Sample Task - Ratio Analysis Report

Xiaohui Li, Yuxin Ma

In order to make more further discussion, at first for the preliminary work, we modified our function to make it much more general, *RatioPower*, including the situation of homoscedasticity or heteroscedasticity for the variance level, minimum, maximum, median or mean for the ratio level, and uniform distribution or log-normal distribution for the another variance level.

1 Homoscedasticity

Firstly, we consider that σ_k has the same variance, which is a homoscedasticity case. We define $r_1 = \frac{Y_1}{\sigma_1}$ and $r_k = \frac{\sigma_k}{\sigma_1}$. And we generate σ_k by Y_1, r_1 and σ_1. Then we do a 100×20 times test where the scaling of r_1 varies from 0 to 2 and the scaling of r_k varies from 0 to 10. We get the plot as follow.

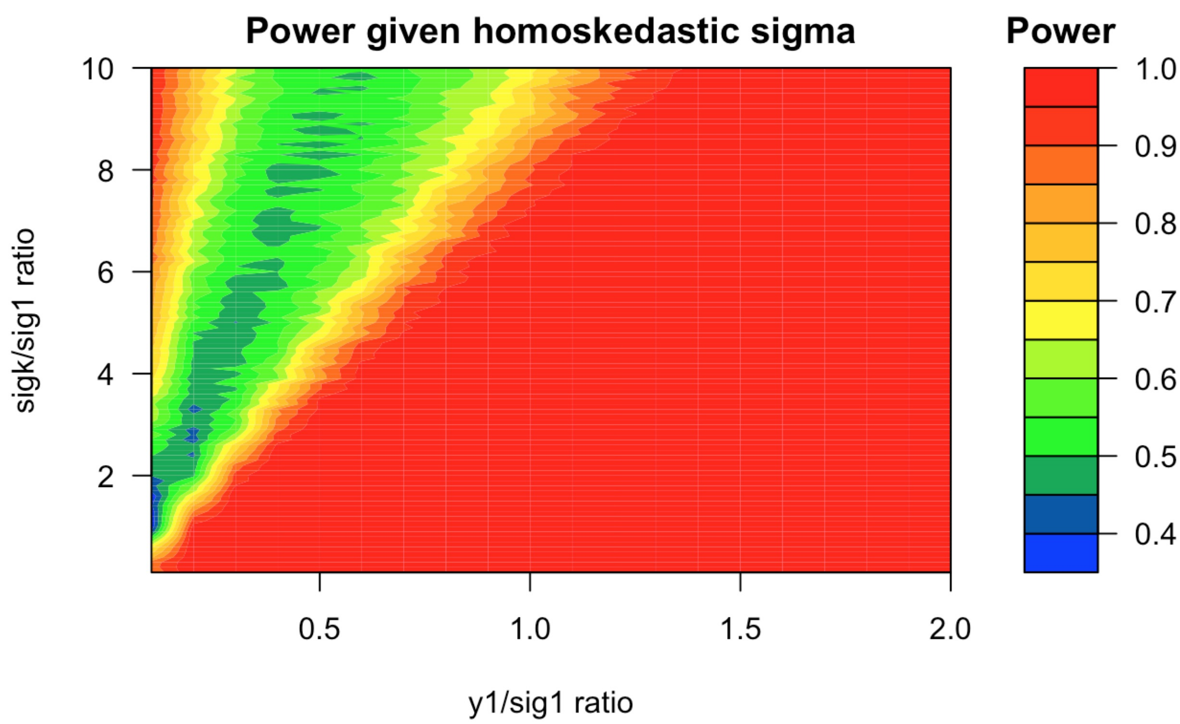

Figure 1.1: Homoscedasticity with r_1 varying from 0 to 2

We can see that there is an increasing linear trend as σ_1 and σ_k increase and the power converges to 1, the increasing trend starts from a ratio $\frac{r_k}{r_1}$ of around 20, which is "blue shade". Also we can see that if σ_1 takes very large values, the power will always be 1 and we will always reject to the null hypothesis, which does not have such significant results.

2 Heteroscedasticity

2.1 Random σ_k

This time, we assume the rest σ_ks are log-normal distributed and we generate them by the normal distribution with the mean equals to $\log(r_k \cdot Y_1/r_1)$ and take its exponential value. Also we do a 50×20×20 times test where the scaling of r_1 varies from 0 to 2 and the scaling of r_k varies from 0 to 10. We plot its power and get a plot having a similar trend as the homoscedasticity case does.

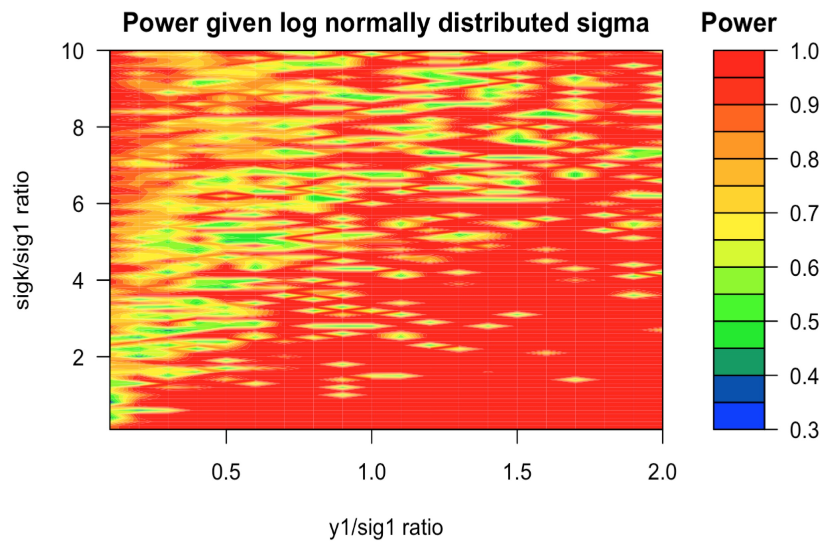

Figure 2.1: Heteroscedasticity

As it is depicted in the figure, there is still an "sparsely" increasing linear trend. But compared it with the homoscedasticity case, we can see that it is more discrete and fluctuates around the trend.

2.2 Random r_k

Unlike previous tests where we generate r_k by the iteration of equidistant increasing, this time we directly generate the random r_k by the uniform distribution or normal distribution to make our tests more general. And for the plot part, we choose a variety of ordinate form, including the minimum, maximum, median and mean of r_k.

2.2.1 Uniform Distribution - Minimum

We choose r_1 varies from 0 to 2. And we generate r_k by the uniform distribution ranging from 0 to 10. Then we choose the minimum of r_k to plot the power as Figure 2.2.

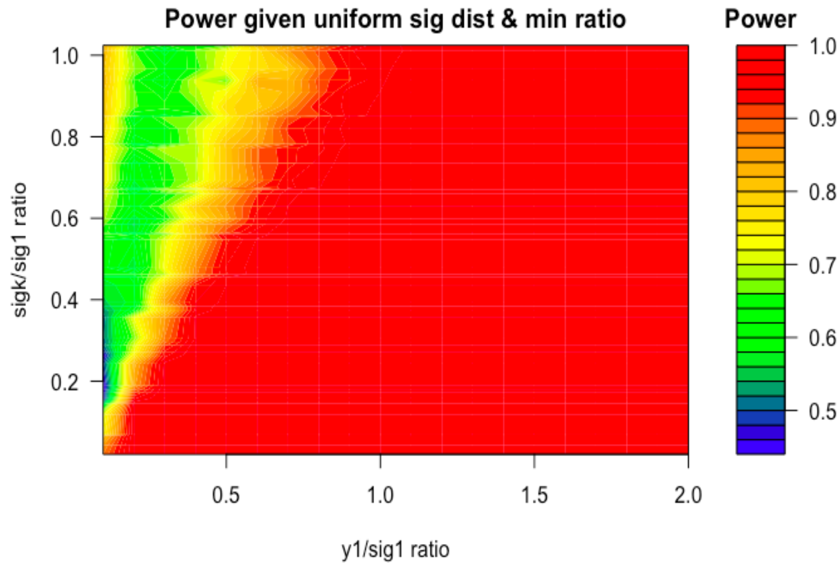

Figure 2.2: Uniform Distribution - Minimum

2.2.2 Normal Distribution - Minimum

We choose r_1 varies from 0 to 2. And we generate r_k by the normal distribution with its mean varying from 0 to 5 and variance equaling to 1 (default). Then we choose the minimum of r_k to plot the power as Figure 2.3.

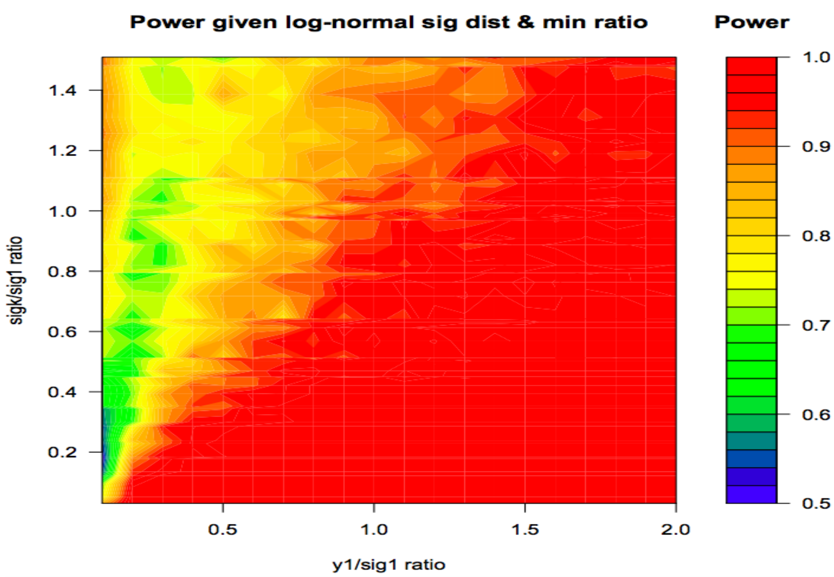

Figure 2.3: Normal Distribution - Minimum

2.2.3 Normal Distribution - Maximum

We choose r_1 varies from 0 to 2. And we generate r_k by the normal distribution with its mean varying from 0 to 5 and variance equaling to 1 (default). Then we choose the maximum of r_k to plot the power as Figure 2.4.

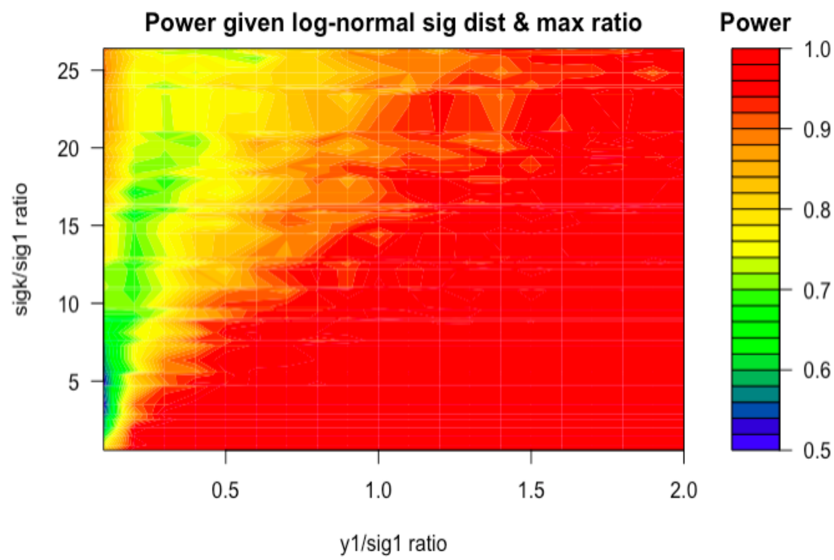

Figure 2.4: Normal Distribution - Maximum

2.2.4 Normal Distribution - Median

We choose r_1 varies from 0 to 2. And we generate r_k by the normal distribution with its mean varying from 0 to 5 and variance equaling to 1 (default). Then we choose the median of r_k to plot the power as Figure 2.5.

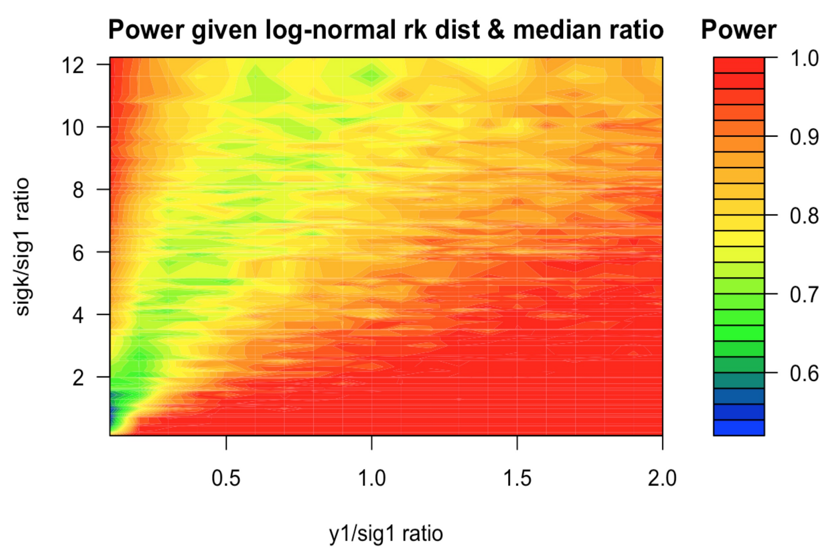

Figure 2.5: Normal Distribution - Median

2.2.5 Normal Distribution - Mean

We choose r_1 varies from 0 to 2. And we generate r_k by the normal distribution with its mean varying from 0 to 5 and variance equaling to 1 (default). Then we choose the mean of r_k to plot the power as Figure 2.6.

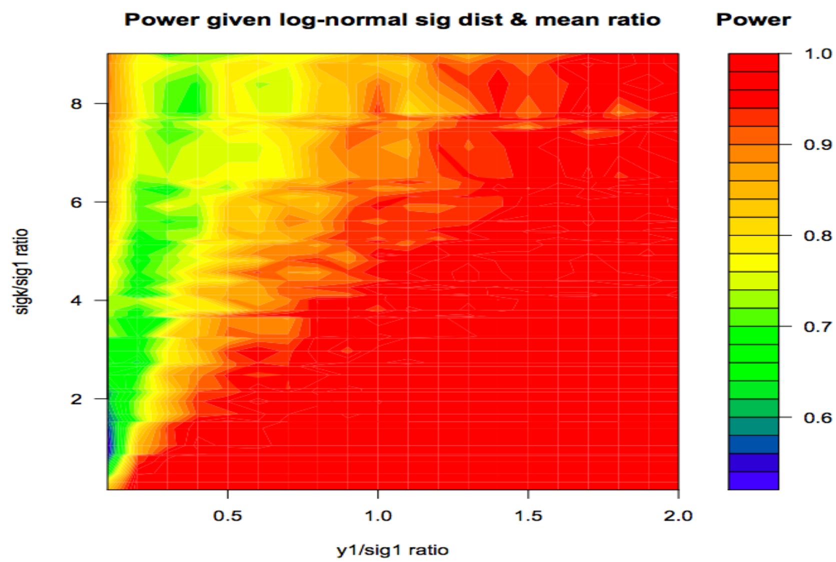

Figure 2.6: Normal Distribution - Mean

2.3 Conclusions of Random r_k

We can see that when we take the random r_k, for either uniform or normal distribution, it will always have a increasing trend like the homoscedasticity case and the heteroscedasticity case for random σ_k. But it is obvious that the power of the uniform distribution increases much more regularly and steadily, unlike the normal distribution, which will be more discrete. But in total, we can see that the general trend is evident as depicted in contour plots.

Specifically, we could find that given the $r_k \sim \log$-Normal assumption, the power increasing more than σ_k log-normal. While the normal-median case is a special case where the power increase slowly. Given fixed $r_k = \frac{\sigma_k}{\sigma_1}$, as $r_1 = \frac{\bar{Y_1}}{\sigma_1}$ increases, the power increase, while given fixed $r_1 = \frac{\bar{Y_1}}{\sigma_1}$, as $r_k = \frac{\sigma_k}{\sigma_1}$ increases, the power decreases. The power converges to 1 with increase of $r_1 = \frac{\bar{Y_1}}{\sigma_1}$.