HYPERGEOMETRIC FUNCTIONS IN SEQUENTIAL ANALYSIS[1]

BY ADNAN F. IFRAM

University of Illinois and Northwestern University

1. Introduction and summary. In several sequential probability ratio tests [9] [12], density ratios may be expressed in terms of hypergeometric functions whose asymptotic behavior is indirectly available in the literature, and is useful in establishing the almost sure termination of these tests [6] [7] [8] [10]. The results of this paper are new for the sequential ordinary and multiple correlation coefficient tests [4] [6]. In addition, they complete the results of [8] and [10] for the sequential F-test [2] [6] as well as those of [7] for the sequential χ^2- and T^2-tests [5] [6].

The generalized hypergeometric function ${}_pF_q$ is defined by:

$$(1.1)\quad {}_pF_q(a_1, \cdots, a_p; c_1, \cdots, c_q; z) = 1 + (a_1 \cdots a_p/c_1 \cdots c_q)z/1! + [a_1(a_1+1) \cdots a_p(a_p+1)/c_1(c_1+1) \cdots c_q(c_q+1)]z^2/2! + \cdots$$

for $p, q \geqq 0$ and $c_i > 0$, $i = 1, \cdots, q$. We shall need in the sequel three such functions: ${}_2F_1(a, b; c; z)$, which is convergent for $|z| < 1$, ${}_1F_1(a; c; z)$, and ${}_0F_1(\ ; c; z)$, which are convergent for all z.

2. Some asymptotic formulae. We use $a_n \sim_{n\to\infty} b_n$ and $a_n = O(b_n)$ to mean respectively that $a_n/b_n \to 1$ and a_n/b_n remains bounded as $n \to \infty$.

(1) Watson [13] treats the asymptotic behavior of ${}_2F_1(a + \epsilon_1\lambda, b + \epsilon_2\lambda; c + \epsilon_3\lambda; z)$ for large values of λ, where each of ϵ_1, ϵ_2, and ϵ_3 can take the values -1, 0 and 1. The asymptotic expansion of two independent solutions of the associated hypergeometric differential equation corresponding to the case $(\epsilon_1, \epsilon_2, \epsilon_3) = (1, -1, 0)$ are given in [13]. A procedure is prescribed for obtaining the expansion of some other related combinations of ϵ_1, ϵ_2 and ϵ_3 by using the relations between different solutions of a hypergeometric differential equation given in Forsyth [3]. Following such a procedure we obtain:

$$(2.1)\qquad {}_2F_1(\lambda, \lambda; c; z) \sim_{\lambda\to\infty} \lambda^{-\frac{1}{2}} B(1 - z^{\frac{1}{2}})^{-2\lambda},$$

where

$$B = [\Gamma(c)\Gamma(\lambda + 1 - c)/2\pi^{\frac{1}{2}}\Gamma(\lambda)]z^{(1-2c)/4}(1 - z^{\frac{1}{2}})^c \qquad \text{for } z \in (0, 1).$$

(2) It follows directly from its definition that

$$(2.2)\qquad {}_2F_1(a, b; c + \lambda; z) \sim_{\lambda\to\infty} 1.$$

(3) We shall need the asymptotic behavior of ${}_1F_1(a_n; c; z_n)$, where c is fixed and $a_n, z_n \to \infty$ as $n \to \infty$, such that $z_n = O(a_n)$. The asymptotic behavior of a function related to ${}_1F_1$, called Whittaker's function, is given in [11]. Using the

Received 7 December 1964; revised 27 May 1965.

[1] This paper is part of a doctoral dissertation submitted at the University of Illinois, June, 1963. The research was supported by the National Science Foundation, grant G-21507.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

results in [11] (p. 9 and p. 75), we obtain

$$
\begin{aligned}
&{}_1F_1(\lambda + a; c; \lambda z) \sim_{\lambda\to\infty} d(\lambda)D(z) \\
&\qquad\qquad \cdot\{[(z^{\frac{1}{2}} + (z + 4)^{\frac{1}{2}})/2] \exp [(z + z^{\frac{1}{2}}(z + 4)^{\frac{1}{2}})/4]\}^{2\lambda}, \\
&D(z) = [\Gamma(c)/(2\pi)^{\frac{1}{2}}]z^{-c/2}[z/(z + 4)]^{\frac{1}{4}} \\
&\qquad\qquad \cdot\{[(z^{\frac{1}{2}} + (z + 4)^{\frac{1}{2}})/2] \exp [(z + z^{\frac{1}{2}}(z + 4)^{\frac{1}{2}})/4]\}^{a-c}, \\
&d(\lambda) = \lambda^{-c/2}(\lambda + a - c)^{(1-c)/2} \\
&\qquad\qquad \cdot\exp [(c - 2)(c - 1)c/48(\lambda + a - c)^2].
\end{aligned}
\tag{2.3}
$$

(4) In order to study the limiting behavior of ${}_0F_1(\ ; c; z)$ for large z, we use a certain relation between ${}_0F_1$ and ${}_1F_1$ known as Kummer's second theorem and given in [11] (p. 12):

$$
{}_0F_1(\ ; c + \tfrac{1}{2}; z^2/16) = e^{-z/2}\, {}_1F_1(c; 2c; z). \tag{2.4}
$$

The following relevant result is given in [11] (p. 60):

$$
{}_1F_1(a; c; z) \sim_{z\to\infty} [\Gamma(c)/\Gamma(a)]e^z z^{a-c}. \tag{2.5}
$$

Combining (2.4) and (2.5) we have:

$$
{}_0F_1(\ ; c + \tfrac{1}{2}; z^2/16) \sim_{z\to\infty} e^{z/2}[\Gamma(2c)/\Gamma(c)]z^{-c}. \tag{2.6}
$$

3. Applications. The applications of this section are special cases of sequential probability ratio tests based on a sequence $\{X_n\}$ whose family of distributions satisfies certain sufficiency and monotone likelihood ratio properties [6]. Let $q_{\theta n}(x)$ denote the density of X_n and let $r_n = q_{\theta_2 n}/q_{\theta_1 n}$ where $\theta_1 < \theta_2$. It was shown in [6] that the almost sure termination of such a test is implied by Conditions B and A_3; where Condition B states that $q_{\theta n}(x) \sim_{n\to\infty} K(n)C(\theta, x)e^{nh(\theta,x)}$ for some K, C and h; while Condition A_3 requires $r_n(\theta_0 + (t/n)) \to_{n\to\infty} \alpha e^{t\beta}$ for some $\alpha, \beta > 0$, all $t \neq 0$, and θ_0 satisfying $h(\theta_1, \theta_0) = h(\theta_2, \theta_0)$. In the following we shall freely use the notation of [6]. The expressions appearing in Conditions B and A_3 will not be displayed here since they coincide with those obtained in [6].

(i) *The sequential F-test.* It follows from [6] and p. 268 of [9] that

$$
\begin{aligned}
q_{\theta n}(x) &= \Gamma(\lambda + a)[\Gamma(c)\Gamma(\lambda + a - c)]^{-1} \\
&\qquad \cdot x^{c-1}(1 + x)^{-(\lambda+a)}e^{-\lambda\theta}\, {}_1F_1(\lambda + a; c; \lambda z),
\end{aligned}
\tag{3.1}
$$

where $2a = s - q$, $2c = q$, $2\lambda = kn$, and $z = \theta x/(1 + x)$. Condition B follows from (2.3). We remark that another asymptotic relation, known as Perron's formula, was used in [8]. This is not justified, since Perron's formula describes the asymptotic behavior of ${}_1F_1(a_n; c, z)$, where $a_n \to \infty$ while c and z remain fixed. It follows from [11] (p. 74) that (2.3) continues to hold if z is replaced throughout by $z(\lambda)$, where $z(\lambda)$ has a finite limit as $\lambda \to \infty$. Thus, Condition A_3 follows by replacing x by $x + (t/n)$ in (3.1) and (2.3).

(ii) *The sequential* χ^2*-test.* It follows from [6] and p. 312 of [9] that

$$
q_{\theta n}(x) = (n/2)^c x^{c-1}e^{-n(x+\theta)/2}\, {}_0F_1(\ ; c; z^2/16), \tag{3.2}
$$

where $z = 2n(\theta x)^{\frac{1}{2}}$ and $2c = q$. Condition B follows from (2.6). Condition A_3 holds since x may be replaced by $x + (t/n)$ in (2.6).

It is easily seen from [6] that the limiting behavior of the sequential T^2-test is reducible to that of the F-test rather than the χ^2-test as contended in [7]. In the ordinary and correlation coefficient test [6], Conditions B and A_3 are easy to establish by virtue of (2.2).

(iii) *Multiple correlation coefficient test.* In Application VI of [6] take $\varphi = \rho^2$, $Y_n^2 = X_n/(1 + X_n)$ and conclude from p. 320 or [9] that

$$(3.3)\quad q_{\varphi n}(y) = \Gamma(\lambda)[\Gamma(c)\Gamma(\lambda - c)]^{-1}(1 - y)^{\lambda-c-1}y^{c-1}(1 - \varphi)^{\lambda}\ {}_2F_1(\lambda, \lambda; c; \varphi y),$$

where $2\lambda = n - 1$, $2c = p - 1$. Condition B follows from (2.1). Since we are not able to justify the use of (2.1) with y replaced by $y + (t/n)$, we resort to the following reasoning for establishing Condition A_3. It follows from (1.1) that the derivative of ${}_2F_1$ with respect to z is given by

$$(3.4)\quad {}_2F_1'(a, b; c; z) = (ab/c)\ {}_2F_1(a + 1, b + 1; c + 1; z).$$

From (2.1) and (3.4) we conclude

$$(3.5)\quad {}_2F_1'(\lambda, \lambda; c; z)/{}_2F_1(\lambda, \lambda; c; z) \sim_{\lambda\to\infty} \lambda^{\frac{3}{2}}(1 + \lambda)^{-\frac{1}{2}}[(z)^{\frac{1}{2}}(1 - (z)^{\frac{1}{2}})]^{-1}.$$

Condition A_3 follows by expanding $\ln q_{\varphi n}(y + (t/n))$ and using (3.3)–(3.5).

4. Acknowledgment. Thanks are due to Professor R. A. Wijsman for his guidance during the preparation of the dissertation of which this paper is a part.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] COX, D. R. (1952). Sequential tests for composite hypotheses. *Proc. Cambridge Phil. Soc.* **48** 290–299.

[3] FORSYTH, A. R. (1933). *A Treatise on Differential Equations* (5th ed.). Macmillan, London.

[4] GHOSH, J. K., HALL, W. J., and WIJSMAN, R. A. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[5] HOEL, P. G. (1955). On a sequential test for the general linear hypothesis. *Ann. Math. Statist.* **26** 136–139.

[6] IFRAM, A. F. (1965). On the asymptotic behavior of densities with applications to sequential analysis. *Ann. Math. Statist.* **36** 615–637.

[7] JACKSON, J. E. and BRADLEY, R. A. (1961). Sequential χ^2- and T^2-tests. *Ann. Math. Statist.* **32** 1063–1077.

[8] JOHNSON, N. L. (1953). Some notes on the applications of sequential methods in the analysis of variance. *Ann. Math. Statist.* **24** 614–623.

[9] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[10] RAY, W. D. (1957). A proof that the sequential probability ratio test (SPRT) of the general linear hypothesis terminates with probability unity. *Ann. Math. Statist.* **28** 521–523.

[11] SLATER, L. J. (1960). *Confluent Hypergeometric Functions.* Cambridge Univ. Press.

[12] WALD, A. (1947). *Sequential Analysis.* Wiley, New York.

[13] WATSON, G. N. (1912). Asymptotic expansions of hypergeometric function. *Cambridge Phil. Soc. Trans.* **22** 277–308.