Math Foundations for ML

10-606

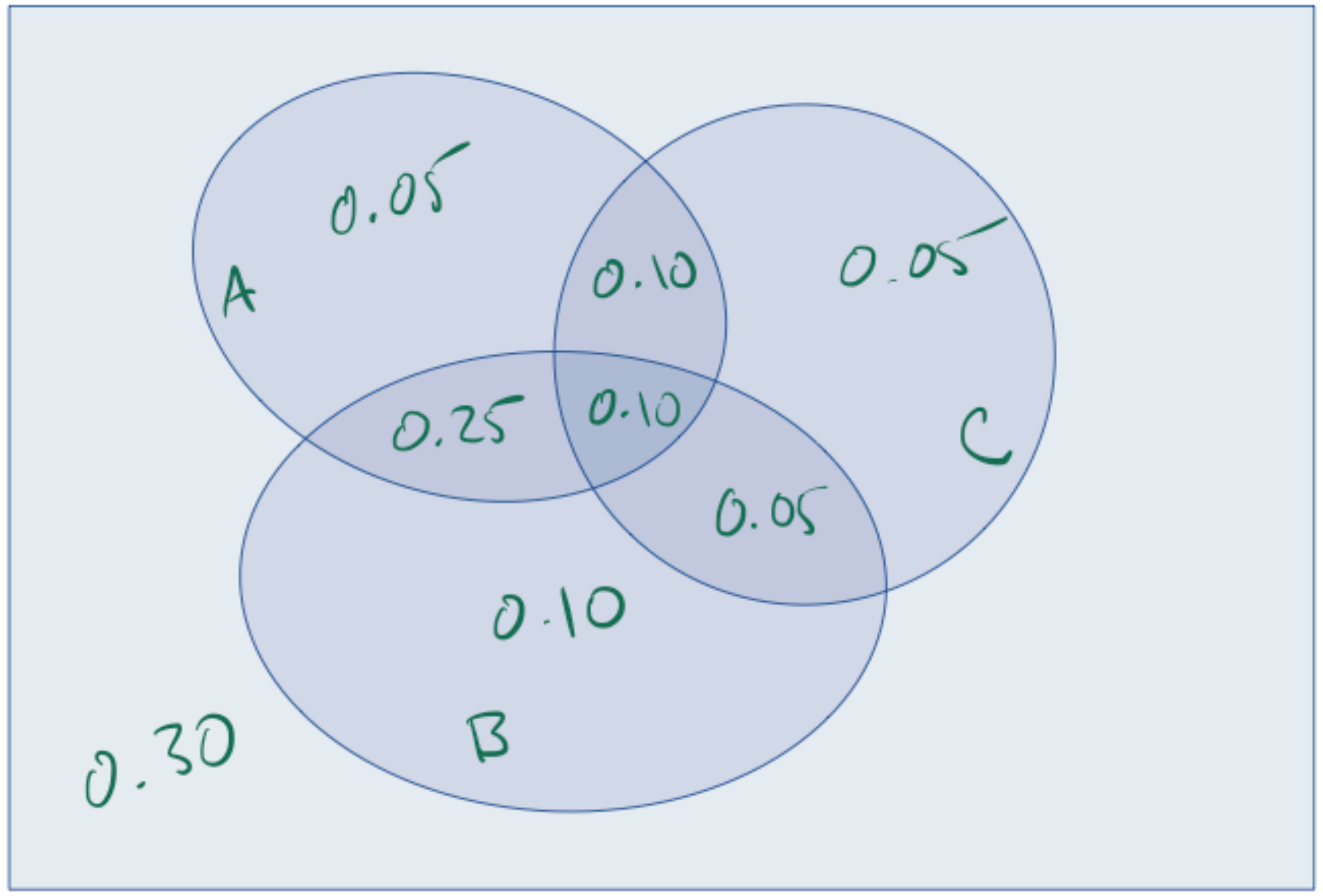
Geoff Gordon

Notes and reminders

5:30 p - 8:30 p
Fri Mar 4

- Upcoming: Lab 4 (Monday)
- Exam schedule should post soon (for time and place of Quiz 2)

Marginal, conditional



A	B	C	P	
0	0	0	0.30	0.35
0	0	1	0.05	
0	1	0	0.10	0.15
0	1	1	0.05	
1	0	0	0.05	0.15
1	0	1	0.10	
1	1	0	0.25	0.35
1	1	1	0.10	

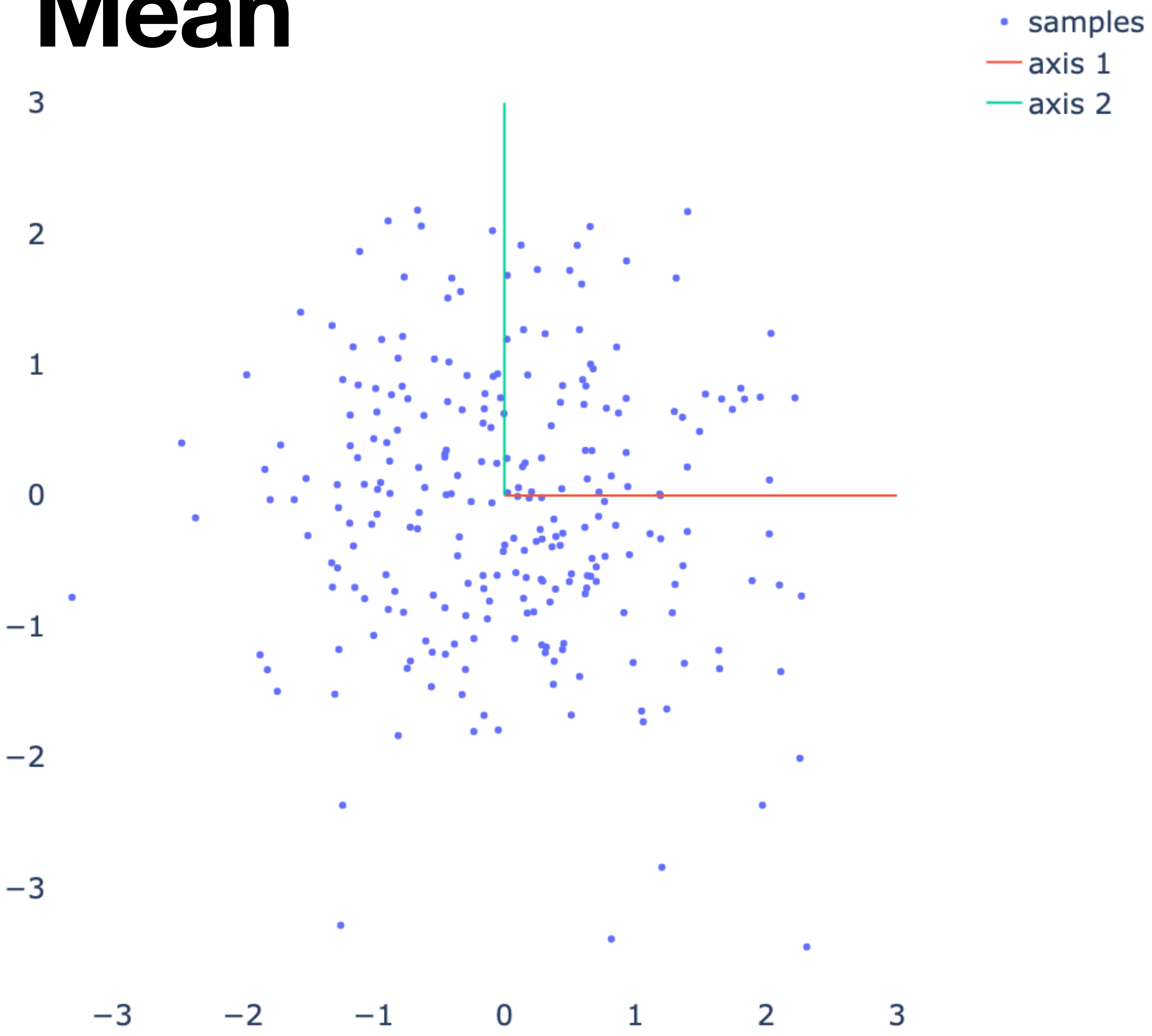
A	B	C	
~~0~~	~~0~~	~~0~~	~~.3~~
~~0~~	~~0~~	~~1~~	~~.05~~
0	1	0	.1
0	1	1	.05
~~1~~	~~0~~	~~0~~	~~.05~~
~~1~~	~~0~~	~~1~~	~~.1~~
1	1	0	.25
1	1	1	.1

given $B=1$

$P(A, C \mid B=1)$

A	C	
0	0	$.1/.5 = .2$
0	1	$.05/.5 = .1$
1	0	$.25/.5 = .5$
1	1	$.1/.5 = .2$

Mean



$\bar{X} = E(X) = \sum_x x \, P(X = x)$

$\frac{1}{6} 1 + \frac{1}{6} 2 + \ldots \frac{1}{6} \cdot 6 = 3\frac{1}{2}$

$E(aX + bY + c)$

$= aE(X) + bE(Y) + c$

$$V(X) = E\left[(X - \bar{X})^2\right]$$

$\uparrow E(X)$

$$V(\text{coin}) = \frac{1}{2}\left(-\frac{1}{2}\right)^2 + \frac{1}{2}\left(+\frac{1}{2}\right)^2$$

$$= \frac{1}{4}$$

$$\sigma(X) = \sqrt{V(X)}$$

height is 160 cm $\pm$ 20 cm

$$E(|X - \bar{X}|)$$

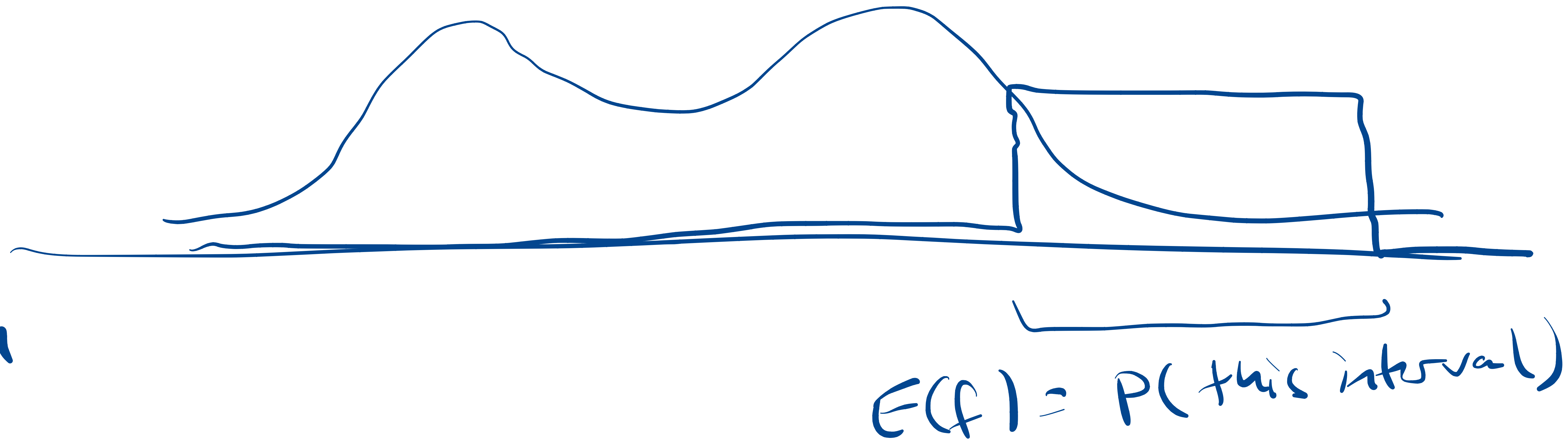
MAD

$$E\left(X^2 - 2X\bar{X} - \bar{X}^2\right)$$

$$= E(X^2) - \underbrace{2\bar{X}E(X)}_{2\bar{X}^2} - \bar{X}^2$$

$$= E(X^2) - \bar{X}^2$$

Moment: $E(f(x))$

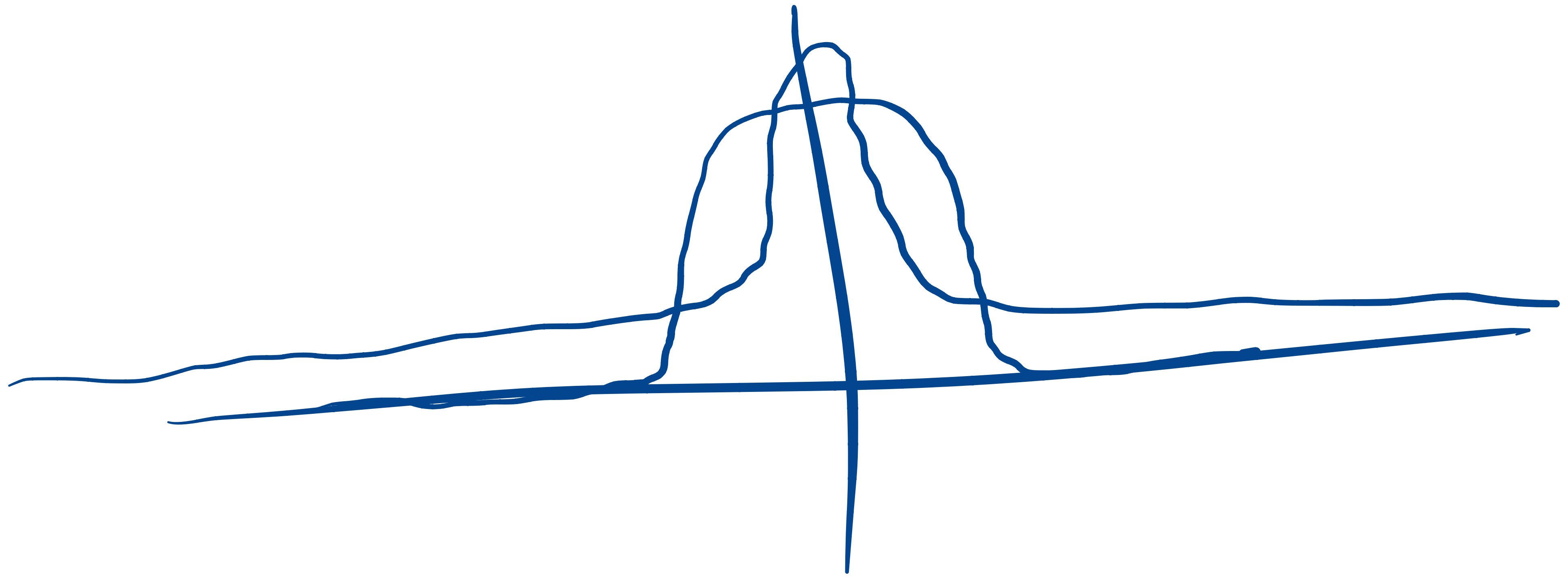


$f =$ identity $\rightarrow$ mean

$f(x) = (x-\bar{x})^2 \rightarrow$ variance

$f(x) = (x-\bar{x})^k$

$f(x) = x^k$



$X, Y \in \mathbb{R}$

$$\mathrm{Cov}(X,Y) = E[(X-\bar{X})(Y-\bar{Y})]$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)} = E\left(\frac{X-\bar{X}}{\sigma_x} \cdot \frac{Y-\bar{Y}}{\sigma_Y}\right)$$

$\in [-1, 1]$

$$\mathrm{Var}\left(\frac{X-\bar{X}}{\sigma_x}\right) = 1$$

$$E\left(\left(\frac{X-\bar{X}}{\sigma_x}\right)^2\right) = \frac{1}{\sigma_x^2} \underbrace{E((X-\bar{X})^2)}_{V(X) = \sigma_x^2} = 1$$

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

$tr(V(X))$

avg squared dist from mean

$$\Sigma = V(X) = E[\underbrace{(X-\bar{X})(X-\bar{X})^T}_{\in \mathbb{R}^{n \times n}}]$$

$$\Sigma_{ii} = E((X_i - \bar{X}_i)^2) = V(X_i)$$

$i \neq j$ $\quad \Sigma_{ij} = E((X_i - \bar{X}_i)(X_j - \bar{X}_j)) = Cov(X_i, X_j)$

psd A

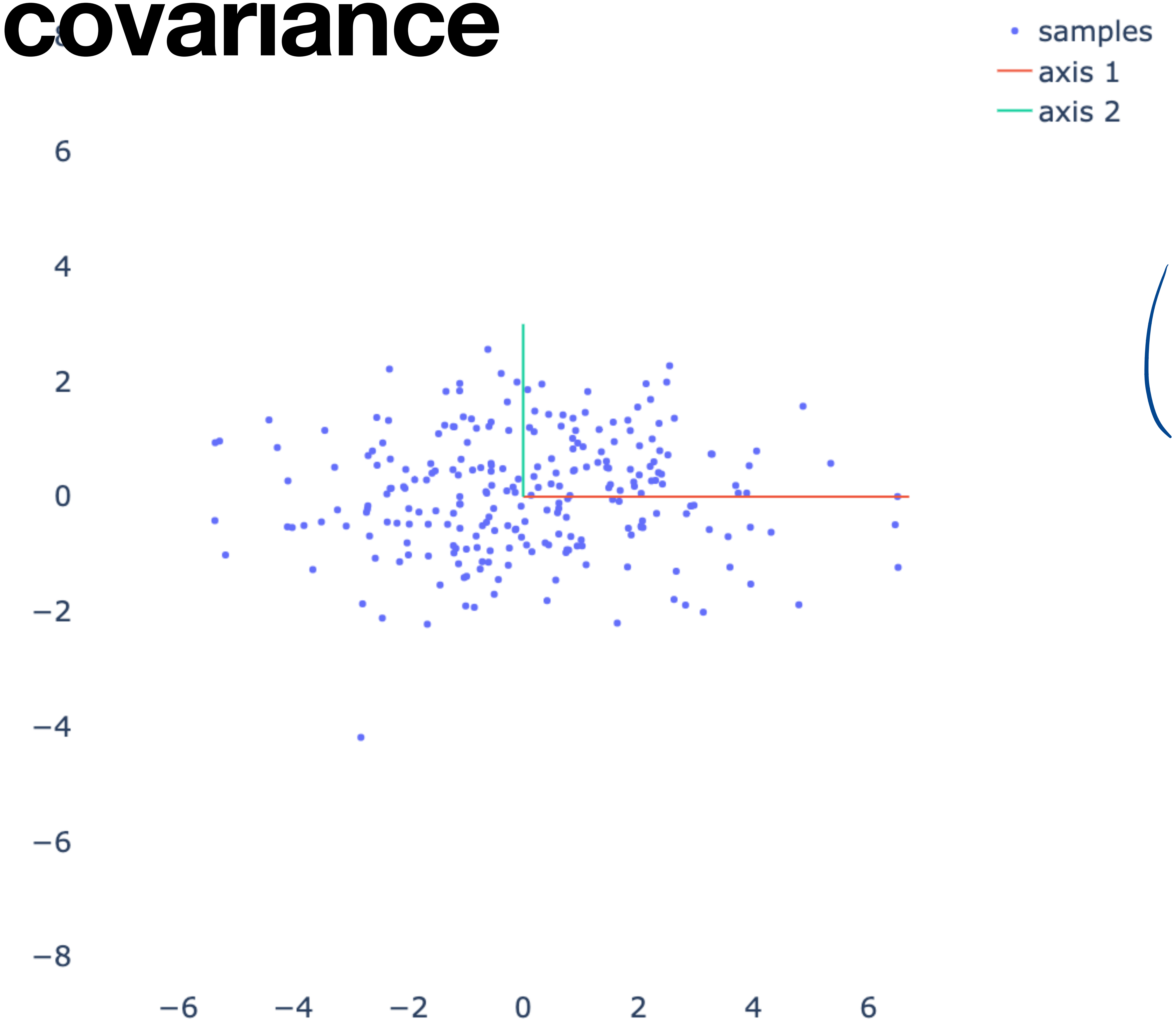
$y^T A y \geq 0$

$$[y^T (x-\bar{x})][(x-\bar{x})^T y]$$

$$= (y^T (x-\bar{x}))^2 \geq 0$$

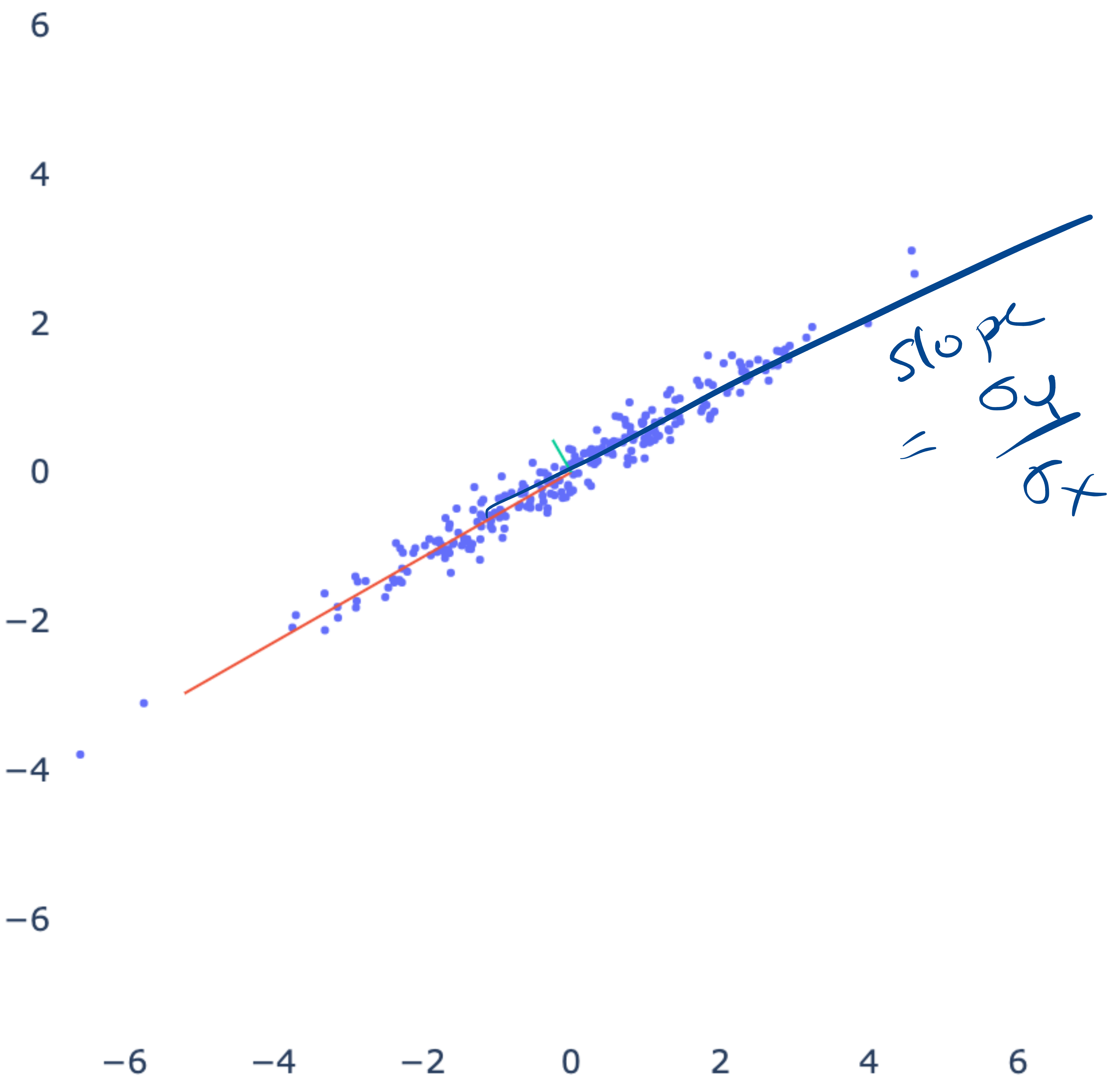
numpy.random.randn(10,1)

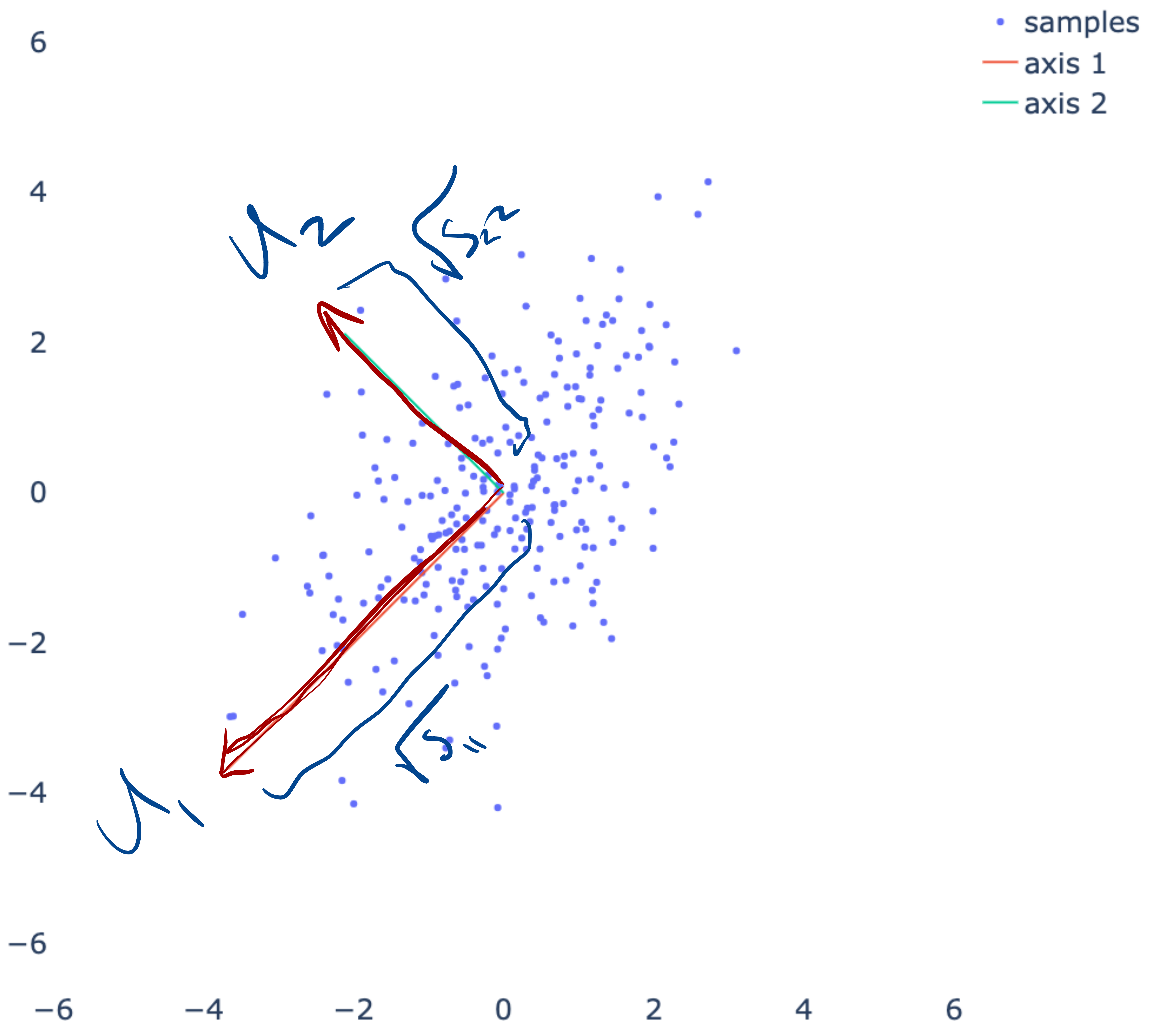
Diagonal covariance



$$\begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

General covariance





$$\Sigma = U S U^T$$

orthonormal

diagonal

$$P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$$

$$\left.\begin{aligned} P(Y \mid X)\,P(X) &= p(X,Y) = P(X \mid Y)\,P(Y) \\ P(Y \mid X) &= P(X \mid Y)\,P(Y) \big/ P(X) \end{aligned}\right\} \text{Bayes' Rule}$$

$$\text{fair} \qquad \overbrace{0.5 \cdot 0.5}^{P(\text{fair} \wedge H)} = 0.25 \,/\, 0.6 = 5/12 \qquad P(\text{fair} \mid H)$$

$$\neg\text{fair} \qquad 0.5 \cdot 0.7 = \underset{P(\neg\text{fair} \wedge H)}{0.35} \,/\, \underbrace{0.6}_{P(H)} = 7/12 \qquad P(\neg\text{fair} \mid H)$$

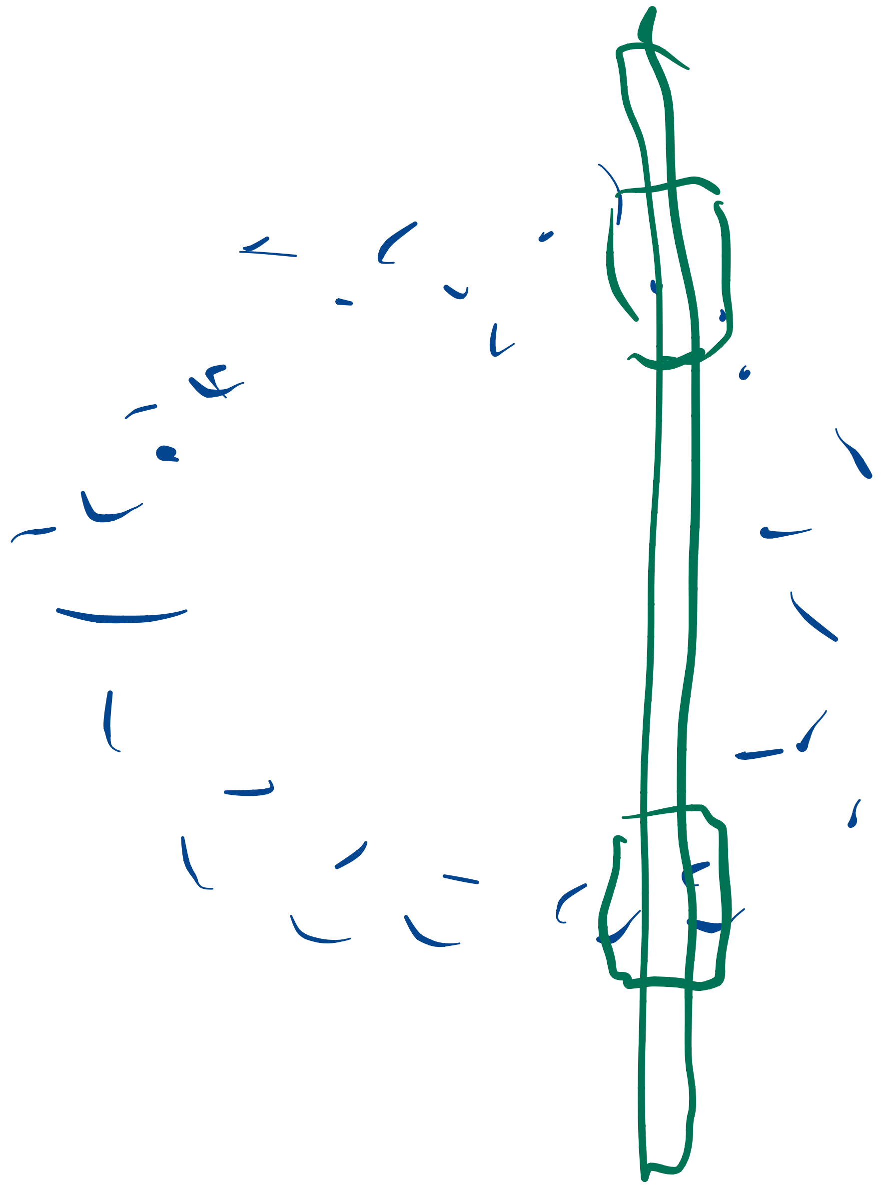
$$P(C \mid \text{flip} = H)$$

$$P(C \mid \text{flip})$$

$$P(x, y, z) = P(x, y \mid z)\, P(z)$$

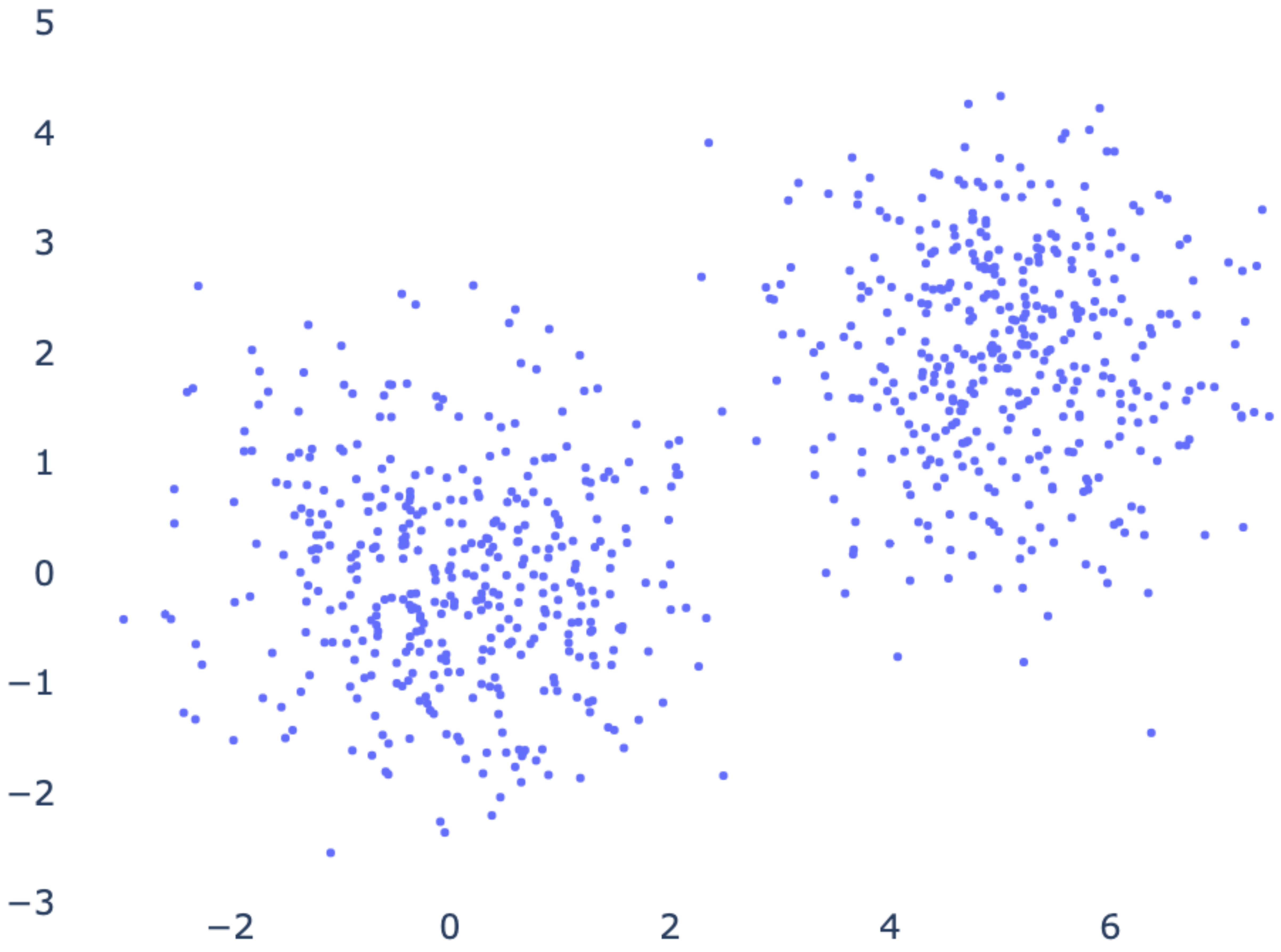
$$= P(x \mid y, z)\, P(y \mid z)\, P(z)$$

$$P(\text{wet} \mid \text{sprinkler}, \text{rain})\, P(\text{sprinkler})\, P(\text{rain})$$



$$P(\text{sprinkler})\, P(\text{rain})$$

Clusters



$$\frac{u}{\|u\|} \qquad \|u\| = (u \cdot u)^{1/2}$$

$$d \frac{u}{(u \cdot u)^{1/2}} = \frac{1}{\sqrt{u \cdot u}} du + \left(-\cancel{\frac{1}{2}}\right) \underbrace{(u \cdot u)^{-3/2}}_{\|u\|^{-3}} (\cancel{2} u \cdot du)\, u$$

$$\frac{1}{\|u\|} \left(\frac{u}{\|u\|} \cdot du\right) \frac{u}{\|u\|}$$