Transformation of random variables

1. Suppose $X \sim U(1, 5)$, and that we consider two transformations of X: $Y_1 = \sqrt{X}$ and $Y_2 = \log X$.
   (a) Use R to simulate 1000 values on each of Y_1 and Y_2, and store the results in the vectors `Y1` and `Y2` respectively.
   (b) Derive the probability density function (PDF) for each of Y_1 and Y_2.
   (c) On a 1×2 panel of plots in R (i.e. 1 row, 2 columns), produce a histogram of your samples from Y_1 and Y_2. On each histogram, overlay the PDF you derived in part (b) to confirm your simulations. Make sure you give your histograms suitable titles, axis labels etc; produce the histograms on the density scale by using the argument `probability=TRUE`; and experiment with different colours and/or line types, as demonstrated in lectures. [Hint: You will need to use the `par(mfrow = c(1, 2))` command to partition the plotting space, and you may need to experiment with the `curve` command for overlaying your derived densities.]

2. Suppose $Z \sim N(0, 1)$, and that we consider the transformation $Y = Z^2$.
   (a) Use R to simulate 1000 values on Y, and store the results in the vector `Y`.
   (b) Produce a histogram of your sample from Y, and comment on the shape of the distribution. Make sure you use the argument `probability = TRUE` when constructing your histogram.
   (c) The *chi-squared distribution on 1 degree of freedom* has PDF

$$f_X(x) = \frac{1}{2^{1/2}\Gamma(1/2)} x^{-1/2} e^{-x/2}, \quad x \geq 0.$$

   Superimpose the PDF of the chi-squared distribution onto your histogram of realised values of Y, and comment. [Hint: the Gamma function $\Gamma(\cdot)$ can be evaluated using the built-in R command `gamma`]
   (d) Use R to find the probability $\Pr(Y > 3.84)$ by considering the proportion of your simulated values in `Y` that exceed 3.84.
   (e) R has built-in functions for random simulation, and computing cumulative probabilities, from the chi-squared distribution. Find out what these are so you can simulate 1000 realisations from the chi-squared distribution directly, and so you can find a more accurate answer to the probability calculation in part (d). [Hint: recall the prefixes `r` and `p` for random generation, and obtaining cumulative probabilities, from standard models such as the Normal distribution; i.e. `rnorm`, `pnorm`. What do you *think* you might use for the chi-squared distribution?]

3. Suppose $X \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$ and $Y = \tan(X)$. Use simulation to study the distribution of Y and explicitly compute the PDF. Some pointers:
   – Produce a histogram of a sample of realisations from Y, and find some numerical summaries
   – Y actually follows *Cauchy* distribution. Is there a built-in command in R to sample directly from this distribution?