# 模擬研究與討論

## 估計式整理

整理上一章所推倒之估計式，以及過去文獻中所存在的估計式。藉由電腦模擬的方式，在不同模型的設定下，比較三種取後不放回以及兩種取後放回的估計方式之表現，並同時估計其標準差。以下為不同估計方法之整理：

|  |  |
| --- | --- |
| 取後放回的兩群落共同種 | |
| 本文方法 |  |
| 現有方法 |  |
| 取後不放回的兩群落共同種 | |
| 本文方法 |  |
|  |
| 現有方法 |  |

且在95%信賴區間樣本涵蓋率方面，由於、、屬於下界估計式，因此使用物種數的信賴區間通過假設 符合對數常態分佈的方法計算估計物種數的95%信賴區間 (Chiu et al., 2014)，並求得該樣本真值於信賴區間的比例，以作為評估估計式穩定性的指標之一。物種數之95%信賴區間以計算，代入其中。

而在 與 的模擬中，則是依照物種數的信賴區間通過假設 符合常態分佈，計算計算估計物種數的95%信賴區間。物種數之95%信賴區間為：。

樣本涵蓋率 (Sample Coverage) 為樣本中檢測到的物種的總發生率的百分比。於出現型資料的樣本中，樣本涵蓋率被定義為樣本中已檢測物種，其物種相對出現率的總和， (Chao et.al, 1992)。

而在群落物種出現率組成之變異係數 () 估計中，為了估計在有限的樣本中，觀測到的物種數量之變異，並評估該群落之變異程度 (Chao et.al, 1992)。其估計式為： 。為變異係數平方的估計式，即。

## 模型模擬設定

### 模擬模型假設

經由電腦模擬，可以設定不同母體並產生不同物種豐富度的結構，透過重複抽樣資料以評估估計式之估計表現。在模擬的過程中，首先需決定兩群落的共同種與特有種的數量，再分別根據群落物種出現率組成設定不同分布模型，以下為四種模擬模型之設定：

1. 同質性模型 (homogeneity model)， ，()
2. 均勻模型 (uniform model)， ， ()，其中服從均勻分佈，為調整常數。
3. Broken-stick 模型， ， ()，其中服從指數函數分佈，為調整常數。
4. 對數常態模型 (log-normal model)，， () ，其中服從對數常態分佈，為調整常數。

### 物種與區塊數的假設

在電腦模擬時，需針對真實的群落物種與區塊數進行假設，以評估估計式~~結果~~之優劣程度。假設群落可劃分為T的區塊，並根據物種的出現率決定每個物種存在的區塊數，且假設不同物種間的空間分布是相互獨立。根據以下四種組合做為電腦模擬的群落物種空間分佈的假設母體。

情境一：總物種數為500種物種 (500)，其中群落一與第二群落皆存在400 種物種 (400)，兩群落間共有300種共同種(300)。且兩群 落皆可劃分為100個區塊 (100)，並從中依比例抽取 與 個區塊作為樣本使用。

情境二：總物種數皆為700種物種 (700)，其中群落一與第二群落分別存在 400與600種物種 (400，600)，兩群落間共有300種共同種 (300)。兩群落皆可劃分為100個區塊 (100)，並從中依比例 抽取與個區塊作為樣本使用，且。

情境三：假設總物種數為500種物種 (500)，其中群落一與第二群落皆存在 400種物種 (400)，兩群落間共有300種共同種(300)；兩族 群皆為100區塊 (100)，並從中依1:2比例抽取與個區塊作為 樣本使用，而 。

情境四：總物種數為500種物種 (500)，其中群落一與第二群落皆存在400 種物種 (400)，兩群落間共有300種共同種(300)。兩群落 皆可分別為100與200區塊 (100，200)，並從中依比例抽取 與 個區塊作為樣本使用。

根據每種情境假設，依不同樣本數進行1000次的模擬。以計算估計式的偏誤、樣本變異數、估計式變異數平均數，估計式之均方根誤差和95% 信賴區間涵蓋率。

## 表格中名詞定義

* q：為抽樣比例，。
* Sample size：抽樣區塊數，t。
* Estimator：所採用之估計方法。
* Obs：兩樣本觀測到的共同物種。
* AVG Estimate：次模擬之平均估計值，。
* Bias：次模擬之平均偏誤，。
* Sample SE：次模擬之估計式樣本標準差。
* Est. SD：次模擬之標準差估計值之平均。
* RMSE：次模擬估計量之估計式之均方根誤差。
* 95% CI Coverage：次模擬中真值於信賴區間的比例。

## 取後放回的模擬結果

首先針對樣本涵蓋率進行計算，在多種組合模擬下，在樣本大小為10個區塊時，各模型的樣本涵蓋率 () 介於0.6至0.84之間；而當樣本大小大於10個區塊時，任何模型的樣本涵蓋率 () 皆大於0.9。

並且分別使用本文所提出的方法 與現有方法 估計兩群落之間的共同物種數。依據模擬結果可以得知 (表 1.1-表 1.4)，在情境一的物種與區塊的設定下 ， 相較於的偏差結果方面更接近真值，尤其在同質與Broken-stick群落組合下具有更好的結果。且無論何種模型組合下，估計皆有低估的現象發生。而在樣本標準差 (Sample SE) 以及估計標準差 (Est. SD) 的結果中，在所有模型組合下，本文所提出的 與現有方法 皆呈現低估的現象。並且在樣本大小為10個區塊時，低估情形更加明顯。

然而，單憑偏誤與標準差的估計的結果無法從中判定估計式的好壞，在評估一個估計方法的優劣之時，還需要考慮該估計量的準確度以及精確度，因此將以均方根誤差 (root-mean-square error，RMSE) 以及95% 信賴區間道真值含太率 (95% confidence interval Coverage，95% CI Coverage) 作為衡量估計式好壞之準則。在RMSE的結果方面，在同質與Broken-stick的模型組合下，本文所提出的 與現有方法 的結果並無明顯差異之外，在其餘的模型組合下本文所提出的 皆有較小的RMSE。而在95% CI Coverage的結果中，則是本文所提出的 皆有更精確的樣本涵蓋率。

在情境二、情境三與多數情境四的物種區塊數的假設下所進行的模擬結果 (表S.1 - 表 S.8，表S.10 - 表 S.12)，在所有估計結果方面，皆與情境一物種與區塊數的假設並無太大的差異。但在情境四下以同質與Broken-stick模型組合下 (表S.9)，使用本文所提出的 進行估計時，在小樣本的情況下雖然偏誤結果相較現有方法 更接近真值，卻容易發生高估的現象，此外在RMSE的評估指標上， 會有較大的RMSE，但具有更精確的95% CI Coverage。

表 1.1、取後放回的抽樣方式在情境一下，群落一為同質模型，群落二為Broken-stick模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 10 | 12.MoRE | 125.12 | 270.79 | -29.21 | 61.1 | 46.64 | 67.69 | 0.86 |
| Pan | 246.81 | -53.19 | 40.24 | 27.93 | 66.68 | 0.82 |
| 30 | 12.MoRE | 240.32 | 287.18 | -12.82 | 19.64 | 18.34 | 23.44 | 0.91 |
| Pan | 277.62 | -22.38 | 13.64 | 11.92 | 26.2 | 0.89 |
| 50 | 12.MoRE | 267.3 | 294.68 | -5.32 | 16.04 | 14.36 | 16.89 | 0.87 |
| Pan | 288.29 | -11.71 | 11.15 | 9.45 | 16.16 | 0.86 |
| 70 | 12.MoRE | 277.28 | 297.93 | -2.07 | 14.03 | 12.42 | 14.18 | 0.83 |
| Pan | 292.66 | -7.34 | 9.78 | 8.19 | 12.22 | 0.82 |

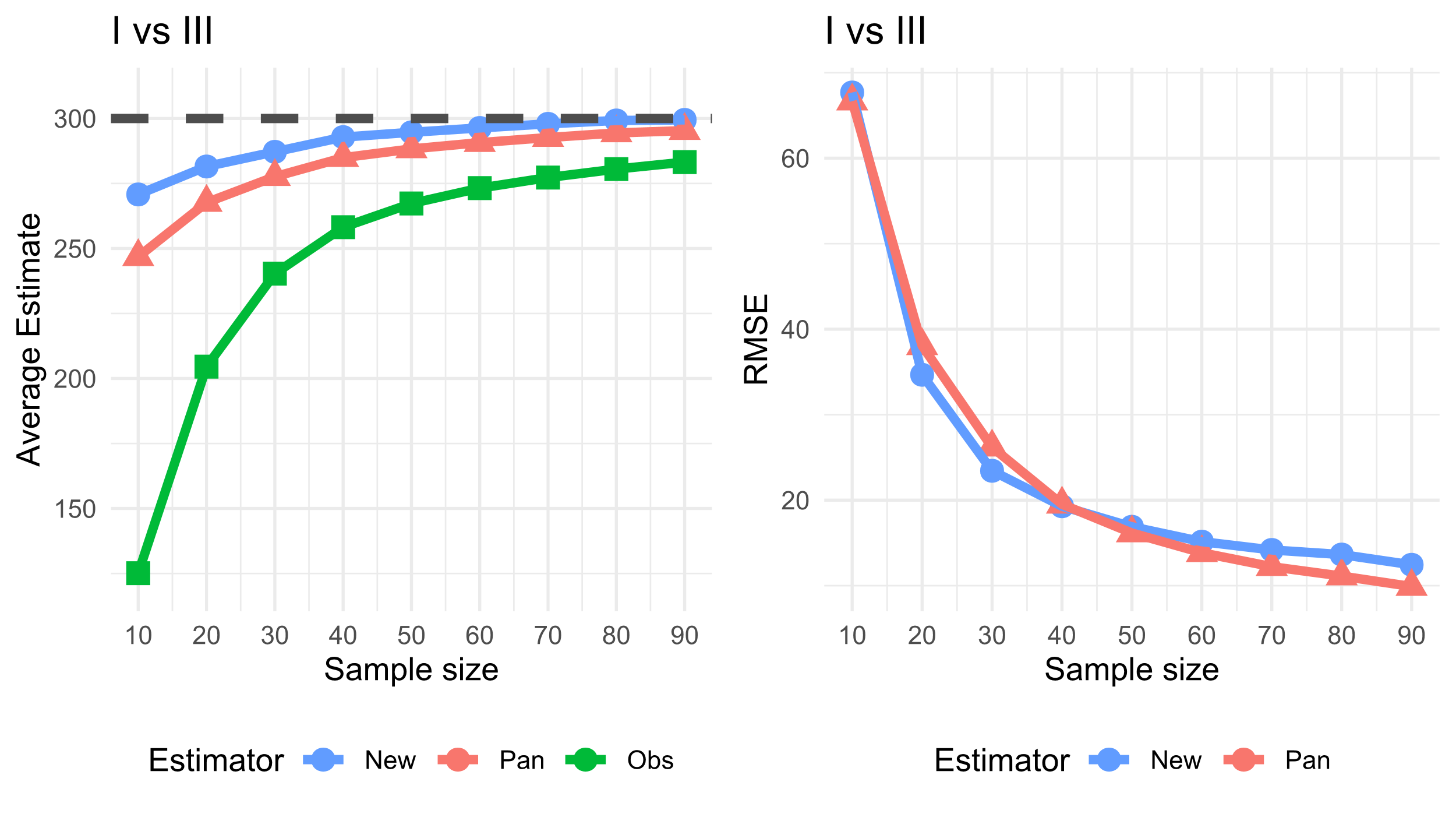


圖 1.1取後放回的抽樣方式在情境一下， 群落一為同質模型，群落二為Broken-stick模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

表 1.2、取後放回的抽樣方式在情境一下，兩群落皆為均勻模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample size | Estimator | Obs. | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 10 | 12.MoRE | 123.34 | 244.33 | -55.67 | 52.13 | 40.13 | 76.25 | 0.86 |
| Pan | 220.02 | -79.98 | 32.53 | 22.51 | 86.34 | 0.83 |
| 30 | 12.MoRE | 225.85 | 277.64 | -22.36 | 22.77 | 19.67 | 31.9 | 0.89 |
| Pan | 265.65 | -34.35 | 15.31 | 12.33 | 37.6 | 0.88 |
| 50 | 12.MoRE | 255.45 | 289.5 | -10.5 | 16.01 | 15.42 | 19.14 | 0.9 |
| Pan | 281.66 | -18.34 | 11.37 | 10.15 | 21.58 | 0.89 |
| 70 | 12.MoRE | 268.74 | 295.39 | -4.61 | 13.68 | 13.72 | 14.43 | 0.91 |
| Pan | 289.14 | -10.86 | 9.76 | 9.21 | 14.6 | 0.9 |

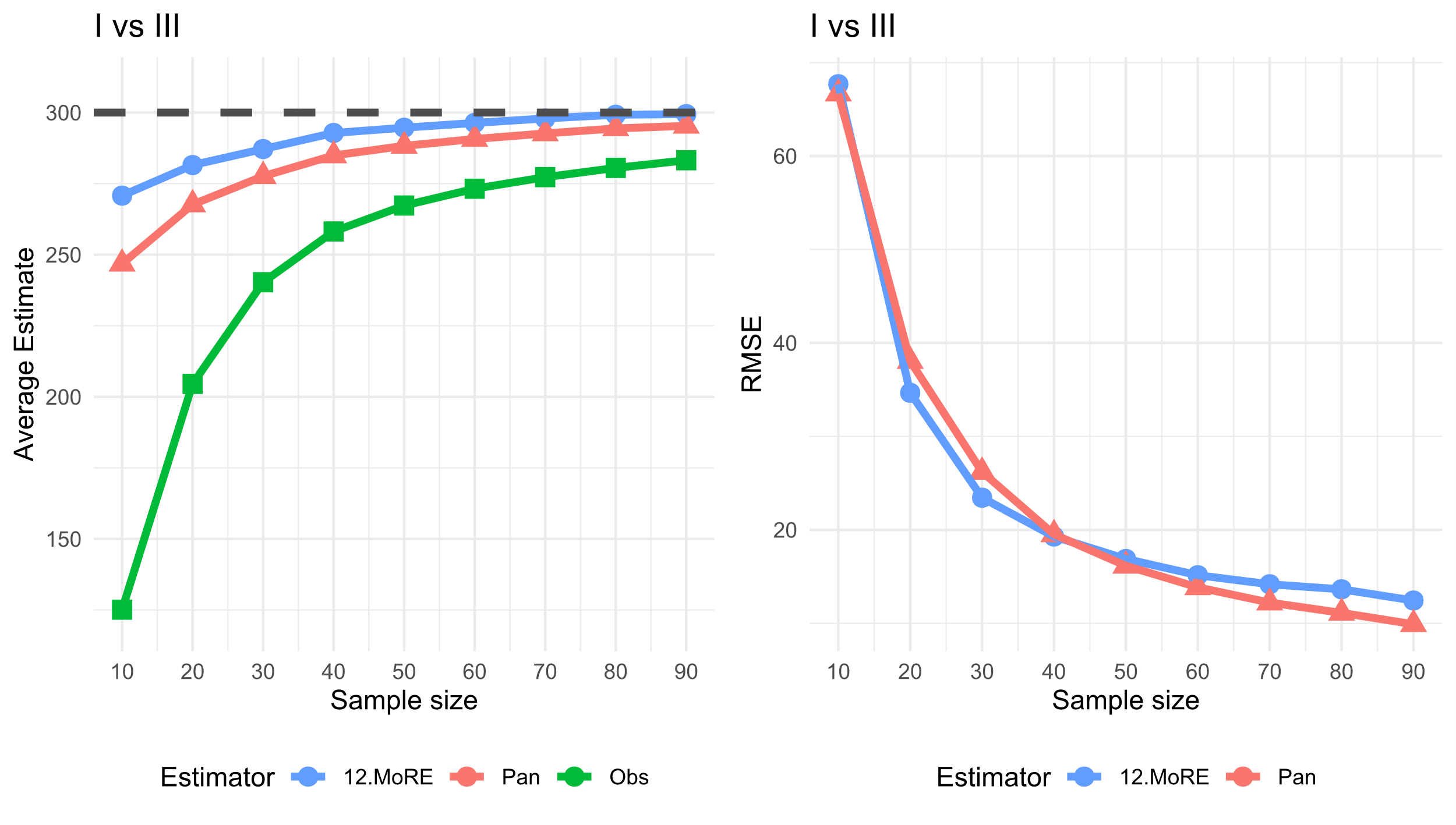


圖 1.2、取後放回的抽樣方式在情境一下，兩群落皆為均勻模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

表 1.3、取後放回的抽樣方式在情境一下，群落一為均勻模型，群落二為Broken-stick模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 10 | 12.MoRE | 123.27 | 232.31 | -67.69 | 47.01 | 37.17 | 82.4 | 0.85 |
| Pan | 207.78 | -92.22 | 29.49 | 20.99 | 96.82 | 0.83 |
| 30 | 12.MoRE | 218.54 | 276.15 | -23.85 | 24.65 | 21.91 | 34.29 | 0.89 |
| Pan | 262.17 | -37.83 | 16.27 | 13.71 | 41.18 | 0.88 |
| 50 | 12.MoRE | 248.57 | 290.11 | -9.89 | 19.9 | 17.98 | 22.21 | 0.88 |
| Pan | 280.02 | -19.98 | 13.76 | 11.77 | 24.26 | 0.88 |
| 70 | 12.MoRE | 263.41 | 295.6 | -4.4 | 17.15 | 15.73 | 17.69 | 0.88 |
| Pan | 287.96 | -12.04 | 12.17 | 10.46 | 17.11 | 0.86 |

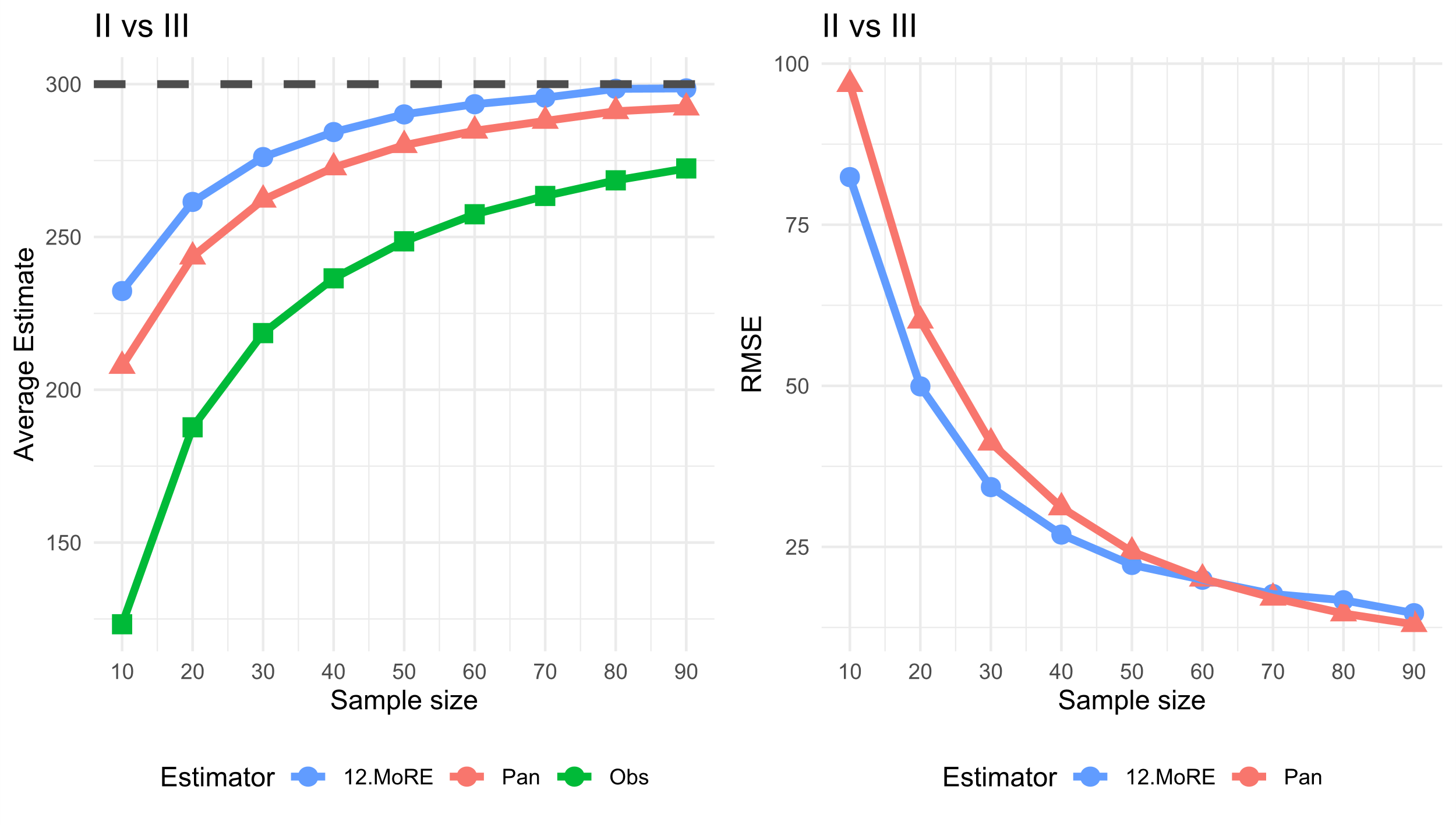


圖 1.3、取後放回的抽樣方式在情境一下，群落一為均勻模型，群落二為Broken-stick模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

表 1.4、取後放回的抽樣方式在情境一下，群落一為Broken-stick模型，群落二為對數常數模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 10 | 12.MoRE | 115.62 | 228.28 | -71.72 | 52.12 | 40.34 | 88.64 | 0.85 |
| Pan | 199.49 | -100.51 | 31.57 | 22.73 | 105.34 | 0.83 |
| 30 | 12.MoRE | 208.96 | 277.73 | -22.27 | 29.47 | 25.47 | 36.93 | 0.9 |
| Pan | 261.16 | -38.84 | 18.92 | 15.57 | 43.2 | 0.88 |
| 50 | 12.MoRE | 242.64 | 292.56 | -7.44 | 22.86 | 20.58 | 24.03 | 0.88 |
| Pan | 280.66 | -19.34 | 15.44 | 13.18 | 24.74 | 0.87 |
| 70 | 12.MoRE | 260.08 | 298.13 | -1.87 | 19.15 | 17.44 | 19.23 | 0.88 |
| Pan | 289.36 | -10.64 | 13.26 | 11.47 | 16.99 | 0.88 |

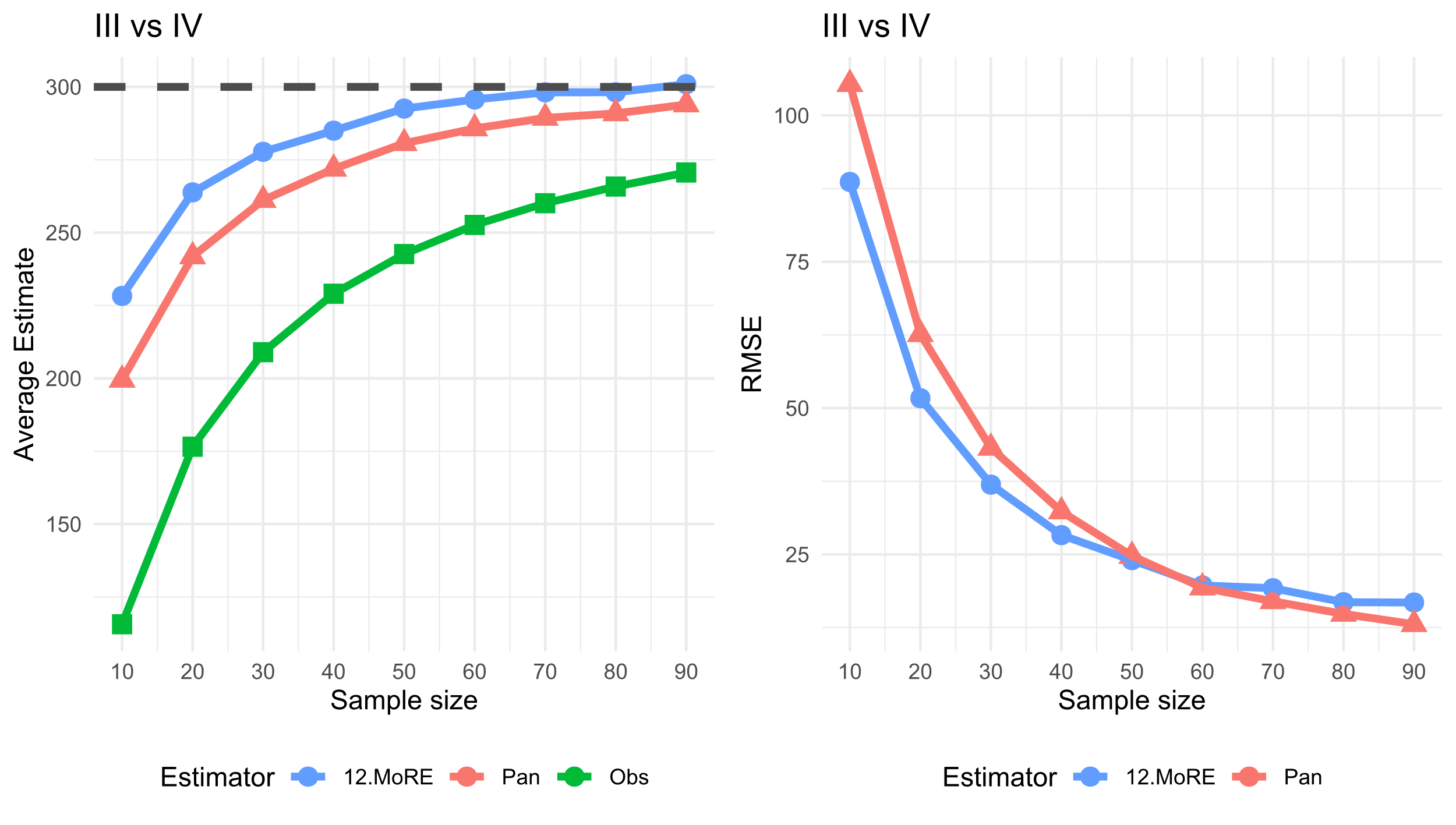


圖 1.4、取後放回的抽樣方式在情境一下，群落一為Broken-stick模型，群落二為對數常數模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

## 取後不放回的模擬結果

在取後不放回的模擬針對樣本涵蓋率型計算，同樣於在多種組合模擬下，除了在抽樣比例為0.1的情境底下，其餘樣本大小所抽樣本，無論是在何種模型或模擬組合下，所估計之樣本涵蓋率 () 皆大於0.9。且模擬同質群落或均勻群落在抽樣比例為0.1的情境下，樣本所估計的群落變異係 () 數通常小於0.4。而隨抽樣比例的增加將越接近母體變異係數。

依據情境一物種假設的模擬結果可以得知 (表 4.5 - 表4.8)，在偏誤的表現上， 使用本文所提出的方法 與 估計共同物種，在小樣本 () 的估計中，同質與Broken-stick以及兩群落皆為均勻的兩種模擬組合下有高估的表現；而在均勻與Broken-stick以及Broken-stick與常對數的兩種模擬組合下則出現低估的表現。但在大樣本的情境下，所有群落組合則是三種估計方法皆呈現低估的表現。

在樣本標準差以及估計標準差的結果中，可以發現無論在何種模擬情形下，無論是在本文所提出的、，或是現有方法所呈現的估計標準差結果皆有低估樣本標準差的情境發生。除此之外，在RMSE的結果中，使用本文所提出的估計方法 與 所估計的結果，兩種方式間並無明顯差異，且都小於 。 而在95% CI Coverage方面，由於 已知為下界估計式，因此在95% CI Coverage的估計上使用log修正的方式進行估計，而在 與 ，則使用未修正的估計方式。並且無論在何種模型組合的結果下，的估計結果具有較精確的95% CI Coverage。

在情境二、情境三與情境四的物種與區塊數的假設底下，所進行的模擬結果，無論是在何種估計值的結果，皆與情境一的物種與區塊數的假設並無太大的差異，皆是在使用本文的估計方法 的結果種具有更接近真值的偏誤、較小的RMSE與較精確的% CI Coverage，且估計標準差皆有低估的現象發生 (表S. 3 - 表S.24)。

表 1.5、取後不放回的抽樣方式在情境一下，群落一為同質模型，群落二為Broken-stick模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| q | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95%CI Coverage |
| 0.1 | 12.wMoRE1 | 130.68 | 363.12 | 63.12 | 57.16 | 51.9 | 85.14 | 0.93 |
| 12.wMoRE2 | 367.93 | 67.93 | 67.03 | 56.69 | 95.41 | 0.92 |
| wChao2.12 | 246.53 | -53.47 | 30.75 | 24.33 | 61.67 | 0.89 |
| 12.MoRE | 282.03 | -17.97 | 57.47 | 45.49 | 60.19 | 0.89 |
| 0.3 | 12.wMoRE1 | 251.5 | 311.16 | 11.16 | 12.8 | 9.77 | 16.98 | 0.88 |
| 12.wMoRE2 | 313.95 | 13.95 | 16.78 | 15 | 21.81 | 0.93 |
| wChao2.12 | 281.55 | -18.45 | 9.7 | 9.15 | 20.85 | 0.92 |
| 12.MoRE | 300.26 | 0.26 | 19.47 | 19.11 | 19.46 | 0.91 |
| 0.5 | 12.wMoRE1 | 277.66 | 299.44 | -0.56 | 6.72 | 7.5 | 6.74 | 0.98 |
| 12.wMoRE2 | 297.72 | -2.28 | 6.06 | 5.89 | 6.47 | 0.94 |
| wChao2.12 | 292.87 | -7.13 | 6.15 | 5.81 | 9.41 | 0.89 |
| 12.MoRE | 313.19 | 13.19 | 16.35 | 16.84 | 21 | 0.91 |
| 0.7 | 12.wMoRE1 | 289.19 | 299.95 | -0.05 | 3.67 | 4.06 | 3.67 | 0.97 |
| 12.wMoRE2 | 298.72 | -1.28 | 3.32 | 3.51 | 3.56 | 0.97 |
| wChao2.12 | 298.19 | -1.81 | 3.62 | 3.74 | 4.05 | 0.9 |
| 12.MoRE | 324.51 | 24.51 | 14.65 | 17.23 | 28.55 | 0.94 |

表 1.6、取後不放回的抽樣方式在情境一下，兩群落皆為均勻模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| q | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95%CI Coverage |
| 0.1 | 12.wMoRE1 | 127.72 | 306.18 | 6.18 | 44.17 | 44.69 | 44.58 | 0.95 |
| 12.wMoRE2 | 310.94 | 10.94 | 51.89 | 43.22 | 53.01 | 0.92 |
| wChao2.12 | 215.98 | -84.02 | 23.97 | 19.41 | 87.37 | 0.87 |
| 12.MoRE | 246.21 | -53.79 | 46.24 | 38.37 | 70.92 | 0.89 |
| 0.3 | 12.wMoRE1 | 234.14 | 295.51 | -4.49 | 14.16 | 14.3 | 14.85 | 0.95 |
| 12.wMoRE2 | 302.6 | 2.6 | 19.56 | 17.19 | 19.72 | 0.93 |
| wChao2.12 | 269.61 | -30.39 | 11.25 | 9.92 | 32.4 | 0.89 |
| 12.MoRE | 294.95 | -5.05 | 24.75 | 21.59 | 25.25 | 0.9 |
| 0.5 | 12.wMoRE1 | 266.2 | 297.26 | -2.74 | 8.56 | 8.55 | 8.98 | 0.94 |
| 12.wMoRE2 | 300.7 | 0.7 | 9.85 | 9.02 | 9.87 | 0.92 |
| wChao2.12 | 287.64 | -12.36 | 7.82 | 6.8 | 14.62 | 0.89 |
| 12.MoRE | 315.69 | 15.69 | 20.48 | 19.29 | 25.79 | 0.9 |
| 0.7 | 12.wMoRE1 | 282.94 | 299.07 | -0.93 | 5.17 | 4.97 | 5.25 | 0.94 |
| 12.wMoRE2 | 299.6 | -0.4 | 5.23 | 4.84 | 5.24 | 0.93 |
| wChao2.12 | 296.33 | -3.67 | 5.08 | 4.55 | 6.26 | 0.89 |
| 12.MoRE | 337.83 | 37.83 | 20.16 | 21.22 | 42.87 | 0.94 |

表 1.7、取後不放回的抽樣方式在情境一下，群落一為均勻模型，群落二為Broken-stick模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| q | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95%CI Coverage |
| 0.1 | 12.wMoRE1 | 126.47 | 290.73 | -9.27 | 43.62 | 38.75 | 44.58 | 0.92 |
| 12.wMoRE2 | 299.38 | -0.62 | 55.43 | 44.92 | 55.4 | 0.92 |
| wChao2.12 | 208.8 | -91.2 | 24.3 | 19.13 | 94.38 | 0.86 |
| 12.MoRE | 243.7 | -56.3 | 47.86 | 38.89 | 73.87 | 0.88 |
| 0.3 | 12.wMoRE1 | 226.97 | 293.64 | -6.36 | 15.81 | 14.89 | 17.03 | 0.94 |
| 12.wMoRE2 | 297.21 | -2.79 | 20.82 | 17.73 | 21 | 0.93 |
| wChao2.12 | 266.41 | -33.59 | 12.31 | 10.79 | 35.77 | 0.89 |
| 12.MoRE | 296.35 | -3.65 | 27.04 | 24.18 | 27.27 | 0.88 |
| 0.5 | 12.wMoRE1 | 261.52 | 297.11 | -2.89 | 8.34 | 9.2 | 8.82 | 0.97 |
| 12.wMoRE2 | 297.61 | -2.39 | 9 | 9.31 | 9.31 | 0.96 |
| wChao2.12 | 286.59 | -13.41 | 7.55 | 7.44 | 15.38 | 0.92 |
| 12.MoRE | 320.96 | 20.96 | 20.55 | 21.94 | 29.35 | 0.94 |
| 0.7 | 12.wMoRE1 | 281.06 | 299.6 | -0.4 | 5.32 | 5.33 | 5.34 | 0.94 |
| 12.wMoRE2 | 299.04 | -0.96 | 5.3 | 5.1 | 5.38 | 0.93 |
| wChao2.12 | 296.51 | -3.49 | 5.24 | 4.9 | 6.3 | 0.9 |
| 12.MoRE | 342.9 | 42.9 | 20.83 | 23.26 | 47.69 | 0.95 |

表 1.8、取後不放回的抽樣方式在情境一下，群落一為Broken-stick模型，群落二為對數常數模型之情境下的估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| q | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95%CI Coverage |
| 0.1 | 12.wMoRE1 | 121.47 | 282.46 | -17.54 | 44.72 | 41.29 | 48.02 | 0.94 |
| 12.wMoRE2 | 290.68 | -9.32 | 56.03 | 47.67 | 56.77 | 0.93 |
| wChao2.12 | 202.82 | -97.18 | 25.04 | 19.94 | 100.36 | 0.87 |
| 12.MoRE | 243.48 | -56.52 | 50.92 | 41.92 | 76.06 | 0.87 |
| 0.3 | 12.wMoRE1 | 223.77 | 300.8 | 0.8 | 17.43 | 19.64 | 17.44 | 0.98 |
| 12.wMoRE2 | 304.2 | 4.2 | 22.7 | 19.45 | 23.08 | 0.93 |
| wChao2.12 | 269.92 | -30.08 | 13.26 | 11.78 | 32.87 | 0.91 |
| 12.MoRE | 304.82 | 4.82 | 31.38 | 27.43 | 31.74 | 0.91 |
| 0.5 | 12.wMoRE1 | 262.89 | 301.49 | 1.49 | 8.75 | 9.85 | 8.87 | 0.97 |
| 12.wMoRE2 | 303.69 | 3.69 | 10.26 | 9.95 | 10.9 | 0.94 |
| wChao2.12 | 289.26 | -10.74 | 8.02 | 7.55 | 13.4 | 0.91 |
| 12.MoRE | 321.77 | 21.77 | 22.08 | 21.99 | 31 | 0.94 |
| 0.7 | 12.wMoRE1 | 283.16 | 300.77 | 0.77 | 4.73 | 5.18 | 4.79 | 0.96 |
| 12.wMoRE2 | 301.25 | 1.25 | 4.89 | 5.12 | 5.04 | 0.96 |
| wChao2.12 | 296.82 | -3.18 | 4.69 | 4.58 | 5.66 | 0.9 |
| 12.MoRE | 329.86 | 29.86 | 17.54 | 19.58 | 34.62 | 0.96 |

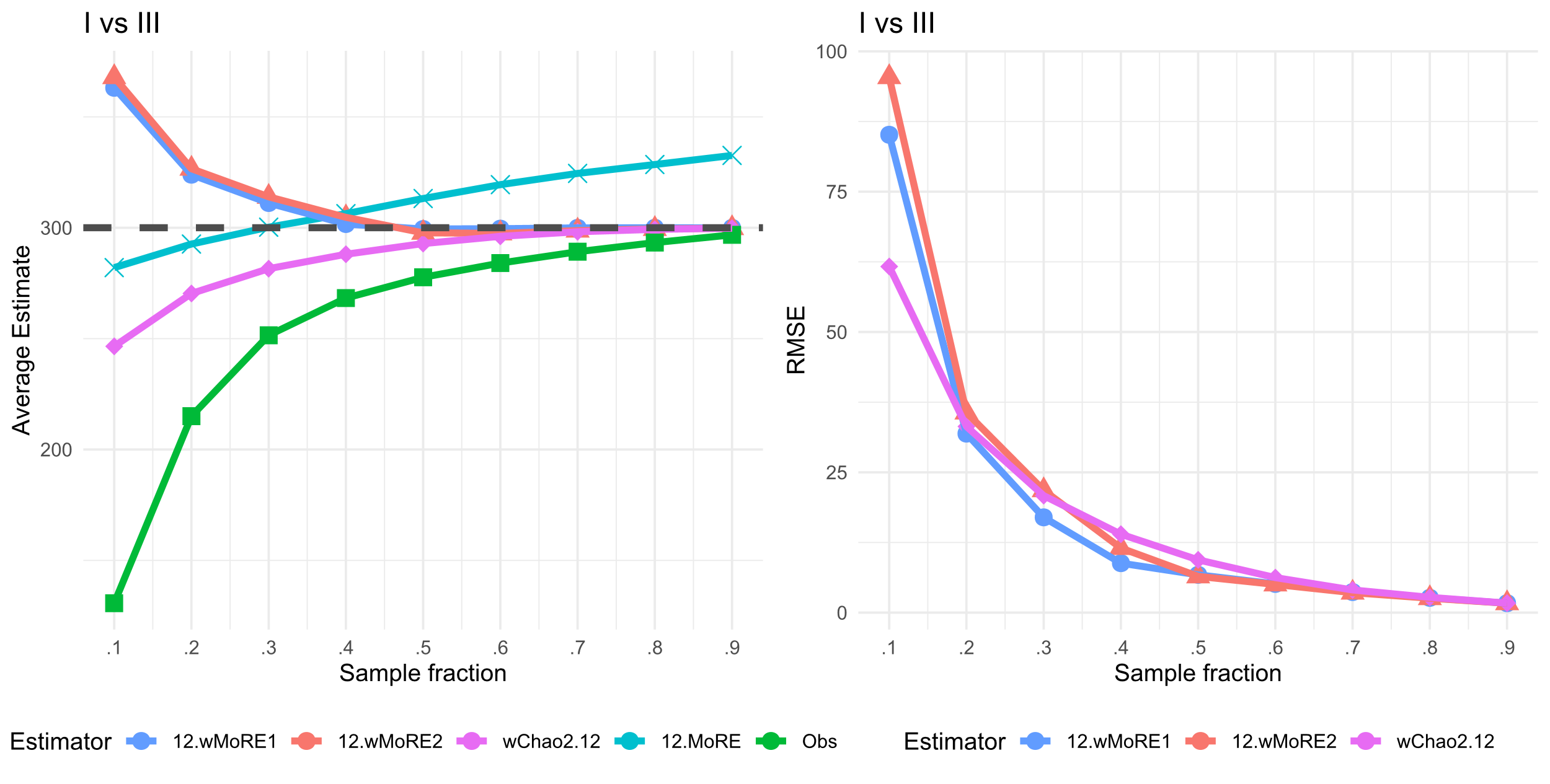


圖 1.5、取後不放回的抽樣方式在情境一下， 群落一為同質模型，群落二為Broken-stick模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

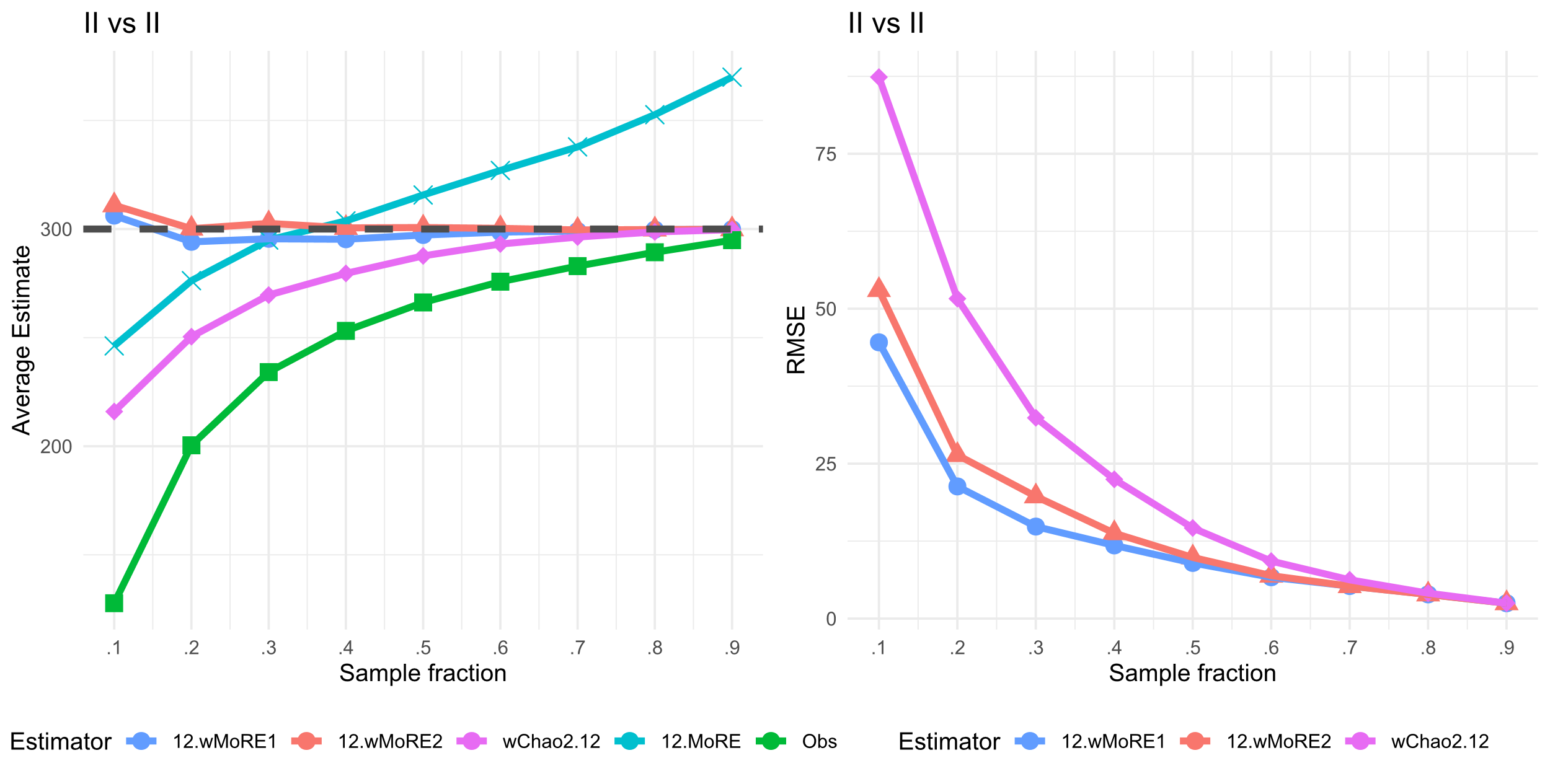


圖 1.6、取後不放回的抽樣方式在情境一下，兩群落皆為均勻模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

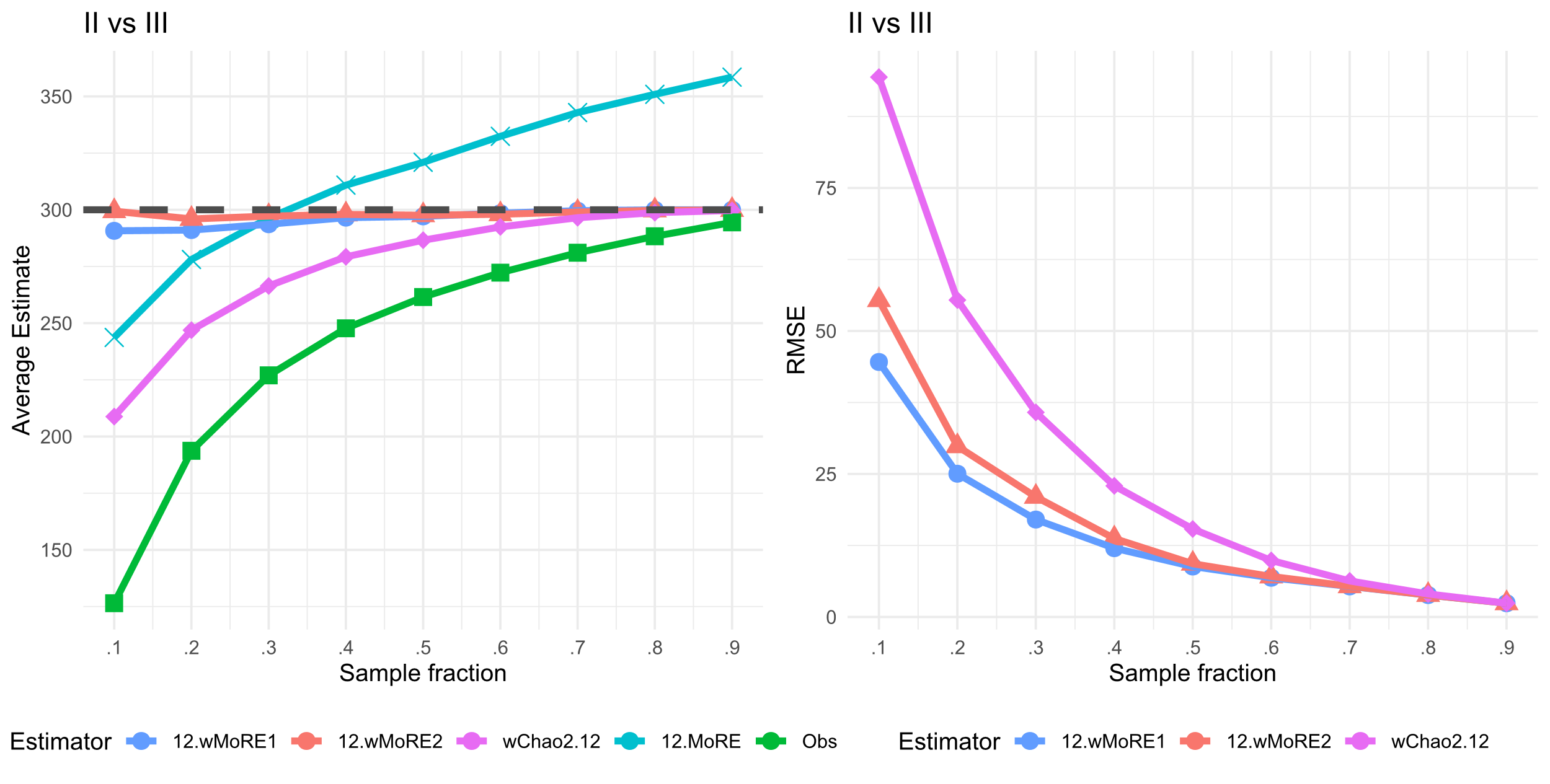


圖 1.7、取後不放回的抽樣方式在情境一下，群落一為均勻模型，群落二為Broken-stick模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

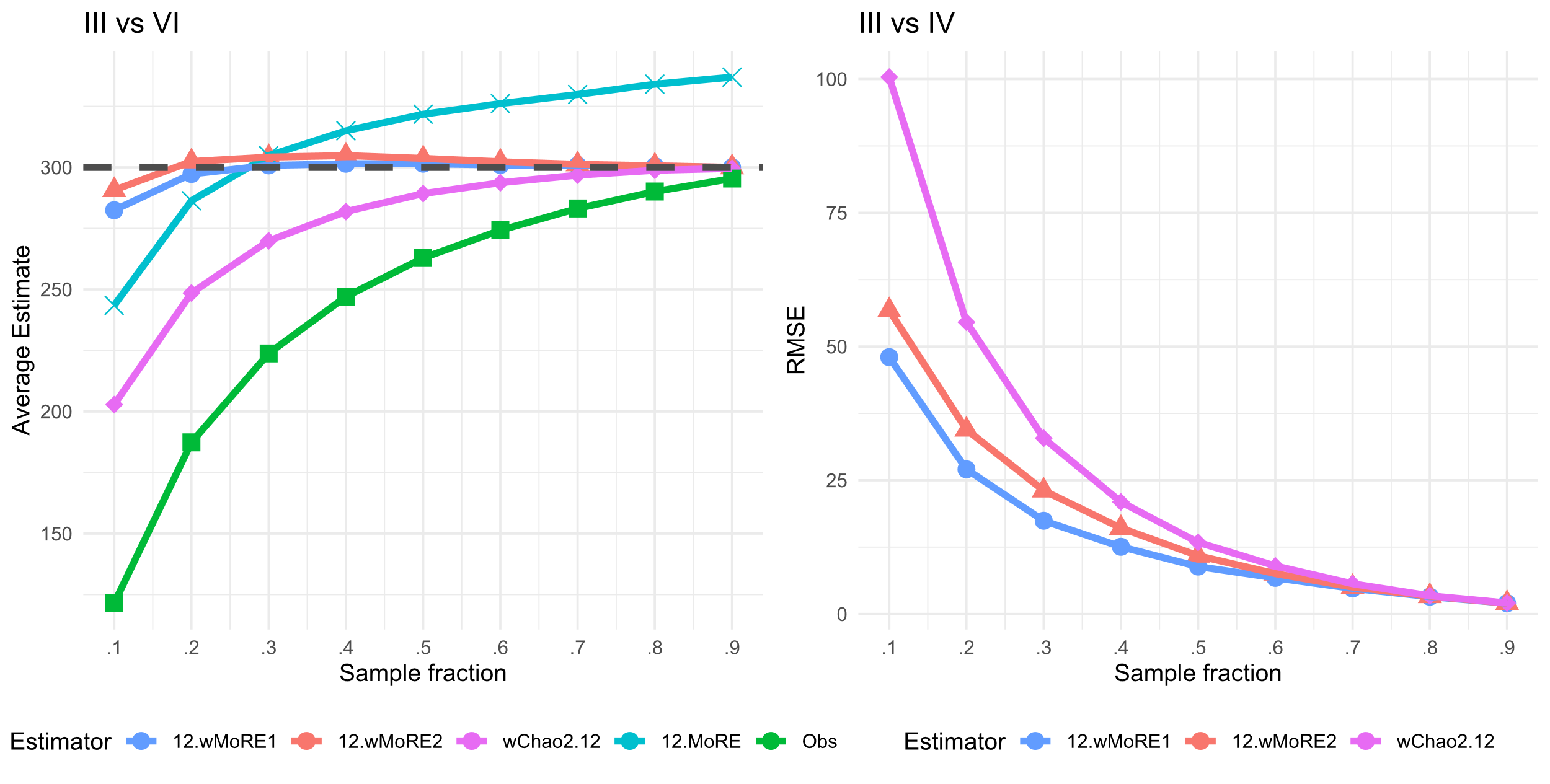


圖 1.8、取後不放回的抽樣方式在情境一下，群落一為Broken-stick模型，群落二為對數常數模型之情境下的AVG Estimate (左圖) 與RMSE (右圖) 估計結果。

## 真實資料模擬

除了使用模擬母體進行重複抽樣的模擬之外，亦使用兩筆真實數據做為母體進行1000次的重複抽樣，以評斷估計式的結果優劣。

### 澳洲三種極端氣候鳥類資料

**資料描述**

針對澳洲東南部半乾旱地區的尤加利樹林地，量化三個氣候時期178 個地點的極端氣候對鳥類物種出現 (species occurrence)、物種豐富度(species richness) 和出現型 (incidence) 的影響——聖嬰現象相關的乾旱 (Big Dry)、反聖嬰現象破壞性降雨 (Big Wet)，以及反聖嬰現象事件三年後 (Post-Big Wet)。並使用Big Dry 作為群落一使用 (Mean = 0.08, CV = 1.32)，其中共存在55種物種；Big Wet 作為群落二使用 (Mean = 0.16, CV = 1.39)，其中共存在65種物種。Post-Big Wet作為群落三使用 (Mean = 0.15, CV = 1.42)，其中共存在56種物種。

**模擬結果**

在該筆資料中使用取後放回的估計方法下，假設該筆資料為母體，並抽取其中的區塊做為樣本所使用。在偏誤結果方面，三個群落之間所估計的共同物種數，使用本文所提出的 的估計結果皆比現有的方法 更接近真值。而在估計標準差方面，與上一小節電腦膜你的結果相同，三群落間的結果皆呈現低估的現象。最後是在兩個模型評估指標上，在以該資料集作為樣本的情況下，的結果皆獲得更小的RMSE與更精確的95% 信賴區間的真值涵蓋率。然而，雖然在兩評估指標中 獲得較好的表現，但在結果上與 的結果沒有明顯的差異。

表 1.9、取後放回澳洲三種極端氣候鳥類資料集中的Big Dry與Big Wet估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 20 | 12.MoRE | 25.54 | 40.84 | -8.16 | 15.08 | 12.58 | 17.13 | 0.8 |
| Pan | 37.54 | -11.46 | 10.53 | 8.79 | 15.56 | 0.81 |
| 60 | 12.MoRE | 37.39 | 46.03 | -2.97 | 9.69 | 8.18 | 10.13 | 0.75 |
| Pan | 44.16 | -4.84 | 6.97 | 5.89 | 8.48 | 0.77 |
| 100 | 12.MoRE | 41.48 | 47.89 | -1.11 | 8.92 | 6.78 | 8.98 | 0.7 |
| Pan | 46.44 | -2.56 | 6.49 | 4.94 | 6.97 | 0.72 |
| 140 | 12.MoRE | 43.66 | 48.91 | -0.09 | 7.21 | 6.13 | 7.21 | 0.69 |
| Pan | 47.72 | -1.28 | 5.35 | 4.46 | 5.49 | 0.71 |

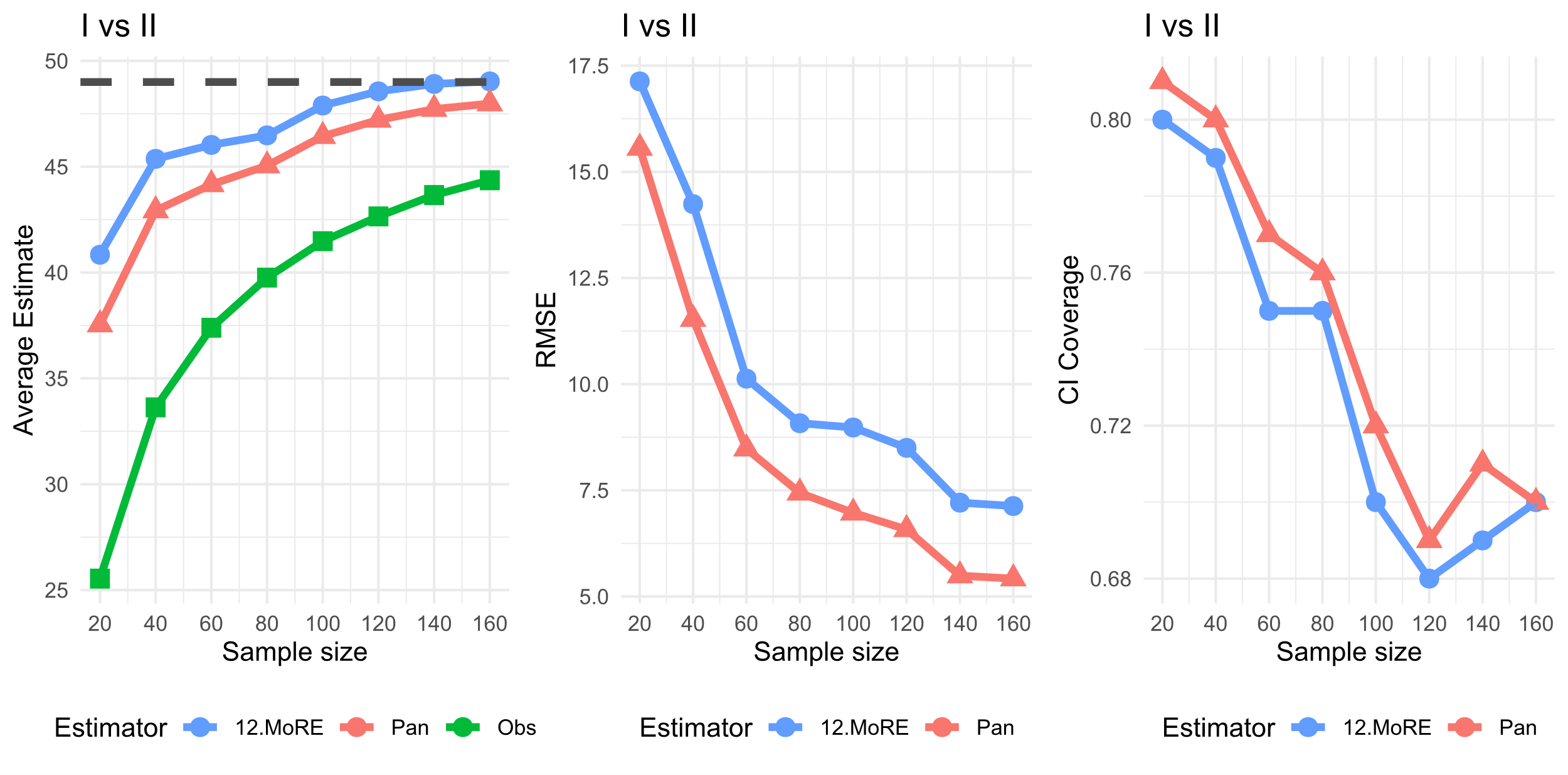


圖 1.9、取後放回澳洲三種極端氣候鳥類資料集中的Big Dry與Big Wet估計結果。

表 1.10、取後放回澳洲三種極端氣候鳥類資料集中的Big Dry與Post-Big Wet估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 20 | 12.MoRE | 24.27 | 40.19 | -11.81 | 14.49 | 13.1 | 18.69 | 0.82 |
| Pan | 36.69 | -15.31 | 10.14 | 9.13 | 18.36 | 0.84 |
| 60 | 12.MoRE | 37.16 | 47.37 | -4.63 | 10.54 | 9.4 | 11.51 | 0.8 |
| Pan | 45.27 | -6.73 | 7.85 | 6.67 | 10.33 | 0.79 |
| 100 | 12.MoRE | 41.93 | 49.97 | -2.03 | 9.39 | 8 | 9.6 | 0.76 |
| Pan | 48.09 | -3.91 | 6.9 | 5.76 | 7.93 | 0.75 |
| 140 | 12.MoRE | 44.6 | 51.59 | -0.41 | 8.53 | 7.41 | 8.53 | 0.77 |
| Pan | 49.88 | -2.12 | 6.22 | 5.31 | 6.56 | 0.76 |

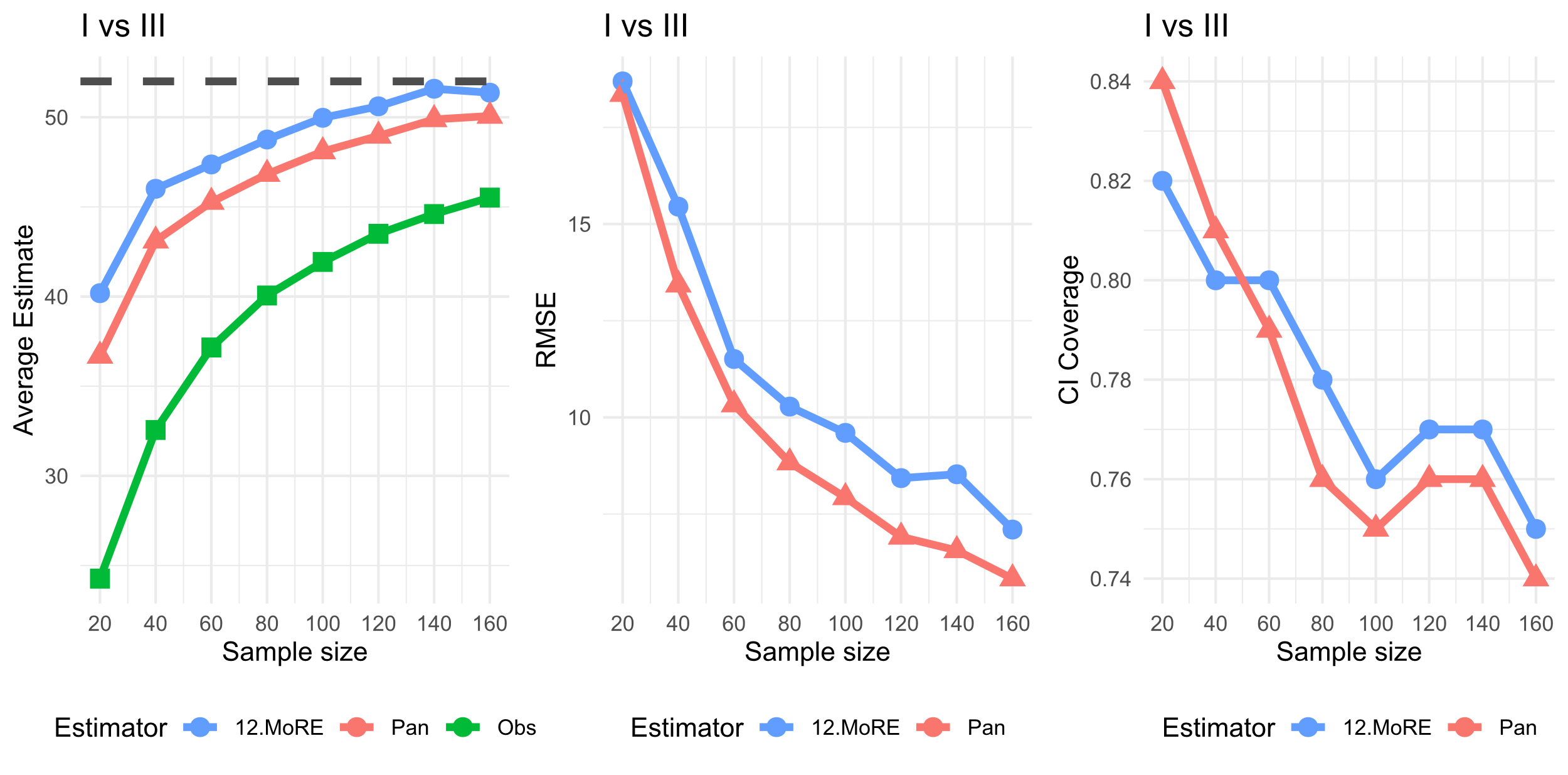


圖 1.10、取後放回澳洲三種極端氣候鳥類資料集中的Big Dry與Post-Big Wet估計結果。

表 1.11、取後放回澳洲三種極端氣候鳥類資料集中的Big Wet與Post-Big Wet估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| sample size | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 20 | 12.MoRE | 30.82 | 45.86 | -10.14 | 14.51 | 12.74 | 17.7 | 0.8 |
| Pan | 42.88 | -13.12 | 10.5 | 8.99 | 16.8 | 0.81 |
| 60 | 12.MoRE | 42.91 | 51.52 | -4.48 | 9.09 | 8.42 | 10.13 | 0.8 |
| Pan | 49.88 | -6.12 | 6.76 | 6.13 | 9.12 | 0.8 |
| 100 | 12.MoRE | 47.31 | 54.12 | -1.88 | 8.64 | 7.23 | 8.84 | 0.76 |
| Pan | 52.73 | -3.27 | 6.3 | 5.29 | 7.1 | 0.77 |
| 140 | 12.MoRE | 49.53 | 55.47 | -0.53 | 7.6 | 6.77 | 7.62 | 0.74 |
| Pan | 54.15 | -1.85 | 5.59 | 4.97 | 5.88 | 0.75 |

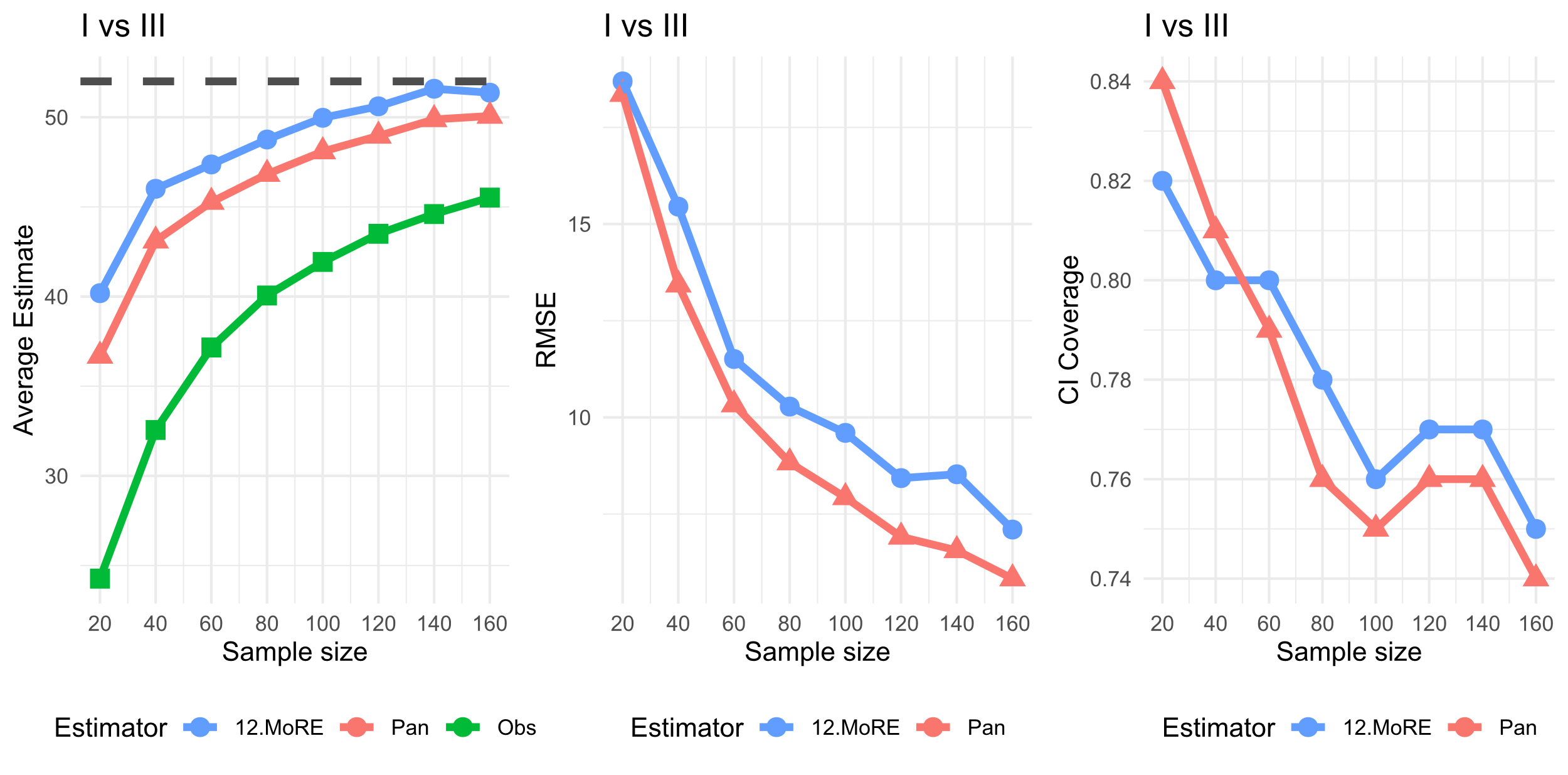


圖 1.11、取後放回澳洲三種極端氣候鳥類資料集中的Big Wet與Post-Big Wet估計結果。

### Barro Colorado Island 樹種生長狀況

**資料描述**

本資料集中，針對1981-2016年間對Barro Colorado Island進行八次調查，紀錄該地區樹種的生長狀況。將該地區非為1250個區塊進行調查，每塊區塊的大小為400平方公尺。以第一次調查 (1981至1983年間) 與第八次調查 (2013至2016年間) 的結果作為兩群落之母體資料，並使用母體的10%、30%、50% 以及70%作為抽樣樣本大小。在兩母體中，第一次調查作為群落一 (Mean =0.18, CV = 1.34)，第八次調查作為群落二 (Mean =0.18, CV = 1.35)，在群落一中，包含了307種物種，而群落二中則擁有299種物種，兩群落的共同物種為284種。

**模擬結果**

在該筆資料中使用取後不放回的估計方法下，假設該筆資料為母體，並抽取其中的區塊做為樣本所使用。在偏誤的模擬結果中，使用本文所提出的 在估計結果方面，相較於現有方法 更接近真值。而在估計標準差方面，則是在兩種估計方法下，與樣本標準差相比皆呈現低估的結果。最後是在兩個模型評估指標上，在以該資料集作為樣本的情況下，的結果相較於 皆獲得更小的RMSE與更精確的95% 信賴區間的真值涵蓋率。

表 1.12、取後不放回BCI資料集中的兩群落估計結果。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| q | Estimator | Obs | AVG Estimate | Bias | Sample SE | Est. SD | RMSE | 95% CI Coverage |
| 0.1 | 12.wMoRE1 | 225.24 | 273.94 | -10.06 | 23.46 | 26.09 | 25.51 | 0.97 |
| 12.wMoRE2 | 298.91 | 14.91 | 46.77 | 34.39 | 49.07 | 0.92 |
| wChao2.12 | 253.31 | -30.69 | 19.56 | 13.09 | 36.38 | 0.81 |
| 0.3 | 12.wMoRE1 | 256.57 | 283.48 | -0.52 | 11.3 | 10.55 | 11.31 | 0.94 |
| 12.wMoRE2 | 288.77 | 4.77 | 14.42 | 11.73 | 15.18 | 0.9 |
| wChao2.12 | 272.93 | -11.07 | 9.05 | 7.26 | 14.3 | 0.84 |
| 0.5 | 12.wMoRE1 | 269.18 | 284.51 | 0.51 | 6.35 | 5.87 | 6.37 | 0.93 |
| 12.wMoRE2 | 285.69 | 1.69 | 6.79 | 5.86 | 6.99 | 0.92 |
| wChao2.12 | 279.89 | -4.11 | 5.77 | 4.84 | 7.08 | 0.84 |
| 0.7 | 12.wMoRE1 | 276.61 | 284.09 | 0.09 | 3.37 | 3.31 | 3.37 | 0.95 |
| 12.wMoRE2 | 284.32 | 0.32 | 3.41 | 3.27 | 3.43 | 0.94 |
| wChao2.12 | 282.57 | -1.43 | 3.3 | 3.01 | 3.59 | 0.87 |

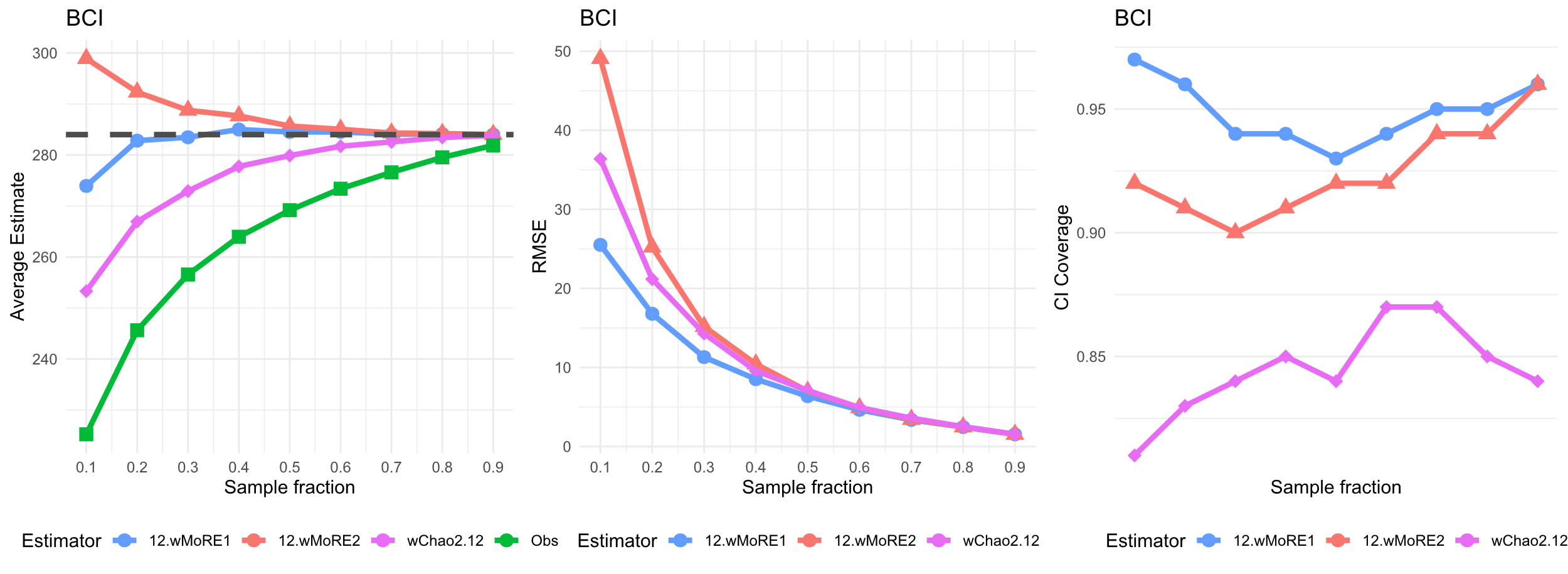


圖 1.12、取後不放回BCI資料集中的兩群落估計結果。

## 討論

總結上述結果，在第4章中使用了電腦模擬建立不同模型假設下的四種群落搭配四種不同的物種分配方式，以及分別使用兩筆真實資料作為群落母體。並針對不同的群落組合與物種區塊設定下，從中隨機以取後放回以及取後不放回兩種不同的抽樣方式，重複抽取1000次，以計算平均的估計值、偏誤、樣本標準差、估計標準差、RMSE以及95% CI Coverage的估計結果。

電腦模擬群落在偏誤結果方面，於取後放回的抽樣與估計下，本文所提出的 相較於現有方法 皆具有更小的偏誤。且在所有的模型組合與物種區塊設定下皆呈現低估的結果，除了在情況四的同質與Broken-stick的模型組合的小樣本下有高估的結果發生。而在取後不放回的抽樣模擬，無論何種物種區塊假設下的結果中，同質與Broken-stick的模型組合在抽樣比例小於0.3時，結果皆高估，同時，使用現有的估計方法 與本文所提出的方法相比皆具有更小的偏誤。但在其他的取後不回的估計中，則是本文所提出的 擁有更接近真值的估計表現。

因此針對多次模擬的結果中，針對取後不放回的抽樣估計在同質與Broken-stick的模型組合發生高估的情況中，發現在估計的物種出現率的變異係數小於0.4時，建議採用本文所提出的取後放回方法 加以估計。

在樣本標準差與估計標準差的結果中，無論是利用何種抽樣方式、物種區塊假設、模型組合以及估計方法下，估計標準差皆低於樣本標準差。且在兩種抽樣估計的模擬中，使用本文所提出的估計式 、 與 的的標準差結果皆較現有方法 與 具有更大的變異。

在兩個模型評估指標RMSE與95% CI Coverage上，取後放回或是取後不放回的估計方法，分別是本文所提出的與估計方式具有較小的RMSE與較精確的樣本涵蓋率。

而除了電腦模擬的群落之外，也使用了兩筆真實資料分別針對兩種抽樣方式作為群落模擬。在偏誤結果方面，使用本文所提出的與方法估計皆得到更接近真值的結果；在標準差估計方面，皆是本文所提出的方法具有更大的變異，且所有方法中皆有低估的現象發生；最後是在在兩個模型評估的指標下，取後放回的模擬估計中，現有方法 具有較佳的表現，但與本文所提出的 之間並沒有明顯差異。而在取後不放回的模擬估計中，則是本文所提出的方法 具有較小的RMSE與較精確的95% CI Coverage。

綜觀上述的所有模擬結果，無論是在何種模型組合與物種區塊假設下，皆是本文所提出的 與 分別在兩種抽樣方式下，偏誤結果上更接近真值，且在模型評估上，普遍具有更小的RMSE與更精確的95% CI Coverage。故在針對出現型樣本估計兩群落的共同物種時，推薦使用本文所提出的 與 ，分別針對取後放回與取後不放回的樣本進行估計。