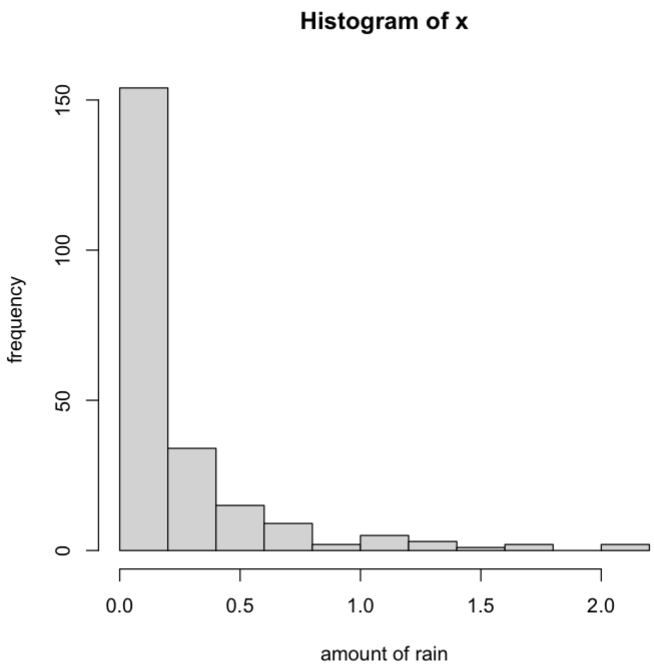
1. Histogram of data (amount of rain)

- It is known that our data follows a gamma distribution, so we want to estimate the parameters α and β.

In [1]:
```
gamma = read.table("gamma.txt", header=T)
x = gamma$x

# Histogram
options(repr.plot.width = 6, repr.plot.height = 6)
hist(x,xlab="amount of rain",ylab="frequency")
```



2. MME of gamma distribution

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta)$. If $X \sim \text{Gamma}(\alpha, \beta)$, then $E[X] = \alpha\beta$ and $E[X^2] = \beta^2(\alpha + \alpha^2)$. So the method of moments estimators $\hat{\alpha}, \hat{\beta}$ solve the equations

$$\hat{\alpha}\hat{\beta} = \hat{\mu}_1$$
$$\hat{\beta}^2(\hat{\alpha} + \hat{\alpha}^2) = \hat{\mu}_2.$$

Substituting the first equation into the second,

$$\left(\frac{1}{\hat{\alpha}} + 1\right)\hat{\mu}_1^2 = \hat{\mu}_2$$

so

$$\hat{\alpha} = \frac{1}{\frac{\hat{\mu}_2}{\hat{\mu}_1^2} - 1} = \frac{\hat{\mu}_1^2}{\hat{\mu}_2 - \hat{\mu}_1^2} = \frac{\bar{X}^2}{\frac{1}{n}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}.$$

The first equation then yields

$$\hat{\beta} = \frac{\hat{\mu}_1}{\hat{\alpha}} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{\bar{X}}.$$

In [2]:
```
# MME
n = length(x) # sample size
xbar = mean(x) # sample mean
alp = xbar^2/(sum((x- xbar)^2)/n) # MME of alpha
bet = (sum((x-xbar)^2)/n)/xbar  # MME of beta
```

3. MLE of gamma distribution

- $\hat{\beta}^{MLE} = \bar{X}/\hat{\alpha}^{MLE}$
- But $\hat{\alpha}^{MLE}$ does not have an analytic form. Hence, use Newton-Raphson method.

First derivative of the log-likelihood (Score Function)

$$l'(\alpha) = n\log\left(\frac{\alpha}{\bar{X}}\right) - n\psi(\alpha) + \sum_{i=1}^{n}\log(X_i).$$

- Here, $\psi(\alpha)$ is the digamma function .

Second derivative of the log-likelihood

$$l''(\alpha) = \frac{n}{\alpha} - n\psi'(\alpha).$$

Newton-Raphson update equation

$$\hat{\alpha}^{(t+1)} = \hat{\alpha}^{(t)} - \frac{l'(\hat{\alpha}^{(t)})}{l''(\hat{\alpha}^{(t)})}, \quad t = 0, 1, \ldots.$$

Convergence criterion

The iterations continue until the following condition is met:

$$\epsilon = |\hat{\alpha}^{(t+1)} - \hat{\alpha}^{(t)}| > 10^{-7}.$$

In [3]:
```
# MLE: Newton-Raphson
epsilon = 1 # initial valeu of epsilon
alpha = alp  # initial value of alpha

while (epsilon > 0.0000001) {
    temp = alp
    # first derivative of log-likelihood
    l1 = n*log(temp/xbar) - n*psigamma(temp,0) + sum(log(x))
    # second derivative
    l2 = n/temp - n*psigamma(temp,1)
    # updating equation
    alp = temp - l1/l2
    # calculating epsilon
    epsilon = abs(alp - temp)
    alpha = append(alpha,alp)
}
```

```
alpha_l = alpha[length(alpha)]    # MLE of alpha
```

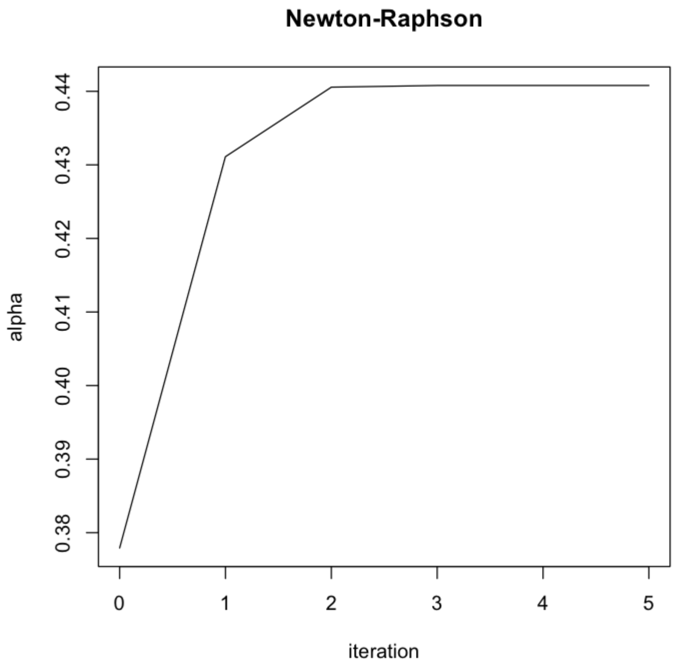
4. MME vs MLE

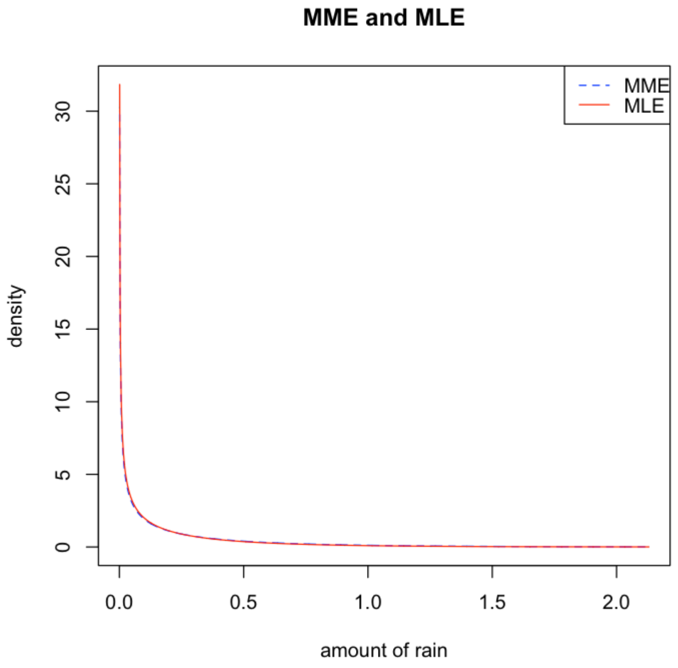
In [4]:
```
xgrid = 1:length(alpha)-1
plot(xgrid,alpha,xlab="iteration",ylab="alpha",main="Newton-Raphson",type="

beta_l = xbar/alpha_l    # MLE of beta


# Gamma fitting
xgrid = seq(from=min(x),to=max(x),by=0.001)    # x point
ym_gam = dgamma(xgrid, shape=alp, scale=bet) # MME density
yl_gam = dgamma(xgrid, shape=alpha_l, scale=beta_l) # MLE density

plot(xgrid,ym_gam,col="blue",type="l",lty=2,xlab="amount of rain",
ylab="density",main="MME and MLE",xlim=c(min(x),max(x)),ylim=c(0,max(max(ym_
par(new=T)
plot(xgrid, yl_gam, col="red",type="l",lty=1,axes=F,xlab="",
     ylab="",,xlim=c(min(x),max(x)),ylim=c(0,max(max(ym_gam,yl_gam))))
legend("topright", c("MME","MLE"),col=c("blue","red"),lty=c(2,1))
```





Zoom

In [5]:
```
xgrid = seq(from=min(x),to=0.1,by=0.001)    # x point
ym_gam = dgamma(xgrid, shape=alp, scale=bet) # MME density
yl_gam = dgamma(xgrid, shape=alpha_l, scale=beta_l) # MLE density

plot(xgrid,ym_gam,col="blue",type="l",lty=2,xlab="amount of rain",
ylab="density",main="MME and MLE",xlim=c(0,0.1),ylim=c(0,35))
par(new=T)
plot(xgrid, yl_gam, col="red",type="l",lty=1,xlab="",ylab="",axes=F,
xlim=c(0,0.1),ylim=c(0,35))
legend("topright", c("MME","MLE"),col=c("blue","red"),lty=c(2,1))
```