

ON THE APPROXIMATE DISTRIBUTION OF RATIOS

By P. L. Hsu

National University of Peking

The purpose of this paper is to apply Cramer's theorem of asymptotic expansion¹ and Berry's theorem² to study the approximate distribution of ratios of the following two types:

$$(I) \quad Z = \frac{1}{n} (Y_1 + \cdots + Y_n) \bigg/ \frac{1}{m} (\bar{X}_1 + \cdots + \bar{X}_m) = \bar{Y}/\bar{X},$$

$$(II) \quad Z = Y \bigg/ \frac{1}{m} (X_1 + \cdots + X_m) = Y/\bar{X}.$$

In (I) the X_i , Y_j are independent, the Y_j are equi-distributed,³ and the X_i are equi-distributed and positive. In (II) X_1, \cdots, X_n , Y are independent and positive, and the X_i are equi-distributed.

1. **The ratio (I).** Assume that (I1) the absolute k th moment of X_i and that of Y_j are finite and positive, where k is a fixed integer ≥ 3 , (I2) the distribution of X_i and that of Y_j are non-singular.

Let

$$\xi = \epsilon(X_i), \quad \eta = \epsilon(Y_j), \quad \sigma^2 = \epsilon(X_i^2) - \xi^2, \quad \tau^2 = \epsilon(Y_j^2) - \eta^2$$

and

$$U = \frac{\sqrt{m}}{\sigma} (\bar{X} - \xi), \quad V = \frac{\sqrt{n}}{\tau} (\bar{Y} - \eta).$$

Let $F(x)$, $G(x)$ and $H(x)$ be respectively the distribution functions of Z , U and V . Let

$$b = \left(\frac{\sigma^2 x^2}{m} + \frac{\tau^2}{n} \right)^{\frac{1}{2}}, \quad u = \frac{\xi n - \eta}{b}.$$

Then the relation $Z \leq x$ is equivalent to

$$-\frac{x\sigma U}{b\sqrt{m}} + \frac{+V}{b\sqrt{n}} \leq u.$$

¹ H. CRAMÉR. *Random Variables and Probability Distributions* (1937), Chap. 7.

² A. C. BERRY. "The accuracy of the Gaussian approximation to the sum of independent variates", *Trans. Amer. Math. Soc.*, Vol. 49 (1941), pp. 122-136.

³ The Y_j are said to be equi-distributed if all Y_j have the same distribution function.



For simplicity we shall assume $x > 0$; the results are, however, general. Then the distribution functions of $-\frac{x\sigma U}{b\sqrt{m}}$ and $\frac{\tau V}{b\sqrt{n}}$ are

$$Pr\left\{-\frac{x\sigma U}{b\sqrt{m}} < y\right\} = 1 - G\left(-\frac{b\sqrt{m}y}{\sigma x}\right), \quad Pr\left\{\frac{\tau V}{b\sqrt{n}} \leq y\right\} = H\left(\frac{b\sqrt{ny}}{\tau}\right).$$

Hence, by the theorem of convolution,

$$(1) \quad F(x) = \int_{-\infty}^{\infty} \left\{1 - G\left(-\frac{b\sqrt{m}(u-y)}{\sigma x}\right)\right\} dH\left(\frac{b\sqrt{ny}}{\tau}\right).$$

Here we recall the theorems of Cramér and Berry: Under the conditions (I1) and (I2)

$$(2) \quad G(x) = \Phi(x) + \sum_{\nu=1}^{k-3} \frac{P_{\nu}(x)}{m^{\nu/2}} + \frac{D_k}{m^{\frac{1}{2}(k-2)}},$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy, \quad P_{\nu}(x) = \sum_{j=1}^{\nu} c_{j\nu} \Phi^{(\nu+2j)}(x),$$

and $|D_k|$ is less than a positive number which depends only on k and the distribution of X_i . If $k = 3$, condition (I2) may be removed.⁴

Analogously,

$$(3) \quad H(x) = \Phi(x) + \sum_{\nu=1}^{k-3} \frac{Q_{\nu}(x)}{n^{\nu/2}} + \frac{D'_k}{m^{\frac{1}{2}(k-2)}},$$

where

$$Q_{\nu}(x) = \sum_{j=1}^{\nu} d_{j\nu} \Phi^{(\nu+2j)}(x).$$

In the sequel we shall use the letter Δ_k to denote an unspecified quantity such that $|\Delta_k|$ is less than a positive number which depends only on k , the distribution of X_i and the distribution of Y_j .

Using (2) we have

$$(4) \quad 1 - G(-x) = \Phi(x) + \sum_{\nu=1}^{k-3} \frac{(-1)^{\nu} P_{\nu}(x)}{m^{\nu/2}} + \frac{D_k}{m^{\frac{1}{2}(k-2)}}$$

and this making this substitution in (1) we get

$$F(x) = \int_{-\infty}^{\infty} \Phi\left(\frac{b\sqrt{m}(u-y)}{\sigma x}\right) dH\left(\frac{b\sqrt{ny}}{\tau}\right) + \sum_{\nu=1}^{k-3} \frac{(-1)^{\nu}}{m^{\nu/2}} \int_{-\infty}^{\infty} P_{\nu}\left(\frac{b\sqrt{m}(u-y)}{\sigma x}\right) dH\left(\frac{b\sqrt{ny}}{\tau}\right) + \frac{\Delta_k}{m^{\frac{1}{2}(k-2)}},$$

⁴ This last assertion constitutes Berry's theorem.

and so by partial integration,

$$F(x) = \int_{-\infty}^{\infty} H\left(\frac{b\sqrt{n}(u-y)}{\tau}\right) d\Phi\left(\frac{b\sqrt{m}y}{\sigma x}\right) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} \int_{-\infty}^{\infty} H\left(\frac{b\sqrt{n}(u-y)}{\tau}\right) dP_\nu\left(\frac{b\sqrt{m}y}{\sigma x}\right) + \frac{\Delta_k}{m^{\frac{1}{2}(k-2)}}.$$

Making the transformation $y = \sigma xv/b\sqrt{m}$ and writing

$$(5) \quad \alpha = \frac{b\sqrt{n}}{\tau}, \quad \beta = \frac{\sigma\sqrt{n}x}{\tau\sqrt{m}}$$

we get

$$F(x) = \int_{-\infty}^{\infty} H(\alpha u - \beta v)\Phi'(v) dv + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} \int_{-\infty}^{\infty} H(\alpha u - \beta v)P'_\nu(v) dv + \frac{\Delta_k}{m^{\frac{1}{2}(k-2)}} \\ = I_0 + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} I_\nu + \frac{\Delta_k}{m^{\frac{1}{2}(k-2)}}.$$

For I_0 we use (3) and obtain

$$I_0 = \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v)\Phi'(v) dv + \sum_{\nu=1}^{k-3} \frac{1}{n^{\nu/2}} \int_{-\infty}^{\infty} Q_\nu(\alpha u - \beta v)\Phi'(v) dv + \frac{\Delta_k}{n^{\frac{1}{2}(k-2)}}.$$

For I_ν we use (3) with k replaced by $k - \nu$. Thus

$$I_\nu = \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v)P'_\nu(v) dv + \sum_{\mu=1}^{k-3-\nu} \frac{1}{n^{\mu/2}} \int_{-\infty}^{\infty} Q_\mu(\alpha u - \beta v)P'_\nu(v) dv + \frac{\Delta_k}{n^{\frac{1}{2}(k-2-\nu)}}.$$

Combining these results we get

$$(6) \quad F(x) = \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v)\Phi'(v) dv + \sum_{\nu=1}^{k-3} \frac{1}{n^{\nu/2}} \int_{-\infty}^{\infty} Q_\nu(\alpha u - \beta v)\Phi'(v) dv \\ + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v)P'_\nu(v) dv \\ + \sum_{\nu=1}^{k-3} \sum_{\mu=1}^{k-3-\nu} \frac{(-1)^\nu}{m^{\nu/2} n^{\mu/2}} \int_{-\infty}^{\infty} Q_\mu(\alpha u - \beta v)P'_\nu(v) dv + R_k,$$

where

$$R_k = \frac{\Delta_k}{m^{\frac{1}{2}(k-2)}} + \frac{\Delta_k}{n^{\frac{1}{2}(k-2)}} + \sum_{\nu=1}^{k-3} \frac{\Delta_k}{m^{\nu/2} n^{\frac{1}{2}(k-2-\nu)}} = \Delta_k \left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right)^{k-2}.$$

Now by (5), $\alpha > 0$ and $\alpha^2 - \beta^2 = 1$. For such values of α and β , however, it follows easily from the theorem of convolution that

$$\int_{-\infty}^{\infty} \Phi(\alpha u - \beta v)\Phi'(v) dv = \Phi'(u).$$

As differentiation under the integration sign is justified by the boundedness of the derivatives of Φ we have

$$\alpha^p \int_{-\infty}^{\infty} \Phi^{(p)}(\alpha u - \beta v) \Phi'(v) dv = \Phi^{(p)}(u).$$

Repeated partial integration then gives

$$\begin{aligned} \int_{-\infty}^{\infty} \Phi^{(p)}(\alpha u - \beta v) \Phi^{(q)}(v) dv &= \beta^{q-1} \int_{-\infty}^{\infty} \Phi^{(p+q-1)}(\alpha u - \beta v) \Phi'(v) dv \\ &= \frac{\beta^{q-1}}{\alpha^{p+q-1}} \Phi^{(p+q-1)}(u). \end{aligned}$$

Hence

$$\begin{aligned} \int_{-\infty}^{\infty} Q_r(\alpha u - \beta v) \Phi'(v) dv &= \sum_{j=1}^r d_{jr} \int_{-\infty}^{\infty} \Phi^{(r+2j)}(\alpha u - \beta v) \Phi'(v) dv \\ &= \sum_{j=1}^r \frac{d_{jr}}{\alpha^{r+2j}} \Phi^{(r+2j)}(u), \\ \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v) P'_r(v) dv &= \sum_{j=1}^r c_{jr} \int_{-\infty}^{\infty} \Phi(\alpha u - \beta v) \Phi^{(r+2j+1)}(v) dv \\ &= \sum_{j=1}^r \frac{\beta^{r+2j} c_{jr}}{\alpha^{r+2j}} \Phi^{(r+2j)}(u), \\ \int_{-\infty}^{\infty} Q_\mu(\alpha u - \beta v) P'_r(v) dv &= \sum_{i=1}^\mu \sum_{j=1}^r d_{i\mu} c_{jr} \int_{-\infty}^{\infty} \Phi^{(\mu+2j)}(\alpha u - \beta v) \Phi^{(r+2j+1)}(v) dv \\ &= \sum_{i=1}^\mu \sum_{j=1}^r d_{i\mu} c_{jr} \frac{\beta^{r+2j}}{\alpha^{\mu+r+2i+2j}} \Phi^{(\mu+r+2i+2j)}(u). \end{aligned}$$

Making all these substitutions in (6) we obtain the final result

$$\begin{aligned} F(x) = \Phi(u) &+ \sum_{r=1}^{k-3} \frac{(-1)^r}{m^{r/2}} \sum_{j=1}^r \frac{\beta^{r+2j} c_{jr}}{\alpha^{r+2j}} \Phi^{(r+2j)}(u) + \sum_{r=1}^{k-3} \frac{1}{n^{r/2}} \sum_{j=1}^r \frac{d_{jr}}{\alpha^{r+2j}} \Phi^{(r+2j)}(u) \\ &+ \sum_{r=1}^{k-3} \sum_{\mu=1}^{k-3-r} \frac{(-1)^r}{m^{r/2} n^{\mu/2}} \sum_{i=1}^\mu \sum_{j=1}^r d_{i\mu} c_{jr} \frac{\beta^{r+2j}}{\alpha^{\mu+r+2i+2j}} \Phi^{(\mu+r+2i+2j)}(u) \\ &+ \Delta_k \left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right)^{k-2}. \end{aligned}$$

If $k = 3$, the result remains true without the condition (I2).

2. The ratio (II). Here we make the following assumptions:

(II1) The k th moment of X_i is finite and positive, where k is a fixed integer $\geq k$, $\epsilon(X_i) = 1$,⁵ $\epsilon(X_i^2) - 1 = \sigma^2$.

(II2) The distribution of X_i is non-singular.

⁵ As the case $\epsilon(X_i) = 0$ is excluded, there is no loss of generality in this assumption.

Let $U = \sqrt{m}(\bar{X} - 1)/\sigma$, and $F(x)$, $G(x)$ and $H(x)$ be respectively the distribution functions of Z , U and Y . Then

$$F(x) = Pr \left\{ Y - \frac{\sigma x U}{\sqrt{m}} \leq x \right\}$$

Because of the positiveness of X_i and Y we may always assume $x > 0$. Then, by the theorem of convolution,

$$F(x) = \int_{-\infty}^{\infty} \left\{ 1 - G \left(-\frac{\sqrt{m}(x-y)}{\sigma x} \right) \right\} dH(y).$$

Using (4) we have

$$F(x) = \int_{-\infty}^{\infty} \left\{ \Phi \left(\frac{\sqrt{m}(x-y)}{\sigma x} \right) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} P_\nu \left(\frac{\sqrt{m}(x-y)}{\sigma x} \right) \right\} dH(y) + \frac{A_k}{m^{\frac{1}{2}(k/2)},}$$

where, as throughout the rest of this paper, A_k represents an unspecified quantity such that $|A_k|$ is less than a positive number depending only on k , the distribution of X_i and the distribution of Y . By partial integration we get

$$\begin{aligned} (7) \quad F(x) &= \int_{-\infty}^{\infty} H(x-y) d \left\{ \Phi \left(\frac{\sqrt{m}y}{\sigma x} \right) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu P_\nu \left(\frac{\sqrt{m}y}{\sigma x} \right)}{m^{\nu/2}} \right\} + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \int_{-\infty}^{\infty} H \left(x - \frac{\sigma x z}{\sqrt{m}} \right) \left(\Phi'(z) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu P'_\nu(z)}{m^{\frac{1}{2}(k-2)}} \right) dz + \frac{A_k}{m^{\frac{1}{2}(k-2)}}. \end{aligned}$$

An interesting special case is the following: Suppose that (II3) $H^{(k-2)}(x)$ exists and is continuous for all $x \geq 0$; (II4) the functions

$$\xi_\nu(x) = x^\nu H^{(\nu)}(x) \quad (\nu = 1, \dots, k-3)$$

are bounded, i.e.

$$\xi_\nu(x) = A_k;$$

(II3) there is a positive constant $c < 1$ such that

$$x^{k-2} H^{(k-2)}(y) = A_k$$

for all $x \geq 0$ and $(1-c)x \leq y \leq (1+c)x$. Under these conditions we have

$$\begin{aligned} H \left(x - \frac{\sigma x z}{\sqrt{m}} \right) &= \sum_{\nu=0}^{k-3} \frac{(-1)^\nu \sigma^\nu x^\nu z^\nu H^{(\nu)}(x)}{\nu! m^{\nu/2}} \\ &\quad + \frac{(-1)^{k-2} \sigma^{k-2} x^{k-2} z^{k-2}}{(k-2)! m^{\frac{1}{2}(k-2)}} H^{(k-2)} \left(x + \frac{\delta \sigma x z}{\sqrt{m}} \right) \quad (|\delta| \leq 1), \end{aligned}$$

and so, for $|z| \leq \frac{c\sqrt{m}}{\sigma}$ we have

$$(8) \quad H \left(x - \frac{\sigma x z}{\sqrt{m}} \right) = \sum_{\nu=0}^{k-3} \frac{(-1)^\nu \sigma^\nu \xi_\nu z^\nu}{\nu! m^{\nu/2}} + \frac{A_k z^{k-2}}{m^{\frac{1}{2}(k-2)}}.$$

Separate now the integral in (7) into two parts:

$$I_1 = \int_{|z| \leq c\sqrt{m}/\sigma}, \quad I_2 = \int_{|z| > c\sqrt{m}/\sigma}.$$

Now

$$|I_2| \leq \int_{|z| > c\sqrt{m}/\sigma} \left| \Phi'(z) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu P'_\nu(z)}{m^{\nu/2}} \right| dz.$$

Evidently this last integral is exponentially small and so is $A_k/m^{\frac{1}{2}(k-2)}$. By (8),

$$\begin{aligned} I_1 &= \int_{|z| \leq c\sqrt{m}/\sigma} \left(\sum_{\nu=0}^{k-3} \frac{(-1)^\nu \sigma^\nu \xi_\nu z^\nu}{\nu! m^{\nu/2}} \right) \left(\Phi'(z) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu P'_\nu(z)}{m^{\nu/2}} \right) dz + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \int_{-\infty}^{\infty} \left(\sum_{\nu=0}^{k-3} \frac{(-1)^\nu \sigma^\nu \xi_\nu z^\nu}{\nu! m^{\nu/2}} \right) \left(\Phi'(z) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu P'_\nu(z)}{m^{\nu/2}} \right) dz + \frac{A_k}{m^{\frac{1}{2}(k-2)}}. \end{aligned}$$

Combining these results we obtain

$$\begin{aligned} F(x) &= \int_{-\infty}^{\infty} \left(\sum_{\nu=0}^{k-3} \frac{(-1)^\nu \sigma^\nu \xi_\nu z^\nu}{\nu! m^{\nu/2}} \right) \left(\Phi'(z) + \sum_{\nu=1}^{k-3} \frac{(-1)^\nu}{m^{\nu/2}} \sum_{j=1}^{\nu} c_{j\nu} \Phi^{(\nu+2j+1)}(z) \right) dz + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \sum_{\nu=0}^{k-3} \frac{d_\nu \xi_\nu}{m^{\nu/2}} I_{\nu 1} + \sum_{\nu=0}^{k-3} \sum_{\mu=1}^{k-3} \sum_{j=1}^{\mu} \frac{d_{j\mu} \xi_\nu}{m^{\frac{1}{2}(\mu+\nu)}} \Phi_{\nu, \mu+2j+1} + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \sum_1 + \sum_2 + \frac{A_k}{m^{\frac{1}{2}(k-2)}}, \end{aligned}$$

where

$$I_{\alpha\beta} = \int_{-\infty}^{\infty} z^\alpha \Phi^{(\beta)}(z) dz.$$

Now the following facts can easily be established by means of partial integration:

(9) $I_{\alpha\beta} = 0$ when $\alpha - \beta$ is even,

(10) $I_{\alpha\beta} = 0$ when $\beta - \alpha > 1$.

By (9), the non-vanishing terms in \sum_1 are the even terms and the non-vanishing terms in \sum_2 are those for which $\mu + \nu$ is even. Hence

$$\begin{aligned} \sum_1 &= \sum_{\nu=0}^{[\frac{1}{2}(k-3)]} \frac{e_\nu \xi_{2\nu}}{m^\nu}, \\ \sum_2 &= \sum_{\nu=0}^{[\frac{1}{2}(k-3)]} \sum_{\mu=1}^{[\frac{1}{2}(k-3)]} \sum_{j=1}^{2\mu} \frac{e_{j\mu\nu} \xi_{2\nu}}{m^{\mu+\nu}} I_{2\nu, 2\mu+2j+1} + \sum_{\nu=0}^{[\frac{1}{2}(k-4)]} \sum_{\mu=0}^{[\frac{1}{2}(k-4)]} \sum_{j=1}^{2\mu+1} \frac{e'_{j\mu\nu} \xi_{2\nu+1}}{m^{\mu+\nu+1}} I_{2\nu+1, 2\mu+2j+2}. \end{aligned}$$

Using (10) to reduce \sum_2 further we get

$$\begin{aligned} \sum_2 &= \sum_{r=2}^{[\frac{1}{2}(k-3)]} \sum_{\mu=1}^{r-1} \sum_{j=1}^{2\mu} \frac{e_{j\mu\nu} \xi_{2\nu}}{m^{\mu+\nu}} I_{2r, 2\mu+2j+1} + \sum_{r=1}^{[\frac{1}{2}(k-4)]} \sum_{\mu=0}^{r-1} \sum_{j=1}^{2\mu+1} \frac{e'_{j\mu\nu} \xi_{2\nu+1}}{m^{\mu+r+1}} I_{2r+1, 2\mu+2j+2} \\ &= \sum_{r=0}^{[\frac{1}{2}(k-7)]} \sum_{\mu=0}^r \frac{g_{\mu\nu} \xi_{2\nu+4}}{m^{\mu+r+3}} + \sum_{r=0}^{[\frac{1}{2}(k-6)]} \sum_{\mu=0}^r \frac{g'_{\mu\nu} \xi_{2\nu+3}}{m^{\mu+r+2}} \\ &= \sum_{\alpha=0}^{[\frac{1}{2}(k-9)]} \frac{1}{m^{\alpha+3}} \sum_{\beta=[\frac{1}{2}(\alpha+1)]}^{\alpha} h_{\alpha\beta} \xi_{2\beta+4} + \sum_{\alpha=0}^{[\frac{1}{2}(k-6)]} \frac{1}{m^{\alpha+2}} \sum_{\beta=[\frac{1}{2}(\alpha+1)]}^{\alpha} h'_{\alpha\beta} \xi_{2\beta+3} + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \sum_{i=3}^{[\frac{1}{2}(k-3)]} \frac{1}{m^i} \sum_{j=[\frac{1}{2}(i-2)]}^{i-2} l_{ij} \xi_{2j+4} + \sum_{i=2}^{[\frac{1}{2}(k-3)]} \frac{1}{m^i} \sum_{j=[\frac{1}{2}(i-1)]}^{i-2} l'_{ij} \xi_{2j+3} + \frac{A_k}{m^{\frac{1}{2}(k-2)}}. \end{aligned}$$

Hence

$$\begin{aligned} \sum_1 + \sum_2 &= \xi_0 + \frac{e_1 \xi_2}{m} + \frac{e_2 \xi_4 + l'_{20} \xi_3}{m^2} + \sum_{r=3}^{[\frac{1}{2}(k-3)]} \frac{1}{m^r} \\ &\quad \cdot \left(e_r \xi_{2r} + \sum_{\mu=[\frac{1}{2}(r-2)]}^{r-3} l_{\mu\nu} \xi_{2\nu+4} + \sum_{\mu=[\frac{1}{2}(r-2)]}^{r-2} l'_{\mu\nu} \xi_{2\nu+3} \right) + \frac{A_k}{m^{\frac{1}{2}(k-2)}} \\ &= \xi_0 + \sum_{r=1}^{k-3} \frac{1}{m^r} \sum_{j=r+1}^{2r} p_{jr} \xi_j + \frac{A_k}{m^{\frac{1}{2}(k-2)}}. \end{aligned}$$

Hence

$$F(x) = \xi_0 + \sum_{r=1}^{k-3} \frac{1}{m^r} \sum_{j=r+1}^{2r} p_{jr} \xi_j + \frac{p_k}{m^{\frac{1}{2}(k-2)}}.$$

Our final conclusion is: Under the conditions (II1)–(II5) formula (11) is true; if $k = 3$, (11) remains true without the condition (II2).