Algorithm 12.3 Likelihood weighting with a data-dependent stopping rule

**Procedure** Data-Dependent-LW ( 

\[ \mathcal{B}, \quad \text{Bayesian network over } \mathcal{X} \]

\[ \mathcal{Z} = z, \quad \text{Instantiation of interest} \]

\[ u, \quad \text{Upper bound on CPD entries of } \mathcal{Z} \]

\[ \epsilon, \quad \text{Desired error bound} \]

\[ \delta, \quad \text{Desired probability of error} \]

)

1. \[ \gamma \leftarrow 4(1+\epsilon) \frac{\ln 2}{\epsilon^2} \ln \frac{2}{\delta} \]
2. \[ k \leftarrow |\mathcal{Z}| \]
3. \[ W \leftarrow 0 \]
4. \[ M \leftarrow 0 \]
5. While \( W < \gamma u^k \)
6. \[ \xi, w \leftarrow \text{LW-Sample}(\mathcal{B}, \mathcal{Z} = z) \]
7. \[ W \leftarrow W + w \]
8. \[ M \leftarrow M + 1 \]
9. return \( W/M \)

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 \( P_{\mathcal{B}}(\mathcal{Z} = z) \) which, with probability at least \( 1 - \delta \), has a relative error of \( \epsilon \).

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

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

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

The intuition behind this result is straightforward. The algorithm terminates when \( W \geq \gamma u^k \).

The expected contribution of each sample is \( E_{Q(\mathcal{X})}[w(\xi)] = P_{\mathcal{B}}(z) \). Thus, the total number of samples required to achieve a total weight of \( W \geq \gamma u^k \) is \( M \geq \gamma u^k / P_{\mathcal{B}}(z) \). 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**

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