Question: *Given a random variable X define another random variable Y with distribution $\mathbb{P}(Y = y) = \frac{\mathbb{P}(X>y)}{\mathbb{E}[X]}$. Find the distribution corresponding to*

(a) Uniform on $1, 2, ..., n$,

(b) Poisson(λ); and

(c) give at least one example of a distribution invariant under this operation.

Solution: (a)

$$\mathbb{P}(X > y) = \begin{cases} 1 & \text{for } y < 0 \\ \frac{n-y}{n} & \text{for } 0 \leq y \leq n \\ 0 & \text{for } y > n \end{cases}$$

Recall the tail formula for the expectation of any non-negative random variable:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}(X > k)$$

This implies that in order to have $\sum_{y=-\infty}^{\infty} \mathbb{P}(Y = y) = 1$ we must set the range of Y to be the set $\{0, 1, 2, \ldots\}$ (or even $\{0, 1, \ldots, n-1\}$). Since $\mathbb{E}[X] = \frac{1}{2}(n+1)$ we have

$$\mathbb{P}(Y = y) = \begin{cases} \frac{2(n-y)}{n(n+1)} & \text{for } 0 \leq y \leq n-1 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$\mathbb{P}(X > y) = \begin{cases} 1 & \text{for } y \leq 0 \\ 1 - \sum_{k=0}^{y} \frac{\lambda^k e^{\lambda}}{k!} & \text{for } y > 0 \end{cases}$$

As before, we must set the range of Y to be the set $\{0, 1, 2, \ldots\}$. Since $\mathbb{E}[X] = \lambda$ we have

$$\mathbb{P}(Y = y) = \begin{cases} 1 - \sum_{k=0}^{y} \frac{\lambda^k e^{\lambda}}{k!} & \text{for } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(c) The answer is that X should be *geometric*, and here's why we should guess this. Recall that Y is the distribution of the amount *remaining* after time zero until the next renewal event in a *stationary* renewal process. So, we need to come up with a renewal process whose intial interval, at stationarity, has the same distribution as the intervals between events. This will be the case if the interval distributions are *memoryless*, i.e. if knowing that a certain amount of time has passed so far doesn't affect the distribution of the remaining amount. Concretely, imagine an infinite sequence of independent coin flips, and let T_k be the (renewal) sequence of times that a head occurs. The process is stationary, as can be seen by adding on another infinite, independent sequence stretching back in time. Also, the time until the first head T_1 has the same distribution as the time between subsequent heads $T_{k+1} - T_k$. By our description of stationary renewal processes, in the notation of the problem, if $T_{k+1} - T_k$ has distribution X then T_1 has distribution Y.

Now we check this. Let X be a geometric random variable on $\{0, 1, 2, \ldots\}$ with parameter p, i.e. $\mathbb{P}(X = k) = p(1-p)^k$ for $k \geq 0$. Then

$$\mathbb{P}(X > y) = \sum_{k=y+1}^{\infty} \mathbb{P}(X = k) = \sum_{k=y+1}^{\infty} p(1-p)^k = p(1-p)^{y+1} \sum_{k=0}^{\infty} (1-p)^k = (1-p)^{y+1}$$

But $E[X] = \frac{1-p}{p}$ and hence for $k \geq 0$,

$$\mathbb{P}(Y = y) = p(1-p)^k$$