Conditional independence: the algebra behind the models

"Algebraic & Geometric Methods in Statistics"

Sonja Petrović
Created for Math/Stat 561

Jan 27, 2026.

Objective

Understand how to translate conditional independence statements into polynomials, and what these polynomials mean.

Example Macaulay2 code

```
Macaulay2, version 1.18
i1 : loadPackage "GraphicalModels";
i2 : R = markovRing (3,2,2,2);
o2 =  R
o2 :  PolynomialRing
i3 : rvNames = {gender,hair,soccer,brazil}
o3 =  {gender,hair,soccer,brazil}
o3 :  List
i4 : CIstatements = { {{gender},{hair},{soccer}} }
-- this says gender indep. of hair given soccer
o4 =  {{{gender},{hair},{soccer}}}
o4 :  List
i5 : conditionalIndependenceIdeal(R,CIstatements,rvNames)
```

You can compute this online yourself.. The lines starting with "i" are input lines that you type into the editor to execute them.

$\texttt{ideal}(-p_{1,2,1,1}p_{2,1,1,1} - p_{1,2,1,2}p_{2,1,1,1} - p_{1,2,1,1}p_{2,1,1,2} - p_{1,2,1,2}p_{2,1,1,2} + p_{1,1,1,1}p_{2,2,1,1} + p_{1,1,1,2}p_{2,2,1,1} + p_{1,1,1,1}p_{2,2,1,2} + p_{1,1,1,2}p_{2,2,1,2},$

$-p_{1,2,1,1}p_{3,1,1,1} - p_{1,2,1,2}p_{3,1,1,1} - p_{1,2,1,1}p_{3,1,1,2} - p_{1,2,1,2}p_{3,1,1,2} + p_{1,1,1,1}p_{3,2,1,1} + p_{1,1,1,2}p_{3,2,1,1} + p_{1,1,1,1}p_{3,2,1,2} + p_{1,1,1,2}p_{3,2,1,2},$

$-p_{2,2,1,1}p_{3,1,1,1} - p_{2,2,1,2}p_{3,1,1,1} - p_{2,2,1,1}p_{3,1,1,2} - p_{2,2,1,2}p_{3,1,1,2} + p_{2,1,1,1}p_{3,2,1,1} + p_{2,1,1,2}p_{3,2,1,1} + p_{2,1,1,1}p_{3,2,1,2} + p_{2,1,1,2}p_{3,2,1,2},$

$-p_{1,2,2,1}p_{2,1,2,1} - p_{1,2,2,2}p_{2,1,2,1} - p_{1,2,2,1}p_{2,1,2,2} - p_{1,2,2,2}p_{2,1,2,2} + p_{1,1,2,1}p_{2,2,2,1} + p_{1,1,2,2}p_{2,2,2,1} + p_{1,1,2,1}p_{2,2,2,2} + p_{1,1,2,2}p_{2,2,2,2},$

$-p_{1,2,2,1}p_{3,1,2,1} - p_{1,2,2,2}p_{3,1,2,1} - p_{1,2,2,1}p_{3,1,2,2} - p_{1,2,2,2}p_{3,1,2,2} + p_{1,1,2,1}p_{3,2,2,1} + p_{1,1,2,2}p_{3,2,2,1} + p_{1,1,2,1}p_{3,2,2,2} + p_{1,1,2,2}p_{3,2,2,2},$

$-p_{2,2,2,1}p_{3,1,2,1} - p_{2,2,2,2}p_{3,1,2,1} - p_{2,2,2,1}p_{3,1,2,2} - p_{2,2,2,2}p_{3,1,2,2} + p_{2,1,2,1}p_{3,2,2,1} + p_{2,1,2,2}p_{3,2,2,1} + p_{2,1,2,1}p_{3,2,2,2} + p_{2,1,2,2}p_{3,2,2,2})$

Class work:

Determine why these are correct. worksheet 2.

Homework 1, problem 3 [due soon!]

Example 3.1.6 (Marginal independence)**.** The (marginal) independence statement $X_1 \perp\!\!\!\perp X_2$, or equivalently, $X_1 \perp\!\!\!\perp X_2 \mid X_\emptyset$, amounts to saying that the matrix

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1r_2} \\ p_{21} & p_{22} & \cdots & p_{2r_2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r_1 1} & p_{r_1 2} & \cdots & p_{r_1 r_2} \end{pmatrix}$$

has rank one. The independence ideal $I_{1 \perp\!\!\!\perp 2}$ is generated by the 2×2-minors:

$$I_{1 \perp\!\!\!\perp 2} = \left\langle p_{i_1 i_2} p_{j_1 j_2} - p_{i_1 j_2} p_{i_2 j_1} \mid i_1, j_1 \in [r_1], i_2, j_2 \in [r_2] \right\rangle.$$

For marginal independence, we already saw these quadratic binomial constraints in Chapter 1. □

Figure 1: From the Lectures on Algebraic Statistics book

Algebraic varieties and polynomial ideals

We have already seen these in Lecture 1, and in homework 1 (problem 4). Here is a brief overview of what you need to know.

- A variety is the solution set to a simultaneous system of polynomial equations.
- If I is an ideal[1], then V (I) is the variety defined by the vanishing of *all* polynomials in I.
- Hilbert basis theorem: even if I is infinite (it is!), there exists a finite basis for every I.

Important questions for statistics [digest next 2 slides]:

- what are points in a variety?
- how do you check if a point is on a variety?
- what if you are given an observation of 3 binary random variables, can you use polynomials to check some CI statements?

[1] an *ideal* is the infinite set of polynomial combination of some generating set.

Recall Example 1.1.2 from the book: 3-step Markov chain.

$p_{ijk} = P(X_1 = i, X_2 = j, X_3 = k)$ and $P(X_3 = k | X_1 = i, X_2 = j) = \frac{p_{ijk}}{p_{ij+}}$.

You verified that a probability distribution, represented by a vector of probabilities $p = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) \in \mathbb{R}^8$, being in this model is *equivalent to* the following four conditions:

$$p_{ijk} \geq 0 \text{ for all } i, j, k \in \{0, 1\}, \qquad \sum_{i,j,k} p_{ijk} = 1,$$

$$p_{000}p_{101} - p_{001}p_{100} = 0, \text{ and } p_{010}p_{111} - p_{011}p_{110} = 0.$$

- In this example: **what is the variety**?

Recall Example 1.1.2 from the book: 3-step Markov chain.

$p_{ijk} = P(X_1 = i, X_2 = j, X_3 = k)$ and $P(X_3 = k | X_1 = i, X_2 = j) = \frac{p_{ijk}}{p_{ij+}}$.

You verified that a probability distribution, represented by a vector of probabilities $p = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) \in \mathbb{R}^8$, being in this model is *equivalent to* the following four conditions:

$$p_{ijk} \geq 0 \text{ for all } i, j, k \in \{0, 1\}, \qquad \sum_{i,j,k} p_{ijk} = 1,$$

$$p_{000}p_{101} - p_{001}p_{100} = 0, \text{ and } p_{010}p_{111} - p_{011}p_{110} = 0.$$

- In this example: **what is the variety**?
- Is the point $(1/8, 1/8, \ldots, 1/8)$ **on** this variety? (That is, is this joint probability vector in the model?)
  - [Verify right now, by hand.]

Recall Example 1.1.2 from the book: 3-step Markov chain.

$p_{ijk} = P(X_1 = i, X_2 = j, X_3 = k)$ and $P(X_3 = k | X_1 = i, X_2 = j) = \frac{p_{ijk}}{p_{ij+}}$.

You verified that a probability distribution, represented by a vector of probabilities $p = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) \in \mathbb{R}^8$, being in this model is *equivalent to* the following four conditions:

$$p_{ijk} \geq 0 \text{ for all } i, j, k \in \{0, 1\}, \qquad \sum_{i,j,k} p_{ijk} = 1,$$

$$p_{000}p_{101} - p_{001}p_{100} = 0, \text{ and } p_{010}p_{111} - p_{011}p_{110} = 0.$$

- In this example: **what is the variety**?
- Is the point $(1/8, 1/8, \ldots, 1/8)$ **on** this variety? (That is, is this joint probability vector in the model?)
  - [Verify right now, by hand.]
- Find an example of a point on the variety, which is a point in this model.
  - [Discussion of HW 1.b).]

Varieties

Points on the model $\leftrightarrow$ points on the variety.

What are model *ideals*?

If $p_{000}p_{101} - p_{001}p_{100} = 0$ and $p_{010}p_{111} - p_{011}p_{110} = 0$ both hold for a probability vector $p = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) \in \mathbb{R}^8$, then **what other polynomial equations also hold**?

Varieties

Points on the model $\leftrightarrow$ points on the variety.

What are model *ideals*?

If $p_{000}p_{101} - p_{001}p_{100} = 0$ and $p_{010}p_{111} - p_{011}p_{110} = 0$ both hold for a probability vector $p = (p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111}) \in \mathbb{R}^8$, then **what other polynomial equations also hold**?

- *All* such equations vanish on the points of the model.
  - [Discussion.]
- The (infinite) collection of all of these polynomials is the ideal of this statistical model.
- It is an equivalent description of the model.
  - [$\rightarrow$ algebraic geometry.]

Ideals

The (defining) ideal of a model is the set of all polynomials that simultaneously vanish on all points in the model.

Luckily, there is a finite basis for each ideal.

How to combine several CI statements?

Sum of ideals.

Example 3.1.10. Let X_1, X_2, X_3, X_4 be binary random variables, and consider the conditional independence model

$$\mathcal{C} = \{1 \perp\!\!\!\perp 3 \,|\, \{2,4\}, 2 \perp\!\!\!\perp 4 \,|\, \{1,3\}\}.$$

These are the conditional independence statements that hold for the graphical model associated to the four cycle graph with edges $\{12, 23, 34, 14\}$; see Section 3.2. The conditional independence ideal is generated by eight quadratic binomials:

$$\begin{aligned} I_{\mathcal{C}} &= I_{1 \perp\!\!\!\perp 3 \,|\, \{2,4\}} + I_{2 \perp\!\!\!\perp 4 \,|\, \{1,3\}} \\ &= \langle p_{1111}p_{2121} - p_{1121}p_{2111}, p_{1112}p_{2122} - p_{1122}p_{2112}, \\ &\quad p_{1211}p_{2221} - p_{1221}p_{2211}, p_{1212}p_{2222} - p_{1222}p_{2212}, \\ &\quad p_{1111}p_{1212} - p_{1112}p_{1211}, p_{1121}p_{1222} - p_{1122}p_{1221}, \\ &\quad p_{2111}p_{2212} - p_{2112}p_{2211}, p_{2121}p_{2222} - p_{2122}p_{2221}\rangle\,. \end{aligned}$$

Figure 2: From the Lectures on Algebraic Statistics book

Appendix

Here are some additional examples you may wish to explore, to familiarize yourself with conditional independence:

- This is where the Martin&Normal example came from.
- This informal website has some additional interesting examples.
- Here is a set of slides with several real-world examples of CI random variables.

License

This document is created for Math/Stat 561, Spring 2023 (2026 update), at Illinois Tech.

The first example is from Kaie Kubjas' course. Other online sources are cited throughout.