Application of Normal Distribution

Huining Kang

HuKang@salud.unm.edu

August 10, 2020

Content

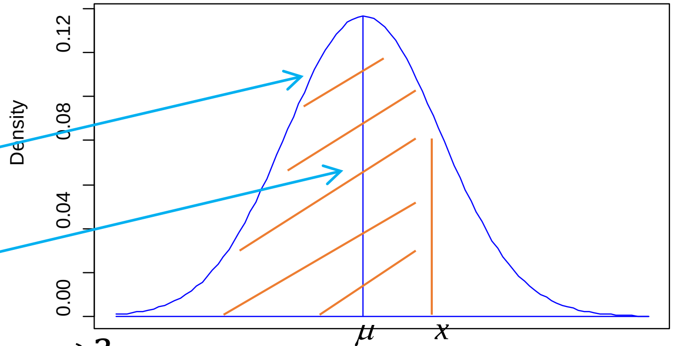
- Application of normal distribution
  - Review of normal distribution
  - Relation between a normal distribution and standard normal distribution
    - Transformation of a random variable
    - Z-score transformation
    - Calculation of probabilities of normal distribution through the Z-score transformation
  - Examples of normal distribution applications
  - Inverse cumulative standard normal distribution

Objective

- Understand the relationship between normal distribution and standard normal distribution and know how to use standard normal distribution to calculate the probabilities of a general normal distribution
- Know how to apply normal distribution to solve real world problems

Review of Normal distribution

- $X \sim N(\mu, \sigma^2)$
  - pdf $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
  - cdf $F(x) = P(X \leq x)$

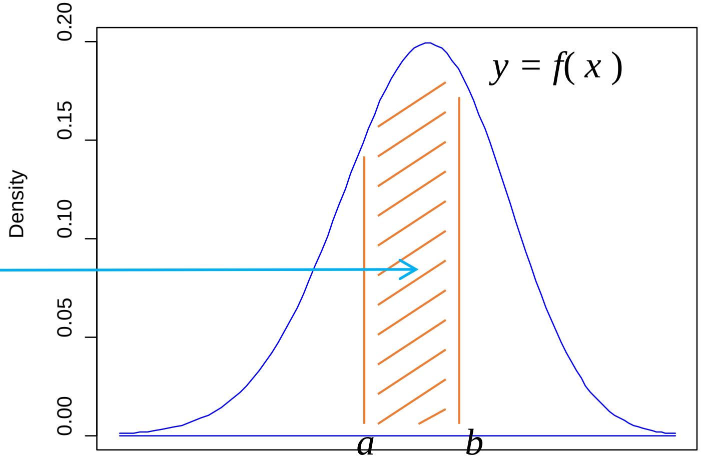


$$= \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

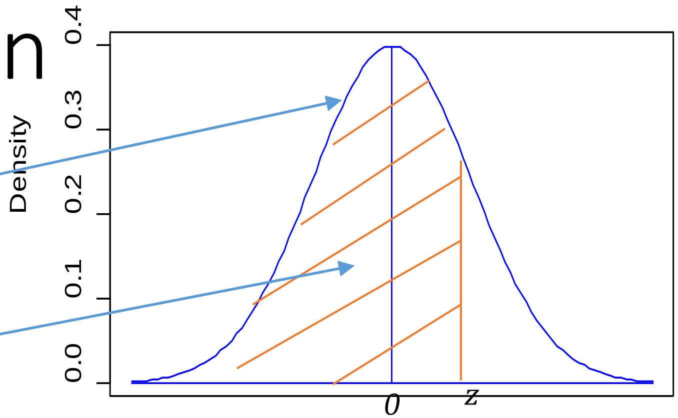
  - Characteristics
    - Mean: $\mu = \int_{-\infty}^{\infty} x f(x)dx = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$
    - Variance: $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$
  - Probability
    - $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$

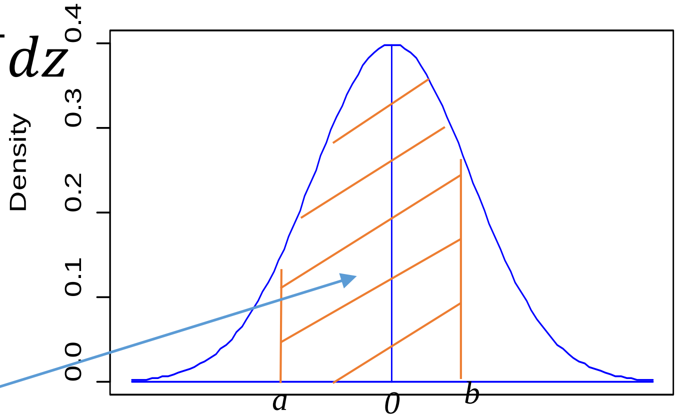
$$= \int_{a}^{b} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= F(b) - F(a)$$



Review of Normal distribution





- $Z \sim N(0,\ 1)$
  - pdf $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$
  - cdf $\Phi(z) = P(Z \leq z)$

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

  - Characteristics
    - Mean: $\mu = 0$
    - Variance: $\sigma^2 = 1$
  - Probability
    - $P(a \leq Z \leq b) = \int_a^b \phi(z)dz = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \Phi(b) - \Phi(a)$
  - Calculation using Stata
    - Calculate $\Phi(z)$: normal(z)
    - Calculate $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$: normalden(z)

Normal distribution & Standard Normal Distribution

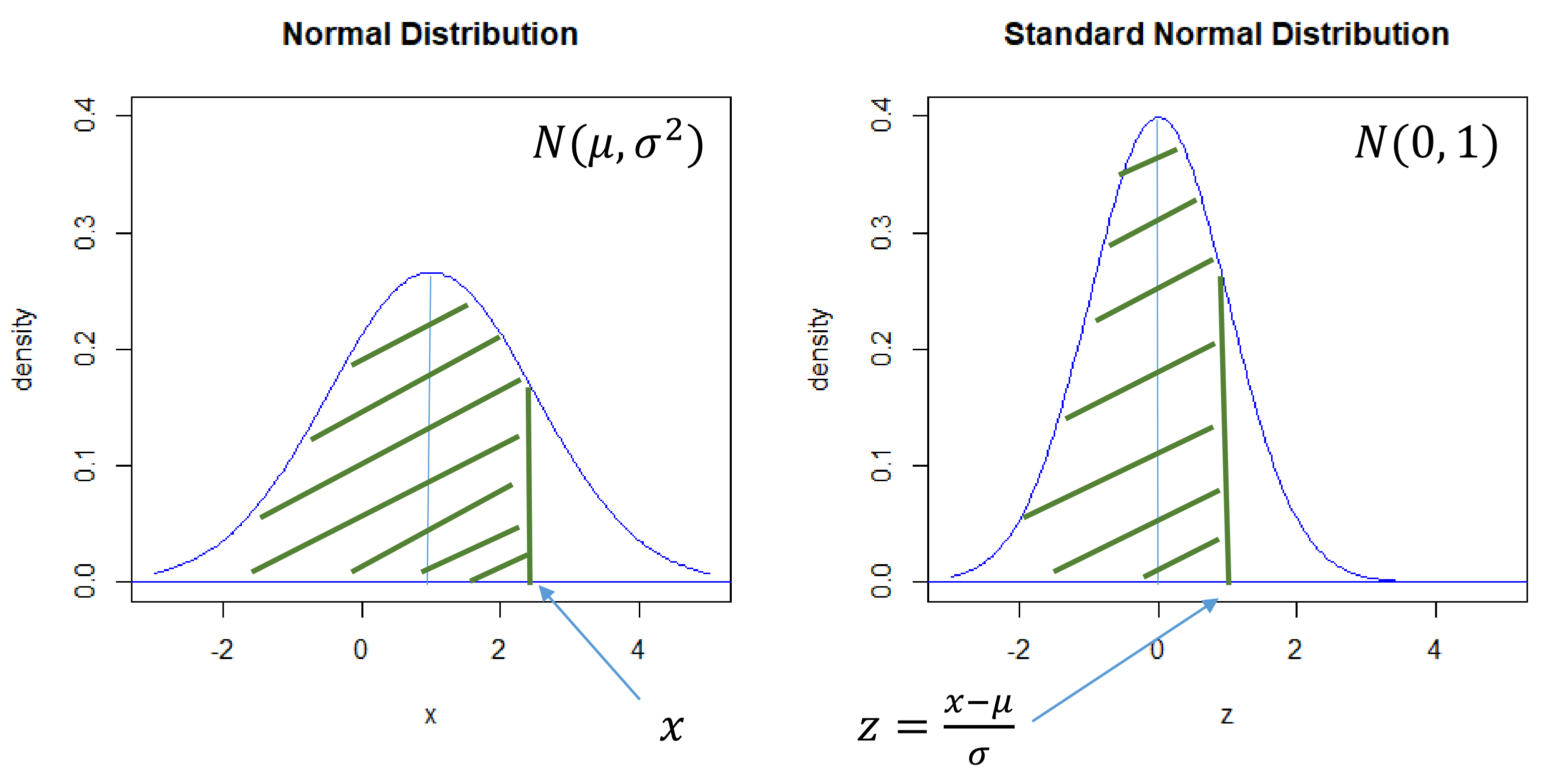
- Transformation of a random variable
  - Suppose X is a r.v. with mean μ and variance σ^2 (or standard deviation σ
    - $X - \mu$ is called centered (at μ).
    - X/σ is called scaled (by σ)
    - $Z = \frac{X-\mu}{\sigma}$ is called Z-score transformation.
    - Z is a new r.v. with mean 0 and variance 1.
  - Suppose $X \sim N(\mu, \sigma^2)$.
    - Then $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$
    - $F(x) = P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi(\frac{x-\mu}{\sigma})$
    - $P(a \leq X \leq b) = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi(\frac{a-\mu}{\sigma})$
  - Conclusion
    - Calculation of probabilities of r.v. X can be done using the standard normal distribution function $\Phi(z)$.

Normal distribution & Standard Normal Distribution

- Graphical interpretation
  - Suppose $X \sim N(\mu, \sigma^2)$.
    - Then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

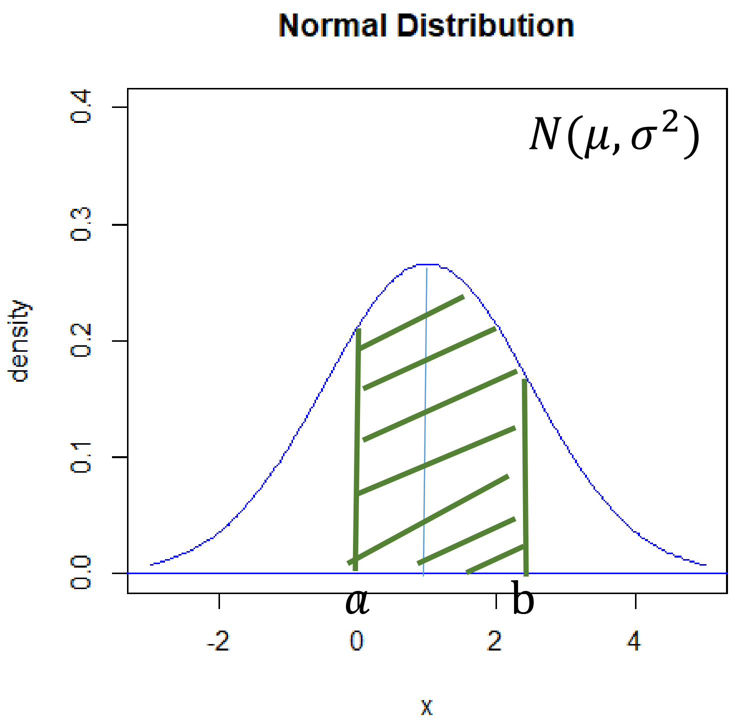
Z

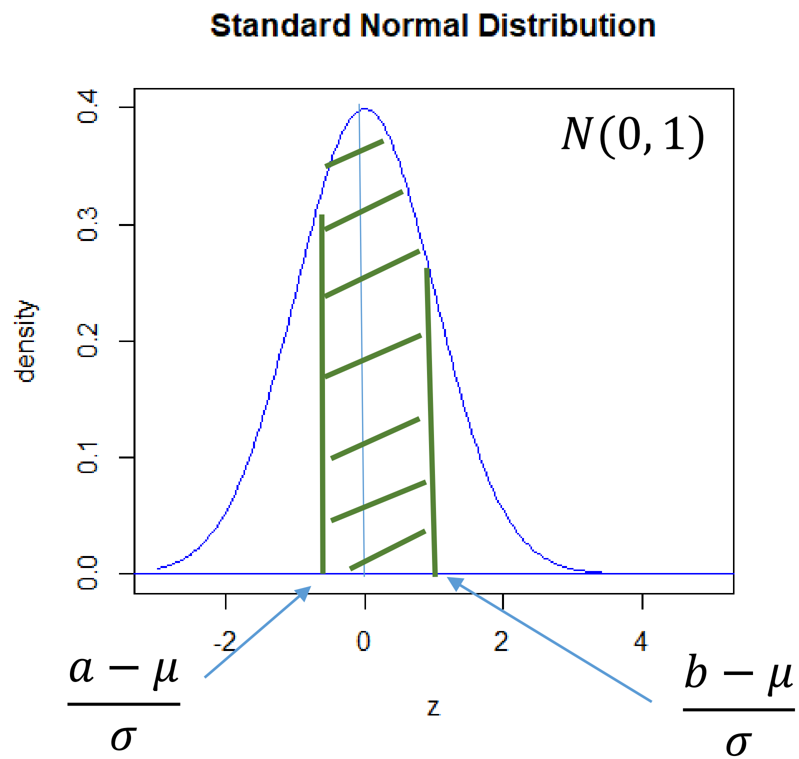
    - $F(x) = P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$



Normal distribution & Standard Normal Distribution

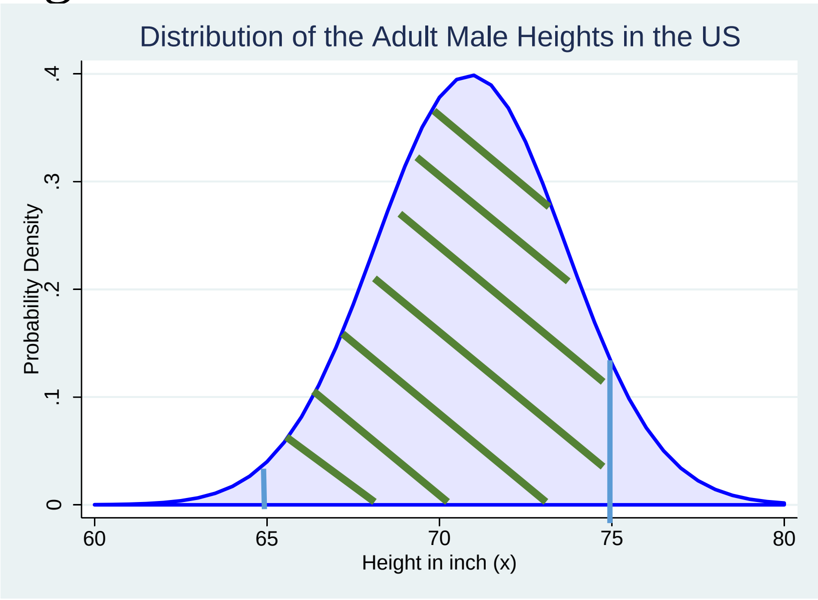
- Graphical interpretation
  - Suppose $X \sim N(\mu, \sigma^2)$.
    - Then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$
    - $P(a \le X \le b) = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$





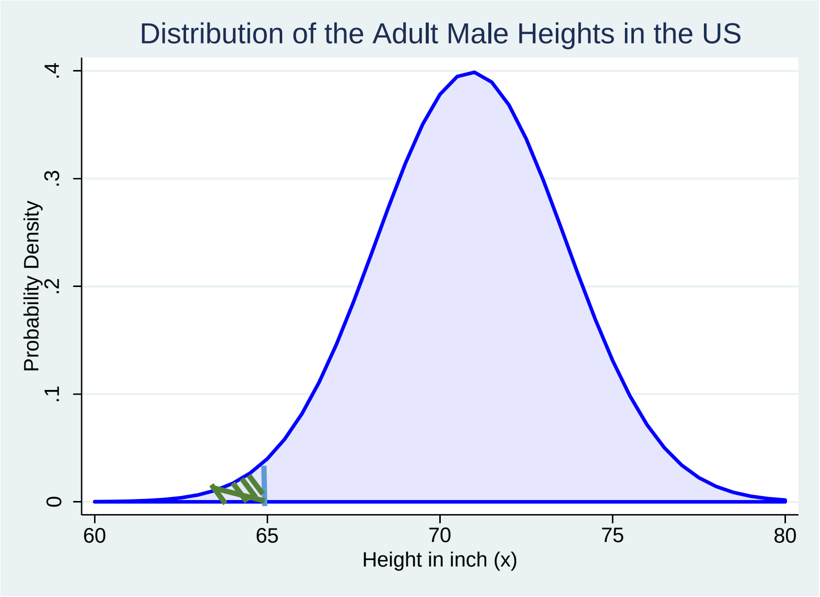
Application of Normal Distribution

- Importance of Normal Distribution
  - A great number of random variables in the real word problems can be approximately modeled using normal distribution.
- Example 1
  - The height (X) of a randomly selected adult male from the US follows a normal distribution with mean $\mu = 70.9$ inches and standard deviation $\sigma = 2.75$.
  - What is the probability that the height of a randomly selected adult male is between 65 in. and 75 in.



- $P(65 \le X \le 75) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) = \Phi\left(\frac{75-70.9}{2.75}\right) - \Phi\left(\frac{65-70.9}{2.75}\right) = .916$
- Stata command:
  - disp normal((75-70.9)/2.75) - normal((65-70.9)/2.75)
  - .91604906
- Interpretation
  - 95% of adult males in the US has a height between 65 inches and 75 inches.

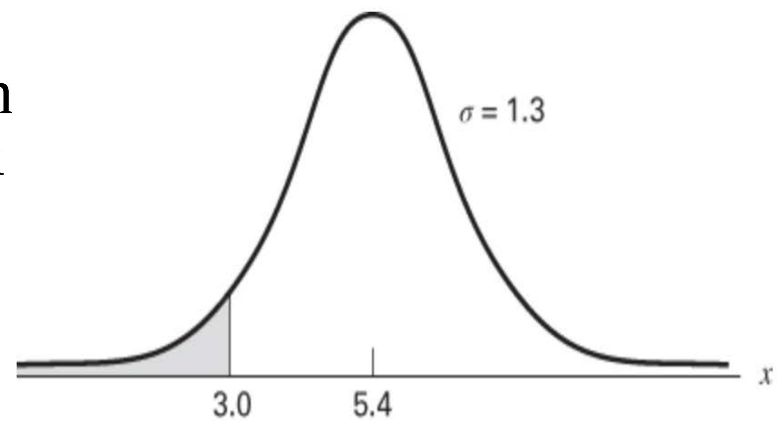
Application of Normal Distribution



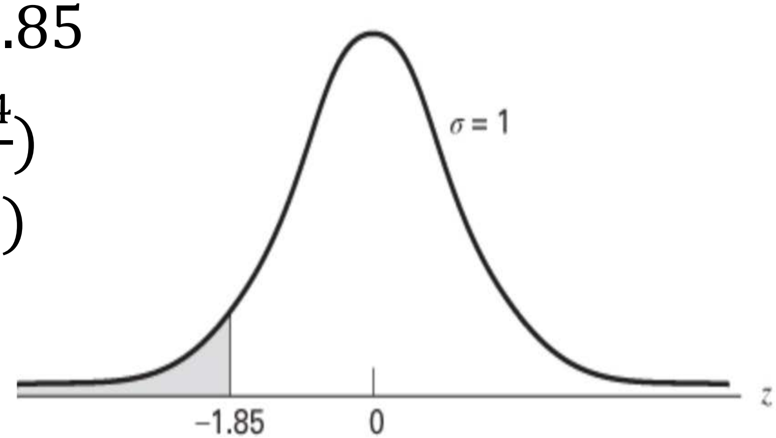
- Example 1 (Cont.)
  - What is the probability that a randomly selected adult male is shorter than 65 in.?
    - $P(X \leq 65) = \Phi\left(\frac{65-70.9}{2.75}\right) < 0.0001$
    - Stata Command
    - disp normal((65-70.9)/2.75)
    - .00003691
  - $P(X \leq 0) = \Phi\left(\frac{0-70.9}{2.75}\right) = 10^{-147}$
    - Stata Command
    - disp normal((0-70.9)/2.75)
    - 7.09e-147
  - Comment
    - Theoretically speaking X assume negative values with a non-zero probability. But the probability is negligible.

Application of Normal Distribution

- Example 4.7.1 (page 123)
  - Children ages 8 – 15 years
  - $X =$ amount of time children spend in the upright position
  - $X \sim N(5.4, 1.3^2)$
  - $P(X \leq 3.0) = ?$



- Solution
  - $x = 3 \;\rightarrow\; z = \frac{3.0-5.4}{1.3} = -1.85$
  - $P(X \leq 3.0) = P(Z \leq \frac{3.0-5.4}{1.3})$
  $= P(Z \leq -1.85) = \Phi(-1.85)$
  $= .0324$

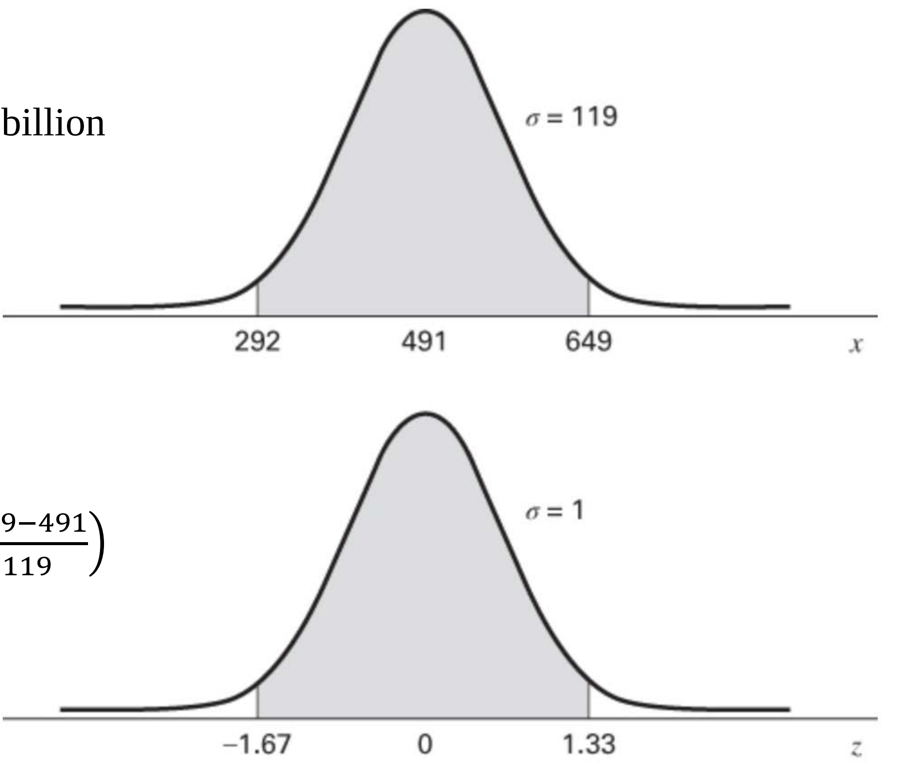


- Stata command
  - disp normal((3-5.4)/1.3)
  - .03243494

Application of Normal Distribution

- Example 4.7.2 (page 125)
    - A study of common breath metabolites
    - A 27-year-old female
    - 30-day observation
    - $X =$ the ammonia concentration in parts per billion (ppb)
    - $X \sim N(491, 119^2)$
    - $P(292 \leq X \leq 649) = ?$
- Solution
    - $x = 292 \;\rightarrow\; z = \frac{292-491}{119} = -1.67$
    - $x = 649 \;\rightarrow\; z = \frac{649-491}{119} = 1.33$
    - $P(292 \leq X \leq 649) = P\left(\frac{292-491}{119} \leq Z \leq \frac{649-491}{119}\right)$

$$= \Phi\left(\frac{649-491}{119}\right) - \Phi\left(\frac{292-491}{119}\right)$$
$$= \Phi(1.33) - \Phi(-1.67) = .8606$$



- Stata command
    - disp normal((649-491)/119) - normal((292-491)/119)
    - .86063087

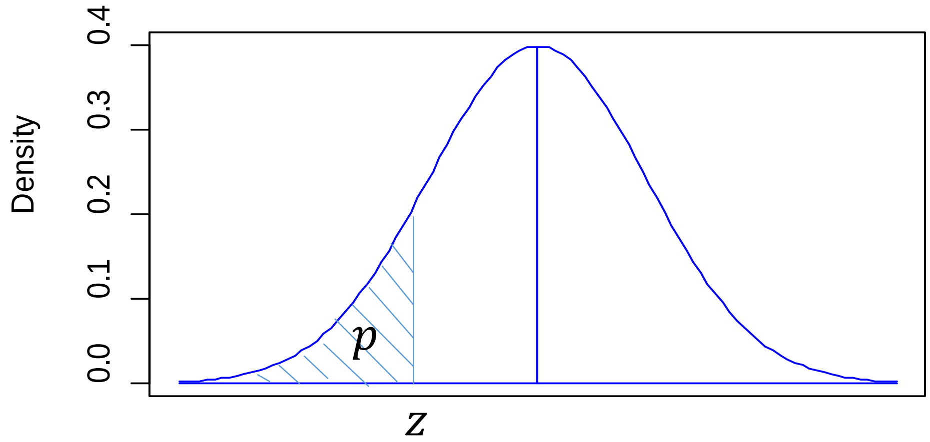
Application of Normal Distribution

- Example 4.7.3 (page 123)
  - A population of 10,000 children ages 8 – 15 years
  - $X =$ amount of time children spend in the upright position
  - $X \sim N(5.4, 1.3^2)$
  - How many children are expected to be upright more than 8.5?
- Solution
  - First find $P(X > 8.5)$
    - $P(X > 8.5) = 1 - P(X \leq 8.5) = 1 - P\left(Z \leq \frac{8.5-5.4}{1.3}\right) = 1 - \Phi\left(\frac{8.5-5.4}{1.3}\right) = 1 - .9915 = .0085$
  - The expected number would be $10000 \times .0085 = 85$
- Stata commands

```
. disp normal((8.5-5.4)/1.3)
.99145151
. disp 1 - normal((8.5-5.4)/1.3)
.00854849
. disp 10000*(1 - normal((8.5-5.4)/1.3))
85.484928
```

Inverse cumulative standard normal distribution

- Definition
  - $p = P(Z \leq z) = \Phi(z)$, then $z = \Phi^{-1}(p)$
  - Given p, find z such that $p = \Phi(z)$, i.e. $p = P(Z \leq z)$



- Stata function
  - Invnormal(p)

Inverse cumulative standard normal distribution

- Examples
  - Exercises 4.6.11 – 13 on page 122 (10th Ed). Given the following probabilities of standard normal r.v. Z, find z_1

Exercise 4.6.11 find z such that $0.0055 = P(Z \leq z_1) = \Phi(z_1)$ $z_1 = \Phi^{-1}(0.0055)$	. disp invnormal(.0055) -2.5426988
Exercise 4.6.12 $P(-2.67 \leq Z \leq z_1) = .9718$ $\Phi(z_1) - \Phi(-2.67) = .9718$ $\Phi(z_1) = .9718 - \Phi(-2.67)$ $z_1 = \Phi^{-1}(.9718 + \Phi(-2.67))$	. disp invnormal(.9718+normal(-2.67)) 1.970205
Exercise 4.6.13 $P(Z > z_1) = .0384$ $1 - P(Z \leq z_1) = .0384$ $P(Z \leq z_1) = 1 - 0.384$ $\Phi(z_1) = 1 - 0.384$	. disp invnormal(1-0.0384) 1.7695628

Inverse cumulative standard normal distribution

- Example
  - Give $X \sim N(\mu, \sigma^2)$, find k such that
$$P(\mu - k\sigma \le X \le \mu + k\sigma) = .754$$
- Solution

$$\mu - k\sigma \le X \le \mu + k\sigma \longrightarrow -k\sigma \le X - \mu \le k\sigma \longrightarrow -k \le \frac{X-\mu}{\sigma} \le k$$

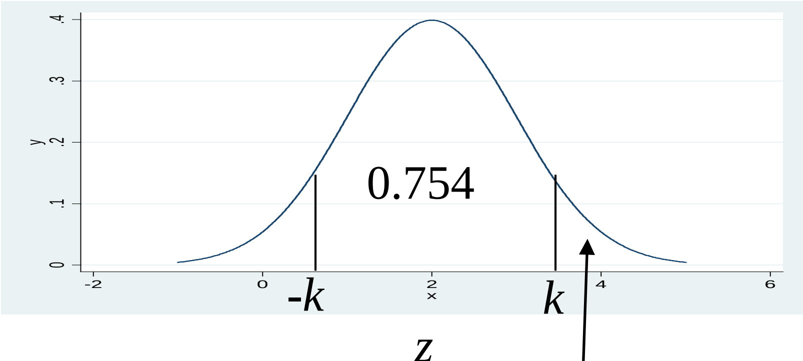
$$P(\mu - k\sigma \le X \le \mu + k\sigma) = P\left(-k \le \frac{X-\mu}{\sigma} \le k\right) = P(-k \le Z \le k) = .754$$

$$2\big(1 - \Phi(k)\big) = 1 - 0.754$$

$$1 - \Phi(k) = \frac{1-0.754}{2}$$

$$\Phi(k) = 1 - \frac{1-0.754}{2} = \frac{1+0.754}{2}$$

$$k = \Phi^{-1}\left(\frac{1+0.754}{2}\right) = 1.1528$$



$P(Z \ge k)$
$= 1 - P(Z \le k)$
$= 1 - \Phi(k)$

- Stata command
  - disp invnormal((1+0.754)/2)
  - 1.1527817

Summary

- What we have learned
  - Relationship between a general and standard normal distributions
  - How to calculate the probabilities of a general normal distribution using standard normal distribution.
  - Application of normal distribution
  - Inverse cumulative standard normal distribution
- Focus
  - Understanding of Relationship between a general and standard normal distributions is most important

Homework 07

- Due August 24, 2020
- Exercises 4.7.3 and 4.7.5 (Page 128)
- Review exercise 27 (page 132)
- Supplement exercise. Based on a survey conducted in 1988 – 1994, the distribution of height of males in the 20 – 29 age bracket (U.S. Census Bureau 1999) has the mean 69.3 (inches) and standard deviation 2.92 and follows a normal distribution reasonably well.
    - If a male is picked at random from that population, what is the probability that

      (1) he is higher than 71 inches, (2) his height is somewhere between 67 and 69 inches?
    - If 10 males were picked at random, (3) what is the probability that at least 5 of them are higher than the average height and (4) what is the probability that at least 5 of them are higher than 71 inches?

  (hint: you need to use both normal and binomial distributions)