บทที่ 3

วิธีดำเนินการวิจัย

กระบวนการสโตแคสติกสำหรับการวิเคราะห์ตัวแบบอนุกรมเวลาโดยประยุกต์ใช้กับข้อมูลราคาจริงและราคาในตลาดล่วงหน้าของข้าวในประเทศไทยในครั้งนี้ มีวิธีดำเนินการในแต่ละข้อต่อไปนี้

ข้อมูลและแหล่งข้อมูล
ขอบเขตการวิจัย
ขั้นตอนการดำเนินการวิจัย
สถานที่ทำวิจัย
ระยะเวลาในการทำวิจัย

ข้อมูลและแหล่งข้อมูล

ข้อมูลราคาปิดของข้าวขาว5% เฉลี่ยรายเดือน ตั้งแต่ปี 2548-2557 จำนวน 120 เดือน จากตลาดสินค้าเกษตรล่วงหน้าแห่งประเทศไทย ข้อมูลราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน ตั้งแต่ปี 2540-2557 จำนวน 216 เดือนจากสำนักงานสถิติการเกษตร

ขอบเขตการวิจัย

1. ประชากรและกลุ่มตัวอย่าง

ประชากร คือ ราคาปิดของข้าวขาว5% รายเดือนเฉลี่ย ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

และกลุ่มตัวอย่าง คือ ราคาปิดของข้าวขาว5% รายเดือนเฉลี่ย จำนวน 120 เดือน ราคาข้าวเปลือกเจ้าความชื้น 15% และราคาข้าวเปลือกเจ้าหอมมะลิ 105 เฉลี่ยรายเดือน

2. ตัวแปรสำหรับการวิจัย

ตัวแปรต้น แนวโน้ม (Trend) ฤดูกาล (Seasonal) ค่าผิดปกติ (Irregular Variation) ค่าอัตตสหสัมพันธ์ (Autocorrelation)

ตัวแปรตามคือ ราคาข้าวแต่ละชนิดรายเดือน

ขั้นตอนการดำเนินการวิจัย

1. ศึกษาหัวเรื่อง ต่อไปนี้

1.1 ราคาข้าวของไทย

1.2 กระบวนการสโตแคสติก ตัวแบบพยากรณ์ข้อมูลอนุกลมเวลาด้วยวิธีการของเบย์ และใช้การประมาณค่าพารามิเตอร์ด้วยวิธี Markov Chain Monte Carlo (MCMC) โดยสุ่มตัวอย่างแบบ Gibbs sampling ที่มีตัวแปรตามมีค่าต่อเนื่อง จึงสมมติให้มีการแจกแจงแบบปกติ และตัวแปรต้นเป็นฟังก์ชั่นต่างๆ

1.3 การประมาณค่าด้วยวิธีการของเบย์

1.4 งานวิจัยที่เกี่ยวข้องกับการประยุกต์ใช้ตัวแบบทางด้านสโตแคสติก ตัวแบบพยากรณ์ต่างๆ และงานวิจัยที่เกี่ยวข้องกับข้าว

2. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลใช้โปรแกรมวิเคราะห์ข้อมูลทางสถิติ วิเคราะห์ลักษณะทั่วไปของข้อมูลจากตัวอย่าง วิเคราะห์ราคาข้าวแต่ละชนิด

2.1 การวิเคราะห์ลักษณะทั่วไปของข้อมูลจากตัวอย่าง ใช้ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน

2.2 การวิเคราะห์ราคาข้าว โดยใช้หลักการทางกระบวนการสโตแคสติก ด้วยวิธีการของเบย์ในการประมาณค่าพารามิเตอร์วิธี Markov Chain Monte Carlo (MCMC) โดยสุ่มตัวอย่างแบบ Gibbs sampling ที่มีตัวแปรตามมีค่าต่อเนื่อง จึงกำหนดให้มีการแจกแจงแบบปกติ และตัวแปรต้นเป็นฟังก์ชั่นต่างๆ

รายละเอียดของตัวแบบแสดงดังต่อไปนี้

$$Y_t \sim N\left(\gamma(\Delta W(t\,|\,\alpha,\delta)+A_t)+\sum_{i=1}^{s-1}\omega_i S_{it},\left[\gamma(1+\zeta_t)\sigma_Y\right]^2\right)$$

ซึ่ง ค่าเฉลี่ยของ Y_t คือ

$$E(Y_t)=\gamma(\Delta W(t\,|\,\alpha,\delta)+A_t)+\sum_{i=1}^{s-1}\omega_i S_{it} \qquad (54)$$

ค่าความแปรปรวนของ Y_t คือ

$$Var(Y_t)=\left[\gamma(1+\zeta_t)\sigma_Y\right]^2 \qquad (55)$$

เมื่อ

γ ค่าคาดหวังของ Z ซึ่งค่า Z คือผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษา

$W(t\,|\,\alpha,\delta)$ คือการแจกแจงสะสมของแนวโน้ม ได้แก่เวย์บูล (Weibull) และ เอ็กซ์โพเนนเชียล (Exponential)

A_t ค่าอัตตสหสัมพันธ์ที่ซ่อนเล้นอยู่ ณ เวลา t

ζ_t ค่าผิดปกติ ณ เวลา t

$\omega_1, \omega_2, ..., \omega_{s-1}$ ค่าสัมประสิทธิ์การถดถอยของตัวแปรดัมมี่ของฤดูกาล

$S_{1t}, ..., S_{s-1t}$ ค่าฤดูกาล ณ เวลา t และ σ_Y^2 คือค่าความแปรปรวนโดยทั่วๆไปของ Y_t

กำหนดค่าการแจกแจงของ prior ในตัวแบบด้วยวิธีการของเบย์คือ

$$\omega_1, \omega_2, ..., \omega_{s-1} \sim N(0, 1.0\text{E}06)$$
$$p(\sigma_Y^2) \sim \text{Inv}Gamma(0.1, 0.0001)$$

แนวโน้มคือ

$$\Delta W(t \mid \alpha, \delta) = W(t \mid \alpha, \delta) - W(t-1 \mid \alpha, \delta)$$

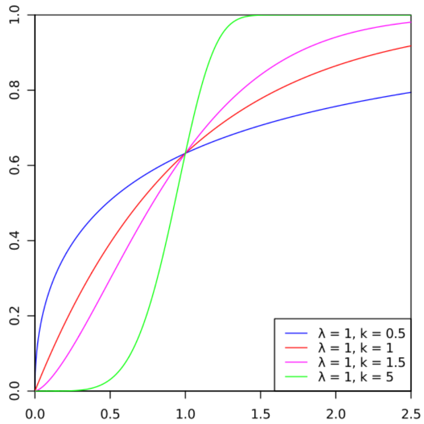
เมื่อ

$$\alpha \sim N_{[0,\infty)}(\mu_\alpha, \sigma_\alpha^2),\ p(\mu_\alpha) \sim N(0, 1.0\text{E}06),$$
$$p(\sigma_\alpha^2) \sim \text{Inv}Gamma(0.1, 0.0001)$$
$$\delta \sim N_{[0,\infty)}(\mu_\delta, \sigma_\delta^2),\ p\left(\mu_\delta\right) \sim N\left(0, 1.0\text{E}06\right),$$
$$p(\sigma_\delta^2) \sim \text{Inv}Gamma(0.1, 0.0001)$$

$CDF(Weibul)$ ซึ่งเป็นฟังก์ชั่นการแจกแจงความน่าจะเป็นสะสมของเวย์บูลคือ

$$F\left(y; \alpha, \delta\right) = \begin{cases} 1 - e^{-(y/\delta)^\alpha}, y \geq 0 \\ 0,\ y < 0 \end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่างๆกันดังภาพที่ 11

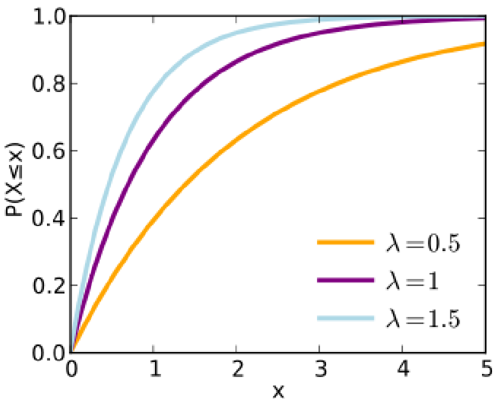


ภาพที่ 11 กราฟฟังก์ชั่นการแจกแจงความน่าจะเป็นสะสมของเวย์บูลเมื่อกำหนดค่ต่างๆ ของพารามิเตอร์ (ค่า $k=\alpha$ และ $\lambda=\delta$ สำหรับในงานวิจัยนี้)

$CDF(\exp)$ เป็นฟังก์ชั่นการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล คือ

$$F(y;\alpha)=\begin{cases}1-e^{-\alpha y}, y\geq 0\\ 0,\ y<0\end{cases}$$

และ แสดงด้วยภาพกราฟเมื่อกำหนดให้พารามิเตอร์มีค่าต่างๆกันดังภาพที่ 12



ภาพที่ 12 กราฟฟังก์ชั่นการแจกแจงความน่าจะเป็นสะสมของเอ็กซ์โพเนนเชียล เมื่อกำหนดค่าต่างๆของพารามิเตอร์ (ค่า $\lambda=\alpha$ สำหรับในงานวิจัยนี้)

ค่าอัตตสหสัมพันธ์ที่ซ่อนเล้นอยู่: AR(i) คือ

$$A_t \sim N(\lambda A_{t-i}, \sigma_A^2),\ p(\sigma_A^2) \sim \text{Inv}Gamma(0.1, 0.001)$$
$$\lambda \sim \text{N}(0, 1.0\text{E}06),\ A_0 = 0$$

ค่าผิดปกติคือ

$$\zeta_t \sim \text{Bern}(0.05)$$

ค่าคาดหวังของผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษาคือ

$$\gamma \sim N_{[0,\infty)}(\mu_\gamma, \sigma_\gamma^2),\ p(\mu_\gamma) \sim N(0, 1.0\text{E}06)$$
$$p(\sigma_\gamma^2) \sim \text{InvGamma}(0.1, 0.0001)$$

ผลรวมของข้อมูลอนุกรมเวลาในช่วงเวลาที่ศึกษา

$$Z \sim N(\gamma, \sigma_Z^2),\ p(\sigma_Z^2) \sim \text{InvGamma}(0.1, 0.0001)$$

สมการพยากรณ์ข้อมูลในอนาคตสามารถอธิบายในรูปของฟังก์ชันการแจกแจงของ posterior คือ

$$p(Y_{t+1} \mid Y_t, Y_{t-1}, \ldots, Y_1) = \int \ldots \int p(Y_{t+1} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid Y_t, Y_{t-1}, \ldots, Y_1)\, d\boldsymbol{\theta} \qquad (56)$$

หรือ

$$p(Y_{t+1} \mid Y_t, Y_{t-1}, \ldots, Y_1) \propto \int \ldots \int p(Y_{t+1} \mid \boldsymbol{\theta}) p(Y_t, Y_{t-1}, \ldots, Y_1 \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})\, d\boldsymbol{\theta} \qquad (57)$$

จากสมการที่ 56 และ 57 สามารถอธิบายสมการพยากรณ์ข้อมูลไปข้างหน้า โดยการอินทิกรัลซึ่งทำการประมาณค่าคำตอบโดยใช้ MCMC แบบ Gibbs sampling ซึ่งจะได้ค่าประมาณของ $\hat{Y}_{t+1}$

3. สรุปผลการวิจัย

นำผลการวิเคราะห์ที่ได้ มาพิจารณาเพื่อหาข้อสรุป อภิปรายผล และข้อเสนอแนะต่อไป

สถานที่ใช้ในการทำวิจัย

สาขาวิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์ ซึ่งตั้งอยู่ที่ศูนย์พระนครเหนือ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ระยะเวลาในการวิจัย

เริ่มตั้งแต่ 1 ตุลาคม 2557 สิ้นสุดการวิจัย 30 กันยายน 2558