ADA03 - 9am Thu 15 Sep 2022

Brief Review
Spurious Correlations
Correlation vs Causation

Non-Linear Transformations
Bias corrections

Transforming random numbers
Uniform -> Lorentzian
Uniform -> Gaussian

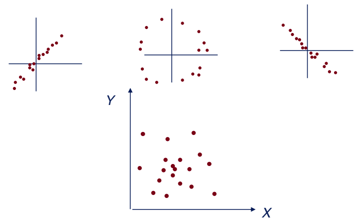
60

Review: Correlation vs Independence

- R > 0
- R = 0
- *R < 0*

Dependent :

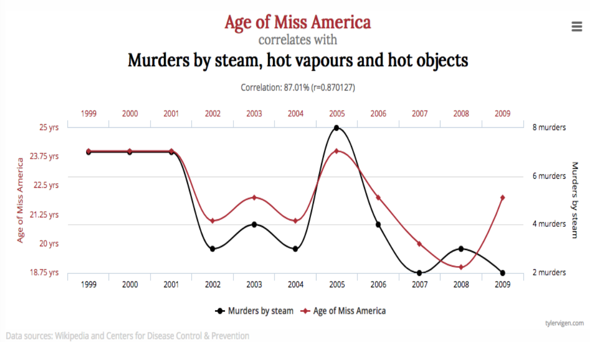
Independent :



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

61

Example of a Spurious Correlation



62

1: Beware Spurious Correlations

- Two variables may appear to be strongly correlated.
- But, can be spurious if you look at many variables, to find the strongest correlations, then pretend you only looked at those.

2 : Correlation is not Causation

- Correlation of 2 variables does not mean that one causes the other. Both could be side effects of something else.

63

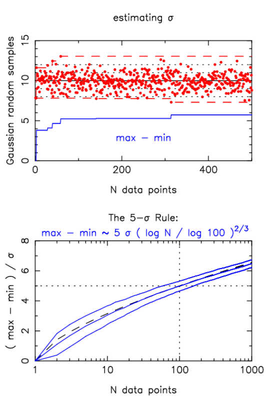
Misleading Significance Claims

If we look at 100 points, we typically find 2 that are 5-sigma apart.

If we pull out those 2

(and omit the others)

we can't honestly claim to have a 5-sigma result.



64

Review: Algebra of Random Variables

$$\langle a \rangle = a \qquad \mathrm{Var}[a] = 0$$

$$\langle a X \rangle = a \langle X \rangle \qquad \mathrm{Var}[a X] = a^2\, \mathrm{Var}[X]$$

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

Co - variance :

$$\mathrm{Cov}[X,Y] \equiv \Big\langle \big(X - \langle X \rangle \big)\big(Y - \langle Y \rangle \big) \Big\rangle \qquad \mathrm{Var}[X] = \mathrm{Cov}[X,X]$$

Linear transformations :

$$\Big\langle \sum_i a_i X_i \Big\rangle = \sum_i a_i \langle X_i \rangle \qquad \mathrm{Var}\Big[\sum_i a_i X_i \Big] = \sum_i \sum_j a_i\, a_j\, \sigma_i\, \sigma_j\, R_{ij}$$

Correlation Matrix :

$$R_{ij} \equiv \frac{\mathrm{Cov}(X_i, X_j)}{\sigma_i\, \sigma_j} \qquad \sigma_i \equiv \sigma(X_i)$$

65

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 …

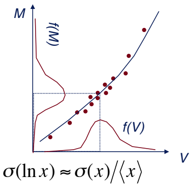
66

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

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

(valid for ***small*** and ***independent*** errors in *V* and *P*).

How do error bars propagate through non-linear functions?

67

Functions of a Random Variable

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

Conserve probability:

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

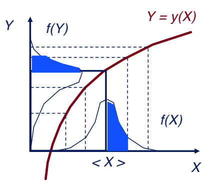
$f(Y) = f(X) \left|\frac{dX}{dY}\right| = \frac{f(X)}{|y'(X)|}$

mean value (biased)

$\langle Y \rangle = y(\langle X \rangle) + \frac{1}{2} y''(\langle X \rangle)\, \sigma_X^2 + \ \ldots$

standard deviation (stretched)

$\sigma_Y = \sigma_X \left|\frac{dy}{dx}\right|_{X = \langle X \rangle} + \ \ldots$



Negative curvature:

Long tail for *Y* < *y*(< *X* >)

Bias: <Y> < y(< X >).

Median is not biased:
Med*(Y)* = *y(* Med*(X))*

68

Examples of Non-linear Transformations

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

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

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

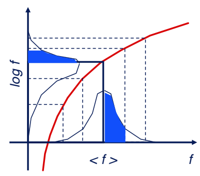
Converting a **flux** to a **magnitude**:

- Measure Flux: Gaussian distribution: $f \sim G(<f>, \sigma_f^2)$
- 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, σ_m in terms of $<f>$, σ_f .



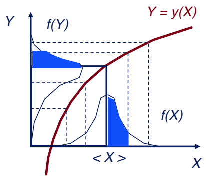
69

Nonlinear Transformations: A Bias from Curvature + Noise

Taylor expand *Y = y(X)* around *X = < X >* :

$$y(X) = y(\langle X \rangle) + y'(\langle X \rangle)\, \varepsilon + \frac{1}{2} y''(\langle X \rangle)\, \varepsilon^2 + \ldots$$

where $\varepsilon \equiv X - \langle X \rangle, \quad \langle \varepsilon \rangle = 0, \quad \langle \varepsilon^2 \rangle = \sigma_X{}^2 .$



Hence (using the algebra of random variables):

$$\langle y(X) \rangle = \left\langle y(\langle X \rangle) + y'(\langle X \rangle)\, \varepsilon + \frac{1}{2} y''(\langle X \rangle)\, \varepsilon^2 + \ldots \right\rangle$$

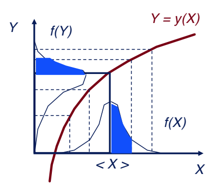
$$= y(\langle X \rangle) + y'(\langle X \rangle)\, \langle \varepsilon \rangle + \frac{1}{2} y''(\langle X \rangle)\, \langle \varepsilon^2 \rangle + \ldots$$

$$= y(\langle X \rangle) + \quad 0 \quad + \frac{1}{2} y''(\langle X \rangle)\ \sigma_X{}^2 + \ldots$$

This is the bias.

70

Variance of a Transformed Variable



Tangent-curve approximation :

σ(y(x)) = σ(x) **stretched** by a factor *|dy/dx|*.

$$\sigma^2(Y) \equiv \left\langle (Y - \langle Y \rangle)^2 \right\rangle = \left\langle \left[y(\langle X \rangle) + y'(\langle X \rangle)\, \varepsilon + \frac{1}{2} y''(\langle X \rangle)\, \varepsilon^2 + \ldots \right.\right.$$

$$\left.\left. - y(\langle X \rangle) \quad - \quad 0 \quad - \frac{1}{2} y''(\langle X \rangle)\, \sigma^2(X) - \ldots \right]^2 \right\rangle$$

Using the algebra of random variables:

$$= \left\langle \left[y'(\langle X \rangle)\, \varepsilon + O(\varepsilon^2) \right]^2 \right\rangle = \left[y'(\langle X \rangle) \right]^2 \sigma_X^2 \quad + \quad \ldots$$

Could extend to higher-order terms (skew, kurtosis) if needed,
but fast computers make it easier to use Monte-Carlo error propagation.

71

Example : Magnitude Bias

Observe flux: $f = (f_0 \pm \sigma_f)$

Convert to a magnitude: $m(f) \equiv m_0 - 2.5\log f = m_0 - (2.5\log e)\ln f$

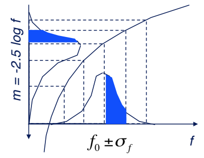
Derivatives: $(\log f = \log e \ln f)$ $m'(f) = -\frac{2.5\log e}{f}$, $m''(f) = \frac{2.5\log e}{f^2}$.

$$\sigma_m \approx |m'(f_0)|\sigma_f = \frac{2.5\log e}{f_0}\sigma_f \approx 1.08\frac{\sigma_f}{f_0}.$$

$$\langle m \rangle = m(f_0) + \frac{m''(f_0)}{2}\sigma_f^2 + \ldots$$

$$= m_0 - 2.5\log(f_0) + \frac{2.5\log e}{2\,f_0^2}\sigma_f^2$$

$$= m_0 - 2.5\log(f_0) + \frac{\sigma_m^2}{5\log e}$$



Note the bias toward faint magnitudes.

72

Example : Magnitude Bias

converting noisy fluxes to magnitudes:

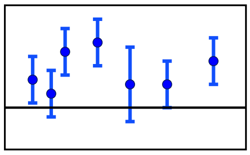
$f = f_0 \pm \sigma_f$ $\quad m(f) \equiv m_0 - 2.5\log f$

$\sigma_m = (2.5\log e)\frac{\sigma_f}{f_0} \approx 1.08\frac{\sigma_f}{f_0}$.

$\langle m \rangle = m(f_0) + \text{bias}$

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

15% uncertainty -> 1% bias
50% uncertainty -> 10% bias



Given noisy fluxes, you could first average the fluxes and then compute the magnitude:

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

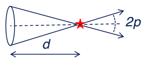
or, first convert each flux to a magnitude and then average the magnitudes:

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

Which method gives the smaller bias ?

73

Example: Distance from Parallax measurements



Parallax is the apparent motion of stars as the Earth orbits the Sun.

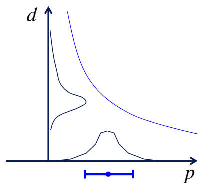
$$\frac{d}{\text{parsec}} = \left(\frac{p}{\text{arcsec}}\right)^{-1}$$

Measure a parallax, with Gaussian error,

$$p = p_0 \pm \sigma_p$$

Estimate the distance and its uncertainty:

$$d = \frac{1}{p_0} + bias \pm \sigma_d$$



Include a correction for the bias due to the non-linear transformation.

74

Example : Cartesian -> Polar coordinates e.g. Amplitude and Phase

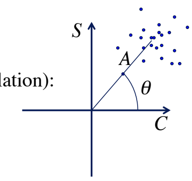
Independent measurements of C and S
(e.g. cos and sin amplitudes of an oscillation):

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

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

Transform to amplitude and phase:

$A \;=\; ? \;\pm\; ? \qquad \theta \;=\; ? \;\pm\; ?$



75

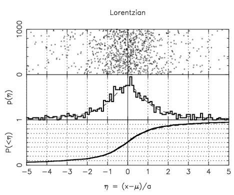
How to Transform Random Numbers

Uniform $\rightarrow$ Lorentzian

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

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

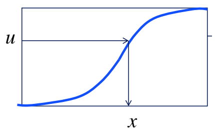
$$x \;=\; F^{-1}(u) = \mu + \sigma\tan\left[\pi\left(u - \frac{1}{2}\right)\right]$$



Practice :

Uniform $\rightarrow$ Exponential

Uniform $\rightarrow$ Power - law



76

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:

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

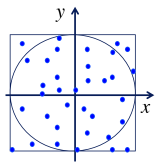
Keep if $\quad r^2 = x^2 + y^2 < 1 \quad$ and $\quad r > 0.$

Two independent gaussian random numbers:

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

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

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



G_1 and G_2 have mean 0 and variance 1:
$G_1 \sim G(0,1)$ $\quad G_2 \sim G(0,1)$

77

Fini -- ADA 03

78