**TRANG THÔNG TIN VỀ LUẬN ÁN**

Tên đề tài luận án: **Thống kê Bayes và ứng dụng trong tài chính**

Chuyên ngành: **Toán Ứng dụng**

Mã số: **62460112**

Họ tên nghiên cứu sinh: **Lê Thanh Hoa**

Khóa đào tạo: 2015

Người hướng dẫn khoa học: **TS. Phạm Hoàng Uyên, TS. Nguyễn Thanh Bình**

Cơ sở đào tạo: **Trường Đại học Khoa học Tự Nhiên- ĐHQG.HCM**

1. **TÓM TẮT NỘI DUNG LUẬN ÁN**

 Luận án nghiên cứu được trình bày vấn đề lý thuyết của thống kê Bayes và ứng dụng bài toán thực tiễn về phân tích giá chứng khoán theo hai hướng tiếp cận:

* 1. Hướng tiếp cận thứ nhất: xác định miền mật độ xác suất cao nhất (HPD) trong thống kê Bayes là bài toán quan trọng.

 Trong khi các tham số của tổng thể được xem xét như hằng số với quan điểm từ thống kê tần suất, thì trong thống kê Bayes, các tham số được mô hình dưới dạng biến ngẫu nhiên nhằm mô tả chính xác sự biến động của tổng thể khi dữ liệu được cập nhật theo thời gian hoặc trên các trường không gian khác nhau. Chính vì vậy, việc xem xét các tham số dạng mô hình của biến ngẫu nhiên thông qua thống kê Bayes là cần thiết. Một trong những bài toán thống kê quan trọng là bài toán ước lượng miền tin cậy của tham số ngẫu nhiên trong bối cảnh sự đa dạng của các phân phối xác suất sao cho đảm bảo miền ước lượng có xác suất bằng $(1-α)$ cho trước với độ dài ngắn nhất. Trong luận án này, chúng tôi đề xuất phương pháp ước lượng miền HPD trong bối cảnh các kết quả nghiên cứu trước đó chỉ dừng ở ở ước lượng miền HPD có dạng một khoảng, chưa ước lượng được dưới dạng nhiều khoảng. Kết quả ứng dụng trong ước lượng mức phân vị dạng hỗn hợp các phân phối xác suất, dự báo xu hướng giá chứng khoán và dự báo khoảng giá chứng khoán.

* 1. Hướng tiếp cận thứ hai: với dữ liệu thu thập thì khả năng dữ liệu bị nhiễu rất cao nên ước lượng hàm mật độ xác suất mờ trong thống kê Bayes là hướng tiếp cận hợp lý.

Thật vậy, khi nghiên cứu các dữ liệu thực, không phải lúc nào cũng thu thập được dữ liệu dưới dạng số chính xác, thông qua một điểm hoặc một vec tơ. Khi đó, dữ liệu được biểu diễn dưới dạng số mờ; vì vậy, bài toán hiệu chỉnh dữ liệu mờ trong thống kê Bayes được nghiên cứu trong luận án. Kết quả nghiên cứu được ứng dụng trong phân tích tài chính thông qua dự báo giá chứng khoán với ước lượng điểm Bayes mờ và hàm giá trị mờ không âm.

1. **NHỮNG KẾT QUẢ MỚI CỦA LUẬN ÁN**
2. Hướng tiếp cận thứ nhất: tác giả đã chứng minh được công thức để ước lượng miền HPD, với sai số nhỏ hơn trong trường hợp miền HPD có dạng một khoảng so với các nghiên cứu trước, trong trường hợp miền HPD có dạng nhiều khoảng, chúng tôi đã ước lượng được miền HPD trong khi các nghiên cứu trước mới chỉ dừng lại ở cách chỉ ra khả năng có thể ước lượng miền HPD có dạng nhiều khoảng. Các kết quả trên thu được khi tiếp cận dựa vào sự sai khác giữa các giá trị của hàm mật độ xác suất tại các cận thay vì tại đúng giá trị của các cận đó. Kết quả lý thuyết này được công bố trong tạp chí Communications in Statistics - Simulation and Computation thuộc SCIE (Hoa Le, Uyen Pham, Phuong Nguyen & Pham The Bao. (2018). "Improvement on Monte Carlo Estimation of HPD intervals".  *Communication in Statistics - Simulation and Computation*, 1-17, doi:10.1080/03610918.2018.1513141).

Kết quả lý thuyết được ứng dụng trong bài toán sau:

 Ước lượng phân vị của hỗn hợp các phân phối xác suất được đăng trong Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á (Lê Thanh Hoa. (2018). "Ước lượng VaR và CVaR dạng hỗn hợp các phân phối xác suất thông qua mô phỏng Monte Carlo và ứng dụng trong phân tích giá chứng khoán Việt Nam". *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 29(9), 53–72).

 Các dự báo giá chứng khoán trong Robustness in Econometrics (Tran, H. D., Nguyen, S. P., Le, H. T., & Pham, U. H. (2017). An Alternative to p-Values in Hypothesis Testing with Applications in Model Selection of Stock Price Data. *In Robustness in Econometrics (pp. 305-319). Springer, Cham.*); Tạp chí Phát triển Kinh tế (Lê Thanh Hoa, Phạm Hoàng Uyên & Nguyễn Đình Thiên. (2017). "Một phương pháp mới tìm khoảng mật độ hậu nghiệm cao nhất và ứng dụng". *Tạp chí Phát triển Kinh tế*, 28(10), 79-120); Hội nghị Toàn quốc lần thứ IV về Ứng dụng Toán học, Hà Nội (Phạm Hoàng Uyên, Nguyễn Đình Thiên & Lê Thanh Hoa. (2015). "Thống kê Bayes trong dự báo giá chứng khoán Việt Nam". *Hội nghị Toàn quốc lần thứ IV về Ứng dụng Toán học, Hà Nội*, 12/2015, 437-464).

1. Hướng tiếp cận thứ hai: tác giả đã tìm được phương pháp ước lượng hàm mật độ xác suất mờ thông qua cận trên và cận dưới của các ($δ-cut$). Từ đây, tác giả đã có các kết quả được công bố về thống kê Bayes với dữ liệu mờ cũng như các ứng dụng trong tài chính thông qua dự báo giá chứng khoán trên các tạp chí: Annals of Biometrics & Biostatistics (Le H, Pham U, Nguyen NT, Bao PT. (2017). "Statistical Methods of Handling Noise in Data Processing". *Ann Biom Biostat* 4(1): 1027); Tạp chí Công nghệ Ngân hàng (Lê Thanh Hoa, Phạm Hoàng Uyên & Nguyễn Phúc Sơn. (2017). "Sử dụng Thống kê Bayes mờ trong dự báo tỷ giá và một số chỉ số kinh tế". *Tạp chí Công nghệ Ngân hàng*, 140 (11), 92- 100); Tạp chí Phát triển Khoa học và Công nghệ (Phạm Hoàng Uyên, Lê Thanh Hoa & Nguyễn Đình Thiên. (2017). "Chọn mô hình tốt nhất trong Thống kê Bayes mờ và ứng dụng trong phân tích Tài chính". *Tạp chí Phát triển Khoa học và Công nghệ*, 20 (2), 144-155).
2. **CÁC ỨNG DỤNG KHẢ NĂNG ỨNG DỤNG TRONG THỰC TIỄN HAY NHỮNG VẤN ĐỀ CÒN BỎ NGỎ CẦN TIẾP TỤC NGHIÊN CỨU**

Khi áp dụng thống kê Bayes vào thực tiễn, chúng tôi quan tâm đến bài toán với dữ liệu lớn và thực hiện để có kết quả với thời gian thực, mối quan hệ giữa các hàm mật độ xác suất thông qua bài toán về copula, các mở rộng trong xác suất lượng tử sẽ là hướng mở rộng của luận án.

|  |  |
| --- | --- |
| **CÁN BỘ HƯỚNG DẪN****TS. Phạm Hoàng Uyên TS. Nguyễn Thanh Bình** | **NGHIÊN CỨU SINH****Lê Thanh Hoa** |

**XÁC NHẬN CỦA CƠ SỞ ĐÀO TẠO**

**PHÓ HIỆU TRƯỞNG**

 **Tran Le Quan**

**THESIS INFORMATION**

Thesis title: **Bayesian Statistics and Applications in Finance**

Speciality: **Applied Mathematics**

Code: **62460112**

PhD Student: **Le Thanh Hoa**

Academic year: **2015**

Supervisor: **PhD. Pham Hoang Uyen, PhD. Nguyen Thanh Binh**

At: **UNIVERSITY OF SCIENCE – VNU.HCMC**

1. **ABSTRACT**

The thesis presents the theory of Bayesian statistics and its application on practical problems of stock price analysis in two approaches:

* 1. The first approach: determining the highest probability density region (HPD) in Bayesian statistics is an important issue.

While the parameters of the population are considered as constants from the perspective of frequency statistics, in Bayesian statistics, parameters are modeled as random variables to accurately describe the fluctuations of the population which is updated over time or across different space fields. Therefore, it is necessary to consider the model parameters as random variables through Bayesian statistics. One of the most important statistical problems is estimating the credible interval of random parameters, which has a probability of $(1-α)$ with given the shortest length. In this thesis, we propose a method of estimating HPD in multi intervals since the previous research results only stop at estimating HPD in a one interval. The results of this thesis suggest applications in estimating the quantiles of the mixed form of probability distributions, forecasting stock price trends and intervals.

* 1. The second approach: In practical the collected data usually contain noises. Therefore, the estimation of the fuzzy probability density function in Bayesian statistics is a suitable approach to address this issue.

Indeed, when studying practical data, it is not always possible to collect data in the proper numerical form. The data also exist in different forms such as a single point or a vector. At that time, data could be expressed as fuzzy numbers explaining the adjusting fuzzy data using Bayes statistics in the thesis. The research results are applied in financial analysis such as stock price prediction using fuzzy Bayesian point estimates and non-negative fuzzy value functions.

1. **NEW RESULTS**
	1. First approach: the authors proved the formula to estimate the HPD with a smaller error for a single range compared to previous studies. Furthermore, we propose the estimation of the HPD in the form of multiple intervals while previous studies just suggested the idea. The above results are estimated from the difference between the values ​​of the probability density function at the neighborhoods of a specific value instead of using solely itself. In addition, this approach is published in Communications in Statistics - Simulation and Computation of SCIE (Hoa Le, Uyen Pham, Phuong Nguyen & Pham The Bao. (2018). "Improvement on Monte Carlo Estimation of HPD intervals". Communication in Statistics - Simulation and Computation, 1-17, doi: 10.1080 / 03610918.2018.1513141).

This research is applied in the following problems:

 Estimating the quantiles of a mixture of probability distributions is published in Journal of Asian Business and Economic Studies (Le Thanh Hoa. (2018). "Estimating VaR and CVaR in a mixture of probability distributions through Monte Carlo simulation and application of Vietnam stock price analysis". *Journal of Asian Business and Economic Studies*, 29(9), 53–72; the stock price forecast in Robustness in Econometrics (Tran, H. D., Nguyen, S. P., Le, H. T., & Pham, U. H. (2017). An Alternative to p-Values in Hypothesis Testing with Applications in Model Selection of Stock Price Data. *In Robustness in Econometrics (pp. 305-319). Springer, Cham.*); Journal of Economic Development (Le Thanh Hoa, Pham Hoang Uyen & Nguyen Dinh Thien. (2017). "A new method of estimating the highest posterior density interval application". *Journal of Economic Development*, 28(10), 79-120); Fourth National Conference on Applied Mathematics, Hanoi (Pham Hoang Uyen, Nguyen Dinh Thien & Le Thanh Hoa. (2015). " Bayesian statistics in Vietnam stock price forecast". *Fourth National Conference on Applied Mathematics, Hanoi*, 12/2015, 437-464).

* 1. The second approach: the authors propose a method for estimating the fuzzy probability density function through the upper and lower bound ($δ-cut$). As a result, the author has published applications of Bayes statistics on fuzzy data as well as stock price forecasts in academic magazines:

Annals of Biometrics & Biostatistics (Le H, Pham U, Nguyen NT, Bao PT. (2017). "Statistical Methods of Handling Noise in Data Processing". *Ann Biom Biostat* 4(1): 1027); Banking Technology Review (Le Thanh Hoa, Pham Hoang Uyen & Nguyen Phuc Son. (2017). " Using fuzzy Bayesian statistics forecasting exchange rate and economic index". Banking Technology Review, 140 (11), 92- 100); Science & Technology Development (Pham Hoang Uyen, Le Thanh Hoa & Nguyen Dinh Thien. (2017). "Choosing the best model in fuzzy Bayesian statistics and its application in financial analysis". Science & Technology Development, 20 (2), 144-155).

1. **POSSIBLE APPLICATIONS**

When applying Bayesian statistics into practice, we are interested in handling and real time processing big data, the relationship between probability density functions through copula problems, the expansion in quantum probability will be the extension of the thesis.

|  |  |
| --- | --- |
| **SUPERVISOR****PhD. Pham Hoang Uyen PhD. Nguyen Thanh Binh** | **PhD STUDENT****Lê Thanh Hoa** |

**CONFIRMATION OF THE UNIVERSITY OF SCIENCE**

**VICE PRESIDENT**

**Tran Le Quan**