Statistics & Probability Letters 14 (1992) 235–241
North-Holland 24 June 1992

On the binary expansion of a random integer

A.D. Barbour
University of Zürich, Switzerland

L.H.Y. Chen
National University of Singapore, Singapore

Received August 1991
Revised October 1991

Abstract: It is shown that the distribution of the number of ones in the binary expansion of an integer chosen uniformly at random from the set $0, 1, \ldots, n-1$ can be approximated in total variation by a mixture of two neighbouring binomial distributions, with error of order $(\log n)^{-1}$. The proof uses Stein's method.

Keywords: Random integer, binary expansion, Stein's method, binomial mixtures.

1. Introduction

Choose a random number from the integers $0, 1, \ldots, n-1$: how many 1's are there in its binary expansion? Clearly, if $n = 2^k$ for some k, the distribution of the number of 1's is $\mathrm{Bi}(k, \frac{1}{2})$, but for $2^{k-1} < n < 2^k$ there is some dependence between the digits in the expansion, and the distribution is no longer so obvious. In fact, if $e(n)$ is the expected number of 1's, the function e has strict local maxima whenever $n = 2^k$ for $k \geqslant 3$, and, as k gets larger, at ever more points intermediate between 2^{k-1} and 2^k. A discussion of the irregular properties of $e(n)$ can be found in Delange (1975).

Nonetheless, $e(n)$ does not change appreciably in magnitude in the interval $2^{k-1} \leqslant n \leqslant 2^k$, as follows from Delange's results or from Lemma 2 below. It thus also seems plausible that the distribution of the number of 1's should not change much either, even if in an irregular manner: indeed, the difference between $\mathrm{Bi}(k-1, \frac{1}{2})$ and $\mathrm{Bi}(k, \frac{1}{2})$ in total variation is only of order $k^{-1/2}$. Diaconis and Stein were able to show that a difference of this magnitude is in fact typical, in the sense that, if $W = W(n)$ denotes the random number of 1's and $2^{k-1} \leqslant n \leqslant 2^k$, the distribution of $W(n)$ is only order $k^{-1/2}$ away from $\mathrm{Bi}(k, \frac{1}{2})$. In Diaconis (1977) this is demonstrated by way of normal approximation, whereas in Stein

Correspondence to: A.D. Barbour, Institut für Angewandte Mathematik, Universität Zürich, Rämistrasse 74, CH-8001 Zürich, Switzerland.

This work was supported in part by Schweiz. Nationalfonds Grant No. 21-25579.88 and by NSF Grant DMS 90-05833

0167-7152/92/$05.00 © 1992 – Elsevier Science Publishers B.V. All rights reserved

(1986) the accuracy of the approximation for point probabilities is investigated. A formulation more in keeping with the aims of this paper is given in Loh (1991, Theorem 5), who shows that

$$d_{\mathrm{TV}}(\mathscr{L}(W(n)), \mathrm{Bi}(k, \tfrac{1}{2})) = \sup_{A \subset \mathbb{Z}^+} \left|\mathbb{P}[W(n) \in A] - \mathrm{Bi}(k, \tfrac{1}{2})\{A\}\right| \leqslant ck^{-1/2} \tag{1.1}$$

for a fixed constant c: Chen and Soon (1991) improve the value of the constant.

In this note, it is shown that the error obtained in the above papers is as big as order $k^{-1/2}$ only because of the discrepancy between $e(n)$ and the mean $\frac{1}{2}k$ of the approximating binomial distribution. If, instead, a mixture of the $\mathrm{Bi}(k-1, \frac{1}{2})$ and either $\mathrm{Bi}(k, \frac{1}{2})$ or $\mathrm{Bi}(k-2, \frac{1}{2})$ distributions is used as an approximation, the mixture being chosen so as to have mean $e(n)$, the error is reduced to order k^{-1}. Stein's method is used once again to prove the result, now by way of approximation with a $\mathrm{Bi}(2e(n), \frac{1}{2})$ distribution, where, for $m > 0$ not necessarily integral,

$$\mathbb{P}[\mathrm{Bi}(m, \tfrac{1}{2}) = j] = \binom{m}{j} \Big/ \sum_{k=0}^{\lceil m \rceil} \binom{m}{k}, \quad 0 \leqslant j \leqslant \lceil m \rceil, \tag{1.2}$$

$\lceil m \rceil$ denoting the smallest integer greater than or equal to m.

2. Details

The first in the argument is to note that the distribution $\mathrm{Bi}(m, \frac{1}{2})$ given by (1.2) is the equilibrium distribution of the birth and death process with transition rates given by

$$q_{j,j+1} = (m-j)_+, \qquad q_{j,j-1} = j. \tag{2.1}$$

It then follows from Lemma 9.2.1 of Barbour, Holst and Janson (1992) that, for any bounded function f, the solution g of the equations

$$(m-j)_+ g(j+1) - jg(j) = f(j) - \mathrm{Bi}(m, \tfrac{1}{2})\{f\}, \quad 0 \leqslant j \leqslant \lceil m \rceil, \tag{2.2}$$

with $g(0) = 0$ and $g(j) = 0$, $j > \lceil m \rceil$, satisfies

$$\Delta g = \sup_{j \geqslant 0} |g(j+1) - g(j)| \leqslant 4m^{-1} \| f \|, \tag{2.3}$$

where $\| f \| = \sup_j |f(j)|$, and hence that, if $f(j) = I[j \in A] - \frac{1}{2}$ for any $A \subset \{0, 1, \dots, \lceil m \rceil\}$,

$$\Delta g \leqslant 2m^{-1}. \tag{2.4}$$

It is also easy to modify the proof of Lemma 1.1.1 of Barbour, Holst and Janson (1991) to show that

$$\| g \| \leqslant c(1 \wedge m^{-1/2}) \| f \|, \tag{2.5}$$

for a universal constant c. As a consequence of these observations, if W is any non-negative random variable, A is any subset of $\{0, 1, \dots, \lceil m \rceil\}$ and $g = g_{m,A}$ is the solution to (2.2) with $f(j) = I[j \in A]$, then

$$\begin{aligned} &\mathbb{P}[W \in A] - \mathrm{Bi}(m, \tfrac{1}{2})\{A\}\mathbb{P}[0 \leqslant W \leqslant \lceil m \rceil] \\ &\quad = \mathbb{E}\{(m-W)_+ g_{m,A}(W+1) - Wg_{m,A}(W)\}. \end{aligned} \tag{2.6}$$

By considering all A with $\mathrm{Bi}(m, \frac{1}{2})\{A\} \leqslant \frac{1}{2}$, this leads to the estimate

$$\begin{aligned} &d_{\mathrm{TV}}(\mathscr{L}(W), \mathrm{Bi}(m, \tfrac{1}{2})) \\ &\quad \leqslant \sup_A \left|\mathbb{E}\{(m-W) g_{m,A}(W+1) - Wg_{m,A}(W)\}\right| + \tfrac{1}{2}\mathbb{P}[W > \lceil m \rceil]: \end{aligned} \tag{2.7}$$

note that $g_{m,A}(w+1)=0$ if $w>m$. This inequality, together with (2.4) and (2.5), forms the basis of the bounds that are derived in the paper: we show that the right hand side of (2.7) is small when $W=W(n)$ and $m=2e(n)$, and that it is also small for the same m if W is that mixture of $\mathrm{Bi}(k-1, \frac{1}{2})$ and either $\mathrm{Bi}(k, \frac{1}{2})$ or $\mathrm{Bi}(k-2, \frac{1}{2})$ which has mean $e(n)$.

The approach used is different from those of the papers referred to earlier, inasmuch as we base the argument upon the observation that $\mathscr{L}(W(n))$ can be expressed as a mixture of shifted binomials with $p=\frac{1}{2}$. For n between 2^{k-1} and 2^k, write

$$n=\sum_{j=1}^{k} n_j 2^{k-j}=\sum_{i=1}^{K} 2^{k-l(i)}, \tag{2.8}$$

where $l(i)$ is the index of the ith 1 in the sequence $(n_j)_{j=1}^{k}$, of which $K=K(n)$ are 1's, and $l(1)=1$ since $n_1=1$. Partition the integers $0, 1, \dots, n-1$ into sets $P_1, \dots, P_K$, where $P_1=\{m\colon 0\leqslant m<2^{k-1}\}$ and

$$P_i=\left\{m\colon \sum_{r=1}^{i-1} 2^{k-l(r)}\leqslant m<\sum_{r=1}^{i} 2^{k-l(r)}\right\},\quad 2\leqslant i\leqslant K.$$

Then the probability that a randomly chosen integer from $0, 1, \dots, n-1$ belongs to P_i is given by

$$\theta_i=n^{-1}|P_i|=2^{k-l(i)}\bigg/\left\{\sum_{r=1}^{K} 2^{k-l(r)}\right\},$$

and, conditional on the integer belonging to P_i, the number of ones in its binary expansion is distributed as

$$\mathrm{Bi}\left(k-l(i), \tfrac{1}{2}\right)+i-1,$$

since there are $i-1$ ones in the first $l(i)$ digits common to all elements of P_i, and the remaining $k-l(i)$ digits can be any arrangement of 0's and 1's. Thus, using the notation $L_k(\theta;\, a, b)$ to denote the mixture according to the probabilities $(\theta_i, i\geqslant 1)$ of the binomial $\mathrm{Bi}(k-a_i, \frac{1}{2})$ distributions, shifted by b_i units, it follows that $W(n)\sim L_k(\theta;\, a, b)$ with

$$\theta_i=2^{k-l(i)}\bigg/\left\{\sum_{r=1}^{K} 2^{k-l(r)}\right\},\qquad a_i=l(i)\geqslant i,\qquad b_i=i-1. \tag{2.9}$$

Hence we are able to deduce the approximation of $\mathscr{L}(W(n))$ by $\mathrm{Bi}(2e(n), \frac{1}{2})$ from a more general result for mixtures of shifted binomials, given in Theorem 1 below.

Before doing so, we shall need the following technical lemma.

Lemma 1. *Let $f\colon \mathbb{Z}\to\mathbb{R}$ satisfy $\|f\|<\infty$, and let $X\sim\mathrm{Bi}(t, \frac{1}{2})$, for some integral t. Suppose that U is an integer valued random variable independent of X, such that $0\leqslant \mathbb{E}U=\gamma\leqslant 1$. Then, if the domain of f is extended to $\mathbb{R}$ by linear interpolation,*

$$\begin{aligned}&|\mathbb{E}\{f(X+U-\gamma)-f(X)\}|\\&\quad\leqslant 8t^{-1}\|f\|\{\mathrm{Var}\ U+\gamma(1-\gamma)\}+2\|f\|(1-\mathbb{P}[1\leqslant X+U\leqslant t]).\end{aligned}$$

Proof. Let g be the solution of (2.2) with $m=t$. Then

$$\begin{aligned}&\mathbb{E}\{f(X+U-\gamma)-f(X)\}\\&\quad=(1-\gamma)\mathbb{E}[f(X+U)-f(X)]+\gamma\mathbb{E}[f(X+U-1)-f(X)].\end{aligned}$$

Now, from (2.2), if V is integer valued and independent of X,

$$\begin{aligned}
&\mathbb{E}[f(X+V)-f(X)]\\
&\quad=\mathbb{E}\{(t-X-V)g(X+V+1)-(X+V)g(X+V)\}\\
&\qquad+\mathbb{E}\{[f(X+V)-f(X)]I[X+V\notin[0,t]]\}\\
&\quad=-\mathbb{E}\{V[g(X+V+1)+g(X+V)]\}\\
&\qquad+\mathbb{E}\{[f(X+V)-f(X)]I[X+V\notin[0,t]]\},
\end{aligned}$$

and applying this equality with $V=U$ and $U-1$ gives

$$\begin{aligned}
&\mathbb{E}[f(X+U-\gamma)-f(X)]\\
&\quad=-(1-\gamma)\mathbb{E}\{U[g(X+U)+g(X+U+1)]\}\\
&\qquad-\gamma\mathbb{E}\{(U-1)[g(X+U-1)+g(X+U)]\}+\eta\\
&\quad=-(1-\gamma)\mathbb{E}\{U[g(X+U)+g(X+U+1)-g(X)-g(X+1)]\}\\
&\qquad-\gamma\mathbb{E}\{(U-1)[g(X+U-1)+g(X+U)-g(X)-g(X+1)]\}+\eta,
\end{aligned}$$

where

$$|\eta|\leqslant 2\|f\|\mathbb{P}[X+U\notin[1,t]].$$

Hence

$$|\mathbb{E}[f(X+U-\gamma)-f(X)]|\leqslant 2\Delta g\{(1-\gamma)\mathbb{E}U^2+\gamma\mathbb{E}(U-1)^2\}+|\eta|,$$

and the result follows from (2.3). □

We are now able to prove the following approximation theorem, comparing the mixture $L_k(\theta;\, a,\, b)$ of shifted binomial random variables with the $\mathrm{Bi}(m,\frac{1}{2})$ distribution with the same mean $m=k-\bar{a}+2\bar{b}$, where $\bar{a}=\sum_{i\geqslant 1}\theta_i a_i$ and $\bar{b}=\sum_{i\geqslant 1}\theta_i b_i$.

Theorem 1. *Let k, $(a_i)_{i\geqslant 1}$ and $(b_i)_{i\geqslant 1}$ satisfy $1\leqslant b_i\leqslant a_i\leqslant k$, $i\geqslant 1$. Then, if m is chosen as above,*

$$d_{\mathrm{TV}}\big(L_k(\theta;\, a,\, b),\, \mathrm{Bi}(m,\tfrac{1}{2})\big)\leqslant m^{-1}\delta_1+m^{-3/2}\delta_2+\delta_3,$$

where

$$\begin{aligned}
\delta_1&=\delta_1(\theta;\, a,\, b)=2\bar{b}+\sum_{i\geqslant 1}\theta_i\big(a_i-2b_i-\bar{a}+2\bar{b}\big)^2,\\
\delta_2&=\delta_2(\theta;\, a,\, b,\, k)=c\sum_{i\geqslant 1}\theta_i(1+\varepsilon_i)|a_i-2b_i-\bar{a}+2\bar{b}|(|a_i-[\bar{a}]|+1),\\
\delta_3&=2cm^{-1/2}\sum_{i\geqslant 1}\theta_i|a_i-2b_i-\bar{a}+2\bar{b}|\mathrm{Bi}(k,\tfrac{1}{2})\{[0,\tfrac{1}{2}([\bar{a}]+a_i)]\cup[\lceil m\rceil,k]\}\\
&\quad+\tfrac{1}{2}L_k(\theta;\, a,\, b)\{(\lceil m\rceil,\infty)\},
\end{aligned}$$

and where

$$\varepsilon_i=m/\big(k-[\bar{a}]\vee a_i\big)-1$$

and c is the constant from (2.5).

Proof. Let $J_1,\dots,J_{k-1}$ and K be independent, $J_1,\dots,J_{k-1} \sim \mathrm{Be}(\frac{1}{2})$ and $\mathbb{P}[K=i]=\theta_i$, $i \geqslant 1$. Set $W=\sum_{i\geqslant 1} I[K=i](W_i+b_i)$, where $W_i=\sum_{r=1}^{k-a_i} J_r \sim \mathrm{Bi}(k-a_i, \frac{1}{2})$: then W has the mixture distribution $L_k(\theta;\, a,\, b)$. Now, in order to use (2.7), note that

$$\mathbb{E}\{Wg(W)\} = \sum_{i\geqslant 1} \theta_i \mathbb{E}\{(W_i+b_i)g(W_i+b_i)\}, \tag{2.10}$$

and that, since $W_i \sim \mathrm{Bi}(k-a_i, \frac{1}{2})$,

$$\mathbb{E}\{W_i g(W_i+b_i)\} = \mathbb{E}\{(k-a_i-W_i)g(W_i+b_i+1)\}, \tag{2.11}$$

for instance from (2.6). On the other hand,

$$\mathbb{E}\{(m-W)g(W+1)\} = \sum_{i\geqslant 1} \theta_i \mathbb{E}\{(m-b_i-W_i)g(W_i+b_i+1)\}. \tag{2.12}$$

Combining (2.10)–(2.12) gives

$$\begin{aligned}
&\mathbb{E}\{(m-W)g(W+1)-Wg(W)\}\\
&\quad= \sum_{i\geqslant 1} \theta_i \mathbb{E}\{[(m-b_i)-(k-a_i)]g(W_i+b_i+1)-b_i g(W_i+b_i)\}\\
&\quad= \sum_{i\geqslant 1} \theta_i \mathbb{E}\Big\{\big(2\bar{b}-\bar{a}-b_i+a_i\big)[g(W_i+b_i+1)-g(W_i^*)]-b_i[g(W_i+b_i)-g(W_i^*)]\Big\}\\
&\quad= \sum_{i\geqslant 1} \theta_i \mathbb{E}\Big\{\big(a_i-2b_i-\bar{a}+2\bar{b}\big)[g(W_i+b_i+1)-g(W_i^*)]\\
&\qquad\qquad\qquad\qquad +b_i[g(W_i+b_i+1)-g(W_i+b_i)]\Big\},
\end{aligned} \tag{2.13}$$

by choice of m, for any identically distributed random variables W_i^*.

Define $\alpha=[\bar{a}]$ and $\beta=\bar{b}-\frac{1}{2}\{\bar{a}-\alpha\}$, and take

$$W_i^* = \sum_{r=1}^{k-\alpha} J_r + \beta. \tag{2.14}$$

Then the expectations in (2.13) can be estimated by

$$\begin{aligned}
&\mathbb{E}|g(W_i^*)-g(W_i+b_i+1)|\\
&\quad\leqslant \tfrac{1}{2}\Delta g\,|a_i-2b_i-\bar{a}+2\bar{b}|+\mathbb{E}\big\{g(W_i^*)-g\big(W_i+\tfrac{1}{2}(a_i-\bar{a})+\bar{b}\big)\big\},
\end{aligned} \tag{2.15}$$

and

$$\mathbb{E}|g(W_i+b_i+1)-g(W_i+b_i)| \leqslant \Delta g, \tag{2.16}$$

where, if β is not integral, values of $g(x)$ for x not integral are obtained by linear interpolation. Finally, the expectation remaining in (2.15) can be estimated using Lemma 1, choosing

$$X=\sum_{r=1}^{k-(\alpha\vee a_i)} J_r, \qquad U=\sum_{r=k-(\alpha\vee a_i)+1}^{k-(\alpha\wedge a_i)} \big(J_r-\tfrac{1}{2}\big)+\gamma,$$

with $\gamma=\frac{1}{2}$ if $|a_i-\alpha|$ is odd and $\gamma=0$ otherwise, yielding the inequality

$$\begin{aligned}
&\Big|\mathbb{E}\big\{g(W_i^*)-g\big(W_i+\tfrac{1}{2}(a_i-\bar{a})+\bar{b}\big)\big\}\Big|\\
&\quad\leqslant \frac{2\,\|g\|}{k-(\alpha\vee a_i)}\big(|a_i-\alpha|+1\big)+2\,\|g\|\Big(1-\mathrm{Bi}\big(k,\tfrac{1}{2}\big)\big\{\big(\tfrac{1}{2}(\alpha+a_i),\,\lceil m\rceil\big]\big\}\Big),
\end{aligned}$$

after a little calculation. Hence

$$|\mathbb{E}\{(m-W)g(W+1)-Wg(W)\}|$$
$$\leqslant \Delta g \sum_{i\geqslant 1} \theta_i\left\{\tfrac{1}{2}\left(a_i-2b_i-\bar{a}+2\bar{b}\right)^2+b_i\right\}$$
$$+2\|g\|\sum_{i\geqslant 1}\theta_i|a_i-2b_i-\bar{a}+2\bar{b}|$$
$$\times\left\{m^{-1}(1+\varepsilon_i)(|\alpha-a_i|+1)+\left(1-\mathrm{Bi}\left(k,\tfrac{1}{2}\right)\left\{\left(\tfrac{1}{2}(\alpha+a_i),\lceil m\rceil\right]\right\}\right)\right\}, \quad (2.17)$$

and the theorem follows by taking $g=g_{m,A}$ and using (2.4), (2.5) and (2.7). □

Remark. In the application to the binary expansion of a random integer, k is large and $\Sigma_{i\geqslant 1}\theta_i(a_i^2+b_i^2)$ is of order 1. In such circumstances, the first term is the term of leading order. Since

$$\delta_1=4\left[\mathrm{Var}\ L_k(\theta;\ a,\ b)-\mathrm{Var}\ \mathrm{Bi}\left(m,\ \tfrac{1}{2}\right)\right],$$

the error estimate is effectively just the relative difference between the variances,

$$\left\{\mathrm{Var}\ L_k(\theta;\ a,\ b)/\mathrm{Var}\ \mathrm{Bi}\left(m,\ \tfrac{1}{2}\right)\right\}-1.$$

This is of a form precisely analogous to that frequently obtained in Poisson approximation problems using Stein's method.

Returning to the original problem, let

$$0\leqslant Z=\sum_{j=1}^{k}2^{k-j}Y_j<n$$

be the binary expansion of a randomly chosen integer between 0 and $n-1$, where $2^{k-1}<n\leqslant 2^k$. Our aim is to approximate the distribution of $W(n)=\Sigma_{j=1}^{n}Y_j$. To do so, we first need the following result, bounding $e(n)=\mathbb{E}W(n)$.

Lemma 2. *For all n such that* $2^{k-1}\leqslant n\leqslant 2^k$,

$$\tfrac{1}{2}(k-1)-\tfrac{1}{8}\leqslant e(n)\leqslant\tfrac{1}{2}k.$$

Proof. The upper bound follows because $\mathbb{E}Y_j\leqslant\frac{1}{2}$ for each $1\leqslant j\leqslant k$. This is clear because, when listing the integers in ascending order, equal length blocks of 0's and 1's alternate at any given place in the binary expansion, and the first block is of 0's.

For the lower bound, we use induction on k. The claim is easily checked for $k=0, 1, 2$. Now suppose that the lower bound holds for all $s<k$. Let $2^{k-1}\leqslant n\leqslant 2^k$, and let s be such that $2^{k-1}+2^{s-1}\leqslant n\leqslant 2^{k-1}+2^s$, so that $1\leqslant s\leqslant k-1$. Then it follows from the induction hypothesis that

$$e(n)\geqslant\frac{2^{k-1}}{n}\frac{k-1}{2}+\frac{n-2^{k-1}}{n}\left(\frac{s-1}{2}-\tfrac{1}{8}+1\right)$$
$$=\frac{k-1}{2}-\tfrac{1}{8}+\frac{n-2^{k-1}}{n}\left(1-\frac{k-s}{2}\right)+\frac{2^{k-1}}{8n}.$$

If $s = k-1$ or $k-2$, $e(n) \geqslant \frac{1}{2}(k-1) - \frac{1}{8}$ is immediate. For $1 \leqslant s \leqslant k-3$,

$$2^{k-1} + 4(n - 2^{k-1})(2-k+s) \geqslant 2^{k-1} + 2^{s+2}(2-k+s)$$
$$= 2^{k-1}\{1 - 2(k-s-2)2^{-(k-s-2)}\} \geqslant 0,$$

so that $e(n) \geqslant \frac{1}{2}(k-1) - \frac{1}{8}$ for such s also. □

If $n = 2^k$, $W(n) \sim \mathrm{Bi}(k, \frac{1}{2}) = \mathrm{Bi}(2e(n), \frac{1}{2})$, and so binomial approximation is exact. So suppose that n has binary expansion $n = \sum_{i=1}^{K} 2^{k-l(i)}$ as in (2.8); then $W(n) \sim L_k(\theta;\, a,\, b)$ as in (2.9), with

$$\theta_i = 2^{k-l(i)} \bigg/ \left\{ \sum_{r=1}^{K} 2^{k-l(r)} \right\}, \qquad a_i = l(i) \geqslant i, \qquad b_i = i-1,$$

and Theorem 1 can be applied. It is a straightforward matter to show that δ_1 and δ_2 are uniformly bounded. Furthermore, Lemma 2 shows that $\lceil m \rceil + 1 \geqslant k$, where, as before, $m = k - \bar{a} + 2\bar{b} = 2e(n)$, and $\mathbb{P}[W(n) = k] = 0$ for $2^{k-1} < n < 2^k$. The remaining part of δ_3 is also very small when k is large. Hence it follows from Theorem 1 that

$$d_{\mathrm{TV}}\left(\mathscr{L}(W(n)), \mathrm{Bi}(2e(n), \tfrac{1}{2})\right) = \mathrm{O}(k^{-1}).$$

The mixture of $\mathrm{Bi}(k-1, \frac{1}{2})$ with either $\mathrm{Bi}(k, \frac{1}{2})$ or $\mathrm{Bi}(k-2, \frac{1}{2})$ chosen to have mean $e(n)$ can likewise be shown, using Theorem 1, to differ in total variation from $\mathrm{Bi}(2e(n), \frac{1}{2})$ by at most an amount of order k^{-1}. Thus, by using this mixture of two neighbouring binomial distributions, the distribution of $W(n)$ is approximated to an accuracy of order k^{-1}.

Acknowledgement

We are indebted to Charles Stein for many helpful discussions about this work.

References

Barbour, A.D., L. Holst and S. Janson (1992), *Poisson Approximation* (Oxford Univ. Press, Oxford).

Chen, L.H.Y. and Y.T. Soon (1991), On the number of ones in the binary expansion of a random integer, in preparation.

Delange, H. (1975), Sur la fonction sommatoire de la fonction "somme des chiffres", *L'Enseign. Math. Sér. 2* **21**, 31–47.

Diaconis, P. (1977), The distribution of leading digits and uniform distribution mod 1, *Ann. Probab.* **5**, 72–81.

Loh, W.-L. (1991), Stein's method and multinomial approximation, in preparation.

Stein, C. (1986), *Approximate Computation of Expectations*, IMS Lecture Notes No. 7 (IMS, Hayward, CA).