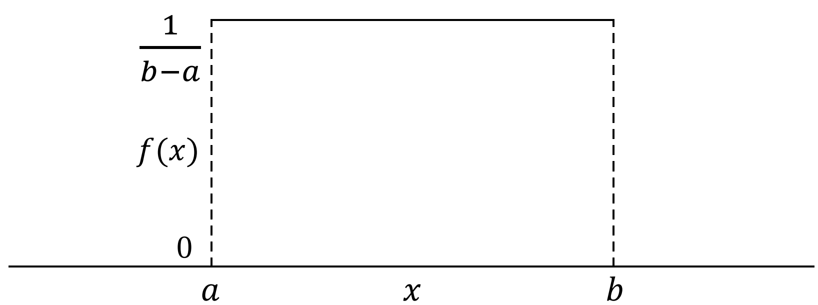
Continuous distributions

Uniform distribution

A continuous random variable X is said to have uniform distribution if it has the following pdf:

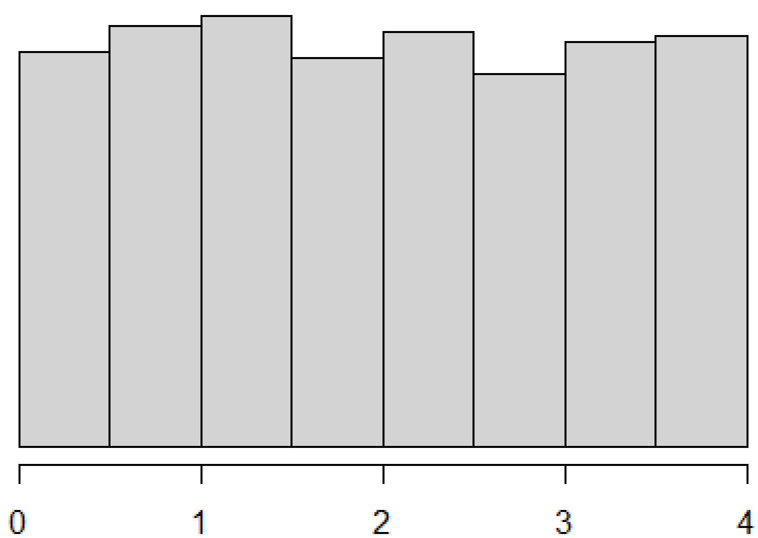
$$f(x) = \frac{1}{b-a} \;;\;\; a < x < b$$



We write $X \sim \text{uniform}(a, b)$.

Example

Let X be the amount (L) of milk you consume in a week. You collected data for 100 weeks and constructed the following histogram. Here, you can reasonably assume that the population follows uniform$(a = 0, b = 4)$ distribution.



For the above problem, $f(x) = \frac{1}{4-0} = \frac{1}{4}; \;\; 0 < x < 4$.

Mean and variance

$$E(X) = \frac{1}{2}(a + b)$$

$$V(X) = \frac{1}{12}(b - a)^2$$

Exercise

Your brother calls you at 9:00 P.M. every day. The duration of the call is uniformly distributed over 10 to 30 minutes. What is the probability that tomorrow's call duration will be (a) less than 20 minutes (b) more than 25 minutes (c) between 15 and 20 minutes?

Solution

X = duration of tomorrow's call

$X \sim \text{uniform}(a = 10, b = 30)$

$$f(x) = \frac{1}{30 - 10} = \frac{1}{20}; \;\; 10 < x < 30$$

(a)

$$P(X < 20) = \int_{10}^{20} \frac{1}{20}\, dx = 0.50$$

(b)

$$P(X > 25) = \int_{25}^{30} \frac{1}{20}\, dx = 0.25$$

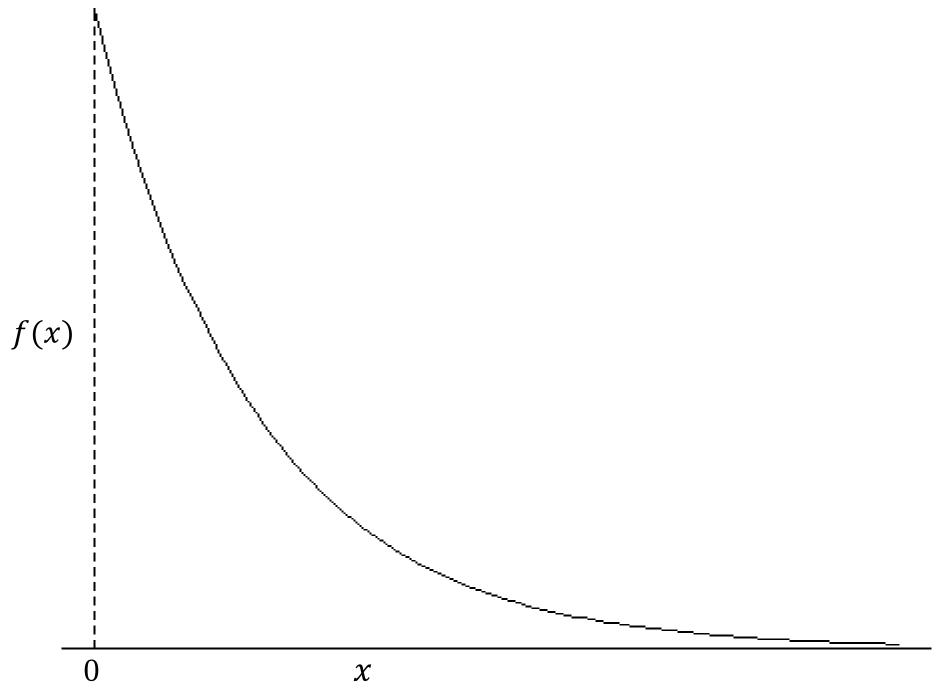
(c)

$$P(15 < X < 20) = \int_{15}^{20} \frac{1}{20}\, dx = 0.25$$

Exponential distribution

A continuous random variable X is said to have exponential distribution if it has the following pdf:

$$f(x) = \lambda\, e^{-\lambda x}\,; \;\; 0 < x < \infty$$



We write $X \sim \text{exponential}(\lambda)$.

Example

Lifetime of a machine, waiting time for an event, etc. sometimes follow exponential distribution.

Mean and variance

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

That is, mean and SD are equal.

Note

$$P(X > k) = \int_{k}^{\infty} \lambda\, e^{-\lambda x}\, dx = e^{-\lambda k}$$

Exercise

Survival time after a particular operation follows exponential distribution with an average survival time of 3 years. What is the probability that a patient will survive at least 2 years?

Solution

X = survival time of the patient

$X \sim$ exponential$(\lambda = 1/3)$

$$P(X > 2) = \int_{2}^{\infty} \frac{1}{3}\, e^{-\frac{1}{3}x}\, dx = e^{-\frac{1}{3}(2)} = 0.5134$$

Memoryless property

$$P(X > s + t \mid X > s) = P(X > t)$$

Proof

$$P(X > s + t \mid X > s)$$

$$= \frac{P(X > s + t)}{P(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$= e^{-\lambda t}$$

$$= P(X > t)$$

Explanation

Suppose, lifetime (X) of television of a particular brand follows exponential distribution. You bought a 2-year-old TV while your friend bought a new one. According to the memoryless property:

$P(X > 2 + 3 \mid X > 2) = P(X > 3)$

This means, the probability that your 2-year-old TV will survive at least 3 more years is equal to the probability that your friend's new TV will survive at least 3 years. That is, the system has 'forgotten' that your TV has already worked for 2 years.

Exercise

Survival time after a particular operation follows exponential distribution with an average survival time of 3 years. If a patient has already survived 1 year, what is the probability that (s)he will survive at least 2 more years?

Solution

X = survival time of the patient

$X \sim$ exponential$(\lambda = 1/3)$

$P(X > 1 + 2 \mid X > 1)$

$= P(X > 2)$ [memoryless property]

$$= \int_2^\infty \frac{1}{3}\, e^{-\frac{1}{3}x}\, dx = e^{-\frac{1}{3}(2)}$$

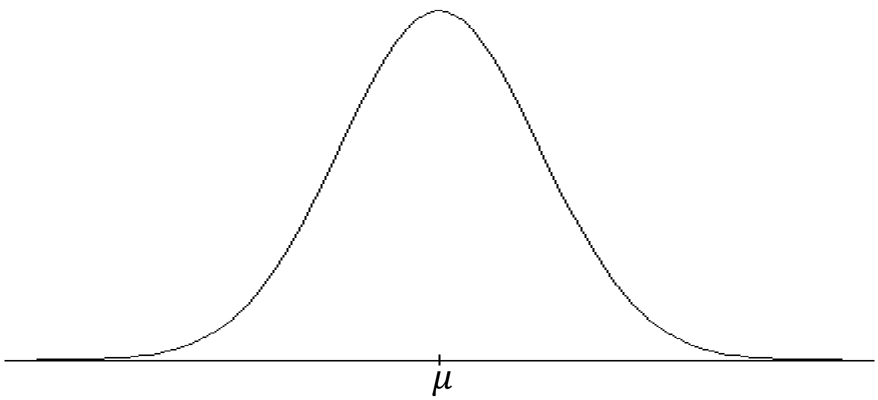
$= 0.5134$

Normal distribution

A continuous random variable X is said to have normal distribution if it has the following pdf:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty, -\infty < \mu < \infty, \ \ \sigma > 0$$

The distribution is symmetric around μ and bell-shaped (see the plot below).



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

- Note that ‘normal’ is the name of the distribution. The word ‘normal’ is not an adjective. There are no ‘abnormal’ distributions.

Example

Height, weight, measurement error, etc. sometimes follow normal distribution.

Mean and variance

$E(X) = \mu$

$V(X) = \sigma^2$

Properties of the normal distribution

1. The normal curve is symmetric around $X = \mu$ and bell-shaped.
2. Mean = median = mode = μ.

Theorem

Let $X \sim N(\mu, \sigma^2)$. Then, $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

Standard normal distribution

Let $X \sim N(\mu, \sigma^2)$. Then,

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

[By using the above theorem, $Z = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \sim N\left(\frac{1}{\sigma}\mu - \frac{\mu}{\sigma}, \left(\frac{1}{\sigma}\right)^2 \sigma^2\right)$. That is, $Z \sim N(0, 1)$.]

The variable Z is called the standard normal random variable. The symbol Z is used only when the variable follows $N(0, 1)$ distribution. The cdf of Z is denoted by $\Phi(z)$, that is, $\Phi(z) = P(Z \leq z)$.

Calculation of probability

We use "standard normal table" (also called Z-table) to calculate probabilities of normal distribution.

Example

Let X follow a normal distribution with mean 50 and standard deviation 10. What is the probability that the value of X is (a) smaller than 40 (b) larger than 40 (c) between 40 and 70?

Solution

Here, $\mu = 50$, $\sigma = 10$

(a)

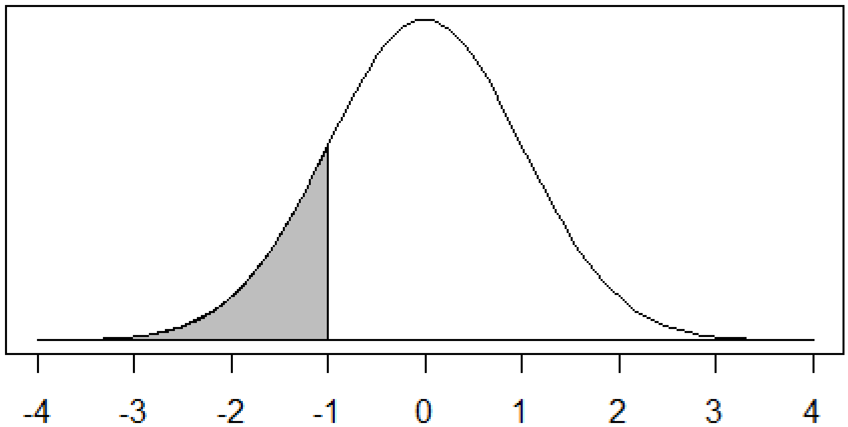
$P(X < 40)$

$= P\left(\frac{X - 50}{10} < \frac{40 - 50}{10}\right)$

$= P(Z < -1.00)$

$= \Phi(-1.00)$

$= 0.1587$ (from Z-table)

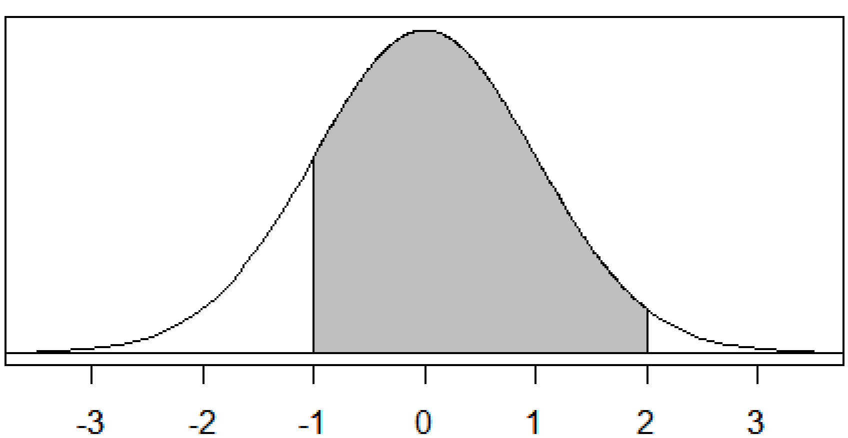


(b)

$P(X > 40)$
$= P(Z > -1.00)$
$= 1 - \Phi(-1.00)$
$= 1 - 0.1587$
$= 0.8413$

(c)

$P(40 < X < 70)$
$= P\left(\frac{40-50}{10} < \frac{X-50}{10} < \frac{70-50}{10}\right)$
$= P(-1 < Z < 2)$
$= \Phi(2) - \Phi(-1)$
$= 0.9772 - 0.1587$
$= 0.8185$



Theorem

Let $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, independent. Then,

$Y = X_1 + X_2 \sim N(\mu_1 + \mu_2,\ \ \sigma_1^2 + \sigma_2^2)$.

* The above theorem can be extended to three or more random variables.

Exercise

The weight (in pounds) of a particular type of fish is normally distributed with mean 10 and variance 1. What is the probability that the total weight of four fish will be more than 44?

Solution

Let $X_i =$ weight of the ith fish; $i = 1, 2, 3, 4$.

$X_i \sim N(10, 1)$.

$$Y = \sum_{i=1}^{4} X_i \sim N(10 + 10 + 10 + 10 = 40,\ \ 1 + 1 + 1 + 1 = 4)$$

(By using the above theorem)

$P(Y > 44)$

$= P\left(\frac{Y - 40}{2} > \frac{44 - 40}{2}\right)$

$= P(Z > 2.00)$

$= 1 - \Phi(2.00)$

$= \ 1 - 0.9772$ (from Z-table)

$= 0.0228$

Exercise

The price of an egg (in BDT) is normally distributed with mean 10 and variance 1. What is the probability that the total price of four eggs will be more than 44?

Solution

Let $X =$ price of an egg.

$X \sim N(10, 1)$.

$Y = 4X \sim N(4 \times 10 = 40,\ \ 4^2 \times 1 = 16)$.

[When $X \sim N(\mu, \sigma^2)$, then $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.]

$P(Y > 44)$

$= P\left(\frac{Y - 40}{4} > \frac{44 - 40}{4}\right)$

$= P(Z > 1.00)$

$= 1 - \Phi(1.00)$

$= 1 - 0.8413$ (from Z-table)

$= 0.1587$

(There is an easier way to solve this problem. Can you find it?)

Exercise

$P(Z < k) = 0.95$. Determine the value of k.

Solution

$\Phi(k) = 0.9500$. Therefore, $k = 1.65$ (from table).

Exercise

$P(-k < Z < k) = 0.95$. Determine the value of k.

Solution

$\Phi(k) - \Phi(-k) = 0.9500$

$\therefore \Phi(k) - (1 - \Phi(k)) = 0.9500$

$\therefore \Phi(k) = 0.9750$

Therefore, $k = 1.96$ (from table).

Gamma distribution

A continuous random variable X has gamma distribution if it has the following pdf:

$$f(x) = \frac{\lambda^k}{\Gamma(k)} e^{-\lambda x} x^{k-1}\ ;\ x > 0,\ \ \lambda > 0,\ \ k > 0.$$

Here, $\Gamma(k)$ is the gamma function. For all positive integers, $\Gamma(k) = (k-1)!$

We write $X \sim \text{gamma}(k, \lambda)$.

- When $k = 1$, $\text{gamma}(k, \lambda)$ reduces to $\text{exponential}(\lambda)$.
- When k is an integer, $\text{gamma}(k, \lambda)$ random variable is the sum of k independent $\text{exponential}(\lambda)$ random variables.

Example

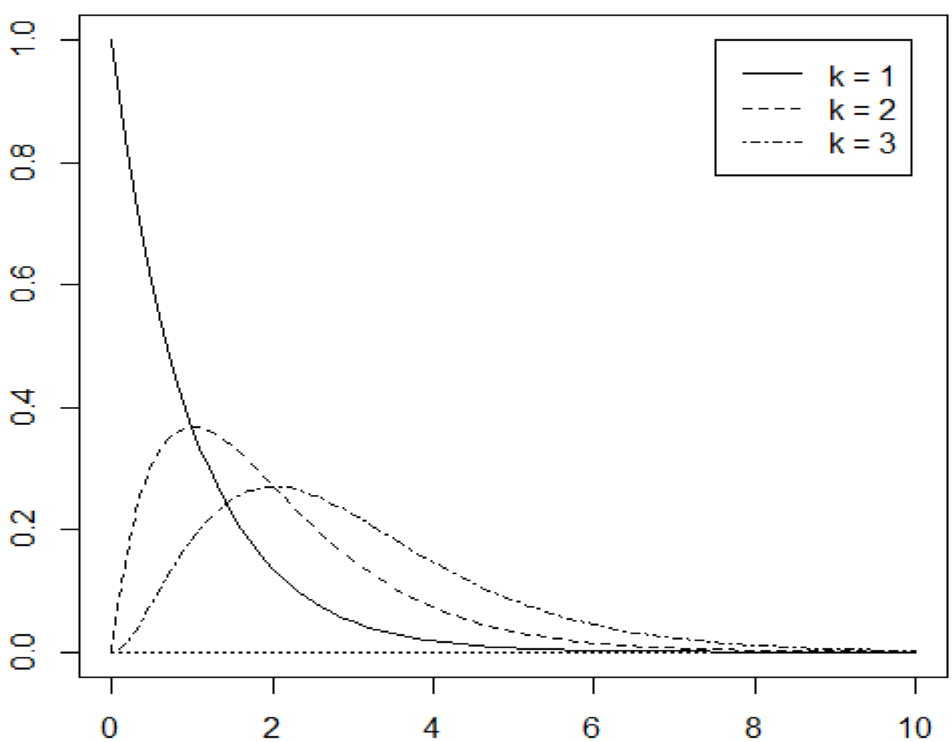
Lifetime of a machine, waiting time for an event, etc. sometimes follow gamma distribution.

Mean and variance

$$E(X) = \frac{k}{\lambda}$$

$$V(X) = \frac{k}{\lambda^2}$$

The figure below shows gamma distributions with $\lambda = 1$ and different values of k.



Exercise

Survival time of a patient after a particular operation follows exponential distribution with an average survival time of 3 years. Specify the distribution of the total survival time of 6 randomly chosen patients.

Solution

Let X_i = survival time of a patient after the operation; $i = 1, 2, \cdots, 6$.

$X_i \sim$ exponential$(\lambda = 1/3)$.

$$Y = \sum_{i=1}^{6} X_i \sim \text{gamma}\left(k = 6,\ \ \lambda = \frac{1}{3}\right).$$