CESA 2006

Using the Intermeans Parameter to Model the Dispersion of Demand

pierre.douillet@ensait.fr

besoa.rabenasolo@ensait.fr

École Nationale Supérieure des Arts et Industries Textiles

Roubaix, France

the newsboy paradigm

introduction

- (Scarf's notations) y : order quantity, $\Phi(\xi)$: demand cdf, c : unit cost, r : unit selling price
- unsold units are discarded
- the satisfied demand is : $\xi^+ = \min(y, \xi)$
- naive solution : $\widehat{G} \doteq G(\mu, \mu)$... but actual gain (ex post) :

$$G(y, \xi) = r\,\xi^+ - c\,y$$

cost of a given choice

- define $\theta_y \doteq \int_y^\infty \mathrm{d}\Phi\,(\xi)$ together with:

$$\xi_y^a \doteq \mathrm{E}\,(\xi \mid y < \xi) \quad ; \quad \xi_y^b \doteq \mathrm{E}\,(\xi \mid \xi < y)$$

- define $G\,(y,\,\Phi) \doteq \mathrm{E}\,(G\,(y,\,\xi))\,\xi$, obtain:

$$\widehat{G} - G\,(y,\,\Phi) =$$

$$\theta_y\,\left(\xi_y^a - y\right)\,(r - c) + (1 - \theta_y)\,\left(y - \xi_y^b\right)\,c$$

- and conclude:

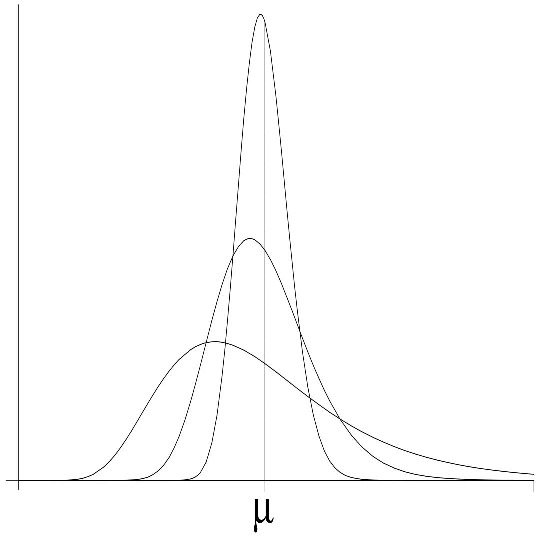
$$\forall\Phi\,\forall y\;:\;G\,(y,\,\Phi) \leq \widehat{G}$$

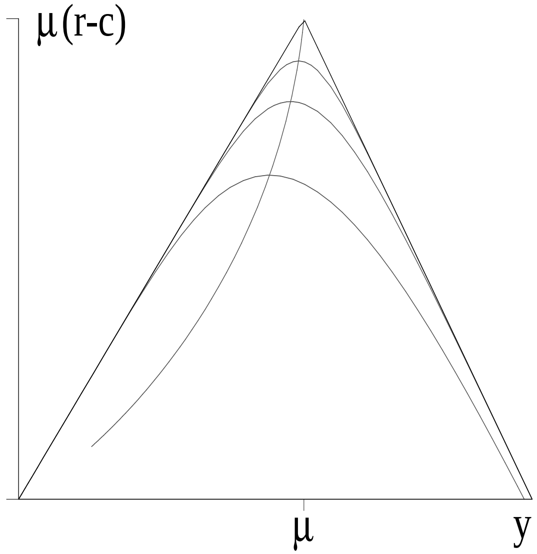
cost of uncertainties

- knowing Φ, we have $y^* = \arg\max_y G(y, \Phi)$
- exact analytical solution : $\theta_{y^*} = \theta_* = 1 - \Phi(y^*) = c/r$
- the corresponding (least) miss to gain can be rewritten as:

$$\widehat{G} - G(y^*, \Phi) = \theta_* (1 - \theta_*) \left(\xi_*^a - \xi_*^b\right) r$$

example : lognormal Φ, given μ





- positive values, but assume multiplicative independence
- curve $\sigma \mapsto (y^*, G^*)$, assuming $c/r < 1/2$

robust solutions

knowledge versus facilities

- probability distributions can be used to impersonate our actual knowledge about the real world... or about the limits of our knowledge
- but, too often, side assumptions are introduced that does not come from the actual framework, but only from computing easiness

basic questions

- does Φ model a lack of knowledge due e.g. to their cost or model the intrinsic wild behavior of the markets ?
- is Φ guessed from many parallel independent worlds or induced from historical data (questionable ergodicity) ?
- can ξ be ever measured, even afterwards, when the demand overflows the inventory ?
- robust solution against a family $\mathcal{F}$ of distributions :

$$G^* \doteq G\,(y^*,\,\mathcal{F}) \doteq \max_{y}\ \min_{\Phi\in\mathcal{F}} G\,(y,\,\Phi)$$

Scarf's theorem

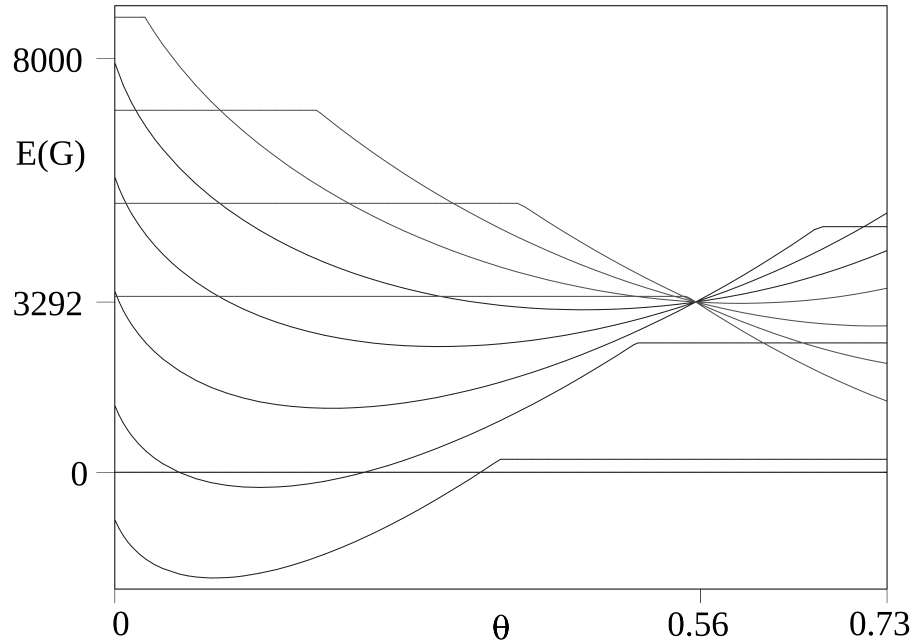
- when playing against $\mathcal{F}(\mu, \sigma)$, the worst case is a Dirac two points distribution where ξ is either ξ^a or ξ^b
- when $(c/r)^{-1} < 1 + \sigma^2/\mu^2$, then better buy nothing
- otherwise, the robust decision is:

$$\begin{cases} y^* = \mu + \sigma\left(r/2 - c\right)/\sqrt{c\left(r-c\right)} \\ G^* = \mu\left(r-c\right) - \sigma\sqrt{c\left(r-c\right)} \end{cases}$$

graphical proof

playing against the $\mathcal{F}(\mu, \sigma, Dirac)$ family is:

- chose an y, i.e. a curve
- wait for answer $G(y, \xi)$
- $\mathrm{E}(G)$ depends on θ



all curves are going through the same point $\theta = c/r$

the intermeans parameter

cost of mean

- we have $\widehat{G} - G\,(\mu,\,\Phi) = \theta_\mu\,(1-\theta_\mu)\left(\xi^a_\mu - \xi^b_\mu\right) \times r$
- this kind of factorization applies only to y^* from $\theta_* = c/r$ and to μ from the obvious $\forall y\ :\ \theta_y\,\xi^a_y + (1-\theta_y)\,\xi^b_y = \mu$
- thus $\widehat{G} - G^* \leq \delta\, r$, independent of c/r, where

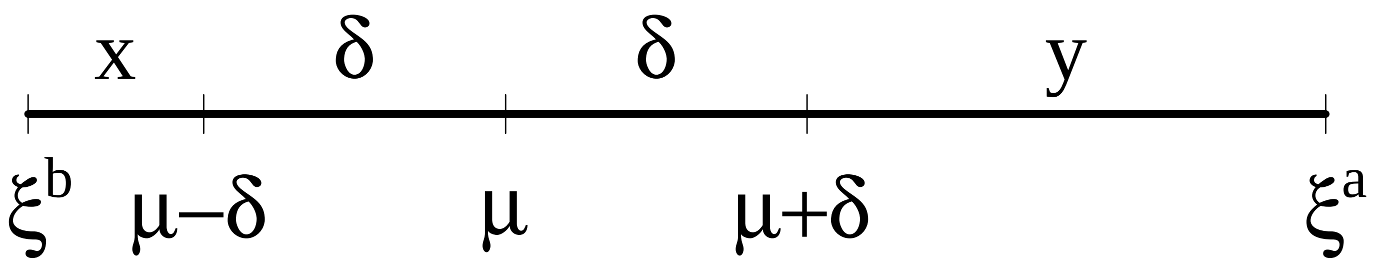
$$\delta \doteq \theta_\mu\,(1-\theta_\mu)\left(\xi^a_\mu - \xi^b_\mu\right)$$

a measure of dispersion

- from now on, all θ, ξ^a, ξ^b are relative to μ

$$\delta \doteq \theta\,(1-\theta)\left(\xi^a - \xi^b\right)$$

- this δ has some similarities with the interquartile range
- properties : $\xi^b < \mu - \delta < \mu + \delta < \xi^a$ and $x\,y = \delta^2$



comparison δ versus σ

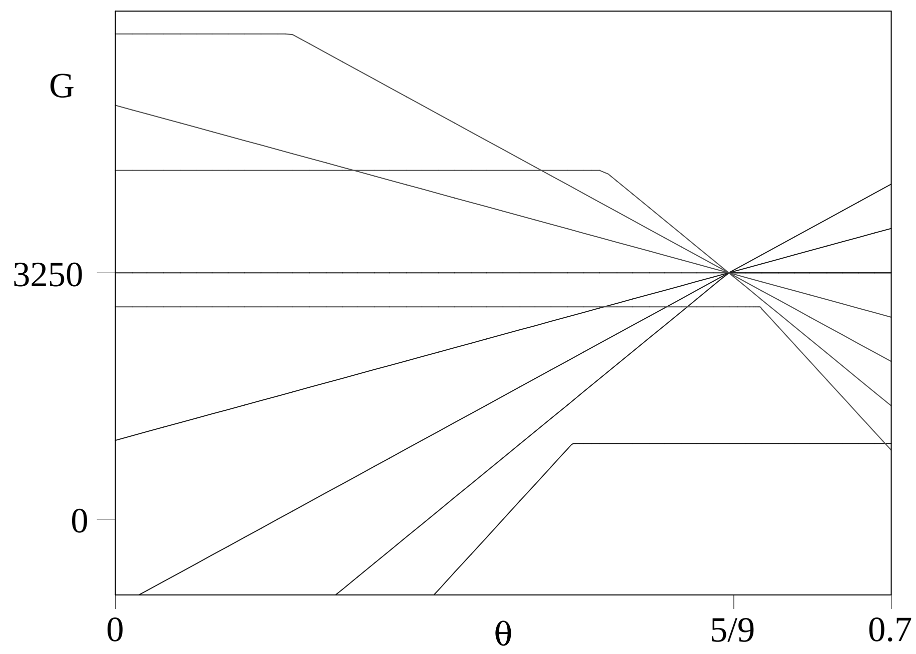
Φ	δ/σ(exact)	δ/σ(approx)	0.25 gap
uniform	$\sqrt{3}/4$	≈ 0.433	
triangular	$1/\sqrt{6} \cdots 8\sqrt{2}/27$	$.408 \cdots .419$	
normal	$1/\sqrt{2\,\pi}$	≈ 0.399	
exp	$1/\mathrm{e}$	≈ 0.368	
Dirac	$\sqrt{\theta\,(1-\theta)}$	$0 \cdots 0.5$	$\theta = 7\%,\ 93\%$
lognormal	$\leq 1/\sqrt{2\,\pi}$	$0 . \cdots 0.399$	$\sigma/\mu \approx 2$

in all realistic situations : $0.25\,\sigma \leq \delta \leq 0.50\,\sigma$

a surprising result

- when playing against $\mathcal{F}(\mu, \delta)$, the worst case is not necessarily a two points distribution
- when playing against $\mathcal{F}(\mu, \delta, Dirac)$, all curves are going through the same point $\theta = c/r$
- and now

$$G^*_{Dirac} = \mu$$



sampling properties

sampling properties of variance

- obtain sample ω by n independent drawings from Φ
- put $s^2 \doteq \text{var}(\omega)$ and define $S^2 \doteq s^2 \times n/(n-1)$. Then:

$$\text{E}\left(S^2\right) = \sigma^2 \quad ; \quad \text{var}\left(S^2\right) = \frac{1}{n}\left(M_4 - \frac{n-3}{n-1}\,\sigma^4\right)$$

$$\text{n} \times \frac{\text{var}}{\text{E}^2}\left(S^2\right) = \left(\frac{M_4}{\sigma^4} - 1\right) + \frac{2}{n-1}$$

sampling properties of δ

- $d \doteq \delta(\omega)$ is well defined, even if $x_n = m$

$$d(\omega) = d(\omega \setminus \{x_n\}) \times (n-1)/n$$

- define $D = d \times bias\_factor$ so that $\mathrm{E}(D) = \delta(\Phi)$:

Φ	D/d	$\mathrm{n} \times \frac{\mathrm{var}}{\mathrm{E}^2}(D)$	$\mathrm{n} \times \frac{\mathrm{var}}{\mathrm{E}^2}\left(S^2\right)$
Dirac's	$\frac{n}{n-1}$	$\frac{1}{\theta(1-\theta)} - 4 + \frac{2}{n-1}$	idem
unif.	$\frac{n}{n-2/3}$	$\frac{1}{3} + \frac{2/3}{n-26/45} + \cdots$	$\frac{4}{5} + \frac{2}{n-1}$

- bias factor depends on Φ

experimental behavior

- exact bias factors have not been found for other Φ
- experiments : $n = 4,\ 7,\ 10,\ 13$ and each time $N = 1600$
- using $m\left(d\right) \approx \mathrm{E}\left(d\right)$ and $S^2\left(d\right) \approx \mathrm{var}\left(d\right)$ leads to:

Φ	$\mathrm{n} \times \frac{\mathrm{var}}{\mathrm{E}^2}\left(d\right)$	$\mathrm{n} \times \frac{\mathrm{var}}{\mathrm{E}^2}\left(S\right)$	$\mathrm{n} \times \frac{\mathrm{var}}{\mathrm{E}^2}\left(S^2\right)$
unif.	≈ 0.4	≈ 0.3	$\frac{4}{5} + \frac{2}{n-1}$
gauss	≈ 0.6	≈ 0.5	$2 + \frac{2}{n-1}$
exp	≈ 1.5	≈ 1.7	$8 + \frac{2}{n-1}$
log	≈ 1.5	≈ 1.7	≈ 10

conclusion

- the best decision for the newsboy problem depends heavily on the choice of the dispersion measure
- the well known "Scarf's rule" follows when assuming an exact knowledge of μ, σ
- but another strategy follows when assuming an exact knowledge of μ, δ
- this happens while $0.3 \leq \delta/\sigma \leq 0.5$ for all relevant Φ and while estimator d doesn't behave worse than estimator S

- an exact identification of reduced parameters concerning the demand models seems to be questionable
- extraction of knowledge from history cannot be model-free (ξ^a is beyond any experience)
- a description using larger families of pdf such as $\mathcal{F}\left(\mu \pm \Delta\mu, \sigma \pm \Delta\sigma\right)$ or $\mathcal{F}\left(\mu \pm \Delta\mu, \delta \pm \Delta\delta\right)$ seems a better way for a robust description