Computer Practical: Transformation Methods and Rejection Sampling

In this computer practical, you can use either Python or R. There is a list of useful Python and R functions at the end of this handout.

Theoretical background: exponential distribution and gamma distribution

This computer practical focuses on sampling from the exponential distribution and from the gamma distribution. The density of the exponential distribution is for $x > 0$ ($\lambda > 0$)

$$f_{(\lambda)}(x) = \lambda \exp(-\lambda x),$$

the corresponding c.d.f. is

$$F_{(\lambda)}(x) = 1 - \exp(-\lambda x).$$

The density of the gamma distribution is for $x > 0$ ($\alpha, \lambda > 0$):

$$f_{(\alpha,\lambda)}(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^{\alpha} \exp(-\lambda x)$$

The exponential distribution is a special case of the gamma distribution with $\alpha = 1$. Furthermore, if $X_i \overset{\text{i.i.d.}}{\sim} \mathsf{Expo}(\lambda)$ and

$$Y = X_1 + \ldots + X_k,$$

then $Y \sim \mathsf{Gamma}(k, \lambda)$.

Sampling from the exponential distribution

We have seen in the lectures that we can sample from the exponential distribution using the inversion method. The c.d.f. of the $\mathsf{Expo}(\lambda)$ distribution is for $x > 0$

$$F_{(\lambda)}(x) = 1 - \exp(-\lambda x).$$

The (generalised) inverse is thus

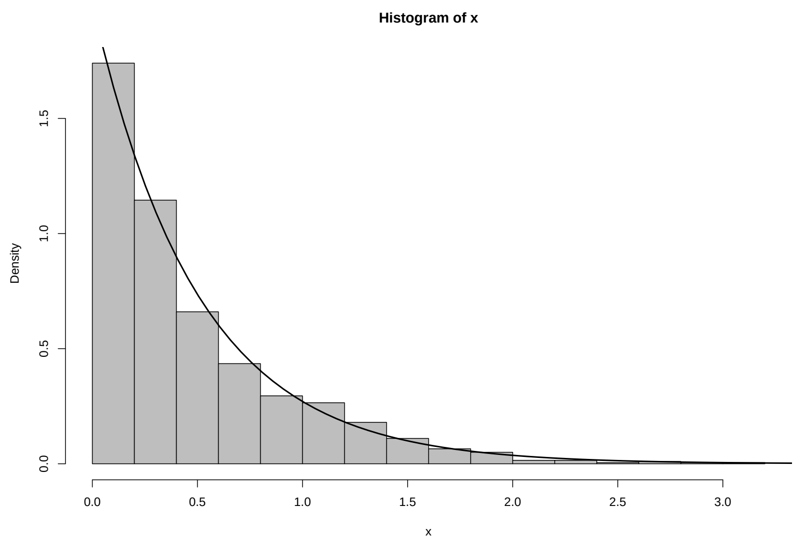
$$F^{-}_{(\lambda)}(u) = F^{-1}_{(\lambda)}(u) = -\log(1-u)/\lambda$$

Thus if $U \sim \mathsf{U}[0,1]$, then $-\log(1-U)/\lambda \sim \mathsf{Expo}(\lambda)$, and also

$$-\log(U)/\lambda \sim \mathsf{Expo}(\lambda).$$

Task 1. Write a function `sampleExpo` which draws a sample of size n from the $\mathsf{Expo}(\lambda)$ distribution using the inversion method. Your function should take λ and n as arguments.

Task 2. To check your code from task 1, generate a sample of size $n = 1000$ from the $\mathsf{Expo}(\lambda)$ distribution with $\lambda = 2$, and create a histogram of it. Superimpose the density function $f_{(\lambda)}(x) = \lambda \exp(-\lambda x)$. Your figure should look like the one below.



Sampling from the gamma distribution (integer shape parameter)

If we want to draw one realisation from the Gamma(k, λ) distribution with $k \in \mathbb{N}$, we can exploit the fact that if $X_1, \ldots, X_k \overset{\text{i.i.d.}}{\sim}$ Expo(λ), then

$$Y = X_1 + \ldots + X_k \sim \text{Gamma}(k, \lambda)$$

Task 3. Write a function `sampleGammaInteger`, which draws one realisation from the Gamma(k, λ) distribution by summing up k exponentially distributed random numbers (generated using your function `sampleExpo`). Your function should take λ and k as arguments.

Sampling from the gamma distribution (non-integer shape parameter)

The method of drawing from the gamma distribution by summing up exponentials only works for integer k. If we want to sample from the Gamma(α, λ) distribution with $\alpha \geq 1$ and $\lambda > 1$ we can use rejection sampling using a Gamma$(k, \lambda - 1)$ distribution with $k = \lfloor \alpha \rfloor$ as instrumental distribution. Then we have for the ratio that

$$\frac{f_{(\alpha,\lambda)}(x)}{f_{(k,\lambda-1)}(x)} = \frac{x^{\alpha-1}\lambda^{\alpha}\exp(-\lambda x)/\Gamma(\alpha)}{x^{k-1}(\lambda-1)^k \exp(-(\lambda-1)x)/\Gamma(k)} = \frac{\Gamma(k)\lambda^{\alpha}}{\Gamma(\alpha)(\lambda-1)^k} x^{\alpha-k}\exp(-x)$$

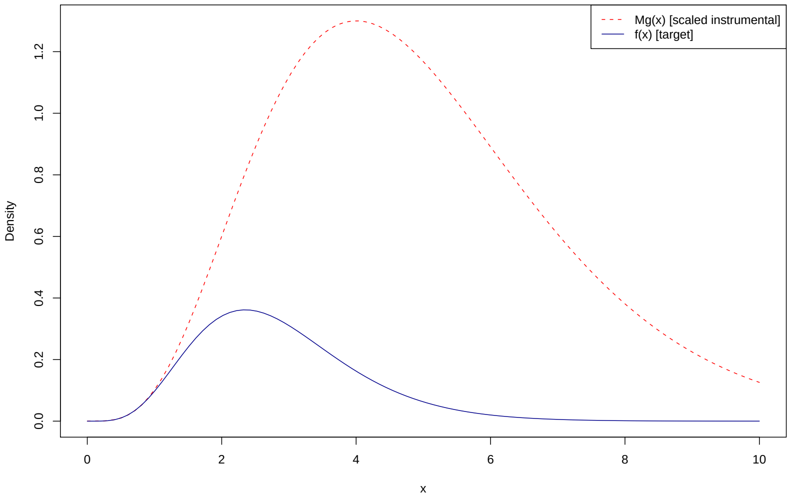
The ratio attains its maximum of $M = \frac{\Gamma(k)\lambda^{\alpha}}{\Gamma(\alpha)(\lambda-1)^k}(\alpha-k)^{\alpha-k}\exp(k-\alpha)$ at $x = \alpha - k$. Thus

$$M = \frac{f_{(\alpha,\lambda)}(\alpha-k)}{f_{(k,\lambda-1)}(\alpha-k)}.$$

Task 4. In this task we look at an example to check the M we have derived above.

(a) For $\alpha = 5.7$ and $\lambda = 2$ compute M.

(b) For $\alpha = 5.7$ and $\lambda = 2$ create a figure containing both $f_{(\alpha,\lambda)}(x)$ (target) and $M \cdot f_{(k,\lambda-1)}(x)$ (instrumental distribution) for $x \in (0, 10)$. The two curves should touch at $\alpha - k = 5.7 - 5 = 0.7$, as shown in the plot below.



Task 5. Write a function `sampleGamma` which samples from the Gamma(α, λ) distribution using rejection sampling using the Gamma$(k, \lambda - 1)$ distribution as instrumental distribution. Your function should take α and λ as arguments and should implement the following steps:

i. Set $k = \lfloor \alpha \rfloor$.

ii. Compute M.

iii. Repeat until a value is accepted …

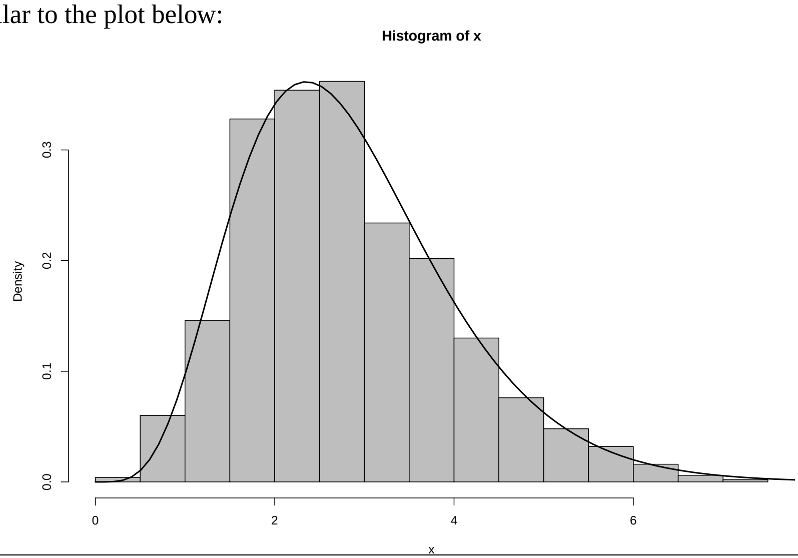
   α. Draw one realisation X from the Gamma$(k, \lambda - 1)$ distribution using the function `sampleGammaInt`.

   β. Compute the probability of acceptance

$$\frac{f_{(\alpha,\lambda)}(X)}{M \cdot f_{(k,\lambda-1)}(X)}$$

   γ. With this probability return X as accepted value, otherwise go back to step α.

Task 6. Check your code from task 5 by drawing a sample of size 1000 from the Gamma(α, λ) distribution for $\alpha = 5.7$ and $\lambda = 2$. Draw a histogram of your sample and superimpose the density of the Gamma(α, λ) distribution. Your plot should look similar to the plot below:



Useful Python functions

The functions listed below are available if you have imported the following libraries:

```
from scipy import *
from scipy import stats
from scipy import random
import pylab
```

The following Python functions might be of use:

random.random_sample The function `random.random_sample(`*n*`)` draws a sample of *n* pseudo-random numbers from the U$[0,1]$ distribution.

```
x = random.random_sample(10)
print x
```

```
array([ 0.32928742,  0.45686706,  0.99910906,  0.98468631,  0.215091  ,
        0.37646124,  0.25244065,  0.86446197,  0.87439556,  0.173766  ])
```

pylab.hist The function `pylab.hist(`*x*`, normed=True)` draws a histogram of the data in *x* using relative frequencies.

```
pylab.hist(x, normed=True)
pylab.show()
```

floor The function `floor(`*x*`)` find the largest integer not larger than *x*.

```
floor(3.9)
```

```
3
```

Gamma density Use the function below to evaluate the density $f_{(\alpha,\lambda)}$ of the Gamma(α,λ) distribution.

```
def gammaPDF(x, alpha, lam):
    return stats.gamma(alpha).pdf(x*lam)*lam
```

Useful R functions

The following Python functions might be of use:

runif The function `runif(`*n*`)` draws a sample of *n* pseudo-random numbers from the U$[0,1]$ distribution.

```
x <- runif(10)
x
```

```
 [1] 0.8331899 0.4151701 0.4832408 0.1252134 0.1989751 0.1101298 0.1448359
 [8] 0.3183492 0.5570466 0.2372144
```

hist The function `hist(`*x*`, freq=FALSE)` draws a histogram of the data in *x* using relative frequencies.

```
hist(x, freq=FALSE)
```

floor The function `floor(`*x*`)` find the largest integer not larger than *x*.

```
floor(3.9)
```

```
[1] 3
```

dgamma The function `dgamma(`*x*`,` *alpha*`,` *lambda*`)` evaluates the density $f_{(\alpha,\lambda)}$ of the Gamma(α,λ) distribution.