# Moments of Sums of Independent and Identically Distributed Random Variables

Daniel M. Packwood

Department of Chemistry
Graduate School of Science
Kyoto University, Kyoto, 606-8502, Japan

June 1, 2011

**Abstract**

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables. We present an analytic method for computing the moments of $S_n = \sum_{i=1}^{n} X_i$. The method is illustrated with a simple example, and is used to prove the strong law of large numbers.

Key words: *iid* random variables, sums of *iid* random variables.

# 1 Introduction

We consider the moments of the sum $S_n$ of the independent and identically distributed (*iid*) random variables $X_1, X_2, \ldots, X_n$. The distribution of these random variables is assumed to be the same as the random variable $X$. Such sums often appear in probability and statistics (e.g., in the law of large numbers) and in stochastic analysis (e.g., in the study of random walks). There are many papers which prove bounds for the moments (see e.g., [1], [2] and [3] and the references therein), however to the best of my knowledge there are no reports of a general analytic approach to their calculation.

# 2 Central result

Fix a $p \in \{1, 2, \ldots\}$ and define

$$\mathcal{Q}^p = \left\{ \begin{array}{ll} E(X^r)E(X^s)\cdots E(X^t): & r, s, \ldots, t \in \{1, 2, \ldots, p\} \\ & r + s + \cdots + t = p \end{array} \right\}.$$

The $p$th moment of $S_n = \sum_{i=1}^n X_i$ is

$$E(S_n^p) = \sum_{q_i \in \mathcal{Q}^p} a_i q_i, \tag{1}$$

where, for $q_i = E(X^{p_1})E(X^{p_2})\cdots E(X^{p_m})$,

$$a_i = \frac{1}{l_1! l_2! \cdots l_h!} \frac{n!}{(n-m)!} \frac{p!}{p_1! p_2! \cdots p_m!}. \tag{2}$$

In equation (2), $h$ is the number of distinct constants in the sequence $\{p_1, p_2, \ldots, p_m\}$, $l_1$ the number of elements equal to the first constant, $l_2$ the number equal to the second constant, ..., and $l_h$ the number equal to the $h$th constant (e.g., for the sequence $\{1, 1, 1, 1\}$, $h = 1$ and $l_1 = 4$, and for $\{1, 2, 2, 1\}$, $h = 2$, $l_1 = 2$, and $l_2 = 2$). Note that $E(S_n^p) < \infty$ if an only if $E(X^\alpha) < \infty$ for all integers $\alpha \leq p$.

# 3 Example calculation

The third and fourth moments work out to be

$$\begin{aligned}
E(S_n^3) &= \left(\frac{1}{1!}\frac{n!}{(n-1)!}\frac{3!}{3!}\right)E(X^3) + \left(\frac{1}{1!1!}\frac{n!}{(n-2)!}\frac{3!}{2!1!}\right)E(X^2)E(X) \\
&+ \left(\frac{1}{3!}\frac{n!}{(n-3)!}\frac{3!}{1!1!1!}\right)E(X)^3 \\
&= nE(X^3) + 3n(n-1)E(X^2)E(X) + n(n-1)(n-2)E(X)^3
\end{aligned}$$

$$\begin{aligned}
E(S_n^4) &= \left(\frac{1}{1!}\frac{n!}{(n-1)!}\frac{4!}{4!}\right)E(X^4) + \left(\frac{1}{1!1!}\frac{n!}{(n-2)!}\frac{4!}{3!1!}\right)E(X^3)E(X) \\
&+ \left(\frac{1}{2!}\frac{n!}{(n-2)!}\frac{4!}{2!2!}\right)E(X^2)^2 + \left(\frac{1}{2!1!}\frac{n!}{(n-3)!}\frac{4!}{2!1!1!}\right)E(X^2)E(X)^2 \\
&+ \left(\frac{1}{4!}\frac{n!}{(n-4)!}\frac{4!}{1!1!1!1!}\right)E(X)^4 \\
&= nE(X^4) + 4n(n-1)E(X^3)E(X) + 3n(n-1)E(X^2)^2 \\
&+ 6n(n-1)(n-2)E(X^2)E(X)^2 + n(n-1)(n-2)(n-3)E(X)^4
\end{aligned}$$

If $X_1, X_2, \ldots, X_n$ are independent normal random variables with mean 1 and variance 1, the above formulae give $E(S_{10}^3) = 1300$ and $E(S_{10}^4) = 16300$. These values, as well as various others calculated with (1) and (2), agree with estimates from numerical simulations performed in *R 2.12.0* [4]. These simulations involved generating $10^4$ to $10^6$ samples of $S_n$ and calculating the average of their $p$th power.

A certain amount of tedium can be avoided by supposing that $X$ has a symmetric distribution about zero. In this case, all factors $q_i$ in the expansion in equation (1) involving an odd moment of $X$ vanish, and therefore all odd moments of $S_n$ also vanish. The even moments take on relatively simple forms, for example

$$E(S_n^4) = nE(X^4) + 3n(n-1)E(X^2)^2.$$

For certain applications involving sums of random variables, inequalities for the moments and coefficients $\{a_i\}$ may be more useful than the moments themselves. Equations (1) and (2) allow for these to be obtained in a straightforward way. For example, in general we have

$$E(S_n^p) \leq nE(X^p), \tag{3}$$

and we can see from equation (2) that

$$a_i \leq n!p!. \tag{4}$$

# 4 Application: The law of large numbers

The various laws of large numbers involve sums of *iid* random variables. For example, Cantelli's theorem states that if $X_1, X_2, \ldots, X_n$ are *iid* random variables with mean zero and finite fourth moments, then $S_n/n \to 0$ almost surely. The results derived above provide new ways to investigate the laws of large numbers, and here we use them to give a straightforward proof of Cantelli's theorem.

Fix a positive integer $n$ and $\epsilon > 0$. Applying Chebyshev's inequality to the random variable $S_n^2$ gives $P\left(|S_n^2| > n^2\epsilon^2\right) \leq E(|S_n^2|^2)/n^4\epsilon^4$, or equivalently,

$$P\left(|S_n| > n\epsilon\right) \leq \frac{E(|S_n|^4)}{n^4\epsilon^4}. \tag{5}$$

Summing both sides of equation (6) and substituting in the inequality (3) gives

$$\sum_{n=1}^{\infty} P\left(|S_n| > n\epsilon\right) \leq \sum_{n=1}^{\infty} \frac{E(X^4)}{n^3\epsilon^4} < \infty. \tag{6}$$

The Borel-Cantelli lemma therefore implies that

$$P(\omega \text{ that are in infinitely many } \{|S_n| > n\epsilon\}) = 0,$$

and so there exists an $m$ such that for $n > m$, $P(|S_n| < n\epsilon) = 1$. Letting $\epsilon \to 0$ then gives the result.

# 5 Proof of equations (1) and (2)

The $p$th moment of $S_n$ is

$$E(S_n^p) = \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}\cdots\sum_{k=1}^{n}}_{p \text{ sums}} E(X_i X_j \cdots X_k), \tag{7}$$

Because $X_1, \ldots, X_n$ are *iid* random variables, each term in the sum can be factored into the form

$$E(X^{p_1})E(X^{p_2})\cdots E(X^{p_n}),$$

where $p_1, p_2, \ldots, p_n$ are positive integers that sum to $p$. (7) can therefore be written as (1). To determine the constants $a_1, a_2, \ldots$ fix a positive value of $p$ and positive non-zero integers $p_1, p_2, \ldots, p_m$ such that $p_1 + p_2 + \ldots + p_m = p$. For each choice of $i, j, \ldots, k \in \{1, 2, \ldots, n\}$, with $i \neq j \neq, \ldots \neq k$, define the collection

$$\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}} = \left\{ \begin{array}{c} \text{all unique permutations of the sequence} \\ \underbrace{X_i, X_i, \ldots X_i}_{p_1 \text{ factors}}, \underbrace{X_j, X_j, \ldots X_j}_{p_2 \text{ factors}}, \ldots, \underbrace{X_k, X_k, \ldots X_k}_{p_m \text{ factors}} \end{array} \right\}$$

Each of the sequences in $\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}}$ corresponds to an expectation of the form $E(X_i^{p_1} X_j^{p_2} \cdots X_k^{p_m})$. These are each equivalent to

$$q_m = E(X^{p_1})E(X^{p_2})E(X^{p_m})$$

The subscript $m$ on $q_m$ is only for bookkeeping purposes. The number of elements in $\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}}$ is the number of ways in which $q_m$ can be created from $p_1$ copies of $X_i$, $p_2$ copies of $X_j$, $\ldots$, $p_m$ copies of $X_k$. For any choice of $i, j, \ldots, k$, this number is

$$|\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}}| = \frac{p!}{p_1! p_2! \cdots p_m!}. \tag{8}$$

The number of elements in the union

$$\mathcal{Q}^p_{p_1, \ldots, p_m} = \bigcup_{\substack{i, j, \ldots, k \in \{1, \ldots, n\} \\ i \neq j \neq \cdots \neq k}} \mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}} \tag{9}$$

is the number of ways $q_m$ can be created from the $p_1$ copies of $X_i$, $p_2$ copies of $X_j$, $\ldots$, $p_m$ copies of $X_k$, $n$ choices of $i$, $n-1$ choices of $j$, $\ldots$, $n-(m-1)$ choices of $k$. This gives the identity

$$a_m = |\mathcal{Q}^p_{p_1, \ldots, p_m}|. \tag{10}$$

If there are no equalities between the constants $p_1, p_2, \ldots, p_m$, then the collections $\{\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}}\}$ are mutually exclusive and so $|\mathcal{Q}^p_{p1,\ldots,p_m}|$ is given by (8) times the number of collections in the union, namely

$$n(n-1)\cdots(n-(m-1))\frac{p!}{p_1!p_2!\cdots p_m!}.$$

If there are equalities between $p_1, p_2, \ldots p_m$, then some of the collections in $\{\mathcal{C}_{i_{p_1}, j_{p_2}, \ldots, k_{p_m}}\}$ are equivalent. For example, if $p_1 = p_2 = p_3$, then

$$\begin{aligned} \mathcal{C}_{a_{p_1}, b_{p_2}, c_{p_3}, \ldots, k_{p_m}} &= \mathcal{C}_{b_{p_1}, c_{p_2}, a_{p_3}, \ldots, k_{p_m}} \\ &= \mathcal{C}_{c_{p_1}, a_{p_2}, b_{p_3}, \ldots, k_{p_m}} \\ &= \mathcal{C}_{c_{p_1}, b_{p_2}, a_{p_3}, \ldots, k_{p_m}} \\ &= \mathcal{C}_{a_{p_1}, c_{p_2}, b_{p_3}, \ldots, k_{p_m}} \\ &= \mathcal{C}_{b_{p_1}, a_{p_2}, c_{p_3}, \ldots, k_{p_m}} \end{aligned}$$

In the above equalities, the indices given by the ellipses and $k_{p_m}$ do not change. Furthermore, if we also had $p_4 = p_5 = p_6 = p_7$ in addition to $p_1 = p_2 = p_3$, then we could also write an additional 4! equivalent collections for each of the 3! equivalent collections given above, leading to 4!3! equivalent collections in the union (9). In general, if there are $h$ distinct constants in the sequence $\{p_1, p_2, \ldots p_m\}$, with $l_1$ of the elements equal to one of these constants, $l_2$ equal to another of these constants, ..., $l_h$ equal to the remaining of these constants, then

$$l_1!l_2!\cdots l_h!$$

of the collections in the union (9) are equivalent. The number of elements in $\mathcal{Q}^p_{p_1,\ldots,p_m}$ for the general case is therefore found by taking the mutually exclusive result given above and dividing through by $l_1!l_2!\cdots l_h!$, namely

$$|\mathcal{Q}^p_{p_1,\ldots,p_m}| = \frac{1}{l_1!l_2!\cdots l_h!}\frac{n!}{(n-m)!}\frac{p!}{p_1!p_2!\cdots p_m!}.$$

Equation (10) then gives the results of section 2.

**Acknowledgements**
D.M.P is supported by a Japan Society for the Promotion of Science Postdoctoral Fellowship. An anonymous reviewer of an earlier version of this manuscript is thanked for providing the strong law of large numbers application in section 2, as well as various other suggestions.

## References

[1] R. Latala, *Ann. Probab.* **25**, 1997, 1502 - 1513.

[2] P. Hitczenko, S. J. Montgomery-Smith and K. Oleszkiewiez, *Stud. Math.* **123**, 1997, 15 - 45.

[3] P. Hitczenko and S. J. Montgomery-Smith, *Ann. Probab.* **29**, 2001, 447 - 466.

[4] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2009.