The data in a two-sample problem are two independent SRSs, each drawn from a separate population.

Tests and confidence intervals for the difference between the means $\mu_1$ and $\mu_2$ of two Normal populations start from the difference $\bar{x}_1 - \bar{x}_2$ between the two sample means. Because of the central limit theorem, the resulting procedures are approximately correct for other population distributions when the sample sizes are large.

Draw independent SRSs of sizes $n_1$ and $n_2$ from two Normal populations with parameters $\mu_1, \sigma_1$ and $\mu_2, \sigma_2$. The two-sample t statistic is

$$ t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} $$

The statistic $t$ has approximately a t distribution.

Note: Usually $H_0$ is that $\mu_1 = \mu_2$, so $\mu_1 - \mu_2 = 0$.

Thus $\bar{x}_1 - \bar{x}_2$ is used.

There are two choices for the degrees of freedom of the two-sample $t$ statistic.
Option 1: software produces accurate probability values using degrees of freedom calculated from the data. Option 2: for conservative inference procedures, use degrees of freedom equal to the smaller of $n_1 - 1$ and $n_2 - 1$.

The confidence interval for $\mu_1 - \mu_2$ is

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The critical value $t^*$ from Option 1 gives a confidence level very close to the desired level C. Option 2 produces a margin of error at least as wide as is needed for the desired level C.

Significance tests for $H_0: \mu_1 = \mu_2$ are based on

$$ t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} $$

P-values calculated from Option 1 are very accurate. Option 2 $P$-values are always at least as large as the true $P$.

**Option 2** used on test $\implies$ $df = \min(n_1-1, n_2-1)$ and look up $p$-value in Table C, $t$-distribution critical values.