සංඛ්‍යානය විෂයයෙහි සම්භවය හා විකාශය පිළිබඳ සංක්ෂිප්ත විමර්ශනයක්

ර.මු.මු. බණ්ඩාර රත්නතිලක[1]

සංක්ෂේපය

දත්ත එක් රැස්කිරීම, සංවිධානය, විශ්ලේෂණය හා විශ්ලේෂිත දත්ත ඇසුරෙන් යුක්ති සහගත වූත් සාධාරණ වූත් නිගමනවලට එළැඹීම සංඛ්‍යානයේ ප්‍රමුඛත ම කාර්යයභාරයෝ වෙති. සංඛ්‍යාන විෂයය පිළිබඳ වත්මන් අධ්‍යයනවල දී සෙසු විෂයයනට සාපේක්ෂ ව ඉතිහාසය දෙස හැරී නො බලන තරම් ය. මෙහෙයින් ඇතැම් විට සංඛ්‍යාන විෂයයට ආධුනික වූවෝ එහි විෂය පථය පිළිබඳ නිසි අවබෝධයකින් තොර වෙත්. සංඛ්‍යානය විෂයයෙහි ඉතිහාසය විමර්ශනය කිරීමෙන් තද් විෂයය හදාරණ සිසුන් වෙත එහි විෂය පථය හා සීමා කවරේ ද යනු අවබෝධ කැරැවීම මේ ලිපියේ අරමුණ යි. මෙය සම්පාදනයේ දී ද්විතීයක මූලාශ්‍රය වශයෙන් ලිපි ලේඛන සමුදායයක් ඇසුරු කැරිණි. සංඛ්‍යාන විෂයයෙහි ප්‍රාරම්භ ලක්ෂ්‍යය ලෙස සැලැකෙන දේශපාලන අංක විද්‍යාව ආරම්භ වූ දාහත් වැනි සියවසේ මැද භාගයේ සිට සංඛ්‍යාන විෂයය විකාශය වූ ආකාරය ප්‍රධාන වශයෙන් වකවානු හතරකට ප්‍රභේද කළ හැකිය. 1650 - 1750 වකවානුව, 1750 - 1820 වකවානුව, 1820 - 1900 වකවානුව හා 1900 - 1950 වකවානුව එම වකවානු වන අතරමෙහි දී ඒ එක් එක් වකවානුවෙහි සුවිශේෂතා වෙන් වෙන් ව ගෙන විෂයයෙහි ක්‍රමික විකාශය විමර්ශනය කැරේ.

මුඛ්‍ය පද: *සම්භාවිතික නො වන දත්ත විශ්ලේෂණය, සම්භාවිතා සංකල්පය, ගණිතමය සංඛ්‍යානය, සංඛ්‍යානමය ආකෘති, අක්ෂීය ප්‍රවේශය*

හැඳින්වීම

සංඛ්‍යාන විද්‍යාවේ ඉතිහාසය හැදෑරීමට මත්තෙන් සංඛ්‍යානය යන්න සඳහා යෙදෙන 'Statistics' යන වචනයේ සම්භවය සිදු වූ ආකාරය විමැසීම වැදගත් ය. 'Statistics' යනු ලතින භාෂාවේ 'Status' (ස්ටෙයිටස්) යන වචනයෙන් ද ඉතාලි භාෂාවේ 'Statista' (ස්තතීස්තා) යන වචනයෙන් ද නව ලතින භාෂාවේ 'Statisticum Collegium' (ස්ටටිස්ටිකම් කොලීජියම්) යන වචනයෙන් ද භින්න වූවෙකැ යි සැලැකේ (Harper, 2001). 'ස්ටෙයිටස්' යන්න 'දේශපාලන ජනපදය' (Political State) යන අර්ථයෙන් ද 'ස්තතීස්තා' යන්න 'රාජ්‍යතාන්ත්‍රික කටයුතුවලට දක්ෂයා' (One Skilled

[1]සමාජ සංඛ්‍යානය ශාස්ත්‍රවේදී (විශේෂ) ප්‍රථම වසර,
muthuminibandara@gmail.com

in Statecraft) යන අර්ථයෙන් ද 'ස්ටටිස්ටිකම් කොලීජියම්' යන්න 'ජනපද කටයුතු සම්බන්ධ දේශන පාඨමාලාව' (Lecture Course on State Affairs) යන අර්ථයෙන් ද එකල භාවිත කොට තිබේ (Harper, 2001). පසු කාලීන ව මේ 'Statistics' යන්න සෙසු භාෂා කරා ද ව්‍යාප්ත විය. ඒ අනුව 1749 දී ගොඩ්ෆ්‍රිඩ් ඇක්න්වෝල් (Gottfried Achenwall) විසින් 'Statistik' (ස්ටටිස්ටික්) යනුවෙන් ජර්මන් භාෂාවට ද 1791 දී ශ්‍රීමත් ජෝන් සින්ක්ලෙයාර් (Sir John Sinclair) විසින් 'Statistics' (ස්ටැටිස්ටික්ස්) යනුවෙන් ඉංගිරිසි භාෂාවට ද මෙය හඳුන්වා-දෙන ලදී.

'Statistics' යන වචනය අතීතියේදී 'ජනපදය' යන නැතහොත් 'ජනපද වැසියා' යන අර්ථවලට සීමා වී තුබුණ ද වර්තමානයේ එය අර්ථ විස්තරණයකට (Semantic Expansion) පත් ව තිබේ. දත්ත එක් රැස් කිරීම, දත්ත සංවිධානය, දත්ත විශ්ලේෂණය හා විශ්ලේෂිත දත්ත පදනම් කැර-ගෙන යුක්ති සහගත වූත් සාධාරණ වූත් නිගමන කරා එළැඹීම යන අර්ථ සමූහය සූචනය කිරීම පිණිස වර්තමානයේ දී 'Statistics' යන වචනය භාවිත කැරේ. 'Statistics' යන්නට සමානාර්ථයෙන් යෙදෙන සිංහල පාරිභාෂිකය වන 'සංඛ්‍යානය' යන්න ද ඉහත අර්ථ සූචනය කිරීමට යෙදූ සංස්කෘත වචනයකි.

සංඛ්‍යානයේ බොහෝ කාර්යයන් සිදු කරනු ලබන්නේ සම්භාවිතික භාෂාවක් (Language of Probability) භාවිත කරමිනි. අවිනිශ්චිත තත්ත්වයන් විස්තර කිරීම සඳහා සම්භාවිතා න්‍යායය (Probability Theory) භාවිත කිරීම ද සංඛ්‍යානයේ දී සිදුවේ. සම්භාවිතා න්‍යායය පැහැදිලි ව ම ගණිතයේ ශාඛාවකි. ගණිතයෙන් විද්‍යාව වෙන් වන ප්‍රදේශය ලෙස සංඛ්‍යානය හැඳින්විය හැකි ය. කෙසේ වෙතත් සම්භාවිතාව හා සංඛ්‍යානය අතර පැහැදිලි අන්තර්සම්බන්ධතාවක් දැකිය හැකි ය. සම්භාවිතාව හා සංඛ්‍යානය අතර මෙසේ අන්තර්සම්බන්ධතාවක් ඇති නිසා ම සම්භාවිතා න්‍යායය සංවර්ධනයට සමාන්තර ව සංඛ්‍යාන විෂයය ද වර්ධනය විය. එබැවින් සංඛ්‍යානයේ ඉතිහාසය අධ්‍යයනය කරන විට සම්භාවිතා න්‍යායයේ සංවර්ධනය හා සංඛ්‍යානයට එයින් ලැබුණු දායකත්වය පසෙකට දැමිය නොහැකි ය. මෙහෙයින් සම්භාවිතා න්‍යායය සංවර්ධනය වූ කාල සීමාවන් පදනම් කැර-ගෙන සංඛ්‍යාන විෂයයේ ඉතිහාසය ප්‍රධාන වශයෙන් වකවානු හතරකට බෙදා දැක්විය හැකි ය. මෙහිදී ඒ එක් එක් වකවානුව වෙන් වෙන් ව විමර්ශනය කැරේ.

1650 - 1750 වකවානුව: ප්‍රාරම්භය

"සැබැවින් ම සංඛ්‍යාන විෂයය ආරම්භ වූයේ කවරදා ද?" යන ප්‍රශ්නයට පිළිතුරු සැපැයීමට උත්සාහ දරන මෞරිස් කෙණ්ඩෝල් (Maurice Kendall) 1960 දී "Where shall the history of statistics begin?" යන ලිපිය ලියමින් තර්ක කරන්නේ කිසි යම් විෂයයක ඉතිහාසය සිහින්

තන්තුවට ම සොයා-යෑම භයානක වන බවට යි. තව දුරටත් අදහස් දක්වන ඔහු පවසන්නේ ක්‍රි. ව. 1660ට ප්‍රථම සංඛ්‍යානයේ ඉතිහාසය සෙවීම කිසිසේත් කළ නොහැක්කක් වන අතර ජෝන් ග්‍රාන්ට් (John Grant) ගේ කටයුතු ආශ්‍රිත කාලය සංඛ්‍යානයේ ආරම්භක ලක්ෂ්‍යය ලෙස දැක්විය හැකි බවයි (Fienberg, 1992). ස්ටෙෆන් එම්. ස්ටිග්ලර් (Stephen M. Stigler) ද ඒ අදහසට ම සමාන අදහසක් දරයි. 1990 වර්ෂයේ දී "The History of Statistics: The Measurement of Uncertainty before 1900" ග්‍රන්ථය ලියමින් ඔහු පවසන්නේ සම්භාවිතාව හා බැඳුණු සංඛ්‍යාන ක්‍රම සංවර්ධනය වීම ඇරඹුණු ස්ථානයෙන් සංඛ්‍යානයේ ඉතිහාසය සීමා කිරීම සහේතුක බවයි (Fienberg, 1992). සම්භාවිතාව සංවර්ධනය වූ අවධිය හා සම්භාවිතික නො වන (Non-probabilistic) දත්ත විශ්ලේෂණ ක්‍රම ආරම්භ වූ අවධිය පූර්ව ඉතිහාසය (Pre-history) හෙවත් සංඛ්‍යානයේ ප්‍රාරම්භය ලෙස සැලැකිය හැක්කේ එබැවිනි.

සම්භාවිතික නො වන දත්ත විශ්ලේෂණ ක්‍රම ආරම්භ වූයේ 17 වැනි සියවසේ මැද භාගයේ දී ය. ඒ විලියම් පෙටී (William Petty) විසින් 1690 දී "Political Arithmetic" නම් ග්‍රන්ථය ලියමින් දේශපාලන අංක විද්‍යාව (Political Arithmetic) හඳුන්වා දීමත් සමඟ ය. හෙතෙමෙ ඒ ග්‍රන්ථයෙහි මෙසේ සඳහන් කරයි.

> "...*දේශපාලන අංක විද්‍යාවේ දී සලකා-බලනුයේ භූමිය, මිනිසුන්, වගාබිම්, පුද්ගල ආදායම්, බදු, පොලිය, බැංකු ආදි වශයෙන් රටක පවතින විවිධ තත්ත්වයන් ගේ අගයන් හා ප්‍රමාණයන් පිළිබඳ ව ය*...(Stochastikon, n.d.), "

මේ ප්‍රකාශය අනුව පෙටී-යන කරුණ නම් දේශපාලනික වශයෙන් වඩාත් නිවැරැදි තීරණ ගැනීම සඳහා ස්වාභාවික, සාමාජික හා දේශපාලනික වශයෙන් රටකට වැදැගත් වන අගයයන් හා ප්‍රමාණයන් පිළිබඳ වාර්තා තබා-ගෙන ඇති බවයි. පෙටී ගේ මේ අදහස් ඔහු ගේ මිත්‍රයකු වූ ජෝන් ග්‍රාන්ට් (John Grant) ගේ අදහස්වල ආභාසය ලබා ඇත. 1692 දී "Natural and Political Observation on the London Bills of Mortality" ග්‍රන්ථය ලියූ ජෝන් ග්‍රාන්ට් දත්ත විශ්ලේෂණ ප්‍රවේශ ගණනාවක් එමගින් හඳුන්වා-දුන්නේ ය.

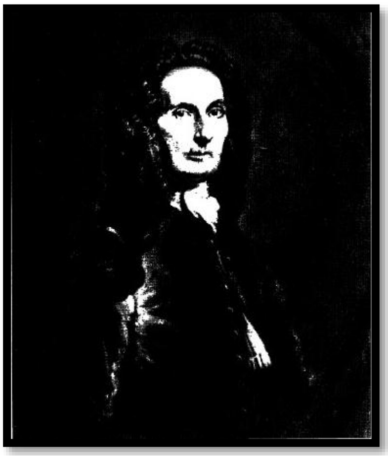
- වසර 60ක කාලයක් තිස්සේ නිකුත් කරන ලද මරණ සහතිකවල (Bills of Mortality) පිළිගතහැකි බව හෙවත් වාස්තවිකත්වය (Acceptability) පරීක්ෂා කිරීම,
- මහාමාරිය වසංගතයෙන් සිදු වූ මරණ පිළිබඳ විශ්ලේෂණය කිරීම,
- පුම්තිරි අනුපාතිකය සම්බන්ධව සවිස්තරාත්මකව විස්තර කිරීම හා විශ්ලේෂණය,

- ආනුභවික පදනමක් මත ජීවන වගුවක් (Life Table) සකස් කිරීම ආදිය ඔහු අතින් සිදු විය (Fienberg, 1992).

ග්‍රාන්ට් ගේ මූලික අදහස වූයේ සාමාන්‍යය (Average) සෙවීම වන අතර සැහෙන කාලයක් යන තුරු එය සංඛ්‍යානයේ මූල අදහසව පැවැතිණි. නියැදියේ ප්‍රමාණය හා ස්ථාවරත්වය අතර සම්බන්ධතාව පිළිබඳව සැලැකිල්ලක් හෝ සම්භාවිතික සංකල්ප පිළිබඳ අවබෝධයක් ඔහුට නොතිබිණි. කෙසේ වෙතත් සසම්භාවී ලෙස ජනිත කැර-ගත් දත්ත සමග ගනුදෙනු කිරීම සඳහා පොදු ප්‍රවේශයක් හෙතෙමෙ හඳුන්වා-දුන්නේය. සාමාන්‍යයන් පදනම් කැර-ගනිමින් ස්ථාවර නියමයන් (Stable Laws) නැත හොත් සංඛ්‍යානමය නියමයන් (Statistical Laws) හඳුනා-ගැනීම එහි දී සිදු විය.

ග්‍රාන්ට්ගේ සහ පෙටීගේ අදහස් ඉදිරියට ගෙන-යෑමට ඒබ්‍රහම් දි මොරේ (Abraham de Moivre) හා ජේකබ් බර්නූලි (Jacob Bernoulli) වැන්නෝ කටයුතු කළහ. මහාද්වීපික යුරෝපය පුරා ඒවා ව්‍යාප්ත කිරීමට ද ඔවුහු සමත් වූහ.

රූපයː 01
ජේකබ් බර්නූලිː 1654 - 1705



රූපයː 02
ඒබ්‍රහම් දි මෝරේː 1667 - 1754

මූලාශ්‍රයː Fienberg, Stephen E. A Brief History of Statistics in Three and One-Half Chapters: A Review Essay. Statistical Science 7 (1992), no. 2, Pg. 213.

ජේකබ් බර්නූලිගේ මරණයෙන් පසු 1713 දී ඔහුගේ 'විශාල සංඛ්‍යාවන් සඳහා වන නියමය' (Law of Large Numbers/ LLN Theorem) ඉදිරි පත් වීමත් සමග සම්භාවිතා සංකල්පය පිළිබඳ විෂයානුබද්ධ සාකච්ඡා ප්‍රථම වතාවට ඇති විය. 1718 සිට 1738 දක්වා කාලයේ දී සම්භාවිතා සංකල්පය සැලැකිය යුතු මට්ටමේ දියුණුවක් අත් කැර-ගත්තේ ය. 1718 මෝරේ ගේ

"Doctrine of Chances" ග්‍රන්ථයේ දෙවන සංස්කරණයෙන් ද්විපද ව්‍යාප්තිය (Binomial Distribution) සඳහා විවිධ ගණිතමය සන්නිකර්ෂණ (Mathematical Approximations) ගොඩනැඟිණි (Fienberg, 1992). ඒ අතර සාමාන්‍ය සන්නිකර්ෂණය (Normal Approximation) සුවිශේෂ වේ. අද වන විට සම්භාව්‍ය සම්භාවිතා න්‍යායය ලෙස හඳුනා-ගනු ලබන්නේ මෝර්රෙගේ කටයුතු පසුකාලීනව සාර්ථක ලෙස සංවර්ධනය කිරීමේ ප්‍රතිඵලයකි.

1750 - 1820 වකවානුව: අනුමානය හඳුන්වා-දීම හා ගණිතමය සංඛ්‍යානයේ ආරම්භය

1750න් ආරම්භ වන වකවානුව ප්‍රධාන සංඛ්‍යානමය ක්‍රියාකාරකම් ද්වයයකින් නිර්මාණය වී ඇත. පළමුවැන්න රේඛීය සමීකරණවල නො දන්නා සංගුණක නිමානය කිරීම සඳහා අඩුතම වර්ග ක්‍රමයේ (Method of Least Squares) සංවර්ධනය යි; දෙ වැන්න සම්භාවිතාව පදනම් කැරැ-ගත් අනුමාන සංවර්ධනය වීම යි.

සම්භාවිතාව පදනම් කැරැ-ගත් අනුමාන සංවර්ධනය

ඉහත කී පරිදි ඒබ්‍රහම් දි මෝර්රේ විසින් ද්විපද ව්‍යාප්තිය සඳහා සාමාන්‍ය සන්නිකර්ෂණය හඳුන්වා-දෙනු ලැබිණි. එහෙත් ද්විපද පරාමිතිය (Binomial Parameter) පිළිබඳ අනුමානයන් සිදු කිරීමට සාමාන්‍ය සන්නිකර්ෂණය භාවිත කැරුණේ නැත. ඒබ්‍රහම් දි මෝර්රේ හා ජේකබ් බර්නූලිට ද්විපද පරාමිතිය පිළිබඳ අනුමාන කරා එළැඹීමට නොහැකි වුව ද 1764 දී එයට පිළිතුරක් සැපැයීමට තෝමස් බේයස් (Thomas Bayes) සමත් විය. හෙතෙමෙ ප්‍රතිලෝම සම්භාවිතා ප්‍රවේශයක් භාවිත කරමින් ද්විපද පරාමිතය පිළිබඳ අනුමානයන් ලබා-ගත්තේ ය. මේ ප්‍රතිලෝම සම්භාවිතා ප්‍රවේශය පසුකාලීන ව බේයස් ප්‍රමේයය (Bayes Theorem) නමින් ප්‍රසිද්ධ විය. බේයස් ගේ මාර්ගයේ ම ගමන් ගත් පියරේ සිමොන් ලප්ලාස් (Piere Simon Laplace) ප්‍රතිලෝම සම්භාවිතා තර්ක ක්‍රමයක් භාවිත කරමින් සෙසු සම්භාවිතා ව්‍යාප්ති සඳහා ද අනුමානයන් ලබා-ගත්තේ ය. තව ද 1810 දී මධ්‍ය සීමා ප්‍රමේයය (Central Limit Theory) නිර්මාණය කරමින් ඕනෑ ම සම්භාවිතා ව්‍යාප්තියක එකතුව/සාමාන්‍යය ව්‍යාප්ත වී ඇති ආකාරය සන්නිකර්ෂණය කිරීම සඳහා ප්‍රමත ව්‍යාප්තිය (Normal Distribution) භාවිත කිරීම සාධාරණීකරණය කරනු ලැබිණි.

අඩුතම වර්ග ක්‍රමයේ සංවර්ධනය

වසර 1805 දී ලජාන්ඩ්ර් (Legendre) විසින් අඩුතම වර්ග ක්‍රමය සංවර්ධනය කර ප්‍රථමයෙන් එය ප්‍රසිද්ධ කරන ලදී (Stigler, 1981). එහි දී ඔහු විසින්

ආවේණික වූ සීමා කරන ලද සමීකරණ භාවිතයෙන් නිරීක්ෂණය පදනම් කැර-ගෙන නොදන්නා සංගුණකයන් පිළිබඳ අනුමාන ලබා-ගන්නා ලදී.

එහෙත් පර්යේෂණ ක්ෂේත්‍රවල නිර්මාණය වී තුබුණු අවිනිශ්චිතතාවන් (විශේෂයෙන් 'සඳ නිදහස් කිරීම' "Liberation of Moon") අඩුතම වර්ග ක්‍රමය යොදා-ගෙන ප්‍රමාණීකරණය කරන්නේ කෙසේ ද? යන්න එකල අඩුතම වර්ග ක්‍රමය සම්බන්ධයෙන් පැවැති ප්‍රබල ගැටලුවක් විය. දක්ෂ ජර්මානු ගණිතඥයකු වූ කාර්ල් ෆ්‍රීද්රික් ගෑස් (Karl Friedrich Gaus) 1809 දී "Theoria motus corporum coelestium in sectionibus conicus solem ambientium" ග්‍රන්ථය ලියමින් රේඛීය සමීකරණ (Linear Equations) පද්ධතිවල සාමාන්‍ය වශයෙන් ව්‍යාප්ත වුණු දෝෂ පද භාවිත කිරීම සාධාරණීකරණය කිරීම සඳහා වක්‍රාකාර තර්කයක් ගොඩනගමින් පෙන්වා-දෙන්නේ දෝෂයන්ගේ පොස්ටීරියර් ව්‍යාප්තිය උපරිමකරණය (Maximising the Posterior distribution of errors) අඩුතම වර්ග ක්‍රමය භාවිත කිරීමට සමාන වන බවයි (Fienberg, 1992). ඒ සමඟ ම ඉහත ගැටලුව නිරාකරණය වේ.

මේ සොයා-ගැනීම ලාප්ලාස් වෙත බලපෑම් එල්ල කළ අතර 1810 දී සාමාන්‍ය දෝෂ පදය (Normal Error Term) තම මධ්‍ය සීමා ප්‍රමේයය යොදා-ගෙන සාධාරණීකරණය කළ හැකි බව ඔහු හඳුනා-ගත්තේ ඒ හේතුවෙනි. මේ හඳුනා-ගැනීමත් සමඟ ම සම්භාවිතාව පදනම් කැර-ගත් අනුමාන සංවර්ධනය හා අඩුතම වර්ග ක්‍රමය සංවර්ධනය යනුවෙන් මාර්ග ද්වයයක් ලෙස සංවර්ධනය වූ සංඛ්‍යානය ඒකාබද්ධ විය. පුළුල් පරාසයක් භෞතික ගැටලුවලට විසැඳුම් සෙවීමට උපයෝගී කැර-ගත හැකි ප්‍රබල සංඛ්‍යානමය ප්‍රවේශයක් මේ ඔස්සේ නිර්මාණය විය. ස්ටෙෆන් එම්. ස්ටිග්ලර්ගේ "The History of Statistics: The Measurement of Uncertainty Before 1900" යන ග්‍රන්ථයෙහි මේ ඒකාබද්ධ වීම 'ගෑස්-ලප්ලාස් සංයෝගය' (Gauss-Laplace Synthesis) ලෙස හඳුන්වා තිබේ (Fienberg, 1992).

1820 - 1900 වකවානුව: සංඛ්‍යානය සමාජගතවීම සහසහසම්බන්ධතාව හා සංඛ්‍යානමය ආකෘති සංවර්ධනය

දහ නව වැනි සියවසේ සංඛ්‍යානය සංවර්ධනය වූ ආකාරය අධ්‍යයනය කිරීමේ දී ඇඩොල්ෆ් කියුටලට් (Adolphe Quetelet) නම් බෙල්ජියම් ජාතිකයා වැදැගත් වේ. ඔහු ගේ අග්‍රගණ්‍ය කෘතිය ලෙස සැලැකෙන "Sur I'homme et le developpment de ses aculties ou Essai de physique sociale" කෘතිය මඟින් සංඛ්‍යානය යොදා-ගෙන සාමාජීය ප්‍රපඤ්චයන්, සාමාන්‍ය මිනිසා (Average Man) ආදිය නිර්වචනය කරමින් ප්‍රමාණාත්මක සමාජ විද්‍යාව ආරම්භ කෙළේ ය. එකල ඔහු එය හැඳින්වූයේ සමාජ

භෞතික විද්‍යාව (Social Physics) යනුවෙනි (Coven, 2003). ජෝන් ග්‍රාන්ට්ගේ සාමාන්‍යය සෙවීම (Averaging) කියුටලට් ගේ සමයයේ දී නවමු මානයකට ගමන් කෙළේ ය. සාමාන්‍ය මිනිසා නිර්වචනය කිරීමේ දී සාමාන්‍යය යොදා-ගන්නා කියුටලට් සාමාන්‍ය මිනිසාගේ සිට දැකිය හැකි ඕනෑ ම විචලතාවක් (Deviation) සසම්භාවී දෝෂයක් ලෙස සලකයි. එවැනි මිනිසුන් දෝෂ නීතියේ (Error Law) මධ්‍යනයන් යොදා-ගෙන විග්‍රහ කරනු ලබයි. දෝෂයන් ගේ ව්‍යාප්තිය විස්තර කිරීම සඳහා පමණක් නො ව සාමාන්‍ය කටයුතු සඳහා ද කියුටලට් විසින් සාමාන්‍ය සම්භාවිතා න්‍යායය (Normal Probability Law) යොදා-ගන්නා ලදි. ඒ නිසා ප්‍රමත ව්‍යාප්තිය සංඛ්‍යාන විෂයය පුරා ආධිපත්‍යය පැතිරවීය.

රූපයඃ 03

ඇඩොල්ෆ් කියුටලට් 1796 - 1874

මූලාශ්‍රයයඃ http://apprendre-math.info/history/photos/Quetelet.jpeg

සංඛ්‍යානය සමාජගත වීම

සංඛ්‍යාන විෂයයේ ප්‍රවර්ධනය සඳහා කියුටලට් සුවිශාල දායකත්වයක් සැපැයූ බව අවිවාදයෙන් කිය හැකි ය. ඔහු සතුව පැවැතුණු ආනුභවික සමාජ දත්ත (Empirical Social Data) එකතුව හේතුවෙන් 1829 දී ලොව ප්‍රථම ජාතික සංගණන ඒකකය බෙල්ජියම හා නෙදර්ලන්තය තුළ පිහිටුවීමට ඔහුට හැකි විය. කියුටලට්ගේ මේ පරමාදර්ශී කටයුතු හේතුවෙන් යුරෝපය පුරා සංඛ්‍යානයට ඇති අවධානය ඉහළ ගොස් සමස්ත යුරෝපය ම ආවරණය වන පරිදි සංඛ්‍යාන කාර්යාලයයන් (Statistical Bureaus) පිහිටුවීමට ඔහුට හැකි විය. එමෙන්ම අන්තර්ජාතික වශයෙන් දත්ත එක් රැස් කිරීමට හා දත්ත විශ්ලේෂණය කිරීමට ඒකාකාර ක්‍රමවේදයන් හා වචන මාලාවක් ව්‍යාප්ත කිරීමට ද හෙතෙමෙ වෙහෙස විය. ඔහු ගේ නායකත්වය යටතේ ලොව ප්‍රථම අන්තර්ජාතික සංඛ්‍යාන සම්මේලනය 1853 දී බෙල්ජියමේ බ්‍රසල්ස් නුවර පැවැත්විණි.

සහසම්බන්ධතාව හා සංඛ්‍යානමය ආකෘති සංවර්ධනය

1880 වසරේ සිට විෂයය වර්ධනය වන වේගයෙහි පැහැදිලි වෙනසක් දැකිය හැකි විය. විශේෂයෙන් එංගලන්තයෙහි මෙය පැහැදිලි ව පෙනෙන්නට තිබිණි. ෆ්‍රැන්සිස් ගැල්ටන් (Francis Galton), ෆ්‍රැන්සිස් එඩ්ග්වර්ත් (Francis Ysidro Edgeworth), කාල් පියර්සන් (Karl Pearson), ජෝජ් උද්නි යුලේ (George Udny Yule) වැන්නන්ගේ නායකත්වය යටතේ සංඛ්‍යානයට සුවිශාල දායකත්වයක් ලැබීම එයට හේතුවිය. එංගලන්තය තුළ සංඛ්‍යාන ප්‍රබෝධය ආරම්භ කෙළේ ෆ්‍රැන්සිස් ගැල්ටන් ය. 1869 දී " Hereditary Genius" නමින් අධ්‍යයනයක යෙදෙන හෙතෙමෙ ප්‍රතිපායනය (Regression) සම්බන්ධ ව මුල්ම අදහස් ඉදිරිපත් කෙළේ ය (Fienberg, 1992). ඉන් වසර දහසයකට පසු අදහස් තීව්‍ර කැර-ගත් හේ 1885 දී ප්‍රපඤ්චයක් පිළිබඳ වූ ප්‍රථම ප්‍රතිපායන විශ්ලේෂණය ද සිදු කෙළේය. ආනුභවික නිදසුනක් යොදා-ගනිමින් ඔහු විසින් එය සිදු කරන ලදී. එහි දී ළමයින්ගේ උස සහ ඔවුන්ගේ දෙමාපියන්ගේ උසෙහි සාමාන්‍යය යන විචල්‍ය දෙක පදනම් කැර-ගෙන මේ ප්‍රතිපායන විශ්ලේෂණය සිදු කැරිණි.

ප්‍රතිපායනය හා සමගාමී සංකල්පයක් වන සහසම්බන්ධතාව (Correlation) පිළිබඳ ව 1888 දී ගැල්ටන් අදහස් දැක්වී ය. ඔහුගේ මේ අදහස් එංගලන්තයේ වුසූ සෙසු විද්‍යාඥයන් විසින් නොපමා ව ම තව දුරටත් වර්ධනය කැරිණි. ඒ කටයුත්තෙහි නියැලුණු අය අතර ෆ්‍රැන්සිස් එඩ්ග්වර්ත් කැපී-පෙනේ.

ඉන්පසු කාල් පියර්සන්ගේ සහ ජෝජ් උද්නි යුලේගේ දායකත්වය විශේෂ අවධානයට ලක්විය යුතු වේ. පියර්සන් 1890 දී කුටිකතා වක්‍ර (Skew Curves) විශ්ලේෂණය කිරීමේ ක්‍රමවේද හඳුන්වා-දුන්නේ ය. 1750 - 1820 අතර කාලයේ දී වර්ධනය වූ අඩුතම වර්ග ක්‍රමය හා දෝෂ පිළිබඳ න්‍යායය (Theory of Errors) තවදුරටත් වර්ධනය කිරීමට යුලේ කටයුතු කළ අතර 1897 දී ආංශික සහසම්බන්ධතාව (Partial Correlation) හා බහුගුණ සහසම්බන්ධතාව (Multiple Correlation) පිළිබඳ සංකල්ප ලොවට හඳුන්වා-දිනි. එයින් දස වසරකට පසු එනම් 1907 දී හෙතෙමෙ වර්තමානයේ පවා ප්‍රතිපායන විශ්ලේෂණයේ දී භාවිත වන නූතන සංකේත (Modern Notations) සමුදායය හඳුන්වා-දුන්නේ ය.

මේ ශතවර්ෂය අවසන් වන්නේ කාල් පියර්සන් විසින් කයි වර්ග පරීක්ෂාව (Chi-Square Test) හඳුන්වා-දීමත් ස්වාධීන ක්‍රමවේද බිහිකිරීම මුල්කැර-ගත් ලොව ප්‍රථම සංඛ්‍යාන සඟරාව වූ 'බයෝමෙට්‍රිකා' (Biometrika) නම් සඟරාව එළිදැක්වීමත් සමඟ ය.

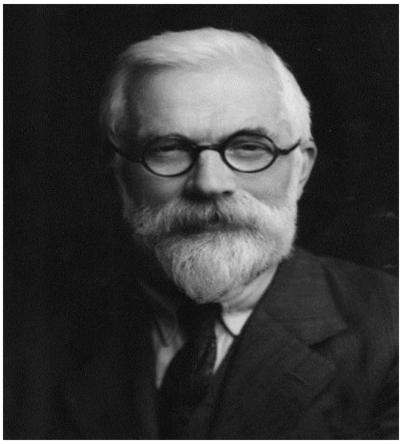
1900 - 1950 වකවානුව: නූතන සංඛ්‍යාන අවධිය

නූතන සංඛ්‍යාන අවධියේ ප්‍රමුඛයා ලෙස සැලැකෙන්නේ ඉංගිරිසි ජාතික සංඛ්‍යාන විද්‍යාඥයකු වන රොනල්ඩ් ඒ. ෆිෂර් (Ronald A. Fisher) ය. සංඛ්‍යාන විෂයයෙහි ඔහු සක්‍රීය ව කටයුතු කළ කාලයෙන් නූතන සංඛ්‍යාන අවධිය ආරම්භ වේ. 1750න් පසු ගණිතමය සංඛ්‍යානය ආරම්භ වුව ද එහි දෙවැනි රැල්ල ක්‍රියාත්මක වූයේ මේ අවධියේ දී ය. එහි පුරෝගාමියා වූයේ රොනල්ඩ් ෆිෂර් ය. 1909 දී කේම්බ්‍රිජ් විශ්වවිද්‍යාලයයට ඇතළත් වූ ෆිෂර් 1912 දී තම මුල් ම පර්යේෂණ වාර්තාව එළිදැක්වී ය (Fienberg, 1992). ඔහු ගේ සංඛ්‍යානමය අදහස් විද්‍යාත්මක දත්ත විශ්ලේෂණය සඳහා අප්‍රමාද ව බලපෑම් කෙළේ ය. උදාහරණයක් ලෙස පර්යේෂණ සැලැස්ම (Design of Experiment) හා විචලතා විශ්ලේෂණය (Analaysis of Variance) යන සංකල්ප දත්ත විශ්ලේෂණය කෙරහි සිදු කළ බලපෑම දැක්විය හැකි ය. නව සංඛ්‍යාන ක්‍රම බිහිකිරීමේ අරමුණින් හෙතෙමෙ සැබෑ දත්ත විශ්ලේෂණය සඳහා ක්‍රමානුකූල ප්‍රවේශයක් සකස් කෙළේය. ඔහු ගේ පළමු වැනි කෘතිය වූයේ 1925 දී රචිත ‘Statistical Methods for Research Workers’ නම් ග්‍රන්ථයයි.

කෙසේ වුව ද ඔහු හඳුන්වා-දුන් ඇතැම් සංකල්ප දැඩි ලෙස මත භේදයට තුඩු දුන් ඒවා විය. ඒ අතරෙහි වෙසෙසියාව පරීක්ෂා කිරීමට ඔහු ගොඩනැගූ සූත්‍ර, සම්භාවිතික පදනමක් මත පිහිටා නොදන්නා සංගහන පරාමිතියක් සඳහා ප්‍රාන්තර උපකල්පනය (Interval Estimates) ආදිය එලෙස දැඩි මත භේදයට තුඩු දුන් සංකල්ප ය (Fienberg, 1992).

රූපය: 04

රොනල්ඩ් ඒ. ෆිෂර්: 1890 - 1962



මූලාශ්‍රය: http://trailblazing.royalsociety.org/photos/1922SA.jpg

චාල්ස් ඊ. ස්පියර්මන් (Charles E. Spearman) විසින් 1904 දී තරා සහසම්බන්ධතා සංගුණයක (Rank Correlation Coefficient) හඳුන්වා-දෙනු ලැබීම (Encyclopedia.com, 2008) හා 1908 දී විලියම් සීලී ගොසෙට් (William Sealy Gosset) විසින් 'ස්ටුඩන්ට්' නමින් හඳුන්වන t ව්‍යාප්තිය (Students' t Distribution) හඳුන්වා-දෙනු ලැබීම ආදිය ද මේ වකවානුවේ සිදු විණි (Encyclopedia Britannica, 2014). කෙසේ වෙතත් මේ යුගයට අයත් වැදගත් ම කටයුත්ත සිදු කරන ලද්දේ 1930 දශකයේ ආරම්භයේ දී රුසියානු ගණිතඥ ඒ.එන්. කොල්මොගොරෝව් (A. N. Kolmogorov) විසිනි. සම්භාවිතාව සඳහා අක්ෂීය ප්‍රවේශයක් (Axiomatic Approach) නිර්මාණය කළ හෙතෙමෙ කුලක න්‍යාය (Set Theory) හා ශ්‍රිතයන් පිළිබඳ න්‍යායය (Theory of Functions) ද හඳුන්වා-දුන්නේය. එමගින් සම්භාවිතාව පැහැදිලි ව හඳුනා ගත හැකි පරිදි ගණිතයේ උපක්ෂේත්‍රයක් බවට පත් විය. මේ හඳුන්වා-දීම සම්භාව්‍ය සම්භාවිතා න්‍යායයට ගණිතමය පදනමක් එක් කෙළේ ය. එමෙන් ම එමගින් සංඛ්‍යානමය න්‍යායයන් (Statistical Theory) ගෙන් ලැබෙන ගණිතමය ප්‍රතිඵල සනාථ කිරීම සඳහා ප්‍රබල පදනමක් එක් විය.

මේ යුගයේ තවත් කැපී පෙනෙන මෙහෙවරක් සංඛ්‍යානය විෂයයට ඉටු කිරිමට කාල් පියර්සන් ගේ පුත් ඒගන් පියර්සන් (Egon Pearson) හා ජර්සි නේමාන් (Jerzy Neyman) සමත් වූ හ. ඔවුහු දෙදෙනා එක් ව දෙ වැනි ප්‍රරූපයේ දෝෂය (Type II error), සංඛ්‍යාන කල්පිත පරික්ෂාවන්හි බලය (The power of test) සහ විශ්‍රම්භ ප්‍රාන්තර (Confidence Intervals) යන සංකල්ප හඳුන්වා-දුන්හ (Statisticians in History, Year Unknown). මේ සංකල්ප ඔස්සේ වැඩි දුර යෙමින් පසු කාලීන ව ඔවුහු 'නේමාන්-පියර්සන් න්‍යායය' (Neyman-Pearson Theory) හඳුන්වා-දුන්හ.

1937 දී ජර්සි නේමාන් ඇමෙරිකාවට සංක්‍රමණය වී හැරල්ඩ් හොට්ලින්ග් (Harold Hotelling), ඒබ්‍රහම් වෝල්ඩ් (Abraham Wold) සහ සැමුවෙල් විල්ක්ස් (Samuel Wilks) සමඟ එක්වී ඇමෙරිකා එක්සත් ජනපදය තුළ සංඛ්‍යාන විෂයයෙහි පිබිදීමක් ඇති කිරිමට දායක විය. 1938න් පසු දෙ වැනි ලෝක යුද්ධ සමයයේ දී හමුදා ශික්ෂණයක් තුළ සංඛ්‍යානය විෂය වර්ධනය විය.

සමාලෝචනය

මෙතෙක් විමර්ශනය කැරුණු ඉතිහාස ප්‍රවෘත්ති පිඬු කොට-ගත් කල සංඛ්‍යානය විෂයය ස්වාධීන අධ්‍යනයක් ලෙස දහ හත් වැනි සියවසේ දී ආරම්භ වූ බව පැහැදිලි ය. එවක සංඛ්‍යානය සීමා වී තුබුණේ දත්ත එක් රැස් කිරීමට පමණි. එලෙස දත්ත එක් රැස් කිරීමෙන් එළැඹෙන නිගමනවල

වාස්තවිකත්වය සම්බන්ධ ව ගැටලුකාරී තත්ත්වයන් තුබුණු නමුත් පසුකාලීන ව සම්භාවිතා න්‍යායයේ වර්ධනයත් සමඟ සංඛ්‍යානමය නිගමනවල වාස්තවිකත්වය ද වර්ධනය වූ බව පෙනේ. සංඛ්‍යාන විෂයයෙහි ප්‍රභවය හා විකාශය පිළිබඳ අධ්‍යයනයේ දී විලියම් පෙටි, ජෝන් ග්‍රාන්ට්, ඒබ්‍රහම් ඩී මෝරේ, ජේකබ් බර්නූලි, ඇඩොල්ෆ් කියුටලට්, කාල් පියර්සන්, රොනල්ඩ් ඒ. ෆිෂර්, ඒ.එන්. කොල්මොගොරෝව් යන විද්‍යාඥයන් ගෙන් ඉටු වූ මෙහෙය ප්‍රශංසනීය බව අවසන් විසින් සඳහන් කළ යුතු වේ.

මූලාශ්‍රය

Coven, V. (2003). A History of Statistics in the Social Sciences. *An Academic Journal on the web*, 1-10. Retrieved from http://grad.usask.ca/gateway/art_Coven_spr_03.pdf

Encyclopedia Britannica. (2014). *Student's t-test*. Retrieved from http://www.britannica.com/EBchecked/topic/569907/Students-t-test

Encyclopedia.com. (2008). *Spearman Rank Correlation Coefficient*. Retrieved from http://www.encyclopedia.com/doc/1G2-3045302562.html

Fienberg, Stephen E. (1992). A Brief History of Statistics in Three and One-Half Chapters: A Review Essay. *Statistical Science*, 7(2), 208 - 225. doi:10.1214/ss/1177011360

Harper, D. (2001). *Online Etymology Dictionary*. Retrieved from http://www.etymonline.com/index.php?allowed_in_frame=0&search=statistics&searchmode=none

Roussas, G. G. (n.d.). Year Unknown.Probability and Statistics throughout the Centuries. University of California. Retrieved from http://anson.ucdavis.edu/~roussas/Probability_And_Statistics_Throughout_TheCenturies.pdf

Statisticians in History, (n.d.). *Jerzy Neyman 1894-1981*. Retrieved from http://www.amstat.org/about/statisticiansinhistory/index.cfm?fuseaction=biosinfo&BioID=11

Stigler, Stephen M. (1981). Gauss and the Invention of Least Square.*The Annals of Statistics*, 9(3), 465-474. doi:10.1214/aos/1176345451

Stochastikon. (n.d). *History of statistics*. Retrieved from 132.187.98.10:8080/encyclopedia/en/statisticsHistory.pdf