Lab3 - Generating Random Variables

Martina Martini s306163

Politecnico di Torino

Assignment Write a function that generates R.V.s distributed as follows:

- Rayleigh(sigma)
- Lognormal(mu, sigma^2)
- Beta(alpha, beta) distributed random variable, with alpha ¿ 1 and beta ¿ 1,
- Chi-square (n) n¿=1
- Rice distribution (nu, sigma) for nu¿=0 sigma¿=0

For one of the previously listed R.V.s, test your generator by evaluating the empirical first two moments you obtain after n in 100,10000,100000 random extractions, and comparing them with analytical predictions. In addition, compare also the empirical Cdf/pdf to the analytical one.

---

1 Rayleigh distribution

The Rayleigh distribution is a continuous probability distribution for nonnegative-valued random variables. In order to generate such a distribution, it was used the inverse-transform technique. This approach consists in the following operating phases:

Step 1 Generate a sequence of random variables uniformly distributed between 0 and 1.

Step 2 Compute the inverse of the cumulative distribution function (CDF now on) of the desired distribution.

Step 3 Apply such an inverse function to the uniform random variables to obtain the probability density function (PDF now on) of the desired distribution.

Following these guidelines, the Rayleigh random variables were generated. In particular, the over-cited inverse function is represented by the following formula:

$$\sqrt{2\sigma^2 \cdot ln(\frac{1}{1-x})}$$

To evaluate the goodness of the generated Rayleigh distributed R.V.s, the analytical PDF is displayed in Fig. 5. As you can notice, they are perfectly plotted in the case of $n = 10\_000$ and $n = 100\_000$, but not exactly in the case of $n = 100$.

Finally, the first and the second moments are computed for evaluation purposes. The expected value is computed as

$$\sigma\sqrt{\frac{\pi}{2}}$$

while the variance as

$$\frac{4-\pi}{2} \cdot \sigma^2$$

Since they are very close, you can assume that this generating method is efficient for this type of distribution.

2 Lognormal distribution

The log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. In order to generate such a distribution the following approach was followed:

Step 1 Generate two sequences of random variables uniformly distributed U and V.

Step 2 Compute $B = \sqrt{-2ln(u)}$ for each R.V. u in U.

Step 3 Compute $\theta = 2\pi v$ for each R.V. v in V.

Step 4 Compute the standard normally distributed random variables z's such as the union of the polar coordinates $Z_1 = B \cdot cos(\theta)$ for each θ, B in the first half of the array and $Z_2 = B \cdot sin(\theta)$ for each θ, B in the second half.

Step 5 Generate the normally distributed random values by applying $\mu + \sigma z$ for each z in Z.

Step 6 Generate the log-normal R.V.s by taking the exponential of each normal values.

To evaluate the goodness of this generating function, the log-normal PDF is displayed (Fig. 6). Also, the first and the second moments are computed. In particular, the expected value is equal to

$$e^{\mu+\frac{1}{2}\sigma^2}$$

while the variance is computed as

$$e^{2\mu+\sigma^2} \cdot (e^{\sigma^2} - 1)$$

Since both the evaluation metrics respect the analytical parameters, you can assume that this generating function is valid, even if in the case of $n = 100$, it is not working as expected.

3 Beta distribution

The Beta distribution is a family of continuous probability distributions that depend on two positive parameters α and β. For such a distribution, the Acceptance/Rejection method was used. It consists in the following phases:

Step 1 Generate two sequences on random variables X and Y such that the R.V.s in X are distributed uniformly between 0 and 1, while the ones in Y are inside an interval between 0 and a chosen k .

Step 2 Apply the PDF of the Beta distribution to the x's in X with the random chosen a and b as parameters.

Step 3 Check if the i-th generated variable is greater than the i-th corresponding y: if the function applied to the x is greater than y, then x is returned. This step is done for each value in X.

Note that the previous-cited k depends on the chosen probability to accept $perc$ (input parameter) since they are connected by the following relation: $perc = \frac{1}{k}$.

To evaluate this generating function, the analytical PDF of the Beta distribution is plotted and the first and the second moments are computed. In particular, we define the expected value as

$$\frac{a}{a+b}$$

and the variance as

$$\frac{ab}{(a+b)^2 \cdot (a+b+1)}$$

As you notice by looking at the plots in Fig. 7 , both in the cases of $n = 10\_000$ and $n = 100\_000$ the histogram is well shaped by the analytical PDF of the Beta distribution, but in case of $n = 100$ some issues come up. However mean and variance are well approximated, so we can conclude that also this generating function work well.

4 Chi squared distribution

The Chi-squared distribution, given k degrees of freedom can be interpreted as the distribution of a sum of the squares of k random variables distributed as independent standard normal variables. Moreover, the Chi-squared distribution is a special case of the Gamma distribution. For such a distribution, the following phases are followed:

Step 1 Generate two arrays of a certain number of sequences of uniformly distributed R.V.s U and V. Each of them is constructed with a chosen length (in the code dof variable is used for simplicity, but it can be modified). The total number of generated sequences is equal to the size.

Step 2 Compute the Bs by using the formula $B = \sqrt{-2ln(u)}$ for each r.v. u in U.

Step 3 Compute the thetas by using the formula $\theta = 2\pi v$ for each r.v. v in V.

Step 4 Compute the stardard normally distributed Zs by using the formula for the R.V.s z's such as the union of the polar coordinates $Z_1 = B \cdot cos(\theta)$ for each θ, B in the first half of the array and $Z_2 = B \cdot sin(\theta)$ for each θ, B in the second half.

Step 5 Elevate to the square these normally distributed R.V.s and sum up them sequence by sequence.

To evaluate the goodness of the function, the analytical PDF of the Chi-squared distributed is plotted as in Fig. 8. Moreover, the first and the second momentts are computed as follows:

$$E(X) = k$$

$$var(X) = 2k$$

where k represents the chosen (randomly in the code) degrees of freedom.

Comparing them, you can notice that this generating function works well, as expected, even if it is not precise for low size as $n = 100$.

5 Rice distribution

The rice distribution is the probability distribution of the magnitude of a bivariate normal random variable. Given that, the distance between the reference point and the center of the bivariate distribution ν and the scale σ are chosen randomly for better evaluations. The following steps are presented in the code:

Step 1 Generate two sequences of uniformly distributed R.V.s U and V with sizes equal to the actual one.

Step 2 Compute the Bs as $B = \sqrt{-2ln(u)}$

for each r.v. u in U.

Step 3 Compute the thetas as $\theta = 2\pi v$ for each r.v. v in V.

Step 4 Compute the z's such as the union of the polar coordinates $Z_1 = B \cdot cos(\theta)$ for each θ, B in the first half of the array and $Z_2 = B \cdot sin(\theta)$ for each θ, B in the second half.

Step 5 Generate the normally distributed R.V.s as $X \sim N(\nu cos(\theta), \sigma^2)$ and $Y \sim N(\nu sin(\theta), \sigma^2)$ by using the formula $X = \nu cos(\theta) + \sigma z$ for each z in teh first half of the array Z and $Y = \nu sin(\theta) + \sigma z$ for each z in the second half.

Step 6 Considering that we can assume $cos(\theta) = 1$ and $sin(\theta) = 0$ we can directly generate X's as $X = \sigma z$ for each z in the first half of the array and $Y = \nu + \sigma z$ for each z in the second half.

Step 7 Compute the Rician R.V.s as $R = \sqrt{X^2 + Y^2}$.

For purposes of evaluation, the analytical PDF of the Rice distribution is plotted as in Fig. 9. Given that it represent well the shape of the empirical histogram, you can consider this function as a valid generating method for the Rice distributed R.V.s.

```
For size = 100: mean = 5.053216936649351 and variance = 6.764890810749082 vs expectation = 5.0132565492620005 and variance = 6.867258771281655
For size = 10000: mean = 5.0062557540642665 and variance = 6.713830114061332 vs expectation = 5.0132565492620005 and variance = 6.867258771281655
For size = 100000: mean = 5.020967080563512 and variance = 6.901274807476562 vs expectation = 5.0132565492620005 and variance = 6.867258771281655
```

Figure 1: First and second moments compared with mean and variance (Rayleigh distribution) - please refer to the relative graph below

```
For size = 100: mean = 56.073915553852046 and variance = 1492.258674403237 vs expectation = 63.851400915853134 and variance = 2825.271043667387
For size = 10000: mean = 63.92597086533854 and variance = 2781.1579480309533 vs expectation = 63.851400915853134 and variance = 2825.271043667387
For size = 100000: mean = 63.75113218258596 and variance = 2856.773316424608 vs expectation = 63.851400915853134 and variance = 2825.271043667387
```

Figure 2: First and second moments compared with mean and variance (Log-normal distribution) - please refer to the relative graph below

```
For size = 100: mean = 0.7738140402046912 and variance = 0.018260386273857197 vs expectation = 0.7777777777777778 and variance = 0.01728395061728395
For size = 10000: mean = 0.7585688733876561 and variance = 0.020544843793914563 vs expectation = 0.7777777777777778 and variance = 0.01728395061728395
For size = 100000: mean = 0.7622619031273011 and variance = 0.01983442065094824 vs expectation = 0.7777777777777778 and variance = 0.01728395061728395
```

Figure 3: First and second moments compared with mean and variance (Beta distribution) - please refer to the relative graph below

```
For size = 100: mean = 7.6972469135841495 and variance = 13.800864351667755 vs expectation = 8 and variance = 16

For size = 10000: mean = 8.055111174119642 and variance = 16.317446550894324 vs expectation = 8 and variance = 16

For size = 100000: mean = 7.999016843880139 and variance = 16.045488461102284 vs expectation = 8 and variance = 16
```

Figure 4: First and second moments compared with mean and variance (Chi-squared distribution) - please refer to the relative graph below

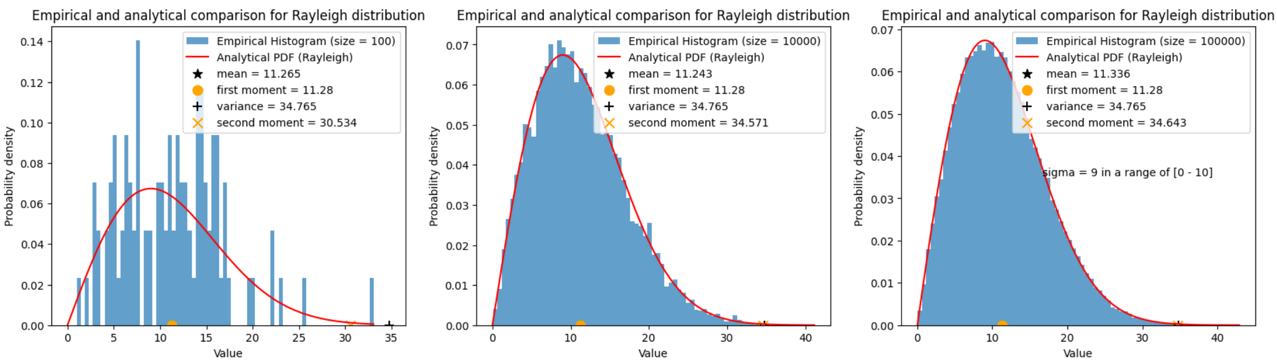

Figure 5: Rayleigh PDF and moments comparisons

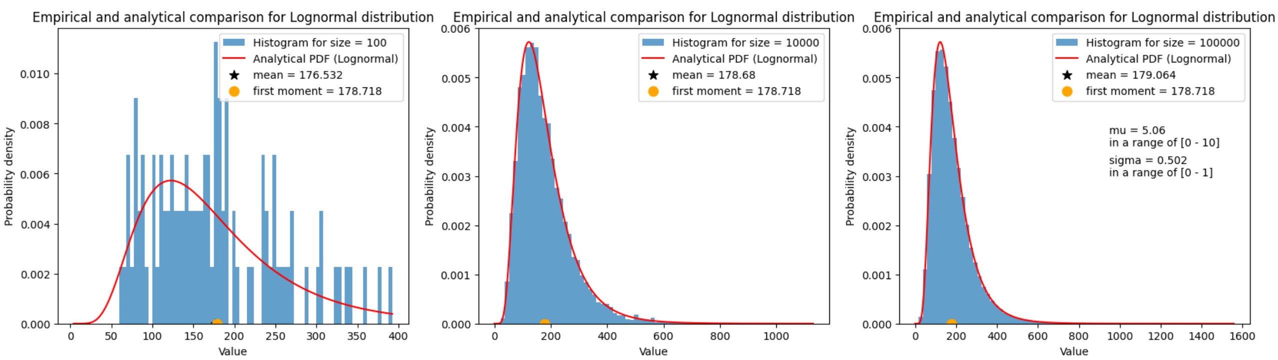

Figure 6: Log-normal PDF and moments comparisons

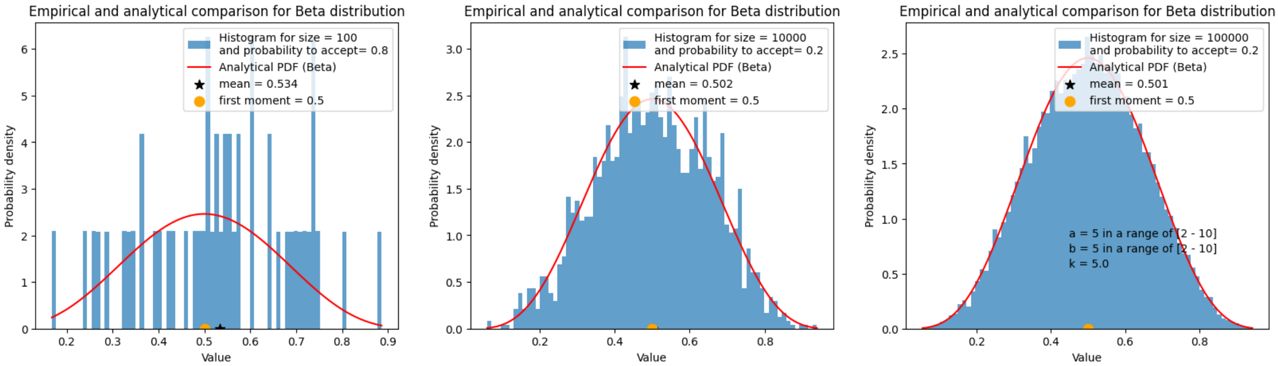

Figure 7: Beta PDF and moments comparisons

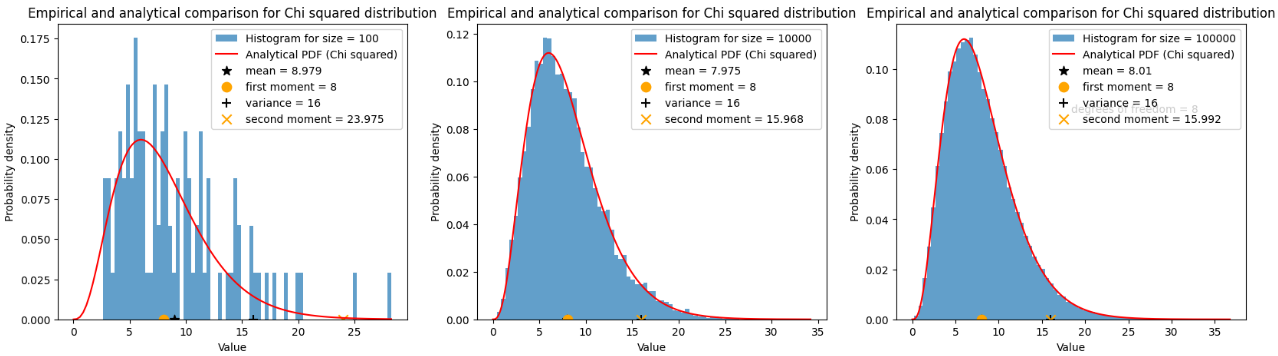

Figure 8: Chi-squared PDF and moments comparisons

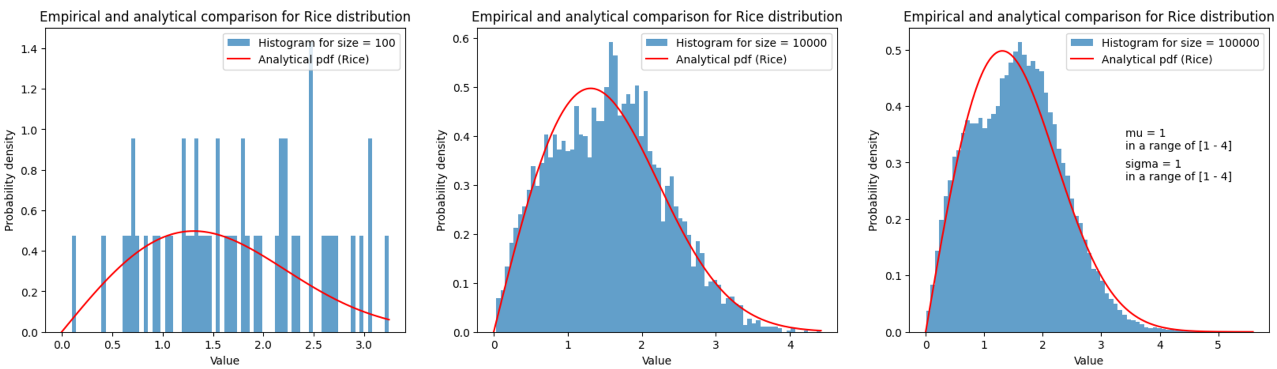

Figure 9: Rice PDF and moments comparisons