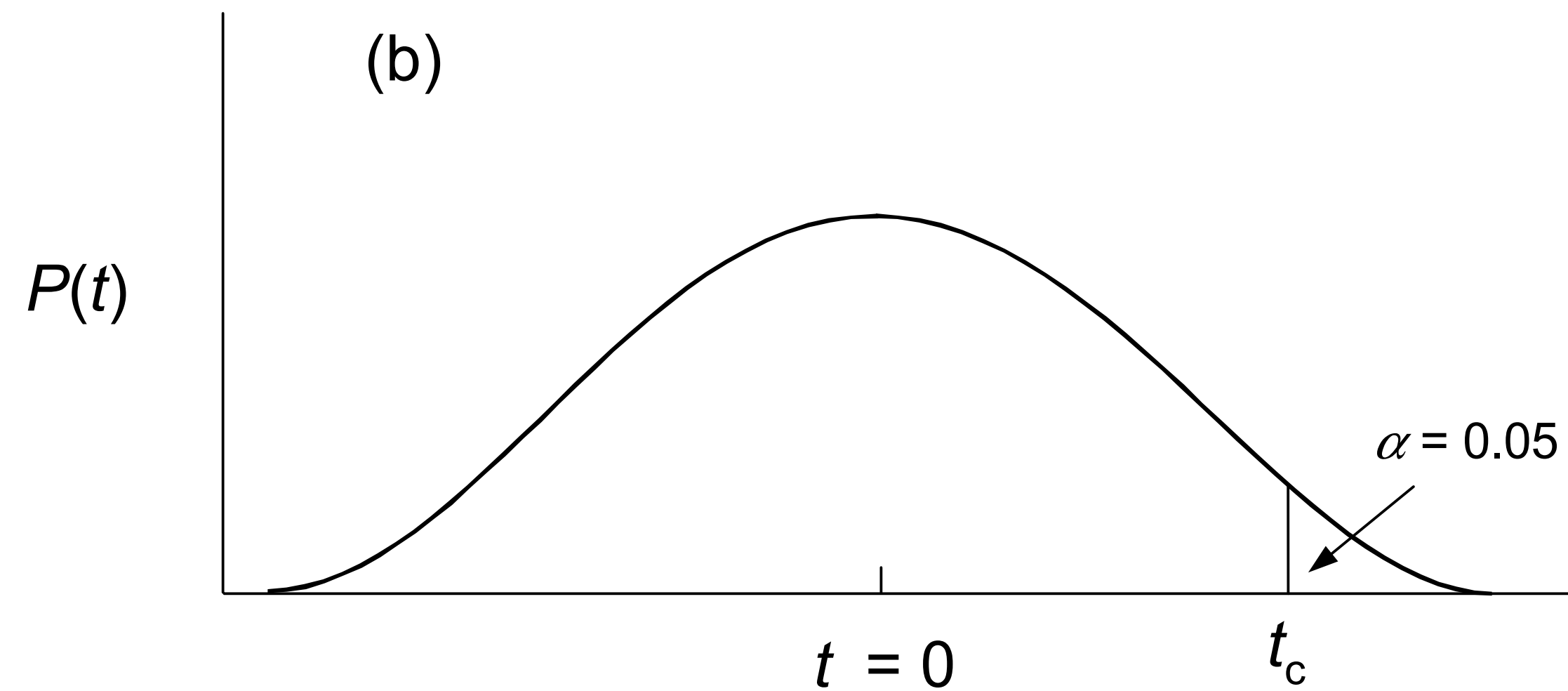
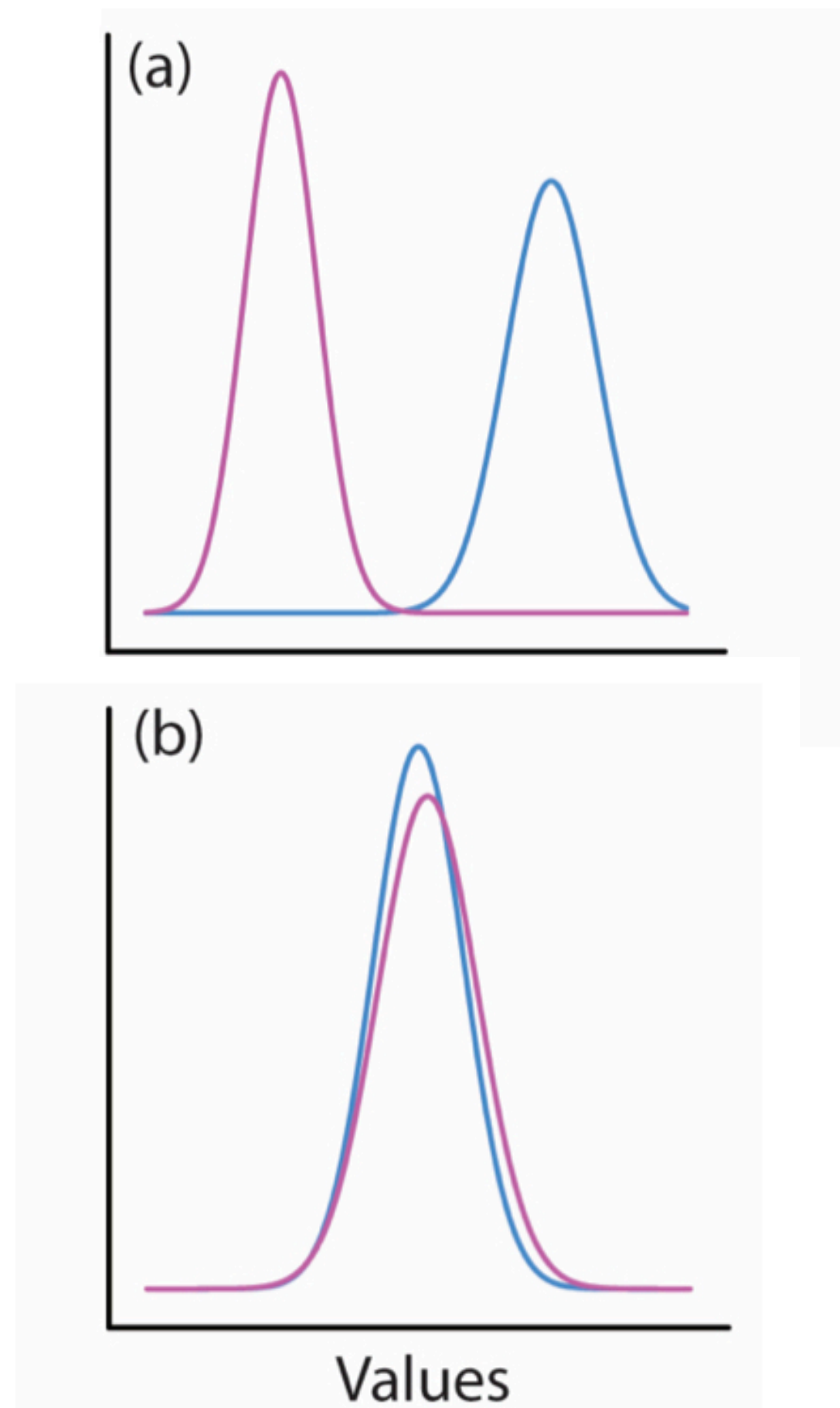


Foundational Statistics

Comparing two means: the t -test and nonparametric alternatives



What p -values don't tell us

- 1. Whether the observed effect is meaningful in the context of your study system.**
- 2. Whether your data meet the assumptions of the statistical test.**
- 3. The level of absolute plausibility of your hypothesis.**
(the results of your hypothesis tests have to be evaluated in the context of prior and similar work)

Comparing 2 means

Null Hypothesis (H_0): Douglas fir trees are **not taller**, on average, than Ponderosa pine trees.

Alternative Hypothesis (H_A): Douglass fir trees are **taller**, on average, than Ponderosa pine trees.

$$H_0 : \mu_1 \leq \mu_2$$

Doug fir



Ponderosa



diff. in means:
basis for test
statistic

$$\mu_1 - \mu_2 \text{ is } 0 \text{ or } < 0$$

$$H_A : \mu_1 > \mu_2$$



$$\mu_1 - \mu_2 \text{ is } > 0$$

The t -statistic and its sampling distribution

$$H_0 : \mu_1 \leq \mu_2$$

$$H_A : \mu_1 > \mu_2$$

Under $H_0 : t \not\gg 0$

Under $H_A : t \gg 0$

We can use the t -statistic to test these hypotheses

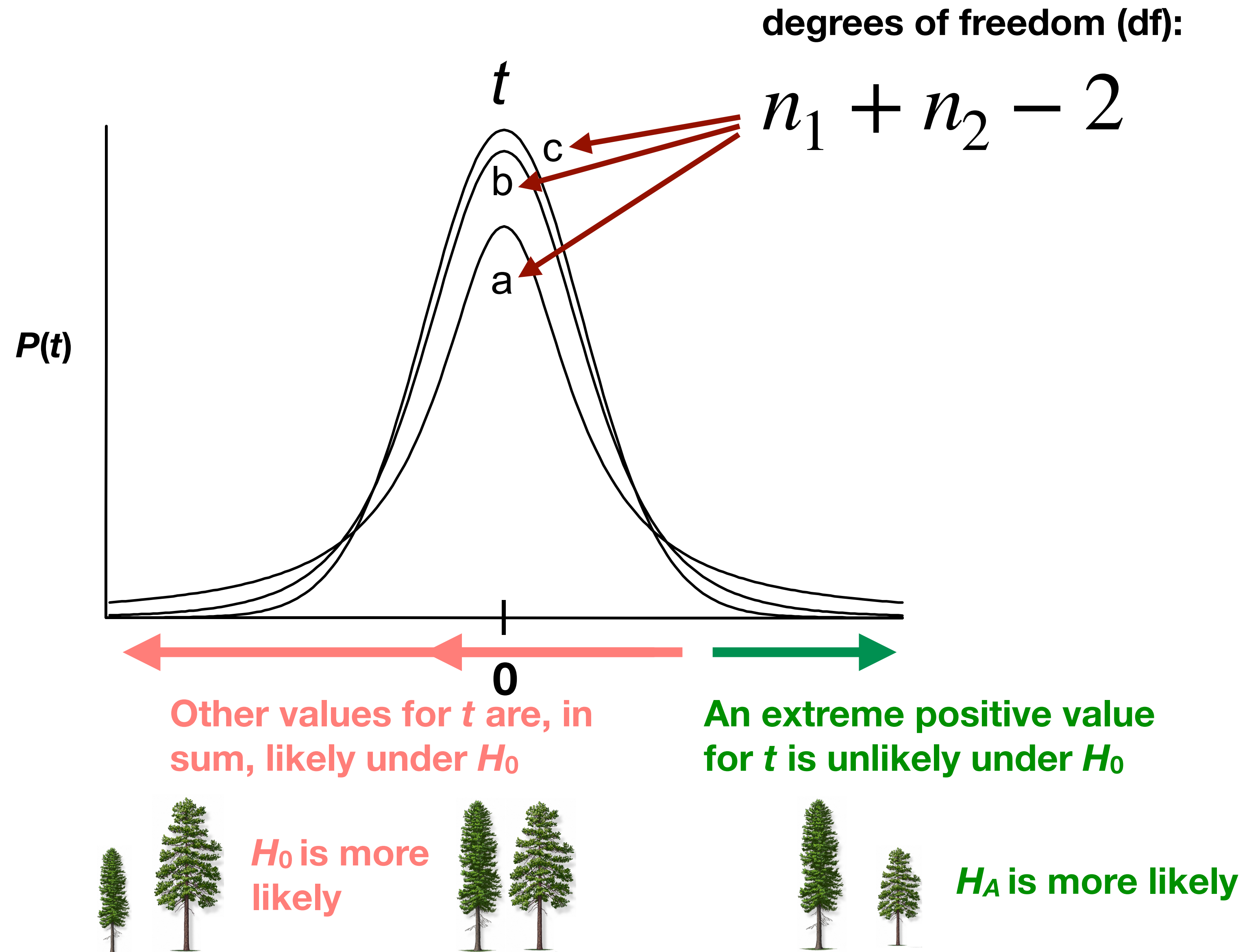
$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_{\bar{y}_1 - \bar{y}_2}}$$

$$s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

(Calculation for the std. error of mean difference)

Let's think about the sampling process for our t -statistic

The t -statistic and t -distribution

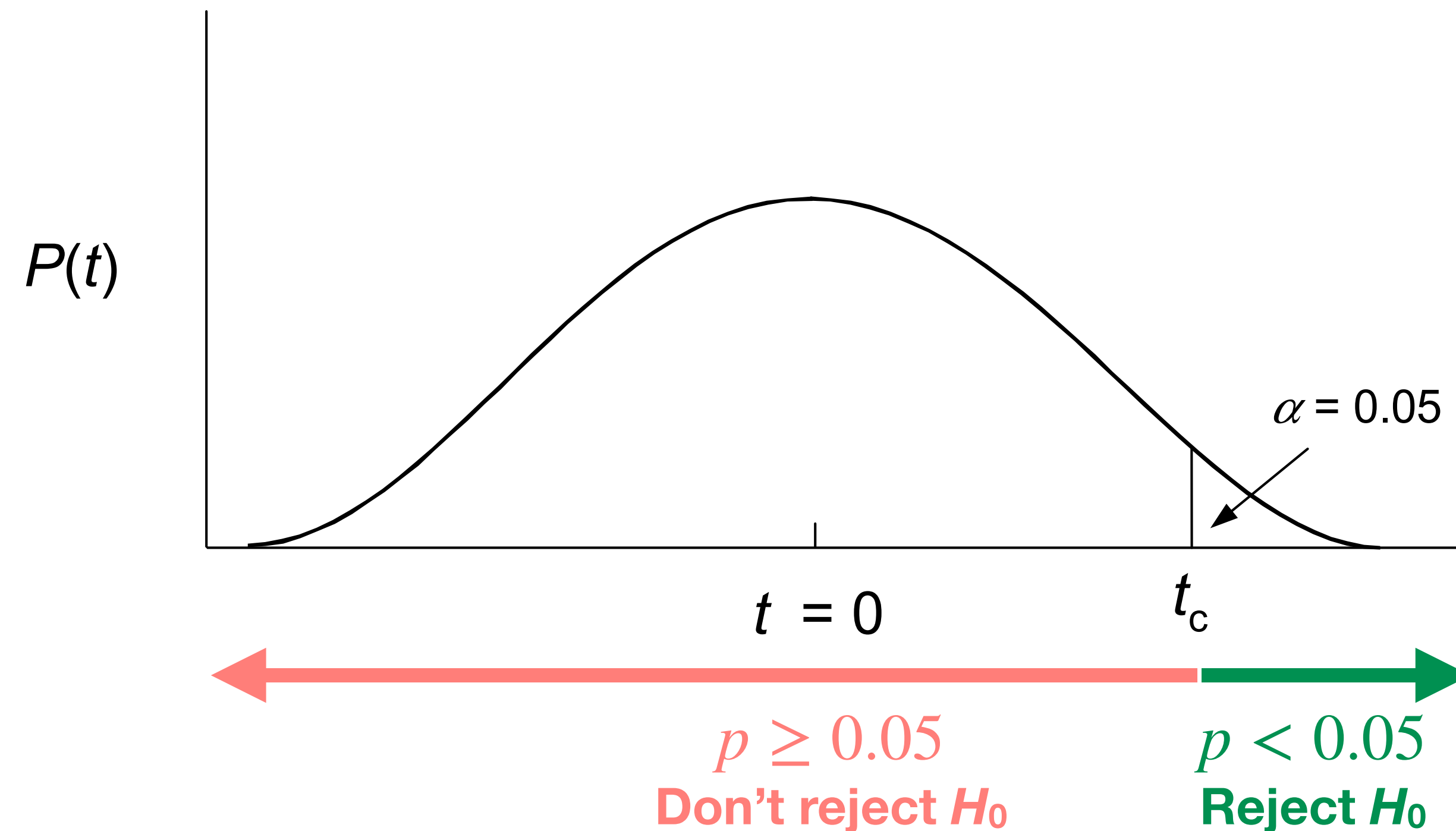


The t -statistic and two-sample t -test (**one-sided** test)

$$H_0 : \mu_1 \leq \mu_2$$

$$H_A : \mu_1 > \mu_2$$

Douglass fir trees are **taller** than
Ponderosa pine trees

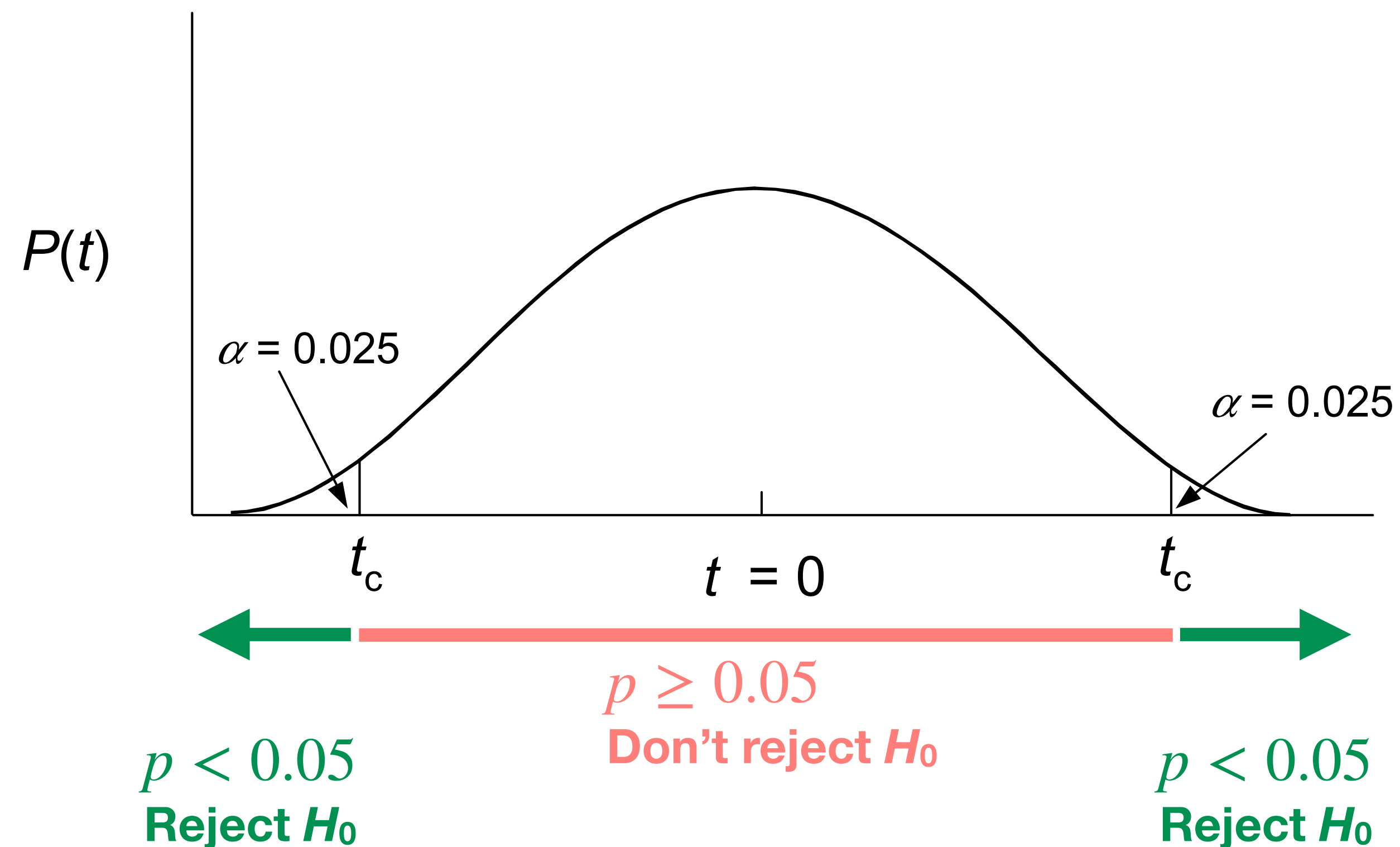


The t -statistic and two-sample t -test (**two-sided** test)

$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Douglas fir trees are **taller or shorter**
than Ponderosa pine trees



Assumptions for parametric t -tests that your sample(s) should meet

- The theoretical t -distributions for each degree of freedom were calculated for variables that meet the following criteria:

1) **normally** distributed in both populations

2) have **approximately equal variances** between populations

3) samples are **random**, composed of **independent** observations

- **Non-normality?**: can use nonparametric tests such as Mann-Whitney-Wilcoxon, or randomization / resampling tests

Randomization to generate a “null” distribution for t

Example: The number of successful broods of pseudoscorpion females that were mated twice to either a single male (SM) or two different males (DM).

SM: 6 5 4 4 2 3 4 2 4 2 0 2 0 1 2 6 0 2 3 3

DM: 2 0 2 6 4 3 4 4 2 7 4 1 6 3 6 4

Calculate the observed t -statistic: $t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_{\bar{y}_1 - \bar{y}_2}} = \mathbf{-1.39}$

Steps of the randomization test:

1. Create a randomized data set in which the measurements are randomly reassigned (shuffled) to the two groups (color coding indicates original treatment assigned)

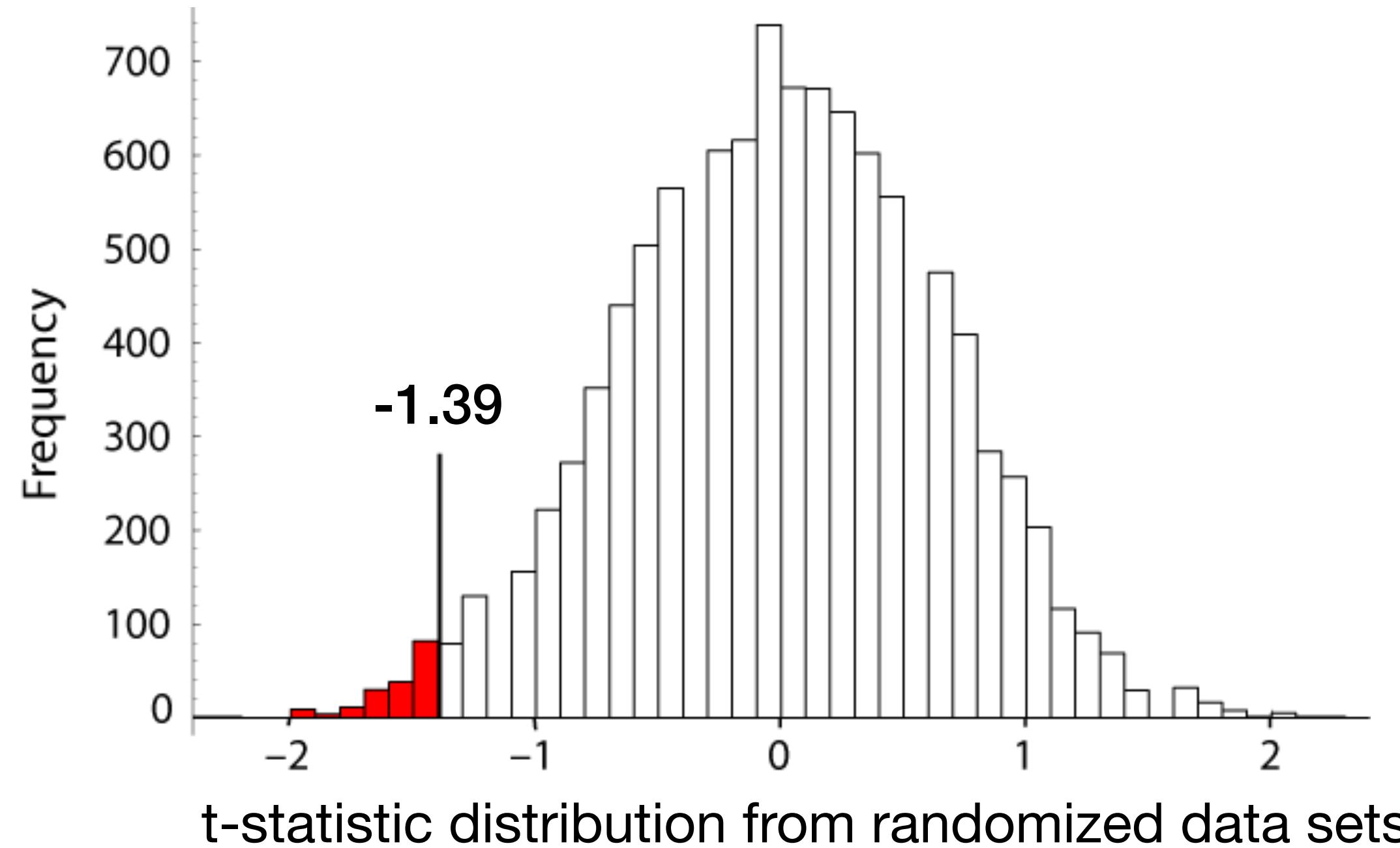
SM: 4 0 7 4 2 2 2 1 6 4 0 3 3 4 6 2 4 6 0 0

DM: 2 2 3 3 2 4 3 4 4 2 1 4 6 6 2 5

2. Calculate the t -statistic for the randomized sample: $t = \frac{(\bar{y}_1 - \bar{y}_2)}{s_{\bar{y}_1 - \bar{y}_2}} = \mathbf{-0.485}$

3. Repeat steps 1 and 2 many times to get an empirical null distribution for t

Randomization to generate a “null” distribution for t



The null distribution of the test statistic from 10,000 replications of the randomization process. (-1.39 is the observed value from the data)

Of these 10,000 randomizations, only 176 had a test statistic equal to or less than the observed value, -1.39 , so $p = 176/10000 = 0.0176$

Reject H_0 for a one-sided test: Singly-mated (SM) females have fewer successful broods than multiply mated (DM) females

Resampling (bootstrapping) to generate a “null” distribution for t

- 1. Use the computer to take two random samples of the data from all observations (with replacement, same sample sizes as 2 groups).**
- 2. Calculate a “t statistic” using the sample pair from step 1.**
- 3. Repeat steps 1 and 2 a very large number of times (≥ 1000)**
- 4. Calculate the proportion of bootstrap replicates with a value at least as extreme as your actual t from the two observed samples.**

This is the p-value for a one-sided test

