The Geometric Random Variable

The geometric random variable concerns a sequence of Bernoulli trials (those resulting in either success (Σ) or failure (F)), but we do <u>not</u> here concern ourselves with the <u>number of successes</u>, but <u>rather with the random</u> variable

$$\boxed{\begin{array}{l} X : S \to \mathbb{R} \\ X(\tau_1, \tau_2, \dots) \stackrel{\text{def}}{=} \overset{\tau_n = T}{n} = \overbrace{\min\{ i \in \mathbb{N} \mid \tau_i = T \}}^{\text{first success's number}} \end{array}}$$

The sample space here is the set of all binary sequences,

$$\boxed{S = \{ \underbrace{(\tau_i)_{i \in \mathbb{N}}}_{\text{sequence of Bernoulli trials}} \mid \overbrace{\tau_i \in \{\Sigma, F\}}^{\text{Bernoulli trial}} \}}$$

If we are given

$$P(S) = p$$
$$P(F) = 1-p$$

then obviously

$$\boxed{p(n) = P(X=n) = (1-p)^{n-1}\cdot p}$$ } defining property of a geometric random variable

and as an immediate consequence

Prop. 0

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

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

$k = n-1$

$\Longleftrightarrow$ $n = k+1$

reason for the name $\longleftrightarrow$ geometric series $\sum_{n=0}^{\infty} r^k = \frac{1}{1-r}$

if $|r|<1$

$$= p\cdot\frac{1}{1-(1-p)}$$

$$= 1$$

Prop. 1
$$\boxed{\begin{aligned} E[X] &= \frac{1}{p} \\ Var(X) &= \frac{1-p}{p^2} \end{aligned}}$$

for a geometric random variable.

pf:

$$E[X] = \sum_{i=0}^{\infty} i\,p(i)$$

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

$$= p \sum_{i=1}^{\infty} (\underbrace{i-1}_{j=i-1}+1)(1-p)^{i-1}$$

$$= p \sum_{j=0}^{\infty} j(1-p)^{j} + \underbrace{p \underbrace{\sum_{j=0}^{\infty} (1-p)^{j}}_{=\frac{1}{1-(1-p)}=\frac{1}{p}}}_{=1}$$

factor out a $1-p$

$$= (1-p) \underbrace{\sum_{j=1}^{\infty} j(1-p)^{j-1} p}_{=E[X]} + 1$$

$$= (1-p)E[X] + 1$$

$$E[X] = (1-p)E[X] + 1$$

$$\Rightarrow E[X](1-(1-p)) = 1$$

$$\Rightarrow E[X] = \frac{1}{p}$$

Similarly, with $q = 1-p$ (to make things easier)

$$E[X^2] = \sum_{i=1}^{\infty} i^2 P(X=i)$$

$$= \sum_{i=1}^{\infty} i^2 (1-p)^{i-1} p$$

$$= \sum_{i=1}^{\infty} i^2 q^{i-1} p \qquad q = 1-p$$

$$= \sum_{i=1}^{\infty} (i-1+1)^2 q^{i-1} p$$

$$= \sum_{i=1}^{\infty} (i-1)^2 q^{i-1} p + 2\sum_{i=1}^{\infty} (i-1) q^{i-1} p + \underbrace{\sum_{i=1}^{\infty} q^{i-1} p}_{= p \cdot \frac{1}{1-q} = 1}$$

$$= \underbrace{\sum_{j=0}^{\infty} j^2 q^j p}_{= qE[X^2]} + \underbrace{2\sum_{j=0}^{\infty} j q^j p}_{= 2qE[X] = \frac{2q}{p}} + 1$$

factor out q factor out q

$$\underbrace{= qE[X^2] + \frac{2q}{p} + 1}$$

$$E[X^2] = qE[X^2] + \frac{2q}{p} + 1 \implies E[X^2] = \frac{2q+p}{p^2}$$

$$= \frac{q+1}{p}$$

Therefore,

$$Var(X) = E[X^2] - E[X]^2$$

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

$$= \frac{q}{p^2} \text{ or } \frac{1-p}{p^2}$$

QED

Examples of Geometric Random Variables

ex. From an urn with N white & M black balls we draw, one at a time and with replacement, a ball randomly until we get a black one.

(a) What is $P(X=n)$?

(b) What is $P(X \geq k)$?

A: Clearly $p = \frac{M}{N+M}$ & $1-p = \frac{N}{N+M}$, so

(a) $$\boxed{P(X=n) = (1-p)^{n-1} \cdot p = \left(\frac{N}{N+M}\right)^{n-1} \cdot \frac{M}{N+M} = \frac{N^{n-1}M}{(N+M)^n}}$$

ex. $M=10$, $N=20 \Rightarrow p = \frac{1}{3}$, $1-p = \frac{2}{3}$

$$\Rightarrow P(X=5) = p \cdot (1-p)^4$$

$$= \frac{1}{3} \cdot \left(\frac{2}{3}\right)^4 = \frac{16}{243} \approx 0.0658$$

(b)
$$P(X \geq k) = \sum_{i=k}^{\infty} P(X = i)$$
$$= \sum_{i=k}^{\infty} (1-p)^{i-1} \cdot p$$
$$= p(1-p)^{k-1} \underbrace{\sum_{i=0}^{\infty} (1-p)^{i}}_{= \frac{1}{1-(1-p)} = \frac{1}{p}}$$
$$= \cancel{p}(1-p)^{k-1} \cdot \frac{1}{\cancel{p}}$$
$$\underbrace{= (1-p)^{k-1}}$$

i.e. $P(X \geq k)$ = prob. that 1st $k-1$ are failures

Hence, with $1-p = \frac{N}{N+M}$, we have

$$\boxed{P(X \geq k) = \left(\frac{N}{N+M}\right)^{k-1}}$$

ex. $M = 10$, $N = 20$, $k = 5 \Longrightarrow$

$$P(X \geq 5) = \left(\frac{2}{3}\right)^4 = \frac{16}{81} \approx 0.1975$$