



The Comparison of Confidence Intervals Under Six Major Sampling Methods

Liqian Situ

Background: A Little Knowledge of Olive Oil

- + Data: Olive Oil, 572 observations and 10 variables
- + Olive Oil:
 - + A fat obtained from the olive tree;
 - + Used in cooking, cosmetics, pharmaceuticals, and soaps
 - + Italy is the second largest producer;
 - + Palmitic acid: up to 20%
 - + Palmitoleic acid: 0.3% - 3.5%



Variables and Sample Size

- + Variables:
 - + Y – Palmitoleic Acid
 - + X – Palmitic Acid
- + Determination of Sample Size:
 - + Marginal of error: 2
 - + Variance: 169 (Guessed)
 - + Confidence level: 95%
 - + Sample size: 162

Sampling Method I: Simple Random Sample

- + Simple random sample 162 observations

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	126.14	13.21	[119.02,133.27]

Sampling Method IIa: Stratified Sample with Proportional & Neyman Allcotions

+ 9 Areas = 9 Stratum

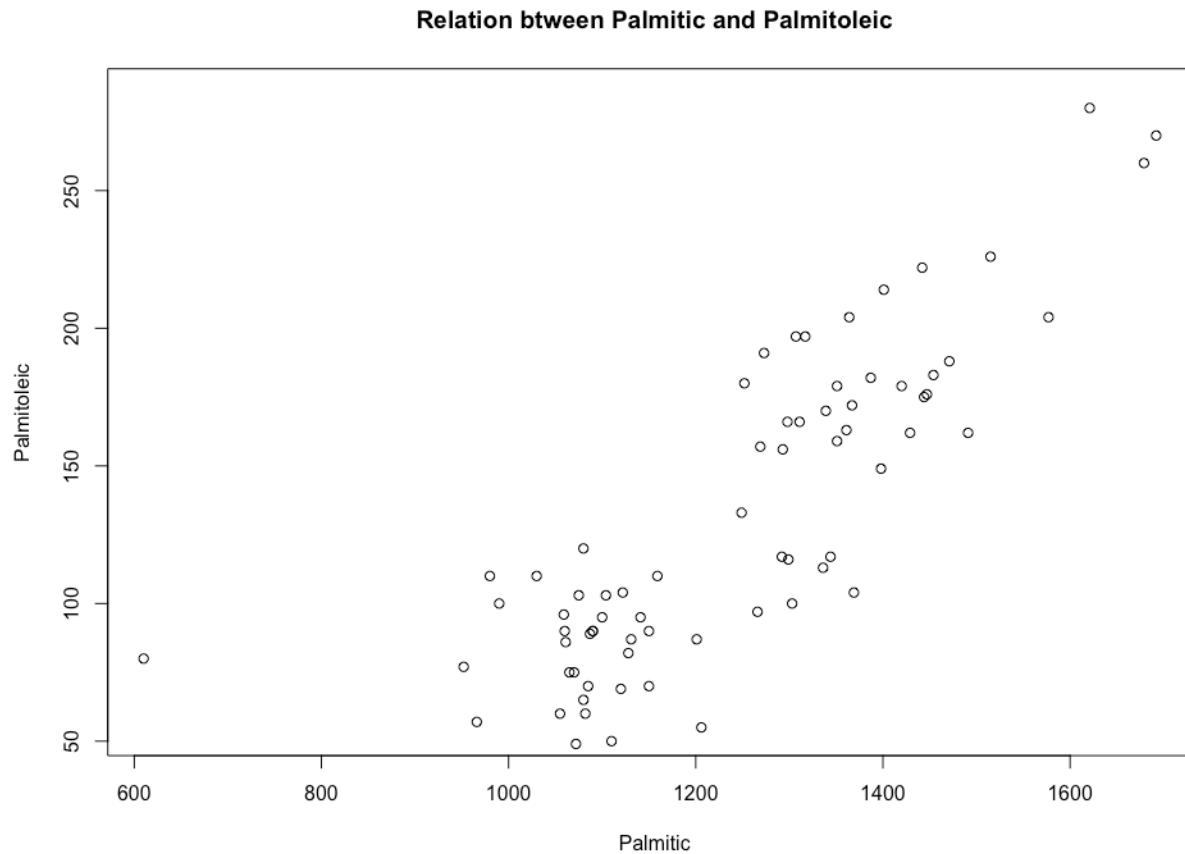
Area	Population	Prop. Allocation	Neyman Allocation
1	25	7	5
2	56	16	14
3	206	58	73
4	36	10	21
5	65	18	12
6	33	9	4
7	50	14	8
8	50	15	15
9	51	15	10
Total	572	162	162

Sampling Method lib: Stratified Sample with Neyman Allocation

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	126.14	13.21	[119.02,133.27]
Stratified Sample: Proportional Allocation	125.31	3.27	[121.77,128.86]
Stratified Sample: Neyman Allocation	127.44	2.32	[124.45,130.43]

Sampling Method III: Ratio Estimates

- + Auxiliary variable: Palmitic; Response variable: Palmitoleic



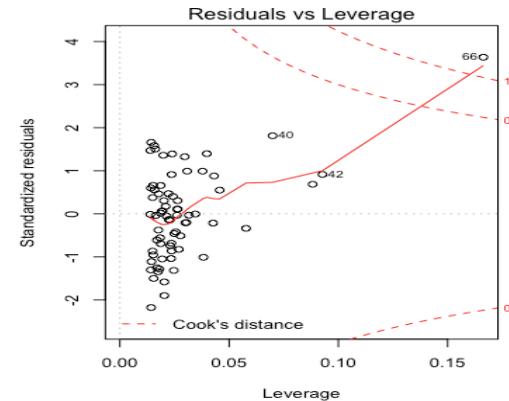
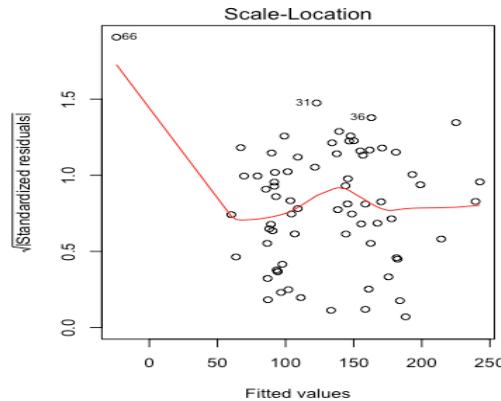
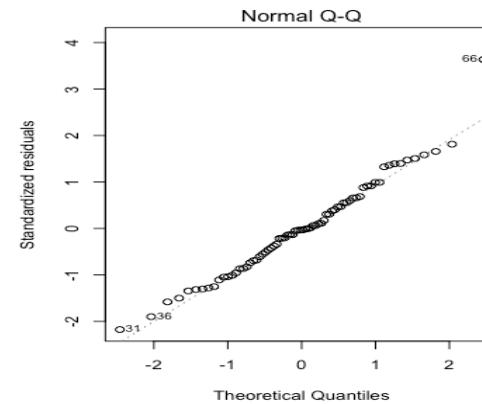
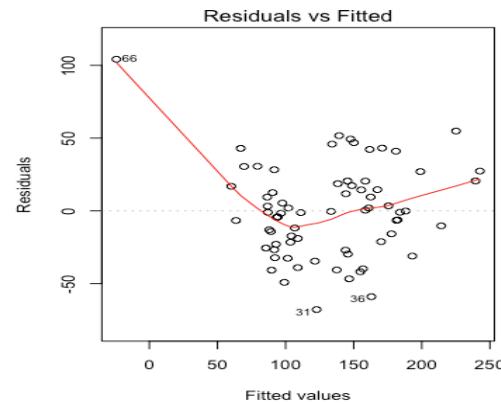
Sampling Method III: Ratio Estimates

- + Sample Correlation R=0.84
- + $CV(X)=0.14$ and $CV(Y)=0.04$
- + $R < 0.5 * CV(X)/CV(Y)$, so we should be cautious of using Ratio

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	126.14	13.21	[119.02,133.27]
Stratified Sample: Proportional Allocation	125.31	3.27	[121.77,128.86]
Stratified Sample: Neyman Allocation	127.44	2.32	[124.45,130.43]
Ratio Estimates	126.07	7.30	[120.77,131.37]

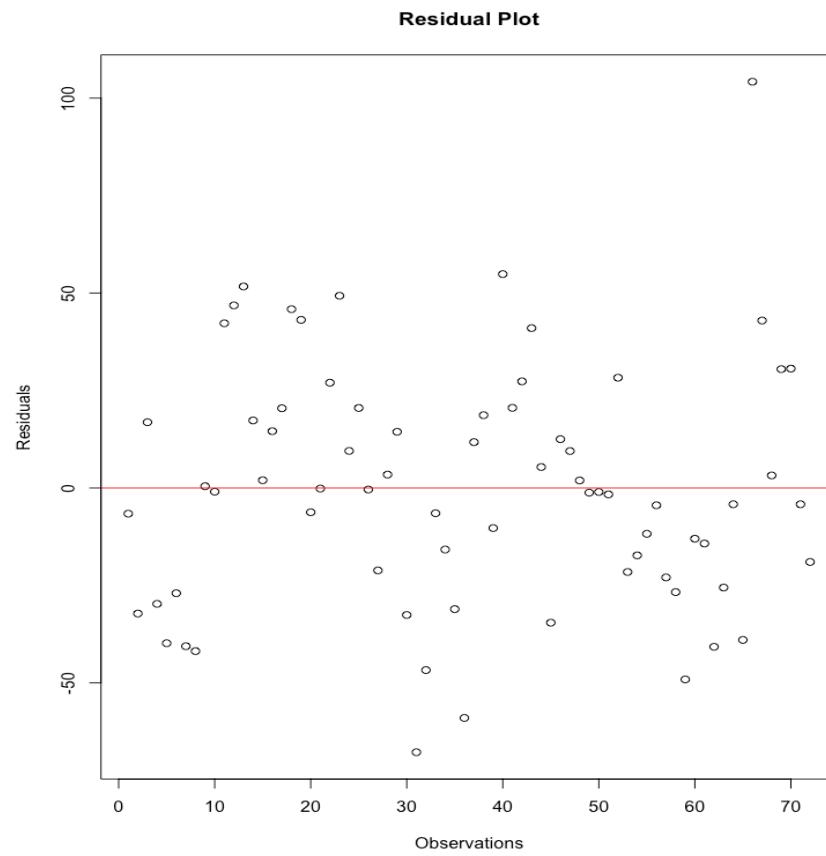
Sampling Method IV: Regression Estimates

- + Check the assumptions of normality and constant residual



Sampling Method IV: Regression Estimates

- + Also check the residual plot



Sampling Method IV: Regression Estimates

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	126.14	13.21	[119.02,133.27]
Stratified Sample: Proportional Allocation	125.31	3.27	[121.77,128.86]
Stratified Sample: Neyman Allocation	127.44	2.32	[124.45,130.43]
Ratio Estimates	126.07	7.30	[120.77,131.37]
Regression Estimates	125.96	3.84	[122.12,129.80]

Sampling Method V: Domains Estimates

- + We will only show the result for **Area 5** here:

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	126.14	13.21	[119.02,133.27]
Stratified Sample: Proportional Allocation	125.31	3.27	[121.77,128.86]
Stratified Sample: Neyman Allocation	127.44	2.32	[124.45,130.43]
Ratio Estimates	126.07	7.30	[120.77,131.37]
Regression Estimates	125.96	3.84	[122.12,129.80]
Domains Estimates	98.87	135.25	[76.07,121.66]

Sampling Method VI: Cluster Estimates

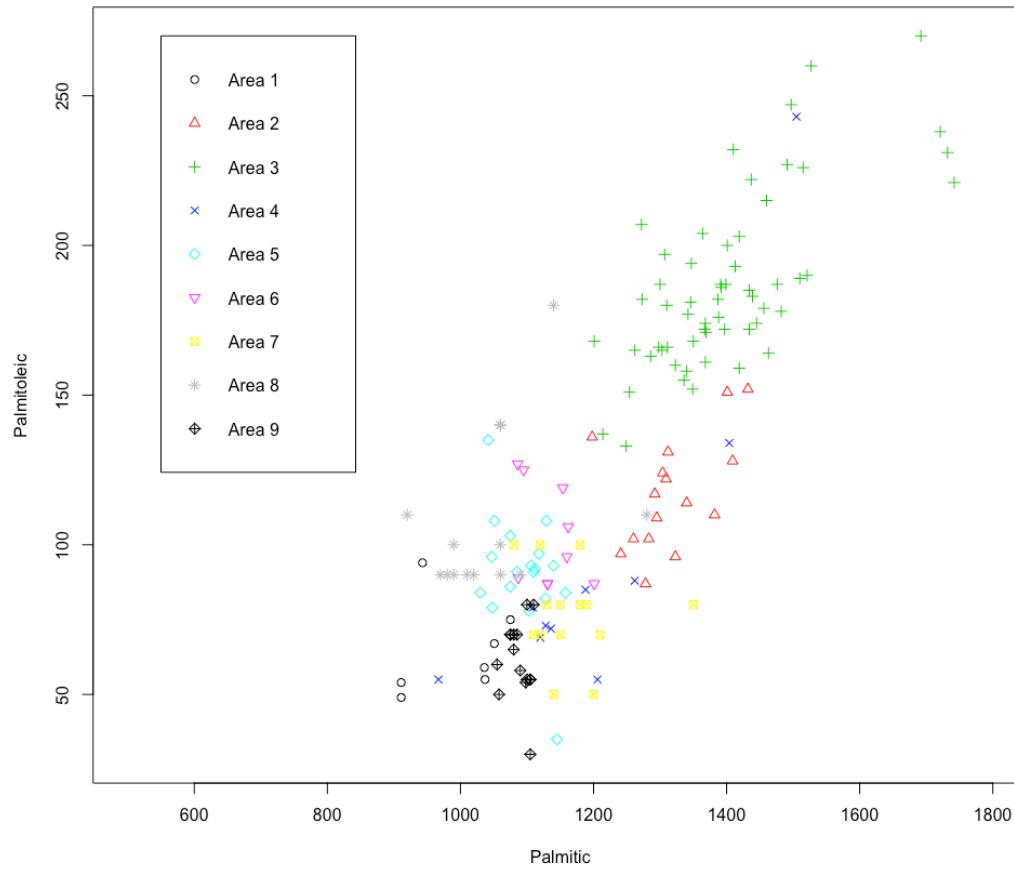
- + The clusters were chosen randomly are: Area 1, 2, 5 and 9

Area	1	2	5	9
Cluster Total	1540	6796	6146	3054
Cluster Size	25	56	65	51

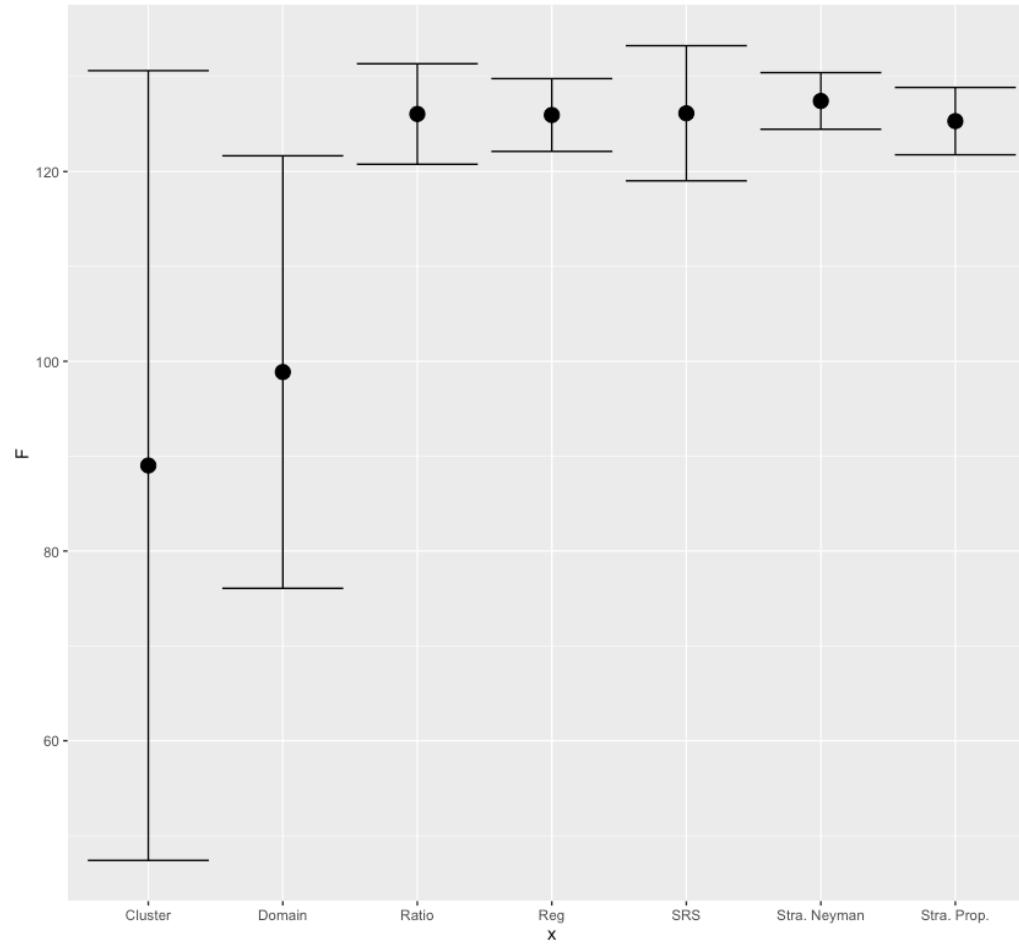
Sampling Method VI: Cluster Estimates

Method	Mean of Palmitoleic	Variance of Mean Palmitoleic	95% CI
Simple Random Sample	131.04	33.47	[119.70,142.38]
Stratified Sample: Proportional Allocation	129.03	8.25	[123.40,134.65]
Stratified Sample: Neyman Allocation	124.70	6.90	[123.88,134.18]
Ratio Estimates	130.23	20.61	[121.33,139.13]
Regression Estimates	129.14	11.79	[122.41,135.87]
Domains Estimates	90.89	137.55	[67.90,113.88]
Cluster Estimates	89.02	450.90	[47.40,130.63]

Conclusions:



Conclusions:



Conclusions:

- + The estimated mean of palmitoleic from Stratified sample with Neyman allocation is closest to the population mean
- + The estimated mean from clustering method is very different, comparing to the population mean
- + Ratio sampling method is not appropriate because the correlation between palmitoleic & palmitic is less than $CV(X) / (2 * CV(Y))$
- + Before any sampling analysis, check the assumptions and criterion of the methods.

References:

- + Lohr, S. L. (1999). Sampling: Design and analysis. Pacific Grove, CA: Duxbury Press.
- + Olive oil. (2016, November 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 17:56, November 5, 2016, from https://en.wikipedia.org/w/index.php?title=Olive_oil&oldid=747992438