Evaluation Error Requirements For Generating Random Variates Using Dominated Rejection Algorithms

Timothy Hall[*]

Abstract

This paper provides an implementing analyst with an error analysis framework by which the likelihood and extent of improper, missed, proper, and spurious variate candidates may be generated by a dominated rejection algorithm. If the actual distribution from which the variate values (on a bounded real interval) are being produced significantly differs from the intended distribution (as given by a density function), then whatever statistical inference is attempted for the resulting distribution is invalid. The analytical methods found in this paper provide for objectively quantifying the extent to which spurious and missed variate candidates may either be eliminated, minimized, or tolerated.

Key Words: Dominated Rejection Algorithms, Numerical Error Estimation, Random Variates

1. Introduction

When a dominated rejection algorithm is used to generate random variates of a continuous distribution (with a density function) on a (closed) bounded region of real values, the extent to which the calculated continuous uniform variates, and the calculated density function evaluation at one of those continuous uniform variates, introduce evaluation error into the acceptance or rejection of the candidate variate values, could possibly interfere with any statistical inference based on the objectivity of those accepted variate values. If the *actual* distribution from which the variate values are being produced significantly differs from the *intended* distribution (as given by the density function), then whatever statistical inference is attempted for the resulting distribution is invalid. The analytical methods found in this paper provide for objectively quantifying the extent to which spurious (invalid variates due to calculation error) and missed (valid variates left out due to calculation error) variate candidates may either be eliminated, minimized, or tolerated.

This paper provides an implementing analyst with an error analysis framework by which only proper (valid variates rightfully accepted) and improper (invalid variates rightfully rejected) variate candidates occur in a particular Bounded Accuracy Dominated Rejection Algorithm (BADRA) implementation.

2. The BADRA Algorithm

The following steps generate a set of n-many bounded accuracy random variates $\{x_1, x_2, \ldots, x_n\}$ that follow the continuous probability density function $f(x)$ with least upper bound $M > 0$.

Algorithm 1 *(Bounded Accuracy Dominated Rejection Algorithm [BADRA]) Given precision* $tol > 0$*, and the bounded (one-dimensional real) closed set* $[a, b > a]$ *completely within the domain of a continuous probability density function* f *with least upper bound* $M > 0$*, the following steps generate random variates* $\{x_1, x_2, \ldots, x_{n \geq 1}\}$ *on* $[a, b]$ *to within tolerance* tol *that follow the distribution given by* f.

[*]PQI Consulting, P. O. Box 425616, Cambridge, MA, USA, 02142-0012

1. Set $i = 0$.

2. Generate a continuous uniform random variate x on $[a, b]$ to within tolerance tol; this is called the *Candidate Value*.

3. Generate a continuous uniform random variate u on $[0, 1]$ to within tolerance tol independently from Step 2.

4. If $uM < f(x)$, then assign $x_{i+1} = x$ and increment i by 1 (this is called "accepting x"); otherwise continue (this is called "rejecting x").

5. If $i = n$, then skip to Step 6; otherwise skip to Step 2.

6. Report $\{x_1, x_2, \ldots, x_n\}$ as the *Acceptance Set* of bounded accuracy random variates within tolerance tol that follow the distribution given by f.

The most efficient choice of M (the choice that produces the minimum expected number of rejected candidates) is $\sup\limits_{x \in [a,b]} f(x)$ (see Appendix 1).

3. Error Analysis Analytical Environment

Let f_x be the errorless value of f at x, and let f_* be the calculated value of f_x with error $\varepsilon_x \neq 0$ (which could be positive or negative), so that

$$f_x = f_*(x) + \varepsilon_x$$

Generally the value of $|\varepsilon_x|$ is relatively small compared to f_x and absolutely smaller than some invariant value; in fact, it is usually the case that substantial effort is made towards minimizing this value. However, it will be convenient in the analytical development of the error analysis for there to be a specific, if not generous, upper bound for $|\varepsilon_x|$, namely

$$|\varepsilon_x| < \frac{1}{2} \sup\limits_{x \in [a,b]} f(x)$$

Since ε_x may be positive or negative, then $f_*(x)$ may be greater than f_x (when $\varepsilon_x < 0$) or less than f_x (when $\varepsilon_x > 0$), so that it is possible for $f_*(x) > \sup\limits_{x \in [a,b]} f(x)$ and $f_*(x) < 0$, respectively; in fact, we have

$$f_*(x) \in \begin{cases} \left[-\varepsilon_x, \sup\limits_{x \in [a,b]} f(x) - \varepsilon_x\right], & \text{when } \varepsilon_x \text{ is known}[1] \\[2em] \left[-|\varepsilon_x|, \sup\limits_{x \in [a,b]} f(x) + |\varepsilon_x|\right], & \text{when } \varepsilon_x \text{ is unknown} \end{cases}$$

3.1 Calculation Errors

For the purposes of this paper a BADRA implementation will be assumed to be implemented in such a manner that cases where $f_*(x) < 0$ and $f_*(x) > \sup\limits_{x \in [a,b]} f(x)$ shall be detected and eliminated from further consideration, i.e., any candidate x found in Step 1 with a calculated value of $f_*(x)$ found in Step 4 to be in $[-|\varepsilon_x|, 0) \cup \left(\sup\limits_{x \in [a,b]} f(x), \sup\limits_{x \in [a,b]} f(x) + |\varepsilon_x|\right]$.

When $\varepsilon_x > 0$, this means $f_*(x) \in \left[0, \sup_{x \in [a,b]} f(x) - \varepsilon_x\right] \subset \left[0, \sup_{x \in [a,b]} f(x)\right]$, since no value greater than $\sup_{x \in [a,b]} f(x) - \varepsilon_x$ can be calculated. However, when $\varepsilon_x < 0$, this means $f_*(x) \in \left[-\varepsilon_x, \sup_{x \in [a,b]} f(x)\right] \subset \left[0, \sup_{x \in [a,b]} f(x)\right]$, as no value less than $-\varepsilon_x > 0$ can be calculated. When the sign of ε_x is known, the appropriate bounds for $f_*(x)$ shall be used; when the sign of ε_x is unknown, then $f_*(x) \in \left[|\varepsilon_x|, \sup_{x \in [a,b]} f(x) - |\varepsilon_x|\right] \subset \left[0, \sup_{x \in [a,b]} f(x)\right]$ shall be used.

With the corresponding definitions for calculating $M = \sup_{x \in [a,b]} f(x)$, we have

$$M = M_* + \varepsilon_M$$

where ε_M could be non-zero or zero, and for generating u we have

$$u_{ex} = u_{gen} + \varepsilon_u$$

where u_{ex} is the exact value of u in $[0, 1]$, u_{gen} is the calculated value of u, and $\varepsilon_u \neq 0$ is the difference.

In a similar manner as for ε_x, since ε_u may be positive or negative, then u_{gen} may be greater than u_{ex} (when $\varepsilon_u < 0$) or less than u_{ex} (when $\varepsilon_u > 0$), so that it is possible for $u_{gen} > 1$ and $u_{gen} < 0$, respectively; in fact, we have

$$u_{gen} \in \begin{cases} [-\varepsilon_u, 1 - \varepsilon_u], & \text{when } \varepsilon_u \text{ is known}^2 \\ [-|\varepsilon_u|, 1 + |\varepsilon_u|], & \text{when } \varepsilon_u \text{ is unknown} \end{cases}$$

For the purposes of this paper a BADRA implementation will be assumed to be implemented in such a manner that cases where $u_{gen} < 0$ and $u_{gen} > 1$ shall be detected and eliminated from further consideration, i.e., any u_{gen} found in Step 2 with a calculated value to be in $[-|\varepsilon_u|, 0) \cup (1, 1 + |\varepsilon_u|]$.

When $\varepsilon_u > 0$, this means $u_{gen} \in [0, 1 - \varepsilon_u] \subset [0, 1]$, since no value greater than $1 - \varepsilon_u$ can be calculated. However, when $\varepsilon_u < 0$, this means $u_{gen} \in [-\varepsilon_u, 1] \subset [0, 1]$, as no value less than $-\varepsilon_u > 0$ can be calculated. When the sign of ε_u is known, the appropriate bounds for u_{gen} shall be used; when the sign of ε_u is unknown, then $u_{gen} \in [|\varepsilon_u|, 1 - |\varepsilon_u|] \subset [0, 1]$ shall be used.

3.2 Mistaken Accept/Reject

For a particular candidate value x, the test we would *like* to perform in Step 4 of the BADRA, the *Exact Test*, is

$$u_{ex} M < f_x$$

yet the *actual* test that is performed, the *Practical Test*, is

$$u_{gen} M_* < f_*(x)$$

In particular, when the exact test accepts a candidate, that value is said to have been *Exactly Accepted*; otherwise it was *Exactly Rejected*. Similarly a candidate value may be either *Practically Accepted* or *Practically Rejected*. This leads to the following definitions of a *Candidate Class*.

Definition 2 *Given a candidate x generated in Step 2 of the algorithm, the candidate class of x is called ...*

1. *Proper* if it is practically and exactly accepted;

2. *Spurious* if it is practically accepted and exactly rejected;

3. *Missed* if it is practically rejected and exactly accepted; and

4. *Improper* if it is practically and exactly rejected.

Definition 3 *The* Strength *of an implementation of the BADRA is its propensity to generate only proper and improper candidates. A BADRA implementation is called* Perfect *if it only generates proper and improper candidates, and it is called* Useless *if it only generates spurious and missed candidates. The* Accuracy *of a BADRA implementation is measured by* $(|\varepsilon_u|, |\varepsilon_x|)$, *where* Accurate *implementations are near* $(0,0)$ *and* Inaccurate *ones have large* $|\varepsilon_u|$ *and/or large* $|\varepsilon_x|$.

The goal of this paper is to find necessary and sufficient conditions on ε_x and ε_u so that a BADRA implementation may be perfect, i.e., only generate proper and improper candidates regardless of the particular values of ε_x and ε_u. In the alternative, the goal would be to find necessary and sufficient conditions on ε_x and ε_u so that a BADRA implementation may be as "perfect as possible," i.e., be implemented with as much accuracy as possible, which would generate as few spurious and missed candidates, both relatively and absolutely, as possible. If trade-off conditions on ε_x and ε_u are available to minimize particular undesirable candidate classes, i.e., spurious and missing, in favor of promoting particular desirable candidate classes, i.e., proper (especially) and improper, those conditions shall also be a goal of this paper.

3.3 The BADRA Calculation Error

We have

$$u_{ex}M = (u_{gen} + \varepsilon_u)(M_* + \varepsilon_M)$$

so that

$$u_{ex}M = u_{gen}M_* + \left(\underbrace{\varepsilon_u M_* + \varepsilon_M u_{gen} + \varepsilon_u \varepsilon_M - \varepsilon_x}_{\text{The } BADRA \text{ Calculation Error } \varepsilon_C(x)} \right)$$

Note that if $0 > u_{gen}M_* - f_*(x) \geq \varepsilon_x$ and $\varepsilon_x \geq -\varepsilon_C(x)$, then

$$\varepsilon_u M_* + \varepsilon_M u_{gen} + \varepsilon_u \varepsilon_M \geq 0$$

and

$$f_x = f_*(x) + \varepsilon_x \leq u_{gen}M_* \leq u_{gen}M_* + \varepsilon_u M_* + \varepsilon_M u_{gen} + \varepsilon_u \varepsilon_M = u_{ex}M$$

yet

$$u_{gen}M_* < f_*(x)$$

This proves that a candidate x would have been accepted under the practical test ($u_{gen}M_* < f_*(x)$) when it would have been rejected under the exact test ($u_{ex}M \geq f_x$).

Likewise, if $0 \leq u_{gen}M_* - f_*(x) < \varepsilon_x$ and $\varepsilon_x < -\varepsilon_C(x)$, then

$$\varepsilon_u M_* + \varepsilon_M u_{gen} + \varepsilon_u \varepsilon_M < 0$$

and

$$f_x = f_* \left(x \right) + \varepsilon_x > u_{gen} M_* > u_{gen} M_* + \varepsilon_u M_* + \varepsilon_M u_{gen} + \varepsilon_u \varepsilon_M = u_{ex} M$$

yet

$$u_{gen} M_* \geq f_* \left(x \right)$$

This proves that a candidate x would have been rejected under the practical test ($u_{gen} M_* \geq f_* \left(x \right)$) when it would have been accepted under the exact test ($u_{ex} M < f_x$).

These calculation demonstrate that the interplay between $u_{gen} M_* - f_* \left(x \right)$ and $-\varepsilon_C \left(x \right)$ determine the candidate class of x given u_{gen}.

3.4 Static Versus Dynamic Errors

Finally, it is reasonable to expect ε_x and ε_u to be static, i.e., be the same value (yet not necessarily equal) regardless of how many times x and u_{gen} are chosen in Steps 1 and 2, respectively, in a BADRA implementation. If, however, these values depend on the sequence number, or time, or any other external factor, then separate models for ε_x and ε_u under such circumstances would be needed. It is common to accept ε_x and ε_u as static and not necessarily equal, even though they may be functions of the particular values of x and u_{gen} within $[a, b]$ and $[0, 1]$, respectively.

4. Using An Approximate M

Instead of using the calculated value M_* for $M = \sup\limits_{x \in [a,b]} f \left(x \right)$ in the practical tests, suppose we use a rational number $M_0 \geq M$ that is arbitrarily close to M. This value ($M_0 > 0$, since f is a probability density function) is exact since it is rational, and if M were rational, then $\varepsilon_M = 0$. This means $\varepsilon_C \left(x \right)$ would become

$$\varepsilon_C^* \left(x \right) = \varepsilon_u M_0 - \varepsilon_x$$

so that

$$\varepsilon_x < -\varepsilon_C^* \left(x \right) \Longrightarrow \varepsilon_x < \varepsilon_x - \varepsilon_u M_0 \Longrightarrow \varepsilon_u < 0$$
$$\varepsilon_x \geq -\varepsilon_C^* \left(x \right) \Longrightarrow \varepsilon_x \geq \varepsilon_x - \varepsilon_u M_0 \Longrightarrow \varepsilon_u \geq 0$$

since $M_0 > 0$. This establishes the relationship between ε_u and ε_x for a given value of $M_0 \geq M$.

In general, the algorithm efficiency lost by using a larger M value than is necessary is gained by decreasing the likelihood of including spurious candidates in the acceptance set and missed candidates that should have been accepted.

5. Avoiding Spurious Candidates

Criterion 4 *A candidate x is spurious only if $\varepsilon_u M \geq \varepsilon_x$.*

Proof. If candidate x were spurious, then x is practically accepted, which means

$$u_{gen} M_0 < f_* \left(x \right) \Longrightarrow 0 > u_{gen} M_0 - f_* \left(x \right) \geq u_{gen} M - f_* \left(x \right) \tag{1}$$

since $M_0 \geq M$ and[3] $u_{gen} > 0$. We also have that x is exactly rejected, which means

$$u_{ex} M \geq f_x = f_* \left(x \right) + \varepsilon_x \Longrightarrow u_{ex} M - f_* \left(x \right) \geq \varepsilon_x \tag{2}$$

---

[3] When $\varepsilon_u > 0$, then $u_{gen} \in [0, 1 - \varepsilon_u]$, and when $\varepsilon_u < 0$, then $u_{gen} \in [-\varepsilon_u, 1]$. In either case, $u_{gen} > 0$.

Therefore, we have

$$
\begin{aligned}
\varepsilon_x &\leq u_{ex} M - f_*(x) \quad \text{(from (2))} \\
&= (u_{gen} + \varepsilon_u) M - f_*(x) \quad \text{(by definition)} \\
&= (u_{gen} M - f_*(x)) + \varepsilon_u M \\
&\leq \varepsilon_u M \quad \text{(from (1))}
\end{aligned}
$$

∎

This means we may ensure no spurious candidates are generated by enforcing $\varepsilon_u M < \varepsilon_x$. This is called the *Standard Spurious Condition*.

Criterion 5 *A BADRA implementation will not generate any spurious candidates if $\varepsilon_u M < \varepsilon_x$.*

When it is not possible to ensure $\varepsilon_u M < \varepsilon_x$, a spurious candidate may or may not be generated, depending on the relationship between u_{gen}, ε_u, M, M_0, $f_*(x)$, and ε_x. The following theorem makes this relationship explicit.

Theorem 6 *A candidate x is spurious if and only if $0 > u_{gen} M_0 - f_*(x) \geq u_{gen} M - f_*(x) \geq \varepsilon_x - \varepsilon_u M$.*

Proof. Note first that $\varepsilon_u M \geq \varepsilon_x$ is part of this condition; however, a case-by-case component is also needed to make the equivalence.

($\Longrightarrow$) Suppose candidate x were spurious. This means x is practically accepted, i.e.,

$$
u_{gen} M_0 < f_*(x) \Longrightarrow 0 > u_{gen} M_0 - f_*(x) \geq u_{gen} M - f_*(x)
$$

since $M_0 \geq M$ and $u_{gen} \geq 0$. Furthermore, x is exactly rejected, i.e.,

$$
u_{ex} M \geq f_x \Longrightarrow 0 > u_{gen} M - f_*(x) \geq \varepsilon_x - \varepsilon_u M
$$

($\Longleftarrow$) Suppose $0 > u_{gen} M_0 - f_*(x) \geq u_{gen} M - f_*(x) \geq \varepsilon_x - \varepsilon_u M$. Since $0 > u_{gen} M_0 - f_*(x) \Longrightarrow u_{gen} M_0 < f_*(x)$, then x is practically accepted. Furthermore, we have

$$
u_{gen} M \geq f_*(x) + \varepsilon_x - \varepsilon_u M
$$

so that

$$
\begin{aligned}
u_{ex} M &= u_{gen} M + \varepsilon_u M \\
&\geq (f_*(x) + \varepsilon_x - \varepsilon_u M) + \varepsilon_u M \\
&= f_*(x) + \varepsilon_x \\
&= f_x
\end{aligned}
$$

so that x is exactly rejected. ∎

6. Avoiding Missed Candidates

Criterion 7 *A candidate x is missed only if $(1 + \varepsilon_u) M < M_0 + \varepsilon_x$.*

Proof. If candidate x were missed, then x is practically rejected yet exactly accepted. This means

$$u_{gen} M_0 \geq f_* (x) \quad (x \text{ is practically rejected})$$
$$u_{ex} M < f_x \quad (x \text{ is exactly accepted})$$

so that

$$u_{gen} M_0 \geq f_* (x)$$
$$-u_{ex} M > -f_* (x) - \varepsilon_x$$

which means

$$u_{ex} M - u_{gen} M_0 < \varepsilon_x$$

and

$$\begin{aligned}
(1 + \varepsilon_u) M - M_0 &= (M - M_0) + \varepsilon_u M \\
&\leq u_{gen} (M - M_0) + \varepsilon_u M \quad (\text{since } M - M_0 \leq 0, u_{gen} \in [0, 1]) \\
&= u_{ex} M - u_{gen} M_0 \\
&< \varepsilon_x
\end{aligned}$$

∎

This means we may ensure no missed candidates are generated by enforcing $(1 + \varepsilon_u) M \geq M_0 + \varepsilon_x$. This is called the *Standard Missed Condition*.

Criterion 8 *A BADRA implementation will not generate any missed candidates if* $(1 + \varepsilon_u) M \geq M_0 + \varepsilon_x$.

When $(1 + \varepsilon_u) M < M_0 + \varepsilon_x$, a missed candidate may or may not be generated, depending on the relationship between u_{gen}, ε_u, M, M_0, $f_* (x)$, and ε_x. The following theorem makes this relationship explicit.

Theorem 9 *A candidate x is missed if and only if* $0 \leq u_{gen} M_0 - f_* (x)$ *and* $u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$.

Proof. Note how this conditions is "the reverse" of the condition for spurious candidates; however, the relationship between M_0 and M prevents the condition from being stated as a single cascading set of inequalities.

($\Longrightarrow$) Suppose candidate x were missed. This means x is practically rejected, i.e.,

$$u_{gen} M_0 \geq f_* (x) \Longrightarrow 0 \leq u_{gen} M_0 - f_* (x)$$

Furthermore, x is exactly accepted, i.e.,

$$u_{ex} M < f_x \Longrightarrow u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$$

($\Longleftarrow$) Suppose $0 \leq u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$. Since $0 \leq u_{gen} M_0 - f_* (x) \Longrightarrow u_{gen} M_0 \geq f_* (x)$, then x is practically rejected. Furthermore, we have

$$u_{gen} M < f_* (x) + \varepsilon_x - \varepsilon_u M$$

so that

$$\begin{aligned}
u_{ex} M &= u_{gen} M + \varepsilon_u M \\
&< (f_* (x) + \varepsilon_x - \varepsilon_u M) + \varepsilon_u M \\
&= f_* (x) + \varepsilon_x \\
&= f_x
\end{aligned}$$

so that x is exactly accepted. ∎

7. Perfect BADRA Implementations

The standard conditions for preventing spurious and missed candidates may be used individually; however, they cannot be used together.

Corollary 10 *A BADRA implementation using the standard conditions cannot always generate both spurious-free and missed-free candidates.*

Proof. From Criterion 5, the condition $\varepsilon_u M < \varepsilon_x$ avoids generating spurious candidates. So we have

$$(1 + \varepsilon_u) M = M + \varepsilon_u M < M_0 + \varepsilon_x$$

since $M_0 \geq M$. Therefore the condition $(1 + \varepsilon_u) M \geq M_0 + \varepsilon_x$ is not available.

Similarly, Criterion 8 shows the condition $(1 + \varepsilon_u) M \geq M_0 + \varepsilon_x$ avoids generating missed candidates. Then we have

$$\varepsilon_u M \geq (M_0 - M) + \varepsilon_x \geq \varepsilon_x$$

since $M_0 \geq M$. Therefore the condition $\varepsilon_u M < \varepsilon_x$ is not available. $\blacksquare$

However, there are conditions, based on the relationship between u_{gen} and $f_* (x)$, where a BADRA implementation may produce a spurious-free and missed-free candidate.

Claim 11 *A BADRA implementation generates a spurious-free and missed-free candidate if and only if either the $0 > u_{gen} M_0 - f_* (x)$ or $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$ condition applies, but not when both occur at the same time.*

Proof. ($\Longrightarrow$) By Theorem 9, to be missed-free we must have $0 > u_{gen} M_0 - f_* (x)$ or $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$. Furthermore, to be spurious-free we cannot have *both* $0 > u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$, for then

$$0 > u_{gen} M_0 - f_* (x) \geq u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$$

since $M_0 \geq M$ and $u_{gen} \geq 0$, which means x would be spurious by Theorem 6.

($\Longleftarrow$) Suppose $0 > u_{gen} M_0 - f_* (x)$ xor $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$ (using the logical "exclusive or"). Then by Theorem 6, x cannot be spurious since $0 > u_{gen} M_0 - f_* (x) \Longrightarrow u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$ and $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M \Longrightarrow 0 \leq u_{gen} M_0 - f_* (x)$. Furthermore, by Theorem 9, x cannot be missed since either condition contradicts the required conditions. $\blacksquare$

Remark 12 *An important consequence of Claim 11 is that a perfect BADRA implementation must always either have $0 > u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$, or $0 \leq u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$. The first case accepts a candidate practically and exactly (a proper candidate), and the second case rejects a candidate practically and exactly (an improper candidate).*

Claim 13 *Let $\mathcal{R} (f_*)$ be the range of (real) calculated values of f. When $\mathcal{R} (f_*) \cap \left[|\varepsilon_u| M_0, \frac{1}{2} M - |\varepsilon_x| \right] \neq \varnothing$, then there is no single value of $M_0 > M$ that makes a BADRA implementation perfect regardless of ε_u and ε_x.*

Proof. By Claim 11, for a BADRA implementation to be perfect, the choice of M_0 must make either $0 > u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) < \varepsilon_x - \varepsilon_u M$ (to avoid a missed candidate), or $0 \leq u_{gen} M_0 - f_* (x)$ and $u_{gen} M - f_* (x) \geq \varepsilon_x - \varepsilon_u M$ (to avoid a spurious candidate) be valid for all u_{gen} and x, regardless of ε_u and ε_x.

Given $M_0 > M > 0$, for $x_0 \in [a, b]$ such that $f_*(x_0) \in \left[|\varepsilon_u| M_0, \frac{1}{2}M - |\varepsilon_x|\right]$, which exists by supposition, define

$$u^*_{gen} = \frac{f_*(x_0)}{M_0}$$

which means

$$|\varepsilon_u| = \frac{|\varepsilon_u| M_0}{M_0} \leq u^*_{gen} = \frac{f_*(x_0)}{M_0} \leq \frac{\frac{1}{2}M - |\varepsilon_x|}{M_0} < \frac{1}{2} < 1 - |\varepsilon_u|$$

so that

$$u^*_{gen}M - f_*(x_0) = \left(\frac{f_*(x_0)}{M_0}\right)M - f_*(x_0) = \left(\frac{M}{M_0} - 1\right)f_*(x_0) < 0$$

yet

$$u^*_{gen}M_0 - f_*(x_0) = 0$$

Hence, we must have

$$u^*_{gen}M - f_*(x_0) \geq \varepsilon_x - \varepsilon_u M$$

which means $\varepsilon_x - \varepsilon_u M < 0$ since $u^*_{gen}M - f_*(x_0) < 0$.

Now let $\omega \geq 2$ be the smallest positive integer such that

$$0 < \frac{\frac{1}{M}\left(1 - \frac{M}{M_0}\right)f_*(x_0)}{\omega} < M - |\varepsilon_x| - \frac{f_*(x_0)}{M_0}$$

and define

$$u^{**}_{gen} = u^*_{gen} + \frac{1}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0)$$

so that

$$u^{**}_{gen} = \frac{f_*(x_0)}{M_0} + \frac{1}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0) < M - |\varepsilon_x|$$

We have

$$\begin{aligned}
u^{**}_{gen}M - f_*(x_0) &= \left(u^*_{gen} + \frac{1}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0)\right)M - f_*(x_0) \\
&= \left(u^*_{gen}M - f_*(x_0)\right) + \frac{1}{\omega}\left(1 - \frac{M}{M_0}\right)f_*(x_0) \\
&= \left(\frac{M}{M_0} - 1\right)f_*(x_0) + \frac{1}{\omega}\left(1 - \frac{M}{M_0}\right)f_*(x_0) \\
&= \left(1 - \frac{1}{\omega}\right)\left(\frac{M}{M_0} - 1\right)f_*(x_0) \\
&< 0
\end{aligned}$$

yet

$$\begin{aligned}
u^{**}_{gen}M_0 - f_*(x_0) &= \left(u^*_{gen} + \frac{1}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0)\right)M_0 - f_*(x_0) \\
&= \left(u^*_{gen}M_0 - f_*(x_0)\right) + \frac{M_0}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0) \\
&= 0 + \frac{M_0}{\omega M}\left(1 - \frac{M}{M_0}\right)f_*(x_0) \\
&> 0
\end{aligned}$$

which means we cannot have $u^{**}_{gen}M_0 - f_*(x_0) < \varepsilon_x - \varepsilon_u M$ as required, since $\varepsilon_x - \varepsilon_u M < 0$. ∎

Claim 14 *It is always the case that* $\mathcal{R}\left(f_*\right) \cap \left[\left|\varepsilon_u\right| M_0, \frac{1}{2}M - \left|\varepsilon_x\right|\right] \neq \varnothing$ *when* $\left|\varepsilon_x\right| > \left|\varepsilon_u\right| M_0 > \frac{1}{b-a} - \varepsilon_x.$

Proof. Suppose $\mathcal{R}\left(f_*\right) \cap \left[\left|\varepsilon_u\right| M_0, \frac{1}{2}M - \left|\varepsilon_x\right|\right] = \varnothing$. Then

$$\left|\varepsilon_x\right| > \left|\varepsilon_u\right| M_0 \Longrightarrow \mathcal{R}\left(f_*\right) \cap \left[\left|\varepsilon_u\right| M_0, \frac{1}{2}M - \left|\varepsilon_x\right|\right] = \mathcal{R}\left(f_*\right) \cap \left[\left|\varepsilon_x\right|, \frac{1}{2}M - \left|\varepsilon_x\right|\right]$$

This means $\varepsilon_u M_0 + \varepsilon_x < \left|\varepsilon_x\right| + \varepsilon_x \leq f_*\left(x\right) + \varepsilon_x = f_x$ for all $x \in [a, b]$. Hence, we have

$$1 = (b - a)\left(\frac{1}{b - a}\right) < (b - a)\left(\left|\varepsilon_u\right| M_0 + \varepsilon_x\right) < \int_a^b f\left(x\right)\,dx$$

which is a contradiction. $\blacksquare$

Conclusion 15 *A BADRA implementation with* $\left|\varepsilon_x\right| < \frac{1}{2}M$ *and* $\left|\varepsilon_u\right| < \frac{1}{2}$ *and* $\left|\varepsilon_x\right| > \left|\varepsilon_u\right| M_0 > \frac{1}{b-a} - \varepsilon_x$ *cannot be perfect.*

Now consider the possibility that

$$\left|\varepsilon_x\right| \geq \frac{1}{2}M$$

contrary to the intuitive desire to minimize $\left|\varepsilon_x\right|$; the condition $\left|\varepsilon_u\right| < \frac{1}{2}$ remains, and there is now no restriction on the relationship between $\left|\varepsilon_x\right|$ and $\left|\varepsilon_u\right| M_0$, nor between any of these values and a and b.

We have

$$M \geq f_*\left(x\right) + \left|\varepsilon_x\right| \geq f_*\left(x\right) + \frac{1}{2}M \Longrightarrow f_*\left(x\right) \leq \frac{1}{2}M$$

and

$$\varepsilon_x > 0 \Longrightarrow \varepsilon_u M < \frac{1}{2}M \leq \varepsilon_x \Longrightarrow \varepsilon_x - \varepsilon_u M > 0$$

$$\varepsilon_x < 0 \Longrightarrow \varepsilon_u M > -\frac{1}{2}M \geq \varepsilon_x \Longrightarrow \varepsilon_x - \varepsilon_u M < 0$$

so that

$$\varepsilon_x > 0 \Longrightarrow \varepsilon_x - \varepsilon_u M \geq \left(\begin{array}{c} \left\{\begin{array}{c} -\varepsilon_u M \geq -\left|\varepsilon_u\right| M \\ \varepsilon_x \geq \frac{1}{2}M \Longrightarrow \varepsilon_x \geq M - \varepsilon_x \end{array}\right\} \\ \Longrightarrow \varepsilon_x - \varepsilon_u M \geq \left(1 - \left|\varepsilon_u\right|\right) M - \varepsilon_x \end{array}\right) \geq u_{gen} M - f_*\left(x\right)$$

$$\varepsilon_x < 0 \Longrightarrow u_{gen} M - f_*\left(x\right) \geq \left|\varepsilon_u\right| M - \frac{1}{2}M \geq \left(\begin{array}{c} \left\{\begin{array}{c} \left|\varepsilon_u\right| M \geq -\varepsilon_u M \\ -\frac{1}{2}M \geq \varepsilon_x \end{array}\right\} \\ \Longrightarrow \left|\varepsilon_u\right| M - \frac{1}{2}M \geq \varepsilon_x - \varepsilon_u M \end{array}\right)$$

which means when $\varepsilon_x > 0$ all candidates are accepted exactly ($\varepsilon_x - \varepsilon_u M > u_{gen} M - f_*\left(x\right)$) except when $\varepsilon_x - \varepsilon_u M = u_{gen} M - f_*\left(x\right)$, when x would be rejected, and accepted/rejected practically depending on the sign of $u_{gen} M_0 - f_*\left(x\right)$. Furthermore, when $\varepsilon_x < 0$ all candidates are rejected exactly ($\varepsilon_x - \varepsilon_u M \leq u_{gen} M - f_*\left(x\right)$) and again accepted/rejected practically depending on the sign of $u_{gen} M_0 - f_*\left(x\right)$. Note that the particular value of M_0 here is not relevant.

Conclusion 16 *A BADRA implementation with* $\left|\varepsilon_x\right| \geq \frac{1}{2}M$ *and* $\left|\varepsilon_u\right| < \frac{1}{2}$ *cannot be perfect regardless of* M_0.

Finally, consider $|\varepsilon_x| \geq \frac{1}{2}M$ and $|\varepsilon_u| \geq \frac{1}{2}$; there is still no restriction on the relationship between $|\varepsilon_x|$ and $|\varepsilon_u| M_0$, nor between any of these values and a and b.

We still have

$$|\varepsilon_x| \geq \frac{1}{2}M \implies f_*(x) \leq \frac{1}{2}M$$

and

$$|\varepsilon_u| \geq \frac{1}{2} \implies 1 - |\varepsilon_u| \leq \frac{1}{2}$$

so that

$$\frac{1}{2}(M_0 - M) \leq |\varepsilon_u| M_0 - \frac{1}{2}M \leq u_{gen}M_0 - f_*(x) \leq (1 - |\varepsilon_u|) M_0 - |\varepsilon_x| \leq \frac{1}{2}(M_0 - M)$$

which means

$$u_{gen}M_0 - f_*(x) = \frac{1}{2}(M_0 - M) \geq 0$$

since $M_0 \geq M$, which means all candidates are practically rejected.

Conclusion 17 *A BADRA implementation with $|\varepsilon_x| \geq \frac{1}{2}M$ and $|\varepsilon_u| \geq \frac{1}{2}$ only generates improper and missed candidates regardless of M_0.*

Conclusion 18 *A BADRA implementation cannot possibly be perfect unless $|\varepsilon_x| < \frac{1}{2}M$ and $|\varepsilon_u| < \frac{1}{2}$ and either $0 < |\varepsilon_x| \leq |\varepsilon_u| M_0$ or $\varepsilon_x + |\varepsilon_u| M_0 \leq \frac{1}{b-a}$ (or $\varepsilon_u = 0 = \varepsilon_x$).*

Note that the condition $0 < |\varepsilon_x| \leq |\varepsilon_u| M_0$ would allow for $|\varepsilon_u| M_0 \geq \varepsilon_u M \geq \varepsilon_x = |\varepsilon_x| > 0$ when, e.g., $\varepsilon_u > 0$ and $\varepsilon_x > 0$, since $M_0 \geq M$, which, in turn, allows for spurious candidates. Furthermore, it shall usually be the case that a and b are far enough apart, and ε_u and ε_x are small enough values (in absolute value), so that the condition $\varepsilon_x + |\varepsilon_u| M_0 \leq \frac{1}{b-a}$ is easily satisfied.

Therefore, the focus herein is on the two necessary conditions given by

$$|\varepsilon_u| < \frac{1}{2} \text{ and } |\varepsilon_x| < \frac{1}{2}M$$

to characterize a perfect BADRA implementation.

8. Efficient BADRA Implementations

We first need a special measure on the candidate set $[a, b]$ since f is defined outside of $[a, b]$, but all inference must be to the set $[a, b]$ in a BADRA implementation.

Definition 19 *Let $([a, b], \mathcal{S}, \mu)$ be a probability measure space defined on $[a, b]$ with measurable sets $\mathcal{S}$, where μ is given by*

$$\mu(S) = \frac{\int_{x \in S} f(x) \, dx}{\int_a^b f(x) \, dx}$$

for a measurable set $S \in \mathcal{S}$.

Theorem 20 *A BADRA implementation with $\varepsilon_u M < \varepsilon_x$ and $-|\varepsilon_u| M_0 \leq \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$, given u_{gen}, will generate a proper candidate with probability $\mu(S_P)$, where*

$$S_P = \{x \in [a, b] : M \geq f_x > u_{gen}M_0 + \varepsilon_x \geq 0\}$$

and will generate a missed candidate with probability $\mu(S_T)$, where

$$S_T = \left\{x \in [a, b] : M \geq u_{gen}M_0 + \varepsilon_x \geq f_x > \frac{1}{2}(u_{gen}(M + M_0) + \varepsilon_u M + \varepsilon_x) \geq 0\right\}$$

and an improper candidate otherwise with probability $1 - \mu(S_P) - \mu(S_T)$.

Proof. The restrictions on the size of ε_x, given ε_u and M_0, are

$$0 \leq \frac{1}{2} \left(u_{gen} \left(M + M_0 \right) + \varepsilon_u M + \varepsilon_x \right) \leq M \quad \text{and} \quad 0 \leq u_{gen} M_0 + \varepsilon_x \leq M$$

and since $u_{gen} \in [|\varepsilon_u|, 1 - |\varepsilon_u|]$, we have

$$- \left(|\varepsilon_u| + \varepsilon_u \right) M - |\varepsilon_u| M_0 \leq \varepsilon_x \leq M - \left(1 - |\varepsilon_u| \right) M_0 + \left(|\varepsilon_u| - \varepsilon_u \right) M$$

and

$$- |\varepsilon_u| M_0 \leq \varepsilon_x \leq M - \left(1 - |\varepsilon_u| \right) M_0$$

respectively, so that

$$- |\varepsilon_u| M_0 \leq \varepsilon_x \leq |\varepsilon_u| M_0 + \left(M - M_0 \right)$$

satisfies both restrictions.

By Criterion 5, $\varepsilon_u M < \varepsilon_x$ means such a BADRA implementation is spurious-free. Then given $u_{gen} \in [|\varepsilon_u|, 1 - |\varepsilon_u|]$ and $\varepsilon_u M < \varepsilon_x$, we have

$$\frac{1}{2} \left(\varepsilon_u M + \varepsilon_x \right) < \varepsilon_x \quad \text{and} \quad \frac{1}{2} u_{gen} \left(M + M_0 \right) \leq \frac{1}{2} u_{gen} \left(M_0 + M_0 \right) = u_{gen} M_0$$

since $M_0 \geq M$ and $u_{gen} > 0$, so that

$$u_{gen} M_0 + \varepsilon_x > \frac{1}{2} \left(u_{gen} \left(M + M_0 \right) + \varepsilon_u M + \varepsilon_x \right)$$

Hence, we have

$$f_x > u_{gen} M_0 + \varepsilon_x \implies f_x > \frac{1}{2} \left(u_{gen} \left(M + M_0 \right) + \varepsilon_u M + \varepsilon_x \right)$$

which applies if and only if

$$u_{gen} \left(M + M_0 \right) - 2 \left(f_* \left(x \right) + \varepsilon_x \right) < \left(\varepsilon_x - \varepsilon_u M \right) - 2 \varepsilon_x$$

which is valid if and only if

$$u_{gen} M - f_* \left(x \right) + \left(u_{gen} M_0 - f_* \left(x \right) \right) < \varepsilon_x - \varepsilon_u M$$

Since we also have

$$f_x > u_{gen} M_0 + \varepsilon_x \implies u_{gen} M_0 - f_* \left(x \right) < 0$$

then x is practically accepted.

However, we also have

$$u_{gen} M - f_* \left(x \right) \leq u_{gen} M_0 - f_* \left(x \right) < \varepsilon_x - \varepsilon_u M$$

since $M_0 \geq M$ and $u_{gen} > 0$, and $u_{gen} M_0 - f_* \left(x \right)$ is negative and $\varepsilon_x - \varepsilon_u M$ is positive, so that x is exactly accepted. This means x is proper.

Furthermore, we have $f_x \leq u_{gen} M_0 + \varepsilon_x \implies u_{gen} M_0 \geq f_* \left(x \right)$, i.e., x is practically rejected, while $f_x > \frac{1}{2} \left(u_{gen} \left(M + M_0 \right) + \varepsilon_u M + \varepsilon_x \right) \implies u_{gen} M - f_* \left(x \right) \leq u_{gen} M - f_* \left(x \right) + \left(u_{gen} M_0 - f_* \left(x \right) \right) < \varepsilon_x - \varepsilon_u M$, i.e., x is exactly accepted. This means x is missed.

The only remaining choice is an improper candidate (since spurious candidates were eliminated by $\varepsilon_u M < \varepsilon_x$). ∎
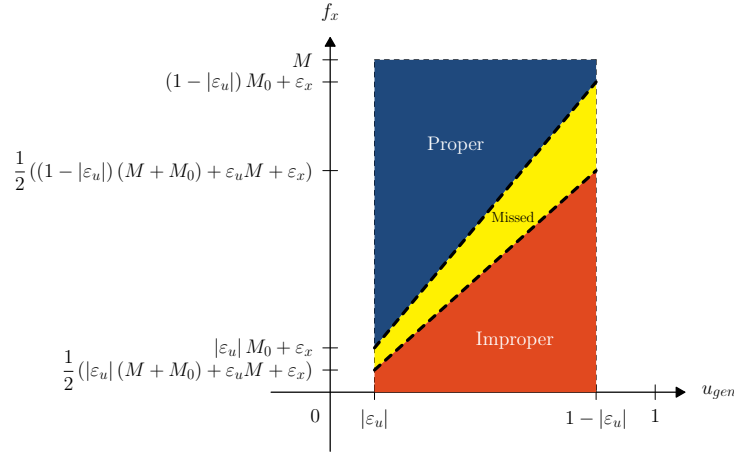
Figure 1: Efficient BADRA Implementation Candidate Classes

Note how the condition $\varepsilon_x + |\varepsilon_u| M_0 \leq \frac{1}{b-a}$ is not explicitly used in the proof of Theorem 20; it is necessary only through Conclusion 18.

A BADRA implementation conforming to the requirements of Theorem 20 is called an *Efficient BADRA Implementation*. See Figure 1 for a depiction of the three possible candidate classes available through Theorem 20.

Note how $\varepsilon_u = 0 = \varepsilon_x$ means $u_{gen} = u_{ex}$, so that $M_0 = M$ would reduce S_P to $\{x \in [a, b] : f_x > u_{ex} M\}$ and S_T to $\{x \in [a, b] : u_{ex} M \geq f_x > u_{ex} M\} = \varnothing$, which means $1 - \mu(S_P) - \mu(S_T) = \mu(\{x \in [a, b] : f_x \leq u_{ex} M\})$.

Note also that the slope of the line dividing the proper and missed regions has slope M_0, which is necessarily no less than the slope of the line dividing the missed and improper regions, namely $\frac{1}{2}(M + M_0)$ (since $M_0 \geq M$). Since these two lines start at $|\varepsilon_u| M_0 + \varepsilon_x$ and $\frac{1}{2}(|\varepsilon_u|(M + M_0) + \varepsilon_u M + \varepsilon_x)$, respectively, i.e., when $u_{gen} = |\varepsilon_u|$, and

$$\left. \begin{array}{r} |\varepsilon_u| M_0 > \frac{1}{2}|\varepsilon_u|(M + M_0) \\ \varepsilon_u M < \varepsilon_x \implies \varepsilon_x > \frac{1}{2}(\varepsilon_u M + \varepsilon_x) \end{array} \right\} \implies |\varepsilon_u| M_0 + \varepsilon_x > \frac{1}{2}(|\varepsilon_u|(M + M_0) + \varepsilon_u M + \varepsilon_x)$$

then these two lines do not intersect.

9. Ensuring Minimally Missed Candidates

The objective of finding necessary and sufficient conditions for a perfect, or "as perfect as possible," BADRA implementation is therefore to choose ε_u, ε_x, and M_0 (to the extent possible in a particular implementation) subject to the conditions of Theorem 20. This will minimize $\mu(S_T)$, thereby making an efficient BADRA implementation "as perfect as possible" when $\varepsilon_u \neq 0$ and $\varepsilon_x \neq 0$.

For non-zero ε_u and ε_x, this minimization would be optimal if we may choose $M_0 = M$ (when the value of M may be errorlessly calculated) and $\varepsilon_u M = \varepsilon_x$, for together these values ensure that

$$-|\varepsilon_u| M_0 = -|\varepsilon_u| M \leq \varepsilon_u M = \varepsilon_x \leq |\varepsilon_u| M = |\varepsilon_u| M_0 = |\varepsilon_u| M_0 + (M - M_0)$$

as required in Theorem 20, and this would completely "squeeze out" the missed region (the middle region in Figure 1) by achieving

$$\frac{1}{2}(u_{gen}(M + M_0) + \varepsilon_u M + \varepsilon_x) = \frac{1}{2}(u_{gen}(M_0 + M_0) + \varepsilon_x + \varepsilon_x) = u_{gen} M_0 + \varepsilon_x$$

Indeed, by Theorem 6, $\varepsilon_u M = \varepsilon_x$ would not admit spurious candidates, since the condition

$$0 \underbrace{\quad > \quad}_{\text{accept practically}} u_{gen} M_0 - f_*(x) = u_{gen} M - f_*(x) \underbrace{\quad \geq \quad}_{\text{reject exactly}} \varepsilon_x - \varepsilon_u M = 0$$

is a contradiction.

However, if M_0 cannot be chosen as M, i.e., we must have $M_0 > M$, then $\varepsilon_u M = \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$ would still not admit spurious candidates (by Theorem 6), and such a choice would still provide

$$- |\varepsilon_u| M_0 < - |\varepsilon_u| M \leq \varepsilon_u M = \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$$

as required in Theorem 20, and it would minimize the missed region (see Figure 1) by uniformly minimizing, i.e., regardless of ε_u and ε_x, the difference between $u_{gen} M_0 + \varepsilon_x$ and $\frac{1}{2} (u_{gen} (M + M_0) + \varepsilon_u M + \varepsilon_x)$, i.e., we would have

$$(u_{gen} M_0 + \varepsilon_x) - \left(\frac{1}{2} (u_{gen} (M + M_0) + \varepsilon_u M + \varepsilon_x) \right) = \frac{1}{2} (u_{gen} (M_0 - M) + \varepsilon_x - \varepsilon_u M)$$

$$= \frac{1}{2} u_{gen} (M_0 - M)$$

Conclusion 21 *With the understanding that $\varepsilon_x + |\varepsilon_u| M_0 \leq \frac{1}{b-a}$, an efficient BADRA implementation is perfect when $M_0 = M$ and $\varepsilon_u M = \varepsilon_x$; otherwise, if M_0 must be chosen greater than M, then an efficient BADRA implementation is as perfect as possible when $\varepsilon_u M = \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$.*

10. Marginal Penalties And Sacrifices

Note also that we may write

$$\frac{1}{2} (u_{gen} (M + M_0) + \varepsilon_u M + \varepsilon_x) = u_{ex} M + \frac{1}{2} (u_{gen} (M_0 - M) + \varepsilon_x - \varepsilon_u M)$$

and

$$u_{gen} M_0 + \varepsilon_x = u_{ex} M + u_{gen} (M_0 - M) + \varepsilon_x - \varepsilon_u M$$

For fixed non-zero ε_u and ε_x, define

$$\xi_{u_{gen}, M_0} = u_{gen} (M_0 - M) + \varepsilon_x - \varepsilon_u M > 0$$

to be the *Marginal Proper Penalty Given u_{gen} And M_0*. Note that ξ_{u_{gen}, M_0} is necessarily positive, since, in an efficient BADRA implementation, we have $\varepsilon_u M < \varepsilon_x$, $u_{gen} \geq 0$, and $M_0 \geq M$. This measures the reduction in the likelihood of proper candidates in an efficient BADRA implementation as a "penalty" for choosing $M_0 > M$ with fixed non-zero ε_u and ε_x.

Claim 22 *In an efficient BADRA implementation, $\xi = \min\limits_{M_0 \geq M} \xi_{u_{gen}, M_0}$ occurs when $M_0 = M$, and at this value we have*

$$S_P = \{ x \in [a, b] : f_x > u_{ex} M + \xi \}$$

$$S_T = \left\{ x \in [a, b] : u_{ex} M + \xi \geq f_x > u_{ex} M + \frac{1}{2} \xi \right\}$$

Proof. Clearly $\xi = \min\limits_{M_0 \geq M} \xi_{u_{gen}, M_0} = \varepsilon_x - \varepsilon_u M$ occurs when $M_0 = M$ for fixed ε_u and ε_x, so that

$$\max \left\{ \begin{array}{c} \frac{1}{2}\left(u_{gen}\left(M + M_0\right) + \varepsilon_u M + \varepsilon_x\right), \\ u_{gen} M_0 + \varepsilon_x \end{array} \right\} \Bigg|_{M_0 = M} = \max \left\{ \begin{array}{c} u_{ex} M + \frac{1}{2}\left(\varepsilon_x - \varepsilon_u M\right), \\ u_{ex} M + \left(\varepsilon_x - \varepsilon_u M\right) \end{array} \right\}$$

$$= \max \left\{ u_{ex} M + \frac{1}{2}\xi, u_{ex} M + \xi \right\}$$

$$= u_{ex} M + \xi$$

since $\varepsilon_u M < \varepsilon_x$ (which means $\xi > 0$) in an efficient BADRA implementation. Therefore, we have

$$S_P = \{x \in [a, b] : f_x > u_{ex} M + \xi\}$$

and

$$S_T = \left\{ x \in [a, b] : u_{ex} M + \xi \geq f_x > u_{ex} M + \frac{1}{2}\xi \right\}$$

∎

Since $f_x > u_{ex} M$ means x is exactly accepted, then define ξ to be the *Marginal Proper Sacrifice*. This measures the reduction in the likelihood of proper candidates in an efficient BADRA implementation that are "sacrificed" by having non-zero ε_u and ε_x.

Note that ξ_{u_{gen}, M_0} and ξ are only defined for efficient BADRA implementations.

The marginal proper sacrifice ξ is therefore the minimum marginal proper penalty available for a given set of non-zero ε_u and ε_x regardless of u_{gen} and $M_0 \geq M$. Note that

$$|\varepsilon_u| \left(M_0 - M\right) \leq \xi_{u_{gen}, M_0} - \xi \leq \left(1 - |\varepsilon_u|\right) \left(M_0 - M\right)$$

by definitions, since $M_0 \geq M$ and $u_{gen} \in [|\varepsilon_u|, 1 - |\varepsilon_u|]$.

Finally, note that when $\varepsilon_u M = \varepsilon_x$ in an efficient BADRA implementation we automatically have a zero marginal proper sacrifice, since then $\xi_{u_{gen}, M_0} = u_{gen}\left(M_0 - M\right)$ so that $\xi = 0$.

11. The Minimum Expected Number Of Improper Candidates

The choice of M that minimizes the expected number of rejections per unit candidates is the value of M that maximizes the expectation of accepting an individual candidate.

If T is the random variable that chooses candidate $\mathbf{t}$ uniformly on $[a, b]$, then $P\left(T = \mathbf{t}\right) = \frac{1}{b-a}$ (since $\int_a^b \frac{1}{b-a} dx = 1$), and if U is the random variable that chooses u uniformly on $[0, 1]$, then $P\left(U \leq \frac{f(\mathbf{t})}{M} \mid T = \mathbf{t}\right) = \frac{f(\mathbf{t})}{M}$, where $M \geq f(\mathbf{t})$ for all $\mathbf{t} \in \mathbf{V}$.

We have

$$P\left(\mathbf{t} \text{ accepted}\right) = P\left(u < \frac{f(\mathbf{t})}{M} \mid \{T = \mathbf{t}, U = u\}\right)$$

$$= P\left(U < \frac{f(\mathbf{t})}{M} \mid T = \mathbf{t}\right) P\left(T = \mathbf{t}\right)$$

$$= \frac{f(\mathbf{t})}{\left(b - a\right) M}$$

Therefore, we have

$$E\left[\mathbf{t} \text{ accepted}\right] = \int_{\mathbf{V}} \frac{f(\mathbf{t})}{\left(b - a\right) M} \mathbf{t} \, d\mathbf{V} = \left(\frac{1}{\left(b - a\right)} \int_{\mathbf{V}} f(\mathbf{t}) \mathbf{t} \, d\mathbf{V}\right) \frac{1}{M} < \infty$$

and viewing $E\left[\mathbf{t} \text{ accepted}\right]$ strictly as a function of M (with fixed a, b, and finite $\int_{\mathbf{V}} f\left(\mathbf{t}\right)\mathbf{t}\, d\mathbf{V}$), we have that $\frac{1}{M}$ is maximized at the smallest available value of M, which is the greatest lower bound for M, i.e.,

$$\max_{M \geq f} E\left[\mathbf{t} \text{ accepted}\right] = \sup_{\mathbf{t} \in \mathbf{V}} f\left(\mathbf{t}\right)$$

Note that these arguments apply to $\mathbf{t}$ and $\mathbf{V}$ of arbitrary (real) dimensions.

12. Example BADRA Implementations

Several calculation examples will demonstrate the dynamics of various BADRA implementations. The first example shows how easily spurious candidates may occur in a BADRA implementation unless specific condition on ε_u and ε_x are imposed. The other six examples use the same $M_0 = \frac{2}{5}$ value justified in the first example for generating standard normal variates on $[-6, 6]$ with various conditions on the two qualifying criteria of Theorem 20 for an efficient BADRA implementation.

For the purposes of the six calculation examples, Criterion I shall be $\varepsilon_u M < \varepsilon_x$ and Criterion II shall be $-\left|\varepsilon_u\right| M_0 \leq \varepsilon_x \leq \left|\varepsilon_u\right| M_0 + \left(M - M_0\right)$ with the understanding that Criterion III ($\varepsilon_x + \left|\varepsilon_u\right| M_0 \leq \frac{1}{b-a}$) shall always be satisfied since the largest absolute value for ε_u and ε_x shall be $\frac{1}{100}$ where $a = -6$ and $b = 6$.

12.1 Spurious Candidates

Our goal in this example is to generate at least $1,000,000$ random variates from the standard normal distribution. This means we must generate at least one million proper candidates x according to the algorithm from the pairs $\left(u_{gen}, f_*\left(x\right)\right)$ *without spurious candidates*.

Suppose we use the approximate[4] M value $M_0 = \frac{2}{5}$, and are using a continuous uniform random variate generator on $[0, 1]$ such that $\varepsilon_u = \frac{1}{1000} > 0$. This means the value of u_{gen} is always slightly smaller (by 0.001) than it should be. As an example, if $u_{ex} = 0.616$ (using a three decimal place generator) comes from a true continuous uniform variate generator, yet $u_{gen} = 0.615$ is used, then $\varepsilon_u = 0.001$.

Suppose further that $\varepsilon_x = -\frac{1}{100} < 0$, so that $f_*\left(x\right)$ is calculated to a value slightly larger (by 0.01) than it should be.

This means[5]

$$\varepsilon_x - \varepsilon_u M < 0$$

so that if we have

$$0 > u_{gen} M_0 - f_*\left(x\right) > u_{gen} M - f_*\left(x\right) \geq \varepsilon_x - \varepsilon_u M$$

then x is a spurious candidate. The exact probability of this happening is therefore[6]

$$P\left(\text{Spurious}\right) = \int_{\left|\varepsilon_u\right|}^{1-\left|\varepsilon_u\right|} \left(\phi^{-1}\left(M\left(x + \varepsilon_u\right)\right) - \phi^{-1}\left(M_0 x + \varepsilon_x\right)\right)\, dx$$

---

[4]Since the standard normal distribution has density function $f\left(x\right)$ given by $\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, then the maximum occurs at its mean value $x = 0$, namely

$$M = \frac{1}{\sqrt{2\pi}} \approx 0.398942 < 0.4 = \frac{2}{5} = M_0$$

This choice for M_0 is very close to M.

[5]Such a BADRA implementation is not efficient by Theorem 20.

[6]This result generalizes to any inverse cumulative probability function.

where ϕ^{-1} is the truncated (to $[-6, 6]$) inverse cumulative distribution function of the standard normal distribution.

With $M_0 = \frac{2}{5}$, $M = \frac{1}{\sqrt{2\pi}}$, $\varepsilon_x = -\frac{1}{100}$, and $\varepsilon_u = \frac{1}{1000}$, whenever u_{gen} and $f_*(x)$ have the relationship[7]

$$\frac{1}{1000}\left(10 + \frac{1}{\sqrt{2\pi}}\right) + \frac{1}{\sqrt{2\pi}}u_{gen} \geq f_*(x) > \frac{2}{5}u_{gen} \tag{3}$$

then x is spurious. As an example, when $u_{gen} = 0.615$, then we have

$$0.25575 \approx \frac{1}{1000}\left(10 + \frac{1}{\sqrt{2\pi}}\right) + \frac{1}{\sqrt{2\pi}}(0.615) \geq f_*(x) > \frac{2}{5}(0.615) = 0.246$$

so that any x where $f_*(x) \in (0.246, 0.255]$, such as when $x = 0.97$ ($f_*(0.97) = f(0.97) + \frac{1}{100} \approx 0.24923 + 0.01 = 0.25023$), is spurious.

The width of the spurious-generating interval from (3) is

$$0.0093423 \lesssim \frac{1}{1000}\left(10 + \frac{1}{\sqrt{2\pi}}\right) + \left(\frac{1}{\sqrt{2\pi}} - \frac{2}{5}\right)u_{gen} \lesssim 0.010398$$

and is shifted upwards according to $\frac{2}{5}u_{gen}$ and $\frac{1}{\sqrt{2\pi}}u_{gen}$, as u_{gen} ranges from $|\varepsilon_u|$ to $1 - |\varepsilon_u|$.

Every time (3) is satisfied for an individually calculated $(u_{gen}, f_*(x))$ pair, then x becomes part of the acceptance set when it should have been rejected. Note also that by the symmetry of the standard normal distribution, if x qualifies as spurious, then so does $-x$.

We may closely approximate the exact probability of a spurious candidate in this example using[8]

$$\int_{\frac{1}{40}}^{1 - \left|\frac{1}{1000}\right|}\left(\phi^{-1}\left(\frac{1}{\sqrt{2\pi}}\left(x + \frac{1}{1000}\right)\right) - \phi^{-1}\left(\frac{2}{5}x - \frac{1}{100}\right)\right)dx \approx 0.8309621$$

where $\frac{1}{40}$ is used instead of $\left|\frac{1}{1000}\right|$ since

$$1 \geq \frac{2}{5}x - \frac{1}{100} > 0 \implies \frac{101}{40} > 1 \geq x > \frac{1}{40} > \left|\frac{1}{1000}\right|$$

and using $1 - \left|\frac{1}{1000}\right|$ as the upper limit is valid since

$$1 \geq \frac{1}{\sqrt{2\pi}}\left(x + \frac{1}{1000}\right) > 0 \implies \sqrt{2\pi} - \frac{1}{1000} > 1 - \left|\frac{1}{1000}\right| \geq x > 0 > -\frac{1}{1000}$$

The remaining part of the exact probability, namely on $x \in \left[\left|\frac{1}{1000}\right|, \frac{1}{40}\right]$, is positive and significantly smaller than the approximation on $\left[\frac{1}{40}, 1 - \left|\frac{1}{1000}\right|\right]$. In particular, its value is approximately 0.1439139, so that the combined value for the entire $\left[\frac{1}{1000}, 1 - \frac{1}{1000}\right]$ interval is approximately 0.9748761.

This means for every $1,000$ candidates the expected number of spurious candidates is approximately $(1000)(0.9748761)/(6 - (-6)) \approx 81$, or 8.1% of all candidates.

---

[7]Since $\varepsilon_x = -\frac{1}{100}$, then we also have

$$\frac{1}{\sqrt{2\pi}}\left(u_{gen} + \frac{1}{1000}\right) \geq f_x = f_*(x) - \frac{1}{100} > \frac{2}{5}u_{gen} - \frac{1}{100}$$

[8]See Appendix 3 for the MAPLE code used to calculate this approximation. MAPLE® is a registered trademark of Maplesoft (a division of Waterloo Maple, Inc.), 615 Kumpf Drive, Waterloo, Ontario, Canada, N2V 1K8. The MAPLE version used to produce the results found in this paper is 2017.1, June, 19, 2017, Maple Build ID 1238644.

12.2 Example 1: Not Efficient Due To Criterion I

Standard normal random variates are chosen from a BADRA implementation (with initial seed 105661067) where $\varepsilon_u > 0$ and $\varepsilon_x < 0$, and $-|\varepsilon_u| M_0 \leq \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$, yet $\varepsilon_u M \geq \varepsilon_x$, where $M = \frac{1}{\sqrt{2\pi}}$ and $M_0 = \frac{2}{5}$. This means Criterion I is not true, even though Criterion II is valid. This is accomplished by setting $\varepsilon_u = -10\varepsilon_x$, for then we always have $\varepsilon_u M \geq \varepsilon_x$ (a positive is always greater than a negative), and

$$-|\varepsilon_u| M_0 = -|-10\varepsilon_x| M_0 = (10\varepsilon_x)\frac{2}{5} = 4\varepsilon_x \leq \varepsilon_x$$

since $\varepsilon_x < 0$, and

$$\frac{M - M_0}{1 + 10M_0} = \frac{\frac{1}{\sqrt{2\pi}} - \frac{2}{5}}{1 + 10\left(\frac{2}{5}\right)} \approx -0.00021154 > -\frac{1}{1000}$$

so that

$$\varepsilon_x \leq -\frac{1}{1000} \implies \varepsilon_x \leq \frac{M - M_0}{1 + 10M_0} \implies \varepsilon_x \leq -10\varepsilon_x M_0 + (M - M_0) = |\varepsilon_u| M_0 + (M - M_0)$$

since $\varepsilon_u > 0$.
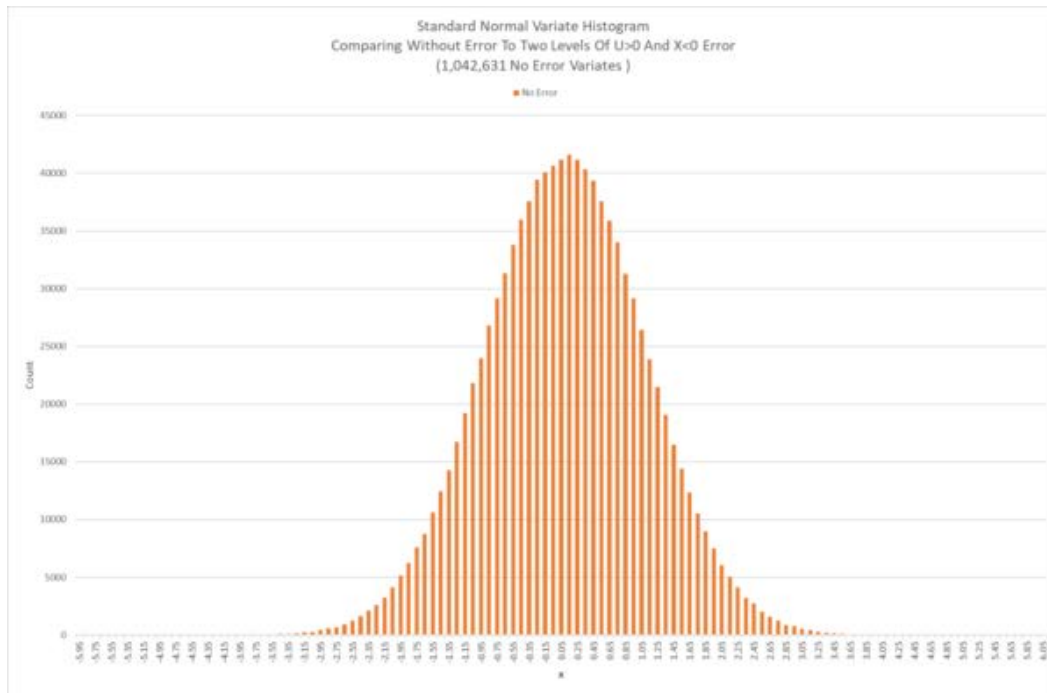
(a) Histogram Of No Error Acceptance Set

(b) Very Large Error ($\varepsilon_u = 0.01$) Differences
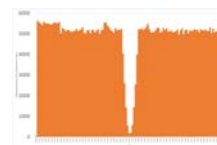
(c) Huge ($\varepsilon_u = 0.1$) Error Differences
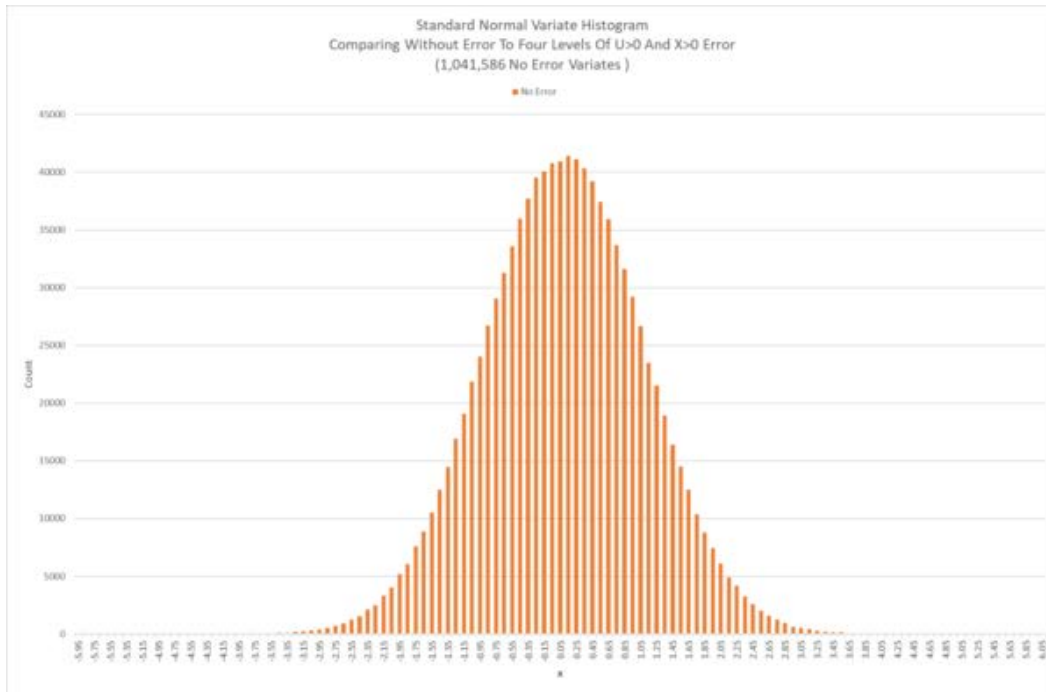
Figure 2: Imperfect BADRA Implementation Due To Criterion I

In particular, two acceptance sets were generated from such a BADRA implementation with $\varepsilon_u = -10\varepsilon_x = \frac{1}{100}$ (Very Large Error) and $\frac{1}{10}$ (Huge Error). In this case, $1,042,631$ proper variates were generated from $5,000,000$ candidates without error ($\varepsilon_u = 0 = \varepsilon_x$). The numbers of proper variates generated with each error value were $1,104,272$ ($61,641 \iff 5.58\%$ spurious) and $1,634,300$ ($591,669 \iff 36.2\%$ spurious), respectively, from $5,000,000$ candidates. Note the large and extremely large, respectively, number/percentage of spurious candidates found in both acceptance sets. Referring to Figure 2, subfigures $(b) - (c)$ demonstrate how the number of spurious candidates (positive count) increases as the size of ε_u increases, and the distribution of the spurious candidates is essentially uniform, so that variates away from the mean are more and more over-represented. There are no missed candidates (negative count) since

$$(1 + \varepsilon_u)\, M \geq \left(1 + \frac{1}{100}\right) \frac{1}{\sqrt{2\pi}} \approx 0.40293 \geq 0.399 = \frac{2}{5} - \frac{1}{1000} \geq M_0 + \varepsilon_x$$

12.3 Example 2: Not Efficient Due To Criterion II

Standard normal random variates are chosen from a BADRA implementation (with initial seed 295013173) where $\varepsilon_u > 0$ and $\varepsilon_x > 0$, and $\varepsilon_u M < \varepsilon_x$, yet $\varepsilon_x > |\varepsilon_u|\, M_0 + (M - M_0)$, where $M = \frac{1}{\sqrt{2\pi}}$ and $M_0 = \frac{2}{5}$. This means Criterion II is not true, even though Criterion I is valid. This is accomplished by setting $\varepsilon_u = \varepsilon_x$, for then we always have $\varepsilon_u M < \varepsilon_x$ (since $M < 1$), and

$$\varepsilon_x > 0 > \frac{\frac{1}{\sqrt{2\pi}} - \frac{2}{5}}{1 - \frac{2}{5}} = \frac{M - M_0}{1 - M_0}$$

which means

$$\varepsilon_x > \varepsilon_x M_0 + (M - M_0) = |\varepsilon_u|\, M_0 + (M - M_0)$$

since $\varepsilon_u > 0$.

In particular, four acceptance sets were generated from such a BADRA implementation with $\varepsilon_u = \varepsilon_x = \frac{1}{1000}$ (Large Error), $\frac{1}{10000}$ (Medium Large Error), $\frac{1}{100000}$ (Small Error), and $\frac{1}{1000000}$ (Very Small Error). In this case, $1,041,586$ proper variates were generated from $5,000,000$ candidates without error ($\varepsilon_u = 0 = \varepsilon_x$). The numbers of proper variates generated with each error value were $1,036,560$ / $1,040,917$ / $1,041,409$ / $1,041,497$, respectively, from $5,000,000$ candidates. Referring to Figure 3, subfigures $(b) - (e)$ demonstrate how the number of missed candidates (negative count) decreases with the size of $\varepsilon_u = \varepsilon_x$, yet they do not completely disappear. Furthermore, the tails of the acceptance set contain the highest portion of the missed candidates, especially when the error is very small. There are no spurious candidates (positive count) as ensured by Criterion I.

12.4 Example 3: Not Efficient Due To Criteria I And II

Standard normal random variates are chosen from a BADRA implementation (with initial seed 84985431) where $\varepsilon_u > 0$ and $\varepsilon_x > 0$, with $\varepsilon_u M \geq \varepsilon_x$ and $\varepsilon_x > |\varepsilon_u|\, M_0 + (M - M_0)$, where $M = \frac{1}{\sqrt{2\pi}}$ and $M_0 = \frac{2}{5}$. This means both Criteria I and II are not true. This is accomplished by setting $\varepsilon_u = \frac{1}{1000}$ and $\varepsilon_x = \gamma \varepsilon_u$, for $\frac{1}{10000} \leq \gamma \leq \frac{1}{10}$, for then we always have
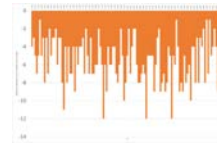
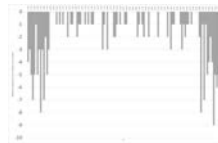$$\varepsilon_u M > \varepsilon_u = \frac{1}{\gamma}\varepsilon_x > \varepsilon_x$$

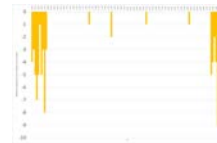(a) Histogram Of No Error Acceptance Set
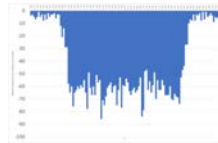
(b) Large Error ($\varepsilon_u = 0.001$) Differences

(c) Medium Large ($\varepsilon_u = 0.0001$) Error Differences

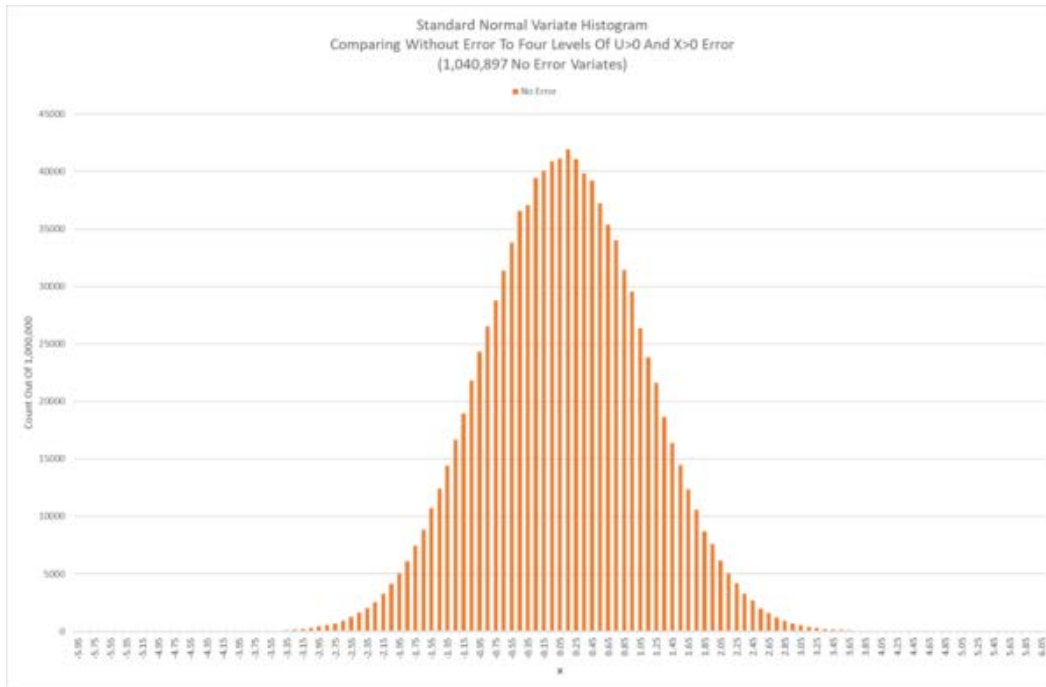(d) Small Error ($\varepsilon_u = 0.00001$) Differences

(e) Very Small Error ($\varepsilon_u = 0.000001$) Differences

Figure 3: Imperfect BADRA Implementation Due To Criterion II

since $\varepsilon_u, \varepsilon_x > 0$, $M > 1$, and $\frac{1}{\gamma} \geq 10$, and

$$\varepsilon_u = \frac{1}{1000} \implies |\varepsilon_u| M_0 + (M - M_0) = \left| \frac{1}{1000} \right| \frac{2}{5} + \left(\frac{1}{\sqrt{2\pi}} - \frac{2}{5} \right)$$
$$\approx -0.00065772$$
$$< \left(\frac{1}{10000} \right) \left(\frac{1}{1000} \right)$$
$$< \gamma \varepsilon_u = \varepsilon_x$$

In particular, four acceptance sets were generated from such a BADRA implementation with $\gamma = \frac{1}{10}$ (Medium Large Error), $\frac{1}{100}$ (Small Error), $\frac{1}{1000}$ (Very Small Error), and $\frac{1}{10000}$ (Tiny Error). In this case, $1,040,897$ proper variates were generated from $5,000,000$

(a) Histogram Of No Error Acceptance Set

(b) Medium Large Error ($\varepsilon_x = 0.0001$) Differences
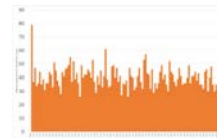
(c) Small ($\varepsilon_x = 0.00001$) Error Differences

(d) Very Small Error ($\varepsilon_x = 0.000001$) Differences

(e) Tiny Error ($\varepsilon_x = 0.0000001$) Differences

Figure 4: Imperfect BADRA Implementation Due To Criteria I And II

candidates without error ($\varepsilon_u = 0 = \varepsilon_x$). The numbers of proper variates generated with each error value were $1,044,553$ / $1,045,695$ / $1,045,824$ / $1,045,868$, respectively, from $5,000,000$ candidates. Referring to Figure 4, subfigures $(b) - (e)$ demonstrate how the spurious candidates are essentially uniformly distributed over $[-6, 6]$ (with notable gaps and peaks are particular values), which means the acceptance sets with non-zero errors have values away from the mean that over-represent the no-error counts.

There are no missed candidates (negative count) in the four calculated acceptance sets.

However, missed candidates are not necessarily excluded (by Criterion 8) since

$$\left(1+\varepsilon_u\right)M = \left(1+\frac{1}{1000}\right)\frac{1}{\sqrt{2\pi}} \approx 0.39934$$

$$< 0.4000001 = \frac{2}{5} + \frac{1}{10000}\left(\frac{1}{1000}\right)$$

$$\leq M_0 + \gamma\varepsilon_u$$

$$= M_0 + \varepsilon_x$$

Furthermore, by Theorem 9, a missed candidate occurs only when $f_*\left(x\right) \leq u_{gen}M_0 = \frac{2}{5}u_{gen}$ and

$$u_{gen}M - f_*\left(x\right) = u_{gen}\frac{1}{\sqrt{2\pi}} - f_*\left(x\right)$$

$$< \varepsilon_x - \varepsilon_u M$$

$$= \gamma\varepsilon_u - \varepsilon_u M$$

$$\leq \left(\gamma - M\right)\varepsilon_u < 0$$

since $\varepsilon_u > 0$ and $\gamma < \frac{1}{10} < M$, so that a superset of missed candidates satisfies $\frac{1}{\sqrt{2\pi}}u_{gen} < f_*\left(x\right) \leq \frac{2}{5}u_{gen}$: $0.398\,94$, which occurs only when $f_*\left(x\right)$ is between $\frac{1}{\sqrt{2\pi}}u_{gen}$ and $\frac{2}{5}u_{gen}$. Since $\frac{2}{5} - \frac{1}{\sqrt{2\pi}} \approx 0.0010577$, the occurrence of such a condition is exceptionally rare.

12.5 Example 4: Efficient Yet Not As Perfect As Possible

Standard normal random variates are chosen from a BADRA implementation (with initial seed 176123807) where $\varepsilon_u < 0$ and $\varepsilon_x > 0$, and $\varepsilon_u M < \varepsilon_x$ and $-\left|\varepsilon_u\right|M_0 \leq \varepsilon_x \leq \left|\varepsilon_u\right|M_0 + \left(M - M_0\right)$, where $M = \frac{1}{\sqrt{2\pi}}$ and $M_0 = \frac{2}{5}$. This is accomplished by setting $\varepsilon_u = -\frac{1}{100}$ and $\varepsilon_x = -\gamma\varepsilon_u$, for $\frac{1}{10000} \leq \gamma \leq \frac{1}{10}$, for then $\varepsilon_u M < 0 < \varepsilon_x$, and

$$-\left|\varepsilon_u\right|M_0 = \varepsilon_u M_0 = -\frac{1}{\gamma}\varepsilon_x M_0 < -\varepsilon_x M_0 < 0 < \left|-\frac{1}{100}\right|\frac{2}{5} + \left(\frac{1}{\sqrt{2\pi}} - \frac{2}{5}\right) = \left|\varepsilon_u\right|M_0 + \left(M - M_0\right)$$

since $-\frac{1}{\gamma} < -1$, $\varepsilon_u < 0$, and $\varepsilon_x > 0$.

In particular, four acceptance sets were generated from such a BADRA implementation with $\gamma = \frac{1}{10}$ (Large Error), $\frac{1}{100}$ (Medium Large Error), $\frac{1}{1000}$ (Small Error), and $\frac{1}{10000}$ (Very Small Error). In this case, $1,042,611$ proper variates were generated from $5,000,000$ candidates without error ($\varepsilon_u = 0 = \varepsilon_x$). The numbers of proper variates generated with each error value were $1,008,366\,/\,1,014,068\,/\,1,014,653\,/\,1,014,705$, respectively, from $5,000,000$ candidates. Referring to Figure 5, subfigures $\left(b\right) - \left(e\right)$ demonstrate how the number of missed candidates (negative count) is essentially constant over most of $\left[-6,6\right]$, excluding the extreme tails, which means the acceptance sets with non-zero errors have values away from the mean that under-represent the no-error counts. There are no spurious candidates (positive count) since $\varepsilon_u M < \varepsilon_x$.
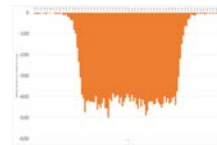
Note the reason there are so many missed candidates in this example is the distance from $\varepsilon_u M$ to ε_x; in this example, they even differ in sign. The closer $\varepsilon_u M$ is to ε_x, while still maintaining $\varepsilon_u M < \varepsilon_x$ (to exclude spurious candidates), the fewer missed candidates would be allowed.
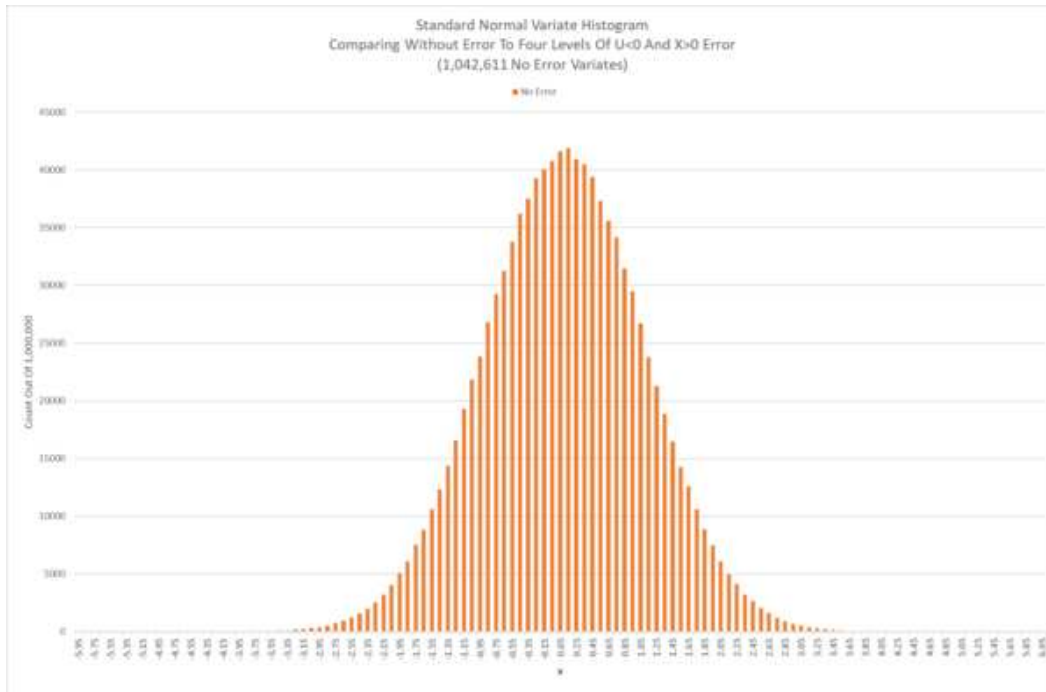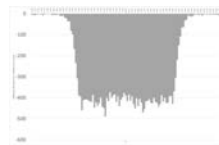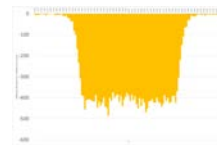
(a) Histogram Of No Error Acceptance Set
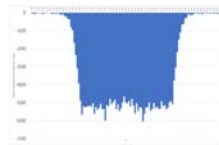
(b) Large Error ($\varepsilon_x = 0.001$) Differences

(c) Medium Large ($\varepsilon_x = 0.0001$) Error Differences

(d) Small Error ($\varepsilon_x = 0.00001$) Differences

(e) Very Small Error ($\varepsilon_x = 0.000001$) Differences

Figure 5: Efficient BADRA Implementation Yet Not As Perfect As Possible

12.6 Example 5: As Perfect As Possible

Standard normal random variates are chosen from a BADRA implementation (with initial seed 420134114) where $\varepsilon_u < 0$ and $\varepsilon_x < 0$, and $\varepsilon_u M = \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$, where $M = \frac{1}{\sqrt{2\pi}}$ and $M_0 = \frac{2}{5}$. This is accomplished by setting $\varepsilon_u = -\frac{1}{100}$ and[9] $\varepsilon_x = \varepsilon_u M$, for then we have
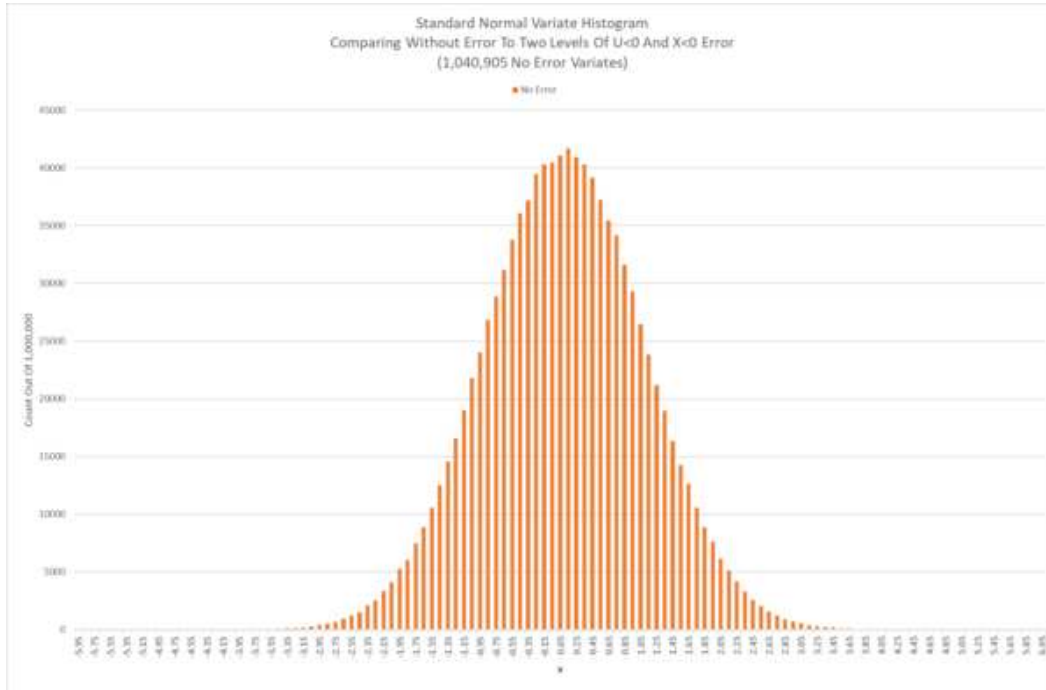
$$\varepsilon_u = -0.01 \leq -0.0013239 \approx \frac{\frac{1}{\sqrt{2\pi}} - \frac{2}{5}}{\frac{1}{\sqrt{2\pi}} + \frac{2}{5}} = \frac{M - M_0}{M + M_0}$$

---

[9]For this example calculation, the IEEE 754 floating point value of $M = \frac{1}{\sqrt{2\pi}}$ used in the BADRA implementation is considered exact.

so that

$$\varepsilon_u \leq \frac{M - M_0}{M + M_0} \implies \varepsilon_u M - |\varepsilon_u| \, M_0 \leq M - M_0 \implies \varepsilon_u M \leq |\varepsilon_u| \, M_0 + (M - M_0)$$
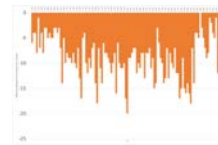
since $\varepsilon_u < 0$.

(a) Histogram Of No Error Acceptance Set

(b) Very Large Error ($\varepsilon_u = -0.01$) Differences

(c) Huge ($\varepsilon_u = -0.1$) Error Differences

Figure 6: Efficient BADRA Implementation As Perfect As Possible

In particular, two acceptance sets were generated from such a BADRA implementation with $\varepsilon_u = -\frac{1}{100}$ (Very Large Error) and $-\frac{1}{10}$ (Huge Error). In this case, $1,040,905$ proper variates were generated from $5,000,000$ candidates without error ($\varepsilon_u = 0 = \varepsilon_x$). The numbers of proper variates generated with each error value were $1,040,683$ and $1,039,851$, respectively, from $5,000,000$ candidates. Referring to Figure 6, subfigures $(b) - (c)$ demonstrate how the number of missed candidates (negative count) is essentially constant over most of $[-6, 6]$, which means the acceptance sets with non-zero errors have values away from the mean that under-represent the no-error counts. There are no spurious candidates (positive count) since $\varepsilon_u M = \varepsilon_x$.

Note that the reason there are still so many missed candidates for an efficient BADRA implementation that is reputedly "as perfect as possible" is the distance between M_0 and M: The closer M_0 is to M, the smaller $\varepsilon_u < 0$ may be in absolute value, which would then allow for fewer missed candidates while still preventing any spurious ones.

12.7 Example 6: Perfect

Standard normal random variates are chosen from a BADRA implementation (with initial seeds 212135135 and 531531212 for negative and positive ε_u, respectively) where ε_u and ε_x have the same sign, and $\varepsilon_u M = \varepsilon_x$, with $M_0 = \frac{1}{\sqrt{2\pi}} = M$.

In particular, two acceptance sets were generated from such a BADRA implementation with $\varepsilon_u = \pm\frac{1}{100}$ (Very Large Error) and $\pm\frac{1}{1000000}$ (Very Small Error). In these cases, $1,043,677$ and $1,044,835$ proper variates were generated from $5,000,000$ candidates without error ($\varepsilon_u = 0 = \varepsilon_x$), corresponding to comparisons to negative and positive ε_u, respectively. The numbers of proper variates generated with each error value were the same as for the no-error acceptance sets (which follows from a "perfect" BADRA implementation). Referring to Figures 7 and 8, subfigures $(b) - (c)$ of each figure demonstrate how there are no spurious candidates (positive count) nor missed candidates (negative count). Note finally that both conditions $\varepsilon_u M = \varepsilon_x$ and $M_0 = M$ are the only required conditions for a perfect BADRA implementation, regardless of the validity of any other pertinent condition required under other BADRA implementations.
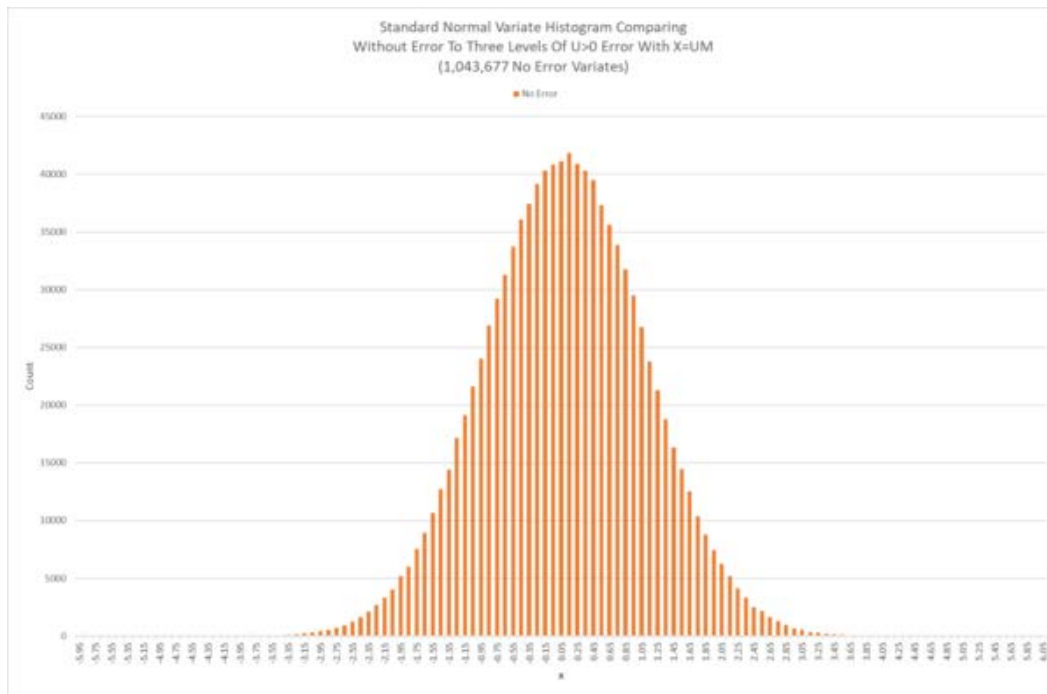
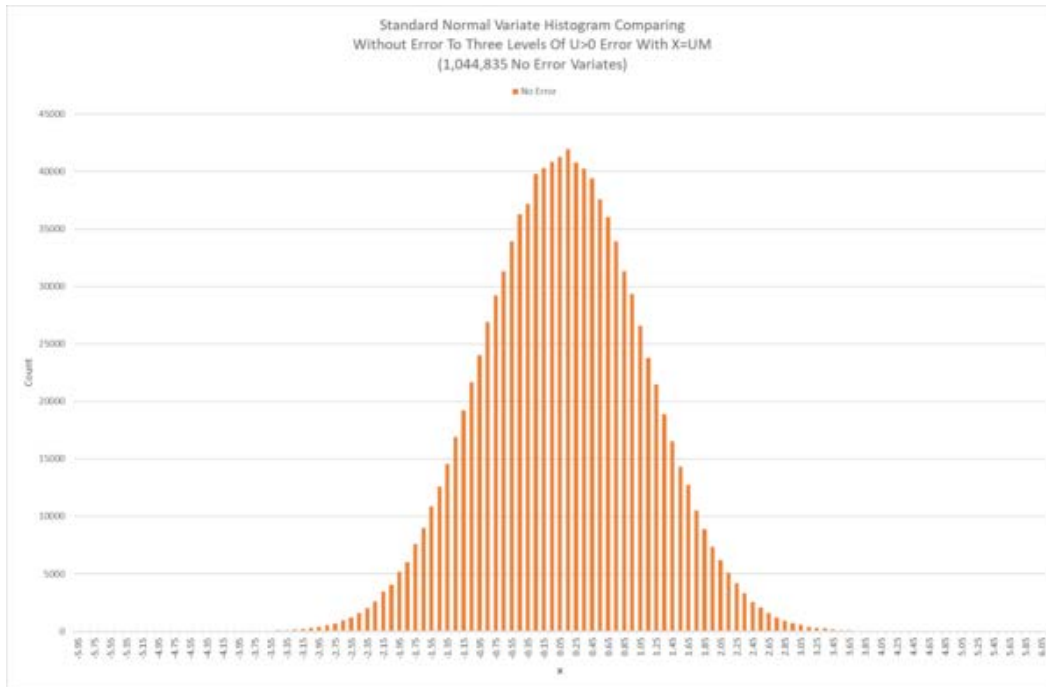
(a) Histogram Of No Error Acceptance Set

(b) Very Large Error ($\varepsilon_u = -0.01$) Differences

(c) Very Small ($\varepsilon_u = -0.000001$) Error Differences

Figure 7: Perfect BADRA Implementation With Negative ε_u (And ε_x)

(a) Histogram Of No Error Acceptance Set

(b) Very Large Error ($\varepsilon_u = 0.01$) Differences

(c) Very Small ($\varepsilon_u = 0.000001$) Error Differences

Figure 8: Perfect BADRA Implementation With Positive ε_u (And ε_x)

12.8 Conclusions From The Examples

The only conditions on the error structure of a BADRA condition that prevents spurious and missed candidates are $\varepsilon_u M = \varepsilon_x$ and $M_0 = M$. If M_0 must be chosen greater than M, then the closer ε_x is to $\varepsilon_u M$ (subject to $\varepsilon_u M < \varepsilon_x$), the fewer missed candidates are introduced into the acceptance set. This is critical in an "efficient yet not as perfect as possible" BADRA implementation, as evidenced in Example 4, where $\varepsilon_u M < \varepsilon_x$ and $-|\varepsilon_u| M_0 \leq \varepsilon_x \leq |\varepsilon_u| M_0 + (M - M_0)$ are valid, yet $\varepsilon_u < 0$ and $\varepsilon_x > 0$ separate the values of $\varepsilon_u M$ and ε_x to the point where many missed candidates are introduced into the acceptance set. Therefore, providing $\varepsilon_u \varepsilon_x > 0$ promotes the minimization of missed candidates in the acceptance set when a BADRA implementation is efficient yet not as perfect as possible.