Unfolding for counting experiments

Marcin Chrząszcz[1,2], Nicola Serra[1]

[1] University of Zurich,
[2] Institute of Nuclear Physics, Krakow

June 30, 2014

Reminder 1 - Constructing Matrix unfolding

- We don't know explicate
- I have proven some time ago that the matrix exist

$$\epsilon(\cos\theta_k, \cos\theta_l, \phi) \tag{1}$$

- I have proven some time ago that the matrix exist
- Now a systemic way to produce it.
- Let's use PHSP MC.
- Moments for PHSP MC are:
  $v_{gen}^T = (2/3, 0, 0, 0, 0, 0, 0, 0)$
- After reconstruction we get(full q^2 range): $v_{rec}^T =$
  $(0.7069, 0.0077, -0.00236466, 0.0005, 0.0007, 0.0011, 0.0011, -0.0012)$

University of Zurich[UZH]

Reminder 2 - Constructing Matrix unfolding

- We got first column of the unfolding matrix $(\frac{3}{2} v_{gen})$.

$$\begin{pmatrix} 1.06 & \cdots & a_{1,8} \\ 0.01157 & \cdots & a_{2,8} \\ -0.003547 & \ddots & \vdots \\ 0.0007841 & \ddots & \vdots \\ 0.001126 & \ddots & \vdots \\ 0.001766 & \ddots & \vdots \\ 0.001664 & \ddots & \vdots \\ -0.001937 & \cdots & a_{8,8} \end{pmatrix}$$

- How about the others?
- We can reweight accordingly to f_x.

Reminder 3 - Constructing Matrix unfolding

- To get S_3 each event i^{th} has has weight $f_{S_3}(\cos\theta_{k_i}, \cos\theta_{l_i}, \phi_i)$
- One can calculate on MC the reweighed moments in PHPS:

$$\int PDF * f_{S_3} = \frac{32}{225} \tag{2}$$

- Our base vector now is:$v^T_{gen} = (0, \frac{32}{225}, 0, 0, 0, 0, 0, 0)$
- So lets see what do we get as reconstructed vector(after multiplying by $\frac{225}{32}$. $v^T_{rec} =$ $(0.042, 1.105, -0.005, 0.003, -0.0023, -0.005, -0.005, -0.006)$
- Please notice that weights are negative, but this is not a problem for the mean.
- Also we are avoiding the negative PDF problem :)

University of Zurich[UZH]

- Now the matrix looks like:

$$\begin{pmatrix} 1.06 & 0.042 & \cdots & a_{1,8} \\ 0.01157 & 1.105 & \cdots & a_{2,8} \\ -0.003547 & -0.005 & \ddots & \vdots \\ 0.0007841 & -0.005 & \ddots & \vdots \\ 0.001126 & 0.003 & \ddots & \vdots \\ 0.001766 & -0.0023 & \ddots & \vdots \\ 0.001664 & -0.005 & \ddots & \vdots \\ -0.001937 & -0.006 & \cdots & a_{8,8} \end{pmatrix}$$

- The others go in the same way.
- Repenting this exercise from 1^{st} year algebra we can get the full matrix

For now:

- We have proven that there has to exists unfolding matrix.
- Shown how to construct transformation matrix: $Gen \longrightarrow Reco$.
- Inverting it we can have transformation matrix of $Reco \longrightarrow Gen$.
- For details: LINK

What is missing?

1. ERROR!

University of Zurich[UZH]

How to?

- So lets say that transformation matrix: $Gen \rightarrow Reco$ is $\epsilon_{i,j}$.
- Each element has an error: $\delta\epsilon_{i,j}$.
- Then we can calculate the matrix: $\epsilon^{-1}_{i,j}$ (assuming it exists).
- The million dollar question is what is the error on inverted matrix?

Answer to 1M dolar quesion

- One can toy it.
- But toying is good for kids and Frequentist.

Answer to 1M dolar quesion

- One can toy it.
- But toying is good for kids and Frequentist.
- Solution comes from τ physics :) hep-ex/9909031
- One can derive(prove in the paper) the general equation:

$$\delta\epsilon^{-1}_{\alpha\ \beta} = [\epsilon^{-1}]^2_{\alpha i}[\delta\epsilon]^2_{ij}[\epsilon^{-1}]^2_{j\beta} \tag{3}$$

$$A_{reco\rightarrow gen} = \begin{pmatrix} 0.9519 & -0.02665 & -0.01432 & 0.002356 & 0.02539 & 0.009878 & -0.01551 & -0.01874 \\ -0.006272 & 0.8122 & -0.00351 & -0.00719 & 0.003585 & 6.784e-05 & 0.02445 & 0.008515 \\ -0.005315 & -0.003716 & 1.048 & 0.01242 & 0.01209 & -0.01478 & -0.001956 & 0.01429 \\ 0.003237 & -0.007177 & 0.01533 & 0.9184 & -0.007548 & -0.0009818 & -0.01874 & 0.009407 \\ 0.01002 & 0.004084 & 0.01391 & -0.006509 & 1.194 & -0.006516 & 0.001536 & -0.02882 \\ 0.002695 & -0.001042 & -0.01721 & -0.001842 & -0.005643 & 0.9264 & 0.02106 & 0.006755 \\ -0.004736 & 0.02346 & -0.002335 & -0.01446 & 0.001169 & 0.01697 & 1.072 & -0.003191 \\ -0.004157 & 0.007576 & 0.01377 & 0.008058 & -0.02219 & 0.005354 & -0.0008608 & 0.8304 \end{pmatrix}$$

$$\delta A_{reco\rightarrow gen} = \begin{pmatrix} 0.005202 & 0.01911 & 0.03258 & 0.02103 & 0.02252 & 0.02145 & 0.03366 & 0.01948 \\ 0.006648 & 0.04654 & 0.03227 & 0.02451 & 0.03602 & 0.02464 & 0.03298 & 0.03397 \\ 0.007557 & 0.03197 & 0.07845 & 0.04272 & 0.04744 & 0.03057 & 0.05698 & 0.03287 \\ 0.007902 & 0.03885 & 0.0678 & 0.04839 & 0.0384 & 0.03464 & 0.04925 & 0.03989 \\ 0.009015 & 0.04122 & 0.06374 & 0.03254 & 0.07349 & 0.03269 & 0.0649 & 0.04202 \\ 0.007939 & 0.0389 & 0.04793 & 0.03433 & 0.03828 & 0.04937 & 0.06985 & 0.04023 \\ 0.007651 & 0.03234 & 0.05611 & 0.03062 & 0.04776 & 0.04388 & 0.08157 & 0.03342 \\ 0.006719 & 0.03345 & 0.03868 & 0.02953 & 0.03633 & 0.03002 & 0.03989 & 0.04827 \end{pmatrix}$$

What did go wrong?

- The errors are $2-3\%$, which is very worrying.
- WG got very worried what is going on with the errors :(
- Started debugging.
- After sleeping with the problem found a stupid:

for(int i=0;i $<$ entries/10;++i)

- Ok, I am an idiot, and used 10% of statistics.

What did go wrong 2 ?

- The errors are tricky. When you re-weight you have negative weights.
- So I change the normal error

$$\hat{\sigma}^2 = \frac{\sum w_i}{(\sum w_i)^2 - \sum w_i^2} \sum w_i (x_i - \hat{\mu})^2 \tag{4}$$

- to:

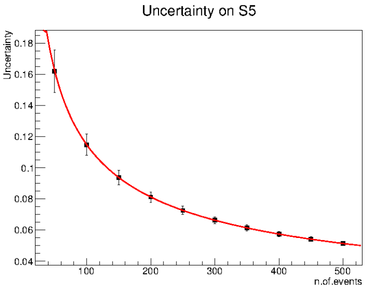
$$\hat{\sigma}^2 = \frac{\sum |w_i|}{(\sum |w_i|)^2 - \sum w_i^2} \sum |w_i| (x_i - \hat{\mu})^2 \tag{5}$$

- And this I am not 100% sure if I is ok =(

University of Zurich[UZH]

What did go wrong 3 ?

- There is a hack of this method:
- "You can cheat on your gf, you can cheat on tax, but you can't cheat on $\sqrt{n}$ "[1].



- We can use this:
- Divide the MC in 10. Then calculate the variance of each matrix element. And divide/multiply by $\sqrt{10}$ and see if the errors are ok.

[1] All rights reserved!

What did go wrong 3 ?

OLD (can be wrong):

$$\delta A_{gen \rightarrow reco} = \begin{pmatrix} 0.005477 & 0.02348 & 0.03125 & 0.02305 & 0.01871 & 0.02307 & 0.03124 & 0.02339 \\ 0.008142 & 0.06734 & 0.03621 & 0.03126 & 0.0352 & 0.03131 & 0.03624 & 0.04767 \\ 0.007168 & 0.0359 & 0.06856 & 0.0423 & 0.03619 & 0.02995 & 0.04856 & 0.03585 \\ 0.008573 & 0.04966 & 0.06736 & 0.05471 & 0.03332 & 0.03886 & 0.04784 & 0.04973 \\ 0.007599 & 0.04063 & 0.04926 & 0.02847 & 0.04998 & 0.02841 & 0.04923 & 0.04059 \\ 0.008582 & 0.04977 & 0.04768 & 0.03878 & 0.03323 & 0.05499 & 0.0676 & 0.04974 \\ 0.007136 & 0.03571 & 0.04833 & 0.02987 & 0.036 & 0.04225 & 0.06843 & 0.0358 \\ 0.008162 & 0.04782 & 0.04294 & 0.03731 & 0.03527 & 0.03738 & 0.04306 & 0.06736 \end{pmatrix}$$

New:

$$\delta A_{gen \rightarrow reco} =$$

$$\begin{pmatrix} 0.006659 & 0.0299 & 0.02207 & 0.01802 & 0.02657 & 0.02196 & 0.02851 & 0.02507 \\ 0.00708 & 0.02046 & 0.007998 & 0.0133 & 0.008828 & 0.01236 & 0.01505 & 0.0149 \\ 0.008469 & 0.00845 & 0.01806 & 0.01442 & 0.009856 & 0.008895 & 0.01389 & 0.01155 \\ 0.008938 & 0.01569 & 0.01798 & 0.01801 & 0.009195 & 0.01097 & 0.01108 & 0.02068 \\ 0.007867 & 0.0109 & 0.01248 & 0.0088 & 0.01104 & 0.0114 & 0.01256 & 0.01097 \\ 0.008078 & 0.01582 & 0.01117 & 0.01093 & 0.01135 & 0.01215 & 0.02122 & 0.01774 \\ 0.008368 & 0.01521 & 0.01391 & 0.008972 & 0.009797 & 0.01702 & 0.0147 & 0.01086 \\ 0.005745 & 0.01561 & 0.0114 & 0.01649 & 0.008631 & 0.01373 & 0.01051 & 0.01792 \end{pmatrix}$$

- I really fu.. this thing ...
- No coding after 3 am form now!