**Motivation:**
Representing distributions more compactly and often more quickly than a bag of samples from MCMC

---

Example: inferring dark matter

http://www.kaggle.com/c/DarkWorlds
http://homepages.inf.ed.ac.uk/imurray2/pub/12kaggle_dark/

These slides are for motivation. Details about dark matter non-examinable!
Answer is obvious

Could optimize likelihood of dark matter position

Faster than MCMC!

Want some way to report error bars though.

Usually dark matter locations *not* obvious...
Summarizing beliefs?

— Average/mean sample?
— Most probable sample?
— Cluster?

I have several answers. But still a research question. For this course:

Some complicated distributions most easily represented by samples

Then predict under each possible world
Lower dimensional example

A posterior over some quantity $\alpha$ from http://iopscience.iop.org/0004-637X/711/2/1157/

Might summarize with the vertical credible intervals containing 95% and 99% of probability mass.

Mean, mode, median?

Weirder example

Red vertical bar is the mean (not a probable point, note log scale)

Median? Mode?

from http://link.springer.com/chapter/10.1007%2F11736790_3

Gaussian approximations

Finite parameter vector $\theta$

$P(\theta \mid \text{lots of data})$ often nearly Gaussian around the mode

Need to identify which Gaussian it is: mean, covariance