Probability: Normal Distributions

Grinnell College

March 4th, 2026

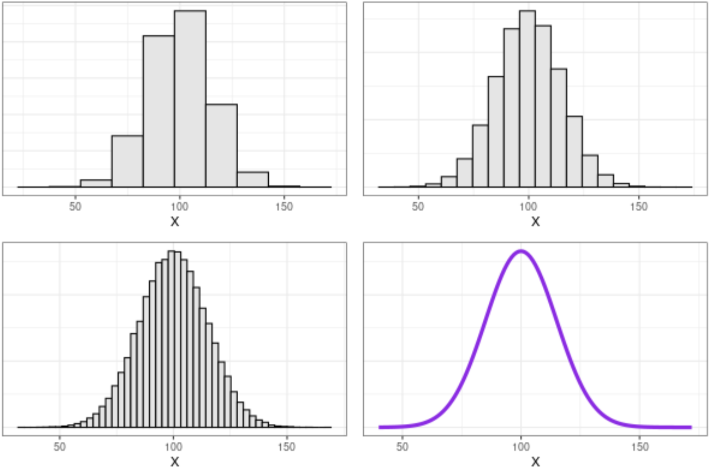
Motivation

We have now briefly seen a couple common distributions

- Uniform
- Binomial (didn't mention the name, number of odd digits out of 10 digits)
- Hypergeometric (not tested on this one, don't worry)
  - Grabbing 1 black juror and 11 white jurors from a pool of 14 and 21 respectively

It's now time to formally talk about the all time most normal distribution ever...the Normal Distribution!

The Normal Distribution



Normal Distribution

Some general info...

- First(?) person to write about it in useful way was Freidrich Gauss and is occasionally called the Gaussian distribution for this
- Second(?) person to write about it was Laplace who...
  - was the first person to prove it's normalization factor does lead to a total probability of 1
  - introduced the Central Limit Theorem
- Comes up often in nature
  - Not as often as people think

$$\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2}\left(\frac{x\ -\mu}{\sigma}\right)^2}$$

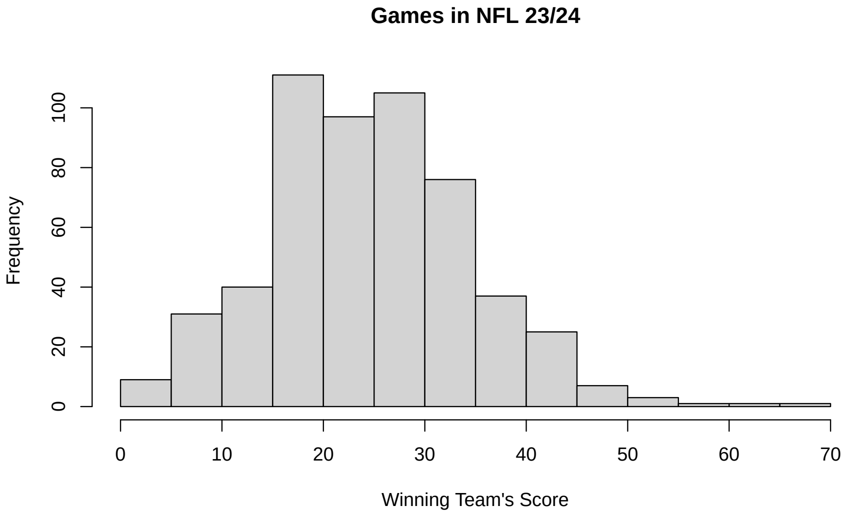
Normal out in Nature?

Often misidentified Binomial or Poisson distributions

- Binomail distribution is the distribution of how many (successes) out of (number of trials)
  - The distribution of the number of odd digits in 10 digits
- Poisson distribution is distribution that is often used for counting
  - Number of cars sold in a day at a dealership
  - Prussians soldiers kicked to death by horses
    - Better at estimating this than we should be
    - Like waaaay better than we should be

Normal or...Poison?!

Sorry, I mean Poisson (it's Halloween....)



Normal Approximation

Despite these differences in data generating mechanisms, we can usually approximate both of those distributions using a Normal Distribution

- Unit 3 is normal approximation for a binomial (proportion stuff)
- We (statisticians and scientists generally) don't even acknowledge when we approximate the Poisson usually
  - Total Sales for a grocery store recorded daily but in pennies.....

Normal Distribution

It turns out we only need to know two things in order to completely describe the Normal distribution

1. the mean (μ)
2. the standard deviation (σ) or variance (σ^2)

These will tell us where the center of the normal distribution is and how stretched out it should be.

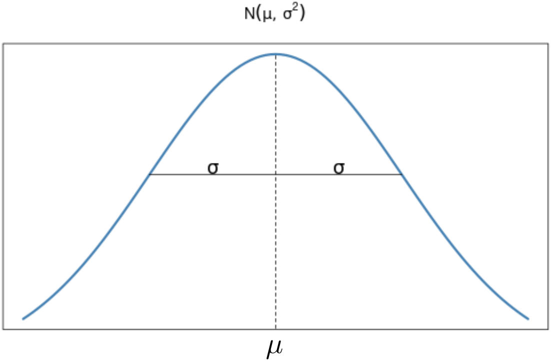
If a variable looks like a normal distribution, we will often use the following notation to say that:

- X $\sim$ N(μ, σ^2)

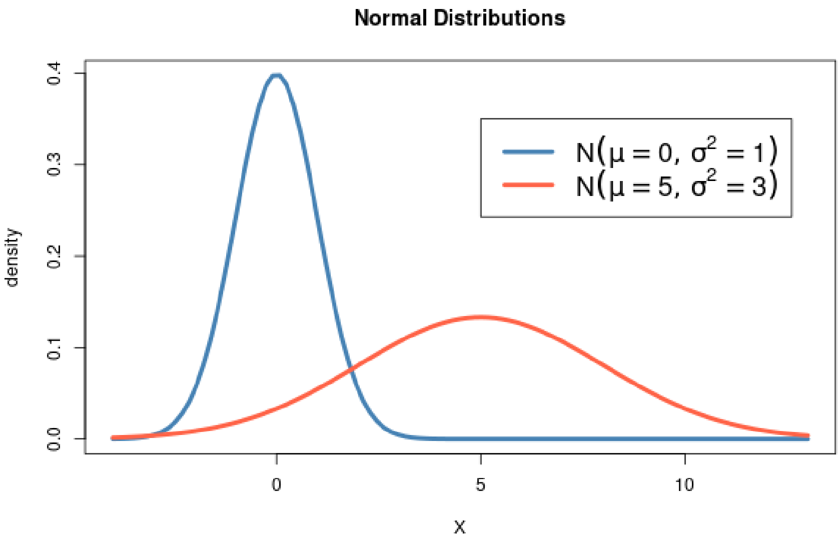
Normal Distribution

X $\sim$ N(μ, σ^2)

- the mean tells us where the center of the normal distribution is
- the variance tells us how spread out the distribution is



Examples



Normal Distributions
0.4
0.3
0.2
0.1
0.0
density
0
5
10
X
N(μ = 0, σ² = 1)
N(μ = 5, σ² = 3)

Standard Normal Distribution

When a normal distribution has mean zero and variance equal to 1, we call it a **Standard Normal Distribution** and write X $\sim$ N(0, 1).

Why? It's related to standardizing variable

Suppose the variable X $\sim$ N(μ, σ^2),
then $Y = \frac{X-\mu}{\sigma} \sim$ N($\mu = 0$, $\sigma^2 = 1$)

In other words, if we standardize a normal variable (with any mean and variance) then we get back a normal variable that has $\mu = 0$ and $\sigma^2 = 1$

What is standardizing?

Standardizing has different definitions to different fields in different settings. To us, it means subtracting the mean and dividing by the standard deviation.

$$\frac{X - \mu}{\sigma} = \frac{\text{Observation - Mean}}{\text{st. dev.}} \tag{1}$$

What does this number tell us?

What is standardizing?

Standardizing has different definitions to different fields in differen settings. To us, it means subtracting the mean and dividing by the standard deviation.

$$\frac{X - \mu}{\sigma} = \frac{\text{Observation - Mean}}{\text{st. dev.}} \tag{2}$$

What does this number tell us? It tells us how many standard deviations away from the mean our observation is.

Ie gives a sense of how far from average something is

Standardizing Example

Men have a mean height of approximately 70 inches with a standard deviation of 3.

I'm 68 inches.

Standardize my height

Standardizing Example

Men have a mean height of approximately 70 inches with a standard deviation of 3. I'm 68 inches. Standardize my height.

Note:

$$\begin{aligned} \mu = & \text{ mean } &= 70 \\ \sigma = & \text{ st. dev. } &= 3 \end{aligned}$$

1. 68 - 70 = -2

Standardizing Example

Men have a mean height of approximately 70 inches with a standard deviation of 3. I'm 68 inches. Standardize my height

Note:

$$\begin{aligned} \mu = & \text{ mean } & = 70 \\ \sigma = & \text{ st. dev. } & = 3 \end{aligned}$$

1. 68 - 70 = -2
2. -2 / 3 = -.66666

My standardized height is -.6666 so I'm 2/3 a standard deviation below the mean

Standardizing: Why?

The mean (μ) and the st. dev. (σ) are measured on the same units so the resulting value is unitless

- This allows us to compare observations on wildly different things
- (My ACT score) vs (My friend's SAT score)
- (My weight with respect to the average adult male) vs (Your mom's weight with respect to the average adult blue whale)
- Also comes up a whole lot in statistics like for Normal Distributions

ACT test mean: 19, st.dev.: 5
SAT test mean: 1500, st.dev.: 220

Standardizing: Why?

The mean (μ) and the st. dev. (σ) are measured on the same units so it's unitless

- This allows us to compare observations on wildly different things
- (My ACT score) vs (My friend's SAT score)
- (My weight with respect to the average adult male) vs (Your mom's weight with respect to the average adult blue whale)
- Also comes up a whole lot in statistics like for Normal Distributions

ACT test mean: 2, st.dev.: 5
SAT test mean: 1500, st.dev.: 300

Any brave soul want to compare their two exams?

Standardizing: Why?

We defined probability as the long term frequency of events happening; that is it is a proportion of (EVENT) out of (NUMBER OF TRIALS) with many many many trials

- Our proportions are invariant
  - We can shuffle our distribution up and down the number line (subtracting the mean)
  - Or make it thinner or fatter (dividing by st. dev.)
  - but the same proportion of data will below a given (transformed) observation at all steps
    - Eg the median will always be our median
    - Eg our bottom 10th percentile will be our bottom 10th percentile
- So we only need to understand the standard normal distribution at the end of the day
  - Because we can backtransform to the original scale

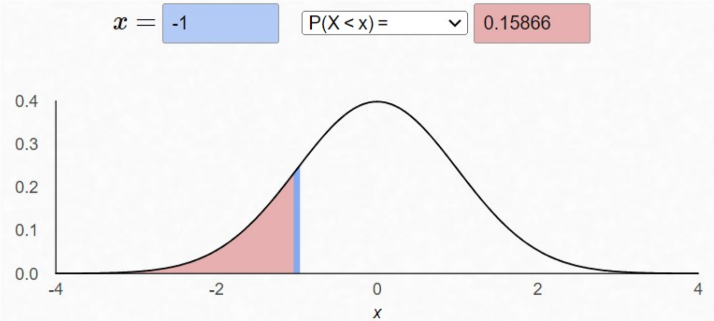
Probabilities

If our population follows a normal distribution... we can pick a case at random from our population

- probability the observation is less/greater than some value?
- probability the observation is between two values?

Note: It turns out that using a normal distribution, or any other continuous distribution, we cannot find the probability of the case having a *specific* value, we can only use ranges of values.

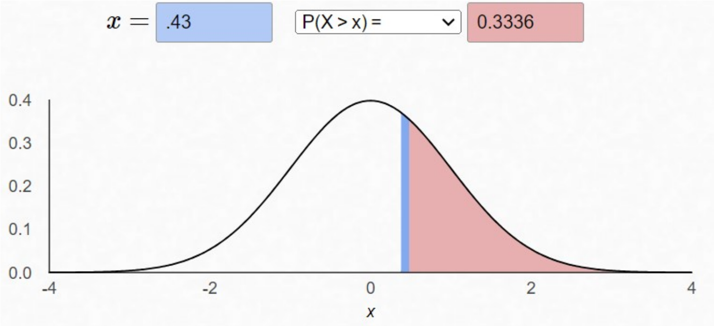
Standard Normal: X $\sim$ N(0, 1)
Probability a randomly selected observation is below (less than) -**1**?



We can write this using our probability notation: P(X $<$ -1) = 0.15866

Probabilities – Greater than

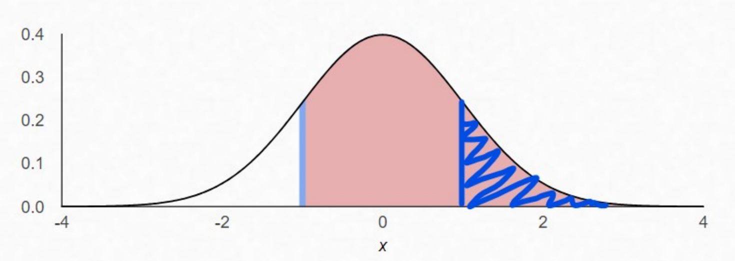
Standard Normal: X $\sim$ N(0, 1)
Probability a randomly selected observation is above (greater than) **0.43**?



P(X > 0.43) = 0.3336

Probabilities – Between

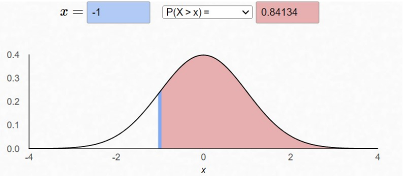
Standard Normal: $X \sim N(0, 1)$
What about the probability that a case falls *between* **-1** and **1**?

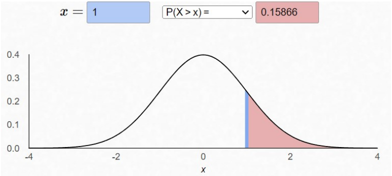


We need to do a bit more work...

Probabilities – Between

Standard Normal: X $\sim$ N(0, 1)
What about the probability that a case falls *between* -**1** and **1**?





We can chop off the extra probability we don't need that is above **1**.

P(X is between -1 and 1) = P(-1 $<$ X $<$ 1) = P(X $>$ -1) - P(X $>$ 1)
= 0.84134 - 0.15866 = 0.68286

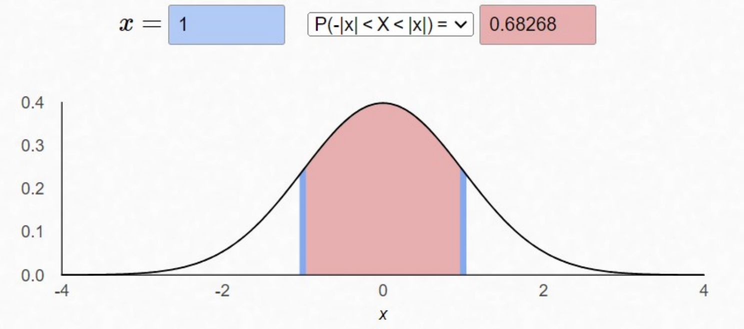
Probabilities – Between

When the values we are looking at are the same but just with different signs (like -1 and $+1$)

- We can write them in a specific way

Probabilities – Between

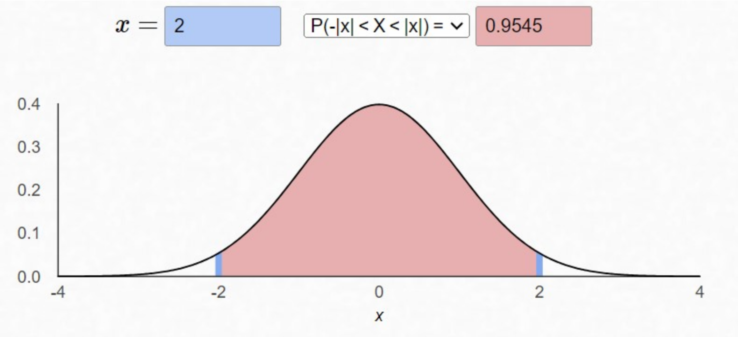
Standard Normal: X $\sim$ N(0, 1)
What about the probability that a case falls *between* **-1** and **1**?



$$P(|X| < 1) = 0.68286$$

Probabilities – Between

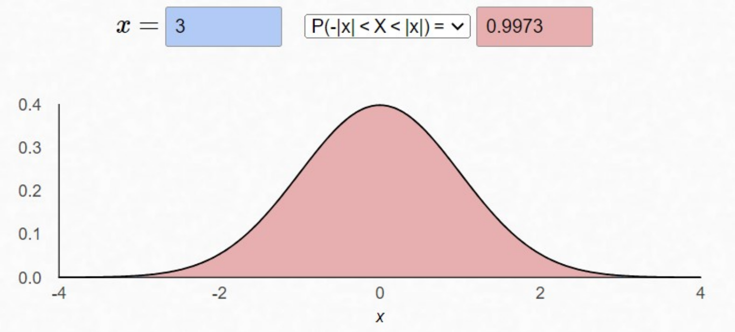
Standard Normal: X $\sim$ N(0, 1)
What about the probability that a case falls *between* -**2** and **2**?



$$P(|X| < 2) = 0.9545$$

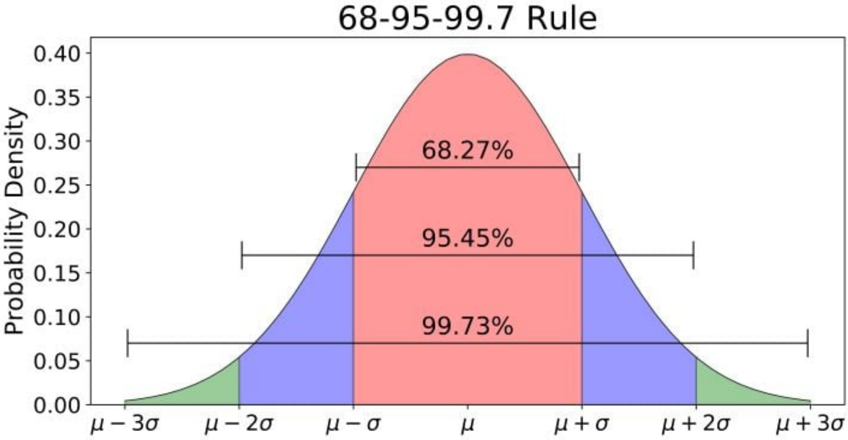
Probabilities – Between

Standard Normal: X $\sim$ N(0, 1)
What about the probability that a case falls *between* -**3** and **3**?



$$P(|X| < 3) = 0.9973$$

Summary


68-95-99.7 Rule
Probability Density
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
68.27%
95.45%
99.73%
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$

Probabilities from R

We can use the "pnorm()" function in R to get these probabilities.

- tell the function what number you are trying to find the probability more/less than
- Need to tell the function the value of the mean and st. dev.
- Doesn't require standard normal actually

Note: By default R will try to give you 'less than' probabilities (also called lower tail probabilities). To get 'greater than' probabilities, put "lower.tail=FALSE" into the pnorm() function.

```
> pnorm(-1, mean=0, sd=1)
[1] 0.1586553
> pnorm(-1, mean=0, sd=1, lower.tail = FALSE)
[1] 0.8413447
> pnorm(-1, mean=0, sd=1, lower.tail = FALSE)
- pnorm(1, mean=0, sd=1, lower.tail = FALSE)
[1] 0.6826895
```