

Cox model with distributed data using the WebDISCO algorithm

Estimates and confidence intervals for the parameters of the Cox model $\lambda(t|Z) = \lambda_0(t) \exp(\beta^T Z)$ are obtained by making use of the Newton-Raphson algorithm using local gradients and Hessian matrices. This allows for the recreation of the estimates and confidence intervals from the centralised setting. For the procedure, all nodes must have the same predictors.

The initial step consists of each node sharing all their unique event times to the coordination center. The global server then identifies and consolidates all unique event times across all sites. Each site then initializes parameters that remain constant throughout the process. Next, the coordination center initializes the first beta using the inverse variance method. If the inverse variance method cannot be used, the coordination will instead do a simple average of each node's beta.

The iterative process starts with each site computing data aggregates. These aggregates are then used by the central server to calculate gradients and Hessian matrices, which in turn are used to update the parameter estimates. This iterative process continues until convergence is achieved (which yields the centralised estimate).

In the following procedure, k is the site number, t is the current iteration number ($t = 0, 1, \dots, \tau$), which increments every time the coordination node is reached, and M represents the number of sites ($k = 1, 2, \dots, M$). N^k is the number of observations in a site, p is the number of predictors, i represents the event time number, and D represents the total number of distinct event time ($i = 1, 2, \dots, D$). r and q are the indices of the element in the parameter vector β (r and $q = 1, 2, \dots, p$), z^l is the variable z (predictors) for an individual subject l . $W^{(k)}$ is the optional local vector of weights, whereas w^l is the weight for an individual subject l . If no weights are provided, a vector of 1s will be used instead, as this represents uniform weights across observations.

Of note, two algorithms are described below:

- Steps 1, 2, 3, 4, 5a, 6a and 7 describe the case where no weights (or uniform weights) are used. This corresponds to using R's coxph function in this way:
`coxph(formula, data, ties = "breslow")`
- Steps 1, 2, 3, 4, 5a, 5b, 6a, 6b and 7 describe the case where non uniform weights are used. This corresponds to using R's coxph function in this way:
`coxph(formula, data, ties = "breslow", weights, robust = TRUE)`

Commenté [RB1]: On ne leur donne pas vraiment le choix. Dans la présente implémentation, la méthode «inverse variance» est utilisée à moins que la matrice soit singulière et alors la méthode du moyennage simple est utilisée.

Commenté [RB2]: Peut-être mentionner que l'on alterne toujours entre les nœuds de données et le serveur central. Peut-être mentionner que le compteur d'itération change au serveur central. Ça aiderait à être moins confus de voir l'itération 0 dans les sites, mais pas d'itération 0 au serveur central.

Commenté [JM3R2]: Comme ça?

Commenté [RB4]: Peut-être indiquer que dans la suite du document les étapes a sont pour le cas sans poids et les étapes b pour le cas avec poids.

Commenté [JM5R4]: Mieux comme ça?

Commenté [RB6R4]: Oui :)

Example.

Suppose the following dataset at node k , with 5 observations ($N = 5$) and 2 predictors ($p = 2$):

Id	time	status	age	sex	weights
1	3	1	42	0	2
2	6	0	38	0	1
3	11	1	37	1	3
4	11	1	51	0	4
5	14	1	36	1	6

where time represents the time of event occurrence or censoring, status indicates whether the event has occurred (1) or the data is censored (0) and age and sex are the predictors. The last column is the optional weight vector, which is provided in this example.

Data node (initial phase)

- Each node identifies unique event times and saves this data to a CSV file (Times_k_output.csv). Each site also computes their local Cox models to obtain local betas (Beta_local_k.csv) and variance-covariance matrices (V_k_k.csv), saving them to a CSV file along with the number of observations N^k (N_node_k.csv). For verification purposes, all predictor variable names and the robust estimation flag (set to TRUE for a robust variance estimation and set to FALSE for a classic variance estimation) are also saved in a CSV file (Local_Settings_k.csv). All files are then sent to the coordinating node.

Commenté [RB7]: Est-ce que les matrices de variance-covariance sont enregistrées dans les fichiers « V_k_k.csv »?

Commenté [JM8]: Ceci pourrait devenir « local_parameters ». Ça permettrait d'également ajouter si on utilise des poids (et par conséquent si le nœud k s'enligne pour une estimation robuste ou non).

Commenté [JM9R8]: Attention: On aura probablement besoin d'un toggle pour robust vs classique.

Si on se fit juste aux poids, on pourrait avoir un problème si un site a des poids uniforme (voulu ou par chance) et pas les autres.

Commenté [RB10]: Est-ce que c'est le même exemple que plus haut ou pas?
Si oui, ça devrait être 3, 6, 11, 14.

Commenté [JM11R10]: Non, car on partage les *unique event times* et il n'y a pas d'évènement au temps 6 (le status est à 0).

Commenté [RB12R10]: ok :)

Example (continued).

Estimates for the Cox proportional hazard model can be found using functions such as *coxph* in R, and we obtain:

$$\hat{\beta}^k = \begin{bmatrix} -0.1654 \\ -3.6567 \end{bmatrix}$$

The following data is shared to the coordinating node:

time
3
11
14

$$\hat{\beta}^k = \begin{bmatrix} -0.1654 \\ -3.6567 \end{bmatrix}, V_k = \begin{bmatrix} 0.0189 & 0.2607 \\ 0.2607 & 4.1247 \end{bmatrix}, \text{ and } N^k = 5.$$

Coordinating node (initial phase)

2. The coordinating node gathers all unique times from every site, generates a list of unique times across all sites, then sends it back to the data nodes (Global_times_output.csv). The global server also computes the inverse variance weighted initial estimator, $\hat{\beta}_0 = (\sum_{k=1}^M V_k^{-1})^{-1} (\sum_{k=1}^M V_k^{-1} \hat{\beta}^k)$, and saves it to a CSV file (Beta_0_output.csv).¹ Should any of the matrix V_k be singular, the initial estimator will be replaced by a simple average $\hat{\beta}_0 = (\sum_{k=1}^M n_k \hat{\beta}^k) / \sum_{k=1}^M n_k$.

Example (continued).

Assume that the data is stored in two nodes and that the coordinating node received these time lists:

time
3
11
14

time
1
3
10
11

The following data is shared to the data nodes:

time
1
3
10
11
14

The estimate $\hat{\beta}_0$ is also shared to the data nodes.

Data node (iteration t=0)

3. Each node initialises the following parameters:
 - R_i^k (Rikk.csv), a list containing the IDs of the at-risk subjects for all times i ,
 - $R_i^{k'}$ (Rik_compk.csv), a list containing the IDs of subjects that died at or by time i ,
 - D_i^k , a list containing the IDs of subjects that had an event at time i .
 Those files will be used by the node throughout the analysis and are not shared with the coordinating node.

Each node also computes the following: $\sum_{l \in D_i^k} w^l z_r^l$ (sumWZrk.csv), which is the weighted

¹ Duan, R., Luo, C., Schuemie, M. J., Tong, J., Liang, C. J., Chang, H. H., ... & Chen, Y. (2020). Learning from local to global: An efficient distributed algorithm for modeling time-to-event data. *Journal of the American Medical Informatics Association*, 27(7), 1028-1036.

Commenté [SL13]: Détail : est-ce qu'ici ça devrait être $(\sum V_k^{-1})^{-1} * (\sum V_k^{-1} \beta^k)$? Vu que les V_k sont des matrices

(Si je comprends bien l'estimateur, c'est l'estimateur distributed estimating equations, cool)

Commenté [RB14]: (Beta_0_output.csv)

Commenté [RB15]: Peut-être mentionner que si la matrice n'est pas inversible, alors la méthode du moyenage simple est utilisée et donner la formule.

Commenté [RB16]: « at or by time i » pourrait peut-être porter à confusion. Je crois qu'il faudrait dire seulement « at ». On veut seulement les individus ayant eu l'événement au temps i, pas ceux qui ont eu l'événement avant i.

Commenté [JM17R16]: Le $\{R_i^k\}$ est un peu particulier. Ce n'est pas le vrai complémentaire du risk set. C'est vraiment les gens qui ont eu l'événement au temps i ou avant le temps i.

Commenté [JM18R16]: Plus clair comme ça?

Commenté [JM19]: À vérifier:

On ne l'enregistre plus. On pourrait probablement le retirer de la docu.

Commenté [RB20R19]: Comme on en a besoin dans les formules, je crois qu'on doit le laisser dans ce document, mais ne pas en parler dans le README.

Commenté [JM21R19]: Je pensais avoir remplacer tout les D_i^k par d'autres quantités qui font intervenir les poids..

Commenté [JM22R19]: On le garde alors!

sum for each component r of the predictor for subjects with an event at time i , and $W^{(k)'} = \sum_{l \in D_i^k} w^l$ (Wprimek.csv), which is the list containing the total weight of events at time i for node k . All of these CSV files are sent to the coordinating node.

Commenté [SL23]: Est-ce qu'on veut dire : for each component r ? Ou all r predictors ? Ou quelque chose du genre (dans l'exemple, on a une même tableau pour $r=1$ et $r=2$, deux prédicteurs différents)

Commenté [SL24]: Les écrire

Example (continued).

The following data is computed but not shared to the coordinating node:

i	R_i^k	$R_i^{k'}$	D_i^k
1	1,2,3,4,5	NA	NA
3	1,2,3,4,5	1	1
10	3,4,5	1	NA
11	3,4,5	1,3,4	3,4
14	5	1,3,4,5	5

The following data is shared to the coordinating node:

i	$W^{(k)'}$
1	0
3	2
10	0
11	7
14	6

i	$\sum_{l \in D_i^k} w^l z_r^l$	
	$r = 1$ (age)	$r = 2$ (sex)
1	0	0
3	84	0
10	0	0
11	315	3
14	216	6

Then, each data node computes aggregated statistics used for the gradient and Hessian:

- $\sum_{l \in R_i^k} w^l \exp(\beta^T z^l)$, a $D \times 1$ matrix (sumWExpk_output_t.csv)
- $\sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l)$, a $D \times p$ matrix (sumWZqExpk_output_t.csv)
- $\sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l)$, a $p \times p \times D$ matrix (sumWZqZrExpk_output_t.csv)

These quantities are sent to the coordinating center using CSV files.

Example (continued).

Suppose the coordinating node shared $\hat{\beta} = \begin{bmatrix} -0.1654 \\ -3.6568 \end{bmatrix}$.

For the first event time ($i = 1$):

$$\sum_{l \in R_i^k} w^l \exp(\beta^T z^l) = 2 \cdot \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 42 \\ 0 \end{bmatrix}\right) + 1 \cdot \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 38 \\ 0 \end{bmatrix}\right) + \dots$$

$$\sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l) = 2 \cdot [42 \quad 0] \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 42 \\ 0 \end{bmatrix}\right)$$

$$+ 1 \cdot [38 \quad 1] \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 38 \\ 0 \end{bmatrix}\right) + \dots$$

$$\sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l) = 2 \cdot \begin{bmatrix} 42^2 & 42 \cdot 0 \\ 0 \cdot 42 & 0^2 \end{bmatrix} \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 42 \\ 0 \end{bmatrix}\right)$$

$$+ 1 \cdot \begin{bmatrix} 38^2 & 38 \cdot 0 \\ 0 \cdot 38 & 0^2 \end{bmatrix} \exp\left([-0.1654 \quad -3.6568] * \begin{bmatrix} 38 \\ 0 \end{bmatrix}\right) + \dots$$

The following quantities are **shared** to the coordinating node:

i	$\sum_{l \in R_i^k} w^l \exp(\beta^T z^l)$
1	0.0052
3	0.0052
10	0.0014
11	0.0014
14	0.0004

Commenté [JM25]: À rouler avec les données de l'exemple plutôt que de reproduire ça à la main.

i	$\sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l)$	
	$q = 1$	$q = 2$
1	0.2165	0.0006
3	0.2165	0.0006
10	0.0650	0.0006
11	0.0650	0.0006
14	0.0145	0.0004

i	$\sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l)$	
	$q = 1$ (age)	$q = 2$ (sex)
1	$r = 1$ (age) 9.0899	0.0208
	$r = 2$ (sex) 0.0208	0.0006
3	$r = 1$ (age) 9.0899	0.0208
	$r = 2$ (sex) 0.0208	0.0006

		$r = 1$ (age)	3.0098	0.0208
		$r = 2$ (sex)	0.0208	0.0006
	10	$r = 1$ (age)	3.0098	0.0208
	11	$r = 2$ (sex)	0.0208	0.0006
	14	$r = 1$ (age)	0.5205	0.0145
		$r = 2$ (sex)	0.0145	0.0004

Coordinating node (iteration t=1)

4. The global server aggregates $\sum_{k=1}^M \sum_{l=1}^D \sum_{l \in D_i^k} w^l z_r^l$ (sumWZrGlobal.csv), which is the sum of all predictors for subjects that had an event across all sites, and calculates $W' = \sum_{k=1}^M \sum_{l \in D_i^k} w^l$ (WprimeGlobal.csv), the list containing the total weight of events at time i .

The coordinating center also computes the global gradient and Hessian:

$$l'_r(\beta) = \sum_{k=1}^M \sum_{l=1}^D \sum_{l \in D_i^k} w^l z_r^l - \sum_{l=1}^D \left(\sum_{k=1}^M \sum_{l \in D_i^k} w^l \right) \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)}$$

$$l''_{r,q}(\beta) = - \sum_{l=1}^D \left(\sum_{k=1}^M \sum_{l \in D_i^k} w^l \right) \left\{ \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \right. \\ \left. - \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \right\}$$

It then computes the Newton-Raphson iteration $\beta^\tau = \beta^{\tau-1} - [l''(\beta^{\tau-1})]^{-1} l'(\beta^{\tau-1})$ and sends the new estimate to the data nodes (Beta_1_output.csv).

Data node (iteration t=1,2,3,...)

5. Classic estimation & robust estimation

- a. Each data node computes aggregated statistics used for the gradient and Hessian:
- $\sum_{l \in R_i^k} w^l \exp(\beta^T z^l)$, a $D \times 1$ matrix (sumWExpk_output_t.csv)
 - $\sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l)$, a $D \times p$ matrix (sumWZqExpk_output_t.csv)
 - $\sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l)$, a $p \times p \times D$ matrix (sumWZqZrExpk_output_t.csv)

These quantities are sent to the coordinating center using CSV files.

Commenté [RB26]: Devrais être plus haut?
On parle longtemps juste des data nodes avant d'arriver ici.

(Robust estimation only)

- b. When using *case weights*, Therneau & Grambsch² suggest using a robust variance estimation, based on score residuals³. In order to do so, more intermediate statistics must be shared between nodes and coordinating node.

Therefore, for iteration t>1, each data node computes aggregated statistics used for the robust variance estimation:

$$\sum_{r \in R_i^k} \frac{w^r \bar{z}_{R_r}}{\sum_{l=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)} \text{ (zbarri_inverseWExp_k_output_(t-1).csv),}$$

and

$$\sum_{r \in R_i^k} \frac{w^r}{\sum_{l=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)} \text{ (inverseWExp_k_output_(t-1).csv).}$$

The CSV files are then shared to the coordinating node.

Example (continued).

Suppose the coordinating node shared the following:

i	\bar{z}_{R_i}	
	age	sex
1	44.9297	0.4697
3	45.2442	0.4997
10	46.0140	0.5440
11	46.1964	0.6419
14	36	1

and

i	$\sum_{l=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)$
1	481.1013
3	452.2446
10	396.2411
11	335.8129
14	66.7125

Commenté [JM27]: À rouler avec les données de l'exemple plutôt que de reproduire ça à la main.

Then, the following is shared to the coordinating node:

$$\sum_{r \in R_i^k} \frac{w^r \bar{z}_{R_r}}{\sum_{l=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)}$$

² T.M. Therneau and P. M. Grambsch. *Modeling Survival Data: Extending the Cox Model*. Statistics for Biology and Health. Springer New York, 2013.

³ Collett, D. (2023). *Modelling Survival Data in Medical Research* (4th ed.). Chapman and Hall/CRC.
<https://doi.org/10.1201/9781003282525>

i	age	sex
1	0	0
3	0.2001	0.0022
10	0.2001	0.0022
11	1.1630	0.0156
14	4.4008	0.1055

i	$\sum_{r \in R_i^k} \frac{w^r}{\sum_{l=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)}$
1	0
3	0.0044
10	0.0044
11	0.0253
14	0.1152

Commenté [JM28]: À rouler avec les données de l'exemple plutôt que de reproduire ça à la main.

Additionally, for iteration $t > 2$, each node should now be able to compute the score residuals of their subjects. Using this, each node will compute the following aggregated statistic used for the robust variance estimation.

First, each node will compute their subjects' *Schoenfeld residuals*:

$$r_{Schoenfeld_l} = w^l \delta_l (z^l - \bar{z}_{R_l}),$$

where δ_l is the event indicator for subject l (this is what we call the *status* variable).

Afterwards, each node will compute their subjects' *score residuals*:

$$r_{score_l} = r_{Schoenfeld_l} + \exp(\beta^T z^l) \sum_{t_r \leq t_l} \frac{w^r \delta_r (\bar{z}_{R_r} - z^l)}{\sum_{k=1}^M \sum_{j \in R_r^k} w^j \exp(\beta^T z^j)}.$$

Finally, each node will compute their contribution to the estimation of the jackknife estimate of variance $D_k^t D_k$, where:

$$D_k = r_{score} \hat{\Sigma}.$$

Note that only the diagonal of the matrix $D_k^T D_k$ will be shared to the coordinating node through a CSV (DDk_output_(t-2).csv) as non-diagonal entries are left unused.

The CSV files are then shared to the coordinating node.

Example (continued).

Suppose the coordinating node shared the following:

i	age	sex
1	0.1868	0.0019
3	0.4869	0.0052
10	0.8353	0.0094
11	2.3485	0.0304
14	5.5863	0.1204

i	$\sum_{k=1}^M \sum_{r \in R_i^{k'}} \frac{w^r}{\sum_{k=1}^M \sum_{l \in R_r^k} w^l \exp(\beta^T z^l)}$
1	0.0042
3	0.0108
10	0.0184
11	0.0511
14	0.1411

Note: The coordinating node, starting at iteration 1, will also compute and share Fisher's information related to the current estimate. This information is not available for the iteration 0.

Commenté [JM29]: J'imagine qu'on ne veut pas laisser ça ici? Mais info importante tout de même.

$$\hat{\Sigma}_1 = \begin{bmatrix} 0.0061 & 0.0778 \\ 0.0778 & 1.2713 \end{bmatrix}$$

Note that the data node should already have access to past estimates of β , such as, say:

$$\hat{\beta}_1 = \begin{bmatrix} 0.0667 \\ 0.0064 \end{bmatrix}$$

This previous estimate of β is needed in order to compute the factor $\exp(\beta^T z^l)$ in the formula of the score residual.

Then, each data node computes but does not share the following:

i	$r_{\text{Schoenfeld}}$	
i	age	sex
1	-4.182535	-0.6364
3	0	0
10	-25.2514	1.4250
11	22.3314	-2.0100
14	0	0

i	r_{Score}	
	age	sex
1	-3.0712	-0.4627
3	0.9704	0.0665
10	-8.9507	0.6867
11	-8.7531	1.5570
14	33.9073	-1.3813

$$D_k = \begin{bmatrix} -0.0549 & -0.8272 \\ 0.0111 & 0.1600 \\ -0.0015 & 0.1767 \\ 0.0674 & 1.2985 \\ 0.1008 & 0.8819 \end{bmatrix}, \quad D_k^T D_k = \begin{bmatrix} 0.0178 & 0.2232 \\ 0.2232 & 3.2048 \end{bmatrix}.$$

However, note that only the following is shared to the coordinating node:

age	sex
$\sqrt{0.0178}$	$\sqrt{3.2048}$

Coordinating node (iteration t=2,3,4,...)

6.

(Classic estimation & robust estimation)

a. The coordinating center computes the global gradient and Hessian:

$$l'_r(\beta) = \sum_{k=1}^M \sum_{l=1}^D \sum_{l \in R_i^k} w^l z_r^l - \sum_{l=1}^D \left(\sum_{k=1}^M \sum_{l \in R_i^k} w^l \right) \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)}$$

$$l''_{r,q}(\beta) = - \sum_{l=1}^D \left(\sum_{k=1}^M \sum_{l \in R_i^k} w^l \right) \left\{ \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l z_q^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \right. \\ \left. - \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_r^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \frac{\sum_{k=1}^M \sum_{l \in R_i^k} w^l z_q^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \right\}$$

It then computes the Newton-Raphson iteration $\beta^\tau = \beta^{\tau-1} - [l''(\beta^{\tau-1})]^{-1} l'(\beta^{\tau-1})$ and sends the new estimate to the data nodes (Beta_t_output.csv).

The coordinating node also computes the confidence intervals. The (estimated) covariance matrix of the coefficients is $\hat{\Sigma} = -(l''(\beta^{\tau-1}))^{-1}$.

Lower and upper bounds for the confidence intervals of the model parameters are also calculated at the coordinating node:

$$CI(\beta) = \left[\beta^{\tau-1} \pm \frac{z_\alpha}{2} \sqrt{diag(\hat{\Sigma})} \right]$$

Commenté [JM30]: Éventuellement, on aimerait que cette matrice ne soit pas partager lorsqu'on est dans le mode classique.

The outputs of the procedure are the coefficients, the upper and lower bounds of the confidence intervals for the exponential of the model parameters, the standard error and the p values, for the previous iteration.

Final results will be located in the Results_t.csv file.

(Robust estimation only)

- b. The (estimated) covariance matrix of the coefficients is $\hat{\Sigma} = -(l''(\beta^{T-1}))^{-1}$ needs to be saved (Fisher_t.csv) and shared with local nodes in order to get robust variance estimates.

The coordinating center computes and stores in a CSV file both:

$$\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l) \text{ (sumWExpGlobal_output_(t-1).csv),}$$

and

$$\bar{z}_{R_i} = \frac{\sum_{k=1}^M \sum_{l \in R_i^k} z^l w^l \exp(\beta^T z^l)}{\sum_{k=1}^M \sum_{l \in R_i^k} w^l \exp(\beta^T z^l)} \text{ (zbarri_(t-1).csv).}$$

The CSV files are then shared to the local nodes.

Additionally, for iteration $t > 2$, the coordinating center computes both

$$(\text{zbarri_inverseWExp_Global_output_(t-2).csv}),$$

and

$$(\text{inverseWExp_Global_output_(t-2).csv}).$$

The CSV files are then shared to the local nodes.

Commenté [RB31]: Coordinating node (iterations t>2)

Additionally, for iteration $t > 3$, the coordinating node has now all the needed information to compute the robust standard error of β . These standard errors are the squared root of the diagonal of the variance matrix $DD = \sum_{k=1}^M D_k^T D_k$. This allows the coordinating node to produce a new results file (RobustResults_(t-3).csv) which is an updated version of Results_(t-3).csv.

Commenté [RB32]: Coordinating node (iterations t>3)

Final results will be located in the RobustResults_t.csv file.

Convergence

- 7. Steps 5 and 6 are repeated manually until convergence.