Sample Task - Treatment Size Analysis Report

Xiaohui Li, Yuxin Ma

For further discussion, we do tests to figure out the impact of the treatment size K to our analysis. Firstly, we try to find out the general pattern for the power as $Y_1(r_1)$ and K increase when r_k is fixed. Secondly, we do the research to find that as K increases, the variation of the ratio r_1. We make the new power equal to 80% of the origin power to find the variation of its respective $Y_1(r_1)$, and also, transform the form of K like $\log(K)$ and $\sqrt{K}$.

1 General Pattern

We let K respectively equal to 10, 50 and 100 to find that as r_1 increases from 0 to 1 as σ is fixed. We can find that it always looks like an "S" shape but as K increases, it moves rightward and for the power equaling to 0.8, we can see the corresponding σ_1 becoming larger.

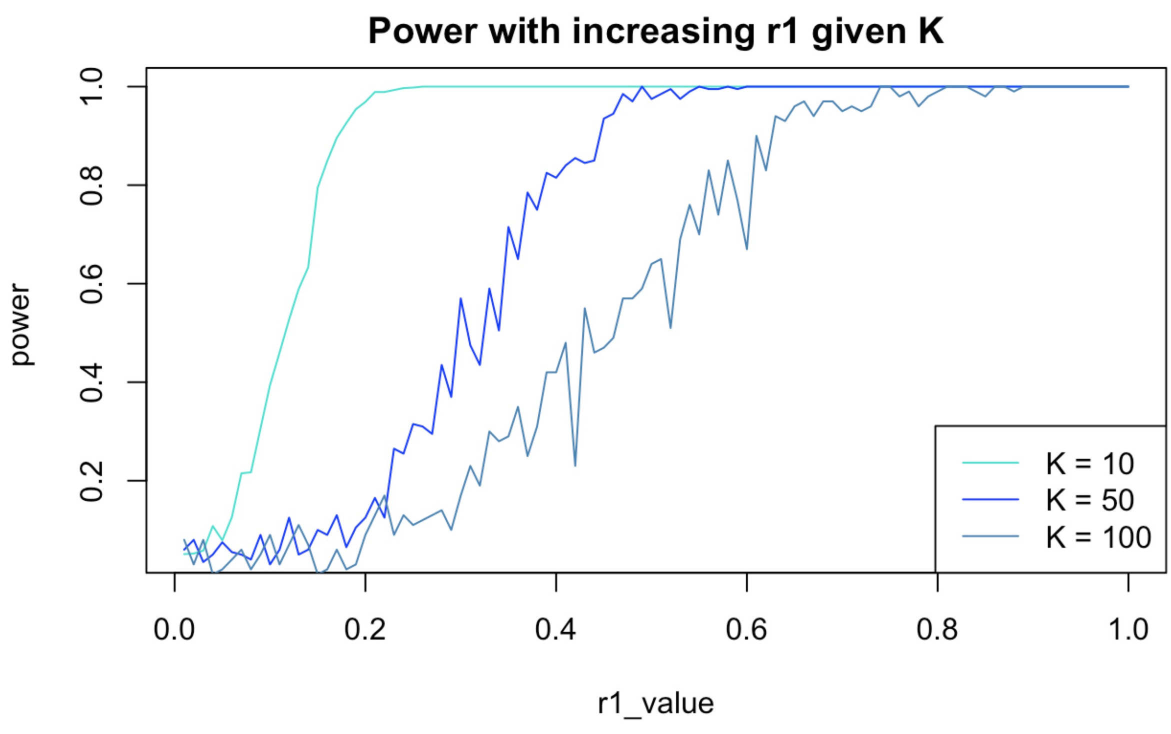

Figure 1.1: General Pattern

We could find that as K size increases, the power increasing trend gets weak, i.e, it increases more slowly. This suggest a negative effect of K size on power.

2 Power with K size

From the general we grasp the general outline for power, further we analyze the specific numerical effect from treatment size K to power. We control for the r_1 and r_k to partial out the exact treatment size effect. We fix the r_k to be 1 and list two r_1 cases.

2.1 $r_1 = 1$

We take the r_1 equals to 1, and see the size K effects upon power from 2 to 100.

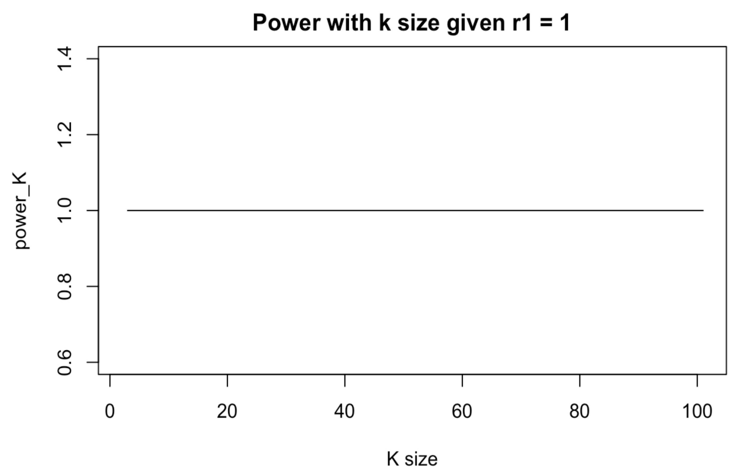

Figure 2.1: Power for size K given r_1 equals to 1

Up till size K to 100, still the power is 1, there is no negative effect of K size. This may be attributed to the limit effect of K neutralized by strong force of r_1.

2.2 $r_1 = 0.5$

We reduce the r_1 effect by changing it to 0.5 and plot the power against K.

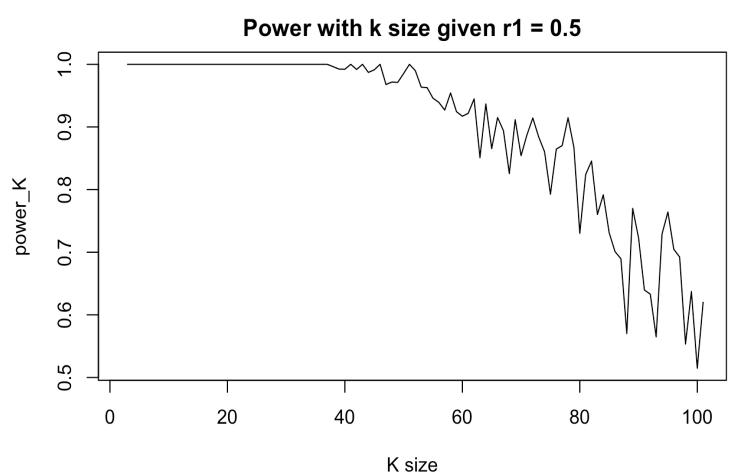

Figure 2.2: Power for size K given r_1 equals to 0.5

The power shows a "sudden" downward sloping trend after size over around 40. And it decreases to 0.

3 Critical Ratio for 80% Power

3.1 Critical ratio calculation mechanism and algorithm

From the general pattern and previous contour plot, we could get that given fixed treatment size, as r_1 increases, the power will firstly decreases and then increases.

As K increases from 0 to 100, we get the critical ratio r_1 by summing the r_1 if it is qualified where the power is either less than 0.8 or the power goes downward. We can find that as K increases, the critical ratio increases gradually but kind of fluctuating as well. This will help us to find the appropriate ratio for σ_1 when given K since only when power equals to 0.8 is somehow significant for our research.

To make the calculation more efficient, we apply the binary tree algorithm to derive the approximate $\Theta(\log n)$ running time and get the result

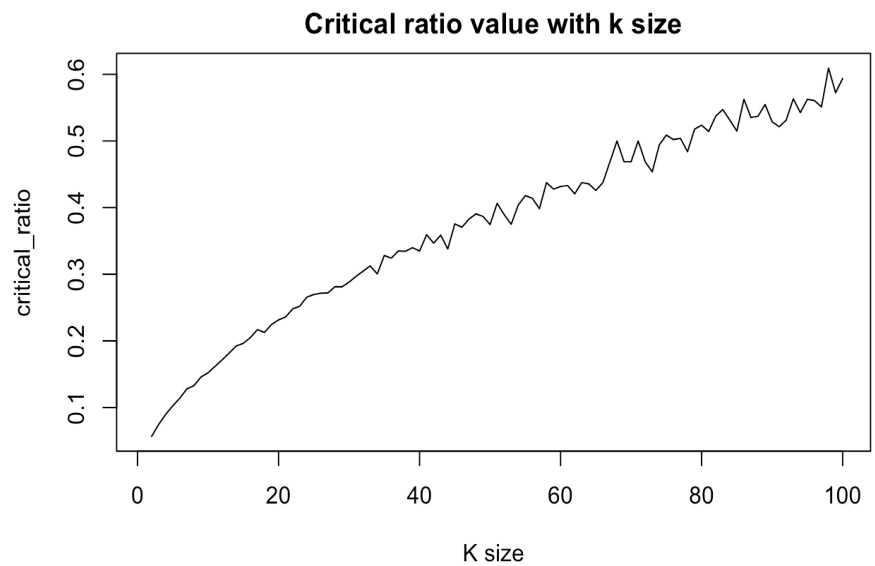

Figure 3.1: Critical Ratio for Origin K

The 80% power critical ratio increase with the treatment size, it increases at rate lower than the constant linear trend, so it may suggest a squared or logarithm increasing rate.

3.2 $\sqrt{K}$

Test for the critical ratio with $\sqrt{k}$.

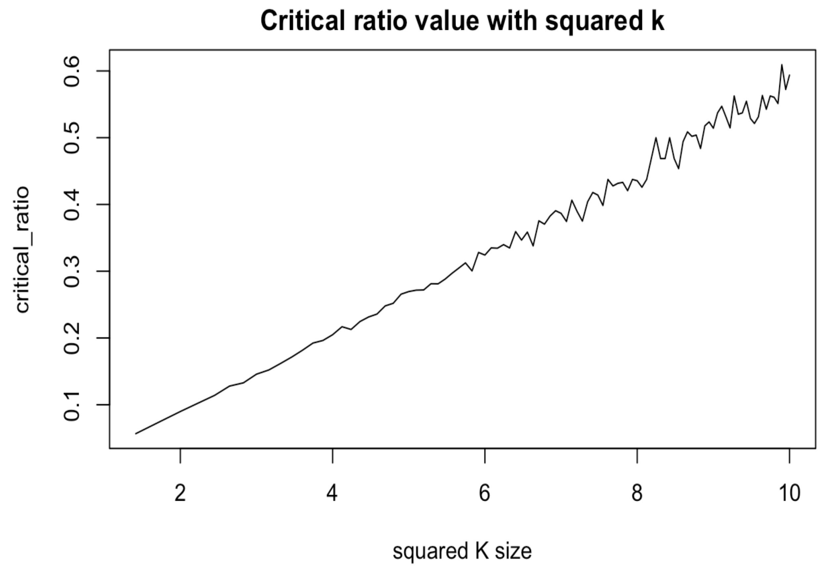

Figure 3.2: Critical Ratio for $\sqrt{K}$

It shows a clear sigh of linear relationship with $\sqrt{K}$.

3.3 $\log K$

Test for the critical ratio with $\log k$.

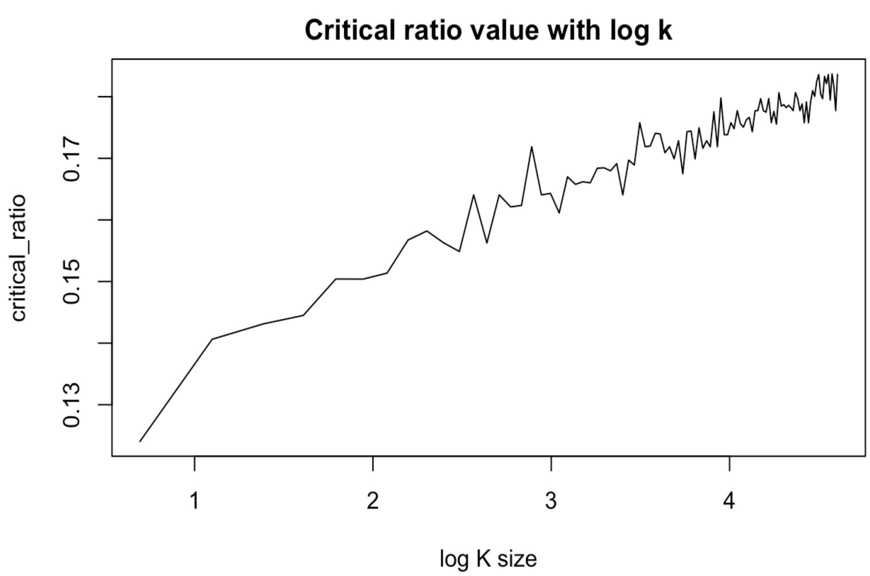

Figure 3.3: Critical Ratio for $\log k$

The $\log K$ case looks more like a linear relation, indicating $\log K$ increasing rate.

3.4 Regression

We further apply the linear regression to compare $\sqrt{K}$ and $\log K$.

```
Call:
lm(formula = critical_ratio ~ logk)

Residuals:
     Min       1Q   Median       3Q      Max
-0.05399 -0.03174 -0.01112  0.02862  0.15321

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.206102   0.017507  -11.77   <2e-16 ***
logk         0.158025   0.004642   34.05   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03936 on 97 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9228,    Adjusted R-squared:  0.922
F-statistic:  1159 on 1 and 97 DF,  p-value: < 2.2e-16


Call:
lm(formula = critical_ratio ~ sqrtk)

Residuals:
      Min        1Q    Median        3Q       Max
-0.030777 -0.005996  0.000198  0.006513  0.042270

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0426060  0.0042328  -10.07   <2e-16 ***
sqrtk        0.0615901  0.0005927  103.91   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01336 on 97 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9911,    Adjusted R-squared:  0.991
F-statistic: 1.08e+04 on 1 and 97 DF,  p-value: < 2.2e-16
```

Figure 3.4: Regression Summary

Note the $\sqrt{K}$'s β coefficient is more statistically significant and larger, while the r^2 of log is also larger than $\log k$, this supports the reasoning that critical ratio is logarithm increase. In particular, the r squared for $\sqrt{K}$ model is high as 0.9911, indicating a potential theoritical relationship:

3.5 Conclusions

Treatment size K has a negative effect upon power, and the 80% power critical ratio increase with the treatment size K at the rate of squared K.