with
\[ W_i = \frac{1}{2} [(x_i - x_0)^2 f_x(x_i, y_i) + (x_i - x_0)^2 f_y(x_i, y_i) + 2(x_i - x_0)(y_i - y_0) f_{xy}(x_i, y_i)]. \]

Corresponding formulas can be derived in this way for any value of \( n; \) in fact, several alternatives may be obtained in each case. In all cases the error \( f(x_0, y_0) \) is given in terms of the derivatives of \( g \) alone if a polynomial of a certain type is used for the interpolating function. For equation (4), the suitable polynomial would be \( h(x, y) = a + bx + cy; \) for (5), \( h(x, y) = a + bx + cy + dx^2 + exy + fy^2; \) for (6), \( h(x, y) = a + bx + cy + dx^3. \) If the interpolating function \( h(x, y) \) is not so chosen, the formulas remain valid, but derivatives of \( h \) will appear.

The same procedure is applicable to functions of any number of independent variables.

---

**ON A LEMA BY KOLMOGOROFF**

**BY KAI-LAI CHUNG**

*Princeton University*

The following lemma was proved by Kolmogoroff [1]:

If \( e_1, e_2, \ldots, e_n \) are independent events and \( U \) an arbitrary event such that \( W(X) \) denoting the probability of \( X \) and \( W_e(X) \) the conditional probability of \( X \) under the hypothesis of \( e \)

\[ W_{e_k}(U) \geq u, \quad W(e_1 + \cdots + e_n) \geq u. \]

Then

\[ W(U) \geq \frac{1}{n} u^n. \]

This result seems of some interest in itself and may also have practical applications, for it is easily seen that [2] in general if \( e_1, e_2, \ldots, e_n \) are arbitrary no information about \( W_{e_1+\cdots+e_n}(U) \) can be obtained from that about \( W_{e_k}(U), \ k = 1, \cdots, n. \) From this point of view the constant 1/9 is interesting, though it is important in Kolmogoroff’s proof of the law of large numbers. Using his original method this constant can easily be improved to 1/8. However, the following method will give a better result. At the same time we shall put it into a more general form.

Let

\[ W_{\alpha}(U) \geq \alpha, \quad \sum_{k=1}^{n} W(e_k) \geq \beta. \]
ON A LEMMA BY KOLMOGOROFF

Then we have for $1 \leq k \leq n$,

\[(1) \quad W(U) \geq W(U(e_1 + \cdots + e_k))) = W(Ue_1 + \cdots + Ue_k).\]

Now a simple case of certain inequalities due to Bonferroni and Frechet [3] states that for arbitrary events $E_1, \cdots, E_k$ we have

\[(2) \quad W(E_1 + \cdots + E_k) \geq \sum_{i=1}^{k} W(E_i) - \sum_{1 \leq i < j \leq k} W(E_i E_j).\]

Applying this to (1), we obtain

\[
W(U) \geq \sum_{i=1}^{k} W(Ue_i) - \sum_{1 \leq i < j \leq k} W(Ue_i e_j)
\]

\[
\geq \sum_{i=1}^{k} W(e_i) W_{e_i}(U) - \sum_{1 \leq i < j \leq k} W(e_i) W(e_j),
\]

using the independence of $e_1, \cdots, e_k$. Hence

\[
W(U) \geq \alpha \sum_{i=1}^{k} W(e_i) - \frac{1}{2} \left( \sum_{i=1}^{k} W(e_i) \right)^2 + \frac{1}{2} \sum_{i=1}^{k} W^2(e_i).
\]

By Cauchy's inequality,

\[
\sum_{i=1}^{k} W^2(e_i) \geq \frac{1}{k} \left( \sum_{i=1}^{k} W(e_i) \right)^2.
\]

Writing $\sum_k = \sum_{i=1}^{k} W(e_i)$, we have

\[(3) \quad W(U) \geq \left[ \alpha - \left( \frac{1}{2} - \frac{1}{2k} \right) \sum_k \right] \sum_k.
\]

Now let $0 < \gamma < \gamma_0 \leq 1$ where $\gamma$ and $\gamma_0$ are to be determined later. If there is an $e_i$, $1 \leq i \leq n$ such that $W(e_i) \geq \gamma \beta$, then

\[(4) \quad W(U) \geq W(U e_i) = W(e_i) W_{e_i}(U) \geq \gamma \alpha \beta.
\]

If every $W(e_i) < \gamma \beta$, we determine $k(> 1)$ such that

\[\Sigma_{k-1} < \gamma_0 \beta \leq \Sigma_k;
\]

thus

\[\Sigma_k < \Sigma_{k-1} + \gamma \beta < (\gamma_0 + \gamma) \beta.
\]

And (3) yields

\[(5) \quad W(U) \geq \left[ \alpha - \frac{1}{2} \left( 1 - \frac{1}{k} \right) (\gamma_0 + \gamma) \beta \right] \gamma \beta.
\]
Now we choose $\gamma$ so that the last terms in (4) and (5) be equal. This gives

$$\gamma = \frac{2\alpha - \left(1 - \frac{1}{k}\right) \gamma_0 \beta}{2\alpha + \left(1 - \frac{1}{k}\right) \gamma_0 \beta}.$$

To maximize $\gamma$, we put $\frac{d\gamma}{d\gamma_0} = 0$ and find

$$\gamma_0 = \frac{2(\sqrt{2} - 1)\alpha}{\beta}.$$

If $2(\sqrt{2} - 1)\alpha \leq \beta$, this choice of $\gamma_0$ is admissible, and we obtain

$$\gamma = \frac{2 - \sqrt{2} + \frac{1}{k} (\sqrt{2} - 1)}{\sqrt{2} - \frac{1}{k} (\sqrt{2} - 1)} \frac{2(\sqrt{2} - 1)\alpha}{\beta}.$$

Thus we get (the first inequality being retained for small values of $n$)

$$W(U) \geq \frac{2 - \sqrt{2} + \frac{1}{n} (\sqrt{2} - 1)}{\sqrt{2} - \frac{1}{n} (\sqrt{2} - 1)} \frac{2(\sqrt{2} - 1)\alpha^2}{\beta} \geq 2(\sqrt{2} - 1)^2 \alpha^2 > \frac{3\alpha}{100} \alpha^2.$$

In case $2(\sqrt{2} - 1)\alpha > \beta$, we choose $\gamma_0 = 1$, and we obtain

$$\gamma = \frac{2\alpha - \left(1 - \frac{1}{k}\right) \beta}{2\alpha + \left(1 - \frac{1}{k}\right) \beta}.$$

Thus we get

$$W(U) \geq \frac{2\alpha - \left(1 - \frac{1}{n}\right) \beta}{2\alpha + \left(1 - \frac{1}{n}\right) \beta} \alpha \beta \geq \frac{2\alpha - \beta}{2\alpha + \beta} \alpha \beta.$$

If we write $\beta = \eta \alpha$, we have

$$W(U) \geq \frac{2 - \eta}{2 + \eta} \eta \alpha^2.$$
We summarize (6) and (7) in the following table:

\[
\begin{array}{c|c|c}
\beta/\alpha & \geq 2(\sqrt{2} - 1) & = \eta < 2\sqrt{2} - 1 \\
W(U) & \geq 2(\sqrt{2} - 1)^2 \alpha^2 & = \frac{2 - \eta}{2 + \eta} \eta \alpha^2
\end{array}
\]

Thus for Kolmogoroff's case (\(\eta = 1\)) we have \(W(U) \geq \frac{1}{2} \alpha^2\).

REFERENCES


APPROXIMATE WEIGHTS

By John W. Tukey

Princeton University

1. Summary. The greatest fractional increase in variance when a weighted mean is calculated with approximate weights is, quite closely, the square of the largest fractional error in an individual weight. The average increase will be about one-half this amount.

The use of weights accurate to two significant figures, or even to the nearest number of the form: 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 95, that is to say, of the form 10(1/20(2)50(5)100 \times 10^0 can thus reduce efficiency by at most \(1/4\) percent, which is negligible in almost all applications.

2. Proof. Let the optimum weights be \(W_i, i = 1, 2, \ldots, n\), with \(W_i \geq 0\), where it is convenient to choose the normalization \(\Sigma W_i = 1\). Let \(\sigma^2\) be the variance of \(\Sigma W_i x_i\), then the variance of each \(x_i\) must be \(\sigma^2/W_i\), and since this is a weighted mean, the means of the \(x_i\) are the same.

Let the approximate weights be \(W_i(1 + \lambda \theta_i), \) where \(0 < \lambda < 1\) and \(|\theta_i| \leq 1, i = 1, 2, \ldots, n\). Thus \(\lambda\) is the largest fractional error which may be made in the situation considered. We need the weak requirement \(\lambda < 1\). The approximately weighted mean is

\[
\frac{\sum W_i(1 + \lambda \theta_i) x_i}{\sum W_i(1 + \lambda \theta_i)} = \sum W_i \frac{1 + \lambda \theta_i}{1 + \lambda \theta},
\]