



General moments of roundoff error

DOI:

[10.1080/03610918.2019.1615627](https://doi.org/10.1080/03610918.2019.1615627)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Wiegand, M., & Nadarajah, S. (2019). General moments of roundoff error. *Communications in Statistics: Simulation and Computation*. <https://doi.org/10.1080/03610918.2019.1615627>

Published in:

Communications in Statistics: Simulation and Computation

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



1.

General moments of roundoff error

Martin Wiegand and Saralees Nadarajah

School of Mathematics, University of Manchester, Manchester M13 9PL, UK

March 30, 2019

Abstract

Li and Nadarajah [Signal Processing 127 (2016) 185-190] derived expressions for mean and variance of roundoff error for any continuous random variable. Here, we derive expressions for general moment of the roundoff error, allowing one to study other aspects of roundoff error than just mean and variance. The expressions are specialised for ten commonly used distributions in signal processing. Numerical studies checking the correctness of the derived expressions are given.

Keywords— Continuous random variable; Distribution; Finite interval; Floor

1 Introduction

Round off errors arise in many areas of signal processing: wave digital-filters, fast state-space decimator and polynomial FIR predictors and predictive FIR differentiators; to mention a few. See Li and Nadarajah [7] for other areas and references.

Two measures of roundoff errors are their mean and variance. Gadzhiev [4] derived expressions for these when X is a centered uniform or a centered Gaussian random variable. Li and Nadarajah [7] extended Gadzhiev [4] for any continuous random variable X defined on either the real line or a finite interval. But often higher order moments are of interest, not just the mean and variance. Examples include skewness and kurtosis. Skewness can be used to know if roundoff errors are more likely to be positive than negative or if they are more likely to be negative than positive. Kurtosis can be used to know how concentrated roundoff errors are around zero.

The aim of this paper is to derive general moments of roundoff errors for any continuous random variable X . By rounding we mean here the largest integer that does not exceed the true value. The derived expressions for moments of roundoff errors are simple. Simple computer programs have been written by the authors that implement the derived expressions for any continuous random variable X . The programs can be obtained from the corresponding author.

Various distributions have been used to model roundoff errors in the signal processing area: uniform distribution in Press [8], Barnes et al. [2], Wong [11], Vladimirov and Diamond [9], Csordas et al. [3]; normal distribution in Ardalan and Alexander [1], Yu and Lim [12]; trapezoidal distribution in Kawarai and Murakami [6]; triangular distribution in Csordas et al. [3]; the sinusoidal distribution, the convolution of triangular and uniform distributions and the convolution of triangular and triangular distributions in Widrow and Kollar [10]. The derived formulas for the general moment can be used to provide basic measures of roundoff error for each of these distributions.

The contents of this paper are organized as follows: four theorems deriving expressions for general moments of $X - \text{floor}(X)$ and $X - \text{floor}(X + \frac{1}{2})$ are given in Section 2; specific forms of the general moments for some special distributions are derived in Section 3; a numerical study showing the use of the theorems and checking correctness of their derivations are given in Section 4.

2 Theoretical considerations

Theorems 2.1 and 2.2 derive the general moments of $X - \text{floor}(X)$ and $X - \text{floor}(X + \frac{1}{2})$ when X is a random variable on the real line. Theorems 2.3 and 2.4 derive the general moments of $X - \text{floor}(X)$ and $X - \text{floor}(X + \frac{1}{2})$ when X is a bounded random variable. The proofs of the theorems are given in the appendix.

We adapt the notation below: for a random variable X with probability density function f (PDF) and cumulative distribution function F (CDF),

$$M_t(x) = \int_{-\infty}^x z^t f(z) dz = x^t F(x) - t \int_{-\infty}^x z^{t-1} F(z) dz,$$

$$M_t^*(x) = \int_{-\infty}^{x-\frac{1}{2}} z^t f(z) dz = \left(x - \frac{1}{2}\right)^t F\left(x - \frac{1}{2}\right) - t \int_{-\infty}^{x-\frac{1}{2}} z^{t-1} F(z) dz.$$

Theorem 2.1. Let X be a continuous random variable on the domain $(-\infty, \infty)$ with PDF f and CDF F . Then for $k \in \mathbb{N}$,

$$\mathbb{E} \left[(X - \text{floor}(X))^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \sum_{j=-\infty}^{\infty} j^i [M_{k-i}(j+1) - M_{k-i}(j)].$$

Theorem 2.2. Let X be a continuous random variable on the domain $(-\infty, \infty)$ with PDF f and CDF F . Then for $k \in \mathbb{N}$,

$$\mathbb{E} \left[\left(X - \text{floor} \left(X + \frac{1}{2} \right) \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \sum_{j=-\infty}^{\infty} j^i [M_{k-i}^*(j+1) - M_{k-i}^*(j)].$$

Theorem 2.3. Let X be a continuous random variable on the domain (a, b) with $-\infty < a < b < \infty$ with PDF f and CDF F . Then for $k \in \mathbb{N}$,

$$\mathbb{E} \left[(X - \text{floor}(X))^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[(q-1)^i M_{k-i}(b) - \sum_{j=p+1}^{q-1} (j^i - (j-1)^i) M_{k-i}(j) \right]$$

with $p = \lfloor a \rfloor$ and $q = \lceil b \rceil$.

Theorem 2.4. Let X be a continuous random variable on the domain (a, b) with $-\infty < a < b < \infty$ with PDF f and CDF F . Then for $k \in \mathbb{N}$,

$$\mathbb{E} \left[\left(X - \text{floor} \left(X + \frac{1}{2} \right) \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[(q-1)^i M_{k-i}^*(b) - \sum_{j=p+1}^{q-1} (j^i - (j-1)^i) M_{k-i}^*(j) \right]$$

with $p = \lfloor a + \frac{1}{2} \rfloor$ and $q = \lceil b + \frac{1}{2} \rceil$.

3 Commonly used distributions in practice

We address the explicit forms of the previously derived general moment formulas for common distributions. We use the notation $k_1 = k - i + 1$ and $k_2 = k - i + 2$ as well as $c(i, j) = j^i - (j-1)^i$ throughout the following paragraphs.

3.1 Uniform distribution (Widrow and Kollar [10], I.7, page 679)

For a uniform random variable X with PDF and CDF specified by

$$f_X(x) = \frac{1}{b-a} \mathbb{I}_{\{x \in [a, b]\}} \quad F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } x \in [a, b] \\ 1 & \text{if } x > b \end{cases}$$

for $-\infty < a < b < \infty$, we have

$$\mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] = \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{k_1(b-a)} \left[q^i b^{k_1} - p^i a^{k_1} - \sum_{j=p+1}^q c(i, j) j^{k_1} \right]$$

for $k \in \mathbb{N}$, where $q = \lceil b \rceil$ and $p = \lfloor a \rfloor$. Furthermore,

$$\mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] = \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{k_1(b-a)} \left[q^i b^{k_1} - p^i a^{k_1} - \sum_{j=p+1}^q c(i, j) \left(j - \frac{1}{2} \right)^{k_1} \right]$$

for $k \in \mathbb{N}$, where $q = \lceil b + \frac{1}{2} \rceil$ and $p = \lfloor a + \frac{1}{2} \rfloor$.

3.2 Triangular distribution (Widrow and Kollar [10], I.8, page 680)

For a triangular random variable X with PDF and CDF specified by

$$f_X(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{2(x-a)}{(b-a)(c-a)} & \text{if } a < x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{if } c < x \leq b \\ 0 & \text{if } x > b \end{cases} \quad F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{(x-a)^2}{(b-a)(c-a)} & \text{if } x \in [a, b] \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)} & \text{if } x > b \\ 1 & \text{if } x \geq b \end{cases}$$

for $-\infty < a \leq c \leq b < \infty$, we have

$$\begin{aligned} \mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] &= \sum_{i=0}^k \frac{2(-1)^i \binom{k}{i}}{(b-a)(c-a)(b-c)} \left[q^i (c-a) b^{k_2} \left(\frac{1}{k_1} - \frac{1}{k_2} \right) - r^i (c-a) \left(b \frac{c^{k_1}}{k_1} - \frac{c^{k_2}}{k_2} \right) \right. \\ &\quad \left. + r^i (b-c) \left(\frac{c^{k_2}}{k_2} - a \frac{c^{k_1}}{k_1} \right) \right. \\ &\quad \left. - p^i (b-c) a^{k_2} \left(\frac{1}{k_2} - \frac{1}{k_1} \right) - \sum_{j=r+1}^q c(i, j) (c-a) \left(b \frac{j^{k_1}}{k_1} - \frac{j^{k_2}}{k_2} \right) \right. \\ &\quad \left. - \sum_{j=p+1}^r c(i, j) (b-c) \left(\frac{j^{k_2}}{k_2} - a \frac{j^{k_1}}{k_1} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil b \rceil$, $r = \lfloor c \rfloor$ and $p = \lfloor a \rfloor$. Furthermore,

$$\begin{aligned} \mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] &= \sum_{i=0}^k \frac{2(-1)^i \binom{k}{i}}{(b-a)(c-a)(b-c)} \left[q^i (c-a) b^{k_2} \left(\frac{1}{k_1} - \frac{1}{k_2} \right) - r^i (c-a) \left(b \frac{c^{k_1}}{k_1} - \frac{c^{k_2}}{k_2} \right) \right. \\ &\quad \left. + r^i (b-c) \left(\frac{c^{k_2}}{k_2} - a \frac{c^{k_1}}{k_1} \right) \right. \\ &\quad \left. - p^i (b-c) a^{k_2} \left(\frac{1}{k_2} - \frac{1}{k_1} \right) - \sum_{j=r+1}^q c(i, j) (c-a) \left(b \frac{(j-\frac{1}{2})^{k_1}}{k_1} - \frac{(j-\frac{1}{2})^{k_2}}{k_2} \right) \right. \\ &\quad \left. - \sum_{j=p+1}^r c(i, j) (b-c) \left(\frac{(j-\frac{1}{2})^{k_2}}{k_2} - a \frac{(j-\frac{1}{2})^{k_1}}{k_1} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil b + \frac{1}{2} \rceil$, $r = \lfloor c + \frac{1}{2} \rfloor$ and $p = \lfloor a + \frac{1}{2} \rfloor$.

3.3 Trapezoidal distribution (Hou et al. [5])

For a trapezoidal random variable X with PDF and CDF specified by

$$f_X(x) = \frac{2}{d+b-a-c} \begin{cases} 0 & \text{if } x < -a \\ \frac{x-a}{c-a} & \text{if } a \leq x < c \\ 1 & \text{if } c \leq x < d \\ \frac{b-x}{b-d} & \text{if } d \leq x < b \\ 0 & \text{if } x \geq a \end{cases} \quad F_X(x) = \frac{2}{d+b-a-c} \begin{cases} 0 & \text{if } x < -a \\ \frac{x(\frac{x}{2}-a) + \frac{a}{2}}{c-a} & \text{if } a \leq x < c \\ (x-c) + \frac{c-a}{2} & \text{if } c \leq x < d \\ \frac{x(b-\frac{x}{2}) - d(b-\frac{d}{2})}{b-d} + (d-c) + \frac{c-a}{2} & \text{if } d \leq x < b \\ 1 & \text{if } x \geq b \end{cases}$$

for $-\infty < a < c \leq d < b < \infty$, we have

$$\begin{aligned} \mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] &= \sum_{i=0}^k \frac{2(-1)^i \binom{k}{i}}{b+d-a-c} \left[\frac{q^i}{b-d} \left(\frac{b^{k_2}}{k_1} - \frac{b^{k_2}}{k_2} \right) - \frac{s^i}{b-d} \left(\frac{d^{k_2}}{k_2} - b \frac{d^{k_1}}{k_1} \right) + s^i \frac{d^{k_1}}{k_1} - r^i \frac{c^{k_1}}{k_1} \right. \\ &\quad + \frac{r^i}{c-a} \left(\frac{c^{k_2}}{k_2} - a \frac{c^{k_1}}{k_1} \right) \\ &\quad - \frac{p^i}{c-a} \left(\frac{a^{k_2}}{k_2} - \frac{a^{k_2}}{k_1} \right) - \sum_{j=p+1}^r \frac{c(i,j)j^{k_1}}{c-a} \left(\frac{j}{k_2} - \frac{a}{k_1} \right) - \sum_{j=r+1}^s c(i,j) \frac{j^{k_1}}{k_1} \\ &\quad \left. - \sum_{j=s+1}^q \frac{c(i,j)j^{k_1}}{b-a} \left(\frac{b}{k_1} - \frac{j}{k_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil b \rceil$, $r = \lfloor c \rfloor$, $s = \lfloor d \rfloor$ and $p = \lfloor a \rfloor$. Furthermore,

$$\begin{aligned} \mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] &= \sum_{i=0}^k \frac{2(-1)^i \binom{k}{i}}{b+d-a-c} \left[\frac{q^i}{b-d} \left(\frac{b^{k_2}}{k_1} - \frac{b^{k_2}}{k_2} \right) - \frac{s^i}{b-d} \left(\frac{d^{k_2}}{k_2} - b \frac{d^{k_1}}{k_1} \right) + s^i \frac{d^{k_1}}{k_1} - r^i \frac{c^{k_1}}{k_1} \right. \\ &\quad + \frac{r^i}{c-a} \left(\frac{c^{k_2}}{k_2} - a \frac{c^{k_1}}{k_1} \right) \\ &\quad - \frac{p^i}{c-a} \left(\frac{a^{k_2}}{k_2} - \frac{a^{k_2}}{k_1} \right) - \sum_{j=p+1}^r \frac{c(i,j) \left(j - \frac{1}{2} \right)^{k_1}}{c-a} \left(\frac{j - \frac{1}{2}}{k_2} - \frac{a}{k_1} \right) - \sum_{j=r+1}^s c(i,j) \frac{\left(j - \frac{1}{2} \right)^{k_1}}{k_1} \\ &\quad \left. - \sum_{j=s+1}^q \frac{c(i,j)j^{k_1}}{b-a} \left(\frac{b}{k_1} - \frac{j - \frac{1}{2}}{k_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil b + \frac{1}{2} \rceil$, $r = \lfloor c + \frac{1}{2} \rfloor$, $s = \lfloor d + \frac{1}{2} \rfloor$ and $p = \lfloor a + \frac{1}{2} \rfloor$.

3.4 House distribution ([10], 3.9, page 55)

For a house random variable X with PDF and CDF specified by

$$f_X(x) = \begin{cases} 0 & \text{if } x < -A \\ (1+\alpha)B + \frac{\alpha Bx}{A} & \text{if } -A \leq x \leq 0 \\ (1+\alpha)B - \frac{\alpha Bx}{A} & \text{if } 0 < x \leq A \\ 0 & \text{for } x > A \end{cases} \quad F_X(x) = \begin{cases} 0 & x < -A \\ (1+\alpha)B(A+x) + \frac{\alpha B(x^2 - A^2)}{2A} & -A \leq x \leq 0 \\ (1+\alpha)B(x) - \frac{\alpha Bx^2}{2A} + \frac{1}{2} & 0 < x \leq A \\ 1 & x > A \end{cases}$$

for $A > 0$ and $\alpha > 0$, where $B = \frac{1}{2A + A\alpha}$, we have

$$\begin{aligned} \mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] &= \sum_{i=0}^k (-1)^i \binom{k}{i} B \left[q^i \frac{A^{k_1}}{k_1} (1+\alpha) - q^i \frac{A^{k_1}}{k_2} \alpha - p^i \frac{(-A)^{k_1}}{k_1} (1+\alpha) - p^i \frac{\alpha (-A)^{k_2}}{A k_2} \right. \\ &\quad \left. - \sum_{j=p+1}^q c(i,j) \frac{j^{k_1}}{k_1} (1+\alpha) - \frac{\alpha}{Ak_2} \left(\sum_{j=1}^q c(i,j)j^{k_2} - \sum_{j=p+1}^0 c(i,j)j^{k_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil A \rceil$ and $p = \lfloor -A \rfloor$. Furthermore,

$$\begin{aligned} \mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] &= \sum_{i=0}^k (-1)^i \binom{k}{i} B \left[q^i \frac{A^{k_1}}{k_1} (1+\alpha) - q^i \frac{A^{k_1}}{k_2} \alpha - p^i \frac{(-A)^{k_1}}{k_1} (1+\alpha) - p^i \frac{\alpha (-A)^{k_2}}{A k_2} \right. \\ &\quad \left. - \sum_{j=p+1}^q c(i,j) \frac{\left(j - \frac{1}{2} \right)^{k_1}}{k_1} (1+\alpha) - \frac{\alpha}{Ak_2} \left(\sum_{j=1}^q c(i,j) \left(j - \frac{1}{2} \right)^{k_2} - \sum_{j=p+1}^0 c(i,j) \left(j - \frac{1}{2} \right)^{k_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil A + \frac{1}{2} \rceil$ and $p = \lfloor -A + \frac{1}{2} \rfloor$.

3.5 Curved trapezoidal distribution (Widrow and Kollar [10], 7.11, page 168; Kawarai and Murakami [6], page 884)

For a curved trapezoidal random variable X with PDF and CDF specified by

$$f_X(x) = \begin{cases} a \left(\frac{\rho}{x} - 1 \right) & \text{if } -\rho \leq x < -\frac{\rho}{2} \\ a & \text{if } -\frac{\rho}{2} \leq x < \frac{\rho}{2} \\ a \left(\frac{\rho}{x} - 1 \right) & \text{if } \frac{\rho}{2} \leq x \leq \rho \\ 0 & \text{else} \end{cases} \quad F_X(x) = \begin{cases} -a \left[x + \rho + \rho \log \left(\frac{-x}{\rho} \right) \right] & \text{if } -\rho \leq x < -\frac{\rho}{2} \\ ax + a\rho \log(2) & \text{if } -\frac{\rho}{2} \leq x < \frac{\rho}{2} \\ a \left[(\rho - x) + \log \left(\frac{x}{\rho} \right) + \rho \log(4) \right] & \text{if } \frac{\rho}{2} \leq x \leq \rho \\ 1 & x > \rho \end{cases}$$

for $\rho > 0$ and $a = \frac{1}{\rho \log(4)}$, we have

$$\begin{aligned} \mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] &= \sum_{i=0}^{k-1} (-1)^i \binom{k}{i} a \left[q^i \frac{\rho^{k_1}}{k-i} - q^i \frac{\rho^{k_1}}{k_1} - d^i \rho \frac{\left(\frac{\rho}{2}\right)^{k-i}}{k-i} + d^i \frac{\left(\frac{\rho}{2}\right)^{k_1}}{k_1} + d^i \frac{\left(\frac{\rho}{2}\right)^{k_1}}{k_1} - c^i \frac{\left(\frac{-\rho}{2}\right)^{k_1}}{k_1} \right. \\ &\quad \left. - c^i \rho \frac{\left(\frac{-\rho}{2}\right)^{k-i}}{k-i} - c^i \frac{\left(\frac{-\rho}{2}\right)^{k_1}}{k_1} + p^i \rho \frac{(-\rho)^{k-i}}{k-i} + p^i \frac{(-\rho)^{k_1}}{k_1} \right. \\ &\quad \left. - \sum_{j=d+1}^q c(i, j) \left(\rho \frac{j^{k-i}}{k-i} - \frac{j^{k_1}}{k_1} \right) - \sum_{j=c+1}^d c(i, j) \frac{j^{k_1}}{k_1} + \sum_{j=p+1}^c c(i, j) \left(\rho \frac{j^{k-i}}{k-i} + \frac{j^{k_1}}{k_1} \right) \right] \\ &\quad + a \left[q^k - \sum_{j=d+1}^q c(k, j) \left((\rho - j) + \rho \log(4) + \log \left(\frac{j}{\rho} \right) \right) - \sum_{j=c+1}^d c(k, j) (j + \rho \log(2)) \right. \\ &\quad \left. + \sum_{j=p+1}^c c(k, j) \left(j + \rho \log \left(\frac{-j}{\rho} \right) \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil \rho \rceil, c = \lfloor -\rho/2 \rfloor, d = \lfloor \rho/2 \rfloor$ and $p = \lfloor -\rho \rfloor$. Furthermore,

$$\begin{aligned} \mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] &= \sum_{i=0}^{k-1} (-1)^i \binom{k}{i} a \left[q^i \frac{\rho^{k_1}}{k-i} - q^i \frac{\rho^{k_1}}{k_1} - d^i \rho \frac{\left(\frac{\rho}{2}\right)^{k-i}}{k-i} + d^i \frac{\left(\frac{\rho}{2}\right)^{k_1}}{k_1} + d^i \frac{\left(\frac{\rho}{2}\right)^{k_1}}{k_1} - c^i \frac{\left(\frac{-\rho}{2}\right)^{k_1}}{k_1} \right. \\ &\quad \left. - c^i \rho \frac{\left(\frac{-\rho}{2}\right)^{k-i}}{k-i} - c^i \frac{\left(\frac{-\rho}{2}\right)^{k_1}}{k_1} + p^i \rho \frac{(-\rho)^{k-i}}{k-i} + p^i \frac{(-\rho)^{k_1}}{k_1} \right. \\ &\quad \left. - \sum_{j=d+1}^q c(i, j) \left(\rho \frac{\left(j - \frac{1}{2}\right)^{k-i}}{k-i} - \frac{\left(j - \frac{1}{2}\right)^{k_1}}{k_1} \right) - \sum_{j=c+1}^d c(i, j) \frac{\left(j - \frac{1}{2}\right)^{k_1}}{k_1} \right. \\ &\quad \left. + \sum_{j=p+1}^c c(i, j) \left(\rho \frac{\left(j - \frac{1}{2}\right)^{k-i}}{k-i} + \frac{\left(j - \frac{1}{2}\right)^{k_1}}{k_1} \right) \right] \\ &\quad + a \left[q^k - \sum_{j=d+1}^q c(k, j) \left(\rho - \left(j - \frac{1}{2}\right) + \rho \log(4) + \log \left(\frac{\left(j - \frac{1}{2}\right)}{\rho} \right) \right) \right. \\ &\quad \left. - \sum_{j=c+1}^d c(k, j) \left(\left(j - \frac{1}{2}\right) + \rho \log(2) \right) + \sum_{j=p+1}^c c(k, j) \left(\left(j - \frac{1}{2}\right) + \rho \log \left(\frac{-\left(j - \frac{1}{2}\right)}{\rho} \right) \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $q = \lceil \rho + \frac{1}{2} \rceil, c = \lfloor -\frac{\rho}{2} + \frac{1}{2} \rfloor, d = \lfloor \frac{\rho}{2} + \frac{1}{2} \rfloor$ and $p = \lfloor \rho + \frac{1}{2} \rfloor$.

3.6 Hexagonal distribution (Widrow and Kollar [10], E3.13.1, page 56)

For a curved trapezoidal random variable X with PDF and CDF specified by

$$f_X(x) = \begin{cases} 0 & \text{if } x < -3a \\ \frac{x+3a}{6a^2} & \text{if } -3a \leq x < -2a \\ \frac{1}{6a} & \text{if } -2a \leq x < -a \\ \frac{1}{6a} + \frac{x+a}{6a^2} & \text{if } -a \leq x < 0 \\ \frac{1}{6a} + \frac{-x+a}{6a^2} & \text{if } 0 \leq x < a \\ \frac{1}{6a} & \text{if } a \leq x < 2a \\ \frac{-x+3a}{6a^2} & \text{if } 2a \leq x \leq 3a \\ 0 & \text{if } x > 3a \end{cases} \quad F_X(x) = \begin{cases} 0 & \text{if } x < -3a \\ \frac{x^2}{12a^2} + \frac{x}{2a} + \frac{3}{4} & \text{if } -3a \leq x < -2a \\ \frac{x}{6a} + \frac{5}{12} & \text{if } -2a \leq x < -a \\ \frac{x^2}{12a^2} + \frac{x}{3a} + \frac{1}{2} & \text{if } -a \leq x < 0 \\ \frac{x}{3a} - \frac{x^2}{12a^2} + \frac{1}{2} & \text{if } 0 \leq x < a \\ \frac{x}{6a} + \frac{7}{12} & \text{if } a \leq x < 2a \\ \frac{-x^2}{12a^2} + \frac{x}{2a} + \frac{1}{4} & \text{if } 2a \leq x \leq 3a \\ 1 & \text{if } x > 3a \end{cases}$$

for $a > 0$ and $a = \frac{1}{\rho \log(4)}$, we have

$$\begin{aligned} \mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] &= \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{6a} \left[q^i \left(\frac{3(3a)^{k_1}}{k_1} - \frac{(3a)^{k_2}}{ak_2} \right) - g^i \left(\frac{3(2a)^{k_1}}{k_1} - \frac{(2a)^{k_2}}{ak_2} \right) + g^i \frac{(2a)^{k_1}}{k_1} + f^i \frac{a^{k_1}}{k_1} - f^i \frac{a^{k_2}}{ak_2} \right. \\ &\quad - d^i \frac{(-a)^{k_1}}{k_1} - d^i \frac{(-a)^{k_2}}{ak_2} - c^i \frac{(-2a)^{k_1}}{k_1} + c^i \left(\frac{(-2a)^{k_2}}{ak_2} + \frac{3(-2a)^{k_1}}{k_1} \right) - p^i \left(\frac{(-3a)^{k_2}}{ak_2} + \frac{3(-3a)^{k_1}}{k_1} \right) \\ &\quad - \sum_{j=p+1}^c c(i, j) \left(\frac{j^{k_2}}{ak_2} + \frac{3j^{k_1}}{k_1} \right) - \sum_{j=c+1}^d c(i, j) \frac{j^{k_1}}{k_1} - \sum_{j=d+1}^0 c(i, j) \left(\frac{2j^{k_1}}{k_1} + \frac{j^{k_2}}{ak_2} \right) \\ &\quad \left. - \sum_{j=1}^f c(i, j) \left(\frac{2j^{k_1}}{k_1} - \frac{j^{k_2}}{ak_2} \right) - \sum_{j=f+1}^g c(i, j) \frac{j^{k_1}}{k_1} - \sum_{j=g+1}^q c(i, j) \left(\frac{3j^{k_1}}{k_1} - \frac{j^{k_2}}{ak_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $p = \lfloor -3a \rfloor$, $c = \lfloor -2a \rfloor$, $d = \lfloor -a \rfloor$, $f = \lfloor a \rfloor$ and $g = \lfloor 2a \rfloor$. Furthermore,

$$\begin{aligned} \mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] &= \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{6a} \left[q^i \left(\frac{3(3a)^{k_1}}{k_1} - \frac{(3a)^{k_2}}{ak_2} \right) - g^i \left(\frac{3(2a)^{k_1}}{k_1} - \frac{(2a)^{k_2}}{ak_2} \right) \right. \\ &\quad + g^i \frac{(2a)^{k_1}}{k_1} + f^i \frac{a^{k_1}}{k_1} - f^i \frac{a^{k_2}}{ak_2} \\ &\quad - d^i \frac{(-a)^{k_1}}{k_1} - d^i \frac{(-a)^{k_2}}{ak_2} - c^i \frac{(-2a)^{k_1}}{k_1} + c^i \left(\frac{(-2a)^{k_2}}{ak_2} + \frac{3(-2a)^{k_1}}{k_1} \right) \\ &\quad - p^i \left(\frac{(-3a)^{k_2}}{ak_2} + \frac{3(-3a)^{k_1}}{k_1} \right) \\ &\quad - \sum_{j=p+1}^c c(i, j) \left(\frac{(j - \frac{1}{2})^{k_2}}{ak_2} + \frac{3(j - \frac{1}{2})^{k_1}}{k_1} \right) - \sum_{j=c+1}^d c(i, j) \frac{(j - \frac{1}{2})^{k_1}}{k_1} \\ &\quad - \sum_{j=d+1}^0 c(i, j) \left(\frac{2(j - \frac{1}{2})^{k_1}}{k_1} + \frac{(j - \frac{1}{2})^{k_2}}{ak_2} \right) - \sum_{j=1}^f c(i, j) \left(\frac{2(j - \frac{1}{2})^{k_1}}{k_1} - \frac{(j - \frac{1}{2})^{k_2}}{ak_2} \right) \\ &\quad \left. - \sum_{j=f+1}^g c(i, j) \frac{(j - \frac{1}{2})^{k_1}}{k_1} - \sum_{j=g+1}^q c(i, j) \left(\frac{3(j - \frac{1}{2})^{k_1}}{k_1} - \frac{(j - \frac{1}{2})^{k_2}}{ak_2} \right) \right] \end{aligned}$$

for $k \in \mathbb{N}$, where $p = \lfloor -3a + \frac{1}{2} \rfloor$, $c = \lfloor -2a + \frac{1}{2} \rfloor$, $d = \lfloor -a + \frac{1}{2} \rfloor$, $f = \lfloor a + \frac{1}{2} \rfloor$, $g = \lfloor 2a + \frac{1}{2} \rfloor$ and $q = \lfloor 3a + \frac{1}{2} \rfloor$.

3.7 Sinusoidal distribution (Widrow and Kollar [10], I.6, page 677)

For a sinusoidal random variable X with PDF and CDF specified by

$$f_X(x) = \begin{cases} \frac{1}{\pi\sqrt{A^2 - x^2}} & \text{if } -a < x < a \\ 0 & \text{else} \end{cases} \quad F_X(x) = \begin{cases} 0 & \text{if } x \leq -a \\ \frac{1}{2} + \frac{\arctan\left(\frac{x}{\sqrt{a^2 - x^2}}\right)}{\pi} & \text{if } -a < x < a \\ 1 & \text{if } x \geq a \end{cases}$$

for $a > 0$, we have

$$\mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[\frac{\Gamma\left(\frac{3+(k-i)}{2}\right)}{(1+k-i)\sqrt{\pi}\Gamma\left(1+\frac{k-i}{2}\right)} (q^i a^{k-i} + p^i (-a)^{k-i}) \right. \\ \left. - \sum_{j=p+1}^q c(i, j) \frac{j^{1+k-i} \sqrt{1 - \frac{j^2}{a^2}}}{(1+k-i)\pi\sqrt{a^2 - j^2}} {}_2F_1\left(\frac{1}{2}, \frac{1+k-i}{2}, \frac{3+k-i}{2}, \frac{j^2}{a^2}\right) \right]$$

for $k \in \mathbb{N}$, where $p = \lfloor -a \rfloor$ and $q = \lceil a \rceil$. Furthermore,

$$\mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[\frac{\Gamma\left(\frac{3+(k-i)}{2}\right)}{(1+k-i)\sqrt{\pi}\Gamma\left(1+\frac{k-i}{2}\right)} (q^i a^{k-i} + p^i (-a)^{k-i}) \right. \\ \left. - \sum_{j=p+1}^q c(i, j) \frac{\left(j - \frac{1}{2}\right)^{1+k-i} \sqrt{1 - \frac{\left(j - \frac{1}{2}\right)^2}{a^2}}}{(1+k-i)\pi\sqrt{a^2 - \left(j - \frac{1}{2}\right)^2}} {}_2F_1\left(\frac{1}{2}, \frac{1+k-i}{2}, \frac{3+k-i}{2}, \frac{\left(j - \frac{1}{2}\right)^2}{a^2}\right) \right]$$

for $k \in \mathbb{N}$, where $p = \lfloor -a + \frac{1}{2} \rfloor$, $q = \lceil a + \frac{1}{2} \rceil$ and ${}_2F_1$ denotes the Gauss hypergeometric function.

3.8 Convolutions of uniform and triangular distributions (Widrow and Kollar [10], 19.6.3, page 501)

For X a random variable obtained by convoluting uniform and triangular random variables with PDF specified by

$$f_X(x) = \frac{1}{(b-a)(e-d)} \begin{cases} \frac{b-c}{(e-d)\frac{-2b+d+e+2x}{(b-c)}} & e > b-c+d \wedge b-e < x < c-d \\ \frac{(c+e+x-2b)(c-e-x)}{(b-c)} & (e=b-c+d \wedge x=b-e) \vee (d < e < b-c+d \wedge c-d < x \leq b-e) \\ \frac{(d+x-b)^2}{(b-c)} & \begin{cases} (e=b-c+d \wedge c-e < x < b-e) \vee (e > b-c+d \wedge c-e < x \leq b-e) \vee \\ (d < e < b-c+d \wedge c-e < x \leq c-d) \end{cases} \\ 0 & \text{else} \end{cases}$$

$$+ \frac{1}{(b-a)(e-d)} \begin{cases} \frac{c-a}{(e-d)\frac{d+e+2x-2a}{(a-c)}} & a > 0 \wedge e > -a+c+d \wedge c-e < x < a-d \\ \frac{-(-a+e+x)^2}{(a-c)} & \begin{cases} (a > 0 \wedge e = -a+c+d \wedge a-e < x < c-e) \vee (a > 0 \wedge d < e < -a+c+d \wedge a-d < x \leq c-e) \\ (a > 0 \wedge e = -a+c+d \wedge a-e < x < c-e) \vee (a > 0 \wedge e > -a+c+d \wedge a-e < x \leq c-e) \vee \\ (a > 0 \wedge d < e < -a+c+d \wedge a-e < x \leq a-d) \end{cases} \\ \frac{(2a-c-d-x)(c-d-x)}{(a-c)} & \begin{cases} (a > 0 \wedge e = -a+c+d \wedge c-e < x < c-d) \vee (a > 0 \wedge e > -a+c+d \wedge a-d \leq x < c-d) \vee \\ (a > 0 \wedge d < e < -a+c+d \wedge c-e < x < c-d) \end{cases} \\ 0 & \text{else} \end{cases}$$

we have

$$\mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[q^i \int_{-\infty}^{\infty} f_X(x) - \sum_{j=p+1}^q c(i, j) \int_{-\infty}^j x^{k-i} f_X(x) dx \right],$$

where $p = \lfloor \inf_x (f_X(x) \neq 0) \rfloor$ and $q = \lceil \sup_x (f_X(x) \neq 0) \rceil$. Furthermore,

$$\mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \left[q^i \int_{-\infty}^{\infty} f_X(x) - \sum_{j=p+1}^q c(i, j) \int_{-\infty}^{j-\frac{1}{2}} x^{k-i} f_X(x) dx \right],$$

where $p = \lfloor \inf_x (f_X(x) \neq 0) + \frac{1}{2} \rfloor$ and $q = \lceil \sup_x (f_X(x) \neq 0) + \frac{1}{2} \rceil$.

3.9 Convolutions of two triangular distributions (Widrow and Kollar [10], 19.6.4, page 502)

For X a random variable obtained by convoluting two triangular random variables with PDF specified by

$$\begin{aligned}
 f_X(x) = & \frac{2}{3(b-a)(e-d)} \left\{ \begin{array}{l} -\frac{(e-f)(-3a+e+2f+3x)}{(a-c)} \\ \frac{(a-c)(a+2c-3(e+x))}{(e-f)} \\ \frac{(a-e-x)^3}{(a-c)(e-f)} \\ -\frac{(c-f-x)(-2c^2+(3e-2f+x)c+3a(c-2e+f-x)+(3e-2f+x)(f+x))}{(a-b)(a-c)(d-e)(e-f)} \\ 0 \end{array} \right. & \left\{ \begin{array}{l} (a > 0 \wedge f = a-c+e \wedge x = a-f) \vee \\ (a > 0 \wedge a-c+e < f < e \wedge a-f \leq x \leq c) \\ (a > 0 \wedge f < a-c+e \wedge x = a-f) \vee \\ (a > 0 \wedge f < a-c+e \wedge c-e < x < a-f) \\ (a > 0 \wedge f = a-c+e \wedge a-e < x < a-f) \\ (a > 0 \wedge a-c+e < f < e \wedge a-e < x < a-f) \\ (c > a \wedge f < a-c+e \wedge a-e < x \leq c-e) \\ (a > 0 \wedge f = a-c+e \wedge a-f < x < c-f) \\ (a > 0 \wedge a-c+e < f < e \wedge c-e < x < c-f) \\ (a > 0 \wedge f < a-c+e \wedge a-f < x < c-f) \\ \text{else} \end{array} \right. \\
 + \frac{2}{3(b-a)(e-d)} & \left\{ \begin{array}{l} \frac{(a-c)(a+2c-3(d+x))}{(d-f)} \\ \frac{(d-f)(-3a+d+2f+3x)}{(a-c)} \\ \frac{(a-3d+2f-x)(-a+f+x)^2}{(a-c)(d-f)} \\ -\frac{(3a-2c-d-x)(-c+d+x)^2}{(a-c)(d-f)} \\ 0 \end{array} \right. & \left\{ \begin{array}{l} a > 0 \wedge f > -a+c+d \wedge c-f < x < a-d \\ (a > 0 \wedge f = -a+c+d \wedge x = c-f) \vee \\ (a > 0 \wedge d < f < -a+c+d \wedge a-d < x \leq c-f) \\ (a > 0 \wedge f = -a+c+d \wedge a-f < x < c-f) \vee \\ (a > 0 \wedge f > -a+c+d \wedge a-f < x \leq c-f) \vee \\ (a > 0 \wedge d < f < -a+c+d \wedge a-f < x \leq a-d) \\ (a > 0 \wedge f = -a+c+d \wedge c-f < x < c-d) \vee \\ (a > 0 \wedge f > -a+c+d \wedge a-d \leq x < c-d) \vee \\ (a > 0 \wedge d < f < -a+c+d \wedge c-f < x < c-d) \\ \text{else} \end{array} \right. \\
 + \frac{2}{3(b-a)(e-d)} & \left\{ \begin{array}{l} -\frac{(b-c)(b+2c-3(d+x))}{(d-f)} \\ \frac{(d-f)(-3b+d+2f+3x)}{(b-c)} \\ \frac{(c-f-x)(-2c^2+(3d-2f+x)c+3b(c-2d+f-x)+(3d-2f+x)(f+x))}{(b-c)(d-f)} \\ -\frac{(b-d-x)^3}{(b-c)(d-f)} \\ 0 \end{array} \right. & \left\{ \begin{array}{l} f > b-c+d \wedge b-f < x < c-d \\ (f = b-c+d \wedge x = b-f) \vee \\ (d < f < b-c+d \wedge c-d < x \leq b-f) \\ (f = b-c+d \wedge c-f < x < b-f) \vee \\ (f > b-c+d \wedge c-f < x \leq b-f) \vee \\ (d < f < b-c+d \wedge c-f < x \leq c-d) \\ (f = b-c+d \wedge b-f < x < b-d) \vee \\ (f > b-c+d \wedge x = c-d) \vee \\ (f > b-c+d \wedge c-d < x < b-d) \vee \\ (d < f < b-c+d \wedge b-f < x < b-d) \\ \text{else} \end{array} \right. \\
 + \frac{2}{3(b-a)(e-d)} & \left\{ \begin{array}{l} -\frac{(b-c)(b+2c-3(e+x))}{(e-f)} \\ -\frac{2(e-f)(-3b+e+2f+3x)}{(b-c)} \\ \frac{(3b-2c-e-x)(-c+e+x)^2}{(b-c)(e-f)} \\ -\frac{2(b-3e+2f-x)(-b+f+x)^2}{3(a-b)(b-c)(d-e)(e-f)} \\ 0 \end{array} \right. & \left\{ \begin{array}{l} f < -b+c+e \wedge b-e < x \leq c-f \\ (f = -b+c+e \wedge x = c-f) \vee \\ (-b+c+e < f < e \wedge c-f \leq x \leq b-e) \\ (f = -b+c+e \wedge c-e < x < c-f) \vee \\ (-b+c+e < f < e \wedge c-e < x < c-f) \vee \\ (f < -b+c+e \wedge c-e < x \leq b-e) \\ (f = -b+c+e \wedge c-f < x < b-f) \vee \\ (-b+c+e < f < e \wedge b-e < x < b-f) \vee \\ (f < -b+c+e \wedge c-f < x < b-f) \\ \text{else} \end{array} \right.
 \end{aligned}$$

for $-\infty < d \leq f \leq e < \infty$ and $-\infty < a \leq c \leq b < \infty$, the moments are the same as those in Section 3.8.

3.10 Normal distribution ([10], I.1, page 633)

For a normal random variable X with PDF and CDF specified by

$$f_X(x) = \frac{1}{\sqrt{2\pi}a} \exp\left(-\frac{x^2}{2a^2}\right) \quad F_X(x) = \Phi\left(\frac{x}{a}\right)$$

for $-\infty < x < \infty$ and $a > 0$, we have

$$\mathbb{E} \left[(X - \lfloor X \rfloor)^k \right] = \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{2\sqrt{2\pi}a} \sum_{j=-\infty}^{\infty} \left[j^i (j+1)^{k_1} E_{k_1/2} \left(\frac{(j+1)^2}{2a^2} \right) - j^{k+1} E_{k_1/2} \left(\frac{j^2}{2a^2} \right) \right]$$

and

$$\mathbb{E} \left[\left(X - \left\lfloor X + \frac{1}{2} \right\rfloor \right)^k \right] = \sum_{i=0}^k \frac{(-1)^i \binom{k}{i}}{2\sqrt{2\pi}a} \sum_{j=-\infty}^{\infty} \left[j^i (j+1)^{k_1} E_{k_1/2} \left(\frac{(j+\frac{1}{2})^2}{2a^2} \right) - j^{k+1} E_{k_1/2} \left(\frac{(j-\frac{1}{2})^2}{2a^2} \right) \right]$$

for $k \in \mathbb{N}$, where E_ν denotes the parabolic cylinder function with index ν .

4 Numerical results

4.1 Scatterplots

In this section, we present the results of our numerical computations. The plots show the first four moments, computed by the previously derived formulas and via the numerical moments, given by the estimator $\sum_{i=1}^n X_i^k$, $k = 1, \dots, 4$ for a randomly generated sample of size $n = 10000$.

For the sample distributions we have decided upon centred versions of the uniform, triangular and normal distributions. Therefore, we have X a uniform random variable with $b = -a$, or X a triangular random variable with $b = -a$, $c = 0$ or X a normal random variable with zero mean and standard deviation a . We let the parameter a increase by increments of 0.1 from 1 to 10 ($a = 0.1, \dots, 10$), leaving us with 100 plotted points for each moment. The tight grouping and overall shape of the plots gives us reason to believe in the correctness of our previous results.

The experiment was repeated for the shifted error moments of $X - \lfloor X + \frac{1}{2} \rfloor$, with the same distributions and sample size.

4.1.1 Plots for the moments of $X - \text{floor}(X)$

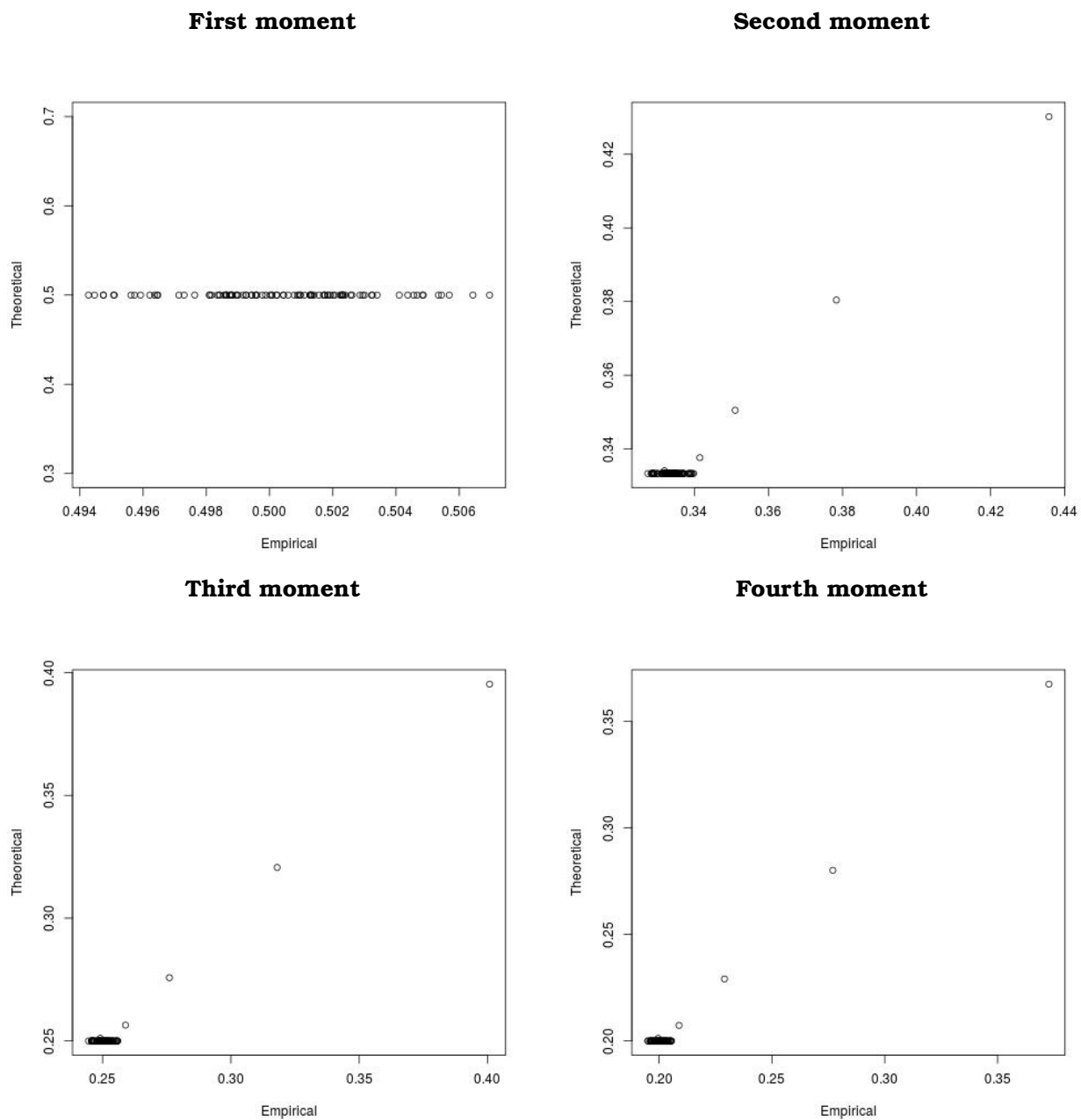


Figure 1: First four moments of $X - \text{floor}(X)$ for the normal distribution, with parameters $\mu = 0$ and standard deviation $a = 0.1, \dots, 10$.

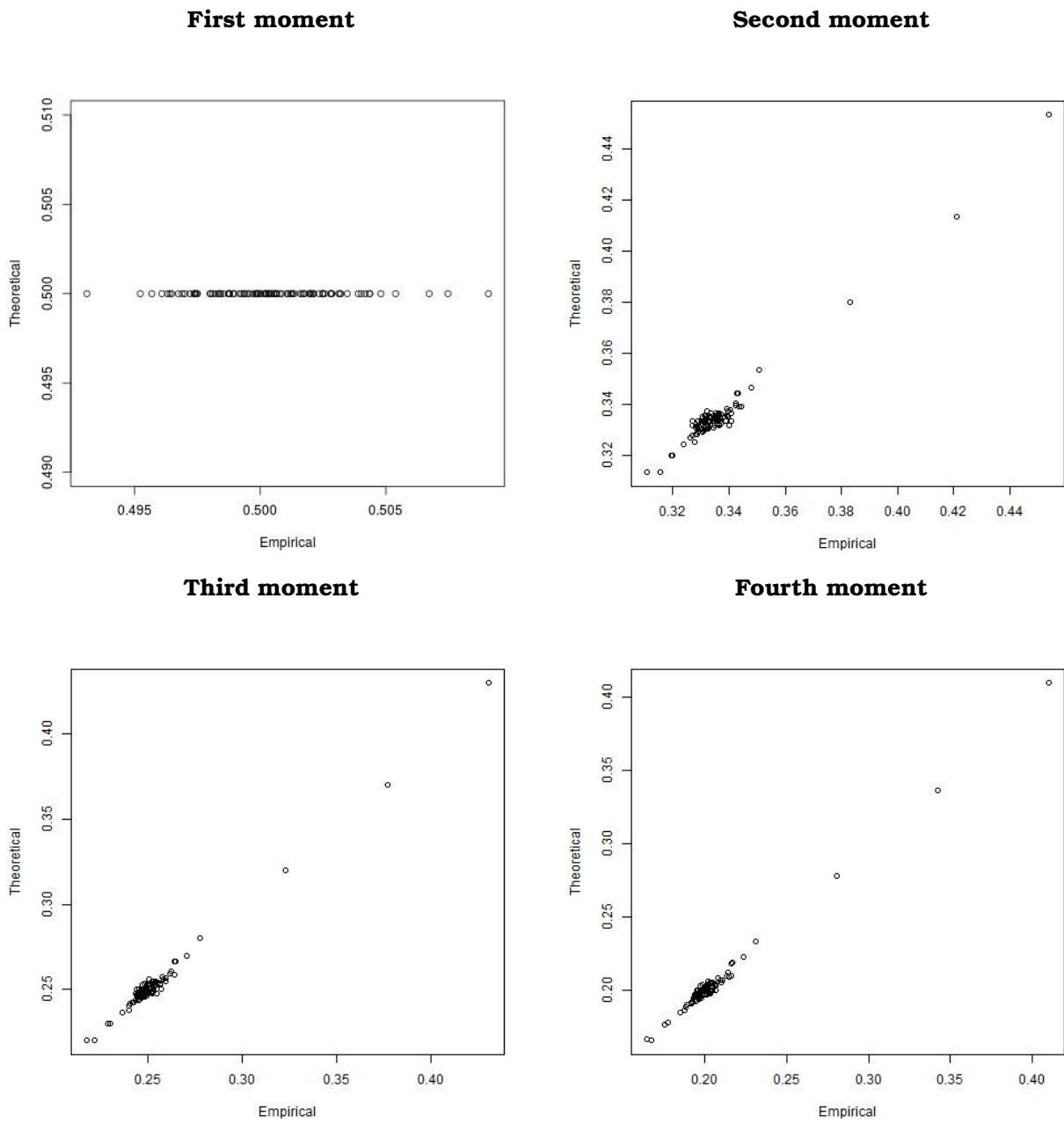


Figure 2: First four moments of $X - \text{floor}(X)$ of the uniform distribution, with parameters $-a = b$ for $a = 0.1, \dots, 10$.

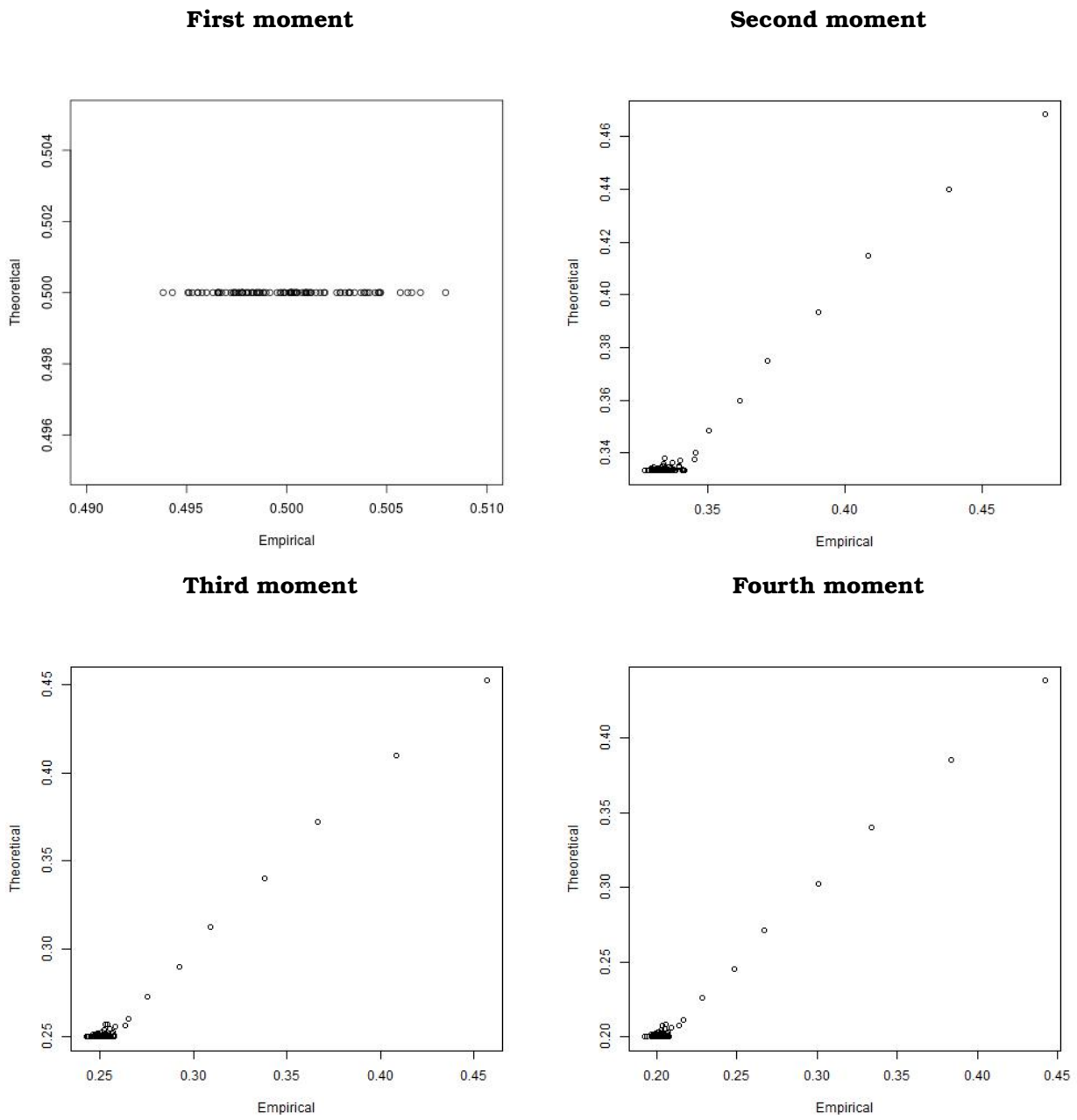


Figure 3: First four moments of $X - \text{floor}(X)$ of the triangular distribution, with parameters $c = 0, -a = b$ for $a = 0.1, \dots, 10$.

4.1.2 Plots for the moments of $X - \text{floor}(X + \frac{1}{2})$

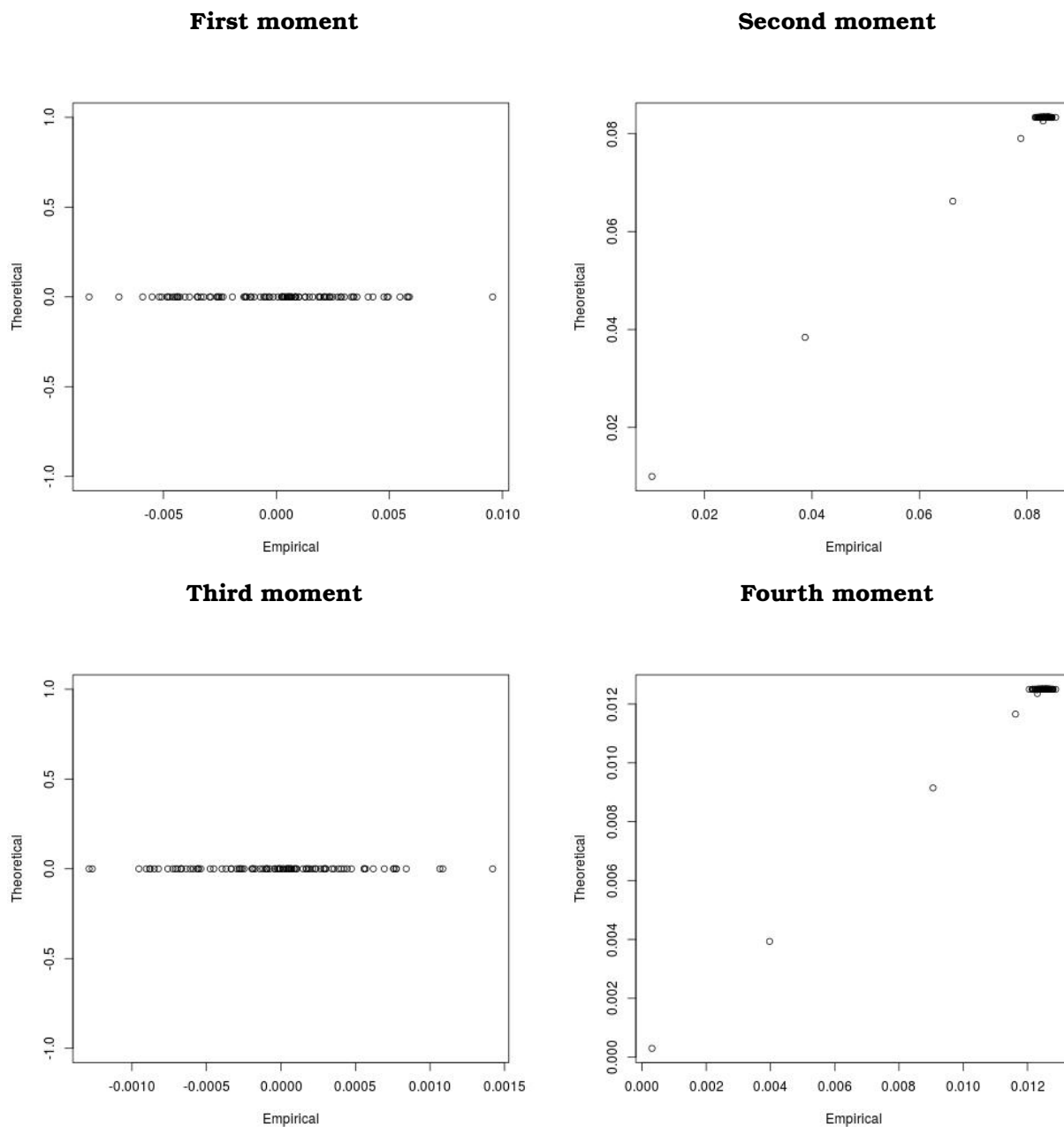


Figure 4: First four moments of $X - \text{floor}(X + \frac{1}{2})$ for the normal distribution, with parameters $\mu = 0$ and standard deviation $a = 0.1, \dots, 10$.

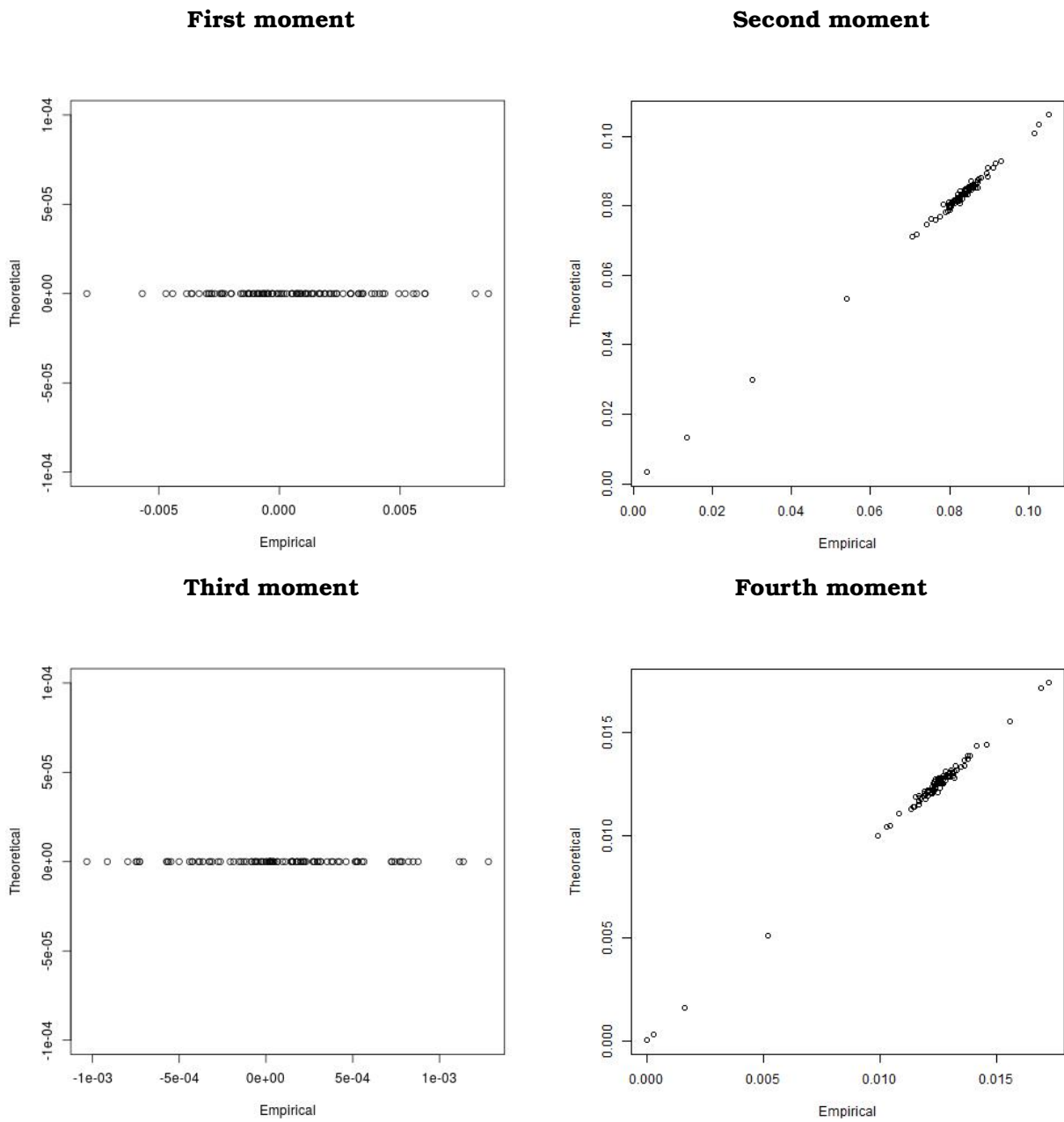


Figure 5: First four moments of $X - \text{floor}\left(X + \frac{1}{2}\right)$ of the uniform distribution, with parameters $-a = b$ for $a = 0.1, \dots, 10$.

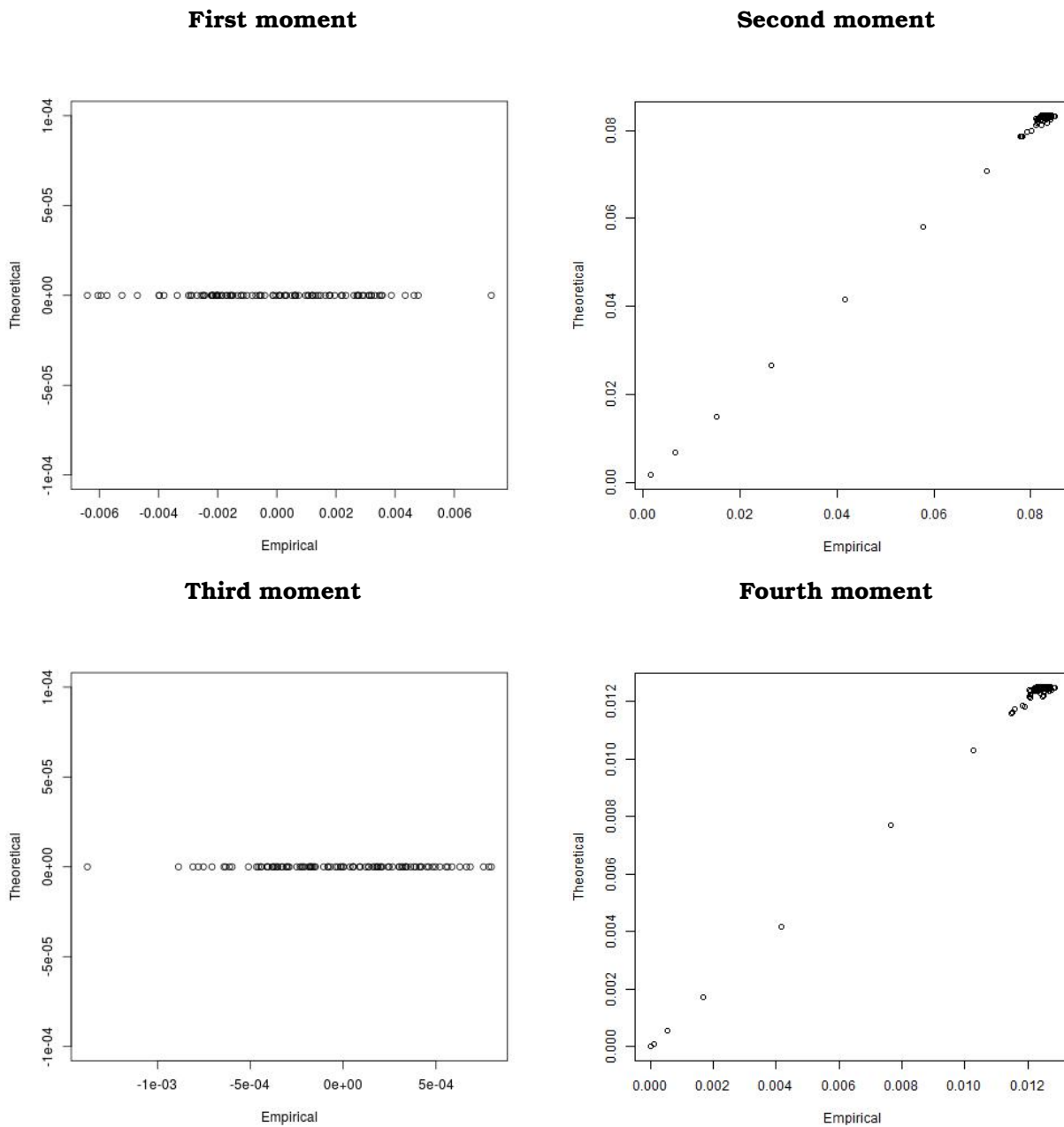


Figure 6: First four moments of $X - \text{floor}(X + \frac{1}{2})$ of the triangular distribution, with parameters $c = 0, -a = b$ for $a = 0.1, \dots, 10$.

4.2 Error tables

In addition to the visualised results of Section 4.1, we provide tables of the mean squared error between numerical and theoretical values of the k th moment for $k = 1, 2, 3, 4$. The distributions remain the same while the deviation has been computed for selected parameter values of a .

Overall the difference between numerical and theoretical approaches is fairly close to naught, for both regular rounded error and shifted error moments. This puts forth further evidence for the feasibility of the theoretical results in Sections 2 and 3.

4.2.1 Mean squared errors for the moments of $X - \text{floor}(X)$

Table 1: MSE of the empirical and theoretical values for the normal distribution.

a	k	1	2	3	4
0.1		2.364196e-05	3.090310e-05	2.984747e-05	2.722596e-05
1		9.033631e-06	1.318285e-06	5.234990e-08	1.142558e-06
2		1.884854e-06	7.064957e-07	2.127054e-07	6.517809e-08
3		8.582170e-07	9.349346e-08	8.328996e-08	4.806649e-07
4		2.728113e-06	4.661411e-06	5.100822e-06	4.918637e-06
5		1.683766e-06	9.190482e-07	6.206288e-07	5.312952e-07
6		3.359889e-06	1.850607e-06	4.870644e-07	2.729104e-08
7		1.595674e-06	8.675618e-07	4.078100e-07	1.846421e-07
8		1.468702e-06	2.132272e-06	2.906684e-06	3.418431e-06

Table 2: MSE of the empirical and theoretical values for the uniform distribution.

a	k	1	2	3	4
0.1		5.900041e-08	2.340302e-07	3.294760e-07	4.138349e-07
1		8.992801e-06	1.018269e-05	9.367884e-06	8.361929e-06
2		2.610486e-06	2.310805e-06	1.476468e-06	8.137844e-07
3		1.240942e-05	7.234127e-06	4.356404e-06	3.142111e-06
4		1.130977e-07	4.766217e-06	7.697295e-06	7.981190e-06
5		1.500870e-06	8.881806e-07	4.367888e-07	2.420640e-07
6		4.756723e-05	4.185227e-05	3.364696e-05	2.674131e-05
7		4.012009e-06	3.343534e-06	3.480836e-06	3.739969e-06
8		3.143884e-06	3.913011e-06	2.297650e-06	1.124660e-06

Table 3: MSE of the empirical and theoretical values for the triangular distribution.

a	k	1	2	3	4
0.1		2.175290e-05	2.027581e-05	1.796827e-05	1.586509e-05
1		1.188664e-05	9.232077e-06	6.423183e-06	4.522853e-06
2		2.291893e-06	1.470480e-06	1.168993e-06	1.098933e-06
3		2.512014e-07	8.698671e-09	1.147041e-08	5.433561e-08
4		9.296816e-07	1.047689e-06	9.158490e-07	6.850873e-07
5		1.058201e-05	6.332940e-06	3.538913e-06	2.173855e-06
6		1.148803e-05	1.189744e-05	1.022145e-05	8.669491e-06
7		1.789967e-05	4.095275e-05	5.251031e-05	5.476888e-05
8		7.053805e-06	9.674796e-06	9.798778e-06	8.776999e-06

4.2.2 Mean squared errors for the moments of $X - \text{floor}(X + \frac{1}{2})$

Table 4: MSE of the empirical and theoretical values for the normal distribution.

a	k	1	2	3	4
0.1		1.897385e-07	5.128911e-08	1.438813e-09	1.504597e-10
1		4.848491e-05	3.929998e-07	5.765655e-07	4.875014e-09
2		6.846606e-06	3.016127e-10	1.352258e-07	1.778393e-09
3		9.816704e-07	2.411716e-08	2.606065e-10	3.236768e-09
4		3.340776e-05	6.215697e-08	1.633690e-07	2.822030e-10
5		2.103549e-05	8.081596e-07	6.759402e-07	2.861358e-08
6		2.237473e-05	4.477350e-09	9.073246e-07	4.855412e-12
7		7.358063e-06	2.029926e-07	8.169570e-08	7.196841e-09
8		2.259615e-05	1.060460e-07	3.841394e-07	1.398618e-09

Table 5: MSE of the empirical and theoretical values for the uniform distribution.

a	k	1	2	3	4
0.1		2.592975e-07	2.208311e-09	1.219367e-11	1.742729e-13
1		6.867934e-07	1.472720e-08	3.942234e-08	3.159564e-09
2		3.571091e-06	3.178441e-09	1.243143e-07	5.069440e-11
3		4.005817e-07	3.658153e-07	2.387804e-08	1.567003e-08
4		8.359841e-08	2.490602e-07	2.355163e-09	4.251432e-08
5		1.188630e-06	7.675209e-08	1.423636e-07	1.764381e-08
6		5.980640e-07	5.767026e-08	1.649205e-09	3.222833e-09
7		1.957862e-05	2.150672e-09	1.065707e-06	1.170508e-08
8		7.498646e-05	1.437785e-08	1.293426e-06	8.886361e-10

Table 6: MSE of the empirical and theoretical values for the triangular distribution.

a	k	1	2	3	4
0.1		1.565134e-08	4.913611e-10	3.399670e-13	3.169243e-14
1		2.221216e-05	5.532335e-07	7.866869e-07	1.350012e-08
2		3.734955e-07	7.588791e-07	6.265540e-08	2.387643e-08
3		4.633213e-06	2.142906e-08	3.093012e-08	2.244004e-10
4		2.193109e-06	9.394391e-07	4.695382e-08	4.935507e-08
5		2.607279e-07	1.200193e-08	3.504874e-08	3.221698e-09
6		1.133684e-05	1.881535e-09	2.081667e-07	2.630884e-10
7		1.451247e-06	3.338605e-07	2.714815e-07	4.224998e-13
8		3.292962e-05	1.449909e-07	5.660277e-07	1.530374e-09

5 Conclusions

We have derived explicit expressions for general order moments of regular rounded error and shifted rounded error. The general expressions have been specialised to ten distributions arising in the signal processing area (references to applications of these distributions in signal processing are given): uniform distribution, triangular distribution, trapezoidal distribution, house distribution, curved trapezoidal distribution, hexagonal distribution, sinusoidal distribution, convolutions of uniform and triangular distributions, convolutions of two triangular distributions and the normal distribution. The specialised expressions are all simple except for the normal distribution.

We have checked the correctness of the derivations through two numerical studies: i) by plotting values of the derived expressions versus those obtained by simulation; ii) by tabulating values of the derived expressions versus those obtained by simulation. Both studies show that the derived expressions are accurate.

A future work is to extend this work to other forms of rounding and also to consider rounding of more than one variable of interest.

References

- [1] S. Ardalan, S. T. Alexander, Fixed-point roundoff error analysis of the exponentially windowed RLS algorithm for time-varying systems, *IEEE Transactions on Acoustics, Speech and Signal Processing* 35 (1987) 770-783.
- [2] C. W. Barnes, B. Tran, S. Leung, On the statistics of fixed-point roundoff error, *IEEE Transactions on Acoustics, Speech and Signal Processing* 33 (1985) 595-606.
- [3] P. Csordas, A. Mersich, I. Kollar, Digital dither: Decreasing round-off errors in digital signal processing, in: *Proceedings of the IEEE International Symposium on Intelligent Signal Processing, 2003*, pp. 9-14.
- [4] Y. A. Gadzhiev, On the variance of a centered random value roundoff error, *Signal Processing* 106 (2015) 30-40.
- [5] J. Hou, J. H. Ge and J. Li, Trapezoidal companding scheme for peak-to-average power ratio reduction of OFDM signals, *State Key Lab of Integrated Services Networks, Xidian University* (2009).
- [6] S. Kawarai, T. Murakami, An optimization procedure to minimize the roundoff noise in cascade floating-point digital filters, in: *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1989*, pp. 884-887.
- [7] R. Li, S. Nadarajah, Mean and Variance of round off error, *Signal Processing* 127 (2016) 185-190.
- [8] M. Press, Round-off error of floating-point digital filters, in: *Papers on Digital Signal Processing, volume 1, 1969*, pp. 94-102.
- [9] I. G. Vladimirov, P. Diamond, A uniform white-noise model for fixed-point roundoff errors in digital systems, *Automation and Remote Control* 63 (2002) 753-765.
- [10] B. Widrow, I. Kollar, *Quantization Noise: Round-off Error in Digital Computation*, *Signal Processing, Control and Communications*, Cambridge University Press, New York, 2008.
- [11] P. W. Wong, Quantization noise, fixed-point multiplicative roundoff noise, and dithering, *IEEE Transactions on Acoustics, Speech and Signal Processing* 38 (1990) 286-300.
- [12] Y. J. Yu, Y. C. Lim, Roundoff noise analysis of signals represented using signed power-of-two terms, in: *Proceedings of the 14th European Signal Processing Conference, 2006*, pp. 1-4.

Appendix: Proofs of Theorems 2.1 to 2.4

Proof of Theorem 2.1 Note that

$$\mathbb{E} [(X - \text{floor}(X))^k] = \sum_{i=0}^k (-1)^i \binom{k}{i} \mathbb{E} [X^{k-i} \text{floor}(X)^i].$$

We investigate the expected values, and follow the basic principle of the proof for Theorems 1 and 3 in Li and Nadarajah [7]:

$$\begin{aligned}
\mathbb{E} [X^{k-i} \text{floor}(X)^i] &= \int_{-\infty}^{\infty} x^{k-i} \text{floor}(x)^i f(x) dx \\
&= \sum_{j=-\infty}^{\infty} j^i \int_j^{j+1} x^{k-i} f(x) dx \\
&= \sum_{j=-\infty}^{\infty} j^i \left[\int_{-\infty}^{j+1} x^{k-i} f(x) dx - \int_{-\infty}^j x^{k-i} f(x) dx \right] \\
&= \sum_{j=-\infty}^{\infty} j^i [M_{k-i}(j+1) - M_{k-i}(j)].
\end{aligned}$$

By inserting the result above into the first equation, we have completed the proof.

Proof of Theorem 2.2 Once more, by the binomial theorem:

$$\mathbb{E} \left[\left(X - \text{floor} \left(X + \frac{1}{2} \right) \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \mathbb{E} \left[X^{k-i} \text{floor} \left(X + \frac{1}{2} \right)^i \right].$$

Set $Y = X + \frac{1}{2}$. Then

$$\begin{aligned}
\int_{-\infty}^{\infty} x^{k-i} \text{floor} \left(x + \frac{1}{2} \right)^i f(x) dx &= \int_{-\infty}^{\infty} \left(y - \frac{1}{2} \right)^{k-i} \text{floor}(y)^i f \left(y - \frac{1}{2} \right) dy \\
&= \sum_{j=-\infty}^{\infty} j^i \int_j^{j+1} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\
&= \sum_{j=-\infty}^{\infty} j^i \left[\int_{-\infty}^{j+\frac{1}{2}} (y)^{k-i} f(y) dy - \int_{-\infty}^{j-\frac{1}{2}} (y)^{k-i} f(y) dy \right] \\
&= \sum_{j=-\infty}^{\infty} j^i [M_{k-i}^*(j+1) - M_{k-i}^*(j)].
\end{aligned}$$

Proof of Theorem 2.3 Note that

$$\mathbb{E} [(X - \text{floor}(X))^k] = \sum_{i=0}^k (-1)^i \binom{k}{i} \mathbb{E} [X^{k-i} \text{floor}(X)^i].$$

Let $i \geq 1$ in the next lines, since $i = 0$ is already clear. We have

$$\begin{aligned}
\int_a^b x^{k-i} \text{floor}(x)^i f(x) dx &= \sum_{j=p+1}^q j^i \int_j^{j+1} x^{k-i} f(x) dx \\
&= \sum_{j=0}^q j^i \int_j^{j+1} x^{k-i} f(x) dx - \sum_{j=0}^p j^i \int_j^{j+1} x^{k-i} f(x) dx \\
&= \sum_{j=0}^q [j^i - (j-1)^i] \int_j^b x^{k-i} f(x) dx - \sum_{j=0}^p [j^i - (j-1)^i] \int_j^a x^{k-i} f(x) dx \\
&= \sum_{j=0}^q [j^i - (j-1)^i] \left[\int_a^b x^{k-i} f(x) dx - \int_a^j x^{k-i} f(x) dx \right] \\
&\quad + \sum_{j=0}^p [j^i - (j-1)^i] \int_a^j x^{k-i} f(x) dx \\
&= \sum_{j=0}^q [j^i - (j-1)^i] [M_{k-i}(b) - M_{k-i}(j)] + \sum_{j=0}^p [j^i - (j-1)^i] \int_a^j x^{k-i} f(x) dx \\
&= q^i M_{k-i}(b) - \sum_{j=p+1}^q [j^i - (j-1)^i] M_{k-i}(j).
\end{aligned}$$

Since $M_{k-i}(b) = M_{k-i}(q)$ the last addend of the sum reads as $[q^i - (q-1)^i] M_{k-i}(b)$ and thus we have our expression.

Proof of Theorem 2.4 Once more, by the binomial theorem:

$$\mathbb{E} \left[\left(X - \text{floor} \left(X + \frac{1}{2} \right) \right)^k \right] = \sum_{i=0}^k (-1)^i \binom{k}{i} \mathbb{E} \left[X^{k-i} \text{floor} \left(X + \frac{1}{2} \right)^i \right].$$

Set $Y = X + \frac{1}{2}$. Again let $i \geq 1$. We have

$$\begin{aligned} \int_a^b \left(y - \frac{1}{2} \right)^{k-i} \text{floor}(y) f \left(y - \frac{1}{2} \right) dy &= \sum_{j=p+1}^q j^i \int_j^{j+1} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &= \sum_{j=0}^q j^i \int_j^{j+1} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &\quad - \sum_{j=0}^p j^i \int_j^{j+1} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &= \sum_{j=0}^q [j^i - (j-1)^i] \int_j^{b+\frac{1}{2}} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &\quad - \sum_{j=0}^p [j^i - (j-1)^i] \int_j^{a+\frac{1}{2}} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &= \sum_{j=0}^q [j^i - (j-1)^i] \left[\int_{a+\frac{1}{2}}^{b+\frac{1}{2}} \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \right. \\ &\quad \left. - \int_{a+\frac{1}{2}}^j \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \right] \\ &\quad + \sum_{j=0}^p [j^i - (j-1)^i] \int_{a+\frac{1}{2}}^j \left(y - \frac{1}{2} \right)^{k-i} f \left(y - \frac{1}{2} \right) dy \\ &= q^i M_{k-i}(b) - \sum_{j=p+1}^q [j^i - (j-1)^i] M_{k-i}^*(j). \end{aligned}$$

Equivalently to the previous proofs, we have thus shown the expression.