

620-374 Sampling

Lab task week 4

Obtain the data `ufc.csv` from the web site. It contains observations of five variables for a population of trees: Plot (plot where tree was grown); Tree (label of tree within a plot); Species (DF for Douglas Fir, etc); Dbh (diameter at breast height); and Height.

Only a sample of values is available for Height.

- 1) Plot Height against Dbh for all sampled trees. Does it look as if a ratio or regression estimator of mean Height would do any better than an SRS estimator?

Repeat your plot for Douglas Firs only. For this reduced population, does it look as if a ratio or regression estimator of mean Height would do any better than an SRS estimator?

- 2) Calculate $\hat{\mu}_{SRS}$, $\hat{\mu}_{ratio}$ and $\hat{\mu}_{lr}$ (using the optimal b) estimators for μ_{Height} , using DBh as the auxiliary variable. Using sample estimates of the variance, which do you think is the most accurate?
- 3) What assumptions are we making for this analysis to be valid?