ECO 602 Analysis of Environmental Data

FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Today's Agenda

1. Assignment 2 discussion
2. T-tests
3. Hypothesis testing in a t-test context
4. Group quiz/activity
5. Assignment 3 group formation

Announcements

- Moodle updates:
  - Lecture and reading schedule
  - Assignment schedule to be added to above

Assignment 2 discussion

- Why do we use both deterministic and stochastic models to describe associations in environmental data?
- Describe a deterministic model the authors used to characterize an association between model variables.
- Describe the stochastic component of the model.

T-tests

- You have all seen t-tests before, what are they?

T-test terms

- 1-sample, 2-sample
- 1-talied, 2-tailed
- Student’s t distribution
- Degrees of freedom
- T-statistic, test-statistic, critical value
- Standard error
- Univariate

T-test data

- Let’s draw some univariate data!

T-test null and alternative hypotheses

- 1-sample:
- 2-sample:
- Using iris data:
  - 3 species: setosa, virginica, versicolor
  - What are some possible 1-sample hypotheses?
  - What are some possible 2-sample hypotheses?

Hypothesis Testing

We will examine hypothesis testing from a t-test perspective.

A way to quantify the strength of evidence for or against a null hypothesis.

Null Hypotheses

We've considered these before in more conceptual model terms.

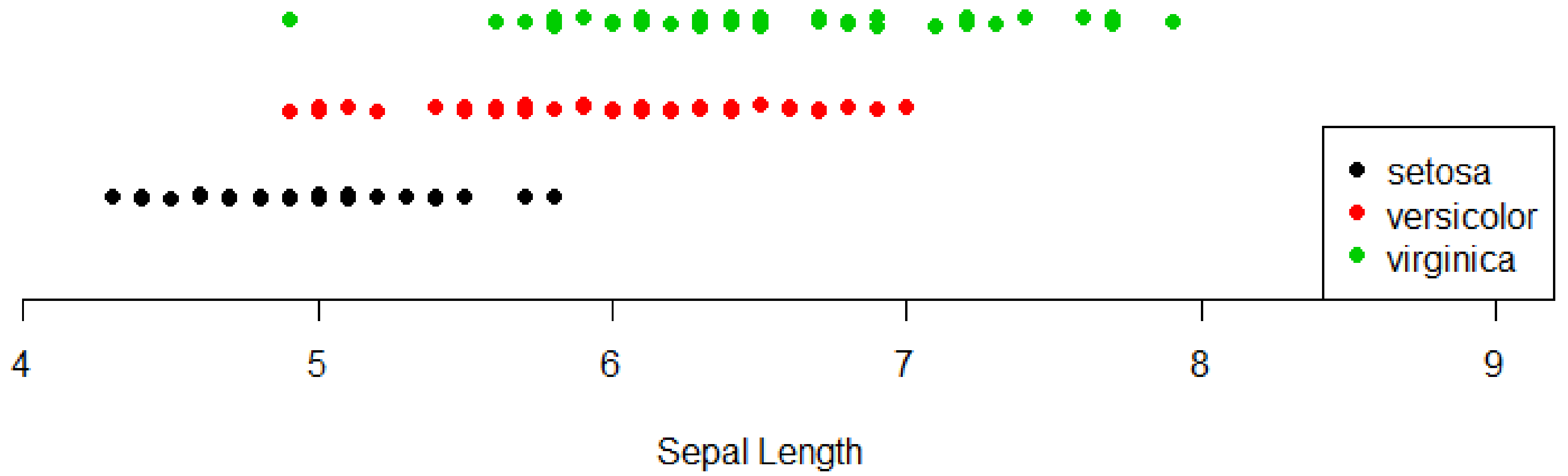
We need statistical models to conduct hypothesis testing.

1-sample: mean of group differs from a constant

2-sample: means of two groups differ

Iris T-tests

Sketch 1- and 2- sample tests



P-Values: Frequentist paradigm

P-values have the same interpretation difficulties as other frequentist ideas, that stem from the repeated-sampling paradigm.

P-values can be controversial!

P-values are meant to quantify the strength of evidence against the null hypothesis.

P-Values

P-values are the estimated Type I error:

- False positive rate
- Probability of falsely rejecting a true null hypothesis

Related to confidence intervals (which we'll talk about later)

The proportion of repeated-sampling events for which we would falsely reject the null hypothesis.

Clarification of terminology

Alpha: significance level, specified in advance

Beta: false negative rate, estimated after data collection

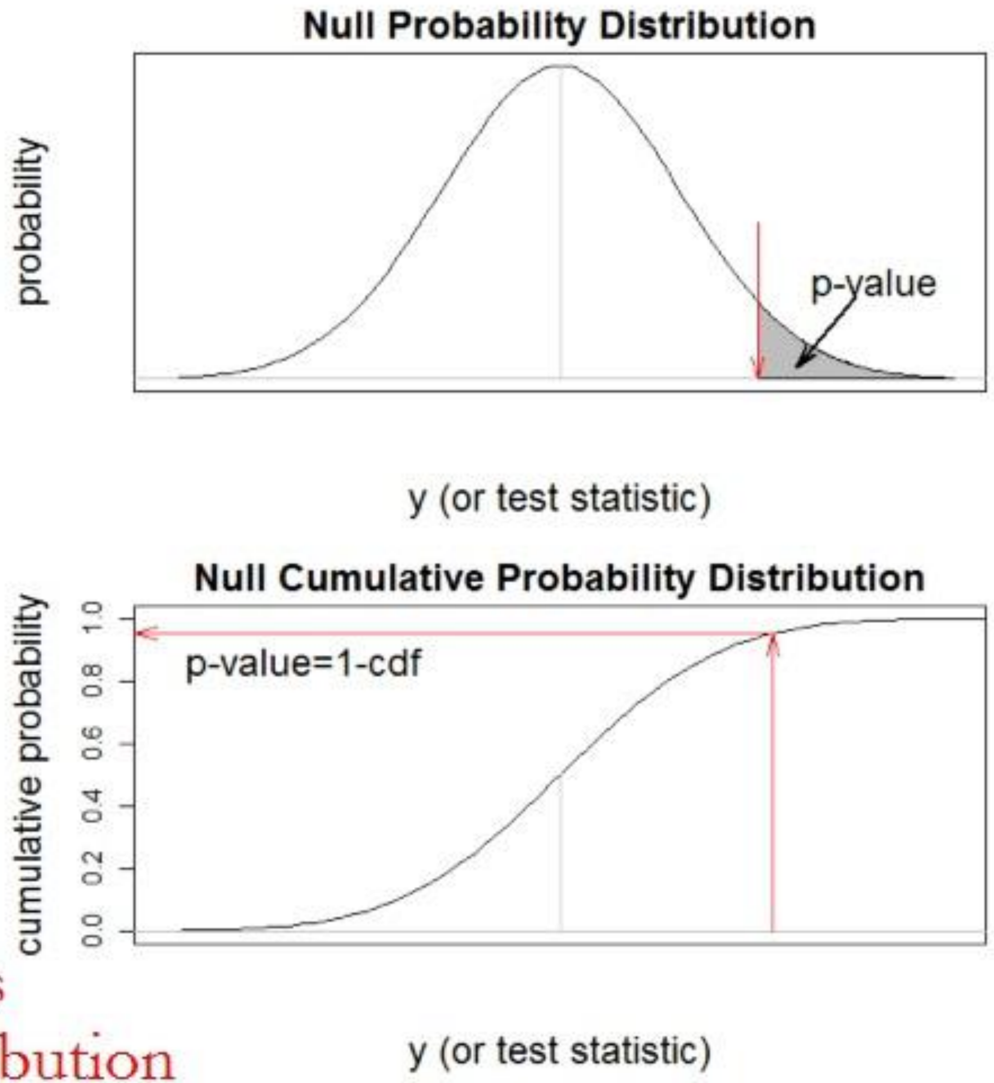
Critical value: test statistic must be more extreme than this value to reject null.

Hypothesis Testing Concepts

P-values

- *Probability* of observing data (Y, or a statistic derived from it, e.g., slope, mean) as large or larger (one-sided evaluation) if the null hypothesis is true (i.e., data was derived from the null probability distribution)
- Strength of evidence against the null hypothesis

Remember, p-values are always calculated under the Null distribution



McGarigal p-value diagrams

Two curves:

- Null hypothesis curve
  - Mean specified in advance, spread estimated from data
- Alternative hypothesis curve
  - Mean and spread estimated by data

- Sketch both curves on different plots first:

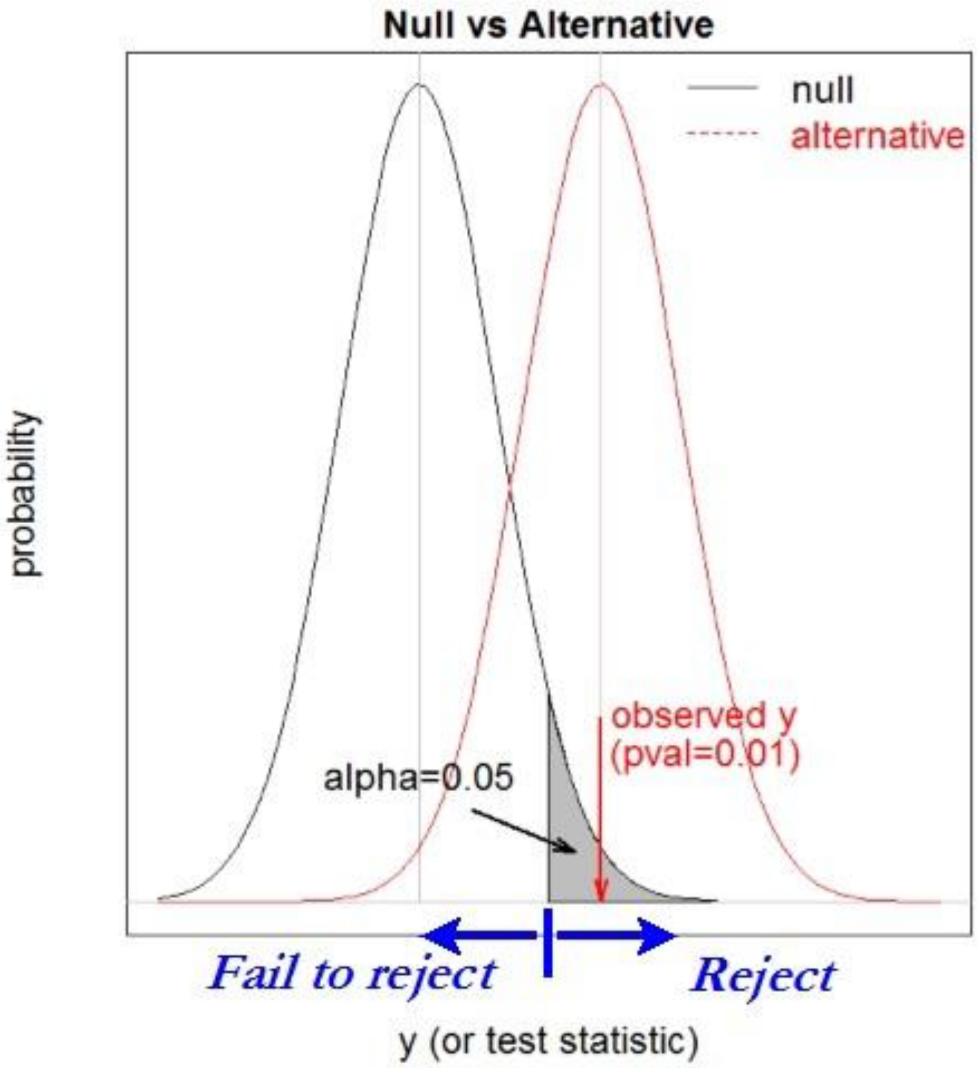
Hypothesis Testing Concepts

Neyman-Pearson decision framework

- *Reject* the null hypothesis if the *p*-value is less than a critical value (*alpha*), by convention usually ≤ 0.05

- *Fail to reject* the null hypothesis if the *p*-value is greater than *alpha* (i.e., there is insufficient evidence to disprove the null)

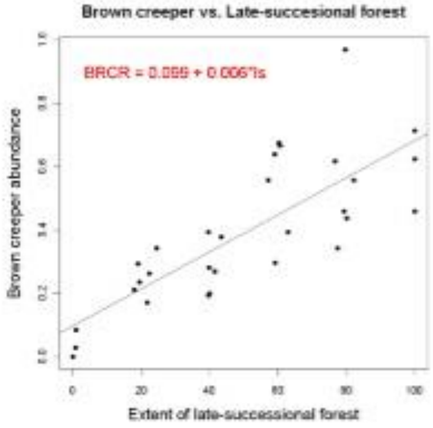
Remember, this applies to any probability distribution



Hypothesis Testing Concepts

P-values

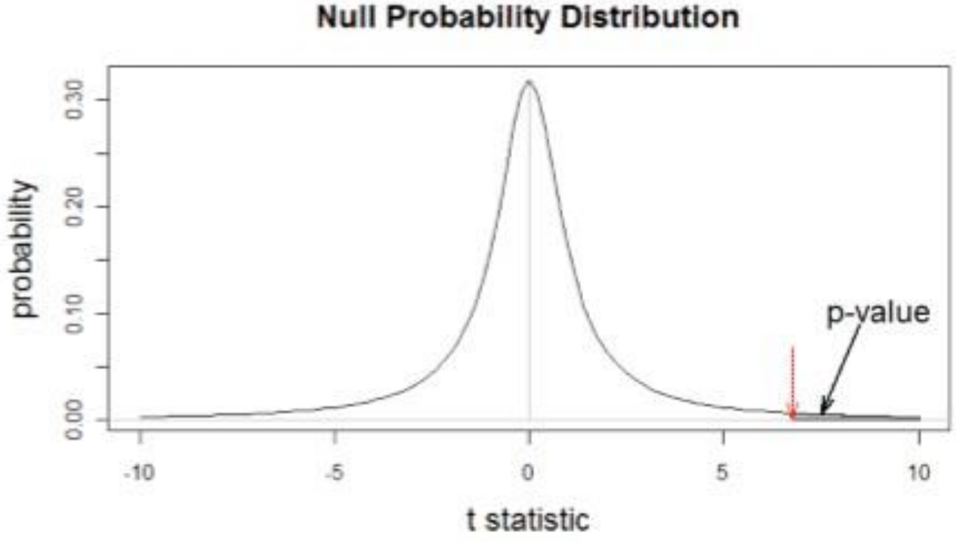
- *Parameters...*
  Probability of observing the value of φ (parameter estimate) under the null hypothesis (typically $\varphi = 0$), for any parameter with a sampling distribution.

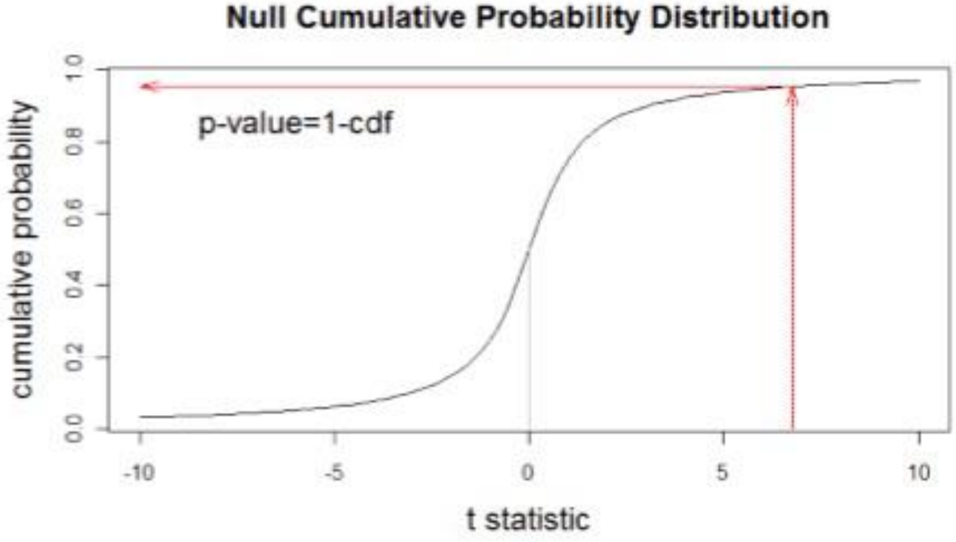


$$H_A\text{: } \Pr(\varphi \neq 0)$$

$$t_\varphi = \frac{\varphi_{obs} - \varphi_{null}}{SE_\varphi}$$

$$t_\varphi \sim t(\text{df}=1)$$





False negatives: beta

Beta is the type II error rate: failing to reject a false null hypothesis.

We select a p-value cutoff ahead of time: alpha

The false negative rate depends on our choice of alpha and the data.

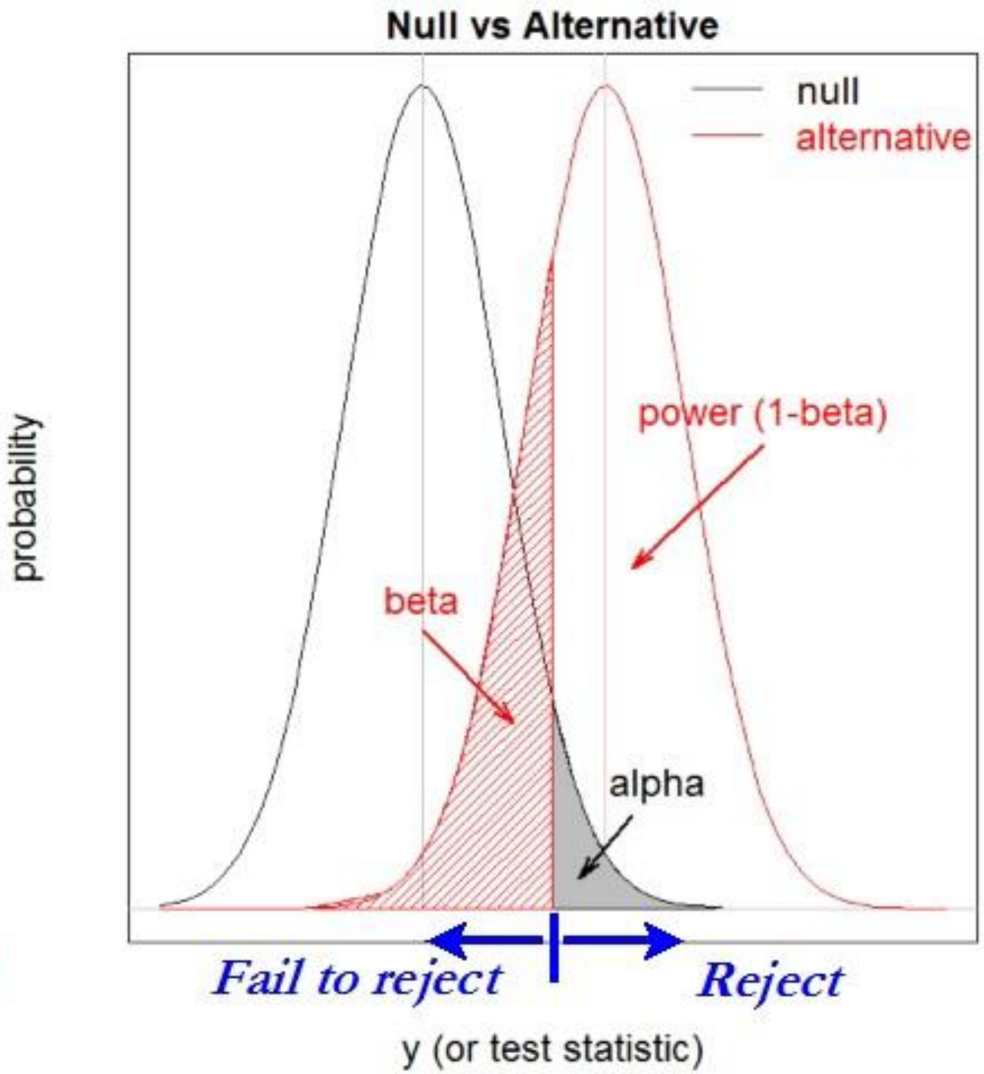
We cannot know beta until after we have collected data ☹

Hypothesis Testing Concepts

Neyman-Pearson decision framework

- *alpha* = probability of wrongly rejecting the null hypothesis (Type I error)
- *beta* = probability of wrongly accepting the null hypothesis (Type II error)
- *power* = probability of correctly rejecting the null hypothesis

alpha is under the null; *beta* and *power* are under the alternative



Power Analysis

Statistical Power: the probability that we **correctly reject** a **false null** hypothesis.

Statistical power is 1 – beta

We can't know our statistical power until after we collect data…

Factors that influence statistical power

Sampling error, sample size

Population variability

Effect size

Our choice of alpha

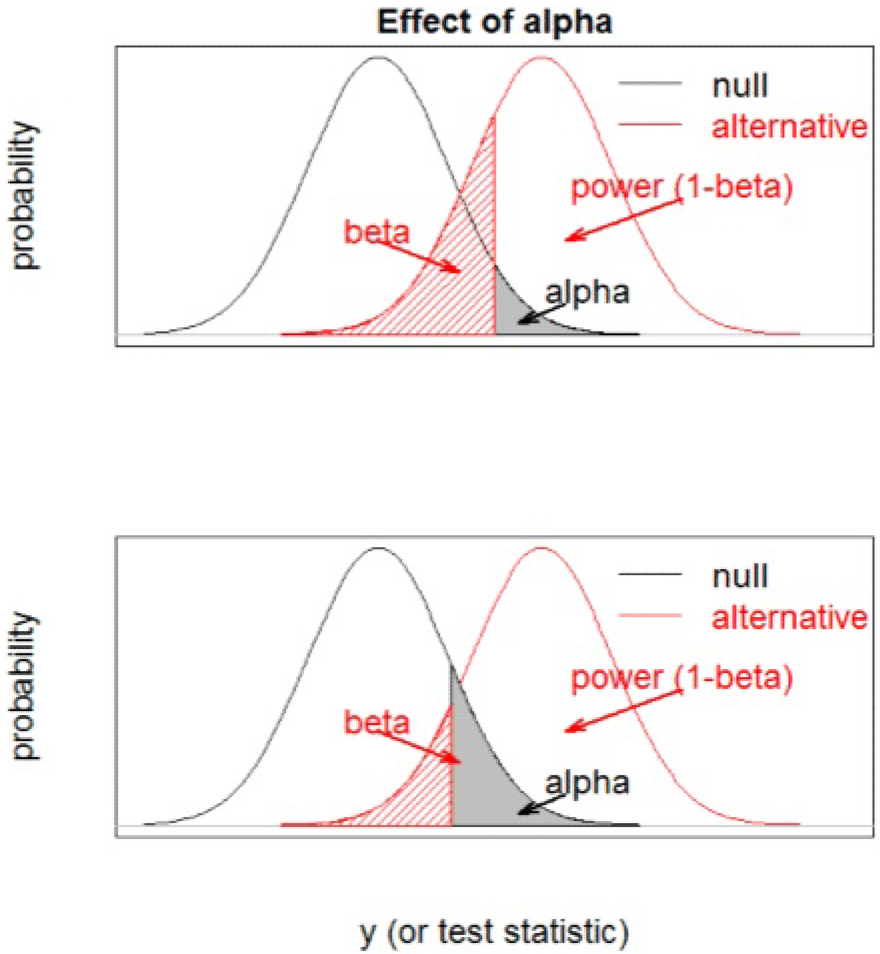
You cannot simultaneously decrease the false positive rate and increase statistical power!

Hypothesis Testing Concepts

Neyman-Pearson decision framework

Effect of alpha?

- Increasing alpha, increases power, all other things being equal

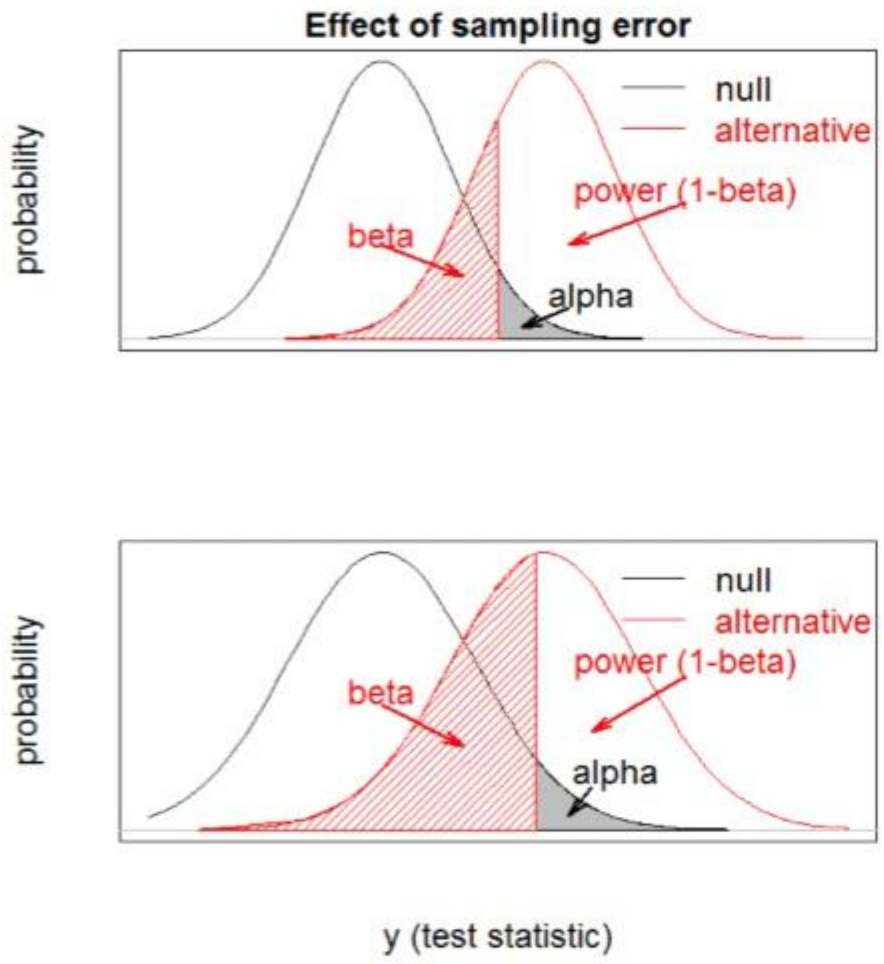


Hypothesis Testing Concepts

Neyman-Pearson decision framework

Effect of sampling variability (standard error)?

- Increasing sampling variability, either by increasing the variance in the underlying distribution or decreasing sample size (both effect sampling precision), decreases power, all other things being equal

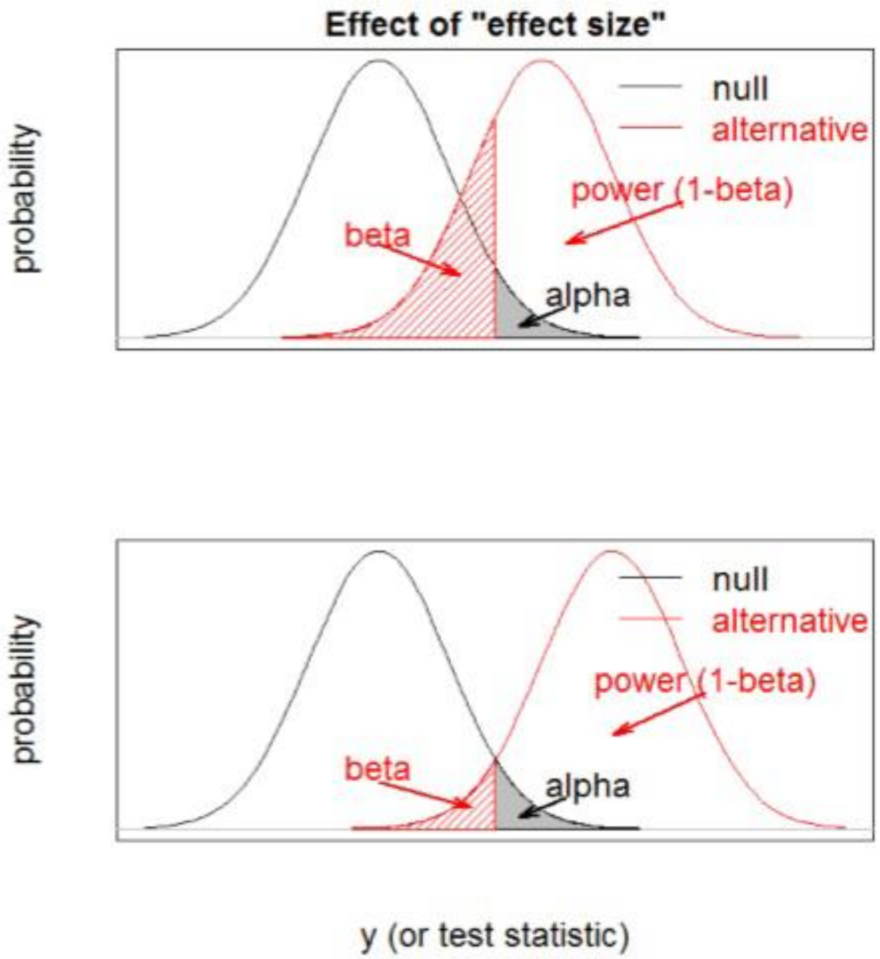


Hypothesis Testing Concepts

Neyman-Pearson decision framework

Effect of effect size?

- Increasing the effect size, increases power, all other things being equal



In-class group quiz

Sketch a uniform Probability Density Function

- Don’t worry about y-axis scale

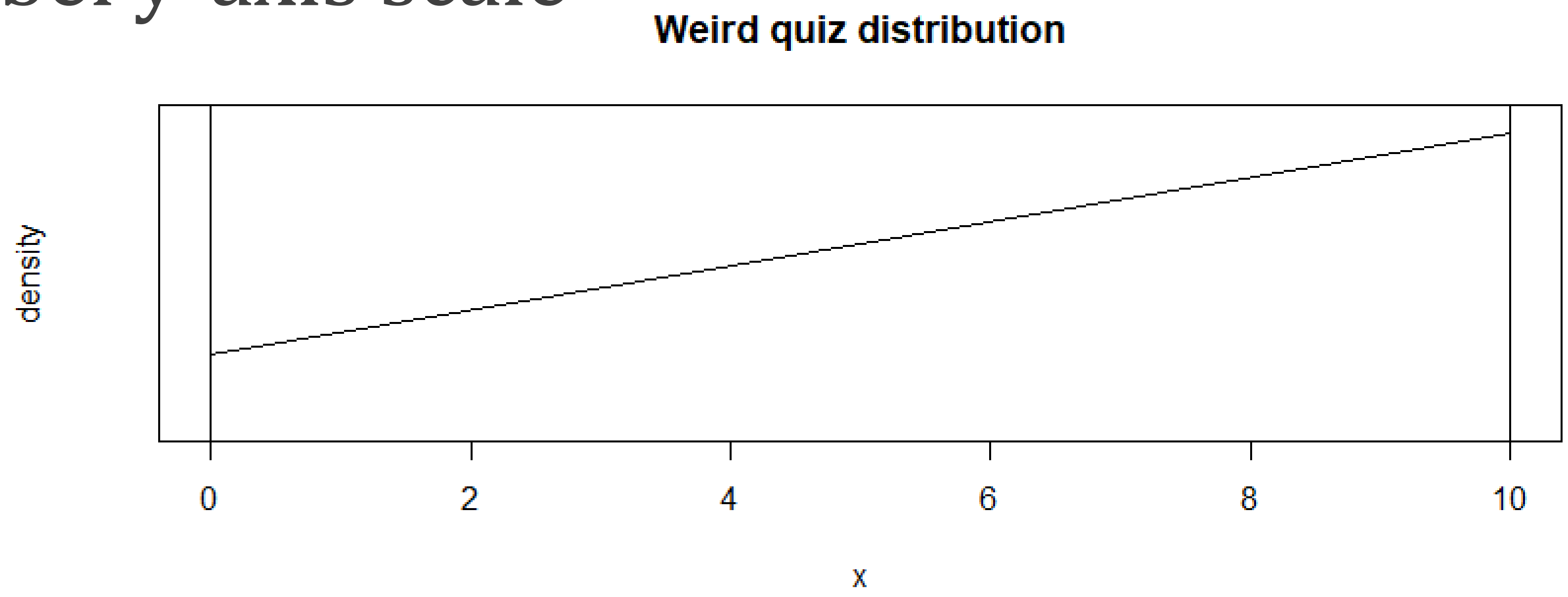
Sketch the corresponding Cumulative Density Function

- Label y-axis scale

In-class group quiz

Sketch the corresponding Cumulative Density Function of the weird distribution

- Label y-axis scale



In-class group quiz

Sketch curves for null and alternative hypothesis for a t-test.

- Alpha = 0.5
- 1-talied, upper tail
- H0: mean = 0

Assignment 3

Please form groups

I will add a group self-selection page

Assignment will be posted later today

We’ll talk about it on Thursday

For next time

Confidence intervals: McGarigal 6b