

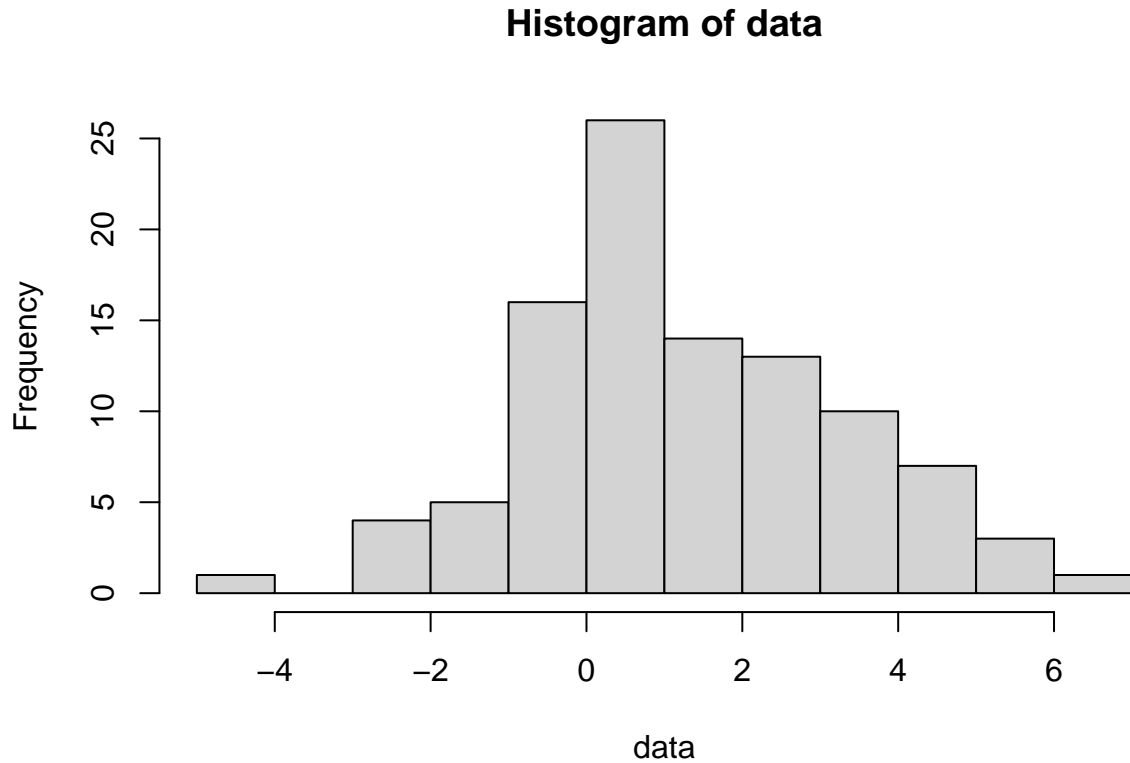
## lab2 Problem2

Jaskirat S Marar

11/12/2021

### Log Likelihood estimators $\mu$ & $\sigma$

The given data is plotted for observation.



As we can see that the data seems to be normally distributed with its mean lying somewhere close to 1, which confirms what is given to us in the problem statement about the data being distributed normally. Let us proceed further with solving this problem.

Given our data, we can write the likelihood as the joint density of the sample i.e.

$$L(\mu, \sigma^2) = \prod_{i=1}^n f(y_i | \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y_i - \mu}{2\sigma^2}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2\right)$$

Taking logs on both sides

$$\ln[L(\mu, \sigma^2)] = -\frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

For MLE of  $\mu$  &  $\sigma$ , we set the partial derivatives w.r.t.  $\mu$  &  $\sigma$  and set them to zero.

First, derivating w.r.t  $\mu$ , the first and second term become 0, and we derivate the 3rd term leading to:

$$\frac{\partial(\ln[L(\mu, \sigma^2)])}{\partial\mu} = \frac{1}{\sigma^2} \left( \sum_{i=1}^n (y_i) - n\mu \right)$$

If we equate this to 0, and isolate  $\mu$ , we get:

$$\mu = \frac{1}{n} \sum_{i=1}^n y_i$$

This we know to be the sample mean, we can summarize the result as follows:

$$\hat{\mu}_{MLE} = \bar{y}$$

Similarly for  $\sigma$ , now we derivate the log likelihood function w.r.t.  $\sigma$ , the second term becomes zero and we get:

$$\frac{\partial(\ln[L(\mu, \sigma^2)])}{\partial\sigma} = -\frac{n}{2} \cdot \frac{2}{\sigma} - \left( \frac{\sum_{i=1}^n (y_i - \mu)^2}{2} \right) \cdot \left( -\frac{2}{\sigma^3} \right)$$

Simplifying:

$$\frac{\partial(\ln[L(\mu, \sigma^2)])}{\partial\sigma} = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n (y_i - \mu)^2}{\sigma^3}$$

Setting partial derivative to 0 and solving for  $\sigma$ , we get

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \mu)^2}{n}$$

And this expression we know is the sample variance, so we can summarize as:

$$\hat{\sigma}_{MLE}^2 = s^2$$

In terms of sample values, the variance can be written as follows:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \frac{1}{n} \sum_{i=1}^n y_i)^2}{n}$$

Now that we have simplified the functions for calculating the estimators, we can use the built in R functions for mean and variance for our given data:

```
mu_data = mean(data)
sigma_sq_data = var(data)

cat("mu = ", mu_data, "\n")
```

```
## mu = 1.275528
```

```
cat("sigma = ", sqrt(sigma_sq_data), "\n")
```

```
## sigma = 2.016082
```

## Writing the loglikelihood function

We now write the function for calculating the log-likelihood which will be then used to optimize with initial parameters,  $\theta$

```
loglikelihood <- function(theta, x) {  
  mu = theta[1]  
  sig = theta[2]  
  n = length(x)  
  lgllhood = -(n/2)*(log(2*pi*sig^2)) + (-1/(2*sig^2))*sum((x-mu)^2)  
  return(-lgllhood)  
}
```

We are passing the negative value of the log-likelihood as the `optim()` minimizes that are passed into it. Hence, in order to maximize log-likelihood, we return the negative value of the log-likelihood.

Next we pass the initial parameters  $\mu = 0$  and  $\sigma = 1$  and optimize as follows:

```
optim_norm_BFGS_withoutgradient <- optim(par = c(0,1), loglikelihood, x = data, method = "BFGS")  
optim_norm_BFGS_withoutgradient
```

```
## $par  
## [1] 1.275528 2.005977  
##  
## $value  
## [1] 211.5069  
##  
## $counts  
## function gradient  
##      37      15  
##  
## $convergence  
## [1] 0  
##  
## $message  
## NULL
```

```
optim_norm_CG_withoutgradient <- optim(par = c(0,1), loglikelihood, x = data, method = "CG")  
optim_norm_CG_withoutgradient
```

```
## $par  
## [1] 1.275528 2.005977  
##  
## $value  
## [1] 211.5069  
##  
## $counts  
## function gradient  
##      297      45  
##  
## $convergence  
## [1] 0  
##  
## $message  
## NULL
```