การประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงโดยวิธีเคอร์เนล

นางละเอียด ปรารถนาดี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต

ภาควิชาสถิติ

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

พ.ศ. 2531

ISBN 974-569-395-2

ลิขสิทธิ์ของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

013918

i 10300922

THE ESTIMATION OF PROBABILITY FUNCTIONS FOR THE POISSON DISTRIBUTION BY KERNEL METHOD

Mrs. La-ied Prathnadee

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science

Department of Statistics

Graduate School

Chulalongkorn University

1988

ISBN 974-569-395-2

หัวข้อวิทยานิพนธ์	การประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงโดยวิธีเคอร์เนล
โดย	นางละเอียด ปรารถนาดี
ภาควิชา	สถิติ
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.อำพล ธรรมเจริญ
อาจารย์ที่ปรึกษาร่วม	อาจารย์สุวาณี สุรเสียงสังข์

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีบัณฑิตวิทยาลัย

(ศาสตราจารย์ ดร.ถาวร วัชราภัย)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(รองศาสตราจารย์ ดร.สรชัย พิศาลบุตร)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.อำพล ธรรมเจริญ)

..... กรรมการ

(อาจารย์สุวาณี สุรเสียงสังข์)

..... กรรมการ

(รองศาสตราจารย์ ดร.สุชาดา กีระนันทน์)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ธีระพร วีระถาวร)

ละเอียด ปรารถนาดี : การประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงโดยวิธีเคอร์เนล (THE ESTIMATION OF PROBABILITY FUNCTIONS FOR THE POISSON DISTRIBUTION BY KERNEL METHOD) อ.ที่ปรึกษา : ผศ.ดร.อำพล ธรรมเจริญ, อ.สุวาณี สุรเสียงสังข์, 89 หน้า.

การวิจัยครั้งนี้เป็นการศึกษาวิธีการประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงด้วยวิธีเคอร์เนล ซึ่งวิธีนี้ใช้ได้ดีกับการประมาณฟังก์ชันความหนาแน่นของตัวแปรสุ่มต่อเนื่อง โดยทำการจำลองข้อมูลให้เป็นกลุ่มตัวอย่างที่มีขนาด 30 50 และ 100 จากประชากรที่มีการแจกแจงปัวส์ซงพารามิเตอร์ λ เป็น 1 2 3 4 และ 5 สำหรับพารามิเตอร์แต่ละค่าและขนาดกลุ่มตัวอย่างแต่ละขนาด ทำการจำลองข้อมูลซ้ำ 200 ครั้ง แต่ละครั้งคำนวณตัวประมาณเคอร์เนล ($\hat{f}$) โดยใช้ฟังก์ชันเคอร์เนลและความกว้างต่าง ๆ กัน แล้วเลือกรูปแบบของฟังก์ชันเคอร์เนลและความกว้างที่ทำให้ค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณต่ำสุด

ผลการวิจัยพบว่า การประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงควรใช้วิธีเคอร์เนลที่มีฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล และความกว้างเป็น 0.5 เมื่อ $\lambda = 1$ และเมื่อ λ เป็น 2 3 4 และ 5 ใช้ความกว้างเป็น $1.06\hat{\sigma}n^{-(1/5)}$ โดยรูปแบบของการประมาณดังกล่าวใช้ได้สำหรับขนาดกลุ่มตัวอย่าง 30 50 และ 100 เมื่อนำผลการวิจัยนี้ไปประยุกต์ใช้กับข้อมูลจริงเพื่อประมาณฟังก์ชันความน่าจะเป็นของจำนวนเซลล์ที่มีสารอัลคาลอยด์ในใบหญ้าเกล็ดปลา ใช้วิธีเคอร์เนลที่มีฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล และความกว้างเป็น 0.5 เมื่อ $\hat{\lambda}$ เป็น 0.74 และ 0.76 ส่วนเมื่อ $\hat{\lambda}$ เป็น 0.5 ได้ใช้การประมาณฟังก์ชันความน่าจะเป็นโดยวิธีหาความถี่สัมพัทธ์ซึ่งให้ผลดีกว่า

ภาควิชา สถิติ
สาขาวิชา สถิติ
ปีการศึกษา 2530

ลายมือชื่อนิสิต .....
ลายมือชื่ออาจารย์ที่ปรึกษา .....

LA-IED PRATHNADEE : THE ESTIMATION OF PROBABILITY FUNCTIONS FOR THE POISSON DISTRIBUTION BY KERNEL METHOD. THESIS ADVISORS : ASST. PROF. AMPON DHAMACHAROEN : Ph.D., SUWANEE SURASIENGSUNK : M.S. 89 pp.

This research is an experiment on estimating the Poisson probability funtion by the kernel method which is generally appropriate for estimating density functions of the continuous random variables. The samples of size 30, 50 and 100 were generated from the data simulated from the Poisson population with parameter λ = 1, 2, 3, 4 and 5. For each sample size and each parameter λ, the simulation was repeated 200 times. Each time the kernel estimator $\hat{f}$ with various window widths and kernel functions were computed. The form of window width and kernel function with the minimum mean sum square error was selected.

The results are as follows: For all cases, the kernel function is the Gaussian density function, for Poisson distribution with parameter λ = 1, the window width is 0.5; for λ = 2, 3, 4 and 5, the window width are $1.06\,\sigma n^{-(1/5)}$. These findings were applied to estimate the probability function of the number of the alkaloid containing leaf cells of Phyla nodiflora Greene. It was found that when $\hat{\lambda}$ = 0.74 and 0.76, the kernel method with Gaussian density function and window width 0.5 is the estimated probability function. When $\hat{\lambda}$ = 0.5, the relative frequency is more appropriate.

ภาควิชา สถิติ
สาขาวิชา สถิติ
ปีการศึกษา 2530

ลายมือชื่อนิสิต .....
ลายมือชื่ออาจารย์ที่ปรึกษา .....

กิตติกรรมประกาศ

ผู้เขียนขอกราบขอบพระคุณคณาจารย์ภาควิชาสถิติทุกท่านที่ได้ให้ความรู้แก่ผู้เขียน และขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.อำพล ธรรมเจริญ อาจารย์ที่ปรึกษา อาจารย์สุวาณี สุรเสียงสังข์ อาจารย์ที่ปรึกษาร่วม ที่ได้กรุณาให้ข้อแนะนำอย่างดียิ่ง

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดำรัส วงศ์สว่าง แห่งบัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล ที่กรุณาให้ข้อแนะนำและความช่วยเหลือเกี่ยวกับการเขียนโปรแกรมคอมพิวเตอร์

ขอขอบคุณ พันเอกวิโรจน์ ปรารถนาดี ลูก ๆ และเพื่อนร่วมงานทุกท่านที่ได้ให้กำลังใจ และความช่วยเหลือ จนทำให้วิทยานิพนธ์สำเร็จด้วยดี

ละเอียด ปรารถนาดี

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญตาราง ..... ช

สารบัญรูป ..... ฏ

บทที่ 1 บทนำ 1

1.1 ความเป็นมาของปัญหา 1

1.2 วัตถุประสงค์ของการวิจัย 4

1.3 ขอบเขตของการวิจัย 5

1.4 ประโยชน์ที่คาดว่าจะได้รับ 6

บทที่ 2 ทฤษฎีที่เกี่ยวข้องในการวิจัย 7

บทที่ 3 วิธีดำเนินการวิจัย 17

บทที่ 4 ผลการวิจัย 20

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 65

5.1 สรุปผลการวิจัย 65

5.2 ข้อเสนอแนะ 72

บรรณานุกรม 74

ภาคผนวก 76

ภาคผนวก ก. 77

ภาคผนวก ข. 78

ประวัติผู้เขียน 89

สารบัญตาราง

ตารางที่			หน้า
4.1.1	4.1.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 1$ และ $n = 30$	21
4.2.1	4.2.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 1$ และ $n = 50$	23
4.3.1	4.3.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 1$ และ $n = 100$	25
4.4.1	4.4.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 2$ และ $n = 30$	27
4.5.1	4.5.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 2$ และ $n = 50$	29

สารบัญตาราง (ต่อ)

ตารางที่			หน้า
4.6.1	4.6.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 2$ และ $n = 100$	32
4.7.1	4.7.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 3$ และ $n = 30$	35
4.8.1	4.8.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 3$ และ $n = 50$	38
4.9.1	4.9.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 3$ และ $n = 100$	41
4.10.1	4.10.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 4$ และ $n = 30$	44

สารบัญตาราง (ต่อ)

ตารางที่			หน้า
4.11.1	4.11.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 4$ และ $n = 50$	47
4.12.1	4.12.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 4$ และ $n = 100$	50
4.13.1	4.13.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 5$ และ $n = 30$	53
4.14.1	4.14.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 5$ และ $n = 50$	56
4.15.1	4.15.2	แสดงค่าคาดหวัง (ค่าเอียงเฉ)^2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความแปรปรวนของ $\hat{f}$ เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 5$ และ $n = 100$	59
4.16.1		แสดงผลการประมาณค่า $\hat{\lambda}$ ของข้อมูลจริง	62
4.16.2		แสดงผลการทดสอบภาวะสารูปดีของข้อมูลจริง	63

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4.16.3	แสดงผลการประมาณฟังก์ชันความน่าจะเป็นของ จำนวนเซลล์ที่มีสารอัลคาลอยด์ในใบหญ้าเกล็ดปลา	63
5.1	แสดงวิธีการประมาณฟังก์ชัน $\hat{f}$ ที่ทำให้ค่าผลรวม ความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด	65
5.2	แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 1$ n = 30 50 และ 100	66
5.3	แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 2$ n = 30 50 และ 100	67
5.4	แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 3$ n = 30 50 และ 100	68
5.5	แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 4$ n = 30 50 และ 100	69
5.6	แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 5$ n = 30 50 และ 100	70

สารบัญรูป

รูปที่ หน้า

2.1 ฮิสโตแกรมแสดงระยะเวลา (วัน) ในการทดลองกับคนไข้กลุ่มควบคุม ที่ใช้ศึกษาเรื่องการฆ่าตัวตาย 8

2.2 ฟังก์ชันความหนาแน่นที่แท้จริง 11

2.3 (ก) (ข) และ (ค) เป็นฟังก์ชันความหนาแน่นที่ประมาณจากข้อมูลตัวอย่างขนาด 200 โดยวิธีเคอร์เนล ใช้ความกว้าง : (ก) 0.1 (ข) 0.3 (ค) 0.6 11

2.4 การประมาณฟังก์ชันความหนาแน่นโดยวิธีเคอร์เนล โดยแสดงให้เห็นถึงแต่ละเคอร์เนล (ความกว้าง 0.4) 12

2.5 (ก) (ข) การประมาณฟังก์ชันความหนาแน่น โดยวิธีเคอร์เนล โดยแสดงให้เห็นถึงแต่ละเคอร์เนล ความกว้าง (ก) 0.2 (ข) 0.8 ... 13