Hypothesis Testing:

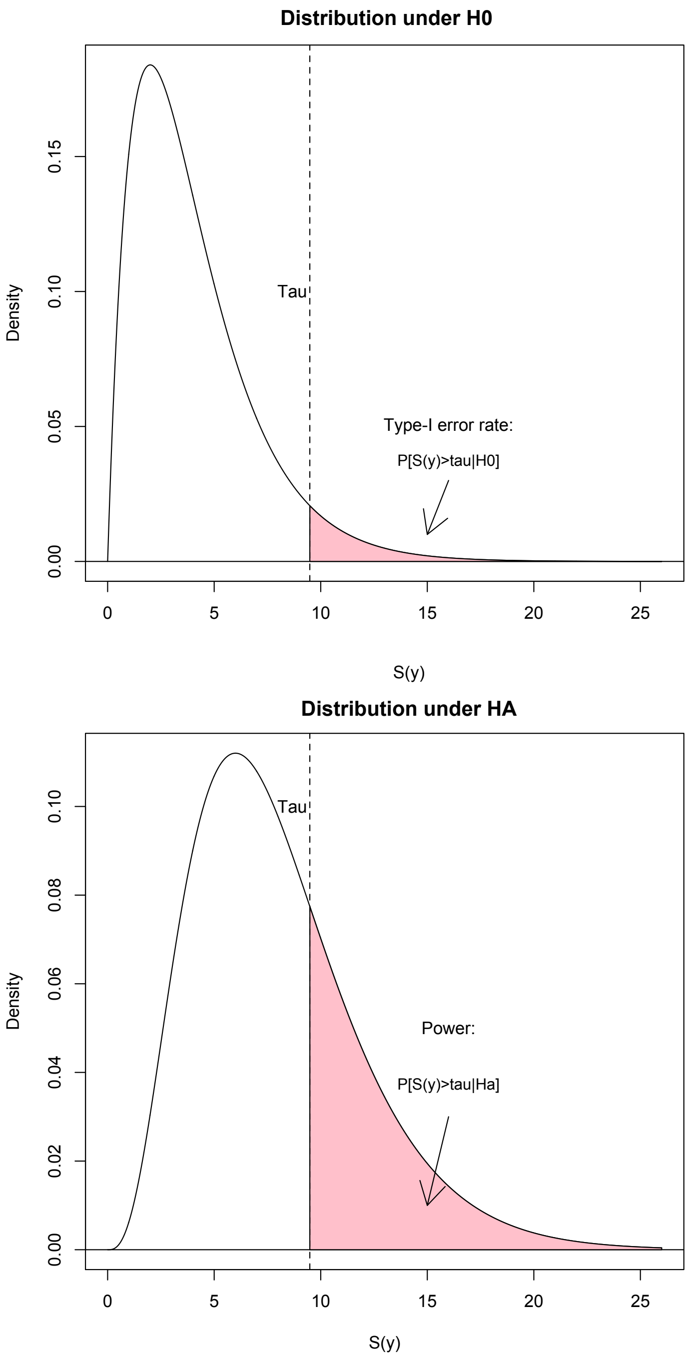
Power, Type-I Error Rate, and Sample Size Calculation

(Neyman-Pearson) Hypothesis Testing Framework

- A statistical model $p(\boldsymbol{y}|\boldsymbol{\theta})$
- Model parameters $\boldsymbol{\theta}$=[$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$]
- Hypotheses: H_0: $\boldsymbol{\theta}_1$=0 ; H_A: $\boldsymbol{\theta}_1 \neq$0 [e.g., a regression coef. =0]
- Test statistic: $S(\boldsymbol{y})$ [e.g., t-stat, F-stat, Chi-sq. stat]
- Decision rule: Reject H_0 if $S(\boldsymbol{y}) > \tau$ [e.g., reject if | t-stat |>1.96
- Possible cases

	Don't reject	Reject
H0	OK	Type-I error
Ha	Type-II error	OK

- The test-statistic, $S(\boldsymbol{y})$, is a function of the data; thus it is random
- It has a sampling distribution (which describes how it varies over conceptual repeated sampling).



Type-I error rate & Power

Type-I error rate:

- Definition: Probability of rejecting, given that the null holds.
- We choose tau to control Type-I error rate low.
- P-values are estimates of type-I error rate

Power:

- Definition: Probability of rejecting, given that Ha holds.
- The distribution of the test statistic under H0 is shifted (in this example, shifted towards the right)
- The area above the threshold is the power of the test.
- What factors affect power: Sample size, effect size, & and the test used.

Optimal decision rule:

- There are trade-offs between Type-I error rate and Power
- If we move tau “right” we reduce type-I error rate but this also reduces power.
- Approach: chose a rule that minimizes Type-II error while controlling Type-I error rate smaller than a given threshold (significance level).
- To achieve this we reject if p-value is smaller than the desired significance level.

Estimation of Type-I error rate & Power

Analytical methods:

- In same cases, based on either assumptions or asymptotic theory, we can have a good guess of the distribution of the test statistic under H0 and Ha.

- In these cases we can estimate power and type-I error rate analytically

Monte Carlo Simulations:

- However, in many cases the distribution of the test statistic under either H0 or Ha may not have a closed form
- In these cases we can estimate power and type-I error rate using simulations

Sample size calculation

We can use results from power analysis to estimate the minimum sample size that would be required for an experiment to achieve certain power.