International Journal of Pure and Applied Mathematics
Volume 79 No. 1 2012, 67-85
ISSN: 1311-8080 (printed version)
url: http://www.ijpam.eu
ijpam.eu

ON MOMENTS OF SAMPLE MEAN AND VARIANCE

Jordanka A. Angelova
Department of Mathematics
University of Chemical Technology and Metallurgy
8, Kliment Ohridsky, Blvd., Sofia, 1756, BULGARIA

Abstract: First four initial moments of the sample variance are derived. The four central moments of the sample mean are represented, and values are checked via characteristic functions. Obtained results are verified for a normal population. We numerically obtain probability density functions of the sample variance of random variables exponentially distributed via Pearson family. In discrete case, for Bernoulli distributed random variables, some example concerning probability mass functions are presented. Graphical representation and comparison with standard approximation are performed.

AMS Subject Classification: 62H10, 60E05, 62E17
Key Words: sample mean, sample variance, moments, Pearson system

1. Introduction

Statistical inference uses sample data to derive conclusions about population's peculiarities represented by random samples. Widespread inferences are: numerical estimations of population's parameters, hypothesis testing, confidence intervals and related to them prediction and tolerance intervals, deriving relations within the population - regression and correlation analysis.

Common practice, when we have an a-priori information about the type of distribution function (d.f.) F of the random variable (r.v.) X describing the population, is to estimate parameters of this function by random sample, obtain and proceed with the estimated d.f. $\widehat{F}$. Usually, numerical measures of

Received: May 12, 2012

a "central tendency" of the population is the mean and of the variability is the variance.

Let $x_1, x_2, \ldots, x_n$ be independent observations on a r.v. X, i.e. the r.v.'s $X_1, X_2, \ldots, X_n$ are considered to be independent identically distributed (i.i.d.) with a d.f. F and finite absolute moments.

Statistics $\bar{X}$ - sample mean (average) and S^2 - sample variance, based on the sample $X_1, X_2, \ldots, X_n$, are introduced traditionally:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,, \tag{1}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right)^2\right] \tag{2}$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - \frac{1}{n-1}\sum_{i=1}^{n} X_i X_j\right). \tag{3}$$

They are usual estimators for the expected value EX and variance DX of the r.v. X. Using tests or/and confidence intervals, built on their sample analogues, or coefficients of them, we have to draw a statistical inference. Thus it is necessary to find d.f.'s of $\bar{X}$ or/and S^2. If we know F or $\widehat{F}$, then applying some statistical transformations we can obtain d.f.'s of the sample mean, $F_{\bar{X}}$, and F_{S^2} or their approximations. Otherwise Pearson systems or some asymptotic expansions have to be used.

There are abundance of articles, books and e-materials devoted to the moments (of order till 8) of the sample mean and first two moments of the sample variance.

As the Pearson continuous density satisfies the first order ordinary differential equation (ODE) with parameters built on first four moments, we shall derive $E\bar{X}^k$ and $E(S^2)^k$, $k = 1, ..., 4$. However the moments of $\bar{X}$ are checking via characteristic function (c.f.) and moments of S^2 are received by expanding the corresponding sums. Verifications and some numerical example are presented too.

2. Notations and Preliminaries

Let X denote a real valued r.v. with finite absolute moments and mean a. The central moments of X are the moments of the r.v. $X - EX = X - a$.

Since $(X-a)^n$, $n \in \mathbb{N}$, may be expanded in powers of X and a, then the central moments are presented as sums of the initial moments and vice versa, for example see [3]. Hence, for $n \in \mathbb{N}$, we set nth, initial moment and nth central moment of a r.v. X as:

$$a_n = EX^n\,, \quad |a_n| < \infty\,, \qquad a = a_1 = EX\,, \tag{4}$$

$$\mu_n = E(X-a)^n = E\sum_{j=0}^{n} C_n^j X^{n-j}(-a)^j = \sum_{j=0}^{n} C_n^j (-a)^j a_{n-j}\,, \tag{5}$$

where $C_n^j = \binom{n}{j}$. It is obvious, that:

— the zero initial and central moments of X are equal to one, $a_0 = \mu_0 = 1$;

— the first central moment of X is zero, $\mu_1 = E(X-a) = 0$ and the second central moment - variance is $\mu_2 = a_2 - a^2$.

From (5) it follows

$$a_n = \sum_{j=0}^{n} C_n^j \mu_{n-j}\, a^j\,, \quad n \in \mathbb{N}\,. \tag{6}$$

So, if $X_1, X_2, \dots, X_n$ are i.i.d. with a d.f. F and finite absolute moments, then for each r.v. X_j is fulfilled $EX_j^n = a_n$ and $E(X_j - a)^n = \mu_n$, $n \in \mathbb{N} = \{1, 2, \dots\}$.

Let a r.v. X possesses finite absolute moments, then the c.f. φ is introduced as $\varphi(t) = Ee^{itX}$. The c.f. uniquely determines the initial moments via derivatives, according to identities

$$\varphi(0) = 1\,, \quad \varphi^{(k)}(0) = i^k a_k\,, \quad i^2 = -1\,, \quad k \in \mathbb{N}\,. \tag{7}$$

If a r.v. X is continuously distributed, then the inverting of c.f. gives the probability density function (p.d.f.) f of X.

Hence, if $X_1, X_2, \dots, X_n$ are i.i.d. with a c.f. φ, then the c.f. of the sample mean is

$$\varphi_{\bar{X}}(t) = E\exp(it\bar{X}) = \prod_{j=1}^{n} E\exp\left(i(t/n)X_j\right) = \varphi^n\left(t/n\right)\,, \tag{8}$$

and applying inversion formula for c.f., we can obtain the p.d.f. of $\bar{X}$.

The analytic expression for the p.d.f. of the sample variance, F'_{S^2}, from a continuous population with p.d.f. f one can obtain by the following equality

$$F'_{S^2}(x) = \frac{d}{dx} P\{S^2 < x\} = \frac{d}{dx} \int\limits_{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2/(n-1) < x} \cdots \int \prod_{i=1}^{n} f(x_i)\, dx_1 \dots dx_n\,.$$

The diagonalization of the quadratic form $\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ by orthogonal transformation with eigenvalues $\lambda_1=0$ and $\lambda_2=\cdots=\lambda_n=1$ and corresponding eigenvectors leads to successful receiving of F'_{S^2} only for normally distributed r.v.'s (see [3], [17]). So, for non-normal initial population the distribution of the sample variance is difficult to find, but moments of S^2 by direct calculations, cumulants, k-statistics and polykays (generalized k-statistics) can be derived.

The expectation $W^k_{2,(n)}$ of $[\sum_{i=1}^{n}(X_i-\bar{X})^2]^k$, $k=1,...,4$, is derived in [24] via expectations of products of type $[\sum(X_i-a)]^r[\sum(X_i-a)]^s$ and $[\sum(X_i-\bar{X})]^r[\sum(X_i-a)]^s$. Expanding $\sum(X_i)^k\sum(X_i^2)^l$, $k+2l\le 8$ and direct computing averages yields the first four moments of the origin, ${}_2M'_k=E[\sum_{i=1}^{n}X_i^2/n-(\sum_{i=1}^{n}X_i/n)^2]^k=E[(1-1/n)S^2]^k$ ($(1-1/n)S^2$ is a biased statistic for μ_2) and about the mean in [2]. In [27] two algorithms for direct calculation of the product moments of order (l,m) of two sums of r.v.'s $X_1,\ldots,X_n$, $E[(\sum X_i^2)^l(\sum X_i)^m]$, without additional converting formulas and tables are proposed. The methods are applicable for computer implementation. Both algorithms need additional procedures, relating partitions of $2l+m$ and the number of elements of permutations without repetition (first algorithm) and number of elements of applied combinations for the second algorithm. This methodology can be used to calculate $E(S^2)^k$ after expanding $(S^2)^k$ in sum of elements $(\sum X_i^2)^l(\sum X_i)^m$, $l+2m=k$.

The augmented symmetric functions in $x_1,x_2,...,x_n$ is defined as (see [5], [9]):

$$[p_1^{\pi_1}p_2^{\pi_2},\ldots,p_s^{\pi_s}]=\sum x_{i_1}^{p_1}x_{i_2}^{p_1}\ldots x_{j_1}^{p_2}x_{j_2}^{p_2}\ldots x_{l_1}^{p_s}\ldots x_{l_s}^{p_s}\,, \qquad (9)$$

where: π_1 is a number of power p_1, ..., π_s is a number of power p_s; suffixes $i_1,\ldots,j_1,\ldots,l_s$ are different. For example

$$\sum x_i^2x_jx_k^3x_l=[1^2 23]=[321^2]\,,\quad \sum x_i^2x_j^2x_k^2=[2^3]\,.$$

For samples of i.i.d. r.v.'s by (9) follows

$$E[p_1^{\pi_1}p_2^{\pi_2}\ldots p_s^{\pi_s}]=n(n-1)\ldots(n-p+1)a_{p_1}^{\pi_1}a_{p_2}^{\pi_2}\ldots a_{p_s}^{\pi_s}\,,\quad p=\sum p_i\,. \qquad (10)$$

In [5] are presented tables for products of power-sums in terms of the augmented symmetric functions and vice versa up to order 12.

Polykays (parantheses - generalized k-statistics) were introduced in [25] (named in [26]) as homogeneous polynomial symmetric functions via symmetric means or angle brackets, e.g.

$$k2=(2)=\langle 2\rangle-\langle 11\rangle\,,\quad \langle pq\rangle=\sum_{i\neq j}x_i^px_j^q/n/(n-1)\,.$$

The development of cumulants, k-statistics, polykays and their generalizations can be find for instance in [7], [10], [16].

Computer oriented approaches for methods of recursive or direct calculation of the moments of moments or the product moments have been developed for decades. The *Mathematica* application package *mathStatica* has programs for computing polykays and k-statistics, also there is a *Maple* algorithm for these objects. For some approximations of p.d.f. of S^2 see [6], [15], [18], [19], [22], [23] and for historical review on the sample variance for non-normal populations refer to [14].

That is why, we shall apply the Pearson system for deriving p.d.f's, although in extreme cases the p.d.f. of S^2 could have more than one mode (Pearson curves do not have such forms). In continuous case, the system of Pearson distributions is established following the properties of normal distribution, which p.d.f. satisfies ODE of first order. For discrete populations instead of an ODE a difference equation is used. Parameters of both continuous and discrete systems are based on first four moments of the r.v. describing the investigated population. So, the p.d.f. f of continuous r.v. X satisfies the next ODE, centered about the mean, for instance see [13]:

$$\frac{df(x)}{dx} = \frac{c_0 + (x-a)}{b_0 + b_1(x-a) + b_2(x-a)x^2} f(x)\,, \quad EX = a\,. \tag{11}$$

Parameters c_0, b_0, b_1 and b_2 are completely defined by first four moments of X, as follows:

$$\begin{aligned} c_0 &= -b_1 = \mu_3(\mu_4 + 3\mu_2^2)/A\,, \quad b_0 = -\mu_2(4\mu_2\mu_4 - 3\mu_3^2)/A\,, \\ b_2 &= -2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3/A\,, \quad A = 10\mu_2\mu_4 - 12\mu_3^2 - 18\mu_2^3\,. \end{aligned} \tag{12}$$

In case $b_2 = 0$ the simpler ODE can be used, namely

$$\frac{df(x)}{dx} = \frac{c_0 + (x-a)}{b_0 + b_1(x-a)} f(x)\,, \tag{13}$$

where: $c_0 = \mu_3/(2\mu_2) = -b_1$ and $b_0 = -\mu_2$.

3. First Four Moments of Sample Mean and Variance

In this section we shall find the expected values of $\bar{X}^k$ and $(S^2)^k$, $k = 1, \ldots, 4$.

Theorem 1. *Let: the r.v's $X_1, X_2, \ldots, X_n$, $n \geq 2$, be i.i.d. with finite absolute fourth moments and c.f. $\varphi(t)$; the r.v. $\bar{X}$ is introduced by* (1).

Then the following statements for initial moments, $a_{k,\bar{X}} = E\bar{X}^k$, and central moments, $\mu_{k,\bar{X}} = E(\bar{X} - E\bar{X})^k$, $k = 1, \ldots, 4$, of $\bar{X}$ are valid:

$$\begin{aligned} a_{1,\bar{X}} &= a\,, \\ a_{2,\bar{X}} &= a^2 + \mu_2/n\,, \\ a_{3,\bar{X}} &= a^3 + 3a\mu_2/n + \mu_3/n^2\,, \\ a_{4,\bar{X}} &= a^4 + 6a^2\mu_2/n + (3\mu_2^2 + 4\mu_3 a)/n^2 + (\mu_4 - 3\mu_2^2)/n^3\,, \end{aligned} \tag{14}$$

and

$$\begin{aligned} \mu_{2,\bar{X}} &= \mu_2/n\,, \\ \mu_{3,\bar{X}} &= \mu_3/n^2\,, \\ \mu_{4,\bar{X}} &= 3\mu_2^2/n^2 + (\mu_4 - 3\mu_2^2)/n^3\,, \end{aligned} \tag{15}$$

Profs are presented in many books and papers. The expressions for the fourth initial and central moments of $\bar{X}$ one can find, for instance, in [20, 21] and $\mu_{5,\bar{X}} = 10\mu_3\mu_2/n^3 + (\mu_5 - 10\mu_3\mu_2)/n^4$ is given in [21]. The same result follows by sequential differentiating of the c.f. of $\bar{X}$, (7) and (5), that gives initial moments. Equalities (14) and (5) yield central moments of the sample mean. For example, fourth derivative of $\varphi_{\bar{X}}$ is

$$\begin{aligned} \varphi_{\bar{X}}^{(4)}(t) = (1/n^3)\varphi^{n-4}(t/n) \Big[&(n-1)(n-2)(n-3)\left(\varphi'(t/n)\right)^4 \\ &+ 6(n-1)(n-2)\varphi(t/n)\left(\varphi'(t/n)\right)^2 \varphi''(t/n) \\ &+ 3(n-1)\varphi^2(t/n)\left(\varphi''(t/n)\right)^2 \\ &+ 4(n-1)\varphi^2(t/n)\,\varphi'(t/n)\,\varphi'''(t/n) + \varphi^3(t/n)\,\varphi^{(4)}(t/n) \Big]\,. \end{aligned}$$

So, to calculate any momenta of $\bar{X}$ we have to expand $(\sum X_i)^k$ as sum $\sum\limits_{i_1, i_2, \ldots, i_k = 1}^{n} \prod_{j=1}^{k} X_{i_j}$, $k = 2, 3, \ldots$, and find all terms with k equal indices, terms in which $k-1$ indices are equal, and so on, till terms with all different indices. This can be done, for example, using a multinomial formula

$$\left(\sum_{i=1}^{n} x_i\right)^k = \sum_{\substack{k_1, k_2, \ldots, k_n \\ k_1 + k_2 + \cdots + k_n = k}} \binom{k}{k_1, k_2, \ldots, kn} \prod_{j=1}^{n} x_j^{k_j} \tag{16}$$

where $\binom{k}{k_1,k_2,\ldots,kn} = \frac{k!}{k_1!k_2!\ldots k_n!}$, $k_j \in \{0, 1, \ldots, k\}$, or applying the recursive formula $(\bar{X})^{k+1} = (\bar{X})^k \bar{X}$. Deriving sums in products of $\prod_{j=1}^{m} x_{i_j}^{k_j}$, not ordered lexicographically we have to find all partitions of k in positive integers, so that $k_1 + k_2 + \cdots + k_m = k$, and $i_1, i_2, \ldots, i_m$ to be a permutation of order m of integers $\{1, 2, \ldots, n\}$. Basic results of partitioning of an integer one can find in many books and works devoted to combinatorics, e.g. see [4], [12], [28].

Therefore, for $k = 2, 3, \ldots, n$, we rewrite (16) by using (9) as

$$\left(\sum_{i=1}^{n} x_i\right)^k = \sum_{i_1=1}^{n} x_{i_1}^k + \cdots + C_k^2 \sum_{\substack{i_1, i_2, \ldots, i_{k-1}=1 \\ i_1 \neq i_2 \neq \cdots \neq i_{k-1}}}^{n} x_{i_1}^2 \prod_{j=2}^{k-1} x_{i_j} + \sum_{\substack{i_1, i_2, \ldots, i_k=1 \\ i_1 \neq i_2 \neq \cdots \neq i_k}}^{n} \prod_{j=1}^{k} x_{i_j} = [k] + \cdots + C_k^2 [21^{k-2}] + [1^k]. \quad (17)$$

Coefficients before the sums are obtained by the formula

$$\left(\frac{k!}{k_1! k_2! \ldots k_n!}\right) / (p_1! p_2! \ldots p_l!)$$

where p_j is a number of terms with equal magnitude k_j in the partition $k_1 + k_2 + \cdots + k_n$.

In expansion of forth power of (2) there is $(\sum X_j)^8$. That is why the exact representation via sums of products of x_j not ordered lexicographically is provided. By example, for the partition $3+2+1+1+1$ coefficient's value before the sum $\sum x_{i_1}^3 x_{i_2}^2 x_{i_3} x_{i_4} x_{i_5} = [321^3]$ is $\frac{8!}{3!2!1!1!1!}/3! = 560$. Therefore

$$\begin{aligned}\left(\sum x_i\right)^8 &= [8] + 8\,[71] + 28\,[62] + 28\,[61^2] + 56\,[53] + 168\,[521] \\ &+ 56\,[51^3] + 35\,[4^2] + 280\,[431] + 210\,[42^2] + 420\,[421^2] \\ &+ 70\,[41^4] + 280\,[3^2 2] + 280\,[3^2 1^2] + 840\,[32^2 1] + 560\,[321^3] \\ &+ 56\,[31^5] + 105\,[2^4] + 420\,[2^3 1^2] + 210\,[2^2 1^4] + 28\,[21^6] + [1^8].\end{aligned}$$

Theorem 2. *Let: the r.v's $X_1, X_2, \ldots, X_n$, $n \geq 2$, be i.i.d. with finite absolute eighth moments and the r.v. S^2 is introduced by (2).*

Then the following statements for initial moments $a_{k,S^2} = E(S^2)^k$ and *central moments* $\mu_{k,S^2} = E(S^2 - \mu_2)^k$, $k = 1, \ldots, 4$, *of* S^2 *are valid:*

$$\begin{aligned}
a_{1,S^2} &= \mu_2 \,, \\
a_{2,S^2} &= \mu_2^2 + (\mu_4 - \mu_2^2)/n + 2\mu_2^2/n/(n-1) \,, \\
a_{3,S^2} &= \mu_2^3 + 3\mu_2(\mu_4 - \mu_2^2)/n \\
&\quad +(\mu_6 - 3\mu_4\mu_2 - 6\mu_3^2 + 8\mu_2^3)/n^2 \\
&\quad +2\mu_2(6\mu_4 - 7\mu_2^2)/n^2/(n-1) - 4(\mu_3^2 - 2\mu_2^3)/n^2/(n-1)^2 \,, \\
a_{4,S^2} &= \mu_2^4 + 6\mu_2^2(\mu_4 - \mu_2^2)/n \\
&\quad +(4\mu_6\mu_2 + 3\mu_4^2 - 18\mu_4\mu_2^2 - 24\mu_3^2\mu_2 + 23\mu_2^4)/n^2 \\
&\quad +(\mu_8 - 4\mu_6\mu_2 - 24\mu_5\mu_3 - 3\mu_4^2 + 72\mu_4\mu_2^2 + 96\mu_3^2\mu_2 - 86\mu_2^4)/n^3 \\
&\quad +4(6\mu_6\mu_2 + 6\mu_4^2 - 39\mu_4\mu_2^2 - 40\mu_3^2\mu_2 + 45\mu_2^4)/n^3/(n-1) \\
&\quad +4(36\mu_4\mu_2^2 - 8\mu_5\mu_3 + 52\mu_3^2\mu_2 - 61\mu_2^4)/n^3/(n-1)^2 \\
&\quad +8(\mu_4^2 - 6\mu_4\mu_2^2 - 12\mu_3^2\mu_2 + 15\mu_2^4)/n^3/(n-1)^3 \,.
\end{aligned} \tag{18}$$

and

$$\begin{aligned}
\mu_{2,S^2} &= (\mu_4 - \mu_2^2)/n + 2\mu_2^2/n/(n-1) \,, \\
\mu_{3,S^2} &= (\mu_6 - 3\mu_4\mu_2 - 6\mu_3^2 + 2\mu_2^3)/n^2 \\
&\quad +4(3\mu_4\mu_2 - 5\mu_2^3)/n^2/(n-1) - 4(\mu_3^2 - 2\mu_2^3)/n^2/(n-1)^2 \,, \\
\mu_{4,S^2} &= 3(\mu_4 - \mu_2^2)^2/n^2 \\
&\quad +(\mu_8 - 4\mu_6\mu_2 - 24\mu_5\mu_3 - 3\mu_4^2 + 24\mu_4\mu_2^2 + 96\mu_3^2\mu_2 - 18\mu_2^4)/n^3 \\
&\quad +12(2\mu_6\mu_2 + 2\mu_4^2 - 17\mu_4\mu_2^2 - 12\mu_3^2\mu_2 + 18\mu_2^4)/n^3/(n-1) \\
&\quad -4(8\mu_5\mu_3 - 36\mu_4\mu_2^2 - 56\mu_3^2\mu_2 + 69\mu_2^4)/n^3/(n-1)^2 \\
&\quad +8(\mu_4^2 - 6\mu_4\mu_2^2 - 12\mu_3^2\mu_2 + 15\mu_2^4)/n^3/(n-1)^3 \,,
\end{aligned} \tag{19}$$

The proof is given in Appendix A. Following this approach any moment of $\bar{X}$ or S^2 can be obtained, although the algebraic transformations and calculations are severe for the higher moments.

4. Verifications

In this section the Pearson system will apply to check the correctness of the results obtained in Sections 3.

4.1. Continuous Case

The continuous Pearson system is described by ODE (11), centered about the mean.

Distribution of $\bar{X}$ from gamma distributed population. Let r.v.'s X_j, $j = 1, 2, \ldots, n$, be independent gamma distributed, $X_j \sim \Gamma(\alpha, \beta)$ $(\alpha, \beta > 0)$ with: p.d.f. $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ for $x > 0$ and $f(x) = 0$ for $x \le 0$; $a_k = \alpha(\alpha+1)\ldots(\alpha+k-1)/\beta^k$, $k \in \mathbb{N}$; $\mu_2 = \alpha/\beta^2$, $\mu_3 = 2\alpha/\beta^3$ and $\mu_4 = 3\alpha(\alpha+2)/\beta^4$.

Formulas (15) yield that central moments $\mu_{k,\bar{X}}$ of the sample mean are:

$\mu_{2,\bar{X}} = \alpha/(n\beta^2)$, $\mu_{3,\bar{X}} = 2\alpha/(n^2\beta^3)$ and $\mu_{4,\bar{X}} = 3\alpha(\alpha+2/n)/(n^2\beta^4)$. As the gamma distributed r.v. obeys an ODE of type (11) with parameter $b_2 = 0$, so we apply ODE (13), with parameters: $c_0 = -b_1 = 1/(n\beta)$, $b_0 = -\alpha/(n\beta^2)$, $a = \alpha/\beta$, which solution is $f_{\bar{X}}(x) = Ce^{-n\beta x} x^{n\alpha-1}$. The domain of $f_{\bar{X}}$ and the constant of integration C (normalizing constant) are determined by the fact that it is a p.d.f., thus $C \int\limits_0^\infty e^{-n\beta x} x^{n\alpha-1}\, dx = 1$, therefore $f_{\bar{X}}(x) = \frac{(n\beta)^{n\alpha}}{\Gamma(n\alpha)} x^{n\alpha-1} e^{-n\beta x}$ for $x > 0$ and $f_{\bar{X}}(x) = 0$ for $x \le 0$, i.e., $\bar{X} \sim \Gamma(n\alpha, n\beta)$ as it known.

Distribution of S^2 from normal population. Let r.v.'s X_j, $j = 1, 2, \ldots, n$, be independent normally distributed, $X_j \sim N(a, \sigma^2)$ with: p.d.f. $f_{X_j}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$, $x \in \mathbb{R}$; $\mu_{2k} = (2k-1)!!\sigma^{2k}$ and $\mu_{2k+1} = 0$, $k \in \mathbb{N}$.

By (18) we obtain moments of the sample variance S^2: $a_{1,S^2} = \sigma^2$, $a_{2,S^2} = \sigma^4(n+1)/(n-1)$, $a_{3,S^2} = \sigma^6(n+1)(n+3)/(n-1)^2$ and $a_{4,S^2} = \sigma^8(n+1)(n+3)(n+5)/(n-1)^3$. Therefore for the central moments of sample variance we have:

$$\mu_{2,S^2} = 2\sigma^4/(n-1)\,,\ \mu_{3,S^2} = 8\sigma^6/(n-1)^2\,,\ \mu_{4,S^2} = 12\sigma^8(n+3)/(n-1)^3\,.$$

Hence, the constants (12) of ODE (11) are:

$$A = 96\sigma^{12}(n+1)/(n-1)^4\,, \quad c_0 = 2\sigma^2/(n-1) = -b_1\,,$$

$$b_0 = -2\sigma^4/(n-1)\,, \quad b_2 = 0\,.$$

Rewriting differential equation as

$$\frac{df_{S^2}(x)}{f_{S^2}(x)} = -\frac{n-1}{2\sigma^2}\left[\frac{(x-\sigma^2) + 2\sigma^2/(n-1)}{(x-\sigma^2) + 2\sigma^2}\right] dx\,,$$

we obtain a solution $f_{S^2}(x) = Cx^{\frac{n-1}{2}-1} e^{-\frac{n-1}{2\sigma^2}x}$. We have $C\int\limits_0^\infty f_{S^2}(x)\, dx = 1$, which yields C and finally derive

$$f_{S^2}(x) = \begin{cases} \dfrac{((n-1)/(2\sigma^2))^{(n-1)/2}}{\Gamma((n-1)/2)} x^{\frac{n-1}{2}-1} e^{-\frac{n-1}{2\sigma^2}x}\,, & x > 0 \\ 0\,, & x \le 0 \end{cases}. \qquad (20)$$

This function is a density of a gamma distributed r.v. with parameters $(n-1)/2$ and $(n-1)/(2\sigma^2)$, i.e. $S^2 \sim \Gamma(\frac{n-1}{2}, \frac{n-1}{2\sigma^2})$. The same result follows from the fact that the r.v. $(n-1)S^2/\sigma^2$ is chi-squared (χ^2) distributed with $n-1$ degree of freedom, for example see [17], $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. So, for $x > 0$, the d.f. of the sample variance is

$$P\{S^2 < x\} = P\{(n-1)S^2/\sigma^2 < (n-1)x/\sigma^2\} = F_{\chi^2_{n-1}}((n-1)x/\sigma^2)\,.$$

Therefore for the p.d.f. of S^2 is valid the following equality $f_{S^2}(x) = f_{\chi^2_{n-1}}((n-1)x/\sigma^2)(n-1)/\sigma^2$, that is scaled χ^2 function, and performing some transformation the above function is equal to the p.d.f. from (20).

Distribution of S^2 from exponential population. Let r.v.'s X_j, $j = 1, 2, \dots, n$, be independent exponentially distributed with: p.d.f. $f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$, $x > 0$, and $f(x) = 0$ for $x \le 0$; $a_k = k!/\lambda^k$, $k \in \mathbb{N}$; $\mu_2 = 1/\lambda^2$, $\mu_3 = 2/\lambda^3$, $\mu_4 = 9/\lambda^4$, $\mu_5 = 44/\lambda^5$, $\mu_6 = 265/\lambda^6$, $\mu_7 = 1854/\lambda^7$ and $\mu_8 = 14833/\lambda^8$. As the expressions (18) for central moments of S^2 and corresponding ODE (11) with coefficients defined by (12) are very complicated, we provide some numerical examples for $\lambda = 2$.

n	P.d.f. f_{S^2}	Gamma approximation
10	$\frac{(x-0.0472)^{0.0732}}{0.0096\,(x+1.2916)^{9.1578}}$	$\frac{38289.70498\,x^{7/2}}{\exp(18x)}$
20	$\frac{(x-0.0332)^{1.8364}}{0.1705(x+0.5983)^{12.0974}}$	$\frac{8.537508833\,10^9 x^{17/2}}{\exp(38x)}$
100	$\frac{\exp[43.8510\arctan(9.0448x+0.1796)]}{2.399993489\,10^{43}(x^2+0.0397x+0.0126)^{9.9830}}$	$\frac{5.585061585\,10^{51} x^{97/2}}{\exp(198x)}$
500	$\frac{\exp[108.2066\arctan(7.4426x-0.2367)]}{2.289029132\,10^{87}(x^2-0.0636x+0.0191)^{34.3156}}$	$\frac{2.342833211\,10^{259} x^{497/2}}{\exp(998x)}$

Table 1: P.d.f.'s of S^2 in exponential case for $\lambda = 2$

For $n \ge 54$, the quadratic function $b_0 + b_1(x-a) + b_2(x-a)^2$ has negative discriminant D and the solution of ODE (11) is of Pearson IV type, otherwise ($D > 0$, $n < 53$) we have curve of type Pearson VI. For $n \ge 100$ there are numbers of order grater than 10^{40}, thus the sequence of operations is important for the numerical precision and derivations of normalizing constants. The calculations are summarized in Table 1, where we also give standard gamma approximation (χ^2 scaled) by (20) to p.d.f. of S^2.

The shapes of graphics and approximations via gamma distributions are shown on Figures 1 and 2. Visually comparing note, that for $n \ge 100$ gamma

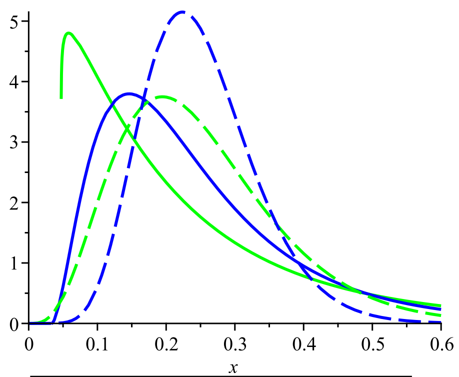

Figure 1: f_{S^2}, $n = 10, 20$ and χ^2-approximation

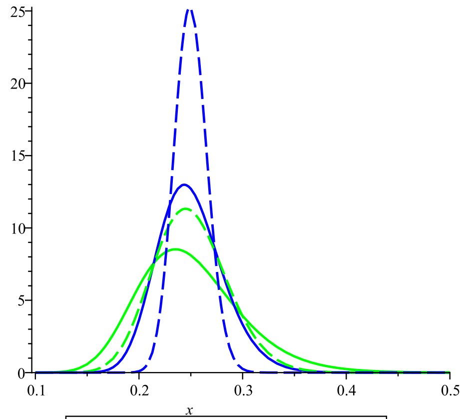

Figure 2: f_{S^2}, $n = 100, 500$ and χ^2-approximation

approximation is fitted better to p.d.f. of the sample variance, although the mode of S^2 is less than mode of scaled χ^2 and corresponding values of maxima are in the same relation.

4.2. Discrete Population

A system of discrete Pearson distributions is used, when the r.v. Y takes on values x_j with probability f_j, and instead of the ODE (11), a difference equation of the form is applied:

$$\frac{f_j - f_{j-1}}{x_j - x_{j-1}} = \frac{P_k(y_j) f_{j-1}}{Q_l(y_j)}, \quad y_j = x_j - a, \quad x_j \in T \subseteq \mathbb{R}, \tag{21}$$

where P_k and Q_l are polynomials of degrees k and l, respectively, centered about the mean $EY = a$. This difference equation depends on the values x_j. Most discrete systems are developed for integers x_j or equally-spaced values, $x_j = x_0 + jh$, $x_0 \in \mathbb{R}$, $h > 0$, for example see [1], [8].

Let X be a r.v. binomially distributed with parameters $n \in \mathbb{N}$ and $p \in (0,1)$, $X \sim Bi(n,p)$, i.e. $P\{X = j\} = C_n^j p^j (1-p)^{n-j} = f_j$, $j = 0, 1, 2, \ldots, n$. For $j = 1, 2, \ldots, n$ we consider the following difference equation

$$f_j - f_{j-1} = \frac{(c_0 + y_{j-1}) f_{j-1}}{b_0 + b_1 y_j + b_2 y_j y_{j-1}}, \tag{22}$$

where: $y_0 = -a$, $y_j = j - a$, $a = np$. Solving system

$$f_j/f_{j-1} - 1 = (c_0 + y_{j-1})/(b_0 + b_1 y_j + b_2 y_{j-1} y_j)$$

we obtain: $c_0 = 1 - p$, $b_0 = -(1-p)np$, $b_1 = -c_0$, $b_2 = 0$.

Distribution of $\bar{X}$ from Bernoulli distributed population. Let r.v.'s X_j, $j = 1, 2, \dots, n$, be i.i.d. via Bernoulli law with parameter p, $p \in (0,1)$, $X_j \sim Be(p)$, $P\{X_j = 0\} = 1 - p$ and $P\{X_j = 1\} = p$. From (8) we find, that $\varphi_{\bar{X}}(t) = (1 - p + e^{it/n})^n$, and hence the distribution of the r.v. $\bar{X}$ is similar to $Bi(n,p)$, and therefore for $j = 1, 2, \dots, n$ satisfies a difference equation

$$\frac{f_j - f_{j-1}}{1/n} = \frac{(c_0 + y_{j-1}) f_{j-1}}{b_0 + b_1 y_j}\,, \quad y_0 = -p\,, y_j = j/n - p\,, \tag{23}$$

with probability mass function (p.m.f.) $f_j = P\{\bar{X} = j/n\} = C_n^j p^j (1-p)^{n-j}$. We have, by Theorem 1: $E\bar{X} = p$, $\mu_{2,\bar{X}} = p(1-p)/n$ and $\mu_{3,\bar{X}} = p(1-p)(1-2p)/n^2$. The coefficients of (23) are: $c_0 = (1-p)/n = -b_1$, $b_0 = -\mu_{2,\bar{X}}$, which slightly differ from continuous analogue (13).

Distribution of S^2 from Bernoulli distributed population Let r.v.'s X_j, $j = 1, 2, \dots, n$, be independent Bernoulli distributed, $X_j \sim Be(p)$, $p \in (0,1)$. The

n	p	Distribution table					χ^2 approximation
4	0.7	S^2	0	1/4	1/3		$\frac{2.072964894\sqrt{x}}{e^{1.5x}}$
		f_j	0.2482	0.4872	0.2646		
5	0.7	S^2	0	1/5	3/10		$4xe^{-2x}$
		f_j	0.1705	0.3885	0.441		
6	0.4	S^2	0	1/6	4/15	3/10	$\frac{7.433850479x^{3/2}}{e^{2.5x}}$
		f_j	0.0533	0.2080	0.4880	0.2507	
7	0.4	S^2	0	1/7	5/21	2/7	$13.5x^2e^{-3x}$
		f_j	0.0312	0.1400	0.3528	0.4760	

Table 2: Distribution table for S^2 in Bernoulli case

sample variance takes on values $x_j = j(n-j)/(n(n-1))$ with probabilities $f_j = P(S^2 = x_j)$, $j = 0, 1, \dots, [n/2]$, where $[n/2]$ is the integer part of positive number $n/2$. Using the fact that it is a probability mass function, i.e. $\sum f_j = 1$ and $\sum x_j^l f_j = a_{l,S^2}$, where a_{l,S^2} are initial moments of S^2 we can obtain their distributions tables, at most for $n = 9$. The method proposed in [1] is impossible to apply because values x_j are not equally-spaced. The p.m.f. of S^2 is numerically obtained for $n = 4, 5, 6, 7$.

The obtained results are presented in Table 2, where we give their standard approximation by χ^2 continuous distribution via formula (20) with $\sigma^2 = 1$. On Figures 3 and 4 are shown p.m.f.'s of S^2 and corresponding χ^2 approximations. Although $n = 7$ is small, p.m.f. of S^2 is enough closed to corresponding χ^2 approximation.

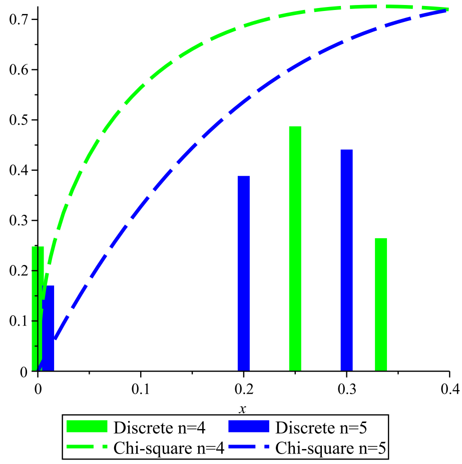

Figure 3: P.m.f. for $p = 0.7$ and χ^2-approximation

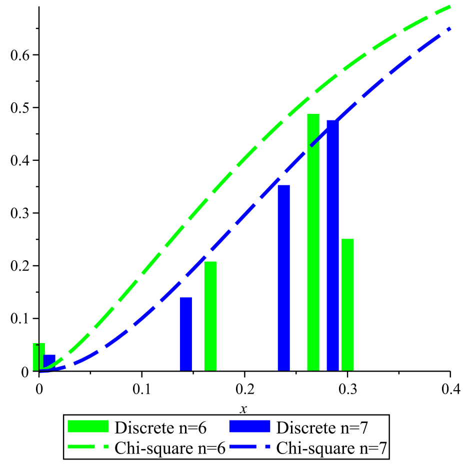

Figure 4: P.m.f. for $p = 0.4$ and χ^2-approximation

5. Applications and Conclusion

Obtained results can be used for:

- estimating population parameters, when the d.f. is known,
- deriving p.d.f's of the sample mean or/and sample variance in analytic form via Pearson system, when they are impossible to derive in closed expression by differential or integral transforms,
- finding quantiles and constructing interval estimators,
- determining skewness, kurtosis or others parameters,

- performing confirmatory analysis, based on $\bar{X}$, S^2 or some coefficients of them, and etc.

By example, standardized r.v. $T = \frac{X-\bar{X}}{S}$, where $X, X_1, \ldots, X_n$ are i.i.d., are functions of X and statistics $\bar{X}$ and S^2, and is often used to provide Student's t-tests or other hypothesis testing and computing confidence intervals. Therefore we have to know d.f. of T. By first and second order Taylor series expansion at the point (a, a, μ_2) we obtain linear and second order approximations to T, namely: $\frac{X-\bar{X}}{S} \sim \frac{X-\bar{X}}{\sqrt{\mu}_2} = T_1$, $\frac{X-\bar{X}}{S} \sim \frac{X-\bar{X}}{2\sqrt{\mu}_2}\left(3 - \frac{S^2}{\mu_2}\right) = T_2$. The same approximations are obtained by binomial series expansion for degree $-n/2$ of $\frac{X-\bar{X}}{\sqrt{\mu_2}} \cdot [1 + (S^2/\mu_2 - 1))]^{-1/2}$.

The coefficient of variation $CV = S/\bar{X}$ is also common and applied in queueing theory, reliability analysis and others, where the exponential distribution is more important than the normal distribution. By first order Taylor expansion at the point (a, μ_2) to CV and geometric series representation for $\frac{S}{a} \cdot \frac{1}{1+(\bar{X}/a-1)}$ we have respectively: $S/\bar{X} \sim \frac{\sqrt{\mu_2}}{a}\left(\frac{3}{2} - \frac{\bar{X}}{a} + \frac{S^2}{2\mu_2}\right) = CV_1$ and $S/\bar{X} \sim S/a = CV_2$.

If we know, from a-priori information, a or μ_2, we can find exactly or via Person systems d.f.'s of statistics T_1 or CV_2, respectively. Otherwise, substituting exact moments by their sample analogues as moments, maximum likelihood, unbiased (as proposed in [11]) or etc estimators, applying (11) we can receive their p.d.f.'s. As concerns T_2, CV_1 or other complicated statistics, product moments $E\bar{X}^m(S^2)^l$ of higher order (m, l) are needed to proceed with Pearson curves or other approximations.

References

[1] K.O. Bowman, L.R. Shenton, M.A. Kastenbaum, Discrete Pearson distributions, *ORNL Report* (1991).

[2] A.E.R. Church, On the moments of the distribution of squared standard-deviations for samples of N drawn from an indefinitely large population, *Biometrika*, **17**, No. 1/2 (1925), 79-83.

[3] H. Cramer, *Mathematical Methods of Statistics*, 19th edn., Princeton University Press (1999).

[4] L.E. Dickson, *History of the Theory of Numbers, Volume II: Diophantine Analysis*, Chelsea Publishing Co., New York (1971).

[5] F.N. David, M.G. Kendall, Tables of symmetric functions-Part I, *Biometrika*, **36**, No. 3/4 (1949), 431-449.

[6] C.S. Davis, A new approximation to the distribution of Pearson's chi-square, *Statistica Sinica*, **3** (1993), 189-196.

[7] P.S. Dwyer, D. S. Tracy, A combinatorial method for products of two polykays with some general formulae, *The Annals of Mathematical Statistics*, **35** (1964), 1174-1185.

[8] N.L. Johnson, S. Kotz, A.W. Kemp, *Univariate Discrete Distributions*, 2nd edn, John Wiley & Sons (1992).

[9] M.G. Kendall, A. Stuart, *Theory of Statistics, Volume 1: Distribution Theory*, 2nd edn, Nauka, Moscow (1966).

[10] J. Kinney, The k-statistics, polykays and randomized sums, *Sankhya: The Indian journal of statistics, Series A*, **38**, No. 3 (1976), 271-286.

[11] B. Klemens, *Modeling with Data: Tools and Techniques for Scientific Computing*, Princeton University Press (2009).

[12] D. Knuth, A note on solid partitions, *Mathematics of computation*, **24**, No. 112 (1970), 955-961.

[13] V.S. Koroluk et al, *Handbook of Theory of Probability and Mathematical Statistics*. Nauka, Moscow (1985).

[14] S. Kotz et al, The sample variance for nonnormal population, In *Encyclopedia of Statistical Sciences*, Volume 1, 2nd edn, John Wiley & Sons (2006), 8974-8975.

[15] G.S. Mudholkar, M.C. Trivedi, A Gaussian approximation to the distribution of the sample variance for nonnormal populations, *Journal of the American Statistical Association*, **76**, No. 374 (1981), 479-485.

[16] E. Di Nardo, G. Guarino, D. Senato, A unifying framework for k-statistics, polykays and their multivariate generalizations, *Bernoulli*, **14**, No. 2 (2008), 440-468.

[17] G. Roussas, *Introduction to Probability and Statistical Inference*. Academic Press, Elsevier Science (2003).

[18] J.M. Le Roux, A Study of the distribution of the variance in small samples, *Biometrika*, **23**, No. 1/2 (1931), 134-190.

[19] T. Royen, Exact distribution of the sample variance from a Gamma parent distribution, arXiv:0704.1415v1 [math.ST] (2007).

[20] T.A. Severini, *Elements of Distribution Theory*, Cambridge University Press (2005).

[21] L.R. Shenton, K.O. Bowman, The development of techniques for the evaluation of sampling moments, *International Statistical Review*, **43**, No. 3 (1975), 317-334.

[22] H. Solomon, M.A. Stephens, An approximation to the distribution of the sample variance, *The Canadian Journal of Statistics*, **11**, No. 2 (1983), 149-154.

[23] W.Y. Tan, S. P. Wong, On the Roy-Tiku approximation to the distribution of sample variances from nonnormal universes, *Journal of the American Statistical Association*, **72**, No. 360 (1977), 875-880.

[24] Al. A. Tchouproff, On the mathematical expectation of the moments of frequency distributions, *Biometrika*, **12** (1919), 185-210.

[25] J.W. Tukey, Some sampling simplified, *Journal of the American Statistical Association*, **45**, No. 252 (1950), 501-519.

[26] J.W. Tukey, Keeping moment-like sampling computation simple, *The Annals of Mathematical Statistics*, **27**, No. 1 (1956), 37-54.

[27] G. Vegas-Sánchez-Ferrero et al, A direct calculation of moments of the sample variance, *Mathematics and Computers in Simulation*, **82** (2012), 790-804.

[28] A. Zoghbi, I. Stojmenović, Fast algorithms for generating integer partitions, *International Journal of Computer Mathematics*, **70** (1998), 319-332.

A. Proof of Theorem 2

Using (3) and the recursive formula $(S^2)^{k+1} = (S^2)^k S^2$ or expand formula (2) by degrees of $\sum X_i^2$ an degrees of $(\sum X_i)^2$ and (9) we obtain:

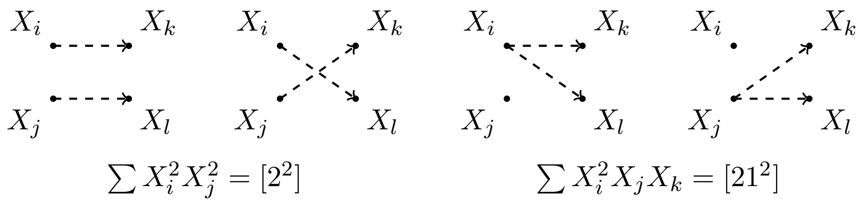

Figure 5: A graph representation of possible combinations of coincidences of indices in product $[1^2][1^2]$

$$\begin{aligned}(S^2)^2 =& n^{-2} \Big[[4] - 4/(n-1)\,[31] + \big(1 + 2/(n-1)^2\big)\,[2^2] \\ & -2\big(1/(n-1) - 2/(n-1)^2\big)\,[21^2] + 1/(n-1)^2\,[1^4]\Big]\ .\end{aligned}$$

During this procedure we have to compute products of power sums as $[1^2][1^2] = (\sum X_i X_j)(\sum X_k X_l) = 2[2^2] + 4[21^2] + [1^4]$, see Figure 5 and etc.

The expressions for first two moments of S^2 for example see in [17], [27].

$$\begin{aligned}(S^2)^3 =& n^{-3} \Big[[6] - 6/(n-1)\,[51] + 3\big(1 + 4/(n-1)^2\big)\,[42] \\ & - 3\big(1/(n-1) - 4/(n-1)^2\big)\,[41^2] - 2\big(3/(n-1) + 2/(n-1)^3\big)\,[3^2] \\ & - 12\big(1/(n-1) - 2/(n-1)^2 + 2/(n-1)^3\big)\,[321] \\ & + 4\big(3/(n-1)^2 - 2/(n-1)^3\big)\,[31^3] \\ & + \big(1 + 6/(n-1)^2 - 8/(n-1)^3\big)\,[2^3] \\ & - 3\big(1/(n-1) - 4/(n-1)^2 + 10/(n-1)^3\big)\,[2^21^2] \\ & +3\big(1/(n-1)^2 - 4/(n-1)^3\big)\,[21^4] - 1/(n-1)^3\,[1^6]\Big]\ , \\ (S^2)^4 =& n^{-4} \Big[[8] - 8/(n-1)\,[71] + 4\big(1 + 6/(n-1)^2\big)\,[62] \\ & - 4\big(1/(n-1) - 6/(n-1)^2\big)\,[61^2] - 8\big(3/(n-1) + 4/(n-1)^3\big)\,[53] \\ & - 24\big(1/(n-1) - 2/(n-1)^2 + 4/(n-1)^3\big)\,[521] \\ & + 8\big(3/(n-1)^2 - 4/(n-1)^3\big)\,[51^3] \\ & + \big(3 + 24/(n-1)^2 + 8/(n-1)^4\big)\,[4^2] \\ & - 8\big(3/(n-1) - 12/(n-1)^2 + 12/(n-1)^3 - 8/(n-1)^4\big)\,[431] \\ & + 6\big(1 + 10/(n-1)^2 - 16/(n-1)^3 + 8/(n-1)^4\big)\,[42^2] \\ & - 12\big(1/(n-1) - 6/(n-1)^2 + 20/(n-1)^3 - 8/(n-1)^4\big)\,[421^2] \\ & + 2\big(3/(n-1)^2 - 24/(n-1)^3 + 8/(n-1)^4\big)\,[41^4]\end{aligned}$$

$$
\begin{aligned}
&- 8\left(3/(n-1) - 6/(n-1)^2 + 14/(n-1)^3 - 12/(n-1)^4\right)\,[3^2 2] \\
&- 24\left(1/(n-1) - 4/(n-1)^2 + 14/(n-1)^3 - 16/(n-1)^4\right)\,[32^2 1] \\
&+ 8\left(9/(n-1)^2 - 12/(n-1)^3 + 14/(n-1)^4\right)\,[3^2 1^2] \\
&+ 16\left(3/(n-1)^2 - 14/(n-1)^3 + 18/(n-1)^4\right)\,[321^3] \\
&- 8\left(3/(n-1)^3 - 4/(n-1)^4\right)\,[31^5] \\
&+ \left(1 + 12/(n-1)^2 - 32/(n-1)^3 + 60/(n-1)^4\right)\,[2^4] \\
&- 4\left(1/(n-1) - 6/(n-1)^2 + 30/(n-1)^3 - 68/(n-1)^4\right)\,[2^3 1^2] \\
&+ 6\left(1/(n-1)^2 - 8/(n-1)^3 + 26/(n-1)^4\right)\,[2^2 1^4] \\
&-4\left(1/(n-1)^3 - 6/(n-1)^4\right)\,[21^6] + 1/(n-1)^4\,[1^8]\big]\ .
\end{aligned}
$$

To calculate the expectations we have to determine the number of summands in each sum. For instance, in the sum $\sum X_{i_1}^3 X_{i_2}^2 X_{i_3} X_{i_4} X_{i_5} = [321^3]$ we have $\prod_{j=0}^{4}(n-j)$ terms (see formula (10)), therefore $E([321^3])/n^4/(n-1)^2 = \prod_{j=0}^{4}(n-j)/n^4/(n-1)^2 a_3 a_2 a^3 = (1/n - 8/n^2 + 18/n^3 - 6/n^3/(n-1))a_3 a_2 a^3$ and etc.

Finally, expanding polynomials $n(n-1)$, $n(n-1)(n-2)$,$\dots$, $\prod_{j=0}^{7}(n-j)$ by powers of $n-1$ and performing division by $n^4(n-1)^j$, $j = 0, 1, \dots, 4$, we obtain:

$$
\begin{aligned}
a_{1,S^2} =& a_2 - a^2\,, \\
a_{2,S^2} =& a_2^2 - 2a_2a^2 + a^4 + (a_4 - 4a_3a - a_2^2 + 8a_2a^2 - 4a^4)/n \\
&+ 2(a_2^2 - 2a_2a^2 + a^4)/n/(n-1)\,, \\
a_{3,S^2} =& a_2^3 - 3a_2^2a^2 + 3a_2a^4 - a^6 \\
&+ 3(a_4a_2 - a_4a^2 - 4a_3a_2a + 4a_3a^3 - a_2^3 + 9a_2^2a^2 - 12a_2a^4 + 4a^6)/n \\
&+ (a_6 - 6a_5a - 3a_4a_2 + 18a_4a^2 - 6a_3^2 + 48a_3a_2a - 56a_3a^3 + 8a_2^3 \\
&- 96a_2^2a^2 + 138a_2a^4 - 46a^6)/n^2 \\
&+ 2(6a_4a_2 - 6a_4a^2 - 24a_3a_2a + 24a_3a^3 - 7a_2^3 + 57a_2^2a^2 \\
&- 75a_2a^4 + 25a^6)/n^2/(n-1) \\
&- 4(a_3^2 - 6a_3a_2a + 4a_3a^3 - 2a_2^3 + 15a_2^2a^2 - 18a_2a^4 + 6a^6)/n^2/(n-1)^2\,, \\
a_{4,S^2} =& a_2^4 - 4a_2^3a^2 + 6a_2^2a^4 - 4a_2a^6 + a^8 \\
&+ 6\left(a_4a_2^2 - 2a_4a_2a^2 + a_4a^4 - 4a_3a_2^2a + 8a_3a_2a^3 - 4a_3a^5 - a_2^4\right. \\
&\left.+10a_2^3a^2 - 21a_2^2a^4 + 16a_2a^6 - 4a^8\right)/n \\
&+ \left(4a_6a_2 - 4a_6a^2 - 24a_5a_2a + 24a_5a^3 + 3a_4^2 - 24a_4a_3a - 18a_4a_2^2\right.
\end{aligned}
$$

$$
\begin{aligned}
&+ 132a_4a_2a^2 - 96a_4a^4 - 24a_3^2a_2 + 216a_3a_2^2a + 72a_3^2a^2 - 608a_3a_2a^3 \\
&+320a_3a^5 + 23a_2^4 - 416a_2^3a^2 + 1080a_2^2a^4 - 880a_2a^6 + 220a^8\big) /n^2 \\
&+ \big(a_8 - 8a_7a - 4a_6a_2 + 32a_6a^2 - 24a_5a_3 + 96a_5a_2a - 128a_5a^3 - 3a_4^2 \\
&+ 144a_4a_3a + 72a_4a_2^2 - 600a_4a_2a^2 + 460a_4a^4 + 96a_3^2a_2 - 864a_3a_2^2a \\
&- 384a_3^2a^2 + 2720a_3a_2a^3 - 1456a_3a^5 - 86a_2^4 + 1640a_2^3a^2 - 4500a_2^2a^4 \\
&+3728a_2a^6 - 932a^8\big) /n^3 \\
&+ 4 \big(6a_6a_2 - 6a_6a^2 - 36a_5a_2a + 36a_5a^3 + 6a_4^2 - 48a_4a_3a - 39a_4a_2^2 \\
&+ 240a_4a_2a^2 - 165a_4a^4 - 40a_3^2a_2 + 396a_3a_2^2a + 136a_3^2a^2 - 1120a_3a_2a^3 \\
&+580a_3a^5 + 45a_2^4 - 774a_2^3a^2 + 2001a_2^2a^4 - 1624a_2a^6 + 406a^8\big) /n^3/(n-1) \\
&+ 4 \big(-8a_5a_3 + 24a_5a_2a - 16a_5a^3 + 40a_4a_3a + 36a_4a_2^2 - 192a_4a_2a^2 \\
&+ 116a_4a^4 + 52a_3^2a_2 - 456a_3a_2^2a - 132a_3^2a^2 + 1128a_3a_2a^3 - 544a_3a^5 \\
&-61a_2^4 + 928a_2^3a^2 - 2238a_2^2a^4 + 1764a_2a^6 - 441a^8\big) /n^3/(n-1)^2 \\
&+ 8 \big(a_4^2 - 8a_4a_3a - 6a_4a_2^2 + 24a_4a_2a^2 - 12a_4a^4 - 12a_3^2a_2 + 96a_3a_2^2a \\
&+ 28a_3^2a^2 - 216a_3a_2a^3 + 96a_3a^5 + 15a_2^4 - 204a_2^3a^2 \\
&+468a_2^2a^4 - 360a_2a^6 + 90a^8\big) /n^3/(n-1)^3 \,.
\end{aligned}
$$

Combining above formulas and (6) we obtain (18). By (5) follows (19) - the second statement of the theorem.