AN APPROACH TO THE APPROXIMATE COMPUTATION OF PROBABILITIES AND EXPECTATIONS

I. THE BASIC IDEA AND A SIMPLE EXAMPLE

BY

CHARLES M. STEIN

TECHNICAL REPORT NO. 172
SEPTEMBER 1981

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MCS 80-24649

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
AN APPROACH TO THE APPROXIMATE COMPUTATION
OF PROBABILITIES AND EXPECTATIONS

I. THE BASIC IDEA AND A SIMPLE EXAMPLE

BY

CHARLES M. STEIN

TECHNICAL REPORT NO. 172
SEPTEMBER 1981

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT MCS80-24649

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
An Approach to the Approximate Computation
of Probabilities and Expectations

I. The basic idea and a simple example

By

Charles M. Stein
Stanford University

1. Introduction.

This is intended to be the first of a series of papers developing
in a leisurely and systematic way the method introduced in my paper on
normal approximation in the Sixth Berkeley Symposium and continued,
with appropriate innovations, by Chen (1974, 1975, 1978) Erickson
(1974), Tikhomirov (1980) and by the author in Stein (1978). In the
present paper I shall give only an abstract formulation of the approach
and a detailed treatment of a very special, almost trivial, problem, the
distribution of the number of ones in the binary expansion of a random
integer. The special problem will be discussed in Section 2 and the
abstract approach in Section 3. In Section 4 I shall try to compare
the other two sections.

It may be useful to give a brief description of my approach to the
problem, although the reader may find this impossible to understand with-
out referring to later sections. My aim is to study the distribution of
a random variable \( W = \omega(X) \) where \( X \) is a random point of the under-
lying sample space. First we introduce additional randomness by
constructing an exchangeable pair \((X, X')\) where \( X \) is a random point
of the original sample space. We wish to choose this pair in such a way that, roughly speaking, $X'$ differs only slightly from $X$ but any particular aspect of $X$ has positive probability of being changed in the transition from $X$ to $X'$. Second we introduce the space $\mathcal{F}$ of measurable anti-symmetric functions, say real-valued, on $\Omega^2$ and observe that for $F \in \mathcal{F}$

$$(1) \quad 0 = EF(X,X') = E[XF(X,X')] .$$

Under appropriate conditions, all random variables depending only on $X$ and having expectation $0$ are expressible in the form $XF(X,X')$ with $F \in \mathcal{F}$.

The third essential part of the approach is a systematic method for searching among the $XF(X,X')$ for a random variable that is nearly of the form $g(X) - C_g$ where $g(X)$ is a given random variable depending only on $X$ and $C_g$ is a constant. Then $Eg(X) = C_g$. Such a method is described around the diagram in (3.32) and occurs in a very special form in the transition (2.13)-(2.15) and in the elaboration of these formulas later in Section 2. I shall not attempt to describe this method here beyond saying that, roughly speaking, it depends on our ability to approximate the operator $T$ that associates with each $F \in \mathcal{F}$ the random variable $XF(X,X')$ by a simpler operator $T_0$, with which a different expectation operator $E_0$ is associated in the same way that $E$ is associated with $T$.

The problem treated in Section 2 was discussed from a similar point of view by Diaconis (1977). For background material on this question the reader may consult Delange (1975).
Except for Delange (1975), the references listed at the end of this paper use a method related to that of the present work. However I shall not try to survey this work, at least not in the first few papers of this series. Here I shall try to simplify the work as much as possible. This simplicity can be achieved either by considering abstract problems, where all complicating details are omitted or by considering very special concrete problems with no attempt at generality. I hope that in this way I shall be able gradually to develop the technique required to deal with more complicated questions.

This is a written version of a paper delivered in Victoria B.C. in June, 1981.
2. The number of ones in the binary expansion of a random natural number.

Let \( n \) be a natural number and \( X \) a random variable uniformly distributed over the set \( \{0, \ldots, n-1\} \). For the binary expansions of \( n-1 \) and \( X \) I shall write

\[
(1) \quad a = n-1 = \sum_{i=1}^{k} a_i 2^{k-i}
\]

and

\[
(2) \quad X = \sum_{i=1}^{k} x_i 2^{k-i},
\]

where

\[
(3) \quad 2^{k-1} < n \leq 2^k
\]

and the \( a_i \) and \( x_i \) take on values in the set \( \{0,1\} \). We are interested in the distribution of

\[
(4) \quad W = \sum_{i=1}^{k} x_i.
\]

When \( n = 2^k \), with \( k \) a positive integer, the distribution of \( W \) is the binomial distribution for \( k \) trials with probability \( \frac{1}{2} \). It is intuitively plausible that this should also hold approximately for all \( n \) satisfying (3) with \( k \) large and we shall see that this is true in a fairly strong sense.
In order to prove this let \( I \) be a random variable uniformly distributed over the set \( \{1, \ldots, k\} \) independent of \( X \), and let the random variable \( X' \) be defined by

\[
X' = \sum_{i=1}^{k} x'_i 2^{k-i},
\]

where

\[
X'_i = \begin{cases} 
X_i & \text{if } i \neq I \\
1-X_i & \text{if } i = I \text{ and this choice yields } X' < n \\
0 & \text{if } i = I, X_i = 0, \text{ and } X + 2^{k-I} \geq n.
\end{cases}
\]

Also let

\[
W' = \sum_{i=1}^{k} x'_i.
\]

The ordered pair \((X, X')\) of random variables is exchangeable, that is, it has the same distribution as the reversed pair \((X', X)\). It follows that the pair \((W, W')\) is also exchangeable.

One way to use these exchangeable pairs is to look at the identity

\[
P(W = w)P(W' = w+1|W = w)
\]

\[
= P(W = w \& W' = w+1) = P(W' = w \& W = w+1)
\]

\[
= P(W = w+1)P(W' = w|W = w+1).
\]
This is useful because, at least to a first approximation, the conditional probabilities occurring in (8) are easy to compute. In fact, for the conditional probability on the extreme right hand side we have

\[(9) \quad P(W' = w' | W = w+1) = P(X_I = 1 | W = w+1) = \frac{w+1}{k} .\]

In order to express the probability on the extreme left hand side of (8) in a convenient way I introduce the random variable

\[(10) \quad Q = |\{ j : X_j = 0 \& X + 2^{k-j} \geq n\}| ,\]

that is the number of \( j \in \{1, \ldots, k\} \) such that if \( I = j \) then according to (6) we have \( X'_I = X_I = 0 \). Then

\[(11) \quad P(W' = w+1 | W = w) = P(X_I = 0 \& X'_I = 1 | W = w) = 1 - \frac{w + E(Q | W = w)}{k} .\]

Thus (8) becomes

\[(12) \quad P(W = w) \left( 1 - \frac{w + E(Q | W = w)}{k} \right) - P(W = w+1) \frac{w+1}{k} = 0 .\]

Multiplying by \( kf(w) \) where \( f: \{0, \ldots, k\} \rightarrow \mathbb{R} \) is an arbitrary function, and summing over \( w \), we obtain

\[(13) \quad 0 = E[(k - W - E(Q)f(W) - Wf(W-1)] .\]
In order to compare the distribution of \( W \) with the binomial distribution we define the function \( h \) by

\[
(14) \quad h(w) = (k-w)f(w) - wf(w-1) .
\]

Then (13) becomes

\[
(15) \quad 0 = E[h(W) - (E^W f(W))] .
\]

It will be convenient to define the mapping \( \beta_{k,1/2} : \mathcal{Y} \rightarrow \mathcal{F} \) (where \( \mathcal{Y} \) is the space of functions on \( \{0, \ldots, k\} \) to \( \mathcal{F} \)) of expectation under the binomial distribution corresponding to \( k \) independent trials with probability \( 1/2 \):

\[
(16) \quad \beta_{k,1/2} = \frac{1}{2^k} \sum_{w=0}^{k} \binom{k}{w} g(w) .
\]

I shall think of (14) as an equation to be solved for \( f \) with \( h \) given and I shall need the following

**Lemma**: In order that there exist a function \( f : \{0, \ldots, k-1\} \rightarrow \mathcal{F} \) such that (14) holds for all \( w \in \{0, \ldots, k\} \) it is necessary and sufficient that

\[
(17) \quad \beta_{k,1/2} h = 0 .
\]

When this condition holds, the unique solution \( f \) of (14) is given by
\[(18) \quad f(w) = \sum_{v=0}^{w} \binom{k}{v} \frac{h(v)}{k-w} \]
\[= - \sum_{v=w+1}^{k} \binom{k}{v} \frac{h(v)}{k-w} . \]

for all \( w \in \{0, \ldots, k-1\} \).

**Proof.** First we observe that the values \( f(-1) \) and \( f(k) \), which are undefined, are multiplied by 0 when they occur in (14) so that no ambiguity arises. The necessity of (17) follows from the case \( n = 2^k \) of (13), since \( Q = 0 \) identically in this case. When (17) holds, we can verify that the first form of (18) satisfies (14) by direct substitution: For \( w < k \),

\[(19) \quad (k-w)f(w) - wf(w-1) \]
\[= \sum_{v=0}^{w} \binom{k}{v} h(v) - \sum_{v=0}^{w-1} \binom{k}{v} \frac{h(v)}{k-w+1} \cdot \frac{w}{w-1} h(v) = h(w) , \]

while for \( w = k \)

\[(20) \quad (k-w)f(w) - wf(w-1) = -kf(k-1) \]
\[= - \sum_{v=0}^{k-1} \binom{k}{v} h(v) = h(k) , \text{ by (17)} . \]

The uniqueness of the solution \( f \) is readily verified by induction.

Now let us look at the special case where \( g = g_{w_0} \), an indicator function defined by
\[ g_{w_0}(w) = \begin{cases} 
1 & \text{if } w = w_0 \\
0 & \text{otherwise} .
\end{cases} \]

Of course \( w_0 \in \{0, \ldots, k\} \). It is not difficult to verify that \( f_{w_0} \), related to \( g_{w_0} \), as \( f \) is related to \( g \) by (14) with

\[ h = g - \mathbb{E}_{w_0} \mathbb{E}_{w \neq w_0} g, \]

satisfies

\[ |f_{w_0}(w)| \leq \frac{2}{k} \]

for all \( w \). It is also not difficult to verify that

\[ \mathbb{E}|Q| \leq C, \]

an absolute constant. See Diaconis (1977) and Delange (1975). It follows from (15) and (21)-(24) that

\[ |p(w = w_0) - \frac{1}{2^k} \binom{k}{w_0}| = |\mathbb{E} g_{w_0}(w) - \mathbb{E} f_{w_0} - \mathbb{E} h| \leq \frac{2}{k} \mathbb{E}|Q| \leq \frac{2C}{k}. \]

Thus the probability that \( W \) has a given value differs from the corresponding binomial probability by \( O\left( \frac{1}{k} \right) \).
3. The abstract problem.

Let \((\Omega, \mathcal{F}, P)\) be a probability space, \((\mathcal{B}, \mathcal{F})\) a measurable space and \(\omega: (\Omega, \mathcal{F}) \rightarrow (\mathcal{B}, \mathcal{F})\) a measurable mapping. We are interested in approximating the distribution of

\[
W = \omega(X)
\]

where \(X\) is a random point of \(\Omega\) distributed according to \(P\), in other words the probability measure \(P \circ \omega^{-1}\). First we introduce additional randomness in the form of an exchangeable pair \((X, X')\) where \(X\) has the same distribution as before. The exchangeability means of course that the ordered pair \((X, X')\) has the same distribution as the reversed pair \((X', X)\). Ordinarily it will be desirable to choose this pair in such a way that, in an appropriate sense, \(X'\) differs only slightly from \(X\).

Let \(\mathcal{F}\) be the set of all real numbers and let \(\mathcal{F}\) be the linear space of all \(\mathcal{B}^2\)-measurable antisymmetric functions \(F: \Omega^2 \rightarrow \mathcal{F}\), antisymmetric in the sense that

\[
F(x, x') = -F(x', x)
\]

for all \(x, x' \in \Omega\), and having finite expectation \(\text{EF}(X, X')\). Let

\(T: \mathcal{F} \rightarrow \mathcal{B}\), where \(\mathcal{B}\) is the linear space of all real-valued \(\mathcal{B}\) measurable functions on \(\Omega\), be a linear operator such that, for all \(F \in \mathcal{F}\)

\[
(TF)(X) = \mathcal{F}^X F(X, X')
\]
Then for all $F \in \mathcal{F}$ we have

\begin{equation}
(4) \quad E(TF)X = EEF(X,X') = EF(X,X') = 0
\end{equation}

since

\begin{equation}
(5) \quad EF(X,X') = -EF(X',X) = -EF(X,X')
\end{equation}

The first equality in (5) uses the antisymmetry of $F$ and the second equality uses the exchangeability of $X$ and $X'$. Thus in the diagram

\begin{equation}
(6) \quad \mathcal{J} \rightarrow \mathcal{B} \rightarrow \mathbb{R}
\end{equation}

we have

\begin{equation}
(7) \quad \text{Im } T \subset \text{Ker } E.
\end{equation}

It will be instructive, although not really essential, to observe that, in fairly general circumstances, we have equality in (7). In the case where $\Omega$ is finite I shall state and prove this as a theorem.

**Theorem 1:** Let $(X,X')$ be an exchangeable pair of random variables taking values in a finite set $\Omega$ and suppose $\Omega$ is connected in the sense that for every pair $(x,x^*) \in \Omega^2$ there exists a sequence $x_0, \ldots, x_k$ such that $x_0 = x$, $x_k = x^*$ and, for every $j \in \{1, \ldots, k\}$
Let $\mathcal{G}$ be the linear space defined below (2), and $\mathcal{J}$ the linear space of antisymmetric functions defined in (2), and $T: \mathcal{J} \to \mathcal{G}$ the linear operator defined by (3). Then

\begin{equation}
\text{Im } T = \text{Ker } E.
\end{equation}

**Proof.** This turns out to be a disguised version of the well-known fact that the zeroth homology group of a connected simplicial complex is 0. For completeness I state and prove this as a lemma.

**Lemma:** Let $\mathcal{G}$ be a finite set and $\mathcal{J}$ a non-empty set of two-element subsets of $\mathcal{G}$ such that $(\mathcal{G}, \mathcal{J})$ is a connected graph, that is, for every $(s, s^*) \in \mathcal{G}^2$ there exists a sequence $s_0, \ldots, s_k$ with $s_0 = s$ and $s^* = s_k$ such that for every $j \in \{1, \ldots, k\}$, $(s_{j-1}, s_j) \in \mathcal{J}$. Let $\mathcal{G}$ be an additive abelian group, $\mathcal{G}_0$ the set of all functions $h: \mathcal{G} \to \mathcal{G}$ and $\mathcal{G}_1$ the set of all functions $\phi$ on the set $\overline{\mathcal{J}}$ of all $(s, s') \in \mathcal{G}^2$ such that $(s, s') \in \mathcal{J}$ taking values in $\mathcal{G}$ that are antisymmetric in the sense that for all $(s, s') \in \overline{\mathcal{J}},$

\begin{equation}
\phi(s, s') = -\phi(s', s).
\end{equation}

Let $S: \mathcal{G}_0 \to \mathcal{G}$ and $U: \mathcal{G}_1 \to \mathcal{G}_0$ be defined by

\begin{equation}
Sh = \sum_{s \in \mathcal{G}} h(s)
\end{equation}

and
\[(12) \quad (U\phi)(s) = \sum_{s' \text{ with } (s,s') \in \mathcal{U}} \phi(s,s') . \]

Then

\[(13) \quad \text{Im } U = \text{Ker } S . \]

**Proof.** Clearly \( \text{Im } U \subset \text{Ker } S \) since

\[(14) \quad S(U\phi) = \sum_{s \in \mathcal{S}} (U\phi)(s) = \sum_{s \in \mathcal{S}} \sum_{s' \text{ with } (s,s') \in \mathcal{U}} \phi(s,s') = 0 \quad \text{as in } (4) \text{ and } (5) . \]

I shall prove by induction on the number \( \nu \) of elements of \( \mathcal{S} \) that \( \text{Im } U \supset \text{Ker } S \). We want to prove that for any \( h \in \mathcal{C}_{0} \) such that \( Sh = 0 \) there exists \( \phi \in \mathcal{C}_{1} \) such that \( U\phi = h \). For \( \nu = 2 \) this is clear since, with \( \mathcal{S} = \{s_{1},s_{2}\} \) we need only take

\[(15) \quad \phi(s_{1},s_{2}) = h(s_{1}) \]

and

\[(16) \quad \phi(s_{2},s_{1}) = -h(s_{1}) = h(s_{2}) . \]

Assume the result true for \( \nu = \nu_{0} \), and suppose \( \mathcal{S} = \{s_{1},\ldots,s_{\nu_{0}+1}\} \). Choose \( k \) such that \( \{s_{k},s_{\nu_{0}+1}\} \in \mathcal{J} \) and, for a given \( h \in \mathcal{C}_{0} \) with \( Sh = 0 \) define \( h^{*}: \{s_{1},\ldots,s_{\nu_{0}}\} \to \mathcal{C} \) by
\[
(17) \quad h^*(s) = \begin{cases} 
  h(s) & \text{if } s \neq s_k \\
  h(s_k) + h(s_{0+1}) & \text{if } s = s_k.
\end{cases}
\]

Then \( Sh^* = 0 \) and consequently by the induction assumption there exists \( \phi^* \) such that \( U\phi^* = h^* \). Of course \( \phi^* \) is a function on the subset of \( \mathcal{I} \) consisting of those \( (s, s') \in \mathcal{I} \) with \( s_{0+1} \notin \{s, s'\} \), and \( U\phi^*, Sh^* \) are to be interpreted similarly. Now define \( \phi: \mathcal{I} \rightarrow \mathbb{Q} \) by

\[
(18) \quad \phi(s, s') = \begin{cases} 
  \phi^*(s, s') & \text{if } \{s, s'\} \cap \{s_k, s_{0+1}\} = \emptyset \\
  h(s_{0+1}) & \text{if } s = s_{0+1}, s' = s_k \\
  -h(s_{0+1}) & \text{if } s = s_k, s' = s_{0+1} \\
  0 & \text{otherwise}.
\end{cases}
\]

Then

\[
(19) \quad (U\phi)(s) = \begin{cases} 
  (U\phi^*)(s) = h^*(s) = h(s) & \text{if } s \notin \{s_k, s_{0+1}\} \\
  h(s_{0+1}) & \text{if } s = s_{0+1} \\
  (U\phi^*)(s_k) - h(s_{0+1}) = h^*(s_k) - h(s_{0+1}) = h(s_k) & \text{if } s = s_k.
\end{cases}
\]

Thus

\[
(20) \quad U\phi = h.
\]

This completes the proof of the lemma by induction.

Now let us return to the main theorem, (8). Define \( h: \Omega \rightarrow \mathbb{R} \) by

\[
(21) \quad h(x) = g(x) \mathbb{P}(X = x)
\]
and, using the fact that

(22) \[ 0 = \mathbb{E}_g(X) = \sum_{x \in \Omega} h(x) , \]

choose \( \phi \), in accordance with the lemma, such that

(23) \[ h = U\phi , \]

that is, for all \( x \in \Omega \)

(24) \[ h(x) = \sum_{(x,x') \in \overline{\mathcal{S}}} \phi(x,x') , \]

where \( \overline{\mathcal{S}} \) is the set of all ordered pairs \((x,x')\) of elements of \( \Omega \) such that

(25) \[ P\{X=x \& X'=x'\} \geq 0 , \]

and \( \phi \) is antisymmetric. But we can rewrite (24) as

(26) \[ g(x) = \sum_{(x,x') \in \overline{\mathcal{S}}} \frac{\phi(x,x')}{P\{X=x\}} \]

\[ = \sum_{(x,x') \in \overline{\mathcal{S}}} \frac{\phi(x,x')}{P\{X=x \& X'=x'\} \cdot P\{X'=x'|X=x\}} \]

\[ = \mathbb{E}\{F(X,X')|X=x\} \]

where \( F \) is the antisymmetric function defined by
\[
F(x, x') = \begin{cases} 
\frac{\phi(x, x')}{P[X=x' \& X'=x']} & \text{if } (x, x') \in \mathfrak{U} \\
0 & \text{otherwise .}
\end{cases}
\]

From (26) it follows that

\[(28) \quad g(x) = E^X F(X, X') .\]

Next I shall try to indicate a way in which the basic identity

\[(29) \quad E(E^X F(X, X')) = 0\]

can be used. Ordinarily we choose a manageable subspace \(\mathfrak{I}_0\) of \(\mathfrak{I}\) and try to approximate the operator \(T \circ \mathcal{I}: \mathfrak{I}_0 \to \mathfrak{B}\), where \(\mathcal{I}: \mathfrak{I}_0 \to \mathfrak{I}\) is the inclusion mapping, by a simpler operator \(T_0: \mathfrak{I}_0 \to \mathfrak{B}\), where \(\mathfrak{B}\) is a subspace of \(\mathfrak{B}\), typically consisting of the functions of \(\mathfrak{W}\), which was introduced in (1). Usually there will exist an expectation operator \(E_0: \mathfrak{B} \to \mathfrak{R}\) (not that associated with the true distribution of \(\mathfrak{W}\)) such that

\[(30) \quad \text{Im} \ T_0 = \ker E_0 .\]

Finally we choose \(V_0: \mathfrak{B} \to \mathfrak{I}_0\) in such a way that for all \(g \in \mathfrak{B}\)

\[(31) \quad T_0(V_0 g) = g - E_0 g .\]
The mappings considered here are displayed in the following diagram

\[
\begin{array}{ccc}
\mathcal{F} & \xrightarrow{T} & \bar{\mathcal{G}} & \xrightarrow{E} & \mathcal{L} \\
\uparrow & & \uparrow & & \\
\mathcal{I} & \xrightarrow{\mathcal{L}_0} & \mathcal{I}_0 & \xrightarrow{E_0} & \\
\mathcal{F}_0 & \xleftarrow{T_0} & \mathcal{G}_0 & \xleftarrow{V_0} & \\
\end{array}
\]

(32)

Of course the square and triangle in the diagram are not ordinarily commutative.

We then have

\[
0 = E[(T \circ \mathcal{L} \circ V_0)g](x)
\]

\[
= E[(\mathcal{L}_0 \circ T_0 \circ V_0)g](x) + E[(T \circ \mathcal{L} - \mathcal{L}_0 \circ T_0)(V_0g)](x)
\]

\[
= Eg(W) - E_0g + E[(T \circ \mathcal{L} - \mathcal{L}_0 \circ T_0)(V_0g)](x).
\]

(33)

If we can show that the remainder on the extreme right hand side of (33) is small, we have succeeded in approximating $Eg(W)$. Methods of accomplishing this typically depend on special features of the problem although a crude bound often follows easily from simple properties of the mapping $V_0$. Of course there is the possibility of applying the same method to the evaluation of the remainder but this tends to be complicated.
4. **The special problem in the abstract framework.**

It may be useful to describe explicitly the way the special problem fits into the abstract framework. The set \( \Omega = \{0, \ldots, n-1\} \), the probability measure \( P \) is uniform in \( \{0, \ldots, n-1\} \) and the random variable \( W \) is the number of ones in the binary expansion of the random point \( X \) of \( \{0, \ldots, n-1\} \). The exchangeable pair \((X,X')\) is constructed from \( X \) by choosing a random number \( I \), uniformly distributed in \( \{1 \ldots k\} \) independent of \( X \) where \( k \) is related to \( n \) by \( 2^{k-1} < n \leq 2^k \), and changing the coefficient of \( 2^{k-I} \) in the binary expansion of \( X \) to obtain \( X' \), provided this is less than \( n \). Otherwise \( X \) remains unchanged.

The full space \( \mathcal{J} \) of all antisymmetric functions of \((X,X')\) was not introduced explicitly in Section 2. Instead I considered only functions \( F \in \mathcal{J} \) having the form

\[
(1) \quad F(X,X') = f(W) \mathbb{1}_{\{W'=W+1\}} - f(W') \mathbb{1}_{\{W=W+1\}} .
\]

The basic identity (2.13) could have been obtained from (3.4), that is

\[
(2) \quad 0 = \mathbb{E}^X F(X,X') ,
\]

although in fact the argument was given a slightly different form. The operator \( T_0 \), obtained by dropping one term from the right hand side of (2.13), is given by

\[
(3) \quad (T_0 f)(w) = (k-w)f(w) - wf(w-1) .
\]
The expectation operator $E_0$ associated with $T_0$ by (3.30) is merely that corresponding to the binomial distribution with $k$ observations and probability $\frac{1}{2}$, denoted in Section 2 by $\Theta_{k,\frac{1}{2}}$. The operator $V_0$, related to $T_0$ by

\begin{equation}
T_0 \circ V_0 = I_{lb} - E_0,
\end{equation}

as indicated in (3.31), where $I_{lb}$ is the identity operator of the space $lb$ of real-valued functions of $W$, is given by

\begin{equation}
(V_0g)(w) = \frac{1}{k-w} \sum_{v=0}^{W} \prod_{t=v+1}^{W} \frac{t}{k+1-t} (g(v) - \Theta_{k,\frac{1}{2}})
\end{equation}

\begin{equation}
= -\frac{1}{w+1} \sum_{v=w+1}^{k} \prod_{t=w+2}^{k+1-t} \frac{v}{t} (g(v) - \Theta_{k,\frac{1}{2}}).
\end{equation}

Thus $V_0g$ is the function that is denoted by $f$ in (2.21).

The identity (2.19) is the special form taken by the abstract identity (4) above.
References


