ADA02 - 9am Tue 13 Sep 2022

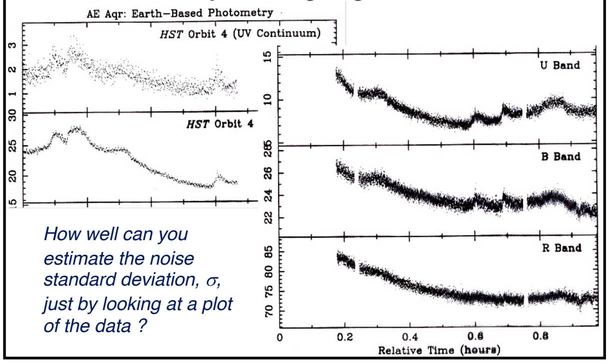
Eyeballing sigma (5-sigma rule)

Joint Probability Distributions
Independence vs Correlation

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

43

Eyeballing Sigma



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

44

Eyeballing Sigma

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

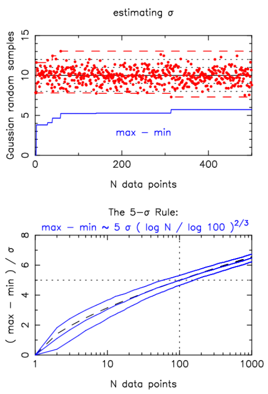
Estimate "by eye" the range (max-min) of the data (~100 data points).

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

Usually good to 20% by eye.

2. **If only a few data points:**

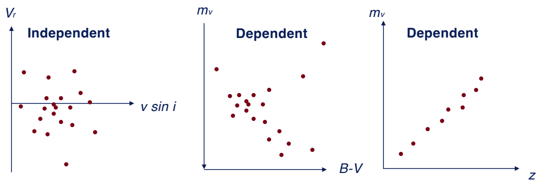
~2/3 of the data points should be inside +/- 1 sigma



45

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



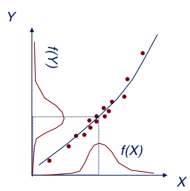
- Does knowing the value of one random variable X inform you about the other?

46

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

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



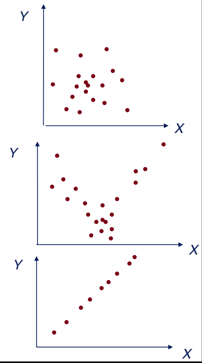
47

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



48

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?

49

Linear Transformations: Scaling

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

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:

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

$$\begin{aligned}\mathrm{Var}(a\,X) &= \langle [a\,X - \langle a\,X \rangle]^2 \rangle \\ &= \langle [a\,X - a\langle X \rangle]^2 \rangle \\ &= \langle a^2 [X - \langle X \rangle]^2 \rangle \\ &= a^2\, \mathrm{Var}(X)\end{aligned}$$

"Stretch the paper" by a factor ***a***.
Location μ and width σ then increase by factor ***a***.

50

Linear Transformations: Addition

- Adding two random variables X and Y :

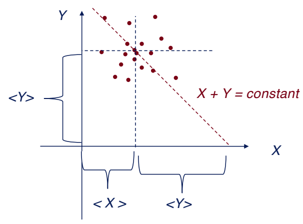
$$\begin{aligned}\langle X + Y \rangle &= \iint (X+Y) f(X,Y)\, dX\, dY \\ &= \iint X\, f(X,Y)\, dX\, dY + \iint Y\, f(X,Y)\, dX\, dY \\ &= \int X \left[\int f(X,Y) dY \right] dX + \int Y \left[\int f(X,Y) dX \right] dY \\ &= \int X\, f(X)\, dX + \int Y\, f(Y)\, dY \\ &\equiv \langle X \rangle + \langle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

- True for **any** joint PDF! $\quad \langle X + Y \rangle = \langle X \rangle + \langle Y \rangle$

51

Why it works...

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



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

52

Variance and Co-variance

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

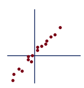
$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$
$$\mathrm{Cov}(X,Y) \equiv \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle$$

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

53

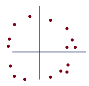
Co-variance vs Independence

- **Cov > 0**

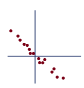


- **Cov = ?**

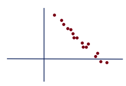
Independent?



- **Cov < 0**



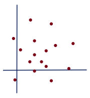
- **Cov = ?**



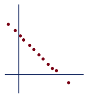
54

Practice !

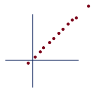
$X = 1 \pm 1$
$Y = 1 \pm 1$
$X + Y = ? \pm ?$



$X = 1 \pm 1$
$Y = 2 - X$
$X + Y = ? \pm ?$



$X = 1 \pm 1$
$Y = X$
$X + Y = ? \pm ?$



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

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

55

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**:

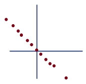
$$\mathrm{Var}\left[\sum_i a_i X_i\right] = \sum_{i,j} a_i C_{ij} a_j$$

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

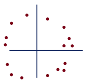
56

Correlation Coefficient *R(X, Y)*

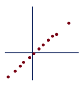
- **Correlation coefficient**: $R(X,Y) \equiv \dfrac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$



R = -1



R = 0



R = +1

- **Correlation matrix**:

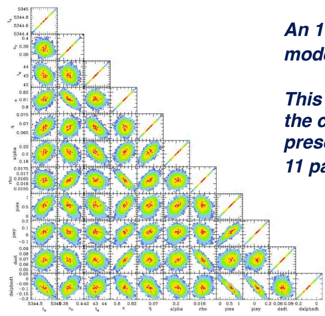
$$R_{ij} \equiv \frac{\mathrm{Cov}(X_i, X_j)}{\sigma(X_i)\,\sigma(X_j)} = \begin{bmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & 1 \end{bmatrix}$$

- **Variance**:

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

57

Example: Correlation Matrix



An 11-parameter model fitted to data.

This matrix shows the correlations present among the 11 parameters.

58

Fini -- ADA 02

59