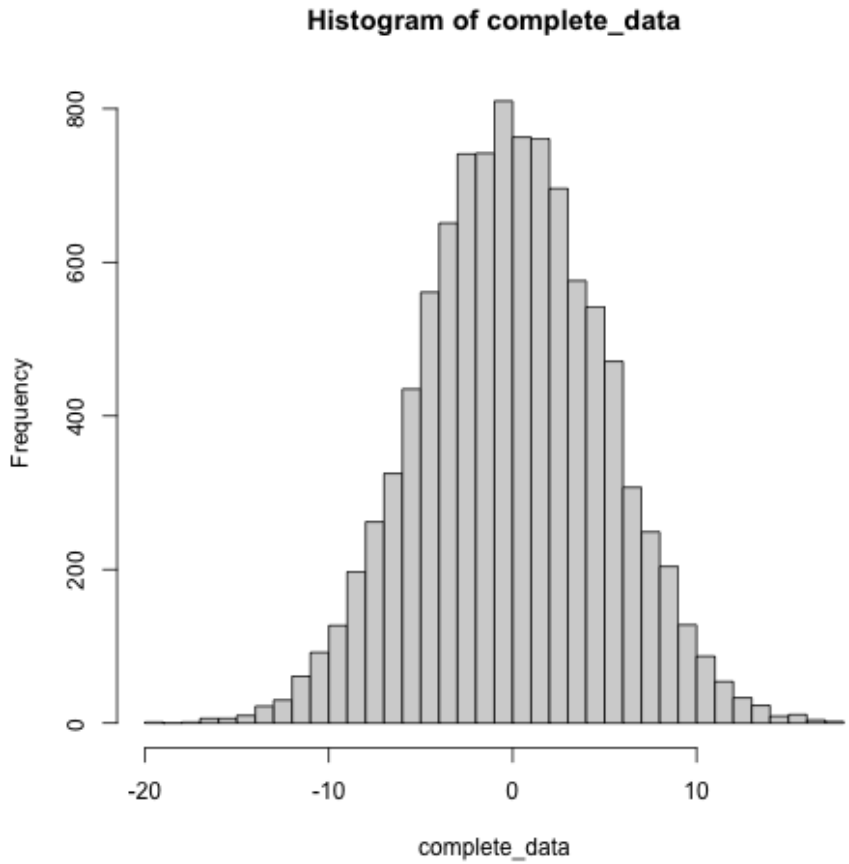
Batch Normalization

Naresh Manwani

Machine Learning Lab, IIIT Hyderabad

Covariate Shift: Example 1

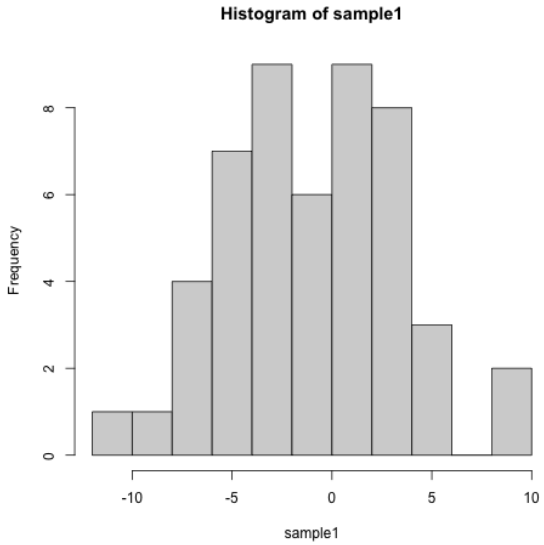
- Consider the data coming from a Gaussian distribution $\mathcal{N}(0, 5)$
- Generate 10000 samples from this ditribution. Call it complete data.

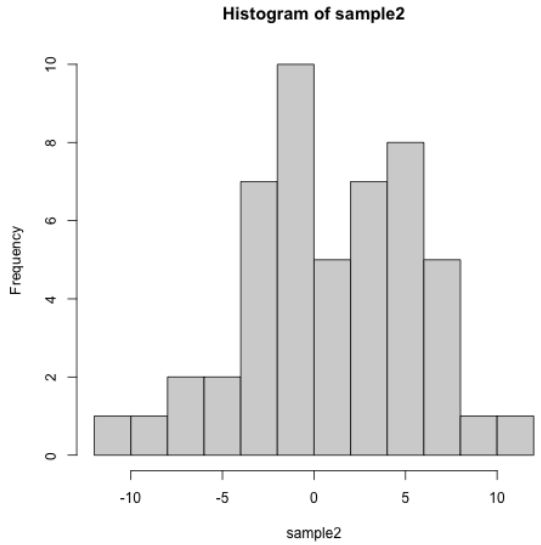


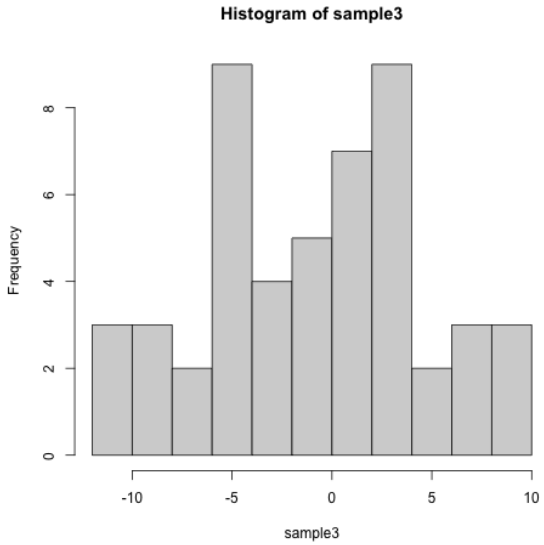
Covariate Shift: Example 1

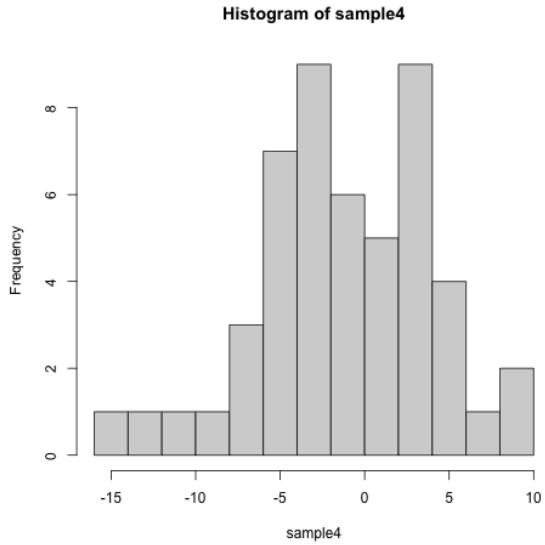
- Now consider minibatches of size 50 randomly sampled from complete data.
- Each minibatch is sampled without replacement.
- Four such minibatches were sampled.

Example 1: Density contour plots of minibatches



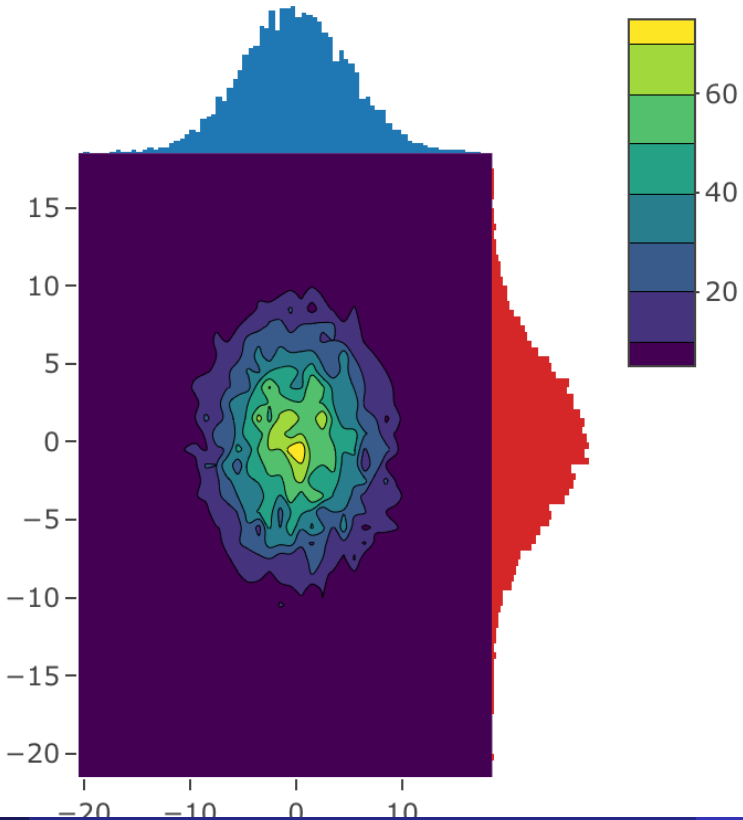






Covariate Shift: Example 2

- Consider the data coming from a Gaussian distribution $\mathcal{N}\mu, \Sigma)$ where $\mu = [0 \;\; 0]$ and $\Sigma = \begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}$
- Generate 10000 samples from this ditribution. Call it complete data.



Covariate Shift: Example 2

- Now consider minibatches of size 50 randomly sampled from complete data.
- Each minibatch is sampled without replacement.
- Four such minibatches were sampled.

Covariate Shift: Example 2