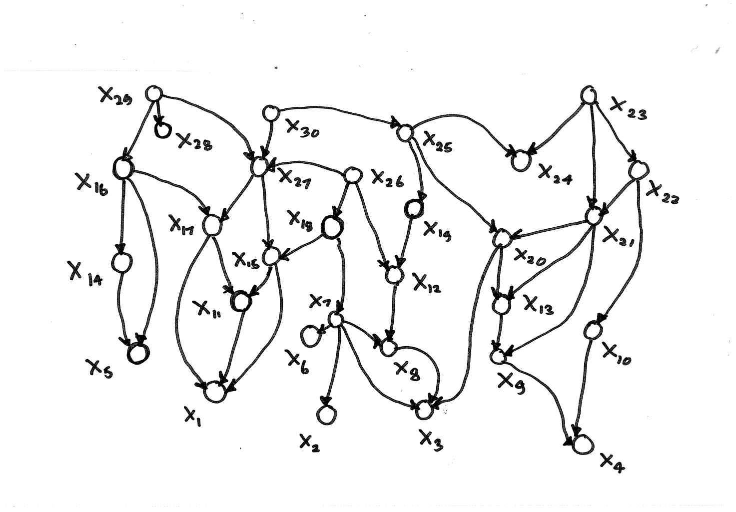
37457

# Advanced Bayesian Methods

Professor Matt Wand

University of Technology Sydney




## Distribution Theory and Uni Stats Subjects

| | | |
|---|---|---|
| 1st subject | single random variable $X$ | density function $f_X(x)$ |
| 2nd subject | two random variables $X, Y$ | density function $f_{X,Y}(x,y)$ |
| this subject | many random variables<br>maybe even hundreds of r.v.s | density function<br>$f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$<br>$k \approx 300$, say |

With dozens, maybe even hundreds, of random variables useful to have a way to keep their distributional structure organised.

$\Longrightarrow$ PROBABILISTIC GRAPHS

## Streamlined Density Function Notation

First subjects in Stats:

$X$ is a random variable.

$f_X(x)$ is probability mass function ($X$ discrete) or density function ($X$ continuous) at $x$.

$f_{X,Y}(x,y)$ is joint density function of $X, Y$.

$X = x$ corresponds to $X$ having observed value $x$.

## Streamlined Density Function Notation

First subjects in Stats:

$X$ is a random variable.

$f_X(x)$ is probability mass function ($X$ discrete) or density function ($X$ continuous) at $x$.

$f_{X,Y}(x, y)$ is joint density function of $X, Y$.

$X = x$ corresponds to $X$ having observed value $x$.

---

Remainder of this subject (in keeping with advanced stats literature):

$x$ is a random variable.

$p(x)$ is density function of $x$ at $x$ (admitted abuse!).

$p(x, y)$ is joint density function of $x, y$.

$x = \overset{\circ}{x}$ corresponds to $x$ having observed value $\overset{\circ}{x}$.

Hand out alternative Assignment 1.

## Important to Always Keep in Mind

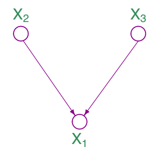
We are just doing Distribution Theory

just like you did in earlier subjects

and Assignment 1!!

## Do You Remember...

**Jessica the DAG**




Not to be confused with...

## Jessica the PROBABILISTIC DAG

$$p(x_1, x_2, x_3) = \prod_{i=1}^{3} p(x_i|\text{parents of } x_i) = p(x_1|x_2, x_3)\, p(x_2)\, p(x_3)$$

## SECOND SPECIFIC EXAMPLE (ANOTHER Jessica the PROBABILISTIC DAG)

$$x_1|x_2, x_3 \sim N(x_2, x_3), \quad x_2 \sim \text{Beta}(7, 9), \quad x_3 \sim \text{Gamma}(7, 19)$$

$$p(x_1|x_2, x_3) = (2\pi x_3)^{-1/2} \exp\{-(x_1 - x_2)/(2x_3)\}$$
$$p(x_2) = 45045\, x_2^6 (1 - x_2)^8, \quad 0 < x_2 < 1.$$
$$p(x_3) = \frac{19^7}{\Gamma(7)} x_3^6 \exp(-19x_3), \quad x_3 > 0.$$

$$\implies p(x_1, x_2, x_3) = (2\pi x_3)^{-1/2} \exp\{-(x_1 - x_2)/(2x_3)\} \times 45045\, x_2^6 (1 - x_2)^8 \times \frac{19^7}{\Gamma(7)} x_3^6 \exp(-19x_3), \quad 0 < x_2 < 1,\ x_3 > 0$$

## Jessica the PROBABILISTIC DAG

$$p(x_1, x_2, x_3) = \prod_{i=1}^{3} p(x_i|\text{parents of } x_i) = p(x_1|x_2, x_3)\, p(x_2)\, p(x_3)$$

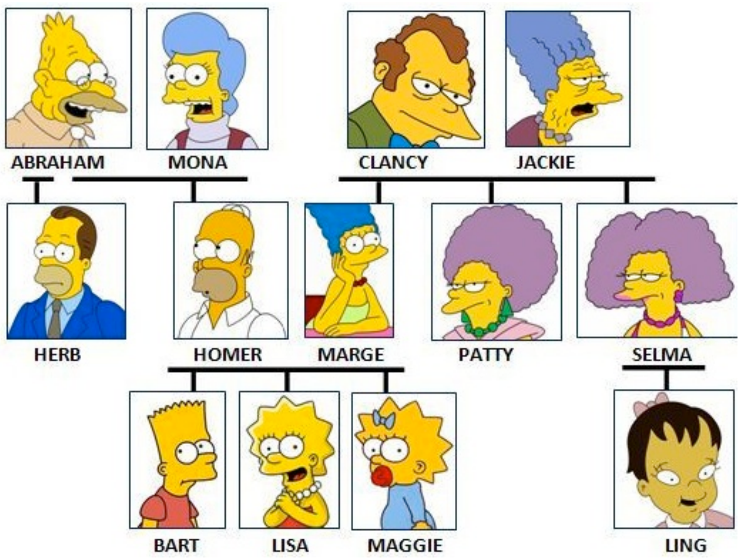
SPECIFIC EXAMPLE:

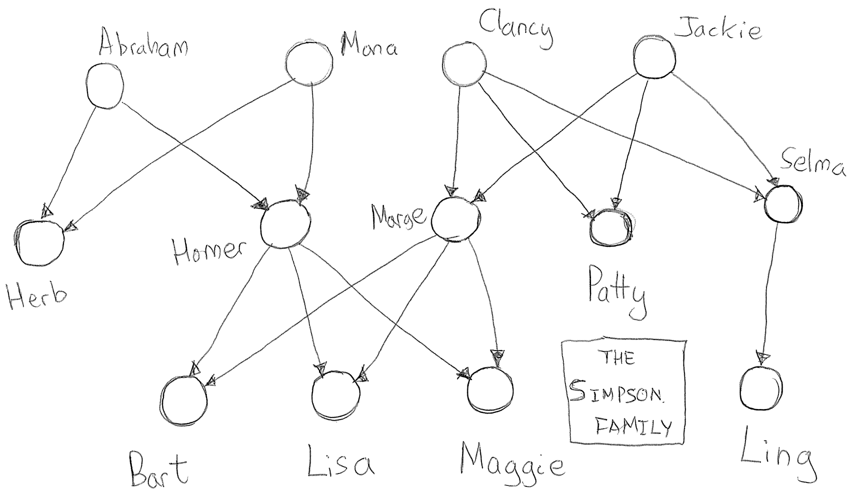
$$p(x_1|x_2, x_3) = \frac{2x_1 + 14x_2 + 5x_3}{252 + 18x_1 + 45x_3}, \quad x_1, x_2, x_3 = 1, 2, 3$$

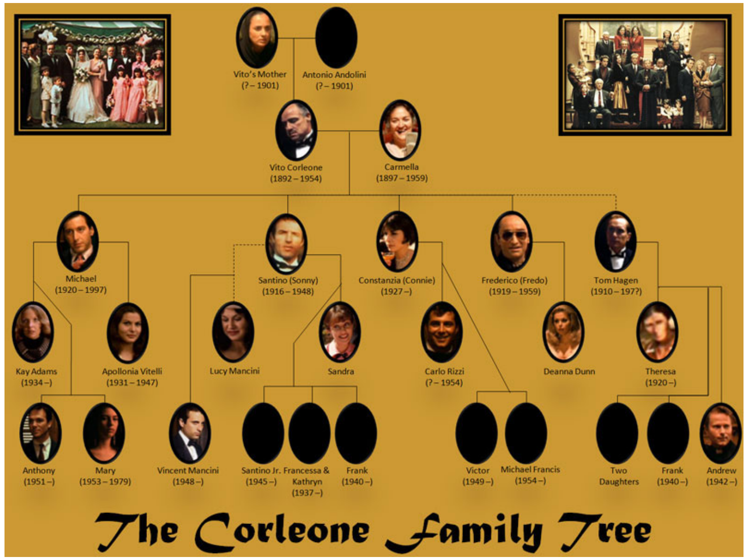
$$p(x_2) = \frac{3 + 7x_2^2}{107}, x_2 = 1, 2, 3$$

$$p(x_3) = \frac{12}{13(1 + x_3)}, x_3 = 1, 2, 3$$

$$\implies p(x_1, x_2, x_3) = \frac{2x_1 + 14x_2 + 5x_3}{252 + 18x_1 + 45x_3} \times \frac{3 + 7x_2^2}{107} \times \frac{12}{13(1 + x_3)}, \quad x_1, x_2, x_3 = 1, 2, 3$$

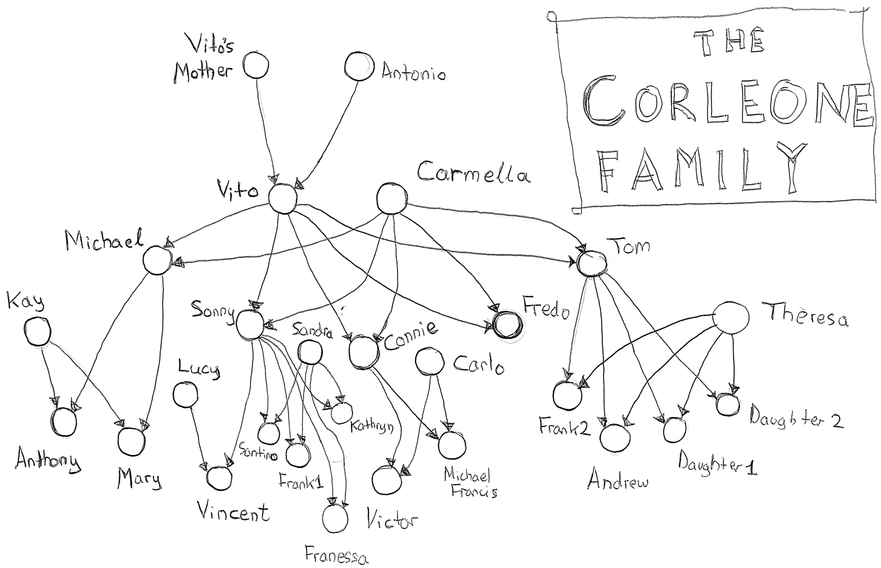




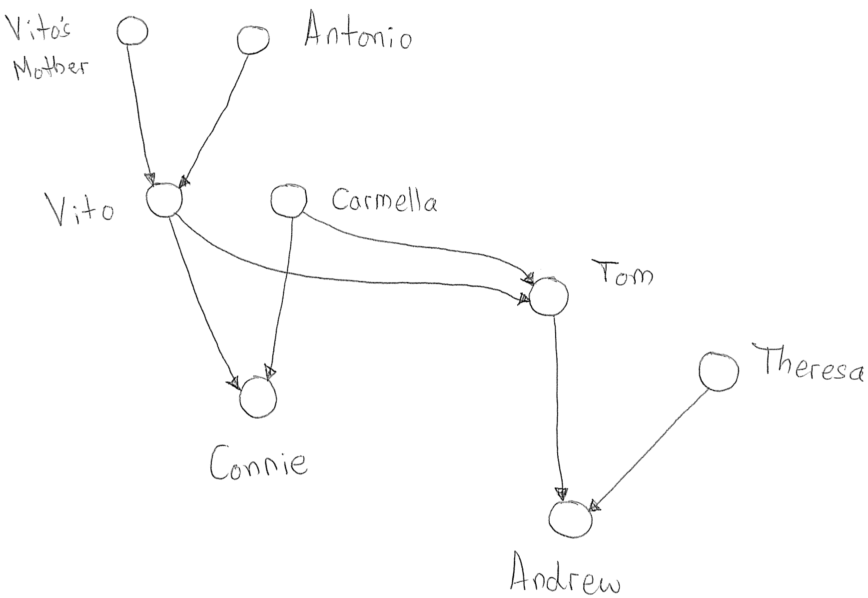

Do smallest ancestral sub-graph exercise with

two randomly chosen Simpson family members.



## Class Exercise

Draw the smallest ancestral sub-graph containing Connie and Andrew.




## Graph Theoretic Concepts

- Maximal cliques.
- Moralisation.
- Small ancestral sub-graphs.
- Separation.
- Markov blankets.

You need to be on top of all of the above from Chapter 1 of the *Graph Theory and Statistics Notes*
(which you will now start reading obsessively).

Hand out Assignment 2.

## Housekeeping

Getting set up for Laboratory 1 in the second half of Class 3.

See the relevant link on subject web-site.

ONE LAST REMINDER ABOUT:

## Class 3 Location

Class 3 is back in

Room 15, Level 2, Building 7

(i.e. next door and where
we held Class 1 last week).