Formulas for the model built on catch declarations

1 Notations

We model observations at the level of catch declarations D_j as a sum of Y_{ij} lognormal observations each one realised at one fishing position x_{ij}.

We note :

- $\mathcal{P}_j$: the vector of all fishing positions related to the j^th declaration
- $j \in \{1, ..., n\}$ with n the number of declarations
- $i \in \{1, ..., m_j\}$ with m_j the number of fishing positions belonging to the j^{th} declaration.

$$D_j = \sum_{i \in \mathcal{P}_j} Y_{ij}$$

2 Alternative observation models accounting for reallocation

2.1 First assumption about convolution of lognormal random variables

We assume a sum of lognormal random variables is still lognormal and that $D_j \sim \mathcal{LN}(M(D_j), \Phi(D_j))$ with $M(.)$ being a function representing the central value of D_j and $\Phi(.)$ the variance function of D_j.

2.2 Reparameterization of the lognormal distribution

The standard lognormal distribution is written as :

$D \sim \mathcal{LN}(\rho; \sigma^2)$ with $D = e^{\rho + \sigma N}$ and $N \sim \mathcal{N}(0, 1)$

In this case, $E(D) = e^{\rho + \frac{\sigma^2}{2}}$ and $Var(D) = (e^{\sigma^2} - 1)e^{2\rho + \sigma^2}$.

We choose to slightly reparameterize the lognormal distribution so that $E(D)$ and $Var(D)$ have more simple expressions.

Let's define $\rho = ln(\mu) - \frac{sigma^2}{2}$, then :

- $D = \mu e^{\sigma N - \frac{sigma^2}{2}}$
- $E(D) = \mu$
- $Var(D) = \mu^2(e^{\sigma^2} - 1) \Leftrightarrow \sigma^2 = ln(\frac{Var(D)}{E(D)^2} + 1)$

Then, $D \sim \mathcal{LN}(M(D), \Phi(D))$ with $M(D) = E(D)$ and $\Phi(D) = ln(\frac{Var(D)}{E(D)^2} + 1)$.

2.3 D_j probability distribution and moments

We have to express the probability distribution of D_j and its moments as a function of Y_{ij} with $i \in \mathcal{P}_j$.

Let's assume $Y_{ij} = C_{ij}.Z_{ij}$ is a zero-inflated lognormal distribution with C_{ij} a binary random variable and Z_{ij} a lognormal random variable.

$$C_{ij} \sim \mathcal{B}(1 - p_{ij})$$

with $p_{ij} = exp(-e^{\xi}.S(x_{ij}))$ the probability to obtain a zero value

$$Z_{ij} \sim \mathcal{LN}(\frac{S(x_{ij})}{1 - p_{ij}}, \sigma^2)$$

Here, Y_{ij}, C_{ij} and Z_{ij} are observations of a latent field $S(x_{ij})$ at a sampled point x_{ij}.

2.3.1 Probability of obtaining a zero declaration

$$\begin{aligned} P(D_j = 0|S, X) &= \prod_{i \in \mathcal{P}_j} P(Y_{ij} = 0|S, X), \\ &= \exp\left\{ -\sum_{i \in \mathcal{P}_j} e^{\xi}.S(x_{ij}) \right\} = \pi_j. \end{aligned}$$

2.3.2 Expectancy of a positive declaration

Following calculations are supposed to be conditionnal on S and X.

$$E(D_j|D_j > 0) = \sum_{i \in \mathcal{P}_j} E(C_{ij} Z_{ij} | \exists i \in \mathcal{P}_j, C_{ij} = 1)$$

As C_{ij} and Z_{ij} are assumed to be independant.

$$\begin{aligned} E(D_j|D_j > 0) &= E(D_j 1_{\{D_j > 0\}}) / P\left(D_j > 0\right), \\ &= E(D_j 1_{\{D_j > 0\}}) / \left(1 - \pi_j\right). \end{aligned}$$

As $E(D_j 1_{\{D_j > 0\}}) = E(D_j)$,

$$\begin{aligned} E(D_j|D_j > 0) &= (1 - \pi_j)^{-1} E(D_j), \\ &= (1 - \pi_j)^{-1} \sum_{i \in \mathcal{P}_j} E(C_{ij} Z_{ij}), \\ &= (1 - \pi_j)^{-1} \sum_{i \in \mathcal{P}_j} (1 - p_{ij}) \frac{S(x_{ij})}{1 - p_{ij}}, \\ &= (1 - \pi_j)^{-1} \sum_{i \in \mathcal{P}_j} S(x_{ij}). \end{aligned}$$

2.3.3 Variance of a positive declaration

$$Var(D_j|D_j > 0) = E(D_j^2|D_j > 0) - E(D_j|D_j > 0)^2.$$

$$E(D_j^2|D_j > 0) = (1-\pi_j)^{-1}E(D_j^2 1_{\{D_j>0\}}) = (1-\pi_j)^{-1}E(D_j^2)$$

$$E(D_j|D_j > 0)^2 = ((1-\pi_j)^{-1}E(D_j 1_{\{D_j>0\}}))^2 = (1-\pi_j)^{-2}E(D_j)^2$$

And then,

$$Var(D_j|D_j > 0) = (1-\pi_j)^{-1}E(D_j^2) - (1-\pi_j)^{-2}E(D_j)^2 = (1-\pi_j)^{-1}Var(D_j) - \frac{\pi_j}{(1-\pi_j)^2}E(D_j)^2.$$

As the $(Y_{ij})_{i\in\mathcal{P}_j}$ are independent, $Var(D_j) = \sum_{i\in\mathcal{P}_j} Var(Y_{ij}) = \sum_{i\in\mathcal{P}_j} Var(C_{ij}.Z_{ij})$.

$$\begin{aligned}
Var(C_{ij}Z_{ij}) &= E(C_{ij}^2 Z_{ij}^2) - E(C_{ij}Z_{ij})^2, \\
&= E(C_{ij}^2)E(Z_{ij}^2) - E(C_{ij})^2 E(Z_{ij})^2, \\
&= (1-p_{ij})E(Z_{ij}^2) - (1-p_{ij})^2 E(Z_{ij})^2, \\
&= (1-p_{ij})(Var(Z_{ij}) + E(Z_{ij})^2) - (1-p_{ij})^2 E(Z_{ij})^2, \\
&= \frac{S(x_{ij})^2}{1-p_{ij}}(e^{\sigma^2}-1) + \frac{S(x_{ij})^2}{1-p_{ij}} - S(x_{ij})^2, \\
&= \frac{S(x_{ij})^2}{1-p_{ij}}(e^{\sigma^2}-(1-p_{ij}))
\end{aligned}$$

2.3.4 Conclusion

$$P(D_j = 0|S, X) = \exp\left\{-\sum_{i\in\mathcal{P}_j} e^{\xi}.S(x_{ij})\right\} = \pi_j$$

$$D_j|D_j > 0 \sim \mathcal{LN}(E(D_j|D_j > 0); ln(\frac{Var(D_j|D_j > 0)}{E(D_j|D_j > 0)^2} + 1))$$

$$E(D_j|D_j > 0) = \frac{\sum_{i\in\mathcal{P}_j} S(x_{ij})}{1-\pi_j}$$

$$Var(D_j|D_j > 0) = \frac{\sum_{i\in\mathcal{P}_j} Var(Y_{ij})}{1-\pi_j} - \frac{\pi_j}{(1-\pi_j)^2}E(D_j)^2$$

$$Var(Y_{ij}) = \frac{S(x_{ij})^2}{1-p_{ij}}(e^{\sigma^2}-(1-p_{ij}))$$

2.3.5 Simulations