

Phương Pháp Phân Tích Định Lượng 1: Học Phần Xác Suất – Thống Kê

Giới Thiệu Giảng Viên

- Giảng viên phụ trách Đỗ Hoàng Phương.
- Quá trình đào tạo: cử nhân ĐH Ngoại Thương, Thạc sĩ tại University of Trento (Italy), and PhD in Economics tại Virginia Tech (USA).
- Lĩnh vực nghiên cứu chính: kinh tế học vĩ mô, chính sách công nghiệp và kinh tế chính trị Việt Nam.
- Địa chỉ email: hoangphuong.do@fulbright.edu.vn

Giờ Tiếp Sinh Viên

- Đỗ Hoàng Phương:
 - Thứ Sáu 15:00 – 17:00 hoặc email để hẹn.
- Lê Khánh Hưng:
 - Thứ Ba 13:30 - 15:00
 - Thứ Sáu 15:30 - 17:00 hoặc email để hẹn.

Mục Tiêu Môn Học

- Mục tiêu: đọc, hiểu nghiên cứu định lượng về chính sách công; phát triển khả năng thực hiện nghiên cứu độc lập.
- Sau khóa học, học viên được kỳ vọng có khả năng:
 - Đọc hiểu các đại lượng thống kê thông dụng;
 - Phân tích và phản biện các số liệu báo cáo định lượng trình bày trên báo chí, tài liệu nghiên cứu, hội thảo;
 - Nhận định được vai trò của mô hình định lượng trong phân tích kinh tế;
 - Từng bước xây dựng các mô hình định lượng từ đơn giản đến phức tạp trong các điều kiện giả định khác nhau với các phần mềm thông dụng.
- Nửa phần đầu sẽ tập trung vào kiến thức nền tảng để tìm hiểu các mô hình phân tích định lượng ở phần 2.

Phát Họa Nội Dung

Nội dung	Tuần
Lý thuyết về xác suất.	1 – 2
Giới thiệu về môn học thống kê. Thống kê mô tả.	2
Thống kê suy luận: - Ước lượng điểm. - Ước lượng khoảng. - Giả thuyết thống kê.	3 – 4
Phân tích ANOVA. Sơ lược về thiết kế thí nghiệm.	4 – 5
Một số vấn đề thực hành.	5

Phương Pháp Đánh Giá

- Bài tập hàng tuần phát vào thứ 5 hàng tuần và nộp vào sáng thứ 3 tuần tiếp theo (40%).
- Bài giữa kì sẽ phát vào ngày 05/12/2023 (40%).
- Câu hỏi quiz có thể diễn ra bất ngờ tùy vào diễn tiến môn học (10%).
- Tham gia trả lời (10%).

Liêm Chính Học Thuật

- Liêm chính học thuật (academic integrity) được tuân thủ nghiêm ngặt tại Đại học Fulbright Việt Nam.
- Các bạn sinh viên nếu không hiểu rõ các quy định có thể liên hệ giảng viên, trợ giảng hoặc bộ phận đào tạo để trả lời cho những thắc mắc cụ thể.
- Các vi phạm sẽ bị xử lý theo quy định.

Sách Tham Khảo

- **Basic Business Statistics: Concepts and Applications** (14th edition) bởi Mark L. Berenson, David M. Levine, Kathryn A. Szabat và David F. Stephan viết tắt **BLSS14**.
- **Probability and Statistics** (4th edition) bởi Morris H. DeGroot và Mark J. Schervish viết tắt **DS4**.

Lý Thuyết Xác Suất (1)

Giới Thiệu

- Tại sao chúng ta cần hiểu lý thuyết xác suất và thống kê (probability and statistics theory) trong mối quan hệ với phương pháp phân tích định lượng (quantitative analysis)?
- Thống kê là ngành nghiên cứu liên quan đến việc thu thập, tổ chức, đánh giá, phân tích và trình bày dữ liệu.
 - Thống kê mô tả (descriptive statistics).
 - Thống kê suy luận (inferential statistics).
- Lý thuyết toán học về xác suất (mathematical probability theory) là nền tảng trọng yếu để học phương pháp về thống kê suy luận cũng như các mô hình định lượng có tính nhân quả.

Khái Quát Nội Dung

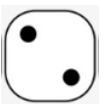
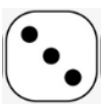
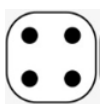
- Các khái niệm cơ bản để hiểu về xác suất (probability).
- Quy luật tính toán xác suất từ phép thử (experiment) cho biến cố (event).
- Xác suất có điều kiện và định lý Bayes.
- Biến cố độc lập.

Các Khái Niệm Căn Bản

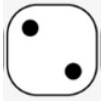
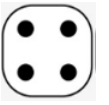
- **Phép thử (experiment):** quá trình (thực hay tưởng tượng) được thực hiện có thể lặp lại nhiều lần, kết quả biết được trước.
- **Biến cố đơn giản (simple event):** tập hợp (a set of) các kết quả đơn giản từ một phép thử.
- **Biến cố (event):** tập hợp của một hay nhiều biến cố đơn giản.
- **Không gian mẫu (sample space):** tập hợp tất các biến cố (đơn giản) có thể xảy ra khi thực hiện phép thử.

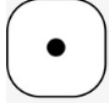
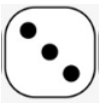
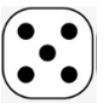
Ví Dụ 1: Tung Xí Ngẫu

- Phép thử: hành động tung xí ngẫu.

- Biến cố đơn giản:  ,  ,  ,  ,  , 

- Biến cố (ký hiệu A):

- Số chẵn diễn ra: $A_C = \{2, 4, 6\}$  +  + 

- Số lẻ diễn ra: $A_L = \{1, 3, 5\}$  +  + 

- Không gian mẫu (ký hiệu S): $S = \{1, 2, 3, 4, 5, 6\}$

$$\begin{array}{|c|} \hline \bullet \\ \hline \end{array} + \begin{array}{|c|} \hline \bullet \quad \bullet \\ \hline \end{array} + \begin{array}{|c|} \hline \bullet \quad \bullet \quad \bullet \\ \hline \end{array} + \begin{array}{|c|} \hline \bullet \quad \bullet \quad \bullet \quad \bullet \\ \hline \end{array} + \begin{array}{|c|} \hline \bullet \quad \bullet \quad \bullet \quad \bullet \quad \bullet \\ \hline \end{array} + \begin{array}{|c|} \hline \bullet \quad \bullet \quad \bullet \quad \bullet \quad \bullet \quad \bullet \\ \hline \end{array}$$

Ví Dụ 2:

- Phép thử: rút 1 lá bài từ bộ bài Tây 52 lá.
 - Biến cố đơn giản?
 - Không gian mẫu?



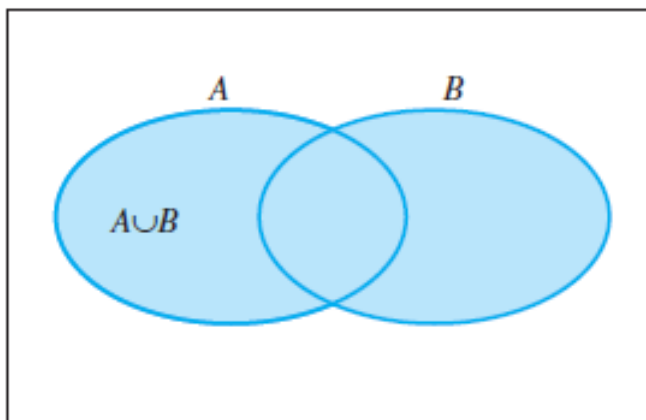
Định Nghĩa Toán Học Về Xác Suất

- **Biến cố ngẫu nhiên** có đặc điểm chung là chúng không thể dự đoán. Tuy nhiên, chúng ta có lượng hóa khả năng xảy ra của chúng bằng một con số **xác suất**:

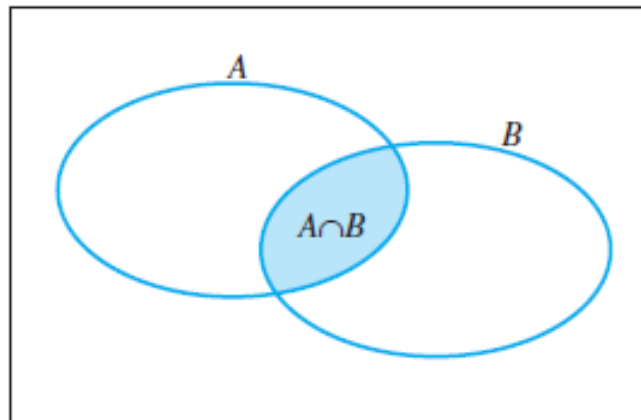
$$P(A) \text{ hoặc } \Pr(A)$$

- Ta quy ước:
 - $P(S) = 1$: Biến cố chắc chắn.
 - $P(A) \geq 0$: Biến cố ngẫu nhiên.
 - $\sum_{i=1}^n P(A_i) = 1$: Tổng của tất cả các xuất phải bằng 1

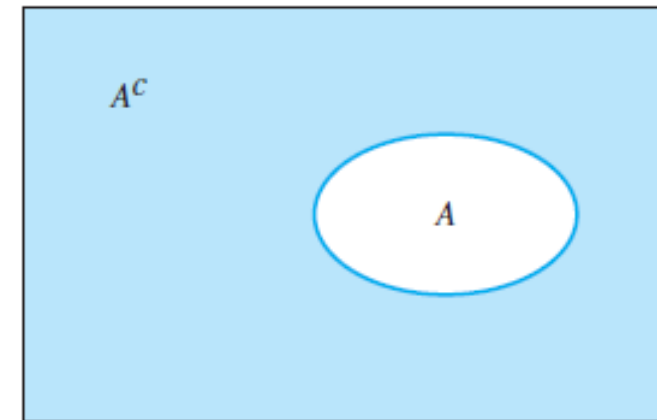
Lý Thuyết Tập Hợp Và Xác Suất



Hợp của A và B: hoặc là A hoặc là B diễn ra



Giao của A và B: cả A và B đều diễn ra



Phần bù của A: A không diễn ra

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Quy luật xác suất cho phép hợp

$$P(A \cap B) = P(A)P(B|A)$$

$$P(A \cap B) = P(B)P(A|B)$$

Quy luật xác suất cho phép giao

$$P(A^c) = 1 - P(A)$$

$$A \cup A^c = S$$

Quy luật xác suất phần bù

Ví Dụ 3:

- Tung một đồng xu hai lần liên tiếp. Hãy tính:
 - Xác suất quan sát ít nhất 1 sấp.
 - Xác suất 2 mặt khác nhau.



Ví Dụ 3: Bài Giải

- Biến cố đơn giản (ngửa N, Sấp S):

$$P(SS) = P(SN) = P(NS) = P(NN) = 1/4$$

- Biến cố:

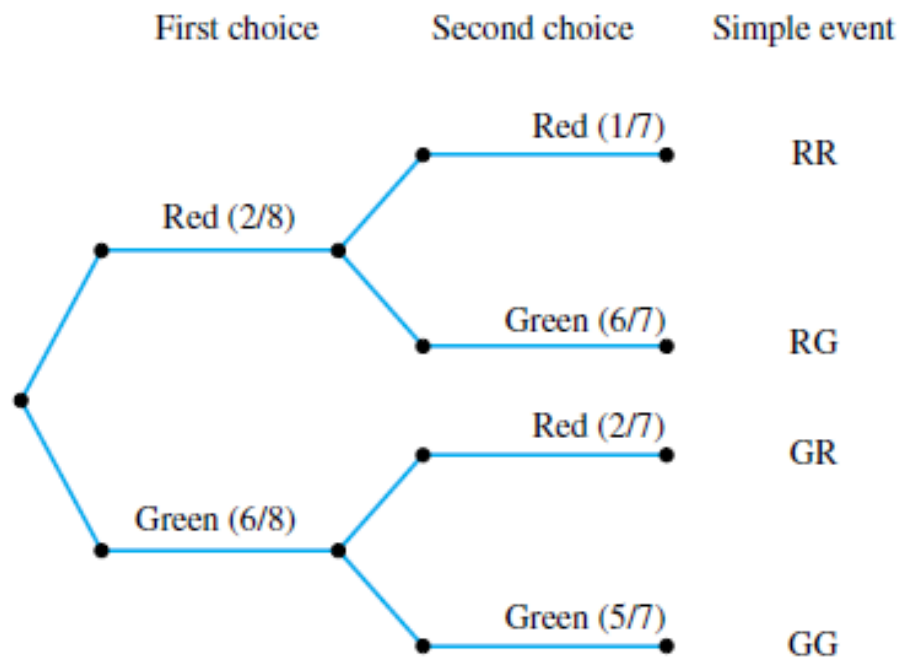
$$A = \{SS, SN, NS\}: P(A) = \frac{3}{4}$$

$$B = \{SN, NS\}: P(B) = \frac{2}{4}$$

- $A \cup B = \{SS, SN, NS\}: P(A \cup B) = 3/4$
- $A \cap B = \{SN, NS\}: P(A \cap B) = 1/2$
- $A^c = \{NN\}, P(A^c) = 1/4$

Ví Dụ 3:

- Có tám quả bóng trong hộp, bao gồm 2 đỏ và 6 xanh. Lấy ngẫu nhiên 2 quả.
- Biến cố A : xác suất lấy ngẫu nhiên 2 quả đỏ?



$$P(A) = P(\text{Đỏ lần 1})P(\text{Đỏ lần 2} \mid \text{Đỏ lần 1 đã xảy ra})$$

$$P(A) = \left(\frac{2}{8}\right)\left(\frac{1}{7}\right) = 1/28$$

Ví Dụ 4:

- Cho một nhóm người nam và nữ ở ba độ tuổi khác nhau như trong bảng. Chọn một người ngẫu nhiên từ nhóm này. Hãy tính:
 - Xác suất người từ 50 tuổi đổ xuống (biến cố A).
 - Xác suất người nam (biến cố B).
 - Xác suất người nữ (biến cố C).
 - Xác suất người từ 50 tuổi đổ xuống và là nữ ($A \cap B$).

	Nam (A)	Nữ (A_N)	
Dưới 16 (B_1)	30	18	48
17-50 tuổi (B_2)	50	55	105
Trên 50 tuổi (B_3)	20	27	47
	100	100	200

Xác Suất Có Điều Kiện

- **Xác suất có điều kiện (conditional probability)** của biến B khi biến cố A đã xảy ra, kí hiệu:

$$P(B | A)$$

- Ví dụ: Rút hai lá bài liên tiếp từ bộ bài 52 lá, lần một chúng ta rút được 1 lá ách. Câu hỏi: xác suất rút tiếp được lá ách ở lần hai là bao nhiêu?

$$P(\text{ách lần 2} | \text{ách lần 1}) = ?$$

- Công thức tính xác suất có điều kiện:

$$P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A | B) P(B)}{P(A)} \text{ khi } P(A) \neq 0$$

Biến Cố Xung Khắc

- **Biến cố xung khắc/ loại trừ (mutual exclusive events)** là các biến cố không thể xảy ra đồng thời với nhau.
 - Ví dụ: tung 1 đồng xu thì biến cố sấp và ngửa không thể xảy ra đồng thời.
- Nếu A và B xung khắc:
 - $P(A \cap B) = 0,$
 - $P(A \cup B) = P(A) + P(B)$

Định Lý Bayes

- Định lý Bayes (Bayes' rule) về xác suất có điều kiện:

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^k P(B_j)P(A|B_j)} \text{ for } i = 1, 2, \dots, k$$

với $B_1, B_2, \dots, B_j$ là xung khắc với nhau từng đôi một và $\sum_{j=1}^k P(B_j) = 1$

Ví Dụ 5:

- Một hộp ít có 2 đồng xu: một đồng xu 2 mặt xấp/ngửa và một đồng xu chỉ có mặt xấp. Bốc một đồng xu bất kì và thả, thu được mặt xấp. Hỏi kết quả xác suất bốc đồng xu 2 mặt là bao nhiêu?
- Gọi B_1 là biến cố bốc đồng 2 mặt xấp/ngửa, B_2 là chỉ mặt xấp, H_1 là thu được được mặt xấp.
- Theo định lý Bayes:

$$\begin{aligned} P(B_1|H_1) &= \frac{P(B_1) \cdot P(H_1|B_1)}{P(B_1) \cdot P(H_1|B_1) + P(B_2)P(H_1|B_2)} \\ &= \frac{(1/2) (1/2)}{(1/2) (1/2) + (1/2) (1)} = \frac{1}{3} \end{aligned}$$

Ví Dụ 6:

- Tính xác suất một người 16 tuổi (B_1) biết rằng người đó là nam (A)?

$$P(B_1 | A) = \frac{P(B_1 \cap A)}{P(A)} = ?$$

$$\begin{aligned} P(B_1 | A) &= \frac{P(B_1)P(A|B_1)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)} \\ &= \frac{0.24 \cdot (30/48)}{0.24 \cdot (30/48) + 0.525 \cdot (50/105) + 0.235 \cdot (20/47)} = 0.3 \end{aligned}$$

	Nam (A)	Nữ (A_N)		Nam (A)	Nữ (A_N)	
Dưới 16 tuổi (B_1)	30	18	48	Dưới 16 tuổi (B_1)	0.15	0.09
17-55 tuổi (B_2)	50	55	105	17-55 tuổi (B_2)	0.25	0.275
Trên 55 tuổi (B_3)	20	27	47	Trên 55 tuổi (B_3)	0.10	0.135
	100	100	200		0.5	0.5
						1.00

Biến Cố Độc Lập

- **Biến cố độc lập** (independent events) là các biến cố xác suất có điều của biến này không phụ thuộc vào biến kia:

$$P(A|B) = P(A)$$

- Ví dụ: tung 2 xúc sắc đồng thời và xác suất xảy giá trị ở riêng từng xúc sắc là độc lập.
- Khi A và B là hai biến cố độc lập thì ta có các tính chất sau:
 - $P(A \cap B) = P(A)P(B)$
 - $P(A \cup B) = P(A) + P(B) - P(A)P(B)$
 - $P(A \cap B \cap C) = P(A)P(B)P(C)$

Kiểm Tra Tính Độc Lập (1)

- Tung một xúc sắc và nghiên cứu hai biến cố:
 - Biến cố C: xảy ra một số chẵn.
 - Biến cố B: xảy ra một số chia hết cho 3.
- Hai biến cố C và B có độc lập hay không?
- Cách trả lời:
 - Tìm xác suất của $P(C)$.
 - Tìm xác suất của $P(B)$.
 - Tìm xác suất $P(C \cap B)$.
 - So sánh hai xác suất $P(C) \times P(B)$ và $P(C \cap B)$.

Kiểm Tra Tính Độc Lập (2)

- Tung một xúc sắc và nghiên cứu hai biến cố:
 - Biến cố C: xảy ra một số chẵn.
 - Biến cố A: xảy ra 1 số con trong khoảng từ 1 đến 4.
- Hai biến cố A và C có độc lập hay không?

Cách Hiểu Về Xác Suất Trong Thực Tế (Interpretation of Probability)

- **Frequentist interpretation – cách hiểu “tần suất”**: thực hiện một thí nghiệm trong điều kiện hoàn toàn tương đồng. Xác suất là tỷ lệ tần xuất hiện tương đối – k lần của n thí nghiệm:

$$P(A) = \frac{k}{n}$$

- **Bayesian (or subjective) interpretation – cách hiểu “chủ quan”**: xác suất là sự tự nhận định khả năng (likelihood) xảy ra chủ quan của từng người dựa vào niềm tin và thông tin cá nhân của họ. Cách hiểu này phù hợp với sự kiện xảy ra ít lặp lại.
 - Ví dụ: xác suất bầu cử tổng thống.

Tổng Kết

- Phép thử, biến cố đơn giản, biến cố.
- Xác suất cho biến cố.
- Xác suất có điều kiện, định lý Bayes, xác suất độc lập.