**DRAFT 0.1**

Glen Cowan
21 June, 2014

# Error analysis with weighted events

In an experiment we may count a certain number of events $n$, which is treated as following a Poisson distribution. The standard procedure is to use $n$ as an estimate of the mean of the Poisson distribution. The value $n$ also gives an estimate for its variance, or equivalently $\sqrt{n}$ is used for the standard deviation or "statistical error". The rationale behind this procedure is reviewed in Sec. 1. If the events are assigned weights, then we find that sum of the weights plays the role of $n$ and the sum of the weights squared can be used to estimate the variance. These results were derived in a draft note [1] and for completeness are reviewed here in Sec. 2. In Sec. 3 the results are generalized to the case where one has a set of $n$ events with two possible sets of weghts, which are correlated.

## 1 Estimate for mean of a Poisson variable

Suppose we select $n$ events and we want to assign a statistical error. Often we regard $n$ as following a Poisson distribution with expectation (mean) value $E[n] = \nu$. The likelihood function is therefore

$$L(\nu) = \frac{\nu^n}{n!} e^{-\nu} \; . \tag{1}$$

By setting the derivative of $\ln L$ equal to zero one finds the maximum-likelihood estimator for $\nu$ (here estimators will be written with hats),

$$\hat{\nu} = n \; . \tag{2}$$

Because $n$ follows a Poisson distribution, its variance is equal to its mean, so the variance of $\hat{\nu}$ is $V[\hat{\nu}] = V[n] = \nu$, and the estimator of the variance of the estimator is $\widehat{V}[\hat{\nu}] = n$. That is, the estimate of the standard deviation (the "statistical error") is $\hat{\sigma}_{\hat{\nu}} = \sqrt{n}$.

## 2 Weighted events

Now suppose each event carries a statistically independent weight, $w_1, \ldots, w_n$, the expectation value of the weights is $E[w] = \omega$, and the variance of the weights is $V[w] = \sigma_w^2$. Let us suppose as well the expectation value of $n$ is $E[n] = \nu/\omega$. Since $n$ is still Poisson distributed, the variance of $n$ is $V[n] = \nu/\omega$.

We have not yet made an assumption about the distribution of the weights except to say $E[w] = \omega$ and $V[w] = \sigma_w^2$. Suppose we use the sum of the weights as an estimator for $\nu$, i.e.,

$$\hat{\nu} = \sum_{i=1}^{n} w_i \; . \tag{3}$$

To compute the expectation value and variance of $\hat{\nu}$ we must take into account that $n$ and the $w_i$ are random quantities. Since the weights are independent, the conditional expectation value and variance of $\hat{\nu}$ given $n$ are

$$E[\hat{\nu}|n] = n\omega , \qquad (4)$$

$$V[\hat{\nu}|n] = n\sigma_w^2 . \qquad (5)$$

Using the law of total expectation [2], the unconditional expectation value is

$$E[\hat{\nu}] = E[E[\hat{\nu}|n]] = E[n\omega] = \frac{\nu}{\omega}\omega = \nu , \qquad (6)$$

and therefore the estimator $\hat{\nu}$ is unbiased.

The unconditional variance is found using the law of total variance [3] to be

$$V[\hat{\nu}] = E[V[\hat{\nu}|n]] + V[E[\hat{\nu}|n]] = E[n\sigma_w^2] + V[n\omega] = \frac{\nu}{\omega}(\sigma_w^2 + \omega^2) . \qquad (7)$$

To estimate the variance $V[\hat{\nu}]$ we can use the sum of the squares of the weights,

$$\hat{\sigma}_{\hat{\nu}}^2 = \sum_{i=1}^{n} w_i^2 . \qquad (8)$$

and we show below that this estimator is unbiased. Since $\sigma_w^2 = E[w^2] - \omega^2$, the conditional expectation of $\hat{\sigma}_{\hat{\nu}}^2$ given $n$ is

$$E\left[\hat{\sigma}_{\hat{\nu}}^2|n\right] = \sum_{i=1}^{n} E[w_i^2] = n(\sigma_w^2 + \omega^2) , \qquad (9)$$

and therefore the expectation value of $\hat{\sigma}_{\hat{\nu}}^2$ is

$$E\left[\hat{\sigma}_{\hat{\nu}}^2\right] = E\left[E\left[\hat{\sigma}_{\hat{\nu}}^2|n\right]\right] = \frac{\nu}{\omega}(\sigma_w^2 + \omega^2) . \qquad (10)$$

This is the same as the true variance from Eq. (7), and therefore the sum of the squares of the weights (3) is an unbiased estimator for the variance of $\hat{\nu}$. That is, we can take the sum of the weights to estimate the number of events that would be present if all had a weight of unity, and we use the square root of the sum of the squares of weights for the corresponding standard deviation (i.e., the statistical error).

## 3  Different sets of weights

Now suppose that we select $n$ events and there are two possible sets of weights, $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$, in which pairs of values having the same index are correlated, i.e., $\mathrm{cov}[u_i, v_i]$ may be nonzero, but otherwise the values are independent. Suppose we use the estimators

$$\hat{\mu} = \sum_{i=1}^{n} u_i , \tag{11}$$

$$\hat{\nu} = \sum_{i=1}^{n} v_i , \tag{12}$$

to estimate the corresponding mean values $\mu$ and $\nu$. From the results of Sec. 2 we know that the estimators are unbiased, and that their variances can be estimated using the sum of squares of the weights. In some problems, we may also need their covariance, $\mathrm{cov}[\hat{\mu}, \hat{\nu}]$. This can be related to conditional covariances and expectation values given $n$ using the law of total covariance [4],

$$\mathrm{cov}[\hat{\mu}, \hat{\nu}] = E\left[\mathrm{cov}[\hat{\mu}, \hat{\nu}|n]\right] + \mathrm{cov}\left[E[\hat{\mu}|n], E[\hat{\nu}|n]\right] . \tag{13}$$

To compute the terms on the right-hand side we need

$$E[\hat{\mu}|n] = E\left[\sum_{i=1}^{n} u_i|n\right] = nE[u] , \tag{14}$$

$$E[\hat{\nu}|n] = E\left[\sum_{i=1}^{n} v_i|n\right] = nE[v] , \tag{15}$$

as well as

$$\mathrm{cov}[\hat{\mu}, \hat{\nu}|n] = \sum_{i,j=1}^{n} \mathrm{cov}[u_i, v_j] = \sum_{i,j=1}^{n} \delta_{ij}\mathrm{cov}[u_i, v_i] = \sum_{i=1}^{n} \mathrm{cov}[u_i, v_i] = n\mathrm{cov}[u, v] . \tag{16}$$

Substituting the ingredients into Eq. (13) gives

$$\begin{aligned}
\mathrm{cov}[\hat{\mu}, \hat{\nu}] &= E\left[n\mathrm{cov}[u, v]\right] + \mathrm{cov}\left[nE[u], nE[v]\right] \\
&= E[n]\mathrm{cov}[u, v] + E[u]E[v]\mathrm{cov}[n, n] .
\end{aligned} \tag{17}$$

We can use the fact that $n$ is Poisson distributed so $\mathrm{cov}[n, n] = V[n] = E[n]$. Furthermore by definition we have $\mathrm{cov}[u, v] = E[uv] - E[u]E[v]$ and therefore we find the covariance of $\hat{\mu}$ and $\hat{\nu}$ is

$$\mathrm{cov}[\hat{\mu}, \hat{\nu}] = E[n]E[uv] . \tag{18}$$

To estimate the covariance we can estimate $E[n]$ with $n$ and $E[uv]$ with the corresponding average. After canceling the factors of $n$ this gives the simple estimator

$$\widehat{\mathrm{cov}}[\hat{\mu}, \hat{\nu}] = \sum_{i=1}^{n} u_i v_i . \tag{19}$$

Using again the law of total expectation one can verify that this estimator is unbiased, since

$$E\left[\widehat{\text{cov}}[\hat{\mu}, \hat{\nu}]\right] = E\left[\sum_{i=1}^{n} u_i v_i\right] = E\left[E\left[\sum_{i=1}^{n} u_i v_i | n\right]\right] = E\left[nE[uv]\right] = E[n]E[uv] , \qquad (20)$$

which is the same as the true covariance found in Eq. (18).

For the special case where the weights are the same for the two samples, i.e., $u_i = v_i$, then $\hat{\mu}$ and $\hat{\nu}$ are equal. In this case the expressions found for the covariance (18) and its estimator (19) reduce to the corresponding results for the variance found above in Eqs. (7) and (8).

## 4   Use with error propagation

As an example using the covariance, suppose we have a function $f(\hat{\mu}, \hat{\nu})$ and we want to estimate its variance. Using error propagation (see, e.g., [5] Sec. 37.2.5), this can be approximated using

$$V[f] \approx \left(\frac{\partial f}{\partial \hat{\mu}}\right)^2 V[\hat{\mu}] + \left(\frac{\partial f}{\partial \hat{\nu}}\right)^2 V[\hat{\nu}] + 2\left(\frac{\partial f}{\partial \hat{\mu}}\frac{\partial f}{\partial \hat{\nu}}\right) \text{cov}[\hat{\mu}, \hat{\nu}] . \qquad (21)$$

For example, for the ratio $f = \hat{\mu}/\hat{\nu}$ one obtains

$$V[f] \approx \frac{V[\hat{\mu}]}{\hat{\nu}^2} + \frac{\hat{\mu}^2 V[\hat{\nu}]}{\hat{\nu}^4} - 2\frac{\hat{\mu}}{\hat{\nu}^3}\text{cov}[\hat{\mu}, \hat{\nu}] . \qquad (22)$$

To obtain an numerical estimate for this quantity, one simply replaces the (co)variances on the right-hand side by their estimators as derived above.

As noted in Ref. [5], the approximation of the error propagation procedure is only valid if the function $f$ is close to linear in a region over which the arguments of $f$ vary. Because the ratio $f = \hat{\mu}/\hat{\nu}$ is a nonlinear function of $\hat{\nu}$, the approximation will break down if the standard deviation of $\hat{\nu}$ is comparable to $\nu$ itself. From Eq. (7), the standard deviation of $\hat{\nu}$ is

$$\sigma_{\hat{\nu}} = \sqrt{\frac{\nu}{\omega}(\sigma_w^2 + \omega^2)} . \qquad (23)$$

If the condition $\sigma_{\hat{\nu}} \ll \nu$ does not hold, the the approximation for the variance in Eq. (22) will not be valid.

## References

[1] G. Cowan, *Statistical tests with weighted Monte Carlo events* (draft), www.pp.rhul.ac.uk/~cowan/stat/notes/weights.pdf (2012).

[2] *Law of Total Expectation*, Wikipedia entry en.wikipedia.org/wiki/Law_of_total_expectation.

[3] *Law of Total Variance*, Wikipedia entry en.wikipedia.org/wiki/Law_of_total_variance.

[4] *Law of Total Covariance*, Wikipedia entry en.wikipedia.org/wiki/Law_of_total_covariance.

[5] J. Beringer et al. (Particle Data Group), Phys. Rev. D86, 010001 (2012).