$\underbrace{\text{Chapter 1}}$
Intro

$\underbrace{\text{Chapters 2 - 4}}$
"Philosophy"

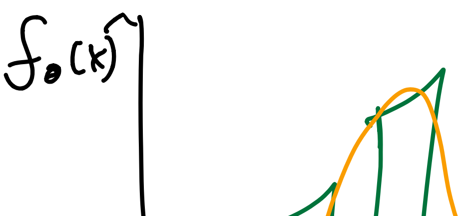
$\underbrace{\text{Chap 5}}$
Classical statistics
Foundations.

$\underbrace{\text{Ch 7, 8, 10}}$
Algorithms

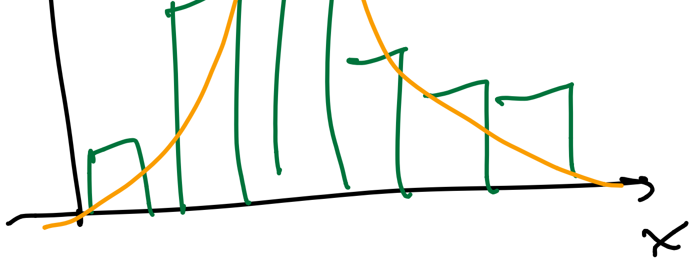
Chap 5. Parametric models and Exponential Families

Freq
Bayesian
Fisherian
→ 1 aspect in common:
- Low dimensionality problems.
(usually Θ is very low dimensional).

- The focus is to estimate parameters. → we call this parametric estimation



① Proposing a model
② Given that model, what is the



best value of its parameters given the data?

- Univariate Families ($x \in \mathbb{R}^1$)

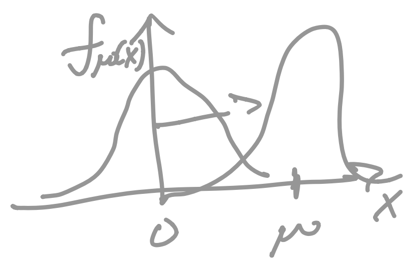
5 most familiar univariate families:

Normal: You can write any Normal as:

$$N(\mu, \sigma^2) \sim \mu + \sigma N(0,1)$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

(Any Normal is a shifted by μ and scaled by σ version of $N(0,1)$).



Gamma:

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\beta x}$$

$$\boxed{\frac{1}{\sigma^\nu \Gamma(\nu)} x^{\nu-1} e^{-x/\sigma}}$$

* χ^2 is special case of Gamma

→ $\chi^2_n \sim 2 \cdot Ga(n/2, 1)$

* Exponential (λ) is special case of Gamma

$Exp(\lambda) \sim Ga(1, \lambda)$

$$\frac{1}{\lambda^1 \Gamma(1)} \cdot x^0 \cdot e^{-x/\lambda} = \frac{1}{\lambda} \cdot e^{-x/\lambda},$$

the exponential pdf.

* Erlang-k assumes ν is an integer.

Poisson

$Poi(\mu)$

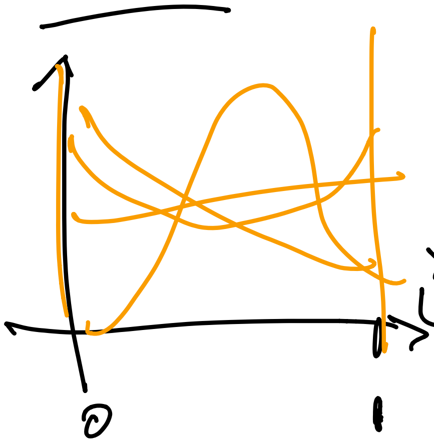
- Usually employed with count data.

$$f_\lambda(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Beta

- Usually employed to model a probability.

- Very flexible in the $(0,1)$ support.



$$X \sim Be(\alpha, \beta) \iff f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}$$

$$E(x) = \frac{\alpha}{\alpha+\beta}$$

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

Binomial : Prob. distribution of the # of Bernoulli successes in a sequence of trials.

Recall : If $X \sim Ber(p)$ then $X = \begin{cases} 1 & \text{with prob } p \\ 0 & \text{o.w.} \end{cases}$

$$\Rightarrow f_p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$\rightarrow$ There is a wealth of relationships between these distributional families,

For example:

(1) $Be(\alpha, \beta) \sim \dfrac{Ga(\alpha, \sigma)}{Ga(\alpha, \sigma) + Ga(\beta, \sigma)}$

(2) $Bin(n, p) \overset{\text{aprox}}{\sim} Poisson(n \cdot p)$

large n
+
Small p

Multivariate Normal Distribution

$X \in \mathbb{R}^p$, p-dimensional real #'s.

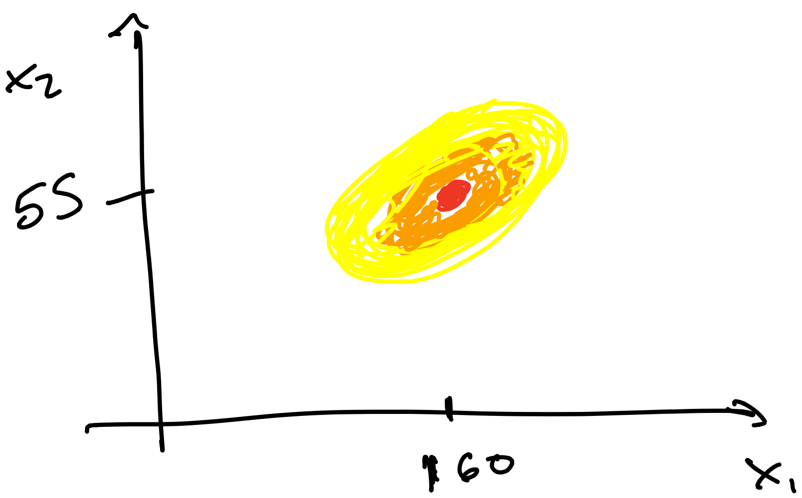
Example: $(\underset{x_1}{\text{weight}}, \underset{x_2}{\text{Height}})$

women in MX

$X \in \mathbb{R}^2$

$f: X \to Y$

$Y \in \mathbb{R}^1$



A random vector $X = (X_1, X_2, \ldots, X_p)^T$

has mean vector

$$\mu = E(x) = [E(x_1), E(x_2), \dots, E(x_p)]^T$$

and pxp covariance matrix:

$$\Sigma = E[\overbrace{(x-\mu)(x-\mu)^T}^{\text{outer product}}] = (E[(x_i-\mu_i)(x_j-\mu_j)])$$

in 1-d this is $(x-\mu)^2$

$\longrightarrow$ p columns

p rows $\downarrow$

V(x₁)

$$= \begin{bmatrix} E[(x_1-\mu_1)(x_1-\mu_1)] & E[(x_1-\mu_1)(x_2-\mu_2)] & \cdots \\ E[(x_2-\mu_2)(x_1-\mu_1)] & \ddots & \\ \vdots & & \end{bmatrix}$$

Main Diagonal

We'll use notation

$$x \sim (\mu, \Sigma)$$

"x has mean vector μ and covariance matrix Σ"!

Notice that in 1-d this reduces to

$$x \sim (\mu, \sigma^2).$$

Denote the entries of Σ by σ_{ij} with diagonal elements being variances

$$\sigma_{ii} = Var(x_i) \text{ for } i = 1, \dots, p$$

Moreover, we can use the off-diagonal elements to compute

$$r(x_i, x_j) = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\,\sigma_{jj}}}$$

(Correlation between x_i and x_j)

Multivariate Normal extends the definition of a univariate Normal.

Standard Multivariate Normal pdf

Let $z = (z_1, z_2, \dots, z_p)^T$ be a vector of p Independent $N(0,1)$ variates, with pdf

$$f(z) = (2\pi)^{-p/2} \cdot e^{-\frac{1}{2}\sum_{i=1}^{p} z_i^2} = (2\pi)^{-p/2} \cdot e^{-\frac{1}{2} z^T z}$$

→ A general Multivariate Normal family is obtained by letting

$x = \mu + \sigma z$

$$\underset{p\times 1}{x} = \underbrace{\underset{p\times 1}{\mu}}_{\text{shift}} + \underbrace{\underset{p\times p}{T} \cdot \underset{p\times 1}{z}}_{p\times 1}$$

univariate $\sigma^2 = \sigma \cdot \sigma$

multivariate

$$\Sigma = T\,T^T$$

Then, following usual transformations of prob. distributions...

$$f_{\mu,\Sigma}(x) = \frac{(2\pi)^{-p/2}}{\underbrace{|\Sigma|^{1/2}}_{\text{Det}(\Sigma)}} \cdot e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

multivariate Normal pdf

we denote a MVN random variable by:

$$x \sim N_p(\mu, \Sigma)$$

- we'll illustrate the bivariate Normal distr. with $\mu = (0,0)^T$ and $\Sigma = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$

Further we'll draw iso density lines for this example.

The lines show combinations of (x_1, x_2) that have the same density.