

ភាពធរមាននៃបំណែងចែកកម្រិត

Normality of a Distribution

ម៉ូឌ ម៉ាត

មន្ត្រីស្រាវជ្រាវនៃផ្នែកគណិតវិទ្យានិងស្ថិតិ

វិទ្យាស្ថានវិទ្យាសាស្ត្រនិងបច្ចេកវិទ្យា

រាជបណ្ឌិត្យសភាកម្ពុជា

អត្ថបទនេះ ត្រូវបានត្រួតពិនិត្យដោយលោក **យឹម រោយវឌ្ឍនៈវិជ្ជា** ប្រធានផ្នែកគណិតវិទ្យានិងស្ថិតិ

វិទ្យាស្ថានវិទ្យាសាស្ត្រនិងបច្ចេកវិទ្យា រាជបណ្ឌិត្យសភាកម្ពុជា

ខែមេសា ឆ្នាំ២០២០

សេចក្តីផ្តើម

នៅក្នុងស្ថិតិវិទ្យា *ភាពណរម៉ាល់* (normality) គឺជាលក្ខណៈមួយបែបដែលស្ថិតិវិទ្យាត្រូវពិនិត្យមើល ចំពោះសំណុំទិន្នន័យមុននឹងសម្រេចជ្រើសរើសប្រើវិធីសមស្របមួយចំនួន ដើម្បីធ្វើការវិភាគ។ សៀវភៅស្ថិតិ វិទ្យាកម្រិតមូលដ្ឋានជាច្រើន ដូចជា Bluman (2018) Lind, Marchal, and Wathen (2019) Moore, McCabe, Alwan, and Craig (2016) ជាដើម បានលើកឡើងអំពីលក្ខខណ្ឌនៃភាពណរម៉ាល់របស់បំណែង ចែកទិន្នន័យនៅក្នុងការប៉ាន់ស្មានមធ្យមស្ថិតិសាកលដោយប្រើចន្លោះជឿជាក់នៅក្នុងតេស្ត z, តេស្ត t, ការវិភាគ វ៉ារ្យង់ និង ការសិក្សាអំពីកូរ៉េឡាស្យុងកាសិន (Pearson correlation) ជាដើម។ នៅក្នុងការវិភាគទិន្នន័យ សម្រាប់ការសិក្សាស្រាវជ្រាវ ការពិនិត្យមើលភាពណរម៉ាល់របស់សំណុំទិន្នន័យមានសារៈសំខាន់ណាស់។ ដើម្បី ឱ្យលទ្ធផលនៃការវិភាគទិន្នន័យ មានភាពត្រឹមត្រូវអាចទទួលយកជាការបាន អ្នកសិក្សាស្រាវជ្រាវ ត្រូវចេះ កំណត់យកវិធីវិភាគ សម្រាប់ទិន្នន័យដែលមានភាពណរម៉ាល់ និង ទិន្នន័យដែលមិនមានភាពណរម៉ាល់ឱ្យបាន ច្បាស់លាស់។ Moore et al. (2016, p. 43) ក៏បានលើកឡើងអំពីសារៈសំខាន់ពីរបំណែងចែកណរម៉ាល់ បន្ថែមពីលើនេះទៀត គឺភាពដែលអាចពិពណ៌នាបានយ៉ាងល្អអំពីបំណែងចែកទិន្នន័យពិតជាក់ស្តែង និង ការ ប៉ាន់ស្មានបានយ៉ាងល្អនូវលទ្ធផលដែលអាចកើតមាននៅក្រោមលក្ខខណ្ឌនៃភាពចៃដន្យ (chance)។

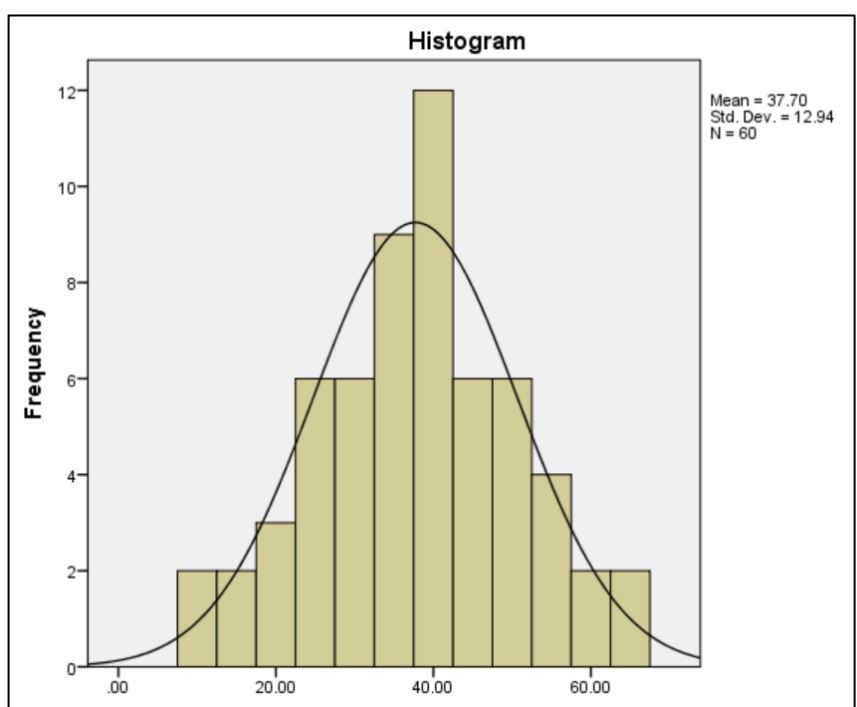
អត្ថបទនេះ នឹងធ្វើការសិក្សាលើកឡើងវិញនូវវិធីកំណត់ភាពណរម៉ាល់នៃបំណែងចែកទិន្នន័យរាប់តាំង ពីវិធីបែបពិពណ៌នា ការប្រើក្រាហ្វិក និង ការធ្វើតេស្តសម្មតិកម្មស្ថិតិមួយចំនួន។

ភាពណរម៉ាល់ នៃបំណែងចែកទិន្នន័យ មានប្រវត្តិកើតឡើងយូរលង់មកហើយ ដែលដំបូងគេនិយាយ ត្រឹមតែអំពីភាពឆ្លុះ។ រឿងរ៉ាវអំពីកំណើតនិងការវិវឌ្ឍនៃបំណែងចែកណរម៉ាល់ មានការជាប់ទាក់ទងនឹងគណិត វិទ្យាជាច្រើន មានជាអាទិ៍ដូចជា Galileo Galilei, Abraham de Moivre, Adrien-Marie Legendre និង Pierre-Simon Laplace (" Distribution, Normal ")។ ទើបតែមកដល់សតវត្សទី១៩ទេ ដែលមានគណិតវិទូពីររូបគឺ Robert Adrain និង Carl Friedrich Gauss បានបង្កើតរូបមន្តសម្រាប់បំណែងចែកណរម៉ាល់ឡើងនៅក្នុងកិច្ច ការដាច់ដោយឡែកពីគ្នា (Lane, n.d.; Mong & Vong, 2019)។

អត្ថបទជាច្រើនដែលមានដូចជា Elliott and Woodward (2007), Ghasemi and Zahediasl (2012), Thode (2002) ជាដើមបានចងក្រងនូវវិធីកំណត់ភាពណរម៉ាល់នៃបំណែងចែកទិន្នន័យ។ ក្នុងចំណោមវិធីដែលលើកមកបង្ហាញទាំងអស់នោះមានចាប់តាំងការបង្ហាញក្រាហ្វិក រង្វាស់ស្ថិតិតិពណ៌នា រហូតដល់តេស្តសម្មតិកម្មស្ថិតិ។

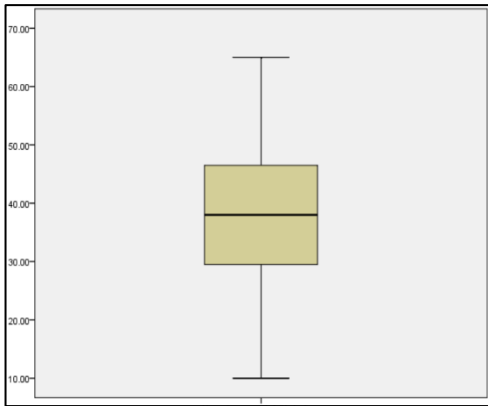
១. ការវាយតម្លៃភាពណរម៉ាល់តាមក្រាហ្វិក

- អ៊ីស្តូក្រាម៖ ជាមធ្យោបាយមួយដែលលឿន សម្រាប់វាយតម្លៃភាពណរម៉ាល់តាមការមើលឃើញ (Korstanje, 2019)។ នៅពេលដែលសំណង់អ៊ីស្តូក្រាមបង្ហាញភាពឆ្លុះនិងរាងជួងនោះ គេអាចសន្និដ្ឋានសំណុំទិន្នន័យអាចមានប្រភពមកពីស្ថិតិសាកលណរម៉ាល់ (normal population)(Elliott & Woodward, 2007, p. 28) (សូមមើលរូបទី១)។ សំណុំទិន្នន័យមួយត្រូវចាត់ទុកថា «ទេរខាងឆ្វេង» ប្រសិនបើប្រេកង់មានកំណើនយឺតៗឡើងដល់តម្លៃខ្ពស់បំផុត រួចចុះយ៉ាងគំហុកមកវិញ។ Frost (2020) លើកឡើងអំពីភាពខ្វះចន្លោះនៅក្នុងការវាយតម្លៃភាពណរម៉ាល់ដោយប្រើអ៊ីស្តូក្រាម។

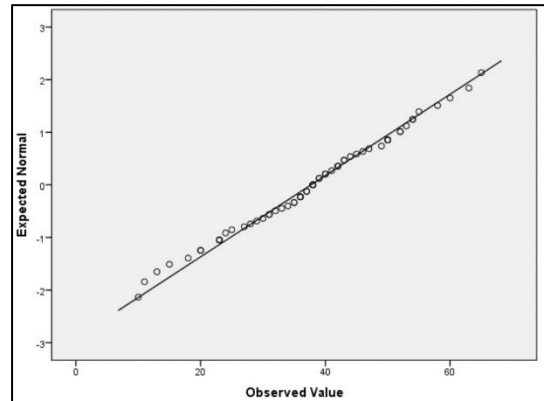


រូបទី១ អ៊ីស្តូក្រាមតាងឱ្យអថេរមួយដែលអាចទុកជាប្រហាក់ប្រហែលបំណែងចែកណរម៉ាល់

- ក្រាហ្វប្រអប់ (boxplot)៖ ខ្សែមេដ្យានកាត់ប្រអប់ចំពាក់កណ្តាល និង ម្យ៉ាងទៀតខ្សែដែលតភ្ជាប់ពីប្រអប់ទៅចុងទាំងសងខាងមានប្រវែងស្មើគ្នា វែងជាងប្រវែងកន្លះប្រអប់បន្តិច នោះគេអាចសន្និដ្ឋានថាសំណុំទិន្នន័យមានប្រភពមកពីស្ថិតិសាកលណរម៉ាល់ (រូបទី២)។
- ក្រាហ្វ Q-Q (quantile-quantile plot)៖ នៅពេលដែលសំណុំចំណុចខិតទៅរកភាពត្រង់ជួរតាមបន្ទាត់មួយ នោះសំណុំទិន្នន័យអាចត្រូវសន្មតថាមានភាពណរម៉ាល់ (រូបទី៣)។
- ក្រាហ្វមែកនិងស្លឹក (Stem-and-leaf plot)៖ មានលក្ខណៈប្រហាក់ប្រហែលនឹងអ៊ីស្តូក្រាមដែរ។ អថេរមួយដែលចាត់ទុកថាប្រហាក់ប្រហែលណរម៉ាល់ គឺលុះត្រាតែនៅក្នុងបំណែងចែកថ្នាក់ដំបូងមានប្រេកង់ទាប ថ្នាក់បន្ទាប់ប្រេកង់កើនឡើងរហូតដល់ចុងក្រោយថយចុះទាបមកវិញ (រូបទី៤)។



រូបទី២ ក្រាហ្វប្រអប់ ដែលបង្ហាញថាអថេរមានភាពណរម៉ាល់



រូបទី៣ ក្រាហ្វ Q-Q ដែលបង្ហាញថាអថេរមានភាពណរម៉ាល់

Frequency	Stem & Leaf
5.00	1 . 01358
10.00	2 . 0033345789
19.00	3 . 0112345566677888899
13.00	4 . 0012223345679
10.00	5 . 0002234458
3.00	6 . 035
Stem width: 10.00	
Each leaf: 1 case(s)	

រូបទី៤ ក្រាហ្វមែកនិងស្លឹក ដែលបង្ហាញថាអថេរមានភាពណរម៉ាល់

ក្រៅពីការសង្កេតមើលលើរូបភាពក្រាហ្វិកទាំងអស់នេះ គេអាចវាយតម្លៃអំពីភាពណរម៉ាល់នៃសំណុំ

ទិន្នន័យតាមរយៈច្បាប់ណរម៉ាល់^១ ដែលចែងថា ចំពោះបំណែងចែកណរម៉ាល់មាន៖

ក) ប្រហែល៦៨%នៃតម្លៃក្នុងបំណែងចែកសំណុំទិន្នន័យស្ថិតនៅក្នុងចន្លោះ១គម្លាតស្តង់ដារពីមធ្យម។

ខ) ប្រហែល៩៥%នៃតម្លៃក្នុងបំណែងចែកសំណុំទិន្នន័យស្ថិតនៅក្នុងចន្លោះ២គម្លាតស្តង់ដារពីមធ្យម។

គ) ប្រហែល៩៩,៧%នៃតម្លៃក្នុងបំណែងចែកសំណុំទិន្នន័យស្ថិតនៅក្នុងចន្លោះ៣គម្លាតស្តង់ដារពីមធ្យម។

នៅក្នុងស្ថិតិវិទ្យា ដើម្បីកំណត់ថាតើបំណែងចែកទិន្នន័យឆ្លុះ ឬ ទេរ គេប្រើប្រាស់មេគុណភាពទេរ (coefficient of skewness)។ មេគុណភាពទេរ មានច្រើនប្រភេទទៅតាមការកំណត់ផ្សេងៗគ្នាដូចជា មេគុណភាពទេរបស់ Pearson (Pearson coefficient of skewness) មេគុណភាពទេរដែលកំណត់នៅក្នុង Excel ឬ Google sheet ឬ មេគុណភាពទេរបស់ Galton ឬ Bowley (" Measures of Skewness and Kurtosis " n.d.) ។

ដើម្បីកំណត់ថាតើកន្ទុយបំណែងចែកទិន្នន័យមានតម្លៃប្រមូលផ្តុំគ្នាតិចឬច្រើន គេប្រើរង្វាស់kurtosis។ លក្ខណៈដែលកំណត់ដោយរង្វាស់ទាំងពីរ (មេគុណភាពទេរ និង kurtosis) ជាប់ទាក់ទងនឹងការវាយតម្លៃអំពីភាពណរម៉ាល់របស់សំណុំទិន្នន័យ។ ចំពោះបំណែងចែកណរម៉ាល់ មេគុណភាពទេរខិតរកសូន្យ^២ ។ ប្រសិនបើតម្លៃមេគុណភាពទេរនៅក្នុងចន្លោះ - 0.5 និង + 0.5 នោះសំណុំទិន្នន័យមានលក្ខណៈស្មើរតែឆ្លុះ។ ចំពោះបំណែងចែកណរម៉ាល់ស្តង់ដារ^៣ kurtosis ស្មើ 3 (Jain, 2018; " Normality Testing - Skewness and Kurtosis, " 2019)។

២. តេស្តស្ថិតិសម្រាប់វាយតម្លៃភាពណរម៉ាល់

^១ ឈ្មោះផ្សេងទៀតសម្រាប់ហៅ មានដូចជា « ច្បាប់តាមពិសោធន៍ (Empirical rule) » ឬ « ច្បាប់ 68-95-99.7 »
^២ ស្មើ ០ តាមបែបទ្រឹស្តី
^៣ បំណែងចែកណរម៉ាល់ដែលមានមធ្យមស្មើ ០ និង រ៉ាវៀងស្មើ ១

បន្ទាប់ពីការសង្កេតលើក្រាហ្វិកនិងរង្វាស់ស្ថិតិមួយចំនួនដូចលើកឡើងនៅក្នុងផ្នែកទី១ខាងលើ តេស្តស្ថិតិមួយត្រូវបានប្រើប្រាស់សម្រាប់កំណត់ភាពណរម៉ាល់របស់បំណែងចែកសំណុំទិន្នន័យ។

២.១. តេស្ត Z ដែលប្រើប្រាស់មេគុណភាពទេរ និង kurtosis

តេស្តនេះមានលើកឡើងនៅក្នុងអត្ថបទរបស់ Kim (2013)។ សម្មតិកម្មនៅក្នុងតេស្តកំណត់ដោយ

$$H_0 : \text{ទិន្នន័យគំរូតាងមានប្រភពអំពីបំណែងចែកណរម៉ាល់}$$

$$H_1 : \text{ទិន្នន័យគំរូតាងមានប្រភពអំពីបំណែងចែកមិនណរម៉ាល់}$$

តម្លៃស្ថិតិតេស្ត z សម្រាប់មេគុណភាពទេរ និង kurtosis កំណត់រៀងគ្នាដោយ

$$z = \frac{\textit{skewness}}{SE_{\textit{skewness}}} \quad (១)$$

និង
$$z = \frac{\textit{kurtosis}}{SE_{\textit{kurtosis}}} \quad (២) ។$$

បើយកកម្រិតសំខាន់ $\alpha = 0.05$ នោះតម្លៃវិនិច្ឆ័យ $z = \pm 1.96$ ហើយបើតម្លៃ z តាមរយៈសមីការ (១) និង (២) ស្ថិតនៅក្នុងចន្លោះ ± 1.96 នោះ H_0 នឹងមិនត្រូវបានបដិសេធ នាំឱ្យភាពណរម៉ាល់អាចទទួលយកបាន។ តារាងទី១ បង្ហាញអំពីលទ្ធផលវិភាគស្ថិតិពណ៌នាដោយប្រើកម្មវិធី IBM SPSS 21 ដែលមេគុណភាពទេរ (skewness) និង kurtosis រួមទាំងភាពល្អៀងស្តង់ដារ (SE) របស់វាស្ថិតនៅជួរដេកខាងក្រោមបង្អស់នៃតារាងនេះ។

តារាងទី១ តម្លៃស្ថិតិពណ៌នា

		Statistic	Std. Error	
Variable	Mean	37.7000	1.67050	
	95% Confidence Interval for Mean	Lower Bound	34.3573	
		Upper Bound	41.0427	
	5% Trimmed Mean	37.7778		
	Median	38.0000		
	Variance	167.434		
	Std. Deviation	12.93963		
	Minimum	10.00		
	Maximum	65.00		
	Range	55.00		
	Interquartile Range	17.50		
	Skewness	-.124	.309	
	Kurtosis	-.364	.608	

២.២. តេស្តកុលម៉ូហ្គោរ៉ូវ-ស្យែណូវ និង តេស្តស្វាញីវី-វីលក

(Kolmogorov-Smirnov Test and Shapiro- Wilk Test)

អត្ថបទជាច្រើនមានដូចជា Ghasemi and Zahediasl (2012), Elliott and Woodward (2007) និង Korstanje (2019) " Normality Tests " n.d.) ជាដើម បានលើកឡើងអំពីការប្រើប្រាស់តេស្តទាំងពីរនេះសម្រាប់វាយតម្លៃភាពណរម៉ាល់។ Hopkins and Weeks (1990) ក៏បានបង្ហាញអំពីការប្រើប្រាស់ដ៏ទូលំទូលាយនៃតេស្តកុលម៉ូហ្គោរ៉ូវ-ស្យែណូវ ប៉ុន្តែបានបន្ថែមអំពីភាពនៅមានកម្រិតនៃប្រសិទ្ធភាពរបស់តេស្តនេះបើធៀបនឹងតេស្តផ្សេងខ្លះទៀត។ សម្មតិកម្មរបស់តេស្តទាំងពីរនេះគឺ

H_0 : ទិន្នន័យគំរូតាងមានប្រភពអំពីបំណែងចែកណរម៉ាល់

H_1 : ទិន្នន័យគំរូតាងមានប្រភពអំពីបំណែងចែកមិនណរម៉ាល់

ឧទាហរណ៍មួយអំពីការប្រើប្រាស់កម្មវិធី IBM SPSS 21 ដើម្បីធ្វើតេស្តបង្ហាញលទ្ធផលដូចក្នុងតារាងទី២។

តារាងទី២ លទ្ធផលតេស្តកុលម៉ូហ្គោវ៉ែ-ស្វៀណូវ និង តេស្តស្វាពីរ៉ូ-វីលកី

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Variable	.067	60	.200 [*]	.988	60	.802

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

ចំពោះការសន្និដ្ឋានអំពីភាពណរម៉ាល់គឺ ត្រូវពិនិត្យមើលតម្លៃប្រូបាប៊ីលីតេដែលតាងដោយ Sig. នៅក្នុងតារាងខាងលើ (តារាងទី២)។ ប្រសិនបើតម្លៃប្រូបាប៊ីលីតេនៅក្នុងតេស្តទាំងពីរធំលើលើកម្រិតសំខាន់^៤ (α) នោះសម្មតិកម្ម H_0 មិនត្រូវបានបដិសេធ ហើយសន្មតថាបំណែងចែកទិន្នន័យមានភាពណរម៉ាល់។

២.៣. តេស្តហ្សាក-បេរ៉ា (Jarque-Bera Test)

ជាតេស្តដែលគេប្រើ ដើម្បីរកមើលភាពណរម៉ាល់នៃទិន្នន័យសេរីពេល ទិន្នន័យភាពលំអៀងនៅក្នុងម៉ូដែលក្រែសស្យុង និង ទិន្នន័យជារៀងរាល់ ជាពិសេសនៅពេលដែលគំរូតាងមានទំហំធំ (Stephanie, 2016)។ សុសវែរដែលអាចអនុវត្តកិច្ចការតេស្តនេះបាន មានដូចជា Matlab, Eviews, STATA និង R ជាដើម។

សម្មតិកម្មសម្រាប់តេស្តនេះគឺ

H_0 : បំណែងចែកទិន្នន័យមានភាពណរម៉ាល់

H_1 : បំណែងចែកទិន្នន័យមិនមានភាពណរម៉ាល់

តេស្តនេះប្រើប្រាស់បំណែងចែកប្រូបាប៊ីលីតេ χ^2 ដែលមានតម្លៃស្ថិតិតេស្តកំណត់ដោយ

$$JB = n \left[\frac{s^2}{6} + \frac{(k-3)^2}{24} \right] \quad (៣)$$

^៤ នៅក្នុងការធ្វើតេស្ត កម្រិតសំខាន់ (significance level) ដែលប្រើប្រាស់អាច 0.01, 0.05 ឬ 0.10 ប៉ុន្តែជាញឹកញាប់គេប្រើនយក 0.05 ។

ដែល k ជា kurtosis , s ជាមេគុណភាពទេរ (skewness), ដឺក្រេសេរី (df) គឺ 2 និង តេស្តនេះ ជាប្រភេទ តេស្តកន្ទុយស្តាំ។

២.៤. តេស្តអានដឺសុន-ដារលីង (Anderson-Darling Test for Normality)

តេស្តអានដឺសុន-ដារលីង ជាប្រភេទតេស្តលើភាពស៊ីគ្នា (goodness-of-fit) រវាងបំណែងចែកសំណុំ ទិន្នន័យមួយនិងបំណែងចែកសំណុំទិន្នន័យមួយទៀត (" The Anderson-Darling statistic, " 2019; NIST, 2013; Stephanie, 2014; STEPHENS, 1979)។ ដូច្នេះ តេស្តនេះក៏អាចត្រូវប្រើប្រាស់ ដើម្បីកំណត់ភាព ណរម៉ាល់នៃសំណុំទិន្នន័យផងដែរ។ សម្មតិកម្មសម្រាប់តេស្តនេះ ដើម្បីពិនិត្យមើលភាពណរម៉ាល់ ស្រដៀងនឹង សម្មតិកម្មនៃតេស្តដែលបានរៀបរាប់ខាងលើគឺ

H_0 : បំណែងចែកទិន្នន័យមានភាពណរម៉ាល់

H_1 : បំណែងចែកទិន្នន័យមិនមានភាពណរម៉ាល់

ការលើកឡើងមួយចំនួនដូចជា Lee, Park, and Jeong (2016) និង Jäntschi and Bolboacă (2018) ជាដើម ថាតេស្តនេះមានអានុភាពជាងតេស្តស្វាពីរ៉ូ-វីលក៏ ទៅទៀត។

៣. សេចក្តីសន្និដ្ឋាន

ការពិនិត្យមើលភាពណរម៉ាល់នៅសំណុំទិន្នន័យមានសារៈសំខាន់ ដើម្បីសម្រេចថាតើត្រូវជ្រើសរើសវិធី ស្ថិតិមួយណាមកប្រើប្រាស់ឱ្យបានត្រឹមត្រូវ និង ដើម្បីប្រើប្រាស់ជាផ្នែកមួយសម្រាប់កំណត់ភាពដែលអាចយក ទៅប្រើបាននៃម៉ូដែលមួយ ដូចជា ម៉ូដែលរ៉ូក្រេសស្យុងជាដើម។ នៅក្នុងដំណើរការនេះ ជាបឋមរូបភាពជា ក្រាហ្វិកតាងឱ្យសំណុំទិន្នន័យតែងត្រូវសង់ឡើង ហើយនិងត្រូវបន្តដោយការធ្វើតេស្តសម្មតិកម្មដែលសមស្រប ជាការបន្ថែមទៀត ដើម្បីវាយតម្លៃអំពីភាពណរម៉ាល់របស់សំណុំទិន្នន័យ។

ការធ្វើតេស្តពិនិត្យមើលភាពណរម៉ាល់ដែលត្រូវបានអនុម័តរួចក្នុងការសិក្សាស្រាវជ្រាវ នៅមានចំនួន ច្រើនលើសពីអ្វីដែលបានលើកឡើងនៅក្នុងអត្ថបទនេះទៅទៀត។ ដូច្នេះ អត្ថបទស្រាវជ្រាវក្រោយៗ អាចនឹង លើកយកវិធីតេស្តមើលភាពណរម៉ាល់មកបង្ហាញបន្ថែម។

၆. နမူနာရွေးချယ်မှု

- The Anderson-Darling statistic. (2019). Retrieved from <https://support.minitab.com/en-us/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/normality/the-anderson-darling-statistic/>
- Bluman, A. G. (2018). *Elementary Statistics: A step by step Approach*. USA: McGraw-Hill Education.
- Distribution, Normal. *International Encyclopedia of the Social Sciences*. Retrieved from <https://www.encyclopedia.com/science-and-technology/mathematics/mathematics/normal-distribution>
- Elliott, A. C., & Woodward, W. A. (2007). *STATISTICAL ANALYSIS Quick Reference Guidebook With SPSS Examples*. USA: Sage Publications, Inc.
- Frost, J. (2020). Assessing Normality: Histograms vs. Normal Probability Plots. Retrieved from <https://statisticsbyjim.com/basics/assessing-normality-histograms-probability-plots/>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2), 486.
- Hopkins, K. D., & Weeks, D. L. (1990). Tests for Normality and Measures of Skewness and Kurtosis: Their Place in Research Reporting. *Educational and Psychological Measurement*, 50(4), 717-729. doi:10.1177/0013164490504001
- Jain, D. (2018). Skew and Kurtosis: 2 Important Statistics terms you need to know in Data Science. Retrieved from <https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis-388fef94eeaa>
- Jäntschi, L., & Bolboacă, S. D. (2018). Computation of Probability Associated with Anderson–Darling Statistic.
- Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics*, 38(1), 52-54. doi:10.5395/rde.2013.38.1.52
- Korstanje, J. (2019). 6 ways to test for a Normal Distribution — which one to use? Retrieved from <https://towardsdatascience.com/6-ways-to-test-for-a-normal-distribution-which-one-to-use-9dcf47d8fa93>
- Lane, D. M. (n.d.). History of the Normal Distribution. Retrieved from http://onlinestatbook.com/2/normal_distribution/history_normal.html
- Lee, C., Park, S., & Jeong, J. (2016). Comprehensive comparison of normality tests: Empirical study using many different types of data. *Journal of the Korean Data and Information Science Society*, 27, 1399-1412. doi:10.7465/jkdi.2016.27.5.1399
- Lind, D. A., Marchal, W. G., & Wathen, S. A. (2019). *Basic Statistics for BUSINESS & ECONOMICS*. United States of America: McGraw-Hill Education.

- Measures of Skewness and Kurtosis. (n.d.). Retrieved from <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm#:~:text=The%20skewness%20for%20a%20normal,data%20that%20are%20skewed%20right.>
- Mong, M., & Vong, B. (2019). ស្តីពីវិទ្យា ដើមកំណើតនិងការវិវឌ្ឍ .Retrieved from <http://rac.gov.kh/royal-academy/research/attachments/original/51.pdf?1549012616>
- Moore, D. S., McCabe, G. P., Alwan, L. C., & Craig, B. A. (2016). *The Practice of Statistics for Business and Economics*. USA: W. H. Freeman and Company.
- NIST. (2013). Anderson-Darling Test. Retrieved from <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm>
- Normality Testing - Skewness and Kurtosis. (2019). Retrieved from <https://help.gooddata.com/doc/en/reporting-and-dashboards/maql-analytical-query-language/maql-expression-reference/aggregation-functions/statistical-functions/predictive-statistical-use-cases/normality-testing-skewness-and-kurtosis>
- Normality Tests. (n.d.). Retrieved from <https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Normality Tests.pdf>
- Stephanie. (2014). Anderson-Darling Test & Statistic: Definition, Examples.
- Stephanie. (2016). Jarque-Bera Test. Retrieved from <https://www.statisticshowto.datasciencecentral.com/jarque-bera-test/>
- STEPHENS, M. A. (1979). THE ANDERSON-DARLING STATISTIC. (Technical Report No. 278.).
- Thode, H. C. (2002). *TESTING FOR NORMALITY*. USA: Marcel Dekker, Inc.