

Визначення гіпотези нормальності розподілу вибірки технологічних змінних за допомогою MatLab

Л.О. Власенко, А.П. Ладанюк, С.А. Буров

Національний університет харчових технологій

MatLab полегшує роботу з вибірками технологічних змінних за допомогою пакету Statistics Toolbox, який включає велику кількість допоміжних функцій. Ці функції призначені для визначення статистичних характеристик вибірок: дисперсії, математичного сподівання, середньоквадратичного відхилення, моди, медіани, кореляційних функцій тощо.

Для коректної роботи із статистичними вибірками необхідно проводити перевірку гіпотези нормальності розподілу. Закон нормального розподілу, так званий закон Гаусса, один з найпоширеніших законів – є типом безперервного розподілу. Головна особливість, яка виділяє нормальний закон серед інших законів, полягає в тому, що він є граничним законом, до якого наближаються інші закони розподілу в певних (типових) умовах. Будь-яка випадкова величина, яка є сумою великої кількості окремих числових значень, кожне з яких підпорядковується різним законам розподілу і несуттєво впливає на суму, розподілена майже за нормальним законом. Значення, які відхиляються від нормального розподілу за допомогою спеціальних перетворень, можуть бути наближені до нормального.

Відсутність перевірки на нормальність розподілу вихідних даних ставить під сумнів надійність підсумкових результатів розрахунку вартості та коректність отриманих оцінок точності.

На першому етапі роботи з MatLab було проведене перенесення даних з програми Excel в MatLab. Використано два найпоширеніших варіанти транспонування даних.

Варіант №1: За допомогою функції *putmatrix*.

Особливістю є обов'язкове попереднє налаштувати Excel на спільну роботу з MatLab.

Варіант №2: За допомогою функції MatLab *Import data*

В тезі наведено результати перевірки гіпотези нормальності розподілу вибірки по вмісту сухих речовин в сиропі після випарної станції цукрового заводу за допомогою наступних функцій MatLab:

- **HISTFIT** – Гістограма з накладеним графіком функції щільності розподілу ймовірностей нормального закону

Результати виконання команди для досліджуваної вибірки свідчить, що розподіл генеральної сукупності суперечить нормальному закону ($H=1$). Підтвердження наведено на рис. 1:

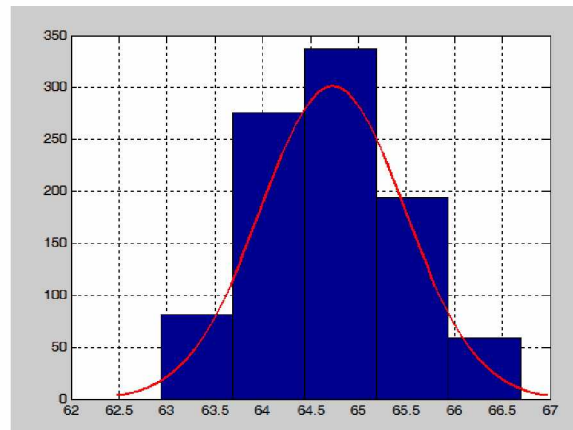


Рис. 1. Перевірка гіпотези на нормальність досліджуваного розподілу за допомогою функції histfit

- **JBTEST** – тест Яркі-Бера на непротириччя розподілу генеральної сукупності значень випадкової величини нормального закону.

За критерієм Яркі-Бера також відкидаємо нульову гіпотезу ($H=1$) про нормальність розподілу на рівні значущості $p < 0,05$ (рис. 2):

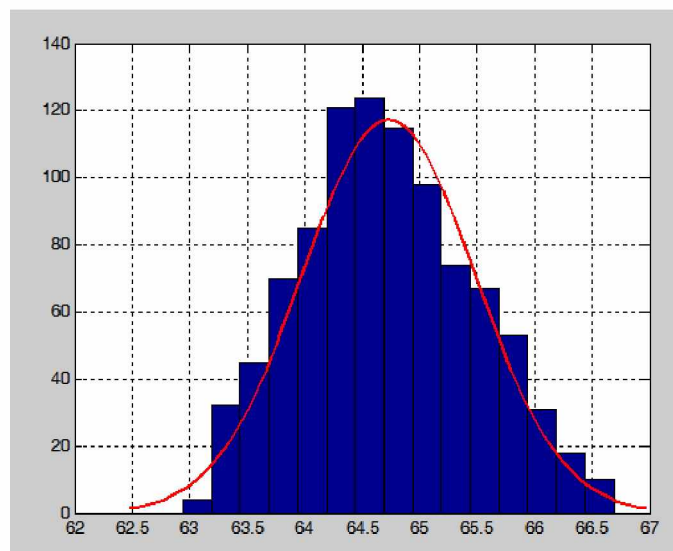


Рис. 2. Перевірка гіпотези на нормальність досліджуваного розподілу за допомогою теста Яркі-Бера

- **LILLIETEST** – тест Ліллієфорса на непротириччя розподілу генеральної сукупності значень випадкової величини нормального закону

За критерієм Яркі-Бера також відкидаємо нульову гіпотезу ($H=1$) про нормальність розподілу на рівні значущості $p < 0,05$.

Виконання статистичних задач стає набагато простіше завдяки пакету Statistics Toolbox в програмі MatLab. Побудова гістограм дає наглядне уявлення про зміни вибірки під час перевірки гіпотези на нормальність розподілу за допомогою різних тестів і функцій пакету Statistics Toolbox.

Проаналізувавши результати проведених тестів можна зробити висновок, що закон розподілу даної вибірки не є нормальним і потребує проведення фільтрації та згладжування для подальшої роботи з ним.