# STATISTICS - III

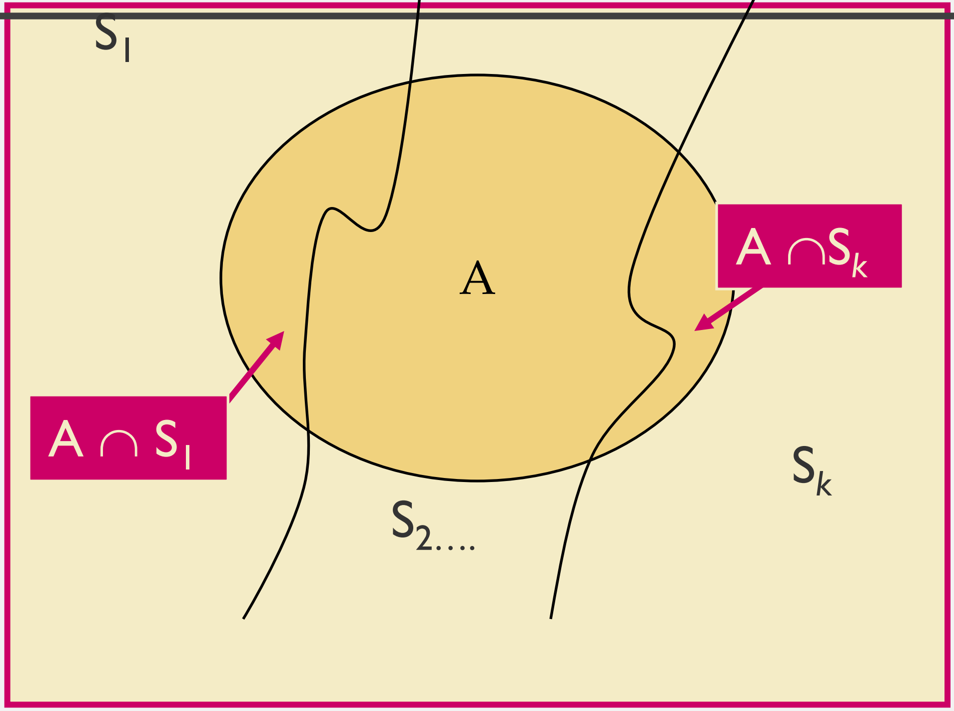
# THE LAW OF TOTAL PROBABILITY

Let $S_1$ , $S_2$ , $S_3$ ,..., $S_k$ be mutually exclusive and exhaustive events (that is, one and only one must happen). Then the probability of any event A can be written as

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \dots + P(A \cap S_k)$$

$$= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \dots + P(S_k)P(A|S_k)$$

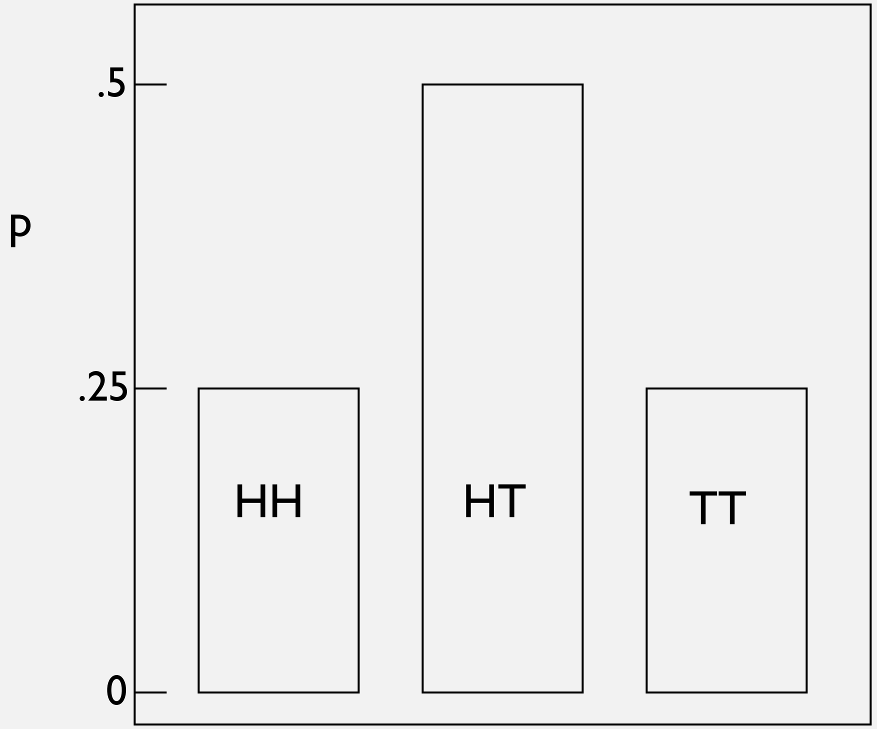
# THE LAW OF TOTAL PROBABILITY



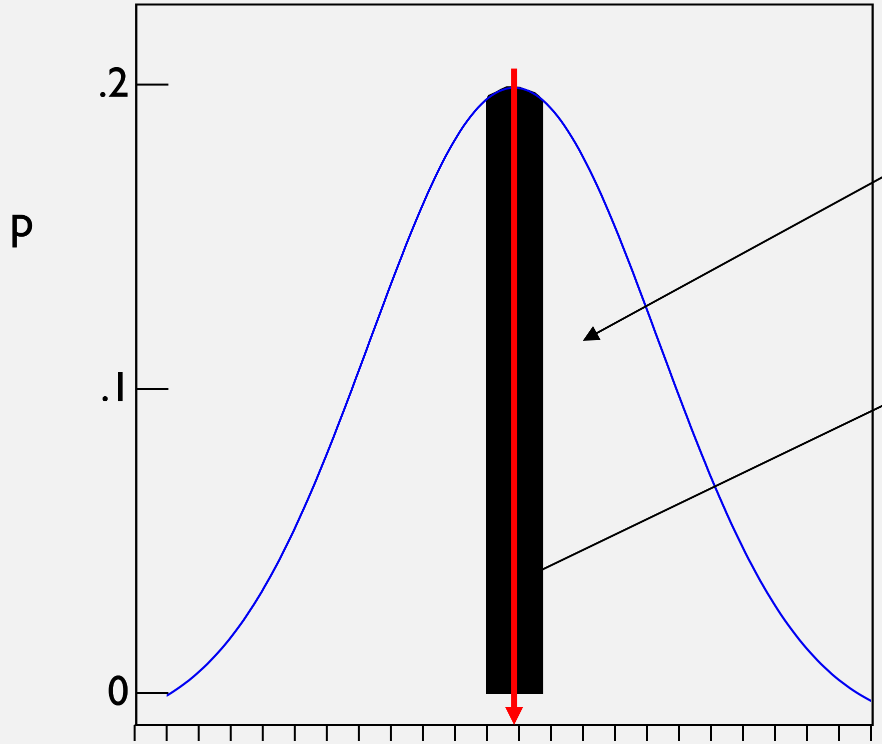

$$P(A) = P(A \cap S_1) + P(A \cap S_2) + \ldots + P(A \cap S_k)$$

$$= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \ldots + P(S_k)P(A|S_k)$$

# DISCRETE PROBABILITIES



# CONTINUOUS PROBABILITIES




total area under curve = 1

<u>but</u>

the probability of any ***single*** value = 0

∴ interested in the probability assoc. w/ <u>intervals</u>

# INDEPENDENT EVENTS

- one event has no influence on the outcome of another event
- if events A & B are independent

  then P(A&B) = P(A)*P(B)
- if P(A&B) = P(A)*P(B)

  then events A & B are independent
- coin flipping

  if P(H) = P(T) = .5 then

  P(HTHTH) = P(HHHHH) =

  $.5 * .5 * .5 * .5 * .5 = .5^5 = .03$

# CONDITIONAL PROBABILITY

- concern the odds of one event occurring, given that another event has occurred

- $P(A|B)$=Prob of A, given B

# BAYES' RULE

Let $S_1$ , $S_2$ , $S_3$ ,..., $S_k$ be mutually exclusive and exhaustive events with prior probabilities $P(S_1)$, $P(S_2)$,…,$P(S_k)$. If an event A occurs, the posterior probability of $S_i$, given that A occurred is

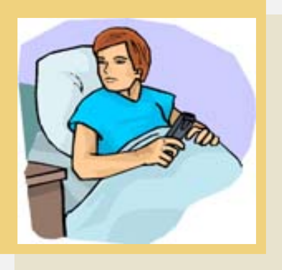
$$P(S_i \mid A) = \frac{P(S_i)P(A \mid S_i)}{\sum P(S_i)P(A \mid S_i)} \text{ for } i = 1, 2, \ldots k$$

Proof

$$P(A \mid S_i) = \frac{P(AS_i)}{P(S_i)} \longrightarrow P(AS_i) = P(S_i)P(A \mid S_i)$$

$$P(S_i \mid A) = \frac{P(AS_i)}{P(A)} = \frac{P(S_i)P(A \mid S_i)}{\sum P(S_i)P(A \mid S_i)}$$

# EXAMPLE



From a previous example, we know that 49% of the population are female. Of the female patients, 8% are high risk for heart attack, while 12% of the male patients are high risk. A single person is selected at random and found to be high risk. What is the probability that it is a male?

Define H: high risk   F: female   M: male

We know:

P(F) = .49

P(M) = .51

P(H|F) = .08

P(H|M) = .12

$$P(M \mid H) = \frac{P(M)P(H \mid M)}{P(M)P(H \mid M) + P(F)P(H \mid F)}$$

$$= \frac{.51(.12)}{.51(.12) + .49(.08)} = .61$$

# EXAMPLE

Suppose a rare disease infects one out of every 1000 people in a population. And suppose that there is a good, but not perfect, test for this disease: if a person has the disease, the test comes back positive 99% of the time. On the other hand, the test also produces some false positives: 2% of uninfected people are also test positive. And someone just tested positive. What are his chances of having this disease?

# EXAMPLE

Define A: has the disease B: test positive

We know:

$P(A) = .001$ $P(A^c) = .999$

$P(B|A) = .99$ $P(B|A^c) = .02$

We want to know $P(A|B) = ?$

$$P(A \mid B) = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^c)P(B|A^c)}$$

$$= \frac{.001 \times .99}{.001 \times .99 + .999 \times .02} = .0472$$

# EXAMPLE

A survey of job satisfaction[2] of teachers was taken, giving the following results

| | | Job Satisfaction | | |
|---|---|---|---|---|
| | | Satisfied | Unsatisfied | Total |
| LEVEL | College | 74 | 43 | 117 |
| | High School | 224 | 171 | 395 |
| | Elementary | 126 | 140 | 266 |
| | Total | 424 | 354 | 778 |

2 "Psychology of the Scientist: Work Related Attitudes of U.S. Scientists" (*Psychological Reports* (1991): 443 – 450).

# EXAMPLE

If all the cells are divided by the total number surveyed, 778, the resulting table is a table of empirically derived probabilities.

| | | Job Satisfaction | | |
|---|---|---|---|---|
| | | Satisfied | Unsatisfied | Total |
| LEVEL | College | 0.095 | 0.055 | 0.150 |
| | High School | 0.288 | 0.220 | 0.508 |
| | Elementary | 0.162 | 0.180 | 0.342 |
| | Total | 0.545 | 0.455 | 1.000 |

# EXAMPLE

|  |  | Job Satisfaction |  |  |
|---|---|---|---|---|
|  |  | Satisfied | Unsatisfied | Total |
| LEVEL | College | 0.095 | 0.055 | 0.150 |
|  | High School | 0.288 | 0.220 | 0.508 |
|  | Elementary | 0.162 | 0.180 | 0.342 |
|  | Total | 0.545 | 0.455 | 1.000 |

For convenience, let C stand for the event that the teacher teaches college, S stand for the teacher being satisfied and so on. Let's look at some probabilities and what they mean.

$P(C) = 0.150$ is the proportion of teachers who are college teachers.

$P(S) = 0.545$ is the proportion of teachers who are satisfied with their job.

$P(C \cap S) = 0.095$ is the proportion of teachers who are college teachers and who are satisfied with their job.

# EXAMPLE

| | | Job Satisfaction | | |
|---|---|---|---|---|
| | | Satisfied | Unsatisfied | Total |
| L E V E L | College | 0.095 | 0.055 | 0.150 |
| | High School | 0.288 | 0.220 | 0.508 |
| | Elementary | 0.162 | 0.180 | 0.342 |
| | Total | 0.545 | 0.455 | 1.000 |

$$P(C \mid S) = \frac{P(C \cap S)}{P(S)}$$

$$= \frac{0.095}{0.545} = 0.175$$

is the proportion of teachers who are college teachers given they are satisfied. Restated: This is the proportion of satisfied that are college teachers.

$$P(S \mid C) = \frac{P(S \cap C)}{P(C)}$$

$$= \frac{P(C \cap S)}{P(C)} = \frac{0.095}{0.150}$$

$$= 0.632$$

is the proportion of teachers who are satisfied given they are college teachers. Restated: This is the proportion of college teachers that are satisfied.

# EXAMPLE

|  |  | Job Satisfaction |  |  |
|---|---|---|---|---|
|  |  | Satisfied | Unsatisfied | Total |
| LEVEL | College | 0.095 | 0.055 | 0.150 |
|  | High School | 0.288 | 0.220 | 0.508 |
|  | Elementary | 0.162 | 0.180 | 0.342 |
|  | Total | 0.545 | 0.455 | 1.000 |

Are C and S independent events?

$$P(C) = 0.150 \text{ and } P(C\,|\,S) = \frac{P(C \cap S)}{P(S)} = \frac{0.095}{0.545} = 0.175$$

$P(C|S) \neq P(C)$ so C and S are dependent events.

# EXAMPLE

|  |  | Job Satisfaction |  |  |
|---|---|---|---|---|
|  |  | Satisfied | Unsatisfied | Total |
| LEVEL | College | 0.095 | 0.055 | 0.150 |
|  | High School | 0.288 | 0.220 | 0.508 |
|  | Elementary | 0.162 | 0.180 | 0.658 |
|  | Total | 0.545 | 0.455 | 1.000 |

$P(C \cup S)$?

$P(C) = 0.150$, $P(S) = 0.545$ and

$P(C \cap S) = 0.095$, so

$$P(C \cup S) = P(C)+P(S) - P(C \cap S)$$
$$= 0.150 + 0.545 - 0.095$$
$$= 0.600$$

# EXAMPLE

Tom and Dick are going to take a driver's test at the nearest DMV office. Tom estimates that his chances to pass the test are 70% and Dick estimates his as 80%. Tom and Dick take their tests independently.

Define D = {Dick passes the driving test}

T = {Tom passes the driving test}

T and D are independent.

$P(T) = 0.7$, $P(D) = 0.8$

# EXAMPLE

What is the probability that at most one of the two friends will pass the test?

**P(At most one person pass)**

$= P(D^c \cap T^c) + P(D^c \cap T) + P(D \cap T^c)$

$= (1 - 0.8)(1 – 0.7) + (0.7)(1 – 0.8) + (0.8)(1 – 0.7)$

$= .44$

**P(At most one person pass)**

$= 1 - P(\text{both pass}) = 1 - 0.8 \times 0.7 = .44$

# EXAMPLE

What is the probability that at least one of the two friends will pass the test?

P(At least one person pass)

$= P(D \cup T)$

$= 0.8 + 0.7 - 0.8 \times 0.7$

$= .94$

P(At least one person pass)

$= 1 - P(\text{neither passes}) = 1 - (1-0.8) \times (1-0.7) = .94$

# EXAMPLE

Suppose we know that only one of the two friends passed the test. What is the probability that it was Dick?

**P(D | exactly one person passed)**

**= P(D $\cap$ exactly one person passed) / P(exactly one person passed)**

**= $P(D \cap T^c) / (P(D \cap T^c) + P(D^c \cap T))$**

**= $0.8 \times (1-0.7)/(0.8 \times (1-0.7)+(1-.8) \times 0.7)$**

**= .63**

# RANDOM VARIABLES

- A quantitative variable *x* is a **random variable** if the value that it assumes, corresponding to the outcome of an experiment is a chance or random event.
- Random variables can be **discrete** or **continuous.**

- **Examples:**
  - ✓ $x$ = SAT score for a randomly selected student
  - ✓ $x$ = number of people in a room at a randomly selected time of day
  - ✓ $x$ = number on the upper face of a randomly tossed die

# PROBABILITY DISTRIBUTIONS FOR DISCRETE RANDOM VARIABLES

The **probability distribution for a discrete random variable *x*** resembles the relative frequency distributions we constructed in Chapter 2. It is a graph, table or formula that gives the possible values of *x* and the probability *p*(*x*) associated with each value.

We must have

$$0 \leq p(x) \leq 1 \text{ and } \sum p(x) = 1$$

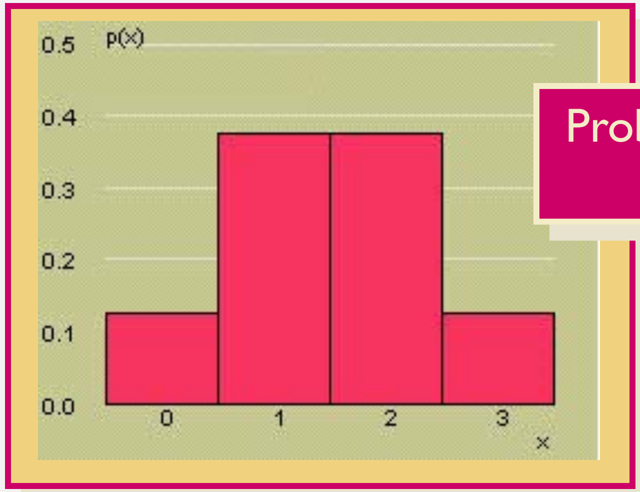
# EXAMPLE

Toss a fair coin three times and define ***x* = number of heads.**

| Outcome | Probability | $x$ |
|---|---|---|
| HHH | 1/8 | 3 |
| HHT | 1/8 | 2 |
| HTH | 1/8 | 2 |
| THH | 1/8 | 2 |
| HTT | 1/8 | 1 |
| THT | 1/8 | 1 |
| TTH | 1/8 | 1 |
| TTT | 1/8 | 0 |

$P(x = 0) = 1/8$
$P(x = 1) = 3/8$
$P(x = 2) = 3/8$
$P(x = 3) = 1/8$

| $x$ | $p(x)$ |
|---|---|
| 0 | 1/8 |
| 1 | 3/8 |
| 2 | 3/8 |
| 3 | 1/8 |

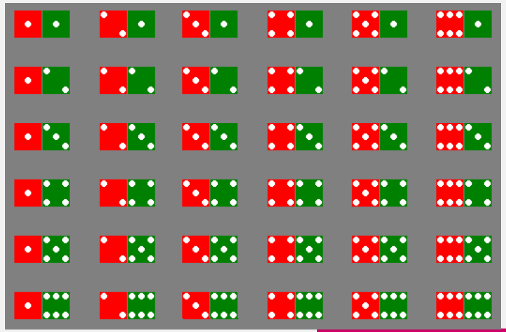


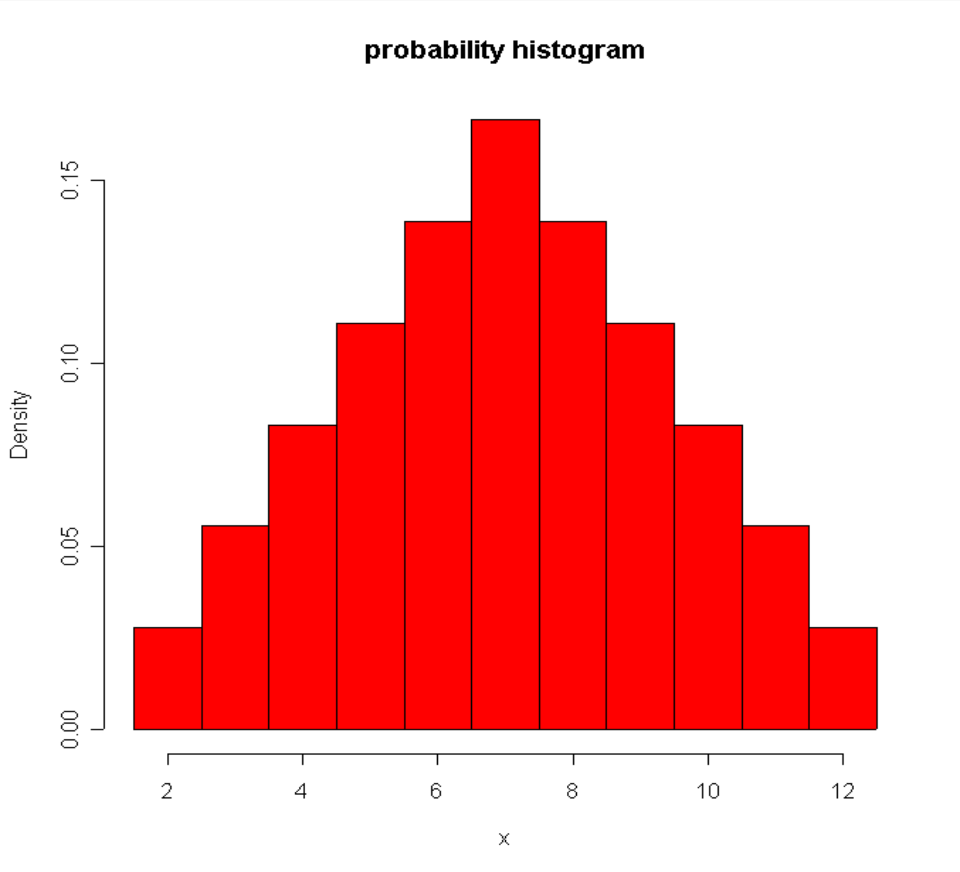

Probability Histogram for *x*

# EXAMPLE

Toss two dice and define

***x* = sum of two dice.**






| $x$ | $p(x)$ |
|---|---|
| 2 | 1/36 |
| 3 | 2/36 |
| 4 | 3/36 |
| 5 | 4/36 |
| 6 | 5/36 |
| 7 | 6/36 |
| 8 | 5/36 |
| 9 | 4/36 |
| 10 | 3/36 |
| 11 | 2/36 |
| 12 | 1/36 |

# PROBABILITY DISTRIBUTIONS

Probability distributions can be used to describe the population, just as we described samples in Chapter 2.

- **Shape:** Symmetric, skewed, mound-shaped…
- **Outliers:** unusual or unlikely measurements
- **Center and spread:** mean and standard deviation. A population mean is called $\mu$ and a population standard deviation is called $\sigma$.

# THE MEAN AND STANDARD DEVIATION

Let $x$ be a discrete random variable with probability distribution $p(x)$. Then the mean, variance and standard deviation of $x$ are given as

$$\text{Mean}: \mu = \sum xp(x)$$

$$\text{Variance}: \sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\text{Standard deviation}: \sigma = \sqrt{\sigma^2}$$

# EXAMPLE

Toss a fair coin 3 times and record *x* the number of heads.

| $x$ | $p(x)$ | $xp(x)$ | $(x-\mu)^2p(x)$ |
|---|---|---|---|
| 0 | 1/8 | 0 | $(-1.5)^2(1/8)$ |
| 1 | 3/8 | 3/8 | $(-0.5)^2(3/8)$ |
| 2 | 3/8 | 6/8 | $(0.5)^2(3/8)$ |
| 3 | 1/8 | 3/8 | $(1.5)^2(1/8)$ |

$$\mu = \sum xp(x) = \frac{12}{8} = 1.5$$

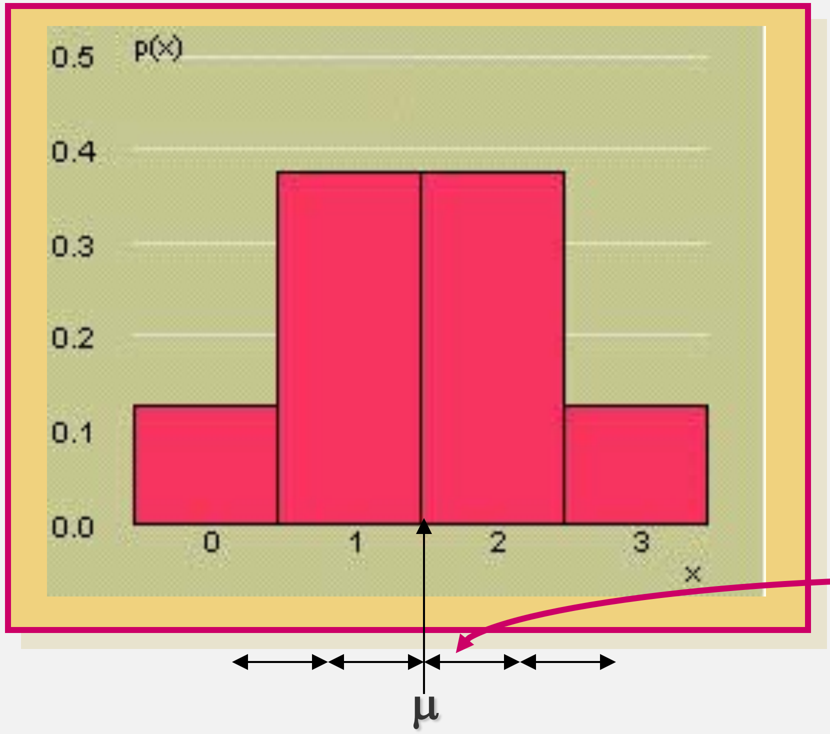
$$\sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\sigma^2 = .28125 + .09375 + .09375 + .28125 = .75$$

$$\sigma = \sqrt{.75} = .688$$

# EXAMPLE

The probability distribution for $x$ the number of heads in tossing 3 fair coins.




- Shape? — Symmetric; mound-shaped
- Outliers? — None
- Center? — $\mu = 1.5$
- Spread? — $\sigma = .688$

# SUMMARY

# KEY CONCEPTS

## I. Experiments and the Sample Space

1. Experiments, events, mutually exclusive events, simple events
2. The sample space

## II. Probabilities

1. Relative frequency definition of probability
2. Properties of probabilities
    a. Each probability lies between 0 and 1.
    b. Sum of all simple-event probabilities equals 1.
3. $P(A)$, the sum of the probabilities for all simple events in $A$

# KEY CONCEPTS

## III. Counting Rules

1. *mn* Rule; extended *mn* Rule
2. Permutations: $$P_r^n = \frac{n!}{(n-r)!}$$
3. Combinations: $$C_r^n = \frac{n!}{r!(n-r)!}$$

## IV. Event Relations

1. Unions and intersections
2. Events
   a. Disjoint or mutually exclusive: $P(A \cap B) = 0$
   b. Complementary: $P(A) = 1 - P(A^C)$

# KEY CONCEPTS

3. Conditional probability: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

4. Independent and dependent events

5. Additive Rule of Probability:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

6. Multiplicative Rule of Probability:

$$P(A \cap B) = P(A)P(B \mid A)$$

7. Law of Total Probability

8. Bayes' Rule