Random sampling plans

Blocks population

Consider again the `blocks` data, as our study population $\mathcal{P}_{Study}$ consisting of $N = 100$ blocks labelled $u = 1, 2, 3, \ldots, 100$.

---

20 marks

In this question, you will investigate different sampling plans and estimation procedures.

a. Simple random sampling.

i. *(4 marks)* Collect the sample average block weight from each of 1000 samples, where each sample consists of 10 blocks selected at random (without replacement) from all 100 blocks.

Before sampling, `set.seed(314159)`

Save the results on the R variable `randomSampleAves`.

Show your code.

```
meanbar <- rep(NA, 1000)
variability <- rep(NA, 1000)
sampleErrors <- rep(NA, 1000)
randomSampleAves = list()
set.seed(314159)
for(i in 1:1000)
{
  S <- sample(blocks$weight,size = 10, replace = FALSE)
  randomSampleAves = c(S,randomSampleAves)
  meanbar[i]<- mean(S)
  variability[i] <- (mean(S)-mean(blocks[,'weight']))^2
  sampleErrors[i] <- mean(S) - mean(blocks[,'weight'])
}
mean_sample_weight <- mean(meanbar)
mean_sample_weight
```

```
## [1] 32.441
```

ii. *(3 marks)* Using `randomSampleAves`, estimate the sampling bias, the sampling variability, and the sampling mean squared error of this sampling plan.

Show your code.

```
samplingBias <- mean(meanbar) - mean(blocks[,'weight'])
samplingBias
```

```
## [1] 0.041
```

```
samplingVariability <- sum(variability)/1000
samplingVariability
```

```
## [1] 23.1042
```

```
samplingMSE <- samplingVariability + (samplingBias)^2
samplingMSE
```

```
## [1] 23.10588
```

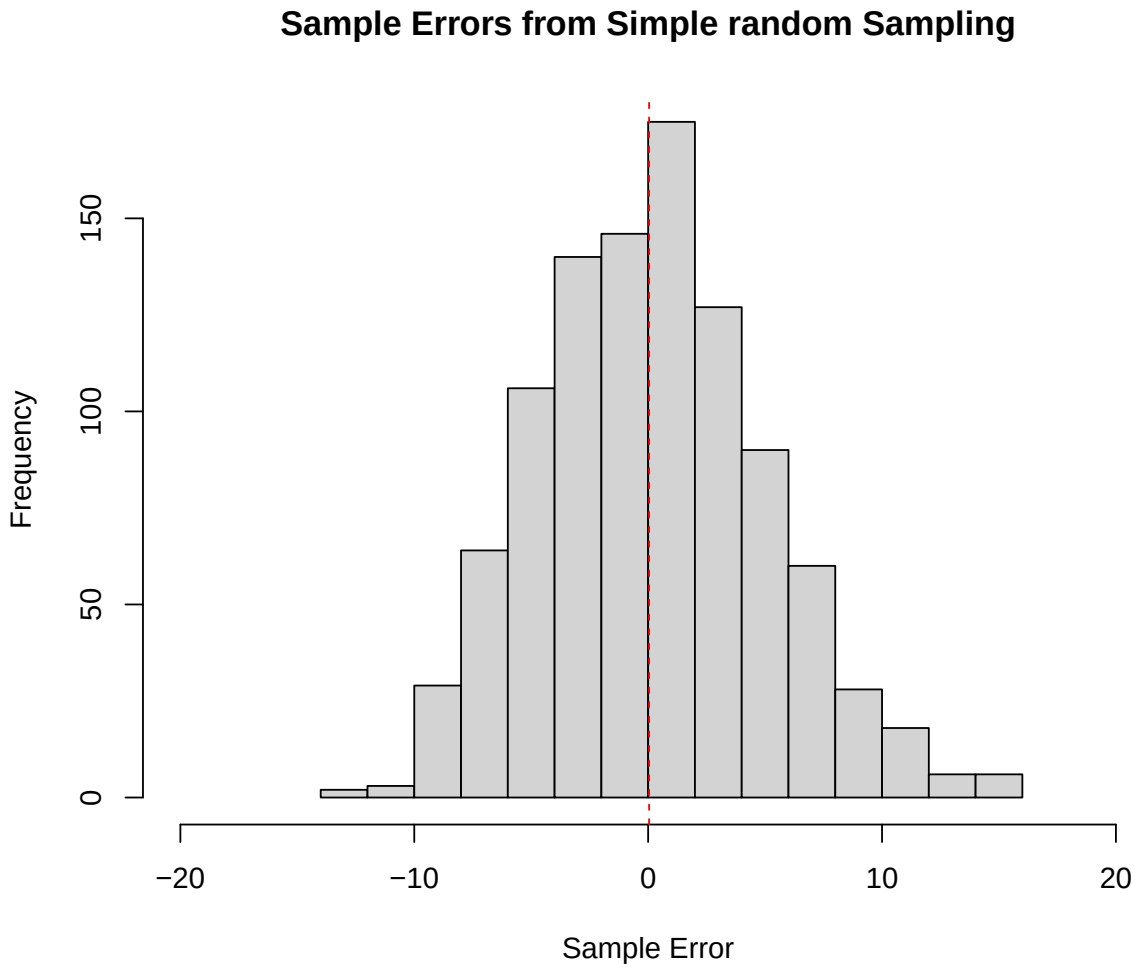
iii. *(3 marks)* Construct a (suitably labelled) histogram of the sample **errors** from this sampling plan.

Use `xlim = c(-20,20)`.

Add a vertical red dashed line of `lwd = 2` at the average error.

Show your code.

```
xlim = c(-20,20)
hist(sampleErrors, xlim = xlim, main = "Sample Errors from Simple random Sampling", xlab = "Sa
abline(v = mean(sampleErrors), col = "red", lty = 2)
```



b. Stratified random sampling.

i. *(4 marks)* Collect the sample average block weight from each of 1000 samples, where now each sample consists of 5 blocks selected at random (without replacement) from each of group "A" and group "B".

Before sampling, `set.seed(314159)`

Save the results on the R variable `stratifiedSampleAves`.

Show your code.

```
groups <- with(blocks, unique(group))
stratifiedSample <- lapply(groups, FUN = function(samp){with(blocks,blocks[group == samp,])})

stratifiedmeanbar <- rep(NA, 1000)
stratifiedvariability <- rep(NA, 1000)
stratifiedsampleErrors <- rep(NA, 1000)
stratifiedSampleAves = list()
set.seed(314159)
for(i in 1:1000)
{
  A <- sample(stratifiedSample[[2]]$weight, size= 5, replace = FALSE)
  B <- sample(stratifiedSample[[1]]$weight, size= 5, replace = FALSE)
  A_B <- c(A,B)
  stratifiedSampleAves = c(A_B, stratifiedSampleAves)
  stratifiedmeanbar[i] <- mean(A_B)
  stratifiedvariability[i] <- (mean(A_B) - mean(blocks[,'weight']))^2
  stratifiedsampleErrors[i] <- mean(A_B) - mean(blocks[,'weight'])
}
mean_stratified_weight <- mean(stratifiedmeanbar)
mean_stratified_weight
```

```
## [1] 32.4515
```

ii. *(3 marks)* Using `stratifiedSampleAves`, estimate the sampling bias, the sampling variability, and the sampling mean squared error of this sampling plan.

Show your code.

```
stratifiedBias <- mean(stratifiedmeanbar) - mean(blocks[,'weight'])
stratifiedBias
```

```
## [1] 0.0515
```

```
stratifiedVariability <- sum(stratifiedvariability)/1000
stratifiedVariability
```

```
## [1] 11.49405
```

```
stratifiedMSE <- stratifiedVariability + (stratifiedBias)^2
stratifiedMSE
```

```
## [1] 11.4967
```

iii. *(3 marks)* Construct a (suitably labelled) histogram of the sample **errors** from this sampling plan.

Use `xlim = c(-20,20)`.

Add a vertical red dashed line of `lwd = 2` at the average error.

Show your code.

```
xlim = c(-20,20)
hist(stratifiedsampleErrors, xlim = xlim, main = "Sample Errors from Stratified Random Samplin
abline(v = mean(stratifiedsampleErrors), col = "red", lty = 2)
```