Q&A - episode 1

Data Analytics

Jerzy Baranowski

What is prior probability?

Is it only a term about prior data set that we have?
Is it ONLY data, based on which we are creating model and trying to predict next results?

Prior probability

What we know before we see the data

- Prior probability encodes our information that we have about our process without looking at the data
- If we have good info it can strongly influence our inference
- We shouldn’t assign prior looking at the data, but…

Does Bayesian statistics always have two possibilities {0,1}?
For
example: success&failure, man&woman, ...?

Everything is possible!

Bayesian methods are used for

- Images
- Counting data
- Continuous data
- Time series data
- Natural language processing
- And much much more

Is „prior predictive distribution” done ONLY on Bayesian models?

Yes

Is Poisson some king of subgroup of Bayesian models?

Poisson distribution

When you have to count something

- Poisson is one of the distributions that are useful for modelling integer valued data
- They are not the only one
- They are used in classical statistics also, however often approximated by normals

What is going on with:

$$\lambda + 3\sqrt{\lambda} \approx 365?$$

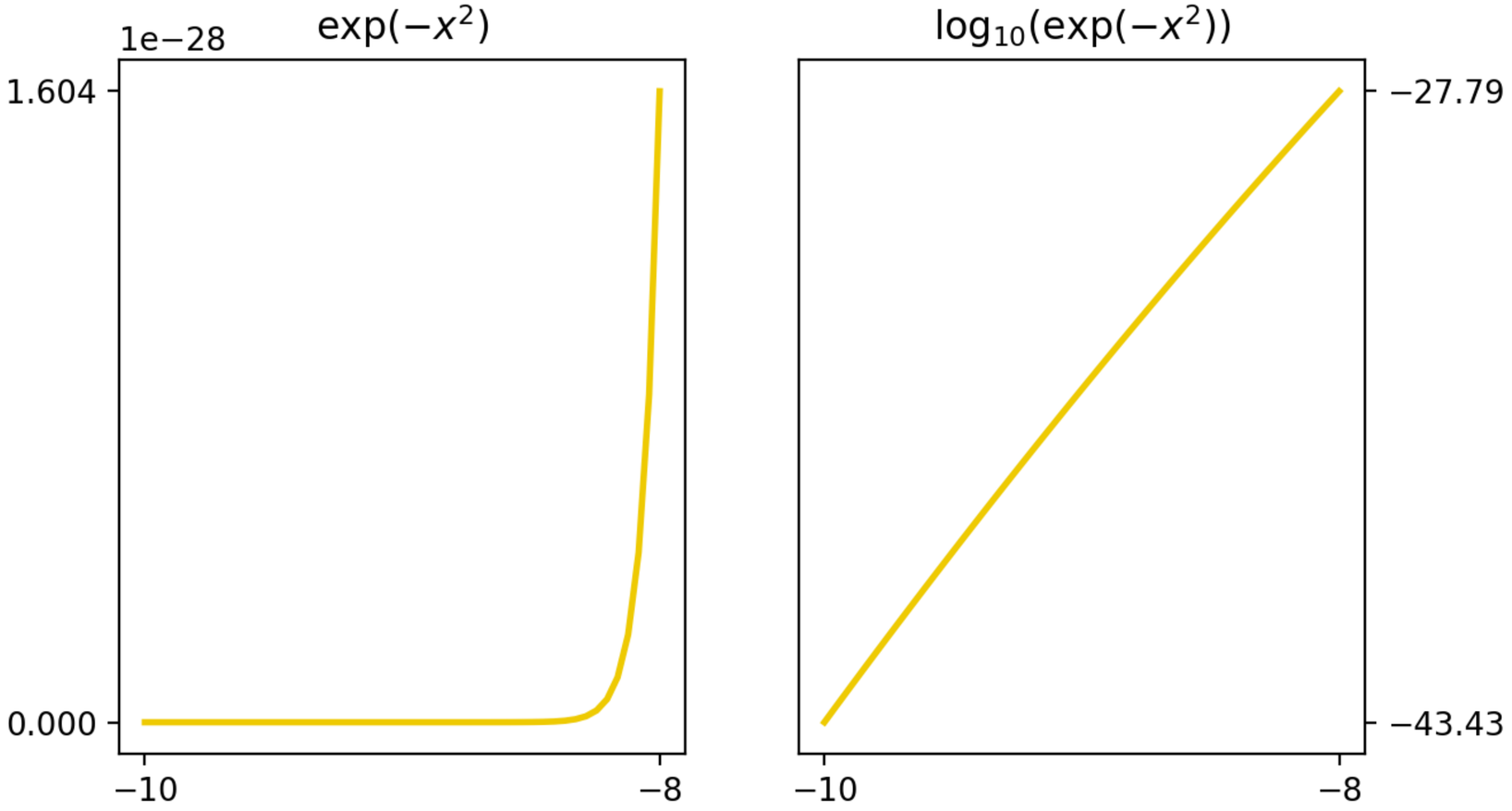
It’s generally from normal approximation

- There is a well known rule, about normal distributions that is, that:
  - 67% of probability mass is in the interval $(\mu - \sigma, \mu + \sigma)$
  - 95% of probability mass is in the interval $(\mu - 2\sigma, \mu + 2\sigma)$
  - 99% of probability mass is in the interval $(\mu - 3\sigma, \mu + 3\sigma)$
- Poisson is well aproximated by normal, and for Poisson $\mu = \lambda,\ \sigma = \sqrt{\lambda}$

Why we use logarithms in computation?

Generally because of rounding errors

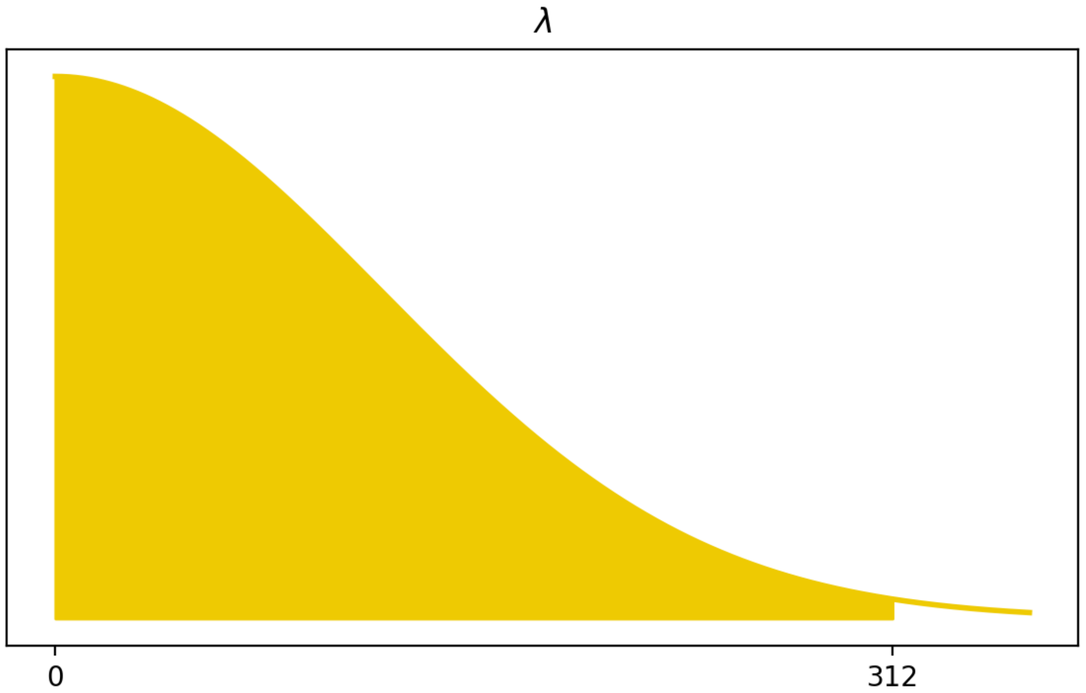
And flatness



Why we set up only upper bound when tuning prior probability distribution for λ or θ?

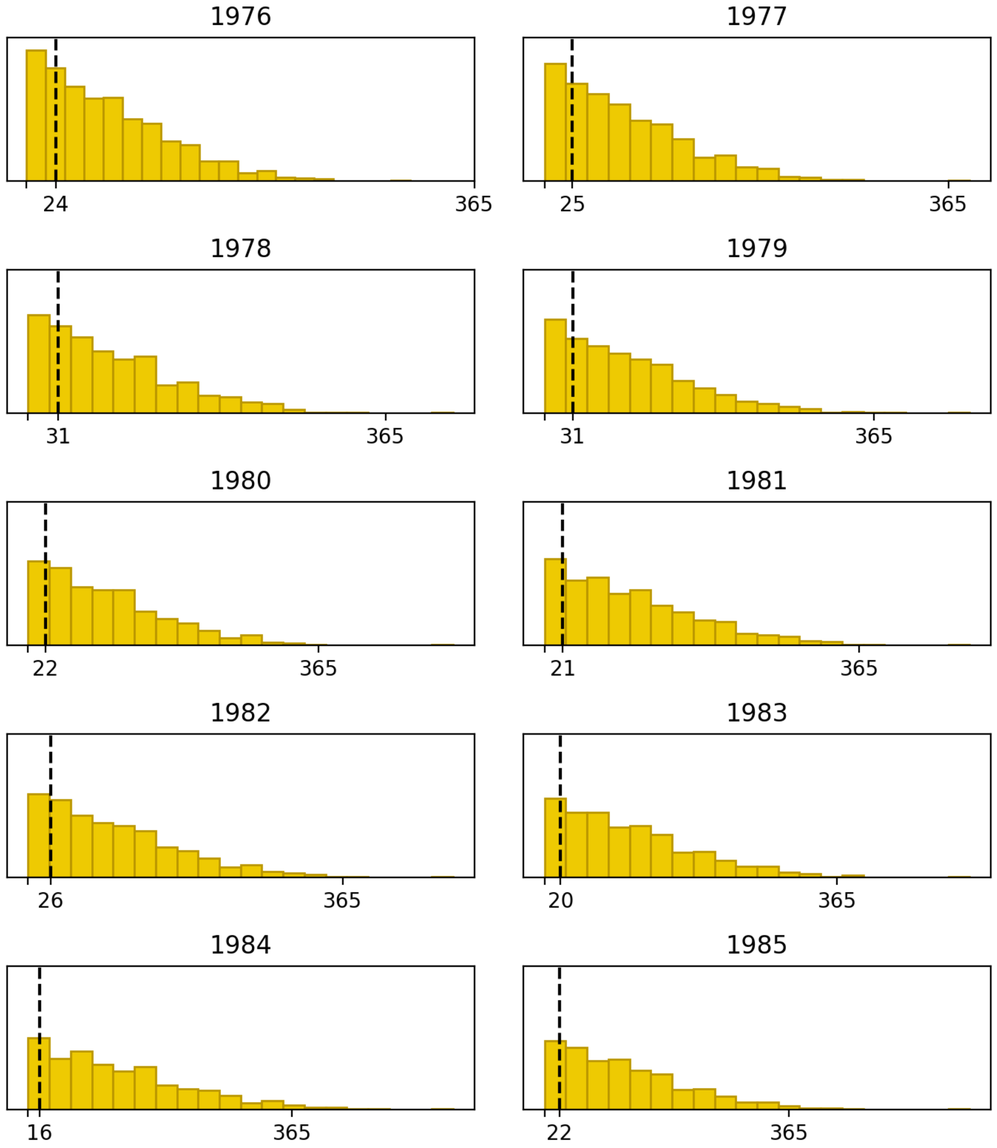
Because its a half normal distribution

Distribution is naturally constrained from zero, so no need for lower bound

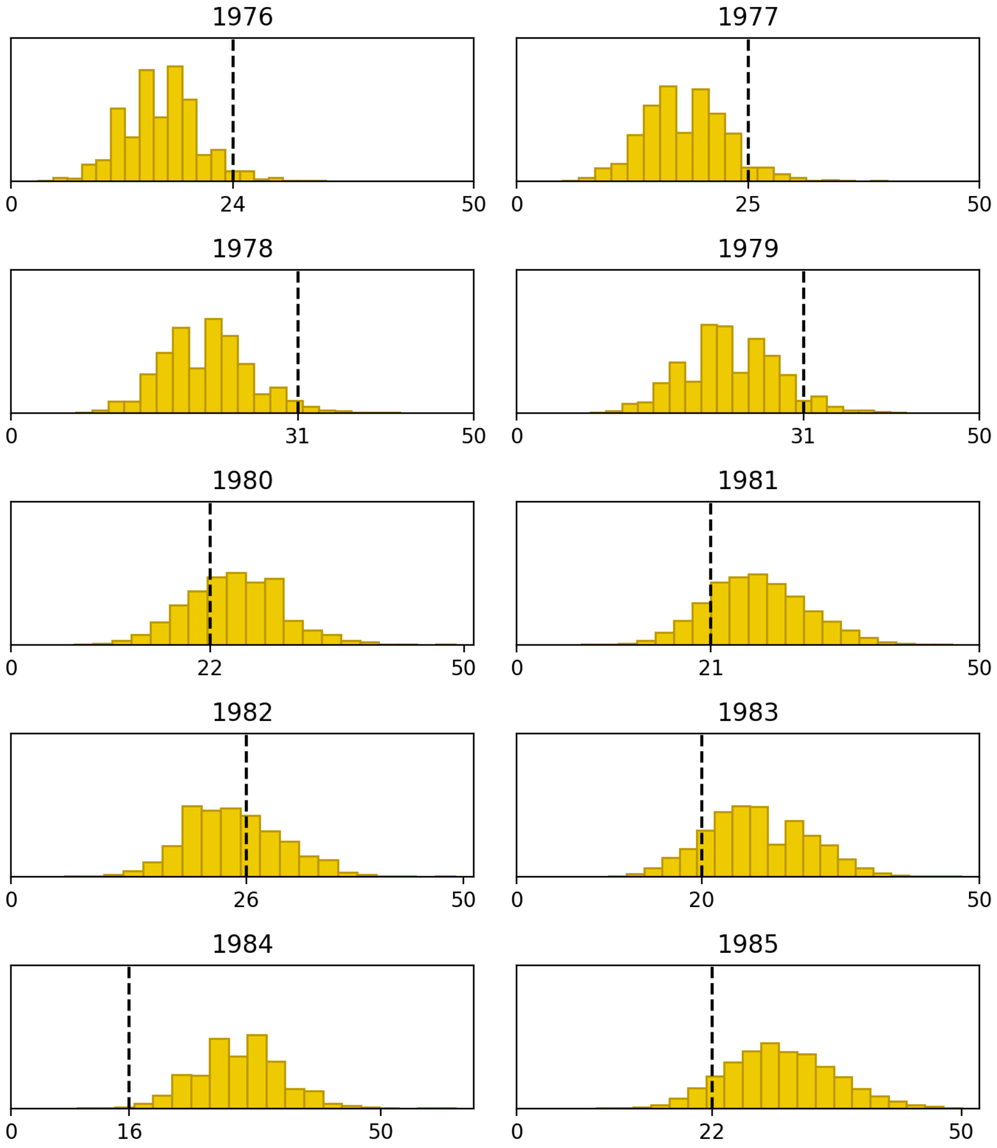


What is the difference between prior and posterior predictions?

Prior predictive distribution



Posterior predictive distribution

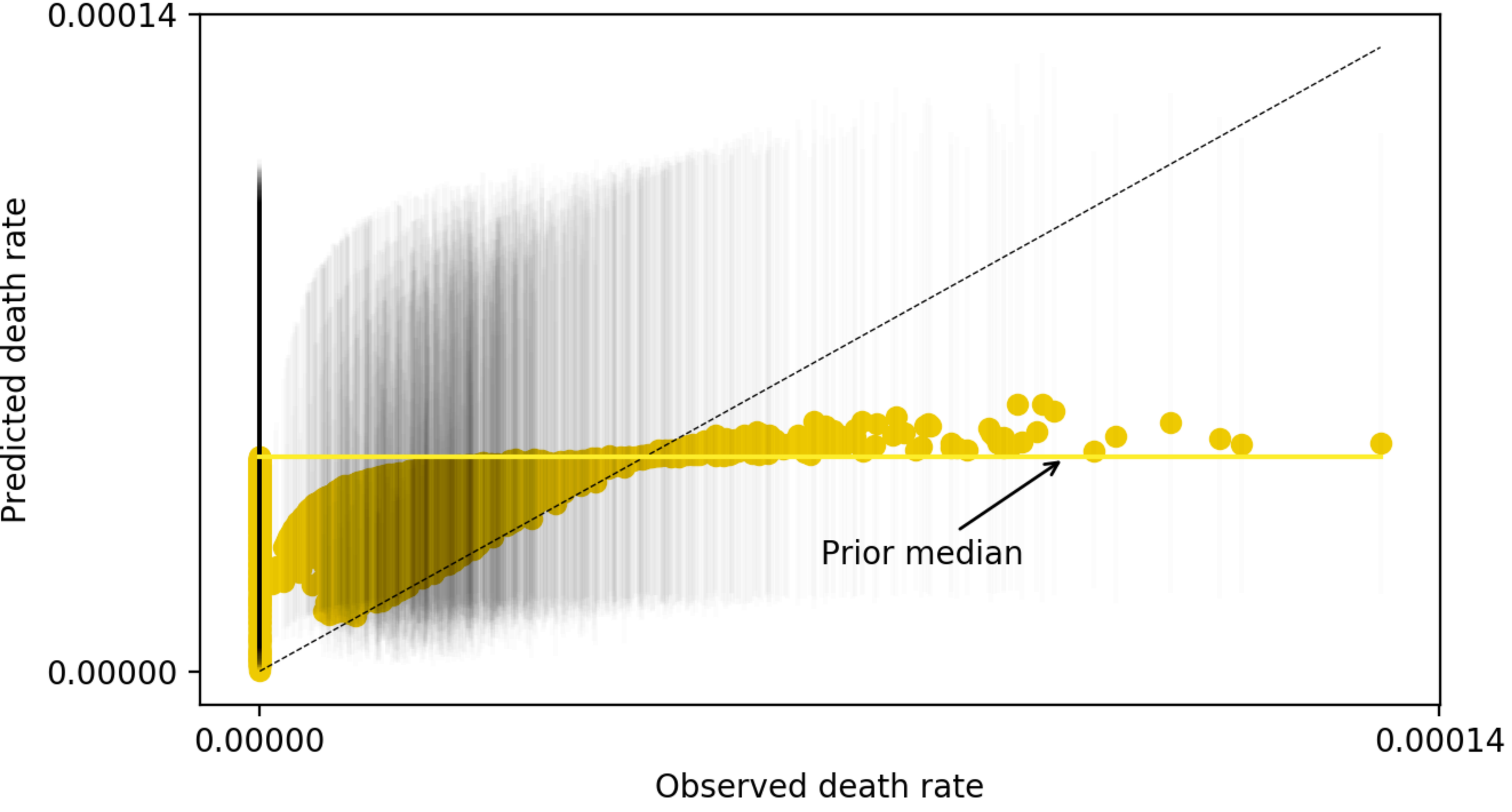


Where we can see the „more flights - more accidents” effect?

First of all from the data

	Fatal accidents	Passenger deaths	Death rate	Miles flown [100 mln miles]
Year				
1976	24	734	0.19	3863.0
1977	25	516	0.12	4300.0
1978	31	754	0.15	5027.0
1979	31	877	0.16	5481.0
1980	22	814	0.14	5814.0
1981	21	362	0.06	6033.0
1982	26	764	0.13	5877.0
1983	20	809	0.13	6223.0
1984	16	223	0.03	7433.0
1985	22	1066	0.15	7107.0

How to interpret the „predicted vs actual” plots?

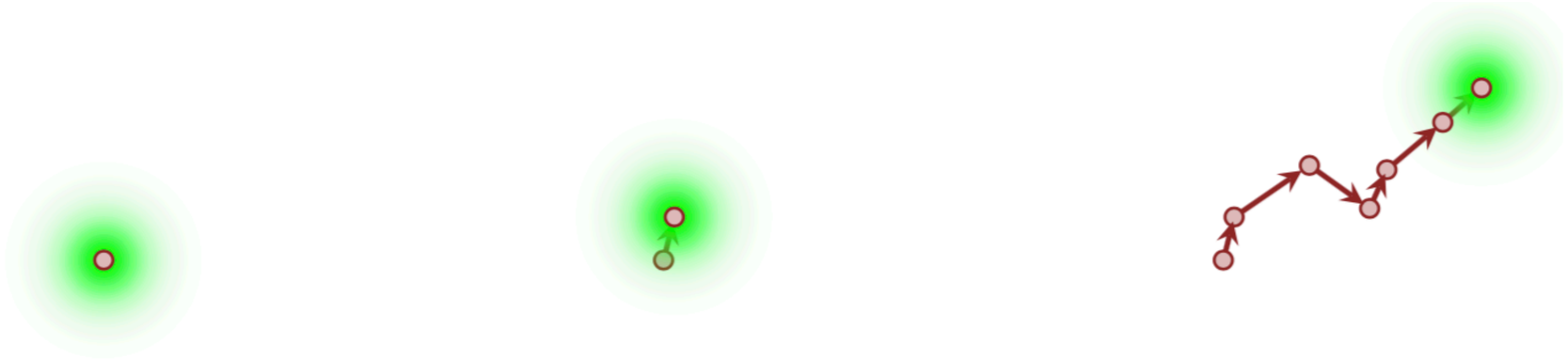

0.00014
0.00000
Predicted death rate
0.00000
0.00014
Observed death rate
Prior median

Can you explain Markov Chain Monte Carlo?

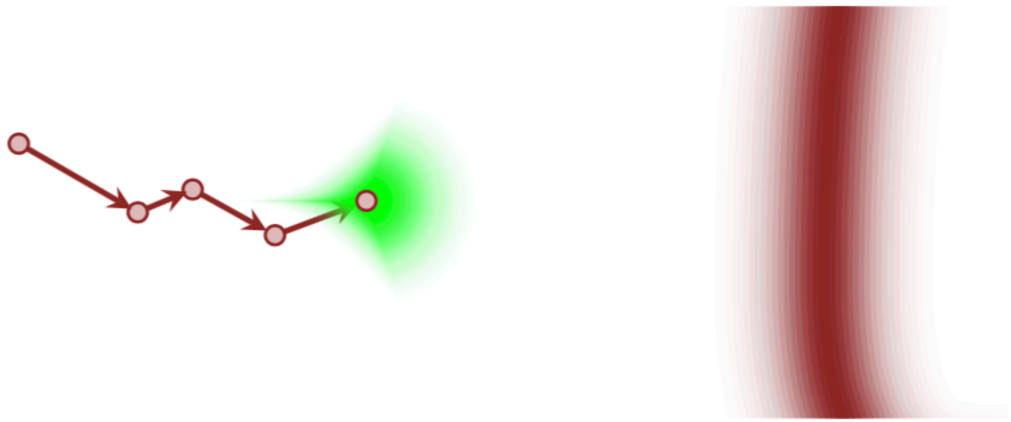
Exploration of typical set

- We need to generate samples from the typical set, which will allow us to get the expectations (Monte Carlo)
- Markov chain is a progression of points through parameter space that transit from one point to another
- Markov chains can preserve distributions (i.e. samples from chain are distributed as the target distribution)

Markov Chain evolution



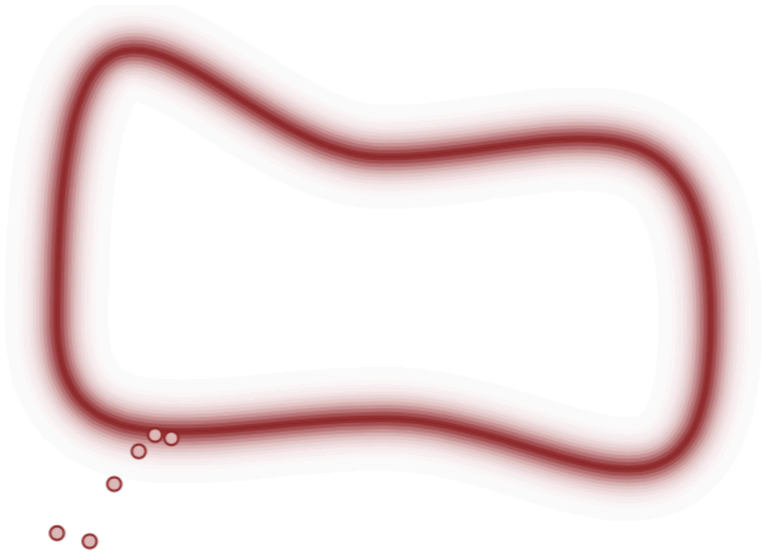
Distribution preserving MC converges towards typical set

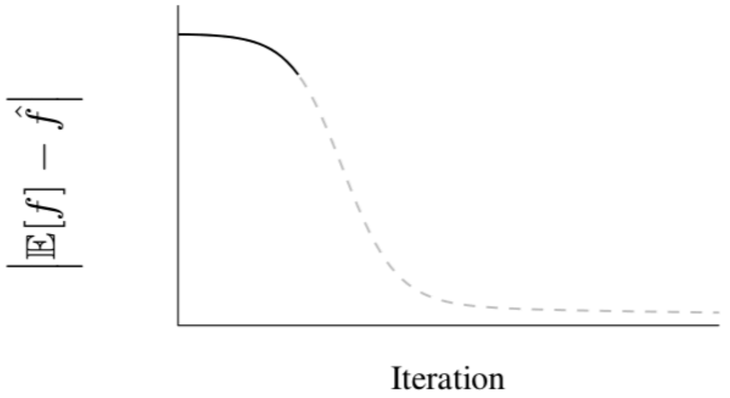


Images (c) Michael Betancourt

Warm up

Reaching the typical set

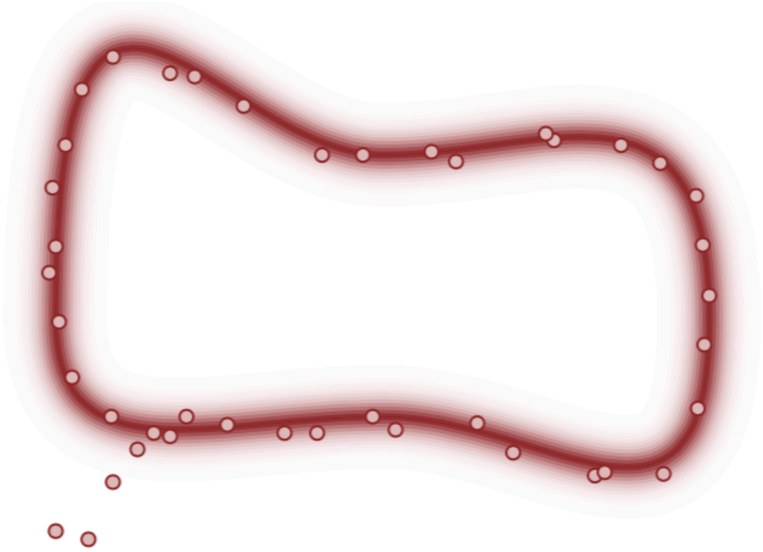


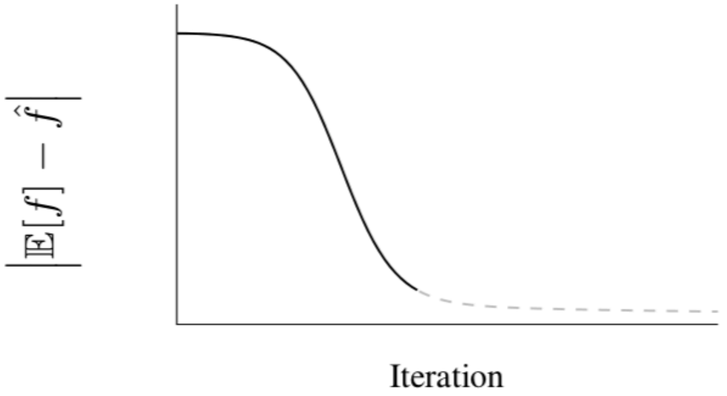


Images (c) Michael Betancourt

Exploration

Covering the most of typical set

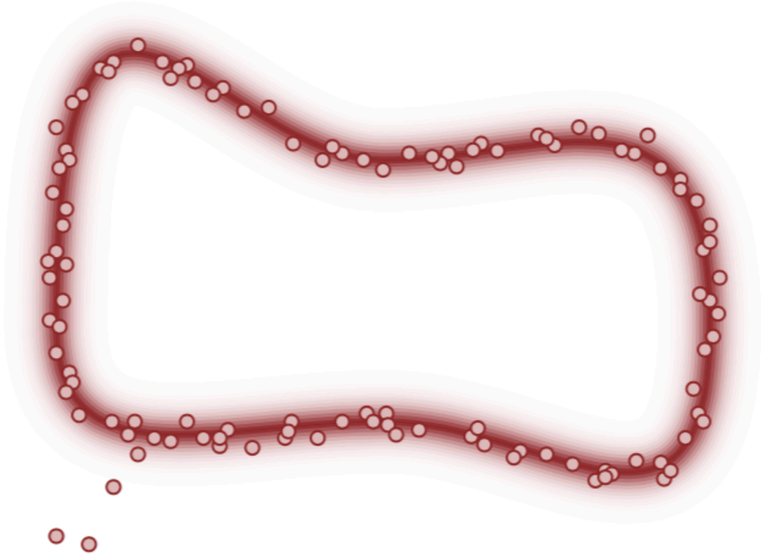


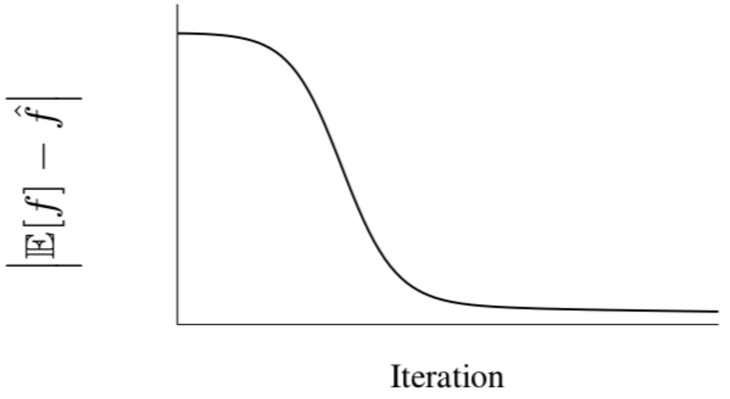


Images (c) Michael Betancourt

Refinement

Improving expectations





Images (c) Michael Betancourt

Why there are no units?

There are units

When they are needed

- Probability is non-dimensional number from 0 to 1, sometimes out of convenience expressed as percentage
- Count data are also dimensionless, as they are integers representing number of events
- In multiparameter lectures we have examples with units

What questions will be on the exam?

Exam questions

And other things you are afraid to ask

- There will be a list of questions available before the exam (at least a week before)
- Everyone will be randomly assigned 3 questions from that list
  - During the exam if you answer the 1st you have passed.
  - If you fail the 1st and 2nd you have failed.
  - If you answered the 2nd and 3rd you have passed.
  - If you have failed the 3rd you have failed.
- Everyone will have 3 approaches.
- Exam will be on MS Teams and probably will be recorded (formal matters to check).

How laboratories will be graded?

What will be with projects?

- You will have 2 more excercise topics that will be graded.
  - First one a poisson model
  - Second one a regression model
- Projects will be severely reduced from the intended plan. There will be an extra video about them.