On the Moments of Functions of Random Variables Using Multivariate Taylor Expansion, Part I

Noel van Erp ¹, Pieter van Gelder ¹
¹ Structural Hydraulic Engineering and Probabilistic Design, TU Delft, Delft, The Netherlands

Introduction
In this article we give a general procedure to find the mean, variance and higher order centralized moments for arbitrary functions of random variables, provided these functions have partial derivatives of order \((n+1)\) equal to zero. We will demonstrate this procedure by finding the general equations for the mean, variance and skewness for sample means; of which the general equations for the mean and variance respectively, equations (13) and (16) given below, are well-known from the various statistical textbooks. Van Erp and Van Gelder (2007) present an algorithm for the developed models in this paper.

Multivariate Taylor Expansion
The following Taylor’s Theorem in \(m\) variables is a generalization of Taylor’s Theorem in 2 variables as given in the standard calculus text by Thomas and Finney (1996). We simply state the result here, and for a proof we refer to (Thomas and Finney, 1996).

Taylor’s Theorem
Suppose that \(g(x_1, \ldots, x_m)\) and all its partial derivatives of order less than or equal to \((n+1)\) are continuous on \(D = \{(x_1, \ldots, x_m) | a_i \leq x_i \leq b_i, \ldots, a_m \leq x_m \leq b_m\}\), and let
For every \((c_1, \ldots, c_m) \in D\), there exists a point \((\xi_1, \ldots, \xi_m) \in D\) lying in between \((x_1, \ldots, x_m)\) and \((c_1, \ldots, c_m)\) with

\[
g(x_1, \ldots, x_m) = P_n(x_1, \ldots, x_m) + R_n(x_1, \ldots, x_m) \tag{1}
\]

where

\[
P_n(x_1, \ldots, x_m) = g(c_1, \ldots, c_m) + \sum_{j=1}^{m} \frac{\partial g}{\partial x_j}(c_1, \ldots, c_m) (x_j - c_j)
\]

\[+ \frac{1}{2} \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{\partial^2 g}{\partial x_{j_1} \partial x_{j_2}}(c_1, \ldots, c_m) (x_{j_1} - c_{j_1})(x_{j_2} - c_{j_2})
\]

\[+ \cdots
\]

\[+ \frac{1}{n!} \sum_{j_1=1}^{m} \cdots \sum_{j_n=1}^{m} \frac{\partial^n g}{\partial x_{j_1} \cdots \partial x_{j_n}}(c_1, \ldots, c_m) (x_{j_1} - c_{j_1}) \cdots (x_{j_n} - c_{j_n})
\]

and

\[
R_n(x_1, \ldots, x_m) = \frac{1}{(n+1)!} \sum_{j_1=1}^{m} \cdots \sum_{j_{n+1}=1}^{m} \frac{\partial^{n+1} g}{\partial x_{j_1} \cdots \partial x_{j_{n+1}}}(\xi_1, \ldots, \xi_m) (x_{j_1} - c_{j_1}) \cdots (x_{j_{n+1}} - c_{j_{n+1}})
\]

The function \(P_n\) is called the \(n\)th order Taylor polynomial in \(m\) variables for the function \(g\) about \((c_1, \ldots, c_m)\), and \(R_n\) is the remainder term associated with \(P_n\).

**Corollary**

If all the partial derivatives of order \((n+1)\) of \(g(x_1, \ldots, x_m)\) equal zero then, we have that
\[ g(x_1, \ldots, x_m) = P_n(x_1, \ldots, x_m) \]  \hspace{1cm} (4)

The proof for this Corollary follows trivially from the fact that the remainder term \( R_n \) must equal zero if all the partial derivatives of order \( (n+1) \) also equal zero, as can be seen in (3).

**Functions of Random variables**

We now want to determine for an arbitrary function

\[ Z = g(X_1, \ldots, X_m) \]  \hspace{1cm} (5)

of \( m \) known random variables, \( (X_1 \cdots X_m) \sim p(x_1 \cdots x_m \mid \theta) \), what the centralized moments are of the probability density function of \( Z \sim p(z \mid \theta) \).

If all the partial derivatives of order \( (n+1) \) of \( g(x_1, \ldots, x_m) \) equal zero, then we have with our Corollary that equality (5) may be written as \( Z = P_n(X_1, \ldots, X_m) \), which in turn implies the following equality

\[ E(Z^k) = E(P_n(X_1, \ldots, X_m)^k), \quad k = 1, 2, \ldots \]  \hspace{1cm} (6)

Now if we let \( P_n \) be about the point \( (\mu_1, \ldots, \mu_m) \), where the \( \mu_j \)'s, \( j = 1, \ldots, m \), are the expectation values of the known random variables \( (X_1 \cdots X_m) \sim p(x_1 \cdots x_m \mid \theta) \), then we may rewrite (3) as a function of centralized random variables
\[ P_n(x_1, \ldots, x_m) = g(\mu_1, \ldots, \mu_m) + \sum_{j=1}^{m} \frac{\partial g}{\partial x_j} (\mu_1, \ldots, \mu_m) (x_j - \mu_j) \]
\[ + \frac{1}{2} \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{\partial^2 g}{\partial x_{j_1} \partial x_{j_2}} (\mu_1, \ldots, \mu_m) (x_{j_1} - \mu_{j_1})(x_{j_2} - \mu_{j_2}) \]
\[ + \cdots \]
\[ + \frac{1}{n!} \sum_{j_1=1}^{m} \cdots \sum_{j_n=1}^{m} \frac{\partial^n g}{\partial x_{j_1} \cdots \partial x_{j_n}} (\mu_1, \ldots, \mu_m) (x_{j_1} - \mu_{j_1}) \cdots (x_{j_n} - \mu_{j_n}) \]
(7)

What we have accomplished in (5) – (7) is the following. The Corollary tells us that if all the partial derivatives of order \((n+1)\) of \(g(x_1, \ldots, x_m)\) equal zero we may, without any loss of generality, make the transformation from our original function \(g\) to its \(n\)th order Taylor expansion \(P_n\). This Taylor expansion \(P_n\), if developed about the point \((\mu_1, \ldots, \mu_m)\), is a function of centralized random variables. So by taking expectation values of \(P_n\), or powers of \(P_n\), we find that the expectation value of \(g\), or powers of \(g\), can be written as a function of (known) centralized moments of the (known) random variables \((X_1 \cdots X_m) \sim p(x_1 \cdots x_m | \{\theta\})\).

Procedure (5) – (7) can be computationally very tedious, even in the most simplest of cases, as shall be demonstrated below when we find the mean, variance and skewness for the a sample mean of \(m\) known random variables. For more complicated functions of random variables one would probably like to use symbolic computing packages like Maple or Mathematica. In Van Erp and Van Gelder (2007), we will give such a routine for Mathematica.

**Example: Finding the Mean, Variance and Skewness of the Sample Mean**

One of the most well known functions of random variables is the sample mean:
\[
Z = g(x_1, \ldots, x_m) = \frac{1}{m} \sum_{j=1}^{m} x_j
\]  

(8)

Since all the second order partial derivatives of (8) equal zero we have, by our corollary, that

\[
Z = g(\mu_1, \ldots, \mu_m) + \sum_{j=1}^{m} \left[ \frac{\partial g}{\partial x_j} (\mu_1, \ldots, \mu_m) \right] (x_j - \mu_j)
\]

\[
= \frac{1}{m} \sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \frac{1}{m} (x_j - \mu_j)
\]

(9)

And it follows that

\[
Z^2 = \left\{ \frac{1}{m} \sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \frac{1}{m} (x_j - \mu_j) \right\}^2
\]

\[
= \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 + 2 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right) \sum_{j=1}^{m} \frac{1}{m} (x_j - \mu_j) + \sum_{j=1}^{m} \sum_{j=1}^{m} \frac{1}{m^2} (x_j - \mu_j) (x_j - \mu_j)
\]

(10)

As well as
\[ Z^3 = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \frac{1}{m} (X_j - \mu_j) \right)^3 \]

\[ = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^3 + 3 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 \sum_{j=1}^{m} \frac{1}{m} (X_j - \mu_j) \]

\[ + 3 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right) \sum_{j=1}^{m} \frac{1}{m} (X_j - \mu_j) (X_{j_2} - \mu_{j_2}) \]

\[ + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{j_3=1}^{m} \frac{1}{m^3} (X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})(X_{j_3} - \mu_{j_3}) \]

(11)

In what follows we will need the following properties of expectation values

\[ E(c) = c, \quad (12a) \]

\[ E(X_j - \mu_j) = 0, \quad (12b) \]

\[ E[(X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})] = \text{cov}(X_{j_1}, X_{j_2}), \quad (12c) \]

\[ E[(X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})(X_{j_3} - \mu_{j_3})] = M_3(X_{j_1}, X_{j_2}, X_{j_3}). \quad (12d) \]

where we let the symbol \( M_3 \) in (12d) stand for the third order mixed centralized moments. Using the results (12a) and (12b), we find the expectation value of (9) to be
\begin{equation}
E(Z) = E \left[ \frac{1}{m} \sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \frac{1}{m} (X_j - \mu_j) \right]
\end{equation}

\begin{equation}
= E \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right) + \sum_{j=1}^{m} \frac{1}{m} E(X_j - \mu_j)
\end{equation}

\begin{equation}
= \frac{1}{m} \sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \frac{1}{m} E(X_j - \mu_j)
\end{equation}

\begin{equation}
= \frac{1}{m} \sum_{j=1}^{m} \mu_j
\end{equation}

Using the results (12a), (12b) and (12c), we find the expectation value of (10) to be

\begin{equation}
Z^2 = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 + 2 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right) \sum_{j=1}^{m} \frac{1}{m} E(X_j - \mu_j) + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{1}{m^2} E[(X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})]
\end{equation}

\begin{equation}
= \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{1}{m^2} \text{cov}(X_{j_1}, X_{j_2})
\end{equation}

(14)

Using the results (12a), (12b), (12c) and (12d), we find the expectation value of (11) to be
\[ E(Z^3) = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^3 + 3 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 \sum_{j=1}^{m} \frac{1}{m} E(X_j - \mu_j) \]

\[ + 3 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right) \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{1}{m^2} E[(X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})] \]

\[ + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{j_3=1}^{m} \frac{1}{m^3} E[(X_{j_1} - \mu_{j_1})(X_{j_2} - \mu_{j_2})(X_{j_3} - \mu_{j_3})] \]

\[ = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^3 + 3 \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{1}{m^2} \text{cov}(X_{j_1}, X_{j_2}) + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{j_3=1}^{m} \frac{1}{m^3} M_3(X_{j_1}, X_{j_2}, X_{j_3}) \]

(15)

If we substitute (13) and (14) in the equation for the variance of \( Z \), then we find

\[ \text{var}(Z) = E(Z^2) - [E(Z)]^2 \]

\[ = \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 + \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{1}{m^2} \text{cov}(X_{j_1}, X_{j_2}) - \left( \frac{1}{m} \sum_{j=1}^{m} \mu_j \right)^2 \]

(16)

\[ = \frac{1}{m^2} \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \text{cov}(X_{j_1}, X_{j_2}) \]

Likewise if we substitute (13), (14) and (15) in the equation for the third centralized moment of \( Z \), then it can be verified that we find

\[ M_3(Z) = E(Z^3) - 3E(Z^2)E(Z) + 2[E(Z)]^3 \]

(17)

\[ = \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \sum_{j_3=1}^{m} \frac{1}{m^3} M_3(X_{j_1}, X_{j_2}, X_{j_3}) \]

We now define the skewness of \( Z \) to be its standardized third moment
\[
\text{skew}(Z) = \frac{M_3(Z)}{\text{var}(Z)^{3/2}}
\]  

(18)

If the \( X_i, \ i = 1, \ldots, m \), are independent and identically distributed with common distribution \( X_i \sim N(\mu, \sigma^2) \), \( i = 1, \ldots, m \), or equivalently, if

\[
E(X_i) = \mu, \quad \text{cov}(X_i, X_i) = \sigma^2, \quad M_3(X_i, X_i, X_i) = 0
\]  

(19)

for \( i = 1, \ldots, m \), then we find, using (9) – (18), the mean, variance and skewness of \( Z \) to be, respectively

\[
E(Z) = \mu, \quad \text{var}(Z) = \frac{\sigma^2}{m}, \quad \text{skew}(Z) = 0
\]  

(20)

But if the \( X_i, \ i = 1, \ldots, m \), are independent and identically distributed with common distribution \( X_i \sim \text{Exp}(\mu) \), \( i = 1, \ldots, m \), or equivalently, if

\[
E(X_i) = \mu, \quad \text{cov}(X_i, X_i) = \mu^2, \quad M_3(X_i, X_i, X_i) = 2\mu^3
\]  

(21)

for \( i = 1, \ldots, m \), then we find, using (9) – (18), the mean, variance and skewness of \( Z \) to be, respectively

\[
E(Z) = \mu, \quad \text{var}(Z) = \frac{\mu^2}{m}, \quad \text{skew}(Z) = \frac{2}{\sqrt{m}}
\]  

(22)

**Summary**

In this article, a general procedure to find the mean, variance and higher order centralized moments for arbitrary functions of random variables has been presented. A case study
and algorithm of the general procedure will be described in part II of this paper (Van Erp and Van Gelder, 2007).

**Acknowledgements**

This work was supported by the European Community's Sixth Framework Programme through the grant to the budget of the integrated Project FLOODsite, Contract GOCE-CT-2004-505420 and the Delft Cluster Project C71D28 ‘Safety against Flooding’, sub-workpackage WP-A3-3.

**References**
