Sampling Distributions When the Population Variance Is Unknown

The t-Distribution

One-Sample Case

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal population with mean μ and unknown variance σ^2.
When σ is unknown, we use the sample standard deviation S.
Then,

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t-distribution with

$$df = n - 1.$$

That is,

$$\boxed{T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)}$$

Example

An electrical firm manufactures light bulbs whose lifetimes are approximately normally distributed with a mean of 800 hours.
A random sample of 9 bulbs has a standard deviation of 50 hours.
Find $P(\overline{X} < 769)$.

Solution

Since the population standard deviation is unknown, we use:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1).$$

Therefore,

$$\begin{aligned} P(\overline{X} < 769) &= P\left(T < \frac{769 - 800}{50/\sqrt{9}}\right) \\ &= P\left(T < \frac{-31}{50/3}\right) \\ &\approx P(T < -1.860). \end{aligned}$$

The degrees of freedom are

$$df = n - 1 = 9 - 1 = 8.$$

Solution

From the t-table, with $df = 8$,

$$P(T > 1.860) = 0.05.$$

Since the t-distribution is symmetric about zero,

$$P(T < -1.860) = P(T > 1.860) = 0.05.$$

Therefore,

$$\boxed{P(\overline{X} < 769) \approx 0.05}$$

Two-Sample Case: Unknown but Equal Variances

Let two independent random samples of sizes n_1 and n_2 be selected from two normal populations with means μ_1 and μ_2. Suppose the population variances are unknown but equal:

$$\sigma_1^2 = \sigma_2^2 = \sigma^2.$$

We estimate it using the pooled sample variance

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

Then,

$$\boxed{T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)}$$

with $df = n_1 + n_2 - 2$.

Example

A company compares the battery life of two types of rechargeable batteries. Independent random samples of 10 Type A batteries and 12 Type B batteries are tested.
Assume that battery life is normally distributed for both types and that the population variances are unknown but equal. The sample variances are

$$S_1^2 = 9 \qquad \text{and} \qquad S_2^2 = 4.$$

Assuming that the two types of batteries have the same mean battery life, find

$$P\left(\overline{X}_1 - \overline{X}_2 > 2.233\right),$$

where battery life is measured in hours.

Solution

The pooled variance:

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

$$S_p^2 = \frac{(10 - 1)(9) + (12 - 1)(4)}{10 + 12 - 2}$$

$$= \frac{81 + 44}{20}$$

$$= 6.25.$$

Therefore,

$$S_p = \sqrt{6.25} = 2.5.$$

Since the two population means are equal,

$$\mu_1 - \mu_2 = 0.$$

Solution

Using the t-distribution,

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

Thus,

$$P(\overline{X}_1 - \overline{X}_2 > 2.233) = P\left(T > \frac{2.233 - 0}{2.5\sqrt{\frac{1}{10} + \frac{1}{12}}}\right)$$

$$\approx P(T > 2.086).$$

The degrees of freedom are

$$df = n_1 + n_2 - 2 = 10 + 12 - 2 = 20.$$

From the t-table,

$$P(T > 2.086) = 0.025.$$