Suppression of bad news in markets:

Equilibrium analysis of correlated optimal data censors

Adam Ostaszewski, Mathematics Dept.,
London School of Economics.

Joint work with:

Miles Gietzmann, Cass Business School, London.

30th Nov. 2011, Topics on Control (EPSRC Workshop), Inst. Mathematics, Warwick University

Alternative Subtitle:

Filtering with selectively censored data (news)

Averaging, bandwagon and quality effects from correlation

Motivation: a disclosure game

1. At the first '**ex-ante date**' Nature selects a probabilistic strategy ('action') X from a known space of actions. Actions are represented by a family of distributions.

2. At the **interim date (**a known later date), as a result of an independent draw with some probability q, this action is observed noisily by an agent ('observer').

3. At the '**terminal date**' (a still later date), there is a publicly observed vector of outcomes F_i dependent on the action X.

The 'public' comprises the agents and a disjoint set of principals (e.g. investors).

At the interim date a **pre-assessment**/evaluation of the outcome F_i may be formed from the observation.

What is the disclosure game? What is news? Answer (T for a transform):

$$T_i = T(X, Y_i) = \text{private signal about } X \text{ involving the observer's noise } Y_i,$$

received at the 'interim date' prior to public (common) knowledge of X at the terminal date.

The effect of X is to yield an outcome, e.g. via

$$F_i := f_i \cdot T(X, Z_i) = \text{effect of } X \text{ with uncertainties from } Z_i.$$

Leads to a public interim re-assessment of any disclosed signals from the agents.

This could be the evaluation of some underlying complex system based on partial noisy observation.

The ex-ante assessment is modelled as

$$\mathbb{E}[F_i] = f_i \cdot \mathbb{E}[T(X, Z_i)].$$

Game objective: maximization at the interim date of the re-assessment of F_i.

Disclosure option: opportunity to suppress the reporting of the signal T_i, if

$$\mathbb{E}[F_i| \text{ report } T_i] < \mathbb{E}[F_i|\text{no report/no disclosure}]$$

equivalently, on using $F_i = f_i \cdot T(X, Z_i)$,

$$\mathbb{E}[T(X, Z_i)| \text{ report } T_i] < \mathbb{E}[T(X, Z_i)|\text{ND}],$$

assuming there is a positive probability that the observer is unable to observe T_i.

A basic question: When is a censor γ optimal?

Answer: it is the 'indifferent censor' γ: indifference as to reporting when $T = \gamma$.

Note for later that

$$\mathbb{E}[T|\text{ND using } \gamma] := \frac{(1-q)\mathbb{E}[T] + q\mathbb{E}[T \cdot 1_{T<\gamma}]}{(1-q) + q\mathbb{E}[1_{T<\gamma}]}$$

We assume:

(i) $0 < q < 1$ and q is public (common) knowledge,

(ii) the observer does not lie, and cannot directly announce credibly absence of an observation.

The Equity-valuation model

Take $X = Y_0, Y_i, Z_i$ all log-normal with unit-mean, so in sttochastic-exponential format:

$$Y_i = e^{\sigma_i v_i - \frac{1}{2}\sigma_i^2}, \text{ for } i = 0, 1, 2, ..., n,$$

with v_i all independent, standard normal, and

$$T_i = XY_i \text{ and } F_i = f_i X Z_i.$$

The observers are called firm-managers and identified with Y_i.

Easy to include individual **dependency loading** index α_i of firm i on X :

$$T_i = XY_i \text{ and } F_i = f_i X^{\alpha_i} Z_i.$$

Corollaries of the model:

1. $T_i = e^{\sigma_{0i} w_i - \frac{1}{2}\sigma_{0i}^2}$, with $\sigma_{0i} w_i = \sigma_0 v_0 + \sigma_i v_i$ and $\sigma_{0i}^2 = \sigma_0^2 + \sigma_i^2$.

So v_0 is the only source of all the correlation.

Useful to refer to $p_i = 1/\sigma_i^2$, the **precision** of Y_i.

2.

$$\mathbb{E}[F_i|\text{data}] = f_i \mathbb{E}[X|\text{data}].$$

Noiseless Dye Cutoff: the Censor equation

For $T = X$, i.e. true value rather than a nosiy signal is observed

Dye indifference equation, or *Dye Censor Equation* is

$$\gamma = \mathbb{E}[X|ND(\gamma)].$$

It is equivalent to:

$$\lambda(m_X - \gamma) = \mathbb{E}[(\gamma - X)^+], \text{ with odds } \lambda = \frac{1-q}{q},$$

where

$$\mathbb{E}[(\gamma - X)^+] = \int (\gamma - t)^+ dF_X(t) = \int_{t \le \gamma} F_X(t)dt.$$

Alternative characterizations of the Dye censor: Minimized valuation consistent with available information:

$$\gamma = \arg\min_{\gamma} \mathbb{E}[X|ND(\gamma)].$$

No-arbitrage valuation: γ such that $\mathbb{E}[X]$ values X consistently with the possibility of further γ-censored information becoming available later.

The hemi-mean function

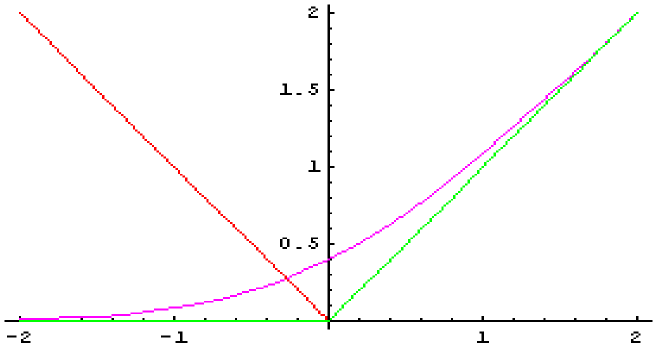
This put-payoff is valued under an expectation, and we call

$$H_X(\gamma) := \int_{t \leq \gamma} F_X(t)dt,$$

the *hemi-mean function* of X. Since $H'' = f_X \geq 0$ that itself is an increasing convex function of γ and so has a smoothed out hockey-stick shape: it looks like the valuation of a call (dual to the put). Examples below! Dye equation standardizes to:

$$\lambda(1 - \gamma) = H_X(\gamma).$$

The Normal Censor



The pink/red intersection identifies the normal Dye censor (here $\lambda = 1$).

A corresponding dual call payoff $(X - x)^+$ is in green.

Location-scale cutoff standardization theorem. *For the location and scale family of distributions* $\Phi_F(\frac{x-\mu}{\sigma})$*, with mean* μ *and variance* σ^2*, the Dye cutoff* $\gamma(\mu, \sigma, \lambda)$ *satisfies*

$$\gamma(\mu, \sigma, \lambda) = \mu - \sigma\xi(\lambda).$$

So:

$$p_{\text{Low}} < p_{\text{High}} \implies \gamma(p_{\text{Low}}) < \gamma(p_{\text{High}}),$$

i.e. more disclosure from the low-precision firm.

This will be altered by the pressence of additional information sources.

***Location-scale cutoff standardization theorem**. *Let* $\Phi_F(x)$ *be an arbitrary zero-mean, unit-variance, cumulative distribution for* F *defined on* $\mathbb{R}$. *For the location and scale family of distributions* $\Phi_F(\frac{x-\mu}{\sigma})$, *with mean* μ *and variance* σ^2, *the Dye cutoff* $\gamma(\mu, \sigma, \lambda)$ *satisfies*

$$\gamma(\mu, \sigma, \lambda) = \mu - \sigma\xi(\lambda), \text{ where } \lambda = \frac{1-q}{q},$$

so that

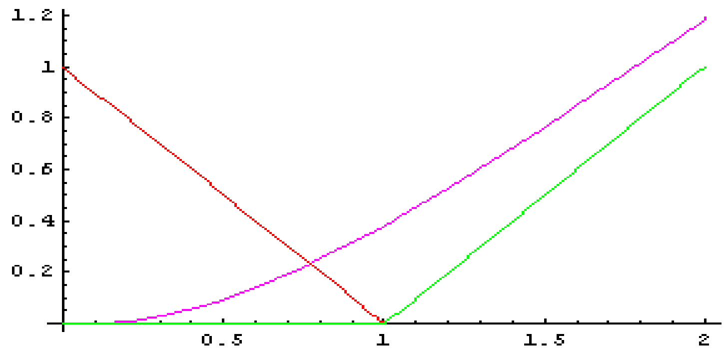
$$\xi(\lambda) = -\gamma(0, 1, \lambda) < 0$$

is the cutoff when standardizing to zero mean and unit variance and is a function only of the odds λ. *The standardized cutoff* $\xi(\lambda)$ *is a convex and decreasing function of* λ *satisfying*

$$\lambda = H_F(-\xi)/\xi,$$

where $H_F(x) = \int_{-\infty}^{x} \Phi_F(t)dt$ *is the corresponding hemi-mean function.*

Black-Scholes Censor



The red-pink intersection identifies the log-normal Dye censor (for $\lambda = 1$).
Green indicates the dual call payoff.

Noisy Dye Cutoff: Estimator-Censor equation

For $T = T(X, Y)$, put

$$\begin{aligned} \mu_X(t) &:= \mathbb{E}[X|T = t], \text{ the regression function,} \\ S &:= \mu_X(T), \text{ the estimator, or } X^{\text{est}}. \end{aligned}$$

Since

$$\mathbb{E}[F] = f_i \mathbb{E}[X],$$

then, provided $\mu_X(.)$ is strictly increasing, the *Dye Equation* holds in the form:

$$\mu_X(\gamma_T) = \gamma_S = E[S|ND(\gamma_S)],$$

where γ_S is the censor for S and γ_T is the equivalent censor for T.

Equivalently, as S is an unbiased estimator of X one has

$$\lambda(m_X - \gamma_S) = H_S(\gamma).$$

By the conditional mean formula (tower law/iterated expecation):

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[X|T]] = \mathbb{E}[X] = m_X.$$

So the hemi-mean function rules OK.

Multi-Censor Equilibrium equation

One has n simultaneous equations corresponding to a simultaneous interim-report date:

$$\begin{aligned} \mathbb{E}[X|T_j &= \gamma_j \text{ for all } j] = \mathbb{E}[X|ND_i(\gamma)], \\ \text{with } \gamma &= (\gamma_1, ..., \gamma_n) \text{ and } ND_i = \text{only } i \text{ makes no disclosure.} \end{aligned}$$

We call these the *Marginal Dye equations*.

Log-normal Marginal Dye equations

Recall the Estimator version of the Dye equation:

$$\lambda(m_X - \gamma_S) = H_S(\gamma).$$

Conditioning on the other disclosures, yields for some K and $\kappa_i = p_i/p$

$$\mu_X(\gamma_1, ..., \gamma_n) = E[X|T_i = \gamma_i \text{ all } i] = K\gamma_1^{\kappa_1} .... \gamma_n^{\kappa_n},$$

(see below). Change of random variable, and change of variable:

$$S := \mu_X(T_1, \gamma_2, ..., \gamma_n), \text{ and } s = \mu_X(\gamma, \gamma_2, ..., \gamma_n)$$

yields a conditioned format, in which $m_{S|\gamma_2....}$ replaces m_S :

$$\lambda(\mathbb{E}[S|\gamma_2, ..., \gamma_n] - s) = H_S(s|\gamma_2, ..., \gamma_n).$$

Principal findings for the Equity Valuation case:

Preparatory Step. Replace the n firm-managers Y_i by n *hypothetical* observers/managers $\hat{Y}_i$ which are uncoupled – acting as though all the competitors had vanished – but with refined precision parameters

$$\kappa_i \sigma_{0i} \sqrt{1-\rho_i^2}, \text{ with } \kappa_i := \frac{p_i}{p} \text{ and } \sigma_{0i}^2 = \sigma_0^2 + \sigma_i^2,$$

and

$$p = p_0 + ... + p_n, \text{ total precision.}$$

Here ρ_i measures the dependence of T_i on the remaining T_j (more properly: **partial co-variance** of w_i on the remaining w_j).

Conclusion. If the corresponding Dye censors for $\hat{T}_i = X\hat{Y}_i$ are $\hat{\gamma}_i$, then the true managers have censors γ_i given by the weighted average:

$$\log \gamma_i = \frac{\log g_i}{\kappa_{-i}} + \frac{1}{\kappa_0}\left(\frac{\kappa_1}{\kappa_{-1}}\log g_1 + \frac{\kappa_2}{\kappa_{-2}}\log g_2 + ... + \frac{\kappa_n}{\kappa_{-n}}\log g_n\right),$$

with

$$\kappa_{-i} = p_i/(p - p_i),$$

and where g_j is the hypothetical firm-j censor.

In fact

$$
\begin{aligned}
g_i &= \log\left(\hat{\gamma}_{\mathsf{LN}}\left(\lambda_i, \kappa_i \sigma_{0i}\sqrt{1-\rho_i^2}\right) L_{-i}\right), \ \ \lambda_i = \frac{1-q_i}{q_i}, \\
L_{-i} &= \exp\left(\frac{n-1}{2(p-p_i)} - \frac{1}{2}\frac{n}{p}\right) = \exp\frac{1}{2}\left(\frac{1}{p_{\mathsf{av,\text{-}}i}} - \frac{1}{p_{\mathsf{av}}}\right),
\end{aligned}
$$

where L_{-i} is a mean adjustment.

Bandwagon effect

Bandwagon Inflator Theorem. *The presence of correlation increases the precision parameter of the cutoff and hence raises the cutoff:*

$$\hat{\gamma}_{\mathsf{LN}}(\lambda_i, \sigma_{0i}) < \hat{\gamma}_{\mathsf{LN}}(\lambda_i, \kappa_i \sigma_{0i}) < \hat{\gamma}_{\mathsf{LN}}\left(\lambda_i, \kappa_i \sigma_{0i}\sqrt{1-\rho_i^2}\right).$$

Proof. Clear since $\hat{\gamma}_{\mathsf{LN}}(\lambda, .)$ is increasing in precsioin, and also ρ_i^2 is increasing in p_i.

Estimator-quality effect

Estimator-Quality Theorem. *The mean-adjustor for firm i is increasing in p_i with*

$$\exp\left(-\frac{1}{2(p-p_i)}\right) < L_{-i} < \exp\left(+\frac{1}{2p_{\mathsf{av},-i}}\right),$$

and in particular

$$L_{-i} < L_{-j} \text{ iff } p_i < p_j.$$

The adjustor is a strict deflator, i.e. $L_{-i} < 1$, *iff* p_i *is below the sector average, equivalently below the competitor average, i.e.*

$$p_i < \frac{p}{n}, \text{ equivalently } p_i < \frac{p-p_i}{n-1}.$$

Tools:

Basic Tools: **Isomorphism**. Equity a log-normal variate, but it is easy to move back and forth from log-normal to normal via the isomorphism $\exp : (\mathbb{R}, +) \to (\mathbb{R}_+, \cdot)$

Explicit Normal and Black-Scholes **put-option formulas**.

Main Tools: Linear regression easily computed via a **Hilbert space approach**: view $\mathbb{E}[..]$ as a projection and use P the **precision matrix**.

Strategy: **Uncoupling** the co-dependency and **solving the uncoupled censor equations** via P.

Some simple algebra: the precision matrix

Put

$$P_n := \begin{bmatrix} p_1 & p_2 & \cdots & p_n \\ p_1 & p_2 & & p_n \\ \vdots & & \ddots & \vdots \\ p_1 & p_2 & \cdots & p_n \end{bmatrix}.$$

and

$$P_n(x) = P_n - xI.$$

Recall that for σ_i^2 a variance parameter, $p_i = 1/\sigma_i^2$ is the **precision** parameter.

Proposition 1. *For any n, the characteristic function of the matrix P_n is*

$$\det(P_n - xI) = (-1)^n x^{n-1}(x - p_1 - ... - p_n),$$

Proof. Easy exercise. [Hint: P_n has nullity $n-1$.]

Proposition 2. *For any non-zero parameter q such that $p_q := q + p_1 + ... + p_n \neq 0$, the simultaneous system of equations*

$$(P_n + qI)x = s,$$

i.e.

$$p_1 x_1 + ... + (p_i + q)x_i + ... + p_n x_n = s_i,$$

has the unique solution

$$x_i = \frac{s_i}{q} + c, \text{ with } c = \frac{1}{qp_q}(p_1 s_1 + ... + p_n s_n).$$

Proof. Easily checked; by Prop. 1, $\det(P_n + qI) = q^{n-1}(p_1 + ... + p_n + q) \neq 0$, so the solution is unique. $\square$

*Example 1: Normal put-option formula

Notation

$$F_X(t) := \mathsf{Pr}[X \leq t]$$

Cases: $X = u$ ~$\mathsf{N}(0, \sigma^2)$ normal

$$\Phi(t) = F_u(t) := \mathsf{Pr}[u \leq t].$$

with density $\varphi(t) = \Phi'(t)$. Here

$$\mathbb{E}[(t - X)^+] = t\Phi\left(\frac{t + \frac{1}{2}\sigma^2}{\sigma}\right) + \varphi(t/\sigma^2).$$

*Example 2: Black-Scholes put-option formula

For X log-normal

$$X = e^{\sigma u - \frac{1}{2}\sigma^2} \text{ with } u \sim N(0,1),$$

$$\mathbb{E}[(t - X)^+] = t\Phi\left(\frac{\log t + \frac{1}{2}\sigma^2}{\sigma}\right) - \Phi\left(\frac{\log t - \frac{1}{2}\sigma^2}{\sigma}\right).$$

Simplification:

Again use the conditional mean formula:

$$\begin{aligned}
\mathbb{E}[S|\gamma_2,....,\gamma_n] &= \mathbb{E}[\mathbb{E}[X|\gamma_2,...\gamma_n]|\gamma_2,...\gamma_n], \text{ defn of } S \\
&= \mathbb{E}[\mathbb{E}[\mathbb{E}[X|T_1,\gamma_2,...\gamma_n]]|\gamma_2,...\gamma_n]]|\gamma_2,...\gamma_n], \text{ refine} \\
&= \mathbb{E}[KT_1^{\kappa_1}\gamma_2^{\kappa_2}....\gamma_n^{\kappa_n}|\gamma_2,...\gamma_n], \text{apply formula} \\
&= K\gamma_2^{\kappa_2}....\gamma_n^{\kappa_n}\cdot\mathbb{E}[T_1^{\kappa_1}|\gamma_2,...\gamma_n].
\end{aligned}$$

Theorem (Conditional hemi-mean formula).

$$\mathbb{E}[T_1^{\kappa_1}|T_2,...,T_n] = L_{-1} T_2^{\bar{h}_2-\kappa_2} ... T_n^{\bar{h}_n-\kappa_n},$$

where, with $p = p_0 + ... + p_n$ the total precision,

$$L_{-1} = \exp\left(\frac{n-1}{2(p-p_1)}\right)\exp\left(-\frac{n}{2p}\right), \text{ and } \bar{h}_j = \frac{p_j}{p-p_1}, \text{ for } j > 1.$$

Proof uses conditional mean formula and yields $L_{-1} = K_{-1}/K$.

Uncoupling Theorem

Uncoupling Theorem. *The substitution*

$$y_1 = \gamma_1^{\kappa_1}/L_{-1}\gamma_2^{\bar{h}_2^1 - \kappa_2}..\gamma_n^{\bar{h}_n^1 - \kappa_n}$$

reduces the marginal Dye equation, namely

$$\begin{aligned} & \lambda_1 \left(\mathbb{E}[X|\gamma_2, ..., \gamma_n] - \mu_X(\gamma, \gamma_2, ..., \gamma_n)\right) \\ = & \int_{t_1<\gamma_1} \left[\mu_X(\gamma_1, \gamma_2, ..., \gamma_n) - \mu_X(t_1, \gamma_2, ..., \gamma_n)\right] dF_{T_1}(t_1|\gamma_2, ..., \gamma_n), \end{aligned}$$

to the standard form

$$\lambda_1(1 - y_1) = H_{\mathsf{LN}}(y_1, \kappa_1\sigma_{01}\sqrt{1 - \rho_1^2}),$$

where $1 - \rho_1^2$ *is the partial covariance, or Schur complement, of* w_1 *given* $w_1, ..., w_n$.

Notational convention for shifting from LN to N

$\eta_i = \log Y_i + \frac{1}{2}\sigma_i^2 = \sigma_i v_i$ the underlying normal variate, etc

Background: a little linear regression

Proposition (Geometric weighted-average)

$$\begin{aligned} E[X|T_1 &= t_1, ..., T_n = t_n] = K t_1^{\kappa_1} ... t_n^{\kappa_n}, \text{ with } \kappa_i = \frac{p_i}{p}, \text{ and} \\ K &= e^{\frac{n}{2p}} = \exp\left(\frac{1}{2p_{\text{av}}}\right) t_1^{\kappa_1} ... t_n^{\kappa_n}, \text{ with } p_{\text{av}} := \frac{p_0 + ... + p_n}{n}. \end{aligned}$$

Sketch Proof. Put $\xi = \log X$, $\tau_i = \log T_i$ (+ take off constants), do classical linear regrssion with normal variates, transform back via exp , finally compute the constant K using the tower law.

Remarks. 1. The preceding shows why log-normals are as easy as normals.

2. The normal regression arguments need only P, so some simple algebra.

Reprise: a little linear regression

Lemma (Arithmetic weighted-average). *One has*

$$\mathbb{E}[\xi|\tau_1, \tau_2] = \kappa_1\tau_1 + ... + \kappa_n\tau_n, \text{ with } \kappa_i = \frac{p_i}{p}.$$

Proof. Method: write

$$\xi^{\text{est}} = \mathbb{E}[\xi|\tau_1, ..., \tau_n] = \kappa_1\tau_1 + ... + \kappa_n\tau_n.$$

By the conditional mean formula,

$$\begin{aligned}\mathbb{E}[\tau_1\xi^{\text{est}}] &= \mathbb{E}[\tau_1\mathbb{E}[\xi|\tau_1, ..., \tau_n]] = \mathbb{E}[\mathbb{E}[\tau_1\xi|\tau_1, ..., \tau_n]] \\ &= \mathbb{E}[\tau_1\xi]\end{aligned}$$

Recall, v_i independent so $E[v_i v_j] = \delta_{ij}$ and

$$\tau_i = (v_0 + v_i)$$

Compute to obtain

$$\mathbb{E}[\tau_1 \xi^{\text{est}}] = \mathbb{E}[\tau_1 \xi]$$

equivalent to:

$$\kappa_1(\sigma_0^2 + \sigma_1^2) + \kappa_2 \sigma_0^2 + ... + \kappa_n \sigma_0^2 = \sigma_0^2.$$

Setting $k_i = \kappa_i / p_i$, obtain

$$k_1(p_0 + p_1) + k_2 p_2 + ... + k_n p_n = 1.$$

More generally,

$$k_1 p_1 + ... + k_i(p_0 + p_i) + ... + k_n p_n = 1.$$

Solution now obviously: $k_i = 1/(p_0 + ... + p_n)$.

Covariance: the Hilbert space view

Recall that each w_i has mean-zero and that

$$\mathbb{E}[w_i w_i] = 1, \text{ and } \mathbb{E}[w_i w_j] = \frac{\sigma_0^2}{\sigma_{0i}\sigma_{0j}} > 0.$$

So any combination of $w_1, ..., w_n$ has mean zero, i.e. they span a vector space W. For $w, w' \in W$ write

$$\langle w, w' \rangle := cov(w, w') = \mathbb{E}[ww'].$$

This is an inner product (so W is a Hilbert space under $\langle ., . \rangle$) iff the following *covariance matrix* is non-singular

$$Q = (\rho_{ij}) \text{ where } \rho_{ij} = \mathbb{E}[w_i w_j].$$

It turns out that Q is related to the precision matrix.

Theorem. *For* $p_i > 0$ *the covariance matrix is non-singular and*

$$\begin{aligned} \det Q &= (p_0 + p_1)...(p_0 + p_m) \det[P + p_0 I] \\ &= \bar{p} p_0^{m-1} (p_0 + p_1)...(p_0 + p_m). \end{aligned}$$

Appendix: the Schur complement: 1

Aim: find the variance of $\mathbb{E}[w_i|w_j \forall j \neq i]$. NB. Requires first to solve e.g.

$$\mathbb{E}[w_n|w_1, ..., w_{n-1}] = \sum_{j<n} \beta_j w_j.$$

Answer: put $\bar{Q}_i = Q$ omitting the i-th row and column; likewise,

$\vec{\rho}_i = i$-th row $(\rho_{i1}, ..., \rho_{i,n})$ omitting i-th entry.

The *Schur complement* (of $\bar{Q}_i$ in Q) is given by

$$\rho_{ii} - \vec{\rho}_i \bar{Q}_i^{-1} \vec{\rho}_i^T.$$

Putting

$$\rho_i := \sqrt{\vec{\rho}_i \bar{Q}_i^{-1} \vec{\rho}_i^T},$$

the Schur complement becomes

$$1 - \rho_i^2.$$

(This notation permits specialization to the $n = 2$ case to yield $\bar{Q}_i = (1)$ and $\vec{\rho}_i = (\rho)$, so that $\rho_i = \rho = \rho_{12}$.)

The conditional distribution of w_i given all the w_j for $j \neq i$ is normal with variance given by the Schur complement.

The Schur complement: 2

Consider the distribution of $\mathbb{E}[T_n|T_1, ..., T_{n-1}]$, or equivalently that of $E[w_n|w_1, ..., w_{n-1}]$. Recall that

$$T_i = e^{\sigma_{0i} w_i - \frac{1}{2}\sigma_{0i}^2}, \text{ with } \sigma_{0i} w_i = \sigma_0 w_0 + \sigma_i v_i.$$

Put

$$w_n^{n-1} = \mathbb{E}[w_n|w_1, ..., w_{n-1}] = \sum_{j<n} \beta_j w_j.$$

Then, by definition and by the conditional mean formula,

$$\rho_{in} = \mathbb{E}[w_i w_n] = \mathbb{E}[w_i w_n^{n-1}] = \sum_{j<n} \beta_j \rho_{ij}.$$

We solve the system of $m := n - 1$ equations for $i < n$

$$\sum_{j<n} \rho_{ij} \beta_j = \rho_{in},$$

or, in matrix form with $\vec{\rho}_n := (\rho_{1n}, ..., \rho_{n-1,n})$

$$Q_{n-1}\beta = \vec{\rho}_n,$$

by computing explicitly $\beta = Q_{n-1}^{-1}\vec{\rho}_n$. Here we have denoted the principal submatrix of the covariance matrix Q_n by:

$$Q_{n-1} = (\rho_{ij})_{i,j<n}.$$

Using the precision matrix one may easily find the β_j *explicitly*. WE have an important corollary.

Monotonicity Theorem (Own precision refined by presence of others) *The Schur complement*

$$1-\rho_n^2,$$

corresponding to conditioning w_n *on* $w_1,..,w_{n-1}$ *as a factor in the conditional variance, acts to increase the precision; increasing the precision of the competitors refines one's own conditional precision. Indeed, one has the explicit formula with* $m=n-1$ and $\bar{p}=p-p_n=p_0+...+p_{n-1}$,

$$\rho_n^2=\frac{p_m}{p_0\bar{p}(p_0+p_m)}\left[\sum\nolimits_{i=1}^{m}p_i(\bar{p}-p_i)+\sum\nolimits_{i<j\leq m}(p_i+p_j)\sqrt{\frac{p_ip_j}{(p_0+p_i)(p_0+p_j)}}\right],$$

which is increasing in p_i *for each* $i<n$, *and so the Schur complement itself decreases with* p_i.

In fact one has:

Theorem 1. *Provided all the precisions p_i are finite and positive, the regression equations*

$$\mathbb{E}[w_n|w_1, ..., w_{n-1}] = \beta_1 w_1 + ... + \beta_{n-1} w_{n-1},$$

which are equivalent to the solution of the system $Q_{n-1}\beta = \vec{\rho}_n$, have non-singular matrix Q_{n-1} and the equivalent system of equations, for $i = 1, 2, ..., m =: n-1$,

$$\rho_{i1}\beta_1 + ... + \beta_i + ... = \rho_{in},$$

has the unique solution:

$$\beta_i = \frac{p_i + p_0}{p}\rho_{in}.$$

Proof of the averaging effect

In the setting of the Uncoupling Theorem, the equations

$$\gamma_i^{\kappa_i} = \hat{\gamma}_i L_{-i} \prod_{j \neq i} \gamma_j^{\bar{h}_j^i - \kappa_j},$$

imply

$$x_i - \sum_{j \neq i} \bar{h}_j^i \, x_j = B_i := \frac{1}{\kappa_i} \log \left(\hat{\gamma}_i L_{-i} \right) = \frac{p}{p_i} \log \left(\hat{\gamma}_i L_{-i} \right),$$

with

$$x_i = \log \gamma_i.$$

Proof. Cross-multiply take logs and note

$$\begin{aligned}\bar{h}_j^i \kappa_i &= \bar{h}_j^i - \kappa_j \\ &= \frac{p_j}{p - p_i} - \frac{p_j}{p} = p_j \frac{p - (p - p_i)}{p(p - p_i)} = \frac{p_i}{p} \frac{p_j}{(p - p_i)}.\end{aligned}$$

The more revealing re-statement is

$$(\kappa_i - 1)x_i + \sum_{j \neq i} \kappa_j x_j = b_i := \frac{(p_i - p)}{p_i} \log\left(\hat{\gamma}_i L_{-i}\right).$$

Conditional hemi-mean formula

The following identifies the hemi-mean function.

Theorem (Conditional hemi-mean formula).

$$\mathbb{E}[T_1^{\kappa_1}|T_2,...,T_n] \;=\; L_{-1}T_2^{\bar{h}_2-\kappa_2}...T_n^{\bar{h}_n-\kappa_n}, \text{ where } L_{-1} = \exp\left(\frac{n-1}{2(p-p_1)}\right)\exp\left(-\frac{n}{2p}\right)$$

$$\text{and } \bar{h}_j \;=\; \frac{p_j}{p-p_1}, \text{ for } j>1.$$

Hence, for any γ,

$$\mathbb{E}[T_1^{\kappa_1}1_{T_1<\gamma}|(T)_{-1}] = L_{-1}\prod_{j>1}T_j^{\bar{h}_j-\kappa_j}\Phi_{\mathsf{LN}}(\gamma^{\kappa_1}/L_{-1}\prod_{j>1}T_j^{\bar{h}_j-\kappa_j}, \kappa_1\sigma_{01}\sqrt{1-\rho_1^2}).$$

Proof. The random variable $S = T_1^{\kappa_1}$ has mean $m = m(\kappa_1, \sigma_{01})$ and volatility $\kappa_1 \sigma_{01}$. Hence, by the Exponent Effect Theorem,

$$H_S(\gamma^{\kappa}) = mH_{\mathsf{LN}}(\gamma^{\kappa_1}/m, \kappa_1\sigma_{01}).$$

The distribution of S conditional on $T_2 = t_2, ..., T_2 = t_n$ (for any $t_2, ..., t_n$) has a mean $\xi = \xi_{-1}$ (depending on $t_2, ..., t_n$ to be determined below) and a volatility $\kappa_1\sigma_{01}\sqrt{1-\rho_n^2}$, with $1-\rho_n^2$ the 'Schur complement' of T_n in $(T_2, .., T_n)$, because that is the effect on normal variates of conditioning (see Bingham & Fry (2010)).

Thus putting $\eta = \eta_{-1} := m\xi_{-1}$ we have for any $\gamma > 0$ that

$$\begin{aligned} H_{S|t_2\ldots}(\gamma^{\kappa_1}) &= E[(\gamma^{\kappa_1} - T_1^{\kappa_1})1_{T_1<\gamma}|T_2 = t_2, ..., T_n = t_n] \\ &= m\xi H_{\mathsf{LN}}(\gamma^{\kappa_1}/m\xi, \kappa_1\sigma_{01}\sqrt{1-\rho_n^2}) \\ &= \gamma^{\kappa_1}\Phi_{\mathsf{N}}\left(\frac{\log(\gamma^{\kappa_1}/\eta) + \frac{1}{2}\kappa_1^2\sigma_{01}^2\rho_n}{\kappa_1\sigma_{01}\sqrt{1-\rho_n^2}}\right) \\ &\quad -\eta\Phi_{\mathsf{N}}\left(\frac{\log(\gamma^{\kappa_1}/\eta) - \frac{1}{2}\kappa_1^2\sigma_{01}^2\rho_n}{\kappa_1\sigma_{01}\sqrt{1-\rho_n^2}}\right). \end{aligned} \tag{CH}$$

This leaves open the determination of the 'constant' $\eta = \eta_{-1}$. But minus the second term has the value

$$E[T_1^{\kappa_1}1_{T_1<\gamma}|T_2 = t_2, ..., T_n = t_n].$$

So taking the limit as $\gamma \to +\infty$ we obtain

$$\eta = \eta_{-1} = E[T_1^{\kappa_1} | T_2 = t_2, ..., T_n = t_n].$$

Now, by the conditional mean formula, with

$$\bar{h}_i = \bar{h}_i^1 = \frac{p_i}{p - p_1}$$

$$\begin{aligned}
H_{-1} t_2^{\bar{h}_2} ... t_n^{\bar{h}_n} &= \mathbb{E}[X | T_2 = t_2, ..., T_n = t_n] \\
&= \mathbb{E}[\mathbb{E}[X | T_1, T_2 = t_2, ..., T_n = t_n] | T_2 = t_2, ..., T_n = t_n] \\
&= \mathbb{E}[K T_1^{\kappa_1} t_2^{\kappa_2} ... t_n^{\kappa_n} | T_2 = t_2, ..., T_n = t_n] \\
&= K t_2^{\kappa_2} ... t_n^{\kappa_n} \mathbb{E}[T_1^{\kappa_1} | T_2 = t_2, ..., T_n = t_n]
\end{aligned}$$

and so

$$\begin{aligned}
\eta_{-1} &= (H_{-1}K^{-1})t_2^{\bar{h}_2-\kappa_2}...t_2^{\bar{h}_n-\kappa_n} \\
&= \exp\left(\frac{n-1}{2(p_0+p_2+...+p_n)}\right)\exp\left(-\frac{n}{2p}\right)t_2^{\bar{h}_2-\kappa_2}...t_2^{\bar{h}_n-\kappa_n} \\
&= \exp\left(\frac{n-1}{2(p-p_1)}\right)\exp\left(-\frac{n}{2p}\right)t_2^{\bar{h}_2-\kappa_2}...t_2^{\bar{h}_n-\kappa_n},
\end{aligned}$$

as required. The rests is now clear from (CH) above.

Postscript: Log-normal vs normal: standardization

Normal x with mean m and variance σ^2 transforms to $v = (x - m)/\sigma \sim$N$(0, 1)$, i.e. zero-mean unit-variance. Note the moment generating function for $x \sim$N$(0, 1)$ is

$$E[e^{sx}] = e^{\frac{1}{2}s^2}.$$

General **log-normal**

$$X = m_X e^{\sigma x - \frac{1}{2}\sigma^2} \text{ with } x \sim N(0, 1).$$

Consider now the power transformation $Y = X^\kappa$ for $0 < \kappa < 1$, then with $s = \kappa\sigma$

$$\begin{aligned} Y &= e^{\kappa\sigma x - \frac{1}{2}\kappa\sigma^2} = e^{\frac{1}{2}\kappa(\kappa-1)\sigma^2} e^{sx - \frac{1}{2}s^2} \\ &= e^{\frac{1}{2}\kappa(\kappa-1)\sigma^2} Z. \end{aligned}$$

That is, the new variable has reduced mean

$$m = m(\kappa, \sigma) := e^{\frac{1}{2}\kappa(\kappa-1)\sigma^2}.$$

(Smart reason: derive this from from Ito's Lemma! via the second derivative of y^κ.)

Log-normal X with mean m_X and variance σ^2 transforms using $\kappa = 1/\sigma$ to $Y = X^\kappa$ with unit variance and mean

$$m_Y = m_X e^{\frac{1}{2}(1-\sigma)}$$

and so we arrive at $Z = X^\kappa/m_Y = (Y/m_Y)$ $\sim$LN$(1, 1)$, i.e. unit-mean unit-variance.