Computational Physics

Topic 02 — Computational Problems Involving Probability

Lecture 10 — Discrete Random Variables

Dr Kieran Murphy

Department of Computing and Mathematics,
Waterford IT.
(kmurphy@wit.ie)

Autumn Semester, 2021

Outline

- Discrete random variables
- Uniform, Bernoulli, Binomial, Poisson

Motivation

- We want to develop a more mathematical and consistent framework to help us deal with more complicated probability experiments.
- This framework allows us reduce a probability problem from a calculation using our probability rules (arbitrary difficult) to one of identification (easy) and then using preexisting results (relatively easy, trivial with computers).
- The main idea here is that we represent the output of a probability experiment by a random variable — depending on the experiment this is either discrete or continuous — and look at the resulting probability distribution.
- Some probability distributions occur with such regularity in real-life applications that they have been given their own names.

  Here we survey and study basic properties of some of them.

Discrete Random Variable

Definition 1 (Discrete Random Variable)

A discrete random variable, X, takes on (can be equal to) a *countable* number of possible values, e.g.,

- total of roll of two dice: $2, 3, \ldots, 12$
- number of desktops sold: $0, 1, \ldots$
- customer count: $0, 1, \ldots$

Notation

If X is a random variable then (uppercase roman)

- $\mathcal{X}$ is the set of possible values (the alphabet/sample space) (calligraphic)
- x or x_k is any particular value in the alphabet (lowercase roman)
- $p_k = \Pr(X = x_k)$

Probability Mass Function (PMF)

The probability that a discrete random variable X takes on a particular value x, that is, $\Pr(X = x)$, is specified by a function, $f(x)$, called the probability mass function.

Definition 2 (Probability Mass Function (PMF))

The probability mass function

$$\Pr(X = x) = f(x)$$

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

- $0 \leq f(x_k) \leq 1$ for all $x_k \in \mathcal{X}$ *(probabilities are in $[0, 1]$)*
- $\sum_{x_k \in \mathcal{X}} f(x_k) = 1$ *(sum of all probabilities equal one)*

Since f(x) is a function, it can be presented in tabular form, in graphical form, or as a formula.

Probability Mass Function — Defined via Table

Empirical data can be used to estimate the probability mass function.

Example 3

Consider, for example, the number of TVs in a household …

No. of TVs	**No. of Households**
0	1,218
1	32,379
2	37,961
3	19,387
4	7,714
5	2,842

$\rightarrow$

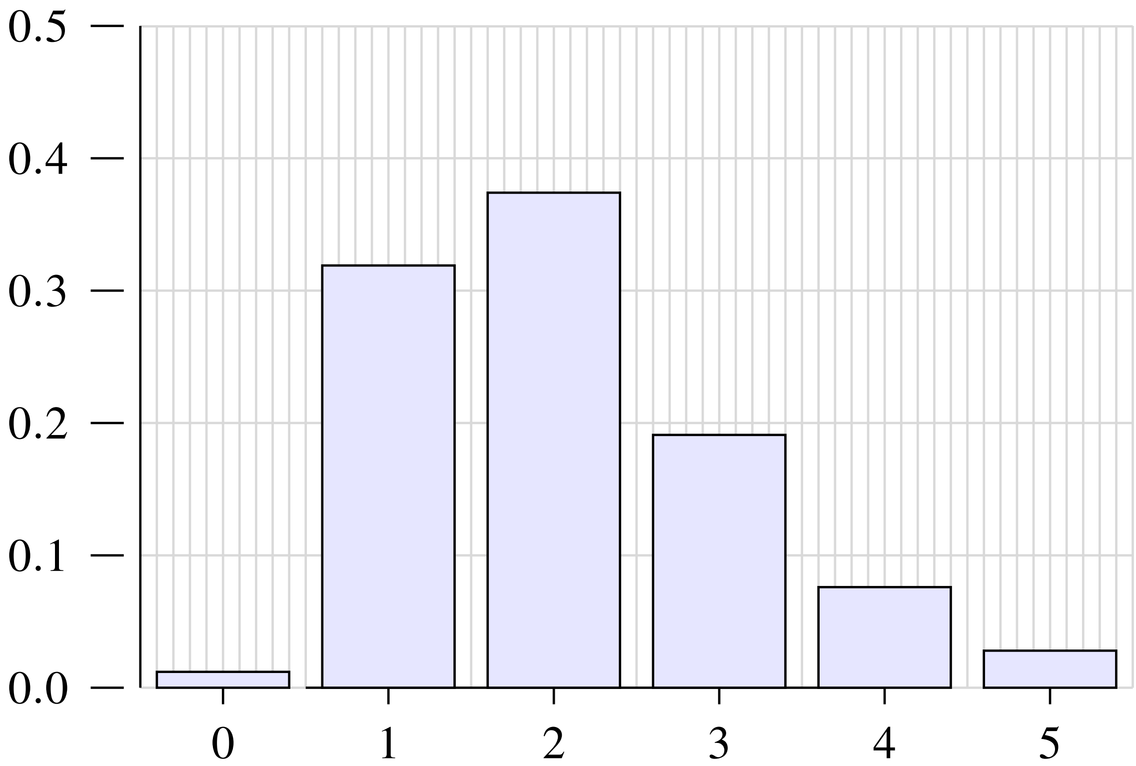
x	$f(x)$
0	0.012
1	0.319
2	0.374
3	0.191
4	0.076
5	0.028
	1.000

Similarly we can construct the probability mass function from a bar plot.

Example 3

x	$f(x)$
0	0.012
1	0.319
2	0.374
3	0.191
4	0.076
5	0.028
	1.000

$\leftrightarrow$



Cumulative Distribution Function (CDF)

We often wish to work with cumulative probability ...

Definition 4 (Cumulative Distribution Function (CDF))

The cumulative distribution function (CDF) of the random variable X has the following definition.

$$F(x) = \Pr(X \leq x) \tag{1}$$

Example 5

Let X be a discrete random variable with PMF

$$f(x) = \frac{5 - x}{10}, \quad \text{for } x = 1, 2, 3, 4$$

Determine its CDF.

Example 5

- $x = 1$

$$\Pr(X \le 1) = \Pr(X = 1) = f(1) = \frac{5-1}{10} = \frac{4}{10}$$

- $x = 2$

$$\begin{aligned} \Pr(X \le 2) = \Pr(X \le 1 \text{ OR } X = 2) &= \Pr(X \le 1) + \Pr(X = 2) \\ &= \Pr(X \le 1) + f(2) = \frac{4}{10} + \frac{3}{10} = \frac{7}{10} \end{aligned}$$

- $x = 3$

$$\begin{aligned} \Pr(X \le 3) = \Pr(X \le 2 \text{ OR } X = 3) &= \Pr(X \le 2) + \Pr(X = 3) \\ &= \Pr(X \le 2) + f(3) = \frac{7}{10} + \frac{2}{10} = \frac{9}{10} \end{aligned}$$

- $x = 4$

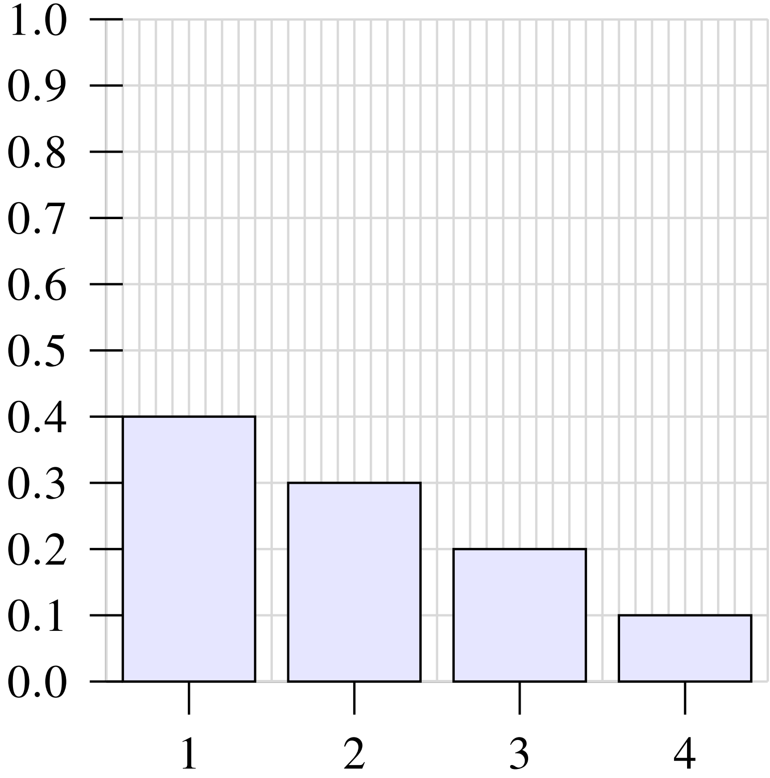
$$\begin{aligned} \Pr(X \le 4) = \Pr(X \le 3 \text{ OR } X = 4) &= \Pr(X \le 3) + \Pr(X = 4) \\ &= \Pr(X \le 3) + f(4) = \frac{9}{10} + \frac{1}{10} = \frac{10}{10} \end{aligned}$$

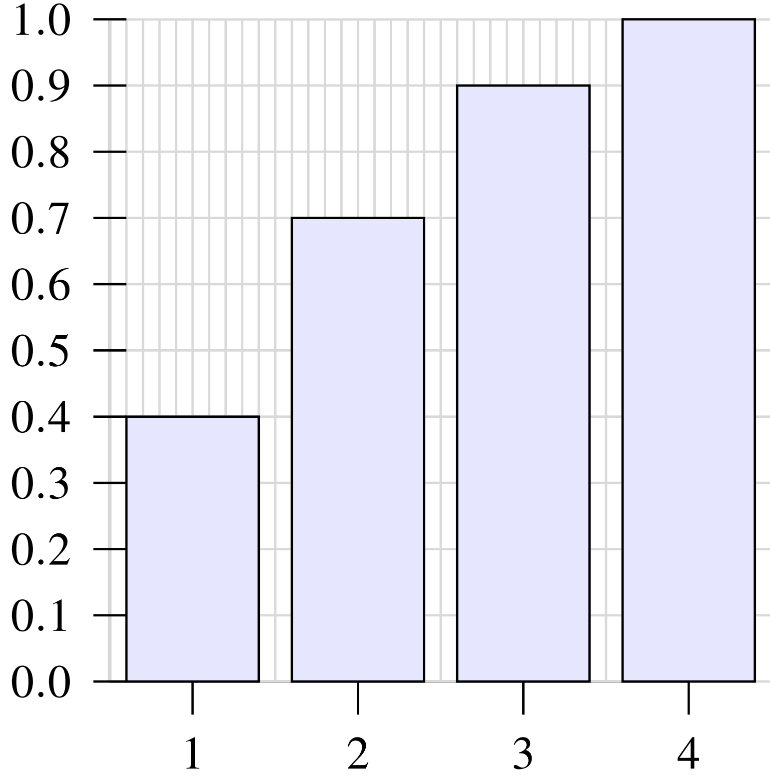
Example 5

So the random variable X with probability mass function

$$f(x) = \frac{5-x}{10}$$

has probability mass function and cumulative distribution function



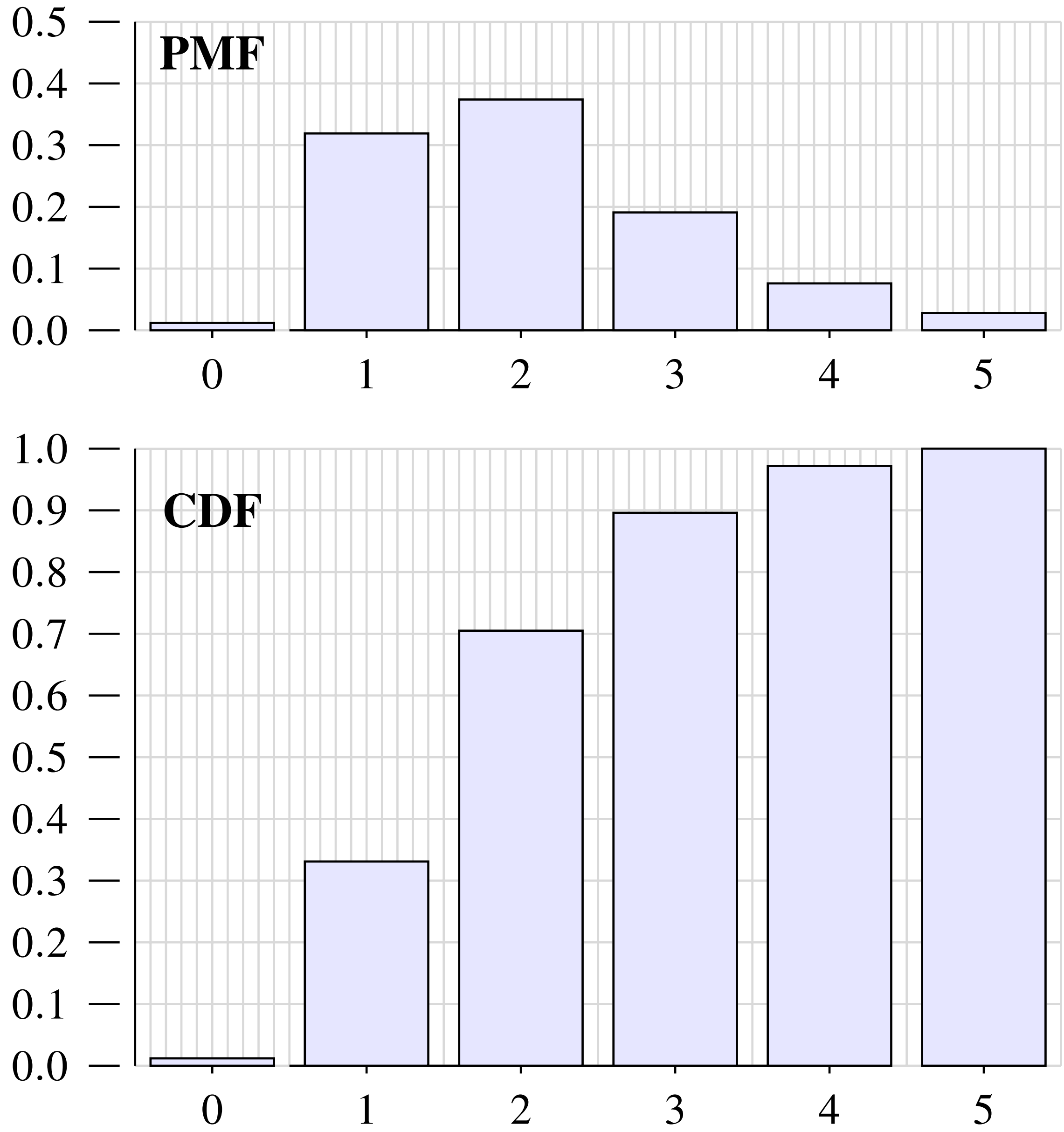


Example 3 (revisited)

The CDF for the TV in households example is

x	**PMF**	**CDF**
0	0.012	0.012
1	0.319	0.331
2	0.374	0.705
3	0.191	0.896
4	0.076	0.972
5	0.028	1.000
	1.000	

$\leftrightarrow$



Expected Value (EV)

The mean or expected value of a discrete random variable is a weighted average of the outcomes.

Definition 6 (Expected Value (EV))

If a discrete random variable, X, takes outcomes $x_1, \ldots, x_n$ with probabilities $\Pr(X = x_1) = f(x_1), \ldots, \Pr(X = x_n) = f(x_n)$, then the mean or expected value of X is the sum of each outcome multiplied by its corresponding probability:

$$\mathrm{E}[X] = \mu = x_1 P(X = x_1) + \cdots + x_n P(X = x_n) = \sum_{k=1}^{n} x_k \cdot f(x_k) \tag{2}$$

More generally, for any function $g(X)$, we have

$$\mathrm{E}\left[g(X)\right] = \sum_{k=1}^{n} g(x_k) \cdot f(x_k)$$

Example 7

Example 7

Calculate the expected value for the number of TVs per household from Example 3

x	$f(x)$	$x \cdot f(x)$
0	0.012	0.000
1	0.319	0.319
2	0.374	0.748
3	0.191	0.573
4	0.076	0.304
5	0.028	0.140
		2.084

Hence the expected number of TVs per household is

$$\mathrm{E}[X] = \sum_k x_k \cdot f(x_k) = 2.084$$

Variance of a Random Variable

The variance of random variable is based on the population variance, i.e., it is the weighted (by probabilities) average of the squared deviations from the mean.

Definition 8

If X is a random variable with mean μ then, its variance is

$$\mathrm{Var}[X] = \sigma^2 = \sum_k (x_k - \mu)^2 \cdot f(x_k) \tag{3}$$

- The variance can be written in terms of the expected value

$$\mathrm{Var}[X] = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \cdots \begin{matrix}\text{some}\\ \text{algebra}\end{matrix} \cdots = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$$

- The rightmost formula is significant faster to use since it does not need to know μ first.

Example 9

Example 9

Calculate the variance for the number of TVs per household from Example 3

x	$f(x)$	$x \cdot f(x)$	$(x-\mu)^2$	$(x-\mu)^2 \cdot f(x)$
0	0.012	0.000	?	?
1	0.319	0.319	?	?
2	0.374	0.748	0.00705	0.002638
3	0.191	0.573	?	?
4	0.076	0.304	?	?
5	0.028	0.140	?	?
		2.084		1.106944
				$\text{Var}[X]$

OR

x	$f(x)$	$x \cdot f(x)$	$x^2 \cdot f(x)$
0	0.012	0.000	0.000
1	0.319	0.319	0.319
2	0.374	0.748	1.496
3	0.191	0.573	1.719
4	0.076	0.304	1.216
5	0.028	0.140	0.700
		2.084	5.450
		$\text{E}[X]$	$\text{E}\left[X^2\right]$

Hence the expected number of TVs per household is

$$\text{Var}[X] = \text{E}\left[X^2\right] - \text{E}[X]^2 = 5.450 - (2.084)^2 = 1.106944$$

Discrete Probability Distributions

Next we will look at some of the more important discrete probability distributions.

Bernoulli Distribution

Generalised coin toss experiment — two outcomes with probability p and $1 - p$.

Discrete Uniform Distribution

Generalised coin toss/ die roll experiment — n outcomes with all equally likely.

Geometric Distribution

Number of repeated Bernoulli experiments until a success occurs.

Binomial Distribution

Experiment consisting of number of success in repeated Bernoulli experiments.

Poisson Distribution

Experiment consisting of number of rare events during a fixed time interval or the number of rare successes in a very large number of trials.

Bernoulli Trial / Distribution

Motivation / Application

- A Bernoulli trial is a generalisation of the coin toss experiment.
- We consider an experiment with two outcomes — failure/success, no/yes, lose/win, false/true, 0/1, etc.
- Generalise to allow for unequal probabilities — parameter $0 \leq p \leq 1$.

Definition

Definition 10 (Bernoulli Distribution)

The Bernoulli distribution, is a **one-parameter** probability distribution, defined by the following probability mass function

Parameter $p \in [0, 1]$ is probability of getting one (success)

x	$f(x)$
0	$1 - p$
1	p

If X is a Bernoulli random variable, or follows the Bernoulli distribution we write

$$X \sim \text{Bernoulli}(p)$$

Bernoulli — Properties

- Rather than using a lookup table, the Bernoulli distribution can be defined via a formula:

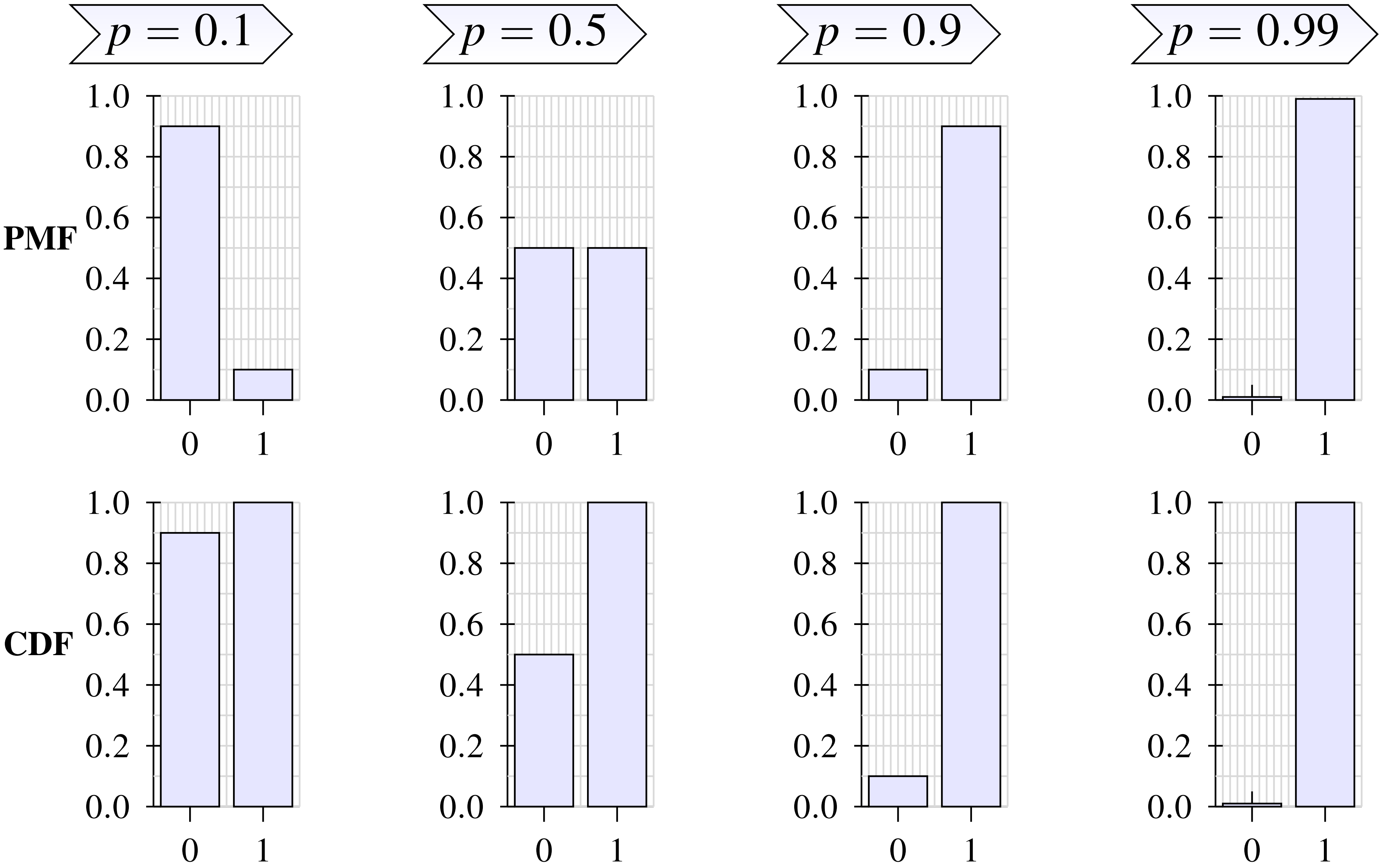
$$f(x) = p^x(1-p)^{1-x} \quad \text{for } x = 0, 1$$

- A Bernoulli has expected value and variance of

$$\mathrm{E}[X] = \mu = p \qquad\qquad \mathrm{Var}[X] = \sigma^2 = p(1-p)$$

- Bernoulli trials are often used as building blocks for more complicated experiments:
  - Geometric — Number of Bernoulli trials needed before first ''success'.
  - Binomial — Number of success in a given number of Bernoulli trials.

Bernoulli PMF and CDF — Effect of parameter p



Discrete Uniform

Motivation / Application

- The Bernoulli distribution is a generalisation of a fair coin toss by allowing unequal probabilities. In contrast the uniform distribution is a generalisation of a fair coin toss by varying the number of sides.
- Applications: coin toss, dice roll, lottery ticket, etc.

Definition

Definition 11 (Discrete Uniform Distribution)

The discrete uniform distribution is a **one-parameter** probability distribution, defined by the probability mass function

$$f(x) = 1/n \quad \text{for } x = 1, 2, \ldots, n$$

Parameter $n \geq 1$, is number of outcomes

If X follows the Uniform distribution we write

$$X \sim \text{Uniform}(n)$$

Discrete Uniform — Properties

- The expected value and variance of the uniform is

$$\mathrm{E}[X] = \mu = \frac{n+1}{2} \qquad \mathrm{Var}[X] = \sigma^2 = \frac{n^2-1}{12}$$

- Examples:
  - Toss a fair coins represented by Uniform(2)

$$\text{Outcomes} = \{1, 2\} \qquad \Pr(\text{any outcome}) = 0.5$$

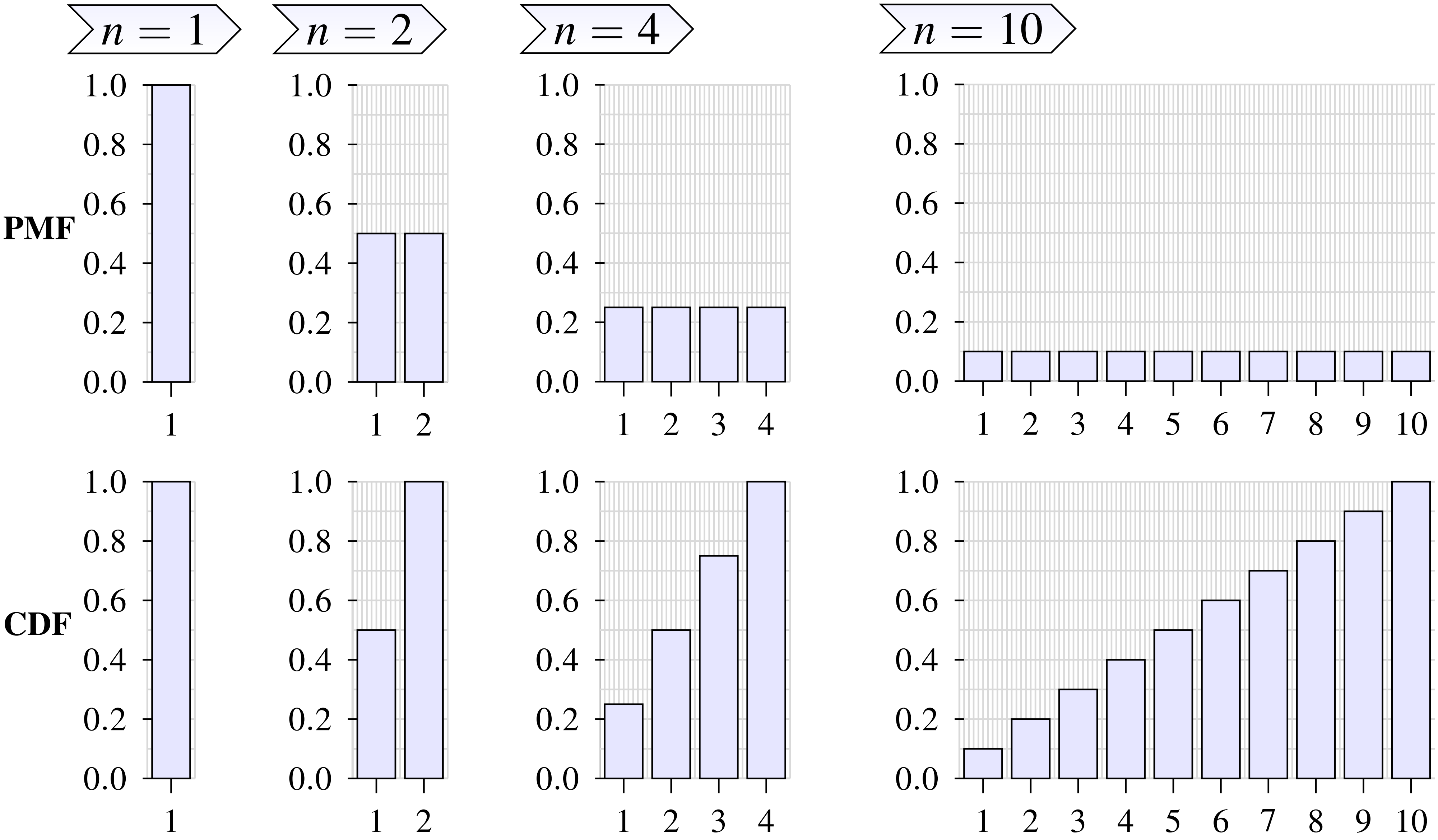
  - Rolling a die is represented by Uniform(6)

$$\text{Outcomes} = \{1, 2, 3, 4, 5, 6\} \qquad \Pr(\text{any outcome}) = 0.16666$$

  - Picking the winner in, say the Irish National lottery, is represented by Uniform(10737573)

$$\Pr(\text{any outcome}) = 0.000000093\ldots$$

Discrete Uniform — Effect of Parameter, n



Geometric

Motivation / Application

- Consider an experiment of running **identical**, **independent** Bernoulli trials until first 'success' occurs.
- How long (number of trials) will occur?

Definition

Definition 12 (Geometric Distribution)

The Geometric Distribution, is a **one-parameter** probability distribution defined by the probability mass function

$$f(x) = (1-p)^{x-1}p \qquad x = 1, 2, \ldots$$

returns the probability of seeing the first 'success' in the x^{th} Bernoulli trial where parameter

- p is the probability of 'success', $0 \leq p \leq 1$

Geometric Distribution — Properties

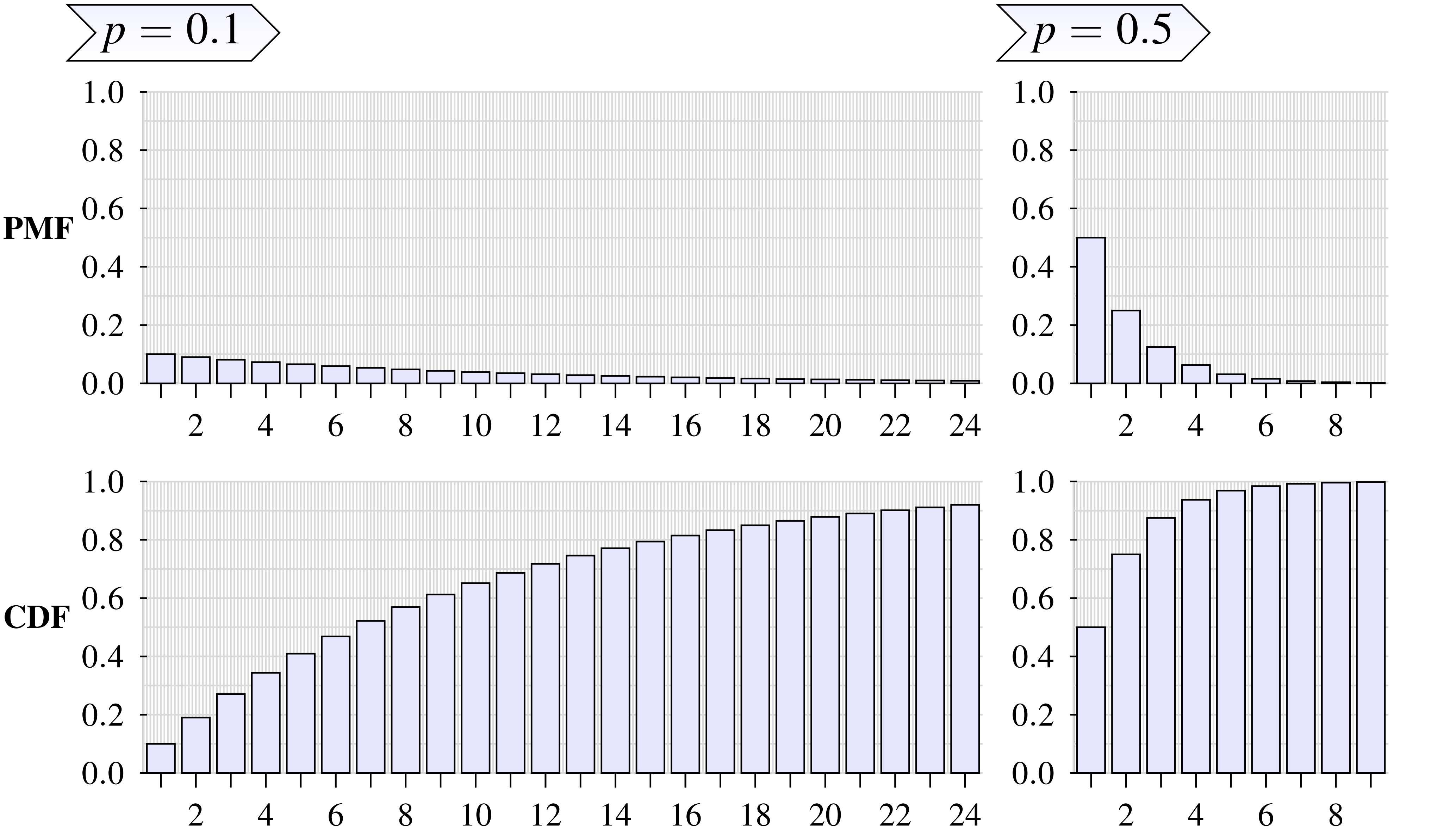
- The Geometric distribution has expected value and variance of

$$\mathrm{E}[X] = \mu = \frac{1}{p} \qquad \mathrm{Var}[X] = \sigma^2 = \frac{1-p}{p^2}$$

- While $\Pr(X = x)$ for large values of x gets smaller and smaller it never reaches zero (unless $p = 0$).
- Has cumulative distribution function (CDF) of

$$F(x) = 1 - (1-p)^x$$

Bernoulli PMF and CDF — Effect of parameter p



Example 13

Example 13

A newly-wed couple (both of whom are non-red heads but are carriers) plan to have children, and will continue until the first red-haired girl is born. How many children can they expect to have, given that probability of a red-head to non-ginger, carriers is 25% and probability of a girl is 50%?

- success = red-haired girl
- $\Pr(\text{success}) = 0.25 \times 0.5 = 0.125$
- Expected number of children

$$\mu = \frac{1}{0.125} = 8$$

Binomial Distribution — Motivation / Applikation

The Binomial distribution is a generalisation of the Bernoulli, in which we consider arbitrary number of repeated Bernoulli experiments:

- The number of heads/tails in a sequence of coin flips.
- Vote counts for two different candidates in an election.
- The number of male/female employees in a company.
- The number of defective products in a production run.
- The number of days in a month your company's computer network experiences a problem.

Checklist for Binomial

- The trials independent and identical.
- The number of trials, n, is fixed in advance.
- Each trial outcome can be classified as a success or failure.
- The probability of a success, p, is the same for each trial.

Binomial Distribution — Definition

Definition 14 (Binomial distribution)

The binomial distribution is a **two-parameter** probability distribution defined by the probability mass function

$$f(x) = \mathrm{C}\,(n, x)\, p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, \ldots, n$$

returns the probability of x 'success' where parameters

- n is the number of trials. $\quad n = 0, 1, 2, \ldots$
- p is probability of success in a trial. $\quad 0 \leq p \leq 1$

and

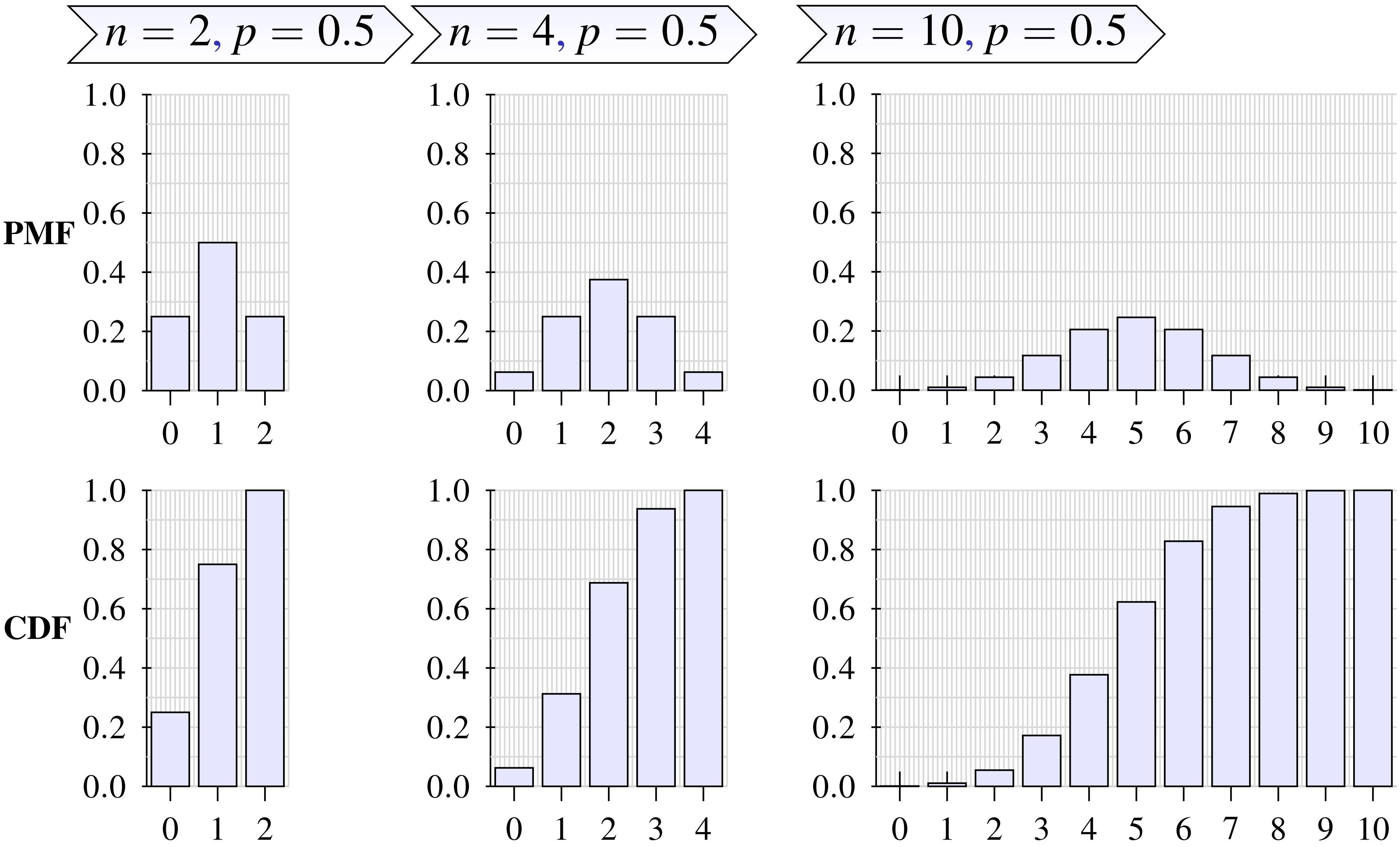
$$\mathrm{C}\,(n, x) = \frac{n!}{(n-x)!x!}$$

- The Binomial distribution has expected value and variance of

$$\mathrm{E}[X] = \mu = np \qquad \mathrm{Var}[X] = \sigma^2 = np(1-p)$$

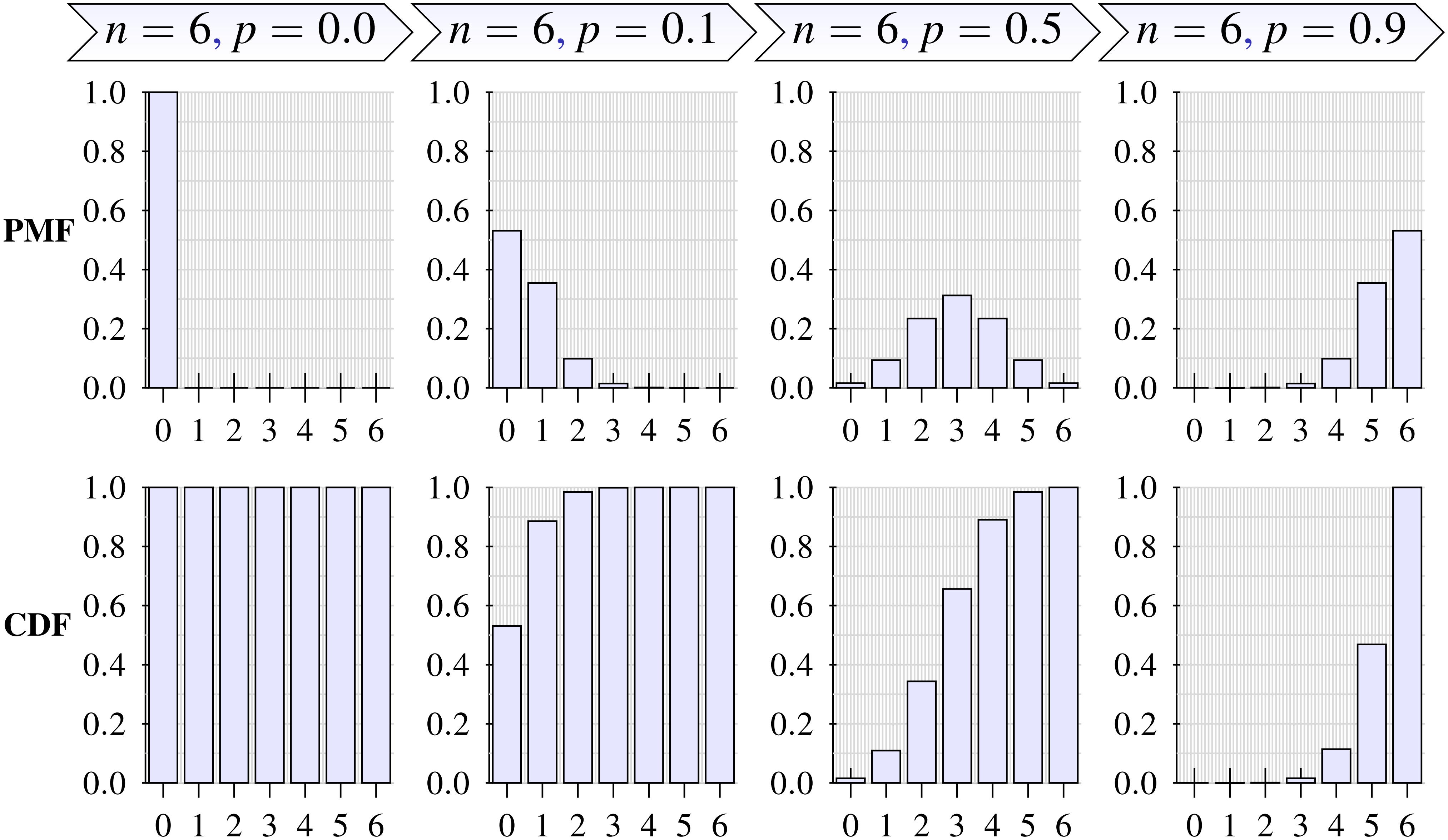
Binomial Distribution — Effect of Parameter n

Parameter n represents the number of trials.



Binomial Distribution — Effect of Parameter p

Parameter p represents the probability of success in a trial.



Example 15

Example 15

A sales person calls to six doors selling his product. At each customer, the probability that he will succeed in making a sale is 0.3.
What is the probability that he will make

(a) no sale
(b) at least one sale
(c) two sales
(d) at most two sales
(e) at least two sales

MODEL
Binomial: success = making a sale with a customer.
PARAMETERS

- n = number of doors/customers $\quad n = 6$
- p = probability of sale $\quad p = 0.3$

Example 15

Hence the probability mass function is

$$\Pr(X = x) = f(x) = \mathrm{C}\,(6, x) \cdot (0.3)^x \cdot (1 - 0.3)^{6-x}$$

(a) *no sale*

$$\Pr(X = 0) = f(0) = 11.7649\%$$

(b) *at least one sale*

$$\begin{aligned}\Pr(X > 0) &= \Pr(X = 1 \text{ OR } X = 2 \text{ OR } X = 3 \cdots \text{ OR } X = 6) \\ &= f(1) + f(2) + \cdots + f(6)\end{aligned}$$

… this is a lot of work … instead use total law of probability …

$$\begin{aligned}\Pr(X > 0) &= 1 - \Pr(X \leq 0) \\ &= 1 - \Pr(X = 0) \\ &= 1 - f(0) = 1 - 0.117649 = 88.2351\%\end{aligned}$$

Example 15

(c) *probability that he will make two sales*

$$\Pr(X = 2) = f(2) = 32.4135\%$$

(d) *probability that he will at most two sales*

$$\begin{aligned}\Pr(X \leq 2) &= f(0) + f(1) + f(2) \\ &= 0.117649 + 0.302526 + 0.324135 = 74.431\%\end{aligned}$$

(e) *probability that he will make at least two sales*

$$\begin{aligned}\Pr(X \geq 2) = 1 - \Pr(X < 2) &= 1 - \Pr(X \leq 1) \\ &= 1 - (f(0) + f(1)) \\ &= 1 - (0.117649 + 0.302526) = 57.9825\%\end{aligned}$$

Poisson Distribution — Motivation / Applikation

The Poisson distribution is used for modelling the occurrence of events in a fixed interval ((his could be time, length, area or volume) or the number of rare successes in a very large number of trials, such as:

- the number of misprints on a book page
- the number of goals during a football match
- the number of telephone calls during a fixed time interval
- the number of incorrectly send messages in a channel.

Checklist for Poisson

- “Events are rare” $\approx$ “assume events don’t happen simultaneously”
- The probability of more than one event occurring during a short interval/length/area/volume must be small relative to the occurrence of only one event.
- The time interval/length/area/volume is fixed and we are interested in number of events that occur.
- Each occurrence of an event is independent of all other occurrences. For example, in a queueing context, the number of people who arrive in the first hour is independent of the number who arrive in any other hour.

Poisson Distribution — Definition

Definition 16 (Poisson distribution)

The Poisson distribution is a **one-parameter** probability distribution defined by the probability mass function

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad \text{for } x = 0, 1, \ldots$$

returns the probability of x occurrences of a rare event per **unit time interval** where parameter

- λ is the average number of occurrences of the rare event per unit time interval $\qquad 0 \leq \lambda$

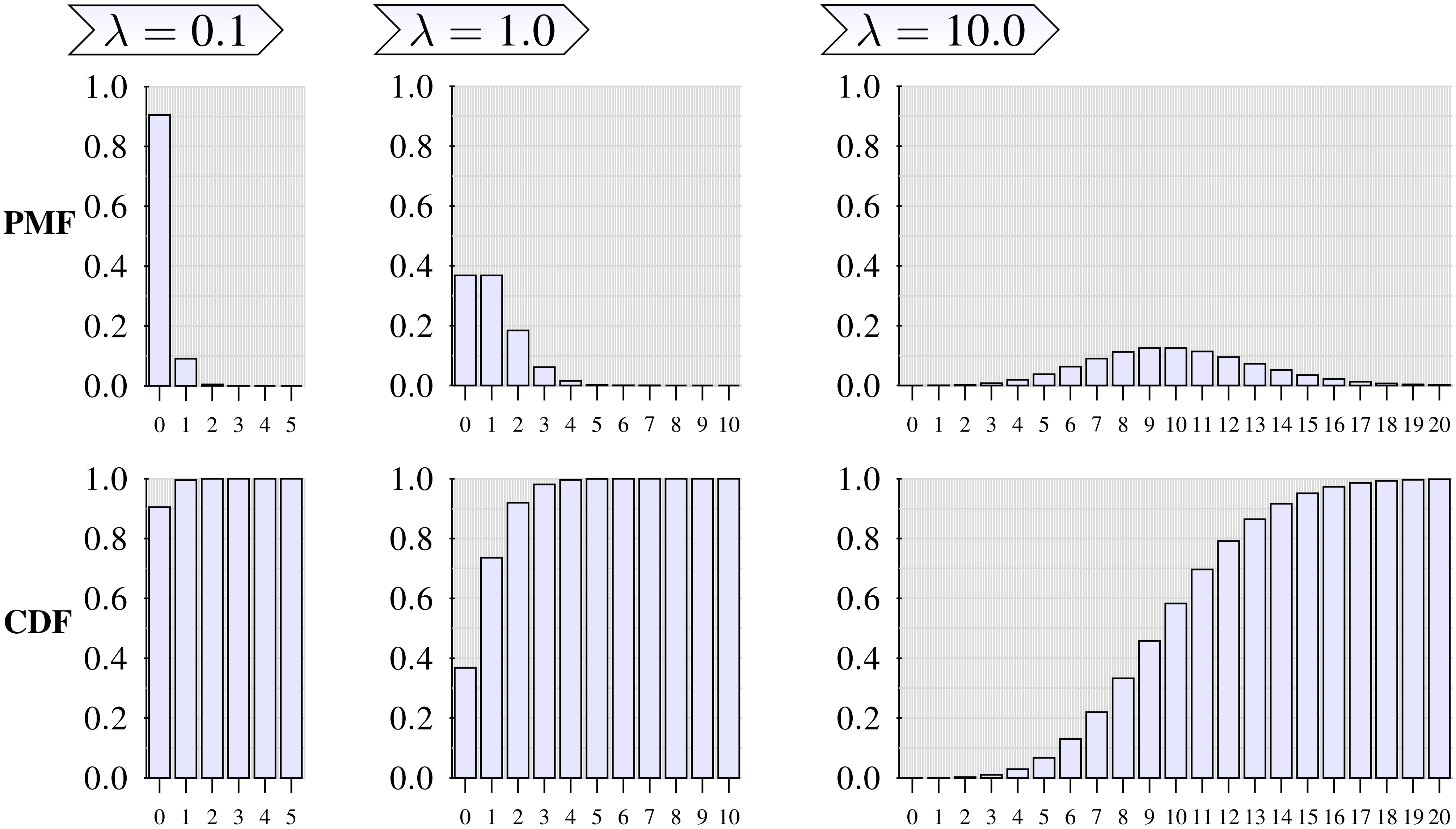
- The Poisson distribution has expected value and variance of

$$\mathrm{E}[X] = \mu = \lambda \qquad \mathrm{Var}[X] = \sigma^2 = \lambda$$

Make sure the you are consistent with the “unit time interval” !

Poisson Distribution — Effect of Parameter λ

Parameter λ represents average number of occurrences of the rare event per unit time interval.



Example 17

Example 17

A service repair centre receives on average of 18 call-outs per 5-day week. Calculate the probability that on any given day the centre will receive 0, 1,2,3,4,5,6 etc.. call outs.

MODEL

Poisson: measuring **number of call-outs** per **day**.

PARAMETERS

- λ = average number of call-outs per day. $\lambda = 18/5$

Scaling parameter vs probabilities

In this problem we were given data for a 5-day interval but wanted probabilities relating to a 1-day interval.
We always scale the parameter, λ, not the resulting probabilities. Why?

Example 17

Hence the probability mass function is

$$\Pr(X = x) = f(x) = \frac{(18/5)^x}{x!} e^{-18/5} \qquad \text{for } x = 0, 1, \ldots$$

Zero call-outs ($x = 0$)

$$\Pr(X = 0) = f(0) = 2.7324\%$$

One call-outs ($x = 1$)

$$\Pr(X = 1) = f(1) = 9.8365\%$$

and so one …

$\Pr(X = 2) = f(2) = 17.7\%$
$\Pr(X = 3) = f(3) = 21.2\%$
$\Pr(X = 4) = f(4) = 19.1\%$
$\Pr(X = 5) = f(5) = 13.8\%$
$\Pr(X = 6) = f(6) = 8.3\%$

$\Pr(X = 7) = f(7) = 4.2\%$
$\Pr(X = 8) = f(8) = 1.9\%$
$\Pr(X = 9) = f(9) = 0.76\%$
$\Pr(X = 10) = f(10) = 0.28\%$
$\Pr(X = 11) = f(11) = 0.09\%$

Example 19

Example 18

On a given morning the average number of telephone calls per minute received by a switchboard of a company is 2.3. Assuming a Poisson distribution, find the probability that in any one-minute period there will be

(a) at least one call
(b) six calls
(c) at most two calls.

Example 19

A radioactive source emits 4 particles on average during a five-second period. Calculate the probability that

(a) it emits 3 particles during a 5-second period.
(b) it emits at least one particle during a 5-second period.
(c) during a ten-second period, it emits 6 particles.

Poisson Approximation to Binomial

The Poisson distribution can be used as an approximation for the binomial distribution if the probability of success is small and the number of trials is large. We used this approximation because the Poisson probability mass function is easier to compute than the Binomial.

Probability Law: (Poisson Approximation to Binomial)

The Poisson distribution can be used as an approximation for the binomial distribution If p is "small" and n is large: (rule of thumb)

$$\underbrace{p \leq 0.05 \quad \text{and} \quad n \geq 20}_{\text{Binomial}} \qquad \approx \qquad \underbrace{\lambda = np}_{\text{Poisson}}$$

- Note d ifferent rules of thumb are used, for example, the rule of thumb $n \geq 100$, $p \leq 0.01$, and $np \leq 20$ is often used.

Example 20

Example 20

Suppose 1 in 5000 light bulbs are defective.

(a) What is the probability that at least 3 bulbs are defective in a batch of 10,000?

(b) Can the number of defective bulbs be approximated by a Poisson distribution and if yes, generate the corresponding Poisson approximation for the probability of a least 3 defective bulbs in a batch of 10,000.

BINOMIAL MODEL ‘success’ = a defective bulb

- n = number of bulbs $n = 5000$
- p = probability of a defective bulb $p = 0.0002$

POISSON MODEL

- $\lambda = np$ average number of defective bulb per batch of 5000. $\lambda = 1$