ON A DIFFERENTIAL EQUATION AND ONE STEP RECURSION FOR POISSON MOMENTS

by

Anirban DasGupta
Purdue University

Technical Report #98–02

Department of Statistics
Purdue University
West Lafayette, IN USA

January 1998
ON A DIFFERENTIAL EQUATION AND
ONE STEP RECURSION FOR POISSON MOMENTS

by
Anirban DasGupta*

Purdue University

Abstract

Let \( X \sim \text{Poisson} (\lambda) \), and \( h(X) \) a given function of \( X \) such that \( E_\lambda h(X) \) exists for all \( \lambda \). We show that for each \( n \geq 1 \), \( E_\lambda (X^n h(X)) \) satisfies an \( n \)th order linear differential equation. The coefficients of this equation are explicit, and remarkably, they do not depend on the function \( h \). A consequence is that from an expression for \( E_\lambda (h(X)) \), one can derive a closed form expression for \( E_\lambda (X^n h(X)) \) for all \( n \geq 1 \). In addition, these lead to exact expressions for the \( n \)th moment and the \( n \)th central moment of a Poisson random variable and in particular show that the \( n \) moment of a Poisson random variable with mean 1 is the \( n \)th Bell number \( B_n \). These also characterize all functions \( h(X) \) that are positively correlated with \( X \).

We also present a general one step recursion formula for \( E_\lambda (X^n h(X)) \). These results may also facilitate computation of \( E_\lambda (X^n h(X)) \) as compared to direct computation from definition.

1. Introduction

The purpose of this article is to show that if \( X \) has a Poisson distribution with mean \( \lambda \), and \( h(X) \) is any function of \( X \), then for all positive integers \( n \), \( E_\lambda (X^n h(X)) \) admits an exact formula in terms of \( f(\lambda) = E_\lambda (h(X)) \) and its first \( n \) derivatives \( f^{(j)}(\lambda) \), \( j = 1, 2, \ldots, n \). Equivalently, one can assert that for each \( n \geq 1 \), \( E_\lambda (h(X)) \) itself satisfies an \( n \)th order linear differential equation, and it is remarkable that the coefficients of this differential equation do not depend on the function \( h \). This exact formula also leads to a one-step recursion formula for the moment sequence \( \{E_\lambda (X^n h(X))\}_{n \geq 1} \). Both of these facilitate closed form computation of \( E_\lambda (X^n h(X)) \) as compared to direct evaluation from

*Research supported by a National Security Agency grant.
definition. Our results, in addition, imply exact expressions for $E_\lambda(X^n)$, $E_\lambda(X - \lambda)^n$, and in particular, imply the fact that if $X \sim$ Poisson (1), then the $n$th moment of $X$ equals the $n$th Bell number $B_n$.

The preceding results also have some interesting covariance implications. For instance, it follows that if $X$ has zero correlation with $h(X)$ under every $\lambda$, then $h(X)$ must be a constant.

2. Notation and Preliminary Useful Facts

Throughout this article $S_2(n, i)$ will denote the Stirling second number defined as the number of partitions of a set of $n$ elements into $i$ nonempty disjoint subsets. Also, given any function $h$, $\Delta_1 h$ will denote the difference operator $\Delta_1 h(X) = h(X + 1) - h(X)$ and for $k \geq 2$, $\Delta_k h$ will denote the $k$th order iterated difference operator $\Delta_k h(X) = \Delta_1(\Delta_{k-1} h(X))$. As usual, $\Delta_0 h(X) = h(X)$. With this notation, we first state some lemmas that will be subsequently used.

**Lemma 1.** For all $n \geq 1$, $X^n = \sum_{i=1}^{n} S_2(n, i) \prod_{k=0}^{i-1} (X - k)$.

**Proof:** This is well known; see, e.g., pp. 125–126 in Bryant (1993).

**Lemma 2.** $S_2(n, i) = S_2(n - 1, i - 1) + i S_2(n - 1, i)$.

**Proof:** This is also well known; see Bryant (1993) again.

**Lemma 3.** For all $i \geq 0$, $h(X + i) = \sum_{j=0}^{i} \binom{i}{j} \Delta_j h(X)$

**Proof:** Fix any $n \geq 1$; then, the iterated differences $\Delta_0, \Delta_1, \ldots, \Delta_{n-1}$ are linear combinations of $h(X), h(X + 1), \ldots, h(X + n - 1)$, i.e.,

\[
\begin{pmatrix}
\Delta_0 h(X) \\
\Delta_1 h(X) \\
\vdots \\
\Delta_{n-1} h(X)
\end{pmatrix} = A_{n \times n} \cdot 
\begin{pmatrix}
h(X) \\
h(X + 1) \\
\vdots \\
h(X + n - 1)
\end{pmatrix},
\]

(2.1)
where the elements of $A$ are $a_{ij} = (-1)^{i+j} \binom{i-1}{j-1} 1 \leq i, j \leq n$. From (2.1),

$$
\begin{pmatrix}
    h(X) \\
    h(X + 1) \\
    \vdots \\
    h(X + n - 1)
\end{pmatrix} = A^{-1} \cdot 
\begin{pmatrix}
    \Delta_0 h(X) \\
    \Delta_1 h(X) \\
    \vdots \\
    \Delta_{n-1} h(X)
\end{pmatrix}.
\tag{2.2}
$$

One can directly verify that the elements of $A^{-1}$ are $a^{ij} = \binom{i-1}{j-1}$ and so the lemma follows.

**Lemma 4.** Let $X \sim \text{Poisson}\left(\lambda\right)$. Then, for all $i \geq 1$, $E\lambda(h(X)) \cdot \{ \prod_{k=0}^{i-1} (X - k) \} = \lambda^i E\lambda(h(X + i)).$

**Proof:** See Hwang (1982).

**Lemma 5.** Let $X \sim \text{Poisson}\left(\lambda\right)$. Then for all $k \geq 1$, $E\lambda(\Delta_k h(X)) = \frac{\partial^k}{\partial \lambda^k} E\lambda(h(X)).$

**Proof:** Let $p(\lambda, x)$ denote the Poisson pmf $\frac{e^{-\lambda} \lambda^x}{x!}$. Note that

$$
\frac{d}{d\lambda} p(\lambda, x) = \frac{x}{\lambda} p(\lambda, x) - p(\lambda, x) \tag{2.3}
$$

Therefore,

$$
\frac{d}{d\lambda} E\lambda(h(X)) = \sum_{x=0}^{\infty} h(x) \frac{x e^{-\lambda} \lambda^x}{\lambda x!} - \sum_{x=0}^{\infty} h(x) \frac{e^{-\lambda} \lambda^x}{x!} \\
= \sum_{x=1}^{\infty} h(x) \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} - \sum_{x=0}^{\infty} h(x) \frac{e^{-\lambda} \lambda^x}{x!} \\
= \sum_{x=0}^{\infty} h(x+1) \frac{e^{-\lambda} \lambda^x}{x!} - \sum_{x=0}^{\infty} h(x) \frac{e^{-\lambda} \lambda^x}{x!} \\
= E\lambda(\Delta_1 h(X)). \tag{2.4}
$$

Now the lemma follows on using the fact $\Delta_k h = \Delta_1 \Delta_{k-1} h$ and by induction.

3. A Linear Differential Equation for $Eh(X)$

For a general function $h(X)$, we will now present a linear differential equation satisfied by $E\lambda(h(X))$ in the following sense: fix any integer $n \geq 1$, and a function $h(X)$.
Denote $E_{\lambda}(h(X))$ by $f(\lambda)$. Then $f(\lambda)$ satisfies an $n$th order linear differential equation $c_{0,n}(\lambda)f(\lambda) + c_{1,n}(\lambda)f'(\lambda) + \cdots + c_{n,n}(\lambda)f^{(n)}(\lambda) = E_{\lambda}(X^n h(X))$. The coefficients $c_{0,n}, c_{1,n}, \ldots, c_{n,n}$ are explicit and they do not depend on the function $h$. In other words, if $X \sim \text{Poisson}(\lambda)$, then, one has the rather remarkable fact that one can write an explicit expression for $E(X^n h(X))$ for all $n \geq 1$ by only knowing an expression for $E(h(X))$!

**Theorem 1.** Let $X \sim \text{Poisson}(\lambda)$ and let $h(X)$ be such that $f(\lambda) = E_{\lambda}(h(X))$ exists for all $\lambda$. Then, for all $n \geq 1$, $E_{\lambda}(X^n h(X))$ also exists, and furthermore,

$$E_{\lambda}(X^n h(X)) = E_{\lambda}(X^n)E_{\lambda}(h(X)) + \sum_{j=1}^{n} c_{j,n}(\lambda) \frac{d^j}{d\lambda^j} f(\lambda),$$

(3.1)

where

$$c_{j,n}(\lambda) = \sum_{i=j}^{n} S_2(n, i) \left( \begin{array}{c} i \\ j \end{array} \right) \lambda^i.$$

(3.2)

**Proof:**

$$E_{\lambda}(X^n h(X)) = E_{\lambda}(h(X)) \cdot \sum_{i=1}^{n} S_2(n, i) \{ \prod_{k=0}^{i-1} (X - k) \} \quad \text{(By Lemma 1)}$$

$$= \sum_{i=1}^{n} S_2(n, i) E_{\lambda}(h(X)) \cdot \{ \prod_{k=0}^{i-1} (X - k) \}$$

$$= \sum_{i=1}^{n} S_2(n, i) \lambda^i E_{\lambda}(h(X + i)) \quad \text{(By Lemma 4)}$$

(3.3)

$$= \sum_{i=0}^{n} S_2(n, i) \lambda^i E_{\lambda}(h(X + i)) \quad (\because S_2(n, 0) = 0)$$

$$= \sum_{i=0}^{n} S_2(n, i) \lambda^i \sum_{j=0}^{i} \left( \begin{array}{c} i \\ j \end{array} \right) E_{\lambda}(\Delta_j h(X)) \quad \text{(By Lemma 3)}$$

$$= \sum_{i=0}^{n} \sum_{j=0}^{i} S_2(n, i) \left( \begin{array}{c} i \\ j \end{array} \right) \lambda^i \frac{d^j}{d\lambda^j} E_{\lambda}(h(X)) \quad \text{(By Lemma 5)}$$

(3.4)

$$= \sum_{j=0}^{n} \sum_{i=j}^{n} S_2(n, i) \left( \begin{array}{c} i \\ j \end{array} \right) \lambda^i \frac{d^j}{d\lambda^j} E_{\lambda}(h(X))$$

$$= \sum_{i=0}^{n} S_2(n, i) \lambda^i E_{\lambda}(h(X)) + \sum_{j=1}^{n} \sum_{i=j}^{n} S_2(n, i) \left( \begin{array}{c} i \\ j \end{array} \right) \lambda^i \frac{d^j}{d\lambda^j} E_{\lambda}(h(X)).$$

(3.5)
From (3.4), on using \( h(X) = 1 \), one gets

\[
E_\lambda(X^n) = \sum_{i=0}^{n} S_2(n, i)\lambda^i.
\]

(3.6)

Now substituting (3.6) into (3.5), for a general \( h \),

\[
E_\lambda(X^n h(X)) = E_\lambda(X^n) E_\lambda(h(X)) + \sum_{j=1}^{n} c_{j,n}(\lambda) \frac{d^j}{d\lambda^j} f(\lambda),
\]

as claimed.

This derivation of (3.1) itself implies that if \( E_\lambda(h(X)) \) exists for all \( \lambda \), then \( E_\lambda(X^n h(X)) \) also exists for all \( \lambda \) and \( n \geq 1 \). This proves Theorem 1.

From (3.6), one immediately gets the following fact as a corollary.

**Corollary 1.** Let \( X \sim \text{Poisson} (1) \). Then \( E(X^n) = B_n \) = the \( n \)th Bell number = Total number of partitions of a set of \( n \) elements into disjoint nonempty subsets.

**4. A General Moment Recursion Formula**

The expansion in (3.3) will be now used to obtain an interesting one step recursion relation for \( E_\lambda(X^n h(X)) \).

**Theorem 2.** Let \( X \sim \text{Poisson} (\lambda) \) and let \( h(X) \) be such that \( E_\lambda(h(X)) \) exists for all \( \lambda \). Then, for all \( n \geq 1 \),

\[
E_\lambda(X^n h(X)) = \lambda \{ E_\lambda(X^{n-1} h(X)) + \frac{d}{d\lambda} E_\lambda(X^{n-1} h(X)) \}
\]

(3.6)

**Proof:** From (3.3),

\[
E_\lambda(X^n h(X))
\]

\[
= \sum_{i=1}^{n} S_2(n, i)\lambda^i E_\lambda(h(X + i))
\]

\[
= \sum_{i=1}^{n} S_2(n - 1, i - 1)\lambda^i E_\lambda(h(X + i)) + \sum_{i=1}^{n} i S_2(n - 1, i)\lambda^i E_\lambda(h(X + i)) \quad \text{(By Lemma 2)}
\]

\[
= \lambda \sum_{i=0}^{n-1} S_2(n - 1, i)\lambda^i E_\lambda(h(X + i + 1)) + \lambda \sum_{i=1}^{n} S_2(n - 1, i)(\lambda^i)' E_\lambda(h(X + i))
\]
\[
= \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i E_\lambda(h(X + i + 1)) + \lambda \sum_{i=1}^{n-1} S_2(n-1,i)(\lambda^i)' E_\lambda(h(X + i))
\]
\[\quad (\because S_2(n-1,0) = S_2(n-1,n) = 0)\]
\[
= \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i E_\lambda(h(X + i + 1)) + \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\frac{d}{d\lambda}(\lambda^i E_\lambda(h(X + i)))
\]
\[
- \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i \frac{d}{d\lambda} E_\lambda(h(X + i))
\]
\[
= \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i E_\lambda(h(X + i + 1)) + \lambda \frac{d}{d\lambda} E_\lambda(X^{n-1}h(X))
\]
\[
- \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i \frac{d}{d\lambda} E_\lambda(h(X + i)) \quad \text{(By (3.3))}
\]
\[
= \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i \{E_\lambda(h(X + i + 1)) - E_\lambda(h(X + i))\}
\]
\[
\quad \quad \quad \text{(By Lemma 5)}
\]
\[
= \lambda \frac{d}{d\lambda} E_\lambda(X^{n-1}h(X)) + \lambda \sum_{i=1}^{n-1} S_2(n-1,i)\lambda^i E_\lambda(h(X + i))
\]
\[
= \lambda \{\frac{d}{d\lambda} E_\lambda(X^{n-1}h(X)) + E_\lambda(X^{n-1}h(X))\}. \quad \text{(By (3.3) again)}
\]

This proves Theorem 2.

5. Some Applications

We close this article with three specific applications of the results given in the previous sections.

**Theorem 3.** Let \( X \sim \text{Poisson} (\lambda) \) and \( h(X) \) is such that \( f(\lambda) = E_\lambda(h(X)) \) exists for all \( \lambda \). Then,

(a) \( \text{Cov}_{\lambda_0}(X, h(X)) \geq 0 \) at a specified \( \lambda_0 \) if and only if \( h(X) \) has an increasing expectation  
locally at \( \lambda_0 \), i.e., \( f'(\lambda_0) \geq 0 \);

(b) There is no nontrivial function \( h \) such that \( \text{Cov}_{\lambda}(X, h(X)) = 0 \) for all \( \lambda \).
Proof: (a) By Theorem 1, \( E_\lambda(X h(X)) = E_\lambda(X) E_\lambda(h(X)) + c_{1,1}(\lambda) f'(\lambda) \), where \( c_{1,1}(\lambda) = \lambda S_2(1, 1) = \lambda \). Thus,

\[
\text{Cov}_\lambda(X, h(X)) = \lambda f'(\lambda) \geq 0 \tag{5.1}
\]

if and only if \( f'(\lambda) \geq 0 \).

(b) From (5.1), \( \text{Cov}_\lambda(X, h(X)) = 0 \ \forall \lambda \)

\[ \iff f'(\lambda) = 0 \ \forall \lambda \]

\[ \iff f(\lambda) = \text{constant} \]

\[ \iff h(X) = \text{constant}, \]

by the completeness of the Poisson family (see Lehmann and Casella (1998)). This proves Theorem 3.

The next result is on a formula for the central moments of a Poisson distribution.

**Theorem 4.** Let \( X \sim \text{Poisson} (\lambda) \). Then,

(a) For any \( n \geq 1 \),

\[ E_\lambda(X - \lambda)^n = \sum_{k=0}^{n} a_{k,n} \lambda^k, \]

where

\[ a_{k,n} = \sum_{i=0}^{k} (-1)^i \binom{n}{i} S_2(n-i, k-i). \tag{5.2} \]

in the above, \( S_2(0,0) = 1 \);

(b) For any \( n \geq 1 \), the leading coefficient \( a_{n,n} = 0 \);

(c) For \( n \geq 3 \), \( a_{n-1,n} \) is also 0.

Proof:

(a) By Theorem 1 and Binomial expansion,
\[ E_\lambda(X - \lambda)^n = \sum_{k=0}^{n} \sum_{i=0}^{k} (-1)^{n-k} \binom{n}{k} S_2(k, i) \lambda^{n-k+i} \]
\[ = \sum_{i=0}^{n} \sum_{k=i}^{n} (-1)^{n-k} \binom{n}{k} S_2(k, i) \lambda^{n-k+i} \]
\[ = \sum_{i=0}^{n} \sum_{k=i}^{n-i} (-1)^{n-k-i} \binom{n}{k+i} S_2(k+i, i) \lambda^{n-k} \quad \text{(write } k \text{ for } k - i) \]
\[ = \sum_{i=0}^{n} \sum_{k=i}^{n-i} (-1)^{k-i} \binom{n}{n-k+i} S_2(n-k+i, i) \lambda^{k} \quad \text{(write } k \text{ for } n-k) \]
\[ = \sum_{i=0}^{n} \sum_{k=i}^{n-k-i} (-1)^{k-i} \binom{n}{k-i} S_2(n-k+i, i) \lambda^{k} \]
\[ = \sum_{k=0}^{n} \sum_{i=0}^{k} (-1)^{k-i} \binom{n}{k-i} S_2(n-k+i, i) \lambda^{k} \quad \text{(write } i \text{ for } k - i) \]
\[ = \sum_{k=0}^{n} a_{k,n} \lambda^{k}, \]

as claimed.

(b) From (5.2), the coefficient of \( \lambda^n \) is
\[ a_{n,n} = \sum_{i=0}^{n} (-1)^i \binom{n}{i} S_2(n-i, n-i) \]
\[ = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \]
\[ = 0. \]

(c) Again, from (5.2), the coefficient of \( \lambda^{n-1} \) is
\[ a_{n-1,n} = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} S_2(n - i, n - i - 1) \]

\[ = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} \binom{n - i}{2} \]

(\because S_2(m, m - 1) = \binom{m}{2}; \text{ see pp. 18 in Tomescu (1985)})

\[ = \sum_{i=0}^{n-2} (-1)^i \binom{n}{i} \binom{n - i}{2} \]

\[ = \frac{n(n - 1)}{2} \sum_{i=0}^{n-2} (-1)^i \binom{n - 2}{i} \]

\[ = 0 \text{ if } n > 2, \]

as claimed

It turns out that for the Poisson distribution, the third central moment is also \( \lambda \). This is stated below.

**Corollary 2.** If \( X \sim \text{Poisson} (\lambda) \),

\[ E_\lambda(X - \lambda)^2 = E_\lambda(X - \lambda)^3 = \lambda, \quad E_\lambda(X - \lambda)^4 = \lambda + 3\lambda^2. \]

**Proof:** Follows from (5.2)

**References**


