

With larger delta (= bigger bin sizes) like in case 1& 2, all sample means are going to the 0.4-0.5 and 0.5-0.6 two bins. The PMFs are large in these two cases and close to 0.5 for both groups. With smaller delta (= smaller bin sizes), the sample means are more separated like in case 3& 4 because a smaller bin sizes return with more separated PMFs. Also, the PMFs are much smaller compare to case 1& 2, with no bigger than 0.03 in case 3 and 0.06 in case 4. From the above observation, we can expect a bigger delta and larger sample sizes return with higher PMF. The reason is because with a bigger delta, sample means are grouped more centralized and the PMFs for such centralized groups are always higher than the decentralized one. With only two bins containing sample means in case 1&2, the PMFs for both groups are much higher than in case 3&4. In case 3&4, the sample size ( $n$ ) plays a significant role in the PMFs. With a bigger sample size, the PMFs are higher because the sample means will become higher as it groups more samples in one group which the groups are less separated compare to a small sample size in case 3.

A smaller delta and larger sample size will return with a closer PMF to the true sampling distribution of the mean, as we can see from case 3&4, the distribution is close to a normal distribution while we lower the delta. The standard deviation in case 3 is lower than case 4 as we increased the sample size where the sample means are more squeezed in under a larger sample size.

The true population mean in the data is at 0.50030, where case 3&4 return us with a much closer result in  $E[X]$  (around 0.4998) compare to case 1&2 (around 0.449). As we decrease the delta and getting a more decentralized distribution of sample means, the odd of getting far away from the population mean is lower compare to case 1&2. We can have a reasonable group of sample means with a lower delta.