KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0132

ON THE DISTRIBUTION OF CASH-FLOWS USING ESSCHER TRANSFORMS

by
D. VYNCKE
M. GOOVAERTS
A. DE SCHEPPER
R. KAAS
J. DHAENE

D/2001/2376/32

On the distribution of cash-flows using Esscher transforms

D. Vyncke[1], M.J. Goovaerts[1,2],
A. De Schepper[3], R. Kaas[1,2], J. Dhaene[1,2]

Abstract

In their seminal paper, Gerber and Shiu (1994) introduced the concept of the Esscher transform for option pricing. As examples they considered the shifted Poisson process, the random walk, a shifted gamma process and a shifted inverse Gaussian process to describe the logarithm of the stock price. In the present paper it is shown how upper and lower bounds in convex order can be obtained when we use these types of models to describe the financial stochasticity for a given cash-flow.

1 Introduction

In their seminal paper, H. Gerber and E. Shiu (1994) advocated the Esscher transform as a tool to deal with stock-price movements for a family of processes. With $M(h)$ denoting the moment generating function of a random variable X, i.e.

$$M(h) = \mathrm{E}\left[e^{hx}\right] \tag{1}$$

the Esscher transform (with parameter h) of the density $f(x)$ is obtained in case the function

$$f(x,h) = \frac{e^{hx} f(x)}{M(h)} \tag{2}$$

[1] University of Leuven, Belgium
[2] University of Amsterdam, the Netherlands
[3] University of Antwerp, Belgium

is a density. In Goovaerts et al. (1984) it is shown how the Esscher transform evolves from utility theory in measuring the price of a random variable. Indeed, one has the following theorem:

Theorem 1. *Assume an insurer has an exponential utility function with risk aversion α. If he charges a premium of the form $E[\varphi(X)X]$ where $\varphi(\cdot)$ is a continuous increasing function with $E[\varphi(X)] = 1$, his utility is maximized if $\varphi(x) \propto e^{\alpha x}$, i.e. if he uses the Esscher premium principle with parameter α.*

For a proof of this theorem, we refer to the Appendix. If the utility function u is an exponential one, i.e. $u(x) = \frac{1}{h}\left(1 - e^{-hx}\right)$, then

$$\varphi(x) = \frac{e^{hx}}{M(h)}, \tag{3}$$

so the Esscher transform of the risk X evolves. If $u(x)$ is quadratic, hence e.g. $u(x) = ax^2 + bx$, we get $\varphi(x) \propto 2ax + b$, and $E[\varphi(X)] = 1$ gives

$$\begin{aligned} E[X\varphi(X)] &= E[X(2aX+b)/E[2aX+b]] = \frac{2aE[X^2] + bE[X]}{2aE[X]+b} \\ &= E[X] + \frac{1}{E[X] + b/2a} Var[X], \end{aligned}$$

which is a premium of type variance premium, if risks with a given expectation are considered.

These rather simple results indicate a relationship between the actuarial approach of premium principles and the financial approach of pricing risks by means of a measure transformation. Gerber and Shiu (1994) considered a stochastic process $\{X(t)\}_{t\geq 0}$ with stationary and independent increments, $X(0) = 0$, such that

$$S(t) = S(0) \cdot e^{X(t)} \qquad (t \geq 0). \tag{4}$$

To make sure that the stock prices of the model are internally consistent, they seek for a $h = h^*$ so that the discounted price process

$$\left\{e^{-\delta t} S(t)\right\}_{t\geq 0} \tag{5}$$

is a martingale with respect to the probability measure corresponding to h^*. In particular,

$$S(0) = e^{-\delta t} E^* \left[S(t)\right] \tag{6}$$

Table 1: Esscher transforms for some types of stochastic processes

Stock-price model	$F(x,t;h^*)$	h^*
Wiener process	$N(x;(\mu+h^*\sigma^2)t,\sigma^2 t)$	$\delta=(\mu+h^*\sigma^2)+\frac{1}{2}\sigma^2$
Shifted Poisson process	$\Lambda\left(\frac{x+ct}{k};\lambda e^{h^*k}t\right)$	$\delta=\lambda e^{h^*k}\left(e^{-k}-1\right)-c$
Random walk	$B\left(\frac{x-at}{b-a};t,\pi(h^*)\right)$	$\pi(h^*)=\frac{e^{\delta}-e^{a}}{e^{b}-e^{a}}$
Shifted Gamma process	$G\left(x+ct;\alpha t,\beta-h^*\right)$	$e^{\delta}=\left(\frac{\beta-h^*}{\beta-h^*-1}\right)^{\alpha}e^{-c}$
Shifted inverse Gaussian	$J\left(x+ct;at,b-h^*\right)$	$\delta=a\left(\sqrt{b-h^*}-\sqrt{b-h^*-1}\right)-c$

where δ denotes the constant risk-free force of interest. Based on the theorem above, the price of the risk is calculated by choosing the coefficient of risk aversion such that the premium coincides with the market price.

The results presented by Gerber and Shiu (1994) can be summarized as in Table 1, where $F(x,t;h^*)$ is the cumulative distribution function of the Esscher transform of the process $X(t)$. A definition of the stochastic processes as well as an overview of the notations for the functions in the second column can be found in the Appendix.
In this paper, we are interested in deriving upper and lower bounds in convex order for a (discrete version of the) cash-flow

$$C(t)=\int_0^t c(s)e^{X(t)-X(s)}ds, \tag{7}$$

where $X(t)$ is assumed to be one of the stochastic price models given in Table 1, using the approach described by Kaas, Dhaene & Goovaerts (2000). Note that the problem of pricing an Asian option under the stochastic processes of Table 1 can be solved by using the same approach, because the bounds will have higher, resp. lower, stop-loss premiums.

In the following section, we explain the concept of convex order and describe a methodology to obtain upper bounds. As we will also construct lower bounds, the results in this paper extend the results in Goovaerts et al. (2000). To calculate the lower bound, and to improve the upper bound, the methodology requires the knowledge of the conditional distribution of the process $\{X(t)\}$, conditionally on some random variable Z. A potential conditional distribution is derived in section 3. Finally, in section 4 we apply the techniques to the problem at hand and in section 5 we illustrate the obtained bounds graphically.

2 Convex order and comonotonicity

The distribution function of (7) is very hard, or even impossible, to obtain due to the dependency structure among the different random variables. Therefore, instead of calculating the exact distribution, we will look for bounds, in the sense of "more favourable/less dangerous" and "less favourable/more dangerous", with a simpler structure. This technique is common practice in the actuarial literature. When lower and upper bounds are close to each other, together they can provide reliable information about the original and more complex variable. The notion "less favourable" or "more dangerous" variable will be defined by means of the convex order.

Definition 1. *A random variable V is smaller than a random variable W in convex order if*

$$E\left[u(V)\right] \leq E\left[u(W)\right], \tag{8}$$

for all convex functions $u : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto u(x)$, provided the expectations exist. This is denoted as

$$V \leq_{cx} W. \tag{9}$$

Since convex functions are functions that take on their largest values in the tails, the variable W is more likely to take on extreme values than the variable V, and thus W is more dangerous.

The convex order can also be interpreted in terms of utility theory. Indeed, if $V \leq_{cx} W$, then V is preferred to W by all risk averse decision makers, see e.g. [4]. This means that replacing the unknown distribution function of the variable V by the distribution function of the variable W is a prudent strategy.

Since the functions $u(x) = x$, $u(x) = -x$ and $u(x) = x^2$ are all convex functions, it follows immediately that $V \leq_{cx} W$ implies $E[V] = E[W]$ and $Var[V] \leq Var[W]$.

The following lemma provides an interesting and useful characterization of convex order, a proof of which can be found in [6] :

Lemma 1. *For any two random variables V and W, we have the following equivalence:*

$$V \leq_{cx} W \Leftrightarrow \begin{cases} E[(V-k)_+] \leq E[(W-k)_+] & \text{for all } k, \\ E[V] = E[W] \end{cases} \tag{10}$$

where $(x)_+ = \max\{0, x\}$.

Now, if V consists of a sum of random variables $X_1, \ldots, X_n$, then replacing the copula of $(X_1, \ldots, X_n)$ by the comonotonic copula yields an upper bound for V in the convex order. On the other hand, applying Jensen's inequality to V provides us with a lower bound. Finally, if we combine both ideas, then we end up with an improved upper bound. This is formalized in the following proposition.

Proposition 1. *Consider an arbitrary sum of random variables*

$$V = X_1 + X_2 + \ldots + X_n, \tag{11}$$

and define the related stochastic quantities

$$V_u = F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + \ldots + F_{X_n}^{-1}(U) \tag{12}$$

$$V_{iu} = F_{X_1|Z}^{-1}(U) + F_{X_2|Z}^{-1}(U) + \ldots + F_{X_n|Z}^{-1}(U) \tag{13}$$

$$V_\ell = E[X_1|Z] + E[X_2|Z] + \ldots + E[X_n|Z], \tag{14}$$

with U an arbitrary random variable, uniformly distributed on $[0,1]$, and with Z an arbitrary random variable, independent of U. The following relations then hold :

$$V_\ell \leq_{cx} V \leq_{cx} V_{iu} \leq_{cx} V_u. \tag{15}$$

Proof: see [1] and [5].

For each $j = 1, \ldots, n$, the terms in the original variable V and the corresponding terms in the upper bounds V_u and V_{iu} are all mutually identically distributed, i.e.

$$X_j \stackrel{d}{=} F_{X_j}^{-1}(U) \stackrel{d}{=} F_{X_j|Z}^{-1}(U). \tag{16}$$

For the lower bound, the equalities of the distributions of X_j and $E[X_j|Z]$ only hold in case all X_j, given $Z = z$, are constant for each z.

These results can be extended to the case where V consists of a sum of monotonic functions ϕ_j of random variables X_j, simply by substituting Y_j for $\phi_j(X_j)$ and applying Proposition 1, see [4, 5, 7].

3 The conditional Esscher transform of a process with stationary and independent increments

Because we consider stationary and independent increments, the unconditional Esscher transform of the process $\{X(s)\}_{0\leq s\leq t}$ equals

$$M[z, s; h^*] = E^*\left[e^{z\, X(s)}\right] = M[z, 1; h^*]^s. \tag{17}$$

The application of Jensen's inequality requires the knowledge of the conditional distribution of the process $\{X(t) - X(s)\}_{0\leq s\leq t}$, conditionally on some random variable Z. To simplify the computations, we will choose $Z = X(t)$. Then, we have for the conditional Esscher transform

$$\widetilde{M_c}[z, s; h^*] = E^*\left[e^{z(X(t)-X(s))}\,\middle|\, X(t) = c\right] \frac{d}{dc} Prob^*\left(X(t) \leq c\right) \tag{18}$$

$$= E^*\left[e^{z(X(t)-X(s))} \cdot I\left(X(t) = c\right)\right] , \tag{19}$$

where $I(\cdot)$ denotes the indicator function, i.e. $I(A) = 1$ if A is true and $I(A) = 0$ otherwise. Hence,

$$\widetilde{M_c}[z, s; h^*] = e^{zc} \int_{-\infty}^{+\infty} dx\, e^{-zx} f(x, t-s; h^*) f(c-x, s; h^*) , \tag{20}$$

where $f(x,s;h^*) = \frac{d}{dx}F(x,s;h^*)$. Inversion with respect to z gives us the density of the conditional random variable $X(t) - X(s)|X(t) = c$:

$$\tilde{f}_c(x,s;h^*) = \frac{d}{dx}Prob^*\left(X(t) - X(s) \le x | X(t) = c\right) \tag{21}$$

$$= \frac{f(x,t-s;h^*)f(c-x,s;h^*)}{f(c,t;h^*)} \tag{22}$$

and of course

$$\widetilde{F}_c(x,s;h^*) = F_{X(t)-X(s)|X(t)=c}(x,s;h^*) \tag{23}$$

$$= \int_{-\infty}^{x} \frac{f(y,t-s;h^*)f(c-y,s;h^*)}{f(c,t;h^*)}\,dy. \tag{24}$$

Consequently, for given s the inverse conditional distribution can be calculated by solving

$$u = \int_{-\infty}^{\widetilde{F}_c^{-1}(u,s;h^*)} \frac{f(x,t-s;h^*)f(c-x,s;h^*)}{f(c,t;h^*)}\,dx. \tag{25}$$

Example: Shifted inverse Gaussian process

The conditional distribution can be calculated for any of the above distributions with the right Gerber-Shiu parameterization. As an example, we consider the case of the shifted inverse Gaussian process

$$X(t) = Y(t) - \alpha t \tag{26}$$

where $\{Y(t)\}$ is an inverse Gaussian process with cumulative probability function

$$Prob[Y(t) \le y] = J(y;a,b) \qquad (y > 0) \tag{27}$$

$$= \Phi\left(\frac{-a}{\sqrt{2y}} + \sqrt{2by}\right) + e^{2a\sqrt{b}}\Phi\left(\frac{-a}{\sqrt{2y}} - \sqrt{2by}\right) \tag{28}$$

and with probability density function

$$\frac{d}{dy}Prob[Y(t) \le y] = j(y;a,b) \qquad (y > 0) \tag{29}$$

$$= \frac{a}{2\sqrt{\pi}}y^{-3/2}e^{-\frac{\left(a\sqrt{b}-2by\right)^2}{4by}}. \tag{30}$$

This gives

$$F(x,t;h^*) = J(x+\alpha t; at, b^*) \qquad (x > -\alpha t) \tag{31}$$

and

$$f(x,t;h^*) = j(x+\alpha t; at, b^*) \qquad (x > -\alpha t), \tag{32}$$

with $b^* = b - h^*$.

For the conditional distribution, applying (25) yields, for $0 \le u \le 1$,

$$u = \int_{-\alpha(t-s)}^{\widetilde{F}_c^{-1}(u,s;h^*)} \frac{j(x+\alpha(t-s); a(t-s), b^*) \;\cdot\; j(c-x+\alpha s; as, b^*)}{j(c+\alpha t; at, b^*)}\, dx \tag{33}$$

where, taking into account the support of $j(\cdot)$, the following restriction applies

$$-\alpha(t-s) \le \widetilde{F}_c^{-1}(u,s;h^*) \le c + \alpha s \tag{34}$$

or

$$0 \le \widetilde{F}_c^{-1}(u,s;h^*) + \alpha(t-s) \le c + \alpha t. \tag{35}$$

Hence,

$$u = \int_{0}^{\widetilde{F}_c^{-1}(u,s;h^*)+\alpha(t-s)} \frac{j(x; a(t-s), b^*) \;\cdot\; j(c-x+\alpha t; as, b^*)}{j(c+\alpha t; at, b^*)}\, dx\,. \tag{36}$$

4 Bounds

We now derive the upper and lower bounds in convex order for the discrete cash-flow (the continuous cash-flow arises by taking appropriate limits)

$$C(t) = \sum_{j=1}^{n} c_j e^{X(t)-X(t_j)} = c_n + \sum_{j=1}^{n-1} c_j e^{X(t)-X(t_j)}, \tag{37}$$

with $t = t_n$, using the approach described in section 2. Henceforth, we will assume that $c_j \ge 0$, $(j = 1, \ldots, n)$, merely to facilitate notation.

4.1 Upper bound

Applying (12) yields

$$E\left[(C(t)-k)_+\right] \le E\left[(C_u(t)-k)_+\right] \tag{38}$$

with

$$C_u(t) = c_n + \sum_{j=1}^{n-1} c_j e^{F^{-1}(U,t-t_j;h^*)}. \tag{39}$$

Since $\frac{d}{dk}E[(Y-k)_+] = F_Y(k)-1$ for any random variable Y and any retention k, the distribution of the upper bound follows as

$$F_u(x) = 1 - \int_0^1 du\; I\left(c_n + \sum_{j=1}^{n-1} c_j e^{F^{-1}(u,t-t_j;h^*)} geqx\right) \tag{40}$$

where $I(\cdot)$ is the indicator function, i.e. $I(A)=1$ if A holds, $I(A)=0$ if not. Hence, let u_x be defined as the value for which

$$c_n + \sum_{j=1}^{n-1} c_j e^{F^{-1}(u_x,t-t_j;h^*)} = x, \tag{41}$$

then

$$F_u(x) = u_x. \tag{42}$$

4.2 Improved upper bound

Applying (13) with $Z = X(t)$ yields

$$E\left[(C(t)-k)_+\right] \le E\left[(C_{iu}(t)-k)_+\right] \le E\left[(C_u(t)-k)_+\right] \tag{43}$$

with

$$E\left[(C_{iu}(t)-k)_+\right] = E_{X(t)}E_U\left[\left(c_n + \sum_{j=1}^{n-1} c_j e^{\widetilde{F}_{CO}^{-1}(U,t_j;h^*|X(t))} - k\right)_+\right], \tag{44}$$

where the distribution function $\widetilde{F}_{CO}(u,t_j;h^*|X(t))$ is defined by its realizations

$$\widetilde{F}_{CO}(u,s;h^*|X(t)=c) = \widetilde{F}_c(u,s;h^*). \tag{45}$$

Since the stop-loss premium for the improved upper bound can be written as

$$E\left[(C_{iu}(t)-k)_+\right] = \int_{-\infty}^{+\infty} dc\, f(c;t,h^*) \times \int_0^1 du \left(c_n + \sum_{j=1}^{n-1} c_j e^{\widetilde{F}_c^{-1}(u,t_j;h^*)} - k\right)_+ \tag{46}$$

the distribution of the improved upper bound follows as

$$F_{iu}(x) = \int_{-\infty}^{+\infty} dc\, f(c;t,h^*)\, u_x(c), \tag{47}$$

where $u_x(c)$ is defined as the root of

$$c_n + \sum_{j=1}^{n-1} c_j e^{\widetilde{F}_c^{-1}(u_x(c),t_j;h^*)} = x. \tag{48}$$

4.3 Lower bound

Finally, applying (14) with $Z = X(t)$ yields

$$E\left[(C_\ell(t)-k)_+\right] \le E\left[(C(t)-k)_+\right] \tag{49}$$

with

$$E\left[(C_\ell(t)-k)_+\right] = E_{\{X(t)\}}\left[\left(c_n + \sum_{j=1}^{n-1} c_j E^*\left[e^{X(t)-X(t_j)} | X(t)\right] - k\right)_+\right]. \tag{50}$$

The stop-loss premium for the lower bound equals

$$E\left[(C_\ell(t)-k)_+\right] = \int_{-\infty}^{+\infty} dc\, f(c;t,h^*) \times \left(c_n + \sum_{j=1}^{n-1} c_j \int_{-\infty}^{+\infty} dx\, e^x\, \tilde{f}_c(x,t_j;h^*) - k\right)_+ \tag{51}$$

so the distribution of the lower bound follows as

$$F_\ell(x) = \int_G dc\, f(c;t,h^*)\,, \tag{52}$$

where $G \subset \mathbb{R}$ is defined as the collection of all values c for which

$$c_n + \sum_{j=1}^{n-1} c_j \int_{-\infty}^{\infty} dy\, e^y\, \tilde{f}_c(y, t_j; h^*) \leq x. \tag{53}$$

5 Numerical illustration

In this section, we illustrate the upper and lower bounds by plotting their distribution functions. We assume that the process $\{X(t)\}$ is a shifted inverse Gaussian process with parameters $a = 3\sqrt{1.2}$, $b = 7.5$ and $\alpha = 0.5$ (see [2], p. 118). The parameter b^* corresponding to a risk-free force of interest $\delta = 0.1$, equals 961/120.

The distribution functions for C_u, C_{iu} and C_ℓ corresponding to a cash-flow $c_j = 10$, $j = 1, \ldots, 10$, are depicted in Figure 1. Since the upper and lower bounds appear to be rather close to each other, they prove to be quite good approximations for the unknown distribution of $C(t_n)$. The improved upper bound C_{iu} indeed improves the upper bound C_u, albeit slightly.

In order to assess the influence of the cash-flow, we change it to $c_j = j$ and $c_j = 11 - j$, $j = 1, \ldots, 10$, in Figures 2 and 3 respectively. Taking into account the scale of the price axis, the bounds appear to behave very similarly in both cases.

Corresponding author :

David Vyncke
University of Leuven, Department of Applied Economics
Minderbroedersstraat 5, B-3000 Leuven
tel : +32-16-32.37.45, fax : +32-16-32.37.40
e-mail : david.vyncke@econ.kuleuven.ac.be

References

[1] Dhaene J., Denuit M., Goovaerts M.J., Kaas R. & Vyncke D. (2001). "The Concept of Comonotonicity in Actuarial Science and Finance: Theory", *North American Actuarial Journal*, to appear.

[2] Gerber H. & Shiu E. (1994). "Option pricing by Esscher transforms", *Transactions of the society of actuaries*, vol.46, p.99–139.

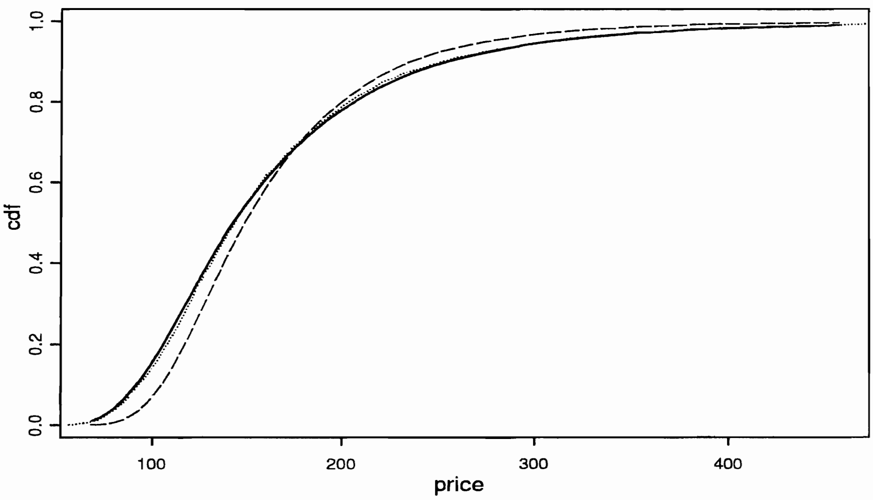
[3] Goovaerts M., De Vylder F., & Haezendonck J. (1984). *Insurance Premiums*, North-Holland, pp.XI+406.

[4] Goovaerts M., Dhaene J., & De Schepper A. (2000). "Stochastic upper bounds for present value functions", *Journal of Risk and Insurance*, vol.67(1), p.1–14.

[5] Kaas R.., Dhaene J., & Goovaerts M. (2000). "Upper and lower bounds for sums of random variables", *Insurance: Mathematics and Economics*, vol.27(2), p.151–168.

[6] Shaked M. & Shanthikumar J.G. (1994). *Stochastic orders and their applications*, Academic Press, pp.545.

[7] Vyncke D., Goovaerts M. & Dhaene J. (2001). "Convex upper and lower bounds for present value functions", *Applied Stochastic Models for Business and Industry*, 17, 149–164.



gure 1: Distribution functions of the lower bound C_ℓ (——), the improved upper ›und C_{iu} ($\cdots$) and the upper bound C_u (—) for $c_j = 10$ ($j = 1, \ldots, 10$).

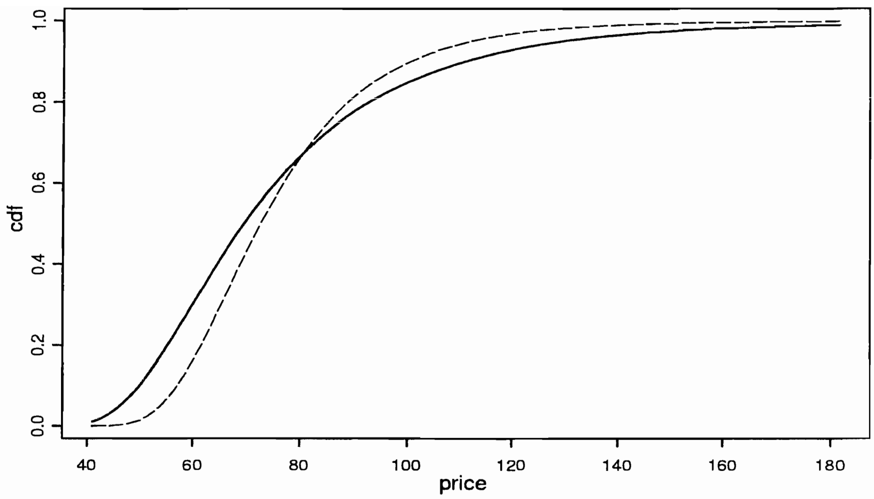


Figure 2: Distribution functions of C_ℓ (——) and C_u (—) for $c_j = 1, \ldots, 10$.

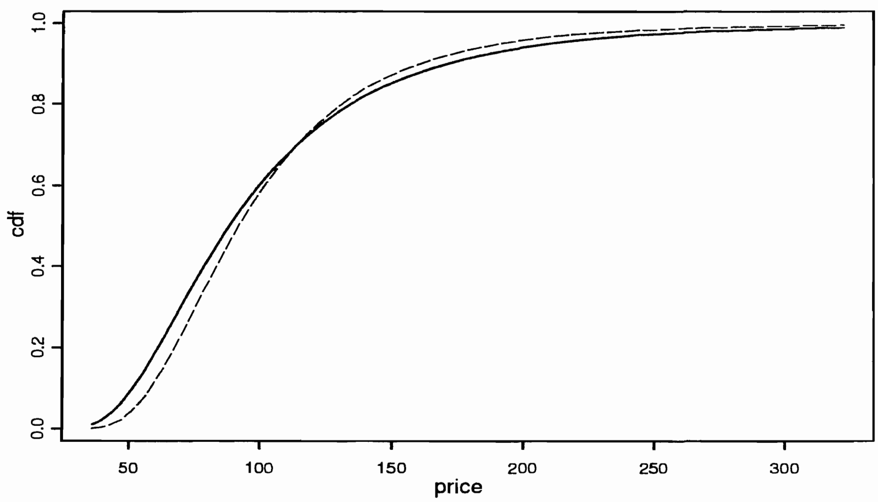


Figure 3: Distribution functions of C_ℓ (——) and C_u (—) for $c_j = 10, \ldots, 1$.

Appendix

Overview of the stochastic processes of Table 1

- Wiener process :
  $X(t) = \sigma Z(t) + \mu t$
  where $\{Z(t)\}$ is a standard Brownian motion.
- Shifted Poisson process :
  $X(t) = kN(t) - ct$
  where $\{N(t)\}$ is a Poisson process with parameter λ, k and c are positive constants.
- Random walk :
  $X(t) = X_1 + X_2 + \ldots + X_t$
  where X_j is such that $P(X_j = b) = p = 1 - P(X_j = a)$, $a < \delta < b$.
- Shifted Gamma process :
  $X(t) = Y(t) - ct$
  where $\{Y(t)\}$ is a Gamma process with parameters α and β; c is a positive constant.
- Shifted inverse Gaussian process :
  $X(t) = Y(t) - ct$
  where $\{Y(t)\}$ is an inverse Gaussian process with parameters a and b; c is a positive constant.

Overview of the functional notations of Table 1

- $N(x; \mu, \sigma^2) = \Phi\left(\frac{x-\mu}{\sigma}\right)$
- $\Lambda(x; \theta) = \sum_{k=0}^{x} \frac{e^{-\theta}\theta^k}{k!} \quad (x \geq 0)$
- $B(x; n, \theta) = \sum_{k=0}^{x} \binom{n}{k} \theta^k (1-\theta)^{n-k} \quad (x \geq 0)$
- $G(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^x y^{\alpha-1} e^{-\beta y} dy \quad (x \geq 0)$
- $J(x; a, b) = \Phi\left(\frac{-a}{\sqrt{2x}} + \sqrt{2bx}\right) + e^{2a\sqrt{b}} \Phi\left(\frac{-a}{\sqrt{2x}} - \sqrt{2bx}\right) \quad (x > 0)$

Proof of Theorem 1

The proof of Theorem 1 is based on the technique of variational calculus and adapted from Goovaerts et al. (1984). Let $u(\cdot)$ be a convex increasing utility function, and introduce $Y = \varphi(X)$. Then, because $\varphi(\cdot)$ increases continuously, we have $X = \varphi^{-1}(Y)$. Write $f(y) = \varphi^{-1}(y)$. To derive a condition for $\mathrm{E}[u(-f(Y)+\mathrm{E}[f(Y)Y])]$ to be maximal for all choices of continuous increasing functions when $\mathrm{E}[Y] = 1$, consider a function $f(y) + \varepsilon g(y)$ for some arbitrary continuous function $g(\cdot)$. A little reflection will lead to the conclusion that the fact that $f(y)$ is optimal, and this new function is not, must mean that

$$\frac{d}{d\varepsilon}\mathrm{E}\big[u\big(-f(Y) + \mathrm{E}[f(Y)Y] + \varepsilon\{-g(Y) + \mathrm{E}[g(Y)Y]\}\big)\big]\Big|_{\varepsilon=0} = 0.$$

But this derivative is equal to

$$\begin{aligned}&\mathrm{E}\big[u'\big(-f(Y) + E[f(Y)Y] + \varepsilon\{-g(Y) + \mathrm{E}[g(Y)Y]\}\big)\\ &\quad\{-g(Y) + \mathrm{E}[g(Y)Y]\}\big].\end{aligned}$$

For $\varepsilon = 0$, this derivative equals zero if

$$\begin{aligned}&\mathrm{E}\left[u'(-f(Y) + \mathrm{E}[f(Y)Y])g(Y)\right] =\\ &\mathrm{E}\left[u'(-f(Y) + \mathrm{E}[f(Y)Y])\right]\mathrm{E}[g(Y)Y].\end{aligned}$$

Writing $c = \mathrm{E}\left[u'(-f(Y) + \mathrm{E}[f(Y)Y])\right]$, this can be rewritten as

$$\mathrm{E}\left[\{u'(-f(Y) + \mathrm{E}[f(Y)Y]) - cY\}\{g(Y)\}\right] = 0.$$

Since the function $g(\cdot)$ is arbitrary, by a well-known theorem from variational calculus we find that necessarily

$$u'(-f(y) + \mathrm{E}[f(Y)Y]) - cy = 0.$$

Using $x = f(y)$ and $y = \varphi(x)$, we see that

$$\varphi(x) \propto u'(-x + \mathrm{E}[X\varphi(X)]).$$

Now, if $u(x)$ is exponential(α), so $u(x) = -\alpha e^{-\alpha x}$, then

$$\varphi(x) \propto e^{-\alpha(-x+\mathrm{E}[X\varphi(X)])} \propto e^{\alpha x}.$$

Since $\mathrm{E}[\varphi(X)] = 1$, we obtain $\varphi(x) = e^{\alpha x}/\mathrm{E}[e^{\alpha X}]$ for the optimal standardized weight function. The resulting premium is an Esscher premium with parameter $h = \alpha$. □