Week6 Report

Xiaohui Li, Yuxin Ma

This week we make the further discussion towards the high-dimensional situation. Firstly, we do the scaling check for the ϵ to the power. Secondly, we keep the ϵ fixed to see the relationship between N and power. Then, we check the scaling behavior of Y_{max} in ϵ under the dense alternative. Finally, we generate the plots of ϵ versus K.

1 Scaling check for ϵ

Firstly, we do the scaling test for the ϵ to check whether the power pattern exists differences when the scaling of ϵ varies. We take ϵ, 1.5ϵ and 2ϵ into consideration.

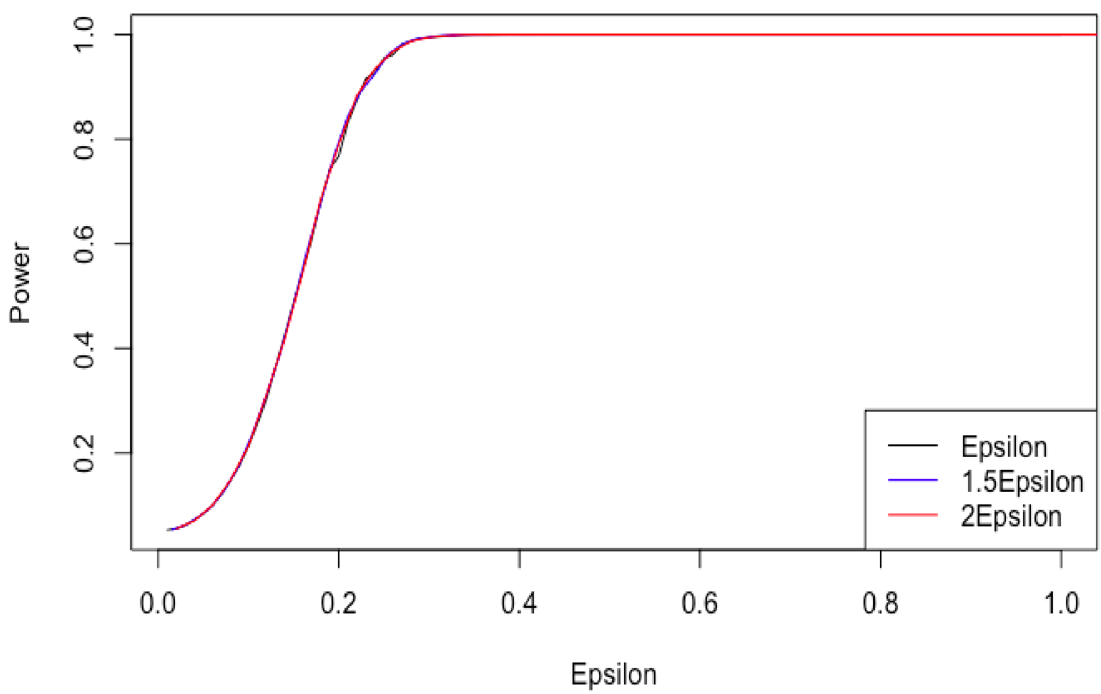

Figure 1.1: Scaling check for ϵ

From the plot we can see that all three of them have similar pattern and they are really closed to each other. So we can conclude that when ϵ scales differently, the power pattern keeps fixed.

2 Relationship between N and power

Secondly, we check the relationship between N and power where we control N to increase from 1000 to 100,000. Also, we do the repetitive test for minimum, median and mean statistics for p_combined value.

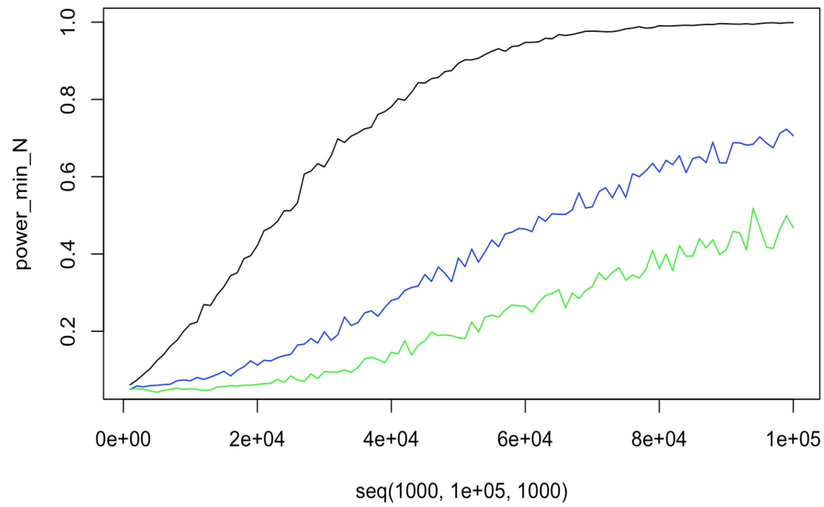

Figure 2.1: Relationship between N and power

We can see that as N increases, the power will increase as well. And also, the minimum statistics, the black curve, performs best while the median and mean (blue and green ones) do not behave such well.

3 Scaling behavior of Y_{max}

Next, we check the scaling behavior of Y_{max} in ϵ under the dense alternative. Since last week we found that for the minimum p_combined, the critical epsilon of K is a linear curve, and the minimum p_combined corresponds to the maximum of Y, this week we do the scaling behavior of Y_{max} to check whether Y_{max} fits in the $\sqrt{K}$ trend and then makes the critical epsilon curve linear under the minimum p_combined.

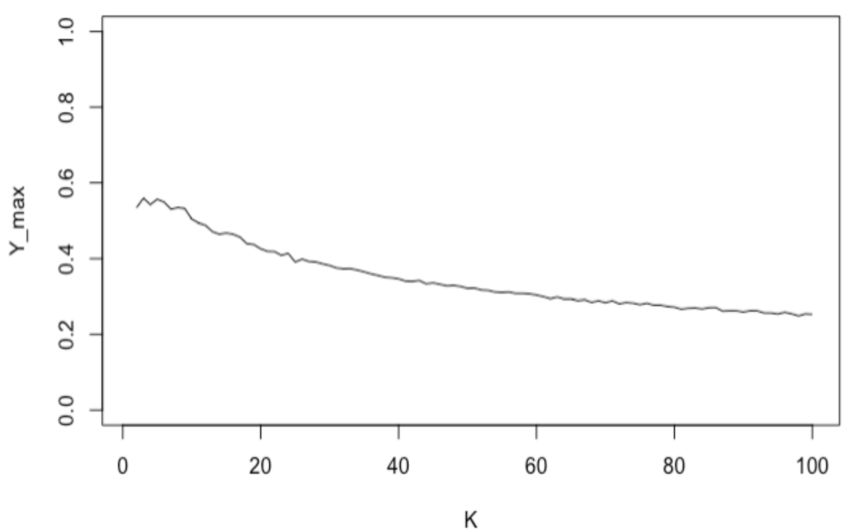

Figure 3.1: Y_{max} versus K

Theoretically, We think that it should fit the $\epsilon/\sqrt{K}$ trend and as we take value 1 for the ϵ, Y_{max} should fit $1/\sqrt{K}$. But the result does not look like that. No coefficients for $1/\sqrt{K}$ and R^2 equals to 0.

```
Call:
lm(formula = Y_max[2:100] ~ 1/sqrt(K_max))

Residuals:
     Min       1Q   Median       3Q      Max 
-0.09884 -0.06798 -0.02672  0.04372  0.21127 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.348042   0.008712   39.95   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08668 on 98 degrees of freedom

$r.squared
[1] 0

$adj.r.squared
[1] 0
```

Figure 3.2: Fit $1/\sqrt{K}$ to Y_{max}

Meanwhile, we try other form of K to fit the Y_{max}, including $\sqrt{K}$, $1/K$ and $\ln K$ etc., and finally we find that $\ln K$ has a significant fitting to the Y_{max} where the coefficient is minus and the R^2 equals to around 0.96.

```
Call:
lm(formula = Y_max[2:100] ~ K_max_ln)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.116593 -0.004604 -0.002592  0.005529  0.035532 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.706244   0.007686   91.88   <2e-16 ***
K_max_ln    -0.097600   0.002038  -47.89   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01728 on 97 degrees of freedom
Multiple R-squared:  0.9594,	Adjusted R-squared:  0.959 
F-statistic:  2294 on 1 and 97 DF,  p-value: < 2.2e-16
```

Figure 3.3: Fit $\ln K$ to Y_{max}

It is not expected according to our intuition, so the result is really surprising.

4 Critical ϵ for K

Last week we check the critical ϵ for K as the p_combined take the minimum statistics. This week for further discussion, we also take the median and mean statistics for p_combined to check their trends.

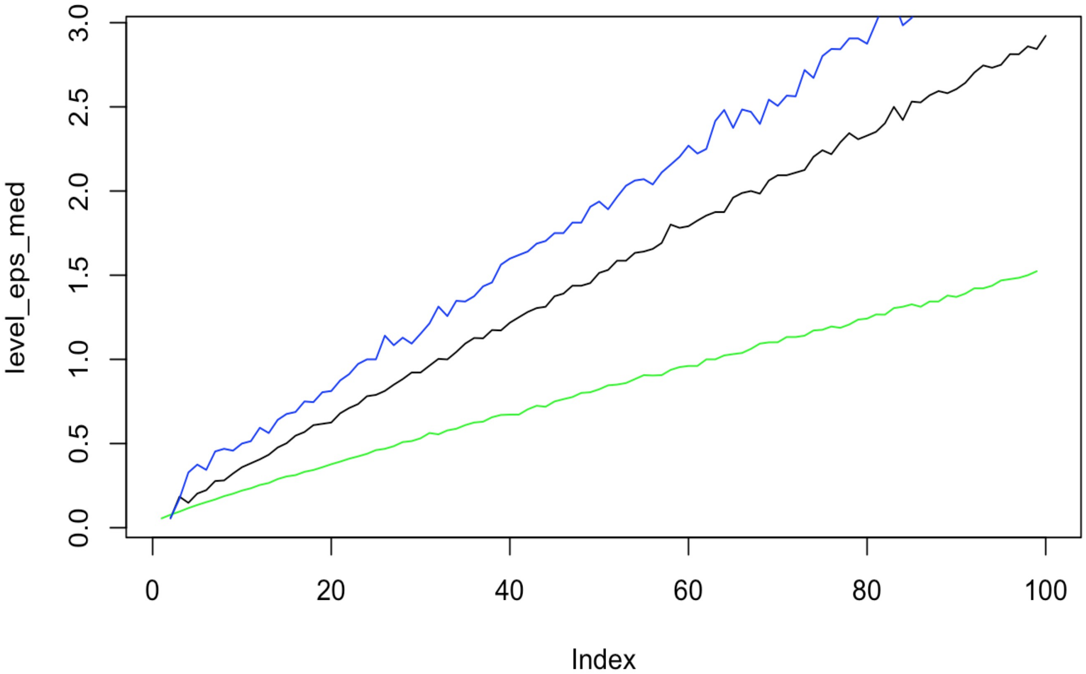

Figure 4.1: Critical ϵ for K

We can see that all three curves, green to the minimum, black to the median, blue to the mean, follow an obvious linear trend and the minimum is the lowest curve, which corresponds to our intuition.