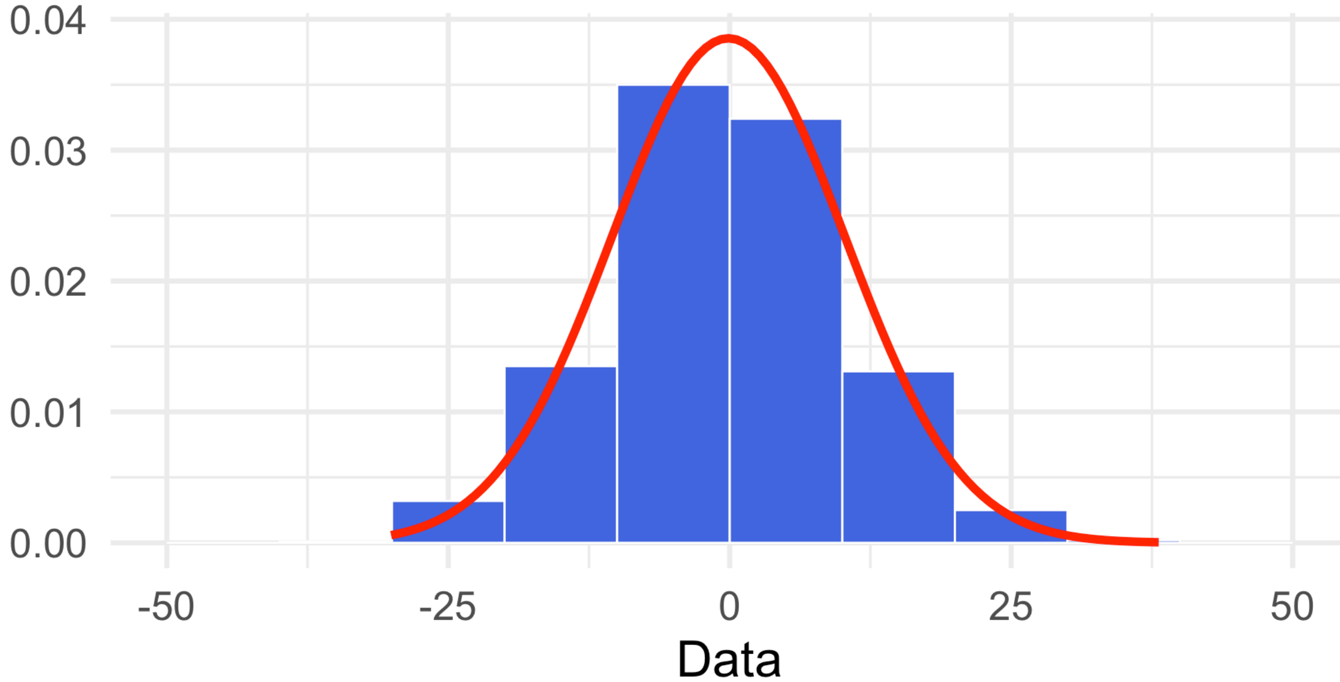
A Comprehensive List of Normality Tests in R

Fernando Corrêa

2024-07-09

Testing for normality is very common

Question: Does the data comes from a normal distribution?



- Use cases: regression, ANOVAs etc
- We don't need to rely only on graphics. There are more than 50 normality tests (and counting!)
- Many of them are (or where) implemented in R. e.g. `PoweR` had many implementations, but is no longer in CRAN
- Usual syntax: `x_sample <- rnorm(100); TEST(x_sample)$p.value`

Strategies for testing normality

Tests based on distances between ECDF/smoothed ECDFs vs expected deviations under gaussian sampling

- Kolmogorov-Smirnov (KS) `stats::ks.test`
- Lilliefors (L) `nortest::lillie.test`
- Anderson-Darling (AD) `nortest::ad.test`
- Cramer-von Mises (CVM) `nortest::cvm.test`
- Jin Zhang's revised versions (Z-K, Z-A, Z-C)
  - `DistributionTest::za.test(..., "norm")`

Tests based on the measured $\bar{X}$, $AVG((X-\bar{X})^2), AVG((X-\bar{X})^3), \ldots$ vs expected deviations under gaussian sampling

- Jarque-Bera (JB) `moments::j.test`
- Anscombe-Glynn (AG) `moments::anscombe.test`
- D'Agostino-skewness (DA) `moments::agostino.test`

Strategies for testing normality

Tests based on correlations

- Shapiro-Wilk (SW) `stats::shapiro.test`
- Shapiro-Francia (SF) `DescTools::ShapiroFranciaTest`

Tests based on the entropy measures

- Vasicek-Song tests `DescTools::vs.test(..., densfun = 'dnorm')`

Tests based on the χ^2 distance between histogram counts and expected counts

- Pearson's χ^2 test `DescTools::PearsonTest`

Many other tests are not on CRAN anymore, some procedures might be recovered from older versions releases

- `PoweR`, `normtest`

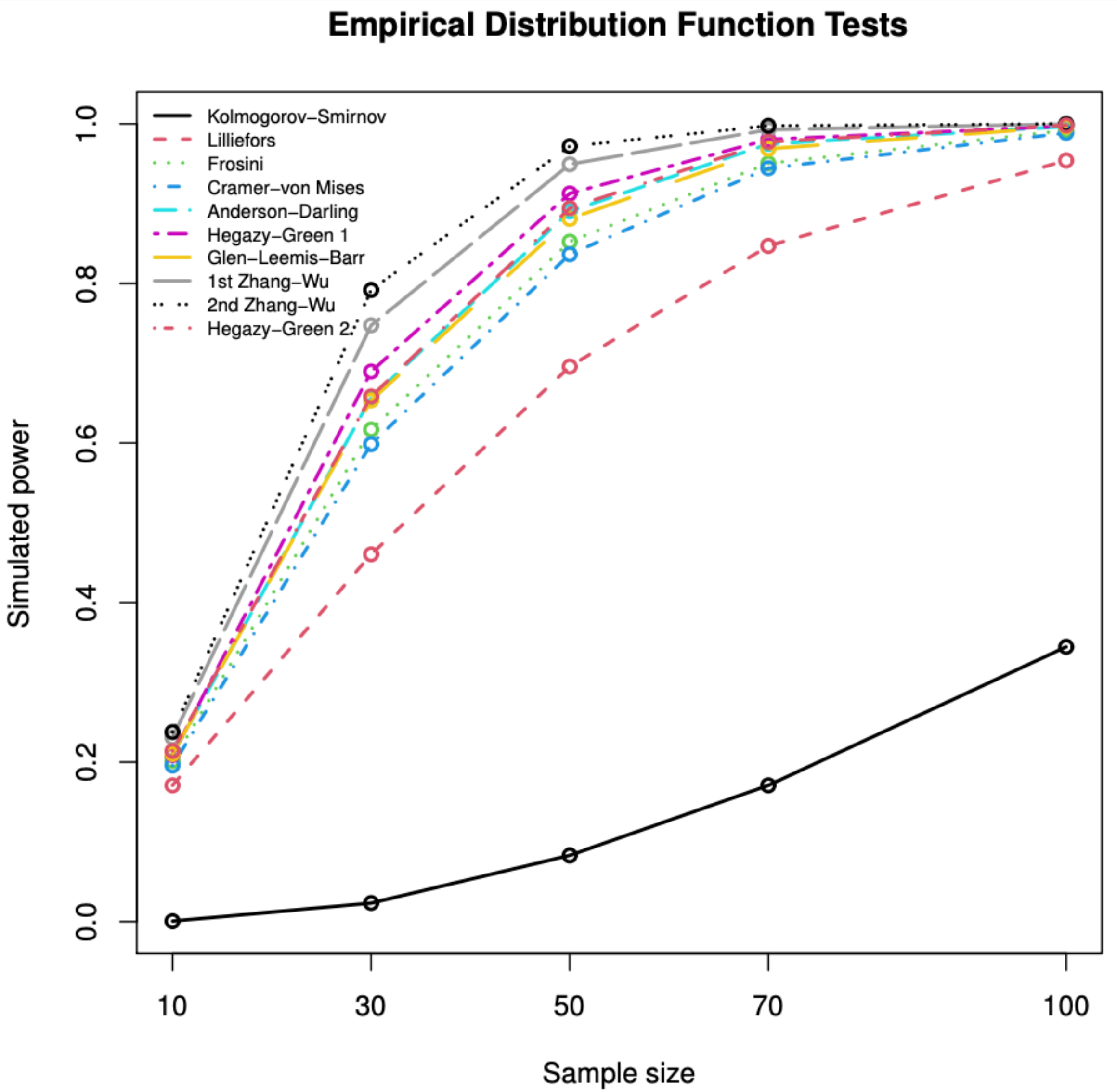
Which one should I pick?

There are many papers on this matter. We'll base our conclusions mainly in a 2022 bibliographical review [1]

- No test can dominate all the others, but some of them are better on *common* situations
- Most papers compare tests based on statistical power comparisons: being able to detect deviations from normality with high probability
- All the mentioned packages run simulations to ensure that Type I errors are nominal i.e. rejection normality on $\mathrm{p} < 5\%$ ensures the probability of rejecting a true gaussian distribution incorrectly is 5%

Which one should I pick?

% of rejection of H_0 when true data generating process is $\mathrm{Unif}\,(0,1)$



Final remarks

- Kolmogorov-Smirnov and Shapiro-Wilk tests are low powered in common situations
  e.g. distributions that are unimodal and symmetric, but tails are slightly heavier than gaussian tails
- For low sample sizes $n < 100$:
- (adjusted) Jarque-Bera shows good performance for symmetric distributions
- Jin Zhang's version of traditional KS, AD and CVM are the most powerful tests otherwise
- Something to remember: there are many implementations of the same tests, but some packages lacks maintenance

THANKS!

Follow me on:

Curso-R: https://curso-r.com

My Github:https://github.com/azeloc