King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 422

June 2011

The Cumulative Distribution Function of the Ratio of Correlated Sample Variances

M. Hafidz Omar, Anwar H. Joarder, and Muhammad Riaz

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: c-math@kfupm.edu.sa

The Cumulative Distribution Function of the Ratio of Correlated Sample Variances

M. Hafidz Omar, Anwar H. Joarder, and Muhammad Riaz
Department of Mathematics and Statistics
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia

Emails: omarmh@kfupm.edu.sa, anwarj@kfupm.edu.sa, riazm@kfupm.edu.sa

Abstract The probability density function of the ratio of correlated sample variances is important in testing equality of variances under correlation. In this technical note, we derive the median and the cumulative distribution function of the distribution of the ratio of correlated sample variances.

1. Introduction

Let S_1^2 and S_2^2 be sample variances based on a sample of size $N = m+1$ from a bivariate normal distribution with unknown means, unknown variances σ_1^2 and σ_2^2, and correlation coefficient ρ $(-1 < \rho < 1)$. The ratio of sample variances S_1^2 / S_2^2 or $H = U/V$ where $U = mS_1^2/\sigma_1^2$ and $V = mS_2^2/\sigma_2^2$ is important in testing equality of true variances under correlation. In this technical report , we study the median and Cumulative Distribution Function (CDF) of the distribution of H. We also provide the percentage points of H.

2. Mathematical Preliminaries

In what follows we will be using the following formulae.

We will be using the following product of k consecutive integers:

$$k_{\{a\}} = k(k+1)\cdots(k+a-1), \quad k^{\{a\}} = k(k-1)\cdots(k-a+1), \tag{2.1}$$

with $k_{\{0\}} = 1$ and $k^{\{0\}} = 1$.

The hypergeometric function ${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z)$ is defined by

$${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_{\{k\}}(a_2)_{\{k\}}\cdots(a_p)_{\{k\}}}{(b_1)_{\{k\}}(b_2)_{\{k\}}\cdots(b_q)_{\{k\}}} \frac{z^k}{k!}, \tag{2.2}$$

(Gradshteyn and Ryzhik, 1994), #9.14, p 1071)

The incomplete beta function is given by

$$B(x;\alpha,\beta)=\int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt, \tag{2.3}$$

which can also be represented by

$$B(x;\alpha,\beta)=\frac{x^{\alpha}}{\alpha}\,{}_2F_1(\alpha,1-\beta;\alpha+1;x). \tag{2.4}$$

The CDF of central variance ratio distribution with m and n degrees of freedom is given by

$$\Upsilon(z:m,m)=B^{-1}\left(\frac{m}{2},\frac{n}{2}\right)B\left(\frac{mz}{mz+n};\frac{m}{2},\frac{n}{2}\right). \tag{2.5}$$

Theorem 2.1 The joint density function of the random variables U and V is given by

$$f_{U,V}(u,v)\propto \exp\left(-\frac{u+v}{2-2\rho^2}\right)\ {}_0F_1\left(\frac{m}{2};\frac{\rho^2uv}{(2-2\rho^2)^2}\right), \tag{2.6}$$

where $m>2$ and $-1<\rho<1$ and ${}_0F_1(;b;z)$ is defined in (2.2).

3. The Probability Density Function

The following density function of H was derived by Bose (1935) and Finney (1938).

Theorem 3.1 The density function of $H=U/V$ is given by

$$f_H(h)\propto\frac{h^{(m-2)/2}}{(1+h)^m}\left(1-\frac{4\rho^2h}{(1+h)^2}\right)^{-(m+1)/2},\ h>0 \tag{3.1}$$

where $m>2$ and $-1<\rho<1.$

We will denote the distribution of H by $F(m,m;\rho)$. For two common choices of significance levels (say $\alpha=0.01,\ 0.05$) used in common applications, the critical values based on the correlated $F(m,m;\rho)$ are provided in Appendix Tables 1 and 2 for some representative values of m and ρ. These values are obtained through Monte Carlo simulations based on 10^6 repetitions.

4. The Median of the Distribution

The median of the distribution of H is given by $\tilde{h}$ where $P(H<\tilde{h})=1/2.$

Theorem 4.1 Let H have a correlated F distribution with density function (3.1). Then

$$P(H<1)=1/2. \tag{4.1}$$

Proof. By definition, it follows from (3.1) that

$$P(H<1)=\int_0^1 \frac{2^{m-1}(1-\rho^2)^{m/2}}{B(1/2,m/2)}\frac{h^{(m-2)/2}}{(1+h)^m}\left(1-\frac{4\rho^2 h}{(1+h)^2}\right)^{-(m+1)/2} dh, \tag{4.2}$$

Completing the integral in (4.2) followed by some algebraic simplification, we have (4.1).

That is, for the correlated F distribution, the median $\tilde{h}$ is 1.

In the context of reliability, the stress-strength model describes the life of a component which has a random strength V and is subjected to random stress U. The component fails if the stress (U) applied to it exceeds the strength (V) and the component will function satisfactorily whenever $V>U$. Thus $P(U<V)$ is a measure of component reliability.

Theorem 4.2 Let the random variables U and V have a bivariate chi-square distribution with density function (2.6). Then the reliability of the distribution is given by $P(U<V)=1/2.$

Proof. Since $H=U/V$, $P(U<V)=P(H<1)$ which is $\frac{1}{2}$ by Theorem 4.1.

5. The CDF of the Distribution

Theorem 5.1 Let H have the correlated variance ratio distribution given by (3.1). Then the Cumulative Distribution Function (CDF) of H is given by

$$\Upsilon(h;m,m,\rho)=\frac{\Gamma(m)(1-\rho^2)^{m/2}}{\Gamma^2(m/2)\Gamma((m+1)/2)} \times\sum_{k=0}^{\infty}\Gamma\left(k+\frac{m+1}{2}\right)B\left(\frac{h}{1+h};k+\frac{m}{2};k+\frac{m}{2}\right)\frac{(4\rho^2)^k}{k!}du, \tag{5.1}$$

where $-1<\rho<1$ and $m>2.$

Proof. The cumulative distribution function of H is given by

$$\Upsilon(h;m,m,\rho)=\frac{2^{m-1}(1-\rho^2)^{m/2}}{B\left(\frac{1}{2},\frac{m}{2}\right)}\int_{y=0}^{h}\frac{y^{(m-2)/2}}{(1+y)^m}\left(1-\frac{4\rho^2 y}{(1+y)^2}\right)^{-(m+1)/2}dy.$$

Completing the above integral followed by a bit of algebraic manipulation we have (5.1).

In case $\rho=0$, then $\Upsilon(h;m,m,\rho)$ reduces to $\Upsilon(h;m,m)$ which is the CDF of central variance ratio with m and m degrees of freedom (cf. 2.5).

Acknowledgements

The first two authors gratefully acknowledge the excellent research facility provided by King Fahd University of Petroleum & Minerals especially through the Project FT 100007. The third author acknowledges the excellent research facilities available at King Fahd University of Petroleum & Minerals.

References

Bose, S.S. (1935). On the distribution of the ratio of variances of two samples drawn from a given normal bivariate correlated population, *Sankhya*, 2, 65-72.

Finney, D.J. (1938). The distribution of the ratio of the estimates of the two variances in a sample from a bivariate normal population, *Biometrika*, 30, 190-192.

Gradshteyn, I.S. and Ryzhik, I.M. (1995). *Table of Integrals*, *Series and Products*, Academic Press.

Table 1: Percentile Points at $\alpha = 0.05$

						ρ_{yx}						
n	***0.0***	***0.1***	***0.2***	***0.3***	***0.4***	***0.5***	***0.6***	***0.7***	***0.8***	***0.9***	***0.95***	***0.99***
2	160.75	158.41	152.85	146.20	134.992	120.31	101.79	83.47	59.52	32.05	17.36	4.964
3	18.88	18.69	18.24	17.39	16.38	14.57	12.82	10.66	7.97	5.03	3.35	1.78
4	9.290	9.204	8.927	8.595	8.056	7.409	6.548	5.577	4.443	3.072	2.288	1.465
5	6.399	6.355	6.205	5.969	5.633	5.202	4.698	4.066	3.350	2.457	1.923	1.349
6	5.039	5.037	4.917	4.758	4.510	4.199	3.808	3.355	2.818	2.154	1.744	1.289
7	4.278	4.266	4.161	4.056	3.847	3.612	3.308	2.950	2.506	1.966	1.634	1.251
8	3.782	3.759	3.701	3.586	3.424	3.236	2.973	2.671	2.302	1.848	1.559	1.224
9	3.440	3.421	3.355	3.267	3.126	2.960	2.744	2.479	2.160	1.759	1.502	1.203
10	3.176	3.163	3.112	3.033	2.908	2.755	2.569	2.334	2.053	1.694	1.461	1.188
11	2.980	2.973	2.922	2.841	2.739	2.603	2.427	2.218	1.960	1.641	1.429	1.175
12	2.824	2.803	2.759	2.690	2.607	2.476	2.316	2.133	1.892	1.596	1.398	1.165
13	2.687	2.679	2.639	2.579	2.480	2.370	2.225	2.051	1.837	1.561	1.377	1.156
14	2.578	2.569	2.527	2.472	2.391	2.286	2.151	1.993	1.787	1.529	1.358	1.149
15	2.482	2.473	2.440	2.385	2.314	2.213	2.086	1.935	1.747	1.502	1.341	1.142
16	2.406	2.389	2.366	2.313	2.238	2.149	2.035	1.890	1.709	1.480	1.325	1.136
17	2.336	2.326	2.296	2.249	2.184	2.094	1.985	1.847	1.680	1.459	1.312	1.131
18	2.277	2.260	2.239	2.192	2.129	2.044	1.939	1.812	1.649	1.443	1.302	1.126
19	2.219	2.213	2.180	2.144	2.081	2.006	1.899	1.777	1.625	1.425	1.290	1.122
20	2.168	2.164	2.136	2.097	2.042	1.962	1.867	1.748	1.603	1.410	1.280	1.118
30	1.863	1.855	1.837	1.813	1.767	1.717	1.647	1.564	1.458	1.316	1.218	1.093
50	1.607	1.604	1.591	1.573	1.547	1.512	1.464	1.405	1.331	1.231	1.161	1.070
100	1.394	1.391	1.385	1.373	1.356	1.334	1.304	1.268	1.222	1.156	1.110	1.048

Table 2: Percentile Points at $\alpha = 0.01$

n	ρ_{yx} 0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
2	4002.7	3922.9	3746.9	3741.2	3421.7	3059.3	2495.2	2045.3	1492.4	760.8	398.8	83.6
3	97.63	99.48	95.04	90.33	82.77	74.52	64.12	51.79	36.34	20.46	11.40	3.67
4	29.66	29.17	28.59	26.79	25.04	22.58	19.42	16.02	11.75	7.04	4.46	2.07
5	16.02	15.85	15.61	14.74	13.69	12.34	10.86	8.97	6.98	4.42	3.05	1.68
6	10.93	10.91	10.60	10.14	9.56	8.67	7.63	6.47	5.09	3.43	2.49	1.52
7	8.439	8.434	8.183	7.862	7.395	6.808	6.022	5.181	4.141	2.898	2.186	1.434
8	6.997	6.918	6.779	6.548	6.139	5.686	5.101	4.397	3.566	2.596	2.002	1.375
9	6.029	5.984	5.850	5.643	5.321	4.956	4.460	3.891	3.193	2.376	1.878	1.333
10	5.370	5.311	5.197	5.047	4.758	4.420	4.014	3.525	2.936	2.222	1.782	1.303
11	4.854	4.832	4.724	4.559	4.340	4.040	3.679	3.239	2.744	2.107	1.714	1.278
12	4.480	4.430	4.343	4.208	4.009	3.740	3.434	3.052	2.574	2.012	1.656	1.260
13	4.159	4.136	4.049	3.923	3.738	3.514	3.241	2.872	2.458	1.941	1.611	1.244
14	3.911	3.882	3.830	3.704	3.540	3.316	3.054	2.738	2.354	1.879	1.577	1.230
15	3.684	3.673	3.615	3.501	3.359	3.150	2.920	2.628	2.262	1.826	1.544	1.219
16	3.522	3.500	3.452	3.348	3.195	3.021	2.803	2.527	2.189	1.780	1.517	1.209
17	3.369	3.358	3.288	3.215	3.085	2.910	2.704	2.444	2.129	1.742	1.491	1.200
18	3.255	3.231	3.173	3.082	2.959	2.798	2.608	2.372	2.073	1.711	1.473	1.192
19	3.134	3.120	3.067	2.989	2.865	2.725	2.534	2.299	2.026	1.679	1.452	1.185
20	3.029	3.010	2.962	2.899	2.787	2.627	2.467	2.242	1.981	1.649	1.435	1.178
30	2.428	2.409	2.383	2.338	2.254	2.169	2.046	1.901	1.723	1.486	1.330	1.137
50	1.960	1.958	1.937	1.904	1.861	1.801	1.721	1.623	1.505	1.349	1.239	1.102
100	1.602	1.597	1.585	1.566	1.541	1.504	1.457	1.401	1.329	1.230	1.160	1.069