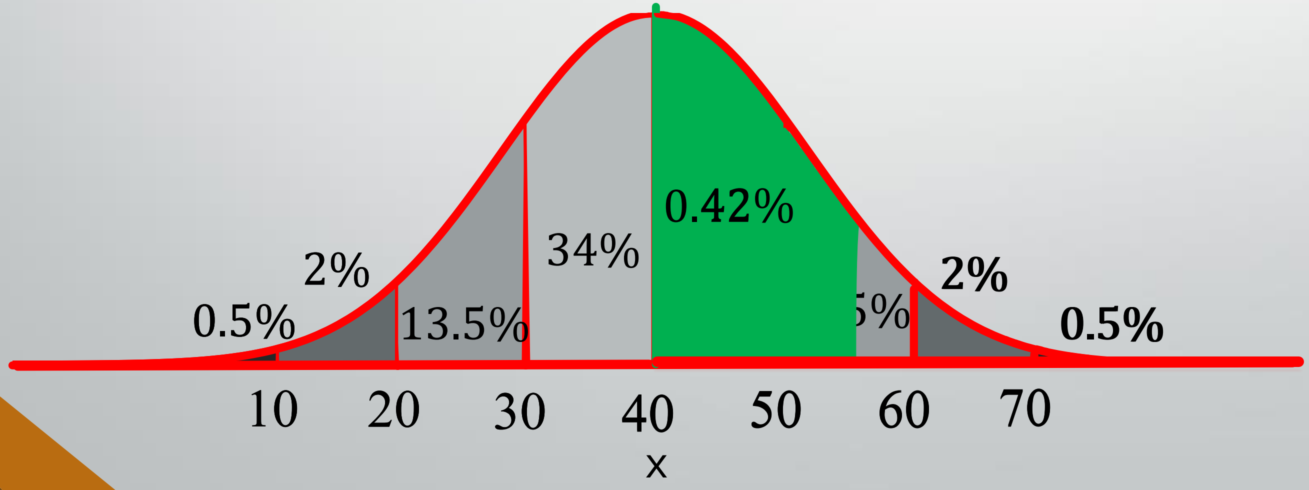
Standard Normal Curve

ID1050– Quantitative & Qualitative Reasoning

Using a table instead of a graphic

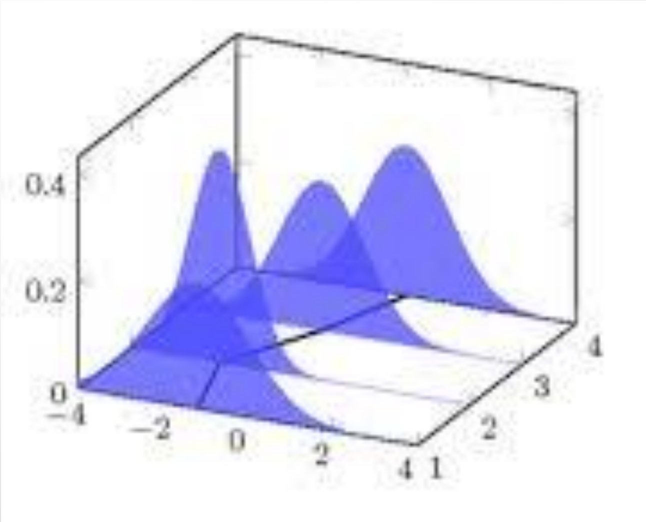
- So far, we have been limited to 0, +-1, +-2, +-3 std. dev. away from the mean.
- We might want more precision, so we could use a tabular format, instead.
  - Say that we have data with mean=40, std. dev.=10.
  - Since the curve has mirror-image symmetry, we can build a table for just the right half.
  - In our new higher precision table, we let Column A hold the x-value, which will run from the mean (here, that's 40) to three standard deviations above (here, that's 70).
  - We can use finer step sizes, like 1 instead of 10 (or even smaller).
  - Column B could be the percentage or fraction under the curve between the mean and the x-value in column A.
  - Using the fraction instead of percentage lets us drop the '%' symbol.



(A) x	(B) Fraction of area between mean and x
40	0.000
41	0.040
42	0.079
…	…

Tables for Every Set of Data?

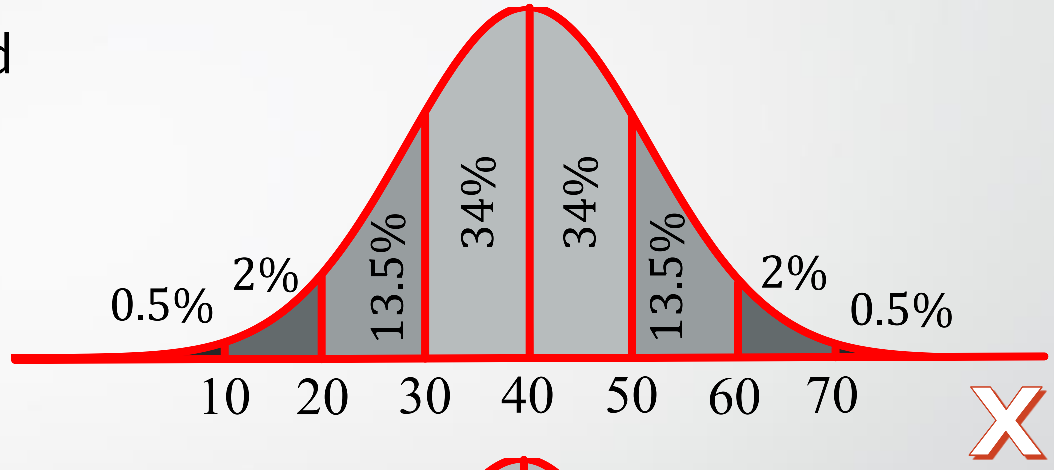
- What if we have a different set of normally distributed data, with a different mean and standard deviation?
  - For any given normal data, the curve has the same shape, but it will be shifted left or right (due to the different mean) and compressed or stretched (due to a different standard deviation).
  - We would have to build a new table, since all of our x-values will be different.
- We would rather not have to re-calculate our table for every new set of data.
- So, we pick standard values for the mean and the standard deviation, build a standard table (called a z-table), then figure out how to use this one table for any set of normal data.

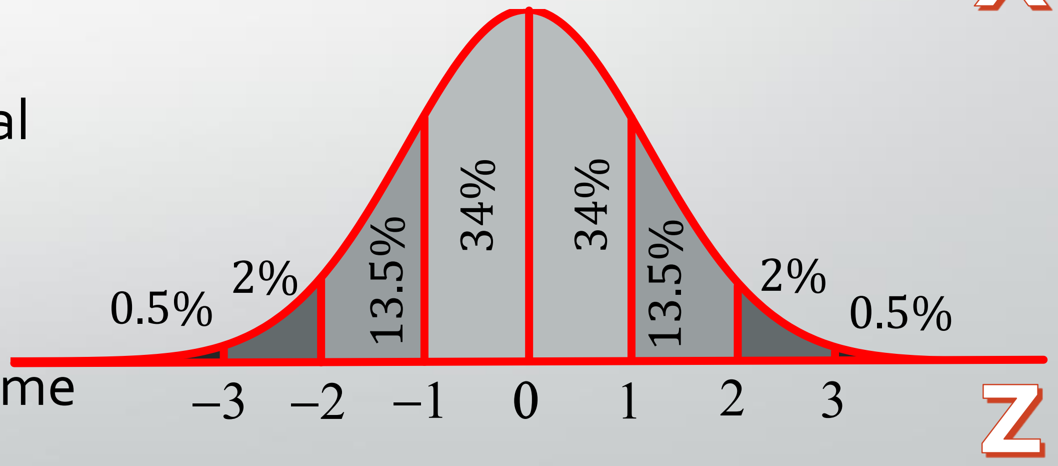


The Standard Normal Curve and the Z-table

So what standard values for mean and standard deviation should we choose?

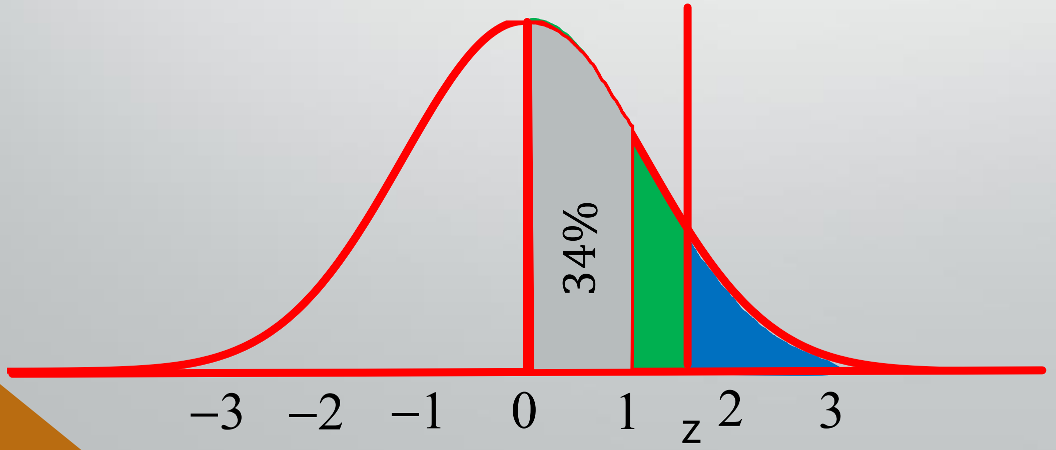
- The standard mean will be zero.
- The standard deviation will be one.
- We'll call the new variable 'z' to distinguish it from 'x', which we are using for the actual normal data we are using.
- We'll need a formula for converting the actual data, 'x', into the standard variable, 'z'.
  - $z = \frac{(x-\mu)}{\sigma}$
- Since the shape of the normal curve is the same for all normal data, we don't have to back-convert the results of the table.





The Z-Table

- Here is the new table:
  - Column A is the z-value
  - Column B is the fraction of area under the curve between z and the mean (0).
    - This is the area from z toward the 'hump'.
  - For convenience, we can also add a third column, C, that is the area *beyond* z.
    - This is the area from z toward the 'tail'.
  - The values in columns B and C sum to 0.5, since together they represent half of the area under the normal curve.
  - When z=1, we see the familiar 34% at one standard deviation from the mean.



(A) z	(B) Area between mean and z	(C) Area beyond z
0.0	0.000	0.500
0.1	0.040	0.460
0.2	0.079	0.421
0.3	0.118	0.382
0.4	0.155	0.345
0.5	0.192	0.309
0.6	0.226	0.274
0.7	0.258	0.242
0.8	0.288	0.212
0.9	0.316	0.184
1.0	0.341	0.159
…	…	…
2.9	0.498	0.002
3.0	0.499	0.001

Conclusion

- We can increase our precision when working with statistical data by replacing the graphic with regions labeled with percentages with a table. Our table can have as many rows as we wish.
- To prevent having to generate a new table any time the mean and standard deviation of the data changes, we develop a standard table:
  - This table uses a new variable, z, which has a mean of zero and a standard deviation of 1.
  - We can convert from our data variable, x, to z using the formula $z = \frac{(x-\mu)}{\sigma}$.
- Our table has three columns:
  - Column A is the z-value. Due to symmetry, we only need positive values of z.
  - Column B is the area under the standard normal curve between z and zero (toward the hump.)
  - Column C is the area under the standard normal curve beyond z (toward the tail.)
- There are other ways to form a z-table to capture the same information, but this is the style this course will use.