ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 141, 1997

*Andrzej Czajkowski, Dariusz Parys**

THE DISTRIBUTION AND PARAMETERS OF THE DISTRIBUTION OF THE QUOTIENT OF RANDOM VARIABLES

Abstract. This paper presents some important earlier results of the research concerning the properties of the distribution of the quotient of random variables.

We present also our own ideas and results of the research concerning the mean and the variance of the distribution of the quotient of random quadratic forms.

Key words: quotient of random variables, random quadratic forms, moment generating function.

***

In this paper we present the most important results concerning the distribution and the parameters of the distribution of the quotient of random variables.

Our results concerning the power of the random quadratic forms are also outlined. The research aimed at fixing the form of the distribution of the quotient of random variables and what had been done so far was assessed as insuficient. Most of the basic results were achieved in the middle of this century and the number of papers published in recent years is relatively small.

The main reason for this situation are difficulties which arise during estimating the density function (or distribution function) of a given distribution or its basic parameters (mainly mean and variance). Formulated theorems require assumptions concerning random variables in the numerator and the denominator (about the form of the distributions or independence of variables).

Let us consider random variable $Z = \frac{Y}{X}$ where X, Y are random variables and $X \neq 0$. It is well known that the density function of variable Z is of the form:

* University of Łódź, Chair of Statistical Methods.

$$g(z) = \int_{-\infty}^{+\infty} f(x, zx)/x/dx \tag{1}$$

where $f(x, zx)$ is the density function of bivariate random variable (X, Y) for $X = X$ and $Y = ZX$.

If random variables X and $Y = ZX$ are independent then:

$$g(z) = \int_{-\infty}^{+\infty} |x| f_1(x) f_2(xz) dx \tag{1'}$$

where $f_1(x)$ and $f_2(xz)$ are the density functions of X and Y, respectively.

Assume that random variables X and Y are independent and both have normal distributions with mean equal 0 and variance equal 1.

Because the distribution is symmetric the density function has the from:

$$g(z) = 2\int_{0}^{+\infty} xf(x, xz)dx = \frac{1}{\Pi\sigma_x\sigma_y}\int_{0}^{+\infty} x \exp\left[-\frac{1}{2}\frac{\sigma_y^2 + \sigma_x^2 z^2}{\sigma_x^2\sigma_y^2}\right]dz =$$

$$= \frac{1}{\Pi\sigma_x\sigma_y}\frac{\sigma_y^2\sigma_x^2}{\sigma_y^2 + \sigma_x^2 z^2} = \frac{1}{\Pi}\frac{\frac{\sigma_y}{\sigma_x}}{\left(\frac{\sigma_y}{\sigma_x}\right)^2 + z^2} \tag{2}$$

and the distribution function of the quotient has the following form:

$$F(x) = \frac{1}{2} + \frac{1}{\Pi}\operatorname{arc\,tg}\frac{z}{\left(\frac{\sigma_y}{\sigma_x}\right)} \tag{3}$$

Then, the distribution of the quotient of random variables when the variables have the standardized normal distributions is the Cauchy distribution and has no moment of any degree.

In 1930 G e a r y introduced formulas for two independent random variables with normal distributions. Let $W = \frac{X}{Y}$ be the quotient of two independent random variables i.e. $X \sim N(\mu_1, \sigma_1)$ and $Y \sim N(\mu_2, \sigma_2)$ and random variable Y in the denominator is positive. Then the density function of random variable W is of the form:

$$f(w) = \frac{1}{\sqrt{2\Pi}}\frac{\mu_2\sigma_1^2 + \mu_1\sigma_2^2 w}{(\sigma_1^2 + \sigma_2^2 w^2)^{3/2}}\exp\left\{-\frac{1}{2}\frac{(\mu_1 - \mu_2 w)^2}{\sigma_1^2 + \sigma_2^2 w^2}\right\} \tag{4}$$

M. G r e e n (1965), allowing the same assumptions (see G e a r y 1930) about normality and independence of the variables from the numerator and the denominator and additionally assuming that the coefficient of variability of variable Y is not too large (in practice not greater then 1/3, has found estimates of the mean $E\left(\frac{X}{Y}\right)$ and the variance $D^2\left(\frac{X}{Y}\right)$:

$$E\left(\frac{X}{Y}\right) \approx \frac{\mu_1}{\mu_2}[1 + V_2^2 + 3V_2^4] \tag{5}$$

$$D^2\left(\frac{X}{Y}\right) \approx \frac{\sigma_1^2}{\mu_2^2}[1 + 3V_2^2 + 15V_2^4] + \frac{\mu_1^2}{\mu_2^2} V_2^2[1 + 8V_2^2] \tag{6}$$

where $V_2 = \frac{\sigma_2}{\mu_2}$ is the coefficient of variability of random variable Y.

M. Green showed further that formulas (5) and (6) are approximations for $\frac{1}{0} \leqslant V_2 \leqslant \frac{1}{3}$ and exact values for $V_2 < \frac{1}{10}$.

The above results have been obtained with the assumption of independence of random variables X and Y.

R. C. Feller (see Fieller 1993) has found formulas for the distribution function and the density function of the quotient of random variables with bivariate binormal distributions without the assumption of independence of variables from the nominator and the denominator.

The probability $P(W \geqslant w)$ has the form:

$$P(W \geqslant w) = \int_h^{+\infty} \int_k^{+\infty} \int_{-h}^{+\infty} \int_{-k}^{+\infty} \frac{1}{2\Pi\sqrt{1-q^2}} \exp\left\{-\frac{1}{2} \cdot \frac{1}{1-q^2} \cdot (x_1^2 - 2\rho x_1 y_1 + \right.$$

$$\left. + y_1^2)\right\} dx_1 dy_1 \tag{7}$$

where

$$h = \frac{\bar{x}}{\sigma_x}$$

$$k = \frac{\bar{y}}{(\sigma_y^2 - 2rw\sigma_x\sigma_y + w^2\sigma_x^2)^{1/2}}$$

$$\rho = \frac{r\sigma_y - w\sigma_x}{(\sigma_y^2 - 2rw\sigma_x\sigma_y + w^2\sigma_x^2)^{1/2}},$$ where r is the correlation coefficient,

$$x_1 = \frac{x}{\sigma_x},$$

$$y_1 = \frac{y - wx}{(\sigma_y^2 - 2rw\sigma_x\sigma_y + w^2\sigma_x^2)^{1/2}}.$$

Special tables with values of the function (7) were compiled by Feller.

In further papers some generalizations of the above results can be found. J. Kotlarski considers more general case of the quotient of two random variables of the form (see Kotlarski 1960):

$$\gamma = \frac{X_1^{q_1}}{X_2^{q_2}} \tag{8}$$

where X_1, X_2 are independent random variables with the gamma distributions; $q_1, q_2 \in R - \{0\}$.

J. Kotlarski shows that for all pairs (q_1, q_2) i.e. the parameters of the distributions of X_1 and X_2, there exist independent positive random variables Y_1 and Y_2 with distribution essentially different from $X_1^{q_1}$, $X_2^{q_2}$, such that the quotient $\frac{Y_1}{Y_2}$ has the same distribution as γ.

The next generalizations were found by G. Marsaglia (see Marsaglia 1965). He considers the quotient of random variables:

$$W = \frac{a+X}{b+Y} \tag{9}$$

where X, Y are independent random variables $X \sim N(0, 1)$, $Y \sim N(0, 1)$, a, b are nonnegative constans.

He shows further that the distribution function $F(t)$ of random variable W of the form:

$$F(t) = P\left[\frac{a+X}{b+Y} < t\right] \tag{10}$$

can be expressed by the bivariate normal distribution or by means of the Nicholson function:

$$F(t) = L\left[\frac{a-bt}{\sqrt{1+t^2}},\ -b,\ \frac{t}{\sqrt{1+t^2}}\right] + L\left[\frac{-a+bt}{\sqrt{1+t^2}},\ b,\ \frac{t}{\sqrt{1+t^2}}\right] \tag{11}$$

$$F(t) = \int_0^{(bt-a)/\sqrt{1+t^2}} \varphi(x)dx + \int_0^b \rho(x)dx + 2L\left(\frac{bt-a}{\sqrt{1+t^2}},\ b,\ \frac{t}{\sqrt{1+t^2}}\right) \tag{12}$$

$$F(t) = \frac{1}{2} + \frac{1}{\Pi} tan^{-1} t + 2V\left(\frac{bt-a}{\sqrt{1+t^2}},\ \frac{b+at}{\sqrt{1+t^2}}\right) - 2V(b, a) \tag{13}$$

where $L(h, k, \varphi) = P(\xi > h, \eta > k)$ and ξ are normal standardized variables with covariance ρ, V is the Nicholson function of the form:

$V(h, g) = \int_0^h \int_0^{qx/h} \varphi(x)\varphi(y)dydx$, where φ is the density function of the standarized normal distribution.

For large b the second and the third component of (12) may be replaced by 0.5 and 0, hence

$$P\left[\frac{a+X}{b+Y}<t\right] \cong + \int_{0}^{(bt-a)/\sqrt{1+t^2}} \varphi(x)dx = \int_{-\infty}^{(bt-a)/\sqrt{1+t^2}} \varphi(x)dx \tag{14}$$

This formula is a very good approximation for estimating values of the distribution function $F(t)$. The density function of variable (9) is of the form:

$$F(W) = \frac{e^{-\frac{1}{2}(a^2+b^2)}}{\Pi(1+W^2)}\left[1+\frac{q}{\varphi(q)}\int_0^q \varphi(y)dy\right] \tag{15}$$

where

$$q = \frac{b+aw}{\sqrt{1+w^2}}.$$

The density function (15) is unimodal or bimodal depending on the values of a and b. If $a > 2{,}257$ then the density is bimodal, but one maximum is insiqnificant.

In the same paper G. Marsaglia shows the formula for the distribution function of the quotient of the sum of independent random variables with the uniform distribution on interval $\langle 0, 1\rangle$:

$$P\left[\frac{u_1+u_2+\ldots+u_n}{v_1+v_2+\ldots+v_m}<a\right] \cong \varphi\left[\frac{\sqrt{3}(am-n)}{\sqrt{a^2m+n}}\right] \tag{16}$$

where φ is the distribution function of the normal distribution with mean $0{,}5(n+ma)$ and variance $(a^2m+n)^{1/2}$.

Approximate formulas for the case where $E(X)$ and $D^2(X)$ are known can be found in the literature of the field. In particular when $Z = \frac{1}{X}$, we have:

$$E\left(\frac{1}{X}\right) \approx \frac{1}{\mu} \tag{17}$$

$$D^2\left(\frac{1}{X}\right) \approx \frac{\sigma^2}{\mu^4} \tag{18}$$

where: $u = E(X) \neq 0$ and $\sigma^2 = D^2(X)$.

The approximate formulas cannot be used in some cases, for example, when random variable X has the normal distribution ($f(x)$ is the density-function of the variable X) the mean $E\left(\frac{1}{X}\right)$ does not exist.

The authors (see Pietcold et al. 1974) suggest assuming in this case, that the variable under consideration has the truncated normal distribution. In this paper the authors show the approximate formulas to calculate the mean and the variance of random variables $\frac{1}{X}$ with truncated normal distribution. With the use of published tables the value of parameters can be obtained. The authors suggest assuming the so called critical truncation level

$$a_{kr} = \frac{\mu}{2} - \sqrt{\left(\frac{\mu}{2}\right)^2 - 1} \quad (\text{for } \mu > 2) \tag{19}$$

because, in practice, there are no situations where we have no clear conditions at what level to fix the truncation.

If $\mu \leqslant 2$, then the truncation level must be choosen only on the grounds of the essential conditions.

J. R. Magnus (1986) presents important results concerning the integral representation of the moment generation function and s-th moment of quadratic forms and the ratio of these forms (see Milo and Parys 1989).

Let x be a $n \times 1$ vector with the normal distribution with mean μ and positively defined covariance matrix $\Omega = LL'$. Let A be a $n \times n$ symetric matrix. Then:

$$E(x'Ax)^s = \sum_v \gamma_s(v) \prod_{j=1}^{s} \{\mathrm{tr}(L'AL)^j + j\mu' L'^{-1}(L'AL)^j L^{-1}u\}^{nj} \tag{20}$$

where the summing is over all $1 \times s$ vectors $v = (n_1, \ldots, n_s)$ with coordinates fulfilling the conditions:

$$\sum_{j=1}^{s} jn_j = S \text{ and } \gamma_s = s!2^s \prod_{j=1}^{s} [n_j!(2j)^{n_j}]^{-1} \tag{21}$$

Let W_1 and W_2 be random variables and $P(W_w > 0) = 1$. Assume that the joint moment generation function exist for W_1 and W_2 and has the form:

$$\phi(\Theta_1, \Theta_2) = E[\exp(\Theta_1 W_1 + \Theta_2 W_2)] \tag{22}$$

for all $|\Theta_1| < \varepsilon$ and $\Theta_2 < e$, $\varepsilon > 0$.

Then (see Marsaglia 1965) for $s \in N$ the following relation holds:

$$E\left(\frac{W_1}{W_2}\right)^s = \frac{1}{(s-1)!} \int_0^{+\infty} t^{s-1} \left[\frac{\partial^s}{\partial \theta^s} \phi(\theta, -t)\right]_{\theta=0} dt \tag{23}$$

Magnus (see Magnus 1986) has proved the folloving theorema.

Assume that x a $n \times 1$ vector with the normal distribution with mean μ and positively defined covariance matrix $\Omega = LL'$.

Let A be a $n \times n$ symetric quadratic matrix and $B - n \times n$ quadratic matrix half positively defined. Assume that P is a $n \times n$ quadratic orthogonal matrix such that:

$$P'L'BLP = D$$

where $P'P = I_n$ and D is a $n \times n$ quadratic diagonal matrix.

Define $A^* = P'L'ALP$, $\mu^* = P'L^{-1}\mu$ and $\Delta = (I_n + 2tD)^{-1/2}$, $R = \Delta A^* \Delta$, $\xi = \Delta\mu^*$ where ξ is a nx vector.

Then:

$$E\left[\frac{x'Ax}{x'Bx}\right]^s = \frac{d}{(s-1)!}\sum_{v}\gamma(v)\int_0^\infty t^{s-1}|\Delta|\exp\left(\frac{1}{2}\xi'\xi\right) \prod_{j=1}(\mathrm{tr}R^j + j\xi'R^j\xi)^{n_j}dt \tag{24}$$

where the summing is over all $1 \times s$ vectors $v = (n_1, \ldots, n_s)$ with coordinates fulfilling the conditions (21)
and

$$d = \exp\left(-\frac{1}{2}\mu'\Omega^{-1}\mu\right).$$

These results allowed us to formulate some conclusions concerning the moments of random quadratic forms, their ratios and modifications. From (20) we have:

$$\begin{aligned} E(x'Ax) &= \mathrm{tr}\,L'AL + \mu'A\mu \\ E(x'Ax)^2 &= (\mathrm{tr}\,L'AL)^2 + 2\mathrm{tr}\,(L'AL)^2 \\ D^2(x'Ax) &= 2\mathrm{tr}\,(L'AL)^2 + 4\mu'ALL'A\mu \end{aligned} \tag{25}$$

It's easy to verify that for $\mu = 0$ above conclusions have simply from:

$$\begin{aligned} E(x'Ax) &= \mathrm{tr}\,L'AL \\ E(x'Ax)^2 &= \mathrm{tr}(L'AL)^2 + 2\mathrm{tr}\,(L'AL)^2 \\ D^2(x'Ax) &= 2\mathrm{tr}\,(L'AL)^2 \end{aligned} \tag{26}$$

In the case $\mu = 0$ we have obtained some results concerning independence of random quadratic forms.

It is well known that in the case of the independence of random variables $\frac{Y}{X}$ and X the equality $E\left(\frac{Y}{X}\right) = \frac{E(Y)}{E(X)}$ is true.

Then from (25) and from the independence of $\dfrac{x'Ax}{x'Bx}$ and $x'Bx$ we have

$$E\left(\frac{x'Ax}{x'Bx}\right)=\frac{\operatorname{tr}L'AL}{\operatorname{tr}L'BL} \tag{27}$$

If $\dfrac{(x'Ax)^2}{(x'Ax)^2+x'Bx}$ and $(x'Ax)^2+x'Bx$ are independent then

$$E\left(\frac{(x'Ax)^2}{(x'Ax)^2+x'Bx}\right)=\frac{(\operatorname{tr}L'AL)^2+2\operatorname{tr}(L'AL)^2}{(\operatorname{tr}L'AL)^2+2\operatorname{tr}(L'AL)^2+\operatorname{tr}L'BL} \tag{28}$$

From (24) for $s=1$ we have obtained the following results:

$$E\left[\frac{x'Ax}{x'Bx}\right]=d\int_0^\infty t|\Delta|\exp\left(\frac{1}{2}\xi'\xi\right)(\operatorname{tr}R+\xi'R\xi)dt. \tag{29}$$

The properties of the variance imply the following formulas:

$$D^2\left[\frac{x'Ax}{x'Bx}\right]=3d\int_0^\infty t|\Delta|\exp\left(\frac{1}{2}\xi'\xi\right)(\operatorname{tr}R+\xi'R\xi)(\operatorname{tr}R^2+2\xi R^2\xi)dt-\left(E\left[\frac{x'Ax}{x'Bx}\right]\right)^2 \tag{30}$$

We intend to continue our studies in the field of the quotient of random quadratic forms. In particular we will use them to find some properties of the distribution and parameters of different types of quotient (28) concerning the problem of regularizing estimator (see M i l o, P a r y s 1989).

REFERENCES

F i e l l e r, B. A. (1993): *The distribution of the Index in Normal Bivariate Population*, „Biometrica", p. 428–440.

G e a r y, R. C. (1930): *The Frequency Distribution of the Quotient of Two Normal Variates*, „Journal of the Royal Statistical Society", p. 442–446.

G r e e n, M. (1965): *The parameters of the Distribution of the Quotient of Two Normal Variables*, „Growth Project Paper" 238, Cambridge.

K o t l a r s k i, J. (1962): *On Pairs of Independent Random Variables Whose Quotient Follows Some Known Distribution*, „Collegium Mathematicum", Vol. IX, p. 151.

K o t l a r s k i, J. (1960): *On Random Variables Whose Quotient Follows the Cauchy Law*, „Collegium Mathematicum", Vol. VII.

M a g n u s, J. R. (1986): *The Exact Moments of a Ratio of Quadratic Forms in Normal Variables*, „Annales d'Economic et de Statistique", No 4.

Marsaglia, G. (1966): *Ratios of Normal Variables and Ratios of Sums of Uniform Variables*, „American Statistical Association Journal", Vol. 60, p. 195–204.

Milo, W., Parys, D. (1989): *Całkowe przedstawienia skalarnych funkcji macierzy losowych i ich zastosowania*, Cz. 1 (*Estymacja modeli losowych*), CPBP 10.09.III.6, Łódź (maszynopis powielony).

Nicholson, C. (1943): *The Probability Integral for Two Variables*, „Biometrica".

Pietcold, K., Tomaszewicz, A., Żółtowska, E. (1974): *Momenty odwrotności zmiennej losowej o uciętym rozkładzie normalnym*, „Przegląd Statystyczny", z. 4.

Andrzej Czajkowski, Dariusz Parys

ROZKŁAD I PARAMETRY ROZKŁADU ILORAZU ZMIENNYCH LOSOWYCH

W pracy tej prezentujemy wcześniejsze rezultaty badań dotyczących własności rozkładów ilorazów zmiennych losowych.

Ponadto prezentujemy własne wyniki badań dotyczące momentów rozkładów ilorazów losowych form kwadratowych.