In class, we claimed that one (weird) way to simulate a binomial random variable with parameters n and p is to

1. Simulate a sequence $E_1, E_2 \ldots \overset{iid}{\sim} exp(rate = 1)$.
2. Return the smallest integer X such that

$$\sum_{i=1}^{X+1} \frac{E_i}{n-i+1} > -ln(1-p).$$

Proof:

Recall that if $Y \sim bin(n,p)$, the pdf for Y is

$$P(Y=y) = \binom{n}{y} p^y (1-p)^{n-y} \quad \cdot \quad I_{\{0,1,2,\ldots,n\}}(y).$$

Consider the random variable X produced by the proposed algorithm:

$$\begin{aligned}
P(X=0) &= P\left(\tfrac{E_1}{n} > -\ln(1-p)\right) \\
&= P(E_1 > -n\ln(1-p)) \\
&= e^{-[-n\ln(1-p)]} \qquad \text{since } E_1 \sim exp(rate = 1) \\
&= (1-p)^n = P(Y=0) \ \checkmark
\end{aligned}$$

Now:

$$\begin{aligned}
P(X=1) &= P\left(\tfrac{E_1}{n} \leq -\ln(1-p), \tfrac{E_1}{n} + \tfrac{E_2}{n-1} > -\ln(1-p)\right) \\
&= \int_0^\infty P\left(\tfrac{E_1}{n} \leq -\ln(1-p), \tfrac{E_1}{n} + \tfrac{E_2}{n-1} > -\ln(1-p) \Big| E_1 = x\right) \cdot e^{-x}\, dx
\end{aligned}$$

Plugging in x for E_2, and using the fact that E_1 and E_2 are independent (so that we may drop the conditional probability when calculating a probability involving E_2), we have

$$\begin{aligned}
P(X=1) &= \int_0^\infty P\left(\tfrac{x}{n} \leq -\ln(1-p), \tfrac{x}{n} + \tfrac{E_2}{n-1} > -\ln(1-p)\right) \cdot e^{-x}\, dx \\
&= \int_0^{-n\ln(1-p)} P\left(E_2 > (n-1)[-\ln(1-p) - \tfrac{x}{n}]\right) \cdot e^{-x}\, dx \\
&= \int_0^{-n\ln(1-p)} e^{-(n-1)[-\ln(1-p)-\frac{x}{n}]} \cdot e^{-x}\, dx \\
&= (1-p)^{n-1} \int_0^{-n\ln(1-p)} e^{-x/n}\, dx \\
&= (1-p)^{n-1} \left[-ne^{-x/n}\right]_0^{-n\ln(1-p)} = np(1-p)^{n-1} = P(Y=1) \ \checkmark
\end{aligned}$$

Continuing in this manner, we get $P(X = k) = P(Y = k)$, for $k = 0, 1, 2, \ldots, n-1$. (We can formally show this by conditioning on k variables and using convolution densities.)

Note that the summand is undefined for $X = n$, however, for $X > n$, the sum given in the algorithm starts decreasing since the denominator of the summand becomes negative. So, if we do not exceed $-\ln(1-p)$ for $X = 0, 1, \ldots, n-1$, we will never exceed it for $X = n+1, n+2, \ldots$. We are forced to put the remaining mass on n. That is,

$$P(X = n) = 1 - \sum_{k=0}^{n-1} P(X = k) = p^n.$$