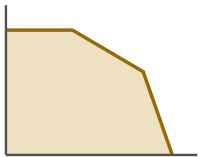
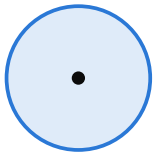
Decisions under uncertainty

208424 Lecture 11

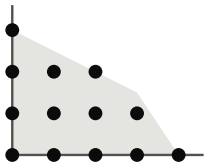
Lectures 8 to 10



LP



SOCP

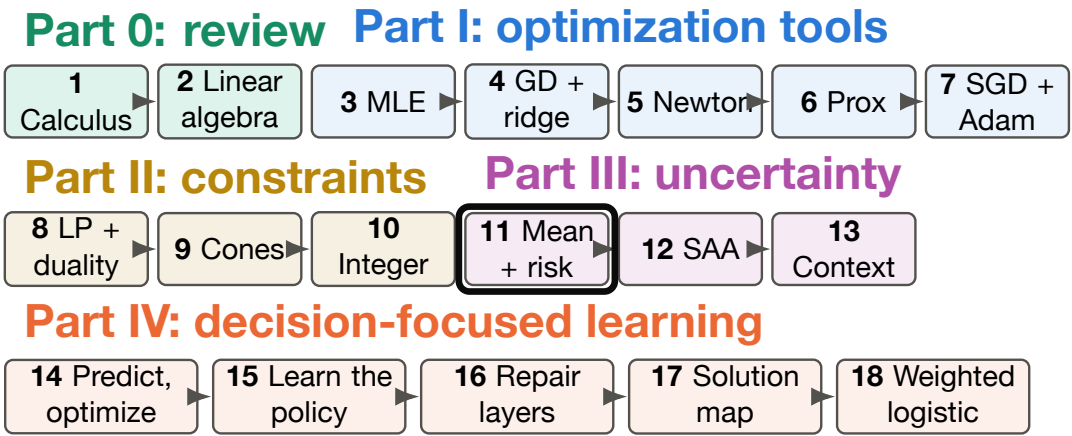


MILP

In Lectures 8 to 10, every number was known.

Today, the demand is random.

Course map


Part 0: review
Part I: optimization tools
1 Calculus
2 Linear algebra
3 MLE
4 GD + ridge
5 Newton
6 Prox
7 SGD + Adam
Part II: constraints
Part III: uncertainty
8 LP + duality
9 Cones
10 Integer
11 Mean + risk
12 SAA
13 Context
Part IV: decision-focused learning
14 Predict, optimize
15 Learn the policy
16 Repair layers
17 Solution map
18 Weighted logistic

Lecture 11: Decisions under uncertainty

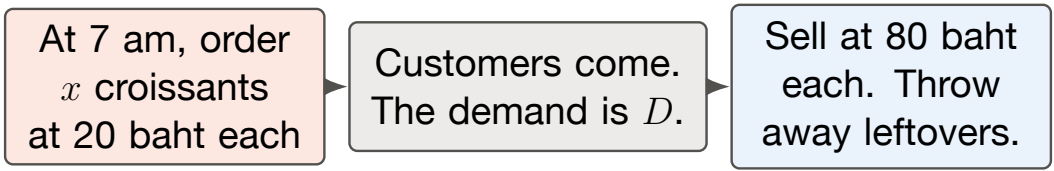


Photo: Thomon, CC BY-SA 4.0, Wikimedia Commons

A café orders croissants early in the morning.

It does not know the demand for that day.

A café on campus

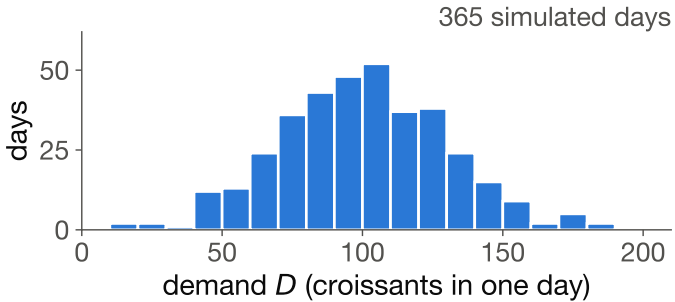


x is the number of croissants the café orders.

D is the number of croissants customers want.

How many croissants should the café order?

Demand changes every day

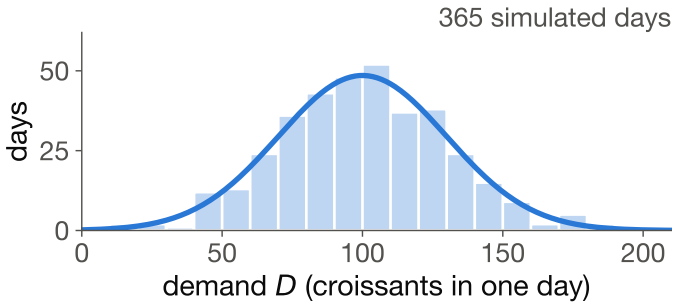


The demand D is different every day.

We do not know D before the day starts.

D is called a random variable.

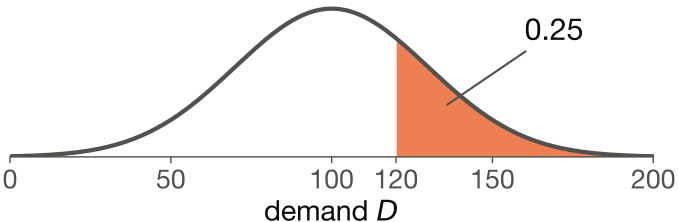
A normal distribution



We model D with a normal distribution.

Its mean is 100, and its SD is 30 (Lecture 5).

Probability as an area



$$P(D > 120) = 0.25$$

This is the probability that D is above 120.

In our data, 93 of the 365 days have $D > 120$.

Sales on one day

order $x = 120$; one dot is 10 croissants

$D = 90$ 

sell 90,
30 left over

$D = 150$ 

sell 120, 30 customers
get no croissant

The café sells the smaller of D and x.

We write this number as $\min(D, x)$.

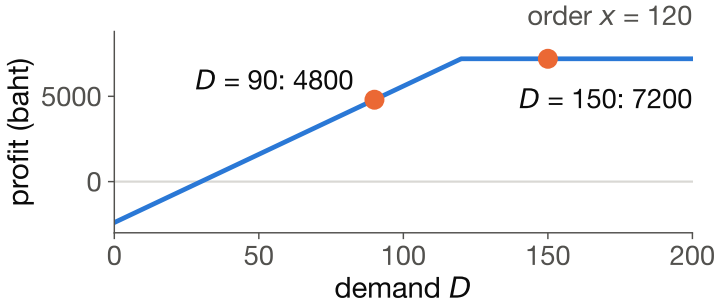
Profit on one day

$$\text{profit} = \underbrace{80 \min(D, x)}_{\text{money from sales (baht)}} \underbrace{- 20x}_{\text{cost of the order (baht)}}$$

The café gets 80 baht for each croissant sold.

It pays 20 baht for each croissant ordered.

Profit depends on demand

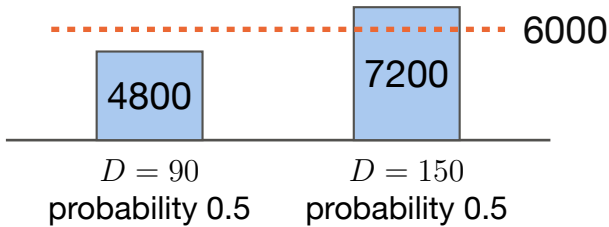


With $x = 120$, the profit is 4800 if $D = 90$,

and 7200 if $D = 150$.

So the profit is also a random variable.

The expected profit

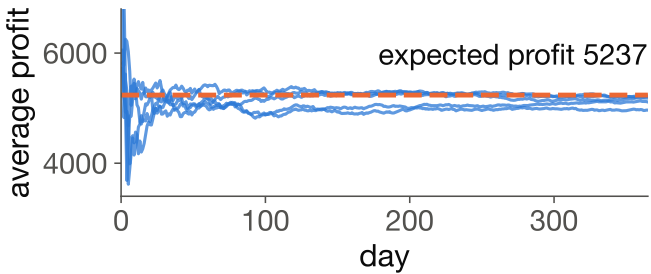


Suppose D takes the values 90 and 150.

$$\mathbb{E}[\text{profit}] = 0.5 \times 4800 + 0.5 \times 7200 = 6000$$

This weighted average is the expected profit.

Expected profit as a long-run average

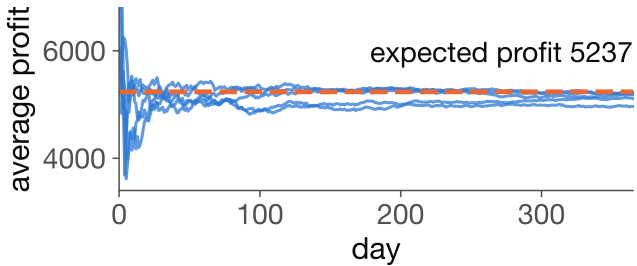


Each line is the average profit of one café.

Over many days, the average gets close to 5237.

5237 is the expected profit of ordering 120.

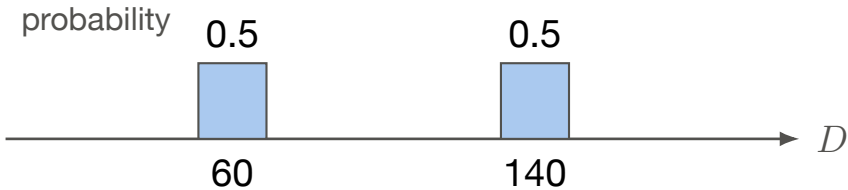
Summary: expected profit



$\mathbb{E}[\text{profit}]$ is the long-run average profit.

We want the order x with the largest $\mathbb{E}[\text{profit}]$.

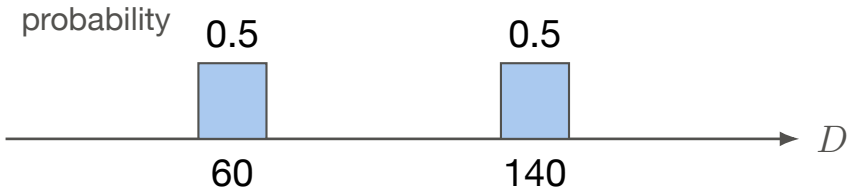
Checkpoint 1



D is 60 or 140, each with probability 0.5.

The café orders 100. What is $\mathbb{E}[\text{profit}]$?

Checkpoint 1



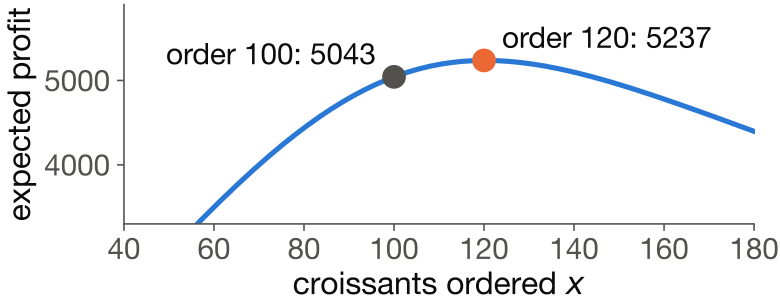
D is 60 or 140, each with probability 0.5.

The café orders 100. What is $\mathbb{E}[\text{profit}]$?

The profit is 2800 or 6000.

$$\mathbb{E}[\text{profit}] = 0.5 \times 2800 + 0.5 \times 6000 = 4400$$

The expected profit of each order



Ordering the mean demand, 100, gives 5043 baht.

The maximum is 5237 baht, at $x = 120$.

So the optimal order is more than the mean.

Too few or too many croissants

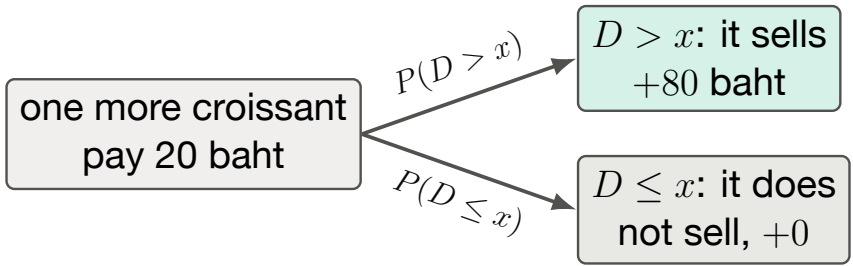
one croissant too few	one croissant too many
-60 baht	-20 baht

One croissant too few loses a profit of 60 baht.

One croissant too many loses its cost of 20 baht.

So the café should order more than the mean.

One more croissant

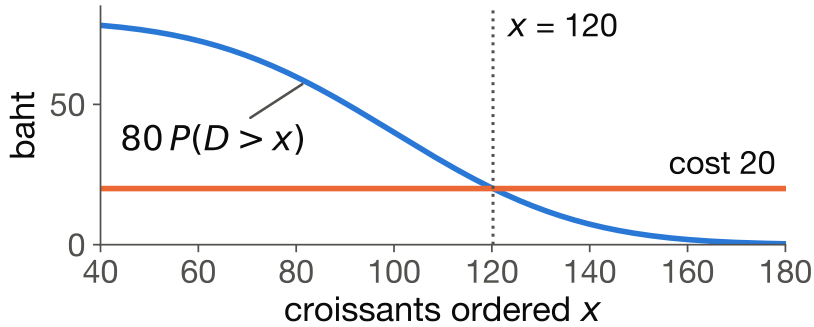


One more croissant costs 20 baht.

That croissant sells only if $D > x$.

expected extra profit $= 80\,P(D > x) - 20$

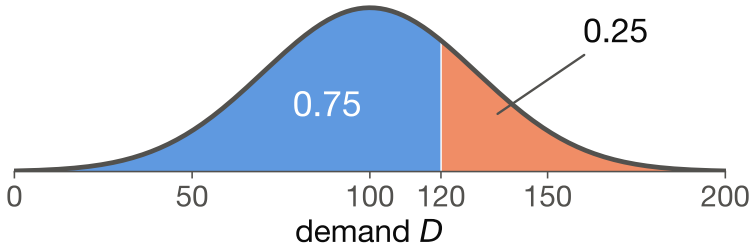
Where to stop



We add croissants while $80\,P(D > x) > 20$.

We stop at $x = 120$, where $80\,P(D > x) = 20$.

Enough croissants on 75% of days

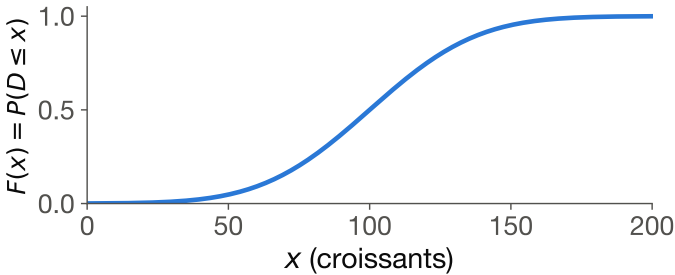


At $x = 120$, $P(D > x) = 20/80 = 0.25$.

So $P(D \leq 120) = 0.75$.

The café has enough croissants on 75% of days.

Cumulative distribution function (CDF)

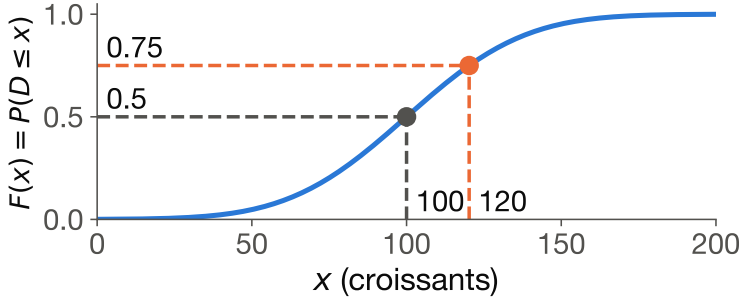


The CDF of D is $F(x) = P(D \leq x)$.

$F(x)$ is the probability that D is at most x.

For example, $F(120) = 0.75$.

Quantiles

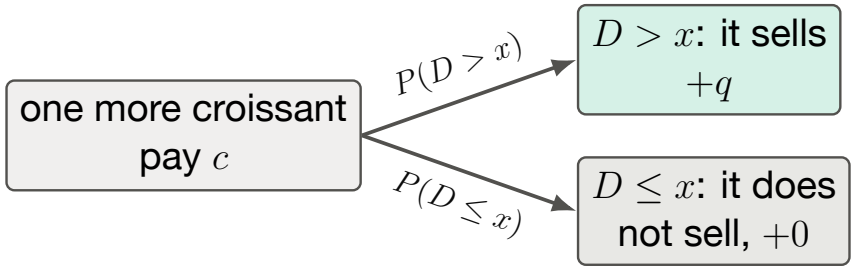


The 0.75 quantile of D is the x with $F(x) = 0.75$.

Since $F(120) = 0.75$, the 0.75 quantile is 120.

The 0.5 quantile, 100, is called the median.

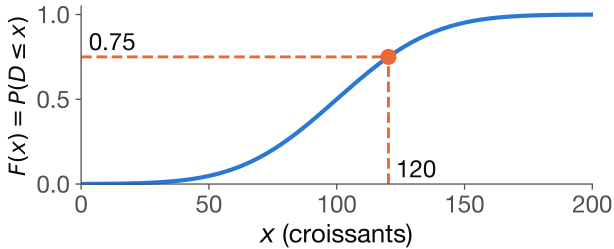
The optimal order in general



Let q be the price and c the cost.

The optimal order $x^\star$ satisfies $q\,P(D > x^\star) = c$.

The optimal order is a quantile



$$F(x^\star) = \frac{q - c}{q}$$

So $x^\star$ is the $(q - c)/q$ quantile of D.

Here $(80 - 20)/80 = 0.75$, so $x^\star = 120$.

The newsvendor problem



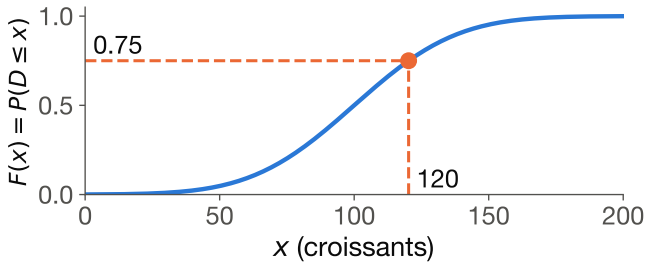
Photo: KF, public domain, Wikimedia Commons.

This problem is called the newsvendor problem.

A newsvendor buys newspapers each morning.

Unsold newspapers are worth nothing.

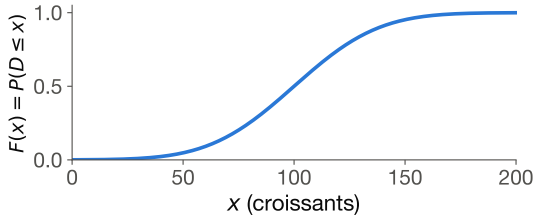
Summary: the newsvendor problem



The optimal order is the $(q - c)/q$ quantile.

Ordering the mean demand is not optimal.

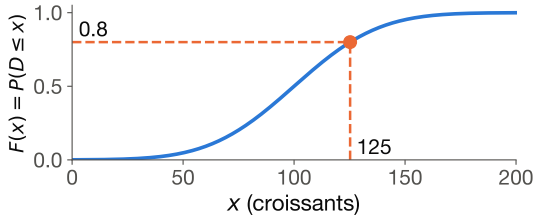
Checkpoint 2



The cost is 20 baht, and the price is 100 baht.

Which quantile of D should the café order?

Checkpoint 2



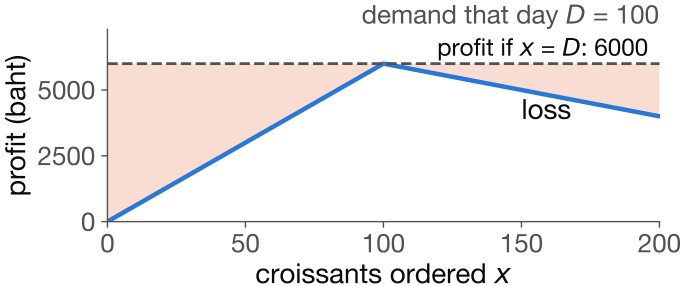
The cost is 20 baht, and the price is 100 baht.

Which quantile of D should the café order?

$(100 - 20)/100 = 0.8$, so the 0.8 quantile.

A higher price gives a larger order.

The loss of an order

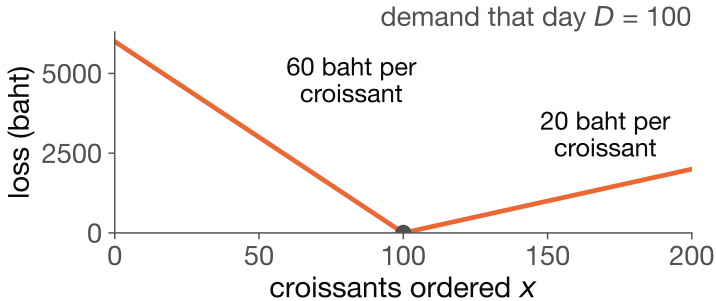


We can also find $x^\star$ by minimizing a loss.

If we knew D, we would order exactly D.

The loss is the profit lost by ordering x.

The pinball loss

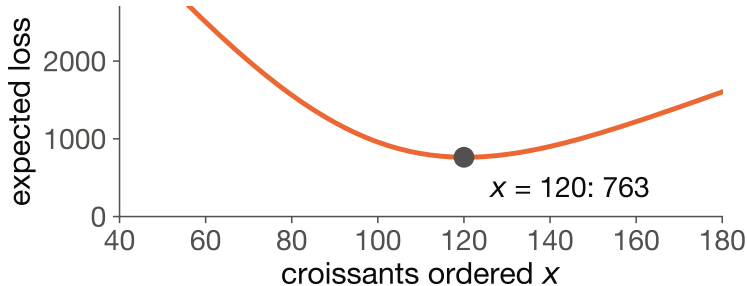


Each missing croissant adds 60 baht to the loss.

Each leftover croissant adds 20 baht to the loss.

This tilted V is called the pinball loss.

The expected loss

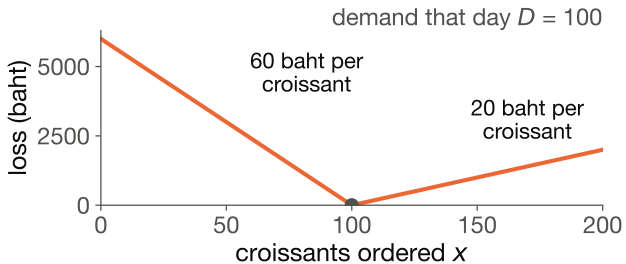


We choose x to minimize the expected loss.

The expected loss is smallest at $x = 120$.

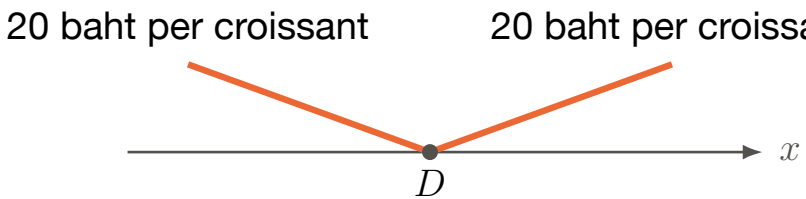
Again, the optimal order is the 0.75 quantile.

Summary: the pinball loss



Minimizing the expected pinball loss gives the $(q-c)/q$ quantile of D.

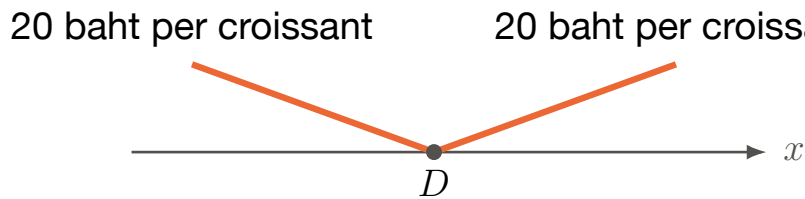
Checkpoint 3



D is still normal with mean 100 and SD 30.

Which order has the smallest expected loss?

Checkpoint 3



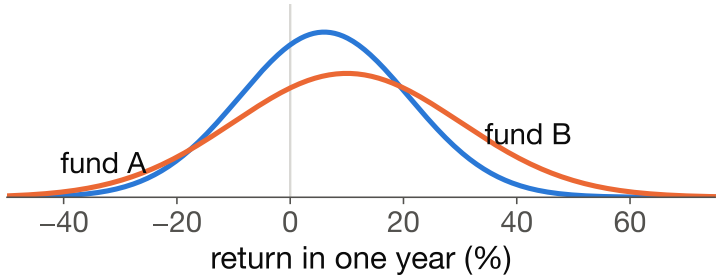
D is still normal with mean 100 and SD 30.

Which order has the smallest expected loss?

The V is symmetric, so $x^\star$ is the median, 100.

In Lecture 8, absolute errors also gave the median.

Two funds



We put savings in two funds for one year.

With a 6% return, 100 baht becomes 106 baht.

The return of each fund is a random variable.

Expected return and risk

	expected return	SD
• fund A	6%	15%
• fund B	10%	20%

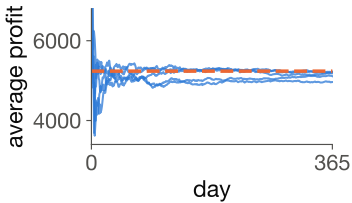
$\mathbb{E}[\text{return}]$ is the long-run average return.

Fund B has the higher $\mathbb{E}[\text{return}]$ and the larger SD.

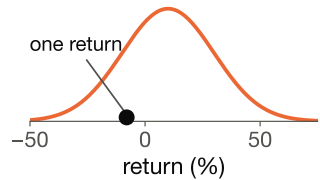
The SD of a return is called its risk.

Why the SD matters

the café: 365 days



the investor: one year

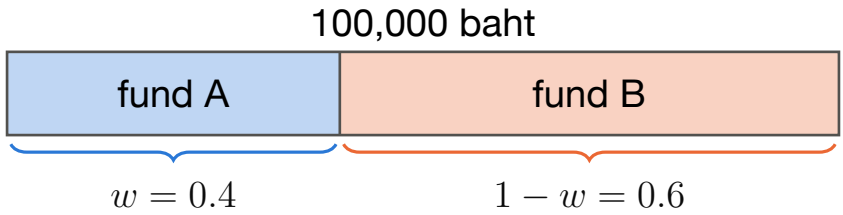


The café repeats its order every day.

An investor gets only one return this year.

So the investor also cares about the SD.

A portfolio



We put a share w of the money in fund A,

and the share $1 - w$ in fund B.

This is a portfolio with weight w on fund A.

The return of a portfolio

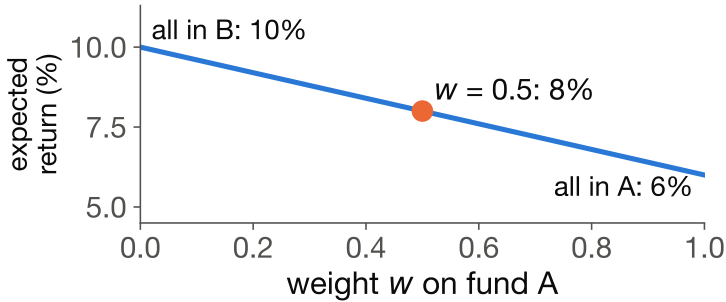
fund A, $w = 0.4$	fund B, $1 - w = 0.6$
10% return this year	5% return this year

$$0.4 \times 10\% + 0.6 \times 5\% = 7\%$$

The portfolio return is a weighted average.

The weights are w and $1 - w$.

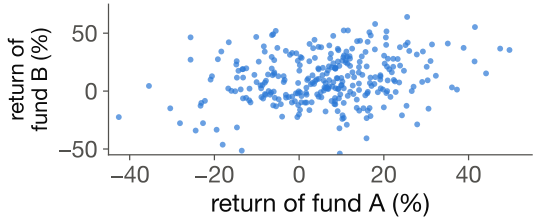
The expected return of a portfolio



expected return $= 6w + 10(1 - w)$

With $w = 0.5$, the expected return is 8%.

Covariance



The SD of a portfolio depends on the covariance.

The returns have correlation 0.2 (Lecture 9).

$$\text{covariance} = \underset{\text{correlation}}{0.2} \times \underset{\text{SD of A}}{15} \times \underset{\text{SD of B}}{20} = 60$$

The covariance matrix

$\Sigma =$

	A	B
A	225	60
B	60	400

The variance is the square of the SD.

The diagonal entries are the variances, 15^2 and 20^2.

Σ is called the covariance matrix.

The variance of a portfolio

$$\text{Var} = v^\top \begin{array}{cc} \text{A} & \text{B} \\ \hline 225 & 60 \\ 60 & 400 \end{array} v, \quad \text{with } v = (w,\, 1-w)$$

Lecture 9 used $v^\top C v$ for standardized sales.

For returns in percent, we use Σ.

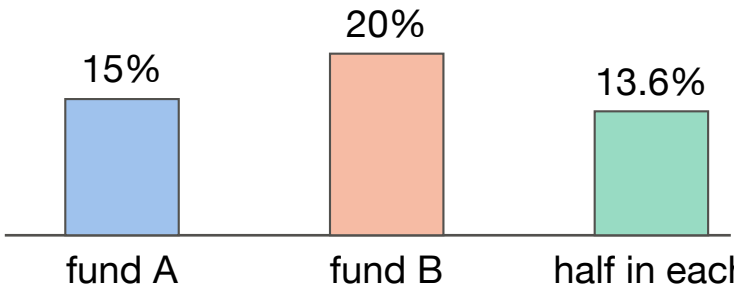
The variance with two funds

	w	$1-w$
w	$225\,w\,w$	$60\,w\,(1-w)$
$1-w$	$60\,(1-w)\,w$	$400\,(1-w)(1-w)$

$v^\top \Sigma v$ adds the four terms in the grid.

$$v^\top \Sigma v = \underbrace{225w^2}_{\text{fund A}} + \underbrace{120w(1-w)}_{\text{both funds}} + \underbrace{400(1-w)^2}_{\text{fund B}}$$

Half in each fund

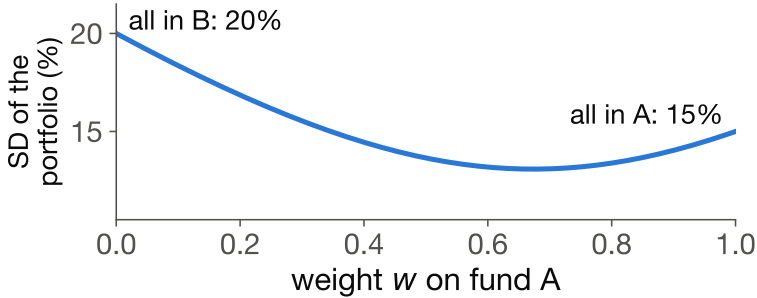


$$\text{Var} = 56.25 + 30 + 100 = 186.25$$

The SD is $\sqrt{186.25} \approx 13.6\%$.

This is less than the SD of each fund.

Diversification

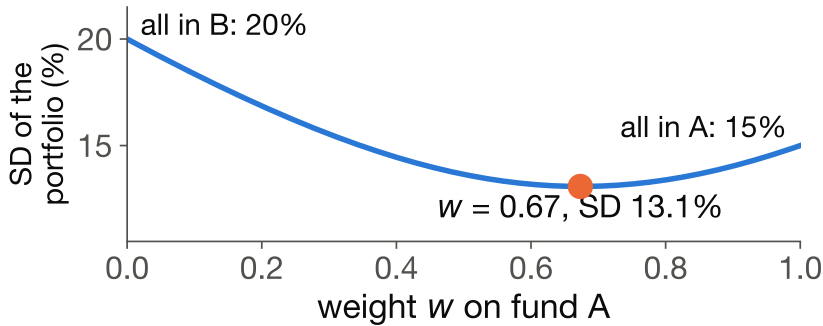


Each point is the SD of one portfolio.

Spreading money over funds can lower the risk.

This is called diversification.

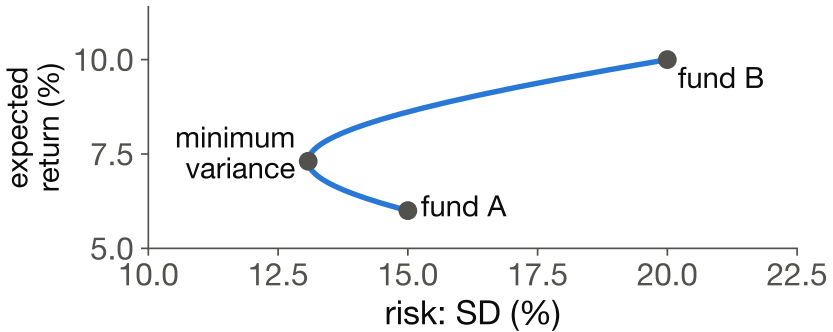
The minimum-variance portfolio



The smallest SD is 13.1%, at $w = 0.67$.

This is called the minimum-variance portfolio.

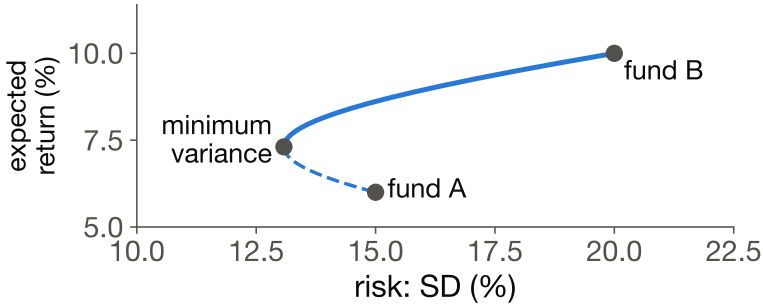
Return and risk of each portfolio



Each point on the curve is one portfolio.

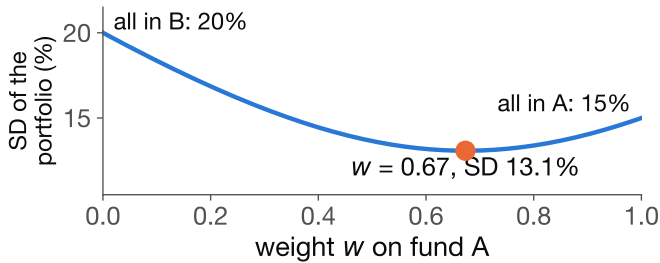
We plot the expected return against the SD.

The efficient frontier



Beyond the minimum-variance portfolio,
a higher expected return means a higher SD.
The solid part is called the efficient frontier.

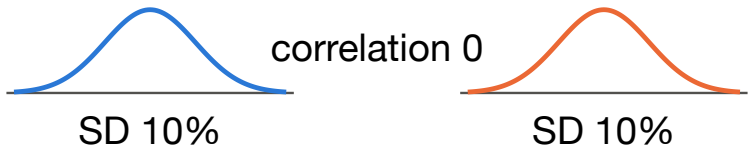
Summary: portfolio risk



The variance of a portfolio is $v^\top \Sigma v$.

A portfolio can have less risk than each fund.

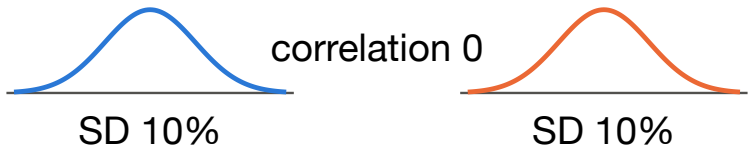
Checkpoint 4



Both SDs are 10%, and the correlation is 0.

What is the SD of the portfolio with $w = 0.5$?

Checkpoint 4



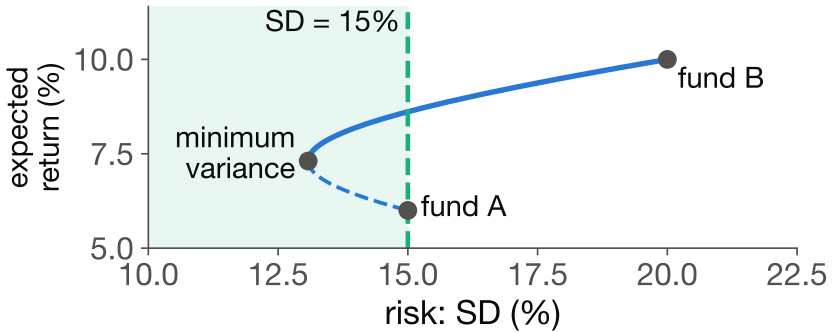
Both SDs are 10%, and the correlation is 0.

What is the SD of the portfolio with $w = 0.5$?

$$\text{Var} = 0.25 \times 100 + 0.25 \times 100 = 50$$

The SD is $\sqrt{50} \approx 7.1\%$, less than 10%.

A risk constraint



The investor accepts an SD of at most 15%.

SD $\leq 15\%$ is called a risk constraint.

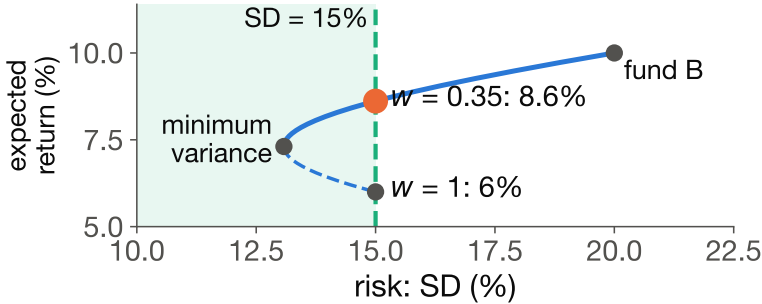
The portfolio problem

maximize $6w + 10(1 - w)$ (expected return)

subject to SD $\leq 15\%$ (risk constraint)

$$0 \leq w \leq 1$$

The best portfolio



The best portfolio has $w = 0.35$.

Its expected return is 8.6%, and its SD is 15%.

At $w = 1$, the expected return is 6% at the same SD.

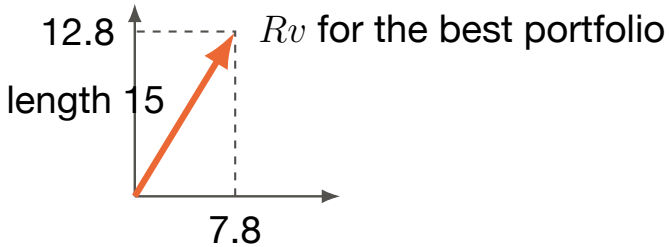
Writing Σ as $R^\top R$

A cone constraint needs the SD as a length.

$$R = \begin{pmatrix} 15 & 4 \\ 0 & \sqrt{384} \end{pmatrix} \qquad R^\top R = \begin{pmatrix} 225 & 60 \\ 60 & 400 \end{pmatrix} = \Sigma$$

For example, $4^2 + (\sqrt{384})^2 = 400$.

The SD as a length



$$v^\top \Sigma v = v^\top R^\top R v = \|Rv\|^2 \quad \text{(Lecture 2)}$$

So the SD of the portfolio is the length $\|Rv\|$.

For the best portfolio, the length of Rv is 15.

A cone constraint

$$\underbrace{\|Rv\|}_{\text{the length of a linear expression}} \leq \underbrace{15}_{\text{a number}}$$

So SD $\leq 15\%$ is a cone constraint (Lecture 9).

The expected return is linear, so this is an SOCP.

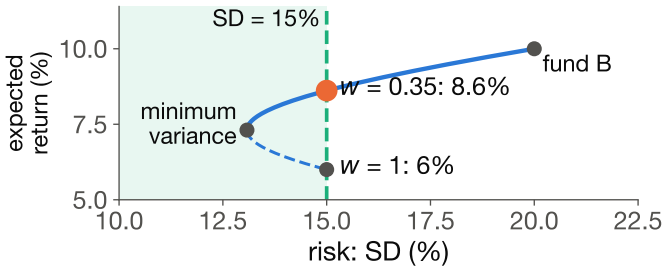
The portfolio in Python

```
import numpy as np, cvxpy as cp
mu = np.array([6, 10])
R = np.array([[15, 4], [0, np.sqrt(384)]])
v = cp.Variable(2)
cons = [cp.norm(R @ v) <= 15,
        cp.sum(v) == 1, v >= 0]
cp.Problem(cp.Maximize(mu @ v), cons).solve()
v.value     # [0.35, 0.65]
```

`cp.sum(v) == 1` means that $v_2 = 1 - v_1$.

cvxpy gives $v = (0.35, 0.65)$, so $w = 0.35$.

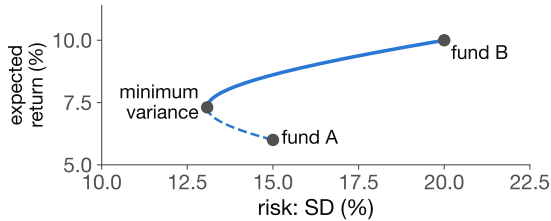
Summary: a risk constraint



The SD of a portfolio is the length $\|Rv\|$.

So a risk constraint is a cone constraint.

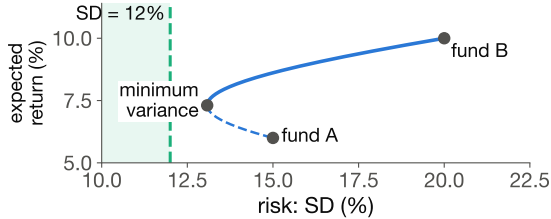
Checkpoint 5



The investor now accepts an SD of at most 12%.

Which portfolios are feasible?

Checkpoint 5



The investor now accepts an SD of at most 12%.

Which portfolios are feasible?

None. The smallest SD is 13.1%.

The problem is infeasible (Lecture 8).

Decisions under uncertainty in practice



fresh food
to stock



seats
to sell



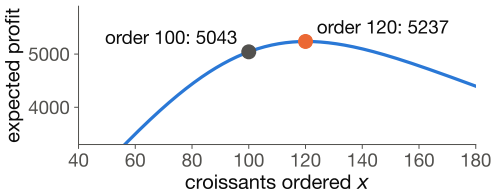
medicine
to stock

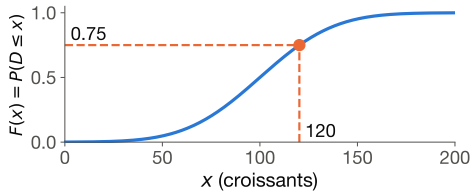


a portfolio
of funds

Photos (Wikimedia Commons, cropped): Andy Li, CC0; Aykleinman, public domain; Mattes, public domain; Francisco Gonzalez, CC BY 2.0.

Lecture summary



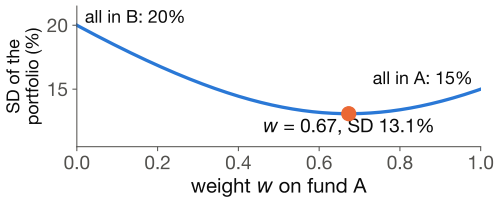


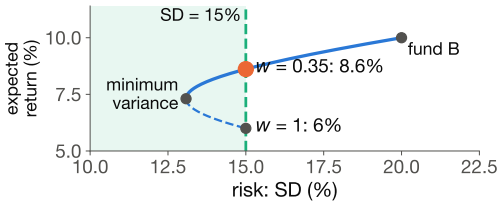
$\mathbb{E}[\text{profit}]$ is the long-run average profit.

The optimal order is the $(q - c)/q$ quantile.

The pinball loss gives the same quantile.

Lecture summary (continued)

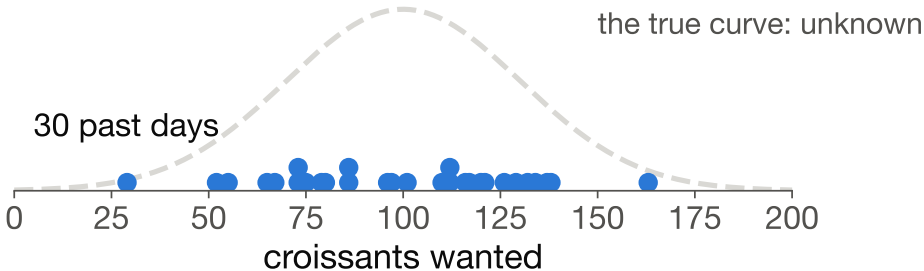




Diversification can lower the risk.

With a risk constraint, the problem is an SOCP.

Next lecture



Today, we used the true distribution of D.

In Lecture 12, we know D only on past days.

Exercises

1. Find the optimal quantile for $c = 30$, $q = 90$.
2. Which w minimizes the SD in Checkpoint 4?
3. Solve the portfolio problem with SD $\leq 18\%$.