

Atelier JGEM - REES
Hôpitaux universitaires Pitié Salpêtrière - 75013 PARIS
10 Novembre 2017

Les Nouvelles Métriques de l'Evaluation Economique sous EXCEL

Anastasiia Kabeshova, Robert Launois
RÉSEAU D'EVALUATION EN ECONOMIE DE LA SANTE

28, rue d'Assas
75006 Paris – France
Tel. 01 44 39 16 90 – Fax 01 44 39 16 92
E-mail : launois.reesfrance@wanadoo.fr – Web : www.rees-france.com



Feuille de route

2

- » Introduction à Microsimulation
- » Techniques d'ajustement paramétriques des courbes de survie publiées
- » Application des ajustements paramétriques des courbes de survie dans le modèle de Microsimulation (Monte-Carlo d'ordre 1)
- » Choix des distributions de probabilité pour les paramètres du modèle
- » Analyse de sensibilité probabiliste (Monte-Carlo d'ordre 2)

Quelques questions avant le départ:

3

- ❖ Quelle est votre rôle dans la vie du modèle?

- Je le construit;
- Je le reçois et je l'utilise;
- Je le reçois et j'essaie de le comprendre;
- Je ne travaille pas avec le modèles.

- ❖ Quelle est votre niveau de connaissance d'Excel?

- Je suis SuperMan en VBA;
- Je sait faire les macros simples sur VBA;
- J'utilise les fonctions de base sur les feuilles Excel;
- Je n'utilise pas les formules.

- ❖ Vous êtes là parce que:

- Vous voulez appliquer les nouvelles méthodes dès la fin du séminaire;
- Vous voulez comprendre qu'est-ce que font les autres;
- Simple curiosité;
- Cette journée permet de m'absenter au travail.

- ❖ Dans quelle domaine travaillez vous actuellement?

- Industrie pharmaceutique;
- Consulting;
- Academia;
- Sans emploi stable / étudiant(e) / autres.



MODELE DE MICROSIMULATION EN TEMPS CONTINU

Typologie des modèles

5

				A	B	C	D
				Modèle de cohorte agrégée		Modèle individus-centrés	
				Déterministe	Probabiliste	Markovien (Sans mémoire)	Non Markovien
1	Prise en compte du temps	Implicite	Sans interactions	Arbre de décision avec chainage arrière	Arbre de décision simulation Monte-Carlo d'ordre 2 (rare)	Arbre de décision simulation; Monte-Carlo d'ordre 1 (rare)	
2		Explicite		Modèle de Markov déterministe	Macrosimulation de Markov; Monte-Carlo d'ordre 2 analyse paramétrique MC2	Microsimulation de Markov; simulation par échantillonnage individuel. ISM, MC 1 Analyse stochastique discrète ou continue; (cf ci-dessous)	
3	Prise en compte du temps	Discret	Avec interactions	Modèle épidémiologique à compartiments, systèmes d'équations différentielles	Chaîne de Markov à temps; discret MC d'ordre 1 et 2	Microsimulation à temps discret des changements d'état individuels MC1	Simulation en temps discret orientée agents Modèle Multi-Agents
4		Continu			Chaîne de Markov à temps; continu MC d'ordre 1 & 2	Microsimulation à temps continu de la survenue des événements sur F(.) inverse†	Simulation en temps continu à événements discrets :SED/DES

MNM : modèle Non Markovien; MC : Monte-Carlo ; ISM : individual sampling model ; DES : Discrete event simulation . †Inverse Cumulative distribution function : Fréquence des contacts , probabilité d'être infecté et fonction inverse

* A Brennan et al. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 15: 1295–1310 (2006)

Introduction à la Microsimulation

6

Entité



Biographie:

- Facteurs de risqué;
- Maladies;
- ...

Population



- Population ouverte/fermée;
- Interactions;
- Facteurs externes:
 - Environnement;
 - Economie...

Analyse

Pour évaluer à court et à long terme:

- Santé;
 - Interventions;
 - technologies; ...
- de la population entière.

Modèle Multi État *et* Microsimulation

7

Modèle de transition Multi-états ou *State-Transition Model*:

- « C'est une approche intuitive, souple et transparente de la modélisation en termes d'analyse décisionnelle qui recouvre à la fois les simulations de cohorte avec les modèles de Markov et les Microsimulations individuelles (Monte Carlo d'ordre 1) »*.

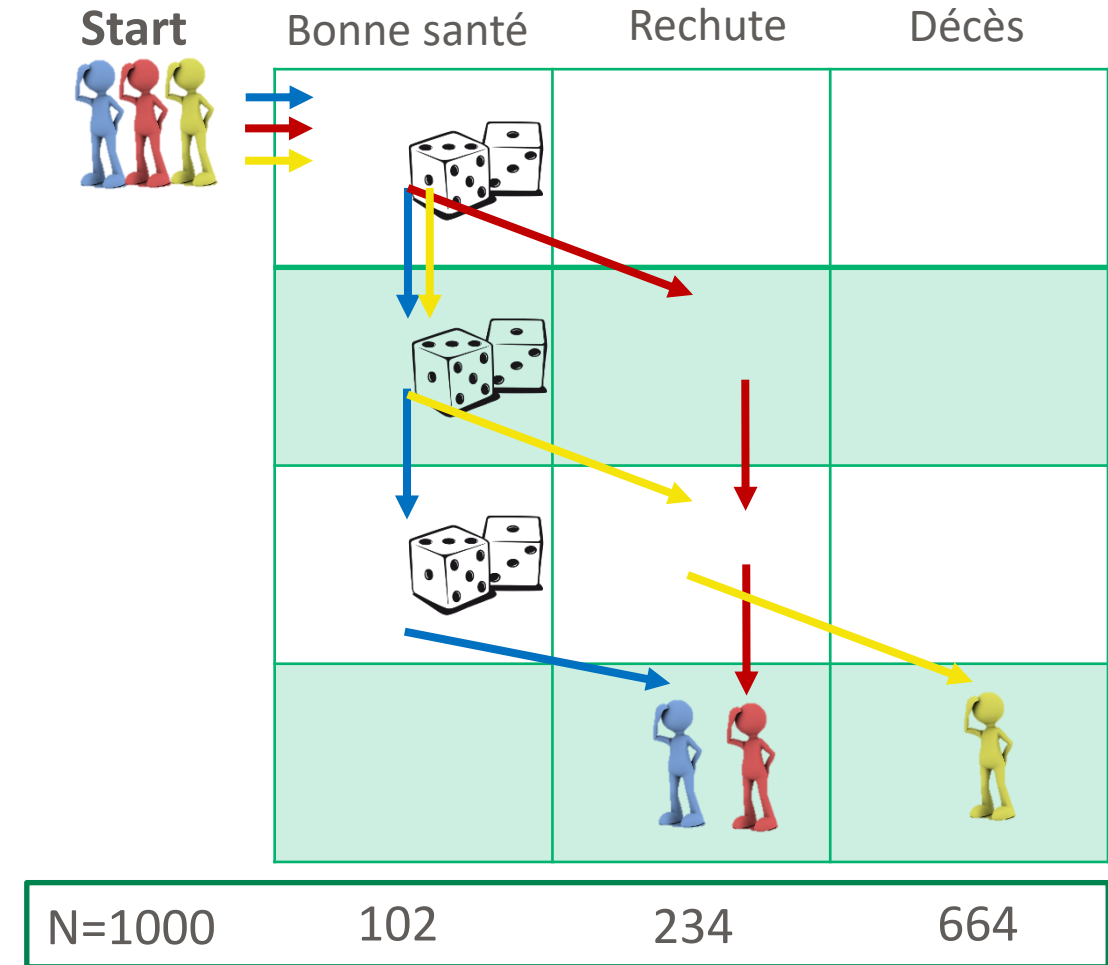
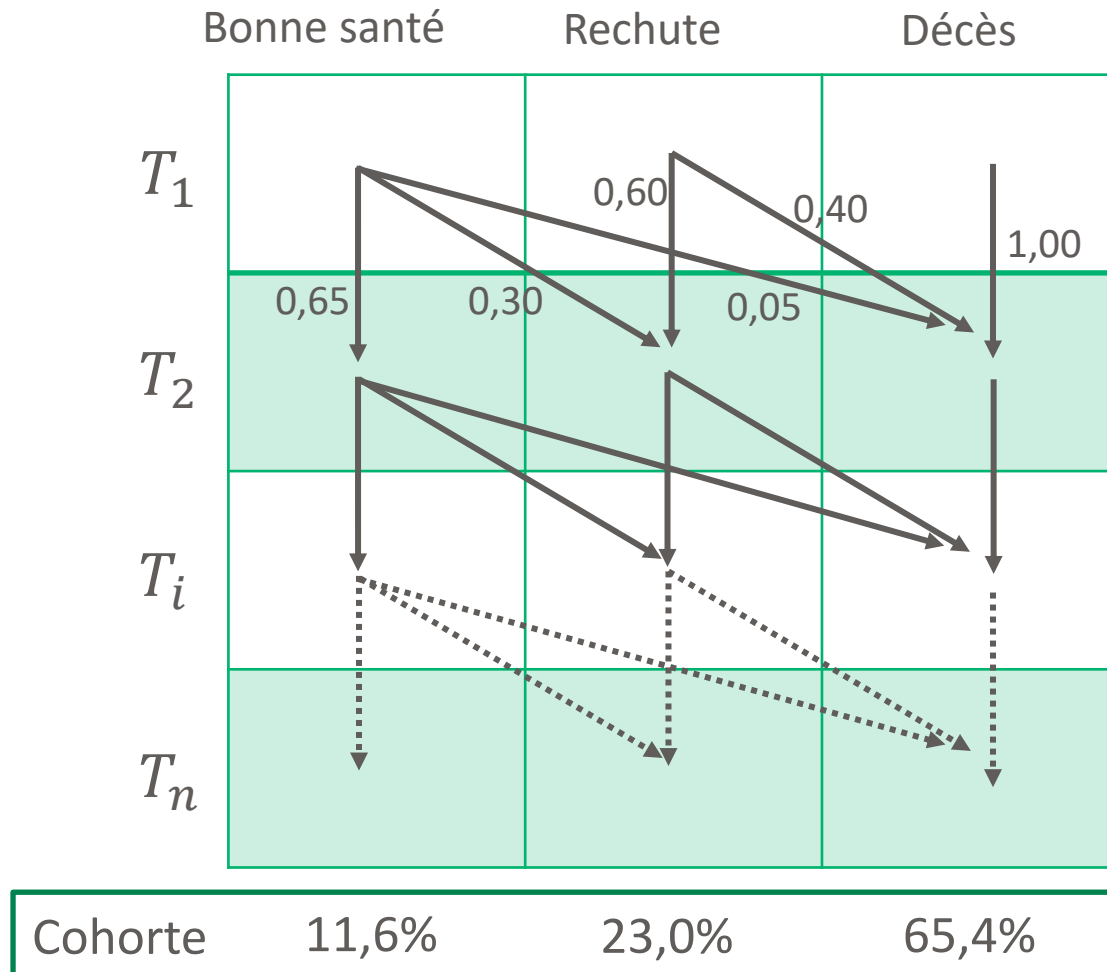
Microsimulation ou *Individual-Level State-Transition model* :

- Fonctionne au niveau de l'entité individuelle (virus, personne, famille, etc.) ;
- Simule une grande population représentative de ces entités;
- Et vise à tirer des conclusions sur les niveaux agrégés (population cible, pays, sous-population des patients)

* Siebert et al. 2012 *State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3*

Modèle de Cohorte vs Microsimulation

8



Modèle de Cohorte vs Microsimulation

9

	Modèle de cohorte	Microsimulation
Possibilités de développement du modèle	Supérieure	Inférieure
Facilité de débogage du modèle	Supérieure	Inférieure
Facilité de communication avec les non-experts	Supérieure	Inférieure
Hypothèse de Markov d'absence de mémoire	Oui	Non
Possibilités de modéliser des sous-groupes	Supérieure	Inférieure
Risque d'explosion du nombre d'états	Oui	Non
Présentation des résultats désagrégés (pas seulement des moyennes)	Possible, mais techniquement plus difficile	Oui
Prise en compte des antécédents des patients	Non	Oui
Logiciel disponible	Oui	Oui (besoin de connaissances avancées)

* Siebert et al. 2012 *State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3* (table 1)

Microsimulation en temps continu vs en temps discret

10

La microsimulation en temps continu résout **trois problèmes importants** de microsimulation en temps discret:

- ❖ Déterminer la **séquence des événements quand, pendant un intervalle, plusieurs événements peuvent se produire**. Dans la microsimulation en temps continu, le modèle de simulation détermine la séquence.
- ❖ Déterminer la **longueur des épisodes entre les événements**. En microsimulation à temps discret, les durées ne peuvent être déterminées qu'approximativement, alors qu'elles peuvent être déterminées avec précision en microsimulation en temps continu.
- ❖ Estimer le **nombre d'événements dans un intervalle**.

Modèle de Microsimulation en temps continu

11

Modèle de survenue des événements individuels en temps continu

L'unité de modélisation est l'individu

La variable d'intérêt est le temps jusqu'à l'événement

Le temps est une variable continue

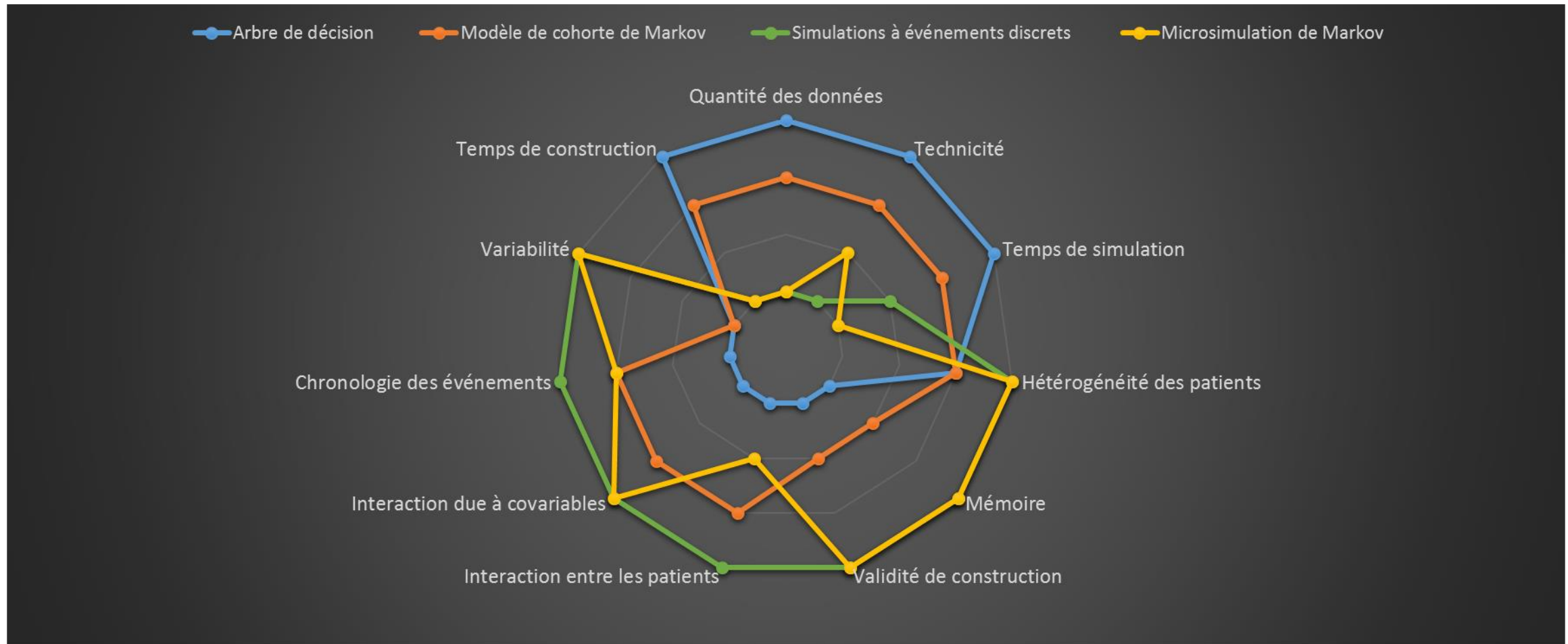
Les probabilités de transitions individuelles sont obtenues par tirage au sort en tirant un temps d'attente aléatoire à partir d'une distribution paramétrique

La distribution des valeurs est décrite par des **fonctions de survie**.

La **fonction de distribution inverse** (ou la fonction quantile) traduit une probabilité en nombre réel, alors que la fonction de survie le plus couramment utilisée traduit un nombre réel en probabilité.

Forces et faiblesses des différentes modélisations

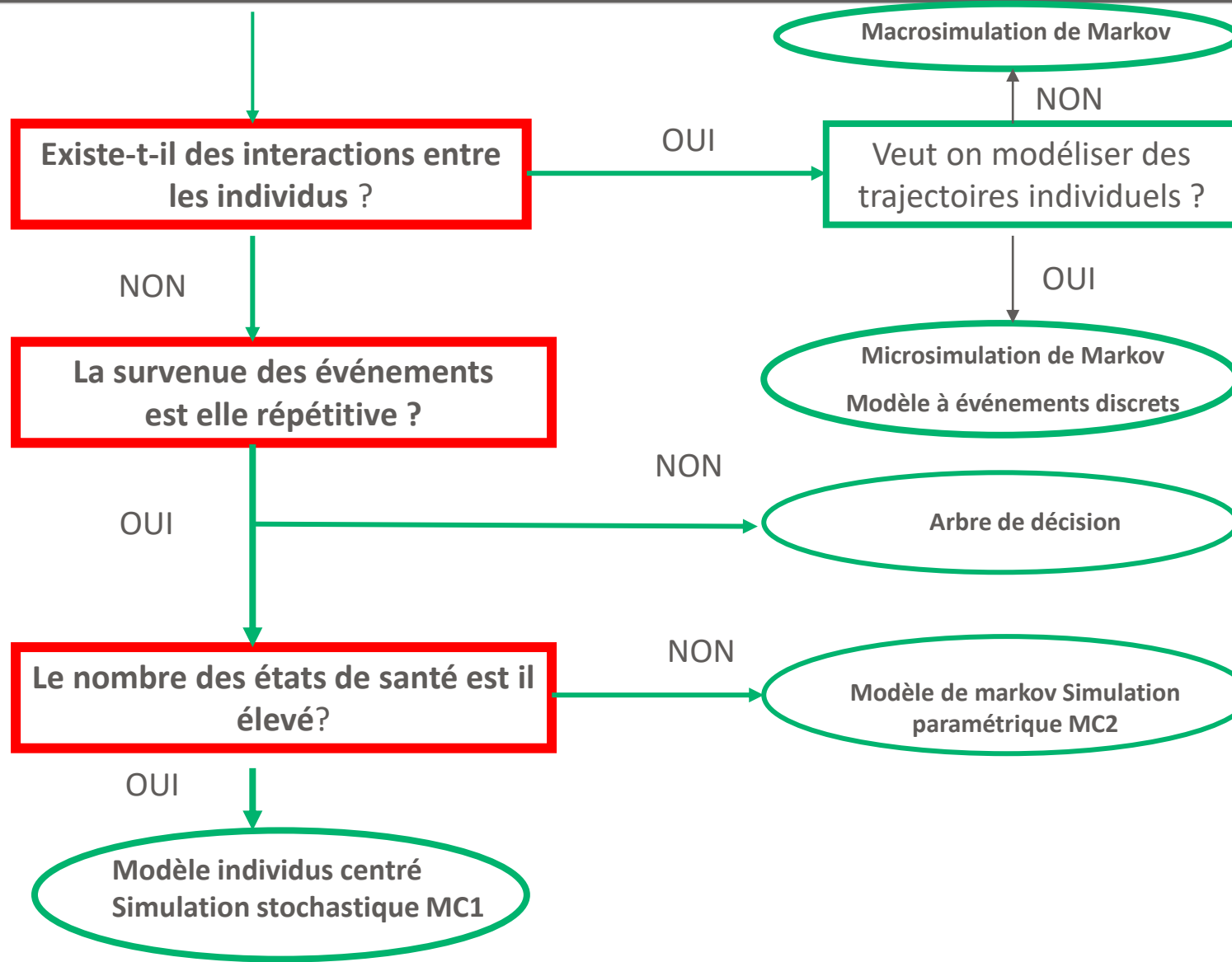
12



*Heeg BM1, Damen J, Buskens E, Caleo S, de Charro F, van Hout BA; Pharmacoeconomics. 2008;26(8):633-48.

Comment choisir ?

13



Choix du modèle : Gérer la Complexité

14

- Les modèles mathématiques sont de plus en plus utilisés pour éclairer les décisions en santé
- Tous les modèles requièrent des *hypothèses simplificatrices* => il y a des compromis entre la création du modèle le plus «réaliste» et ce qui repose sur des données mieux documentées
- **Les modèles complexes ne sont pas nécessairement plus précis ou fiables**
- À l'inverse, **la complexité peut être importante à inclure** lorsque des facteurs de confusion sont au centre d'un processus de maladie ou d'une question de recherche
- L'analyse de sensibilité sur l'**incertitude structurelle** en testant les structures de modèles alternatifs peut aider à déterminer les facteurs particulièrement importants

Solution?

15

- Les modèles doivent être suffisamment transparents pour que leurs résultats puissent être reproduits;
- Les modèles servent d'outils même lorsqu'ils sont de simples représentations du monde réel;
- De nouvelles techniques peuvent nous aider à trouver le juste équilibre entre la parcimonie et le réalisme de manière objective en utilisant les données pour construire le modèle à partir de la meilleure information disponible pour toute question donnée.

As Albert Einstein stated:

“Everything should be made as simple as possible, but not simpler.”

Variabilité – Incertitude - Hétérogénéité

16

Type d'analyse	Nature de l'incertitude	Autre terminologie	Analogie avec la régression
<i>Analyse stochastique</i>	Incertain par essence : variation purement aléatoire entre les mesures successives d'un même phénomène, →irréductible	Variabilité intrinsèque. Erreur de Monte Carlo. Incertitude d'ordre 1 ou à une dimension 1D	Erreur aléatoire
<i>Analyse paramétrique</i>	<i>Incertain sur la valeur des paramètres</i> incertitude par ignorance, →réductible par recherche nouvelles données	Incertain du second ordre ou à deux dimensions 2D	Erreur standard de l'estimateur
<i>Analyse de l'hétérogénéité</i>	Variabilité interindividuelle attribuable à la diversité des tableaux cliniques, →réductible par stratification	Variabilité. expliquée	Coefficients beta (la valeur de la variable dépendante change selon les patients)
<i>Analyse structurelle</i>	Bien fondé des hypothèses sur la structure du modèle →réductible par mélange de modélisations	Incertain sur le construit de la modélisation	Forme du modèle de régression (linéaire, log linéaire).

* A Briggs, 2012. *Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6*

TECHNIQUES D'AJUSTEMENT PARAMÉTRIQUES DES COURBES DE SURVIE PUBLIÉES



PRINCIPES DES MODÈLES DE DURÉE

Analyse de survie

19

- Les méthodes les plus souvent utilisées pour analyser les données du temps jusqu'à l'événement :
 1. **Kaplan-Meier + Test de Log-Rank / Wilcoxon** : Produit une estimation empirique de la distribution du temps jusqu'à l'événement et compare les groupes
 2. **Régression de Cox (risque proportionnel)** : Mesure l'effet de plusieurs covariables sans modéliser la distribution sous-jacente, en supposant des risques proportionnels entre les niveaux de la variables dépendante
- Aucune de ces méthodes donne de l'estimation de la forme fonctionnelle de la distribution sous-jacente

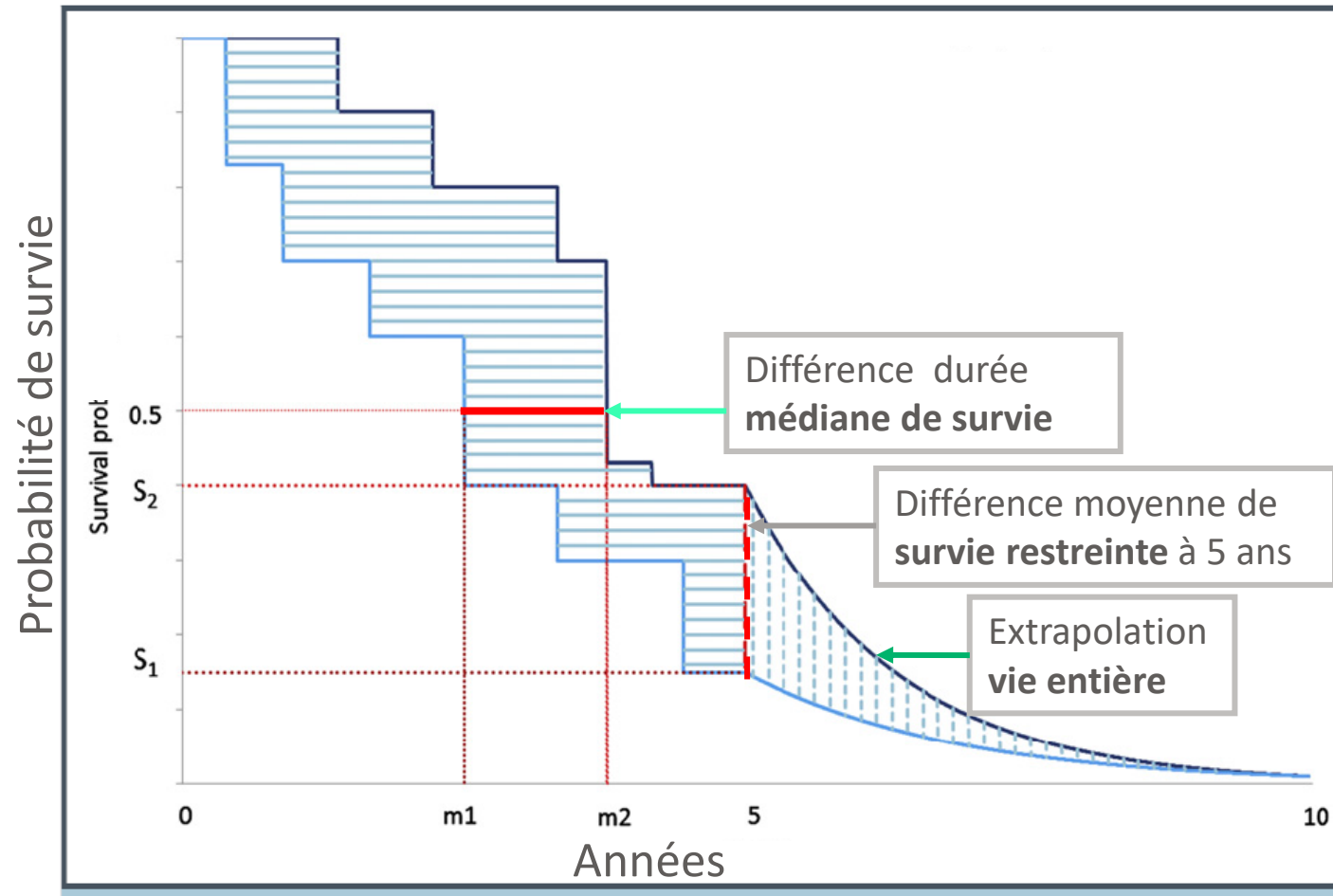
Méthode d'Estimation de Kaplan-Meier

20

- La méthode de Kaplan-Meier (KM) est utilisée pour estimer à partir des données individuelles patients (DIP) documenté dans un ERC **la proportion des patients initialement inclus (S) qui survivent à un instant donné (t)** conditionnellement au fait qu'ils ont survécu juste avant cet instant.
- La méthode résume les DIP sous la forme d'une série d'intervalles de temps débutant à l'instant où un événement se produit: $[0, t_1), [t_1, t_2), \dots, [t_r, \infty)$
- Pour chaque intervalle de temps $m = 1, 2, \dots, r$:
 - $n_{m+1} = n_m - d_m - c_m$
 - $S^{KM}(t_m) = \prod_{j=1}^m \frac{n_j - d_j}{n_j} = S^{KM}(t_{m-1}) * \frac{n_m - d_m}{n_m}$
 - n_m - le nombre de patients à risque juste avant le début de l'intervalle
 - d_m - le nombre d'événements survenant au début de l'intervalle;
 - c_m - le nombre d'individus censurés au cours de l'intervalle;
 - n_{m+1} - le nombre de survivants exposés au risque de décéder juste avant l'instant $m+1$

Extrapolation survie vie entière et calcul des différences

21



Analyse de Survie

22

- Fonction de risque instantané $h(t)$ est le risque que l'événement se produise au cours d'un intervalle de temps $[t; t+\Delta]$ sachant que jusqu'à t cet événement n'a pas eu lieu :

$$h(t) = \frac{P[t \leq T < t + \Delta | T > t]}{\Delta} \text{ pour } \Delta \rightarrow 0$$

- Quand le temps passe, le risque augmente => Fonction de risque cumulative :

$$H(t) = \int_0^t h(u) du$$

- Fonction de survie est la probabilité que l'événement ne se produise pas avant le temps t :

$$S(t) = P[T > t] = e^{-H(t)}$$

- Probabilité de transition, ou la probabilité d'occurrence d'un événement dans un cycle de modèle :

$$p(t) = 1 - \frac{S(t)}{S(t-1)} = 1 - e^{H(t-1)-H(t)}$$



RECONSTRUCTION DES DONNÉES INDIVIDUELLES BASÉE SUR DES COURBES DE KM

Etapes de la reconstruction des données individuelles

24

1

- Extraction des coordonnées des courbes de KM publiées

2

- Vérification de la cohérence des extractions

3

- Groupement des points numérisés par intervalles du temps

4

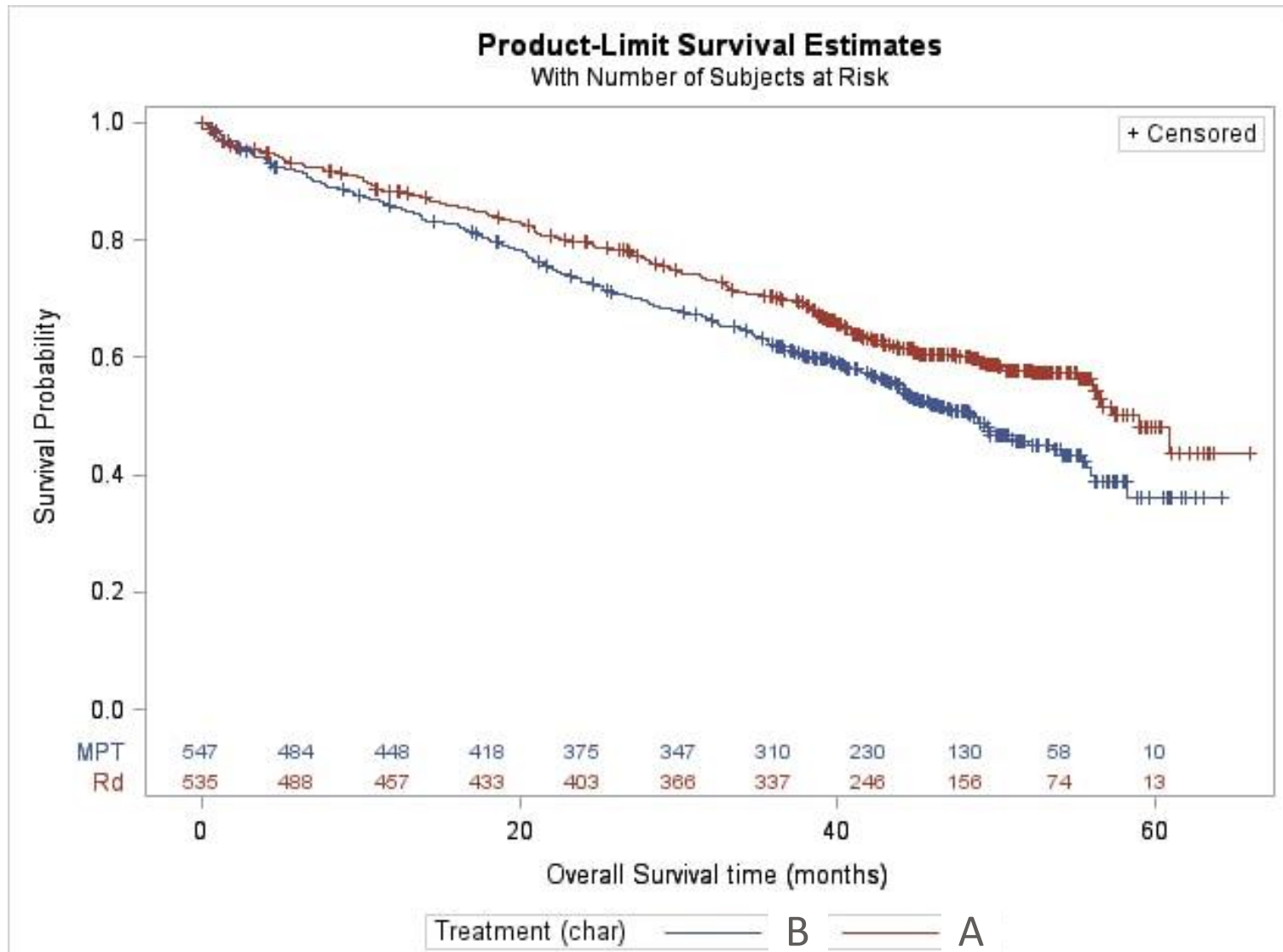
- Reconstruction des données initiales des patients sous R (algorithme de Guyot *)

* Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol.* 2012;12:9.

Courbe de Kaplan-Meier :

Survie Globale en Myélome multiple

25



Logiciels de numérisation des courbes de survie

26

	PLATFORM	COST	AUTOMATIC DETECTION	FILES SUPPORTED	POST ACQUISITION ANALYSIS	YEAR
Dagra: Digitize graphical data	Windows	\$49.95	yes	~ all image formats	no	2012
DataThief	Windows, MacOS, Unix	\$25	no	JPG, PNG, GIF	no	2006
dcsDigitiser	Windows	\$423	yes	~ all image formats	yes	2015
Digitizelt	Windows, MacOS, Unix	\$49	yes	~ all image formats	no	2014
Engauge	Windows, MacOS, Unix	Free	yes	~ all image formats	no	2015
g3data	Windows	Free	no	~ all image formats	no	2011
Get Data	Windows	Free	yes	~ all image formats	no	2013
Graph Click	MacOS	Free	yes	~ all image formats	no	2014
im2graph	Windows, Linux	Free	yes	~ all image formats	no	2015

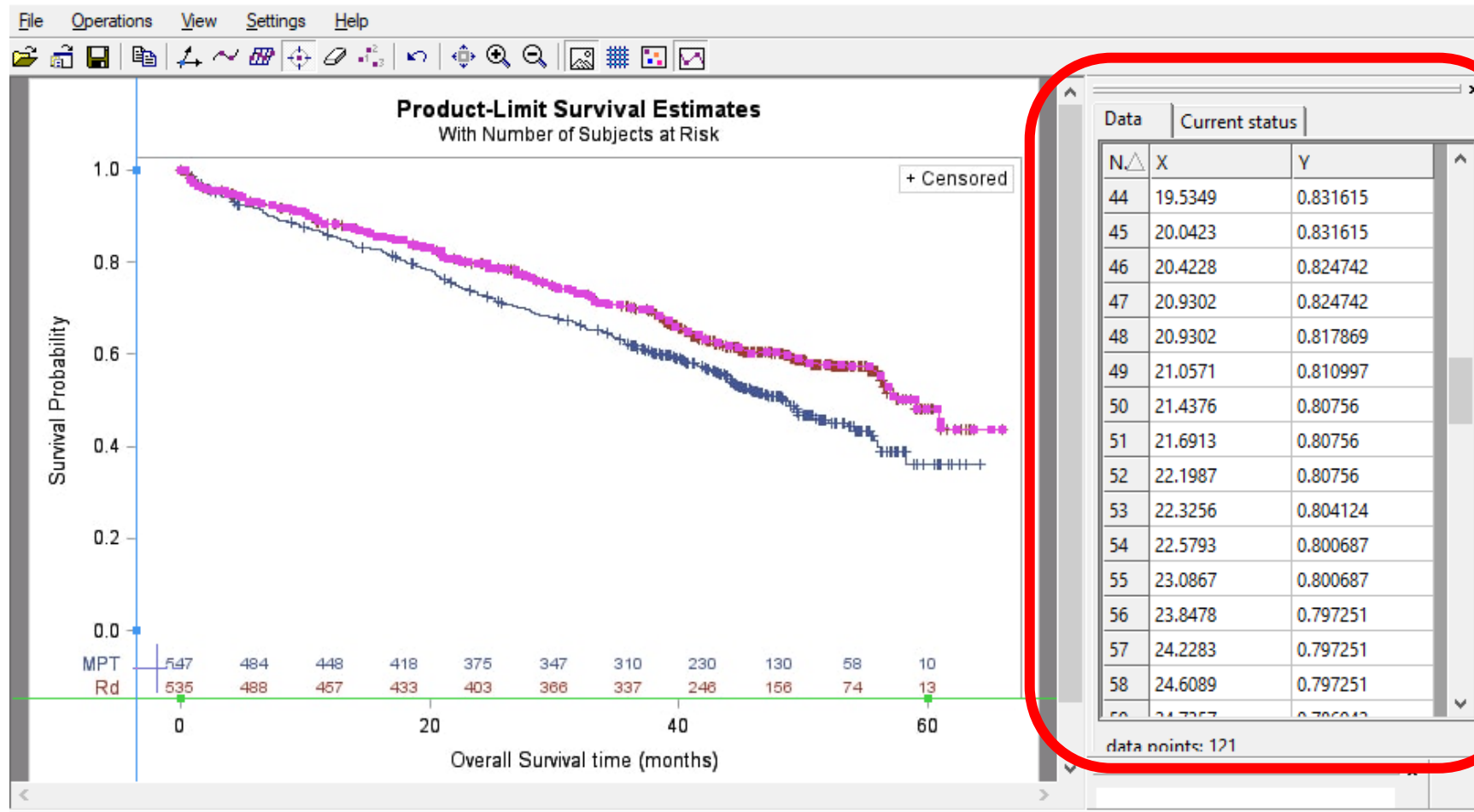
Tous les logiciels suivent un processus similaire pour convertir histogrammes, diagrammes de dispersion, et tracés linéaires en séries de chiffres

1. Accéder au graphique
2. Régler l'échelle
3. Numériser les points de données
4. Exporter les données

* <http://connectedresearchers.com/graph-digitizer-comparison-16-ways-to-digitize-your-data/>

1. Numérisation de la courbe de survie avec GetData Graph Digitizer

27



Coordonnées extraites

2. Vérification de la cohérence des coordonnées extraites

- ☒ Les données individuelles (DIP) colligées sont extrapolées sous Excel.
- ☒ Il est important de vérifier que la survie diminue avec le temps pour éviter que des messages d'erreurs apparaissent à l'étape suivante de l'algorithme qui traite les données extraites.
- ☒ Il est également important d'exprimer les données de survie en termes de proportions de patients et non pas sous forme de pourcentages.

3. Groupement des points numérisés par intervalles du temps

29

Tableau 1

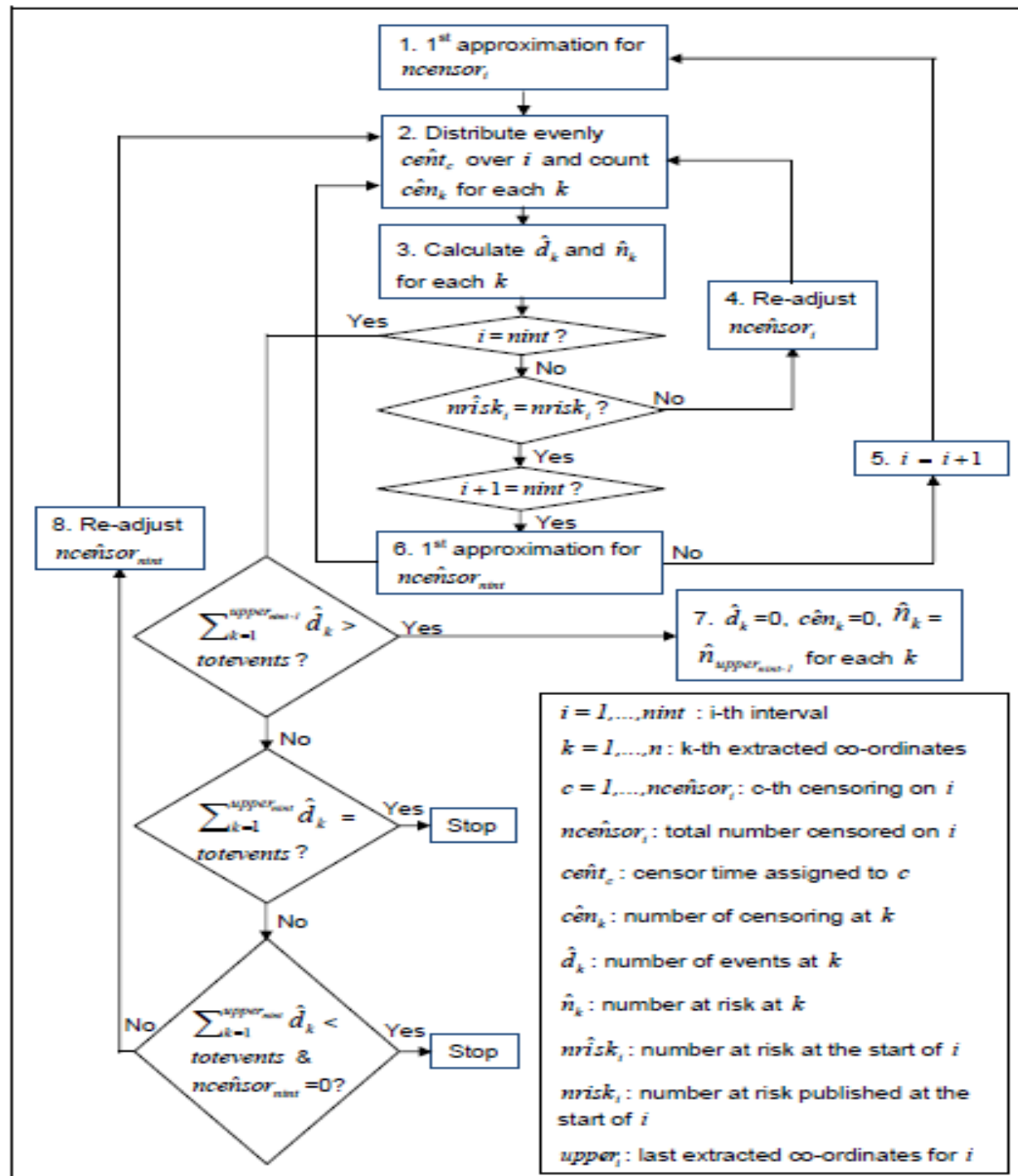
Rang coordonnées k	Temps (Mois) Tk	Proportion des survies Sk
1	0	1
2	0,26	0,9967
3	0,56	0,9843
4	0,75	0,9818
5	0,86	0,9777
6	0,97	0,974
7	1,61	0,9627
8	1,88	0,9581
9	2,03	0,9555
10	2,36	0,954
11	3,79	0,9524
12	4,09	0,9488
13	4,69	0,9432
14	4,88	0,9427
15	4,99	0,9401
16	5,33	0,9355
17	5,55	0,9314
18	5,85	0,9303
19	6,23	0,9288
20	6,46	0,9278
21	6,49	0,9252
22	6,61	0,9231
23	7,06	0,9226
24	7,39	0,9221
25	7,58	0,9206
26	7,81	0,9185
27	8,07	0,9164

Création d'une 2^{ème} série de données
composée des intervalles sur la période
de suivi de 5 ans (60 mois)

Tableau 2

Intervalle	Temps (mois)	Borne sup	Borne inf	#Pts à risque
1	0	1	18	535
2	6	19	57	488
3	12	58	215	457
4	18	216	418	433
5	24	419	589	403
6	30	590	782	366
7	36	783	820	337
8	42	821	842	246
9	48	842	863	156
10	54	864	896	74
11	60	897	920	13

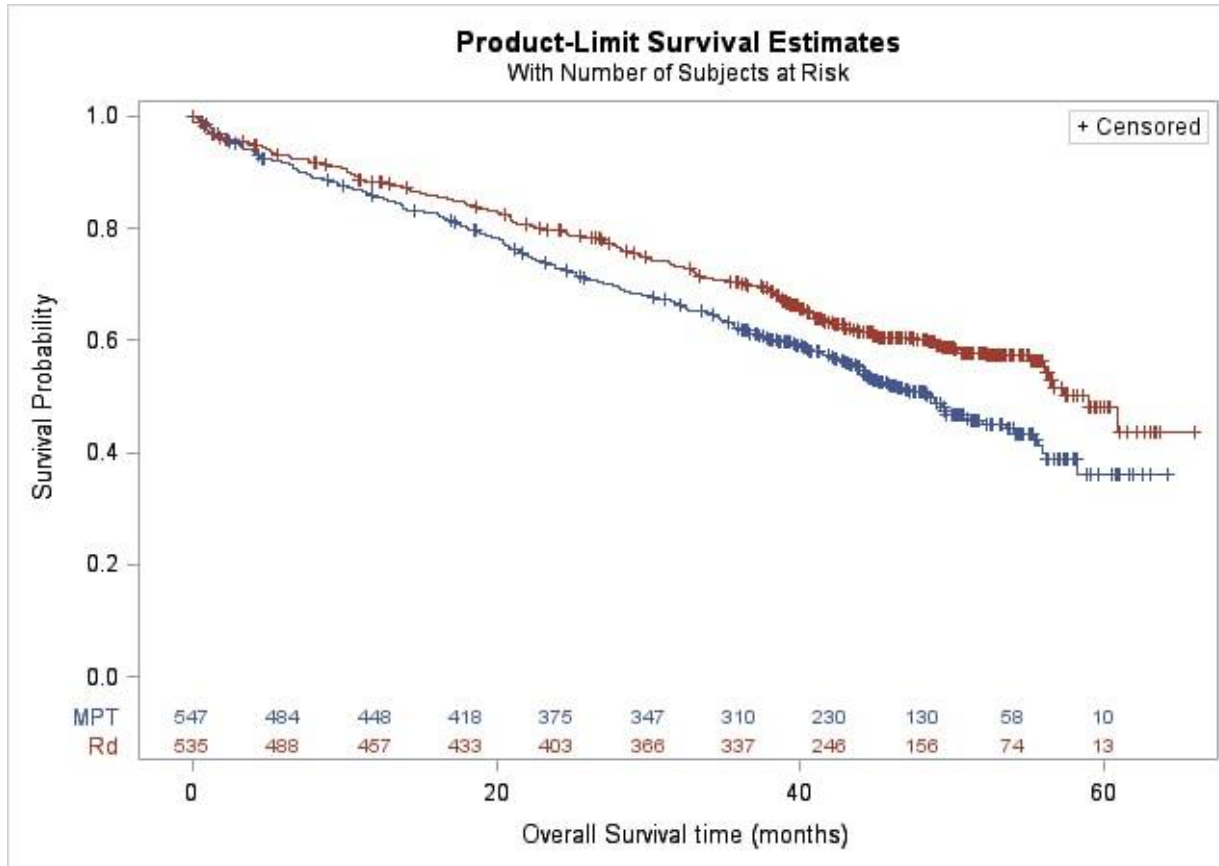
4. Flowchart de l'Algorithme de reconstruction de données KM sur R



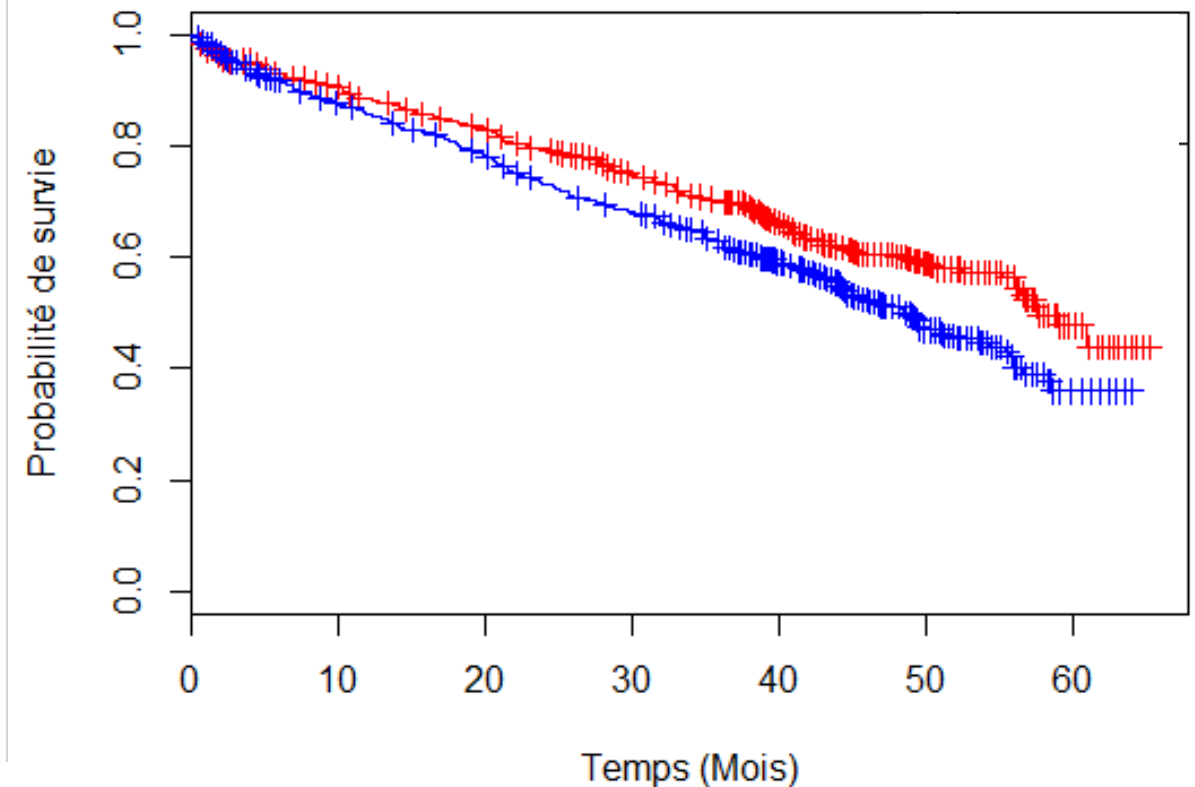
Courbes de Kaplan-Meier : Survie Globale

31

Initiales



Reconstituées





LES MODÈLES DE SURVIE À TEMPS CONTINU

Choix des Ajustements Paramétriques

33

- Lois de distributions utilisées pour les ajustements:
 - **Exponentiel**
 - **Weibull**
 - **Gompertz**
 - **Log-normal**
 - **Log-logistique**
 - **Modèles plus flexibles** (poly-Weibull model, spline-based modèle paramétrique, generalized gamma, generalized F-model, etc.)
- Impliquent généralement deux paramètres :
 - paramètre d'échelle (λ) et de forme (γ)
- Modèle de survie choisi doit être justifié avec les :
 - mesures statistiques (validité interne)
 - plausibilité clinique (validité externe)

Principales propriétés des fonctions de survie paramétriques

34

	Exponential	Weibull	Gompertz	Log-Logistic	Log-Normal
Risk pattern	Constant hazard	Monotonically increasing ($\gamma > 1$) or decreasing ($0 < \gamma < 1$) hazard	Monotonically increasing ($\gamma > 1$) or decreasing ($\gamma < 0$) hazard	Monotonically decreasing when $\gamma \leq 1$, or increasing followed by a gradually decreasing hazard when $\gamma > 1$	Log of the event time has a normal distribution
Type: PH vs. AFT	PH or AFT	PH or AFT	PH	PO or AFT	AFT
Survival function: $S(t) =$	$\exp(-\lambda t)$	$\exp(-\lambda t^\gamma)$	$\exp\left[\frac{-\exp(\lambda)(\exp(\gamma t)-1)}{\gamma}\right]$	$\frac{1}{1+\lambda t^\gamma}$	$1 - \phi\left(\frac{\log(t)-\lambda}{\gamma}\right)$ where ϕ is the cumulative standard normal distribution
Hazard function: $h(t) =$	λ	$\lambda \gamma t^{\gamma-1}$	$\exp(\exp(\lambda) + \gamma t)$	$\frac{\lambda \gamma t^{\gamma-1}}{1+\lambda t^\gamma}$	Hazard increases to a maximum and then decreases to 0 as time tends to infinity, and $h(0) = 0$
Linear function	$\log(S(t)) = -\lambda t$	$\log(-\log(S(t))) = \log(\lambda) + \gamma \log(t)$	$\log(h(t)) = \exp(\lambda) + \gamma t$	$-\log\left(\frac{\hat{S}(t)}{1-\hat{S}(t)}\right) = \log(\lambda) + \gamma \log(t)$	$\log(t) = \lambda + \frac{\phi^{-1}(1-\hat{S}(t))}{\gamma}$
Graphical test	$\log(\hat{S}(t))$ vs. t	$\log(-\log(\hat{S}(t)))$ vs. $\log(t)$	$\log(\hat{h}(t))$ vs. t	$\log\left(\frac{\hat{S}(t)}{1-\hat{S}(t)}\right)$ vs. $\log(t)$	$\phi^{-1}(1 - \hat{S}(t))$ vs. $\log(t)$
Regression function	PH: $\log(\lambda_i) = -X_i \beta$ AFT: $\log(\lambda_i) = X_i \beta$	PH: $\frac{-\log(\lambda_i)}{\gamma} = -X_i \beta$ AFT: $-\log(\lambda_i) = X_i \beta$	$\frac{-\log(\lambda_i)}{\gamma} = -X_i \beta$	PO: $\frac{-\log(\lambda_i)}{\gamma} = -X_i \beta$ AFT: $-\log(\lambda_i) = X_i \beta$	$\lambda = X_i \beta$

AFT accelerated failure time, PH proportional hazards, PO proportional odds, Log natural logarithmic function, Exp natural exponential function.

Further details on these distributions and their properties can be found in Collett

Distribution Exponentielle

35

La distribution exponentielle suppose un taux de risque constant au fil du temps :

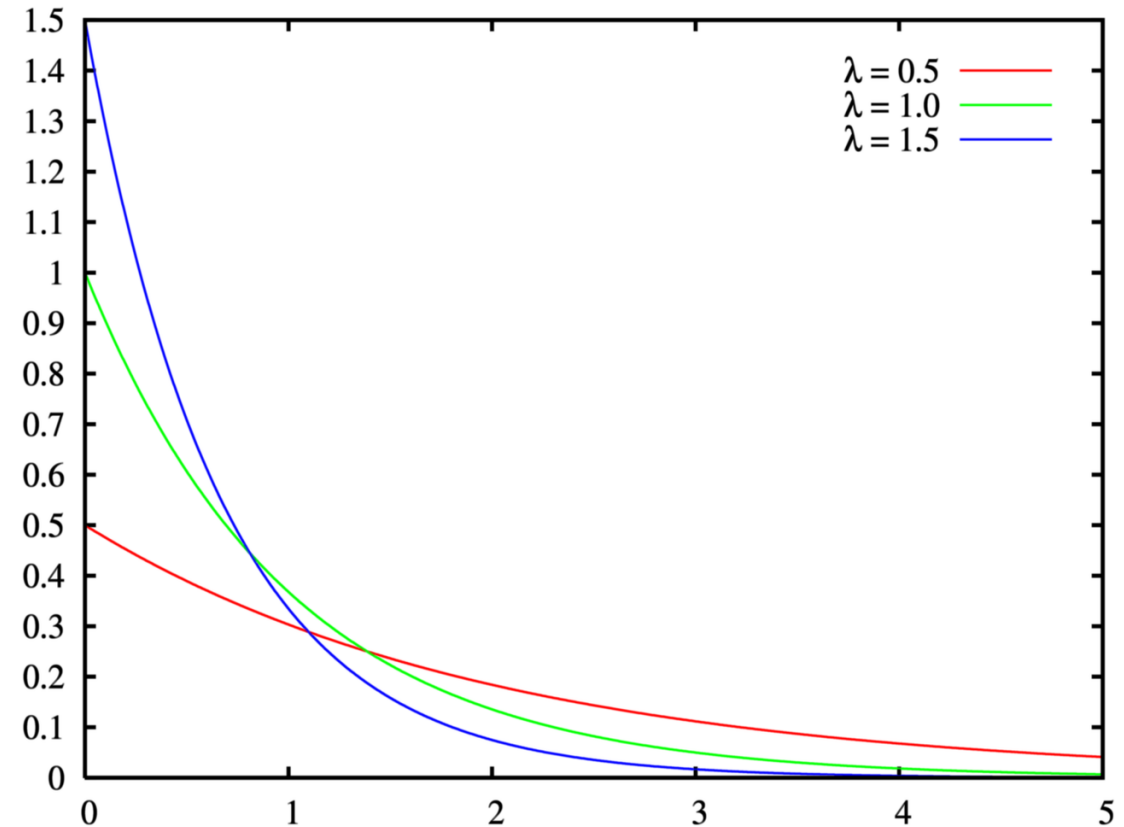
$$f(t) = \lambda e^{-\lambda t}$$

$$S(t) = e^{-\lambda t}$$

$$h(t) = \lambda, \lambda > 0$$

$$H(t) = \lambda t$$

Densité de probabilité



Distribution de Weibull

36

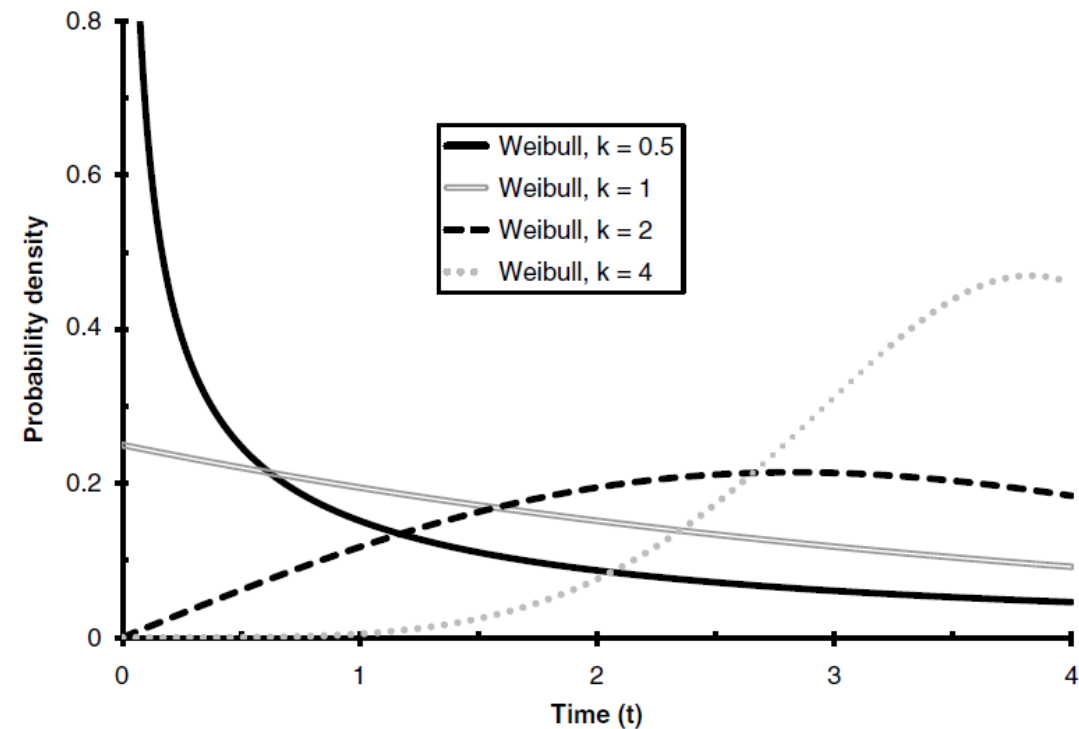
- $\lambda > 0$ paramètre de l'échelle
- $\gamma > 0$ paramètre de forme

$$S(t) = e^{-(\lambda t)^\gamma}$$
$$h(t) = \lambda \gamma (\lambda t)^{\gamma-1}$$

$\lambda \uparrow$, fonction de survie $\searrow$.

γ définit le changement du risque dans le temps :

- $\gamma = 1$, le risque est constant (équivalent à la fonction exponentielle)
- $\gamma > 1$, le risque augmente avec le temps
- $\gamma < 1$, le risque diminue avec le temps



Distribution de Gompertz

37

- $\lambda > 0$ paramètre de l'échelle, $\gamma > 0$ paramètre de la forme

$$S(t) = \exp\left\{-\frac{\lambda}{\gamma}(1 - e^{\gamma t})\right\}$$

$$h(t) = \lambda e^{\gamma t}$$

$$H(t) = -\frac{\lambda}{\gamma}(1 - e^{\gamma t})$$

$$tp(t, u) = 1 - \exp\left\{\frac{\lambda}{\gamma}(e^{\gamma(t-u)} - e^{\gamma t})\right\}$$

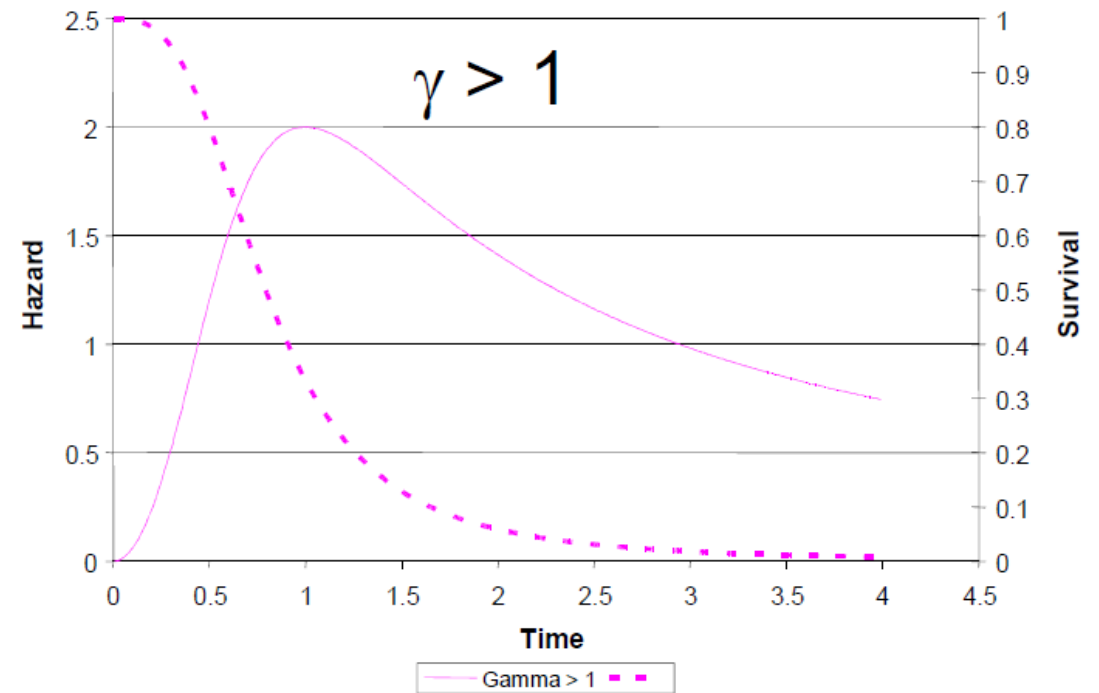
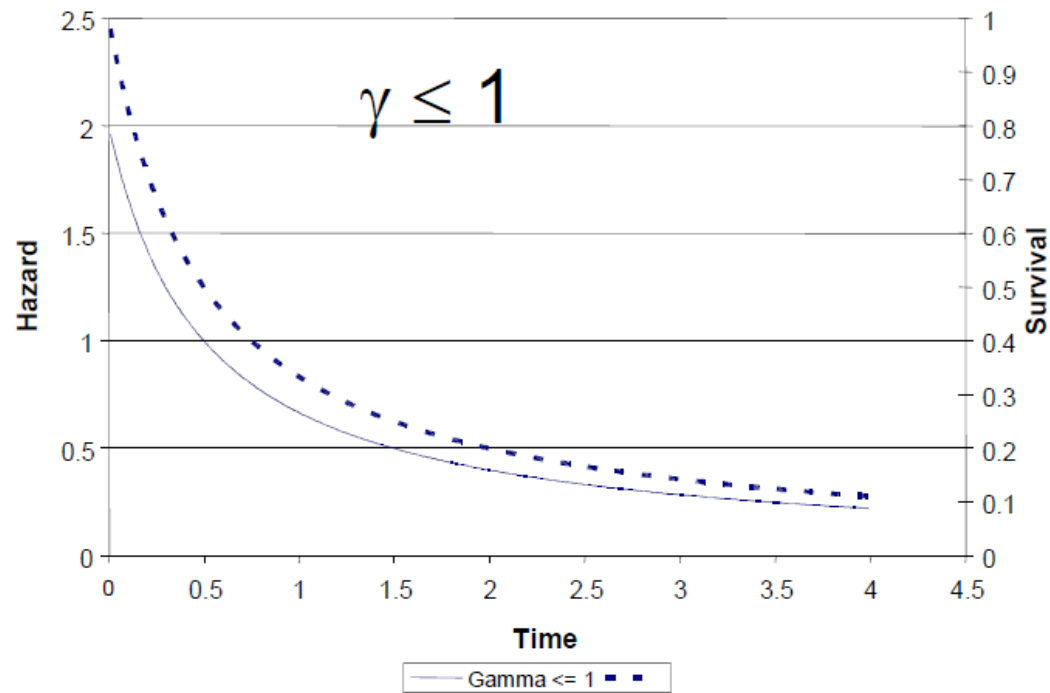
$tp(t, u) = 1 - \exp\left\{\frac{\eta\lambda}{\gamma}(e^{\gamma(t-u)} - e^{\gamma t})\right\}$ pour le bras d'intervention, où η est le rapport de risque

Distribution Log-Logistique

38

$$S(t) = \frac{1}{1 + \lambda t^\gamma}$$

$$h(t) = \frac{\lambda \gamma t^{\gamma-1}}{(1 + \lambda t^\gamma)^2}$$



Distribution Log-Normale

39

- Très similaire à la distribution log-logistique avec $\gamma > 1$

- Spécification :

$$\text{Log}(T) \sim \text{Normal}(\mu, \sigma)$$

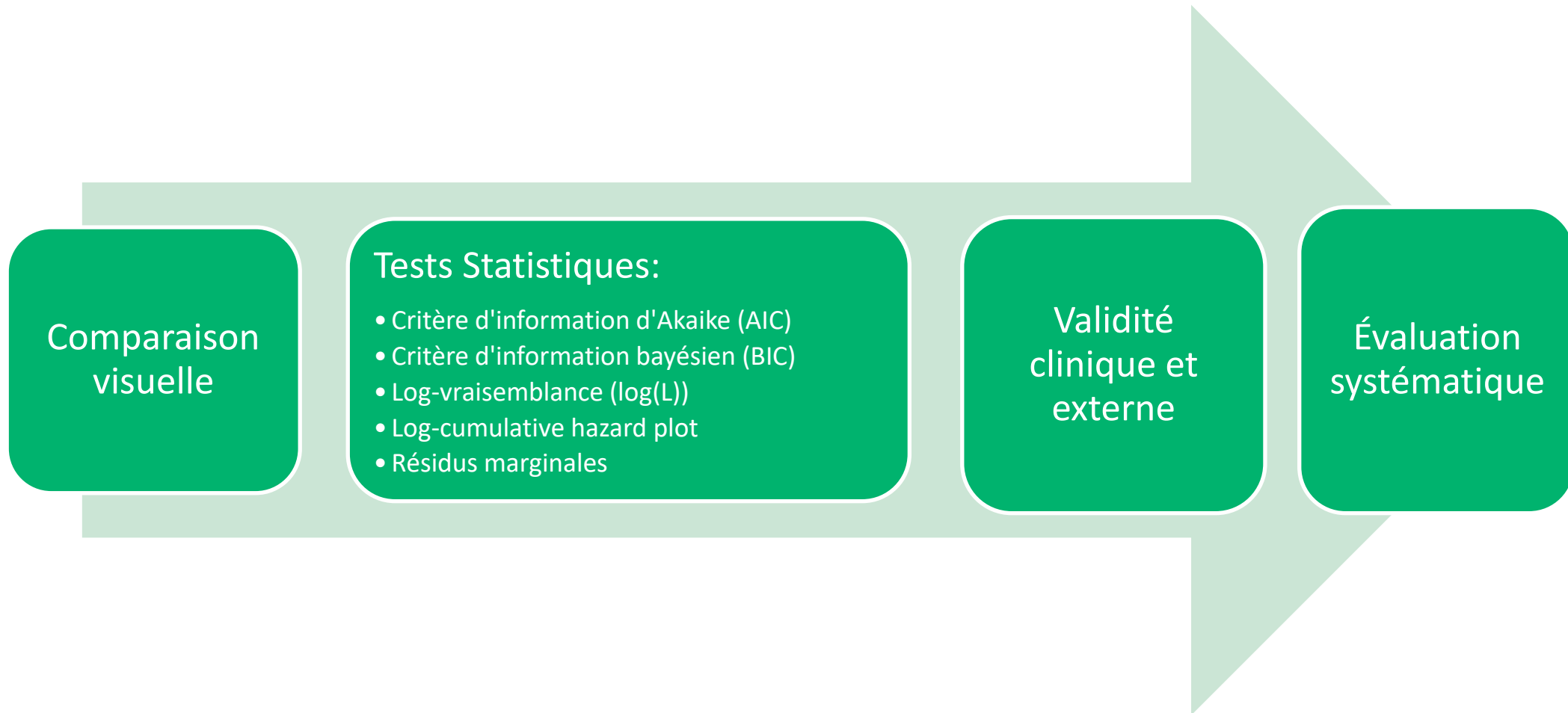
- Distribution de survie :

$$S(t) = 1 - \Phi\left[\frac{\log(t) - \mu}{\sigma}\right]$$

Où Φ est la distribution normale cumulative standard

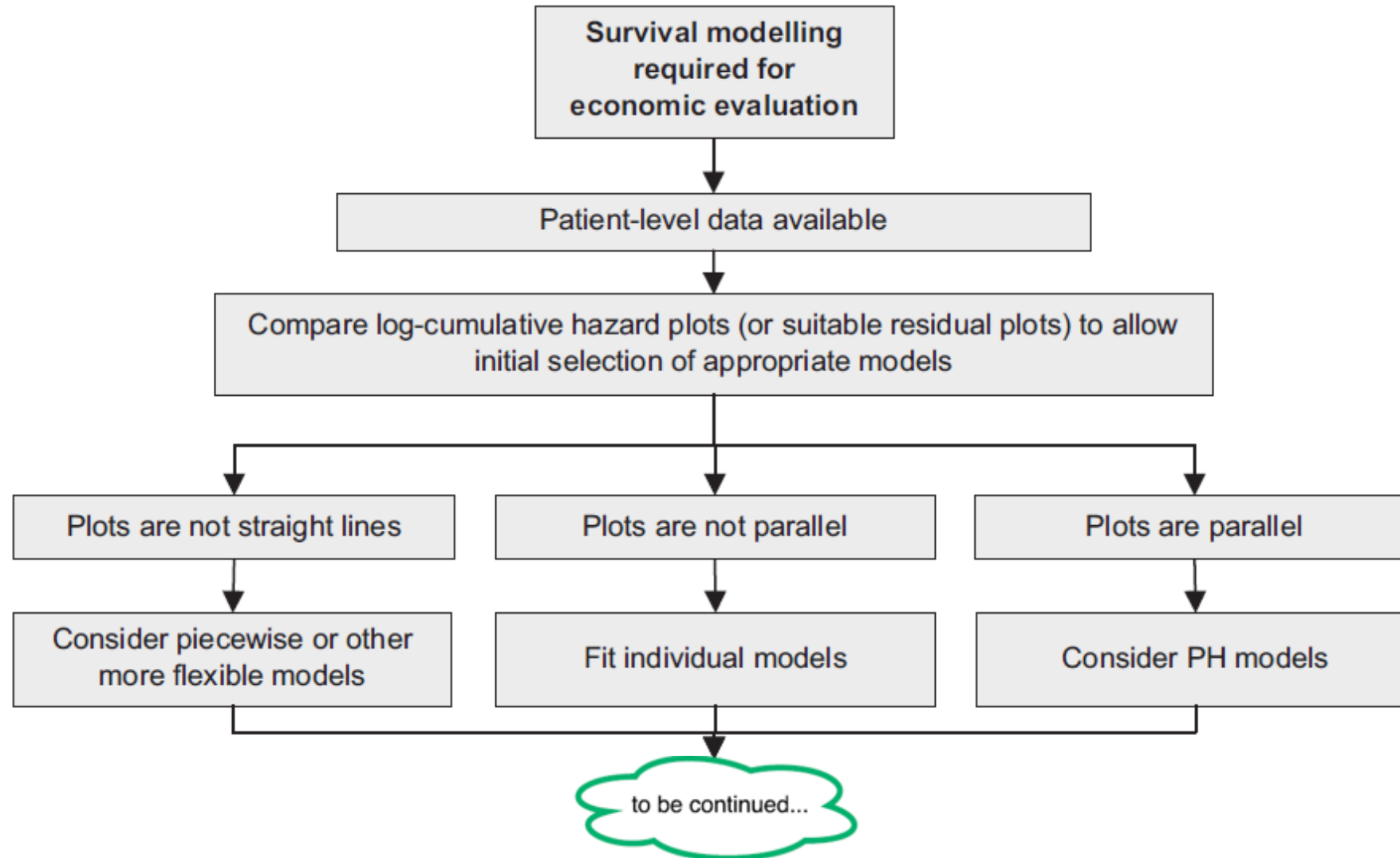
Choix de la distribution pour les données de survie

40



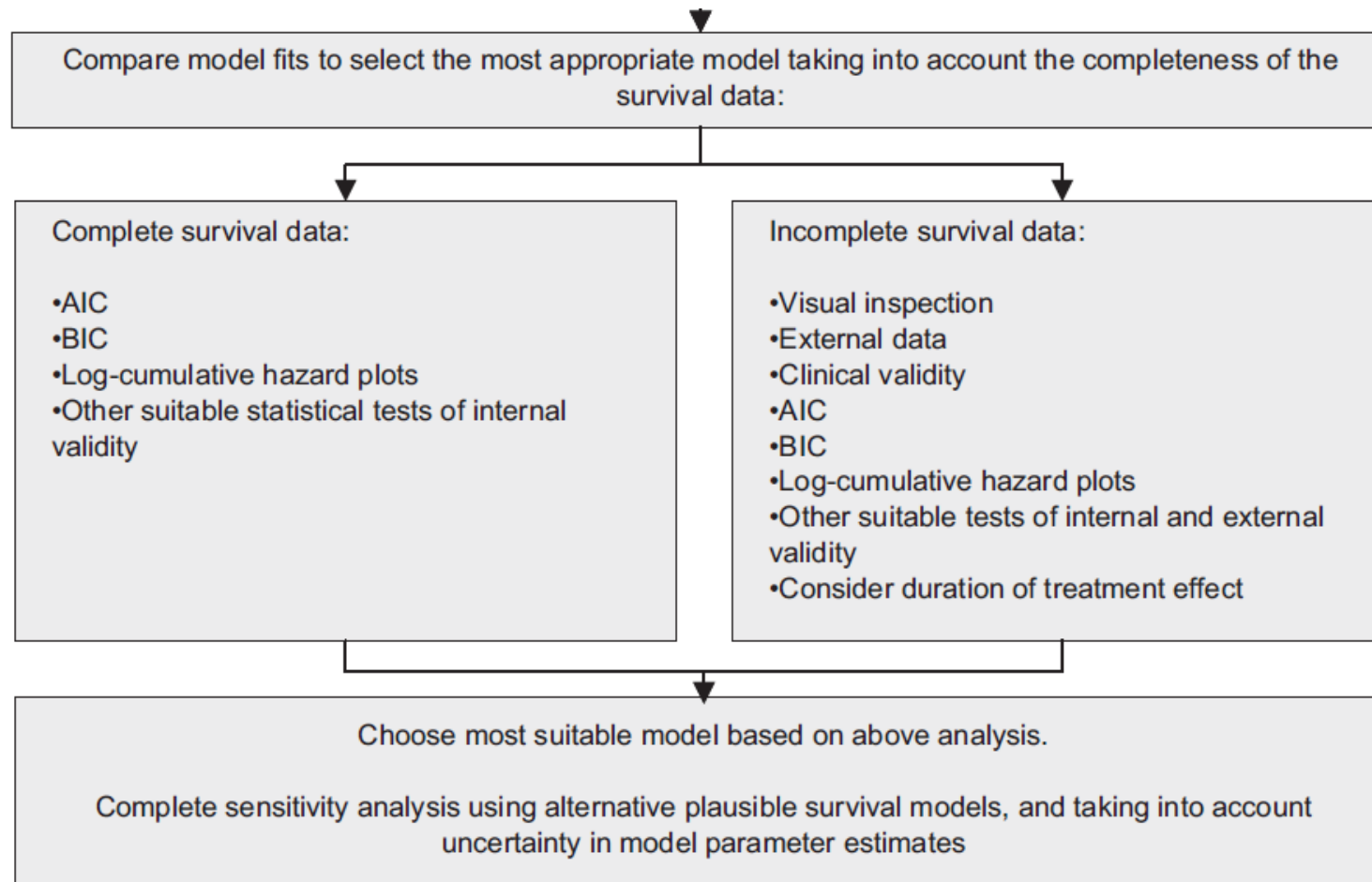
Algorithme de sélection du modèle de survie

41



Algorithme de sélection du modèle de survie

42

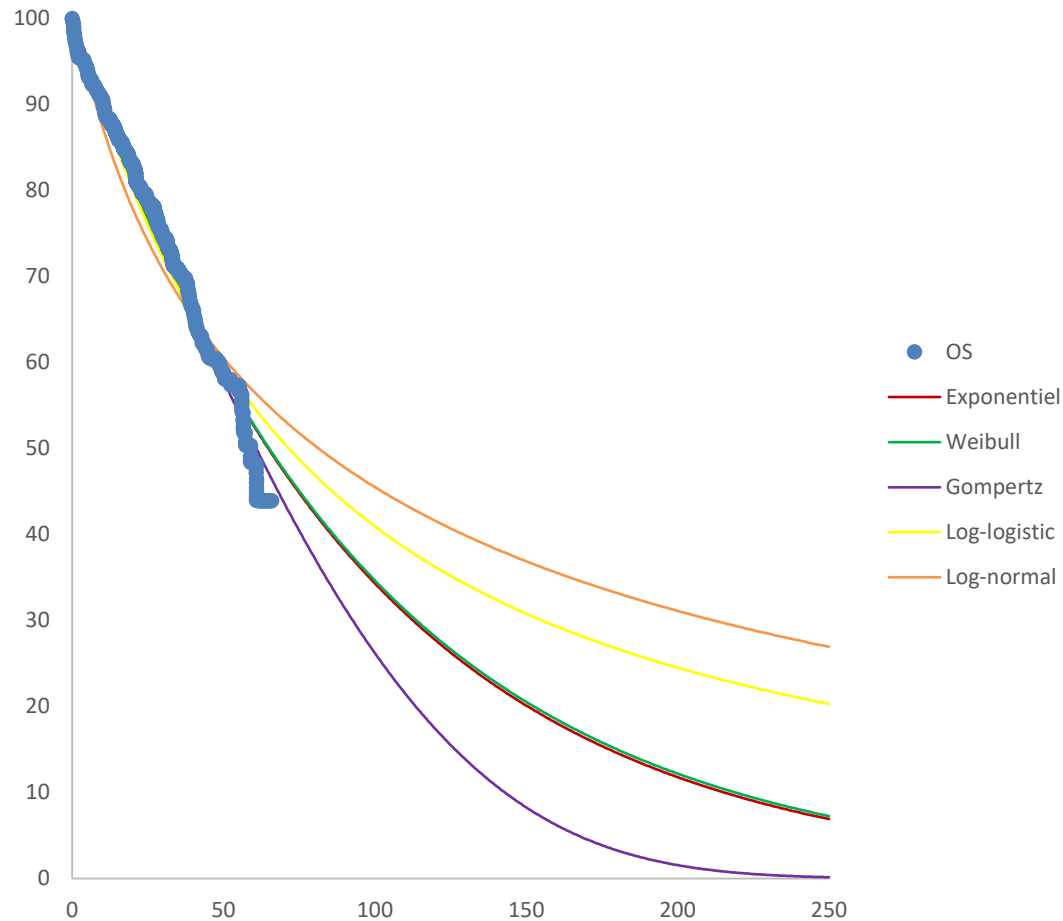


* Latimer et al., 2013, *Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide.*

Exercice 1

Ajustement paramétrique de courbe de survie

44

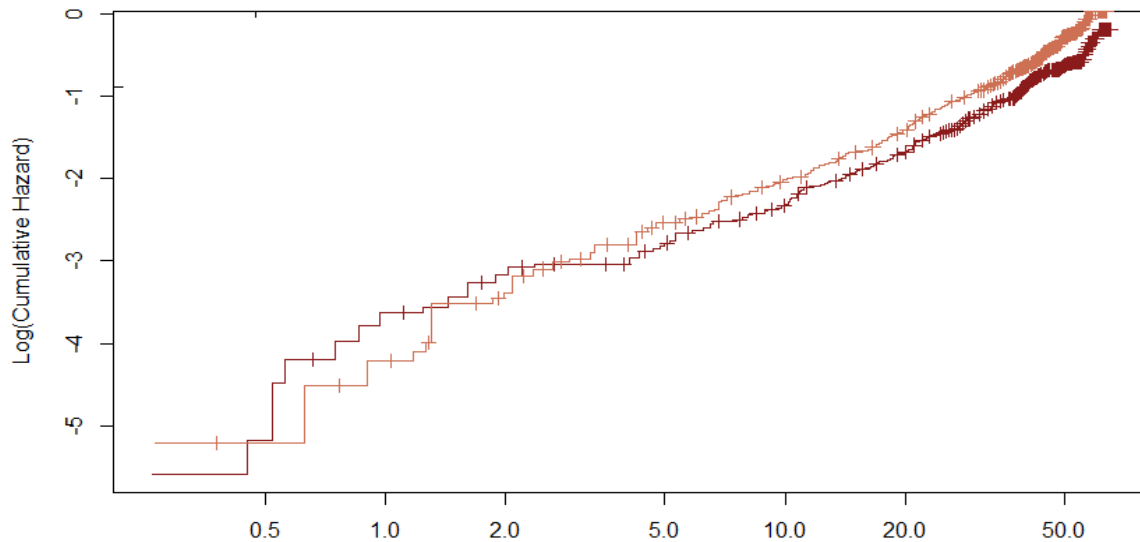


Loi de distribution	AIC	Log-Vraisemblance	Nb paramètres
Exponentiel	2305,79	-1151,89	1
Weibull	2307,76	-1151,88	2
Gompertz	2308,22	-1152,61	2
Log-Normale	2335,83	-1165,92	2
Log-Logistique	2316,53	-1156,26	2
Gamma	2307,57	-1151,78	2

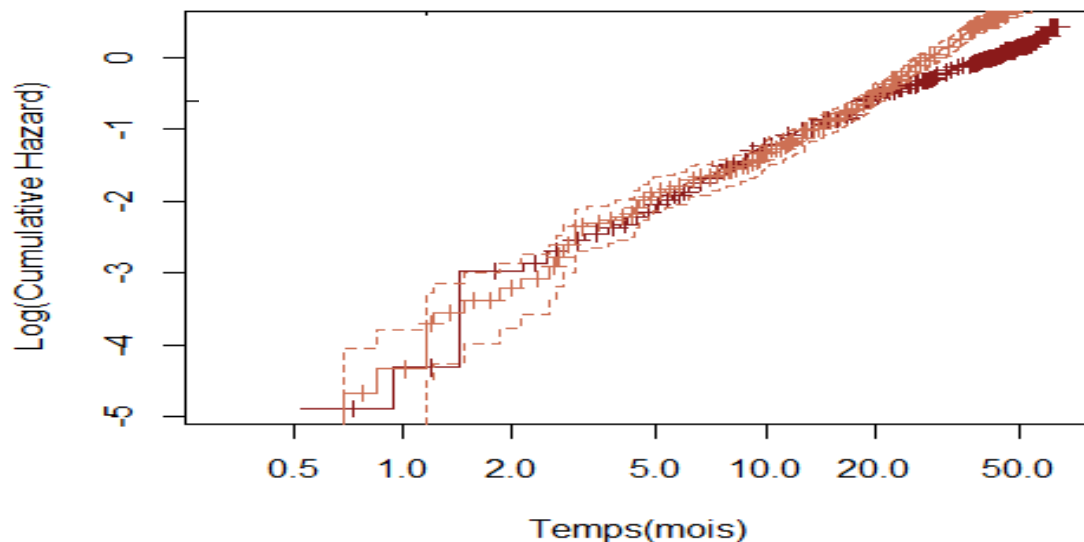
$$AIC = 2k - 2\ln(L)$$

Modèle à risque proportionnel vs modèle à risque accéléré

45



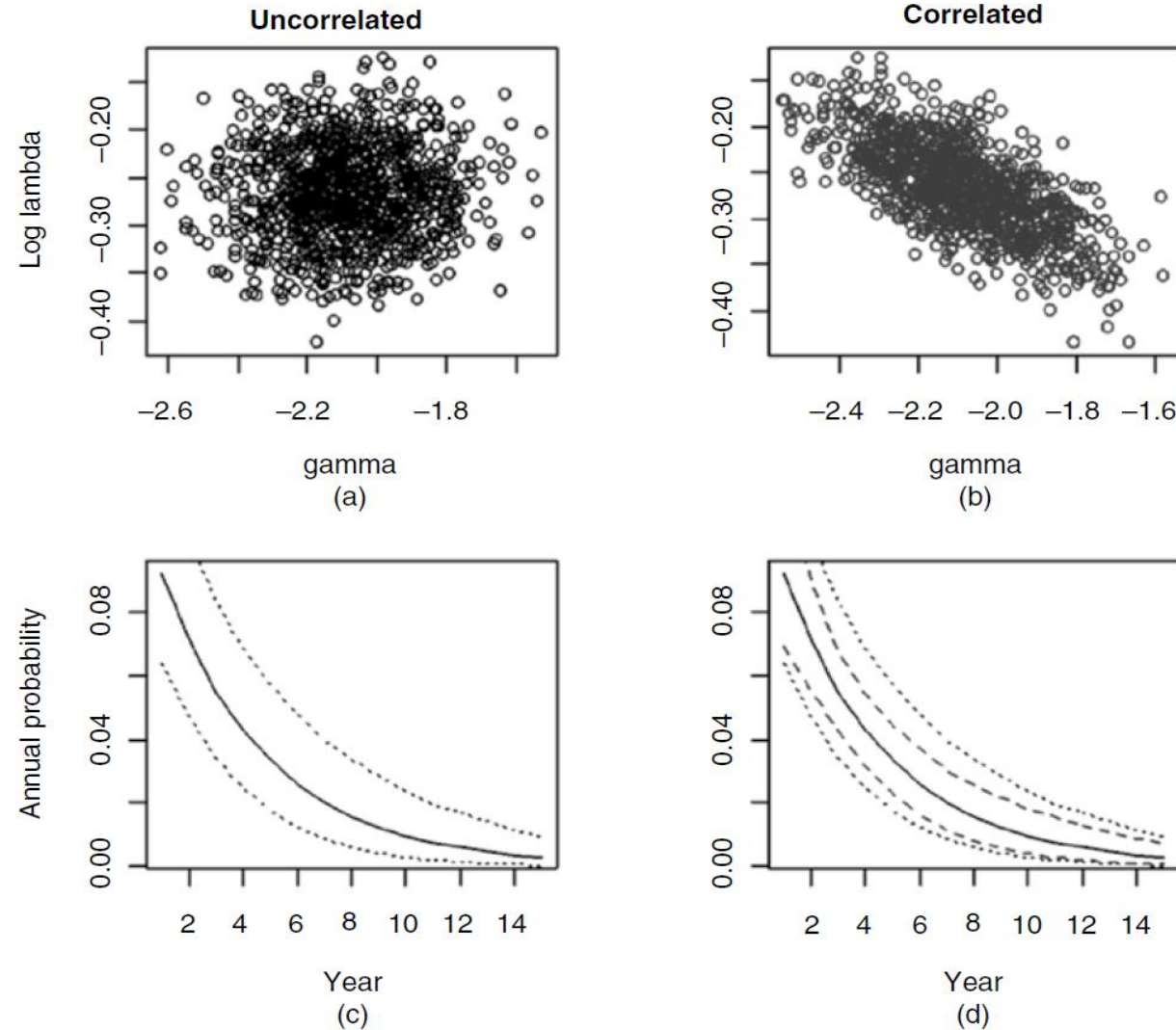
- Courbes représentant les Log des fonctions de risque cumulé des courbes KM sont **parallèles** :
 - **Proportionnalité des risques** pour les deux traitements
 - Choix d'une seule forme fonctionnelle et estimation dérivée de la courbe pour deuxième traitement



- Courbes ne sont pas **parallèles** :
 - Rejet de l'**hypothèse de risque proportionnel**; les risques entre les traitements ne sont pas proportionnels
 - Ajustement des courbes de survie doit se faire séparément pour les deux traitements

Corrélation des paramètres dans l'analyse de survie

46



Exercice 1 : étape 3



ANALYSE STOCHASTIQUE

Monte Carlo d'ordre 1



3

Propagation des Incertitudes

49

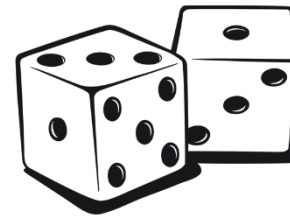
La technique du **Monte-Carlo d'ordre 1** est réalisée en quatre étapes :

1. Tirage aléatoire à partir des distributions de probabilité du **groupe test** d'une valeur pour chaque paramètre caractéristique et calcul du coût total et de l'efficacité totale.
2. Tirage aléatoire à partir des distributions de probabilité **du groupe contrôle** d'une valeur pour chaque paramètre caractéristique puis calcul du coût total et de l'efficacité totale correspondants.
3. **Calcul par différence du ratio différentiel coût résultat** correspondant.
4. **Répétition de l'opération N fois** pour obtenir une estimation de la distribution du RDCR dans le plan ACR.

Simulation d'une loi de probabilité

50

- ❖ Pour pouvoir **simuler** le comportement d'un système faisant intervenir des **événements probabilisés**, il va falloir simuler l'apparition de ces événements; c'est-à-dire générer des événements dont la fréquence observée sur un grand nombre de simulations doit être proche de la loi de probabilité théorique.
- ❖ La simulation d'une loi de probabilité quelconque peut se ramener à la **simulation d'une loi uniforme sur l'intervalle $[0;1[$** .



Valeur d'une variable aléatoire de la loi uniforme sur l'intervalle $[0;1[$

51

- **Fonction pseudo aléatoire** : une fonction (déterministe) qui permet de simuler une loi uniforme sur l'intervalle $[0;1[$. Cette fonction doit avoir les propriétés suivantes :
 - Les valeurs prises par cette fonction doivent être uniformément réparties sur l'intervalle $[0;1[$
 - Des résultats consécutifs doivent être indépendants
- Dans la pratique, ces fonctions sont réalisées par des méthodes de congruence, ce qui signifie que les résultats sont périodiques, mais la période est suffisamment longue (plusieurs milliards) pour que cela ne gêne pas la réalisation de simulations.
- Formule sur Excel =ALEA()

Simulation individuelle

52

#	Baseline		Sequence						TTP			DoT			Death			TTD
	Line	Seq	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5	Line 1	Line 2	Line 3	Line 1	Line 2	Line 3	Line 1	Line 2	Line 3	BSC
1	0,106	0,533	0,580	0,290	0,302	0,775	0,014	0,761	0,814	0,709	0,045	0,414	0,863	0,790	0,374	0,962	0,871	0,056
2	0,950	0,364	0,525	0,767	0,054	0,592	0,469	0,298	0,623	0,648	0,264	0,279	0,830	0,825	0,589	0,986	0,911	0,227
3	0,695	0,980	0,244	0,534	0,106	0,999	0,676	0,016	0,775	0,100	0,103	0,599	0,284	0,046	0,296	0,382	0,301	0,949
4	0,980	0,401	0,278	0,160	0,163	0,647	0,410	0,413	0,713	0,326	0,633	0,208	0,186	0,583	0,081	0,458	0,906	0,261
5	0,785	0,379	0,290	0,919	0,632	0,628	0,428	0,098	0,561	0,694	0,914	0,835	0,023	0,543	0,916	0,430	0,678	0,502
6	0,514	0,463	0,353	0,405	0,270	0,056	0,244	0,979	0,061	0,390	0,365	0,490	0,156	0,474	0,257	0,629	0,542	0,156
7	0,939	0,654	0,506	0,390	0,107	0,784	0,460	0,754	0,596	0,833	0,019	0,210	0,074	0,105	0,332	0,128	0,000	0,537
8	0,657	0,544	0,827	0,082	0,192	0,679	0,454	0,357	0,150	0,704	0,929	0,530	0,090	0,758	0,402	0,462	0,492	0,208
9	0,330	0,095	0,590	0,170	0,928	0,098	0,444	0,273	0,873	0,751	0,273	0,674	0,257	0,090	0,031	0,323	0,790	0,297
10	0,235	0,480	0,255	0,341	0,045	0,482	0,206	0,865	0,589	0,755	0,928	0,331	0,543	0,081	0,634	0,410	0,960	0,115
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Simulation d'une variable discrète

53

- Soit une loi de probabilité discrète définie par $P(X = x_i) = p_i$ pour $i = 1, \dots, n$ (ex. attribution du traitement en première ligne)

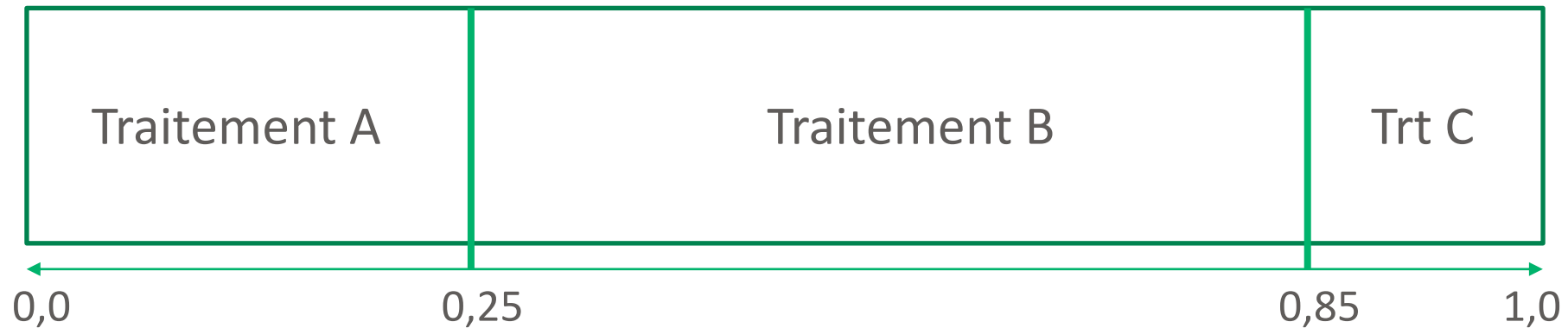


- La règle d'affectation est suivante:
 - $0 \leq x < p_1$ alors $X = x_1$
 - $p_1 \leq x < p_2$ alors $X = x_2$
 - ...

Exemple

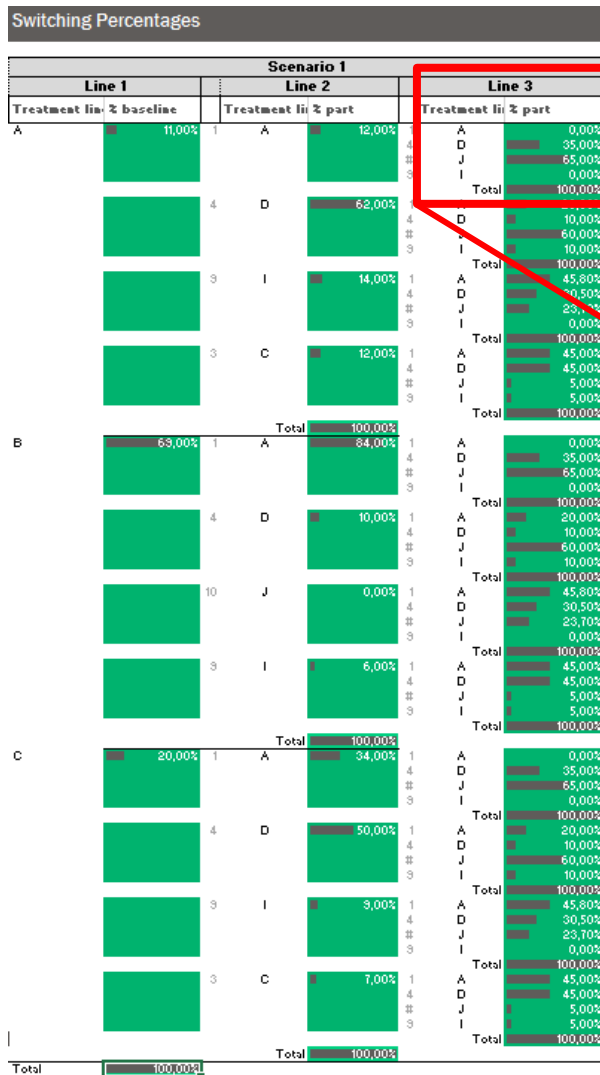
54

- Pour le patient N la probabilité aléatoire simulée = 0,48
- En fonction de cette valeur, définir le traitement qu'il va prendre, sachant que les distribution des parts de marché sont les suivantes:



Distribution des traitements par ligne (parts de marché)

55



Line 3	
Treatment line	% part
A	0,00%
D	35,00%
J	65,00%
I	0,00%
Total	100,00%

Quelle est la probabilité de suivre une séquence: A – A – D?

56

Line 1		Line 2		Line 3	
Treatment line	% baseline		Treatment line % part		Treatment line % part
A		1	A		1 A 0,00%
		4	D		4 D 35,00%
		9	I		# J 65,00%
					9 I 0,00%
					Total 100,00%
		1	D		1 A 20,00%
		4	D		4 D 10,00%
		9	I		# J 60,00%
					9 I 10,00%
					Total 100,00%
		1	I		1 A 45,80%
		4	D		4 D 30,50%
		9	I		# J 23,70%
					9 I 0,00%
					Total 100,00%
		1	C		1 A 45,00%
		4	D		4 D 45,00%
		9	I		# J 5,00%
					9 I 5,00%
					Total 100,00%
		Total		100,00%	

Identification des séquences

57

First line	Second line	Third line	% L1	% L2	% L3	% Total
A	A	A	11,00%	12,00%	0,00%	0,00%
A	A	D	11,00%	12,00%	35,00%	0,46%
A	A	J	11,00%	12,00%	65,00%	0,86%

Sachant qu'en L1 il y a 3 traitements, en L2 4 et en L3 4 traitements aussi, combien y a-t-il de séquences au total?

A	D	I	11,00%	62,00%	0,00%	0,00%
A						0,00%
A						0,54%
...	...	...	...	...	...	...

$$N = 3 * 4 * 4 = 48$$

Séquençage

58

First line	Second line	Third line	% Total
1	1	1	0,00%
1	1	4	0,46%
1	1	10	0,86%
1	1	9	0,00%
1	4	1	0,00%
1	4	4	2,39%
1	4	10	4,43%
1	4	9	0,00%
1	9	1	0,00%
1	9	4	0,54%
...	...	...	...

Entry Line	Traitement		
	Line 1	Line 2	Line 3
1	B	D	I
1	B	D	I
1	C	C	J
1	B	D	I
1	B	D	I
1	B	D	I
1	B	I	I
1	B	D	I
1	A		
1	B	D	I
1	B	D	I
1	B	D	J
1	B	D	J
...	...	...	...
1000	A	C	J

Loi de distribution des paramètres de l'efficacité

59

Progression-free Survival

Treatment	A	B	C
Method of input	Distribution	Distribution	Distribution
Distribution used	Gompertz	Weibull	Weibull
λ	0,006	0,007	0,008
γ	0,907	1,240	1,207
Max. duration (years)	5,00	5,00	5,00

Duration of Treatment

Treatment	A	B	C
Max. duration (years)	5,000	5,000	5,000
Distribution	Weibull	Weibull	Weibull
HR Value	1,550	2,100	1,473
λ	0,009	0,014	0,011
γ	0,907	1,240	1,207

Overall Survival

Distribution	Weibull
λ	0,011
γ	0,989
Max. duration (years)	15

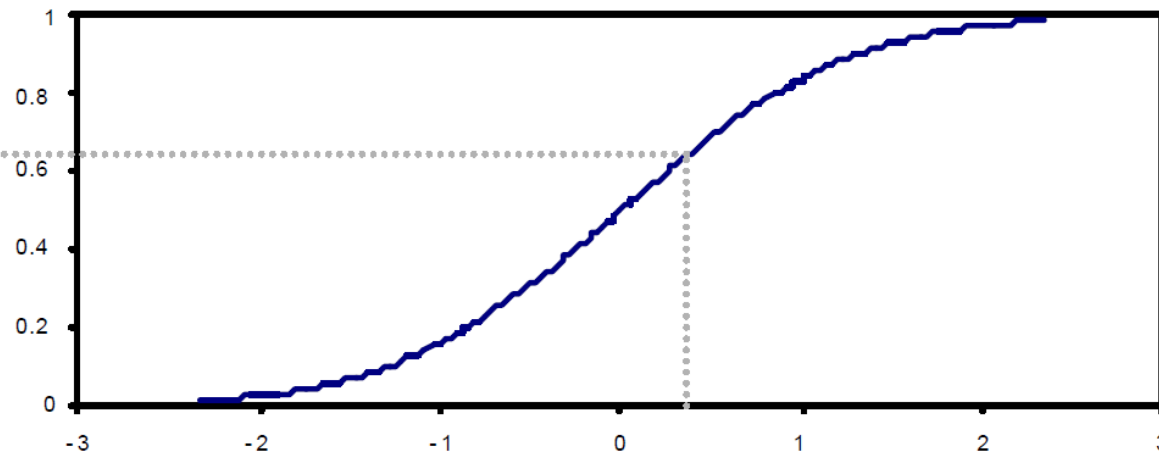
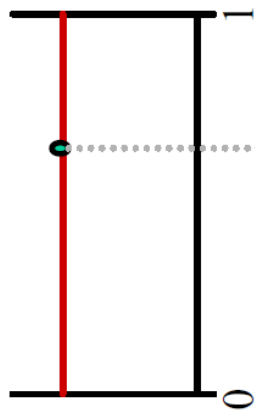
Utilisation de la fonction de Répartition d'une distribution

60

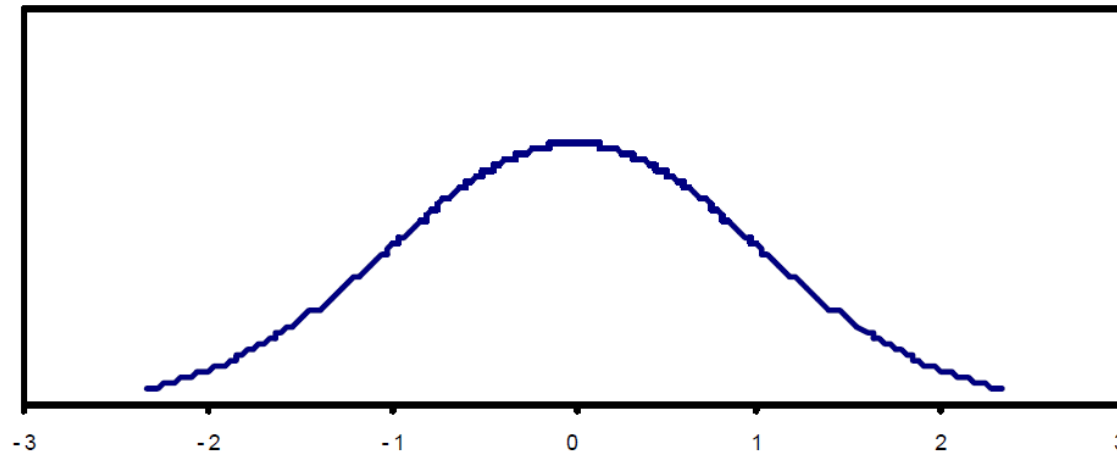
- La **fonction de densité** de probabilité (PDF)
- La caractéristique de la fonction de densité est que son intégrale est égale à 1 (c.à.d. que la surface sous la courbe est égale à l'unité)
- La **fonction de répartition cumulée (CDF)** est l'intégrale de la fonction de densité (PDF) et son domaine est défini sur l'intervalle [0-1]
- On peut donc utiliser l'inverse de la fonction **de répartition cumulée** pour générer des tirages aléatoires à partir de la fonction correspondante

Illustration graphique

61



Fonction de répartition



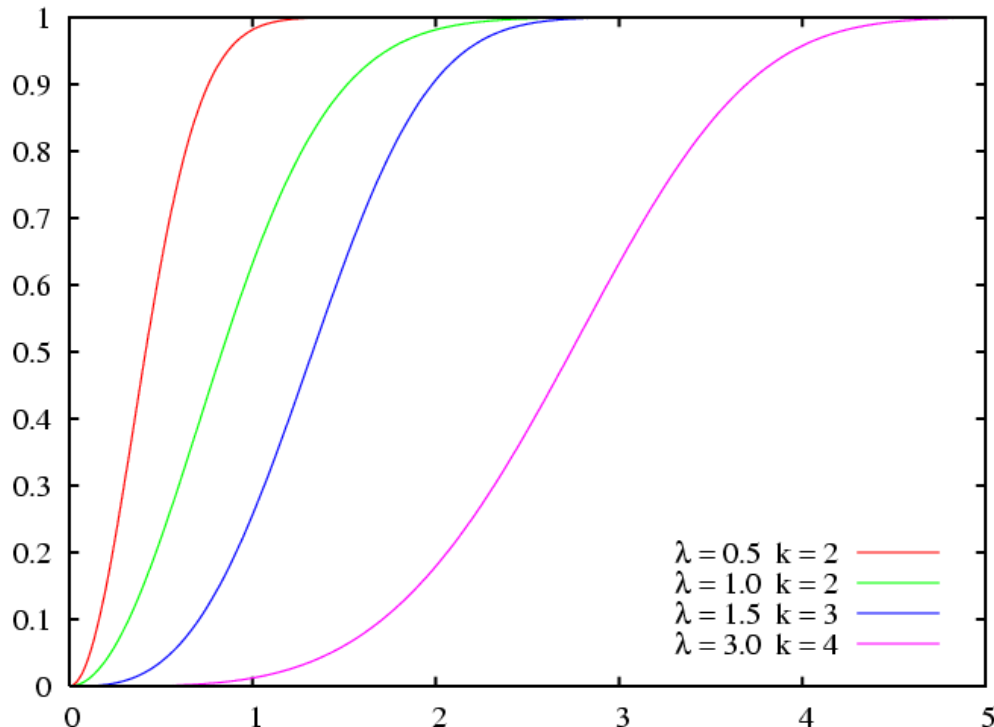
Densité de probabilité

Fonction de répartition de la loi de Weibull

62

Fonction de répartition:

$$F(x; k; \lambda) = 1 - \exp(-(x/\lambda)^k)$$



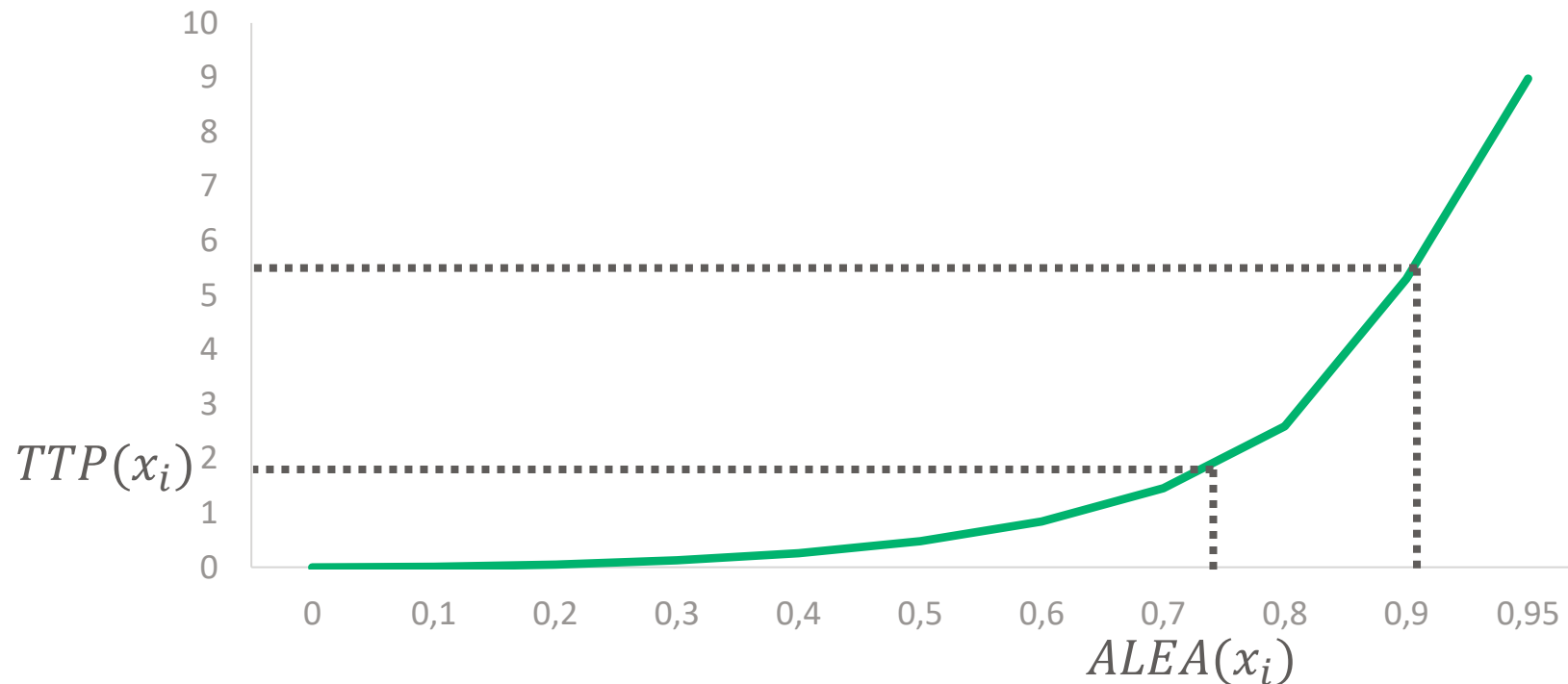
- La fonction quantile (**fonction de répartition inverse cumulée**) pour la distribution de Weibull:

$$Q^{-1}(p; k; \lambda) = \lambda(-\ln(1 - p))^{\frac{1}{k}} \\ 0 \leq p < 1$$

Fonction de répartition inverse

63

- Un nombre aléatoire uniforme sur l'intervalle (0, 1) est utilisé pour sélectionner les valeurs de la fonction de répartition inverse cumulée
- L'abscisse représente la fonction de répartition inverse cumulée de la loi de Weibull



Simulation de la loi continue

64

- Pour simuler un tirage aléatoire dans une loi normale $N(\mu, \sigma)$, on utilisera la formule :

LOI.NORMALE.INVERSE(ALEA(); μ ; σ)

- Pour une loi exponentielle de paramètre 1, on utilisera le fait que la loi exponentielle de paramètre λ est la loi gamma particulière de paramètres 1 et $1/\lambda$. L' inverse de la fonction de répartition de la loi gamma étant donnée dans Excel, on utilisera la formule

LOI.GAMMA.INVERSE(ALEA(); 1; λ)

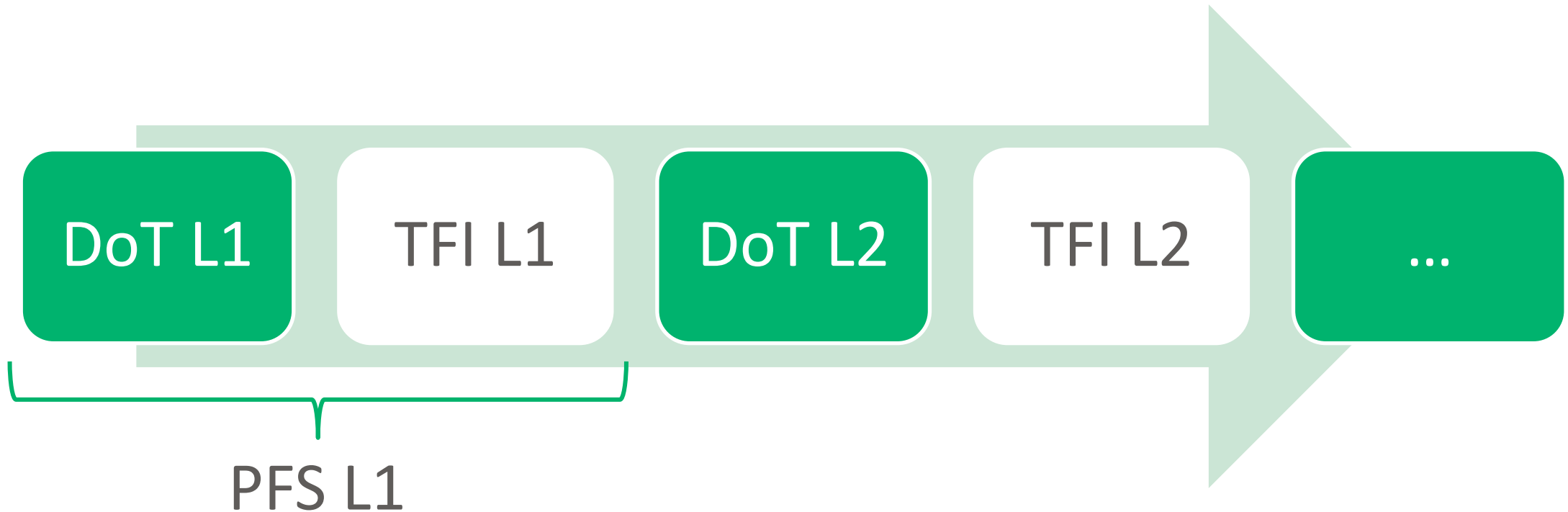
1000 Simulations Individuelles

65

Entry Line	Traitement			TTP			DoT			Death			TTD	
	Line 1	Line 2	Line 3	Line 1	Line 2	Line 3	Line 1	Line 2	Line 3	Line 1	Line 2	Line 3	Time	Line
1	B	D	I	0,72	1,07	2,34	0,72	0,88	1,04	FAUX	FAUX	FAUX	7,789	
1	B	D	I	1,41	1,84	2,60	1,41	1,61	1,75	FAUX	FAUX	FAUX	5,572	
1	C	C	J	0,74	2,39	4,58	0,74	1,69	4,14	FAUX	FAUX	FAUX	5,042	
1	B	D	I	1,08	2,10	2,49	1,08	2,10	2,36	FAUX	FAUX	FAUX	5,213	
1	B	D	I	1,66	2,02	2,17	0,36	2,02	2,17	FAUX	FAUX	FAUX	3,74	
1	B	D	I	5,92	6,79	7,43	1,08	2,70	3,02	FAUX	FAUX	FAUX	11,06	
1	B	I	I	1,52	1,65	3,14	1,52	1,65	2,28	FAUX	FAUX	VRAI	3,137	3
1	B	D	I	4,33	4,68	4,81	0,98	3,04	3,21	FAUX	FAUX	FAUX	7,936	
1	A			0,38			0,38			VRAI			0,378	1
1	B	D	I	1,55	1,84	1,97	1,54	1,84	1,97	FAUX	FAUX	FAUX	6,147	
1	B	D	I	7,16	8,13	8,60	0,19	0,78	1,14	FAUX	FAUX	FAUX	9,537	
1	B	D	J	0,72	1,31	2,28	0,72	1,31	1,84	FAUX	FAUX	FAUX	4,389	
1	B	D	J	2,02	3,48	3,98	0,52	1,00	1,31	FAUX	FAUX	FAUX	4,433	
1	A	D	I	11,57	11,91	12,39	1,15	2,50	2,82	FAUX	FAUX	FAUX	12,94	
1	A			0,03			0,03			VRAI			0,032	1
1	B	D	I	1,95	1,98	2,80	1,38	1,98	2,42	FAUX	FAUX	FAUX	3,239	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	

Durée de traitement et intervalle libre

66



PFS = Progression-Free Survival (Survie sans progression)

DoT = Duration of Treatment (Durée de traitement)

TFI = Treatment-Free Interval (Intervalle sans traitement)

Exercice 2



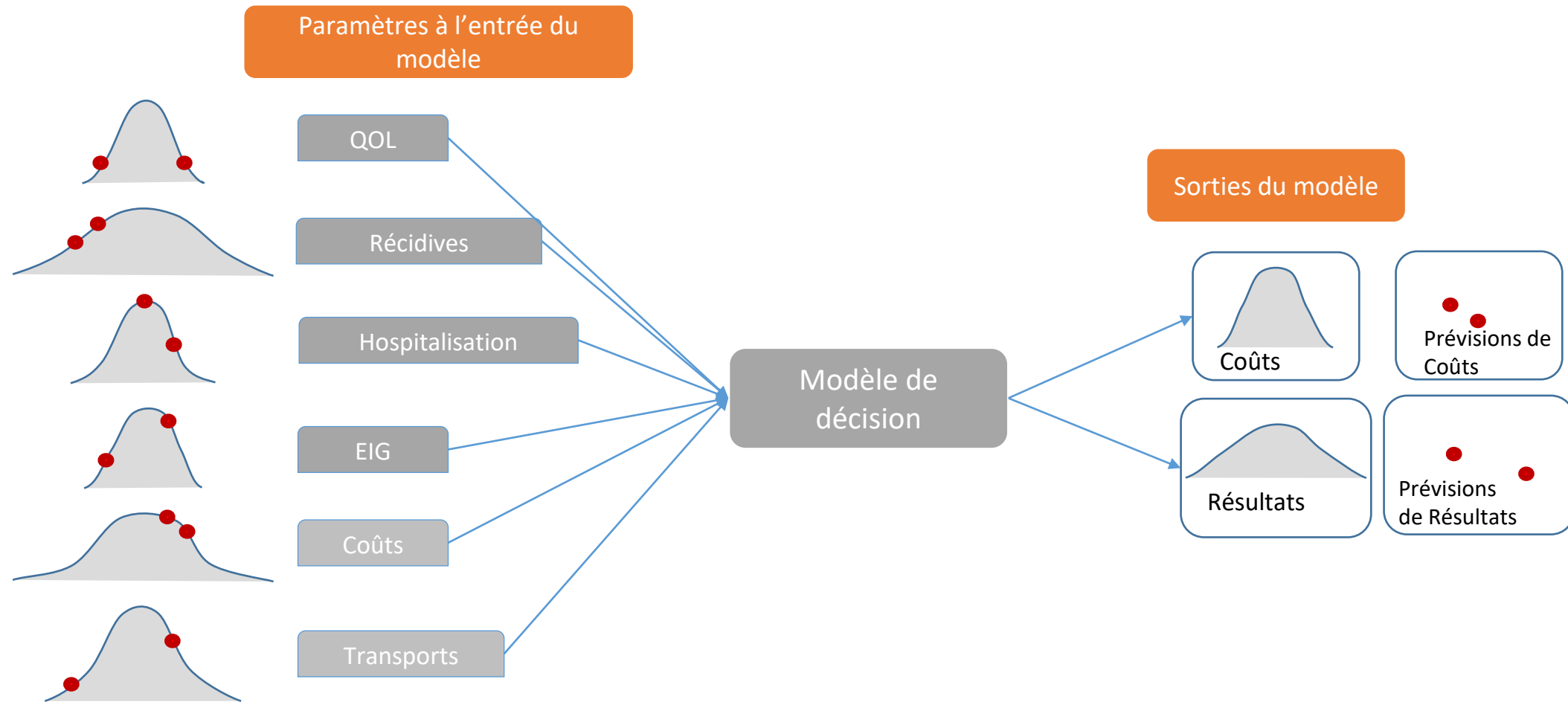
4



ANALYSE PARAMETRIQUE : Monte Carlo d'ordre 2

Réunir l'Incertitude dans de grands sacs d'ignorance

69



Analyse probabiliste

70

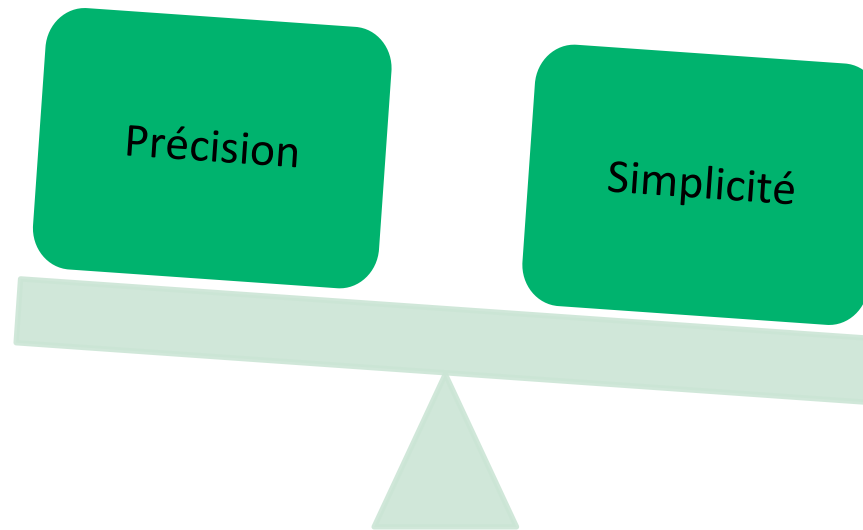
- A chaque variable aléatoire utilisée, on associe non pas une **probabilité moyenne** mais une **distribution de valeurs possibles** associées à leur probabilité de survenue
- Pour une famille arrêtée de loi de probabilité, on « **cale** » la **valeur de ses paramètres théoriques** qui simule le mieux la réalité observée
- Après **avoir spécifier la loi de distribution** de chaque variable, on **tire au sort la réalisation** de chacune d'entre elles et l'incertitude **se propage** dans tout le modèle.
- Le résultat d'une **analyse probabiliste des risques** est une distribution de probabilité. Sur un grand nombre de tirages, la moyenne des sorties du modèle approche leur espérance.

Choix des distributions pour les paramètres d'efficacité

71

Choisir la distribution appropriée

La **précision** de l'ajustement améliorera la précision du modèle



La **simplicité** protège contre les erreurs des utilisateurs et facilite la révision, la reproductibilité et la faisabilité du calcul.

Distributions couramment utilisées

72

Distribution	Parameters	Application
Normal	Mean (μ); standard deviation (σ)	Effectiveness, utilities
Beta	α, β	Effectiveness, utilities
Dirichlet	α, β	Effectiveness
Gamma	α, β	Costs, utilities
LogNormal	Mean (μ); standard deviation (σ)	Costs, effectiveness
Weibull	λ, κ	Effectiveness
Gompertz	γ, λ	Effectiveness
Exponential	λ	Effectiveness

Estimation des paramètres de la population par la Méthode des Moments

73

La méthode des moments implique l'égalisation des **moments empiriques** avec des **moments théoriques**

- $E(X^k), E\left((X - E(X))^k\right)$ k^{th} moments théoriques de la distribution, $k = 1, 2, \dots$
- $M_k = \frac{1}{n} \sum_{i=1}^n X_i^k, M_k^* = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$, sont les moments empiriques, $k = 1, 2, \dots$

L'idée: estimer les paramètres recherchés en égalisant les premiers moments théoriques avec leurs contreparties empiriques:

$$M_1 = E(X) = \bar{X} \text{ avec } E(X) - \text{espérance}$$

$$M_2 = E\left((X - E(x))^2\right) = s^2 \text{ avec } E\left((X - E(x))^2\right) - \text{variance de la variable}$$

Moments de la distribution Bêta

74

$$\theta \sim \text{Beta}(\alpha, \beta)$$
$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbf{1}_{[0,1]}(x)$$

$$E(X) = \frac{\alpha}{\alpha + \beta}$$

$$V(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Méthode des moments : loi Bêta

75

- Définir les moments théoriques et les moments empiriques à partir de l'échantillon:

$$\bar{x} = \frac{\alpha}{\alpha + \beta} \qquad s^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

- Et réorganiser:

$$(\alpha + \beta) = \frac{\bar{x}(1 - \bar{x})}{s^2} - 1$$

$$\alpha = \bar{x}(\alpha + \beta) \qquad \beta = (\alpha + \beta) - \alpha$$

Exemple 1: $\bar{x} = 0,65; s = 0,12$

$$\alpha = 0,65 * (\alpha + \beta) = 0,65 * \left(\frac{0,65 (1 - 0,65)}{0,12^2} - 1 \right) = 0,65 * 14,80 = 9,62$$
$$\beta = 14,80 - 9,62 = 5,18$$

Méthode des moments : loi Gamma

76

- Définir les moments théoriques et les moments empiriques à partir de l'échantillon:

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i & E(X) &= \frac{\alpha}{\lambda} \\ s^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 & V(X) &= \frac{\alpha}{\lambda^2}\end{aligned}$$

- Et réorganiser:

$$\begin{aligned}\alpha &= \frac{\bar{x}^2}{s^2} & \lambda &= \frac{\bar{x}}{s^2}\end{aligned}$$

Méthode des moments : loi Log-Normale

77

- Moyenne : $E(X) = e^{\mu + \frac{\sigma^2}{2}}$
- Variance : $V(X) = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2}$
- Par la méthode des moments:

$$\bar{x} = e^{\mu + \frac{\sigma^2}{2}}$$
$$\mu = \ln(\bar{x}) - \frac{\sigma^2}{2}$$
$$\sigma = \sqrt{\ln\left(1 + \frac{s^2}{\bar{x}^2}\right)}$$

Exercice 3

Bibliographie

- Briggs, A.H., 2004. Statistical approaches to handling uncertainty in health economic evaluation. Eur J Gastroenterol Hepatol 16, 551–561.
- **Briggs, A., Sculpher, M., Claxton, K., 2006. Decision Modelling for Health Economic Evaluation, 1 edition. ed. OUP Oxford.**
- Davies, C., Briggs, A., Lorgelly, P., Garellick, G., Malchau, H., 2013. The “Hazards” of Extrapolating Survival Curves. Med Decis Making 33, 369–380.
- Demiris, N., Lunn, D., Sharples, L.D., 2015. Survival extrapolation using the poly-Weibull model. Stat Methods Med Res 24, 287–301.
- **Grieve, R., Hawkins, N., Pennington, M., 2013. Extrapolation of survival data in cost-effectiveness analyses: improving the current state of play. Med Decis Making 33, 740–742.**
- **Guyot, P., Ades, A.E., Ouwers, M.J., Welton, N.J., 2012. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC medical research methodology 12, 9.**
- **Latimer, N.R., 2013. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making 33, 743–754.**

» **Adresse:** REES France
28, rue d'Assas
75006 Paris, France

» **Téléphone:** +33 (0)1 44 39 16 90

» **Email:** launois.reesfrance@wanadoo.fr

» **Web:** www.rees-france.com