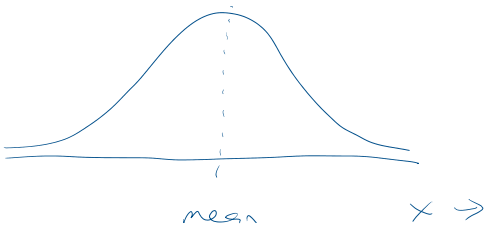
Section 9.2/9.3: The Normal Distribution

Wednesday, November 29, 2023 10:50 AM

we've looked at bell-shaped curves quite a lot:

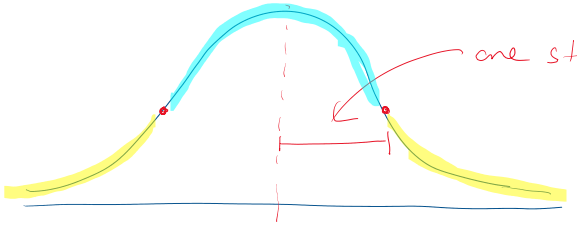
normal distribution



unimodal
symmetrical

we'll see later why this shape is so common, but for now we'll say that you can see this distribution whenever your continuous random variable is the result of many chance outcomes

note: you can estimate the standard deviation from the graph of the normal distribution



you look for the two points on the curve where the curvature goes from "concave up" ∪ to "concave down" ∩

digression: will not be tested

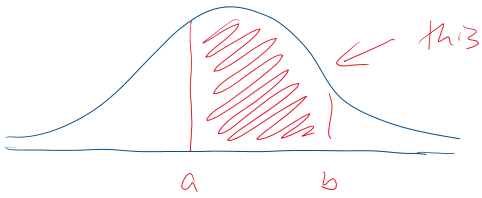
what is the shape? it's given by the formula

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad \text{for } -\infty < x < \infty$$

---

the probability of the data point x being between values a and b is equal

the probability of the data point x being between values a and b is equal to the area under the curve between points a and b



this area is equal to the probability that $a < x < b$

but how do you calculate this area?

1. use a computer or calculator (PREFERRED!)
2. look it up in a table of values

for this course, we will use an online calculator rather than using a table of values

---

if using a table, there's a problem! you'd need an infinite number of tables.

- one for every combination of μ (mean) and σ (std dev)

solution: standardize it

instead of using μ and σ, you use

$$z = \frac{x - \mu}{\sigma}$$

↗ the same z-score we looked at in Section 6.4

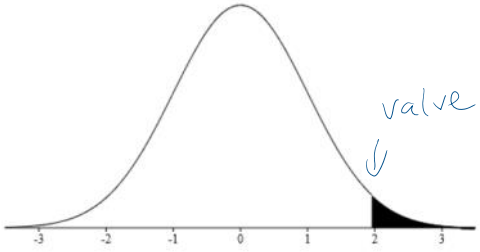
---

the online calculator I recommend using is the Hyperstat online Statistics Textbook

The online calculator I recommend using is the HyperStat Online Statistics Textbook

- link on Lectures page under Handouts and Resources
- link on D2L

HyperStat Online Home Page



◉ Area from a value (Use to compute p from Z) ← default option
○ Value from an area (Use to compute Z for confidence intervals)

Specify Parameters:
Mean 0 } input these values (mean and std dev)
SD 1
◉ Above 1.96
○ Below 1.96
○ Between -1.96 and 1.96 } choose one of these options
○ Outside -1.96 and 1.96

Results:
Area (probability) 0.025 ← this is your answer
Recalculate

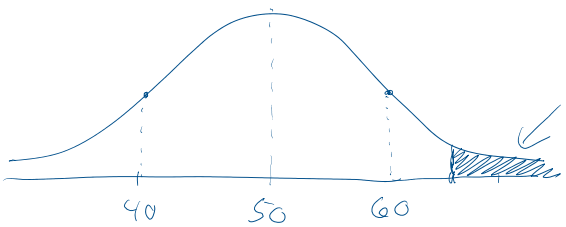
Screen clipping taken: 2023-11-30 3:38 PM

example: suppose we know that the time it takes to commute between campuses during rush hour is normally distributed with a mean of 50 minutes and std dev of 10 minutes.

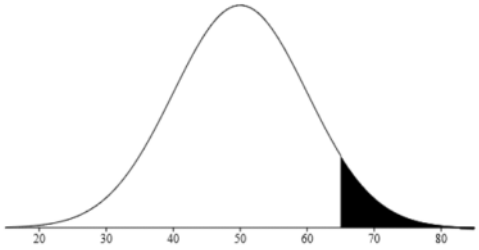
a) what is the probability that the next time you make this commute in rush hour that you will take over 65 minutes?

mean = 50 min
std dev = 10 min

$x > 65$ min (x is the data point of interest)



probability we want is equal to this area



Area from a value (Use to compute p from Z)
Value from an area (Use to compute Z for confidence intervals)

Specify Parameters:
Mean 50
SD 10
Above 65
Below 1.96
Between -1.96 and 1.96
Outside -1.96 and 1.96

Results:
Area (probability) 0.0668
Recalculate

Screen clipping taken: 2023-11-30 3:50 PM

this is our probability

$p = 0.0668$ or 6.68%

okay to round to 6.7% or 7% because original question didn't have much precision

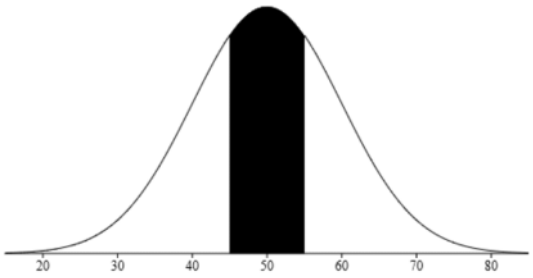
b) what is the probability that the next time you make this commute in rush hour it will take you less than 65 minutes?

you could use the online calculator again, or just say

note: probability of $x =$ exactly 65 is zero

$$P(x < 65) = 1 - P(x > 65)$$
$$= 1 - 0.0668$$
$$= 0.9332$$
or 93%

c) what is the probability that the next time you make this commute in rush hour that it will take between 45 and 55 minutes?



◉ Area from a value (Use to compute p from Z)
○ Value from an area (Use to compute Z for confidence intervals)

Specify Parameters:
Mean 50
SD 10
○ Above 65
○ Below 1.96
◉ Between 45 and 55
○ Outside -1.96 and 1.96

Results:
Area (probability) 0.3829
Recalculate

Screen clipping taken: 2023-11-30 4:00 PM

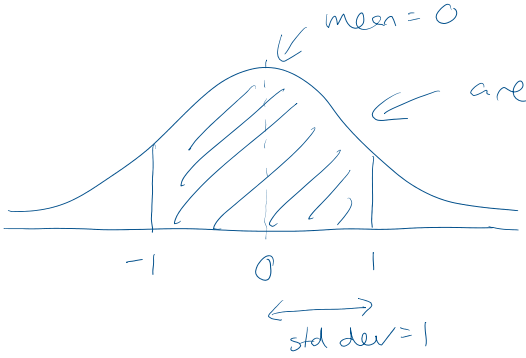
$p = 0.3829$ or 38%

---

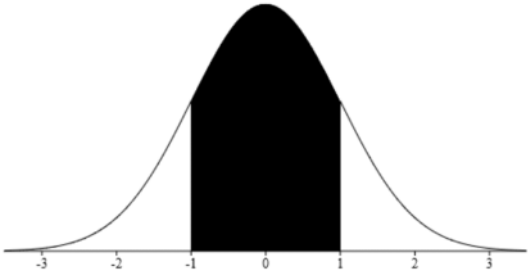
what is the probability that a data point taken at random from a normal distribution is within

a) one standard deviation of the mean?
b) two " " " " " "?

answer:



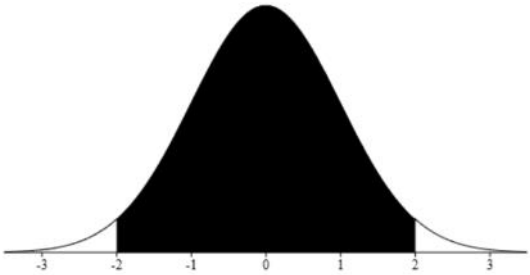
a)



◉ Area from a value (Use to compute p from Z)
○ Value from an area (Use to compute Z for confidence intervals)

Specify Parameters:
Mean 0
SD 1
○ Above 65
○ Below 1.96

b)



◉ Area from a value (Use to compute p from Z)
○ Value from an area (Use to compute Z for confidence intervals)

Specify Parameters:
Mean 0
SD 1
○ Above 65
○ Below 1.96

SD 1

- Above 65
- Below 1.96
- Between -1 and 1
- Outside -1.96 and 1.96

Results:
Area (probability) 0.6827
Recalculate

Screen clipping taken: 2023-11-30 4:08 PM

Mean 0
SD 1

- Above 65
- Below 1.96
- Between -2 and 2
- Outside -1.96 and 1.96

Results:
Area (probability) 0.9545
Recalculate

Screen clipping taken: 2023-11-30 4:09 PM

a) $P(-1 < z < 1) = 0.6827$ or 68.27%

b) $P(-2 < z < 2) = 0.9545$ or 95.45%

Oh, look! It's the Empirical Rule!

---

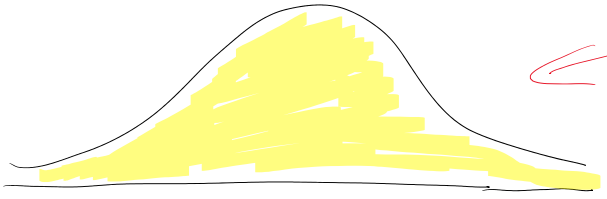
two last thoughts:

1. if you know the z-score of a value and want to compute probabilities related to that z-score, then use

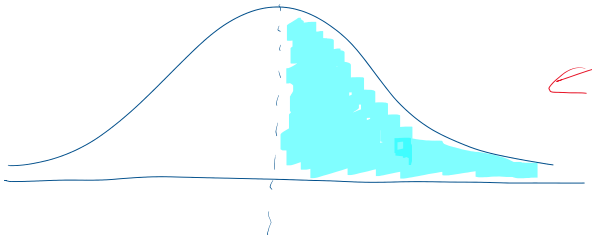
   mean = 0
   std dev = 1

   and then plug in your z-score where you'd put x in the calculator

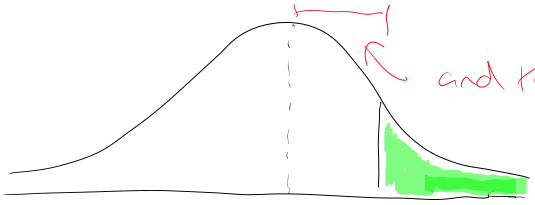
2. when visualizing your answer, it can be helpful to use symmetry:



total shaded area is equal to one



this shaded area is equal to 0.5



and this shaded area = 0.2

if this shaded area = 0.3

then this shaded area = 0.7