Probability and Statistics I

16. The Normal Distribution

Suppose that X is a continuous random variable with pdf $f(x)$ and cdf $F(x)$.

a) TRUE or FALSE: $f(x) \leq 1$ for all $x \in \mathbb{R}$

b) TRUE or FALSE: $f(x)$ is continuous

c) TRUE or FALSE: $E(X^2) \geq E(X)^2$

d) Which of the following is not equal to all of the others?

   A) $P(3 \leq X \leq 5)$ B) $P(3 < X \leq 5)$

   C) $P(3 < X < 5)$ D) $P(3 < X < 4) + P(4 < X < 5)$

4.3 The Normal Distribution

The Normal Distribution

We say that a random variable X has a normal distribution with mean μ and variance σ^2 if the pdf of X is

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

for all $x \in \Re$.

Mathematically, we write $X \sim \text{Normal}(\mu, \sigma^2)$.

Properties

CDF: No closed form

Mean: $E(Z) = \mu$

Variance: $V(Z) = \sigma^2$

Calculator

`https://stattrek.com/online-calculator/normal`

Example 16.1

The adult heights of people assigned to be male and female at birth can be modelled amazingly well by a normal distribution. Suppose that the adult height people assigned to be female at birth is normally distributed with mean 64 inches and standard deviation 3 inches.

$$X \sim \text{Normal}(64, 9).$$

a) What is the density of X? Sketch the density.

b) What is the probability that someone assigned to be female at birth will be:

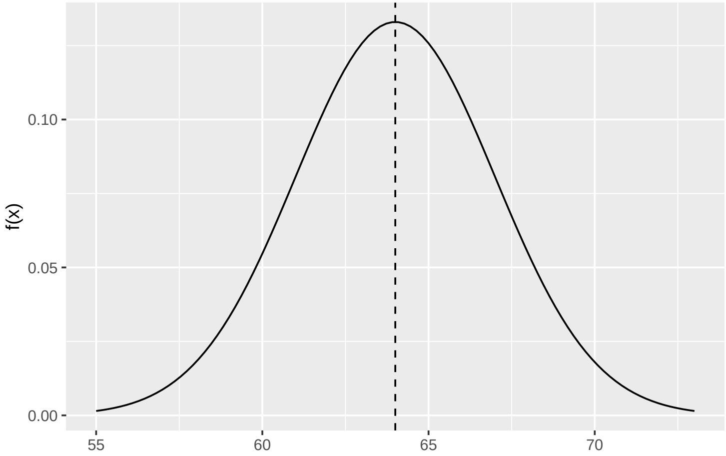
   i) less than 5 feet tall?

   ii) greater than 6 feet tall?

   iii) between 5 and 6 feet tall?

c) Find values l and u such that $P(l < X < u) \approx .95$.

Example 16.1 ctd



Example 16.1 ctd

The adult heights of people assigned to be male and female at birth can be modelled amazingly well by a normal distribution. Suppose that the adult height people assigned to be female at birth is normally distributed with mean 64 inches and standard deviation 3 inches:

$$X \sim \text{Normal}(64, 9).$$

a) What is the density of X? Sketch the density.

b) What is the probability that someone assigned to be female at birth will be:

   i) less than 5 feet tall?

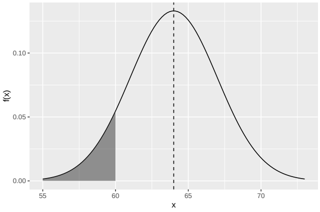
   ii) greater than 6 feet tall?

   iii) between 5 and 6 feet tall?

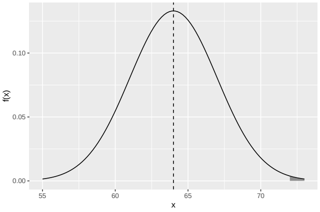
c) Find values l and u such that $P(l < X < u) \approx .95$.

Example 16.1 ctd

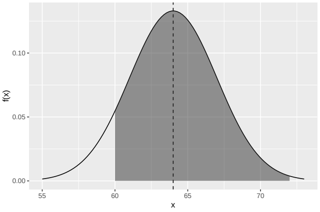
$P(X < 60)$



$P(X > 72)$



$P(60 < X < 72)$



The Standard Normal Distribution

We say that a random variable Z has a **standard** normal distribution if

$$Z \sim \text{Normal}(0, 1).$$

The pdf of the standard normal is

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

for all $x \in \Re$.

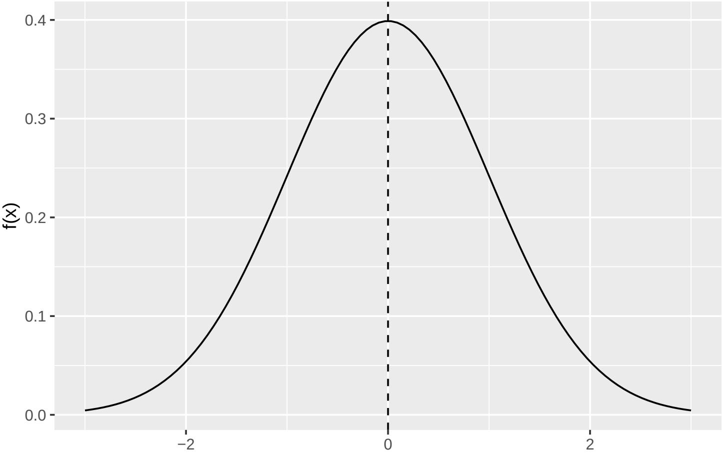
Properties

CDF: No closed form

Mean: $E(Z) = 0$

Variance: $V(Z) = 1$

Example 16.1 ctd



Standardization

If $X \sim \text{Normal}(\mu, \sigma^2)$ then

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

has a standard normal distribution.

Standardization

If $X \sim \text{Normal}(\mu, \sigma^2)$ then

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

has a standard normal distribution.

We can use this fact to compute probabilities for any normal random variable from the probabilities of a standard normal random variable:

$$P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

where $Z \sim \text{Normal}(0, 1)$.

Example 16.1 ctd

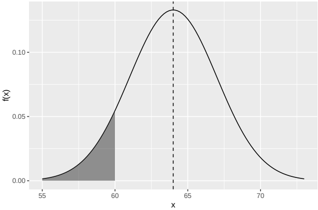
The adult heights of people assigned to be male and female at birth can be modelled amazingly well by a normal distribution. Suppose that the adult height people assigned to be female at birth is normally distributed with mean 64 inches and standard deviation 3 inches:

$$X \sim \text{Normal}(64, 9).$$

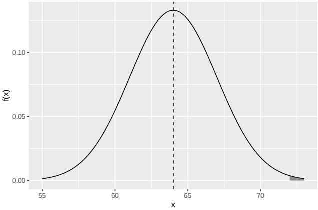
d) Repeat part b) using standardization.

Example 16.1 ctd

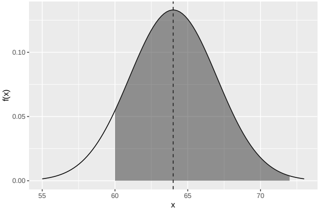
$P(X < 60)$



$P(X > 72)$

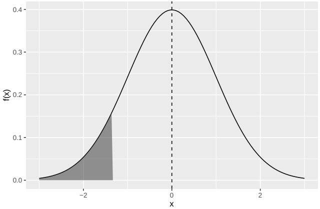


$P(60 < X < 72)$

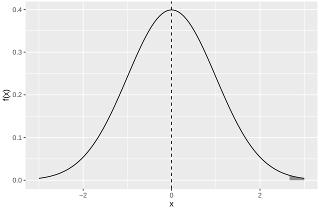


Example 16.1 ctd

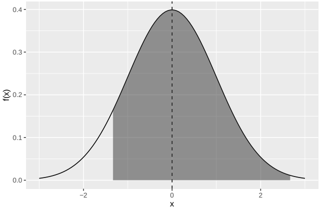
$$P\left(Z < \frac{(60-64)}{3}\right) = P(Z < -4/3)$$



$$P\left(Z > \frac{(72-64)}{3}\right) = P(Z > 8/3)$$



$$P\left(\frac{(60-64)}{3} < Z < \frac{(72-64)}{3}\right) = P(-4/3 < Z < 8/3)$$



Standardization

If we measure in terms of standard deviations from the mean then probabilities for all normal distributions are the same.

If

$$X \sim \text{Normal}(\mu, \sigma^2)$$

and Z is standard normal then

$$P(X \leq \mu + c\sigma) = P\left(Z \leq \frac{(\mu + c\sigma) - \mu}{\sigma}\right) = P(Z \leq c)$$

for all possible values of μ and σ^2.

Standardization – Percentiles

We will denote the $100(1-\alpha)$-th percentile of the standard normal by z_α:

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

If $X \sim \text{Normal}(\mu, \sigma^2)$ then

$$P(X > \mu + z_\alpha \sigma) = P\left(Z > \frac{\mu + z_\alpha \sigma - \mu}{\sigma}\right) = \alpha.$$

The Empirical Rule

If $X \sim \text{Normal}(\mu, \sigma^2)$ then

$$P(\mu - \sigma < X < \mu + \sigma) \approx .68$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx .95$$
$$P(\mu - 3\sigma < X < 3\mu + \sigma) \approx .997$$

Example 16.1 ctd

The adult heights of people assigned to be male and female at birth can be modelled amazingly well by a normal distribution. Suppose that the adult height people assigned to be female at birth is normally distributed with mean 64 inches and standard deviation 3 inches:

$$X \sim \text{Normal}(64, 9).$$

e) Repeat part c) using standardization.

Normal Approximation to the Binomial

Suppose that $X \sim \text{Binomial}(n, p)$ with $np \geq 10$ and $n(1-p) \geq 10$. Then

$$P(X \leq x) \approx P\left(Z \leq \frac{x + .5 - np}{\sqrt{np(1-p)}}\right)$$

where $Z \sim \text{Normal}(0, 1)$.

Example 16.2

A standard roulette wheel has 37 pockets in which the ball may land. Of these, 18 pockets are red, 18 are black, and 1 is green.

Suppose that you place \$1 bets that the ball will land in a black pocket on **200** consecutive games. Let X be the number of times you win.

a) What is the exact probability that you win between 95 and 105 games inclusive?

b) Approximate this probability with the normal distribution?

Questions?

Exercise 16.1

Suppose that the amount of time a cell phone battery lasts normally distributed with mean 28 hours and standard deviation 4 hours depending on the use.

a) Sketch the probability density function.
b) Shade the area defining the probability that the battery lasts for more than 34 hours.
c) What is the probability that the battery lasts for more than 34 hours? Compute the value using the calculator and by standardizing.
d) Compute the probability that the battery lasts for more than 36 hours without using the calculator or standardizing.