~~Geometric Distributions~~ Hypergeometric Distribution

Number of successes in a dependent trials (sampling - without replacements) with fixed sample size.

↳ X: Number of success in sample size 'n'.

The hypergeometric distribution, the probability of 'x' successes when sampling without replacement 'n' items from a population with 'M' successes and 'N-M' failures, is

$$p(x) = P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

where $0 \leq x \leq M$, $0 \leq n-x \leq N-M$.

(i) Mean $E(X) = n \cdot \frac{M}{N}$.

(ii) Variance $V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$.

Example. Five individuals from an animal population thought to be near extinction in a certain region have been caught, tagged and released to mix into the population. After they have had an opportunity to mix, a random sample of 10 of these animals is selected. Let X = number of tagged animals in a second sample. If there are actually 25 animals of this type in the region, What is the probability that

a) $X = 2$

b) $X \leq 2$

c) $E(X)$

d) $V(X)$

Sol^n $n = 10$ (fixed)

$M = 5$

$N = 25$ $\quad p(x) = \dfrac{\binom{5}{x}\binom{20}{10-x}}{\binom{25}{10}}$

for $0 \leq x \leq 5$, $0 \leq 10-x \leq 20$

a) $P(X=2) = \dfrac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} = 0.385$

b) $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

$= 0.57 + 0.257 + 0.385 = 0.699.$

c) $E(X) = 10 \cdot \frac{5}{25} = 2.$

d) $V(X) = \left(\frac{25-10}{25-1}\right) \cdot 10 \cdot \frac{5}{25} \cdot \left(1 - \frac{5}{25}\right).$

$= \frac{15}{24} \cdot 10(0.2)(0.8) = 1.$

Geometric Distribution.

Conditions 1) An experiment consists of repeating trials until first success.

2) Each trial has two possible outcomes

a) A success with propability p.

b) A failure with probability $q = 1-p$.

3) Repeated trials are independent.

X = number of trials to first success.

X is a geometric random variable.

Let X be a discrete random variable, then X is said to have geometric distribution, if it's pmf is given by

$$P(X=x) = P_X(x) = p \cdot q^{x-1} \quad \text{for } x = 1, 2, \cdots$$

$X \sim Geo(p)$; X follows geometric distribution.

Q. Verify that it's proper pmf. Find $E(X)$, $V(X)$.

Soln

a) $$\sum_{x=1}^{\infty} P_X(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p$$

$$= p \sum_{x=0}^{\infty} (1-p)^x$$

$$= p \frac{1}{1-(1-p)} = 1 \quad //$$

b) $$E(X) = \sum_{x=1}^{\infty} x \cdot P_X(x)$$

$$= \sum_{x=1}^{\infty} x \cdot p \cdot q^{x-1} = \cancel{p \sum_{x=1}^{\infty} x (1-p)^{x-1}}$$

$$\cancel{= p(1 + 2q + 3q^2 + 4q^3 + \cdots)}$$

$$\therefore E(X) = p(1 + 2q + 3q^2 + 4q^3 + \cdots)$$

$$q E(X) = p(q + 2q^2 + 3q^3 + 4q^4 + \cdots)$$

$$E(X) - q E(X) = p[(1 + 2q + 3q^2 + 4q^3 + \cdots) - q - 2q^2 - 3q^3 - 4q^4 - \cdots)$$

$$E(X)(1-q) = p(1 + q + q^2 + q^3 + q^4 + \cdots)$$

$$= \cancel{p} \frac{1}{\cancel{1-q}}$$

$$\Rightarrow E(X) = \frac{1}{1-q} = \frac{1}{p} \quad //$$

c) $V(X) = E(X^2) - (E(X))^2$

$= E(X(X-1)+X) - (E(X))^2$

$$E(X(X-1)+X) = \sum_{x=1}^{\infty} x(x-1)\, P_X(x)$$

$$= \sum_{x=1}^{\infty} x(x-1)\, p q^{x-1}$$

$$= p \sum_{x=1}^{\infty} x(x-1)\, q^{x-1}$$

$$= p\left[2\cdot1\cdot q + 3\cdot2\cdot q^2 + 4\cdot3\cdot q^3 + \cdots\right]$$

$$= 2pq\left[1 + 3q + 6q^2 + \cdots\right]$$

$$= 2pq\,(1-q)^{-3}$$

$$= \frac{2q}{p^2}.$$

$$\left(\because (1-x)^{-n} = 1 + (-n)(-x) + \frac{(-n)(n-1)x^2}{2!} + \cdots\right).$$

$$\therefore V(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{2q+p-1}{p^2} = \frac{2q-q}{p^2} = \frac{q}{p^2}.$$

Example. you play a game of chance, that you can either win or lose (there are no other possibilites), and you play this game, until you lose. your probability of losing 0.57. What is the probability that it takes 5 games until you lose.

Soln $P_X(x) = p(1-p)^{x-1}$ for $x = 1, 2, 3, \cdots$

The probability of losing is 0.57.

X = number of game you play until you lose.

$$P(X=5) = P_X(5) = (0.57)(1-0.57)^{5-1}$$

$$= 0.0194.$$

$$\therefore E(X) = \frac{1}{p} = \frac{1}{0.57}, \quad V(X) = \frac{q}{p^2} = \frac{1-0.57}{(0.57)^2}.$$