EC435
หัวข้อ 1: อนุกรมเวลาทางการเงินและคุณลักษณะ

เฉลิมพงษ์ คงเจริญ ©2556-2562

คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์

วันที่ 16 สิงหาคม 2562

การคำนวณผลตอบแทน (return)

- ในการศึกษาทางการเงินส่วนมากเรามักจะสนใจผลได้ตอบแทน(return)ของทรัพย์สินมากกว่าราคา (price)

การคำนวณผลตอบแทน (return)

- ในการศึกษาทางการเงินส่วนมากเรามักจะสนใจผลได้ตอบแทน(return)ของทรัพย์สินมากกว่าราคา (price)
- ผลได้ตอบแทนของทรัพย์สินมีความสมบูรณ์ (complete) และปราศจากผลของหน่วยวัดเพื่อที่จะใช้ในการประเมินโอกาสในการลงทุน เช่นการระบุว่าราคาเพิ่มขึ้น 10 บาทไม่ได้บอกว่าผลได้ตอบแทนนั้นดีหรือไม่ จำเป็นต้องระบุในรูปของผลได้ตอบแทนเป็นเปอร์เซนต์

การคำนวณผลตอบแทน (return)

- ในการศึกษาทางการเงินส่วนมากเรามักจะสนใจผลได้ตอบแทน(return)ของทรัพย์สินมากกว่าราคา (price)
- ผลได้ตอบแทนของทรัพย์สินมีความสมบูรณ์ (complete) และปราศจากผลของหน่วยวัดเพื่อที่จะใช้ในการประเมินโอกาสในการลงทุน เช่นการระบุว่าราคาเพิ่มขึ้น 10 บาทไม่ได้บอกว่าผลได้ตอบแทนนั้นดีหรือไม่ จำเป็นต้องระบุในรูปของผลได้ตอบแทนเป็นเปอร์เซนต์
- อนุกรม (series) ของผลได้ตอบแทนมีคุณสมบัติทางสถิติที่ดีกว่าอนุกรมของราคาในการดำเนินการทางสถิติ

นิยามของผลได้ตอบแทนของทรัพย์สิน

สมมุติให้การซื้อสินทรัพย์ เช่นหุ้น พันธบัตร หรือกองทุนรวม ณ เวลา t_0 ด้วยราคา P_{t_0} บาทและขายสินทรัพย์ ณ เวลา t_1 ด้วยราคา P_{t_1} บาท ร้อยละของการเปลี่ยนแปลงของราคา

$$R(t_0, t_1) = \frac{P_{t_1} - P_{t_0}}{P_{t_0}} \tag{1}$$

เราจะเรียกระยะเวลาระหว่าง t_0 และ t_1 ว่าระยะเวลาการถือสินทรัพย์ (holding period)

เราจะสมมุติให้ระยะเวลาการถือมีลักษณะเพิ่มขึ้นเช่นเดียวกับปฏิทิน เช่น รายนาที รายวัน รายเดือน หรือรายปี

ขาย ซื้อ P_{t-1} P_t $t-1$ t

One-month simple return

กำหนดให้ P_t เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน t และไม่มีการจ่ายเงินปันผล และกำหนดให้ P_{t-1} เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน $t-1$

- **ผลได้ตอบแทนรวมอย่างง่ายหนึ่งเดือน (one-month simple gross return)** สามารถคำนวณได้ดังนี้

$$1 + R_t = \frac{P_t}{P_{t-1}} \tag{2}$$

One-month simple return

กำหนดให้ P_t เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน t และไม่มีการจ่ายเงินปันผล และกำหนดให้ P_{t-1} เป็นราคาของสินทรัพย์ ณ เวลาสิ้นเดือน $t-1$

- **ผลได้ตอบแทนรวมอย่างง่ายหนึ่งเดือน (one-month simple gross return)** สามารถคำนวณได้ดังนี้

$$1 + R_t = \frac{P_t}{P_{t-1}} \tag{2}$$

- **ผลได้ตอบแทนสุทธิอย่างง่ายหนึ่งเดือน (one-month simple net return)** หรือ ผลได้ตอบแทนอย่างง่ายหนึ่งเดือน (one-month simple return)

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \boxed{\frac{P_t - P_{t-1}}{P_{t-1}}} \tag{3}$$

ตัวอย่างที่ 1.1

สมมุติว่าเราพิจารณาการลงทุนในหุ้น APPLE โดยที่เราซื้อหุ้นเมื่อเดือน $t-1$ ด้วยราคา $P_{t-1} = 190$ ดอลลาร์และขายไปในเดือนต่อไปด้วยราคา $P_t = 200$ ดอลลาร์และไม่มีการจ่ายเงินปันผลในระหว่างที่เราถือหุ้น ดังนั้นผลได้ตอบแทนสุทธิและผลได้ตอบแทนอย่างง่ายรวมหนึ่งเดือนจะเท่ากับ

one-month Simple Gross return

$$\underline{1+R_t} = \frac{P_t}{P_{t-1}} = \frac{200}{190} \approx 1.0526$$

one-month Simple (net) return

$$R_t = \frac{200-190}{190} \approx 0.0526$$

ผลได้ตอบแทนหลายเดือน

ผลได้ตอบแทนจากการลงทุนสองเดือนสามารถคำนวณได้จากการหาการเปลี่ยนแปลงของราคา ณ เดือน P_t และ P_{t-2} หรือผลได้ตอบแทนอย่างง่ายสองเดือนจะเท่ากับ

2 months

$$R_t(\underline{2}) = \frac{P_t - P_{t-2}}{P_{t-2}}$$

$$R_t(2) = \frac{P_t}{P_{t-2}} - \frac{P_{t-2}}{P_{t-2}}$$

$$\frac{P_{t-2}}{P_{t-2}} = 1$$

simple gross return 2-เดือน

$$\underline{\underline{1 + R_t(2)}} = \frac{P_t}{P_{t-2}} = \boxed{\frac{P_t}{P_{t-1}}} \cdot \boxed{\frac{P_{t-1}}{P_{t-2}}}$$

$$\boxed{1 + R_t(2) = (1 + R_t(1))(1 + R_{t-1}(1))}$$

ขาย

ระยะเวลาถือ

two-month simple gross return = ผลคูณของ one-month simple gross return แต่ละเดือน

ผลได้ตอบแทนหลายเดือน

Net Return ของ

$\neq R_t(2)$

ผลได้ตอบแทนรวมอย่างง่ายสองเดือนจะเท่ากับ

$$\underline{1+R_t(2)=(1+R_t)(1+R_{t-1})}=1+\underline{R_t+R_{t-1}}+\boxed{R_tR_{t-1}} \tag{4}$$

ซึ่งคือการรวมเรขาคณิต (geometric sum) ของผลได้ตอบแทนรวมอย่างง่ายหนึ่งเดือนของเดือน t และ $t-1$ ดังนั้นจะเห็นได้ว่า $R_t(2)$ จะไม่เท่ากับผลรวมของ R_t และ R_{t-1}

simple net return ของการลงทุนหลายเดือน (2)

≠ ผลรวมของ simple net return แต่ละเดือน.

ตัวอย่างที่ 1.2

พิจารณาต่อเนื่องจากตัวอย่าง 1.1 สมมุติให้เราซื้อหุ้น ณ เดือนที่ $t-2$ ด้วยราคา $P_{t-2} = 180$ ดอลลาร์และไม่มีการจ่ายเงินปันผล ผลได้ตอบแทนสุทธิสองเดือนจะเท่ากับ

$t-2 \to t$

Simple net return $R_t(2) = \frac{P_t - P_{t-2}}{P_{t-2}} = \frac{200-180}{180} \approx 0.1111$ (11.11%)

Simple Gross return. $1 + R_t(2) = 1.1111$

ตัวอย่างที่ 1.2

โดยที่ผลได้ตอบแทนหนึ่งเดือนของแต่ละเดือนจะเท่ากับ

$t-2 \to t-1$

$$1 + R_{t-1} = \frac{190}{180} = 1.0556$$

$t-1 \to t$

$$1 + R_t = \frac{200}{190} = 1.0526$$

และผลได้ตอบแทนรวมสองเดือนจะเท่ากับ

$$1 + R_t(2) = (1+R_{t-1})(1+R_t) = (1.0556)(1.0526) = 1.1111$$

$$R_t + R_{t-1} = 0.0526 + 0.0556$$

$$= 0.1082 \neq 0.1111 = R_t(2)$$

ผลได้ตอบแทนของกลุ่มสินทรัพย์ลงทุน (Portfolio)

- ลงทุนด้วยเงินจำนวน V บาทในสินทรัพย์สองตัวคือ A และ B
- สัดส่วนการลงทุนในสินทรัพย์ทั้งสองคือ x_A และ x_B $x_A + x_B = 1$
- ผลได้ตอบแทนอย่างง่ายหนึ่งเดือนของ A และ B คือ $R_{A,t}$ และ $R_{B,t}$
- มูลค่าของกลุ่มสินทรัพย์ลงทุนเมื่อสิ้นเดือนจะเท่ากับ

$$V \times [x_A(1 + R_{A,t}) + x_B(1 + R_{B,t})]$$

- ผลได้ตอบแทนรวมของกลุ่มสินทรัพย์ลงทุน และอัตราผลตอบแทนจากการลงทุนจะเท่ากับ

$$x_A R_{A,t} + x_B R_{B,t}$$

การปรับกรณีเงินปันผล

- ถ้าสินทรัพย์มีการจ่ายเงินปันผลเท่ากับ D_t ในช่วงเวลาระหว่างเดือน t และ $t-1$

การปรับกรณีเงินปันผล

- ถ้าสินทรัพย์มีการจ่ายเงินปันผลเท่ากับ D_t ในช่วงเวลาระหว่างเดือน t และ $t-1$
- การคำนวณผลได้ตอบแทนสุทธิทั้งหมด (total net return) สามารถทำได้โดย

$$R_t^{total} = \frac{(P_t + D_t) - P_{t-1}}{P_{t-1}} = \frac{P_t - P_{t-1}}{P_{t-1}} + \frac{D_t}{P_{t-1}} \tag{5}$$

การปรับกรณีเงินปันผล

- ถ้าสินทรัพย์มีการจ่ายเงินปันผลเท่ากับ D_t ในช่วงเวลาระหว่างเดือน t และ $t-1$
- การคำนวณผลได้ตอบแทนสุทธิทั้งหมด (total net return) สามารถทำได้โดย

$$R_t^{total} = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \boxed{\frac{P_t - P_{t-1}}{P_{t-1}}} + \boxed{\frac{D_t}{P_{t-1}}} \qquad (5)$$

- โดยที่ส่วนแรกเป็นกำไรส่วนเกินทุน (capital gain) และส่วนที่สองคืออัตราผลตอบแทนเงินปันผล (dividend yield)

1-year

หุ้น Simple return 6%

1 เดือน ?

การแปลงผลได้ตอบแทนเป็นผลได้ตอบแทนรายปี

- การแปลงผลได้ตอบแทนที่มีความถี่อื่นๆเป็นรายปี สมมุติว่าเราต้องการหาผลได้ตอบแทนหนึ่งปี $(1+R_A)$ จากข้อมูลผลได้ตอบแทนรายเดือน เราสามารถคำนวณได้โดย

$$\begin{aligned} 1+R_A = 1+R_t(12) &= \frac{P_t}{P_{t-12}} = \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-11}}{P_{t-12}} \\ &= (1+R_t)(1+R_{t-1})\cdots(1+R_{t-11}) \\ R_A &= R_t(12) \end{aligned}$$

การแปลงผลได้ตอบแทนเป็นผลได้ตอบแทนรายปี

- การแปลงผลได้ตอบแทนที่มีความถี่อื่นๆเป็นรายปี สมมุติว่าเราต้องการหาผลได้ตอบแทนหนึ่งปี $(1 + R_A)$ จากข้อมูลผลได้ตอบแทนรายเดือน เราสามารถคำนวณได้โดย

$$1 + R_A = 1 + R_t(12) = \frac{P_t}{P_{t-12}} = \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-11}}{P_{t-12}}$$

$$= (1 + \underbrace{R_t}_{= R})(1 + \underbrace{R_{t-1}}_{= R})\cdots(1 + \underbrace{R_{t-11}}_{= R})$$

$$R_A = R_t(12)$$

- การคำนวณผลได้ตอบแทนรายเดือนภายใต้ข้อสมมุติว่าผลได้ตอบแทนคงที่ทุกเดือนเท่ากับ R เราจะได้ผลได้ตอบแทนรายหนึ่งปีเท่ากับ

$R_A = 6\%$

รายเดือน $R = 0.048$

$$\underbrace{1 + \overset{1) \; 6\%}{R_A}} = 1 + R_t(12) = \underbrace{(1 + R)^{12}}$$

$(1+R_A) = (1+R)^{12} \Longrightarrow 1+R = (1+R_A)^{1/12}$

$1+R = (1+0.06)^{1/12}$

ซึ่งก็คือผลได้ตอบแทนทบต้น 12 เดือนนั่นเอง

ผลได้ตอบแทนทบต้นอย่างต่อเนื่อง

การคิดผลตอบแทนทบต้น(compounding) สมมุติว่าธนาคารจ่ายดอกเบี้ยเงินฝากในอัตรา 10 % ต่อปี และมีเงินต้นเท่ากับ 100 บาท

- จ่ายดอกเบี้ยหนึ่งครั้งเมื่อสิ้นปี มูลค่าสุทธิของเงินฝากเท่ากับ $100(1+0.1) = 110$ บาท

ผลได้ตอบแทนทบต้นอย่างต่อเนื่อง

การคิดผลตอบแทนทบต้น(compounding) สมมุติว่าธนาคารจ่ายดอกเบี้ยเงินฝากในอัตรา 10 % ต่อปี และมีเงินต้นเท่ากับ 100 บาท

- จ่ายดอกเบี้ยหนึ่งครั้งเมื่อสิ้นปี มูลค่าสุทธิของเงินฝากเท่ากับ $100(1+0.1) = 110$ บาท
- จ่ายดอกเบี้ยออกเป็นสองครั้งครั้งละ 5 % ทุกครึ่งปี มูลค่าสุทธิจะเท่ากับ $100(1+0.1/2)^2 = 110.25$ บาท

100

$100(1+0.05)$

$100(1+0.05)(1+0.05) =$

$100\left(1+\frac{0.1}{2}\right)^{2}$

ผลได้ตอบแทนทบต้นอย่างต่อเนื่อง

การคิดผลตอบแทนทบต้น(compounding) สมมุติว่าธนาคารจ่ายดอกเบี้ยเงินฝาก ในอัตรา 10 % ต่อปี และมีเงินต้นเท่ากับ 100 บาท

- จ่ายดอกเบี้ยหนึ่งครั้งเมื่อสิ้นปี มูลค่าสุทธิของเงินฝากเท่ากับ $100(1+0.1) = 110$ บาท
- จ่ายดอกเบี้ยออกเป็นสองครั้งครั้งละ 5 % ทุกครึ่งปี มูลค่าสุทธิจะเท่ากับ $100(1+0.1/2)^2 = 110.25$ บาท
- จ่ายดอกเบี้ย m ครั้ง มูลค่าสุทธิ ณ สิ้นปีที่หนึ่งจะมีค่า $100(1+0.1/m)^m$

$$\lim_{m\to\infty} 100\left(1+\frac{0.1}{m}\right)^m = 100 \boxed{\lim_{m\to\infty}\left(1+\frac{0.1}{m}\right)^m}$$

$$\exp(0.1)$$

มูลค่าสุทธิ $= 100 \cdot \exp(0.1)$

ผลได้ตอบแทนทบต้นอย่างต่อเนื่อง

การคิดผลตอบแทนทบต้น(compounding) สมมุติว่าธนาคารจ่ายดอกเบี้ยเงินฝาก ในอัตรา 10 % ต่อปี และมีเงินต้นเท่ากับ 100 บาท

- จ่ายดอกเบี้ยหนึ่งครั้งเมื่อสิ้นปี มูลค่าสุทธิของเงินฝากเท่ากับ $100(1+0.1)=110$ บาท
- จ่ายดอกเบี้ยออกเป็นสองครั้งครั้งละ 5 % ทุกครึ่งปี มูลค่าสุทธิจะเท่ากับ $100(1+0.1/2)^2=110.25$ บาท
- จ่ายดอกเบี้ย m ครั้ง มูลค่าสุทธิ ณ สิ้นปีที่หนึ่งจะมีค่า $100(1+0.1/m)^m$
- จ่ายดอกเบี้ยทบต้นอย่างต่อเนื่อง (continuous compounding) มูลค่าสุทธิ ณ สิ้นปี จะเท่ากับ $\boxed{100}(\exp(0.1))=\boxed{110.517}$
  Note: $\exp(x)=\lim_{n\to\infty}(1+x/n)^n$

เงินต้น

P_{t-1}

เงินที่ได้สิ้น

P_t

r_t แทน continuous compounding return

$$\boxed{P_{t-1}(\exp(r_t))=P_t}$$

$$\underbrace{\ln(\exp(r_t))}_{r_t}=\ln\left(\frac{P_t}{P_{t-1}}\right) \Leftarrow \exp(r_t)=\frac{P_t}{P_{t-1}}$$

กำหนดให้ R_t เป็นผลได้ตอบแทนอย่างง่ายรายเดือนของการลงทุน เนื่องจากมูลค่าในอนาคตจะเท่ากับมูลค่าปัจจุบันคูณกับตัวคูณทบต้น ($P_t = P_{t-1}\exp(r_t)$) เราสามารถคำนวณผลได้ตอบแทนทบต้นต่อเนื่องหนึ่งเดือน (one-month continuously compounding return) หรือผลได้ตอบแทนในรูปลอการิทึม(one-month log return) ได้โดย

log return → $r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(1+R_t)$ ← simple (net) return.

$$r_t = \ln(1+R_t) \overset{①}{=} \ln\left(\frac{P_t}{P_{t-1}}\right)$$
$$\overset{②}{=} \ln(P_t) - \ln(P_{t-1}) = p_t - p_{t-1} \quad (6)$$

โดยที่ $p_t = \ln(P_t)$

ตัวอย่างที่ 1.3

$P_t = 200$

$P_{t-1} = 190$

ใช้ข้อมูลจากตัวอย่างที่ 1.1 ผลได้ตอบแทนทบต้นอย่างต่อเนื่องหนึ่งเดือนหรือผลได้ตอบแทนในรูปลอการิทึมหนึ่งเดือนจะเท่ากับ

① $$r_t = \ln(1.0526) = 0.0513 \quad (5.13\%)$$

② $$r_t = \ln(200) - \ln(190) = 5.2983 - 5.2470 = 0.0513 \quad (5.13\%)$$

simple return 0.0526

$$r_t = \ln\left(\frac{200}{190}\right)$$

ความสัมพันธ์ระหว่าง R_t กับ r_t

$$\boxed{r_t} = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(1+R_t)$$

$$\exp(r_t) = \exp(\ln(1+R_t))$$

$$\exp(r_t) = 1+R_t$$

$$\boxed{\exp(r_t) - 1 = R_t}$$

ความสัมพันธ์ระหว่าง simple return กับ log return

- เราสามารถคำนวณผลตอบแทนอย่างง่ายสุทธิได้จากผลได้ตอบแทนในรูปลอการิทึมโดย

$$R_t = \exp(r_t) - 1$$

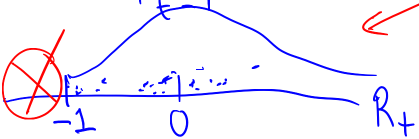
$$P_{t-1} = 100 \qquad P_t = 0$$

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = -1 \quad (-100\%)$$

ค่าของตัวแปร $(-\infty, \infty)$

ถ้าเอาจะเป็น Normal distribution มาอธิบาย simple net return ก็ไม่เหมาะสม

ไม่มีโอกาส ที่ $R_t < -1$



ความสัมพันธ์ระหว่าง simple return กับ log return

- เราสามารถคำนวณผลตอบแทนอย่างง่ายสุทธิได้จากผลได้ตอบแทนในรูปลอการิทึมโดย

$$R_t = \exp(r_t) - 1$$

- หากผลได้ตอบแทนมีค่าต่ำ(ในกรณีเราพิจารณาผลได้ตอบแทนรายวันหรือรายเดือน) ผลได้ตอบแทนในรูปลอการิทึมและผลได้ตอบแทนอย่างง่ายจะมีค่าใกล้เคียงกัน

ความสัมพันธ์ระหว่าง simple return กับ log return

- เราสามารถคำนวณผลตอบแทนอย่างง่ายสุทธิได้จากผลได้ตอบแทนในรูปลอการิทึมโดย

$$R_t = \exp(r_t) - 1$$

- หากผลได้ตอบแทนมีค่าต่ำ(ในกรณีเราพิจารณาผลได้ตอบแทนรายวันหรือรายเดือน) ผลได้ตอบแทนในรูปลอการิทึมและผลได้ตอบแทนอย่างง่ายจะมีค่าใกล้เคียงกัน

$R_t \geq -1$

- ค่าต่ำสุดที่จะเป็นไปได้ของผลได้ตอบแทนอย่างง่ายคือ -1 ดังนั้นขอบเขตล่างของผลได้ตอบแทนในรูปลอการิทึมจะเท่ากับ $-\infty$

ค่าที่ต่ำที่สุด ของ r_t (log return)

$r_t = \ln(1+R_t) = \ln(1+(-1)) = \ln(0) = -\infty$

$R_t = \exp(-2)-1$

ค่าต่ำที่สุด (-1)

ค่าของ $r_t \to (-\infty, \infty)$

$r_t = -2$

เหตุผลว่าทำไมเราใช้ นิยาม log return ในการ
คำนวณ → แพร่หลาย

ความสัมพันธ์ระหว่าง simple return กับ log return

- เราสามารถคำนวณผลตอบแทนอย่างง่ายสุทธิได้จากผลได้ตอบแทนในรูปลอการิทึมโดย
$$R_t = \exp(r_t) - 1$$
- หากผลได้ตอบแทนมีค่าต่ำ(ในกรณีเราพิจารณาผลได้ตอบแทนรายวันหรือรายเดือน) ผลได้ตอบแทนในรูปลอการิทึมและผลได้ตอบแทนอย่างง่ายจะมีค่าใกล้เคียงกัน
- ค่าต่ำสุดที่จะเป็นไปได้ของผลได้ตอบแทนอย่างง่ายคือ -1 ดังนั้นขอบเขตล่างของผลได้ตอบแทนในรูปลอการิทึมจะเท่ากับ $-\infty$
- การวิเคราะห์เชิงสถิติหรือการสร้างแบบจำลองเรามักจะใช้ผลได้ตอบแทนในรูปลอการิทึม

$\boxed{\text{log - return}}$

การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่ม Y_t คือฟังก์ชันที่ใช้อธิบายค่าของเหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยที่เรารู้ค่าทั้งหมดที่จะเป็นไปได้แต่ไม่รู้ว่าจะเกิดอะไรขึ้นแน่นอน เช่น ราคาของหลักทรัพย์(P)

การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่ม Y_t คือฟังก์ชันที่ใช้อธิบายค่าของเหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยที่เรารู้ค่าทั้งหมดที่จะเป็นไปได้แต่ไม่รู้ว่าจะเกิดอะไรขึ้นแน่นอน เช่น ราคาของหลักทรัพย์(P)
- ตัวแปรสุ่มจะมีการแจกแจงอย่างไร โดยที่หนึ่งในทางเลือกคือการแจกแจงแบบ log-normal

การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่ม Y_t คือฟังก์ชันที่ใช้อธิบายค่าของเหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยที่เรารู้ค่าทั้งหมดที่จะเป็นไปได้แต่ไม่รู้ว่าจะเกิดอะไรขึ้นแน่นอน เช่น ราคาของหลักทรัพย์(P)
- ตัวแปรสุ่มจะมีการแจกแจงอย่างไร โดยที่หนึ่งในทางเลือกคือการแจกแจงแบบ log-normal
- การลงทุนของเราในหนึ่งเดือนข้างหน้าจะมีผลได้ตอบแทน(R_t)เป็นอย่างไร ดังนั้น R_t ก็จะเป็นตัวแปรสุ่ม

การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่ม Y_t คือฟังก์ชันที่ใช้อธิบายค่าของเหตุการณ์ที่จะเกิดขึ้นในอนาคต โดยที่เรารู้ค่าทั้งหมดที่จะเป็นไปได้แต่ไม่รู้ว่าจะเกิดอะไรขึ้นแน่นอน เช่น ราคาของหลักทรัพย์(P)
- ตัวแปรสุ่มจะมีการแจกแจงอย่างไร โดยที่หนึ่งในทางเลือกคือการแจกแจงแบบ log-normal
- การลงทุนของเราในหนึ่งเดือนข้างหน้าจะมีผลได้ตอบแทน(R_t)เป็นอย่างไร ดังนั้น R_t ก็จะเป็นตัวแปรสุ่ม
- การแจกแจงที่ใช้ประมาณค่าสำหรับผลได้ตอบแทนอย่างง่ายรายเดือนที่ดีคือการแจกแจงแบบปกติ

การแจกแจงของอนุกรมเวลา

ขึ้น +1

คง −1

- ราคาจะเปลี่ยนแปลงอย่างไร โดยอาจจะสนใจกรณีที่ราคาปรับตัวขึ้นแทนด้วย 0 และกรณีที่ราคาคงที่หรือลดลงแทนด้วย 1 ในกรณีนี้ sample space ของ discrete random variable

$$P(X) = \begin{cases} 0.5 & x = -1 \\ 0.5 & x = +1 \end{cases}$$

การแจกแจงของอนุกรมเวลา

- ราคาจะเปลี่ยนแปลงอย่างไร โดยอาจจะสนใจกรณีที่ราคาปรับตัวขึ้นแทนด้วย 0 และกรณีที่ราคาคงที่หรือลดลงแทนด้วย 1 ในกรณีนี้ sample space ของ discrete random variable
- ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function: *pdf*) ของตัวแปรสุ่มวิยุตสามารถเขียนแทนด้วย $p(y)$ จะเป็นฟังก์ชัน
$p(y) = Pr(Y = y)$

การแจกแจงของอนุกรมเวลา

- ราคาจะเปลี่ยนแปลงอย่างไร โดยอาจจะสนใจกรณีที่ราคาปรับตัวขึ้นแทนด้วย 0 และกรณีที่ราคาคงที่หรือลดลงแทนด้วย 1 ในกรณีนี้ sample space ของ discrete random variable
- ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function: *pdf*) ของตัวแปรสุ่มวิยุตสามารถเขียนแทนด้วย $p(y)$ จะเป็นฟังก์ชัน $p(y) = Pr(Y = y)$
- *pdf* จะต้องมีคุณสมบัติคือ (1) $p(y) \geq 0$ สำหรับทุกค่า $y \in S_y$ (2) $p(y) = 0$ สำหรับทุกค่า $y \notin S_y$ และ (3) $\sum_{y \in S_y} p(y) = 1$

การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่มแบบต่อเนื่อง Y จะมีค่าเป็นเท่าใดก็ได้บนเส้นจำนวนจริง ดังนั้น $S_Y = \{y : y \in \mathbb{R}\}$

$r_t \in (-\infty, \infty)$

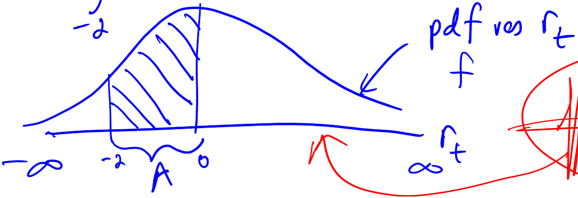
การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่มแบบต่อเนื่อง Y จะมีค่าเป็นเท่าใดก็ได้บนเส้นจำนวนจริง ดังนั้น $S_Y = \{y : y \in \mathbb{R}\}$
- ฟังก์ชันความหนาแน่นของความน่าจะเป็น *pdf* ของตัวแปรสุ่มต่อเนื่อง *Y* จะเป็นฟังก์ชันที่ไม่เป็นค่าลบ *f* นิยามบนเส้นจำนวนจริงโดยที่สำหรับช่วง *A* ใดๆ

$$\boxed{Pr(Y \in A) = \int_A f(y)dy}$$

ความน่าจะเป็นของเหตุการณ์ $Y \in A$ เท่ากับพื้นที่ใต้ function (pdf) ตลอดช่วง A.

$$\boxed{P(-2 < r_t < 0)} = \int_{-2}^{0} f(r_t)\, dr_t$$



การแจกแจงของอนุกรมเวลา

- ตัวแปรสุ่มแบบต่อเนื่อง Y จะมีค่าเป็นเท่าใดก็ได้บนเส้นจำนวนจริง ดังนั้น $S_Y = \{y : y \in \mathbb{R}\}$
- ฟังก์ชันความหนาแน่นของความน่าจะเป็น *pdf* ของตัวแปรสุ่มต่อเนื่อง Y จะเป็นฟังก์ชันที่ไม่เป็นค่าลบ f นิยามบนเส้นจำนวนจริงโดยที่สำหรับช่วง A ใดๆ

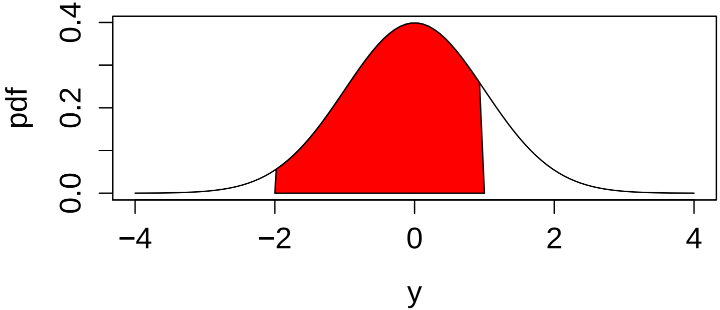
$$Pr(Y \in A) = \int_A f(y)dy$$

- $Pr(Y \in A)$ คือพื้นที่ใต้กราฟความน่าจะเป็นตลอดช่วง A โดยที่ *pdf* $f(y)$ จะมีคุณสมบัติดังนี้ (1) $f(y) \geq 0$ และ (2) $\int_{-\infty}^{\infty} f(y)dy = 1$

การแจกแจงของอนุกรมเวลา:การแจกแจงต่อเนื่อง

ตัวอย่างเช่นกราฟรูประฆังรูปที่ 1 เป็น pdf ฟังก์ชันโดยที่พื้นที่ใต้กราฟระหว่าง $y = -2$ ถึง $y = 1$ จะแสดงถึง $Pr(-2 \leq Y < 1)$

Figure: ฟังก์ชัน pdf ของการแจกแจงต่อเนื่อง



การแจกแจงของอนุกรมเวลา:การแจกแจงต่อเนื่อง

ฟังก์ชันการแจกแจงสะสม (*cdf*) จะมีคุณสมบัติดังนี้

1. ถ้า $y_1 < y_2$ แล้ว $F_Y(y_1) \leq F_Y(y_2)$
2. $F_Y(-\infty) = 0$ และ $F_Y(\infty) = 1$
3. $Pr(Y > y) = 1 - F_Y(y)$
4. $Pr(y_1 < X \leq y_2) = F_Y(y_2) - F_Y(y_1)$
5. $\frac{\partial F_Y(y)}{\partial y} = f_Y(y)$ ถ้า Y เป็นตัวแปรสุ่มต่อเนื่องและ $F_Y(y)$ เป็นฟังก์ชันต่อเนื่องและหาอนุพันธ์ได้

ควอนไทล์ของการแจกแจงความน่าจะเป็น

- หากเราพิจารณาตัวแปรสุ่ม Y ที่มีฟังก์ชันการแจกแจงสะสม $F_Y(y)$ หากค่า $0 \leq \alpha \leq 1$ แล้วค่าควอนไทล์ที่ $100\alpha\%$ ของการแจกแจง Y คือค่า q_α ที่ตรงกับเงื่อนไข

$$F_Y(q_\alpha) = Pr(Y \leq q_\alpha) = \alpha$$

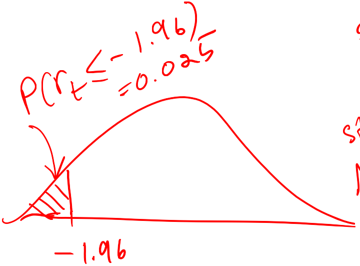
ควอนไทล์ของการแจกแจงความน่าจะเป็น

- หากเราพิจารณาตัวแปรสุ่ม Y ที่มีฟังก์ชันการแจกแจงสะสม $F_Y(y)$ หากค่า $0 \leq \alpha \leq 1$ แล้วค่าควอนไทล์ที่ $100\alpha\%$ ของการแจกแจง Y คือค่า q_α ที่ตรงกับเงื่อนไข

$$F_Y(q_\alpha) = Pr(Y \leq q_\alpha) = \alpha$$

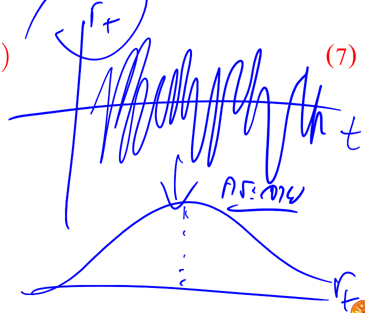
- กำหนดให้ $Y \sim N(0,1)$ ค่าควอนไทล์ของการแจกแจงแบบปกติมาตรฐานจะเท่ากับ

$$q_\alpha = \Phi^{-1}(\alpha) \quad (7)$$



standard Normal

r_t



คุณลักษณะเรื่องรูปร่างของการแจกแจงความน่าจะเป็น

คุณลักษณะด้านรูปร่างของการแจกแจงความน่าจะเป็นใดๆจะวัดได้ด้วยคุณลักษณะ 4 ประการ

1. ค่าคาดหมาย (expected value) หรือค่าเฉลี่ยเป็นการวัดค่ากลางของการแจกแจง

คุณลักษณะเรื่องรูปร่างของการแจกแจงความน่าจะเป็น

คุณลักษณะด้านรูปร่างของการแจกแจงความน่าจะเป็นใดๆจะวัดได้ด้วยคุณลักษณะ 4 ประการ

1. ค่าคาดหมาย (expected value) หรือค่าเฉลี่ยเป็นการวัดค่ากลางของการแจกแจง
2. ความแปรปรวนและค่าเบี่ยงเบนมาตรฐานวัดการแผ่จากค่ากลาง

คุณลักษณะเรื่องรูปร่างของการแจกแจงความน่าจะเป็น

คุณลักษณะด้านรูปร่างของการแจกแจงความน่าจะเป็นใดๆจะวัดได้ด้วยคุณลักษณะ 4 ประการ

1. ค่าคาดหมาย (expected value) หรือค่าเฉลี่ยเป็นการวัดค่ากลางของการแจกแจง
2. ความแปรปรวนและค่าเบี่ยงเบนมาตรฐานวัดการแผ่จากค่ากลาง
3. ความเบ้ (skewness) วัดความสมมาตรรอบๆค่ากลาง

คุณลักษณะเรื่องรูปร่างของการแจกแจงความน่าจะเป็น

คุณลักษณะด้านรูปร่างของการแจกแจงความน่าจะเป็นใดๆจะวัดได้ด้วยคุณลักษณะ 4 ประการ

1. ค่าคาดหมาย (expected value) หรือค่าเฉลี่ยเป็นการวัดค่ากลางของการแจกแจง
2. ความแปรปรวนและค่าเบี่ยงเบนมาตรฐานวัดการแผ่จากค่ากลาง
3. ความเบ้ (skewness) วัดความสมมาตรรอบๆค่ากลาง
4. ค่าความโด่ง (kurtosis) วัดความอ้วนของหาง

ค่าคาดหมาย

ฟังก์ชันค่าคาดหมาย (mean function)ของตัวแปรสุ่ม Y ใดๆ ใช้สัญลักษณ์ $E(Y)$ สามารถคำนวณได้ดังนี้ ในกรณีที่ Y เป็น discrete r.v. ค่าคาดหมายจะเท่ากับ

$$\mu_Y = E(Y) = \sum_{y \in S_Y} y \times P(Y = y) \tag{8}$$

หรือในกรณี continuous r.v. ค่าคาดหมายจะเท่ากับ

$$\mu_Y = E(Y) = \int_{-\infty}^{\infty} y f(y) dy \tag{9}$$

โดยที่ E คือสัญลักษณ์แทนค่าคาดหมาย (Expected value)

ค่าความแปรปรวนและค่าเบี่ยงเบนมาตรฐาน

ค่าแปรปรวนของตัวแปรสุ่ม Y เขียนแทนด้วย $Var(Y)$ หรือ σ_Y^2 วัดการแผ่ของการแจกแจงจากค่าเฉลี่ย โดยที่ค่าแปรปรวนสามารถนิยามได้โดย

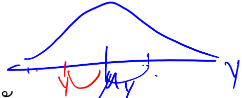
$$\sigma_Y^2 = Var(Y) = E\left[(Y - \mu_Y)^2\right] \tag{10}$$

หรือ

$$\sigma_Y^2 = Var(Y) = E\left[Y^2\right] - \mu_Y^2 \tag{11}$$

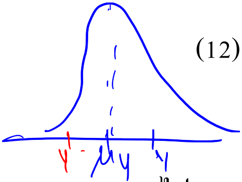
เนื่องจากหน่วยวัดของค่าความแปรปรวนมีหน่วยที่แตกต่างจากตัวแปรสุ่ม Y ดังนั้น เรามักจะพิจารณาค่าเบี่ยงเบนมาตรฐานเขียนแทนด้วย $sd(Y)$ หรือ σ_Y ซึ่งเท่ากับค่ารากที่สองของค่าความแปรปรวน ($\sqrt{\sigma_Y^2}$)

ค่าความเบ้(skewness)



ค่าความเบ้ซึ่งสามารถเขียนแทนด้วย ($S(Y)$) วัดความสมมาตรรอบๆค่ากลาง สามารถทำได้โดย

$$S(Y) = \frac{E[(Y-\mu_Y)^3]}{\sigma_Y^3} \tag{12}$$



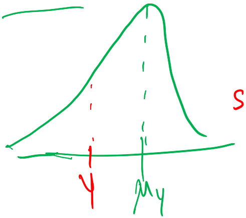
$S(Y)=0$

- ค่าความเบ้เป็นศูนย์แสดงว่าการแจกแจงมีความสมมาตร
- ค่าความเบ้เป็นบวกแสดงว่าข้อมูลถูกดึงไปด้านขวาของการแจกแจง(หางไปทางขวา)

  $S(Y) > 0$ Positive Skewness [ถ้ามี positive skewness

  ↓

  มีโอกาสเกิด
- ค่าความเบ้เป็นลบแสดงว่าข้อมูลถูกดึงไปทางซ้าย(หางไปทางซ้าย)



$S(Y) < 0$ Negative Skewness

[มีโอกาสที่ $y < \mu_Y$ มากกว่า $y > \mu_Y$]

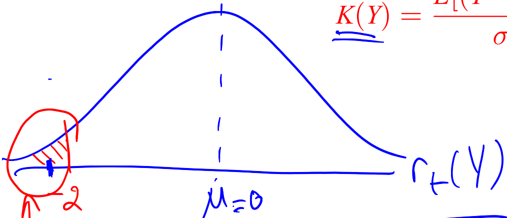
ค่าความโด่ง(kurtosis)

- ค่าความโด่งใช้วัดความหนาของหางของการแจกแจงซึ่งสามารถเขียนแทนได้ด้วย $K(Y)$ และสามารถคำนวณได้จาก

$$K(Y) = \frac{E[(Y-\mu_Y)^4]}{\sigma_Y^4} \tag{13}$$

$(Y-\mu_Y)^4$ มาก

ก็ Y ห่างจาก μ_Y



พื้นที่ปลายหาง (tail)

$\rightarrow P(r_t < -2)$

ค่าความโด่ง(kurtosis)

- ค่าความโด่งใช้วัดความหนาของหางของการแจกแจงซึ่งสามารถเขียนแทนได้ด้วย $K(Y)$ และสามารถคำนวณได้จาก

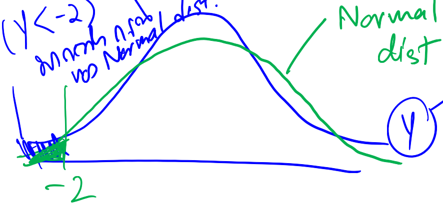
$$\boxed{K(Y) = \frac{E[(Y-\mu_Y)^4]}{\sigma_Y^4}} \tag{13}$$

K(Y) > 3

ตัวแปร Y มีหางที่อ้วนกว่า Normal dist.

- ค่าความโด่งของตัวแปรสุ่มที่มีการแจกแจงปกติเท่ากับ $\boxed{3}$

P(Y<-2) มากกว่ากรณีของ Normal dist.



K(Y) > 3

ค่าความโด่ง(kurtosis)

- ค่าความโด่งใช้วัดความหนาของหางของการแจกแจงซึ่งสามารถเขียนแทนได้ด้วย $K(Y)$ และสามารถคำนวณได้จาก

$$K(Y) = \frac{E[(Y - \mu_Y)^4]}{\sigma_Y^4} \tag{13}$$

- ค่าความโด่งของตัวแปรสุ่มที่มีการแจกแจงปกติเท่ากับ 3
- เราใช้ค่าความโด่งดังกล่าวเป็นมาตรฐานความหนาของหาง

ค่าความโด่ง(kurtosis)

- ค่าความโด่งใช้วัดความหนาของหางของการแจกแจงซึ่งสามารถเขียนแทนได้ด้วย $K(Y)$ และสามารถคำนวณได้จาก

$$K(Y) = \frac{E[(Y-\mu_Y)^4]}{\sigma_Y^4} \tag{13}$$

- ค่าความโด่งของตัวแปรสุ่มที่มีการแจกแจงปกติเท่ากับ 3
- เราใช้ค่าความโด่งดังกล่าวเป็นมาตรฐานความหนาของหาง
- ความโด่งส่วนเกิน (excess kurtosis) โดย excess $K(Y) = K(Y) - 3$

ดูแนวรูปร่างของ dist

Normal dist

$E(Y)$

$var(Y)$

$S(Y) = 0$

$K(Y) = 3$

Population Parameters

สมมุติ dist → parameters.

การแจกแจงแบบที

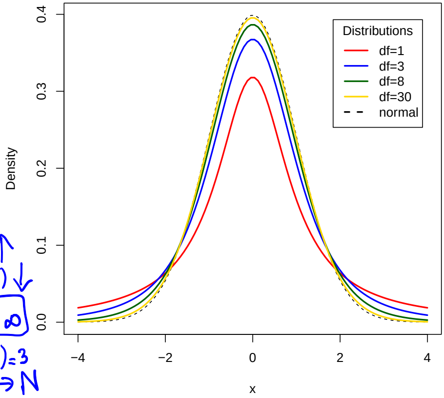
- การแจกแจงอื่นที่มีลักษณะสมมาตรและมีหางที่อ้วนกว่าการแจกแจงแบบปกติ คือ **การแจกแจงแบบที (Student's t)**

การแจกแจงแบบที

- การแจกแจงอื่นที่มีลักษณะสมมาตรและมีหางที่อ้วนกว่าการแจกแจงแบบปกติ คือ **การแจกแจงแบบที (Student's t)**
- ถ้า Y มีการแจกแจกแบบทีด้วยองศาเสรี (degree of freedom) v จะมีค่าเฉลี่ยเท่ากับ 0, ค่าความแปรปรวนเท่ากับ $v/v-2$ โดยที่ $v>2$, ค่าความเบ้เท่ากับ 0, และค่าความโด่งเท่ากับ $\frac{6}{(v-4)}-4$ โดยที่ $v>4$

การแจกแจงแบบที

Figure: การแจกแจงแบบที



df ↑
$K(\nu)$ ↓
$\boxed{df \to \infty}$
$K(\nu) = 3$
$t \to N$

ตัวประมาณค่า

สมมุติให้เราสุ่มตัวอย่าง $y_1, \ldots, y_T$ ซึ่งมีขนาดเท่ากับ T

- ค่าเฉลี่ยของตัวอย่าง (sample mean) $\hat{\mu}_Y = \frac{1}{T}\sum_{t=1}^{T} y_t$

$\bar{r}_t = 0.001 \rightarrow \mu(r_t) = 0$

sample

population mean

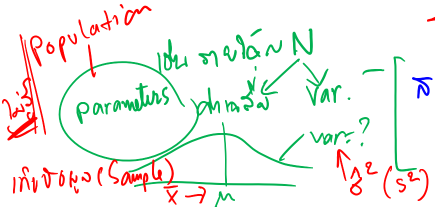
ในโลกความเป็นจริง

เราไม่ทราบทั้ง distribution + parameters ของประชากร (population)

$\Downarrow$

สมมุติ รูปแบบ ของการแจกแจง (distribution)

แล้ว ประมาณค่า parameters



ตัวประมาณค่า

$\bar{X}$ (sample) อ้างอิง μ (ทดสอบสมมติฐาน)

y_1, y_2, ⋮, y_T

สมมุติให้เราสุ่มตัวอย่าง $y_1, \ldots, y_T$ ซึ่งมีขนาดเท่ากับ T

estimator

- ค่าเฉลี่ยของตัวอย่าง (sample mean) $\hat{\mu}_Y = \frac{1}{T}\sum_{t=1}^{T} y_t$
- ค่าความแปรปรวนของตัวอย่าง $\hat{\sigma}_Y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^2$

sample SD $= \sqrt{\hat{\sigma}_Y^2}$

ตัวประมาณค่า

สมมุติให้เราสุ่มตัวอย่าง $y_1, \ldots, y_T$ ซึ่งมีขนาดเท่ากับ T

- ค่าเฉลี่ยของตัวอย่าง (sample mean) $\hat{\mu}_Y = \frac{1}{T}\sum_{t=1}^{T} y_t$
- ค่าความแปรปรวนของตัวอย่าง $\hat{\sigma}_Y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^2$
- ค่าความเบ้ของตัวอย่าง (sample skewness) $\hat{S}(Y) = \frac{1}{(T-1)\hat{\sigma}_Y^3}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^3$

ตัวประมาณค่า $\longrightarrow$ อ้างอิง ถึง population parameters

สมมุติให้เราสุ่มตัวอย่าง $y_1, ..., y_T$ ซึ่งมีขนาดเท่ากับ T

- ค่าเฉลี่ยของตัวอย่าง (sample mean) $\hat{\mu}_Y = \frac{1}{T}\sum_{t=1}^{T} y_t$
- ค่าความแปรปรวนของตัวอย่าง $\hat{\sigma}_Y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^2$
- ค่าความเบ้ของตัวอย่าง (sample skewness) $\hat{S}(Y) = \frac{1}{(T-1)\hat{\sigma}_Y^3}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^3$
- ค่าความโด่งของตัวอย่าง (sample kurtosis) $\hat{K}(Y) = \frac{1}{(T-1)\hat{\sigma}_Y^4}\sum_{t=1}^{T}(y_t - \hat{\mu}_Y)^4$

การทดสอบสมมุติฐาน: ค่าเฉลี่ย (population mean).

ภายใต้ข้อสมมุติที่ว่า Y_t มีการแจกแจงแบบปกติ $N(0, \sigma^2)$ แล้ว $\bar{\mu}$ จะมีการแจกแจงแบบปกติ $N(0, \frac{\sigma^2}{T})$

$H_0 : \mu_Y = 0$ vs $H_1 : \mu_Y \neq 0$

$$t = \frac{\hat{\mu}_Y - 0}{\sqrt{\hat{\sigma}^2_Y / T}} \sim N(0,1)$$

เกณฑ์ ปฏิเสธ H_0 ถ้า $|t| > c.v. (Z_\alpha)$

ตย. หุ้น PTT จากตัวอย่าง $2088 = T$, $\hat{\mu} = 0.00077$, $\hat{\sigma} = 0.0227$

$$t = \frac{\hat{\mu} - 0}{\hat{\sigma}/\sqrt{T}} = 1.579$$

$|t| < cv \ (1.96, \alpha = 0.05)$

การทดสอบสมมุติฐาน:ค่าความเบ้(skewness) $\Rightarrow S(Y) = 0$

เมื่อกลุ่มตัวอย่างใหญ่ ค่าความเบ้ของตัวอย่าง $\hat{S}(Y) \sim N(0, 6/T)$

$H_0: S(Y) = 0$ vs $H_1: S(Y) \neq 0$

ตัวสถิติ $t = \dfrac{\hat{S}(Y) - 0}{\sqrt{6/T}} \sim N(0,1)$

เกณฑ์ ปฏิเสธ H_0 ถ้า $|t| > C.V. \ (Z_\alpha)$

ต.ย. $\hat{S} = -0.0689$

$$t = \frac{-0.0689}{\sqrt{6/2088}} = -1.2846$$

$|t| < 1.96$ (C.V ที่ระดับ $\alpha = 0.05$) เราไม่สามารถ ปฏิเสธ H_0 ที่ว่า skewness $= 0$.

การทดสอบสมมุติฐาน:ค่าความโด่ง(kurtosis)

$H_0: K = \overset{\text{Normal}}{3}$ vs $H_1: K \neq 3$

$H_0: \underbrace{K-3}_{\text{Excess Kurtosis}} = 0$ vs $H_1: K-3 \neq 0$

เมื่อกลุ่มตัวอย่างใหญ่ ค่าความโด่งส่วนเกินของตัวอย่าง

$\underbrace{\hat{K}(Y) - 3}_{\text{sample excess kurtosis}} \sim N(0, 24/T)$

$$t = \frac{[\hat{K}(Y) - 3] - 0}{\sqrt{24/T}} \sim N(0,1)$$

เกณฑ์ $|t| > CV.$ ปฏิเสธ H_0

ตย. $\hat{K}(Y) - 3 = 6.2285$

$$t = \frac{(\hat{K}-3)}{\sqrt{24/T}} = \frac{6.2285}{\sqrt{24/2088}} = 58.0959$$

เนื่องจาก $|t| > C.V.\ (1.96)$ เราสามารถปฏิเสธ H_0 และยอมรับ $K \neq 3$

$\alpha = 0.05$

ทดสอบสมมุติ r_t ของ PTT เป็น Normal ไม่เหมาะสม

[ถ้าเรา สมมุติ Normal P (ของเหตุการณ์สุดโต่ง) น้อยกว่าที่ควร]

Probability

หางไม่เท่ากับ Normal (อ้วนกว่า Normal)

การทดสอบสมมุติฐาน:การแจกแจงแบบปกติ

$S(Y)=0$

$K(Y)-3=0$

- การทดสอบว่าตัวแปรสุ่ม Y_t มีการแจกแจงแบบปกติหรือไม่ โดยหากตัวแปรสุ่มมีการแจกแจงแบบปกติค่าความเบ้และค่าความโด่งส่วนเกินจะต้องเท่ากับศูนย์
- Jarque and Bera (1987) ได้รวมการทดสอบของค่าสถิติทั้งสอบเข้าด้วยกันและเสนอตัวสถิติ

$$\left(t=\frac{\hat{S}}{\sqrt{6/T}}\right)^2 \sim N$$

$$JB = \frac{\hat{S}^2(Y)}{6/T} + \frac{[\hat{K}(Y)-3]^2}{24/T} \sim \chi^2_{df=2}$$

$$\left(t=\frac{\hat{K}-3}{\sqrt{24/T}}\right)^2$$

ซึ่งมีการแจกแจงเมื่อกลุ่มตัวอย่างใหญ่เข้าใกล้ $\chi^2_{df=2}$

การทดสอบสมมุติฐาน:การแจกแจงแบบปกติ

- การทดสอบว่าตัวแปรสุ่ม Y_t มีการแจกแจงแบบปกติหรือไม่ โดยหากตัวแปรสุ่มมีการแจกแจงแบบปกติค่าความเบ้และค่าความโด่งส่วนเกินจะต้องเท่ากับศูนย์
- Jarque and Bera (1987) ได้รวมการทดสอบของค่าสถิติทั้งสอบเข้าด้วยกันและเสนอตัวสถิติ

$$JB = \frac{\hat{S}^2(Y)}{6/T} + \frac{[\hat{K}(Y) - 3]^2}{24/T}$$

ซึ่งมีการแจกแจงเมื่อกลุ่มตัวอย่างใหญ่เข้าใกล้ $\chi^2_{df=2}$

- ปฏิเสธสมมุติฐานหลักที่ระดับนัยสำคัญ α ถ้า $JB > \chi^2_{(1-\alpha),df=2}$

C.V.

$S=0$

$K-3=0$

H_0: $Y \sim$ Normal dist

H_1: Y ไม่เป็น Normal dist

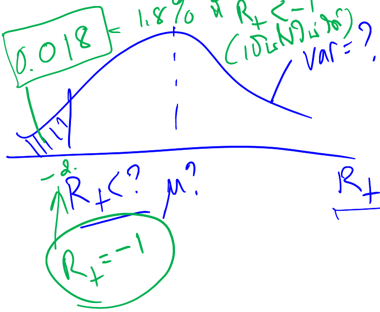
ต.ย.

$$JB = \left(\frac{-0.0689}{\sqrt{6/2088}}\right)^2 + \left(\frac{6.2285}{\sqrt{24/2088}}\right)^2 = 33.76 > \chi^2_{df=2,\alpha=0.05} = 5.99$$

JB > CV เพราะฉะนั้นปฏิเสธ Ho และสรุปว่า r_t ของ PTT ไม่เป็น Normal distribution

การแจกแจงของผลได้ตอบแทน

- เรามักจะพิจารณาผลได้ตอบแทนในรูปของล็อกและมักจะสมมุติให้มีการแจกแจงแบบปกติ
- ปัญหาในกรณีผลได้ตอบแทนอย่างง่ายมีการแจกแจงแบบปกติ $R_t \sim N(0.05, (\cancel{0.05})^2)$ 0.5 ซึ่งเราทราบว่าราคาสินทรัพย์ใดๆจะมีค่าติดลบไม่ได้ดังนั้น R_t จะต้องมีค่ามากกว่า -1



$R_t \sim N(0.05, 0.05^2)$

$P(R_t < -1)$ 100%

$= P\left(\frac{R_t - 0.05}{0.05} < \frac{-1-0.05}{0.05}\right)$ Z

dist เป็น $N(0,1)$

$= P\left(Z < \frac{-1.05}{\cancel{0.05}\,0.50}\right)$

$= P(Z < -2.1) = 0.018$

การแจกแจงของผลได้ตอบแทน

- เหมาะสมมากกว่าหากสมมุติให้ผลได้ตอบแทนในรูปล็อกมีการแจกแจงแบบปกติ $\ln(1+R_t) = r_t \sim N(0.05, (0.5)^2)$ โดยในกรณีผลได้ตอบแทนในรูปล็อกสามารถจะมีค่าน้อยกว่า -1 ได้ เช่น หาก $r_t = -2$

$P(r_t < -2) = ?$

1) แปลง $N(0,1)$

2) $P(r_t < -2) = P(R_t < -0.865)$

$P\left(\frac{r_t - 0.05}{0.5} < \frac{-2-0.05}{0.5}\right)$

$P\left(Z < \frac{-2.05}{0.5}\right)$

$R_t = \exp(r_t) - 1$

$\boxed{-0.865 = \exp(-2) - 1}$

ฟังก์ชันเส้นตรงของตัวแปรสุ่ม

ค่าคาดหวัง variance

หากกำหนดให้ X เป็นตัวแปรสุ่มที่มี $E[X] = \mu_X$ และ $Var(X) = \sigma_X^2$ และ a และ b เป็นค่าคงที่ หากเรากำหนดตัวแปรสุ่มใหม่ Y เป็นฟังก์ชันเส้นตรงกับตัวแปรสุ่ม X โดยที่ $\boxed{Y = a + bX}$ แล้ว

$$E(a+bX) = E(a)+E(bX) = a+b\,E(X)$$

- $\mu_Y = E(Y) = a + bE(X) = a + b\mu_X$
- $\sigma_Y^2 = Var(Y) = b^2 Var(X) = b^2\sigma_X^2$

$$
\begin{aligned}
Var(Y) &= E(Y-E(Y))^2 \\
&= E(\cancel{a}+bX-(\cancel{a}+b\mu_X))^2 \\
&= E[b(X-\mu_X)]^2 \\
&= E[b^2(X-\mu_X)^2] \\
&= b^2 \underbrace{E[X-\mu_X]^2}_{Var(X)}
\end{aligned}
$$

Autocovariance function

ฟังก์ชันที่ใช้ในการสรุปความสัมพันธ์ขึ้นอยู่ต่อกันตามเวลา (temporal dependence) ในข้อมูลอนุกรมเวลาคือ ฟังก์ฟันค่าแปรปรวนร่วมในตนเอง(autocovariance) ซึ่งสามารถคำนวณได้โดย

gamma

ย้อนไป k period

lag ที่ k ของ Autocovariance

$$
\begin{aligned}
\gamma_{k,t} &= Cov(Y_t, Y_{t-k}) = E\left((y_t - E(Y_t))(y_{t-k} - E(Y_{t-k}))\right) \\
&= E(y_t y_{t-k}) - E(Y_t)E(Y_{t-k}) \quad \text{for} \quad k = 0, 1, 2, \ldots
\end{aligned}
\tag{14}
$$

โดยสมการดังกล่าวสอดคล้องกับค่าแปรปรวนร่วม (covariance) ระหว่างตัวแปรสุ่มสองตัวใดๆ

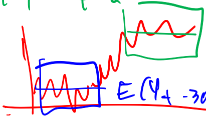
P_t & P_{t-1} $\rightarrow$ $\gamma_{1,t}$

P_{t-1} & P_{t-2} $\rightarrow$ $\gamma_{1,t-1}$

ค่าเฉลี่ย

$E(Y_t)$

$E(Y_{t-300})$



$E(Y_t)$

$E(Y_{t-300})$

→ เราไม่ต้องการ series แบบนี้ในการสร้างแบบจำลอง

Stationary

— อนุกรมเวลานิ่ง

นิยาม

Definition 1 (Strict stationary)

อนุกรมเวลา Y_t ใดๆจะถูกเรียกว่า **strictly stationary** ถ้าการแจกแจงร่วม (joint distribution) ของ $(y_{t_1}, ..., y_{t_k})$ เหมือนกันกับการแจกแจงร่วมของ $(y_{t_1+t}, ..., y_{t_k+t})$ สำหรับทุกค่าของ t

เหตุ: ใช้นิยาม

Definition 2 (weakly stationary หรือ covariance stationary)

ข้อมูลอนุกรมเวลา Y_t ใดๆ จะเรียกว่าเป็น **Weakly stationary** ถ้ามีคุณสมบัติดังนี้

1. $E(Y_t) = \mu$ ค่าเฉลี่ยคงที่ $E(Y_t) = E(Y_{t-1}) = E(Y_{t-300}) = \mu$
2. $Var(Y_t) = \sigma^2 < \infty$ ค่า variance คงที่ $Var(Y_t) = Var(Y_{t-1}) = Var(Y_{t-300})$
3. $\gamma_{k,t} = \gamma_k = Cov(Y_t, Y_{t-k})$ for all k and t

Autocovariance ที่ lag เดียวกันมีค่าเท่ากัน $[\gamma_{1,t} = \gamma_{1,t-1}]$

- แต่ γ_1, γ_2 สามารถแตกต่างกันได้

Autocorrelation function; ACF

ฟังก์ชันสหสัมพันธ์ในตัวเองที่ช้ากว่ากัน k คาบเวลา สำหรับอนุกรมที่เป็น weakly stationary จะคำนวณได้จาก

lag ที่ k rho
ของ Autocorrelation

$$\rho_k = \frac{Cov(Y_t, Y_{t-k})}{[Var(Y_t)Var(Y_{t-k})]^{1/2}}, \qquad k = 0, 1, 2, \ldots \tag{15}$$

โดยที่ $\rho_0 = 1$ และ $|\rho_k| \leq 1$ สำหรับทุกค่า k. สำหรับข้อมูลที่เป็น weakly stationary $Var(Y_t) = Var(Y_{t-k}) = \gamma_0$ ดังนั้น

Autocov ที่ lag k ห่างกัน k

$$* \boxed{\rho_k = \frac{\gamma_k}{\gamma_0}}, \qquad k = 0, 1, 2, \ldots \tag{16}$$

var $Cov(Y_t, Y_{t-k})$ $Cov(Y_t, Y_{t-(-k)})$ ห่างกัน k

เนื่องจากฟังก์ชัน Autocovariance เป็นสมมาตรดังนั้น $\gamma_k = \gamma_{-k}$ และ $\rho_k = \rho_{-k}$

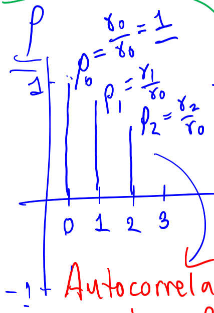
กราฟที่แสดงค่า ACF ในแกนตั้งและค่าล่า k ในแกนนอนเราจะเรียกว่า *โครีโลแกรม (Correlogram)*

sample autocovariance และ autocorrelation

- ค่าแปรปรวนร่วมในตนเองที่ช้ากว่ากัน k คาบเวลาของตัวอย่าง (lag k^{th} sample autocovariance)

sample estimator.

$$\hat{\gamma}_k = \frac{1}{T}\sum_{t=k+1}^{T}(y_t - \bar{Y})(y_{t-k} - \bar{Y}) \qquad (17)$$



$-1 \le \rho_k \le 1$

$\boxed{\rho_k = \frac{\gamma_k}{\gamma_0}}$

Autocorrelation Function (ACF)

lag (k) → ρ_k

กราฟของ ACF → Correlogram

sample autocovariance และ autocorrelation

- ค่าแปรปรวนร่วมในตนเองที่ช้ากว่ากัน k คาบเวลาของตัวอย่าง (lag k^{th} sample autocovariance)

$$\hat{\gamma}_k = \frac{1}{T}\sum_{t=k+1}^{T}(y_t - \bar{Y})(y_{t-k} - \bar{Y}) \tag{17}$$

- สหสัมพันธ์ร่วมในตนเองที่ช้ากว่ากัน k คาบเวลาของตัวอย่าง (lag k^{th} sample autocorrelation) ได้จากสูตรดังต่อไปนี้

$$\boxed{\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}} \tag{18}$$

sample $\hat{\rho}_1$

ρ_t & ρ_{t-1}

0.2

$H_0: \rho_1 = 0$ vs $H_1: \rho_1 \neq 0$

population parameter

$H_0: \rho_k = 0$ vs $H_1: \rho_k \neq 0$

โดยที่ $\bar{Y} = \frac{1}{T}\sum_t^T y_t$ คือค่าเฉลี่ยของตัวอย่าง

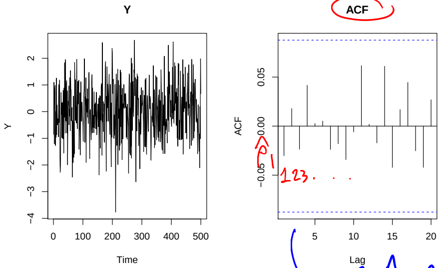
สำหรับ k ใดๆ $\hat{\rho}_k \sim N(0, \frac{1}{T})$

$t = \hat{\rho}_k / \sqrt{\frac{1}{T}} \sim N(0,1)$

ปฏิเสธ H_0 ถ้า $|t| = \left|\frac{\hat{\rho}_k}{\sqrt{1/T}}\right| = |\sqrt{T}\hat{\rho}_k| > c.v.$ จาก N $(\alpha = 0.05,\ 1.96)$

Figure: ข้อมูลเกาซเซียนไวทนอซและฟังก์ชันสหสัมพันธ์ในตัวเอง

$\sqrt{T}|\hat{\rho}_k| > c.v. \Rightarrow |\hat{\rho}_k| > \frac{c.v.}{\sqrt{T}}$



$\hat{\rho}_k > \frac{cv}{\sqrt{T}}$ หรือ $\hat{\rho}_k < -\frac{cv}{\sqrt{T}}$

ถ้า $\hat{\rho}_k$ ไปสูงต่ำ เกินเส้นประเหล่านี้: ปฏิเสธ $H_0(\rho_k = 0)$ และสรุปว่า $\rho_k \neq 0$ แสดงว่า y_t กับ y_{t-k} มีความสัมพันธ์กัน

ในโปรแกรม R แผนภาพสำหรับฟังก์ชันสหสัมพันธ์ในตัวเองจะมีเส้นประแสดงความเชื่อมั่น 95% จาก 0 ให้เสมอ โดยที่เส้นประดังกล่าวมีพื้นฐานจากการที่ $y_t \sim iidN(0, \sigma^2)$ ดังนั้น

$$\hat{\rho}_k \overset{a}{\sim} N(0, \frac{1}{T}), \qquad \text{for} \qquad k > 0$$

ตัวอย่างการคำนวณ ACF

เบื้องต้น ⇒ ตัวแปร Y_t มีความสัมพันธ์กับ Y_{t-k}

พิจารณาผลได้ตอบแทนในรูปล็อกของหุ้น PTT เราสามารถเรียกใช้ฟังก์ชัน `acf` หลังจากเรียก package `library(TSA)`

Figure: ACF ของผลได้ตอบแทนในรูปล็อกของ PTT

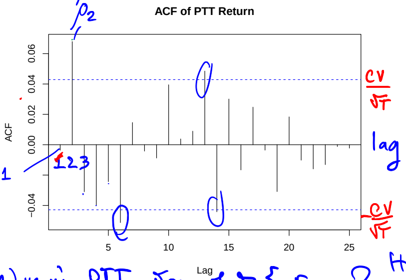
หลังจากเรียก TSA − acf(pttstock$lret)

คำสั่ง r

acf(pttstock$lret)

คำนวณ $\hat{\rho}_0 = 1$

ตัด lag 0



r_t (log return) ของหุ้น PTT มีความสัมพันธ์ r_{t-1} ?

$H_0: \rho_1 = 0$ vs $H_1: \rho_1 \neq 0$

$\hat{\rho}_1$ อยู่ในเส้นประ ไม่สามารถปฏิเสธ H_0

$r_t \sim r_{t-2}$

$H_0: \rho_2 = 0$ vs $H_1: \rho_2 \neq 0$

$\hat{\rho}_2 > \frac{cv}{\sqrt{T}}$ ปฏิเสธ $H_0 \to \rho_2 \neq 0$

การทดสอบพอรทแมนโท (Portmanteau test)

- การทดสอบว่าข้อมูลมีสหสัมพันธ์ในตัวเองในหลายๆคาบ (k) ว่าเท่ากับ 0 พร้อมๆกันหรือไม่ Box and Pierce (1970) ได้เสนอตัวสถิติพอรทแมนโทที่คำนวณดังสูตรต่อไปนี้

$H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0$

$H_1: \rho_j \neq 0$

$$Q^*(m) = T \sum_{k=1}^{m} \hat{\rho}_k^2 \sim \chi^2_{df=m} \tag{19}$$

การทดสอบพอรทแมนโท (Portmanteau test)

- การทดสอบว่าข้อมูลมีสหสัมพันธ์ในตัวเองในหลายๆคาบ (k) ว่าเท่ากับ 0 พร้อมๆกันหรือไม่ Box and Pierce (1970) ได้เสนอตัวสถิติพอรทแมนโทที่คำนวณดังสูตรต่อไปนี้

$$Q^*(m) = T\sum_{k=1}^{m}\hat{\rho}_k^2 \tag{19}$$

- ในการทดสอบสมมุติฐานหลัก $H_0 : \rho_1 = ... = \rho_m = 0$ กับสมมุติฐานทางเลือก $H_1 : \rho_i \neq 0$ สำหรับบางคาบย้อนหลังใน $i \in 1, 2, ..., m$

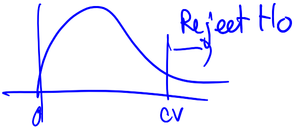
การทดสอบพอรทแมนโท (Portmanteau test)

- การทดสอบว่าข้อมูลมีสหสัมพันธ์ในตัวเองในหลายๆคาบ (k) ว่าเท่ากับ 0 พร้อมๆกันหรือไม่ Box and Pierce (1970) ได้เสนอตัวสถิติพอรทแมนโทที่คำนวณดังสูตรต่อไปนี้

$$Q^*(m) = T\sum_{k=1}^{m} \hat{\rho}_k^2 \tag{19}$$

- ในการทดสอบสมมุติฐานหลัก $H_0 : \rho_1 = \ldots = \rho_m = 0$ กับสมมุติฐานทางเลือก $H_1 : \rho_i \neq 0$ สำหรับบางคาบย้อนหลังใน $i \in 1, 2, \ldots, m$
- ภายใต้ข้อสมมุติว่า Y_t เป็นลำดับที่แจกแจงเป็นอิสระและเหมือนกัน $Q^*(m) \sim \chi^2_m$

ปฏิเสธ $H_0: Q(m) > c.v. \chi^2$



การทดสอบพอรทแมนโท (Portmanteau test)

- Ljung and Box (1978) ได้ปรับปรุงตัวสถิติ $Q^*(m)$ โดยการเพิ่มพลัง(power)ในการทดสอบเมื่อมีตัวอย่างจำกัดโดยเสนอตัวสถิติ equation Q(m)=T(T+2) $\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k}$equation

การทดสอบพอรทแมนโท (Portmanteau test)

- Ljung and Box (1978) ได้ปรับปรุงตัวสถิติ $Q^*(m)$ โดยการเพิ่มพลัง(power)ในการทดสอบเมื่อมีตัวอย่างจำกัดโดยเสนอตัวสถิติ equation Q(m)=T(T+2) $\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k}$equation
- ตัดสินใจปฏิเสธสมมุติฐานหลัก $H_0 : \rho_1 = ... = \rho_m = 0$ ถ้า $Q(m) > \chi^2_{1-\alpha,m}$ โดยที่ $\chi^2_{1-\alpha,m}$ แสดงถึงควอนไทล์ที่ $100(1-\alpha)$ ของ χ^2_m

การทดสอบพอรทแมนโท (Portmanteau test)

- Ljung and Box (1978) ได้ปรับปรุงตัวสถิติ $Q^*(m)$ โดยการเพิ่มพลัง(power)ในการทดสอบเมื่อมีตัวอย่างจำกัดโดยเสนอตัวสถิติ equation Q(m)=T(T+2) $\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k}$equation
- ตัดสินใจปฏิเสธสมมุติฐานหลัก $H_0 : \rho_1 = ... = \rho_m = 0$ ถ้า $Q(m) > \chi^2_{1-\alpha,m}$ โดยที่ $\chi^2_{1-\alpha,m}$ แสดงถึงควอนไทล์ที่ $100(1-\alpha)$ ของ χ^2_m
- ในทางปฏิบัติการเลือกค่า m จะส่งผลต่อความสามารถในการทดสอบ ดังนั้นงานศึกษาเชิงประจักษ์ส่วนใหญ่มักจะรายงานค่า Ljung-Box $Q(m)$ หลายๆค่าเช่น $m = 5, 10, 20$ หรืองานวิจัยบางงานพบว่าค่า $m = \ln(T)$ ให้ผลการทดสอบที่ดี

ตัวอย่างการทดสอบพอรทแมนโทของผลได้ตอบแทนของ PTT

ฟังก์ชันที่ใช้ในการทดสอบพอรทแมนโทคือ `Box.test` โดยเราต้องกำหนด argument คือ

- ข้อมูลที่ใช้ทดสอบ - *lret*
- จำนวนคาบที่รวมมาทดสอบ (m)
- ชนิดของการทดสอบ (`type="Ljung"`) สำหรับ Ljung and Box (1978)

m=5 วิธีคำนวณ Q

```
> Box.test(lret, lag=5, type="Ljung")
^^IBox-Ljung test
data:  lret
X-squared = 16.2609, df = 5, p-value = 0.006137
```

$Q(m=5) = Q(5) = 16.26$ $H_0: \rho_1=\rho_2=\rho_3=\rho_4=\rho_5=0$

$CV(df=5, \alpha=0.05) = 11$

$Q(m=5) > CV.$ ปฏิเสธ H_0 แสดงว่ามี $\rho_j \ (j=1,\dots,5)$ ตัวใดตัวหนึ่ง $\neq 0$

หรือ พิจารณา p-value $= 0.006 < \alpha(0.05)$ ปฏิเสธ H_0