Probability Distributions and Moment Matching

Bayesian Modeling for Socio-Environmental Data

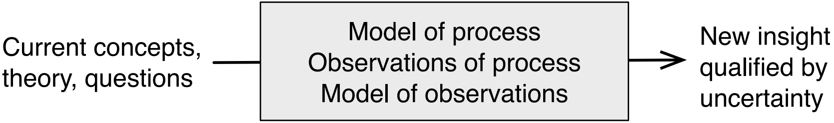
N. Thompson Hobbs

August 2, 2016

Roadmap

- The rules of probability
  - conditional probability
  - independence
  - the law of total probability
- Factoring joint probabilities
- Marginal distributions
- Probability distributions for discrete and continuous random variables
- Moment matching

Motivation: A general approach to scientific research



Motivation: Why do we need to know this stuff?

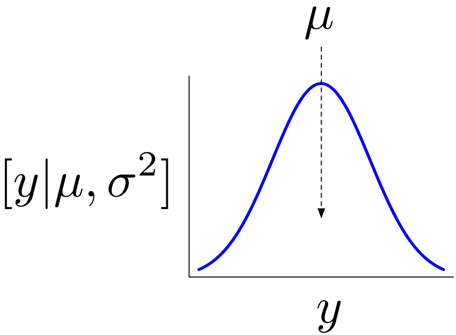
Concept to be taught	Why do you need to understand this concept?
Conditional probability	It is the foundation for Bayes' Theorem and all inferences we will make.
The law of total probability	Basis for the denominator of Bayes' Theorem $[y]$
Factoring joint distributions	This is the procedure we will use to build models.
Independence	Allows us to simplify fully factored joint distributions.
Marginal distributions	Bayesian inference is based on marginal distributions of unobserved quantities.
Statistical distributions	Our toolbox for representing uncertainty and for understanding unobserved quantities based on observed observed ones
Moments	Basis for inference from MCMC
Moment matching	Allows us to embed the predictions of models into any statistical distribution, use scientific literature to inform priors

Motivation: The essence of Bayes

Bayesian analysis is the *only* branch of statistics that treats all unobserved quantities as random variables. We seek to understand the characteristics of the probability distributions governing the behavior of these random variables.

Motivation: linking models to data

$$\mu = g(\theta, x)$$



Motivation: flexibility in analysis

Deterministic models	Probability models	Types of data
general linear	normal	real numbers
non-linear	lognormal	non-negative real numbers
difference equations	gamma	counts
differential equations	beta	0 or 1
occupancy	binomial	0 to 1
state-transition	multinomial	counts in categories
	Poisson	
	negative-binomial	
	Dirichlet	
	Cauchy	

Board work on probability distributions

1. probability density function
   1.1 notation $[z], f(z)$
   1.2 requirements
      1.2.1 $[z] \geq 0$
      1.2.2 $\int_{-\infty}^{\infty} [z]dz = 1$
      1.2.3 $\Pr(a < z < b) = \int_a^b [z]dz$
      1.2.4 Support of random variable z is defined as all values of z for which $[z] > 0$ and defined.
   1.3 cumulative distribution function
   1.4 quantile function
   1.5 moments
      1.5.1 first moment, the expected value (or mean) = $\mathsf{E}(z) = \mu = \int_{-\infty}^{\infty} z[z]dx$, approximated from many (n) random draws from $[z]$ using $\mathsf{E}(z) \simeq \frac{1}{n}\sum_{i=1}^{n} z_i$
      1.5.2 second central moment, the variance = $\mathsf{E}(z-\mu)^2 = \sigma^2 = \int_{-\infty}^{\infty}(z-\mu)^2[z]dx$, approximated from many (n) random draws from $[z]$ using $\mathsf{E}(z-\mu)^2 \simeq \frac{1}{n}\sum_{i=1}^{n}(z_i-\mu)^2$

Board work on discrete distributions

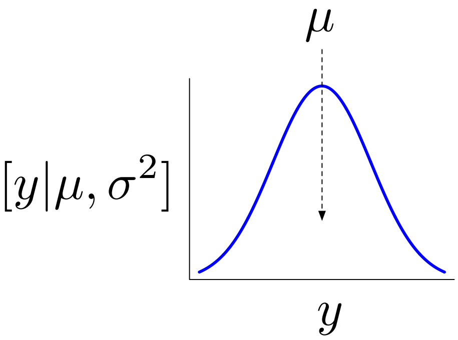
1. probability density function
   1.1 notation $[z], f(z), z \sim$ normal(),
   1.2 requirements
      1.2.1 $[z] \geq 0$
      1.2.2 $\int_{-\infty}^{\infty}[z]dz = 1$
      1.2.3 $\Pr(a < z < b) = \int_a^b [z]dz$
      1.2.4 Support of random variable z is defined as all values of z for which $[z] > 0$ and defined.
   1.3 cumulative distribution function
   1.4 quantile function
   1.5 moments
      1.5.1 first moment, the expected value (or mean) $=$ $\text{E}(z) = \mu = \int_{-\infty}^{\infty} z[z]dx$, approximated from many (n) random draws from $[z]$ using $\text{E}(z) \simeq \frac{1}{n}\sum_{i=1}^{n} z_i$
      1.5.2 second central moment, the variance $=$ $\text{E}(z-\mu)^2 = \sigma^2 = \int_{-\infty}^{\infty}(z-\mu)^2[z]dx$, approximated from many (n) random draws from $[z]$ using $\text{E}(z-\mu)^2 \simeq \frac{1}{n}\sum_{i=1}^{n}(z_i-\mu)^2$

Problems

1. Introduce cheat sheet
2. Do problems 1-8 in Section VII, Models and Data of Probability Lab

Motivation: linking models to data

$$\mu = g(\theta, x)$$



Motivation: flexibility in analysis

Deterministic models	Probability models	Types of data
general linear	normal	real numbers
non-linear	lognormal	non-negative real numbers
matrix models	gamma	counts
differential equations	beta	0 or 1
occupancy	binomial	0 to 1
	multinomial	counts in categories
	Poisson	
	negative-binomial	
	Dirichlet	
	Cauchy	

A familiar approach

$$\begin{aligned} y_i &= \beta_0 + \beta + \varepsilon_i \\ \varepsilon_i &\sim \text{normal}(0, \sigma^2) \end{aligned}$$

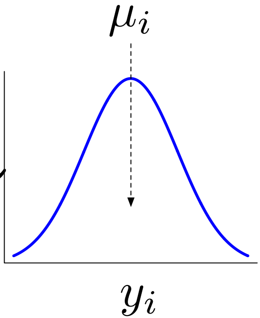
which is identical to

$$\begin{aligned} \mu_i &= \beta_0 + \beta_1 x_i \\ y_i &\sim \text{normal}(\mu_i, \sigma^2) \end{aligned}$$

A familiar approach

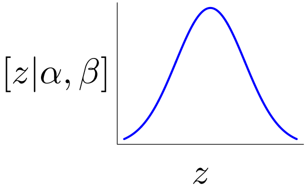
$$\mu_i = \beta_0 + \beta_1 x_i$$

$$\underbrace{[y_i|\mu_i, \sigma^2]}_{\text{normal}}$$



The problem

All distributions have parameters:



α and β are parameters of the distribution of the random variable z.

Types of parameters

Parameter name	Function
intensity, centrality, location	sets position on x axis
shape	controls dispersion and skew
scale, dispersion parameter	shrinks or expands width
rate	scale^{-1}

The problem

The normal and the Poisson are the only distributions for which the parameters of the distribution are the *same* as the moments. For all other distributions, the parameters are *functions* of the moments.

$$
\begin{aligned}
\alpha &= m_1(\mu, \sigma^2) \\
\beta &= m_2(\mu, \sigma^2)
\end{aligned}
$$

We can use these functions to “match”' the moments to the parameters.

Moment matching

$$
\begin{aligned}
\boldsymbol{\mu}_i &= g(\boldsymbol{\theta}, x_i) \\
\boldsymbol{\alpha} &= m_1(\boldsymbol{\mu}_i, \boldsymbol{\sigma}^2) \\
\boldsymbol{\beta} &= m_2(\boldsymbol{\mu}_i, \boldsymbol{\sigma}^2) \\
& [y_i|\boldsymbol{\alpha}, \boldsymbol{\beta}]
\end{aligned}
$$

Moment matching the gamma distribution

The gamma distribution: $[z|\alpha,\beta] = \frac{n^\alpha z^{\alpha-1} e^{-\beta z}}{\Gamma(\alpha)}$
The mean of the gamma distribution is

$$\mu = \frac{\alpha}{\beta}$$

and the variance is

$$\sigma^2 = \frac{\alpha}{\beta^2}.$$

Discover functions for α and β in terms of μ and σ^2.

Note: $\Gamma(\alpha) = \int_0^\infty t^\alpha e^{-t} \frac{\mathrm{d}t}{t}$

Answer

$1)\mu = \frac{\alpha}{\beta}$
$2)\sigma^2 = \frac{\alpha}{\beta^2}$
Solve 1 for β, substitue for β in 2), solve for α :
$3)\ \alpha = \frac{\mu^2}{\sigma^2}$
Substitute rhs 3) for α in 2), solve for β :
$4)\ \beta = \frac{\mu}{\sigma^2}$

Moment matching the beta distribution

The beta distribution gives the probability density of random variables with support on $0, ..., 1$.

$$[z|\alpha,\beta] = \frac{z^{\alpha-1}(1-z)^{\beta-1}}{B(\alpha,\beta)}$$

$$B = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

$$\mu = \frac{\alpha}{\alpha+\beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\alpha = \frac{\mu^2 - \mu^3 - \mu\sigma^2}{\sigma^2}$$

$$\beta = \frac{\mu - 2\mu^2 + \mu^3 - \sigma^2 + \mu\sigma^2}{\sigma^2}$$

You need some functions...

```
#BetaMomentMatch.R
# Function for parameters from moments
shape_from_stats <- function(mu, sigma){
 a <-(mu^2-mu^3-mu*sigma^2)/sigma^2
 b <- (mu-2*mu^2+mu^3-sigma^2+mu*sigma^2)/sigma^2
shape_ps <- c(a,b)
return(shape_ps)
}
# Functions for moments from parameters
beta.mean=function(a,b)a/(a+b)
beta.var = function(a,b)a*b/((a+b)^2*(a+b+1))
```

Problems continued

Do problems 9 - 13 in Section VII, Models and Data