TMA4300 Computer Intensive Statistical Methods
Interactive lecture 2, Spring 2018

Problem A: Rejection sampling for the beta distribution

In this problem we will consider how to make samples from a beta distribution. The density function of the beta distribution is given by

$$f(x;\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } x \in [0,1], \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where the parameters $\alpha, \beta > 0$.

1. Familiarise yourself with the density of the beta distribution by plotting it for different values of α and β.

2. We want to use rejection sampling to generate samples from the beta distribution with parameters α and $\beta \geq 1$. As proposal distribution we simply use a uniform distribution on the unit interval $[0,1]$. Find a formula for the acceptance probability as function of the proposed value x and the parameters α and β. For what values of α and β can this rejection sampling algorithm be used? [*Remember that to find for which value a function has its maximum value, it may give easier calculations if you in stead consider the logarithm of the function of interest.*]

3. Find also a formula for the average number of proposals needed to generate one realisation.

4. Implement the rejection sampling algorithm developed above. Check your calculations and implementation by comparing realisations with known theoretical properties of the beta distribution. Moreover, use also your simulations to check the formula you found in 3.

Problem B: Sampling using known relations

1. The probability mass function and cumulative distribution function of the geometric distribution is
$$f(x;p) = p(1-p)^{x-1} \quad \text{and} \quad F(x;p) = 1-(1-p)^x \tag{2}$$
for $x = 1, 2, \ldots$, respectively. Using the general technique for simulating from a discrete distribution, write an R function generating a realisation (or n independent realisations) from a geometric distribution with parameter p.

2. Using the situation leading to the geometric distribution, write an R function generating a realisation (or n independent realisations) from a geometric distribution with parameter p. Compare the efficiency of your two functions simulating from the geometric distribution.

3. Use how the Poisson distribution is defined from a Poisson process and its relation to the exponential distribution to write an R function simulating n independent realisations from a Poisson distribution with parameter λ.

4. The density of a gamma distribution with parameters $\alpha = k$ and $\beta = \lambda$ is
$$f(x;n) = \begin{cases} \frac{1}{\lambda^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\lambda}} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$
Use that the time to event number k in a Poisson process with intensity λ is gamma distributed with parameters $\alpha = k$ and $\beta = \lambda$ to write an R function which returns n independent realisations this distribution when k is an integer.

Problem C: More rejection sampling for the beta distribution

Consider again the beta distribution with density function

$$f(x;\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } x \in [0,1], \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In the following we consider the case $\alpha, \beta \in (0,1)$ and use the proposal distribution

$$g(x;\alpha,\beta,\gamma) = \begin{cases} \gamma\alpha x^{\alpha-1} + (1-\gamma)\beta(1-x)^{\beta-1} & \text{for } x \in [0,1], \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where $\gamma \in [0,1]$ is a parameter we can choose.

1. Familiarise yourself with $g(x;\alpha,\beta,\gamma)$ by plotting it up for some sets of parameter values. Plot also $f(x;\alpha,\beta)$ in the same plot so that you can compare $f(x;\alpha,\beta)$ and $f(x;\alpha,\beta,\gamma)$.

2. Find a formula for the acceptance probability as function of the proposed value x and the parameters α, β and γ. [*Hint: Start by plotting up $f(x;\alpha,\beta)/g(x;\alpha,\beta,\gamma)$ for some values of the parameters to get a rough idea of how this ratio behaves as function of x. Thereafter find the necessary analytical formulas.*]

3. Find also a formula for the average number of proposals needed to generate one realisation.

4. Assuming the beta distribution to be given, i.e. the values of α and β are fixed, find the optimal value for γ.

5. Implement this new rejection sampling algorithm. [*Hint: to see how to sample from $g(x,\alpha,\beta,\gamma)$ note that it is a mixture distribution, $g(x;\alpha,\beta,\gamma) = \gamma g_0(x;\alpha) + (1-\gamma)g_1(x;\beta)$.*] Again check your calculations and implementation by comparing realisations with known theoretical properties.

6. What proposal distribution would you have used if $\alpha > 1$ and $\beta < 1$, or if $\alpha < 1$ and $\beta > 1$. How would you have sampled from the beta distribution if $\alpha = 1$ and $\beta \neq 1$, or if $\alpha \neq 1$ and $\beta = 1$?