Question 3. We have introduced the idea of Bernoulli distribution and geometric distribution during the lecture and tutorial. The geometric distribution describes the number of trails until the first success, what does the distribution looks like if we further generalize the situation. In this case, suppose X denote the number of independent Bernoulli trials with the same probability of success p, until we have r successes. This is known as the negative binomial distribution.

(a) **[Non-coding Question]** Please derive the probability mass function of X (i.e., P(X=k; r, p)) of the above negative binomial distribution. (Hint: recall the binomial distribution when get r successes out of X trials). Furthermore, please derive the mean and variance of X (i.e., E(X) and Var(X))? (10')

Solution: X ~ Negative Binomial (r,p); Since the last trial must be a success and the total number of trials is x, there are C(x-1, r-1) possibilities, i.e. selecting r-1 successful trials out of x-1 trials ; E(X) = r/p; Var(X) = r(1-p)/(p^2)

$$P(X = x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r, \ x = r, r+1, \ldots.$$

(b) **[Non-coding Question]** It is a well-attested fact that cats have 9 lives. Montgomery is a cat who lives in a house on a busy main road. On the other side of the road is a fish shop. Every Friday, *Montgomery* sprints across the road during the rush hour to steal a haddock, and then sprints back. The probability that he is being hit by a car in any week is 1/20, independently from week to week; if he is being hit, he loses one of his lives. Find his life expectancy and the standard deviation of his lifespan in weeks if he has 9 lives left. (i.e., E(X), std(X) where X denotes the number of weeks he will survive). (10')

Solution:

E(Xr) = 20r = 180 weeks

Var(Xr) = r* (19/20)/(1/20)^2 = 380r

Std(Xr) = sqrt(Var(Xr)) = 58.48 weeks

(d) **[Non-coding Question]** What is the probability that *Montgomery* will survive for another 2 years (104 weeks) if he has 1 and 9 lives left, respectively. Please provide the theoretical results as well as use R simulation to verify your results. (10')

Solution:

Let Y denote the number of lives lost by an cat in 104 weeks, Y follows a binomial distribution B(n = 104,p = 0.05)

(1) P(X1 > 104 | one life) = P(Y = 0) = (19/20)^104 = 0.005

(2) P(X1 > 104 | nine lives) = P(Y = 0) + ... + P(Y=8) = 0.923