Tail probability

Known variance

Tail probability

$$v_\alpha(Y_i, \sigma_i) = P(\theta_i > \theta_\alpha | Y_i, \sigma_i^2) = \frac{\int_{\theta_\alpha}^{\infty} f(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}{\int_{-\infty}^{\infty} f(y_i|\theta_i, \sigma_i^2) dG(\theta_i)}$$

Capactiy constraint

$$P(v_\alpha(Y_i, \sigma_i) > \lambda) \leq \alpha$$

Since

$$v_\alpha(Y_i, \sigma_i) > \lambda \Leftrightarrow Y_i > t(\lambda, \sigma_i)$$

We have

$$P(Y_i > t(\lambda, \sigma_i)) \leq \alpha$$

which can be written explicitly as

$$\int \int (1 - \Phi(\frac{t(\lambda, \sigma_i) - \theta_i}{\sigma_i})) dG(\theta_i) dF(\sigma_i) \leq \alpha$$

FDR constraint

The marginal/local false discovery rate is defined as

$$\text{mFDR} = P(\theta_i \leq \theta_\alpha | \delta_i = 1, \sigma_i) \approx \text{lFDR} = \frac{E[\sum (1 - h_i)\delta_i]}{E[\sum \delta_i]} \leq \gamma$$

The left hand side of $\approx$ can be written as

$$P(\theta_i \leq \theta_\alpha | \delta_i = 1, \sigma_i) = \frac{P(\theta_i \leq \theta_\alpha, v_\alpha(Y_i, \sigma_i) > \lambda | \sigma_i)}{P(v_\alpha(Y_i, \sigma_i) > \lambda | \sigma_i)} = \frac{P(\theta_i \leq \theta_\alpha, Y_i > t(\lambda, \sigma_i) | \sigma_i)}{P(Y_i > t(\lambda, \sigma_i) | \sigma_i)}$$

which is

$$\frac{\int \int_{-\infty}^{\theta_\alpha} 1 - \Phi(t(\lambda, \sigma_i) - \theta_i / \sigma_i) dG(\theta_i) dF(\sigma_i)}{\int \int_{-\infty}^{\infty} 1 - \Phi(t(\lambda, \sigma_i) - \theta_i / \sigma_i) dG(\theta_i) dF(\sigma_i)} \leq \gamma$$

The right hand side of $\approx$ can be written as

$$\frac{\sum_i P(\theta_i \leq \theta_\alpha, Y_i > t(\lambda, \sigma_i) | \sigma_i)}{\sum_i P[1\{v_\alpha(Y_i, \sigma_i) > \lambda\}]}$$

It is left to shown that

$$P(\theta_i \leq \theta_\alpha, Y_i > t(\lambda, \sigma_i) | \sigma_i) = E[(1 - v_\alpha(Y_i, \sigma_i)) 1\{v_\alpha(Y_i, \sigma_i) > \lambda\}]$$

The LHS can be rewritten as

$$\begin{aligned}
& E_Y[E[1\{\theta_i \leq \theta_\alpha, Y_i > t(\lambda, \sigma_i)\} | Y_i, \sigma_i]] \\
&= E_Y[1\{Y_i > t(\lambda, \sigma_i)\} E[1\{\theta_i \leq \theta_\alpha | Y_i, \sigma_i\}]] \\
&= E_Y[1\{Y_i > t(\lambda, \sigma_i)\}(1 - v_\alpha(Y_i, \sigma_i))] \\
&= E[1\{v_\alpha(Y_i, \sigma_i) > \lambda\}(1 - v_\alpha(Y_i, \sigma_i))]
\end{aligned}$$

Unknown variance

Tail probability

We have

$$Y_{it} = \theta_i + \sigma_i \epsilon_{it}$$

The sufficient statistics for θ_i is

$$Y_i = \frac{1}{T_i} \sum_{t=1}^{T_i} Y_{it}$$

and for σ_i^2 is

$$S_i = \frac{1}{T_i - 1} \sum_{t=1}^{T_i} (Y_{it} - Y_i)^2$$

Conditional on the true parameter (θ_i, σ_i^2), the sufficient statistics follow

$$Y_i|\theta_i, \sigma_i^2 \sim N(\theta_i, \sigma_i^2/T_i) S_i|\theta_i, \sigma_i^2 \sim \Gamma(r_i = (T_i - 1)/2, 2\sigma_i^2/(T_i - 1))$$

Therefore, the tail probability is

$$v_\alpha(Y_i, S_i) = P(\theta_i > \theta_\alpha | Y_i, S_i) = \frac{\int_{\theta_\alpha}^{\infty} \Gamma(s_i|r_i, \sigma_i^2) f(y_i|\theta_i, \sigma_i^2) dG(\theta_i, \sigma_i^2)}{\int_{-\infty}^{\infty} \Gamma(s_i|r_i, \sigma_i^2) f(y_i|\theta_i, \sigma_i^2) dG(\theta_i, \sigma_i^2)}$$

Capactiy constraint

$$\int\int P(v_\alpha(Y_i, S_i) > \lambda) dG(\theta_i, \sigma_i^2) \leq \alpha$$

This is a bit tricky with $v_\alpha(Y_i, S_i)$ since we don't have the monoticiy as before (given σ_i, $v_\alpha(Y_i, \sigma_i)$ is increasing in Y_i)

The best we can do is to write

$$P(v_\alpha(Y_i, S_i) > \lambda) = P((Y_i, S_i) \in \mathcal{S}) \leq \alpha$$

If $v_\alpha(Y_i, S_i)$ is increasing in Y_i for a given S_i, we can draw *level curve* for $v_\alpha(Y_i, S_i)$ and find the region $\mathcal{S}$ that satisfies the capacity constraint. But they are under special cases.

FDR constraint

lFDR

$$\int\int \frac{E[1\{v_\alpha(Y_i, S_i) > \lambda\}(1 - v_\alpha(Y_i, S_i))]}{E[1\{v_\alpha(Y_i, S_i) > \lambda\}]} dG(\theta_i, \sigma_i^2) \leq \gamma$$