



GUTENBERG SCHOOL OF
MANAGEMENT
& ECONOMICS



Johannes Gutenberg-Universität Mainz
Master in International Economics and Public Policy

Chair in Macroeconomics

Economic Methods

2022/23 winter term

Wanda Schleder

Contents

1	Introduction	1.0
2	Preliminaries: formulating a problem in economics	2.0
I	Static Utility Maximisation	3.0
3	Typical static maximisation problem	3.0
3.1	Setup	3.0
3.2	Example: Optimal goods allocation (Prof. Schunk: PPE)	3.2
3.3	Exercises	3.5
II	Cost Minimisation	4.0
4	Typical minimisation problem	4.0
4.1	Setup	4.0
4.2	Example: firm cost minimisation (Prof. Harms: Int'l Trade)	4.3
4.3	Exercises	4.6

III	Profit Maximisation	5.0
5	Profit maximisation and factor demands	5.0
5.1	Setup	5.0
5.2	Example: Capital demand (Prof. Barbaro: Devp't + Growth)	5.2
5.3	Exercises	5.4
6	Profit maximisation under monopoly	6.0
6.1	Setup	6.0
6.2	Example: A Monopolist (Prof. Schunk: PPE)	6.2
6.3	Exercises	6.6
IV	Envelope Theorem	7.0
7	Definition and application	7.0
7.1	Preliminaries	7.0
7.2	Example: Shephard's Lemma (Prof. Harms: Int'l Trade)	7.2
7.3	Exercises	7.4
V	Game Theory	8.0

8	Nash Games and Repeated Nash Games	8.0
8.1	Standard Nash Game	8.0
8.2	Repeated Nash Games	8.1
8.3	Example: Costs of default (Prof. Sauré: Fin Economics II)	8.2
8.4	Exercises	9.0
VI	Intertemporal Maximisation Problem	9.0
9	Discrete time model	9.1
9.1	Setup	9.1
9.2	Example: Sovereign default (Prof. Sauré: Fin. Economics I)	9.3
9.3	Exercises	9.4
10	Continuous time problem	10.0
10.1	Setup	10.0
10.2	Example: the central planner (Prof. Wälde: Advanced Macro)	10.2
10.3	The Shell Model	10.4
10.4	Exercises	10.7
VII	Ordinary Least Squares	11.0

11 Typical OLS regression	11.0
11.1 Preliminaries: method of moments for 1 explanatory variable	11.0
11.2 Example: Multiple Regression (Prof. van Ewijk: Intro Econometrics)	11.3
11.3 Exercises	11.4
 VIII Probability Theory	 12.0
12 Distribution	12.0
12.1 Discrete random variable	12.0
12.2 Continuous random variable	12.1
12.3 Example: Distribution of discrete and continuous RV (Prof. Schank: Stats Methods)	12.2
12.4 Exercises	12.5
 13 Mean and variance	 13.0
13.1 Discrete random variable	13.0
13.2 Continuous random variable	13.1
13.3 Example: Mean and variance of discrete and continuous RV (Prof. Schank: Stats Methods)	13.2
13.4 Exercises	13.3



GUTENBERG SCHOOL OF
MANAGEMENT
& ECONOMICS



Johannes Gutenberg-Universität Mainz
Master in International Economics and Public Policy

Chair in Macroeconomics

Economic Methods

2022/23 winter term

Wanda Schleder

Contact info:

Wanda Schleder

Office: 01-413 (Container building behind old ReWi)

E-mail: waschled@uni-mainz.de

Office hours: by appointment

1 Introduction

This *Economic Methods* course

- covers
 - Static utility maximisation
 - Cost minimisation
 - Profit maximisation
 - * Profit maximisation and factor demands
 - * Profit maximisation under monopoly
 - Envelope theorem
 - Intertemporal maximisation
 - Basic probability theory
 - Ordinary least squares
- lasts 7 weeks and ends before Christmas break
- has **no** examination at the end

- is intended for
 - students who have not done economic theory before (or a long time ago)
 - students who need a little more help connecting abstract maths concepts to economic problems
- builds on
 - *Maths Review* (by the Chair of Public and Behavioural Economics): covers basic linear algebra, real analysis, and some probability theory
 - *MIEPP courses* (compulsory and elective): pose economic problems to be solved
 - \Rightarrow *this course helps connect the two*

Therefore we begin with the Core Modules

Table 1: MIEPP Timetable Winter Semester 2022/2023

Week #	Dates	Courses
1	24/10 to 28/10	<i>“Maths Review”</i>
2	31/10 to 04/10	<i>“Core Modules”</i> lectures start
2	31/10 to 04/11	<i>“Economic Methods”</i> starts
...
8	12/12 to 16/12	<i>“Economic Methods”</i> ends
...

The course format will be structured as follows...

- Classes twice a week: 1 lecture + 1 tutorial (see [Macroeconomics Chair website](#) for details)
- Lecture (90 mins):
 - (i) We look at a typical economic problem (e.g. utility maximisation).
 - (ii) A solution method is introduced for this class of problems (e.g. Lagrangian).
 - (iii) We discuss and explain an example from one of the other lectures (e.g. PPE).
- Tutorial (90 mins):
 - (i) You (the students) solve problems on the board in class.
 - (ii) At the end, we check your steps/results together.

⇒ Course material available online on the [Macroeconomics Chair website](#) before classes start.

Table 2: Mapping of topics with MIEPP courses

Part	Topic	Professor (<i>course</i>)	Semester	Section
I	Static maximisation	Schunk (<i>Principles Pub Econ</i>)	Winter	3
II	Cost minimisation	Harms (<i>International Trade</i>)	Winter	4
III	Profit max.: - <i>factor demand</i> - <i>monopoly power</i>	Barbaro (<i>Dev't + Growth</i>)	Winter	5
		Schunk (<i>Principles Pub Econ</i>)	Winter	6
IV	Envelope theorem	Harms (<i>International Trade</i>)	Winter	7
V	Game Theory	Sauré (<i>Fin Econonomics II</i>)	Winter	8
VI	Intertemp. max.: - <i>discrete time</i> - <i>continuous time</i>	Sauré (<i>Fin Econonomics I</i>)	Summer	9
		Wälde (<i>Advanced Macro</i>)	Winter	10
VII	Ordinary least squares	van Ewijk (<i>Intro Econometrics</i>)	Winter	11
VIII	Probability: - <i>distribution</i> - <i>mean and variance</i>	Schank (<i>Stats Methods</i>)	Summer	12
		Schank (<i>Stats Methods</i>)	Summer	13

2 Preliminaries: formulating a problem in economics

Simplest possible definition:

Economic theory studies the optimal behaviour of agents.

How do we do this? We define

- an **objective function** (or payoff in game theory), i.e. the goal the agent is trying to achieve,
- a set of possible *choices*,
- and a **trade-off**, i.e. some cost to the choices made.

⇒ no trade-off = no problem!

Next, we apply mathematical tools to ...

- ... find the best possible *choice* given the **objective** and the **trade-off** faced by the agent.

Why use maths?

- Forces us to have clear definitions.
- Helps to have clear relationships between various elements of the model.
- Helps to apply logic to find the “true” best possible solution to a problem.
 - \Rightarrow 1 economic model tackles 1 problem, simplified to its bare minimum.

Empirical economics uses statistics (a sub-branch of mathematics) to

- find out if a model accurately represents reality (i.e. data),
- and test predictions of models.
- (*also to test relationships between variables without model \rightsquigarrow bad research*)

\Rightarrow holds economic theory accountable, while economic theory gives problems for empirical economics to investigate, i.e. they complement each other.

Part I

Static Utility Maximisation

3 Typical static maximisation problem

3.1 Setup

Consider the following consumption allocation problem,

$$\max_{\{c_i\}_{i=1}^N} u(c_1, \dots, c_N),$$

subject to

$$W = \sum_{i=1}^N p_i c_i = p_1 c_1 + p_2 c_2 + \dots + p_N c_N.$$

- An individual makes *choices* over many consumption goods $(\{c_i\}_{i=1}^N := c_1, \dots, c_N)$ given a budget set $(W > 0)$ to maximise *utility* at a given point in time (micro perspective \rightsquigarrow static). Each good i costs p_i .
- The problem can be solved using the ***Lagrangian*** method

- Lagrangian method consists in (i) setting up the Lagrangian function and (ii) taking the first-order conditions (FOCs) to determine the optimal solution.

(i) First we set up the Lagrangian function,

$$\mathbb{L}(c_1, \dots, c_N; \lambda) = u(c_1, \dots, c_N) + \lambda \left[W - \sum_{i=1}^N p_i c_i \right],$$

(ii) next we take the FOCs,

$$\frac{\partial \mathbb{L}}{\partial c_i} = \frac{\partial u(c_1, \dots, c_N)}{\partial c_i} - \lambda p_i = 0, \quad \forall i.$$

- From this, we can also compute the marginal rate of substitution (MRS) between two goods $i, j \in [1, N]$ defined as

$$MRS_{ij} = \frac{\partial u(c_1, \dots, c_N) / \partial c_i}{\partial u(c_1, \dots, c_N) / \partial c_j}.$$

\Rightarrow Useful for computing slope of indifference curves ($= -MRS$)

3.2 Example: Optimal goods allocation (Prof. Schunk: PPE)

Optimal allocation of goods

- Formally:
Maximize the utility of A for a given utility of B and the resource constraint:

$$\begin{aligned} \max_{x_1^A, x_2^A, x_1^B, x_2^B} u(x_1^A, x_2^A) \quad & s.t. \quad u(x_1^B, x_2^B) \geq \bar{u} \\ & x_1^A + x_1^B = e_1^A + e_1^B \\ & x_2^A + x_2^B = e_2^A + e_2^B \end{aligned}$$

- Lagrange function:

$$\begin{aligned} L = & u(x_1^A, x_2^A) + \lambda(u(x_1^B, x_2^B) - \bar{u}) + \mu_1(x_1^A + x_1^B - e_1^A - e_1^B) \\ & + \mu_2(x_2^A + x_2^B - e_2^A - e_2^B) \end{aligned}$$

Optimal allocation of goods

- Lagrange function:

$$L = u(x_1^A, x_2^A) + \lambda(u(x_1^B, x_2^B) - \bar{u}) + \mu_1(x_1^A + x_1^B - e_1^A - e_1^B) + \mu_2(x_2^A + x_2^B - e_2^A - e_2^B)$$

- First order conditions:

$$x_1^A : \frac{\partial u^A}{\partial x_1^A} + \mu_1 = 0$$

$$x_2^A : \frac{\partial u^A}{\partial x_2^A} + \mu_2 = 0$$

$$x_1^B : \lambda \frac{\partial u^B}{\partial x_1^B} + \mu_1 = 0$$

$$x_2^B : \lambda \frac{\partial u^B}{\partial x_2^B} + \mu_2 = 0$$

Optimal allocation of goods

- From the FOCs, it follows:

$$\frac{\mu_1}{\mu_2} = \frac{\frac{\partial u^A}{\partial x_1^A}}{\frac{\partial u^A}{\partial x_2^A}} = \frac{\frac{\partial u^B}{\partial x_1^B}}{\frac{\partial u^B}{\partial x_2^B}}$$

$$\boxed{-MRS^A = -MRS^B}$$

- Interpretation:
An efficient allocation is achieved when the marginal rates of substitution (in consumption) are equal for all individuals
- See also: Contract curve

3.3 Exercises

In a world with two goods, 1 and 2, and two individuals A and B , what is the rule governing the optimal allocation of goods between A and B (i.e. the MRS_{AB})? Use the same setup as in Example 3.2.

1. Use the following functional form: $u(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, with $0 < \alpha < 1$.
2. Use the following functional form: $u(x_1, x_2) = \ln x_1 + \ln x_2$.
3. Use the following functional form: $u(x_1, x_2) = \frac{x_1^{1-\sigma}-1}{1-\sigma} + \frac{x_2^{1-\sigma}-1}{1-\sigma}$, with $\sigma > 0$ and $\sigma \neq 1$.
4. Use the following functional form: $u(x_1, x_2) = \sum_{i=1}^2 x_i^\theta$, with $\theta > 0$.

Part II

Cost Minimisation

4 Typical minimisation problem

4.1 Setup

Consider the following problem where costs of production are a function of inputs (v_i) and prices (w_i) for each factor of production i

$$\min_{\{v_i\}_{i=1}^N} c(v_1, \dots, v_N) = \sum_{i=1}^N v_i w_i = v_1 w_1 + v_2 w_2 + \dots + v_N w_N$$

The firm uses factor inputs $\{v_i\}_{i=1}^N := v_1, \dots, v_N$ to produce a given level of output $y > 0$ using a technology $f(\cdot)$, where $\partial f(\cdot) / \partial v_i > 0$ and $\partial^2 f(\cdot) / \partial v_i^2 < 0$, i.e. production increases in the factors of production at a decreasing rate,

$$y = f(v_1, \dots, v_N).$$

- For a given output level y , a firm chooses a combination of factor inputs $\{v_i\}_{i=1}^N$ to minimise the costs at given factor prices $\{w_i\}_{i=1}^N$.

- Solving this problem yields *conditional factor demands*
- Once more, we can use the Lagrangian method,

$$\mathbb{L}(v_1, \dots, v_N; \lambda) = \underbrace{\sum_{i=1}^N v_i w_i}_{=c(v_1, \dots, v_N)} + \lambda [y - f(v_1, \dots, v_N)],$$

and once again obtain the FOCs

$$\frac{\partial \mathbb{L}}{\partial v_i} = w_i - \lambda \frac{\partial f(v_1, \dots, v_N)}{\partial v_i} = 0, \quad \forall i$$

such that looking at two separate factors of production i and j , where $i \neq j$, we would have *conditional* factor demands fixed by their respective price

$$w_i = \lambda \frac{\partial f(v_1, \dots, v_N)}{\partial v_i},$$

$$w_j = \lambda \frac{\partial f(v_1, \dots, v_N)}{\partial v_j},$$

and dividing one by the other yields the optimality condition for the relative demand for both factors of production as the ratio of their respective price being equal to the ratio

of their marginal effect on the production function

$$\frac{w_i}{w_j} = \frac{\partial f(v_1, \dots, v_N) / \partial v_i}{\partial f(v_1, \dots, v_N) / \partial v_j}.$$

- Where are the N equations and N unknowns?
Note that we only have N-1 First Order Conditions
- We have N+1 unknowns (v_1 - v_N and λ), N First-Order Conditions and one constraint, therefore N+1 equations.

4.2 Example: firm cost minimisation (Prof. Harms: Int'l Trade)

Harms/International Trade



Supply

- The firm's cost minimization problem (contd.):
 - Special case: Two factors of production (K and L) with factor prices r (rental rate of capital) and w (wage).

$$\underset{K_i, L_i}{Min} \quad c_i = r K_i + w L_i$$

subject to ("s.t.") the constraint: $f_i(K_i, L_i) = y_i$

- **Iso-cost line**: $K_i = \frac{c_i}{r} - \frac{w}{r} L_i$ (slope: - w/r)

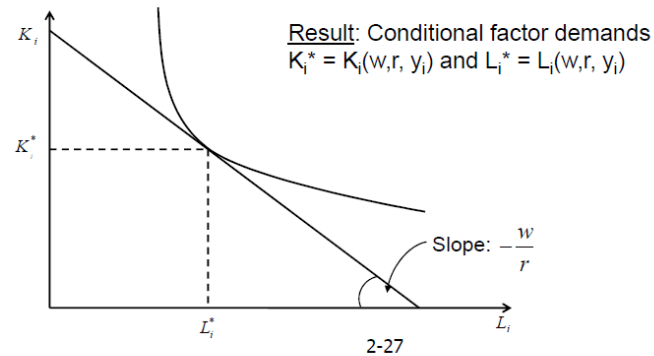
- Optimality condition:

$$\frac{w}{r} = \frac{\partial f_i(K_i, L_i) / \partial L_i}{\partial f_i(K_i, L_i) / \partial K_i}, \text{ i.e. optimum defined by } \textbf{point of tangency} \text{ between isoquant and iso-cost line.}$$

2-26

- The firm's cost minimization problem (contd.):

Figure 2.4: The firm's cost minimization problem with 2 factor inputs



- Minimization of c_i moves iso-cost line inwards

The problem above gives us the following Lagrangian function

$$\mathbb{L}(K_i, L_i; \lambda) = rK_i + wL_i + \lambda [y_i - f_i(K_i, L_i)],$$

with FOCs

$$\begin{aligned}\frac{\partial \mathbb{L}}{\partial K_i} &= r - \lambda \frac{\partial f_i(K_i, L_i)}{\partial K_i} = 0 \Leftrightarrow r = \lambda \frac{\partial f_i(K_i, L_i)}{\partial K_i}, \\ \frac{\partial \mathbb{L}}{\partial L_i} &= w - \lambda \frac{\partial f_i(K_i, L_i)}{\partial L_i} = 0 \Leftrightarrow w = \lambda \frac{\partial f_i(K_i, L_i)}{\partial L_i}.\end{aligned}$$

Dividing the second FOC by the first one we obtain the optimality condition from the slides above

$$\frac{w}{r} = \frac{\partial f_i(K_i, L_i) / \partial L_i}{\partial f_i(K_i, L_i) / \partial K_i}.$$

- Where are the N equations and N unknowns?
- N=2:
 - 2 unknowns: capital and labour
 - 2 equations: FOC and constraint

4.3 Exercises

Consider a firm i producing output $y_i \geq 0$ using technology $f_i(K_i, L_i) \geq 0$ using two factors of production, $K_i \geq 0$ and $L_i \geq 0$. What is the optimality condition governing factors demand? Use the same setup as in Example 4.2.

1. Use the following functional form: $f_i(K_i, L_i) = K_i^\alpha L_i^{1-\alpha}$, with $0 < \alpha < 1$.
2. Use the following functional form: $f_i(K_i, L_i) = [\alpha K_i^\epsilon + (1 - \alpha) L_i^\epsilon]^{\frac{1}{\epsilon}}$, with $0 < \alpha < 1$ and $\epsilon < 1$.

Part III

Profit Maximisation

5 Profit maximisation and factor demands

5.1 Setup

Consider the following problem where a firm tries to maximise profits π

$$\max_{\{v_i\}_{i=1}^N} \pi = y - c(v_1, \dots, v_N)$$

subject to

$$y = f(v_1, \dots, v_N)$$

and as before costs are given by

$$c(v_1, \dots, v_N) = \sum_{i=1}^N w_i v_i.$$

- Given factor prices $\{w_i\}_{i=1}^N$, a firm chooses a combination of factor inputs $\{v_i\}_{i=1}^N$ to maximise its profits.

- Solving this problem yields *unconditional factor demands*
- We can solve this problem by the Lagrangian method once more,

$$\mathbb{L}(v_1, \dots, v_N; \lambda, \mu) = y - c(v_1, \dots, v_N) + \lambda [f(v_1, \dots, v_N) - y] + \mu \left[c(v_1, \dots, v_N) - \sum_{i=1}^N w_i v_i \right]$$

... but this becomes a little complicated.

- Instead, we can employ a simpler method, known as the **substitution** method, where (i) we replace terms by their definitions and (ii) solve directly for optimal factor allocation:

$$\max_{\{v_i\}_{i=1}^N} \pi = y - c(v_1, \dots, v_N) \Leftrightarrow \max_{\{v_i\}_{i=1}^N} \pi = f(v_1, \dots, v_N) - \sum_{i=1}^N w_i v_i,$$

giving us the FOCs,

$$\frac{\partial \pi}{\partial v_i} = \frac{\partial f(v_1, \dots, v_N)}{\partial v_i} - w_i = 0 \Leftrightarrow w_i = \frac{\partial f(v_1, \dots, v_N)}{\partial v_i}, \quad \forall i,$$

such that each factor of production is demanded up until its marginal benefit to production equals its price (i.e. its marginal cost).

5.2 Example: Capital demand (Prof. Barbaro: Devp't + Growth)

3. Openness and Factor Accumulation

- Growth with Capital Mobility
 - Production function: $y = Ak^\alpha$
 - Marginal Product of Capital: $MPK = \alpha Ak^{\alpha-1}$
 - Rental rate of capital: $r = MPK = \alpha Ak^{\alpha-1}$
 - Perfect openness to capital flows: $r = r_w$
 - Level of capital per worker: $r_w = \alpha Ak^{\alpha-1} \Leftrightarrow k = (\alpha A / r_w)^{1/(1-\alpha)}$
 - Result:
 - If $r_{\text{home}} < r_w$: capital mobility leads to a lower capital stock
 - If $r_{\text{home}} > r_w$: capital mobility leads to an increasing capital stock
- If the country is small with perfect capital mobility, domestic savings are negligible for the growth rate of a country

This model starts from a representative firm producing aggregate output Y using aggregate capital K and labour L with a Cobb-Douglas production function with parameter $0 < \alpha < 1$ and total factor productivity (TFP) A

$$Y = AK^\alpha L^{1-\alpha}.$$

representative firm maximises profits, which are given by

$$\pi = Y - rK - wL.$$

Maximising profits by choosing capital optimally, using the substitution method, we obtain

$$\begin{aligned} \frac{\partial}{\partial K}\pi &= \frac{\partial}{\partial K} (AK^\alpha L^{1-\alpha} - rK - wL) = 0 \\ \Leftrightarrow K &= \left(\frac{\alpha A}{r}\right)^{\frac{1}{1-\alpha}} L. \end{aligned}$$

Note that we can rewrite K/L as k , which is interpreted as capital per worker. If r is fixed because the country is small in comparison to the world (so is a price taker for capital, which is mobile, and $r = r_w$), we can rewrite the condition above

$$k^* = \left(\frac{\alpha A}{r_w}\right)^{\frac{1}{1-\alpha}}.$$

5.3 Exercises

Consider a country producing aggregate output $Y \geq 0$ using technology $F(K, L) \geq 0$ using aggregate capital, $K \geq 0$ and aggregate labour $L \geq 0$. What will capital per worker (k) be equal to? Use the same setup as in Example 5.2.

1. Use the following functional form: $Y = K^\alpha (AL)^{1-\alpha}$, with $0 < \alpha < 1$.
2. Use the following functional form: $Y = \left[\alpha K^\theta + (1 - \alpha) (AL)^\theta \right]^{\frac{1}{\theta}}$, with $0 < \alpha < 1$ and $\theta < 1$.

6 Profit maximisation under monopoly

6.1 Setup

Consider the following profit maximisation problem for some firm with monopoly power, where we now consider that the price in this market will depend on the output of the firm (*was not the case before with perfect competition, because each firm was too small to affect the price* → *not the case with monopoly*)

$$\max_y \pi = p(y) y - c(w_1, \dots, w_N, y)$$

- For a given cost function $c(w_1, \dots, w_N, y)$, the monopolistic firm chooses output y in order to maximise profits, thus setting optimal output production.
- Solving this problem yields the optimal monopolistic price as the reciprocal of the optimal output: $y(p) \Leftrightarrow p := y^{-1}(p)$. In other words: choosing output fixes the price in the market, because the firm has “monopoly power”.

- To solve this problem, we can immediately take the derivative w.r.t. output y and solve for the optimal level...

$$\begin{aligned}\frac{\partial \pi}{\partial y} &= p'(y)y + p(y) - c_y(w_1, \dots, w_N, y) = 0 \\ \Leftrightarrow c_y(w_1, \dots, w_N, y) &= p'(y)y + p(y) \\ c_y(w_1, \dots, w_N, y) &= p(y) \left[\underbrace{\frac{dp(y)}{dy} \frac{y}{p(y)}}_{\text{= elasticity of price in output}} + 1 \right]\end{aligned}$$

- Rent creation is one possible application of monopoly power, whereby a firm creates a monopoly in order to extract "rent" from consumers.
 \Rightarrow We can assume that $c_y(w_1, \dots, w_N, y)$ has the functional form $c(w_1, \dots, w_N)y$ (see exercises chapter 7). Then solve for y

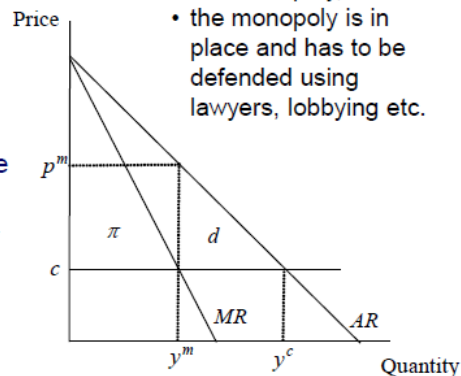
6.2 Example: A Monopolist (Prof. Schunk: PPE)

Social Cost of Monopoly

- Monopoly is a cause of inefficiency
 - Output is set below the competitive level to maximize profit
- Some consumer surplus becomes profit, some becomes deadweight loss
- Standard analysis views **only** the deadweight loss as the cost of monopoly power
- The application of rent-seeking suggests that the cost may be **greater**
 - Resources are used to obtain or defend the monopoly position
 - → See next slide

Social Cost of Monopoly

- Consider a monopoly producing with constant marginal cost c and no fixed cost
- The standard deadweight loss is area d and monopoly profit is π
- If there are competitors for the monopoly position value, π will be dissipated
- The total social cost of monopoly is then $\pi + d$
- Calculations show d may be small (as % GDP) but $\pi + d$ can be large
- The social costs of monopoly may be much greater than measurements through deadweight loss would suggest



- Competitors lobby the government for getting the monopoly, or
- the monopoly is in place and has to be defended using lawyers, lobbying etc.

- The maximization problem

Behind this figure, there is a simple model of profit maximization. It reads

$$\max_y \pi = p(y)y - cy = \text{revenue} - \text{cost}. \quad (6.1)$$

The dependence of the price p on output y of the monopolist is captured by the function $p(y)$. This function can be thought of as reflecting a goods demand function by utility-maximizing households. The costs of producing output y are given by cy . Marginal costs are therefore c . Such a cost function results e.g. from a linear production function $y = Al$ where A is labour productivity and l is labour input. Given a wage w , marginal costs c would be w/A .

- Average and marginal revenue

Average revenue is $\frac{p(y)y}{y} = p(y)$ is the demand curve. In the figure, it is visible as a straight line, i.e. it takes the form $p(y) = a - by$ where a and b are constants. Note that such a (inverse) demand curve rarely comes out of standard utility functions (but try the quadratic utility function). Marginal revenue is

$$MR \equiv \frac{d(p(y)y)}{dy} = \frac{dp(y)}{dy}y + p(y)$$

With the linear functional form $p(y) = a - by$ from above, this reads

$$MR = -by + a - by = a - 2by.$$

It is also a line and falls more steeply than the average cost curve, as visible in the figure as well.

- Optimality

Optimal output is described by the first-order condition

$$MR = MC,$$

where the latter is c with our profit function (6.1). This intersection point gives the optimal quantity y^m and optimal price p^m . Optimal means here of course from the perspective of the monopolist but not from the perspective of consumers or social welfare (i.e. society).

When we express optimality in terms of the condition (??) from above, we employ $c_y(w_1, \dots, w_N, y) = c$ and can express it as

$$c = p(y) \left[\frac{dp(y)}{dy} \frac{y}{p(y)} + 1 \right] \Leftrightarrow p = \frac{c}{1 + \frac{dp(y)}{dy} \frac{y}{p(y)}}.$$

In this expression, $\frac{1}{1 + \frac{dp(y)}{dy} \frac{y}{p(y)}}$ is called the mark-up of the price p over marginal costs c . Note that the mark-up is larger than one as the demand function falls in the price, i.e. $\frac{dp(y)}{dy} < 0$ and therefore the denominator is smaller than one.

6.3 Exercises

Consider a monopolist from Example 6.2 and help them to maximize profits – even though this might be considered as a morally questionable task. Let the (inverse) demand function be given by $p(y) = y^{-1/\varepsilon}\Phi$, where ε is price elasticity of demand and Φ is a constant (capturing, inter alia, total expenditure, a price index and preference parameters).

1. Illustrate the link between the demand function as specified above, $p(y) = y^{-1/\varepsilon}\Phi$, and Dixit-Stiglitz or Cobb-Douglas demand functions known from other lectures in trade or macro.
2. Solve the profit maximization problem using a cost function that is cubic, i.e. $C(y) = \frac{1}{3}y^3$. What is the optimal level of output y^* ? What are profits π^* given by at this production level?
3. Solve the problem for a firm under perfect competition.

Part IV

Envelope Theorem

7 Definition and application

7.1 Preliminaries

Theorem Let $f(x_1, \dots, x_N; a)$ be a continuous function of $(x_1, \dots, x_N) \in \mathbb{R}^N$ and the scalar a . For each choice of parameter a , consider the unconstrained maximisation problem

$$\max_{\{x_i\}_{i=1}^N} f(x_1, \dots, x_N; a).$$

Let $x_1^*(a), \dots, x_N^*(a)$ be solutions to this problem. Suppose that $x_1^*(a), \dots, x_N^*(a)$ are continuous functions of a . Then,

$$\frac{d}{da} f(x_1^*(a), \dots, x_N^*(a); a) = \frac{\partial}{\partial a} f(x_1^*(a), \dots, x_N^*(a); a).$$

In other words, $\frac{\partial f(\cdot)}{\partial x_i^*(a)} = 0$ and we only need to look at the partial derivative of $f(\cdot)$ w.r.t. a as a direct argument.

Intuition:

$$\begin{aligned}\frac{d}{da}f(x_1^*(a), \dots, x_N^*(a); a) &= \underbrace{\frac{\partial f(x_1^*(a), \dots, x_N^*(a); a)}{\partial x_1^*(a)}}_{=0} \frac{\partial x_1^*(a)}{\partial a} + \dots + \frac{\partial f(x_1^*(a), \dots, x_N^*(a); a)}{\partial a} \\ &= \frac{\partial f(x_1^*(a), \dots, x_N^*(a); \textcolor{red}{a})}{\partial \textcolor{red}{a}}\end{aligned}$$

$x_i^*(a)$ are solutions to the problem \rightarrow increasing a has no effect on the overall function because each $x_i^*(a) \rightarrow$ is already the optimal choice for $f(\cdot)$, thus only direct partial effect of a matters

- For graphical illustration check chapter 4.2 of the book ” Applied intertemporal optimization” by Klaus Wälde (2012)

7.2 Example: Shephard's Lemma (Prof. Harms: Int'l Trade)

Harms/International Trade

Supply



- The firm's cost minimization problem (contd.):

- Shephard's Lemma:

$$\frac{\partial c_i}{\partial \omega_j} = v_{ij}(\omega, y_i)$$

Note: This is an application of the **Envelope Theorem**.

(The cost function is an **optimal-value function**, i.e. it is based on solution of an optimization problem. As a consequence, only the *direct* effect of a (marginal) change in factor prices on costs has to be considered.)

Harms/International Trade

Supply

- The firm's cost minimization problem (contd.):
 - The solution to the firm's cost minimization problem yields the **cost function**

$$c_i = c_i(\omega, y_i) = \sum_{j=1}^M \omega_j v_{ij}(\omega, y_i) \quad \text{— Conditional demand for factor } j$$

The *minimal* costs of producing y_i given the vector of factor prices ω .

- **Marginal costs:** $MC_i = \frac{\partial c_i}{\partial y_i}$ **Average costs:** $AC_i = \frac{c_i}{y_i}$
- With constant returns to scale

$$c_i(\omega, y_i) = \varphi_i(\omega) \cdot y_i$$

Implication: For given factor prices, $MC_i = AC_i$ independent of output level.

2-28

7.3 Exercises

Using the solution to the cost minimisation of Section 4, i.e. $\frac{w}{r} = \frac{\partial f(K^*, L^*) / \partial L}{\partial f(K^*, L^*) / \partial K}$, and the definition of the cost function $c = rK^* + wL^*$ (where K^* and L^* are the conditional factor demands), use the following functional forms to rewrite the cost function as $c = \phi(w, r)y$, where $\phi(\cdot)$ is some function of factor prices.

1. Using a Cobb-Douglas function, $y = f(K, L) = K^\alpha L^{1-\alpha}$, where $0 < \alpha < 1$.
2. Show that marginal costs $\partial c / \partial y$ equals average costs c / y for the Cobb-Douglas case.
3. Using a CES function, $y = f(K, L) = [\alpha K^\theta + (1 - \alpha) L^\theta]^{\frac{1}{\theta}}$, where $0 < \alpha < 1$ and $\theta < 1$.
4. Show that marginal costs $\partial c / \partial y$ equals average costs c / y for the CES case.

Part V

Game Theory

8 Nash Games and Repeated Nash Games

8.1 Standard Nash Game

In a standard Nash Game, we have two players A and B that can choose between two strategies. Call these strategies *cooperate* and *defect*. Depending on the choice of the two players, these scenarios will lead to different payoffs for the players. Let's take the Prisoner's Dilemma as an example:

		Player B	
		<i>Cooperate</i>	<i>Defect</i>
Player A	<i>Cooperate</i>	(1, 1)	(-1, 2)
	<i>Defect</i>	(2, -1)	(0, 0)

Although cooperating yields the highest total payoff, cooperate is a **dominated strategy**. It is always optimal to Defect, no matter what the other player does. The Strategy (*Defect, Defect*) is therefore a **Nash Equilibrium**.

8.2 Repeated Nash Games

The optimal strategy can change once you repeat the Nash Game.

- What happens if you repeat the Nash Game a finite number of times?

Answer: Nothing changes! Rational agents will still defect in every period. Why?

- assume $T=10$ and that Agent A and B cooperate for all 10 periods. Then The payoff for players A and B will be 10.
- Now check for incentive to deviate with backward induction. Assuming that in the first 9 periods the agents cooperated, there is an incentive to deviate in the last period, as the total payoff will be 11 for the person who defected.
- Both will think so and therefore both will defect in the last period.
- same logic holds for the second to last period ...
- In the end, the agents will always defect.
- What happens if you repeat the Nash Game an infinite number of times (no last period, backward induction not possible)?

8.3 Example: Costs of default (Prof. Sauré: Fin Economics II)

Detour: Costs of default – Prisoner's Dilemma and Reputation

Resolving the Dilemma: Repetition of the game

Assume now

- ▶ the game is repeated indefinitely in periods $t = 0, 1, 2, \dots$
- ▶ both players discount future benefits with the factor $\beta < 1$.

Reputation-linked strategies:

In period t player i plays

- ▶ C if player $j \neq i$ has played C in all periods $t' < t$
- ▶ D if not.

Detour: Costs of default – Prisoner's Dilemma and Reputation

Comments (Reputation-linked strategies):

- ▶ If player j plays D for the first time, player i still plays C .
 - ▶ A player can *defect* by playing uncooperatively surprising the other player exactly once.
 - ▶ The other player who is defected on is 'too slow' to react within the defection period and is exposed to the uncooperative behaviour.
- ▶ A player's *reputation* depends on his/her past actions:
 - ▶ as long as a player has a history of cooperation, he/she has a good reputation
 - ▶ as soon as he/she defects, he/she has a bad reputation; the partner's trust erodes and cooperation ends.

Detour: Costs of default – Prisoner's Dilemma and Reputation

Value of cooperation:

The value of sustained cooperation equals the net present value of all future gains from cooperation. In our example, this is:

$$V(C) = \sum_{t=0}^{\infty} \beta^t \times 1 = \frac{1}{1 - \beta}$$

Value of defection:

The value of (one-time) defection at period $t = 0$ is

$$V(D) = 2 + \sum_{t=1}^{\infty} \beta^t \times 0 = 2$$

8.4 Exercises

		Player 2		
		L	C	R
Player 1	T	6, 6	0, 0	0, 7
	B	0, 0	-1, -1	0, -1
	M	7, 0	-1, 0	1, 1

1. Take the table above. are there any dominated strategies? What is the Nash equilibrium?
2. Suppose the game is played infinitely many times. What is the value of defecting and what is the value of cooperating (cooperating in this case means playing (T,L))? What does discount factor have to be in order for cooperating to be the optimal strategy?

Part VI

Intertemporal Maximisation Problem

9 Discrete time model

9.1 Setup

Consider a typical macroeconomic model within this framework, the **Social Planner Problem**:

$$\max_{\{C_t\}_{t=0}^{\infty}} U_0 = \sum_{t=0}^{\infty} \beta^t u(C_t)$$

for $0 < \beta < 1$, subject to

$$K_{t+1} = (1 + r) K_t - \delta K_t + Y_t - C_t = (1 + r - \delta) K_t + Y_t - C_t.$$

The capital stock increase with a return rate $r > 0$ on its previous level, minus depreciation at rate $0 < \delta < 1$, as well as with production Y_t , but decreases with consumption C_t . What is not consumed, $Y_t - C_t$, gets re-invested into capital.

- This problem can be solved using the Lagrangian method,

$$\mathbb{L}(C_t, K_{t+1}; \lambda_t) = \sum_{t=0}^{\infty} \beta^t \{u(C_t) + \lambda_t [(1+r-\delta)K_t + Y_t - C_t - K_{t+1}]\}$$

Taking the FOCs w.r.t. C_t and K_{t+1} yields

$$\begin{aligned} \frac{\partial \mathbb{L}}{\partial C_t} &= \beta^t \{u'(C_t) + \lambda_t [-1]\} = 0 \\ \Leftrightarrow u'(C_t) &= \lambda_t, \end{aligned}$$

$$\begin{aligned} \frac{\partial \mathbb{L}}{\partial K_{t+1}} &= \beta^t \{\lambda_t [-1]\} + \beta^{t+1} \{\lambda_{t+1} (1+r-\delta)\} = 0 \\ \Leftrightarrow \lambda_t &= \beta \lambda_{t+1} (1+r-\delta). \end{aligned}$$

Combining both FOCs together yields the Euler equation

$$u'(C_t) = \beta (1+r-\delta) u'(C_{t+1}).$$

9.2 Example: Sovereign default (Prof. Sauré: Fin. Economics I)

Costs of Sovereign default – 2. Exclusion from financial markets – theory

Scenario 1 – no risk of default (very high default costs), the efficient benchmark:

The country can fully commitment to repay and solves

$$\max_{c_t, d_t} \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$

subject to

$$d_t = c_t - y_t + (1 + r)d_{t-1}$$

and

$$\lim_{T \rightarrow \infty} \frac{d_T}{(1 + r)^T} = 0 \quad (\text{No Ponzi condition})$$

9.3 Exercises

1. What is the optimal consumption rule for the Sovereign default model when utility is CRRA, i.e. $u(C_t) = \frac{c_t^{1-\sigma}-1}{1-\sigma}$? With $\sigma \geq 0$ and $\sigma \neq 1$.
2. Derive the optimal consumption rule for an individual. The objective function is given by the utility function of an individual,

$$U_t = \sum_{\tau=t}^{\infty} \beta^{\tau-t} u(c_{\tau}), \quad (9.1)$$

where

$$\beta \equiv (1 + \rho)^{-1}, \quad \rho > 0 \quad (9.2)$$

The budget constraint can be expressed in the intertemporal version by

$$\sum_{\tau=t}^{\infty} (1 + r)^{-(\tau-t)} e_{\tau} = a_t + \sum_{\tau=t}^{\infty} (1 + r)^{-(\tau-t)} w_{\tau}, \quad (9.3)$$

where $e_{\tau} = p_{\tau} c_{\tau}$.

3. What is the optimal consumption rule for the Sovereign default model when utility is exponential, i.e. $u(C_t) = \frac{e^{\alpha c_t}}{\alpha}$? With $\alpha > 0$. [additional exercise, beyond standard course content]

10 Continuous time problem

10.1 Setup

Continuous time problem

$$\max_{\{c(\tau)\}_{\tau=t}^{\infty}} U(t) = \int_t^{\infty} e^{-\rho[\tau-t]} u(c(\tau)) d\tau$$

subject to

$$\dot{a}(\tau) = r(\tau) a(\tau) + w(\tau) - c(\tau)$$

- This problem can be solved using the Hamiltonian approach
- This approach is related to the Lagrangian, but is specific to continuous time models.
- The steps are similar: set up your objective function, take first-order conditions, and solve for the optimal rule.

Setting up the Hamiltonian here gives us

$$\mathbb{H}(c(\tau), a(\tau); \lambda(\tau)) = u(c(\tau)) + \lambda(\tau) [r(\tau) a(\tau) + w(\tau) - c(\tau)].$$

- Thus, we maximise the instantaneous function of the control variable in blue with respect to the equation describing the evolution of the state variable in red.
- $\lambda(\tau)$ is the shadow price as before.
- The FOCs are now given by:

$$\begin{aligned} \frac{\partial \mathbb{H}}{\partial c} &= 0 \Leftrightarrow u'(c(\tau)) = \lambda(\tau) \\ \frac{\partial \mathbb{H}}{\partial a} &= \rho \lambda(\tau) - \dot{\lambda}(\tau) \Leftrightarrow \dot{\lambda}(\tau) = (\rho - r(\tau)) \lambda(\tau) \end{aligned}$$

- The first FOC looks at the optimal choice of the control variable such that the Hamiltonian does not change any more.
- The second FOC looks at the optimal change of the state variable w.r.t. the Hamiltonian such that it equates the difference between the flow value of the shadow price ($\rho \lambda$) and the change of the shadow price over time ($\dot{\lambda}$)
- Next, we take the derivative of the first FOC w.r.t. time, and then plug in the result and the first FOC itself into the second FOC, and rearrange.

10.2 Example: the central planner (Prof. Wälde: Advanced Macro)

- Preferences

- We study a central planner problem
- This allows us to focus on optimally chosen s in the simplest way (as opposed to decentralised economy)
- Objective function is social welfare function $U(t)$

$$U(t) = \int_t^{\infty} e^{-\rho[\tau-t]} u(C(\tau)) d\tau \quad (3.5)$$

where

- $u(C(\tau))$ is the instantaneous utility function,
- $\rho > 0$ is the time preference rate (measuring impatience of individual) and
- $e^{-\rho[\tau-t]}$ is the discount factor (or function)
- Planning/ optimal behaviour starts in t (like 'today') and goes up to infinity ∞

- Instantaneous utility

- Functional form ...

$$u(C(\tau)) = \frac{(C(\tau))^{1-\sigma} - 1}{1-\sigma}, \quad \sigma > 0 \quad (3.6)$$

- ... implies constant intertemporal elasticity of substitution of $1/\sigma$

- (Note that the logarithmic utility function is a special case of this functional form,

$$\lim_{\sigma \rightarrow 1} u(C(\tau)) = \ln C(\tau),$$

apply L'Hôpital's rule to find this)

- The resource constraint

- Maximization problem becomes meaningful only with a constraint

- Constraint here is a resource constraint as we look at the economy as a whole

- Capital evolves according to

$$\dot{K}(t) = Y(K(t), L) - \delta K(t) - C(t) \quad (3.7)$$

- The change in the capital stock (net investment) is given by output minus consumption (gross investment) minus depreciation $\delta K(t)$

3.14

10.3 The Shell Model

- The idea:

An economy needs resources for innovation and growth

Technological progress does not come costlessly (as in Solow growth model)

- A planner setup

Social welfare function reads

$$U(t) = \int_t^\infty e^{-\rho[\tau-t]} u(C(\tau)) d\tau$$

where utility from aggregate consumption is given by

$$u(C(\tau)) = \frac{C(\tau)^{1-\sigma} - 1}{1-\sigma}, \quad \sigma > 0$$

$$C(\tau) = A(\tau) [L - L_A(\tau)]$$

$$\frac{\dot{A}(\tau)}{A(\tau)} = L_A(\tau) \tag{10.1}$$

where $A(\tau)$ is labour productivity in $\tau \geq t$ and labour L is the only (fixed) factor of production. The number of workers in the research sector is given by $L_A(\tau)$.

We now have two versions of the maximization problems, using constraint 1 and using constraint 2. We solve the problem here using constraint 1, i.e. using consumption as a control variable.

Constraint 1:

$$C(\tau) = A(\tau) [L - L_A(\tau)] \Leftrightarrow \frac{C(\tau)}{A(\tau)} = L - L_A(\tau) \Leftrightarrow L_A(\tau) = L - \frac{C(\tau)}{A(\tau)}$$

we obtain as constraint

$$\frac{\dot{A}(\tau)}{A(\tau)} = L - \frac{C(\tau)}{A(\tau)} \Leftrightarrow \dot{A}(\tau) = A(\tau) L - C(\tau).$$

What is the Keynes-Ramsey rule?

Hamiltonian reads

$$H = \frac{C(\tau)^{1-\sigma} - 1}{1-\sigma} + \lambda [A(\tau) L - C(\tau)].$$

Two partial derivatives of the Hamiltonian

$$\begin{aligned} \frac{\partial H}{\partial C(\tau)} &= C(\tau)^{-\sigma} - \lambda = 0 \Leftrightarrow C(\tau)^{-\sigma} = \lambda \\ \frac{\partial H}{\partial A(\tau)} &= \lambda L \end{aligned} \tag{10.2}$$

Second optimality condition requires

$$\begin{aligned}\frac{\partial H}{\partial A(\tau)} &= \rho\lambda - \dot{\lambda} \Leftrightarrow \dot{\lambda} = \rho\lambda - \frac{\partial H}{\partial A(\tau)} = \rho\lambda - \lambda L \\ \Leftrightarrow \dot{\lambda} &= \rho\lambda - \lambda L = (\rho - L)\lambda\end{aligned}\tag{10.3}$$

Replace λ in second optimality condition

$$\dot{\lambda} = (\rho - L) C(\tau)^{-\sigma}$$

Compute time derivative of lambda from first-order condition

$$\dot{\lambda} = -\sigma C(\tau)^{-\sigma-1} \dot{C}(\tau)$$

and plug this into earlier equation

$$\begin{aligned}-\sigma C(\tau)^{-\sigma-1} \dot{C}(\tau) &= (\rho - L) C(\tau)^{-\sigma} \\ \Leftrightarrow -\sigma C(\tau)^{-1} \dot{C}(\tau) &= (\rho - L) \Leftrightarrow \frac{\dot{C}(\tau)}{C(\tau)} = \frac{L - \rho}{\sigma}\end{aligned}$$

10.4 Exercises

1. What is the optimal consumption rule (or Keynes-Ramsey rule), when production is Cobb Douglas, $Y(K(t), L) = K(t)^\alpha L^{1-\alpha}$? With $0 < \alpha < 1$.
2. What is the optimal consumption rule (or Keynes-Ramsey rule), when production is CES, $Y(K(t), L) = [\alpha K(t)^\theta + (1 - \alpha) L^\theta]^{\frac{1}{\theta}}$? With $\theta < 1$.
3. What is the optimal consumption rule (or Keynes-Ramsey rule) in the Shell-Model, when LA is the control variable?

Part VII

Ordinary Least Squares

11 Typical OLS regression

11.1 Preliminaries: method of moments for 1 explanatory variable

Typical OLS regression equation explains a (linear) relationship between a **dependent** variable y and an **independent** variable x (see “Intro Econometrics” by Prof. Dr. Van Ewijk + “Cross Section and Panel Data” by Prof. Dr. Schank for multivariate case)

$$y = \beta_0 + \beta_1 x + u,$$

where β_0 is the intercept, β_1 is the coefficient of interest in the regression, and u is the residual.

In turn, we estimate β_0 and β_1 by solving the following minimization problem, given a sample size n

$$\min_{\beta_0, \beta_1} \frac{1}{N} \sum_{i=1}^N \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$$

which yields the following FOC's

$$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0, \quad (\text{Condition 1})$$

$$\frac{1}{n} \sum_{i=1}^n x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0. \quad (\text{Condition 2})$$

We can rewrite the first expression as

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} \Leftrightarrow \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

where \bar{y} and \bar{x} are the sample averages of y and x . Thus, if we know $\hat{\beta}_1$, we can compute the estimate of the intercept (the sample averages being trivial to obtain).

To obtain this estimate of the slope, we go back to the second expression above, and replace $\hat{\beta}_0$ with our expression in terms of sample averages,

$$\sum_{i=1}^n x_i (y_i - [\bar{y} - \hat{\beta}_1 \bar{x}] - \hat{\beta}_1 x_i) = 0,$$

where we drop $\frac{1}{n}$, since the other side is equal to 0, and thus it does not affect the solution. We can then rewrite this in terms of deviations from the sample averages (i.e. deviations from the mean),

$$\sum_{i=1}^n x_i[y_i - \bar{y}] = \hat{\beta}_1 \sum_{i=1}^n x_i[x_i - \bar{x}].$$

Using properties of sums we can rewrite this as (see Wooldrige (2000) for details)

$$\sum_{i=1}^n [x_i - \bar{x}][y_i - \bar{y}] = \hat{\beta}_1 \sum_{i=1}^n [x_i - \bar{x}]^2 \Leftrightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^n [x_i - \bar{x}][y_i - \bar{y}]}{\sum_{i=1}^n [x_i - \bar{x}]^2}.$$

This last equality tells us that the estimate $\hat{\beta}_1$ is given by the covariance of x and y divided by the variance of x .

11.2 Example: Multiple Regression (Prof. van Ewijk: Intro Econometrics)

Recall: multiple regression

General format:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i, \quad i = 1, \dots, n$$

e.g. We estimate: $Health_i = \beta_0 + \beta_1 Exercise_i + \beta_2 Education_i + u_i$, in which:

- $Health$ = Some variable measuring health
 - $Exercise$ = Amount of exercise a person does
 - Variant A: measured in hours per week
 - Variant B: dummy variable indicating whether person does exercise yes/no
 - $Education$ = Number of years of education
 - Result:
$$\widehat{Health}_i = 5.4 + 2.00 * Exercise_i + 0.05 * Education_i$$

(0.9)

(0.50)

(0.20)
1. Interpret each of the 6 numbers – for variant A & B
 2. Which parameters are significant?
 3. I was interested *only* in the effect of exercise on health. Why would I have chosen to include the variable $Education$?

11.3 Exercises

Given the following samples, compute the estimates of the intercept $\hat{\beta}_0$ and of the slope $\hat{\beta}_1$.

1. The following table gives the number of units produced for a clone army y (in thousands) and the number of hours spent per day producing clones x over a 10-day period on Kamino.

y	5	8	15	12	5	8	2	10	12	17
x	1.5	2	2.5	2.5	2	2	1	1.5	2	2.5

Given your estimates, how many clones would be produced if the Kaminoans spent 4h/day producing them?

2. The following are final grades in an exam y and average number of hours spent studying per week x for 12 students.

y	51	45	44	43	40	35	33	24	23	21	16	2
x	10	12	10	6	9	9	9	7	5	6	4	2

Given your estimates, what grade would a student achieve if they spent 11h/week studying toward their exam?

Part VIII

Probability Theory

12 Distribution

12.1 Discrete random variable

We begin by defining the probability that a random variable X take on a specific value x as

$$f_X(x_i) := \text{Prob}(X = x_i) = p_i,$$

where we call $f_X(x_i)$ the probability mass function (PMF) of X . Correspondingly, we can define the cumulative probability that X take on any value less than or equal to x_i as

$$F_X(x_i) := \text{Prob}(X \leq x_i) = \sum_{x_i \leq x} f_X(x_i),$$

also called the cumulative distribution function (CDF).

12.2 Continuous random variable

Consider a continuous random variable X that can take any value in the interval $S = [a, b]$. The probability that X be smaller than or equal to some $x \in S$ is given by

$$F_X(x) := \text{Prob}(X \leq x),$$

which is its CDF. Correspondingly, we can compute the probability density function (PDF) of X as

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

Thus, the probability that X takes on a value between points c and d such that $a < c < d < b$ is given by

$$\text{Prob}(c < X < d) = F_X(d) - F_X(c) = \int_c^d f_X(x) dx.$$

Notation is often abbreviated to $f(x) := f_X(x)$, when there is no ambiguity about the domain of the variable.

12.3 Example: Distribution of discrete and continuous RV (Prof. Schank: Stats Methods)

Introduction

Probability

Inductive Statistics

Monte Carlo Simulation

Distribution of discrete random variables

- ▶ **Discrete random variable** X : takes only a finite number of different values $\{x_1, x_2, \dots, x_i, \dots, x_k\}$.
- ▶ Each of these values is realized with a certain probability $\{p_1, p_2, \dots, p_i, \dots, p_k\}$.
- ▶ The **probability mass function (PMF)** of the RV is defined for all real numbers x :

$$\begin{aligned} f_X(x) &= \Pr(X = x) \\ &= \begin{cases} p_i & \text{if } x = x_i \\ 0 & \text{if } x \notin \{x_1, \dots, x_k\} \end{cases} \end{aligned}$$

20 / 60

Distribution of discrete RV

- The **cumulative distribution function (CDF)** is

$$\begin{aligned} F_X(x) &= \Pr(X \leq x) \\ &= \sum_{x_i \leq x} p_X(x_i) \end{aligned}$$

Distribution of continuous RV

- ▶ **Continuous random variable X** : takes an infinite number of different values $x \in \mathbb{R}$ (e.g. body height)
- ▶ Each of these values x is realized with probability $\Pr(X = x) = 0$.
- ▶ The **cumulative distribution function (CDF)** is:

$$F_X(x) = \Pr(X \leq x)$$

- ▶ The **probability density function (PDF)** is:

$$f_X(x) = \frac{\partial F_X(x)}{\partial x}$$

- ▶ The probability for a number between a and b is

$$\Pr(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$$

12.4 Exercises

1. Using a fair 20-sided die, what is the probability that you will roll a 17? What is the probability that you will roll a 15 or higher?
2. Given a uniformly distributed variable X on the interval $[a, b]$ with CDF, $F(x) = \frac{x-a}{b-a}$, compute the PDF of this function. Additionally, what is the probability that $x = c$, where $c = \frac{b-a}{3}$? What is the probability that X is greater than c ?

13 Mean and variance

13.1 Discrete random variable

Consider a discrete random variable X that can take any value x_1, x_2, \dots, x_N with corresponding probabilities p_1, p_2, \dots, p_N . The expected value of the random variable X is

$$\mathbb{E}(X) \equiv \mu_X = \sum_{i=1}^N p_i x_i$$

When the number of realizations is infinitely large, we need to add up to ∞ , so N is replaced by ∞ .

The variance is defined as

$$\text{Var}(X) \equiv \sigma_X^2 = \sum_{i=1}^N p_i (x_i - \mu_X)^2$$

Where $(x_i - \mu_X)^2$ is the centered second moment. Note that if we were to define a new random variable Z with realizations $z_i = (x_i - \mu_X)^2$, then we would have $\text{Var}(X) = E(Z)$.

The standard deviation is the square root of the variance

$$\sigma_X = \sqrt{\sum_{i=1}^N p_i (x_i - \mu_X)^2}$$

13.2 Continuous random variable

Given Section 12, we consider a continuous random variable with density $f(x)$ on the support $[a, b]$

- the expected value of X as

$$\mathbb{E}(X) \equiv \int_a^b x f(x) dx,$$

- and the variance of X as

$$\text{Var}(X) \equiv \int_a^b (x - \mathbb{E}(x))^2 f(x) dx.$$

Of course, the standard deviation is also given by the square root of the variance,

$$\sigma_X = \sqrt{\text{Var}(X)}.$$

13.3 Example: Mean and variance of discrete and continuous RV (Prof. Schank: Stats Methods)

Important features of distributions

- **Expected value** (mean): measure of central tendency

$$\begin{aligned}\mu_X = E(X) &= \sum_{j=1}^k x_j \cdot f_X(x_j) \quad \text{if discrete} \\ &= \int_{-\infty}^{\infty} x \cdot f_X(x) \, dx \quad \text{if continuous}\end{aligned}$$

- **Variance**: measure of variability

$$\begin{aligned}\sigma_X^2 = \text{Var}(X) &= E((X - \mu_X)^2) \\ &= E(X^2) - \mu_X^2\end{aligned}$$

13.4 Exercises

1. Let $X \sim \log(k)$ (i.e. logarithmically distributed), compute the expected value of the discrete r.v. X (Note: the pmf of the log-distribution is $p_i(x_i) = \frac{-1}{\ln(1-k)} \frac{k^{x_i}}{x_i}$, where $x_i \in \{1, 2, 3, \dots\}$ is the support of the distribution and $0 < k < 1$ is a parameter).
2. Let $Y \sim \mathcal{N}(0, 1)$ (i.e. standard normally distributed), compute the expected value of the continuous r.v. y . (Note: the pdf of the standard normal distribution is given by: $f(y) = \frac{1}{\sqrt{2\pi}} e^{-(y^2)/2}$, and Y is defined on the entire real line).
3. Let $Z \sim \text{uniform}(0, 2)$ (i.e. uniformly distributed), compute the expected value AND the variance of the continuous r.v. Z . (Note: the pdf of this uniform distribution is given by: $f(z) = \frac{1}{2}$, and Z is defined on the set of real numbers from 0 to 2).