FROM STATISTICS TO ALGEBRA AND BACK AGAIN

Courtney Gibbons

September 9, 2016

Hamilton College

GOALS

- Goal 1: Use algebra to solve a statistics problem.

GOALS

- Goal 1: Use algebra to solve a statistics problem.
- Goal 1.5: Traipse through centuries of mathematics to turn this stats problem into an algebra problem. [1]

GOALS

- Goal 1: Use algebra to solve a statistics problem.
- Goal 1.5: Traipse through centuries of mathematics to turn this stats problem into an algebra problem. [1]
- Goal 2: Convince you that questions from statistics lead to new (and renewed) algebraic lines of inquiry.

Thanks!

[1] Mathias Drton, Bernd Sturmfels, and Seth Sullivant, Lectures on algebraic statistics, Oberwolfach Seminars, vol. 39, Birkhäuser Verlag, Basel, 2009. MR2723140

[2] Wolfram and Greuel Decker Gert-Martin and Pfister, SINGULAR 4-0-2 — A computer algebra system for polynomial computations, 2015. `http://www.singular.uni-kl.de`.

[3] Likelihood Geometry Working Group led by Serkan Hosten, Toric Models, Discriminants, and Homotopies. Work in progress based upon work supported by the National Science Foundation under Grant Number DMS 1321794.

GAMBLER WITH 2 BIASED COINS

- Trial: Flip a coin four times, record # of heads (200 total trials)

- Trial: Flip a coin four times, record # of heads (200 total trials)
- Data vector: $H = (H_0, H_1, H_2, H_3, H_4)$ where, eg, H_0 is the number of trials in which there were 0 observed heads. Note that $H_0 + H_1 + H_2 + H_3 + H_4 = 200$.

We have reason to suspect the gambler is using two biased coins (one up each sleeve) with probability of heads ℓ (left sleeve) and r (right sleeve). He chooses between them with probability p and $1 - p$ respectively.

GAMBLER WITH 2 BIASED COINS

- Trial: Flip a coin four times, record # of heads (200 total trials)
- Data vector: $H = (H_0, H_1, H_2, H_3, H_4)$ where, eg, H_0 is the number of trials in which there were 0 observed heads. Note that $H_0 + H_1 + H_2 + H_3 + H_4 = 200$.

We have reason to suspect the gambler is using two biased coins (one up each sleeve) with probability of heads ℓ (left sleeve) and r (right sleeve). He chooses between them with probability p and $1 - p$ respectively.

- Trial: Flip a coin four times, record # of heads (200 total trials)
- Data vector: $H = (H_0, H_1, H_2, H_3, H_4)$ where, eg, H_0 is the number of trials in which there were 0 observed heads. Note that $H_0 + H_1 + H_2 + H_3 + H_4 = 200$.

We have reason to suspect the gambler is using two biased coins (one up each sleeve) with probability of heads ℓ (left sleeve) and r (right sleeve). He chooses between them with probability p and $1 - p$ respectively.

- Our task: use the data vector H to estimate the parameters p, ℓ, and r.

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$$
\begin{aligned}
P_0 &= && \text{TTTT} \\
P_1 &= && \text{HTTT, THTT, TTHT, TTTH} \\
P_2 &= && \text{HHTT, HTHT, HTTH, THHT, THTH, TTHH} \\
P_3 &= && \text{HHHT, HHTH, HTHH, THHH} \\
P_4 &= && \text{HHHH}
\end{aligned}
$$

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$$
\begin{aligned}
P_0 &= (1-\ell)^4 && \text{TTTT} \\
P_1 &= && \text{HTTT, THTT, TTHT, TTTH} \\
P_2 &= && \text{HHTT, HTHT, HTTH, THHT, THTH, TTHH} \\
P_3 &= && \text{HHHT, HHTH, HTHH, THHH} \\
P_4 &= && \text{HHHH}
\end{aligned}
$$

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$P_0 = (1-\ell)^4$	TTTT
$P_1 = 4\ \ell(1-\ell)^3$	HTTT, THTT, TTHT, TTTH
$P_2 =$	HHTT, HTHT, HTTH, THHT, THTH, TTHH
$P_3 =$	HHHT, HHTH, HTHH, THHH
$P_4 =$	HHHH

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$P_0 = \quad (1-\ell)^4$	TTTT
$P_1 = 4\ \ \ell(1-\ell)^3$	HTTT, THTT, TTHT, TTTH
$P_2 = 6\ \ \ell^2(1-\ell)^2$	HHTT, HTHT, HTTH, THHT, THTH, TTHH
$P_3 = 4\ \ \ell^3(1-\ell)$	HHHT, HHTH, HTHH, THHH
$P_4 = \quad \ell^4$	HHHH

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$P_0 = (1-\ell)^4 + (1-r)^4$		TTTT
$P_1 = 4\ \ell(1-\ell)^3 + 4\ r(1-r)^3$		HTTT, THTT, TTHT, TTTH
$P_2 = 6\ \ell^2(1-\ell)^2 + 6\ r^2(1-r)^2$		HHTT, HTHT, HTTH, THHT, THTH, TTHH
$P_3 = 4\ \ell^3(1-\ell) + 4\ r^3(1-r)$		HHHT, HHTH, HTHH, THHH
$P_4 = \ell^4 + r^4$		HHHH

We can calculate the probabilities of getting 0, 1, 2, 3, or 4 heads as follows:

$$
\begin{aligned}
P_0 &= p(1-\ell)^4 + (1-p)(1-r)^4 && \text{TTTT} \\
P_1 &= 4p\ell(1-\ell)^3 + 4(1-p)r(1-r)^3 && \text{HTTT, THTT, TTHT, TTTH} \\
P_2 &= 6p\ell^2(1-\ell)^2 + 6(1-p)r^2(1-r)^2 && \text{HHTT, HTHT, HTTH, THHT, THTH, TTHH} \\
P_3 &= 4p\ell^3(1-\ell) + 4(1-p)r^3(1-r) && \text{HHHT, HHTH, HTHH, THHH} \\
P_4 &= p\ell^4 + (1-p)r^4 && \text{HHHH}
\end{aligned}
$$

The likelihood of observing our data vector H after 200 trials is given by

$$\frac{200!}{H_0!H_1!H_2!H_3!H_4!} P_0^{H_0} P_1^{H_1} P_2^{H_2} P_3^{H_3} P_4^{H_4}. \tag{MLE}$$

The likelihood of observing our data vector H after 200 trials is given by

$$\frac{200!}{H_0!H_1!H_2!H_3!H_4!} P_0^{H_0} P_1^{H_1} P_2^{H_2} P_3^{H_3} P_4^{H_4}. \tag{MLE}$$

Find values of P_0, P_1, P_2, P_3, P_4 (and hence p, ℓ, r) that maximize Equation (MLE) subject to $0 < p, \ell, r < 1$.

The likelihood of observing our data vector H after 200 trials is given by

$$\frac{200!}{H_0!H_1!H_2!H_3!H_4!} P_0^{H_0} P_1^{H_1} P_2^{H_2} P_3^{H_3} P_4^{H_4}. \tag{MLE}$$

Find values of P_0, P_1, P_2, P_3, P_4 (and hence p, ℓ, r) that maximize Equation (MLE) subject to $0 < p, \ell, r < 1$.

However, this is not an algebraic problem (yet).

There's a relationship between the P_i's that we can capture in a matrix.
Define

$$P = 12p \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix} \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix}^T + 12(1-p) \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix} \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix}^T \quad (\text{rank} < 3)$$

There's a relationship between the P_i's that we can capture in a matrix.
Define

$$P = 12p \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix} \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix}^T + 12(1-p) \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix} \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix}^T \quad (\text{rank} < 3)$$

$$P_0 = p(1-\ell)^4 + (1-p)(1-r)^4,$$

$$P_1 = 4p\ell(1-\ell)^3 + 4(1-p)r(1-r)^3,$$

$$P_2 = 6p\ell^2(1-\ell)^2 + 6(1-p)r^2(1-r)^2, \text{etc.}$$

There's a relationship between the P_i's that we can capture in a matrix.
Define

$$P = 12p \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix} \begin{bmatrix} (1-\ell)^2 \\ \ell(1-\ell) \\ \ell^2 \end{bmatrix}^T + 12(1-p) \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix} \begin{bmatrix} (1-r)^2 \\ \ell(1-r) \\ r^2 \end{bmatrix}^T \quad (\text{rank} < 3)$$

$$= \begin{bmatrix} 12P_0 & 3P_1 & 2P_2 \\ 3P_1 & 2P_2 & 3P_3 \\ 2P_2 & 3P_3 & 12P_4 \end{bmatrix}.$$

MAXIMIZING THE MLE AFTER IMPLICITIZATION

Better goal: maximize

$$\ln\left(\frac{200!}{H_0!H_1!H_2!H_3!H_4!}P_0^{H_0}P_1^{H_1}P_2^{H_2}P_3^{H_3}P_4^{H_4}\right)$$

subject to $P_0 + P_1 + P_2 + P_3 + P_4 = 1$ and $\det(P) = 0$.

Better goal: maximize

$$\ln\left(\frac{200!}{H_0!H_1!H_2!H_3!H_4!}P_0^{H_0}P_1^{H_1}P_2^{H_2}P_3^{H_3}P_4^{H_4}\right)$$

subject to $P_0+P_1+P_2+P_3+P_4=1$ and $\det(P)=0$.

$$\det P=\begin{vmatrix}12P_0 & 3P_1 & 2P_2\\ 3P_1 & 2P_2 & 3P_3\\ 2P_2 & 3P_3 & 12P_4\end{vmatrix}=288P_0P_2P_4-108P_0P_3^2+36P_1P_2P_3-108P_1^2P_4-8P_2^3$$

Why P?

Why ln?

SOME REMARKS

Why P? We can think of the image $(p, \ell, r) \mapsto (P_0, P_1, P_2, P_3, P_4) \in \mathbb{P}^4$ as lying on the hyperplane $\{\det(P) = 0\}$. (Now that's algebra!)

Why ln?

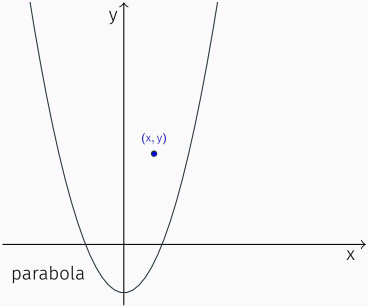
Why P? We can think of the image $(p, \ell, r) \mapsto (P_0, P_1, P_2, P_3, P_4) \in \mathbb{P}^4$ as lying on the hyperplane $\{\det(P) = 0\}$. (Now that's algebra!)

Why ln? To make taking derivatives easier:

$$\ln\left(\frac{200!}{H_0!H_1!H_2!H_3!H_4!}P_0^{H_0}P_1^{H_1}P_2^{H_2}P_3^{H_3}P_4^{H_4}\right)$$
$$= \text{Constant} + H_0\ln(P_0) + H_1\ln(P_1) + H_2\ln(P_2) + H_3\ln(P_3) + H_4\ln(P_4).$$

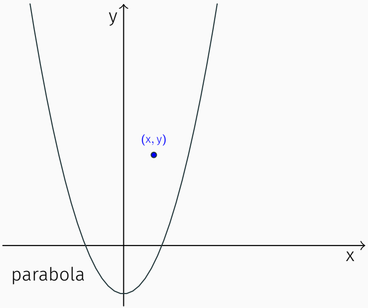
WHAT THE HECK IS PROJECTIVE SPACE?

Euclidean 2-Space

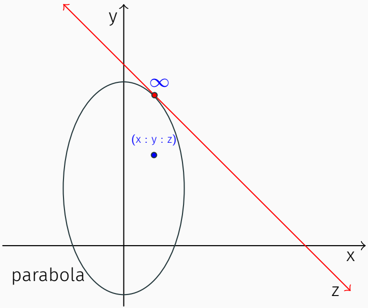


WHAT THE HECK IS PROJECTIVE SPACE?

Euclidean 2-Space



Projective 2-Space



Recall

$$\det P = 288P_0P_2P_4 - 108P_0P_3^2 + 36P_1P_2P_3 - 108P_1^2P_4 - 8P_2^3.$$

To maximize subject to the constraint $\det(P) = 0$, we use Lagrange Multipliers:

$$\nabla(H_0 \ln(P_0) + H_1 \ln(P_1) + H_2 \ln(P_2) + H_3 \ln(P_3) + H_4 \ln(P_4)) = \lambda \det P$$

($P_0 + P_1 + P_2 + P_3 + P_4 = 1$ comes into play later).

LAGRANGE MULTIPLIER EQUATIONS

Assume none of the P_i are zero.

Using Lagrange Multipliers produces a system of rational equations in P_0 through P_4:

$$
\begin{aligned}
\frac{H_0}{P_0} &= \lambda(288P_2P_4 - 108P_3^2) \\
\frac{H_1}{P_1} &= \lambda(36P_2P_3 - 216P_1P_4) \\
\frac{H_2}{P_2} &= \lambda(288P_0P_4 + 36P_1P_3 - 24P_2^2) \\
\frac{H_3}{P_3} &= \lambda(-216P_0P_3 + 36P_1P_2) \\
\frac{H_4}{P_4} &= \lambda(288P_0P_2 - 108P_1^2).
\end{aligned}
$$

$$
\begin{aligned}
H_0 &= \lambda(288P_0P_2P_4 - 108P_0P_3^2) \\
H_1 &= \lambda(36P_1P_2P_3 - 216P_1^2P_4) \\
H_2 &= \lambda(288P_0P_2P_4 + 36P_1P_2P_3 - 24P_2^3) \\
H_3 &= \lambda(-216P_0P_3^2 + 36P_1P_2P_3) \\
H_4 &= \lambda(288P_0P_2P_4 - 108P_1^2P_4).
\end{aligned}
$$

Exercise: Add these together and use $P_0 + P_1 + P_2 + P_3 + P_4 = 1$ to show that $\lambda = H_0 + H_1 + H_2 + H_3 + H_4 = 200$.

Observed: $H = (4, 18, 25, 80, 73)$; total 200 trials.

By subtracting the H_i, we obtain 5 polynomials that we will set to zero to find critical numbers:

$$\begin{aligned}
q_1 &= 200(288P_0P_2P_4 - 108P_0P_3^2) - 4 \\
q_2 &= 200(36P_1P_2P_3 - 216P_1^2P_4) - 18 \\
q_3 &= 200(288P_0P_2P_4 + 36P_1P_2P_3 - 24P_2^3) - 25 \\
q_4 &= 200(-216P_0P_3^2 + 36P_1P_2P_3) - 80 \\
q_5 &= 200(288P_0P_2P_4 - 108P_1^2P_4) - 73
\end{aligned}$$

Critical numbers: points in $\mathbb{P}^4$ where $q_1 = q_2 = q_3 = q_4 = q_5 = 0$.

An **algebraic variety** $V(q_1, \ldots, q_n)$ is a set of points (in $\mathbb{P}^{n-1}$ for today) where the polynomials $q_1, \ldots, q_n$ are simultaneously zero.

Algebraic geometers have software for calculating varieties, which I shamelessly used to find all twelve points in this variety. [2]

	Sol'n 1	2	3	4	5	6
P_0	−0.008	−0.041	−1.406	−0.499	0.002734006	0.026462824
P_1	−1.699	−6.754	0.218	0.067	0.054826753	0.071480775
P_2	2.482	13.110	3.686	1.290	0.18659817	0.1598897
P_3	1.163	−4.798	−1.404	0.218	0.45138739	0.35992698
P_4	−0.939	−0.519	−0.095	−0.075	0.30445369	0.38223972
	Sol'n 7	8	9	10	11	12
P_0	0.432	0.203	0.002 − 0.001i	0.002 + 0.001i	0.020 + 0.025i	0.020 − 0.025i
P_1	−0.657	0.040	0.033 + 0.024i	0.033 − 0.024i	0.099 − 0.055i	0.099 + 0.055i
P_2	0.493	1.119	0.267 + 0.001i	0.267 − 0.001i	0.099 − 0.007i	0.099 + 0.007i
P_3	0.256	−2.172	0.359 − 0.065i	0.359 + 0.066i	0.423 + 0.079i	0.423 − 0.079i
P_4	0.476	1.809	0.338 + 0.043i	0.338 − 0.043i	0.357 − 0.042i	0.357 + 0.042i

An **algebraic variety** $V(q_1, \ldots, q_5)$ is a set of points (in $\mathbb{P}^4$ for today) where the polynomials $q_1, \ldots, q_5$ are simultaneously zero.

Algebraic geometers have software for calculating varieties, which I shamelessly used to find all twelve points in this variety. [2]

	Sol'n 1	2	3	4	5	6
P_0	-0.008	-0.041	-1.406	-0.499	0.002734006	0.026462824
P_1	-1.699	-6.754	0.218	0.067	0.054826753	0.071480775
P_2	2.482	13.110	3.686	1.290	0.18659817	0.1598897
P_3	1.163	-4.798	-1.404	0.218	0.45138739	0.35992698
P_4	-0.939	-0.519	-0.095	-0.075	0.30445369	0.38223972
	Sol'n 7	8	9	10	11	12
P_0	0.432	0.203	$0.002 - 0.001i$	$0.002 + 0.001i$	$0.020 + 0.025i$	$0.020 - 0.025i$
P_1	-0.657	0.040	$0.033 + 0.024i$	$0.033 - 0.024i$	$0.099 - 0.055i$	$0.099 + 0.055i$
P_2	0.493	1.119	$0.267 + 0.001i$	$0.267 - 0.001i$	$0.099 - 0.007i$	$0.099 + 0.007i$
P_3	0.256	-2.172	$0.359 - 0.065i$	$0.359 + 0.066i$	$0.423 + 0.079i$	$0.423 - 0.079i$
P_4	0.476	1.809	$0.338 + 0.043i$	$0.338 - 0.043i$	$0.357 - 0.042i$	$0.357 + 0.042i$

An **algebraic variety** $V(q_1, \ldots, q_5)$ is a set of points (in $\mathbb{P}^4$ for today) where the polynomials $q_1, \ldots, q_5$ are simultaneously zero.

Algebraic geometers have software for calculating varieties, which I shamelessly used to find all twelve points in this variety. [2]

	Sol'n 1	2	3	4	5	6
P_0	−0.008	−0.041	−1.406	−0.499	0.002734006	0.026462824
P_1	−1.699	−6.754	0.218	0.067	0.054826753	0.071480775
P_2	2.482	13.110	3.686	1.290	0.18659817	0.1598897
P_3	1.163	−4.798	−1.404	0.218	0.45138739	0.35992698
P_4	−0.939	−0.519	−0.095	−0.075	0.30445369	0.38223972
	Sol'n 7	8	9	10	11	12
P_0	0.432	0.203	0.002 − 0.001i	0.002 + 0.001i	0.020 + 0.025i	0.020 − 0.025i
P_1	−0.657	0.040	0.033 + 0.024i	0.033 − 0.024i	0.099 − 0.055i	0.099 + 0.055i
P_2	0.493	1.119	0.267 + 0.001i	0.267 − 0.001i	0.099 − 0.007i	0.099 + 0.007i
P_3	0.256	−2.172	0.359 − 0.065i	0.359 + 0.066i	0.423 + 0.079i	0.423 − 0.079i
P_4	0.476	1.809	0.338 + 0.043i	0.338 − 0.043i	0.357 − 0.042i	0.357 + 0.042i

An **algebraic variety** $V(q_1, \ldots, q_5)$ is a set of points (in $\mathbb{P}^4$ for today) where the polynomials $q_1, \ldots, q_5$ are simultaneously zero.

Algebraic geometers have software for calculating varieties, which I shamelessly used to find all twelve points in this variety. [2]

	Sol'n 1	2	3	4	5	6
P_0	-0.008	-0.041	-1.406	-0.499	0.002734006	0.026462824
P_1	-1.699	-6.754	0.218	0.067	0.054826753	0.071480775
P_2	2.482	13.110	3.686	1.290	0.18659817	0.1598897
P_3	1.163	-4.798	-1.404	0.218	0.45138739	0.35992698
P_4	-0.939	-0.519	-0.095	-0.075	0.30445369	0.38223972
	Sol'n 7	8	9	10	11	12
P_0	0.432	0.203	$0.002 - 0.001i$	$0.002 + 0.001i$	$0.020 + 0.025i$	$0.020 - 0.025i$
P_1	-0.657	0.040	$0.033 + 0.024i$	$0.033 - 0.024i$	$0.099 - 0.055i$	$0.099 + 0.055i$
P_2	0.493	1.119	$0.267 + 0.001i$	$0.267 - 0.001i$	$0.099 - 0.007i$	$0.099 + 0.007i$
P_3	0.256	-2.172	$0.359 - 0.065i$	$0.359 + 0.066i$	$0.423 + 0.079i$	$0.423 - 0.079i$
P_4	0.476	1.809	$0.338 + 0.043i$	$0.338 - 0.043i$	$0.357 - 0.042i$	$0.357 + 0.042i$

$$X = (0.002734006, 0.054826753, 0.18659817, 0.45138739, 0.30445369)$$
$$Y = (0.026462824, 0.071480775, 0.1598897, 0.35992698, 0.38223972)$$

Number sense says Y should be the maximum. Recall, the Likelihood Equation is

$$\mathrm{MLE}(P_0, P_1, P_2, P_3, P_4) = \frac{200!}{H_0!H_1!H_2!H_3!H_4!} P_0^{H_0} P_1^{H_1} P_2^{H_2} P_3^{H_3} P_4^{H_4}.$$

$$X = (0.002734006, 0.054826753, 0.18659817, 0.45138739, 0.30445369)$$
$$Y = (0.026462824, 0.071480775, 0.1598897, 0.35992698, 0.38223972)$$

Number sense says Y should be the maximum. Recall, the Likelihood Equation is

$$\mathrm{MLE}(P_0, P_1, P_2, P_3, P_4) = \frac{200!}{H_0!H_1!H_2!H_3!H_4!} P_0^{H_0} P_1^{H_1} P_2^{H_2} P_3^{H_3} P_4^{H_4}.$$

Indeed, $\mathrm{MLE}(X) = 3.14338 * 10^{-9} < \mathrm{MLE}(Y) = 1.52677 * 10^{-5}$

BACK TO THE ORIGINAL PROBLEM

Our original probalities give us:

$$\begin{aligned} P_0 &= p(1-\ell)^4 + (1-p)(1-r)^4 \\ P_1 &= 4p\ell(1-\ell)^3 + 4(1-p)r(1-r)^3 \\ P_2 &= 6p\ell^2(1-\ell)^2 + 6(1-p)r^2(1-r)^2 \\ P_3 &= 4p\ell^3(1-\ell) + 4(1-p)r^3(1-r) \\ P_4 &= p\ell^4 + (1-p)r^4 \end{aligned}$$

Pulling the same stunts, we consider the algebraic variety for five polynomials in the variables p, ℓ, and r:

$p = 0.13867$, $\ell = 0.345438$, and $r = 0.815136$ (or $p = 0.861324$, $\ell = 0.815136$, $r = 0.345438$).

Our original probalities give us:

$$\begin{aligned}
0.026462824 &= p(1-\ell)^4 + (1-p)(1-r)^4 \\
0.071480775 &= 4p\ell(1-\ell)^3 + 4(1-p)r(1-r)^3 \\
0.1598897 &= 6p\ell^2(1-\ell)^2 + 6(1-p)r^2(1-r)^2 \\
0.35992698 &= 4p\ell^3(1-\ell) + 4(1-p)r^3(1-r) \\
0.38223972 &= p\ell^4 + (1-p)r^4
\end{aligned}$$

Pulling the same stunts, we consider the algebraic variety for five polynomials in the variables p, ℓ, and r:

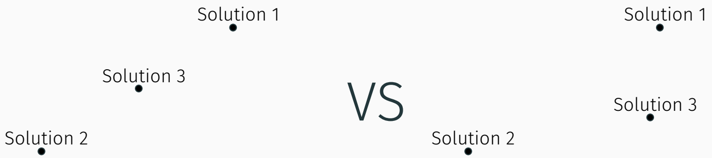
$p = 0.13867$, $\ell = 0.345438$, and $r = 0.815136$ (or $p = 0.861324$, $\ell = 0.815136$, $r = 0.345438$).

- Algebraic question: Define the **MLE Degree** to be the number of points in the variety obtained from simplifying the Lagrange Multiplier equations. What are the possible MLE Degrees for a given model? E.g., if we tinker with the H-vector, how many solutions can we get? What if we tinker with the implicitization or parametrize the model differently somehow?

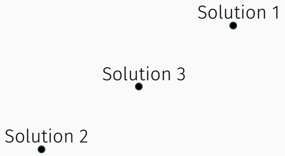
WE SOLVED ONE PROBLEM. NOW WHAT?

- Algebraic question: Define the **MLE Degree** to be the number of points in the variety obtained from simplifying the Lagrange Multiplier equations. What are the possible MLE Degrees for a given model? E.g., if we tinker with the H-vector, how many solutions can we get? What if we tinker with the implicitization or parametrize the model differently somehow?
- Geometric question: How can the MLE solutions in the variety be arranged in projective space?

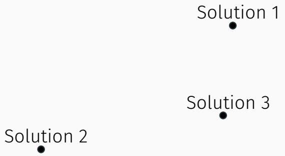
MLE DEGREE 3, BUT...



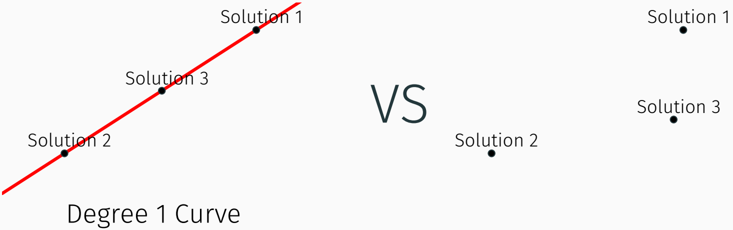
MLE DEGREE 3, BUT...



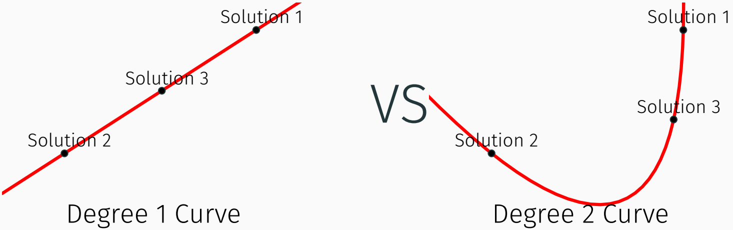
VS

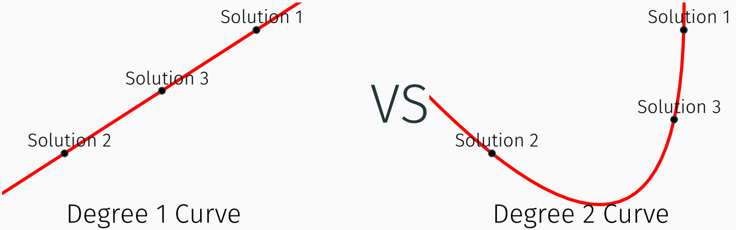


MLE DEGREE 3, BUT...



MLE DEGREE 3, BUT...





How does MLE Degree relate to the degree of the curve through the solutions?

PARTIAL ANSWERS (AND MORE QUESTIONS)

For statistical models that yield **Toric Varieties**, the MLE Degree can change based on the parametrization of the toric variety. (Segre, Veronese, Rational Normal Scrolls, etc.) [3]

Mixture models (like in our coin tossing example) can be examined using **Secant Varieties**.

REFERENCES

[1] Mathias Drton, Bernd Sturmfels, and Seth Sullivant, Lectures on algebraic statistics, Oberwolfach Seminars, vol. 39, Birkhäuser Verlag, Basel, 2009. MR2723140

[2] Wolfram and Greuel Decker Gert-Martin and Pfister, Singular 4-0-2 — A computer algebra system for polynomial computations, 2015. `http://www.singular.uni-kl.de`.

[3] Likelihood Geometry Working Group led by Serkan Hosten, Toric Models, Discriminants, and Homotopies. Work in progress based upon work supported by the National Science Foundation under Grant Number DMS 1321794.