The Patchwork Rejection Technique for Sampling from Unimodal Distributions *

ERNST STADLOBER and HEINZ ZECHNER

Technical University Graz, Graz, Austria

Dedicated to Professor Ulrich Dieter on the occasion of his 66^{th} *birthday*

19th October 1998

Abstract

We report on both theoretical developments of and computational experience with the *patchwork rejection* technique as studied in [20] and [21]. The basic approach is due to Minh [13] who suggested a special sampling method for the gamma distribution. Its general objective is to rearrange the area below the density or histogram $f(x)$ in the body of the distribution by certain point reflections such that variates may be generated efficiently within a large center interval. This is carried out via uniform hat functions combined with minorizing rectangles for immediate acceptance of one transformed uniform deviate. The remaining tails of $f(x)$ are covered by exponential functions. Experiments show that patchwork rejection algorithms are in general faster than its competitors at the cost of higher set–up times.

Categories and Subject Descriptors: G.3 [**Probability and Statistics**]: Random number generation; G.3 [**Probability and Statistics**]: Statistical Software; I.6.1 [**Simulation and Modeling**]: Simulation Theory

General Terms: Random Variate Transformations, Algorithms, Stochastic Simulation

Additional Key Words and Phrases: Random variate generation, patchwork rejection, sampling techniques, unimodal distributions

1 Introduction

Sampling methods of the acceptance–rejection type are available for most of the well–known continuous and discrete distributions. Competitive procedures have time com-

*This research was supported by the Austrian Science Foundation (FWF), project P8257–PHY.
Authors' address: Lessingstraße 27, A–8010 Graz, Austria. web: http://www.cis.tu-graz.ac.at/stat/
e-mail: <stadlober@stat.tu-graz.ac.at>.

plexity $O(1)$, i. e. they are uniformly fast, and remain efficient even if the (shape) parameters of the target distribution change frequently (see [1], [2], [4], [5], [15], [17], [18]). Performance measures for the valuation of rejection methods are the expected number of uniforms per variate, the set–up time, sampling time (excluding set–up time) and memory requirements.

Suggestions for improvements of the acceptance–rejection methodology are given in [12]. However, its applications are limited to continuous distributions with fixed parameters. Universal rejection algorithms for log–concave discrete distributions are suggested in [3] and [6]. Recently, an automatic generator for an even larger class of unimodal discrete distributions has been proposed in [9]. It is based on the rejection–inversion approach of [8] where the corresponding hat function is constructed by the principle of transformed density rejection [7], having the feature that only one uniform random number is needed per iteration. It is remarkable that all these universal methods have uniformly bounded sampling time for the given class of distributions.

In this paper we discuss a rejection technique for unimodal distributions in which the area below the density or histogram $f(x)$ is rearranged in the center such that sampling with uniform hat functions becomes extremely efficient. The crucial problem is to establish feasible *reflection points* $P_l = (x_2, f(x_2))$ left from the mode m and $P_r = (x_4, f(x_4))$ right from the mode m having two properties: the reflected regions are within the support of $f(x)$ and they contain no points below $f(x)$. General conditions for the abscissas x_2 and x_4 — which change to $k_2 + \frac{1}{2}$ and $k_4 + \frac{1}{2}$ in the discrete case — are the basis of determining suitable choices for specific distributions.

In case of beta(p, q) distributions with $p > 1$, $q > 1$ the parameter space is partitioned into three parts, utilizing optimal (two regions) and nearly optimal (one region) closed form representations of x_2 and x_4. The resulting sampling method has even performance: the expected number of uniforms per variate is bounded by $2(2-e^{-1}) = 3.264241\ldots$ ($p \downarrow 2,\ q \to 1$ or $q \downarrow 2,\ p \to 1$), but for a wide range of pairs (p, q) this figure is well below 2; it decreases to $6/\sqrt{2\pi e} = 1.451824\ldots$ as $p \to \infty$, $q \to \infty$ (see [20]).

In the discrete case specific procedures are developed for Poisson, binomial and hypergeometric distributions with modes $m \geq 4$. It is possible to employ the optimal choices $k_2 \approx m - \sigma$ and $k_4 \approx m + \sigma$ as easily calculated values (σ = standard deviation). Then the expected number of uniforms per variate is rather low; it ranges from 1.31 to 2.65 (see Section 3). The most remarkable improvement is observed in the hypergeometric case where both the sampling time and the average number of uniforms decrease to half of the two published uniformly fast methods [10], [17] and [18].

Section 2 contains the general description of the patchwork technique, including Theorem 1 for feasible reflection centers k_2 in the discrete case and its application to binomial distributions (Lemma 1). A general procedure for sampling from log–concave discrete distributions is stated in Section 3. Computational experience is reported in the final Section 4. Appendix A includes Theorem 2 for feasible reflection centers k_4 as well as Lemma 2 for the hypergeometric case.

2 The Patchwork Technique

2.1 Continuous distributions

Let $f(x)$ be a density which is bounded and unimodal with the mode at m. With the notations of Figure 1 below the basic idea may be described as follows. The left part T_l of the top region under $f(x)$ is transformed into the congruent area T_l' by using a feasible reflection point $P_l = (x_2, f(x_2))$:

$$(x, y) \in T_l \;\mapsto\; (v, w) = (2x_2 - x, 2f(x_2) - y) \in T_l' \,. \tag{1}$$

Similarly the right part T_r is reflected at $P_r = (x_4, f(x_4))$ into T_r':

$$(x, y) \in T_r \;\mapsto\; (v, w) = (2x_4 - x, 2f(x_4) - y) \in T_r' \,. \tag{2}$$

Both abscissas x_2 and x_4 (called *reflection centers*) must be determined such that

(i) T_l' and T_r' are within the support of $f(x)$,

(ii) T_l' and T_r' contain no points below $f(x)$.

Requirement (i) says that the boundaries $x_1 = 2x_2 - m$ and $x_5 = 2x_4 - m$ have to be in the support of $f(x)$, (ii) is equivalent to the following conditions on x_2 and x_4:

$$f(x_2 + z) - f(x_2) \,\le\, f(x_2) - f(x_2 - z)\,, \qquad 0 \le z \le m - x_2 \tag{3}$$

$$f(x_4 - z) - f(x_4) \,\le\, f(x_4) - f(x_4 + z)\,, \qquad 0 \le z \le x_4 - m\,. \tag{4}$$

Inequalities (3) and (4) may be replaced by simpler conditions if $g(x) = \log|f'(x)|$ is differentiable. Then we get *sufficient conditions* for valid reflection centers x_2 and x_4:

$$g'(x_2 + z) + g'(x_2 - z) \,\le\, 0\,, \qquad 0 \le z \le m - x_2 \tag{5}$$

$$g'(x_4 - z) + g'(x_4 + z) \,\ge\, 0\,, \qquad 0 \le z \le x_4 - m\,. \tag{6}$$

Applications

Let us consider the beta(p, q) distribution with density

$$f(x) = f(x; p, q) = \frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}\,, \;\; 0 \le x \le 1,\;\; p > 1,\, q > 1$$

where $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$ is the complete beta function and where the mode is at $m = (p-1)/(p+q-2)$. Condition (5) may be written as

$$A(x_2; p, q)\, z^4 - B(x_2; p, q)\, z^2 + C(x_2; p, q) \le 0\,, \quad 0 \le z \le m - x_2 \tag{7}$$

where the coefficients $A(x; p, q)$, $B(x; p, q)$ and $C(x; p, q)$ are polynomials in x of degrees 1, 3 and 4. Careful analysis of (7) in [20] called for a partition of the (p, q)–plane into three parts having the following closed form reflection centers

$$x_2 = \begin{cases} \frac{m}{2} &= x_2^{min}\,, \quad 1 < p \le 3 \text{ and } x_l \le \frac{m}{2} \\ x_l &= x_2^{min}\,, \quad p > 2,\; x_l \ge \frac{m}{2} \text{ and } (p+q-8)x_l - p + 4 \le 0 \\ \frac{m+z_2}{2} &> x_2^{min}\,, \quad \text{otherwise}\,. \end{cases}$$

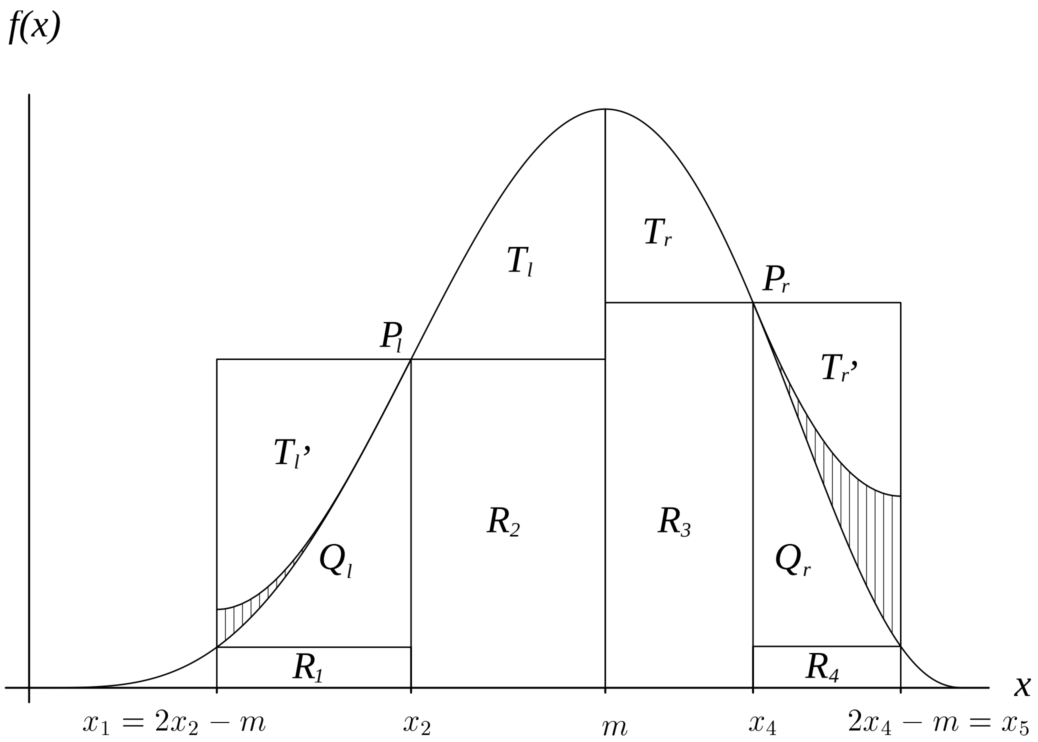

Figure 1: Patchwork technique for continuous distributions.

x_l denotes the left point of inflection $x_l = m - \sqrt{m(1-m)/(p+q-3)}$ if it is real and in $[0,1]$, otherwise $x_l = 0$. $z_2 = x_l - \frac{x_l(1-x_l)}{(p+q-2)(m-x_l)}$ is the point at which the tangent of $f(x)$ to the left point of inflection x_l crosses the x–axis. x_2^{min} is the smallest x_2–value which fulfils (7) and for which $x_1 = 2x_2 - m \geq 0$ is in the support of $f(x)$. The corresponding values $x_4 > m$ are easily obtained from the fact that $f(x;p,q) = f(1-x;q,p)$ by interchanging p and q, replacing m by $1-m$, x_l by $1-x_r$, z_2 by $1-z_4$ and x_2 by $1-x_4$.

2.2 Discrete Distributions

Let $f(x)$ be the histogram function $f(x) = p_k$, $k = \lfloor x \rfloor = 0, 1, \ldots n$ $(n \in \mathbb{N})$ of a unimodal discrete distribution with probabilities $p_k \leq p_m$. The situation is illustrated in Figure 2. Reflections are carried out at the points $P_l = (k_2 + \frac{1}{2}, f(k_2))$ and $P_r = (k_4 + \frac{1}{2}, f(k_4))$. Then the reflection rules (1) and (2) of the continuous case turn over to

$$(x,y) \mapsto (v,w) = (2k_2 + 1 - x, 2f(k_2) - y)\,, \quad k_2 + \tfrac{1}{2} \leq x < m\,, \tag{8}$$

$$(x,y) \mapsto (v,w) = (2k_4 + 1 - x, 2f(k_4) - y)\,, \quad m < x \leq k_4 + \tfrac{1}{2}\,. \tag{9}$$

Note that $k_2 = m-1$ $(k_4 = m)$ is a feasible reflection center if $m \geq 1$ $(n - m \geq 0)$, having the property that $f(x)$ is not altered by (8), respectively (9). Conditions (3) and (4) of the continuous case change to

$$
\begin{aligned}
f(k_2+i)-f(k_2) &\le f(k_2)-f(k_2-i)\,, \quad 1\le i\le m-k_2-1 && (10)\\
f(k_4-i)-f(k_4) &\le f(k_4)-f(k_4+i)\,, \quad 1\le i\le k_4-m\,. && (11)
\end{aligned}
$$

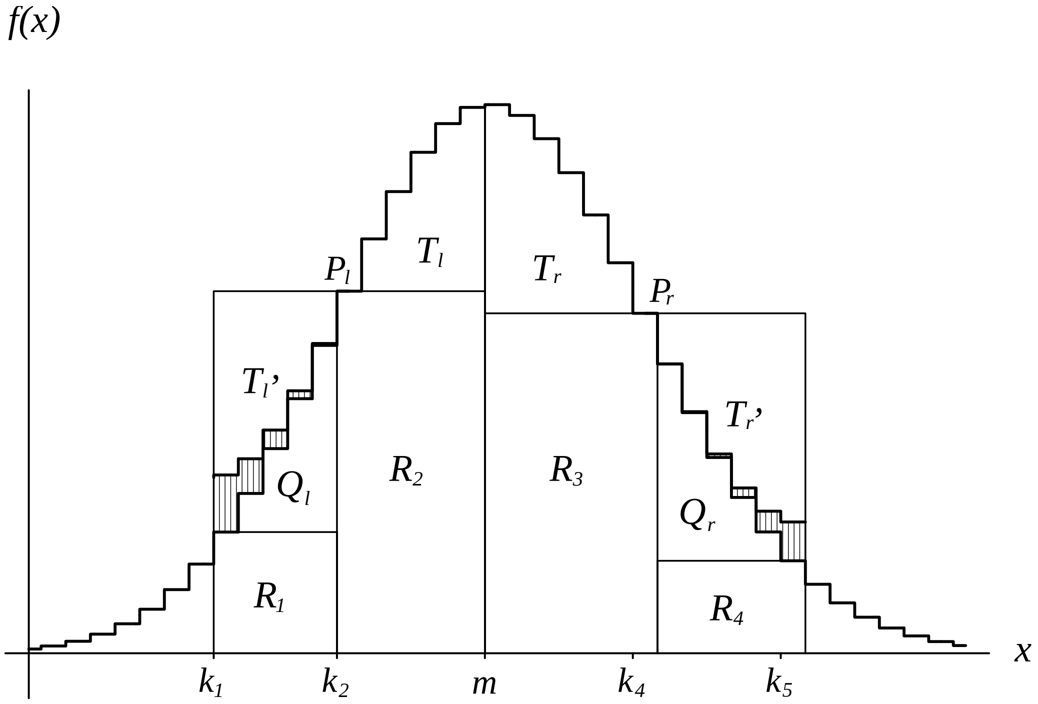

Figure 2: Modification of the histogram by transformation of the top region.

We replace (10) and (11) by inequalities comparing differences of adjoining probabilities

$$
\begin{aligned}
f(k_2+i)-f(k_2+i-1) &\le f(k_2-i+1)-f(k_2-i)\,, \quad 1\le i\le m-k_2-1 && (12)\\
f(k_4-i)-f(k_4-i+1) &\le f(k_4+i-1)-f(k_4+i)\,, \quad 1\le i\le k_4-m\,. && (13)
\end{aligned}
$$

These are *sufficient conditions* for k_2 and k_4 to be valid. For $i=1$, (12) is identical to (10) ((13) conforms to (11)) and this can be used to check whether $k_2=m-2$ ($k_4=m+1$) is valid or not, provided $m\ge 3$ ($n-m\ge 2$). With $r(k)=f(k)/f(k-1)$ and

$$N(k)=r(k+1)\,r(k)-2\,r(k)+1 \tag{14}$$

(12), $i=1$, ((13), $i=1$) can be written as $N(k_2)\le 0$ ($N(k_4)\le 0$). However, this is only a *necessary condition* for $k_2\le m-3$ ($k_4\ge m+2$) to be valid. So, the feasibility of candidates $k_2\le m-3$ ($k_4\ge m+2$) satisfying $N(k_2)\le 0$ ($N(k_4)\le 0$) has to be verified by a *sufficient condition* stated in part 4(b) of Theorem 1 below (Theorem 2 in Appendix A .

Theorem 1. (Valid reflection centers k_2)
Let m be the mode and $N(k)$ as given in (14).

1. *If $m \geq 1$ then there exists at least one valid reflection center $k_2 = m - 1$.*
2. *If $m \geq 3$ and $N(m-2) \leq 0$ then $k_2 = m - 2$ is a valid reflection center.*
3. *If $m \geq 5$ and $N(m-2) \leq 0\,,\ N(m-3) \leq 0$ then there exists a set*
$$I = \{k_2^{min}, \ldots, m-3\} \quad \text{of candidates } k_2 \text{ where}$$
$$k_2^{min} = \min_{k_2} \Big(N(k) \leq 0 \ \text{ for all } k \in \{k_2, \ldots, m-3\} \Big)\,.$$
4. *$k_2 \in I$ is a valid reflection center if the following holds:*
   (a) *$k_2 \geq \frac{m-1}{2}$ ($k_1 = 2k_2 + 1 - m \geq 0$ is in the support of $f(x)$).*
   (b) *$S(i, k_2) \leq 0$ for all $1 \leq i \leq m - k_2 - 2$ where*
$$S(i,k) = r_-(k+i+1)\, r_-(k-i+1)\, r(k+i)\, r(k-i) - r_-(k+i)\, r_-(k-i) \qquad (15)$$
   and $r_-(k) = r(k) - 1$.

Proof. Parts 1, 2 and 3 have been discussed earlier and 4(a) is trivial. It remains to show 4(b). Note that $k_2 \in I$ fulfills (12), $i = 1$, i.e. $N(k_2) \leq 0$ which is equivalent to the initial condition $S(0, k_2) \leq 0$.
Let $i = j$, $1 \leq j \leq m - k_2 - 2$ and assume that k_2 fulfills (12).
Since $r(k) > 1$ for $k \leq m-1$, condition (12) may be re–expressed as

$$f(k_2 + j) \leq \frac{r(k_2+j)}{r(k_2+j)-1} \Big(r(k_2 - j + 1) - 1 \Big) f(k_2 - j)\,. \qquad (16)$$

For $i = j + 1$ we have to check (12) which now reads

$$\Big(r(k_2 + j + 1) - 1 \Big) f(k_2 + j) - \left(1 - \frac{1}{r(k_2 - j)} \right) f(k_2 - j) \leq 0\,. \qquad (17)$$

Since k_2 fulfills (16) we obtain a *stronger condition* than (17) replacing $f(k_2 + j)$ in (17) by the right–hand side of (16). Multiplying the resulting expression by the positive factor $(r(k_2+j)-1)\, r(k_2-j)/f(k_2-j)$ brings us to the requirement

$$S(j, k_2) \leq 0\,, \quad 1 \leq j \leq m - k_2 - 2 \qquad (18)$$

as stated in 4(b). Hence, all $k_2 \in I$ fulfilling $S(j, k_2) \leq 0$ satisfy also (17), but then $f(k_2 + j + 1)$ is bounded according to (16), $j+1$, which leads to condition $S(j+1, k_2) \leq 0$. Thus the following chain of implications holds for $k_2 \leq m - 3$:

$$(18) \Longrightarrow (17) \Longleftrightarrow (12)_{i=j+1} \Longrightarrow (10)_{i=j+1} \Longleftrightarrow k_2 \text{ is a valid reflection center.} \qquad \square$$

Applications

Let us consider the binomial distribution with histogram function

$$f(x) = f(x;n,p) = \binom{n}{\lfloor x \rfloor} p^{\lfloor x \rfloor}(1-p)^{n-\lfloor x \rfloor}\,,\ 0 \le x < n+1, \quad 0 < p < 1,\ n \ge 1\,,$$

and mode at $m = \lfloor \nu \rfloor = \lfloor (n+1)p \rfloor$. Analysis can be restricted to $p \le \frac{1}{2}$, because of $f(x;n,p) = f(n-x;n,1-p)$. The recursive function $r(k)$ reads

$$r(k) = \frac{f(k)}{f(k-1)} = \frac{1}{1-p}\left(\frac{\nu}{k} - p\right)\,,\ \ k = 1,\ldots,n\,. \tag{19}$$

Normalized forms of $N(k)$ (14) and $S(i,k)$ (15) are defined as

$$N^*(k) = (1-p)^2 k(k+1)N(k)\,, \quad S^*(i,k) = (1-p)^4(k^2-i^2)\left((k+1)^2-i^2\right)S(i,k)\,.$$

Now (19) yields

$$N^*(k) = k^2 - (2\nu - 1)\,k + \nu\,(\nu + p - 2) \tag{20}$$

and

$$S^*(i,\,k) = -(1-2p)\,i^4 + B(k)\,i^2 + N^*(k)\,C(k)\,, \tag{21}$$

where

$$B(k) = 2(1-2p)k^2 + 2\Big((3p-1)\nu + (1-2p)\Big)\,k - 2p\nu(\nu-p) + 1 - 2p$$

and

$$C(k) = -(1-2p)k^2 - \Big(2p\nu + (1-2p)\Big)\,k + \nu\,(\nu - p)\,. \tag{22}$$

Lemma 1. (Binomial distribution: $p \le \frac{1}{2}$, $\nu = (n+1)p > 2 + \sqrt{2}$)

(a) All $k_2 \in I = \{k_2^{min},\ldots,m-1\}$ are valid reflection centers with

$$k_2^{min} = \left\lceil \nu - \tfrac{1}{2} - \sqrt{(1-p)\nu + \tfrac{1}{4}}\,\right\rceil \ge \tfrac{m}{2}\,.$$

(b) All $k_4 \in J = \{m,\ldots,k_4^{max}\}$ are valid reflection centers with

$$k_4^{max} = \left\lfloor \nu - \tfrac{1}{2} + \sqrt{(1-p)\nu + \tfrac{1}{4}}\,\right\rfloor \le \tfrac{m+n-1}{2}\,.$$

Proof. Part (a). Reflection center k_2. First it is verified that $N^*(k_2) \le 0$ for all $k_2 \in I\backslash\{m-1\}$ (Theorem 1, part 2 and 3). For this note that $N^*(x)$ (20) is strictly decreasing if $x < \nu - \frac{1}{2}$ and observe

$$N^*(x_2) = 0 \quad \text{for} \quad x_2 = \nu - \tfrac{1}{2} - \sqrt{(1-p)\nu + \tfrac{1}{4}}\,.$$

The restrictions $\nu > 2+\sqrt{2}$ and $0 < p \le \frac{1}{2}$ lead to

$$1 < \nu - \tfrac{1}{2} - \sqrt{\nu + \tfrac{1}{4}} < x_2 \le \nu - \tfrac{1}{2} - \sqrt{\tfrac{1}{2}\nu + \tfrac{1}{4}} \le \lfloor \nu \rfloor - 1 = m-1\,,$$

and it is easy to show that $k_2^{min} \ge \lceil \frac{\nu}{2} \rceil \ge \frac{m}{2}$, i. e. $k_1^{min} = 2k_2^{min} + 1 - m \ge 1$ (Theorem 1.4(a)).

Secondly, the feasibility of $k_2 \in \{k_2^{min}, \ldots, m-3\}$ is proved according to Theorem 1.4(b). For $m=5$ we have $k_2^{min} \ge \frac{m}{2} > m-3$, hence we can assume that $m \ge 6$.
In order to verify $S^*(i,k_2) \le 0$ for $1 \le i \le m - k_2 - 2$, we write (21) as

$$S^*(i,k_2) = R(i,k_2)\, i^2 + N^*(k_2)\, C(k_2)\,.$$

Since $N^*(k_2) \le 0$ for $k_2 \in I\backslash\{m-1\}$, it remains to show that $R(i,k_2) \le 0$ and $C(k_2) \ge 0$. For the proof of the first inequality note that

$$R(i,k_2) = -(1-2p)\, i^2 + B(k_2) \;\le\; R(1,k_2) \quad \text{since } p \le \tfrac{1}{2}\,,$$

and

$$R(1,k_2) = -2(1-2p)\, k_2 \underbrace{(\nu - k_2 - 1)}_{\ge 0} - 2p\nu \underbrace{(\nu - k_2 - p)}_{\ge 0} \le 0\,.$$

By $N^*(k_2) \le 0$ we have $k_2^2 \le (2\nu-1)k_2 - \nu(\nu+p-2)$; applying it to (22) we arrive finally at

$$C(k_2) \ge 2(1-p)\nu \underbrace{\Big(\nu - k_2 - (1-p)\Big)}_{\ge 0} \ge 0$$

which establishes the second inequality and hence part (a).

Part (b). Reflection center k_4**.** To verify that $N^*(k_4) \le 0$ for all $k_4 \in J\backslash\{m\}$ (Appendix A, Theorem 2, part 2 and 3) note that $N^*(x)$ (20) is strictly increasing if $x > \nu - \frac{1}{2}$ and observe

$$N^*(x_4) = 0 \quad \text{for} \quad x_4 = \nu - \tfrac{1}{2} + \sqrt{(1-p)\nu + \tfrac{1}{4}}\,.$$

For $\nu \ge 3$ and due to $0 < p \le \frac{1}{2}$ we have by elementary analysis

$$m+1 = \lfloor \nu \rfloor + 1 \le \nu - \tfrac{1}{2} + \sqrt{\tfrac{1}{2}\nu + \tfrac{1}{4}} \le x_4 < \left\lfloor \tfrac{\nu+n-1}{2} \right\rfloor + 1\,,$$

implying $k_4^{max} \le \frac{m+n-1}{2}$, i. e. $k_5^{max} = 2k_4^{max} - m \le n-1$ (Appendix A, Theorem 2.4(a)).

In order to examine Theorem 2.4(b) we can assume that $n - m \ge 5$. Parallel to part (a) it remains to show that $R(i,k_4) = -(1-2p)i^2 + B(k_4) \ge 0$ and $C(k_4) \le 0$ for $k_4 \in J\backslash\{m\}$. Now we have due to $p \le \frac{1}{2}$

$$R(i,k_4) \ge R(k_4 - m - 1, k_4) \ge R(k_4 - \nu, k_4)\,,\; 1 \le i \le k_4 - m - 1\,,$$

where

$$R(k_4-\nu,k_4) = (1-2p) \underbrace{\Big((k_4-\nu)(k_4+\nu) + 2k_4 + 1\Big)}_{\ge 0} + 2p\nu \underbrace{(k_4 - \nu + p)}_{\ge 0} \ge 0\,,$$

and finally

$$C(k_4) = -(1-2p)k_4\,\underbrace{(k_4-\nu+1)}_{\geq 0} - \nu\,\underbrace{(k_4-\nu+p)}_{\geq 0} \leq 0$$

which completes the proof of Lemma 1. □

Remarks

(i) The restriction $\nu > 2+\sqrt{2}$ guarantees the existence of a left tail ($k_1 \geq 1$) and of a right tail ($k_5 \leq n-1$) which are both parts of the general sampling method PRD in Section 3.

(ii) The Poisson distribution

$$f(x) = f(x;\mu) = \frac{\mu^{\lfloor x \rfloor}}{\lfloor x \rfloor !}\, e^{-\mu}\,, \quad x \geq 0\,,\ \mu > 0\,,$$

is covered as limiting case $n \to \infty$, $p \to 0$, $\mu = np$ fixed. For $\mu > 2+\sqrt{2}$ and $m = \lfloor \mu \rfloor$ the sets of valid reflection centers are

$$\begin{array}{lll} I = \{k_2^{min}, \ldots, m-1\} & \text{where} & k_2^{min} = \left\lceil \mu - \frac{1}{2} - \sqrt{\mu + \frac{1}{4}} \right\rceil \geq \frac{m}{2}\,, \\ J = \{m, \ldots, k_4^{max}\} & \text{where} & k_4^{max} = \left\lfloor \mu - \frac{1}{2} + \sqrt{\mu + \frac{1}{4}} \right\rfloor\,. \end{array}$$

(iii) A similar result (see Appendix A, Lemma 2) holds for the hypergeometric case

$$f(x) = f(x; N, M, n) = \frac{\binom{M}{\lfloor x \rfloor}\binom{N-M}{n-\lfloor x \rfloor}}{\binom{N}{n}}\,, \quad 0 \leq x < n+1\,,\ 7 \leq n \leq M \leq N/2\,.$$

With $\nu = (n+1)p = (n+1)\frac{M+1}{N+2} > 2+\sqrt{2}$, $m = \lfloor \nu \rfloor$ and $t = (1-p)\left(1 - \frac{n+2}{N+3}\right)$ one has

$$\begin{array}{ll} I = \{k_2^{min}, \ldots, m-1\} \text{ where } k_2^{min} = \min\left(\left\lceil \nu - \frac{1}{2} - \sqrt{t\,\nu + \frac{1}{4}} \right\rceil, m-1\right) \geq \frac{m}{2}\,, \\ J = \{m, \ldots, k_4^{max}\} \quad \text{where } k_4^{max} = \left\lfloor \nu - \frac{1}{2} + \sqrt{t\,\nu + \frac{1}{4}} \right\rfloor \leq \frac{m+n-1}{2}\,. \end{array}$$

3 Sampling Method

We describe the sampling technique for discrete distributions only; a thorough discussion of the continuous case may be found in [20]. For convenience, let $f(x) = p_{\lfloor x \rfloor}/p_m \leq 1$ be a normalized histogram. The support of the *reflected* histogram $f(x)$ is split up into three parts (see Figure 2 in Section 2): the left tail ($x < k_1$), the center ($k_1 \leq x \leq k_5+1$) and the right tail ($x > k_5+1$). This suggests a sampling routine based on acceptance–rejection with a uniform–exponential envelope

$$h(x) = \begin{cases} f(k_1)\exp(\lambda_l(x-k_1)) & -\infty < x < k_1 \\ f(k_2) & k_1 \leq x < m \\ f(k_4) & m \leq x \leq k_5+1 \\ f(k_5)\exp(\lambda_r(k_5+1-x)) & k_5+1 < x < \infty\,, \end{cases} \tag{23}$$

where $\lambda_l = \log(f(k_1)/f(k_1-1)) = \log r(k_1)$ and $\lambda_r = -\log r(k_5+1)$, i. e. the exponential tails of $h(x)$ touch $f(x)$ at k_1, k_1-1 and at k_5+1, k_5+2. They majorize $f(x)$ if $r(k) \geq r(k_1)$, $k \leq k_1$, and if $r(k_5+1) \geq r(k)$, $k \geq k_5+1$. Log–concave distributions fulfil these requirements: the condition of log–concavity $p_k^2 \geq p_{k-1}p_{k+1}$ is equivalent to $r(k) \geq r(k+1)$ implying $r(k) \geq r(k+i)$, $i \geq 1$. Exponentials are not appropriate for distributions with heavier tails; in these cases power functions could be employed instead.

Sampling from the center (uniform part of $h(x)$) is substantially improved by the immediate acceptance regions R_1, R_2, R_3 and R_4 under the graph of $f(x)$, displayed in Figure 2. Within $\bigcup R_i$ only one uniform U is needed for one variate K from p_k (see steps 2 and 4 of procedure PRD). If immediate acceptance does not apply in the left center (step 2), then in step 3 a new uniform deviate V is generated and the pair (V, Y) is changed to the random point $(K^*, W) \in C_l = \{(x, w) \mid k_2+1 \leq x < m, f_2 < w < 2f_2 - f_1\}$. If $W \leq f(K^*)$ then (K^*, W) is in the top region T_l of Figure 2, and K^* can be accepted. ($W \leq 1$ is a pre–test, because of $f(x) \leq 1$.) In case of $W > f(K^*)$ the final acceptance test in step 8 is carried out with the transformed point (K, Y): K is accepted whenever $Y \leq f(K)$, i. e. (K, Y) is in Q_l. Otherwise (K, Y) belongs to the shaded area of Figure 2, rejection is indicated, and a new sample from $h(x)$ is required. Additional remarks are included in the formal statement of procedure PRD below.

Procedure PRD

Definition

$f(x)$ `standardized histogram:` $f(x) \leq f(m) = 1$ `where` m `is the mode.`
k_2, k_4 `valid reflection centers,` $r(k) = f(k)/f(k-1)$,
$f_i = f(k_i), i = 1, 2, 4, 5,\quad r_1 = r(k_1),\ r_5 = r(k_5+1)$.

Set–up

`Set` $d_l = m - k_2 - 1$, $d_r = k_4 - m$, $k_1 = k_2 - d_l$, $k_5 = k_4 + d_r$;
`Set` $p_1 = f_2(d_l+1)$, $p_2 = f_2 d_l + p_1$, $p_3 = f_4(d_r+1) + p_2$, $p_4 = f_4 d_r + p_3$;
`Set` $\lambda_l = \log r_1$, $p_5 = f_1/\lambda_l + p_4$, $\lambda_r = -\log r_5$, $p_6 = f_5/\lambda_r + p_5$;

Generation

1. Case Distinction:
   `Generate` $U \sim U(0, p_6)$; `If` $U \leq p_2$ `then Goto 2` [Left Center]
   `else if` $(U \leq p_4)$ `then Goto 4` [Right Center]
   `else if` $(U \leq p_5)$ `then Goto 6` [Left Tail]
   `else Goto 7.` [Right Tail];
2. Left Center — Imm. Accept. of K, $k_1 \leq K \leq m-1$:
   `Set` $V = U - p_1$; $V \sim U(-f_2(d_l+1), +f_2 d_l)$
   `If` $(V < 0)$ `then Deliver` $K = k_2 + \lfloor U/f_2 \rfloor$; $k_2 \leq K \leq m-1$ [Region R_2]
   `Set` $Y = V/d_l$; $Y \sim U(0, f_2)$
   `If` $(Y < f_1)$ `then Deliver` $K = k_1 + \lfloor V/f_1 \rfloor$; $k_1 \leq K \leq k_2 - 1$ [Region R_1]
3. Left Center — Accept. Test of K^*, $k_2+1 \leq K^* \leq m-1$:
   `Generate` $V \sim U(0, 1)$ `and Set` $D = \lfloor d_l V \rfloor + 1$; $1 \leq D \leq d_l = k_2 - k_1$
   `Set` $W = 2f_2 - Y$, $K^* = k_2 + D$; $W \sim U(f_2, 2f_2 - f_1)$
   `If` $(W \leq 1$ `and` $W \leq f(K^*))$ `then Deliver` K^* Acceptance of K^*
   `else Set` $K = k_2 - D$ `and Goto 8.` Rejection of K^*

4. Right Center — Imm. Accept. of K, $m \leq K \leq k_5$:

```
Set V = U − p3;                                   V ∼ U(−f4(dr + 1), +f4 dr)
If (V < 0) then Deliver K = k4 − ⌊(U − p2)/f4⌋;   m ≤ K ≤ k4        [Region R3]
Set Y = V/dr;                                     Y ∼ U(0, f4)
If (Y < f5) then Deliver K = k5 − ⌊V/f5⌋;         k4 + 1 ≤ K ≤ k5   [Region R4]
```

5. Right Center — Accept. Test of K^*, $m \leq K^* \leq k_4 - 1$:

```
Generate V ∼ U(0, 1) and Set D = ⌊dr V⌋ + 1;      1 ≤ D ≤ dr = k5 − k4
Set W = 2 f4 − Y, K* = k4 − D;                    W ∼ U(f4, 2f4 − f5)
If (W ≤ 1 and W ≤ f(K*)) then Deliver K*          Acceptance of K*
    else Set K = k4 + D and Goto 8.               Rejection of K*
```

6. Left Tail — Truncated Exponential K, $0 \leq K \leq k_1 - 1$:

```
Set U = (U − p4)/(p5 − p4) and X = k1 + log U/λl;
If (X < 0) then Goto 1. [Continue];               Fast Rejection
Generate Y ∼ U(0, 1) and Set Y = f1 U Y;          Y ∼ U(0, h(X))
Set K = ⌊X⌋ and Goto 8. [Final acceptance test];
```

7. Right Tail — Truncated Exponential K, $k_5 + 1 \leq K$:

```
Set U = (U − p5)/(p6 − p5) and X = k5 + 1 − log U/λr;
Generate Y ∼ U(0, 1) and Set Y = f5 U Y;          Y ∼ U(0, h(X))
Set K = ⌊X⌋ and Goto 8. [Final acceptance test];
```

8. Final Accept. Test — Accept. of K, whenever $Y \leq f(K)$, $Y \sim U(0,\, h(X))$:

```
If (Y ≤ f(K)) then Deliver K;
               else Goto 1. [Continue];           Rejection of K
```

The performance of PRD may be characterized by three measures. The probability $P(I)$ of accepting K immediately, given that K is accepted as a sample from p_k, can be written as

$$P(I) = p_m\left((k_2 - k_1)f_1 + (m - k_2)f_2 + (k_4 + 1 - m)f_4 + (k_5 - k_4)f_5\right) . \tag{24}$$

The average number α_T, of trials until K is accepted reads

$$\alpha_T = p_m\left(f_1/\lambda_l + (m - k_1)f_2 + (k_5 + 1 - m)f_4 + f_5/\lambda_r\right) .$$

Finally, the expected number of uniforms α_U for one variate from p_k results as

$$\alpha_U = 2\alpha_T - P(I) - p_m\, f_1\, e^{-\lambda_l k_1}/\lambda_l\,, \tag{25}$$

where the last term denotes the probability $P(X < 0)$ of fast rejection in step 6. Values of $P(I)$ and α_U for our binomial and Poisson algorithms BPRS and PPRS [21] are listed as first and second entries of Table 1 below. They are based on the reflection centers $k_2 = k_2^{min}$ and $k_4 = k_4^{max}$ as stated in the Lemma 1, Section 2. For $\mu \geq 5$ the probability $P(I)$ (24) is mostly between .592 (as $\mu \to \infty$) and .784 ($\mu = 5$). The even performance of the methods is notable. Even for small μ the average number of uniforms α_U (25) is always below 2.0 and approaches the limiting normal value $6/\sqrt{2\pi e} = 1.451824\ldots$.

The range of the average number of uniforms α_U (25) were established by extensive numerical search and illustrated by graphical displays. In case of PPRS the value of α_U is between $1.381363\ldots$ ($\mu = 4.354250\ldots$) and $1.773784\ldots$ ($\mu = 4.732051\ldots$). For BPRS the value of α_U is always below $1.987286\ldots$ ($n = 14$, $\mu = 6.05050\ldots$) and approaches its minimum $1.311525\ldots$ at $n = 7$, $\mu = 3.5$. For the hypergeometric generator HPRS ($k_2 = k_2^{min}$ and $k_4 = k_4^{max}$ from Lemma 2, Appendix A) we observed the maximal α_U as $2.642622\ldots$ for the parameter combination $N = 30$, $M = 15$, $n = 15$ ($\mu = 7.5$).

Table 1. Probabilities $P(I)$ and expected number of uniforms α_U per variate.

		Binomial						Poisson
n		10	20	40	100	400	2000	20000	$n \to \infty$
$\mu = 5$	$P(I)$	.732	.784	.773	.766	.763	.762	.762	.762
	α_U	1.841	1.508	1.542	1.560	1.568	1.571	1.571	1.571
$\mu = 10$	$P(I)$		.749	.726	.733	.734	.734	.734	.734
	α_U		1.467	1.428	1.459	1.474	1.478	1.478	1.478
$\mu = 20$	$P(I)$			.701	.703	.709	.710	.710	.691
	α_U			1.433	1.448	1.475	1.483	1.484	1.450
$\mu = 50$	$P(I)$				.651	.642	.646	.658	.659
	α_U				1.423	1.430	1.444	1.446	1.447
$\mu = 200$	$P(I)$					.622	.624	.626	.627
	α_U					1.435	1.445	1.448	1.447
$\mu = 10^3$	$P(I)$						.610	.605	.602
	α_U						1.447	1.448	1.447
$\mu = 10^4$	$P(I)$							.610	.605
	α_U							1.451	1.451
limiting normal case	$P(I)$								.592
	α_U								1.452

Limiting values: $n \to \infty$, $\mu \to \infty$, p fixed

$P(I) \to 2(e^{-1/2} + e^{-2})/\sqrt{2\pi} = .591923\ldots$

$\alpha_U \to 6e^{-1/2}/\sqrt{2\pi} = 1.451824\ldots$

4 Computational Experience

Continuous distributions

Minh [13] compares `Fortran` implementations of his gamma generator with the four point rejection method [16] of similar complexity and reports some improvements in speed. The observed sampling times of our beta method [20] are smaller than those of the competing rejection methods [14] and [15], unless $p \approx 1 + \epsilon$, $q > 4$ or $p > 4$, $q \approx 1 + \epsilon$. The programs coded in `C` have similar set–up times and require about the same space of memory.

Discrete distributions

The general procedure PRD of Section 3 was specialized to specific algorithms for Poisson (PPRS), binomial (BPRS) and hypergeometric (HPRS) distributions. In each algorithm a staircase function $L(x) \leq f(x)$ was added to speed up sampling via fast pretests $W \leq L(K^*)$ in steps 3 and 5, and $Y \leq L(K)$ in step 8 of PRD, thereby avoiding the slow acceptance tests $W \leq f(K^*)$ and $Y \leq f(K)$ most of the time. The details of this squeezing are described in [21]. Efficient implementations in `C` are contained in the package WinRand 1.0/95 [19], a collection of random variate generators for more than thirty discrete and continuous univariate distributions.

The `C` codes of the algorithms were compiled and linked in `Borland Turbo C++/4.5` under Windows NT on a Pentium II–333MHz. Uniform random numbers U were produced by the linear congruential method with large modulus $m = 2^{54}$ and multiplicator $a = 2783377640906189$ which needs 1.7 μsec per U–sample. This generator conforms to the full accuracy of the employed (double precision) arithmetic and has also a good lattice structure. For the evaluation of $\log f(k)$ an external function $\gamma_k = \log k!$ was used. It falls back to tabulated values ($0 \le k \le 29$) or utilizes a Stirling approximation ($k \ge 30$) with truncation error $|\epsilon| < 4.27 \times 10^{-17}$.

The following statements refer to means $\mu \ge 10$; for smaller means inversion methods may be preferred. For the Poisson distribution algorithm PPRS is compared with its competitors PTRD (transformed rejection [4]), PD (acceptance complement [2]), PRUE (ratio of uniforms [17]) and the automatic generator PARI [9] (implementation: log–transformation, with squeeze, no table). The sampling times of PPRS are about 4.0 μsec per variate which is 1.1 to 1.5 times faster than those of PTRD (4.3–6.3 μsec) or those of PD (4.8–5.2 μsec) and about two times faster than those of PRUE (8.2–8.7 μsec). But PPRS is not competitive in case of changing parameters. Here it suffers from its slow set–up: 12–18 μsec in contrast to 1.5–2 μsec for PTRD and only 1 μsec for PD. The universal method PARI samples as fast as PTRD (4.6–6.0 μsec), but has of course the most expensive set–up between 40 and 49 μsec.

Similar conclusions can be drawn in the binomial case. Generator BPRS samples in 3.9–4.3 μsec, and the alternative methods BTRD (transformed rejection [5]), BTPE (triangle–parallelogram–exponential rejection [11]) and BARI (automatic rejection inversion [9]) need 4.4–6.9 μsec, 5.4–10.4 μsec and 4.7–7.2 μsec, respectively. The figures for set–up are 13–23 μsec, 1–2 μsec, 2–3 μsec and 44–66 μsec.

However, a remarkable result is observed in the hypergeometric case where our proposal HPRS compares favourably with H2PE (uniform–exponential rejection [10]), HRUE (ratio of uniforms [17],[18]) and HARI (automatic rejection inversion [9]). It samples with its 4.2–5.8 μsec per variate 1.3 to 2.5 times faster than H2PE (7.3–10.2 μsec), 1.4 to 3.0 times faster than HRUE (8.0–12.6 μsec) and also faster than HARI (4.6–8.0 μsec). The set–up times of HPRS (15–35 μsec) are comparable with those of H2PE (10–24 μsec), but much less than those of HARI (51–89 μsec). It occupies 8032 bytes of compiled code which is slightly less than the 8136 bytes of H2PE and less than the 10098 bytes of HARI. The convenient method HRUE (4832 bytes) has moderate set–up times of 6–14 μsec.

A Appendix

Theorem 2. (Valid reflection centers k_4)
Let m be the mode and $N(k)$ as given in (14).

1. *If $n - m \ge 0$ then there exists at least one valid reflection center $k_4 = m$.*
2. *If $n - m \ge 2$ and $N(m+1) \le 0$ then $k_4 = m + 1$ is a valid reflection center.*
3. *If $n - m \ge 4$ and $N(m+1) \le 0$, $N(m+2) \le 0$ then there exists a set*

$$J = \{m+2, \ldots, k_4^{max}\} \quad \textit{of candidates } k_4 \textit{ where}$$

$$k_4^{max} = \max_{k_4} \Big(N(k) \leq 0 \ \textit{for all} \ k \in \{m+2, \ldots, k_4\} \Big) .$$

4. *$k_4 \in J$ is a valid reflection center if the following holds:*

   (a) *$k_4 \leq \frac{n+m}{2}$ ($k_5 = 2k_4 - m \leq n$ is in the support of $f(x)$).*

   (b) *$S(i, k_4) \geq 0$ for all $1 \leq i \leq k_4 - m - 1$ where $S(i,k)$ from (15).*

Proof. Parts 1, 2 and 3 have been discussed in 2.2 and 4(a) is trivial. The proof of 4(b) follows the same logic as that of Theorem 1.4(b) in Section 2. Note that $k_4 \in J$ satisfies (13), $i = 1$, i.e. $N(k_4) \leq 0$ which is equivalent to the initial condition $S(0, k_4) \geq 0$.
Let $i = j$, $1 \leq j \leq k_4 - m - 1$ and assume that k_4 fulfills (13).
Since $r(k) = f(k)/f(k-1) < 1$ for $k \geq m+1$, condition (13) may be re–expressed as

$$f(k_4 - j) \leq \frac{1 - r(k_4 + j)}{r(k_4 + j)\Big(1 - r(k_4 - j + 1)\Big)} \, f(k_4 + j) \, . \tag{26}$$

For $i = j + 1$ (13) may be written as

$$f(k_4 - j) \left(\frac{1}{r(k_4 - j)} - 1 \right) - f(k_4 + j)\Big(1 - r(k_4 + j + 1)\Big) \leq 0 \, . \tag{27}$$

Since k_4 fulfills (26) we obtain a *stronger condition* than (27) replacing $f(k_4 - j)$ in (27) by the right–hand side of (26). Multiplying the resulting expression by the negative factor $(-1)\,(1 - r(k_4 - j + 1))\, r(k_4 - j)\, r(k_4 + j)/f(k_4 + j)$ yields

$$S(j, k_4) \geq 0 \, , \quad 1 \leq j \leq k_4 - m - 1 \tag{28}$$

as stated in 4(b). Therefore the following chain of implications holds for $k_4 \geq m + 2$:

$$(28) \Longrightarrow (27) \Longleftrightarrow (13)_{i=j+1} \Longrightarrow (11)_{i=j+1} \Longleftrightarrow k_4 \text{ is a valid reflection center.} \qquad \square$$

Application to the Hypergeometric Case

The hypergeometric distribution with parameters N, M and n, where $1 \leq M \leq N$, $1 \leq n \leq N$, has histogram function

$$f(x) = f(x; N, M, n) = \frac{\binom{M}{\lfloor x \rfloor}\binom{N-M}{n - \lfloor x \rfloor}}{\binom{N}{n}} \, , \ \ \max(0, n - N + M) \leq x < \min(n, M) + 1 \, .$$

We can concentrate our attention to the restricted parameter range $1 \leq n \leq M \leq N/2$, using the well-known relationships $f(x; N, M, n) = f(x; N, n, M) = f(n - x; N, N - M, n) = f(M - x; N, M, N - n) = f(N - M - n + x; N, N - M, N - n)$. For convenience we introduce the notation

$$N_2 = N + 2 \, , \quad n_1 = n + 1 \, , \quad p = (M + 1)/N_2 \, , \quad \nu = n_1 p \tag{29}$$

where the restriction $1 \le n \le M \le N/2$ is transformed to

$$0 < p \le \tfrac{1}{2}\,, \quad 2 \le n_1 \le pN_2\,. \tag{30}$$

There is a mode at $m = \lfloor \nu \rfloor$ and the recursive function $r(k)$ reads

$$r(k) = \frac{f(k)}{f(k-1)} = \left(\frac{\nu}{k} - p - \frac{n_1 - k}{N_2}\right) \Big/ \left(1 - p - \frac{n_1 - k}{N_2}\right)\,, \quad k = 1, \ldots, n\,. \tag{31}$$

Normalized forms of $N(k)$ (14) and $S(i,k)$ (15) are defined as

$$N^*(k) = c_N(k)\, N(k)\,, \quad S^*(i,k) = c_S(i,k)\, S(i,k)$$

where the positive constants $c_N(k)$ and $c_S(i,k)$ read as

$$\begin{aligned} c_N(k) \;&=\; \left(1 - p - \frac{n_1 - k}{N_2}\right)\left(1 - p - \frac{n_1 - k - p}{N_2 + 1}\right) k(k+1) \\ c_S(i,\,k) \;&=\; (N_2+1)(k+i+1)(k+i)(k-i+1)(k-i) \\ &\qquad \times \left(1 - p - \frac{n_1 - (k+i+p)}{N_2+1}\right)\left(1 - p - \frac{n_1 - (k+i)}{N_2}\right) \\ &\qquad \times \left(1 - p - \frac{n_1 - (k-i+p)}{N_2+1}\right)\left(1 - p - \frac{n_1 - (k-i)}{N_2}\right). \end{aligned}$$

Using (31) we get after some tedious, but straightforward calculations

$$N^*(k) = k^2 - (2\nu - 1)\,k + \nu\left(\nu + p - 2 + \frac{n_1 - \nu - p + 1}{N_2 + 1}\right) \tag{32}$$

and

$$S^*(i,\,k) = A(k)\, i^4 + B(k)\, i^2 + N^*(k)\, C(k)\,, \tag{33}$$

with

$$A(k) = \,-\,2k \,+\, 2\nu \,-\, 1 \,-\, (1-2p)(N_2 - 2n_1)\,, \tag{34}$$

$$\begin{aligned} B(k) \;=\;& \quad 4k^3 \,+\, 2\Big((1-2p)N_2 - 2n_1 - 2\nu + 3\Big)k^2 \\ &+\, 2\bigg(\Big((3p-1)\nu + 1 - 2p\Big)N_2 + 3\nu n_1 - \nu^2 \\ &\qquad - 2n_1 - (3p-1)\nu + 2 - \frac{3\nu(n_1 - \nu + 1 - p)}{N_2+1}\bigg)k \\ &-\, \Big(2p\nu^2 - 2p^2\nu - (1-2p)\Big)N_2 - 2\nu^2 n_1 + 2\nu^3 + 2\nu^2 \\ &\qquad - 2n_1 \,+\, 2(1-p^2)\nu \,+\, 1 \,+\, \frac{2\nu(n_1^2 - \nu^2 - (1-p^2))}{N_2+1} \end{aligned} \tag{35}$$

and

$$\begin{aligned} C(k) \;=\; & -2k^3 - \Big((1-2p)\,N_2 - 2n_1 - 2\nu + 3\Big)k^2 \\ & - \Big((2p\nu + 1 - 2p)\,N_2 + 2\nu n_1 - 2n_1 - 2\nu + 1\Big)k \\ & + \nu(\nu - p)\,N_2 \;-\; \nu n_1 \;+\; \nu\,. \end{aligned} \tag{36}$$

Lemma 2. (Hypergeometric distribution: $7 \le n \le M \le N/2\,,\ \nu > 2+\sqrt{2}$)

(a) All $k_2 \in I = \{k_2^{min}, \dots, m-1\}$ are valid reflection centers with

$$k_2^{min} = \min\Big(\lceil x_2 \rceil, m-1\Big) \ge \tfrac{m}{2} \quad \text{with } x_2 = \nu - \tfrac{1}{2} - \sqrt{(1-p)\left(1 - \tfrac{n+2}{N+3}\right)\nu + \tfrac{1}{4}}\,,$$

(b) All $k_4 \in J = \{m, \dots, k_4^{max}\}$ are valid reflection centers with

$$k_4^{max} = \lfloor x_4 \rfloor \le \tfrac{m+n-1}{2} \quad \text{with } x_4 = \nu - \tfrac{1}{2} + \sqrt{(1-p)\left(1 - \tfrac{n+2}{N+3}\right)\nu + \tfrac{1}{4}}\,.$$

Proof. Part (a). Reflection center k_2. To examine $N^*(k_2) \le 0$ for all $k_2 \in I \backslash \{m-1\}$ (Theorem 1, part 2 and 3 in Section 2) we mention that $N^*(x)$ (32) is strictly decreasing if $x < \nu - \frac{1}{2}$ and $N^*(x_2) = 0$. It is easy to see that $x_2 \le \nu - 1$, implying $\lceil x_2 \rceil \le m$, but $\lceil x_2 \rceil \le m-1$ does not hold for all parameter combinations, so we have to define $k_2^{min} = \min(\lceil x_2 \rceil, m-1)$. With the restriction $\nu > 2+\sqrt{2}$ by considering $0 < p \le \frac{1}{2}$, and noting that the corresponding $k_2^{min} = k_2^{min}(\nu, p)$ of the binomial case (see Lemma 1 (a) in Section 2) is less equal $\lceil x_2 \rceil$ we obtain $\lceil x_2 \rceil \ge \frac{m}{2}$ and consequently $k_1^{min} \ge 1$.

For the proof of the feasibility of $k_2 \in \{k_2^{min}, \dots, m-3\}$ (Theorem 1.4(b) in Section 2) we can assume $m \ge 6$. Because of (33) and $N^*(k_2) \le 0$ it remains to show
(i) $R(i,k_2) = A(k_2)i^2 + B(k_2) \le 0$ for $1 \le i \le m - k_2 - 2$, and (ii) $C(k_2) \ge 0$.
Inequality (i) can be derived as follows.
Let

$$B_f(k_2) \;=\; -\,\frac{6\nu(n_1 - \nu + 1 - p)}{N_2 + 1}\,k_2 \;+\; \frac{2\nu(n_1^2 - \nu^2 - (1-p^2))}{N_2 + 1}$$

be the fractional part of $B(k_2)$ (35) which satisfies (in consequence of (30))

$$B_f(k_2) \;\le\; 0 \;+\; \frac{2(1-p^2)\nu(\nu^2 - p^2)}{(N_2+1)\,p^2} \;\le\; 2(1-p^2)(\nu^2 - p^2) \;\le\; 2(1-p^2)\nu^2\,. \tag{37}$$

Since $A(k_2)$ (34) may attain positive or negative values, we discuss the two cases $A(k_2) \le 0$ and $A(k_2) \ge 0$ separately.

Suppose $A(k_2) \le 0$, then we have $R(i,k_2) \le R(1,k_2) = R_-(k_2)$. Substituting $k_2(z) = \nu - 2 - z\,,\ z \ge 1$ for k_2 in $R_-(k_2) - B_f(k_2)$ and rewriting the resulting expression in terms of powers of z yields

$$
\begin{aligned}
R_-(\nu-2-z) = & -4z^3 + 2\Big((1-2p)N_2 - 2n_1 + 4\nu - 9\Big)z^2 \\
& - 2\bigg(\Big((1-2p)(\nu-3)+p\nu\Big)N_2 - \nu n_1 + \nu^2 + 6n_1 - 3(3+p)\nu + 13\bigg)z \\
& - 2\bigg(\Big((1-p^2)(\nu-2)+2p(2-p)\Big)N_2 \\
& \qquad - 3(1-p)\nu(\nu-2) + 4n_1 - (9-p^2)\nu + 6\bigg) + B_f(k_2)\,.
\end{aligned}
$$

Then $R_-(k_2) \le 0$ follows from (37) and the inequalities below which are simple consequences of (29), (30) and the fact $\nu - 2 = z + k_2 \ge z + 3$.

$$
\begin{aligned}
-4z^3 - 18z^2 - 2(\nu^2+13)z - 12 &\le 0\,,\\
2(1-2p)N_2 z^2 - 2(1-2p)N_2(\nu-3)z &\le 0\,,\\
-2(2n_1-4\nu)z^2 - 2\Big(6n_1 - 3(3+p)\nu\Big)z &\le 0\,,\\
-2(p\nu N_2 - \nu n_1)z &\le 0\,,\\
-2\Big((1-p^2)(\nu-2)N_2 - (4+p)(1-p)\nu(\nu-2)\Big) &\le 0\,,\\
-2\Big(2p(2-p)N_2 + 4n_1 - (11-3p^2)\nu\Big) &\le 0\,.
\end{aligned}
$$

Suppose $A(k_2) \ge 0$, then we have $R(i,k_2) \le R(m-k_2-1,k_2) \le R(\nu-k_2-1,k_2) = R_+(k_2)$. With the same arguments as before it can be verified that $R_+(k_2) \le 0$ (for details see [21]) completing the proof of (i).
To show (ii) $C(k_2) \le 0$ we represent $C(k_2) = C(\nu-2-z)$ (36) as

$$
\begin{aligned}
C(\nu-2-z) = \quad & 2z^3 - \Big((1-2p)N_2 - 2n_1 + 4\nu - 9\Big)z^2 \\
& + \bigg(\Big(2(1-p)\nu - 3(1-2p)\Big)N_2 - 2\nu n_1 + 2\nu^2 + 6n_1 - 12\nu + 13\bigg)z \\
& + \Big(3(1-p)\nu - 2(1-2p)\Big)N_2 - 3\nu n_1 + 3\nu^2 + 4n_1 - 8\nu + 6\,.
\end{aligned}
$$

This expression may be represented as sum of six positive terms stated below.

$$
\begin{aligned}
2z^3 + 9z^2 + (2\nu^2+13)z + (3\nu^2+6) &\ge 0\,,\\
(2n_1-4\nu)z^2 + (6n_1-12\nu)z + (2pN_2+2n_1-8\nu) &\ge 0\,,\\
-(1-2p)N - 2z^2 + 2(1-2p)N_2(\nu-2)z &\ge 0\,,\\
\Big((2p\nu+(1-2p))N_2 - 2\nu n_1\Big)z &\ge 0\,,\\
(1-p)(3\nu-2)N_2 - (3\nu-2)n_1 &\ge 0\,.
\end{aligned}
$$

Thus $C(k_2) \ge 0$, which establishes part (a).

Part (b). Reflection center k_4. To verify that $N^*(k_4) \leq 0$ for all $k_4 \in J\backslash\{m\}$ (Theorem 2, part 2 and 3) we remark that $N^*(x)$ (32) is strictly increasing if $x > \nu - \frac{1}{2}$ and $N^*(x_4) = 0$. Define $k_4^{max} = \lfloor x_4 \rfloor$, since $x_4 \geq \nu$ and hence $k_4^{max} \geq \lfloor \nu \rfloor = m$. Observing that $k_4^{max} = k_4^{max}(\nu, p)$ of the binomial case (see Lemma 1 (b) in Section 2) is greater equal k_4^{max} we have $k_4^{max} \leq \frac{m+n-1}{2}$ and thus $k_5^{max} \leq n-1$, provided that $\nu \geq 3$ (Theorem 2.4(a)).

To establish the feasibility of $k_4 \in \{m+2, \ldots, k_4^{max}\}$ (Theorem 2.4(b)) we can assume $n - m \geq 5$ and have to show (i) $R(i, k_4) = A(k_4)i^2 + B(k_4) \geq 0$ and (ii) $C(k_4) \leq 0$.

Here, a sketch of the proof will do, since similar analysis as in part (a) applies. Considering (i) we mention that the fractional part $B_f(k_4)$ of $B(k_4)$ fulfills $B_f(k_4) \geq -6(1-p)\nu(\nu+p)$ and that $R(i, k_4) \geq R(k_4 - m - 1, k_4) \geq R(k_4 - \nu, k_4)$, due to $A(k_4) \geq 0$. Substituting $k_4(z) = \nu + 1 + z$ in $R(k_4 - \nu, k_4) - B_f(k_4)$ yields $R(k_4 - \nu, k_4)$ as polynomial of degree 3 in z for which $R(k_4 - \nu, k_4) \geq 0$ is verified by elementary arguments.

To show (ii), $C(k_4)$ is written as polynomial in z which can be split into three negative terms, concluding part (b) of Lemma 2. □

Acknowledgements

The authors are grateful to the Area Editor and both referees for their helpful comments and constructive suggestions.

References

[1] AHRENS, J. H. AND DIETER, U. 1982a. Generating gamma variates by a modified rejection technique. *Commun. ACM 25*, 1, 47–54.

[2] AHRENS, J. H. AND DIETER, U. 1982b. Computer generation of Poisson deviates from modified normal distributions. *ACM Trans. Math. Softw. 8*, 2, 163–179.

[3] DEVROYE, L. 1987. A simple generator for discrete log–concave distributions. *Computing 39*, 87–91.

[4] HÖRMANN, W. 1992. The transformed rejection method for generating Poisson random variables. *Insurance: Math. Econom. 11*, 1–7.

[5] HÖRMANN, W. 1993. The generation of binomial variates. *J. Statist. Comput. Simul. 46*, 1-2, 101–110.

[6] HÖRMANN, W. 1994. A universal generator for discrete log–concave distributions. *Computing 52*, 89–96.

[7] HÖRMANN, W. 1995. A rejection technique for sampling from T–concave distributions. *ACM Trans. Math. Softw. 21*, 182–193.

[8] HÖRMANN, W. AND DERFLINGER, G. 1996. Rejection–inversion to generate variates from monotone discrete distributions. *ACM Trans. Model. Comput. Simul. 6*, 169–184.

[9] HÖRMANN, W. AND DERFLINGER, G. 1997. An automatic generator for a large class of unimodal discrete distributions. In *ESM´97, Gaining Competitive Advantage Through Simulation Technologies*, A.R. Kaylan and A. Lehmann EDS., 139–144.

[10] KACHITVICHYANUKUL, V. AND SCHMEISER, B. W. 1985. Computer generation of hypergeometric random variates. *J. Statist. Comput. Simul. 22*, 127–145.

[11] KACHITVICHYANUKUL, V. AND SCHMEISER, B. W. 1988. Binomial random variate generation. *Commun. ACM 31*, 2, 216–222.

[12] KEMP, A.W. 1990. Patchwork rejection algorithms. *J. Comput. Appl. Math. 31(1)*, 127–131.

[13] MINH, D. L. 1988. Generating gamma variates. *ACM Trans. Math. Softw. 14*, 3, 261–266.

[14] SAKASEGAWA, H. 1983. Stratified rejection and squeeze method for generating beta random numbers. *Ann. Inst. Statist. Math. 35*, B, 291–302.

[15] SCHMEISER, B. W. AND BABU, A. J. G. 1980. Beta variate generation via exponential majorizing functions. *Oper. Research 28*, 4, 917–926, (Errata: *Oper. Research 31*, 802, 1983.)

[16] SCHMEISER, B. W. AND LAL, R. 1980. Squeeze methods for generating gamma variates. *J. Amer. Statist. Assoc. 75*, 371, 675–682.

[17] STADLOBER, E. 1989. Sampling from Poisson, binomial and hypergeometric distributions: ratio of uniforms as a simple and fast alternative. *Math. Statist. Sektion 303*, 93 pp., Forschungsgesellschaft Joanneum, Graz, Austria.

[18] STADLOBER, E. 1990. The ratio of uniforms approach for generating discrete random variables. *J. Comput. Appl. Math. 31(1)*, 181–189.

[19] STADLOBER, E. AND PIOK, M. 1995. WinRand 1.0/95. Accessible by anonymous ftp from `statistik.tu-graz.ac.at\winrand`.

[20] ZECHNER, H. AND STADLOBER, E. 1993. Generating beta variates via patchwork rejection. *Computing 50*, 1–18.

[21] ZECHNER, H. 1997. Efficient sampling from continuous and discrete unimodal distributions. Doctoral dissertation, 156 pp., Technical University Graz, Austria.