Continuous uniform random variables

Example

It is known that a bus will arrive at random at the bus stop in between $8:00$ am to $8:30$ am. A man decides that he will go at random to the bus stop between the above period of time and will wait at most 5 minutes for the bus. Find the probability that he can take on the bus.

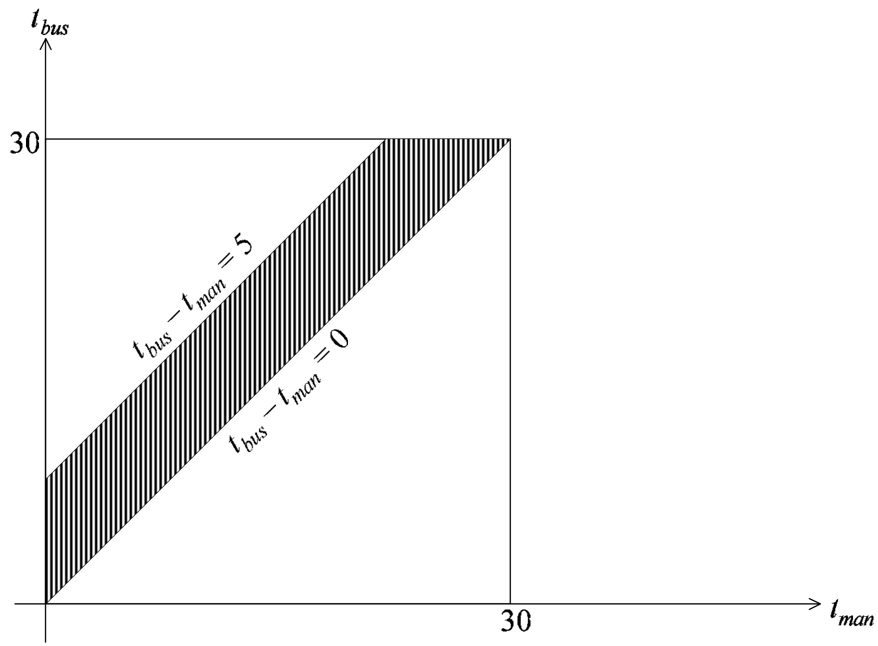
Solution The time durations (measured in minutes) of the arrival times of the bus and the man counting from 8 : 00 am are numbers randomly chosen from the interval $[0, 30]$. Let t_{bus} and t_{man} denote the random variables representing the arrival times of the bus and the man lapsed from 8 : 00 am. The man can catch the bus if

$$0 \leq t_{bus} - t_{man} \leq 5.$$

The probability that he can take on the bus

= ratio of the shaded area to the area of the rectangle

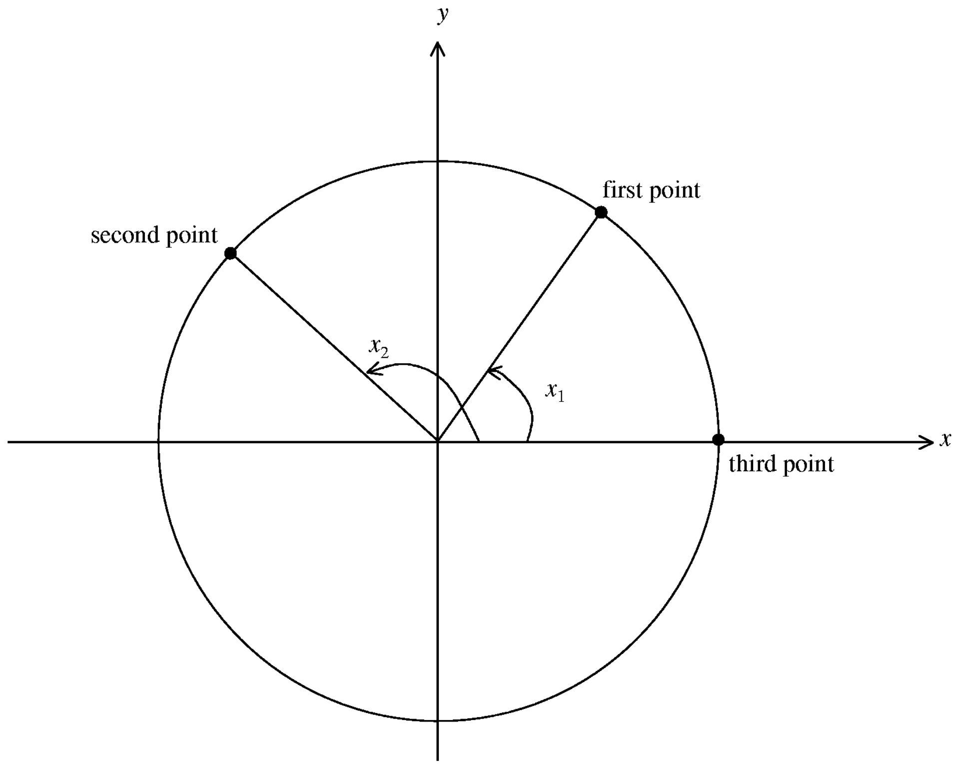
$$= 1 - \frac{\frac{25\times 25}{2} + \frac{30\times 30}{2}}{30 \times 30} = \frac{11}{72}.$$



Example
Find the probability that three points chosen at random on the circumference of a circle lie on a semi-circle.

Solution
Without loss of generality, we take the third point to lie on the x-axis with radian measure zero.

Let X_1 and X_2 denote the random variables representing the radian measure of the first point and second point; $-\pi < X_1 \leq \pi$ and $-\pi < X_2 \leq \pi$.



We consider four mutually exclusive events: both points on the upper semi-circle or lower semi-circle or they lie on different semi-circles. These events form a partition. They are also exhaustive. The three points lie on a semi-circle if

(i) $0 \le X_1 \le \pi$ and $0 \le X_2 \le \pi$

(both the first and second points lie on the upper semi-circle)

(ii) $-\pi \le X_1 \le 0$ and $-\pi \le X_2 \le 0$

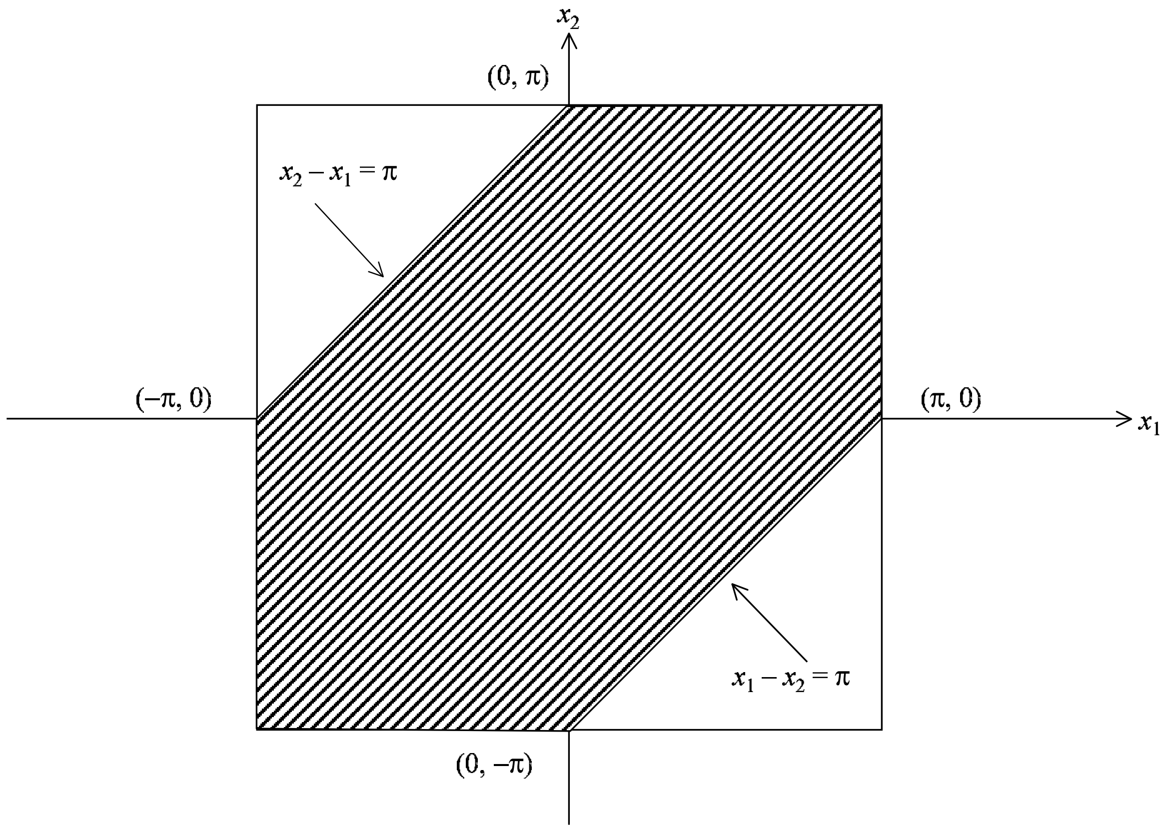
(both the first and second points lie on the lower semi-circle)

(iii) $0 \le X_1 - X_2 \le \pi, \quad 0 \le X_1 \le \pi$ and $-\pi \le X_2 \le 0$

(first point in the upper semi-circle, second point in the lower semi-circle)

(iv) $0 \le X_2 - X_1 \le \pi, \quad 0 \le X_2 \le \pi$ and $-\pi \le X_1 \le 0$.

(first point in the lower semi-circle, second point in the upper semi-circle)

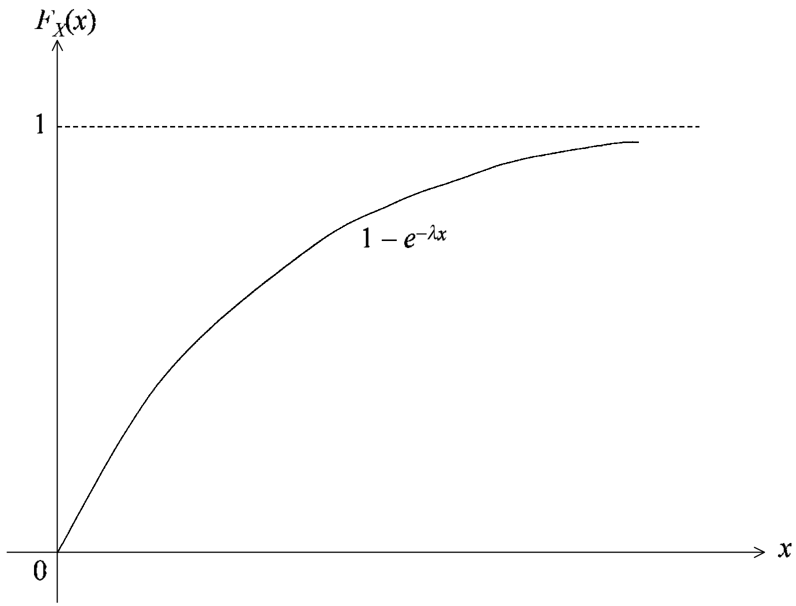


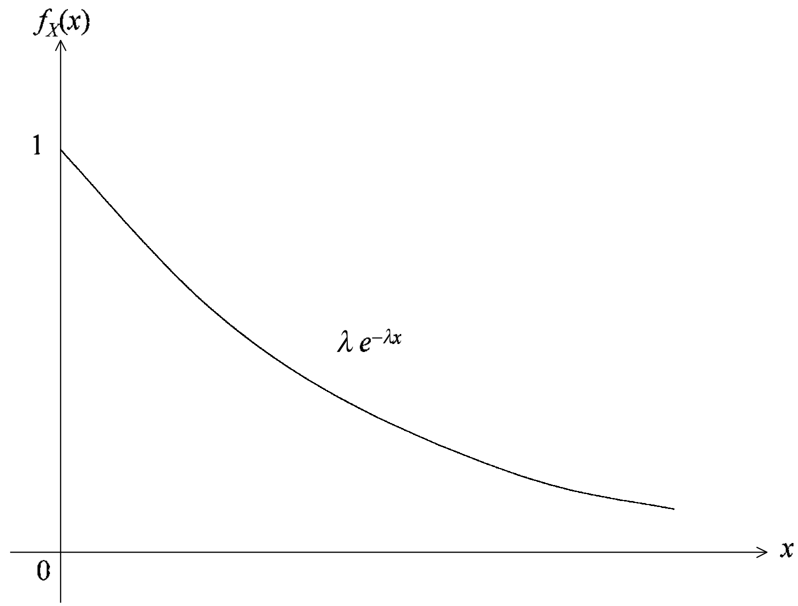
The above 4 conditions are satisfied by points inside the shaded area.

$$\text{Required probability} = \frac{\text{area of shaded region}}{\text{total area of the rectangular region}} = \frac{3}{4}.$$

Exponential random variable (with parameter λ)

cdf $\quad F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$ $\qquad$ *pdf* $\quad f_X(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0 \end{cases}$





λ is the rate at which events occur, that is, the average number of event occurrences over unit time.

Uses of the exponential distribution

Example

1. Arrivals of customers to a servicing station, including the waiting time between event occurrences.

2. Earthquakes or terror attack.

3. Life of a device (ageing or human error).

Query The device "forgets" how long it has been running, and its eventual breakdown is the result of some sudden arrival of failure event (not of gradual deterioration).

Occurrence of the first success in a Poisson process

$Y =$ Poisson variable that counts the number of occurrences over time t

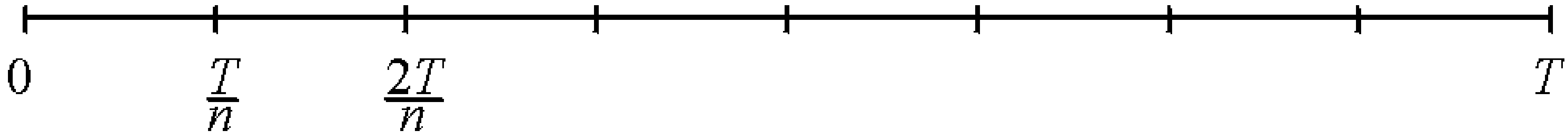
$\lambda =$ mean number of occurrences per unit time

$$\begin{aligned} & \text{Probability of at least one occurrence over time } t \\ = \;& P[Y \geq 1] = 1 - P[Y = 0] \\ = \;& 1 - e^{-\lambda t} = P[X \leq t] \end{aligned}$$

where X is the waiting time for the arrival of the first customer.

The exponential random variable is obtained as the limiting form of the goemetric random variables

An interval of duration T is divided into n subintervals, where n is infinitely large.



Each subinterval may be visualized as a Bernoulli trial if

(i) at most one event can occur in a subinterval (probability of more than one event occurrence is negligible);

(ii) outcomes in different subintervals are independent.

Let λ be the average number of events per unit time,

$\alpha =$ average number of events over T-period $= \lambda T$.

Let p be the probability of event occurrence in each subinterval,

where $p = \frac{\alpha}{n} = \frac{\lambda T}{n}$.

Let X be the random time for the first occurrence of event and M be the geometric random variable for the associated (independent) Bernoulli trials. The continuous random variable X and the discrete random variable M are related by

$$X = M\frac{T}{n}.$$

The probability that this time exceeds t seconds is

$$\begin{aligned} P[X > t] &= P\left[M > n\frac{t}{T}\right] = (1-p)^{nt/T} \\ &= \left[\left(1 - \frac{\alpha}{n}\right)^n\right]^{t/T} \longrightarrow e^{-\alpha t/T} \text{ as } n \to \infty. \end{aligned}$$

Since $\lambda = \alpha/T, P[X > t] \longrightarrow e^{-\lambda t}$ as $n \to \infty$.

Expected value of exponential random variable

$F_X(x) = (1 - e^{-\lambda x})u(x)$ and $f_X(x) = \lambda e^{-\lambda x}u(x)$.

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x)\ dx = \int_0^{\infty} \lambda x e^{-\lambda x}\ dx \\ &= \int_0^{\infty} y e^{-y} \frac{dy}{\lambda}, \qquad y = \lambda x, \\ &= \frac{1}{\lambda} \qquad\qquad \text{since } \int_0^{\infty} y e^{-y}\ dy = 1. \end{aligned}$$

Example

Suppose a device has an average life of 5 years, what is the probability that it lasts more than 7 years?

Hence, average life $= \dfrac{1}{\lambda} = 5$ so that $\lambda = \dfrac{1}{5}$.

$P[X > 7] = F_X(7) = e^{-7/5}$.

Memoryless property

Let X be the time for the arrival of the first customer. Given that there is no customer arriving for t units of time, the probability that no customer arrives for the next h units of time is given by

$$P[X > t + h | X > t].$$

The memoryless property means

$$P[X > t + h | X > t] = P[X > h].$$

Proof

$$\begin{aligned} P[X > t + h | X > t] &= \frac{P[\{X > t + h\} \cap \{X > t\}]}{P[X > t]}, \quad h > 0 \\ &= \frac{P[X > t + h]}{P[X > t]} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} = P[X > h]. \end{aligned}$$

The probability of the first arrival of an event is independent of how long you have been waiting.

Inter-event waiting time

For a Poisson random variable, the time between events is an exponentially distributed random variable with parameter λ.

Why?

From the memoryless property, the arrival of a new customer is independent of the number of earlier arrivals. Essentially, we can treat each new arrival as if it is the first customer. Hence, the inter-event waiting time can be modeled by the random variable representing the waiting time of the first customer.

Average inter-event waiting time $= \dfrac{1}{\lambda}$.

Example Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with an average value of $10,000$ miles. If a person desires to take a 5000-mile trip, what is the probability that he or she will be able to complete the trip without having to replace the car battery? What can be said when the distribution is not exponential?

Solution Let one unit of distance be $1,000$ miles. We have an average one battery failure over 10 units of distance. It follows by the *memoryless property* of the exponential distribution that the remaining lifetime (in thousands of miles) of the battery is exponential with parameter $\lambda = \frac{1}{10}$. Hence, the desired probability is

$$P[\text{remaining lifetime } > 5] = 1 - F(5) = e^{-5\lambda} = e^{-1/2} \approx 0.604.$$

Query Does it matter if different mileage amount is chosen for one unit of distance?

For example, take $10,000$ miles as one unit of distance. The new λ becomes 1 and we are asked to find $1 - F(0.5)$. We again obtain $e^{-0.5}$, the same answer.

Remark If the lifetime distribution F is not exponential, then the relevant probability is

$$P[\text{lifetime } > x + 5 | \text{lifetime } > x] = \frac{1 - F(x+5)}{1 - F(x)}$$

where x is the number of units of distance that the battery had been in use prior to the start of the trip. Therefore, if the distribution is not exponential, additional information is needed (namely, x) before the desired probability can be calculated.

Query With wearing effect in battery operation, is it reasonable to assume exponential distribution for the life time of a battery?

Example Consider a post office that is staffed by two clerks. Suppose that when Mr. Smith enters the system, he discovers that Ms. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Jones or Brown leaves. If the amount of time that a clerk spends with every customer is exponentially distributed with the same parameter λ, what is the probability that Mr. Smith is the last to leave the post office?

Solution Consider the time at which Mr. Smith first finds a free clerk. At this point either Ms Jones or Mr. Brown would have just left and the other one would still be in service. However, by the lack of memory of the exponential distribution, it follows that the additional amount of time that this other person (either Jones or Brown) would still have to spend in the post office is exponentially distributed with parameter λ. That is, it is the same as if service for this person were just starting at this point. Hence, by symmetry, the probability that the remaining person finishes before Smith must equal $\frac{1}{2}$.

Exponential density and memoryless property

The exponential density is the **only** continuous probability density that has the *memoryless property*.

Proof

We would like to find $F_X(t)$ such that

$$\frac{1 - F_X(t+h)}{1 - F_X(t)} = 1 - F_X(h).$$

Write $g(t) = 1 - F_X(t)$ for convenience, we find $g(t)$ which satisfies

$$\frac{g(t+h)}{g(t)} = g(h).$$

Note that $g(0) = \dfrac{g(t)}{g(t)} = 1$.

The only function which satisfies such property is $g(t) = e^{-\lambda t}$.

To deduce the result, we consider

$$\begin{aligned}\frac{g(t+h) - g(t)}{h} &= \frac{g(h)g(t) - g(t)}{h} \\ &= g(t)\left[\frac{g(h) - 1}{h}\right] = g(t)\left[\frac{g(h) - g(0)}{h}\right].\end{aligned}$$

Taking the limit $h \to 0$

$$g'(t) = g(t)g'(0).$$

Note that $g'(0)$ is just a constant. Recall that only the exponential function has the property that the derivative of the function equals a multiple of that function.

Since we must observe $0 \leq 1 - g(t) \leq 1$ and $g(0) = 1$, we then obtain

$$g(t) = e^{-\lambda t}, \quad \lambda > 0 \quad \text{for} \quad t \geq 0.$$

Gaussian (normal) random variable

Two-parameter density function with mean m and standard deviation σ

pdf

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

Note that $\int_{-\infty}^{\infty} f_X(x)\ dx = 1$. The proof relies on $\int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi}$.

$$\begin{aligned} \text{Mean } = E[X] &= \int_{-\infty}^{\infty} x f_X(x)\ dx \\ &= \int_{-\infty}^{\infty} (x-m) \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2}\ dx \\ &\quad + m \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2}\ dx = m. \end{aligned}$$

When $m = 0$ and $\sigma = 1$, we call it the standard Gaussian random variable Z, where

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

To compute the variance, we start from $\int_{-\infty}^{\infty} e^{-(x-m)^2/2\sigma^2}\, dx = \sqrt{2\pi}\sigma$

so that

$$\frac{d}{d\sigma}\int_{-\infty}^{\infty} e^{-(x-m)^2/2\sigma^2}\, dx = \frac{d}{d\sigma}\sqrt{2\pi}\sigma = \sqrt{2\pi}.$$

On the other hand, we differentiate under the integral sign and obtain

$$\frac{1}{\sigma^2}\int_{-\infty}^{\infty} \frac{(x-m)^2}{\sigma} e^{-(x-m)^2/2\sigma^2}\, dx = \sqrt{2\pi},$$

so that

$$\text{VAR}[X] = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} (x-m)^2 e^{-(x-m)^2/2\sigma^2}\, dx = \sigma^2.$$

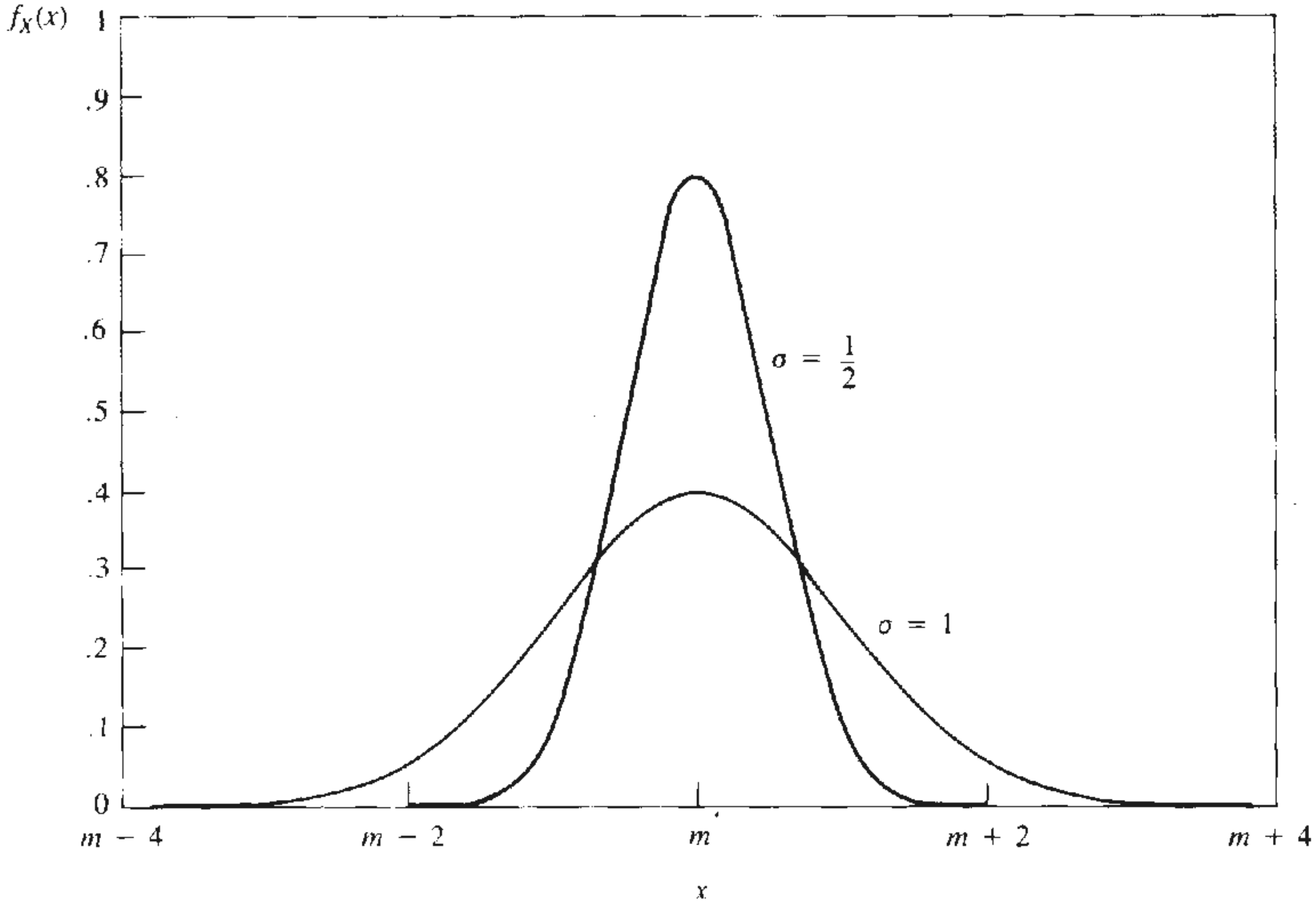
cdf $\quad F_X(x) = P[X \leq x] = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-(x'-m)^2/2\sigma^2}\, dx'.$

Define $Z = \dfrac{X-m}{\sigma}$, the Z becomes the standard normal variable.

Let $t = (x' - m)/\sigma$, so that

$$F_X(x) = \int_{-\infty}^{(x-m)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = N\left(\frac{x-m}{\sigma}\right) = F_Z(z),$$

where $z = \dfrac{x-m}{\sigma}$ and $N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt.$



Example The mean weight of 500 male students at a certain college is 151 lb and the standard deviation is 15 lb. Assuming that the weights are normally distributed, find how many students weigh (a) between 120 and 155 lb, (b) more than 185 lb.

Solution

In this problem, we have $m = 151$ and $\sigma = 15$. Let W denote the weight of a student chosen at random.

(a) Weights recorded as being between 120 and 155 lb can actually have any value from 119.5 to 155.5 lb, assuming they are recorded to the nearest pound. The required probability is given by $P[119.5 \leq W \leq 155.5]$. Define $Z = \dfrac{W - m}{\sigma}$ so that Z is the standard normal random variable with zero mean and unit standard deviation.

When $W = 119.5, Z = (119.5 - 151)/15 = -2.10$;

when $W = 155.5, Z = (155.5 - 151)/15 = 0.30$.

Statistics table gives values to the standard cumulative normal distribution function

$$N(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\ dz.$$

$$\begin{aligned}
\text{Required proportion of students} &= P[119.5 \le W \le 155.5] \\
&= P[-2.10 \le Z \le 0.30] \\
&= \int_{2.10}^{0.30} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\
&= \int_{-\infty}^{0.30} \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ dz \\
&\quad - \int_{-\infty}^{-2.10} \frac{1}{\sqrt{2\pi}}\ e^{-z^2/2}\ dz \\
&= 0.6179 - 0.0179 = 0.6000.
\end{aligned}$$

The number of students weighing between 120 and 155 lb is $500 \times 0.6 = 300$.

(b) Students weighing more than 185 lb must weigh at least 185.5 lb. We first compute $P[W \geq 185.5]$.

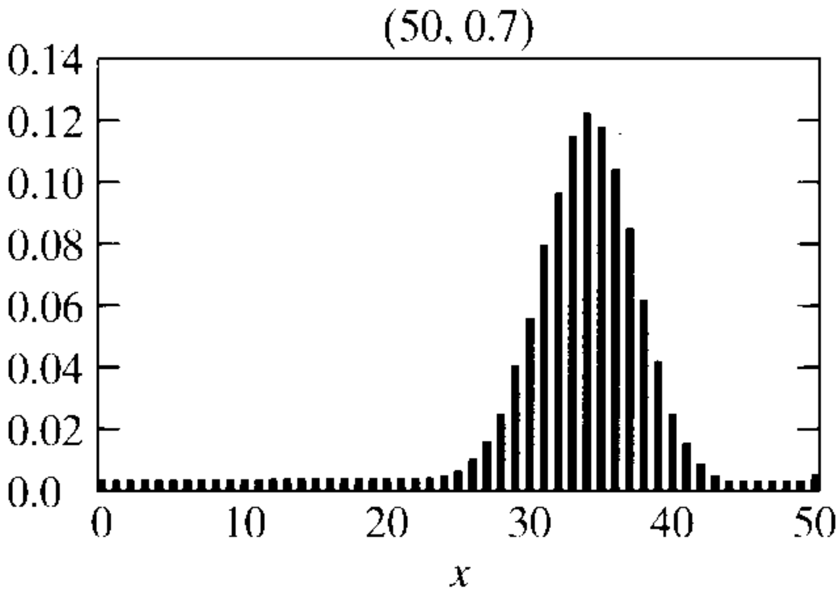
$$\text{When } W = 185.5, Z = (185.5 - 151)/15 = 2.30$$

$$\begin{aligned}\text{Required proportion of students} &= \int_{2.30}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}\, dz \\ &= 1 - \int_{-\infty}^{2.30} \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}\, dz \\ &= 1 - 0.9893 = 0.0107.\end{aligned}$$

Number of students weighing more than 185 lb is $500 \times 0.0107 = 5.35$.

Binomial distribution $\xrightarrow{np(1-p)\gg 1}$ normal distribution.

Approximation is well acceptable if $np(1-p) \geq 10$.



Binomial distribution with $n = 50$ and $p = 0.7$.

The probability histogram is computed by

$$P_N(k) = {}_{50}C_k (0.7)^k (0.3)^{50-k}.$$

The distribution looks almost like "normal".

Example

Find the probability of getting between 3 and 6 heads inclusive in 10 tosses of a fair coin by using (a) the binomial distribution, (b) the normal approximation to the binomial distribution.

Solution

Let X be the random variable giving the number of heads in 10 tosses.

$$
\begin{aligned}
P[X=3] &= {}_{10}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \frac{15}{128}, P[X=4] = \frac{105}{512}, \\
P[X=5] &= \frac{63}{256}, P[X=6] = {}_{10}C_6 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4 = \frac{105}{512} = P[X=4] \\
P[3 \leq X \leq 6] &= P[X=3] + P[X=4] + P[X=5] + P[X=6] \\
&= \frac{99}{128} = 0.7734.
\end{aligned}
$$

Normal approximation

Treating the data as continuous, 3 to 6 heads can be considered as 2.5 to 6.5 heads. The mean and variance of X are $np = 5$ and $\sigma = \sqrt{npq} = \sqrt{10\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 1.58$, respectively. Define $Z = \dfrac{X-5}{1.58}, Z$ is the standard normal variable.

When $X = 2.5, Z = \dfrac{2.5-5}{1.58} = -1.58$

$$X = 6.5, Z = \frac{6.5-5}{1.58} = 0.95.$$

$$\begin{aligned} P[2.5 \leq X \leq 6.5] &= P[-1.58 \leq Z \leq 0.95] = N(0.95) - N(-1.58) \\ &= 0.8289 - 0.0571 = 0.7718. \end{aligned}$$