On the Distribution of the Inverted Linear Compound of Dependent F-Variates and its Application to the Combination of Forecasts

Kuo-Yuan Liang, Jack C. Lee & Kurt S.H. Shao

Polaris Research Institute and Department of Economics, National Taiwan University, Taiwan
National Chiao Tung University, Taiwan
Polaris Research Institute, Taiwan

Published online: 01 Dec 2006.

To cite this article: Kuo-Yuan Liang, Jack C. Lee & Kurt S.H. Shao (2006) On the Distribution of the Inverted Linear Compound of Dependent F-Variates and its Application to the Combination of Forecasts, Journal of Applied Statistics, 33:9, 961-973, DOI: 10.1080/02664760600744330

To link to this article: http://dx.doi.org/10.1080/02664760600744330
On the Distribution of the Inverted Linear Compound of Dependent F-Variates and its Application to the Combination of Forecasts

KUO-YUAN LIANG* , JACK C. LEE** & KURT S.H. SHAO†

*Polaris Research Institute and Department of Economics, National Taiwan University, Taiwan,
**National Chiao Tung University, Taiwan, †Polaris Research Institute, Taiwan

ABSTRACT This paper establishes a sampling theory for an inverted linear combination of two dependent F-variates. It is found that the random variable is approximately expressible in terms of a mixture of weighted beta distributions. Operational results, including rth-order raw moments and critical values of the density are subsequently obtained by using the Pearson Type I approximation technique. As a contribution to the probability theory, our findings extend Lee & Hu’s (1996) recent investigation on the distribution of the linear compound of two independent F-variates. In terms of relevant applied works, our results refine Dickinson’s (1973) inquiry on the distribution of the optimal combining weights estimates based on combining two independent rival forecasts, and provide a further advancement to the general case of combining three independent competing forecasts. Accordingly, our conclusions give a new perception of constructing the confidence intervals for the optimal combining weights estimates studied in the literature of the linear combination of forecasts.

KEY WORDS: Combining weights, critical values, error-variance minimizing criterion, inverted F-variates, Pearson Type I approximation

Introduction

In this paper, we study the distribution of an inverted linear compound of dependent F-variates in the form:

\[
\frac{1}{1 + a_1 F_1(T, T) + a_2 F_2(T, T)}
\]

with degrees of freedom as indicated. Here, the two constants \( a_1 \) and \( a_2 \) lie in the interval \((0, 1]\). This distribution is useful in constructing confidence intervals of the
minimum variance weights that can be attached to the components of the linear composite forecasts.

From Reid (1969), Dickinson (1973) or Newbold & Granger (1974), it is well known that given a history of unbiased forecast errors for \( k \) \((k \geq 2)\) models, under the error-variance minimizing criterion, the optimal weighting vector \((W)\) of the combined forecasts becomes

\[
W = \frac{\sum^{-1} u}{u' \sum^{-1} u}
\]  
(2)

where \( u \) is a \((k \times 1)\) vector of ones and \( \sum \) a \((k \times k)\) positive definite covariance matrix of forecasting errors between the \( k \) models.

It is worth noting that despite the popularity of equation (2), very little is known about the sampling properties of its estimator. A notable exception to this issue is the work of Dickinson (1973). When \( k = 2 \), based on the maximum likelihood estimator \((S)\) of \( \sum \) with zero off-diagonal elements and normally distributed forecasting errors, Dickinson (1973) demonstrated that each component of the estimated weight vector \( \hat{W} \) is expressible as a weighted beta or beta (if homoscedasticity is further imposed) distribution.

Although the combining procedure may involve more than two competing forecasts, we will restrict our attention to the \( k = 3 \) set-up with \( \sum \) having zero off-diagonal elements and normally distributed forecasting errors only. This proves necessary as use of the general \( k \times k \) set-up is technically difficult. Our restricted set-up hence extends the initial work of Dickinson (1973).

From the statistical viewpoint, a corollary of Dickinson’s (1973) result is that an inverted \( F \)-variante of the form: \( \frac{1}{1+aF(T, T)} \) where \( a \) is an arbitrary constant, is expressible as a weighted beta \((a \neq 1)\) or as a beta \((a = 1)\) distribution, i.e.,

\[
\frac{1}{1+aF(T, T)} \sim \text{weighted beta}
\]  
(3)

Another relevant theoretical contribution to our investigation is the work of Lee & Hu (1996). According to them, an arbitrary linear combination of two independent \( F \)-variates can be expressed approximately as a suitable constant \((c)\) times an \( F \) density function, i.e.,

\[
a_1F(u_1, u_2) + a_2F(u_1, u_2) \sim cF(m_1, m_2)
\]  
(4)

where \( m_1 \) and \( m_2 \) are two positive constants.

Using the restricted set-up (detailed above), this paper derives the sampling distribution of the estimated combining weights. To achieve this goal, we begin in the next section with a reformulation of equation (2) by replacing \( \sum \) with \( S \), and show that each estimated weight is an inverted linear compound of dependent \( F \)-variates. We then relax the independence assumption on equation (4), and demonstrate in the third section that an expression of the right-hand side of equation (4) approximately still holds. Using this result and random variables transformation techniques, we also show in the third section that the distribution of an inverted linear compound of dependent \( F \)-variates of equation (1) is approximately expressible in terms of a mixture of weighted beta distributions. Additionally, we conduct extensive simulations to assess the accuracy of these approximations. Our results thus generalize those of Lee & Hu (1996) as well as Dickinson (1973). A notable implication of our theoretical results to the equal weighting scheme is elaborated as well.
Owing to the complexity of the derived distribution, the fourth section presents several operational results, including rth-order raw moments and critical values of the density based on the Pearson Type I approximation technique (Johnson et al., 1963). The fifth section summarizes our findings and indicates future research directions.

Model and Related Results

Consider equation (2) in the restricted case where \( \sum = \text{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33}) \). In practice, the parameters \( \sigma_{ii} \) are unknown, and \( \sum \) is estimated by:

\[
S = \text{diag} \left( \frac{\sum_{t=1}^{T} e_{1t}^2}{T}, \frac{\sum_{t=1}^{T} e_{2t}^2}{T}, \frac{\sum_{t=1}^{T} e_{3t}^2}{T} \right)
\]

where \( e_{it} \) is the error in the \( t \)-th forecast value, using the \( i \)-th forecasting method. Assuming \( e_{it} \) to be normally distributed with zero mean and variance \( \sigma_{ii} \), the maximum likelihood estimator of \( T \sum \) is given by:

\[
TS = \text{diag} \left( \sum_{t=1}^{T} e_{1t}^2, \sum_{t=1}^{T} e_{2t}^2, \sum_{t=1}^{T} e_{3t}^2 \right)
\]

It follows that:

\[
\sum_{t=1}^{T} e_{it}^2 \sim \sigma_{ii} \chi^2(T)
\]

From above, the \( i \)-th weight in equation (2) is estimated by:

\[
\hat{w}_i = \frac{1}{(1/\sum_{t=1}^{T} e_{1t}^2) + (1/\sum_{t=1}^{T} e_{2t}^2) + (1/\sum_{t=1}^{T} e_{3t}^2)}, \quad i = 1, 2, 3
\]

Based on equation (7), each estimated weight can thus be written as an inverted linear compound of dependent \( F \)-variates of the form stated in equation (1). For example, equation (7) implies that another expression for the first estimated weight is:

\[
\hat{w}_1 = \frac{1}{1 + (\sum_{t=1}^{T} e_{1t}^2/\sum_{t=1}^{T} e_{2t}^2) + (\sum_{t=1}^{T} e_{1t}^2/\sum_{t=1}^{T} e_{3t}^2)}
\]

Using equation (6) in equation (8), we can conclude that:

\[
\hat{w}_1 \sim \frac{1}{1 + a_1 F_1(T, T) + a_2 F_2(T, T)}
\]

with degrees of freedom as indicated, and \( a_1 = \sigma_{11}/\sigma_{22}, \ a_2 = \sigma_{11}/\sigma_{33} \) are two positive constants. Comparing this set-up with equation (1), we see that \( \sigma_{11} \leq \sigma_{22} \) and \( \sigma_{11} \leq \sigma_{33} \) are assumed here for illustrational convenience. Similar expressions for \( \hat{w}_2 \) and \( \hat{w}_3 \) can also be readily derived. Since \( \sum_{t=1}^{T} e_{1t}^2 \) appears in the second and third denominator terms of the right-hand expression of equation (8), it can be shown, that if \( T > 4 \) these
two \( F \)-variates are indeed dependent and their correlation is given by:

\[
\text{corr}(F_1, F_2) = \frac{T - 4}{2(T - 1)}
\]  

(10)

See Appendix A.1 for the proof.

**Theorems, Simulations and Implications**

Having verified the dependency between \( F_1 \) and \( F_2 \) in equation (10), we now turn our attention to the problem of finding the probability density of its linear compound of \( a_1 F_1(T, T) \) and \( a_2 F_2(T, T) \). As the following result indicates, this linear compound can be approximated by a constant (\( \eta \)) times an \( F(m_1, m_2) \) variate, with the degrees of freedom as indicated.

**Theorem 1**

\[
a_1 F_1(T, T) + a_2 F_2(T, T) \sim \eta F(m_1, m_2)
\]  

(11)

where the right-hand expression of equation (11) comes from the denominator of equation (9), and the parameters \( \eta, m_1 \) and \( m_2 \) can be expressed in explicit forms in terms of \( a_1, a_2 \) and \( T \). Specifically (Lee & Hu, 1996),

\[
\eta = \frac{2A^2C - 2AB^2}{A^2B + 3AC - 4B^2}
\]

\[
m_1 = \frac{4A^2C - 4AB^2}{AB^2 - 2A^2C + BC}
\]

and

\[
m_2 = \frac{2A^2B + 6AC - 8B^2}{A^2B + AC - 2B^2}
\]  

(12)

where

\[
A = \frac{(a_1 + a_2)T}{T - 2}
\]

\[
B = \frac{T(T + 2)}{T - 2} \left( a_1^2 + a_2^2 + \frac{2a_1a_2}{T - 2} \right)
\]

\[
C = \frac{T(T + 2)(T + 4)}{(T - 2)(T - 4)} \left( a_1^3 + a_2^3 + \frac{3a_1^2a_2}{T - 6} + \frac{3a_1a_2^2}{T - 2} \right)
\]  

(13)

and \( T > 6 \).

Similar to Lee & Hu (1996), we conduct an extensive simulation study to assess the accuracy of this approximation. The results of our study are summarized in Table 1.

In the simulation, we conduct 15,000 runs for each linear compound of the form \( a_1 F_1(T, T) + a_2 F_2(T, T) \) and compute the probabilities of exceeding the 1%, 5% and 10% points. From Table 1, we see that the approximation to the assigned probability
Distribution of the Inverted Linear Compound of Dependent F-Variates

Table 1. Simulated results for Theorem 1

<table>
<thead>
<tr>
<th>Linear compound</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1(10,10) + F_2(10,10)$</td>
<td>0.0101</td>
<td>0.0510</td>
<td>0.1019</td>
</tr>
<tr>
<td>$F_1(10,10) + 0.5F_2(10,10)$</td>
<td>0.0101</td>
<td>0.0510</td>
<td>0.1024</td>
</tr>
<tr>
<td>$0.3F_1(10,10) + 0.01F_2(10,10)$</td>
<td>0.0106</td>
<td>0.0491</td>
<td>0.0994</td>
</tr>
<tr>
<td>$0.7F_1(10,10) + 0.01F_2(10,10)$</td>
<td>0.0103</td>
<td>0.0493</td>
<td>0.0993</td>
</tr>
<tr>
<td>$0.9F_1(10,10) + 0.01F_2(10,10)$</td>
<td>0.0097</td>
<td>0.0491</td>
<td>0.1000</td>
</tr>
<tr>
<td>$F_1(10,10) + 0.01F_2(10,10)$</td>
<td>0.0098</td>
<td>0.0500</td>
<td>0.0984</td>
</tr>
<tr>
<td>$0.9F_1(10,10) + 0.7F_2(10,10)$</td>
<td>0.0106</td>
<td>0.0507</td>
<td>0.1039</td>
</tr>
<tr>
<td>$0.9F_1(30,30) + 0.01F_2(30,30)$</td>
<td>0.0098</td>
<td>0.0493</td>
<td>0.0978</td>
</tr>
<tr>
<td>$F_1(30,30) + 0.01F_2(30,30)$</td>
<td>0.0098</td>
<td>0.0493</td>
<td>0.0972</td>
</tr>
</tbody>
</table>

Table entries are the simulated probabilities in the right-hand tail of the listed linear compound of dependent $F$-variates.

$(\alpha)$ in the right-hand tail of the listed linear compound of dependent $F$-variates is generally quite accurate.

By virtue of Theorem 1, $\hat{w}_1$ in equation (9) can thus be reasonably approximated by:

$$\hat{w}_1 \sim \frac{1}{1 + \eta F(m_1, m_2)}$$

(14)

Likewise, similar expressions for $\hat{w}_2$ and $\hat{w}_3$ can be obtained. Using the approximations derived in equation (14), we are ready to apply the variable transformation techniques to derive the probability density of $f(\hat{w}_i) (i = 1, 2, 3)$.

**Theorem 2**

Let $e_i(t; i = 1, 2, 3; t = 1, 2, \ldots, T)$ be the error in the $t$th forecast value using the $i$th forecasting model. Assume at a particular point of time, $\epsilon^t = (e_{1t}, e_{2t}, e_{3t}) \sim N(0, \Sigma)$ with $\Sigma = \text{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33})$, then under the error-variance minimizing criterion, the distribution of the optimal combining weight estimator $\hat{w}_i$ is approximately a mixture of beta random variables with the probability density function of the form:

$$f(\hat{w}_i) = \frac{(1 - b)^{m_2/2}}{B(m_2/2, m_1/2)} \sum_{j=0}^{\infty} \left[ C_j^{m+j-1} b^j B(m_2/2 + j, \frac{m_1}{2}) \text{Beta}(m_2/2 + j, \frac{m_1}{2}) \right]$$

(15)

where $C_j^{m+j-1} = \frac{(m+j-1)!}{j!(m-1)!}$, $b = 1 - \frac{m_2}{m_1}$, $B(p, q)$ is a beta function and $\text{Beta}(p, q)$ is a beta density function with parameters $p, q$, respectively.

**Proof**

For the proof see Appendix A.2.

The following two theorems show that the condition $|b| < 1$ is sufficient for the integrability of $f(\hat{w}_i)$ over $[0, 1]$, the satisfaction of $\int_0^1 f(\hat{w}_i) d\hat{w}_i = 1$, and the existence of the $r$th order raw moment.
Theorem 3
If \(|b| < 1\), then \(\int_0^1 f(\hat{w}_i)\,d\hat{w}_i = 1\).

Proof
For the proof see Appendix A.3.

Theorem 4
If \(|b| < 1\), then the \(r\)th-order raw moment of \(f(\hat{w}_i)\) exists.

Proof
For the proof see Appendix A.4.

A notable implication of the condition \(|b| < 1\) is elaborated as follows. Suppose the matrix \(\sum\) in equation (2) is further restricted to \(\sum = \text{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33})\) and \(\sigma_{11} = \sigma_{22} = \sigma_{33}\). It would immediately appear that this case would generate \(w_1 = w_2 = w_3 = 1/3\). The case in which each forecast receives equal weight is of particular interest, because it may be reasonable in many realistic applications.

More specifically, the usual rationale for the equal weighting scheme is as follows. First, ‘if (a) there is only a small data base and/or (b) the error covariance structure is not stationary’ (Bunn, 1986, p. 152), then specifying \(\sum\) as an unrestricted real symmetric positive definite matrix tends to cause the robustness problems due to poor estimation of its elements. A resolution is therefore suggested to specify the matrix \(\sum\) in our restricted setup as \(\text{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33})\). Second, if no information is known or no reason to believe a priori on the relative accuracy of the competing forecasts, an even more extreme response is to further impose the constraint \(\sigma_{11} = \sigma_{22} = \sigma_{33}\) into the above diagonal setting and utilize the equal weighting scheme (Bunn, 1986).

This extreme case means that \(a_1 = a_2 = 1\) in equation (9). Substituting \(a_1 = a_2 = 1\) for \(b\) in equation (15) and using the condition \(|b| < 1\) produces:

\[
0 < \frac{4(T - 6)(3T^2 - 10T - 10)}{(T - 2)(3T^2 - 16T + 28)} < 2
\]

Significantly, the above inequality holds only when \(T = 7, 8, 9\). Therefore, as a practical matter, the existing conditions of \(f(\hat{w}_i)\) and its \(r\)th-order moment in this particular equal weighting scheme are extremely hard to satisfy. A cautious approach is suggested when applying this method, where other sources also share this view (Bunn, 1986; Winkler & Clemen, 1992).

Theorem 4 gives the following corollary.

Corollary 1
Each \(r\)th-order moment of \(f(\hat{w}_i)\) is expressible as a monotonically decreasing sequence.

Proof
For the proof see Appendix A.5.

To check the validity of the properties expressed in Theorems 2, 3, 4 and Corollary 1, the raw moments of \(\hat{w}_i\) up to the fourth-order with sample sizes 10, 30 and 100 are studied.
s separately. Because each of the three weights studied leads to similar conclusions, only the numerical results of the first weight ($w_1$) are reported in Table 2. Three important points are noted as follows. First, for each reported inverted linear compound with three different sample sizes, the condition $|b| < 1$ is satisfied. For example, for the case, $1/(1 + 0.7 F_1 + 0.01 F_2)$ with sample sizes 10, 30, and 100, $|b| = 0.5804, 0.4999,$ and 0.4971, respectively. Second, the numerical results are consistent with Corollary 1, displaying a monotonically decreasing sequence pattern. Third, all $E(\hat{w}_1)$ entries have a downward bias, i.e. $E(\hat{w}_1) = E(1/(1 + a_1 F_1 + a_2 F_2)) < (1/(1 + a_1 + a_2)) = w_1$. However, the magnitude of this downward bias shrinks as the sample size increases, implying that $\hat{w}_1$ tends to be asymptotically unbiased for $w_1$.

**Pearson Type I Approximation**

Although the density of $\hat{w}_1$ has been derived in the previous section, the critical values for interval estimation and hypothesis testing purposes are still extremely hard to obtain. However, since the moments are available, we can approximate the distribution of $\hat{w}_1$ by the Pearson type I distribution which is defined as (Lee & Hu, 1996)

$$f(x) = [\beta(a + 1, b + 1)(\sigma_1 - \sigma_0)^{a+b+1}]^{-1}(x - \sigma_0)^a(\sigma_1 - x)^b$$

where $\sigma_0 \leq x \leq \sigma_1$, $a, b \in \mathbb{R}$.

**Table 2.** $w_1$ and the raw moments of $\hat{w}_1$ up to the fourth-order

<table>
<thead>
<tr>
<th>$\hat{w}_1$</th>
<th>$w_1$</th>
<th>$E\hat{w}_1$</th>
<th>$E\hat{w}_1^2$</th>
<th>$E\hat{w}_1^3$</th>
<th>$E\hat{w}_1^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.970874</td>
<td>0.964700</td>
<td>0.931100</td>
<td>0.899077</td>
<td>0.868543</td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.997009</td>
<td>0.996307</td>
<td>0.992634</td>
<td>0.988980</td>
<td>0.985345</td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.763359</td>
<td>0.744231</td>
<td>0.567644</td>
<td>0.441612</td>
<td>0.349256</td>
</tr>
<tr>
<td>$1/(1 + 0.7F_1 + 0.01F_2)$</td>
<td>0.584795</td>
<td>0.576654</td>
<td>0.354011</td>
<td>0.228156</td>
<td>0.152900</td>
</tr>
<tr>
<td>$1/(1 + 0.9F_1 + 0.01F_2)$</td>
<td>0.523561</td>
<td>0.520868</td>
<td>0.293674</td>
<td>0.156437</td>
<td>0.095475</td>
</tr>
<tr>
<td>$1/(1 + F_1 + 0.01F_2)$</td>
<td>0.497512</td>
<td>0.497211</td>
<td>0.269679</td>
<td>0.156437</td>
<td>0.095745</td>
</tr>
<tr>
<td>$1/(1 + 0.9F_1 + 0.7F_2)$</td>
<td>0.384618</td>
<td>0.382487</td>
<td>0.162398</td>
<td>0.075405</td>
<td>0.037481</td>
</tr>
<tr>
<td>$1/(1 + F_1 + 0.5F_2)$</td>
<td>0.400000</td>
<td>0.397275</td>
<td>0.174562</td>
<td>0.083274</td>
<td>0.042524</td>
</tr>
<tr>
<td>Sample size 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.970874</td>
<td>0.968968</td>
<td>0.939000</td>
<td>0.910054</td>
<td>0.882093</td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.997009</td>
<td>0.996798</td>
<td>0.993607</td>
<td>0.990427</td>
<td>0.987259</td>
</tr>
<tr>
<td>$1/(1 + 0.3F_1 + 0.01F_2)$</td>
<td>0.763359</td>
<td>0.756860</td>
<td>0.577213</td>
<td>0.443337</td>
<td>0.342773</td>
</tr>
<tr>
<td>$1/(1 + 0.7F_1 + 0.01F_2)$</td>
<td>0.584795</td>
<td>0.584420</td>
<td>0.346758</td>
<td>0.210638</td>
<td>0.130330</td>
</tr>
<tr>
<td>$1/(1 + 0.9F_1 + 0.01F_2)$</td>
<td>0.523561</td>
<td>0.522624</td>
<td>0.281083</td>
<td>0.155174</td>
<td>0.087736</td>
</tr>
<tr>
<td>$1/(1 + F_1 + 0.5F_2)$</td>
<td>0.497512</td>
<td>0.497424</td>
<td>0.255416</td>
<td>0.134998</td>
<td>0.073263</td>
</tr>
<tr>
<td>Sample size 100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1/(1 + 0.02F_1 + 0.01F_2)$</td>
<td>0.970874</td>
<td>0.970324</td>
<td>0.941556</td>
<td>0.913665</td>
<td>0.886626</td>
</tr>
<tr>
<td>$1/(1 + 0.3F_1 + 0.01F_2)$</td>
<td>0.763359</td>
<td>0.761404</td>
<td>0.581016</td>
<td>0.444323</td>
<td>0.340508</td>
</tr>
<tr>
<td>$1/(1 + 0.7F_1 + 0.01F_2)$</td>
<td>0.584795</td>
<td>0.583922</td>
<td>0.343273</td>
<td>0.203130</td>
<td>0.120970</td>
</tr>
<tr>
<td>$1/(1 + 0.9F_1 + 0.01F_2)$</td>
<td>0.523561</td>
<td>0.523275</td>
<td>0.276255</td>
<td>0.147106</td>
<td>0.078992</td>
</tr>
<tr>
<td>$1/(1 + F_1 + 0.5F_2)$</td>
<td>0.497512</td>
<td>0.497487</td>
<td>0.249945</td>
<td>0.126783</td>
<td>0.064911</td>
</tr>
</tbody>
</table>
Table 3. Critical values for $\hat{w}_1$ based on the Pearson Type I approximation

<table>
<thead>
<tr>
<th>$\hat{w}_1$</th>
<th>$w_1$</th>
<th>$\alpha = 0.005$</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.025$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.95$</th>
<th>$\alpha = 0.975$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 0.995$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/(1 + 0.3F_1 + 0.01F_2)</td>
<td>0.763359</td>
<td>0.361607</td>
<td>0.406223</td>
<td>0.470608</td>
<td>0.524328</td>
<td>0.902542</td>
<td>0.918216</td>
<td>0.932646</td>
<td>0.940401</td>
</tr>
<tr>
<td>1/(1 + 0.7F_1 + 0.01F_2)</td>
<td>0.584795</td>
<td>0.185902</td>
<td>0.223393</td>
<td>0.279248</td>
<td>0.327863</td>
<td>0.810466</td>
<td>0.849467</td>
<td>0.892380</td>
<td>0.920052</td>
</tr>
<tr>
<td>1/(1 + F_1 + 0.5F_2)</td>
<td>0.497512</td>
<td>0.125202</td>
<td>0.142536</td>
<td>0.171396</td>
<td>0.199804</td>
<td>0.625318</td>
<td>0.668873</td>
<td>0.717642</td>
<td>0.749386</td>
</tr>
</tbody>
</table>

Sample size 10

<table>
<thead>
<tr>
<th>$\hat{w}_1$</th>
<th>$w_1$</th>
<th>$\alpha = 0.005$</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.025$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.95$</th>
<th>$\alpha = 0.975$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 0.995$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/(1 + 0.02F_1 + 0.01F_2)</td>
<td>0.970874</td>
<td>0.934439</td>
<td>0.939205</td>
<td>0.945599</td>
<td>0.950575</td>
<td>0.982675</td>
<td>0.984428</td>
<td>0.986272</td>
<td>0.987420</td>
</tr>
<tr>
<td>1/(1 + 0.3F_1 + 0.01F_2)</td>
<td>0.763359</td>
<td>0.584444</td>
<td>0.598797</td>
<td>0.614514</td>
<td>0.642355</td>
<td>0.860101</td>
<td>0.874330</td>
<td>0.888617</td>
<td>0.896978</td>
</tr>
<tr>
<td>1/(1 + 0.7F_1 + 0.01F_2)</td>
<td>0.584795</td>
<td>0.348087</td>
<td>0.371800</td>
<td>0.405815</td>
<td>0.435148</td>
<td>0.721016</td>
<td>0.744503</td>
<td>0.770518</td>
<td>0.787430</td>
</tr>
</tbody>
</table>

Sample size 30

<table>
<thead>
<tr>
<th>$\hat{w}_1$</th>
<th>$w_1$</th>
<th>$\alpha = 0.005$</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.025$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.95$</th>
<th>$\alpha = 0.975$</th>
<th>$\alpha = 0.99$</th>
<th>$\alpha = 0.995$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/(1 + 0.02F_1 + 0.01F_2)</td>
<td>0.970874</td>
<td>0.954803</td>
<td>0.956636</td>
<td>0.959200</td>
<td>0.961292</td>
<td>0.978035</td>
<td>0.979237</td>
<td>0.980566</td>
<td>0.981427</td>
</tr>
<tr>
<td>1/(1 + 0.3F_1 + 0.01F_2)</td>
<td>0.763359</td>
<td>0.651932</td>
<td>0.664615</td>
<td>0.682600</td>
<td>0.697790</td>
<td>0.814163</td>
<td>0.821643</td>
<td>0.829568</td>
<td>0.834512</td>
</tr>
</tbody>
</table>

Sample size 100

Table entries are the critical values with the probability $\alpha$ lying beneath.
In order to utilize the Pearson Type I approximation, we need the first four moments of $\hat{w}_i$, which can be obtained as demonstrated numerically in Table 2. Let $\mu = E(\hat{w}_i)$, $\mu_h = E(\hat{w}_i - \mu)^h$, $h = 2, 3, 4$ and $\beta_1 = \mu_2^2/\mu_3^2$, $\beta_2 = \mu_4/\mu_2^2$. Then the Pearson Type I distribution requires that $6 + 3\beta_1 - 2\beta_2 > 0$, $\beta_2 - \beta_1 - 1 > 0$.

Instead of computing from the density directly, we will make use of the tables produced by Johnson et al. (1963). For this purpose, we need the following double entry interpolations. Linear interpolation is often possible for $\sqrt{\beta_1}$, while second differences are needed for $\sqrt{\beta_2}$. This procedure allows us to interpolate first for $\beta_2$ at each of the nearest four values of $\sqrt{\beta_1}$. Furthermore, it also tabulates first $x_{-1}, x_0, x_1, x_2$, and finally to interpolate for $\sqrt{\beta_1}$, using the formula

$$x(\theta) = (1 - \theta)x_0 + \theta x_1 - \frac{1}{4} \theta (1 - \theta) [\Delta^2 x_0 + \Delta^2 x_1]$$

where $\theta$ is the appropriate fraction in the tabular interval.

Based on the Pearson Type I approximation, as briefed above, Table 3 gives critical values of $\hat{w}_i$ with $\alpha = 0.005, 0.01, 0.025, 0.05, 0.95, 0.975, 0.99, 0.995$ for several cases considered in Table 2. In reference to Table 3, two major results emerged. First, by picking up $\alpha = 0.025$ and $\alpha = 0.975$, it can be seen that with sample sizes 10, 30 and 100 the 95% interval estimates of $w_1$ for $\hat{w}_1 = 1/(1 + 0.3F_1 + 0.01F_2)$ lie in the interval $[0.470608, 0.918216]$, $[0.614514, 0.874330]$ and $[0.6826, 0.821643]$, respectively. Most importantly by using the data in Table 3, the same method also applies to the construction of interval estimates of $\hat{w}_1$ based on particular $\hat{w}_1$ and distinctive width considerations. Second, as expected, we note that, under the preassigned percentage, the larger are the sample sizes the narrower are the interval weight estimates.

**Conclusions**

Among methods of combining forecasts (Liang, 1992), the formula (2) proposed by Reid (1969), Dickinson (1973) or Newbold & Granger (1974) is perhaps the single most extensively used measure of the optimal weights. Despite the popularity of this formula, very little is known about the sampling properties of its estimator. Although Dickinson (1973) has studied this issue, it only dealt with the combination of two forecasts exhibiting no covariance between their errors. Dickinson (1973, p. 259) also mentioned that '...the exact derivation of confidence intervals for the weights ... of the combined forecasts is extremely complex when more forecasts, or covariance between errors, are introduced'.

In this paper, attention has been directed mainly to the combination of three forecasts exhibiting no covariance between their errors. With normally distributed forecasting errors, we show that each estimated weight is expressible as an inverted linear compound of dependent $F$-variates and has approximately a mixture of weighted beta distributions.

Operational results, including $r$th-order raw moments and critical values of the density are subsequently obtained by using the Pearson Type I approximation technique. As a contribution to the probability theory, our findings extend Lee & Hu’s (1996) recent investigation on the distribution of the linear compound of two independent $F$-variates. In terms of relevant applied works, our results refine Dickinson’s (1973) inquiry on the distribution of the optimal combining weights estimates based on combining two independent rival forecasts, and provide a further advancement to the general case of combining three independent competing forecasts. Accordingly,
this paper gives a new perception of constructing the confidence intervals for the optimal combining weights estimates studied in the literature of the linear combination of forecasts. A cautious approach is also suggested when applying the popular equal weighting combining method, because the existing conditions of \( f(\hat{w}_i) \) and its \( r \)-th order moment are practically very hard to satisfy.

In this paper, we have enlarged the forecasting error covariance matrix from Dickinson’s (1973) 2 \( \times \) 2 diagonal setting to 3 \( \times \) 3. We strongly hope that this can serve as a stepping stone into the studies based on the more general formulation of the matrix.

**Appendix: Proofs**

**Appendix A.1. Proof of equation (10)**

Suppose \( X_1, X_2 \) and \( X_3 \) are three independent chi-square distributed random variables with \( T \) degrees of freedom. By independence, we have

\[
f_{X_1X_2X_3}(x_1, x_2, x_3) = f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3)
\]

Using the following transformations of variables

\[
F_1 = \frac{X_1}{X_2}, \quad F_2 = \frac{X_1}{X_3}, \quad F_3 = X_1
\]

we obtain the joint probability density function of \( F_1, F_2 \) and \( F_3 \)

\[
f_{F_1F_2F_3}(f_1, f_2, f_3) = \frac{1}{\Gamma(T/2)^3 2^{3T/2}} e^{-(T/2)-1} f_1^{(T/2)-1} f_2^{(T/2)-1} f_3^{(3T/2)-1} e^{-(1+(1/f_1)+(1/f_2))f_3/2},
\]

and the joint probability density function of \( F_1 \) and \( F_2 \)

\[
f_{F_1F_2} = \frac{\Gamma(3T/2)}{\Gamma(T/2)^3} \frac{f_1^{(T/2)-1} f_2^{(T/2)-1}}{(1 + (1/f_1) + (1/f_2))^{(3T/2)}}
\]

It is easy to verify that

\[
E(F_{rj}^j) = \frac{\Gamma(T/2 + r_j)\Gamma(T/2 - r_j)}{[\Gamma(T/2)]^2}, \quad \forall r_j \in \mathbb{N}
\]

and

\[
E(F_{1r}^j F_{2r}^r) = \frac{\Gamma(T/2 + r)\Gamma(T/2 - r_1)\Gamma(T/2 - r_2)}{[\Gamma(T/2)]^3}, \quad r = r_1 + r_2
\]

Hence,

\[
E(F_j) = \frac{T}{T - 2}
\]
\[ E(F_j^2) = \frac{T(T+2)}{(T-2)(T-4)} \]
\[ E(F_1F_2) = \frac{T(T+2)}{(T-2)^2} \]

and

\[ \text{cov}(F_1, F_2) = \frac{2T}{(T-2)^2} \]
\[ \text{var}(F_j) = \frac{4T(T-1)}{(T-2)^2(T-4)} \]
\[ \text{corr}(F_1, F_2) = \frac{T-4}{2(T-1)} \]

Appendix A.2. Proof of Theorem 2

Since

\[ \hat{w}_i \sim \frac{1}{1 + a_1F_1 + a_2F_2} \sim \frac{1}{1 + \eta F(m_1, m_2)} \]

we have the following approximate probability density function of \( \hat{w}_i \)

\[ f(\hat{w}_i) = \frac{(1-b)^{m_2/2}}{B(m_2/2, m_1/2)} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j \hat{w}_i^{(m_2/2)+j-1} \]

Using a negative binomial expansion we can express \( f(\hat{w}_i) \) as

\[ f(\hat{w}_i) = \frac{(1-b)^{m_2/2}}{B(m_2/2, m_1/2)} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j \hat{w}_i^{(m_2/2)+j-1}(1 - \hat{w}_i)^{(m_1/2)-1} \]

or alternatively as

\[ f(\hat{w}_i) = \frac{(1-b)^{m_2/2}}{B(m_2/2, m_1/2)} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j B\left(\frac{m_2}{2} + j, \frac{m_1}{2}\right) \]

Appendix A.3. Proof of Theorem 3

\[ \int_0^1 f(\hat{w}_i) d\hat{w}_i = \int_0^1 \left[ \frac{(1-b)^{m_2/2}}{B(m_2/2, m_1/2)} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j B\left(\frac{m_2}{2} + j, \frac{m_1}{2}\right) \right] d\hat{w}_i \]
\[ (1 - b)^{m_2/2} \sum_{j=0}^{\infty} \left[ C_j^{(m+j)-1} b^j \frac{B((m_2/2) + j, m_1/2)}{B(m_2/2, m_1/2)} \right] \]

Consider a non-negative infinite series \( \{a_n\} \), and let

\[ R = \limsup_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| \]

If \( R < 1 \), then \( \{a_n\} \) is absolutely convergent, by the ratio test (Apostol, 1974, p.173). By the previous theorem, with

\[ R = \limsup_{n \to \infty} \left| \frac{m_2 + 2n}{2 + 2n} \right| = |b| \]

if \( R = |b| < 1 \), then

\[ \int_0^1 f(\hat{w}_i) d\hat{w}_i = 1 \]

**Appendix A.4. Proof of Theorem 4**

\[ E(\hat{w}_i') = \int_0^1 \hat{w}_i' f(\hat{w}_i) d\hat{w}_i \]

\[ = (1 - b)^{m_2/2} \sum_{j=0}^{\infty} \left[ C_j^{(m+j)-1} b^j \frac{B((m_2/2) + j, m_1/2)}{B(m_2/2, m_1/2)} \right] \]

Again, by the previous theorem (Apostol, 1974, p. 193), with

\[ R_r = \limsup_{n \to \infty} \left| \frac{(m + n)(m_2 + 2r + 2n)}{(1 + n)(2m + 2r + 2n)} b \right| = |b| \]

\[ R_r = |b| < 1 \], then the \( r \)th-order raw moment of \( \hat{w}_i \) exists.

**Appendix A.5. Proof of Corollary 1**

Let

\[ B_j^r = \frac{B((m_2/2) + r + j, m_2/2)}{B(m_2/2, m_1/2)} \]

\[ B_j^{r+1} = \frac{m_2 + 2r + 2j}{m_1 + m_2 + 2r + 2j} B_j^r \]

Then

\[ B_j^{r+1} < B_j^r, \forall r \in \mathbb{N} \]
and therefore
\[
E(\hat{\omega}_{r+1}^2) = (1 - b)^{m/2} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j B_{r+1}^j < (1 - b)^{m/2} \sum_{j=0}^{\infty} C_j^{(m+j)-1} b^j B_r^j = E(\hat{\omega}_r^2), \forall r \in \mathbb{N}
\]

**Appendix A.6. Pearson Type I Approximation**

We must compute some important coefficients such as
\[
\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \beta_2 = \frac{\mu_4}{\mu_2^2}
\]
then check the following conditions
\[
6 + 3\beta_1 - 2\beta_2 > 0, \beta_2 - \beta_1 - 1 > 0
\]
and the interpolation between \(x_0\) and \(x_1\) is
\[
x(\theta) = (1 - \theta)x_0 + \theta x_1 - \frac{1}{4} \theta(1 - \theta)(\Delta^2 x_0 + \Delta^2 x_1)
\]

**References**