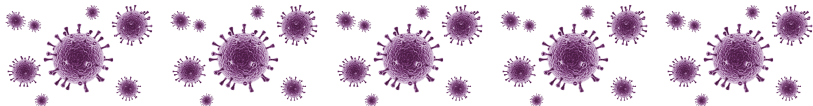


# Modèles Linéaires Appliqués / Régression GLM & Tests

Arthur Charpentier

UQAM

Hiver 2020 - COVID-19 # 18



## Tests & Intervalles de Confiance

La déviance était définie par

$$D(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2\varphi [\log \mathcal{L}(\mathbf{y}) - \log \mathcal{L}(\hat{\boldsymbol{\mu}})]$$

$$D^*(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2 [\log \mathcal{L}(\mathbf{y}) - \log \mathcal{L}(\hat{\boldsymbol{\mu}})] = \frac{D}{\varphi}$$

Si on veut tester  $H_0 : \beta_I = \mathbf{0}$  pour un sous ensemble de variables,  $I \subset \{1, 2, \dots, k\}$ , on fait un test de rapport de vraisemblance

$$LRT = D^*(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}}_I) = D^*(\mathbf{y}, \hat{\boldsymbol{\mu}}_I) - D^*(\mathbf{y}, \hat{\boldsymbol{\mu}})$$

Sous  $H_0$ ,  $LRT \xrightarrow{\mathcal{L}} \chi^2(k - \dim(I))$

# Convergences Asymptotiques

On a vu que

**Proposition** sous les hypothèses mentionnées auparavant

- l'estimateur du maximum de vraisemblance  $\hat{\beta}$  existe et est unique,
- $\hat{\beta} \xrightarrow{p.s.} \beta$  (fortement consistant)
- lorsque  $n \rightarrow \infty$ ,  $\hat{\beta} - \beta \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \varphi \Sigma)$

où  $\Sigma = (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1}$  et  $\mathbf{W} = \text{diag}((V(\mu_i)g'(\mu_i)^2)^{-1})$ .

**Proposition** sous les hypothèses mentionnées auparavant

- lorsque  $n \rightarrow \infty$ ,  $\hat{\eta} - \eta \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \varphi \mathbf{X} \Sigma \mathbf{X}^\top)$
- lorsque  $n \rightarrow \infty$ ,  $\hat{\mu} - \mu \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \varphi \Delta^{-2} \mathbf{X} \Sigma \mathbf{X}^\top)$

## Convergences Asymptotiques

- lorsque  $n \rightarrow \infty$ ,  $\hat{\beta} - \beta \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \varphi \Sigma)$

```
1 > reg = glm(Survived~Sex+Pclass+Age,family="binomial")
2 > confint(reg)
3              2.5 %          97.5 %
4 (Intercept)  3.01450041    4.58890189
5 Sexmale     -2.93866348   -2.12456642
6 Pclass2     -1.86335813   -0.77181573
7 Pclass3     -3.14672951   -2.04176314
8 Age         -0.05229015   -0.02223298
```

- lorsque  $n \rightarrow \infty$ ,  $\hat{\mu} - \mu \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \varphi \Delta^{-2} \mathbf{X} \Sigma \mathbf{X}^\top)$

```
1 > predict(reg,newdata=newbase,se.fit = TRUE,type="
  response")
2 $fit
3           1           2
4 0.9588402 0.1124358
5 $se.fit
6           1           2
7 0.01217853 0.01834374
```

# Convergences Asymptotiques

En réalité, on utilise ici la **vraisemblance profilée**

```
1 > confint(reg)
2 Waiting for profiling to be done...
3           2.5 %           97.5 %
4 Age      -0.05229015 -0.02223298
```

$\beta = (\beta_1, \dots, \beta_{k-1}, \beta_k)$ . Pour  $\beta_k$ , on définit

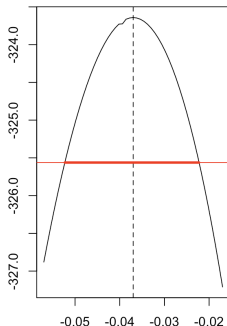
$$\log \mathcal{L}_p(\beta_k) = \log \mathcal{L}((\beta_{-k}^*, \beta_k))$$

où  $\beta_{-k}^* = \operatorname{argmax} \{ \log \mathcal{L}((\beta_{-k}, \beta_k)) \}$

L'intervalle de confiance est en :

$$\max \{ \log \mathcal{L}_p(\beta_k) \} - \frac{1}{2} q_{\chi^2}(.95)$$

cf test du rapport de vraisemblance



## Convergences Asymptotiques

Soit  $\mathbf{R}$  une matrice  $q \times p$  (avec  $q \leq p$ ), alors

$$\frac{1}{\varphi} (\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}))^\top \left( \mathbf{R}(\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{R}^\top \right)^{-1} (\mathbf{R}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})) \xrightarrow{\mathcal{L}} \chi^2(n-p)$$

Si on veut tester  $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}_0$ , contre  $H_1 : \mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}_0$ , on utilise

$$Q = \frac{1}{\varphi} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}_0)^\top \left( \mathbf{R}(\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{R}^\top \right)^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r}_0)$$

et sous  $H_0$ ,  $Q \xrightarrow{\mathcal{L}} \chi^2(n-p)$

# Vraisemblance Profilée

```
1 > p0 = rep(0,4)
2 > profL = function(b){
3   Lp = function(p){
4     eta = X%%c(p,b)
5     mu = exp(eta)/(1+exp(eta))
6     -sum(dbinom(reg$y,size = 1,prob=mu,
7       log=TRUE))}
8   -optim(par=p0,fn=Lp)$value }
9 > m=profL(coefficients(reg)[5])
10 > minf=m-qchisq(.95,df=1)/2
11 > abline(h=minf,col="red")
12 [1] -0.05227511
13 > (b1=uniroot(function(z) Vectorize(
14   profL)(z)-minf,c(-.06,-.04))$root)
15 [1] -0.02225003
```

