1 Calculating Asymmetries and Uncertainties in Asymmetry

In order to study the uncertainty in a measured asymmetry and how to combine independent measurements of the same asymmetry; first consider an asymmetry, A, determined from:

$$A = \frac{N^+ - N^-}{N^+ + N^-} \tag{1}$$

For simplicity define $N = N^+ + N^-$, then:

$$A = \frac{N^+ - N^-}{N} \tag{2}$$

1.1 Correlated Asymmetries

If the data for N^+ and N^- are collected concurrently in the same experiment such as in a left-right asymmetry measurement; then they are correlated and we can define a probability, q, for a given event to be within the set of N^+ events as:

$$q = \frac{N^+}{N} \tag{3}$$

Since an event is either in the set N^+ or the set N^-; the probability for a given event to be within the set N^- is then $1 - q$ and the asymmetry can be written as:

$$A = q - (1 - q) = 2q - 1 \tag{4}$$

The variance of A, σ_A^2, is then simply related to the variance of q by $\sigma_A^2 = 4\sigma_q^2$ and the variance of q is simply related to the variance in N^+ by:

$$\sigma_q^2 = \frac{\sigma_{N^+}^2}{N^2} \tag{5}$$

When N^+ and N^- are correlated the statistics for N^+ is binomial; since an event is either in N^+, with probability q, or it isn't, with probability $1 - q$. For binomial statistics the variance for N^+ is:

$$\sigma_{N^+}^2 = Nq(1 - q) \tag{6}$$

and then the variance for a correlated measured asymmetry is:

$$\sigma_A^2 = \frac{4q(1 - q)}{N} = \frac{4N^+N^-}{N^3} \tag{7}$$

1.2 Uncorrelated Asymetries

If the data for N^+ and N^- are collected at two different times or with different experimental configurations then they are uncorrelated and the variances follow Poisson statistics. However, there now must be a mechanism for normalising the two measurements. Assume that there are normalising factors Q^+ and Q^- such that:

$$R^+ = \frac{N^+}{Q^+} \tag{8}$$

$$R^- = \frac{N^-}{Q^-} \tag{9}$$

where the values R^+ and R^- can now be used to determine the asymmetry:

$$A = \frac{R^+ - R^-}{R^+ + R^-} \tag{10}$$

Since the measurements are not correlated the variance in A can be written as:

$$\sigma_A^2 = \frac{4R^{-2}}{(R^+ + R^-)^4}\sigma_{R^+}^2 + \frac{4R^{+2}}{(R^+ + R^-)^4}\sigma_{R^-}^2 \tag{11}$$

and the variances in the R values take the form:

$$\sigma_R^2 = \frac{1}{Q^2}\sigma_N^2 + \frac{N^2}{Q^4}\sigma_Q^2 \tag{12}$$

In principle the above form for σ_R^2 should be used but if the normalisation is such that $N << Q$ then perhaps the second term with σ_Q^2 can be neglected. Assuming this and knowing that $\sigma_N^2 = N$ from Poisson statistics then:

$$\sigma_A^2 = \frac{4R^+R^-}{(R^+ + R^-)^4}\left(\frac{R^+}{Q^-} + \frac{R^-}{Q^+}\right) \tag{13}$$

$$= \frac{4N^+N^-(N^+ + N^-)(Q^+Q^-)^2}{(N^+Q^- + N^-Q^+)^4} \tag{14}$$

The above expression for σ_A^2 is not as simple as that given by equation 7 when the asymmetry measurement is correlated but the form is similar and the similarity becomes much clearer if we contrive the uncorrelated experiments such that $Q^+ \approx Q^-$ in which case:

$$\sigma_A^2 = \frac{4R^+R^-}{(R^+ + R^-)^3}\left(\frac{1}{Q}\right) \tag{15}$$

$$= \frac{4N^+N^-}{(N^+ + N^-)^3} = \frac{4N^+N^-}{N^3} \tag{16}$$

1.3 Caveat

The expression for σ_A^2 in equations 7 and 16 are the same but remember that equation 16 assumes the variance from the measurement of the normalisation factors can be ignored and also that the normalisation factors for the separate measurements are the same or can be made so.

In the following discussions I will assume equations 7 and 16 can be used and furthermore will assume $Q = 1$ in some system of units so that $R = N$ and thus deal with both correlated and uncorrelated asymmetries as if they were both expressed as equation 1. Furthermore I can still define the probability q as before but must be careful using it.

1.4 Combining Asymmetry Measurements

If there are two independent measurements of A, namely A_1 and A_2 given by:

$$A_1 = \frac{N_1^+ - N_1^-}{N_1^+ + N_1^-} \tag{17}$$

and

$$A_2 = \frac{N_2^+ - N_2^-}{N_2^+ + N_2^-} \tag{18}$$

then the variance weighted average is:

$$A = \frac{\frac{A_1}{\sigma_{A_1}^2} + \frac{A_2}{\sigma_{A_2}^2}}{\frac{1}{\sigma_{A_1}^2} + \frac{1}{\sigma_{A_2}^2}} \tag{19}$$

Substituting for the variances in terms of q_1 and q_2 and simplifying yields:

$$A = \frac{A_1 N_1 q_2(1-q_2) + A_2 N_2 q_1(1-q_1)}{N_1 q_2(1-q_2) + N_2 q_1(1-q_1)} \tag{20}$$

or

$$A = \frac{(N_1^+ - N_1^-)q_2(1-q_2) + (N_2^+ - N_2^-)q_1(1-q_1)}{N_1 q_2(1-q_2) + N_2 q_1(1-q_1)} \tag{21}$$

and the variance of the weighted average is simply:

$$\sigma_A^2 = \frac{1}{\frac{1}{\sigma_{A_1}^2} + \frac{1}{\sigma_{A_2}^2}} \tag{22}$$

$$= \frac{4q_1(1-q_1)q_2(1-q_2)}{N_1 q_2(1-q_2) + N_2 q_1(1-q_1)} \tag{23}$$

While we have two independent measures of the probability for an event to be in the "+" category, namely q_1 and q_2; in principle these should be the same quantity and, assuming all measures are equally good, we can replace q_1 and q_2 with simply q yielding the much more manageable equations:

$$A = \frac{A_1 N_1 + A_2 N_2}{N_1 + N_2} \tag{24}$$

or

$$A = \frac{(N_1^+ - N_1^-) + (N_2^+ - N_2^-)}{N_1 + N_2} \tag{25}$$

and

$$\sigma_A^2 = \frac{4q(1-q)}{N_1 + N_2} \tag{26}$$

where we should use:

$$q = \frac{N_1^+ + N_2^+}{N_1 + N_2} \tag{27}$$

which is the variance weighted average of q_1 and q_2. Note the variance of q is different in the correlated and uncorrelated experiments and the above expression holds for the correlated case in general but is only valid for the uncorrelated case with the approximations again assumed that the σ_Q^2 term is negligible and that $Q^+ \approx Q^-$

The above of course extends to multiple independent measurements in an obvious manner:

$$\begin{aligned}
A &= \frac{\sum A_i N_i}{\sum N_i} &(28)\\
&= \frac{\sum (N_i^+ - N_i^-)}{\sum N_i} &(29)\\
\sigma_A^2 &= \frac{4q(1-q)}{\sum N_i} &(30)\\
q &= \frac{\sum N_i^+}{\sum N_i} &(31)
\end{aligned}$$

1.5 Asymmetry Measurements with Polarisation

The situation is slightly more complicated when the asymmetry to be determined is not measured directly but depends on a polarisation used in the experiment. Then:

$$A = \frac{N^+ - N^-}{N} = PA^{true} \tag{32}$$

or

$$A^{true} = \frac{A}{P} = \frac{N^+ - N^-}{NP} = \frac{2q-1}{P} \tag{33}$$

where A is the measured asymmetry as before, P is the polarisation used in the experiment, and A^{true} is the "true" asymmetry to be determined.

Note that P and A are correlated since A^{true} is presumably a constant. However, if the measurement of P is independent of A then it is possible that their variances are not correlated provided the variations are statistical and not due to an underlying change in polarisation during the course of the measurement. (i.e. the polarisation doesn't actually change during the measurement but the observed polarisation fluctuates statistically due to the limitations in the measurement of the polarisation).

Assuming the variances of P and A are independent then the variance in A^{true} is given by:

$$\begin{aligned}
\sigma^2_{A^{true}} &= \frac{1}{P^2}\sigma^2_A + \frac{A^2}{P^4}\sigma^2_P & (34)\\
&= \frac{4N^+N^-}{N^3P^2} + \frac{(N^+ - N^-)^2}{N^2P^4}\sigma^2_P & (35)\\
&= \frac{4q(1-q)}{NP^2} + \frac{(2q-1)^2}{P^4}\sigma^2_P & (36)
\end{aligned}$$

Now, re-examine the case where there are two or more independent measurements, this time possibly with different polarisations. Consider first the case of just two independent measurements and form the variance weighted average.

$$\frac{A^{true}_1 N_1 P^4_1[4q(1-q)P^2_2 + (2q-1)^2 N_2 \sigma^2_{P_2}] + A^{true}_2 N_2 P^4_2[4q(1-q)P^2_1 + (2q-1)^2 N_1 \sigma^2_{P_1}]}{N_1 P^4_1[4q(1-q)P^2_2 + (2q-1)^2 N_2 \sigma^2_{P_2}] + N_2 P^4_2[4q(1-q)P^2_1 + (2q-1)^2 N_1 \sigma^2_{P_1}]} \tag{37}$$

This is of course the proper calculation but is quite unwieldy and does not immediately lend itself to further simplification or improved insight.

A slight simplification is possible with the assumption that $\sigma^2_{P_i} = kP^2_i$ for some constant k. Then the variance weighted average can be written as:

$$\frac{A^{true}_1 N_1 P^2_1[4q(1-q) + (2q-1)^2 N_2 k] + A^{true}_2 N_2 P^2_2[4q(1-q) + (2q-1)^2 N_1 k]}{N_1 P^2_1[4q(1-q) + (2q-1)^2 N_2 k] + N_2 P^2_2[4q(1-q) + (2q-1)^2 N_1 k]} \tag{38}$$

but the terms within the brackets still can not be factored out because of the N_i within each which varies for each measurement and which dominates unless k is very small. (If the N_i were the same for all measurements then this could be simplified and used but this is an unrealistic limitation for experiments.)

The only useful simplifications are the cases when $\frac{1}{P^2}\sigma_A^2$ or $\frac{A}{P^4}\sigma_P^2$ can be ignored (i.e. when the uncertainty is dominated by the polarisation measurement or the asymmetry measurement).

As an example consider the magnitude of the two terms in the expression for $\sigma^2_{A^{true}}$ for an asymmetry measured with a polarisation $P = 0.7$ and data $N^+ = 6000$ and $N^- = 4000$. The calculated quantities are:

$$A = 0.2 \tag{39}$$
$$q = 0.6 \tag{40}$$
$$\sigma_A^2 = 0.000096 \tag{41}$$
$$\sigma^2_{A^{true}} = 0.000196 + 0.167\sigma_P^2 \tag{42}$$

In this example the measured asymmetry is known to better than 1% and the determination of the "true" asymmetry is dominated by the polarisation measurement unless it is better than $\sim 3\%$ or $\sigma_P^2 = 0.0009$. If the statistics for the asymmetry measurement were $N = 100$ instead of $N = 10,000$ then the asymmetry measurement would dominate so long as the polarisation measurement was better than $\sim 20\%$.

Conclusion: ignoring the terms in the variance of asymmetry or polarisation must be evaluated on a case by case basis and applied with caution. Also, using the full form is always right.

Nevertheless, with the above mentioned caveats, consider the case where $\frac{A}{P^4}\sigma_P^2$ is negligible (i.e. asymmetry uncertainties dominate). Then the variance weighted average can be simplified and generalised as:

$$A^{true} = \frac{\sum A_i^{true} N_i P_i^2}{\sum N_i P_i^2} \tag{43}$$
$$= \frac{\sum (N_i^+ - N_i^-) P_i}{\sum N_i P_i^2} \tag{44}$$

Note the weights when summing the intermediate A_i^{true}'s are $N_i P_i^2$ but if you sum directly from the data, N_i^+ and N_i^-, the weights are just P_i. The variance for the final result is:

$$\sigma^2_{A^{true}} = \frac{4q(1-q)}{\sum N_i P_i^2} \tag{45}$$

1.6 Conclusion

In conclusion one should use the full form for the various variances and form the variance weighted average whenever possible. That's what computers are made for.

$$A^{true} = \frac{\sum \frac{A_i^{true}}{\sigma^2_{A_i^{true}}}}{\sum \frac{1}{\sigma^2_{A_i^{true}}}} \tag{46}$$

and

$$\sigma^2_{A^{true}} = \frac{1}{\sum \frac{1}{\sigma^2_{A_i^{true}}}} \tag{47}$$

For correlated asymmetry measurements:

$$A_i^{true} = \frac{A_i}{P_i} = \frac{N_i^+ - N_i^-}{N_i P_i} = \frac{2q_i - 1}{P_i} \tag{48}$$

and

$$\sigma^2_{A_i^{true}} = \frac{1}{P_i^2}\sigma^2_{A_i} + \frac{A_i^2}{P_i^4}\sigma^2_{P_i} \tag{49}$$

$$= \frac{4N_i^+ N_i^-}{N_i^3 P_i^2} + \frac{(N_i^+ - N_i^-)^2}{N_i^2 P_i^4}\sigma^2_{P_i} \tag{50}$$

$$= \frac{4q_i(1-q_i)}{N_i P_i^2} + \frac{(2q_i - 1)^2}{P_i^4}\sigma^2_{P_i} \tag{51}$$

These can also be used for uncorrelated asymmetry measurements as typically made in BLAST but with the understanding that the variance in normalisation is being ignored and that the normalisation factors have been contrived to be the same which is very restrictive. Caveat emptor ! Use the full calculations.

Equation 51 does illustrate that the variance in asymmetry measurements improves as $1/NP^2$ for asymmetry dominated uncertainties reflecting that the figure of merit is NP^2. When the polarisation uncertainty becomes significant this enters as $1/P^4$ and should not be ignored.