Chapter 2. Discrete Distributions

Math 3215 Spring 2024

Georgia Institute of Technology

Section 1.
Random Variables of the Discrete Type

Random variables

Definition

Given a random experiment with a sample space S, a function X that assigns one and only one real number $X(s) = r$ to each elements in S is called a **random variable**.

The space of X is the set of real numbers $\{x : X(s) = x, s \in S\}$ and denoted by $S(X)$. (= the support of X)

Example $S = \{\text{Male}, \text{Female}\}$

$X : S \longrightarrow \mathbb{R}$ = the set of real numbers

Male $\longmapsto 1$

Female $\longmapsto 2$

$S(X) = \{1, 2\}$

$Y : S \longrightarrow \mathbb{R}$

Male $\longmapsto -1$

Female $\longmapsto 100$

$S(Y) = \{-1, 100\}$

Random variables

Example

A rat is selected at random from a cage and its sex is determined.

The set of possible outcomes is female and male. Thus, the sample space is $S = \{\text{female, male}\}$.

Random variables

Example

Consider a random experiment in which we roll a six-sided die.

The sample space associated with this experiment is $S = \{1, 2, 3, 4, 5, 6\}$.

Let $X(s) = s$. Compute $\mathbb{P}(2 \leq X \leq 4)$.

$$\begin{matrix} 1 & 2 & 3 & \cdots & 6 \end{matrix} \leftarrow X \quad \text{random variable}$$

$$\mathbb{P}(2 \leq X \leq 4)$$

$$= \mathbb{P}(X = 2 \text{ or } 3 \text{ or } 4)$$

$$= \mathbb{P}(X=2) + \mathbb{P}(X=3) + \mathbb{P}(X=4)$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

Discrete random variables

Example uncountable set, $S = [0, 1]$

Definition

Let X be a random variable defined on a sample space S.

$\{1,2,3,4\}$ $\{1,2,3,\ldots\}$

If S consists of finite outcomes or countable outcomes, then X is called a **discrete random variable**.

The **probability mass function (pmf)** of X is $f : \mathbb{R} \to \mathbb{R}$

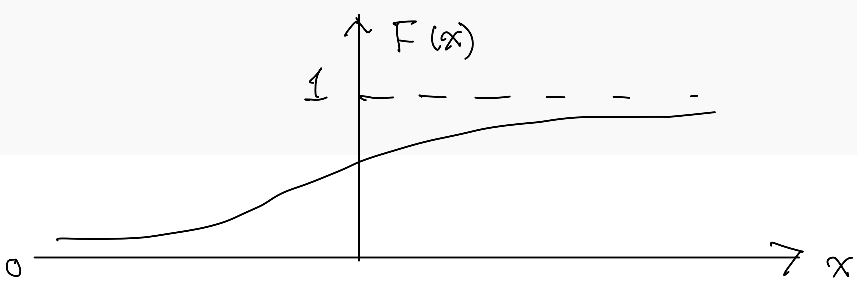
$f(x) = \mathbb{P}(X = x)$

only for discrete RVs

The **cumulative distribution function (cdf)** of X is

for any RVs, non decreasing $F : \mathbb{R} \to \mathbb{R}$

$F(x) = \mathbb{P}(X \leq x)$



($\leq$ "=" $\leqq$)

Discrete random variables

Properties of PMF

The pmf $f(x)$ of a discrete random variable X is a function that satisfies the following properties:

$f(x) = \mathbb{P}(X = x) \geqslant 0$

- $f(x) \geq 0$ for all x,
- $\sum_{x \in S(X)} f(x) = 1$, and
- $\mathbb{P}(X \in A) = \sum_{x \in A} f(x)$.

$= \mathbb{P}(S)$

Discrete random variables

$$S = \{1, 2, 3, 4, 5, 6\}$$
$$\downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \quad X$$
$$1 \ \ 2 \ \ 3 \ \ 4 \ \ 5 \ \ 6$$

Example

Roll a die, let X be the outcome.

Find the pmf and the cdf of X.

$$S(X) = \{1, 2, \cdots, 6\} = S$$

PMF $\quad f(x) = \mathbb{P}(X = x) = \begin{cases} \frac{1}{6}, & x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$

$$\left(\begin{array}{l} f(1) = \mathbb{P}(X=1) = \frac{1}{6} \\ f(0) = \mathbb{P}(X=0) = 0 \end{array} \right)$$

CDF $\quad F(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & \text{for } x < 1 \\ \frac{1}{6} & 1 \leq x < 2 \\ \frac{2}{6} & 2 \leq x < 3 \\ \frac{3}{6} & 3 \leq x < 4 \\ \frac{4}{6} & 4 \leq x < 5 \\ \frac{5}{6} & 5 \leq x < 6 \\ \frac{6}{6} & x \geq 6 \end{cases}$

$$F(1) = \mathbb{P}(X \leq 1) = \frac{1}{6}$$

$$F(0) = \mathbb{P}(X \leq 0) = 0$$

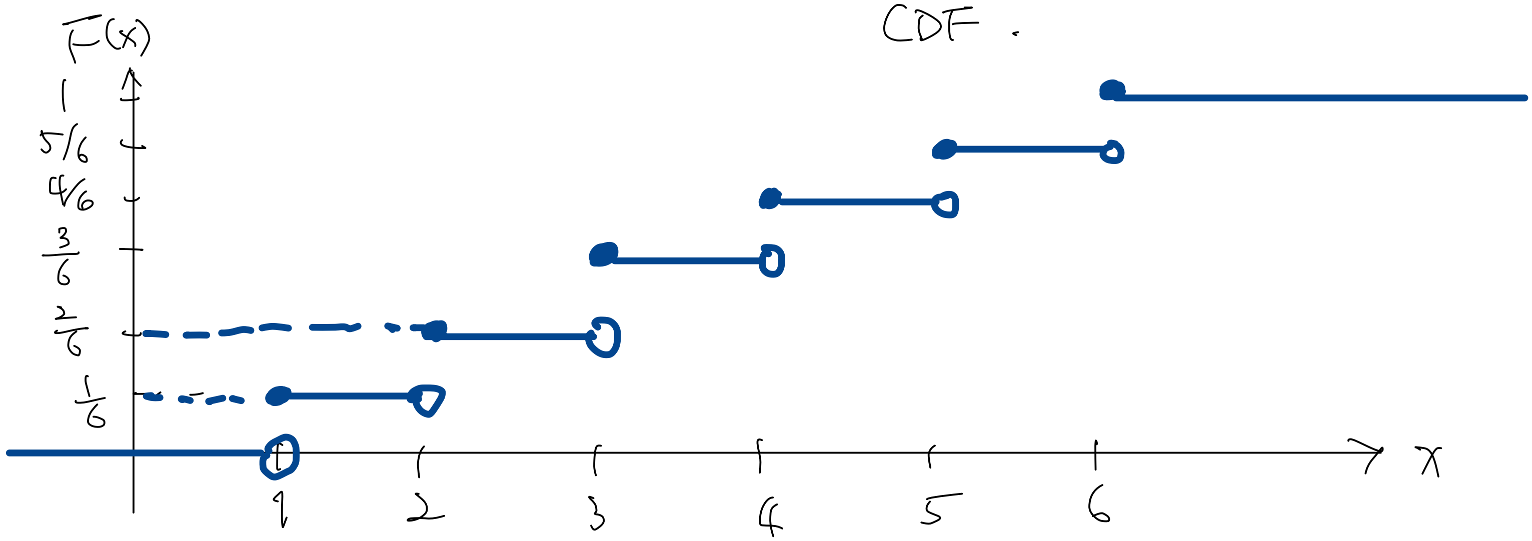
$$F(-1) = \mathbb{P}(X \leq -1) = 0$$

$$F\left(\frac{3}{2}\right) = \mathbb{P}\left(X \leq \frac{3}{2}\right) = \frac{1}{6}$$

$$F(7) = 1$$

$$F(2) = \mathbb{P}(X \leq 2) = \frac{2}{6}$$

CDF .



Discrete random variables

$S = \{ (1,1), (1,2), (1,3), \ldots (2,4) \}$ 16 outcomes

1st → ↖ 2nd

$\downarrow$ 1 $\quad \downarrow$ 2 $\quad \downarrow$ 3 $\quad \downarrow$ 4 $\cdots$ X

Example

Roll a fair four-sided die twice.

Let X equal the larger of the two outcomes if they are different and the common value if they are the same.

Find the pmf and the cdf of X.

$S(X) = \{1, 2, 3, 4\}$

PMF

$$f(x) = \mathbb{P}(X = x) = \begin{cases} 1/16 & x = 1 \\ 3/16 & x = 2 \\ 5/16 & x = 3 \\ 7/16 & x = 4 \\ 0 & \text{otherwise.} \end{cases}$$

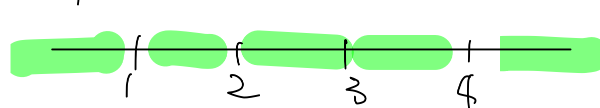
$X = 3$: $(1,3), (2,3), (3,3)$

$(3,2) \quad (3,1)$

7

CDF

$$F(x) = \mathbb{P}(X \leq x) = \begin{cases} 0, & x < 1 \\ \frac{1}{16} & 1 \leq x < 2 \\ \frac{1}{16} + \frac{3}{16} & 2 \leq x < 3 \\ \frac{1}{16} + \frac{3}{16} + \frac{5}{16} & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$



$X : S \to \mathbb{R}$ RV

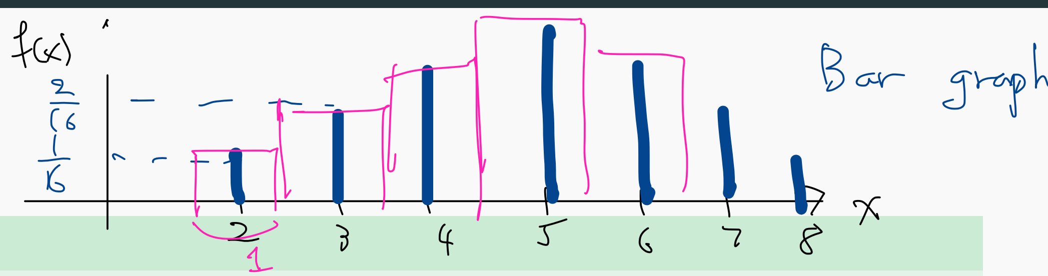
finite, countable — Discrete RV

PMF $f(x) = f_X(x) = P(X = x)$

CDF $F(x) = F_X(x) = P(X \leq x)$

$S(X) = \{ s : X = s \}$

Bar graph, Probability histogram, relative frequency histogram



Example

A fair four-sided die with outcomes 1, 2, 3, and 4 is rolled twice.

Let X equal the sum of the two outcomes.

$S = \{ (1,1), (1,2), (1,3), \ldots \}$

$\downarrow \quad \downarrow \quad \downarrow \quad \ldots \quad X$

$2 \quad 3 \quad 4$

$S(x) = \{ 2, 3, \ldots, 8 \}$

PMF
$$f(x) = \begin{cases} 1/16 & x = 2 \\ 2/16 & x = 3 \\ 3/16 & x = 4 \\ 4/16 & x = 5 \\ 3/16 & x = 6 \\ 2/16 & x = 7 \\ 1/16 & x = 8 \\ 0 & \text{otherwise} \end{cases}$$

(1,2) (2,1)

CDF $\quad F(x) = \begin{cases} 0 & x < 2 \\ 1/16 & 2 \leq x < 3 \\ 3/16 & 3 \leq x < 4 \\ 6/16 & 4 \leq x < 5 \\ \vdots & \end{cases}$

Bar graph, Probability histogram, relative frequency histogram

Example

Two fair four-sided dice are rolled. Write down the sum of the two outcomes. Repeat this 1000 times.

The sum of two outcomes	2	3	4	5	6	7	8
Number of Observations	71	124	194	258	177	122	54

$\frac{71}{1000} \quad \frac{124}{1000} \quad \frac{194}{1000}$

relative frequency

relative frequency histogram.

Section 2.
Mathematical Expectation

Definition of Expectation

Example

Consider the following game. A player roll a fair die.

If the event $A = \{1, 2, 3\}$ occurs, he receives one dollar.

If $B = \{4, 5\}$ occurs, he receives two dollars.

If $C = \{6\}$ occurs, he receives three dollars.

If the game is repeated a large number of times, what is the average payment?

Playing 6000 times

- A will happen about 3000 times $\Rightarrow$ \$ 3000
- B will happen about 2000 times $\Rightarrow$ \$ 4000
- C will happen about 1000 times $\Rightarrow$ \$ 3000

\$ 10,000 for 6000 games

Per game, $\$\frac{10}{6}$

$$\frac{10}{6} = \frac{1}{\cancel{6000}}\left(\underbrace{3000}_{=\cancel{6000}\cdot P(A)} \cdot 1 + \underbrace{2000}_{=\cancel{6000}\cdot P(B)} \cdot 2 + \underbrace{1000}_{=\cancel{6000}\cdot P(C)} \cdot 3\right)$$

$$= 1 \cdot P(A) + 2 \cdot P(B) + 3 \cdot P(C)$$

$$S = \{\underbrace{1, 2, 3}, \underbrace{4, 5}, 6\}$$

X

$$S(X) = \{1 \quad 2 \quad 3\}$$

$$= 1 \cdot P(X=1) + 2\,P(X=2) + 3\,P(X=3)$$

$$= \sum_{x \in S(X)} x \underbrace{P(X=x)}_{=f(x)} = \text{Expectation of } X$$

$$= E[X]$$

Definition of Expectation

Definition

If $f(x)$ is the pmf of a discrete random variable X with the space $S(X)$, and if the summation

$$\sum_{x \in S(X)} u(x) f(x)$$

exists, then the sum is called **the mathematical expectation or the expected value of $u(X)$**, and denoted by $\mathbb{E}[u(X)]$.

$$S = \{\underbrace{1, 2, 3}, \underbrace{4, 5}, 6\}$$

$\cancel{1}$ 1^2 $\cancel{2}$ 2^2 $\cancel{3}$ 3^2 X^2

$X \longrightarrow u(X)$

$$\mathbb{E}[X^2] = 1^2\,P(X=1) + 2^2\,P(X=2) + 3^2\,P(X=3)$$

$$\mathbb{E}[u(X)] = u(1)\,P(X=1) + u(2)\,P(X=2) + u(3)\,P(X=3)$$

$$= \sum_{x \in S(X)} u(x) \cdot f(x)$$

Definition of Expectation

Example

Let the random variable X have the pmf $f(x) = \frac{1}{3}$ for $x \in \{-1, 0, 1\} = S(X)$.

Let $Y = u(X) = X^2$.

Find the pmf of Y and $\mathbb{E}[Y] = \mathbb{E}[X^2]$.

$S(Y) = \{0, 1\}$

$$f_Y(y) = \mathbb{P}(Y = y) = \begin{cases} \overset{\mathbb{P}(X=0)}{\frac{1}{3}} & y = 0 \\ \frac{2}{3} & y = 1 \\ 0 & \text{o.w} \end{cases}$$

$\mathbb{P}(X = 1, -1)$

$$\mathbb{E}[Y] = 0 \cdot \underbrace{f_Y(0)}_{= \frac{1}{3}} + 1 \cdot \underbrace{f_Y(1)}_{= \frac{2}{3}} = \frac{2}{3}.$$

$$= \mathbb{E}[X^2] = (-1)^2 \underbrace{f_X(-1)}_{= \frac{1}{3}} + 0^2 \underbrace{f_X(0)}_{= \frac{1}{3}} + 1^2 \underbrace{f_X(1)}_{= \frac{1}{3}}$$

$$= \frac{2}{3}.$$

Properties of Expectation

Theorem

1. If c is a constant, then $\mathbb{E}[c] = c$.
2. If c is a constant and u is a function, then $\mathbb{E}[cu(X)] = c\mathbb{E}[u(X)]$.
3. If c_1 and c_2 are constants and u_1 and u_2 are functions. then

$$\mathbb{E}[c_1 u_1(X) + c_2 u_2(X)] = c_1 \mathbb{E}[u_1(X)] + c_2 \mathbb{E}[u_2(X)].$$

Properties of Expectation

$S(X) = \{1, 2, 3, 4\}$

Example

Let X have the pmf $f(x) = \frac{x}{10}$ for $x = 1, 2, 3, 4$.

Find $\mathbb{E}[X]$, $\mathbb{E}[X^2]$ and $\mathbb{E}[X(5-X)]$.

$$f(x) = \begin{cases} \frac{1}{10} & x=1 \\ \frac{2}{10} & x=2 \\ \frac{3}{10} & x=3 \\ \frac{4}{10} & x=4 \end{cases}$$

$$\mathbb{E}[X] = 1 \cdot f(1) + 2 f(2) + 3 f(3) + 4 f(4)$$

$$= 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{4}{10} = \frac{30}{10} = 3$$

$$\mathbb{E}[X^2] = 1^2 \cdot f(1) + 2^2 f(2) + 3^2 f(3) + 4^2 f(4)$$

$$= \frac{1^3}{10} + \frac{2^3}{10} + \frac{3^3}{10} + \frac{4^3}{10} = \frac{1}{10} \cdot (1+2+3+4)^2$$

$$= 10$$

$$\mathbb{E}[X(5-X)] = 1 \cdot (5-1) f(1) + 2 \cdot (5-2) f(2) + 3 \cdot (5-3) f(3) + 4(5-4) f(4)$$

$$= \mathbb{E}[5X - X^2]$$

$$= \mathbb{E}[5X] - \mathbb{E}[X^2] = 5 \cdot \mathbb{E}[X] - \mathbb{E}[X^2] = 5 \cdot 3 - 10 = 5$$

Note $\mathbb{E}[X^2] \neq (\mathbb{E}[X])^2$

$\mathbb{E}[u(X)] \neq u(\mathbb{E}[X])$

$$\mathbb{E}[X] = \sum_{x \in S(X)} x \cdot \underline{f(x)} \quad \text{PMF}, \ \mathbb{P}(X = x)$$

$$\mathbb{E}[u(X)] = \sum_{x \in S(X)} u(x) f(x)$$

Properties of Expectation

Example

An experiment has probability of success $p \in (0, 1)$ and probability of failure $q = 1 - p$.

This experiment is repeated independently until the first success occurs.

Let X be the number of trials. Find $\mathbb{E}[X]$. $\mathbb{E}[X^2]$

$X = 1$: S $\quad f(1) = \mathbb{P}(X=1) = p$

$X = 2$: F S $\quad f(2) = \mathbb{P}(X=2) = (1-p) \cdot p$

$X = 3$: F F S $\quad f(3) = (1-p)^2 p$

$X = 4$: F F F S $\quad f(4) = (1-p)^3 p$

$\vdots \qquad \vdots$

Geometric RV

$$A = \mathbb{E}[X] = 1 \cdot p + 2(1-p) \cdot p + 3(1-p)^2 \cdot p + 4(1-p)^3 p + \cdots$$

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

$$\underbrace{A - (1-p)A}_{p \cdot A} = 1 \cdot p + 1 \cdot \cancel{(1-p)} \cdot p + 1 \cdot \cancel{(1-p)^2} \cdot p + 1 \cdot \cancel{(1-p)^3} \cdot p + \cdots$$

$$-\ (1-p) \cdot pA = \cancel{(1-p) \cdot p} + \cancel{(1-p)^2 \cdot p} + \cancel{(1-p)^3 \cdot p} + \cdots$$

$$pA = pA - (1-p) \cdot pA = \cancel{p} \Rightarrow pA = 1 \quad \therefore A = \frac{1}{p}.$$

1st trial

S ← Given this

Exp. # of trial = 1

F ← Given this

Exp # of trial $= 1 + A$

$E[X] = A$

$$A = p \cdot 1 + \underline{(1-p)}(1+A)$$

$$\cancel{A} = \underbrace{p + (1-p)}_{} + (1-p) \cdot A = 1 + \cancel{A} - \underline{p \cdot A}$$

$$pA = 1 \quad \therefore \quad A = \frac{1}{p}.$$

Section 3.
Special Mathematical Expectations

Moments

The 1st moment of X about b

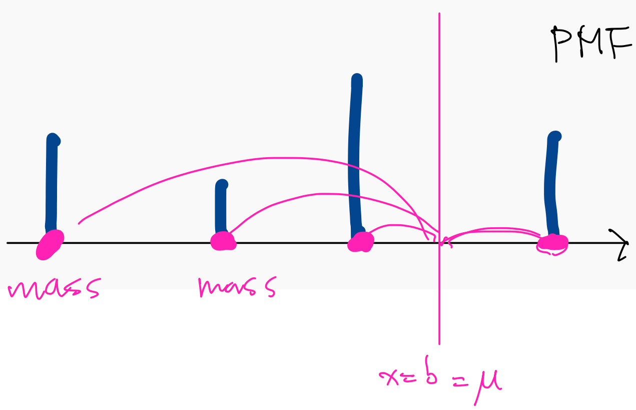
$$= \mathbb{E}[(X-b)^1]$$

The expectation or mean of a random variable X is

$$\mu = \underline{\mathbb{E}[X]} = \sum x f(x).$$

This is also called the first moment about the origin.

The first moment about the mean μ is $\mathbb{E}[X-\mu] = \mathbb{E}[X] - \mathbb{E}[\mu] = \underline{\mathbb{E}[X]} - \mu = 0$



Moments

The second moment of X about b is $\mathbb{E}[(X-b)^2]$.

If $b=\mu$, it is also called **the variance** of X and denoted by Var$(X)=\sigma^2$. $= \sigma_X^2$

Its positive square root is **the standard deviation** of X and denoted by Std$(X)=\sigma$. $= \sigma_X$

$= \sqrt{Var(X)}$

$$Var(X) = \sigma_X^2 = \mathbb{E}[(X-\mu)^2]$$

$$= \mathbb{E}[(X-\mathbb{E}(X))^2]$$

$$Std(X) = \sigma_X = \sqrt{Var(X)}$$

$$= \sqrt{\mathbb{E}[(X-\mu)^2]}$$

Moments

Example

Roll a fair die and let X be the outcome.

Find $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.

$$\mathbb{E}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{1}{6} \cdot (\underbrace{1+2+\dots+6}_{=21}) = \frac{7}{2}$$

$$\mathrm{Var}(X) = \mathbb{E}\left[\left(X - \frac{7}{2}\right)^2 \right]$$

$$= \left(1-\frac{7}{2}\right)^2 \cdot \frac{1}{6} + \left(2-\frac{7}{2}\right)^2 \frac{1}{6} + \left(3-\frac{7}{2}\right)^2 \frac{1}{6}$$

$$+ \left(4-\frac{7}{2}\right)^2 \cdot \frac{1}{6} + \left(5-\frac{7}{2}\right)^2 \frac{1}{6} + \left(6-\frac{7}{2}\right)^2 \frac{1}{6}$$

$$= \frac{1}{6} \cdot 2 \cdot \frac{1}{\cancel{4}_{2}} \cdot (\underbrace{1^2+3^2+5^2}_{=35}) = \frac{35}{12}$$

$$= \mathbb{E}\left[X^2 - \cancel{2} \cdot X \cdot \frac{7}{\cancel{2}} + \left(\frac{7}{2}\right)^2 \right]$$

$$= \mathbb{E}\left[X^2 - 7X + \frac{49}{4} \right] = \underbrace{\mathbb{E}[X^2]}_{=?} - 7 \cdot \overbrace{\mathbb{E}[X]}^{=\frac{7}{2}} + \frac{49}{4}$$

$$\mathbb{E}[X^2] = \frac{1}{6} \cdot 1^2 + \dots + \frac{1}{6} \cdot 6^2 = \frac{1}{6} \cdot \underline{(1^2 + \dots + 6^2)} \,.$$

Moments

In general, the r-th moment of X about b is $\mathbb{E}[(X-b)^r]$.

Definition

Index of skewness is defined by

$$\gamma = \mathbb{E}[(X-\mu)^3]/\sigma^3. \quad = \frac{\mathbb{E}[(X-\mu)^3]}{\left(\mathbb{E}[(X-\mu)^2]\right)^{\frac{3}{2}}}$$

Moments

$$\mu = \mathbb{E}(X) \qquad \sigma^2 = \mathbb{E}[(X-\mu)^2]$$

$$\gamma = \frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3}$$

Example

Let $f(x) = \frac{4-x}{6}$ for $x = 1, 2, 3$ be the pmf of X. Compute the index of skewness.

$$\mu = \mathbb{E}[X] = 1\cdot\frac{3}{6} + 2\cdot\frac{2}{6} + 3\cdot\frac{1}{6} = \frac{1}{6}\cdot(1\cdot 3 + 2\cdot 2 + 3\cdot 1) = \frac{10}{6}$$

$$\sigma^2 = \mathbb{E}\left[\left(X-\frac{5}{3}\right)^2\right] = \Big(\underbrace{1-\frac{5}{3}}_{=-\frac{2}{3}}\Big)^2\cdot\frac{3}{6} + \Big(\underbrace{2-\frac{5}{3}}_{=\frac{1}{3}}\Big)^2\cdot\frac{2}{6} + \Big(\underbrace{3-\frac{5}{3}}_{\frac{4}{3}}\Big)^2\cdot\frac{1}{6}$$

$$= \frac{1}{9\cdot 6}\cdot\left((-2)^2\cdot 3 + 1^2\cdot 2 + 4^2\cdot 1\right) = \frac{\cancel{30}^{\,5}}{9\cdot\cancel{6}} = \frac{5}{9}$$

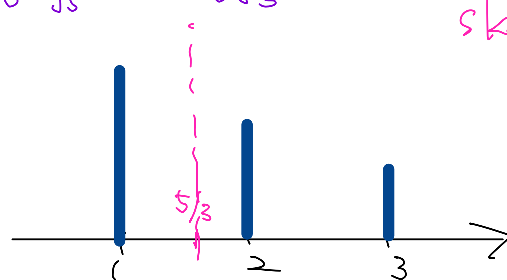
$$\mathbb{E}\left[\left(X-\frac{5}{3}\right)^3\right] = \Big(\underbrace{1-\frac{5}{3}}_{=-\frac{2}{3}}\Big)^3\cdot\frac{3}{6} + \Big(\underbrace{2-\frac{5}{3}}_{=\frac{1}{3}}\Big)^3\cdot\frac{2}{6} + \Big(\underbrace{3-\frac{5}{3}}_{\frac{4}{3}}\Big)^3\cdot\frac{1}{6}$$

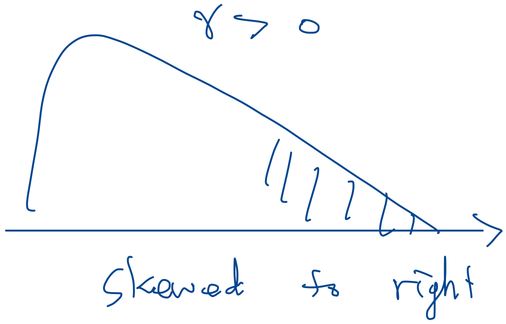
$$= \frac{1}{27\cdot 6}\Big(\underbrace{(-2)^3\cdot 3}_{-24} + \underbrace{1^3\cdot 2}_{+2} + \underbrace{4^3\cdot 1}_{+64}\Big) = \frac{7}{27}$$

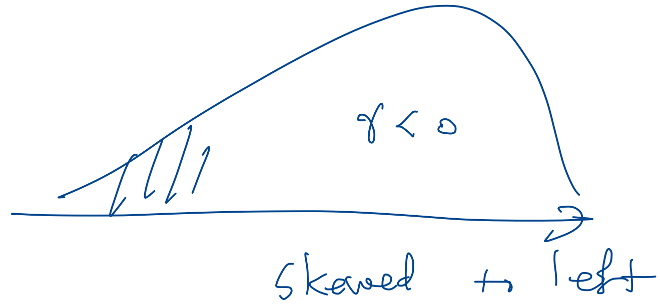
$$-24 + 2 + 64 = 42$$

$$\gamma = \frac{7/27}{(5/9)^{3/2}} = \frac{7}{27}\cdot\frac{27}{5\cdot\sqrt{5}} = \frac{7}{5\sqrt{5}} > 0$$

skewed to right







Moments

(Exp. of Sqr) − (Sqr of Exp)

Theorem

$\sigma^2 = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$= \mathbb{E}[X^2 - 2\mu X + \mu^2]$

$= \mathbb{E}[X^2] - 2\mu \underbrace{\mathbb{E}[X]}_{=\mu} + \mu^2$

$= \mathbb{E}[X^2] - 2\mu^2 + \mu^2$

$= \mathbb{E}[X^2] - \mu^2$.

$\mathrm{Var}(X) = \sum_x \underbrace{(x-\mu)^2}_{\geq 0} \underbrace{f(x)}_{\geq 0} \geq 0$

$\mathbb{E}[X^2] - (\mathbb{E}[X])^2 \geq 0$

$\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$ Jensen's Inequality

$$u(x) = e^{tx}$$

$$\mathbb{E}[u(X)] = \mathbb{E}[e^{tX}]$$

Moment generating functions

Definition

Let X be a discrete random variable and assume that there exists $h > 0$ such that

$$\mathbb{E}[e^{tX}] = \sum e^{tx} f(x) \quad \leftarrow \text{ a function of } t$$

is finite for all $t \in (-h, h)$. Then, $M(t) = \mathbb{E}[e^{tX}]$ is called **the moment generating function (mgf)**.

$$M(0) = \mathbb{E}[e^{0 \cdot X}] = 1$$

$$M'(t)\Big|_{t=0} = \frac{d}{dt}\mathbb{E}[e^{tX}]\Big|_{t=0} = \mathbb{E}\Big[\underbrace{\frac{d}{dt}e^{tX}}_{= Xe^{tX}}\Big]\Big|_{t=0}$$

$$= \mathbb{E}[Xe^{tX}]\Big|_{t=0}$$

$$= \mathbb{E}[X]$$

$$M''(0) = \mathbb{E}[X^2 e^{tX}]\Big|_{t=0}$$

$$= \mathbb{E}[X^2]$$

by $(e^t)' = e^t$

Chain Rule

$$(e^{f(t)})' = e^{f(x)} \cdot f'(t)$$

Moment generating functions

Properties

1. $M(0) = 1$
2. $M'(0) = \mathbb{E}[X]$
3. $M''(0) = \mathbb{E}[X^2]$
4. In general, $M^{(r)}(0) = \mathbb{E}[X^r]$.

$M(t) = \mathbb{E}[e^{tX}]$

Hint: $a + a \cdot r + a \cdot r^2 + \cdots = \dfrac{a}{1-r}$.

Moment generating functions

X = # of trials until first $\underline{\text{success}}$

prob $= p$.

Example PMF $x = 1, 2, 3, \ldots$

Let $f(x) = q^{x-1} p$ where $p \in (0,1)$ and $q = 1 - p$.

Compute $M(t)$.

$e^t \cdot q = r$ $\qquad$ $e^t \cdot q$

$$M(t) = \mathbb{E}[e^{tX}] = \underbrace{\boxed{e^{t \cdot 1} \cdot q^0 \cdot p}}_{a} + e^{2t} q^1 \cdot p + e^{3t} q^2 \cdot p + \cdots$$

$$= \frac{e^t \cdot p}{1 - e^t \cdot q}$$

Exercise

- $M'(0) = \frac{1}{p} = \mathbb{E}[X]$.

- $\mathbb{E}\left[\left(\frac{1}{2}\right)^X\right] = ?$

- $M(t) < \infty$? for all t?

$$\mathrm{Var}(X) = \mathrm{Std}(X)^2 = \mathbb{E}[(X-\mu)^2] \quad , \quad \mu = \mathbb{E}[X]$$

$$= \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$M(t) = \mathbb{E}[e^{tX}] \qquad -h < t < h$$

$$M(0) = 1$$

$$M'(0) = \mathbb{E}[X]$$

$$M''(0) = \mathbb{E}[X^2]$$

X , PMF $\quad f(x) = (1-p)^{x-1} \cdot p \quad , \quad p \in (0,1)$

$x = 1, 2, \cdots$

$$M(t) = \mathbb{E}[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} \cdot f(x)$$

$$= e^t \cdot (1-p)^0 \cdot p + e^{2t}(1-p)^1 \cdot p + e^{3t}(1-p)^2 p + \cdots$$

$\times e^t(1-p) \qquad e^t(1-p) \qquad \cdots$

(Geometric Series)

$$= \frac{\text{First}}{1 - \text{Ratio}} = \frac{e^t \cdot p}{1 - e^t(1-p)} . \qquad t < \ln\left(\frac{1}{1-p}\right)$$

only works when $|e^t(1-p)| < 1$

$\qquad\qquad\qquad\qquad\qquad \| $

$\qquad\qquad\qquad\qquad\quad e^t(1-p)$

$$e^t < \frac{1}{1-p}$$

$$t < \ln\left(\frac{1}{1-p}\right)$$

$$\mathbb{E}[2^X] = 2^1 \cdot (1-p)^0 p + 2^2 (1-p)^1 \cdot p + 2^3 (1-p)^2 \cdot p + \cdots$$

$2(1-p)$

$$= \frac{2p}{1 - 2(1-p)} .$$

$$E[2^X] = M(t) = E[e^{tX}] = M(\log 2)$$

??

$$2^X = e^{tX} \longrightarrow X \cdot \log 2 = t \cdot X$$

$$t = \log 2$$

Section 4.
The Binomial Distribution

Bernoulli random variables

A Bernoulli experiment, more commonly called a Bernoulli trial, is a random experiment with two outcomes.

Say $S = \{\text{success, failure }\}$ and $\mathbb{P}(\text{success}) = p$ for some $p \in (0,1)$. Then $\mathbb{P}(\text{failure}) = q = 1 - p$.

A random variable X is a **Bernoulli random variable** with success probability p is $X = 1$ if success and 0 otherwise.

$$X = \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } (1-p) \end{cases}$$

- PMF $$f(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \\ 0, & \text{otherwise} \end{cases}$$

- $\mathbb{E}[X] = 0 \cdot (1-p) + 1 \cdot p = p$.
- $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p) = p \cdot q$

  $\mathbb{E}[X^2] = 0^2 (1-p) + 1^2 \cdot p = p$
- MGF $M(t) = \mathbb{E}[e^{tX}] = e^{t \cdot 0} \cdot (1-p) + e^{t \cdot 1} p = pe^t + (1-p)$

Bernoulli random variables

Theorem

Let X be a Bernoulli random variable with success probability p.

$\mathbb{E}[X] = p$.

$\mathrm{Var}[X] = p(1-p)$.

Binomial random variables

Consider a sequence of independent Bernoulli experiments with success probability p.

Let X be the number of success trials in the first n experiments.

This is called a **Binomial random variable** with the number of trials n and success probability p.

We use the notation $X \sim b(n, p) = \text{Bin}(n, p)$.

Ex $n = 5$ $p = \frac{1}{2}$ $X \sim \text{Bin}(5, \frac{1}{2})$

$X = \#$ of Heads in 5 exp.

- PMF. $f(x) = \begin{cases} (\frac{1}{2})^5, & x = 0 \\ 5 (\frac{1}{2})^5, & x = 1 \\ \binom{5}{2} (\frac{1}{2})^5 & x = 2 \\ \binom{5}{3} (\frac{1}{2})^5 & x = 3 \\ \binom{5}{4} (\frac{1}{2})^5 & x = 4 \\ \binom{5}{5} (\frac{1}{2})^5 & x = 5 \end{cases}$

TTTTT

HTTTT, THTTT,

$\binom{5}{1} = {}_5C_1$

${}_5C_2 =$

In general, $X \sim \text{Bin}(n, p)$, $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, 2, \cdots, n$

$$\underbrace{\underbrace{H \cdots H}_{x \text{ many}} \quad \underbrace{T \cdots T}_{(n-x) \text{ many}}}_{n \text{ spots}}$$

Binomial random variables

Theorem

Let X a binomial random variable with the number of trials n and success probability p. $X \sim Bin(n,p)$

The pmf of X is $f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \cdots, n$

$\mathbb{E}[X] = n \cdot p$

$\mathrm{Var}[X] = n \cdot p \cdot (1-p)$.

$$\sum_{x=0}^{n} f(x) = 1 = \binom{n}{0} p^0 (1-p)^n + \binom{n}{1} p^1 (1-p)^{n-1} + \cdots + \binom{n}{n} p^n (1-p)^0$$

Binomial Thm $\quad (a+b)^n = \binom{n}{0} a^n \cdot b^0 + \binom{n}{1} a^{n-1} b^0 + \cdots + \binom{n}{n-1} a^1 b^{n-1} + \binom{n}{n} a^0 b^n$

$$1 = ((1-p) + p)^n = \binom{n}{0} p^0 (1-p)^n + \binom{n}{1} p^1 (1-p)^{n-1} + \cdots$$

$$\mathbb{E}[X] = \sum_{x=\cancel{0}\,1}^{n} x \cdot f(x) = \sum_{x=1}^{n} x \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$x \binom{n}{x} = x \cdot \frac{n!}{x!(n-x)!} = \frac{n \cdot (n-1)!}{(x-1)!((n-1)-(x-1))!} = n \cdot \binom{n-1}{x-1}$$

$$= n \cdot \sum_{x=1}^{n} \binom{n-1}{x-1} \overset{p \cdot p^{x-1}}{p^x} (1-p)^{n-x} \quad n-x = (n-1)-(x-1)$$

$$= n \cdot p \underbrace{\sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)}}_{\downarrow \text{ Binomial Thm}}$$

$$= np \left(p + (1-p) \right)^{n-1} = \underline{n \cdot p}.$$

$$Var(X) = E[X^2] - (\underline{E[X]})^2 = E[X^2] - (np)^2$$

$E[X^2] = ?$

$E[X(X-1)]$ is "doable".

$$= E[X^2 - X] = E[X^2] - \underbrace{E[X]}_{np}$$

<u>Binomial RV</u> Repeat $\wedge$ Bernoulli Trials n times

indep. (2 outcomes)

same success prob. p

$(0 < p < 1)$

$X = \#$ of success $\sim Bin(n, p)$

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

$(1-p = q)$

$E[X] = np$ $\quad Var(X) = np(1-p) = npq$

$$MGF = M(t) = E[e^{tX}] = (e^t p + (1-p))^n$$

<u>MGF</u> ① (Heuristic) $e^t = 1 + t + \frac{1}{2!} \cdot t^2 + \frac{1}{3!} t^3 + \cdots$

$$E[e^{tX}] = E\left[1 + (tX) + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right]$$

$$= 1 + t \cdot E[X] + \frac{t^2}{2!} E[X^2] + \frac{t^3}{3!} E[X^3] + \cdots$$

$$\Rightarrow M'(0) = E[X], \quad M''(0) = E[X^2], \cdots$$

② Sometimes, we define a RV X $\rightarrow$ give MGF easily not PMF.

$$M_X(t) \overset{\approx}{=} M_Y(t) \Rightarrow X \overset{d}{\underset{\approx}{=}} Y$$

③ $X + Y$ indep. $\Rightarrow M_{X+Y} = M_X \cdot M_Y$

Repeat B.T. 20. success prob. = 0.3

$$\underbrace{O - - - O \ O \ O \ O \ O \ O}_{10} \ \Big| \ \underbrace{O - - - - O}_{10}$$

$$\Downarrow \qquad\qquad\qquad\qquad \Downarrow$$

X_1 = # of success $\qquad$ X_2 = # of success

$X_1 \sim Bin(10, 0.3)$ $\qquad$ $X_2 \sim Bin(10, 0.3)$

$$\underline{X_1 + X_2 \sim Bin(20, 0.3)}$$

$$M_{X_1} = (e^t p + (1-p))^{10} = M_{X_2}$$

$$M_{X_1+X_2} = (e^t p + (1-p))^{20} = M_{X_1} \cdot M_{X_2}.$$

Binomial random variables

$$X \sim \mathrm{Bin}\left(8, \frac{1}{5}\right)$$

Example

Out of millions of instant lottery tickets, suppose that 20% are winners. If eight such tickets are purchased, what is the probability of purchasing two winning tickets?

$$\mathbb{P}(X = 2) = f(2) = \binom{8}{2} \cdot \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^6$$

Binomial random variables

Example

H5N1 is a type of influenza virus that causes a severe respiratory disease in birds called avian influenza (or "bird flu").

Although human cases are rare, they are deadly; according to the World Health Organization the mortality rate among humans is 60%.

Let X equal the number of people, among the next 25 reported cases, who survive the disease.

+ same prob

Assuming independence, the distribution of X is $b(25, 0.4)$. What is the probability that ten or fewer of the cases survive?

$$P(X \leq 10) = \sum_{x=0}^{10} P(X = x)$$

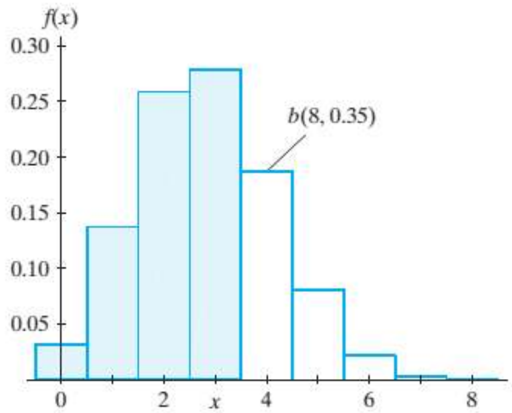
$$= \sum_{x=0}^{10} \binom{25}{x} (0.4)^x \cdot (0.6)^{25-x}$$

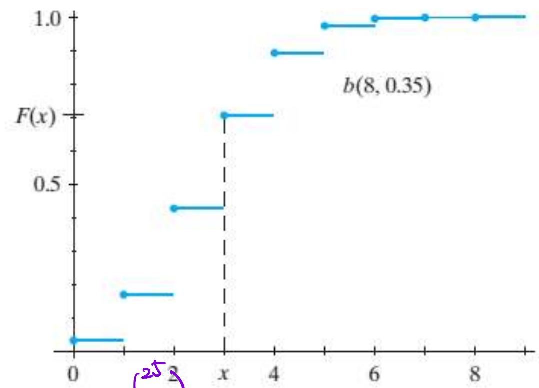
$X \sim Bin(25, 0.4)$

$P(X \le 10)$

$n = 25$. $p = 0.4$, $x = 10$

Table II The Binomial Distribution





CDF $\longrightarrow$ $$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

10, $\binom{25}{k}$, $25-k$, $n=25$

$p = 0.4$

$1-p = 0.6$

		p									
n	*x*	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7182	0.6480	0.5748	0.5000
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0312
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125
	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9688
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.9672	0.8857	0.7765	0.6553	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094
	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438
	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6562
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906
	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917	0.9844
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625

$P(X = 10) = P(X \le 10) - P(X \le 9)$

$\mathbb{P}(X \leq 10) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \cdots + \mathbb{P}(X=10)$

Table II *continued*

		p									
n	*x*	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
20	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.2454	0.1256	0.0553	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.4166	0.2500	0.1299	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.6010	0.4159	0.2520	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.7624	0.5956	0.4143	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.8782	0.7553	0.5914	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.9468	0.8725	0.7507	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9804	0.9435	0.8692	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9940	0.9790	0.9420	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9935	0.9786	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9936	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
25	0	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000	0.0000	0.0000	0.0000
	1	0.6424	0.2712	0.0931	0.0274	0.0070	0.0016	0.0003	0.0001	0.0000	0.0000
	2	0.8729	0.5371	0.2537	0.0982	0.0321	0.0090	0.0021	0.0004	0.0001	0.0000
	3	0.9659	0.7636	0.4711	0.2340	0.0962	0.0332	0.0097	0.0024	0.0005	0.0001
	4	0.9928	0.9020	0.6821	0.4207	0.2137	0.0905	0.0320	0.0095	0.0023	0.0005
	5	0.9988	0.9666	0.8385	0.6167	0.3783	0.1935	0.0826	0.0294	0.0086	0.0020
	6	0.9998	0.9905	0.9305	0.7800	0.5611	0.3407	0.1734	0.0736	0.0258	0.0073
	7	1.0000	0.9977	0.9745	0.8909	0.7265	0.5118	0.3061	0.1536	0.0639	0.0216
	8	1.0000	0.9995	0.9920	0.9532	0.8506	0.6769	0.4668	0.2735	0.1340	0.0539
	9	1.0000	0.9999	0.9979	0.9827	0.9287	0.8106	0.6303	0.4246	0.2424	0.1148
	10	1.0000	1.0000	0.9995	0.9944	0.9703	0.9022	0.7712	0.5858	0.3843	0.2122
	11	1.0000	1.0000	0.9999	0.9985	0.9893	0.9558	0.8746	0.7323	0.5426	0.3450
	12	1.0000	1.0000	1.0000	0.9996	0.9966	0.9825	0.9396	0.8462	0.6937	0.5000
	13	1.0000	1.0000	1.0000	0.9999	0.9991	0.9940	0.9745	0.9222	0.8173	0.6550
	14	1.0000	1.0000	1,0000	1.0000	0.9998	0.9982	0.9907	0.9656	0.9040	0.7878
	15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9971	0.9868	0.9560	0.8852
	16	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9957	0.9826	0.9461
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9942	0.9784
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9927
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9980
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995
	21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

Binomial random variables

$$X \sim Bin(n,p)$$

Theorem

$$(e^t)^x \cdot p^x = (e^t \cdot p)^x$$

The mgf of a binomial random variable X is

$$M(t) =$$

$$M(t) = \mathbb{E}[e^{tX}] = \sum_{x=0}^{n} (e^t)^x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (\underbrace{\boxed{e^t \cdot p}}_{= a})^x (\underbrace{\boxed{1-p}}_{= b})^{n-x}$$

$$= (e^t \cdot p + (1-p))^n.$$

Binomial random variables

Exercise

It is believed that approximately 75% of American youth now have insurance due to the health care law.

Suppose this is true, and let X equal the number of American youth in a random sample of $n = 15$ with private health insurance.

How is X distributed? Find the probability that X is at least 10. Find the mean, variance, and standard deviation of X.

$X \sim Bin(15, 0.75)$

① Indep.

② the same distribution (identical)

$$P(X \geq 10) = \sum_{x=10}^{15} \binom{15}{x} (0.75)^x (0.25)^{15-x}.$$

Use the table!

$P(X \geq 10) = P(Y \leq 5)$

Y " # of people not having insurance.

$Y \sim Bin(15, 0.25)$

$X \sim Bin(15, 0.75)$, $P(X \geq 10) = 1 - P(X \leq 9)$

Table II *continued*

n	x	p = 0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9978	0.9935
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	0	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0016	0.0005	0.0001	0.0000
	1	0.8290	0.5490	0.3186	0.1671	0.0802	0.0353	0.0142	0.0052	0.0017	0.0005
	2	0.9638	0.8159	0.6042	0.3980	0.2361	0.1268	0.0617	0.0271	0.0107	0.0037
	3	0.9945	0.9444	0.8227	0.6482	0.4613	0.2969	0.1727	0.0905	0.0424	0.0176
	4	0.9994	0.9873	0.9383	0.8358	0.6865	0.5155	0.3519	0.2173	0.1204	0.0592
	5	0.9999	0.9978	0.9832	0.9389	0.8516	0.7216	0.5643	0.4032	0.2608	0.1509
	6	1.0000	0.9997	0.9964	0.9819	0.9434	0.8689	0.7548	0.6098	0.4522	0.3036
	7	1.0000	1.0000	0.9994	0.9958	0.9827	0.9500	0.8868	0.7869	0.6535	0.5000
	8	1.0000	1.0000	0.9999	0.9992	0.9958	0.9848	0.9578	0.9050	0.8182	0.6964
	9	1.0000	1.0000	1.0000	0.9999	0.9992	0.9963	0.9876	0.9662	0.9231	0.8491
	10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9972	0.9907	0.9745	0.9408
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9981	0.9937	0.9824
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9989	0.9963
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0010	0.0003	0.0001	0.0000
	1	0.8108	0.5147	0.2839	0.1407	0.0635	0.0261	0.0098	0.0033	0.0010	0.0003
	2	0.9571	0.7892	0.5614	0.3518	0.1971	0.0994	0.0451	0.0183	0.0066	0.0021
	3	0.9930	0.9316	0.7899	0.5981	0.4050	0.2459	0.1339	0,0651	0.0281	0.0106
	4	0.9991	0.9830	0.9209	0.7982	0.6302	0.4499	0.2892	0.1666	0.0853	0.0384
	5	0.9999	0.9967	0.9765	0.9183	0.8103	0.6598	0.4900	0.3288	0.1976	0.1051
	6	1.0000	0.9995	0.9944	0.9733	0.9204	0.8247	0.6881	0.5272	0.3660	0.2272
	7	1.0000	0.9999	0.9989	0.9930	0.9729	0.9256	0.8406	0.7161	0.5629	0.4018
	8	1.0000	1.0000	0.9998	0.9985	0.9925	0.9743	0.9329	0.8577	0.7441	0.5982
	9	1.0000	1.0000	1.0000	0.9998	0.9984	0.9929	0.9771	0.9417	0.8759	0.7728
	10	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9938	0.9809	0.9514	0.8949
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9851	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9991	0.9965	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0002	0.0000	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0021	0.0005	0.0001	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0121	0.0036	0.0009	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0444	0.0160	0.0049	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.1182	0.0510	0.0189	0.0059

Section 5.
The Hypergeometric Distribution

Ex 6 Red balls 4 Blue balls

Choose 4 balls at random with replacement

X = # of red balls chosen. $\sim Bin(4, 0.6)$

Question : without replacement.

$$\left(\begin{aligned} \mathbb{P}(2^{nd} = R) &= \mathbb{P}(BR) + \mathbb{P}(RR) \\ &= \frac{4}{10} \cdot \frac{6}{9} + \frac{6}{10} \cdot \frac{5}{9} = \frac{24+30}{10 \cdot 9} = \frac{6}{10} \end{aligned} \right)$$

$$\mathbb{P}(X=2) = \frac{\binom{6}{2}\cdot\binom{4}{2}}{\binom{10}{4}}$$

$$= \underbrace{\frac{6}{10}\cdot\frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}}_{RRBB}\cdot\binom{4}{2}$$

# of arrangement.

The Hypergeometric Distribution

$N_1 + N_2 = N$.

There is a collection of $\overset{=6}{N_1}$ red balls and $\overset{=4}{N_2}$ blue balls.

Sample $\overset{=4}{n}$ balls at random **without replacement** ($n \leq N_1 + N_2$).

Let X be the number of red balls chosen.

Then, X is called a **hypergeometric random variable** with parameters N_1, N_2, n, and denoted by $\text{HG}(N_1, N_2, n)$.

If with replacement, $\text{Bin}\left(n, \frac{N_1}{N_1+N_2}\right)$

The Hypergeometric Distribution

Example

In a small pond there are 50 fish, ten of which have been tagged.

If a fisherman's catch consists of seven fish selected at random and without replacement, and X denotes the number of tagged fish,

what is the probability that exactly two tagged fish are caught?

$$X \sim HG(\underset{N_1}{10}, \underset{N_2}{40}, \underset{n}{7})$$

$$P(X=2) = \frac{\binom{10}{2}\cdot\binom{40}{5}}{\binom{50}{7}} .$$

$X \sim HG(N_1, N_2, n)$

PMF: $f(x) = \dfrac{\binom{N_1}{x} \cdot \binom{N_2}{n-x}}{\binom{N_1+N_2}{n}}$

$n - x \le N_2$

$x \ge n - N_2$

$(x = \cancel{0}, 1, 2, \dots, \min\{n, N_1\})$

$x = \max\{0, n-N_2\}, \dots, \min\{n, N_1\}$

Ex $N_1 = 5$, $N_2 = 10$, $n = 7$

The Hypergeometric Distribution

without replacement

with replacement

Theorem

$\mathbb{P}(X = k) =$

$\mathbb{E}[X] = n\frac{N_1}{N_1+N_2}$

$\mathrm{Var}[X] = n\frac{N_1}{N_1+N_2}\frac{N_2}{N_1+N_2} \cdot \frac{N_1+N_2-n}{N_1+N_2-1}$

$X \sim HG(N_1, N_2, n)$

$Y \sim Bin(n, \underbrace{\frac{N_1}{N_1+N_2}}_{p})$

$E[Y] = n \cdot \frac{N_1}{N_1+N_2}$

$Var(Y) = n \cdot p \cdot (1-p)$

$= n \cdot \frac{N_1}{N_1+N_2} \frac{N_2}{N_1+N_2}$

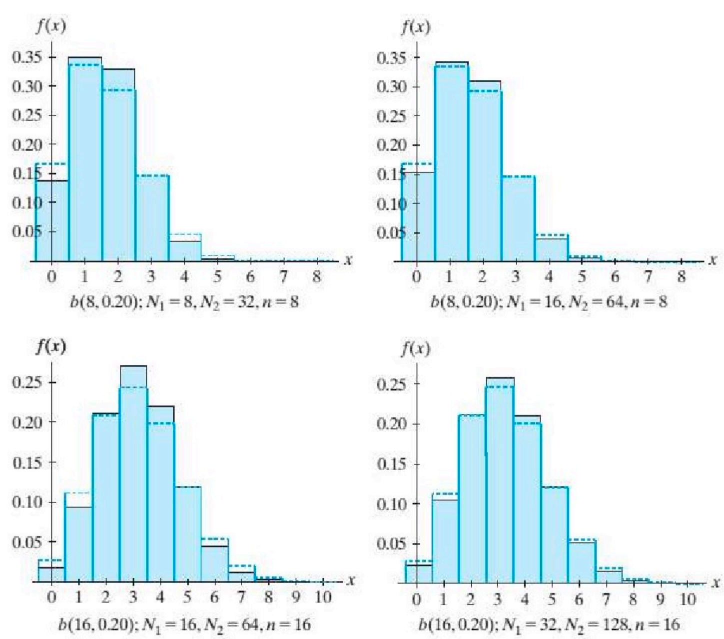


Figure 2.5-2 Binomial and hypergeometric (shaded) probability histograms

$X \simeq Y$

when $N_1, N_2 \to \infty$

$\frac{N_1}{N_1+N_2} = p$

The Hypergeometric Distribution

Exercise

In a lot (collection) of 100 light bulbs, there are five bad bulbs.

An inspector inspects ten bulbs selected at random. without replacement

Find the probability of finding at least one defective bulb.

$X =$ # of Defective Bulbs chosen.

$\sim HG(\underset{N_1}{5}, \underset{N_2}{95}, \underset{n}{10})$

$P(X \geq 1) = 1 - P(X=0)$

$= 1 - \dfrac{\binom{5}{0}\binom{95}{5}}{\binom{100}{10}}$

Section 6.
The Negative Binomial Distribution

Geometric random variables

Consider a sequence of independent Bernoulli trials with success probability $p \in (0,1)$

Let X be the number of trials until the first success.

This random variable is called a **geometric random variable**. $X \sim Geom(p)$

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

$x = 1, 2, 3, \ldots$

$\underbrace{F \; F \; \ldots \; F \; F \; S}_{x \text{ trials}}$

Geometric random variables

$X \sim Geom(p)$ $q = 1-p$

Theorem

The pmf of X is $f(x) = (1-p)^{x-1} \cdot p$. $x = 1, 2, \ldots$

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

$\mathrm{Var}[X] = \frac{q}{p^2}$

$M(t) = \frac{pe^t}{1-(1-p)e^t}$

Geometric random variables

Example

Some biology students were checking eye color in a large number of fruit flies.

For the individual fly, suppose that the probability of white eyes is 1/4 and the probability of red eyes is 3/4, and that we may treat these observations as independent Bernoulli trials.

What is the probability that at least four flies have to be checked for eye color to observe a white-eyed fly?

X = # of observations until 1^{st} white.

$\sim \text{Geom}\left(\frac{1}{4}\right)$.

$$P(X \geq 4) = \sum_{x=4}^{\infty} (1-p)^{x-1} \cdot p \qquad \left(p = \frac{1}{4}\right).$$

$$= (1-p)^3 \cdot p + (1-p)^4 \cdot p + (1-p)^5 \cdot p + \cdots$$

$$= \frac{(1-p)^3 \cdot \cancel{p}}{\underbrace{1-(1-p)}_{=\cancel{p}}} = (1-p)^3.$$

$$P(X \geq 4) = P(FFF) = (1-p)^3$$

In general, $X \sim \text{Geom}(p)$

$$P(X > k) = (1-p)^k$$

Negative Binomial random variables

Note : $\text{Geom}(p) = \text{NegBin}(1, p)$

Consider a sequence of independent Bernoulli trials with success probability

Let X be the number of trials until the r-th success.

This random variable is called a **negative binomial random variable**.

$$X \sim \text{NegBin}(r, p)$$

PMF :

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

$$x = r, r+1, \ldots$$

$$\underbrace{(r-1)\ S\ ,\ (x-r)\ F \ \Big|\ S}_{x}$$

$(x-1)$ positions

Negative Binomial random variables

Theorem $X \sim \text{NegBin}(r, p)$

The pmf of X is

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

for $k = r, r+1, \cdots$ and otherwise zero.

$\mathbb{E}[X] = \frac{r}{p}$

$\text{Var}[X] = \frac{rq}{p^2}$

$M(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$

A negative binomial random variable can be written as a sum of independent geometric random variables.

Ex $X \sim \text{NegBin}(3, p)$

$X_2 \sim \text{Geom}(p)$ $X_3 \sim \text{Geom}(p)$

$\boxed{F\ F\ S}\ \boxed{F\ F\ F\ F\ S}\ \boxed{F\ F\ F\ S}$

$X_1 \sim \text{Geom}(p)$

$X = 12 = X_1 + X_2 + X_3$

Indep.

$$\frac{10}{(12.5)} = 80\%$$

Negative Binomial random variables

Example

Suppose that during practice a basketball player can make a free throw 80% of the time.

Furthermore, assume that a sequence of free-throw shooting can be thought of as independent Bernoulli trials.

Let X equal the minimum number of free throws that this player must attempt to make a total of ten shots.

Find the mean of X.

$$X \sim \text{NegBin}\left(10, \frac{4}{5}\right)$$

$$E[X] = \frac{r}{p} = 10 \cdot \frac{5}{4} = 12.5.$$

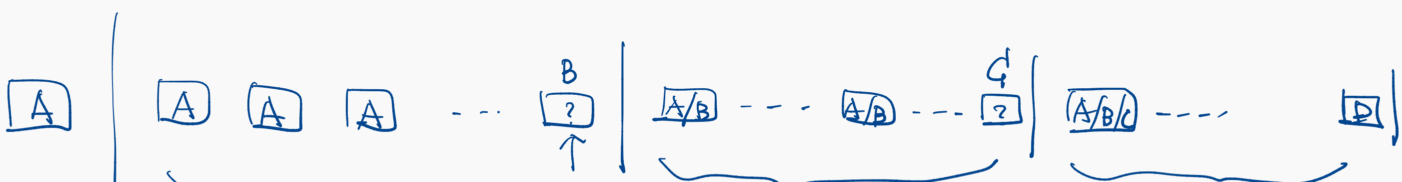
Negative Binomial random variables

Coupon Collection Problem.

Exercise

One of four different prizes was randomly put into each box of a cereal.

If a family decided to buy this cereal until it obtained at least one of each of the four different prizes, what is the expected number of boxes of cereal that must be purchased?



$X_1 = 1$ $\qquad$ $X_2 \sim \text{Geom}\left(\frac{3}{4}\right)$ $\qquad$ $X_3 \sim \text{Geom}\left(\frac{1}{2}\right)$ $\qquad$ $X_4 \sim \text{Geom}\left(\frac{1}{4}\right)$

$1 = \frac{4}{4}$ $\qquad$ $E[X_2] = \frac{4}{3}$ $\qquad$ $E[X_3] = 2 = \frac{4}{2}$ $\qquad$ $E[X_4] = 4 = \frac{4}{1}$

$$\text{ANS} = \frac{4}{4} + \frac{4}{3} + \frac{4}{2} + \frac{4}{1} = 4 \cdot \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)$$

$$= \frac{25}{3}.$$

Section 7.
The Poisson Distribution

Definition

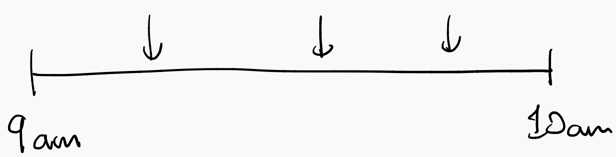
Some experiments result in counting the number of times particular events occur at given times or with given physical objects.

Example

- the number of cell phone calls passing through a relay tower between 9 and 10am.
- the number of flaws in 100 feet of wire
- the number of customers that arrive at a ticket window between noon and 2pm.
- the number of defects in a 100-foot roll of aluminum screen that is 2 feet wide.

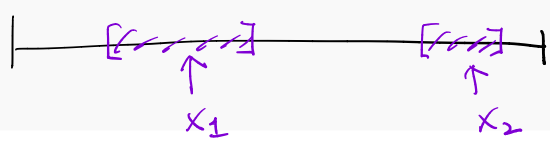
Definition

Counting such events can be looked upon as observations of a random variable associated with an **approximate Poisson process**, provided that the conditions in the following definition are satisfied.



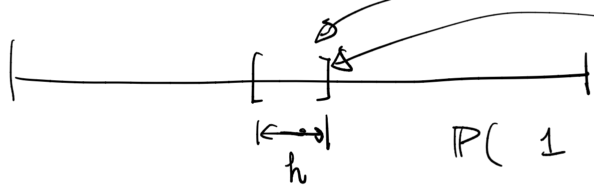
$X = \#$ of customers between $9 - 10$.

①



no overlap $\Longrightarrow$ indep.

②

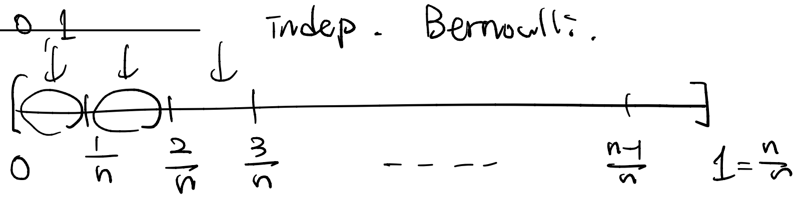


$\mathbb{P}($ 1 customer here $) \approx \lambda \cdot h^2$

③ $\mathbb{P}($ more than 1 customer here $) \approx 0$

0 1

Indep. Bernoulli.



$X = \#$ of customer $\approx \underline{Bin(n, p)}$

$E[X] = \lambda = np$, $p = \frac{\lambda}{n}$

$X \approx Bin(n, \underset{p}{\frac{\lambda}{n}}) \xrightarrow[n \to \infty]{} \underline{\underline{\text{Poisson}}}$

$\binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \longrightarrow e^{-\lambda} \frac{\lambda^x}{x!}$

Definition

Let the number of occurrences of some event in a given continuous interval be counted. Then we have an **approximate Poisson process** with parameter $\lambda > 0$ if

- The numbers of occurrences in nonoverlapping subintervals are independent.
- The probability of exactly one occurrence in a sufficiently short subinterval of length h is approximately λh.
- The probability of two or more occurrences in a sufficiently short subinterval is essentially zero.

Under these assumption, consider the number of occurrences in a time interval $[0, 1]$.

Definition

Split $[0, 1]$ into n subintervales $[0, \frac{1}{n}], [\frac{1}{n}, \frac{2}{n}], \cdots, [\frac{n-1}{n}, 1]$.

In each subinterval, at most one event occurs with probability $\frac{\lambda}{n}$.

Thus, the number of occurrences is a binomial random variable with n and $\frac{\lambda}{n}$.

As $n \to \infty$, the random variable gets close to some random variable X.

We say X is a **Poisson random variable with parameter λ** if its pmf is

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for $k = 0, 1, 2, \cdots$.

Definition

$$X \sim \mathrm{Pois}(\lambda)$$

$$f(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Theorem

$\mathbb{E}[X] = \lambda$

$k = 0, 1, 2, \ldots$

$\mathrm{Var}[X] = \lambda$

$M(t) = e^{\lambda(e^t - 1)}$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$\mathbb{E}[X] = \sum_{k=\cancel{0}\,1}^{\infty} k \cdot \boxed{e^{-\lambda}} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{(k-1)} \cdot \lambda}{(k-1)!}$$

$$= \lambda \cdot e^{-\lambda} \left(\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \right) = \lambda \, \underbrace{e^{-\lambda}}\, \underbrace{e^{\lambda}} = \lambda.$$

$n = k-1$

X = # of customers between 9 – 10. $\sim \text{Pois}(\lambda)$

- $f(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ $k = 0, 1, 2, \cdots$
- $\mathbb{E}[X] = \lambda$, $\text{Var}(X) = \lambda$, $M(t) = e^{\lambda(e^t - 1)}$

Definition

Example

In a large city, telephone calls to 911 come on the average of two every 3 minutes.

If one assumes an approximate Poisson distribution, what is the probability of five or more calls arriving in a 9 minute period?

X = # of calls in 9 min $\sim \text{Pois}(6)$

$$\mathbb{P}(X \geq 5) = \sum_{k=5}^{\infty} f(k) = \sum_{k=5}^{\infty} e^{-6} \cdot \frac{6^k}{k!}$$

use table

Poisson Approximation to Binomial

Supose X is a binomial random variable $b(n, p)$, n is large, and p is small but np converges to some constant, say λ.

In this case, X can be approximated by a Poisson random variable with parameter λ.

This approximation is quite accurate if $n \geq 20$, $p \leq 0.05$ or $n \geq 100$, $p \leq 0.1$.

$$X \sim \mathrm{Bin}(\underset{\substack{\uparrow \\ \text{large}}}{n}, \underset{\substack{\uparrow \\ \text{small}}}{p}) \approx \mathrm{Pois}(\lambda)$$

$\lambda = np$

Ex $X \sim \mathrm{Bin}(1000, 0.99)$ X

$1000 - X = Y \sim \mathrm{Bin}(1000, 0.01) \approx \mathrm{Pois}(10)$

Poisson Approximation to Binomial

Example

A manufacturer of Christmas tree light bulbs knows that 2% of its bulbs are defective.

Assuming independence, X = the number of defective bulbs in a box of 100 bulbs has a binomial distribution with parameters n = 100 and p = 0.02.

Find the probability that a box of 100 of these bulbs contains at most three defective bulbs.

$X \sim Bin(100, 0.02) \approx Pois(2)$, $2 = 100 \cdot 0.02$

$$\mathbb{P}(X \le 3) = \sum_{k=0}^{3} \binom{100}{k} (0.02)^k (0.98)^{100-k}$$

$$\approx \sum_{k=0}^{3} e^{-2} \frac{2^k}{k!}$$

$$= e^{-2} \cdot \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} \right)$$

$$= e^{-2} \cdot \left(1 + 2 + 2 + \frac{8}{6} \right) = \frac{19}{3} e^{-2}.$$

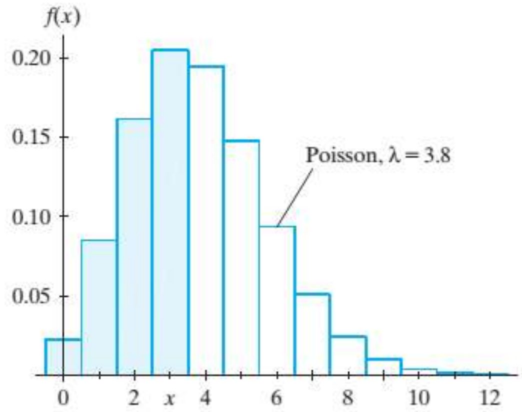
$$\approx 0.857.$$

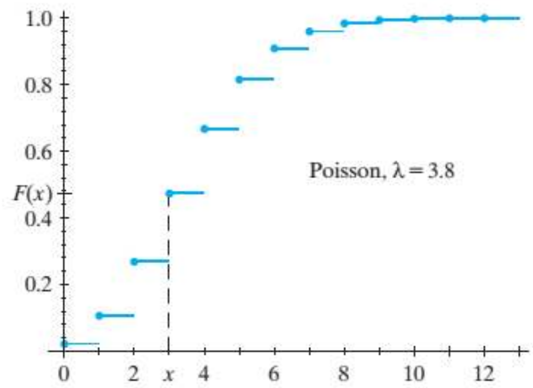
table.

$X \sim \mathrm{Pois}(6)$

$\mathbb{P}(X \geq 5)$

Table III The Poisson Distribution





CDF $\rightarrow$ $F(x) = P(X \leq x) = \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!}$

$X \sim \mathrm{Pois}(\lambda)$

$\lambda = E(X)$

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.333	0.301	0.273	0.247	0.223	0.202	0.183	0.165	0.150	0.135
1	0.699	0.663	0.627	0.592	0.558	0.525	0.493	0.463	0.434	0.406
2	0.900	0.879	0.857	0.833	0.809	0.783	0.757	0.731	0.704	0.677
3	0.974	0.966	0.957	0.946	0.934	0.921	0.907	0.891	0.875	0.857
4	0.995	0.992	0.989	0.986	0.981	0.976	0.970	0.964	0.956	0.947
5	0.999	0.998	0.998	0.997	0.996	0.994	0.992	0.990	0.987	0.983
6	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.997	0.995
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	2.2	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0
0	0.111	0.091	0.074	0.061	0.050	0.041	0.033	0.027	0.022	0.018
1	0.355	0.308	0.267	0.231	0.199	0.171	0.147	0.126	0.107	0.092
2	0.623	0.570	0.518	0.469	0.423	0.380	0.340	0.303	0.269	0.238
3	0.819	0.779	0.736	0.692	0.647	0.603	0.558	0.515	0.473	0.433
4	0.928	0.904	0.877	0.848	0.815	0.781	0.744	0.706	0.668	0.629
5	0.975	0.964	0.951	0.935	0.916	0.895	0.871	0.844	0.816	0.785
6	0.993	0.988	0.983	0.976	0.966	0.955	0.942	0.927	0.909	0.889
7	0.998	0.997	0.995	0.992	0.988	0.983	0.977	0.969	0.960	0.949
8	1.000	0.999	0.999	0.998	0.996	0.994	0.992	0.988	0.984	0.979
9	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.992
10	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$F(x) = P(X \le x)$

$= 1 - 0.285.$

$P(X \ge 5) = 1 - P(X \le 4)$

Table III *continued*

x	4.2	4.4	4.6	4.8	5.0	5.2	5.4	5.6	5.8	6.0
0	0.015	0.012	0.010	0.008	0.007	0.006	0.005	0.004	0.003	0.002
1	0.078	0.066	0.056	0.048	0.040	0.034	0.029	0.024	0.021	0.017
2	0.210	0.185	0.163	0.143	0.125	0.109	0.095	0.082	0.072	0.062
3	0.395	0.359	0.326	0.294	0.265	0.238	0.213	0.191	0.170	0.151
4	0.590	0.551	0.513	0.476	0.440	0.406	0.373	0.342	0.313	0.285
5	0.753	0.720	0.686	0.651	0.616	0.581	0.546	0.512	0.478	0.446
6	0.867	0.844	0.818	0.791	0.762	0.732	0.702	0.670	0.638	0.606
7	0.936	0.921	0.905	0.887	0.867	0.845	0.822	0.797	0.771	0.744
8	0.972	0.964	0.955	0.944	0.932	0.918	0.903	0.886	0.867	0.847
9	0.989	0.985	0.980	0.975	0.968	0.960	0.951	0.941	0.929	0.916
10	0.996	0.994	0.992	0.990	0.986	0.982	0.977	0.972	0.965	0.957
11	0.999	0.998	0.997	0.996	0.995	0.993	0.990	0.988	0.984	0.980
12	1.000	0.999	0.999	0.999	0.998	0.997	0.996	0.995	0.993	0.991
13	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.996
14	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.999
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0
0	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.011	0.007	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000
2	0.043	0.030	0.020	0.014	0.009	0.006	0.004	0.003	0.002	0.001
3	0.112	0.082	0.059	0.042	0.030	0.021	0.015	0.010	0.007	0.005
4	0.224	0.173	0.132	0.100	0.074	0.055	0.040	0.029	0.021	0.015
5	0.369	0.301	0.241	0.191	0.150	0.116	0.089	0.067	0.050	0.038
6	0.527	0.450	0.378	0.313	0.256	0.207	0.165	0.130	0.102	0.079
7	0.673	0.599	0.525	0.453	0.386	0.324	0.269	0.220	0.179	0.143
8	0.792	0.729	0.662	0.593	0.523	0.456	0.392	0.333	0.279	0.232
9	0.877	0.830	0.776	0.717	0.653	0.587	0.522	0.458	0.397	0.341
10	0.933	0.901	0.862	0.816	0.763	0.706	0.645	0.583	0.521	0.460
11	0.966	0.947	0.921	0.888	0.849	0.803	0.752	0.697	0.639	0.579
12	0.984	0.973	0.957	0.936	0.909	0.876	0.836	0.792	0.742	0.689
13	0.993	0.987	0.978	0.966	0.949	0.926	0.898	0.864	0.825	0.781
14	0.997	0.994	0.990	0.983	0.973	0.959	0.940	0.917	0.888	0.854
15	0.999	0.998	0.995	0.992	0.986	0.978	0.967	0.951	0.932	0.907
16	1.000	0.999	0.998	0.996	0.993	0.989	0.982	0.973	0.960	0.944
17	1.000	1.000	0.999	0.998	0.997	0.995	0.991	0.986	0.978	0.968
18	1.000	1.000	1.000	0.999	0.999	0.998	0.096	0.993	0.988	0.982
19	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.994	0.991
20	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.995
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table III *continued*

x	11.5	12.0	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.003	0.002	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000
4	0.011	0.008	0.005	0.004	0.003	0.002	0.001	0.001	0.001	0.000
5	0.028	0.020	0.015	0.011	0.008	0.006	0.004	0.003	0.002	0.001
6	0.060	0.046	0.035	0.026	0.019	0.014	0.010	0.008	0.006	0.004
7	0.114	0.090	0.070	0.054	0.041	0.032	0.024	0.018	0.013	0.010
8	0.191	0.155	0.125	0.100	0.079	0.062	0.048	0.037	0.029	0.022
9	0.289	0.242	0.201	0.166	0.135	0.109	0.088	0.070	0.055	0.043
10	0.402	0.347	0.297	0.252	0.211	0.176	0.145	0.118	0.096	0.077
11	0.520	0.462	0.406	0.353	0.304	0.260	0.220	0.185	0.154	0.127
12	0.633	0.576	0.519	0.463	0.409	0.358	0.311	0.268	0.228	0.193
13	0.733	0.682	0.629	0.573	0.518	0.464	0.413	0.363	0.317	0.275
14	0.815	0.772	0.725	0.675	0.623	0.570	0.518	0.466	0.415	0.368
15	0.878	0.844	0.806	0.764	0.718	0.669	0.619	0.568	0.517	0.467
16	0.924	0.899	0.869	0.835	0.798	0.756	0.711	0.664	0.615	0.566
17	0.954	0.937	0.916	0.890	0.861	0.827	0.790	0.749	0.705	0.659
18	0.974	0.963	0.948	0.930	0.908	0.883	0.853	0.819	0.782	0.742
19	0.986	0.979	0.969	0.957	0.942	0.923	0.901	0.875	0.846	0.812
20	0.992	0.988	0.983	0.975	0.965	0.952	0.936	0.917	0.894	0.868
21	0.996	0.994	0.991	0.986	0.980	0.971	0.960	0.947	0.930	0.911
22	0.999	0.997	0.995	0.992	0.989	0.983	0.976	0.967	0.956	0.942
23	0.999	0.999	0.998	0.996	0.994	0.991	0.986	0.981	0.973	0.963
24	1.000	0.999	0.999	0.998	0.997	0.995	0.992	0.989	0.984	0.978
25	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.991	0.987
26	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.995	0.993
27	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996
28	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.998
29	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
31	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
32	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
33	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
34	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
35	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Poisson Approximation to Binomial

Exercise

Suppose that the probability of suffering a side effect from a certain flu vaccine is 0.005. If 1000 persons are vaccinated, approximate the probability that (a) At most one person suffers. (b) Four, five, or six persons suffer.

Ch1 - #10, 5, 4 , 5 #5.

$\boxed{\#10}$ 2R, 4W Sample 5 with replacement.

P(No two balls drawn consecutively have the same color)

$= P(WRWRW \text{ or } RWRWR)$

$= 1\cdot\left(\frac{2}{6}\right)^2\cdot\left(\frac{4}{6}\right)^3 + 1\left(\frac{2}{6}\right)^3\cdot\left(\frac{4}{6}\right)^2$

$= \left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^3 + \left(\frac{1}{3}\right)^3\cdot\left(\frac{2}{3}\right)^2 = \frac{8+4}{3^5} = \frac{12}{243}$

without replacement

$= P(WRWRW) = \frac{4}{6}\cdot\frac{2}{5}\cdot\frac{3}{4}\cdot\frac{1}{3}\cdot\frac{2}{2} = \frac{1}{15}.$

$\boxed{\#5}$ 5G, 7R $\longrightarrow$ Sample 9.

$A = \{\text{First 3 balls} = G\}$

$B = \{\text{Exactly } 4G\} = \{4G, 5R\}$

(a) With Replacement: $P(B)$, $P(A\cap B) = ?$

$P(B) = P(X=4) = \binom{9}{4}\left(\frac{5}{12}\right)^4\cdot\left(\frac{7}{12}\right)^5.$

$X = \#$ of G in 9 sample $\sim Bin\left(9, \frac{5}{12}\right)$

Note: $P(A) = \left(\frac{5}{12}\right)^3.$

$P(A\cap B) = P(A)\cdot P(B)$? $\leftarrow$ No. A, B Not indep.

$P(A\cap B) = \binom{6}{1}\left(\frac{5}{12}\right)^4\cdot\left(\frac{7}{12}\right)^5$ "

GGG $\boxed{6}$

1G, 5R

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{\binom{6}{1}}{\binom{9}{4}} = \frac{6}{\frac{9\cdot 8\cdot 7\cdot 6}{4!}}$$

$$= \frac{1}{21}.$$

(b) without replacement.

$$\mathbb{P}(B) = \mathbb{P}(Y=4) = \frac{\binom{5}{4}\binom{7}{5}}{\binom{12}{9}}$$

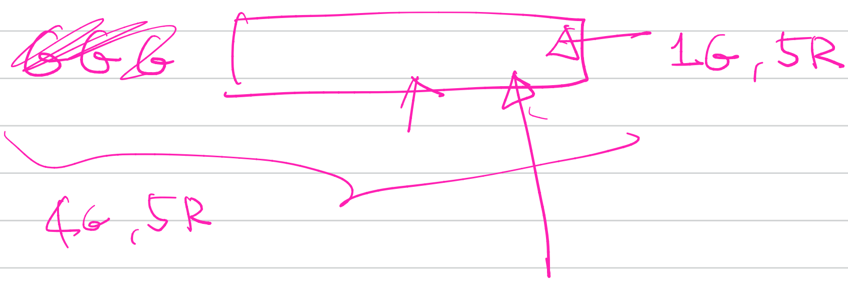
$\uparrow$ 4G, 5R

$Y = \#$ of G in 9 samples

$\sim HG(5, 7, 9)$

$$\mathbb{P}(A) = \frac{\cancel{5}}{\cancel{12}}\cdot\frac{4}{11}\cdot\frac{3}{\cancel{10}_{2}} = \frac{1}{22}$$

$$\mathbb{P}(B|A) = \mathbb{P}(Z=1) = \frac{\binom{2}{1}\binom{7}{5}}{\binom{9}{6}}$$



$Z = \#$ of G in last 6 $\sim HG(2, 7, 6)$

#15 $$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\cdot\mathbb{P}(A)}{\mathbb{P}(B)}$$

#4 WS in baseball A, N

win 4 games first → winner.

(a) Four-game : $\begin{cases} AAAA \\ NNNN \end{cases}$

3A, 0N

$\underbrace{\boxed{}}_{3 \text{ game}} \underline{\underline{A}} \quad \times 2$

(b) 5-game : 3A 1N

$\underbrace{\boxed{}}_{4 \text{ games}} A$

$\binom{4}{3} \times 2$

(c) 6-game : $\underbrace{\boxed{}}_{3A,\ 2N} A$

$\binom{5}{3} \times 2$

(d) 7-game : $\binom{6}{3} \cdot 2$

Sample #5

$$M(t) = \underbrace{c}(1-2t)^{-5}, \quad t<\tfrac{1}{2}$$

(a) $c = ?$

$M(t) = \mathbb{E}[e^{tX}]$

$M(0) = E[e^0] = 1 = c(1-0)^{-5} = c \quad \therefore c=1.$

(b) $E[X] = M'(0) = (-5)\cdot(1-2t)^{-6}\cdot(1-2t)'\Big|_{t=0}$

$= \underbrace{(-5)\cdot(-2)}\cdot\underbrace{(1-2t)^{-6}}_{=1.}\Big|_{t=0}$

$= 10.$

(c) $Var(X) = \underbrace{E[X^2]} - (E[X])^2 = 120 - (10)^2 = 20.$

$E[X^2] = M''(0) = 10\cdot(-6)\cdot(1-2t)^{-7}\cdot(-2)\Big|_{t=0}$

$= 120$

Sample #4

$$f(k) = \begin{cases} c & k=0 \\ \dfrac{1}{3^k\, k!} & k=1,2,\dots \end{cases}$$

$c = ?$

$$1 = \sum_{k=0}^{\infty} f(k) = c + \underbrace{\sum_{k=\cancel{1}\,0}^{\infty} \frac{1}{k!}\left(\frac{1}{3}\right)^k}_{= e^{\frac{1}{3}} - 1} - 1$$

$2 - e^{\frac{1}{3}} = c.$

$$\boxed{e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$ → Binomial

$$a + a \cdot r + ar^2 + \cdots$$ → Geom.

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$ → Poisson.