# ADA 12 - 10am Mon 10 Oct 2022

Diagnosis of $\chi^2_{\min}$ above/below DoF
( Reject model? Clip outliers?
Rescale error bars? )
ML Estimate for Excess Variance

Background Functions
(polynomials, spines, Running Optimal Average (ROA), median filter)

228

## Use $\chi^2_{\min}$ (or AIC,BIC,...) to reject models

Fit $M$ parameters to $N$ data points:

$$\chi^2_{\min} = \sum_{i=1}^{N} \left[ \frac{X_i - \mu_i(\alpha_1,...,\alpha_M)}{\sigma_i} \right]^2 \sim \chi^2_{N-M}$$

$$\left\langle \chi^2_{N-M} \right\rangle = N - M \qquad \sigma^2(\chi^2_{N-M}) = 2(N-M)$$

Why $N$ - $M$ degrees of freedom?
Fitting $M = N$ parameters should fit $N$ points exactly.
If model is good, then the best-fit $\chi^2_{\min}$ should be:

$$\chi^2_{\min} \approx N - M \pm \sqrt{2(N-M)}$$

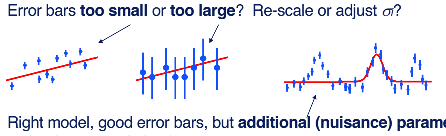
$$\frac{\chi^2_{\min}}{N-M} \approx 1 \pm \sqrt{\frac{2}{N-M}}$$

229

## What if $\chi^2_{\min}$ is too high (or low)?

Several possibilities:

1. Statistical fluke? Use $\chi^2_{N-M}$ distribution to estimate probability ( p-value )
2. A few outliers ? Use e.g. sigma-clipping to identify and reject outliers.
3. Wrong model? Use $\chi^2_{N-M}$ distribution to reject model ( p < threshold )
4. Error bars **too small** or **too large**? Re-scale or adjust $\sigma_i$?



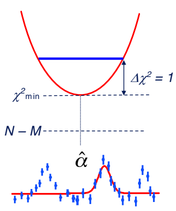
5. Right model, good error bars, but **additional (nuisance) parameters** omitted or not optimised?

Failure to optimise nuisance parameters increases $\chi^2_{\min}$ , but may leave the $\chi^2$ curvature the same, **if the nuisance parameters are orthogonal to the parameters of interest**.

Can then still use $\Delta\chi^2$ to set confidence intervals on parameters **orthogonal to the nuisance parameters**.

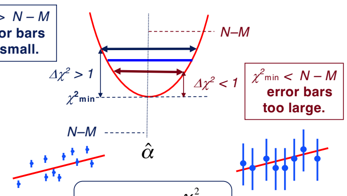
230

## Diagnosis of $\chi^2_{\min}$ too large or small




$\chi^2_{\min} > N - M$ due to failure to optimise nuisance parameters ? If these are **orthogonal** to parameters of interest, then the curvature at the minimum is unchanged. Use $\Delta\chi^2 = 1$.

$\chi^2_{\min} > N - M$
**error bars too small.**




$\chi^2_{\min} < N - M$
**error bars too large.**

Use: $\Delta\chi^2 = \dfrac{\chi^2_{\min}}{N-M}$

Equivalent to **re-scaling the error bars**

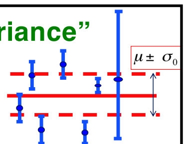
$$\sigma \Rightarrow \sigma \sqrt{\frac{\chi^2_{\min}}{N-M}}$$

231

## ML Estimate for the "Extra Variance"

Assume two independent noise sources:

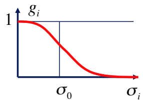
Errors with known $\sigma_i$. Extra variance $\sigma_0^2$ .




$$\mathrm{Var}[X_i] = \sigma_0^2 + \sigma_i^2 = \frac{\sigma_0^2}{g_i} \qquad g_i \equiv \frac{\sigma_0^2}{\sigma_0^2 + \sigma_i^2} = \frac{1}{1 + (\sigma_i/\sigma_0)^2}$$

$$-2\ln L = \sum_{i=1}^{N} \frac{(X_i - \mu)^2}{\sigma_0^2 + \sigma_i^2} + \sum_{i=1}^{N} \ln\left(\sigma_0^2 + \sigma_i^2\right)$$

$g_i$ = "goodness" of $X_i$ for measuring $\sigma_0^2$




$$0 = \frac{\partial(-2\ln L)}{\partial\mu} = -2\sum_{i=1}^{N} \frac{X_i - \mu}{\sigma_0^2 + \sigma_i^2}$$

$$0 = \frac{\partial(-2\ln L)}{\partial\sigma_0^2} = -\sum_{i=1}^{N}\left(\frac{X_i - \mu}{\sigma_0^2 + \sigma_i^2}\right)^2 + \sum_{i=1}^{N} \frac{1}{\sigma_0^2 + \sigma_i^2} = -\sum \frac{(X_i - \mu)^2 g_i^2}{\sigma_0^4} + \sum \frac{g_i}{\sigma_0^2}$$

$$\hat{\mu} = \frac{\sum \frac{X_i}{\sigma_0^2 + \sigma_i^2}}{\sum \frac{1}{\sigma_0^2 + \sigma_i^2}} = \frac{\sum X_i g_i}{\sum g_i} \qquad \mathrm{Var}[\hat{\mu}] = \frac{\sigma_0^2}{\sum g_i} \qquad \hat{\sigma}_0^2 = \frac{\sum (X_i - \mu)^2 g_i^2}{\sum g_i}$$

***Need to iterate.***

Homework : Work out Var[ $\hat{\sigma}_0^2$ ] .

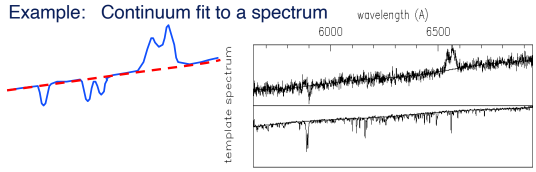
232

## Background Functions

**Smooth functions with adjustable flexibility.**

- Polynomials
- Splines
- Running Optimal Average
- Any of the above – with sigma clipping.
- Running Median

Example: Continuum fit to a spectrum




233

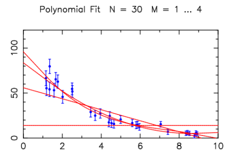
## Polynomials

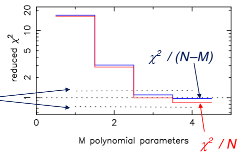
Fit $N = 30$ points with $M = 1, 2, 3, 4$ polynomial coefficients.

Higher $M$ = more flexible model.
Use lowest $M$ that gives good fit.
e.g. minimise AIC or BIC

Reject $M = 1, 2$.
Accept $M = 3, 4$.
Based on Reduced $\chi^2$

$$\frac{\chi^2}{N-M} \approx 1 \pm \sqrt{\frac{2}{N-M}}$$

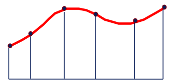



234

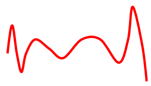
## Splines – e.g. piecewise cubic

***N* nodes**: $x_i$, $y_i$, $i = 1, \dots N$. $x_i$ fixed, $y_i$ adjustable.
**$4(N-1)$ parameters** ( 4 cubic coefficients for each of the $N-1$ segments )
**$3(N-2)$ matching conditions** ( value, slope, curvature at each of the $N-2$ internal nodes )
**$N+2$ degrees of freedom** ( $N$ values $y_i$ plus either slope or curvature at 2 end points).



8-parameter cubic spline

- First, **distribute the nodes** $x_i$, e.g. equally spaced, or equal weight $\Sigma(1/\sigma^2)$ on each segment.
- Then, **fit $N+2$ parameters**, e.g. find $y_i$ to minimise $\chi^2$, with endpoint curvatures (or slopes) set to zero.



8-parameter polynomial

Low-order polys are good for simple background fits.
Splines better than high-order polys. Better control over the $x$ distribution of the degrees of freedom.

235

## Running Optimal Average (ROA)

Time-series data: $X_i \pm \sigma_i$ at times $t_i$

$$\hat{X}(t) = \frac{\sum X_i\, w_i(t)}{\sum w_i(t)} \qquad \sigma^2\left(\hat{X}(t)\right) = \frac{1}{\sum w_i(t)}$$

$$w_i(t) = \frac{G(t-t_i)}{\sigma_i^2}$$

**Memory function $G(t)$** expands the error bars as time-difference increases.

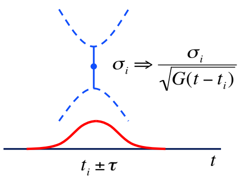
Parameter $\tau$ controls time interval over which the data point retains its $1/\sigma^2$ weight.

Memory functions:

boxcar: $G(t) = \begin{cases} 1 & |t| < \tau \\ 0 & |t| > \tau \end{cases}$

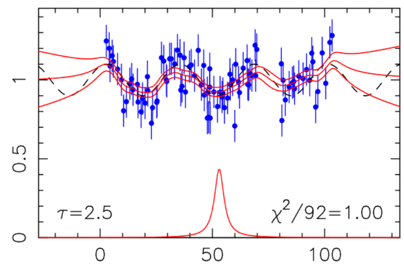
Gaussian: $= \exp\left\{-\frac{1}{2}\left(\frac{t}{\tau}\right)^2\right\}$

Lorentzian: $= \frac{1}{1+(t/\tau)^2}$




236

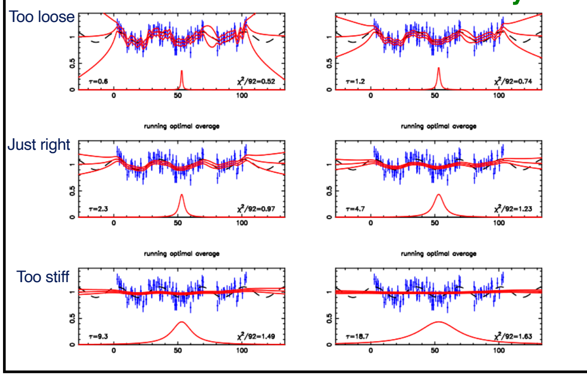
## Running Optimal Average (ROA)




Smooth curve, with error bars, running thru the data.

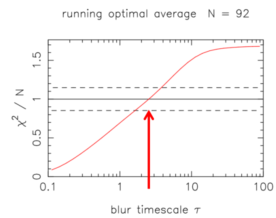
237

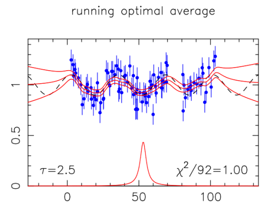
## ROA timescale τ controls flexibility




238

## Running Optimal Average (ROA)




Blur timescale $\tau$ can be chosen to make $\chi^2/N \sim 1$.

Can also define the effective number of parameters, and minimise BIC.

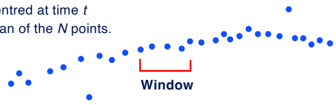
Interpolates across gaps.
Extrapolates past ends.
Averages appropriately.
Error bars provided.
(Almost) model-free.

239

## Median Filter and Sigma-Clip

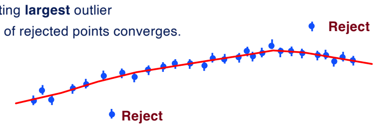
**Median filter**:

- window of *N* points centred at time *t*
- medfilt( *t* ) is the median of the *N* points.




**Sigma-clip**:

- Fit all points by minimising $\chi^2$
- 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.




**240**

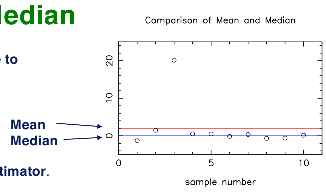
## Mean vs Median

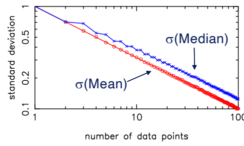
- The median is **less sensitive to outliers** than the mean.
- The median is **unbiased** but **not a minimum-variance estimator**.
- Note how **standard deviation of the median** varies with sample size *N* in comparison to **standard deviation of the mean.**

$$\mathrm{Var}[\bar{X}] = \sigma^2/N$$

$$\mathrm{Var}[X_{\mathrm{med}}] \to (\pi/2)\,(\sigma^2/N)$$

**Variance of the Median exceeds the Variance of the Mean by a factor** ***π / 2 = 1.57*** **( for large** ***N*** **).**




**241**

## Var[Median] = (π/2) Var[Mean]

$N$ gaussian random numbers:

$$\langle X_i \rangle = \mu \quad \mathrm{Var}[X_i] = \sigma^2 \quad i = 1 \ldots N$$

$$f(x) = F'(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

$P$ = fraction of positive errors :

$$p_i = \begin{cases} 1 & X_i > \mu \\ 0 & X_i < \mu \end{cases} \quad \langle p_i \rangle = \frac{1}{2} \quad \sigma^2(p_i) = \frac{1}{4}$$

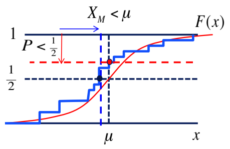
$$P \equiv \frac{1}{N}\sum_{i=1}^{N} p_i \qquad \langle P \rangle = \frac{1}{2} \quad \sigma^2(P) = \frac{1}{4N}$$

$$\text{Median: } X_M - \mu \approx \frac{P - \langle P \rangle}{F'(\mu)} = \frac{P - \frac{1}{2}}{f(\mu)}$$

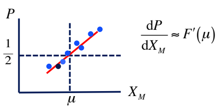
$$\frac{\partial X_M}{\partial P} = \frac{1}{f(\mu)} = (2\pi\sigma^2)^{1/2}$$

$$\sigma^2(X_M) = \sigma^2(P)\left|\frac{\partial X_M}{\partial P}\right|^2 = \frac{1}{4N(f(\mu))^2} = \frac{2\pi\sigma^2}{4N} = \frac{\pi}{2}\frac{\sigma^2}{N}$$

$$\sigma^2(\bar{X}) = \frac{\sigma^2}{N}$$




***P co-varies with $X_M$ :***




**Variance of the Median is larger by a factor** ***π / 2 = 1.57*** **( for large** ***N*** **) than the Variance of the Mean.**

**242**

## Var[Median] = (π/2) Var[Mean]

$N$ gaussian random numbers:

$$\langle X_i \rangle = \mu \quad \mathrm{Var}[X_i] = \sigma^2 \quad i = 1 \ldots N$$

$$f(x) = F'(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

$P$ = fraction of positive errors :

$$p_i = \begin{cases} 1 & X_i > \mu \\ 0 & X_i < \mu \end{cases} \quad \langle p_i \rangle = \frac{1}{2} \quad \sigma^2(p_i) = \frac{1}{4}$$

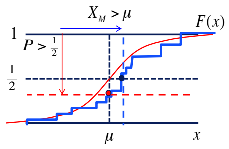
$$P \equiv \frac{1}{N}\sum_{i=1}^{N} p_i \qquad \langle P \rangle = \frac{1}{2} \quad \sigma^2(P) = \frac{1}{4N}$$

$$\text{Median: } X_M - \mu \approx \frac{P - \langle P \rangle}{F'(\mu)} = \frac{P - \frac{1}{2}}{f(\mu)}$$

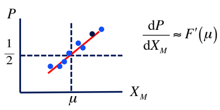
$$\frac{\partial X_M}{\partial P} = \frac{1}{f(\mu)} = (2\pi\sigma^2)^{1/2}$$

$$\sigma^2(X_M) = \sigma^2(P)\left|\frac{\partial X_M}{\partial P}\right|^2 = \frac{1}{4N(f(\mu))^2} = \frac{2\pi\sigma^2}{4N} = \frac{\pi}{2}\frac{\sigma^2}{N}$$

$$\sigma^2(\bar{X}) = \frac{\sigma^2}{N}$$




***P co-varies with $X_M$ :***




**Variance of the Median is larger by a factor** ***π / 2 = 1.57*** **( for large** ***N*** **) than the Variance of the Mean.**

**243**

## Fini -- ADA 12

**244**