Quantitative Methods in Chemistry

Week 3, Lecture 2

This week we have already been introduced to the concepts of:

- ✓ Confidence Level
- ✓ Significance Level
- ✓ Confidence Interval
- ✓ z-statistics
- ✓ t-statistics

Remember, **Confidence Level = (1- α) x 100 %, where α is the significance level**

We will now learn about the following:

- Null Hypothesis and Alternate Hypothesis
- Using z- and t-statistics for hypothesis testing
- Two-tailed test
- One-tailed test
- Errors in Hypothesis testing
- Concept of Outliers
- Identification of Outliers using Q-tables

Hypothesis testing using z- or t-statistics

Null Hypothesis

Null hypothesis assumes that there is no statistical difference between the two measurements being compared. In other words, the minor differences between the observed values are presumed to be occurring due to random fluctuations.

We then employ statistics to test the veracity of this presumption. We set a critical level of significance (α) beyond which the null hypothesis is deemed to be questionable.

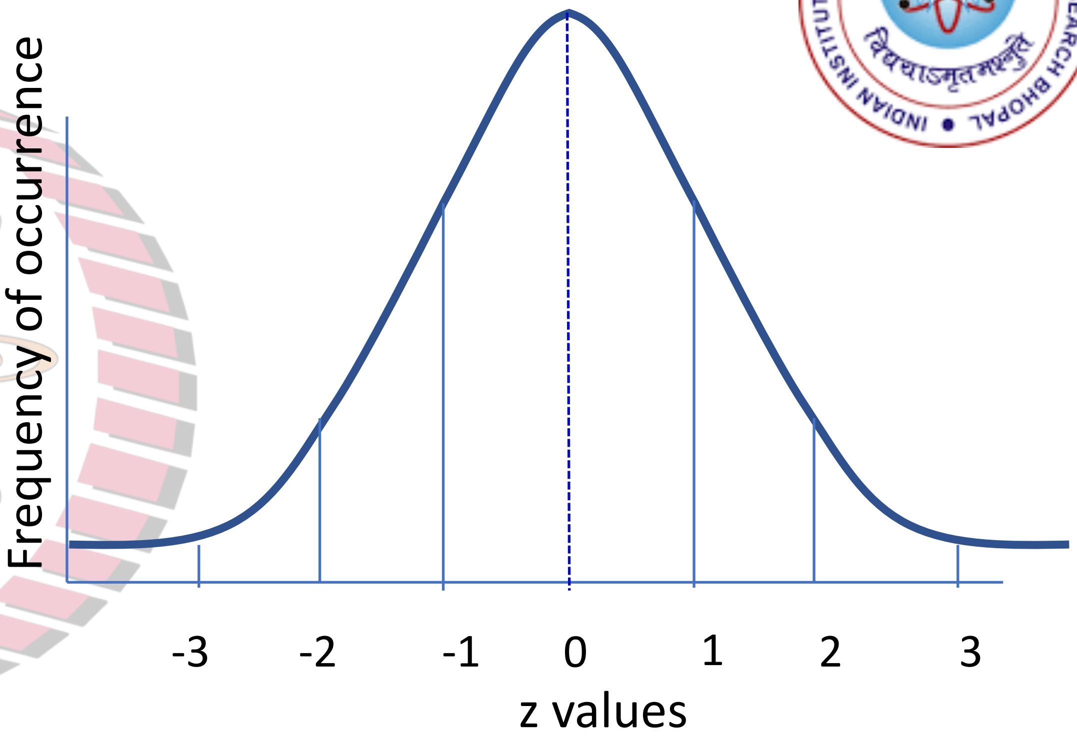
Two-tailed test

Example question to be answered:

1. In city X, the mean level of CO_2 is 1200 ± 100 ppm.
After the recent rains, this value was observed to be 1050 ppm.
At 95% Confidence Level, can we say that the observed CO_2 levels are indeed **different** after the rains?

2. Buffalo milk contains 6% solid not fat (SNF). A sample of milk contained 7 ± 0.2 % SNF. Is the SNF content of this sample **different** from the buffalo milk, at 95% confidence level?

(i.e. from the two tailed test we can only inform if the values are statistically different from the accepted or known values



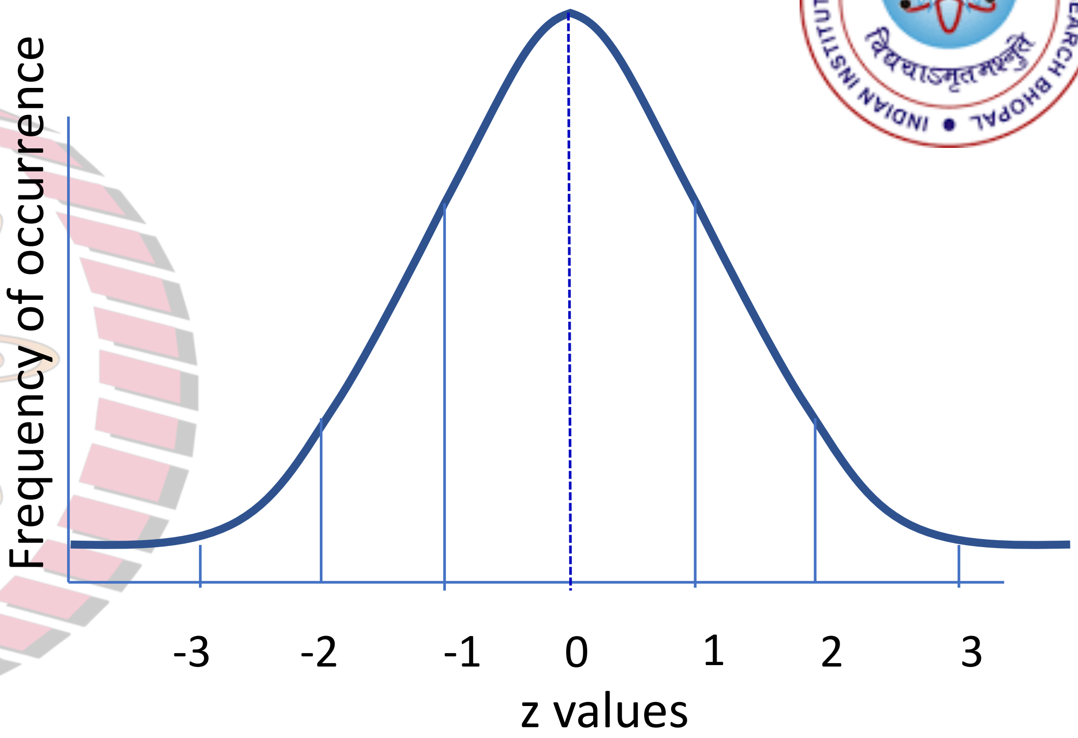
For two tailed test,
Confidence Level = 95%, so
Significance Level = (100 – 95) = 5% or 0.05
Equally distributed on both tails!
=> Z = 1.96

One-tailed test

Suppose we change the question to:

1. In city X, the mean level of CO_2 is 1200 ± 100 ppm. After the recent rains, this value was observed to be 1050 ppm. At 95% Confidence Level, can we say that the observed CO_2 levels are indeed **lower** after the rains?

2. Buffalo milk contains 6% solid not fat (SNF). A sample of milk contained 7.0 ± 0.2 % SNF. Is the SNF content of this sample **higher** from the buffalo milk, at 95% confidence level?



(i.e. from the one tailed test we can gauge if the test values are statistically **greater** or **smaller** than the accepted or known values.

For one tailed test,
Confidence Level = 95%
Significance Level = (100 – 95) = 5%
Only on one of the tails of the curve
=> Z = 1.64

t Distribution: Critical Values of *t*

Degrees of freedom		Significance level					
	Two-tailed test:	10%	5%	2%	1%	0.2%	0.1%
	One-tailed test:	5%	2.5%	1%	0.5%	0.1%	0.05%
1		6.314	12.706	31.821	63.657	318.309	636.619
2		2.920	4.303	6.965	9.925	22.327	31.599
3		2.353	3.182	4.541	5.841	10.215	12.924
4		2.132	2.776	3.747	4.604	7.173	8.610
5		2.015	2.571	3.365	4.032	5.893	6.869
6		1.943	2.447	3.143	3.707	5.208	5.959
7		1.894	2.365	2.998	3.499	4.785	5.408
8		1.860	2.306	2.896	3.355	4.501	5.041
9		1.833	2.262	2.821	3.250	4.297	4.781
10		1.812	2.228	2.764	3.169	4.144	4.587
11		1.796	2.201	2.718	3.106	4.025	4.437
12		1.782	2.179	2.681	3.055	3.930	4.318
13		1.771	2.160	2.650	3.012	3.852	4.221
14		1.761	2.145	2.624	2.977	3.787	4.140
15		1.753	2.131	2.602	2.947	3.733	4.073
16		1.746	2.120	2.583	2.921	3.686	4.015
17		1.740	2.110	2.567	2.898	3.646	3.965
18		1.734	2.101	2.552	2.878	3.610	3.922
19		1.729	2.093	2.539	2.861	3.579	3.883
20		1.725	2.086	2.528	2.845	3.552	3.850
21		1.721	2.080	2.518	2.831	3.527	3.819
22		1.717	2.074	2.508	2.819	3.505	3.792
23		1.714	2.069	2.500	2.807	3.485	3.768
24		1.711	2.064	2.492	2.797	3.467	3.745
25		1.708	2.060	2.485	2.787	3.450	3.725
26		1.706	2.056	2.479	2.779	3.435	3.707
27		1.703	2.052	2.473	2.771	3.421	3.690
28		1.701	2.048	2.467	2.763	3.408	3.674
∞		1.645	1.960	2.326	2.576	3.090	3.291

Using **z-test** for hypothesis testing:

Can be applied for large samples for which sample standard deviation (*s*) is a good estimate of the population standard deviation (σ).

Protocol:

1. State the null hypothesis. Let us call it H_0. It presumes that the test mean and population mean are statistically similar. i.e. $H_0 : x = \mu_o$. The veracity of this presumption is to be tested now.

2. Calculate the z-value: z_{calc} =

3. State the alternate hypothesis, H_a and determine the rejection region:

For $H_a : x \neq \mu_o$, reject H_0 if $z_{calc} \geq z_{crit}$ or, $z_{calc} \leq z_{crit}$ (apply a two-tailed test here!)

For $H_a : x > \mu_o$, reject H_0 if $z_{calc} \geq z_{crit}$ (apply the one-tailed test, use only the **right** tail)

For $H_a : x < \mu_o$, reject H_0 if $z_{calc} \leq z_{crit}$ (apply the one-tailed test, use only the **left** tail)

Using **t-test** for hypothesis testing:

To be applied for small samples for which the population standard deviation (σ) is not known or for which the sample standard deviation (*s*) is not a good estimate of σ, i.e. $s \rightarrow \sigma$.
=> **When in doubt, apply t-test!**

Protocol:

1. State the null hypothesis. Let us call it H_0. It presumes that the sample mean and population mean are statistically similar. i.e. $H_0 : x = \mu_o$. The veracity of this presumption is to be tested now.

2. Calculate the t-value: t_{calc} =

3. State the alternate hypothesis, H_a and determine the rejection region:

For $H_a : x \neq \mu_o$, reject H_0 if $t_{calc} \geq t_{crit}$ or, $t_{calc} \leq t_{crit}$ (apply a two-tailed test here!)

For $H_a : x > \mu_o$, reject H_0 **only if** $t_{calc} \geq t_{crit}$ (apply the one-tailed test, use only the **right** tail)

For $H_a : x < \mu_o$, reject H_0 **only if** $t_{calc} \leq t_{crit}$ (apply the one-tailed test, use only the **left** tail)

Solved example:
Suppose weight of 30 students of a particular age-group was taken. The mean value of their weight was 27.7 kg with a standard deviation value of 5.2 kg. If the national average of weight for this age-group is 30.8 kg, then, can we say that:

(1) the observed mean is **different** from the national mean at **95% confidence level**?

(2) the observed mean is **different** from the national mean at **99% confidence level**?

(a) Presume $s \rightarrow \sigma$.
(b) Do not make the above presumption.

Apply two-tailed test here

In the previous case,
At 99.9% Confidence Level, and not presuming $s \rightarrow \sigma$,
Can we say that the observed mean is indeed **less than** the national average?

apply the one-tailed test, use only the **left** tail

Errors in Hypothesis testing

Type I error (The "False negative" error) – **Rejecting null hypothesis** when it should not be rejected (At 95% Confidence level, there is still 5% chance that the rejection of null hypothesis is erroneous. In other words, the <u>significance level indicates the probability of committing Type I error as well</u>.)

Type II error (The "False Positive" error – Accepting null hypothesis when it should be rejected. Decreasing the Type I error probability (by making a smaller) increases the chances of committing Type II error.

A judicial analogy to Hypothesis testing:

An accused is brought to the court. The court pronounces the verdict after following the due procedure.

1. Null Hypothesis – The accused is **presumed not guilty** unless proven otherwise.
2. Testing the veracity of null hypothesis – Based on the evidence presented.
3. Make an inference based on the evidence presented – Pronounce a **verdict**.
4. Now, there are still finite chances of pronouncing an **innocent person guilty** (committing a Type I error) or pronouncing a **guilty person as being innocent** (committing a Type II error)
5. In this example, a Type I error is considered more serious than Type II error!

The concept of outliers, and how to test them

Suppose we have a set of data and we have a reading which we feel is out of range of the expected values, or is very away from the other readings. Such readings that lie "outside our comfort zone" can be considered as outliers.

- What should we do with such reading/ measurement that seem to be outliers for us?
- Is there a **scientific way to assess** if a reading is to be really considered as an outlier or not?
- What criterion can be applied to **discard or keep** a reading considered an outlier?

Rejection quotient or Q-test for deciding outliers

$$Q_{calc} = \frac{|x_q - x_n|}{w}$$

Where, x_q = the questionable reading
x_n = the reading nearest to the questionable reading
w = the spread of the data (largest value – smallest value)

Table 1. Tabulated values for the Q-test.

n	68%	90%	95%	98%	99%
3	0.822	0.941	0.970	0.988	0.994
4	0.603	0.765	0.829	0.889	0.926
5	0.488	0.642	0.710	0.780	0.821
6	0.421	0.560	0.625	0.698	0.740
7	0.375	0.507	0.568	0.637	0.680
8	0.343	0.468	0.526	0.590	0.634
9	0.319	0.437	0.493	0.555	0.598
10	0.299	0.412	0.466	0.527	0.568
12	0.271	0.375	0.425	0.480	0.518
14	0.250	0.350	0.397	0.447	0.483
16	0.234	0.329	0.376	0.422	0.460
18	0.223	0.314	0.358	0.408	0.438
20	0.213	0.300	0.343	0.392	0.420

Reject the reading if $Q_{calc} > Q_{crit}$ **at that CL**.

Solved example for rejecting/keeping the outlier

Suppose the length of tiles in a box of 5 tiles is measured, and the values were obtained: 85.15, 84.98, 84.67, 84.55 and 84.75 cm.

Apply Q test on this data to find out if there is an outlier that can be rejected in this data set.

Table 1. Tabulated values for the Q-test.

n	68%	90%	95%	98%	99%
3	0.822	0.941	0.970	0.988	0.994
4	0.603	0.765	0.829	0.889	0.926
5	0.488	0.642	0.710	0.780	0.821
6	0.421	0.560	0.625	0.698	0.740
7	0.375	0.507	0.568	0.637	0.680
8	0.343	0.468	0.526	0.590	0.634
9	0.319	0.437	0.493	0.555	0.598
10	0.299	0.412	0.466	0.527	0.568
12	0.271	0.375	0.425	0.480	0.518
14	0.250	0.350	0.397	0.447	0.483
16	0.234	0.329	0.376	0.422	0.460
18	0.223	0.314	0.358	0.408	0.438
20	0.213	0.300	0.343	0.392	0.420