STT 3260
Devoir 3

Mathieu Lemire

2025-10-15

```r
library(survival)
```

Exercice 1

a)

Estimations des paramètres pour le groupe 0

Pour le groupe 0, on a

```r
data <- read.table("C:/Users/14385/Desktop/A25/STT 3260/Devoirs/Devoir 3/kidney.txt",
                   header=TRUE)

groupe_0_weib <- survreg(Surv(time, delta) ~ 1,
                         data = data[data$type == 2,],
                         dist = "weibull")
summary(groupe_0_weib)
```

```
##
## Call:
## survreg(formula = Surv(time, delta) ~ 1, data = data[data$type ==
##     2, ], dist = "weibull")
##             Value Std. Error    z       p
## (Intercept) 5.411      1.024 5.29 1.3e-07
## Log(scale)  0.616      0.265 2.32    0.02
##
## Scale= 1.85
##
## Weibull distribution
## Loglik(model)= -51.6   Loglik(intercept only)= -51.6
## Number of Newton-Raphson Iterations: 7
## n= 76
```

Estimation de α_0

On a

$$\log(\hat{\sigma}_0) = 0.616 \implies \hat{\sigma}_0 = \exp(0.616) \approx 1.852$$

Comme $\sigma = \frac{1}{\alpha}$, on a $\alpha = \frac{1}{\sigma}$. On calcule $\hat{\alpha}_0$

$$\hat{\alpha}_0 = \frac{1}{\hat{\sigma}_0} = \frac{1}{1.852} \approx 0.54$$

```
mu_0 <- unname(groupe_0_weib$coefficients)
sigma_0 <- groupe_0_weib$scale
alpha_0 <- 1/sigma_0
alpha_0
```

```
## [1] 0.540188
```

Écart type de $\hat{\alpha}_0$

On doit utiliser la méthode delta:

$$\widehat{Var(g(\hat{\beta}))} = (g'(\hat{\beta}))^2 Var(\hat{\beta}) \implies \sqrt{\widehat{Var(g(\hat{\beta}))}} = \left|g'(\hat{\beta})\right| \sqrt{Var(\hat{\beta})}$$

Puisque la relation entre $\log(\sigma)$ et α est $\alpha = g(\log(\sigma)) = \exp(-\log(\sigma))$, on a :

$$\frac{\hat{\alpha}}{d(\log \hat{\sigma})} = -\exp(-\log(\hat{\sigma}))$$

$$\begin{aligned}
\sqrt{\widehat{Var(\hat{\alpha}_0)}} &= \frac{\hat{\alpha}_0}{d(\log \hat{\sigma}_0)} \sqrt{Var(\log \hat{\sigma}_0)} \\
&= \left| -\exp(-\log(\hat{\sigma}_0)) \right| \sqrt{Var(\log \hat{\sigma}_0)} \\
&= \left| -\hat{\alpha}_0 \right| \sqrt{Var(\log \hat{\sigma}_0)} \\
&\approx \left| -0.54 \right| \times 0.265 \\
&= 0.1431
\end{aligned}$$

Estimation de λ_0

La relation entre μ et λ est $\mu = \frac{-\log(\lambda)}{\alpha} \implies \lambda = \exp(-\alpha\mu)$. Puisque $\hat{\mu}_0 = 5.411$, on a

$$\hat{\lambda}_0 = \exp(-\hat{\alpha}_0 \hat{\mu}_0) = \exp(-0.54 \times 5.411) \approx 0.0538$$

```
lambda_0 <- exp(-alpha_0 * mu_0)
lambda_0 <- unname(lambda_0)
lambda_0
```

```
## [1] 0.05377752
```

Écart type de $\hat{\lambda}_0$

On pose $\eta = \log(\sigma)$, ce qui implique que $\lambda = \exp\left(-\frac{\mu}{\sigma}\right) = \exp\left(-\mu\exp(-\log\mu)\right) = \exp\left(-\mu e^{-\eta}\right)$.

On calcule les dérivées de $\lambda = g(\mu, \eta)$:

$$\frac{\partial\lambda}{\partial\mu} = -e^{-\eta}\exp(-\mu e^{-\eta}) = -\lambda e^{-\eta} \qquad \frac{\partial\lambda}{\partial\eta} = \mu e^{-\eta}\exp(-\mu e^{-\eta}) = \lambda\mu e^{-\eta}$$

On calcule la variance estimée de $\hat{\lambda}$

$$\widehat{Var(\hat{\lambda})} = \begin{pmatrix} \frac{\partial\hat{\lambda}}{\partial\hat{\mu}} & \frac{\partial\hat{\lambda}}{\partial\hat{\eta}} \end{pmatrix} \begin{pmatrix} Var(\hat{\mu}) & Cov(\hat{\mu},\hat{\eta}) \\ Cov(\hat{\mu},\hat{\eta}) & Var(\hat{\eta}) \end{pmatrix} \begin{pmatrix} \frac{\partial\hat{\lambda}}{\partial\hat{\mu}} \\ \frac{\partial\hat{\lambda}}{\partial\hat{\eta}} \end{pmatrix}$$

```
eta_0 <- log(sigma_0)
vecteur_der_0 <-c(-lambda_0*exp(-eta_0) , lambda_0*mu_0 *exp(-eta_0))
vecteur_der_0
```

```
## [1] -0.02904997  0.15718630
```

```
matrice_var_0 <- groupe_0_weib$var
matrice_var_0
```

```
##             (Intercept) Log(scale)
## (Intercept)   1.0478158 0.22738874
## Log(scale)    0.2273887 0.07022616
```

```
var_lamda_0 <- as.numeric(t(unname(vecteur_der_0)) %*% unname(matrice_var_0) %*% unname(vecteur_der_0))
var_lamda_0
```

```
## [1] 0.0005427366
```

Donc $\widehat{Var(\hat{\lambda}_0)} \approx 0.0005427$ et l'écart type estimé est

```
std_err_lambda_0 <- as.numeric(sqrt(var_lamda_0))
std_err_lambda_0
```

```
## [1] 0.02329671
```

Donc $\widehat{SE(\hat{\lambda}_0)} \approx 0.0233$

Estimations des paramètres pour le groupe 1

```
groupe_1_weib <- survreg(Surv(time, delta) ~ 1, data = data[data$type == 1,], dist = "weibull")
summary(groupe_1_weib)
```

```
##
## Call:
## survreg(formula = Surv(time, delta) ~ 1, data = data[data$type ==
##     1, ], dist = "weibull")
##               Value Std. Error    z      p
## (Intercept)  3.194      0.180 17.7 <2e-16
## Log(scale)  -0.467      0.203 -2.3  0.021
##
## Scale= 0.627
##
## Weibull distribution
## Loglik(model)= -65   Loglik(intercept only)= -65
## Number of Newton-Raphson Iterations: 6
## n= 43
```

Estimation de α_1

On a

$$\log(\hat{\sigma}_1) = -0.467 \implies \hat{\sigma}_1 = \exp(-0.467) \approx 0.627 \implies \hat{\alpha}_1 = \frac{1}{\hat{\sigma}_1} = \frac{1}{0.627} \approx 1.595$$

```r
mu_1 <- unname(groupe_1_weib$coefficients)
sigma_1 <- groupe_1_weib$scale
alpha_1 <- 1/sigma_1
alpha_1
```

```
## [1] 1.594824
```

Écart type de $\hat{\alpha}_1$

$$\begin{aligned}
\sqrt{\widehat{Var(\hat{\alpha}_1)}} &= \frac{\hat{\alpha}_1}{d(\log \hat{\sigma}_1)} \sqrt{Var(\log \hat{\sigma}_1)} \\
&= \Big| -\exp(-\log(\hat{\sigma}_1)) \Big| \sqrt{Var(\log \hat{\sigma}_1)} \\
&= \Big| -\hat{\alpha}_1 \Big| \sqrt{Var(\log \hat{\sigma}_1)} \\
&\approx \Big| -1.595 \Big| \times 0.203 \\
&= 0.323785
\end{aligned}$$

Estimation de λ_1

Puisque $\hat{\mu}_1 = 3.194$, on a

$$\hat{\lambda}_1 = \exp(-\hat{\alpha}_1\hat{\mu}_1) = \exp(-1.595 \times 3.194) \approx 0.00613$$

```r
lambda_1 <- exp(-alpha_1 * mu_1)
lambda_1 <- unname(lambda_1)
lambda_1
```

```
## [1] 0.006135598
```

Écart type de $\hat{\lambda}_1$

On calcule en premier les dérivés partielles de $\hat{\lambda}_1$ par rapport à η_1 et par rapport à μ_1

```
eta_1 <- log(sigma_1)
vecteur_der_weib_1 <-c(-lambda_1*exp(-eta_1) , lambda_1*mu_1 *exp(-eta_1))
vecteur_der_weib_1
```

```
## [1] -0.009785197  0.031252575
```

La matrice de covariance $Cov(\hat{\mu}, \hat{\eta})$ est

```
matrice_var_1 <- groupe_1_weib$var
matrice_var_1
```

```
##             (Intercept) Log(scale)
## (Intercept)  0.03258009 0.01616489
## Log(scale)   0.01616489 0.04102682
```

On calcule à présent $Cov(\hat{\lambda}_1)$

```
var_lambda_1 <- as.numeric(t(unname(vecteur_der_weib_1)) %*% unname(matrice_var_1) %*% unname(vecteur_de
var_lambda_1
```

```
## [1] 3.330455e-05
```

On calcule maintenant l'écart type

```
std_err_lambda_1 <- sqrt(var_lambda_1)
std_err_lambda_1
```

```
## [1] 0.005771009
```

On a $\widehat{SE(\hat{\lambda}_1)} = 0.00577$

b)

On teste l'hypothèse que $\alpha = 1$ pour chacun des groupe. On pose

$$H_0 : \alpha = 1 \qquad H_1 : \alpha \neq 1$$

Mais puisqu'on a $\alpha = \frac{1}{\sigma} \implies \log(\alpha) = -\log(\sigma)$ et pour $\alpha = 1 \implies \log(\sigma) = 0$ On avait déjà posé plus haut que $\eta = \log(\sigma)$. Nos hypothèses sont alors:

$$H_0 : \eta = 0 \qquad H_1 : \eta \neq 0$$

Tests pour le groupe 0

On rappel la sortie:

```r
summary(groupe_0_weib)
```

```
##
## Call:
## survreg(formula = Surv(time, delta) ~ 1, data = data[data$type ==
##     2, ], dist = "weibull")
##             Value Std. Error    z       p
## (Intercept) 5.411      1.024 5.29 1.3e-07
## Log(scale)  0.616      0.265 2.32    0.02
##
## Scale= 1.85
##
## Weibull distribution
## Loglik(model)= -51.6   Loglik(intercept only)= -51.6
## Number of Newton-Raphson Iterations: 7
## n= 76
```

Test de Wald

La statistique de Wald est déjà calculer, $W = 2.32$ et la p-valeur associé à la statistique de Wald pour η est de $0.02 < 0.05$. On peut rejetter l'hypothèse nulle au niveau 5% mais nous ne pouvons pas rejetter l'hypothèse nul au niveau 1%

Test du rapport de vraisemblance

Pour $\alpha = 1$ la distribution Weibull(α, λ) suit une distribution Exponentielle(λ). On peut faire l'ajustement avec l'exponentielle et faire le test du rapport de vraisemblance puisque l'Exponentielle(λ) est emboîté dans la Weibull$(\alpha = 1, \lambda)$

```r
groupe_0_exp <- survreg( Surv(time, delta) ~1,
                         data = data[data$type == 2,],
                         dist = "exp")
summary(groupe_0_exp)
```

```
##
## Call:
## survreg(formula = Surv(time, delta) ~ 1, data = data[data$type ==
##     2, ], dist = "exp")
##             Value Std. Error    z      p
## (Intercept) 4.011      0.302 13.3 <2e-16
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -55.1   Loglik(intercept only)= -55.1
## Number of Newton-Raphson Iterations: 6
## n= 76
```

On calcule la statistitque de test:

$$\Lambda = -2[\ell(\alpha = 1, \hat{\lambda}) - \ell(\hat{\alpha}, \hat{\lambda})] = -2[-55.1 - (-51.6)] = 7$$

Nous avons un paramètres de différents entre les modèles donc on calcule la valeur-p avec une χ^2_1. La valeur-p est

$$\mathbb{P}(\chi^2_1 \geq 7) =$$

```r
valeur_p_rap_vrai_groupe_0 <-pchisq(7,1,lower.tail = FALSE)
names(valeur_p_rap_vrai_groupe_0) <- "Valeur-p"
valeur_p_rap_vrai_groupe_0
```

```
##    Valeur-p
## 0.008150972
```