Kyle Cranmer
Draft 2, April 1, 2011

HistFactory Likelihood

1 The Likelihood Template

The parametrized probability density function constructed by the HistFactory is of a concrete form, but sufficiently flexible to describe many analyses based on template histograms. In general, the HistFactory produces probability density functions of the from

$$\mathcal{P}(n_m, m_i|\mu, \alpha_i) = \prod_{m\in\mathrm{bins}} \mathrm{Pois}(n_m|\nu_m) \cdot G(L_0|L, \Delta_L) \prod_{i\in\mathrm{Syst}}^{5} N(m_i|\alpha_i, 1) \tag{1}$$

where m is an index over the bins of the template histograms, i is an index over systematic effects, ν_m is the expected number of events in bin m given by

$$\nu_m = \sum_{j\in\mathrm{Samp}} LF_j\eta_j(\boldsymbol{\alpha})\, \sigma_{jm}(\boldsymbol{\alpha}), \tag{2}$$

where F_j is a product of unconstrained normalization factors for sample j

$$F_j = \prod_{n\in\mathrm{NormFact}_j} f_n \tag{3}$$

that typically include the parameter of interest (eg. $\mu = \sigma/\sigma_{SM}$). The term $\eta_j(\boldsymbol{\alpha})$ parametrizes relative changes in the overall normalization, and $\sigma_{jm}(\boldsymbol{\alpha})$ contains the nominal normalization and parametrizes uncertainties in the shape of the distribution of the discriminating variable. Here j is an index of contributions from different processes with $j = 1$ being the signal process. The nuisance parameters α_i are associated to the source of the systematic effect (e.g. the muon momentum resolution uncertainty), while $\eta_j(\boldsymbol{\alpha})$ and $\sigma_{jm}(\boldsymbol{\alpha})$ represent the effect of that uncertainty. The α_i are scaled so that $\alpha_i = 0$ corresponds to the nominal expectation and $\alpha_i = \pm 1$ correspond to the $\pm 1\sigma$ variations of the source, thus $N(\alpha_i)$ is the standard normal distribution. The effect of these variations about the nominal predictions $\eta_j(0) = 1$ and σ^0_{jm} is quantified by dedicated studies that provide $\eta^{\pm}_{ij}$ and $\sigma^{\pm}_{ijm}$, which are then used to form

$$\eta_j(\boldsymbol{\alpha}) = \prod_{i\in\mathrm{Syst}} I(\alpha_i; \eta^+_{ij},\ \eta^-_{ij}) \tag{4}$$

and

$$\sigma_{jm}(\boldsymbol{\alpha}) = \sigma^0_{jm} \prod_{i\in\mathrm{Syst}} I(\alpha_i; \sigma^+_{ijm}/\sigma^0_{jm},\ \sigma^-_{ijm}/\sigma^0_{jm}) \tag{5}$$

with

$$I(\alpha; I^+, I^-) = \begin{cases} 1 + \alpha(I^+ - 1) & \text{if } \alpha > 0 \\ 1 & \text{if } \alpha = 0 \\ 1 - \alpha(I^- - 1) & \text{if } \alpha < 0 \end{cases} \tag{6}$$

enabling piece-wise linear interpolation in the case of asymmetric response to the source of systematic. In the next version of HistFactory, a quadratic interpolation will be available, and may become the default, as it avoids discontinuities in the second derivative that often confuse MINUIT and HESSE.

2 Implementation with HistFactory

Note, when using the `HistFactory` the production modes l and backgrounds j correspond to a single XML `Sample` element. The `HistoName` attribute inside each sample element specifies the histogram with the σ^0_{ijm}. The index $j =' J$' is set by the `Name` attribute of the `Sample` element (eg. `<Sample Name='J'>`). Between the open `<Sample>` and close `</Sample>` one can add

- An `OverallSys` element where the `Name='I'` attribute identifies which α_I is the source of the systematic and implies that the Gaussian constraint $N(m_I|\alpha_I, 1)$ is present. The `High` attribute corresponds to η^+_{IJ}, eg when the source of the systematic is at $+1\sigma$ and $\alpha_I = 1$. Similarly, the `Low` attribute corresponds to η^-_{IJ}, eg when the source of the systematic is at -1σ and $\alpha_I = -1$. The nominal value is $\eta^0_{IJ} = 1$ for the ovreall systematics. The distinction between the sign of the source α and the effect η allows one to have anti-correlated systematics. The `HistFactory` is able to deal with asymmetric uncertainties as well, by using a piece-wise linear interpolation for the $\alpha_I > 0$ and $\alpha_I < 0$ regions.

- A `NormFactor` element is used to introduce an overall constant factor into the expected number of events. In the example below, the term $\mu = \sigma/\sigma_{SM}$ corresponds to the line `<NormFactor Name='SigXsecOverSM'>`. In this case, the histograms were normalized to unity, so additional `NormFactor` elements were used to give the overall cross-sections σ_J.

- A `HistoSys` element is used to introduce shape systematics and the `HistoNameHigh` and `HistoNameLow` attributes have the variational histograms σ^+_{ijm} and σ^-_{ijm} corresponding to $\alpha_i = +1$ and $\alpha = -1$, respectively.

Below is an example XML file for the electron channel.

```
<!DOCTYPE Channel  SYSTEM 'HistFactorySchema.dtd'>
  <Channel Name="channelEle" InputFile="./data/central_Ele_5jet_inc_35invpb.root" HistoName="" >
    <!--<Data HistoName="data" HistoPath="" />-->
    <Sample Name="bbAtautau120" HistoPath="" NormalizeByTheory="True" HistoName="bbAtautau120All">
      <OverallSys Name="JES" High="1.05" Low="0.95"/>
      <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
      <OverallSys Name="bbAtautau" High="1.15" Low="0.85"/>
      <NormFactor Name="NEle_bbAtautau120" Val=".83202" Low=".83202" High=".83202" Const="True" />
      <NormFactor Name="SigXsecOverSM" Val="0" Low="-10." High="30." Const="True" />
    </Sample>
    <Sample Name="Atautau120" HistoPath="" NormalizeByTheory="True" HistoName="Atautau120All">
      <OverallSys Name="JES" High="1.05" Low="0.95"/>
      <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
      <OverallSys Name="Atautau" High="1.15" Low="0.85"/>
      <NormFactor Name="NEle_Atautau120" Val=".24224" Low=".24224" High=".24224" Const="True" />
      <NormFactor Name="SigXsecOverSM" Val="0" Low="-10." High="30." Const="True" />
    </Sample>
    <Sample Name="bbAtautau130" HistoPath="" NormalizeByTheory="True" HistoName="bbAtautau130All">
      <OverallSys Name="JES" High="1.05" Low="0.95"/>
      <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
      <OverallSys Name="bbAtautau" High="1.15" Low="0.85"/>
      <NormFactor Name="NEle_bbAtautau130" Val=".01767" Low=".01767" High=".01767" Const="True" />
      <NormFactor Name="SigXsecOverSM" Val="0" Low="-10." High="30." Const="True" />
    </Sample>
    <Sample Name="Atautau130" HistoPath="" NormalizeByTheory="True" HistoName="Atautau130All">
      <OverallSys Name="JES" High="1.05" Low="0.95"/>
      <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
      <OverallSys Name="Atautau" High="1.15" Low="0.85"/>
      <NormFactor Name="NEle_Atautau130" Val=".02441" Low=".02441" High=".02441" Const="True" />
      <NormFactor Name="SigXsecOverSM" Val="0" Low="-10." High="30." Const="True" />
    </Sample>
    <Sample Name="Ztautau" HistoPath="" NormalizeByTheory="True" HistoName="ZtautauAll">
      <OverallSys Name="JES" High="1.05" Low="0.95"/>
      <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
      <OverallSys Name="Alpgen" High="1.131" Low="0.869"/>
      <OverallSys Name="Ztautau" High="1.15" Low="0.85"/>
      <NormFactor Name="NEle_Ztautau" Val="1.26818" Low="1.26818" High="1.26818" Const="True" />
    </Sample>
   <Sample Name="AddOn" HistoPath="" NormalizeByTheory="False" HistoName="AddOnAll">
      <OverallSys Name="AddOn" High="1.173" Low=".827"/>
      <NormFactor Name="NEle_AddOn" Val=".88267" Low=".88267" High=".88267" Const="True" />
    </Sample>
```

```
  <Sample Name="SameSign" HistoPath="" NormalizeByTheory="False" HistoName="SameSignAll">
    <OverallSys Name="SameSign" High="1.06828" Low=".93172"/>
    <NormFactor Name="NEle_SameSign" Val="4.00568" Low="4.00568" High="4.00568" Const="True" />
  </Sample>
  <Sample Name="Others" HistoPath="" NormalizeByTheory="True" HistoName="OthersAll">
    <OverallSys Name="JES" High="1.05" Low="0.95"/>
    <OverallSys Name="EVTEFF" High="1.122" Low="0.878"/>
    <OverallSys Name="QFAC" High="1.03" Low="0.97"/>
    <OverallSys Name="Alpgen" High="1.131" Low="0.869"/>
    <OverallSys Name="Others" High="1.15" Low="0.85"/>
    <NormFactor Name="NEle_Others" Val=".17949" Low=".17949" High=".17949" Const="True" />
  </Sample>
</Channel>
```

One can convert this Gaussian constraints into a Poisson/Gamma systematic by adding lines like

```
<ConstraintTerm Type="Gamma" RelativeUncertainty="0.1">JES</ConstraintTerm>
```

to the `Measurement` element. For example:

```
<Measurement Name="AllSYS" Lumi="35.2" LumiRelErr="0.11" BinLow="0" BinHigh="20" Mode="comb"  ExportOnly="True">
  <POI>SigXsecOverSM</POI>
  <ParamSetting Const="True">NEle_AddOn,NEle_Atautau120,NEle_Atautau130,NEle_Others,
    NEle_SameSign,NEle_Ztautau,NEle_bbAtautau120,NEle_bbAtautau130,NMuo_AddOn,
    NMuo_Atautau120,NMuo_Atautau130,NMuo_Others,NMuo_SameSign,NMuo_Ztautau,NMuo_bbAtautau120,
    NMuo_bbAtautau130
  </ParamSetting>
  <ConstraintTerm Type="Gamma" RelativeUncertainty="0.1">JES</ConstraintTerm>
  <!--<ConstraintTerm Type="LogNormal" RelativeUncertainty="0.1">JES</ConstraintTerm>-->
</Measurement>
```