Assignment 01: Bayes Rule and Computer Generation of Random Variables

1. **A posteriori probability:** *(10 points)*

   There are $K = 11$ urns labeled by $u \in \{0, 1, 2, \ldots, 10\}$, each containing $L = 10$ balls. Urn u contains u black balls and $10 - u$ white balls. Fred selects an urn u at random and draws N times with replacement from that urn, obtaining N_B blacks and $N - N_B$ whites. Fred's friend, Bill, looks on. If after $N = 10$ draws $n_B = 3$ blacks have been drawn, what is the probability that the urn Fred is using is urn u, from Bill's point of view? (Bill doesn't know the value of u.) Calculate the numerical values of the probability for $u = 0, 1, \ldots 10$ and tabulate the probability values. Indicate the value of u for which the probability is maximum.

2. **Computer Generation of Random Variables** *(30 points)*

   Suppose that $p(u)$ denotes a valid probability density function (PDF) for a (continuous) real-valued random variable and let $F(u) = \int_{-\infty}^{u} p(v)dv$ denote the corresponding cumulative distribution function. Let X be a random variable that is uniformly distributed over the interval between 0 and 1, i.e., its probability density function (PDF) is given by

   $$p_X(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

   Now let $G(\cdot)$ denote a function for which the function $F(\cdot)$ is an inverse[1], i.e. for each $t \in [0, 1]$,

   $$F\left(G(t)\right) = t \tag{2}$$

   and define the random variable Y as a function of X by the relation:

   $$Y = G(X) \tag{3}$$

   (a) Show that the PDF $p_Y(y)$ for the random variable Y is given by $p_Y(y) = p(y)$, i.e., the PDF of the random variable matches the probability density function $p(\cdot)$. In order to simplify your proof, you may assume that $G(t)$ is differentiable with a non-zero derivative for all $t \in [0, 1]$ (though this condition is not required).

   Hint: You can solve this problem either by computing the cumulative distribution function for Y and computing the PDF from that or by using the formula for the PDF of a function of a random variable along with the chain rule for derivatives.

   (b) **Matlab:** For computer simulation, the above observation suggests a method for generating a random variable with a desired PDF $p(u)$, using the mapping $Y = G(X)$ on a uniformly distributed random variable X. Consider the PDF

   $$p(u) = \begin{cases} 2u & 0 \leq u \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

   Generate $N = 10^4$ realizations of the random variable with the above PDF using the method suggested above. Plot and compare the *suitably normalized* histogram of these realizations you generate against the above PDF in a single plot. Explain your normalization procedure. Repeat this exercise for $N = 10^6$

---

[1] Note that for a reasonable (i.e. measurable) function $p(u)$ that can be the PDF of a continuous random variable, it can be shown that such a $G(\cdot)$ exists.