A Common Mistake

Quincy has a 93% in his math class, and 94% in physics. Mathematically speaking, of which grade should he be more proud?

The Normal Distribution

Continuous vs. Discrete | Understanding | z-scores and t-scores | Application | Normal Approximation | Confidence intervales

Why the Normal?

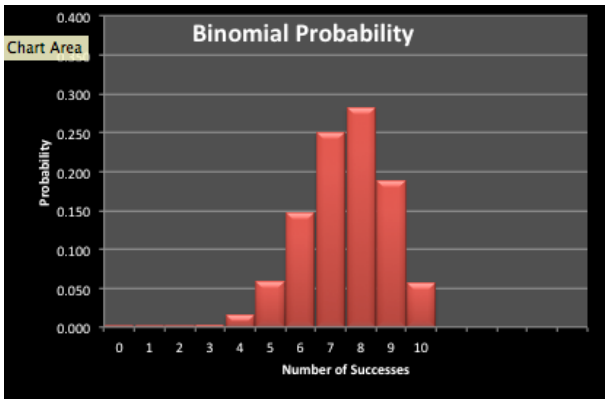
1. Naturally occurring phenomenon
2. Can be used to estimate binomial
3. Can even be used with non-normal distributions

Continuous versus Discrete Data

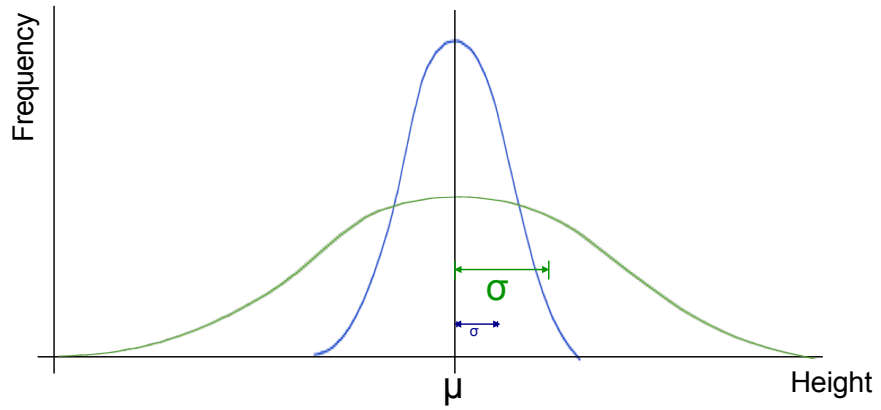
Is this binomial probability distribution a discrete or continuous probability distribution? Why?

n	10
P	0.75
Q	0.25

r	nC_r	P^r	Q^{n-r}	P(x)
0	1	1.00	0.00	0.000
1	10	0.75	0.00	0.000
2	45	0.56	0.00	0.000
3	120	0.42	0.00	0.003
4	210	0.32	0.00	0.016
5	252	0.24	0.00	0.058
6	210	0.18	0.00	0.146
7	120	0.13	0.02	0.250
8	45	0.10	0.06	0.282
9	10	0.08	0.25	0.188
10	1	0.06	1.00	0.056



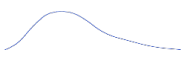
The Symmetric or Normal Distribution



=min() =mean() =max()

=median()

=mode()



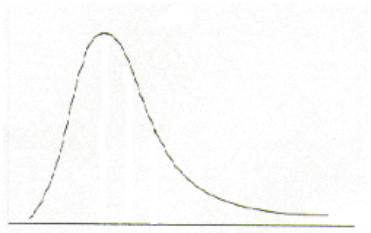
Continuous & Discrete Data

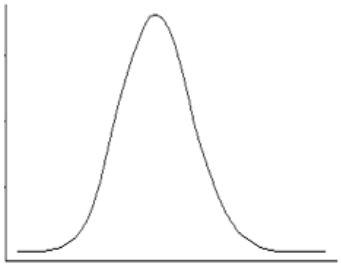
1. Identify each of the following situations as discrete distributions or continuous distributions. Explain your reasoning.

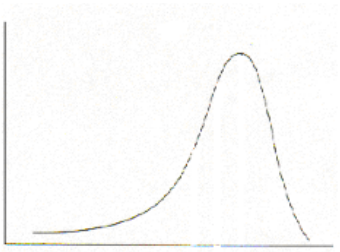
   a) counting the number of outcomes for drawing a card
   b) measuring the time taken to complete a task
   c) counting the number of outcomes when tossing three coins
   d) measuring the maximum distance a ball can be thrown

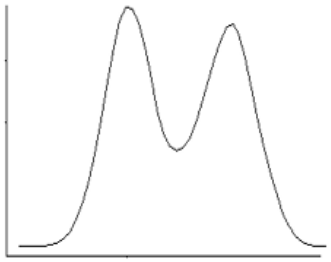
Distributions of Continuous Data

Label each graph as being skewed-left, skewed-right, bimodal, or symmetric.
In small groups develop a list of situations in which the data would produce the 4 different graphs.

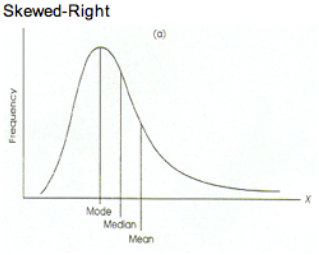






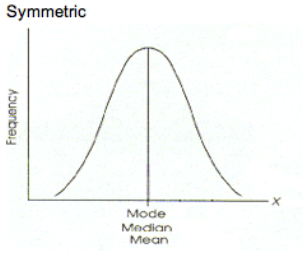


Skewed-Right



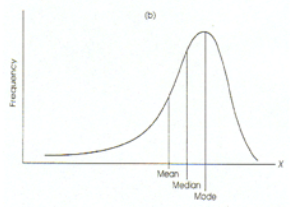
Mean> Median> Mode

Symmetric



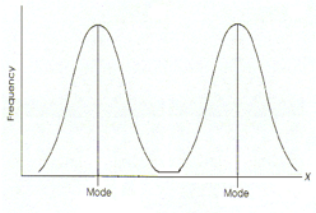
Mean= Median= Mode

Skewed-Left



Mean< Median< Mode

Bimodal



2 modes

2. Using the mean, median and mode, describe the shape of the frequency histogram.
   a) mean: 7.5 median: 6 mode: 5.7
   b) mean: 6 median: 6 mode: 10, 12
   c) mean: 7.5 median: 8.5 mode: 9
   d) mean: 7.5 median: 7.5 mode: 7.5

The Symmetric or Normal Distribution

The Normal Distribution is a special curve that we can use to find probabilities.

Check out this Fathom 2 file:

Note what happens to a symmetric distribution when the standard deviation and mean are changed.

Questions to ponder:

What is the effect of changing the mean?

What is the effect of changing the standard deviation?

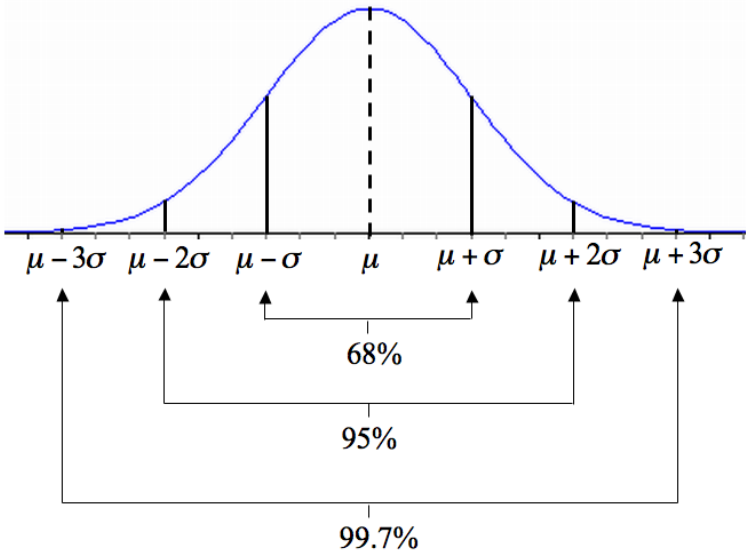
How might you describe standard deviation to a friend?

It is said that the area under the normal curve is related to probability. What might that mean?

68%, 95%, 99.7% Rule

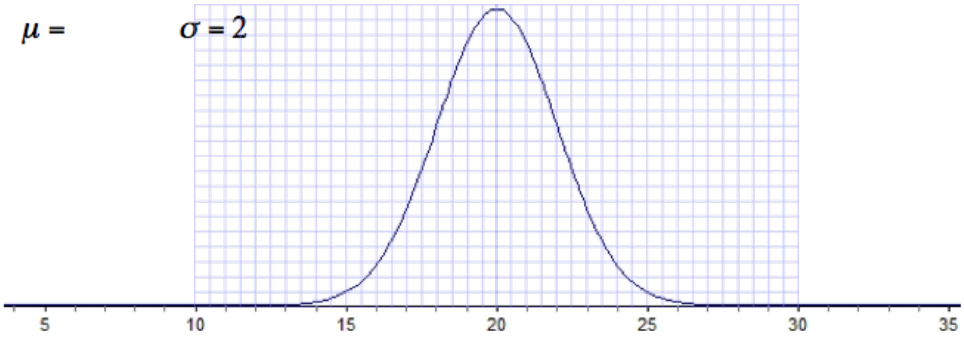
For any given normal distribution X, the percentage of data;

- between $\mu-\sigma$ and $\mu+\sigma$ is approximately 68%. (i.e. within one standard deviation of the mean.)
- between $\mu-2\sigma$ and $\mu+2\sigma$ is approximately 95%
- between $\mu-3\sigma$ and $\mu+3\sigma$ is approximately 99.7%



Features of the Normal Distribution

Approximate the area under the curve by counting the squares - be as accurate as possible in your estimate, but realize that you may have to use some guesswork.



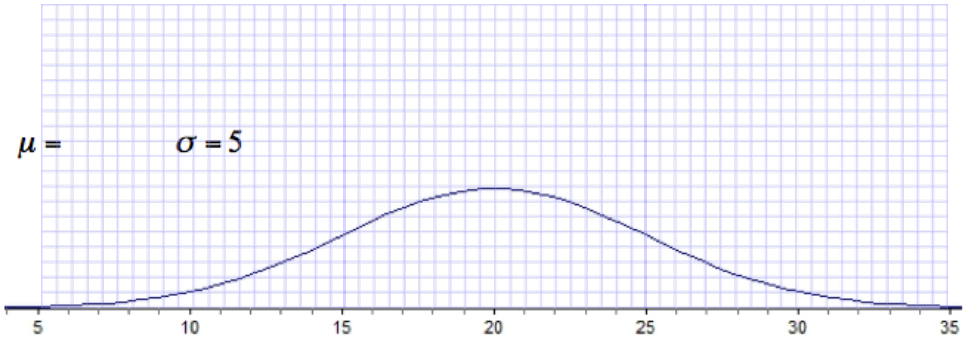
Total number of squares under the curve: _____

Number of squares ± 1 Std. Dev of mean: _____ Fraction of Total: _____

Number of squares ± 2 Std. Dev of mean: _____ Fraction of Total: _____

Number of squares ± 3 Std. Dev of mean: _____ Fraction of Total: _____

Approximate the area under the curve by counting the squares - be as accurate as possible in your estimate, but realize that you may have to use some guesswork.



Total number of squares under the curve: _____

Number of squares ± 1 Std. Dev of mean: _____ Fraction of Total: _____

Number of squares ± 2 Std. Dev of mean: _____ Fraction of Total: _____

Number of squares ± 3 Std. Dev of mean: _____ Fraction of Total: _____

Summarize Your Observations: What do you notice about the number of squares under the curve for each distribution? Write a rule for the percentage of data points within 1, 2 and 3 standard deviations of the mean.

Example Problems

1. Sally is 164 cm tall. In her school, the girls heights are normally distributed with a mean of 168cm and a standard deviation of 4cm.
   a. What is the probability that Sally's friend Joanne is taller than she is?
   b. What is the probability that Joanne is between 164cm and 172cm tall?

2. The daily sales of Gary's chip truck has a mean of $675.00 and a standard deviation of $35.50.
   a. What percent of time will the daily sales be greater than $639.50?
   b. What percent of time will the daily sales be less than $746.00?

3. The mean household income in Kingston is $45000 with a standard deviation of $15000. Household incomes below $30000 will receive a tax credit, household incomes between $30000 and $75000 will have to pay a 2% tax, and household incomes over $75000 will have to pay a 5% tax.
   a. What percentage of households will have to pay a 2% tax?
   b. What percentage of households will not have to pay tax?
   c. What percentage of households will pay tax?

4. An elite university only accepts the top 5% of students within the province to attend their university. Last year the student's average marks were normally distributed with a mean of 75% and a standard deviation of 7.5%. What average is needed in order to attend this university?

Answer Clues

1a) ~84%
b) ~ 68%

2a)~84%
b) ~97.5%

3a)~81.5%
b)~16%
c)~84%

4a)~87.3%