

Course: Applied Business Statistics

Week 15

Chapter 7. Hypothesis Testing (3)

Lecturer: Udam Prang, PhD, MEd

មុខវិជ្ជា៖ ស្ថិតិវិភាគកម្មវិធីអនុវត្ត

សប្តាហ៍ទី១៥

មេរៀនទី៧. ការធ្វើតេស្តសម្របសម្រួល (៣)

គ្រូបង្រៀន៖ បណ្ឌិត ប្រាំង ឧត្តម

៣. តេស្តសម្រាកប្រឆាំងជំងឺ

តេស្តសម្មតិកម្មគំរូពីរ

យើងធ្វើតេស្តសម្មតិកម្មដោយផ្ដោតលើគំរូពីរផ្សេងគ្នានៃស្ថិតិសាកលពីរផ្សេងគ្នា នៅពេលដែលយើងចង់បញ្ជាក់ពីភាពខុសគ្នារវាងប៉ារ៉ាម៉ែត្ររបស់ស្ថិតិសាកលទាំងពីរ។

ក្នុងមេរៀននេះ យើងនឹងសិក្សាលើការធ្វើតេស្តសម្មតិកម្ម សម្រាប់ភាពខុសគ្នា (ឬគម្លាត) រវាងតម្លៃមធ្យមសាកលពីរ (μ_1 និង μ_2) និងភាពខុសគ្នា (ឬគម្លាត) រវាងតម្លៃសមាមាត្រសាកលពីរ (p_1 និង p_2)។ ក្នុងនោះ ភាពខុសគ្នារវាង μ_1 និង μ_2 តាងដោយ μ_0 ហើយភាពខុសគ្នារវាង p_1 និង p_2 តាងដោយ p_0 ។

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

វិធីសាស្ត្រក្នុងការធ្វើតេស្តសម្មតិកម្ម សម្រាប់តម្លៃ μ_0 មានពីរគឺ៖

- ដោយប្រើបំណែងចែកអក្សរ Z
- ដោយប្រើបំណែងចែកអក្សរ T

គេស្តួសប្រតិកម្មគំរូពីសម្រាប់តម្លៃ μ_0

យើងប្រើបំណែងចែកអក្សរ Z នៅពេលដែលយើងស្គាល់តម្លៃគម្លាតស្តង់ដារសាកល (σ)
ហើយយើងប្រើបំណែងចែកអក្សរ T នៅពេលដែលយើងមិនស្គាល់តម្លៃគម្លាតស្តង់ដារ
សាកល (σ)។

តេស្តសម្មតិកម្មកំរិតសម្រាប់តម្លៃ μ_0

ខាងក្រោមនេះ ជាទម្រង់សម្មតិកម្មទៅតាមប្រភេទតេស្តនីមួយៗ៖

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តទាំងសងខាង

- $H_0: \mu_1 - \mu_2 = \mu_0$
- $H_1: \mu_1 - \mu_2 \neq \mu_0$

តេស្តសម្មតិកម្មតំរូវលើសម្រាប់តម្លៃ μ_0

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តខាងឆ្វេង

- $H_0: \mu_1 - \mu_2 \geq \mu_0$
- $H_1: \mu_1 - \mu_2 < \mu_0$

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តខាងស្តាំ

- $H_0: \mu_1 - \mu_2 \leq \mu_0$
- $H_1: \mu_1 - \mu_2 > \mu_0$

ក្នុងនោះ μ_0 ជាតម្លៃរបស់មធ្យមសាកលដែលត្រូវឆ្លងកាត់តេស្តបញ្ជាក់ពីសុពលភាព។

តេស្តសម្រេចចិត្តសម្រាប់តម្លៃ μ_0

ក្នុងករណីយើងប្រើបំណែងចែកអក្សរ Z យើងត្រូវបំប្លែងគម្លាតរវាង \bar{x}_1 និង \bar{x}_2 ទៅជាពិន្ទុស្តង់ដារ z តាមរូបមន្តដូចខាងក្រោម៖

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ទន្ទឹមនេះ p-value ត្រូវបានកំណត់ដូចខាងក្រោម៖

សម្រាប់តេស្តទាំងសងខាង៖

- ប្រសិនបើតម្លៃ z ធំជាងសូន្យ p-value មានតម្លៃស្មើនឹង $2P(Z \geq z)$ ។
- ប្រសិនបើតម្លៃ z តូចជាងសូន្យ p-value មានតម្លៃស្មើនឹង $2P(Z \leq z)$ ។

សម្រាប់តេស្តខាងឆ្វេង៖ p-value មានតម្លៃស្មើនឹង $P(Z \leq z)$ ។

សម្រាប់តេស្តខាងស្តាំ៖ p-value មានតម្លៃស្មើនឹង $P(Z \geq z)$ ។

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ប្រសិនបើយើងប្រើបំណែងចែកអក្សរ T រូបមន្តបំប្លែងគម្លាតរវាង \bar{x}_1 និង \bar{x}_2 ទៅជាពិន្ទុស្តង់ដារ t_{df} ប្រែប្រួលទៅតាមករណីខាងក្រោម៖

- ករណីយើងមិនស្គាល់តម្លៃគម្លាតស្តង់ដាររបស់ស្ថិតិសាកលទាំងពីរ ប៉ុន្តែយើងសន្មតថា តម្លៃគម្លាតស្តង់ដារទាំងពីរនោះស្មើគ្នា
- ករណីយើងមិនស្គាល់តម្លៃគម្លាតស្តង់ដាររបស់ស្ថិតិសាកលទាំងពីរ ហើយយើងមិនអាចកំណត់បានថាតម្លៃគម្លាតស្តង់ដារនោះស្មើគ្នានោះទេ

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ករណីយើងមិនស្គាល់តម្លៃគម្លាតស្តង់ដាររបស់ស្ថិតិសាកលទាំងពីរ ប៉ុន្តែយើងសន្មតថា តម្លៃគម្លាតស្តង់ដារទាំងពីរនោះស្មើគ្នា

$$t_{df} = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{\left(\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{df}\right) \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

ក្នុងនោះ: $df = n_1 + n_2 - 2$

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ករណីយើងមិនស្គាល់តម្លៃគម្លាតស្តង់ដាររបស់ស្ថិតិសាកលទាំងពីរ ហើយយើងមិនអាចកំណត់បានថាតម្លៃគម្លាតស្តង់ដារនោះស្មើគ្នានោះទេ

$$t_{df} = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

ក្នុងនោះ $df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ទន្ទឹមនេះ p-value ត្រូវបានកំណត់ដូចខាងក្រោម៖

សម្រាប់តេស្តទាំងសងខាង៖

- ប្រសិនបើតម្លៃ t_{df} ធំជាងសូន្យ p-value មានតម្លៃស្មើនឹង $2P(T_{df} \geq t_{df})$ ។
- ប្រសិនបើតម្លៃ t_{df} តូចជាងសូន្យ p-value មានតម្លៃស្មើនឹង $2P(T_{df} \leq t_{df})$ ។

សម្រាប់តេស្តខាងឆ្វេង៖ p-value មានតម្លៃស្មើនឹង $P(T_{df} \leq t_{df})$ ។

សម្រាប់តេស្តខាងស្តាំ៖ p-value មានតម្លៃស្មើនឹង $P(T_{df} \geq t_{df})$ ។

តេស្តសម្មតិកម្មគំរូពីរសម្រាប់តម្លៃ μ_0

ឧទាហរណ៍. ការសិក្សាលើគំរូដែលមានបារីម៉ាក A ចំនួន ២០ ដើម បង្ហាញថា សារជាតិដែលប៉ះពាល់សុខភាពមានកម្រិត ១,៧៥ មិល្លីក្រាម ជាមធ្យម។ ទន្ទឹមនេះ ការសិក្សាលើគំរូដែលមានបារីម៉ាក B ចំនួន ២៥ ដើម បង្ហាញថា សារជាតិដែលប៉ះពាល់សុខភាពមានកម្រិត ២,០៥ មិល្លីក្រាម ជាមធ្យម។ ការសិក្សាមុនៗ បានបញ្ជាក់ថា គម្លាតស្តង់ដារសាកល របស់បារីម៉ាក A ស្មើនឹង ០,២២ មិល្លីក្រាម ហើយគម្លាតស្តង់ដារសាកល របស់បារីម៉ាក B ស្មើនឹង ០,២៤ មិល្លីក្រាម។ តើកម្រិតសារជាតិដែលប៉ះពាល់សុខភាព នៅក្នុងបារីទាំងពីរម៉ាកខុសគ្នាដែរឬទេ? ចូរធ្វើតេស្តសម្មតិកម្ម ដោយកំណត់យកកម្រិតសំខាន់ ៥%។

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ μ_0

ដំណោះស្រាយ

តាង μ_A និង μ_B ជាកម្រិតមធ្យមនៃសារជាតិដែលប៉ះពាល់សុខភាពនៅក្នុងបារីម៉ាក A និងម៉ាក B រៀងគ្នា។

ដូច្នេះ តាមប្រធានលំហាត់ សម្មតិកម្មនត្តិភាព និង សម្មតិកម្មអត្តិភាព មានទម្រង់ដូចខាងក្រោម៖

- $H_0: \mu_A - \mu_B = 0$
- $H_1: \mu_A - \mu_B \neq 0$

តេស្តសម្រាប់កម្រិតសម្រាប់តម្លៃ μ_0

ទិន្នន័យនេះ យើងដឹងថា៖

- $\bar{x}_A = 9, ៧៥$
- $n_A = ២០$
- $\sigma_A = 0, ២២$
- $\bar{x}_B = ២, ០៥$
- $n_B = ២៥$
- $\sigma_B = 0, ២៤$
- $\alpha = 0, ០៥$ ($SL = ៥\%$)

តេស្តសម្មតិកម្មតំរូវលើសម្រាប់តម្លៃ μ_0

ដោយសារយើងស្គាល់តម្លៃគម្លាតស្តង់ដារសាកល យើងអាចប្រើបំណែងចែកអក្សរ Z និង ត្រូវបំប្លែងតម្លៃនៃគម្លាតរវាង \bar{x}_1 និង \bar{x}_2 ទៅជាពិន្ទុស្តង់ដារ z តាមរូបមន្តដូចខាងក្រោម៖

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(9, 75 - 2, 05) - 0}{\sqrt{\frac{0, 222}{20} + \frac{0, 242}{25}}} = -4, 38$$

តេស្តសម្មតិកម្មកំរិតសេចក្តីសង្ខេប μ_0

ដោយសារតេស្តនេះ ជាតេស្តទាំងសងខាង ហើយតម្លៃ z តូចជាងសូន្យ យើងបាន៖

$$p\text{-value} = 2 \times P(Z \leq z) = 2 \times P(Z \leq -4,38)$$

តាមតារាង Z យើងកំណត់បានថា $P(Z \leq -4,38)$ ស្មើនឹងសូន្យ។

យើងបាន៖ $p\text{-value} = 2 \times 0 = 0$

តាមលទ្ធផលខាងលើ $p\text{-value}$ តូចជាង α ។

ដូច្នេះ យើងត្រូវបដិសេធសម្មតិកម្មនតិភាព។

តេស្តសម្មតិកម្មតំរូវលើសម្រាប់តម្លៃ p_0

វិធីសាស្ត្រក្នុងការធ្វើតេស្តសម្មតិកម្ម សម្រាប់តម្លៃ p_0 មានតែមួយប៉ុណ្ណោះគឺ ដោយប្រើ
បំណែងចែកអក្សរ Z ។

តេស្តសម្មតិកម្មគំរូពីសេចក្តីស្របគ្នា p_0

ខាងក្រោមនេះ ជាទម្រង់សម្មតិកម្ម ទៅតាមប្រភេទតេស្តនីមួយៗ៖

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តទាំងសងខាង

- $H_0: p_1 - p_2 = p_0$
- $H_1: p_1 - p_2 \neq p_0$

តេស្តសម្មតិកម្មតំរូវលើសម្រាប់តម្លៃ p_0

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តខាងឆ្វេង

- $H_0: p_1 - p_2 \geq p_0$
- $H_1: p_1 - p_2 < p_0$

សម្មតិកម្មដែលត្រូវឆ្លងកាត់តេស្តខាងស្តាំ

- $H_0: p_1 - p_2 \leq p_0$
- $H_1: p_1 - p_2 > p_0$

ក្នុងនោះ p_0 ជាតម្លៃរបស់មធ្យមសាកលដែលត្រូវឆ្លងកាត់តេស្តបញ្ជាក់ពីសុពលភាព។

តេស្តសម្រាប់កម្រិតស្របគ្នា p_0

រូបមន្តបំប្លែងតម្លៃនៃគម្លាតរវាង \bar{p}_1 និង \bar{p}_2 ទៅជាពិន្ទុស្តង់ដារ z មានដូចខាងក្រោម៖

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - p_0}{\sqrt{\frac{\bar{p}_1(1 - \bar{p}_1)}{n_1} + \frac{\bar{p}_2(1 - \bar{p}_2)}{n_2}}}$$

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ p_0

ទន្ទឹមនេះ p -value ត្រូវបានកំណត់ដូចខាងក្រោម៖

សម្រាប់តេស្តទាំងសងខាង៖

- ប្រសិនបើតម្លៃ z ធំជាងសូន្យ p -value មានតម្លៃស្មើនឹង $2P(Z \geq z)$ ។
- ប្រសិនបើតម្លៃ z តូចជាងសូន្យ p -value មានតម្លៃស្មើនឹង $2P(Z \leq z)$ ។

សម្រាប់តេស្តខាងឆ្វេង៖ p -value មានតម្លៃស្មើនឹង $P(Z \leq z)$ ។

សម្រាប់តេស្តខាងស្តាំ៖ p -value មានតម្លៃស្មើនឹង $P(Z \geq z)$ ។

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ p_0

ឧទាហរណ៍. ការស្រាវជ្រាវដោយក្រុមហ៊ុនលក់ទំនិញអនឡាញជាច្រើនបានអះអាងថា អ្នកដែលទិញទំនិញអនឡាញជាស្រ្តី មានច្រើនជាង អ្នកដែលទិញទំនិញអនឡាញជាបុរស។ ការអង្កេតថ្មីមួយ ដែលធ្វើឡើងដោយវិទ្យាស្ថានស្រាវជ្រាវមួយ បានរកឃើញថា បុរស ៥៥០០ នាក់ ក្នុងចំណោម ៧០០០ នាក់ តែងតែទិញទំនិញតាមអនឡាញ ហើយស្រ្តី ៨៥០០ នាក់ ក្នុងចំណោម ១០០០០ នាក់ តែងតែទិញទំនិញតាមអនឡាញ។ ដោយកំណត់យកកម្រិតសំខាន់ ៥% និង ដោយផ្អែកលើលទ្ធផលនៃការអង្កេតដោយវិទ្យាស្ថានស្រាវជ្រាវ តើអ្នកដែលទិញទំនិញអនឡាញជាស្រ្តី ពិតជាច្រើនជាង អ្នកដែលទិញទំនិញអនឡាញជាបុរស មែនឬយ៉ាងណា។ ចូរធ្វើតេស្តសម្មតិកម្ម ដើម្បីបញ្ជាក់។

តេស្តសម្មតិកម្មតំរូវពីសម្រាប់តម្លៃ p_0

ដំណោះស្រាយ

តាង p_w ជាសមាមាត្រនៃចំនួនស្ត្រីដែលទិញទំនិញអនឡាញ ហើយ p_m ជាសមាមាត្រនៃចំនួនបុរសដែលទិញទំនិញអនឡាញ។

តាមប្រធានលំហាត់ យើងត្រូវធ្វើតេស្តខាងឆ្វេង ហើយសម្មតិកម្មនត្តិភាពនិងសម្មតិកម្មអត្តិភាពមានទម្រង់ដូចខាងក្រោម៖

- $H_0: p_w - p_m \geq 0$
- $H_1: p_w - p_m < 0$

តេស្តសម្រាប់កម្រិតសម្រាប់តម្លៃ p_0

ទិន្នន័យនេះ យើងដឹងថា៖

- $\bar{p}_w = 8500/90000 = 0, 094$
- $n_w = 90000$
- $\bar{p}_m = 5500/7000 = 0, 786$
- $n_m = 7000$
- $\alpha = 0, 05$ ($SL = 5\%$)

តេស្តសម្រាប់ភាពខុសគ្នាស្របតាមប្រែប្រួល p_0

តាមរូបមន្តបំប្លែងតម្លៃនៃគម្លាតរវាង \bar{p}_1 និង \bar{p}_2 ទៅជាពិន្ទុស្តង់ដារ z

$$z = \frac{(\bar{p}_w - \bar{p}_m) - p_0}{\sqrt{\frac{\bar{p}_w(1 - \bar{p}_w)}{n_w} + \frac{\bar{p}_m(1 - \bar{p}_m)}{n_m}}}$$

តេស្តសម្រាប់កម្រិតសម្រាប់តម្លៃ p_0

យើងបាន៖

$$z = \frac{(0, ៨៥ - ០, ៧៩) - ០}{\sqrt{\frac{0, ៨៥(១ - ០, ៨៥)}{១០០០០} + \frac{០, ៧៩(១ - ០, ៧៩)}{៧០០០}}} = ៩, ៩៤$$

តេស្តសម្មតិកម្មតំរូវលើសម្រាប់តម្លៃ p_0

ដោយសារតេស្តនេះជាតេស្តខាងឆ្វេង យើងបាន៖

$$p\text{-value} = P(Z \leq z) = P(Z \leq ៩,៩៤)$$

តាមតារាង Z យើងអាចកំណត់បានថា ស្មើនឹងមួយ។

យើងបាន៖ $p\text{-value} = ១$

តាមលទ្ធផលខាងលើ $p\text{-value}$ ធំជាង α ។ ដូច្នេះ យើងមិនអាចបដិសេធសម្មតិកម្ម
នត្តិភាពបានទេ។

បញ្ចប់មេរៀនត្រីមនេះ!

នៅសប្តាហ៍បន្ទាប់ដែលជាសប្តាហ៍ចុងក្រោយ យើងនឹងសាកល្បងអនុវត្តលំហាត់មួយ
ចំនួនដើម្បីរំលឹកមេរៀនដែលបានសិក្សាក្នុងមក។