SPECIAL DISCRETE DISTRIBUTIONS

Definition. A random variable X has a **discrete uniform distribution** if it is equally likely to assume any one of a finite set of possible values.

Examples. Flip a fair coin. Roll a single (fair) die. Choose a winning number in a lottery.

Definition. A random variable X has a **Bernoulli distribution** with parameter p (with $0 < p < 1$) if its has two possible values, 0 and 1, and $P(X = 1) = 0$ The PMF of a Bernoulli random variable is

$$\boxed{m(x) = \begin{cases} 1 - p & \text{if } x = 0 \\ p & \text{if } x = 1 \end{cases}}$$

The outcome 1 is often referred to as "success" while 0 is "failure" and the experiment is often called a Bernoulli trial.

Proposition. *The mean and variance of a Bernoulli random variable are* $\mu = p$ *and* $\sigma^2 = p(1 - p)$.

Examples. Flip a biased coin once. Ask a random person a yes or no question. Play the lottery once.

Definition. The total number of successes in n independent, identically distributed (iid) Bernoulli trials with parameter p is a random variable with a **Binomial distribution**. The PMF of a random variable X having a binomial distribution with parameters n and p is

$$\boxed{b(x) = \binom{n}{x} p^x (1 - p)^{n-x} \text{ for } x = 0, 1, \ldots, n}$$

Proposition. *The mean and variance of a binomial distribution are* $\mu = np$ *and* $\sigma^2 = np(1 - p)$.

Examples. Flip 6 fair coins and count the number of heads ($X \sim \text{binom}(6, 1/2)$). Roll 5 dice and count the number of sixes ($X \sim \text{binom}(5, 1/6)$). Ask 15 random people (selected with replacement) a yes or no question and count yeses ($X \sim \text{binom}(15, p)$ where p is the proportion of yeses in the population).

R Implementation. If $X \sim \text{binom}(n, p)$, then the PMF is `> dbinom(x, n, p)` and the CDF is `> pbinom(x, n, p`.

Definition. Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed (iid) Bernoulli trials, all with probability of success p. Let N be the trial on which the first success occurs. The random variable N is said to have a **geometric distribution** with parameter p and its PMF is

$$\boxed{g(n) = p(1 - p)^{n-1} \text{ for } n = 1, 2, 3, \ldots}$$

Proposition. *The mean and variance of a geometric distribution are* $\mu = \dfrac{1}{p}$ *and* $\sigma^2 = \dfrac{1}{p}\left(\dfrac{1}{p} - 1\right)$.

Examples. Flip a biased coin until the first heads appears ($N \sim \text{geom}(1/2)$). Roll a die until you first get a six ($N \sim \text{geom}(1/6)$). Ask randomly people (selected with replacement) a yes or no question until you first get a yes ($N \sim \text{geom}(p)$ where p is the proportion of yeses in the population).

R Implementation. If $X \sim \text{geom}(p)$, then the PMF is `> dgeom(x, p)` and the CDF is `> pgeom(x, p)`.

Definition. Suppose n elements are to be selected without replacement from a population of size N of which k are successes. The number of successes selected is a **hypergeometric** random variable and its PMF is

$$\boxed{h(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}}$$

Date: February 4, 2020.

Proposition. *The mean and variance of a hypergeometric distribution are* $\mu = \dfrac{nk}{N}$ *and* $\sigma^2 = \dfrac{nk(N-k)(N-n)}{k^2(N-1)}$.

Definition. A random variable with the probability mass function

$$\boxed{p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \text{ for } x = 0, 1, 2, \ldots}$$

is said to have a **Poisson distribution** with parameter λ (where $\lambda > 0$).

Proposition. *The mean and variance of a Poisson distribution are* $\mu = \lambda$ *and* $\sigma^2 = \lambda$.

R Implementation. If $X \sim \text{pois}(\lambda)$, then the PMF is `> dpois(x, λ)` and the CDF is `> ppois(x. λ)`.