Chapter 2: Introduction to extreme-value theory

Stéphane Girard

Inria Grenoble Rhône-Alpes
http://mistis.inrialpes.fr/people/girard/

November-december 2018

1. Convergence in distribution of the maximum

2. Caracterisation of Maximum Domains of Attraction

Maximum Domain of Attraction

Let $X_1, \ldots, X_n$ be n i.i.d. random variables with cumulative distribution function (cdf) F. The maximum is denoted by $X_{n,n} = \max(X_1, \ldots, X_n)$. The cdf of $X_{n,n}$ is given by $\mathbb{P}(X_{n,n} \leq x) = F^n(x)$.

Definition 1

Let H be a non-degenerated cdf (*i.e.* associated with a non-constant random variable). F belongs to the maximum domain of attraction of H if and only if there exist two sequences $a_n > 0$ and b_n such that

$$F^n(a_n x + b_n) \to H(x) \text{ as } n \to \infty \text{ for all } x \in \mathbb{R}.$$

Equivalently, $(X_{n,n} - b_n)/a_n \xrightarrow{d} Y$ where Y has cdf H.

This property is denoted for short by $F \in \text{MDA}(H)$. Except some pathological cases, every cdf F belongs to a maximum domain of attraction.

Proposition 1

$F \in MDA(H)$ if and only if

$$n\bar{F}(a_n x + b_n) \to -\log H(x) \text{ as } n \to \infty \text{ for all } x \in \mathbb{R}.$$

Proof. ($\Longrightarrow$). Let us first assume that $F \in \text{MDA}(H)$. Then, there exist two sequences $a_n > 0$ and b_n such that $F^n(a_n x + b_n) \to H(x)$ as $n \to \infty$ for all $x \in \mathbb{R}$. Taking the logarithm yields

$$n \log(1 - \bar{F}(a_n x + b_n)) \to \log H(x)$$

as $n \to \infty$ for all $x \in \mathbb{R}$. This implies that $\bar{F}(a_n x + b_n) \to 0$ as $n \to \infty$. Thus, since $\log(1+u) \sim u$ as $u \to 0$, it follows that

$$n\bar{F}(a_n x + b_n) \to -\log H(x)$$

as $n \to \infty$ for all $x \in \mathbb{R}$.

($\Longleftarrow$). Conversely, let us assume that there exists two sequences $a_n > 0$ and b_n such that

$$n\bar{F}(a_n x + b_n) \to -\log H(x)$$

as $n \to \infty$ for all $x \in \mathbb{R}$. Therefore, $\bar{F}(a_n x + b_n) \to 0$ as $n \to \infty$ and

$$n \log(1 - \bar{F}(a_n x + b_n)) \sim -n\bar{F}(a_n x + b_n) \to \log H(x)$$

as $n \to \infty$. Taking the exponential concludes the proof. ■

Definition 2

Let G and H be two non-degenerated cdf. G and H are of same type if and only if there exist $a > 0$ and $b \in \mathbb{R}$ such that $H(x) = G(ax + b)$ for all $x \in \mathbb{R}$.

Lemma 1

Let F be a non-degenerated cdf. If there exists $a > 0$ and $b \in \mathbb{R}$ such that $F(ax + b) = F(x)$ for all $x \in \mathbb{R}$ then, necessarily $a = 1$ and $b = 0$.

The next proposition shows that the maximum domain of attraction is unique up to position and scale parameters.

Proposition 2

- *If $F \in MDA(G)$ and $F \in MDA(H)$ then, necessarily, G and H are of the same type i.e. there exists $a > 0$ and $b \in \mathbb{R}$ such that $G(x) = H(ax + b)$.*
- *More precisely, if $a_n > 0$, $u_n > 0$, b_n and v_n are such that $F^n(a_n x + b_n) \to G(x)$ and $F^n(u_n x + v_n) \to H(x)$ as $n \to \infty$ for all $x \in \mathbb{R}$, then, necessarily, $u_n/a_n \to a$ and $(v_n - b_n)/a_n \to b$ as $n \to \infty$.*

Extreme-Value Theorem

Theorem 1

If $F \in MDA(G)$ then, necessarily, G is of the same type as the cdf

$$H_\gamma(x) = \begin{cases} \exp[-(1+\gamma x)_+^{-1/\gamma}] & \text{if } \gamma \neq 0, \\ \exp(-e^{-x}) \text{ if } \gamma = 0, \end{cases}$$

where $y_+ = \max(0, y)$.

H_γ is referred to as the cdf of the Extreme Value Distribution (EVD).
$\gamma \in \mathbb{R}$ is called the extreme-value index.

Proof. Since $F \in MDA(G)$, there exist two sequences $a_n > 0$ and b_n such that

$$F^{[nt]}(a_{[nt]}x + b_{[nt]}) \to G(x) \qquad (1)$$

as $n \to \infty$ for all $t > 0$ and $x \in \mathbb{R}$. Moreover,

$$F^{[nt]}(a_n x + b_n) = (F^n(a_n x + b_n))^{[nt]/n} \to G^t(x) \qquad (2)$$

since $[nt]/n \to t$ as $n \to \infty$. In view of (1) and (2), Proposition 2 implies that

$$\frac{a_n}{a_{[nt]}} \to \alpha(t) > 0, \quad \frac{b_n - b_{[nt]}}{a_{[nt]}} \to \beta(t)$$

and that G and G^t are of the same type:

$$G^t(x) = G(\alpha(t)x + \beta(t)). \qquad (3)$$

The proof consists in solving equation (3) with respect to G.

For all $t > 0$ and $s > 0$, the quantity $G^{st}(x)$ can be rewritten in three different ways:

$$G^{st}(x) = G(\alpha(st)x + \beta(st)) \tag{4}$$
$$= G^{s}(\alpha(t)x + \beta(t)) = G(\alpha(s)\alpha(t)x + \alpha(s)\beta(t) + \beta(s)) \tag{5}$$
$$= G^{t}(\alpha(s)x + \beta(s)) = G(\alpha(s)\alpha(t)x + \alpha(t)\beta(s) + \beta(t)). \tag{6}$$

From (4), (5), (6) and in view of Lemma 1, if follows that

$$\alpha(st) = \alpha(s)\alpha(t) \tag{7}$$
$$\beta(st) = \alpha(s)\beta(t) + \beta(s) \tag{8}$$
$$\beta(st) = \alpha(t)\beta(s) + \beta(t), \tag{9}$$

for all $s > 0$ and $t > 0$. It can be shown that the unique positive solution of equation (7) is $\alpha(t) = t^A$, for all $t > 0$ where $A \in \mathbb{R}$. Three cases appear.

1. If $A = 0$ then $\alpha(t) = 1$ for all $t > 0$ and equations (8) and (9) can be rewritten as

$$\beta(st) = \beta(s) + \beta(t).$$

It can be shown that the unique solution of this equation is $\beta(t) = \beta(e) \log t$, for all $t > 0$. Replacing in equation (3) yields the new equation on G:

$$G^t(x) = G(x + \beta(e) \log t),$$

for all $x \in \mathbb{R}$ and $t > 0$. Besides, since G is non-degenerated, there exists $x_0 \in \mathbb{R}$ such that $0 < G(x_0) < 1$. Thus, for all $t > 1$,

$$G(x_0) > G^t(x_0) = G(x_0 + \beta(e) \log t).$$

This implies that $\beta(e) < 0$. In the following, we thus note $\sigma = -\beta(e) > 0$.

1. *(continued).* Let us now prove that $0 < G(x) < 1$ for all $x \in \mathbb{R}$. To this end, assume there exists $x_0 \in \mathbb{R}$ such that $G(x_0) = 1$. It follows that $G^t(x_0) = 1$ for all $t > 0$ and therefore $G(x_0 - \sigma \log t) = 1$ for all $t > 0$. This implies that G is constant equal to 1, which is not possible. As a consequence, $G(x) < 1$ for all $x \in \mathbb{R}$ and a similar proof can be done for $G(x) > 0$ for all $x > 0$. Let us thus introduce $\theta = -\log G(0) > 0$ and recall that

$$G^t(x) = G(x - \sigma \log t)$$

for all $x \in \mathbb{R}$ and $t > 0$. In particular, for $x = 0$, we have

$$\exp(-\theta t) = G^t(0) = G(-\sigma \log t).$$

The change of variable $u = -\sigma \log t$ thus yields

$$G(u) = \exp(-\theta \exp(-u/\sigma))$$

for all $u \in \mathbb{R}$. This shows that G is of the same type as H_0 *i.e.*

$$G(x) = H_0(x/\sigma + \log \theta).$$

2. If $A < 0$, let us introduce $\gamma = -A > 0$ leading to $\alpha(t) = t^{-\gamma}$ for all $t > 0$. Equations (8) and (9) imply that

$$\alpha(s)\beta(t) + \beta(s) = \alpha(t)\beta(s) + \beta(t)$$

for all $s > 0$ and $t > 0$ or equivalently,

$$(1 - \alpha(s))\beta(t) = (1 - \alpha(t))\beta(s).$$

Excluding the value $t = s = 1$, we thus have

$$\frac{\beta(t)}{1 - \alpha(t)} = \frac{\beta(s)}{1 - \alpha(s)} = c.$$

It follows that $\beta(t) = c(1 - \alpha(t)) = c(1 - t^{-\gamma})$ where $c \in \mathbb{R}$ and $t > 0$. Replacing in equation (3) yields the new equation

$$G^t(x) = G(t^{-\gamma}(x - c) + c)$$

for all $x \in \mathbb{R}$ and $t > 0$.

2. *(continued)*. Letting $y = x - c$ and G_* the translated cdf defined by $G_*(y) = G(y + c)$, we end up with the equation

$$G_*^t(y) = G_*(t^{-\gamma} y)$$

for all $t > 0$ and $y \in \mathbb{R}$. Moreover, since G_* and G are of the same type, it is sufficient to solve the above equation. Let $y_0 \in \mathbb{R}$ such that $G_*(y_0) < 1$. Then, letting $t \to \infty$ in the previous equation yields $G_*(0) = 0$. It follows that $G_*(y) = 0$ for all $y \leq 0$. A proof similar to the case $A = 0$ can be done to prove that $0 < G_*(1) < 1$ and we thus introduce $\sigma = -\log G_*(1) > 0$. Considering $y = 1$ in the functional equation yields

$$G_*(t^{-\gamma}) = G_*^t(1) = \exp(-\sigma t)$$

for all $t > 0$. As a consequence, the change of variable $u = t^{-\gamma}$ entails

$$G_*(u) = \exp(-\sigma u^{-1/\gamma})$$

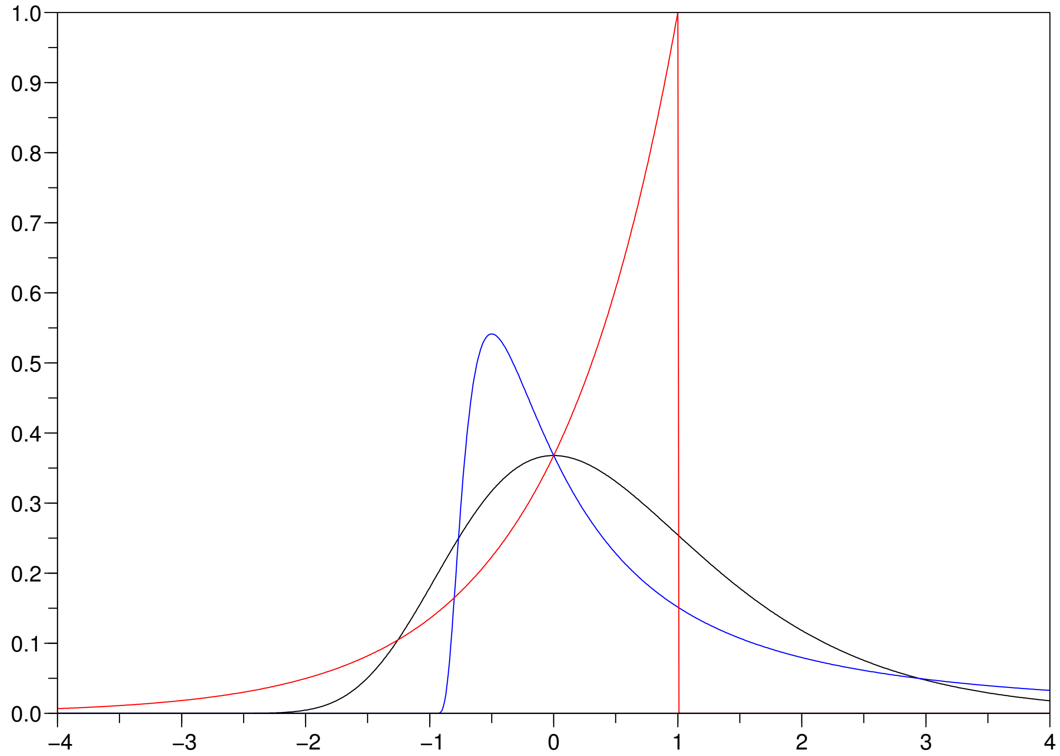
for all $u > 0$. It is easily seen that G_* is of the same type as H_γ *i.e.*

$$G_*(x) = H_\gamma(x\gamma^{-1}\sigma^{-\gamma} - \gamma^{-1})$$

for all $x \in \mathbb{R}$.

3. The case $A > 0$ is similar. ■

The proof clearly shows that there are three very different situations $\gamma < 0$, $\gamma = 0$ and $\gamma > 0$.



Example of densities associated with the extreme-value distribution (black: $\gamma = 0$, blue: $\gamma = 1$ and red: $\gamma = -1$).

Following the sign of γ, three maximum domains of attraction are defined:

- If $\gamma > 0$, F is said to belong to the Fréchet maximum domain of attraction, $F \in \text{MDA(Fréchet)}$. It will be shown later that this domain of attraction includes distribution with heavy tails, *i.e.* their survival distribution function decrease as a power function.
- If $\gamma = 0$, F is said to belong to the Gumbel maximum domain of attraction, $F \in \text{MDA(Gumbel)}$. This domain of attraction includes distributions with light tails, *i.e.* their survival distribution function decrease as an exponential rate.
- If $\gamma < 0$, F is said to belong to the Weibull maximum of attraction, $F \in \text{MDA(Weibull)}$. This domain of attraction includes distributions with short tails, *i.e.* they have a finite endpoint $x_F = \inf\{x,\ F(x) \geq 1\}$.

A classification of numerous distributions by maximum domain of attraction is available in [3, Tables 3.4.2-3.4.4].

Maximum Domain of attraction	**Gumbel** $\gamma = 0$	**Fréchet** $\gamma > 0$	**Weibull** $\gamma < 0$
Distribution	Gaussian Exponential Lognormal Gamma Weibull	Cauchy Pareto Student Burr	Uniform Beta

Maximum domains of attraction associated with usual distributions.

Exercise 1 (Standard exponential distribution)

$F(x) = 1 - \exp(-x)$, $x > 0$. *Extreme-value index* $\gamma = 0$ *(Gumbel Maximum Domain of Attraction), sequences* $a_n = 1$ *and* $b_n = \log n$.

Exercise 2 (Pareto distribution)

$F(x) = 1 - (x/a)^{-\alpha}$, $x > a > 0$ *and* $\alpha > 0$. *Extreme-value index* $\gamma = 1/\alpha > 0$ *(Fréchet Maximum Domain of Attraction), sequences* $a_n = a\alpha^{-1}n^{1/\alpha}$ *and* $b_n = an^{1/\alpha}$.

Exercise 3 (Standard uniform distribution)

$F(x) = x$ *if* $x \in [0, 1]$, $F(x) = 1$ *if* $x > 1$, $F(x) = 0$ *otherwise.* *Extreme-value index* $\gamma = -1 < 0$ *(Weibull Maximum Domain of Attraction), sequences* $a_n = 1/n$ *and* $b_n = 1 - 1/n$.

Solutions

- Standard exponential distribution $F(x) = 1 - \exp(-x)$, $x > 0$.
  $\gamma = 0$, $a_n = 1$ and $b_n = \log n$.
  For all $x > -\log n$,
  $$F^n(a_n x + b_n) = \left(1 - \frac{\exp(-x)}{n}\right)^n \to \exp(-\exp(-x)) = H_0(x).$$
- Pareto distribution $F(x) = 1 - (x/a)^{-\alpha}$, $x > a > 0$ and $\alpha > 0$.
  index $\gamma = 1/\alpha > 0$, $a_n = a\alpha^{-1}n^{1/\alpha}$ and $b_n = an^{1/\alpha}$.
  Indeed, for all $x > \alpha(n^{-1/\alpha} - 1)$,
  $$F^n(a_n x + b_n) = \left(1 - \frac{(1 + x/\alpha)^{-\alpha}}{n}\right)^n \to \exp\left(-(1 + x/\alpha)^{-\alpha}\right) = H_{1/\alpha}(x),$$
  for $x > -\alpha$.
- Uniform distribution $F(x) = x$ if $x \in [0, 1]$, $F(x) = 1$ if $x > 1$, $F(x) = 0$ otherwise. $\gamma = -1 < 0$, $a_n = 1/n$ and $b_n = 1 - 1/n$.
  Indeed, for all $x \in [1 - n, 1]$,
  $$F^n(a_n x + b_n) = \left(1 - \frac{x - 1}{n}\right)^n \to \exp(x - 1) = H_{-1}(x),$$
  for $x \leq 1$.

1. Convergence in distribution of the maximum

2. Caracterisation of Maximum Domains of Attraction

Regularly varying functions

The caracterisation of maximum domains of attractions relies on the theory of regularly-varying functions [1].

Definition 3

An eventually positive function U is regularly-varying with index $\delta \in \mathbb{R}$ at infinity if

$$\lim_{x \to \infty} \frac{U(\lambda x)}{U(x)} = \lambda^{\delta},$$

for all $\lambda > 0$. This property is denoted by $U \in \mathcal{RV}_{\delta}$. If $\delta = 0$, the function U is said to be slowly-varying.

$U \in \mathcal{RV}_{\delta}$ if and only if there exists $L \in \mathcal{RV}_0$ such that $U(x) = x^{\delta} L(x)$.

Example 4

All functions tending to a strictly positive limit are slowly-varying. Every power function of the logarithm is also slowly-varying.

As a consequence, every function in $\mathcal{RV}_{\delta}$ qualitatively behaves as $x \to x^{\delta}$ when x is large.

Proposition 3 (Karamata representation)

$L \in \mathcal{RV}_0$ if and only if L can be written as

$$L(x) = c(x) \exp\left(\int_{x_0}^{x} \frac{\Delta(t)}{t} dt \right),$$

with $x_0 > 0$ and where $c : \mathbb{R}^+ \to \mathbb{R}^+$, $\Delta : \mathbb{R}^+ \to \mathbb{R}$, $c(x) \to c > 0$, $\Delta(x) \to 0$ as $x \to \infty$.

Proposition 4 (Potter bounds)

Suppose $f \in \mathcal{RV}_\alpha$ and let $\delta_1 > 0$, $\delta_2 > 0$. Then, there exists $t_0 \in \mathbb{R}$ such that, for all $t \geq t_0$, $tx \geq t_0$,

$$(1 - \delta_1) x^\alpha \min(x^{\delta_2}, x^{-\delta_2}) < \frac{f(tx)}{f(t)} < (1 + \delta_1) x^\alpha \max(x^{\delta_2}, x^{-\delta_2}).$$

Potter bounds can be used to prove that:

Proposition 5

Let $f \in \mathcal{RV}_\alpha$, $u_n \to \infty$ *and* $v_n \sim u_n$ *as* $n \to \infty$. *Then* $f(v_n) \sim f(u_n)$ *as* $n \to \infty$.

Let us recall that the generalised inverse of a function f is defined as $f^{\leftarrow}(x) = \inf\{y,\ f(y) > x\}$. It coincides with the classical inverse f^{-1} when it exists.

Proposition 6

If $U \in \mathcal{RV}_\alpha$ *with* $\alpha < 0$ *then* $U^{\leftarrow}(1/\cdot) \in \mathcal{RV}_{-1/\alpha}$.

Fréchet Maximum Domain of Attraction

Theorem 2

- *F belongs to MDA(Fréchet) if and only if $\bar{F} := 1 - F$ is regularly varying with index $-1/\gamma$. The associated extreme-value index is γ.*
- *Moreover, a possible choice for the normalizing sequences is $a_n = F^{\leftarrow}(1 - 1/n)$ and $b_n = 0$.*

Let us highlight that necessarily the endpoint of F is infinite.

Proof. ($\Longleftarrow$). Assume that $\bar{F} \in \mathcal{RV}_{-1/\gamma}$, $\gamma > 0$. Let us introduce $a_n = F^{\leftarrow}(1 - 1/n)$ and G the cdf defined by $G(x) = \exp(-x^{-1/\gamma})$ if $x > 0$ and $G(x) = 0$ otherwise.

- Let $x > 0$. Remarking that $\bar{F}(a_n) = 1/n$ and $a_n \to \infty$ as $n \to \infty$, it follows that

$$n\bar{F}(a_n x) = \frac{\bar{F}(a_n x)}{\bar{F}(a_n)} \to x^{-1/\gamma} = -\log G(x),$$

  as $n \to \infty$, from the definition of regular variation.

- If $x < 0$ then $\bar{F}(a_n x) \to 1$ and thus $n\bar{F}(a_n x) \to \infty$ as $n \to \infty$.

As a conclusion, $n\bar{F}(a_n x) \to -\log G(x)$ for all $x \in \mathbb{R}$ and Proposition 1 entails that $F \in \text{MDA}(G)$. Clearly, G and H_γ are of the same type, $G(x) = H_\gamma((x-1)/\gamma)$, and the conclusion follows. $\blacksquare$

Weibull Maximum Domain of Attraction

For all cdf F with finite endpoint x_F, we denote by F_* the cdf defined by $F_*(x) = F(x_F - 1/x)$ if $x > 0$ and $F_*(x) = 0$ otherwise.

Theorem 3

- *F belongs to MDA(Weibull) if and only if x_F is finite and F_* belongs to MDA(Fréchet). Letting $\gamma > 0$ the extreme-value index associated with F_*, the extreme-value index associated with F is $-\gamma$.*
- *Besides, a possible choice of normalizing sequences is $a_n = x_F - F^{\leftarrow}(1 - 1/n)$ and $b_n = x_F$.*

Proof. ($\Longleftarrow$). Assume that x_F is finite and $F_* \in$ MDA(Fréchet). Letting $a_n^* = F_*^{\leftarrow}(1 - 1/n)$, Theorem 2 entails that $F_*^n(a_n^* x) \to \exp(-x^{-1/\gamma})$ if $x > 0$ and $F_*^n(a_n^* x) \to 0$ otherwise, when $n \to \infty$ and for some $\gamma > 0$.

- Let $x > 0$ and introduce $a_n = 1/a_n^*$. The above convergence can be rewritten as
$$F^n(x_F - a_n/x) \to \exp(-x^{-1/\gamma})$$
as $n \to \infty$. Letting $y = -1/x < 0$, it follows that
$$F^n(x_F + a_n y) \to \exp(-(-y)^{1/\gamma})$$
as $n \to \infty$. Remarking that a_n can be rewritten as $a_n = x_F - F^{\leftarrow}(1 - 1/n)$ and introducing $b_n = x_F$ and the cdf G defined by $G(y) = \exp(-(-y)^{1/\gamma})$ if $y < 0$ and $G(y) = 1$ otherwise, we have proven that $F^n(a_n y + b_n) \to G(y)$ as $n \to \infty$ for all $y < 0$.
- The case $y \geq 0$ is straightforward.

As a conclusion, $F \in$ MDA(G), G is of the same type as $H_{-\gamma}$ with $-\gamma < 0$, and thus F belongs to the Weibull Maximum Domain of Attraction. ■

Gumbel Maximum Domain of Attraction

Theorem 4

F belongs to MDA(Gumbel) if and only if there exists $x_0 < x_F \leq \infty$ such that

$$\bar{F}(x) = c(x) \exp\left(- \int_{x_0}^{x} \frac{g(t)}{a(t)} dt \right),$$

where a, c and g are three functions verifying $a'(x) \to 0$, $c(x) \to c > 0$ and $g(x) \to 1$ as $x \to x_F$.

Exercise 4 (Standard Gaussian distribution)

The density is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} \right),$$

and the survival function verifies $\bar{F}(x) \sim f(x)/x$ as $x \to \infty$. Check that $\mathcal{N}(0,1) \in$ MDA(Gumbel).

Solution. Letting,

- $x_0 > 0$,
- $\lambda(x) = x\bar{F}(x)/f(x) \to 1$ as $x \to \infty$,
- $k(x_0) = x_0 \exp(x_0^2/2)$,
- $g(x) = 1$,
- $a(x) = x/(x^2+1)$,
- $c(x) = \lambda(x)/(\sqrt{2\pi}k(x_0))$,

we end up with

$$\bar{F}(x) = c(x) \exp\left(-\int_{x_0}^{x} \frac{g(t)}{a(t)} dt\right).$$