1 Introduction

In this section, we summarized several formulas which were obtained by Neyman with a finite number of plots $m$. The urn model was summarized and pictured at the end of this report as an appendix.

The true yield on plot $k$ and condition $i$ is denoted as

$$u_{ik}, \quad i = 1, \ldots, \nu, \quad k = 1, \ldots, m,$$

and the random yield on condition $i$ is denoted as

$$U_i, \quad i = 1, \ldots, \nu,$$

and the random yield on plot $k$ and condition $i$ is denoted as

$$U_{ik}, \quad i = 1, \ldots, \nu, \quad k = 1, \ldots, m.$$

As Neyman [2] mentioned, if we could repeat the measurement of the yield on the same plot under the same condition, the observed yield would be essentially equal to the true yield $u_{ik}$.

The true average of yields from $i$th variety on all different plots is

$$u_i = \frac{\sum_{k=1}^{m} u_{ik}}{m},$$

and the estimated average of $u_i$ from $\kappa$ measurements (from $\kappa$ plots out of $m$ plots) is

$$U_i = \frac{1}{\kappa} \sum_{j=1}^{\kappa} U_{ikj},$$
where \( \{k_1, \ldots, k_\kappa\} \subseteq \{1, \ldots, m\} \), and

\[
E U_i = u_i,
\]

because \( E U_{ik} = u_i \) for each \( k \in \{1, \ldots, m\} \).

If each of the \( \nu \) varieties is sown on \( \kappa \) plots, then \( m = \nu \kappa \).

Consider the variance of the estimated average \( \bar{\sigma}_i^2 \), we have

\[
\bar{\sigma}_i^2 = E \left[ (U_i - u_i)^2 \right] = E U_i^2 - u_i^2.
\]

Note that

\[
U_i^2 = \frac{1}{\kappa^2} \left( \sum_{j=1}^{\kappa} U_{ikj}^2 + 2 \sum_{j=1}^{\kappa-1} \sum_{r=j+1}^{\kappa} U_{ikj} U_{ikr} \right),
\]

and

\[
E U_{ikj}^2 = \frac{\sum_{k=1}^{m} u_{ik}^2}{m},
\]

\[
E U_{ikj} U_{ikr} = \frac{\sum_{k=1}^{m-1} \sum_{r=k+1}^{m} u_{ik} u_{ir}}{m(m-1)/2},
\]

we have

\[
\bar{\sigma}_i^2 = \frac{1}{\kappa^2} \left( \kappa \cdot \frac{\sum_{k=1}^{m} u_{ik}^2}{m} + 2 \cdot \frac{\kappa(\kappa - 1)}{2} \cdot \frac{\sum_{k=1}^{m-1} \sum_{r=k+1}^{m} u_{ik} u_{ir}}{m(m-1)/2} \right) - u_i^2.
\]

\[
= \frac{m - \kappa}{\kappa m(m-1)} \sum_{k=1}^{m} u_{ik}^2 + \frac{\kappa - 1}{\kappa m(m-1)} \left( \frac{\sum_{k=1}^{m} u_{ik}}{m} \right)^2 - \frac{1}{m} \left( \frac{\sum_{k=1}^{m} u_{ik}}{m} \right)^2
\]

\[
= \frac{m - \kappa}{\kappa m(m-1)} \left( \sum_{k=1}^{m} u_{ik}^2 - \left( \frac{\sum_{k=1}^{m} u_{ik}}{m} \right)^2 \right)
\]

\[
= \frac{m - \kappa}{\kappa(m-1)} \left( \sum_{k=1}^{m} u_{ik} - u_i \right)^2
\]

\[
= \frac{m - \kappa}{\kappa(m-1)} \sigma_{U_i}^2
\]

So we have

\[
\bar{\sigma}_i^2 = \frac{1 - \frac{\kappa}{m} \cdot \sigma_{U_i}^2}{1 - \frac{1}{m} \cdot \kappa}.
\]
Note that the finite correction factor $\frac{1 - \frac{\kappa}{m}}{1 - \frac{1}{m}}$ here, because $m$ is a finite number.

Note that
$$\lim_{\kappa \to +\infty} \hat{\sigma}_i^2 = 0,$$
and
$$\lim_{m \to +\infty} \hat{\sigma}_i^2 = \frac{\sigma_{U_i}^2}{\kappa}.$$

The arithmetic mean $U_i = \frac{1}{\kappa} \sum_{j=1}^{\kappa} U_{ikj}$ is also the best unbiased linear estimator among all estimators with expression
$$\sum_{j=1}^{\kappa} \lambda_j U_{ikj},$$
by achieving the minimal variance.

The estimated variance is
$$\hat{\sigma}_i^2 = \mathbb{E} \left[ (U_{ikj} - U_i)^2 \right]$$
$$= \mathbb{E} \left[ \left( U_{ikj} - \frac{1}{\kappa} \sum_{r=1}^{\kappa} U_{ikr} \right)^2 \right]$$
$$= \mathbb{E} \left[ \left( \frac{\kappa - 1}{\kappa} U_{ikj} - \frac{1}{\kappa} \sum_{r \neq j} U_{ikr} \right)^2 \right]$$
$$= \left( \frac{(\kappa - 1)^2}{\kappa^2} + \frac{1}{\kappa^2} (\kappa - 1) \right) \mathbb{E} U_{ikj}^2 + \left( -2 \frac{\kappa - 1}{\kappa} \frac{\kappa - 1}{\kappa} + \frac{1}{\kappa^2} (\kappa - 1) (\kappa - 2) \right) \mathbb{E} U_{ikj} U_{ikr}$$
$$= \frac{\kappa - 1}{\kappa} \mathbb{E} U_{ikj}^2 - \frac{\kappa - 1}{\kappa} \mathbb{E} U_{ikj} U_{ikr}$$
$$= \frac{\kappa - 1}{\kappa} \frac{1}{m} \sum_{k=1}^{m} u_{ik}^2 - \frac{\kappa - 1}{\kappa} \frac{1}{m} \sum_{k=1}^{m-1} \sum_{r=k+1}^{m} u_{ik} u_{ir}$$
$$= \frac{\kappa - 1}{\kappa(m - 1)} \left( \sum_{k=1}^{m} u_{ik}^2 - \frac{(\sum_{k=1}^{m} u_{ik})^2}{m} \right)$$
$$= \frac{(\kappa - 1) m \left( \sum_{k=1}^{m} u_{ik}^2 - u_i \right)^2}{\kappa(m - 1) \sigma_{U_i}^2}$$
$$= \frac{(\kappa - 1) m}{\kappa(m - 1)} \sigma_{U_i}^2.$$

3
where we assume that the \( k_j \)th plot is in the set of sampled \( \kappa \) plots.

So we have

\[
\hat{\sigma}_i^2 = \frac{1 - \frac{1}{\kappa}}{1 - \frac{1}{m}} \sigma_{U,i},
\]

and note the finite effect of \( m \) here.

So the ratio of the variance of the estimated average \( \bar{\sigma}^2_i \) to the estimated variance \( \hat{\sigma}^2_i \) is

\[
\frac{\bar{\sigma}^2_i}{\hat{\sigma}^2_i} = \frac{(\kappa - 1) \frac{m}{\kappa}}{m - \kappa} = \frac{1 - 1/\kappa}{\kappa - 1/m}.
\]

Next we determined the difference between the yields of two varieties \( i \) and \( j \). First it is clear that

\[
\mathbb{E}[U_i - U_j] = u_i - u_j.
\]

The variance of estimated difference is

\[
\bar{\sigma}^2_{U_i - U_j} = \mathbb{E}[(U_i - U_j) - (u_i - u_j)]^2
\]

\[
= \bar{\sigma}^2_i + \bar{\sigma}^2_j - 2 (\mathbb{E}U_iU_j - u_iu_j).
\]

Note that \( \{k_r\}_{r=1}^{\kappa} \cap \{k_l\}_{l=1}^{\kappa} = \emptyset \) and

\[
\mathbb{E}U_iU_j = \mathbb{E}\left[\left(\frac{1}{\kappa} \sum_{r=1}^{\kappa} U_{ik_r}\right)\left(\frac{1}{\kappa} \sum_{l=1}^{\kappa} U_{jk_l}\right)\right]
\]

\[
= \mathbb{E}_{r \neq l} U_{ir}U_{jl}
\]

\[
= \frac{\sum_{r=1}^{m-1} \sum_{i=r+1}^{m} u_{ir}u_{jl}}{m(m-1)/2}
\]

\[
= \left(\sum_{r=1}^{m} u_{ir}\right)\left(\sum_{l=1}^{m} u_{jl}\right) - \sum_{r=1}^{m} u_{ir}u_{jr}
\]

\[
= \frac{m^2 u_iu_j - \sum_{r=1}^{m} u_{ir}u_{jr}}{m(m-1)},
\]

so

\[
\mathbb{E}U_iU_j - u_iu_j = \frac{u_iu_j - \frac{1}{m} \sum_{r=1}^{m} u_{ir}u_{jr}}{m - 1}.
\]

Define the correlation coefficient \( r_{ij} \) between the yield of two varieties \( i \) and \( j \) on the same plot

\[
r_{ij} = \frac{1}{\sigma_{U,i}\sigma_{U,j}} \sum_{r=1}^{m} u_{ir}u_{jr} - u_iu_j.
\]
we have
\[
\mathbb{E}U_i U_j - u_i u_j = -\frac{1}{m-1} r_{ij} \sigma_i \sigma_j
\]
\[
= -\frac{\kappa r_{ij}}{m - \kappa} \bar{\sigma}_i \bar{\sigma}_j.
\]

So
\[
\bar{\sigma}_{U_i - U_j}^2 = \bar{\sigma}_i^2 + \bar{\sigma}_j^2 + \frac{2 \kappa r_{ij}}{m - \kappa} \bar{\sigma}_i \bar{\sigma}_j,
\]

note that the finite correction term \(\frac{2 \kappa r_{ij}}{m - \kappa} \bar{\sigma}_i \bar{\sigma}_j\) goes to 0 when \(m\) goes to infinity.

Note that \(m = \nu \kappa\), we have
\[
\bar{\sigma}_{U_i - U_j}^2 = \bar{\sigma}_i^2 + \bar{\sigma}_j^2 + \frac{2 r_{ij}}{\nu - 1} \bar{\sigma}_i \bar{\sigma}_j.
\]

So if the number of plots \(m\) is finite, because of the correlation of two varieties on the same plot, actually we have
\[
\bar{\sigma}_i^2 + \bar{\sigma}_j^2 - \frac{2}{\nu - 1} \bar{\sigma}_i \bar{\sigma}_j \leq \bar{\sigma}_{U_i - U_j}^2 \leq \bar{\sigma}_i^2 + \bar{\sigma}_j^2 + \frac{2}{\nu - 1} \bar{\sigma}_i \bar{\sigma}_j,
\]

where \(\nu = m/\kappa\).

2 My Comments

In [2], Neyman obtained formulas for the variance of average yields, and the variance of the difference between the averages of the observed yields of two varieties, where Neyman assumed that the yield of each variety was observed on \(\kappa\) plots, where \(\kappa\) kept the same for different varieties.

A slight generalization would be to consider different \(\kappa\)'s. Say, different varieties would assign to different numbers of plots. Let us assume that variety \(i\) had assigned to \(\kappa_i\) plots, where \(\sum_{i=1}^{\nu} \kappa_i = m\).

This generalization is not difficult, so I think that probably someone has done this work before.

By using the similar mathematical deduction, we have
\[
\bar{\sigma}_i^2 = \frac{m - \kappa_i}{\kappa_i (m - 1)} \sigma_{U_i}^2,
\]
and
\[
\bar{\sigma}^2_{U_i-U_j} = \bar{\sigma}^2_i + \bar{\sigma}^2_j + \frac{2r_{ij}\sqrt{\kappa_i\kappa_j}}{\sqrt{(m-\kappa_i)(m-\kappa_j)}}\bar{\sigma}_i\bar{\sigma}_j,
\]

Now the question is: How to allocate the numbers of plots?

(1) If we want to minimize the sum of \(\nu\) variances of average yields, which is
\[
\sum_{i=1}^{\nu} \bar{\sigma}^2_i,
\]
note that
\[
\sum_{i=1}^{\nu} \bar{\sigma}^2_i = \frac{1}{m-1} \sum_{i=1}^{\nu} \frac{m-\kappa_i}{\kappa_i} \bar{\sigma}^2_{U_i} = \frac{1}{m-1} \sum_{i=1}^{\nu} \left( \frac{m}{\kappa_i} - 1 \right) \bar{\sigma}^2_{U_i}
\]
\[
= \frac{m}{m-1} \sum_{i=1}^{\nu} \frac{\bar{\sigma}^2_{U_i}}{\kappa_i} - \frac{1}{m-1} \sum_{i=1}^{\nu} \bar{\sigma}^2_{U_i},
\]
by using Lagrange Multiplier method, we got that \(\sigma^2_{U_i}/\kappa_i^2\) should be a constant.

Finally the best allocation\(^1\) which makes the minimal \(\sum_{i=1}^{\nu} \bar{\sigma}^2_i\) is
\[
\kappa_{BV}^{i} = \frac{m \sigma_{U_i}}{\sum_{s=1}^{\nu} \sigma_{U_s}},
\]
and the smallest \(\sum_{i=1}^{\nu} \bar{\sigma}^2_i\) is
\[
\frac{1}{m-1} \left( \sum_{i=1}^{\nu} \bar{\sigma}^2_{U_i} \right)^2 - \sum_{i=1}^{\nu} \bar{\sigma}^2_{U_i}.
\]

(2) If we want to minimize the sum of \(\nu(\nu-1)/2\) differences between the averages of the observed yields of two varieties, which is
\[
\sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \bar{\sigma}^2_{U_i-U_j},
\]
note that
\[
\sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \bar{\sigma}^2_i = \sum_{i=1}^{\nu-1} (\nu - i) \bar{\sigma}^2_i = \sum_{i=1}^{\nu} (\nu - i) \bar{\sigma}^2_i,
\]

\(^1\)BV stands for the Best allocation for sum of Variances.
and
\[ \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \sigma_j^2 = \sum_{j=2}^{\nu} \sum_{i=1}^{j-1} \sigma_i^2 \sigma_j^2 = \sum_{j=2}^{\nu} (j-1) \sigma_j^2 = \sum_{j=1}^{\nu} (j-1) \sigma_j^2, \]
then
\[ \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \bar{\sigma}_{U_i-U_j}^2 = \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \left( \sigma_i^2 + \frac{2r_{ij} \sqrt{\kappa_i \kappa_j}}{(m - \kappa_i)(m - \kappa_j)} \bar{\sigma}_i \bar{\sigma}_j \right) \]
\[ = (\nu - 1) \sum_{i=1}^{\nu} \sigma_i^2 + \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \frac{2r_{ij} \sqrt{\kappa_i \kappa_j}}{(m - \kappa_i)(m - \kappa_j)} \bar{\sigma}_i \bar{\sigma}_j \]
\[ = (\nu - 1) \sum_{i=1}^{\nu} \frac{m - \kappa_i}{\kappa_i (m - 1)} \sigma_{U_i}^2 + \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \frac{2r_{ij}}{m - 1} \sigma_{U_i} \sigma_{U_j} \]
\[ = \frac{m(\nu - 1)}{m - 1} \sum_{i=1}^{\nu} \frac{\sigma_{U_i}^2}{\kappa_i} + \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \frac{2r_{ij}}{m - 1} \sigma_{U_i} \sigma_{U_j} - \frac{\nu - 1}{m - 1} \sum_{i=1}^{\nu} \sigma_{U_i}^2, \]

note that \( \sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \frac{2r_{ij}}{m - 1} \sigma_{U_i} \sigma_{U_j} - \frac{\nu - 1}{m - 1} \sum_{i=1}^{\nu} \sigma_{U_i}^2 \) does not depend on \( \kappa_i \), so the best allocation\(^2\) is still
\[ \kappa_i^{BD} = \frac{m \sigma_{U_i}}{\sum_{s=1}^{\nu} \sigma_{U_s}}. \]

### 3 Summary

In Section 1, we summarized the results in Neyman (1923).

The variance of the estimated average \( \bar{\sigma}_i^2 \) is
\[ \bar{\sigma}_i^2 = \frac{1 - \kappa}{1 - \frac{1}{m}} \cdot \frac{\sigma_{U_i}^2}{\kappa}. \]

The variance of estimated difference \( \bar{\sigma}_{U_i-U_j}^2 \) is
\[ \bar{\sigma}_{U_i-U_j}^2 = \sigma_i^2 + \sigma_j^2 + \frac{2r_{ij} \sqrt{\kappa_i \kappa_j}}{(m - \kappa \sigma_i \sigma_j)}. \]

\(^2\)BD stands for the Best allocation for sum of variances of Differences.
In Section 2, we generalized the results from using constant $\kappa$ among different varieties to using different $\kappa_i$ for each variety $i$, where $\kappa$ is the number of plots for a certain variety.

The variance of the estimated average $\bar{\sigma}_i^2$ is

$$\bar{\sigma}_i^2 = \frac{m - \kappa_i}{\kappa_i(m - 1)} \sigma_{U_i}^2.$$  

The variance of estimated difference $\bar{\sigma}_{U_i - U_j}^2$ is

$$\bar{\sigma}_{U_i - U_j}^2 = \bar{\sigma}_i^2 + \bar{\sigma}_j^2 + \frac{2r_{ij}\sqrt{\kappa_i\kappa_j}}{\sqrt{(m - \kappa_i)(m - \kappa_j)}} \bar{\sigma}_i \bar{\sigma}_j.$$  

In order to get the best allocation, we would like to minimize either

$$\sum_{i=1}^{\nu} \bar{\sigma}_i^2,$$

or

$$\sum_{i=1}^{\nu-1} \sum_{j=i+1}^{\nu} \bar{\sigma}_{U_i - U_j}^2,$$

and the best allocations are the same, which are

$$\kappa_i^{best} = \frac{m\sigma_{U_i}}{\sum_{s=1}^{\nu} \sigma_{U_s}}.$$  

References


Appendix: Urn Model
The design of a field experiment involving plots

Field

\[ \sum_{i=1}^{m} U_i \]

... plot

\[ U_1, \ldots, U_m \]

the true yields of a particular variety on each plot

average yield \( \bar{a} = \frac{\sum_{i=1}^{m} U_i}{m} \)
differences among yields

e.g. soil condition

To compare \( z \) varieties, two indices:

one corresponding to the variety to the plot

\( U_{i1}, \ldots, U_{im} \quad (i=1,2,\ldots,z) \)

Take \( z \) urns, each variety is associated with exactly one urn.

Variety 1 \( \leftarrow \) urn 1

Put \( m \) balls in the \( i \)th urn, labeled \( \leftrightarrow \) plot

Variety 2 \( \leftarrow \) urn 2

\( a_i = \frac{\sum_{k=1}^{m} U_{ik}}{m} \)
average number

best estimate of yield from the \( i \)th variety on the field

(why)

Um 1           Um 2
\[ \begin{array}{cc}
0 & o \\
o & 0 \\
o & 0 \\
o & 0 \\
\end{array} \]

The goal of a field experiment is to compare the numbers \( a_1, \ldots, a_z \)
simplest way w/ replacement

Drawn balls will be returned,

independent outcomes

\( a \)
a possible outcome

\( X \)
corresponding realized values

\( \mathbb{E} X = \sum_{k=1}^{m} U_{ik} \cdot \frac{1}{m} = \bar{a_i} \)

↑ probabilities are equal
In practice, w/o replacement

\[ x_1, \ldots, x_k \quad \text{possible outcomes} \quad \text{K trials} \]

\[ X_1, \ldots, X_k \quad \text{true} \]

group the sequence in one urn

\[ U_{i1}, U_{i2}, \ldots, U_{ik}, \ldots, U_{im} \]

\[ \begin{array}{c}
V_{i1}, V_{i2}, \ldots, V_{in} \\
\uparrow & \uparrow & \uparrow \\
m_{p_1} & m_{p_2} & m_{p_n}
\end{array} \]

\[ \sum_{i=1}^{n} P_i = 1 \]

First ball drawing we got \( V_{ik} \)

Second trial \( V_{ir}, r \neq k \)

\[ P_r^1 = \frac{mp_r}{m-1} = Pr + \frac{Pr}{m-1} \]

if \( r = k \)

\[ P_r^1 = \frac{mp_{k-1}}{m-1} = mp_{k-1} + P_k - P_k = P_k - \frac{1 - P_k}{m-1} \]

note \[ \sum_{r \neq k} \left( P_r + \frac{Pr}{m-1} \right) + \left( P_k - \frac{1 - P_k}{m-1} \right) = 1 \]

In the end, after \( K-1 \) trails being carried out

if \( V_{ik} \), \( V_{ik} \) has not been drawn so far for previous trial

\[ p_{k,0}^{K-1} = \frac{mp_k}{m-K+1} = P_k + \frac{(K-1)P_k}{m-K+1} \]

has an effect \( V_{is} \) has been drawn \( l \) times on the subsequent trial

\[ p_{s,l}^{K-1} = \frac{mp_s - l}{m-K+l} = P_s + \frac{(K-1)P_s - l}{m-K+l} \]

not independent

if \( m \gg K \)

\( \nu \) is relatively large nearly independent

Neyman asked: What if \( \nu \) is small?
First, the design w/ one urn i-th urn

True mean \( a = \frac{\sum_{k=1}^{m} U_k}{m} \)

arithmetic mean from \( k \) measurements may be considered

\( x_i = \frac{1}{k} \sum_{k=1}^{k} x_{ik} \) an estimate of \( a \)

From Tchebychev's theory, it's enough to show

\[ \mu^2 = \mathbb{E}(x_i - a)^2 \] trends to 0, as \( k \to \infty \)

We calculate \( \mu^2 \)

\[ \mu^2 = \mathbb{E}x_i^2 - a^2 = \frac{1}{k^2} \mathbb{E} \left[ \sum_{k=1}^{k} x_{ik}^2 + 2 \sum_{k=1}^{k} \sum_{r=k+1}^{k} x_{ik} x_{ir} \right] - a^2 \]

\[ \mathbb{E} x_{ik}^2 = \frac{m}{m} U_k^2 \]

\[ \mathbb{E} x_{ik} x_{ir} = \frac{m}{m} \sum_{r=k+1}^{m} \frac{U_k U_r}{m(m-1)/2} \]

\[ \Rightarrow \mu^2 = \frac{1}{k^2} \left[ k \cdot \sum_{k=1}^{m} U_k^2 + 2 \frac{k(k-1)}{2} \frac{m}{m} \sum_{k=1}^{m} \sum_{r=k+1}^{m} U_k U_r \right] - a^2 \]

\[ = \frac{1}{k} \left[ \frac{\sum_{k=1}^{m} U_k^2}{m} + 2 \frac{k(k-1)}{2} \frac{m}{m(m-1)/2} \sum_{k=1}^{m} U_k U_r \right] - a^2 \]

\[ = \frac{1}{k m(m-1)} \left[ (m-1) \frac{\sum_{k=1}^{m} U_k^2}{m} + 2 (k-1) \sum_{k=1}^{m} U_k U_r \right] - a^2 \]

\[ = \frac{1}{k m(m-1)} \left[ (m-k) \sum_{k=1}^{m} U_k^2 + (k-1) \frac{m}{k-1} U_k^2 + 2 (k-1) \sum_{k=1}^{m} U_k U_r \right] - a^2 \]

\[ = \frac{m-k}{k m(m-1)} \sum_{k=1}^{m} U_k^2 + \frac{k-1}{k m(m-1)} \left( \sum_{k=1}^{m} U_k \right)^2 - a^2 \]

\[ a = \frac{m}{m} U_k \]

\[ = \frac{m-k}{k m(m-1)} \sum_{k=1}^{m} U_k^2 - \frac{m-k}{k m(m-1)} \left( \sum_{k=1}^{m} U_k \right)^2 \]

\[ \frac{k-1}{k m(m-1)} - \frac{1}{m} = \frac{m k-m-k m+k}{k m(m-1) m} \]

\[ \frac{m-k}{k m(m-1)} \left( U_k - a \right)^2 = \frac{m-k}{k m(m-1)} a^2 \]
\( a_u^2 \) is a constant

\[ m > k, \quad \text{so} \quad \lim_{k \to \infty} \frac{m-k}{k(m-1)} = 0 \]

Thus regarded as an estimate.

\[ \mathbb{E} F(x, k) = a \quad \quad \text{look for the best est.} \]

\[ F(x, k) = \lambda_1 x_1 + \ldots + \lambda_k x_k \quad \quad \sum \lambda_k = 1 \]

It's necessary that

\[ M^2 = \mathbb{E}(F(x, k) - a)^2 \quad \text{be a minimum} \]

\[ \mathbb{E}(x_i - a)(x_k - a) = \frac{2 \sum_{i=1}^{m-1} \sum_{k=i+1}^{m} (U_i - a)(U_k - a)}{m(m-1)} \]

\[ = \frac{2 \sum (U_i - a)(U_k - a) + \sum (U_i - a)^2 - \sum (U_i - a)^2}{m(m-1)} \]

\[ = \frac{[\sum (U_i - a)]^2 - \sum (U_i - a)^2}{m(m-1)} \quad \text{note} \quad \sum U_i = m a \]

\[ m^2 = \sum_{i=1}^{k} \lambda_i^2 a_u^2 + 2 \sum_{i=1}^{k} \sum_{k=i}^{k} \lambda_i \lambda_k (- a_u^2 \frac{m}{m-1}) \]

\[ = a_u^2 \left( \sum_{i=1}^{k} \lambda_i^2 - \frac{2}{m-1} \sum_{i=1}^{k} \sum \lambda_i \lambda_k \right) \]

\[ M^2 = a_u^2 \left( \text{(m-1) \sum} \lambda_i^2 - 2 \sum_{i=1}^{k-1} \sum \lambda_i \lambda_k \right) \]

\[ \frac{m^2 \sum \lambda_i^2 - 1}{m-1} \quad \text{best estimate} \]

\[ \Lambda = m \sum_{i=1}^{k} \lambda_i^2 - 1 \quad \Lambda \lambda_i = 1 \quad \text{Lagrange Multiplier} \]

\[ f = m \sum_{i=1}^{k} \lambda_i^2 - 1 \quad \sum \lambda_i = 1 \]

\[ \frac{\partial f}{\partial \lambda_k} = 2m \lambda_k - \delta = 0 \quad \Rightarrow \lambda_k = \frac{\delta}{2m} \quad \sum \lambda_k = 1 \]

\[ \Rightarrow \delta = \frac{2m}{k} \quad \Rightarrow \delta = \frac{2m}{k} \]