Statistics and R short course

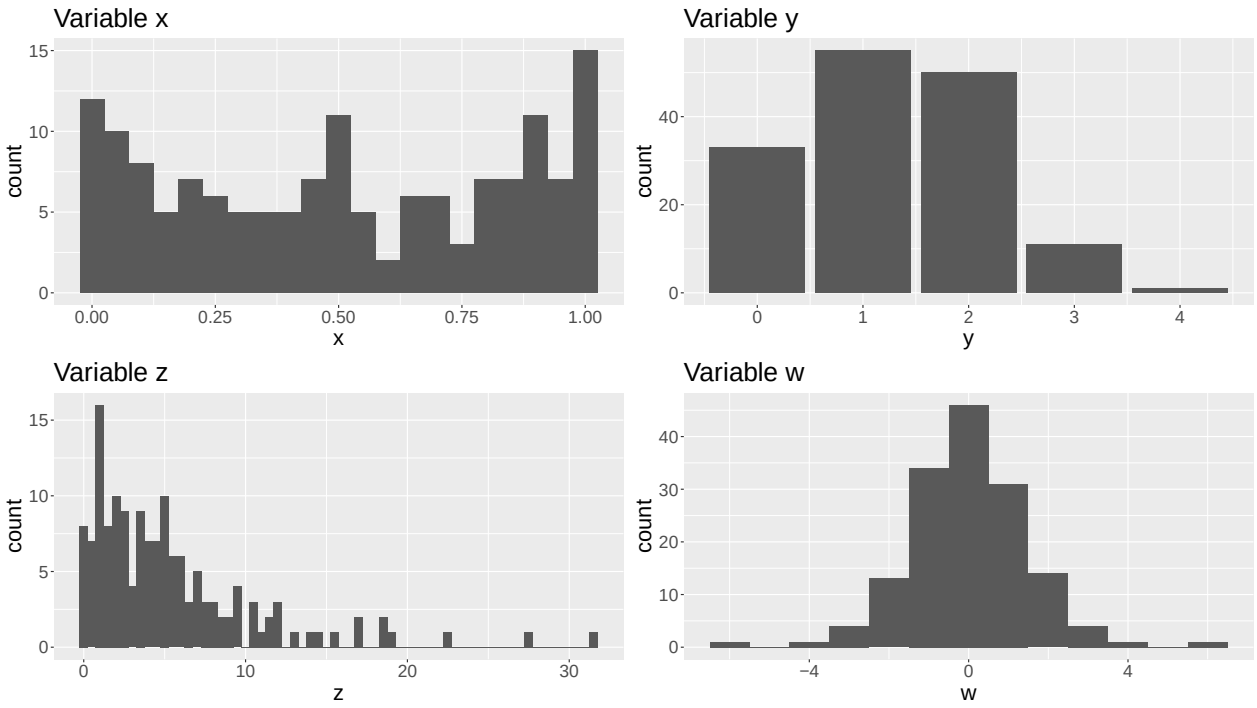
Augustine Choko, Marc Henrion, James Chirombo

01 December 2020

Session 3 - Practical

Exercise 1 - CLT

Which distributions do you think gave rise to each of the variables displayed below?



Exercise 2 - Central Limit Theorem

In the lecture we saw how to empirically prove the CLT for the normal, beta and negative binomial distributions. Do the same now for the exponential and the Poisson distributions.

Exercise 3 - study design

Decide what design could be used to answer the following research questions:

1. What is the prevalence of HIV in urban Blantyre in 2018?
2. Do men experience higher mortality compared to women once they start ART?
3. Does smoking increase the chance of having lung cancer?
4. What is the effect of providing oral HIV self-test kits on the uptake of HIV testing?
5. What interventions may improve linkage to ART following community based HIV testing?

Exercise 4 - sample size calculation

The difference of the mean birth weights between Babies born at Queen Elizabeth Central hospital (QECH) and Kamuzu Central Hospital (KCH) will be determined. Suppose at QECH the mean is 3000 g with a standard deviation of 500 g (from previous/pilot studies). At KCH the mean is 3300 with a s.d of 500 grams. The difference in mean birth weight to be detected is therefore 300 g. The significance level for the test will be 0.05. What would be the sample size required to detect this difference at 80% power?