Law of Large Numbers
Rules for Means and Variances
Random Variable

A random variable is a variable whose value is a numerical outcome of a random phenomenon.
Mean of a Discrete Random Variable

Suppose that $X$ is a discrete random variable whose distribution is

<table>
<thead>
<tr>
<th>Value of $X$: $x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability: $p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_k$</td>
</tr>
</tbody>
</table>

To find the mean of $X$, multiply each possible value by its probability, then add all the products:

$$
\mu_X = x_1p_1 + x_2p_2 + \ldots + x_kp_k \\
= \sum x_ip_i
$$
The best way to make this decision is by calculating the expected value of each possible outcome.

You multiply the... I sense that we're done here.

You must pretend to be dead. I hope the dead sometimes cover their ears.
Variance of a Discrete Random Variable

Suppose that $X$ is a discrete random variable whose distribution is

<table>
<thead>
<tr>
<th>Value of $X$:</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>...</th>
<th>$x_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability:</td>
<td>$p_1$</td>
<td>$p_2$</td>
<td>$p_3$</td>
<td>...</td>
<td>$p_k$</td>
</tr>
</tbody>
</table>

and that $\mu$ is the mean of $X$. The variance of $X$ is

$$\sigma_X^2 = (x_1 - \mu_X)^2 p_1 + (x_2 - \mu_X)^2 p_2 + \ldots + (x_k - \mu_X)^2 p_k$$

$$= \sum (x_i - \mu_X)^2 p_i$$

The standard deviation $\sigma_X$ of $X$ is the square root of the variance.
Law of Large Numbers

Draw independent observations at random from any population with finite mean $\mu$. Decide how accurately you would like to estimate $\mu$. As the number of observations drawn increases, the mean $\bar{x}$ of the observed values eventually approaches the mean $\mu$ of the population as closely as you specified and then stays that close.
Rules for Means

- On the math portion of the SAT, the mean score was 510 with a standard deviation of 95. On the verbal portion, the mean score was 530 with a standard deviation of 85. What was the combined mean?
  - Finding a combined mean is relatively simple.
  - If X and Y are two different random variables (such as math score and verbal score), then:

\[
\mu_{X+Y} = \mu_X + \mu_Y
\]

  - That is, simply add the two means.
  - Then the mean combined score is \(510 + 530 = 1040\)
More Rules for Means

- Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.
  - What if, due to a misplaced measuring stick, we accidentally measured everyone to be 1 inch shorter than they actually were?
  - What would the new mean be?

\[ \mu_{a+X} = a + \mu_X \]

- That is, the mean was just an inch smaller than it should have been, so simply add the inch to the mean.
- Then the new mean is 67 + 1 = 68 inches.
More Rules for Means

- Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.
  - What if we needed to the measurements in centimeters instead of inches (there are 2.54 cm in one inch)?
  - What would the new mean be?

\[ \mu_{bX} = b \mu_X \]

- That is, we can simply convert the mean to cm by multiplying by 2.54.
- Then the new mean is 67 (2.54) = 170.18 cm.
More Rules for Means

• Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.
  – With means you can combine the two previous ideas to arrive at a general rule.

$$\mu_{a+bX} = a + b\mu_X$$

– The moral of the story of means, which seems to be common sense, is:
  • You can add them to one another.
  • You can add a constant to them.
  • You can multiply them by a constant.
Example 7.10 p494

- Linda sells cars and trucks \( X: \# \) cars sold, \( Y: \# \) truck and SUVs sold

- Cars sold: 0 1 2 3
- Probability 0.3 0.4 0.2 0.1

- Vehicles Sold: 0 1 2
- Probability: 0.4 0.5 0.1

- At her commission rate of 25% of gross profits she sells, Linda expects to earn $350 for each car sold and $400 for each truck or SUV sold. Find her expected earnings.

Step 1: Find means for cars sold and vehicles sold
• Step 2: Write equation for earnings:

\[ \$ = 350X + 400Y \]

Combine rules for means:

\[ U\$ = 350Ux + 400Uy \]

Linda’s best estimate for earnings for the day:
Rules for Variances

• On the math portion of the SAT, the mean score was 510 with a standard deviation of 95. On the verbal portion, the mean score was 530 with a standard deviation of 85. What was the combined standard deviation?
  – Finding a combined standard deviation is more complex.
  – If X and Y are two different independent random variables (such as math score and verbal score), then:

\[
\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y
\]

  – That is, you CANNOT add the standard deviations, but you CAN ADD THE VARIANCES!
  – Then the standard deviation of the combined score is:

\[
\sigma^2_{X+Y} = 85^2 + 95^2 \\
\sigma^2_{X+Y} = 7225 + 9025 \\
\sigma^2_{X+Y} = 16250 \\
\sigma_{X+Y} = 127.48
\]
Working with Combined Variables

- On the math portion of the SAT, the mean score was 510 with a standard deviation of 95. On the verbal portion, the mean score was 530 with a standard deviation of 85.
  - We now know the mean of the combined score is 1040 with a standard deviation of 127.48.
  - Given that these are normal distributions, we can do normal calculations with the combined statistics.
  - What is the probability that a random student scored 1100 or better?
    • Normalcdf (1100,9999,1040,127.48) = .3189
  - What score would you need to be in the top 5%?
    • Invnorm (.95,1040,127.48) = 1249.7
More Rules for Variances

• Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.
  – What if, due to a misplaced measuring stick, we accidentally measured everyone to be 1 inch shorter than they actually were?
  – What would the new standard deviation be?

\[ \sigma_{a+X} = \sigma_X \]

– Recall that standard deviation is a measure of spread. How much would the spread change if we just added one to each number in the dataset?
– That is, the spread, and thus standard deviation (and variance) is exactly the same.
– Then the new standard deviation is still 3 inches.
More Rules for Variances

Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.

– What if we needed to the measurements in centimeters instead of inches (there are 2.54 cm in one inch)?
– What would the new standard deviation be?
– Certainly the data’s spread will increase if we multiply each number in the dataset by 2.54.

\[ \sigma_{bX}^2 = b^2 \sigma_X^2 \]

– That is, don’t work with the standard deviation, use the variances. Then, you have to similarly square the constant (2.54) you’re working with as well.
– Then the new standard deviation is 7.62.
More Rules for Variances

• Consider measuring the height of everyone in the class. We obtain a mean of 67 inches and a standard deviation of 3 inches.
  – With standard deviations (and variances) you can combine the two previous ideas to arrive at a general rule.

\[ \sigma_{a+bX}^2 = b^2 \sigma_X^2 \]

– The moral of the story of standard deviations is to work with the variances instead:
  • You can add the variances but you cannot add the standard deviations.
  • Add a constant to the variable does not affect the standard deviation or variance.
  • You can multiply the variance by the square of the constant.
Example 7.13 p 497

- Zadie has invested 20% of her funds in T-bills and 80% in an “index fund” that represents all US common stocks. The rate of return of an investment over a time period is the percent change in the price during that time period, plus any income received. If X is the return on T-bills and Y the return on stocks, the portfolio rate of return R is:

  - \( R = 0.2X \) and \( 0.8Y \)

- The returns X and Y are random variables because they vary from year to year. Based on returns between 1950 and 2000,

  - \( U_x = 5.0\% \quad SD_x = 2.9\% \quad \text{**Note: Which of these has a greater risk???}}
  - \( U_y = 13.2\% \quad SD_y = 17.6\% \)

- Find the mean and SD of the portfolio return.
• Find the mean of the portfolio:

  \[ R = 0.2X + 0.8Y \]

  \[ Ur = 0.2Ux + 0.8Uy \]

  \[ = 0.2(5.0) + 0.8(13.2) = 11.56\% \]

• The expected return on the portfolio is 11.56\%. 
• Can we find the standard deviation of the portfolio?

  – In order to do this we would have to convert to variances and add. This rule is only an option when the events are independent.

  – Are T-bills and stocks independent?