International Journal of Statistics and Applied Mathematics 2019; 4(1): 23-28

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2019; 4(1): 23-28

www.mathsjournal.com
Received: 06-11-2018
Accepted: 10-12-2018

P Rajini
Lecturer, Bhavan's Vivekananda Degree College, Sainikpuri, Secunderabad, Telangana, India

Krishna Reddy M
Professor, Department of Statistics, University College of Science, O.U., Hyderabad, Telangana, India

Jaya lakshmi C
Professor, Department of Statistics, University College of Science, O.U., Hyderabad, Hyderabad, Telangana, India

Correspondence
P Rajini
Lecturer, Bhavan's Vivekananda Degree College, Sainikpuri, Secunderabad, Telangana, India

Comparison of estimators of common mean of two normal populations using Monte Carlo simulation

P Rajini, Krishna Reddy M and Jaya lakshmi C

Abstract
Estimating common mean of several normal populations with unknown variances, which may be equal or unequal is one of the interesting problems in statistical inference. Estimation of common mean based on independent samples has some unresolved problems in literature. In this paper we considered different estimators of common mean available in the literature and compared them with respect to their bias, standard error, skewness and kurtosis.

Keywords: Common mean, normal population, unequal variances, standard error, relative bias, different estimators

1. Introduction
In estimating the common mean of several normal populations with unknown and unequal variances it is clear that the distribution of any combined estimator of μ will involve nuisance parameters. The problem of estimation of common mean μ of two or more univariate normal populations with unknown and unequal variances based on independent samples of fixed sizes has some unresolved problems. Uniformly minimum variance unbiased estimator of μ in this problem does not exist. Estimation of treatment effect in a BIBD with uncorrelated random block effects, by suitably combining the intra block estimate and the inter block estimate is the motivation to march towards the problem of estimation of common mean of several normal populations with unequal and unknown variances. The main objective of this paper is to compare the several estimators of common mean of two normal populations with respect to their Relative bias and standard error for different sample sizes using Monte Carlo Simulation.

2. Different estimators of common mean
We considered eleven estimators proposed by different authors.
2.1. The Graybill – Deal (1959) defined an estimator for common mean as,

$$T_1 = \hat{\mu} = \frac{n_1 \bar{x} s_2^2 + n_2 \bar{y} s_1^2}{n_1 s_2^2 + n_2 s_1^2} \quad \dots (2.1)$$

They defined that it is an unbiased estimator of μ and is uniformly better than sample mean for $n_1, n_2 \geq 10$.

2.2. Brown and Cohen (1974) considered the estimators of the following forms

$$T_2 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x}) \left\{ \frac{a_0 \bar{s}_1^2}{\bar{s}_1^2 + (n_2 - 1)\left(\frac{\bar{s}_2^2}{n_2 + 2}\right) + \left(\frac{(\bar{y} - \bar{x})^2}{n_2 + 2}\right)} \right\} \quad \dots (2.2)$$

and $T_3 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x}) \left\{ \frac{a_0 \bar{s}_1^2}{\bar{s}_1^2 + \bar{s}_2^2} \right\}$ … (2.3)

Where, $a_0 = \frac{(n_1 - 1)(n_2 - 2)}{(n_1 + 1)(n_2 + 2)}$

2.3. Bhattacharya (1980) proposed the following two estimators for the common mean of two normal populations

$$T_4 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x})\left\{\frac{a_0 \bar{s}_1^2}{\bar{s}_1^2 + c\left[\bar{s}_2^2 + \left(\frac{(\bar{y}-\bar{x})^2}{(n_2-1)}\right)\right]}\right\} \quad \dots (2.4)$$

Where, $a_0 = \frac{n_2-1}{n_1+1}$ and $c = 0.5\left(\frac{n_2-1}{n_1-1}\right)$

and $$T_5 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x})\left\{\frac{a_0 \bar{s}_1^2}{\bar{s}_1^2 + c\bar{s}_2^2}\right\} \quad \dots (2.5)$$

where, $a_0 = \frac{n_2-5}{n_1+1}$ and c as defined above for T_4

2.4. Khatri and Shah (1974) proposed the following two estimators of common mean.

$$T_6 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x})\left\{\frac{\bar{s}_1^2}{\bar{s}_1^2 + 0.5\left(\frac{n_2-1}{n_1-1}\right)^2\left[\bar{s}_2^2 + \left(\frac{(\bar{y}-\bar{x})^2}{(n_2-1)}\right)\right]}\right\} \quad \dots (2.6)$$

and $$T_7 = \hat{\mu} = \bar{x} + (\bar{y} - \bar{x})\left\{\frac{\bar{s}_1^2}{\bar{s}_1^2 + 0.5\left(\frac{n_2-1}{n_1-1}\right)^2 \bar{s}_2^2}\right\} \quad \dots (2.7)$$

2.5. Zacks (1966) considered two classes of unbiased estimators of common mean

$$T_8 = \hat{\mu} = I\left(\frac{s_2}{s_1}; \rho^*\right)\mu^* + \left\{1 - I\left(\frac{s_2}{s_1}; \rho^*\right)\right\}\mu^{**};\ 1 \le \rho^* \le \infty \quad \dots (2.8)$$

$$T_9 = \hat{\mu} = I\left(\frac{s_2}{s_1}; \rho^*\right)\mu^* + J_1\left(\frac{s_2}{s_1}; \rho^*\right)\bar{x} + J_2\left(\frac{s_2}{s_1}; \rho^*\right)\bar{y} \quad \dots (2.9)$$

where, $I\left(\frac{s_2}{s_1}; \rho^*\right) = \begin{cases} 1, if\ \frac{1}{\rho^*} \le \left(\frac{s_2^2}{s_1^2}\right) \le \rho^* \\ 0, otherwise \end{cases}$

$$J_1\left(\frac{s_2}{s_1}; \rho^*\right) = \begin{cases} 1, if\ \left(\frac{s_2^2}{s_1^2}\right) > \frac{1}{\rho^*} \\ 0, otherwise \end{cases}$$

$$J_2\left(\frac{s_2}{s_1}; \rho^*\right) = \begin{cases} 1, if\ \left(\frac{s_2^2}{s_1^2}\right) < \frac{1}{\rho^*} \\ 0, otherwise \end{cases}$$

$$\mu^* = \frac{(\bar{x} + \bar{y})}{2},\ \mu^{**} = \left(\frac{\frac{s_2^2}{s_1^2}\bar{x} + \bar{y}}{\frac{s_2^2}{s_2^2} + 1}\right)$$

2.6. Cohen and Sackrowitz (1974) proposed the common mean estimator for the equal sample size as

$$T_{10} = \hat{\mu} = \left(1 - c_n^*\, G^*(\hat{\rho})\right)\bar{x} + \left(c_n^*\, G^*(\hat{\rho})\right)\bar{y} \quad \dots (2.10)$$

where, $c_n^* = \begin{cases} \frac{(n-3)^2}{(n-1)(n+1)}\ ;\ if\ n\ is\ odd \\ \frac{n-4}{n+2}\ ;\ if\ n\ is\ even \end{cases}$

and $G^*(\hat{\rho}) = \frac{1}{1 + \left(\frac{s_2^2}{s_1^2}\right)}$

2.7. Krishnamoorthy and Rohatgi (1988) proposed a new estimator of common mean based on transformed variables u and v as

$$T_{11} = \hat{\mu} = \bar{u} - c\, b_0 \bar{v} \quad \dots (2.11)$$

where, u = x; v = x – y; $c = \frac{n-2}{n+1}$; $s_u = \sum_{i=1}^{n}(u - \bar{u})^2$; $s_v = \sum_{i=1}^{n}(v - \bar{v})^2$;

and $b_0 = \frac{s_u}{(s_v + n\,\bar{v}^2)}$

In all the above estimators $\bar{x} = \sum_{i=1}^{n_1} \frac{x_i}{n_1},\ s_1^2 = \sum_{i=1}^{n_1} \frac{(x_i - \bar{x})^2}{(n_1 - 1)}$, $\bar{s}_1^2 = \frac{s_1^2}{n_1}$ and $\bar{\sigma}_1^2 = \frac{\sigma_1^2}{n_1}$ similarly $\bar{y}, s_2^2, \bar{s}_2^2$ and $\bar{\sigma}_2^2$

All these estimators can be used to estimate the common mean when $n_1 = n_2$. When $n_1 \neq n_2$ only the estimators given from equation (2.1) to (2.7) can be used.

3. Comparison of Estimators

While comparing the two sample means we encounter three different cases about the variance ratios of the populations. They are

a. $\rho = (\sigma_2^2/\sigma_1^2) = 1$
b. $\rho = (\sigma_2^2/\sigma_1^2) > 1$
c. $\rho = (\sigma_2^2/\sigma_1^2) < 1$

In this paper we presented simulation results of the case $\rho = 1$ only. We generated 2000 samples of sizes n_1 and n_2 from Normal $(10, \sigma_i^2)$; i = 1, 2 populations. We considered values of n_1 and n_2 from 5 to 50 with a gap of 5. Through simulation studies the samples are generated from two normal populations with parameters $\mu = 10$, $\sigma_1^2 = 4 = \sigma_2^2$. Keeping sample size n_2 fixed at 5, n_1 is from 5 to 50 with a gap of 5. When $n_1 = n_2$, for each combination of (n_1, n_2) we computed all 11 estimators, where as $n_1 \neq n_2$ we computed only 7 estimators. In both the cases all these estimators we computed common mean, standard error, relative bias, skewness and kurtosis coefficients. The obtained simulation study's results are presented in the Table 1. Due to lack of space we are presenting the results for n1, n2 = 5, 15, 30, 50.

Table 1: Comparison of estimators of common mean when the population variance ratio $\rho = 1$ (n_1 = 5, 15, 30, 50 and n_2 = 5)

n_1	n_2	Estimator	$\hat{T}_i$	SE($\hat{T}_i$)	RB($\hat{T}_i$)	β_1	β_2
5	5	T_1	10.145	0.485	0.014	0.033	2.447
		T_2	10.104	0.572	0.010	0.003	3.082
		T_3	10.111	0.534	0.011	0.009	2.998
		T_4	10.113	0.505	0.011	0.011	2.991
		T_5	10.085	0.771	0.008	-0.002	2.942
		T_6	10.128	0.482	0.013	0.023	2.688
		T_7	10.138	0.499	0.014	0.031	2.434
		T_8	10.111	0.408	0.011	0.069	2.548
		T_9	10.116	0.416	0.012	0.070	2.509
		T_{10}	10.095	0.662	0.009	0.000	2.970
		T_{11}	10.097	0.601	0.010	-0.003	3.183
15	5	T_1	10.127	0.224	0.013	0.527	3.898
		T_2	10.116	0.212	0.012	0.646	4.571
		T_3	10.120	0.209	0.012	0.761	4.640
		T_4	10.114	0.207	0.011	0.693	4.665
		T_5	10.109	0.241	0.011	0.499	4.519
		T_6	10.130	0.600	0.013	0.012	2.478
		T_7	10.137	0.708	0.014	0.023	2.569
30	5	T_1	10.101	0.149	0.010	0.102	4.413
		T_2	10.106	0.122	0.011	-0.022	3.184
		T_3	10.099	0.128	0.010	0.000	3.352
		T_4	10.101	0.118	0.010	-0.034	3.168
		T_5	10.095	0.128	0.010	-0.027	3.015
		T_6	10.129	0.628	0.013	0.015	2.914
		T_7	10.123	0.697	0.012	0.002	3.113
50	5	T_1	10.113	0.076	0.011	-0.018	2.616
		T_2	10.116	0.069	0.012	-0.001	2.465
		T_3	10.113	0.070	0.011	-0.005	2.494
		T_4	10.113	0.068	0.011	-0.004	2.523
		T_5	10.113	0.072	0.011	-0.010	2.591
		T_6	10.086	0.724	0.009	-0.016	2.729
		T_7	10.082	0.764	0.008	-0.037	2.987

Similarly, simulation studies are carried out by fixing n_2 at 15, 30, 50, varying n_1 = 5 from 50 with a gap of 5, in each case. The obtained results are presented in Tables 2 to 4.

Table 2: Comparison of estimators of common mean when the population variance ratio ρ = 1 (n_1= 5, 15, 30, 50 and n_2=15)

n_1	n_2	Estimator	$\hat{T}_i$	SE($\hat{T}_i$)	RB($\hat{T}_i$)	β_1	β_2
5	15	T1	10.163	0.287	0.016	-0.012	2.676
		T2	10.124	0.447	0.012	0.018	2.802
		T3	10.129	0.403	0.013	0.012	2.688
		T4	10.218	0.532	0.022	-0.100	3.266
		T5	10.205	0.406	0.020	-0.078	3.142
		T6	10.113	0.612	0.011	0.026	2.921
		T7	10.129	0.502	0.013	0.050	2.852
15	15	T1	10.139	0.202	0.014	-0.001	3.042
		T2	10.126	0.218	0.013	0.003	2.947
		T3	10.131	0.216	0.013	0.001	2.897
		T4	10.130	0.192	0.013	-0.004	3.130
		T5	10.126	0.205	0.013	0.000	2.942
		T6	10.133	0.190	0.013	-0.008	3.249
		T7	10.139	0.194	0.014	-0.004	3.394
		T8	10.120	0.191	0.012	-0.002	3.140
		T9	10.128	0.205	0.013	0.010	3.442
		T10	10.127	0.229	0.013	0.006	2.848
		T11	10.121	0.225	0.012	0.000	3.071
30	15	T1	10.131	0.093	0.013	-0.030	2.609
		T2	10.127	0.091	0.013	-0.026	2.702
		T3	10.128	0.091	0.013	-0.030	2.670
		T4	10.125	0.089	0.013	-0.030	2.658
		T5	10.121	0.091	0.012	-0.031	2.708
		T6	10.146	0.163	0.015	-0.008	2.573
		T7	10.146	0.172	0.015	-0.008	2.614
50	15	T1	10.123	0.065	0.012	-0.049	2.906
		T2	10.118	0.064	0.012	-0.030	2.935
		T3	10.118	0.063	0.012	-0.031	2.928
		T4	10.111	0.061	0.011	-0.024	2.871
		T5	10.107	0.061	0.011	-0.007	2.934
		T6	10.172	0.191	0.017	-0.050	2.635
		T7	10.173	0.201	0.017	-0.040	2.746

Table 3: Comparison of estimators of common mean when the population variance ratio ρ = 1 (n_1= 5, 15, 30, 50 and n_2=30)

n_1	n_2	Estimator	$\hat{T}_i$	SE($\hat{T}_i$)	RB($\hat{T}_i$)	β_1	β_2
5	30	T1	10.162	0.119	0.016	-0.218	2.865
		T2	10.141	0.244	0.014	-0.015	3.299
		T3	10.143	0.215	0.014	-0.017	2.733
		T4	10.282	2.009	0.028	0.001	3.260
		T5	10.283	2.269	0.028	0.019	4.759
		T6	10.128	0.564	0.013	-0.003	3.392
		T7	10.131	0.502	0.013	0.000	3.107
15	30	T1	10.119	0.107	0.012	-0.221	3.009
		T2	10.108	0.117	0.011	-0.031	3.428
		T3	10.109	0.112	0.011	-0.092	3.163
		T4	10.156	0.164	0.016	-0.102	3.125
		T5	10.147	0.144	0.015	-0.105	3.249
		T6	10.104	0.134	0.010	0.001	3.852
		T7	10.107	0.123	0.011	-0.022	3.368
30	30	T1	10.137	0.067	0.014	-0.133	3.541
		T2	10.134	0.066	0.013	-0.114	3.533
		T3	10.136	0.067	0.014	-0.112	3.496
		T4	10.134	0.069	0.013	-0.153	3.610
		T5	10.135	0.067	0.013	-0.139	3.561
		T6	10.134	0.071	0.013	-0.145	3.600
		T7	10.137	0.074	0.014	-0.114	3.568
		T8	10.133	0.068	0.013	-0.141	3.520
		T9	10.140	0.073	0.014	-0.125	3.292
		T10	10.135	0.068	0.014	-0.098	3.460
		T11	10.128	0.068	0.013	-0.106	3.512
50	30	T1	10.112	0.053	0.011	-0.189	2.743
		T2	10.109	0.054	0.011	-0.178	2.719
		T3	10.109	0.054	0.011	-0.173	2.693
		T4	10.109	0.053	0.011	-0.185	2.735
		T5	10.106	0.053	0.011	-0.165	2.653
		T6	10.131	0.075	0.013	-0.119	2.979
		T7	10.132	0.078	0.013	-0.104	3.034

Table 4: Comparison of estimators of common mean when the population variance ratio ρ = 1 (n_1= 5, 15, 30, 50 and n_2=50)

n_1	n_2	Estimator	$\widehat{T}_i$	SE($\widehat{T}_i$)	RB($\widehat{T}_i$)	β_1	β_2
5	50	T1	10.148	0.089	0.015	0.659	4.024
		T2	10.119	0.236	0.012	0.254	2.814
		T3	10.120	0.214	0.012	0.250	2.625
		T4	10.459	7.359	0.046	0.064	3.321
		T5	10.469	9.033	0.047	-0.019	4.676
		T6	10.081	0.711	0.008	0.074	2.610
		T7	10.084	0.670	0.008	0.061	2.448
15	50	T1	10.141	0.068	0.014	0.056	4.230
		T2	10.131	0.080	0.013	0.075	3.951
		T3	10.132	0.078	0.013	0.066	3.947
		T4	10.237	0.420	0.024	0.050	3.402
		T5	10.233	0.401	0.023	0.064	3.644
		T6	10.109	0.156	0.011	0.124	3.459
		T7	10.112	0.144	0.011	0.125	3.470
30	50	T1	10.129	0.056	0.013	0.000	3.521
		T2	10.121	0.057	0.012	0.000	3.206
		T3	10.124	0.057	0.012	0.000	3.278
		T4	10.160	0.080	0.016	0.053	4.325
		T5	10.159	0.075	0.016	0.056	4.354
		T6	10.118	0.060	0.012	0.000	3.092
		T7	10.123	0.058	0.012	0.000	3.214
50	50	T1	10.116	0.042	0.012	-0.030	3.660
		T2	10.114	0.042	0.011	-0.059	3.439
		T3	10.115	0.042	0.011	-0.049	3.463
		T4	10.118	0.043	0.012	0.000	4.233
		T5	10.117	0.042	0.012	-0.004	4.079
		T6	10.119	0.044	0.012	0.001	4.347
		T7	10.120	0.044	0.012	0.009	4.452
		T8	10.112	0.041	0.011	-0.030	3.636
		T9	10.116	0.046	0.012	0.021	4.224
		T10	10.114	0.042	0.011	-0.059	3.371
		T11	10.111	0.044	0.011	-0.014	3.426

Though, we presented only some combinations of (n_1, n_2) in the above tables due to space constraint. For each combination of (n_1, n_2), in both cases, the estimator with least SE value is identified as the most efficient estimator. The same is presented in the Table 5 below.

Table 5: Efficient estimators of common mean when the population variance ratio ρ = 1

n1\ n2	5	10	15	20	25	30	35	40	45	50
5	T_8	T_1	T_1	T_1	T_1	T_1	T_1	T_1	T_1	T_1
10	T_4	T_8	T_5	T_1	T_1	T_1	T_1	T_1	T_1	T_1
15	T_4	T_4	T_6	T_5	T_1	T_1	T_1	T_1	T_1	T_1
20	T_4	T_4	T_4	T_5	T_5	T_1	T_1	T_1	T_1	T_1
25	T_4	T_4	T_4	T_5	T_4	T_2	T_1	T_1	T_1	T_1
30	T_4	T_2	T_4	T_4	T_5	T_2	T_5	T_6	T_1	T_1
35	T_3	T_4	T_4	T_4	T_4	T_5	T_4	T_2	T_6	T_1
40	T_4	T_4	T_4	T_4	T_5	T_5	T_2	T_8	T_5	T_7
45	T_4	T_4	T_4	T_4	T_4	T_5	T_4	T_5	T_8	T_2
50	T_4	T_4	T_5	T_4	T_4	T_4	T_5	T_5	T_5	T_8

4. Final Conclusions

The following conclusions are drawn from the above tables

i) Larger the sample size smaller the SE of the estimator in all cases.
ii) As the size of the samples is increasing, the relative bias of estimators is decreasing.
iii) Since the relative bias of the estimators is very near to zero and it indicates that the estimators are almost unbiased as $n \rightarrow \infty$.
iv) In case of equal sample sizes, when it is ≤10 or ≥40 T_8 is formed to be the best estimator. When sample size is in between 10 and 40, no estimator has shown efficiency consistently.
v) When $n_1 < n_2$, in majority cases we found T_1 as the best estimator.
vi) Similarly, when $n_1 > n_2$, T_8 as the best estimator.
vii) Since the coefficient of skewness is near to zero and coefficient of kurtosis is near to three, which indicates that the distribution of the estimators is normal distribution.

5. References

1. Bhattacharya CG. Estimation of common mean and recovery of interblock information, Ann. Statist. 1980; 8:205-211.
2. Brown LD, Cohen A. Point and confidence estimation of a common mean and recovery of inter-block information. Ann. Statist. 1974; 2:963-976.
3. Cohen A, Sackrowitz HB. On estimating the common mean of two normal distributions, Ann. Statist. 1974; 2:1274-1282.
4. Graybill FA, Deal RB. Combining unbiased estimators, Biometrics. 1959; 15:543-550.
5. Khatri CG, Shah KR. Estimation of location parameters from two linear models under normality. Comm. Statist. 1974; 3:647-663.
6. Krishnamoorthy K, Rohatgi VK. Unbiased estimation of the common mean of two or more normal populations, Technical Report, Temple University, 1988.
7. Zacks S. Unbiased estimation of the common mean of two normal distributions based on samples of equal size, J. Amer. Statist. Assoc. 1966; 61:467-476.