From Risk Analysis to Adversarial Risk Analysis

Part II. Modeling beliefs

David Ríos

david.rios@icmat.es

AXA-ICMAT Chair in ARA and Royal Academy of Sciences

Eliciting probabilities

Once the graphical model is built, we must elicit the probabilities. Sometimes we have access to good databases and may approximate probabilities based on relative frequencies. If not, we may use expert judgements…

https://www.expertsinuncertainty.net/

Eliciting probabilities. The reference experiment

- A ‘ruler’ to measure beliefs called reference experiment
- An experiment is a reference experiment for somebody if this person finds all the experiment results equally likely.

For me, some reference experiments are:

  - A bag with six identical balls numbered 1,2,…,6. This allows me to measure probabilities with values in between 0, 1/6, 2/6, 3/6,4/6,5/6,6/6=1
  - Throw four balanced coins. This allows me to measure probabilities in between 0, 1/16, 2/16,…., 15/16, 16/16=1
  - A fortune wheel with 14 equal sectors, to measure probabilities in between 0, 1/14, 2/14,…., 14/14=1.

Eliciting probabilities. Protocol

- Once identified the calibration experiment, we use it to calibrate the probability of the event of interest. The idea is to gradually compare the event of interest with the reference event until we find one which is as likely. This is not easy to do for a beginner, but we may appeal to several protocols. One is available at

http://www-math.bgsu.edu/~albert/m115/probability/calibration.html

Exercise

- Probability that elections will be repeated in Spain

Eliciting probabilities. Biases

- At tonite's party, I introduce you to Roman, a shy guy. Is he a salesman or a librarian?

Eliciting probabilities. Biases

- Salespersons tend to be extrovert, say 9 out of 10 are extrovert.
- However, there are many timid librarians, say 5 out of 10.
- But, there are much more salespersons than librarians, say 10 salespersons per librarian.
- Those who said that Roman was a librarian are ignoring such fact.

Eliciting probabilities. Biases

- Multitude of experiments have identified biases in our minds when processing info and making decisions. Pioneered by Kahneman and Tversky
- Bias identified -> Remedy (for practical purposes)
- System 1, System 2. Thinking fast and slow (Kahneman)

Judgements under risk

Consider this case:

In a public health study in Madrid. An individual called David Rios is sampled. Which of these results is more likely?

1. Mr. Ríos has had one or more heart attacks.
2. Mr. Ríos has had one or more heart attacks and is older than 55.

Judgements under risk

In experiments, many individuals choose option 2, which is an error called the conjunction fallacy. Event 1 includes Event 2 and should be found more likely!!!!

REMEDY:

Use rigorous probability assessment methods, which require verifying that such inconsistencies do not hold.

Heuristics and biases…

Both experts and non experts may find difficulties in calibrating what they know and what they do not know. A typical experiment consists of proposing almanac type questions like

Give me a range of values for the population of Mostoles, so that you're 90% sure that the actual value is within such range

Heuristics and biases…

Both experts and non experts may find difficulties in calibrating what they know and what they do not know. A typical experiment consists of proposing almanac type questions like

Give me a range of values for the population of Mostoles, so that you're 90% sure that the actual value is within such range

Heuristics and biases

According to INE (2014) the population was 206589.

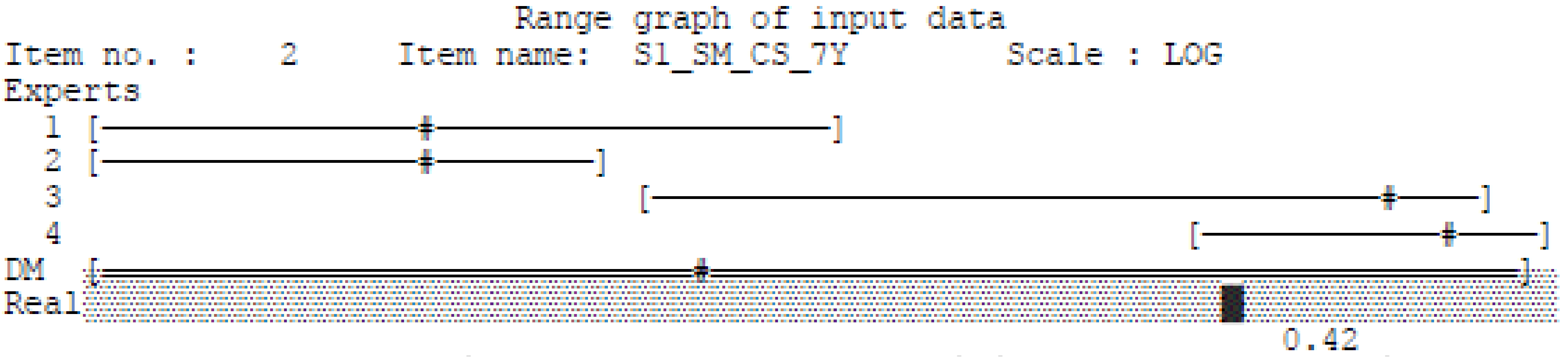
In a series of questions of that type, we may estimate the proportion p of times that the individual hits. If p=0.9 it is well calibrated; if p <0.9, overconfident (too narrow intervals); if p>0.9, underconfident (and not so informative) (too wide intervals)

Calibration studies show that people tend to be overconfident, ie tend to be too sure of what they believe they know.

REMEDY: Rigorous methods for expert judgement assessment.

Heuristics and biases

We want to be calibrated and informative



Cooke's classical method

Updating probabilities

- Many times, we additionally have access to evidence (data) which provides information about the event of interest. Our beliefs are updated through Bayes formula

P(Event|Data)=P(Data|Event) P(Event)/ P(Data)

More on Bayes formula in

http://en.wikipedia.org/wiki/Bayes'_theorem

Forms the basis of Bayesian inference, as we outline now

Exchangeability

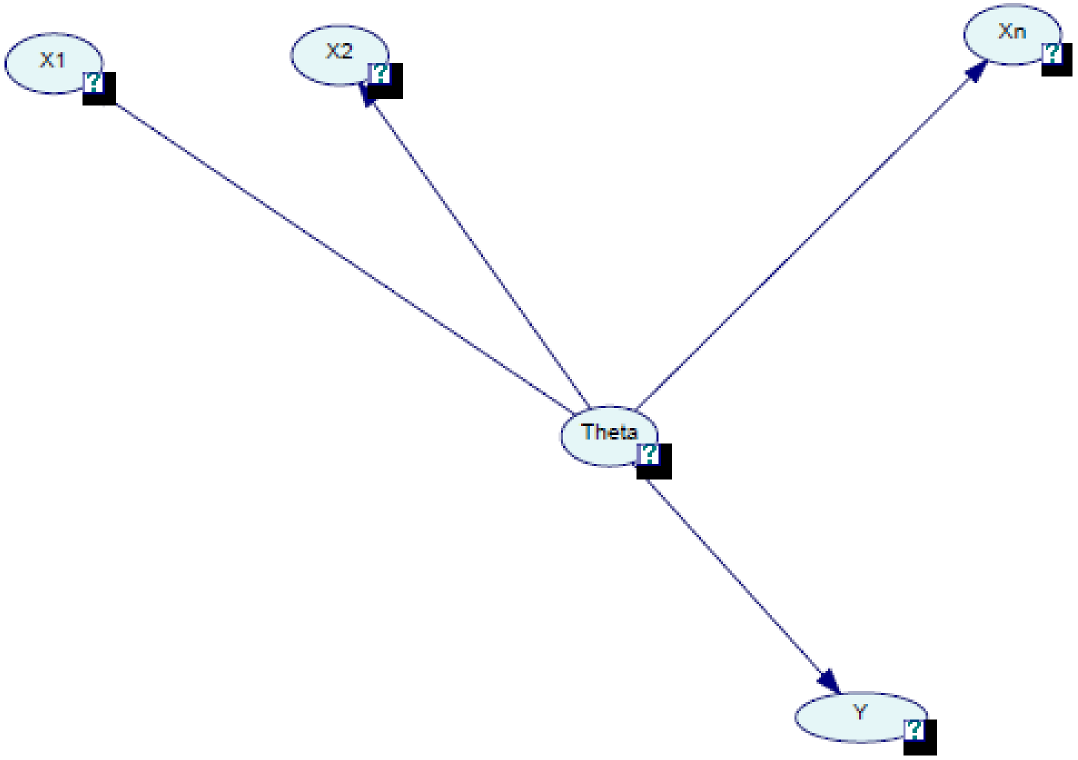
In Stats we deal with observations from random phenomena. Frequently, such observations are assumed to be independent given a certain parameter (AKA conditionally independent) associated with the concept of exchangeability.

http://en.wikipedia.org/wiki/Exchangeable_random_variables

Exchangeability

- A finite set of random variables is exchangeable if any of its permutations have the same joint distribution.
- An infinite set of random variables is exchangeable, if any finite subset is exchangeable.
- A set of rv’s is exchangeable iff are ciid given a certain parameterization (De Finetti’s representation theorem).

Standard parametric model for forecasting



Provide the model

Bayesian inference 1

- We have iid observations which provide info about a parameter of interest
- We also have prior information
- Combine both sources to obaint the posterior
- The posterior summarises all info available and is key for solving the standard inference problems:
  - Point estimation
  - Interval estimation
  - Hypothesis testing

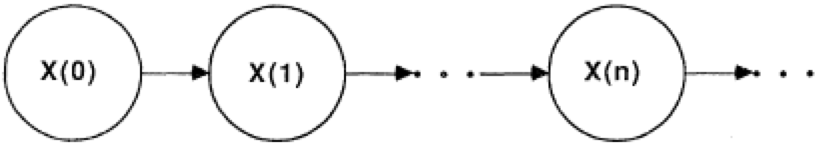
Bayesian Inference 2

- The posterior is also key for solving two key problems in risk analysis (and DM in general):
  - **Forecasting**. Through the predicitive distribution (point forecast, interval forecast, predictive hypothesis testing).
  - **Decision support**. Maximising posterior (or predictive) expected utility.

Bayesian Inference 3

- For the almost ubiquitous operation of integration, use Monte Carlo simulation. (Some info later).
- Sometimes we may sample (almost) directly from relevant distributions (R!)
- Other times (specially if multimdemensional) use Markov chains (MCMC)

Markov chains in discrete time



Transition

$$P_{ij}^{(m,n)} = P\left(X_n = j \mid X_m = i\right)$$

$$P_{ij}^{(m,m+1)} = P\left(X_{m+1} = j \mid X_m = i\right)$$

Homogeneous case

$$P_{ij} = P\left(X_{m+1} = j \mid X_m = i\right)$$

$$P_{ij}^{n} = P\left(X_{n+m} = j \mid X_m = i\right)$$

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^{n} P_{kj}^{m}$$

Stationary distribution, if it exists

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, j \geq 0, \text{ with } \sum_{i=0}^{\infty} \pi_i = 1$$

Inference and forecasting with MCs

- Stochastic process in discrete time with discrete (finite) state space with Markov property
- Assume chain is homogeneous

$p_{ij} = P(X_n = j | X_{n-1} = i)$ $\qquad \pi = \pi \mathrm{P},\ \pi_i \geq 0, \sum p_i = 1$

- Build graphical model for inference and forecasting with MC

Inference and forecasting with MCs

- Assume initial state known. Observe first m transitions

$$X_1 = x_1, \ldots, X_m = x_m$$

- Likelihood

$$l(\mathrm{P}|\mathrm{x}) = \prod_{i=1}^{K} \prod_{j=1}^{K} p_{ij}^{n_{ij}}$$

- MLE

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i.}}$$

Inference and forecasting with MCs

- Prior

$$\mathbf{p}_i \sim \mathrm{Dir}(\boldsymbol{\alpha}_i)$$

- Posterior

$$\mathbf{p}_i|\mathbf{x} \sim \mathrm{Dir}(\boldsymbol{\alpha}'_i) \quad \text{where } \alpha'_{ij} = \alpha_{ij} + n_{ij}$$

Example

- Rain data at a certain station (2, rain; 1, no rain) (Feb 1st- March 20th)

2	2	2	2	2	2	2	2	2	2
1	1	2	1	1	1	1	1	1	1
2	2	1	1	1	1	2	2	2	1
2	1	1	1	1	1	2	1	1	1
1	1	1	1	1	1	1	1	1	

- Model

$$\mathrm{P} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}.$$

- Prior (Jeffreys)

$$p_{ii} \sim \mathrm{Be}(1/2, 1/2).$$

Example

- Conditioning on first day raining, posteriors are

$$p_{11}|\mathrm{x} \sim \mathrm{Be}(25.5, 5.5) \quad p_{22}|\mathrm{x} \sim \mathrm{Be}(12.5, 6.5).$$

- Expected transition matrix

$$E[\mathbf{P}|\mathbf{x}] = \begin{pmatrix} 0.823 & 0.177 \\ 0.342 & 0.658 \end{pmatrix}.$$

Inference and forecasting with MCs

- One step ahead forecast

$$\begin{aligned} P(X_{n+1}=j|\mathrm{x}) &= \int P(X_{n+1}=j|\mathrm{x},\mathrm{P}) f(\mathrm{P}|\mathrm{x})\, d\mathrm{P} \\ &= \int p_{x_n j} f(\mathrm{P}|\mathrm{x})\, d\mathrm{P} \\ &= \frac{\alpha_{x_n j}+n_{x_n j}}{\alpha_{x_n \cdot}+n_{x_n \cdot}} \end{aligned}$$

$$\alpha_{i\cdot} = \sum_{j=1}^{K} \alpha_{ij}.$$

- T steps ahead forecasts

$$P(X_{n+t}=j|\mathrm{x}) = \int \left(\mathrm{P}^t\right)_{x_n j} f(\mathrm{P}|\mathrm{x})\, d\mathrm{P}$$

For $s=1,\ldots,S$:

1. Generate $\mathrm{P}^{(s)}$ from $f(\mathrm{P}|\mathrm{x})$.
2. Generate $x_{n+1}^{(s)},\ldots,x_{n+t}^{(s)}$ from the Markov chain with $\mathrm{P}^{(s)}$ and initial state x_n.

$$P(X_{n+t}=j|\mathrm{x}) \approx \frac{1}{S}\sum_{s=1}^{S} I\left(x_{n+t}^{(s)}=j\right)$$

$$E[X_{n+t}|\mathrm{x}] \approx \frac{1}{S}\sum_{s=1}^{S} x_{n+t}^{(s)}$$

Inference and forecasting with MCs

- Long term forecast

$$\mathrm{P} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}.$$

$$\pi_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}$$

$$E[\pi_1|\mathbf{x}] = \int_0^1 \int_0^1 \frac{1 - p_{22}}{2 - p_{11} - p_{22}} f(p_{11}, p_{22}|\mathbf{x})\, dx$$

- Alternatively, simulation

Example

- No rain March 20th

$$
\begin{aligned}
P(\text{no rain on 21st March}|\mathbf{x}) &= E[p_{11}|\mathbf{x}] = 0.823, \\
P(\text{no rain on 22nd March}|\mathbf{x}) &= E\left[p_{11}^2 + p_{12}p_{21}\middle|\mathbf{x}\right] = 0.742, \\
P(\text{no rain on both}) &= E[p_{11}^2|\mathbf{x}] = 0.681.
\end{aligned}
$$

$$
E[\pi_1|\mathbf{x}] = E\left[\frac{1-p_{22}}{2-p_{11}-p_{22}}\middle|\mathbf{x}\right] = 0.655
$$

Incident Forecasting in Aviation Safety

- (Non-homogeneous) Poisson processes

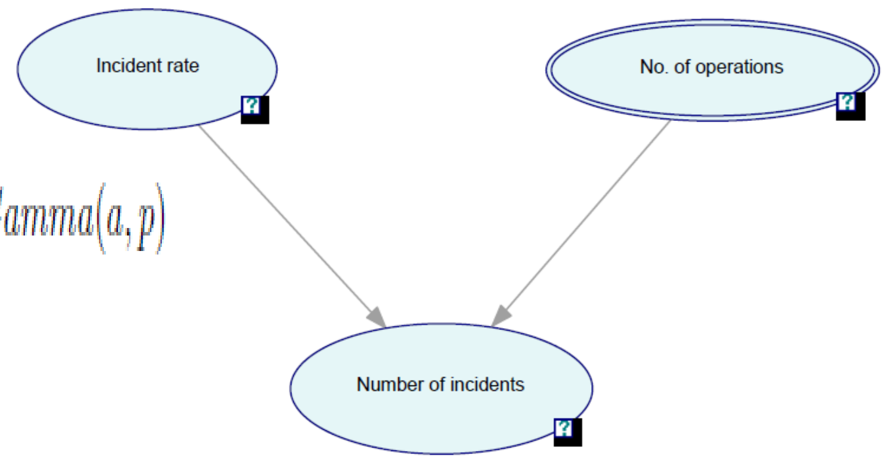
- Dynamic number of operations
- Dynamic rate

- Expert prior elicitation

- Forecasting incidents
  - Annual forecasts for risk assessment and management
  - Monthly, Weekly forecasts for tracking incidents, alarm setting

Basic Model

- **ID**



$$\lambda \sim Gamma(a, p)$$

- **Model**

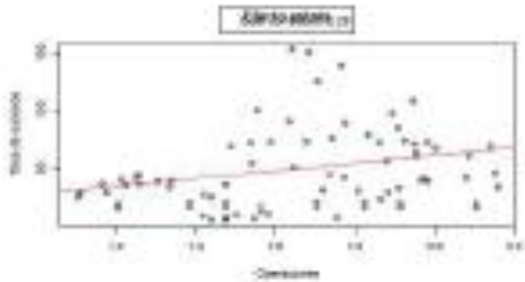
$$X_k | \lambda, n_k \sim Po(\lambda n_k)$$

45

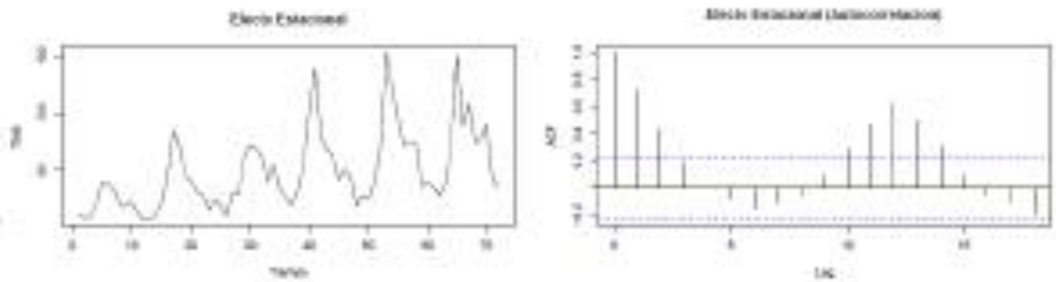
60

75

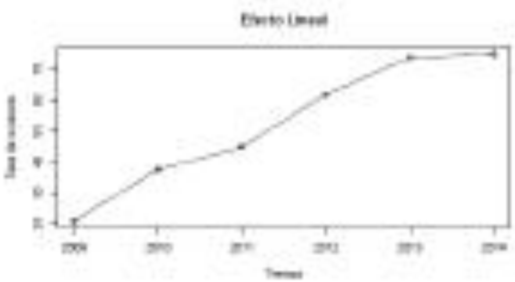
Features of incident rates



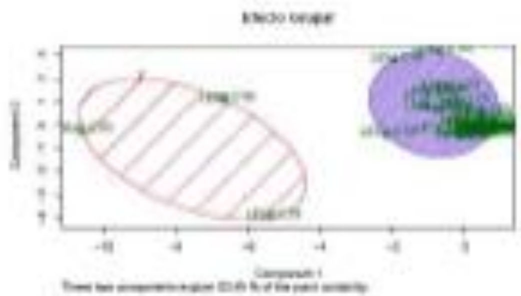
(a)Stress



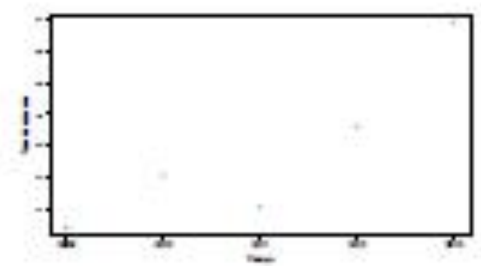
(b)Seasonal



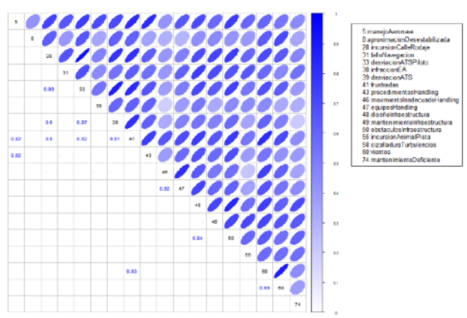
(c)Linear



(d)Group

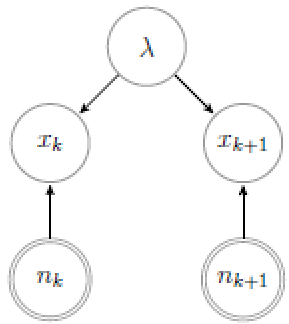


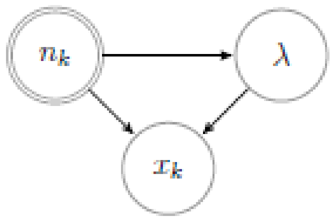
(e)Underreporting



Models

- **ID**





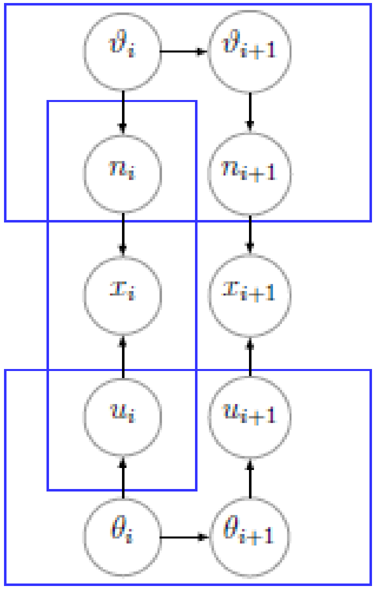
- **Model**

$$\begin{aligned} X_k|\lambda, n_k \sim Po(n_k\lambda) \\ \lambda \sim Ga(a,p). \end{aligned}$$

$$\begin{aligned} X_k|\lambda, n_k &\quad \sim Po(\lambda n_k), \\ \lambda = a n_k + b + \epsilon_k, &\quad \epsilon_k \sim N(0,\sigma^2), \\ p(a,b,\sigma^2) &\quad \propto \tfrac{1}{\sigma^2}. \end{aligned}$$

$$\begin{aligned} x_k^i|\lambda^i, n_k^i &\sim \mathrm{Po}(\lambda^i n_k^i) \\ \lambda^i &\sim \mathrm{Ga}(a^i, p^i) \\ a^i &\sim \mathrm{Ga}(\alpha,\beta) \\ p^i &\sim \mathrm{Ga}(\gamma,\delta) \end{aligned}$$

ID

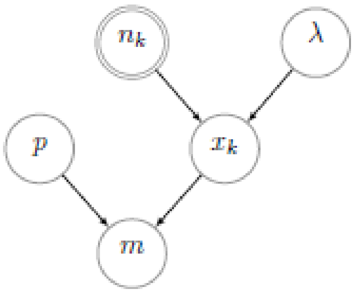


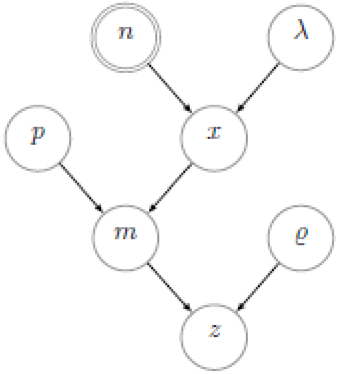
Model

$$\begin{cases} \begin{cases} \begin{cases} n_i = H_i\vartheta_i + z_i,\, z_i \sim N(0, \Sigma_i) \\ \vartheta_i = J_i\vartheta_{i-1} + \xi_i,\, \xi_i \sim N(0, S_i) \end{cases} \\ \vartheta_0 \sim N(\eta_0, S_0) \end{cases} \\ x_i | \lambda_i, n_i \sim Po(\lambda_i n_i) \\ \lambda_i = \exp(u_i) \\ \begin{cases} \begin{cases} u_i = F_i\theta_i + v_i,\, v_i \sim N(0, V_i) \\ \theta_i = G_i\theta_{i-1} + w_i,\, w_i \sim N(0, W_i) \end{cases} \\ \theta_0 \sim N(\mu_0, W_0), \end{cases} \end{cases}$$

Severity Forecasting. Undereporting

- **ID**





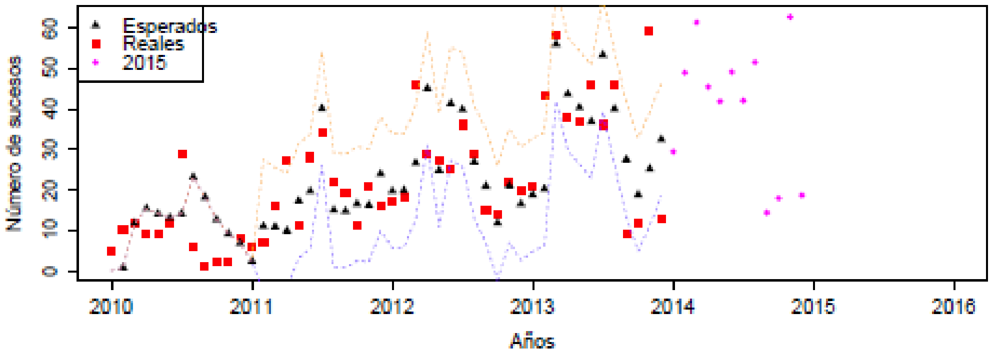
- **Model**

$$\lambda \sim Ga(a, p),$$
$$X_k \sim Po(\lambda n_k),$$
$$p \sim \mathcal{D}ir(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5),$$
$$m|p, X_k \sim \mathcal{M}(X_k; p_1, p_2, p_3, p_4, p_5).$$

$\varrho = (\varrho_1, \varrho_2, \varrho_3, \varrho_4, \varrho_5)$, con $\varrho_i \in [0, 1]$, $\mathcal{B}e(\alpha_i, \beta_i)$.

$z = (z_1, z_2, z_3, z_4, z_5)$ $z_i \sim \mathrm{Bin}(m_i, \varrho_i)$.

Uses: Forecasting and Monitoring



m'	σ'	α'_1	α'_2	α'_3	α'_4	α'_5
19.48	10.68	2	4	27	1055	99