Four

Special Probability Distributions

1 Discrete Distributions

Situations involving uncertainty and probability fall into certain broad classes, and we can use the same set of rules and principles for all situations within a class.

So it is beneficial for us to study whole classes of discrete random variables that arise frequently in applications.

Remark

Recall that for a discrete random variable X, the number of possible values in the range space R_X is either **finite** or **countable**.

Then the elements of R_X can be listed as $x_1, x_2, x_3, \ldots$.

Discrete Uniform Distribution

One of the simplest class of a discrete random variable is the discrete uniform distribution.

Definition 1 (Discrete Uniform Distribution)
*If a random variable X assumes the values $x_1, x_2, \ldots, x_k$ with equal probability, then X follows a **discrete uniform distribution**.*

The probability mass function for X is given by

$$f_X(x) = \begin{cases} \dfrac{1}{k}, & x = x_1, x_2, \ldots, x_k; \\ 0, & \textit{otherwise}. \end{cases}$$

Theorem 2

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

The expectation of X 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 of X 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, a 80-watt bulb, and a 100-watt bulb.

Each bulb has $1/4$ probability of being selected.

Let X be the wattage of the bulb being selected. Identify the distribution of X, and compute its mean and variance.

Solution:
X follows a uniform distribution and

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

Further, the probability mass function for X is given as

$$f_X(x) = \begin{cases} \dfrac{1}{4}, & x = 40, 60, 80, 100; \\ 0, & \text{otherwise.} \end{cases}$$

We can compute the expectation as

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

The variance is also found to be

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

Bernoulli Trial, Bernoulli Random Variable and Bernoulli Process

Numerous experiments have two possible outcomes.

If an item is selected from the assembly line and inspected, it is either defective or not defective. A piece of fruit is either damaged or not damaged.

Such experiments are called Bernoulli trials after the Swiss mathematician Jacob Bernoulli.

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 often code the two outcomes as "1" (success) and "0" (failure).

DEFINITION 4 (BERNOULLI RANDOM VARIABLE) *Let X be the 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, where $0 \leq p \leq 1$, the probability of success for a Bernoulli trial. Then X has the probability mass function

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

This probability mass function can also be written as

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

Remark

We denote a Bernoulli random variable by $X \sim \text{Bernoulli}(p)$, and write $q = 1 - p$.

Then the probability mass function becomes

$$f_X(1) = p, \quad f_X(0) = q.$$

THEOREM 5

For a Bernoulli random variable defined as above, we have

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

Parameters

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

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

For example, p is the parameter for the Bernoulli distribution.

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

Thus 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 "heads" is a success, and "tails" is a failure.

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

Polls
: Choosing a voter at random to ascertain if he 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 and

$$P(X = 1) = 4/10 = 0.4.$$

Furthermore, the probability mass function 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.*

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

Several distributions useful in applications are based on the Bernoulli trial and Bernoulli process. We will look at them in the next subsections:

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

Binomial Distribution

Suppose we have n independent and identically distributed Bernoulli trials. We can use 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?
- A 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 of 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 $X \sim \mathrm{Bin}(n, p)$.

The probability of getting exactly x successes is given as

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \ \ \textit{for } x = 0, 1, 2, \ldots, n.$$

It can be shown that $E(X) = np$, *and* $V(X) = np(1 - p)$.

Remark

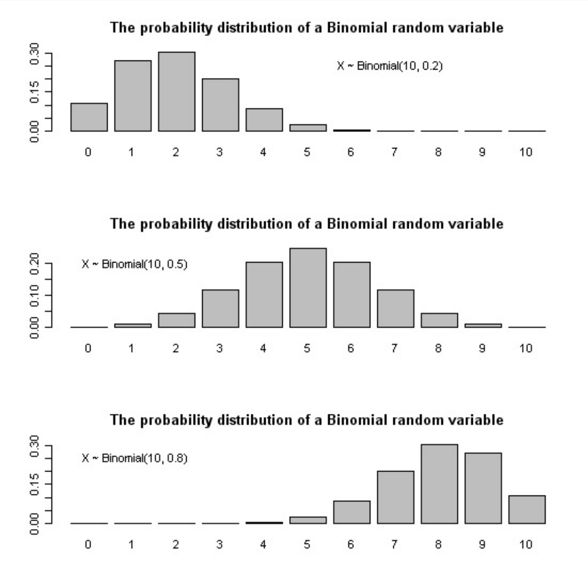
When $n = 1$, the probability mass function for the binomial random variable X is reduced to

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

This is the probability mass function for the Bernoulli distribution. Therefore the Bernoulli distribution is a special case of the binomial distribution.

BINOMIAL PROBABILITY MASS FUNCTION

The probability mass function for $\text{Bin}(10, 0.2)$, $\text{Bin}(10, 0.5)$, and $\text{Bin}(10, 0.8)$.



EXAMPLE 4.4

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

Solution:
Let X be the number of heads in 10 flips of the coin.

Each flip of the coin can be observed as a Bernoulli trial, with the probability of getting head (success) $p = 0.5$. Then X is the number success out of 10 Bernoulli trials; so $X \sim \mathrm{Bin}(10, 0.5)$.

We can compute

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

Negative Binomial Distribution

Consider a Bernoulli process, where the Bernoulli trials can be repeated as many times as desired or necessary.

Suppose we are interested in the number of trials needed so that k number of successes occur.

DEFINITION 8 (NEGATIVE BINOMIAL DISTRIBUTION)
Let X be the number of independent and identically distributed Bernoulli(p) *trials needed until the kth success occurs. Then X follows a* ***Negative Binomial distribution****, denoted by* $X \sim \mathrm{NB}(k, p)$.

The probability mass function of X is given by

$$f_X(x) = P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad \textit{for } x = k, k+1, k+2, \ldots$$

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

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 be the number of rolls needed to get the 6th number 6. Then $X \sim \mathrm{NB}(6, 1/6)$.

Using the probability mass function of the negative binomial distribution:

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

Geometric Distribution

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

Definition 9 (Geometric Distribution) *Let X be the number of independent and identically distributed* Bernoulli(p) *trials needed until the first success occurs. Then X follows a* ***Geometric distribution****, denote by $X \sim \text{Geom}(p)$.*

The probability mass function of X is given by

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

It can be shown that $E(X) = \dfrac{1}{p}$ and $V(X) = \dfrac{1-p}{p^2}$.

Poisson Distribution

A number of probability distributions come about through limiting arguments applied to other distributions. One useful distribution of this type is called the 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 $X \sim \mathrm{Poisson}(\lambda)$, where the parameter $\lambda > 0$ is the expected number of occurrences during the given period/region. Its probability mass function is given by

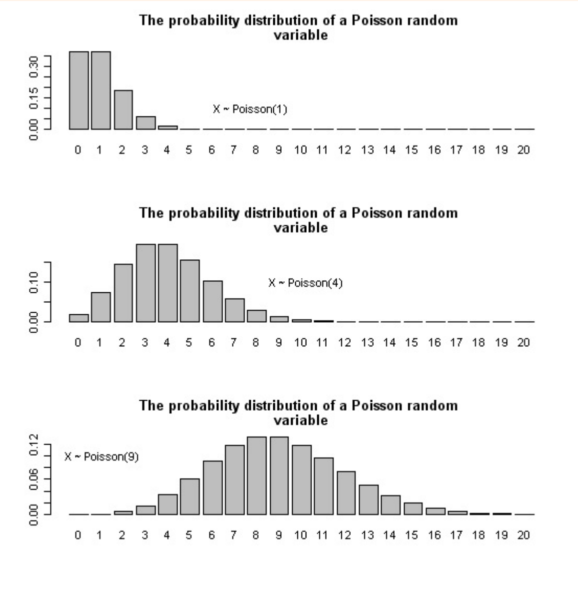
$$f_X(k) = P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!},$$

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

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

POISSON PROBABILITY MASS FUNCTION

The probability mass function for Poisson(1), Poisson(4), and Poisson(9).



EXAMPLE 4.6
The "fixed period of time" given in the definition can be a time period of any length: a minute, a day, a week, a month, etc. The "fixed region" can be of any size.

Here are some 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.

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*

- *the expected number of occurrences in an interval of length T is αT;*
- *there are no simultaneous occurrences;*
- *the number of occurrences in disjoint time intervals are independent.*

The number of occurrences in any interval T of a Poisson process follows a $\text{Poisson}(\alpha T)$ *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 be the number of robberies in one day. Then $X_1 \sim \text{Poisson}(4)$ from the given conditions.

Let X be the number of robberies in two days. Then

$$X \sim \text{Poisson}(2 \times 4) = \text{Poisson}(8).$$

We then have

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

Poisson Approximation to Binomial

The Poisson random variable has a tremendous range of applications in diverse areas because it may be used as an approximation for a binomial random variable under certain conditions.

The following result shows us how.

PROPOSITION 12 (POISSON APPROXIMATION TO BINOMIAL)
Let $X \sim \text{Bin}(n, p)$. *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

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

Remark

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 be the number of accidents among the 1000 cars.

Then $X \sim \text{Bin}(1000, 0.0001)$. If we compute using the binomial distribution,

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

We can also use the Poisson approximation.

We have $n = 1000$ and $p = 0.0001$. Hence $np = \lambda = 0.1$.

Thus

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

2 Continuous Distribution

There are many "natural" random variables whose set of possible values is uncountable. For example, consider

- the lifetime of an electrical appliance; or
- the amount of rainfall we get in a month.

How then, can we model such variables?

To achieve this aim, we shall now study some classes of continuous random variables.

Continuous Uniform Distribution

Intuitively, a uniform random variable on the interval (a, b) is a completely random number between a and b. We formalize the notion of "completely random" on an interval by specifying that the probability density function should be constant over the interval.

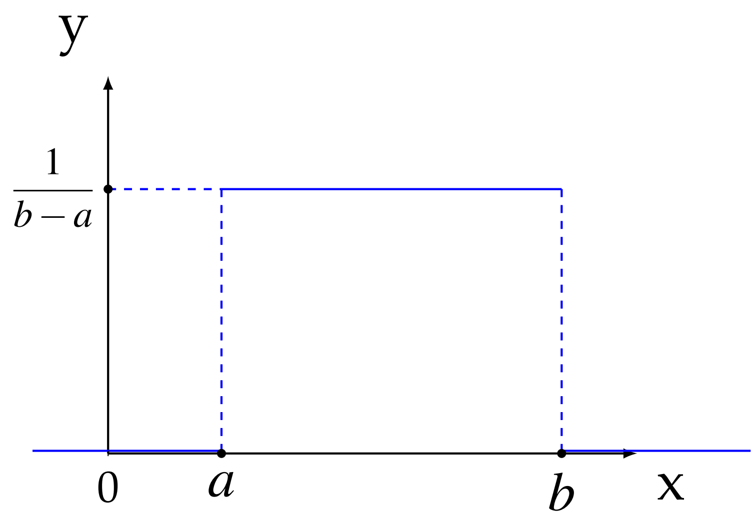
DEFINITION 13 (CONTINUOUS UNIFORM DISTRIBUTION)
A random variable X is said to follow a ***uniform distribution*** *over the interval (a,b) if its probability density function is given by*

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

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

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

The probability density function for the continuous uniform distribution can be drawn as below.



The cumulative distribution function 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 $\frac{3}{2}$?

Solution:
Let X be the position of the point. Then $X \sim \mathrm{U}(0,2)$, and we have

$$f_X(x) = \begin{cases} \frac{1}{2}, & 0 \leq x \leq 2; \\ 0, & \text{otherwise.} \end{cases}$$

Then the required probability is

$$P\left(1 \leq X \leq \tfrac{3}{2}\right) = \int_1^{3/2} \tfrac{1}{2}\,\mathrm{dx} = \tfrac{1}{2}\left[x\right]_1^{3/2} = \tfrac{1}{4}.$$

Exponential Distribution

The exponential distribution is the continuous counterpart to the geometric distribution. It is often used to model the waiting time to the first success in *continuous time*.

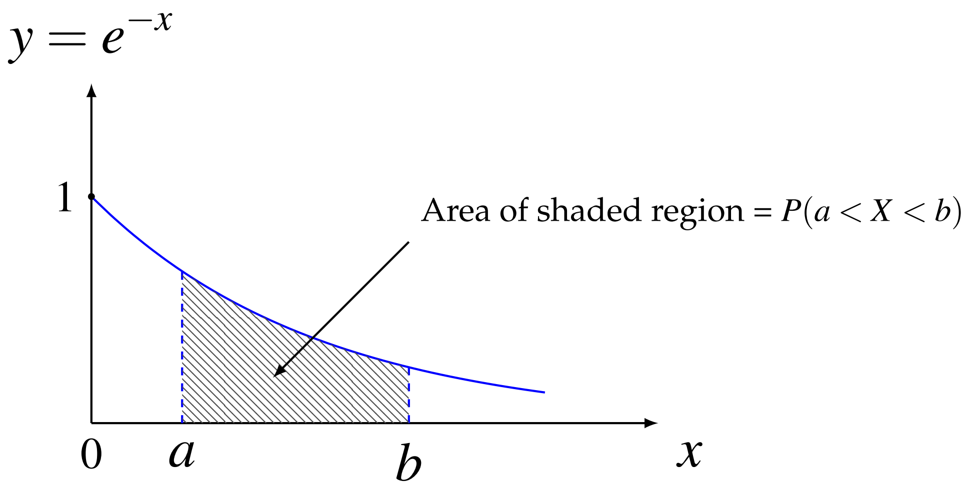
Definition 14 (Exponential Distribution)
A continuous random variable X is said to follow an ***exponential distribution*** *with parameter $\lambda > 0$ if its probability density function is given by*

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

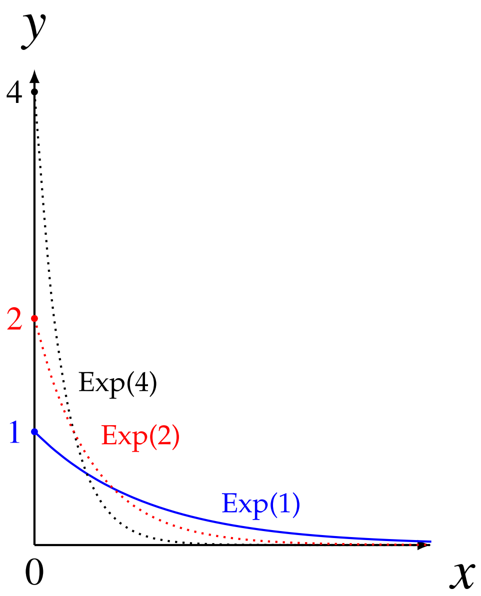
We denote $X \sim \mathrm{Exp}(\lambda)$.

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

The probability density function for Exp(1).



The probability density function for $\text{Exp}(\lambda)$, where $\lambda = 1, 2, 4$.



The cumulative distribution function of $X \sim \text{Exp}(\lambda)$ is given by

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

ALTERNATIVE FORM OF THE EXPONENTIAL

The probability density function of the exponential distribution can be written in the following alternative form

$$f_X(x) = \begin{cases} \frac{1}{\mu} e^{-x/\mu}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

The parameters μ and λ have the relationship $\mu = 1/\lambda$.

We will then have

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

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 then have $T \sim \text{Exp}(1/5)$, and so

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

Now let X be the number of systems out of 5 that are still functioning after 8 years. We see tha $X \sim \text{Bin}(5, 0.2)$. Hence,

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

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**" or is "**memoryless**".

To illustrate, suppose $X \sim \text{Exp}(\lambda)$ is the life length of a bulb. Then

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

This means that, if the bulb has lasted s time units (that is, $X > s$), the probability that it will last for another t units (that is, $X > s+t$), is the same as the probability that it will last for the first t units as a brand new bulb.

Normal Distribution

We next look at one of the most important class of continuous random variables.

Definition 16 (Normal Distribution)

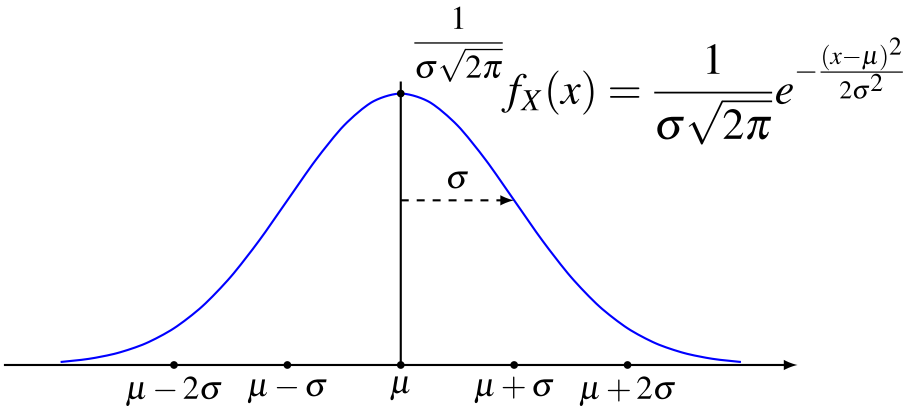
A random variable X is said to follow a ***normal distribution*** *with parameters μ and σ^2 if its probability density function is given by*

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

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

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

The probability density function of the normal distribution is positive over the whole real line, symmetrical about $x = \mu$, and bell-shaped.



Properties of the Normal Distribution

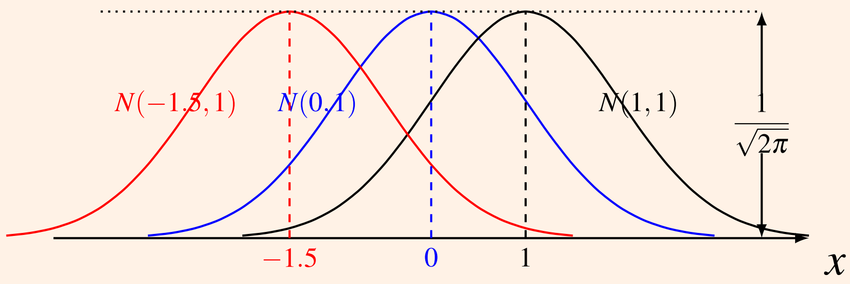
We go through some important properties of the normal distribution.

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

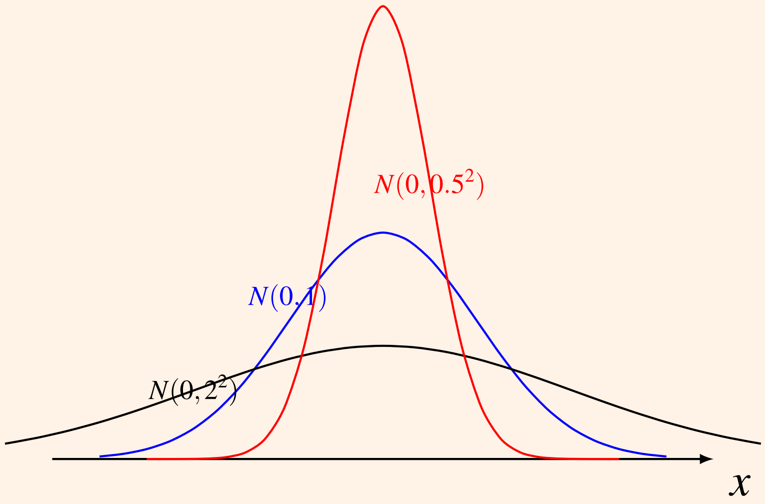
$$\int_{-\infty}^{\infty} f_X(x)\,\mathrm{dx} = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \mathrm{dx} = 1.$$

This validates that $f_X(\cdot)$ is a probability density function.

(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) Given $X \sim N(\mu, \sigma^2)$, let

$$Z = \frac{X - \mu}{\sigma}.$$

Then Z follows the $N(0,1)$ distribution, with $E(Z) = 0$ and $V(Z) = 1$.

We say that Z has a **standardized normal** or **standard normal** distribution, and the probability density function of Z is given by

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

Remark

- Calculating normal probabilities is a challenge because
  - there is no close formula for the integration,
  - and the computation relies on numerical integration.
- Fortunately we can use Property 4 from above.

  Suppose $X \sim N(\mu, \sigma^2)$ and we seek $P(x_1 < X < x_2)$. Consider

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

  Consider the transformation $Z = \frac{X-\mu}{\sigma}$, and let $z_1 = \frac{x_1-\mu}{\sigma}$ and $z_2 = \frac{x_2-\mu}{\sigma}$. 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 probability density function and cumulative distribution function of the standard normal. That is,

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

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

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

- Thus we can use the standard normal to calculate any normal probability.

  To do so, $\Phi(z)$ can be tabulated, or computed using 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, $\Phi(z) = P(Z \leq z) = P(Z \geq -z) = 1 - \Phi(-z)$;
  - If $Z \sim N(0,1)$, then $-Z \sim N(0,1)$;
  - If $Z \sim N(0,1)$, then $\sigma Z + \mu \sim N(\mu, \sigma^2)$.

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

Solution:
We have $\mu = 50$, $\sigma = 10$. Then

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

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

Definition 17 (Quantile)
*The **αth (upper) quantile**, where $0 \leq \alpha \leq 1$, of the random variable X is the number x_α that satisfies*

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

The Z Upper Quantile

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

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

Here are some common values of z_α:

$$z_{0.05} = 1.645, \quad z_{0.01} = 2.326.$$

Since $\phi(z)$, the probability density function of Z, is symmetrical about 0, then

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

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

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

This means that $P(Z > z) = 0.01$. Therefore $z = z_{0.01} = 2.326$.

Normal Approximation to Binomial

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

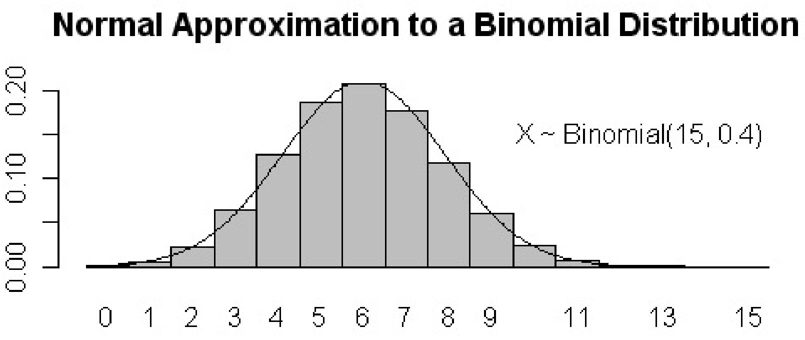
When $n \to \infty$, but p remains a constant (practically, p is not very close to 0 or 1), we can use the normal distribution to approximate the binomial distribution.

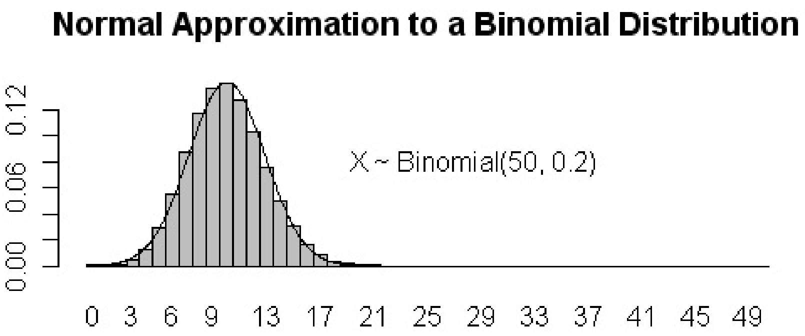
A good rule of thumb is to use the normal approximation when

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

Proposition 18 (Normal Approximation to Binomial)
Let $X \sim \mathrm{Bin}(n, p)$*, 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 a Binomial Distribution
X ~ Binomial(15, 0.4)
0.00
0.10
0.20
0 1 2 3 4 5 6 7 8 9 11 13 15


Normal Approximation to a Binomial Distribution
X ~ Binomial(50, 0.2)
0.00
0.06
0.12
0 3 6 9 13 17 21 25 29 33 37 41 45 49