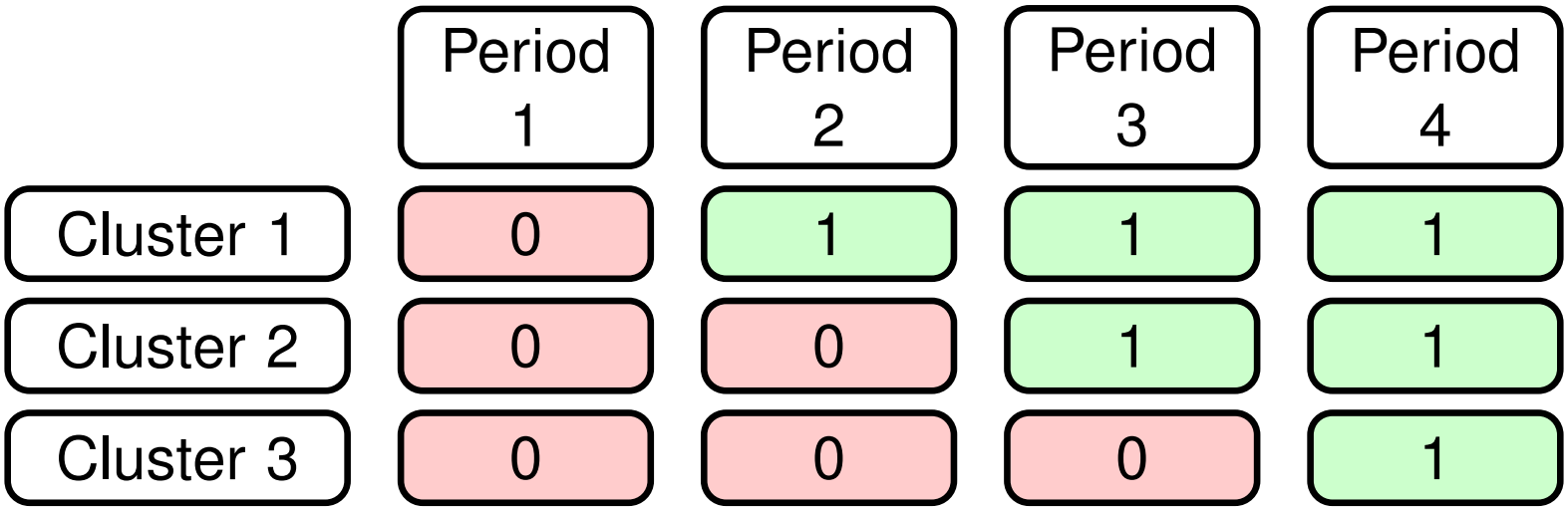
Information content of cluster-periods in stepped wedge trials

Jessica Kasza
Andrew Forbes

jessica.kasza@monash.edu

Department of Epidemiology and Preventive Medicine
Monash University

The usual stepped wedge

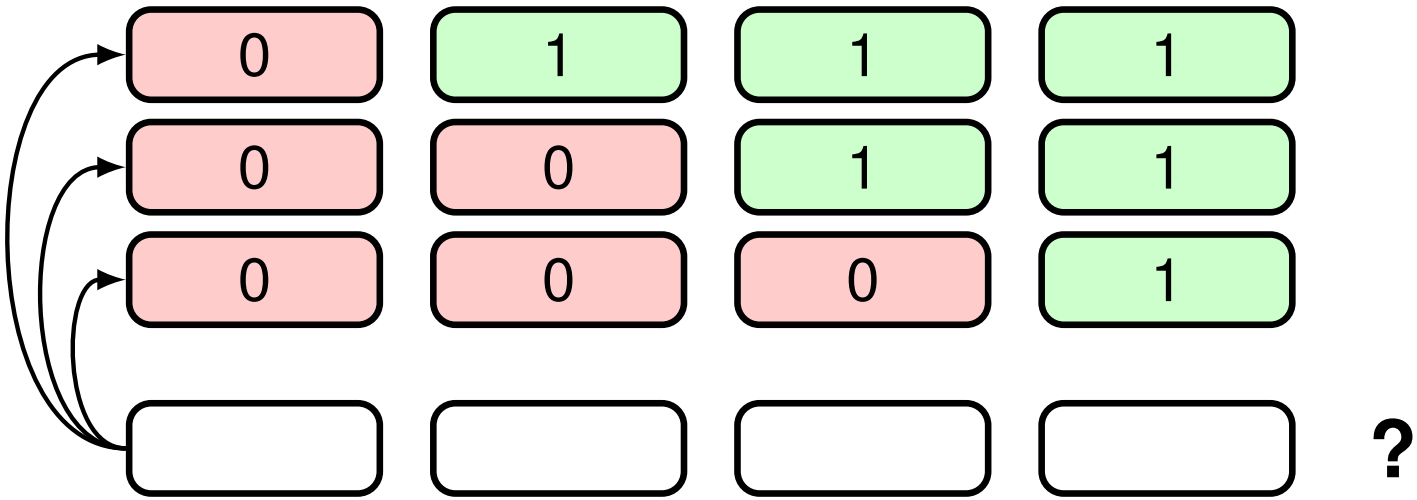


K clusters; T periods; $K \times T$ cluster-period ***cells***
m subjects per cell

Optimal designs: where to allocate?

Optimal designs seek to allocate a *fixed number of subjects* in such a way so as to maximise power.

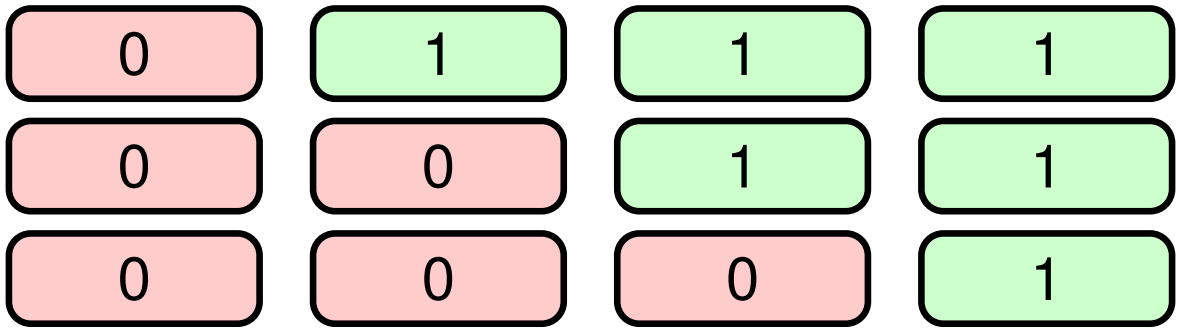
- Example: to which treatment sequence should a cluster be assigned?



Minimal designs: which to measure?

Minimal designs seek to *reduce the total number of subjects* with a minimal decrease in power.

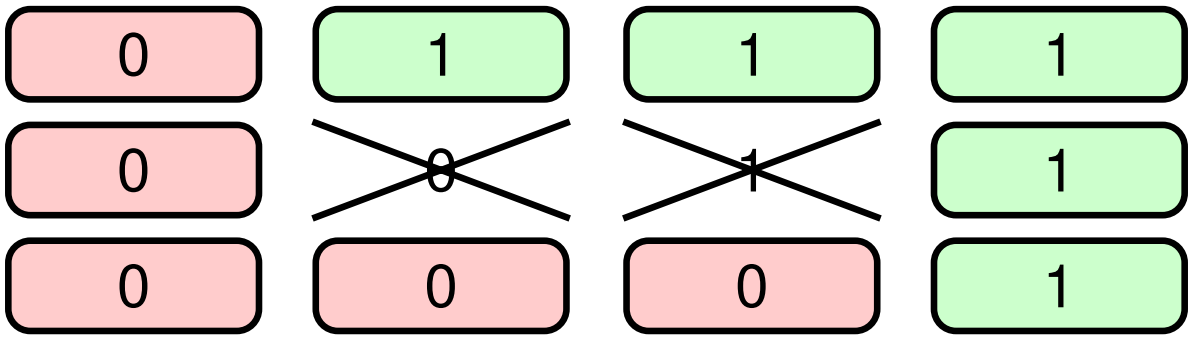
- Which cells can be excluded?



Minimal designs: which to measure?

Minimal designs seek to *reduce the total number of subjects* with a minimal decrease in power.

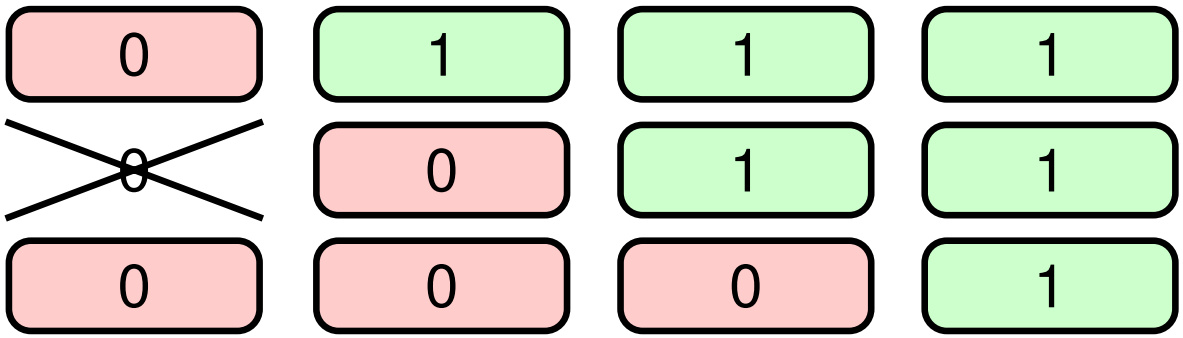
- Which cells can be excluded?



Minimal designs: which to measure?

Minimal designs seek to *reduce the total number of subjects* with a minimal decrease in power.

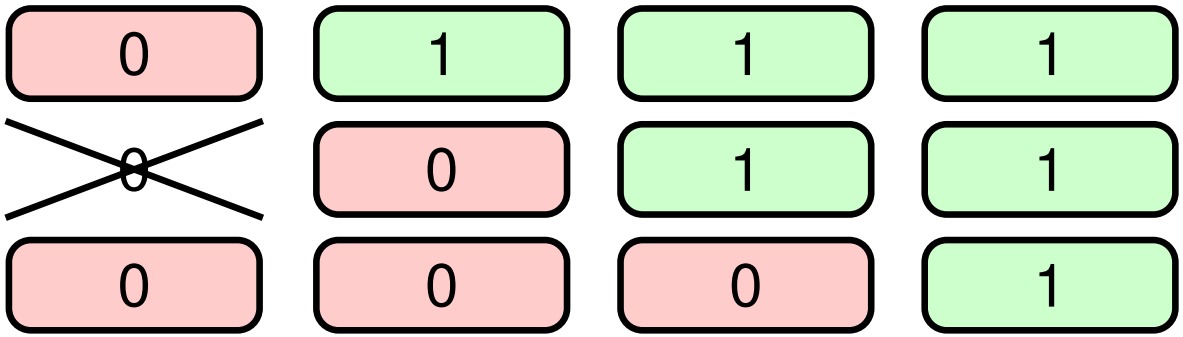
- Which cells can be excluded?



Minimal designs: which to measure?

Minimal designs seek to *reduce the total number of subjects* with a minimal decrease in power.

- Which cells can be excluded?



My focus is on minimal designs:
aim to reduce trial costs by omitting cells.

Which cells can be omitted with the smallest acceptable decrease in power (or precision)?

Models for continuous outcomes

Y_{ikt}: outcome for subject $i = 1, \ldots, m$, in cluster $k = 1, \ldots, K$, during period $t = 1, \ldots, T$

X_{kt}: treatment indicator for cluster k in period t

Hussey and Hughes ('standard' model):

$$Y_{ikt} = \beta_t + X_{kt}\theta + C_k + \epsilon_{ikt}, \quad C_k \sim N(0, \tau^2), \quad \epsilon_{ikt} \sim N(0, \sigma_\epsilon^2)$$

$$\text{Intra-cluster correlation: } \rho_0 = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

$\hat{\theta}$ the weighted least squares estimator of the treatment effect θ.

Models for continuous outcomes

Y_{ikt}: outcome for subject $i = 1, \ldots, m$, in cluster $k = 1, \ldots, K$, during period $t = 1, \ldots, T$

X_{kt}: treatment indicator for cluster k in period t

Hussey and Hughes ('standard' model):

$$Y_{ikt} = \beta_t + X_{kt}\theta + C_k + \epsilon_{ikt}, \quad C_k \sim N(0, \tau^2), \quad \epsilon_{ikt} \sim N(0, \sigma_\epsilon^2)$$

$$\text{Intra-cluster correlation: } \rho_0 = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

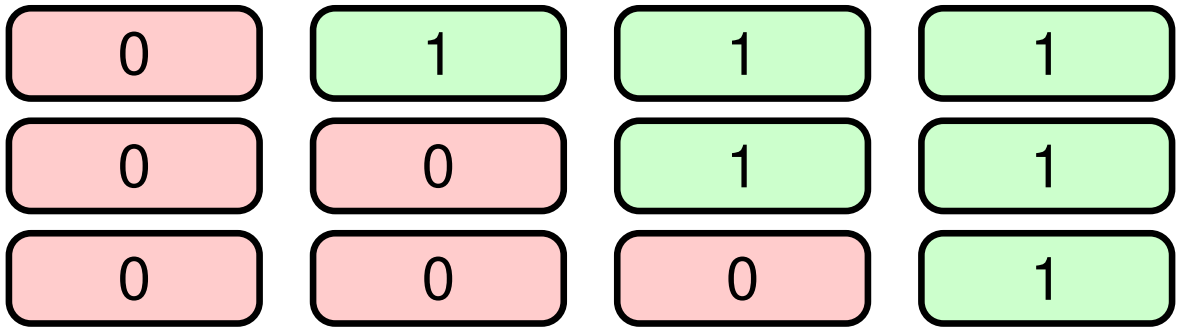
$\hat{\theta}$ the weighted least squares estimator of the treatment effect θ.

- $var(\hat{\theta})$ of interest: used in sample size calculations.

How much does $var(\hat{\theta})$ increase if observations from a given cell are omitted?

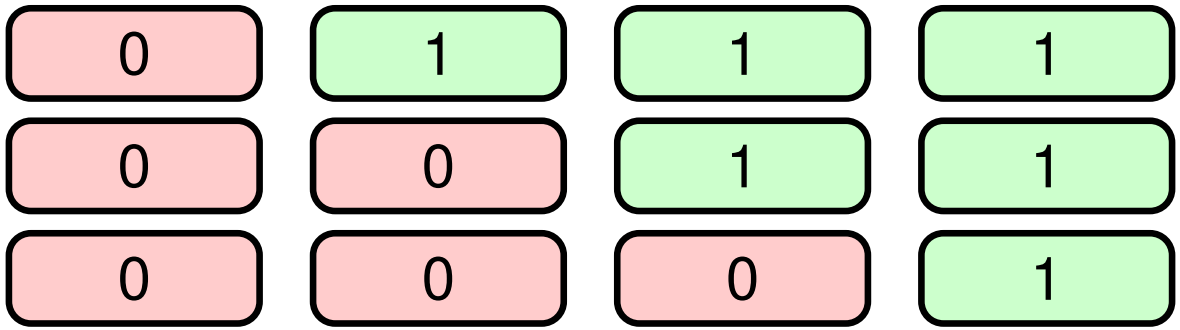
Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:

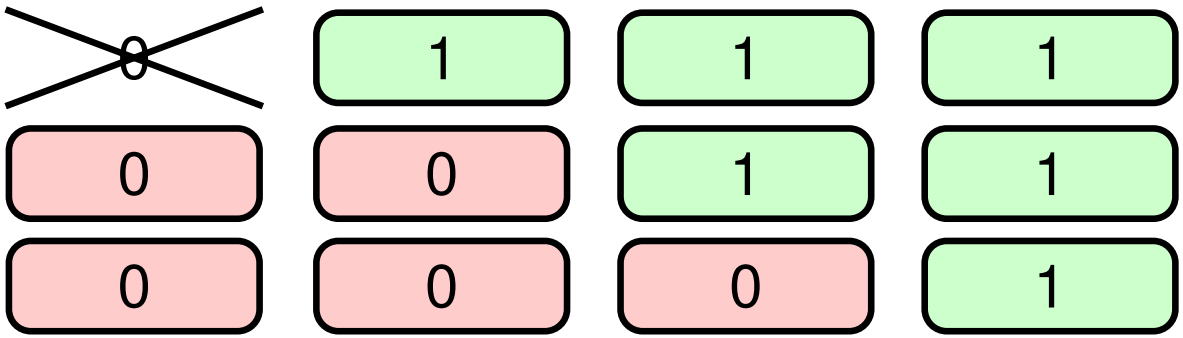


Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:

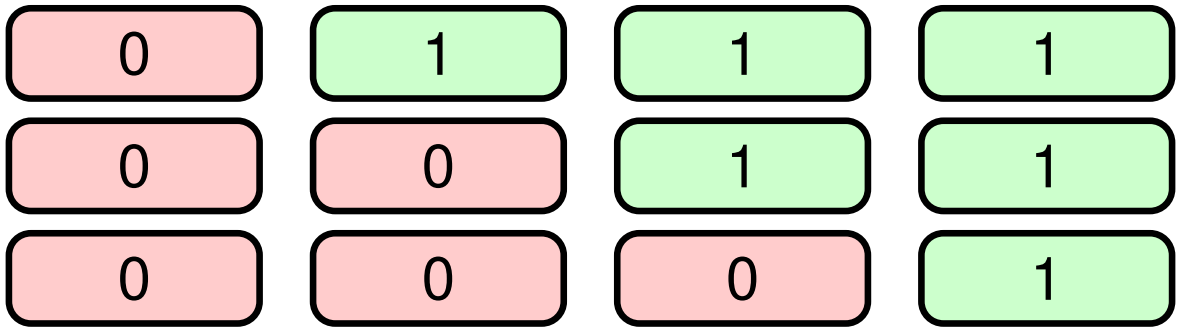


Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:

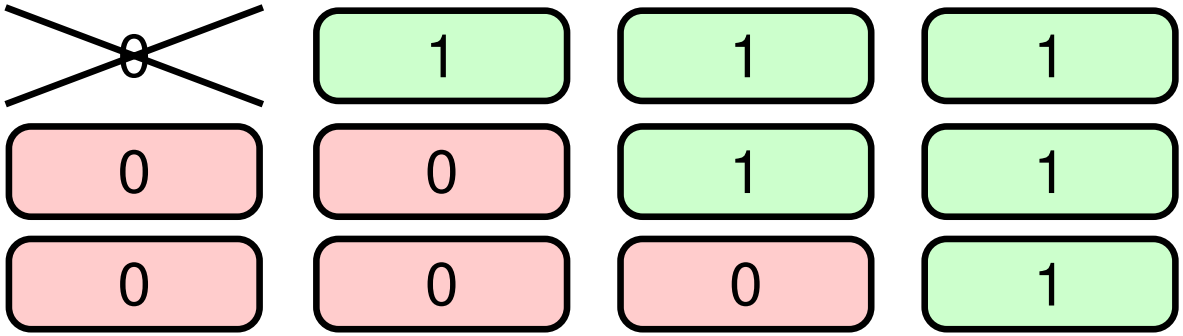


Information content of each cell

Calculate $var(\hat{\theta})$ given the complete design:



Calculate $var(\hat{\theta})_{[kt]}$ from the incomplete design, omitting period t of cluster k:



Information content of cell (k,t) defined as

$$IC(k,t) = var(\hat{\theta})_{[kt]} / var(\hat{\theta})$$

Information content of cells: theoretical results

Can obtain a closed-form expression for $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$ for the Hussey and Hughes model (and for related models)[1]

- I'll spare you the gory details!

[1] Such as those considered in Hooper et al (Stats in Med, 2016), Girling and Hemming (Stats in Med, 2016): an analytical expression for $IC(k,t)$ is available whenever the inverse of the covariance matrix of observations from a cluster has a closed form.

Information content of cells: theoretical results

Can obtain a closed-form expression for $IC(k,t) = var(\hat{\theta})_{[kt]}/var(\hat{\theta})$ for the Hussey and Hughes model (and for related models)[1]

- I'll spare you the gory details!

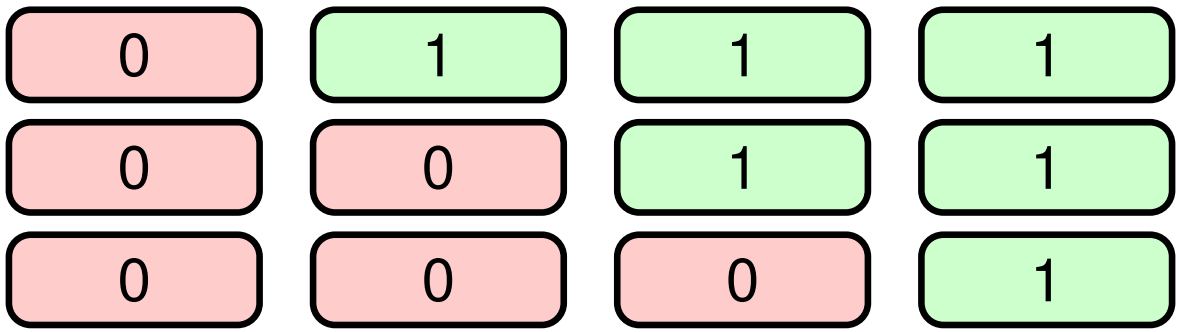
For this (and related) models, $IC(k,t)$ has the following properties:

Centrosymmetry: $IC(k,t) = IC(K+1-k, T+1-t)$

Information-free cells: $IC\left(\frac{K+1}{2}, 1\right) = IC\left(\frac{K+1}{2}, T\right) = 1$

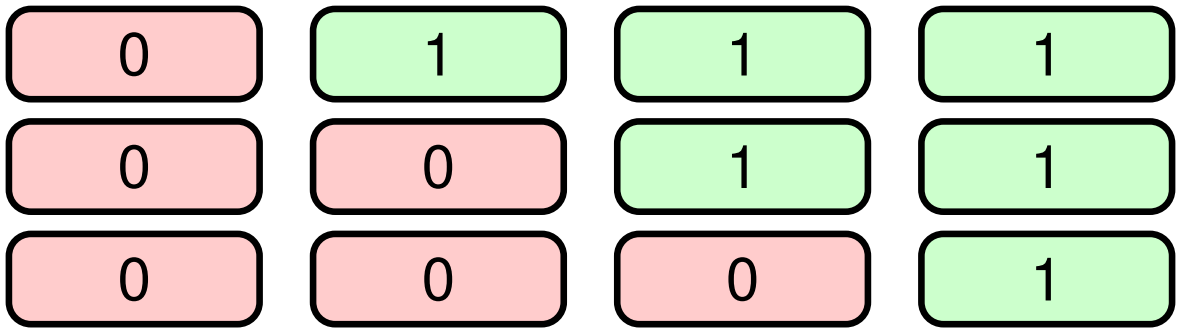
[1] Such as those considered in Hooper et al (Stats in Med, 2016), Girling and Hemming (Stats in Med, 2016): an analytical expression for $IC(k,t)$ is available whenever the inverse of the covariance matrix of observations from a cluster has a closed form.

Information content of cells for $K = 3$ and $T = 4$



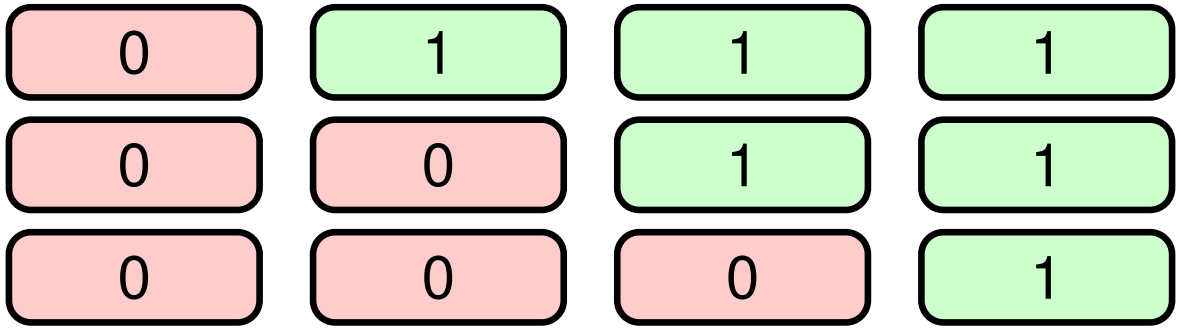
$$IC\left(\frac{K+1}{2}, 1\right) = IC\left(\frac{K+1}{2}, T\right) = 1, \quad IC(k,t) = IC(K+1-k, T+1-t)$$

Information content of cells for $K = 3$ and $T = 4$



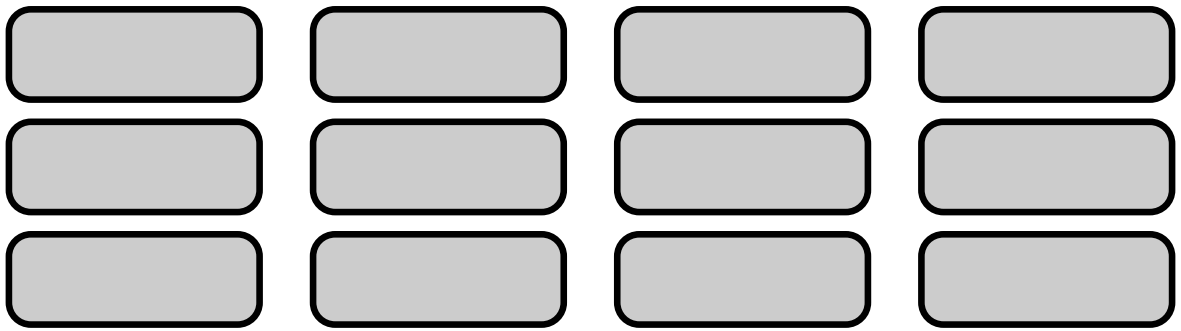
$$IC\,(2,1) = IC\,(2,4) = 1, \quad IC(k,t) = IC(4-k,5-t)$$

Information content of cells for $K = 3$ and $T = 4$

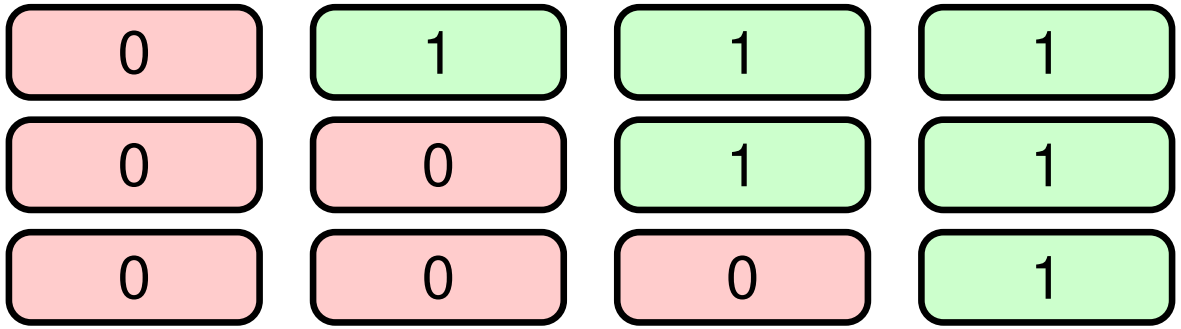


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

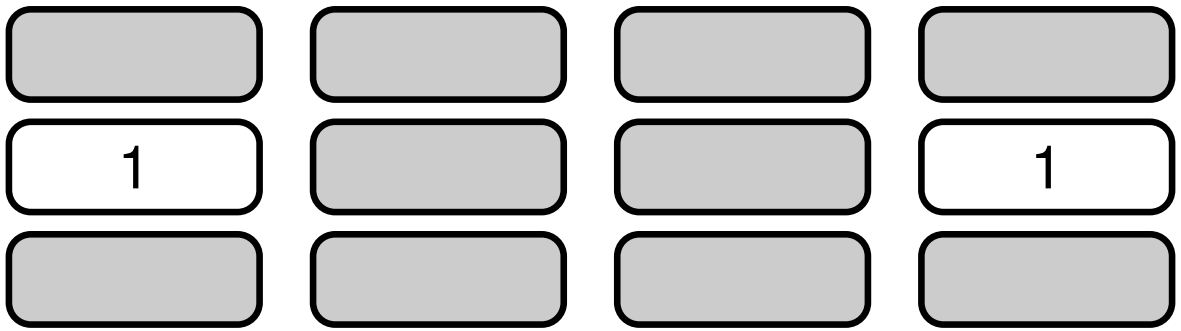


Information content of cells for $K = 3$ and $T = 4$

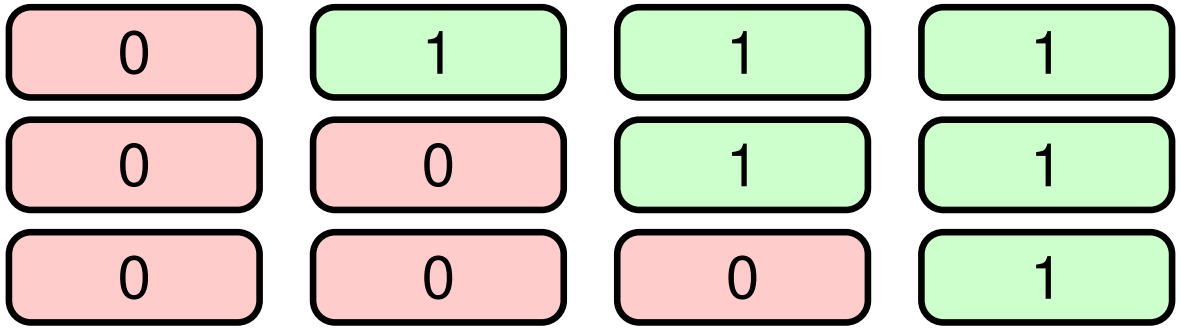


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

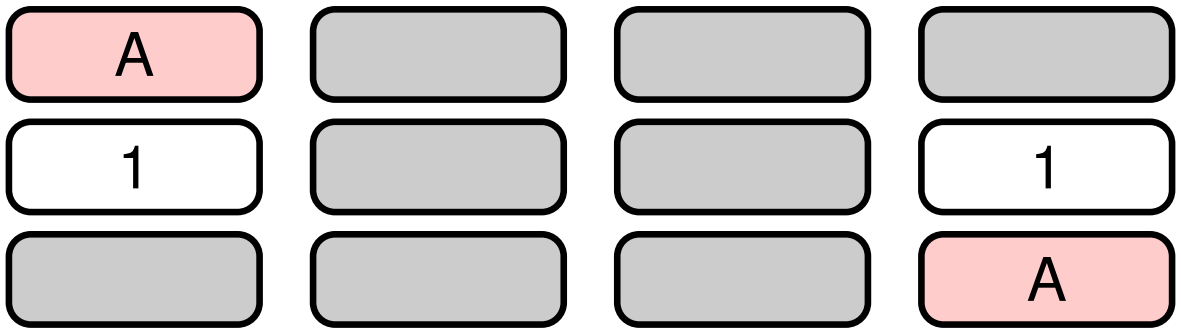


Information content of cells for $K = 3$ and $T = 4$

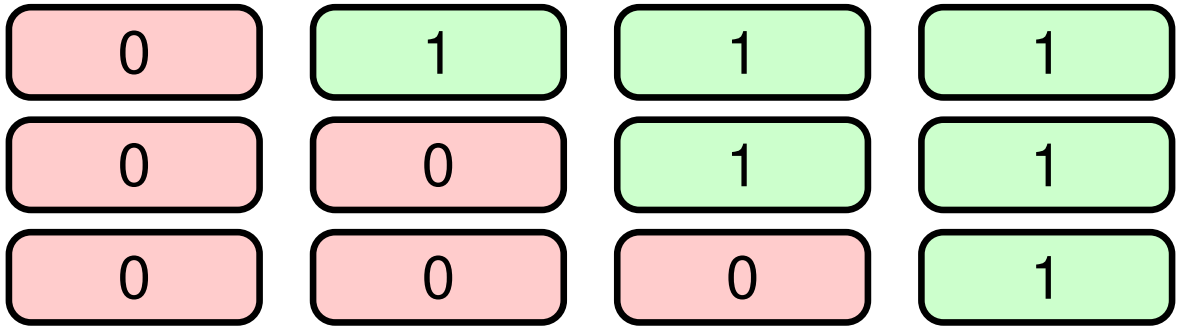


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

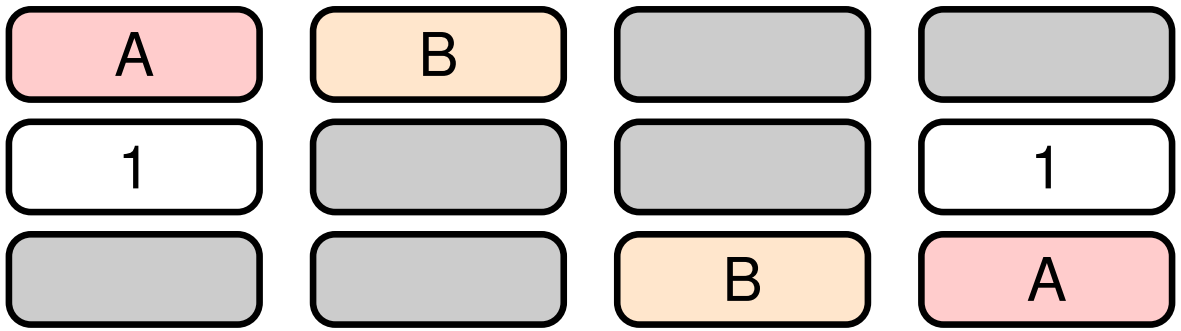


Information content of cells for $K = 3$ and $T = 4$

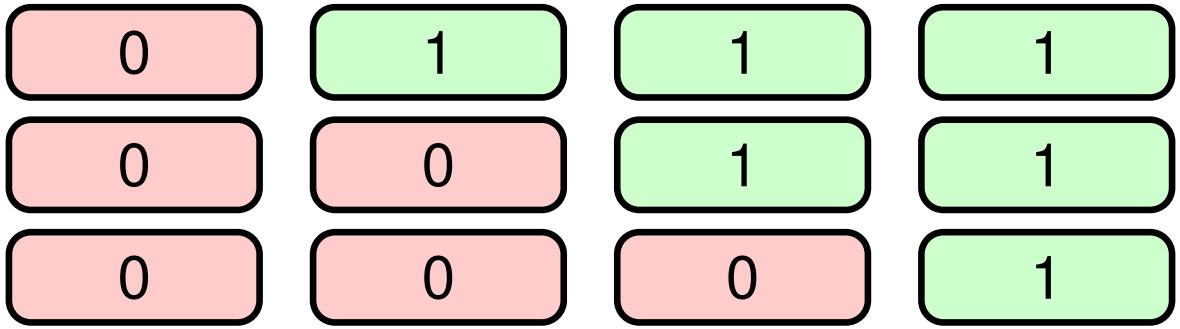


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

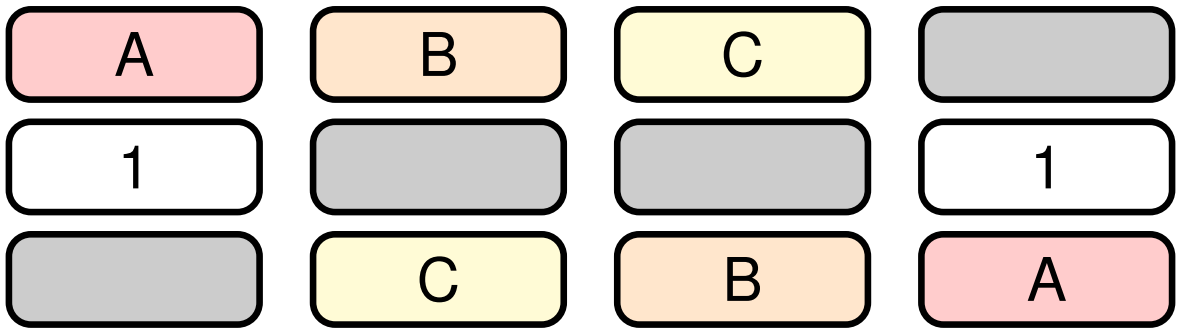


Information content of cells for $K = 3$ and $T = 4$

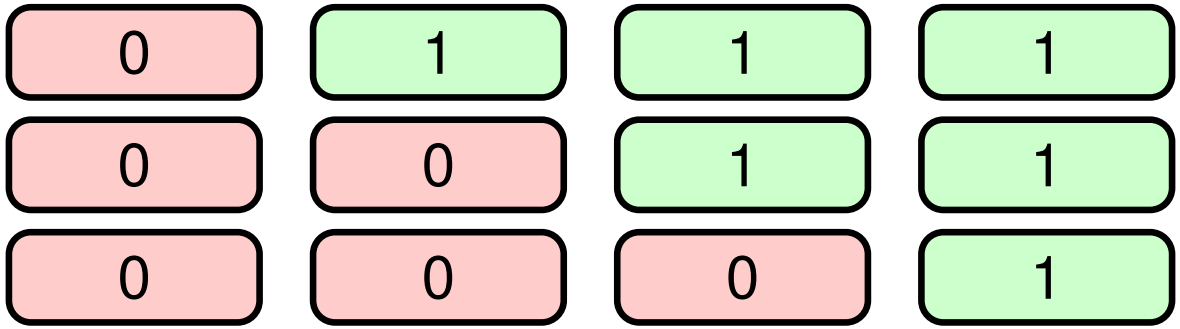


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

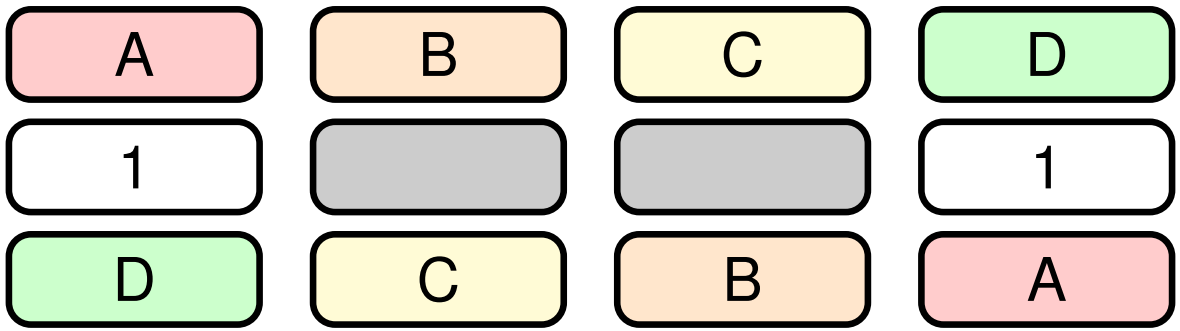


Information content of cells for $K = 3$ and $T = 4$

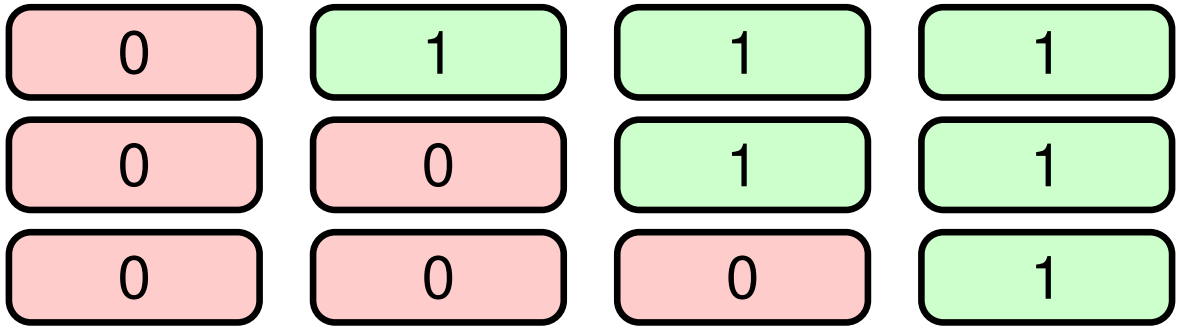


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

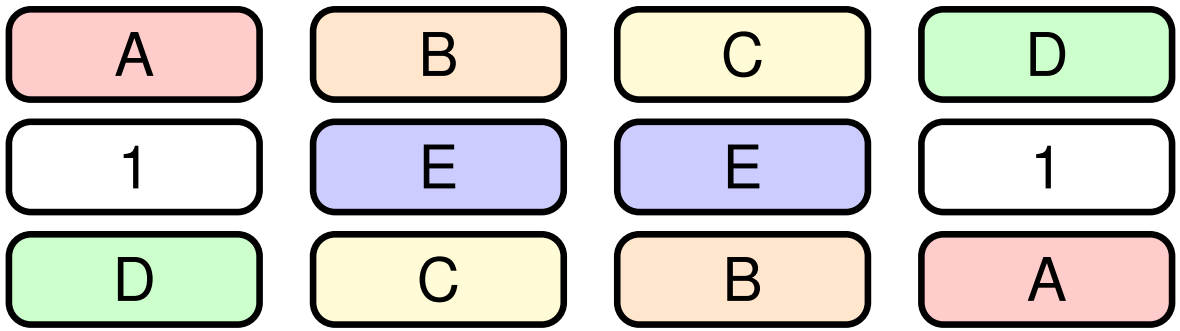


Information content of cells for $K = 3$ and $T = 4$

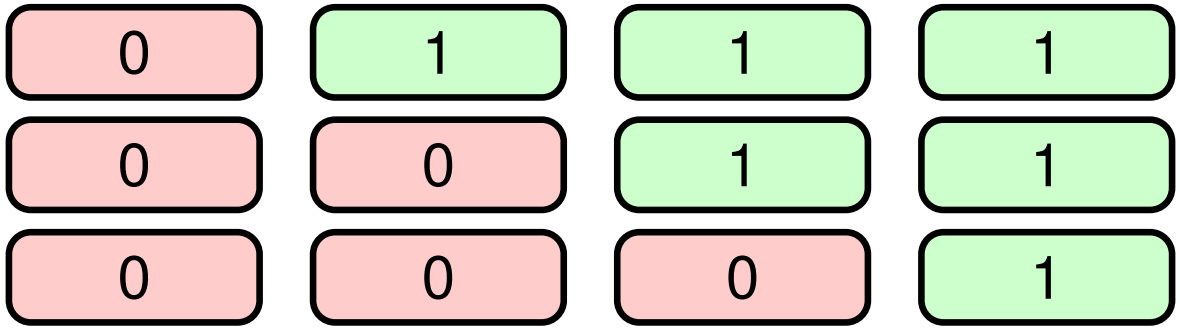


$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:

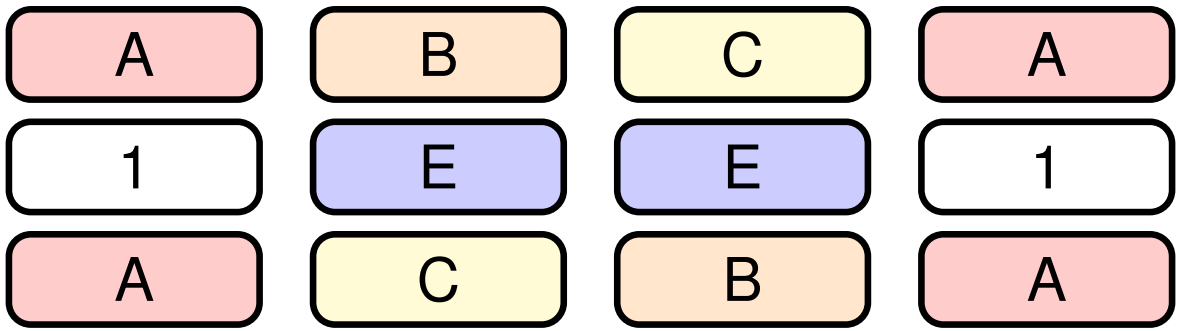


Information content of cells for $K = 3$ and $T = 4$



$$IC(2,1) = IC(2,4) = 1, \quad IC(k,t) = IC(4-k, 5-t)$$

$IC(k,t)$ for all k, t:



Particular examples

Hussey and Hughes model:

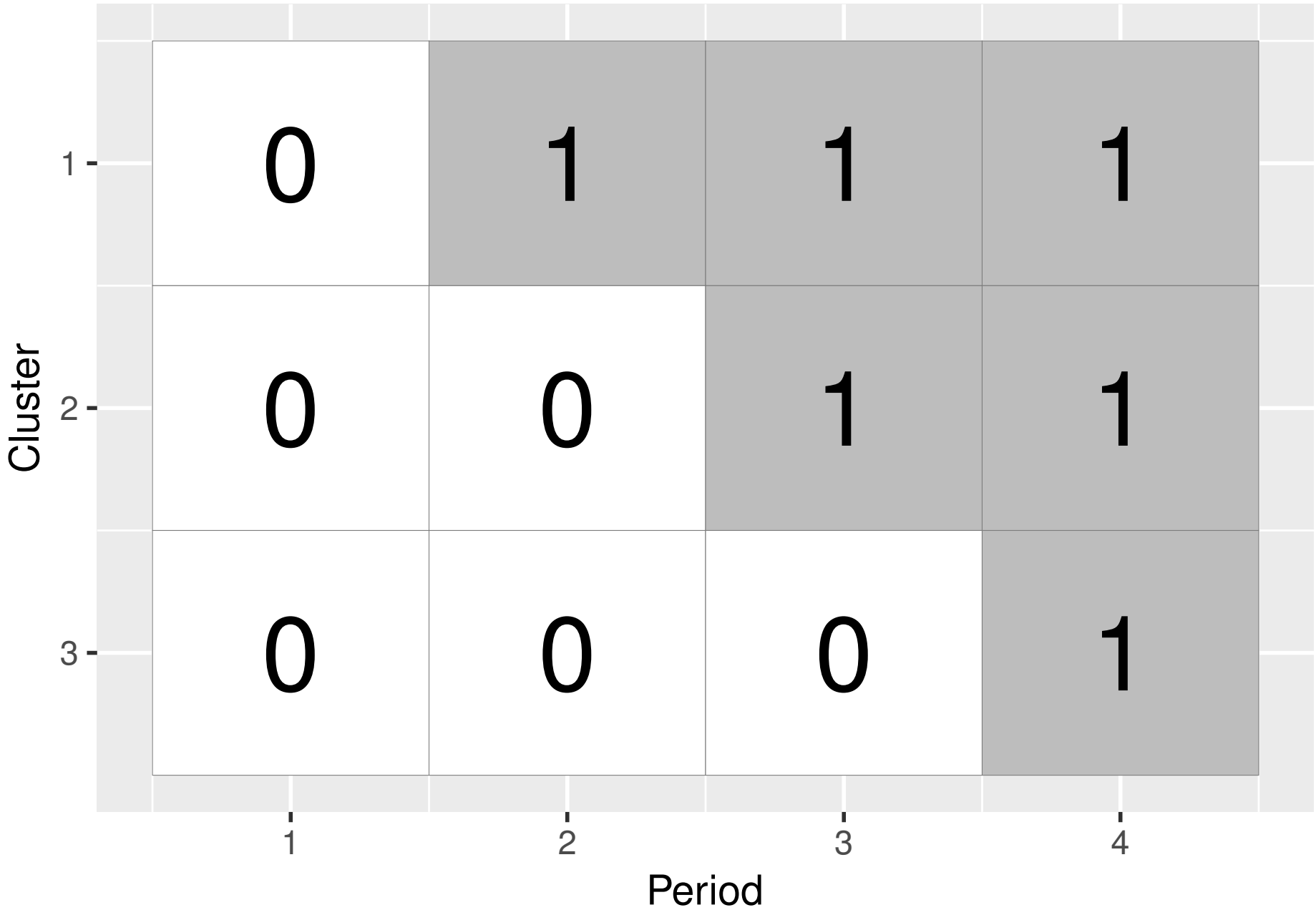
$$Y_{ikt} = \beta_t + X_{kt}\theta + C_k + \epsilon_{ikt}, \quad C_k \sim N(0, \tau^2), \quad \epsilon_{ikt} \sim N(0, \sigma_\epsilon^2)$$

$$\text{Intra-cluster correlation: } \rho_0 = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

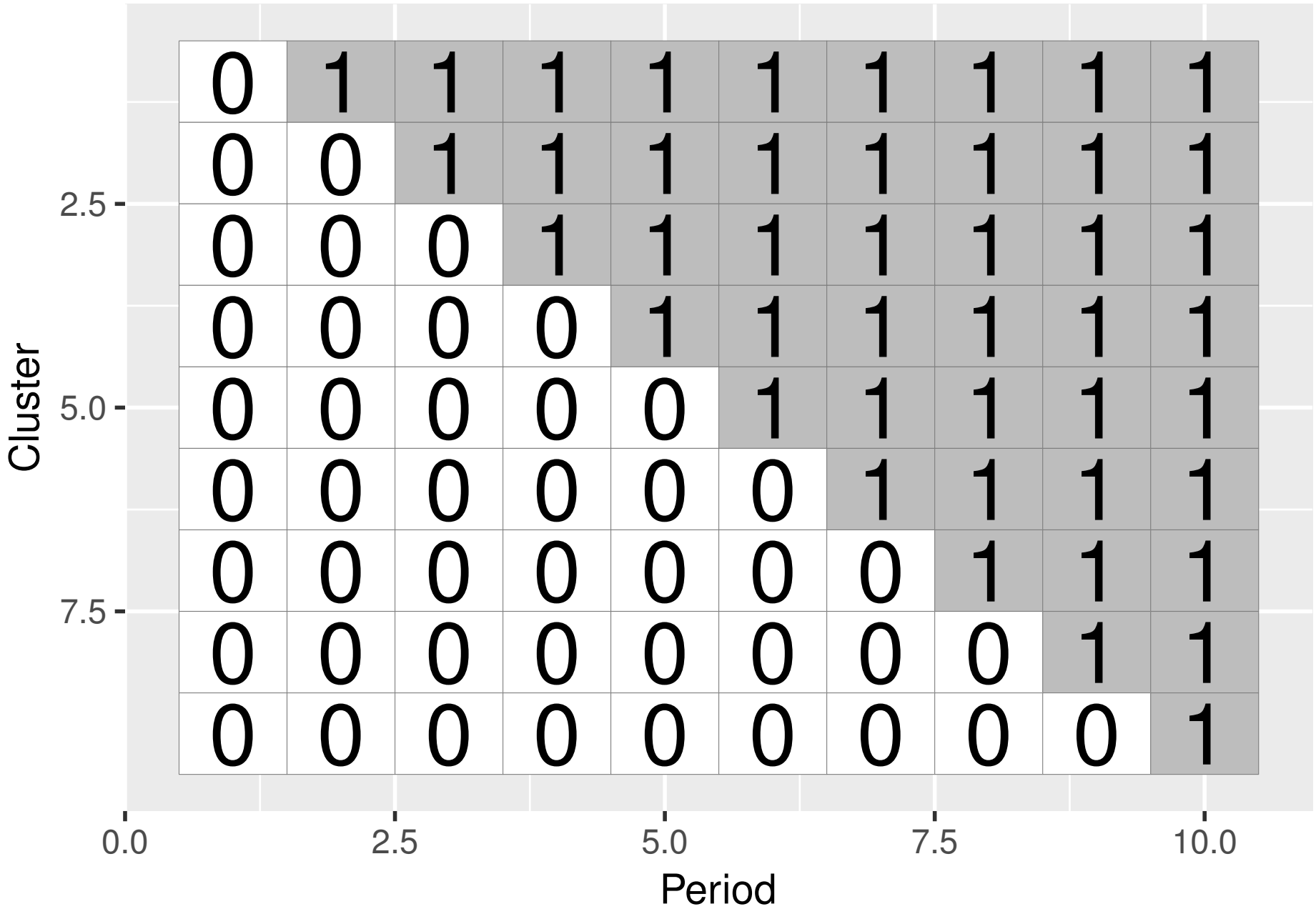
- Fix total variance at unity: $\tau^2 + \sigma_\epsilon^2 = 1 \Rightarrow \rho_0 = \tau^2 = 0.05$
- $m = 100$ subjects per cluster-period cell
- Consider standard SW designs with $T = 4, 10, 15, 20$ periods.

Calculate $IC(k, t)$ for $K = 1, \ldots, K$, $T = 1, \ldots, T$.

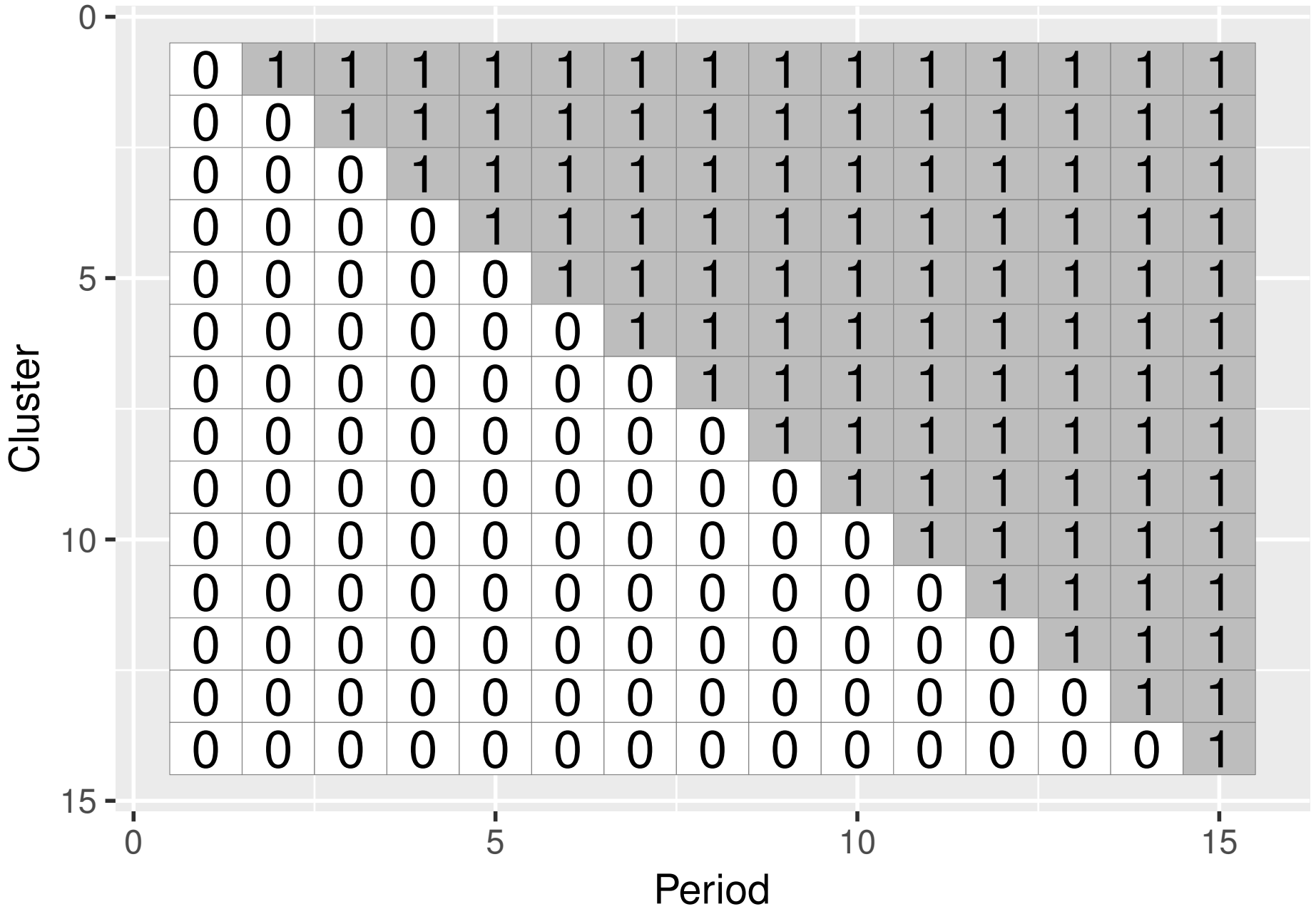
Design matrix: $T = 4$



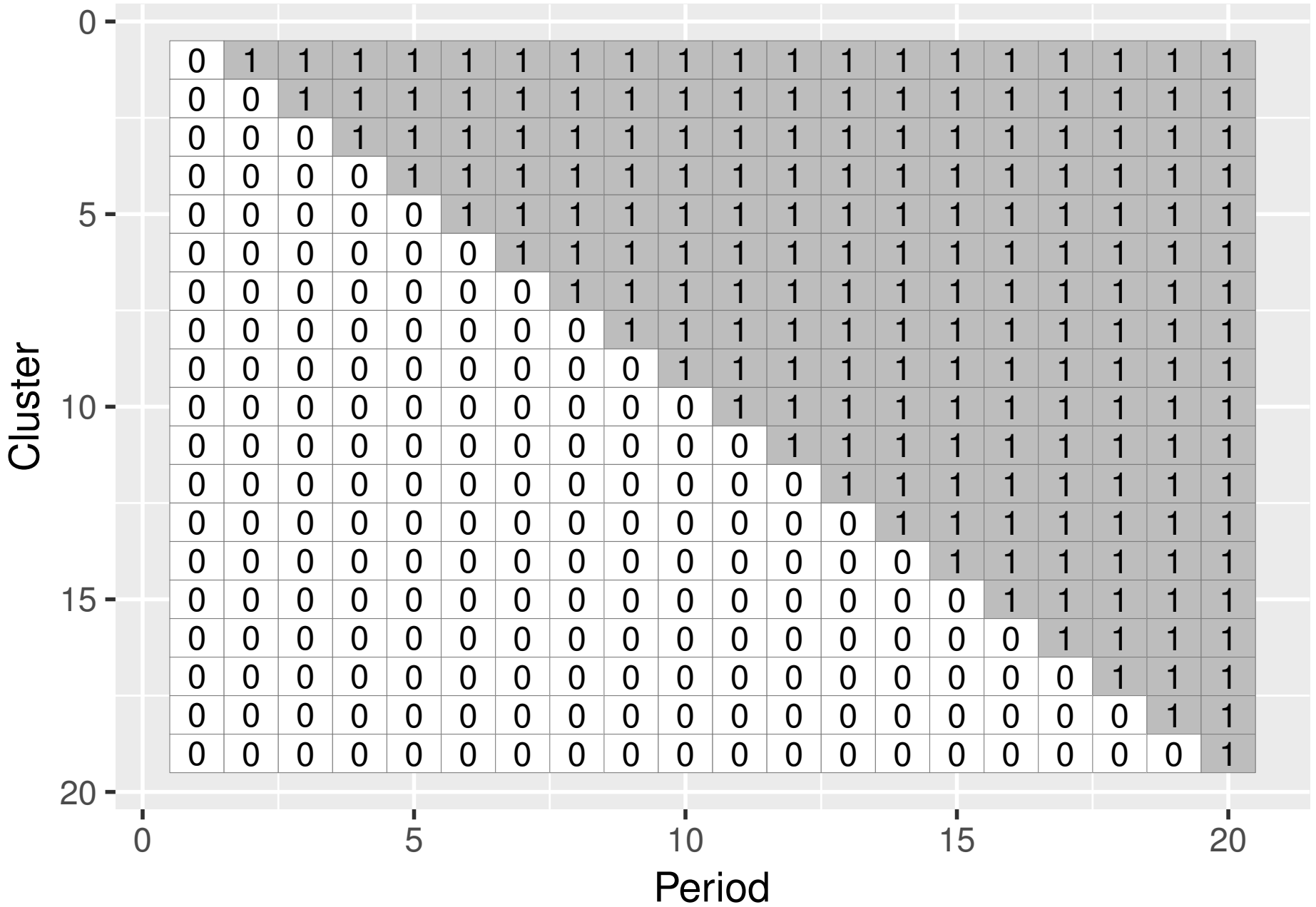
Design matrix: $T = 10$



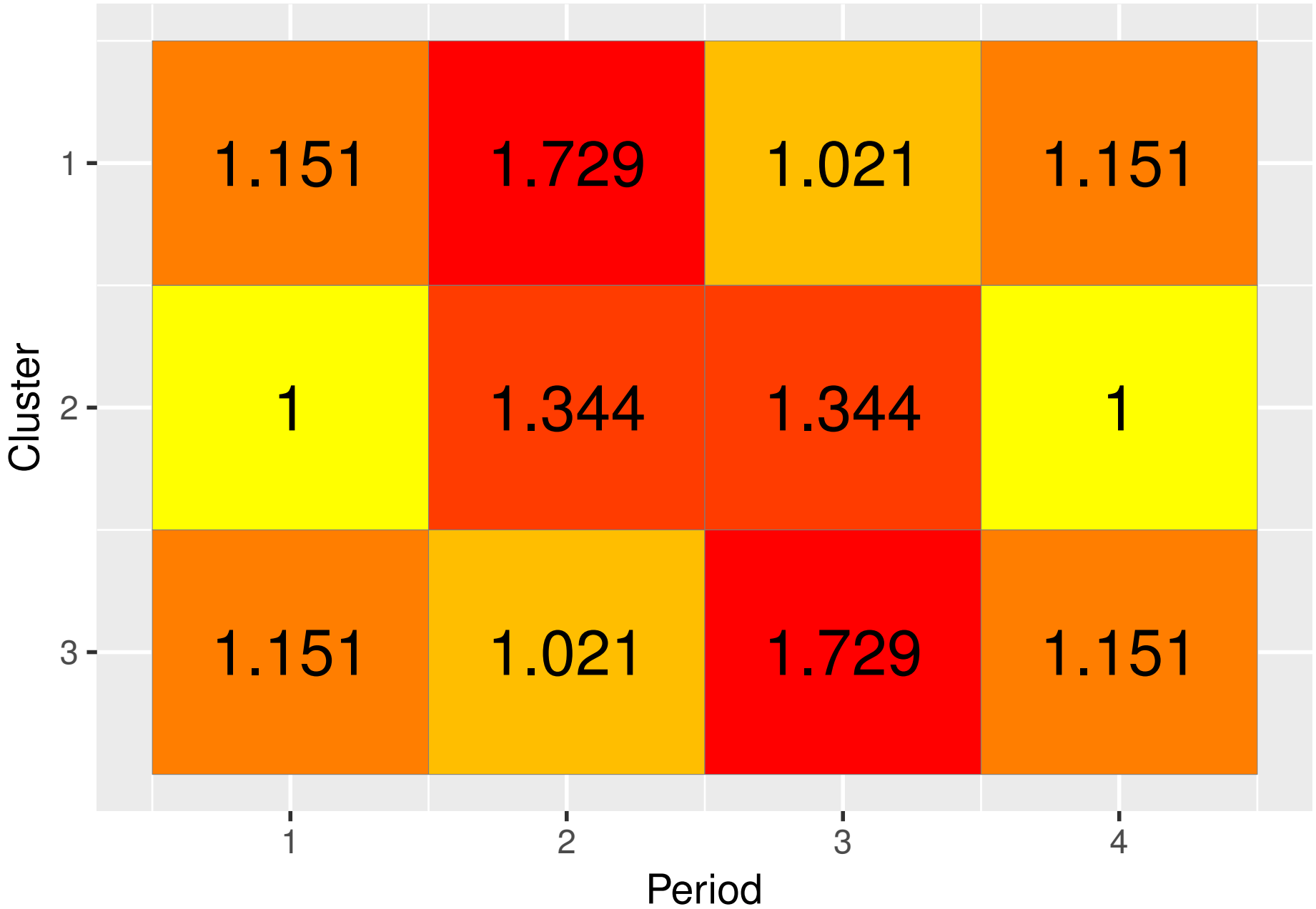
Design matrix: $T = 15$



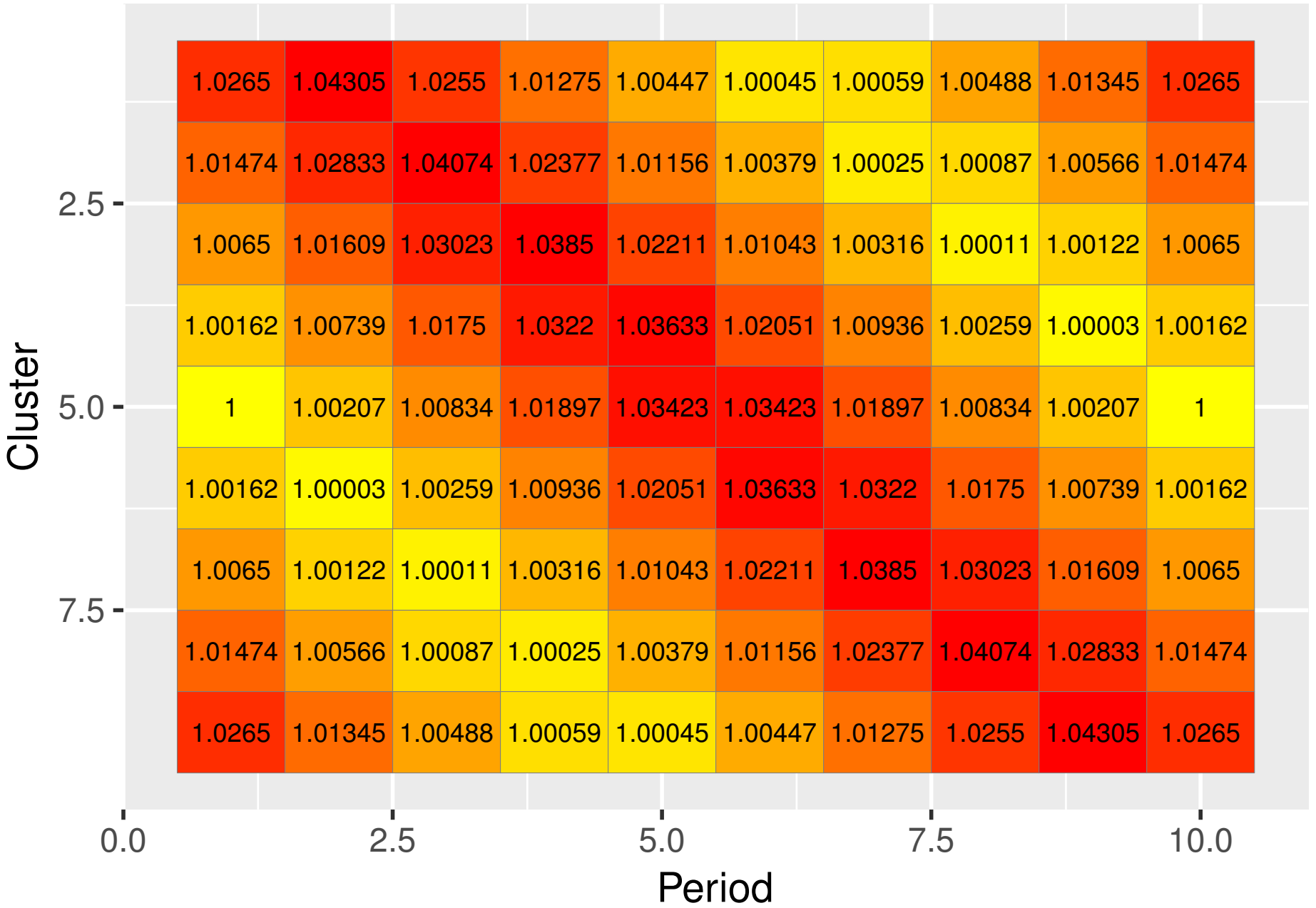
Design matrix: $T = 20$



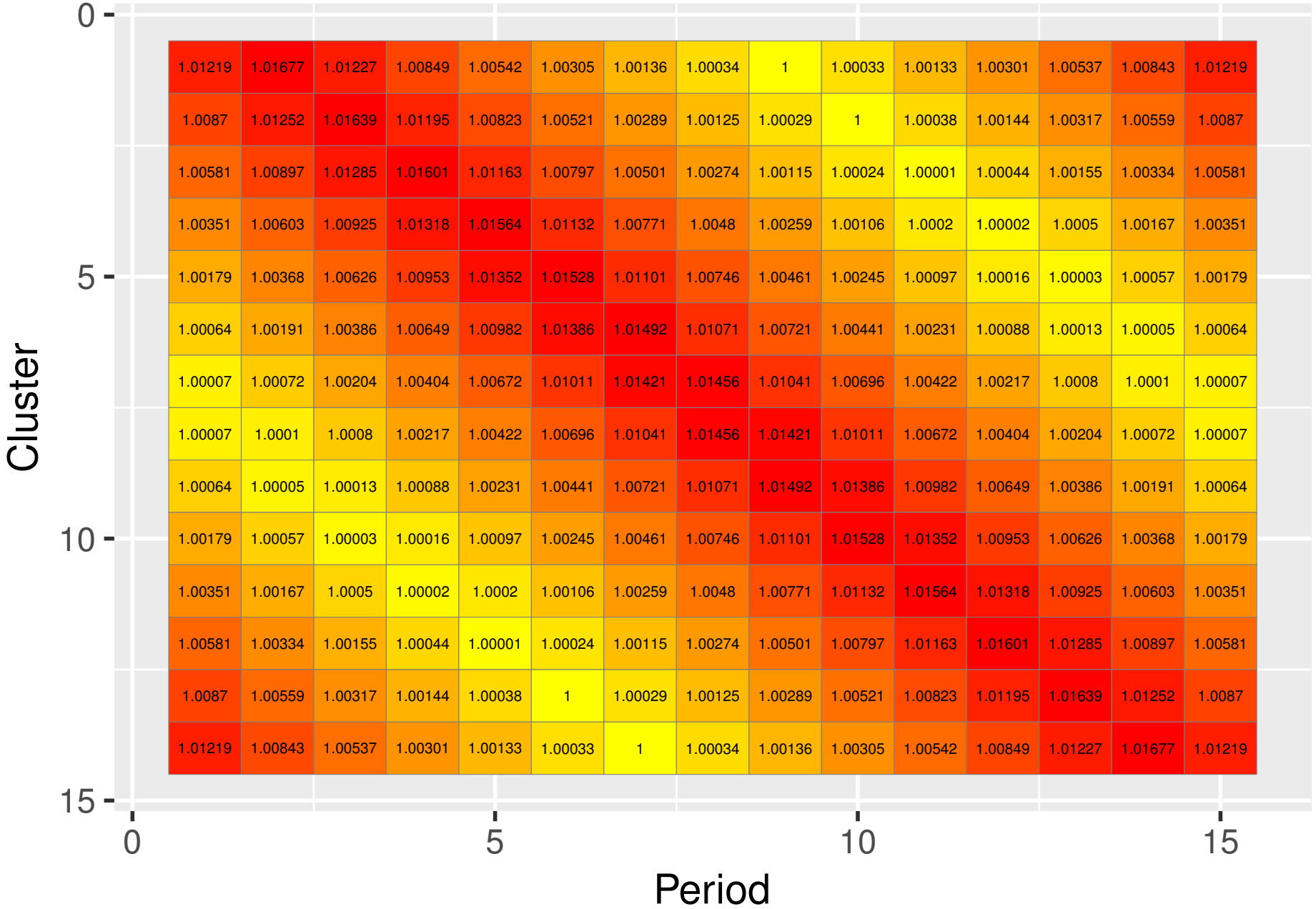
Information content of cells: $m = 100$, $\rho_0 = 0.05$



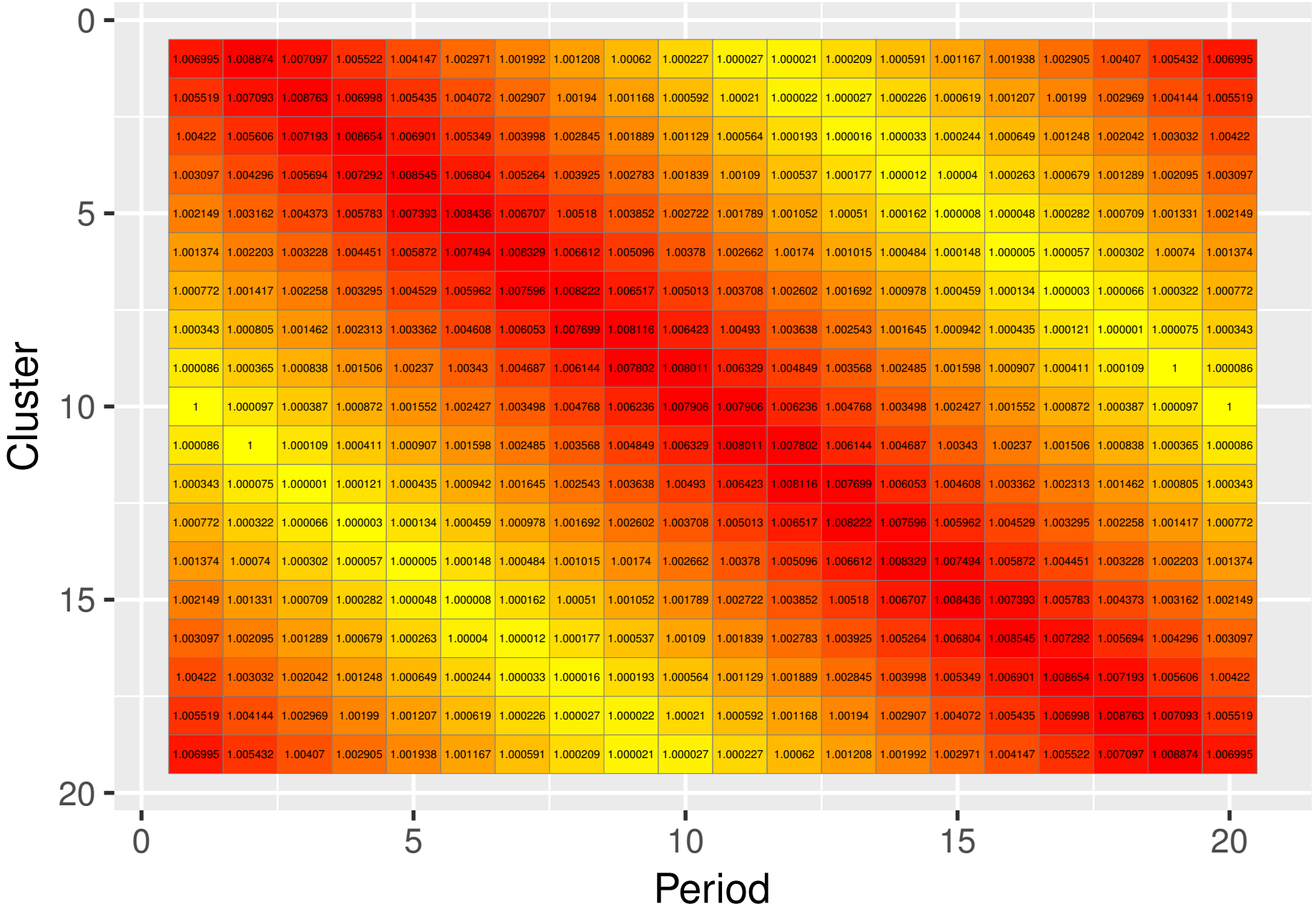
Information content of cells: $m = 100$, $\rho_0 = 0.05$



Information content of cells: $m = 100$, $\rho_0 = 0.05$



Information content of cells: $m = 100$, $\rho_0 = 0.05$



Final points

- Periods near the treatment cross-over tend to be most valuable...
  - But the “hot corners” are also necessary (allow for time effects)
- Logistical vs. statistical value of cells?

Here we assumed a simple structure for within-cluster correlations

- Hussey and Hughes: correlation does not depend on the time between observations from same cluster.
- What if the correlation between observations from the same cluster *decays* over time?

You can explore the information content of cells in your own cluster randomised trial at:

`https://jkasza.shinyapps.io/InformationContentofCells`

SAVE THE DATE

Joint International Society for Clinical Biostatistics and Australian Statistical Conference 26-30 August 2018

ISCB ASC18

26-30 AUGUST 2018

MELBOURNE, AUSTRALIA

A more complex intra-cluster correlation structure

- Hussey and Hughes:

$$Y_{ikt} = \beta_t + X_{kt}\theta + C_k + \epsilon_{ikt}, \quad C_k \sim N(0, \tau^2), \quad \epsilon_{ikt} \sim N(0, \sigma_\epsilon^2)$$

$$corr(Y_{ikt}, Y_{jkt}) = corr(Y_{ikt}, Y_{jks}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

- Exponential decay model:

$$Y_{ikt} = \beta_t + X_{kt}\theta + CP_{kt} + \epsilon_{ikt}, \quad \mathbf{CP}_k \sim N_T(\mathbf{0}, \tau^2 R), \quad \epsilon_{ikt} \sim N(0, \sigma_\epsilon^2)$$

$$R[t, s] = r^{|t-s|} \Rightarrow corr(Y_{ikt}, Y_{jkt}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2}$$

$$\textbf{but } corr(Y_{ikt}, Y_{jks}) = \frac{\tau^2}{\tau^2 + \sigma_\epsilon^2} r^{|t-s|}$$

Key difference: the correlation between two observations in the same cluster now depends on the amount of time between them!

Exponential decay and information content of clusters

$$Y_{ikt} = \beta_t + X_{kt}\theta + CP_{kt} + \epsilon_{ikt}, \quad \mathbf{CP}_k \sim N_T(\mathbf{0}, \tau^2 R), \quad \epsilon_{ikt} \sim N(0, \sigma^2_\epsilon)$$

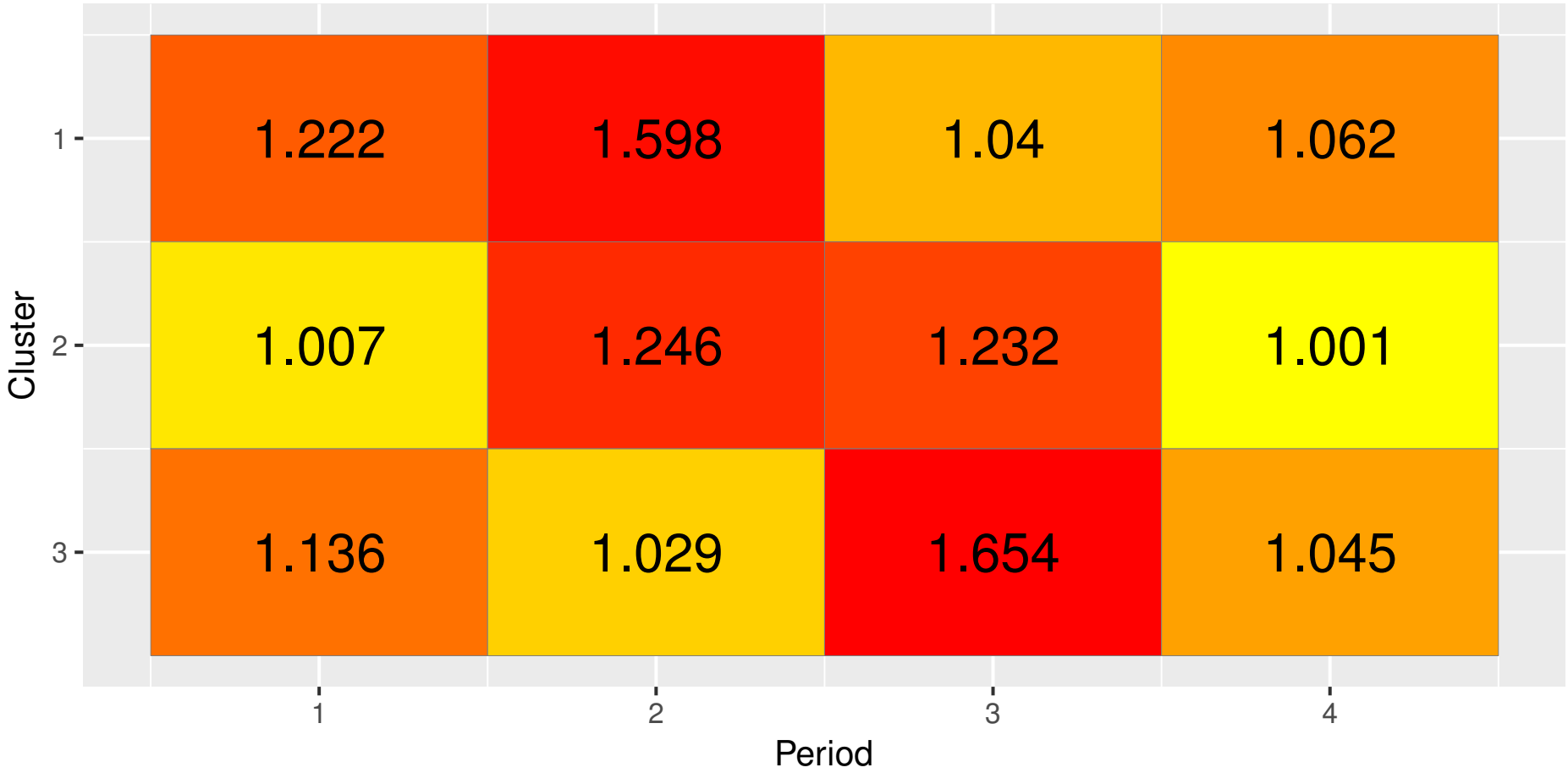
$$R[t,s] = r^{|t-s|} \Rightarrow corr(Y_{ikt}, Y_{jkt}) = \frac{\tau^2}{\tau^2 + \sigma^2_\epsilon}, corr(Y_{ikt}, Y_{jks}) = \frac{\tau^2}{\tau^2 + \sigma^2_\epsilon} r^{|t-s|}$$

Consider same design parameters as previously:

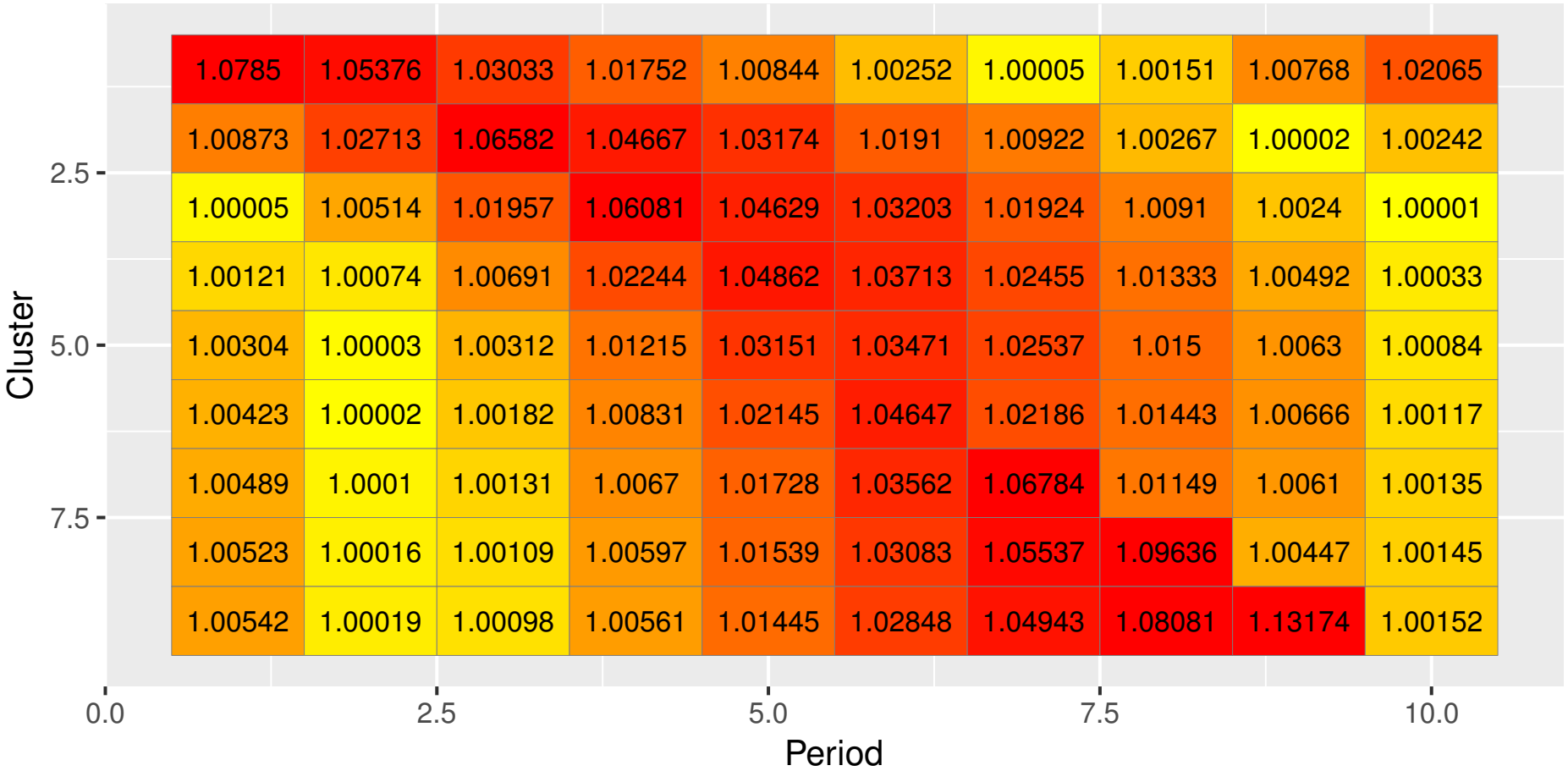
- Fix total variance at unity: $\tau^2 + \sigma^2_\epsilon = 1 \Rightarrow \rho_0 = \tau^2 = 0.05$
- $m = 100$ subjects per cluster-period cell
- Consider standard SW designs with $T = 4, 10, 15, 20$ periods.

What about r? Set $r = 0.95 \Rightarrow$ 5% decay in correlation per period.

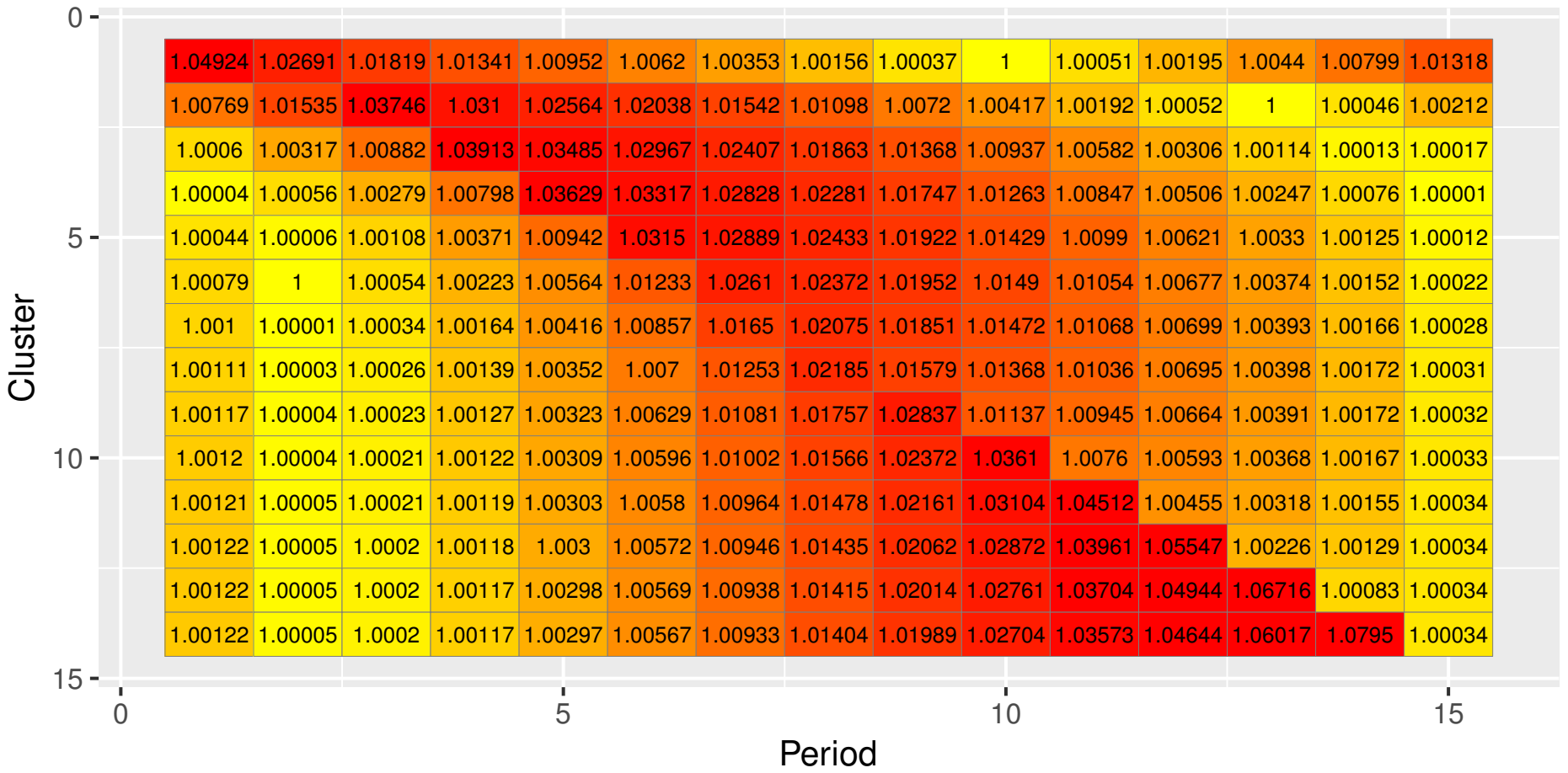
Information content of cells: $m = 100$, $\rho_0 = 0.05$, $r = 0.95$



Information content of cells: $m = 100$, $\rho_0 = 0.05$, $r = 0.95$



Information content of cells: $m = 100$, $\rho_0 = 0.05$, $r = 0.95$



Information content of cells: $m = 100$, $\rho_0 = 0.05$, $r = 0.95$