Modèles de durée / Examen / Janvier 2019

Durée 2h – aucun document n'est autorisé

Corrigé

Quelques propriétés des fonctions de survie et de hasard cumulé

La qualité de la rédaction, des justifications apportées et de la présentation de la copie seront prises en compte dans la notation.

On considère une variable de durée *X* dont la fonction de hasard sous-jacente est notée *h* et la fonction de survie *S*.

Question n°1 (4 points) : Rappelez les définition des fonctions suivantes : survie, survie conditionnelle, hasard, hasard cumulé et les relations entre ces grandeurs (ces relations devront être démontrées).

> Voir le support de cours

Question n°2 (2 points) : on considère une subdivision $0 = t_0 < t_1 < \ldots < t_n = t$ de l'intervalle $[0,t]$. Quelle est la limite lorsque le pas de discrétisation tend vers 0 de $Z_n(t) = \sum_i P(T \leq t_i | T > t_{i-1})$?

> $Z_n(t)$ tend vers $H(t)$; pour le montrer, on écrit
>
> $$\sum_i P(T \leq t_i | T > t_{i-1}) = \sum_i \frac{S(t_{i-1}) - S(t_i)}{S(t_{i-1})} \to -\int_0^t \frac{dS(u)}{S(u)} = H(t)$$

Question n°3 (2 points) : On note, pour $t \leq u$, $S(t,u) = \frac{S(u)}{S(t)}$, la fonction de survie conditionnelle. Montrez que, pour $s \leq t \leq u$, $S(s,u) = S(s,t)S(t,u)$

> $$S(s,u) = \frac{S(u)}{S(s)} = \frac{P(X>u)}{P(X>s)} = \frac{P(X>u)}{P(X>t)}\frac{P(X>t)}{P(X>s)} = \frac{S(u)}{S(t)}\frac{S(t)}{S(s)} = S(s,t)S(t,u)$$

Question n°4 (2 points) : En déduire la loi de X si $S(t,u) = S(u-t)$, c'est-à-dire si $S(t,u)$ ne dépend que de la différence $u-t$?

$f = hS \qquad S(x) = e^{-H(x)}$

> On déduit de la question précédente que la fonction f satisfait l'égalité $S(x+y) = S(x)S(y)$; donc

On pose $x = t-s$ et $y = u-t$

$S(x+y) = S(t-s+u-t) = S(u-s) = S(s,u) = S(s,t)S(t,u)$

$= S(t-s)S(u-t) = S(x)S(y)$

$$\frac{S(x+y)-S(x)}{y} = \frac{S(x)S(y)-S(x)}{y} = S(x)\,\frac{S(y)-1}{y} \xrightarrow[y\to 0]{} S'(x) = -\lambda S(x)$$

Car

$$\frac{S(x+y)-S(x)}{y} = \frac{S(x)S(y)-S(x)}{y} = S(x)\frac{S(y)-1}{y}$$

et en faisant tendre y vers 0 on trouve que $S'(x) = -\lambda S(x)$, soit $S(x) = e^{-\lambda x}$: il s'agit d'une loi exponentielle.

Question n°5 (2 points) : Quelle est la loi de $H(X)$?

$$P\big(H(X) > x\big) = P\big(X > H^{-1}(x)\big) = S\big(H^{-1}(x)\big) = \mathbf{exp}\Big(-H\big(H^{-1}(x)\big)\Big) = \mathbf{exp}(-x)$$

Question n°6 (3 points) : En désignant par dt l'intervalle $[t, t+dt[$ et en utilisant la notation $H(s,t) = H(t) - H(s)$, montrer que $H(dt) = 1 - S(dt)$. Comment peut-on écrire $H(t)$ en fonction de $S(du)$? $S(t)$ en fonction de $H(du)$?

Par définition $H(dt) = H(t+dt) - H(t) = h(t)dt = \dfrac{S(t) - S(t+dt)}{S(t)} = 1 - S(dt)$

Avec la propriété triviale d'additivité de H, on a $H(t) = \int_0^t \big(1 - S(du)\big)$; pour la relation duale pour S, on utilise le fait que S est multiplicative et cela conduit à l'expression $S(t) = \prod_0^t \big(1 - H(du)\big)$, au sens du produit-limite.

Question n°7 (3 points) : Rappelez les expressions des estimateurs de Nelson-Aalen de la fonction de hasard cumulée et de Kaplan-Meier de la fonction de survie, discutez les liens que vous voyez avec les expressions obtenues à la question précédente.

On a $\hat{S}(t) = \prod_{s \le t} \big(1 - \Delta\hat{H}(s)\big)$ et $\hat{H}(t) = \sum_{\{s \le t\}} \big(1 - \Delta\hat{S}(s)\big)$, avec

$$\Delta\hat{H}(s) = 1 - \Delta\hat{S}(s) = \frac{\Delta\bar{N}^1(s)}{\bar{R}(s)}.$$

Question n°8 (3 points) : Comment construire simplement un estimateur de la fonction de survie à partir de l'estimateur de Nelson-Aalen ? Si vous disposez d'un estimateur de la variance de l'estimateur de Nelson-Aalen, comment faites-vous pour en déduire un estimateur de la variance de l'estimateur de la fonction d survie ainsi obtenu ?

$$\hat{S}_{HF}(t) = \mathbf{exp}\big(-\hat{H}_{NA}(t)\big)$$

Sa variance peut être obtenue par la méthode Delta qui, sous des conditions raisonnables de régularité de la fonction f permet d'écrire que

$V\left(f\left(X\right)\right)\approx\left(\frac{df}{dx}\left(E\left(X\right)\right)\right)^2 V\left(X\right)$. En effet, si $X=\mu+\sigma Z$ avec σ petit et Z centrée réduite, on remarque que pour une fonction $x\to f(x)$ suffisamment régulière, en effectuant le développement limité $f\left(\mu+h\right)\approx f\left(\mu\right)+h\frac{df}{dx}\left(\mu\right)$, on trouve que $V\left(f\left(X\right)\right)\approx V\left(f\left(\mu\right)+\sigma Z\frac{df}{dx}\left(\mu\right)\right)=\sigma^2\frac{df}{dx}\left(\mu\right)^2$. En prenant ici $f(x)=e^{-x}$, on trouve que $V\left(\hat{S}\right)\approx e^{-2E\left(\hat{H}\right)}V\left(\hat{H}\right)\approx\hat{S}^2 V\left(\hat{H}\right)$,