

Journées d'études de Deauville 2015

L'épidémie de fièvre Ebola 2014/2015 en Afrique
de l'Ouest comme trajectoire d'un Processus de
Branchement [Working Paper]

Guillaume Ominetti¹ Benjamin Schannes²

24 Septembre 2015

¹SCOR, 5 avenue Kleber, 75795 Paris Cedex 16, France, e-mail:
gominetti@scor.com.

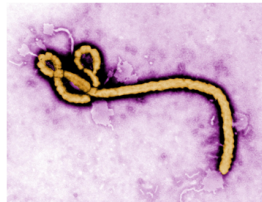
²Mercer, Tour Ariane, 92088 Paris La Defense Cedex, France, e-mail:
benjamin.schannes@mercercor.com.

Motivation

- ▶ Evaluer le risque de propagation d'une maladie virale par une approche statistique
- ▶ Donner un sens mathématique au terme générique propagation
- ▶ Paradigme Bayésien
- ▶ Décision statistique [Sous-critique] vs. [Sur-critique] dans un contexte de PB
- ▶ Comment résoudre le problème de prédiction 'Explosion ou Extinction' ?

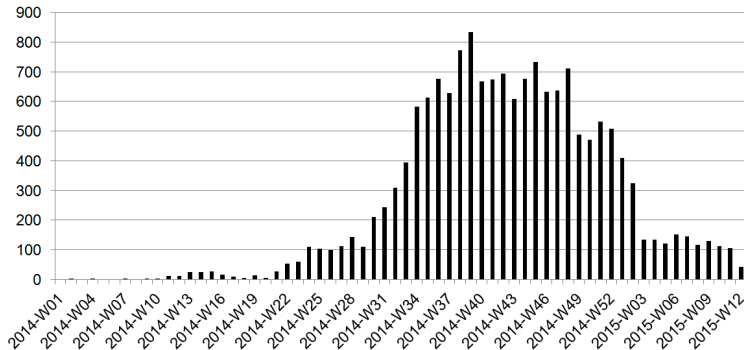
Maladie à Virus Ebola/Ebola Virus Disease (EVD)

- ▶ EVD: Fièvre hémorragique très violente
- ▶ Chauves-souris frugivores: porteuses saines
- ▶ Mécanisme de transmission: par les animaux sauvages (viande de brousse e.g.), par contact direct avec les liquides organiques d'humains infectés
- ▶ Taux moyen de létalité: environ 50 % (a varié de 25% à 90% dans les épisodes antérieurs)
- ▶ Période d'incubation: 2 à 21 jours
- ▶ Symptômes:
 - ▶ Soudaine montée de fièvre avec fatigue physique, douleurs musculaires, céphalées, maux de gorge
 - ▶ Ensuite 'diarrhée rouge', vomissements, éruptions cutanées, insuffisance rénale et hépatique, hémorragies internes et externes.
- ▶ Traitements anti EVD encore insuffisants



L'épidémie actuelle d'Ebola est la plus importante enregistrée depuis la découverte du virus en 1976

Weekly number of confirmed new EVD cases registered across Guinea, Liberia and Sierra Leone



- ▶ Pic de 833 nouveaux cas hebdomadaires confirmés lors de la semaine 39 de 2014
- ▶ L'O.M.S. fait état dans ses derniers bilans de 11,296 morts sur près de 27,898 cas

Philosophie Bayésienne: une méthode probabiliste (subjective) de révision de l'*a priori*

- ▶ Problème d'induction de Hume: Le soleil se lèvera-t-il demain? Mélangeant une approche expérimentale (observations passées) et les principes de la raison (Théorie des Probabilités), Laplace suppose une loi initiale uniforme pour la probabilité inconnue p que le soleil se lève le lendemain. La règle de Succession fournit alors

$$\Pr(r_{\text{demain} \sim n+1} \mid r_{\text{auj.} \sim n}, r_{\text{hier} \sim n-1}, \dots, r_0) = \frac{n+1}{n+2}$$

- ▶ Considérant que le soleil s'est levé chaque jour depuis 5000 ans, Laplace donne une cote de 1826251 : 1 en faveur de l'hypothèse suivant laquelle le soleil se lèvera le lendemain.
- ▶ ⚠ L'hypothèse faite lorsqu'on applique la règle de succession est que nous n'avons aucune connaissance *a priori* quant à la question de savoir si le soleil se lèvera ou non demain, si ce n'est la possibilité des deux cas de figure: $0 < p < 1$. Ce n'est pas le cas pour le lever de soleil ...

Approches Fréquentiste et Bayésienne

DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)

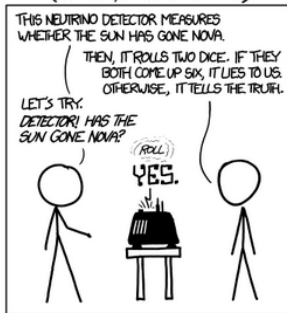
THIS NEUTRINO DETECTOR MEASURES
WHETHER THE SUN HAS GONE NOVA.

THEN, IT ROLLS TWO DICE. IF THEY
BOTH COME UP SIX, IT LIES TO US.
OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.

DETECTOR! HAS THE
SUN GONE NOVA?

(ROLL)
YES.



FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT
HAPPENING BY CHANCE IS $\frac{1}{36} = 0.027$.
SINCE $p < 0.05$, I CONCLUDE
THAT THE SUN HAS EXPLODED.



BAYESIAN STATISTICIAN:

BET YOU \$50
IT HASN'T.



Test Bayésien

- ▶ Suppose $\mathbf{X} \mid \theta \sim f(\mathbf{x} \mid \theta)$
- ▶ Nous voulons tester: $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$.
- ▶ Croyances *a priori*: $\theta \sim \pi_0(\theta)$ on Θ_0 and $\theta \sim \pi_1(\theta)$ on Θ_1 .
- ▶ Probabilités *a priori* $\Pr(H_0)$ and $\Pr(H_1)$ des hyp.
- ▶ Vraisemblances marginales associées aux hypothèses

$$\mathcal{L}(\mathbf{x} \mid H_i) = \int_{\Theta_i} f(\mathbf{x} \mid \theta) \pi_i(\theta) d\theta, \quad i = 0, 1.$$

- ▶ Distributions *a posteriori* des paramètres sous les hypothèses:

$$g_i(\theta \mid \mathbf{x}) = \frac{f(\mathbf{x} \mid \theta) \chi_{\Theta_i}(\theta)}{\mathcal{L}(\mathbf{x} \mid H_i)}, \quad i = 0, 1.$$

- ▶ Probabilités *a posteriori* $\Pr(H_0 \mid \mathbf{x})$ et $\Pr(H_1 \mid \mathbf{x})$ des hyp.
- ▶ Facteur de Bayes de H_0 vs. H_1

$$B_{01}(\mathbf{x}) = \frac{\frac{\Pr(H_0 \mid \mathbf{x})}{\Pr(H_1 \mid \mathbf{x})}}{\frac{\Pr(H_0)}{\Pr(H_1)}} = \frac{\mathcal{L}(\mathbf{x} \mid H_0)}{\mathcal{L}(\mathbf{x} \mid H_1)}$$

Echelle de Jeffreys

B_{01}	Force de la 'Preuve'
1 : 1 to 3 : 1	Anecdotique
3 : 1 to 10 : 1	Substantielle
10 : 1 to 30 : 1	Forte
30 : 1 to 100 : 1	Très Forte
> 100 : 1	Décisive

Processus de Branchement

- ▶ Soit \mathbf{p} une mesure de proba. sur \mathbb{N} , définie par $\{p_k; k = 0, 1, 2, \dots\}$.
- ▶ Soit $(X_n)_{n \in \mathbb{N}}$ la CM sur \mathbb{N} de matrice de transition P

$$P(x, y) = \mathbf{p}^{*x}(y) = p_y^{*x}, \quad \forall x, y \in \mathbb{N},$$

où \mathbf{p}^{*x} est la x -ième puissance de convolution de la loi \mathbf{p} .

- ▶ Fonction Génératrice h de $\xi \sim \mathbf{p}$

$$h(s) = \mathbb{E} \left[s^\xi \right] = \sum_{k=0}^{\infty} p_k s^k, \quad s \geq 0,$$

- ▶ Nombre moyen d'enfants d'un individu typique de la population

$$\mathbb{E}[\xi] = h'(1) \equiv m$$

Suivant que $m < 1$, $= 1$, ou > 1 , le processus de GW est dit *sous-critique*, *critique* ou *surcritique*.

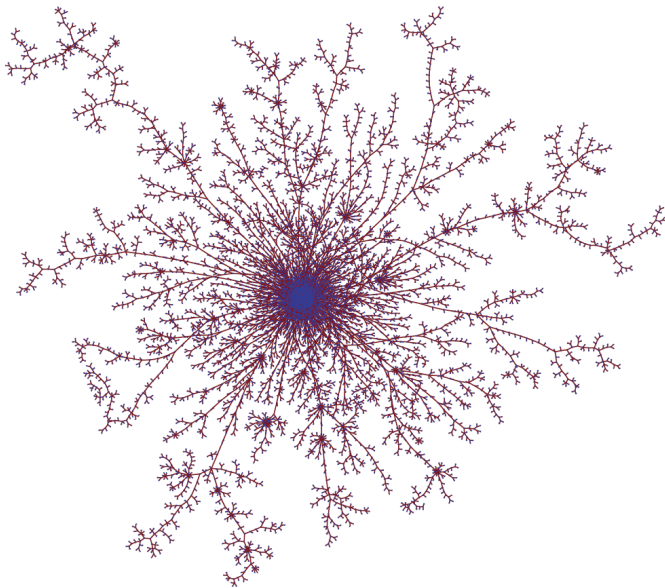
- ▶ L'ensemble d'Extinction et l'ensemble d'Explosion forment une partition de Ω :

$$\mathbf{Extinction} = \{\omega \in \Omega, X_n(\omega) = 0 \text{ pour un } n \geq 1\}$$

$$\mathbf{Explosion} = \{\omega \in \Omega, X_n(\omega) \rightarrow \infty \text{ as } n \rightarrow \infty\}$$

- ▶ La probabilité d'extinction $q := \mathbb{P}(\mathbf{Extinction})$ est la plus petite racine positive de l'équation $h(s) = s$.

Processus de Branchement sous-critique avec phénomène de condensation



Exemple d'approche bayésienne

- ▶ Posons $a_k = 1/k!$, si bien que la loi de la progéniture est de type Poisson(θ)
- ▶ $\exists t^* < \kappa$ tel que

$$H^{\text{sub}} := \{\theta \mid m_\theta < 1\} = (0, t^*)$$

$$H^{\text{sup}} := \{\theta \mid m_\theta > 1\} = (t^*, \kappa),$$

- ▶ Supposons $\theta \sim \text{Gamma}(\alpha, \beta)$

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}$$

où $\alpha > 0$, $\beta > 0$, and $\theta > 0$.

- ▶ Notons $\mathbf{x}_n \equiv (x_1, \dots, x_n)$ les données de nouveaux cas hebdomadaires. Posons $y_n := \sum_{i=0}^n x_i$.
- ▶ Le Théorème de Bayes fournit la distribution *a posteriori*

$$g(\theta \mid \mathbf{x}_n) \propto \frac{\theta^{\alpha+y_n-y_0-1}}{e^{(\beta+y_{n-1})\theta}}.$$

- ▶ A nouveau on retrouve une loi Gamma, mais avec paramètres révisés

$$\theta \mid \mathbf{x}_n \sim \text{Gamma}(\alpha + x_1 + \dots + x_n, \beta + x_0 + \dots + x_{n-1}).$$

Application numérique: Prédire l'Explosion ou l'Extinction pour un processus de GW Poisson/Gamma

- ▶ L'analyse repose sur la série de nombre de nouveaux cas confirmés hebdomadaires, enregistrés en Guinée, au Libéria et en Sierra Leone, entre le début de l'épidémie et le 31 mars 2015.
- ▶ Comme avant, on suppose que la progéniture est $\text{Poisson}(\theta)$, où $\theta \sim \text{Gamma}(\alpha, \beta)$.
- ▶ On ajuste α et β de telle sorte que les chances d'extinction et d'explosion *a priori* soient identiques (égales à 50 %).
- ▶ Nous avons favorisé les petites valeurs de α and β pour garder une relative variabilité dans la distribution *a priori*. Nous avons finalement retenu $\alpha = 1.85$ et $\beta = 1$.
- ▶ Facteur de Bayes

$$B_n^{\text{pred}}(\mathbf{x}_n) = \frac{\left(\frac{\mathbb{P}(\mathbf{Explosion} \mid \mathbf{x}_n)}{\mathbb{P}(\mathbf{Extinction} \mid \mathbf{x}_n)} \right)}{\left(\frac{\mathbb{P}(\mathbf{Explosion})}{\mathbb{P}(\mathbf{Extinction})} \right)}$$

Prédire l'Explosion ou l'Extinction pour un processus de GW Poisson/Gamma

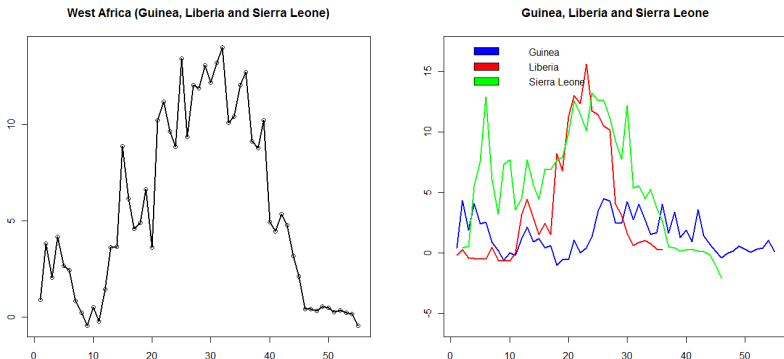












Figure: Evolution hebdomadaire du Facteur de Bayes associé au Problème de Prédiction [Explosion vs. Extinction] (échelle logarithmique, base 10) en Guinée, au Libéria et en Sierra Leone, de la semaine 9 de 2014 à la semaine 12 de 2015. Données: OMS

Conclusion

- ▶ Le besoin d'évaluer le risque de propagation peut être satisfait dans un cadre GW paramétrique et sous un paradigme Bayésien.
- ▶ Les fondations probabilistes de notre modèle assurent son implémentation fiable et rapide, et un diagnostic simple pour les autorités de santé publiques.

Merci de votre attention

-  Athreya, K.B., & Ney, P.E., *Branching Processes*, Springer-Verlag, Berlin, 1972.
-  Jaynes, E.T., Confidence Intervals vs Bayesian Intervals, *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science and Measure*, W.L. Harper and C.A. Hooker (eds.), 175, 1976
-  Guttorp, P., *Statistical inference for branching processes*, New York, Wiley, 1999.
-  Harris, T. E., *The Theory of Branching Processes. Die Grundlehren der Mathematischen Wissenschaften* **119**, Springer, Berlin, 1963.
-  Jagers, P., *Branching processes with biological applications*, Wiley, 1975.
-  Kimmel, M., & Axelrod, D.E., *Branching Processes in Biology*, Springer, 2002.
-  Mode, C. J., *Multitype Branching Processes*, Elsevier, New York, 1971.
-  Olofsson, P., & Sindi, S.S., A Crump-Mode-Jagers branching process model of prion loss in yeast, *Journal of Applied Probability*, Volume **51**, Number 2, 453-465, 2014.
-  Sevastyanov, B. A., *Branching Processes*, Nauka, Moscow, 1971.
-  Watson, H. W., & Galton, F., On the probability of extinction of families. *J. Royal Anthropological Inst.* **6**, 138-144, 1874.

ANNEXES

Ebola 2014: évolution en Guinée

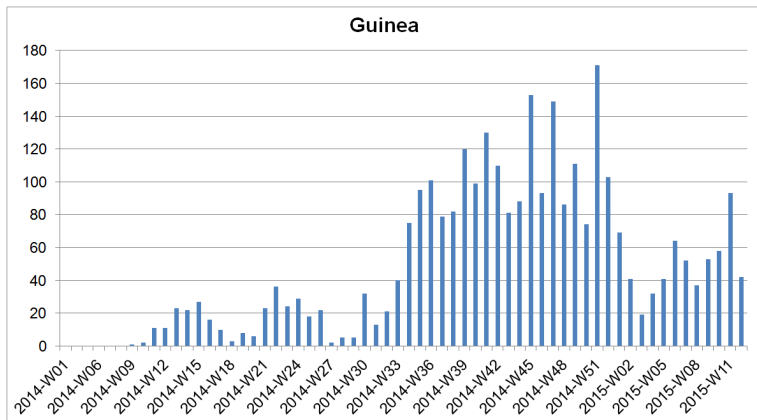


Figure: Nombre hebdomadaire de nouveaux cas d’Ebola confirmés en Guinée. Données:OMS

Ebola 2014: évolution au Liberia

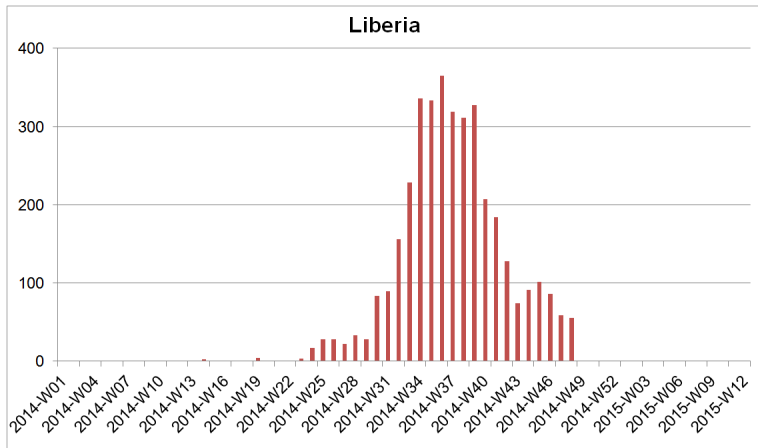


Figure: Nombre hebdomadaire de nouveaux cas d'Ebola confirmés au Liberia. Données:OMS

Ebola 2014: évolution en Sierra Leone

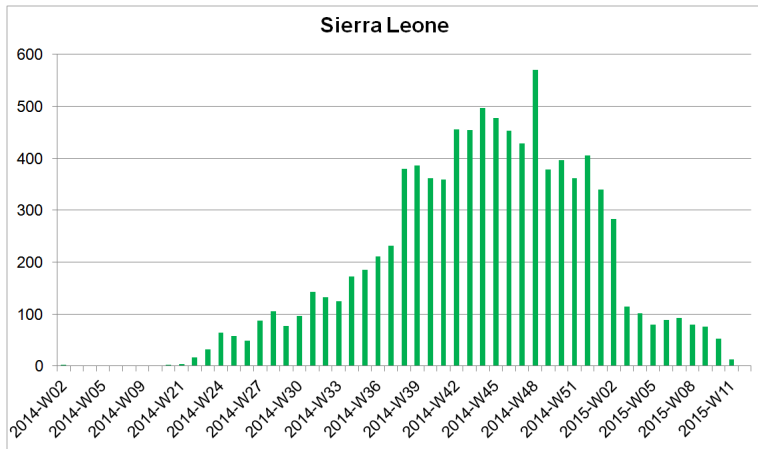


Figure: Nombre hebdomadaire de nouveaux cas d'Ebola confirmés en Sierra Leone. Données:OMS

Application numérique: Test [Super-critique] vs. [Sous-critique] pour un processus GW Poisson/Gamma

- ▶ Privilégiant les valeurs petites de α et β pour préserver la variabilité *a priori* du paramètre inconnu, nous avons finalement retenu $\alpha = 3$ et $\beta = 2.674$.
- ▶ Facteur de Bayes

$$B_n^{\text{crit}}(\mathbf{x}_n) = \frac{\left(\frac{\mathbb{P}(\theta \in H^{\text{sup}} | \mathbf{x}_n)}{\mathbb{P}(\theta \in H^{\text{sub}} | \mathbf{x}_n)} \right)}{\left(\frac{\mathbb{P}(\theta \in H^{\text{sup}})}{\mathbb{P}(\theta \in H^{\text{sub}})} \right)}$$

Test [Super-critique] vs. [Sous-critique] pour un processus GW Poisson/Gamma

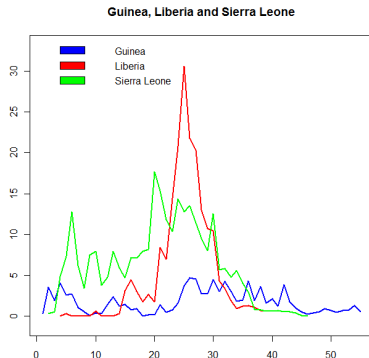
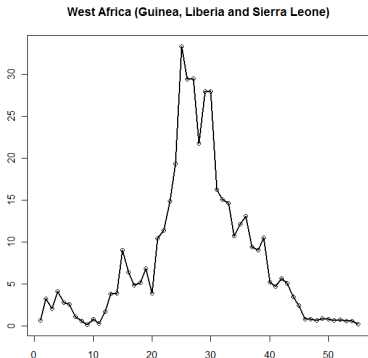


Figure: Evolution hebdomadaire du Facteur de Bayes associé au Problème de catégorisation [Sur-critique vs. Sous-critique] (échelle logarithmique, base 10) en Guinée, au Libéria et en Sierra Leone, entre la semaine 9 de 2014 et la semaine 12 de 2015. Données: OMS