# Extended Weighted Unbinned Maximum Likelihood Fits

Christoph Langenbruch
Heidelberg University

March 20, 2024

This writeup illustrates how to include the extended term in extended weighted unbinned maximum likelihood fits.

## 1 Recap

The asymptotically correct treatment of uncertainties in weighted unbinned maximum likelihood fits is based on a Taylor expansion of the estimating equation [1]

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda} = 0 \quad \text{with} \tag{1}$$

$$\sum_e w_e \left.\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\right|_{\hat{\lambda}} = 0$$

around the true value $\lambda_0$, which results in

$$\sum w_e \left.\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\right|_{\lambda_0} + (\hat{\lambda} - \lambda_0) \sum w_e \left.\frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right|_{\lambda_0} = 0. \tag{2}$$

Resolving to $\hat{\lambda} - \lambda_0$ and neglecting terms of $\mathcal{O}(1/N)$ gives

$$\Rightarrow \hat{\lambda} - \lambda_0 = - \sum w_e \left.\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\right|_{\lambda_0} \Big/ \left[N \cdot E\left(w \frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right)\right]. \tag{3}$$

We can then calculate $\text{Var}(\lambda) = E((\hat{\lambda} - \lambda_0)^2)$, resulting in

$$\begin{aligned}
\text{Var}(\lambda) &= \left[E\left(\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2}\right)\right]^{-1} \left[E\left(\frac{\partial \ln \mathcal{L}}{\partial \lambda}\frac{\partial \ln \mathcal{L}}{\partial \lambda}\right)\right] \left[E\left(\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2}\right)\right]^{-1} \qquad (4)\\
&= \left[N \cdot E\left(w \frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right)\right]^{-1} N \cdot E\left(w_e^2 \frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\right) \left[N \cdot E\left(w \frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right)\right]^{-1}\\
&= \left[\sum w_e \left.\frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right|_{\hat{\lambda}}\right]^{-1} \left[\sum w_e^2 \left.\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\frac{\partial \ln \mathcal{P}_{\text{tot}}}{\partial \lambda}\right|_{\hat{\lambda}}\right] \left[\sum w_e \left.\frac{\partial^2 \ln \mathcal{P}_{\text{tot}}}{\partial \lambda^2}\right|_{\hat{\lambda}}\right]^{-1}
\end{aligned}$$

and analogously for multiple dimensions. The last line here is the sample estimate.

## 2 Extended fit using $N_{\rm sig}$ and $N_{\rm bkg}$

The $\ln \mathcal{L}$ in this case is given by

$$\begin{aligned}
\ln \mathcal{L} &= \sum^{\hat{N}} w_e \ln\left( \frac{N_{\rm sig}}{N_{\rm sig}+N_{\rm bkg}} \mathcal{P}_{\rm sig} + \frac{N_{\rm bkg}}{N_{\rm sig}+N_{\rm bkg}} \mathcal{P}_{\rm bkg} \right) + \hat{N}_{\rm tot} \ln(N_{\rm sig}+N_{\rm bkg}) - (N_{\rm sig}+N_{\rm bkg}) \\
&= \sum^{\hat{N}} w_e \ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big) - \sum^{\hat{N}} w_e \ln\big(N_{\rm sig}+N_{\rm bkg}\big) + \hat{N}_{\rm tot} \ln(N_{\rm sig}+N_{\rm bkg}) - (N_{\rm sig}+N_{\rm bkg}) \\
&= \sum^{\hat{N}} w_e \ln(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}) - (N_{\rm sig}+N_{\rm bkg}) \qquad (5)
\end{aligned}$$

Here we have the following first derivatives:

$$\begin{aligned}
\frac{\partial \ln \mathcal{L}}{\partial \lambda} &= \sum w_e \frac{\partial}{\partial \lambda} \ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big) \qquad (6) \\
\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}} &= \sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - 1 \\
\frac{\partial \ln \mathcal{L}}{\partial N_{\rm bkg}} &= \sum w_e \frac{\mathcal{P}_{\rm bkg}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - 1
\end{aligned}$$

as well as these second derivatives:

$$\begin{aligned}
\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2} &= \sum w_e \frac{\partial^2}{\partial \lambda^2} \ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big) \qquad (7) \\
\frac{\partial^2 \ln \mathcal{L}}{\partial N_{\rm sig}^2} &= -\sum w_e \frac{\mathcal{P}_{\rm sig}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2} \\
\frac{\partial^2 \ln \mathcal{L}}{\partial N_{\rm bkg}^2} &= -\sum w_e \frac{\mathcal{P}_{\rm bkg}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2} \\
\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda \partial N_{\rm sig}} &= \sum w_e \frac{\partial}{\partial \lambda} \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} \\
\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda \partial N_{\rm bkg}} &= \sum w_e \frac{\partial}{\partial \lambda} \frac{\mathcal{P}_{\rm bkg}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} \\
\frac{\partial^2 \ln \mathcal{L}}{\partial N_{\rm sig} \partial N_{\rm bkg}} &= -\sum w_e \frac{\mathcal{P}_{\rm sig}\mathcal{P}_{\rm bkg}}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}.
\end{aligned}$$

More complicated is the numerator, which requires the (cross-) expectation values. For these, we find

$$
\begin{aligned}
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial \lambda}\frac{\partial \ln \mathcal{L}}{\partial \lambda}\bigg) &= E\bigg(\sum w_e^2 \Big[\frac{\partial}{\partial \lambda}\ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big)\Big]^2\bigg) \qquad (8)\\
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\bigg) &= E\bigg(\Big[\sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - 1\Big]\Big[\sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - 1\Big]\bigg)\\
&= E\bigg(\Big(\sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}}\Big)^2 - 2\sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} + 1\bigg)\\
&= E\bigg(\hat{N}(\hat{N}-1)E^2\Big(w\frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}}\Big) + \hat{N}E\Big(w^2\frac{\mathcal{P}_{\rm sig}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}\Big)\\
&\qquad - 2\hat{N}E\Big(w\frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}}\Big) + 1\bigg)\\
&= \frac{N^2}{N^2} + NE\Big(w^2\frac{\mathcal{P}_{\rm sig}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}\Big) - 2\frac{N}{N} + 1\\
&= E\bigg(\sum w_e^2\frac{\mathcal{P}_{\rm sig}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}\bigg)\\
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial N_{\rm bkg}}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm bkg}}\bigg) &= \sum w_e^2\frac{\mathcal{P}_{\rm bkg}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}\\
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial \lambda}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\bigg) &= E\bigg(\Big[\sum w_e \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - 1\Big]\sum w_e\frac{\partial}{\partial \lambda}\ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big)\bigg)\\
&= E\bigg(\sum w_e^2 \frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}}\frac{\partial}{\partial \lambda}\ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big)\bigg)\\
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial \lambda}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm bkg}}\bigg) &= E\bigg(\sum w_e^2 \frac{\mathcal{P}_{\rm bkg}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}}\frac{\partial}{\partial \lambda}\ln\big(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}\big)\bigg)\\
E\bigg(\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm bkg}}\bigg) &= E\bigg(\sum w_e^2\frac{\mathcal{P}_{\rm sig}\mathcal{P}_{\rm bkg}}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}\bigg)
\end{aligned}
$$

These terms, together with the second derivatives in Eq. 7 allow the determination of the full asymptotic covariance. The expectation values can alternatively be calculated using the logarithmic total probability

$$
\begin{aligned}
\frac{\partial \ln \mathcal{P}_{\rm tot}}{\partial N_{\rm sig}} &= \frac{\partial}{\partial N_{\rm sig}}\ln\Big[\frac{N_{\rm sig}}{N_{\rm sig}+N_{\rm bkg}}\mathcal{P}_{\rm sig} + \frac{N_{\rm bkg}}{N_{\rm sig}+N_{\rm bkg}}\mathcal{P}_{\rm bkg}\Big] \qquad (9)\\
&= \frac{\partial}{\partial N_{\rm sig}}\big[\ln(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}) - \ln(N_{\rm sig}+N_{\rm bkg})\big]\\
&= \Big[\frac{\mathcal{P}_{\rm sig}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - \frac{\partial}{\partial N_{\rm sig}}\ln(N_{\rm sig}+N_{\rm bkg})\Big]\\
\frac{\partial \ln \mathcal{P}_{\rm tot}}{\partial N_{\rm bkg}} &= \Big[\frac{\mathcal{P}_{\rm bkg}}{N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg}} - \frac{\partial}{\partial N_{\rm bkg}}\ln(N_{\rm sig}+N_{\rm bkg})\Big].
\end{aligned}
$$

Comparing with the expectation values in Eq. 8 we see that we simply need to add the terms $\partial \ln(N_{\rm sig}+N_{\rm bkg})/\partial N_{\rm sig}$ and $\partial \ln(N_{\rm sig}+N_{\rm bkg})/\partial N_{\rm bkg}$ (note $\partial \ln(N_{\rm sig} +$

$N_{\rm bkg})/\partial\lambda = 0)$ to get the needed factors, *e. g.*

$$E\left(\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\frac{\partial \ln \mathcal{L}}{\partial N_{\rm sig}}\right) = \sum w_e^2 \left[\frac{\partial \ln \mathcal{P}_{\rm tot}}{\partial N_{\rm sig}} + \frac{\partial \ln (N_{\rm sig} + N_{\rm bkg})}{\partial N_{\rm sig}}\right]\left[\frac{\partial \ln \mathcal{P}_{\rm tot}}{\partial N_{\rm sig}} + \frac{\partial \ln (N_{\rm sig} + N_{\rm bkg})}{\partial N_{\rm sig}}\right] \quad (10)$$

$$= \sum w_e^2 \frac{\mathcal{P}_{\rm sig}^2}{(N_{\rm sig}\mathcal{P}_{\rm sig} + N_{\rm bkg}\mathcal{P}_{\rm bkg})^2}$$

## References


[1] C. Langenbruch, Eur. Phys. J. C **82** (2022) no.5, 393 doi:10.1140/epjc/s10052-022-10254-8 [arXiv:1911.01303 [physics.data-an]].