Algorithm 12.3 Likelihood weighting with a data-dependent stopping rule

Procedure Data-Dependent-LW (
  $\mathcal{B}$, // Bayesian network over $\mathcal{X}$
  $\boldsymbol{Z} = \boldsymbol{z}$, // Instantiation of interest
  u, // Upper bound on CPD entries of $\boldsymbol{Z}$
  ϵ, // Desired error bound
  δ // Desired probability of error
)
  $\gamma \leftarrow \frac{4(1+\epsilon)}{\epsilon^2} \ln \frac{2}{\delta}$
  $k \leftarrow |\boldsymbol{Z}|$
  $W \leftarrow 0$
  $M \leftarrow 0$
  while $W < \gamma u^k$
    $\xi, w \leftarrow$ LW-Sample($\mathcal{B}, \boldsymbol{Z} = \boldsymbol{z}$)
    $W \leftarrow W + w$
    $M \leftarrow M + 1$
  return W/M

data-dependent likelihood weighting

For this algorithm, we can provide a similar theoretical analysis with certain guarantees for this *data-dependent likelihood weighting* approach. Algorithm 12.3 shows an algorithm that uses a data-dependent stopping rule to terminate the sampling process when enough weight has been accumulated. We can show that:

Theorem 12.1

Data-Dependent-LW *returns an estimate* $\hat{p}$ *for* $P_{\mathcal{B}}(\boldsymbol{Z} = \boldsymbol{z})$ *which, with probability at least* $1 - \delta$*, has a relative error of* ϵ.

expected sample size

We can also place an upper bound on the *expected sample size* used by the algorithm:

Theorem 12.2

The expected number of samples used by Data-Dependent-LW *is*

$$\frac{u^k}{P_{\mathcal{B}}(\boldsymbol{z})}\gamma \leq \left(\frac{u}{\ell}\right)^k \gamma,$$

where $\gamma = \frac{4(1+\epsilon)}{\epsilon^2} \ln \frac{2}{\delta}$.

The intuition behind this result is straightforward. The algorithm terminates when $W \geq \gamma u^k$. The expected contribution of each sample is $\boldsymbol{E}_{Q(\mathcal{X})}[w(\xi)] = P_{\mathcal{B}}(\boldsymbol{z})$. Thus, the total number of samples required to achieve a total weight of $W \geq \gamma u^k$ is $M \geq \gamma u^k / P_{\mathcal{B}}(\boldsymbol{z})$. Although this bound on the expected number of samples is no better than our bound in equation (12.17), the data-dependent bound allows us to stop early in cases where we were lucky in our random choice of samples, and to continue sampling in cases where we were unlucky.

12.2.3.3 Ratio Likelihood Weighting

ratio likelihood weighting

We now move to the problem of computing a conditional probability $P(\boldsymbol{y} \mid \boldsymbol{e})$ for a specific event $\boldsymbol{y}$. One obvious approach is *ratio likelihood weighting*: we compute the conditional