***CÁC PHƯƠNG PHÁP MÔ PHỎNG SỐ MONTE-CARLO***

***Computer Simulation Methods Monte-Carlo***

***--------------------------------------------------------------------------------------------------------------------***

***1.1 Khái niệm về phương pháp Monte Carlo***

* *Các****phương pháp Monte Carlo****là một lớp các*[*thuật toán*](https://vi.wikipedia.org/wiki/Thu%E1%BA%ADt_to%C3%A1n)*để giải quyết nhiều bài toán trên*[*máy tính*](https://vi.wikipedia.org/wiki/M%C3%A1y_t%C3%ADnh)*theo kiểu*[*không tất định*](https://vi.wikipedia.org/w/index.php?title=Thu%E1%BA%ADt_to%C3%A1n_kh%C3%B4ng_t%E1%BA%A5t_%C4%91%E1%BB%8Bnh&action=edit&redlink=1)*, thường bằng cách sử dụng các*[*số ngẫu nhiên*](https://vi.wikipedia.org/w/index.php?title=S%E1%BB%91_ng%E1%BA%ABu_nhi%C3%AAn&action=edit&redlink=1)*(thường là các*[*số giả ngẫu nhiên*](https://vi.wikipedia.org/w/index.php?title=S%E1%BB%91_gi%E1%BA%A3_ng%E1%BA%ABu_nhi%C3%AAn&action=edit&redlink=1)*).*

*Các phương pháp này đặc biệt hiệu quả khi giải quyết các*[*phương trình vi-tích phân*](https://vi.wikipedia.org/w/index.php?title=Ph%C6%B0%C6%A1ng_tr%C3%ACnh_vi-t%C3%ADch_ph%C3%A2n&action=edit&redlink=1)*; ví dụ như trong mô tả*[*trường bức xạ*](https://vi.wikipedia.org/w/index.php?title=Tr%C6%B0%E1%BB%9Dng_b%E1%BB%A9c_x%E1%BA%A1&action=edit&redlink=1)*hay*[*trường ánh sáng*](https://vi.wikipedia.org/w/index.php?title=Tr%C6%B0%E1%BB%9Dng_%C3%A1nh_s%C3%A1ng&action=edit&redlink=1)*trong mô phỏng hình ảnh 3 chiều trên máy tính, có ứng dụng trong*[*trò chơi điện tử*](https://vi.wikipedia.org/wiki/Tr%C3%B2_ch%C6%A1i_%C4%91i%E1%BB%87n_t%E1%BB%AD)*,*[*kiến trúc*](https://vi.wikipedia.org/wiki/Ki%E1%BA%BFn_tr%C3%BAc)*,*[*thiết kế*](https://vi.wikipedia.org/wiki/Thi%E1%BA%BFt_k%E1%BA%BF)*,*[*phim*](https://vi.wikipedia.org/wiki/Phim_(%C4%91%E1%BB%8Bnh_h%C6%B0%E1%BB%9Bng))*tạo từ máy tính, các hiệu ứng đặc biệt trong điện ảnh, hay trong nghiên cứu*[*khí quyển*](https://vi.wikipedia.org/wiki/Kh%C3%AD_quy%E1%BB%83n)*, và các ứng dụng nghiên cứu vật liệu bằng*[*laser*](https://vi.wikipedia.org/wiki/Laser)*...*

* *Trong*[*toán học*](https://vi.wikipedia.org/wiki/To%C3%A1n_h%E1%BB%8Dc)*,****thuật toán Monte Carlo****là phương pháp tính bằng số hiệu quả cho nhiều bài toán liên quan đến nhiều biến số mà không dễ dàng giải được bằng các phương pháp khác, chẳng hạn bằng*[*tính tích phân*](https://vi.wikipedia.org/wiki/T%C3%ADch_ph%C3%A2n)*.*

*Nhiều khi, phương pháp Monte Carlo được thực hiện hiệu quả hơn với*[*số giả ngẫu nhiên*](https://vi.wikipedia.org/w/index.php?title=S%E1%BB%91_gi%E1%BA%A3_ng%E1%BA%ABu_nhi%C3%AAn&action=edit&redlink=1)*, thay cho số ngẫu nhiên thực thụ, vốn rất khó tạo ra được bởi máy tính. Các số giả ngẫu nhiên có tính tất định, tạo ra từ chuỗi giả ngẫu nhiên có quy luật, có thể sử dụng để chạy thử, hoặc chạy lại mô phỏng theo cùng điều kiện như trước. Các số giả ngẫu nhiên trong các*[*mô phỏng*](https://vi.wikipedia.org/wiki/M%C3%B4_ph%E1%BB%8Fng)*chỉ cần tỏ ra "đủ mức ngẫu nhiên", nghĩa là chúng theo*[*phân bố đều*](https://vi.wikipedia.org/wiki/Ph%C3%A2n_b%E1%BB%91_ng%E1%BA%ABu_nhi%C3%AAn_%C4%91%E1%BB%81u)*hay theo một phân bố định trước, khi số lượng của chúng lớn.*

*Phương pháp Monte Carlo thường thực hiện lặp lại một số lượng rất lớn các bước đơn giản, song song với nhau; một phương pháp phù hợp cho*[*máy tính*](https://vi.wikipedia.org/wiki/M%C3%A1y_t%C3%ADnh)*. Kết quả của phương pháp này càng chính xác (tiệm cận về kết quả đúng) khi số lượng lặp các bước tăng.*

***1.2 Các ví dụ***

***Ví dụ 1. Tính tích phân xác định***

*Ta biết rằng nếu f(x) là hàm không âm tích phân xác định*

*(1.1)*

*sẽ cho ta diện tích phần hình học(S) giới hạn bởi các đường*

*trong mặt phẳng XOY.*

*Giả sử R là hình chữ nhật có cạnh đáy là khoảng (a,b) và chiều cao c, Nếu M là một điểm được ném ngẫu nhiên lên R (tức là các toa độ x và y độc lập và có phân phối đều tương ứng trên đoạn [a,b] và [0,c]. Khi đó xác suất P để điểm M rơi vào miền S sẽ bằng tỷ số giữa diện tích S và diện tích hình chữ nhật*

*(1.2)*

*Nếu ta ném N điểm ngẫu nhiên lên hình chữ nhật, trong đó có n điểm rơi vào miền S, khi đó ta sẽ có*

*Vì vậy S có thể được đánh giá bởi việc tạo ra N cặp độc lập các số ngẫu nhiên*

*, với có phân phối đều trên (a,b) còn có phân phối đều trên (0,c).*

*Tính số n cặp thỏa mãn ta lấy làm giá trị xấp xỉ của S.*

*Viết lại (1.1) dưới dạng*

*(1.3)*

*Trong đó là một hàm mật độ xác suất bất kỳ trên (a,b), nghĩa là nó phải thỏa*

*Từ (1.3) suy ra rằng nếu X là biến ngẫu nhiên với hàm mật độ và Y là biến ngẫu nhiên xác định bởi: thì .*

*Vì vậy một xấp xỉ (hay ta thường gọi là một ước lượng) của S có thể nhận được bởi N số ngẫu nhiên độc lập từ phân phối có mật độ và lấy trung bình*

*Trong trường hợp trên, mật độ của phân phối đều được chọn cho xấp xỉ cần tìm trở thành*

*ở đây các số ngẫu nhiên độc lập và có phân phối đều trên (a.b).*

***Ví dụ 2. Tính giá trị số thông qua bài toán gieo kim Buffon***

*Trên mặt phẳng ta kẻ các đường thẳng song song cách đều nhau . Ta gieo ngẫu*

*nhiên lên mặt phẳng một cây kim có độ dài 2, với Tìm xác suất để cây kim cắt*

*một trong những đường thẳng song song đã kẻ.*

*Mọi vị trí của cây kim đều xác định bởi khoảng cách Y từ tâm của kim tới đường thẳng gần nhất và góc giữa cây kim và các đường thẳng song song. Ta thấy rằng Y và có*

*các liên hệ: Ngoài ra việc tung ngẫu nhiên có thể hiểu rằng Y và độc và có phân phối đều trên các khoảng tương ứng nêu trên.*

*Cây kim cắt một trong các đường thẳng nếu*

*.*

*Từ đó suy ra xác suất để cây kim cắt một trong các đường thẳng bằng:*

*Giả sử cây kim được tung N lần, trong đó có n lần cây kim cắt một đường thẳng*

*theo luật mạnh số lớn ta sẽ có*

*(h.c.c)*

*Suy ra xấp xỉ của số có thể tính theo thuật giải sau:*

1. *Chọn các giá trị và*
2. *Tạo N cặp độc lập các số ngẫu nhiên*

*Trong đó có phân phối đều trên đoạn (0,a); có phân phối đều trên đoạn* (*0,. Tính tỷ số các cặp thỏa mãn*

1. *Chọn tỷ số là một xấp xỉ của P*
2. *Đánh giá xấp xỉ, nếu cần thiết sẽ tăng số N và lặp lại từ bước 2.*

* ***Sự hội tụ của phương pháp***

*Sự hội tụ của phương pháp Monte-Carlo được cho bởi 2 định lý quan trọng trong lý thuyết xác suất:*

* *Luật mạnh số lớn cho đảm bảo về sự hội tụ của phương pháp*
* *Định lý giới hạn trung tâm xác định tốc độ hội tụ, cụ thể như sau:*

*Nếu dãy các biến ngẫu nhiên độc lập cóa cùng phân phối (có momen cho đến bậc hai hữu hạn), khi đó*

*trong đó (Tốc độ hội tụ của mô phỏng Monte-Carlo xác định bởi*

* ***Các bước thực hiện mô phỏng Monte-Carlo***

1. *Xác định bài toán và hệ thống mô phỏng*
2. *Thiết lập mô hình toán học để mô phỏng.*
3. *Thu thập dữ liệu cần thiết cho mô hình.*
4. *Kiểm chứng mô hình và so sánh các kết quả và dữ liệu trong thực tế.*
5. *Chạy chương trình mô phỏng*
6. *Phân tích kết quả mô phỏng, điều chỉnh, thay đổi khi thấy cần thiết.*

*Phương pháp Monte-Carlo thường được sử dụng trong các trường hợp sau*

* *Giải bài toán mà không dễ dàng giải được bằng các phương pháp khác (như tích phân nhiều chiều hoặc phương trình vi phân,…)*
* *Mô phỏng là cách tính nhanh và hiệu quả*

*Hạn chế của phương pháp này là*

* *Không đưa ra lời giải chính xác, khó xác định được sai số*
* *Chỉ tạo ra phương pháp đánh giá chứ không đưa ra kỹ thuật tìm lời giải tối ưu*

***1.3 Một số phương pháp tạo số ngẫu nhiên dưới dạng giả ngẫu nhiên***

*(Phục vụ cho việc giải bài toán bằng phương pháp Monte-Carlo)*

* ***Phương pháp 1.***

*Sử dụng máy tính có gắn thêm thiết bị phát số ngẫu nhiên (RNG:Random Number Generator – bàn quay rulet, máy đếm số, máy gây nhiễu,…)Phương pháp này phiền phức đồng thời khó kiểm tra kết quả vì các số ngẫu nhiên không được lưu và in ra.*

* ***Phương pháp 2.***

*Sử dụng bảng các số ngẫu nhiên (chẳng hạn bảng do Tippet đưa ra năm 1972 dùng trong thống kê dân số; hoặc bảng do Rirma tạo ra gồm một triệu chữ số ngẫu nhiên thập phân)*

* ***Phương pháp 3.***

*Tạo số ngẫu nhiên bằng chương trình.Theo phương pháp này, người ta xây dựng những phần mềm để thiết lập những dãy số bằng các chương trình truy hồi cấp r, với r giá trị ban đầu là các đại lượng ngẫu nhiên đã cho.*

***1.4 Phương pháp nghịch đảo hàm phân phối***

*Đại lượng ngẫu nhiên X có hàm phân phối đã cho, nghĩa là xét những thể hiện trong phép thử mô phỏng của biến ngẫu nhiên X, trong đó là hàm không giảm, liên tục trái và Đối với những hàm có các đặc điểm trên, thì hàm ngược của nó nói chung là một hàm đa trị. Nhằm đơn trị hóa hàm ngược ta dùng phương pháp nghịch đảo hàm phân phối.*

* ***Đinh nghĩa.***

*Cho biến ngẫu nhiên X có hàm phân phối Hàm ngược của hàm phân phối được xác định bởi*

* ***Định lý 1 (****Phương pháp nghịch đảo hàm phân phối cho biến liên tục)*

*Cho là hàm liên tục, thực sự tăng trên (a,b), sao cho*

*và*

*Khi đó phương trình sẽ có nghiệm duy nhất.*

*Từ đặc tính trên ta sẽ rút ra hệ quả:*

* ***Hệ quả***

*Cho là hàm mật độ của đại lượng ngẫu nhiên xác định trên (a,b);*

*có nghiệm duy nhất với X là biến ngẫu nhiên có hàm mật độ*

* ***Định lý 2. (****Phương pháp nghịch đảo hàm phân phối cho biến rời rạc)*

*Cho biến ngẫu nhiên X nhận các giá trị rời rạc (I là tập vô hạn hay đếm được), với các xác suất tương ứng*

*Khi đó biến ngẫu nhiên X, sẽ xác định được bởi*

*trong đó .*

* ***Ví dụ 3.***

*Tạo biến ngẫu nhiên có phân phối mũ với tham số λ =1.*

*Gọi X là biến ngẫu nhiên có phân phối mũ với tham số λ =1, thì nó có hàm phân phối xác suất xác định bởi*

*Đặt thì*

*Lấy logarit hai vế ta được:*

*Vậy để tạo một biến ngẫu nhiên có phân phối mũ λ =1, ta tiến hành*

*Bước 1: Tạo số ngẫu nhiên U có phân phối đều trên khoảng (0,1).*

*Bước 2: Đặt*

*Chú ý:*

*\* U có phân phối đều trên khoảng (0,1), nên cũng có phân phối đều trên khoảng (0,1), do đó và có cùng phân phối.*

*Mà có phân phối mũ λ =1; nên cũng có phân phối mũ λ =1.*

*Do đó ở bước 2 ta chỉ cần đặt*

*\*\* Để tạo một biến ngẫu nhiên có phân phối mũ , ta tiến hành*

*Bước 1: Tạo số ngẫu nhiên U có phân phối đều trên khoảng (0,1).*

*Bước 2: Đặt*

* ***Ví dụ 4.***

*Tạo biến ngẫu nhiên có phân phối Gamma với tham số*

*Ta biết rằng biến ngẫu nhiên có phân phối Gamma sẽ có hàm phân phối*

*Nếu biến đổi trực tiếp hàm ngược giống ví dụ trên thì khá khó do phức tạp.*

*Vì vậy ở đây ta sử dụng tính chất của hàm Gamma “X có phân phối Gamma với tham số thì nó có thể biểu diễn được dưới dạng*

*với là các biến ngẫu nhiên độc lập có cùng phân phối mũ với tham số λ. Vậy ta sẽ sử dụng kết quả của ví dụ 3 để tạo các số ngẫu nhiên*

*Các bước thực hiện để tạo biến ngẫu nhiên có phân phối Gamma với tham số như sau:*

*Bước 1: Tạo số ngẫu nhiên có phân phối mũ với tham số λ.*

*Bước 2: Đặt Từ đó suy ra*

***PHƯƠNG PHÁP BOOTSTRAP TRONG PHÂN TÍCH CHUỖI THỜI GIAN***

***Bootstrap Methods for Time Serie***

*---------------------------------------------------------------------------------------------------------------------*

* 1. ***Khái niệm về phương pháp Bootstrap***
* *Phương pháp thống kê bootstrap, là một phương pháp được sử dụng hiệu quả với hỗ trợ của các công cụ tính toán mở rộng như máy tính.*

*Boostrapping method được xem là phương pháp chuẩn trong phân tích thống kê và đã làm nên một cuộc cách mạng trong thống kê vì nó giải quyết được nhiều vấn đề mà trước đây tưởng như không giải được, đặc biệt là vấn đề về dữ liệu.*

* *Bradley Efron đưa ra khái niệm lần đầu năm 1979, với sự hỗ trợ của máy tính phương pháp bootstrap đã khẳng định được tính hiệu quả trong thống kê Toán học, như hai nhà thống kê đã tuyên bố “Bootstrap là ý tưởng mới quan trọng nhất trong số liệu thống kê được giới thiệu trong 20 năm qua, và có lẽ trong 50 năm tới nữa.” (Jerome H. Friedman, 1998), “Bootstrap đã chỉ cho chúng ta cách sử dụng sức mạnh của máy tính và phương pháp lấy mẫu lặp để đi đến các kết luận thống kê mà tính toán lý thuyết không thể thực hiện, cũng như giới thiệu một cách suy nghĩ khác về số liệu thống kê.” (George Casella, 2003).*

*Sau đây chúng ta sẽ phân tích cụ thể hơn về phương pháp này.*

*Cho mẫu ngẫu nhiên cỡ n, từ hàm phân phối F chưa biết. Trong thống kê ta thường coi là những đại lượng ngẫu nhiên độc lập và có cùng phân phối F. Mỗi phép thử ngẫu nhiên là một thể hiện của mẫu tạo ra một giá trị cụ thể trong*

* ***Ý tưởng chính*** *của phương pháp bootstrap: Các giá trị trong một mẫu đã thu được có thể được lặp lại một cách ngẫu nhiên. Nói rộng hơn phương pháp bootstrap không chỉ xem một mẫu (sample) như một thể hiện mà xét nó như một tổng thể (population) hay đúng hơn là đại diện của tổng thể.*

*Giả sử ta xét thống kê phụ thuộc vào dữ liệu ngẫu nhiên trong mẫu, và phân phối của thống kê thông thường phụ thuộc vào để mô tả mối phụ thuộc đó ta thường viết . Trong nhiều trường hợp khi phân phối F chưa biết hoặc chưa xác định được người ta thường ước lượng bởi hàm phân phối thực nghiệm (EDF – empirical distribution function)*

*trong đó:*

*( thường được gọi là hàm chỉ tiêu của A).*

*Nếu (sao cho P(A) xác định được, nghĩa là A là tập Borel), ta sẽ có:*

*, khi*

*Kết quả được xác định trực tiếp theo luật số lớn, từ đó*

*Và như vậy có thể kết luận rằng:*

*khi*

*trong đó ß là tập các khoảng trong R. Nói một cách khác có thể xấp xỉ bởi đối với mọi*

* 1. ***Lấy mẫu từ phân phối thực nghiệm***

*Giả sử ta muốn tạo mẫu độc lập và có cùng phân phối từ Ta cho mỗi quan trắc một trọng số , và coi có phân phối đều trên tập , như vậy sẽ thu được mẫu lặp khác từ mẫu gốc đã có Quá trình trên được thể hiện theo thuật giải sau.*

**Thuật giải Bootstrap 1: Bootstrap độc lập có cùng phân phối** (i.i.d bootstrap).

1. *Tạo ra mẫu bootstrap ngẫu nhiên (bootstrap sample) có hoàn lại kích thước n từ mẫu ban đầu.*
2. *Tính các thông số đặc trưng cho thống kê được sinh ra (trung bình, khoảng tin cậy, độ lệch chuẩn,…)*
3. *Lặp lại các bước 1 và 2 với số B lần lớn (B thường lớn hơn 1000)*
4. *Sử dụng các thông số thống kê của các thống kê bootstrap thu được từ bước 2 để đánh giá về các thông số thống kê theo mẫu ban đầu (original sample), ta ước lượng bởi .*

Từ cách xác định mẫu lặp như trên ta thấy rằng:

Vậy theo thuật giải trên ta ta sẽ có trung bình ngẫu nhiên không chệch:

Khi đó phương sai mẫu lặp sẽ là:

trong đó là phương sai tương ứng với , và khác với .

**1.3 Phương pháp Bootstrap cho dữ liệu phụ thuộc**

Đối với dữ liệu phụ thuộc ta không thể áp dụng bootstrap tiêu chuẩn như đã trình bày ở phần trên, khi đó ta xét việc thực hiện với một số bootstrap đơn giản như: *Bootstrap chuyển dịch khối, Bootstrap khối phủ* như bên dưới.

### **Bootstrap cho chuỗi thời gian dừng – bootstrap khối**

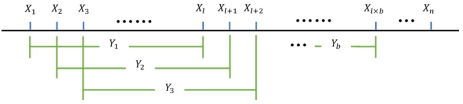
* **Bootstrap chuyển dịch khối MBB** (moving block bootstrap)

*Kỹ thuật bootstrap khối (block bootstrap) được giới thiệu bởi Carlstein E. (1986) và Kunsch H. R. (1989). Với mỗi một khối cấu trúc của chuỗi mang tính xu hướng của chuỗi thời gian dừng ta sẽ thực hiện kỹ thuật bootstrap khối mà dạng đơn giản thường gặp là bootstrap chuyển dịch khối theo thuật giải sau.*

**Thuật giải Bootstrap 2.** *Bootstrap chuyển dịch khối MBB****.***

1. *Tạo ra mẫu bootstrap trong dạng chồng (overlapping) độ dài l, sao cho khối thứ nhất là và khối thứ hai là ta sẽ có b khối độc lập có cùng phân phối và trong đó là phần nguyên của tỷ số*
2. *Các khối được lấy lặp ngẫu nhiên và ta sẽ thu được mẫu lặp (bootstrap sample)*
3. *Tính giá trị của thống kê*
4. *Lặp lại các bước 2 và bước 3 B lần. Gọi là thống kê bootstrap sinh bởi lần lặp b.*

*Ta ước lượng thông qua .*

**

*Bootstrap chuyển dịch khối*

* *Sự khác biệt giữa MBB và bootstrap dừng SB (stationary bootstrap) là việc sử dụng độ dài khối l trong SB là số ngẫu nhiên.*
* *Mặt khác Buhlmann P. (2002) đã chỉ ra rằng nếu*

*+ là một quá trình dừng Gauss,*

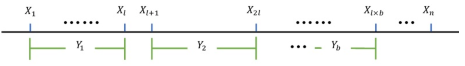
*+ là một hàm trơn của dữ liệu,*

*+ Độ dài khối l của mẫu cỡ n, chặn bởi ;*

*khi đó các sản phẩm của mẫu lặp MBB sẽ hội tụ như từ mẫu gốc. Trong trường hợp này MBB được gọi là tiệm cận hợp lý (Asimptopical valid).*

* ***Bootstrap khối không chồng NBB*** *(Non-overlapping Block Bootstrap)*

*Trước hết ta xét Bootstrap khối không chồng NBB cho trường hợp chuỗi thời đơn biến. Ta thấy rằng MBB và NBB chỉ có điểm khác nhau khi xác định khối bootstrap, các khối là độc lập cùng độ dài và rời nhau trong NBB như sau*

**

*Bootstrap khối không chồng*

*Ước lượng tham số trong dạng này có thể xác định , và trong phần này để đơn giản ta chỉ xét tham số là giá trị kỳ vọng.*

*Để phân biệt MBB và NBB cho ước lượng tham số ta ký hiệu tương ứng là khi đó ta sẽ có:*

*Như vậy xác suất của mỗi khối chọn (theo nghĩa đều) sẽ là:*

*.*

*Điều đó cũng có nghĩa là:*

*.*

*Suy ra ta sẽ có:*

*=*

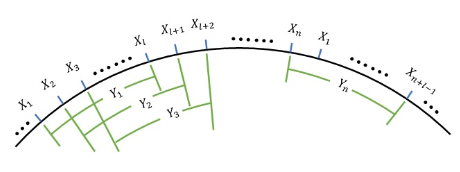
*Mặt khác theo cách xác định NBB ta thu được:*

*Chú ý rằng quá trình thỏa một số điều kiện nêu trên sẽ có độ lệch giữa hai ước lượng là: , và kỳ vọng của bình phương độ lệch quân phương là:*

*Từ đó ta thấy rằng khi cỡ mẫu lớn sự khác biệt giữa hai phương pháp từ hai loại bootstrap nói trên là không đáng kể.*

* ***Bootstrap khối vòng CBB*** *(Circular block bootstrap****)***

*Bootstrap khối vòng do Politis và Romano đưa ra năm 1992, được xác định tương tự như trong phương pháp xây dựng MBB. Với mỗi mẫu ta kéo dài dạng chuỗi thành các khối có độ dài l là*

**

*Bootstrap khối vòng*

*Mẫu với xác suất lặp của mỗi phần tử bằng , vậy ta sẽ có m khối với:*

*Từ đó ta ký hiệu chuỗi dữ liệu mới là , đó chính là bootstrap khối vòng CBB mà trong các tài liệu đã chúng minh tính vững và tính tiệm cận đúng đắn của chúng.*

* ***Bootstrap dừng SB*** *(Stationary Bootstrap****)***

*Cho là dữ liệu gốc là chuỗi thời gian dừng và hàm thống kê*

*Xét khối dữ liệu với thành phần đầu là và độ dài ngẫu nhiên l của nó có phân phối hình học: Khi đó pseudo chuỗi thời gian , là dãy các khối của chuỗi*

*Chú ý rằng chuỗi thời gian mới là không dừng nên ta thường gọi nó là pseudo (giả) chuỗi thời gian, nói một cách khác SP không phải là chuỗi thời gian dừng.*

* ***Bootstrap tự hồi quy sàng*** *(AR-Sieve Bootstrap)*

*Bootstrap tự hồi quy sàng – AR-SB được thiết lập cho mô hình chuỗi thời gian tự hồi quy (autoregressive – AR) . Trước hết ta nói về chuỗi thời gian gốc , và giả sử rằng nó có cấu trúc theo mô hình:*

*trong đó và là dãy các biến ngẫu nhiên độc lập có cùng phân phối có kỳ vọng bằng không và độc lập với*

*Với mô hình ta cần xác định giá trị của bậc p. Sử dụng tiêu chuẩn thông tin Akaike ta sẽ tìm ra bậc tự hồi quy*

*Cho là các ước lượng tham số tự hồi quy Yule-Walker, khi đó ta sẽ có:*

*trong đó:*

*Với các ước lượng Yule-Walker, một số phần của được xác định, còn phần quan trọng là thông tin về các phần dư.*

*Cho: là các phần dư của mô hình phù hợp, khi đó ta có thể xác định được hàm phân phối thực nghiệm của các phần dư độc lập có cùng phân phối.*

*, trong đó:*

*Vấn đề cuối cùng của bootstrap AR-SB là cấu trúc của chuỗi mới trên cơ sở của mô hình phù hợp. Mô hình phù hợp được cho bởi:*

*trong đó là các phần dư độc lập và có cùng phân phối*

* ***Các tiêu chuẩn lựa chọn mô hình***
* ***Tiêu chuẩn thông tin AIC (Akaike Information Criterion)***

*Giả sử rằng ta có mô hình thống kê cho một số dữ liệu. Đặt p là số tham số ước lượng trong mô hình. là giá trị lớn nhất của hàm hợp lý (likelihood function) cho mô hình,*

*trong đó là hàm phân phối đồng thời của trong mô hình nó phụ thuộc vào các tham số thông qua hàm hợp lý ta có thể đánh giá được sự phù hợp của mô hình thống kê với dữ liệu mẫu. Khi đó giá trị AIC của mô hình được tính như sau;*

*Mô hình có giá trị AIC càng nhỏ thì mô hình càng phù hợp. Vì thế, với bộ dữ liệu đươc đưa ra, ta lưa chọn mô hình có AIC nhỏ nhất.*

* ***Tiêu chuẩn thông tin BIC (Bayesian Information Criterion)***

*Tiêu chuẩn thông tin BIC được xác định như sau*

*trong đó;*

*là giá trị lớn nhất của hàm hợp lý,*

*là kích thước mẫu,*

*là số các tham số ước lượng trong mô hình.*

*Tương tự như tiêu chuẩn AIC, mô hình có giá trị BIC càng nhỏ thì càng phù hợp. Do đó được coi là các tiêu chuẩn lựa chọn mô hình.*