

FESTIVAL D'ASTRONOMIE DE FLEURANCE 2009

# LA SCIENCE DE L'INCERTITUDE



**DR ROBERTO TROTTA**  
**IMPERIAL COLLEGE LONDON**

**WWW.ROBERTOTROTTA.COM**

1

## Sommaire

---

Imperial College  
London

- Science et probabilité
- Les lois de la probabilité
- Quelques applications pratiques
- Applications cosmologiques
- Conclusion

---

Roberto Trotta

2

## Science et probabilité

3

## Qu'est-ce que la science?

Imperial College  
London

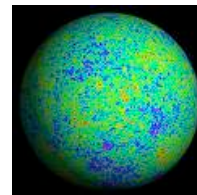
- La science est souvent perçue comme une méthode capable d'établir ce qui est "vrai" dans l'Univers
- Le métier des scientifiques est plutôt celui de réduire (et contrôler) le degré d'incertitude de nos modèles de la réalité.
- Essayez vous-mêmes! Quelle est la probabilité que...
  - ... le soleil se lève demain?
  - ... une pièce est équilibrée si on obtient 10 fois face sur 10 lancements?
  - ... la vie existe ailleurs dans l'Univers?
  - ... le Big Bang s'est produit il y a 13.7 milliards d'années?

Roberto Trotta

4

# L'astronomie comme science ?

- Astronomie, astrophysique et cosmologie s'occupent d'objets qu'on ne peut pas manipuler !



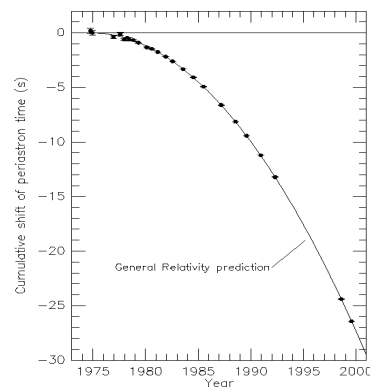
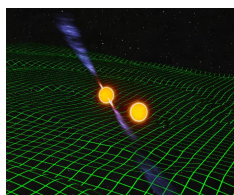
Roberto Trotta

5

# Théorie et observations

- Les lois naturelles sont écrites en langage mathématique : formel et exact
- L'observation des phénomènes est toujours soumise à des erreurs de mesure et au bruit.
- Par exemple :

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

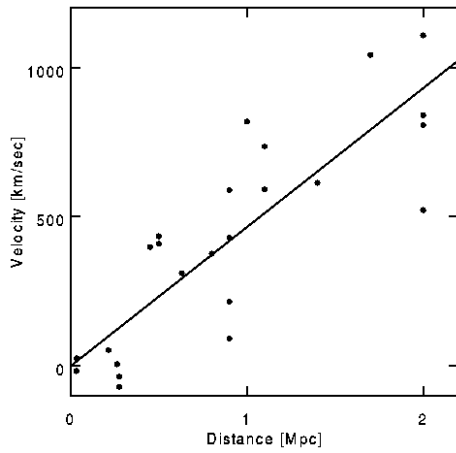


Roberto Trotta

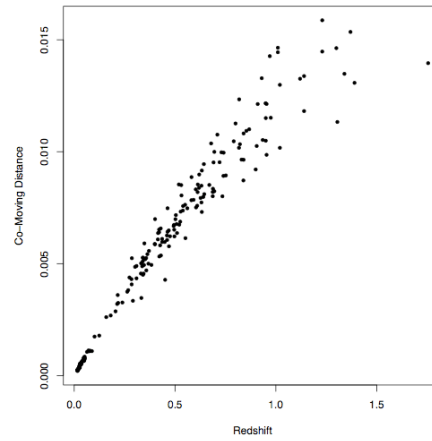
6

# L'expansion de l'Univers

Hubble (1929)



Davis et al (2007)



$$v = H_0 R$$

Roberto Trotta

Les lois de la probabilité

## Déduction: le domaine de la logique pure

Toutes les pommes poussent sur  
des arbres

Ce fruit est une pomme

Donc, elle pousse sur un arbre



Général

Particulier

## Induction: le domaine de la science

Tous les corbeaux observés  
sont noirs

Cet oiseau est un corbeau

Donc, il est noir



Expérience passée

Observation

Induction

Roberto Trotta

9

# Qu'est-ce que la probabilité?

- La notion fréquentiste de probabilité :  
“Probabilité = le nombre de fois qu’un événement se produit dans une séquence infinie de répétitions équiprobables”
- Une définition problématique :
  - circulaire : on assume “équiprobable” pour définir “probable”
  - Elle ne s’applique pas a des situations uniques
  - Valable seulement si l’on a une séquence “infinie”

Roberto Trotta

10

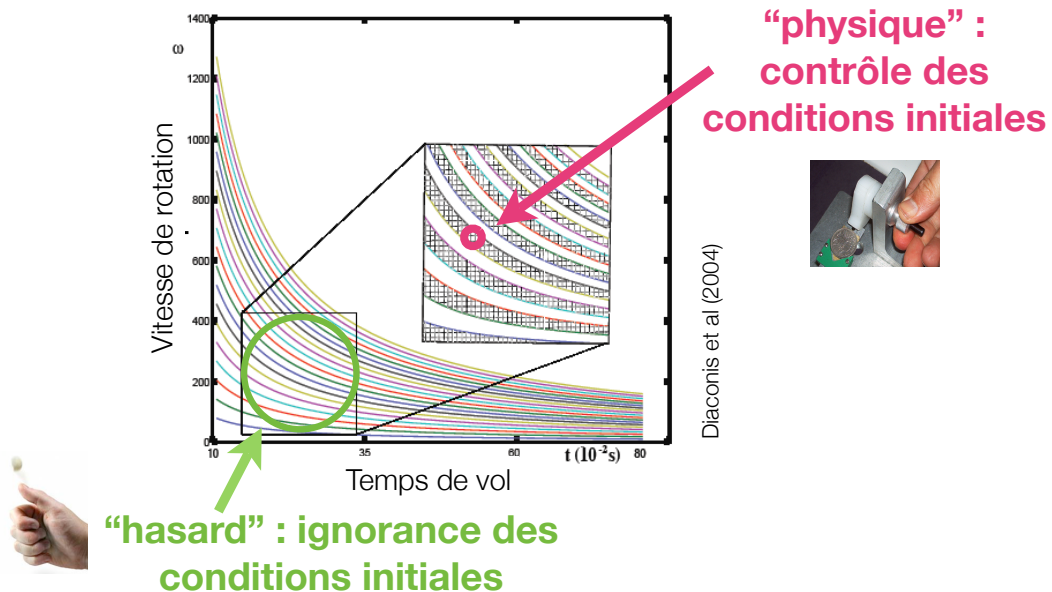
# Qu'est que le hasard?

- Exemple typique: lancement d'une pièce de monnaie (pile ou face)
- Est-ce que la pièce est truquée? But : déterminer avec une séquence de lancements si la probabilité d'obtenir face est  $p_F = 1/2$
- Est-ce qu'on peut dire qu'on veut déterminer une propriété physique de la pièce?
- Qu'est-ce que "lancer au hasard" signifie ?

Roberto Trotta

11

# Hasard ou physique ?



12

- Deux différentes écoles :  
l'école fréquentiste (probabilité = fréquence)  
et l'école Bayésienne
- L'école Bayésienne utilise la probabilité comme un système formel pour traduire numériquement un degré de connaissance sur des propositions, tout en prenant en considération les informations disponibles.

- Cette approche peut exprimer la probabilité d'événements uniques et d'événements répétables
- On reconnaît un degré de subjectivité dans l'expression des degrés de connaissance
- Le hasard devient une expression d'un état d'ignorance/manque d'information
- Possibilité de formuler des prédictions pour le futur

- Notation:  
Soient A et B deux événements.

Soit  $P(A)$  la probabilité de A

Soit  $P(B)$  la probabilité de B

Soit  $P(A|B)$  la probabilité de A en sachant B (= probabilité conditionnelle)

Soit  $P(B|A)$  la probabilité de B en sachant A (= probabilité conditionnelle)

Soit  $P(A,B)$  la probabilité que A et B aient tous les deux lieu

- La définition de probabilité conditionnelle donne:

$$P(B|A)P(A) = P(A,B) = P(A|B)P(B)$$

Roberto Trotta

15

- De la définition  $P(B|A)P(A) = P(A,B) = P(A|B)P(B)$  on obtient, en rajoutant toute information qui est disponible (I) :

$$P(A|B, I) = \frac{P(B|A, I)P(A|I)}{P(B|I)}$$

Le théorème de Bayes



Thomas Bayes (ca. 1701-1761)

*T. Bayes.*

Roberto Trotta

16



- En remplaçant  $A \rightarrow H$  (**l'hypothèse**) et  $B \rightarrow d$  (**les données**) on obtient

The diagram illustrates the Bayesian equation  $P(H|d, I) = \frac{P(d|H, I)P(H|I)}{P(d|I)}$  with four callouts in speech bubbles:

- probabilité a posteriori de H en sachant d** (yellow bubble, points to  $P(H|d, I)$ )
- fonction de vraisemblance de H** (orange bubble, points to  $P(d|H, I)$ )
- probabilité a priori de H** (blue bubble, points to  $P(H|I)$ )
- valeur de plausibilité de d** (green bubble, points to  $P(d|I)$ )

Roberto Trotta

17

$$P(H|d, I) = \frac{P(d|H, I)P(H|I)}{P(d|I)}$$

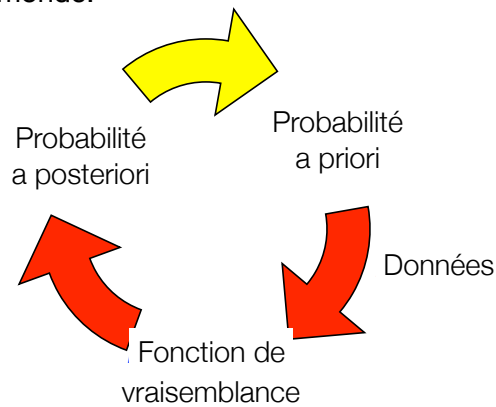
- **Probabilité a priori de H** : ce que l'on sait à propos de  $H$  (avec information  $I$ ) avant de voir les données
- **Fonction de vraisemblance de H** : la probabilité d'obtenir des données  $d$  si l'hypothèse  $H$  est vraie
- **Probabilité a posteriori de H** : notre degré de connaissance sur  $H$  après avoir observé les données  $d$
- **Valeur de plausibilité de d** : une constante indépendante de l'hypothèse  $H$

Roberto Trotta

18

## L'importance du théorème de Bayes

- Le théorème de Bayes est la généralisation unique de la logique formelle dans la situation où il y a de l'incertitude sur les propositions
- Il représente une voie formelle d'exprimer l'apprentissage en conséquence de notre expérience du monde:

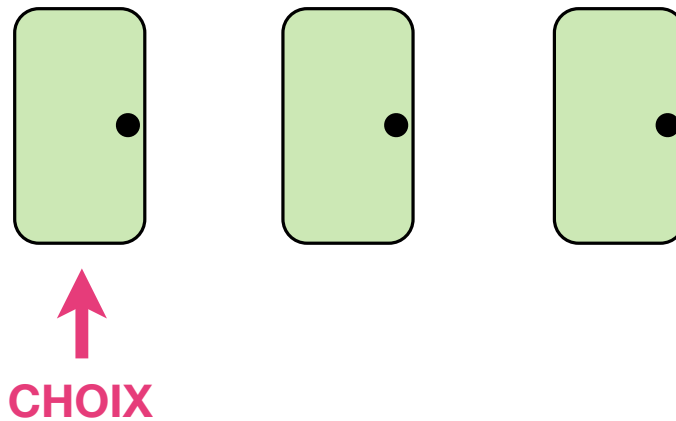


Roberto Trotta

## Quelques applications pratiques

## Application 1 : l'énigme des 3 portes

- Dans un jeu à la télé, le concurrent doit choisir une porte parmi 3, en sachant qu'il y a un prix derrière l'une des trois portes, tandis que les 2 autres n'ont rien du tout :

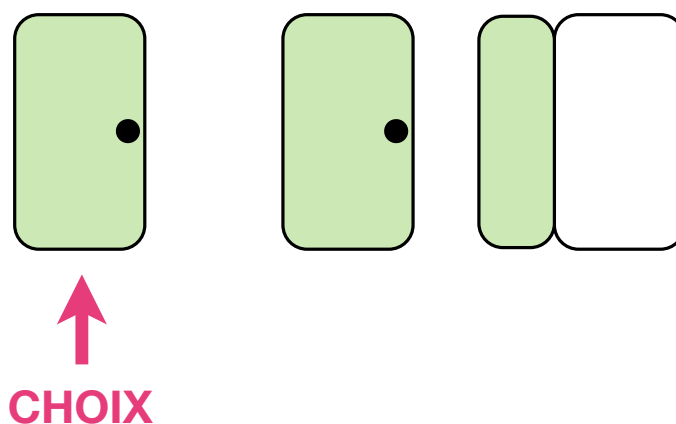


Roberto Trotta

21

## L'énigme des 3 portes (cont.)

- Le présentateur ouvre ensuite une des deux portes que le concurrent *n'a pas choisi*, et il montre qu'il n'y a rien derrière:

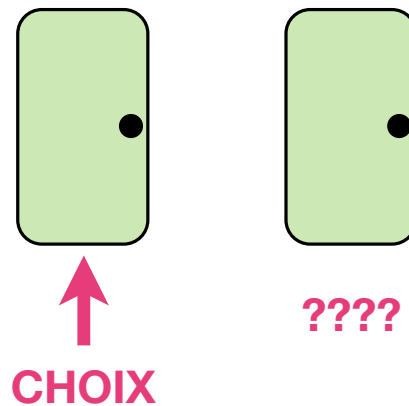


Roberto Trotta

22

## L'énigme des 3 portes (cont.)

- Le concurrent a maintenant l'option de changer son choix, et de choisir la porte qui n'a pas été ouverte par le présentateur. **Est-ce qu'il devrait changer son choix?**
- Est-ce que la réponse change si c'est un tremblement de terre qui ouvre la porte?



Roberto Trotta

23

## Le processus scientifique

- **Probabilité directe** : on donne un processus de génération de données et l'on veut prédire le résultat.

**Exemple** : une boîte contient  $N$  balles.  $B$  sont blanches et  $N-B$  sont rouges. On tire  $M$  balles au hasard (avec remplacement) : quelle est la probabilité d'obtenir  $K$  balles blanches?

- **Probabilité inverse** : on a observé le résultat d'un processus casuel et l'on veut reconstruire la probabilité d'une ou plusieurs variables cachées qu'on n'a pas observées.

**Exemple** : on a 10 boîtes, chacune avec 1, 2, 3, ..., 10 balles blanches (le reste étant des balles rouges). On prend 3 balles blanches d'une boîte quelconque. Quelle est la probabilité que cette boîte contienne 1, 2, 3, ..., 10 balles blanches?

- **Le processus scientifique utilise toujours des probabilités inverses!**

Roberto Trotta

24

## Application 2 : les 3 dés

- On a 3 dés, chacun avec 20 faces. Les numéros sur les faces de chaque dé sont distribués comme suit :

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	5	3	2	1	1	1	1	1	0	0	0	0	0	0	0
B	2	2	2	1	1	1	1	1	1	1	1	0	0	1	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Roberto Trotta

25

## Application 2 : les 3 dés (cont.)

- On prend un des trois dés au hasard, et on le lance 7 fois. On observe la séquence suivante :

2, 5, 1, 8, 4, 7, 3

Quelle est la probabilité qu'il s'agit du dé A?

Un exemple de probabilité inverse!

Roberto Trotta

26

## Application 3 : question de vie ou de mort

- David se rend à l'hôpital pour un contrôle de routine. Le test pour le virus *Frequentistus voracissimus* donne résultat positif.
- Le test a une fiabilité de 95% : si la personne est vraiment positive ( $M=1$ ), le test donne un résultat positif 95% des fois; si la personne est saine ( $M=0$ ), le test donne un résultat négatif 95% des fois.
- On sait aussi que 1% des personnes de l'âge de David ont le virus.
- **Quelle est la probabilité que David a le virus?**
  - **A** : 95%
  - **B** : moins de 95%
  - **C** : plus de 95%



Roberto Trotta

27

## Application 4 : M. Bayes au tribunal

- Mme Dupont a été retrouvée morte dans sa cuisine. M. Dupont se comporte de façon suspecte. Il est arrêté et inculpé de l'avoir tuée.
- La police découvre que M. Dupont avait été dénoncé par sa femme pour maltraitance. Le ministère public utilise cela comme preuve de la culpabilité de M. Dupont.
- Mais la défense a une opinion bien différente : "D'un point de vue statistique, seulement 1 homme sur 1000 qui bat sa femme se rend coupable d'homicide. Donc cela montre plutôt que M. Dupont est innocent!"
- **Qui a raison? Quelle est la probabilité que M. Dupont a tué sa femme?**



Roberto Trotta

28

## M. Bayes au tribunal (cont.)

- Pour résoudre ce problème, il nous est utile de considérer les informations suivantes :
- Aux Etats-Unis (~ 200 millions d'habitants) on estime que 2 millions de femmes sont battues par leur partenaire.
- En 1994, sur 4739 femmes tuées aux Etats-Unis, 1326 ont été tuées par leur partenaire.



???

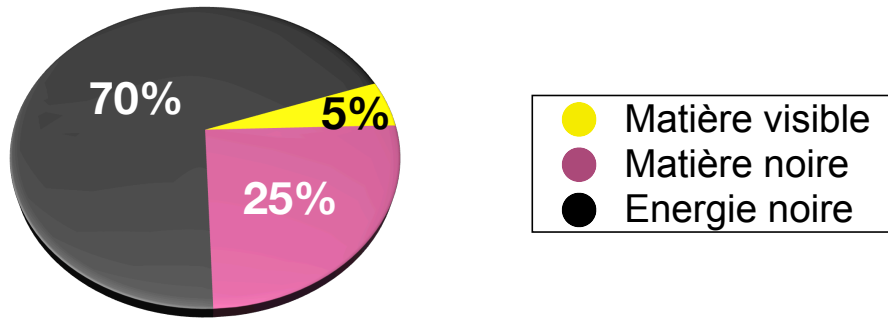


Roberto Trotta

Applications cosmologiques

## L'incertitude en cosmologie

- Les découvertes cosmologiques sont souvent présentées de façon simplifiée. Par exemple :



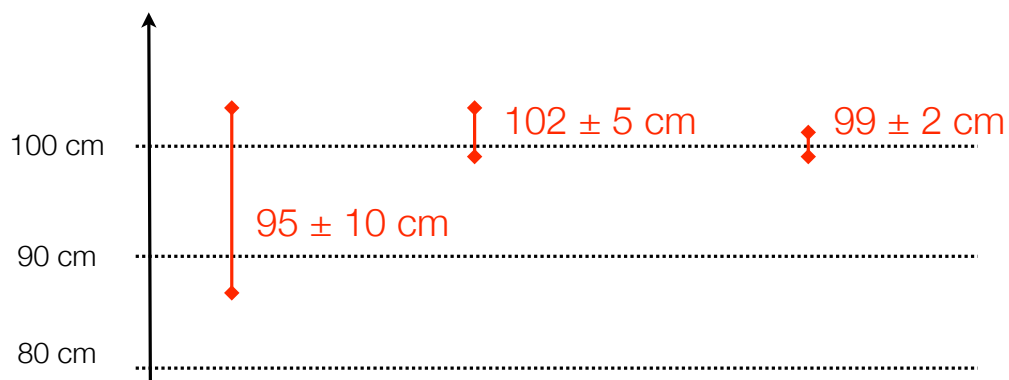
Mais toute mesure scientifique a une marge d'erreur!

Roberto Trotta

31

## La marge d'erreur est importante !

- Pour toute activité scientifique, il est fondamental de quantifier très précisément (!) la marge d'erreur des mesures
- On utilise pour cela les "barres d'erreur" :

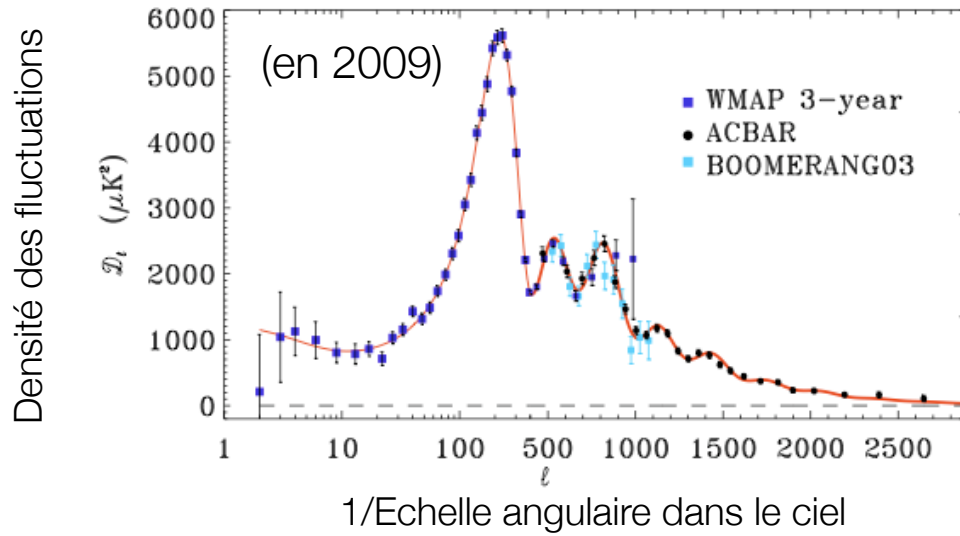
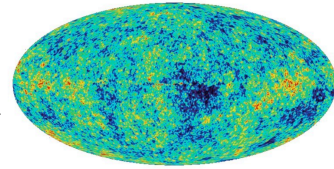


Roberto Trotta

32



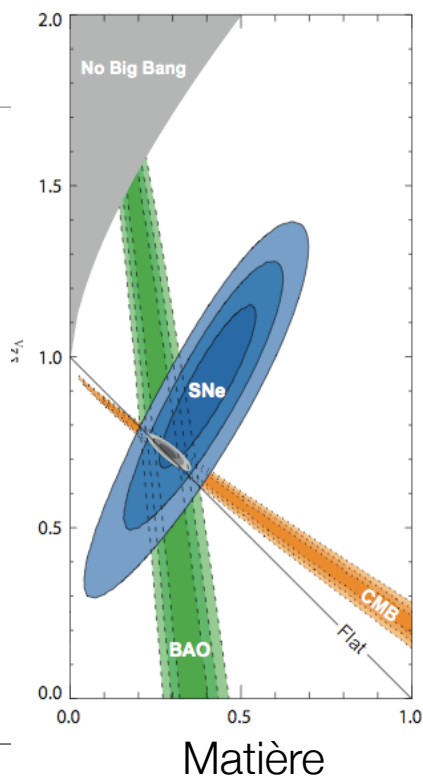
# Le modèle cosmologique standard



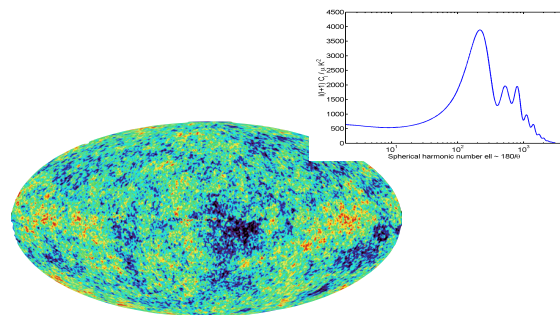
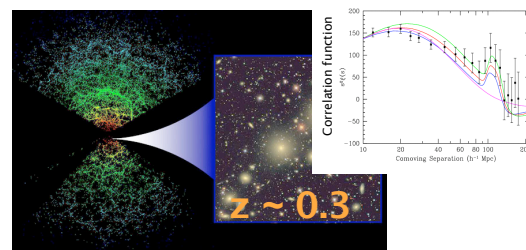
Roberto Trotta

33

Constante cosmologique



Imperial College  
London



Dr Roberto Trotta

34

## La fin de la cosmologie ?

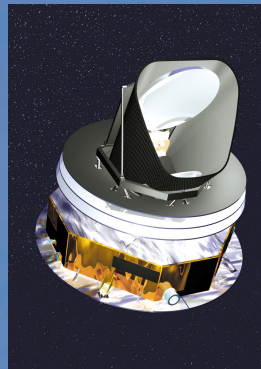
- On peut réduire les barres d'erreurs de deux façons :
- **En utilisant des instruments toujours plus précis** (réduction du bruit) :

cela est équivalent à employer des règles toujours plus précises  
(précision de 10 cm -> 1 cm -> 1 mm -> ...)

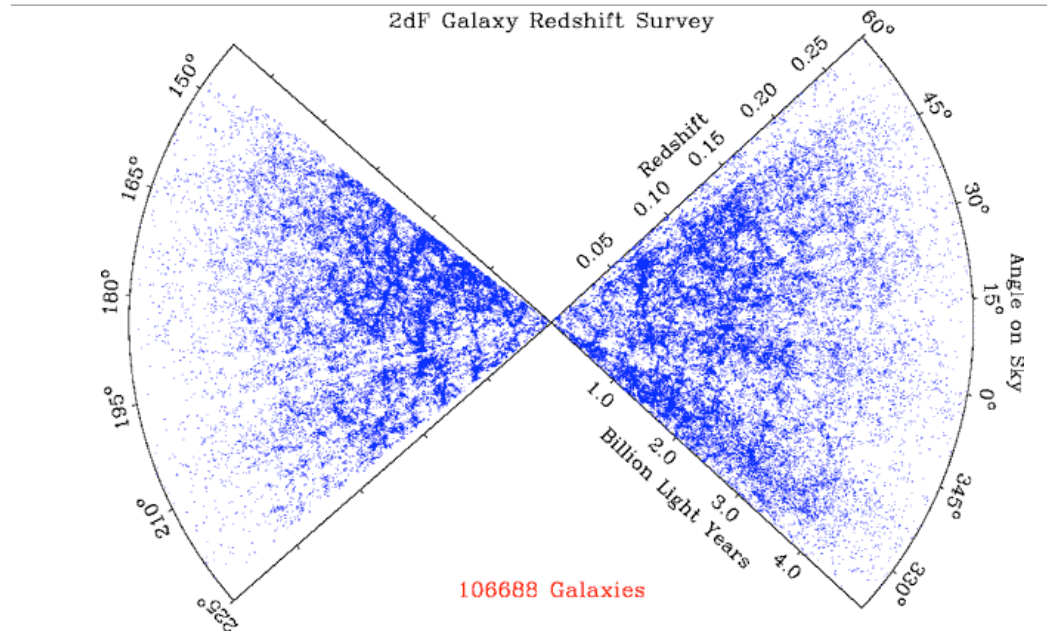
- **En mesurant de plus en plus d'objets** (galaxies, fluctuations dans la radiation cosmique à microondes, explosions de supernovae, etc) :

cela est équivalent à mesurer plusieurs fois le même objet avec la même règle, puis faire une moyenne des mesures  
(100 mesures avec une précision de 10 cm sont équivalentes à 1 mesure avec une précision de 1 cm)

## Lancement du satellite "Planck" en Mai 2009



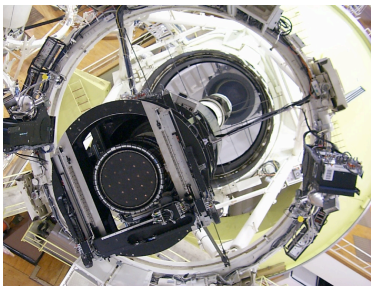
## Les structures à grande échelle



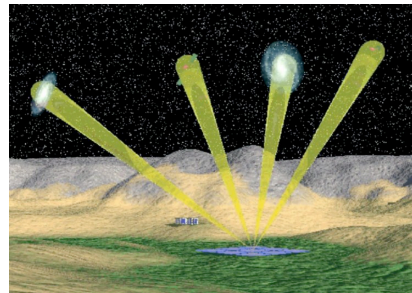
37

## Una carte globale de l'Univers

- 2dF Galaxy Redshift Survey – 250'000 galaxies (terminé)
- Sloan Digital Sky Survey – 1 million de galaxies (terminé)
- Large Synoptic Survey Telescope (2018?) – 1 milliard de galaxies
- Square Kilometer Array (2020?) – dizaines de milliards de galaxies



Télescope 2dF



Square Kilometer Array : radiotélescope

38

- La science procède grâce aux deux piliers de la théorie et de l'observation
- Les deux sont mis en communication par l'analyse statistique
- La formule de Bayes est un instrument mathématique pour formaliser l'apprentissage de l'expérience
- Tous les jours nous sommes confrontés à des situations incertaines : souvent la manipulation correcte des probabilités est très contre-intuitive !
- Dans la pratique de la cosmologie, le fait d'avoir seulement un Univers observable limite (même en principe !) la quantité de données accessibles
- La science de l'incertitude jouera donc un rôle toujours plus important pour notre compréhension de l'Univers.