Financial Mathematics
Basics of piecewise constant random variables
Definition: 
A **piecewise constant random variable** is a piecewise constant function $[0, 1] \rightarrow \mathbb{R}$. 

**PCRV: Finitely many pieces.**

**e.g.:** Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 \leq \omega \leq 1.00 
\end{cases}$$

Five "pieces"
e.g.: Let \( X : [0, 1] \to \mathbb{R} \) be defined by

\[
X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 \leq \omega \leq 1.00
\end{cases}
\]

Intuition:

On \( [0, 1] \), \( X(\omega) \) picks a point \( \omega \) in each interval, \( \Omega \).
e.g.: Let \( X : [0, 1] \rightarrow \mathbb{R} \) be defined by

\[
X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 \leq \omega \leq 1.00 
\end{cases}
\]

\[
\begin{array}{c|c}
0.22 - 0.00 &= 0.22 \\
0.54 - 0.22 &= 0.32 \\
0.65 - 0.54 &= 0.11 \\
0.99 - 0.65 &= 0.34 \\
1.00 - 0.99 &= 0.01 
\end{array}
\]

Intuition:
On each “trial”, Tyche picks a point \( \omega \in \Omega := [0, 1] \) at random, and reports back \( X(\omega) \) to us.

Question: What is the probability that \( 1 < X < 2 \)?
e.g.: Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 1, & \text{if } 0.00 \leq \omega < 0.22 \quad 0.22 - 0.00 = 0.22 \\ 2, & \text{if } 0.22 \leq \omega \leq 0.54 \quad 0.54 - 0.22 = 0.32 \\ -7, & \text{if } 0.54 < \omega \leq 0.65 \quad 0.65 - 0.54 = 0.11 \\ 8, & \text{if } 0.65 < \omega < 0.99 \quad 0.99 - 0.65 = 0.34 \\ 0, & \text{if } 0.99 \leq \omega \leq 1.00 \quad 1.00 - 0.99 = 0.01 \end{cases}$$

Pr[$1 < X < 2$] = 0
Pr[$1 \leq X < 2$] = 0.22 = 22%
Pr[$1 \leq X \leq 2$] = 0.32 = 32%
Pr[$1 < X \leq 2$] = 0.22 + 0.32 = 54%
Pr[$X = 0$] = 0.01 = 1%

**Question:** What is the probability that $1 < X < 2$? **Ans:** 0
Definition: A PCRV is **deterministic** if it’s constant, except (possibly) at finitely many points.

e.g.:
Let $U : \Omega \rightarrow \mathbb{R}$ be defined by $U(\omega) = 27$.

e.g.:
Let $V : \Omega \rightarrow \mathbb{R}$ be defined by $V(\omega) = 12$, except $V(0.75) = 6$ and $V(1) = 3$.

Note: “Pieces” can have zero length.
Let $Y : [0, 1] \rightarrow \mathbb{R}$ be defined by

\[
Y(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega < 0.54 \\
-7, & \text{if } \omega = 0.54 \\
8, & \text{if } 0.54 < \omega < 1 \\
0, & \text{if } \omega = 1.00
\end{cases}
\]

$\Pr[Y > 0] = 1$, so we say:

$Y > 0$ \textbf{almost surely} (a.s.)

It's not true that $Y > 0$ surely.

It's true that $Y \geq -7$ surely.
Modeling two coin flips

\[ C_1 := \begin{cases} 
1, & \text{if first flip is heads} \\
-1, & \text{if first flip is tails} 
\end{cases} \]

\[ C_2 := \begin{cases} 
1, & \text{if second flip is heads} \\
-1, & \text{if second flip is tails} 
\end{cases} \]
Modeling two coin flips

\[ C_1 := \begin{cases} 
1, & \text{if first flip is heads} \\
-1, & \text{if first flip is tails} 
\end{cases} \]

\[ C_2 := \begin{cases} 
1, & \text{if second flip is heads} \\
-1, & \text{if second flip is tails} 
\end{cases} \]
The distribution of a PCRV

Definition:
Let $S : [0, 1] \rightarrow \mathbb{R}$ be a PCRV.
Let $F := \{a \in \mathbb{R} \mid \Pr[S = a] > 0\}$
The distribution of $S$ associates to any $a \in F$,
the value $\Pr[S = a]$.

Note:
Can be thought of as a function $F \rightarrow (0, 1)$,
or as a “measure” on $\mathbb{R}$,
which is “supported” on $F$. 
e.g.: Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 \leq \omega \leq 1.00
\end{cases}$$

The distribution of $X$ is:

1 | 0.22
2 | 0.32
-7 | 0.11
8 | 0.34
0 | 0.01
e.g.: Let $Y : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$Y(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega < 0.54 \\
-7, & \text{if } \omega = 0.54 \\
8, & \text{if } 0.54 < \omega < 1 \\
0, & \text{if } \omega = 1.00 
\end{cases}$$

The distribution of $Y$ is:

- $1$ with probability $0.22$
- $2$ with probability $0.32$
- $8$ with probability $0.46$
Note: $C_1$ and $C_2$ are identically distributed, but are not equal.
Def’n: Let $S$ and $T$ be PCRVs. 
Let $F := \{(a, b) \in \mathbb{R}^2 | \Pr[(S = a) \& (T = b)] > 0\}$. 
The **joint distribution of** $(S, T)$ 
associates, to each element $(a, b) \in F$, 
the value $\Pr[(S = a) \& (T = b)]$.

Note: Can be thought of 
as a function $F \rightarrow (0, \infty)$, 
or as a “measure” on $\mathbb{R}^2$, 
which is “supported” on $F$.

Remark: To compute the distribution of $S + T$, 
you need to know the JOINT distr. of $(S, T)$. 
Knowing both the distribution of $S$ 
and the distribution of $T$ 
is insufficient. Same for $ST$. 

Def’n: Let $S$ and $T$ be PCRVs.
Let $F := \{(a, b) \in \mathbb{R}^2 \mid \Pr[(S = a) \& (T = b)] > 0\}$.

The **joint** distribution of $(S, T)$ associates, to each element $(a, b) \in F$, the value $\Pr[(S = a) \& (T = b)]$.

**e.g.** Let $A := C_1$, $B := C_2$, $A' := C_1$, $B' := C_1$.
Then $A$ and $A'$ have the same distribution, and $B$ and $B'$ have the same distribution, but $A + B$ and $A' + B'$ do not have the same distribution.

**Note:** $(A, B)$ and $(A', B')$ do not have the same joint distribution.

**Remark:** To compute the distribution of $S + T$, you need to know the **JOINT distr.** of $(S, T)$. Knowing both the distribution of $S$ and the distribution of $T$ is insufficient. Same for $ST$. 
Let $A$ be a PCRV, so $A : [0, 1] \to \mathbb{R}$ is piecewise constant. Let $f : \mathbb{R} \to \mathbb{R}$ be a function.

**Definition:** $f(A) := f \circ A : [0, 1] \to \mathbb{R}$

**Note:** $f(A)$ is a PCRV as well.

**e.g.:**

\[
A(\omega) = \begin{cases} 
4, & \text{if } 0.00 \leq \omega < 0.45 \\
2, & \text{if } 0.45 \leq \omega \leq 0.75 \\
-8, & \text{if } 0.75 < \omega \leq 1
\end{cases}
\]

\[
f(x) = x^2
\]

Then $B(\omega) = \begin{cases} 
16, & \text{if } 0.00 \leq \omega < 0.45 \\
4, & \text{if } 0.45 \leq \omega \leq 0.75 \\
64, & \text{if } 0.75 < \omega \leq 1
\end{cases}$
Definition: For any PCRV \( T : \hat{\Omega} \rightarrow \mathbb{R} \),

the **mean** of \( T \) is \( \mathbb{E}[T] := \int_0^1 T(\omega) \, d\omega \).

Note: \( \mathbb{E}[\bullet] \) is linear,

i.e., \( \mathbb{E}[S + T] = (\mathbb{E}[S]) + (\mathbb{E}[T]) \)
and \( \mathbb{E}[cS] = c(\mathbb{E}[S]). \)

For any PCRV \( T \), let \( T^\circ := T - (\mathbb{E}[T]). \)

For any PRCV \( T \), the **variance** of \( T \) is

\( \text{Var}[T] := \mathbb{E}[(T^\circ)^2]. \)

Remark: For any PCRV \( T \), \( \text{Var}[T] \geq 0. \)

\( T \) is deterministic iff \( \text{Var}[T] = 0. \)

Fact: \( \text{Var}[T] = (\mathbb{E}[T^2]) - (\mathbb{E}[T])^2 \)
Pf of fact: $\mu := \mathbb{E}[T]$  
$T^o = T - \mu$

\[ \text{Var}[T] = \mathbb{E}[(T^o)^2] = \mathbb{E}[(T - \mu)^2] = \mathbb{E}[T^2 - 2\mu T + \mu^2] \]

Note: $\mathbb{E}[\bullet]$ is linear, 
\[ i.e., \mathbb{E}[S + T] = (\mathbb{E}[S]) + (\mathbb{E}[T]) \] 
and $\mathbb{E}[cS] = c(\mathbb{E}[S])$.

For any PCRV $T$, let $T^o := T - (\mathbb{E}[T])$.

For any PRCV $T$, the variance of $T$ is 
$\text{Var}[T] := \mathbb{E}[(T^o)^2]$. 

Remark: For any PCRV $T$, $\text{Var}[T] \geq 0$. 
$T$ is deterministic iff $\text{Var}[T] = 0$.

Fact: $\text{Var}[T] = (\mathbb{E}[T^2]) - (\mathbb{E}[T])^2$
Pf of fact: \( \mu := \mathbb{E}[T] \quad T^\circ = T - \mu \)

\[
\text{Var}[T] = \text{Var}[\hat{T}] - 2\mu T + \mu^2 \\
= (E[T^2]) - 2\mu(E[T]) + \mu^2 - 2\mu T + \mu^2
\]

Note: \( E[\bullet] \) is linear,

i.e., \( E[S + T] = (E[S]) + (E[T]) \)

and \( E[cS] = c(E[S]) \).

For any PCRV \( T \), let \( T^\circ := T - (E[T]) \).

For any PRCV \( T \), the **variance of** \( T \) is

\[
\text{Var}[T] := E[(T^\circ)^2] 
\]

Remark: For any PCRV \( T \), \( \text{Var}[T] \geq 0 \).

\( T \) is deterministic iff \( \text{Var}[T] = 0 \).

Fact: \( \text{Var}[T] = (E[T^2]) - (E[T])^2 \)
Pf of fact: $\mu := \mathbb{E}[T]$ \hspace{1cm} $T^\circ = T - \mu$

\[
\text{Var}[T] = \mathbb{E}[T^2 - 2\mu T + \mu^2] = (\mathbb{E}[T^2]) - 2\mu(\mathbb{E}[T]) + \mu^2 = (\mathbb{E}[T^2]) - \mu^2
\]

Note: $\mathbb{E}[\bullet]$ is linear, \hspace{1cm} i.e., $\mathbb{E}[S + T] = (\mathbb{E}[S]) + (\mathbb{E}[T])$ and $\mathbb{E}[cS] = c(\mathbb{E}[S])$.

For any PCRVT $T$, let $T^\circ := T - (\mathbb{E}[T])$.

For any PRCV $T$, the \textbf{variance of $T$} is \hspace{1cm} $\text{Var}[T] := \mathbb{E}[(T^\circ)^2]$.

Remark: For any PCRVT $T$, $\text{Var}[T] \geq 0$. $T$ is deterministic iff $\text{Var}[T] = 0$.

Fact: $\text{Var}[T] = (\mathbb{E}[T^2]) - (\mathbb{E}[T])^2$
Pf of fact: \( \mu := \mathbb{E}[T] \quad T^0 = T - \mu \)

\[ \text{Var}[T] = (\mathbb{E}[T^2]) - \mu^2 \]

\[ = (\mathbb{E}[T^2]) - \mu^2 \]

Note: \( \mathbb{E}[\bullet] \) is linear, i.e., \( \mathbb{E}[S + T] = (\mathbb{E}[S]) + (\mathbb{E}[T]) \)
and \( \mathbb{E}[cS] = c(\mathbb{E}[S]). \)

For any PCRV \( T \), let \( T^0 := T - (\mathbb{E}[T]). \)

For any PRCV \( T \), the variance of \( T \) is \( \text{Var}[T] := \mathbb{E}[(T^0)^2] \).

Remark: For any PCRV \( T \), \( \text{Var}[T] \geq 0. \)
\( T \) is deterministic iff \( \text{Var}[T] = 0. \)

Fact: \( \text{Var}[T] = (\mathbb{E}[T^2]) - (\mathbb{E}[T])^2 \)
Pf of fact: \( \mu := E[T] \quad T^o = T - \mu \)

\[
\text{Var}[T] = (E[T^2]) - \mu^2
= (E[T^2]) - (E[T])^2
\quad \text{QED}
\]

Note: \( E[\bullet] \) is linear,
i.e., \( E[S + T] = (E[S]) + (E[T]) \)
and \( E[cS] = c(E[S]) \).

For any PCRV \( T \), let \( T^o := T - (E[T]) \).
For any PRCV \( T \), the variance of \( T \) is \( \text{Var}[T] := E[(T^o)^2] \).

Remark: For any PCRV \( T \), \( \text{Var}[T] \geq 0 \).
\( T \) is deterministic iff \( \text{Var}[T] = 0 \).

Fact: \( \text{Var}[T] = (E[T^2]) - (E[T])^2 \)
Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 < \omega \leq 1.00
\end{cases}$$

The expectation or mean of $X$ is

$$E[X] := \int_{0}^{1} X(\omega) \, d\omega$$

$$= 1(0.22) + 2(0.32) - 7(0.11) + 8(0.34) + 0(0.01)$$

$$= 2.81$$

Intuition: Measure of (average) size.
Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 < \omega \leq 1.00
\end{cases}$$

$E[X] = 2.81$

$\text{Var}[U] := E[(U^o)^2]$

$E[X] = 2.81$

Intuition: Measure of (average) size.
Let $X : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$X(\omega) = \begin{cases} 
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 < \omega \leq 1.00 
\end{cases}$$

$E[X] = 2.81$

The variance of $X$ is

$$\text{Var}[X] := E[(X^o)^2] = E[(X - (E[X]))^2]$$

$$= (1 - 2.81)^2(0.22) + (2 - 2.81)^2(0.32) + (-7 - 2.81)^2(0.11) + (8 - 2.81)^2(0.34) + (0 - 2.81)^2(0.01) = \text{Exercise}$$

Intuition: Measure of risk.
Let \( X : [0, 1] \rightarrow \mathbb{R} \) be defined by

\[
X(\omega) = \begin{cases}
1, & \text{if } 0.00 \leq \omega < 0.22 \\
2, & \text{if } 0.22 \leq \omega \leq 0.54 \\
-7, & \text{if } 0.54 < \omega \leq 0.65 \\
8, & \text{if } 0.65 < \omega < 0.99 \\
0, & \text{if } 0.99 < \omega \leq 1.00
\end{cases}
\]

\[
\mathbb{E}[X] = 2.81
\]

Key idea: “Most investors are return-loving, but risk-averse.”

If \( X \) is the price, one month from now, of some financial asset, then investors typically hope for \( X \) to have large mean and small variance.
Fact: \( \mathbb{E}[\bullet] \) is linear, i.e., \( \mathbb{E}[S + T] = (\mathbb{E}[S]) + (\mathbb{E}[T]) \) and \( \mathbb{E}[cS] = c(\mathbb{E}[S]) \).

**WARNING:** \( \text{Var}[\bullet] \) is **NOT** linear, but rather quadratic.

\[
(2S)^\circ = 2S - (\mathbb{E}[2S]) = 2S - 2(\mathbb{E}[S]) = 2(S - (\mathbb{E}[S])) = 2(S^\circ) \]

\[
\mathbb{U}^\circ := U - (\mathbb{E}[U])
\]

\[
\mathbb{V}
\]

\[
\mathbb{V}[U] := \mathbb{E}[(U^\circ)^2]
\]

\[
\mathbb{V}[2S] = \mathbb{E}[(2S)^\circ]^2 = \mathbb{E}[(2(S^\circ))^2] = 4(\mathbb{E}[(S^\circ)^2]) = 4(\mathbb{V}[S])
\]
Fact: $E[\bullet]$ is linear,
i.e., $E[S + T] = (E[S]) + (E[T])$
and $E[cS] = c(E[S])$.

**WARNING:** $\text{Var}[\bullet]$ is **NOT** linear,
but rather quadratic.

$\text{Cov}[S, T]$ is defined by:

$$\text{Var}[S + T] = (\text{Var}[S]) + (\text{Var}[T])$$

$$+ 2(\text{Cov}[S, T])$$

**Cauchy-Schwarz:** $-1 \leq \frac{\text{Cov}[S, T]}{(\sqrt{\text{Var}[S]} \sqrt{\text{Var}[T]})} \leq 1$

**WARNING:**
$\text{Corr}[S, T]$ is not defined if $S$ or $T$
is deterministic.

**Def'n:** $S$ and $T$ are **uncorrelated**
if $\text{Cov}[S, T] = 0$. 


Fact: $E[\bullet]$ is linear, i.e., $E[S + T] = (E[S]) + (E[T])$
and $E[cS] = c(E[S])$.

**WARNING:** Var[$\bullet$] is **NOT** linear, but rather quadratic.

**Cov**[$S, T$] is defined by:

$\text{Var}[S + T] = (\text{Var}[S]) + (\text{Var}[T]) + 2(\text{Cov}[S, T])$

Definition:

$S$ and $T$ are **uncorrelated** if $\text{Cov}[S, T] = 0$, i.e., if $\text{Var}[S + T] = (\text{Var}[S]) + (\text{Var}[T])$.

Definition: $S$ and $T$ are **uncorrelated** if $\text{Cov}[S, T] = 0$. 

29
Fact: $E[\bullet]$ is linear, 
\[ i.e., \ E[S + T] = (E[S]) + (E[T]) \] 
and $E[cS] = c(E[S])$.

**WARNING:** $Var[\bullet]$ is **NOT** linear, 
but rather quadratic.

**(Cov)** $[S, T]$ is defined by:
\[ Var[S + T] = (Var[S]) + (Var[T]) \]
\[ + 2(Cov[S, T]) \]

**Definition:**
$S$ and $T$ are **uncorrelated** if $Cov[S, T] = 0$, 
\[ i.e., \ Var[S + T] = (Var[S]) + (Var[T]), \]
\[ i.e., \ E[ST] = (E[S])(E[T]). \]

**Fact:** $Var[T] = (E[T^2]) - (E[T])^2$

**Fact:** $Cov[S, T] = (E[ST]) - (E[S])(E[T])$
Fact: If \( T = S + 3 \), then \( \text{Var}[T] = \text{Var}[S] \).

Proof: \( \text{E}[T] = (\text{E}[S]) + 3 \)

\[
T^\circ = T - (\text{E}[T])
\]

\[
= (S + 3) - ((\text{E}[S]) + 3)
\]

\[
= S - (\text{E}[S])
\]

\[
= S^\circ
\]

\[
\text{Var}[U] := \text{E}[(U^\circ)^2]
\]

Definition:
\( S \) and \( T \) are **uncorrelated** if \( \text{Cov}[S, T] = 0 \), i.e., if \( \text{Var}[S + T] = (\text{Var}[S]) + (\text{Var}[T]) \), i.e., if \( \text{E}[ST] = (\text{E}[S])(\text{E}[T]) \).

Fact: \( \text{Var}[T] = (\text{E}[T^2]) - (\text{E}[T])^2 \)

Fact: \( \text{Cov}[S, T] = (\text{E}[ST]) - (\text{E}[S])(\text{E}[T]) \)
Fact: If $T = S + 3$, then $\text{Var}[T] = \text{Var}[S]$.  

Proof: $E[T] = (E[S]) + 3$

\[
U^{\circ} := U - (E[U])
\]

\[
T^{\circ} = T - (E[T])
\]

\[
= (S + \beta) - ((E[S]) + \beta)
\]

\[
= S - (E[S])
\]

\[
= S^{\circ}
\]

\[
\text{Var}[T] = E[(T^{\circ})^2] = E[(S^{\circ})^2] = \text{Var}[S] \quad \text{QED}
\]

Fact: If $T - S$ is constant, then $\text{Var}[T] = \text{Var}[S]$.  

Fact: If $T - S$ is deterministic, then $\text{Var}[T] = \text{Var}[S]$.  

32
Problem: Suppose $0 \leq p \leq 1$ and $a, b \in \mathbb{R}$.
Suppose $\Pr[S = b] = q := 1 - p$
and $\Pr[S = a] = p$.
Find $\text{Var}[S]$.

Solution: Let $T := S - a$.
Then $\text{Var}[S] = \text{Var}[T]$,
$\Pr[T = b - a] = q$
and $\Pr[T = 0] = p$.
Then $E[T] = q(b - a)$
Problem: Suppose $0 \leq p \leq 1$ and $a, b \in \mathbb{R}$.

Suppose $\Pr[S = b] = q := 1 - p$ and $\Pr[S = a] = p$.

Find $\text{Var}[S]$.

Solution: Let $T := S - a$.

Then $\text{Var}[S] = \text{Var}[T]$,

$\Pr[T = b - a] = q$

and $\Pr[T = 0] = p$.

Then

$\mathbb{E}[T] = q(b - a)$

and

$\mathbb{E}[T^2] = q(b - a)^2$.

Then

$\text{Var}[S] = \text{Var}[T]$

$= (\mathbb{E}[T^2]) - (\mathbb{E}[T])^2$

$= q(b - a)^2 - q^2(b - a)^2$

$= (q - q^2)(b - a)^2$

$= (1 - q)q(b - a)^2$.
Problem: Suppose $0 \leq p \leq 1$ and $a, b \in \mathbb{R}$. Suppose $\Pr[S = b] = q := 1 - p$ and $\Pr[S = a] = p$.

Find $\text{Var}[S]$.

Solution: $\text{Var}[S] = pq(b - a)^2$
Problem: Suppose $0 \leq p \leq 1$ and $a, b \in \mathbb{R}$.

Suppose $\Pr[S = b] = q := 1 - p$
and $\Pr[S = a] = p$.

Find $\text{Var}[S]$.

Solution: $\text{Var}[S] = pq(b - a)^2$.

*e.g.*, coin-flipping:

$\mathbb{E}[C] = 0$ and $\text{Var}[C] = 1$, i.e., $C$ is standard.

**Definition:** $S$ is **standard** means:

$\mathbb{E}[S] = 0$ and $\text{Var}[S] = 1$. 

36
Definition: **Standard deviation** := \( \sqrt{\text{Variance}} \)

\[ \text{SD}[S] := \sqrt{\text{Var}[S]} \]

Fact: \( \forall U, \quad U_\circ := \frac{U - (E[U])}{\text{SD}[U]} \) is standard.

A renormalization of \( U \)

non-deterministic

e.g., coin-flipping:

\[
\begin{array}{c c}
0.5 & 1 \\
0.5 & -1 \\
\end{array}
\]

\( E[C] = 0 \) and \( \text{Var}[C] = 1 \),

i.e., \( C \) is standard.

Definition: \( S \) is **standard** means:

\( E[S] = 0 \) and \( \text{Var}[S] = 1 \).
Definition: **Standard deviation** := $\sqrt{\text{Variance}}$

$$\text{SD}[S] := \sqrt{\text{Var}[S]}$$

Fact: $\forall U$, $U_\circ := \frac{U - (E[U])}{\text{SD}[U]}$ is standard.

Pf: exercise

A renormalization of $U$

non-deterministic

$\mu = E[U]$

$\sigma = \text{SD}[U]$

$U_\circ = \frac{U - \mu}{\sigma} \Rightarrow U = \sigma U_\circ + \mu$

Any (non-deterministic) $U$ is “almost” standard.

Definition: $S$ is **standard** means:

$$E[S] = 0 \quad \text{and} \quad \text{Var}[S] = 1.$$
Definition: Standard deviation := \( \sqrt{\text{Variance}} \)

\[
\text{SD}[S] := \sqrt{\text{Var}[S]}
\]

\[
(2S)^\circ = 2S - (\text{E}[2S]) = 2S - 2(\text{E}[S])
= 2(S - (\text{E}[S])) = 2(S^\circ)
\]

\[
\text{U}^\circ := U - (\text{E}[U])
\]

\[
\text{Var}[2S] = \text{E}[((2S)^\circ)^2] = \text{E}[(2(S^\circ))^2]
= \text{E}[4(S^\circ)^2] = 4(\text{E}[(S^\circ)^2]) = 4(\text{Var}[S])
\]

\[
\text{SD}[2S] = 2(\text{SD}[S])
\]

Intuition: Variance measures risk, but standard deviation measures risk better, because doubling the position really ought only to double the risk, not quadruple it.
Definition: **Standard deviation** := \( \sqrt{\text{Variance}} \)

\[
\text{SD}[S] := \sqrt{\text{Var}[S]}
\]

Remark: identically distributed \( \Rightarrow \)

same mean,
same variance,
same standard deviation.

If I hold a portfolio with variance 4, and if I give you half of each asset, then we both hold portfolios of variance 1.

Has our risk gone from 4 to 1 + 1? **NO!**

Intuition: Variance measures risk, **but** standard deviation measures risk better, because doubling the position really ought only to double the risk, not quadruple it.
Definition: **Standard deviation** := $\sqrt{\text{Variance}}$

$$\text{SD}[S] := \sqrt{\text{Var}[S]}$$

Remark: identically distributed $\Rightarrow$

same mean,  
same variance,  
same standard deviation.

**WARNING**: identical JOINT distribution is needed for same covariance, same correlation.

**Intuition**: Variance measures risk, but standard deviation measures risk better, because doubling the position really ought only to double the risk, not quadruple it.
Definition: **Standard deviation** \( \text{SD}[S] := \sqrt{\text{Var}[S]} \)

Problem: Suppose \( 0 \leq p \leq 1 \) and \( a \leq b \).

Suppose \( \Pr[S = b] = q := 1 - p \) and \( \Pr[S = a] = p \).

Find \( \text{SD}[S] \).

Solution: \( \text{Var}[S] = pq(b - a)^2 \), so \( \text{SD}[S] = \sqrt{pq(b - a)} \).

The geometric mean of \( p \) and \( q \) is \( \sqrt{pq} \).

The arithmetic mean of \( p \) and \( q \) is \( \frac{p + q}{2} \).

Next topic: probabilities are expectations.
Probabilities are expectations!

**Definition:** Let $A \subseteq B$.

The **indicator function** of $A$ (in $B$) is the function $1^B_A : B \to \{0, 1\}$ defined by

$$1^B_A(x) = \begin{cases} 
1, & \text{if } x \in A \\
0, & \text{if } x \in B\setminus A. 
\end{cases}$$

**Fact:**

If $f$ is the indicator function of $S$ in $\mathbb{R}$, then $\mathbb{E}[f(X)] = \Pr[X \in S]$.

**Pf:**

$$f(X) = f \circ X : \Omega \to [0, 1]$$
Probabilities are expectations!  

Definition: Let $A \subseteq B$. The **indicator function** of $A$ (in $B$) is the function $1^B_A : B \to \{0, 1\}$ defined by

$$1^B_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \in B \setminus A. \end{cases}$$

Fact: If $f$ is the indicator function of $S$ in $\mathbb{R}$, then $\mathbb{E}[f(X)] = \Pr[X \in S]$. 

Pf: $\Pr[f(X) = 1]$ 

\[f(X) = 1 \text{ iff } X \in S\] 

QED 

\[f(X) \in \{0, 1\} \text{ a.s.}\] 

\[f(X) = f \circ X : \Omega \to \{0, 1\}\]
Probabilities are expectations!

**Definition:** Let \( A \subseteq B \).

The **indicator function** of \( A \) (in \( B \)) is the function \( 1^B_A : B \to \{0,1\} \) defined by

\[
1^B_A(x) = \begin{cases} 
1, & \text{if } x \in A \\
0, & \text{if } x \in B \setminus A.
\end{cases}
\]

**Fact:**

If \( f \) is the indicator function of \( S \) in \( \mathbb{R} \), then \( \mathbb{E}[f(X)] = \Pr[X \in S] \).

\[
\Pr[(f(X))(g(Y)) = 1] = \Pr[(f(X) = 1) \& (g(Y) = 1)] = \Pr[(X \in S) \& (Y \in T)] \quad \text{QED}
\]

**Fact:**
If \( f \) is the indicator function of \( S \) in \( \mathbb{R} \), and if \( g \) is the indicator function of \( T \) in \( \mathbb{R} \), then

\[
\mathbb{E}[(f(X))(g(Y))] = \Pr[(X \in S) \& (Y \in T)].
\]
Applications to finance: Reduction of risk
Let $X_1, \ldots, X_{100}$ be the dollar prices, one month from now, of 100 assets, all costing 1 dollar now.

Suppose, for all integers $j \in [1, 100]$, that $E[X_j] = 1.1$ and $SD[X_j] = 1$.

Let $S = X_1 + \cdots + X_{100}$. Cost: 100 dollars.

Then $E[S] = 110$.

If, $\forall$ integers $j, k \in [1, 100]$, $Corr[X_j, X_k] = 0$, then $SD[S] = \sqrt{100} = 10$.

If, $\forall$ integers $j, k \in [1, 100]$, $Corr[X_j, X_k] = 1$, then $SD[S] = 100$. 
Let $X_1, \ldots, X_N$ be the dollar prices, one month from now, of $N$ assets, all costing 1 dollar now.

Suppose we have 1 dollar to spend.

Suppose, $\forall$ integers $j \in [1, N]$, $\mathbb{E}[X_j]$ is known.

Suppose we want expected return to be $\rho$.

Suppose, $\forall$ integers $j, k \in [1, N]$, $\text{Cov}[X_j, X_k]$ is known.

**Goal of “Modern Portfolio Theory”:**

Choose $c_1, \ldots, c_N$ such that

$$\text{Var}[c_1 X_1 + \cdots + c_N X_N]$$

is minimized, subject to $\mathbb{E}[c_1 X_1 + \cdots + c_N X_N] = 1 + \rho$

and $c_1 + \cdots + c_N = 1$. 

Problem: We trade in two assets.

\( A := \text{return, one month from now, on the first asset} \)

\( B := \text{return, one month from now, on the second asset} \)

Assume \( \text{SD}[A] = 0.5, \)
\( \text{SD}[B] = 0.3 \)
and \( \text{Corr}[A, B] = 0.8. \)

We need to hold $10 of the first asset.
We want to short the second asset,
so as to reduce our risk as much as possible.

Find \( x \) such that \( \text{SD}[10A - xB] \) is minimized.

Sol’n: Same as minimizing \( (\text{SD}[10A - xB])^2 \),
i.e., \( \text{Var}[10A - xB] \)
\[ = \text{Var}[10A] + \text{Var}[-xB] + 2(\text{Cov}[10A, -xB]) \]
\[
\begin{align*}
\text{min: } & \text{Var}[10A] + \text{Var}[-xB] + 2(\text{Cov}[10A, -xB]) \\
\text{Assume } & \text{SD}[A] = 0.5, \\
& \text{SD}[B] = 0.3 \\
\text{and } & \text{Corr}[A, B] = 0.8.
\end{align*}
\]

We need to hold $10$ of the first asset. We want to short the second asset, so as to reduce our risk as much as possible. Find $x$ such that \text{SD}[10A - xB] is minimized.

\textbf{Sol’n:} Same as minimizing \((\text{SD}[10A - xB])^2\), i.e.,
\[
\begin{align*}
\text{Var}[10A - xB] &= \text{Var}[10A] + \text{Var}[-xB] \\
&+ 2(\text{Cov}[10A, -xB])
\end{align*}
\]
\[
\begin{align*}
\text{min: } & \operatorname{Var}[10A] + \operatorname{Var}[-xB] + 2(\operatorname{Cov}[10A, -xB]) \\
& 100(\operatorname{Var}[A]) + x^2(\operatorname{Var}[B]) - 20x(\operatorname{Cov}[A, B]) \\
& (0.5)^2 + (0.3)^2 - (0.8)(0.5)(0.3)
\end{align*}
\]

Assume \( \operatorname{SD}[A] = 0.5 \), \( \operatorname{SD}[B] = 0.3 \), and \( \operatorname{Corr}[A, B] = 0.8 \).

We need to hold $10 of the first asset.
We want to short the second asset, so as to reduce our risk as much as possible.
Find \( x \) such that \( \operatorname{SD}[10A - xB] \) is minimized.

Sol’n: Same as minimizing \( (\operatorname{SD}[10A - xB])^2 \),
i.e., \( \operatorname{Var}[10A - xB] \)
\[
= \operatorname{Var}[10A] + \operatorname{Var}[-xB] + 2(\operatorname{Cov}[10A, -xB])
\]
\[ \min: \text{Var}[10A] + \text{Var}[\neg xB] + 2(\text{Cov}[10A, \neg xB]) \]
\[ = \underbrace{100(\text{Var}[A])}_{(0.5)^2} + \underbrace{x^2(\text{Var}[B])}_{(0.3)^2} + \underbrace{-20x(\text{Cov}[A, B])}_{(0.8)(0.5)(0.3)} \]

\[ \min: \quad 25 + (0.09)x^2 - (2.4)x = 0 \]

\[ \text{deriv.:} \quad (0.18)x - (2.4) = 0 \]

\[ x = \frac{2.4}{0.18} = \frac{40}{3} = 13.33 \]

Stop $13.33$ of the second asset.

Find \( x \) such that \( \text{SD}[10A - xB] \) is minimized.

Sol’n: Same as minimizing \( (\text{SD}[10A - xB])^2 \), i.e.,
\[ \text{Var}[10A - xB] = \text{Var}[10A] + \text{Var}[\neg xB] \]