Compare Two Population Means $\mu_1 - \mu_2$ Based on Two Independent Random Samples

Setting: Suppose that $x_{11}, x_{12}, \ldots, x_{1n_1}$ are randomly selected from the first population, which has mean μ_1 and variance σ_1^2, and $x_{21}, x_{22}, \ldots, x_{2n_2}$ are randomly selected from the second population, which has mean μ_2 and variance σ_2^2.

The first sample yields a sample average $\bar{x}_1$ and a sample variance s_1^2 and the second sample yields a sample average $\bar{x}_2$ and a sample variance s_2^2.

	Sample 1	Sample 2
Size	n_1	n_2
Average	$\bar{x}_1$	$\bar{x}_2$
Variance	s_1^2	s_2^2

Objective: compare μ_1 with μ_2 or, equivalently, infer about the difference between μ_1 and μ_2.

Statistical Assumptions:

1. Two independent random samples are available.
2. (normality) – Both populations are normally distributed.
3. (homogeneity) – Two populations have the same variances, *i.e.*, $\sigma_1^2 = \sigma_2^2 \equiv \sigma^2$.

In summary, two independent random samples are taken from $\mathcal{N}(\mu_1, \sigma^2)$ and $\mathcal{N}(\mu_2, \sigma^2)$, respectively.

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$:
A natural point estimate of $\mu_1 - \mu_2$ is $\bar{x}_1 - \bar{x}_2$. Under the above statistical assumptions,

$$\bar{x}_1 - \bar{x}_2 \sim \mathcal{N}(\mu_1 - \mu_2, \quad \sigma^2 \cdot (\frac{1}{n_1} + \frac{1}{n_2}),$$

where the common variance σ^2 can be estimated by the *pooled* estimator given by

$$s_p^2 = \frac{\sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1) + \sum_{i=1}^{n_2}(x_{2i} - \bar{x}_2)}{(n_1 - 1) + (n_2 - 1)} = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$$

An $(1-\alpha)$ **Confidence Interval for** $\mu_1 - \mu_2$:

$$(\bar{x}_1 - \bar{x}_2) \ \pm \ t_{\frac{\alpha}{2}}^{(n_1+n_2-2)} \sqrt{s_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)},$$

Hypothesis Testing about $\mu_1 - \mu_2$:

1. Hypotheses

	Two-Sided	Upper-Tailed	Lower-Tailed
Null	$\text{H}_0 : \ \mu_1 - \mu_2 = D_0$	$\text{H}_0 : \ \mu_1 - \mu_2 \leq D_0$	$\text{H}_0 : \ \mu_1 - \mu_2 \geq D_0$
Alternative	$\text{H}_0 : \ \mu_1 - \mu_2 \neq D_0$	$\text{H}_0 : \ \mu_1 - \mu_2 > D_0$	$\text{H}_0 : \ \mu_1 - \mu_2 < D_0$

2. Test Statistic

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

Null Distribution: under H_0, $T \sim t^{(n_1+n_2-2)}$.

3. Decision Rule

One may make a decision using either the rejection region (for a given significance level α) or the p-value.

	Two-Sided	Upper-Tailed	Lower-Tailed
Rejection Region	$\lvert T \rvert > t_{\alpha/2}^{(n_1+n_2-2)}$	$T > t_{\alpha}^{(n_1+n_2-2)}$	$T < -t_{\alpha}^{(n_1+n_2-2)}$
p-value	$2\,\text{P}(t^{(n_1+n_2-2)} > \lvert T \rvert)$	$\text{P}(t^{(n_1+n_2-2)} > T)$	$\text{P}(t^{(n_1+n_2-2)} < T)$

4. Make a Decision

Reject H_0 if the observed or computed T falls into the rejection region or the computed p-value is less than α.

Example
In a study of iron deficiency among infants, samples of infants following different feeding regimens were compared. One group contained breast-fed infants, while the children in another group were fed a standard baby formula without any iron supplements. Here are summary results on blood hemoglobin levels at 12 months of age.

	Formula	Breast-fed
size	19	23
average	12.4	13.3
s.d.	1.8	1.7

1. Set up a test to test that the mean hemoglobin level is higher among breast-fed babies at $\alpha = 0.05$.

   Solution:

$$\begin{aligned} \mathrm{H}_0 : &\quad \mu_1 - \mu_2 \geq 0 \\ \mathrm{H}_a : &\quad \mu_1 - \mu_2 < 0 \end{aligned}$$

   First compute the pooled estimate of the common variance σ^2:

$$s_p^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} = \frac{(19 - 1) \cdot 1.8^2 + (23 - 1) \cdot 1.7^2}{19 + 23 - 2} = 3.05$$

   Then the test statistic is

$$\begin{aligned} T &= \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{(12.4 - 13.3) - 0}{\sqrt{3.05 \left(\frac{1}{19} + \frac{1}{23}\right)}} \\ &= -1.61 \end{aligned}$$

   We reject H_0 if $T < -t_{\alpha}^{(n_1+n_2-2)} = -t_{.05}^{(19+23-2)} = -t_{.05}^{(40)} = -1.684$. So H_0 is rejected at $\alpha = 0.05$.

2. Find the p-value for the above test

   Solution:

   p-value $= P\left(t^{(n_1+n_2-2)} < T\right) = P(t^{(40)} < -1.61) = P(t^{(40)} > 1.61)$ by symmetry of t distribution. According to table VI, It can be found that

   $$.05 < p\text{-value} < .10$$

   since

   $$t_{.10}^{(40)} = 1.303 < T = 1.61 < 1.684 = t_{.05}^{(40)}.$$

3. State the statistical assumptions that are necessary to guarantee the validity of the above analysis.

4. Construct a 95% CI for the difference in the mean hemoglobin levels between the breast-feeding group and the formula-feeding group.

   Solution:

   $$\begin{aligned} &(\bar{x}_1 - \bar{x}_2) \ \pm \ t_{\frac{\alpha}{2}}^{(n_1+n_2-2)} \sqrt{s_p^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \\ &= (12.4 - 13.3) \ \pm \ t_{.025}^{(40)} \sqrt{3.05 \left(\frac{1}{19} + \frac{1}{23}\right)} \\ &= -.9 \ \pm \ 2.201 \cdot .56 = (-2.03, .23) \end{aligned}$$

5. Test if the mean hemoglobin levels between two groups is significantly different at $\alpha = 0.05$.

   (Hint: make use of the 95% CI in part d.)