Statistics 1 Chapter 6: Statistical Distributions

A variable is a collection of values (in statistics)

If each value is assigned a probability, we get a **RANDOM VARIABLE** represented by "X"

x	red	green	blue	orange
P(X=x)	0.3	0.4	0.1	0.2

probability of the random variable being x (ie X=x)
(Also written as p(x))

Mapping OUTCOMES to PROBABILITY:

Function

$$p(x) = \begin{cases} 0.1x & , x = 1,2,3,4 \\ 0 & , \text{otherwise} \end{cases}$$

(when x=1, p(x)=0.1
when x=2, p(x)=0.2 ... etc
when x=5, p(x)=0)

Diagram (discrete)

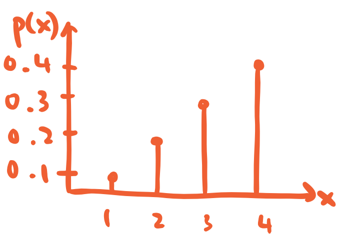
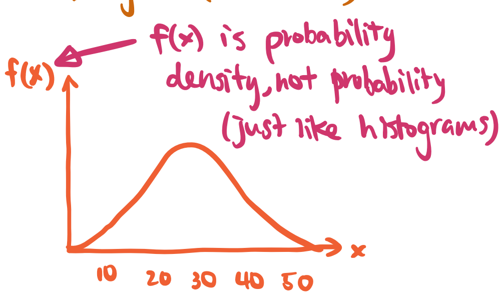


Diagram (continuous)



Table

x	1	2	3	4
p(x)	0.1	0.2	0.3	0.4

EXAMPLE: X = # of heads in 3 coin tosses.

Sample space

{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

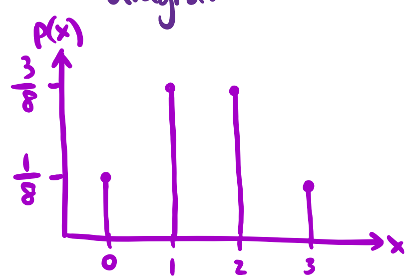
function

$$p(x) = \begin{cases} \frac{1}{8} & x = 0,3 \\ \frac{3}{8} & x = 1,2 \\ 0 & \text{otherwise} \end{cases}$$

table

x	0	1	2	3
P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

diagram



EX 6A (p.86) (Ans p.219)

① a) No, because height is a continuous quantity

b) Yes, because it is always a whole number that can vary

c) No, because the value doesn't vary (7 days a week)

② 4 dice throws, $y = \#$ of 6 Sample space $= \{1, 2, 3, 4\}$

③ Bag ⇒ 2x"2", 2x"3"
draw & return 3 times

a) $\{[2,2], [2,3], [3,2], [3,3]\}$

b) X = sum of 2 #s

i)

x	4	5	6
$p(x)$	$1/4$	$1/2$	$1/4$

ii)
$$p(x) = \begin{cases} 1/4 & x = 4, 6 \\ 1/2 & x = 5 \\ 0 & \text{otherwise} \end{cases}$$

④ $k = 1 - \frac{1}{3} - \frac{1}{3} - \frac{1}{4} = 1 - \frac{8}{12} - \frac{3}{12} = \frac{1}{12}$

6.2: Binomial Distribution

$X \sim B(n,p)$ — "distributed as"

X: number of successful trials
n: total number of trials
p: probability of a trial succeeding

conditions: n is fixed (# of trials don't vary)
p is fixed (trials are independent and identical)
there are only 2 outcomes (success & failure)

$$\boxed{P(X=r) = {}^nC_r\, p^r (1-p)^{n-r}} \quad (X \sim B(n,p))$$

Explaining the formula: $P(X=r) = {}^nC_r\, p^r (1-p)^{n-r}$

- ${}^nC_r$: # of possibilities we can arrange the the success-fail pattern (eg. $\underbrace{101, 110, 011}$) all X=2
- p^r: probability of r successful trials
- $(1-p)^{n-r}$: probability of (n-r) failed trials

example: dice throwing

12 throws: 6 ⇒ success $\quad X \sim B(12, \tfrac{1}{6})$
1-5 ⇒ failure

Solve using 1 of 2 methods:

① Formula:

$P(X=r) = {}^nC_r\, p^r (1-p)^{n-r}$

$P(X=2) = {}^{12}C_2 (\tfrac{1}{6})^2 (\tfrac{5}{6})^{10}$

$= 0.296$ (3 sf)

② Calculator:

Menu → "Distribution" (7)
Choose "Binomial PD" (4)
Choose "Variable" (2)
Substitute x, n and p
⇓
0.296 (3sf)

Ex 6B (p. 90)

① $X \sim B(8, \frac{1}{3})$

a) $P(X=2) = {}^8C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^6 = 0.273$

b) $P(X=5) = {}^8C_5 \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^3 = 0.0683$

c) $P(X \leq 1) = {}^8C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^8 + {}^8C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^7 = 0.039 + 0.156 = 0.195$

② $T \sim B(15, \frac{2}{3})$

a) $P(T=5) = 6.70 \times 10^{-3} = 0.00670$

b) $P(T=10) = 0.214$

c) $P(3 \leq T \leq 4) = P(T=3) + P(T=4) = 2.54 \times 10^{-4} + 1.52 \times 10^{-3} = 1.77 \times 10^{-3}$

③ a) $X \sim B(20, 0.01)$ where X = defective bolts

assumptions: defective bolts aren't made in groups (independent)

b) $X \sim B(6, 0.52)$ where X = number of wait & stops

assumptions: the wait, stop and go lights are completely random

c) $X \sim B(30, 0.125)$ where X = number of aces in the next 30 serves

assumptions: aces are completely random (they don't give confidence and increase prob.)

④ a) Yes, because only 2 outcomes, not in family ($\therefore$ no inheritance)

b) No, because there are more than 2 outcomes

c) Yes, because the trials are random & independent. Only 2 outcomes.

⑤ a) $X \sim B(20, 0.05)$ where X = # of balloons that burst

$P(X=0) = {}^{20}C_0 (0.05)^0 (0.95)^{20} = 0.358$

b) $P(X=2) = {}^{20}C_2 (0.05)^2 (0.95)^{18} = 0.189$

What if X is in a range? (e.g. $P(X \leq 6)$)

let $X \sim B(10, 0.3)$

$$P(X \leq 6) = P(X=0) + P(X=1) + \dots + P(X=6) = \sum_{i=0}^{6} P(X=i)$$

3 ways to solve:

① Summation

Type into calculator
or do it by hand

② Distribution ⇒ Binomial CD ("C" stands for cumulative)

Type x, N and p, and will find $P(X \leq x)$

③ Table (p. 204 of textbook)

Look up corresponding values

If $P(X \geq x)$, you will have to rewrite it!

$P(X \geq x) = 1 - P(X \leq x-1)$

e.g. $P(X \geq 5) = 1 - P(X \leq 4)$

Reverse?

Example: spinner → $\geq r$ reds will win prize, prize win chance < 0.05

% chance of red $= 0.3$, 12 spins what is r?

$P(X \geq r) < 0.05$ ⟵ represent the question

$1 - P(X \leq r-1) < 0.05$ ⟵ change inequality sign

$P(X \leq r-1) > 0.95$ ⟵ rearrange

$r-1 = 6$ ⟵ find in table the closest value to 0.95 but higher (0.9857)

$r = 7$ ⟵ yay!