The University of British Columbia

Computer Science/Data Science 405/505 Modelling and Simulation

Assignment 2

You are encouraged to discuss these problems with one or more classmates, but please make sure to submit your own work.

Exercises

1. Use 1000 pseudorandom numbers generated by the `runif()` function to simulate values from the probability distribution of a random variable X which takes the value 0, with 50% probability, 1 with 30% probability and 2, with 20% probability. To do this exercise, undertake the following steps.

   (a) First, verify for yourself that the cumulative distribution function for the random variable X takes the value 0.5 at 0, 0.8 at 1, and 1.0 at 2. In other words, $P(X \leq 0) = 0.5$, $P(X \leq 1) = 0.8$ and $P(X \leq 2) = 1.0$.

   (b) Next, write a function called `pX`, as in Example 4.6 of the textbook, which takes an argument `x` and returns the value of $P(X \leq x)$.

   (c) Now, imitate Example 4.7 of the textbook to write a function `rX` which takes `n` as an argument and returns a vector of length n consisting of random numbers that follow the distribution of X.

   (d) Assign output from the function `rX(1000)` to an object called `myX`.

   (e) Use the `table()` function to count the numbers of 0's, 1's and 2's in the vector `myX`. How do these counts differ from what you would have expected?

   (f) Construct a bar plot which displays the observed distribution of values from `myX`.

2. Simulate 10000 binomial pseudorandom numbers with parameters 20 and 0.3, assigning them to a vector called `binsim`. Let X be a binomial(20, 0.3) random variable. Use the simulated numbers to estimate

   (a) $P(X \leq 5)$.

   (b) $P(X = 5)$.

   (c) $E[X]$.

   (d) $\text{Var}(X)$.

3. Simulate vectors of 10000 pseudorandom Poisson variates with mean 5, 25, 125 and 625, assigning the results to `P1`, `P2`, `P3` and `P4`, respectively.

   (a) Use these simulated datasets to estimate $E[X]$, $E[\sqrt{X}]$, $\text{Var}(\sqrt{X})$ and $\text{Var}(X)$, where X is Poisson with rates $\lambda = 5, 25, 125$ and 625.

   (b) Noting that the variance of X increases with the mean of X, when X is a Poisson random variable, what is the effect of taking a square root of X on the relationship between the variance and the mean? (Statisticians often take square roots of count data to 'stabilize the variance'; based on what you saw in (a), can you explain what this means?)