UNFOLDING / DECONVOLUTION

- Direct and inverse process

Suppose we want to measure some variable y using a detector that has a response function that "deforms" or "transforms" in some way the variable y into x

If $f(x)$ is the PDF of x, and $r(x,y)$ is the response function (also called kernel function), the PDF of y will be:

$$g(x) = \int r(x,y)\, f(y)\, dy \; [+ b(x)]$$

→ Fredholm integral

↳ possible noise or bkg

↓ convolution of f and r

Typically, r induces a "smearing" which broadens the peak structures present in the original distributions.

Examples: – Energy resolution for an X-ray detector
– Space resolution induced by lens aberrations on telescope image

Possible "distortions" are:

→ Statistical fluctuations (e.g. due to Poisson statistics) if the data are binned
→ Smearing due to limited resolution = migration of counts between bins
→ Reduced efficiency
→ Non-linear detector response

See next page

• Convolution and Fourier Transform

In most cases, r depends only on the difference of $x-y$, so it depends just on one variable: $r(x,y) = r(x-y)$.

In general, we will denote the convolution as $\otimes$.

Let's take the Fourier Transform of g:

$$\hat{g}(k) = \int_{-\infty}^{+\infty} g(y)\, e^{-iky}\, dy$$

Conversely: $g(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{g}(k)\, e^{iky}\, dk$

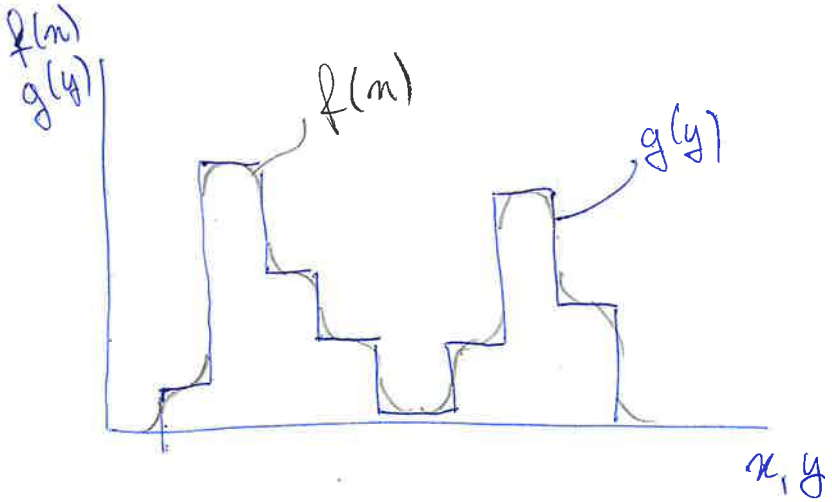
One can demonstrate that: $\widehat{g \otimes r} = \hat{g} \cdot \hat{r}$

Conversely: $\widehat{g \cdot r} = \hat{g} \otimes \hat{r}$

→ Example: Gaussian smearing

Take $r(x-y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right)$

$$\hat{r}(k) = e^{-iky} \exp\left(-\frac{\sigma^2 k^2}{2}\right)$$



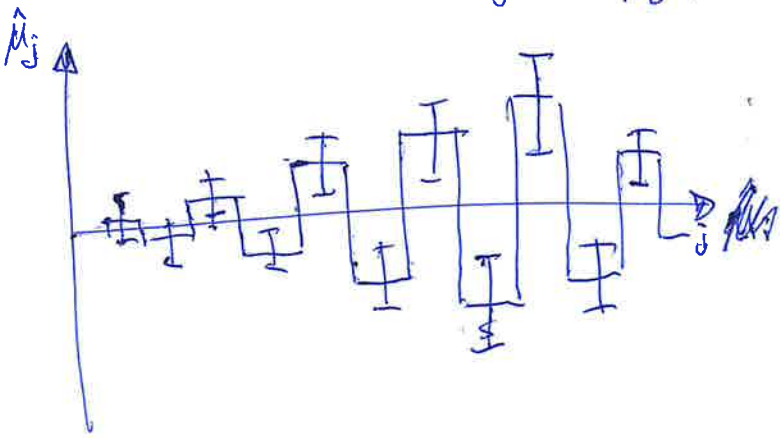
The covariance matrix of the expected bin entries $\hat{\mu}_i$ is:

$$U = R^{-1} V (R^{-1})^T \quad \text{where} \quad V = \text{covariance matrix of data } n_i$$

$$V_{ij} = \delta_{ij} v_j \quad \text{because of Poisson process.}$$

The issue is that each bin $\hat{\mu}_j$ is strongly anti-correlated with its neighbors $\hat{\mu}_{j\pm1}$ because of the off-diagonal entries of U. This is caused by statistical fluctuations in the data reflecting into strong anti-correlations rather than into the uncertainty on $\hat{\mu}_j$, as one would expect.

One gets something like:



- Bin-by-bin correction

One possible solution is to correct n_i using the ratio of expected values before and after unfolding:

$$\hat{\mu}_i = \frac{\mu_i^{est}}{\nu_i^{est}} (n_i - b_i) \quad \text{where: } \mu_i^{est}, \nu_i^{est} \text{ are obtained from simulations.}$$

ν_i^{est} has no background

Problem: This approach introduces a bias that drives $\hat{\mu}_i$ towards μ_i^{exp}:

$$\langle \hat{\mu}_i \rangle - \mu_i = \left(\frac{\mu_i^{est}}{\nu_i^{est}} - \frac{\mu_i}{\nu_i - b_i} \right) (\nu_i - b_i)$$

In this way, we risk to hide real discrepancies!

- Tikhonov Regularization

Let's choose: $S(\vec{\mu}) = (L\vec{\mu})^T L\vec{\mu}$

where L is a $n \times k$ matrix.

→ In the simplest case, $L = \mathbb{1}$, Therefore: $S(\vec{\mu}) = \sum_j \mu_j^2$

In this way, the term $\gamma^2 S(\vec{\mu})$ dumps the cases with large deviations of μ_j from zero.

→ Another choice is:

$$L = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & & \vdots \\ 0 & 1 & -2 & \ddots & 1 & 0 \\ \vdots & & & 1 & -2 & 1 \\ 0 & \cdots & 0 & & 1 & -1 \end{pmatrix}$$

In fact, if we have a function approximated by a histogram h with bin size δ, the first two derivatives are:

$$h_i' = \frac{h_i - h_{i-1}}{\delta}$$

$$h_i'' = \frac{h_{i-1} - 2h_i + h_{i+1}}{\delta^2}$$

So L is an approximation of the 2nd derivative of $\vec{\mu}$ as approximation of $g(y)$.

$S(\vec{\mu})$ would correspond to:

$$S \approx \int \left(\frac{d^2 g(y)}{dy^2} \right)^2 dy$$

In this way, large 2nd derivatives that induce high-frequency oscillations are dumped.

The experimental distribution $\vec{n} = (n_1, \dots, n_N)$ is the number of occurences of each effect E_i: $n_i = n(E_i)$

The expected number of events $\mu_j^{(1)}$ assigned to each cause C_j is:

$$\mu_j^{(1)} = \hat{n}(C_j) = \sum_{i=1}^{N} n(E_i) \frac{P(C_j|E_i)}{\varepsilon_j} = \mu_j^{(0)} \sum_{i=1}^{N} \frac{R_{ij}}{\varepsilon_j} \frac{n_i}{\sum_{k=1}^{N} R_{ik} \mu_k^{(0)}}$$

- Other methods

This was just a brief introduction.

Other possible values are:

→ Singular Value Decomposition (SVD)

→ Shape constrained unfolding

→ Fully Bayesian unfolding