The residues modulo \( m \) of products of random integers

Baryshnikov, Yu.; Stadje, W.

Published: 01/01/1999

Citation for published version (APA):
Report 99-046

The Residues modulo m
of Products of Random Integers
Yuli Baryshnikov, Wolfgang Stadje
ISSN 1389-2355
The Residues modulo \( m \) of Products of Random Integers

Yuliy Baryshnikov and Wolfgang Stadje
University of Osnabrück

Abstract

For two (possibly stochastically dependent) random variables \( X \) and \( Y \) taking values in \( \{0, \ldots, m - 1\} \) we study the distribution of the random residue \( U = XY \mod m \). In the case of independent and uniformly distributed \( X \) and \( Y \) we provide an exact solution in terms of generating functions that are computed via \( p \)-adic analysis. We show also that in the uniform case it is stochastically smaller than (and very close to) the uniform distribution. For general dependent \( X \) and \( Y \) we prove an inequality for the distance \( \sup_{x \in [0,1]} |F_U(x) - x| \).

1 Introduction

Let \( X \) and \( Y \) be two (possibly dependent) random variables taking values in \( \{0,1,\ldots,m-1\} \), where \( m \geq 2 \) is some fixed integer. In this note we study the distribution of the random residue of the product

\[ U = XY \mod m. \]

We consider first the case when \( X \) and \( Y \) are independent and uniformly distributed, i.e., \( P(X = i, Y = j) = m^{-2} \) for \( i, j \in \{0,\ldots,m-1\} \). In Section 2 it is shown that the problem for general \( m \) can be reduced to that for \( m = p^n \), where \( p \) is some prime number and \( n \in \mathbb{N} \), and that in this case it is sufficient to determine the cardinalities

\[ N_p(l,n) = \#\{(x,y) \in (\mathbb{Z}/p^n\mathbb{Z}) \times (\mathbb{Z}/p^n\mathbb{Z}) \mid xy = p^{n-l}\}. \]
We prove that for every prime number \( p \) the generating function \( H_p(T, Z) = \sum_{n,l} N_p(l, n)T^nZ^l \) of the double sequence \( N_p(l, n) \) is given by

\[
H_p(T, Z) = \frac{(1 - pT)^2(1 - p^{-1}Z) - p^2(1 - p^{-1}T)T(1 - Z)}{(1 - Z)(1 - p^{-1}Z)(1 - pT)^2(1 - p^2T)}.
\] (1.1)

In the case \( p = 2 \) we derive a neat explicit formula for the distribution function of \( U \). It is given by

\[
P(U \leq k) = (k + 1)2^{-n} + 2^{-n+1}\sum_{i=0}^{n-1}(1 - \delta_i)
\] (1.2)

for \( k = 0, \ldots, 2^{n-1} \), where \( \delta_0, \ldots, \delta_{n-1} \in \{0, 1\} \) are the binary digits of \( k \), defined by \( k = \delta_0 + 2\delta_1 + 4\delta_2 + \cdots + 2^{n-1}\delta_{n-1} \).

It follows from (1.2) that the random 'fractional residue' \( 2^{-n}U \) is stochastically smaller than a uniform random variable on \([0, 1)\), i.e. \( P(U/2^n < u) \geq u \) for all \( u \in [0, 1] \) and that the maximal deviation is given by

\[
\sup_{0 < u \leq 1} (P(2^{-n}U < u) - u) = (n + 2)2^{-(n+1)},
\] (1.3)

so that the distribution of \( 2^{-n}U \) tends to the uniform distribution on \([0, 1]\) at an exponential rate (given by (1.3)), as \( n \to \infty \). In fact, these stochastic dominance and convergence remain valid for arbitrary \( m \).

The rest of the paper is devoted to an extension of this asymptotic equidistribution result to general \( m \) and dependent, non-uniform random variables \( X \) and \( Y \).

We will show that

\[
\sup_{0 \leq u \leq 1} |P(U/m < u) - u| \leq C\left(\frac{\log m}{m}\right)^{1/2}
\] (1.4)

if the distribution of \( Y \) and the conditional distribution of \( X \) given \( Y \) do not deviate too much from uniformity and if the latter distribution satisfies a certain Lipschitz condition. Specifically, we assume that

\[
P(Y = k) \leq C_0/m,
\]

\[
p(j|k) = P(X = j \mid Y = k) \leq C_1/m
\]

\[
\left|\frac{p(j_1|k)}{p(j_2|k)} - 1\right| \leq C_2|j_1 - j_2|/m
\]

2
for some constants $C_0, C_1, C_2$. Then (1.4) holds for a certain constant $C$ which depends only on $C_0, C_1$ and $C_2$. From (1.4) we can conclude that $U/m$ is for a large class of joint distributions of $X$ and $Y$ 'almost' uniformly distributed on $[0,1]$ in the sense of weak convergence.

Deterministic sequences of integers whose residues are uniformly distributed are treated in Narkiewicz [10] and Kuipers and Niederreiter [8]. They play an important role in random number generation (Ripley [12]). In the realm of stochastic sequences already Dvoretzky and Wolfowitz [5] studied weak convergence of residues for sums of independent, $\mathbb{Z}_+$-valued random variables; more recent papers on related questions are Brown [3], Barbour and Grubel [1], and Grubel [6]. The distribution of the fractional part of continuous random variables, in particular its closeness or convergence to the uniform distribution on $[0,1)$, has been studied by many authors (e.g. Schatte [13], Stadje [14, 15], Qi and Wilms [11]).

2 The uniform case

We start by deriving the exact probability distribution of $U$ in the case $m = 2^n$, $n \in \mathbb{N}$. For $x \in \mathbb{R}_+$ let $\text{frac}(x)$ be the fractional part of $x$.

**Proposition 1** We have

$$P(U \leq k) = (k + 1)2^{-n} + 2^{-(n+1)} \sum_{i=0}^{n-1} (1 - \delta_i),$$

for every $k \in \{0, 1, \ldots, 2^n - 1\}$, where $\delta_0, \ldots, \delta_{n-1} \in \{0, \ldots, n-1\}$ are the binary digits of $k$, i.e. $k = \delta_0 + 2\delta_1 + 4\delta_2 + \cdots + 2^{n-1}\delta_{n-1}$.

**Proof.** Obviously,

$$P(U = k) = \sum_{i=0}^{2^n-1} 2^{-2n} \text{card}\{j \in I_n \mid \text{frac}(ij2^{-n}) = k2^{-n}\}. \quad (2.2)$$

Let

$$A_m = \begin{cases} \{i \in I_n \mid i2^{-m} \text{ is odd}\}, & \text{if } m < n \\ \{0\}, & \text{if } m = n. \end{cases}$$

It is easily seen that

$$\text{card } A_m = \begin{cases} 2^{n-m-1}, & \text{if } m \in \{0, \ldots, n-1\} \\ 1, & \text{if } m = n. \end{cases}$$
Consider $i \in A_m$ and $k \in A_l$ for some $m, l \in \{0, \ldots, n-1\}$, say $i = (2p+1)2^m$ and $k = (2q+1)2^l$. Then for any $j \in I_n$,

$$\frac{(ij)2^{-n}}{k2^{-n}} = \frac{k2^{-n}}{k2^{-n}} \tag{2.3}$$

is equivalent to

$$(2p+1)j - (2q+1)2^{l-m} = N2^{n-m} \text{ for some integer } N. \tag{2.4}$$

For $l < m$ the lefthand side of (2.4) is not integer, so there is no solution $j$ of (2.3). Now let $l \geq m$. Since $2p+1$ and $2^n$ are relatively prime, a simple result on residues implies that the numbers $(2p+1)j - (2q+1)2^{l-m}$ run through a complete set of residues mod $2^n$ if $j$ runs through (the complete set of residues) $0, 1, \ldots, 2^n - 1$. But $N2^{n-m}$ gives different residues mod $2^n$ for $N = 0, \ldots, 2^m - 1$, while for larger values of $N$ one only gets replications of these residues. Thus, the number of solutions $j$ of (2.3) is $2^n$ if $l \geq m$. The same result also holds for $m \in A_s$, i.e. $m = 0$.

From (2.2) it now follows that if $k \in A_l$ for some $l < n$ we obtain

$$P(U = k2^{-n}) = \sum_{m=0}^{n-1} 2^{-2n} \sum_{i \in A_m} \text{card}\{j \in I_n \mid \text{int}(ij2^{-n}) = k2^{-n}\} + 2^{-n}\delta_{0k}$$

$$= \sum_{m=0}^{l} 2^{-2n} \text{card}(A_m)2^n$$

$$= \sum_{m=0}^{l} 2^{-n}2^{n-m-1}$$

$$= (l + 1)2^{-(n+1)}, \tag{2.5}$$

while if $k \in A_n$,

$$P(U = 0) = \sum_{m=0}^{n-1} 2^{-2n} \text{card}(A_m)2^n + 2^{-n}$$

$$= (n + 2)2^{-(n+1)}. \tag{2.6}$$

In particular, $k \mapsto P(U = k)$ is constant on $A_l$ for every $l$. Therefore, the probability $P(U \in (2^m\alpha, 2^m\alpha + 2^{m-1}])$ is the same for every $\alpha \in \{0, \ldots, 2^{n-m}-1\}$. \[4\]
It follows that
\[
P(U \leq k) = P(U = 0) + P(0 < U < n^{-1}2^n) + \sum_{l=1}^{n-1} P\left(\sum_{i=0}^{l-1} \delta_i2^i < U \leq \sum_{i=l}^{n-1} \delta_i2^i\right)
\] (2.7)
\[
= P(U = 0) + \sum_{l=0}^{n-1} P(0 < U \leq \delta_l2^l).
\]

To compute the righthand side of (2.7), note that the number of integers \(i \in A_m\) satisfying \(0 < i \leq 2^l\) is equal to \(2^{l-m-1}\) for \(m = 0, \ldots, l-1\) and equal to 1 for \(m = l\). Hence, by (2.5),
\[
P(0 < U \leq 2^l) = \sum_{m=0}^{l} P(U \in A_m \cap \{0, \ldots, 2^l\})
\]
\[
= \sum_{m=0}^{l-1} (l+1)2^{-(n+1)}2^{l-m-1} + (l+1)2^{-(n+1)}
\]
\[
= 2^{-(n+1)}(2^{l+1} - 1).
\] (2.8)

Inserting (2.8) and (2.6) in (2.7) now yields (2.1).

**Proposition 2**
1) For arbitrary \(m\) \(U\) is stochastically smaller than a uniform random variable on \([0, 1]\);
2) For arbitrary \(m\)
\[
\sup_{0 < u \leq 1} (P(U < u) - u) = O(m^{-1+\epsilon}),
\] (2.9)
for any \(\epsilon > 0\);
and
3) For \(m = 2^n\),
\[
\sup_{0 < u \leq 1} (P(U < u) - u) = (n + 2)2^{-(n+1)}.
\] (2.10)

**Proof.** We start with 1). It is clear that
\[
\#\{0 \leq j < m : ij \mod m \leq k\} = \gcd(i, m) \left(\left\lfloor \frac{k}{\gcd(i, m)} \right\rfloor + 1\right).
\] (2.11)
This implies
\[
P(U \leq k) = \frac{1}{m^2} \sum_{i=0}^{m-1} \gcd(i, m) \left\lfloor \frac{k}{\gcd(i, m)} \right\rfloor + 1 > k/m \tag{2.12}
\]
for all \(0 \leq k < m\), and hence proves 1).

Further, estimating (2.12) in an obvious way from above, we obtain
\[
P(U \leq k) \leq \frac{1}{m^2} \sum_{i=0}^{m-1} \gcd(i, m) \left( \frac{k}{\gcd(i, m)} + 1 \right)
\leq \frac{k}{m} + \frac{1}{m^2} \sum_{i=0}^{m-1} \gcd(i, m)
= \frac{k}{m} + \frac{1}{m^2} \sum_{i|m} \# \{0 \leq i < m : \gcd(i, m) = t\}
\leq \frac{k}{m} + \frac{1}{m} \sum_{i|m} \frac{d(m)}{m}
= \frac{k}{m} + d(m)/m,
\tag{2.13}
\]
where \(d(m)\) denotes the number of divisors of \(m\). It is known that \(d(m) = O(m^\epsilon)\) for all \(\epsilon > 0\), which implies 2).

To prove 3) define for \(0 < u \leq 1\) the integer \(k(u)\) by \(k(u)2^{-n} < u \leq (k(u) + 1)2^{-n}\) and let \(\delta_0, \ldots, \delta_{n-1}\) be its binary digits. By (2.1) we can write
\[
P(U < u) - u = (k(u)2^{-n} + 2^{-n} - u) + 2^{-(n+1)} \sum_{i=0}^{n-1} (1 - \delta_i), \tag{2.14}
\]
which is nonnegative by the definition of \(k(u)\). Further it is clear from (2.14) that \(\sup_{0 < u \leq 1} (P(U < u) - u)\) is approached as \(u \downarrow 0\), yielding (2.10).

Now we derive the exact formulae for \(P(U = k)\) in the case of general \(m \in \mathbb{N}\).

Let \(X\) and \(Y\) be independent and uniform on the set \(\{0, \ldots, m - 1\}\), which we identify with \(\mathbb{Z}/m\mathbb{Z}\). Then \(P(U = a)\) is equal to \(m^{-2}\) times the number of solutions \((x, y) \in (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})\) of the equation
\[
x y \equiv a \mod m.
\]

Let \(m = \prod p_i^{n_i}\) be the prime factorization of \(m\) (\(p_i\) primes, \(n_i \in \mathbb{N}\)). For \(a \in \mathbb{Z}/m\mathbb{Z}\) we define \(a(i) \in \mathbb{Z}/p_i^{n_i}\mathbb{Z}\) as the (unique) solution of
\[
a(i) \equiv a \mod p_i^{n_i}.
\]

Then as \(\mathbb{Z}/m\mathbb{Z} = \prod (\mathbb{Z}/p_i^{n_i}\mathbb{Z})\) (the Chinese remainder theorem), we have the following decomposition.
Lemma 1  The number of pairs \((x, y) \in (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})\) satisfying
\[
xy \equiv a \mod m
\] (2.15)
is equal to the product of the numbers of solutions \((x, y) \in (\mathbb{Z}/p_i^n\mathbb{Z}) \times (\mathbb{Z}/p_i^n\mathbb{Z})\) of
\[
xy \equiv a(i) \mod p_i^n.
\] (2.16)

By the Lemma, we only have to determine the number of solutions of (2.15) for \(m\) of the form \(m = p^n\).

Fix a prime number \(p\) and a natural number \(n\). Observe first that the number of solutions \((x, y) \in (\mathbb{Z}/p^n\mathbb{Z}) \times (\mathbb{Z}/p^n\mathbb{Z})\) of \(xy \equiv a \mod p^n\) depends on \(a\) only through the \(p\)-adic norm of \(a\), that is, through the exponent of the maximal power of \(p\) that divides \(a\). Indeed, if there exists an invertible \(b\) in \(\mathbb{Z}/p^n\mathbb{Z}\) satisfying
\[
\# \{(x, y) \in (\mathbb{Z}/p^n\mathbb{Z}) \times (\mathbb{Z}/p^n\mathbb{Z}) \mid xy \equiv a \mod p^n\}
= \# \{(x, y) \mid xyb \equiv p^{n-l} \mod p^n\}
= \# \{(x, z) \in (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z}) \mid xz \equiv p^{n-l} \mod p^n\}
= N_p(l, n).
\]

To compute \(N_p(l, n)\), we use the following well-known formula from the theory of \(p\)-adic integration (Christol [4, Sect. 7.2.2, p. 466]). Let \(f(x_1, \ldots, x_r)\) be a polynomial with coefficients in \(\mathbb{Z}_p\), the ring of \(p\)-adic integers, and let \(| \cdot |_p\) denote the \(p\)-adic norm. Then for any real \(s > 0\),
\[
\int \left(\mathbb{Z}_p\right)^r |f(x_1, \ldots, x_r)|_p^s \mu(dx_1) \cdots \mu(dx_r) = p^s - (p^s - 1)Q(p^{r-s}), \quad (2.17)
\]
where \(\mu\) is the Haar measure on \(\mathbb{Z}_p\) and \(Q(T)\) is a Poincaré series:
\[
Q(T) = \sum_{k=0}^{\infty} T^k \# \{(x_1, \ldots, x_r) \in (\mathbb{Z}/p^k\mathbb{Z})^r \mid f(x_1, \ldots, x_r) \equiv 0 \mod p^k\}.
\]

Theorem 1  The generating functions
\[
G_{p,l}(T) = \sum_{n=0}^{\infty} N_p(l, n) T^n, \quad H_{p}(T, Z) = \sum_{n=0}^{\infty} \sum_{l=0}^{n} N_p(l, n) T^n Z^l
\]
are given by

\[
G_{p,i}(T) = \frac{p^l(1-pT)^2 - p^2(1-p^{-1})^2T}{p^l(1-pT)^2(1-p^2T)}
\] (2.18)

\[
H_p(T, Z) = \frac{(1-pT)^2(1-p^{-1}Z) - p^2(1-p^{-1}T)(1-Z)T}{(1-Z)(1-p^{-1}Z)(1-p^2T)(1-p^2T)}
\] (2.19)

**Proof.** We use formula (2.17) for \( r = 2 \) and \( f(x, y) = f_l(x, y) = p^lxy \). For the lefthand side of (2.17) we obtain

\[
\int \left| f_l(x, y) \right|_p^s \mu(dx) \mu(dy) = \int \left| x \right|_p^s \left| y \right|_p^s \mu(dx) \mu(dy)
\]

\[
= p^{-1} \left( \int \left| x \right|_p^s \mu(dx) \right)^2.
\]

By (2.17),

\[
\int \left| x \right|_p^s \mu(dx) = p^s - (p^s - 1) \frac{1}{1-p^{-1-s}} = \frac{1-p^{-1}}{1-p^{-1-s}}.
\]

(Note that here \( Q(T) = 1/(1-T) \), since \( \#\{x \in \mathbb{Z}p^n/Z \mid x \equiv 0 \mod p^n\} = 1 \) for all \( n \)). Furthermore,

\[
xy \equiv p^{-l} \mod p^n \iff \ p^lxy \equiv 0 \mod p^n.
\]

Thus, the coefficients on the righthand side of (2.17) are just the \( N_p(l, n) \). It follows that

\[
p^s - (p^s - 1) \sum_n N_p(l, n)(p^{-2-s})^n = p^{-1} \left( \frac{1-p^{-1}}{1-p^{-1-s}} \right)^2.
\]

Setting \( T = p^{-2-s} \), so that \( p^{-s} = p^2T \) we get

\[
\frac{1}{p^2T} - \left( \frac{1}{p^2T} - 1 \right) G_{p,i}(T) = p^{-1} \left( \frac{1-p^{-1}}{1-pT} \right)^2
\] (2.20)

and (2.18) follows from (2.20) by a short calculation. Similarly, multiplying (2.20) by \( Z^l \) and summing over \( l \) yields (2.19).
For example, if \( p = 2 \) the numbers \( N_p(0, n) \) of solutions \( (x, y) \) of \( (x, y) \equiv 0 \mod 2^n \) is \((n + 2)2^{n-1}\), as

\[
G_{2,0}(T) = \sum_{n=0}^{\infty} N_p(0, n)T^n = \frac{(1 - 2T)^2 - T}{(1 - 2T)^2(1 - 4T)} = \frac{1 - T}{(1 - 2T)^2} = \sum_{n=0}^{\infty} (n + 2)2^{n-1}T^n.
\]

3 The inequality for dependent random variables

We will now prove (1.4). For this we need some basic theory of continued fractions (see e.g. Hardy and Wright [7], Billingsley [2]) and a probability estimate due to Lévy [9]).

Any \( x \in [0, 1] \) has a continued fraction expansion \( x = [a_1(x), a_2(x), \ldots] \) providing a sequence of fractions usually denoted by

\[
p_n(x)/q_n(x) = [a_1(x), \ldots, a_n(x)].
\]

For two positive numbers \( \rho_0 < \rho_1 \) let

\[
B(\rho_0, \rho_1) = \{x \in [0, 1] \mid \rho_0 < q_k(x) < \rho_1 \text{ for some } k \in \mathbb{N}\}.
\]

**Lemma 2** \( \lambda(B(\rho_0, \rho_1)) \geq 1 - \frac{2\rho_0}{\rho_1 - \rho_0}(1 + 2 \log_2 \rho_0) - \rho_1^{-1} \).

**Proof.** Let \( Q \) be the set of all finite sequences \( \vec{q} = (q_1, \ldots, q_k), \ k \in \mathbb{N}, \) of denominators of possible continued fraction expansions satisfying \( q_k \leq \rho_0 \). We set \( x(\vec{q}) = p_k/q_k \), where \( p_k \) is the \( k \)th numerator corresponding to \( q_1, \ldots, q_k \), and

\[
I(\vec{q}) = \{x \in [0, 1] \mid (q_1(x), \ldots, q_k(x)) = \vec{q}\}
\]

\[
J(\vec{q}) = I(\vec{q}) \cap \{x \in [0, 1] \mid q_{k+1}(x) \geq \rho_1 \text{ or } x = x(\vec{q})\}
\]

\[
J(0) = \{x \in [0, 1] \mid q_1(x) \geq \rho_1\}.
\]

The sets \( J(\vec{q}), \vec{q} \in Q, \) and \( J(0) \) are pairwise disjoint intervals and

\[
B(\rho_0, \rho_1) = [0, 1]\setminus \left( J(0) \cup \bigcup_{\vec{q} \in Q} J(\vec{q}) \right).
\]
Thus,
\[
\lambda([0, 1] \setminus B(\rho_0, \rho_1)) = \lambda(J(0)) + \sum_{q \in Q} \lambda(J(q))
\]
\[
= \lambda(J(0)) + \sum_{k=1}^{k_0} \sum_{q \in Q, |q| = k} \lambda(J(q)),
\]
(3.1)

where $|q|$ denotes the length of the sequence $q$ and $k_0$ is the maximum length of sequences in $Q$. Since
\[
\rho_0 > q_k \geq 2^{(k-1)/2}
\]
for every $(q_1, \ldots, q_k) \in Q$,

it follows that
\[
k_0 < 1 + 2 \log_2 \rho_0.
\]

(3.2)

Now let $U$ be a random variable that is uniformly distributed on $[0, 1]$. Then if $q \in Q, |q| = k$, it follows that
\[
\lambda(J(q)) = P(q_{k+1}(U) \geq \rho_1, U \in I(q))
\]
\[
= P(U \in I(q))P(q_{k+1}(U) \geq \rho_1 | U \in I(q))
\]
\[
\leq P(U \in I(q))P(a_{k+1}(U) > \frac{\rho_1 - \rho_0}{\rho_0} | U \in I(q))
\]
\[
\leq P(U \in I(q))2 \left( \frac{\rho_1 - \rho_0}{\rho_0} \right)^{-1}.
\]

(3.3)

For the first inequality in (3.3) we have used the recursion $q_{k+1} = q_k a_{k+1} + q_{k-1}$ which for $q \in Q, |q| = k$, implies that $a_{k+1} > (\rho_1 - \rho_0)/\rho_0$. The second inequality follows from a result of Lévy [9, p. 296].

To estimate $\lambda(J(0))$, note that $q_1(x) \geq \rho_0$ implies that $x \leq p_1(x)/q_1(x) = 1/\rho_1$. Thus, by (3.1), (3.2) and (3.3).
\[
\lambda([0, 1] \setminus B(\rho_0, \rho_1)) \leq \rho_1^{-1} + k_0 \frac{2\rho_0}{\rho_1 - \rho_0} \sum_{q \in Q} P(U \in I(q))
\]
\[
\leq \rho_1^{-1} + (1 + 2 \log_2 \rho_0) \frac{2\rho_0}{\rho_1 - \rho_0}.
\]

The Lemma is proved.
Lemma 3 Let $X$ be uniformly distributed on $\{0, 1, \ldots, m - 1\}$. Then
\[
P(X/m \notin B(\rho_0, \rho_1)) \leq 2\rho_0 (1 + 2 \log \rho_0) \left( \frac{1}{\rho_1 - \rho_0} + \frac{\rho_0}{m} \right) + \rho_1^{-1} + m^{-1}.
\] (3.4)

Proof. For every half-open or open interval $I$ in $[0, 1]$ we have
\[
|P(X/m \in I) - \lambda(I)| \leq m^{-1}.
\] (3.5)
As $J(0)$ and $J(\tilde{q})$ are half-open intervals, (3.1) and (3.4) yield
\[
P(X/m \notin B(\rho_0, \rho_1)) \leq \lambda(J(0)) + \sum_{\tilde{q} \in \tilde{Q}} \lambda(J(\tilde{q}))
+ m^{-1}(1 + \text{card } Q).
\] (3.6)

It remains to find an upper bound for $\text{card } Q$. Let $\tilde{Q}$ be the set of sequences in $Q$ having maximal length, i.e., the set of those $(q_1(x), \ldots, q_k(x)) \in Q$ for which $q_{k+1}(x) \geq \rho_0$. Since
\[
\lambda(I(q_1, \ldots, q_k)) = \frac{1}{q_k(q_k + q_{k-1})} > \frac{1}{2q_k^2} \geq \frac{1}{2\rho_0^2}
\]
for $(q_1, \ldots, q_k) \in \tilde{Q}$, we clearly have $\text{card } \tilde{Q} < 2\rho_0^2$. Inequality (3.4) now follows from (3.6), Lemma 2 and
\[
\text{card } Q \leq k_0 \text{card } \tilde{Q} < (1 + \log \rho_0)(2\rho_0^2).
\]

Lemma 4 Let
\[
p(j, k) = P(X = j, \ Y = k), \ j, k \in \{0, \ldots, m - 1\}
\]
be the joint distribution of $X$ and $Y$. Assume that there are constants $C_1$ and $C_2$ such that
\[
p(j|k) = P(X = j|Y = k) \leq C_1/m
\] (3.7)
\[
\left| \frac{p(j_1|k)}{p(j_2|k)} - 1 \right| \leq C_2|j_1 - j_2|/m
\] (3.8)
for all $j, k, j_1, j_2 \in \{0, \ldots, m - 1\}$. Then
\[
|P(U/m < u|Y = k) - u| \leq \frac{3C_2}{m} + \inf_{n \geq 1} f \left( q_n \left( \frac{k}{m} \right) \right)
\]
for all $k \in \{0, \ldots, m-1\}$, where

$$f(q) = \frac{3}{q} + \frac{(C_1 + C_2)q}{m}, \quad q \in \mathbb{N}.$$  

Proof. Let $p/q$ be an arbitrary fraction from the continued fraction expansion of $k/m$. Let

$$J_i = \{(i-1)q, (i-1)q + 1, \ldots, iq - 1\}$$

$$J_i(u) = \{j \in J_i \mid \text{frac}(jk/m) < u\},$$

where $\text{frac}(x)$ denotes the fractional part of $x \geq 0$. Then

$$P(U/m < u \mid Y = k) = \sum_{i=1}^{[m/q]} \sum_{j \in J_i(u)} P(X = j \mid Y = k) + \sum_{k \in J_i([m/q]+1)} P(X = j \mid Y = k) \quad (3.9)$$

$$= I + II.$$

Clearly, (3.7) yields

$$II \leq C_1q/m. \quad (3.10)$$

Regarding the sum $I$, we can write

$$I = \sum_{i=1}^{[m/q]} \sum_{j \in J_i(u)} p(j \mid k) \quad (3.11)$$

$$\leq \sum_{i=1}^{[m/q]} \frac{A_i \text{card } J_i(u)}{a_i \text{card } J_i} \sum_{j \in J_i} p(j \mid k),$$

where $A_i = \max_{j \in J_i} p(j \mid k)$ and $a_i = \min_{j \in J_i} p(j \mid k)$. From (3.8) we can conclude that

$$A_i/a_i \leq 1 + (C_2q/m). \quad (3.12)$$

Obviously, $\text{card } J_i = q$. We need an upper bound for $\text{card } J_i(u)$. Note that

$$\left| \frac{k}{m} - \frac{p}{q} \right| < q^{-2}.$$
For arbitrary \( j \in J_i(u) \) write \( j = (i - 1)q + h \), where \( h \in J_1 \); we obtain

\[
\frac{jk}{m} = \frac{\left((i - 1)q \frac{k}{m} + \frac{hk}{m}\right)}{m} = \frac{\left((i - 1)q \frac{k}{m} + \frac{hk}{m}\right)}{m}
\]

and

\[
\frac{hk}{m} = \frac{h\left(\frac{k}{m} - \frac{p}{q}\right) + \frac{hp}{q}}{m} = \frac{\left(\alpha + \frac{hp}{q}\right)}{m}
\]

where \(|\alpha| < q^{-1}\). Recall that \( p \) and \( q \) are relatively prime. Thus, as \( h \) runs through \( J_1 \), \( \frac{hk}{m} \) runs through the set of all values \( \frac{l}{q} + \alpha \), \( l \in J_1 \). Let \( \beta_i = (i - 1)qk/m \).

Let \( \tilde{j}_i(u) \) be the number of values \( \frac{\beta_i + (l/q)}{m} \) in \([0, u)\) for which \( l \in J_1 \). Clearly, we have \( \tilde{j}_i(u) \in \{[qu], [qu] + 1\} \). Since \(|\alpha| < q^{-1}\), it now follows easily that

\[
|\tilde{j}_i(u) - \text{card } J_i(u)| \leq 2,
\]

so that

\[
|qu - \text{card } J_i(u)| \leq 3. \quad (3.13)
\]

By (3.12) and (3.13),

\[
\frac{A_i \text{ card } J_i(u)}{a_i \text{ card } J_i} \leq \left(1 + \frac{C_1q}{m}\right) \frac{qu + 3}{q} \leq u + \frac{C_1q}{m} + \frac{3}{q} + \frac{3C_2}{m}. \quad (3.14)
\]

Inserting (3.14) and (3.10) in (3.9) we find that

\[
P(U/m < u) \leq u + \frac{C_2q}{m} + \frac{3}{q} + \frac{3C_2}{m} + \frac{C_1q}{m}
= u + \frac{3C_2}{m} + f(q).
\]

Minimizing with respect to all possible denominators \( q = q_n(k/m) \) we arrive at

\[
P(U/m < u) - u \leq \frac{3C_2}{m} + \inf_{n \geq 1} f \left(\frac{q_n(k/m)}{m}\right).
\]

The analogous lower bound \( P(U/m < u) \geq u - (3C_2/m) - f(q) \) is derived along the same lines.
Theorem 2 Assume that the joint distribution of $X$ and $Y$ satisfies conditions (3.7) and (3.8) and that

$$P(Y = k) \leq C_0/m, \ k = 0, \ldots, m - 1.$$  \hspace{1cm} (3.15)

for some constant $C_0$. Then there is a constant $C$ depending only on $C_0, C_1, C_2$ such that

$$\sup_{0 \leq u \leq 1} |P(U/m < u) - u| \leq C \left( \frac{\log m}{m} \right)^{1/2}. \hspace{1cm} (3.16)$$

Proof. By the formula of total probability and Lemma 4, we obtain

$$P(U/m < u) = \sum_{k=0}^{m-1} P(Y = k)P(U/m < u|Y = k)$$

$$\leq u + 3C_2m^{-1} + \sum_{k=0}^{m-1} P(Y = k) \min \left[ 1, \min_{n \geq 1} f \left( q_n \left( \frac{k}{m} \right) \right) \right]$$

$$= u + 3C_2m^{-1} + \min \left[ 1, \min_{n \geq 1} f \left( q_n \left( \frac{Y}{m} \right) \right) \right]. \hspace{1cm} (3.17)$$

Note that the right side of (3.17) is equal to \( \int_0^1 (1 - G(x))dx \), where

$$G(x) = P \left( \min_{n \geq 1} f \left( q_n \left( \frac{Y}{m} \right) \right) < x \right).$$

Let $C_3 = C_1 + C_2$. The function $f(t) = 3t^{-1} + C_3m^{-1}t$, $t > 0$, is strictly convex, has the unique minimum $t_0 = (3m/C_3)^{1/2}$ and $x_0 = f(t_0) = 2t_0^{-1}$. Thus the equation $f(t) = x$ has no solution for $x < x_0$ and exactly two solutions $t_1(x) < t_2(x)$ for $x > x_0$. If $x > x_0$, a short calculation yields

$$f(6/x) = f(mx/2C_3) = \frac{x}{2} + \frac{6C_3}{mx} < x,$$

and consequently $t_1(x) < 6/x < mx/2C_3 < t_2(x)$. These observations show that

$$G(x) = P \left( t_1(x) < q_n(Y/m) < t_2(x) \mbox{ for some } n \in \mathbb{N} \right)$$

$$\geq P(6/x < q_n(Y/m) < mx/2C_3 \mbox{ for some } n \in \mathbb{N})$$

$$= P(Y/m \in B(6/x, \ mx/2C_3)). \hspace{1cm} (3.18)$$

From (3.15) and Lemma 3 it now follows that

$$1 - G(x) \leq H(x) + m^{-1}, \ x \in (0,1]$$

14
where the function $H$ is defined by
\[
H(x) = \frac{2C_3}{mx} + 2C_0 \left( \frac{(6/x)^2 m^{-1} + \frac{12C_3}{mx^2 - 12C_3}}{1 + 2\log_2(6/x)} \right), \quad x > x_0.
\]
Thus, for any $y \in (x_0, 1]$ we have the following estimate:
\[
E(\min[1, f(qn(Y/m))]) = \int_0^1 (1 - G(x)) \, dx \leq y + \int_y^1 H(x) \, dx. \quad (3.19)
\]
On $(x_0, \infty)$ the function $H(x)$ is positive and strictly decreasing from infinity at zero. Further,
\[
H(x) \geq 2 \left( \frac{36}{mx^2} + \frac{12C_3}{mx^2} \right) \left( 1 + 2\log_2(6/x) \right) \geq 12 \cdot \frac{48}{mx^2}, \quad x \in (x_0, 1] \quad (3.20)
\]
as $C_0 \geq 1$ and $C_3 \geq 1$. Let $x_1$ be the solution of $H(x) = 1$ in $(x_0, \infty)$. For sufficiently large $m$ we have $x_1 < 1$ and then, by (3.20),
\[
x_1 \geq \max[12(C_3/m)^{1/2}, \ (576/m)^{1/2}].
\]
Hence if $x_1 \leq x \leq 1$, $H(x)$ can be bounded as follows:
\[
H(x) \leq \frac{2C_3}{mx} + 2C_0 \left( \frac{36}{mx^2} + \frac{12C_3}{mx^2(1 - (12C_3/mx_1^2))} \right) \left( 1 + \log_2(36/x_1^2) \right)
\leq \frac{2C_3}{mx} + \frac{2C_0}{mx} \left( 36 + \frac{144}{11}C_3 \right) \left( 1 + \log_2(36m/576) \right)
\leq \frac{2C_3}{mx} + \frac{2C_0}{mx} (36 + 14C_3)(\log_2 m - 3).
\]
For any $y \in [x_1, 1]$ we now find that
\[
y + \int_y^1 H(x) \, dx \leq y + \frac{2C_3}{my} + \frac{2C_0(36 + 14C_3)(\log_2 m - 3)}{my}. \quad (3.21)
\]
Over $y \in (0, \infty)$ the right-hand side of (3.21) is minimized for
\[
y_0 = \left[ 2C_3 + 2C_0(36 + 14C_3)(\log_2 m - 3) \right]^{1/2} m^{-1/2},
\]
the corresponding minimum being equal to $2y_0$. A short calculation shows that $H(y_0) \to (9 + 3C_3)/(9 + 4C_3) < 1$, as $m \to \infty$. Thus, $y_0 > x_1$ for sufficiently large $m$. Hence we may insert the value $y_0$ in (3.21) for all but finitely many $m$. To summarize, it is now proved that
\[
P(U/m < u) \leq u + C \sqrt{\frac{\log m}{m}}.
\]
for some constant $C$ depending only on $C_0, C_1,$ and $C_2$. Similarly it can be shown that $P(U/m < u) \geq u - C((\log m)/m)^{1/2}$.

References