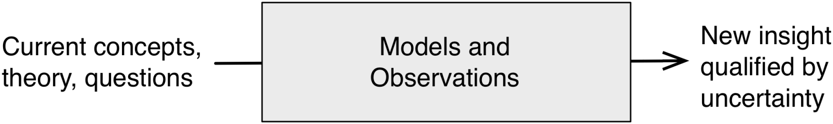
Probability Concepts and Distributions

Models for Socio-Environmental Data

Chris Che-Castaldo, Mary B. Collins, N. Thompson Hobbs

August 1, 2019

Motivation: A general approach to scientific research



Roadmap

- The rules of probability
  - conditional probability and independence
  - the law of total probability
  - the chain law of probability
- Directed acyclic graphs (Bayesian networks)
- Probability distributions for discrete and continuous random variables
- Marginal distributions
- Moment matching

What you must know and why

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.
Probability distributions	Our toolbox for fitting models to data and representing uncertainty
Moments	The way we summarize distributions
Marginal distributions	Bayesian inference is based on marginal distributions of unobserved quantities.
Moment matching	Allows us to embed the predictions of models into any statistical distribution

Motivation: The essence of Bayes

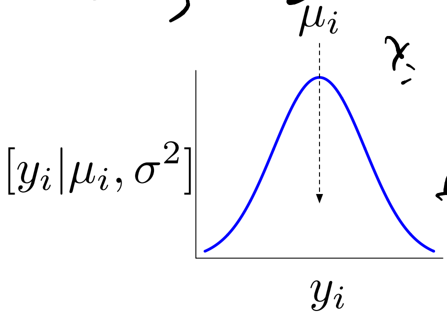
Bayesian analysis is the *only* branch of statistics that treats all unobserved quantities as random variables. We seek to understand the characteristics of the probability distributions governing the behavior of these random variables.

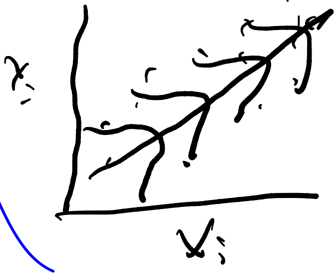
Motivation: models of data

$\underline{\Theta} = (\beta_0, \beta_1)'$

$$\mu_i = g(\boldsymbol{\theta}, x_i) = \beta_0 + \beta_1 x_i$$

$y_i \sim \text{normal}(\mu_i, \sigma^2)$





A model of the data describes our ideas about how the data arise.

Motivation: flexibility in analysis

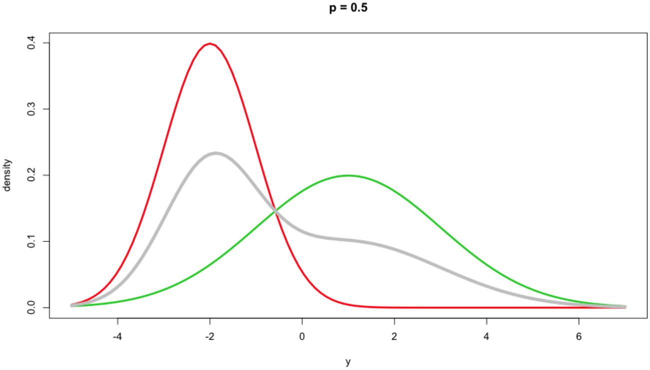
Deterministic models	Types of data
general linear nonlinear differential equations difference equations auto-regressive occupancy state-transition integral-projection	real numbers non-negative real numbers counts 0 to 1 0 or 1 counts in categories proportions in categories ordinal categories

univariate and
multivariate

Motivation: flexibility in analysis

Probability model	Support for random variable
normal	real numbers
multivariate normal	real numbers (vectors)
lognormal	non-negative real numbers
gamma	non-negative real numbers
beta	0 to 1 real numbers
Bernoulli	0 or 1
binomial	counts in 2 categories
Poisson	counts
multinomial	counts in $>$ 2 categories
negative binomial	counts
Dirichlet	proportions in $\geq$ 2 categories
Cauchy	real numbers

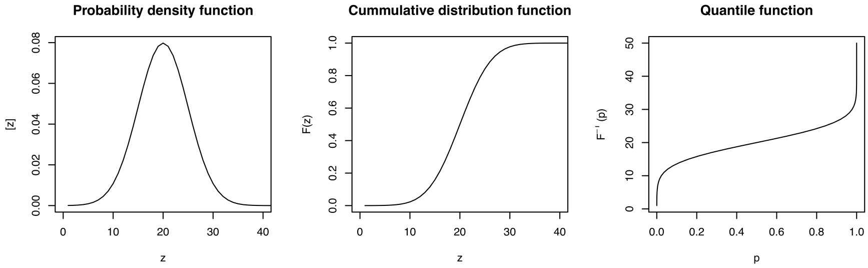
Motivation: flexibility in analysis



Work flow: probability distributions

- General properties and definitions (today)
  - discrete random variables
  - continuous random variables
- Specific distributions (cheat sheet and Probability Lab 2)
- Marginal distributions (Probability Lab 3)
- Moment matching (Probability Lab 4)

How will we use probability distributions?



Used to fit models to data, to represent uncertainty in processes and parameters, and to portray prior information

Used to make inference

Key points for today

1. What makes a function a probability mass function or a probability density function?
2. How to compute moments of distributions?
3. How to approximate moments of distributions from random draws?
4. Relationships among:
    4.1 probability mass function
    4.2 probability density function
    4.3 cumulative distribution function
    4.4 quantile function

5. Concept of support

Board work on general concepts of probability distributions with previous slide on screens.

What is a continuous random variable?

$$a < Z < b$$

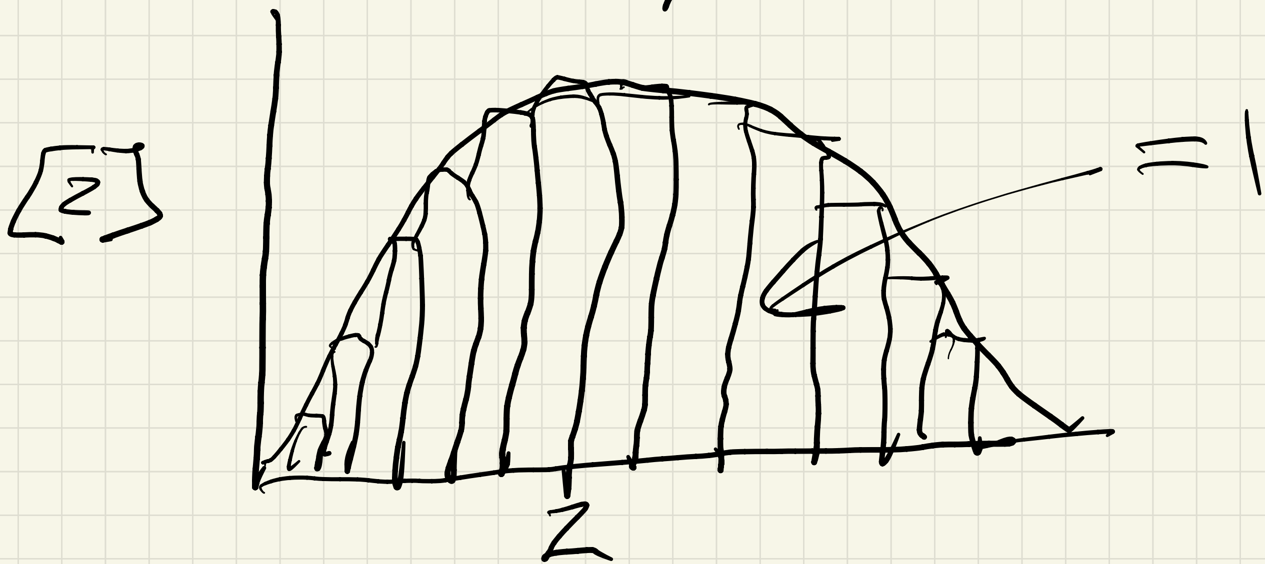
Notation

$[z]$ = probability density of Z

$$f(z) = pr(Z = z)$$

[Z] = probability density of the continuous random variable Z

What is probability density?



scaling of [Z] axis?

Requirements for probability density function (pdf)

① $[z] \geq 0$

② $\displaystyle\int_{-\infty}^{\infty} [z]\,dz = 1$

③ $\displaystyle\int_{a}^{b} [z]\,dz = Pr(a < Z < b)$

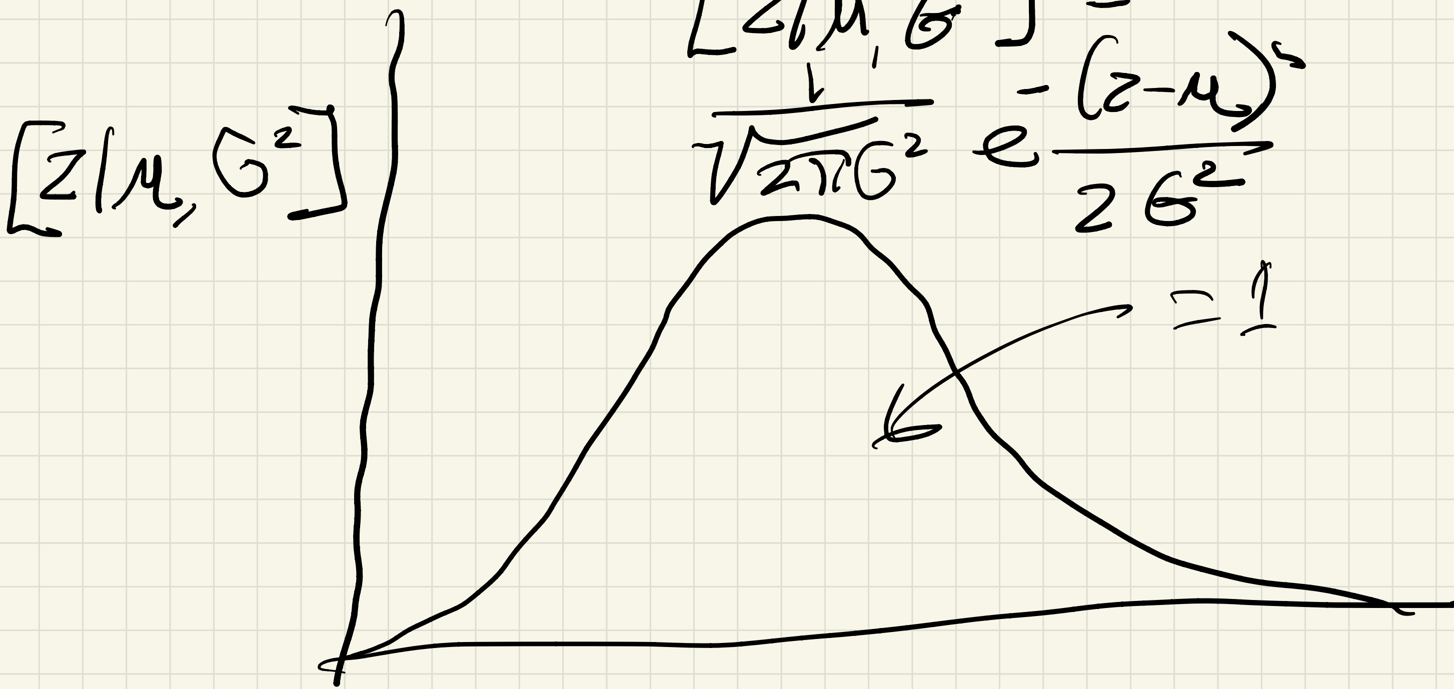
$\leq Z \leq$

Definition of support

$[z] > 0$ and defined

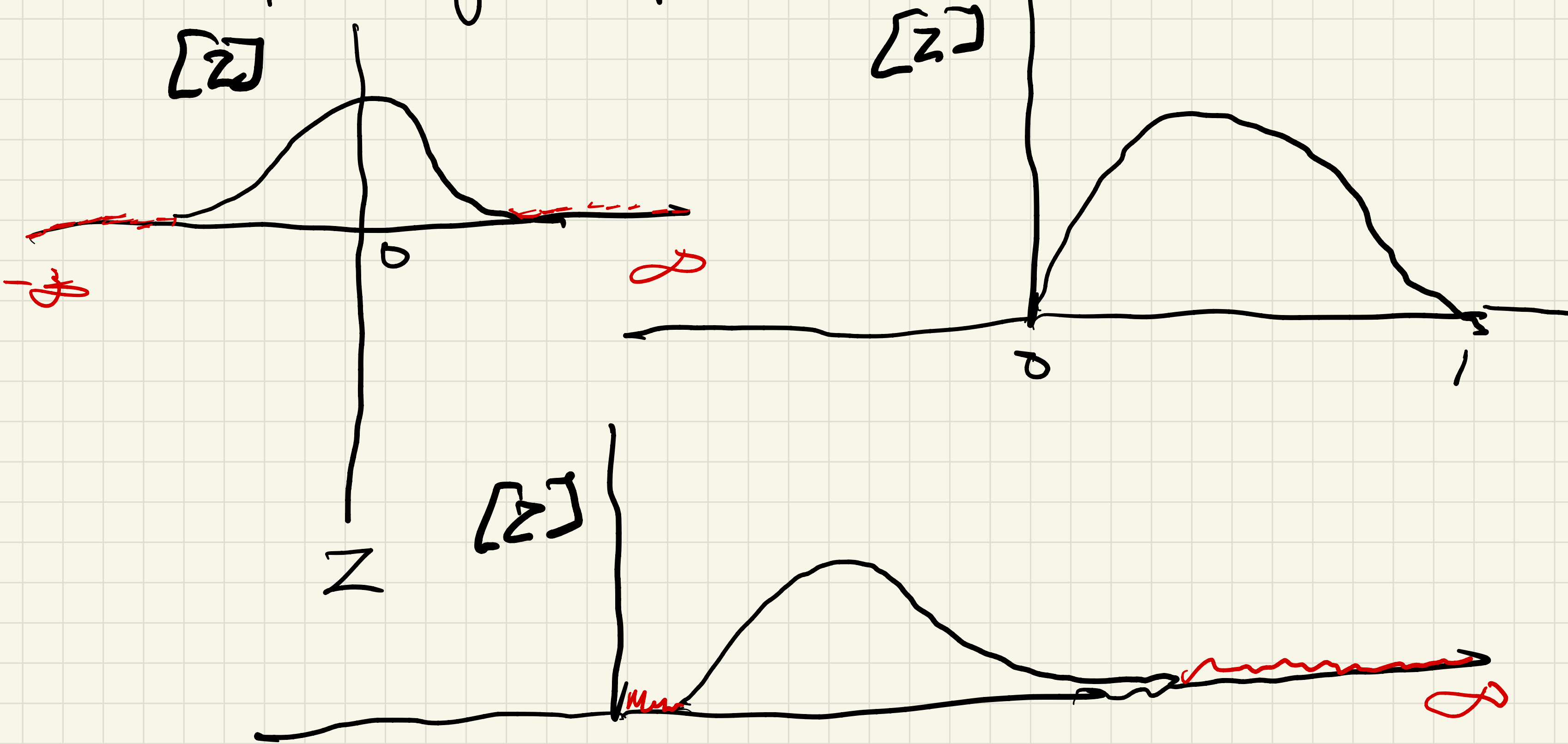
Specific example of PDF

$$[z \mid \mu, \sigma^2] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$



$= 1$

Examples of support



Moments and moment generating function revised

Mean, first moment

$$E(z) = \mu = \int_{-\infty}^{\infty} z\, [z]\, dz$$

Approx

$$\frac{1}{n} \sum_{i=1}^{n} z_i$$

Variance, second central moment

$$E\left((z-\mu)^2\right) = \sigma^2 = \int_{-\infty}^{\infty} (z-\mu)^2 [z]\, dz \qquad \frac{1}{n} \sum_{i=1}^{n} (z_i - \mu)^2$$

Moments and moment generating function revised

Skewness, third standardized moment

$$\gamma = \frac{E((z-\mu)^3)}{\sigma^3} = E\left(\left(\frac{z-\mu}{\sigma}\right)^3\right)$$

approximated by skewness function in R package moments

Kurtosis, fourth standardized moment

$$\kappa = \frac{E((z-\mu)^4)}{\sigma^4} = E\left(\left(\frac{z-\mu}{\sigma}\right)^4\right)$$

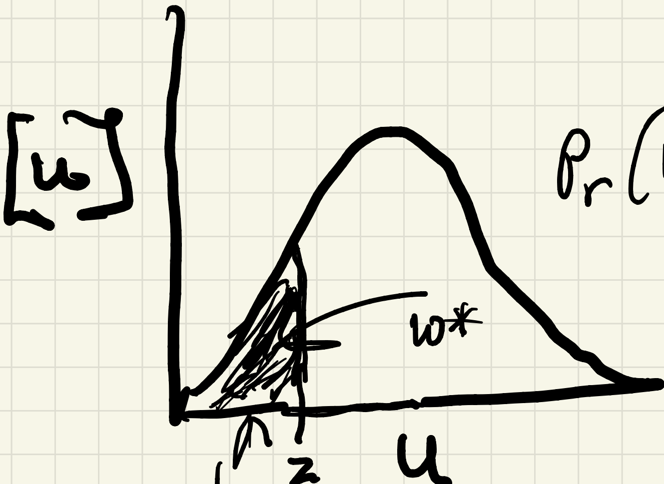
approximated by Kurtosis function in R package moments

Moments and moment generating function revised

Moment generating function turn out to be a deeper topic than we need for this course.

I pasted a url for a nice treatment of the topic in slack general if you want to pursue.

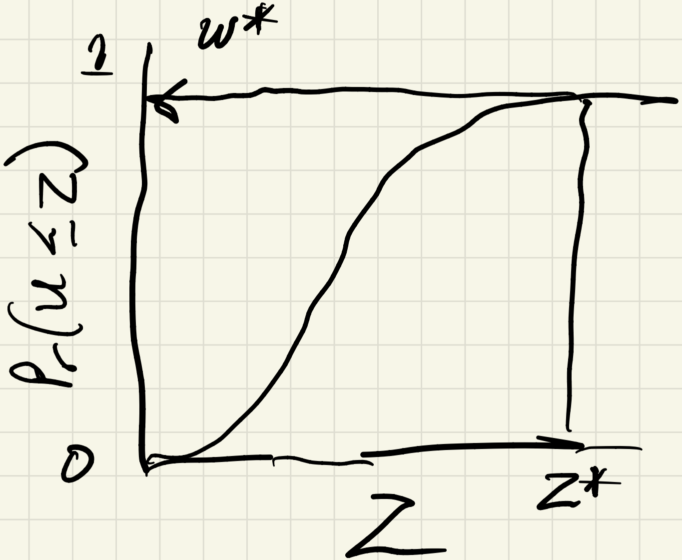
Probability density function (PDF)



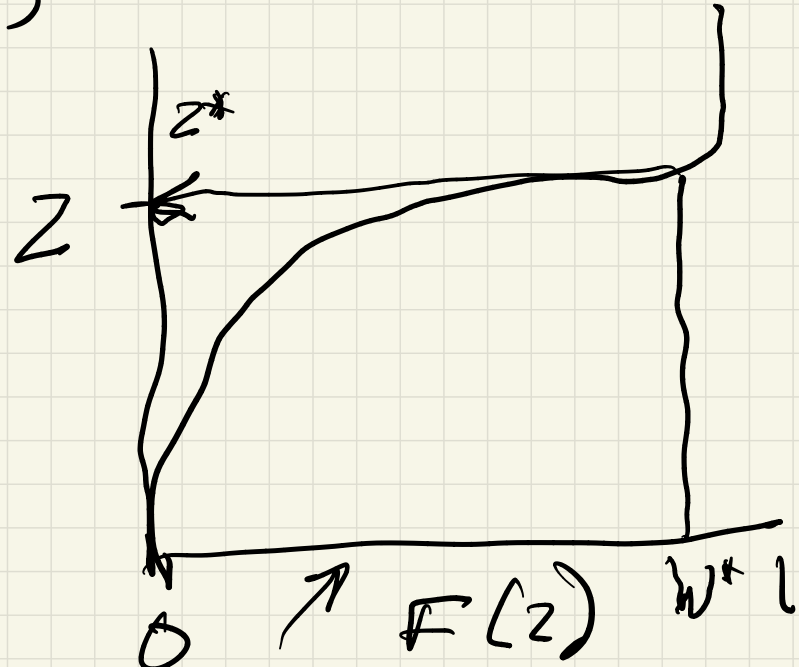
Cummulative distribution function (CDF)

$$F(z) = \int_{-\infty}^{z} [u]\, du$$

Cummalative distribution function (CDF) $F(z)$



Quantile function



$F^{-1}(z)$ = inverse of CDF

What is a discrete radom variable?

$$a < Z < b$$

Probability mass function (PMF)

$[z]$ = probability of z

Alternative notation for PMF

$$f(z) = pr(Z = z) = [z]$$

Definition of support for Z:

$[Z] > 0$ and defined

Set notation for Z:

$S = \{1, 2, 3, \ldots\ldots \infty\}$

Assignments

Group	Main board number	Task
bayes4life	1	Sketch PMF: $[Z]$ as function of z
cream	2	Write the two requirements for a function to be a PMF

Assignments

For discrete random variable:

Group	Main Board Number	Task
jamright	3	write equation for 1st moment and its approximation
Uninformative priors	4	write equation for 2nd central moment and its approximation
group 1	5	Sketch relationship between pdf and CDF

Key points for today

1. What makes a function a probability mass function or a probability density function?
2. How to compute moments of distributions?
3. How to approximate moments of distributions from random draws.
4. Relationships among:
   4.1 probability mass function
   4.2 probability density function
   4.3 cumulative distribution function
   4.4 quantile function

Housekeeping

Errors in text:

```
http://www.stat.colostate.edu/~hooten/papers/pdf/
Hobbs_Hooten_Bayesian_Models_2015_errata.pdf
```

Note in particular that the plot for a cumulative distribution function is wrong.
Break here for marginal distributions and moment matching