Probability Theory

Wu Yanzu

2023-12-02

Table of contents

Preface

This is a book for reviewing the class `Probability Theory`.

Class Matetial: the Professor Hong Yongmiao Class

Book: Probability and Statistics for Economists Yongmiao Hong, World Scientific, 2017

Video: Bilibili

1 Introduce

2 General methodology of modern ecnomic research

2.1 Data collection

- survey
- field studies
- experimental economics
- **Big data**

In most time,we get so-called stylized facts by summarizing from observed economic data.

Engel Curve,Phillips Curve,Okun's Law,volatility clustering in finance,etc.

So the empirical research is the first step of economic research which is so-called **economical intuition**.But for more further understanding,we need to build a model to explain the stylized facts by using the statistic tools.

2.2 Development of economic theories and models

With the emprical stylized facts,we can build a model to explain the facts and make predictions.This process is called **mathematical modeling of economic theroy**.

: An example is the Euler equation for rational expectations in macroeconomics.

Moreover, the objective of economic modeling is not merely to explain the stylized facts but also to understand the economic mechanism behind the facts.

2.3 Empriical validation/inference of economic theories and models

The last step is to test the model by using the observed data and make inference about the model.The key is to transform the model into a **textable empirical econometric model**.

2.4 Applications

After an econometric model passes the empirical evaluation, it can then be used to: - Explain important empirical stylized facts - Test economic theory and/or hypotheses - Forecast future evolution of the economy - Policy evaluation and other application

3 Roles of Econometrics

Modern economy is essentially built by upon the following three fundamental axioms:

- Any economy can be viewed as a stochastic process governed by some probability laws.
- Economic phenomenon, often summarized in form of data, can be reviewed as a realization of this stochastic data generating process.

So we establish the econometrics to infer the probability laws from the observed data that reflect the stochastic economic system, and then use the inferred probability laws for economic application e.g. to make predictions and conduct policy analysis.

4 Illustrative Examples

4.1 The simple keynesian consumption function Modle

For keynesian consumption function, we have the following model:

$$Y_t = C_t + I_t + G_t C_t = \alpha + \beta Y_t + \epsilon_t$$

Where Y_t is the aggregate income,C_t is the private consumption,I_t is the private investment,G_t is the government expenditure,ϵ_t is the random error term.

The parameters α and β can have appealing economic interpretations:

- α is the survival level consumption even if income is zero.
- β is the marginal propensity to consume (MPC), i.e. the amount of additional consumption for each additional unit of income.

Multiplier of income with respect to govenrnment spending

$$\frac{\partial Y_t}{\partial G_t} = \frac{1}{1-\beta}$$

which depends on the maginal propensity to consume β.So when assess the effect of fiscal policy, it is important to know the magnitude of β. ## Rational Expectations and Dynamic Asset Pricing Models

The rational expectations hypothesis (REH) is a key assumption in modern macroeconomics and finance. It states that agents' expectations about the future value of an economic variable are not systematically wrong. In other words, agents' expectations are not biased.

Suppose a representative agent has a constant relative risk aversion utility function:

$$U = \sum_{t=0}^{n} \beta^t u(C_t) = \sum_{t=0}^{n} \beta^t \frac{C_t^\gamma - 1}{\gamma}$$

where $\beta \geq 0$ is the discount factor and $\gamma \geq 0$ is the coefficient of relative risk aversion.$\mu(\cdot)$ is the agent's utility function in each time period.C_t is consumption in period t.

So, obviously the agent's optimization problem is: choosing a sequence of consumption $\{C_t\}_{t=0}^{\infty}$ to maximize the expected utility

$$\max_{\{C_t\}} E(u)$$

subject to the budget constraint:

$$C_t + P_t q_t \leq W_t + P_t q_{t-1}$$

where P_t is the price of the consumption good in period t, q_t is the quantity of the asset held at the end of period t, and W_t is the wage income in period t.

So we can define this marginal rate of intertemporal substitution (MRIS) as:

$$MRS_{t+1}(\theta) = \frac{\frac{\partial u(C_{t+1})}{\partial C_{t+1}}}{\frac{\partial u(C_t)}{\partial C_t}} = (\frac{C_{t+1}}{C_t})^{\gamma-1}$$

where model parameter vector $\theta = (\beta, \gamma)^{'}$.

So the First order condition is:

$$E[\beta MRS_{t+1}(\theta) R_{t+1} | I_t] = 1$$

where R_{t+1} is the gross return on the asset in period $t+1$ and I_t is the information set available at the beginning of period t.And the FOC is usually called the Euler equation.And we can estimate this model by using the generalized method of moments (GMM) method.

4.1.1 Production Function and Hypothsis on constant returns to scale

Production function

$$Y_t = \exp(\epsilon_t) F(L_t, K_t)$$

where Y_t is the output,L_t is the labor input,K_t is the capital input,ϵ_t is the random error term.

So we can get the constant return to scale hypothesis:

$$\lambda F(L_i, K_i) = F(\lambda L_i, \lambda K_i) \ for \ all \lambda > 0$$

CRS is a necessary condition for the existence of a long-run equilibrium in a competitive market.If CRS does not hold, and the technology displays the increasing returns to scale, then the market will lead to natural monopoly.

In practical, a conventional approach to estimate the production function is to use the Cobb-Douglas production function:

$$Y_i = F(L_i, k_i) = A \exp(\epsilon_i) L_i^{\alpha} K_i^{\beta} =$$

Then CRS becomes a mathematical restriction on the parameters α and β:

$$\mathbb{H}_0 : \alpha + \beta = 1$$

So if $\alpha + \beta > 1$, the production function displays increasing returns to scale; if $\alpha + \beta < 1$, the production function displays decreasing returns to scale.

In statistics, we can use the **F-test** to test the null hypothesis $\mathbb{H}_0$ (one dimensional restriction).

unfortunately, this test is not suitable for many cross-sectional economic data, which usually display conditional heteroskedasticity.One needs to use a robust, heteroskedasticity-consistent test procedure, such as the **White test**.

5 Roles of Probability and Statistics

6 Foundation of Probability Theory

6.1 Random Experiments

[**Definition 1. Random Experiment**] A random experiment is an experiment whose outcome is not known in advance.

there are two essential elements of a random experiment:

> **!** Important
>
> the set of all possible outcomes
> the likelihood of each outcome

we need to know the sample space and the probability of each outcome in the sample space.

The purpose of mathematical statistics is to provide mathematical models for random experiments ofinterest.

Once a model for such an experiment is provided and the theory worked out in detail, the statistician may,within this framework, make inference about the probability law of the random experiment.

6.2 Basic Concepts of Probability

[**Definition 2. Sample Space**] The possible outconmes of a random experiment are called the sample space and is denoted by S.

When an experiment is performed, the realization of the experiment will be one **(and only one)** outcome in thesample space.(Mutually exclusive)

If the experiment is performed a number of times, a different outcome may occur each time or some outcomes may repeat.

a sample space S can be countable or uncountable.

[**Definition 3. Event**] An event A is a collection of basic outcomes from the sample space S that share certain common features or equivalently obey certain restrictions.

The event is the **subset** of the sample space.

! Important

- *Basicoutconme* $\subseteq$ *Event* $\subseteq$ *Sample space*

6.3 Review of Set Theory

Use Venn Diagram to represent the relationship between sets(or sample point,event, all the related concepts).

[**Definition 4. Intersection**] Intersection of A and B, denoted $A \cap B$ is the set of basic outcomes in S that belong to both A and B.The intersection of A and B is the event that both A and B occur. is also called the **logicall product** of A and B.

[**Definition 5. Exclusiveness**] If A and B have no common basic outcomes, they are called mutually exclusive and their intersection is empty set $\emptyset$, i.e., $A \cap B = \emptyset$ where $\emptyset$ denotes an empty set that contains nothing.

- mutually exclusive events are also called disjoint events.

[**Definition 6. Union**] The union of A and B, denoted $A \cup B$, is the set of basic outcomes in S that belong to either A or B or both. The union of A and B is the event that either A or B or both occur. is also called the **logical sum** of A and B.

[**Definition 7. Collective Exhaustiveness**] Suppose that $A_1, A_2, A_3, \cdots$ are events in S. If $A_1 \cup A_2 \cup A_3 \cup \cdots = S$, then the events $A_1, A_2, A_3, \cdots$ are collectively exhaustive.

[**Definition 8. Complement**] The complement of A, denoted A^c, is the set of basic outcomes in S that do not belong to A. The complement of A is the event that A does not occur.

- $A \cap A^c = \emptyset$ *and* $A \cup A^c = S$

[**Definition 9. Difference**] The difference of A and B, denoted $A - B = A \cap B^c$, is the set of basic outcomes in S that belong to A but not to B. The difference of A and B is the event that A occurs but B does not occur.

Distributivity Laws :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

In more general, we have

$$B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (B \cap A_i) B \cup \left(\bigcap_{i=1}^{n} A_i \right) = \bigcap_{i=1}^{n} (B \cup A_i)$$

De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c \, (A \cap B)^c = A^c \cup B^c$$

In more general, we have

$$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c \left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c$$

: Suppose the events A and B are disjoint.Under what condition are A^c and B^c also disjoint?

: A^c and B^c are disjoint if and only if $A \cup B = S$.That is say A and B is exhaustive.

: • Are $A \cap B$ and $A^c \cap B$ mutually exclusive?

- Is $(A \cap B) \cup (A^c \cap B) = B$?
- Are A and $A^c \cap B$ mutually exclusive?
- Is $A \cup (A^c \cap B) = A \cap B$?

Yes, Yes, Yes, No

: Let the set of events $\{A_i = 1, \dots, n\}$ be mutually exclusive and collectively exhaustive, and let A be an event in S - Are $A_1 \cap A, \dots, A_n \cap A$ mutually exclusive? - Is the union of $A_i \cap A, i = 1, \dots, n$, equal to A? That is, do we have:

$$\bigcup_{i=1}^{n} (A_i \cap A) = A$$

: Yes, Yes

> In the linear algebra, we can use the projection matrix to understand this.The $\{A_1, A_2, \dots, A_n\}$ is the orthogonal bases of the specific space, and the A is the vector in this space. The $A_i \cap A$ is the projection of A on the A_i direction.
>
> - A sequence of collective and mutually exclusive events forms a partition of sample space S.
> - A set of collectively exhaustive and mutually exclusive events can be viewed as a complete set of orthogonal bases.
> - The projection of a vector on a subspace is the sum of the projections of the vector on the orthogonal bases of the subspace.

- A complete set of orthogonal bases can represent any event A in the sample space S, and could be viewed as the projection of event A on the base A_i.

6.4 Fundamental Probability Laws

To assign a probability to an event $A \in S$, we shall introduce a **probability function**, which is a function or a mapping from an event to a real number (0,1).

To assign probabilities to events, complements of events,unions and intersections of events, we want our collection of events to include all these combinations of events.

Such a collection of events is called an **σ-field(σ algaebra)** of subsets of the sample space S,which constitude the domain of the probability function.

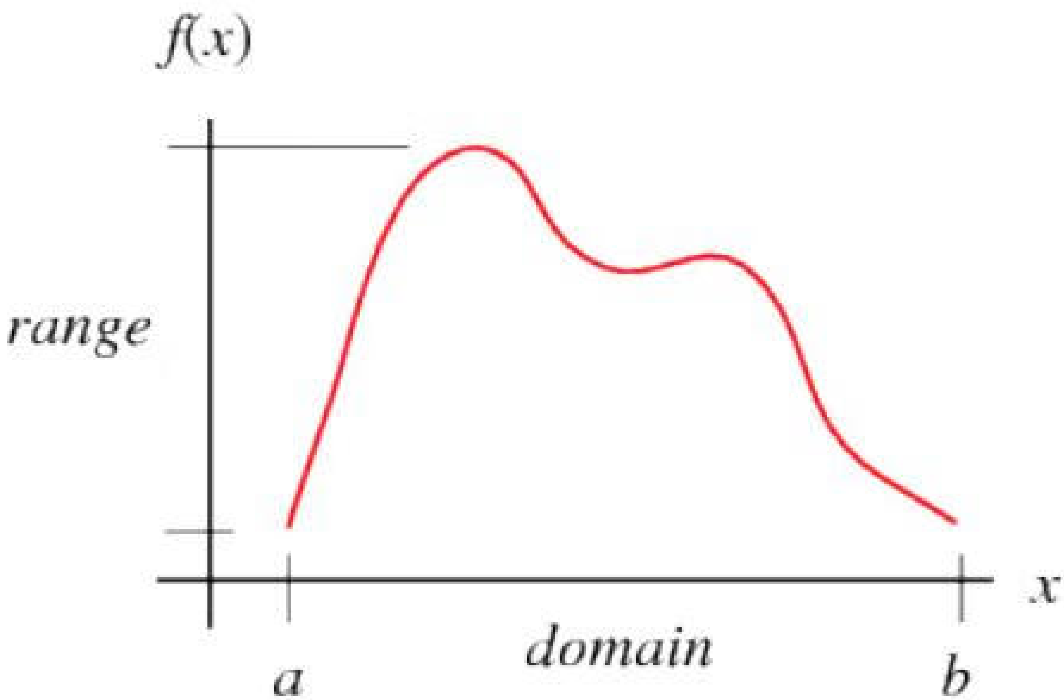

Figure 6.1: image.png

[**Definition 10. Sigma Algebra**] A $sigma(\sigma)algebra$,denoted by $\mathbb{B}$, is a collection of subsets(events) of S that satisfies the following three conditions:

1. $\emptyset \in \mathbb{B}$i.e., the empty set is in $\mathbb{B}$.
2. If $A \in \mathbb{B}$, then $A^c \in \mathbb{B}$.(i.e., $\mathbb{B}$ is closed under complements)
3. If $A_1, A_2, \ldots \in \mathbb{B}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathbb{B}$.(i.e., $\mathbb{B}$ is closed under countable unions)

! Important

- σ-algebra is a collection of events in S(subset) that satisfies certain properties and constitutes the domain of a probability function.
- A σ -field is a collection of subsets in S , but itself is not a subset of S . In contrast, the sample space S is **only an element of a σ -field**.
- The pair $(S, \mathbb{B})$ is called a measurable space.So for a specific sample space S, we can have different σ-algebra $\mathbb{B}$.

[**Definition 11. Probability Function**] Suppose a random experiment has a sample space S and an associated σ -field $\mathbb{B}$. A probability function

$$P : \mathbb{B} \to [0, 1]$$

is defined as a mapping that satisfies the following three conditions:

(1) $0 \leq P(A) \leq 1 for all A \in \mathbb{B}$.

(2) $P(S) = 1$.

(3) If $A_1, A_2, \ldots \in \mathbb{B}$ are mutually exclusive, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$. The probability of the union of mutually exclusive events is the sum of the probabilities of the individual events.

! Important

A probability function tell how the probability of occurrence is distributed over the set of events $\mathbb{B}$.In this sense we speak of a distribution of probability.

6.5 Methods of Counting

6.6 Conditional Probability

6.7 Bayes' Theorem

6.8 Independence

6.9 Conclusion

7 Summary

In summary, this book has no content whatsoever.

References