La philosophie et l'expérience de la statistique bayésienne

Andrew Gelman

Département de Statistique et Département des Sciences Politiques, Columbia University
En visite à Sciences Po, Paris

En collaboration avec Cosma Shalizi (Carnegie Mellon University)

15 fevrier 2010

La statistique et la philosophie de la science

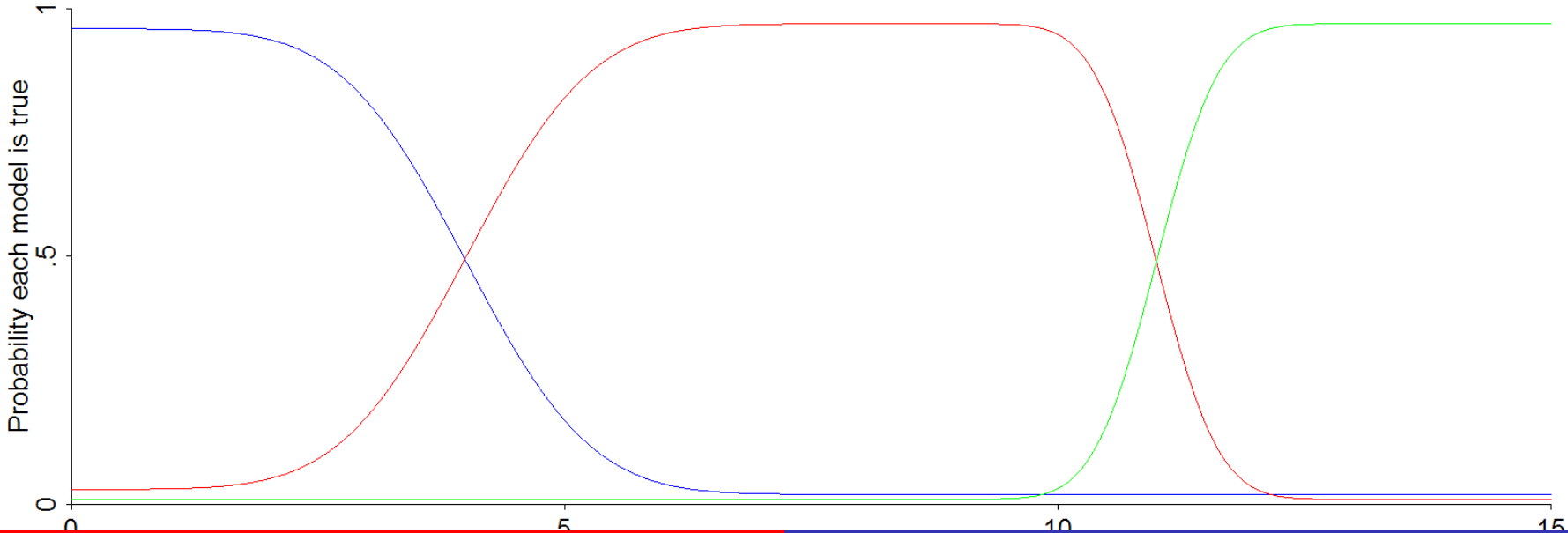
- L'induction
- La méthode hypothètico-déductive (Popper)
- La science normale et des révolutions scientifiques (Kuhn)
- Les programmes de recherche scientifique (Lakatos)
- Le rôle central de la statistique
- La correspondance usuelle:
  - L'induction = l'inférence bayésienne
  - L'inference deductive = les tests d'hypothèse classique

La philosophie bayésienne conventionnelle: 1

- La science avance par l'induction
- Dans la notation statistique: du cas particulier (les "données," y) aux généralités (les "paramètres," θ)
- L'inference bayésienne *dans* un modèle: $p(\theta|y) \propto p(\theta)p(y|\theta)$
  - C'est un *exemple* de l'inference inductive, et un *modèle* de toute l'inference inductive et rationnelle
  - Le progrès de la probabilité avant ("a priori") à la probabilité après ("a posteriori")

La philosophie bayésienne conventionnelle: 2

- L'inference bayésienne *dans* un modèle: $p(\theta|y) \propto p(\theta)p(y|\theta)$
- L'inference bayésienne *entre* des modèles:
  - Les modèles (hypothèses) $H_1, \ldots, H_K$
  - Pour chaque hypothèse:
    $p(H_k|y) \propto p(H_k)p(y|H_k) = p(H_k) \int p(y|\theta, H_k)p(\theta|H_k)d\theta$
- Le progrès scientifique, c'est le changement des probabilités des hypothèses:



La philosophie bayésienne conventionnelle: pour et contre

- Pour
  - C'est une histoire agréable du progrès
  - Elle est en accord avec la statistique bayésienne
- Contre
  - On ne peut pas prévoir tous les modèles possibles en avance
  - Un problème technique de définir les probabilités des modèles après les données
  - Dans notre expérience, la science avance par la déduction et la réfutation, pas par l'induction

Les pas de l'analyse des données bayésienne

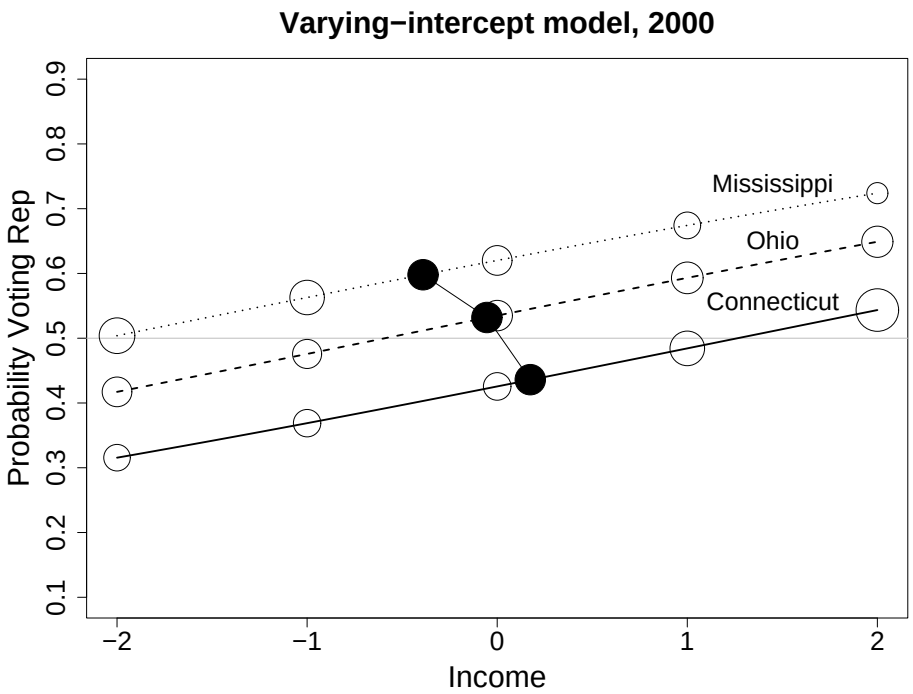
- Faire et proposer un modèle
- L'inférence
- Comparer les inferences du modèle avec les données et autre information ("les tests posterieurs prédictifs")
- La philosophie
  - C'est hypothètico-déductive, pas inductive
  - Plus d'explications après un exemple

Un exemple de l'analyse des données bayésienne

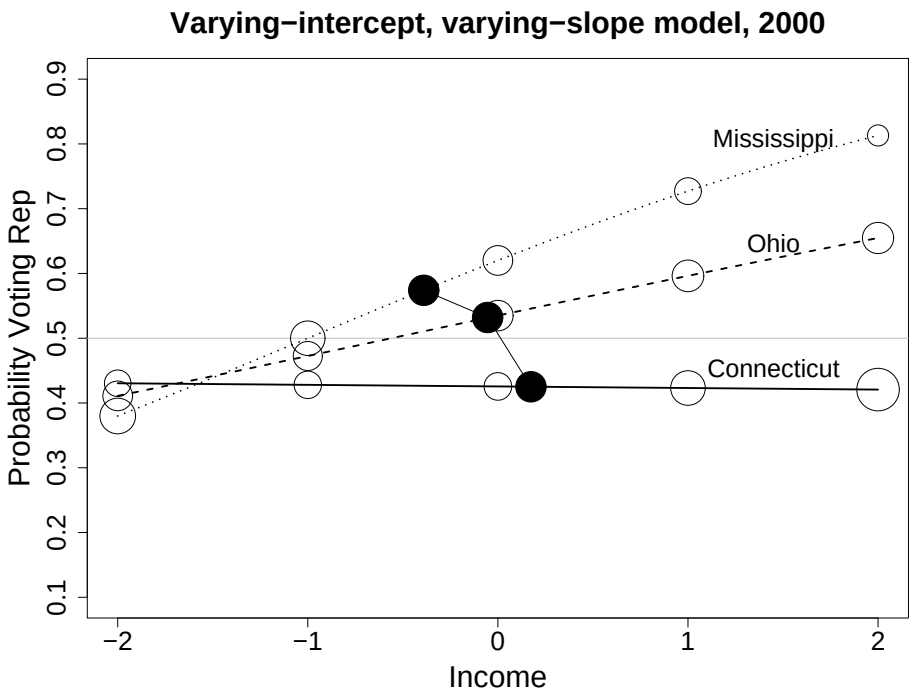
- On a intérêt aux connections entre la politique et l'inégalité economique
- Nous avons étudié les votes des riches et pauvres dans les 50 états de l'Amerique
- Je vous montrerai une série des modèles

États riches et pauvres, électeurs riches et pauvres

Le premier modèle

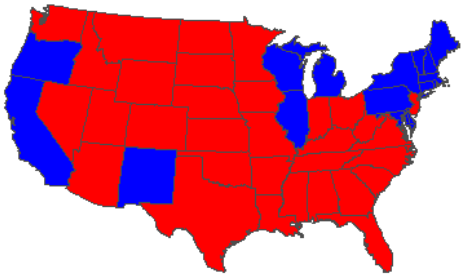


Le deuxième modèle

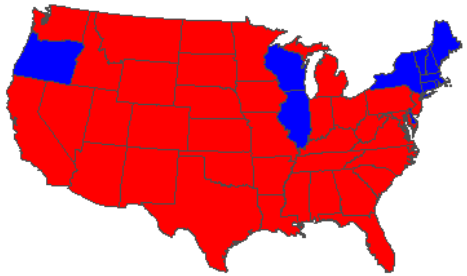


L'inférence . . .

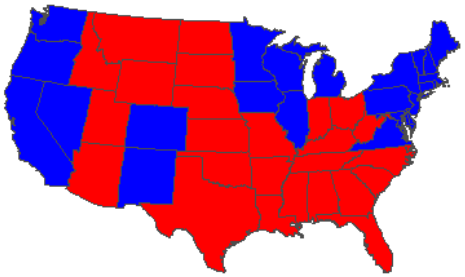
State winners in 2008 (rich voters only)



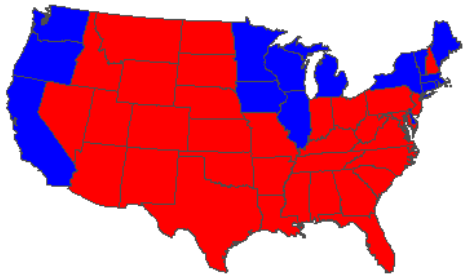
State winners in 2008 (rich Whites only)



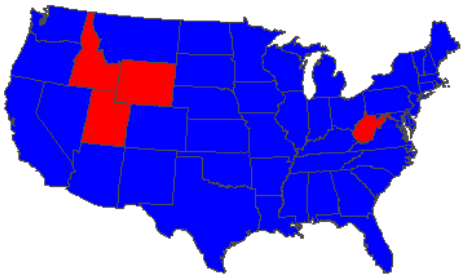
State winners in 2008 (middle-income voters)



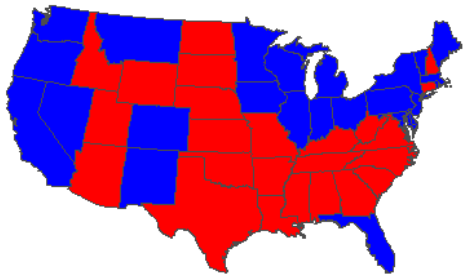
State winners in 2008 (middle-income Whites)



State winners in 2008 (poor voters only)



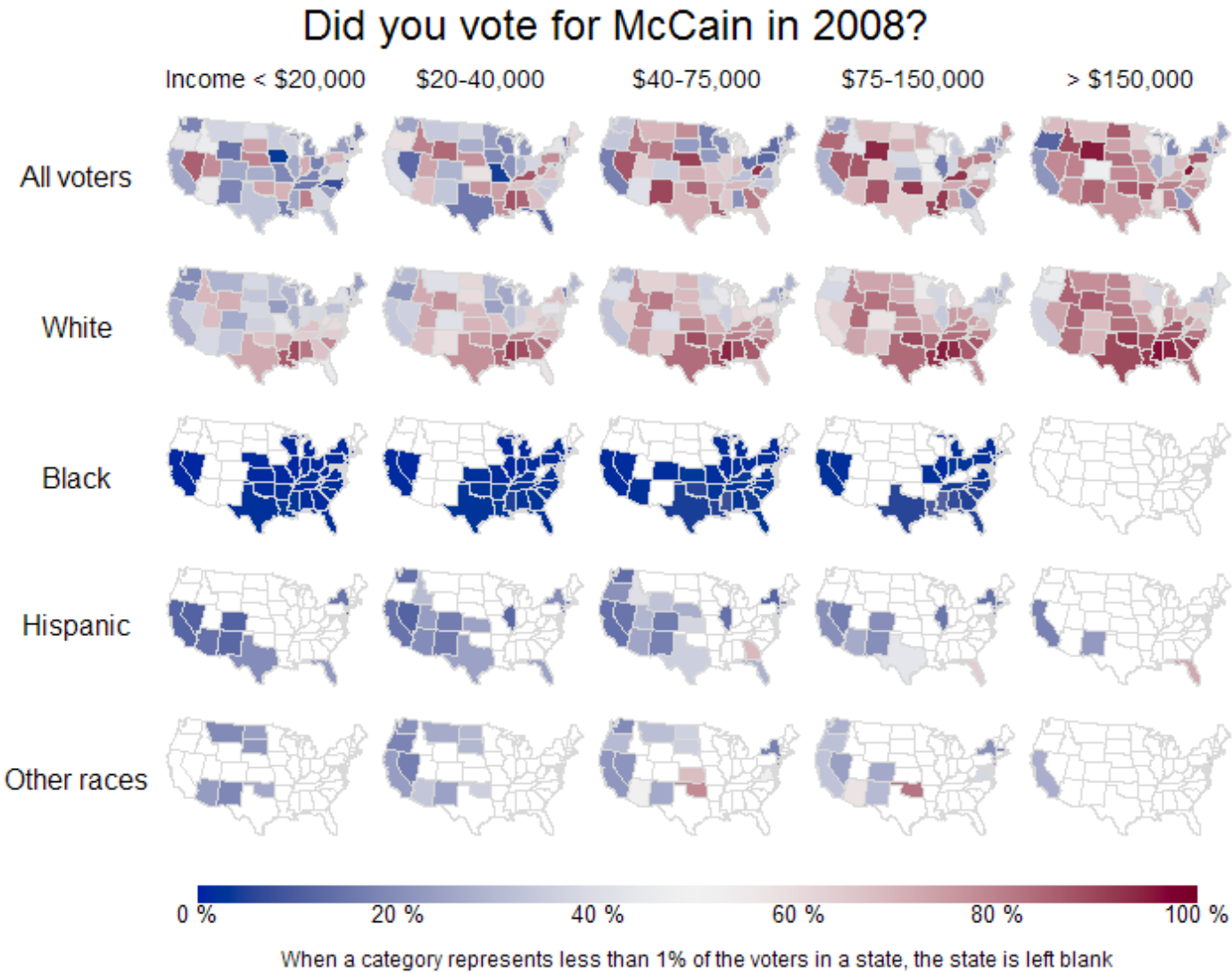
State winners in 2008 (poor Whites only)



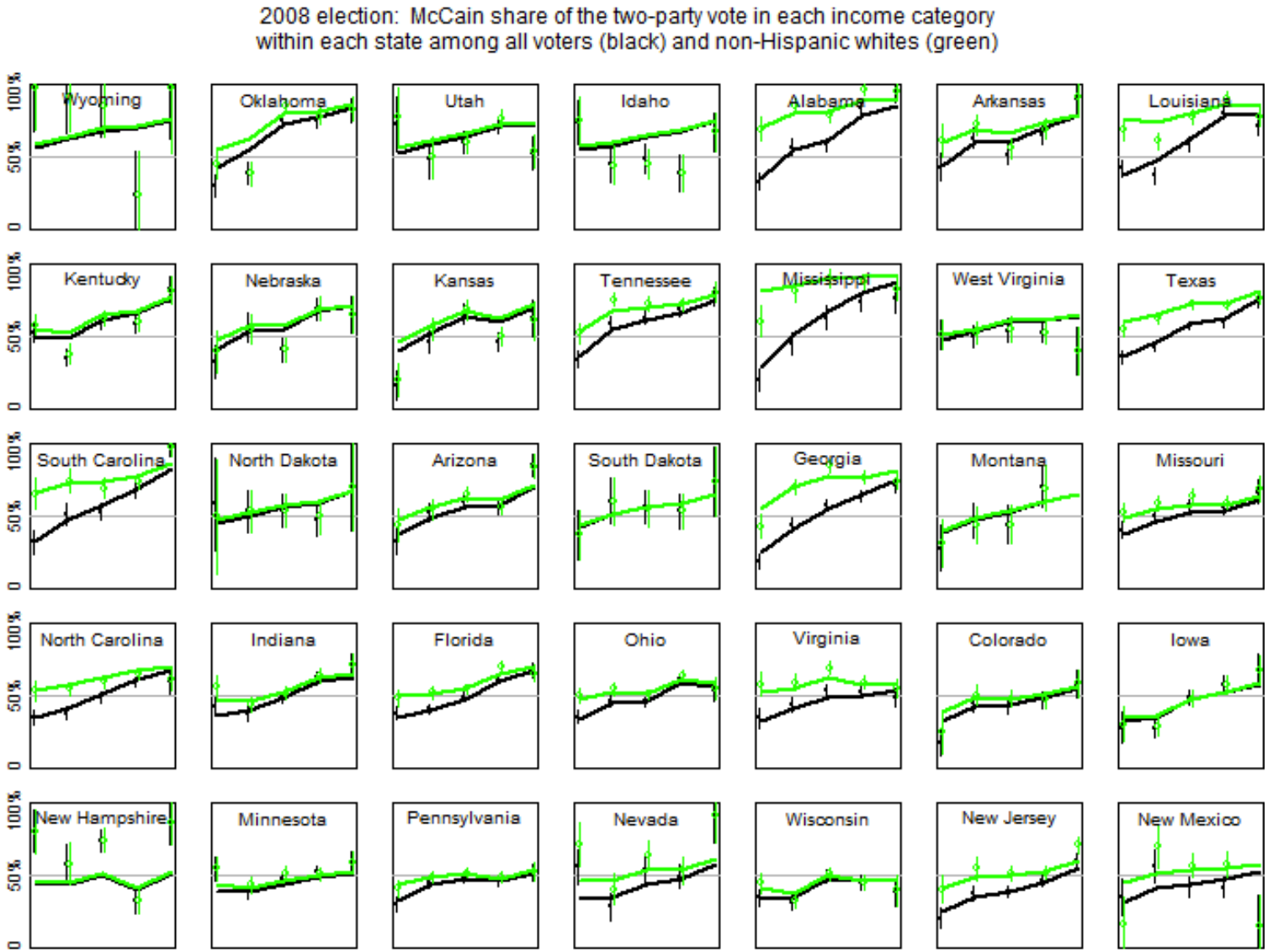
. . . et la réfutation!

- Les critiques du blogueur "Daily Kos":
  - Les critiques des inférences:
    "While Gelman claims only the under-$20K white demo went for Obama, the results were far different. Per the exit poll – real voters – Obama won all whites: 54-45 percent for those making under $50K, and 51-47% for those making over$50K. . . . New Hampshire is solidly Blue unlike Gelman's maps, 58-40 – one of the most obvious misses in Gelman's analysis. . . ."
  - Les critiques de la méthode:
    "Gelman inexplicably avoids using exit poll data . . . while exit polls have their own margin of errors and sample composition problems, they sure as heck beat anything done over the telephone."

Après le changement du modèle



Un graphe pour comprendre et critiquer nos inférences



Notre philosophie

- L'analyse des données bayésienne, c'est déductif:
  - On fait un modèle, on explore ses implications, on le teste
  - Quand le modèle a des problèmes serieux, on doit l'augmenter ou le remplacer
  - C'est comme la philosophie des tests classiques (Mayo, 1996), aussi les idées de Popper sur la réfutation
- La loi avant ("a priori")
  - Pour nous, la loi "a priori" n'est pas subjective; en revanche, elle est une hypothèse pour utiliser jusqu'a un meilleur modèle est necessaire
  - On l'*utilize*, on ne la *croit* pas!

Résumé de la philosophie bayésienne conventionnelle

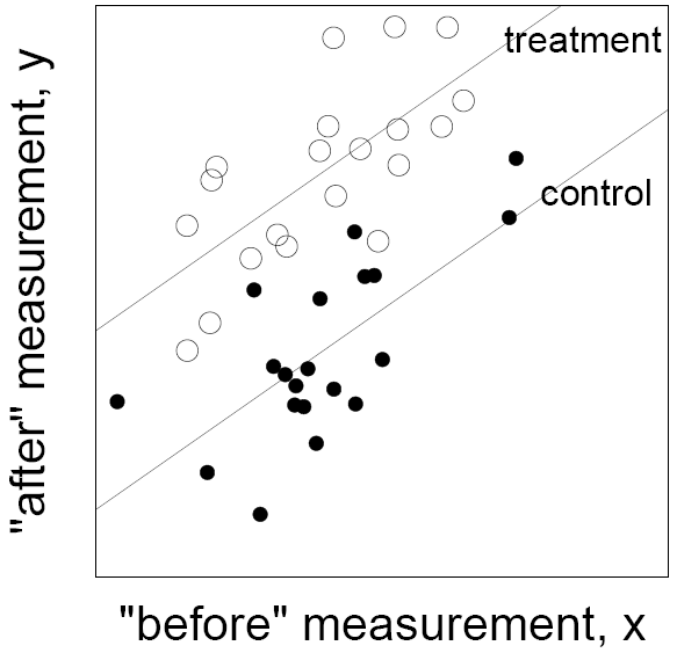
- De Wikipédia française sur "inférence bayésienne":
  "On nomme inférence bayésienne la démarche logique permettant de calculer ou réviser la probabilité d'une hypothèse. . . . L'inférence bayésienne est particulièrement utile dans les problèmes d'induction."
- De Wikipédia anglaise sur "Bayesian inference":
  "Bayesian inference uses aspects of the scientific method, which involves collecting evidence that is meant to be consistent or inconsistent with a given hypothesis. As evidence accumulates, the degree of belief in a hypothesis ought to change. With enough evidence, it should become very high or very low. . . . Bayesian inference uses a numerical estimate of the degree of belief in a hypothesis before evidence has been observed and calculates a numerical estimate of the degree of belief in the hypothesis after evidence has been observed."

Pourquoi nous ne sommes pas en accord avec la philosophie conventionelle

- Nous faisons l'inference déductive *dans* un modèle
- Nous pouvons trouver des problèmes d'un modèle, sans avoir une alternative
- Même si on peut écrire une liste des modèles possibles, il y a un problème serieux et complet de définir et calculer les probabilités des modèles après les données
  - La loi "a posteriori" de chaque modèle dépend entièrement sur les paramètres qui influencerent l'inférence dans le modèle: $p(H_k|y) \propto p(H_k)p(y|H_k) \propto p(H_k) \int p(y|\theta, H_k)p(\theta|H_k)d\theta$
  - Si on change $p(\theta|H_k)$ à l'extérieur de la région de θ où $p(y|\theta, H_k)$ est grand, on peut changer $p(H_k|y)$ sans changer essentiellement l'inférence pour $\theta|H_k$

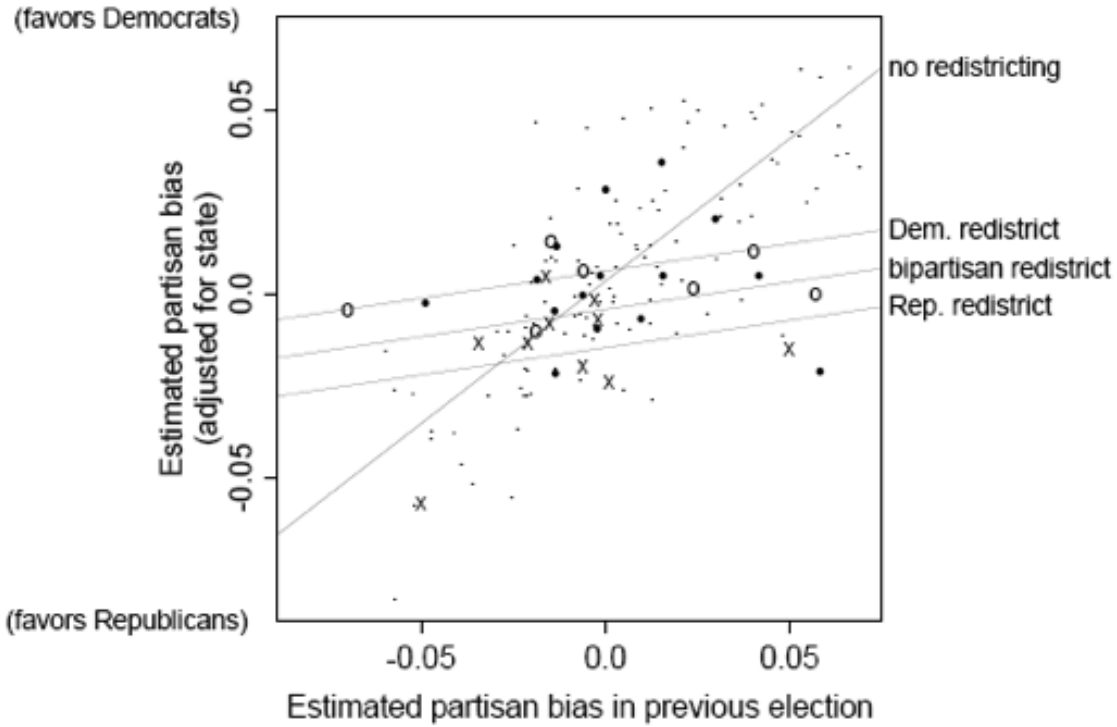
Un autre exemple

- On s'intéresse aux effets du redécoupage électoral
- Nous avons trouvé les données sur la partialité des redécoupages et leurs contrôles partisans
- Nous avons commencé avec le modèle défaut:



Le progrès depuis la déduction et la réfutation

- Nous avons fait un graphe:



- En fait, l'effet n'est pas constant
- La nouvelle histoire, c'est important: le redécoupage diminue la partialité partisane

L'erreur du mélange discret des hypothèses

- La *comparaison des hypothèses*, c'est un problème moins important que ce que tout le monde pense
- Examiner les exemples précédents
- Un exemple plus: les modèles de l'idéologie politique et les votes
  - On a comparé deux hypothèses:
    - Le modèle de la proximité: on préfère le partie qui est plus proche de lui-même
    - Le modèle de la direction: on préfère le partie qui est dans la même direction, mais plus loin du centre
  - Au lieu de comparer les deux hypothèses, ou en faire un mélange, nous recommandons faire un modèle plus grand qui comprend les deux explications des préfèrences politiques

Popper + Kuhn + Lakatos = Bayes

- On fait un modèle et l'utilise pour les inférences déductives (Popper)
- On utilise les déductions pour faire des déclarations fortes qu'on essaie de réfuter par les données (Popper)
- On améliore ou ajoute au modèle, si nécessaire
- Dans la philosophie de Kuhn, ce sont la *science normale* et la possibilité d'une *révolution*
- Dans la philosophie de Lakatos, ce sont les pas d'un *programme de récherche* avec les hypothèses *centrales* et *auxiliare*
- On peut utiliser l'inférence bayésienne pour changer les modèles *sans* avoir besoin de calculer (ou imaginer) les probabilités des modèles!

La nature fractale des révolutions scientifiques

- Dans la vie scientifique, on a des révolutions a toutes les périodes temporales
  - Dans un paradigme (un modèle, H), on fait la science normale ($p(\theta|y, H)$)
  - Les tests (les comparaisons des données y à la loi des prédictions, $p(y^{\text{rep}}|H, y)$) sont les possibilités de révolution
  - Le nouveau modèle, c'est une amélioration (dans l'ancien paradigme) ou, de temps en temps, un nouveau paradigme
- Les révolutions sont fractales:
  - Les micro-révolutions chaque 5 minutes
  - Les révolutions plus grandes chaque semaine, mois, année, ...

Résumé philosophique

- On peut accepter l'importance de la statistique bayésienne, sans accepter l'idée de l'induction scientifique
- L'analyse des données bayésienne n'est pas la même chose de l'inference bayésienne des modèles discrètes
- Dans l'analyse bayésienne, on peut (on doit!) faire les tests d'hypothèse
- Les idées de Popper/Kuhn/Lakatos correspondent aux bonnes pratiques statistiques

Résumé statistique

- Un modèle bayésien, c'est une hypothèse, pas une croyance
- Le but d'un test statistique, c'est la compréhension des problèmes d'un modèle, pas un rejet
- Nous recommandons les tests graphiques, pas les "p-values"
- Quand on considère quelques hypothèses, nous recommandons faire un grand modèle continu, pas un mélange discret