Web-appendix of 'Guidance for Practitioners on the Choices of Software Implementation for Frailty Models: Simulations and An Application in Determining the Birth Interval Dynamics'

1 Mathematical Form of the Frailty Model

In a simplified setting, we assume that each subject $i, (i = 1, 2, ..., n)$ is a member of a single family $j, (j = 1, 2, ..., q)$, having an inherent and unmeasured risk (frailty) ν_i, such that–

$$\begin{aligned} \lambda_i(t) &= \lambda_0(t) \times \nu_i \times \Psi(\boldsymbol{X}_i, \boldsymbol{\beta}) \\ &= \lambda_0(t) \times \nu_i \times e^{\boldsymbol{X}_i' \boldsymbol{\beta}} \\ &= \lambda_0(t) \times \exp(\boldsymbol{X}_i' \boldsymbol{\beta} + \boldsymbol{Z}_i' \boldsymbol{\omega}), \end{aligned} \quad (1.1)$$

where $\boldsymbol{Z}$ is a design matrix, $Z_{ij} = 1$ if and only if the subject i is a member of family j, otherwise 0, $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are the ith row of covariate matrices $\boldsymbol{X}_{n\times p}$, $\boldsymbol{Z}_{n\times p}$ respectively, $\boldsymbol{X}$ and $\boldsymbol{\beta}$ correspond to p fixed effects in the model (covariates may be binary or continuous), $\boldsymbol{\omega}$ is a vector containing the q unknown random effects (frailties) (Therneau and Grambsch, 1998).
We assume that $\boldsymbol{\omega}$'s are independent sample from an arbitrary distribution with the property that the mean is zero and the unit variance. As a consequence, ν's are independently and identically distributed from an arbitrary distribution with the property that the mean is one, with some unknown finite variance. If this is true, then interpretation becomes easier: assuming $\nu_i = 1$ as standard or average hazard of the subjects (this makes the usual proportional hazards model), for $\nu_i > 1$, individual fails faster, or said to be more 'frail' than an average subject, while for $\nu_i < 1$, individual lives longer in general, therefore, being less 'frail' than an average subject. However, we need to make assumptions about this scale factor being independent of both survival times and covariates. This model further assumes that survival times are conditionally independent given the frailty and the frailty term and the covariates are independent. Frailty and the covariate effects (function of $\boldsymbol{X}$) are assumed to act multiplicatively on the baseline hazard and the value of the frailty is constant within the group. Therefore, each study subject is susceptible to the study outcome and will eventually experience that event if they are followed for a sufficiently long time. As the scale factor 'frailty' is unobserved, estimating survivorship function and the density function is problematic, as both of them are mathematically related to hazard function that clearly depends on the scale factor. The trick researchers use is to specify a parametric distributional form of this unobserved scale factor in terms of fairly small number of parameters (the distribution is chosen in such a way that the parameters can be estimated in some way) and then for continuous case, integrate the frailty out of the expression.

2 Theoretical Issues Regarding the Likelihood Construction

2.1 Generalization of Cox Model with Random Effects

We will consider clustered failure-time data with q clusters, each with size n (this is obviously the special case of all n_j being equal for each cluster), being adherent to the notations in the literature

(Fleming and Harrington, 1991; Andersen, 1993). The failure-time variable corresponding to subject i $(i = 1, 2, ..., n)$ from cluster j $(j = 1, 2, ..., q)$ will be denoted by Y_{ij}. It is assumed that the observations of Y_{ij} can be right-censored. Thus, for subject i in the cluster j we observe

$$T_{ij} = \min(C_{ij}, Y_{ij})$$

where C_{ij} is a random censoring time independent of Y_{ij}. Additionally, a censoring indicator δ_{ij} is observed, with

$$\begin{aligned} \delta_{ij} &= 1 \qquad \text{if } T_{ij} = Y_{ij} \\ &= 0 \qquad \text{if } T_{ij} = C_{ij}. \end{aligned}$$

We consider a general mixed-effects proportional hazards model for T_{ij} as:

$$\begin{aligned} \lambda(t_{ij}|\boldsymbol{\beta}_j, \boldsymbol{\omega}_j) &= \lambda_0(t_{ij})\nu_{ij}e^{\boldsymbol{X}'_{ij}\boldsymbol{\beta}_j} \\ &= \lambda_0(t_{ij})e^{\boldsymbol{X}'_{ij}\boldsymbol{\beta}_j + \boldsymbol{Z}'_{ij}\boldsymbol{\omega}_j} \\ &= \lambda_0(t_{ij}) \times \exp(\boldsymbol{X}'_{ij}\boldsymbol{\beta}_j + \boldsymbol{Z}'_{ij}\boldsymbol{\omega}_j), \end{aligned} \tag{2.1}$$

where, $\lambda_0(t)$ is the baseline hazard function, $\boldsymbol{\beta}_j$ is a vector of cluster-specific fixed-effects corresponding to a vector of covariates $\boldsymbol{X}_{ij}$ and $\boldsymbol{\omega}_j$ is a vector of random effects associated with a vector of unobserved covariates $\boldsymbol{Z}_{ij}$. The key assumptions for this model are:

1. The random effects $\boldsymbol{\omega}_j$ are assumed to be randomly distributed with mean 0 and variance-covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\Theta})$, which depends in a d-dimentional vector of parameters $\boldsymbol{\Theta} = (\theta_1, \theta_2, ..., \theta_d)$. In our shared frailty case, we consider simplest case $d = 1$. Therefore, we consider, $\boldsymbol{\Theta} = \theta$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\Theta}) = \boldsymbol{\Sigma}(\theta)$

2. $\boldsymbol{\Sigma}$ is a positive definite matrix. This provides a rich class of models for the random effects. For example, setting $\boldsymbol{\Sigma} = \theta \times \mathbb{I}$ results in a shared frailty model.

3. The density function of the $\boldsymbol{\omega}_j$ which, except for θ, is assumed to be known, and will be denoted by $f_\omega(\boldsymbol{\omega}_j)$. Different models available in the literature are gamma, Log-normal, Positive stable, Power variance function, Inverse Gaussian, Compound Poisson, Uniform, Threshold distribution and Franks Family, respectively. However, in the current work, we considered the Gamma and Log-normal frailty model.

4. Each individual can belong to only one family or cluster.

2.2 Issues on Construction of the Likelihoods

Model (2.1) can be seen as a linear mixed-effects model on the log-hazard scale. The estimation of the parameters $\boldsymbol{\beta}_j$ and θ from the observed data on T_{ij} is our main interest.

2.2.1 Conditional Likelihood

Assuming the conditional independence of the observations within cluster given $\boldsymbol{\omega}_j$, one might modify the conditional log-likelihood for the usual Cox proportional Hazards model (denoted by $\log \mathfrak{L}^{\mathfrak{C}}$) given the observed data as

$$
\begin{aligned}
\log[\mathfrak{L}^{\mathfrak{C}}(\boldsymbol{\beta},\lambda_0,\boldsymbol{\omega})] &= \sum_{j=1}^{q}\sum_{i=1}^{n}\Bigg[\delta_{ij}\Bigg(\log\big(\lambda_0(t_{ij})+\boldsymbol{X}'_{ij}\boldsymbol{\beta}_j+\boldsymbol{Z}'_{ij}\boldsymbol{\omega}_j\big)\Bigg)\\
&\quad -\Lambda_0(t_{ij})\times\exp(\boldsymbol{X}'_{ij}\boldsymbol{\beta}_j+\boldsymbol{Z}'_{ij}\boldsymbol{\omega}_j)\Bigg]\\
&= \sum_{j=1}^{q}\log\Big[\mathfrak{L}_j^{\mathfrak{C}}(\boldsymbol{\beta}_j,\lambda_0,\boldsymbol{\omega}_j)\Big], \qquad (2.2)
\end{aligned}
$$

is the conditional log-likelihood for the observed data in the jth cluster, and $\boldsymbol{\beta}$, $\boldsymbol{\omega}$ denote the vectors resulting from "stacking" vectors $\boldsymbol{\beta}_j$ and $\boldsymbol{\omega}_j$ for all clusters, respectively. Also, $\Lambda_0(t_{ij})$ is the cumulative hazard function.

2.2.2 Marginal Likelihood

The marginal likelihood (denoted by $\mathfrak{L}^{\mathfrak{M}}$) of the observed data for all clusters can then be expressed as

$$
\begin{aligned}
\mathfrak{L}^{\mathfrak{M}}(\boldsymbol{\beta},\theta,\lambda_0) &= \prod_{j=1}^{q}\int \mathfrak{L}_j^{\mathfrak{A}}(\boldsymbol{\beta}_j,\theta,\lambda_0,\boldsymbol{\omega}_j)d\boldsymbol{\omega}_j\\
&= \prod_{j=1}^{q}\int f(\boldsymbol{\omega}_j)\times e^{\log[\mathfrak{L}_j^{\mathfrak{C}}(\boldsymbol{\beta}_j,\lambda_0,\boldsymbol{\omega}_j)]}d\boldsymbol{\omega}_j, \qquad (2.3)
\end{aligned}
$$

where

$$
\mathfrak{L}_j^{\mathfrak{A}}(\boldsymbol{\beta}_j,\theta,\lambda_0,\boldsymbol{\omega}_j) = f(\boldsymbol{\omega}_j)\times\exp\Bigg(\log\Big[\mathfrak{L}_j^{\mathfrak{C}}(\boldsymbol{\beta}_j,\lambda_0,\boldsymbol{\omega}_j)\Big]\Bigg). \qquad (2.4)
$$

Equation (2.4) can be regarded as the likelihood of the augmented data (denoted by $\mathfrak{L}^{\mathfrak{A}}$) for cluster j, treating $\boldsymbol{\omega}_j$ as additional observations.

2.2.3 Augmented Likelihood

In consequence of the equation (2.4), we have–

$$
\mathfrak{L}^{\mathfrak{A}}(\boldsymbol{\beta},\theta,\lambda_0,\boldsymbol{\omega}) = \prod_{j=1}^{q}\mathfrak{L}_j^{\mathfrak{A}}(\boldsymbol{\beta}_j,\theta,\lambda_0,\boldsymbol{\omega}_j), \qquad (2.5)
$$

is the likelihood of the augmented data for all clusters.
One might consider using the likelihood function (2.3) in the inference on $\boldsymbol{\beta}$ and θ. There are two major problems with using (2.3) for inference: it depends on the baseline hazard function λ_0 and the integral will usually be multi-dimentional, unless a very simple model is considered, and in general will not be available in a closed form.

3 Software Implementation Details

We used `genfrail` function in `frailtySurv` package to generate the dataset suitable for frailty model fitting. As for the model fitting, we used the following implementations:

1. **`Cox.EM.G`** and **`Cox.REML.N`**: `coxph( ..., frailty(distribution="gamma", method = "em", ...))` and `coxph( ..., frailty(distribution="gaussian", method = "reml" ...))` functions respectively in `survival` package

2. **`Cox.ML.N`**: `coxme` function in `coxme` package

3. **`Cox.Sp.G`** and **`Cox.Sp.N`**: `fitfrail(..., frailty="gamma")` and `fitfrail(..., frailty="lognormal")` functions respectively in `frailtySurv` package

4. **`Weib.JM.G`**: `weibull.frailty` function in `JM` package

5. **`Weib.ML.G`** and **`Weib.ML.N`**: `parfm(..., frailty = "gamma")` and `parfm(..., frailty = "lognormal")` functions respectively in `parfm` package

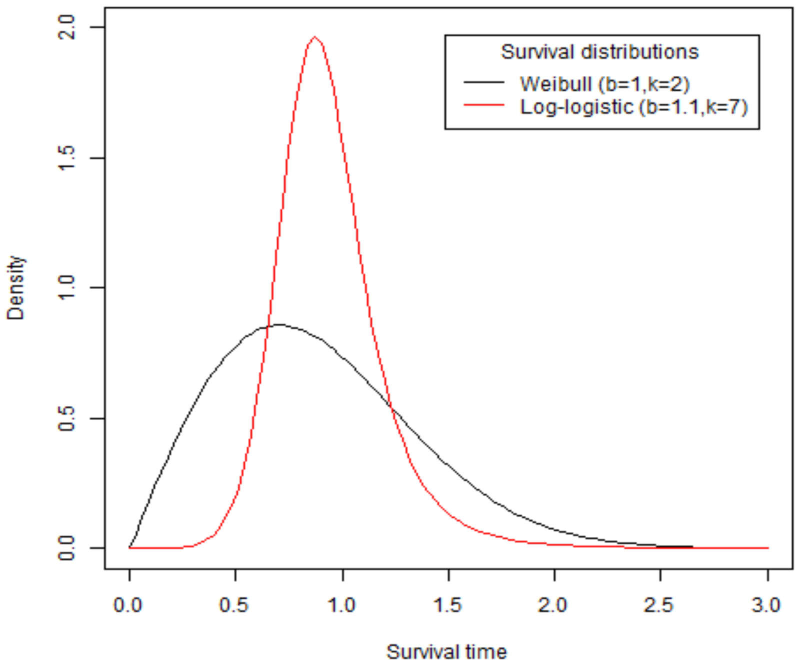
4 Simulation Characteristics

Web-Table 4.1: Simulating survival times from Weibull and log-logistic cumulative baseline hazard.

Characteristics	Weibull distribution	Log-logistic distribution
Shape parameter	$k = 2$	$k = 7$
Scale parameter	$b = 1$	$b = 1.1$
Density $f(t)$	$bkt^{k-1}\exp(-bt^k)$	$\frac{bk(bt)^{k-1}}{[1+(bt)^k]^2}$
Hazard function $h(t)$	bkt^{k-1}	$\frac{bk(bt)^{k-1}}{[1+(bt)^k]}$
Cumulative hazard function $H(t)$	bt^k	$\log[1 + (bt)^k]$

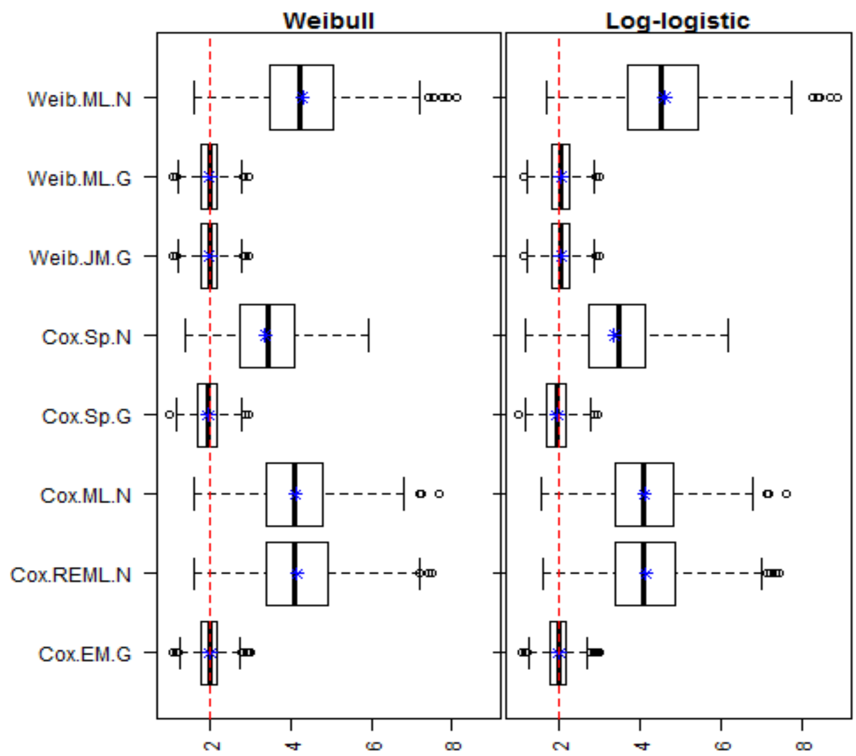
Web-Table 4.2: Simulating frailties from gamma, log-normal or inverse Gaussian distributions.

Distribution	$f(\theta)$	Parameters
gamma	$\beta^\alpha \theta^{\alpha-1} \exp(-\beta\theta)/\Gamma(\alpha)$	shape α rate β
log-normal	$1/(\sigma\sqrt{2\pi}) \exp[-(\log\theta - \mu)^2/2\sigma^2]$	mean μ SD σ
inverse Gaussian	$(\lambda/2\pi\theta^3) \exp(-\lambda(\theta-\mu)^2/2\mu^2\theta)$	shape λ, mean μ

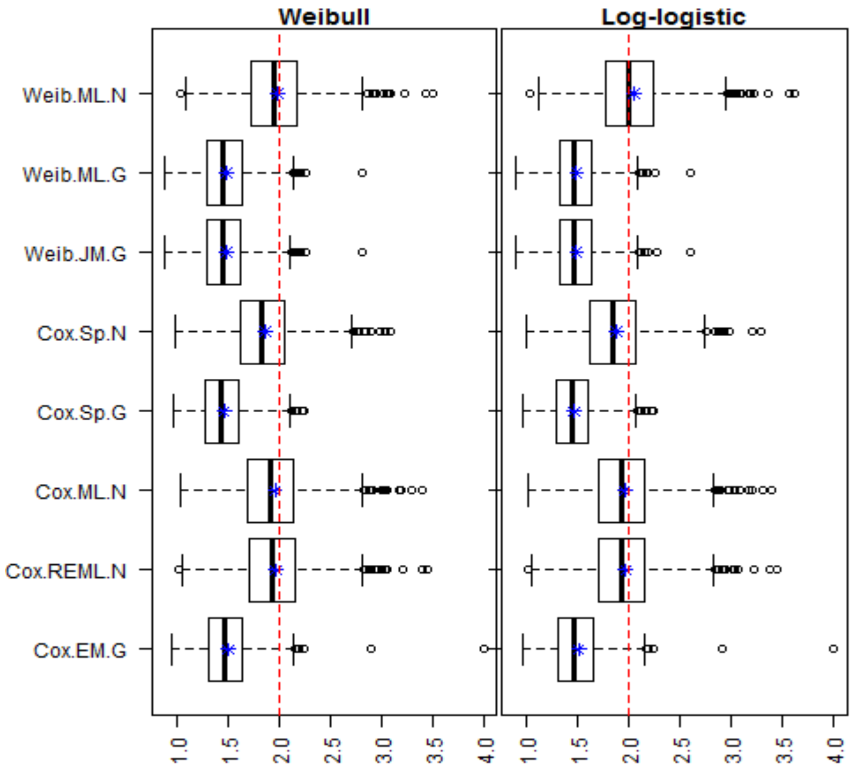


Web-Figure 4.1: Survival time densities

5 Chance in Frailty distributions and their effect on θ

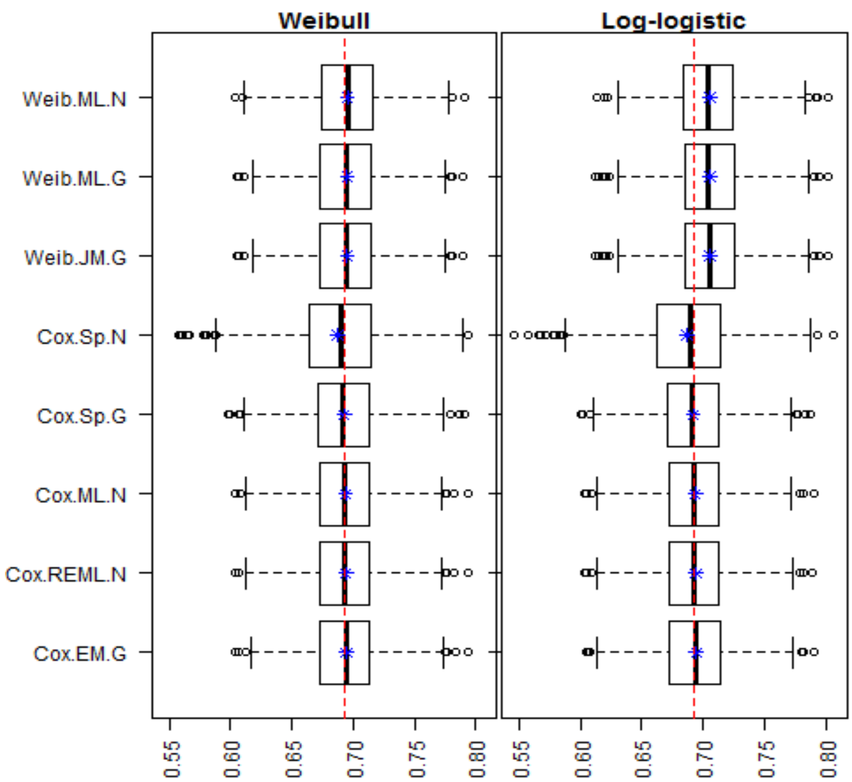


Web-Figure 5.1: θ estimates when the frailty was generated from a Gammma distribution ($N = 60$, censoring 15%).

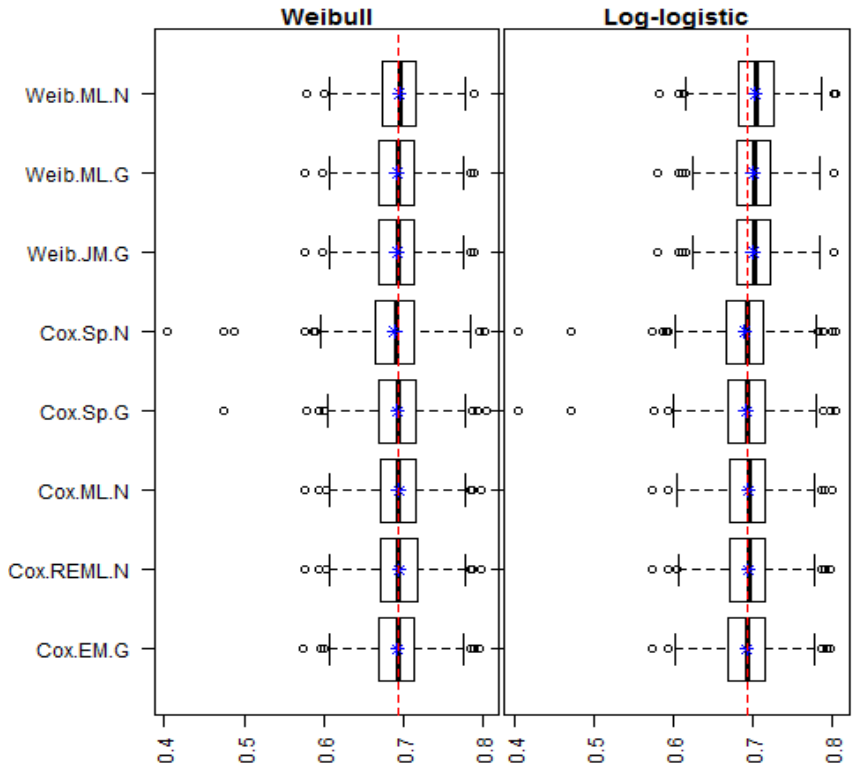


Web-Figure 5.2: θ estimates when the frailty was generated from a log-normal distribution ($N = 60$, censoring 15%).

6 Chance in Frailty distributions and their effect on β_1

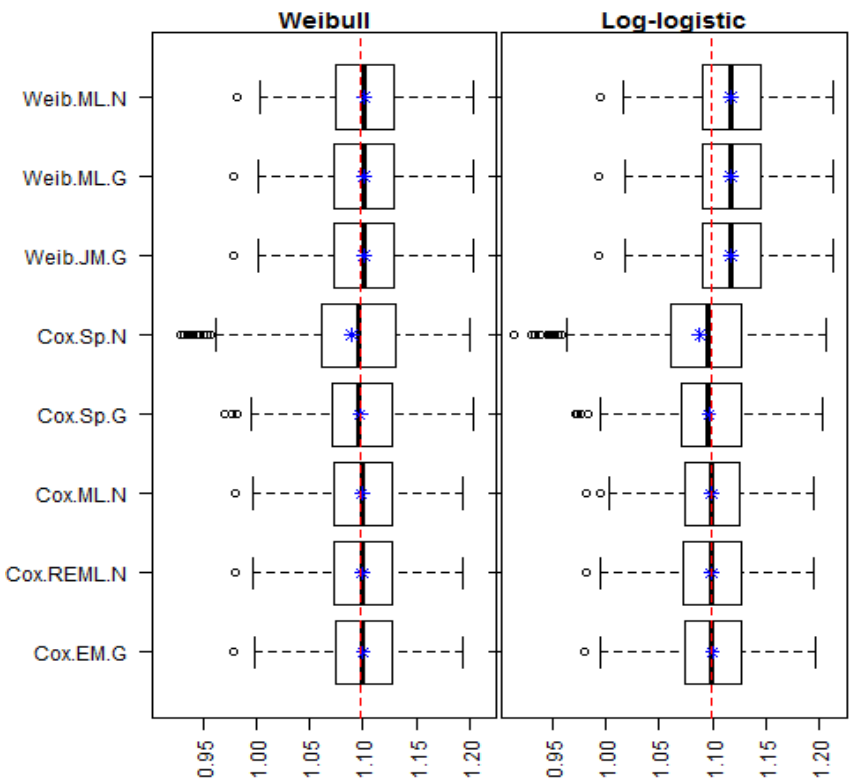


Web-Figure 6.1: β_1 estimates when the frailty was generated from a Gammma distribution ($N = 60$, censoring 15%).

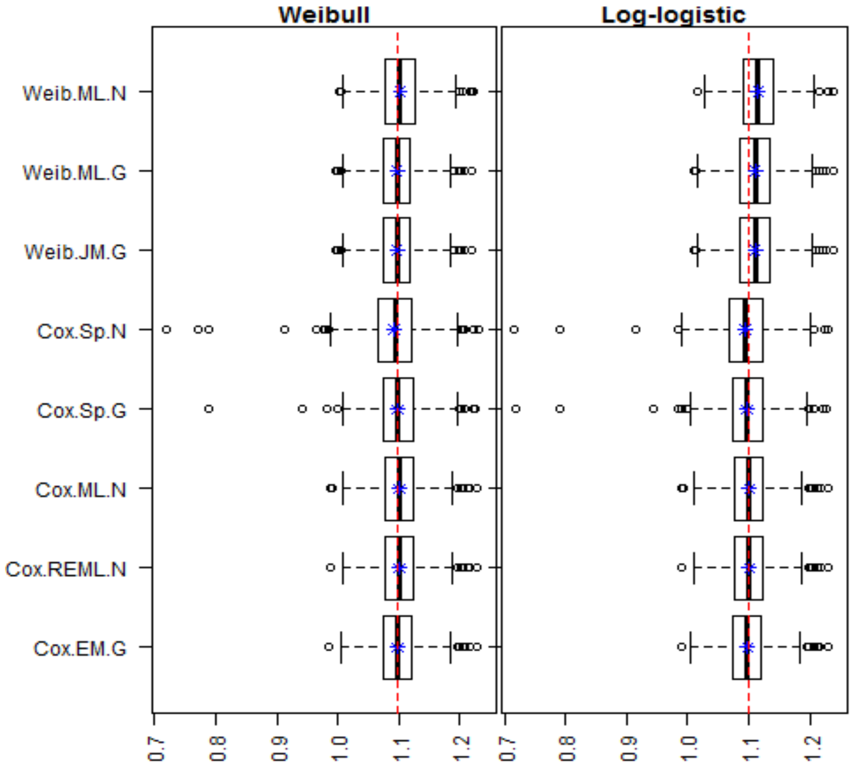


Web-Figure 6.2: β_1 estimates when the frailty was generated from a log-normal distribution ($N = 60$, censoring 15%).

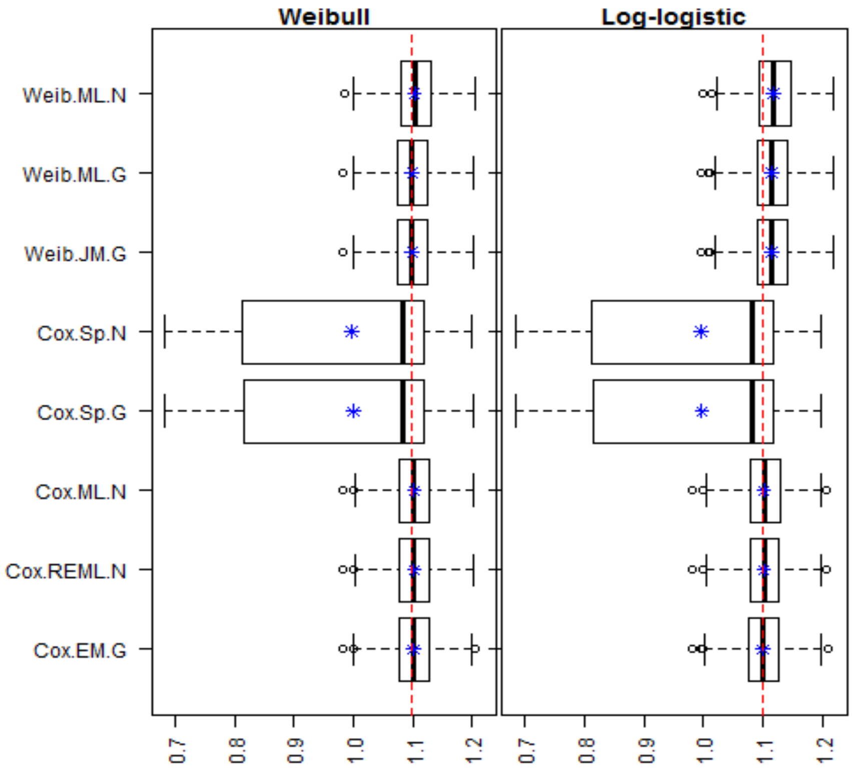
7 Chance in Frailty distributions and their effect on β_2



Web-Figure 7.1: β_2 estimates when the frailty was generated from a Gammma distribution ($N = 60$, censoring 15%).

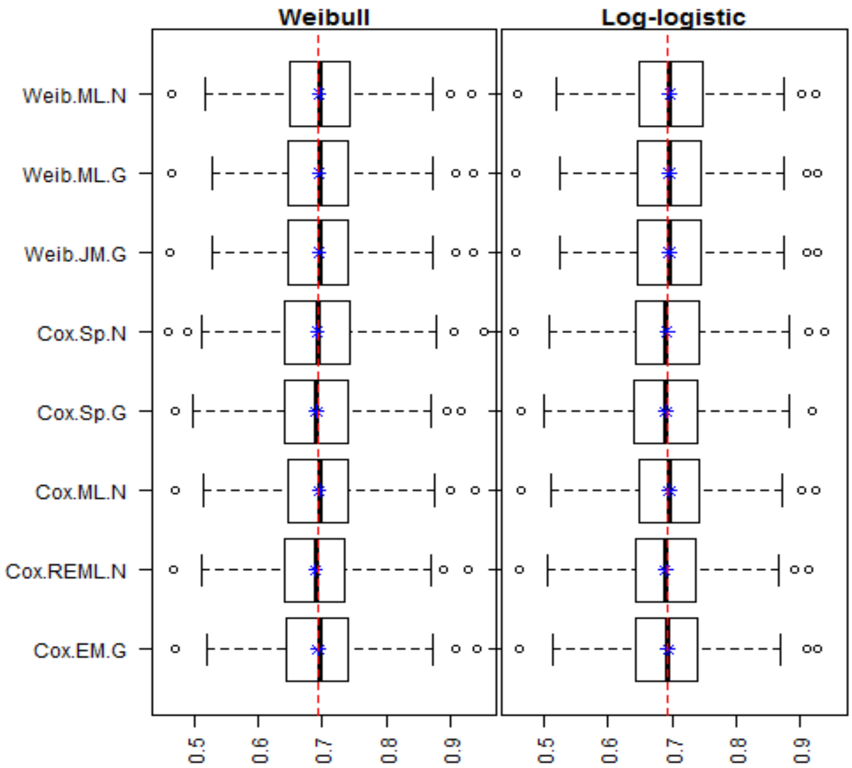


Web-Figure 7.2: β_2 estimates when the frailty was generated from a log-normal distribution ($N = 60$, censoring 15%).

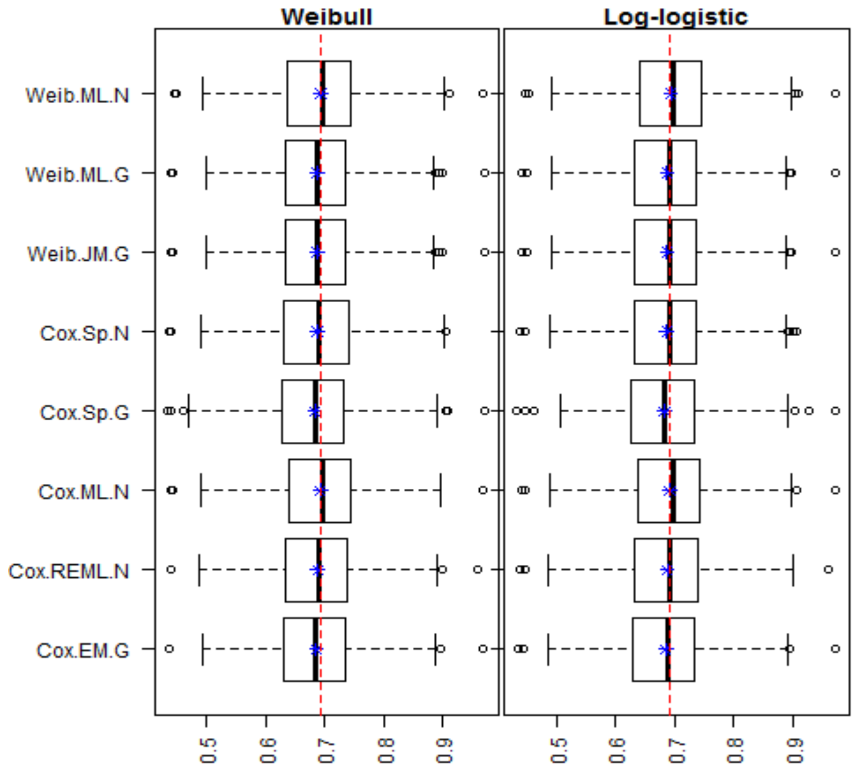


Web-Figure 7.3: β_2 estimates when the frailty was generated from an Inverse-Gaussian distribution ($N = 60$, censoring 15%).

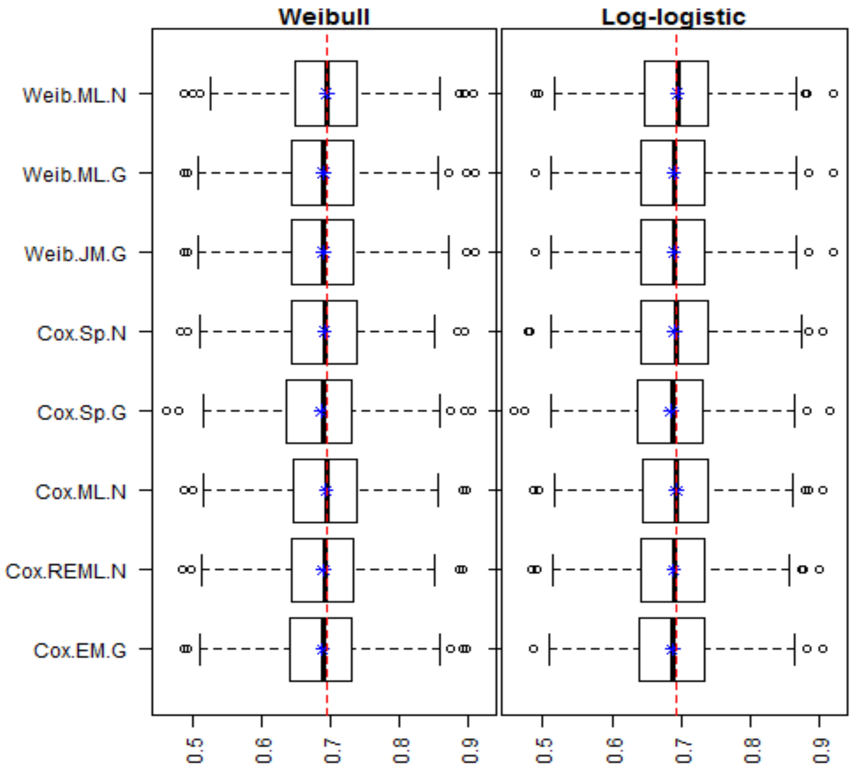
8 Chance in Frailty distributions and their effect on β_1 when censoring is severe



Web-Figure 8.1: β_1 estimates when the frailty was generated from a Gammma ($N = 60$, censoring 85%).

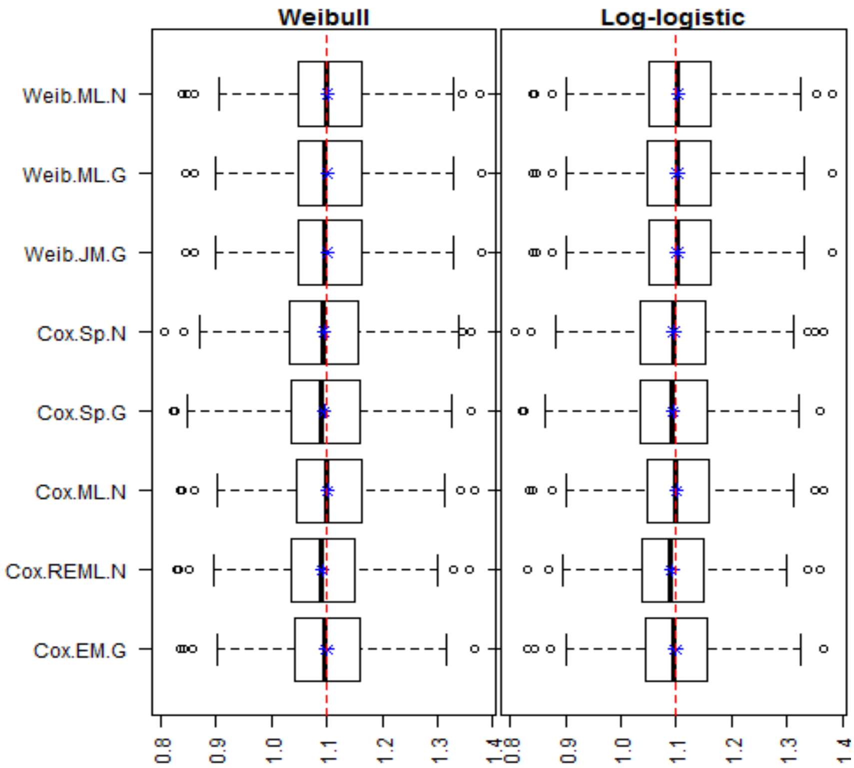


Web-Figure 8.2: β_1 estimates when the frailty was generated from a log-normal ($N = 60$, censoring 85%).

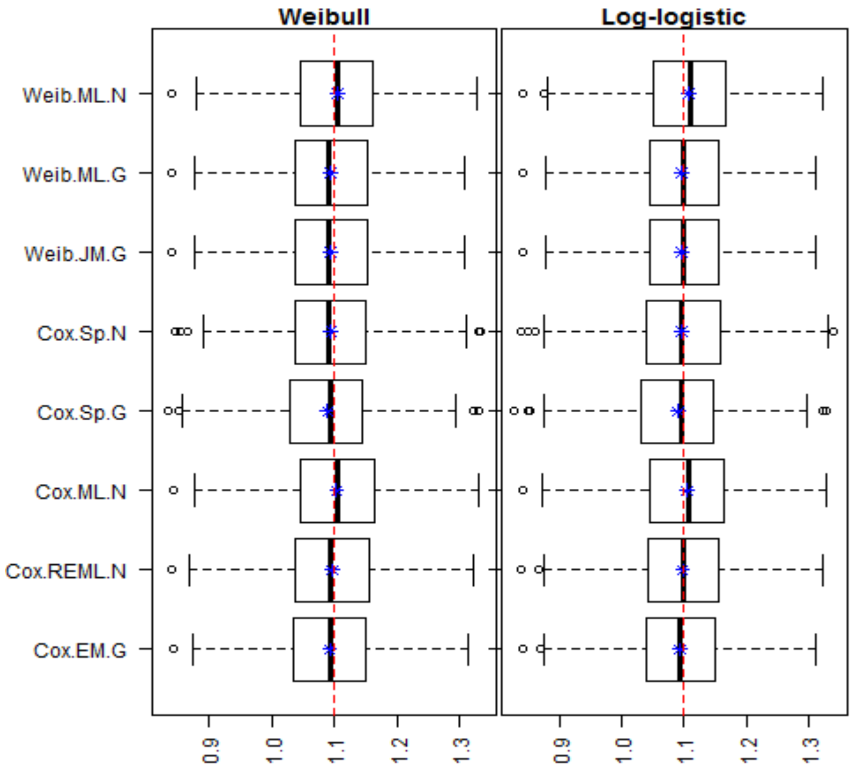


Web-Figure 8.3: β_1 estimates when the frailty was generated from an inverse-Gaussian ($N = 60$, censoring 85%).

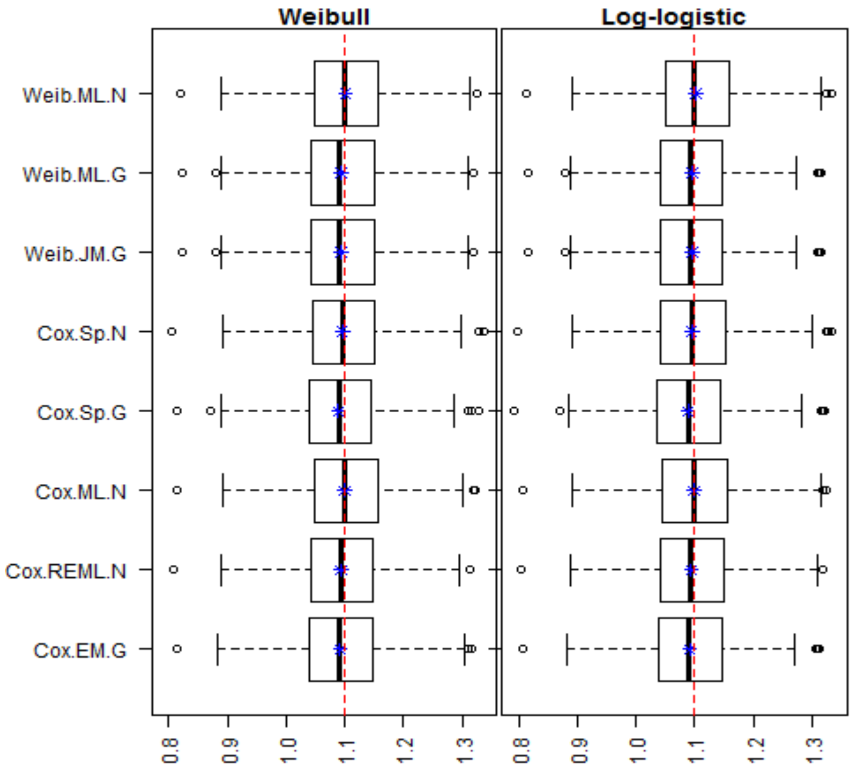
9 Chance in Frailty distributions and their effect on β_2 when censoring is severe



Web-Figure 9.1: β_2 estimates when the frailty was generated from a Gammma ($N = 60$, censoring 85%).

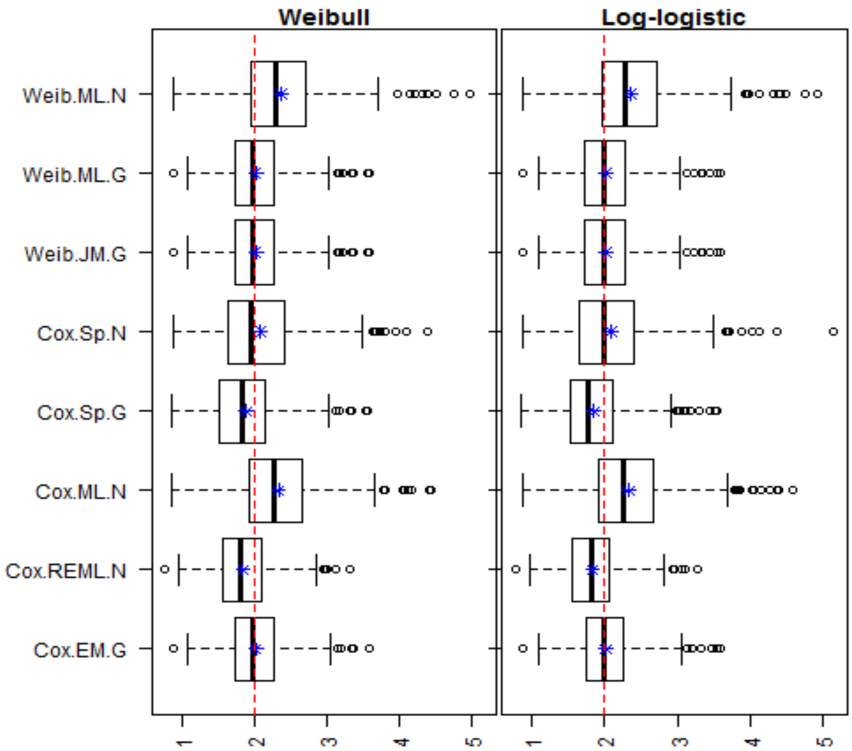


Web-Figure 9.2: β_2 estimates when the frailty was generated from a log-normal ($N = 60$, censoring 85%).

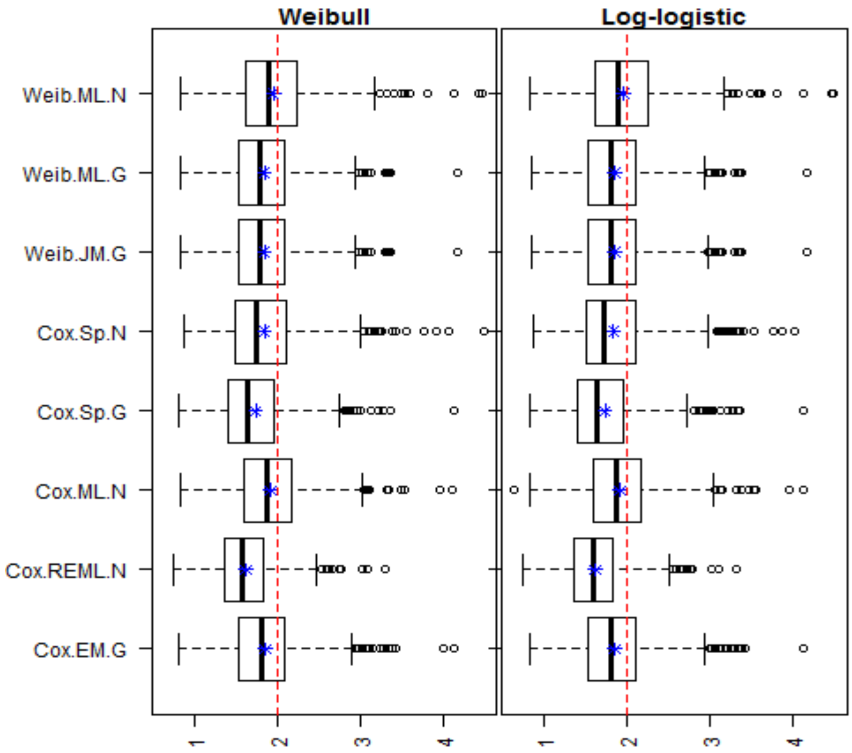


Web-Figure 9.3: β_2 estimates when the frailty was generated from an inverse-Gaussian ($N = 60$, censoring 85%).

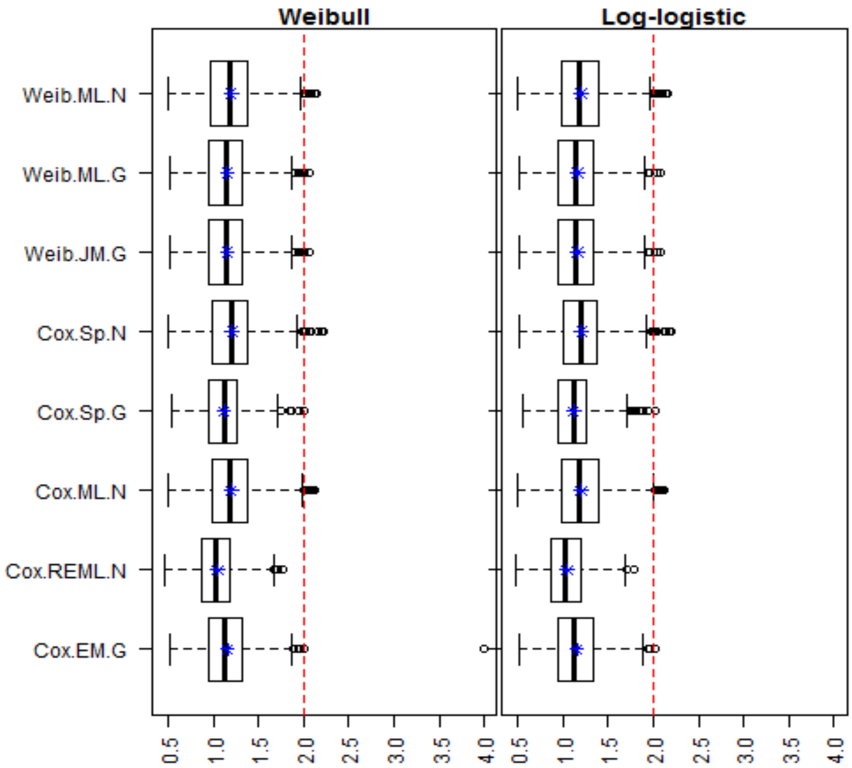
10 Chance in Frailty distributions and their effect on θ when censoring is severe



Web-Figure 10.1: θ estimates when the frailty was generated from a Gammma ($N = 60$, censoring 85%).

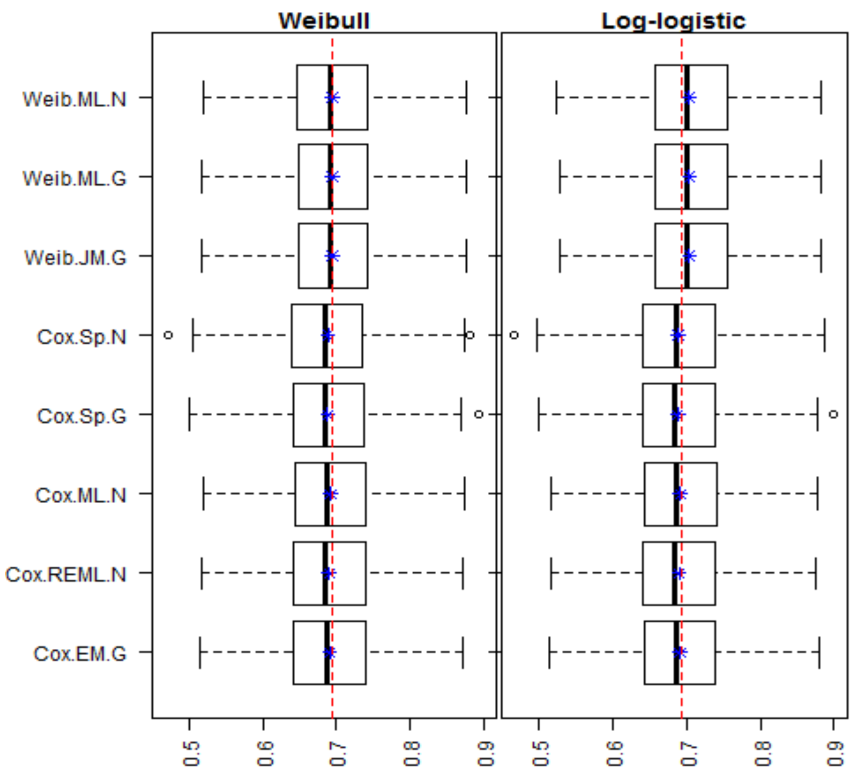


Web-Figure 10.2: θ estimates when the frailty was generated from a log-normal ($N = 60$, censoring 85%).

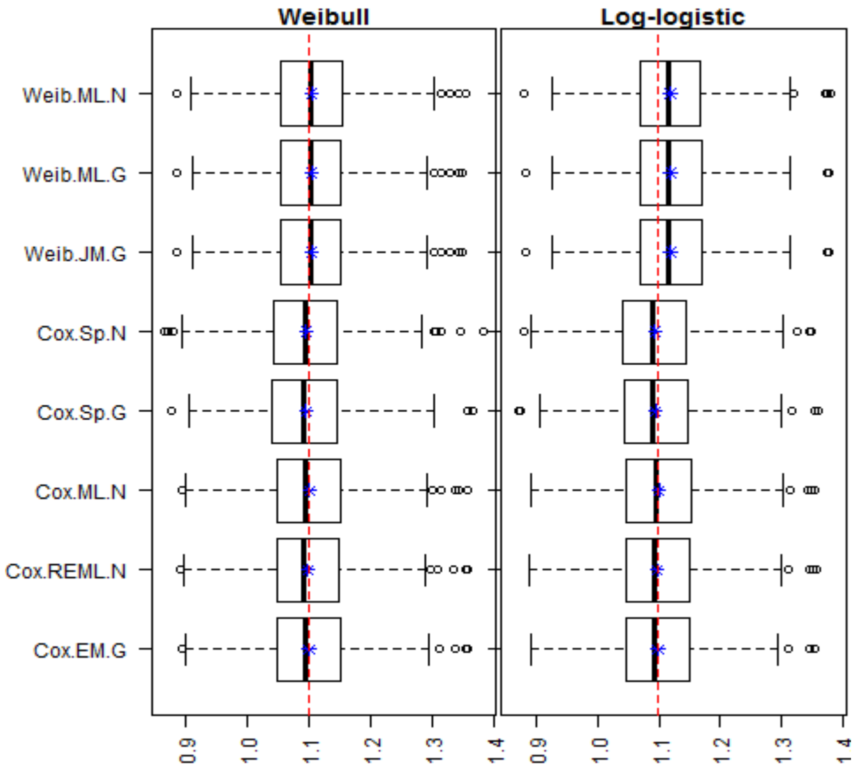


Web-Figure 10.3: θ estimates when the frailty was generated from an inverse-Gaussian ($N = 60$, censoring 85%).

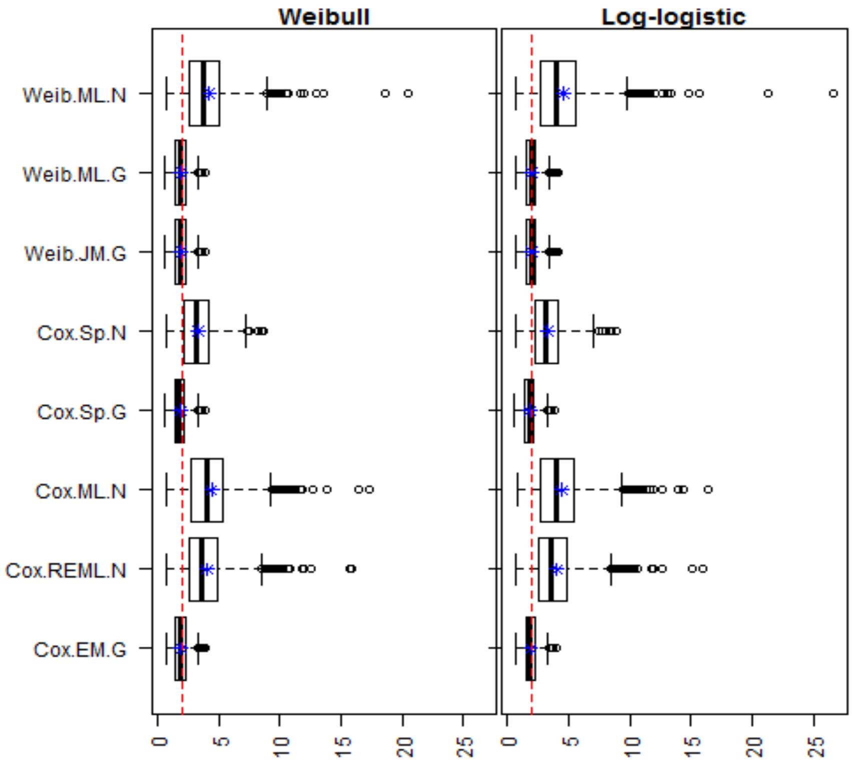
11 Chance in Frailty distributions and their effect on parameters when number of clusters is small



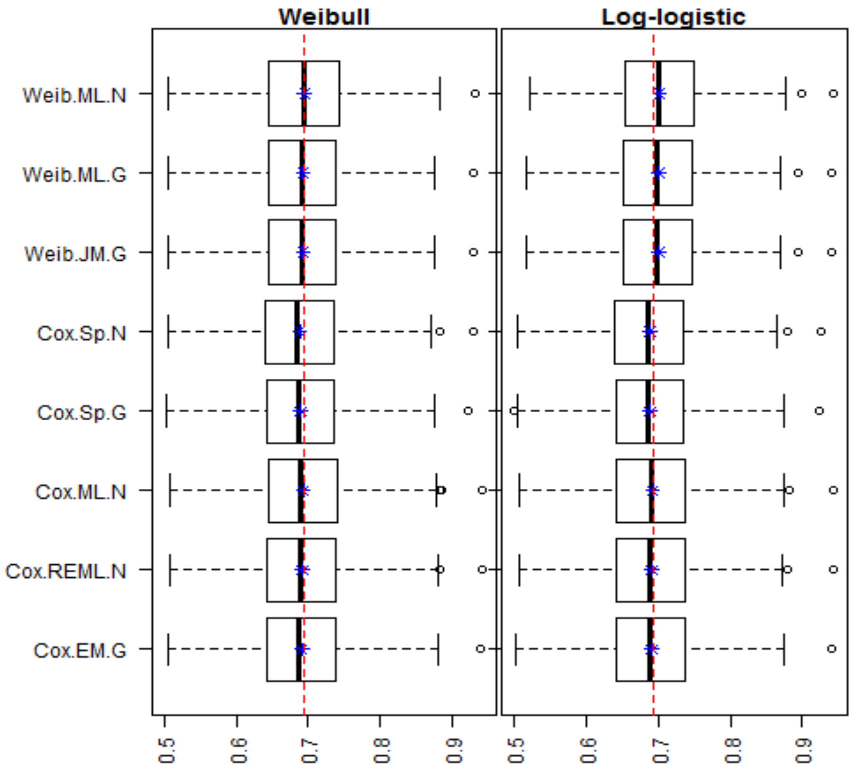
Web-Figure 11.1: β_1 estimates when the frailty was generated from a gamma distribution and number of clusters is small ($N = 15$, censoring 15%).



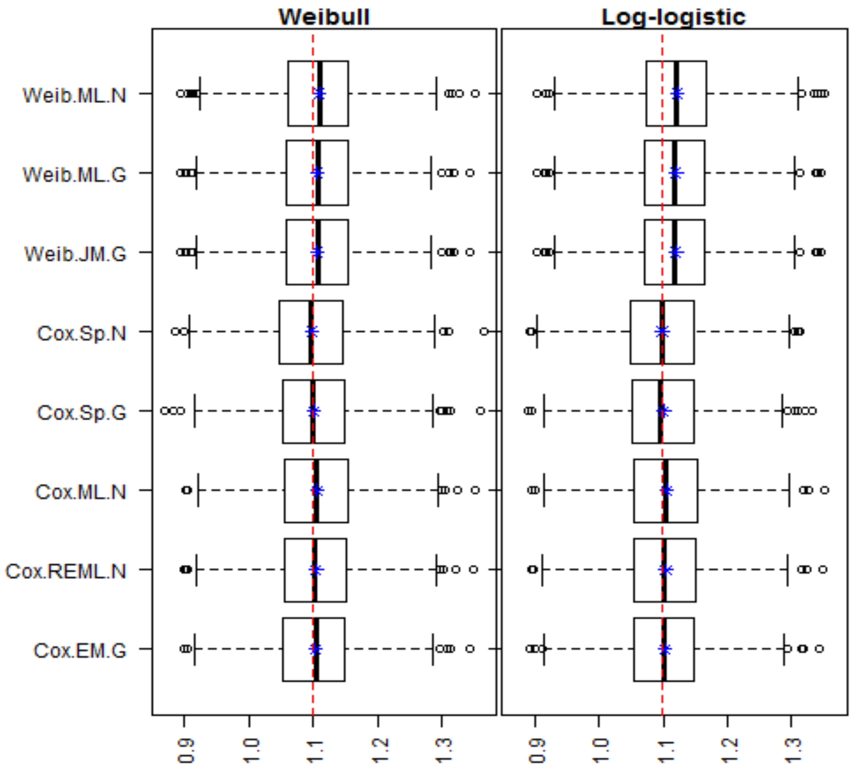
Web-Figure 11.2: β_2 estimates when the frailty was generated from a gamma distribution and number of clusters is small ($N = 15$, censoring 15%).



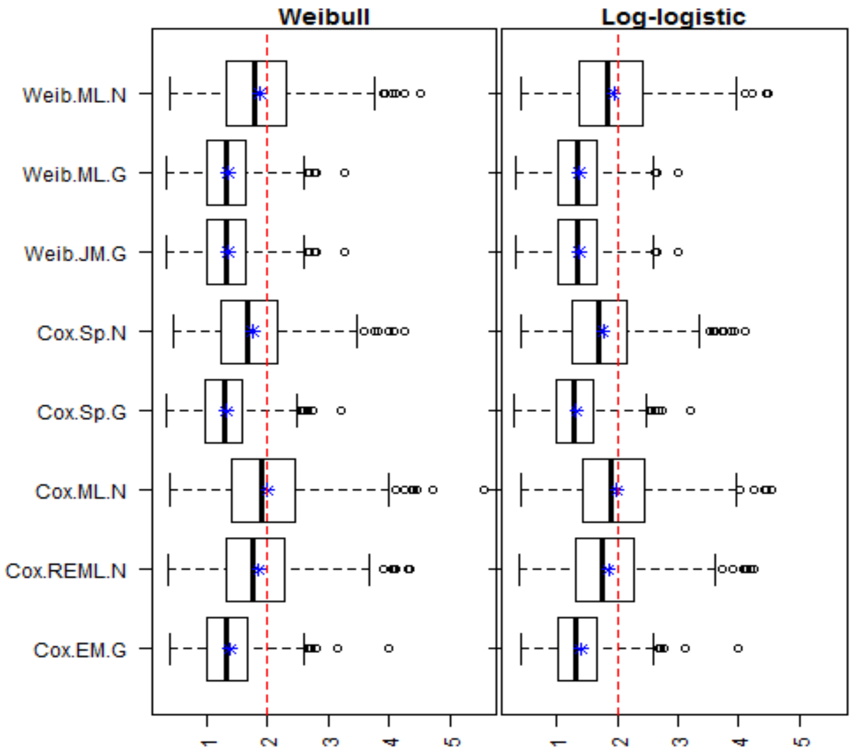
Web-Figure 11.3: θ estimates when the frailty was generated from a gamma distribution and used small number of clusters ($N = 15$, censoring 15%).



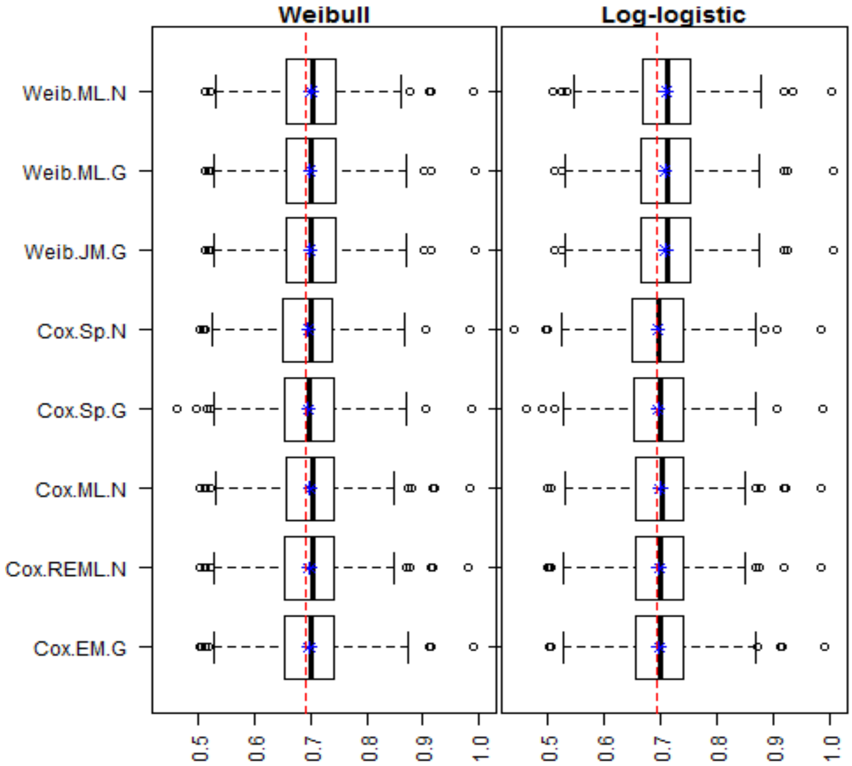
Web-Figure 11.4: β_1 estimates when the frailty was generated from a log-normal and number of clusters is small ($N = 15$, censoring 15%).



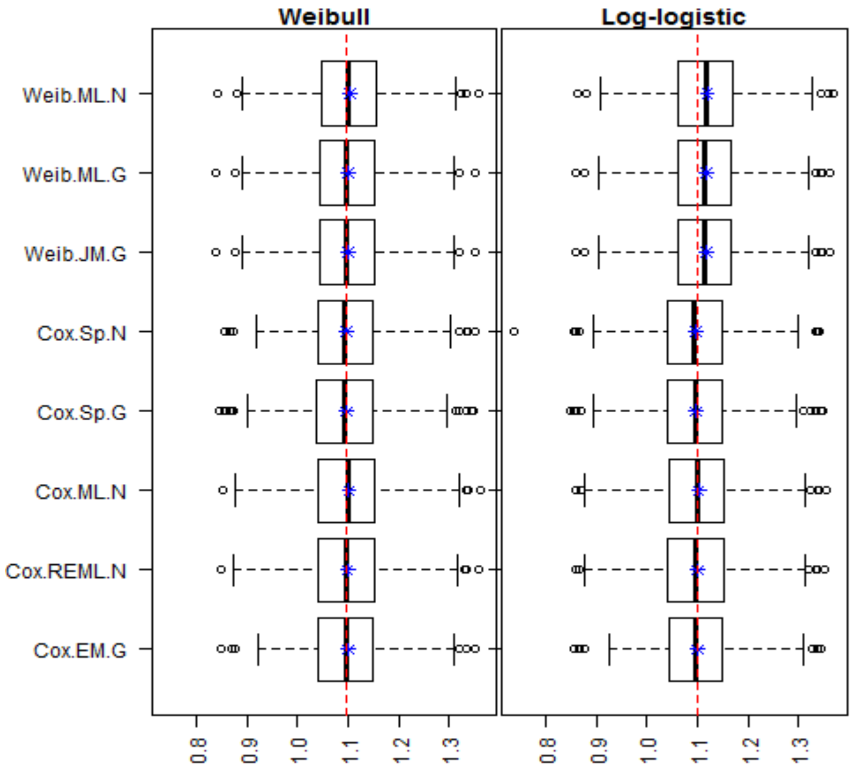
Web-Figure 11.5: β_2 estimates when the frailty was generated from a log-normal and number of clusters is small ($N = 15$, censoring 15%).



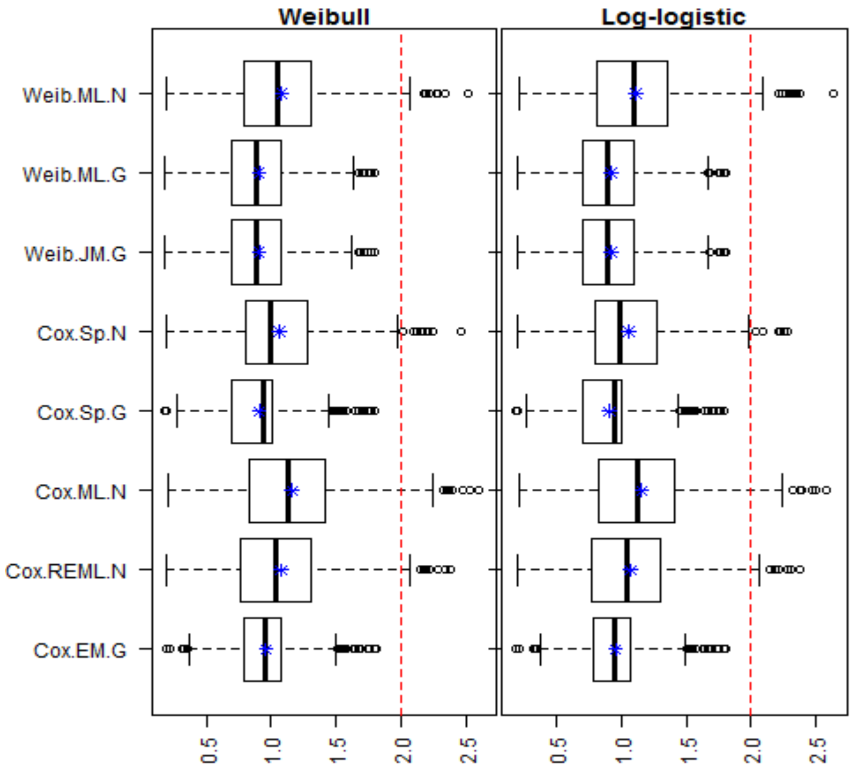
Web-Figure 11.6: θ estimates when the frailty was generated from a log-normal and used small number of clusters ($N = 15$, censoring 15%).



Web-Figure 11.7: β_1 estimates when the frailty was generated from a Inverse Gaussian and number of clusters is small ($N = 15$, censoring 15%).

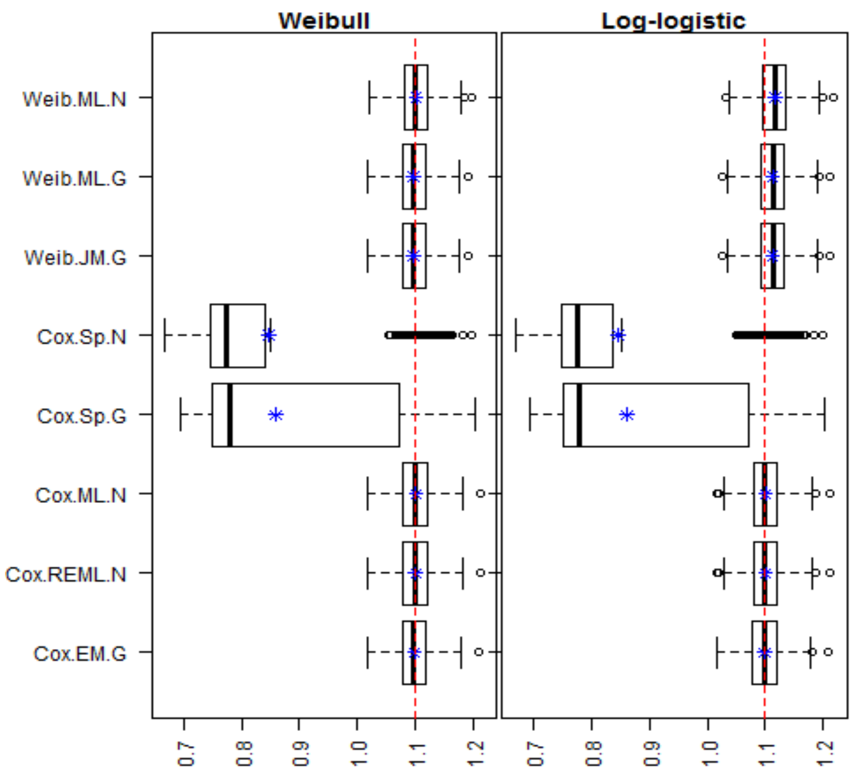


Web-Figure 11.8: β_2 estimates when the frailty was generated from an Inverse Gaussian and number of clusters is small ($N = 15$, censoring 15%).

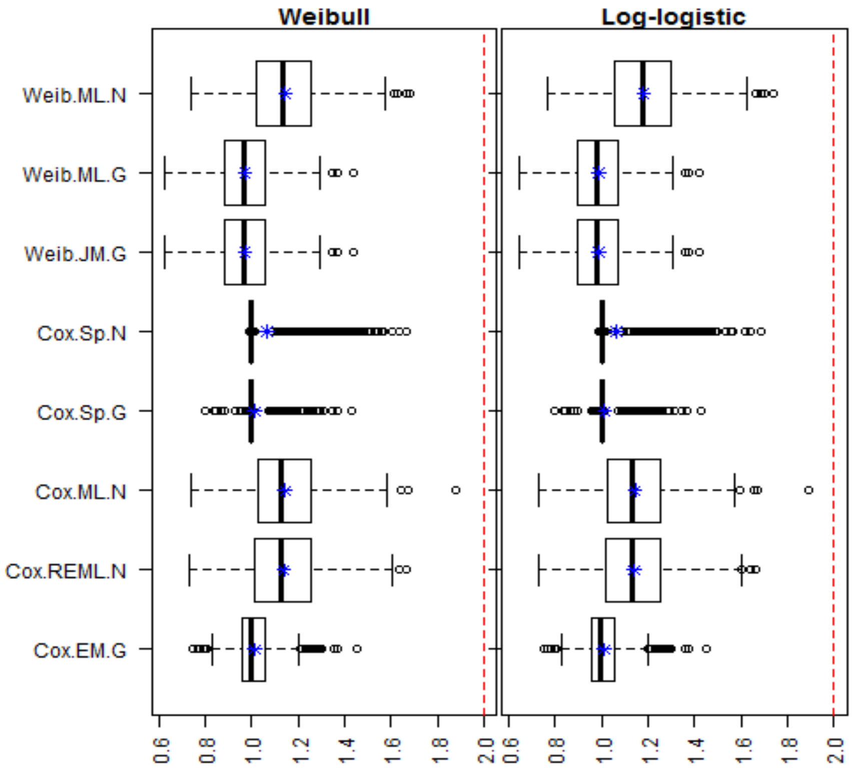


Web-Figure 11.9: θ estimates when the frailty was generated from an Inverse Gaussian and used small number of clusters ($N = 15$, censoring 15%).

12 Chance in Frailty distributions and their effect on parameters when number of clusters is very large



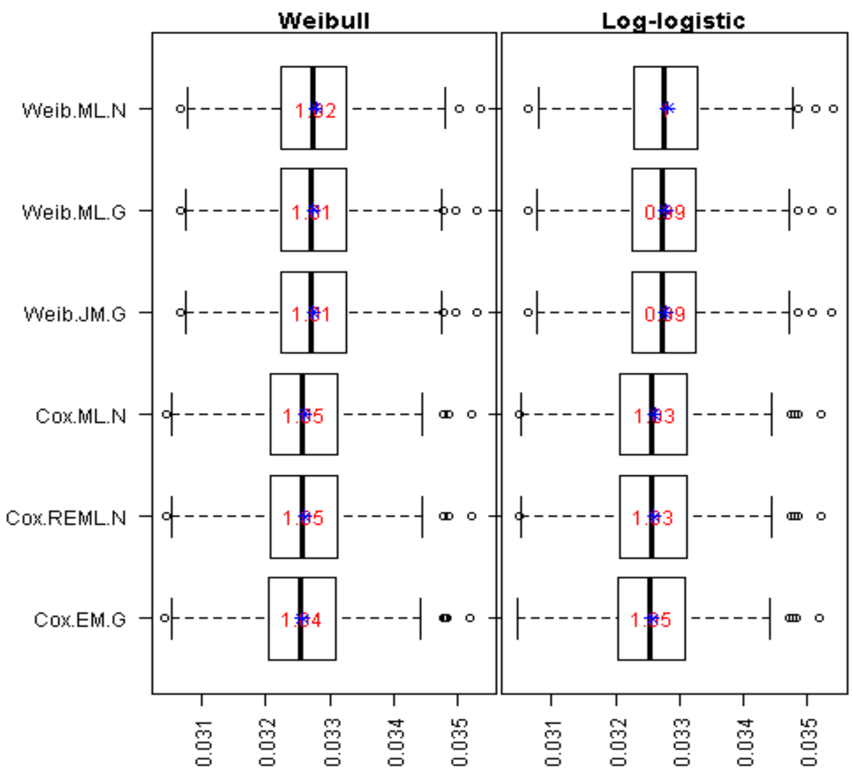
Web-Figure 12.1: β_2 estimates when the frailty was generated from an Inverse Gaussian and used large number of clusters ($N = 90$, censoring 15%).



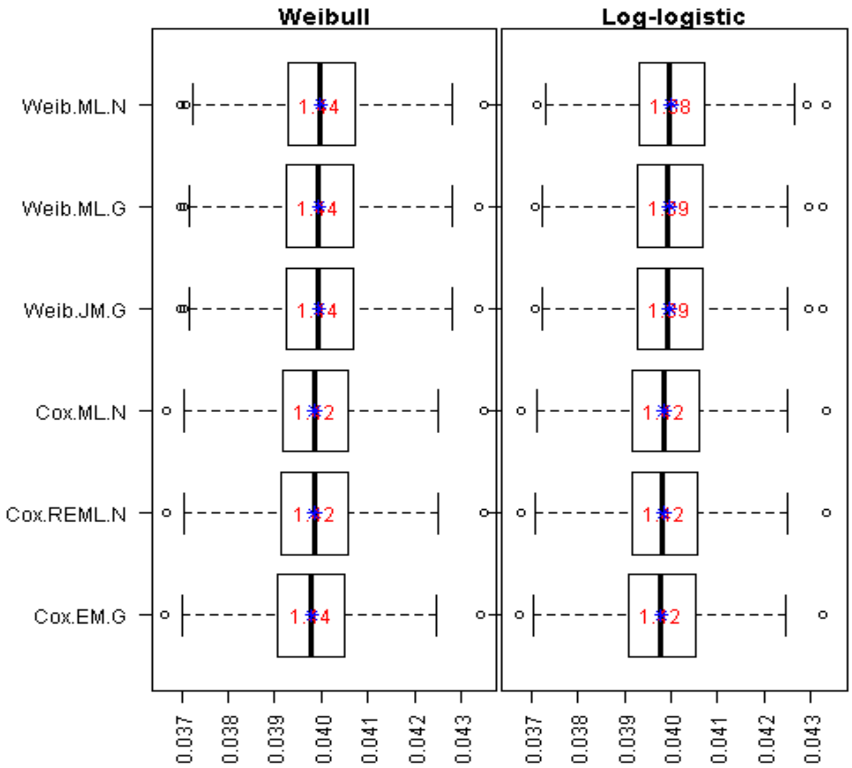
Web-Figure 12.2: θ estimates when the frailty was generated from an Inverse Gaussian and used large number of clusters ($N = 90$, censoring 15%).

13 Effects on estimated SEs

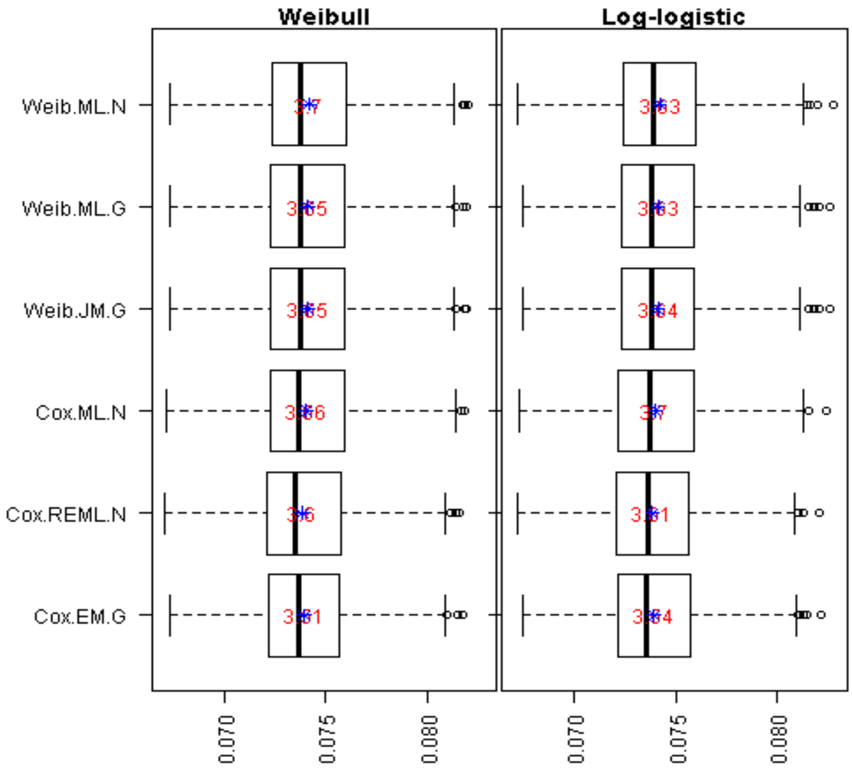
In the following figures, the number marked red are IQR $\times\ 1,000$.



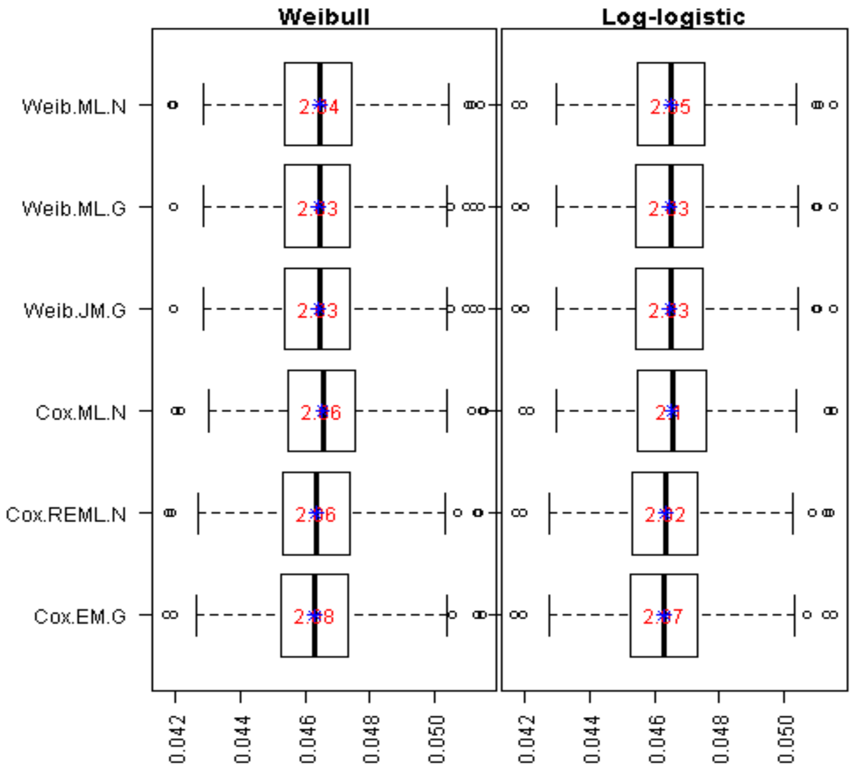
Web-Figure 13.1: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage ($N = 60$, censoring 15%).



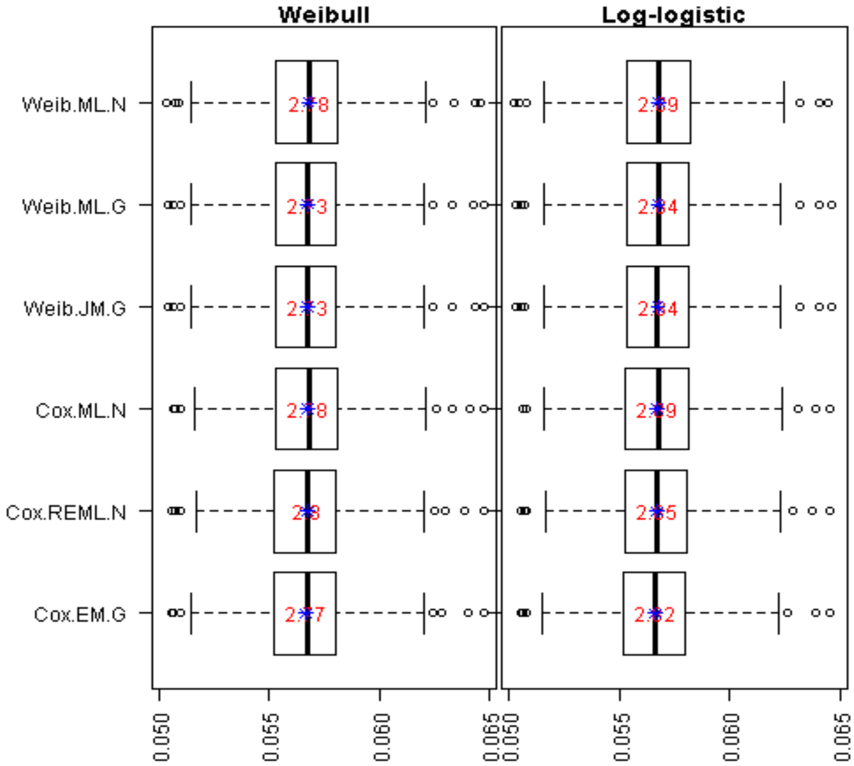
Web-Figure 13.2: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage ($N = 60$, censoring 45%).



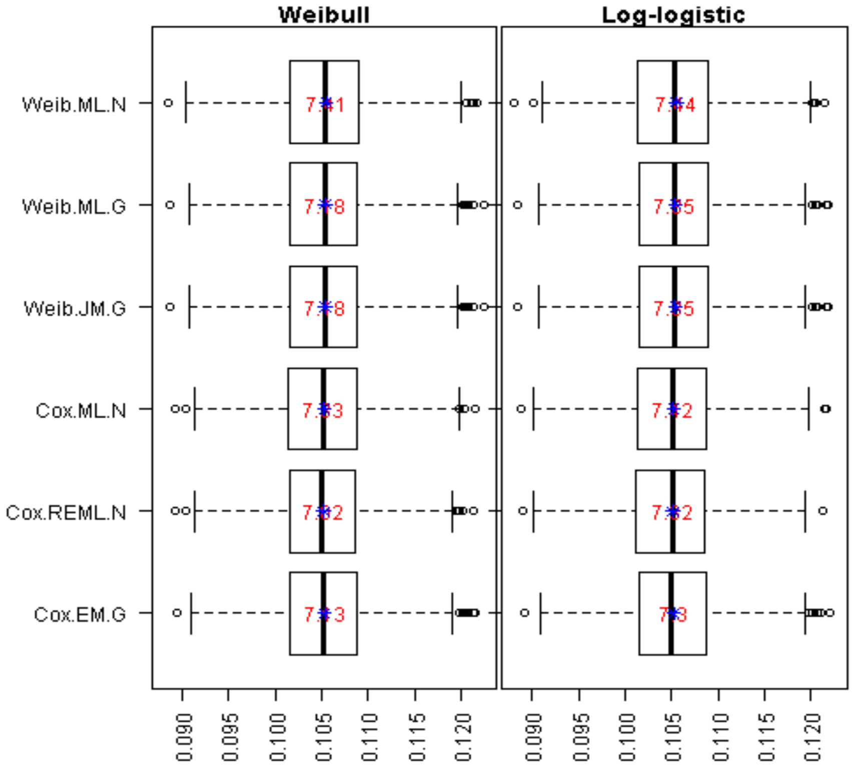
Web-Figure 13.3: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used large censoring percentage (N = 60, censoring 85%).



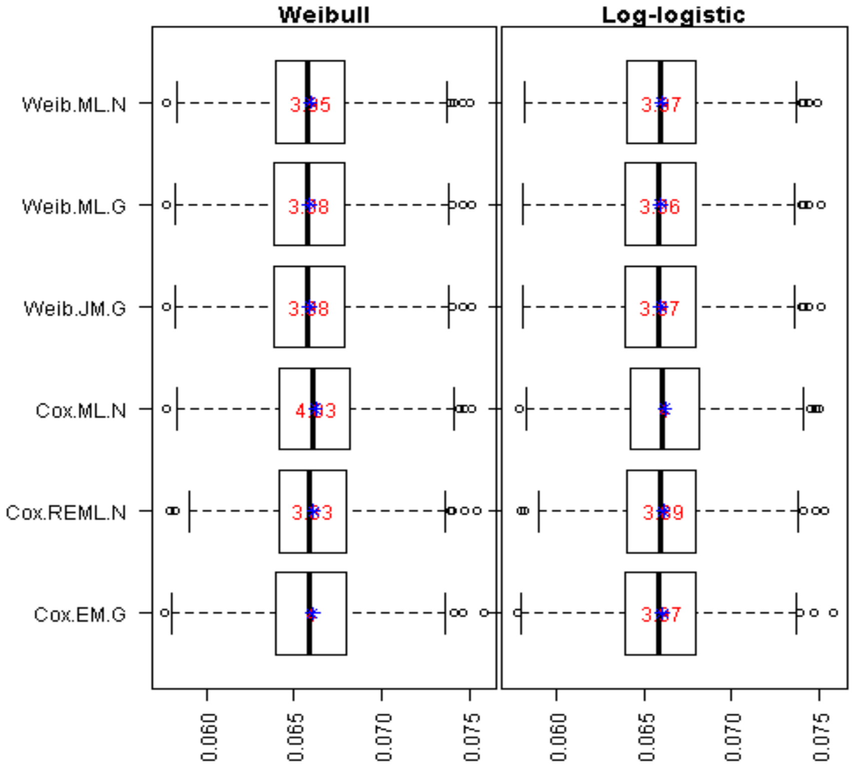
Web-Figure 13.4: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage and smaller number of clusters ($N = 30$, censoring 15%).



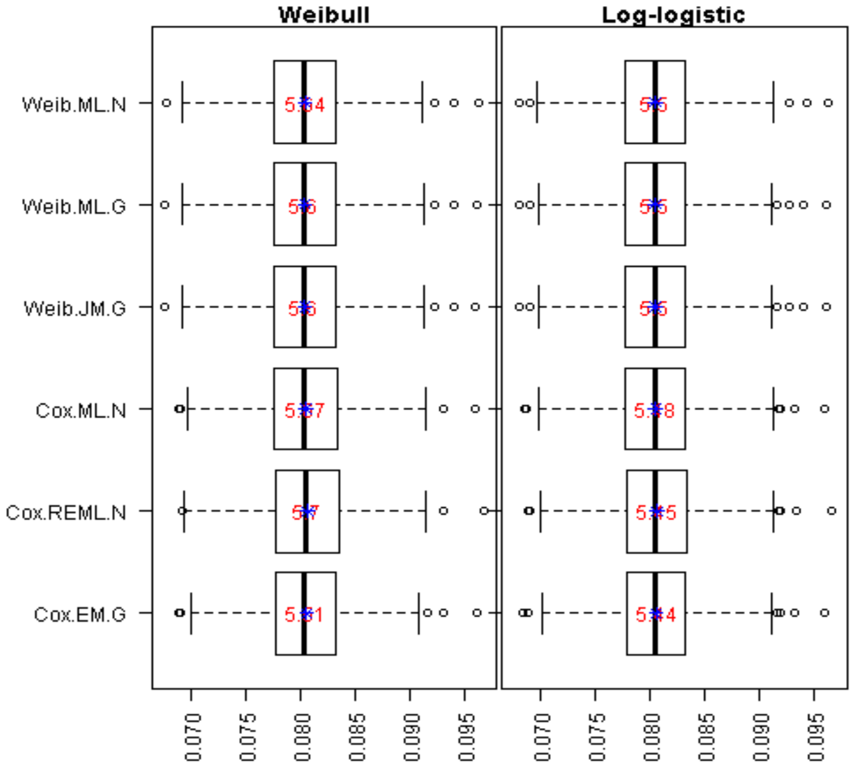
Web-Figure 13.5: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage and smaller number of clusters ($N = 30$, censoring 45%).



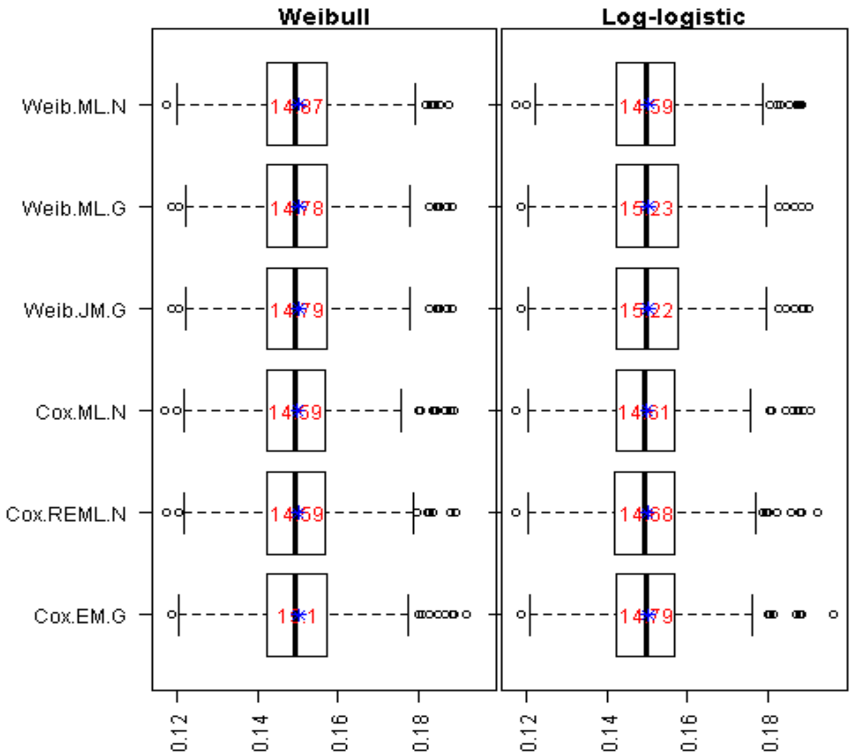
Web-Figure 13.6: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used large censoring percentage and smaller number of clusters ($N = 30$, censoring 85%).



Web-Figure 13.7: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage and smaller number of clusters ($N = 15$, censoring 15%).



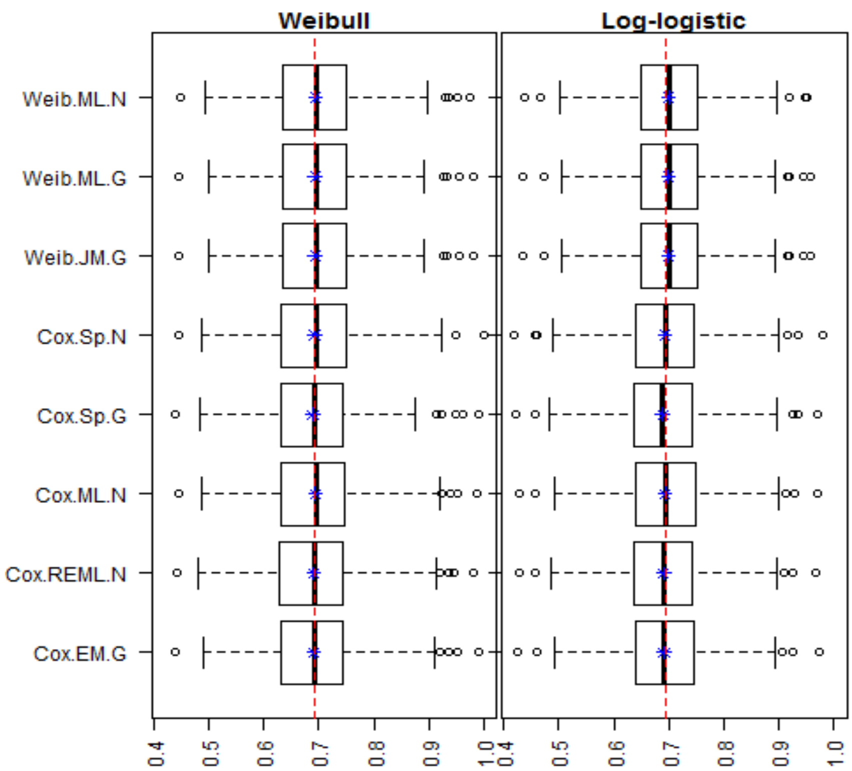
Web-Figure 13.8: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used small censoring percentage and smaller number of clusters ($N = 15$, censoring 45%).



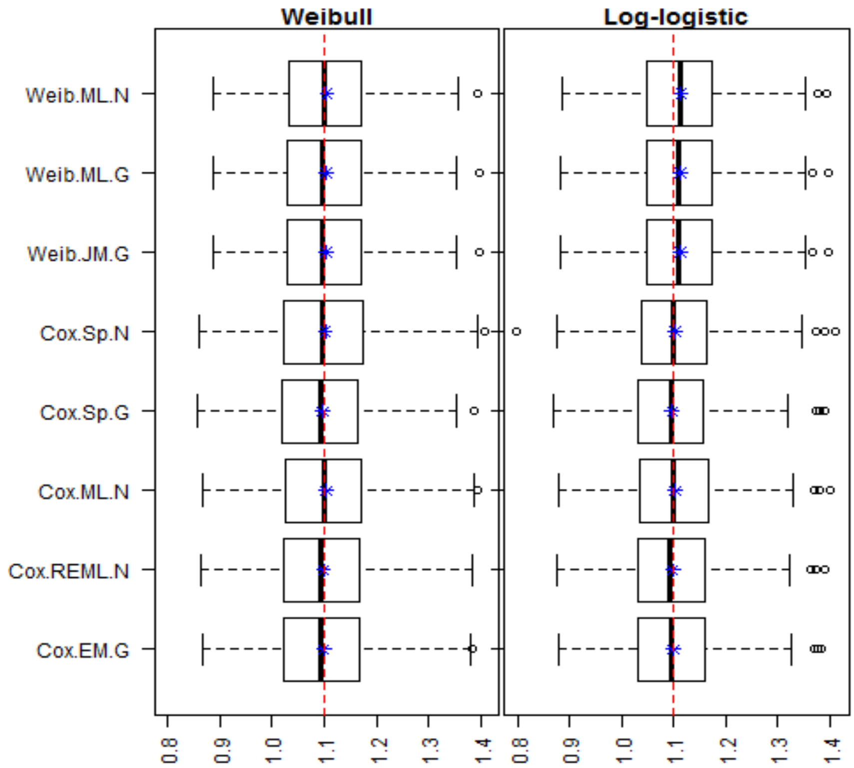
Web-Figure 13.9: SE(β_1) estimates when the frailty was generated from an Inverse Gaussian and used large censoring percentage and smaller number of clusters ($N = 15$, censoring 85%).

13.1 Effect of Changing Number of Clusters

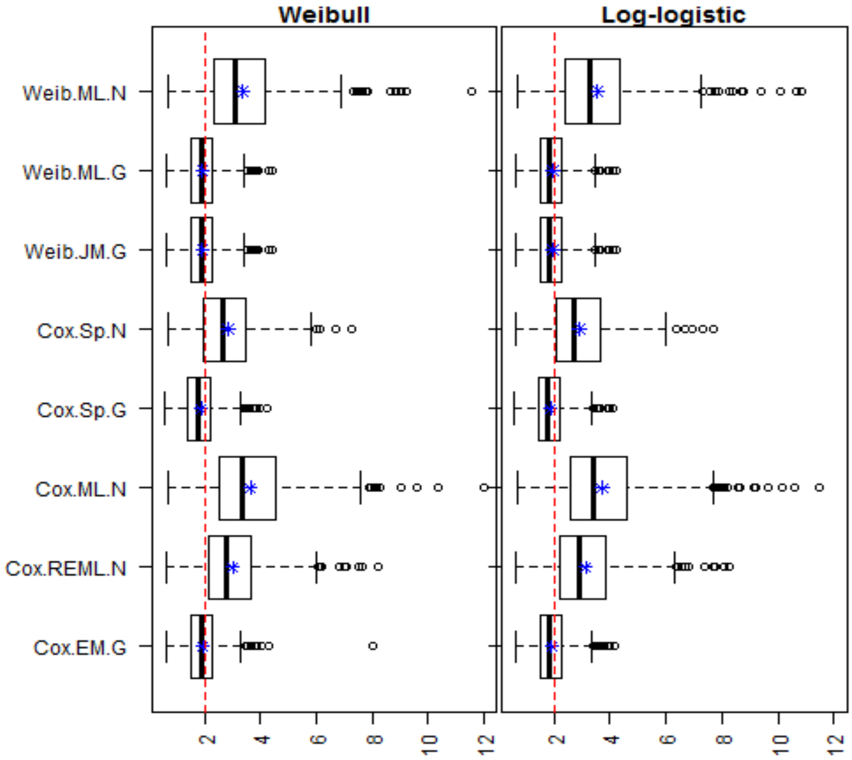
13.1.1 Number of clusters is small ($N = 15$)



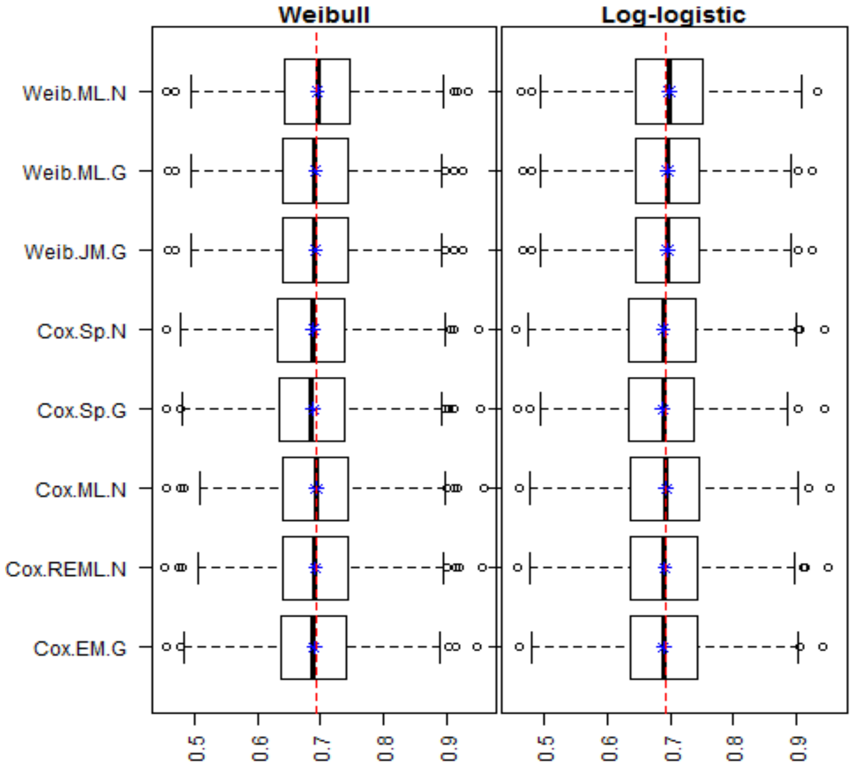
Web-Figure 13.10: β_1 estimates when small number of clusters were used under moderate censoring and frailty generated from a Gamma distribution.



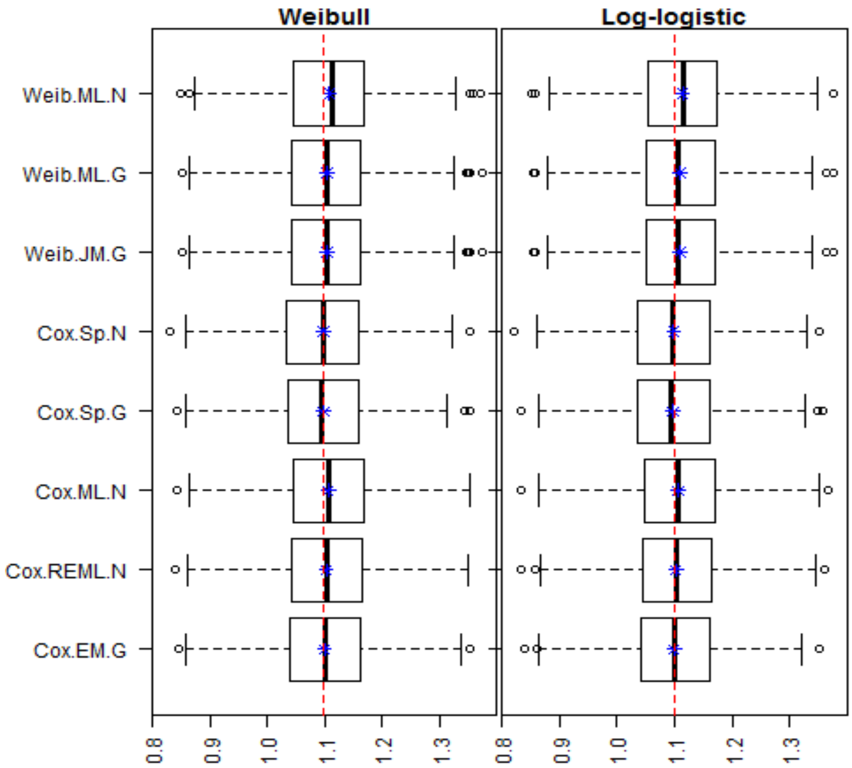
Web-Figure 13.11: β_2 estimates when small number of clusters were used under moderate censoring and frailty generated from a Gamma distribution.



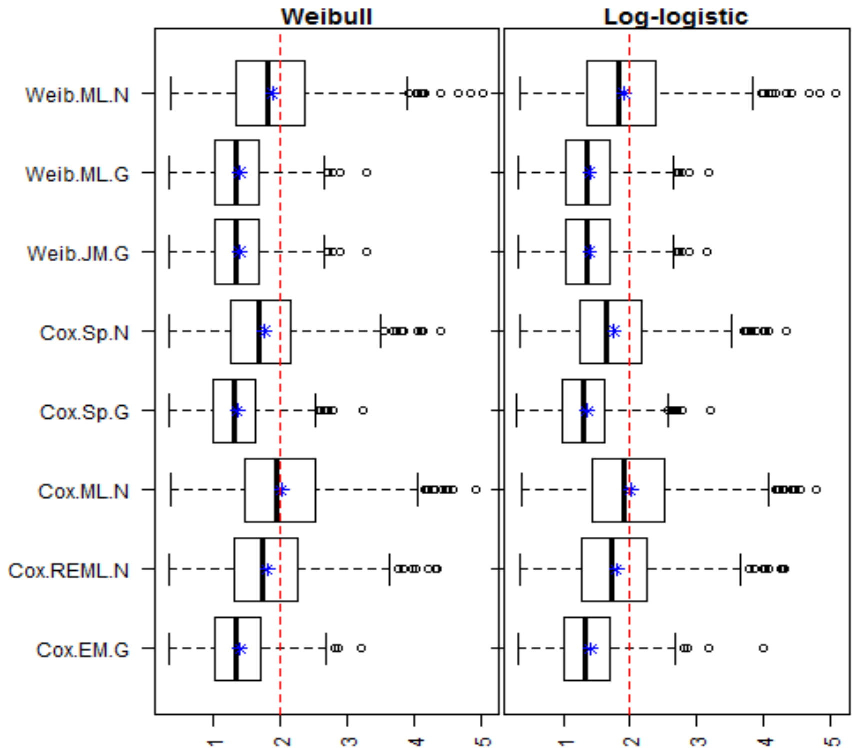
Web-Figure 13.12: θ estimates when small number of clusters were used under moderate censoring and frailty generated from a Gamma distribution.



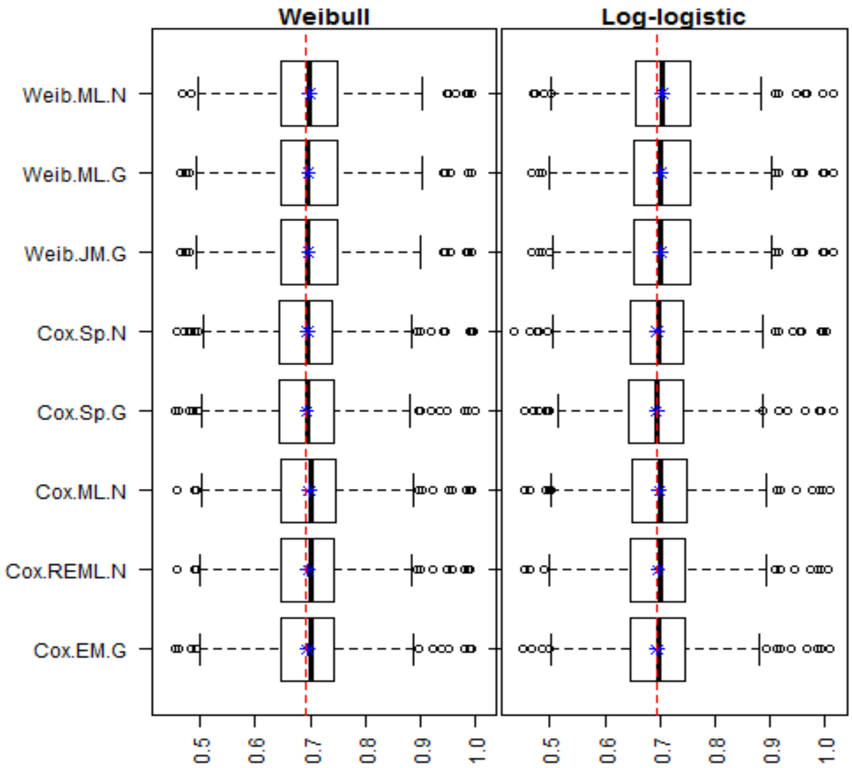
Web-Figure 13.13: β_1 estimates when small number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



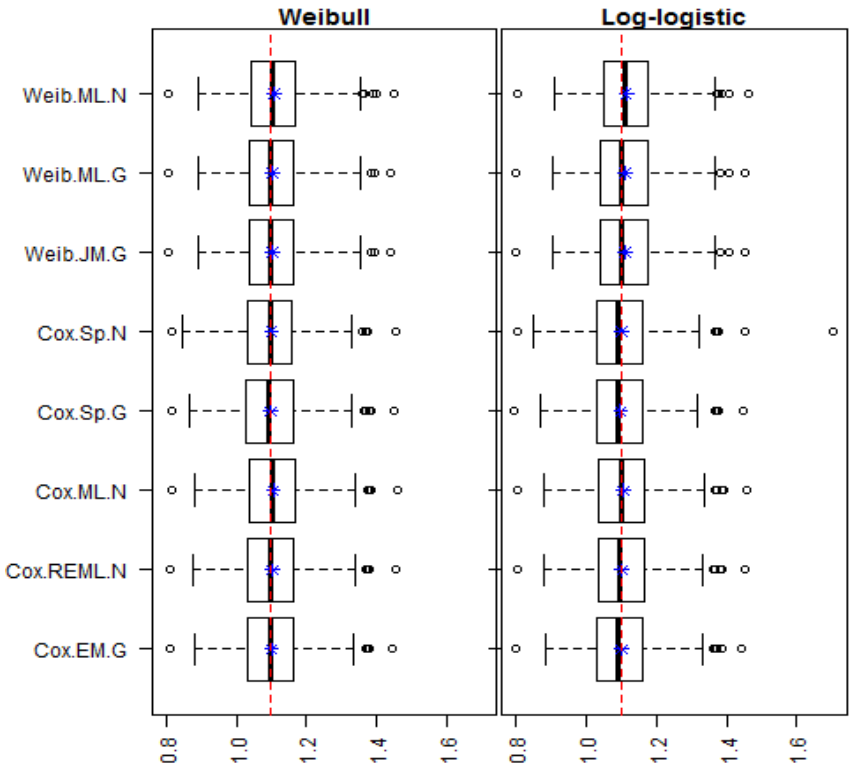
Web-Figure 13.14: β_2 estimates when small number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



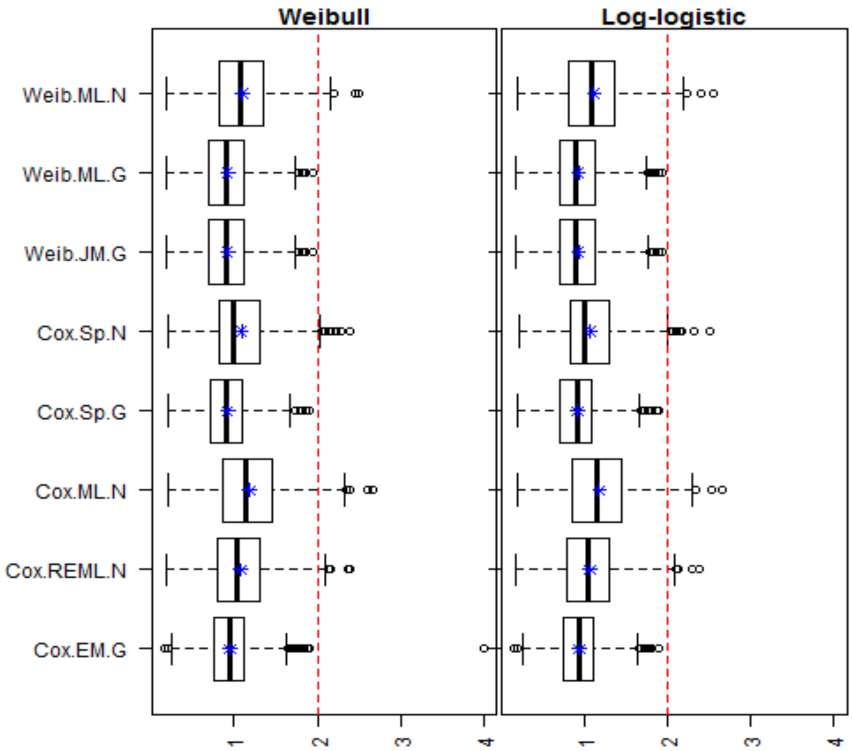
Web-Figure 13.15: θ estimates when small number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



Web-Figure 13.16: β_1 estimates when small number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

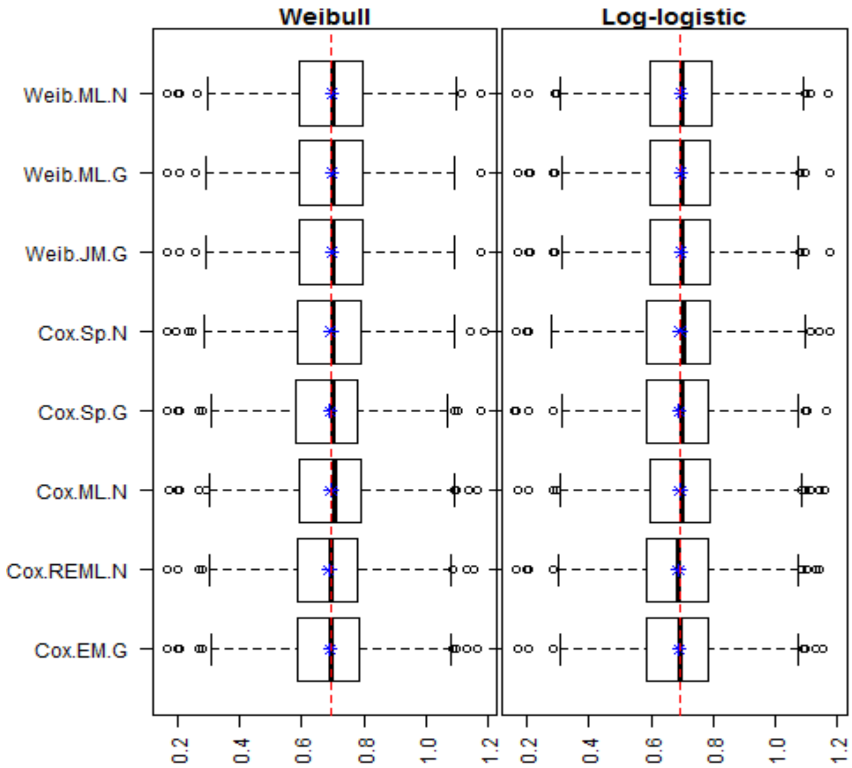


Web-Figure 13.17: β_2 estimates when small number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

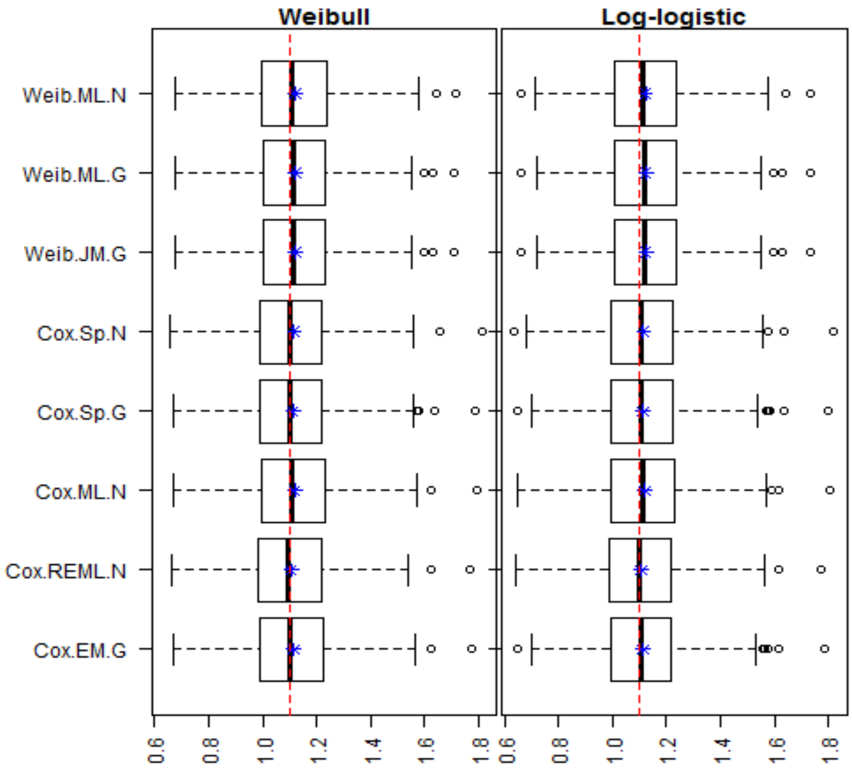


Web-Figure 13.18: θ estimates when small number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

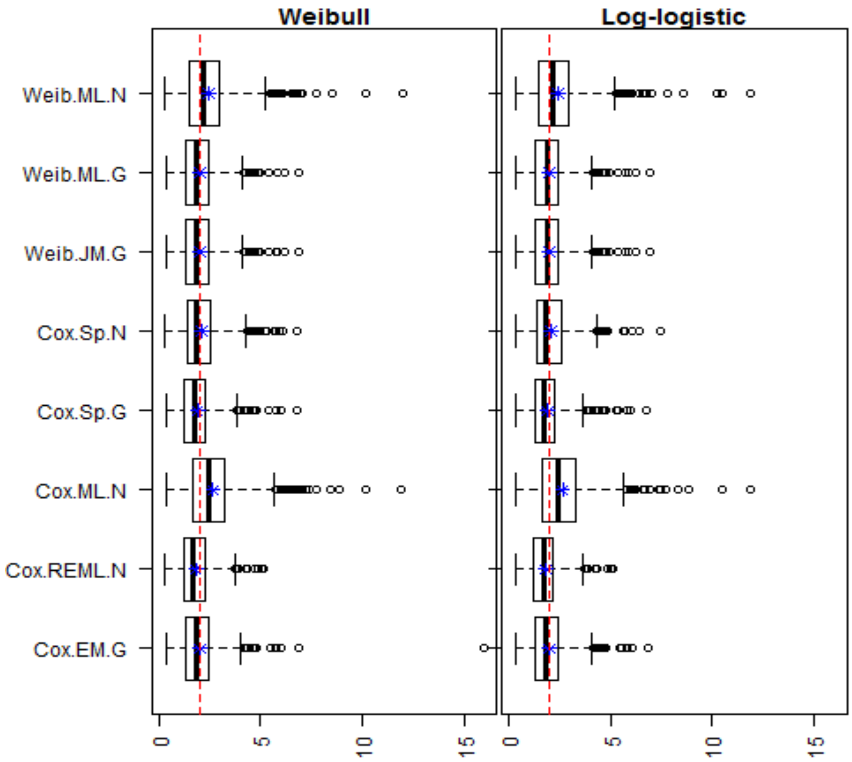
Censoring is moderate (45%) **when** $N = 15$



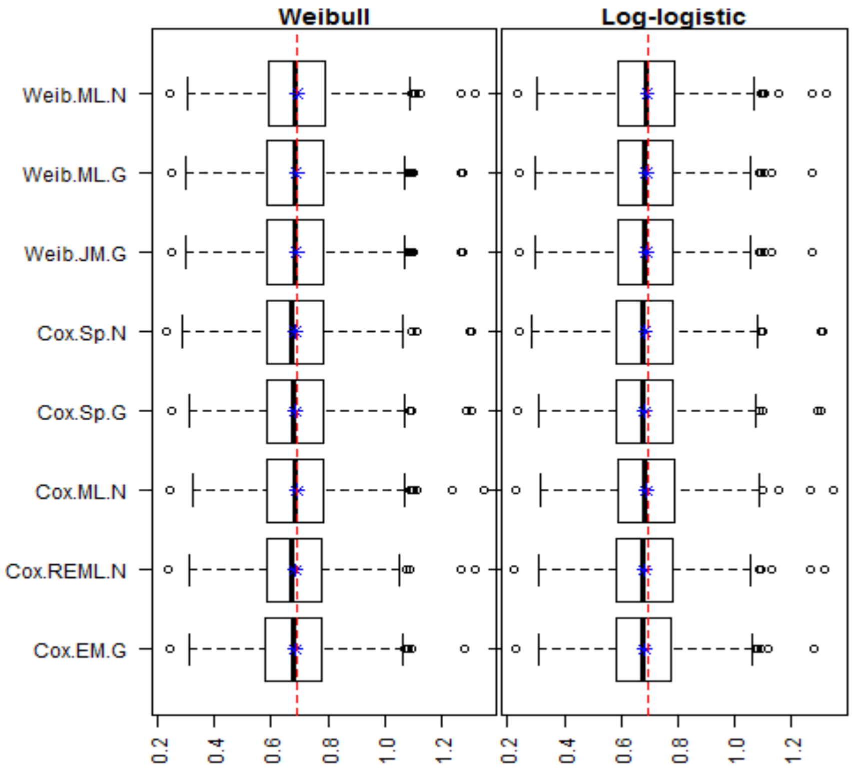
Web-Figure 13.19: β_1 estimates when small number of clusters were used under severe censoring and frailty generated from a gamma distribution.



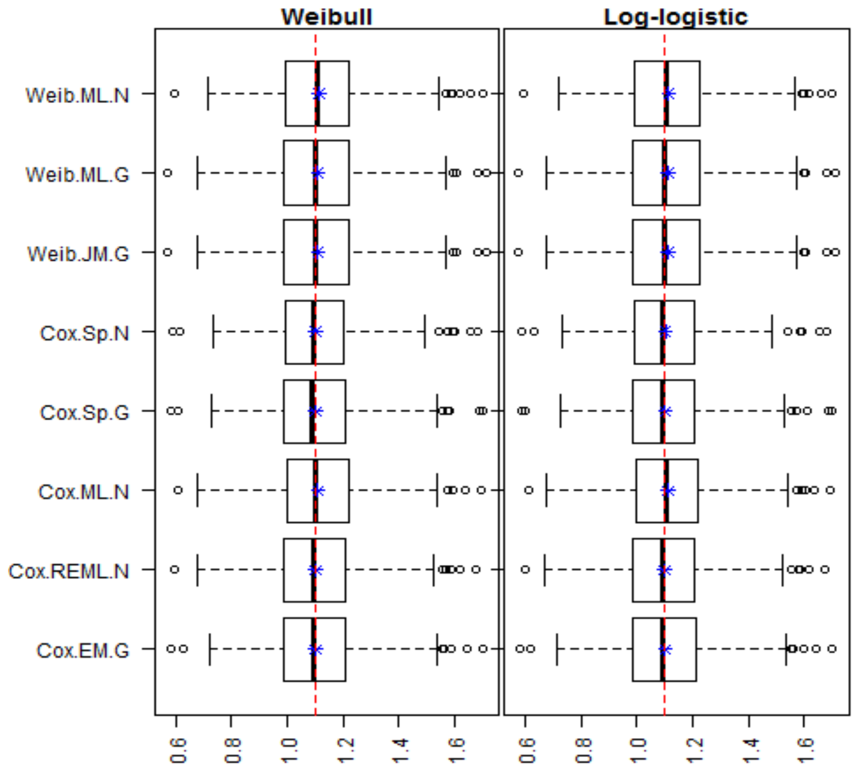
Web-Figure 13.20: β_2 estimates when small number of clusters were used under severe censoring and frailty generated from a gamma distribution.



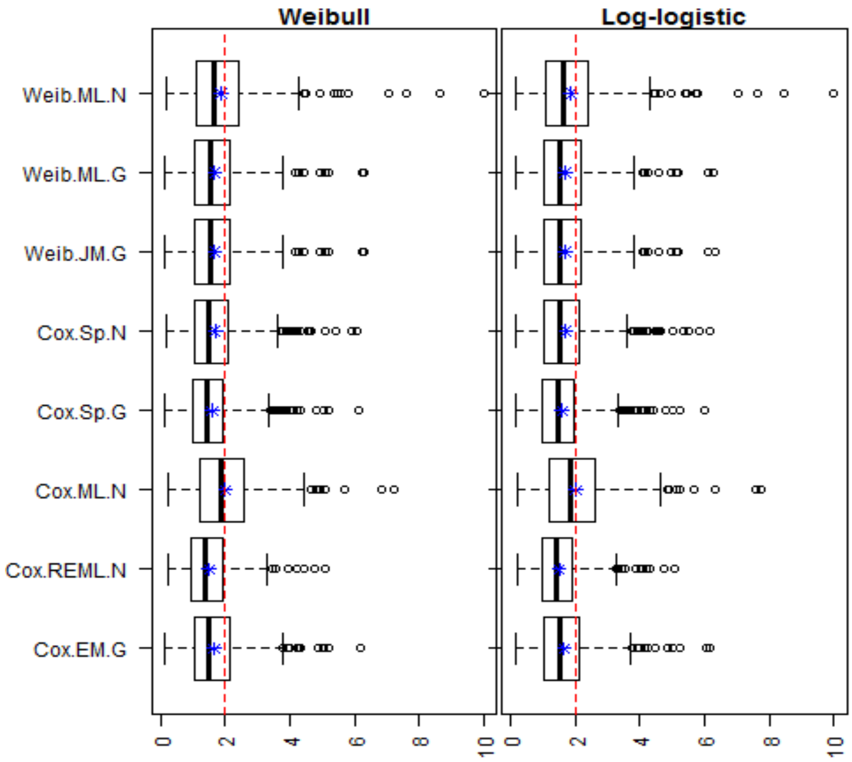
Web-Figure 13.21: θ estimates when small number of clusters were used under severe censoring and frailty generated from a gamma distribution.



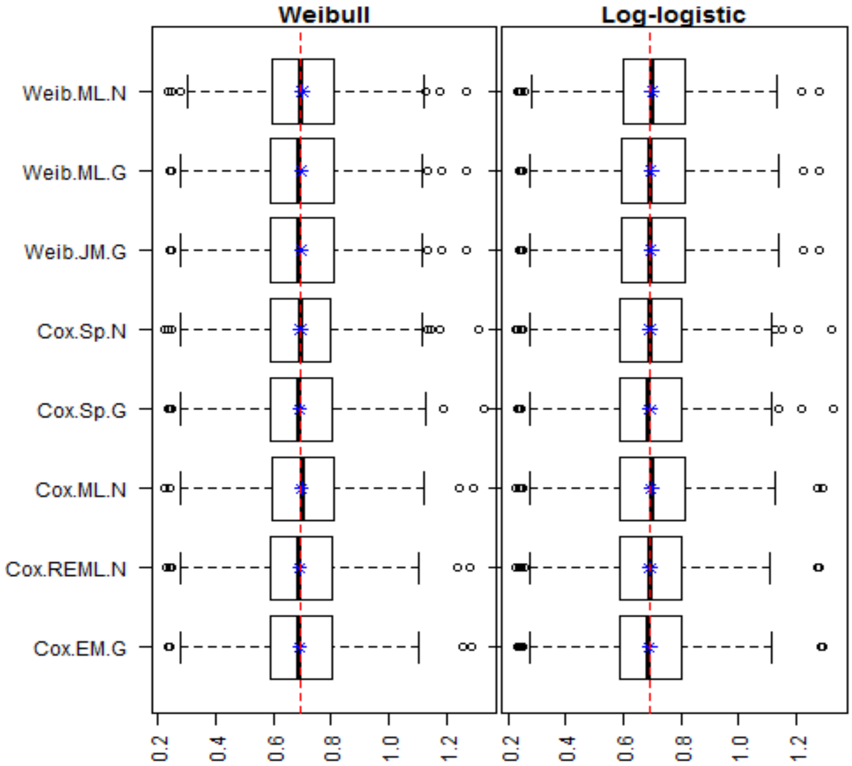
Web-Figure 13.22: β_1 estimates when small number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



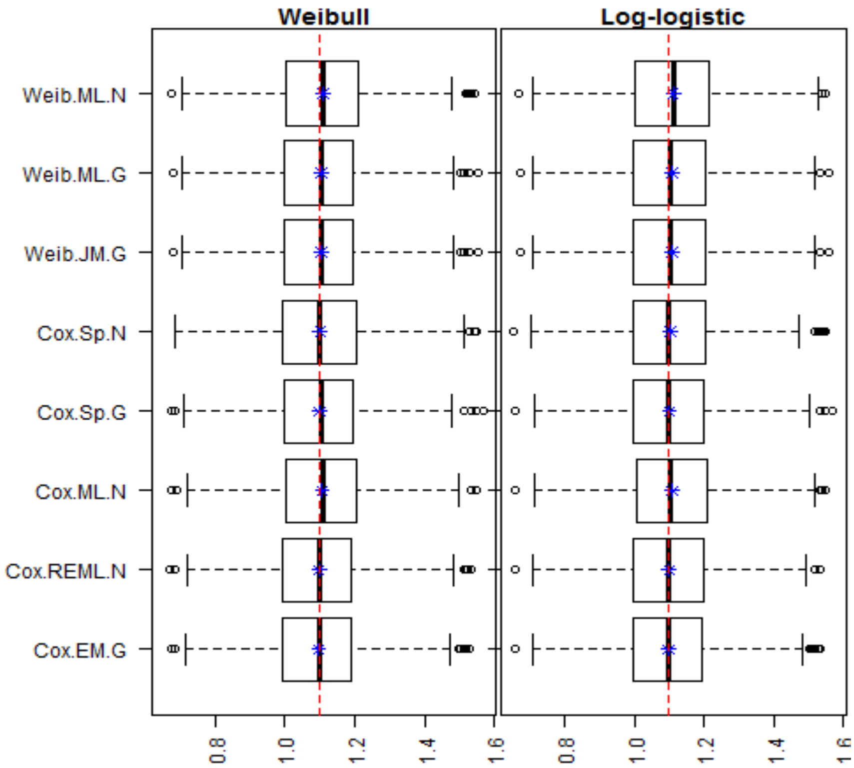
Web-Figure 13.23: β_2 estimates when small number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



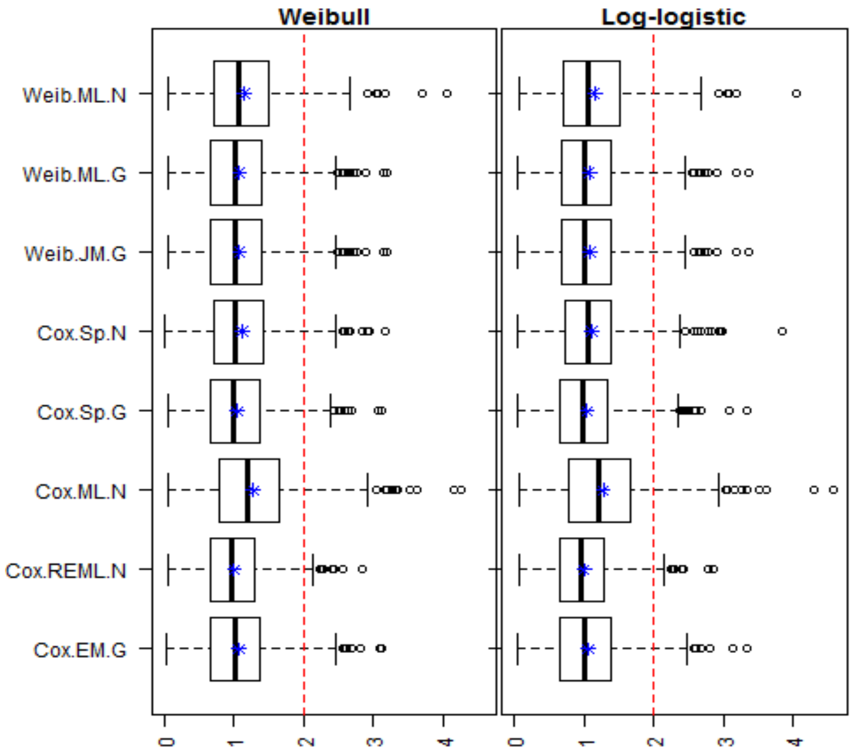
Web-Figure 13.24: θ estimates when small number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



Web-Figure 13.25: β_1 estimates when small number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.



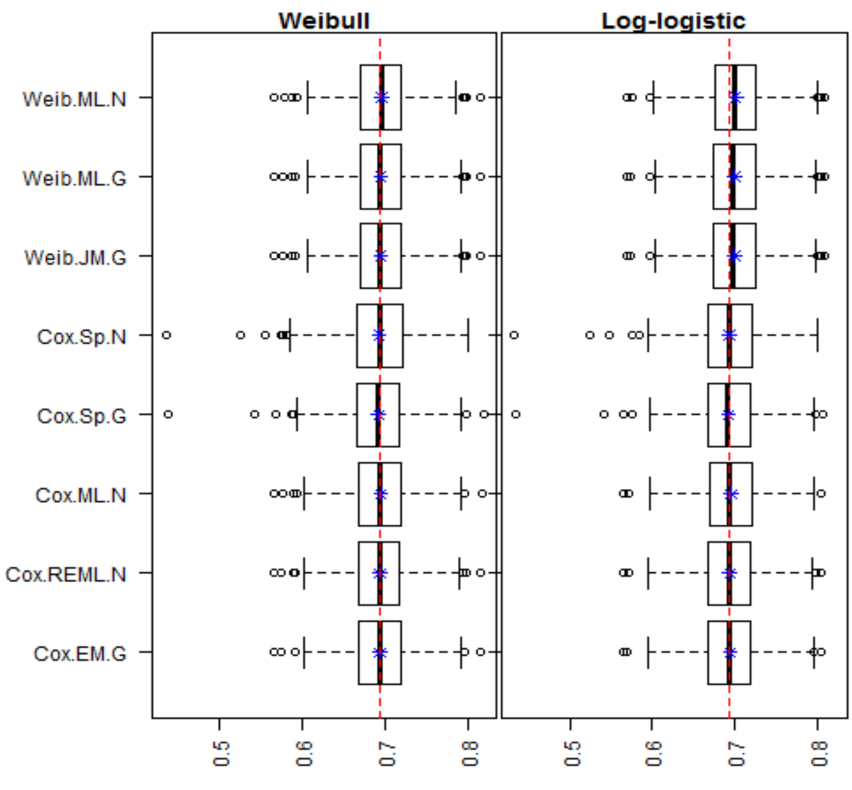
Web-Figure 13.26: β_2 estimates when small number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.



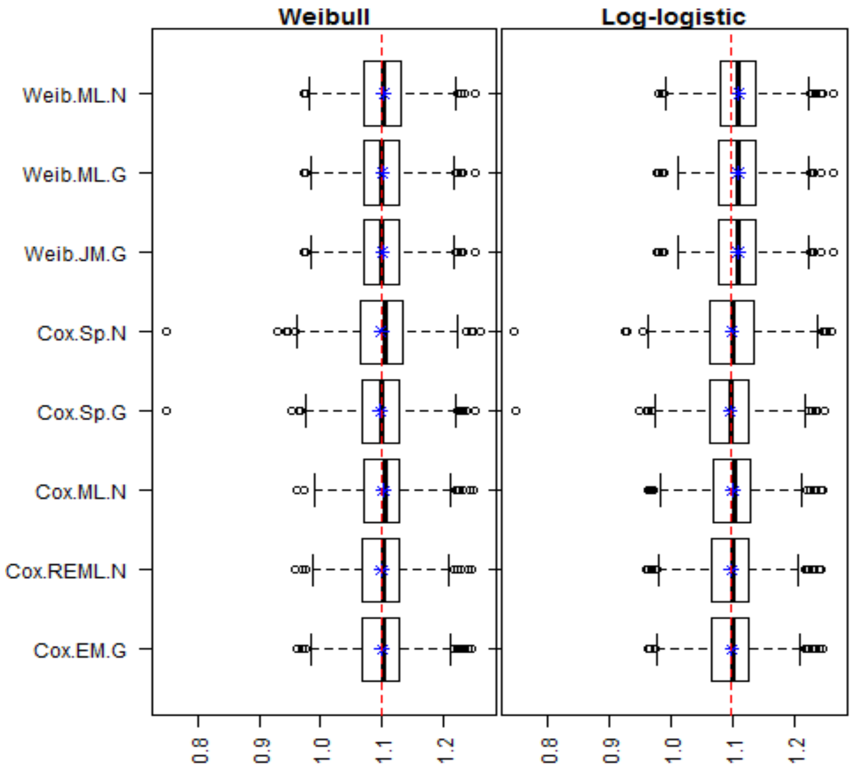
Web-Figure 13.27: θ estimates when small number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.

Censoring is severe (85%) **when** $N = 15$

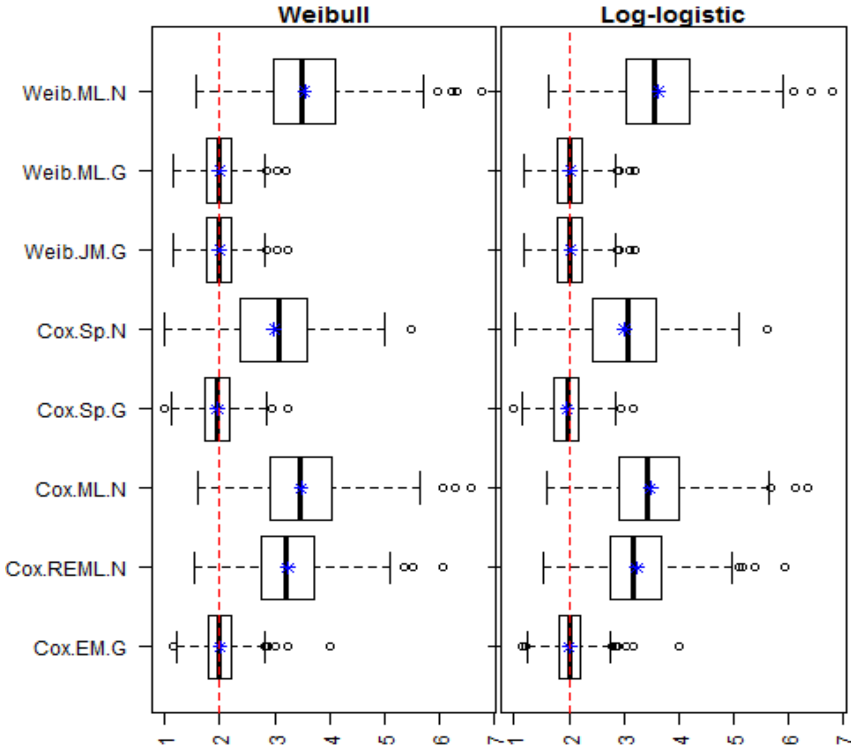
13.1.2 Number of clusters is large ($N = 60$)



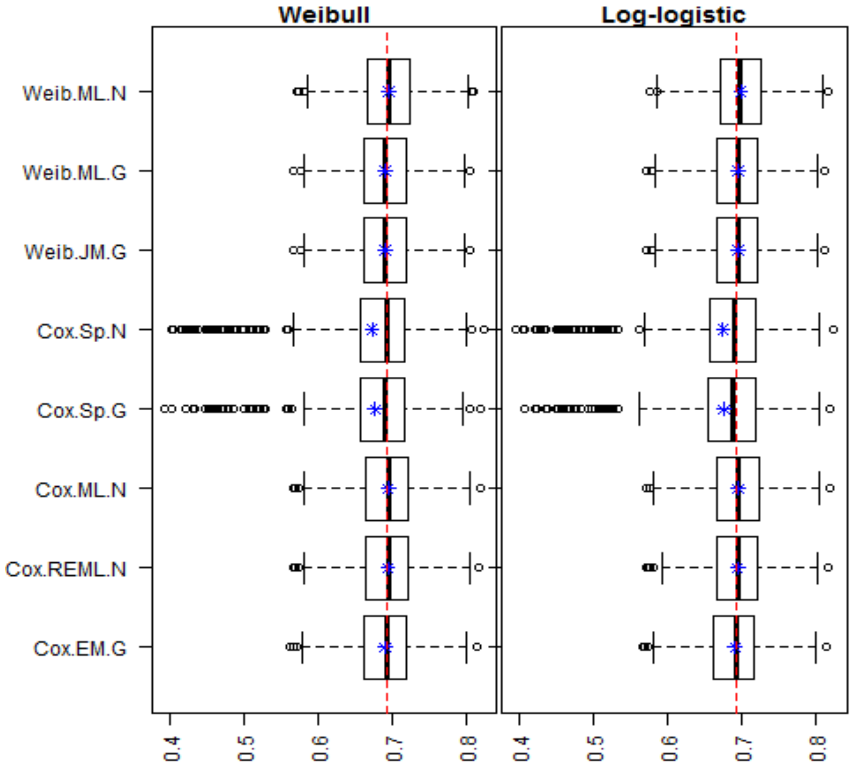
Web-Figure 13.28: β_1 estimates when large number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



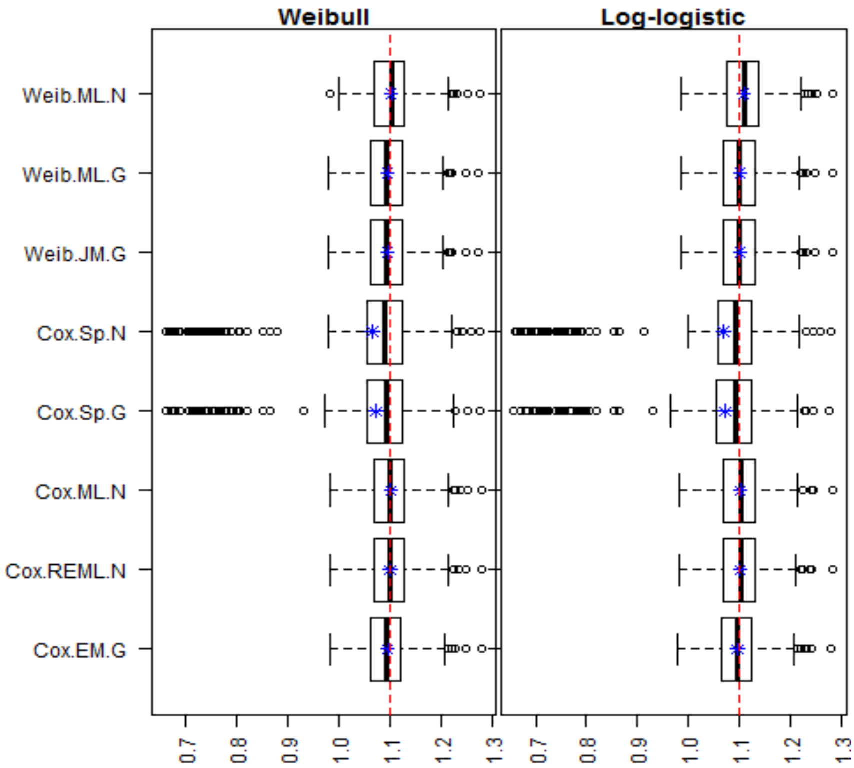
Web-Figure 13.29: β_2 estimates when large number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



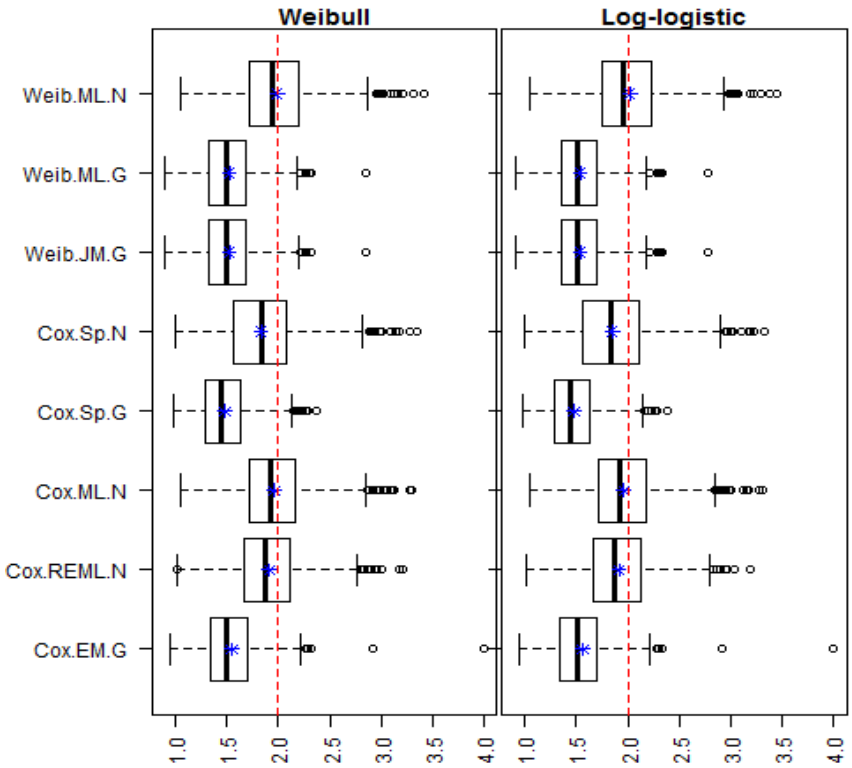
Web-Figure 13.30: θ estimates when large number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



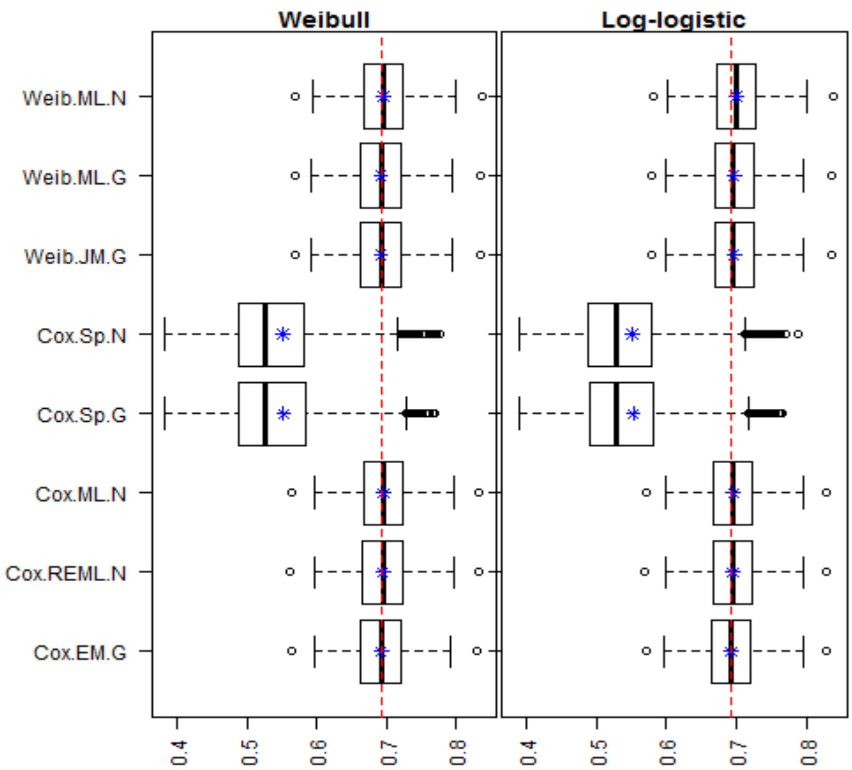
Web-Figure 13.31: β_1 estimates when large number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



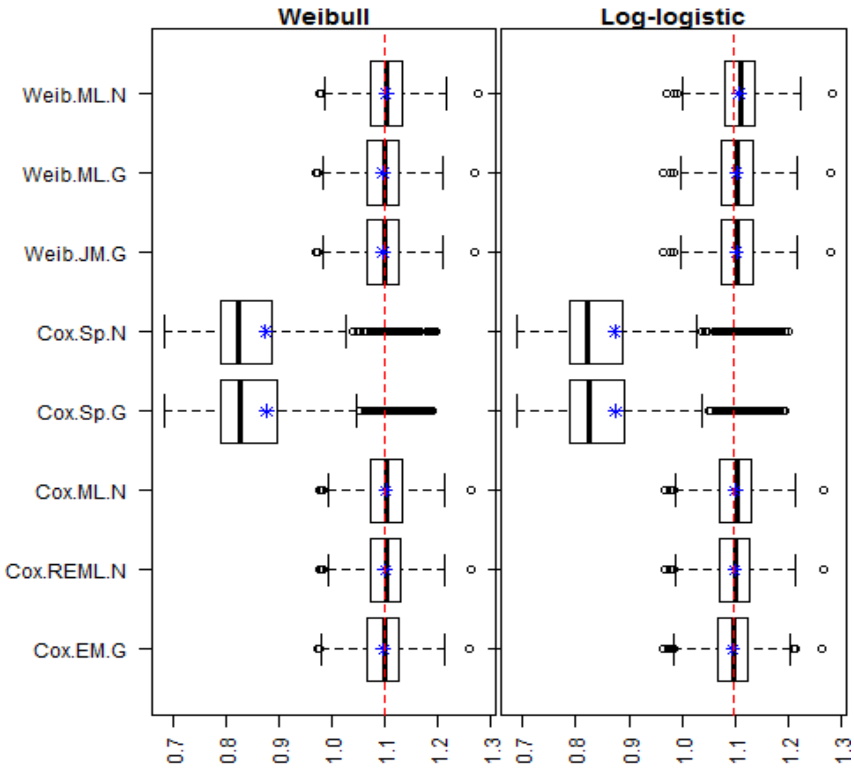
Web-Figure 13.32: β_2 estimates when large number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



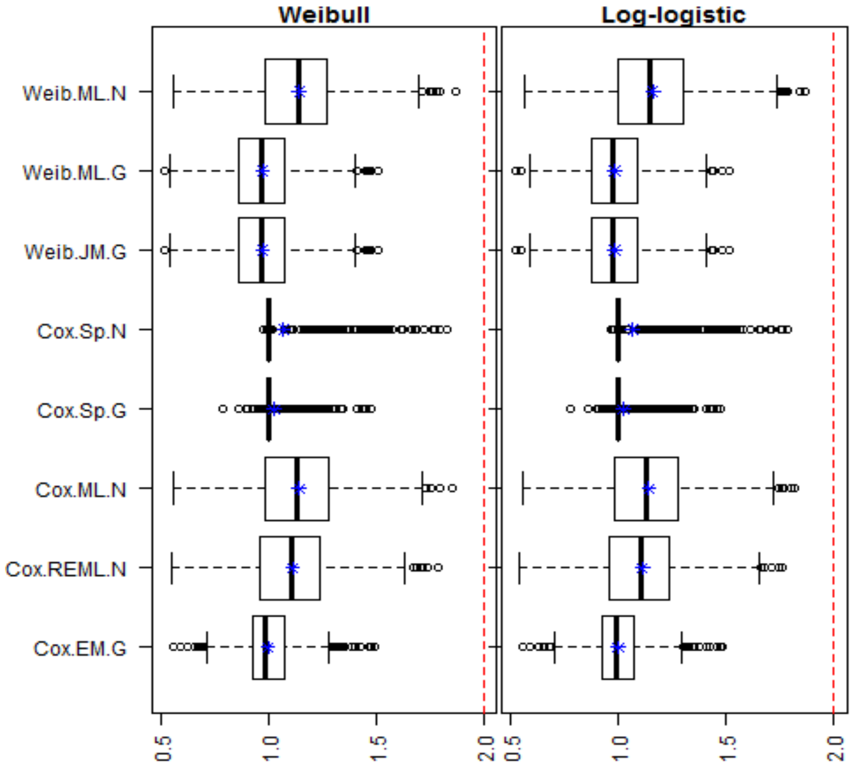
Web-Figure 13.33: θ estimates when large number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



Web-Figure 13.34: β_1 estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

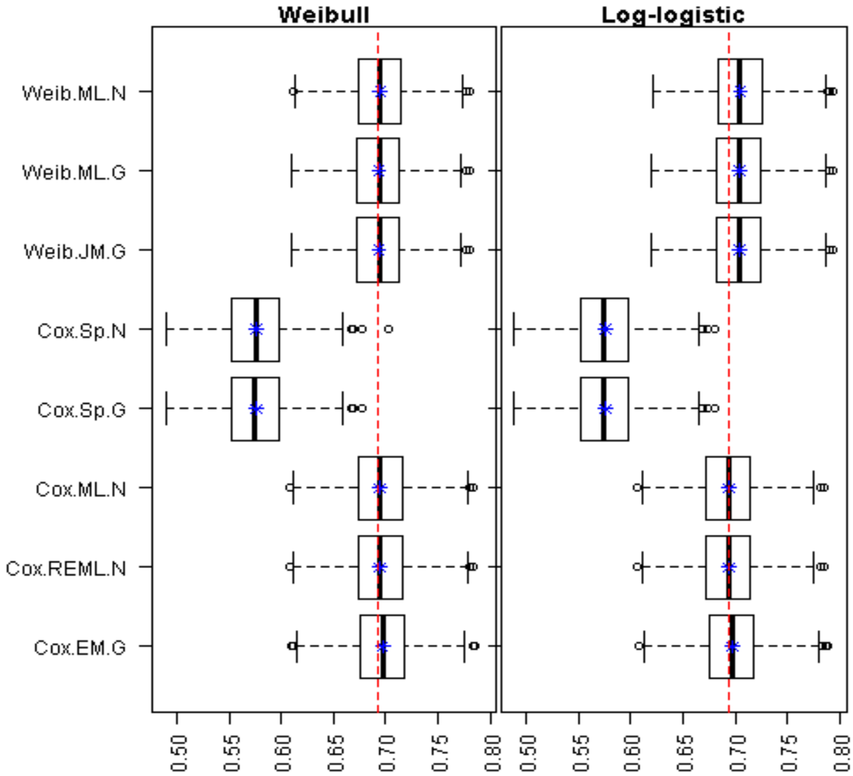


Web-Figure 13.35: β_2 estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

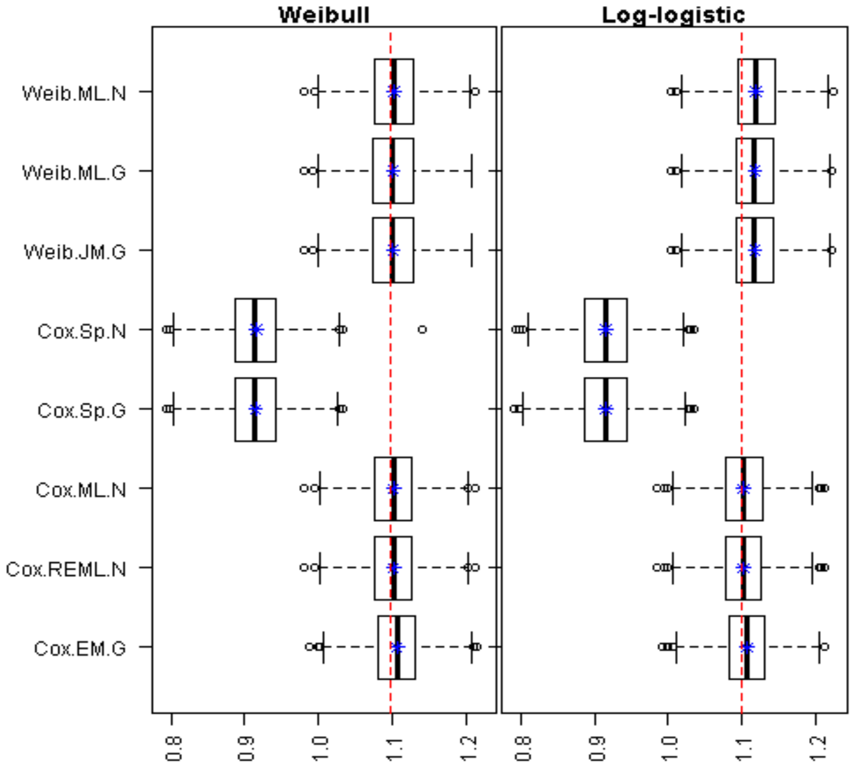


Web-Figure 13.36: θ estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

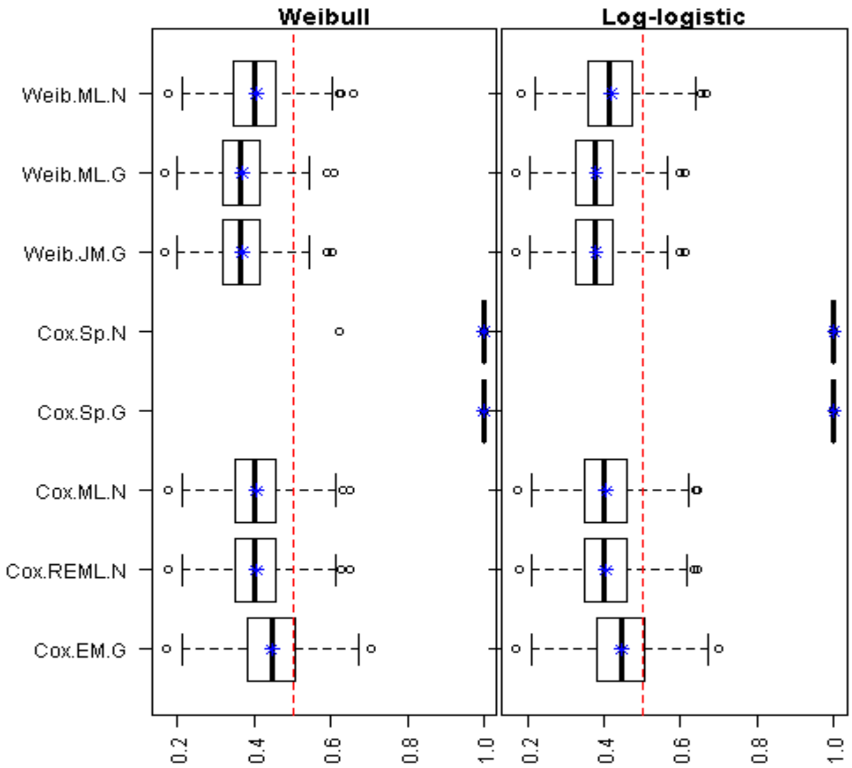
Censoring is moderate (45%) **when** $N = 60$



Web-Figure 13.37: β_1 estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.



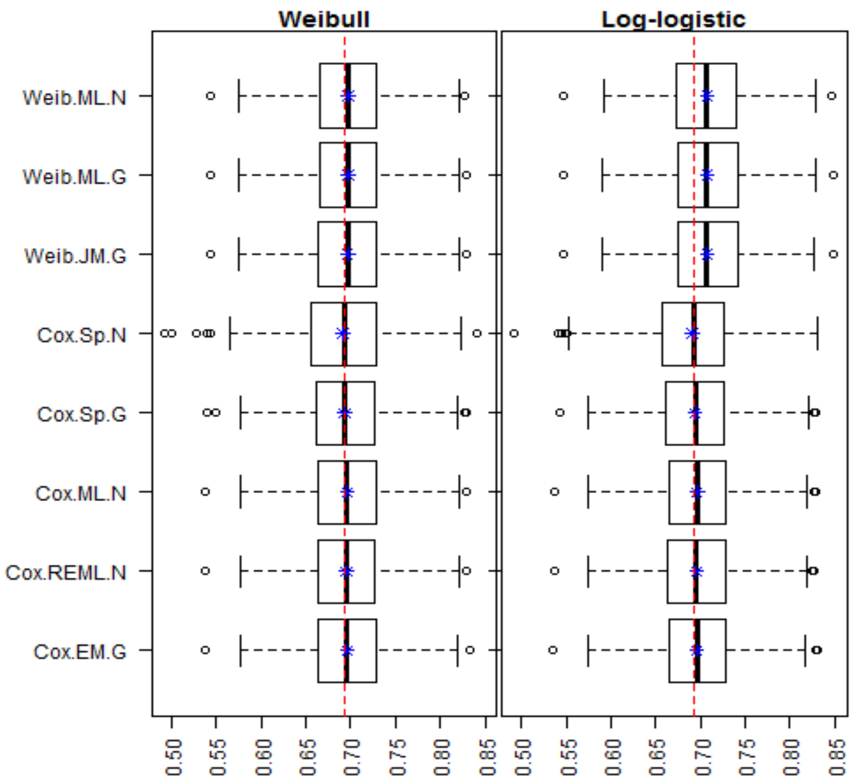
Web-Figure 13.38: β_2 estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.



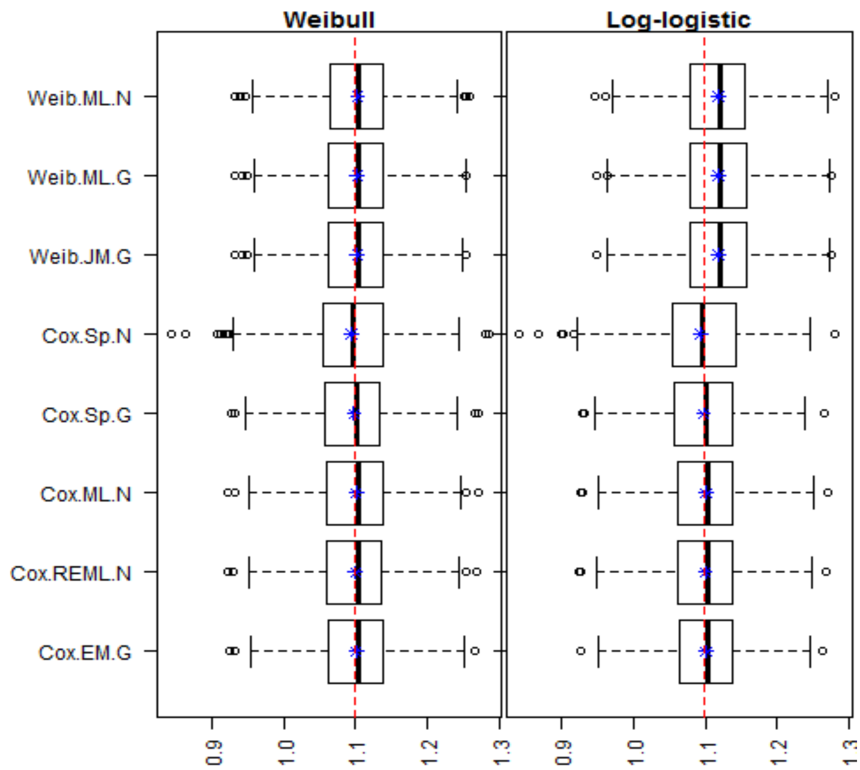
Web-Figure 13.39: θ estimates when large number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

Censoring is mild (15%) **when** $N = 60$

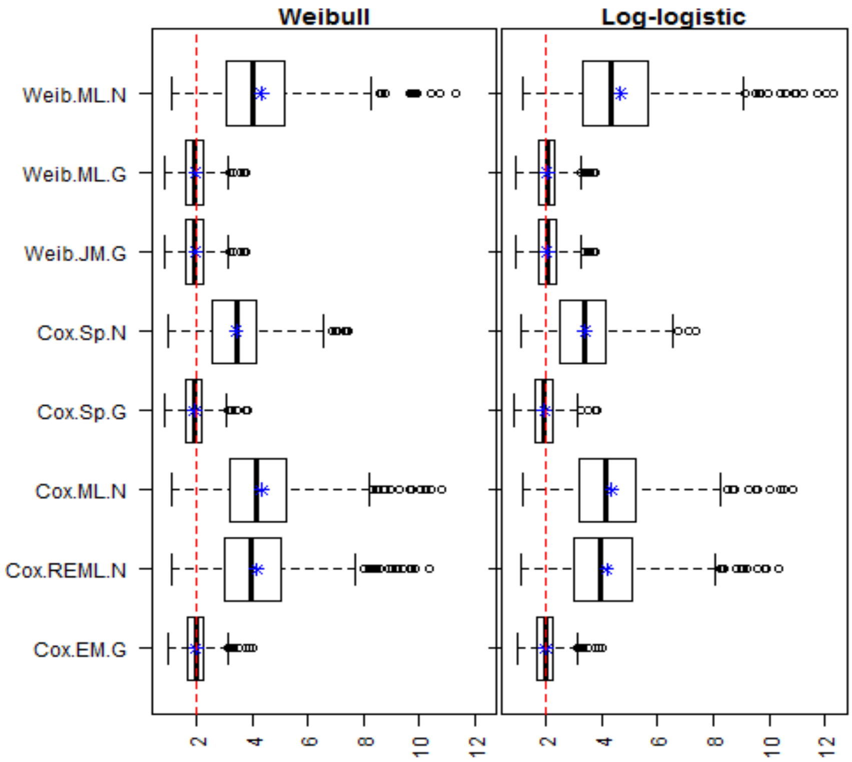
13.1.3 Number of clusters is moderate ($N = 30$)



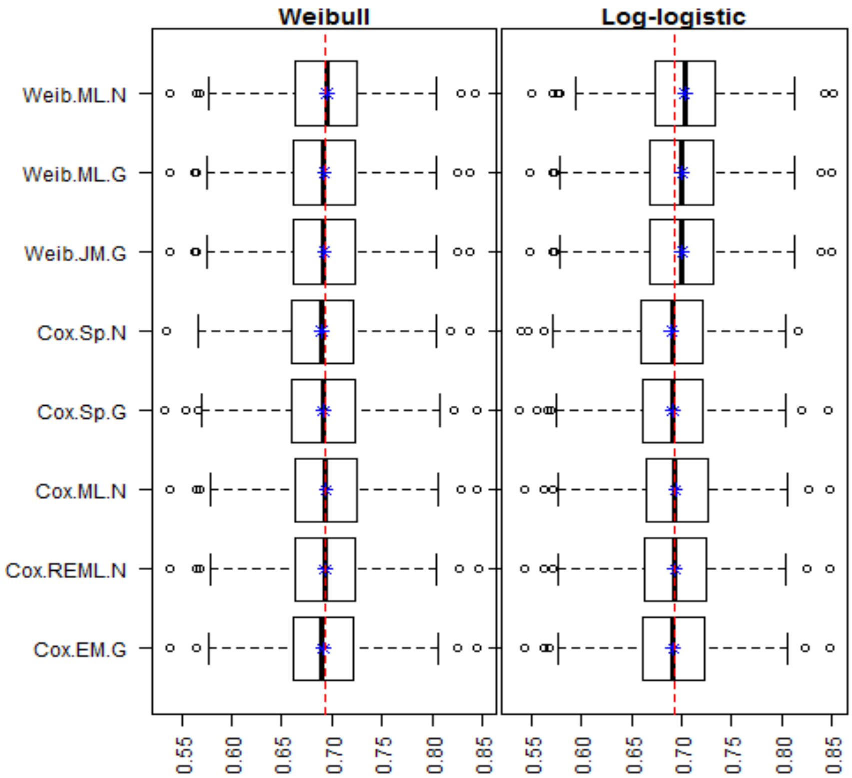
Web-Figure 13.40: β_1 estimates when moderate number of clusters were used under mild censoring and frailty generated from a gamma distribution.



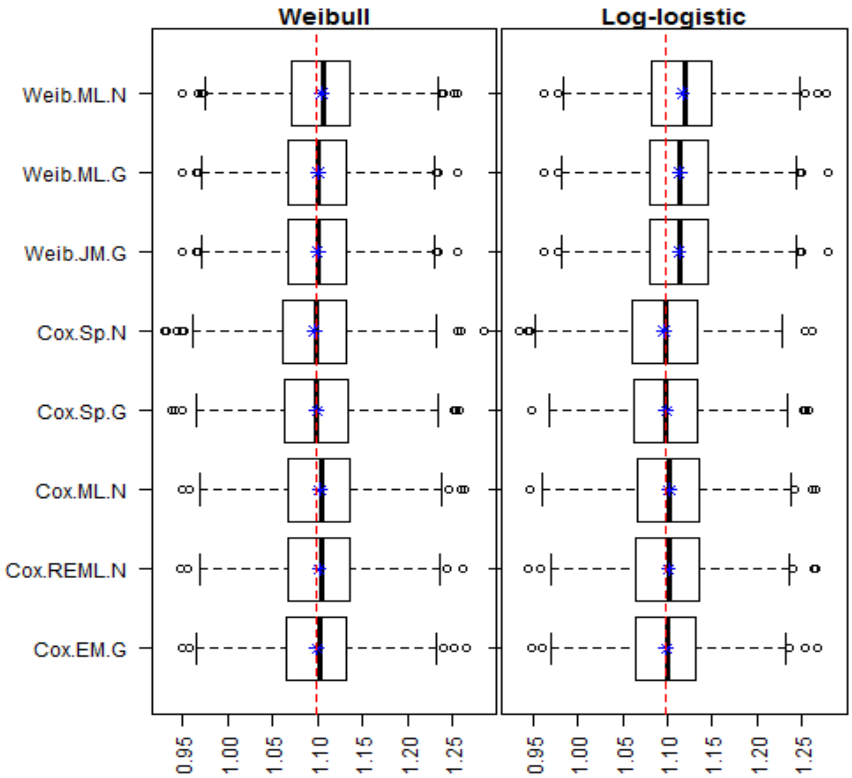
Web-Figure 13.41: β_2 estimates when moderate number of clusters were used under mild censoring and frailty generated from a gamma distribution.



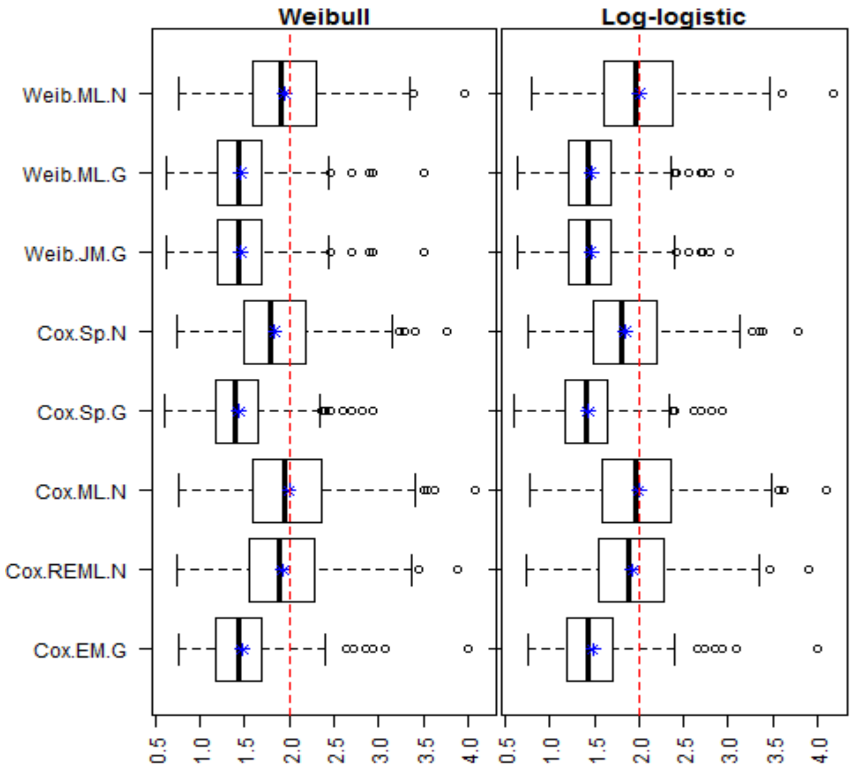
Web-Figure 13.42: θ estimates when moderate number of clusters were used under mild censoring and frailty generated from a gamma distribution.



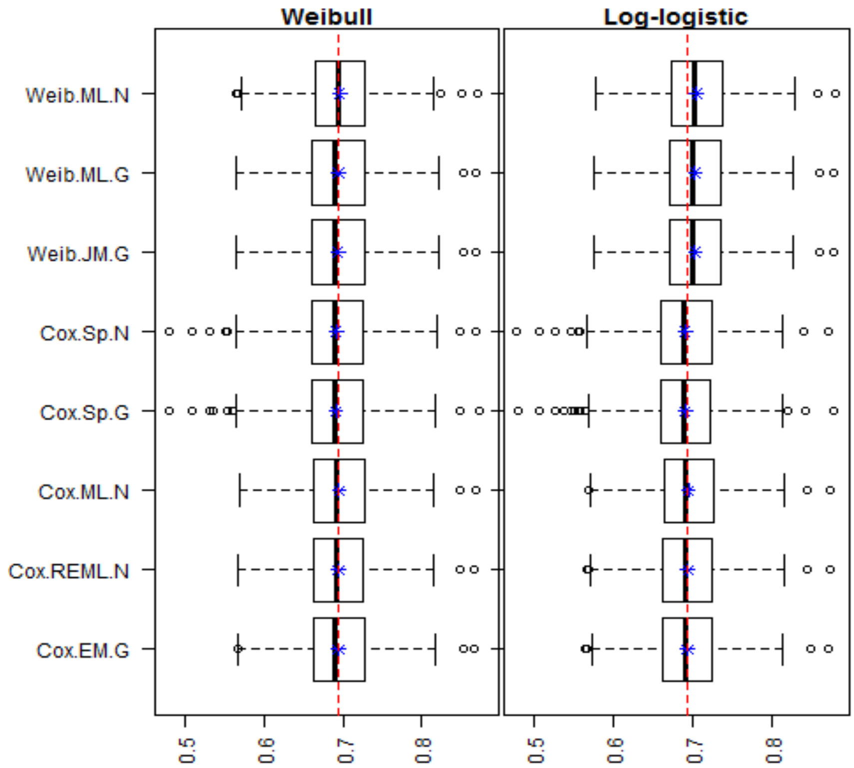
Web-Figure 13.43: β_1 estimates when moderate number of clusters were used under mild censoring and frailty generated from a log-normal distribution.



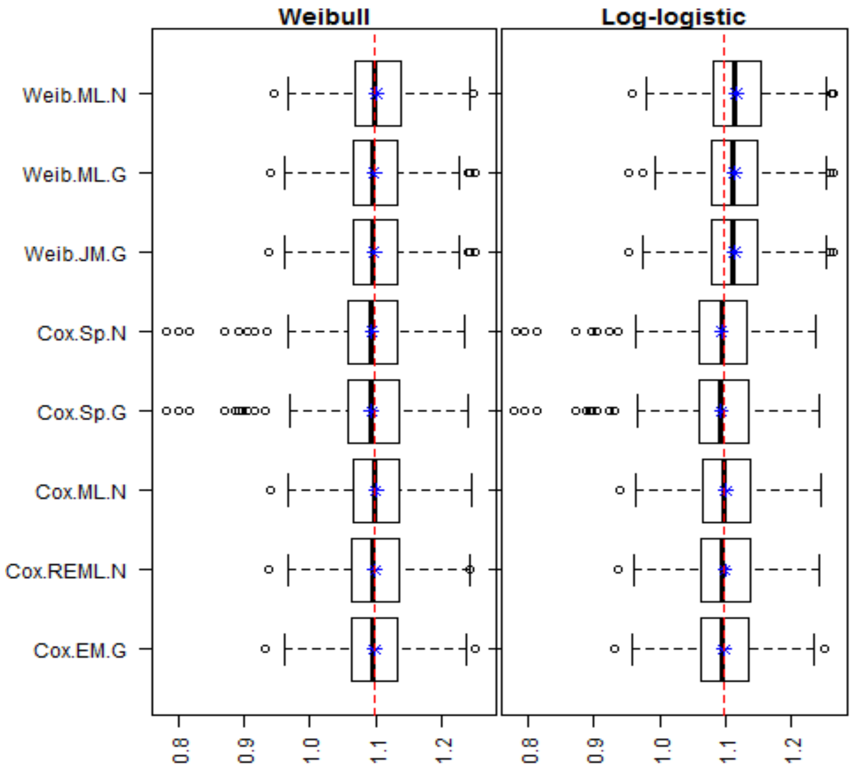
Web-Figure 13.44: β_2 estimates when moderate number of clusters were used under mild censoring and frailty generated from a log-normal distribution.



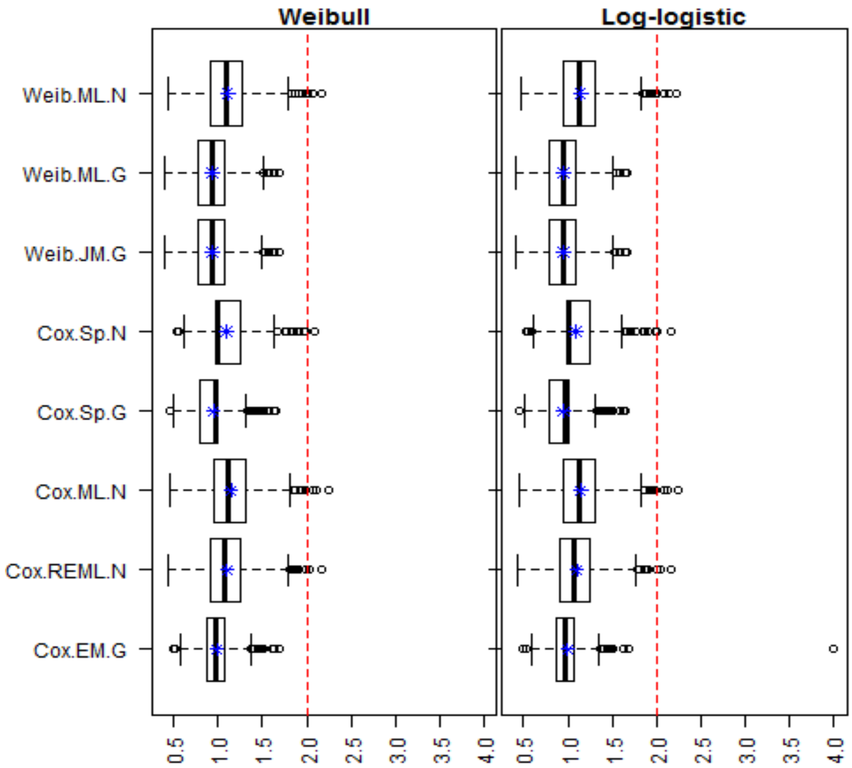
Web-Figure 13.45: θ estimates when moderate number of clusters were used under mild censoring and frailty generated from a log-normal distribution.



Web-Figure 13.46: β_1 estimates when moderate number of clusters were used under mild censoring and frailty generated from an inverse Gaussian distribution.

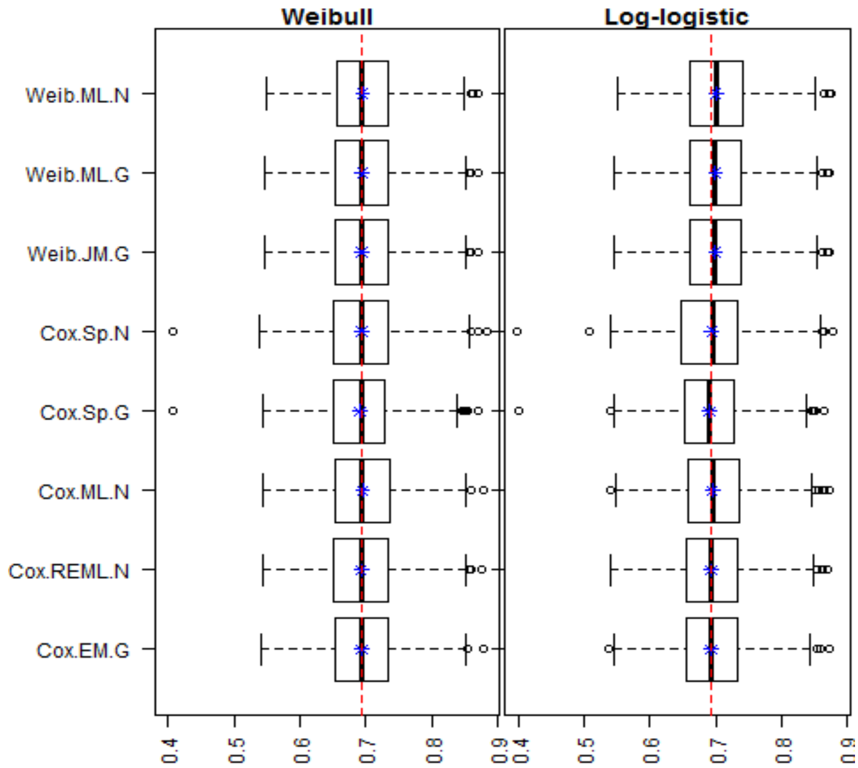


Web-Figure 13.47: β_2 estimates when moderate number of clusters were used under mild censoring and frailty generated from an inverse Gaussian distribution.

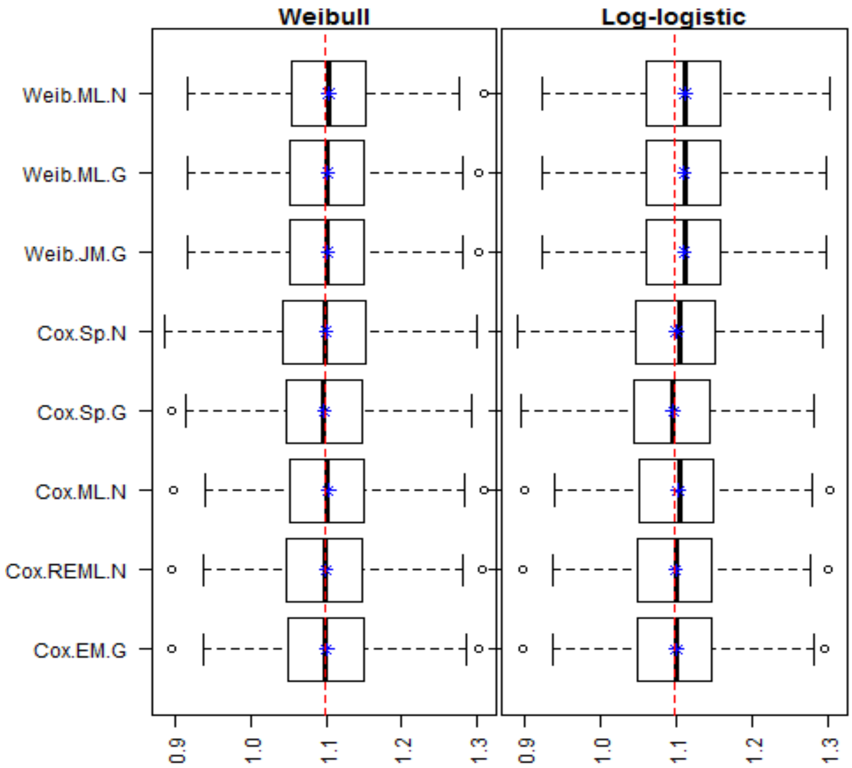


Web-Figure 13.48: θ estimates when moderate number of clusters were used under mild censoring and frailty generated from an inverse Gaussian distribution.

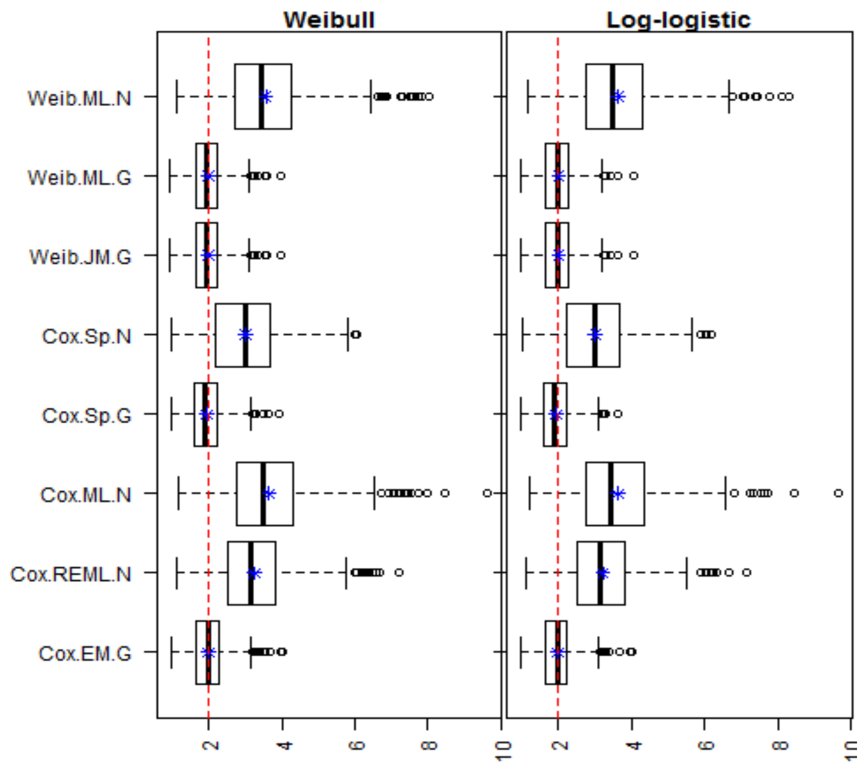
Censoring is mild (15%) **when** $N = 30$



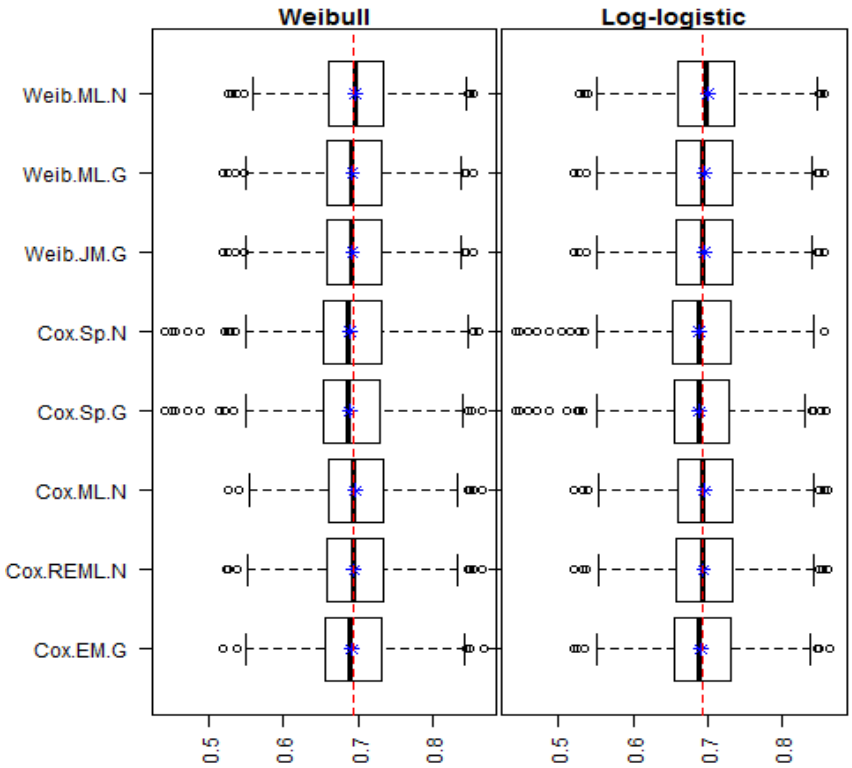
Web-Figure 13.49: β_1 estimates when moderate number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



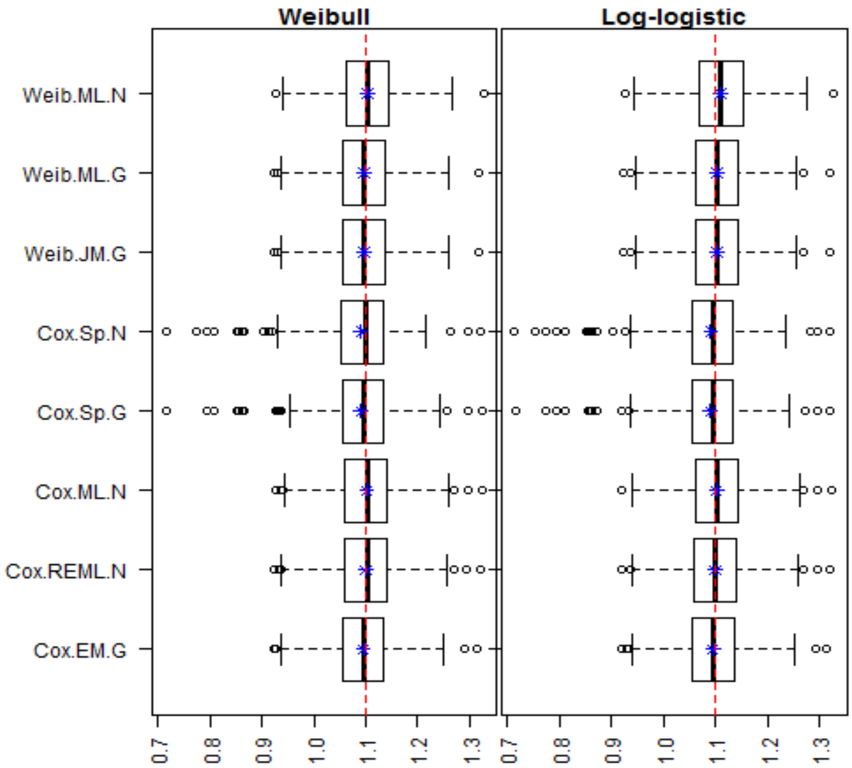
Web-Figure 13.50: β_2 estimates when moderate number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



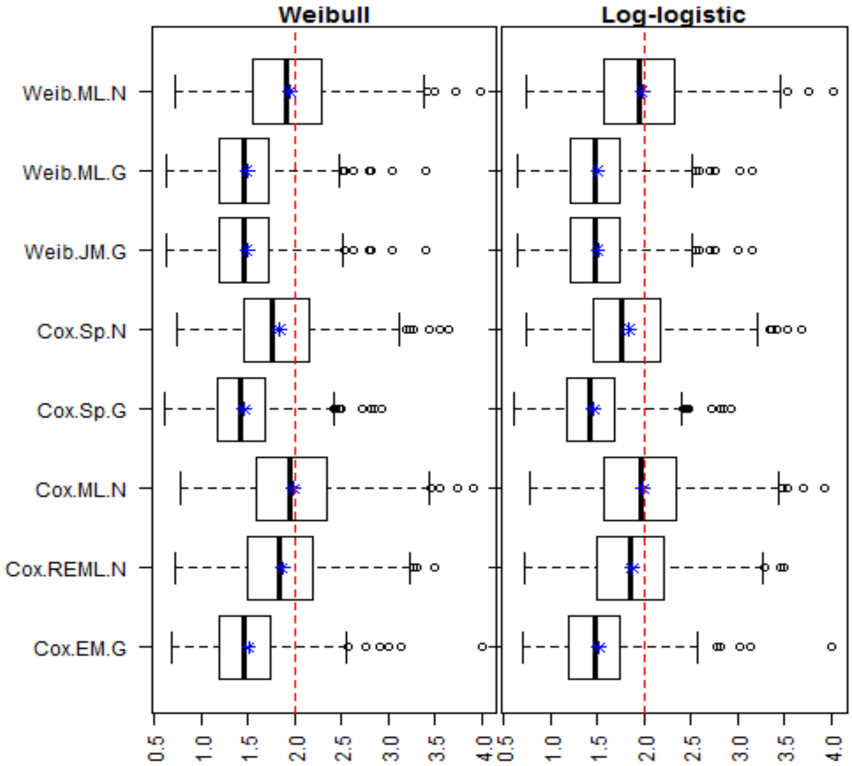
Web-Figure 13.51: θ estimates when moderate number of clusters were used under moderate censoring and frailty generated from a gamma distribution.



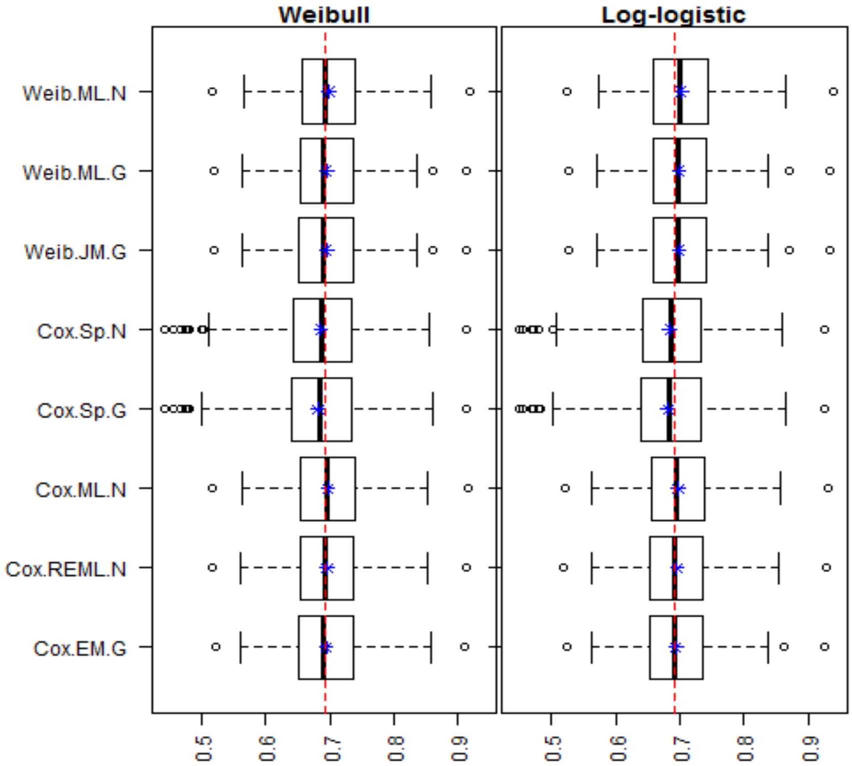
Web-Figure 13.52: β_1 estimates when moderate number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



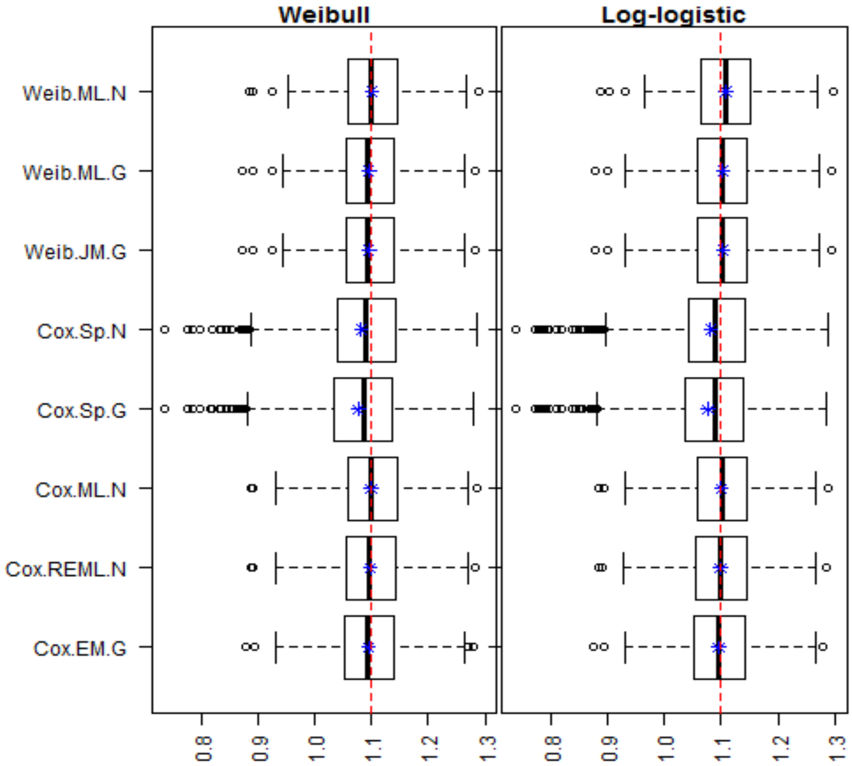
Web-Figure 13.53: β_2 estimates when moderate number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



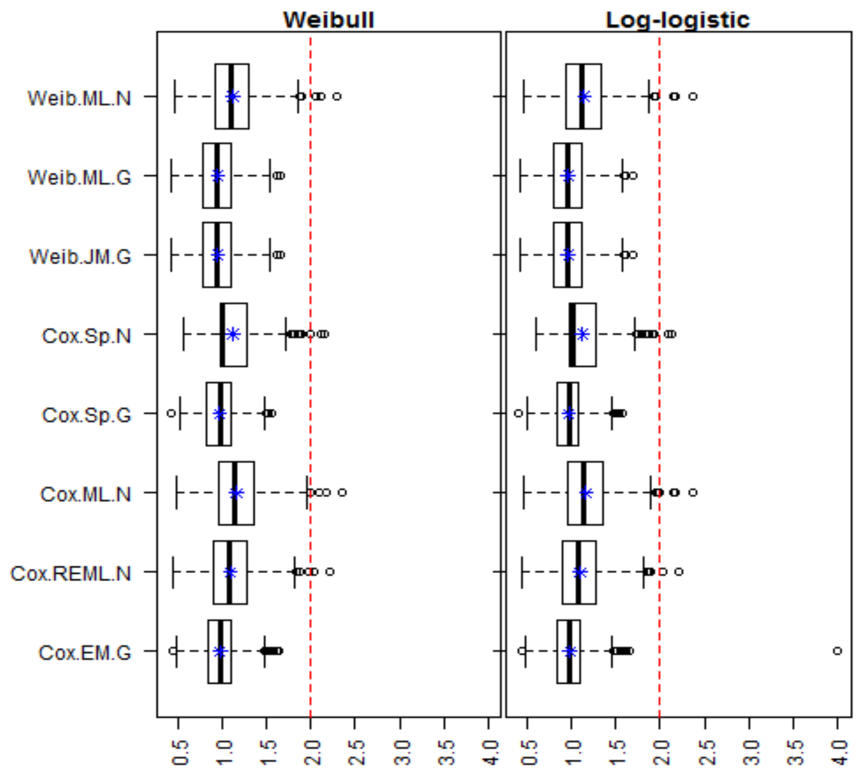
Web-Figure 13.54: θ estimates when moderate number of clusters were used under moderate censoring and frailty generated from a log-normal distribution.



Web-Figure 13.55: β_1 estimates when moderate number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

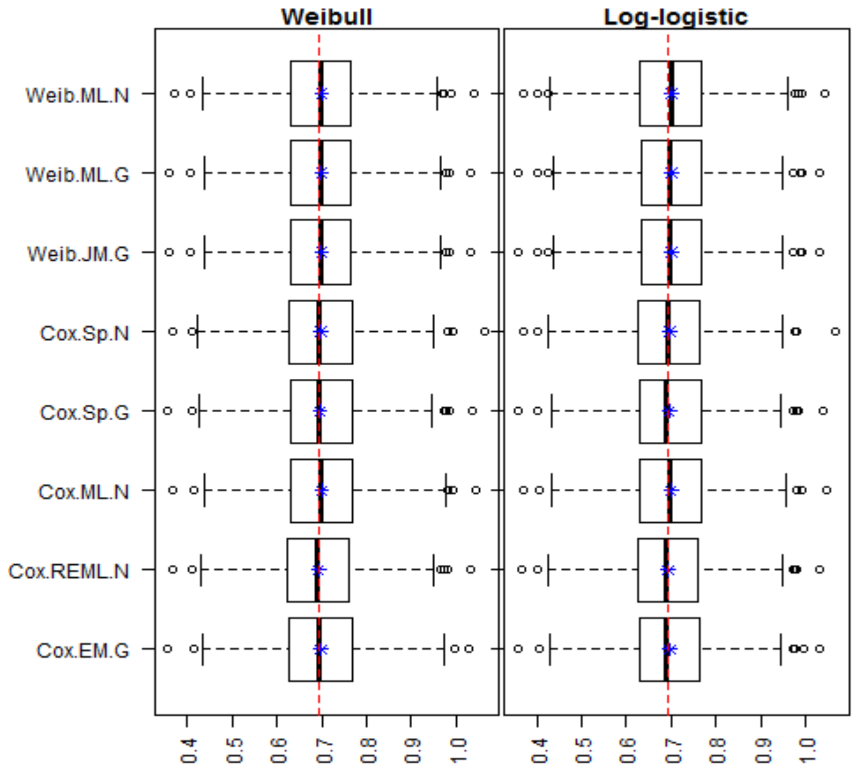


Web-Figure 13.56: β_2 estimates when moderate number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

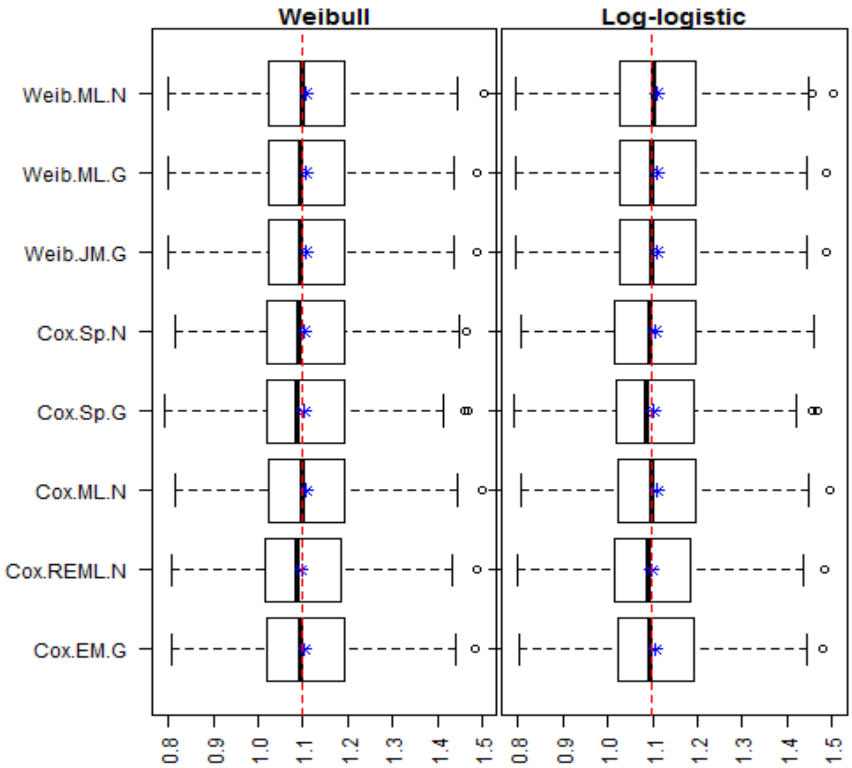


Web-Figure 13.57: θ estimates when moderate number of clusters were used under moderate censoring and frailty generated from an inverse Gaussian distribution.

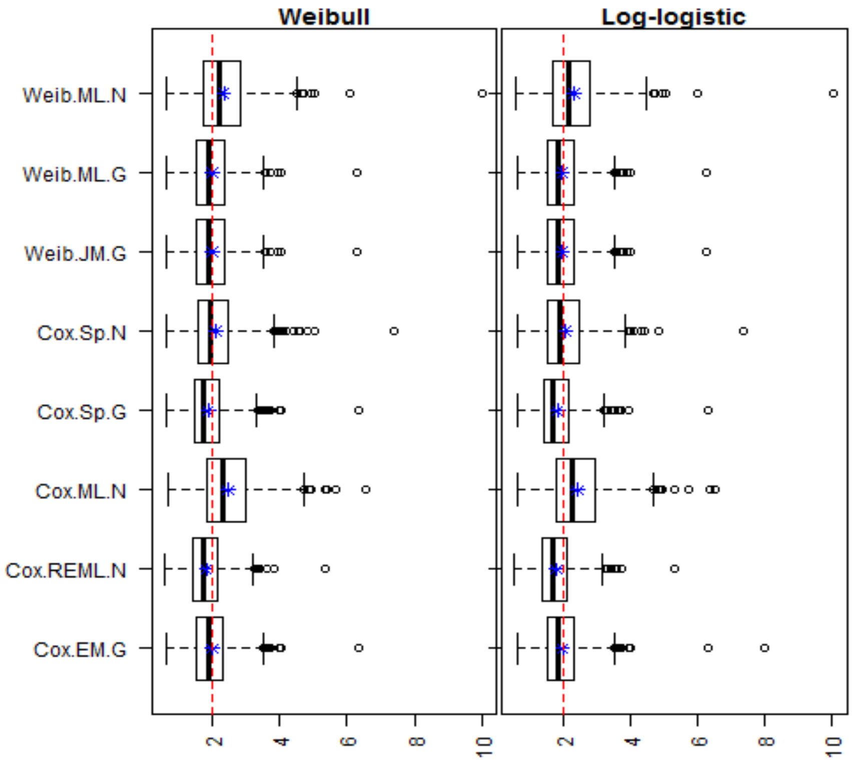
Censoring is moderate (45%) **when** $N = 30$



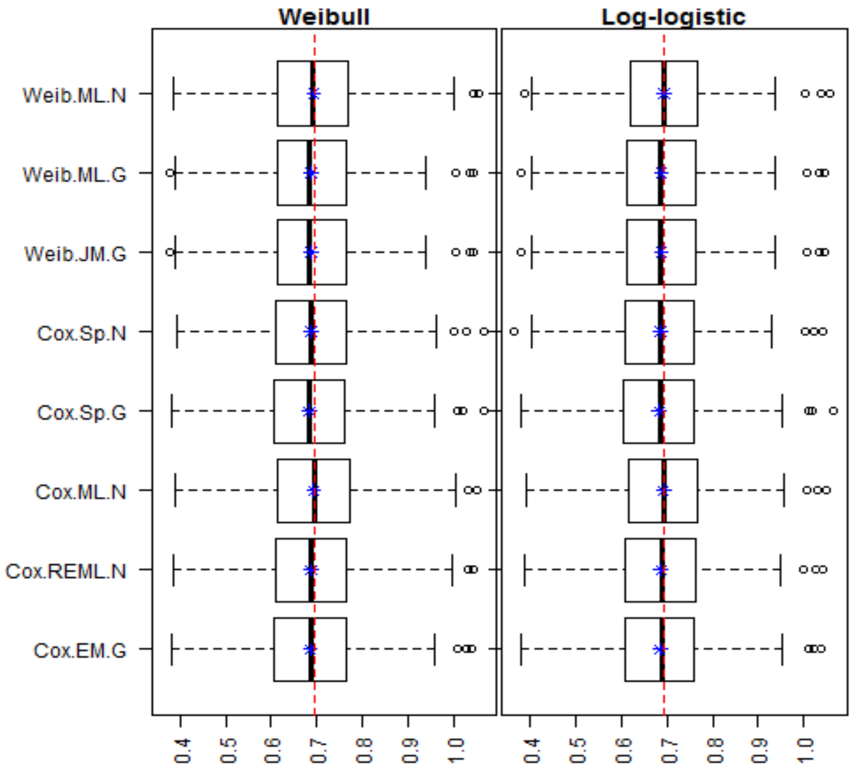
Web-Figure 13.58: β_1 estimates when moderate number of clusters were used under severe censoring and frailty generated from a gamma distribution.



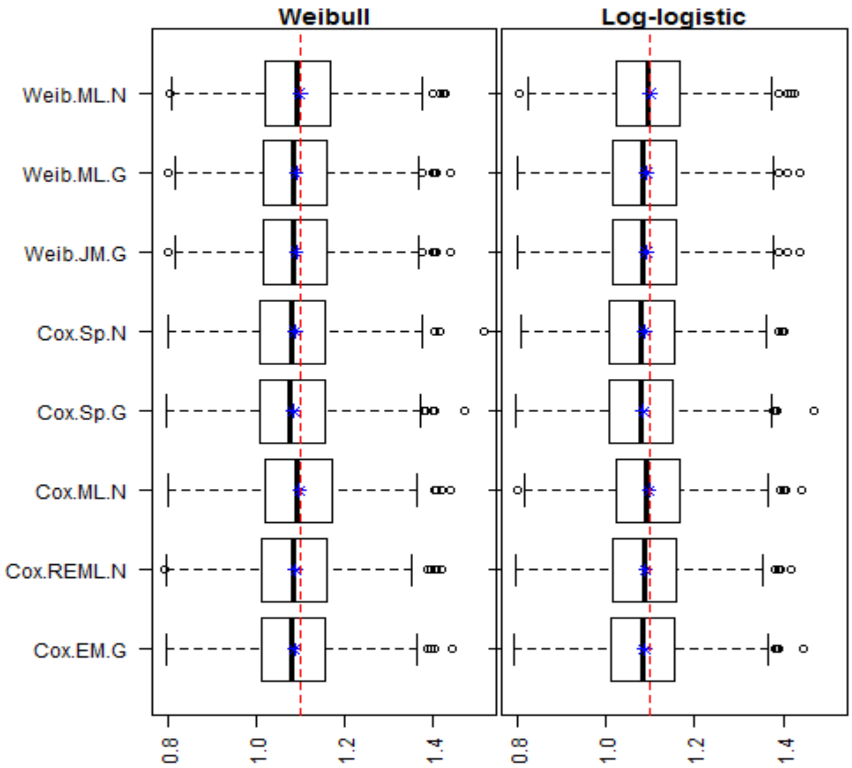
Web-Figure 13.59: β_2 estimates when moderate number of clusters were used under severe censoring and frailty generated from a gamma distribution.



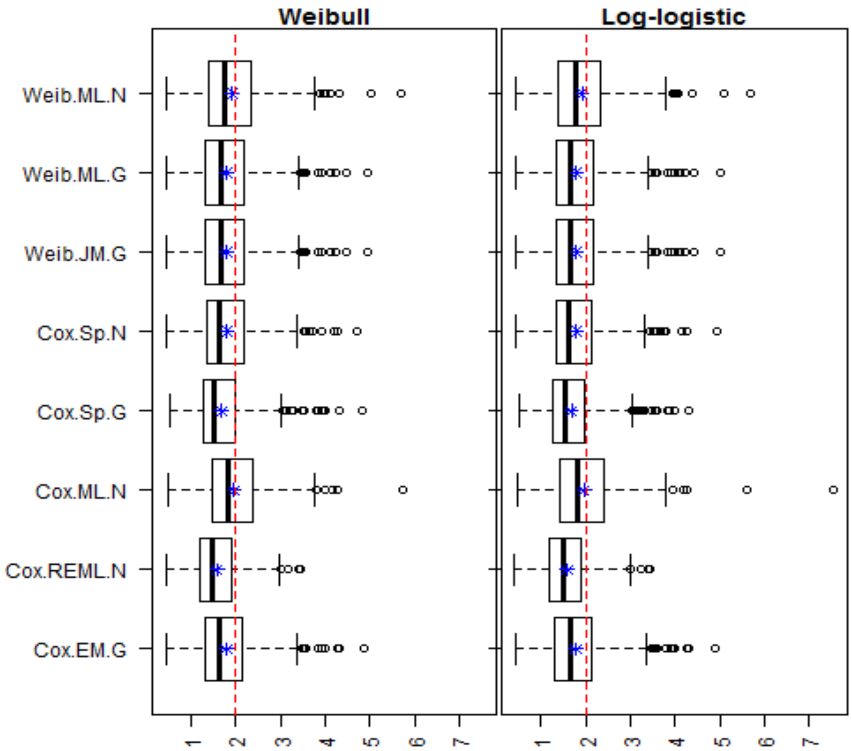
Web-Figure 13.60: θ estimates when moderate number of clusters were used under severe censoring and frailty generated from a gamma distribution.



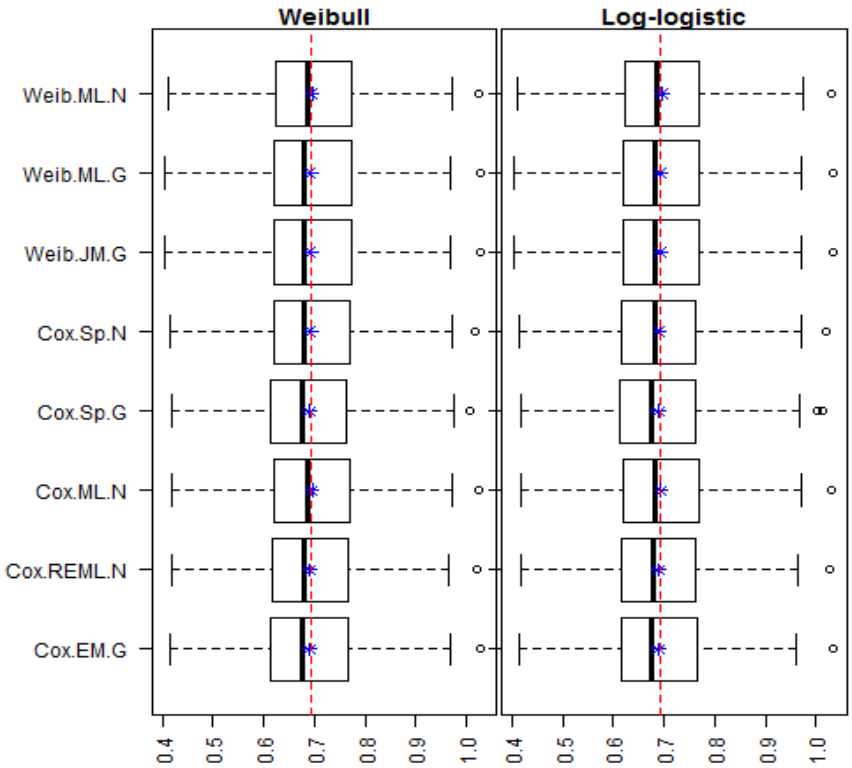
Web-Figure 13.61: β_1 estimates when moderate number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



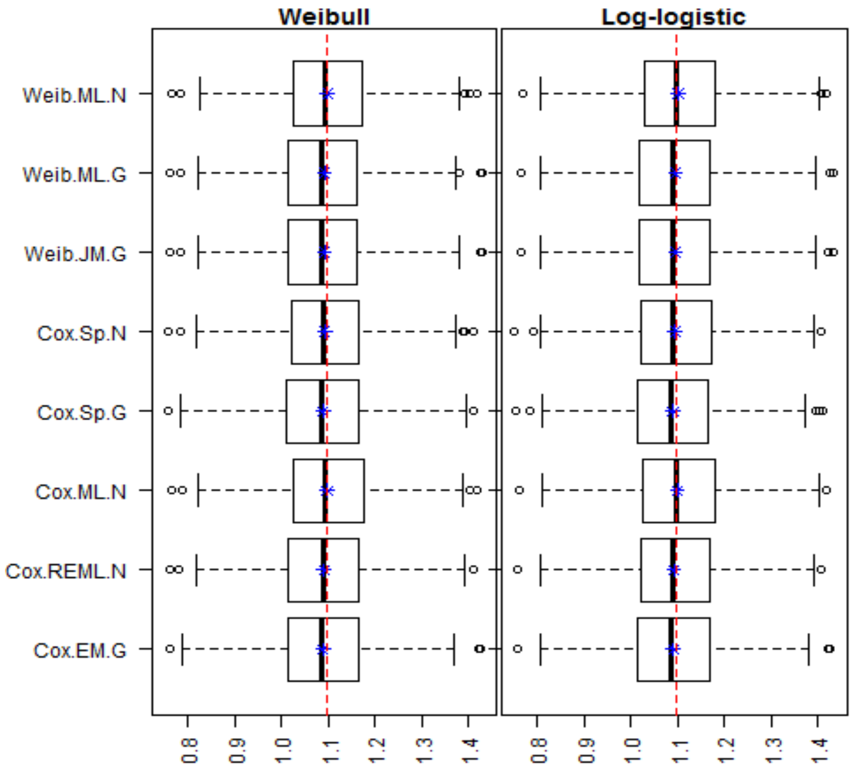
Web-Figure 13.62: β_2 estimates when moderate number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



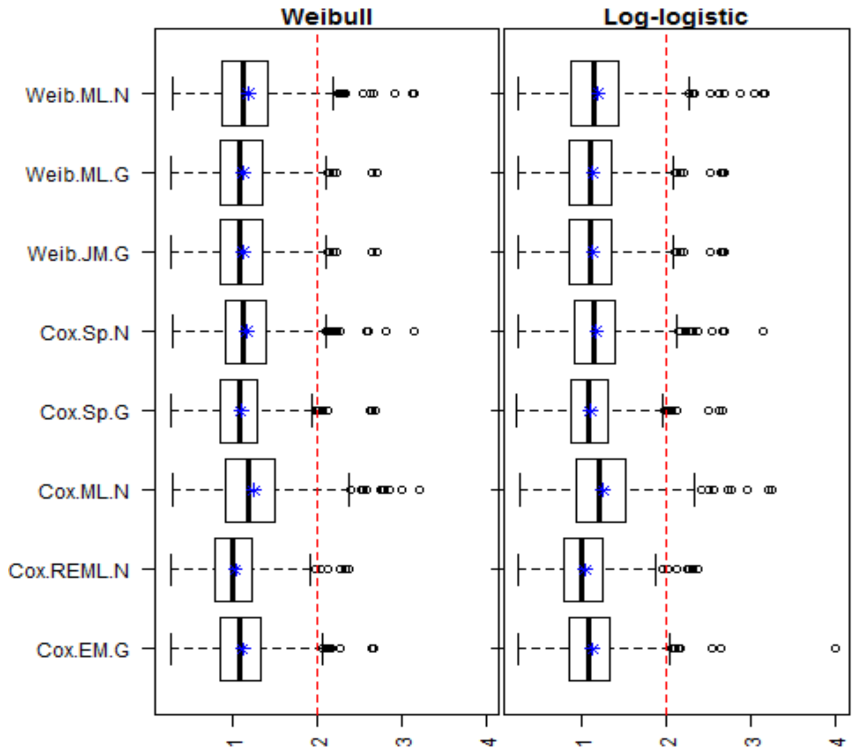
Web-Figure 13.63: θ estimates when moderate number of clusters were used under severe censoring and frailty generated from a log-normal distribution.



Web-Figure 13.64: β_1 estimates when moderate number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.



Web-Figure 13.65: β_2 estimates when moderate number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.

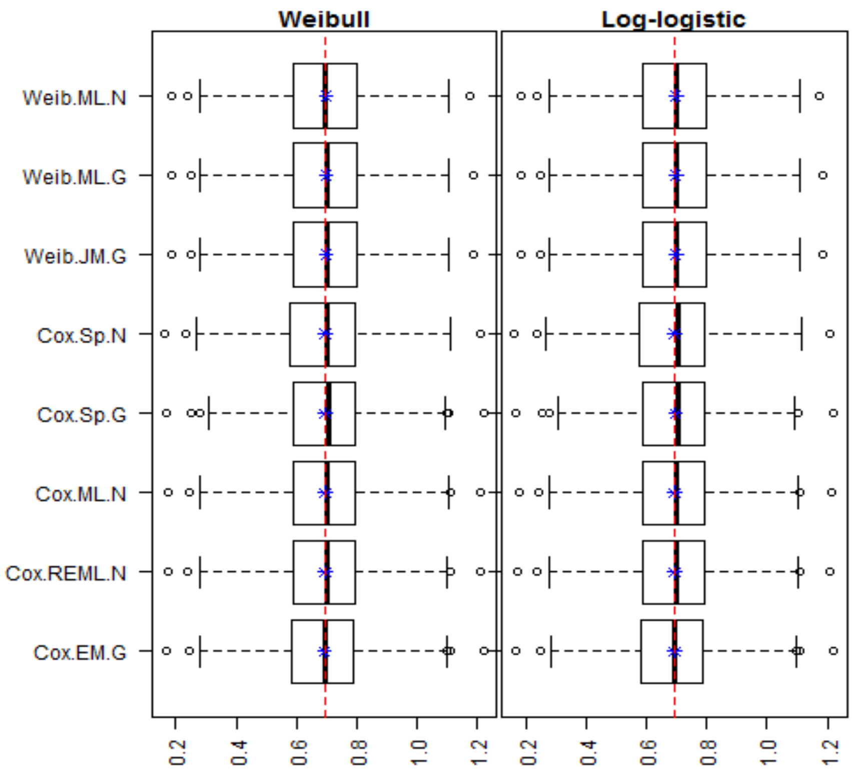


Web-Figure 13.66: θ estimates when moderate number of clusters were used under severe censoring and frailty generated from an inverse Gaussian distribution.

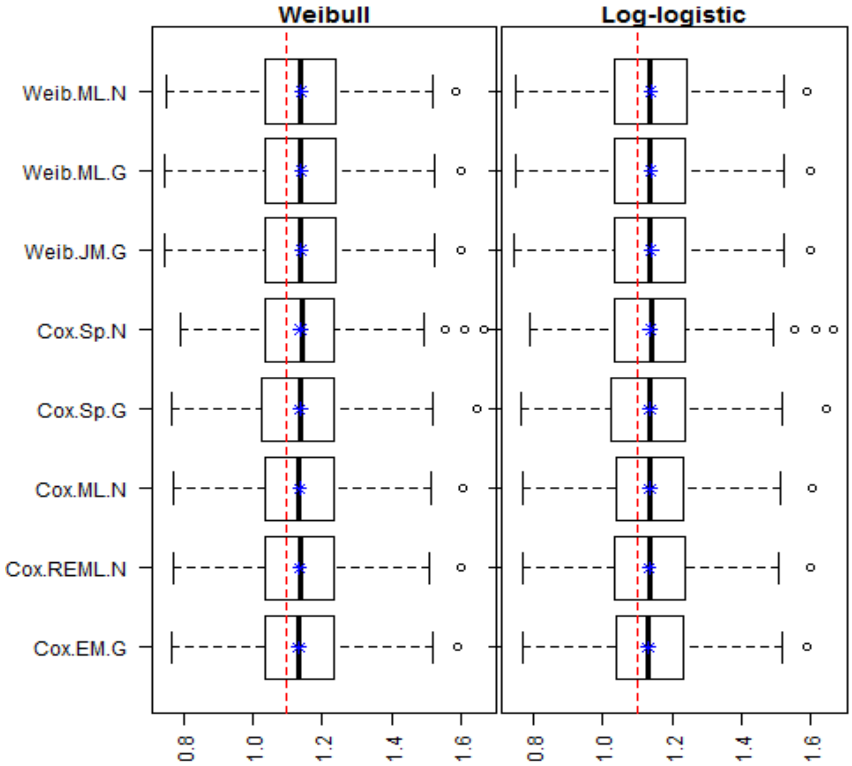
Censoring is severe (85%) **when** $N = 30$

13.2 Effects of changing θ for $N = 15$

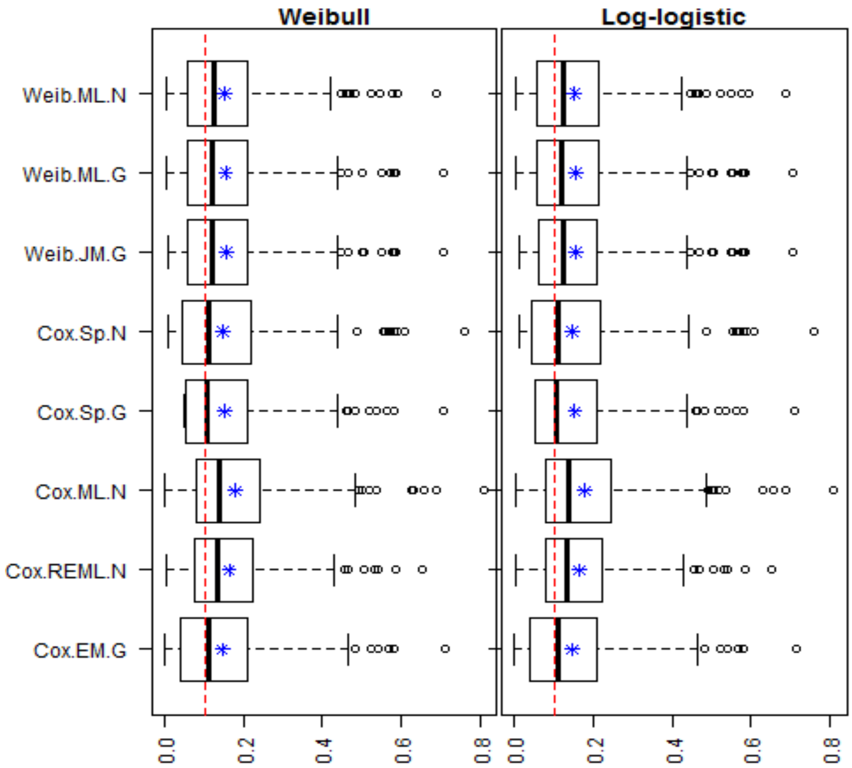
13.2.1 Heterogeneity Parameter, $\theta = 0.1$



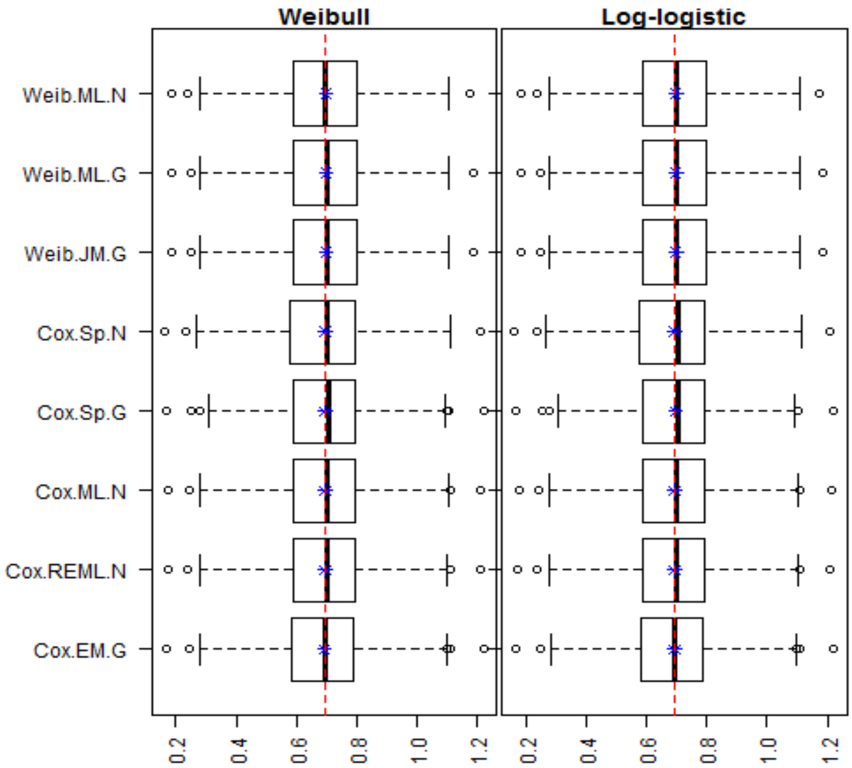
Web-Figure 13.67: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



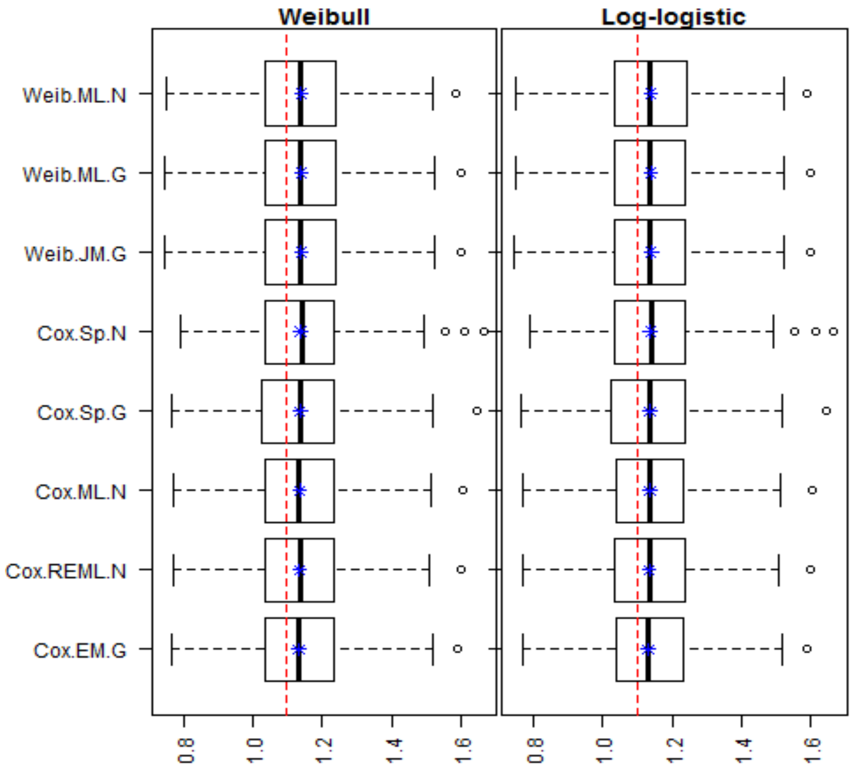
Web-Figure 13.68: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



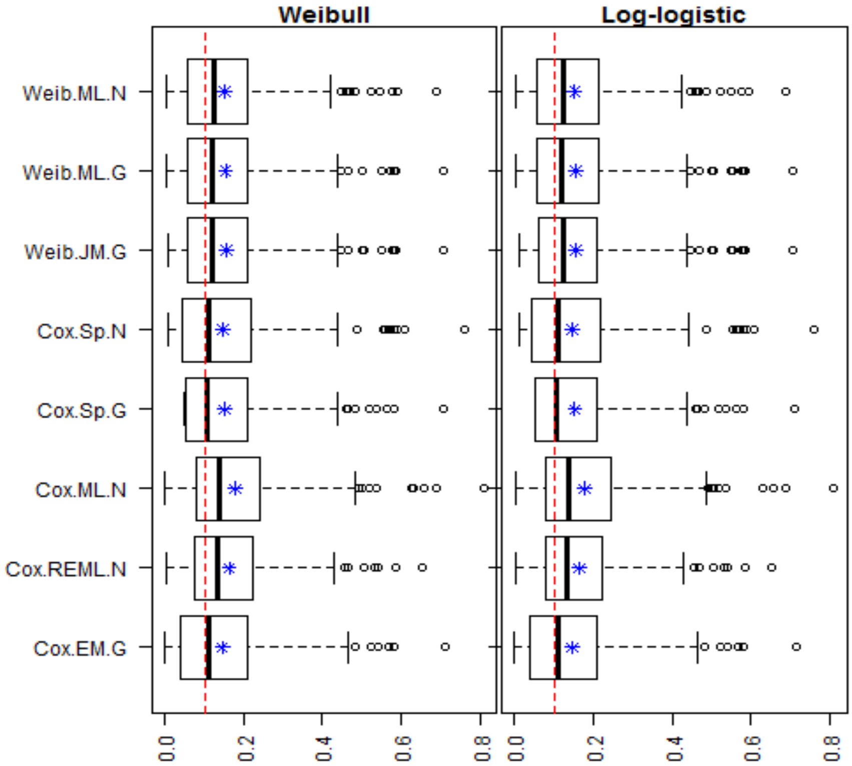
Web-Figure 13.69: θ estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



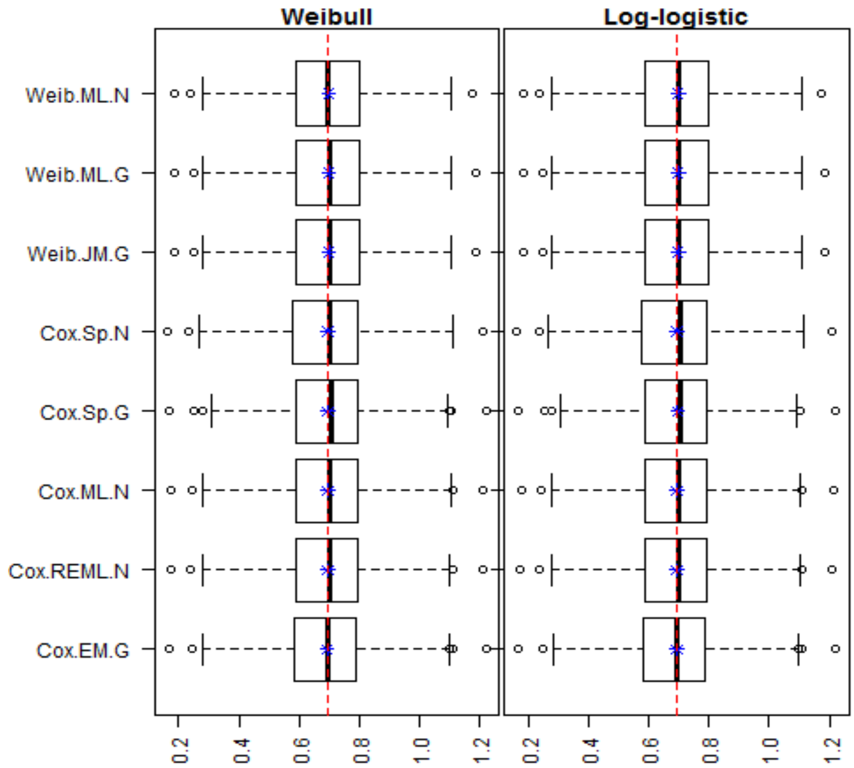
Web-Figure 13.70: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



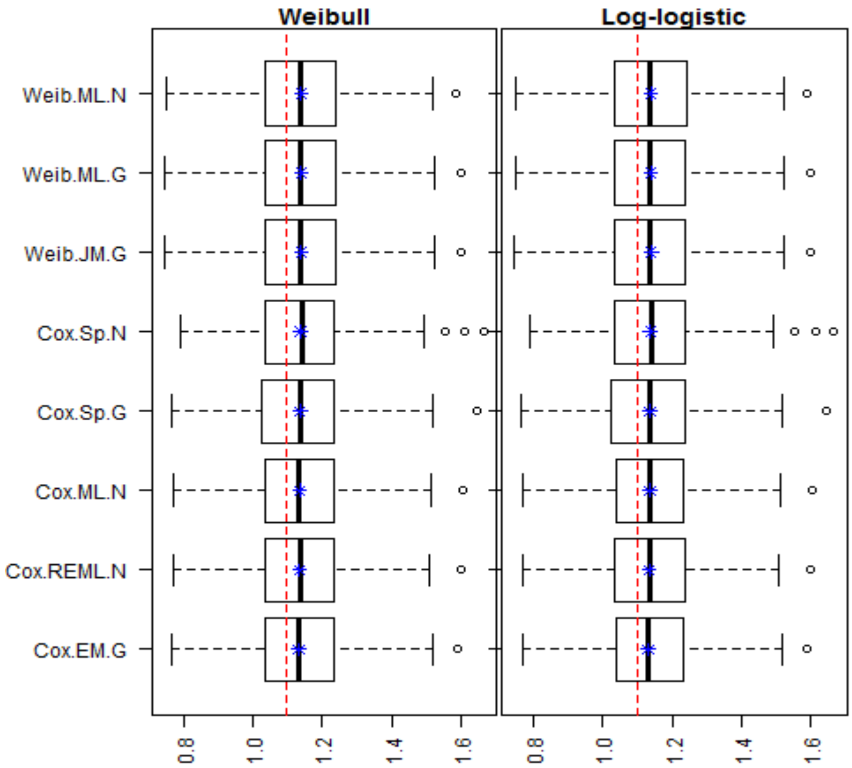
Web-Figure 13.71: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



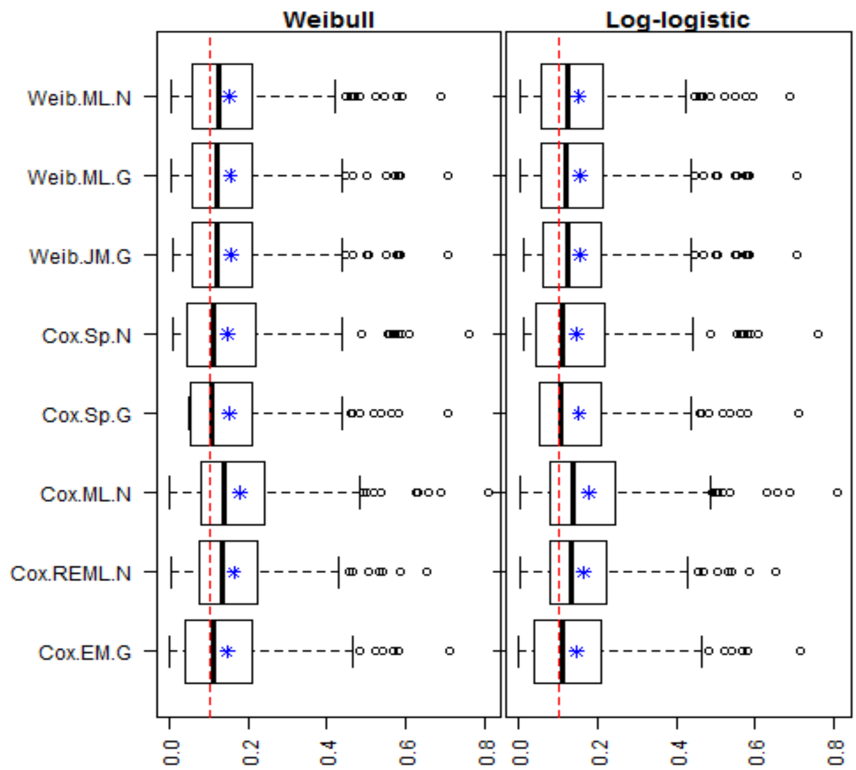
Web-Figure 13.72: θ estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



Web-Figure 13.73: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

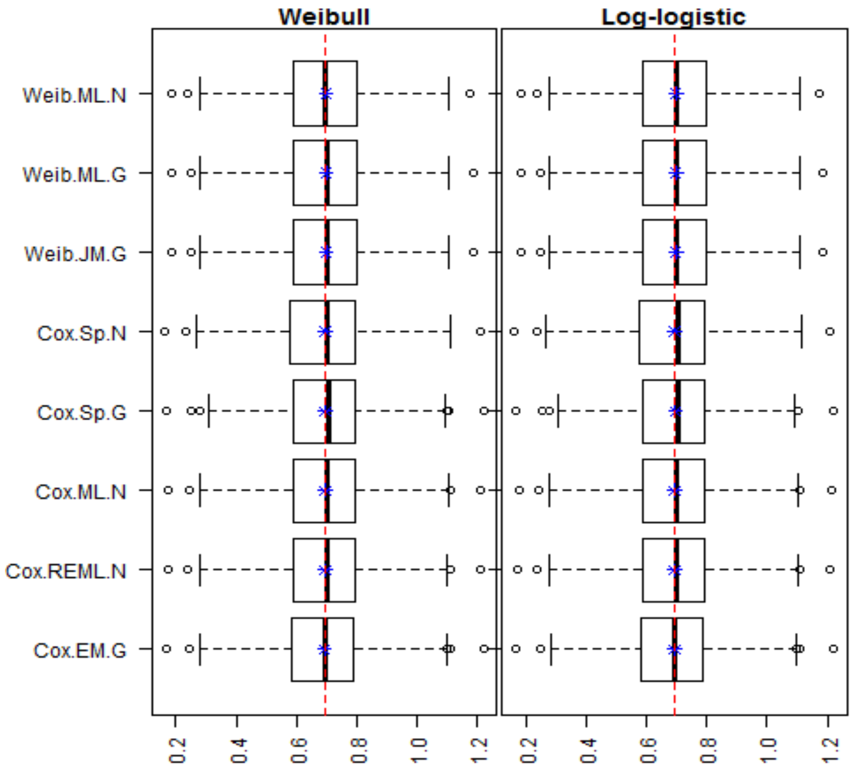


Web-Figure 13.74: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

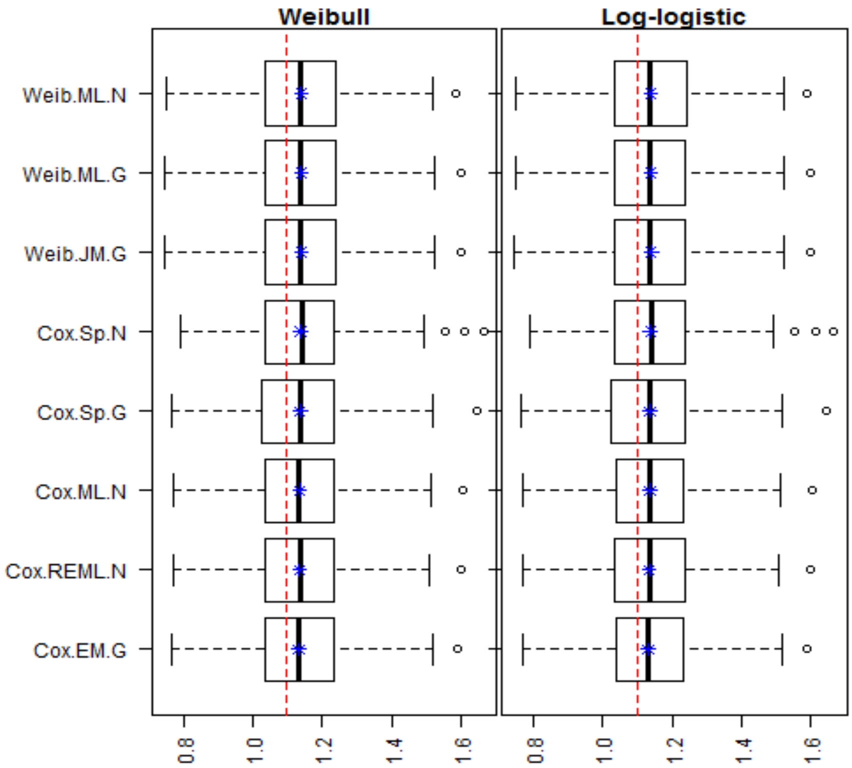


Web-Figure 13.75: θ estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

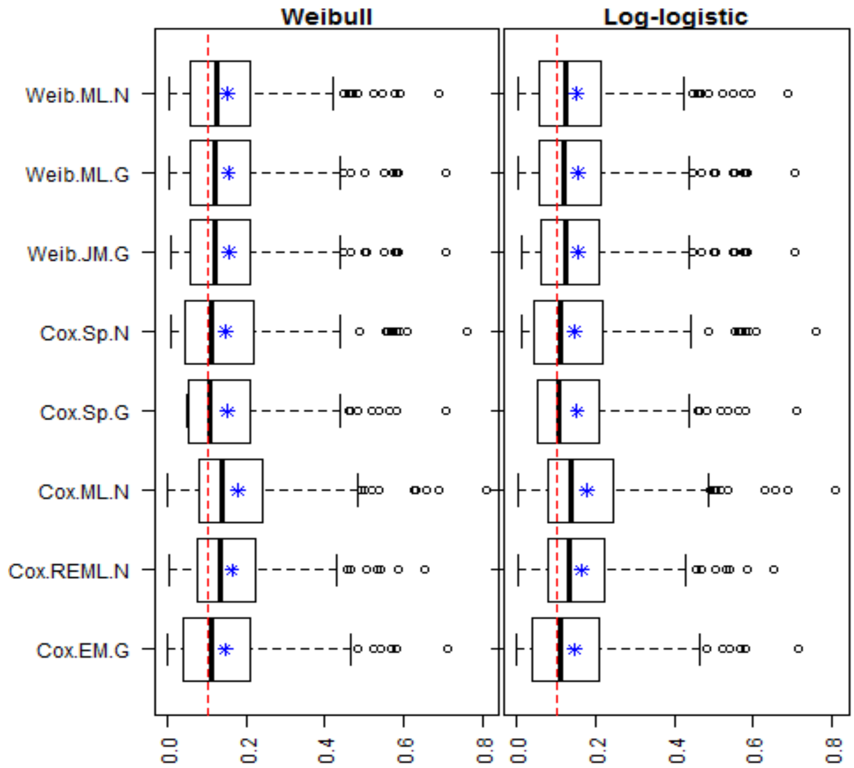
Mild censoring (15%) **when** $N = 15$



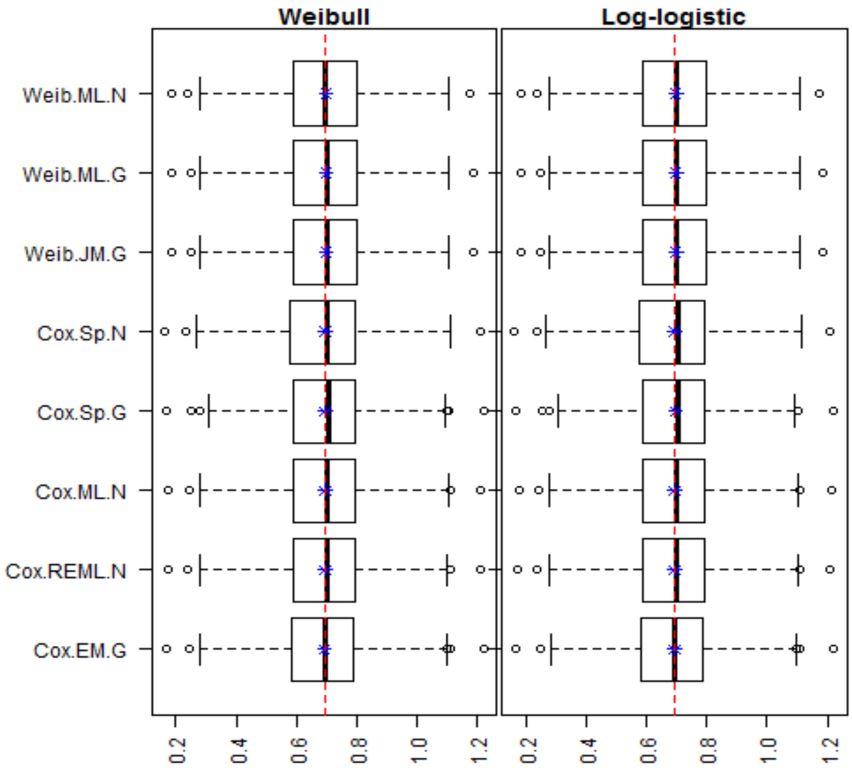
Web-Figure 13.76: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



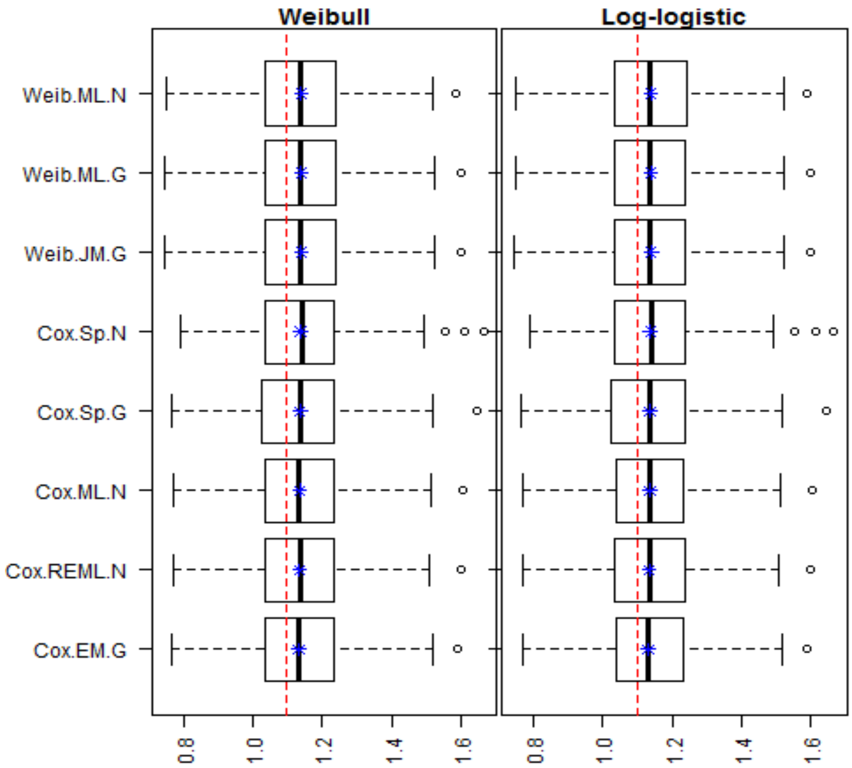
Web-Figure 13.77: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



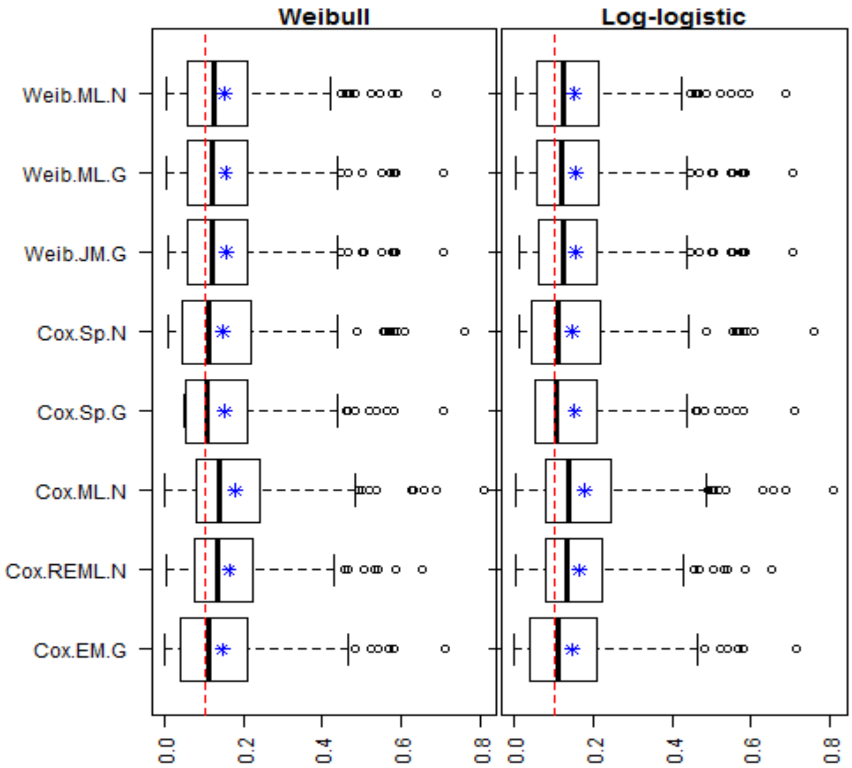
Web-Figure 13.78: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



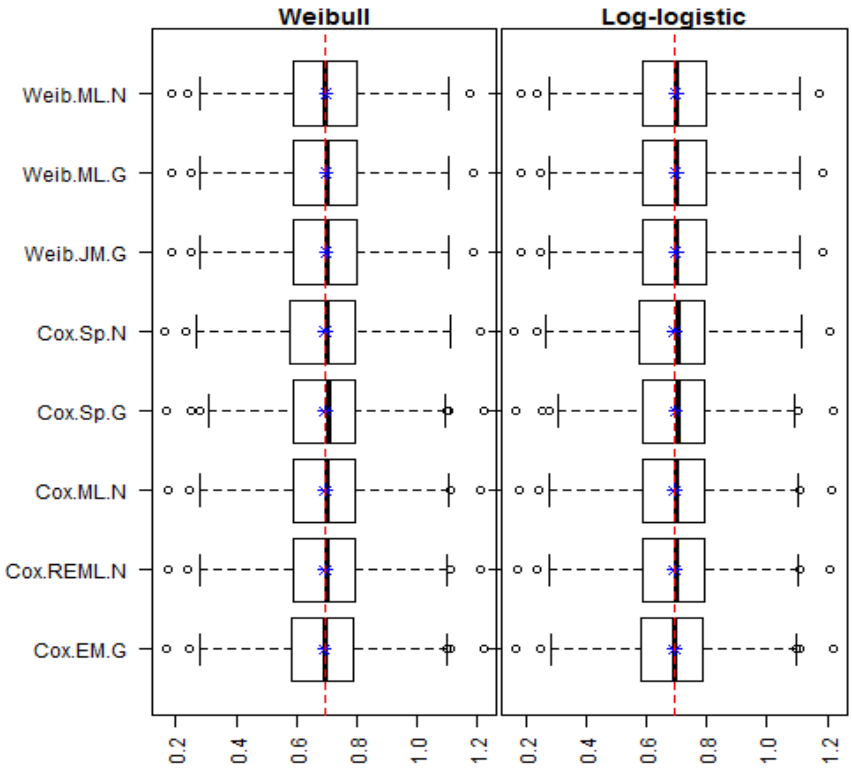
Web-Figure 13.79: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



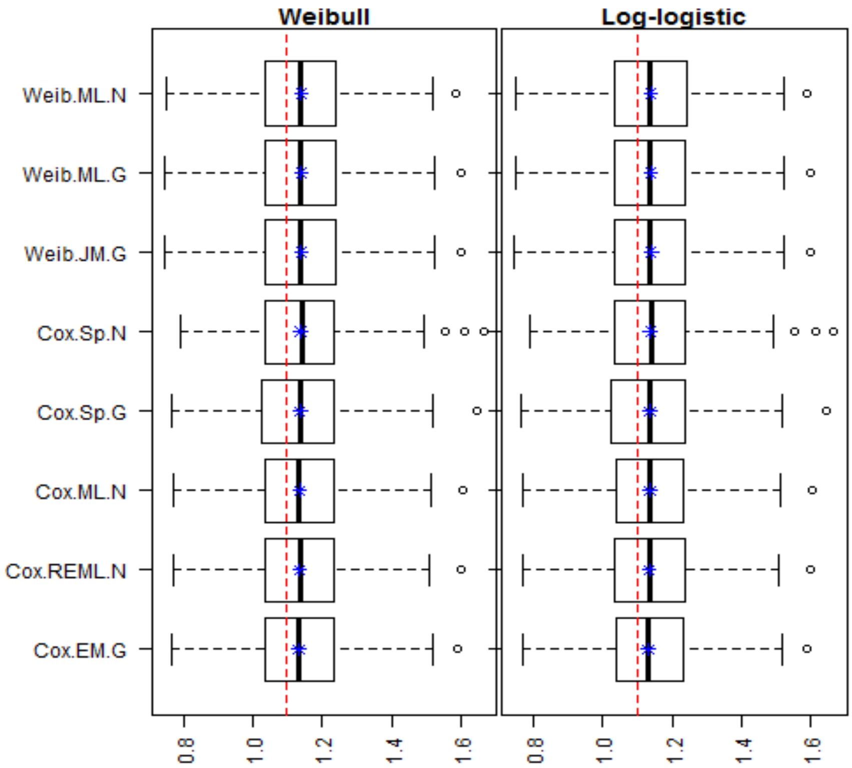
Web-Figure 13.80: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



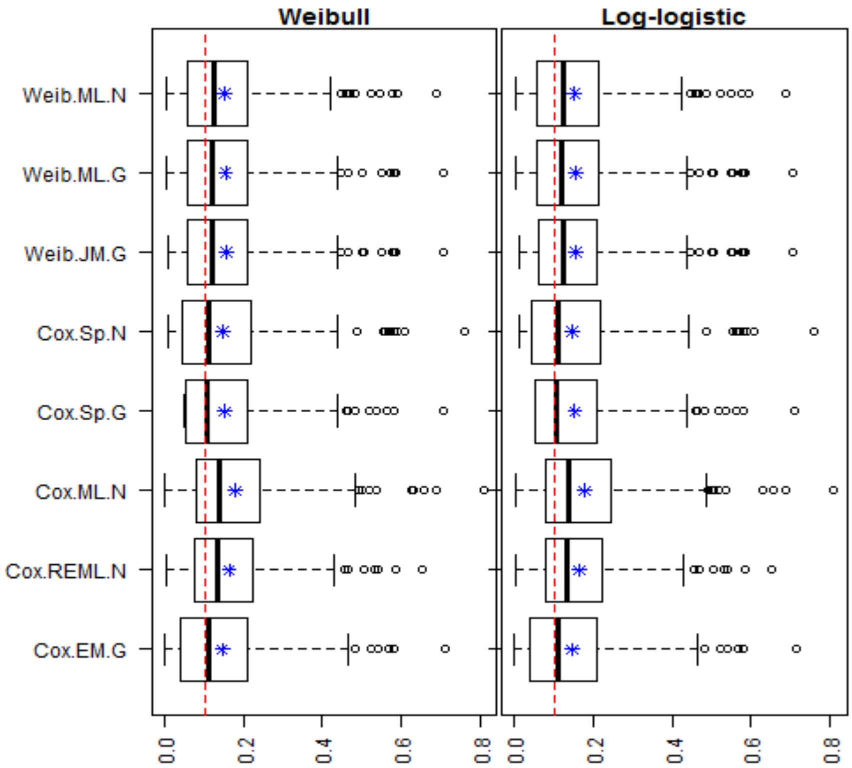
Web-Figure 13.81: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



Web-Figure 13.82: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

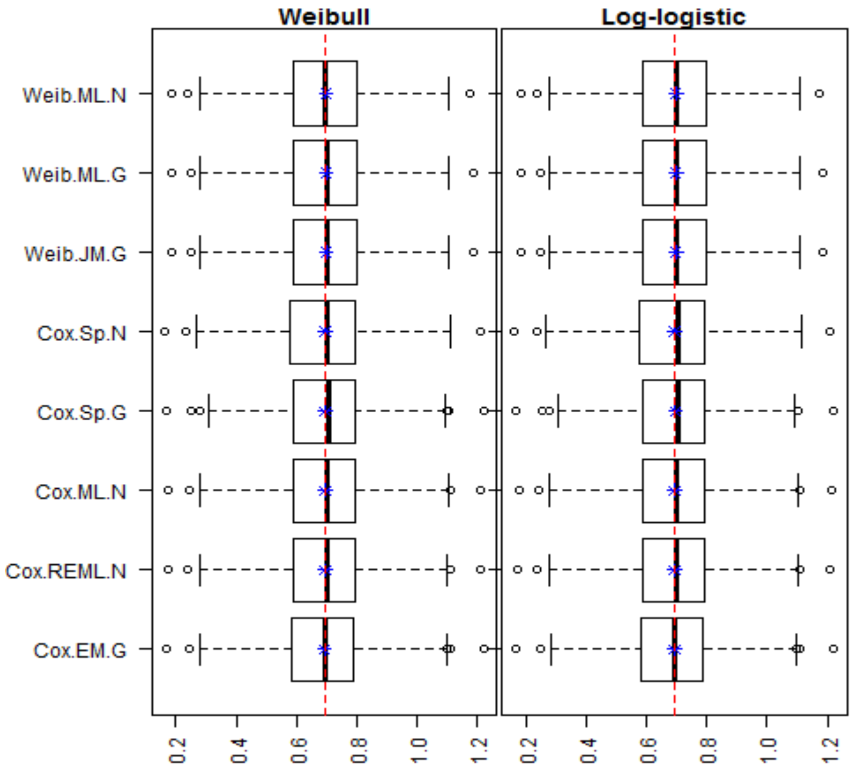


Web-Figure 13.83: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

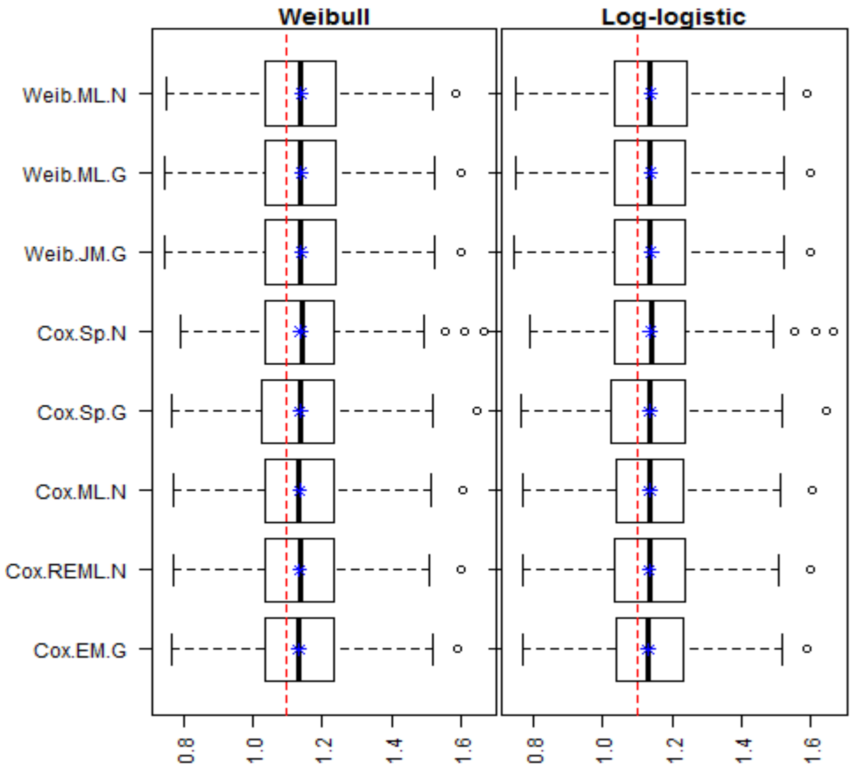


Web-Figure 13.84: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

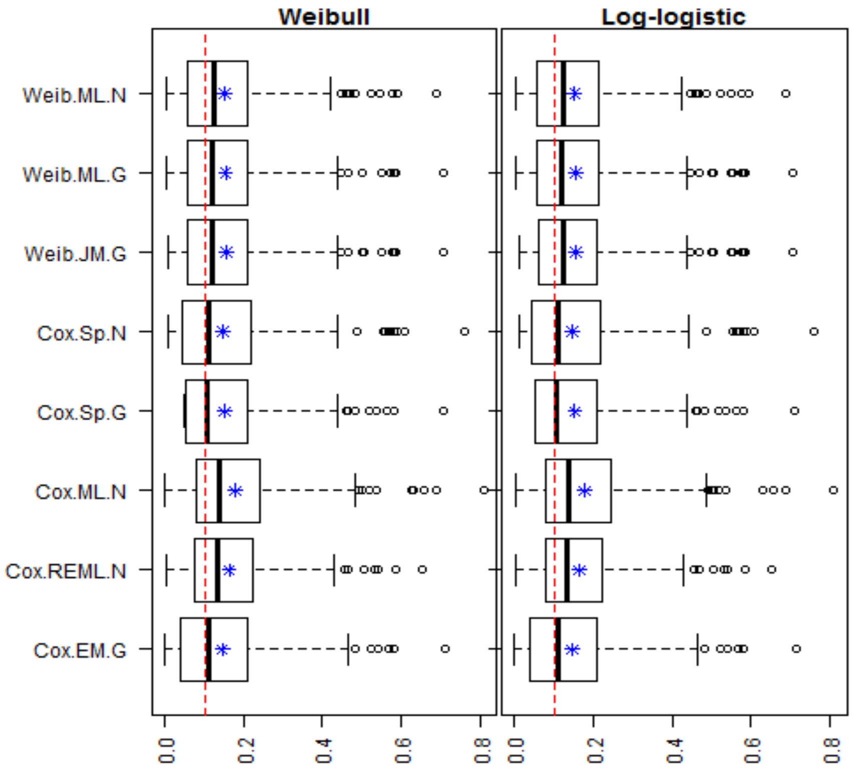
Moderate censoring (45%) **when** $N = 15$



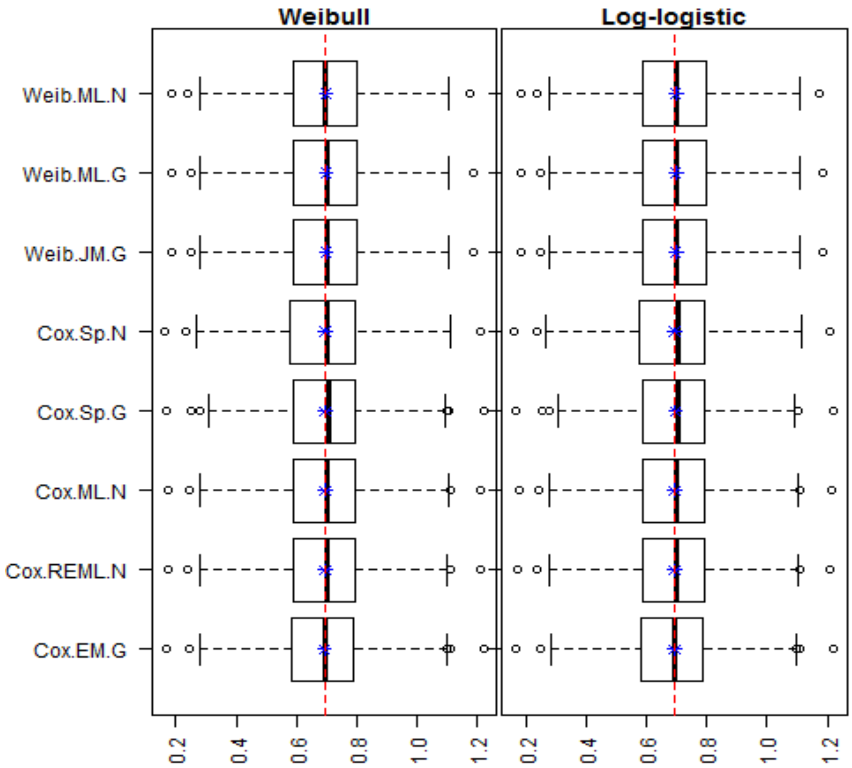
Web-Figure 13.85: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



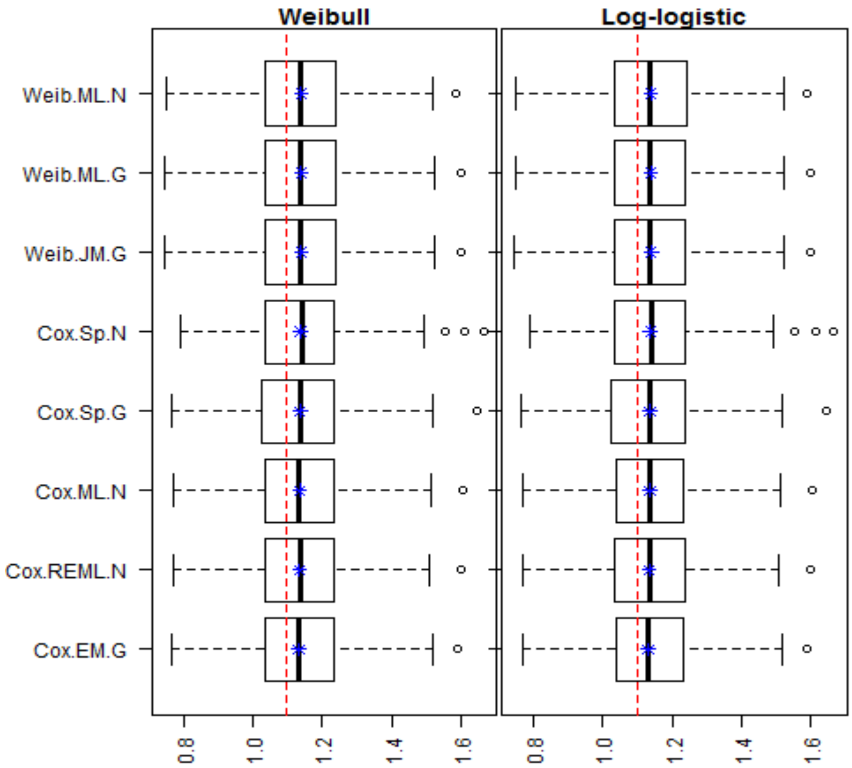
Web-Figure 13.86: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



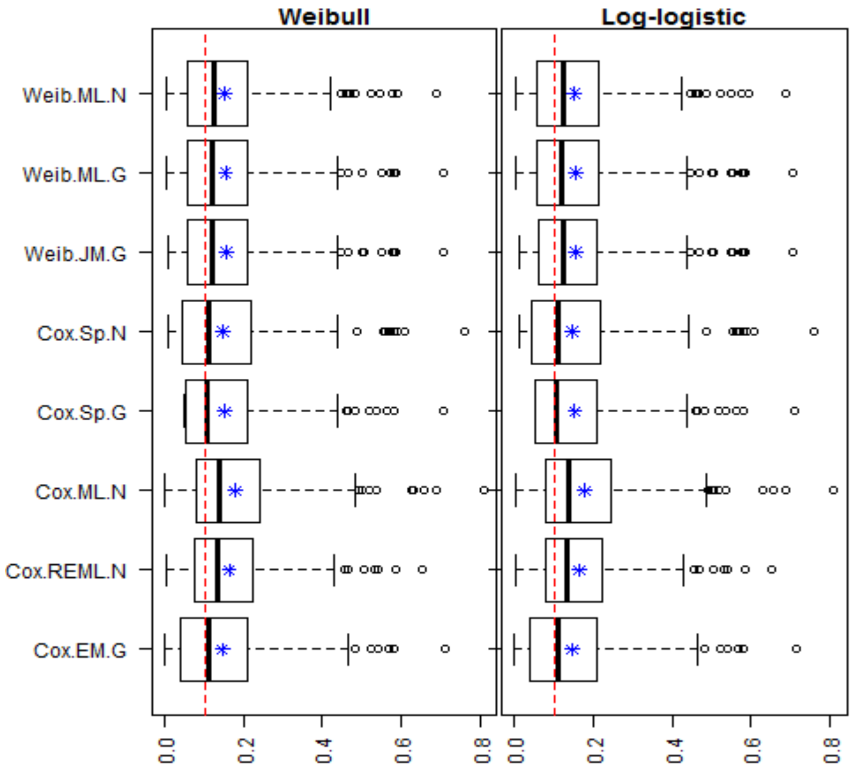
Web-Figure 13.87: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.1$.



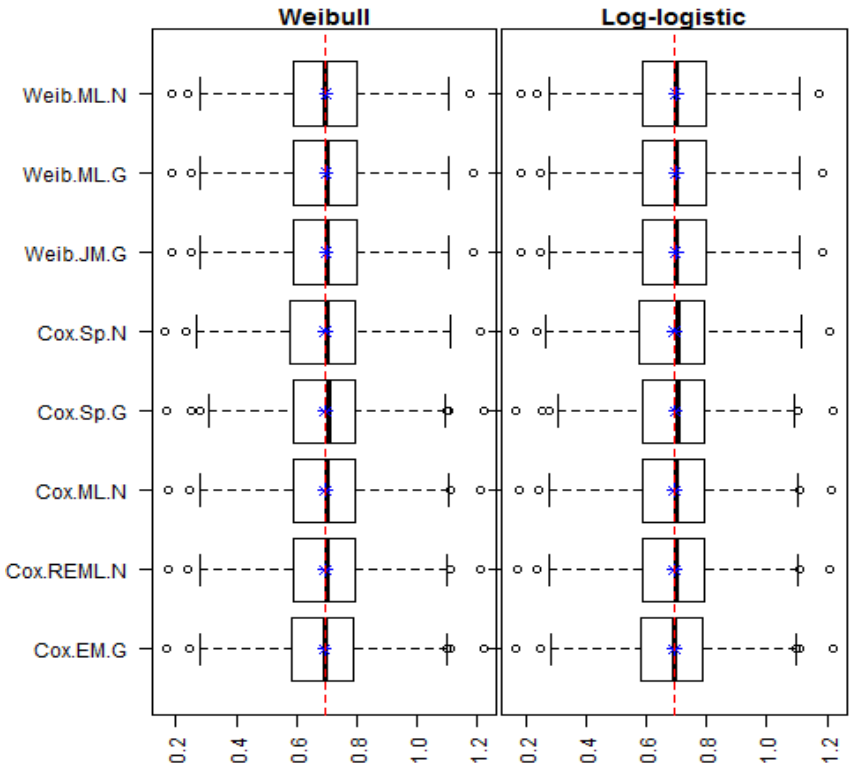
Web-Figure 13.88: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



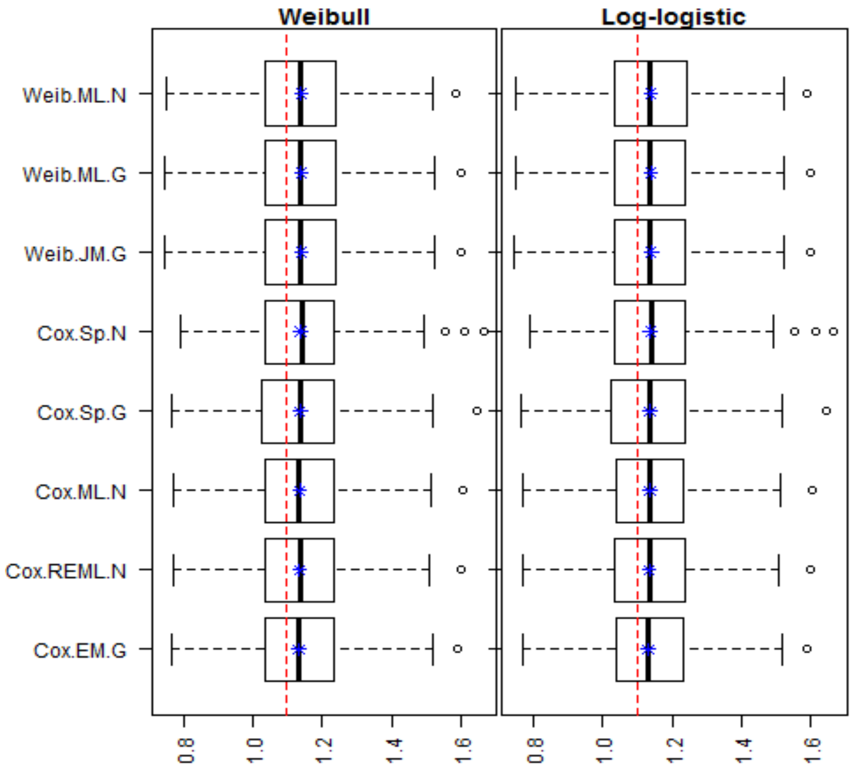
Web-Figure 13.89: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



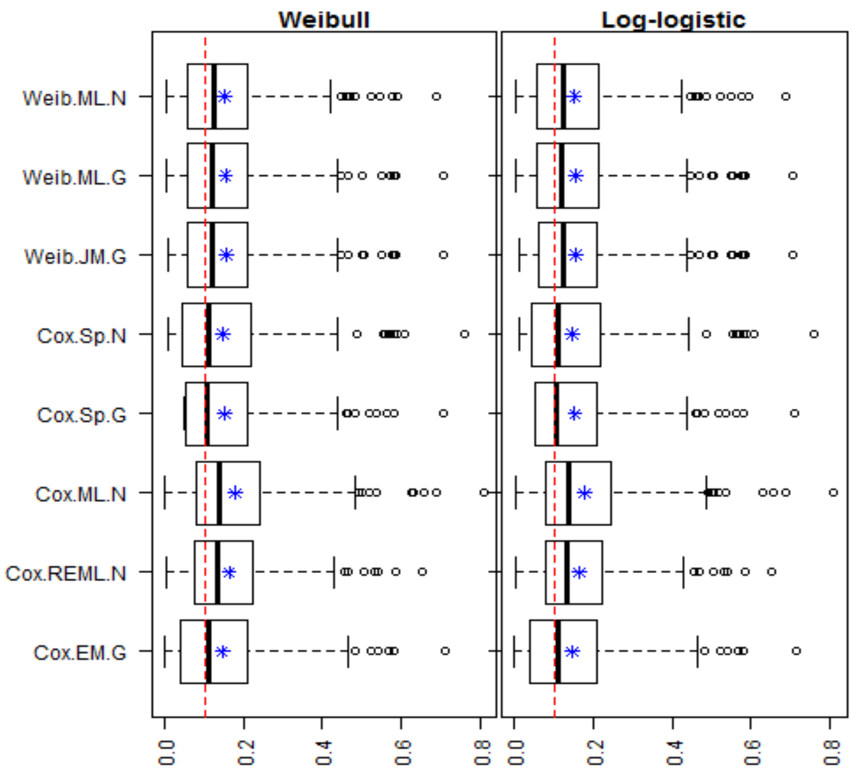
Web-Figure 13.90: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.1$.



Web-Figure 13.91: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.



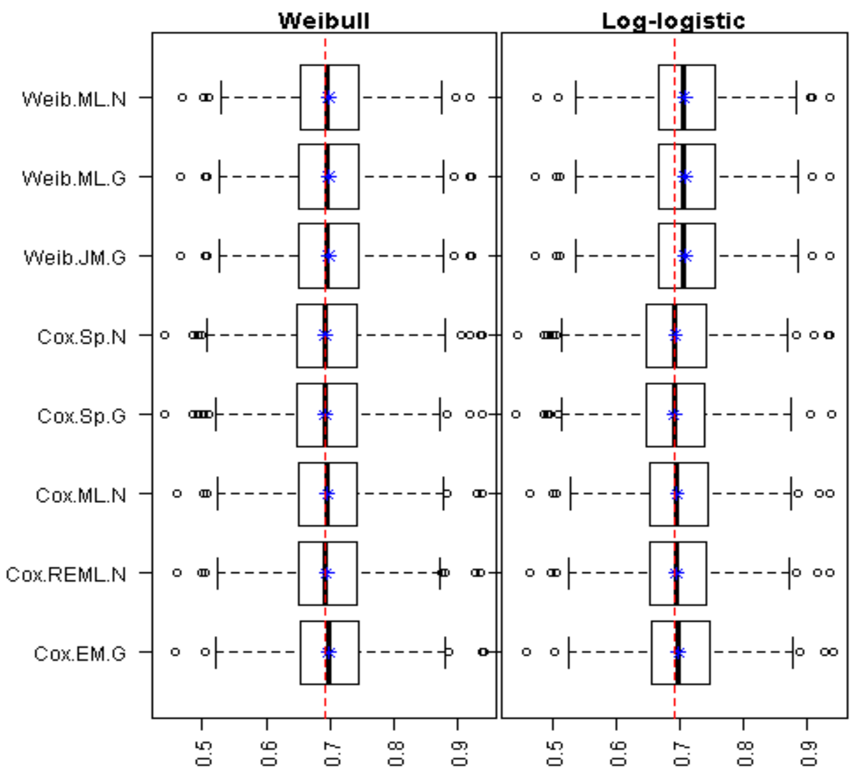
Web-Figure 13.92: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.



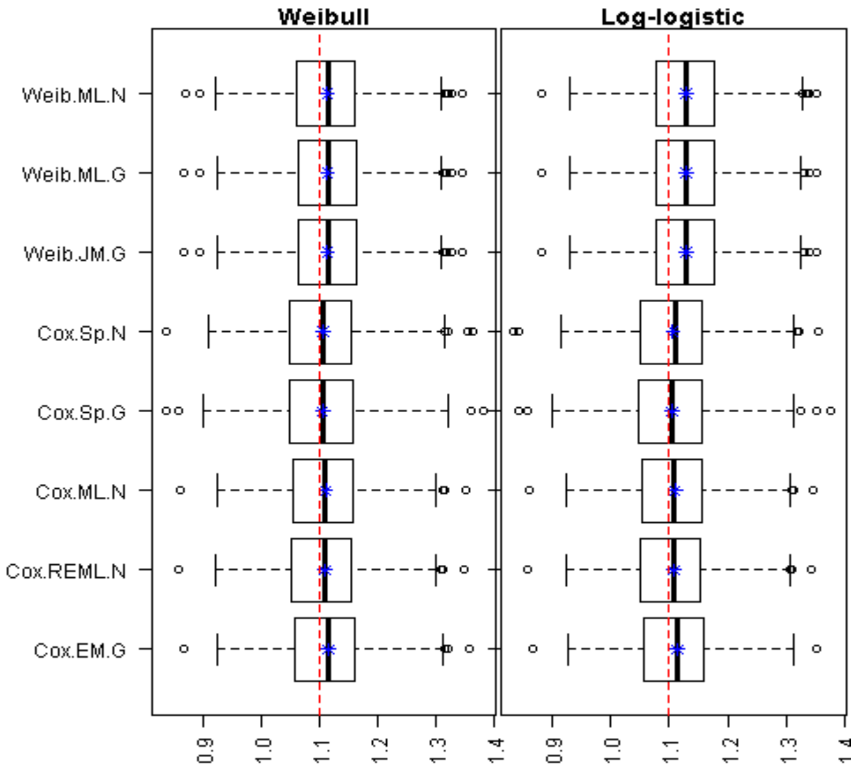
Web-Figure 13.93: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.1$.

Severe censoring (85%) **when** $N = 15$

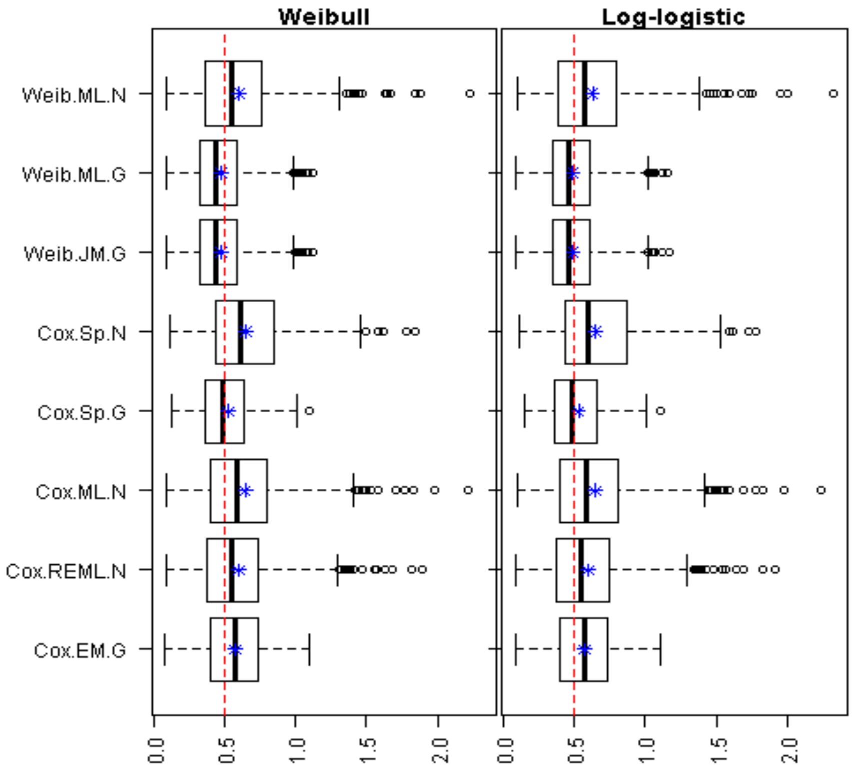
13.2.2 Heterogeneity Parameter, $\theta = 0.5$



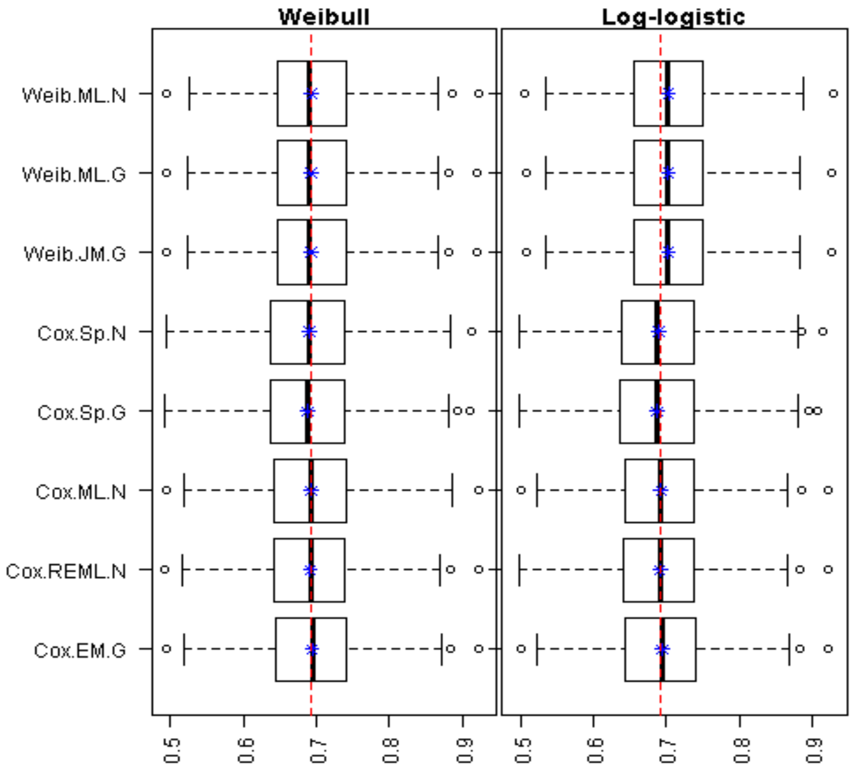
Web-Figure 13.94: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



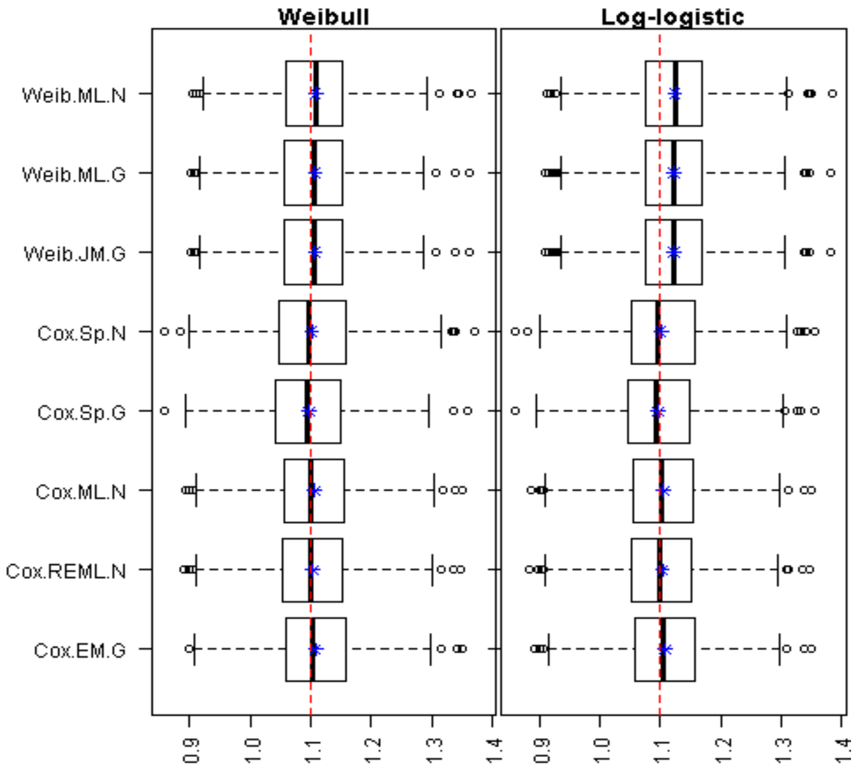
Web-Figure 13.95: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



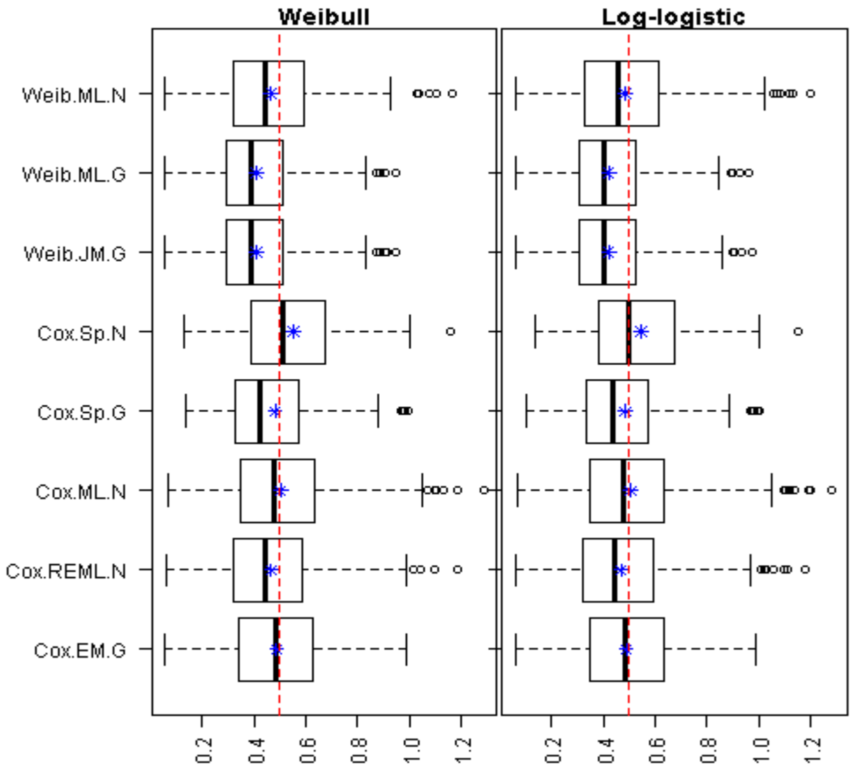
Web-Figure 13.96: θ estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



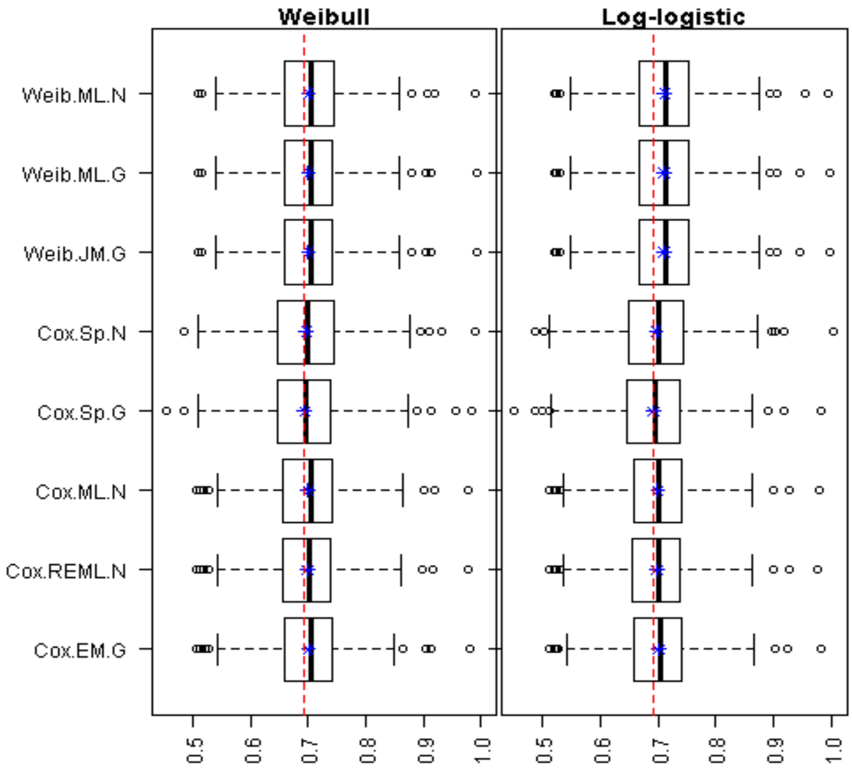
Web-Figure 13.97: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



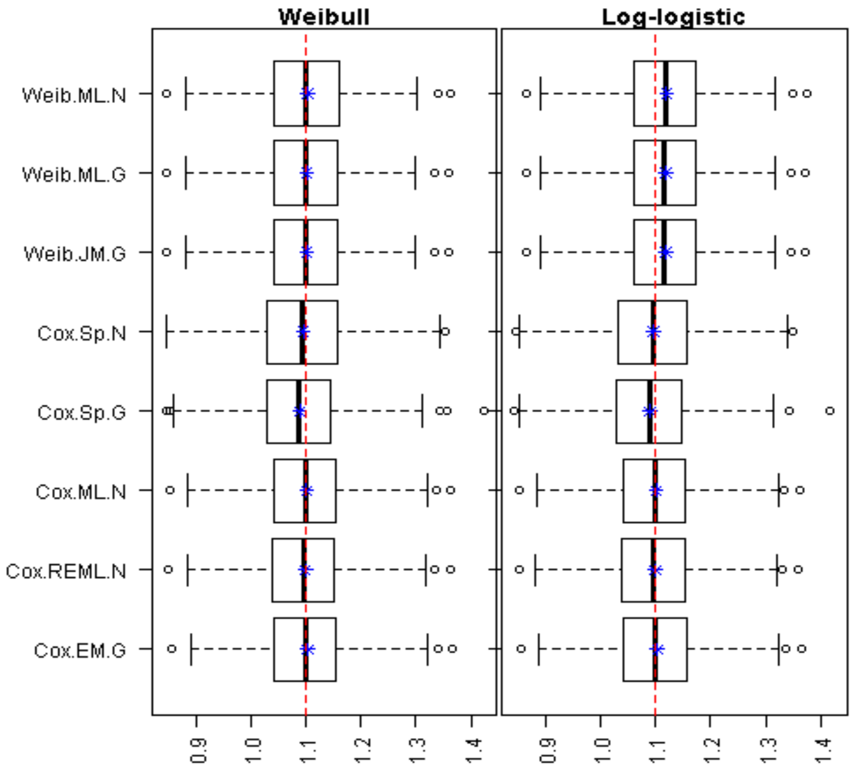
Web-Figure 13.98: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



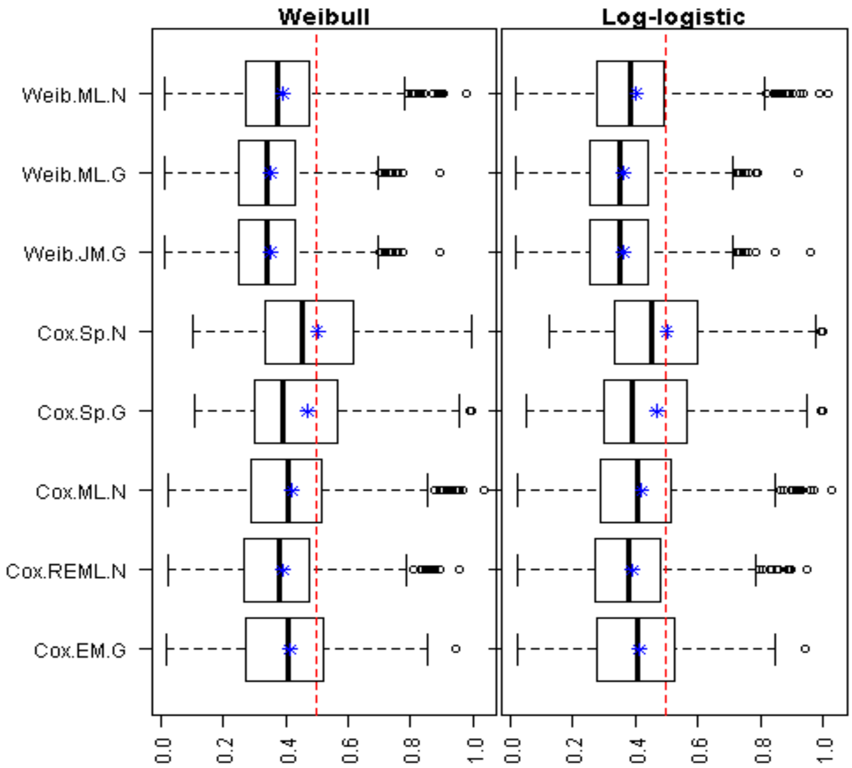
Web-Figure 13.99: θ estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



Web-Figure 13.100: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

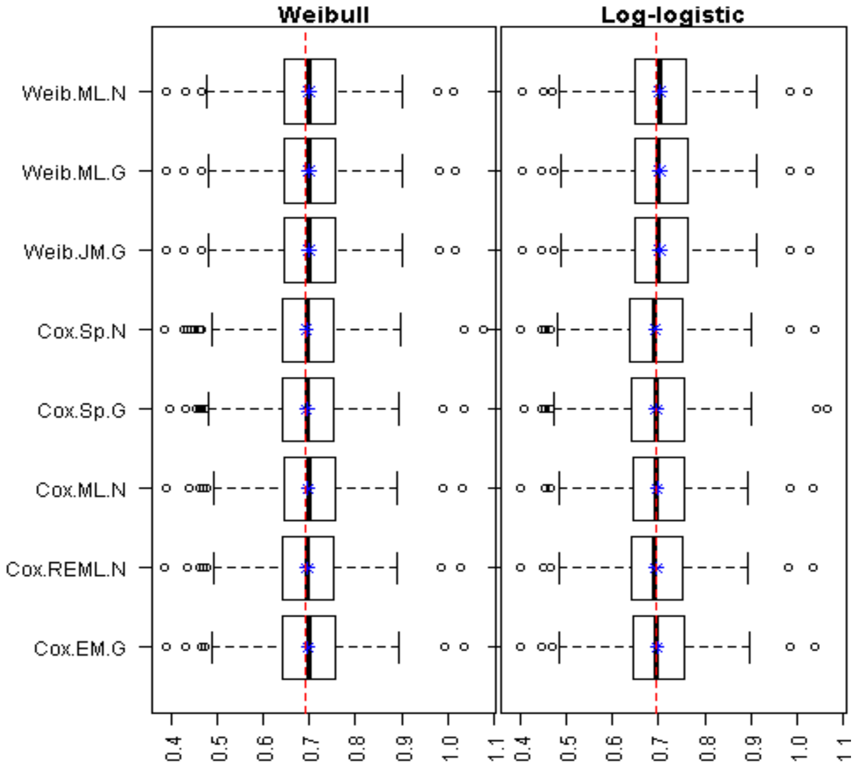


Web-Figure 13.101: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

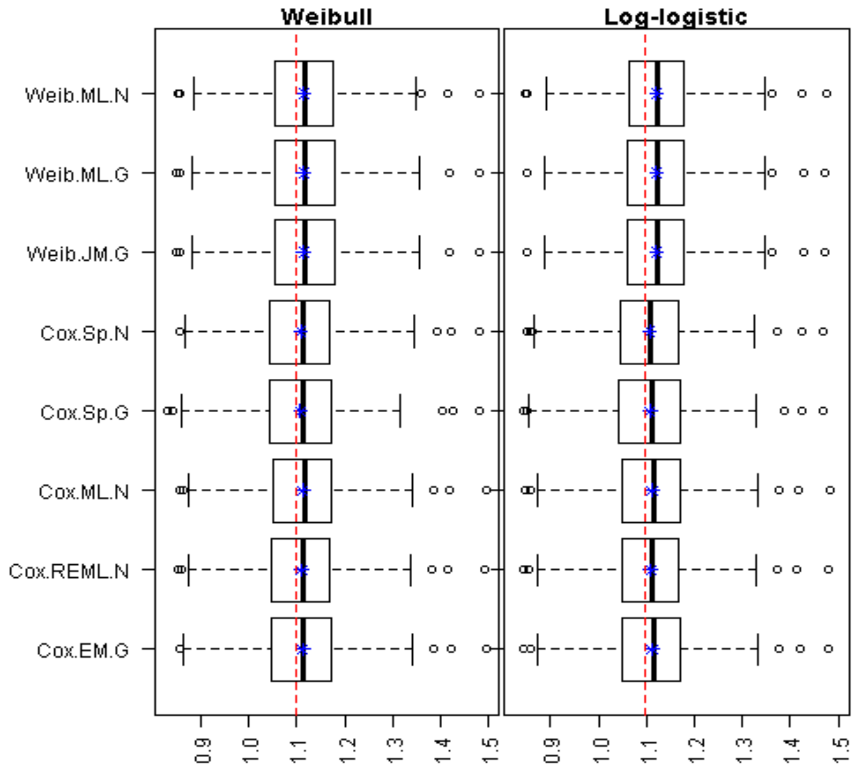


Web-Figure 13.102: θ estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

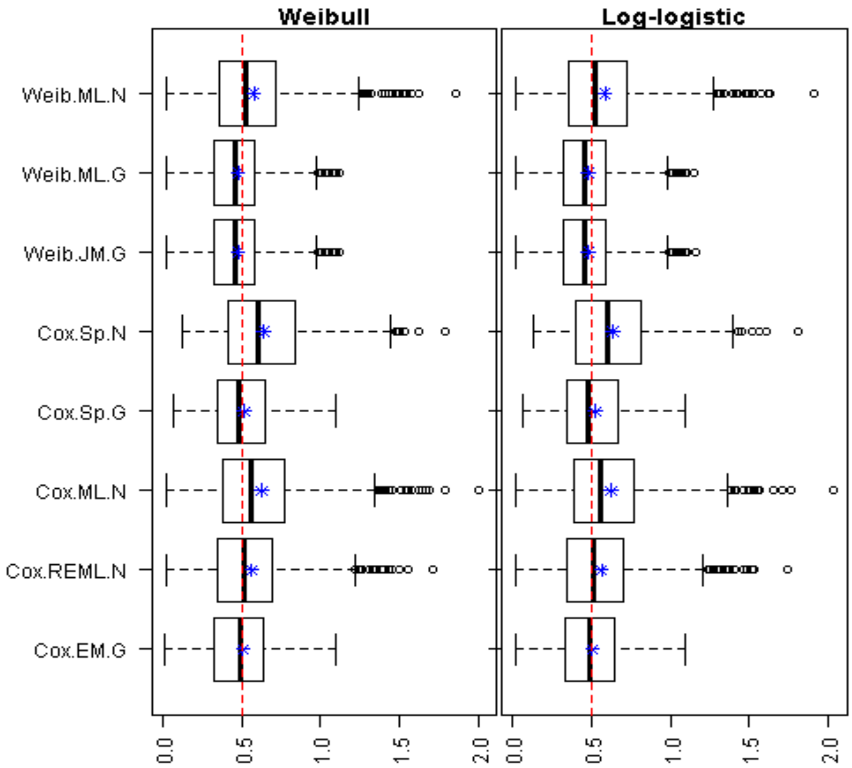
Mild censoring (15%) **when** $N = 15$



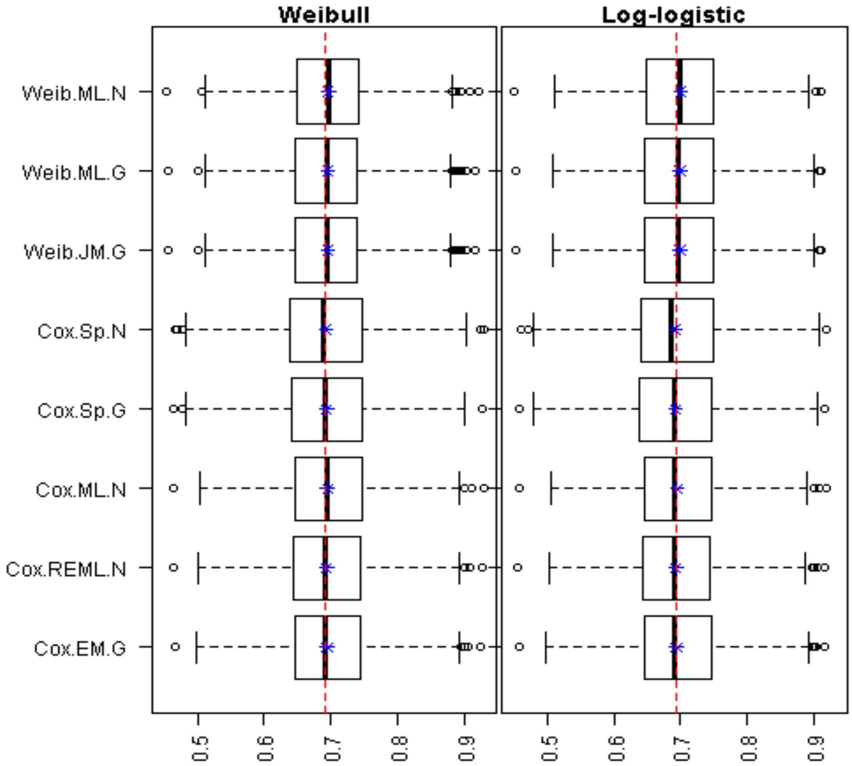
Web-Figure 13.103: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



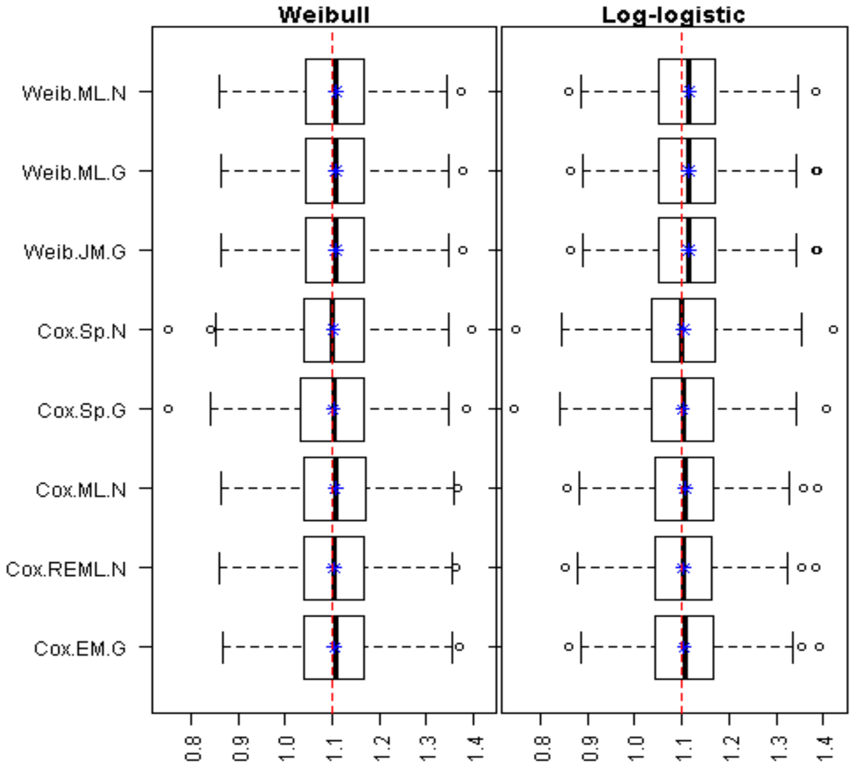
Web-Figure 13.104: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



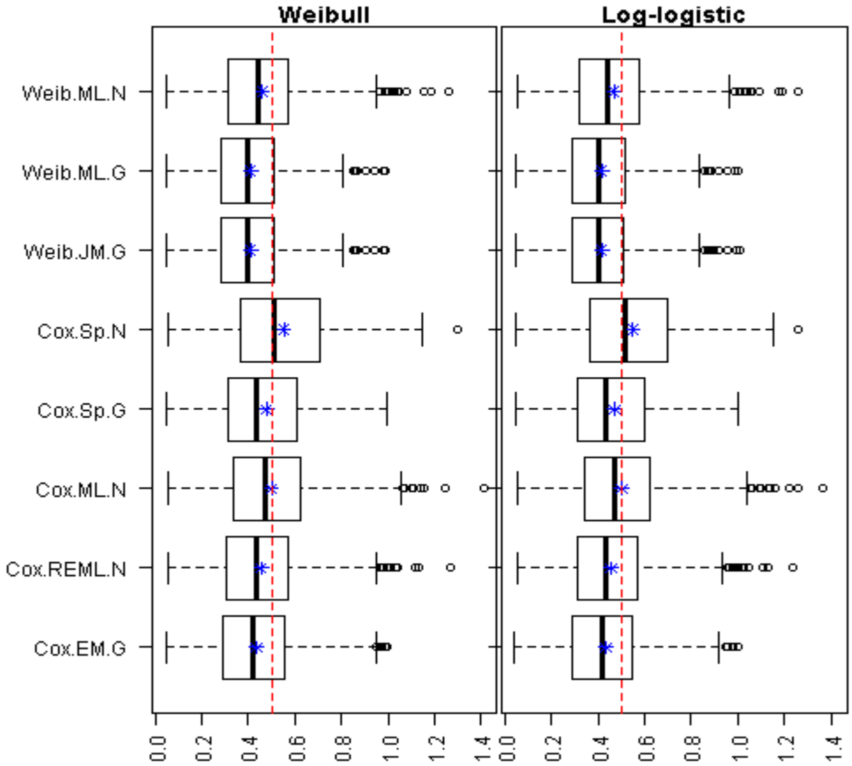
Web-Figure 13.105: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



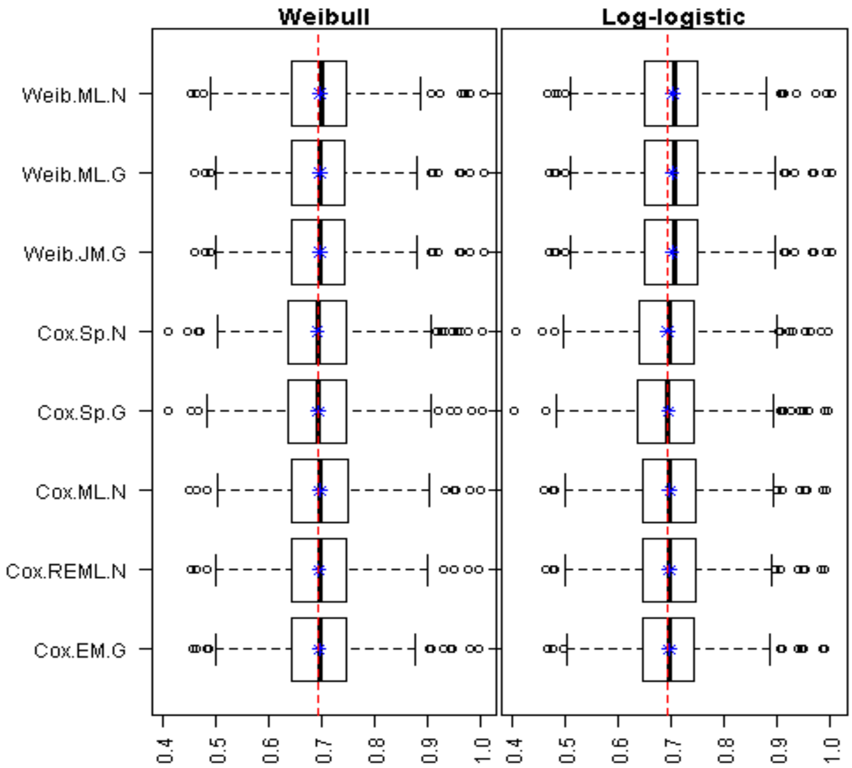
Web-Figure 13.106: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



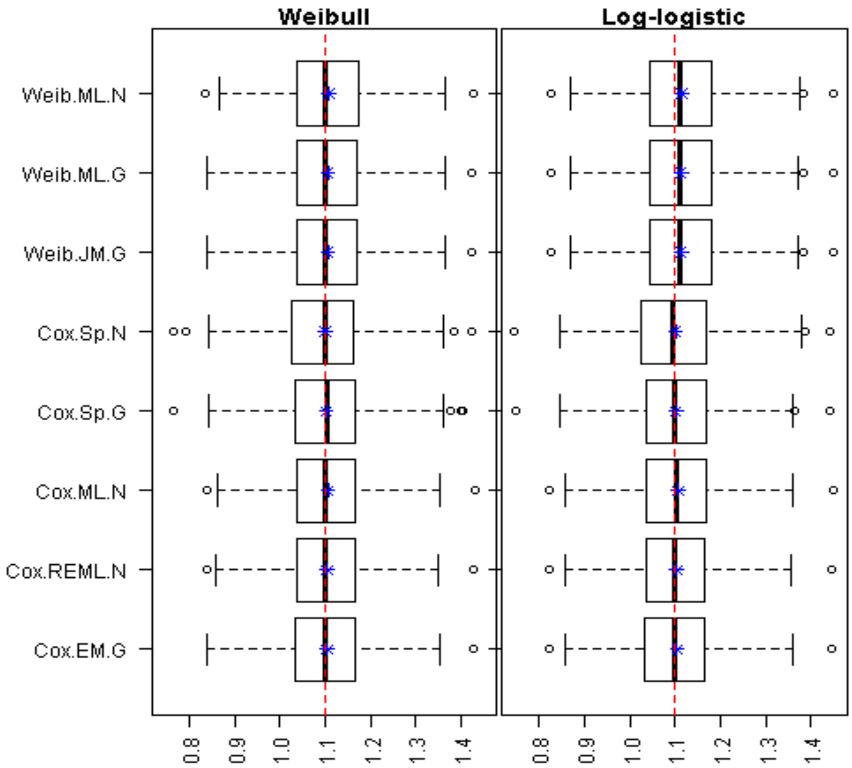
Web-Figure 13.107: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



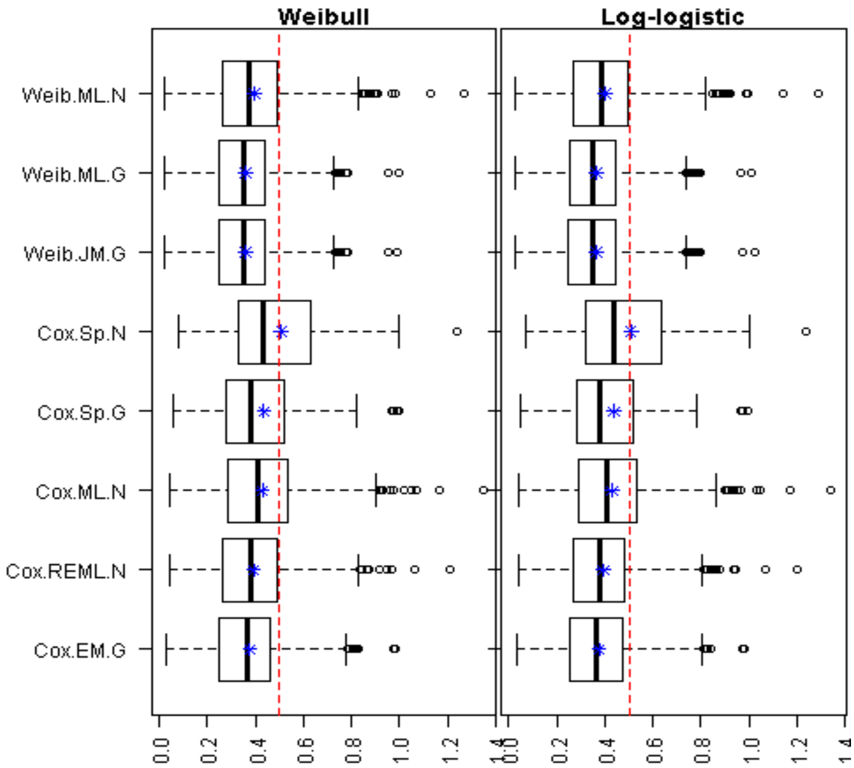
Web-Figure 13.108: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



Web-Figure 13.109: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

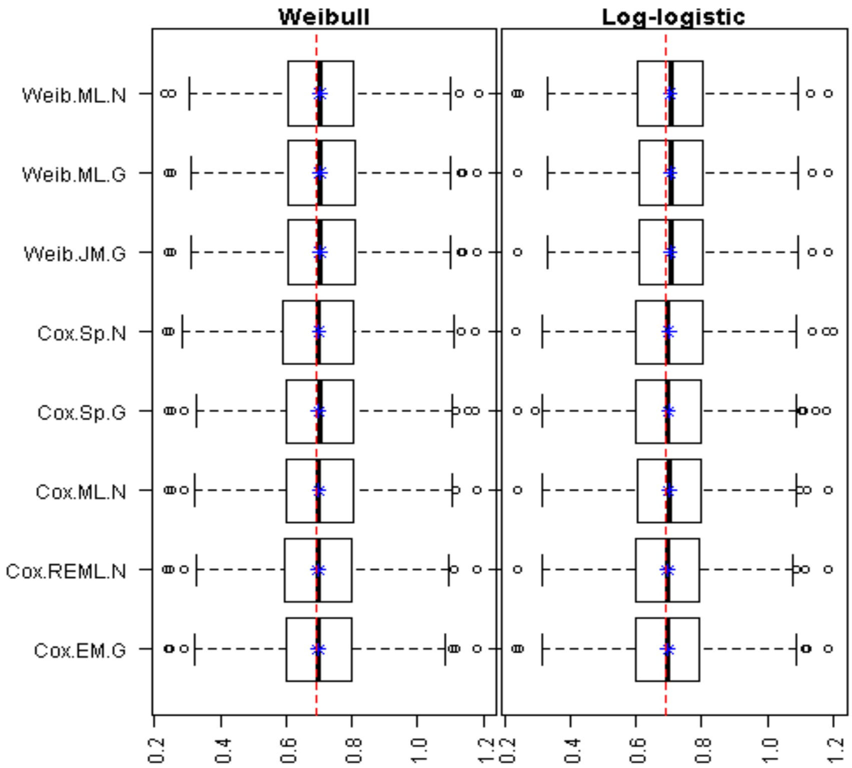


Web-Figure 13.110: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

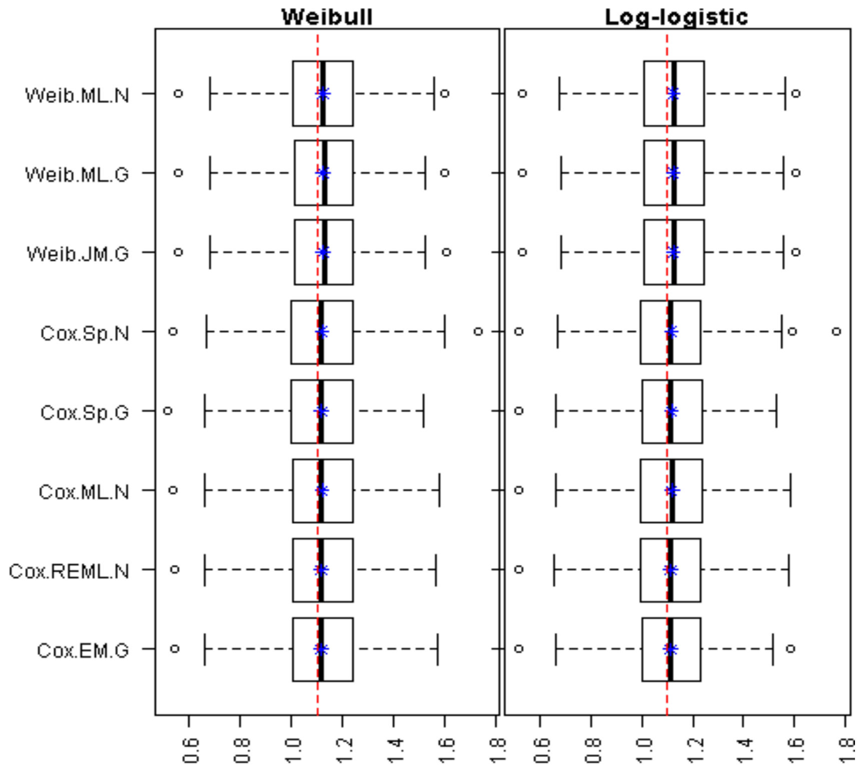


Web-Figure 13.111: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

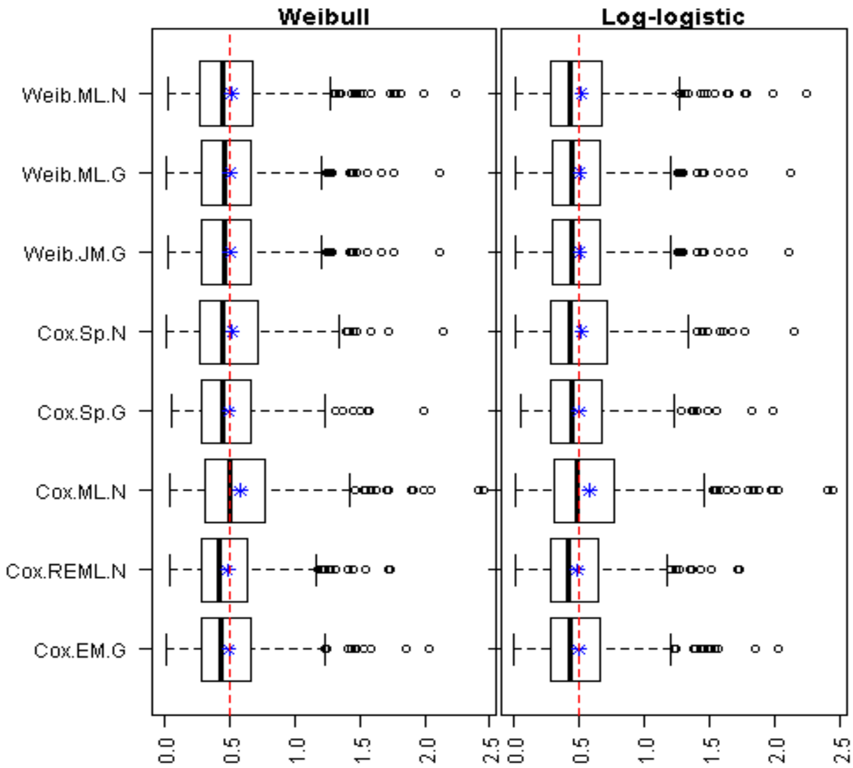
Moderate censoring (45%) **when** $N = 15$



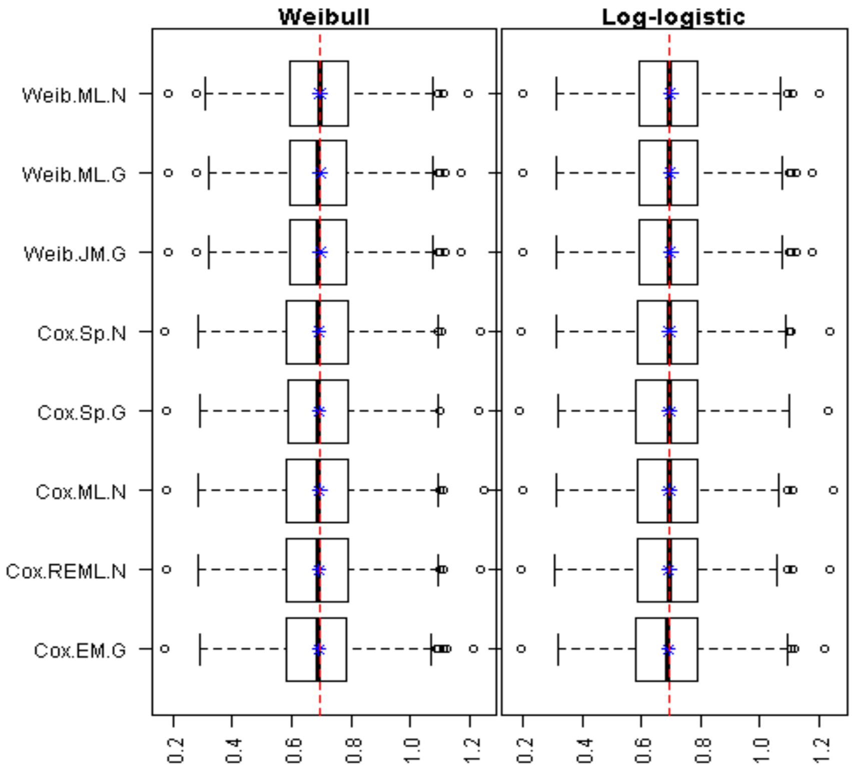
Web-Figure 13.112: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



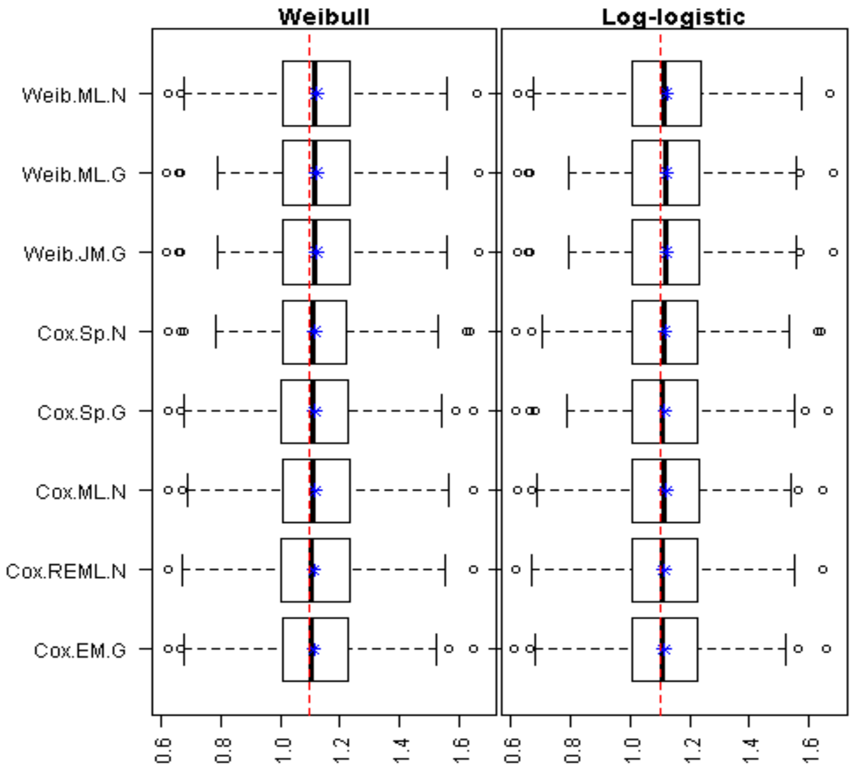
Web-Figure 13.113: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



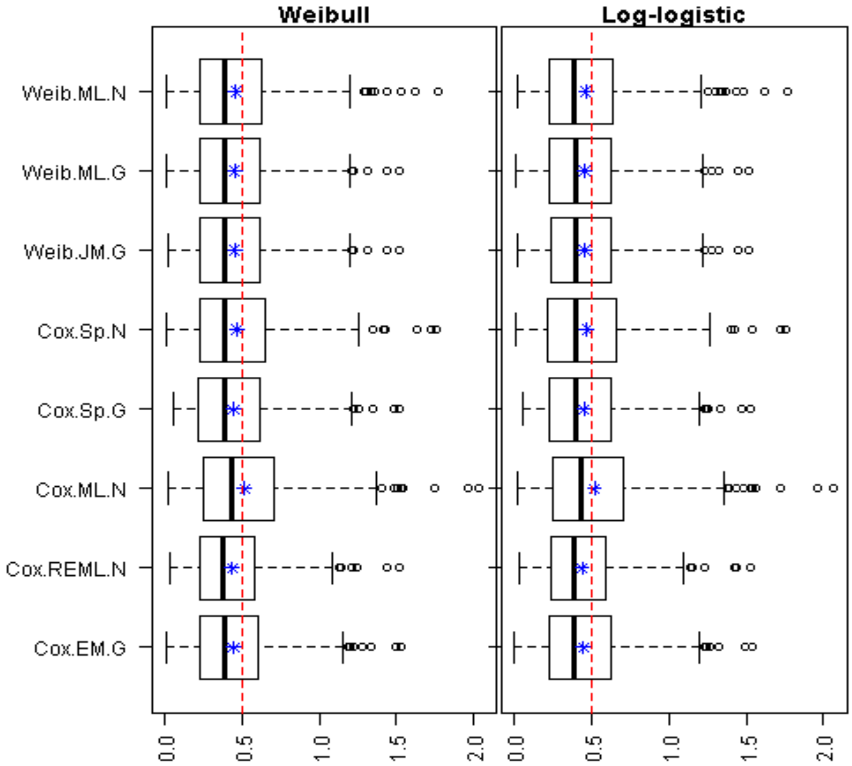
Web-Figure 13.114: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 0.5$.



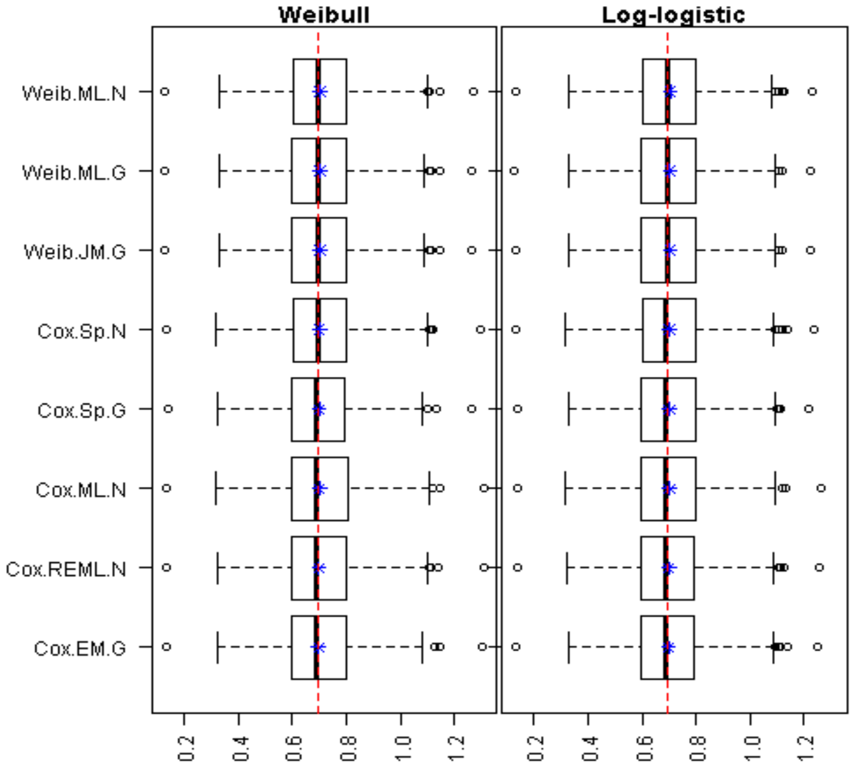
Web-Figure 13.115: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



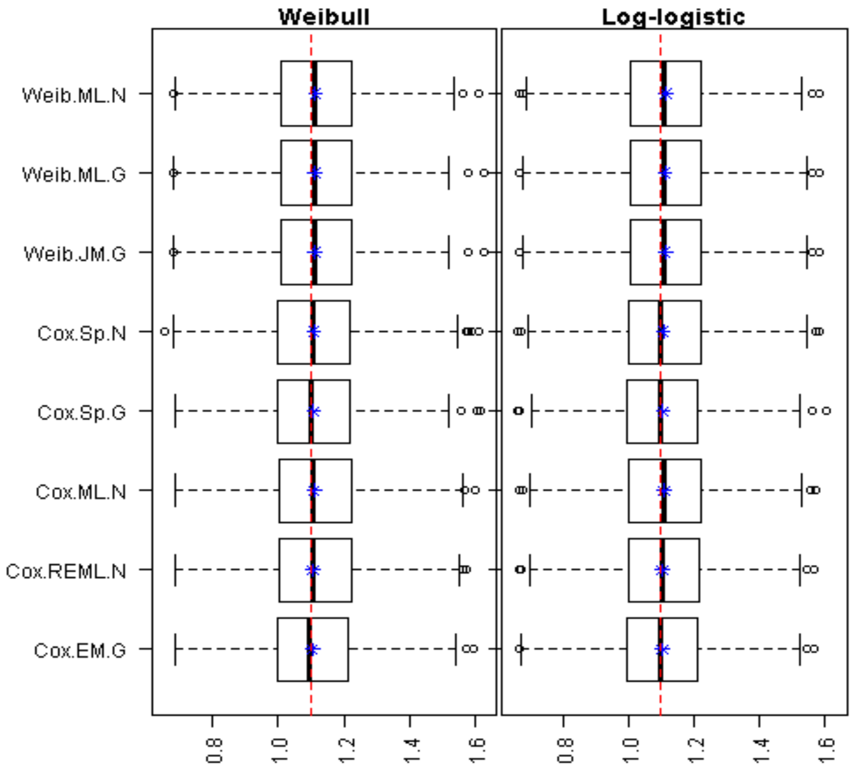
Web-Figure 13.116: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



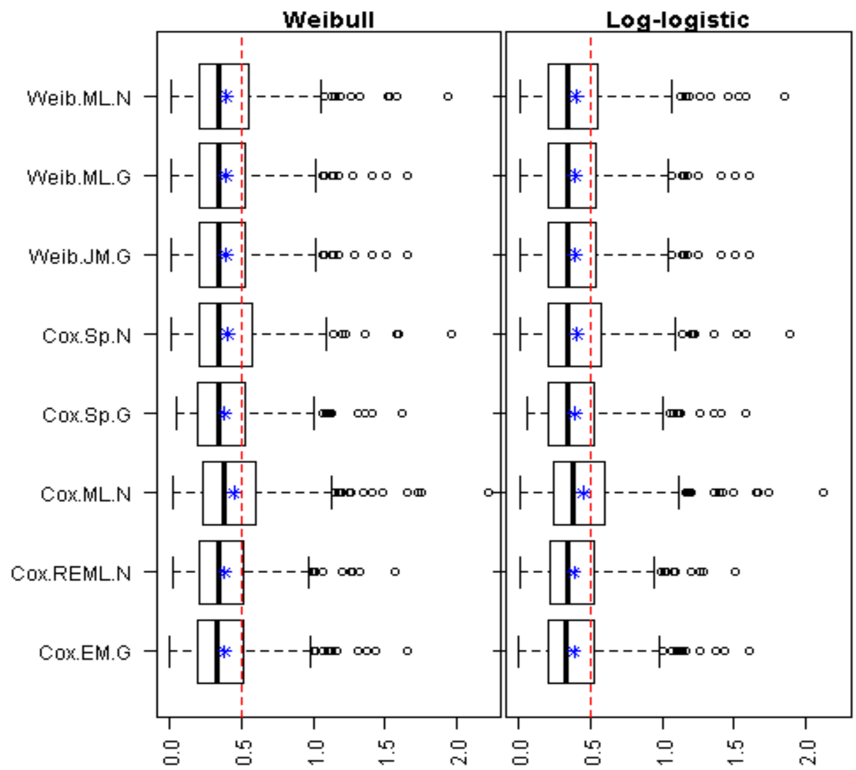
Web-Figure 13.117: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 0.5$.



Web-Figure 13.118: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.



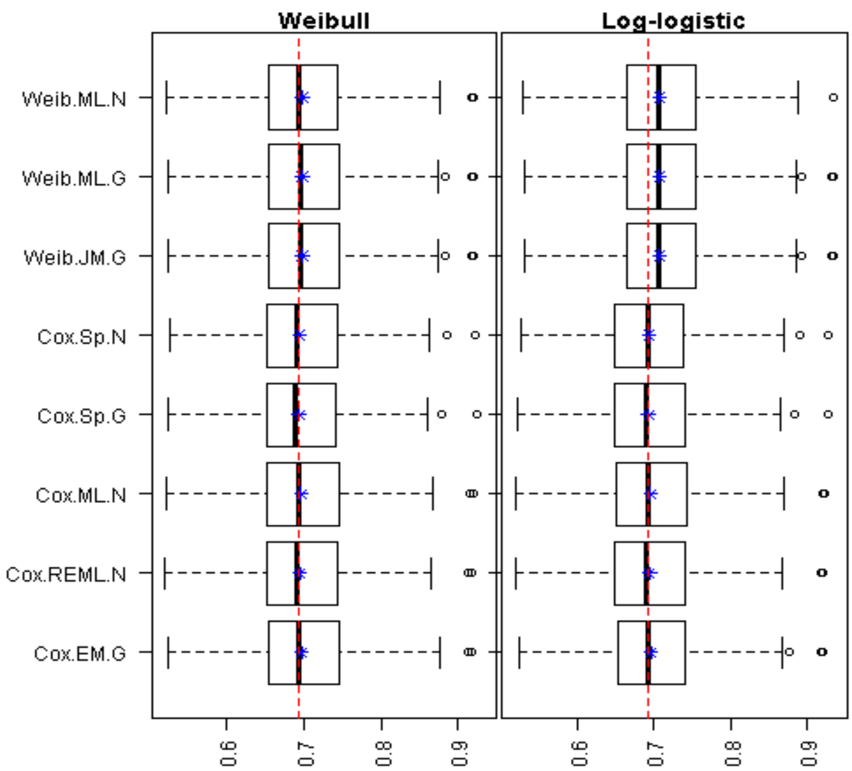
Web-Figure 13.119: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.



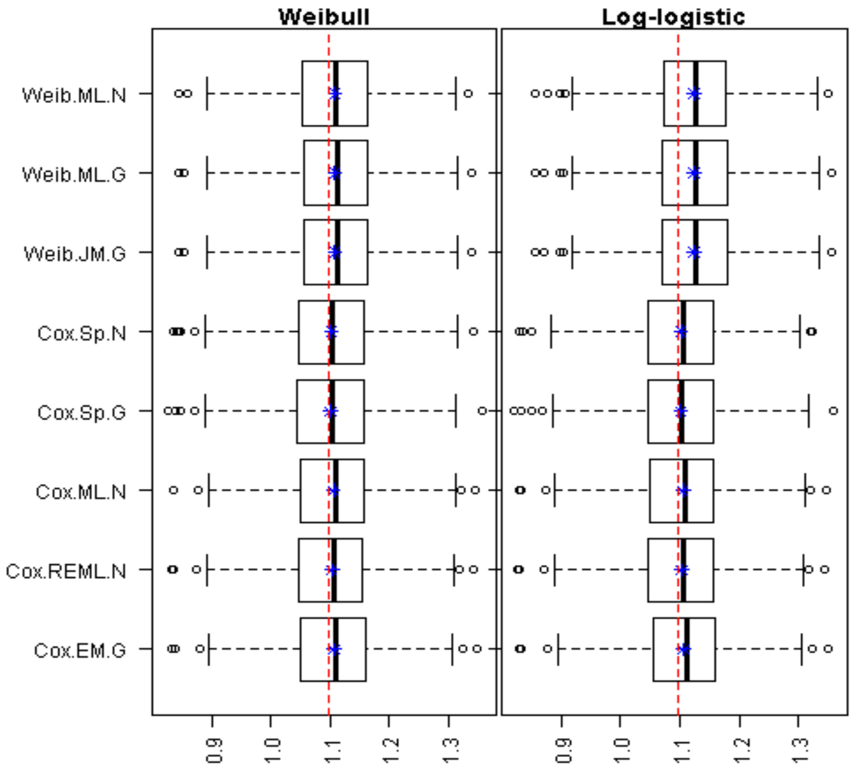
Web-Figure 13.120: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 0.5$.

Severe censoring (85%) **when** $N = 15$

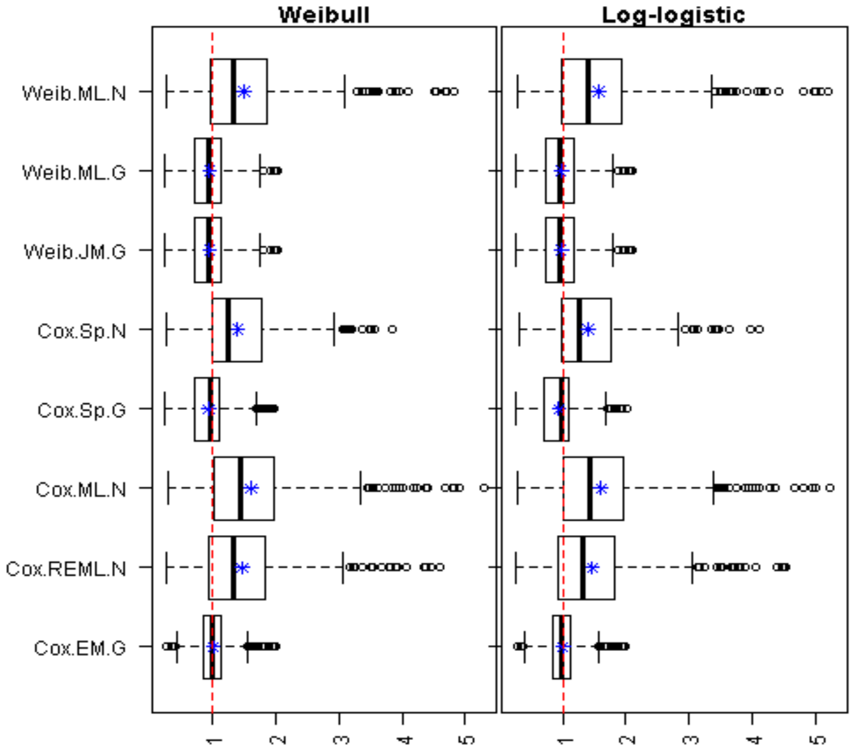
13.2.3 Heterogeneity Parameter, $\theta = 1$



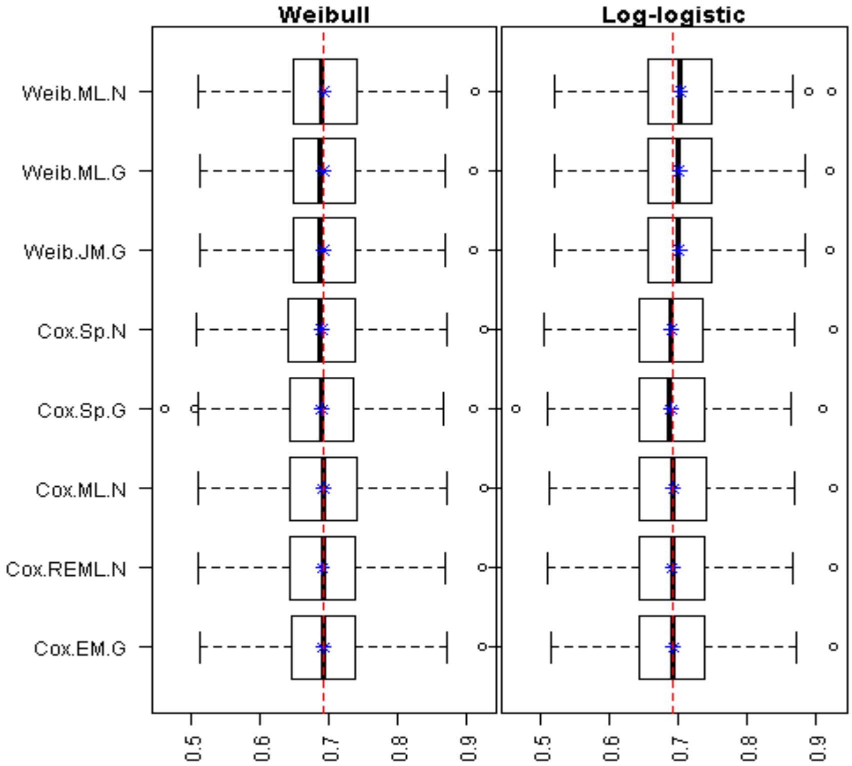
Web-Figure 13.121: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 1$.



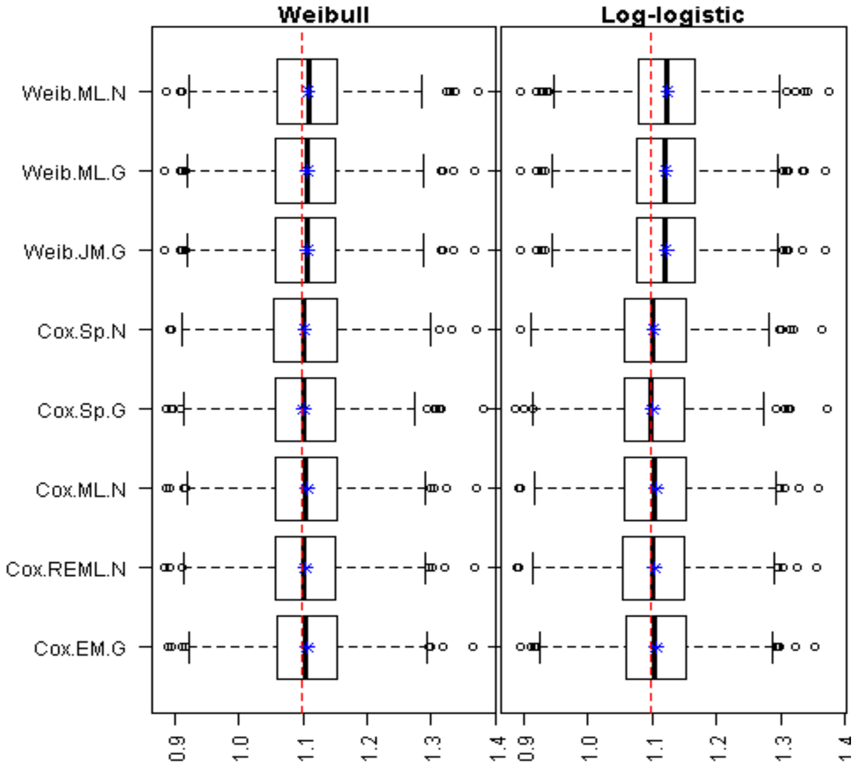
Web-Figure 13.122: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 1$.



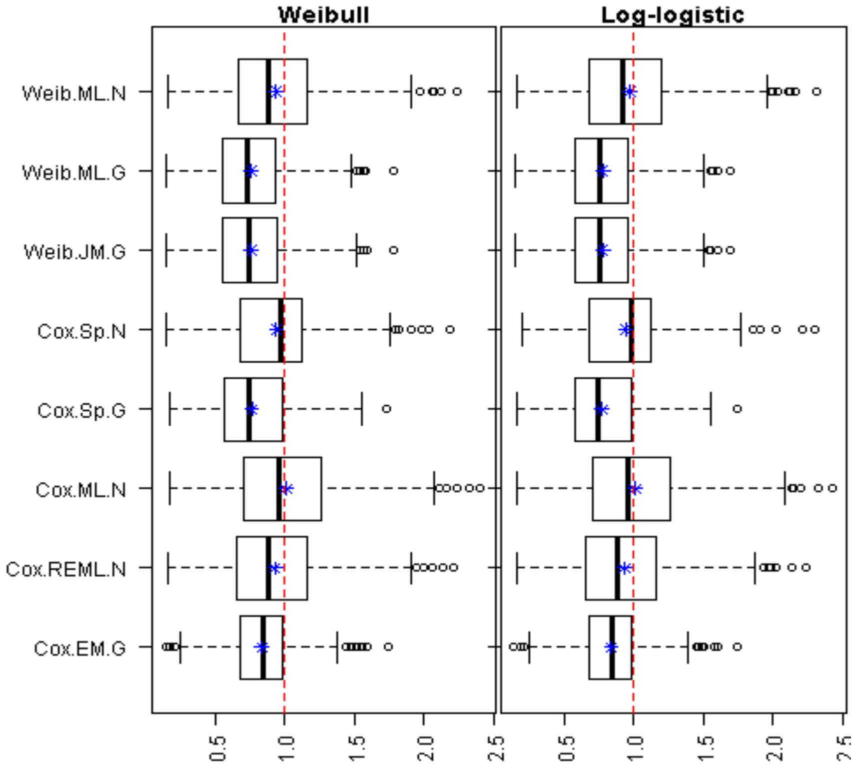
Web-Figure 13.123: θ estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 1$.



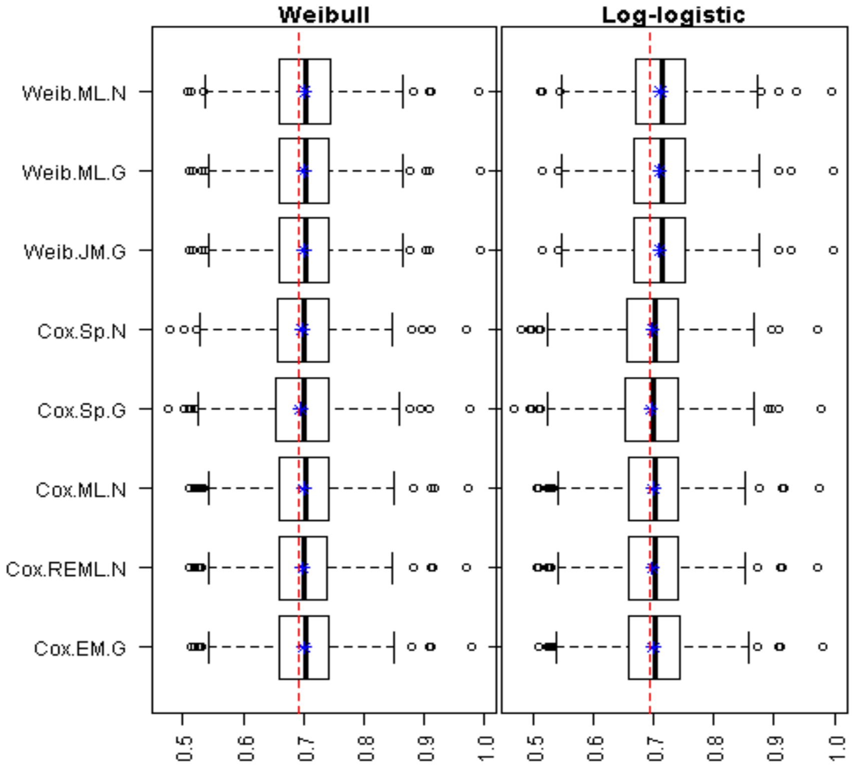
Web-Figure 13.124: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 1$.



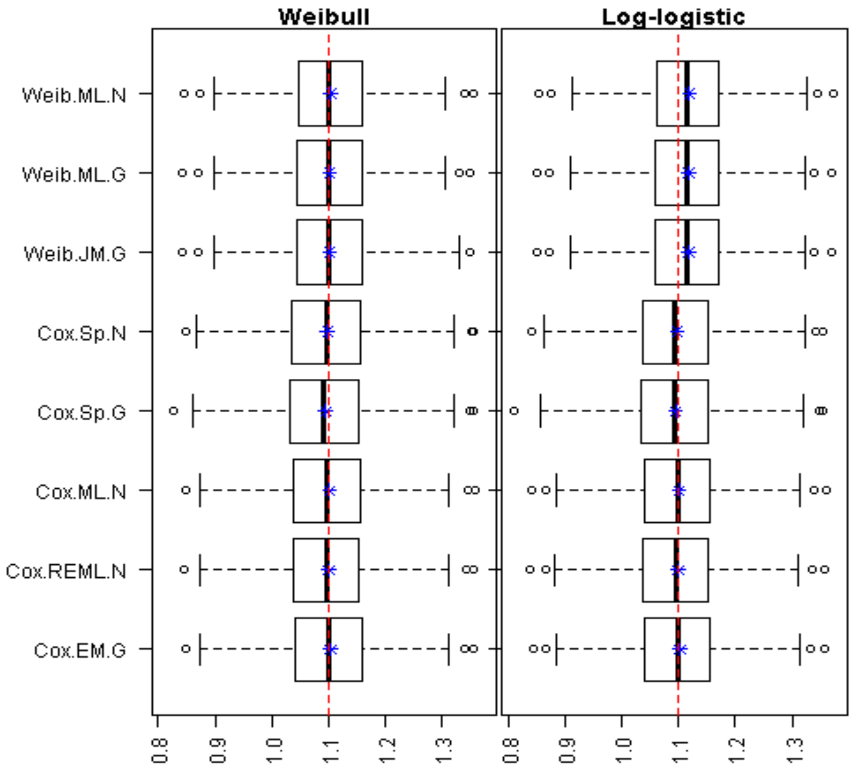
Web-Figure 13.125: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 1$.



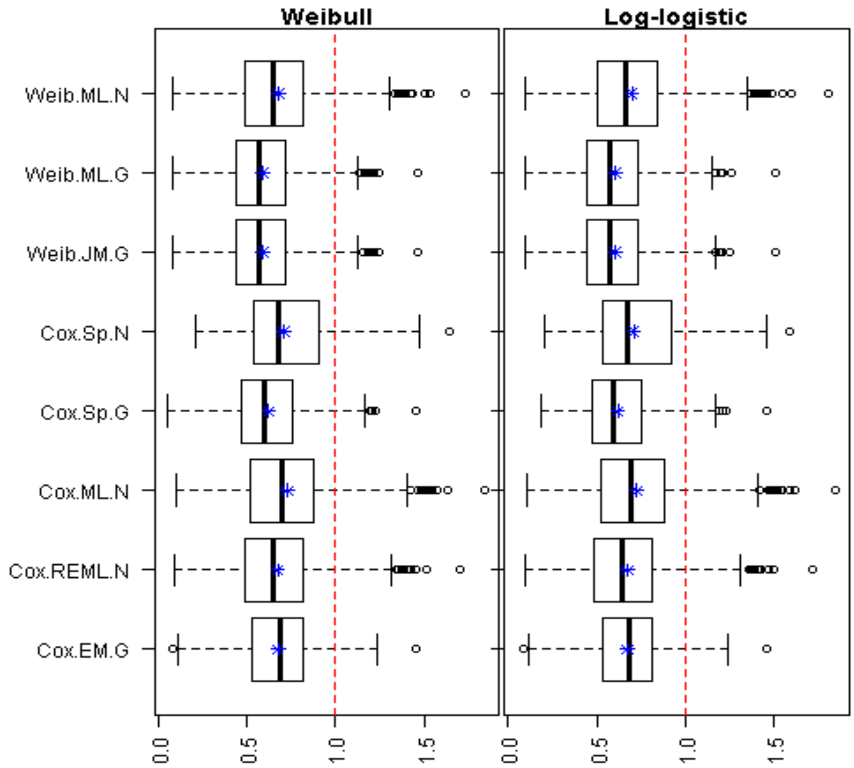
Web-Figure 13.126: θ estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 1$.



Web-Figure 13.127: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

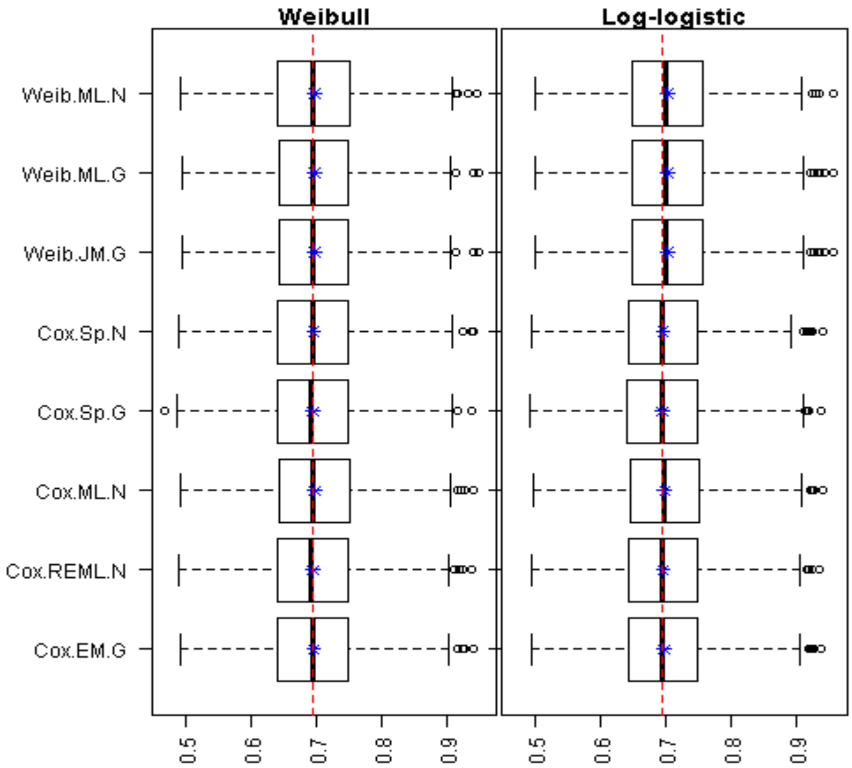


Web-Figure 13.128: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

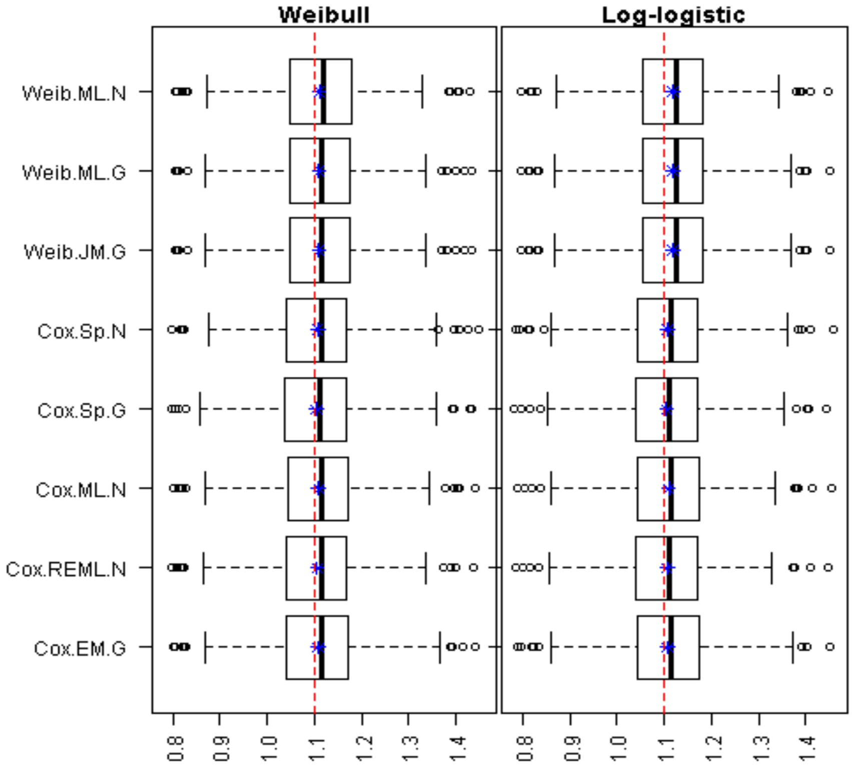


Web-Figure 13.129: θ estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

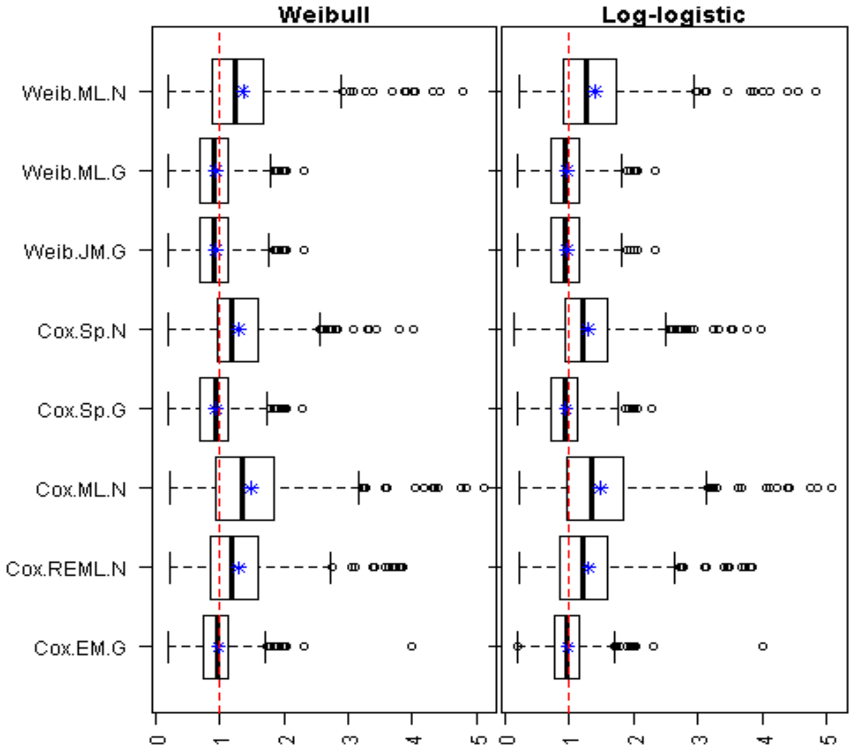
Mild censoring (15%) **when** $N = 15$



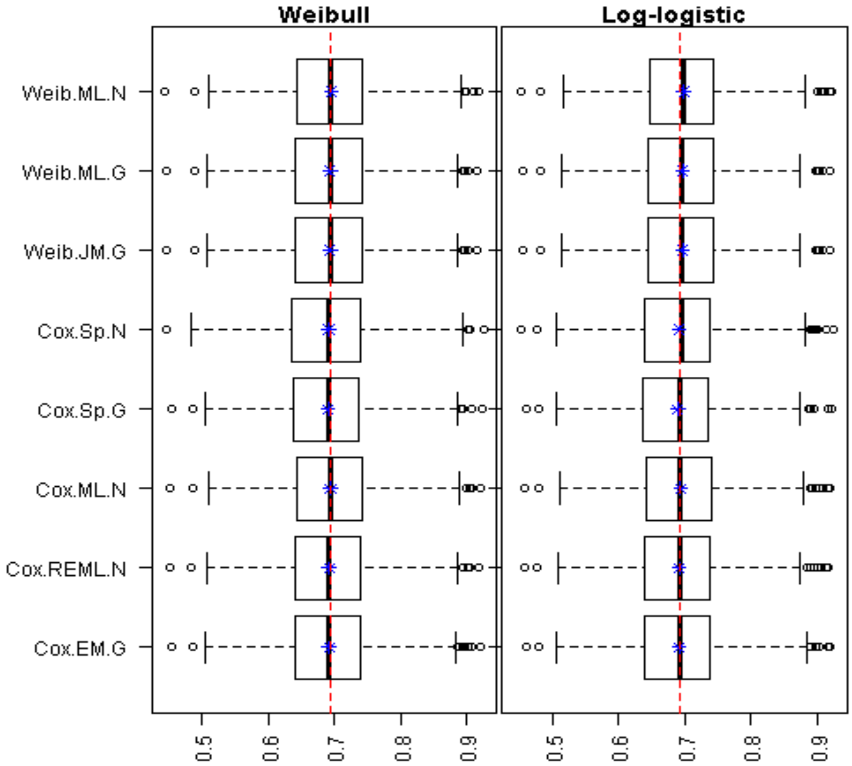
Web-Figure 13.130: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



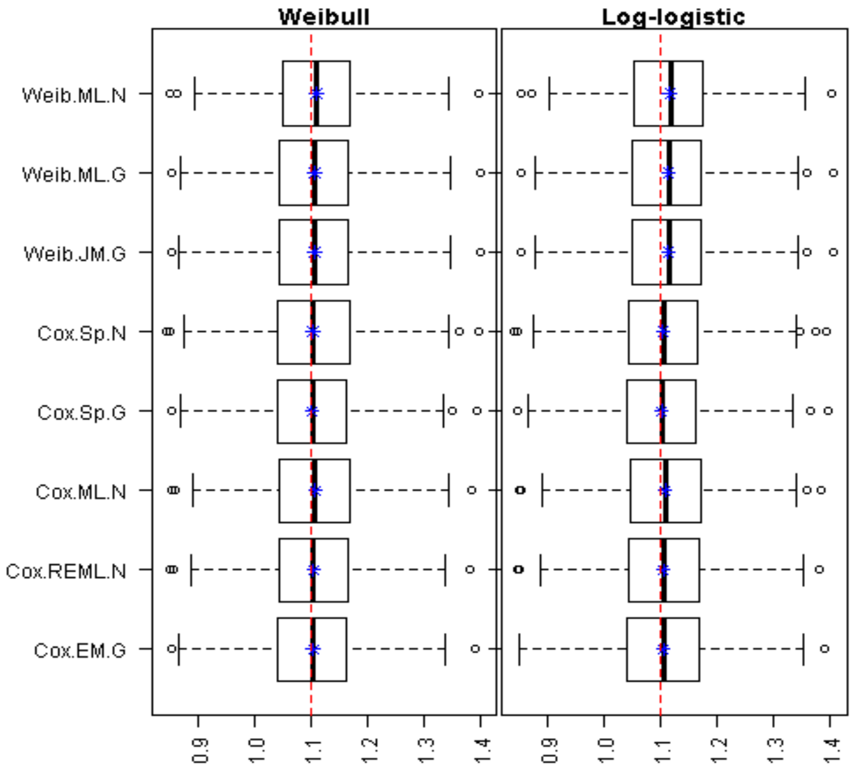
Web-Figure 13.131: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



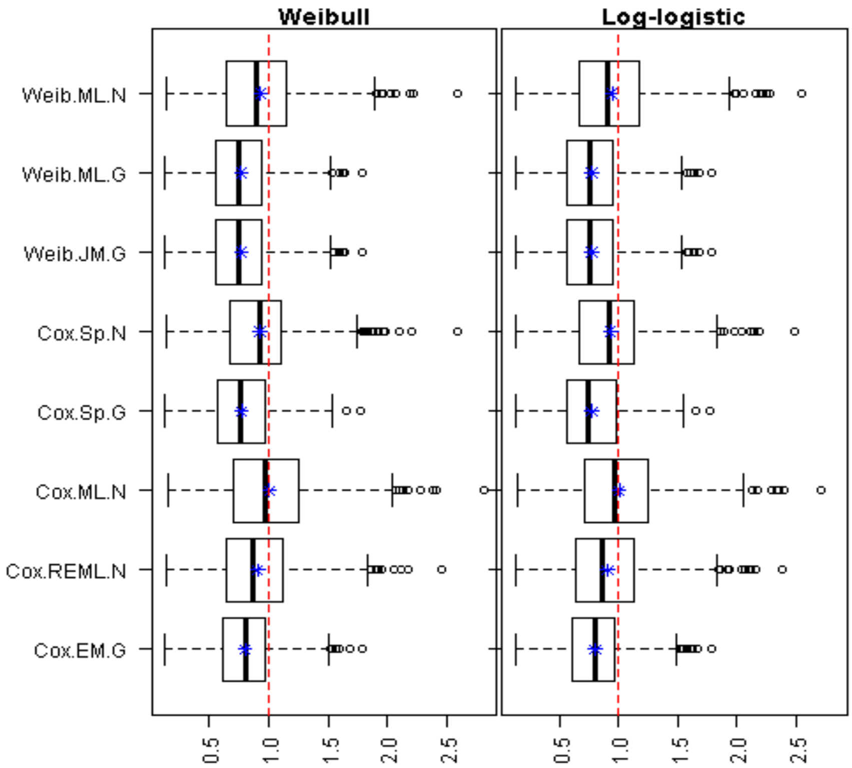
Web-Figure 13.132: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



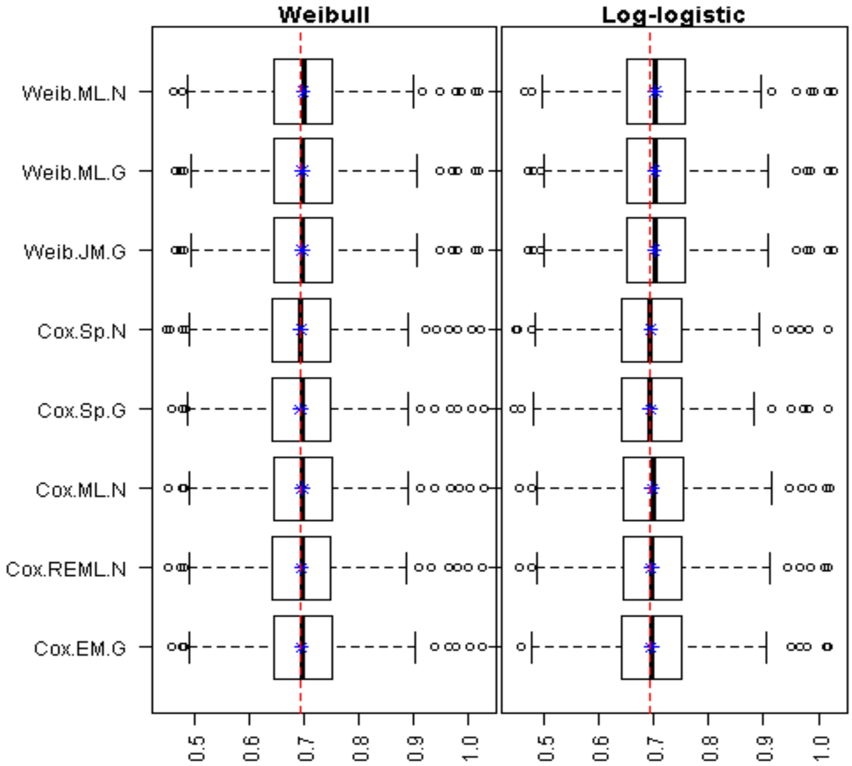
Web-Figure 13.133: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



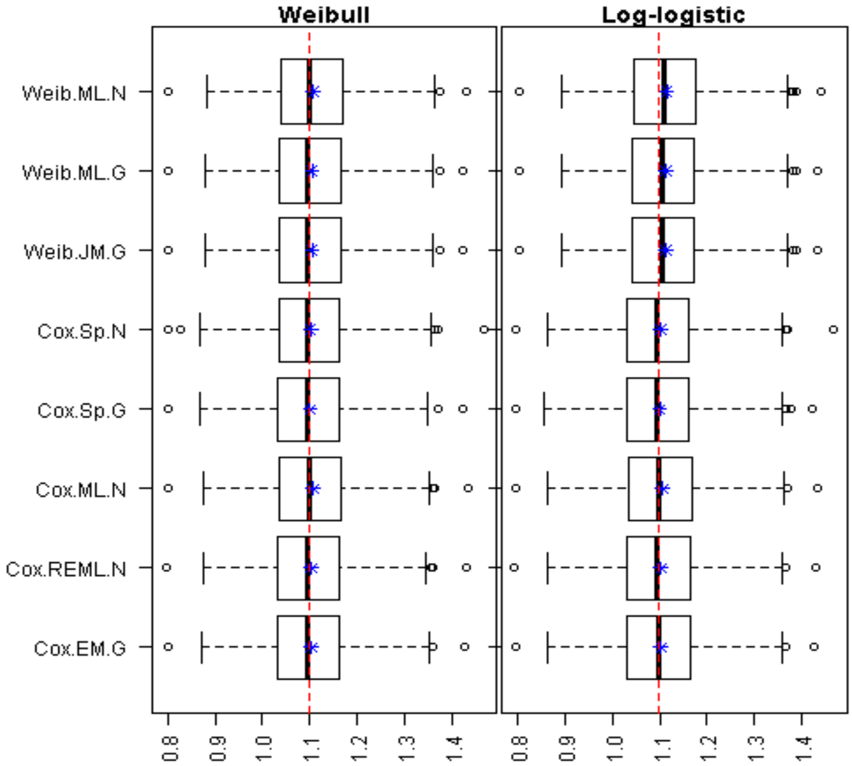
Web-Figure 13.134: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



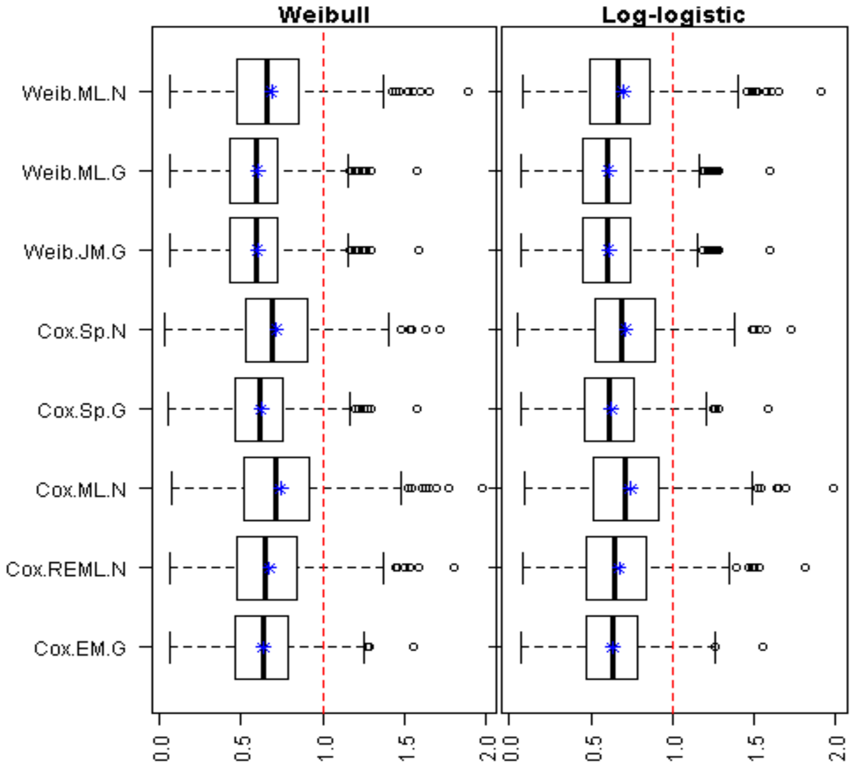
Web-Figure 13.135: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



Web-Figure 13.136: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

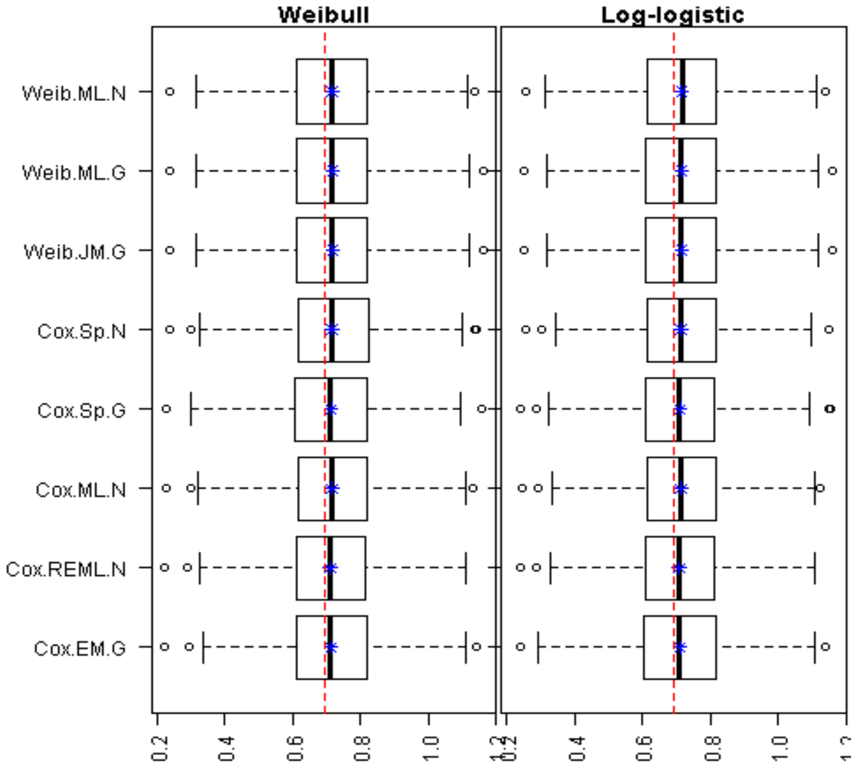


Web-Figure 13.137: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

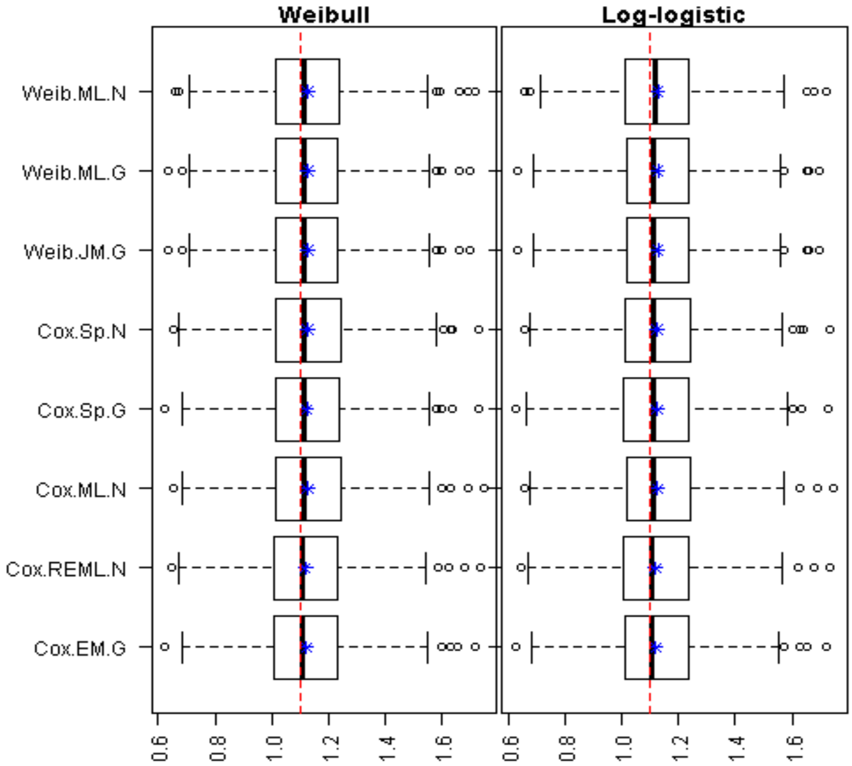


Web-Figure 13.138: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

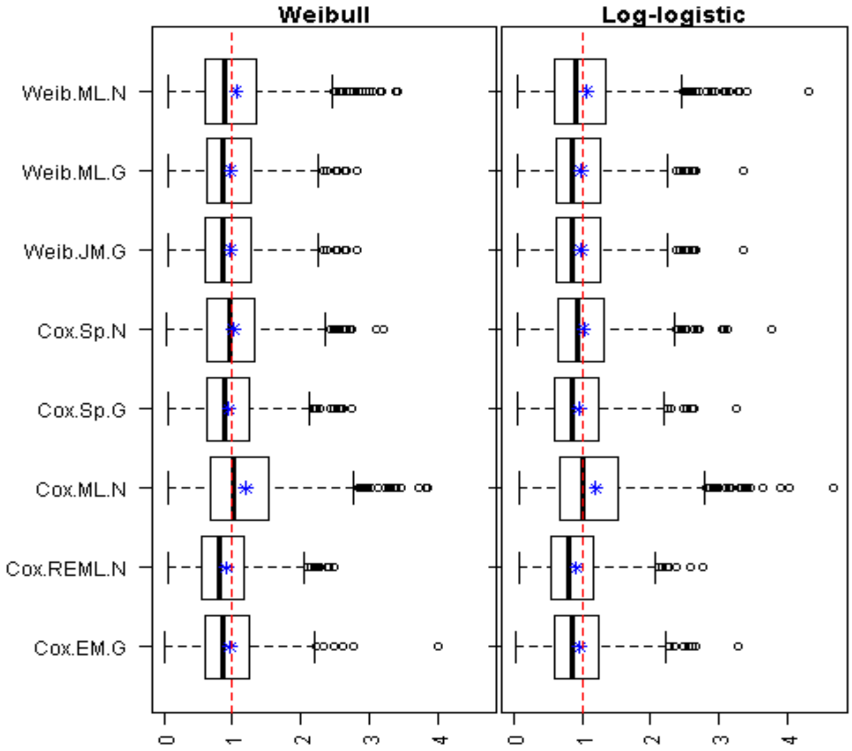
Moderate censoring (45%) **when** $N = 15$



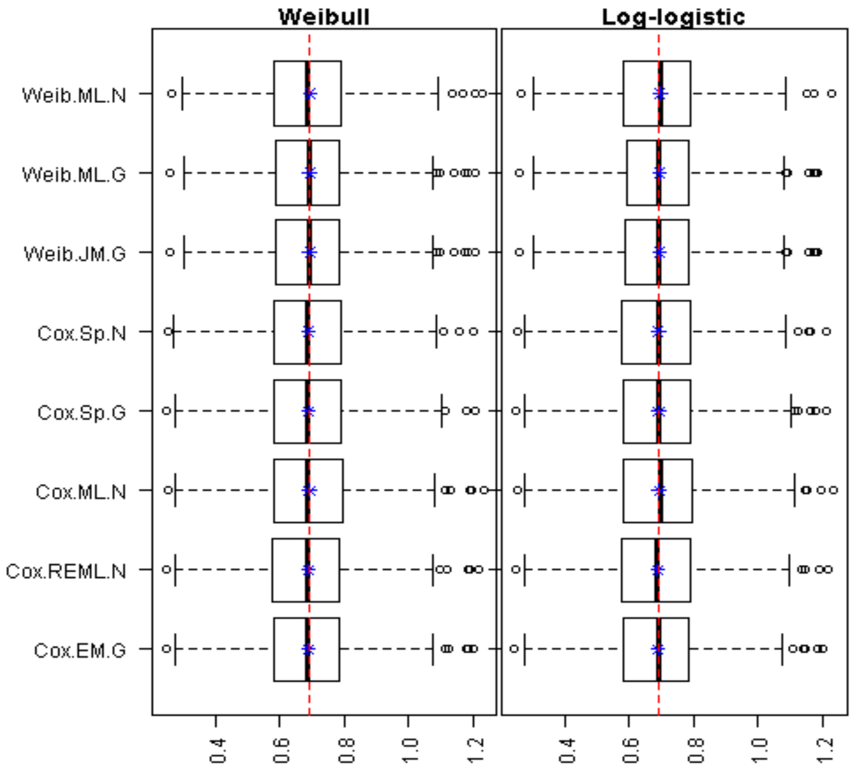
Web-Figure 13.139: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



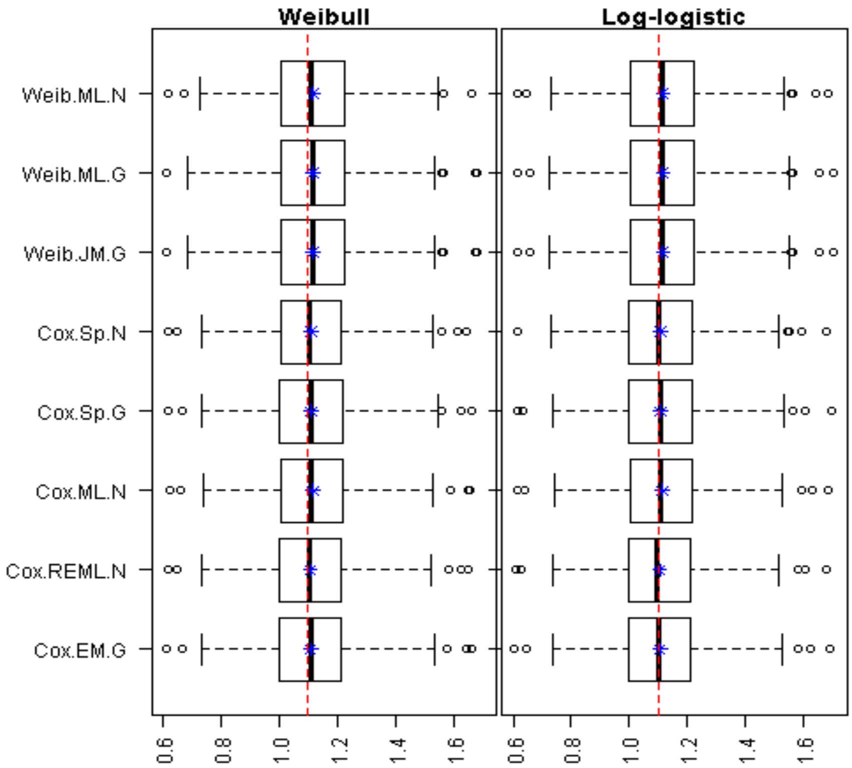
Web-Figure 13.140: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



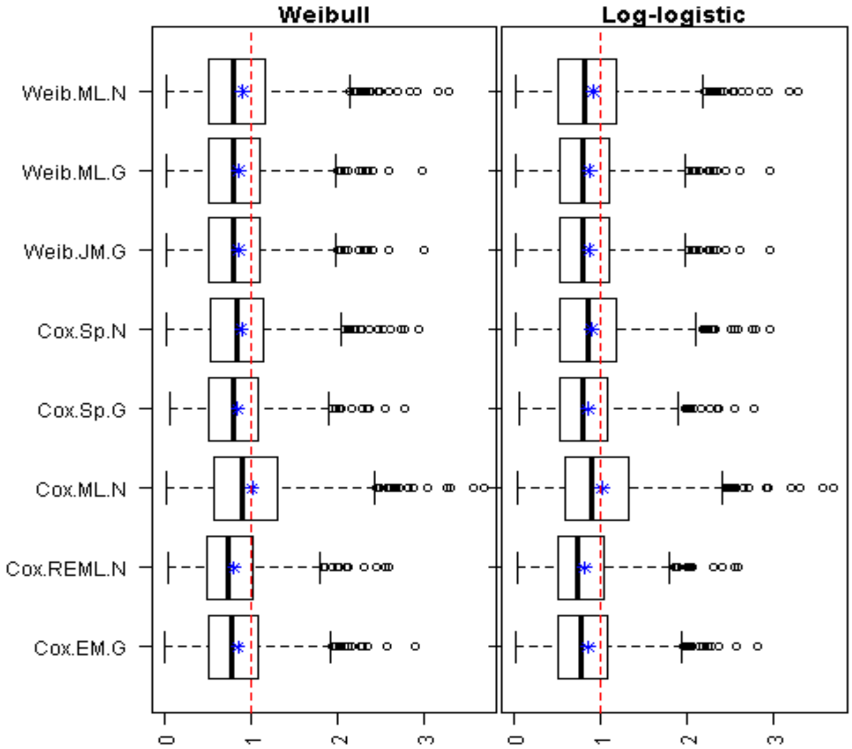
Web-Figure 13.141: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a gamma distribution and $\theta = 1$.



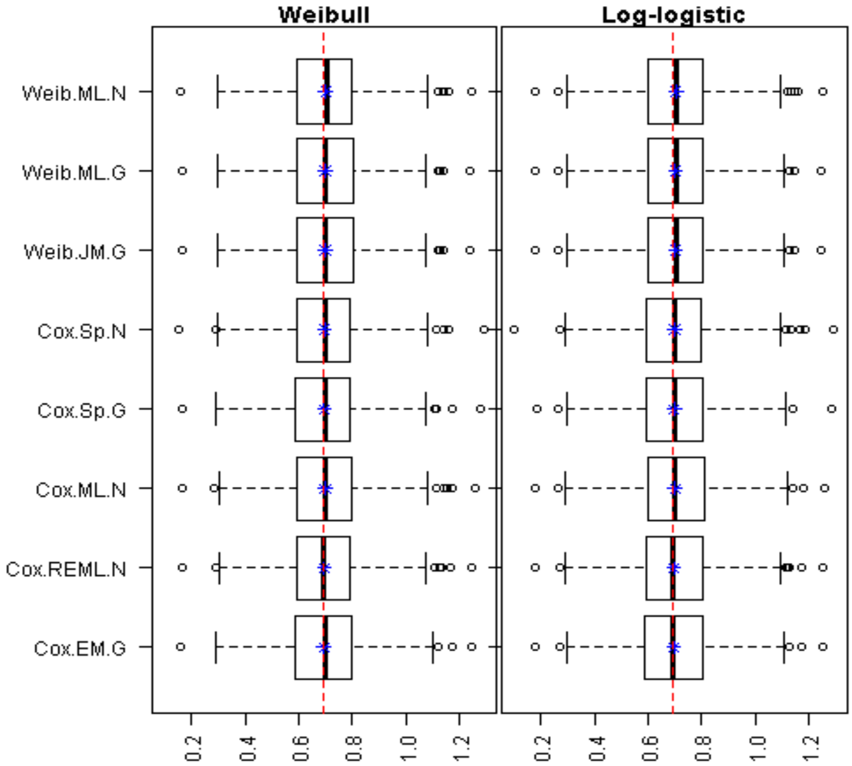
Web-Figure 13.142: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



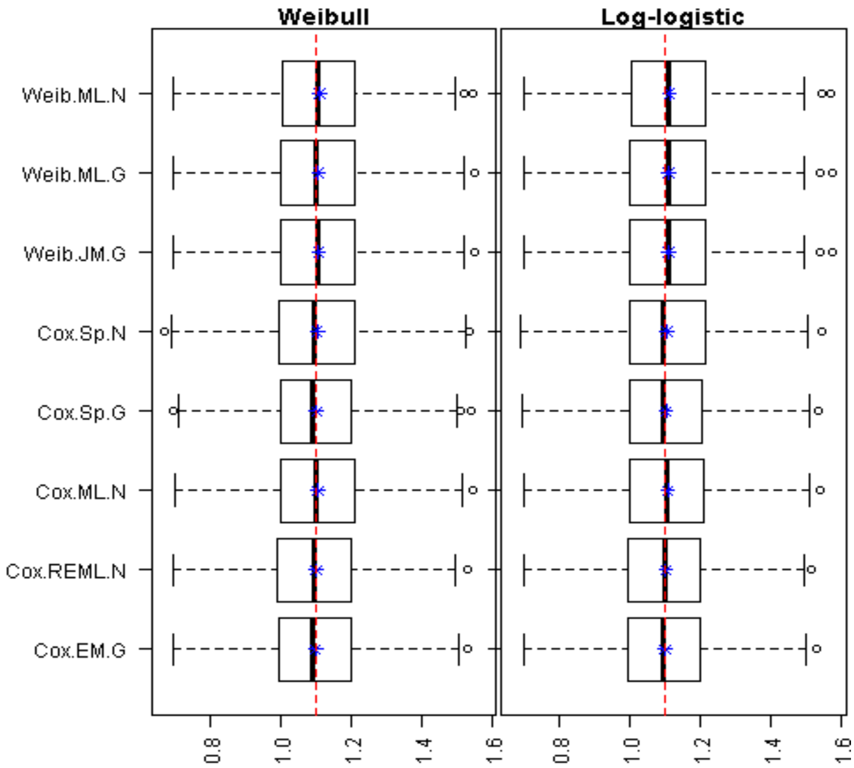
Web-Figure 13.143: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



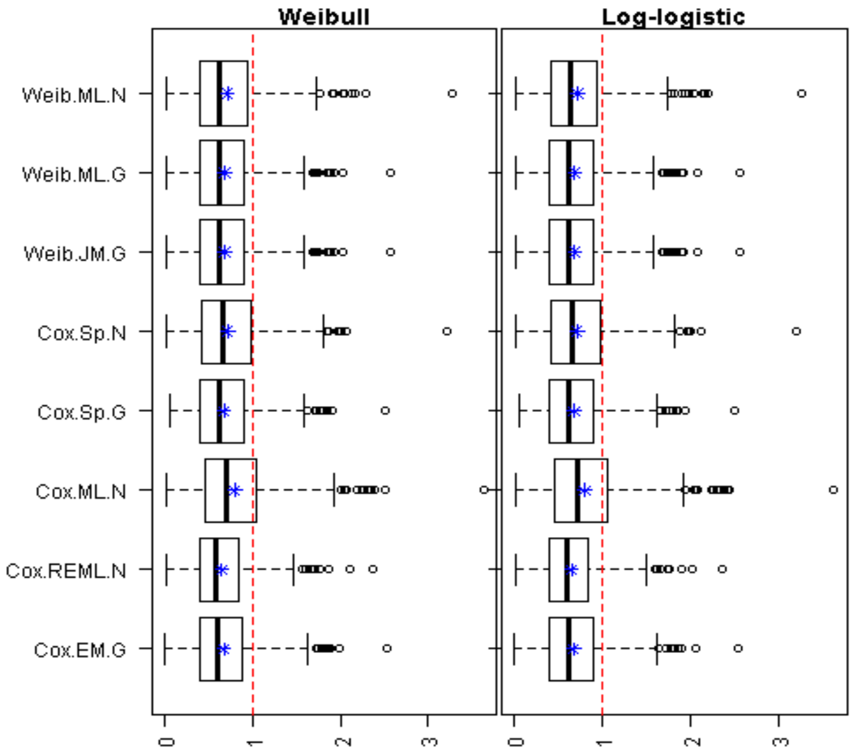
Web-Figure 13.144: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from a log-normal distribution and $\theta = 1$.



Web-Figure 13.145: β_1 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.



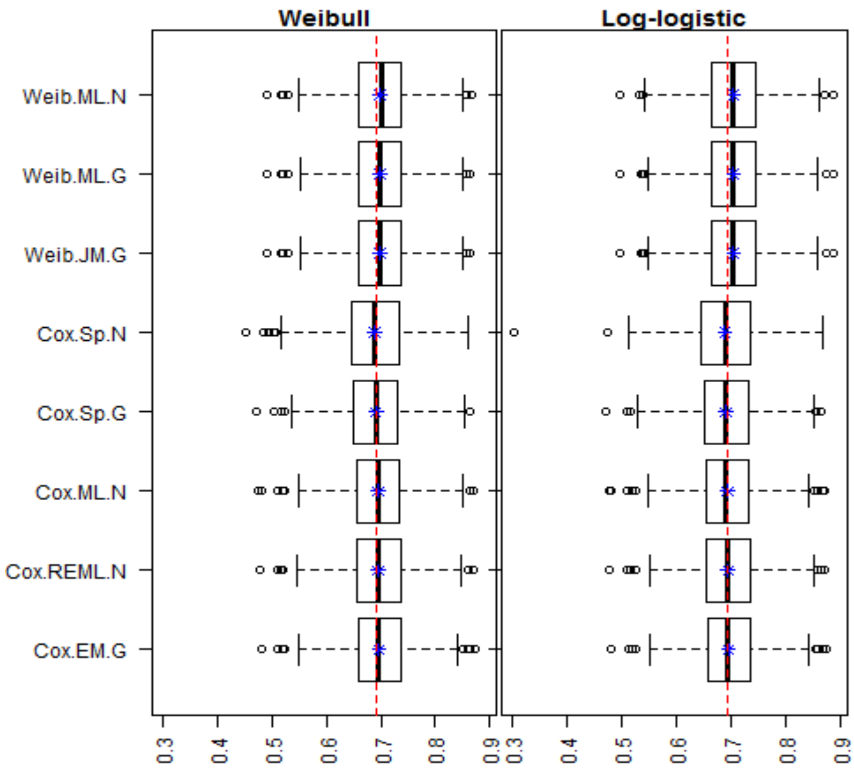
Web-Figure 13.146: β_2 estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.



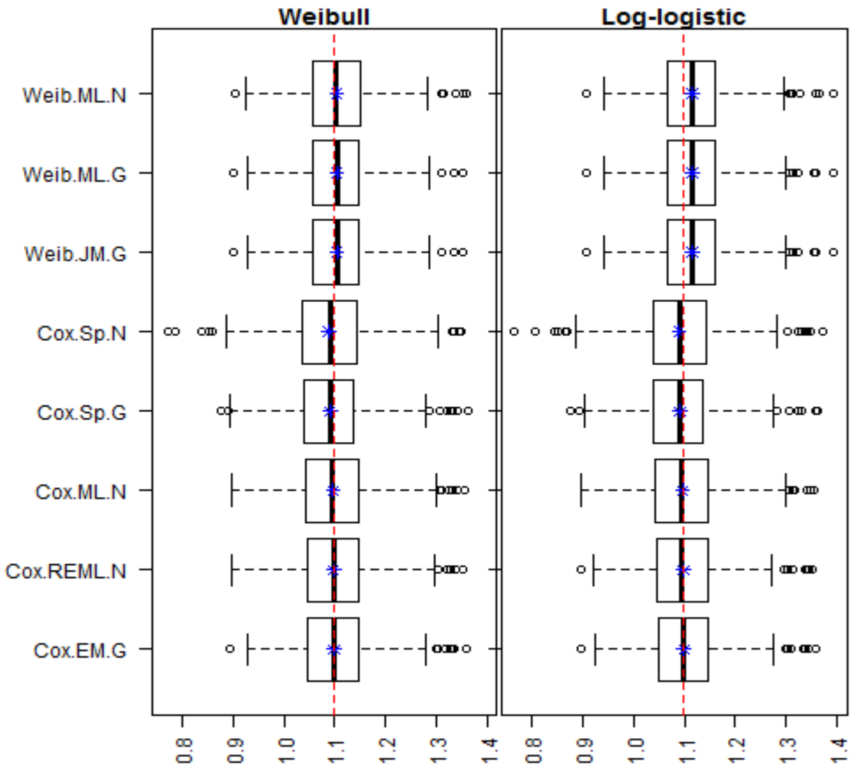
Web-Figure 13.147: θ estimates when small number of clusters were used under severe censoring censoring, frailty generated from an inverse Gaussian distribution and $\theta = 1$.

Severe censoring (85%) **when** $N = 15$

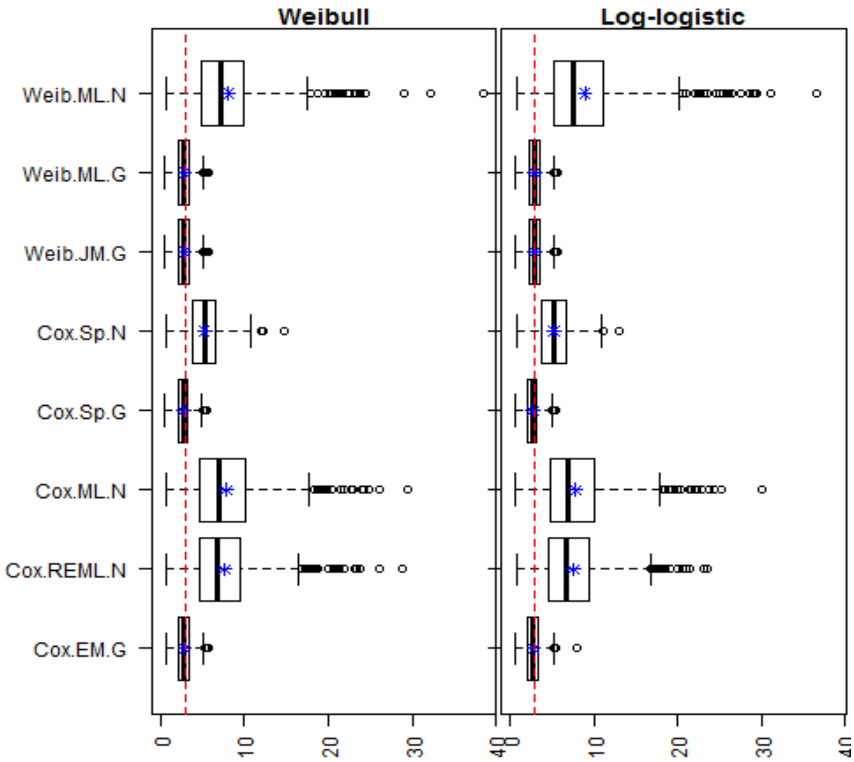
13.2.4 Heterogeneity Parameter, $\theta = 3$



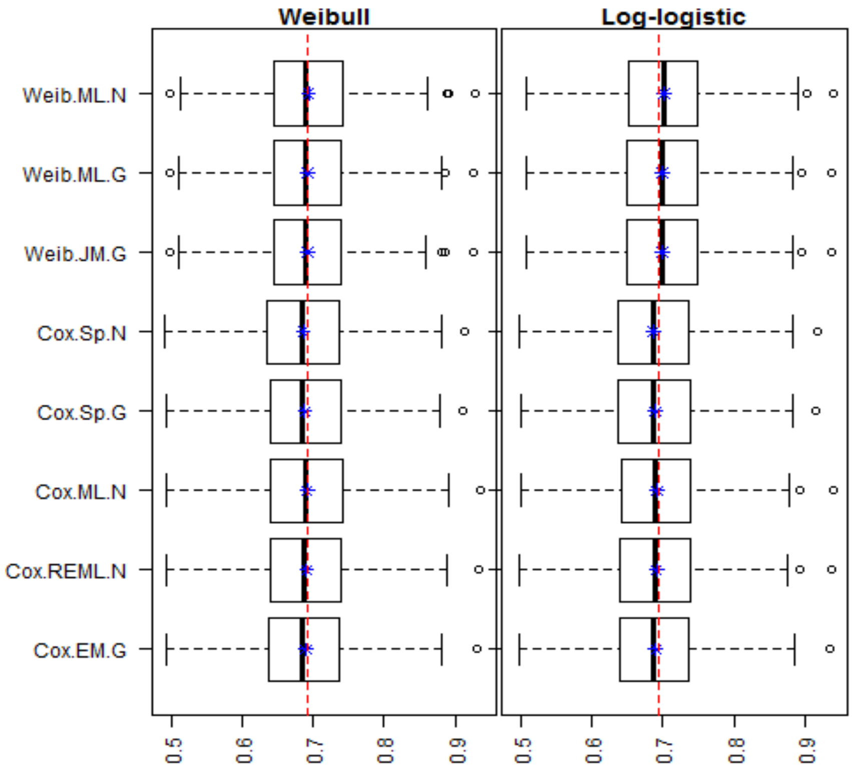
Web-Figure 13.148: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 3$.



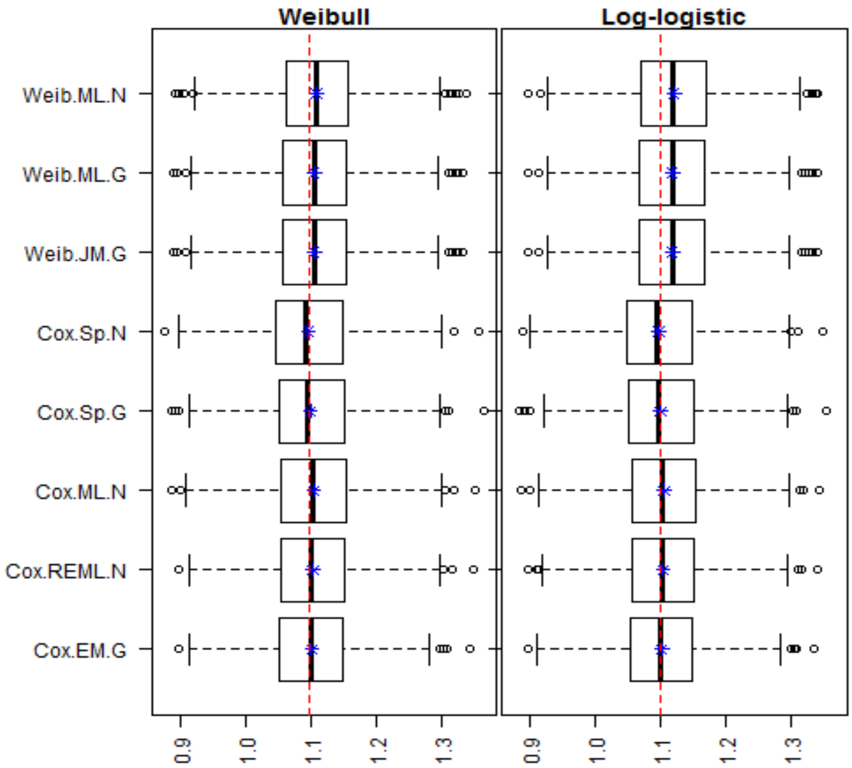
Web-Figure 13.149: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 3$.



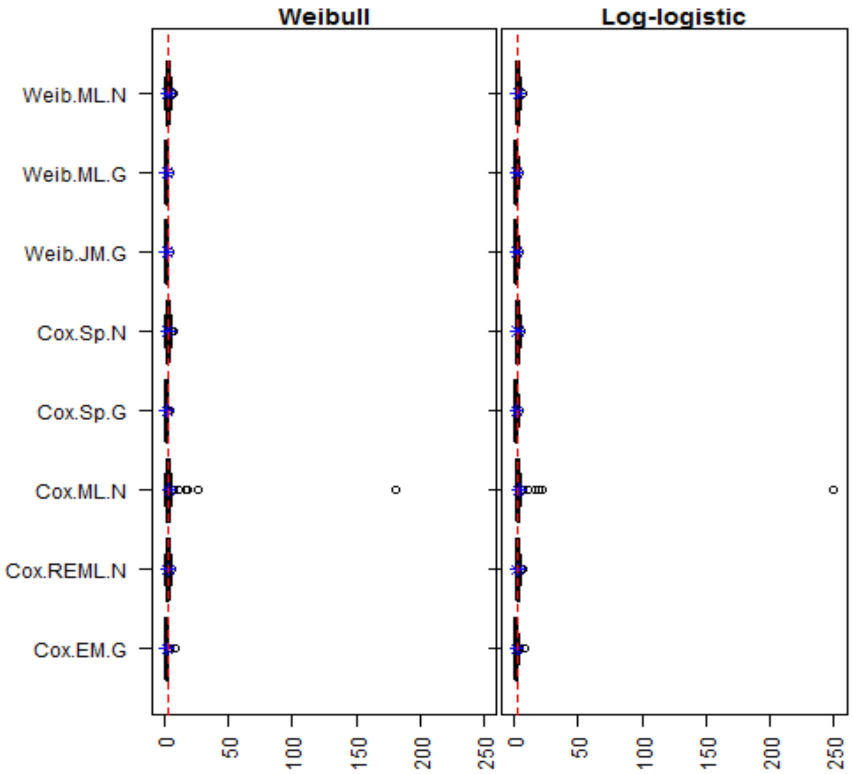
Web-Figure 13.150: θ estimates when small number of clusters were used under mild censoring, frailty generated from a gamma distribution and $\theta = 3$.



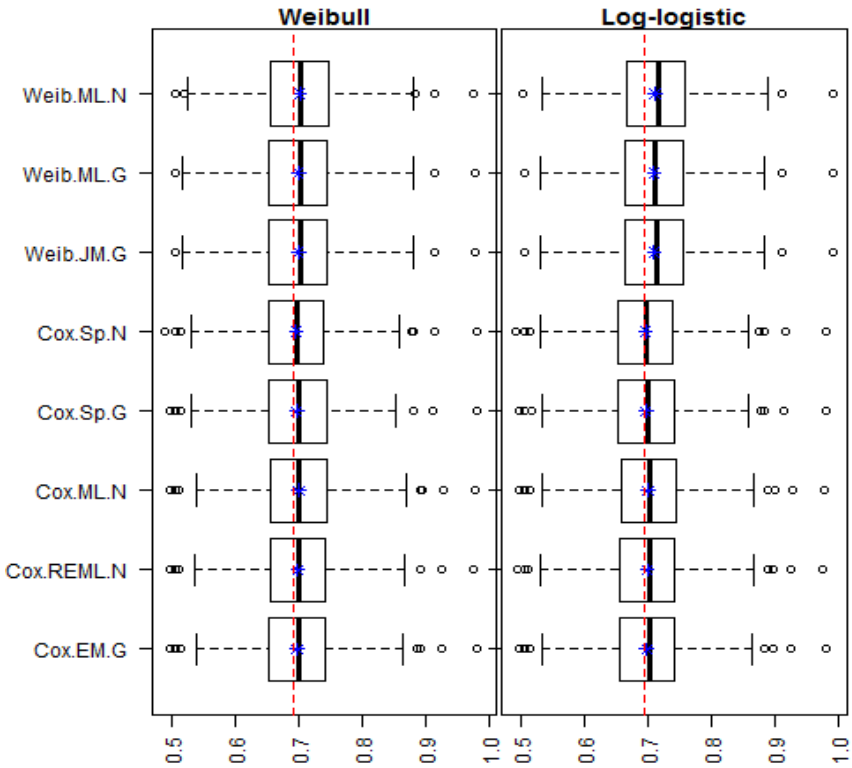
Web-Figure 13.151: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 3$.



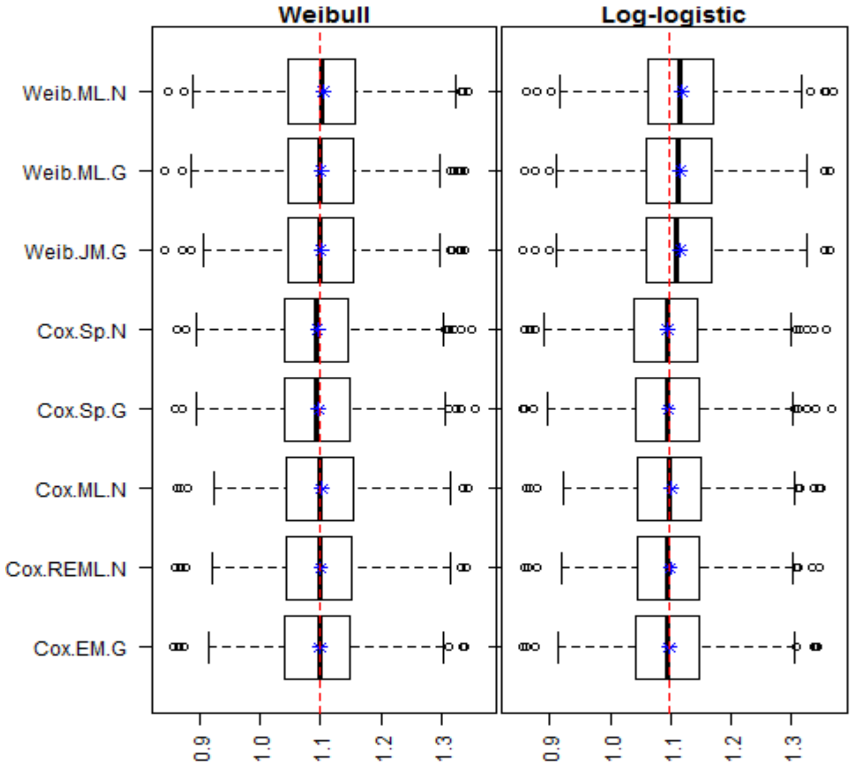
Web-Figure 13.152: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 3$.



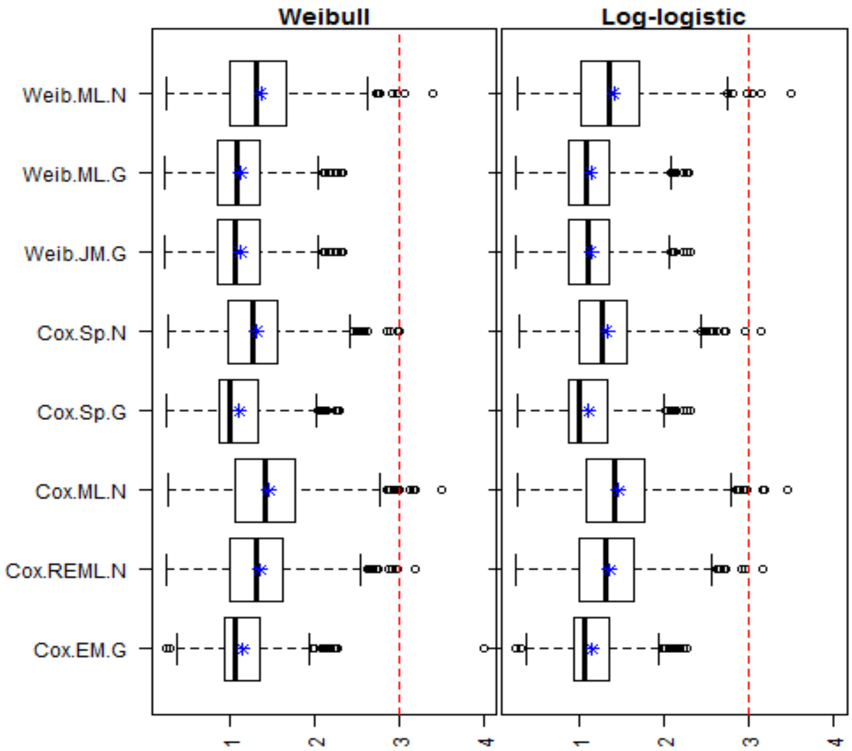
Web-Figure 13.153: θ estimates when small number of clusters were used under mild censoring, frailty generated from a log-normal distribution and $\theta = 3$.



Web-Figure 13.154: β_1 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 3$.

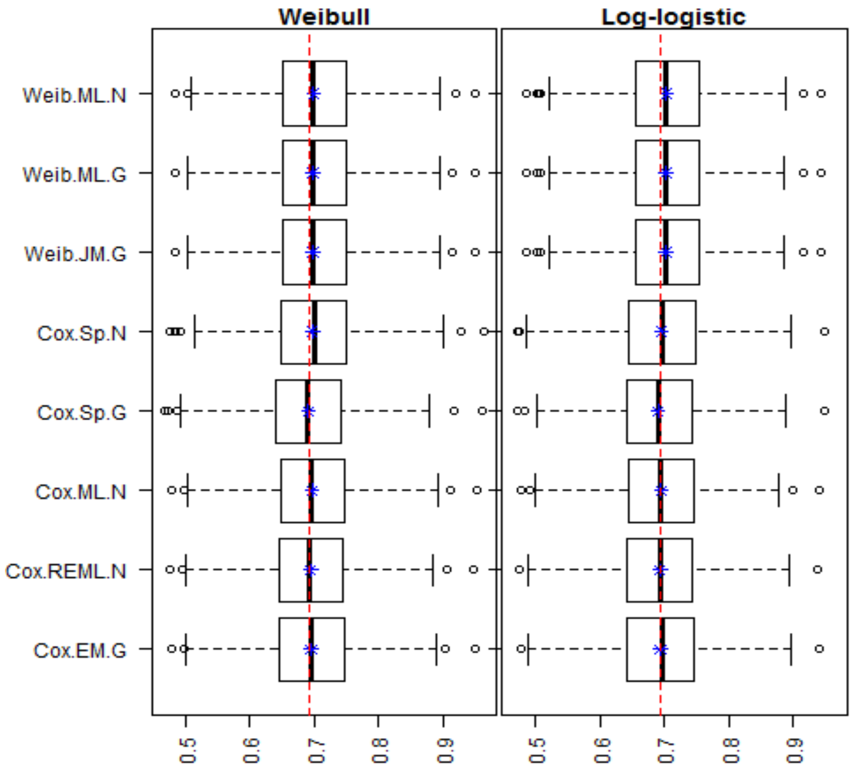


Web-Figure 13.155: β_2 estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 3$.

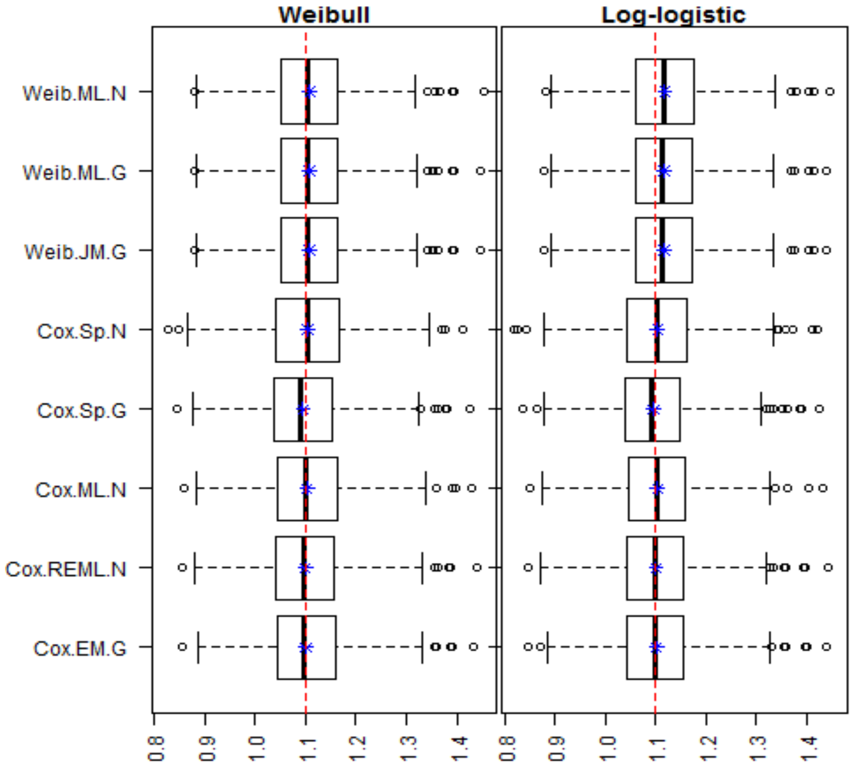


Web-Figure 13.156: θ estimates when small number of clusters were used under mild censoring, frailty generated from an inverse Gaussian distribution and $\theta = 3$.

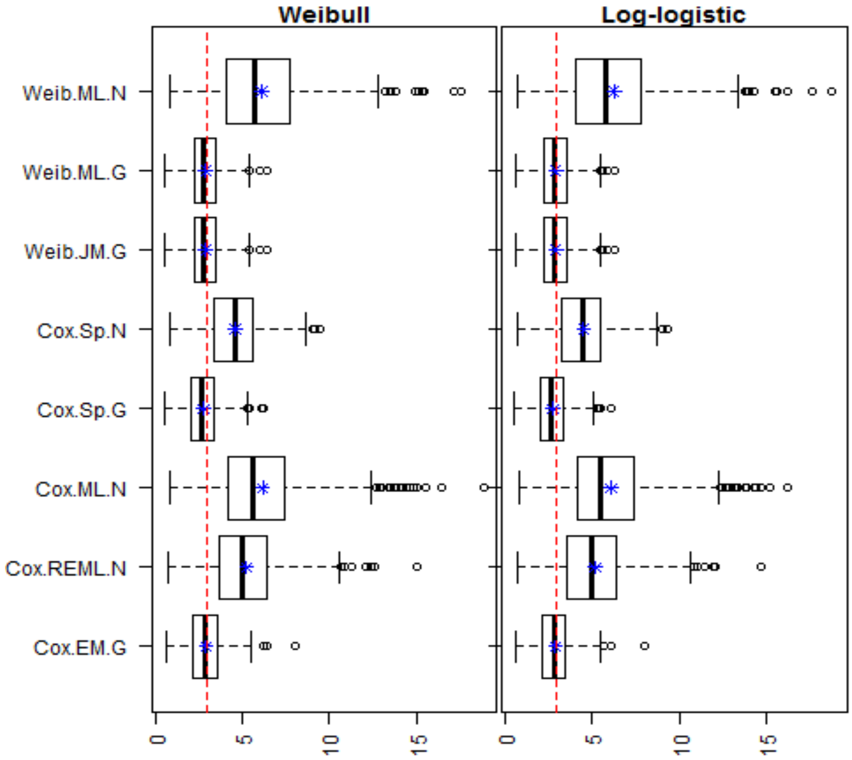
Mild censoring (15%) **when** $N = 15$



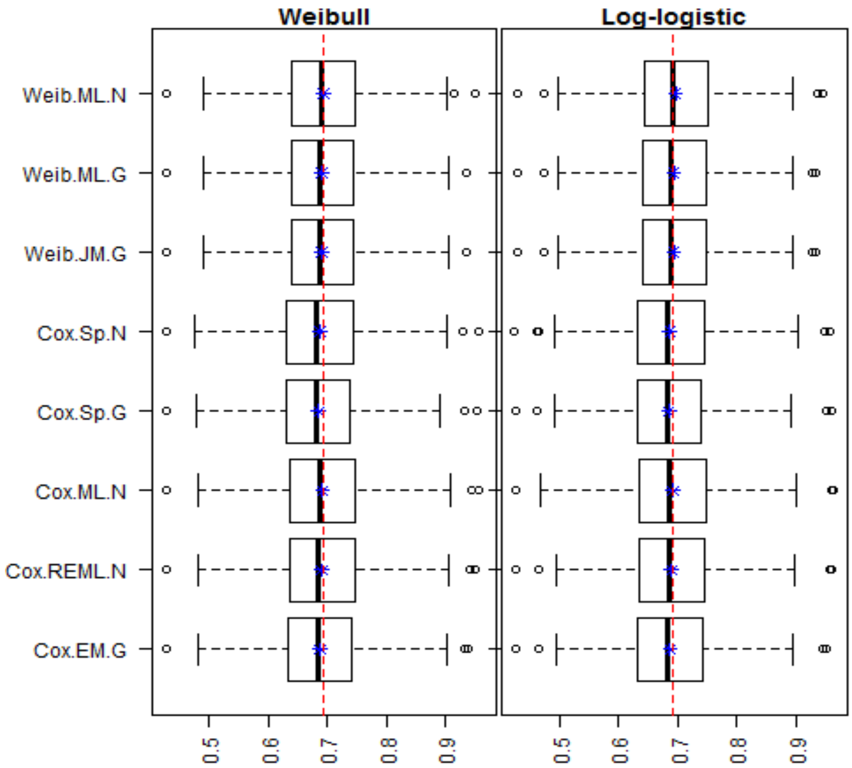
Web-Figure 13.157: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 3$.



Web-Figure 13.158: β_2 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 3$.



Web-Figure 13.159: θ estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a gamma distribution and $\theta = 3$.



Web-Figure 13.160: β_1 estimates when small number of clusters were used under moderate censoring censoring, frailty generated from a log-normal distribution and $\theta = 3$.