Likelihood

Theory

Given 3 observations .8, 1.2, 1.1 and a model $y_i \sim Norm(1, .2)$

```
obs <- c(0.8, 1.2, 1.1) # observations

# probability of seeing  these  in n(1, .2)
p1 <- prod(dnorm(obs, mean = 1, sd = 0.2))
p1
```

```
[1] 2.576671
```

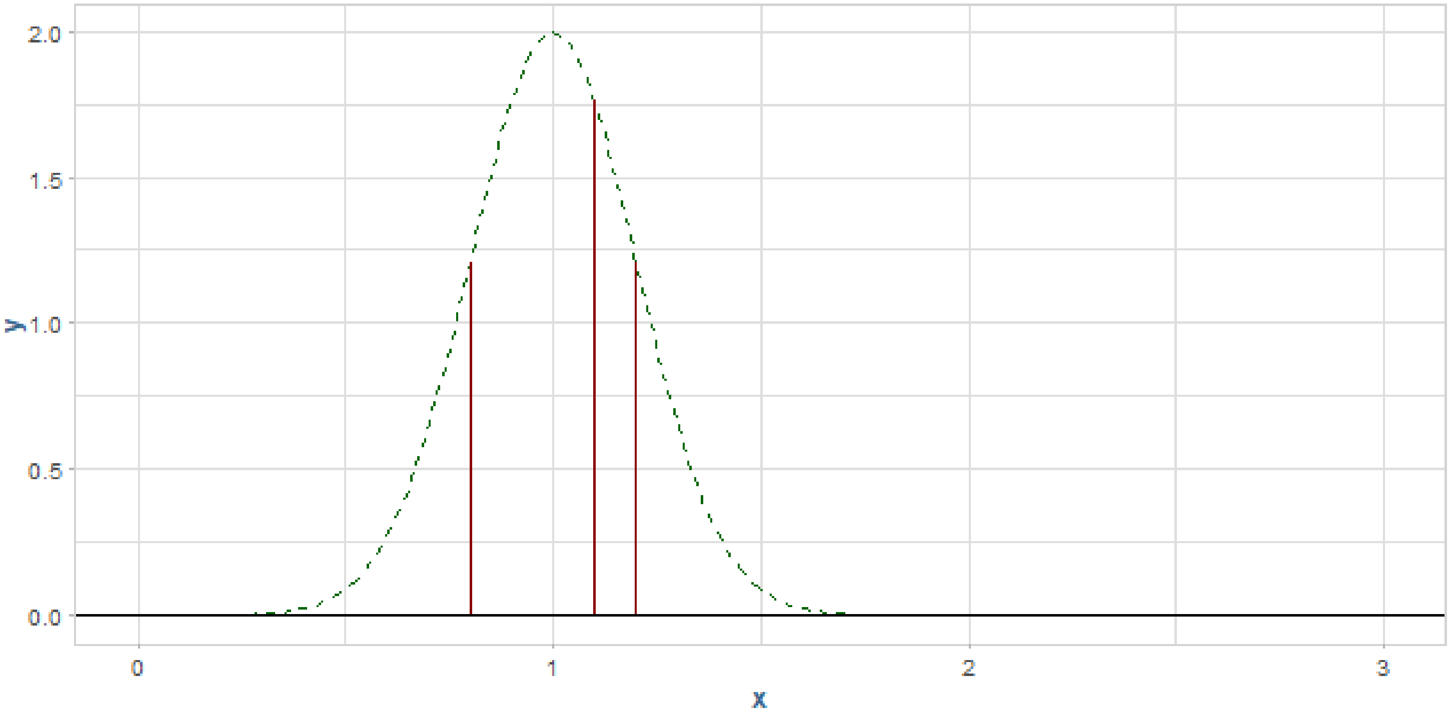
```
# probability of seeing  these  in n(1.2, .4)
p2 <- prod(dnorm(obs, mean = 1.2, sd = 0.4))
p2
```

```
[1] 0.5832185
```

```
p1/p2 # 4x liklier to see under model 1 than model 2
```

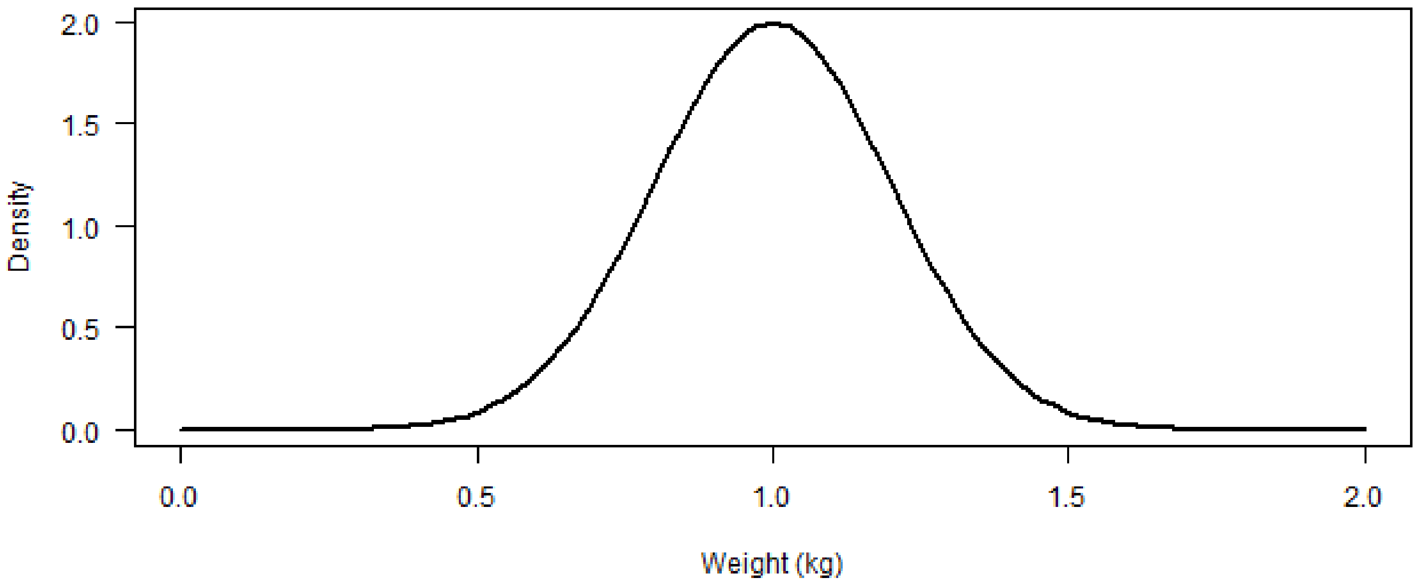
```
[1] 4.41802
```

```
# graphically
xs <- seq(from=0, to=3, by = .01)
ys <- dnorm(xs, mean = 1, sd = .2)

dat <- data.table(x = xs, y = ys)

obs_marks <- data.table(x = obs, y = dnorm(obs, mean = 1, sd = .2))

ggplot(dat) +
   geom_line(aes(x,y), col = "darkgreen", lty=2) +
   geom_hline(yintercept = 0) +
   geom_segment(data = obs_marks, aes(x, y, xend=x, yend=0), col="darkred")
```



Quick Look at the Normal Distribution

```
x <- seq(0, 2, length = 100)
dx <- dnorm(x, mean = 1, sd = 0.2)
plot(x, dx, type = "l", xlab = "Weight (kg)", ylab = "Density", lwd = 2, las = 1)
```




```r
rnorm(5, 1, 0.2) # draws 5 random numbers from Norm(1, 0.2)
```

```
[1] 1.1477332 1.0755553 0.8027062 1.3199443 0.9106785
```

```r
pnorm(q = .8, 1, 0.2)
```

```
[1] 0.1586553
```

$L(\mu, \sigma|y) = \prod_{i=1}^{n} p(y_i, |\mu, \sigma)$

```r
y <- c(0.8, 1.2, 1.1)
lf <- function(mu, sigma) prod(dnorm(y, mu, sigma))

lf(1, .2)
```

```
[1] 2.576671
```

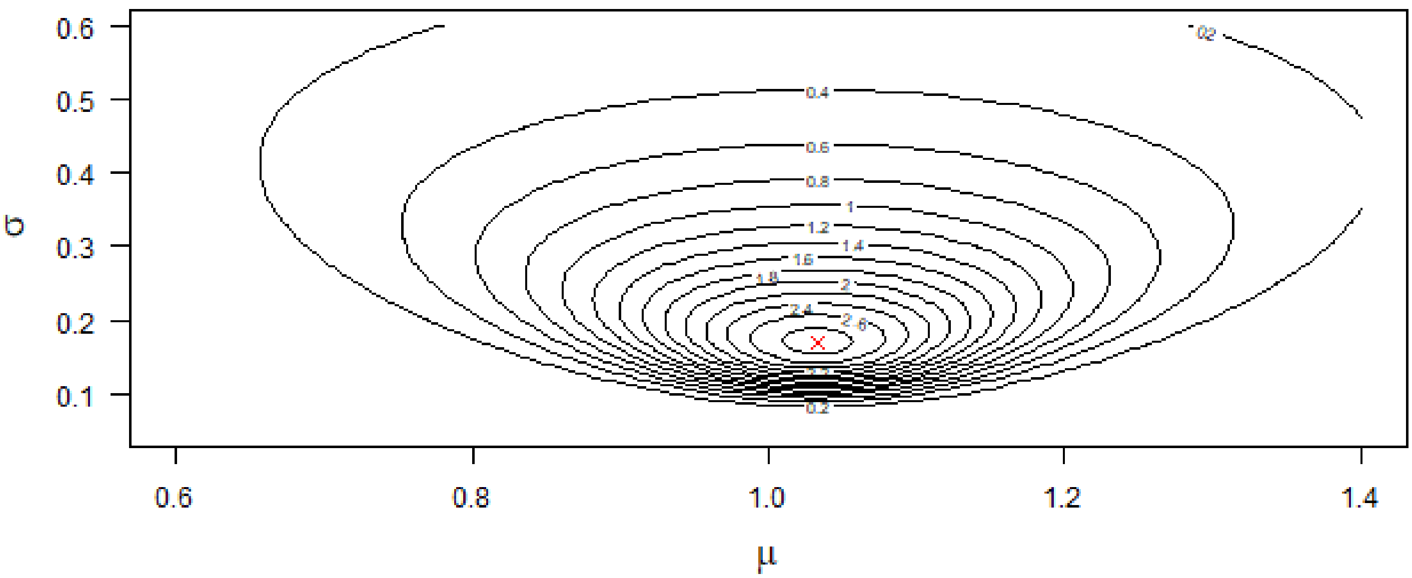
The Maximum Likelihood Method

```r
mu <- seq(0.6, 1.4, length=100)
sigma <- seq(0.05, 0.6, length=100)

lik <- matrix(nrow=length(mu), ncol = length(sigma))

for(i in 1:length(mu)) {
   for(j in 1:length(sigma)) {
```

```
      lik[i, j] <- lf(mu[i], sigma[j])
   }
}

contour(mu, sigma, lik, nlevels=20, xlab = expression(mu),
        ylab =expression(sigma), las=1, cex.lab = 1.4)

neglf <- function(x) -prod(dnorm(y, x[1], x[2]))
MLest <- optim(c(1, 0.2), neglf)
MLest$par
```

```
[1] 1.0333272 0.1699747
```

```
# optimal values
points(MLest$par[1], MLest$par[2], col="red", pch=4)
```



Likelihood ratio = p1/p2

The Log Pointwise Predictive Density

$llpd = \sum_{i=1}^{n} log \int p(y_i|\theta)p(\theta|y)d\theta$

```
mod <- lm(y~1) # fit model by LS method
nsim <- 2000
bsim <- sim(mod, n.sim=nsim) # simulate from posterior dist. of parameters
```

```r
pyi <- matrix(nrow=length(y), ncol=nsim)

for(i in 1:nsim) pyi[, i] <- dnorm(y, mean=bsim@coef[i, 1],
                                   sd=bsim@sigma[i])

mpyi <- apply(pyi, 1, mean)

sum(log(mpyi))
```

```
[1] 0.1595696
```

The log posterior density can be used as a measure of model fit.