CM 3.2:
Probability Functions

The probability triple
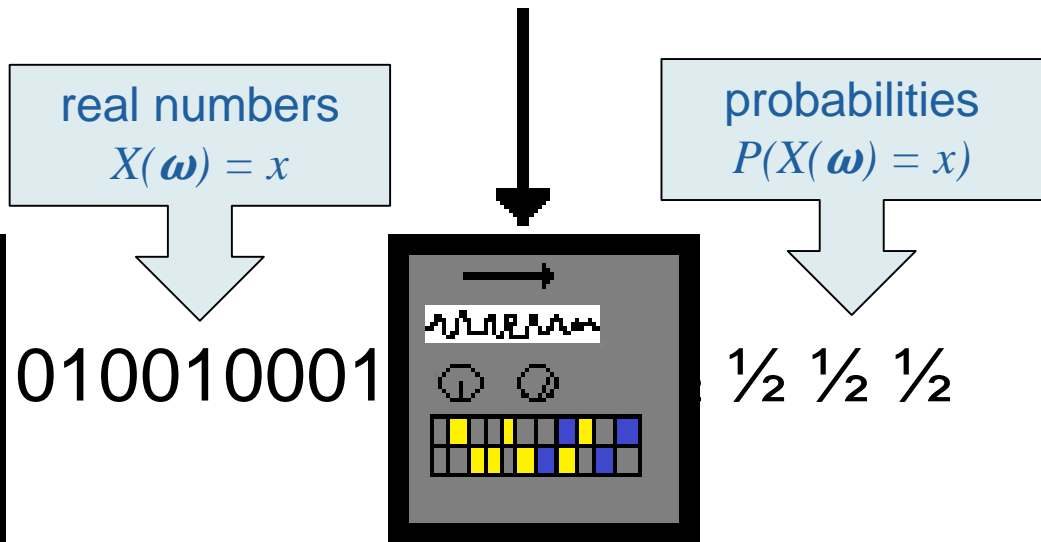
- Given a random experiment…

| Ω	Sample space	What can possibly happen?
$\mathcal{F}$	Set of events	What are the sane questions I can ask about this probability distribution?
$\mathbf{P}$	Probability measure	Function that maps elements from $\mathcal{F}$ to the interval [0,1]

RANDOM VARIABLE
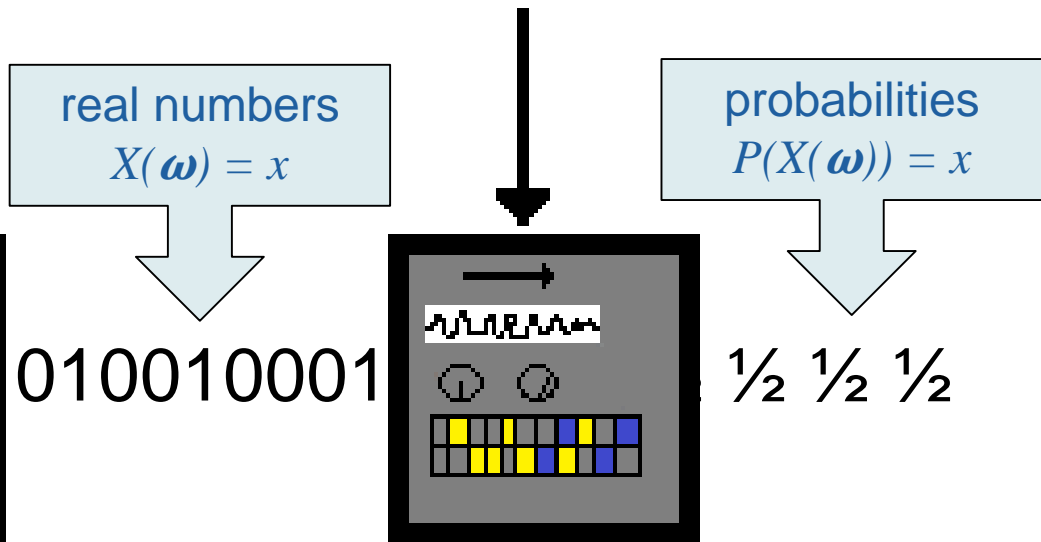
PROBABILITY DISTRIBUTION FUNCTION

ω

real numbers
$X(\omega) = x$

probabilities
$P(X(\omega) = x)$

TTHHTH

010010001

½ ½ ½
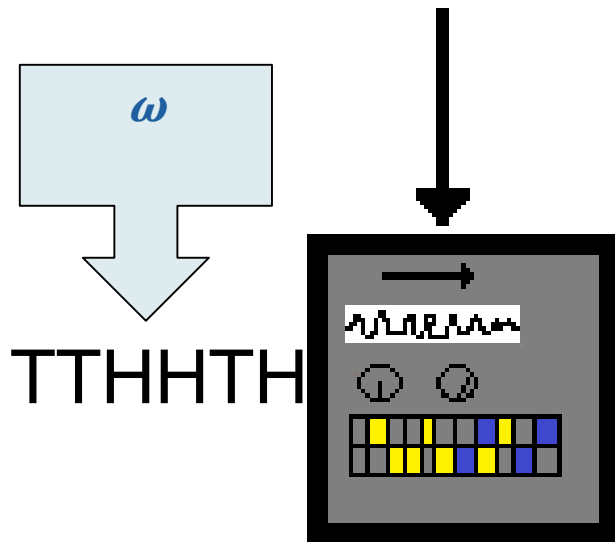
A function that associates an outcome with a real number.

Random Variable

RANDOM VARIABLE

PROBABILITY DISTRIBUTION FUNCTION

$\boldsymbol{\omega}$

real numbers
$X(\boldsymbol{\omega}) = x$

probabilities
$P(X(\boldsymbol{\omega})) = x$

TTHHTH

010010001

½ ½ ½
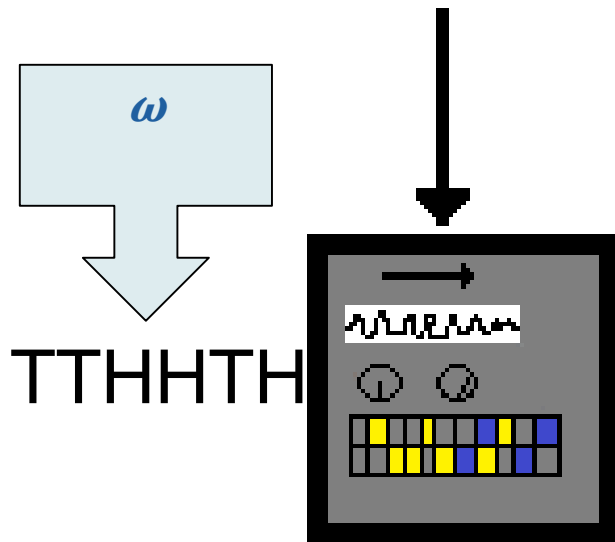
Function that maps the output of a random variable to the interval [0, 1] in probability space.

Probability Distribution Function

Discrete or continuous?

- A continuous rv can take any value in an interval or collection of intervals

- A discrete rv can take one of a countable list of distinct values
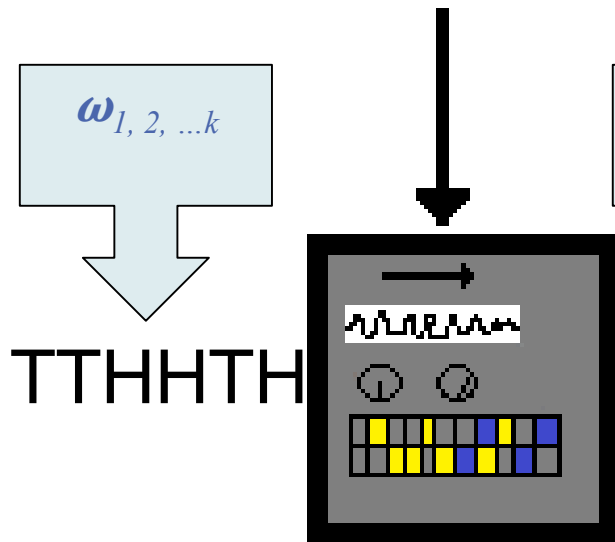
- Different things → different math required

pmf or pdf?

- Every discrete rv has a probability mass function (pmf)

- Every continuous rv has a probability density function (pdf)

- Different ways of defining the function that says how probable an event is.

Discrete rv

- Remember X is our random variable
- We feed it all the $\boldsymbol{\omega}_{1, 2, \ldots k}$ and it transforms each simple event into a value, call that $x_{1, 2, \ldots k}$
- Now each of these values is fed to the probability distribution function, which takes these $x_{1, 2, \ldots k}$ and tallies the number of values it sees, then divides by the total, such that we now have P($X = x_k$)
- So far we have only seen this as a column in a table, but it is a function! P($X = x_k$) = $f(x_k)$
- Called the *probability mass function* (pmf)
- What must be true about this function?
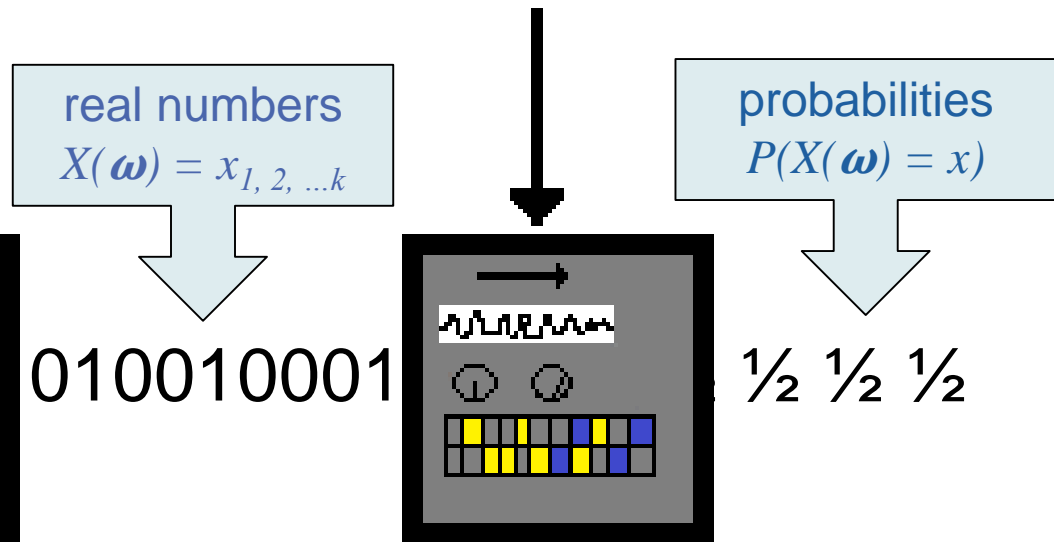
$$\sum_{x_i \in \Omega} f(x_i) = 1$$

$$f(x_i) \geq 0, \forall x_i \in \Omega$$

$$f(x)$$

For discrete X, $f(x)$ is a

Probability Mass Function*

*It gives you a probability! (always between 0 to 1)

Which of these is a pmf?

x	f(x)
-1	0.3
0	0.3
2	0.3

(crossed out)

x	f(x)
10	-0.1
20	0.9
30	0.2

(crossed out)

x	f(x)
1	0.35
2	0.25
3	0.2
4	0.1
5	0.1

x	f(x)
10	0.1
20	0.9
30	0.2

(crossed out)

x	f(x)
5	1

What does a pmf look like?

- Table/list of numbers
  - Rows = $x_{1, 2, ...k}$

- Function (large n)

- A plot
  - x-axis?
  - y-axis?

Bernoulli rv

- Special type of simple discrete rv

- Two possible outcomes: success ($X = 1$) or failure ($X = 0$)
  - P(1) = p
  - P(0) = $1 - p$

Sequences of independent Bernoulli trials → other special types of discrete rvs with discrete distributions

Distribution	Definition
Binomial	Number of successes in n trials
Negative binomial	Number of failures before the x-th success
Geometric	Number of failures before the 1st success

Notation

- The tilde or twiddle: "~"
  - English: "is distributed as"

- $X \sim$ DistributionName(parameters)

- Coin tossing example: $X \sim Bin(n, p)$
  - English: rv X is distributed as a binomial rv with…
  - n number of trials and*…
  - p as the **single trial** probability of success

*n here can be number of coins tossed, or you could have 1 coin tossed n times

$X \sim Bin(2, .5)$: 1 experiment

In R, rbinom(number of experiments (**n**), number of trials per experiment (**size**), probability of success (**prob**). We'll start with one experiment (flipping 2 coins once):

```
> num_exp <- 1
> num_trials <- 2
> p <- .5
> rbinom(num_exp, num_trials, p)
[1] 2
> rbinom(num_exp, num_trials, p)
[1] 0
> rbinom(num_exp, num_trials, p)
[1] 1
```

A little boring- but seems to work!
Let's try flipping 2 coins 10 times…

$X \sim Bin(2, .5)$: 10 experiments

```
> num_exp <- 10
> num_trials <- 2
> x <- rbinom(num_exp,
num_trials, .5)
> length(x)
[1] 10
> table(x)
x
0 1 2
1 6 3
```

x	Expected $P(x)$	Observed $P(x)$
0		10%
1		60%
2		30%

$X \sim Bin(2, .5)$: 1,000 experiments

```
> num_exp <- 1000
> num_trials <- 2
> x <- rbinom(num_exp,
num_trials, .5)
> length(x)
[1] 1000
> table(x)
x
  0   1   2
234 512 254
```

x	Expected $P(x)$	Observed $P(x)$
0	25%	23.4%
1	50%	51.2%
2	25%	25.4%

Let X now count the number of heads (k) in $n = 30$ independent coin flips…

$$X(\omega) = \begin{cases} 2, & \text{if } \omega \text{ is 2 heads} \\ 1, & \text{if } \omega \text{ is 1 head} \\ 0, & \text{if } \omega \text{ is 0 heads} \end{cases}$$

$X(\omega) = k$, where k = total number of heads

In R:
```
x <- rbinom(num_exp, 2, .5)
```

In R:
```
k <- rbinom(num_exp, 30, .5)
```

What is the range of k?

Binomial pmf: $f_X(k; n, p)$ (equivalent notation)

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

1. How many ways are there to get k heads? (combination)

2. How likely is any particular combination of k heads and $n - k$ tails? (multiplicative rule)

When we use the Binomial distribution we know *p* and *n* and the distribution describes the probabilities for all the *k*s we could have.

$X \sim Bin(30, .5)$: 1 experiment, 30 flips

- The distribution that describes the number of heads (k) after n tosses of a fair coin. Given that we have $n = 30$ trials, and we want k heads, and the probability of getting a head is $p = .5$, the probability of each k is:

$$P(X = k) = \binom{30}{k} .5^k (1 - .5)^{30-k}$$

Now we can ask about the probability for any k!

```
> k <- rbinom(1, 30, .5)
> k # in my 1 experiment, I got 17 heads
[1] 17
```

In R…

R FUNCTION	DISCRETE RV	CONTINUOUS RV
`r`	Generates random deviates	
`d`	Probability (pmf)	Probability density (pdf)
`p`	Cumulative probability (cdf)	
`q`	Inverse cumulative probability (quantiles)	

Questions we can ask the pmf

- "What is the probability of getting exactly 17 heads?" 0.112

- "What is the probability of getting 18 to 24 heads?" 0.181

- "What is the probability that the number of heads ends in 7?" 0.113

- "What is the probability of getting 1.5 heads?" Huh?
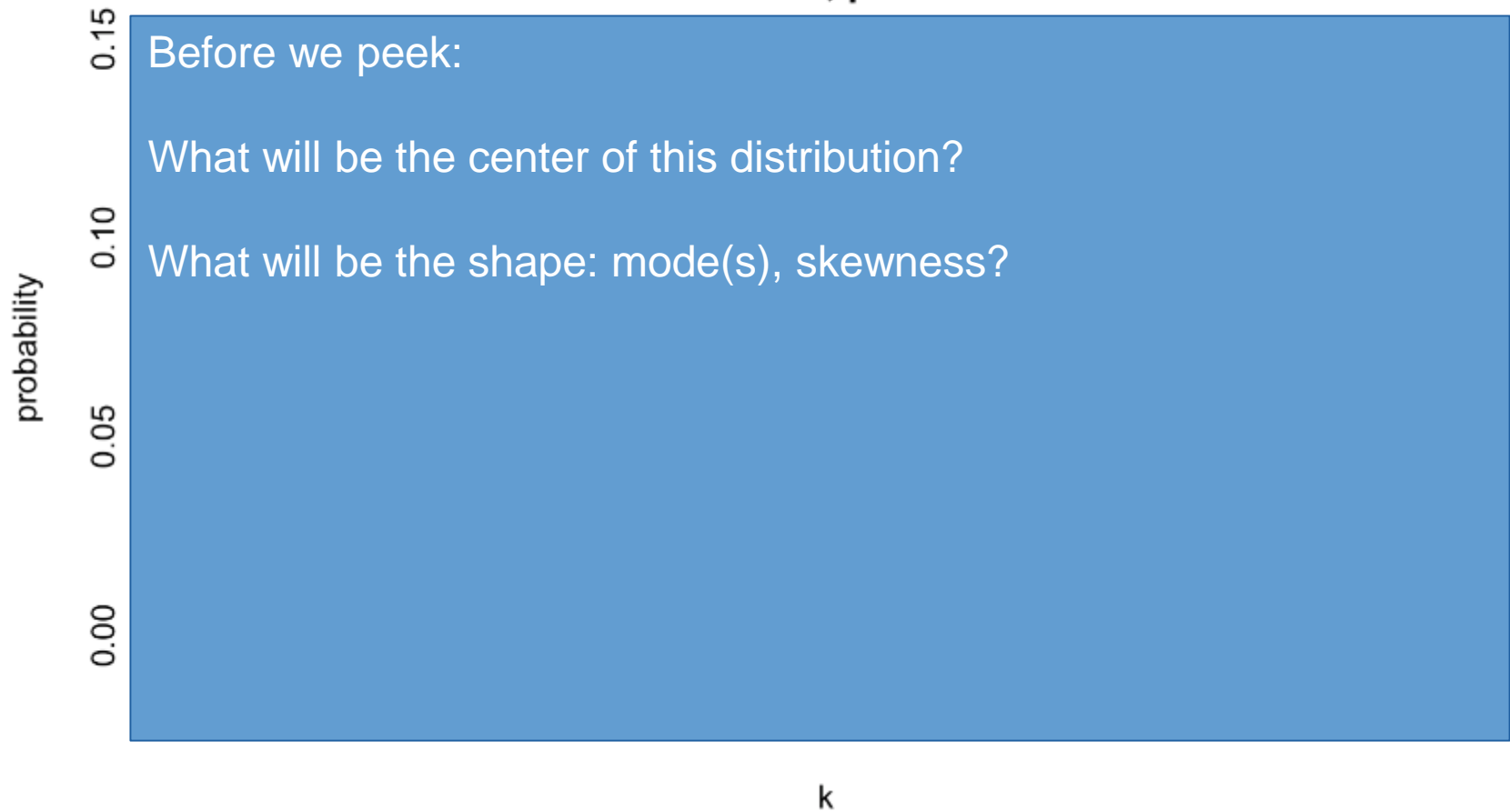
Questions we can ask the pmf in R

- "What is the probability of getting exactly 17 heads?" 0.112

  `dbinom(17, 30, .5)`

- "What is the probability of getting 18 to 24 heads?" 0.181

  `sum(dbinom(18:24, 30, .5))`

- "What is the probability that the number of heads ends in 7?" 0.113

  `sum(dbinom(c(7, 17, 27), 30, .5))`

- "What is the probability of getting 1.5 heads?" Huh?
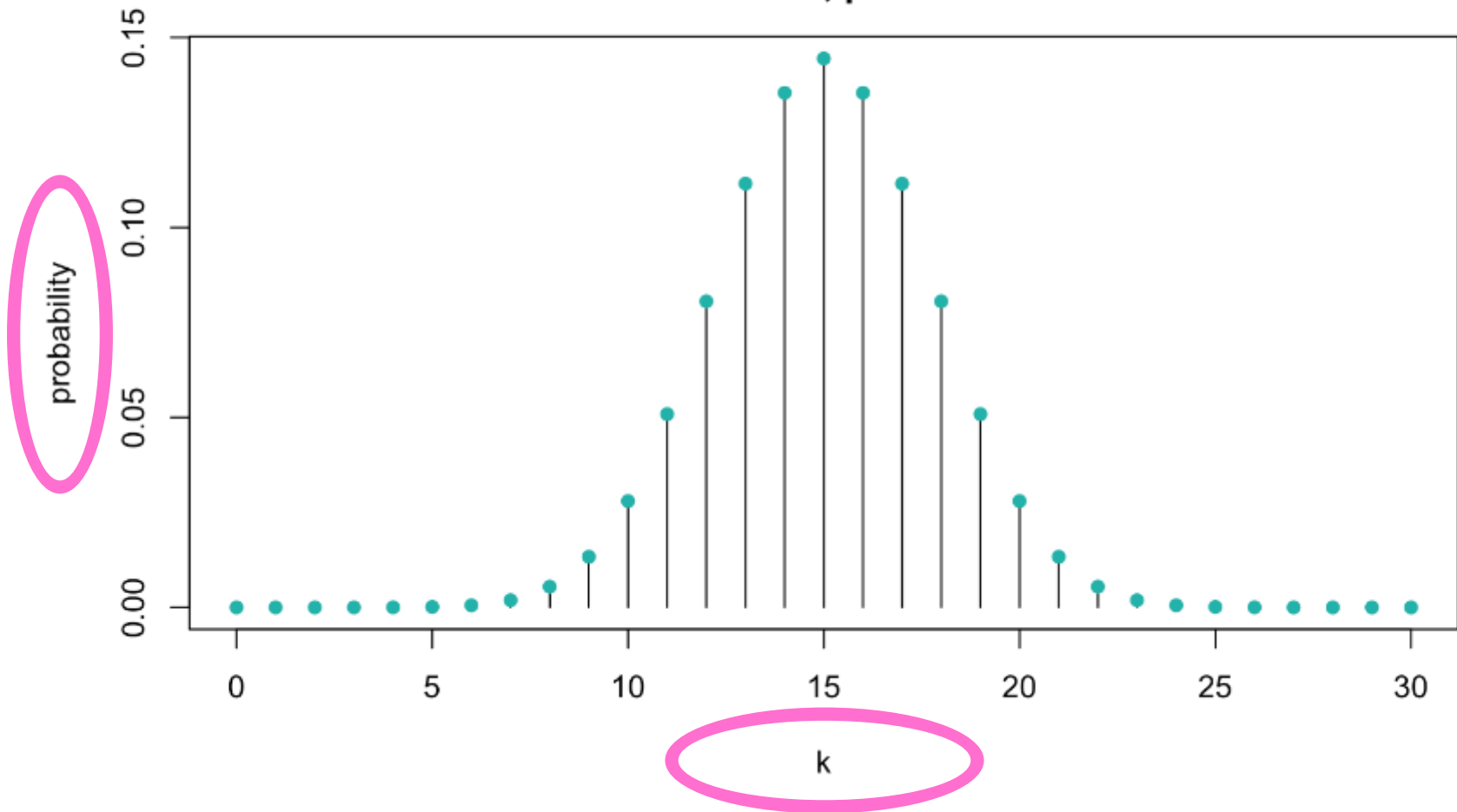
  `Warning message:`
  `In dbinom(1.5, 30, 0.5) : non-integer x = 1.500000`

binomial distribution pmf
n = 30, p = .5

Before we peek:

What will be the center of this distribution?

What will be the shape: mode(s), skewness?

probability

0.00 0.05 0.10 0.15

k

binomial distribution pmf
n = 30, p = .5

$$F(x)$$

For all X, $F(x)$ is a

Cumulative Distribution Function

Discrete rv: cumulative distribution function

$$F(x) = \sum_{t \leq x} f(t)$$

(since x is used as a variable in the summation, we use "t" just as some other variable)

Cumulative distribution function

1. $F(x) \to 1 \; as \; x \to \infty$

2. $F(x) \to 0 \; as \; x \to -\infty$

3. $F(x) \; is \; monotonic; \; never \; decreasing$

1. $F(x) \to 1$ as $x \to \infty$

$$\lim_{x \to +\infty} F(x) = 1$$

$2. \, F(x) \rightarrow 0 \; as \; x \rightarrow -\infty$

$$\lim_{x \to -\infty} F(x) = 0$$

3. $F(x)$ is monotonic; never decreasing
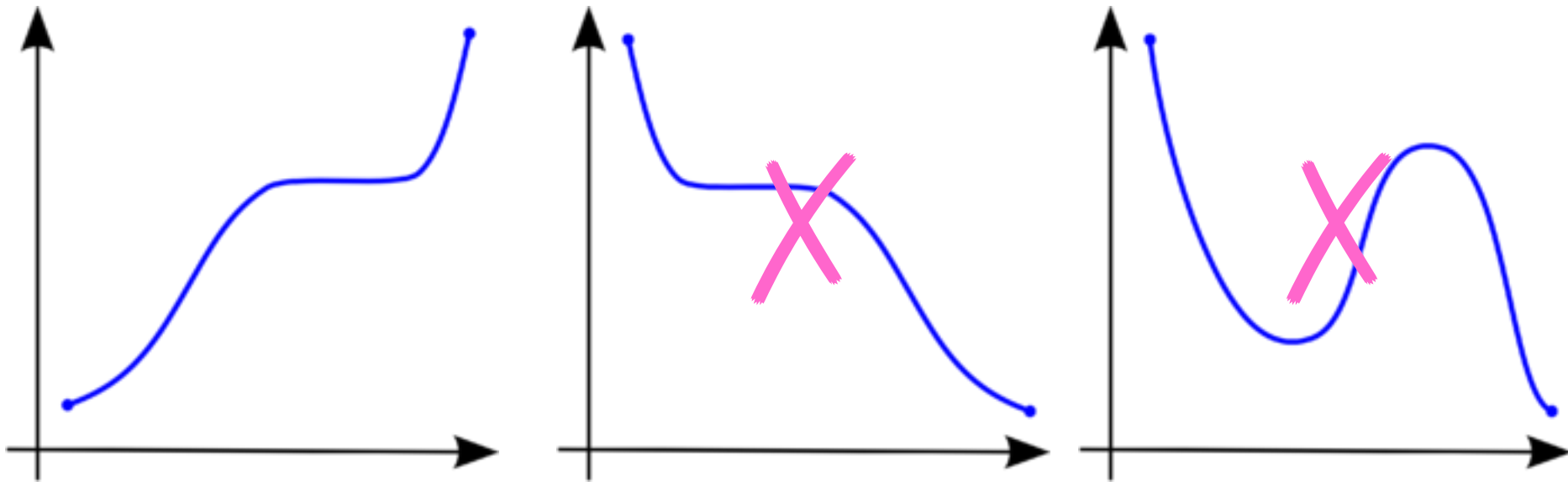
Let X count the number of heads *(k)* in n independent coin flips (again)…

$X(\omega) = k,$ where k = total number of heads

binomial distribution cdf
n = 30, p = .5

Before we peek:

What is the range on the y-axis?
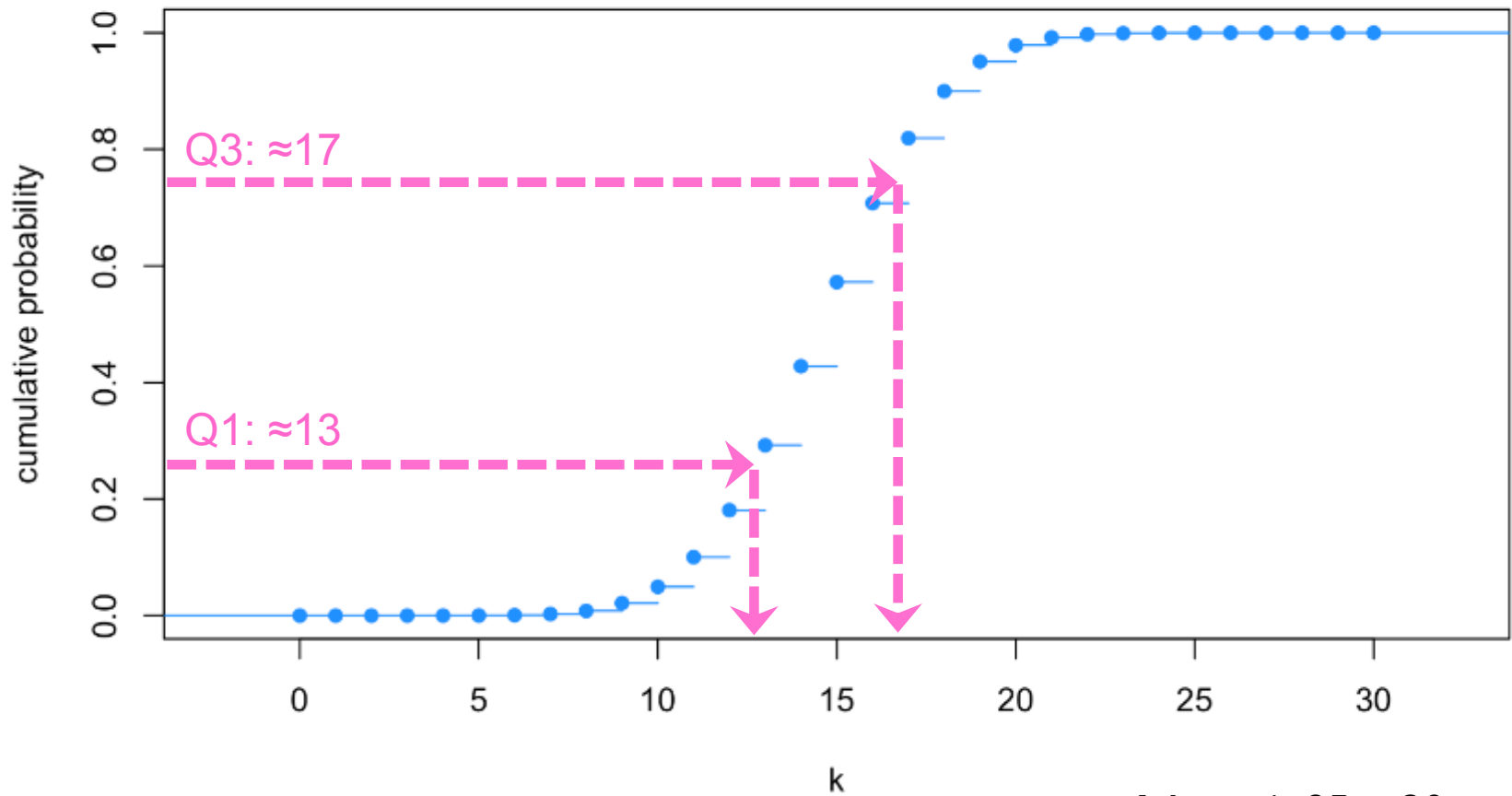
Will it be smooth? Mound-shaped?

cumulative probability

0   5   10   15   20   25   30

k

binomial distribution cdf
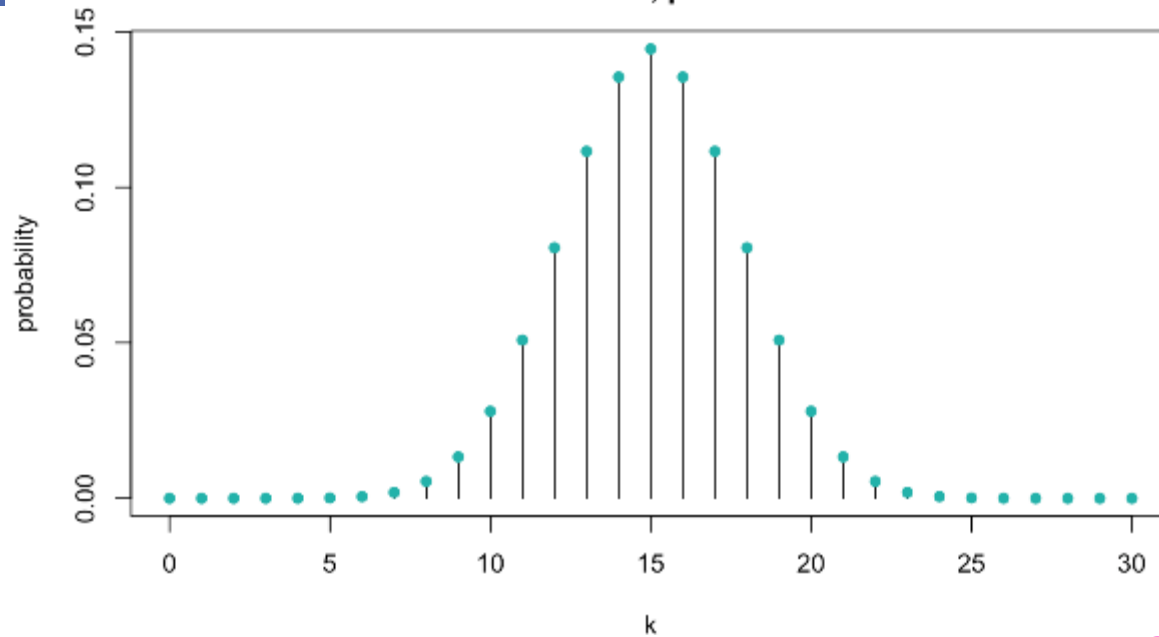## n = 30, p = .5
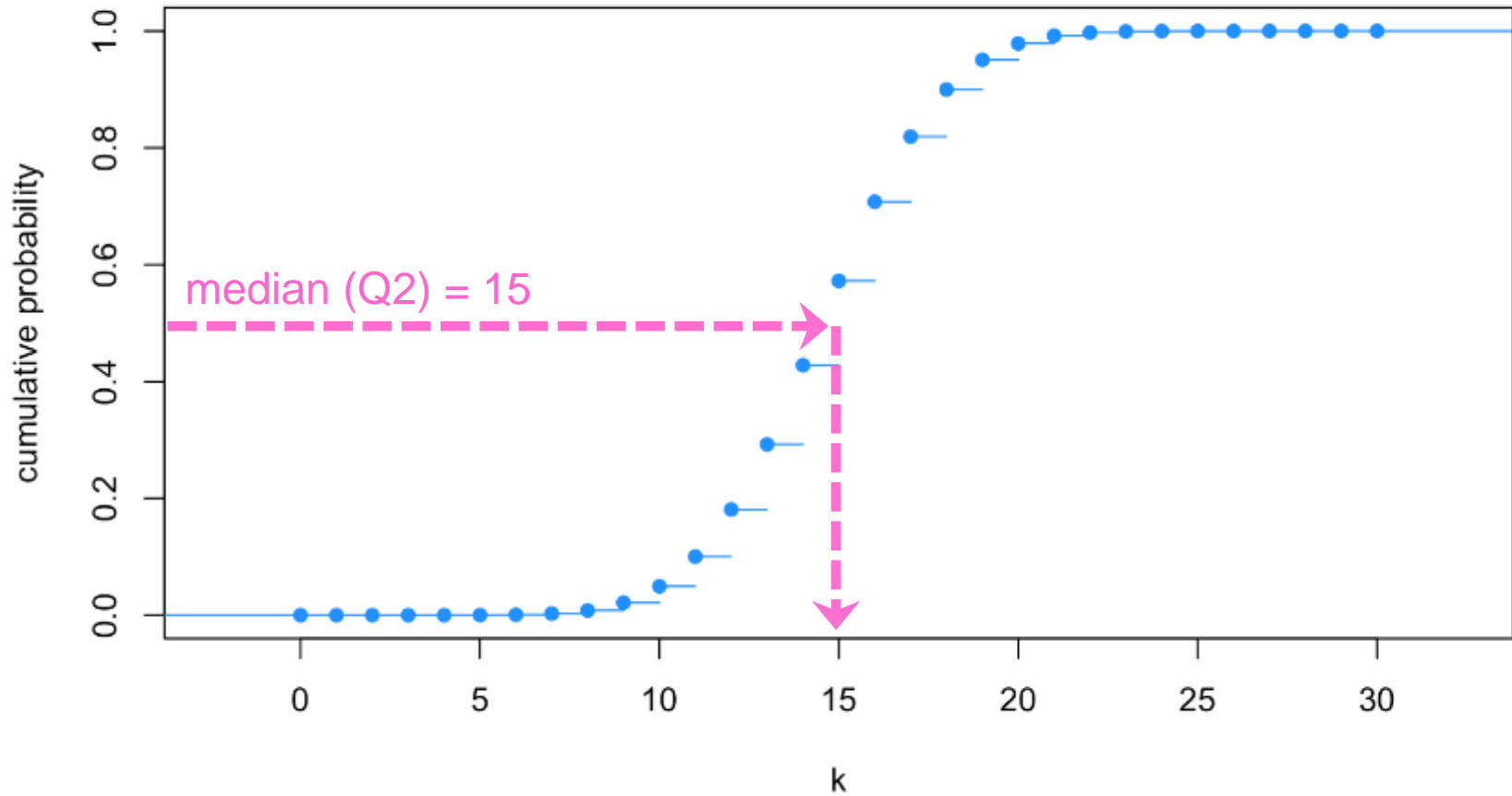
> qbinom(.5, 30, .5) #Q2 (median)
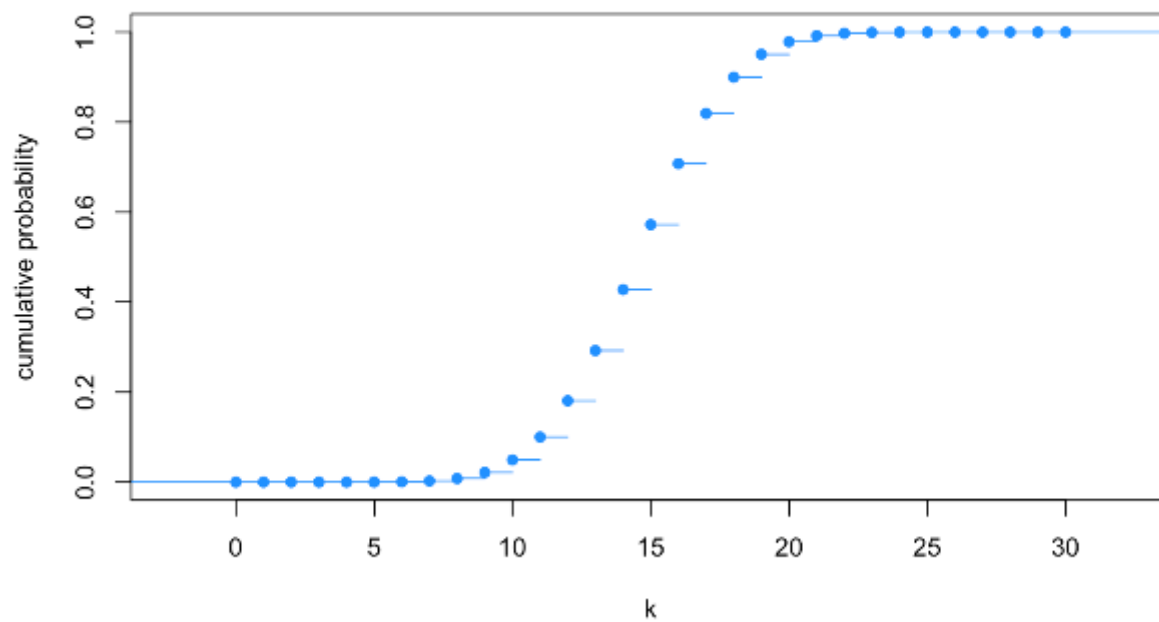[1] 15

binomial distribution cdf
n = 30, p = .5

```
> qbinom(.25, 30, .5) #Q1
[1] 13
> qbinom(.75, 30, .5) #Q3
[1] 17
```

binomial distribution pmf
n = 30, p = .5

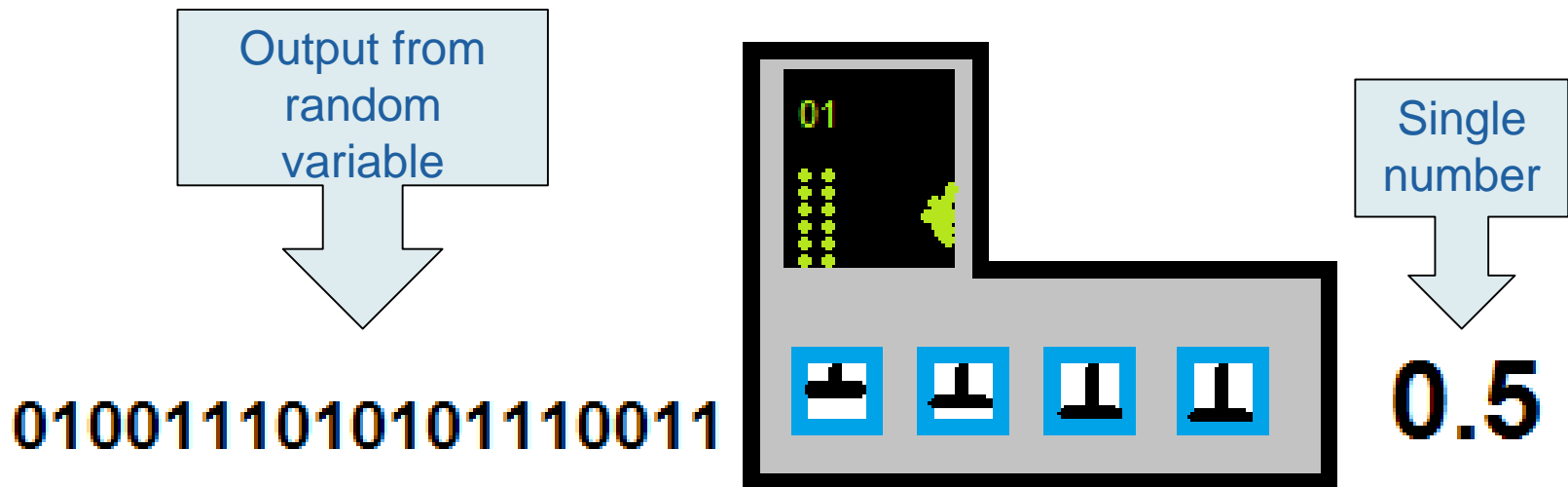**binomial distribution cdf**
n = 30, p = .5

cumulative
sum

Expectation of a
discrete random variable

Expectation

The expectation of a random variable (X) is the sum of its values (x_i) weighted by their probability (p_i).

Expectation!

Output from random variable

Single number

0100111010101110011

0.5

Expectation: Toss a fair coin 2 times

$$X(\omega) = \begin{cases} 2, & \text{if } \omega \text{ is 2 heads} \\ 1, & \text{if } \omega \text{ is 1 head} \\ 0, & \text{if } \omega \text{ is 0 heads} \end{cases}$$

$$E(X) = (1 \times .5) + (2 \times .25) + (0 \times .25) = 1$$

x	$P(x)$	$E(x)$
1 × .5		.5
2 × .25		.5
0 × .25		0
$E(X)$		1.0

Expected values

- The expected value estimates the population mean, μ_X, for an rv with a given distribution

- Example: families with fraternal twins
  - Random variable X: # of girls (discrete!)
  - Population: all families with fraternal twins
  - Expected value: the mean number of girls per families across all families with fraternal twins

- Given that boys and girls are equally likely, what is μ_X?

$$E(X) = (1 \times .5) + (2 \times .25) + (0 \times .25) = 1$$

x	P(x)	E(x)
1 × .5		.5
2 × .25		.5
0 × .25		0
E(X)		1.0

Expected values: discrete distributions
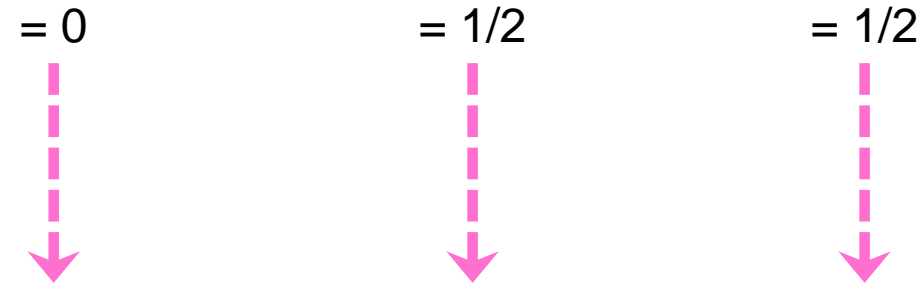
Population mean

μ_X

pmf

$$E(X) = \sum_{i=1}^{n} x_i f_X(x_i)$$

For binomial, plug in pmf: $\binom{n}{k} p^k (1-p)^{n-k}$
$n = 2$
$p = .5$

Expected value of a binomial rv: the long way

= 0   = 1/2   = 1/2

$$E(X) = \sum_{i=0}^{2} x_i \binom{2}{x_i} \left(\frac{1}{2}\right)^2 = \left(0 \times \left(1 \times \frac{1}{4}\right)\right) + \left(1 \times \left(2 \times \frac{1}{4}\right)\right) + \left(2 \times \left(1 \times \frac{1}{4}\right)\right)$$

```
R can help:
> choose(2,0)
[1] 1
> choose(2,1)
[1] 2
> choose(2,2)
[1] 1
```

$$E(X) = \sum_{i=0}^{2} x_i \binom{2}{x_i} \left(\frac{1}{2}\right)^2 = \frac{1}{2} + \frac{1}{2} = 1$$

Population variance for discrete rv

- The variance is the expected value of the square of the deviation of X from its own mean

$$Var(X) = \sigma_X^2 = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Population variance for discrete rv

- The variance is the expected value of the square of the deviation of X from its own mean

$$Var(X) = \sigma_X^2 = \sum_{i=1}^{n} [x_i - \mu_X]^2 f_X(x_i)$$

pmf →
probabilities!

$$Var(X) = \sigma_X^2 = \sum_{i=1}^{n} [x_i - \mu_X]^2 \times p(x_i)$$

Expected value of a binomial rv: the short way

$$E(X) = np$$

$$Var(X) = np(1 - p)$$

Expected values

- If you sample from a distribution, and your sample is sufficiently large, then you have a reasonable expectation that your sample mean will be close to the population mean

- We can calculate the expected value for the underlying distribution and use this to form hypotheses about our current sample

$$\mu_X \neq \overline{x}$$