THE CENTRAL LIMIT THEOREM & CONFIDENCE INTERVALS

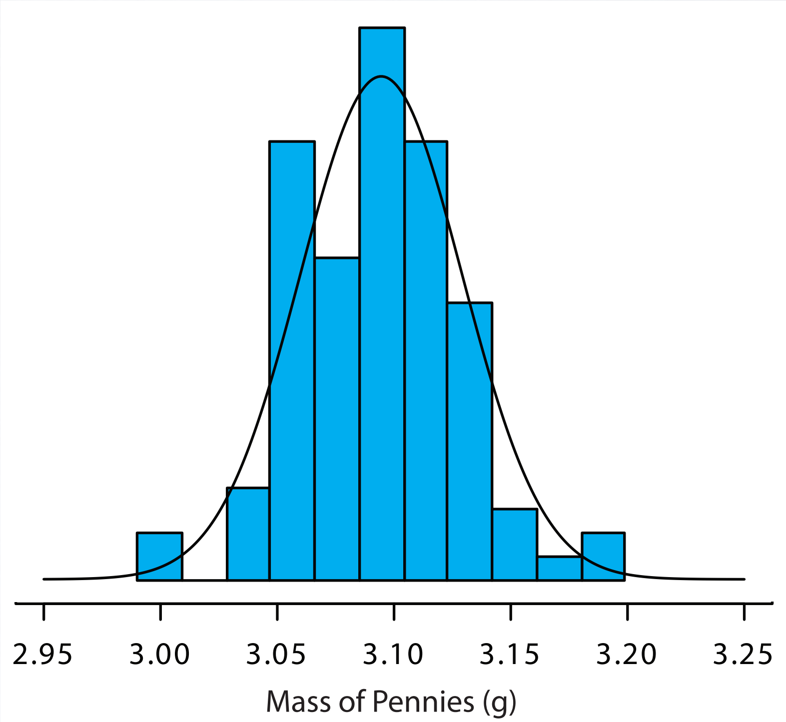
CHEM 251 SDSU

NORMAL DISTRIBUTION

- When we do measurements of a **sample** from the population we are unlikely to do enough measurements to fully represent the **population**.

- However, if we make enough measurements, are resulting **mean** and **variance** can approximate the normal distribution of the **population**.

Histogram of data with approximated normal distribution.



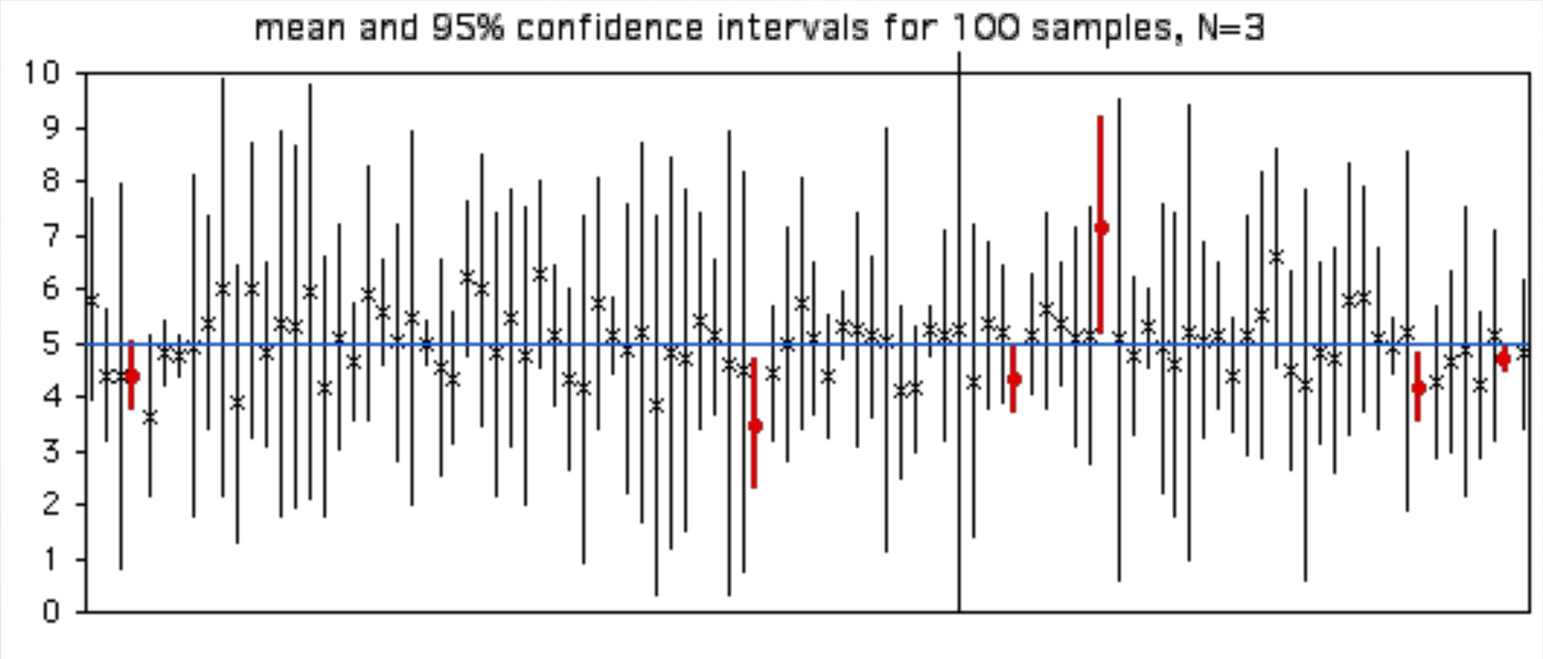
THE CENTRAL LIMIT THEOREM

- The reason that the approximation of the normal distribution works is due to the **Central Limit Theorem.**
- The **Central Limit Theorem** states that when a system is subject to a variety of indeterminate errors, the results of multiple measurements approximate a **normal distribution**.
- As such **samples** can reflect, with some degree of confidence, attributes of the **population**, such as the **mean** and **variance**.

CONFIDENCE INTERVALS

- As the sample mean does not truly represent the population mean, we can use confidence intervals to indicate the likely range where the true mean might lie.

- Confidence intervals can be determined with different levels of certainty (e.g. 95%, 90%, 50%,…)



CALCULATING CONFIDENCE INTERVALS

- The determination of the confidence interval can be done with a few key pieces of data from the sample:
- The sample mean ($\overline{x}$)
- The sample standard deviation (*s*)
- The number of measurements (n)
- A *t* value, based on the degrees of freedom (*n*-*1*) and the desired level of certainty (90%, 95%,…)

$$\mu = \overline{X} \pm \frac{ts}{\sqrt{n}}$$

$$s = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}}$$

Table 4.15 Values of *t* for a 95% Confidence Interval

Degrees of Freedom	*t*	Degrees of Freedom	*t*
1	12.706	12	2.179
2	4.303	14	2.145
3	3.181	16	2.120
4	2.776	18	2.101
5	2.571	20	2.086
6	2.447	30	2.042
7	2.365	40	2.021
8	2.306	60	2.000
9	2.262	100	1.984
10	2.228	∞	1.960

SAMPLE CALCULATIONS

Two students (A & B) have made measurements of samples taken from the same population.

Determine the 95% confidence interval for each of their sample means.

Table 4.15 Values of *t* for a 95% Confidence Interval

Degrees of Freedom	*t*	Degrees of Freedom	*t*
1	12.706	12	2.179
2	4.303	14	2.145
3	3.181	16	2.120
4	2.776	18	2.101
5	2.571	20	2.086
6	2.447	30	2.042
7	2.365	40	2.021
8	2.306	60	2.000
9	2.262	100	1.984
10	2.228	∞	1.960

Trails	Student A	Student B
1	14.602	14.408
2	14.782	14.517
3	14.668	14.322
4	14.534	14.477
5	14.721	14.398
6	14.596	
Average	14.6505	14.4244
Std. Dev.	0.091	0.075