Commun.Fac.Sci.Univ.Ank.Series A1
Volume 54, Number 1, Pages 21- 28(2005)
ISSN 1303-5991

AN EASY WAY TO CALCULATE THE MOMENTS OF SAMPLE CORRELATION COEFFICIENTS

EL-HOUSSAINY A. RADY, HALA A. FERGANY AND ADEL M. EDRESS

ABSTRACT. The purpose of this paper is to introduce an easy way to calculate the general moment of the sample correlation coefficients. A straightforward method of doing this is introduced, and applied to obtain more moments, and more terms, than those have been introduced by Ghosh (1966) and Hotelling (1953). Moreover we deduce the first six moments, skewness coefficient and kurtosis coefficient.

1. INTRODUCTION

Let $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$ be a random sampling from a bivariate normal distribution with standard deviations σ_x, σ_y and a correlation coefficient ρ. The sample standard deviations are denoted by s_x and s_y in the sense that $n\, s_x^2$ and $n\, s_y^2$ $(n = N - 1)$ are the sums of squares of the deviations of x and y respectively from their sample means, and the sample correlation coefficient is denoted by r. The maximum likelihood estimate of the correlation coefficients is

$$r = \frac{s_{xy}}{s_x s_y}, \tag{1}$$

where $s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x - \overline{x})(y - \overline{y})$.

Fisher (1915) has obtained the basic distribution of the correlation coefficients r. The probability density function as introduced by Hotelling (1953) in a form of hypergeometric function, is given by the following equation

$$f_n(r, \rho) = A(1 - \rho^2)^{n/2}(1 - r)^{\frac{n-3}{2}}(1 - \rho r)^{-n+\frac{1}{2}}\, {}_2F_1(\frac{1}{2}, \frac{1}{2}, n + \frac{1}{2}, \frac{1 + \rho r}{2}), \tag{2}$$

where $A = \dfrac{(n-1)\Gamma(n)}{\sqrt{2\pi}\Gamma(n + \frac{1}{2})}$ and ${}_2F_1$ is the hypergeometric series.

Bell (1967) showed the following properties of the hypergeometric function:

$${}_2F_1(a, b, c, t) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\sum_{i=0}^{\infty}\frac{\Gamma(a+i)\Gamma(b+i)}{\Gamma(c+i)}\frac{t^i}{i!} \tag{3}$$

$$\begin{aligned}{}_2F_1(a, b, c, t) = \quad & 1 + \frac{ab}{c}\frac{t}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{t^2}{2!} \\ & + \frac{a(a+1)(a+2)b(b+1)(b+2)}{c(c+1)(c+2)}\frac{t^3}{3!} + \cdots\end{aligned} \tag{4}$$

Received by the editors June 25, 2004, Revised: Nov. 18, 2004, Accepted: April 04, 2005.
Key words and phrases. Moments, Correlation, Hypergeometric function.

$$
{}_2F_1(a,b,c,t) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 z^{c-1}(1,z)^{b-c-1}(1-tz)^{-a} dz. \qquad (5)
$$

This series is convergent if $|t| < 1$ and divergent if $|t| > 1$. For $t = 1$ the series converges if $c > a + b$, while for $t = -1$ the series converges if $c > a + b - 1$.

Some important properties for the distribution of r were given in the co-operative study (1917).

Ghosh (1966) gave the first four moments of r as a power series of order $(n+1)^{-1}$ to order $(n+1)^{-7}$. In Hotelling (1953) method, it was required to know all the $(n-1)^{th}$ moments to calculate the n^{th} moment. In the recent paper a direct method to get the n^{th} moment without knowing information about other moments.

2. General Moment

We shall deal with the moments of r around ρ instead of zero. The following theorem illustrate the h moment of the sample correlatiion coefficient of r about ρ.

Theorem. The h moment of the sample correlation coefficient of r about ρ in a series of powers n^{-1} can be given by

$$
\begin{aligned}
E[(r-\rho)^h] &= (1-\rho^2)^h \Big[\alpha_{h,0} + \frac{1}{8n}\alpha_{h,1} + \frac{1}{128n^2}\{9\alpha_{h,2} - 8\alpha_{h,1}\} \\
&\quad + \frac{1}{1024n^3}\{75\alpha_{h,3} - 144\alpha_{h,2} + 32\alpha_{h,1}\} + \ldots\Big],
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_{h,0} &= \frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{n+h}{2})} A[1 + \frac{1}{8n}(2h+1)(2h-1)(h+1)\rho^2 \\
&\quad + \frac{(2h+1)(2h-1)(h+1)\rho^2}{3(128)n^2}\{(2h+3)(2h+5)(h+3)\rho^2 - 48h\} \\
&\quad + \frac{(2h+1)(2h-1)(h+1)\rho^2}{15(1024)n^3}\{\rho^2(2h+3)(2h+5)(h+3)\{(2h+5) \\
&\quad \times(2h+7)(h+5)\rho^2 - 80(h+1)\} + (15)(128)h^2\} + \ldots] \\
\alpha_{h,1} &= \frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{n+h}{2})} A[(1+\rho^2) + \frac{1}{8n}(2h+1)(h+1)\rho^2\{(1+\rho^2)(2h-1) \\
&\quad -4(1-\rho^2)\} + \frac{(2h+1)(h+1)\rho^2}{3(128)n^2}\{\rho^2[(1+\rho^2)(2h-1) - 8(1-\rho^2)] \\
&\quad \times[(2h+3)(2h+5)(h+3)] + 24h(1-\rho^2) - 48h(2h-1)(1+\rho^2)\} + \ldots] \\
\alpha_{h,2} &= \frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{n+h}{2})} A(1+\rho^2)[(1+\rho^2) + \frac{1}{8n}(2h+1)(h+1)\rho^2 \\
&\quad \times\{(1+\rho^2)(2h-1) - 8(1-\rho^2)\} + \ldots] \\
\alpha_{h,3} &= \frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{n+h}{2})} A[(1+\rho^2)^3 + \ldots].
\end{aligned}
$$

Proof. Let us define

$$\lambda_h = E[(r-\rho)^h] = \int_{-1}^{1} f_n(r,\rho)(r-\rho)^h dr. \tag{6}$$

By substituting Equation (2) in Equation (6) we get

$$\begin{aligned}\lambda_h = A(1-\rho^2)^{n/2} \textstyle\int_{-1}^{1} (r-\rho r)^h (1-r^2)^{\frac{n-3}{2}} (1-\rho r)^{-n+\frac{1}{2}} \\ \times {}_2F_1(\tfrac{1}{2}, \tfrac{1}{2}, n+\tfrac{1}{2}, \tfrac{1+\rho r}{2}) dr.\end{aligned} \tag{7}$$

Using Equations (4) and (7), we have

$$\begin{aligned}\lambda_h &= (1-\rho^2)^h [\alpha_{h,0} + \frac{1}{8(n+1/2)}\alpha_{h,1} + \frac{9}{128(n+1/2)(n+3/2)}\alpha_{h,2} \\ &+ \frac{75}{1024(n+1/2)(n+3/2)(n+5/2)}\alpha_{h,3} + ...],\end{aligned} \tag{8}$$

where

$$\alpha_{h,k} = A(1-\rho^2)^{\frac{n-2h}{2}} \int_{-1}^{1} (1-r^2)^{\frac{n-3}{2}} (r-\rho)^k (1-\rho r)^{-n+1/2} (1+\rho r)^k dr. \tag{9}$$

Using the binomial theorem the term $\frac{1}{n+1/2}$ can be expanded as

$$\frac{1}{n+1/2} = \frac{1}{n}(1+\frac{1}{2n})^{-1} = \frac{1}{n}(1-\frac{1}{2n}+\frac{1}{4n^2}+...). \tag{10}$$

In a same way, one can obtain

$$\frac{1}{(n+1/2)(n+3/2)} = \frac{1}{n^2}(1+\frac{2}{n}+...) \tag{11}$$

and

$$\frac{1}{(n+1/2)(n+3/2)(n+5/2)} = \frac{1}{n^3}(1+...). \tag{12}$$

Inserting Equations (10), (11) and (12) in Equation (8) we get

$$\begin{aligned}\lambda_h &= (1-\rho^2)^h [\alpha_{h,0} + \frac{1}{8n}\alpha_{h,1} + \frac{1}{128n^2}(9\alpha_{h,2} - 8\alpha_{h,1}) \\ &+ \frac{1}{1024n^3}(75\alpha_{h,3} - 144\alpha_{h,2} + 32\alpha_{h,1}) + ...].\end{aligned} \tag{13}$$

Substitution of $k+1$ for k in $\alpha_{h,k}$ is equivalent to multiplying the integrand by $1+\rho r$, that is, by

$$1+\rho^2+\rho(r-\rho)$$

which gives

$$\alpha_{h,k+1} = (1+\rho^2)\alpha_{h,k} + \rho(1-\rho^2)\alpha_{h+1,k}\,. \tag{14}$$

Let $k=0$ in Equation (9), then we get

$$\alpha_{h,0} = A(1-\rho^2)^{\frac{n-2h}{2}} \int_{-1}^{1} (1-r^2)^{\frac{n-3}{2}} (r-\rho)^h (1-\rho r)^{-n+1/2} dr. \tag{15}$$

Define

$$r = \frac{\rho+\omega}{1+\rho\omega},\ -1<\omega<1.$$

Then substituting it in Equation (15), we get

$$\alpha_{h,0} = A\int_{-1}^{1}(1-\omega^2)^{\frac{n-3}{2}}\omega^h(1+\rho\omega)^{\frac{1}{2}-h}d\omega. \tag{16}$$

Taking the partial derivative with respect to ρ, we get

$$\alpha_{h+1,0} = \frac{-2}{2h-1}\frac{\partial\alpha_{h,0}}{\partial\rho}. \tag{17}$$

Using the binomial theorem, then we get

$$\alpha_{h,0} = A\sum_{j=0}^{\infty}\frac{\Gamma(\frac{3}{2})\rho^{2j}}{2j!\Gamma(\frac{3}{2}-2j)}\int_{-1}^{1}(1-\omega^2)^{\frac{n-3}{2}}(\omega^2)^{j+\frac{h}{2}}d\omega.$$

Let $u = \omega^2$,

$$\begin{aligned}
\alpha_{h,0} &= A\sum_{j=0}^{\infty}\frac{\Gamma(\frac{3}{2})\rho^{2j}}{2j!\Gamma(\frac{3}{2}-2j)}\int_{-1}^{1}(1-u)^{\frac{n-3}{2}}u^{j-\frac{1}{2}}du \\
&= A\sum_{j=0}^{\infty}\frac{\rho^{2j}\,\Gamma(\frac{3}{2})\Gamma(\frac{n-1}{2})\Gamma(j+\frac{1}{2})}{2j!\Gamma(\frac{3}{2}-2j)\Gamma(\frac{n}{2}+j)} \\
&= A\frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(\frac{n+h}{2})}\sum_{j=0}^{\infty}\frac{(\frac{2h-1}{4})_j\,(\frac{2h+1}{4})_j\,(\frac{h+1}{2})_j}{(\frac{1}{2})_j(\frac{n+h}{2})_j}\frac{\rho^{2j}}{j!},
\end{aligned}$$

then

$$\alpha_{h,0} = A'\,{}_3F_2(\frac{2h-1}{4},\frac{2h+1}{4},\frac{h+1}{2};\frac{1}{2}\frac{n+h}{2},\rho^2), \tag{18}$$

where $A' = \dfrac{\Gamma(\frac{n-1}{2})\Gamma(\frac{h+1}{2})}{\Gamma(n-1)\Gamma(\frac{n+h}{2})}$. After expanding Equation (18), we get

$$\begin{aligned}
\alpha_{h,0} &= A'[1 + \tfrac{1}{8(n+h)}(2h+1)(2h-1)(h+1)\rho^2 \\
&+ \frac{(2h+1)(2h-1)(h+1)}{3(128)(n+h)(n+h+2)}(2h+3)(2h+5)(h+3)\rho^4 \\
&+ \frac{(2h+1)(2h-1)(h+1)}{15(1024)(n+h)(n+h+2)(n+h+4)}(2h+3)(2h+5) \\
&\times(h+3)(2h+5)(2h+7)(h+5)\rho^6 + ...] \\
&= A[1 + \frac{1}{8n}(2h+1)(2h-1)(h+1)\rho^2 \\
&+ \frac{(2h+1)(2h-1)(h+1)\rho^2}{3(128)n^2}\{(2h+3)(2h+5)(h+3)\rho^2 - 48h\} \\
&+ \frac{(2h+1)(2h-1)(h+1)\rho^2}{15(1024)n^3}\{\rho^2(2h+3)(2h+5)(h+3) \\
&\times\{(2h+5)(2h+7)(h+5)\rho^2 - 80(h+1)\} + (15)(128)h^2\} + ...]
\end{aligned} \tag{19}$$

by using Equations (17) and (19), we get

$$\alpha_{h+1,0} = -A'\frac{(2h+1)(h+1)\rho}{2(n+h)}\,{}_3F_2(\frac{2h+3}{4}, \frac{2h+5}{4}, \frac{h+3}{2}; \frac{3}{2}\frac{n+h+2}{2}, \rho^2)$$

which gives

$$\begin{aligned}
\alpha_{h+1,0} &= A'[\tfrac{-1}{2n}(2h+1)(h+1)\rho - \frac{1}{48n^2}\{(2h+1)(2h+3)(2h+5) \\
&\times(h+1)(h+3)\rho^3 - 24(2h+1)h(h+1)\rho + \frac{(2h+1)(h+1)}{15(1024)n^3} \\
&\times\{(2h+3)(2h+5)(h+3)\{6(2h+5)(2h+7)(h+5)\rho^5 \\
&-320(2h+1)\rho^3\} + (30)(128)h^2\rho\} + ...] \\
&= A'(2h+1)(h+1)\rho[\frac{-1}{2n} - \frac{1}{48n^2}\{(2h+3)(2h+5)(h+3)\rho^2 \\
&-24h + \frac{1}{15(1024)n^3}\{(2h+3)(2h+5)(h+3)\{6(2h+5) \\
&\times(2h+7)(h+5)\rho^4 - 320(2h+1)\rho^2\} + (30)(128)h^2\} + ...].
\end{aligned} \tag{20}$$

By using Equations (14) and (20), we get

$$\begin{aligned}\alpha_{h,1} = \quad & A'[(1+\rho^2) + \frac{1}{8n}(2h+1)(h+1)\rho^2\{(1+\rho^2)(2h-1) - 4(1-\rho^2)\} \\ & + \frac{(2h+1)(h+1)\rho^2}{3(128)n^2}\{\rho^2[(1+\rho^2)(2h-1) - 8(1-\rho^2)][(2h+3) \\ & \times(2h+5)(h+3)] + 24h(1-\rho^2) - 48h(2h-1)(1+\rho^2)\} + ...],\end{aligned} \tag{21}$$

$$\begin{aligned}\alpha_{h+1,1} &= A'[\frac{\rho(h+1)}{2n}\{-(1+\rho^2)(2h+1) + 2(1-\rho^2)\} \\ & + \frac{\rho(h+1)}{96n^2}\{-24h(1-\rho^2) + \rho^2(2h+3)(2h+5)(h+3) \\ & \times\{3(1-\rho^2) - 2(1+\rho^2)(2h+1)\}\} + ...], \\ \alpha_{h+2,1} &= A'[\frac{(1+\rho^2)(h+1)}{n} + ...], \\ \alpha_{h+1,2} &= A'[\frac{\rho(1+\rho^2)(h+1)}{n}\{-(1+\rho^2)(2h+1) + 4(1-\rho^2)\} + ...],\end{aligned}$$

$$\begin{aligned}\alpha_{h,2} = \quad & A'(1+\rho^2)[(1+\rho^2) + \frac{1}{8n}(2h+1)(h+1)\rho^2 \\ & \times\{(1+\rho^2)(2h-1) - 8(1-\rho^2)\} + ...],\end{aligned} \tag{22}$$

$$\alpha_{h,3} = A'[(1+\rho^2)^3 + ...]. \tag{23}$$

3. The First Six Moments

From the previous results one can deduce

$$\begin{aligned}\lambda_1 &= (1-\rho^2)[-\frac{\rho}{2n} + \frac{\rho - 9\rho^3}{8n^2} + \frac{\rho + 42\rho^3 - 75\rho^5}{16n^3} + ...] \\ \lambda_2 &= (1-\rho^2)^2[\frac{1}{n} + \frac{23\rho^2}{4n^2} + \frac{-97\rho^2 + 309\rho^4}{8n^3} + ...] \\ \lambda_3 &= (1-\rho^2)^3[-\frac{15\rho}{2n^2} + \frac{123\rho - 798\rho^3}{8n^3} + ...] \\ \lambda_4 &= (1-\rho^2)^4[-\frac{3}{n^2} + \frac{-12 + 237\rho^2}{2n^3} + ...] \\ \lambda_5 &= (1-\rho^2)^5[-\frac{-135\rho}{2n^3} + ...] \\ \lambda_6 &= (1-\rho^2)^6[\frac{15}{n^3} + ...].\end{aligned}$$

For the moments of r about its mean λ_1, we find

$$\begin{aligned}
\mu_2 &= (1-\rho^2)^2[\frac{1}{n}+\frac{11\rho^2}{2n^2}+\frac{-24\rho^2+75\rho^4}{2n^3}+...] \\
\mu_3 &= (1-\rho^2)^3[-\frac{6\rho}{n^2}+\frac{15\rho-88\rho^3}{n^3}+...] \\
\mu_4 &= (1-\rho^2)^4[\frac{3}{n^2}+\frac{-6+105\rho^2}{n^3}+...] \\
\mu_5 &= (1-\rho^2)^5[-\frac{60\rho}{n^3}+...] \\
\mu_6 &= (1-\rho^2)^6[\frac{15}{n^3}+...].
\end{aligned}$$

Taking the square root of μ_2, the variance of r, σ_r is obtained as

$$\sigma_r = (1-\rho^2)[\frac{1}{n}+\frac{11\rho^2}{2n^2}+\frac{-24\rho^2+75\rho^4}{2n^3}+...]^{1/2}.$$

While skewness and kurtosis are indicated by

$$\begin{aligned}
\sqrt{\beta_1} &= 6\rho n^{-\frac{1}{2}}[1+\frac{-30+77\rho^2}{12n}+...] \\
\beta_2-3 &= \frac{6(12\rho^2-1)}{n}+...
\end{aligned}$$

4. Conclusion

In each of the previous equations of λ_h and μ_h the row of dots stands for a function of higher order in n^{-1} than the term next preceeding; and the series are asymptotic but presumably divergent.

The stepwise nature of the sets of equations for λ_h and μ_h, with the order in n^{-1} increasing by unity on alternate occasions as h increases, and with the coefficients of the leading terms of the even moments proportional to the even moments of a normal distribution, is typical of skew distributions of statistics approaching normality for large samples. It provides in fact one of the proofs of asymptotic normality of a large class of statistics including the correlation coefficient.

ÖZET:Bu çalışmanın amacı, örnek korelasyon katsayısının momentlerini hesaplamak için kolay bir hesaplama yolu sunmaktır. Bunu yapmak için direkt bir yöntem tanımlanmış ve Ghosh (1966) ve Hotelling (1953) in elde ettiklerinden daha çok sayıda terim içeren yüksek dereceden momentler elde etmek için kullanılmıştır. Ayrıca, ilk altı moment, çarpıklık ve basıklık katsayılarına ulaşılmıştır.

References

[1] A. Zellner, (1963), Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results, Journal of the American Statistical Association, Vol. 58, pp. 977-992.

[2] H. Hotelling, (1953), New Light on the Correlation Coefficient and Its Transforms, J. Roy. Stat. Soc., B 15, 193-225.

[3] H. E. Soper, A. W. Young, B. M. Cave, A. Lee, and K. Pearson, (1916), On the Distribution of the Correlation Coefficient in Small Sample, Appendix II to the Papers of Student and R. A. Fisher. "A Cooperative Study", Biometrika vol 11 pp 328-413. (Cited as the Cooperative Study.)

[4] I. S. Gradshteyn, and Ryzhik I. M., (1980), Table of Integrals, Series, and Products. Academic Press, New York.
[5] R. A. Fisher, (1915), Frequency Distribution of the Values of the Correlation Coefficient Sample From Indefinitely Large Population, Biometrika vol 10 pp 507-521.
[6] Stwrt, A. and Ord, K, (1994), Kendall's Advanced Theory of Statistics, Vol 1, Distribution Theory, Edward Arrold, London.
[7] T. W. Anderson, (1984), An Introduction to Multivariate Statistical Analysis. Wiley Eastern private limited, New Delhi.
[8] W. W. Bell, (1968), Special Functions for Scientists and Engineers, D. Van Nostrand Company LTD, London.
[9] X. Rong Li, (1999), Probability, Random Signals, and Statistics, CRC Press, New York.

Current address: El-Houssainy A. Rady.Institute of Statistical Studies and Research Cairo University., Hala A. Fergany. Department of Mathematics, Faculty of Science, Tanta University, Adel M. Edress. Department of Mathematics, Faculty of Science, Tanta University