STAT 406: HW7

- All computer code should be written using the language R. Type ALL your code into one PLAIN Text format file. Plain text format is available by default in R. Please do not use Microsoft Word .doc format or .rtf format of .pdf format. Inside your plain text file, make sure you identify each problem in a comment placed at the beginning of the problem. The file name should match your name as in 'JohnDoe.R'.
- Submit your R code file online (under Assignments) at or before the due date, and hand in a hard copy of the code as well as a printed copy of your answers to the questions. The hard copy is due at the beginning of your respective lab sessions.
- I recommend that before submitting your homework, you also create a new directory and run your R code, to make sure that it is self-contained and runs as you intended.

1. Consider the discrete distribution on $\{1, 2, \ldots\}$ with probability mass function (pmf)
$$p(k) = \frac{-(1-\alpha)^k}{k \log(\alpha)},$$
for parameter $\alpha \in (0, 1)$.
   (a) Show that p is indeed a valid pmf on $\{1, 2, \ldots\}$.
   (b) Present a clear inversion algorithm to draw samples from f.
   (c) Write a R function that takes two arguments n and α, and returns n random variables each with density f.
   (d) For $\alpha = 0.3$, use your function to fill up the following table.
2. For a constant $a > 0$, we consider the problem of simulating from the a-truncated standard normal density
$$f(x) = \begin{cases} \frac{1}{c} e^{-x^2/2} & \text{if } x \geq a \\ 0 & \text{otherwise,} \end{cases}$$

x_i	1	2	3	4	5	6
True pmf value						
Estimated pmf value						

where $c = \sqrt{2\pi}\,\Pr(Z \geq a)$, $Z \sim N(0,1)$, is the normalizing constant.

(a) A naive approach to simulating from f is to generate standard normal random variables Z_i until $Z_i \geq a$. Using this naive approach, write a R function that takes arguments n and a and returns n random samples with density f.

(b) Another possibility is by rejection sampling. A possible choice of envelop density is the translated exponential density

$$g_\alpha(x) = \begin{cases} \alpha e^{-\alpha(x-a)} & \text{if } x \geq a \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha = 3a/2$. Show that if $X \sim \mathcal{E}\mathrm{xp}(\alpha)$, then $X+a$ has density g_α.

(c) Describe clearly the best possible rejection algorithm one can get to sample from f using the density g_α.

(d) Write a R function that takes arguments n, a and return n random samples from f using your rejection algorithm.

(e) For $a = 3$, and $n = 1000$ compare the running time of the two algorithms.