Application of statistical tools and hypothesis testing of adsorption data obtained for removal of heavy metals from aqueous solutions

Achla Kaushal[1,*], Prof.S.K.Singh[2]
[1]Faculty in Chemical Engineering Department, Directorate of Training and Technical Education, Delhi. India
[2]Department of Environmental Engineering, Delhi Technological University, Delhi, India.

Article Info
Article history:
Received January 2016
Received in revised form
20 February 2016
Accepted March 2016
Available online June 2016

Keywords
Adsorption, experimental data, statistical analysis, hypothesis testing,

Abstract

Hypothesis testing is useful statistical tool understood as a critical testing of the experimental data. The experiment for the removal of heavy metal ions from aqueous solution by adsorption was carried out with the deliberate object of testing hypothesis through different tests meant for the purpose using available data obtained from experiments. Decisions about the validity and significance of the data were made on the basis of hypothesis testing. It enabled us to make probability statements about the adsorption parameters. In the present study, dried mango leaf powder was used to remove Cu, Zn and Pb from heavy metal samples prepared in lab. the experimental data was tested using different statistical tools namely Student's t-tests, F-test to test the equality of variance of two normal populations, Chi-square test to test the effectiveness of the adsorption in removing heavy metal ions, ANOVA for the analysis of significance of difference between means of multiple samples at the same time, all within 5% level of significance.

1. Introduction

Adsorption is a known phenomenon in water treatment since ancient times. It is a common phenomenon in gaseous phase, but is used effectively in treatment of waste water. Literature has shown the use of charcoal prepared from wood char for reducing ores of Cu, Zn and Sn to manufacture bronze by Egyptians and Sumerians during 3750 BC. 450 BC onwards, purification of drinking water drinking came into practice using sand and charcoal filters. [1]. During several decades, concentrations of natural organic materials (NOMs), synthetic organic chemicals (SOCs) and trace heavy metals have increased in all water bodies and drinking water supplies to harmful levels. Lead, mercury, nickel, chromium, thallium, zinc, copper, cobalt, selenium, arsenic and cadmium are the heavy metals known for their toxicity and persistency, are mostly non-biodegradable, unlike organic pollutants. They find their way into water bodies through industrial, domestic and municipal discharges. Their presence in water is a serious threats to all forms of life, because they are mutagenic and carcinogenic in nature. Their presence above prescribed limits in body can cause severe damages to vital organs of the body, such as kidney, liver and brain, reproductive and nervous system [3].

Of all the methods of waste water treatment, adsorption with low cost agricultural waste is one of the trusted methods in heavy metal removal for its cost effectiveness, availability of wide range of adsorbents and reliable metal removal efficiency. A large number of adsorbents prepared from agricultural wastes have been exploited by the researchers for this purpose, such as sawdust, tea waste, palm shell, coconut shell, bamboo activated charcoal, olive cake etc.. Experimental data for the adsorption process obtained in terms of %age removal of heavy metal ions w.r.t. pH, time, temperature, adsorbent dose and metal ion concentration in the sample is analysed by fitting them into non-linear forms such as Langmuir and Freundlich isotherm models[3].

Corresponding Author,
***E-mail address:** achla001@hotmail.com
+919811200429

Efforts are put to describe the process qualitatively and quantitatively statistically. The hypothesis testing is based on the null hypothesis (H_o), keeping an alternative hypothesis (H_a). It is assumed that the null hypothesis is true, probabilities are assigned to different possible sample results, but this is possible only if we proceed with an alternative hypothesis. Level of significance is normally determined before testing the hypothesis. It is the maximum value of the probability of rejecting null hypothesis. Our sample results decide if the null hypothesis can be accepted or rejected. If null hypothesis is rejected, alternative hypothesis is accepted.

Various hypothesis tests known as tests of significance developed by statisticians are classified as parametric tests, the standard tests and non-parametric tests, the distribution-free tests of hypotheses.

The important parametric tests, based on the assumption that the data is distributed normally, are z-test, t-test, F-test and χ^2-test.

Z-test is used for judging the significance of significance of mean, median and mode etc by comparing the sample mean with some hypothesized value of the population mean when sample is large (n>30). For a normal and infinite population and known population variance,

$z = (\bar{X} - \mu_{H_o}) / (\sigma_p/\sqrt{n})$, H_a may be one sided or two sided.

T-test, based on distribution is the appropriate test to judge the significance of differences between the two sample means for small samples when the population variance is unknown. For a normal and infinite population, small sample size and unknown population variance,

$t = (\overline{X} - \mu_{Ho})/(\sigma_s/\sqrt{n})$ for (n-1) degree of freedoms and H_a may be one sided or two sided.

$\sigma_s = \sqrt{[\{\Sigma(X_i - \overline{X})^2\}/(n-1)]}$

F-test based on the F-distribution is used to compare the variance of two samples. It is also used in analysis of variance (ANOVA) to judge the significance of means of multiple samples at the same time.

$F = \sigma_{s1}^2 / \sigma_{s2}^2$

$\sigma_{s1}^2 = \{\Sigma(X_{1i} - \overline{X}_1)^2\}/(n_1-1)$

$\sigma_{s2}^2 = \{\Sigma(X_{2i} - \overline{X}_2)^2\}/(n_2-1)$

While performing calculation, F is always taken > 1.

χ^2-test is a technique used to test the goodness of fit, independence and significance of population variance.

$\chi^2 = \sum\{(O_{ij} - E_{ij})^2/E_{ij}\}$

O_{ij}= Observed frequency of the cell in the i^{th} row and j^{th} column.

E_{ij}= Expected frequency of the cell in the i^{th} row and j^{th} column. [4]

2. Materials and Method

Analytical grade Zn metal chips were dissolved in a few drops of conc. H_2SO_4 and then diluted with distilled water to prepare the stock solution. Stock solution was further diluted to concentrations in the range 10-100ppm with the help of distilled water. Mango tree leaves collected from the University campus were washed, dried and crushed to 60 mesh size. They were thoroughly washed to remove all the colour, bone dried in a hot air oven and then stored in air tight container. It was used to determine the % removal of Zn ions from the solutions prepared as mentioned above in the range 1-10 g/L. 12 Erlenmeyer flasks were used at a time in the 'Orbital' constant temperature variable speed shaker at 22°C and 150 rpm for 100 minutes to carry out batch experiment. The concentration of heavy metal was determined by Atomic Absorption Spectrophotometer using air-acetylene flame. Percentage removal of zinc from the solution after the batch adsorption was calculated as

% Removal = $[(C_i - C_o)/C_i]*100$

Where C_o represents the final zinc ion concentration and C_i represents the initial zinc ion concentration in the solution.

The equilibrium adsorption capacity is calculated as

$q_e = [(C_i - C_e)V]/m$

Where V is the volume of the solution in liters and m is the mass of the adsorbent used [5].

3. Results and Discussion

Adsorption of zinc metal ions was carried out on mango leaves powder batch wise. The experiment was conducted for zinc ion metal concentration in the range 10-100 ppm, adsorbent dose 1-10 g/L for the pH between 2-8 at 20°C and at 150 rpm. After each batch experiment, the solution was filtered with Whitman filter paper no. 40 and the concentration of zinc in the filtrate was examined with the help of AAS. Data collected from samples was used for hypotheses testing 1) to judge the maximum lead ion removal from solution 2) to judge the time required to attain equilibrium 3) To test the equality of variance of two n-populations 4) To test the effectiveness of the experiment 5) to judge the significance of means of multiple sample at the same time.

3.1. **Hypothesis testing to judge the maximum removal of the zinc ions from the samples** containing 100 ppm was done with two-tailed t-test within 5% level of significance. Calculations in Table 1 showed the standard deviation was 21.74747, t-observed was -2.41959 and t-tabulated was -2.447 at 5% level of significance for 6 degrees of freedom (d.f.) for 2-tailed t-distribution.

Since $t_{observed} < t_{tabulated}$, null hypothesis was accepted.

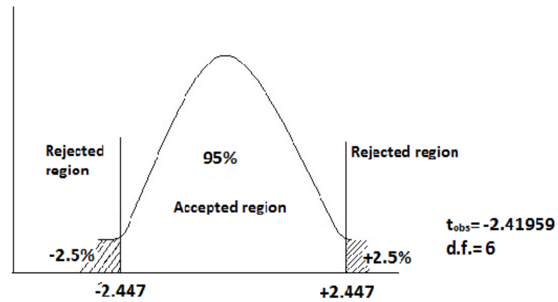

Fig. 1 t-Distribution for two–tailed test

2. **Hypothesis testing to judge the time required to attain equilibrium** was done with one-tailed t-test within 5% level of significance. Calculations (Table 2) showed the standard deviation was 19.467, $t_{observed}$ was -1.99227 and $t_{tabulated}$ was 2.015 at 5% level of significance for 6 degrees of freedom (d.f.) for 1-tailed t-distribution.

Since $t_{observed} < t_{tabulated}$, null hypothesis was accepted.

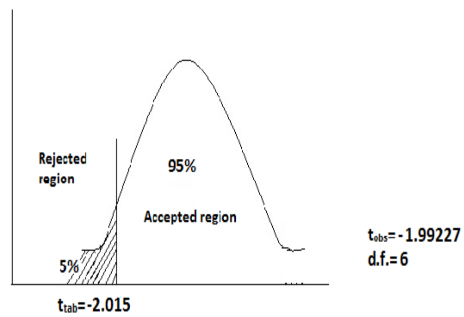

Fig 2. t- Distribution for one – tailed test

3. **Hypothesis testing to test the equality of variance of two n-populations,** samples collected for initial zinc ion concentrations 100ppm and 50 ppm, was done with F-test within 5% level of significance. It was assumed that the samples have been drawn randomly from two normal populations, observations are independent and there are no measurement errors.

Calculations in Table 3 showed that the variances for two samples were 520.1429 and 423.5714, $F_{calculated}$ was 1.227993 and $F_{tabulated}$ was 4.28 at 5% level of significance for 6 degrees of freedom (d.f.) for F-distribution. Since $F_{calculated} < F_{tabulated}$, null hypothesis was accepted.

4. **Hypothesis testing to test the effectiveness of the experiment** was done with χ^2-test within 5% level of significance. 60 experiment were conducted varying values of pH [2-8], metal ion concentration (10-100ppm) and adsorbent dose (1-20)g/L. Metal ion removal frequencies were observed for % removal between groups 0-20% [1], 20-40% [2], 40-60% [3], 60-80% [4] and 80-100% [5]. Since the frequencies in groups 1 and 5 were less than 10, they were combined with groups 2 and 4 respectively to

make the frequencies greater than 10. Expected frequencies for each group were equal and 12 each.
Calculations (Table 4) showed that$\chi^2_{calculated}$ was 3.583 and $\chi^2_{tabulated}$ was 9.488 at 5% level of significance for 4 degrees of freedom (d.f.). Since $\chi^2_{calculated} < \chi^2_{tabulated}$, null hypothesis was accepted.

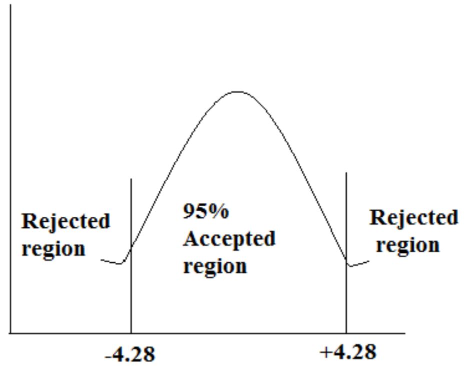

Fig. 3 F- Distribution for 100ppm and 50ppm samples

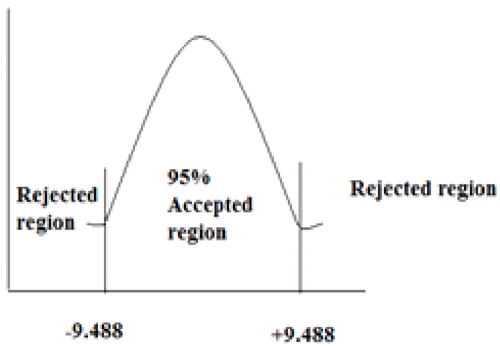

Fig. 4 Chi Square Distribution

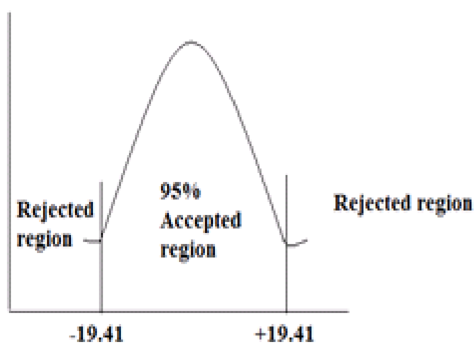

Fig. 4 ANOVA: F-Distribution

4. **Analysis of variance (ANOVA)** is a useful statistical tool used to test the differences in the means of multiple populations by studying the variations within each sample, since the application t-test and z-test is is limited to two samples only. We compared three populations i.e. adsorption capacities of mango tree leaves for three different heavy metal ions, Zn, Cu and Pb for varying values of adsorbent dose (table 5). F-ratio calculated was 0.018187 and tabulated was 19.41 for (2,12) degrees of freedom. Since F-ratio $_{calculated}$<F-ratio $_{tabulated}$, null hypothesis was accepted.

Conclusions

Hypotheses are the clear and concise statements, assumptions or formal questions for the available data, capable to be tested, which are meant to be accepted or rejected. A hypothesis testing validates these assumptions. In the above study, our assumptions, called the null hypothesis, framed for different situations a) to judge the maximum removal of Zn with by mango tree leaves b) to judge the equilibrium time c) to compare the variances of two different samples of zinc ion concentration d) to check the effectiveness of the experiment and e) to test the difference of means of three different samples at the same time were tested at 5% level of confidence each. All the calculated values were found to be within the accepted region of the probability chart and hence the null hypotheses were accepted.
Paired t-test to compare two related samples, statistical tests for correlation coefficients and adsorption isotherm are recommended for future studies.

References

[1.] Anderson T.W., 1958, An Introduction to Multivariate Analysis, New York: John Wiley & sons.
[2.] Bowley A.L., 1937, Elements of Statistics, 6th ed. London: P.S.King and Staples Ltd.
[3.] Cafer S, Omer S, Ahin M. M., 2012, Applications on agricultural and forest waste adsorbents for the removal of lead (II) from contaminated waters, int. J. Environ. Sci. Technol., 9: 379-394.
[4.] Chance W.A., 1975, Statistical Methods for Decision Making, Bombay: D.B>Taraporevala Sons & co. Pvt Ltd.
[5.] Dabrowski A, 2001, Adsorption—from theory to practice, Adv. Colloid Interface Sci.93, 135–224.
[6.] Ferhan C. and Ozgur A., 2011, Activated Carbon for Water and Wastewater treatment: Integration of Adsorption and Biological Treatment, First Edition, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
[7.] Goel P.K., 2006, Water pollution, Causes, Effects and Control, N. Delhi: New age international publishers, Revised second edition.
[8.] Kothari C.R., 2005, Research Methodology: Methods & Techniques, New Delhi: New Age International (P) Limited, Publishers.
[9.] Qaiser S, Saleemi A.R. and Ahmad M.M., 2007, Heavy metal uptake by agro based waste materials, Electronic j of Biotech, 10(3).