

Math Foundations for ML

10-606

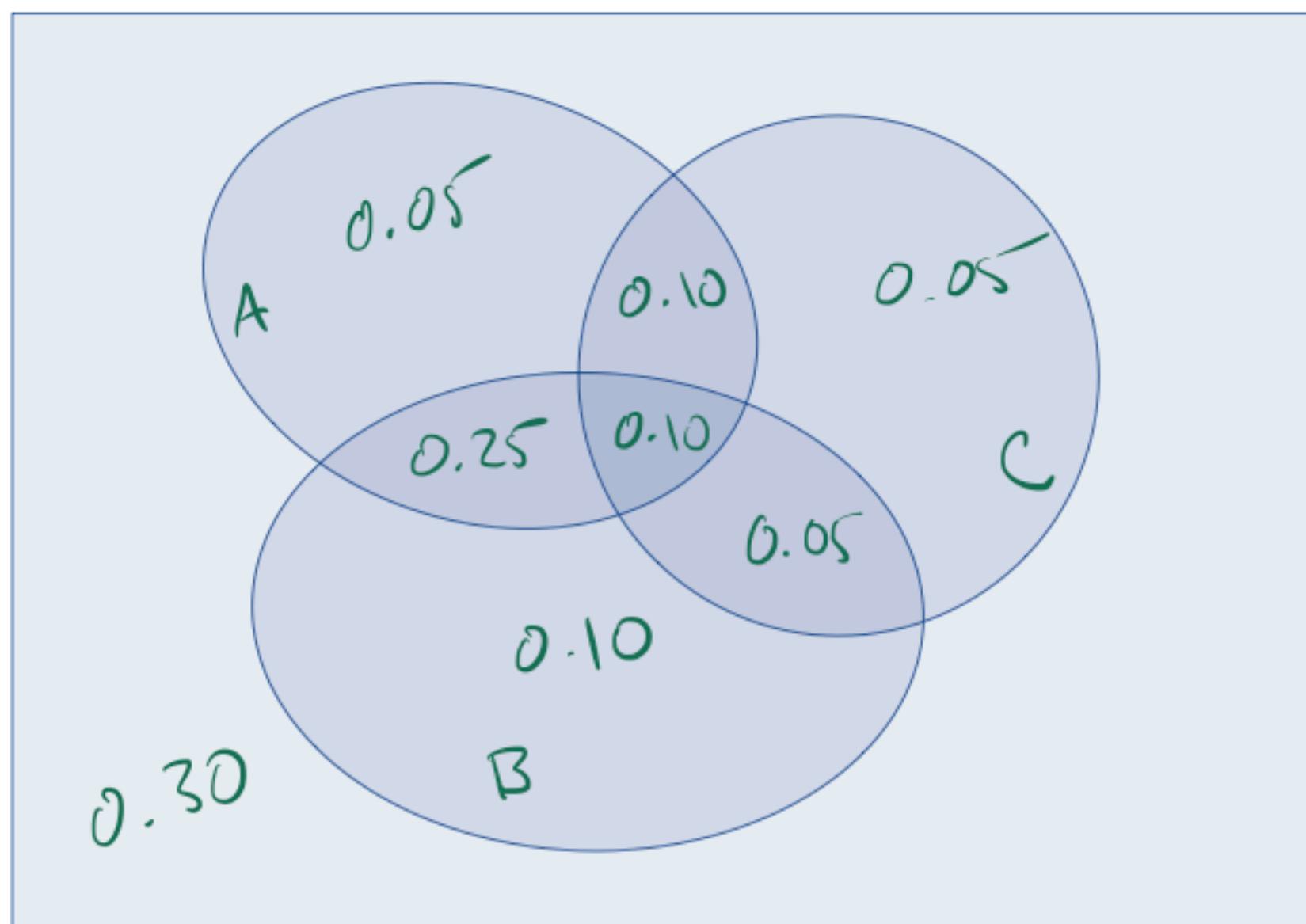
Geoff Gordon

Notes and reminders

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

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

Marginal, conditional

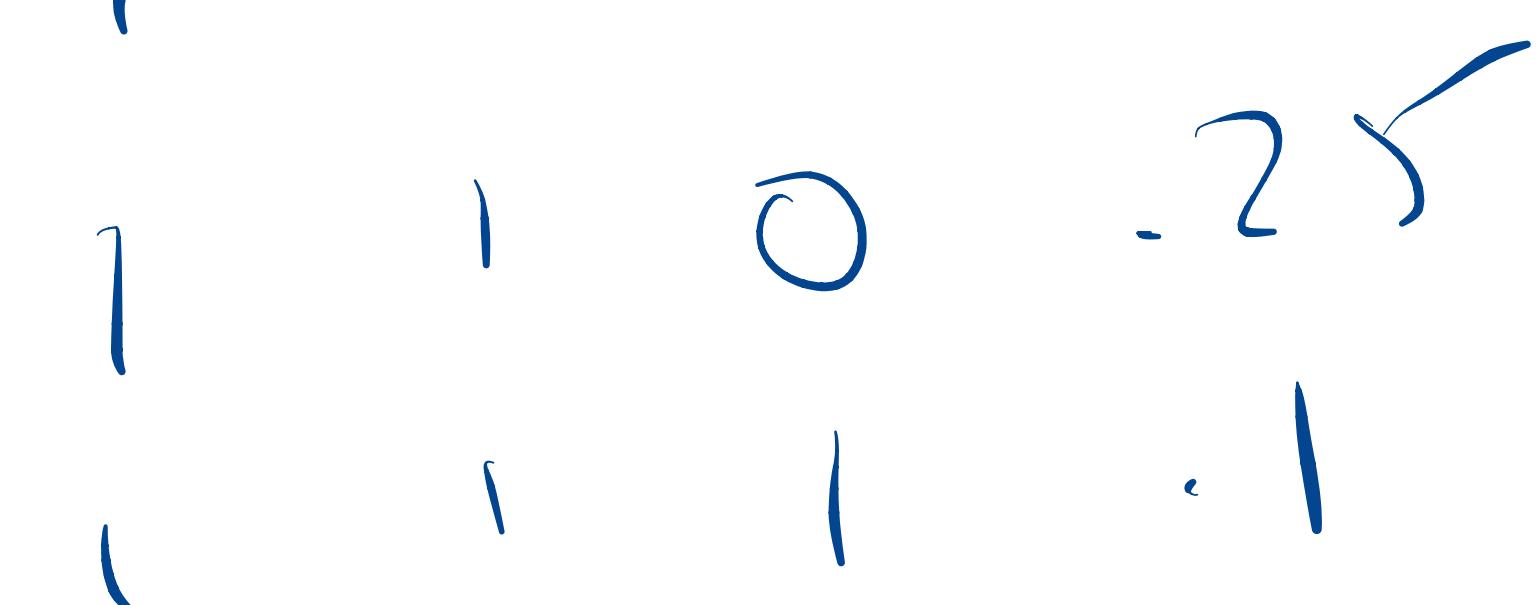
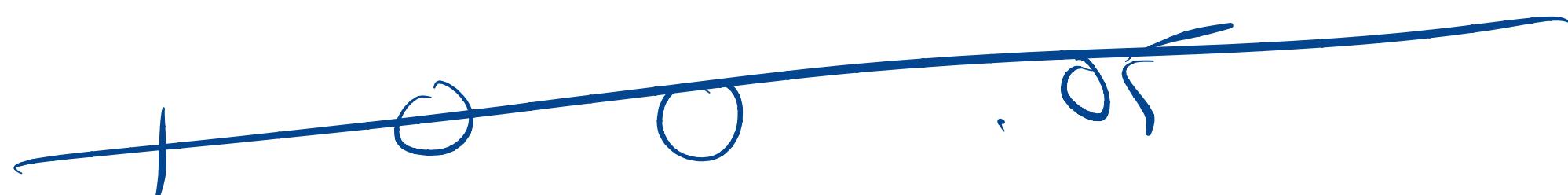
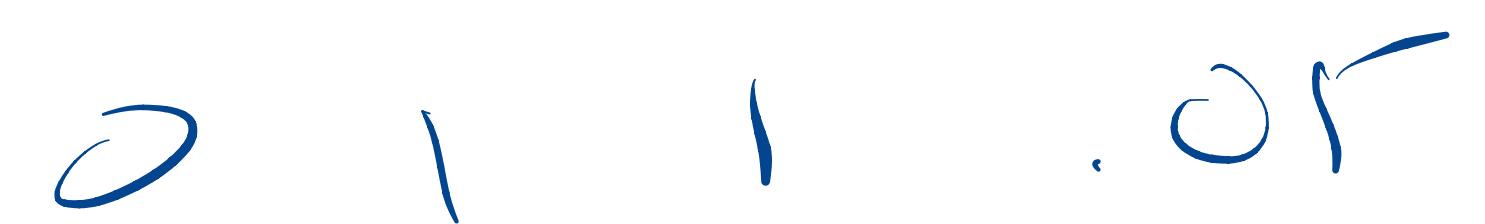


| A | B | C | P | |
|---|---|---|------|--|
| 0 | 0 | 0 | 0.30 | |
| 0 | 0 | 1 | 0.05 | |
| 0 | 1 | 0 | 0.10 | |
| 0 | 1 | 1 | 0.05 | |
| 1 | 0 | 0 | 0.05 | |
| 1 | 0 | 1 | 0.10 | |
| 1 | 1 | 0 | 0.25 | |
| 1 | 1 | 1 | 0.10 | |

Handwritten annotations in blue ink show the marginal probabilities for columns A, B, and C:

- Column A: 0.35 (written vertically next to the first two rows)
- Column B: 0.15 (written vertically next to the third and fourth rows)
- Column C: 0.15 (written vertically next to the fifth and sixth rows)
- Column P: 0.35 (written vertically next to the last two rows)

A B C



given $B = 1$

$P(A, C | B = 1)$

A C

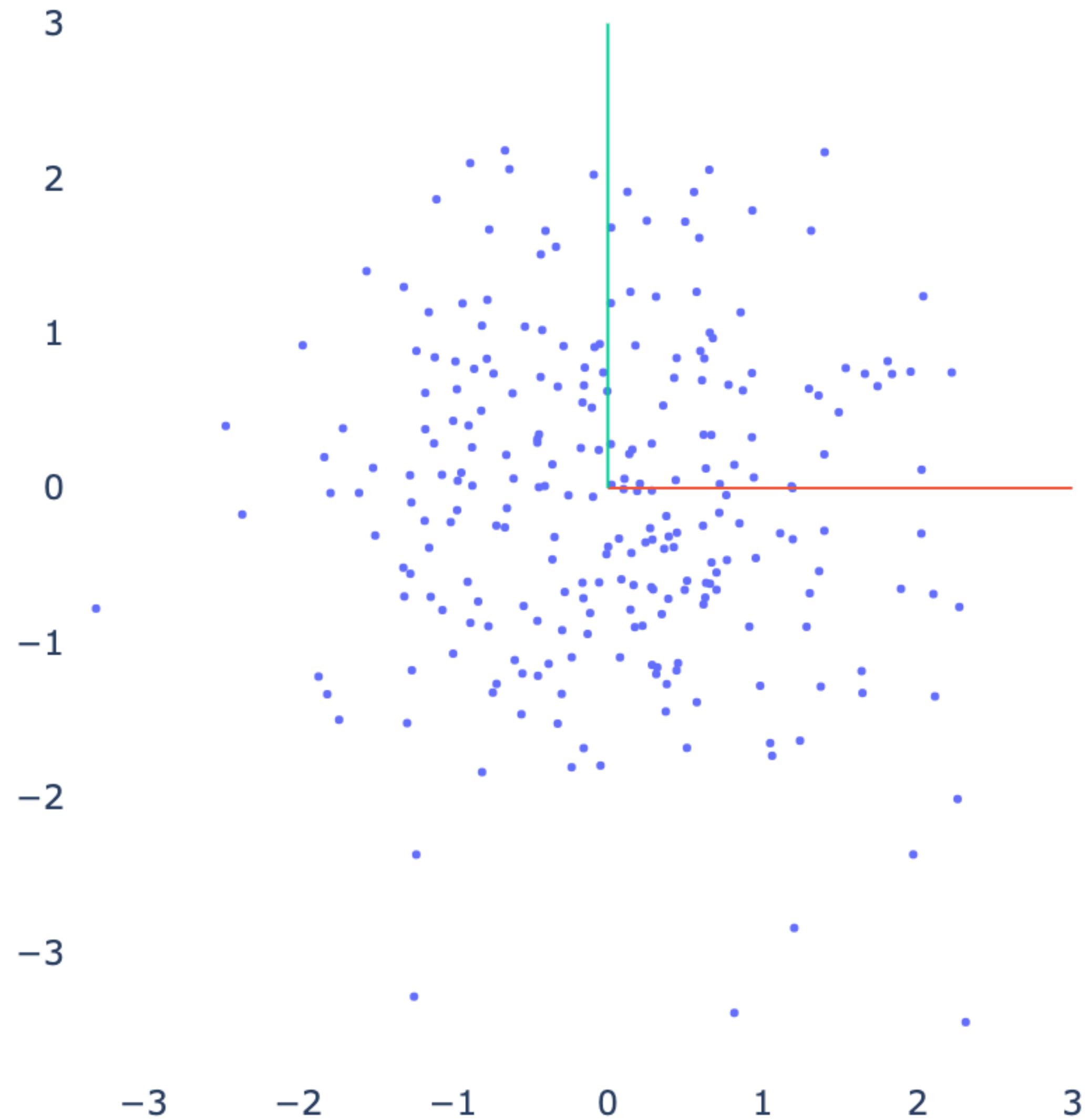
0 0 $\cdot 1/5 = .2$

0 1 $\cdot 0/5 = .1$

1 0 $.25/5 = .5$

1 1 $\cdot 1/5 = .2$

Mean



- samples
- axis 1
- axis 2

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

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

$$\begin{aligned} E(aX + bY + c) \\ = aE(X) + bE(Y) + c \end{aligned}$$

$$V(x) = E[(x - \bar{x})^2]$$

$\hat{\epsilon}_x$

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

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

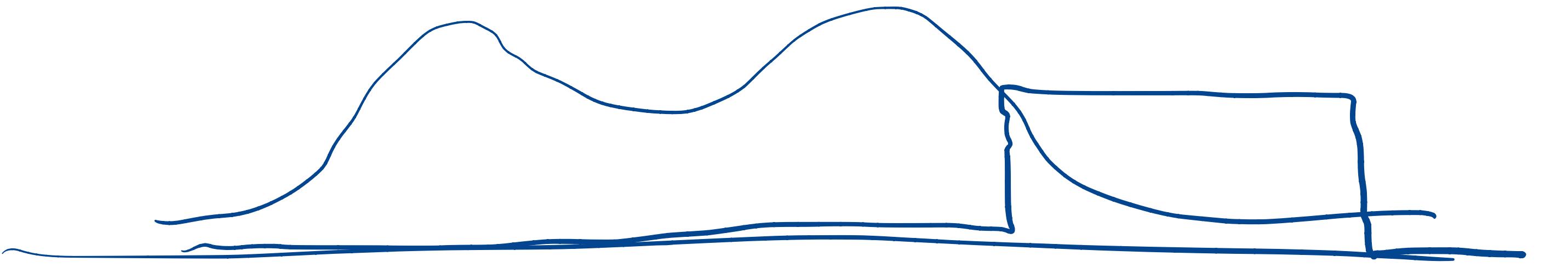
$$\sigma(x) = \sqrt{V(x)}$$

height is 160 cm \pm 20 cm

$$\begin{aligned} & E(x^2 - 2x\bar{x} - \bar{x}^2) \\ &= E(x^2) - \underbrace{2\bar{x}E(x)}_{2\bar{x}^2} - \bar{x}^2 \\ &= E(x^2) - \bar{x}^2 \end{aligned}$$

$$\begin{aligned} & E(|x - \bar{x}|) \\ & \text{MAD} \end{aligned}$$

Moment: $E(f(x))$



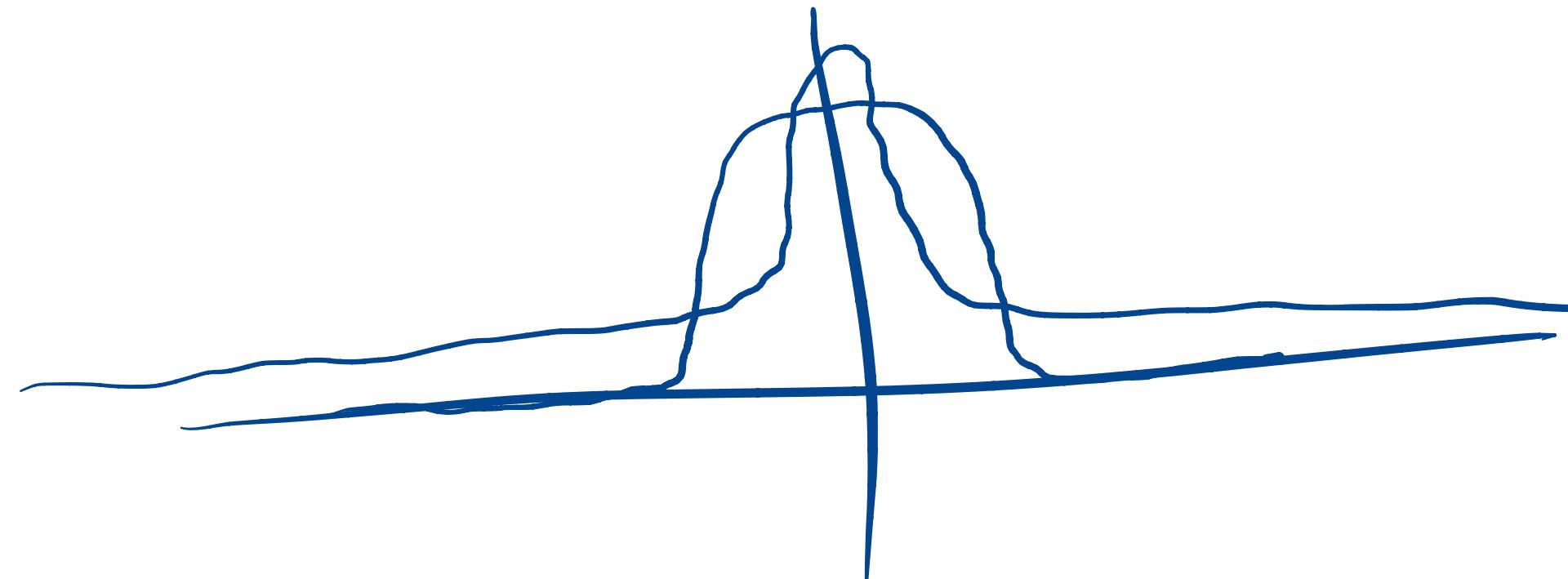
$f = \text{identity} \rightarrow \text{mean}$

$$E(f) = P(\text{this interval})$$

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

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

$f(x) = x^k$



$x, y \in \mathbb{R}$

$$\text{Cov}(x, y) = E[(x - \bar{x})(y - \bar{y})]$$

$$\text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sigma(x)\sigma(y)} = E\left(\frac{x - \bar{x}}{\sigma_x} \cdot \frac{y - \bar{y}}{\sigma_y}\right) \in [-1, 1]$$

$$\text{Var}\left(\frac{x - \bar{x}}{\sigma_x}\right) = 1$$

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

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

$$\Sigma = V(x) = E \left[\underbrace{(x - \bar{x})(x - \bar{x})^\top}_{\in \mathbb{R}^{n \times n}} \right]$$

$\text{tr}(V(x))$
avg squared dist
from mean

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

$$\text{if } i \neq j \quad \Sigma_{ij} = E((x_i - \bar{x}_i)(x_j - \bar{x}_j)) = \text{Cov}(x_i, x_j)$$

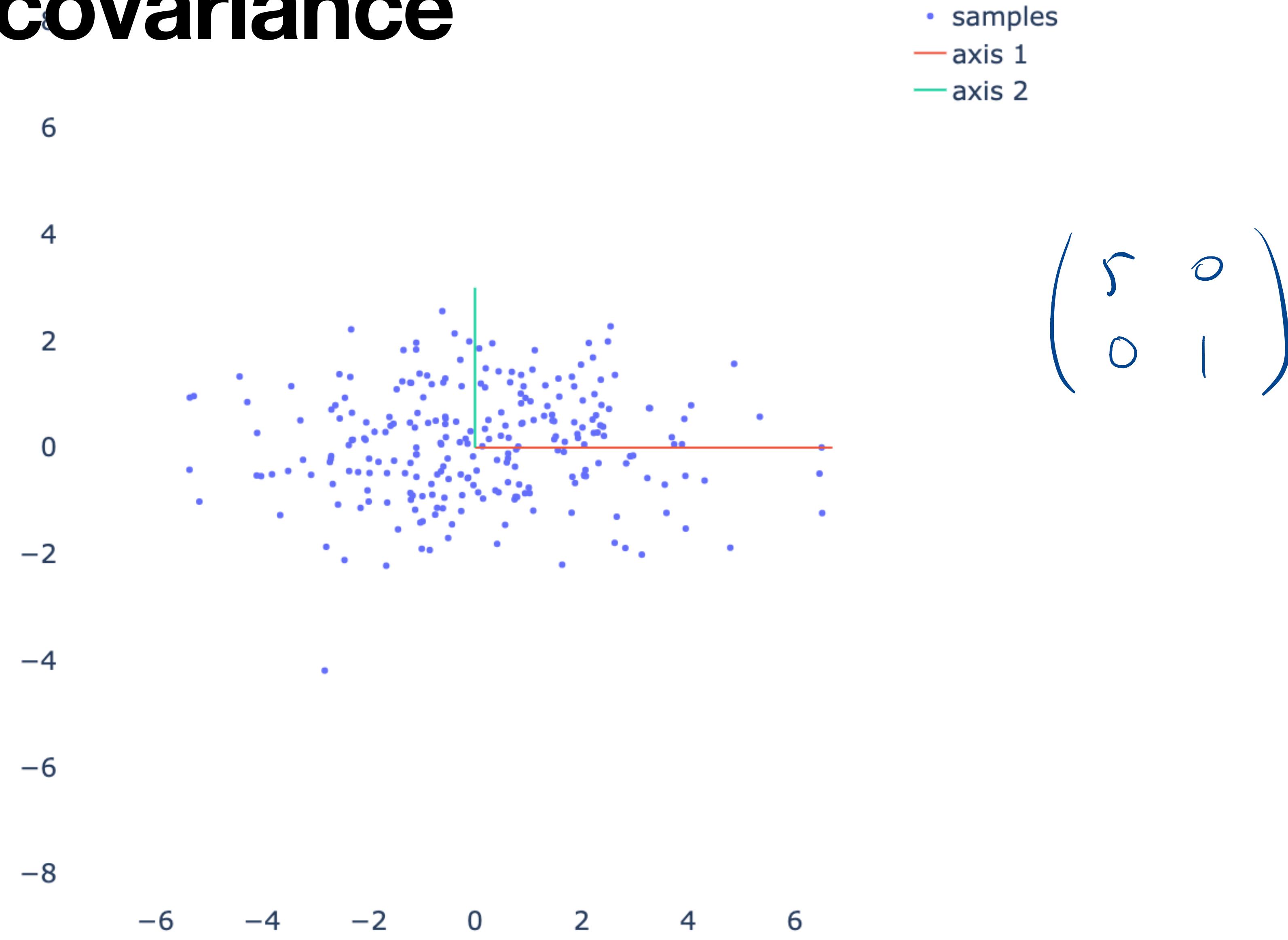
psd A

$$y^\top A y \geq 0$$

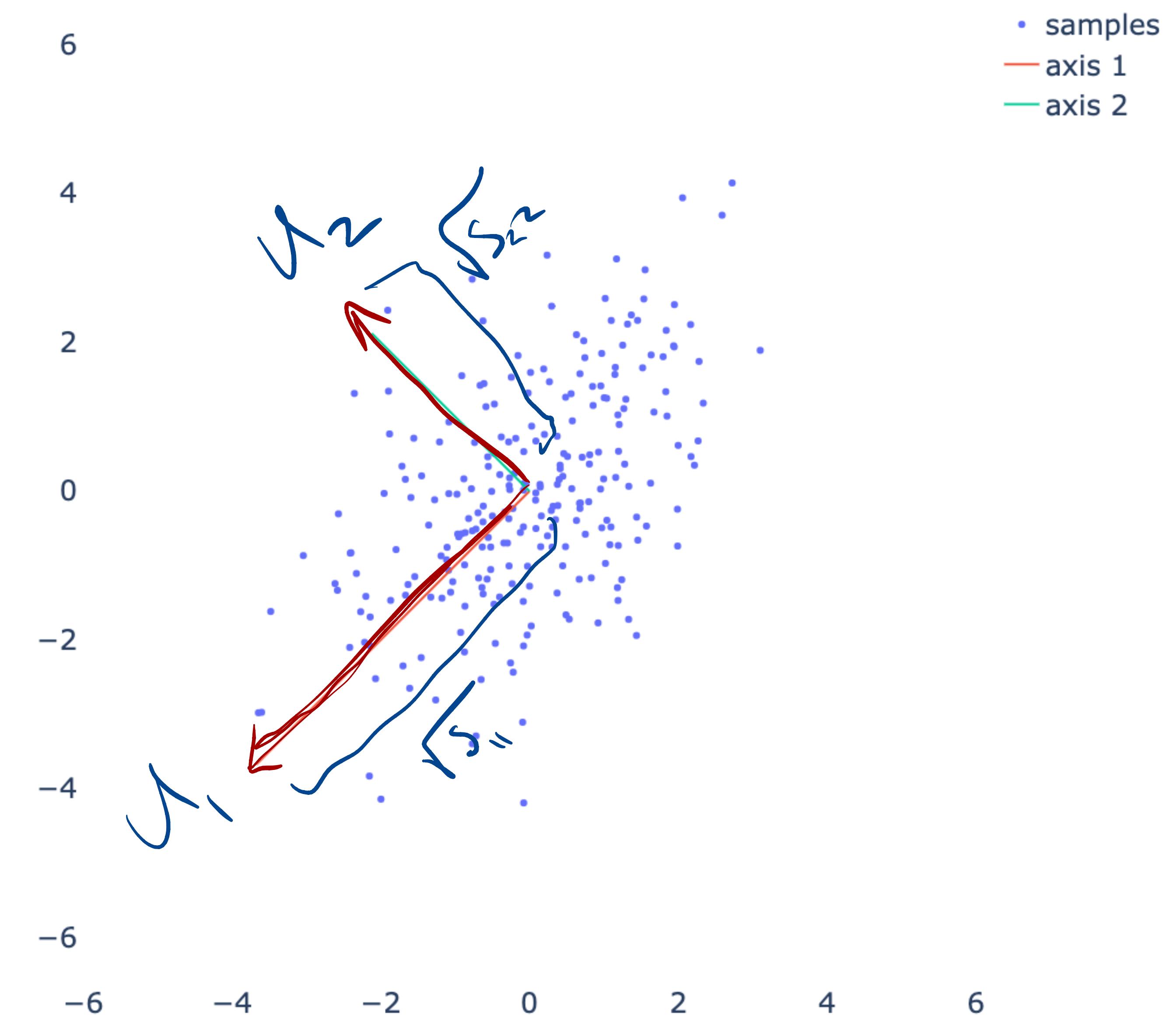
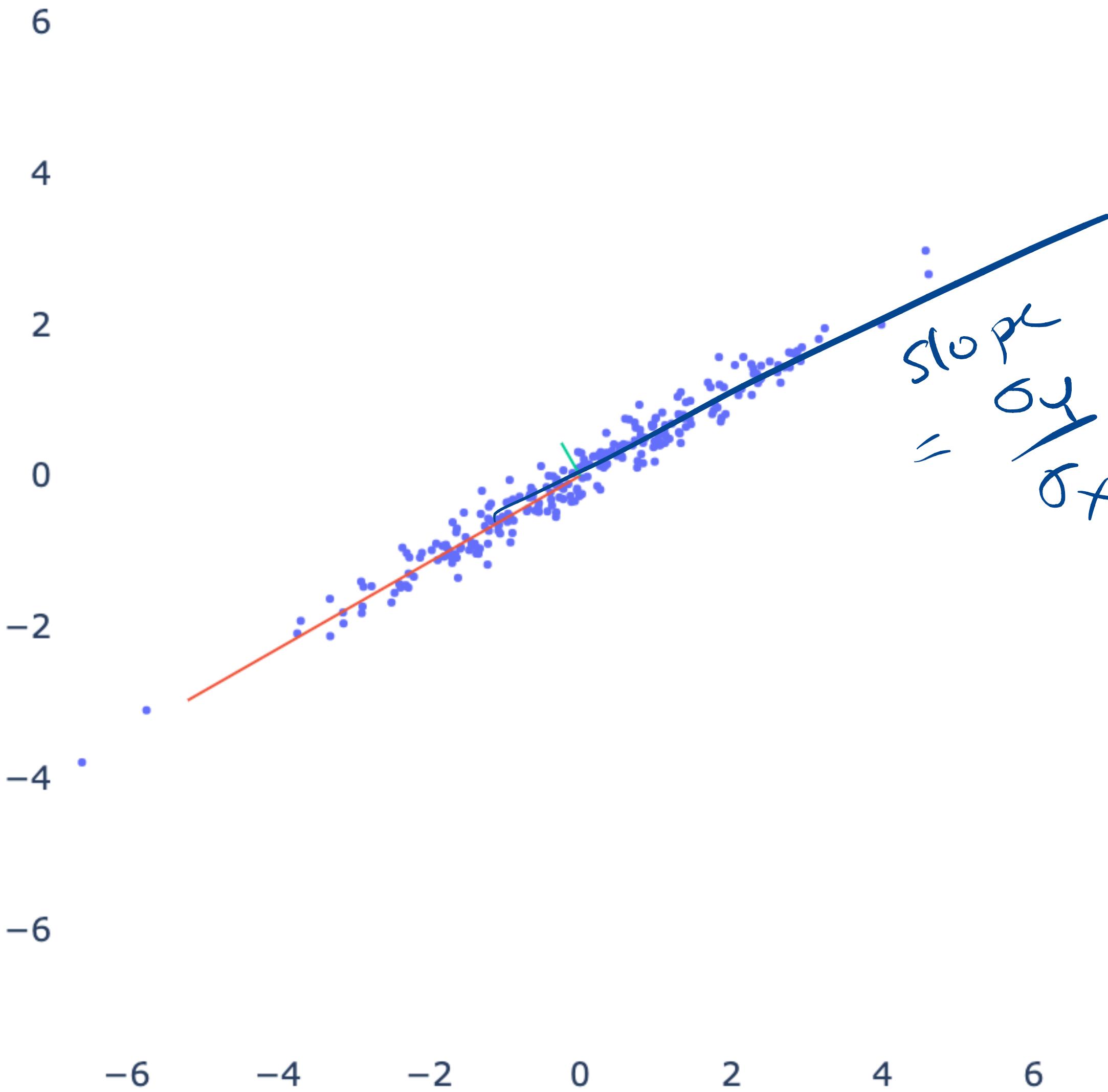
$$\begin{aligned} & y^\top (x - \bar{x})(x - \bar{x})^\top y \\ &= (y^\top (x - \bar{x}))^2 \geq 0 \end{aligned}$$

`numpy.random.rand(10, 1)`

Diagonal covariance



General covariance



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

↗ ↑ diagonal
 orthonormal

$$P(x|y) = \frac{P(x,y)}{P(y)}$$

$$P(y|x)p(x) = p(x,y) = P(x|y)p(y)$$

$$P(y|x) = P(x|y)p(y) / P(x)$$

} Bayes' Rule

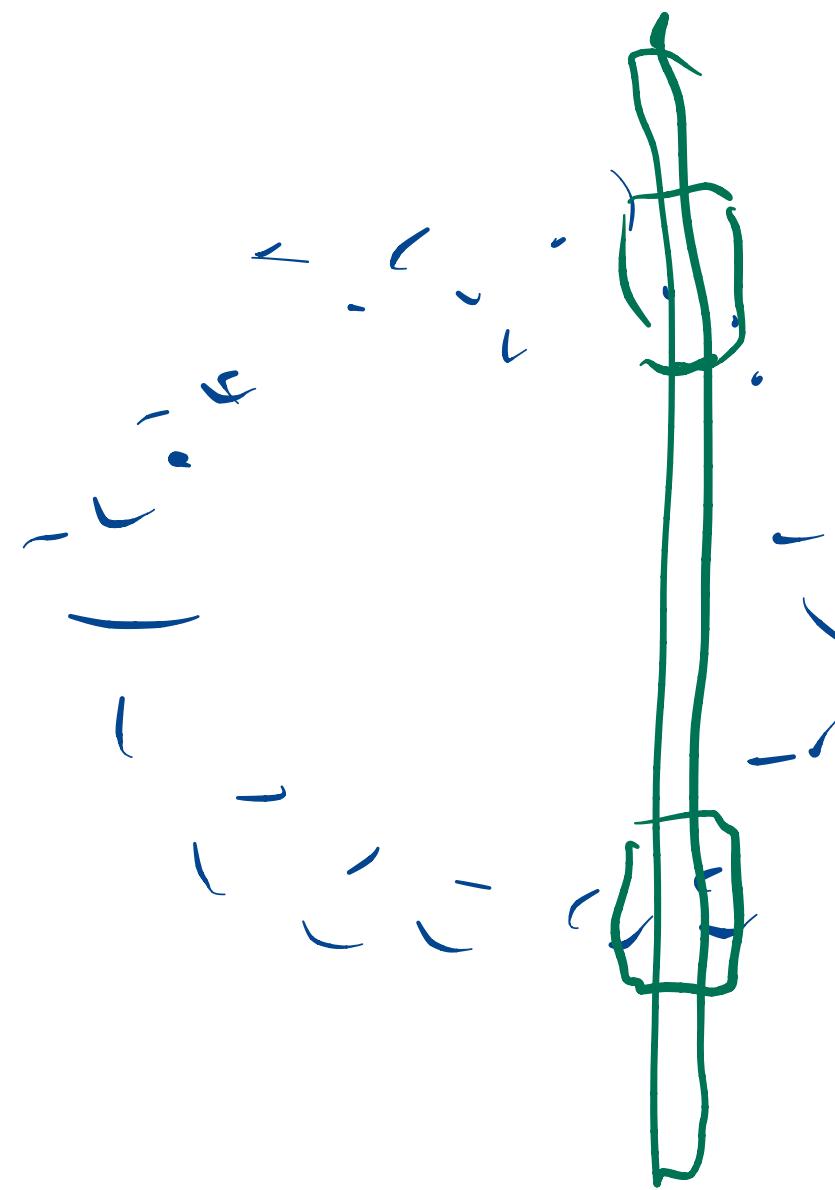
$$\begin{array}{l} \text{fair} \\ \text{\hspace{1cm}} P(\text{fair} \wedge H) \\ 0.5 \cdot 0.5 = 0.25 / 0.6 = 5/12 \\ \text{\hspace{1cm}} P(\text{fair} | H) \\ \\ \sim \text{fair} \\ \text{\hspace{1cm}} P(\sim \text{fair} \wedge H) \\ 0.5 \cdot 0.7 = 0.35 / 0.6 = 7/12 \\ \text{\hspace{1cm}} P(\sim \text{fair} | H) \end{array}$$

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

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

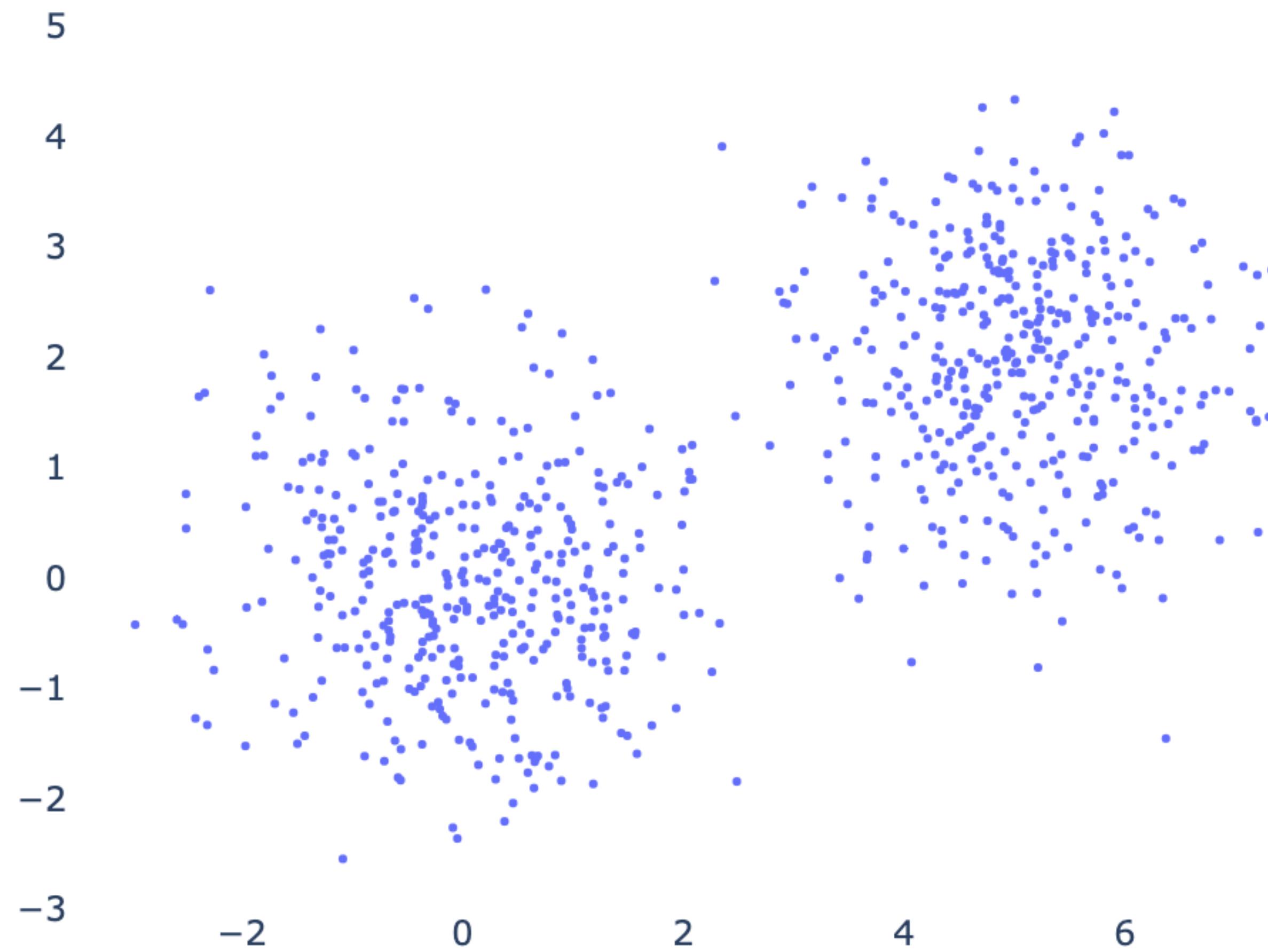
$$P(x, y, z) = P(x, y | z) \underbrace{P(z)}_{= P(x|y, z) P(y|z) P(z)}$$

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



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

Clusters



$$\frac{u}{\|u\|}$$

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

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