# Bayesian Computation

*BIOS719 Generalized Linear Models*

## Binomial-Beta example

Bayes' theorem: $P(\theta|y) = \frac{P(y|\theta)P(\theta)}{\sum P(y|\theta)P(\theta)}$

Likelihood: $P(y|\pi) \sim Bin(n, \pi)$

Prior: $P(\pi) \sim Beta(\alpha, \beta)$

Posterior: $P(\pi|y) \sim Beta(y + \alpha, n - y + \beta)$

Beta Distribution:

$X \sim Beta(\alpha, \beta), x \in [0, 1]$

$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{(\alpha-1)}(1-x)^{\beta-1}$

$E(X) = \frac{\alpha}{\alpha+\beta}$

$Mode = \frac{\alpha-1}{\alpha+\beta-2}$ for $\alpha, \beta > 1$

$Variance = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

**Example** 7 successes out of 10 trials. No prior information about $\pi$.

Prior: $P(\pi) \sim Beta(1, 1) \equiv Unif(0, 1)$

Posterior: $P(\pi|y) \sim Beta(8, 4)$

MLE: $\frac{7}{10}$

Posterior mean: $E(\pi|y) = \frac{8}{8+4}$

Posterior mode: $\frac{7}{10}$

Calculate 95% credible interval

```r
qbeta(p=c(0.025, 1-0.025), shape1=8, shape2=4)
```

```
## [1] 0.3902574 0.8907366
```

Let's compare MLE and posterior distribution of $\pi$ at different data points. We observe that posterior means show shrinkage towards 0.5 as compared to MLE. Note that prior mean of $\pi$ is 0.5.

```r
temp <- cbind( (0:10), (0:10)/10, (1+(0:10))/(12))
colnames(temp) <- c("Data", "MLE", "Posterior Mean")
temp
```

```
##       Data MLE Posterior Mean
##  [1,]    0 0.0     0.08333333
##  [2,]    1 0.1     0.16666667
##  [3,]    2 0.2     0.25000000
##  [4,]    3 0.3     0.33333333
##  [5,]    4 0.4     0.41666667
##  [6,]    5 0.5     0.50000000
##  [7,]    6 0.6     0.58333333
##  [8,]    7 0.7     0.66666667
##  [9,]    8 0.8     0.75000000
```

```
## [10,]    9 0.9     0.83333333
## [11,]   10 1.0     0.91666667
```

Let's draw $\theta$ step by step.

```r
theta <- seq(0, 1, by=0.00001) ## A grid of theta (collection of theta)
theta.prior <- 1/length(theta)  ## Uniform distribution, P(theta)
lik <- dbinom(x=7, size=10, prob=theta)  ## Likelihood at each theta value P(Y|theta)
lik.theta.prior <- lik*theta.prior       ## P(Y|theta)P(theta)
theta.post <- lik.theta.prior/sum(lik.theta.prior)  ## P(Y|theta)P(theta) / sum[P(Y|theta)P(theta)]
theta.post.sample <- rbeta(100000, shape1=8, shape2=4)  ## theta sample from posterior distribution

out.tab <- cbind(theta=theta, theta.prior=theta.prior, lik=lik,
                 lik.theta.prior=lik.theta.prior, theta.post=theta.post)
head(out.tab)
```

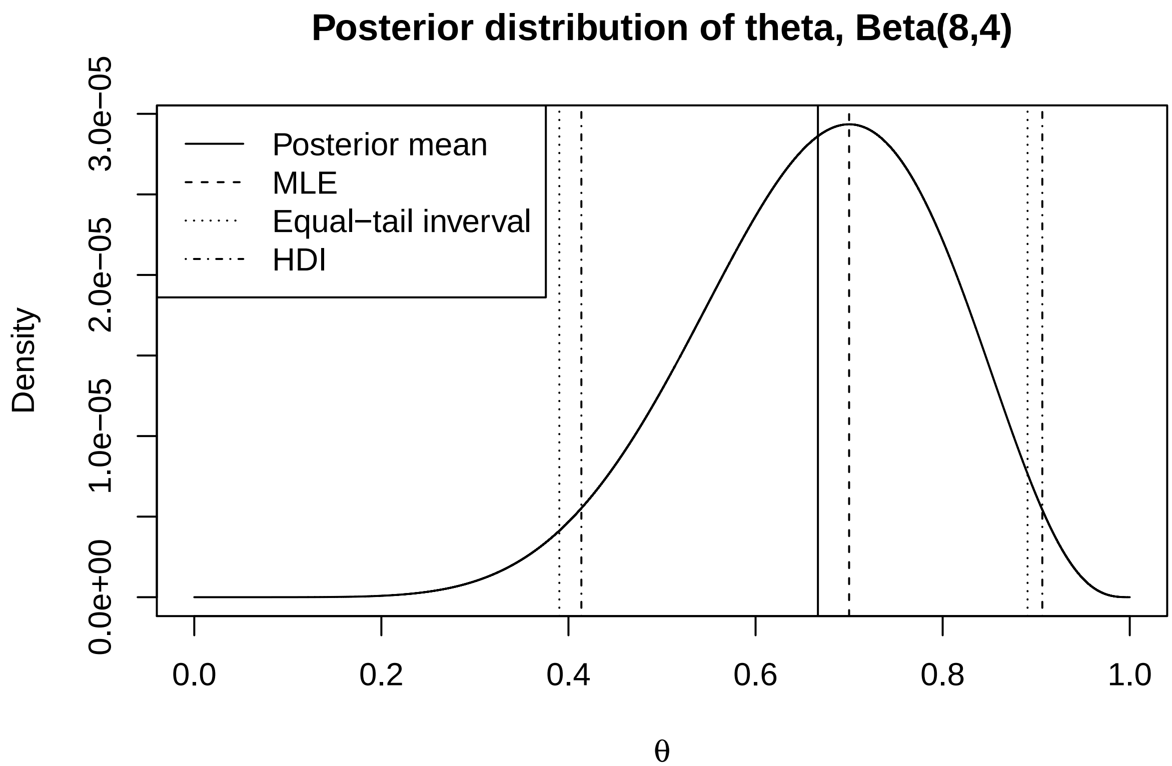
```
##      theta theta.prior          lik lik.theta.prior   theta.post
## [1,] 0e+00  9.9999e-06 0.000000e+00    0.000000e+00 0.000000e+00
## [2,] 1e-05  9.9999e-06 1.199964e-33    1.199952e-38 1.319960e-37
## [3,] 2e-05  9.9999e-06 1.535908e-31    1.535892e-36 1.689499e-35
## [4,] 3e-05  9.9999e-06 2.624164e-30    2.624138e-35 2.886580e-34
## [5,] 4e-05  9.9999e-06 1.965844e-29    1.965824e-34 2.162428e-33
## [6,] 5e-05  9.9999e-06 9.373594e-29    9.373500e-34 1.031095e-32
```

```r
out <- list(mean.prior=sum(theta*theta.prior),
            mean.post=sum(theta*theta.post),
            Equal.tail.interval1=
              c(max(theta[cumsum(theta.post)<=0.05/2]), min(theta[cumsum(theta.post)>1-0.05/2])),
            Equal.tail.interval2=
              c(quantile(theta.post.sample, probs=0.025), quantile(theta.post.sample, probs=0.975)),
            HDI=hdi(theta.post.sample, credMass=0.95)
            )
out
```

```
## $mean.prior
## [1] 0.5
##
## $mean.post
## [1] 0.6666667
##
## $Equal.tail.interval1
## [1] 0.39025 0.89074
##
## $Equal.tail.interval2
##      2.5%     97.5%
## 0.3902681 0.8906650
##
## $HDI
##     lower     upper
## 0.4137897 0.9065233
## attr(,"credMass")
## [1] 0.95
```

```r
plot(x=theta, y=theta.post, type="l", lwd=1,
     xlab=expression(theta), ylab="Density",
     main="Posterior distribution of theta, Beta(8,4)")
#lines(density(theta.post.sample), col=2)
abline(v=8/12)
abline(v=7/10, lty=2)
abline(v=out$Equal.tail.interval1, lty=3)
abline(v=out$HDI, lty=4)

legend("topleft", c("Posterior mean", "MLE", "Equal-tail inverval", "HDI"), lty=c(1,2,3,4))
```



Now, check SAS code and outputs.

```
data example;
input y n;
datalines;
7 10
;

proc mcmc data=example seed=13 nmc=20000 STATS=ALL;
parm p;
prior p ~ beta(1,1);
model y ~ binomial(n, p);
run;
```

| Parameters | | | | |
|---|---|---|---|---|
| Block | Parameter | Sampling Method | Initial Value | Prior Distribution |
| 1 | p | Conjugate | 0.5000 | beta(1,1) |

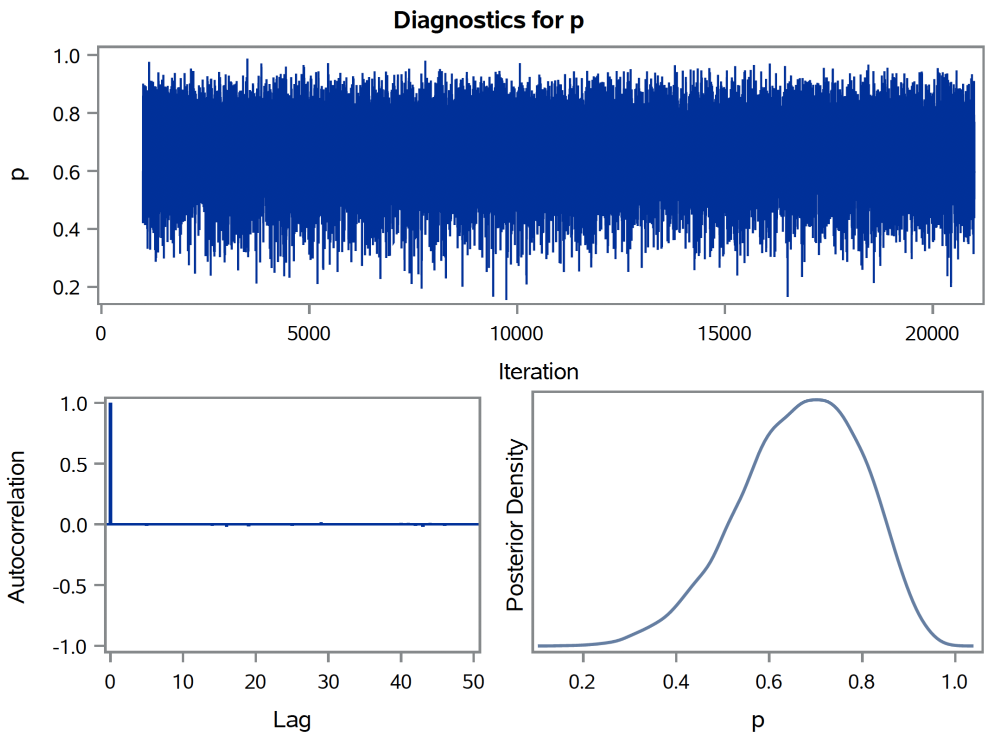
### The MCMC Procedure

| Posterior Summaries | | | | | | |
|---|---|---|---|---|---|---|
| | | | | Percentiles | | |
| Parameter | N | Mean | Standard Deviation | 25 | 50 | 75 |
| p | 20000 | 0.6669 | 0.1305 | 0.5800 | 0.6756 | 0.7642 |

| Posterior Intervals | | | | | |
|---|---|---|---|---|---|
| Parameter | Alpha | Equal-Tail Interval | | HPD Interval | |
| p | 0.050 | 0.3896 | 0.8888 | 0.4136 | 0.9053 |

| Posterior Correlation Matrix | |
|---|---|
| Parameter | p |
| p | 1.0000 |

| Posterior Covariance Matrix | |
|---|---|
| Parameter | p |
| p | 0.0170 |


Diagnostics for p
p
1.0
0.8
0.6
0.4
0.2
0
5000
10000
15000
20000
Iteration
Autocorrelation
1.0
0.5
0.0
-0.5
-1.0
0
10
20
30
40
50
Lag
Posterior Density
0.2
0.4
0.6
0.8
1.0
p