Nouvelle Partie – 09/2015

Problème : simulate censored data such that most with longest time are censored

Title of paper: Dynamic Optimization of Censor Rate for Simulated “Extinction”

Idée: **Apprendre le meilleur censor rate dynamiquement**

Context

- Le contexte général est l’analyse de survie, c’est à dire l’estimation du taux d’apparition d’un événement.
  - Pour estimer un tel taux, idéalement, on suit plusieurs individus (au sens statistique, dans ton cas un l’individu statistique est en fait une population) jusqu’à l’apparition de l’événement pour chaque individu.
- En réalité, il est difficile d’observer tous les individus jusqu’à l’apparition de l’événement parce que le coût d’observation est trop élevé (le coût peut être la durée de l’étude si on a besoin d’une réponse rapide, ça peut-être un coût monétaire dans le cas des études cliniques, ou un coût en CPU dans le cas des serveurs de calculs).
- En conséquence, nous avons des données dites tronquées (ou censurées) c’est à dire que pour certains individus on arrête l’observation avant l’apparition de l’événement. Le fait d’avoir des données tronquées se traduit statistiquement par de la perte d’information.
- En effet, savoir que l’événement à lieu pour $t > T$ (donnée tronquée) est moins informatif que savoir que l’événement à lieu à $t = T$. Les méthodes statistiques d’analyse de survie permettent de faire pour le mieux dans ces conditions, c’est à dire d’utiliser efficacement l’information qu’il reste dans les données tronquées.

Problème

- Notion importante : Il faut distinguer ici les données de l’information (au sens statistique). Les données peuvent être résumée par le nombre d’individus. L’information est ce qui est vraiment utile dans les données pour faire de l’inférence (i.e. estimation).
  - Example (extrême): variable 1 = x, variable 2 = 2*x. Si je considère les variable 1 et 2 à la fois, je double le nombre de données mais la quantité d’information est inchangée (parce que les variables 1 et 2 sont redondantes, non-indépendante).
- Notre problème est un problème d’optimisation.
  - Nous avons une quantité de ressources limitée (par exemple CPU, temps, argent). Dans ces conditions, deux plans d’expériences extrêmes et opposés sont les suivants:
    - (1) suivre peu d’individus pendant longtemps : Dans ce cas, l'avantage est que l’on a peu de perte d’information par données (car peu de données tronquées), par contre l’inconvénient est que l’on a peu de données (car peu d’individus)
    - (2) suivre beaucoup d’individus pendant une durée courte : ici au contraire l’avantage est que l’on a beaucoup de données (car beaucoup d’individus), par contre l’inconvénient est que beaucoup de ces données se trouve tronquées et donc ont une perte d’information.

  Comme nos resources sont limitées, il est impossible de combiner les avantages des deux méthodes : nous sommes face à un compromis (deuxième notion importante). Le problème consiste donc à trouver la meilleure stratégie intermédiaire entre ces deux extrêmes. La solution dépend de la valeur du taux, que l’on ne connaît pas, puisqu’on cherche précisément à l’estimer!

BALANCE entre le nombre de villes et la duration de simulation

ETAT DE L'ART

Censored Data?

1. **Donées censurées I**:

- Pendant les heures de test de T, il y a r échecs (où r peut être n'importe quel nombre de 0 à n).
- Les temps exacts de défaillance sont t1, t2, ..., tr .
- Il y a (n-r) cas qui sont survécus à l'ensemble du test T-heure sans défaillance.
- T est fixé à l'avance et r est aléatoire. On ne sais pas combien échecs va se produire jusqu'à ce que le test est fini.

2. **Donées censurées II** :

- On observe t1, t2, ...,tr , où r est défini à l'avance. Le test se termine au temps t= tr, et (n-r) unités sont survécu.
- Rarement utilisé.

ETAT DE L'ART

Méthodes pour résoudre le problème de balance :

- 1 . **Bandit model**

  Le problème de bandit multi-armés stochastique est un modèle important pour étudier le « tradeoff » exploration-exploitation dans l'apprentissage par renforcement.

- 2. **Adaptive design**

  « Adaptive designs » est un désign qui a pour de maximiser le résultat obtenu et minimiser le résultat rédondant. Pendant le temps de simulation, les agents apprendent pour adapter à ses situation courentes et maximiser le résultat. Le but est non seulement d'identifier efficacement les avantages du traitement cliniques sous enquête, mais aussi pour accroître la probabilité de succès du développement clinique.

- 3. **Algorithme heuristique /Algorithme glouton (gourmand)**

  Les deux sont un choix optimum local.

Définir notre problem

- Problème : simulate censored data such that most with longest time are censored.
- Title of paper: Dynamic Optimization of Censor Rate for Simulated “Extinction”.

Idée: Apprendre dynamiquement le meilleur censor rate.

Solution proposée : par Prf. JEAN DANIEL

ALGO proposée par Prf. JEAN DANIEL
Input: Budget de TS CPU time
N Sub Pop
NUMBER OF SIMU // Je ne comprend pas la rôle de NUMBER SIMU

REPEAT DURING TS/10 // à l'instant t,
Select Random Population of N Sub-pop //random A subPOP dans N subpop (subPOP1, …,subPOPn)
Run and WAIT until Extinction to compute Extinction Time //duration:[0, Ttemp], tenir les subPOP d'extinction global, estimer le taux d'EXT avec ces subPOP. // simuler A subPOP indépendant ou une métapopulation de A subpop dépendantes ?
END REPEAT
COMPUTE OPTIMAL WaitingTime to have NUMBER SIMU→ OPTWAIT//c'est à dire que, après chaque répétition, on recalcule la valeur de OPTWAIT, OPTWAIT ← t?

REPEAT UNTIL END OF BUDGET
RUN until OPTWAIT to cmpute // On arrête quand on arrive NUMBER SIMU ? Comment calculer OPTWAIT ?

Solution proposée : Algo EXP3 Bandit

- Bandit algorithm is a framework to balance the tradeoff of Exploitation and Exploration.
  - Exploration : exploit existing knowledge.
  - Exploitation : explore new knowledge from these unknow regions.

- Multi Armed Bandits:
  - Arms = possible treatments
  - Arm Pulls = application of treatment to inidividual
  - Rewards = outcome of treatment
  - Objective = maximize cumulative reward = maximize benefit to trial population (or find best treatment quickly)

- Le bandit contre un adversaire :
  - 2 types de problèmes selon l’information reçue à chaque round:
    - information parfaite (“full information”)
    - information partielle (“bandit information”).

Modèle de bandit contre un adversaire

On dispose de K bras. A chaque instant t = 1,. . . ,n.

- L’adversaire choisi des récompenses xt(1), . . . ,xt(K) (supposées à valeurs dans [0, 1], sans les dévoiler au joueur
- Le joueur choisit un bras It
- Dans le cas information parfaite, le joueur observe les récompenses de tous les bras xt(k), pour k =1 . . . , K.
- Dans le cas information partielle, le joueur n’observe que la récompense du bras choisi: xt(It).

Modèle de bandit contre un adversaire

- Information parfaite : (négliger)
- Information partielle : Considérons l'algorithme EXP3 (Exploration-Exploitation using Exponential weights):

  - $\eta > 0$ et $\beta > 0$ sont deux paramètres de l'algorithme.
  - On initialise l'algorithme avec $w_1(k) = 1$ pour tout $k = 1, \ldots, n$.
  - A chaque instant $t = 1, \ldots, n$, le joueur choisit le bras $I_t \sim p_t$, où

  $$p_t(k) = (1 - \beta)\frac{w_t(k)}{\sum_{i=1}^{K} w_t(i)} + \frac{\beta}{K},$$

  avec

  $$w_t(k) = e^{\eta \sum_{s=1}^{t-1} \tilde{x}_s(k)},$$

  où $\tilde{x}_s(k) = \frac{x_s(k)}{p_s(k)} \mathbf{1}\{I_s = k\}$.

Questions

1. Est-ce que je déjà comprends vraiment le problème?

   – Si vrai, je commence à coder?

2. En effet, je ne sais pas comment on peux utiliser l'algorithme Bandit dans notre problème. Modéliser les conceptions : BRAS, LOIS à jouer, Récompense?

3. Pour commencer les codes, on commence avec le modèle de \rho=0? Je trouve que, la valeur de “graine” influence fortement le temps de persistence de maladie.

3. J'ai encore la partie “optimiser les politiques de vaccination”, d'abord on finit la partie courante, on va exécuter la partie de vacciantion?

4. Je vais envoyer le mémoire.

5. J'ai envoyé Prf. Christope main aucune réponse, sur le PDI.