Information Economics

Introduction and Review of Optimization

Ling-Chieh Kung

Department of Information Management
National Taiwan University

Road map

- **Course overview**.
- Convexity and optimization.
- Applications.

Welcome!

- This is **Information Economics**, NOT **Information Economy**.
  - We do not put emphasis on IT, IS, information goods, etc.
  - We focus on **information**.
- We focus on the **economics of information**.
  - How people behave with different information?
  - What is the value of information?
  - What information to acquire? How?
  - What are the implications on business and economy?
- **Information asymmetry** is particularly important.

Information asymmetry

- The world is full of asymmetric information:
  - A consumer does not know a retailer's procurement cost.
  - A consumer does not know a product's quality.
  - A retailer does not know a consumer's valuation.
  - An instructor does not know how hard a student works.
- As the world is **decentralized**:
  - There is the **incentive** issue.
  - There is the **information** issue.
- As information asymmetry results in inefficiency, we want to:
  - Analyze its impact. If possible, quantify it.
  - Decide whether it introduces driving forces for some phenomena.
  - Find a way to deal with it if it cannot be eliminated.
- This field is definitely fascinating. However:
  - We need to have some "**weapons**" to explore the world!

Before you enroll...

- Prerequisites:
  - Calculus.
  - Probability.
  - Convex optimization.
  - Game theory.
- This is an **academic methodology** course.
  - It is directly helpful if you are going to write a thesis with this research methodology.
  - It can be indirectly helpful for you to analyze the real world. However, we do not train you to do that in this course.
- This course is about **science**, not **business** or **engineering**.
  - It is about **identifying reasons**.
  - It is not about **solving problems**.
  - It is not about **making decisions**.

The instructing team

- Instructor:
  - Ling-Chieh Kung.
  - Assistant professor.
  - Office: Room 413, Management Building II.
  - Office hour: by appointment.
  - E-mail: lckung@ntu.edu.tw.
- There is no teaching assistant for this course.

Related information

- Classroom: Room 204, Management Building II.
- Lecture time: 9:10am-12:10pm, Monday.
- References:
  - *Information Rules* by C. Shapiro and H. Varian.
  - *Freakonomics* by S. Levitt and S. Dubner.
  - *Contract Theory* by P. Bolton and M. Dewatripont.
  - *Game Theory for Applied Economists* by R. Gibbons.
  - About ten academic papers.

"Flipped classroom"

- Lectures in **videos**, then discussions in classes.
- Before each Monday, the instructor uploads a video of lectures.
  - Ideally, the video will be no longer than one and a half hour.
  - Students must watch the video by themselves before that Monday.
- During the lecture, we do three things:
  - Discussing the lecture materials.
  - Solving lecture problems (to earn points).
  - Further discussions.
- Teams:
  - Students form teams to work on class problems and case studies.
  - Each team should have **two or three students**.
  - Your teammates may be different from week to week.

Pre-lecture problems and class participation

- No homework!
  - Problem sets and solutions will be posted for students to do practices.
- Pre-lecture problems.
  - One problem to submit per set of lecture videos.
  - Submit a hard copy at the beginning of a lecture.
- Class participation:
  - Just say something!
  - Use whatever way to impress the instructor.

Paper presentations and projects

- Paper presentations:
  - Students will form six teams to present six academic papers. The team size will be determined according to the class size.
  - On the date that a team present, they should submit one paper summary and their slides.
- Midterm project:
  - Students form teams to do a midterm project.
  - A topic will be assigned, and each team constructs its own models and generate its own findings.
  - A written report is required.
- Final project:
  - Students form teams to do a midterm project.
  - A direction will be assigned, and each team conducts its own research by defining its own research questions.
  - Each team will submit a proposal for the self-selected topic, make a 30-minute presentation, and submit a report.

Grading

- Not dropping this course: 10%.
- Class participation: 10%.
- Pre-lecture problems: 10%.
- Paper presentations: 20%.
- Midterm project: 20%.
- Final project: 30%.
- The final letter grades will be given according to the following conversion rule:

Letter	Range	Letter	Range	Letter	Range
A+	$[90, 100]$	B+	$[77, 80)$	C+	$[67, 70)$
A	$[85, 90)$	B	$[73, 77)$	C	$[63, 67)$
A−	$[80, 85)$	B−	$[70, 73)$	C−	$[60, 63)$

Important dates, tentative plan, and websites

- Tentative plan:
  - Incentives: 5 weeks.
  - Information: 5 weeks.
  - Student presentations: 4 weeks.
  - Review and preview: 1 week.
- CEIBA.
  - Viewing your grades.
- http://www.ntu.edu.tw/~lckung/courses/IE16/.
  - Downloading course materials.
  - Linking to lecture videos.
- https://piazza.com/ntu.edu.tw/fall2016/im7011/.
  - On-line discussions.
  - Receiving announcements.

Quiz

- Now it is time for a quiz!

Road map

- Course overview.
- **Convexity and optimization**.
- Applications.

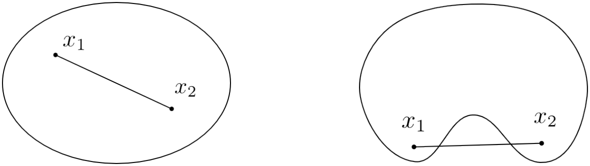
Convex sets

Definition 1 (Convex sets)

A set F is **convex** *if*

$$\lambda x_1 + (1-\lambda)x_2 \in F$$

for all $\lambda \in [0,1]$ and $x_1, x_2 \in F$.



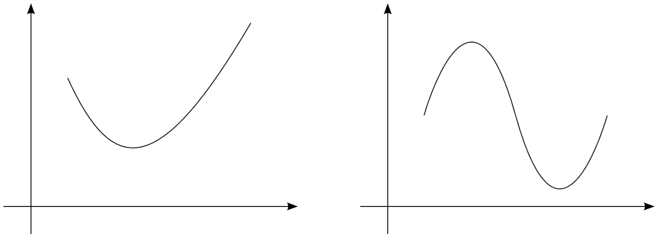
Convex functions

Definition 2 (Convex functions)

For a convex domain F, a function $f(\cdot)$ is **convex** *over F if*

$$f\Big(\lambda x_1 + (1-\lambda)x_2\Big) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

for all $\lambda \in [0,1]$ and $x_1, x_2 \in F$.



Some examples

- Convex sets?
  - $X_1 = [10, 20]$.
  - $X_2 = (10, 20)$.
  - $X_3 = \mathbb{N}$.
  - $X_4 = \mathbb{R}$.
  - $X_5 = \{(x, y) | x^2 + y^2 \leq 4\}$.
  - $X_6 = \{(x, y) | x^2 + y^2 \geq 4\}$.
- Convex functions?
  - $f_1(x) = x + 2, x \in \mathbb{R}$.
  - $f_2(x) = x^2 + 2, x \in \mathbb{R}$.
  - $f_3(x) = \sin(x), x \in (0, 2\pi)$.
  - $f_4(x) = \sin(x), x \in (\pi, 2\pi)$.
  - $f_5(x) = \log(x), x \in (0, \infty)$.
  - $f_6(x, y) = x^2 + y^2, (x, y) \in \mathbb{R}^2$.

Strictly convex and concave functions

Definition 3 (Strictly convex functions)

For a convex domain F, a function $f(\cdot)$ is **strictly convex** *over F if*

$$f\Big(\lambda x_1 + (1-\lambda)x_2\Big) < \lambda f(x_1) + (1-\lambda)f(x_2)$$

for all $\lambda \in (0,1)$ and $x_1, x_2 \in F$ such that $x_1 \neq x_2$.

Definition 4 ((Strictly) concave functions)

For a convex domain F, a function $f(\cdot)$ is **(strictly) concave** *over F if $-f(\cdot)$ is (strictly) convex.*

Derivatives of convex functions

- In this course, most of the functions are twice-differentiable with continuous second-order derivatives.
- Recall a function's gradient and Hessian:
  - For an n-dimensional differentiable function $f(x)$, its **gradient** is the $n \times 1$ vector

$$\nabla f \equiv \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}.$$

  - For an n-dimensional twice-differentiable function $f(x_1, ..., x_n)$, its **Hessian** is the $n \times n$ matrix

$$\nabla^2 f \equiv \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \, \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

- (Calculus) If the second-order derivatives are all continuous, the Hessian is symmetric.

Derivatives of convex functions

- Let f be twice-differentiable with continuous second-order derivatives:

Proposition 1

For $f : \mathbb{R} \to \mathbb{R}$ over an interval $F \subseteq \mathbb{R}$:

- *f is (strictly) convex over F if and only if $f''(x) \geq 0$ (> 0) for all $x \in F$.*
- *f is (strictly) concave over F if and only if $f''(x) \leq 0$ (< 0) for all $x \in F$.*

Proposition 2

For $f : \mathbb{R}^n \to \mathbb{R}$ over a region $F \subseteq \mathbb{R}^n$:

- *f is (strictly) convex over F if and only if $\nabla^2 f(x)$ is positive (semi)definite for all $x \in F$.*
- *f is (strictly) concave over F if and only if $\nabla^2 f(x)$ is negative (semi)definite for all $x \in F$.*

- (Linear Algebra) A symmetric $n \times n$ matrix A is called positive (semi)definite if $y^T A y \geq 0$ (> 0) for all $y \in \mathbb{R}^n$.

Some examples revisited

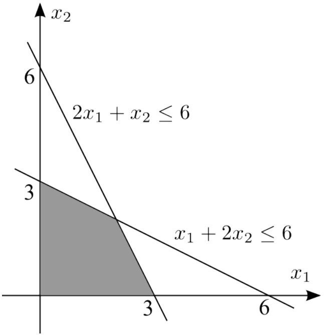
- $f_1(x) = x + 2, x \in \mathbb{R}$: $f_1''(x) = 0$, convex and concave.
- $f_2(x) = x^2 + 2, x \in \mathbb{R}$: $f_2''(x) = 2 > 0$, strictly convex.
- $f_3(x) = \sin(x), x \in (0, 2\pi)$, $f_3''(x) = -\sin(x)$, neither.
- $f_4(x) = \sin(x), x \in (\pi, 2\pi)$, $f_4''(x) = -\sin(x) > 0$, strictly convex.
- $f_5(x) = \log(x), x \in (0, \infty)$: $f_5''(x) = -\frac{1}{x^2} < 0$, strictly concave.
- $f_6(x, y) = x^2 + y^2, (x, y) \in \mathbb{R}^2$: $\nabla^2 f_6(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is positive definite, strictly convex.

Linear programming

- Consider the problem

$$\begin{aligned} z^* = \max \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1 + 2x_2 \leq 6 \\ & 2x_1 + x_2 \leq 6 \\ & x_i \geq 0 \quad \forall i = 1, 2. \end{aligned}$$

  - The feasible region is the shaded area.
  - An optimal solution is $(x_1^*, x_2^*) = (2, 2)$. Is it unique?
  - The corresponding objective value $z^* = 6$.



- An optimization problem is a **linear program** (LP) if the objective function and constraints are all linear.

Nonlinear programming

- An optimization problem is a **nonlinear program** (NLP) if it is not a linear program.
- Consider the problem

$$\begin{aligned} z^* = \ \max \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 16 \\ & x_1 + x_2 \geq 1. \end{aligned}$$

  - What is the feasible region?
  - What is an optimal solution? Is it unique?
  - What is the value of z^*?
- An optimization problem is a **convex program** if in it we maximize a concave function over a convex feasible region.
- All convex programs can be solved efficiently.
- It may not be possible to solve a nonconvex program efficiently.

Infeasible and unbounded problems

- Not all problems have an optimal solution.
- A problem is **infeasible** if there is no feasible solution.
  - E.g., $\max\{x^2 | x \leq 2, x \geq 3\}$.
- A problem is **unbounded** if given any feasible solution, there is another feasible solution that is better.
  - E.g., $\max\{e^x | x \geq 3\}$.
  - How about $\min\{\sin x | x \geq 0\}$?
- A problem may be infeasible, unbounded, or finitely optimal (i.e., having at least one optimal solution).

Set of optimal solutions

- The **set of optimal solutions** of a problem $\max\{f(x)|x \in X\}$ is

$$\operatorname{argmax}\{f(x)|x \in X\}.$$

- For $f(x) = \cos x$ and $X = [0, 2\pi]$, we have

$$\operatorname{argmax}\left\{\cos x \middle| x \in [0, 2\pi]\right\} = \{0, 2\pi\}.$$

- If x^* is an optimal solution of $\max\{f(x)|x \in X\}$, we should write

$$x^* \in \operatorname{argmax}\{f(x)|x \in X\},$$

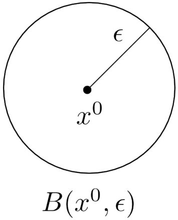
NOT $x^* = \operatorname{argmax}\{f(x)|x \in X\}$!

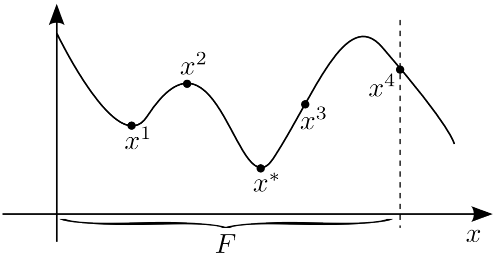
Global optima

- For a function $f(x)$ over a feasible region F:
  - A point x^* is a **global minimum** if $f(x^*) \leq f(x)$ for all $x \in F$.
  - A point x' is a **local minimum** if for some $\epsilon > 0$ we have

$$f(x') \leq f(x) \quad \forall x \in B(x', \epsilon) \cap F,$$

where $B(x^0, \epsilon) \equiv \{x | d(x, x^0) \leq \epsilon\}$ and $d(x, y) \equiv \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$.





- **Global maxima** and **local maxima** are defined accordingly.

First-order necessary condition

- Consider an **unconstrained** problem

$$\max_{x \in \mathbb{R}^n} f(x).$$

Proposition 3 (Unconstrained FONC)

For a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, a point x^* is a local maximum of f only if

- $f'(x^*) = 0$ *if* $n = 1$.
- $\nabla f(x^*) = 0$ *if* $n \geq 2$.

Examples

- Consider the problem

$$\max_{x \in \mathbb{R}} \; x^3 - \frac{9}{2}x^2 + 6x + 2$$

  The FONC yields

$$3(x^2 - 3x + 2) = 0.$$

  Solving the equation gives us 1 and 2 as two candidates of local maxima.
  - It is easy to see that $x^* = 1$ is a local maxima but $\tilde{x} = 2$ is NOT.

- Consider the problem

$$\max_{x \in \mathbb{R}^2} \; f(x) = x_1^2 - x_1 x_2 + x_2^2 - 6x_2.$$

  The FONC yields

$$\nabla f(x) = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 2x_2 - 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

  Solving the linear system gives us $(2, 4)$ as the only candidate of local maxima.
  - Note that it is NOT necessarily a local maximum!

Second-order necessary condition

- Let's proceed further.

Proposition 4 (Unconstrained SONC)

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice-differentiable. For a point x^* to be a local maximum of f, we need:

- $f''(x^*) \leq 0$ *if* $n = 1$.
- $\nabla^2 f(x^*)$ *is negative semidefinite if* $n \geq 2$.

- Note that we do not need the function to be concave; we only need f'' or $\nabla^2 f$ to be negative or negative definite **at the point** x^*.
- In this course, we will not apply the SONC a lot.
- Here our point is that a local maximum requires NOT just

$$\frac{\partial^2 f}{\partial x_i^2} \leq 0 \quad \forall i = 1, ..., n.$$

We want more than candidates!

- The FONC and SONC produce candidates of local maxima/minima.
- What's next?
  - We need some ways to ensure local optimality.
  - We need to find a global optimal solution.
- If the function is **convex or concave**, things are much easier:

Proposition 5

For a differentiable convex (concave) function $f : \mathbb{R}^n \to \mathbb{R}$:

- *x^* is a global minimum (maximum) if and only if $\nabla f(x^*) = 0$.*
- *The global optimum is unique if f is strictly convex or concave.*

Remarks

- When you are asked to solve a problem:
  - First check whether the objective function is convex/concave. If so the problem typically becomes much easier.
- All the conditions for unconstrained problems apply to **interior** points of a feasible region.
- One common strategy for solving constrained problems proceeds in the following steps:
  - **Ignore** all the constraints.
  - Solve the unconstrained problem.
  - Verify that the unconstrained optimal solution satisfies all constraints.
- If the strategy fails, we seek for other ways.

Binding constraints and boundary solutions

- An optimal solution may lie on the boundary of the feasible region.
  - It is a **boundary solution** or a **corner solution**.
- We need to take a look at those **binding** (or **active**) constraints:

Definition 5

Let $g(\cdot) \leq b$ be an inequality constraint and $\bar{x}$ be a solution. $g(\cdot)$ is binding at $\bar{x}$ if $g(\bar{x}) = b$.

  - $x_1 + x_2 \leq 10$ is binding at $(x_1, x_2) = (2, 8)$.
  - $2x_1 + x_2 \geq 6$ is nonbinding at $(x_1, x_2) = (2, 8)$.
  - $x_1 + 3x_2 = 9$ is binding at $(x_1, x_2) = (6, 1)$.
- Remarks:
  - An inequality is nonbinding (inactive) at a point if it is strictly satisfied.
  - An equality constraint is always binding at any feasible solution.

Binding constraints and boundary solutions

- Consider a **single-dimensional** constrained optimization problem

$$\begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0 \quad \forall i = 1, ..., m. \end{aligned}$$

- If $f(\cdot)$ is strictly concave:
  - Apply the FOC to obtain a candidate solution $\bar{x}$.
  - If $\bar{x}$ is feasible, it is optimal.
  - Otherwise, the feasible point that is closest to $\bar{x}$ is optimal.
- In general:
  - Apply the FONC and SONC to obtain a set of candidate solutions.
  - Include all the boundary points as candidate solutions.
  - Compare all the candidate solutions to find an optimal one.
- For a **multi-dimensional** constrained optimization problem, more advanced techniques are required (e.g., the KKT condition).

Road map

- Course overview.
- Convexity and optimization.
- **Applications**.

Monopoly pricing

- Suppose a monopolist sells a single product to consumers.
- Consumers are **heterogeneous** in their **willingness-to-pay**, or **valuation**, of this product.
- One's valuation, θ, lies on the interval $[0, b]$ uniformly.
  - He buys the product if and only if his valuation is above the price.
  - Consumers' decisions are independent.
  - The total number of consumers is a.
  - Given a price p, in expectation the number of consumers who buy the product is
$$a \Pr(\theta \geq p) = a\left(1 - \frac{p}{b}\right).$$
- The unit production cost is c.
- The seller chooses a unit price p to maximize her total expected profit.

Formulation

- The **endogenous** decision variable is p.
- The **exogenous** parameters are a, b, and c.
- The only constraint is $p \geq 0$.
- Let $\pi(p)$ be the profit under price p. We have

$$\pi(p) = (p - c)a\left(1 - \frac{p}{b}\right).$$

- The complete problem formulation is

$$\begin{aligned} \max \quad & (p - c)a\left(1 - \frac{p}{b}\right) \\ \text{s.t.} \quad & p \geq 0. \end{aligned}$$

  - It is without loss of generality to **normalize** the population size a to 1.

Solving the problem

- Given that $\pi(p) = \frac{a}{b}(p-c)(b-p)$, let's show it is strictly concave:
  - $\pi'(p) = \dfrac{a}{b}(b+c-2p)$.
  - $\pi''(p) = -2\left(\dfrac{a}{b}\right) < 0$.
- Great! Now let's ignore the constraint $p \geq 0$.
- Applying the FOC, the unconstrained optimal solution is

$$b + c - 2\bar{p} = 0 \quad \Leftrightarrow \quad \bar{p} = \frac{b+c}{2}.$$

- Does $\bar{p}$ satisfy the ignored constraint? Is it globally optimal?

Managerial/economic implications

- The optimal price $\bar{p} = \frac{b+c}{2}$ tells us something:
  - $\bar{p}$ is increasing in the highest possible valuation b. Why?
  - $\bar{p}$ is increasing in the unit cost c. Why?
  - $\bar{p}$ has nothing to do with the total number of consumer a. Why?
- The optimal profit $\pi^* \equiv \pi(\bar{p}) = \frac{a(b-c)^2}{4b}$.
  - π^* is decreasing in c. Why?
  - π^* is increasing in a. Why?
  - How is π^* affected by b?
  - Let's answer it:

$$\frac{\partial}{\partial b}\pi^* = \frac{a(b-c)(b+c)}{4b^2} > 0 \quad \text{(why } b > c\text{?)}.$$

- It is these **qualitative** managerial/economic implications that matters.
- Never forget to verify your solutions with your **intuitions**.

Impact of price control

- Sometimes the price is controlled (e.g., by the government) and has a cap K.
- For the problem

$$\begin{aligned} \max \quad & (p-c)a\left(1-\frac{p}{b}\right) \\ \text{s.t.} \quad & p \in [0, K], \end{aligned}$$

  the first-order solution $\frac{b+c}{2}$ may be infeasible.
- The optimal price is

$$p^* = \begin{cases} \frac{b+c}{2} & \text{if } \frac{b+c}{2} \leq K \\ K & \text{otherwise} \end{cases}.$$

Newsvendor problem

- In some situations, people sell **perishable products**.
  - They become valueless after the **selling season** is end.
  - E.g., newspapers become valueless after each day.
  - High-tech goods become valueless once the next generation is offered.
  - Fashion goods become valueless when they become out of fashion.
- In many cases, the seller only have **one chance** for replenishment.
  - E.g., a small newspaper seller can order only once and obtain those newspapers only at the morning of each day.
- Often sellers of perishable products face **uncertain demands**.
- **How many** products one should prepare for the selling season?
  - Not too many and not too few!

Newsvendor model

- Let D be the uncertain demand.
- Let F and f be the cdf and pdf of D (assuming D is continuous).
- Let r and c be the unit sales revenue and purchasing cost, respectively.
- Let q be the order quantity.
- The (expected) profit-maximizing newsvendor solves

$$\max_{q \geq 0} \; r\mathbb{E}\Big[\min\{q, D\}\Big] - cq.$$

  - Let $\pi(q) = r\mathbb{E}[\min\{q, D\}] - cq$ be the expected profit function.
  - The model can be expanded to include salvage values, disposal fees, shortage costs, etc.

Convexity of the profit function

- The expected profit function $\pi(q)$ is

$$\begin{aligned}\pi(q) &= r\mathbb{E}\Big[\min\{q, D\}\Big] - cq \\ &= r\left\{\int_0^q xf(x)dx + \int_q^\infty qf(x)dx\right\} - cq \\ &= r\left\{\int_0^q xf(x)dx + q\Big[1 - F(q)\Big]\right\} - cq.\end{aligned}$$

- We have

$$\pi'(q) = r\left\{qf(q) + 1 - F(q) - qf(q)\right\} - c = r\Big[1 - F(q)\Big] - c$$

and $\pi''(q) = -rf(q) < 0$. $\pi(q)$ is strictly concave.

Optimizing the order quantity

- Let $\bar{q}$ be the order quantity that satisfies the FOC, we have

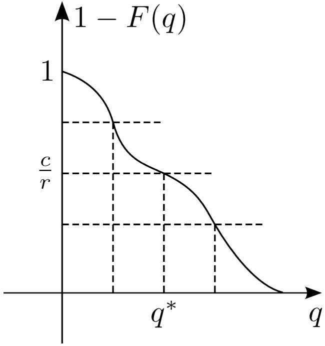
$$r\Big[1 - F(\bar{q}\Big] - c = 0$$
$$\Rightarrow F(\bar{q}) = 1 - \frac{c}{r} \quad \text{or} \quad 1 - F(\bar{q}) = \frac{c}{r}.$$

- Such $\bar{q}$ must be positive (for regular demand distributions).
  - So $\bar{q}$ is optimal.
  - The quantity $\bar{q}$ is called the **newsvendor** quantity.
  - The formula applies to **any** continuous random variable D.

Interpretations of the newsvendor quantity

- The newsvendor quantity $\bar{q}$ satisfies $1 - F(\bar{q}) = \frac{c}{r}$.
  - The probability of having a shortage, $1 - F(q)$, is decreasing in q.
- The optimal quantity $\bar{q}$ is:
  - Decreasing in c.
  - Increasing in r.

  Does that make economic sense?



Impact of capacity limitation

- Sometimes the capacity is limited and at most K units can be ordered.
- For the problem

$$\max_{q \in [0,K]} r\mathbb{E}\Big[\min\{q, D\}\Big] - cq,$$

the newsvendor quantity $\bar{q}$ satisfying $1 - F(\bar{q}) = \frac{c}{r}$ may be infeasible.

- The optimal order quantity is

$$q^* = \begin{cases} \bar{q} & \text{if } r\Big[1 - F(K)\Big] - c \leq 0 \\ K & \text{otherwise} \end{cases},$$

where $r[1 - F(K)] - c \leq 0$ is equivalent to $\bar{q} \leq K$ (why?).