Autumn Semester **September 2, 2020**

DS203:Programming for Data Science

Assignment 2

Exercise 1 *Let X and Y be independent exponential random variables with respective parameters λ_1 and λ_2. Find the distribution of the following.*

- $\min(X, Y)$
- $\max(X, Y)$

Exercise 2 *A bag contains 3 white, 6 red and 5 blue balls. A ball is selected at random, it's color is noted and is then replaced in the bag before making the next selection. In all 6 selections are made. Let X = the number of white balls selected and Y = number of blue balls selected. Find $E[X/Y = 3]$.*

Exercise 3 *If X_1 and X_2 are independent binomial random variables with respective parameters (n_1, p) and (n_2, p). Calculate the conditional probability mass function of X_1 given that $X_1 + X_2 = m$.*

Exercise 4 *Give an example of two random variables X and Y that are uncorrelated but not independent.*

Exercise 5 *Suppose X is a Poisson random variable with mean λ. The parameter λ is itself a random variable whose distribution is exponential with mean 1. Show that $P\{X = n\} = (1/2)^{n+1}$.*

Exercise 6 *Suppose X and Y have joint density function $f_{X,Y}(x, y) = c(1 + xy)$ if $2 \leq x \leq 3$ and $1 \leq y \leq 2$, and $f_{X,Y}(x, y) = 0$ otherwise.*

1. *Find c.*
2. *Find f_X and f_Y.*

Exercise 7 *An insurance company supposes that the number of accidents that each of its policyholders will have in a year is Poisson distributed, with the mean of the Poisson depending on the policyholder. If the Poisson mean of a randomly chosen policyholder has a gamma distribution with density function,*

$$g(\lambda) = \lambda e^{-\lambda}, \qquad \lambda \geq 0$$

what is the probability that a randomly chosen policyholder has exactly n accidents next year?

Exercise 8 *Suppose that the number of people who visit a yoga studio each day is a Poisson random variable with mean* λ*. Suppose further that each person who visits is, independently, female with probability p or male with probability* $1 - p$*. Find the joint probability that exactly n women and m men visit the academy today.*

Exercise 9 *Let* X_1, X_2, X_3 *are RVs and* a, b, c, d *are constants. Show that*

- $Cov(aX_1 + b, cX_2 + b) = acCov(X_1, X_2)$

- $Cov(X_1 + X_2, X_3) = Cov(X_1, X_3) + Cov(X_2, X_3)$.

Exercise 10 *You are given* $n = 100$ *i.i.d. samples generated from a random experiment. Let the estimate of mean from these samples is* $\hat{\mu} = 0.45$*. We know that true mean lies somewhere around* $\hat{\mu}$ *and we would like to find an interval (around* $\hat{\mu}$*) such that the true value lies in the interval with probability at least* 0.95.

- *What would be your (confidence) interval? Specify the method you used to come up with the interval.*

- *If you want the your confidence interval to shrink by half, how many more samples would you need? (the estimate could be different now)*