N(0.70, \frac{0.458}{\sqrt{n}}).$$

7. (a: 5 marks) Please see the attached figure.

(b: 5 marks) The slope is 0.738 with a standard error of 0.028. The R^2 statistic is 0.493.

(c: 3 marks) Yes, there is a moderately strong, positive relationship, as seen from the value of R^2.

(c: 4 marks) The estimated intercept is 3.932 so the predicted value of the interest rate is 3.932 + 0.738 (0.5) = 4.301%.



Canadian Inflation Rate and Interest Rate, 1962-2020
Source: Fisher2021.xls
10-year government bond interest rate
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
-4
0
4
8
12
16
CPI Inflation Rate