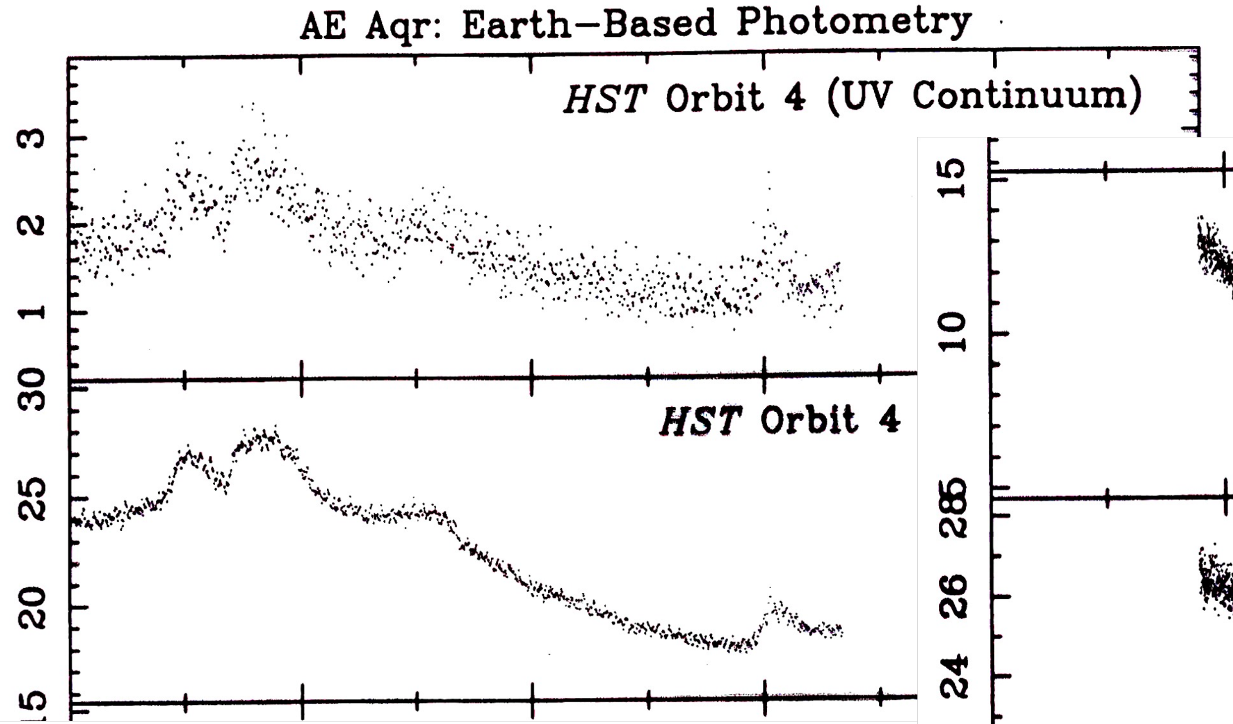
ADA02 - 9am Tue 13 Sep 2022

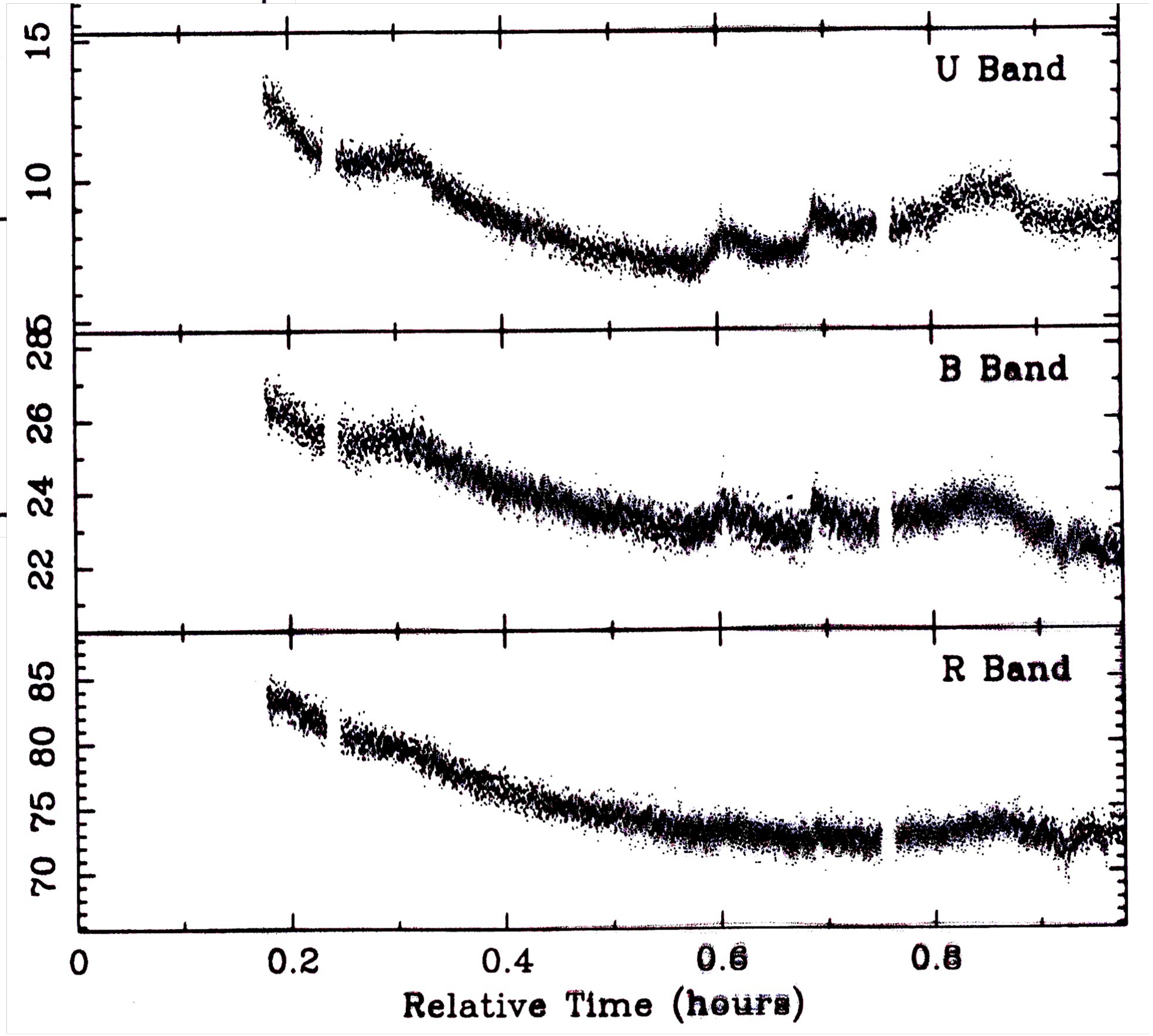
Eyeballing sigma (5-sigma rule)

Joint Probability Distributions
Independence vs Correlation

Algebra of Random Variables:
Linear Transformations
Covariance Matrix
Correlation Coefficient / Matrix

Eyeballing Sigma





How well can you estimate the noise standard deviation, σ, just by looking at a plot of the data ?

Eyeballing Sigma

1. If lots of data points: The 5-sigma rule:

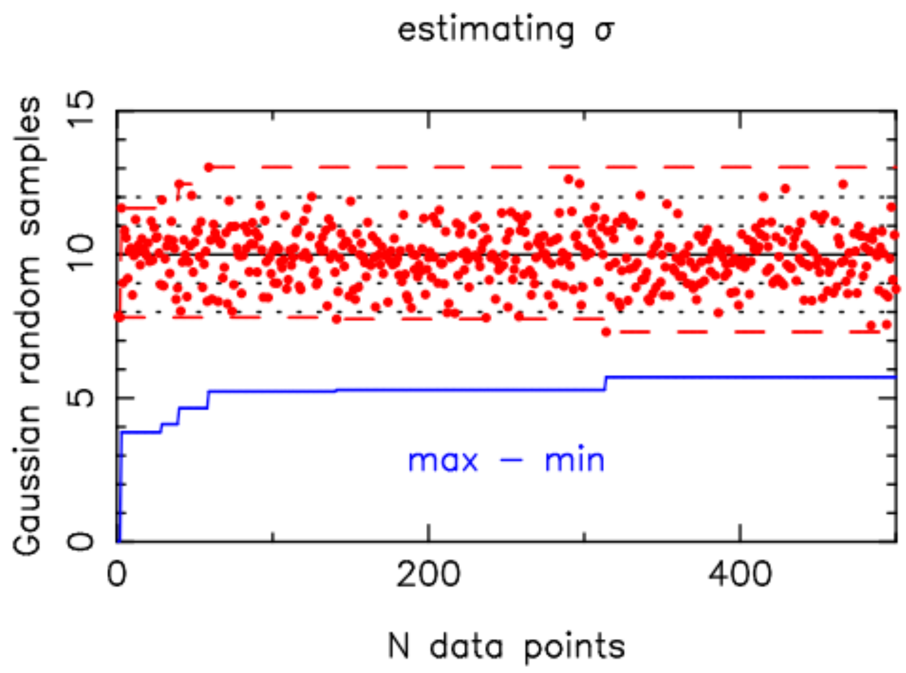
Estimate “by eye” the range (max-min) of the data (~100 data points).

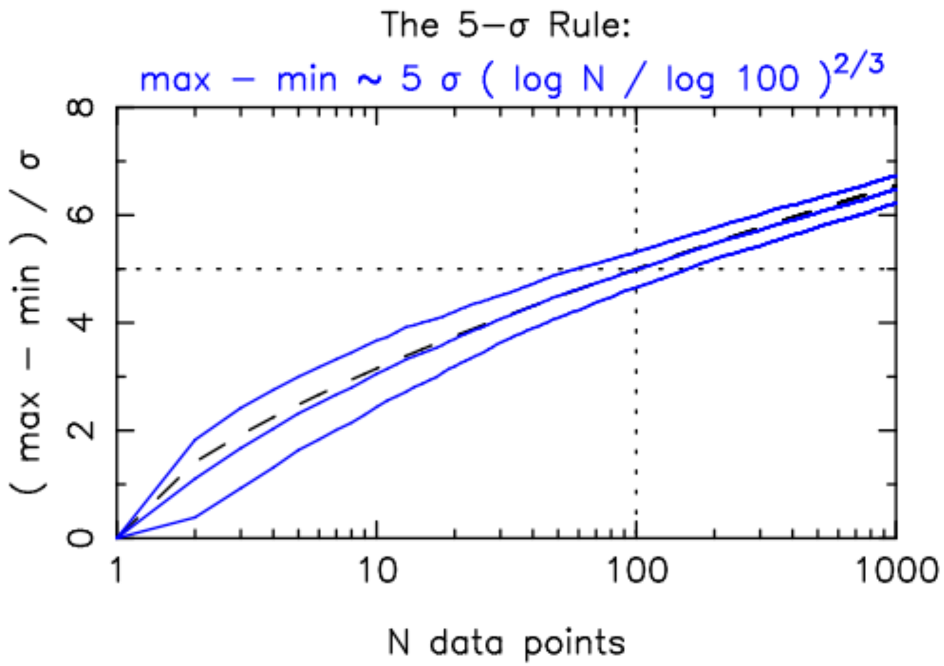
That range (max-min) is about 5-sigma

Usually good to 20% by eye.

2. If only a few data points:

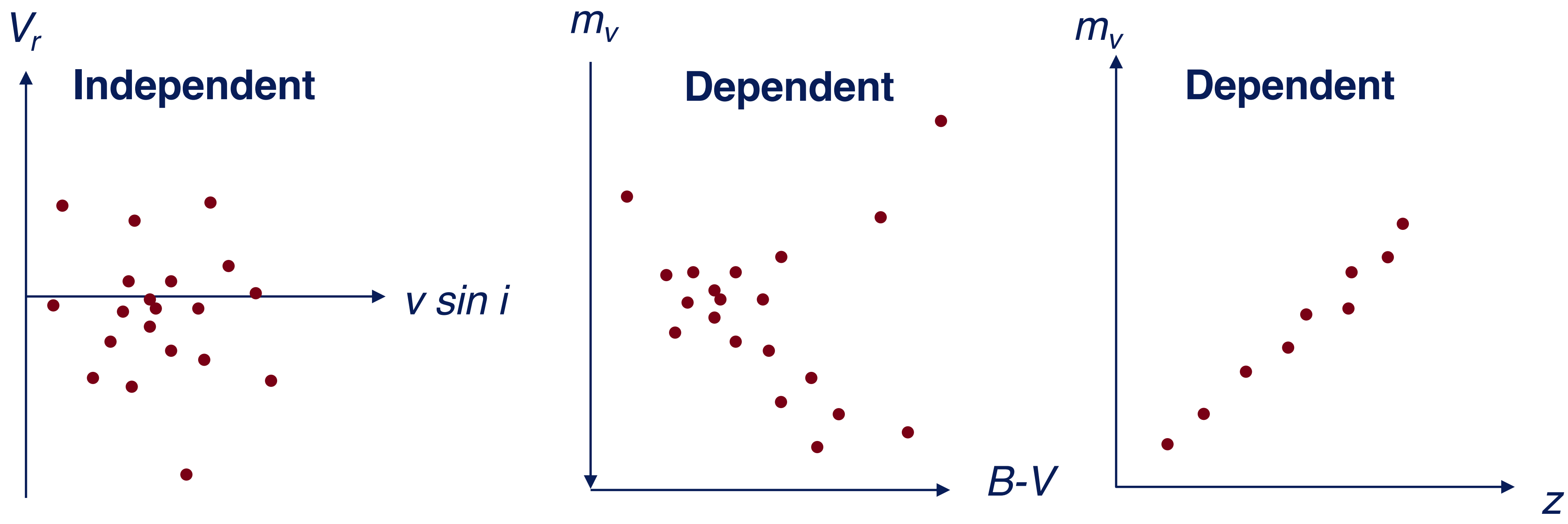
~2/3 of the data points should be inside +/- 1 sigma





Multivariate Distributions

- Suppose we measure 2 (or more) different properties
  - e.g. rotational and radial velocities of stars in a cluster
  - colours and magnitudes of stars in a cluster
  - redshifts and peak apparent magnitudes of distant supernovae



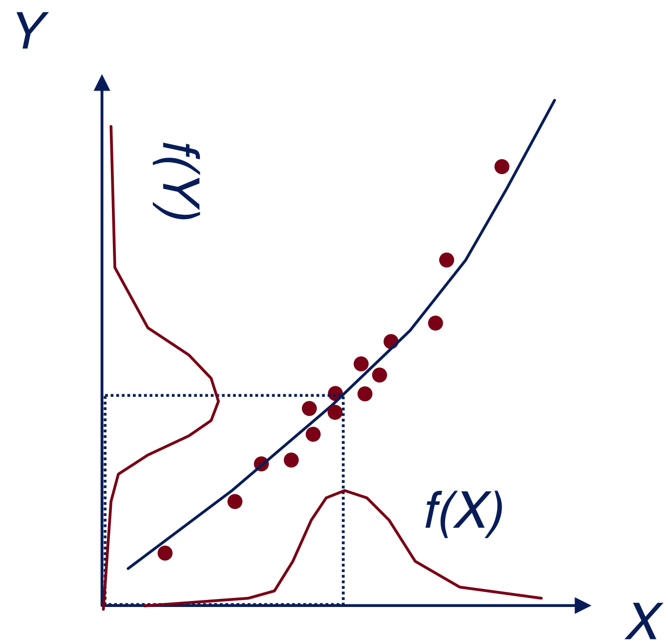
- Does knowing the value of one random variable *X* inform you about the other?

Joint Probability Distribution *f(X, Y)*

- *X* and *Y* are two random variables.
- Their **joint probability distribution** is *f(X, Y)*
- Normalisation: $$\iint f(X,Y)\, dX\, dY = 1$$
- Projection gives *f(X)*, *f(Y)* :

$$f(Y) = \int f(X,Y)\, dX$$

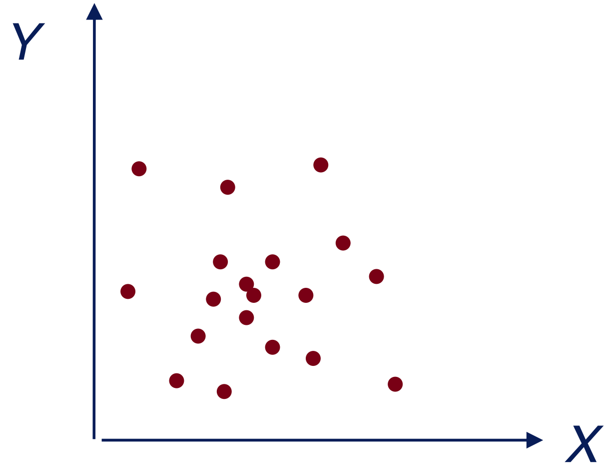
$$f(X) = \int f(X,Y)\, dY$$



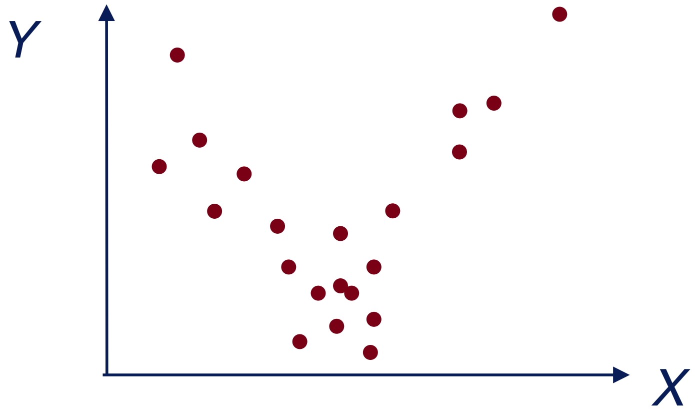
Independence vs Correlation

- **Independent variables:**
  - knowing *X* does not inform about *Y*
  - Definition:

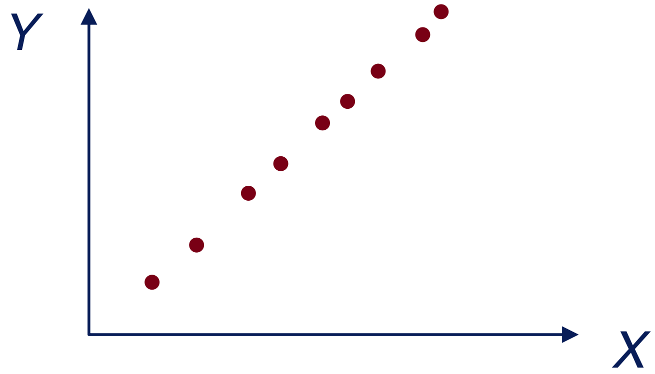
$$f(X,Y) = f(X)\, f(Y)$$



- **Partially correlated:**
  - knowing *X* tells you something about *Y*



- **Perfect correlation:**
  - *X* determines *Y*



The Algebra of Random Variables

Ordinary numbers are "sharp".

$$1 + 1 = 2$$

Random variables are "fuzzy" numbers.

$(\mu \pm \sigma)$ is a shorthand notation giving the mean μ and standard deviation σ of a random variable.

$$(1 \pm 1) + (1 \pm 2) = (? \pm ?)$$

How do the mean and variance change when we add or subtract or multiply fuzzy numbers?

How do the higher moments change?

Linear Transformations: Scaling

Constants: $\langle a \rangle \;=\; ?$ $\qquad \mathrm{Var}(a) \;=\; ?$

Scaling a random variable, *X*, by a constant, *a* :

- **Mean:**

$$\langle a\,X \rangle = a \langle X \rangle$$

$$\langle a\,X \rangle = \int a\,X\; f(X)\;\; dX$$
$$= a \int X\; f(X)\, dX = a \langle X \rangle$$

- **Variance:**

$$\mathrm{Var}(a\,X) = a^2\,\mathrm{Var}(X)$$
$$\sigma\left(a\,X\right) = |\,a\,|\,\sigma\left(X\right)$$

$$\mathrm{Var}(a\,X) = \left\langle [a\,X - \langle a\,X \rangle]^2 \right\rangle$$
$$= \left\langle [a\,X - a\langle X \rangle]^2 \right\rangle$$
$$= \left\langle a^2 [X - \langle X \rangle]^2 \right\rangle$$
$$= a^2\,\mathrm{Var}(X)$$

"Stretch the paper" by a factor ***a.***

Location μ and width σ then increase by factor ***a.***

Linear Transformations: Addition

- Adding two random variables *X* and *Y* :

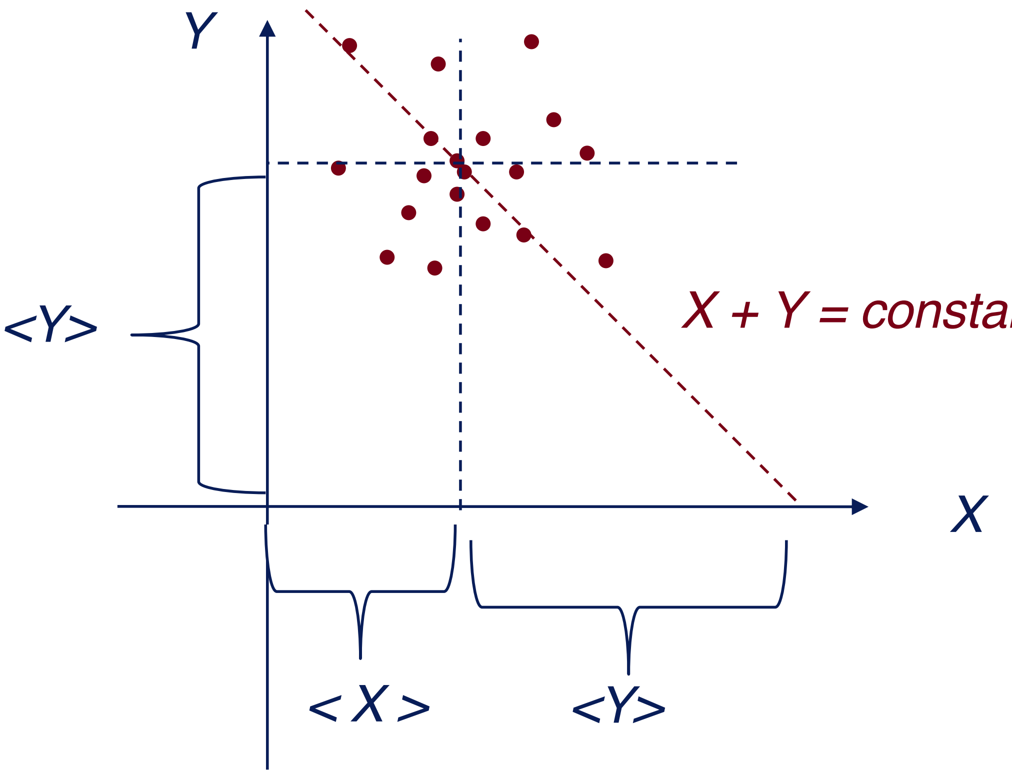
$$
\begin{aligned}
\langle X+Y \rangle &\equiv \iint (X+Y) f(X,Y)\ dX\, dY \\
&= \iint X\ f(X,Y)\, dX\, dY + \iint Y\ f(X,Y)\ dX\, dY \\
&= \int X \left[\int f(X,Y) dY \right] dX + \int Y \left[\int f(X,Y) dX \right] dY \\
&= \int X\ f(X)\ dX + \int Y\ f(Y)\ dY \\
&\equiv \quad \langle X \rangle \quad + \quad \langle Y \rangle
\end{aligned}
$$

- True for **any** joint PDF!

$$\boxed{\langle X+Y \rangle = \langle X \rangle + \langle Y \rangle}$$

Why it works...

- **Centre of mass** is a well-defined position.



$$\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle$$

Variance and Co-variance

- Variance of $X+Y$ depends on how X and Y co-vary:

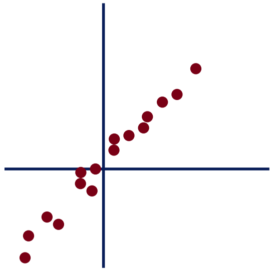
$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(X,Y) \equiv \left\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \right\rangle$$

$$\begin{aligned}
\mathrm{Var}(X+Y) &\equiv \left\langle [X + Y - \langle X + Y \rangle]^2 \right\rangle \\
&= \left\langle [X + Y - \langle X \rangle - \langle Y \rangle]^2 \right\rangle \\
&= \left\langle [(X - \langle X \rangle) + (Y - \langle Y \rangle)]^2 \right\rangle \\
&= \left\langle (X - \langle X \rangle)^2 + (Y - \langle Y \rangle)^2 + 2\,(X - \langle X \rangle)(Y - \langle Y \rangle) \right\rangle \\
&= \left\langle (X - \langle X \rangle)^2 \right\rangle + \left\langle (Y - \langle Y \rangle)^2 \right\rangle + 2\left\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \right\rangle \\
&= \quad \mathrm{Var}(X) \quad + \quad \mathrm{Var}(Y) \quad + \quad 2\,\mathrm{Cov}(X,Y)
\end{aligned}$$

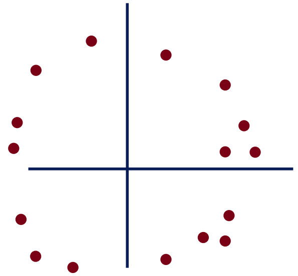
Co-variance vs Independence

- **Cov > 0**

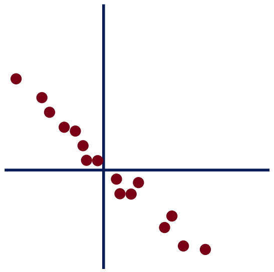


- **Cov = ?**

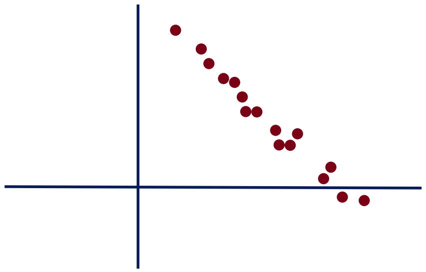
Independent?



- **Cov < 0**



- **Cov = ?**

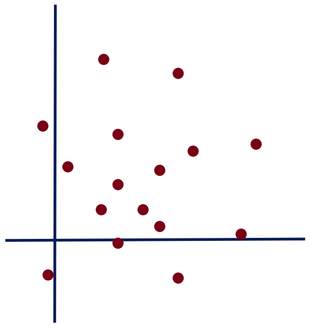


Practice !

$X = 1 \pm 1$

$Y = 1 \pm 1$

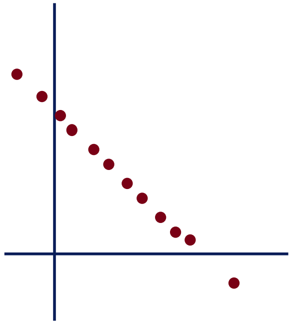
$X + Y = ? \pm ?$



$X = 1 \pm 1$

$Y = 2 - X$

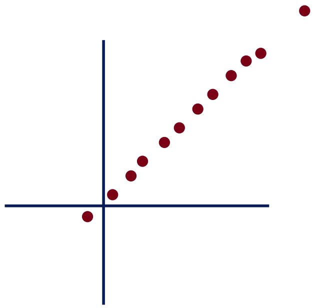
$X + Y = ? \pm ?$



$X = 1 \pm 1$

$Y = X$

$X + Y = ? \pm ?$



$X = 1 \pm 1 \qquad Y = 2 \pm 1 \qquad \mathrm{Cov}[X,Y] = 0 \qquad a = 2 \qquad b = 1$

$Z = aX + bY \qquad \langle Z \rangle = ? \qquad \mathrm{Var}(Z) = ?$

Linear Transformations

- Scale and add any number of random variables:

$$\left\langle \sum_i a_i X_i \right\rangle = \sum_i a_i \langle X_i \rangle \qquad \mathrm{Var}\left[\sum_i a_i X_i\right] = \sum_{i,j} a_i a_j \mathrm{Cov}(X_i, X_j)$$

Or, in terms of the (symmetric) **Co-variance Matrix**:

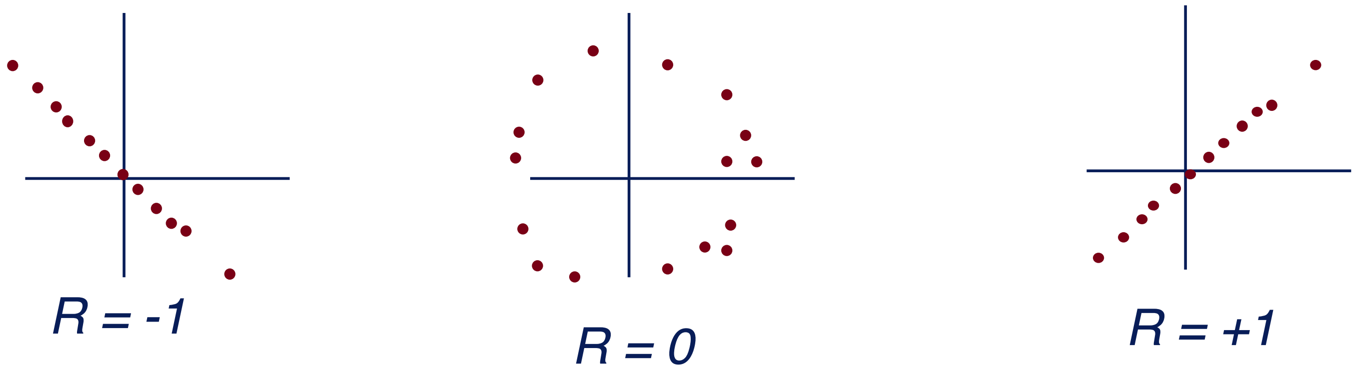
$$\mathrm{Var}\left[\sum_i a_i X_i\right] = \sum_{i,j} a_i C_{ij} a_j$$

$$\mathrm{Var}\left[\begin{pmatrix} a_i & \dots & a_N \end{pmatrix}\begin{pmatrix} X_1 \\ \dots \\ X_N \end{pmatrix}\right] = \begin{pmatrix} a_i & \dots & a_N \end{pmatrix}\begin{pmatrix} C_{11} & \dots & C_{1N} \\ \dots & \dots & \dots \\ C_{N1} & \dots & C_{NN} \end{pmatrix}\begin{pmatrix} a_1 \\ \dots \\ a_N \end{pmatrix}$$

Correlation Coefficient *R(X, Y)*

- **Correlation coefficient:**

$$R(X,Y) \equiv \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$



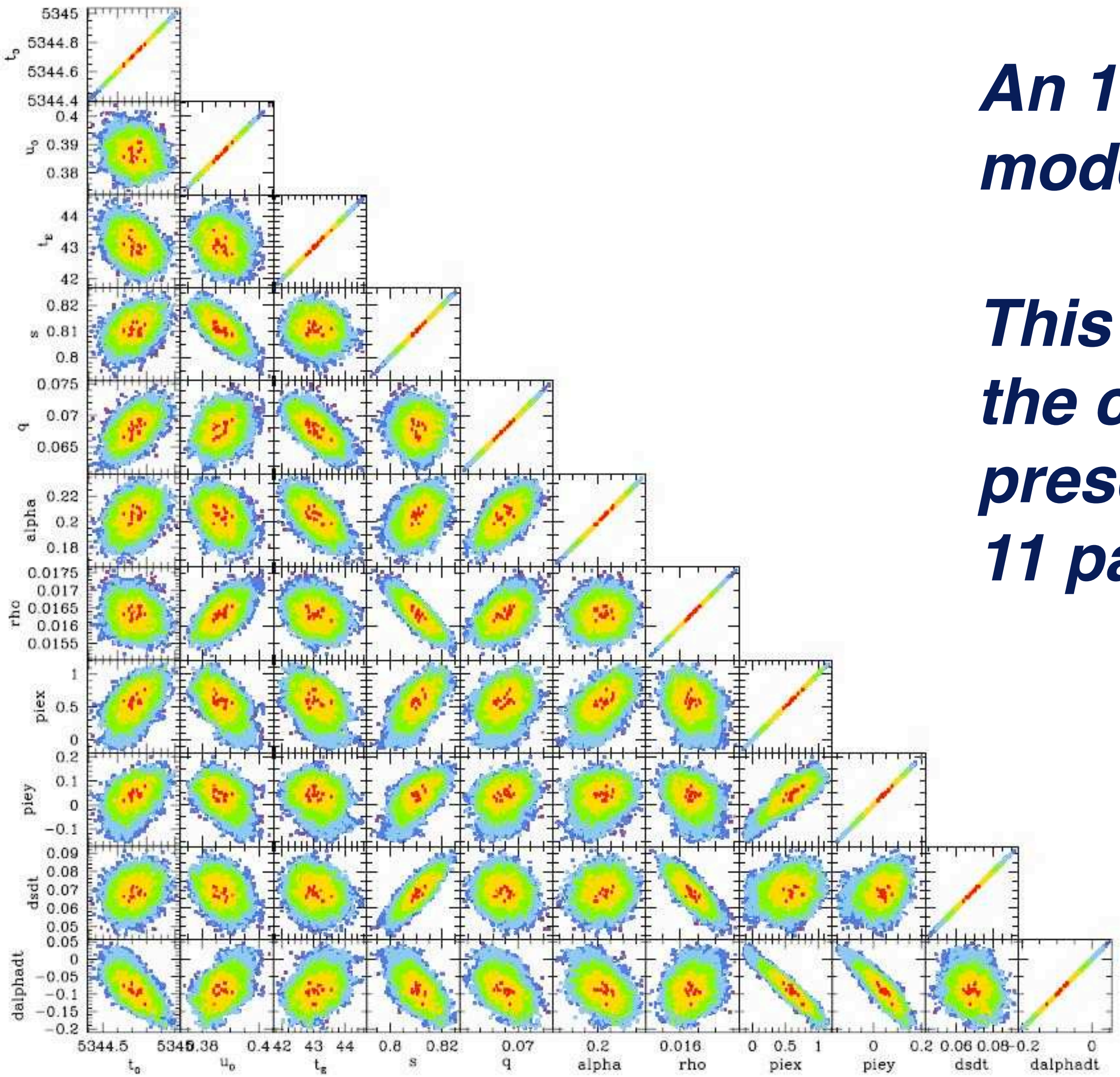
- **Correlation matrix:**

$$R_{ij} \equiv \frac{\mathrm{Cov}(X_i,X_j)}{\sigma(X_i)\,\sigma(X_j)} = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & 1 \end{bmatrix}$$

- **Variance:**

$$\mathrm{Var}\left[\sum_i a_i X_i\right] = \sum_i \sum_j a_i a_j \sigma(X_i)\sigma\left(X_j\right) R_{ij}$$

Example: Correlation Matrix



An 11-parameter model fitted to data.

This matrix shows the correlations present among the 11 parameters.

Fini -- ADA 02