

Vous trouverez le fichier de données *chomage\_tp.rda* dont vous avez besoin pour ce TP sur la page dédiée à ce cours du site web <http://www.andreberchtold.com>.

### Exercice 1 Table de survie avec la méthode actuarielle

Le tableau suivant (fictif) donne pour 202 personnes le temps entre le début et la fin du chômage, en distinguant entre celles qui retrouvent un emploi salarié et celles qui quittent le chômage pour d'autres raisons.

$k$	temps en mois	Durée du chômage	
		nouvel emploi salarié	autres raisons
1	$[0, 2[$	77	30
2	$[2, 4[$	16	28
3	$[4, 6[$	11	14
4	$[6, 8[$	5	6
5	$[8, 15]$	-	15

On s'intéresse au temps de chômage avant de retrouver un emploi salarié. Le "risque" ci-après concerne les chances de retrouver un emploi.

1. Complétez le tableau de survie en donnant pour chaque classe  $k$  :

$n_k$  : nombre de cas exposés au risque après une durée correspondant à la borne inférieure de la classe  $k$ .

$R_k$  : nombre de cas exposés utilisé pour l'estimation du risque.

$\hat{h}_k$  : estimation du risque instantané par intervalle.

$\hat{S}_k$  : estimation de la probabilité de rester au chômage pour une durée égale ou supérieure à celle de la classe  $k$ .

$\hat{p}_k$  : estimation de la probabilité de rester au chômage au moins jusqu'au début et au plus tard jusqu'à la fin de la classe  $k$ .

2. Vous trouverez les 202 données individuelles dans le fichier *chomage\_tp.rda*. Ce fichier contient 3 variables, à savoir le genre, le temps en mois avant qu'une personne quitte le chômage et le statut (événement "avoir retrouvé un emploi salarié" ou censure). Calculez une table de survie pour ces données en utilisant la fonction *lifetab()* du package *KMsurv* de R, puis représentez graphiquement la courbe de survie. Comparez les résultats avec ceux que vous avez obtenus précédemment et commentez-les.

*Remarque : Construisez les variables nécessaires à l'utilisation de *lifetab()* à partir du fichier de données en utilisant les fonctionnalités de construction de variables de R.*

3. Construisez un intervalle de confiance à 95% pour la courbe de survie en utilisant les erreurs standard fournies par *lifetab()*.
4. Construisez deux courbes de survie séparées pour les femmes et les hommes, puis comparez-les à l'aide d'un test du log-rank. Commentez le résultat.

## Exercice 2 Courbes de survie en temps continu

On utilise à nouveau le fichier *chomage\_tp.rda*.

1. En utilisant la fonction *survfit* du package *survival*, construisez la courbe de survie de Kaplan-Meier, puis représentez-la graphiquement. Représentez aussi la fonction de hasard cumulé sur un graphique. Commentez les résultats obtenus.
2. Construisez la courbe de survie selon la méthode de Nelson-Aalen et comparez avec le résultat précédent.
3. Construisez et représentez graphiquement des courbes de survie séparées pour les hommes et les femmes selon la méthode de Kaplan-Meier, puis comparez les deux courbes à l'aide du log-rank test. Commentez les résultats.