

Censored observations

Week2-ex1, solution

Exercise instructions

Suppose you have a $\text{Gamma}(\alpha = 1, \beta = 1)$ prior distribution on the parameter λ which corresponds to the expected number of ship ice besetting events (=events where a ship gets stuck in ice) during 1000 nautical miles in ice infested waters. The number of besetting events, y per distance d (nm) is modeled with a Poisson distribution $\text{Poisson}(\lambda \times d)$. The hyper-parameter α is the shape and β is the inverse scale parameter. You are told that during winters 2013-2017 category A ice breakers traveled in total 6560 nautical miles in the Kara Sea (a sea area in the Arctic Sea). Within this distance they experienced in total more than 2 but less than 7 ice besetting events.

- 1) Write down the equation for the posterior probability density for λ .
- 2) Discretize interval $[0, 3]$ into 100 equally spaced intervals, calculate the unnormalized posterior probability density at each of the discrete bins, normalize the densities to sum up to one and draw the posterior density at the discrete cells.
- 3) Using the discretized values of λ and their corresponding posterior density values
 - a) draw the posterior cumulative distribution function of λ and
 - b) calculate the posterior probability that $\lambda > 1$.
- 4) Draw the posterior probability density for λ in case where you are told that the exact number of besetting events is 6. What are the differences between the posterior densities from 2) and 4)?
- 5) Calculate the posterior predictive probability mass function for number of besetting events, \tilde{y} , within a 500 nm distance by using the posterior predictive density from step 2) and draw it. Note, you can restrict $\tilde{y} \in [0, 4]$

Model answer

1-3)

The posterior probability density function in case of censored observation is

$$p(\lambda | 2 < y < 7, d = 6.56) \propto (\text{Poisson}(3|\lambda \times d) + \text{Poisson}(4|\lambda \times d) + \text{Poisson}(5|\lambda \times d) + \text{Poisson}(6|\lambda \times d)) \text{Gamma}(\lambda | 1, 1)$$

When discretizing λ we can normalize the distribution as done below

```
# vectorize th into 100 bins
lambda = seq(0, 3, length=101)          # The end points of the intervals
lambda = (lambda[1:100]+lambda[2:101])/2 # the middle points of the intervals to be used in the calcul
d = 6.56

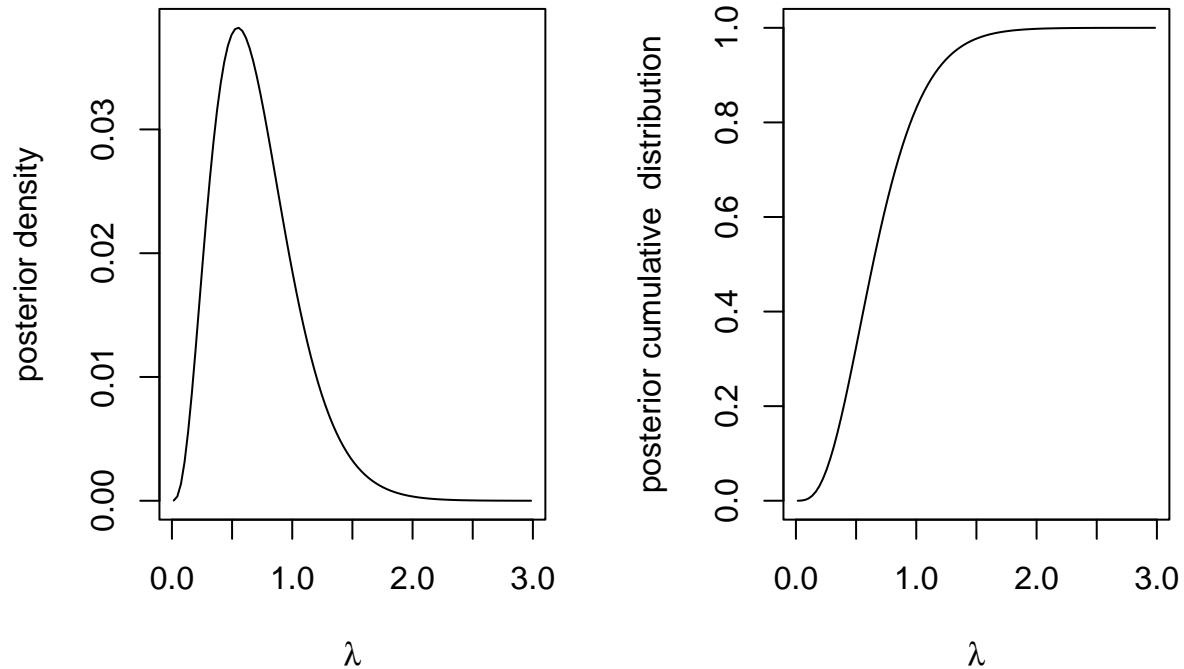
# calculate the unnormalized density at each bin
dens = (dpois(3,lambda*d)+dpois(4,lambda*d)+dpois(5,lambda*d)+dpois(6,lambda*d))*dgamma(lambda,1,1)
# normalize the discretized probability densities
dens=dens/sum(dens)
```

```

# calculate the cumulative distribution function
post_cdf = cumsum(dens)

# plot the posterior density
par(mfrow=c(1,2)) # divide plot into 2 subplots
plot (lambda, dens, type="l", xlab=expression(lambda), ylab="posterior density",cex=2)
# plot the posterior cumulative distribution function
plot (lambda, post_cdf, type="l", xlab=expression(lambda), ylab="posterior cumulative distribution",cex=2)

```



```

# calculate the probability that lambda > 1
1-max(post_cdf[which(lambda<=1)])

```

```
## [1] 0.1852897
```

4)

The posterior density function in case of $y = 6$ observation is

$$p(\lambda|y = 6, d = 6.56) \propto \text{Poisson}(6|\lambda \times d)\text{Gamma}(\lambda, 1, 1)$$

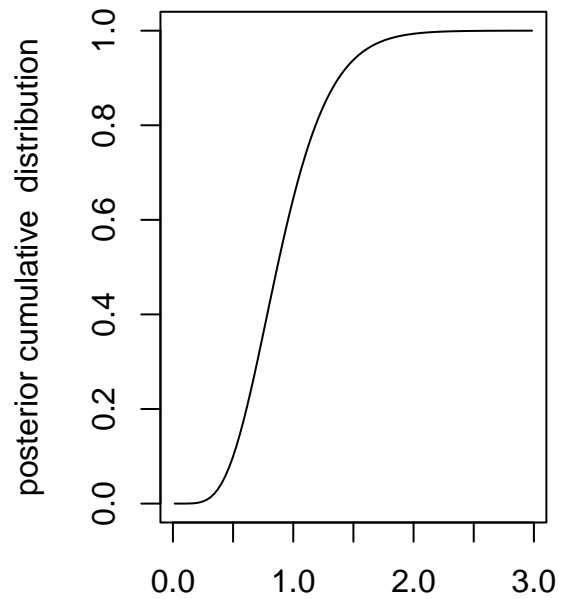
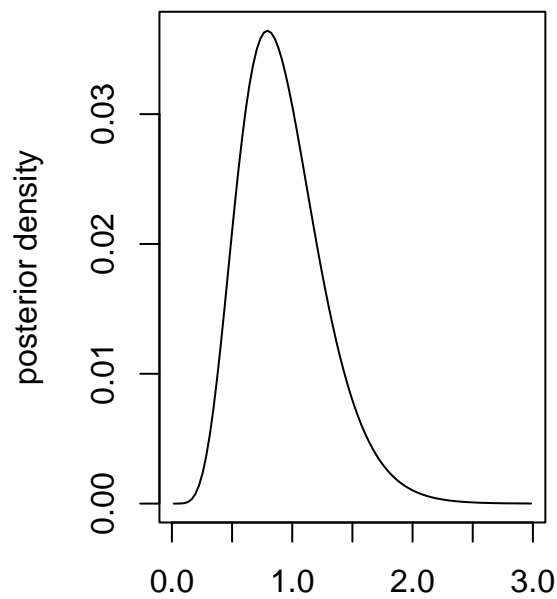
```

# calculate the density at each bin
dens2 = dpois(6,lambda*d)*dgamma(lambda,1,1)
dens2 = dens2 /sum(dens2)

# calculate the cumulative distribution function
post_cdf2 = cumsum(dens2)

# plot the unnormalized posterior
par(mfrow=c(1,2)) # divide plot into 2 subplots
plot (lambda, dens2, type="l", xlab="lambda", ylab="posterior density", cex=2)
plot (lambda, post_cdf2, type="l", xlab=expression(lambda), ylab="posterior cumulative distribution",cex=2)

```

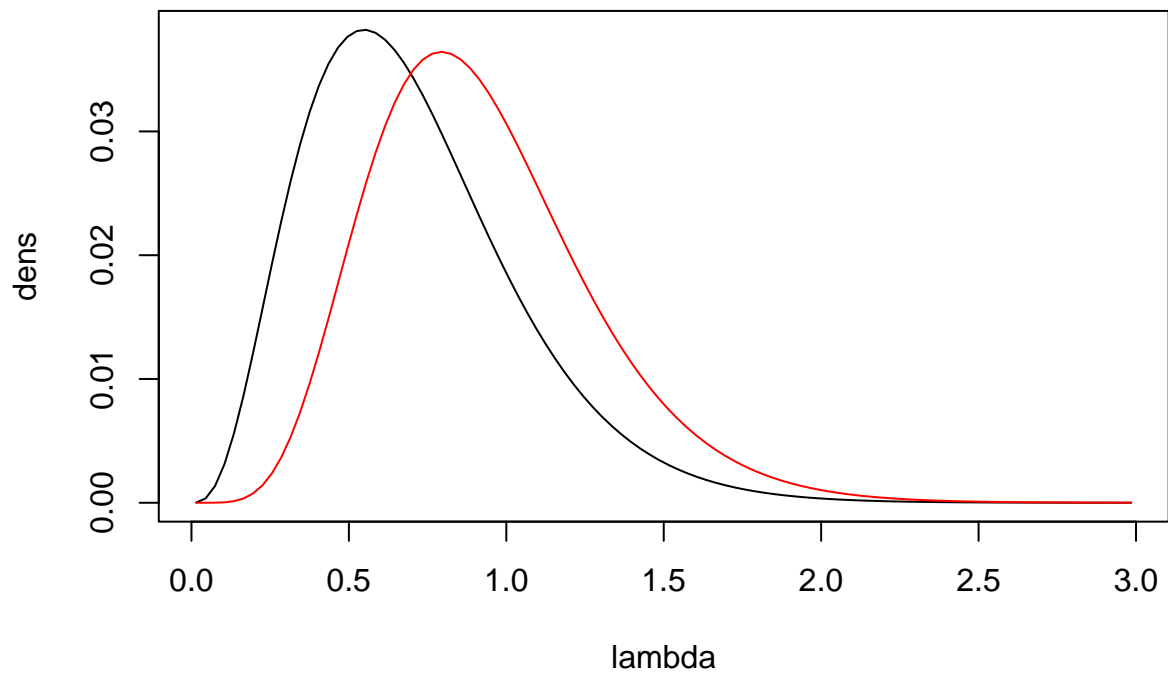


```
# calculate the probability that lambda > 1
1-max(post_cdf2[which(lambda<=1)])
```

```
## [1] 0.3802104
```

Let's compare the two distributions by plotting them together and calculating their central 95% posterior credible interval.

```
plot(lambda,dens, type="l")
lines(lambda,dens2, type="l", col="red")
```



```
min(lambda[ which(post_cdf>=0.975) ])-min(lambda[ which(post_cdf>=0.025) ])
```

```
## [1] 1.29
```

```
min(lambda[ which(post_cdf2>=0.975) ])-min(lambda[ which(post_cdf2>=0.025) ])
```

```
## [1] 1.35
```

The apparent differences are that the latter posterior probability density is centered more to the right than the first one and that the first one is narrower than the latter one.

5

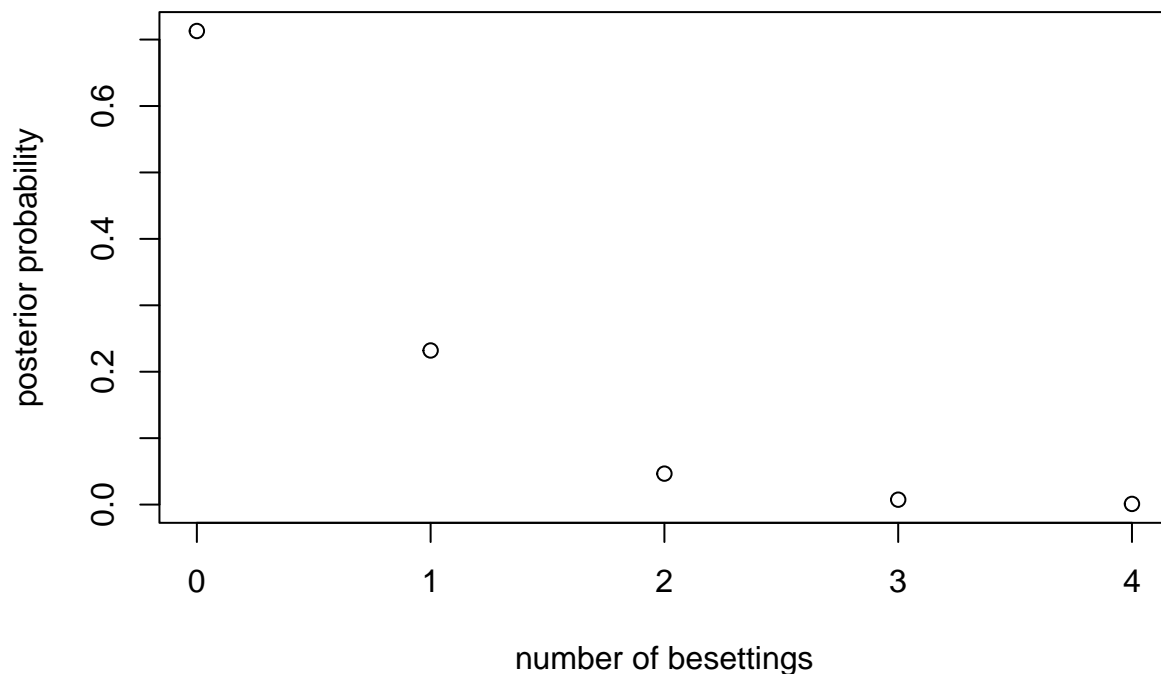
The posterior predictive probability for \tilde{y} is

$$p(\tilde{y}|2 < y < 7, d = 6.56, \tilde{d} = 0.5) \approx \sum_{i=1}^{100} \text{Poisson}(\tilde{y}|\lambda_i \times 0.5)p(\lambda_i|2 < y < 7, d = 6.56)$$

We need to now calculate this sum for each $\tilde{y} \in [0, 4]$.

```
y_tilde = c(0,1,2,3,4)
p_pred = vector(length=length(y_tilde))
for (i in 1:length(y_tilde)){
  p_pred[i] = sum(dpois(y_tilde[i],lambda*0.5)*dens)
}
plot(y_tilde,p_pred, main="posterior predictive distribution",
     ylab="posterior probability", xlab="number of besettings")
```

posterior predictive distribution



Grading

Total 10 points: Two points from correct answer in each step. One point per step if the main idea is correct but the result is erroneous because of a minor bug or misthought.