An *indicator random variable* is a binary variable (0 or 1) that indicates whether an event has occurred

Ex. If X ~ Bin(n, p), an indicator random variable for each trial i=1,2,…,n is

$$X_i = \begin{cases} 0 : \text{failure} \\ 1 : \text{success} \end{cases}$$

Then:

$$E(X_i) = \underset{x}{0}\underset{f(x)}{(1-p)} + (1)(p) = p$$

$$E(X_i^2) = p$$

$$E(X) = \sum_{i=1}^{n} \underbrace{E(X_i)}_{p} = np$$

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = np(1-p)$$

Ex. We have N letters addressed to N different people, and N envelopes addressed to those N people. One letter is randomly put in each envelope.

Find the mean and variance of the number of letters placed in the correct envelope.

$$X_i = \begin{cases} 0 : \text{failure} \\ 1 : \text{success} \end{cases}$$

$\sum_{i=1}^{n} X_i$ is the number of matches

Any match increases the probability of future matches, since an envelope that does not match a future letter is removed from the pool. This means that each X_i is *not* independent.

The probability of a letter being properly placed is 1/N

$$E(X_i) = 0\left(1-\frac{1}{N}\right) + 1\left(\frac{1}{N}\right) = \frac{1}{N}$$

$$E(X_i^2) = \frac{1}{N}$$

$$Var(X_i) = E(X_i^2) - E(X_i)^2 = \frac{N-1}{N^2}$$

So the expected number of matches is

$$E\left(\sum_{i=1}^{n} X_i\right) = (N)\left(\frac{1}{N}\right) = 1$$

Also:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov(X_i, X_j)$$

$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$
$$= \frac{1}{N(N-1)} - \frac{1}{N^2}$$
$$= \frac{1}{N^2(N-1)}$$

$$\Rightarrow \sum_{i=1}^{n} Var(X_i) + 2\binom{N}{2}\frac{1}{N^2(N-1)}$$

$$= \frac{1}{N} - \frac{1}{N^2} + 2\left[\frac{N!}{(N-2)!2!}\right]\left[\frac{1}{N^2(N-1)}\right]$$
$$= 1$$