IMPERIAL

Making decisions when you do not trust your forecasts

Pierre Pinson

Imperial College London, Dyson School of Design Engineering

mail: p.pinson@imperial.ac.uk - **webpage:** www.pierrepinson.com

CMAF Friday Forecasting Talks – 24 May 2024

Forecasts as input to decision-making

runwayml

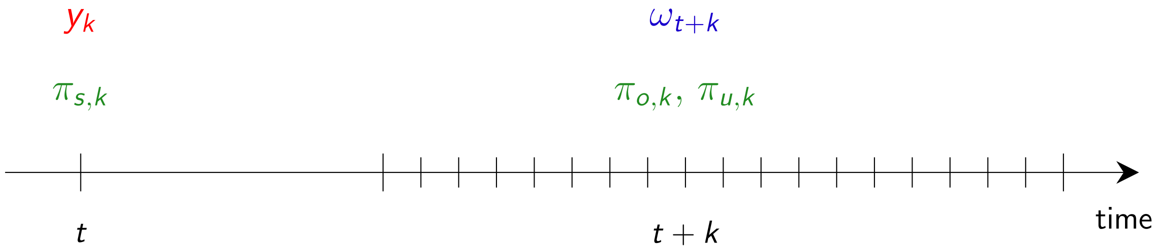
Outline

1. Renewables (wind) offering in electricity markets
2. No forecasts
3. Deterministic forecasts
4. Probabilistic forecasts
5. Concluding thoughts

... though when using forecasts, we will *not necessarily trust* them...

1. **Renewables (wind) offering in electricity markets**

Timeline, important variables and parameters



y_k: Decision at time t for time $t+k$ (energy sold in day-ahead market)

$\pi_{s,k}$: Price per unit when selling energy at time t

ω_{t+k}: Random variable for renewable energy generation at time $t+k$
(which we normalise, to take values in [0,1])

$\pi_{o,k},\ \pi_{u,k}$: Penalties (price per unit) if over- (resp. under-) producing compared to contract y_k
(based on realisation of ω_{t+k})

Revenue function

Let us focus on a given market time unit $t + k$ (e.g., 12-1pm tomorrow – and drop time-related notations)

Write

- π_s the day-ahead price
- π_o the unit penalty if over-producing (referred to as *overage* penalty)
- π_u the unit penalty if under-producing (referred to as *underage* penalty)

The revenue for $t + k$ (after some reformulation, to be maximised) is given by

$$\mathcal{R}(y, \omega, \pi_o, \pi_u) = \pi_s \omega - \pi_o(\omega - y)_+ - \pi_u(y - \omega)_+$$

This yield the equivalent loss function (to be minimised)

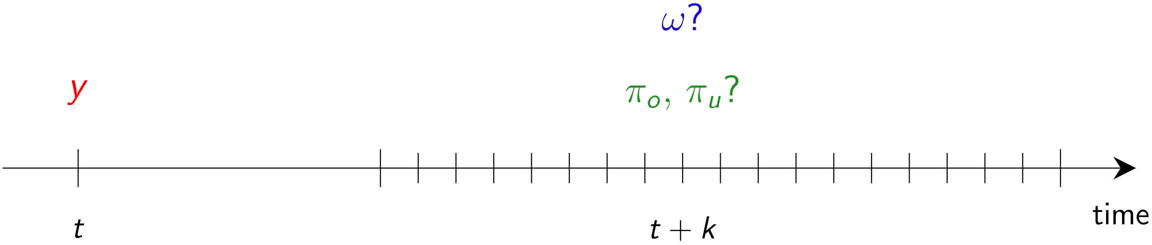
$$\mathcal{L}(y, \omega, \pi_o, \pi_u) = \pi_o(\omega - y)_+ + \pi_u(y - \omega)_+$$

2. **No forecasts**

(the most interesting case...!)

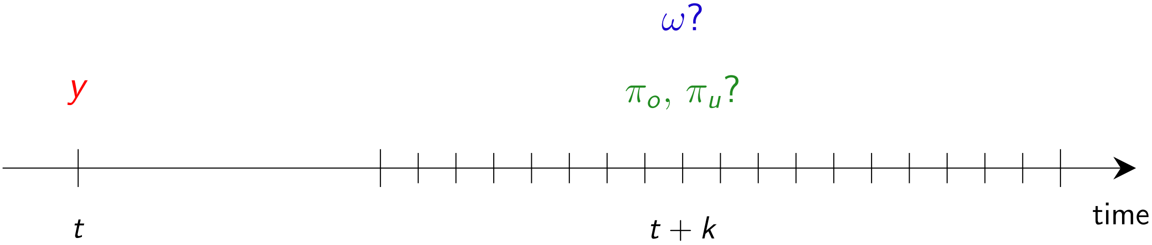
Blind decision-making (1)

What would you do, if you knew nothing about ω and about the penalties $\pi_o,\ \pi_u$??



Blind decision-making (1)

What would you do, if you knew nothing about ω and about the penalties $\pi_o,\ \pi_u$??



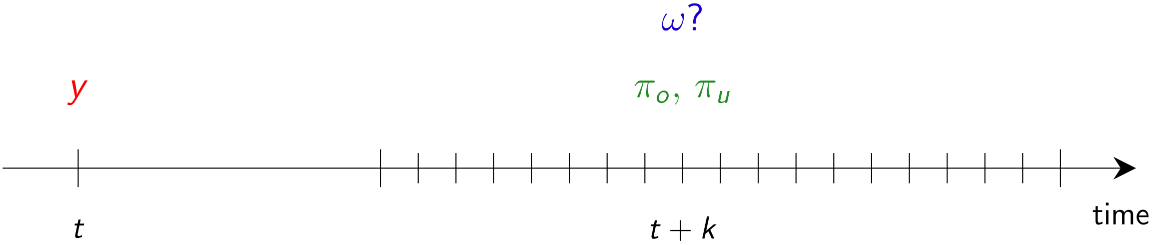
Implicit assumptions:

- Uninformative prior on ω: $\omega \sim U[0,1]$
- Uninformative prior on $\pi_o,\ \pi_u$: $\pi_o = \pi_u$

The minimax regret solution is $y = \frac{1}{2}$

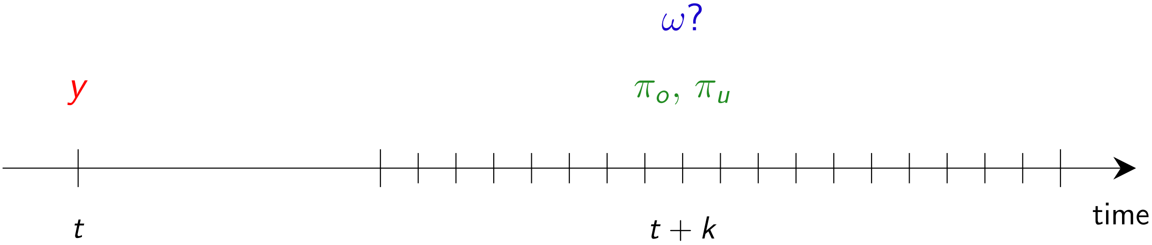
Energy-blind decision-making

What would you do, if you knew nothing about ω, but you knew something about the penalties $\pi_o,\ \pi_u$??



Energy-blind decision-making

What would you do, if you knew nothing about ω, but you knew something about the penalties $\pi_o,\ \pi_u$??



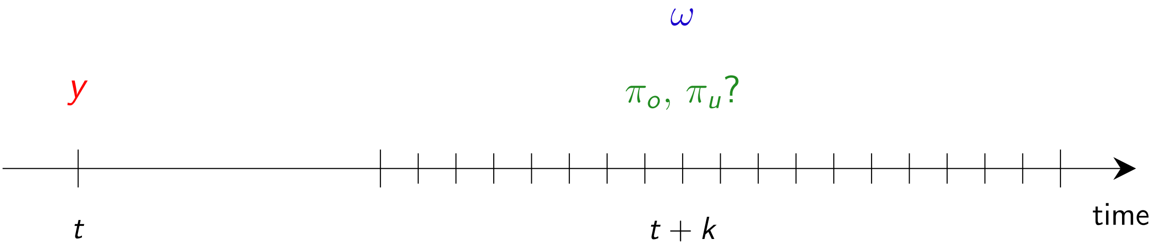
Implicit assumptions:

- Uninformative prior on ω: $\omega \sim U[0,1]$
- Perfect knowledge of $\pi_o,\ \pi_u$

The minimax regret solution is $y = \dfrac{\pi_o}{\pi_o + \pi_u}$

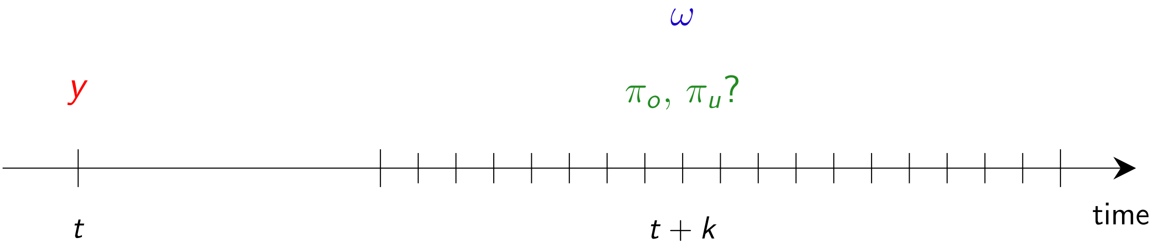
Penalty-blind decision-making

What would you do, if you knew the distribution (or the expectation) of ω, but you knew nothing about the penalties $\pi_o,\ \pi_u$??



Penalty-blind decision-making

What would you do, if you knew the distribution (or the expectation) of ω, but you knew nothing about the penalties $\pi_o,\ \pi_u?$?



Implicit assumptions:

- Perfect knowledge of $F(\omega)$, or of $\mathbb{E}[\omega]$
- No information (whatsoever) about $\pi_o,\ \pi_u$

The minimax regret solution is $y = \mathbb{E}[\omega]$

3. **Deterministic forecasts**

(the least interesting case... use decision rules!)

4. **Probabilistic forecasts**

(the most fun case... though technical!)

The newsvendor problem

- The newsvendor problem is one of the most classical problems in **stochastic optimization** (or **statistical decision theory**)

- It can be traced back to:

FY Edgeworth (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society* **51**(1): 113–127

(even though in this paper the problem is about how much a bank should keep in its reserves to satisfy request for withdrawal (i.e., the *bank-cash-flow problem*))

- It applies to varied problems as long as:
  - one shot possibility to decide on the quantity of interest
  - outcome is uncertain
  - known marginal profit and loss
  - the aim is to maximize expected profit!

Optimal offering as a price-taker

Let us focus on a given market time unit $t+k$ (e.g., 12-1pm tomorrow – and drop time-related notations)

We have

- $\hat{F}_\omega$ the predictive CDF for renewable energy generation
- π_o the unit penalty if over-producing (referred to as *overage* penalty)
- π_u the unit penalty if under-producing (referred to as *underage* penalty)

The loss function for $t+k$ (to be minimised) is

$$\mathcal{L}(y, \omega, \pi_o, \pi_u) = \pi_o(\omega - y)_+ + \pi_u(y - \omega)_+$$

The expected utility maximization offer is

$$y^* = \underset{y}{\operatorname{argmin}}\, \mathbb{E}_\omega \left[\mathcal{L}(y, \omega, \pi_o, \pi_u)\right] = \hat{F}_\omega{}^{-1}\left(\frac{\pi_o}{\pi_o + \pi_u}\right)$$

In practice, π_o and π_u are replaced by estimates (/forecasts).

Generalizing the newsvendor problem

- The overage π_o and underage π_u penalties are unknown!
- Instead of parameters, they should be seen as random variables (now, π_o and π_u)
- Only one of them is active in most cases (two-price imbalance settlement)

One can then define the penalization as a the outcome of the Bernoulli variable s (with chance of success τ):

$$s = \frac{\pi_o}{\pi_o + \pi_u}$$

where

$$\begin{aligned} \pi_o &= (\pi_s - \pi_b)\,\mathbf{1}_{\{s_L \geq 0\}} \\ \pi_u &= (\pi_b - \pi_s)\,\mathbf{1}_{\{s_L < 0\}} \end{aligned}$$

(π_b the balancing price, s_L the overall system imbalance)

s has only 2 potential discrete outcomes, i.e.

$$\begin{aligned} (i) \quad & \pi_o \neq 0,\ \pi_u = 0 \ \Rightarrow \ s = \frac{\pi_o}{\pi_o + \pi_u} = 1,\ 1 - s = 0 \\ (ii) \quad & \pi_o = 0,\ \pi_u \neq 0 \ \Rightarrow \ s = \frac{\pi_o}{\pi_o + \pi_u} = 0,\ 1 - s = 1 \end{aligned}$$

The Bernoulli newsvendor problem

Definition (Bernoulli newsvendor problem)

Based on a Bernoulli random variable s (with chance of success τ), and the random variable ω (with c.d.f. F_ω), the decision y^* minimizing the expected opportunity cost function is

$$y^* = \underset{y}{\operatorname{argmin}}\ \mathbb{E}_{\omega,s}\left[\mathcal{L}(y,\omega,s)\right] , \tag{1a}$$

where the opportunity cost is defined as

$$\mathcal{L}(y,\omega,s) = s(\omega - y)_+ + (1-s)(y-\omega)_+ , \tag{1b}$$

and where $(.)_+$ is for the positive part.

The Bernoulli newsvendor problem

Definition (Bernoulli newsvendor problem)

Based on a Bernoulli random variable s (with chance of success τ), and the random variable ω (with c.d.f. F_ω), the decision y^* minimizing the expected opportunity cost function is

$$y^* = \underset{y}{\operatorname{argmin}}\ \mathbb{E}_{\omega,s}\left[\mathcal{L}(y,\omega,s)\right] , \tag{1a}$$

where the opportunity cost is defined as

$$\mathcal{L}(y,\omega,s) = s(\omega - y)_+ + (1-s)(y-\omega)_+ , \tag{1b}$$

and where $(.)_+$ is for the positive part.

And its solution:

Proposition

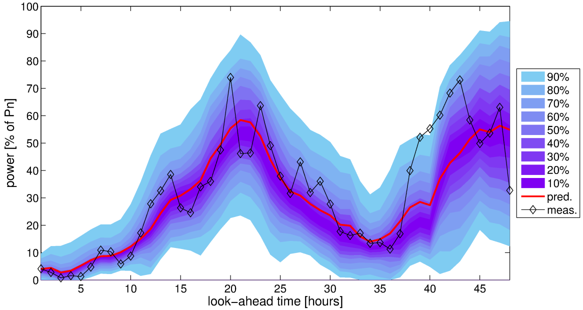
Consider F_ω the c.d.f. for the uncertain parameter ω and τ the chance of success for s. The optimal decision y^* for the Bernoulli newsvendor problem (1a) *is*

$$y^* = F_\omega{}^{-1}(\tau) .$$

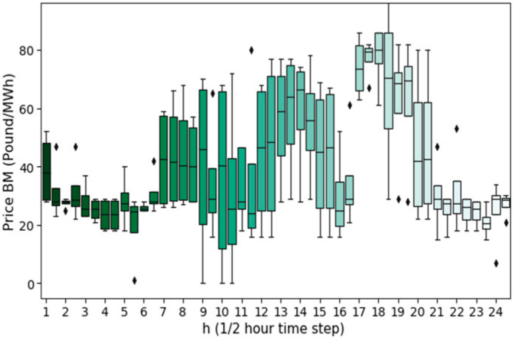
Problem setup

- F_ω and F_s are still unknown and need to be replaced by forecasts, $\hat{F}_\omega$ and $\hat{F}_s$
- These forecasts are necessarily imperfect...

Both renewable energy generation and balancing market outcomes are notoriously difficult to predict...



$\hat{F}_\omega$



(input to obtain) $\hat{F}_s$
(reproduced from Lucas *et al.* (2020))

Ambiguity about $\hat{F}_\omega$

We adapt here the solution approach of Fu *et al.* (2021) to the Bernoulli newsvendor problem...

Definition (distributionally robust Bernoulli newsvendor problem – ambiguity about $\hat{F}_\omega$)

Consider a Bernoulli random variable s with estimated chance of success $\hat{\tau}$, the uncertain production ω with predictive c.d.f. $\hat{F}_\omega$, and an ambiguity set $\mathcal{B}_{\hat{F}_\omega}(\rho)$ with radius ρ. The distributionally robust Bernoulli newsvendor problem, with ambiguity about $\hat{F}_\omega$, is that for which the decision y^* is given by

$$y^* = \operatorname*{argmin}_{y} \sup_{F_\omega \in \mathcal{B}_{\hat{F}_\omega}(\rho)} \mathbb{E}_{\omega,s}\left[\mathcal{L}(y,\omega,\hat{\tau})\right] . \tag{2}$$

How to define $\mathcal{B}_{\hat{F}_\omega}(\rho)$?

Deformation operator and FSD-ambiguity set

A first-order stochastic dominance ambiguity set (FSD-ambiguity set) is such that

$$\underline{F}_\omega(x) \leq F_\omega(x) \leq \overline{F}_\omega(x), \quad \forall x,\ \forall F_\omega \in \mathcal{B}_{\hat{F}_\omega}(\rho)\,.$$

As an example, we introduce here a double-power deformation operator that fulfil the above definition.

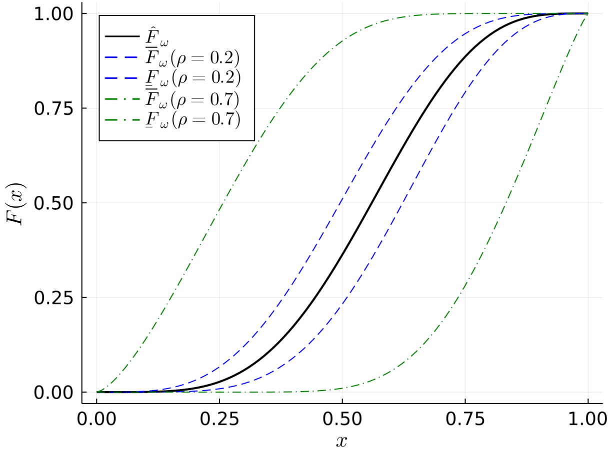
Definition (double-power deformation operator)

Consider a reference c.d.f. F_ω. The upper $\overline{\mathcal{O}}_\rho$ and lower $\underline{\mathcal{O}}_\rho$ double-power deformation operators are defined as

$$\overline{\mathcal{O}}_\rho(F_\omega) = \left(1-(1-F_\omega{}^{\frac{1}{1-\rho}})\right)^{1-\rho}\,,$$

$$\underline{\mathcal{O}}_\rho(F_\omega) = 1-(1-F_\omega{}^{\frac{1}{1-\rho}})^{1-\rho}\,,$$

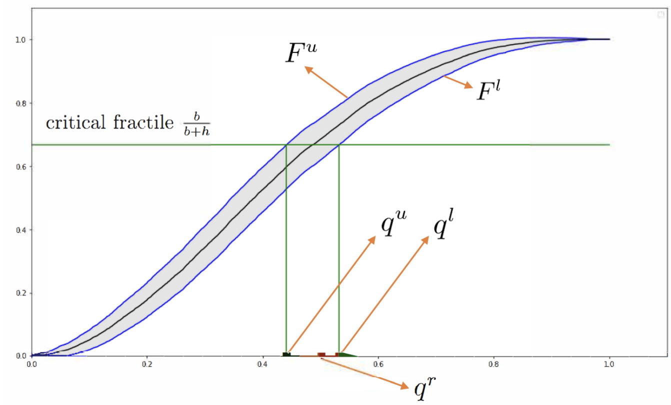
with ρ the deformation parameter.



These deformations readily allow to define $\underline{F}_\omega$ and $\overline{F}_\omega(x)$.

Worst-case distribution

The worst-case distribution F^{ws} for this distributionally robust problem is defined by



(reproduced from Fu *et al.* (2021))

$$F^{\text{ws}}(x) = \begin{cases} \overline{F}_\omega(x), & x < \overline{F}_\omega^{-1}(\hat{\tau}) \\ \tau, & \overline{F}_\omega^{-1}(\hat{\tau}) < x < \underline{F}_\omega^{-1}(\hat{\tau}) \\ \underline{F}_\omega(x), & x > \underline{F}_\omega^{-1}(\hat{\tau}) \end{cases}$$

Distributionally robust solution

Theorem

Consider an FSD-ambiguity set defined by a ball $\mathcal{B}_{\hat{F}_\omega}(\rho)$ with radius ρ, yielding the two bounding distributions $\underline{F}_\omega$ and $\overline{F}_\omega$. For a predicted chance of success $\hat{\tau}$, the solution of the distributionally robust Bernoulli newsvendor problem (2) is

$$y^* = \hat{\tau}\underline{F}_\omega^{-1}(\hat{\tau}) + (1-\hat{\tau})\overline{F}_\omega^{-1}(\hat{\tau})\,.$$

Distributionally robust solution

Theorem

Consider an FSD-ambiguity set defined by a ball $\mathcal{B}_{\hat{F}_\omega}(\rho)$ with radius ρ, yielding the two bounding distributions $\underline{F}_\omega$ and $\overline{F}_\omega$. For a predicted chance of success $\hat{\tau}$, the solution of the distributionally robust Bernoulli newsvendor problem (2) is

$$y^* = \hat{\tau}\underline{F}_\omega^{-1}(\hat{\tau}) + (1-\hat{\tau})\overline{F}_\omega^{-1}(\hat{\tau}) .$$

For sufficiently large values of the radius ρ, we obtain the following (robust) limiting case:

Corollary

For sufficiently large values of ρ, the radius of $\mathcal{B}_{\hat{F}_\omega}(\rho)$, the solution of the distributionally robust Bernoulli newsvendor problem (2) converges to the robust solution $y^* = \hat{\tau}$.

Ambiguity about $\hat{F}_s$

The following solution is not available in the literature...

Definition (distributionally robust Bernoulli newsvendor problem – ambiguity about $\hat{F}_s$)

Consider a Bernoulli random variable s with estimated chance of success $\hat{\tau}$, the uncertain production ω with predictive c.d.f. $\hat{F}_\omega$, and an ambiguity set for $\hat{F}_s$ defined by the ball $\mathcal{B}_{\hat{\tau}}(\varepsilon)$ with radius ε. The distributionally robust Bernoulli newsvendor problem, with ambiguity about $\hat{F}_s$, is that for which the decision y^* is given by

$$y^* = \underset{y}{\mathrm{argmin}} \max_{\tau \in \mathcal{B}_{\hat{\tau}}(\varepsilon)} \mathbb{E}_{\omega,s}\left[\mathcal{L}(y, \omega, \tau)\right]$$

Ambiguity about $\hat{F}_s$

The following solution is not available in the literature...

Definition (distributionally robust Bernoulli newsvendor problem – ambiguity about $\hat{F}_s$)

Consider a Bernoulli random variable s with estimated chance of success $\hat{\tau}$, the uncertain production ω with predictive c.d.f. $\hat{F}_\omega$, and an ambiguity set for $\hat{F}_s$ defined by the ball $\mathcal{B}_{\hat{\tau}}(\varepsilon)$ with radius ε. The distributionally robust Bernoulli newsvendor problem, with ambiguity about $\hat{F}_s$, is that for which the decision y^* is given by

$$y^* = \operatorname*{argmin}_{y} \max_{\tau \in \mathcal{B}_{\hat{\tau}}(\varepsilon)} \mathbb{E}_{\omega,s}\left[\mathcal{L}(y, \omega, \tau)\right]$$

How to define $\mathcal{B}_{\hat{\tau}}(\varepsilon)$?

Definition (uniform and level-adjusted ambiguity sets for $\hat{F}_s$)

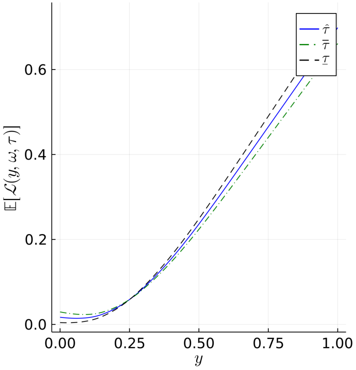
Given a ball radius ε, a uniform ambiguity set for $\hat{F}_s$ is defined by the ball $\mathcal{B}_{\hat{\tau}}$ with

$$\begin{aligned} \underline{\tau} &= \max(\hat{\tau} - \varepsilon, 0) \\ \overline{\tau} &= \min(\hat{\tau} + \varepsilon, 1) \end{aligned}$$

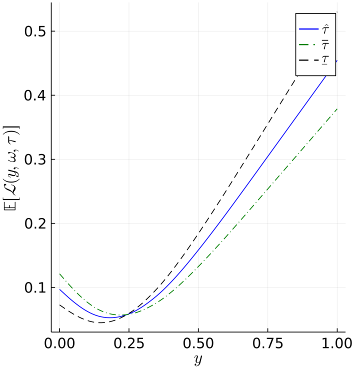
while a level-adjusted ambiguity set (with parameter $\theta \in [0,1)$) for $\hat{F}_s$ is defined by the ball $\mathcal{B}_{\hat{\tau}}$ with

$$\begin{aligned} \underline{\tau} &= \max(\hat{\tau} - \varepsilon(1 - 4\,\hat{\tau}(1-\hat{\tau})), 0) \\ \overline{\tau} &= \min(\hat{\tau} + \varepsilon(1 - 4\,\hat{\tau}(1-\hat{\tau})), 1) \end{aligned}$$

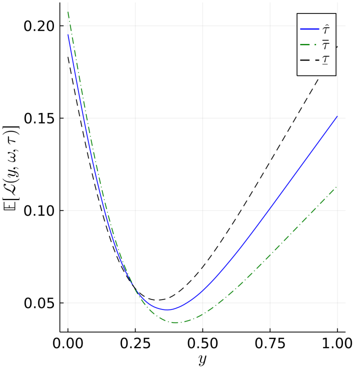
Intuition for the upcoming result



(a) $\hat{\tau} = 0.1$, $\varepsilon = 0.05$



(b) $\hat{\tau} = 0.5$, $\varepsilon = 0.1$



(c) $\hat{\tau} = 0.8$, $\varepsilon = 0.05$

Figure: Expected opportunity cost, as a function the decision y, for different values of $\hat{\tau}$ and ε (using uniform ambiguity sets). The uncertain parameter ω follows a Beta(2,6) distribution (with expected value $\mathbb{E}[\omega] = 0.25$), while the estimate $\hat{\tau}$ is based on 15 samples.

Distributionally robust solution

Theorem

Consider an ambiguity set for $\hat{F}_s$ defined by a ball $\mathcal{B}_{\hat{\tau}}(\varepsilon)$ with radius ε, and the predictive c.d.f. $\hat{F}_\omega$ for the random variable ω. The solution of the distributionally robust Bernoulli newsvendor problem (6) *is*

$$
\begin{aligned}
y^* &= \hat{F}_\omega{}^{-1}(\overline{\tau})\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\overline{\tau})<\mathbb{E}[\omega]\}} + \hat{F}_\omega{}^{-1}(\underline{\tau})\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\underline{\tau})>\mathbb{E}[\omega]\}} \\
&\quad + \mathbb{E}[\omega]\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\overline{\tau})\geq\mathbb{E}[\omega]\}}\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\underline{\tau})\leq\mathbb{E}[\omega]\}}\,.
\end{aligned}
$$

Distributionally robust solution

Theorem

Consider an ambiguity set for $\hat{F}_s$ defined by a ball $\mathcal{B}_{\hat{\tau}}(\varepsilon)$ with radius ε, and the predictive c.d.f. $\hat{F}_\omega$ for the random variable ω. The solution of the distributionally robust Bernoulli newsvendor problem (6) *is*

$$y^* = \hat{F}_\omega{}^{-1}(\overline{\tau})\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\overline{\tau})<\mathbb{E}[\omega]\}} + \hat{F}_\omega{}^{-1}(\underline{\tau})\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\underline{\tau})>\mathbb{E}[\omega]\}} + \mathbb{E}[\omega]\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\overline{\tau})\geq\mathbb{E}[\omega]\}}\,\mathbf{1}_{\{\hat{F}_\omega{}^{-1}(\underline{\tau})\leq\mathbb{E}[\omega]\}}\,.$$

And, for sufficiently large values of ε, we obtain the following limiting (robust) case:

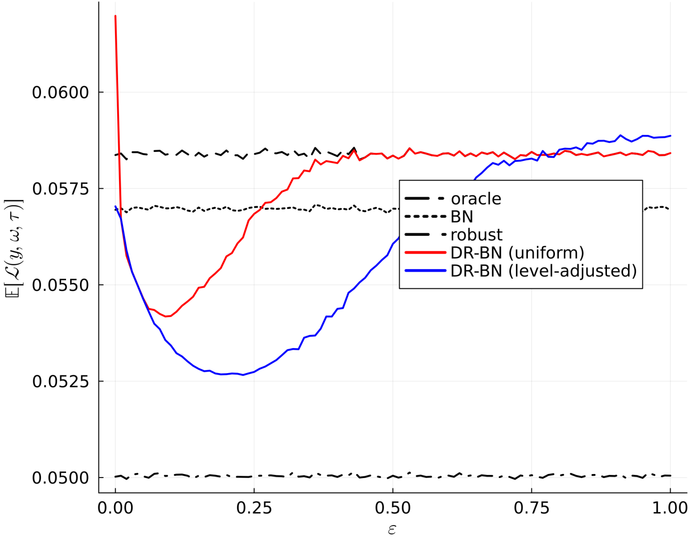
Corollary

For sufficiently large values of ε, the radius of $\mathcal{B}_{\hat{\tau}}(\varepsilon)$, the solution of the distributionally robust Bernoulli newsvendor problem (6) *converges to the robust solution $y^* = \mathbb{E}[\omega]$.*

5. **Simulations and case-study application**

Ambiguity about $\hat{F}_s$ - impact of ball radius ε

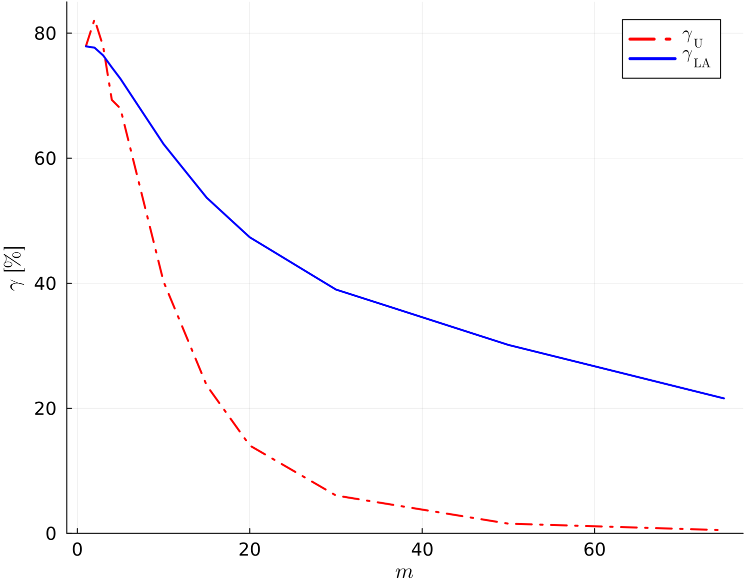
- Monte-Carlo simulation with $N = 10^7$ replicates
- $\omega \sim \text{Beta}(2, 6)$
- $\tau = 0.75$
- $m = 10$ (number of draws of a Bern(τ) to obtain estimate $\hat{\tau}$)
- $\theta = 0.9$ (for the level-adjusted ball)



The performance measure γ (in %) is then defined as $\gamma = \frac{L_{\text{BN}} - L^*_{\text{DR-BN}}}{L_{\text{BN}} - L_{\text{O}}}$, and can be expressed in percents. For the above example, $\gamma_{\text{U}} = 40.3\%$ and $\gamma_{\text{LA}} = 60.2\%$.

Ambiguity about $\hat{F}_s$ - impact of forecast quality

- Monte-Carlo simulation with $N = 10^7$ replicates
- m (number of draws of a Bern(τ) to obtain estimate $\hat{\tau}$) is a proxy for forecast quality
- $\omega \sim$ Beta$(2, 6)$
- $\tau = 0.75$
- $\theta = 0.9$ (for the level-adjusted ball)

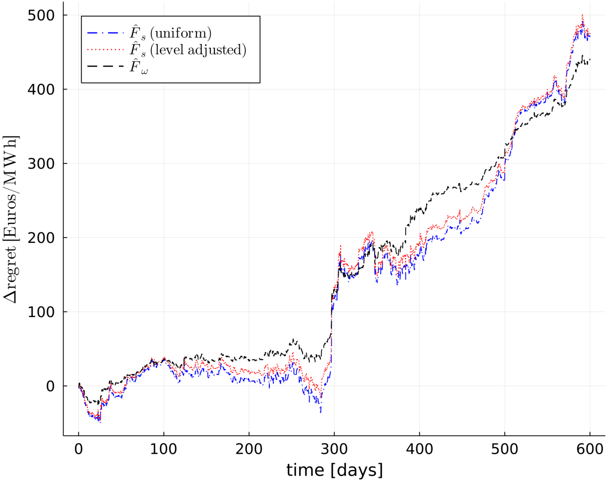


Application to French electricity market

Portfolio of wind farms from Midwest France (confidential) over a 2-year period

- first 131 days for warm start and cross-validation (to decide on m, ρ, ε, θ)
- remaining 600 days for genuine out-of-sample evaluation

Approach	R [€/MWh]	r [€/MWh]
Oracle	31.63	0
BN	29.25	2.38
DR-BN ($\hat{F}_w$)	29.34	2.29
DR-BN ($\hat{F}_s$, uniform)	29.35	2.28
DR-BN ($\hat{F}_s$, level-adj.)	29.36	2.27



Cumulative regret normalized per MWh produced

5. **Concluding thoughts**

- Making decisions from forecasts is not always obvious
- Various aspects of forecast quality are of relevance
- Trust (or lack of) in the forecasts can be handled in a data-driven manner

Thanks for your attention!