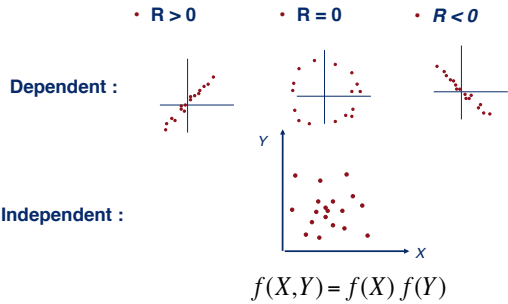
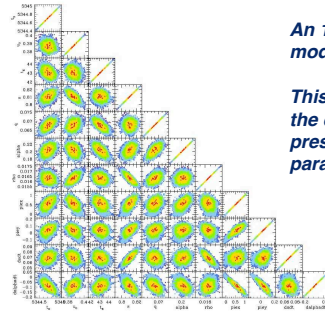


## Review: Correlation vs Independence



## Example: Correlation Matrix



An 11-parameter model fitted to data.

This matrix shows the correlations present among the 11 parameters.

## Review: Algebra of Random Variables

$$\langle a \rangle = a \quad \text{Var}[a] = 0$$

$$\langle aX \rangle = a \langle X \rangle \quad \text{Var}[aX] = a^2 \text{Var}[X]$$

$$\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle \quad \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$$

Co - variance :

$$\text{Cov}[X, Y] = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle \quad \text{Var}[X] = \text{Cov}[X, X]$$

Linear transformations :

$$\left\langle \sum_i a_i X_i \right\rangle = \sum_i a_i \langle X_i \rangle \quad \text{Var} \left[ \sum_i a_i X_i \right] = \sum_i \sum_j a_i a_j \sigma_i \sigma_j R_{ij}$$

Correlation Matrix :

$$R_{ij} = \frac{\text{Cov}(X_i, X_j)}{\sigma_i \sigma_j} \quad \sigma_i = \sigma(X_i)$$

## Practice the “fuzzy” algebra of random variables

$$6(1 \pm 1) =$$

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

$$(1 \pm 2) - (2 \pm 2) =$$

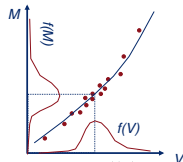
Practice until this becomes automatic ...

## Functions of Random Variables

Often what we can measure is not what we are most interested in!  
Example: mass of binary-star system:

$$M = \frac{V^2 a}{G} = \frac{V^3 P}{2\pi G}$$

We want  $M$ , but can only measure  $V$  and  $P$ .  
 $P$  = accurate, but  $V$  usually less certain.  
What is the uncertainty in  $M$ ?



For power-laws:  $\ln M = 3 \ln V + \ln P + \text{const.}$      $\sigma(\ln x) \approx \sigma(x)/\langle x \rangle$

$$\left( \frac{\sigma_M}{\langle M \rangle} \right)^2 \approx \left( 3 \frac{\sigma_V}{\langle V \rangle} \right)^2 + \left( \frac{\sigma_P}{\langle P \rangle} \right)^2$$

(valid for small and independent errors in  $V$  and  $P$ ).  
How do error bars propagate through non-linear functions?

## Functions of a Random Variable

$$Y = y(X) \quad \frac{dY}{dX} = y'(X)$$

conserve probability:

$$d(\text{Prob}) = f(Y) |dY| = f(X) |dX|$$

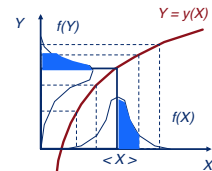
$$f(Y) = f(X) \left| \frac{dX}{dY} \right| = \frac{f(X)}{|y'(X)|}$$

mean value (biased)

$$\langle Y \rangle = y(\langle X \rangle) + \frac{1}{2} y''(\langle X \rangle) \sigma_x^2 + \dots$$

standard deviation (stretched)

$$\sigma_y = \sigma_x \left| \frac{dy}{dx} \right|_{x=\langle x \rangle} + \dots$$



Negative curvature:

Long tail for  $Y < y(\langle X \rangle)$

Bias:  $\langle Y \rangle < y(\langle X \rangle)$ .

Median is not biased:  
 $\text{Med}(Y) = y(\text{Med}(X))$

## Examples of Non-linear Transformations

Spectral Energy Distributions: per unit **wavelength** ( $\text{erg cm}^{-2} \text{Å}^{-1}$ ),  
or per unit **frequency** ( $\text{erg cm}^{-2} \text{s}^{-1} \text{Hz}^{-1}$ )

$$f_\nu(\lambda) |d\nu| = f_\lambda(\lambda) |d\lambda|$$

$$\nu = \frac{c}{\lambda} \quad d\nu = -\frac{c}{\lambda^2} d\lambda \Rightarrow f_\nu(\lambda) = \left| \frac{d\lambda}{d\nu} \right| f_\lambda(\lambda) = \frac{\lambda^2}{c} f_\lambda(\lambda)$$

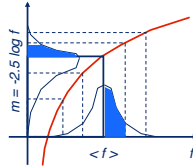
Converting a flux to a magnitude:

- Measure Flux: Gaussian distribution:  $f \sim G(\langle f \rangle, \sigma_f^2)$
- Nonlinear transformation induces a bias:

$$m = m_0 - 2.5 \log f$$

$$\langle m \rangle = m_0 - 2.5 \log \langle f \rangle + a \sigma_m^2$$

- PROBLEM: evaluate  $a$ ,  $\sigma_m$  in terms of  $\langle f \rangle$ ,  $\sigma_f$ .



## Nonlinear Transformations: A Bias from Curvature + Noise

Taylor expand  $Y = y(X)$  around  $X = \langle X \rangle$ :

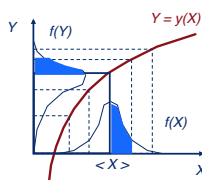
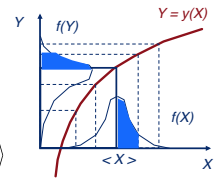
$$y(X) = y(\langle X \rangle) + y'(\langle X \rangle) \epsilon + \frac{1}{2} y''(\langle X \rangle) \epsilon^2 + \dots$$

$$\text{where } \epsilon = X - \langle X \rangle, \langle \epsilon \rangle = 0, \langle \epsilon^2 \rangle = \sigma_X^2.$$

Hence

$$\begin{aligned} \langle y(X) \rangle &= \left\langle y(\langle X \rangle) + y'(\langle X \rangle) \epsilon + \frac{1}{2} y''(\langle X \rangle) \epsilon^2 + \dots \right\rangle \\ &= y(\langle X \rangle) + y'(\langle X \rangle) \langle \epsilon \rangle + \frac{1}{2} y''(\langle X \rangle) \langle \epsilon^2 \rangle + \dots \\ &= y(\langle X \rangle) + 0 + \frac{1}{2} y''(\langle X \rangle) \sigma_X^2 + \dots \end{aligned}$$

This is the bias.



## Variance of a Transformed Variable

Tangent-curve approximation :

$\sigma(y(x)) = \sigma(x)$  stretched by a factor  $dy/dx$ .

$$\begin{aligned} \sigma^2(Y) &= \langle (Y - \langle Y \rangle)^2 \rangle = \left\langle \left[ y(\langle X \rangle) + y'(\langle X \rangle) \epsilon + \frac{1}{2} y''(\langle X \rangle) \epsilon^2 + \dots \right. \right. \\ &\quad \left. \left. - y(\langle X \rangle) - 0 - \frac{1}{2} y''(\langle X \rangle) \sigma_X^2 - \dots \right]^2 \right\rangle \\ &= \left\langle \left[ y'(\langle X \rangle) \epsilon + O(\epsilon^2) \right]^2 \right\rangle = \left[ y'(\langle X \rangle) \right]^2 \sigma_X^2 \end{aligned}$$

Can extend to higher-order terms (skew, kurtosis) if needed.

## Magnitude Bias

Observe flux:  $f \sim G(f_0, \sigma_f^2) = (f_0 \pm \sigma_f)$

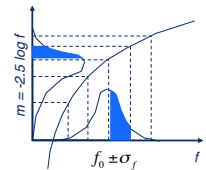
Transform to magnitude:  $m(f) = m_0 - 2.5 \log f$

$$\text{Derivatives: } (\log f = \log e \ln f) \quad \frac{dm}{df} = -\frac{2.5 \log e}{f}, \quad \frac{d^2 m}{df^2} = \frac{2.5 \log e}{f^2}$$

$$\sigma_m \approx \left| \frac{dm}{df} \right| \sigma_f = \frac{2.5 \log e}{f} \sigma_f \approx 1.08 \frac{\sigma_f}{f}$$

$$\begin{aligned} \langle m \rangle &= m(\langle f \rangle) + \frac{1}{2} \frac{d^2 m}{df^2} \sigma_f^2 \\ &= m_0 - 2.5 \log \langle f \rangle + \frac{2.5 \log e}{2 \langle f \rangle^2} \sigma_f^2 \\ &= m_0 - 2.5 \log \langle f \rangle + \frac{\sigma_m^2}{5 \log e} \end{aligned}$$

Note the bias toward faint magnitudes.



## Magnitude Bias

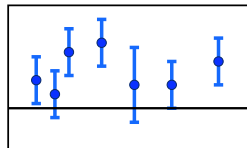
$$m(f) = m_0 - 2.5 \log f$$

$$\sigma_m = (2.5 \log e) \frac{\sigma_f}{\langle f \rangle} \approx 1.08 \frac{\sigma_f}{\langle f \rangle}$$

$$\langle m \rangle = m(\langle f \rangle) + \text{bias}(\langle m \rangle)$$

$$\text{bias}(\langle m \rangle) = \left( \frac{\sigma_m}{\sqrt{5 \log e}} \right)^2 \approx 0.01 \left( \frac{\sigma_m}{0.15} \right)^2$$

15% errors  $\rightarrow$  1% bias  
50% errors  $\rightarrow$  10% bias



Given noisy fluxes, should you average fluxes and then compute the magnitude?

Or, should you convert fluxes to magnitudes and average the magnitudes?

$$m(f) = m_0 - 2.5 \log \langle f \rangle$$

$$\text{or } m(f) = \langle m_0 - 2.5 \log f \rangle$$

## Cartesian $\rightarrow$ Polar coordinates e.g. Amplitude and Phase

Independent measurements of C and S

(e.g. cos and sin amplitudes of an oscillation):

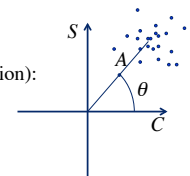
$$S = A \sin \theta \sim (S_0 \pm \sigma_S)$$

$$C = A \cos \theta \sim (C_0 \pm \sigma_C)$$

Transform to amplitude and phase:

$$A = ? \pm ?$$

$$\theta = ? \pm ?$$



## Transforming Random Numbers

Uniform  $\rightarrow$  Lorentzian

$u \sim U(0,1) \rightarrow x \sim L(\mu, \sigma)$

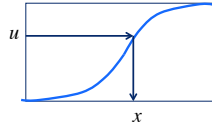
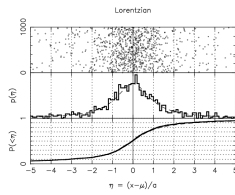
$$u = F(x) = \frac{1}{\pi} \arctan\left[\frac{x-\mu}{\sigma}\right] + \frac{1}{2}$$

$$x = F^{-1}(u) = \mu + \sigma \tan\left[\pi\left(u - \frac{1}{2}\right)\right]$$

Practice :

Uniform  $\rightarrow$  Exponential

Uniform  $\rightarrow$  Power - law



## Box-Muller Transform

For Gaussians, cumulative probability  $F(x)$  has no analytic expression.  
Harder to generate Gaussian random numbers  $x = F^{-1}(u)$   
from Uniform random numbers  $u$ .

Two independent uniform random numbers :

$$x \sim U(-1, +1) \quad y \sim U(-1, +1)$$

Keep if  $r^2 = x^2 + y^2 < 1$

Two independent gaussian random numbers :

$$G_1 = \frac{2x}{r} (-\ln r)^{1/2} \quad G_2 = \frac{2y}{r} (-\ln r)^{1/2}$$

$$r = 0 \rightarrow G = \infty$$

$$r = 1 \rightarrow G = 0$$

$G_1$  and  $G_2$  have mean 0 and variance 1:

$$G_1 \sim G(0,1) \quad G_2 \sim G(0,1)$$

