UE 11 Biomédecine Quantitative

20/03 de 13h30 à 15h30

Pr. Chevret

Ronéotypeur : Mouasseh Maxime

Ronéolecteur : Aurélie Abatuci

COURS N°7 : LES COURBES DE SURVIE

Pour ce cours, aucune formule n’est à apprendre par cœur, les Qcms du CC porteront sur des courbes, tous les points en gras sont ceux que Mme Chevret a souligné pendant ce cours et qui seront donc (normalement) dans le CC.

**SOMMAIRE**

Je remets le même plan que la prof, cependant le cours n’est absolument pas linéaire et aborde énormément de notion au fur et à mesure.

 1 - Qu’est-ce que le pronostic et la survie ?

 2 - Quelle est la définition de la survie ?

 3 -Comment mesurer la survie ?

 4 -Comment estimer la survie ?

 5 -Comment prédire la survie ?

**I & II – Pronostic et Survie : Définition et Histoire.**

Le pronostic est le devenir d’une maladie, son évolution. On peut le quantifier par mort/survie : on survit à la maladie, ou on en meurt.

La notion de pronostic a émergé dans les années 70 pour quantifier les survies en s’intéressant au devenir de la maladie en termes de mortalité.

En médecine on parle de **Survie** et non pas de **Mortalité** : la définition de la survie et de la mortalité est binaire ; en tout ou rien (Ou on survit, ou on meurt).

On peut assimiler la survie à une loi de Bernoulli ou 0 équivaut à vivant et 1 équivaut à mort.

La survie est indexée sur le temps et on peut la mesurer à différents moments : de l’entrée en réanimation jusqu’à J14 (14 jours après l’entrée) ou de l’entrée jusqu’à J90 etc…

Le premier exemple donné dans ce cours est celui d’ALLO-ICU :

On regarde la survie de patients suivis en réanimation. Parmi 497 malades, 194 décèdent en réanimation, 85 décèdent à l’hôpital après être sortis de réanimation et 13 meurent à J90

Autrement dit la mortalité en réanimation est de 39% (194/497), puis la mortalité entre l’entrée en réanimation et la sortie de l’hôpital est de 56% (194/497 + 85/497) et enfin la mortalité entre l’entrée en réanimation et J90 est de 59% (194/497 + 85/497 + 13/497). La mortalité est additive, c’est-à-dire qu’elle se cumule au cours du temps (on prend en compte les décès précédents plus ceux qui surviennent entre 2 étapes.). Et comme on s’intéresse en médecine à la survie on dira que la Survie en réanimation est de 61% car si il y a 39% de morts et il y a 61% de vivants ; tout ça pour dire que **patients morts + patients vivants = 100%**

On voit bien que les chiffres évoluent au cours du temps. Les courbes de survies sont **Dynamiques** et non pas **statiques**.

Cependant ces pourcentages de survie ou de mortalité s’interprètent comme des mesures de **prévalence**. Elles **quantifient** le problème dans une population supposée **homogène**

Les analyses de survie **ignore le temps de survenue des évènements** et **suppose un temps de suivi identique** pour tous les sujets.

Quand on a mesuré les dates de décès on peut calculer un « délai de survie » : T qui est une variable aléatoire positive : T ≥ 0

T mesure un délai de survenue entre 2 évènements, c’est un **délai de transition entre deux états.** Il est nécessaire de disposer d’un suivi (avec toutes les dates).

La loi de probabilité de T est de :

-Si T est une variable aléatoire discrète : f(t) = P(T=t)

-Si T est une variable aléatoire continue c’est une densité de probabilité :



**III- Mesurer la survie**

On peut mesurer la survie via 3 types de fonctions :

**A – La fonction de mortalité**

Si on étudie la fonction de répartition F(t) qui représente la fraction d’individu ayant présenté l’évènement avant ou en t :

 

Ceci est la fonction de répartition F on voit qu’elle est croissante, F(0) =0 ce qui signifie qu’à l’instant t=0 personne ne présent l’évènement étudié et aussi que plus t augmente plus la courbe tend vers 1 (autrement dit plus le temps passe, plus les individus vont présenter l’évènement.)

Cette courbe peut être une courbe de mortalité : après un t très long, tous les sujets sont morts.

Petite formule cadeau pour Calculer la fonction de répartition à un instant t.

**B- La fonction de survie**

On étudiera les courbes de survie avec la **fonction de survie S(t) : S(t) = 1 – F(t) = P(t) > T**Une courbe de survie a cet aspect-là : S(0) = 1, S(t) est décroissante et S(t) temps vers 0 quand t est très grand. En pratique une courbe de survie ressemble à ça : 

Elle a un **aspect en escalier** car on prend **un t DISCRET** et non **pas continu**.

**C- Le risque instantané**

Le risque instantané est une fonction de densité conditionnelle, c’est-à-dire la probabilité de mourir sur un petit intervalle de temps sachant qu’au début on était vivant



Sachant que S(t) = P(t)>T alors 

Le risque instantané peut-être constant (survie des néons en électronique), croissant (leucémie au cours du temps) ou lognormal c’est-à-dire croissant ou décroissant au fil du temps.

**Remarque importante : Connaître une seule fonction : F, S ou h donne les autres (sachant qu’on connait déjà la fonction f)**

**IV & V– Estimer la survie et la prédire**

Pourquoi les courbes de survies sont particulières ? :

La distribution est très asymétrique : allure exponentielle, les courbes n’ont absolument pas une allure de courbe normale donc toutes les hypothèses et propriétés des lois Normales ne s’appliquent pas.

De plus certains délais d’apparition ne sont pas observés au cours de l’étude (tous les sujets ne sont pas forcément morts à la fin de l’étude) : on parle de **« CENSURE »** et plus particulièrement **de censure à droite** (par rapport à la courbe : la droite correspond à un t important). C’est ce qui biaise les estimations de la survie

*La prof a énormément insisté sur la censure au niveau des courbes de survies, il y aura au CC et aux partiels des Qcms dessus.*

La censure est la non survenue de l’évènement en fin d’observation : T (temps de survie ou de non apparition) > tc (Temps d’observation). On dit que le délai de survie est censuré à droite.

Il y existe deux mécanismes importants à connaître, responsable de cette censure à droite ;

**-Interruption de suivi (on parle de censure administrative) qui correspond à la fin de l’étude**

**-Le sujet a été perdu de vue : il ne s’est pas présenté aux rendez-vous médicaux etc…**

**Le mécanisme qui génère la censure doit être indépendant du mécanisme de mortalité (QCM ++++)**

Il existe 3 types de censure :

-Fixe : délai d’observation fixe pour tous les individus (mon étude durera X temps et j’arrêterai à t=X) on prévoit la date de la fin d’étude

-Séquentielle : on arrête l’étude au bout d’un certains nombres d’échecs (J’arrête l’étude quand 100 de mes 500 patients sont morts) 🡺 jamais réalisé en pratique médical

-Aléatoire : cela correspond au délai de censure : c’est une variable aléatoire indépendante de la survie qui fait qu’il y a interruption de suivi (c’est ce qui arrive le plus souvent lors d’essais thérapeutiques) 🡺 sujets perdus de vue quittant l’étude ou sujets arrivant pendant l’étude

Exemple de censure ALEATOIRE : c’est une étude qui s’arrête à T = 10 mois

Le sujet A meurt au cours de l’étude et est arrivé au début (pas de censure)

Le sujet B arrive à T=5 mois et fini l’étude à T=10 : il y a une censure aléatoire et une censure administrative.

Le sujet C est perdu de vue (disparition avant la fin de l’étude) on suppose que ça n’ait rien à voir avec la maladie.

Le sujet D présente également une censure administrative. (Une étoile signifie une censure)

Pour étudier ces données, on dresse un tableau de survie avec 2 colonnes : une pour le délai d’évènement et une deuxième pour l’apparition ou non de l’évènement.



**Retenir les censures à droite ++**

Remarque :

Il existe également des censures à gauche : T < Tc : concrètement je fais une étude sur le VIH j’ai 10 patients qui au début de l’étude ne sont pas positifs au test mais qui en réalité l’ont à cause du temps d’incubation.

Et des censures par intervalle : en gros t1 < T < t2 : concrètement je fais une étude puis un de mes patients tombe malade au début mais se guérit avant la fin.

Une donnée de survie est : « tout délai d’évènement en tout ou rien et tout délai de transition entre 2 états »

Grâce aux données de survie on peut estimer des fonctions de survie à partir d’échantillon et les interpréter. On peut également comparer ces fonctions pour 2 groupes. On peut aussi étudier l’influence d’autres variables sur la survie des patients (régression : *La régression ne sera pas du tout abordée dans ce cours, la prof a passé tout ce qui touchait à la régression car cela a été fait aux cours précédents*)

Les données de survie mesure l’incidence (mesure de la vitesse d’apparition de décès au cours du temps dans la population), pour les courbes de survie on mesure une incidence cumulée : décès survenus avant T

Pour estimer la fonction de survie chez les décédés on observe le délai de survie, chez les vivants le délai de survie est au moins supérieur à leur durée de participation, si la censure est indépendante de la mortalité des individus elle est dite non informative.

La seule manière d’estimer une courbe de survie est la méthode de **Kaplan et Meier :** elle consiste à estimer l’incidence cumulée (des décès/vivants.)

Si on effectue une estimation avec de la censure, l’estimation est évidemment biaisée. Cependant la méthode de KM permet de tenir compte des patients dont le délai de survie est censuré. On estime 1-F(t) (S(t)) à l’aide de fonctions de répartitions empiriques.
Dans la méthode de KM chaque patient contribue à l’estimation de S(t) tant qu’il est dans l’étude.

L'estimateur de Kaplan-Meier, également connu sous le nom de l’estimateur produit-limite, est un estimateur pour estimer la fonction de survie d’après des données de durée de vie. En recherche médicale, il est souvent utilisé pour mesurer la fraction de patients en vie pour une certaine durée après leur traitement

L’estimateur de Kaplan-Meier est l’[estimation du maximum de vraisemblance](http://fr.wikipedia.org/wiki/Maximum_de_vraisemblance) non-paramétrique de *S*(*t*). C’est un produit de la forme :



Lorsqu’il n’y aucune censure, ni est le nombre de survivants juste avant le temps t.

Lorsqu’il y a censure, ni est le nombre de survivants moins le nombre de pertes (cas censurés). Ce sont seulement ces cas survivant qui continuent à être observés (qui n’ont pas encore été censurés) qui sont « à risque » de décès (observé).

Exemple ;

On prend 5 sujets A B C D et E, Le sujet E est resté 4 mois dans l’étude avant de décédé, le sujet A lui est resté 7 mois avant de mourir

Le sujet B est perdu de vue a T=5 mois et le sujet C à T=7mois

P (T>1) = 1 autrement dit à partir du mois n°1 tout le monde est en vie (S(1) = 1)

Idem pour T = 2 et T = 3

La formule pour raisonner sur un tableau de survie est la suivante :

P(T>X) = P(T>X I T>X-1) x P(T>X-1)

La probabilité d’être vivant à ,l’instant X est égale à la probabilité d’être vivant à l’instant X sachant que l’on était vivant à l’instant X-1 multiplié par la probabilité d’être vivant à l’instant X-1

A T=4 P(T>4) = P(T>4 I T>3) x P(T>3) = 4/5 x 1 = 4/5

A T=5 P(T>5) = P(T>5 I T>4) x (T>4) = 4/5 **Bien que le sujet B ait été perdu de vue, on ne le compte pas comme un sujet mort.**

A T=7 P(T>7) = P(T>7 I T>6) x (T>6) = 2/3 x 0.8 = 0.53

**On obtient ce résultat car à T=7 Le sujet A décède et le sujet C est perdu de vue cependant on ne prend pas en compte le cas C comme un décès mais comme un vivant (via l’estimateur de Kaplan Meier) donc à T=7 C et D sont vivants alors que A est mort sachant qu’à T=6 ils étaient encore vivants d’où le P(T>7 I T>6) = 2/3**

A T=10 P(T>10) = 0.53 car **aucun nouveau décès.**



On obtient une courbe avec cette allure la grâce à la méthode **de Kaplan Meier**

E correspond au décès du sujet E.

Lorsqu’un sujet vient à **être perdu de vue on trace un petit trait horizontal** sur la courbe **signifiant la censure**.

On voit **que B est censuré**, pui**s A et C correspondent au décès du sujet A et à la perte de vue du sujet C** (donc petit trait horizontal) et **D correspond à l’arrêt de l’étude : censure administrative** d’où le petit trait horizontal a T=10 Mois

On voit un **aspect en escalier**, **mais pourquoi la marche à T=7 parait plus grande qu’ àT=4 ?**

Car a T=4 on a eu 1 mort parmi 5 sujets donc courbe réduite de 20% on passe de 1 à 0.8 pour P(Survie) puis à T=7 on a eu 1 mort parmi 3 sujets ! 1 étant mort (Sujet A), l’autre étant perdu de vue (sujet C) donc on réduit la courbe de 33% P (Survie à T=7) = 0.8 x 2/3 = 0.53. Donc la courbe se réduit de 0.27 alors qu’à la première marche elle s’est réduite de 0.2

Voilà pourquoi la deuxième marche d’escalier est plus grande que la première

.

**Voilà l’aspect d’une vraie courbe de Kaplan Meier avec de nombreux traits horizontaux synonymes de censure. Chaque marche indique un nouveau décès.**

*Attention (la prof a insisté ++) : Si on vous demande une estimation de la survie APRES l’arrêt de l’étude (censure administrative) on ne sait pas. Il est impossible d’estimer après le dernier suivi. Gros piège au QCM ++*

EXEMPLE 1 comparaison de 2 courbes de survie

Dans cet exemple on compare 2 types de traitements via des courbes de survies avec la méthode de **Kaplan Meier bien sûr** ! On voit clairement que la courbe « sédation » offre une meilleure survavibilité que l’autre « no sédation », à travers cet exemple la prof veut montrer que **ce qui nous intéresse quand on compare 2 courbes de survie, c’est que c’est la MEDIANE qui nous intéresse et pas la moyenne (QCM ++)**



EXEMPLE 2 Une courbe de survie qui chute :



Si une **courbe tombe à 0** soit **tous les sujets sont morts** (pas de censure) ou alors **que le dernier sujet de l’étude est mort** *(QCM +++++).* **L’aire sous la courbe de survie est la moyenne de survie**, **cependant l’estimation est biaisée si la dernière observation est censurée.**

Attention : Quand on **calcule la survie MEDIANE** (jamais la survie moyenne) il faut **le faire sur tous les sujets** autrement dit prendre en compte les morts et les vivants et pas que les morts.

Remarque de la prof : *La censure est l’observation incomplète d’une population complète tandis que la troncature est l’observation complète d’une population incomplète (sélection de sujet). Les estimations n’ont d’intérêt que dans des populations HOMOGENES*

Là on approche de la fin, et ce cours va commencer à partir en sucette : moult formules incompréhensibles vous seront présentées, il ne faut pas les apprendre juste **connaitre le TEST DU LOGRANK**, **qui permet de comparer mathématiquement deux courbes de Kaplan Meier.**

Le but du test LOGRANK est de **tester l’égalité des distributions de survie**, c’est un test **non paramétrique**, **aucune hypothèse à vérifier en dehors des censures**. La puissance du test est maximale si les risques de décès des deux groupes sont proportionnels au cours du temps.

**Le log-Rank est une généralisation du Chi2 :**

La formule du log Rank c’est ça (super hein ?) et à droite la formule développée



Pour finir la formule du log-Rank approché : 

Bon **concrètement**, il y aura pas de calcul à faire juste à savoir qu’on vous dira **« on effectue un log-Rank et on obtient P=une valeur » alors si votre p est inférieur à 5%** (p<0.05) **vous pourrez dire qu’un tel traitement marche mieux sur la survie qu’un autre car on rejette l’hypothèse nulle : que les deux traitements offrent la même courbe de survie.**

Les **limites** de ce test sont : que c’est impossible **pour des comparaisons multiples** et qu’on **ne prend pas en compte les facteurs de confusion**.

Et donc pour éviter cela en pratique on utilise des **modèles de régression logistique** (on binarise les données) et généralement des **modèles de régression pour données de survie : les Cox.**

**Le modelé de Cox effectue une estimation d’un effet de traitement ajusté sur des variables pronostiques**, il explique le risque instantané de décès en fonction des covariables et suppose que l’effet du traitement est stable dans le temps

Petite formule : 

***Donc à partir de là ce cours par en live,* la dernière chose à savoir est que quand on parle de log-Rank et de Cox ou de Kaplan Meier on Part de HAZARD RATIO et non plus D’ODS RATIO**

**Même principe du HR que les OR Si HR>1 association positive, si <1 association négative et si =0 pas d’association**

*Apres la prof a fait les 20 dernières diapos en 2 minutes car elle a très bien vu qu’on comprenait rien et qu’après tout on était qu’en D1*



En gros si X est binaire Le hazard ratio ne dépend que du paramètre Béta (en fait d’exp(b))

Et si X est continu ça se complique : 

Après on peut utiliser différents modèles pour construire les Hazard ratio : 



Fin des formules incompréhensibles, si vous avez vraiment lu et essayé de comprendre ces formules, mes respects.

Dédicace à la team Chir Ortho Larib tonton Woudy, Le petit de Jeffreyyy et Chispoune**. 3 grands bosseurs dans l’âme**. Dédicace à ceux qui vont nous remplacer, vous inquiétez pas les externes savent déjà qu’ils vous verront une fois par semaine.

Sinon dédicace a la comère de la faculté, ce bon vieux **Antar** (jte kiffe ;)) et ses fameux 230 rattrapages en Urgences (T’as éffacé le message Fb mais on s’en souvient tous :p)

Dédicace à ceux resté en P2…

Dédicace aux potos D1: Jéjé, Quentin, Jeff, Rudy (futur D1 again), Saada, Shirley (Stop feed TY), Nora (black sistaa), Aless, Quentin, Malek (T’as droit un cœur BB <3), Thomas T.(qui collectionne les heures de travail et les rattrapages), Mathieu (le mec le plus endetté de la fac), Piccolo (on sfait une game quand tu veux), Raph, Clément, Anfony, Mahmoud, ma ronéolectrice qui a Kiffé ce cours avec moi, et à Cyrielle (ou Cyrille c’est la meme d’apres la scola), Mehdi, Andy, Alex, Gwen et Nounouuuurs (VP basket)

Bisous et travaillez bien ! ;)