ADA06 - 10am Mon 26 Sep 2022

Recap : χ^2 fitting

Sample Variance S^2
bias correction

"Robust" Statistics:
Median and MAD
sigma clipping

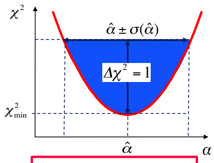
119

Recap : Fitting models by minimizing χ^2

Fit M parameters to N data points.

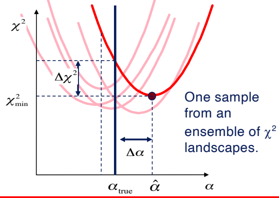
$$\chi^2(X,\sigma,\alpha) \equiv \sum_{i=1}^{N}\left(\frac{X_i-\mu_i(\alpha)}{\sigma_i}\right)^2$$

Best - fit parameters $\hat\alpha$ minimise χ^2.



$$\sigma^2(\hat\alpha) = \frac{2}{\left.\frac{\partial^2\chi^2}{\partial\alpha^2}\right|_{\alpha=\hat\alpha}}$$

Caveat: Assumes orthogonal parameters.
Generalise to correlated parameters later.

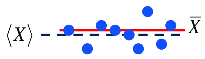


$$\hat\alpha \sim G\left(\alpha_{true}, \sigma^2(\hat\alpha)\right)$$
$$\chi^2(\alpha_{true}) \sim \chi^2_N$$
$$\chi^2_{min} \equiv \chi^2(\hat\alpha) \sim \chi^2_{N-M}$$
$$\Delta\chi^2 \equiv \chi^2(\alpha_{true}) - \chi^2_{min} \sim \chi^2_M$$

$$\langle\chi^2_N\rangle = N\ , \quad \text{Var}\left[\chi^2_N\right] = 2\,N$$

120

Data points with no error bars ☹



N data points: $\langle X_i\rangle = \langle X\rangle$ $\quad \text{Cov}(X_i,X_j) = \sigma^2\,\delta_{ij}$

Sample mean: $\bar X \equiv \frac{1}{N}\sum_i X_i$ $\quad$ unbiased: $\langle\bar X\rangle = \langle X\rangle$. $\quad \text{Var}(\bar X) = \frac{\sigma^2}{N}$

But σ_i^2 are unknown. How can we estimate σ^2 ?

Variance: $\sigma^2(X) \equiv \left\langle\left(X-\langle X\rangle\right)^2\right\rangle$

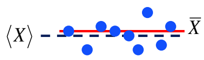
Try: $s^2 \equiv \frac{1}{N}\sum_i\left(X_i-\bar X\right)^2$

Is $\langle s^2\rangle = \sigma^2$?

No. $\langle s^2\rangle < \sigma^2$ We can evaluate and then correct for this bias.

121

Sample Variance S^2 : Unbiased for σ^2



$S^2 \equiv A\sum_{i=1}^{N}(X_i-\bar X)^2$ $\quad$ Pick A so that $\langle S^2\rangle = A\sum_{i=1}^{N}\left\langle(X_i-\bar X)^2\right\rangle = \sigma^2$

$$\left\langle(X_i-\bar X)^2\right\rangle = \left\langle\left[(X_i-\langle X\rangle)-(\bar X-\langle X\rangle)\right]^2\right\rangle$$
$$= \left\langle(X_i-\langle X\rangle)^2 + (\bar X-\langle X\rangle)^2 - 2\,(X_i-\langle X\rangle)\,(\bar X-\langle X\rangle)\right\rangle$$
$$= \sigma^2(X_i) + \sigma^2(\bar X) - 2\,\text{Cov}(X_i,\bar X)$$
$$= \sigma^2 \quad + \frac{\sigma^2}{N} \quad - 2\,\frac{\sigma^2}{N}$$
$$= \left(1-\frac{1}{N}\right)\sigma^2 = \left(\frac{N-1}{N}\right)\sigma^2$$

Note : $\text{Cov}(X_i,\bar X) = \frac{\sigma^2}{N}$

$\therefore\ \langle S^2\rangle = A\sum_{i=1}^{N}\left(\frac{N-1}{N}\right)\sigma^2$ $\quad$ Pick $A = \frac{1}{N-1}$

$$S^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i-\bar X\right)^2$$

122

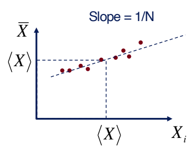
Evaluation of Cov(X_i , $\bar X$)

$$\text{Cov}(X_i,\bar X) \equiv \left\langle(X_i-\langle X_i\rangle)\,(\bar X-\langle\bar X\rangle)\right\rangle$$

Note : $\langle X_i\rangle = \langle\bar X\rangle = \langle X\rangle$

Shift coords to put $\langle X\rangle = 0$:

$$\text{Cov}(X_i,\bar X) = \left\langle(X_i-0)\,(\bar X-0)\right\rangle$$
$$= \left\langle X_i\,\frac{1}{N}\sum_k X_k\right\rangle$$
$$= \frac{1}{N}\sum_k\langle X_i\,X_k\rangle$$
$$= \frac{1}{N}\sum_k\sigma^2\delta_{ik} = \frac{\sigma^2}{N}$$

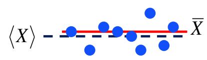


$$\text{Cov}(X_i,X_j) \equiv \sigma^2\,\delta_{ij}$$

123

Sample Variance S^2 : Unbiased for σ^2

$$S^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar X)^2$$



Why $\frac{1}{N\text{-}1}$, not $\frac{1}{N}$?

Because $\bar X$ "chases" the dancing data points,
removing 1 "degree-of-freedom" from the dance.

$$S^2 \sim \frac{\sigma^2}{N-1}\,\chi^2_{N-1}$$
$$\langle S^2\rangle = \frac{\sigma^2}{N-1}\left\langle\chi^2_{N-1}\right\rangle$$
$$= \frac{\sigma^2}{N-1}\left(N-1\right) = \sigma^2$$

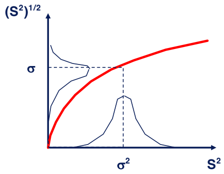
$$\text{Var}\left[S^2\right] = \left(\frac{\sigma^2}{N-1}\right)^2\text{Var}\left[\chi^2_{N-1}\right]$$
$$= \left(\frac{\sigma^2}{N-1}\right)^2 2\left(N-1\right) \quad = \frac{2\,\sigma^4}{N-1}$$
$$\frac{\sigma(S^2)}{\langle S^2\rangle} = \left(\frac{2}{N-1}\right)^{1/2} = \text{fractional accuracy}$$

124

Is $(S^2)^{1/2}$ unbiased for σ ?

- The sample variance S^2 is unbiased for σ^2.
- i.e. $< S^2 > = \sigma^2$
- Is $(S^2)^{1/2}$ unbiased for σ ?
- No. The square root introduces a bias:

$$S^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$$



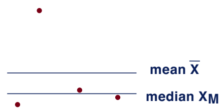
Homework:
Work out the bias correction, as a function of *N*, to construct an unbiased estimate for σ^p.

$$\left\langle \sqrt{S^2} \right\rangle < \sigma, \text{ even though } \sqrt{\sigma^2} = \sigma.$$

125

"Robust" estimation methods

Robust Statistics :
less sensitive to "bad" data
Example: use the **median** rather than the **mean**.



Sample Mean $\bar{X}$ minimizes the **Sample Variance**:

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \mu)^2$$

$$\frac{\partial}{\partial\mu}\left[\sum_{i=1}^{N}(X_i - \mu)^2\right] = 0$$

for $\mu = \bar{X}$

Median X_M minimizes the **"Mean Absolute Deviation"** :

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N}|X_i - \mu|$$

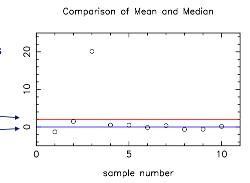
$$\frac{\partial}{\partial\mu}\left[\sum_{i=1}^{N}|X_i - \mu|\right] = 0$$

for $\mu = X_M \equiv \text{Median}(X_i)$

126

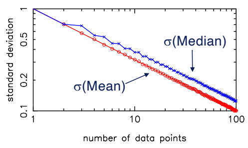
Mean vs Median

- The median is less sensitive to outliers than the mean.



- The median is **unbiased**, but **not** a minimum-variance estimator.
- Note how the standard deviations of the median and of the mean vary with sample size.

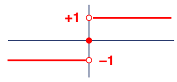
$$\sigma\left(\bar{X}\right) = \frac{\sigma}{\sqrt{N}} \leq \sigma\left(X_{\text{Med}}\right)$$



127

"Proof" that the Median minimises the MAD

$$H(x) \equiv \begin{cases} +1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \qquad \frac{\mathrm{d}H}{\mathrm{d}x} = 2\delta(x)$$



$$MAD \equiv \frac{1}{N}\sum_{i=1}^{N}|X_i - \mu| = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)\,H(X_i - \mu)$$

$$\frac{d\,MAD}{d\,\mu} = \frac{1}{N}\sum_{i=1}^{N}\left[(-1)H(X_i - \mu) + (X_i - \mu)(-1)H'(X_i - \mu)\right] = 0$$

$$= \frac{-1}{N}\sum_{i=1}^{N}H(X_i - \mu) = \frac{-1}{N}\left(\sum_{X_i > \mu}(+1) + \sum_{X_i < \mu}(-1)\right)$$

since $H'(x) = 0$ whenever $x \neq 0$

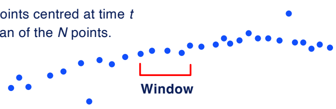
$= 0$ if $\mu = \text{median}(X_i)$

128

Median Filter and Sigma-Clip

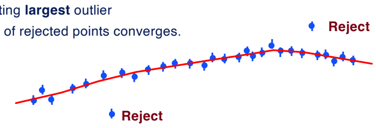
Median filter:

- Window encloses *N* points centred at time *t*
- Medfilt(*t*) is the median of the *N* points.



Sigma-clip:

- Fit all points by minimising χ^2
- Set threshold *K* and check for outliers at $\pm K\sigma$ or more
- Repeat fit omitting **largest** outlier
- Iterate until set of rejected points converges.



129

Various "Badness-of-Fit" Statistics

Sample Variance — mean

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \mu_i)^2 \rightarrow \bar{X}$$

Chi-squared — optimal average

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \rightarrow \hat{X}$$

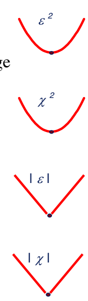
Mean Absolute Deviation — median

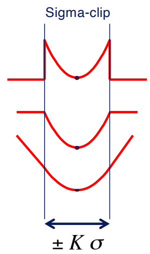
$$MAD = \frac{1}{N}\sum_{i=1}^{N}|X_i - \mu_i| \rightarrow X_M$$

Sum Absolute Normalised Errors:

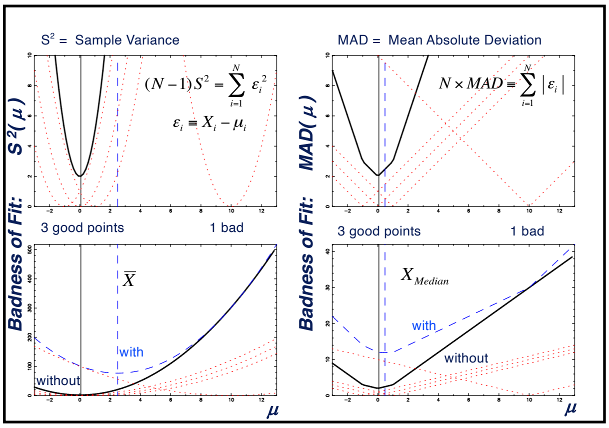
$$SANE = \sum_{i=1}^{N}\left|\frac{X_i - \mu_i}{\sigma_i}\right|$$

Badness functions:

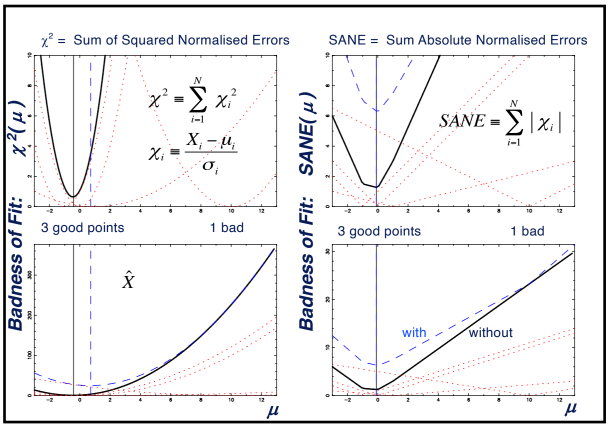




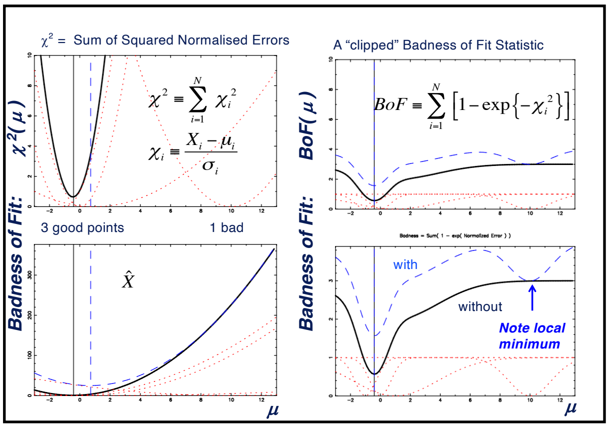
130



131



132



133

Fini -- ADA 06

134