APPLICATIONS OF TWO NON-CENTRAL HYPERGEOMETRIC DISTRIBUTIONS OF BIASED SAMPLING STATISTICAL MODELS

K.O. Adetunji, T. Alanamu, E.J. Adefila

Department of Statistics, University of Ilorin, Ilorin, Nigeria

Corresponding author: Adetunji Kamaldeen Olarewaju, karlmarx4@yahoo.com

ABSTRACT: *Statistical models of biased sampling of two non central hypergeometric distributions Wallenius' and Fisher's distribution. However, not many of the logic of hypergeometric distribution have been investigated by different techniques. The work examined the procedure of the two non-central hypergeometric distributions and investigates the statistical properties of the two non central hypergeometrics which includes the mean and variance that were obtained. The parameters of the distribution were used using the direct inversion method of hyper simulation of biased urn model in the environment of R statistical software with varying odd ratios (w) and group sizes (m_i) were used. It was found that the two non-central hypergeometrics are approximately equal in mean, variance and coefficient of variation and differ as odds ratios (w) becomes higher and differfrom the central hypergeometric distribution with ω = 1. Furthermore, in univariate situation we observed that Fisher distribution at (ω = 0.2, 0.5,0.7, 0.9) is more consistent than Wallenius distribution although central hypergeometric is more consistent than any of them also in multinomial situation, it was observed that Fisher distribution is more consistent at (ω = 0.2, 0.5) Wallenius distribution at (ω = 0.7, 0.9) central hypergeometric at (ω = 0.2)*
KEYWORDS: *Non central hypergeometric, wallenius distribution, fisher distribution, univarate situation, multinomial situation*

1. INRODUCTION

The hypergeometric distribution occupies a place of great significance in statistic theory.
It applies to sampling without replacement from a finite population whose element can be classified into two categories, one which possesses certain characteristics. The category could be male or female, employed or unemployed etc. When random selection are made without replacement from the population, each subsequent drawn is dependent and the probability of success change in each draw. The following conditions are categories of hyper geometric distribution

- The result of each draw can be classified into one or two categories
- The probability of success changes in each draw.

In probability theory and statistics, the hyper geometric distribution is a discrete probability that describes the probability of k successes in n drawn from a finite population of size N containing k successes without replacement.

A hyper geometric random variable with parameter $W+B, W$ and n, given a set consisting of W element of the kind and B element of the second kind, a number of element of the first kind appearing in a randomly chosen subset of n element, where every of such subset is equally likely $h(W+B,W,n)$. for a hyper geometric $H(W+B,W,n)$ random variable X. (i) The sample space s is the set of integers that meet max $(0, n-B) < x < min\,(n,w)$ and (ii) The probability mass function is in

$$P\,(X = x/(W+B,w,n) =$$

$$P(x) = \frac{\binom{w}{x}\binom{B}{n-x}}{\binom{w+B}{n}}$$

The max function chooses the larger of the listed values (since x has to be bigger than both numbers) in the same way min chooses the smaller.
Fog in [3] has suggested that the best way to avoid confusion is to use the name Wallenius non central hyper geometric distribution of a biased urn model where a predetermined number of items are drawn one by one in a competitive manner while the name Fisher's non central hyper geometric distribution is used where items are drawn independently of each other. This is done so that the total number of items drawn is known only after the experiment. The names refer to Kenneth Ted Wallenius and R.A. Fisher who were the first to describe the respective distribution.

2. AIM AND OBJECTIVES

This study is aimed at application of two non-central hyper-geometric distributions under statistical models of biased sampling. It is therefore designed in line with the following objectives:

1. To examine the procedure of applying the two non – central hyper geometric distributions (Wallenius' and Fisher's distributions.)
2. To investigate the statistical properties of the two non – central hypergeometric distributions such as (mean, variance and coefficient of variation).
3. To compare the coefficient of variation of the two non – central hypergeometric distributions in univariate and multinomial cases

APPLICATIONS	
Wallenius	**Fishers'**
1. Wallenius distribution is used in models of natural selection and biased sampling	1. Fisher's non-central hyper-geometric distribution is useful for models of biased sampling or biased.
2. Wallenius' non-central hyper-geometric distribution were used when items are sampled one by one with competition.	2. Fisher's non-central hyper-geometric distribution can also be used to select on items sampled.
3. The distribution is applicable in random number theory.	3. The distribution can also be used for test in contingency tables where a conditional distribution for fixed margin is desired.

CONDITIONS	
Wallenius	**Fishers'**
1. Items are taken randomly from a finite source containing different kinds of items without replacement.	1. Items are taken randomly from a finite source containing different kinds of items without replacement.
2.Items are drawn one by one.	2. Items are taken independently of each other. Whether one item is taken is independent of whether another item is taken. Whether one item is taken before, after, or simultaneously with another item is irrelevant.
3. The probability of taking a particular item at a particular draw is equal to its fraction of the total weight of all items that have not yet been taken at that moment. The weight of an item depends only on its kind.	3. The probability of taking a particular item is proportional to its weight. The weight of an item depends only on its kind.
4. The total number *n* of items to take is fixed and independent of which items happen to be taken first.	4. The total number *n* of items that will be taken is not known before the experiment.
	5. *n* is determined after the experiment and the conditional distribution for *n* known is desired.

3. METHODOLOGY

Central Hypergeometric Distribution

Suppose X_1 and X_2 represent two independent binomial random variables with parameter (m_1, π) and (m_2, π) respectively. Then $X_1 + X_2$ has a binomial distribution with parameters $N = m + m_2$ and π. The conditional distribution of X_1 given $X_1 + X_2 = n$ is the univariate central hypergeometric distribution and is derived as follows;

$$P(X_1 = x_1) = \binom{m_1}{x_1} \pi^{x_1} (1-\pi)^{m_1 - x_1}$$

$$P(X_2 = x_2) = \binom{m_2}{x_2} \pi^{x_2} (1-\pi)^{m_2 - x_2}$$

$$P(X_1 + X_2 = n) = \binom{m_1 + m_2}{n} \pi^{n} (1-\pi)^{m_1 + m_2 - n}$$

Now if we let $x_1 = x$ and $x_2 = n - x$. then the conditional distribution becomes;

$$P(X_1 = x|n) = \frac{\left\{\binom{m_1}{x}\pi^x(1-\pi)^{m_1-x}\right\}\left\{\binom{m_2}{n-x}\pi^x(1-\pi)^{m_2-(n-x)}\right\}}{\binom{m_1+m_2}{n}\pi^n(1-\pi)^{m_1+m_2-n}}$$

Collecting the exponent x together leads to;

$$P(X_1 = x|n) = \frac{\binom{m_1}{x}\binom{m_2}{n-x}}{\binom{m_1+m_2}{n}}$$

which is the required mass function with support; $\max(0, m_1 + n - N) \le X \le \min(m_1, n)$

For univariate central hypergeometric distribution, the above p.m.f has the corresponding mean and variance respectively.

$$E(x) = \frac{nM_1}{N}$$

$$V(x) = \frac{nM_1(N - M_1)(N - m)}{N^2(N - 1)}$$

For the multinomial hypergeometric distribution, the p.m.f is as follow;

$$\prod_{i=1}^{c} \frac{\binom{M_i}{m_i}}{\binom{N}{n}}$$

The corresponding mean and variance are;

$$E(x) = \frac{nM_i}{N}$$

$$V(x) = \frac{nM_i(N - M_i)(N - m)}{N^2(N - 1)}$$

The above p.m.f can be extended based on the violation of equal probability assumption of the two binomial random variables, then this leads to non-central hypergeometric distribution since the distribution of the sum is no longer binomial. The proof according to ([Law03]) is;

$$P(X_1 = x|n) = \frac{\left\{\binom{m_1}{x}\pi_1{}^x(1-\pi_1)^{m_1-x}\right\}\left\{\binom{m_2}{n-x}\pi_2{}^{n-x}(1-\pi_2)^{m_2-(n-x)}\right\}}{\sum_j\left\{\binom{m_1}{j}\pi_1{}^j(1-\pi_1)^{m_1-j}\right\}\left\{\binom{m_2}{n-j}\pi_2{}^{n-j}(1-\pi_2)^{m_2-(n-j)}\right\}}$$

Collecting the exponent x together again leads to;

$$P(X_1 = x|n) = \frac{\binom{m_1}{x}\binom{m_2}{n-x}\left(\frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}\right)^x}{\sum_j\binom{m_1}{j}\binom{m_2}{n-j}\left(\frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}\right)^j}$$

Put

$$\omega = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}$$

then;

$$P(X_1 = x|n) = \frac{\binom{m_1}{x}\binom{m_2}{n-x}\omega^x}{\sum_j\binom{m_1}{j}\binom{m_2}{n-j}\omega^j}$$

Univariate fishers' non-central hypergeometric distribution

The above p.m.f was referred to as the extended hypergeometric distribution by [2], where is the non-centrality parameter. However, the above non-central hypergeometric was given the name fishers hypergeometric distribution by [3].

The corresponding mean and variance according to [6] are;

$$E(X;\omega) = \frac{P_1(\omega)}{P_0(\omega)}$$

$$V(X;\omega) = \frac{P_2(\omega)}{P_0(\omega)} - \{P_1(\omega)/P_0(\omega)\}^2$$

Where $P_r(\omega)$ is the polynomial

$$P_r(\omega) = \sum_j \binom{m_1}{j}\binom{m_2}{n-j} j^r\omega^j$$

The moments about the origin are expressible as;

$$\mu_r(\omega) = \frac{P_r(\omega)}{P_0(\omega)}$$

Multinomial Fisher Non-Central Hypergeometric Distribution

Suppose that $X_1 \sim M(m_1, \pi_1)$ and $X_2 \sim M(m_2, \pi_2)$ are independent random variables on k categories each. Then the conditional distribution of given $X_1 + X_2 = n$ that $X = X_1$ is as follows;

In the above equation equal probability assumption is assumed and thus the probabilities will cancel out.

For the **multinomial fishers hypergeometric** case where the equal probability is not assumed, the resultant p.m.f follows from the univariate case provided above.

$$\mathrm{P}(X = X_1|n) = \frac{\binom{m_1}{x}\binom{m_2}{n-x}}{\binom{m_1+m_2}{n}}$$

([MN89]) defined the p.m.f as;

$$P(X = X_1|n) = \frac{\binom{m_1}{x}\binom{m_2}{n-x}\omega_1^{x_1}\cdots\omega_k^{x_k}}{\sum_j\binom{m_1}{j}\binom{m_2}{n-j}\omega_1^{j_1}\cdots\omega_k^{j_k}}$$

Where

$$\omega_j = \frac{\pi_{1j}\pi_{2k}}{\pi_{2j}\pi_{1k}}$$

The corresponding approximate relationship between the mean and variance according to [6] is

$$\frac{E(X_{1j}, X_{2k})}{E(X_{2j}, X_{1k})} = \omega_j = \frac{\mu_{1j}\mu_{2k} - \sigma_{jk}}{\mu_{2j}\mu_{1k} - \sigma_{jk}}$$

Where,

$$\sigma_{jk} = cov(X_{1j}, X_{1k}) = -cov(X_{1j}, X_{2k})$$

Is negative for $j < k$

The covariance matrix $\sum$ may be approximated quite accurately as follows. If we let the vector ζ with components ζ_j given by;

$$\frac{1}{\zeta_j} = \frac{1}{\mu_{1j}} + \frac{1}{\mu_{2j}}$$

The approximate $\sum$ is then given in terms of ζ by;

$$\hat{\Sigma} = \frac{n}{n-1}\{diag(\zeta) - \zeta\zeta'/\zeta.\}$$

Univariate Wallenius Non-Central Hypergeometric Distribution

The noncentral hypergeometric distribution is the name given by [7] to a distribution constructed by supposing that in sampling without replacement the probability of drawing a white ball given that there are m_1 white and m_2 black balls is not p but $p/[p + \omega(1 - p)]$ with $\omega \neq 1$.

The mathematical analysis following from this assumption is complicated. Starting from the recurrence relationship;

$$P(X = x|m_1, m_2, N) = \frac{pP(X = x-1|m_1 - 1, m_2 - 1, N-1)}{p + \omega(1 - p)} + \frac{\omega(1-p)P(X = x|m_1, m_2 - 1, N-1)}{p + \omega(1 - p)}$$

Wallenius obtained the formular;

$$P(X = x) = \binom{m_1}{x}\binom{m_2}{n-x}\int_0^1 (1-t^c)^x\left(1-t^{\omega/c}\right)^{n-x} dt$$

With $c = [m_1 - x + \omega(N - m_1 - n + x)]^{-1}$

For the univariate case with $\omega > 1$, [3] indicated that it may be more efficient to solve;

$$\left(1-\frac{\mu_1}{m_1}\right)^{1/\omega_1} = \left(1-\frac{\mu_2}{m_2}\right)^{1/\omega_2} = \ldots \left(1-\frac{\mu_k}{m_k}\right)^{1/\omega_k}$$

An approximation to the variance can be obtained by approximating Wallenius' non-central hypergeometric distribution with a Fisher's non-central hypergeometric distribution with the same mean and using an approximate formula given by [5] for the variance of the latter distribution;

$$\sigma^2 \approx \sigma_F^2 = \frac{Nab}{(N-1)(mb + (N-m)a)}$$
$$a = \mu(m-\mu), b = (n-\mu)(\mu + N - n - m)$$

This approximation is good when w is near 1 and n is far from N .

Multinomial Wallenius Non-central Hypergeometric Distribution

Accordingly [1] extended the univariate Wallenius distribution to the multinomial case given by;

$$P(\boldsymbol{X} = \boldsymbol{x}) = \prod_{i=1}^{k}\binom{m_i}{x_i}\int_0^1 \prod_{i=1}^{k}\left(1 - t^{\frac{\omega_i}{d}}\right)^{x_i} dt$$

$$d = \boldsymbol{\omega}(\boldsymbol{m} - \boldsymbol{x})$$

Where, $\boldsymbol{m} = (m_1, m_2, \ldots m_k)$; $\boldsymbol{x} = (x_1, x_2, \ldots x_k)$; $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots \omega_k)$

The approximate mean and variance for Wallenius was derived by [3] using;

$$\mu_{iv} = \mu_{i(v-1)} + p_{iv}(\boldsymbol{\mu}_{v-1})$$

Where,

$$\boldsymbol{\mu}_v = (\mu_1, \mu_2, \ldots \mu_k)$$

Notations

$x = (x_1, x_2, \ldots x_k)$
is the number of balls drawn of each color

$\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots \omega_k)$
is the initial number of balls of each color in the urn

$\boldsymbol{m} = (m_1, m_2, \ldots m_k)$
is the weight or odds of balls of each Color

$n = \sum x$ is the total number of balls drawn
k = is the number of colors
N = is the total number of balls in urn before sampling

4. SIMULATION STUDY

In this study the direct inversion method of *R* statistical software were used with varying Odds ratio and group sizes. The aim is to investigate and compare the statistical properties (mean, variance and coefficient of variation) of the two non-central hypergeometric distributions in relation to the central hypergeometric distribution. Also to investigate the consistency nature of the distributions with Random number generated of (10, 50, 100, 500, 1000) were also considered. The samples generated were replicated 1000 times to ensure stability of the results.

RESULTS AND DISCUSSIONS

Table 1 Univariate Case: **Mean** result of simulation based on; $m_1 = 80$; $m_2 = 20$; **n=20;** $0 < \omega < 1$

Random number generated	Odds ratio ω	Wallenius Distribution	Fisher Distribution	Central Hypergeometric Distribution
10	$\omega = 0.2$	10.12	11.01	16
	$\omega = 0.5$	13.71	14.01	16
	$\omega = 0.7$	14.89	15.02	16
	$\omega = 0.9$	15.69	15.72	16
50	$\omega = 0.2$	10.12	11.01	16
	$\omega = 0.5$	13.71	14.01	16
	$\omega = 0.7$	14.89	15.02	16
	$\omega = 0.9$	15.69	15.72	16
100	$\omega = 0.2$	10.12	11.01	16
	$\omega = 0.5$	13.71	14.01	16
	$\omega = 0.7$	14.89	15.02	16
	$\omega = 0.9$	15.69	15.72	16
500	$\omega = 0.2$	10.12	11.01	16
	$\omega = 0.5$	13.71	14.01	16
	$\omega = 0.7$	14.89	15.02	16
	$\omega = 0.9$	15.69	15.72	16
1000	$\omega = 0.2$	10.12	11.01	16
	$\omega = 0.5$	13.71	14.01	16
	$\omega = 0.7$	14.89	15.02	16
	$\omega = 0.9$	15.69	15.72	16

Table 2

Univariate Case: Coefficient of variation results of simulation $m_1 = 80$; $m_2 = 20$; $n = 20$; $\omega = 0.2$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	17.9505	16.52444	10.05842
50	17.9505	16.52444	10.05842
100	17.9505	16.52444	10.05842
500	17.9505	16.52444	10.05842
1000	17.9505	16.52444	10.05842

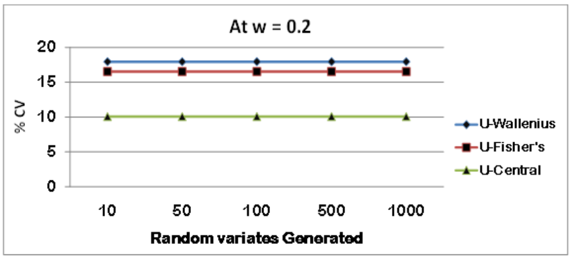

Fig 1: Plot of coefficient of variation at Various Random sample generated of odds ratio = 0.2.

Table 3

Univariate Case: Coefficient of variation results of simulation $m_1 = 80$; $m_2 = 20$; $n = 20$; $\omega = 0.5$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	12.9865	12.6078	10.05842
50	12.9865	12.6078	10.05842
100	12.9865	12.6078	10.05842
500	12.9865	12.6078	10.05842
1000	12.9865	12.6078	10.05842

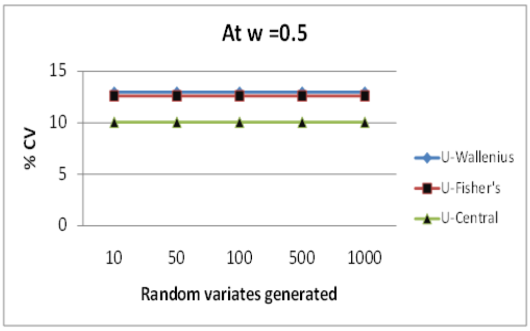

Fig 2: Plot of coefficient of variation at Various Random sample generated of odds ratio = 0.5

Table 4

Univariate Case: Coefficient of variation results of simulation $m_1 = 80; m_2 = 20$; $n = 20$; $\omega = 0.7$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	11.4958	11.33781	10.05842
50	11.4958	11.33781	10.05842
100	11.4958	11.33781	10.05842
500	11.4958	11.33781	10.05842
1000	11.4958	11.33781	10.05842

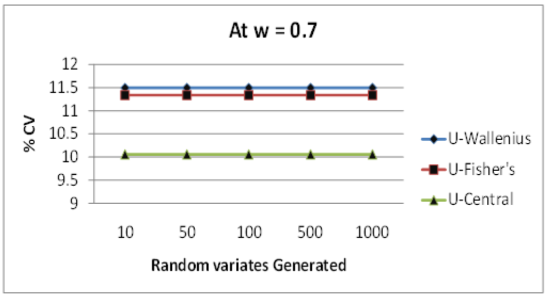

Fig 3: Plot of coefficient of variation at Various Random sample generated of odds ratio = 0.7

Table 5

Univariate Case: Coefficient of variation results of simulation $m_1 = 80; m_2 = 20$; $n = 20$; $\omega = 0.9$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	10.45329	10.41393	10.05842
50	10.45329	10.41393	10.05842
100	10.45329	10.41393	10.05842
500	10.45329	10.41393	10.05842
1000	10.45329	10.41393	10.05842

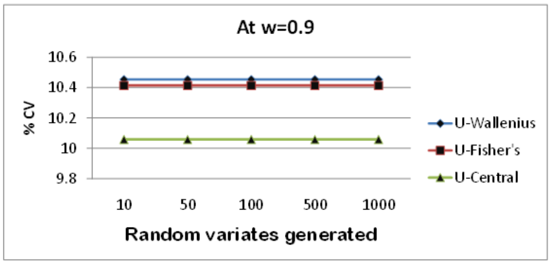

Fig 4: Plot of coefficient of variation at Various Random sample generated of odds ratio = 0.9

Table 6

Multinomial Case: **Mean** result of simulation based on m_i = (80,20,30,**40**); ω = (0.2,0.5,0.7,0.9) ; $n = 20$

Random Number generated	Odds ratio ω	Wallenius Distribution	Fisher Distribution	Central hypergeomatric Distribution
10	$\omega = 0.2$	4.08	4.3	9.41
	$\omega = 0.5$	2.46	2.5	2.35
	$\omega = 0.7$	5.04	5.0	3.53
	$\omega = 0.9$	8.43	8.2	4.71
50	$\omega = 0.2$	4.08	4.3	9.41
	$\omega = 0.5$	2.46	2.5	2.35
	$\omega = 0.7$	5.04	5.0	3.53
	$\omega = 0.9$	8.43	8.2	4.71
100	$\omega = 0.2$	4.08	4.3	9.41
	$\omega = 0.5$	2.46	2.5	2.35
	$\omega = 0.7$	5.04	5.0	3.53
	$\omega = 0.9$	8.43	8.2	4.71
500	$\omega = 0.2$	4.08	4.3	9.41
	$\omega = 0.5$	2.46	2.5	2.35
	$\omega = 0.7$	5.04	5.0	3.53
	$\omega = 0.9$	8.43	8.2	4.71
1000	$\omega = 0.2$	4.08	4.3	9.41
	$\omega = 0.5$	2.46	2.5	2.35
	$\omega = 0.7$	5.04	5.0	3.53
	$\omega = 0.9$	8.43	8.2	4.71

Table 7

Multinomial Case: Variance results of simulation based on m_1 = (80,20,30,40); ω = (0.2,0.5,0.7,0.9) ; $n = 20$

Random number generated	Odds ratio ω	Wallenius Distribution	Fisher Distribution	Central Hypergeometric Distribution
10	$\omega = 0.2$	3.0	3.11	4.42
	$\omega = 0.5$	1.89	1.92	1.84
	$\omega = 0.7$	3.18	3.17	2.58
	$\omega = 0.9$	4.07	4.04	3.19
50	$\omega = 0.2$	3.0	3.11	4.42
	$\omega = 0.5$	1.89	1.92	1.84
	$\omega = 0.7$	3.18	3.17	2.58
	$\omega = 0.9$	4.07	4.04	3.19
100	$\omega = 0.2$	3.0	3.11	4.42
	$\omega = 0.5$	1.89	1.92	1.84
	$\omega = 0.7$	3.18	3.17	2.58
	$\omega = 0.9$	4.07	4.04	3.19
500	$\omega = 0.2$	3.0	3.11	4.42
	$\omega = 0.5$	1.89	1.92	1.84
	$\omega = 0.7$	3.18	3.17	2.58
	$\omega = 0.9$	4.07	4.04	3.19
1000	$\omega = 0.2$	3.0	3.11	4.42
	$\omega = 0.5$	1.89	1.92	1.84
	$\omega = 0.7$	3.18	3.17	2.58
	$\omega = 0.9$	4.07	4.04	3.19

Table 8

Multinominal Case: Coefficient of Variation result of simulation based on $m_i = (80,20,30,\mathbf{40})$; $\omega = (0.2)$ $n = 20$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	42.45223	41.01207	22.34197
50	42.45223	41.01207	22.34197
100	42.45223	41.01207	22.34197
500	42.45223	41.01207	22.34197
1000	42.45223	41.01207	22.34197

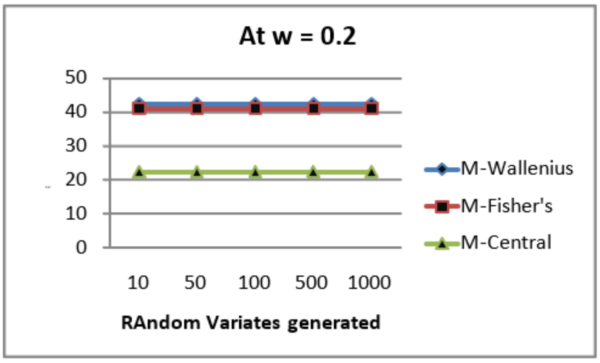

Fig 5: Plot of coefficient of variation for Multinomial Distribution at Various Random sample generated and Odds ratio = 0.2.

Table 9

Multinomial Case: Coefficient of Variation result of simulation based on $m_i = (80,20,30,\mathbf{40})$; $\omega = (0.5)$ $n = 20$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	55.88507	55.42563	57.72196
50	55.88507	55.42563	57.72196
100	55.88507	55.42563	57.72196
500	55.88507	55.42563	57.72196
1000	55.88507	55.42563	57.72196

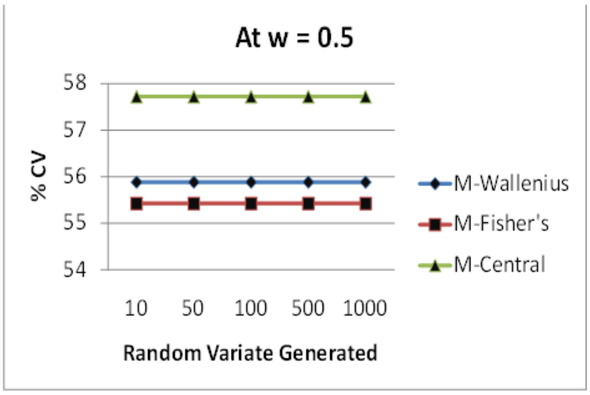

Fig 6: Plot of coefficient of variation for Multinomial Distribution at Various Random sample generated and Odds ratio = 0.5.

Table 10

Multinomial Case: Coefficient of Variation result of simulation based on $m_i = (80,20,30,\mathbf{40})$; $\omega = (0.7)$ $n = 20$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	35.38205	35.60899	45.50249
50	35.38205	35.60899	45.50249
100	35.38205	35.60899	45.50249
500	35.38205	35.60899	45.50249
1000	35.38205	35.60899	45.50249

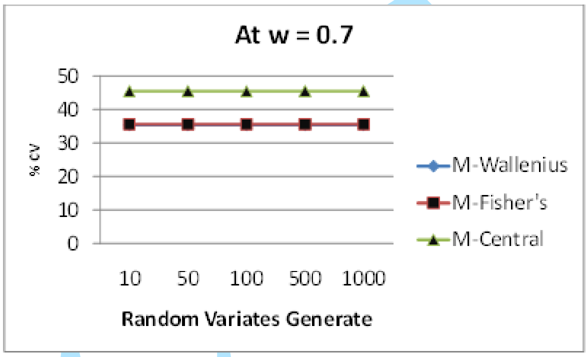

Fig 7: Plot of coefficient of variation for Multinomial Distribution at Various Random sample generated and Odds ratio = 0.7

Table 11

Multinomial Case: Coefficient of Variation result of simulation based on $m_i = (80,20,30,\mathbf{40})$; $\omega = (\mathbf{0.9})$ $n = 20$

Random Number generated	Wallenius distribution	Fisher distribution	Central hypergeometric distribution
10	23.93148	24.51189	37.92053
50	23.93148	24.51189	37.92053
100	23.93148	24.51189	37.92053
500	23.93148	24.51189	37.92053
1000	23.93148	24.51189	37.92053

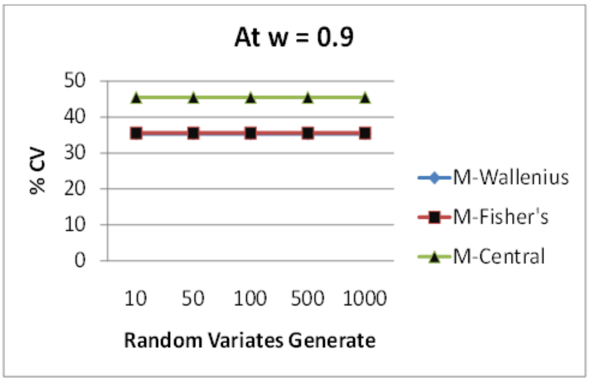

Fig 8: Plot of coefficient of variation for Multinomial Distribution at Various Random sample generated and Odds ratio = 0.9

5. DISCUSSION

The simulation result in table 1 and 6-9 show that with respect to random sample generated (10, 50, 100, 500, 1000) the estimated mean, the variance and coefficient of variation are approximately the same for the two non-central distributions with varying odds ratios (ω = 0.2, 0.5, 0.7, 0.9).

In table 1 and 6 – 9, the non-central hypergeometric distribution (Wallenius and Fishers') posses a closely approximately estimate of mean, variance and coefficient of variation differ from central hypergeometric distribution.

In univariate cases, table 2 – 5 and figure 1 – 4, it was observed that Fishers distribution at (ω = 0.2, 0.5, 0.7, 0.9) is more consistent than Wallenius distribution although central hypeergeometric is more better.

In multinomial cases, table 8 – 11 and figure 5 - 8, it was observed that Fisher distribution is more consistent at ω =0.5, Wallenius distribution at ω = 0.7, 0.9 and central hypergeometric distribution at ω = 0.2.

CONCLUSION

From the aforementioned, it can be concluded that:

- The two non – central hypergeometric distributions (Wallenius and Fishers') are approximately equal in estimated mean, variance and co efficient of variation with respect to random sample generated.
- The difference between the two non – central hypergeometric distributions becomes higher when the odd ratio close to 1
- The two non – central hypergeometric distributions differ from the central hypergeometric when odd ratio is 1
- The two non – central hypergeometric distributions approximately each other when they have same mean than when they have same ratio
- In univariate case, Fisher distribution are more consistent than Wallenius distribution while in multinomial case both distribution perform differently.

REFERENCES

[1] **Chesson J. A. -** A non - central multinomial hypergeometric distribution arising from biased sampling with application to selective predation. Pg (795 - 797), 1976.

[2] **Fishers R. A. -** The mathematical theory of probabilities and its application to frequency - curves and statistical method Vol 1 New York; Macmillan, 1935.

[3] **Fog A. -** "Calculation methods for Wallenius Non-centrall hypergeometric Distribution" communication ins Statistics Simulation and computation 37 (2): 258 – 273, 2008.

[4] **Lawal H. B. -** "Categorical Data Analysis with SAS and SPSS Applications. St Cloud University, 2003.

[5] **Levin B. -** "Compound multinomial likelihood functions are uni - model: proof of a conjecture of I. J. Good, Animals of statistics, 5, 79-87 [5.8.5], 2007.

[6] **Mc cullagh . P. and Nelder J. A. -** "Generalized linear models, London; Chappman & Hall (11.1.1), 1989.

[7] **Wallenius K.T. -** Biased Sampling. The non-central hypergeometric Distribution Ph.D. thesis, Stanford university, Stanford C.A, 1963.

Dr. Olarewaju Kamoldeen Adetunji **University of Ilorin** **Faculty of Physical Sciences** **Department of Statistics** **Researcher**	
Dr. (Mrs.) Taoheedat Alanamu **University of Ilorin** **Faculty of Physical Sciences** **Department of Statistics** **Researcher**	
Mrs. Enilara Justina Adefila **University of Ilorin** **Faculty of Physical Sciences** **Department of Statistics** **Researcher**