Recall:

- $Var(X) = E[(X-\mu)^2]$, where $E[X] = \mu$.
- $E[aX+b] = aE[X] + b, for\ any\ a, b \in \mathbb{R}$.
- $E[X+Y] = E[X] + E[Y]$.
- $E[aX+bY+c] = aE[X] + bE[Y] + c$.

Before the example that we promised, we will prove an important theorem that makes computing variance much simpler.

Theorem: Let X be a r.v. with mean $E[X] = \mu$. Then, $Var(X) = E[X^2] - E^2[X] = E[X^2] - (E[X])^2$.

Proof:

$$Var(X) = E[(X-\mu)^2] \quad [by\ definition]$$
$$= E[X^2 - 2\mu X + \mu^2]$$
$$= E[X^2] - 2\mu E[X] + \mu^2 \ [by\ linearity\ of\ expectation]$$
$$= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 = E[X^2] - E^2[X]$$

Q.E.D.

Remark: When asked to find variance, we only need to find two quantities, $E[X^2]$ and $E[X]$. Never use the definition of variance to calculate it – it is much more difficult. Save yourself the trouble and use this theorem to calculate it.

<u>Example:</u> Suppose we play a game as follows: we flip a coin with $P(heads) = \frac{2}{3}$. If it comes up heads, I win \$3, otherwise I lose \$2. Let X be how much money I win in one coin flip. What is $Var(X)$? [*Note: although the problem only asks for variance, we will find the expectation in the process of doing so*].

Step 1: Find the pmf

The pmf for X is

$$p_X(x) = \begin{cases} 2/3, & x = 3 \\ 1/3, & x = -2 \end{cases}$$

Step 2: Find $\boldsymbol{E[X]}$ and $\boldsymbol{E[X^2]}$

$$E[X] = \frac{2}{3} * (3) + \frac{1}{3} * (-2) = 2 - \frac{2}{3} = \frac{4}{3}$$

$$E[X^2] = \frac{2}{3} * (3)^2 + \frac{1}{3} * (-2)^2 = 6 + \frac{4}{3} = \frac{22}{3}$$

Note: all we did was square the value of x

Step 3: Use $\boldsymbol{E[X]}$ and $\boldsymbol{E[X^2]}$ to find $\boldsymbol{Var(X)}$

$$Var(X) = E[X^2] - E^2[X] = \frac{22}{3} - \left(\frac{4}{3}\right)^2 = \boldsymbol{\frac{50}{9}}$$

Theorem: $Var(aX + b) = a^2 Var(X)$

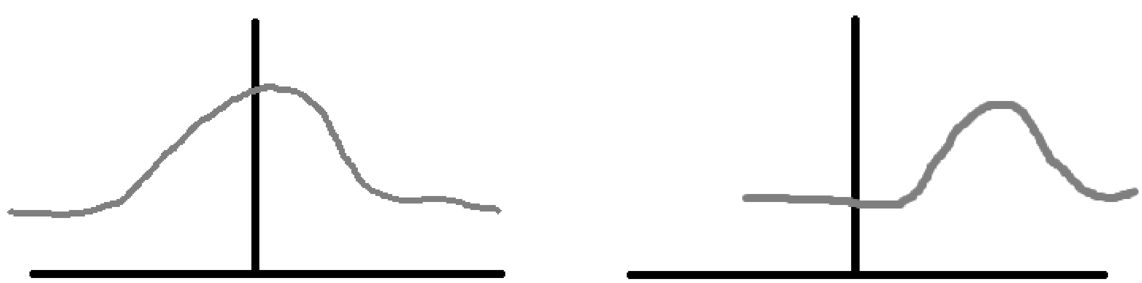
Proof:

$$E[aX + b] = aE[X] + b \; [by\ linearity\ of\ expectation]$$

$$Var(aX + b) = E\left[\left((aX + b) - (a\mu + b)\right)^2\right] \; [by\ definition]$$
$$= E\left[\left(a(X - \mu)\right)^2\right] = E[a^2(X - \mu)^2]$$
$$= a^2 E[(X - \mu)^2] \; [by\ linearity\ of\ expectation]$$
$$= a^2 Var(X) \; [by\ definition]$$

Q.E.D.

Note: $Var(aX + b) = a^2 Var(X)$ ***does not depend on b***. *Variance is greater, but it is not a measure of center or location, but the "variation" or "spread". Let's look at* X *and* $X + 2$. *(they are supposed to look the same, just shifted).*



Note: In general, $Var(X + Y) \neq Var(X) + Var(Y)$.

Example:

$$Var(X + X) = Var(2X) = 4Var(X) \neq Var(X) + Var(X)$$

Recall: Independent events: $P(E \cap F) = P(E)P(F)$

Definition: Two random variables X and Y are independent, $X \perp Y$ iff

$$\forall x, \forall y, P(X = x \cap Y = y) = P(X = x)P(Y = y)$$

Lemma 1: If $X \perp Y$, $E[XY] = E[X]E[Y]$.

Proof:

Let $p(x, y) = P(X = x \cap Y = y)$
($p(x, y)$ is called the **joint pmf**, but we will not be learning about this further).

$$E[XY] = \sum_x \sum_y xyp(x, y) \ [by\ definition]$$
$$= \sum_x \sum_y xyp(x)p(y) \ [independence]$$
$$= \sum_x xp(x) \sum_y yp(y) = E[X]E[Y]$$

Q.E.D.

Definition: Let X, Y be r.v.s. Let $\mu_X = E[X]$, $\mu_Y = E[Y]$. We define the **covariance of *X* and *Y*** as $Cov(X, Y) = E[XY] - \mu_X \mu_Y$.

Lemma 2: If $X \perp Y$, $Cov(X, Y) = 0$. Result follows immediately from previous proof.

Note: We just discussed the ***covariance of X and Y,*** *(look similar to formula for variance?). We will not be studying covariance in this class. In fact, you don't even need to know what it is for this class. This is just in case you're interested.*

Theorem: $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$.

Proof:

Let $\mu_X = E[X]$, $\mu_Y = E[Y]$.

$$\begin{aligned} Var(X+Y) &= E[(X+Y)^2] - (E[X+Y])^2 \; [by\ definition] \\ &= E[X^2 + 2XY + Y^2] - (\mu_X + \mu_Y)^2 \\ &= E[X^2] + \mathbf{2E[XY]} + E[Y^2] - \mu_X^2 - \mathbf{2\mu_X\mu_Y} - \mu_Y^2 \\ &= E[X^2] - \mu_X^2 + E[Y^2] - \mu_Y^2 + \mathbf{2Cov(X,Y)} \\ &= E[X^2] - E^2[X] + E[Y^2] - E^2[Y] + 2Cov(X,Y) \\ &= Var(X) + Var(Y) + 2Cov(X,Y) \end{aligned}$$

Q.E.D.

Theorem (Linearity of Variance): If $X \perp Y$, $Var(X+Y) = Var(X) + Var(Y)$. Proof is trivial based on Lemma 2 and previous theorem.

Note: If $X \perp Y$, $Var(X-Y) = Var(X) + Var(-Y) = Var(X) + (-1)^2 Var(Y) = Var(X) + Var(Y)$.

<u>Note</u>: As I said we will not use covariance in this class, so you will not be asked to find $Var(X+Y)$ unless $X \perp Y$. You will also not need to know what a joint pmf is. You can forget about covariance and joint pmfs for the rest of this quarter.

<u>Summary:</u>

- $Var(X) = E[X^2] - E^2[X]$
  - Common mistakes: not doing it this way (using the first definition instead)
- $E[f(X)] = \sum_x f(x)p(x) \neq f(E[X])$
- $Var(aX+b) = a^2 Var(X)$
  - Common mistakes: not squaring
  - Note: does not depend on b
- Independence ($X \perp Y$) iff
  $\forall x, \forall y, P(X = x \cap Y = y) = P(X = x)P(Y = y)$
- Linearity of Variance
  - If $X \perp Y$, $Var(X+Y) = Var(X) + Var(Y)$
    - Common mistakes: people forget they are allowed to do this
    - $Var(X-Y) \neq Var(X) - Var(Y)$
    - $Var(aX+bY+c) = a^2 Var(X) + b^2 Var(Y)$
- Linearity of Expectation **<u>always</u>** holds, while Linearity of Variance **<u>only</u>** holds with independence

Definition: $E[X|A] = \sum_x xp(x|A)$

Recall (Law of Total Probability):
$P(X) = P(X|A)P(A) + P(X|A^C)P(A^C)$

Proof of the Law of Total Expectation:

$$\begin{aligned}
E[X] &= \sum_x xp(x) \\
&= \sum_x x(p(x \mid A)P(A) + p(x \mid \overline{A})P(\overline{A})) \\
&= \sum_x xp(x \mid A)P(A) + \sum_x xp(x \mid \overline{A})P(\overline{A}) \\
&= \left(\sum_x xp(x \mid A)\right) P(A) + \left(\sum_x xp(x \mid \overline{A})\right) P(\overline{A}) \\
&= E[X \mid A]P(A) + E[X \mid \overline{A}]P(\overline{A})
\end{aligned}$$

The Discrete Uniform r.v.

Denoted: $Unif(a,b)$
Parameters: $a, b \in \mathbb{Z}, a \leq b$
Example: Roll of a fair die $\sim Unif(1,6)$
PMF: $p(x) = \frac{1}{b-a+1}$ $if\ a \leq x \leq b, 0\ otherwise$
$E[X] = \frac{a+b}{2}$
$Var(X) = \frac{(b-a)(b-a+2)}{12}$