Connectomics: Null Hypotheses

Horacio Gómez-Acevedo

November 1, 2022

UAMS

Today's paper

REVIEWS

Check for updates

Null models in network neuroscience

František Váša[1] *and Bratislav Mišić*[2]✉

Main Ideas

- Rigorous identification and quantification of functionally important features of the brain network architecture.
- Using Null models for the presence and relevance of features (sensitivity analysis).
- Feature randomization (or bootstraping) to formally investigate relevance.
- Multiple null models are possible.

Features of Interest

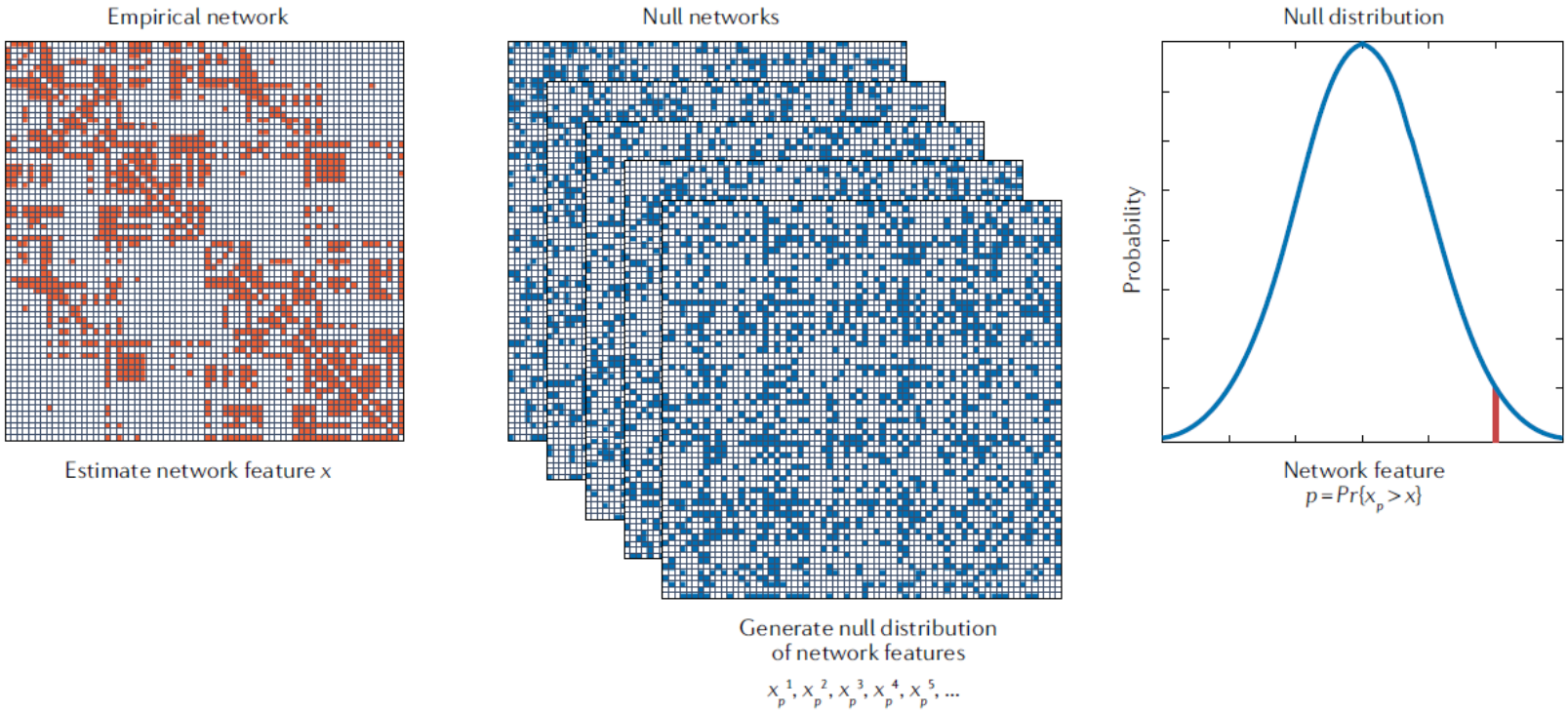
Research on imaging across species have "reconstructed" graph organization of brain networks and have determined that:

- Connection profiles
- Degree distribution of nodes is heavy tailed ("few hyper-connected nodes")
- Densely interconnected network modules

How to determine the importance of features of interest (FOI)? **Null models** provide a way to systematically compare networks and reveal the presence of a FOI arises as a consequence of other features.
What are network (graph) features?

1. Density, which refers to the proportion of edges.
2. Degree sequence, which refers to the number of edges incident on each node
3. Topology

Null Distribution



I

Graph Null networks

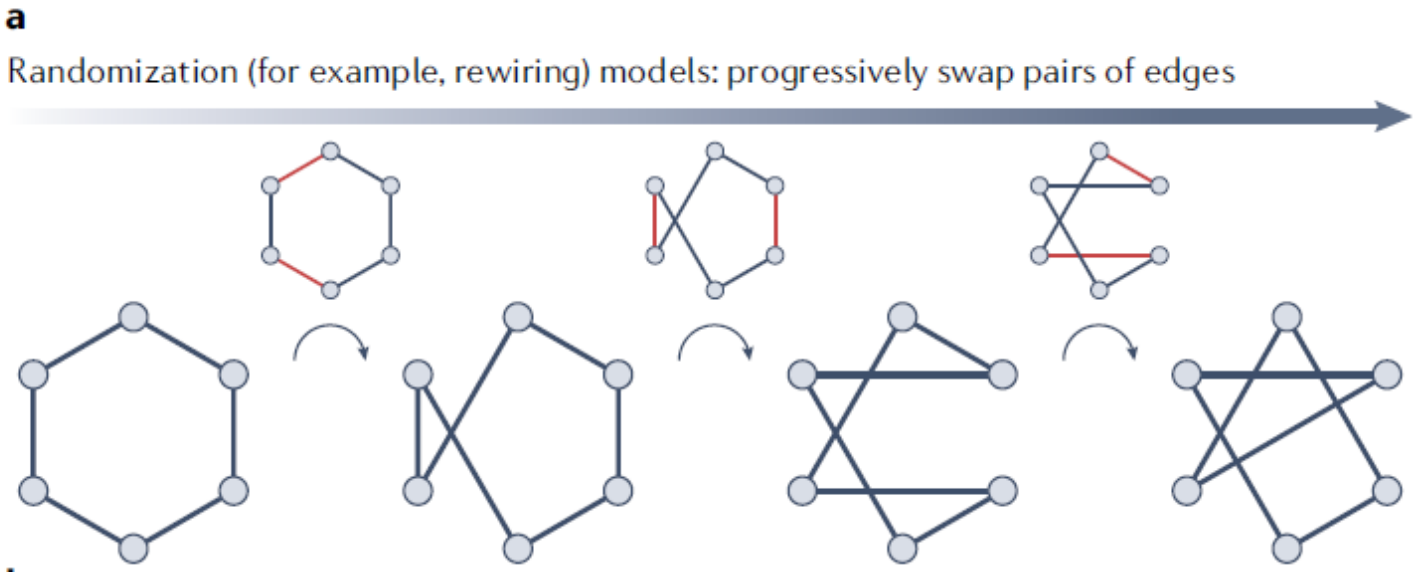
We begin with a network feature F_I (e.g., path length) and an empirical network (adjacency matrix).
We generate a population of null networks by randomizing other feature F_J (e.g., topology) and calculating the same feature F_I for each of the randomized variants $F_I^{J_1}, F_I^{J_2}, \ldots$
Generate the distribution of $\{F_I^{J_k}\}$. The corresponding hypothesis test

$$H_0 : F_I = F_I^0 \qquad H_1 : F_I > F_I^0$$

If you calculate the p-value associated with the feature F_I distribution.
"construct a distribution of features $\{F_I^{J_k}\}$ under the null hypothesis that the magnitude of feature F_I^0 is due to properties that were preserved, and not due to properties that were randomized (and not preserved)"

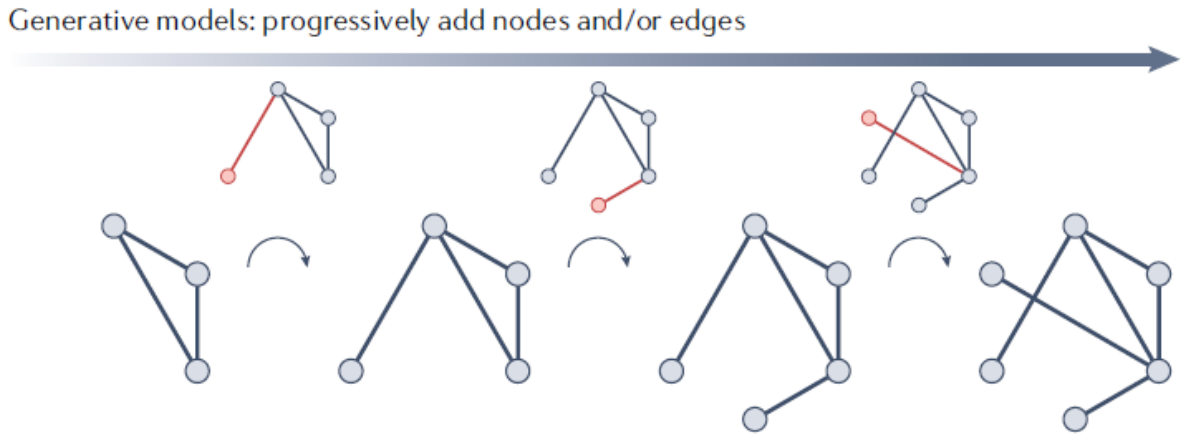
Randomization

A frequently used method is *rewiring* in which pairs of edges are selected at random and then swapped.



Generative Models

These models *build* the network using a predefined wiring rule that embodies the null hypothesis. This process stops when the network has the same size and density as the observed one.

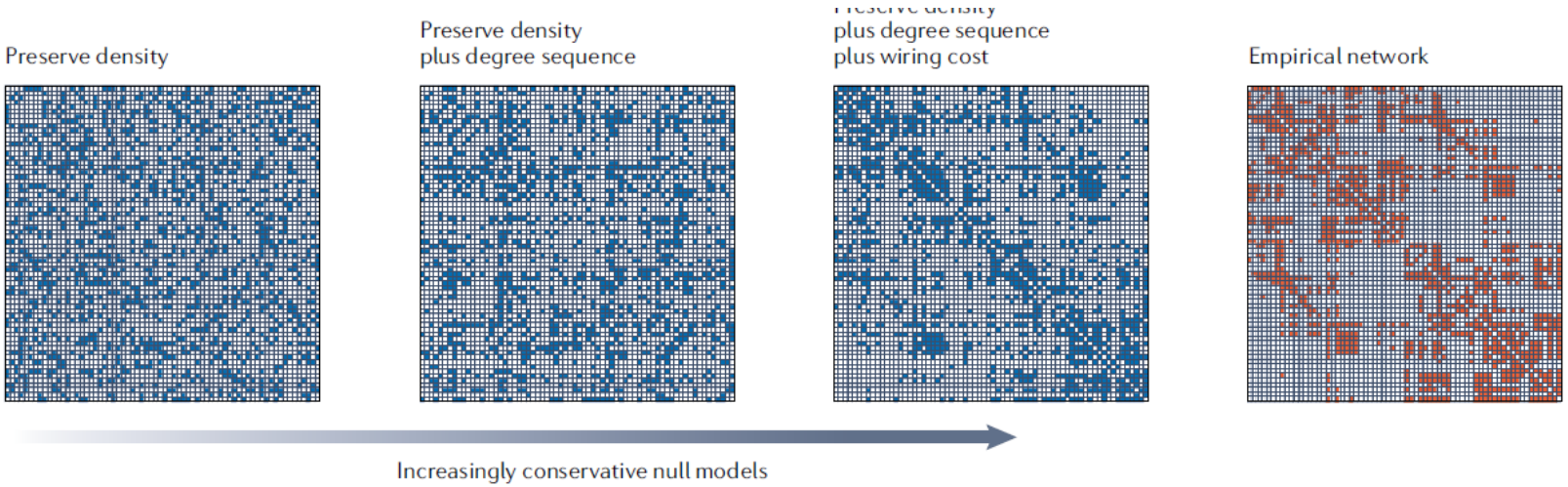


Other uses of Generative Models

These models have been also used for **model identification** or **model comparison**.
As long as you have a **cost function**, the generative models produce a way to select an optimal solution.
"This is conceptually similar to model identification in formulations of brain networks as dynamic systems, such as dynamic causal modelling, in which competing accounts of dynamic neural circuit interactions are tested to identify the best- fitting or most parsimonious model, or when competing hypotheses are grouped into distinct families of mechanisms or families of models"

Null after null...

"An important methodological question is whether null models uniformly sample the target space. The mere fact that a model retains one feature and randomizes another does not mean that it samples the space of all possible realizations exhaustively. "



Null after null... (cont)

"Ultimately, there is no right or wrong null model. The null model should be an implementation of an explicit and falsifiable null hypothesis that is specific to one's research question."

Null models for spatial networks

"Perhaps the most important feature to consider when analysing brain network topology is geometry"
The challenge is the bias in connectivity between "close" neural elements.
Also, the standard rewiring with naive swapping algorithm tend to form networks with unrealistic (and costly) configurations.
One possible alternative is to use **spatially repositioned nulls** that preserve topology but allow locations to be permuted.
"Collectively, geometric nulls enable us to quantify the proportion of topology that comes passively from spatial embedding"

Annotated networks (multi-omics)

"A typical comparison may involve correlating, across brain regions, a region's network attribute (such as degree) and its microscale annotation (such as gene expression)"

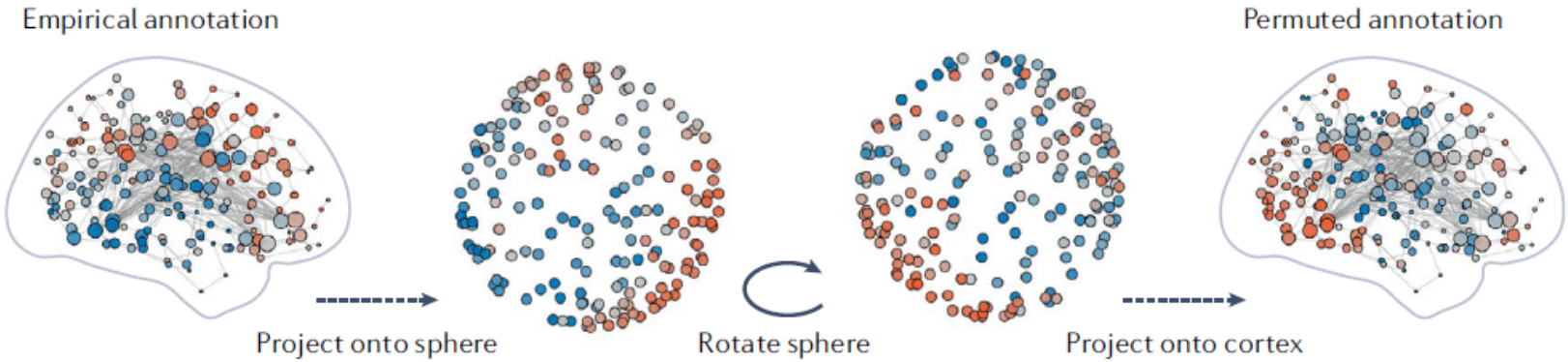
We are violating independence!

"However, an important problem arises when estimating a p value for the correlation coefficient. Namely, the standard parametric method assumes that the two vectors come from an uncorrelated bivariate normal distribution, whereas the non- parametric (naive permutation- based) method assumes that the elements of the vectors are exchangeable. This independence assumption is violated by multiple forms of dependence between data points"

Problems with independence

- Spatial autocorrelation. Similar values of anatomical and physiological measurements between neighbouring locations.
- Homotopic symmetry. Similar measurements between corresponding locations within the left and right hemispheres of the brain.
- Spatial resolution. The effect size on the number of regions, vertices or voxels.

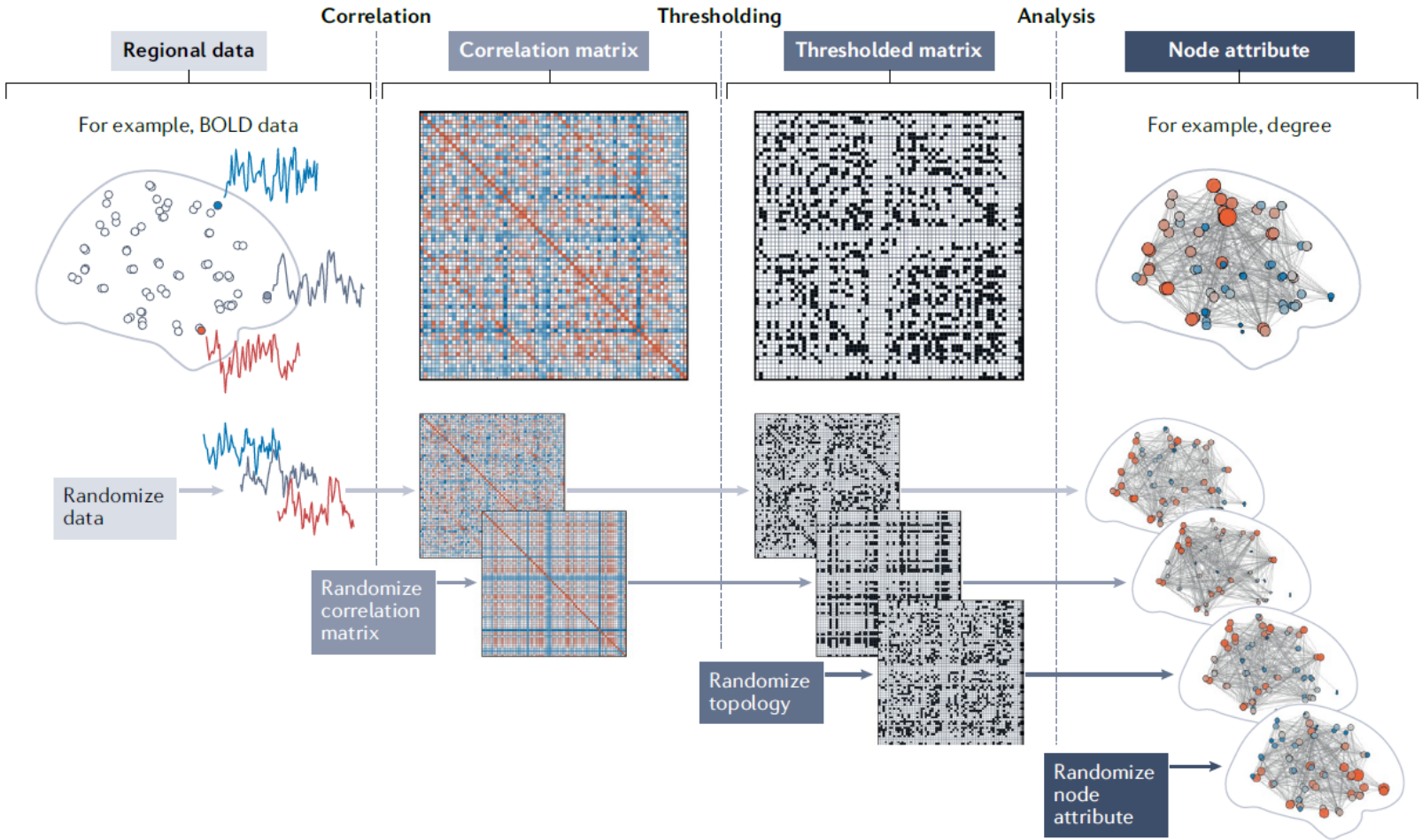
Spin Test


Empirical annotation
Permuted annotation
Project onto sphere
Rotate sphere
Project onto cortex

Correlation Networks

Connectivity is often measured with covariances. Thus, these networks are subject to transitivity issues.
If node A is positively correlated with nodes B and C, nodes B and C may be negatively correlated.
We can randomize the signal itself:
"surrogate time series can be generated by transforming empirical time series to the frequency domain using the Fourier transform, shuffling the phase coefficients and taking the inverse transform to the time domain. The resulting surrogate time series have preserved power spectra but randomized temporal dependencies"
See this link for Fourier Transform

Randomization on steroids



Conclusion

"In network neuroscience, null networks are typically created through complete randomization, and, conversely, generative models are typically fully completed. However, both processes can be paused incrementally to explore the immediate vicinity of an empirical network in null- model space, thereby giving insight into the stochastic neighbourhood of brain networks. Such 'connectome mutagenesis' can offer insight into pathological perturbations involved in psychiatric and neurological disorders"

Hypothesis Testing

There are key components to take full advantage of th

- Likelihood
- Statistical Hypothesis
- Hypothesis Test
- Likelihood Ratio
- Asymptotic Properties of the Normal Distribution
- *p*-values

Likelihood

If we have $X_1, \ldots, X_n$ independent and identically distributed (iid) random variables with a common probability (mass/density) function $f(x; \theta)$ where the parameter θ is unknown ($\theta \in \Omega$). The likelihood of a sample $\vec{x} = (x_1, \ldots, x_n)$ is

$$L(\theta, \vec{x}) = \prod_{i=1}^{n} f(x_i, \theta)$$

Example: If we have $X_i \sim N(\theta, \sigma^2)$ with $\sigma^2 > 0$ known but θ unknown. Then

$$L(\theta, \vec{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2\right) \exp\left(-\frac{1}{2\sigma^2} n(\overline{x} - \theta)^2\right)$$

Statistical Hypothesis

A Statistical Hypothesis is a conjecture about the probability distribution of a population.
Example: We suppose that in an experiment we have a random sample from $N(\theta, 10)$.

$$H_0: \text{ The population is } N(5, 10)\text{-distributed}$$
$$H_1: \text{ The population is } N(1, 10)\text{-distributed}$$

Hypothesis Test

A Hypothesis Test is a tuple $(X_1, \ldots, X_n; H_0, H_1, G)$, where

1. $(x_1, \ldots, x_n)$ is a sample of $(X_1, \ldots, X_n)$ random variables iid.
2. H_0 and H_1 are hypothesis concerning the probability distribution of the population.
3. $G \subset \mathbb{R}^n$ is Borel set (countable unions of open sets).

The **level of significance** is defined as

$$\alpha = P^{H_0}_{X_1,\ldots,X_n}(G)$$

We will consider hypothesis such as

$$H_0 : \theta = \theta_0 (\text{ or } \theta \in \Theta_0) \qquad H_1 : \theta \neq \theta_0 (\text{ or } \theta \in \Theta = \Theta_1 \cup \Theta_0)$$

Maximum Likelihood Test

The **likelihood ratio** function

$$\Lambda(x_1, \ldots, x_n) = \frac{\sup_{\theta \in \Theta_0} L_\theta(x_1, \ldots, x_n)}{\sup_{\theta \in \Theta} L_\theta(x_1, \ldots, x_n)}$$

Let $\widehat{\theta}$ be the maximum likelihood estimate of θ.
If θ_0 is the true value of θ, then $L(\theta_0)$ is the maximum value of $L(\theta)$.
Since $\Lambda \leq 1$, then if H_0 is true Λ should be close to 1, whereas if H_1 is true then Λ should be smaller.
We have the decision rule

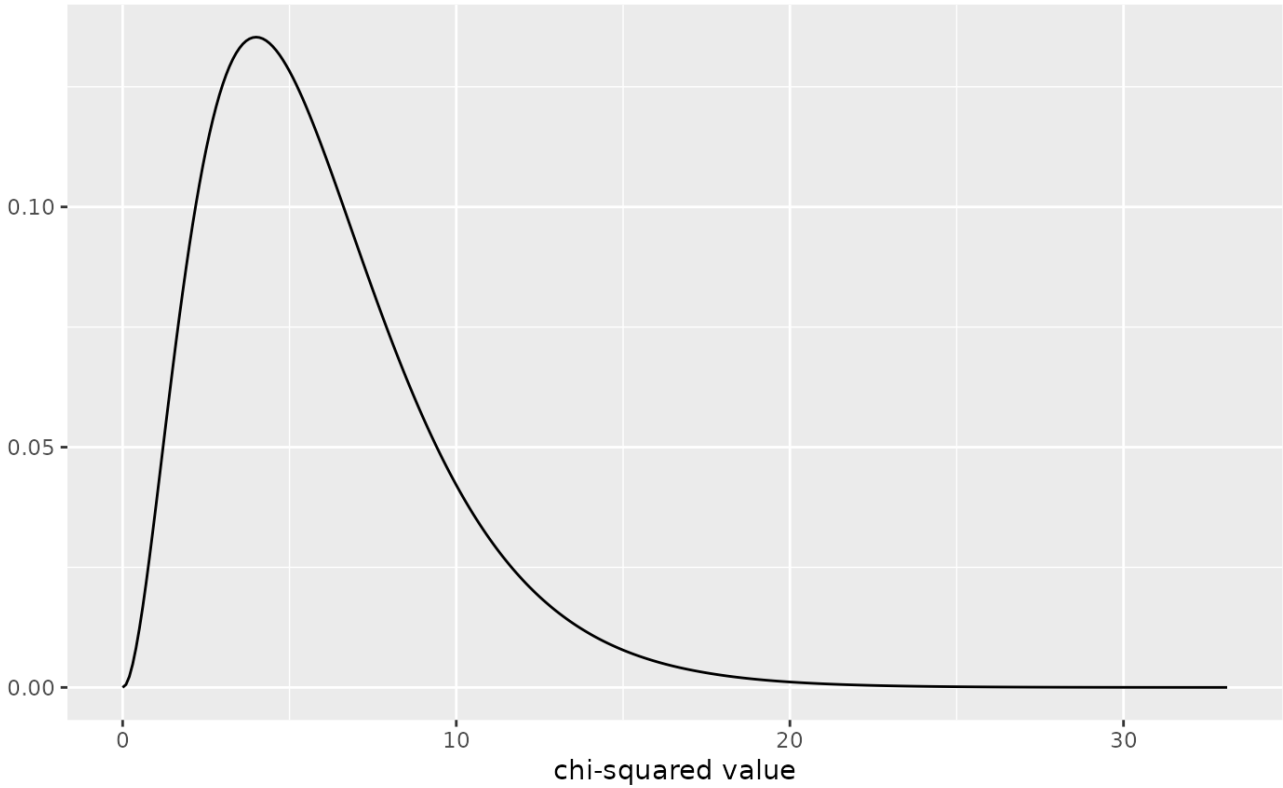
$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } \Lambda \leq c,$$

where $\alpha = P^{\theta_0}(\Lambda \leq c)$.

Asymptotic Properties

Under some (regularity)conditions we have the following result:
If the null hypothesis $H_0 : \theta = \theta_0$.

$$-2 \log \Lambda(X_1, \ldots, X_n) \rightarrow \chi^2(1)$$



Asymptotics for the Normal distribution

When μ and σ are unknown and testing the hypothesis

$$H_0 : \mu = \mu_0 \quad H_1 : \mu \neq \mu_0$$

The likelihood ratio is given by

$$\Lambda(x_1, \ldots, x_n) = \left\{ 1 + \frac{1}{n-1} \left(\frac{\overline{x} - \mu_0}{s/\sqrt{n}} \right) \right\}^{-n/2}$$

where $s^2 = \frac{1}{n-1} \sum_i (x_i - \overline{x})^2$. The critical regions are of the form

$$G = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \left| \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \right| \geq c\}$$

p-values

The test statistic $\frac{\overline{x}-\mu_0}{s/\sqrt{n}}$ is critical to reject H_0 or not. The decision procedure is as follows

$$\text{if } \left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq c \text{ then we assume } H_1$$

$$\text{if } \left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| < c \text{ then we assume } H_0$$

Furthermore, we have the following equivalence

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq u\right) \leq \alpha \iff u \geq c$$

The p-value associated with the outcome u of the test statistic $\frac{\overline{x}-\mu_0}{S/\sqrt{n}}$

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq |u|\right)$$

p-values (cont)

Thus if an outcome u of $\frac{\overline{x}-\mu_0}{S/\sqrt{n}}$ satisfies

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq |u|\right) \leq \alpha \text{ we accept } H_1$$