Reweighting of Events

Benno List

5.6.1999

Reweighting Lifetime Distributions

Assume that a Monte Carlo sample contains particles, say D^0 particles, with a certain lifetime τ. The distribution of the proper lifetime t of these particles is then given by a probability density function $p(t)$ according to

$$p(t) = \frac{1}{\tau} \exp\left(-\frac{t}{\tau}\right).$$

If one wants to reweight the sample in order to simulate a different lifetime τ', the resulting p. d. f. should be given by

$$p'(t) = \frac{1}{\tau'} \exp\left(-\frac{t}{\tau'}\right),$$

so obviously an event containing a single D^0 with a proper lifetime t has to be assigned a weight $w(t)$

$$\frac{p'(t)}{p(t)} = \frac{\tau}{\tau'} \exp\left(\frac{t}{\tau} - \frac{t}{\tau'}\right) = \frac{\tau}{\tau'} \exp\left(-\frac{t}{\tau_{\mathrm{rew}}}\right)$$

with τ_{rew} given by

$$\frac{1}{\tau_{\mathrm{rew}}} = \frac{1}{\tau'} - \frac{1}{\tau}.$$

As should be, the mean weight $\langle w \rangle$ defined as

$$\langle w \rangle = \int_0^\infty p(t) \cdot w(t) \, \mathrm{d}\, t$$

is equal to one.

In cases where $\tau' < \tau$, the maximum weight occurs clearly for $t = 0$ and is given by τ/τ'. Therefore, as long as τ' is smaller than τ, the weights are bounded, and if τ/τ' is not too large, the reweighting will not result in too large a loss of statistics.

On the other hand, if $\tau' > \tau$, the weight is in principle unbounded, although for a finite Monte Carlo sample it will be $\tau/\tau' \exp(t_{\max}/|\tau_{\mathrm{rew}}|)$, where $t_{\max}$ is the largest proper lifetime occuring in the Monte Carlo sample.

Since the probability for a lifetime t to be greater than or equal to some value t_0 is given by $\exp(t_0/\tau)$, one expects less than 0.05 events with a lifetime greater than $\tau \cdot \ln(N/0.05)$ in a Monte Carlo sample with N events. For example, in a MC sample of 10^6 events, the factor $\ln(N/0.05)$ is 16.81. Consequently, there is only a 5 % chance to encounter a weight greater than

$$w_{95} = w\left(\tau \cdot \ln\frac{N}{0.05}\right) = \frac{\tau}{\tau'}\exp\left(\frac{\tau'-\tau}{\tau'}\ln\frac{N}{0.05}\right) = \frac{\tau}{\tau'}\left(\frac{N}{0.05}\right)^{\frac{\tau'-\tau}{\tau'}}.$$

Since $(\tau' - \tau)/\tau' < 1$ is always valid, and $\tau/\tau' < 1$ is valid in the case considered here, this weight will always be smaller than $\frac{N}{0.05}$, a fact that is, however, not very consoling. If the maximum expected weight is to have a reasonable value $w_{\max}$, say, 10, one has to demand

$$\frac{\tau}{\tau'}\left(\frac{N}{0.05}\right)^{\frac{\tau'-\tau}{\tau'}} < \left(\frac{N}{0.05}\right)^{\frac{\tau'-\tau}{\tau'}} < w_{\max}$$

which leads to

$$\tau' < \frac{\tau}{1 - \frac{\ln(N/0.05)}{\ln w_{\max}}}$$

(neglecting the factor τ/τ', which does not change the result much). For $N = 10^6$ and $w_{\max} = 10$ this leads to $\tau' < 1.16\tau$, which means that if one wants to avoid having weights in excess of 10 in a Monte Carlo sample, one may increase the lifetime of a particle by only 16 % through reweighting.

The variance of the weight distribution V_{w} is given by

$$\begin{aligned} V_{\mathrm{w}} &= \left\langle (w - \langle w\rangle)^2 \right\rangle = \langle w^2\rangle - \langle w\rangle^2 \\ &= \int_0^\infty p(t)\cdot w^2(t)\,\mathrm{d}\,t - 1 = \frac{\tau}{\tau'^2}\int_0^\infty \exp\left(-t\left[\frac{2}{\tau'} - \frac{1}{\tau}\right]\right)\mathrm{d}\,t - 1. \end{aligned}$$

The integral converges only if $\tau' < 2\tau$, otherwise the variance is infinitely big. If the condition $\tau' < 2\tau$ holds, the variance is given by

$$V_{\mathrm{w}} = \frac{(\tau' - \tau)^2}{\tau'(2\tau - \tau')}.$$

Error on Efficiency for Reweighted MC

Assume that some tagging technique is applied to the Monte Carlo sample. For the sake of simplicity, let us assume that the tag is fully efficient above some proper particle lifetime t_0, and totally inefficient below that threshold.

Then the efficiency before the reweighting is given by

$$\epsilon = \int_{t_0}^\infty p(t)\,\mathrm{d}\,t = \exp\left(-\frac{t_0}{\tau}\right).$$

The efficiency after reweighting ϵ' is given by

$$\epsilon' = \int_{t_0}^\infty p(t)\cdot w(t)\,\mathrm{d}\,t = \exp\left(-\frac{t_0}{\tau'}\right).$$

Both efficiencies are related through

$$\frac{\epsilon'}{\epsilon} = \exp\left(-t_0\left[\frac{1}{\tau'} - \frac{1}{\tau}\right]\right) = \exp\left(-\frac{t_0}{\tau_{\text{rew}}}\right).$$

For a Monte Carlo sample containing N D^0 particles, the sum of the weights of particles that pass the selection is

$$\sum_{+} w_i = N \int_{t_0}^{\infty} p(t) \cdot w(t) \,\mathrm{d}\, t = \epsilon' N,$$

and the sum of the weights failing the selection is consequently given by

$$\sum_{-} w_i = N \int_{0}^{t_0} p(t) \cdot w(t) \,\mathrm{d}\, t = (1-\epsilon')N.$$

Likewise the sum of the squared weights for events that pass the selection is

$$\sum_{+} w_i^2 = N \int_{t_0}^{\infty} p(t) \cdot w^2(t) \,\mathrm{d}\, t = \frac{\tau^2}{\tau'(2\tau - \tau')} \exp\left(-t_0\left[\frac{2}{\tau'} - \frac{1}{\tau}\right]\right) N = \frac{(tau^2}{\tau'(2\tau-\tau')}\frac{\epsilon'^2}{\epsilon}N.$$

Again, this integral converges only if $\tau' < 2\tau$, demonstrating that otherwise a meaningful reweighting is totally impossible.

Finally, the sum of the squared weights for events that fail the selection is

$$\sum_{-} w_i^2 = N \int_{0}^{t_0} p(t) \cdot w^2(t) \,\mathrm{d}\, t = \frac{(tau^2}{\tau'(2\tau-\tau')}\frac{\epsilon - \epsilon'^2}{\epsilon}N.$$

Inserting this in the formula for the error of the efficiency, which is

$$\delta\epsilon' = \frac{\sqrt{\sum_{+} w_i^2 \left(\sum_{-} w_i\right)^2 + \sum_{-} w_i^2 \left(\sum_{+} w_i\right)^2}}{\left(\sum w_i\right)^2},$$

leads to

$$\delta\epsilon' = \sqrt{\frac{\frac{\epsilon'^2(1-2\epsilon'+\epsilon)}{\epsilon} \cdot \frac{\tau^2}{\tau'(2\tau-\tau')}}{N}}.$$

Inserting

$$\epsilon = \epsilon' \exp\left(\frac{t_0}{\tau_{\text{rew}}}\right)$$

leads to

$$\delta\epsilon' = \sqrt{\frac{\epsilon'(1-\epsilon'+\epsilon'[\exp(t_0/\tau_{\text{rew}})-1]) \cdot \frac{\tau^2}{\exp(t_0/\tau_{\text{rew}}) \cdot \tau'(2\tau-\tau')}}{N}}.$$

In the case that $\tau' = \tau$, i. e. no reweighting has occurred, one gets $\epsilon' = \epsilon$ and therefore the usual result

$$\delta\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{N}}.$$

Table 1 lists some numerical examples for the resulting error and compares the result to the case where the error is determined from an unweighted sample.

τ'/τ	ϵ	0.05	0.10	0.20	0.50	0.90
0.80	ϵ'	0.0236 ± 0.0011	0.0562 ± 0.0018	0.1337 ± 0.0029	0.4204 ± 0.0049	0.8766 ± 0.0036
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	0.71	0.78	0.87	1.00	1.10
0.90	ϵ'	0.0358 ± 0.0016	0.0774 ± 0.0024	0.1673 ± 0.0035	0.4629 ± 0.0050	0.8895 ± 0.0033
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	0.86	0.90	0.94	1.00	1.05
0.95	ϵ'	0.0427 ± 0.0019	0.0886 ± 0.0027	0.1838 ± 0.0038	0.4821 ± 0.0050	0.8950 ± 0.0031
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	0.93	0.95	0.97	1.00	1.02
0.98	ϵ'	0.0470 ± 0.0021	0.0954 ± 0.0029	0.1935 ± 0.0039	0.4930 ± 0.0050	0.8981 ± 0.0031
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	0.97	0.98	0.99	1.00	1.01
0.99	ϵ'	0.0485 ± 0.0021	0.0977 ± 0.0029	0.1968 ± 0.0040	0.4965 ± 0.0050	0.8990 ± 0.0030
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	0.99	0.99	0.99	1.00	1.00
1.00	ϵ'	0.0500 ± 0.0022	0.1000 ± 0.0030	0.2000 ± 0.0040	0.5000 ± 0.0050	0.9000 ± 0.0030
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.00	1.00	1.00	1.00	1.00
1.01	ϵ'	0.0515 ± 0.0022	0.1023 ± 0.0031	0.2032 ± 0.0040	0.5034 ± 0.0050	0.9009 ± 0.0030
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.01	1.01	1.01	1.00	1.00
1.02	ϵ'	0.0530 ± 0.0023	0.1046 ± 0.0031	0.2064 ± 0.0041	0.5068 ± 0.0050	0.9019 ± 0.0030
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.03	1.02	1.01	1.00	0.99
1.05	ϵ'	0.0577 ± 0.0025	0.1116 ± 0.0033	0.2159 ± 0.0042	0.5168 ± 0.0050	0.9045 ± 0.0029
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.07	1.05	1.03	1.00	0.98
1.10	ϵ'	0.0657 ± 0.0028	0.1233 ± 0.0036	0.2315 ± 0.0045	0.5325 ± 0.0050	0.9087 ± 0.0028
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.14	1.10	1.06	1.00	0.96
1.20	ϵ'	0.0824 ± 0.0035	0.1468 ± 0.0043	0.2615 ± 0.0049	0.5612 ± 0.0050	0.9159 ± 0.0026
	$\delta\epsilon'/\delta\epsilon'_{\rm unw}$	1.29	1.20	1.12	1.00	0.93

Table 1: Results for the reweighting of lifetime distributions fro different tagging efficiencies ϵ and different ratios of τ'/τ. For each pair of $(\epsilon, \tau'/\tau)$ the new tagging efficiency ϵ' is given with its error, assuming $N = 10000$ events. Additionally, the ratio of the error for the reweighted sample, $\delta\epsilon'$, to the error $\delta\epsilon'_{\rm unw}$ that would result for an unweighted sample of $N = 10000$ events is given.

Linear Reweighting

Consider the case that some variable x is distributed according to a probability density function (p. d. f.) $p(x)$, and that the values of x all lie within $x_{\max} \leq x \leq x_{\max}$, i. e. $p(x) = 0$ for $x < x_{\min}$ and for $x > x_{\max}$.

Let $\langle x \rangle$ be the mean value of x, which is given by

$$\langle x \rangle = \int_{x_{\min}}^{x_{\max}} x \cdot p(x) \, \mathrm{d}x,$$

from which follows that

$$\int_{x_{\min}}^{x_{\max}} (x - \langle x \rangle) \cdot p(x) \, \mathrm{d}x = 0.$$

Now let us assume that the events are reweighted using a reweighting function f, given by

$$f(x) = 1 + a(x - \langle x \rangle).$$

a is an arbitrary constant, which must, however, fulfill the condition

$$\frac{1}{x_{\min} - \langle x \rangle} < a < \frac{1}{x_{\max} - \langle x \rangle},$$

which ensures that $f(x) > 0$ for all x within $x_{\min}$ and $x_{\max}$.

Let us define $p'(x) = f(x) \cdot p(x)$, which has the normalization

$$\begin{aligned}
\int_{x_{\min}}^{x_{\max}} p'(x) \, \mathrm{d}x &= \int_{x_{\min}}^{x_{\max}} f(x) \cdot p(x) \, \mathrm{d}x \\
&= \int_{x_{\min}}^{x_{\max}} p(x) \, \mathrm{d}x + a \int_{x_{\min}}^{x_{\max}} (x - \langle x \rangle) \cdot p(x) \, \mathrm{d}x \\
&= 1 + a \cdot 0 = 1.
\end{aligned}$$

Therefore, since it is nonnegative everywhere and normalized to 1, $p'(x)$ is a proper p. d. f. itself, namely the p. d. f. of the reweighted event sample.

Now let us evaluate the mean value of x when the new p. d. f. $p'(x)$ is used, which we denote by $\langle x \rangle'$:

$$\begin{aligned}
\langle x \rangle' &= \int_{x_{\min}}^{x_{\max}} x \cdot p'(x) \, \mathrm{d}x \\
&= \int_{x_{\min}}^{x_{\max}} x \cdot p(x) \, \mathrm{d}x + a \int_{x_{\min}}^{x_{\max}} x \cdot (x - \langle x \rangle) \cdot p(x) \, \mathrm{d}x - a \cdot \langle x \rangle \int_{x_{\min}}^{x_{\max}} (x - \langle x \rangle) \cdot p(x) \, \mathrm{d}x \\
&= \int_{x_{\min}}^{x_{\max}} x \cdot p(x) \, \mathrm{d}x + a \int_{x_{\min}}^{x_{\max}} (x - \langle x \rangle)^2 \cdot p(x) \, \mathrm{d}x \\
&= \langle x \rangle + a \left\langle (x - \langle x \rangle)^2 \right\rangle
\end{aligned}$$

(in the second line we have inserted the term $a \cdot \langle x \rangle \int\limits_{x_{\min}}^{x_{\max}} (x - \langle x \rangle) \cdot p\,(x)\,\mathrm{d}x$, which is zero). This means that the reweighting procedure shifts the mean value of x by an amount of $a\langle (x - \langle x \rangle)^2 \rangle$, i. e. a times the variance of x. Therefore, linear reweighting is a simple method to shift the mean value of some variable x in a Monte Carlo sample by a (small) amount, provided that x is bounded[1].

Nonlinear Reweighting

In general, every function $f\,(x)$ can be used for reweighting, provided it fulfills the conditions

$$
\begin{aligned}
f\,(x) &> 0 \qquad \text{for} \quad x_{\min} \le x \le x_{\max}, \\
\langle f \rangle &= \int\limits_{x_{\min}}^{x_{\max}} f\,(x) \cdot p\,(x)\,\mathrm{d}x = 1.
\end{aligned}
$$

Let us write $f\,(x)$ in the general form

$$
f\,(x) = 1 + a \cdot (k\,(x - \langle x \rangle) - \bar{k}),
$$

where $k\,(x)$ is the *kernel* of the reweighting function, and $\bar{k}$ is its mean value, defined as

$$
\bar{k} = \int\limits_{x_{\min}}^{x_{\max}} k\,(x - \langle x \rangle) \cdot p\,(x)\,\mathrm{d}x.
$$

Linear reweighting is a special case of this, with the kernel $k\,(x) = x$, for which we have $\bar{k} = 0$.

With this definition, we have

$$
\begin{aligned}
\int\limits_{x_{\min}}^{x_{\max}} f\,(x) \cdot p\,(x)\,\mathrm{d}x &= \int\limits_{x_{\min}}^{x_{\max}} p\,(x)\,\mathrm{d}x + a \int\limits_{x_{\min}}^{x_{\max}} k(x - \langle x \rangle) \cdot p\,(x)\,\mathrm{d}x - a\bar{k} \int\limits_{x_{\min}}^{x_{\max}} p\,(x)\,\mathrm{d}x \\
&= 1 + a \cdot (\bar{k} - \bar{k}) = 1,
\end{aligned}
$$

which means that $p'\,(x) = f\,(x) \cdot p\,(x)$ is again a proper p. d. f..

The mean of the reweighted distribution of x is given by

$$
\begin{aligned}
\langle x \rangle' &= \int\limits_{x_{\min}}^{x_{\max}} x \cdot p'\,(x)\,\mathrm{d}x \\
&= \int\limits_{x_{\min}}^{x_{\max}} x \cdot p\,(x)\,\mathrm{d}x + a \int\limits_{x_{\min}}^{x_{\max}} x \cdot k\,(x - \langle x \rangle) \cdot p\,(x)\,\mathrm{d}x - a \cdot \bar{k} \int\limits_{x_{\min}}^{x_{\max}} x \cdot p\,(x)\,\mathrm{d}x
\end{aligned}
$$

[1] Which is necessarily the case in a finite Monte Carlo sample. However, in reality $|a(x_{\min} - \langle x \rangle)| \ll 1$ and $|a(x_{\min} - \langle x \rangle)| \ll 1$ should hold to ensure that the weights are not excessively big.

$$
\begin{aligned}
&= \int\limits_{x_{\min}}^{x_{\max}} x \cdot p\left(x\right) \mathrm{d}x + a \int\limits_{x_{\min}}^{x_{\max}} x \cdot k\left(x - \langle x \rangle\right) \cdot p\left(x\right) \mathrm{d}x - a \cdot \langle x \rangle \int\limits_{x_{\min}}^{x_{\max}} k\left(x - \langle x \rangle\right) \cdot p\left(x\right) \mathrm{d}x \\
&= \int\limits_{x_{\min}}^{x_{\max}} x \cdot p\left(x\right) \mathrm{d}x + a \int\limits_{x_{\min}}^{x_{\max}} \left(x - \langle x \rangle\right) \cdot k\left(x - \langle x \rangle\right) \cdot p\left(x\right) \mathrm{d}x \\
&= \langle x \rangle + a \left\langle \left(x - \langle x \rangle\right) \cdot k\left(x - \langle x \rangle\right) \right\rangle
\end{aligned}
$$

This means that the mean value of the reweighted distribution is shifted by an amount which is proportional to $\left\langle \left(x - \langle x \rangle\right) \cdot k\left(x - \langle x \rangle\right) \right\rangle$. This quantity can be determined easily from the distribution of x, and from that a reweighting function can be constructed which shifts the mean value by any (small enough) amount.

A useful choice of the kernel function can be the hyperbolic tangens:

$$
k\left(x\right) = \tanh\left(\frac{x}{b}\right) = \frac{\exp\left(\frac{x}{b}\right) - \exp\left(\frac{x}{b}\right)}{\exp\left(\frac{x}{b}\right) + \exp\left(\frac{x}{b}\right)},
$$

where b is a constant that is characteristic of the spread of the p. d. f., for instance $b = \sigma$ for a Gaussian or $b = \Gamma/2$ for a Breit–Wigner distribution. Since $-1 < \tanh(x) < 1$ is true for all x, this kernel is suitable for reweighting dsitributions where x is not bounded, such as Gaussian or Breit–Wigner distributions. Of course, the shape of the distribution is distored by this reweighting method, but the weights are reasonably small for small values of a, so that a shift in the mean value can be obtained without getting too big weights.