Chapter 4: Special Probability Distributions

1 Discrete Distributions

- Recall that for a discrete random variable X, the number of possible values (i.e., R_X) is **finite** or **countable**.
- The elements of R_X can be listed as $x_1, x_2, x_3, \ldots$.
- In this section, we study some classes of discrete random variables.

Discrete Uniform Distribution

Definition 1

- *If RV X assumes the values $x_1, x_2, \ldots, x_k$ with equal probability, then X follows a* ***discrete uniform distribution****.*

- *The p.f. for X is given by*

$$f_X(x) = \frac{1}{k}, \qquad x = x_1, x_2, \ldots, x_k,$$

and 0 otherwise.

THEOREM 2

Suppose X follows the discrete uniform distribution with $R_X = \{x_1, x_2, \ldots, x_k\}$, we have

- *The expectation is given by*

$$\mu_X = E(X) = \sum_{i=1}^{k} x_i f_X(x_i) = \frac{1}{k}\sum_{i=1}^{k} x_i.$$

- *The variance is given by*

$$\sigma_X^2 = V(X) = E(X^2) - (E(X))^2 = \frac{1}{k}\sum_{i=1}^{k} x_i^2 - \mu_X^2.$$

Example 4.1

- A bulb is selected at random from a box that contains a 40-watt bulb, a 60-watt bulb, an 80-watt bulb, and a 100-watt bulb.
- Each bulb has $1/4$ probability of being selected.
- Let $X =$ the watts of the bulb being selected. Then X follows a uniform distribution, and

$$R_X = \{40, 60, 80, 100\}.$$

$$f_X(x) = 1/4, \qquad \text{for } x = 40, 60, 80, 100,$$

and 0 otherwise.

- We can compute the expectation:

$$E(X) = \sum_i x_i f_X(x_i) = 40 \cdot (1/4) + 60 \cdot (1/4) + 80 \cdot (1/4) + 100 \cdot (1/4) = 70$$

- Variance can also be computed:

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= 40^2 \cdot (1/4) + 60^2 \cdot (1/4) + 80^2 \cdot (1/4) + 100^2 (1/4) - 70^2 \\ &= 500. \end{aligned}$$

L–example 4.1

- Toss a fair die, $X =$ the number on the top face. Then X follows a uniform distribution.
- $R_X = \{1, 2, 3, 4, 5, 6\}$, and

$$f_X(x) = {}^1\!/\!_6, \qquad \text{for } x = 1, 2, 3, 4, 5, 6,$$

and 0 otherwise.

- Expectation can be computed by

$$E(X) = \sum_i x_i f_X(x_i) = \sum_{i=1}^{6} i\left(\frac{1}{6}\right) = 3.5.$$

- Variance can be computed by

$= EX^2 - (EX)^2$

$$\begin{aligned} V(X) &= \sum_i x_i^2 f_X(x_i) - (E(X))^2 \\ &= \sum_{i=1}^{6} i^2\left(\frac{1}{6}\right) - 3.5^2 = \frac{35}{12}. \end{aligned}$$

Bernoulli Trial, Bernoulli Random Variable and Bernoulli Process

Definition 3 (Bernoulli Trial)

- *A **Bernoulli Trial** is a random experiment with only two possible outcomes.*
- *One is called a "success", and the other a "failure".*
- *We code the two outcomes as "1" (success) and "0" (failure).*

Definition 4 (Bernoulli Random Variable)

- *Let $X =$ number of success in a Bernoulli trial; then X has only two possible values: 1 or 0, and is called a* ***Bernoulli random variable****.*

- *Denote by p $(0 \leq p \leq 1)$ the probability of success of the Bernoulli trial. Then X has the p.f.:*

$$f_X(x) = P(X = x) = \begin{cases} p & x = 1 \\ (1-p) & x = 0 \end{cases},$$

and $= 0$ for other values of x.

- *This p.f. can also be written by*

$$p(X=x) = f_X(x) = p^x(1-p)^{1-x}, \quad \text{for } x = 0 \text{ or } 1.$$

- We often denote $X \sim \text{Bernoulli}(p)$, and denote $q = 1 - p$. Then the p.f. becomes $f_X(1) = p$ and $f_X(0) = q$.

THEOREM 5

For a Bernoulli RV defined above, we have

$$\begin{aligned} \mu_X &= E(X) = p \\ \sigma_X^2 &= V(X) = p(1-p) = pq. \end{aligned}$$

Remark (Parameters)

- In occasions, $f_X(x)$ may rely on one or more unknown quantities; different values of the quantities lead to different probability distributions.

- Such a quantity is called a **parameter** of the distribution.

- p is the parameter in the Bernoulli distribution.

- The collection of the distributions that are determined by one or more unknown parameters is called a **family of probability distributions**.

- So the aforementioned Bernoulli distributions determined by the parameter p is a family of probability distributions.

Example 4.2 The following are all examples of Bernoulli trials:

- A coin toss
  Say we want heads, then H="heads" is success, and T="tails" is failure.

- Rolling a die
  Say we only care about rolling a 6. The outcome space is binarized to "success"$= \{6\}$ and "failure" $= \{1, 2, 3, 4, 5\}$.

- Polls
  Choosing a voter at random to ascertain whether that voter will vote "yes" in an upcoming referendum.

Example 4.3

- A box contains 4 blue and 6 red balls.
- Draw a ball from the box at random.
- What is the probability that a blue ball is chosen?

Solution:

- Let $X = 1$ if a blue ball is drawn; and $X = 0$ otherwise.
- Then X is a Bernoulli random variable.
- $P(X = 1) = 4/10 = 0.4$.
- Furthermore, the p.f. for X is given by

$$f_X(x) = \begin{cases} 0.4 & x = 1 \\ 0.6 & x = 0 \end{cases}.$$

Definition 6 (Bernoulli Process)

- *A **Bernoulli process** consists of a sequence of repeatedly performed **independent and identical** Bernoulli trials.*

- *Correspondingly, a Bernoulli process generates a sequence of **independent and identically distributed, i.i.d.** Bernoulli random variables:* $X_1, X_2, X_3, \ldots$.

We are able to define several useful distributions based on Bernoulli trial and Bernoulli process. These distributions include:

- **Binomial distribution**;
- **Negative Binomial distribution**; **Geometric distribution**;
- **Poisson distribution**.

Binomial Distribution

If we have several (say n) i.i.d. Bernoulli trials, we can establish the binomial distribution to address some interesting questions. For example,

- A student randomly guesses at 5 multiple-choice questions. What is the number of questions the student guessed correctly?
- Randomly pick a family with 4 kids. What is the number of girls amongst the kids?
- Urn has 4 black balls and 3 white balls, draw 5 balls with replacement. How many black balls will there be?

DEFINITION 7 (BINOMIAL RANDOM VARIABLE)

*A **Binomial random variable** counts the number of successes in n trials in a Bernoulli Process. That is, suppose we have n trials where*

- *the probability of success for each trial is the same p,*
- *the trials are independent.*

Then the number of successes, denoted by X, in the n trials is a Binomial random variable.

We say X has a binomial distribution and write it as $X \sim B(n, p)$.

> *The probability of getting exactly x successes is given as*
>
> $$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x},\ \ for\ x=0,1,2,\ldots,n.$$
>
> *It can be shown that* $E(X)=np$, *and* $V(X)=np(1-p)$.

The theoretical development for Binomial distribution will be given in a lecture meeting.

L–example 4.2 (Theory of the Binomial Distribution)

- Based on the definition of binomial distribution: "X is the number of successes in n trials in a Bernoulli Process", so $X \sim B(n, p)$ **if and only if**

$$\boxed{X = X_1 + X_2 + \ldots + X_n,}$$

with $X_1, X_2, \ldots, X_n$ being i.i.d. Bernoulli(p) RVs.

- We are able to derive the p.f. for X as follows.

$\mathrm{Exp}(\lambda_1+\lambda_2) \sim Z$

$E\,Z = \boxed{\frac{1}{\lambda_1+\lambda_2}}$

$Y \sim B(m, p)$

$Y = Y_1 + \ldots + Y_m$

$X + Y \sim B(n+m, p)$

X_λ $\quad Y$

$E X = \frac{1}{\lambda_x}$ $\quad E Y = \frac{1}{\lambda_y}$

$E(X+Y) = \boxed{\frac{1}{\lambda_x} + \frac{1}{\lambda_y}}$

$X = X_1 + \ldots + X_n$

$Y = Y_1 + \ldots + Y_m$

- Consider a specific realization of $X_1, X_2, \ldots, X_n$, namely $x_1, x_2, \ldots, x_n$ such that $\sum_{i=1}^{n} x_n = \boxed{x}$.

$1 \quad 1 \quad 0 \; \ldots \; 1$

$1 \quad 0 \quad 1 \; \ldots \; 1$

- Because $X_1, X_2, \ldots, X_n$ are i.i.d. Bernoulli(p) RVs, we have

$$
\begin{aligned}
& P(\boxed{X_1 = x_1}, \boxed{X_2 = x_2}, \ldots, \boxed{X_n = x_n}) \neq P(X = x) \\
= \; & P(X_1 = x_1)P(X_2 = x_2)\ldots P(X_n = x_n) \\
= \; & \prod_{i=1}^{n} p^{x_i} q^{1-x_i} = p^{\sum_{i=1}^{n} x_i} q^{n - \sum_{i=1}^{n} x_i} \\
= \; & p^x q^{n-x}.
\end{aligned}
$$

$P(X_1 = x_1)$
$= p^{x_1}(1-p)^{1-x_1}$

$q = 1 - p$

- Note that $\sum_{i=1}^{n} x_i = x$ means: out of n trials, x are observed as success and the rest as failure.

$\underbrace{T_1, T_2, \ldots T_n}_{x \text{ successes}}$

- For the collection of n trials, how many such outcomes are possible? The answer is $\binom{n}{x}$, since it is equivalently to choosing x trials out of n to take success, and the rest take failure.

- Furthermore, for different choices of $x_1, x_2, \ldots, x_n$,

$$\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\}$$

are mutually exclusive events.

$= \{s \in S : X_1(s) = x_1,\ X_2(s) = x_2, \ldots X_n(s) = x_n\}$

$P(A \cup B) = P(A) + P(B)$ if $A \cap B = \phi$

- We have

$$\begin{aligned} P(X = x) &= P\left(\bigcup_{x_1,\ldots,x_n : \sum x_i = x} \{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\} \right) \\ &= \sum_{x_1,\ldots,x_n : \sum x_i = x} P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) \\ &= \sum_{x_1,\ldots,x_n : \sum x_i = x} \boxed{p^x q^{n-x}} = \binom{n}{x} p^x q^{n-x}. \end{aligned}$$

- We can also derive other characteristics of the binomial distribution based on the expression

$$X = X_1 + X_2 + \ldots + X_n.$$

$E(X_i) = p$

$V(X_i) = pq$

- Expectation is given by

$$E(X) = E(X_1) + E(X_2) + \ldots + E(X_n) = p + p + \ldots + p = np.$$

- Because of the independence of $X_1, X_2, \ldots, X_n$, variance is

$$\begin{aligned} V(X) &= V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n) \\ &= pq + pq + \ldots + pq = npq. \end{aligned}$$

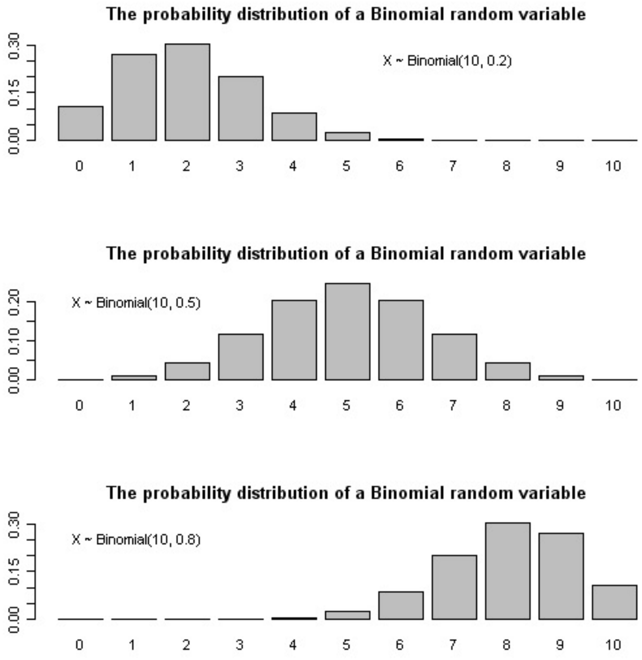
Remark

- When $n = 1$, the p.f. for the binomial RV X is reduced to

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

- It is the p.f. for the Bernoulli distribution. Therefore Bernoulli distribution is a special case of the binomial distribution.

The p.f. for $B(10, 0.2)$, $B(10, 0.5)$, and $B(10, 0.8)$ are compared below.



Example 4.4

- Flip a fair coin 10 independent times.
- What is the probability of observing exactly 6 heads?

Solution:

- Let $X =$ number of heads in 10 flips.
- Each flip of the coin can be observed as a Bernoulli trial, with probability of getting head (success) $p = 0.5$.
- Then X is the number success out of 10 Bernoulli trials; so $X \sim B(10, 0.5)$.
- We can compute

$$P(X = 6) = \binom{10}{6}(0.5)^6(1 - 0.5)^{10-6} = 0.205.$$

L–example 4.3 Pat Statsdud failed to study for the next statistics quiz. Pat's strategy is to rely on luck. The quiz consists of 10 multiple-choice questions. Each question has five possible answers, only one of which is correct. Pat plans to guess the answer to every question.

(a) What is the probability that Pat gets two answers correct?

(b) What is the probability that Pat fails the quiz? (suppose it is considered a failed quiz if a grade on the quiz is less than 50% , i.e. 5 questions out of 10).

< 5

Solution: Let X denote the number of correct answers. Then $X \sim B(10, 0.2)$.

(a) The probability that he gets two correct answers is given by

$$P(X = 2) = \binom{10}{2}(0.2)^2(0.8)^8 \approx 0.302.$$

(b) The probability that he fails is given by

$P(X<5)$

$F_X(4)$

$$\begin{aligned} P(\text{fail quiz}) &= P(X \leq 4) \\ &= P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) \\ &\approx 0.967. \end{aligned}$$

To compute $P(X \leq 4)$ for $X \sim B(10, 0.2)$:

(A) Method 1: use an online `R` compiler:

- Browse to `https://rdrr.io/snippets/`
- Enter the command

  ```
  pbinom(4, 10, 0.2, lower.tail = TRUE)
  ```

  unto the compiler.
- `Ctrl-Enter` or `Run` to obtain the answer.

- For $X \sim B(n, p)$.
  - `pbinom(x, n, p)` gives $P(X \leq x)$.
  - `pbinom(x, n, p, lower.tail=FALSE)` gives $P(X > x)$.
  - `dbinom(x, n, p)` gives $P(X = x)$.

(B) Method 2: use `R` Shiny app `Radiant`:

- Browse to `https://vnijs.shinyapps.io/radiant`
- Select `Basics > Probability Calculator`.
- Select `Binomial` as the `Distribution`.
- Select n as 10, p as 0.2.
- Select `Values` as the `Input type`.
- Select 4 as the upper bound, $P(X = 4)$, $P(X \leq 4)$, $P(X > 4)$ are included.

L–example 4.4

- A man claims to have extrasensory perception (ESP).
- As a test, a fair coin is flipped 10 times, and he is asked to predict the outcome in advance.
- The man gets 7 out of 10 correct.
- What is the probability that he would have done at least this well if he had no ESP? That is, he gets 7 or more out of 10 correct.

Solution:

- Without ESP, the probability that he guesses correctly for each outcome is 0.5.
- Let X = number of correct guesses out of 10 guesses. Then $X \sim B(10, 0.5)$.
- We have

$$P(X \geq 7) = P(X=7)+P(X=8)+P(X=9)+P(X=10)$$
$$= \binom{10}{7}0.5^7 0.5^3 + \binom{10}{8}0.5^8 0.5^2 + \binom{10}{9}0.5^9 0.5^1 + \binom{10}{10}0.5^{10} 0.5^0$$
$$= 0.1719.$$

Negative Binomial Distribution

- Consider a Bernoulli process, where the Bernoulli experiments can be repeated an arbitrary number of times.

- The interest could be how many trials are needed so that a certain number of successes occur.

- Set $X =$ number of trials until the kth success occurs. Then X follows a **negative binomial distribution**; denoted by $X \sim NB(k, p)$, where p is probability of success for each Bernoulli trial.

- In comparison with binomial distribution: the random variable "X" is the number of successes out of a fixed number n of trials.

DEFINITION 8 (NEGATIVE BINOMIAL DISTRIBUTION)

- X = *number of i.i.d. Bernoulli(p) trials until the kth success occurs; then X follows a* ***negative binomial distribution****, denoted by* $X \sim NB(k, p)$.

- *The p.f. of X is given by*

$$f_X(x) = P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k},$$

for $x = k, k+1, k+2, \ldots$

- *It can be shown that* $E(X) = k/p$ *and* $V(X) = (1-p)k/p^2$.

$(T_1 \cdots T_{x-1})(T_x)$

$k-1$ $\qquad$ S

L–example 4.5

- We derive the probability function of the negative binomial distribution.
- We can interpret the event $X = x$ as follows,

$$
\begin{aligned}
\{X = x\} &= \{\text{used } x \text{ trials until the } k\text{th success occurs}\} \\
&= \{\text{observe } k-1 \text{ successes in the first } x-1 \text{ trials}\} \\
&\quad \cap \{x\text{th trial is a success}\} \\
&= A \cap B.
\end{aligned}
$$

- Based on binomial distribution,

$$
\begin{aligned}
P(A) &= P(\text{observe } \boxed{k-1} \text{ successes in the first } \boxed{x-1} \text{ trials}) \\
&= \binom{x-1}{k-1} p^{\cancel{x}\,k-1}(1-p)^{x-k}
\end{aligned}
$$

$\leftarrow$ Binomial$(x-1, p) = Y$

$P(Y = k-1)$

- Since the last trial is the Bernoulli trial,

$$
P(B) = P(x\text{th trial is a success}) = p \leftarrow
$$

- A and B are independent; therefore, we have

$$
P(X = x) = P(A \cap B) = P(A)P(B) = \binom{x-1}{k-1} p^{\cancel{x}\,k-1}(1-p)^{x-k} \cdot p.
$$

Example 4.5

- Keep rolling a fair die, until the 6th time we get the number 6.
- What is the probability that we need to roll the die 10 times?

Solution:

- Let $X =$ number of rolls to get the 6th number 6. $X \sim NB(6, 1/6)$.
- Using the p.f. of negative binomial distribution:

$$P(X = 10) = \binom{10-1}{6-1}(1/6)^6(1-1/6)^4 = 0.001302.$$

L–example 4.6 In an NBA championship series, the team that **wins four games out of seven is the winner**. Suppose that teams A and B face each other in the championship games and that **team A has probability 0.55 of winning a game over team B**.

(a) What is the probability that team A will win the series in 6 games?

(b) What is the probability that team A will win the series?

Solution: Suppose that Teams A and B can continuously play games. Let

$$X = \text{number of games that A needs to win 4 games}$$

and for each game, the chance that A will win is 0.55. Therefore $X \sim NB(4, 0.55)$.

(a) The question is asking

$$P(X = 6) = \binom{6-1}{4-1} 0.55^4 (1 - 0.55)^{6-4} = \boxed{0.1853.}$$

$\{X=4\} = \{s : X(s) = 4\}$ $\{X=5\}$ $= \{s : X(s) = 5\}$

(b) The probability that Team A will win is

$$
\begin{aligned}
& P(X=4)+P(X=5)+P(X=6)+P(X=7) \\
= & \binom{4-1}{4-1} 0.55^4 (1-0.55)^{4-4} + \binom{5-1}{4-1} 0.55^4 (1-0.55)^{5-4} \\
& + \binom{6-1}{4-1} 0.55^4 (1-0.55)^{6-4} + \binom{7-1}{4-1} 0.55^4 (1-0.55)^{7-4} \\
= & 0.6083.
\end{aligned}
$$

Question: Can Part (b) be solved using binomial distribution instead?

Y = # of wins out of 7 by A

$Y \sim B(7, 0.55)$

$P\{Y \geq 4\}$

For $X \sim NB(k, p)$, we can use an online `R` compiler:

- Browse to `https://rdrr.io/snippets/`
- Command:
  - `dnbinom(x-k, k, p)` computes $P(X = x)$;
  - `pnbinom(x-k, k, p)` computes $P(X \leq x)$;
  - `pnbinom(x-k, k, p, lower.tail = F)` computes $P(X > x)$.

$P(X \geq x)$

$P(X \geq 6) = P(X > 7)$

Geometric Distribution

Geometric distribution is a special case of the negative binomial distribution.

Definition 9 (Geometric Distribution)

- X = *number of i.i.d. Bernoulli(p) trials until the first success occurs; then X follows a* ***geometric distribution****, denote by $X \sim G(p)$.*
- *The p.f. of X is given by*

$$f_X(x) = P(X = x) = (1-p)^{x-1}p.$$

- *We have $E(X) = 1/p$ and $V(X) = (1-p)/p^2$.*

L–example 4.7

- At a “busy time", a telephone exchange is very near capacity, so callers have difficulty placing their calls.
- It may be of interest to know the number of attempts necessary in order to make a connection.
- Suppose that we let $p = 0.05$ be the probability of connection during a busy time.

$P(X=5)$

- We are interested in knowing the probability that 5 attempts are necessary for a successful call.

$P(X \le 5)$

Solution:

- Let X = number of attempts needed for the first successful call.
- Then $X \sim G(p)$ or $X \sim NB(1, p)$, where $p = 0.05$.
- We have

$$P(X = 5) = (1-p)^{5-1}p = 0.95^4(0.05) = 0.0407.$$

Poisson Distribution

DEFINITION 10 (POISSON RANDOM VARIABLE)
The ***Poisson random variable*** *X denotes the number of events occurring in a* ***fixed period of time or fixed region****.*

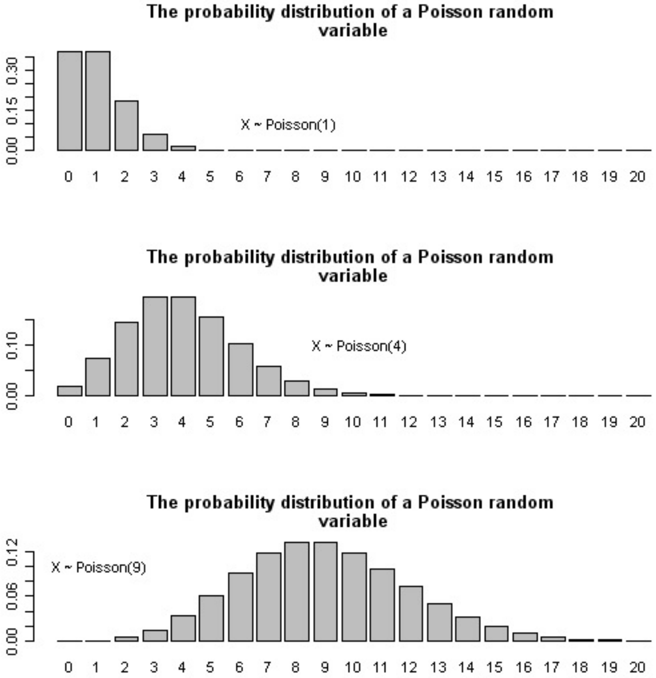
We denote $X \sim \text{Poisson}(\lambda)$ where parameter $\lambda > 0$ is the expected number of occurrences during the given period/region; its p.m.f. is given by

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

where $k = 0, 1, \ldots$ is the number of occurrences of events.

It can be shown that $\underline{E(X)} = \lambda$, and $\underline{V(X) = \lambda}$.

The p.f. for Poisson(1), Poisson(4), and Poisson(9) are compared below.



Example 4.6 The "fixed period of time or fixed region" given in the definition can be time period of any length, e.g., a minute, a day, a week, a month etc., and region of any size.

Examples of events that may be modeled by the Poisson Distribution:

(a) The number of spelling mistakes one makes while typing a single page.

(b) The number of times a web server is accessed per minute.

(c) The number of road kill (animals killed) found per unit length of road.

(d) The number of mutations in a given stretch of DNA after a certain amount of radiation exposure.

(e) The number of unstable atomic nuclei that decayed within a given period of time in a piece of radioactive substance.

(f) The distribution of visual receptor cells in the retina of the human eye.

(g) The number of light bulbs that burn out in a certain amount of time.

L–example 4.8 The number of infections X in a hospital each week has been shown to follow a Poisson distribution with a mean of $\overline{3.0}$ infections per week. What is the probability that

λ

(a) there is **no** infection for a week?

(b) there are **less than** 4 infections for a week?

<u>Solution</u>: It follows that

(a) $P(X=0)=e^{-3}$.

$$\boxed{e^{-x}\,\frac{x^x}{x!}}$$

$0!=1$

(b) $P(X<4)=e^{-3}\left(1+3+\frac{3^2}{2!}+\frac{3^3}{3!}\right)$.

$=P(X=0)+P(X=1)+P(X=2)+P(X=3)$

Numerical computation for $X \sim \text{Poisson}(\lambda)$:

(A) Online `R` compiler: `https://rdrr.io/snippets/`

- `dpois(x, lambda)` computes $P(X = x)$;
- `ppois(x, lambda)` computes $P(X \leq x)$;
- `ppois(x, lambda, lower.tail = F)` computes $P(X > x)$.

(B) use `R` Shiny app `Radiant`: `https://vnijs.shinyapps.io/radiant`; similar steps as Binomial distribution to do the computation.

L–example 4.9

- If the average number of oil tankers arriving each day at a port is known to be 10. $= \lambda$
- The facilities at the port can handle at most 15 tankers per day.
- What is the probability that on a given day tankers will have to be sent away?

Solution:

- Let X = number of tankers arriving each day.
- We have $X \sim \text{Poisson}(\lambda)$, where $\lambda = 10$.

$= P(X \le 15)$

$$
\begin{aligned}
P(X > 15) &= \sum_{x=16}^{\infty} \frac{e^{-10} 10^x}{x!} = 1 - \boxed{\sum_{x=0}^{15} \frac{e^{-10} 10^x}{x!}} \\
&= 1 - e^{-10}\left(1 + 10 + \frac{10^2}{2!} + \ldots + \frac{10^{15}}{15!}\right) \\
&= 0.0487.
\end{aligned}
$$

L–example 4.10 We derive $E(X)$ and $V(X)$, for $X \sim \text{Poisson}(\lambda)$.

- For these derivation, the fundamental idea is to use the fact that for p.m.f. $f_X(x)$, we must have

$$\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1} \qquad \sum_{x \in R_X} f_X(x) = 1. \qquad \boxed{\sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} = 1}$$

- We derive $E(X)$ first.

$$\boxed{\begin{array}{l} Ee^X \\ Ee^{Xt} \\ \text{for } t \in \mathbb{R} \end{array}}$$

$$E(X) = \sum_{x=0}^{\infty} x f_X(x) = \sum_{x=0}^{\infty} x \cdot \boxed{\frac{e^{-\lambda}\lambda^x}{x!}} = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} \qquad = \lambda^{y+1} = \lambda^y \cdot \lambda$$

$$= \lambda \boxed{\sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!}} = \lambda, \quad \text{set } y = x - 1.$$

$x = y + 1$

- We derive $V(X)$ next.

$= EX^2 - (EX)^2 = (\lambda^2+\lambda) - (\lambda)^2$

$$E(X(X-1)) = \sum_{x=0}^{\infty} x(x-1)\boxed{\frac{e^{-\lambda}\lambda^x}{x!}} = \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-2)!} \quad = \lambda^{y+2}$$

$= EX^2 - EX$

$$= \boxed{\lambda^2}\boxed{\sum_{y=0}^{\infty} \frac{e^{-\lambda}\boxed{\lambda^y}}{y!}} = \lambda^2, \quad \text{set } y = x-2.$$

$EX^2 = \lambda^2 + \lambda$

$x = y + 2$

We can compute $V(X)$ by

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 = E(X(X-1)) + E(X) - [E(X)]^2 \\ &= \lambda^2 + \lambda - \lambda^2 = \lambda. \end{aligned}$$

$P(\alpha)$

$P(T\alpha)$

DEFINITION 11 (POISSON PROCESS)

*The **Poisson Process** is a continuous time process. We count the number of occurrences within some interval of time. The defining properties of a Poisson Process with rate parameter α are*

$\boxed{X \sim P(\lambda_1),\ Y \sim P(\lambda_2),\ X \perp Y \quad X+Y \sim P(\lambda_1+\lambda_2)}$

- *the expected number of occurrences in an interval of length T is αT;*

$E[X+Y] = \lambda_1+\lambda_2$ $\quad X+Y \sim \text{Poisson}(\lambda_1+\lambda_2)$

- *there are no simultaneous occurrences;*

$\underbrace{}_{X,\ T_1}\ \underbrace{}_{Y,\ T_2}$ $\quad T_1 : T_2 = \lambda_1 : \lambda_2$

- *the number of occurrences in disjoint time intervals are independent.*

The number of occurrences in any interval T of a Poisson Process follows a Poisson(αT) *distribution.*

Example 4.7

- The average number of robberies in a day is four in a certain big city.
- What is the probability that six robberies occurring in two days?

Solution:

- Let X_1 = number of robberies in one day. Then $X_1 \sim \text{Poisson}(4)$ from the condition.
- Let X = number of robberies in two days. Then $X \sim \text{Poisson}(2 \times 4) = \text{Poisson}(8)$.
- We have

$$P(X = 6) = \frac{e^{-8}8^6}{6!} = 0.1222.$$

L–example 4.11

A can company reports that the number of breakdowns per 8 hour shift on its machine-operated assembly line follows a Poisson distribution, with a mean of 1.5.

(a) What is the probability of exactly two breakdowns during the midnight shift?

(b) What is the probability of fewer than two breakdowns during the afternoon shift?

(c) What is the probability that no breakdown during three consecutive 8-hour shifts?

Solution: Let X = number of breakdowns in an 8 hour shift. We have $X \sim \text{Poisson}(\lambda)$ with $\lambda = 1.5$.

(a) The probability of exactly 2 breakdowns during the night shift is

$$P(X=2) = \frac{e^{-1.5}1.5^2}{2!} = 0.251.$$

(b) The probability of fewer than 2 breakdowns during the afternoon shift is

$$\begin{aligned} P(X<2) &= P(X=0) + P(X=1) \\ &= \frac{e^{-1.5}1.5^0}{0!} + \frac{e^{-1.5}1.5^1}{1!} = 0.5578. \end{aligned}$$

(c)

- Let Y_1 be a Bernoulli RV, where $Y_1 = 1$ if there is no breakdowns in the 1st 8 hour shift; and $Y_1 = 0$ otherwise. The probability of success is

$$p = P(Y_1 = 1) = P(X = 0) = \boxed{\frac{e^{-1.5}1.5^0}{0!}} = 0.2231.$$

- Similarly define Y_2 and Y_3 as Bernoulli RVs, $Y_i = 1$ if no breakdown in the ith hour shift; and $Y_i = 0$ otherwise; for $i = 2, 3$.
- Then Y_1, Y_2, Y_3 are i.i.d. Bernoulli(p) RVs. Set $Y = Y_1 + Y_2 + Y_3$; then $Y \sim B(3, p)$. On the other hand Y is counting the number of 8-hour shifts without breakdowns.

- "$Y = 3$" stands for the practical situation that no breakdown during three consecutive 8-hour shifts.

$$P(Y=3) = \binom{3}{3} p^3 (1-p)^0 = 0.0111.$$

- An alternative method: using Poisson process, the number of breakdowns in $24 = 3 \times 8$ hours, denoted by RV Z, follows a Poisson$(3 \times 1.5) =$ Poisson(4.5) distribution. The question is asking

$$P(Z=0) = \frac{e^{-4.5} 4.5^0}{0!} = 0.0111.$$

PROPOSITION 12 (POISSON APPROX. OF BINOMIAL DISTRIBUTION) *Let $X \sim B(n, p)$. Suppose that $n \to \infty$ and $p \to 0$ in such a way that $\lambda = np$ remains a constant. Then approximately, $X \sim \text{Poisson}(np)$. That is*

$$\lim_{p \to 0; n \to \infty} P(X = x) = \frac{e^{-np}(np)^x}{x!}.$$

The approximation is good when $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$.

Example 4.8

- The probability, p, of an individual car having an accident at a junction is 0.0001.
- If there are 1000 cars passing through the junction during certain period of a day, what is the probability of two or more accidents occurring during that period?

Solution:

- Let $X =$ number of accidents among the 1000 cars.
- Then $X \sim B(1000, 0.0001)$. If we compute using binomial distribution,

$$P(X \geq 2) = \sum_{x=2}^{1000} \binom{1000}{x} 0.0001^x 0.9999^{1000-x}.$$

- Computing these numbers is not easy.

- We solve the question using Poisson approximation.
- $n = 1000$ and $p = 0.0001$, hence, $np = \lambda = 0.1$.
- Thus

$$
\begin{aligned}
P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\
&= 1 - e^{-0.1} - e^{-0.1}(0.1)^1/1! \\
&= 0.0047.
\end{aligned}
$$

L–example 4.12

- In a manufacturing process in which glass items are being produced, defects or bubbles occur, occasionally rendering the piece undesirable for marketing.
- It is known that on the average 1 in every 1000 of these items produced has one or more bubbles.

  $p = 0.001$

- What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

Solution:

- Let $X =$ number of items processing bubbles.
- Then $X \sim B(8000, 0.001)$.

$(0.999)^{8000}$

- Use the Poisson approximation, $\lambda = np = 8000 \times 0.001 = 8$, and hence $X \boxed{\approx} \text{Poisson}(\lambda)$.
- The (approximate) probability is

$$P(X < 7) = 1 - P(X \geq 7) \approx 1 - 0.6866 = 0.3134.$$

2 Continuous Distribution

- For a continuous random variable X, its range R_X is an interval or a collection of multiple intervals.
- In this section, we study some classes of continuous random variables.

Continuous Uniform Distribution

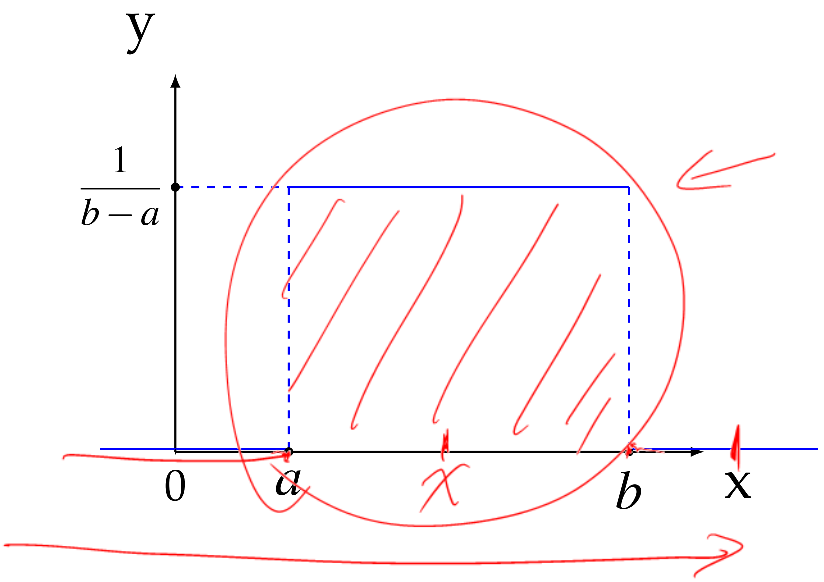
Definition 13 (Continuous Uniform Distribution)
A random variable X is said to follow a ***uniform distribution*** *over the interval (a,b) if its probability density function is given by*

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}.$$

We denote this by $X \sim U(a,b)$.

It can be shown that $E(X) = \dfrac{a+b}{2}$ and $V(X) = \dfrac{(b-a)^2}{12}$.

The p.d.f. for the continuous uniform distribution can be drawn as a figure below.



The c.d.f. for the continuous uniform distribution is given by

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Example 4.9

- A point is chosen at random on the line segment $[0, 2]$.
- What is the probability that the chosen point lies between 1 and $3/2$?

Solution:

- Let X = position of the point. $X \sim U(0,2)$.
- We have

$$f_X(x) = \frac{1}{2}, \quad \text{for } 0 \le x \le 2,$$

and 0 otherwise.

$$P\left(1 \le X \le \frac{3}{2}\right) = \int_1^{3/2} \frac{1}{2} dx = \frac{1}{2} x\Big|_1^{3/2} = 1/4.$$

L–example 4.13

- Buses arrive at a specified stop at 15-minute intervals starting at 7:00 am.
- That is, they arrive at 7:00, 7:15, 7:30, 7:45, and so on.
- If a passenger arrives at the stop at a time that is uniformly distributed between 7:00 and 7:30, find the probability that he waits less than 5 minutes for a bus.

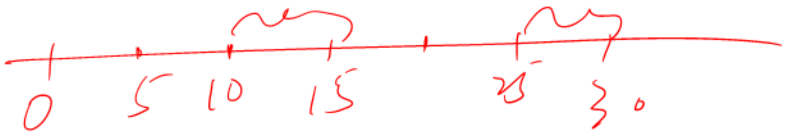
Solution: Let X denote the arrival time of the passenger (after 7:00am, in minutes). Then $X \sim U(0, 30)$.

The passenger waits less than 5 minutes for a bus when and only when he arrives (a) between 7:10-7:15 or (b) 7:25-7:30. So the desired probability is

$$P(10 < X < 15 \text{ or } 25 < X < 30)$$

$$= P(10 < X < 15) + P(25 < X < 30) = \frac{15-10}{30} + \frac{30-25}{30} = \frac{1}{3}.$$

$$= \int_{10}^{15} \frac{1}{30} \, dx$$



L–example 4.14 For the continuous uniform distribution, we derive

$$E(X) = \frac{a+b}{2}; \quad V(X) = \frac{1}{12}(b-a)^2.$$

- We derive $E(X)$ first

$$\begin{aligned} E(X) &= \int_a^b x \cdot \overbrace{\frac{1}{b-a}}^{f(x)} dx = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b \\ &= \frac{1}{b-a} \cdot \frac{b^2-a^2}{2} = \frac{a+b}{2}. \end{aligned}$$

- We derive $V(X)$ next,

$= EX^2 - (EX)^2$

$$E(X^2) = \int_a^b x^2 \cdot \boxed{\frac{1}{b-a}} dx = \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b$$

$$= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}.$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{1}{12}(a^2 - 2ab + b^2) = \frac{(b-a)^2}{12}.$$

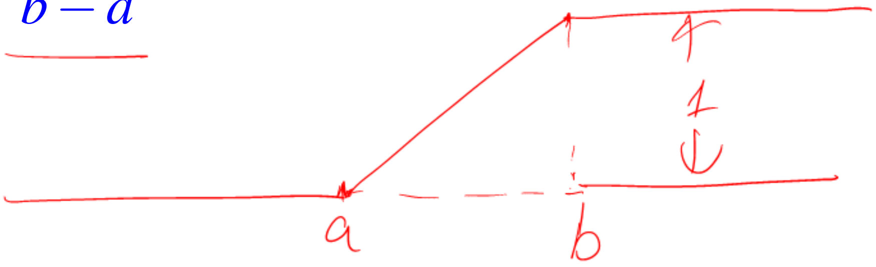
L–example 4.15 We derive the c.d.f. of a continuous uniform distribution.

- We take note that $F_X(x) = 0$ when $x < a$, and $F_X(x) = 1$ when $x > b$.
- When $a \le x \le b$,

1

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \overbrace{\int_{-\infty}^{a} 0\, dt}^{0} + \int_{a}^{x} \frac{1}{b-a}dt$$

$$= \frac{1}{b-a} \cdot t\Big|_a^x = \frac{x-a}{b-a}.$$



Exponential Distribution

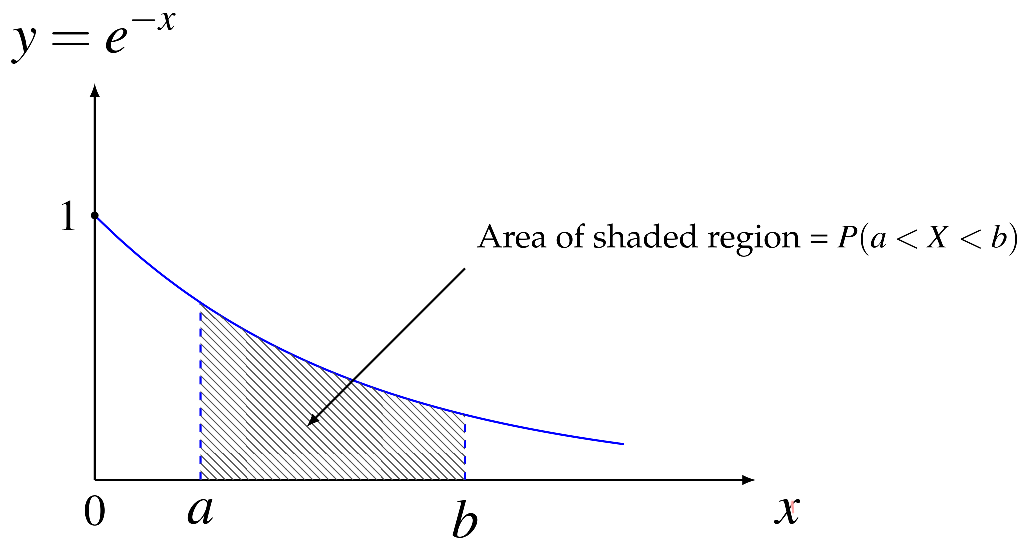
DEFINITION 14 (EXPONENTIAL DISTRIBUTION)
A continuous RV X is said to follow an ***exponential distribution*** *with parameter $\lambda > 0$ if its p.d.f. is given by*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}.$$

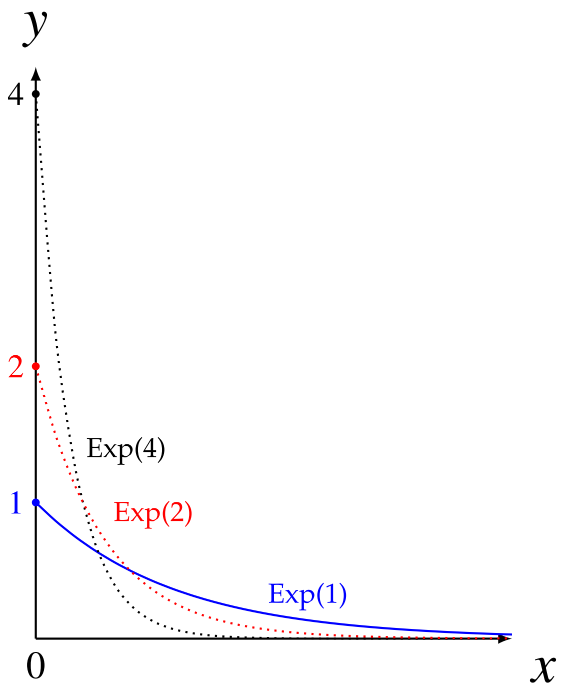
We denote $X \sim \mathrm{Exp}(\lambda)$.

It can be shown that $\boxed{E(X) = \frac{1}{\lambda}}$ and $\boxed{V(X) = \frac{1}{\lambda^2}}$.

The exponential p.d.f. with $\lambda = 1$ is shown below.



The shapes of the p.d.f.s of Exp(λ) for $\lambda = 1, 2, 4$.



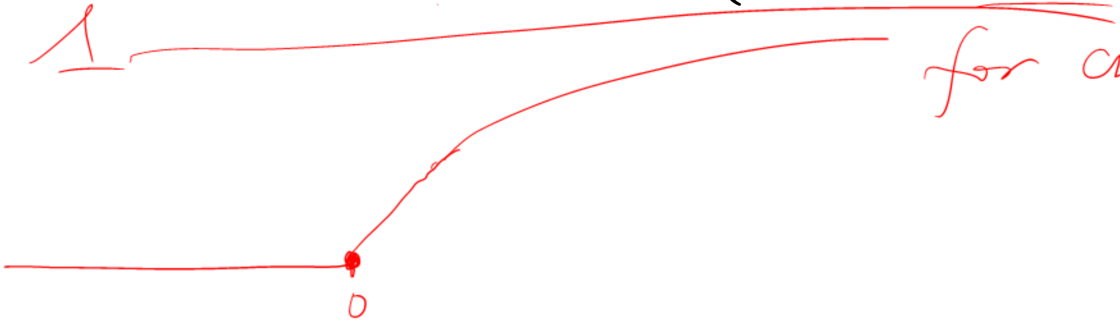
$1 - e^{-\lambda 0} = 0$

The c.d.f. of $X \sim \mathrm{Exp}(\lambda)$ is given by

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$\leftarrow$ $x \geq 0$
$x < 0$

for any x



Remark

- The p.d.f. can be written in the form

$$f_X(x) = \frac{1}{\mu} e^{-x/\mu}, \quad \text{for } x > 0,$$

  and 0 elsewhere.

- The parameters have the relationship $\boxed{\mu = 1/\lambda}$.

- We have

$$\boxed{E(X) = \mu}, \quad V(X) = \mu^2, \quad \text{and} \quad F_X(x) = 1 - e^{-x/\mu} \quad \text{for } x > 0.$$

Example 4.10

- Suppose that the failure time, T, of a system is exponentially distributed, with a mean of 5 years.
- What is the probability that at least two out of five of these systems are still functioning at the end of 8 years?

Solution:

- Since $E(T) = 5$, therefore $\lambda = 1/5$.
- We have $T \sim \text{Exp}(1/5)$,

$$P(T > 8) = 1 - P(T \leq 8) = 1 - F_X(8) = e^{-(1/5)\times 8} = e^{-1.6} \approx 0.2.$$

- Let X = # of systems out of 5 that are still functioning after 8 years.
- Then $X \sim B(5, 0.2)$. Hence,

$$P(X \geq 2) = 0.2627.$$

L–example 4.16 Let X = response time at a certain on-line computer terminal (the elapsed time between the end of a user's inquiry and the beginning of the system's response to that inquiry). X follows an exponential distribution with expected response time equal to 5 seconds.

$EX = 5$

(a) Find the probability that the response time is at most 10 seconds.

(b) Find the probability that the response time is between 5 and 10 seconds.

<u>Solution</u>: Since $E(X) = 5$, we have $X \sim \text{Exp}(1/5)$.

(a)

$$P(X \leq 10) = 1 - e^{-10/5} = 0.8647.$$

$= F_X(10)$

(b)

$$\begin{aligned} P(5 \leq X \leq 10) &= P(X \leq 10) - P(X < 5) \\ &= (1 - e^{-10/5}) - (1 - e^{-5/5}) = 0.2326. \end{aligned}$$

Numerical computation for $\text{Exp}(\lambda)$ distribution:

(A) Online `R` compiler: `https://rdrr.io/snippets/`

- `dexp(x, lambda)` computes $f_X(x)$;
- `pexp(x, lambda)` computes $P(X \leq x)$;
- `pexp(x, lambda, lower.tail = F)` computes $P(X > x)$.

(B) use `R` Shiny app `Radiant`: `https://vnijs.shinyapps.io/radiant`; similar steps as Binomial distribution to do the computation.

L–example 4.17 We derive $E(X)$ and $V(X)$ for the exponential distribution.

$$
\begin{aligned}
E(X) &= \int_0^\infty x\lambda e^{-\lambda x}dx = \int_0^\infty x d\left(-e^{-\lambda x}\right) \\
&= -xe^{-\lambda x}\Big|_0^\infty - \int_0^\infty \left(-e^{-\lambda x}\right)dx = \int_0^\infty \left(e^{-\lambda x}\right)dx \\
&= -\frac{1}{\lambda}e^{-\lambda x}\Big|_0^\infty = \frac{1}{\lambda}.
\end{aligned}
$$

$$\lim_{x\to\infty} x^a e^{-\lambda x} = 0 \qquad a>0$$

$$\begin{aligned} E(X^2) &= \int_0^\infty x^2 \boxed{\lambda e^{-\lambda x}} dx = \int_0^\infty x^2 d\left(-e^{-\lambda x}\right) \\ &= -x^2 e^{-\lambda x}\Big|_0^\infty - \int_0^\infty \left(-e^{-\lambda x}\right) d(x^2) \\ &= \frac{2}{\lambda} \boxed{\int_0^\infty x\lambda e^{-\lambda x} dx} = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2}. \end{aligned}$$

Hence,

$$V(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

L–example 4.18 Find the c.d.f. of the exponential distribution with parameter λ.

<u>Solution</u>:

- For $x \geq 0$,

$$F_X(x) = P(X \leq x) = \int_0^x \boxed{\lambda e^{-\lambda t}}\, dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}, \quad < 1$$

and 0 otherwise. $x < 0$

- Also, we have

$$S(x) = P(X > x) = \boxed{e^{-\lambda x},} \quad \text{for } x > 0.$$

Theorem 15

Suppose that X has an exponential distribution with parameter $\lambda > 0$. Then for any two positive numbers s and t, we have

$$P(X > s + t | X > s) = P(X > t).$$

REMARK

The above theorem states that the exponential distribution has "**no memory**" in the sense:

- Let X denote the life length of a bulb.
- Given that the bulb has lasted s time units, i.e., $X > s$,
- the probability that it will last for the next t units, i.e., $X > s + t$, is the same as the probability that it will last for the first t units as brand new.

L–example 4.19 We verify the no memory property of exponential distribution.

$$P(\underbrace{X > \overset{90+10}{s+t}}_{A} | \underbrace{X > \overset{90}{s}}_{B}) = \frac{P(\overbrace{\{X > s+t\}}^{A} \overset{\cap}{\cap} \overbrace{\{X > s\}}^{B})}{P(\underbrace{X > s}_{B})}$$

$$= \frac{P(X > s+t)}{P(X > s)} = \frac{\boxed{e^{-\lambda(s+t)}}}{\boxed{e^{-\lambda s}}} = e^{-\lambda t} = P(X > \overset{10}{t}).$$

Normal Distribution

DEFINITION 16 (NORMAL DISTRIBUTION)

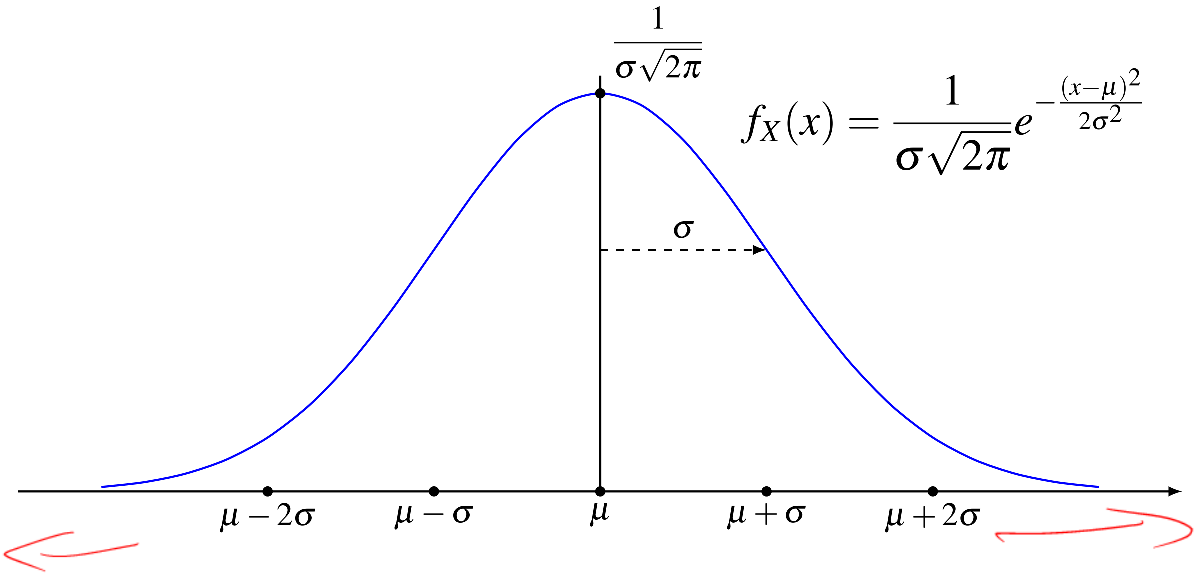
A random variable X is said to follow a ***normal distribution*** *with parameters μ and σ^2 if its probability density function is given by*

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}}, \qquad -\infty < x < \infty.$$

We denote $X \sim N(\mu, \sigma^2)$.

It can be shown that $\underline{E(X) = \mu}$ and $\underline{V(X) = \sigma^2}$.

The p.d.f. of normal distribution is positive over the whole real line, symmetric about $x = \mu$, and bell-shaped; see below.



We give some properties of normal distribution.

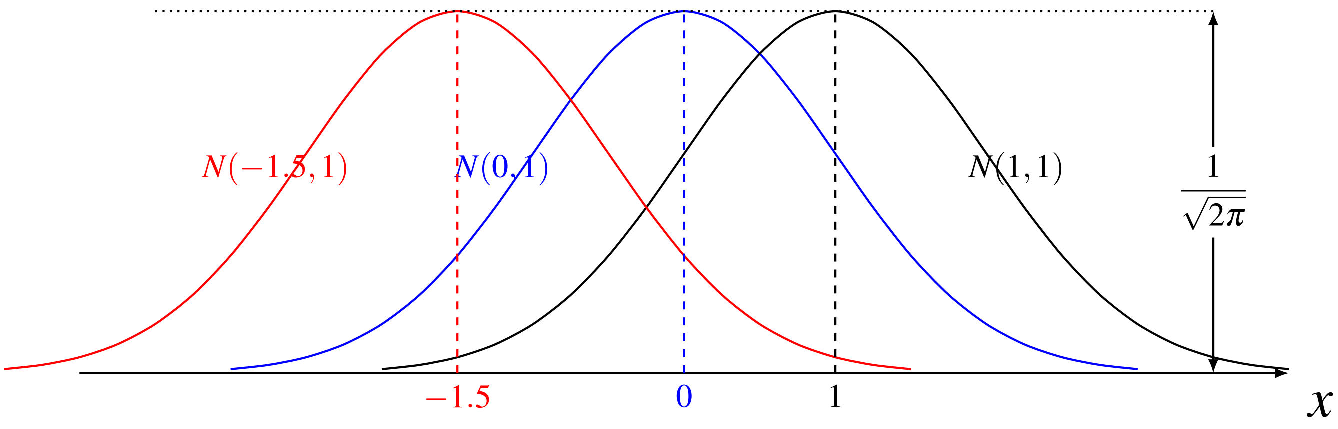
(1) The total area under the curve and above the horizontal axis is equal to 1.

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = 1.$$

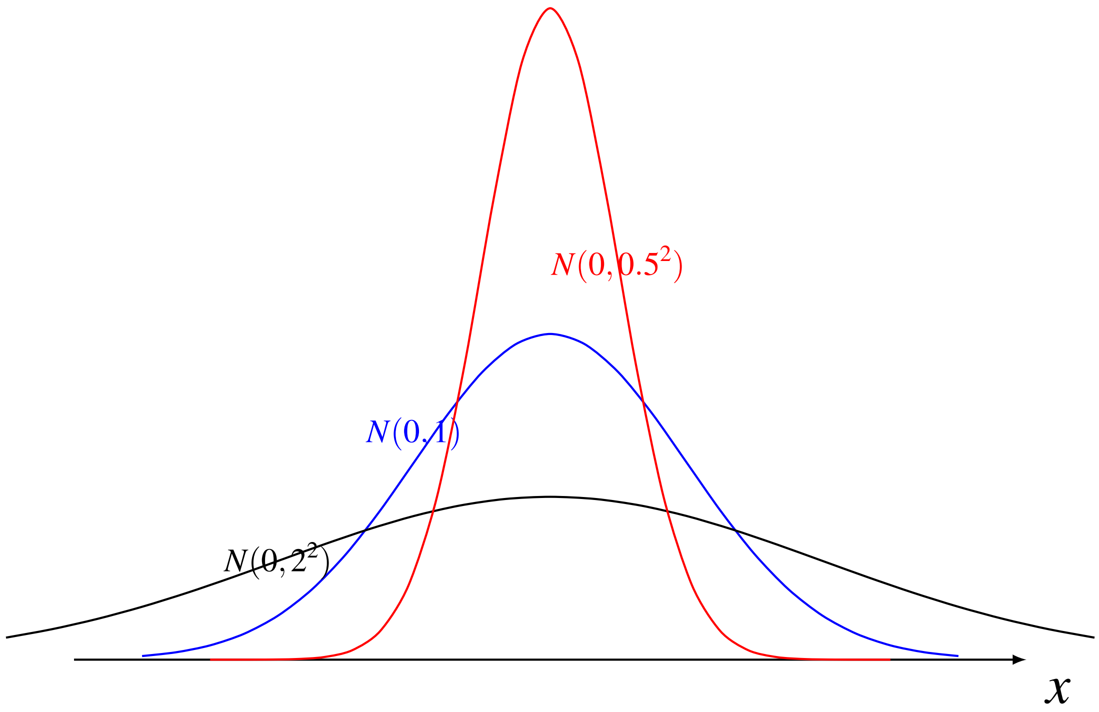
This validates that $f_X(\cdot)$ is a p.d.f.

$$\int_{-\infty}^{+\infty} e^{-x^2}\, dx$$

(2) Two normal curves are identical in shape if they have the same σ^2. But they are centered at different positions when their means are different.



(3) As σ increases, the curve flattens; and vice versa.



(4) If $X \sim N(\mu, \sigma^2)$ and let

$$\boxed{Z = \frac{X - \mu}{\sigma},}$$

then Z follows the $N(0,1)$ distribution. Thus $E(Z) = 0$ and $V(Z) = 1$. We say that Z has a standardized normal distribution; the p.d.f. of Z is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).$$

REMARK

- The importance of the standardized normal distribution is that it can be tabulated.

- Consider $X \sim N(\mu, \sigma^2)$; if we are to compute $P(x_1 < X < x_2)$ for any real values x_1, x_2, we can use the transformation $Z = (X - \mu)/\sigma$. In particular,

$$x_1 < X < x_2 \Longleftrightarrow \frac{x_1 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{x_2 - \mu}{\sigma}$$

- Let $z_1 = (x_1 - \mu)/\sigma$ and $z_2 = (x_2 - \mu)/\sigma$; then

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2).$$

- By convention, we use $\phi(\cdot)$ and $\Phi(\cdot)$ to denote the p.d.f. and c.d.f. of the standard normal distribution respectively. That is,

24150

$$
\begin{aligned}
\phi(z) &= f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \\
\Phi(z) &= \cancel{\frac{1}{\sqrt{2\pi}}} \int_{-\infty}^{z} \boxed{\phi(t)}\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt.
\end{aligned}
$$

- Therefore, for $X \sim N(\mu, \sigma^2)$ and any real numbers x_1, x_2,

$$
P(x_1 < X < x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right).
$$

- However, calculating the probabilities for the normal probabilities is challenging because
  - there is no close formula for $\Phi(z)$;
  - so the computation relies on numerical integration.
- Instead, $\Phi(z)$ can be tabulated, or computed based on some statistical software.

- The standard normal distribution has the following properties:
  - ⋆ $P(Z \geq 0) = P(Z \leq 0) = \Phi(0) = 0.5$;
  - ⋆ For any z, $\Phi(z) = P(Z \leq z) = P(Z \geq -z) = 1 - \Phi(-z)$;
  - ⋆ $-Z \sim N(0, 1)$;
  - ⋆ If $Z \sim N(0, 1)$, then $\sigma Z + \mu \sim N(\mu, \sigma^2)$.

Example 4.11 Given $X \sim N(50, 100)$, find $P(45 < X < 62)$.

<u>Solution</u>: We have $\mu = 50$, $\sigma = 10$.

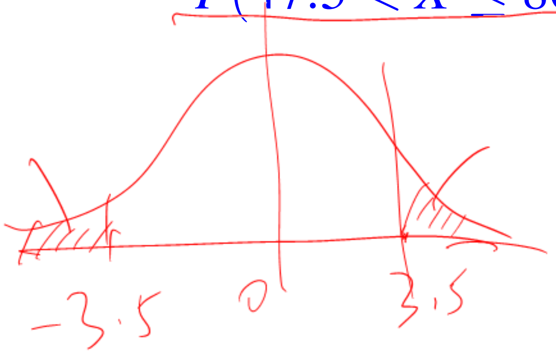
$$\begin{aligned} P(45 < X < 62) &= P\left(\frac{45-50}{10} < \frac{X-50}{10} < \frac{62-50}{10}\right) \\ &= P(-0.5 < Z < 1.2) \\ &= P(Z < 1.2) - P(Z \leq -0.5) \\ &= \Phi(1.2) - \Phi(-0.5), \end{aligned}$$

where $\Phi(1.2)$ and $\Phi(-0.5)$ can either be computed from some statistical software or obtained from a table.

L–example 4.20 When $X \sim N(65, \underset{\sigma^2}{25})$, compute $P(47.5 < X \leq 80)$.

$= \dfrac{X-\mu}{\sigma} = Z$

Solution: we have $\mu = 65$, $\sigma = 5$;



$$
\begin{aligned}
P(47.5 < X \leq 80) &= P\left(\frac{47.5-65}{5} < \boxed{\frac{X-65}{5}} \leq \frac{80-65}{5}\right)\\
&= P(-3.5 < \boxed{Z} \leq 3)\\
&= P(Z \leq 3) - P(Z \leq -3.5)\\
&= P(Z \leq 3) - P(Z \geq 3.5)\\
&= P(Z \leq 3) - (1 - P(Z < 3.5)) \quad \Phi(3.5)\\
&= \underbrace{0.99865}_{\Phi(3)} - 1 + 0.999767 = 0.998417.
\end{aligned}
$$

Numerical computation for $X \sim N(\mu, \sigma^2)$:

(A) Online `R` compiler: `https://rdrr.io/snippets/`

- `dnorm(x, mu, sigma)` computes $f_X(x)$;
- `pnorm(x, mu, sigma)` computes $P(X \leq x)$;
- `pnorm(x, mu, sigma, lower.tail = F)` computes $P(X > x)$.

pnorm(x) = P(Z ≤ x)

(B) use `R` Shiny app `Radiant`: `https://vnijs.shinyapps.io/radiant`; similar steps as Binomial distribution to do the computation.

L–example 4.21

- An expert witnesses in a paternity suit testifies that the length (in days) of pregnancy is approximately normally distributed with parameters $\mu = 270$ and $\sigma = 10$.
- The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth.
- If the defendant was, in fact, the father of the child, what is the probability that the mother could have had a very long or a very short pregnancy indicated by the testimony?

Solution: Let X denote the length of the pregnancy and assume that the defendant is the father; then $X \sim N(270, 10^2)$. The probability of the birth could occur within the indicated duration is

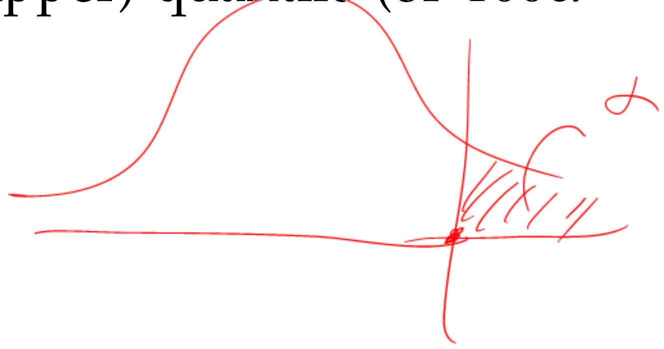
$$
\begin{aligned}
& P(X > 290 \text{ or } X < 240) \\
= \ & P(X > 290) + P(X < 240) \\
= \ & P\left(\frac{X-270}{10} > \frac{290-270}{10}\right) + P\left(\frac{X-270}{10} < \frac{240-270}{10}\right) \\
= \ & 1 - \Phi(2) + \Phi(-3) \\
= \ & 1 - \Phi(2) + [1 - \Phi(3)] = 0.0241.
\end{aligned}
$$

Definition 17 (Quantile)
The αth (upper) quantile ($0 \leq \alpha \leq 1$) of the RV X is the number x_α that satisfies

$$P(X \geq x_\alpha) = \alpha.$$

- Specifically, we denote by z_α the αth (upper) quantile (or 100α percentage point) of $Z \sim N(0,1)$. That is

$$P(Z \geq z_\alpha) = \alpha.$$



- For example, $z_{0.05} = 1.645$, $z_{0.01} = 2.326$.

- Since the p.d.f. of Z, i.e., $\phi(z)$, is symmetrical about 0, therefore

$$P(Z \geq z_\alpha) = P(Z \leq -z_\alpha) = \alpha.$$

Example 4.12 Find z such that

(a) $P(Z < z) = 0.95$;

(b) $P(|Z| \leq z) = 0.98$.

Solution:

(a) We need z such that

$$P(Z > z) = 1 - P(Z < z) = 0.05;$$

therefore $z = z_{0.05} = 1.645$.

(b) We have

$$\begin{aligned} 0.98 &= P(|Z| \leq z) = 1 - P(|Z| > z) \\ &= 1 - P(Z > z) - P(Z < -z) = 1 - 2P(Z > z), \end{aligned}$$

which implies $P(Z > z) = 0.01$; therefore $z = z_{0.01} = 2.326$.

L–example 4.22

- On a common test, the average grade was 74 and the standard deviation was 7. Suppose that the grades are given as integers.

A B C D

- If 12% of the class are given A's, and the grades are assumed to follow a normal distribution,
- what is the lowest possible A and the highest possible B?

$P(\underline{X} > \underline{x}) = \underline{0.12}$

Solution:

- We want to find x such that $P(X > x) = 0.12$.

$$P(X > x) = P\left(Z > \boxed{\frac{x-74}{7}}\right) = 0.12,$$

$z_{0.12}$

where $Z = \boxed{(X - 74)/7}$.

- On the other hand, using a statistical software, $P(Z > z) = 0.12$ implies $z = 1.175$.

- By setting $(x-74)/7 = 1.175$, we obtain

$$x = 74 + (1.175)7 = 82.225.$$

- Hence, the lowest possible A is 83 and the highest possible B is 82.

Compute the αth (upper) quantile of $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$:

(A) Online `R` compiler: `https://rdrr.io/snippets/`

- `qnorm(alpha, mu, sigma, lower.tail = False)` computes x_α;
- `qnorm(alpha, mu, sigma)` computes $x_{1-\alpha}$; $= P(Z > z_\alpha)$
- `qnorm(alpha, lower.tail = False)` computes z_α; $= \alpha$
- `qnorm(alpha)` computes $z_{1-\alpha}$. F

(B) use `R` Shiny app `Radiant`: `https://vnijs.shinyapps.io/radiant`.

L–example 4.23

- Let $X =$ the amount of sugar which a filling machine puts into "500g" packets.
- The actual amount of sugar filled varies from packets to packets.
- Suppose $X \sim N(\mu, 4^2)$.
- If only 2% of the packets contain less than 500g of sugar.
- What is the actual mean fill of these packets?

Solution: We need

$$0.02 = P(X < 500) = P\left(Z < \boxed{\frac{500-\mu}{4}}\right) = P\left(Z > \boxed{-\frac{500-\mu}{4}}\right),$$

where $Z = (X-\mu)/4$.

$= z_{0.02}$

On the other hand, from a statistical software, we have $P(Z > 2.0537) = 0.02$. Therefore

$$-\frac{500-\mu}{4} = 2.0537,$$

which leads to $\mu = 508.2$. That is, the mean fill should be 508.2g.

L–example 4.24 The width of a slot of a duralumin in forging is (in inches) normally distributed with $\mu = 0.9000$ and $\sigma = 0.0030$. The specification limits were given as 0.9000 ± 0.0050.

(a) What percentage of forgings will be defective?

(b) What is the maximum allowable value of σ that will permit no more than 1 in 100 defectives when the widths are normally distributed with $\mu = 0.9000$ and σ?

Solution:

(a) Let X be the width of our normally distributed slot. The probability that a forging is acceptable is given by

$$
\begin{aligned}
P(0.895 < X < 0.905) &= P\left(\frac{0.895-0.9}{0.003} < Z < \frac{0.905-0.9}{0.003}\right) \\
&= P(-1.67 < Z < 1.67) \quad \leftarrow P(Z<1.67) \\
&\qquad\qquad\qquad\qquad\qquad - P(Z<-1.67) \\
&= 2\Phi(1.67) - 1 = 0.905.
\end{aligned}
$$

So that the probability that a forging is defective is $1 - 0.905 = 0.095$. Thus 9.5 percent of forgings are defective.

(b) We need to find the value of σ such that

$$\boxed{P(0.895 < X < 0.905)} \geq \boxed{\frac{99}{100}}.$$

Now

$$\boxed{P(0.895 < X < 0.905)} = \ldots = \boxed{2P\left(Z < \frac{0.005}{\sigma}\right) - 1}.$$

$$P\left(-\frac{0.005}{\sigma} < Z < \frac{0.005}{\sigma}\right)$$

$$= P\left(Z < \frac{0.005}{\sigma}\right) - P\left(Z < -\frac{0.005}{\sigma}\right)$$

We thus have to solve for σ so that

$$2P\left(Z < \frac{0.005}{\sigma}\right) - 1 \geq 0.99.$$

or

$$P\left(Z < \frac{0.005}{\sigma}\right) \geq (1+0.99)/2 = 0.995.$$

$\iff P\left(Z > \boxed{\frac{0.005}{\sigma}}\right) \leq 0.005$

From a statistical software, we have $P(Z \geq \boxed{2.576}) = 0.005$ so we can use $\frac{0.005}{\sigma} \geq 2.576$ which gives $\sigma \leq 0.0019$.

- Recall that when $n \to \infty$, $p \to 0$, and np remains a constant, we can use **Poisson distribution to approximate the binomial distribution**.

  $n = 1000$ $p = 0.001$

- When $n \to \infty$, but p remains a constant (practically, p is not very close to 0 or 1), we can use **normal distribution to approximate the binomial distribution**.

  p 0.1 $n = 100$

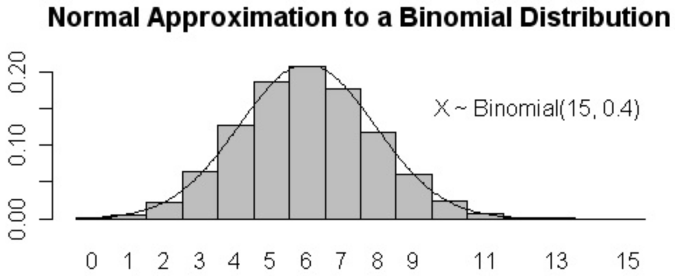
- A good rule of thumb is to use the normal approximation only when

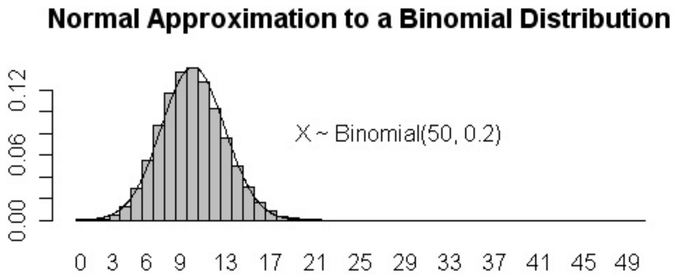
$$\overset{15}{np > 5} \quad \text{and} \quad \overset{15}{n(1-p) > 5}.$$

PROPOSITION 18 (NORMAL APPROX. TO BINOMIAL DISTRIBUTION) *Let $X \sim B(n, p)$; so that $E(X) = np$ and $V(X) = np(1-p)$. Then as $n \to \infty$,*

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - np}{\sqrt{np(1-p)}} \text{ is approximately } \sim N(0,1).$$

Normal Approximation to the Binomial Distribution





L–example 4.25

- If $X \sim B(15, 0.4)$, then

$$P(X = 4) = \binom{15}{4} 0.4^4 (0.6)^{11} = 0.1268.$$

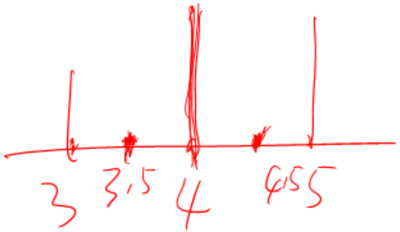
- By normal approximation, we may consider

$$Y \sim N(\mu, \sigma^2),$$

with $\mu = np = 6$ and $\sigma^2 = npq = 3.6.$

Hence,

$$\begin{aligned}
P(X=4) &= P(3.5 < X < 4.5) \approx P(3.5 < Y < 4.5) \\
&= P\left(\frac{3.5-6}{\sqrt{3.6}} < Z < \frac{4.5-6}{\sqrt{3.6}}\right) \\
&\approx P(-1.32 < Z < -0.79) \\
&= \Phi(-0.79) - \Phi(-1.32) \\
&= \boxed{0.1214.}
\end{aligned}$$



In this example, we have made the **continuity correction** to improve the approximation. In general, we have

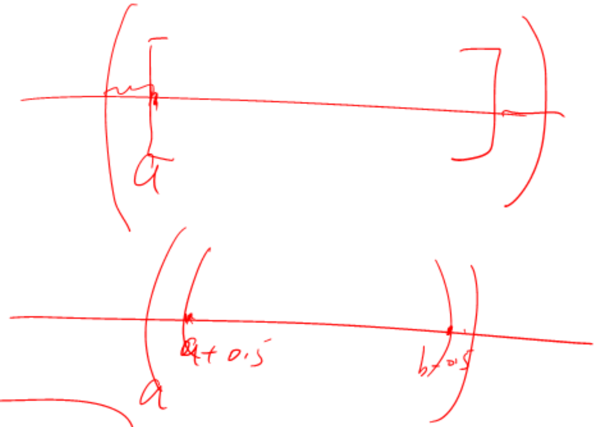
(a) $P(X = k) \approx P\,(k - {}^1\!/_2 < X < k + {}^1\!/_2)$;

(b) $P(a \leq X \leq b) \approx P\,(a - {}^1\!/_2 < X < b + {}^1\!/_2)$;

$P(a < X \leq b) \approx P\,(a + {}^1\!/_2 < X < b + {}^1\!/_2)$;

$P(a \leq X < b) \approx P\,(a - {}^1\!/_2 < X < b - {}^1\!/_2)$;

$P(a < X < b) \approx P\,(a + {}^1\!/_2 < X < b - {}^1\!/_2)$.



(c) $P(X \leq c) = P(0 \leq X \leq c) \approx P\,(-{}^1\!/_2 < X < c + {}^1\!/_2)$.

(d) $P(X > c) = P(c < X \leq n) \approx P\,(c + {}^1\!/_2 < X < n + {}^1\!/_2)$.

L–example 4.26

- A system is made up of 100 components, and each of which has a reliability equal to 0.90.
- These components function independently of one another, and the entire system functions only when at least 80 components function.
- What is the probability that the system functioning?

Solution:

- Let X = number of components functioning.
- Then $X \sim B(100, 0.9)$.

$X \approx N(90, 9)\ 3^2$

- Thus $E(X) = (100)(0.9) = \boxed{90}$ and $V(X) = (100)(0.9)(0.1) = \boxed{9}$.
- The system is functioning if $80 \le X \le 100$,

$$
\begin{aligned}
P(80 \le X \le 100) &\approx P\left(\frac{79.5-90}{3} < \frac{X-90}{3} < \frac{100.5-90}{3}\right) \\
&= P(-3.5 < Z < 3.5) \\
&= \Phi(3.5) - \Phi(-3.5) = 0.9995.
\end{aligned}
$$