Random Variables :-

It is a function from the sample space Ω to the real numbers $\mathbb{R}$.

Usual notation for random variable is X, Y, Z etc.
We often don't mark them as functions: $X(\omega), Y(\omega), Z(\omega)$ etc.

Ex: Flipping 3 coins, let X count the number of heads obtained. Then as a function on Ω

$$X(T,T,T) = 0$$
$$X(T,T,H) = X(T,H,T) = X(H,T,T) = 1;$$
$$X(T,H,H) = X(H,T,H) = X(H,H,T) = 2;$$
$$X(H,H,H) = 3$$

Instead, we will say that X can take on values $0,1,2,3$ with respective probabilities $1/8, 3/8, 3/8, 1/8$.

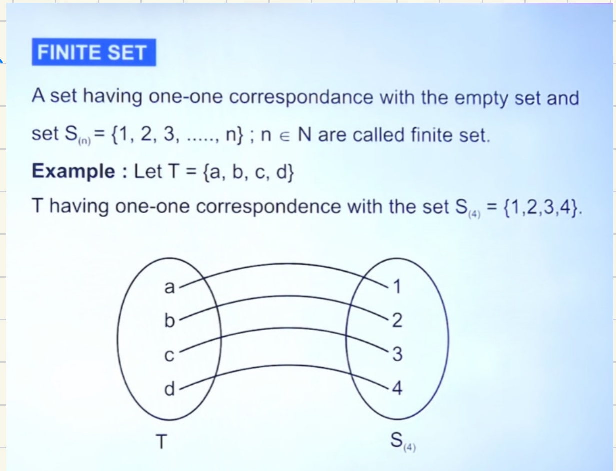
Discrete random variables:

A random variable X that can take on finitely or countably infinite many possible values is called discrete.

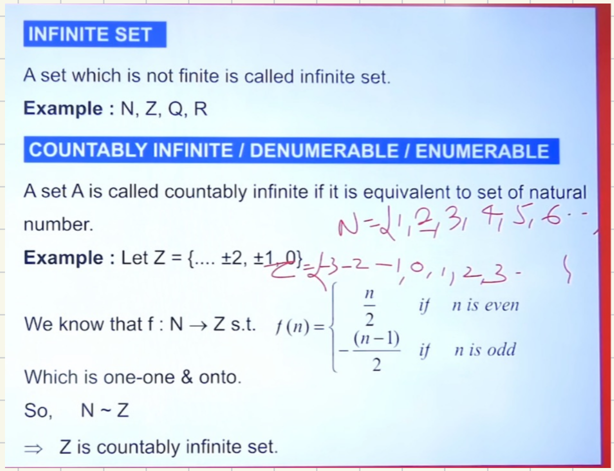
Ex: The number of heads in three coinflips is discrete.

→ The number of coin flips needed to first see a head is discrete, it can be 1, 2, 3 ...

→ The lifetime of a device is not discrete, it can be anything in the real interval $[0, \alpha)$.



Mass function:



Let X be a discrete random variable with possible values $x_1, x_2, \ldots$ The probability mass fn (pmf) or distribution of a random variable tells us the probabilities of those possible values:

$$P(x_i) = P\{X = x_i\} \quad \forall x_i$$

It says what is the probability of random variable X which takes the value x_i.

Proposition:

(1) $P(x_i) \geq 0$

(2) $\sum_i P(x_i) = 1$

Problem:

Example

Fix a positive parameter $\lambda > 0$, and define

$$\mathfrak{p}(i) = c \cdot \frac{\lambda^i}{i!}, \qquad i = 0, 1, 2, \ldots .$$

How should we choose c to make this into a mass function? In that case, what are $\mathbf{P}\{X = 0\}$ and $\mathbf{P}\{X > 2\}$ for the random variable X having this mass function?

$$\boxed{e^x = \sum_{i=0}^{n} \frac{x^i}{\lfloor i}}$$

$$p(i) = c \cdot \frac{\lambda^i}{\lfloor i}$$

$$\Rightarrow c\left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^3}{\lfloor 3} \cdots\right]$$

$$\Rightarrow c e^{\lambda} = 1$$

$$\Rightarrow \boxed{c = e^{-\lambda}}$$

$$\boxed{P(X=0) = e^{-\lambda}}$$

$$p(i) = e^{-\lambda} \frac{\lambda^i}{\lfloor i}$$

$$P(X > 2) = 1 - \left[P(X=0) + P(X=1) + P(X=2)\right]$$

$$= 1 - \left[e^{-\lambda} + e^{-\lambda}\frac{\lambda}{\lfloor 1} + e^{-\lambda}\frac{\lambda^2}{\lfloor 2}\right]$$

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

$$\boxed{P(X>2) = 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right)}$$

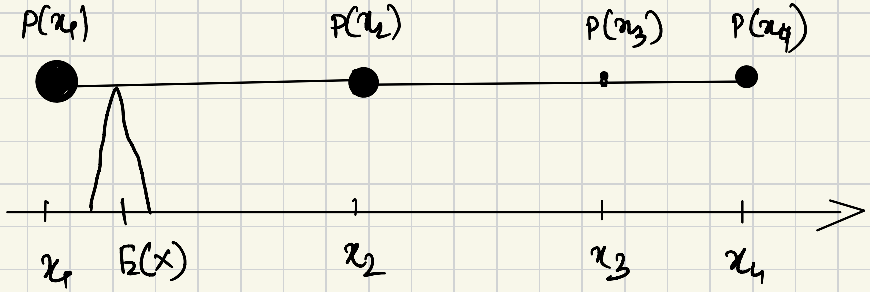
two most useful quantities to define behaviour of a random variable are

1. <u>Expectation:</u>

The expectation or mean or expected value of a discrete random variable X is defined as

$$E(X) = \sum_i x_i \, p(x_i)$$ provided that this sum exists.

the expectation is nothing else than a weighted average of the possible values x_i with weights $p(x_i)$. A center of mass, in other words.



Let X be an indicator variable:

$$X = \begin{cases} 1 & \text{if event } E \text{ occurs} \\ 0 & \text{if event } E^c \text{ occurs.} \end{cases}$$

It's mass f^n $P(1) = P\{E\}$

$P(0) = 1 - P\{E\}$

It's expectation

$$E(X) = 0 \cdot P(0) + 1 \cdot P(1)$$
$$= P\{E\}$$

Example (fair die)

Let X be the number shown after rolling a fair die. Then $X = 1, 2, \ldots, 6$, each with probability $\frac{1}{6}$. The expectation is

$$\mathbf{E}X = \sum_{i=1}^{6} i \cdot p(i) = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{1+6}{2} \cdot 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

The expected value is not necessarily a possible value. Have you ever seen a die showing 3.5…?

Properties of Expectations: (Expectation of a function of a r.v)

(1) Let X be a discrete random variable, and $g: \mathbb{R} \to \mathbb{R}$ function. Then

$$E g(X) = \sum_i g(x_i) \cdot p(x_i)$$

Ex:

X	P
0	1/8
1	3/8
2	3/8
3	1/8

Let define,

$$Y = X^2$$

$$E(Y) = E(X^2)$$

$$= \sum_i i^2 \cdot P(X=i)$$

$$= 0^2 \cdot \tfrac{1}{8} + 1^2 \times \tfrac{3}{8} + 2^2 \times \tfrac{3}{8} + 3^2 \times \tfrac{1}{8}$$

$$= \tfrac{3}{8} + \tfrac{3}{2} + \tfrac{9}{8}$$

$$= \frac{3+12+9}{8}$$

$$= \frac{\cancel{24}\,3}{\cancel{8}} = 3$$

Corollary (This works for linear function of X only)

Let X be a discrete random variable, a & b fixed real numbers. Then

$$\boxed{E(aX+b) = aE(X) + b}$$

Made with Goodnotes

Proof:

$$E(ax+b) = \sum_i (ax_i+b)\cdot p(x_i)$$

$$= a\cdot \sum_i x_i\, p(x_i) + b\cdot \sum_i p(x_i)$$

$$= a\cdot \underline{E(x) + b}$$ ✓

$E(2x+3) = 2E(x)+3$

① Moments

Let n be a positive integer. The n^{th} moment of a random variable X is defined as

$$EX^n.$$

The n^{th} absolute moment of X is

$$\underline{E|X|^n}.$$

**

$$\boxed{EX^n = E(X^n) \neq (EX)^n}$$

② Variance :

The variance & the standard deviation of a random variable is defined as

$$Var(X) = E(X^2) - \{E(X)\}^2$$

$$s.d(X) = \sqrt{Var(X)}$$

$$Y = \begin{cases} 1, & \text{w.p. } 1/2 \\ -1, & \text{w.p. } 1/2 \end{cases} \qquad Z = \begin{cases} 2, & \text{w.p. } \frac{1}{5} \\ -\frac{1}{2} & \text{w.p. } \frac{4}{5} \end{cases} \qquad U = \begin{cases} 10, & \text{w.p. } \frac{1}{2} \\ -10, & \text{w.p. } \frac{1}{2} \end{cases}$$

$E(Y) = E(Z) = E(U) = 0$ So, expection of these random variable can't distinguish among themselves. Let's try variance & s.d.

$$\begin{aligned} var(Y) &= E(Y^2) - \{E(Y)\}^2 \\ &= \sum_i i^2 P(y_i) - \{\sum_i i.P(y_i)\}^2 \\ &= \left(\tfrac{1}{2} + \tfrac{1}{2}\right) - 0 \\ &= 1 \end{aligned}$$

$s.d(Y) = \sqrt{1} = \pm 1$

$$\begin{aligned} var(Z) &= E(Z^2) - \{E(Z)\}^2 \\ &= \sum_i i^2 P(z_i) - \{\sum_i i\, P(z_i)\}^2 \\ &= 4 \cdot \tfrac{1}{5} + \tfrac{1}{5} - \left\{\tfrac{2}{5} - \tfrac{2}{5}\right\} \\ &= 1 \end{aligned}$$

$s.d(Z) = \sqrt{1} = \pm 1$

$$\begin{aligned} var(U) &= E(U^2) - \{E(U)\}^2 \\ &= \sum_i i^2 P(u_i) - \{\sum_i i\, P(u_i)\}^2 \\ &= 2 \times 100 \times \tfrac{1}{2} - 0 \\ &= 100. \end{aligned}$$

$s.d = \sqrt{100} = \pm 10$

Corollary:

for any X, $\boxed{E X^2 \geq \{E(X)\}^2}$ with equality only if $X = \text{Const.}$

variance of any random variable is always ≥ 0.

** $\boxed{E(X^2) = \{E(X)\}^2 \text{ only when } X \text{ is Constant.}}$

** Variance of the indicator random variable X of the event E is –

$$\text{Var}\,X = E(X^2) - \{E(X)\}^2$$
$$= \sum_i i^2 P(X = x_i) - \left\{\sum_i i\, P(X = x_i)\right\}^2$$
$$= P(E) - \{P(E)\}^2$$
$$= P(E)\left(1 - P(E)\right)$$

$$\text{s.d} = \sqrt{P(E)(1 - P(E))}$$

X	P
1	$P\{E\}$
0	$1 - P\{E\}$

** $\text{Var}(aX + b) = a^2\, \text{Var}(X)$

$\text{var}(X + b) = \text{var}(X) = \text{var}(-X)$

So, the variance is invariant to shifting the random variable by a constant b or to reflecting it.

☐ Bernoulli's Distribution:

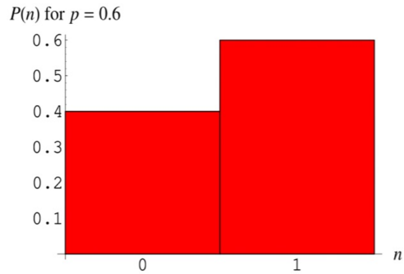
It is an discrete random variable.
when we have two outcomes.

parameter $0 < p < 1$

$X \sim \text{Bernoulli}(p = 0.5)$

Bernoulli Distribution

Download Wolfram Notebook



The Bernoulli distribution is a discrete distribution having two possible outcomes labelled by $n = 0$ and $n = 1$ in which $n = 1$ ("success") occurs with probability p and $n = 0$ ("failure") occurs with probability $q \equiv 1 - p$, where $0 < p < 1$. It therefore has probability [redacted] function
mass

$$P(n) = \begin{cases} 1-p & \text{for } n = 0 \\ p & \text{for } n = 1, \end{cases} \quad (1)$$

which can also be written

$$P(n) = p^n (1-p)^{1-n}. \quad (2)$$

(This is closed form)

The corresponding distribution function is

$$D(n) = \begin{cases} 1-p & \text{for } n = 0 \\ 1 & \text{for } n = 1. \end{cases} \quad (3)$$

Made with Goodnotes

II Binomial Distribution:

Suppose that n independent trials are performed, each succeeding with probability p. Let X count the number of successes within the n trials. Then X has the Binomial Distribution with parameters n & p or

$$X \sim \text{Binom}(n, p)$$

If $n=1$ then ~~it~~ is similar to Bernoulli's distribution.

Let $X \sim \text{Binom}(n, p)$

then it's mass function

$$P(i) = P\{X=i\} = {}^nC_i \, p^i (1-p)^{n-i}$$

$$\sum_{i=0}^{n} {}^nC_i \, p^i (1-p)^{n-i}$$

$$\Rightarrow (1-p)^n + {}^nC_1 p (1-p)^{n-1} + {}^nC_2 p^2 (1-p)^{n-2} + \cdots + p^n$$

$$\Rightarrow \left((1-p) + p\right)^n = 1^n = 1$$

A basketball player takes 4 independent free throws with a probability of 0.7 of getting a basket on each shot. Let X = the number of baskets he gets. Write out the full probability distribution for X.

X	**P**(X)
0	${}^4C_0 (0.7)^0 (0.3)^4$
1	${}^4C_1 (0.7)^1 (0.3)^3$
2	${}^4C_2 (0.7)^2 (0.3)^2$
3	${}^4C_3 (0.7)^3 (0.3)^1$
4	${}^4C_4 (0.7)^4 (0.3)^0$

$P(x=0) = {}^4C_0 (0.7)^0 (.3)^4$

$P(x=1) = {}^4C_1 (0.7)^1 (0.3)^3$

$P(x=2) = {}^4C_2 (0.7)^2 (0.3)^2$

$P(x=3) = {}^4C_3 (0.7)^3 (0.3)^1$

$P(x=4) = {}^4C_4 (0.7)^4 (0.3)^0$

Example

Screws are sold in packages of 10. Due to a manufacturing error, each screw today is independently defective with probability 0.1. If there is money-back guarantee that at most one screw is defective in a package, what percentage of packages is returned?

Define X to be the number of defective screws in a package. Then $X \sim \text{Binom}(10, 0.1)$, and the answer is the chance that a given package has 2 or more faulty screws:

$$\begin{aligned}\mathbf{P}\{X \geq 2\} &= 1 - \mathbf{P}\{X = 0\} - \mathbf{P}\{X = 1\} \\ &= 1 - \binom{10}{0} 0.1^0 0.9^{10} - \binom{10}{1} 0.1^1 0.9^9 \simeq 0.2639.\end{aligned}$$

Ex:

$P_d = 0.1$

$P_{nd} = 0.9$

$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$

or $P\{X > 1\}$

$= 1 - {}^{10}C_0\, p^0 (1-p)^{10} - {}^{10}C_1\, p^1 (1-p)^9$

$= 1 - (0.9)^{10} - 10 \times 0.1 \times (0.9)^9$

$= \underline{0.2639}$

Let $X \sim \text{Binom}(n, p)$ then

$\boxed{E(X) = np}$ $\boxed{Var(X) = np(1-p)}$

$$E(X) = \sum_i i\, p(X = i)$$

$$= \sum_{i=0}^{n} i\, {}^nC_i\, p^i (1-p)^{n-i}$$

$$= \sum_{i=0}^{n} \frac{d}{dt}(t^i)\Big|_{t=1} {}^nC_i\, p^i (1-p)^{n-i}$$

$$= \frac{d}{dt}\left(\sum_{i=0}^{n} {}^nC_i \, (pt)^i \, (1-p)^{n-i} \Big|_{t=1} \right)$$

$$= \frac{d}{dt}\left((tp + 1 - p)^n \right)$$

$$= n(p) = \underline{np}.$$

☐ Poission Distribution:

For a positive real number λ. The random variable X is poission distributed with parameter λ, in shorts $X \sim poi(\lambda)$, if it is non-negative integer valued & its mass function is

$$p(i) = P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{\underline{|i}} \quad i = 0, 1, 2 \ldots$$

**** Poission approximation of Binomial:**

Let's take $Y \sim Binom(n, p)$ with large n & small p, such that $np \simeq \lambda$. Then Y is approximately poission(λ) distributed.

Poisson approximation to binomial

$n \to \infty$
$p \to 0$ $\quad p = \frac{\lambda}{n}$

- Interesting regime: large n, small p, moderate $\underline{\lambda = np}$

- Number of arrivals S in n slots: $p_S(k) = \frac{n!}{(n-k)!\,k!} \cdot p^k (1-p)^{n-k}, \quad k = 0, \ldots, n$

For fixed $k = 0, 1, \ldots,$

$$= \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\xrightarrow[n \to \infty]{} 1 \cdot 1 \cdots 1 \cdot \frac{\lambda^k}{k!} e^{-\lambda} \cdot 1$$

- Fact: $\lim_{n \to \infty} (1 - \lambda/n)^n = e^{-\lambda}$

Expectation & Variance

if $X \sim \text{Poi}(\lambda)$, $E(X) = \text{Var}(X) = \underline{\lambda}$.

$$E(X) = \sum_i i \, P(X = i)$$

$$= \sum_i i \, e^{-\lambda} \frac{\lambda^i}{i!} = \sum_i e^{-\lambda} \frac{\lambda^i}{(i-1)!}$$

$$= \sum_i e^{-\lambda} \lambda \frac{\lambda^{i-1}}{(i-1)!}$$

$$= \lambda e^{-\lambda} \sum_i \frac{\lambda^{i-1}}{(i-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$= \underline{\lambda}.$$

Example

Screws are sold in packages of 10. Due to a manufacturing error, each screw today is independently defective with probability 0.1. If there is money-back guarantee that at most one screw is defective in a package, what percentage of packages is returned?

Define X to be the number of defective screws in a package. Then $X \sim \text{Binom}(10, 0.1)$, and the answer is the chance that a given package has 2 or more faulty screws:

$$\begin{aligned}\mathbf{P}\{X \geq 2\} &= 1 - \mathbf{P}\{X = 0\} - \mathbf{P}\{X = 1\} \\ &= 1 - \binom{10}{0} 0.1^0 0.9^{10} - \binom{10}{1} 0.1^1 0.9^9 \simeq 0.2639.\end{aligned}$$

the above problem can be solved with poisson also,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X=0) - P(X=1) \\ &= 1 - \frac{e^{-\lambda}\lambda^0}{\lfloor 0} - \frac{e^{-\lambda}\lambda^1}{\lfloor 1} \\ &= 1 - e^{-\lambda} - e^{-\lambda}\lambda \\ &\simeq \underline{0.2642.} \end{aligned}$$

$$\begin{aligned} \lambda &= np \\ &= 10 \times 0.1 \\ &= 1 \end{aligned}$$

Example

A book on average has $1/2$ typos per page. What is the probability that the next page has at least three of them?

The number X of typos on a page follows a Poisson(λ) distribution, where λ can be determined from $\frac{1}{2} = \mathbf{E}X = \lambda$. To answer the question,

$$
\begin{aligned}
\mathbf{P}\{X \geq 3\} & \\
&= 1 - \mathbf{P}\{X \leq 2\} \\
&= 1 - \mathbf{P}\{X = 0\} - \mathbf{P}\{X = 1\} - \mathbf{P}\{X = 2\} \\
&= 1 - \frac{(1/2)^0}{0!} \cdot e^{-1/2} - \frac{(1/2)^1}{1!} \cdot e^{-1/2} - \frac{(1/2)^2}{2!} \cdot e^{-1/2} \\
&\simeq 0.014.
\end{aligned}
$$

$\lambda = \frac{1}{2}$ typos/page

$$
\begin{aligned}
& P(x \geqslant 3) \\
&= 1 - P(x < 3) \\
&= 1 - P(x = 0) - P(x = 1) - P(x = 2) \\
&= 1 - \frac{e^{-\lambda}\lambda^0}{\lfloor 0} - \frac{e^{-\lambda}\lambda^1}{\lfloor 1} - \frac{e^{-\lambda}\lambda^2}{\lfloor 2} \\
&= 1 - e^{-\lambda} - e^{-\lambda} \cdot \lambda - e^{-\lambda}\frac{\lambda^2}{2} \\
&= 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right) \\
&= 1 - e^{-1/2}\left(1 + 1/2 + 1/8\right) \\
&= 1 - e^{-1/2}\left(\frac{8+4+1}{8}\right) \approx \underline{0.014} \checkmark
\end{aligned}
$$

Geometric Distribution:

The geometric distribution is a probability distribution that models the number of trials required to achieve the first success in a sequence of independent Bernoulli trials, where each trial has a constant probability of success.

i.e., Geometric distribution that is based on three important assumptions. These are listed as follows.

- → The trials being conducted are independent.
- → There can only be two outcomes of each trial - success or failure.
- → The success probability, denoted by p, is the same for each trial.

Suppose that independent trials, each succeeding with probability p, are repeated until the first success. The total number X of trials made has Geometric(p) distribution.

$$X \sim \text{Geom}(p)$$

$$\text{Pmf} \quad P(i) = (1-p)^{i-1} \cdot p \quad i \geq 1, 2 \ldots$$

for a Geometric random variable and any $k \geq 1$ we have

$P\{X \geq k\} = (1-p)^{k-1}$ [we have atleast $k-1$ failure].

Memoryless property:

$$\boxed{P(X \geq n+k \mid X > n) = P(X \geq k)}$$

$$P(X \geq n+k \mid X > n) = \frac{P(X \geq n+k \text{ AND } X > n)}{P(X > n)}$$

$$= \frac{P(X \geq n+k)}{P(X > n)}$$

$X > n$ is subevent of $X \geq n+k$

$$= \frac{(1-p)^{n+k-1}}{(1-p)^n} = (1-p)^{k-1}$$

$$= P(X \geq k).$$

Made with Goodnotes

Expectation & variance

For a Geometric(p) random variable X,

$$E(X) = 1/p \qquad var(X) = \frac{1-p}{p^2}$$

$$E(X) = \sum_i i \cdot P(X=i) = \sum_i i * (1-p)^{i-1} p$$

$$= 1 \cdot p + 2(1-p)p + 3(1-p)^2 p + \cdots$$

$$(1-p)E(X) = \qquad (1-p) \cdot p + 2(1-p)^2 p + \cdots$$

$$E(X) - (1-p)E(X) = 1 \cdot p + (1-p)p + (1-p)^2 p + \cdots$$

$$= p\left[\frac{1}{1-(1-p)}\right] = \frac{p}{p} = 1$$

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

$$\Rightarrow E(X) = 1/p.$$

Example

To first see 3 appearing on a fair die, we wait $X \sim \text{Geom}(\frac{1}{6})$ many rolls. Our average waiting time is $\mathbf{E}X = \frac{1}{1/6} = 6$ rolls, and the standard deviation is

$$\text{SD}\,X = \sqrt{\mathbf{Var}X} = \sqrt{\frac{1-\frac{1}{6}}{(\frac{1}{6})^2}} = \sqrt{30} \simeq 5.48.$$

Example (…cont'd)

The chance that 3 first comes on the 7^{th} roll is

$$\mathfrak{p}(7) = \mathbf{P}\{X = 7\} = \left(1 - \frac{1}{6}\right)^6 \cdot \frac{1}{6} \simeq 0.056,$$

while the chance that 3 first comes on the 7^{th} or later rolls is

$$\mathbf{P}\{X \geq 7\} = \left(1 - \frac{1}{6}\right)^6 \simeq 0.335.$$

Geometric Distribution

Ex Coming home from work, you always seem to hit every light. You calculate the odds of making it through a light to be 0.2. How many lights can you expect to hit before making it through one? With what std. dev.? What's the prob. of the 3^{rd} light being the first one that is green?

$p = 0.2$

$q = 0.8$

Mean: $\mu = \frac{1}{p} = \frac{1}{0.2} = \boxed{5 \text{ lights}}$

Std. Dev.: $\sigma = \frac{\sqrt{1-p}}{p} = \frac{\sqrt{q}}{p} = \frac{\sqrt{0.8}}{0.2} = \boxed{4.47 \text{ lights}}$

Prob.: $P(X=3) = q^{x-1}p = (0.8)^2(0.2) = \boxed{0.128}$

A+

$p = 0.2$

$\mu = \frac{1}{p} = 5$ lights

std = $\sqrt{\text{var}(x)}$

$= \sqrt{\frac{1-p}{p^2}}$

$= \sqrt{\frac{1-0.2}{.04}}$

$= \sqrt{\frac{0.80}{0.04}} = \sqrt{20}$

$= 4.47$ lights

$P(X=3) = (1-0.2)^2 \cdot 0.2$

$= (0.8)^2 \cdot 0.2$

$= .128$