- **Example:** Measure the sulfate concentration at a local reservoir over 9 consecutive days. Obtain a sample mean of $M_9 = 6.1$ mg/L. The variance is known to be $\sigma^2 = 0.81\ (\text{mg/L})^2$.

→ Is the concentration significantly different from the baseline concentration $\mu = 5.4$ mg/L at a significance level of 0.01?

One dataset with known variance ⇒ One-Sample Z-Test

→ Z-statistic: $Z = \dfrac{\sqrt{n}\,(M_n - \mu)}{\sigma} = \dfrac{\sqrt{9}\,(6.1 - 5.4)}{\sqrt{0.81}} = 2.33$

→ p-value $= 2\,\Phi(-|z|) = 2 \cdot \underbrace{\Phi(-2.33)}_{\text{MATLAB normcdf}(-2.33)} = 0.020$

→ Since p-value $\geq \alpha = 0.01$, fail to reject the null.

"concentration not elevated"

- <u>Example</u>: Measure the sulfate concentration at a local reservoir over 9 consecutive days. Obtain a sample mean of $M_9 = 6.1$ mg/L and a sample variance of $V_9 = 0.36\ (\text{mg/L})^2$.

→ Is the concentration significantly different from the baseline concentration $\mu = 5.4$ mg/L at a significance level of 0.05?

One dataset with unknown variance ⇒ One-Sample T-Test

→ T-statistic : $T = \dfrac{\sqrt{n}\,(M_n - \mu)}{\sqrt{V_n}} = \dfrac{\sqrt{9}\,(6.1 - 5.4)}{\sqrt{0.36}} = 3.50$

→ p-value $= 2\,F_{T_{n-1}}(-|T|) = \underbrace{2\,F_{T_8}(-3.50)}_{\text{MATLAB tcdf}(-3.50,\ 8)} = 0.008$

→ Since p-value $< \alpha = 0.05$, reject the null.

"concentration elevated"

- **Example:** Test a new blood pressure drug with a control group of 27 patients and an experimental group with 25 patients. Sample mean in the control group is $M^{(1)}_{27} = 2.10$ and sample mean in the experimental group is $M^{(2)}_{25} = 2.02$. From prior studies, variance in the control group and experimental group are known to be $\sigma_1^2 = 0.01$ and $\sigma_2^2 = 0.02$, respectively.

→ Does the drug lower cholesterol at a signficance level of 0.05?

Two datasets with known variance ⇒ Two-Sample Z-Test

→ Z-Statistic: $Z = \dfrac{M^{(1)}_{n_1} - M^{(2)}_{n_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{2.10 - 2.02}{\sqrt{\dfrac{0.01}{27} + \dfrac{0.02}{25}}} = 2.34$

→ p-value $= 2\,\Phi(-|Z|) = 2\,\underbrace{\Phi(-2.34)}_{\text{MATLAB normcdf(-2.34)}} = 0.02$

→ Since p-value $< \alpha = 0.05$, reject the null.

"drug lowers cholesterol"

- **Example:** Test a new blood pressure drug with a control group of 27 patients and an experimental group with 25 patients. Sample mean in the control group is $M_{27}^{(1)} = 2.10$ and sample mean in the experimental group is $M_{25}^{(2)} = 2.02$. Sample variances in the control group and experimental group are $V_{27}^{(1)} = 0.31$ and $V_{25}^{(2)} = 0.28$, respectively. Variance is thought to be equal.

→ Does the drug lower cholesterol at a significance level of 0.05?

Two datasets with unknown, equal variance ⇒ Two-Sample T-Test

→ Pooled Sample Variance: $\hat{\sigma}^2 = \dfrac{(n_1 - 1)V_{n_1}^{(1)} + (n_2 - 1)V_{n_2}^{(2)}}{n_1 + n_2 - 2} = \dfrac{26 \cdot 0.31 + 24 \cdot 0.28}{27 + 25 - 2}$

$= 0.30$

→ T-Statistic: $T = \dfrac{M_{n_1}^{(1)} - M_{n_2}^{(2)}}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \dfrac{2.10 - 2.02}{\sqrt{0.30\left(\frac{1}{27} + \frac{1}{25}\right)}} = 0.53$

→ p-value $= 2F_{T_{n_1+n_2-2}}(-|T|) = \underbrace{2F_{T_{50}}(-0.53)}_{\text{MATLAB tcdf(-0.53, 50)}} = 0.60$

→ Since p-value $\geq \alpha = 0.05$, fail to reject the null.

"drug does not lower cholesterol"