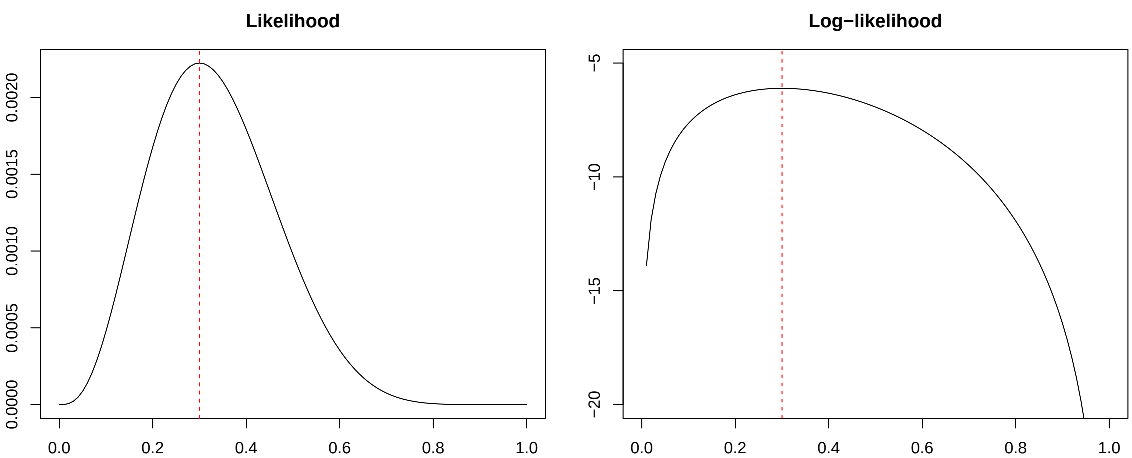
Module 3: Exercises for binomial model

Exercise 1 (solve by inserting code in the Rmd file)

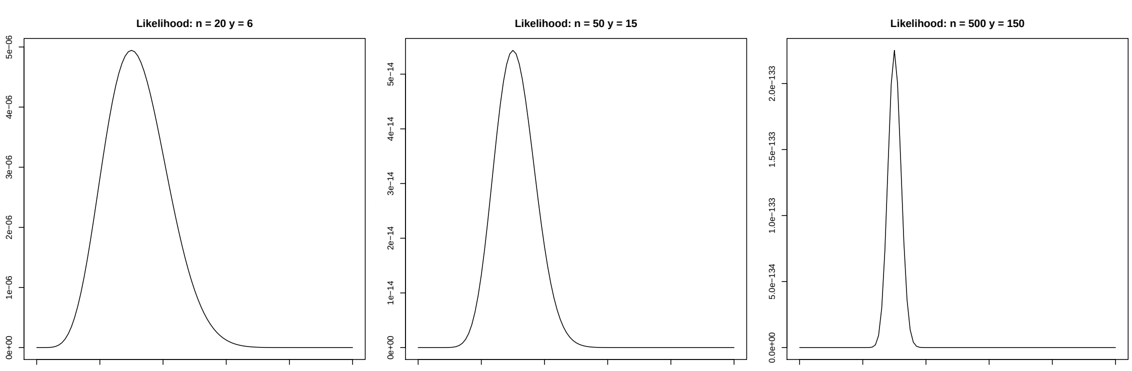
The following figure shows the likelihood and log-likelihood for the binomial model when $n = 10$ and $y = 3$:

```
lik <- function(parm, y, n){parm^y * (1 - parm)^(n - y)}
loglik <- function(parm, y, n){y * log(parm) + (n - y) * log(1 - parm)}
par(mfrow=c(1, 2), mar = c(3,3,3,1))
n <- 10; y <- 3
curve(lik(x, y, n), main = "Likelihood")
abline(v = y/n, lty = 2, col = 2)
curve(loglik(x, y, n), main = "Log-likelihood", ylim = c(-20, -5))
abline(v = y/n, lty = 2, col = 2)
```



Redo the likelihood plot above for $n = 20, y = 6$, for $n = 50, y = 15$ and for $n = 500, y = 150$.

```
par(mfrow=c(1, 3), mar = c(3,3,3,1))
n <- 20; y <- 6
curve(lik(x, y, n), main = paste("Likelihood:", "n =", n, "y =", y))
n <- 50; y <- 15
curve(lik(x, y, n), main = paste("Likelihood:", "n =", n, "y =", y))
n <- 500; y <- 150
curve(lik(x, y, n), main = paste("Likelihood:", "n =", n, "y =", y))
```



- Do the plots surprise you?
- What do you conclude? As n (the number of trials) grows the uncertainty of the parameter estimate becomes smaller.

Exercise 2 (solve by hand with pen and paper)

For the binomial model

- Differentiate $l(\theta)$ to obtain $l'(\theta)$ and verify that the solution to $l'(\theta) = 0$ is $\hat{\theta} = y/n$.

$$l'(\theta) = \frac{y}{\theta} - \frac{n-y}{(1-\theta)}$$

$$l'(\theta) = 0 \Leftrightarrow \frac{y}{\theta} = \frac{n-y}{1-\theta} \Leftrightarrow \theta = y/n$$

Only do the next two bullets if you have solved all other exercises (also Exercise 3):

- Differentiate $l'(\theta)$ to obtain $l''(\theta)$.

$$l''(\theta) = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}$$

- For the MLE it generally holds that the variance of $\hat{\theta}$ is approximately

$$\text{Var}(\hat{\theta}) \approx -1/l''(\hat{\theta}).$$

Verify by a direct computation that this in fact results in the estimated variance found in the text:

$$\hat{\theta}(1-\hat{\theta})/n = y(n-y)/n^3.$$

$$\ell''(\hat{\theta}) = -\frac{y}{(y/n)^2} - \frac{n-y}{(1-y/n)^2} = -\frac{n^2}{y} - \frac{n-y}{(n-y)^2/n^2} = -n^2(\frac{1}{y} + \frac{1}{(n-y)}) = -n^2\frac{n-y+y}{y(n-y)}$$

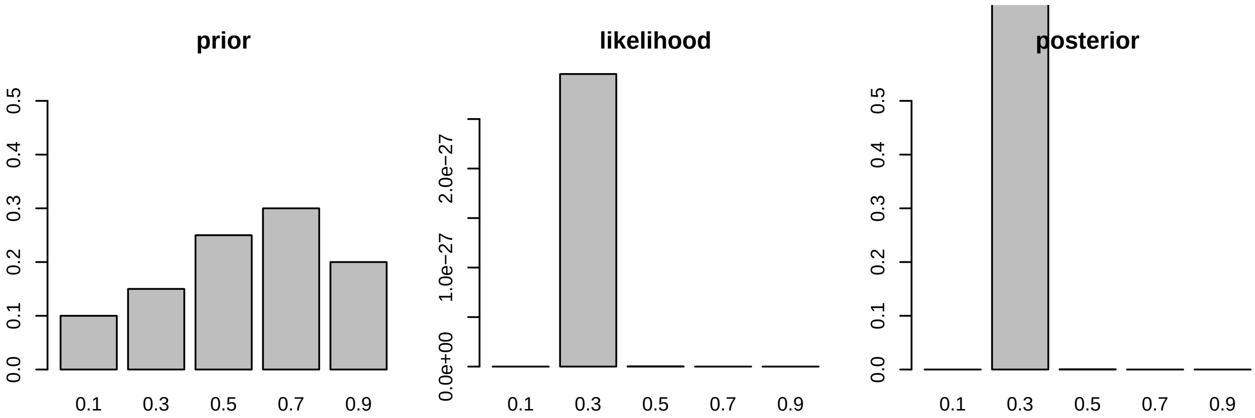
$$-1/l''(\hat{\theta}) = \frac{y(n-y)}{n^3}$$

Exercise 3 (solve by inserting code in the Rmd file)

For the Bayesian example with discrete prior:

1. Think about the effect data has on the posterior when compared to the prior.
2. Repeat the computations (mean and variance of posterior) and plots of the text but with $n = 100, y = 30$. Do the results surprise you?

```
theta <- c(.1, .3, .5, .7, .9)
prior <- c(0.10, 0.15, 0.25, 0.30, 0.20)
n <- 100; y <- 30
likval <- lik(theta, y, n)
posterior <- likval * prior
posterior <- posterior / sum( posterior )
par(mfrow=c(1,3), mar = c(3, 3, 3, 0.5))
barplot(prior, main="prior", names.arg=theta, ylim = c(0, .55))
barplot(likval, main="likelihood", names.arg=theta)
barplot(posterior, main="posterior", names.arg=theta, ylim = c(0, .55))
```




```
sum(theta * posterior) # Posterior mean closer to pure data driven estimate y/n
```

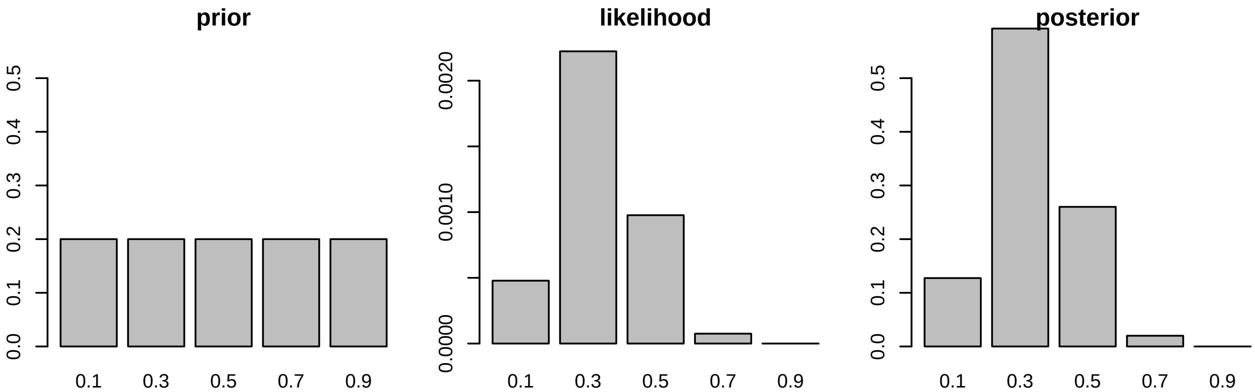
```
## [1] 0.3000889
```

```
sum(theta^2 * posterior) - sum(theta * posterior)^2 # Posterior variance decreases
```

```
## [1] 1.778936e-05
```

3. Repeat the computations and plots for the case where the prior has a uniform distribution (i.e. if all five values have prior probability 0.20), and $n = 10, y = 3$. What is the "relationship" between the posterior and the likelihood in this case? **The posterior and likelihood are now directly proportional:**

```
theta <- c(.1, .3, .5, .7, .9)
prior <- rep(.2, 5)
n <- 10; y <- 3
likval <- lik(theta, y, n)
posterior <- likval * prior
posterior <- posterior / sum( posterior )
par(mfrow=c(1,3), mar = c(3, 3, 3, 0.5))
barplot(prior, main="prior", names.arg=theta, ylim = c(0, .55))
barplot(likval, main="likelihood", names.arg=theta)
barplot(posterior, main="posterior", names.arg=theta, ylim = c(0, .55))
```



```r
sum(theta * posterior)
```

```
## [1] 0.334555
```

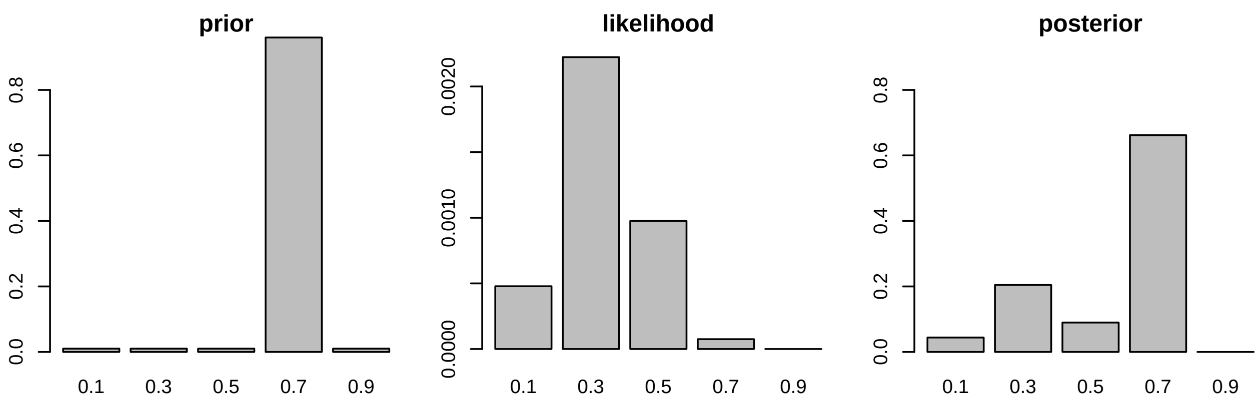
```r
sum(theta^2 * posterior) - sum(theta * posterior)^2
```

```
## [1] 0.01751453
```

4. Lastly, repeat the computations and plots for the case where $\pi(0.1) = \pi(0.3) = \pi(0.5) = \pi(0.9) = 0.01$ and $\pi(0.7) = 0.96$ (still $n = 10, y = 3$). Comment on the result.

Now the strong prior shifts the results away from the pure data driven results:

```r
theta <- c(.1, .3, .5, .7, .9)
prior <- rep(.01, 5)
prior[4] <- .96
n <- 10; y <- 3
likval <- lik(theta, y, n)
posterior <- likval * prior
posterior <- posterior / sum( posterior )
par(mfrow=c(1,3), mar = c(3, 3, 3, 0.5))
barplot(prior, main="prior", names.arg=theta, ylim = c(0, .9))
barplot(likval, main="likelihood", names.arg=theta)
barplot(posterior, main="posterior", names.arg=theta, ylim = c(0, .9))
```




```r
sum(theta * posterior)
```

```
## [1] 0.5739227
```

```r
sum(theta^2 * posterior) - sum(theta * posterior)^2
```

```
## [1] 0.03622129
```