1. Let X be the ratio between a claim payment and insured capital. Let us assume that $X \sim beta(\alpha, \beta)$, i.e.

$$f_X(x \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha+1}(1-x)^{\beta+1},\ 0 < x < 1,\ \alpha, \beta > 0$$

Using software R we get (observed data in array x)

```
> length (x)
[1] 500
> mean(x)
[1] 0.6715272
> sd(x)
[1] 0.1775871
> summary(x); length(x); sd(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.1471  0.5576  0.6811  0.6715  0.8059  0.9945
> # first parameter is alpha
> # second parameter is beta
 > minusloglik_beta=function(theta){
+   -sum(dbeta(x,shape1=theta[1],shape2=theta[2],log=TRUE))
+   }
>
> theta.start=c(3,2.5)
> out=nlm(minusloglik_beta,theta.start,hessian=TRUE)
Warning messages:
1: In dbeta(x, shape1, shape2, log) : NaNs produced
2: In nlm(minusloglik_beta, theta.start, hessian = TRUE) :
  NA/Inf replaced by maximum positive value
3: In dbeta(x, shape1, shape2, log) : NaNs produced
4: In nlm(minusloglik_beta, theta.start, hessian = TRUE) :
  NA/Inf replaced by maximum positive value
> HH=out$hessian
> InvHH=solve(HH)
> InvHH_n=InvHH/500
> out
$minimum
[1] -181.6821
$estimate
[1] 3.963062 1.933768
$gradient
[1] -4.785322e-06  5.532537e-06
$hessian
         [,1]     [,2]
[1,]  51.01073 -92.3800
[2,] -92.38000 243.9896
$code
```

```
[1] 1
$iterations
[1] 13
> InvHH
           [,1]       [,2]
[1,] 0.06236915 0.02361438
[2,] 0.02361438 0.01303947
> InvHH_n
             [,1]         [,2]
[1,] 1.247383e-04 4.722875e-05
[2,] 4.722875e-05 2.607895e-05
```

a) What are the maximum likelihood estimates for α and β? Are these estimates reliable?
b) Determine a 95% confidence interval for β.
c) Using delta method, determine a 95% interval for $\mu = \alpha / (\alpha + \beta)$
d) Without using delta method determine another 95% confidence interval for μ.

2. Assume that X is gamma distributed with unknown parameters α and θ. From this population we collected data set B as a sample.
Using software R we got

```
> x=c(27,82,115,126,155,161,243,294,340,384,457,680,855,877,974,
+     1193,1340,1884,2558,15743)
> minusLogLikGamma=function(param){
+   -sum(dgamma(x,shape=param[1],scale=param[2],log=TRUE))
+   }
> param.start=c(1,1000)
> out=nlm(minusLogLikGamma,param.start,hessian=TRUE)
Warning messages:
1: In dgamma(x, shape, scale, log) : NaNs produced
2: In nlm(minusLogLikGamma, param.start, hessian = TRUE) :
  NA/Inf replaced by maximum positive value
> out
$minimum
[1] 162.2934
$estimate
[1]    0.556156 2561.146495
$gradient
[1] -8.273560e-05 -6.824815e-09
$hessian
             [,1]         [,2]
[1,] 82.443074234 7.808613e-03
[2,]  0.007808613 1.695064e-06
```

```
$code
[1] 1
$iterations
[1] 26
> solve(out$hessian)
            [,1]        [,2]
[1,]  0.02151866     -99.12953
[2,] -99.12952956 1046606.28099
```

a) What are the maximum likelihood estimates for α and θ?
b) Determine a 95% confidence interval for each parameter.
c) Using delta method, determine a 95% interval for $\mu = E(X)$

3. To get a maximum likelihood estimate for the parameter θ we defined the following function using R

```
>minusloglik_f=function(theta){
  -sum(log(x)-2*log(theta)-x/theta)
 >}
```

Write the density function of the observed data.

4. Assume that X is beta distributed with unknown parameters α and β. From this population we collected a random sample (see file exer4.csv).
   a. What is the sample average?
   b. Plot an histogram of the observed sample
   c. What are the maximum likelihood estimates for α and β?
   d. Determine a 95% confidence interval for each parameter.

5. Assume that we observed the random sample presented in file exer5.csv.
   a. Plot an histogram of the observed sample
   b. Assuming a Weibull distribution for the population, compute the maximum likelihood estimate for θ and τ? Obtain a maximum likelihood estimate for $P(X > 100)$.
   c. Assuming an inverse Weibull distribution for the population, compute the maximum likelihood estimate for θ and τ? Obtain a maximum likelihood estimate for $P(X > 100)$.
   d. Obtain a non-parametric estimate for $P(X > 100)$ and compute a 95% confidence interval for this probability.

6. A random sample of 400 claim amounts (thousands of euros)originates the following results:

Claim size	$(0,0.25]$	$(0.25,0.5]$	$(0.5,1.0]$	$(1.0,2.0]$	$(2.0,5.0]$	$(5.0,10]$	Over 10
Nº claims	20	40	110	130	60	30	10

a. Assuming a gamma distribution for the population, compute the maximum likelihood estimate for α and θ? Obtain a maximum likelihood estimate for $P(X > 3)$.

b. Compute a non-parametric estimate for $P(X > 3)$.

7. Repeat example 15.17 using improper prior $\pi(\alpha) = 1/\alpha$, $\alpha > 0$

8. Repeat example 15.17 using the discrete prior $\pi(\alpha) = 1/3$, $\alpha = 1, 2, 3$.

9. Assume that, given θ, X follows a Bernoulli distribution with parameter θ. Using a Bayesian approach obtain, for each case, the posterior distribution for θ and compute the mean and the variance of the posterior distribution. Answer each question assuming that we observed : (i) a sample with size $n = 10$ where $\bar{x} = 0.1$; (ii) a sample with size $n = 100$ where $\bar{x} = 0.1$; (iii) a sample with size $n = 1000$ where $\bar{x} = 0.1$.

   a. $\pi(\theta) = \begin{cases} 0.3 & \theta = 0.05 \\ 0.5 & \theta = 0.1 \\ 0.2 & \theta = 0.2 \end{cases}$

   b. $\pi(\theta) = 1/3$, $\theta = 1, 2, 3$. Compare with the results obtained using the previous prior.

   c. $\pi(\theta) = 30\theta^2 (1-\theta)^2$, $0 < \theta < 1$.

   d. $\pi(\theta) = 1$, $0 < \theta < 1$. Compare with the results obtained using the previous prior.

Solutions:

1. a) 3.963; 1.934 b) (1.710; 2.158) c) (0.6565; 0.6876) d) (0.6560; 0.6871)
2. a) 0.5562; 2561.1 b) (0.2686; 0.8437) (555.99; 4566.3) c) (587.14; 2261.7)
3. deliberately without solution
4. deliberately without solution
5. deliberately without solution
6. deliberately without solution
7. Gamma with parameters 2 and θ
8. Posterior: Gamma with parameters 9 and 0.2631
9. Posterior (0.0187 for $\alpha = 1$, 0.4288 for $\alpha = 2$,0.5525 for $\alpha = 3$)