 pueh-leng.tan@fysik.su.se

Kolmogorov test

The Kolmogorov-Smirnov test of GOF is a simple and widely used test. It considers the maximal difference between the EDF and the CDF for all x: $D_{KS} = \sqrt(N) \cdot sup|\mathrm{EDF}(x) - \mathrm{CDF}(x)|$for all x. The p-value of this test may be computed in e.g. `scipy.stats.kstwobign.sf(D)` [1]

Anderson-Darling

The Anderson-Darling distribution is a more involved test, using the integrated quadratic distance between the EDF and CDF (weighted higher at the tails of the distribution).

Model

We will work with Gaussian distributions in this exercise. For the null hypothesis, H_0, the model is a standard Gaussian of mean, $\mu_0 = 0$ and standard deviation, $\sigma_0 = 1$.

Exercise

In this exercise, you will have a go at estimating the *power* of the K-S test for various alternative hypotheses, H_1.

To compute the power of a test for a certain alternative hypothesis H1 and significance $\alpha = 0.05$, you need to compute how many times you can exclude H0 if H1 is true. This is conveniently done with Monte Carlo:

- Find the distribution of the KS test statistic under the null hypothesis, H_0, by generating a large numbers of observations (e.g. 10000 sets of 50 numbers) under H_0 and computing the KS test p-value for each of the dataset.
- Find the value of the KS test statistic, T^*, at the desired confidence level, $\alpha = 0.05$.
- Given an alternative hypothesis, H_1, of a Gaussian distribution with mean $\mu \neq \mu_0$ and standard deviation $\sigma = \sigma_0$, find the distribution of the KS test statistic under the alternative hypothesis, H_1. Generate a large numbers of observations (e.g. 10000 sets of 50 numbers) under H_1 and compute the KS test p-value for each of the dataset.
- The power of the test is the fraction of p-values greater than T^*- i.e. the fraction of the p-values which will lead to the rejection of H_0 is rejected given that H_1 is true.
- Compute the power for 10 different alternative hypotheses, H_1's with μ's is between -1 and 1, and make a graph of the power as a function of μ.
- Repeat the above exercise, but change σ from 0.5 to 2 instead of μ. How does it compare?
- Bonus exercise: You can compute the Anderson-Darling test (with H_0 being a standard normal distribution) using `scipy.stats.anderson(x,'norm')`. Compare the power of the two tests, both for $\sigma \neq \sigma_0$ and $\mu \neq \mu_0$.

[1] Note that some packages, e.g. `scipy.stats.kstest` may include corrections to the asymptotic formula at low N. In this case, `python` switches over to using `2*scipy.stats.ksone(N).sf(D)`