Less is More: Optimal Contest Design with a Shortlist

Joint Work with

Ningyuan Li[1], Weian Li[2], Qi Qi[3] and Changyuan Yu[4]

[1]Peking University, [2]Shandong University,
[3]Renmin University of China, [4]Baidu Inc

Hanbing Liu,
Renmin University of China (RUC)

Motivation

Contests are everywhere.









Rank-order Contest

Designer

Announce a **rank-order** prize structure $\vec{V} \in \mathbb{R}^n$.

Player

n player register the contest
each associate with a iid private **ability** $x_i \sim F$.

Player

Decide an **effort** e_i to exert, with cost $g(e_i)/x_i$



Prize is allocated based on the **rank-order of efforts**
and utility of player and the designer are realized

$$If \quad e_1 > e_2 > \cdots > e_n$$
$$Then \quad V_1 \geq V_2 \geq \cdots \geq V_n$$

Motivation

Large-scale online competition

1) Players compete for prizes by creating competent machine learning models

2) Organizers get good solutions



kaggle

NeurIPS - Open Polymer Prediction 2025

Predicting polymer properties with machine learning to accelerate sustainable materials research.

Competition Host
University of Notre Dame

Prizes & Awards
$50,000
Awards Points & Medals

Participation
5,899 Entrants
1,019 Participants
928 Teams
7,835 Submissions

Prizes

- 1st Place: $12,000
- 2nd Place: $10,000
- 3rd Place: $10,000
- 4th Place: $8,000
- 5th Place: $5,000

Prize Contenders

#	Team	Members	Score
1	James Day		0.031
2	Martin Kovacevic Buvinic		0.033
3	MT		0.033
4	Yichencai01		0.033
5	naPoly		0.033
6	Halil İbrahim kaya		0.033

Motivation

Large-scale online competition

1) Players compete for prizes by creating competent machine learning models

2) Organizers get good solutions



kaggle

NeurIPS - Open Polymer Prediction 2025

Predicting polymer properties with machine learning to accelerate sustainable materials research.

Competition Host
University of Notre Dame

Prizes & Awards
$50,000
Awards Points & Medals

Participation
5,899 Entrants
1,019 Participants
928 Teams
7,835 Submissions

Prizes

- 1st Place: $12,000
- 2nd Place: $10,000
- 3rd Place: $10,000
- 4th Place: $8,000
- 5th Place: $5,000

Prize Contenders

#	Team	Members	Score
1	James Day		0.031
2	Martin Kovacevic Buvinic		0.033
3	MT		0.033
4	Yichencai01		0.033
5	naPoly		0.033
6	Halil İbrahim kaya		0.033

Requires a lot of Computational resources to evaluate

Motivation

History-based User Rating
Can platform use this information to shortlist players in advance, achieving similar (or even better) performance?

Kaggle Rankings

Competitions | Datasets | Notebooks | Discussions

Learn more about rankings

372 Grandmasters | 2,187 Masters | 10,786 Experts | 67,409 Contributors | 121,348 Novices

Rank	Tier	User		Medals			Points
1		Dieter	joined 8 years ago	46	17	3	165,314
2		c-number	joined 4 years ago	11	2	0	150,059
3		tascj	joined 10 years ago	19	11	3	120,689
4		hyd	joined 4 years ago	13	13	0	120,074
5		yuanzhe zhou	joined 5 years ago	8	11	9	101,035

Motivation



Total air time is limited

Total number of matches are fixed



This Paper [Sun et al., 24]

Designer

Announce the shortlist capacity m
and a rank-order prize structure $\vec{V} \in \mathbb{R}^m$.

Player

n player register the contest
each associate with a iid private ability $x_i \sim F$.

Designer

The designer sees each players' ability
(By submitted CV, ELO Score, International Points, etc.)
and shortlist to m strongest players.

Player

m admitted players
Decide an effort e_i to exert, with cost $g(e_i)/x_i$

Contestants' Game

The symmetric equilibrium strategy is unique.

Theorem 1 [Sun et al., 24]

For any ability distribution F, any size of shortlist m and any rank-order prize structure $\vec{V}$, the unique symmetric Bayesian Nash equilibrium exists if and only if $x/J(F,n,m,x)$ is non-decreasing in the support of ability distribution, and can be expresses as:

$$b^*(x) = g^{-1}\left(\int_0^x \frac{\sum_{l=1}^{m-1}\binom{n-1}{l-1}(n-l)(V_l-V_{l+1})F^{n-l-1}(t)(1-F(t))^{l-1}f(t)}{J(F,n,m,t)}\,tdt\right),$$

where $g(\cdot)$ is the cost function and $J(F,n,m,x)$ is a normalizing factor.

Given a fixed budget B,
how to design contest that maximize designer's *effort* objective?

Contest Design



We focus on two concrete types of effort objectives:

1) Maximum Individual Effort: Candidate Selection, Project Bid, etc.

$$\mathbb{E}[b(X_{(1)})]$$

2) Total Effort: Sport game viewership, Creatives Collection, etc.

$$\mathbb{E}[\sum_{i=1}^{m} b(X_{(i)})]$$



Theorem [Moldovanu and Sela, 01]

The **optimal rank-order contest is the winner-take-all contest**, i.e. $V_1 = B$ and 0 elsewhere, for both objectives (concave cost for Total Effort).

Reseach Questions

- **RQ1:** Will a shortlist improves the contest's outcome?

- **RQ2:** When shortlists are applied,
  what the best contest designs $\left(m, \vec{V}\right)$ will be?

- **RQ3:** Can we measure the performance gain / loss
  from incorporating shortlists?

Maximum Individual Effort

A two-player winner-take-all contest is the best ex-post design.

Theorem 2

The optimal contest that maximizes the ex-post maximum individual effort is **a two-player winner-take-all contest**, where $m = 2$ and $V_1 = B$.

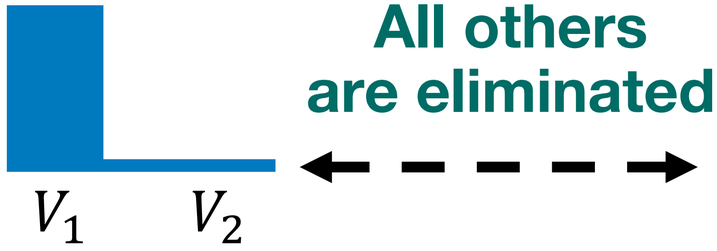
The resulting maximum ex-post individual effort is as follows:

$$g^{-1}\left(B\int_0^{x_1}\frac{(n-1)f(t)t}{F(t)+(n-1)(1-F(t))}dt\right),$$

where x_1 denotes the highest realized ability level among all players.

So as for ex-ante, as a direct corollary (Corollary 3).

$m = 2$



Total Effort

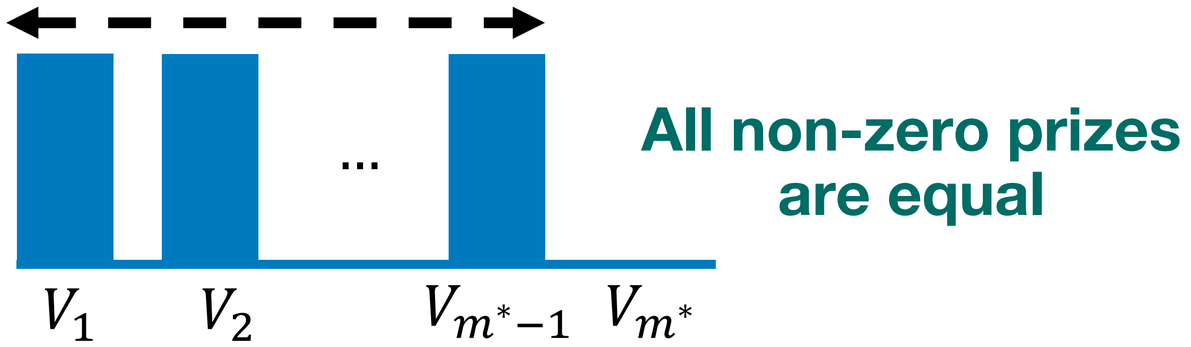
Proposition 3

The optimal contest with a shortlist that maximizes total effort is to **admit m^* contestants and set $m^* - 1$ equal prizes**.

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k - k^2)\frac{1}{q^2}\right)dq = 0.$$



All non-zero prizes are equal

Total Effort

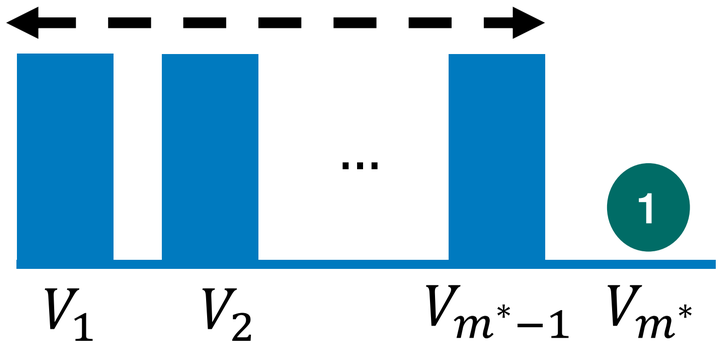
Proposition 3

The optimal contest with a shortlist that maximizes total effort is to **admit m^* contestants and set $m^* - 1$ equal prizes**.

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k - k^2)\frac{1}{q^2}\right)dq = 0.$$



Total Effort

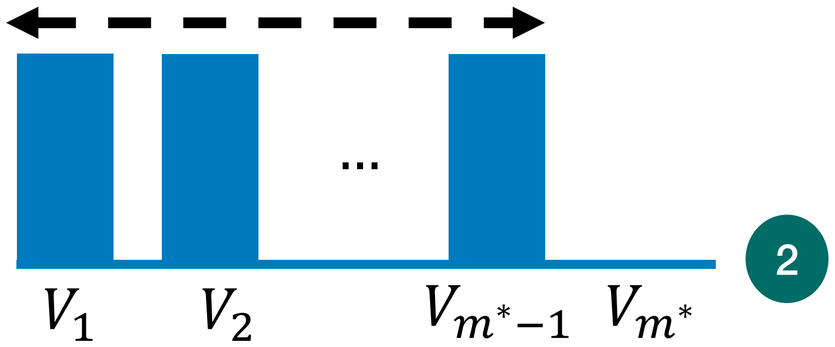
Proposition 3

The optimal contest with a shortlist that maximizes total effort is to **admit m^* contestants and set $m^* - 1$ equal prizes**.

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k - k^2)\frac{1}{q^2}\right)dq = 0.$$



Total Effort

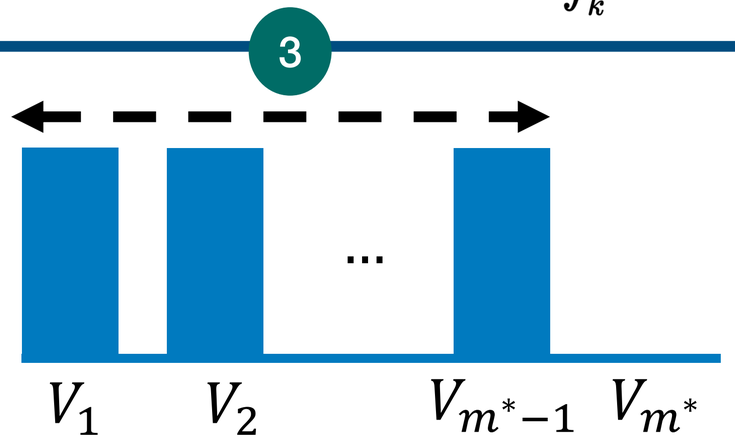
Proposition 3

The optimal contest with a shortlist that maximizes total effort is to **admit m^* contestants and set $m^* - 1$ equal prizes**.

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k - k^2)\frac{1}{q^2}\right)dq = 0.$$



Contestants' Game

The symmetric equilibrium strategy is unique.

Theorem 1 [Sun et al., 24]

For any ability distribution F, any size of shortlist m and any rank-order prize structure $\vec{V}$, the unique symmetric Bayesian Nash equilibrium exists if and only if $x/J(F,n,m,x)$ is non-decreasing in the support of ability distribution, and can be expresses as:

$$b^*(x) = g^{-1}\left(\int_0^x \frac{\sum_{l=1}^{m-1}\binom{n-1}{l-1}(n-l)(V_l - V_{l+1})F^{n-l-1}(t)(1-F(t))^{l-1}f(t)}{J(F,n,m,t)}\,t dt\right),$$

where $g(\cdot)$ is the cost function and $J(F,n,m,x)$ is a normalizing factor.

Given a fixed budget B,
how to design contest that maximize designer's *effort* objective?

Total Effort

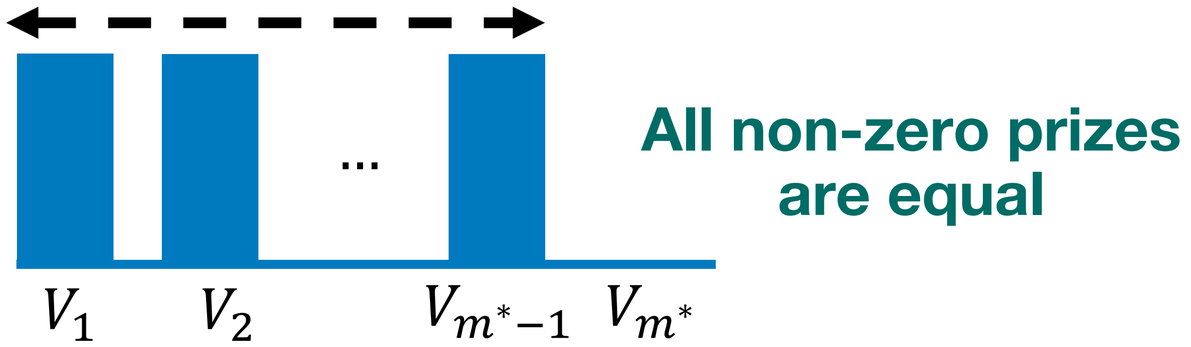
Proposition 3

The optimal contest with a shortlist that maximizes total effort is to **admit m^* contestants and set $m^* - 1$ equal prizes**.

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k - k^2)\frac{1}{q^2}\right)dq = 0.$$



All non-zero prizes are equal

Total Effort

We further show that the ratio k^* is bounded.

Theorem 4

For an arbitrary ability distribution F, when $n \to \infty$, **the optimal shortlist size ratio is upper bounded by $\bar{k} \approx 31.62\%$**, which is the solution to the equation $\ln k = (2-k)(k-1)$.

Moreover, there exists a distribution such that $\frac{m^*(n)}{n} = \bar{k}$, meaning that **the bound is tight**.

Also, we show that for any ability distribution F,
Total Effort monotonically decreses after $k > \bar{k}$.

Total Effort

We further show that the ratio k^* is bounded.

Theorem 4

For an arbitrary ability distribution F, when $n \to \infty$, **the optimal shortlist size ratio is upper bounded by $\bar{k} \approx 31.62\%$**, which is the solution to the equation $\ln k = (2-k)(k-1)$.

Moreover, there exists a distribution such that $\frac{m^*(n)}{n} = \bar{k}$, meaning that **the bound is tight**.

Therefore a shortlist makes the designer better off even if the distribution is unknown!
(by shortlist to $\bar{k}n$ contestants)

Performance Comparison

How much advantage shortlist can bring against optimal normal contest?

Maximum Individual Effort

Theorem 5

For any ability distribution, the two-player winner-take-all contest results in $\Theta(\log n)$ times the effort of the optimal contest without a shortlist.

Total Effort

Theorem 6

Fixed any ability distribution F, the optimal contest with a shortlist can achieve $\Theta(n)$ times the effort of the optimal contest without a shortlist.

Roadmap

Original joint optimization problem

$(\vec{V}, m)$

↓

Discrete univariate problem

m

↓

Asymptotic version
continuous univariate problem

$k = m/n$

↓

Solution of a single equation
in one variable k

Step I: Dimension Reduction
(Prop. 2, Prop. 3)

Step II: Beta Representation
(Lem. 1, Lem. C.2)

Step III: Asymptotic Behavior
(Lem. C.3, Lem. C.4, Lem. 2)

Step IV: Analysis
(Prop. 4, Thm. 4)

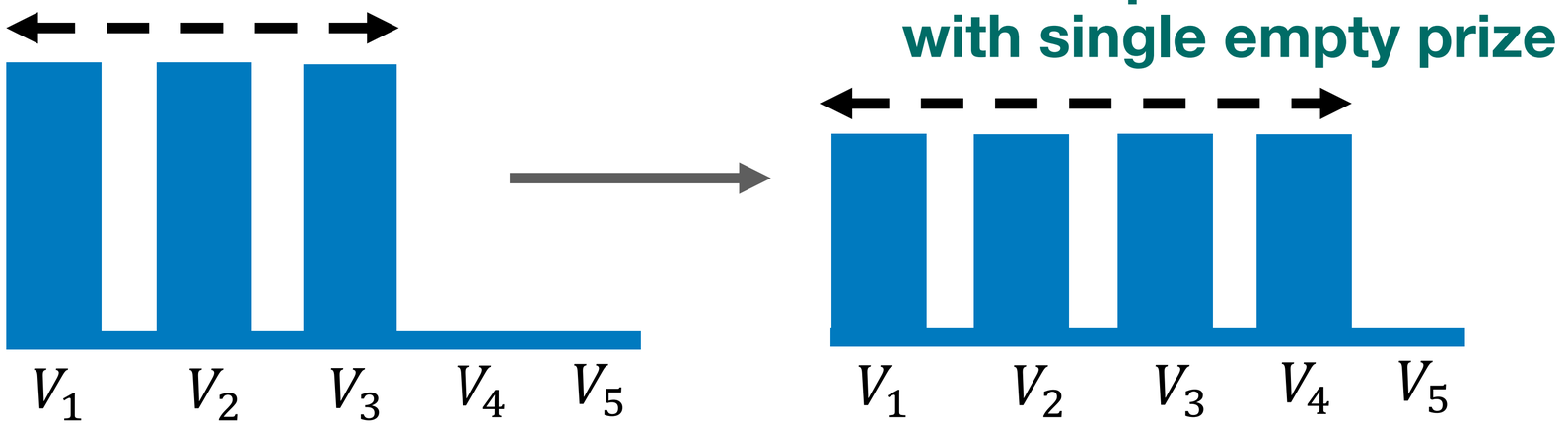
Total Effort

Step I: Dimension Reduction

Definition 3 (Complete Simple Contest)

A simple contest is complete, if it's a non-trivial simple contest with shortlist size m that has $m-1$ prizes, i.e., $V_1 = \cdots = V_{m-1} > V_m = 0$.

$m = 5$



Total Effort

Still intractable even as a univariate optimization problem.

Current Problem

$$\arg\max_{m \le n} \mathbb{E}_{X_{(1)},\ldots,X_{(m)}} \left[\sum_{i=1}^{m} \int_0^{x_{(i)}} \frac{\binom{n-1}{m-1} F^{n-m}(t)(1-F(t))^{m-1}}{\sum_{j=1}^{m} \binom{n-1}{j-1} F^{n-j}(t)(1-F(t))^{j-1}} \frac{f(t)}{1-F(t)} t \, dt \right]$$

Where $X_{(i)}$ is the i^{th}-highest ability level among all contestants, and $x_{(i)}$ is its realization. The prize is set as a complete simple contest, i.e., $l = m - 1$.

However, numerical experiments show strong evidences that the problem has **clean analytical solution** in asymptotic sense.

Total Effort

Step II: Beta Representation

Definition (Beta Distribution)

$\beta(x,a,b) = x^{a-1}(1-x)^{b-1}B(a,b)^{-1}$ is the PDF of beta distribution on support $[0,1]$, where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the beta function.

Current Problem

$$\arg\max_{m\le n} \mathbb{E}_{X_{(1)},\ldots,X_{(m)}} \left[\sum_{i=1}^{m} \int_0^{x_{(i)}} \frac{\binom{n-1}{m-1} F^{n-m}(t)(1-F(t))^{m-1}}{\sum_{j=1}^{m} \binom{n-1}{j-1} F^{n-j}(t)(1-F(t))^{j-1}} \frac{f(t)}{1-F(t)} t\, dt \right]$$

Where $X_{(i)}$ is the i^{th}-highest ability level among all contestants, and $x_{(i)}$ is its realization. The prize is set as a complete simple contest, i.e., $l = m-1$.

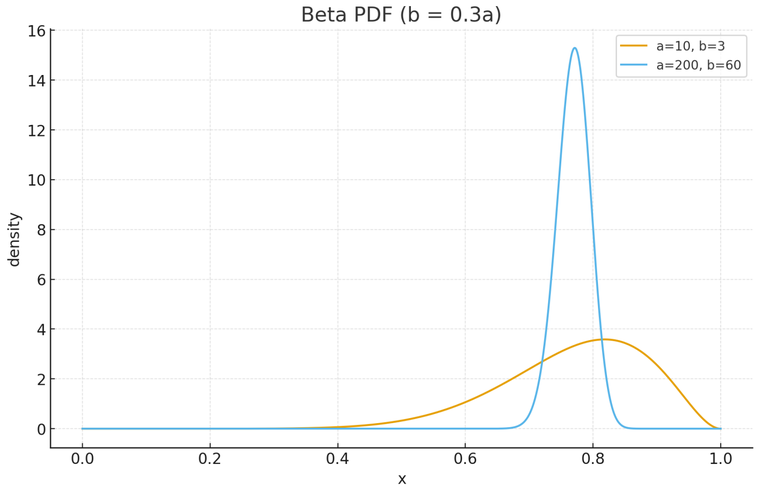
Total Effort

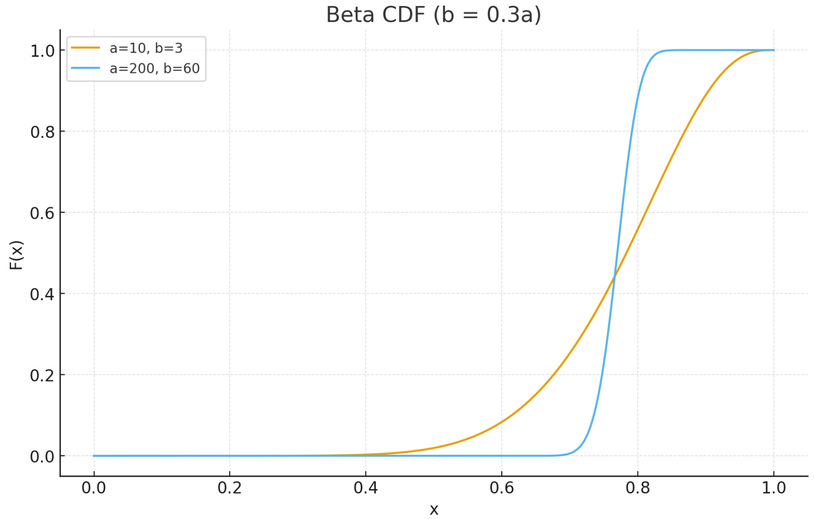
Step II: Beta Representation

Lemma C.2

Assume $\beta(q, n-m+1, m+1)$ attain it maximum at $q=\mu$, then when $|q-\mu|>\delta$, we have $\beta(q, n-m+1, m+1) \leq \epsilon$,

where $\delta=\sqrt{\frac{k(1-k)}{n}(\ln n-2\ln\epsilon-\ln 2\pi k(1-k))}=\Theta(\sqrt{\frac{\ln n}{n}})$.





Total Effort

Step II: Beta Representation

Lemma 1

Total Effort in a complete simple contest, denoted by $S(m,n)$, can be expressed using the beta distribution $\beta(x,a,b)$, as follows:

$$S(m,n) = \int_0^1 F^{-1}(q)\beta(q,n-m+1,m)\,dq + \int_0^1 F^{-1}(q)\frac{q}{1-q}\frac{m}{n-m}\frac{\beta(q,n-m,m)}{\int_0^q \beta(x,n-m,m)\,dx}\int_q^1 \beta(x,n-m,m+1)\,dx\,dq,$$

Total Effort

Step III: Asymptotic Behavior

Lemma 2

The ex-ante total effort in a complete simple contest has the following asymptotic expression with respect to n:

$$\lim_{n\to+\infty}\frac{S(m,n)}{n}=\int_0^{1-k}F^{-1}(q)\frac{q}{1-q}\frac{k}{1-k}\left(\frac{1}{q}-\frac{k}{q(1-q)}\right)dq,$$

where $k=m/n$, and the convergence rate is independent of k.

Moreover, it can be shown that $S(m,n)=o(n)$ if $m=o(n)$, therefore **it suffice to find the best admission rate $k\in(0,1]$**.

Total Effort

Step IV: Analysis

Proposition 4

The optimal shortlist size grows asymptotically linearly with n. Formally, there exists a $k^* \in (0,1)$ such that $\lim_{n\to\infty} m^*(n)/n = k$, where k^* is the solution of the following equation:

$$\int_k^1 F^{-1}(1-q)\left(\frac{1}{q} - (2k-k^2)\frac{1}{q^2}\right)dq = 0.$$

We can instantly get the optimal contest design for any ability distribution, by solving a single equation in one variable.

e.g.

$X \sim \mathrm{Exp}(\lambda) \to k^* \approx 9.70\%$

$X \sim \mathrm{U}[0,1] \to k^* \approx 15.07\%$

$f(x) = 2x \to k^* \approx 20.67\%$

Convexity?

Total Effort

Step IV: Analysis

Theorem 4

For an arbitrary ability distribution F, when $n \to \infty$, the optimal shortlist size has the following linear upper bound with respect to n:

$$\lim_{n\to\infty} \frac{m^*(n)}{n} \leq \bar{k},$$

where $\bar{k} \approx 31.62\%$ is the solution to the equation $\ln k = (2-k)(k-1)$.

Moreover, there exists a distribution such that $\frac{m^*(n)}{n} = \bar{k}$, meaning that the bound is tight.

Proof Sketch

1. $\bar{k}$ is an upperbound: Upperbound first-order condition given by Prop. 4 and show no zeros appears after $\bar{k}$.

2. $\bar{k}$ is binding: Construct a sequence of beta distribution that converge towards a bar distribution, with $k \to \bar{k}$.

Total Effort

Step IV: Analysis

Theorem 4

For an arbitrary ability distribution F, when $n \to \infty$, the optimal shortlist size has the following linear upper bound with respect to n:

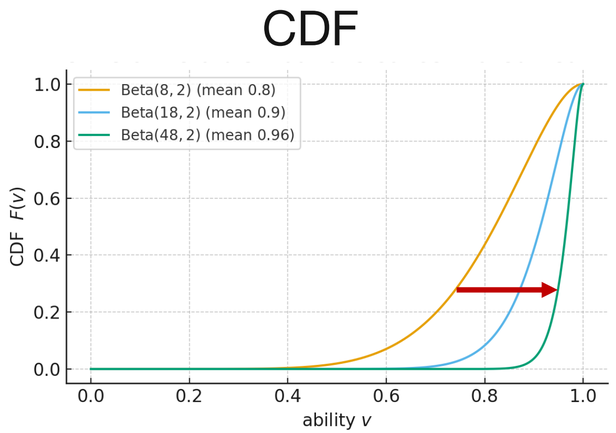
$$\lim_{n\to\infty} \frac{m^*(n)}{n} \leq \bar{k},$$

where $\bar{k} \approx 31.62\%$ is the solution to the equation $\ln k = (2-k)(k-1)$.

Moreover, there exists a distribution such that $\frac{m^*(n)}{n} = \bar{k}$, meaning that the bound is tight.

Worst Case

Contestants are equally strong



Takeaway

1. **Theoretical Aspect**
   Shortlist can be a powerful tool in mechanism design.
   (In our case, improvements can be $\Theta(n)$ times!)

2. **Practical Aspect**
   Only a small fraction of contestants can be enough.
   (First-round rejection in EC makes sense!)

   You can shortlist more aggressively if top candidates are sparse.
   (This justifies why top theory conferences generally have higher acceptance rate!)

Thanks!

27Fall PhD Applicant

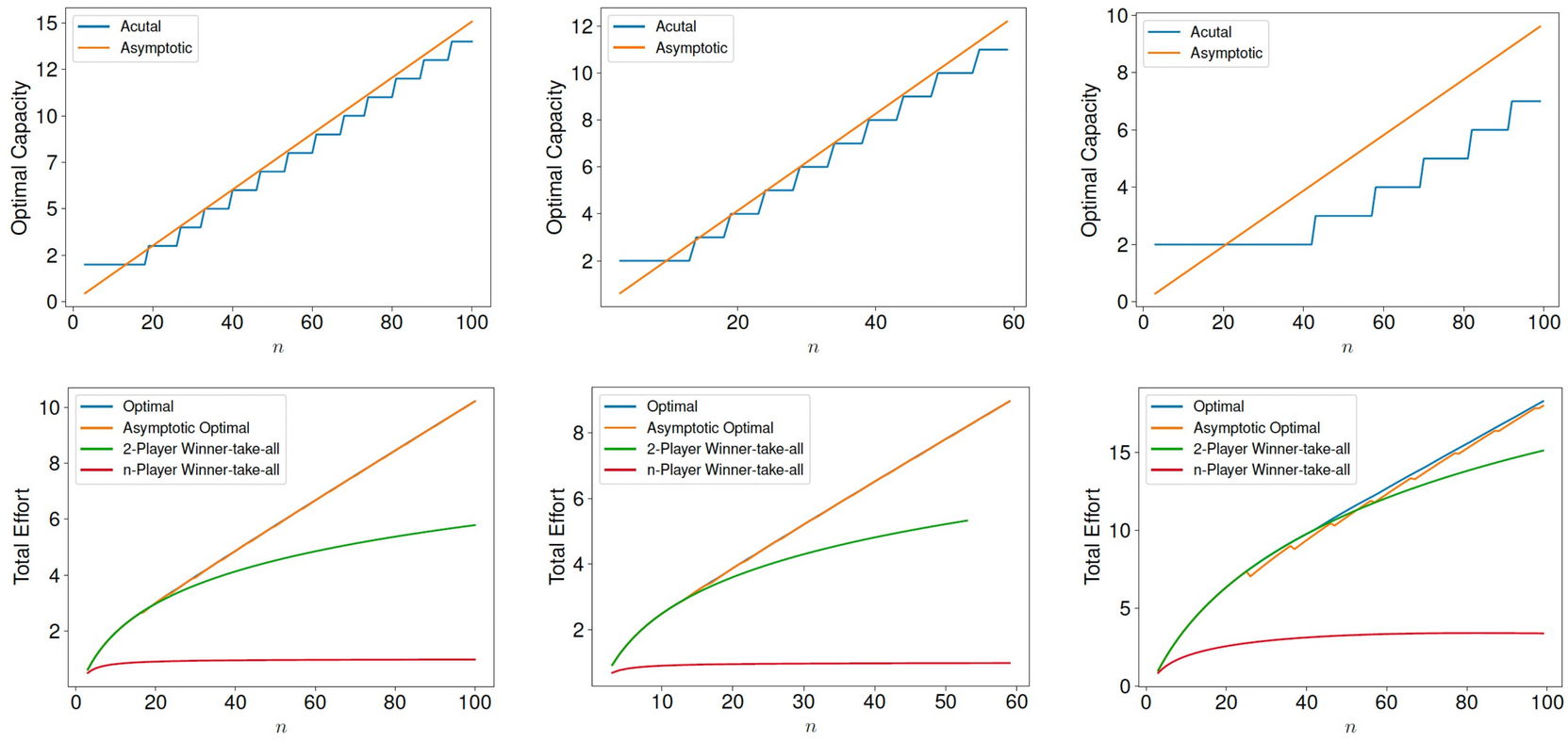
Solid background in LLM research

Boardly interested in AI Economics

Hanbing Liu,
Renmin University of China (RUC)

Towards Practical Applications

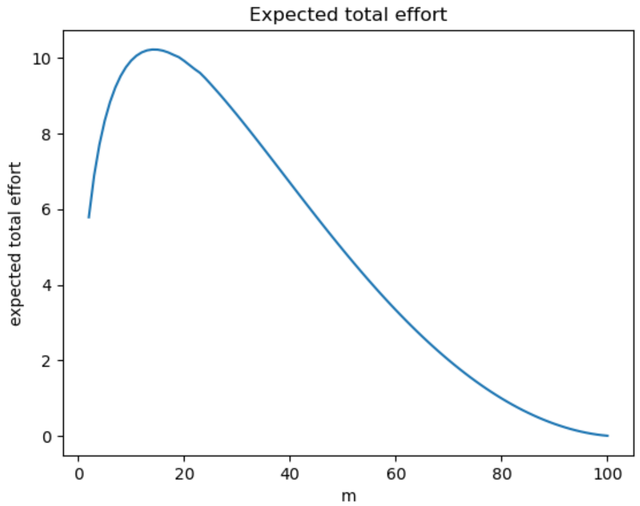
Asymptotic results perform well starting from small n.

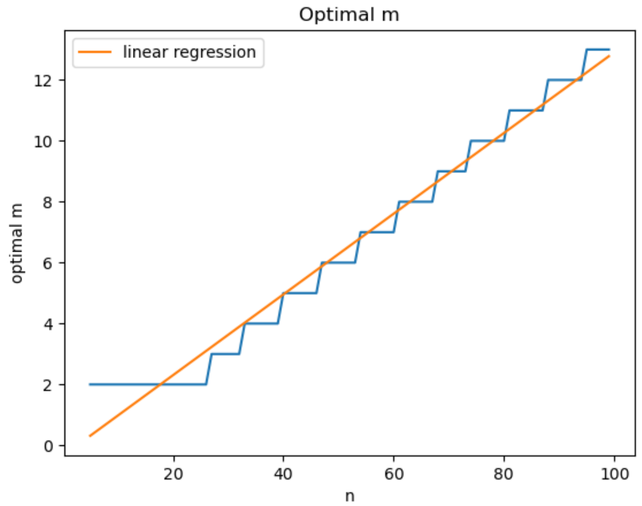


(a) $F(x) = x$ (b) $F(x) = x^2$ (c) $F(x) = 1 - e^{-x}$

Total Effort

$X \sim U[0,1]$





Key Observations

1. Total Effort is **unimodal** as a function of shortlist size, and the peak arrives early.

2. The optimal size grows **linearly** with contestant numbers, and the slope is determined by ability distribution.

Related Literature

Rank-order Contests

- Equilibrium analysis [Barut and Kovenock, 98, Baye et al., 96, 12];
- Different contestant characteristics [Glazer and Hassin, 88, Moldovanu and Sela, 01, Che and Gale, 03];
- Crowdsourcing scenario [Archak and Sundararajan, 09, Chawla et al., 19];
- Special configuration [Ghosh and Kleinberg, 16, Liu et al., 18, Liu and Lu, 23, Elkind et al., 21]

Elimination and Signaling in Contests

- Multi-stage contest [Moldovanu and Sela, 06, Fu and Lu, 12, Robertson, 24];
- Information design [Lu et al., 18, Levy et al., 19, Fu and Wu, 22, Chen, 24, Chen et al., 17, Kuang et al., 24];
- Equlibrium analysis of the same model [Sun et al. 24]

Future Work

1. What is the optimal contest with a shortlist for total effort when cost function $g(\cdot)$ is more general?
2. What will the approximation ratio be, when cost function $g(\cdot)$ is more general?
3. How do contestant equilibria and optimal contest design, if there are multiple contests with a shortlist?