

I ❤️ SMP Discrete Random Variables

A function that assigns a real number to each outcome:

$$X: S \rightarrow \mathbb{R}$$

* Usually an R.V. is denoted

by X, Y, Z (Uppercase)

* The outcomes are denoted

by x, y, z (lower case)

e.g. $P(X=x)$ if $x=2$

$P(X=2)$

$P(Y=y)$ if $y=3$

$P(Y=3)$

Random Variables are either Discrete or Continuous:

Discrete

- The range is countable

$\{1, 2, 3, \dots\}$

* Uniform

* Binomial

* Geometric

* Negative Binomial

* Hypergeometric

* Poisson

Continuous

- The range is continuous

\mathbb{R}_+

* Uniform

* Normal Distribution

* Chi-Squared

* Log-Normal

* Exponential

Definition: Discrete Random Variable

Let X be a discrete R.V. with $R_X = \{x_1, x_2, \dots\}$

The function

$$P_X(x_k) = P(X=x_k), \text{ for } k=1, 2, 3, \dots$$

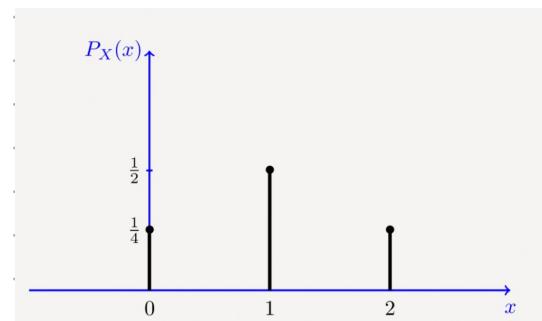
is called the probability mass function (PMF) of X .

Properties:

$$1) f(x_k) = P_X(x_k)$$

$$2) 0 \leq f(x_k) \leq 1$$

$$3) \sum_{k=1}^n f(x_k) = 1$$



Independent Random Variables:

Consider X and Y . We say X and Y are independent if:

$$P(X=x) \cap P(Y=y) = P(X=x) \cdot P(Y=y)$$

It follows:

$$P(Y=y | X=x) = P(Y=y)$$

Example:

I toss a coin twice and define X to be the number of heads I observe. Then, I toss the coin two more times and define Y to be the number of heads that I observe this time. Find $P((X < 2) \text{ and } (Y > 1))$.

$$\begin{aligned} P(X < 2, Y > 1) &= P(X < 2) \cdot P(Y > 1) \\ &= (P(X=0) + P(X=1)) \cdot P(Y=2) \\ &= \left(\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2}\right) \cdot \left(\frac{1}{2} \cdot \frac{1}{2}\right) = \left(\frac{1}{4} + \frac{1}{2}\right) \cdot \frac{1}{4} \\ &= \frac{3}{16} \approx \underline{0.1875} \end{aligned}$$

Bernoulli Distribution:

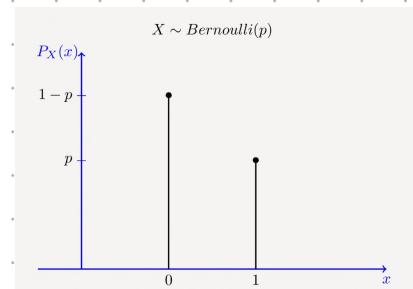
A Bernoulli R.V. can only take two values:

1 : success

0: failure

Bernoulli PMF:

$$P_x(x) = \begin{cases} p & \text{for } x=1 \\ 1-p & \text{for } x=0 \\ 0 & \text{otherwise} \end{cases}$$



Binomial Distribution:

Given multiple independent Bernoulli experiments, the resulting R.V. has a Binomial PMF:

$$P_x(x) = f(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

x = number of successes

p = probability of success

n = number of trials/Experiments

$\binom{n}{x}$ = Binomial Coefficient

↳ Combinatorial of subset:

In how many ways can I choose x elements from n when order does not matter:

$$C_r^n = \frac{n!}{r!(n-r)!}$$

We can summarize as

$$P(X=x) = \frac{n!}{x!(n-x)!} \cdot p^x (1-p)^{n-x}$$

* Flip a coin 12 times

↳ each is a Bernoulli trial

↳ together they form $B \sim B(n, p)$

↳ $n = 12, p = 1/2 \quad X \sim \text{Bin}(n, p)$

$X \sim \text{Bin}(12, 0.5)$

$P(X=2)$ means given 12 coin flips

What is the probability of getting exactly 2 heads (if defined as success)?

$$P(X=2) = \frac{12!}{2! \cdot 10!} \cdot (1/2)^2 \cdot (1/2)^{10} = \underline{\underline{0.016}}$$

* A footballer takes 4 penalties.

Assume each trial has $p = 0.7$

(scores). Then $X \sim B(4, 0.7)$

Examples:

$X \sim \text{Binomial}(n = 10, p = 0.3)$

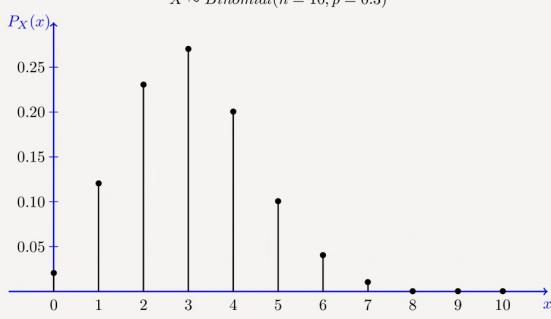
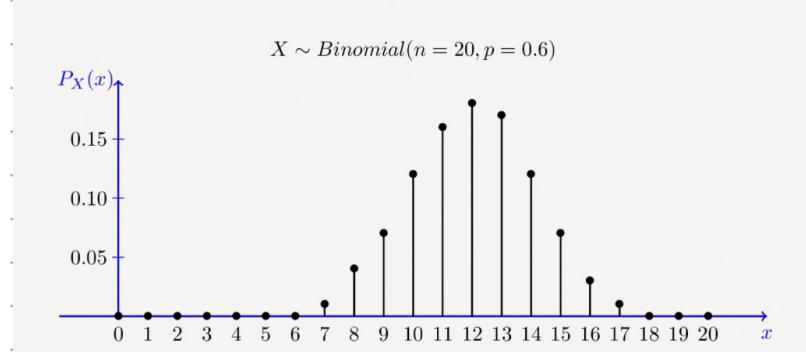


Fig.3.4 - PMF of a $\text{Binomial}(10, 0.3)$ random variable.

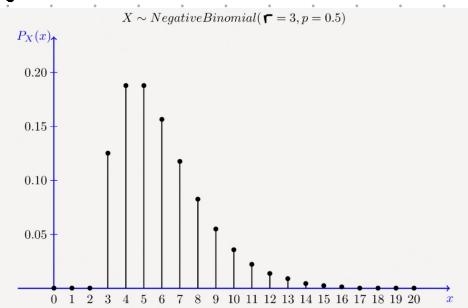
$X \sim \text{Binomial}(n = 20, p = 0.6)$



Negative Binomial Distribution: (Pascal)

Let X denote the number of trials until r successes. Then X has a negative binomial PMF:

$$f(x) = \binom{x-1}{r-1} \cdot (1-p)^{x-r} \cdot p^r$$



Geometric Distribution:

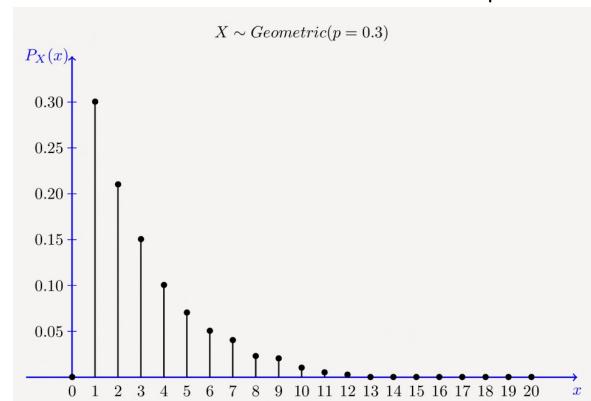
Let X denote the number of trials until the first success. Then X has a geometric PMF:

$$f(x) = (1-p)^{x-1} \cdot p$$

Relation to negative binomial:

Same but $r=1$:

$$f(x) = \binom{x-1}{1-1} \cdot (1-p)^{x-1} \cdot p^1 = (1-p)^{x-1} \cdot p$$



Hypergeometric Distribution

N

- Given a pool of size N

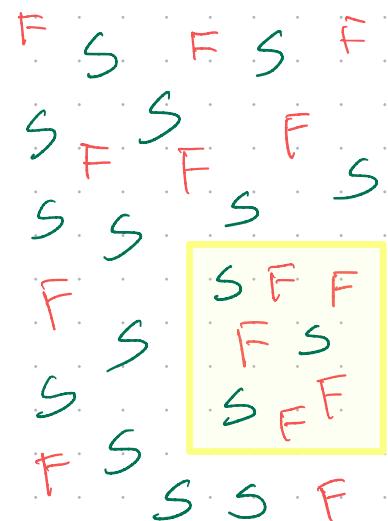
- With exactly r success

- and a random sample n

Now, let X denote the number of successes in n.

Then X is a Hypergeometric R.V. with PMF:

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$



$$N = 30$$

$$r = 16$$

$$n = 8$$

$$x = 3$$

Example: lottery

$$N = 40, r = 7, n = 7$$

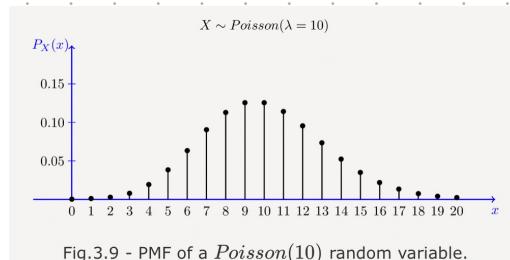
$$P(X=7) = \frac{\binom{7}{7} \cdot \binom{33}{0}}{\binom{40}{7}} = \frac{1}{\frac{40!}{7! 33!}} = \frac{1}{18643560} \approx \underline{\underline{5.36 \cdot 10^{-8}}}$$

$$P(X=4) = \frac{\binom{7}{4} \cdot \binom{33}{3}}{18643560} = \frac{\frac{7!}{4! 3!} \cdot \frac{33!}{3! 30!}}{18643560} = \frac{190960}{18643560}$$

Poisson Distribution:

Used to model the number of events occurring within a specific time/space interval and has PMF:

$$f(x) = P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$



λ indicates the average number of events in a given interval

Important

PMF Summary:

$f(x) = P(X=x)$ Allows us to find probability of exactly x success.

Assume we want to find $P(X \leq 22)$ where X is a Poisson R.V.:

$$P(X=0) + P(X=1) + P(X=2) + \dots + P(X=22)$$
$$P(X \leq 22) = \sum_{k=0}^{22} \frac{e^{-\lambda} \cdot \lambda^k}{k!}$$

Cumulative Distribution Function:

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}$$

Example: Find CDF

Toss a coin twice. Let X denote heads

$$X \sim \text{Binomial}(2, \frac{1}{2})$$

$$\mathcal{R}_x = \{0, 1, 2\}$$

$$P(X=0) = \binom{2}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(1 - \frac{1}{2}\right)^2 = \frac{1}{4}$$

$$P(X=1) = \binom{2}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(1 - \frac{1}{2}\right)^1 = \frac{1}{2}$$

$$P(X=2) = \binom{2}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \frac{1}{2}\right)^0 = \frac{1}{4}$$

$$F(x) = P(X \leq x) = 0 \quad \text{for } x < 0$$

$$F(x) = P(X \leq x) = 1 \quad \text{for } x \geq 2$$

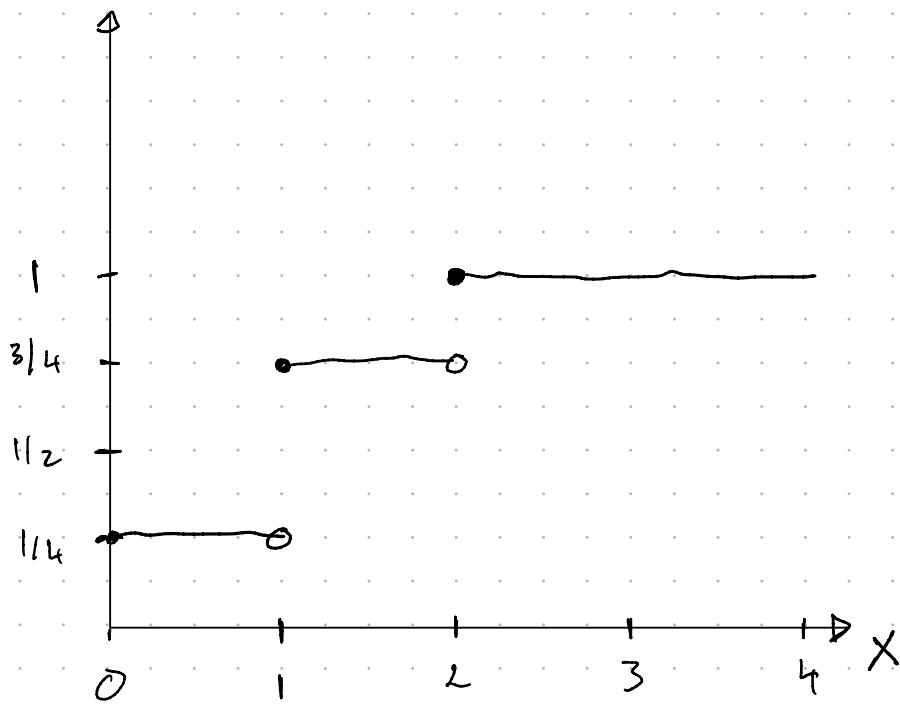
$0 \leq X \leq 1$:

$$P(X \leq x) = \frac{1}{4}$$

$1 \leq X \leq 2$:

$$P(X \leq x) = P(X=0) + P(X=1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x \leq 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$



We get:

$$F_X(x) = \sum_{X_k \leq x} P_X(X_k)$$

Intervals:

$$P(3 \leq X \leq 7) = P(X \leq 7) - P(X \leq 3) = \underbrace{P(X \leq 7) - P(X \leq 2)}_{\text{use this}}$$

$$P(3 \leq X < 7) = P(X < 7) - P(X \leq 3) = P(X \leq 6) - P(X \leq 2)$$

$$P(3 < X \leq 7) = P(X \leq 7) - P(X \leq 3) =$$

$$P(3 < X < 7) = P(X \leq 6) - P(X \leq 3)$$

In general:

$$P(a < X \leq b) = F(b) - F(a)$$

$$P(a \leq X \leq b) = F(b) - F(a-1)$$

Example:

Let X be a discrete R.V with
 $R_X = \{1, 2, 3, \dots\}$ and $f(x) = \frac{1}{2^x}$

① Find CDF

②

$$P(2 < X \leq 5) =$$

③

$$P(X > 4) =$$

Expectation:

Expected value:

$$E(X) = E(x) = \sum_{x_k \in \Omega_x} x_k \cdot P(X=x_k)$$

Recall Flipping two coins:

$$P(X=0) = 1/4, P(X=1) = 1/2, P(X=2) = 1/4$$

$$\begin{aligned} E(X) &= 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 \\ &= 1 \end{aligned}$$

Variance:

$$\begin{aligned} \text{Var}(X) &= E[(X - \mu_x)^2] \\ &= \sum_{k=1}^n x_k^2 \cdot P(X=x_k) - (E(X))^2 \end{aligned}$$

