FACULTY OF ECONOMICS STUDY AIDS 2021

ECT1 Paper 3 Quantitative Methods in Economics

The Faculty Board has agreed to release outline solutions to the 2021 examinations as a study aid for exam revision. They are abridged solutions, and not 'definitive', and should therefore not be considered as an exemplar for 'complete' answers.

Note that the Faculty will not respond to any queries regarding these solutions.

Paper 3

QUANTITATIVE METHODS IN ECONOMICS

Section A

1 Give an example of:

(a) A function $f(x)$ which is differentiable for all $x \in \mathbb{R}$, but not twice differentiable for some $x \in \mathbb{R}$.

(b) A function $g(x)$ which is continuous and differentiable for all $x \in \mathbb{R}$, strictly concave for $x \leq 0$, and simultaneously concave and convex for $x > 0$.

Solution to 1. For example:

(a)

$$f(x) = \begin{cases} -x^2 & \text{for } x \leq 0 \\ 0 & \text{for } x > 0 \end{cases}$$

This function is obviously differentiable for $x > 0$ and for $x < 0$. Further,

$$\lim_{h \to 0} \frac{f(h) - f(0)}{h} = 0$$

and so $f(x)$ is also differentiable at $x = 0$. Indeed,

$$\frac{f(h) - f(0)}{h} = \begin{cases} -h & \text{for } h < 0 \\ 0 & \text{for } h > 0 \end{cases}$$

so the limits from left and from right exist and coinside. The derivative of $f(x)$ is not differentiable at $x = 0$. Indeed,

$$f'(x) = \begin{cases} -2x & \text{for } x \leq 0 \\ 0 & \text{for } x > 0 \end{cases} \quad \text{and } f''(x) = \begin{cases} -2 & \text{for } x < 0 \\ 0 & \text{for } x > 0 \end{cases}$$

But

$$\lim_{x \to -0} f''(x) = -2 \neq 0 = \lim_{x \to +0} f'(x).$$

(b) We can take $g(x) = f(x)$, where $f(x)$ is defined in (a). We have seen that $g(x)$ is continuous and even differentiable everywhere. It is strictly concave for $x < 0$ because $g''(x) = -2 < 0$ for $x < 0$ and it is both concave and convex for $x > 0$ because it is linear for $x > 0$. Very good students may point out that in the course, we only formulated theorem which says that if a function has non-negative second derivative on (a, b), then it is concave on (a, b). So, using that theorem, we only can show the concavity of $g(x)$ on (a, b) for any $a, b \leq 0$. They may point out that this is equivalent to the concavity for $x < 0$. Further, if a function is concave but not strictly concave, it must have a linear part of the graph, which x^2 does not have. So $g(x)$ is strictly concave for $x < 0$. Of course, examples will vary.

2. Answer the following questions:

(a) Consider the function

$$g(x) = \int_{1/x}^{x} \ln u \mathrm{d}u$$

with $x > 1$. For which values of $a > 1$ is this function increasing over $(1, a)$? For which values of $a > 1$ is this function concave over $(1, a)$? Justify your answers.

(b) Find the value of $\int_0^{100} f(x)\mathrm{d}x$, where

$$f(x) = \begin{cases} 2^{-n}(x - 2n) & \text{for } x \in [2n, 2n+1] \\ -2^{-n}(x - 2n - 2) & \text{for } x \in [2n+1, 2n+2] \end{cases}$$

with $n = 0, 1, 2, ...$ Show your work!

Solution to 2. (a) Using integration by parts, we obtain

$$\begin{aligned} g(x) &= \int_{1/x}^{x} \left(\frac{\mathrm{d}}{\mathrm{d}u}u\right) \ln u \mathrm{d}u = u \ln u|_{1/x}^{x} - \int_{1/x}^{x} u \frac{1}{u} \mathrm{d}u \\ &= x \ln x + \frac{1}{x} \ln x - x + 1/x. \end{aligned}$$

Differentiating, we obtain

$$\begin{aligned} g'(x) &= \ln x + 1 - \frac{1}{x^2} \ln x + \frac{1}{x^2} - 1 - \frac{1}{x^2} \\ &= \left(1 - \frac{1}{x^2}\right) \ln x. \end{aligned}$$

Since $g'(x) > 0$ for any $x > 1$, $g(x)$ is increasing over $(1, a)$ for any $a > 1$. Further, computing the second derivative, we obtain

$$g''(x) = \frac{1}{x^3} \ln x + \left(1 - \frac{1}{x^2}\right) \frac{1}{x} > 0$$

for all $x > 1$. Hence, $g(x)$ is strictly convex for $(1, a)$, and thus, for no $a > 1$ it is concave over $(1, a)$.

(b) Note that function $f(x)$ is piece-wise linear and continuous. It equals zero for any even integer, and 2^{-n} for $x = 2n + 1$. It is linear in between even and odd integer points. By the additivity of the definite integral of continuous functions, we have

$$\int_0^{100} f(x)\mathrm{d}x = \sum\nolimits_{n=0}^{49} \int_{2n}^{2n+2} f(x) = \sum\nolimits_{n=0}^{49} \frac{1}{2^n}$$

Using the formula for the sum of geometric progression, we obtain

$$\sum\nolimits_{n=0}^{49} \frac{1}{2^n} = \frac{1 - 2^{-50}}{1 - 2^{-1}} = 2\left(1 - 2^{-50}\right).$$

3. Suppose that $f(\mathbf{x})$ is a strictly concave continuous function defined on a compact convex subset D of $\mathbb{R}^2$.

(a) Show that $f(\mathbf{x})$ does achieve its maximum on D.

(b) Suppose further that D_0 is a compact but not necessarily convex subset of D. Give an example of $f(\mathbf{x})$, D and D_0 that satisfy the above requirements and are such that the maximum of $f(\mathbf{x})$ on D_0 is achieved at two different points.

Solution to 3. (a) By Weierstrass theorem, any continuous function over a compact set achieves both its maximum and its minimum.

(b) Answers may vary. Suppose $\mathbf{x} = (x_1, x_2)$. One possible example is $f(\mathbf{x}) = -x_1^2 - x_2^2$, $D = \left\{\mathbf{x} : x_1^2 + x_2^2 \leq 1\right\}$, and $D_0 = \{(1,0), (-1,0)\}$ so that D_0 consists only of two points where the values of $f(\mathbf{x})$ are obviously the same. These values, of course, are both the maximum and the minimum values of $f(\mathbf{x})$ over D_0.

4. Consider a 2×2 matrix $A = \begin{pmatrix} x & y \\ y & x \end{pmatrix}$

(a) Describe the set of pairs $(x, y) \in \mathbb{R}^2$ such that $x \in [-1, 1]$, $y \in [-1, 1]$, and the corresponding A is positive definite. Is this set convex? Is this set compact? Justify your answers.

(b) Find the maximum value of the determinant of A if it is known that $x \in [-1, 1]$, $y \in [-1, 1]$, that A is positive definite, and $2x = y + 1$. Explain why you think this is indeed the maximum.

Solution to 4. (a) For positive definiteness, we must have $A_{11} > 0$ and $A_{11}A_{22} - A_{12}^2 > 0$. That is,

$$x > 0 \text{ and } x^2 - y^2 > 0.$$

Together with the constraints $x \in [-1, 1]$, $y \in [-1, 1]$, these two inequalities correspond to a triangle in the x, y-plane with vertices at $(0, 0)$, $(1, 1)$, and $(1, -1)$. The sides of this triangle that connect $(0, 0)$ with $(1, 1)$ and with $(1, -1)$ are excluded from this set. Hence, although the set is convex, it is not compact.

(b) The determinant of A equals $x^2 - y^2$. Substituting $y = 2x - 1$, we get

$$\det A = x^2 - 4x^2 + 4x - 1 = -3x^2 + 4x - 1,$$

which is maximized at $-6x + 4 = 0$, so that $x = 2/3$ and $y = 4/3 - 1 = 1/3$. Since $(2/3, 1/3)$ is an internal point of the set satisfying the inequality constraints and the requirement of positive definiteness, this is indeed the constrained maximum point. The value of $\det A$ at this point is $(2/3)^2 - (1/3)^2 = 1/3$. This problem also can be solved using the Lagrange multiplier method. However, one still would need some extra argument about why the found solution candidate is indeed the maximum. The standard method that checks the concavity of the Lagrangian would not work directly.

Section B.

5. Suppose that a consumer has utility function $U(x, y) = x^{\alpha} + y^{\alpha}$ with constant $\alpha \in (0, 1)$, where $x \geq 0$ and $y \geq 0$ are quantities of two goods available for consumption.
   (a) Using the Lagrange multiplier method, find the bundle (x^*, y^*) that maximises the utility subject to the budget constraint $px + y = 1$, where $p > 0$ is the relative price of good x.
   (b) Show that your answer in (a) indeed delivers the maximum utility as opposed to, say, the minimum.
   (c) Express the value of the Lagrange multiplier at the maximum as a function of α and p. Interpret this value and comment on how it depends on p in light of this interpretation.
   (d) Let $U^*(p, \alpha) = U(x^*, y^*)$ denote the indirect utility function. Find $\frac{\partial}{\partial p} U^*(p, \alpha)$ as a function of p and α. How does this quantity behave as $\alpha \to 0$?

Solution to 5. (a) The Lagrangian is

$$L(x, y, \lambda) = x^{\alpha} + y^{\alpha} - \lambda(px + y - 1)$$

The FOC are

$$\begin{aligned} \alpha x^{\alpha-1} - p\lambda &= 0 \\ \alpha y^{\alpha-1} - \lambda &= 0 \\ px + y - 1 &= 0 \end{aligned}$$

Hence, $\lambda = \alpha y^{\alpha-1}$ and $\alpha x^{\alpha-1} - p\alpha y^{\alpha-1} = 0$, which implies that, either $x = y = 0$ (which would not satisfy the third FOC equation), or

$$x = p^{1/(\alpha-1)} y$$

Using this in the third FOC equation, we obtain

$$x^* = \frac{p^{\frac{1}{\alpha-1}}}{p^{\frac{\alpha}{\alpha-1}} + 1}, y^* = \frac{1}{p^{\frac{\alpha}{\alpha-1}} + 1}.$$

(b) The bundle found in (a) delivers the maximum utility because the Lagrangian is strictly concave everywhere. Indeed, its Hessian is

$$\begin{pmatrix} \alpha(\alpha-1)x^{\alpha-2} & 0 \\ 0 & \alpha(\alpha-1)y^{\alpha-2} \end{pmatrix},$$

which is a negative definite matrix for $x > 0$ and $y > 0$.

(c) Since $\lambda^* = \alpha y^{*\alpha-1}$ and $y^* = \frac{1}{p^{\frac{\alpha}{\alpha-1}}+1}$, we have

$$\lambda^* = \frac{\alpha}{\left(p^{\frac{\alpha}{\alpha-1}} + 1\right)^{\alpha-1}}$$

This is the shadow price of the constraint. That is, the marginal increase in the maximum utility when income spent on the goods is increased. Since $\alpha \in (0, 1)$, $p^{\frac{\alpha}{\alpha-1}} + 1$ is decreasing with p. Hence, λ^* is decreasing with p as well, which makes sense because with higher price, a marginal increase in spending will buy less of the 'extra' goods.

(d) By the envelope theorem for the constrained optimisation case, $\frac{\partial}{\partial p}U^*(p,\alpha)$ equals the partial derivative of the Lagrangian with respect to p at the optimum. That is

$$\frac{\partial}{\partial p}U^*(p,\alpha) = -\lambda^* x^* = -\frac{\alpha p^{\frac{1}{\alpha-1}}}{\left(p^{\frac{\alpha}{\alpha-1}}+1\right)^{\alpha}}.$$

For any fixed $p>0$, as $\alpha \to 0$, the numerator of the latter ratio converges to zero whereas the denominator converges to 1. Hence, $\frac{\partial}{\partial p}U^*(p,\alpha) \to 0$. The maximum utility becomes less sensitive to the change in the relative price of the good x.

6. A country's population at time t equals $P(t)$ million people. This function satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) = bP(t) + I(t),$$

where b is a positive constant, t is measured in years, and $I(t)$ equals the rate of immigration to the country from the rest of the world.

(a) Suppose that the rate of immigration is constant $I(t)=a$. Then all solutions to the above differential equation have the form $P(t)=Ae^{bt}-a/b$, where A is some constant. Explain why this does not imply that the higher immigration rate $a_1>a$ would lead to a lower population size $P(t)$.

(b) Let $a=1, b=0.005$, and $P(2000)=50$. Find the value of $P(2020)$.

(c) You would like to approximate $P(t)$ by its Taylor polynomial of second degree around $t=2000$. Suppose that all assumptions made in (a) and (b) hold. Using Taylor's remainder theorem, estimate an upper bound on the maximum of the absolute value of the approximation error for $t\in[2000,2020]$.

(d) Now suppose that $b=0.005$ and $P(2000)=50$, as above, but that the immigration was stopped at 2010 so that $I(t)=1$ for $t\le 2010$ but $I(t)=0$ for $t>2010$. Find the value of $P(2020)$. Show your work!

Solution to 6. (a) Higher a does not lead to lower $P(t)$ because, given an initial population size, the constant A will depend on a. Hence, to evaluate the effect of the change in a on $P(t)$, this dependence should be taken into account.

(b) Substituting the values of the constants into the general form of the solution, we obtain

$$P(t) = Ae^{0.005t} - 200.$$

Using $P(2000)=50$, we get

$$50 = Ae^{10} - 200$$

so that $A = 250e^{-10}$. Therefore,

$$P(2020) = 250e^{0.1} - 200 \approx 75$$

(c) Recall that $P(t) = Ae^{0.005t} - 200$. The second order Taylor's polynomial of $Ae^{0.005t}$ around $t=2000$ is

$$Ae^{10}\left(1+0.005(t-2000)+\frac{(0.005)^2}{2}(t-2000)^2\right).$$

Therefore, the second order Taylor's polynomial for $P(t)$ around $t = 2000$ is

$$Ae^{10} - 200 + Ae^{10} 0.005\,(t - 2000) + \frac{Ae^{10}\,(0.005)^2}{2}\,(t - 2000)^2\,.$$

By Taylor's remainder theorem, the difference between this approximation and $P(t)$ at $t \in [2000, 2020]$ equals

$$\frac{A\,(0.005)^3\,e^{0.005s}}{3!}\,(t - 2000)^3$$

for some $s \in [2000, t]$. Hence, the absolute value of the difference for any $t \in [2000, 2020]$ cannot be larger than

$$\begin{aligned}\frac{A\,(0.005)^3\,e^{0.005\times 2020}}{3!} 20^3 &= \frac{Ae^{10}\,(0.005)^3\,e^{0.1}}{3!} 20^3 \\ &= \frac{250\,(0.005)^3\,e^{0.1}}{3!} 20^3,\end{aligned}$$

here we used the fact that $Ae^{10} = 250$. Evaluating this explicitly, we get the upper bound 0.047 on the approximation error (47 thousand individuals).

(d) There are at least two approaches. One can use the general formula for solutions of non-autonomous differential equations given in class. Alternatively, one can find $P\,(2010)$ similarly to how we found $P(2020)$ in (b), and then repeat the calculation for $P(2020)$ using the initial value $P(2010)$ and the new parameter value $a = 0$. Here we take the second approach. We have

$$P\,(2010) = 250e^{0.05} - 200 \approx 63.$$

For $t \geq 2010$, we have $P(t) = \tilde{A}e^{0.005t}$. Hence,

$$\tilde{A} \approx 63e^{-10.05}.$$

Therefore,

$$P(2020) \approx 63e^{-10.05}e^{10.1} = 63e^{0.05} \approx 66$$

which is about 9 million smaller than for the case when the immigration is not stopped.

Section C

7. A large number of people, N, are subjected to a test for a disease. This can be administered in two ways: (1) each person can be tested separately, in this case N tests are required, (2) the samples of k persons can be pooled and analysed together. If this test is negative, this one test suffices for the k people. If the test is positive, each of the k persons must be tested separately, and, in all, $k+1$ tests are required for the k people. Assume that the probability p that a test for a single person is positive is the same for all people and that these events are independent.

   (a) Find the probability that the test for a pooled sample of k people will be positive.

   (b) What is the expected value of the number X of tests necessary under plan (2)? (Assume that N is divisible by k.)

   (c) Show that the value of k which will minimise the expected number of tests under the second plan when p is small is approximately $1/\sqrt{p}$.

   You may use, if necessary, the approximation $(1-p)^k \approx 1 - kp$ for small p.

   (d) For an airborne infectious disease if we pool by households the assumption of independence is unlikely to be met. What might be the consequence on the number of tests required under pooled testing?

Solution to 7. (a) The test will be negative if and only if all individual tests for the people from the pool are negative. The probability of this is $(1-p)^k$. Hence, the probability of a positive test is $1-(1-p)^k$.

(b) For a pool of k people only one test is needed with probability $(1-p)^k$, and $(k+1)$ tests are needed with probability $1-(1-p)^k$. Hence, the expected value of the number of the tests for this pool is

$$1 * (1-p)^k + (k+1)\left(1-(1-p)^k\right) = k+1-k(1-p)^k.$$

There are N/k groups of k people in the population. Therefore,

$$E(X) = \frac{N}{k}\left(k+1-k(1-p)^k\right)$$

(c) Using the approximation $(1-p)^k \approx 1-kp$, we obtain

$$E(X) = \frac{N}{k}\left(1+k^2p\right),$$

which is minimised at k that satisfies

$$-\frac{1}{k^2}+p=0.$$

That is, the minimum of $E(X)$ should be achieved for k close to $1/\sqrt{p}$.

(d) Consider an extreme case where all pooled people either have or do not have the disease together. Then, the probability of the negative pooled test is just $1-p$, and the probability of positive one is p. Hence,

$$E(X) = \frac{N}{k}\left(k+1-k(1-p)\right) < \frac{N}{k}\left(k+1-k(1-p)^k\right).$$

IMore generally, intuitively, the probability of the negative pooled test is larger when there is dependence and therefore, the probability of administering $k+1$ tests instead of just one test for the group of k people is smaller. The expected number of tests is smaller.

8. Let X and Y be independent random variables with density functions $f_X(x)$ and $f_Y(y)$ respectively. Denote their sum by $Z = X + Y$. The density of Z is then given by the formula

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\, dy$$

An exponentially distributed random variable has density function

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & elsewhere \end{cases}$$

where λ is a positive parameter.

(a) Sketch the density function of an exponentially distributed random variable X for the case $\lambda = 1$.

(b) Show that if X and Y are independent exponentially distributed random variables with the same λ then the density function of their sum Z is given by

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z \geq 0 \\ 0 & elsewhere \end{cases}$$

Sketch the density function of Z in the case $\lambda = 1$.

Hint: In part (b) pay careful attention to the limits and variable that is integrated over.

Solution to 8. (a) Straightforward sketch.

(b) We have

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\, dy$$

where $f_X(z-y) = \lambda e^{-\lambda(z-y)}$ for $z \geq y$ and zero otherwise, and $f_Y(y) = \lambda e^{-\lambda y}$ for $y \geq 0$ and zero otherwise. Hence, $f_Z(z)$ equals zero for $z < 0$, and

$$\begin{aligned} f_Z(z) &= \int_0^z \lambda e^{-\lambda(z-y)} \lambda e^{-\lambda y} dy \\ &= \int_0^z \lambda^2 e^{-\lambda z} dy = \lambda^2 z e^{-\lambda z} \end{aligned}$$

for $z \geq 0$. The sketch of the density is straightforward. In contrast to (a), the density is not decreasing for all $z \geq 0$. Instead, it first increases and then decreases.

9. A researcher has data $X_1, ..., X_N$ with $N = 50$, where X_i are independent and identically distributed random variables with unknown population mean μ and known population variance $\sigma^2 = 1600$. The researcher uses these data to run a standard one-tail test of the null hypothesis $H_0 : \mu = 170$ against the alternative $H_1 : \mu = 190$ at a significance level of $\alpha = 0.05$. Find the power of this test. State clearly any assumptions you make.

Solution to 9. The standard one-sided test is based on the normal approximation of $\bar{X} = (X_1 + ... + X_N)/N$. Specifically, under the null hypothesis, $\left(\bar{X} - 170\right)/\sqrt{\sigma^2/N}$ is approximately standard normal as $N = 50$ is relatively large. We reject the null if

$$\frac{\bar{X} - 170}{\sqrt{\sigma^2/N}} \geq 1.645$$

Using the numerical values of N and σ^2, we get an equivalent rejection criterion

$$\bar{X} \geq 179.305.$$

Therefore, under the alternative, we reject the null with probability

$$\begin{aligned} & \Pr\left(\bar{X} \geq 179.305\right) \\ = \; & \Pr\left(\frac{\bar{X} - 190}{\sqrt{\sigma^2/N}} \geq \frac{179.305 - 190}{\sqrt{\sigma^2/N}}\right) \\ \approx \; & \Pr\left(Z \geq -1.89\right) \approx 0.97. \end{aligned}$$

That is, the power of the test is approximately 97%.

10. Figure 1 summarises the results of a test for Covid-19 applied to a small sample of individuals. Predicted outcomes based on test results are cross-classified with a final diagnosis.

Final Diagnosis

Test Result	**Covid**	**Not Covid**
Covid	TP = 82	FP = 13
Not Covid	FN = 71	TN = 301

Figure 1

True positive (TP): the numbers with Covid-19 who had a *positive* test result.
False positive (FP): the numbers without Covid-19 who had a *positive* test result.
True negative (TN): the numbers without Covid-19 who had a *negative* test result.
False negative (FN): the numbers with Covid-19 who had a *negative* test result.

(a) Based on the data in Figure 1 interpret the probabilities

$$\tau_1 = TP/(TP + FN) \text{ and } \tau_2 = TN/(TN + FP)$$

Comment on the numerical values of τ_1 and τ_2.

(b) Using a large sample, the Covid-19 testing programme in the UK has been verified by Public Health England, with both τ_1 and τ_2 over 95%.

(i) What does this tell us about the extent of false positives and false negatives?

(ii) What is the link between τ_1 and τ_2 and Type I and Type II errors?

(c) An analyst computes the ratio

$$LR = \left(\frac{TP}{TP + FN}\right) / \left(\frac{FP}{FP + TN}\right)$$

(i) Use the data in Figure 1 to compute a value for LR and interpret your findings.

(ii) Use the data in Figure 1 to compute the pre-test odds ratio

$$\Pr(COVID) / (1 - \Pr(COVID))$$

In what sense does the information in LR update the pre-test odds of having Covid-19?

Solution to 10. (a) τ_1 is the conditional probability that a randomly chosen person from the small sample under study has a positive test result given that the person's final diagnosis is covid. τ_2 is the conditional probability that a randomly chosen person from the small sample under study has a negative test result given that the person's final diagnosis is no covid. The numerical values are $\tau_1 = \frac{82}{71+82} \approx 0.54$ and $\tau_2 = \frac{301}{13+301} \approx 0.96$. It looks like the test is pretty bad at detecting covid when it is in fact present. On the other hand, when there is no covid, the probability of false alarm is relatively low. When the test is positive we are pretty sure that there is covid.

(b) (i) The extent of false negatives and false positives is fairly low in this large study.

(ii) Type-I error is the error of rejecting the null hypothesis of no Covid, when it is true. Hence, $1 - \tau_2$ estimates the type-I error. Type-II error is the error of not rejecting the null of no Covid when we should. Thus, it is estimated by $1 - \tau_1$.

(c) (i) $LR = \frac{\tau_1}{1-\tau_2} = \frac{0.54}{1-0.96} = 13.5$. The ratio is interpreted as an estimate of the probability of rejecting the null of no covid when we should divided by the probability of rejecting the null of covid when we should not. Hence, it estimates how much more likely we see positive test results among those with covid than among those with no covid.

(ii) $\Pr(COVID) = \frac{TP+FN}{TP+FN+FP+TN} = \frac{153}{467}$. Therefofre,

$$\frac{\Pr(COVID)}{1 - \Pr(COVID)} = \frac{\frac{153}{467}}{\frac{467-153}{467}} \approx 0.49$$

Interpretation: This is the pre-test odds of Covid: $PRE\_ODDS = 0.49$. Using Bayes Theorem (BT), the post-test odds is given by $PRE\_ODDS \times LR = 0.49 * 13.5 = 6.64$. Key is to realise that LR can be used as a multiplier to convert pre-test odds to post-test odds. Some students may notice that this question is looking at how sample evidence on the pre-test prevalence of Covid is modified by additional evidence based on test data i.e. consider this using BT.

Section D.

11. For two events A, B in an outcome space Ω we say event A *attracts* event B if $Prob(B \mid A) > Prob(B)$ and event A *repels* event B if $Prob(B \mid A) < Prob(B)$.

(a) Prove that A attracts B if and only if B attracts A. Hence we can say that A and B are *mutually attractive* if A attracts B.

(b) Prove that A neither attracts nor repels B if and only if A and B are independent.

(c) Prove that A and B are mutually attractive if and only if $Prob(B \mid A) > Prob\left(B \mid \tilde{A}\right)$ where $\tilde{A}$ is the complement of A (that is those elements of Ω that are not in A so that $\Omega = A \cup \tilde{A}$ and A and $\tilde{A}$ are disjoint).

(d) Prove that if A attracts B, then A repels $\tilde{B}$, where $\tilde{B}$ is the complement of B.

(e) A standard pack of cards contains 52 cards divided into four suits. Let R_i be the event that the i^{th} player in a poker game is dealt a royal flush (A,K,Q,J,10 of one suit). Show that a royal flush attracts another royal flush, that is $Prob(R_2 \mid R_1) > Prob(R_2)$. Assume that, first, five cards are dealt to the first player, and then five cards from what remains in the pack are dealt to the second player.

Solution to 11. (a) Suppose that A attracts B, that is $Prob(B \mid A) > Prob(B)$. We have

$$Prob(B \mid A) = \frac{Prob(B \text{ and } A)}{Prob(A)}$$

Therefore,

$$Prob(B \text{ and } A) > Prob(A)\, Prob(B)$$

This implies that

$$Prob(A \mid B) = \frac{Prob(B \text{ and } A)}{Prob(B)} > Prob(A)$$

that is, B attracts A. Similarly, the statement B attracts A implies that A attracts B.

(b) If A neither attracts nor repels B then, we must have

$$Prob(B \mid A) = Prob(B)$$

That is,

$$Prob(B \text{ and } A) = Prob(A)\, Prob(B)$$

which means that A and B are independent. The implication in the other direction is straightforward.

(c) We have

$$Prob\left(B \mid \tilde{A}\right) = \frac{Prob\left(B \text{ and } \tilde{A}\right)}{Prob\left(\tilde{A}\right)}$$

On the other hand, $Prob\left(B \text{ and } \tilde{A}\right) + Prob(B \text{ and } A) = Prob(B)$ and $Prob\left(\tilde{A}\right) = 1 - Prob(A)$. Therefore,

$$Prob\left(B \mid \tilde{A}\right) = \frac{Prob(B) - Prob(B \text{ and } A)}{1 - Prob(A)},$$

and

$$Prob\,(B \mid A) > Prob\left(B \mid \tilde{A}\right)$$

if and only if

$$\frac{Prob\,(B \text{ and } A)}{Prob\,(A)} > \frac{Prob\,(B) - Prob\,(B \text{ and } A)}{1 - Prob\,(A)}$$

Elementary algebra shows that this is equivalent to

$$Prob\,(B \text{ and } A) > Prob\,(A)\,Prob\,(B)$$

which means that A and B are mutually attractive. Of course, starting from the last inequality, we can go back to

$$Prob\,(B \mid A) > Prob\left(B \mid \tilde{A}\right),$$

which proves the equivalence.

(d) If A attracts B then $Prob\,(B \text{ and } A) > Prob\,(A)\,Prob\,(B)$. Therefore,

$$\begin{aligned} Prob\left(\tilde{B} \text{ and } A\right) &= Prob\,(A) - Prob\,(B \text{ and } A) \\ &< Prob\,(A) - Prob\,(A)\,Prob\,(B) \\ &= Prob\,(A)\,Prob\left(\tilde{B}\right) \end{aligned}$$

Hence, A repels $\tilde{B}$.

(e) The probability of the royal flush of a specific suite dealt to the first player is

$$\frac{5}{52}\frac{4}{51}\frac{3}{50}\frac{2}{49}\frac{1}{48} = \frac{5!47!}{52!}$$

Hence, the probability of R_1 (royal flush of any suite dealt to the first player) is

$$Prob\,(R_1) = 4\frac{5!47!}{52!}.$$

The probability of the royal flush of a specific suite dealt to the second player is

$$\frac{47}{52}\frac{46}{51}\frac{45}{50}\frac{44}{49}\frac{43}{48}\frac{5}{47}\frac{4}{46}\frac{3}{45}\frac{2}{44}\frac{1}{43} = \frac{5!47!}{52!}$$

Hence,

$$Prob\,(R_2) = 4\frac{5!47!}{52!}.$$

The probability of the royal flushes of two specific suites dealt to the first and the second player is

$$\frac{5}{52}\frac{4}{51}\frac{3}{50}\frac{2}{49}\frac{1}{48}\frac{5}{47}\frac{4}{46}\frac{3}{45}\frac{2}{44}\frac{1}{43} = \frac{(5!)^2\,42!}{52!}$$

There are 12 possible ways to choose two specific suites (order matters) from 4. Therefore,

$$Prob\,(R_1 \text{ and } R_2) = 12\frac{(5!)^2\,42!}{52!}.$$

This is larger than

$$Prob\,(R_1)\,Prob\,(R_2) = 16\frac{(5!)^2\,(47!)^2}{(52!)^2}$$

Indeed,

$$\begin{aligned}\frac{12\frac{(5!)^2 42!}{52!}}{16\frac{(5!)^2(47!)^2}{(52!)^2}} &= \frac{3\,(42!)\,(52!)}{4\,(47!)^2} \\ &= \frac{3*48*49*50*51*52}{4*43*44*45*46*47} \approx 1.27\end{aligned}$$

Therefore, R_1 and R_2 are mutually attractive.

12. Consider the regression model

$$Y_i = \beta_1 + \beta_2 X_i + u_i \tag{1}$$

with $i = 1, 2, ..., n$, where β_1 and β_2 are unknown parameters, $E(u_i) = 0$ and $Cov\,(X_i, u_i) = 0$. Let $\bar{Y}, \bar{X}$, and $\bar{u}$ denote the respective sample means.

(a) Express β_2 in terms of $Cov(Y_i, X_i)$ and $Var\,(X_i)$. How is this expression related to $\widehat{\beta_2}$, the OLS estimate of β_2?

(b) We now demean X_i to obtain $X_i^* = X_i - \bar{X}$.

(i) Demonstrate that the OLS estimate of the intercept in the regression of Y_i on X_i^* will be equal to $\bar{Y}$.

(ii) Show that the OLS estimate of the slope coefficient in the regression of Y_i on X_i^* will be the same as in regression (1).

(c) Equation (1) implies

$$\bar{Y} = \beta_1 + \beta_2 \bar{X} + \bar{u}$$

and therefore

$$Y_i^* = \beta_2 X_i^* + u_i^* \tag{2}$$

for $i = 1, 2, ..., n$, where $Y_i^* = Y_i - \bar{Y}$, $X_i^* = X_i - \bar{X}$, and $u_i^* = u_i - \bar{u}$.

(i) Demonstrate that regression (2) will yield the same OLS estimate of β_2 as regression (1).

(ii) Determine the OLS estimate of the intercept if Y_i^* were regressed on X_i^* with an intercept included in the regression specification.

Solution to 12. (a) We have

$$Cov\,(Y_i, X_i) = Cov\,(\beta_1 + \beta_2 X_i + u_i, X_i) = \beta_2 Var\,(X_i)\,.$$

Therefore,

$$\beta_2 = \frac{Cov\,(Y_i, X_i)}{Var\,(X_i)}.$$

The OLS estimate

$$\hat{\beta}_2 = \frac{\frac{1}{n}\sum\left(Y_i - \bar{Y}\right)\left(X_i - \bar{X}\right)}{\frac{1}{n}\sum\left(X_i - \bar{X}\right)^2}$$

is the sample analogue of this expression.

(b) Let us denote the estimate of the intercept as $\hat{\alpha}$ and that of the slope as $\hat{\beta}$. We have

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}^* = \bar{Y}$$

because the sample mean $\bar{X}^*$ of X_i^* equals zero. Further,

$$\hat{\beta} = \frac{\sum \left(Y_i - \bar{Y}\right) X_i^*}{\sum X_i^{*2}} = \frac{\sum \left(Y_i - \bar{Y}\right) \left(X_i - \bar{X}\right)}{\sum \left(X_i - \bar{X}\right)^2} = \hat{\beta}_2.$$

(c) Let $\tilde{\beta}_2$ be the OLS estimate of β_2 from the demeaned regression. We have

$$\tilde{\beta}_2 = \frac{\sum Y_i^* X_i^*}{\sum X_i^{*2}} = \frac{\sum \left(Y_i - \bar{Y}\right) \left(X_i - \bar{X}\right)}{\sum \left(X_i - \bar{X}\right)^2} = \hat{\beta}_2.$$

Had intercept been included in the specification, the OLS estimate of β_2 would stay the same because $\bar{Y}^* = \bar{X}^* = 0$. The estimate $\tilde{\alpha}$ of the intercept would have been

$$\tilde{\alpha} = \bar{Y}^* - \tilde{\beta}_2\bar{X}^* = 0.$$

END OF PAPER