TMA4300 Computer Intensive Statistical Methods

Interactive lecture 1, Spring 2018

Problem A: Probability integral transform

The Weibull distribution with scale parameter $\alpha > 0$ and shape parameter $\beta > 0$ has density function

$$f(x;\alpha,\beta) = \begin{cases} \alpha\beta x^{\beta-1}e^{-\alpha x^\beta} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

and cumulative distribution distribution function

$$F(x;\alpha,\beta) = \begin{cases} 1 - e^{-\alpha x^\beta} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Note that when $\beta = 1$ and $\alpha = \lambda$ the Weibull distribution becomes an exponential distribution with parameter λ.

1. Assuming you know how to generate realisations from a uniform distribution on the unit interval $[0,1]$, use the probability integral transform (inversion) method to find how to generate a realisation from a Weibull distribution with parameters α and β.

2. Write an R function which returns a vector of n independent realisations from the Weibull distribution with parameters α and β. The function should have three input parameters, n, α and β.

3. Use your R function to generate a vector of independent realisations from the Weibull distribution. Check your calculations in 1 and your implementation in 2 by comparing your realisations with known theoretical properties of the Weibull distribution.

4. Extra: Generalise your implementation in 2 so that the R function can generate independent realisations from different Weibull distribution. The R function should then allow the input parameters α and β to be vectors, but should also work properly when they are scalar. If the length of one or both of the input vectors α and β and smaller than n, the values in the input vector(s) should be recycled. This is what is implemented in standard R function for generation of random numbers. **Note**: dependent on how you implemented the algorithm your function may already work like this, try! You may also implement default values for the input parameters n, α and β. Again you should check your implementation by comparing simulation results with know theoretical properties.

Problem B: Box-Muller transform

Let $X \sim N(0,1)$ and $Y \sim N(0,1)$ be independent. Define

$$R = \sqrt{X^2 + Y^2} \quad \text{and} \quad \Theta = \tan^{-1}\left(\frac{Y}{X}\right). \tag{3}$$

1. Visualise the transformation from (X, Y) to (R, Θ) in a plot.
2. Use the multivariate change-of-variables formula to find the joint density function of R and Θ. What distribution is this?
3. Assuming you know how to generate realisations from a uniform distribution on the unit interval $[0, 1]$, how can you generate realisations of R and Θ? When you know how to generate realisations of R and Θ, how can you use this to generate realisations from the standard normal distribution. In turn, how can you generate realisations from a normal distribution with mean μ and standard deviation σ?
4. Write an R function which returns n independent realisations from the normal distribution. The function should have three input parameters, n, μ and σ. Check your calculations and implementation by comparing simulated values with know theoretical properties of the normal distribution.

Problem C: **Some discrete distributions**

1. The probability mass function and cumulative distribution function of the geometric distribution is
$$f(x;p) = p(1-p)^{x-1} \quad \text{and} \quad F(x;p) = 1-(1-p)^x \tag{4}$$
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 $\beta = 1$ is
$$f(x;n) = \begin{cases} \frac{1}{\Gamma(k)} x^{k-1} e^{-x} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$
Use that the time to event number k in a Poisson process with intensity λ is gamma distributed with parameters $\alpha = k$ and $\beta = 1$ to write an R function which returns n independent realisations from a gamma distribution with parameters $\alpha = k$ and $\beta = 1$, where k is an integer. How can you then simulate from a gamma distribution where $\alpha = k$ still is an integer, but where β may be any value larger than zero.