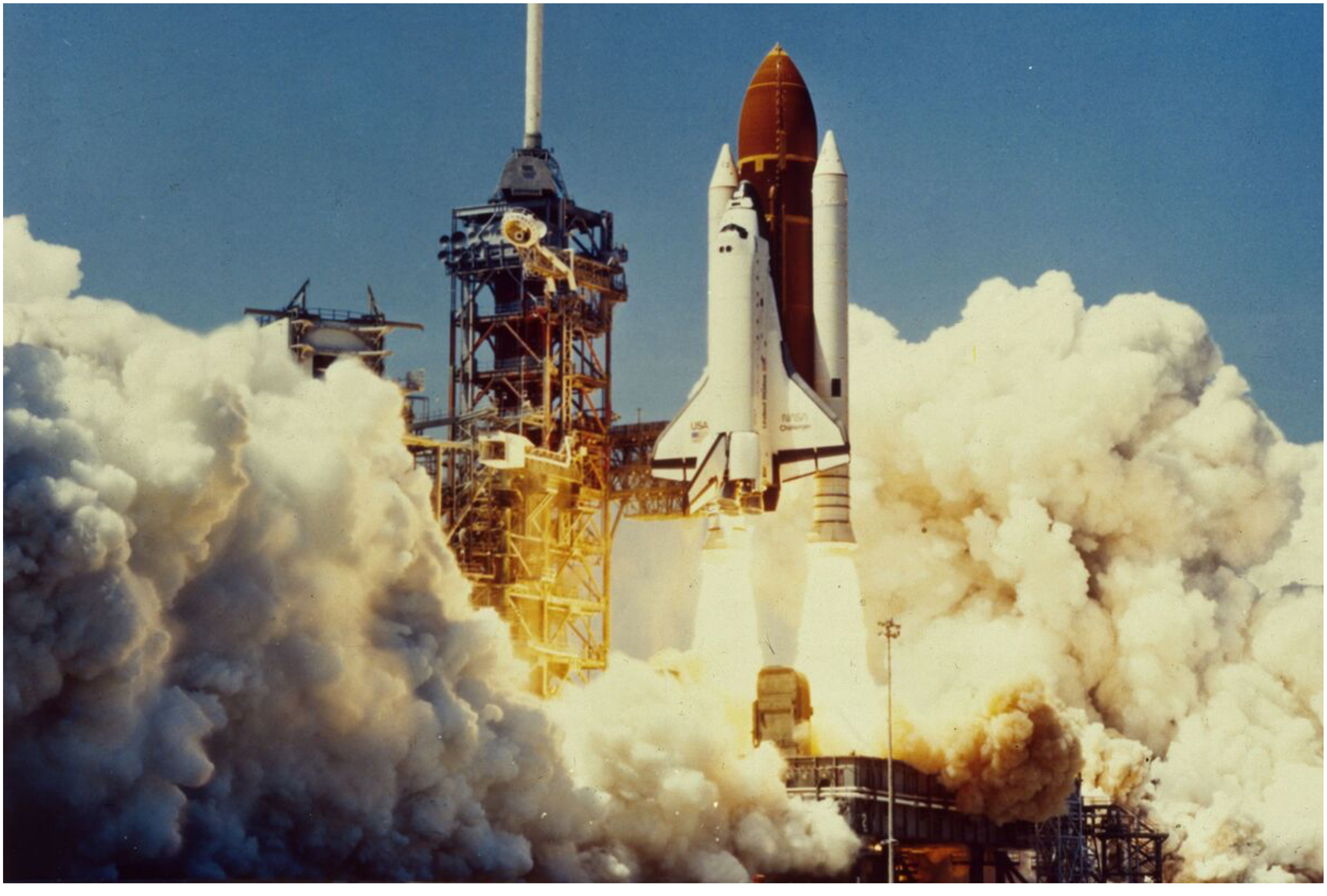
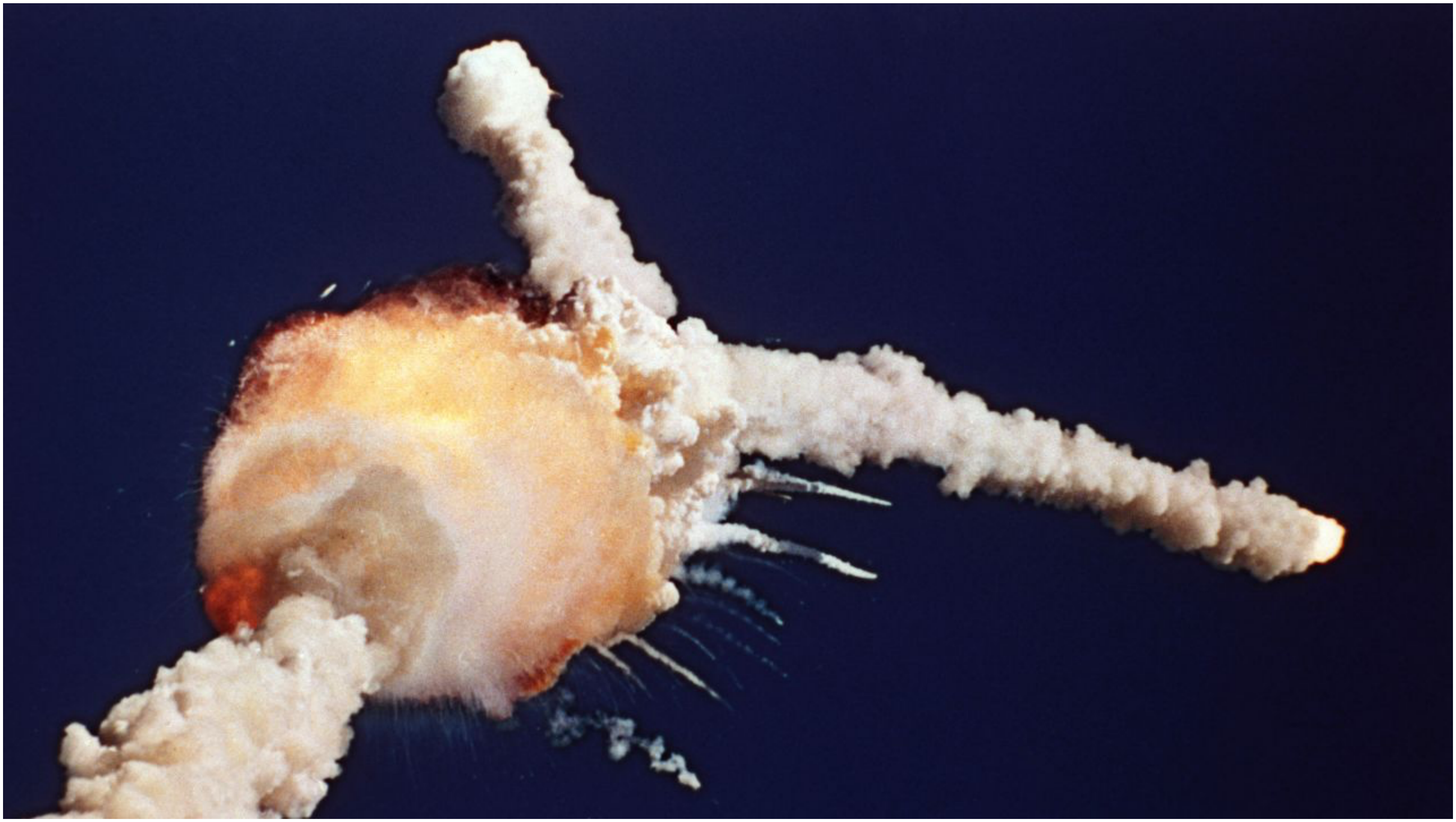
Prof. Alex Rogers
Department of Computer Science
University of Oxford
6th May 2023

Problems

1. Space Shuttle Challenger Disaster



1. Space Shuttle Challenger Disaster



1. Space Shuttle Challenger Disaster

- On January 1986, the Space Shuttle Challenger exploded shortly after launch.

- The subsequent investigation showed that the o-rings in the solid rocket boosters were affected by cold launch temperatures. Previous launches at low temperature had shown evidence of o-ring damage.

- On the day of the launch, the temperature was an unprecedented 36°F.

- Could the risk of o-ring damage have been predicted?

1. Space Shuttle Challenger Disaster

- Plot temperature against occurrence of damage.

- Model the dependence of temperature and occurrence of damage as a logistic function. This function has two parameters - a midpoint t_0 and a rate k.

$$f(t) = \frac{1}{1+e^{-k(t-t_0)}}$$

- Use the value of the logistic function as prior for the probability of success of a Bernoulli random variable. Set appropriate priors for t_0 and k.

- Calculate the posterior distribution of t_0 and k and determine the probability of o-ring damage at 36°F.

2. Ecology Data

2. Ecology Data

- An ecologist is doing a survey to estimate the proportion of a population of bats that come from species A and the proportion that comes from species B. The two bat species are identical in appearance apart from their size and weight.
- Species A is typically smaller, with a mean weight of just 10 grams. The weight of individuals within the species varies, with a standard deviation of 5 grams. Species B is slightly larger and heavier, with a mean weight of 15 grams, and a standard deviation of 2.5 grams.
- Over a night time survey, the ecologist traps and weighs 20 bats.
- What is the proportion of different species in the population?

2. Ecology Data

- Plot a histogram of the measured weight of the bats.

- Propose a model with a Bernoulli random variable describing which species any captured individual comes from.

- Use the model, with appropriate priors, to determine the species of each individual and the proportion of each species within the population.

3. Server Problem

3. Server Problem

- A system administrator is looking at the logs of a server.
- She notices that the number of outgoing messages to a particular IP address seems to have increased at some point during the last 100 days.
- She expects that this must have happened after a software update, and wants to use PyMC to estimate when the software update took place, and what the typical rate was before and after the change.

3. Server Problem

- Plot the number of messages against the day. Note that it is hard to pick the change point by eye.

- You can model this data with a Poisson distribution - a discrete distribution that describes the number of events occurring in a fixed period of time.

- The Poisson distribution has a single parameter that describes the average number of events. The model should include priors for the day on which the change occurred, the average number of events before the change, and the average number of events after the change.