General result for the bias

We wish to compute for $N \to \infty$ the the bias, that is the average

$$J_N = \left\langle \frac{1}{N+1} \sum_{n=0}^{N} f(X_n) - \langle f \rangle_\pi \right\rangle_{MC,0}$$

where $\langle \cdot \rangle_{MC,0}$ indicates the average over all processes that start in X_0.

The probability of being in point y at $t = n$ is P^n_{0y}, so that

$$\langle f(X_n) \rangle_{MC,0} = \sum_y f(y) \times (\text{probability to be in } y \text{ at } t = n) = \sum_y f(y) P^n_{0y}$$

We also write

$$\langle f \rangle_\pi = \frac{1}{N+1} \sum_{n=0}^{N} \langle f \rangle_\pi = \frac{1}{N+1} \sum_{n=0}^{N} \sum_y f(y) \pi_y$$

so that

$$\begin{aligned} J_N &= \frac{1}{N+1} \sum_{n=0}^{N} \sum_y \left[P^n_{0y} f(y) - \pi_y f(y) \right] \\ &= \frac{1}{N+1} \sum_y f(y) \sum_{n=0}^{N} \left[P^n_{0y} - \pi_y \right] \end{aligned}$$

Now, we introduce the projector $\Pi_{xy} = \pi_y$. It has the properties:

$$(\Pi P)_{xy} = \sum_z \Pi_{xz} P_{zy} = \sum_z \pi_z P_{zy} = \pi_y = \Pi_{xy}$$
$$(P\Pi)_{xy} = \sum_z P_{xz} \Pi_{zy} = \sum_z P_{xz} \pi_y = \pi_y = \Pi_{xy}$$
$$(\Pi\Pi)_{xy} = \sum_z \Pi_{xz} \Pi_{zy} = \sum_z \pi_z \pi_y = \pi_y = \Pi_{xy}$$

It follows that P and Π commute: $P\Pi = \Pi P = \Pi$. Moreover, for $m \geq 1$, we have

$$P^n \Pi^m = P^n \Pi = P^{n-1}(P\Pi) = P^{n-1}\Pi = \ldots = \Pi$$

It follows ($k \geq 1$)

$$(P - \Pi)^k = \sum_{n=0}^{k} \binom{k}{n} (-1)^{k-n} P^n \Pi^{k-n} = P^k + \Pi \sum_{n=0}^{k-1} \binom{k}{n} (-1)^{k-n}.$$

Now we include the term with $n = k$ in the sum and subtract it; the remaining series can be resummed, obtaining

$$\sum_{n=0}^{k-1} \binom{k}{n} (-1)^{k-n} = -1 + \sum_{n=0}^{k} \binom{k}{n} (-1)^{k-n} = -1 + (1-1)^k = -1$$

It follows

$$(P-\Pi)^k = P^k - \Pi \quad \Rightarrow \quad P^k = (P-\Pi)^k + \Pi$$

Therefore, for $n \geq 1$

$$P^n_{0y} - \pi_y = (P^n - \Pi)_{0y} = (P-\Pi)^n_{0y}$$

so that, we can write

$$J_N = \frac{1}{N+1} \sum_y f(y) \left[I - \Pi + \sum_{n=1}^{N} (P-\Pi)^n \right]_{0y}$$

where I is the identity matrix: $I_{xy} = 1$ if $x = y$, and 0 otherwise. the first two terms in the previous expression, $I - \Pi$, are the terms that correspond to $n = 0$.

Now, a general theorem states that the eigenvalues λ_i of an irreducible aperiodic transition matrix satisfy the conditions:
1) $|\lambda_i| \leq 1$;
2) there is only one eigenvalue on the unit circle: it has $\lambda_i = 1$, the left eigenvector is π_x, the right eigenvector is $(1, \ldots, 1)$.
Now, Π is the projector onto the eigenvector of eigenvalue 1, so that all eigenvectors of $P - \Pi$ lie strictly inside the unit circle.
Consequences:
1) $(P-\Pi)^n_{0y}$ behaves as $|\lambda_2|^n$ as $n \to \infty$, where λ_2 is the second-largest (in absolute value) eigenvalue. Since $|\lambda_2| < 1$, the matrix element vanishes exponentially for $n \to \infty$.
2) The matrix $I - (P - \Pi)$ has an inverse (there are some caveats for infinite-dimensional systems, which have however little practical interest).
Because of 1), we can extend the sum from N to ∞. Then, we can resum the series, using $(1-x)^{-1} = \sum_n x^n$.
Therefore, discarding exponentially small terms we can write

$$J_N \approx \frac{1}{N+1} \sum_y f(y) \left[I - \Pi + \sum_{n=1}^{\infty} [(P-\Pi)^n_{0y} \right] = \frac{1}{N+1} \sum_y f(y) \left[(I - P + \Pi)^{-1} - \Pi \right]_{0y}$$

Therefore, the sample mean is a **biased** estimate of the sample average. The bias is of order $1/N$.