Bayesian inference day 2

Introduction

Today, we start working with real data! A few quick review of yesterday:

Bayesian inference

1. A model for the data generating mechanism is specified in terms of probability distributions with unknown parameters. The model should be specified in a way that the questions you have of the data may be asked of the model parameters.

2. Your prior belief about the unknown parameters is expressed in a probabilistic way via the **prior distribution**, $p(\theta)$.

3. Data are collected and are modeled via a probability distribution, $p(\boldsymbol{y}|\theta)$ (i.e. the likelihood or sampling model)

4. Your updated belief in θ is expressed via the **posterior distribution**, according to Bayes' rule:
$$p(\theta|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\theta)p(\theta)}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\theta)p(\theta)}{\int_\theta p(\boldsymbol{y}|\theta)p(\theta)d\theta}$$

5. All subsequent inference is based on the posterior distribution.

Bayesian data analysis

1. Generate samples from the posterior distribution (using math, PPLs, etc)

2. Assess model convergence using trace plots and the Gelman-Rubin diagnostic.

3. Obtain posterior summaries of the model (means, medians, credibility intervals)

4. Assess model fit using posterior predictive checks

5. (Optionally) Compute metrics for model comparison (Bayes' factors, information criterion, LOO-CV)

Distribution	Parameters	Probability function	Mean	Variance	MGF
Bernoulli Bern(p)	$p \in [0,1]$	$f(x) = p^x(1-p)^{1-x}$; $x \in \{0,1\}$	p	$p(1-p)$	$pe^t + (1-p)$
Binomial Bin(p)	$p \in [0,1]$	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$; $x \in \{0,1,...,n\}$	np	$np(1-p)$	$\left[pe^t + (1-p)\right]^n$
Geometric Geom(p)	$p \in [0,1]$	$f(x) = p(1-p)^{x-1}$; $x \in \{1,2,\dots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$N \in \{0,1,\dots\}$ $r \in \{0,1,\dots,N\}$ $n \in \{0,1,\dots,N\}$	$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$; $x \in \{0,1,\dots,n\}$ if $n \le r$, $x \in \{0,1,\dots,r\}$ if $n > r$	$\frac{nr}{N}$	$n\left(\frac{n}{r}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	Don't bother
Poisson Pois(λ)	$\lambda > 0$	$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x \in \{0,1,\dots\}$	λ	λ	$\exp\left[\lambda(e^t-1)\right]$
Negative binomial NegBin(r, p)	$r \in \{0,1,\dots\}$ $p \in [0,1]$	$f(x) = \binom{x+r+1}{x} p^r (1-p)^x$; $x \in \{0,1,\dots\}$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Beta beta(a, b)	$a, b > 0$	$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$; $x \in (0,1)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	DNE
Chi-square χ^2_ν	$\nu \in \{1,2,\dots\}$	$f(x) = \frac{1}{2^{\nu/2}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}$; $x > 0$	ν	2ν	$(1-2t)^{-v/2}$
Exponential Exp(λ)	$\lambda > 0$	$f(x) = \lambda e^{-\lambda x}$; $x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$
Gamma Ga(α, β)	$\alpha, \beta > 0$	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$; $x > 0$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(1 - \frac{t}{\beta}\right)^{-\alpha}$
Normal $N(\mu, \sigma^2)$	$\mu \in (-\infty, \infty)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$; $x \in (-\infty, \infty)$	μ	σ^2	$\exp\left\{\mu t + \frac{t^2\sigma^2}{2}\right\}$
Uniform Unif(θ_1, θ_2)	$\{\theta_1, \theta_2 : \theta_1 < \theta_2\}$	$f(x) = \frac{1}{\theta_2 - \theta_1}$; $\theta_1 \le x \le \theta_2$	$\frac{1}{2}(\theta_1 + \theta_2)$	$\frac{1}{12}(\theta_2 - \theta_1)^2$	$\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$

Warm-up: Real-data example

Now let's practice on some real data! We will use a subset of the data is Anna is working with.

```
weeds <- readRDS("bcc_dat.rds") %>% dplyr::select(Number_Tre:first3, Year_)
weeds
```

```
# A tibble: 1,791 x 5
   Number_Tre Code  PCT_ZoneTr first3 Year_
        <dbl> <chr> <chr>      <chr>  <dbl>
 1         50 HIIN  100        PIN     2014
 2         12 HIIN  100        ESR     2014
 3         49 MAVU  100        HWY     2014
 4         51 HIIN  100        CHL     2014
 5         58 COMA  90         CAM     2014
 6         10 CAPY  100        NOW     2014
 7          8 CESO  100        PIN     2014
 8         47 MAVU  25         BOT     2014
 9         31 CAPY  100        BGC     2014
10         18 CAPY  5          MAR     2014
# i 1,781 more rows
```

There are a number of potential questions to investigate for these data, including:

1. `Number_Tre ~ PCT_ZoneTr`
2. `Number_Tre ~ PCT_ZoneTr * Code`
3. `Number_Tre ~ PCT_ZoneTr * Code + first3`

For each, write out the model including the priors, and fit the model using either `nimble` or `stan`. Provide posterior summaries and interpret the results.