2

Mathematical Fundamentals

In this chapter, we present the essential mathematical tools needed in the modelling of portfolio credit derivative products. This includes: doubly-stochastic Poisson processes, also known as Cox processes; point processes and their intensities, on some given filtration; and copula functions.

2.1 Credit Pricing Building Blocks

The key building blocks needed for pricing single-name credit products are: expectations of risky cash-flows, at fixed time horizons, conditional on survival; and expectations of recovery payments at the time of default. Those were derived by Lando (1998) in a doubly-stochastic Poisson process (also known as a Cox process) framework.

2.1.1 Cox Process

An inhomogeneous Poisson process N, with non-negative intensity function $h(.)$ is defined as a process with independent increments such that

$$\mathbb{P}(N_t - N_s = k) = \frac{\left(\int_s^t h(u)\,du\right)^k}{k!} \exp\left(-\int_s^t h(u)\,du\right), \text{ for } k = 0, 1, \ldots$$

Y. Elouerkhaoui, *Credit Correlation*, Applied Quantitative Finance,
DOI 10.1007/978-3-319-60973-7_2

Define τ as the first jump time of the Poisson process N. The probability of survival after time T, which is equivalent to $N_T = 0$, is

$$\mathbb{P}(\tau > T) = \mathbb{P}(N_t = 0) = \exp\left(-\int_0^T h(u)\, du\right).$$

This gives an operational recipe for simulating the first jump time of a Poisson process: let E_1 be a unit exponential random variable, then τ can be determined by setting

$$\tau = \inf\left\{t : \int_0^t h(u)\, du \geq E_1\right\}.$$

A Cox process is a generalized Poisson process where the intensity itself is stochastic: so, instead of having a time-dependent intensity function $h(.)$, the intensity is a stochastic process $h(., \omega)$. Conditional on each realization $\omega \in \Omega$, of the intensity process $h(., \omega)$, the counting process N is an inhomogeneous Poisson process with intensity $h(t, \omega)$.

Furthermore, we assume that the stochastic intensity can be written as the functional form

$$h(t, \omega) = \lambda(X_t),$$

where X is an $\mathbb{R}^d$-valued Ito process representing the background state variables in the economy, and $\lambda(.) : \mathbb{R}^d \to [0, \infty)$ is a non-negative, continuous function. The economic state variables would include: (risk-free) interest rates, equity prices, credit ratings and other macroeconomic variables, which would drive the likelihood of default; but they would exclude the actual default events of the obligor in question and other obligors in the economy. The stochastic intensity $\lambda(X_t)$ can be thought of as a (conditional) instantaneous default probability; in other words, conditional on the firm having survived up to time t, and for a given observed path of the state variables X up to time t, the probability of defaulting in the next instant, between t and $t + dt$, is equal to $\lambda(X_t)\, dt + o(dt)$.

More formally, if we work on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$, where we have an an $\mathbb{R}^d$-valued Ito process X, and a unit exponential variable E_1, independent from X, we define the default time τ as

$$\tau = \inf\left\{t : \int_0^t \lambda(X_u)\, du \geq E_1\right\}.$$

This is the exact analogue to the Poisson construction algorithm, but here we are using a random intensity instead.

Having modeled the default time as the first jump time of a Cox process, we can now write the survival probability, conditional on X, as

$$\mathbb{P}\left(\tau > T \left| (X_t)_{0\leq t\leq T}\right.\right) = \exp\left(-\int_0^T \lambda\left(X_u\right) du\right),$$

which yields the expression of the survival probability by taking the expectation on both sides

$$\mathbb{P}\left(\tau > T\right) = \mathbb{E}\left[\exp\left(-\int_0^T \lambda\left(X_u\right) du\right)\right].$$

Next, we evaluate the basic three pricing building blocks that we need.

2.1.2 Three Building Blocks

First, we need to make precise the various filtrations that we work with. As with all credit modelling problems, we will have three filtrations: (1) the background filtration containing information about the state variables in the economy; (2) the default filtration, which tracks the obligor default events history; (3) the enlarged filtration, which combines both the economic state variables and the default events information; the three filtrations are defined as

$$\begin{aligned}
\mathcal{F}_t &= \sigma\left\{X_s : 0 \leq s \leq t\right\}; \\
\mathcal{H}_t &= \sigma\left\{\mathbf{1}_{\{\tau\leq t\}} : 0 \leq s \leq t\right\}; \\
\mathcal{G}_t &= \mathcal{F}_t \vee \mathcal{H}_t.
\end{aligned}$$

The conditional survival probability is given by the following lemma.

Lemma 5 (Conditional Survival Probability) *The survival probability conditional on (the whole path) $\mathcal{F}_\infty$ and (the default state) $\mathcal{H}_t$ is given by*

$$\mathbb{P}\left(\tau > T \left| \mathcal{F}_\infty \vee \mathcal{H}_t\right.\right) = \mathbf{1}_{\{\tau>t\}} \exp\left(-\int_t^T \lambda_s ds\right).$$

Proof It suffices to observe that the conditional expectation $\mathbb{E}\left[\mathbf{1}_{\{\tau>t\}}\,|\mathcal{F}_\infty \vee \mathcal{H}_t\,\right]$ is 0 on the set $\{\tau \leq t\}$ and that $\{\tau > t\} \in \mathcal{H}_t$; then, using Bayes' rule, we obtain

$$\begin{aligned}
\mathbb{P}\,(\tau > T\,|\mathcal{F}_\infty \vee \mathcal{H}_t\,) &= \mathbf{1}_{\{\tau>t\}}\mathbb{P}\,(\tau > T\,|\mathcal{F}_\infty \vee \mathcal{H}_t\,) \\
&= \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{P}\,(\{\tau > T\} \cap \{\tau > t\}\,|\mathcal{F}_\infty\,)}{\mathbb{P}\,(\tau > t\,|\mathcal{F}_\infty\,)} \\
&= \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{P}\,(\tau > T\,|\mathcal{F}_\infty\,)}{\mathbb{P}\,(\tau > t\,|\mathcal{F}_\infty\,)} \\
&= \mathbf{1}_{\{\tau>t\}}\frac{\exp\left(-\int_0^T \lambda_s ds\right)}{\exp\left(-\int_0^t \lambda_s ds\right)} = \mathbf{1}_{\{\tau>t\}}\exp\left(-\int_t^T \lambda_s ds\right).
\end{aligned}$$

■

To price any credit risky (defaultable) contingent claim, we have to compute expectations of its discounted cash-flows, which can be one of three types:

1. Payment at Maturity $-X_T\mathbf{1}_{\{\tau>T\}}$: a cash-flow payment X_T, which is an $\mathcal{F}_T$-measurable variable, at a fixed time horizon T, if default has not occurred before time T;
2. Coupon Payments $-\mathbf{1}_{\{\tau>s\}}Y_s ds$: a stream of (continuous) payments specified by an $\mathcal{F}_t$-adapted process Y, which terminates when the default event happens;
3. Payment at Default $-Z_\tau$: A recovery rate payment, at the time of default τ, where Z is an $\mathcal{F}_t$-adapted process; the payment at default is the random variable $Z_\tau = Z_{\tau(\omega)}\,(\omega)$.

The key formulas are summarized in the next proposition.

Proposition 6 (Three Building Blocks) *We have the following conditional expectations for the three cash-flow types above.*

1. Payment at Maturity:

$$\mathbb{E}\left[e^{\left(-\int_t^T r_s ds\right)}X_T\mathbf{1}_{\{\tau>T\}}\,|\mathcal{G}_t\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}\left[e^{\left(-\int_t^T (r_s+\lambda_s)ds\right)}X_T\,|\mathcal{F}_t\right].$$

2. Coupon Payments:

$$\mathbb{E}\left[\int_t^T e^{\left(-\int_t^s r_u du\right)}Y_s\mathbf{1}_{\{\tau>s\}}ds\,|\mathcal{G}_t\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}\left[\int_t^T e^{\left(-\int_t^s (r_u+\lambda_u)du\right)}Y_s ds\,|\mathcal{F}_t\right].$$

3. Payment at Default:

$$\mathbb{E}\left[e^{\left(-\int_t^\tau r_u du\right)} Z_\tau \,|\mathcal{G}_t\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}\left[\int_t^T \lambda_s e^{\left(-\int_t^s (r_u+\lambda_u)du\right)} Z_s ds \,|\mathcal{F}_t\right].$$

Note that the expectations for payments types 1 and 2, i.e., the payment at maturity and the coupon payments, are not dissimilar. In fact, the coupon payments integral is just a linear sum of multiple $\mathcal{F}_s$ -measurable payments conditional on survival after time s.

Proof First, we start with expectations conditional on survival at a fixed time horizon.

1. Using the law of iterated expectations and the conditional survival probability from the previous lemma, we have

$$\begin{aligned}\mathbb{E}\left[e^{\left(-\int_t^T r_s ds\right)} X_T \mathbf{1}_{\{\tau>T\}} \,|\mathcal{G}_t\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{\left(-\int_t^T r_s ds\right)} X_T \mathbf{1}_{\{\tau>T\}} \,|\mathcal{F}_\infty \vee \mathcal{H}_t\right] |\mathcal{G}_t\right] \\ &= \mathbb{E}\left[e^{\left(-\int_t^T r_s ds\right)} X_T \mathbb{E}\left[\mathbf{1}_{\{\tau>T\}} \,|\mathcal{F}_\infty \vee \mathcal{H}_t\right] |\mathcal{G}_t\right] \\ &= \mathbf{1}_{\{\tau>t\}}\mathbb{E}\left[e^{\left(-\int_t^T (r_s+\lambda_s)ds\right)} X_T \,|\mathcal{G}_t\right]\end{aligned}$$

The last step that we need is to switch the conditioning on the enlarged filtration $\mathcal{G}_t$ to the conditioning on the background filtration $\mathcal{F}_t$. Notice the following sigma fields' inclusions

$$\mathcal{F}_t \subset \mathcal{F}_t \vee \mathcal{H}_t \subset \mathcal{F}_t \vee \sigma\left(E_1\right);$$

and recall from the Cox process construction the independence between the (threshold) exponential random variable E_1 and the sigma filed $\sigma\left(\widetilde{X_T}\right) \vee \mathcal{F}_t$, where $\widetilde{X_T} = e^{\left(-\int_t^T (r_s+\lambda_s)ds\right)} X_T$, so that we can write

$$\mathbb{E}\left[e^{\left(-\int_t^T (r_s+\lambda_s)ds\right)} X_T \,|\mathcal{F}_t \vee \sigma\left(E_1\right)\right] = \mathbb{E}\left[e^{\left(-\int_t^T (r_s+\lambda_s)ds\right)} X_T \,|\mathcal{F}_t\right],$$

which gives the final result.

2. The proof for expectations of coupon payment streams is exactly identical to the proof for cash-flow payments at a fixed time horizon T.

3. For recovery payments, we use the expression of the default time density conditional on $\mathcal{F}_\infty$: for all $s > t$,

$$\mathbb{P}(\tau \in ds \,|\tau > t, \mathcal{F}_\infty) = \frac{\partial}{\partial s}\mathbb{P}(\tau \leq s \,|\tau > t, \mathcal{F}_\infty) = \lambda_s \exp\left(-\int_t^s \lambda_u du\right).$$

Thus, we can write the recovery expectation in terms of the conditional default density as

$$\begin{aligned}
\mathbb{E}\left[e^{\left(-\int_t^\tau r_u du\right)} Z_\tau \,|\mathcal{G}_t\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{\left(-\int_t^\tau r_u du\right)} Z_\tau \,|\mathcal{F}_\infty \vee \mathcal{H}_t\right] |\mathcal{G}_t\right] \\
&= \mathbf{1}_{\{\tau > t\}}\mathbb{E}\left[\int_t^T \lambda_s e^{\left(-\int_t^s \lambda_u du\right)}\left[e^{\left(-\int_t^s r_u du\right)} Z_s\right] ds \,|\mathcal{G}_t\right] \\
&= \mathbf{1}_{\{\tau > t\}}\mathbb{E}\left[\int_t^T \lambda_s e^{\left(-\int_t^s (r_u + \lambda_u) du\right)} Z_s ds \,|\mathcal{F}_t\right];
\end{aligned}$$

in the last line, we have replaced the conditioning on $\mathcal{G}_t$ with the conditioning on $\mathcal{F}_t$ following the same argument as before.

■

Next, we give a brief overview of the theory of point processes and the general definition of intensity processes with respect to a given filtration. This is a generalization of the results obtained in the Cox process framework. The choice of filtration (and its corresponding intensity process) is of critical importance when we work in a credit portfolio set-up with multiple default times. Single-name default intensities (on the enlarged filtration) can be distorted as (portfolio) default events occur, which creates some interesting default clustering patterns.

2.2 Point Processes, Filtrations and Intensities

We have seen in the previous section a definition of the default event (stopping) time based on a Cox process approach. The default time is constructed, from first principles, as the first jump time of Poisson process whose intensity is stochastic and driven by some economic state variables (Ito) process. It turns out that we can go one step further and model the default event, in a more general framework, as a stopping time with respect to a given filtration. But as soon as we do that, extra care and attention need to be given to the choice of (working) filtration and the exact definition of the intensity process (associated with this default time). Intuitively, when we have a general point process N_t, on some filtration $\mathcal{F}_t$, the heuristic definition of an ($\mathcal{F}_t$-) intensity is simply the conditional instantaneous jump probability:

$$\mathbb{P}\left(dN_t = 1 \,|\mathcal{F}_t\right) = \lambda_t dt + o\left(dt\right).$$

This concept has been formalized in a mathematically rigorous manner by Brémaud (1980). This is the definitive mathematical bible on point processes (and market point processes) that any credit modeler (keen on mathematical rigour) needs to refer it continuously.

Here we follow the pedagogical presentation in Brémaud (1980) and we summarize some of the key results, which we will need later in the sequel.

2.2.1 Counting Process

A point process on the half line $[0, \infty)$ can be represented in one of three ways: (a) as a sequence of non-negative random times; (b) as a discrete random measure; (c) or as the associated counting process. Here we shall use the point process terminology to refer to both the sequence of random times and their counting process interchangeably.

Definition 7 (*Point Process*) A point process on can be described by a sequence of non-negative random times, on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that

$$\begin{aligned} &T_0 = 0, \\ &T_n < \infty \Longrightarrow T_n < T_{n+1}. \end{aligned}$$

A realization of the point process is said to be non-explosive if $T_\infty = \lim_{n \uparrow \infty} T_n = +\infty$. For each realization T_n corresponds a counting function N, defined as

$$N_t = \begin{cases} n & \text{if } t \in \left[T_n, T_{n+1}\right), n \geq 0, \\ +\infty & \text{if } t \geq T_\infty. \end{cases}$$

The sequence T_n is called a point process; but sometimes the associated counting process N is also called a point process by abuse of notation. We say that it is non-explosive if $N_t < \infty$ a.s., for $t \geq 0$. And we say that the point process N_t is integrable if $\mathbb{E}\left[N_t\right] < \infty$, for $t \geq 0$.

We can also have a multivariate point process. The definition is given below.

Definition 8 (*Multivariate Point Process*) Let T_n be a point process, and let Z_n be a sequence of random variables in $\{1, 2, \ldots, k\}$. Define for all $1 \leq i \leq k$,

$$N_t\left(i\right) = \sum_{n \geq 1} \mathbf{1}\left\{T_n \leq t\right\} \mathbf{1}\left\{Z_n = i\right\}, \text{ for } t \geq 0.$$

The k-dimensional vector process $(N_t(1), \ldots, N_t(k))$ and the double sequence $(T_n, Z_n, n \geq 1)$ are called k-variate point processes. The processes $N_t(i)$ have no common jumps, i.e., for $i \neq j$, we have $\Delta N_t(i)\, \Delta N_t(j) = 0$ a.s., for $t \geq 0$.

2.2.2 Doubly Stochastic Poisson Process

Broadly speaking, a doubly stochastic Poisson process is generated with a two-step procedure: first, we simulate a full path of the intensity-driving background process X; then, given this realized path, we generate a Poisson process with intensity $\lambda_t = h(X_t)$. This is the Cox process construction that we have discussed in Sect. 2.1.1.

This definition can be generalized and extended formally to a larger class of intensity processes.

Definition 9 (*Doubly Stochastic Poisson Processes*) Let N_t be a point process adapted to some filtration $\mathcal{F}_t$, and let λ_t be a non-negative measurable process. Suppose that

$$\lambda_t \text{ is } \mathcal{F}_0\text{-measurable, for } t \geq 0,$$

and that

$$\int_0^t \lambda_s ds < \infty \text{ a.s., for } t \geq 0.$$

If we have, for all $0 \leq s \leq t$ and all $u \in \mathbb{R}$,

$$\mathbb{E}\left[e^{iu(N_t - N_s)} | \mathcal{F}_s\right] = \exp\left(\left(e^{iu} - 1\right) \int_s^t \lambda_v dv\right),$$

then the process N_t is called a $(\mathbb{P}, \mathcal{F}_t)$-doubly stochastic Poisson process with (stochastic) $\mathcal{F}_t$-intensity.

If the intensity λ_t is deterministic (and, therefore, is a time-dependent function $\lambda(t)$) then N_t is called an $(\mathbb{P}, \mathcal{F}_t)$-Poisson process.

If the filtration is restricted to its natural filtration $\mathcal{F}_t = \mathcal{F}_t^N$, then we say that N_t is a Poisson process with intensity $\lambda(t)$.

If the intensity is equal to one, $\lambda(t) = 1$, then N_t is a standard Poisson process.

Remark 10 From the definition above, it follows that:

- the Poisson process increments $N_t - N_s$ are independent from $\mathcal{F}_s$ conditionally on (the path) $\mathcal{F}_0$. Indeed, since λ_t is $\mathcal{F}_0$-measurable, we can write

$$\mathbb{E}\left[e^{iu(N_t-N_s)}\,|\mathcal{F}_s \vee \mathcal{F}_0\right] = \mathbb{E}\left[\exp\left(\left(e^{iu}-1\right)\int_s^t \lambda_v dv\right)|\mathcal{F}_0\right] = \mathbb{E}\left[e^{iu(N_t-N_s)}\,|\mathcal{F}_0\right];$$

- for all $0 \leq s \leq t$, the probability distribution of the increment $N_t - N_s$ conditional on $\mathcal{F}_s$, is given by

$$\mathbb{P}\left(N_t - N_s = k\,|\mathcal{F}_s\right) = \frac{\left(\int_s^t \lambda_v dv\right)^k}{k!} e^{-\int_s^t \lambda_v dv}, \text{ for } k \geq 0.$$

An alternative definition of a doubly stochastic process is based on a result due to Watanabe (1964): it offers a more general characterization, which can be extended to define the (stochastic) intensity for any general point process (which is not necessarily Poisson or doubly stochastic Poisson).

2.2.3 Watanabe's Characterization

Let N_t be a doubly stochastic Poisson process with an $\mathcal{F}_t$ -intensity λ_t. Using the $\mathcal{F}_s$-conditional probability distribution of the increment $N_t - N_s$, we can write

$$\mathbb{E}\left[N_t - N_s\,|\mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \lambda_u du\,|\mathcal{F}_s\right].$$

Suppose that the cumulative intensity is integrable, i.e., $\mathbb{E}\left[\int_0^t \lambda_u du\right] < \infty$, for all $t \geq 0$, then from the equation above the process N_t is also integrable $\mathbb{E}\left[N_t\right] < \infty$; hence, the process M_t defined as

$$M_t = N_t - \int_0^t \lambda_u du$$

is an $\mathcal{F}_t$-martingale. Furthermore, for all non-negative $\mathcal{F}_t$-predictable processes C_t, we have

$$\mathbb{E}\left[\int_0^\infty C_s dN_s\right] = \mathbb{E}\left[\int_0^\infty C_s \lambda_s ds\right].$$

This characterization can be used as a definition for the intensity of a doubly stochastic Poisson process.

Theorem 11 (Characterization of Doubly Stochastic Poisson Processes) *Let N_t be a point process adapted to some filtration $\mathcal{F}_t$, and let λ_t be a non-negative measurable process such that: for all $t \geq 0$, λ_t is -measurable, and $\int_0^t \lambda_s ds < \infty$ a.s.*

If the equality

$$\mathbb{E}\left[\int_0^\infty C_s dN_s\right] = \mathbb{E}\left[\int_0^\infty C_s \lambda_s ds\right]$$

holds for all non-negative $\mathcal{F}_t$-predictable processes C_t, then N_t is a doubly stochastic process with $\mathcal{F}_t$-intensity λ_t.

Watanabe in 1964 came up with the first important characterization property, which links point processes and martingales. His characterization result relates to Poisson processes.

Theorem 12 (Watanabe 1964) *Let N_t be a point process adapted to some filtration $\mathcal{F}_t$, and let $\lambda(t)$ be a locally integrable non-negative measurable function. Suppose that $M_t = N_t - \int_0^t \lambda(s)\, ds$ is an $\mathcal{F}_t$-martingale. Then, N_t is an $\mathcal{F}_t$-Poisson process with intensity $\lambda(t)$, i.e., for all $0 \leq s \leq t$, the increment $N_t - N_s$ is a Poisson random variable with parameter $\int_s^t \lambda(u)\, du$, which is independent from $\mathcal{F}_s$.*

2.2.4 Stochastic Intensity

In the general case, to define the $\mathcal{F}_t$-intensity for any point process (adapted to the filtration $\mathcal{F}_t$), we can use the previous doubly stochastic Poisson process characterization theorem.

Definition 13 (*Stochastic Intensity*) Let N_t be a point process adapted to some filtration $\mathcal{F}_t$, and let λ_t be a non-negative $\mathcal{F}_t$-progressive process such that $\int_0^t \lambda_s ds < \infty$ a.s, for all $t \geq 0$. If for all non-negative $\mathcal{F}_t$-predictable processes C_t, the equality

$$\mathbb{E}\left[\int_0^\infty C_s dN_s\right] = \mathbb{E}\left[\int_0^\infty C_s \lambda_s ds\right]$$

holds, then we say that the process N_t admits a $(\mathbb{P}, \mathcal{F}_t)$-intensity (or $\mathcal{F}_t$-intensity) λ_t.

Using this definition, we have the following integration theorem.

Theorem 14 (Integration Theorem) *If N_t admits an $\mathcal{F}_t$-intensity λ_t (where $\int_0^t \lambda_s ds < \infty$ a.s, for all $t \geq 0$), then N_t is non-explosive and*

- *$M_t = N_t - \int_0^t \lambda_s ds$ is an $\mathcal{F}_t$-local martingale;*
- *if X_t is an $\mathcal{F}_t$-predictable process such that $\mathbb{E}\left[\int_0^t |X_s| \lambda_s ds\right] < \infty$, $t \geq 0$, then $\int_0^t X_s dM_s$ is an $\mathcal{F}_t$-martingale;*
- *if X_t is an $\mathcal{F}_t$-predictable process such that $\int_0^t |X_s| \lambda_s ds < \infty$, $t \geq 0$, then $\int_0^t X_s dM_s$ is an $\mathcal{F}_t$-local martingale.*

The next martingale characterization theorem is the main result that we shall use to define the intensity for default (stopping) times.

Theorem 15 (Martingale Characterization of Intensity) *Let N_t be a non-explosive point process adapted to the filtration $\mathcal{F}_t$. Suppose that for some non-negative $\mathcal{F}_t$-progressive process λ_t and for all $n \geq 1$,*

$$M_{t \wedge T_n} = N_{t \wedge T_n} - \int_0^{t \wedge T_n} \lambda_s ds, \text{ is an } (\mathbb{P}, \mathcal{F}_t) \text{-martingale.}$$

Then, λ_t is the $\mathcal{F}_t$-intensity of point process N_t.

One can observe that, using this intensity martingale characterization property, the following equality holds:

$$\mathbb{E}\left[N_{t \wedge T_n} - N_{s \wedge T_n} | \mathcal{F}_s\right] = \mathbb{E}\left[\int_{s \wedge T_n}^{t \wedge T_n} \lambda_u du \, | \mathcal{F}_s\right],$$

which, by letting $n \uparrow \infty$, becomes

$$\mathbb{E}\left[N_t - N_s | \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \lambda_u du \, | \mathcal{F}_s\right].$$

This is reminiscent of one of the classical definitions of intensity. In particular, if we assume that λ_t is right-continuous and bounded, then by applying successively the Lebesgue averaging theorem and the Lebesgue dominated convergence theorem, we can see that the $\mathcal{F}_t$-conditional probability of instantaneous

jumps (in the point process) is equal to the $\mathcal{F}_t$-intensity

$$\lim_{t \downarrow s} \frac{1}{t-s} \mathbb{E}\left[N_t - N_s \,|\mathcal{F}_s \right] = \lambda_s, \text{ a.s.}$$

2.2.5 Predictable Intensities

So far, we have given a characterization of an $\mathcal{F}_t$-intensity process for a point process N_t, but we have not said anything about its uniqueness. In general, the $\mathcal{F}_t$-intensity, as defined previously, is not unique. But we can always find a predictable version of the intensity, which is made unique by the predictability constraint.

The formal results regarding uniqueness and existence of predictable versions are given in the next two theorems.

Theorem 16 (Uniqueness of Predictable Intensities) *Let N_t be a point process adapted to the filtration $\mathcal{F}_t$. Let λ_t and $\widetilde{\lambda}_t$ be two $\mathcal{F}_t$-intensities of the point process N_t, which are $\mathcal{F}_t$-predictable, then*

$$\lambda_t(\omega) = \widetilde{\lambda}_t(\omega), \quad \mathbb{P}(d\omega)\, dN_t(\omega) \text{-a.e.}$$

In particular, for $n \geq 1$, we have

$$\begin{aligned} \lambda_{T_n} &= \widetilde{\lambda}_{T_n}, \qquad \text{on } \{T_n < \infty\}, \\ \lambda_t(\omega) &= \widetilde{\lambda}_t(\omega), \quad \lambda_t(\omega)\, dt \text{ and } \widetilde{\lambda}_t(\omega)\, dt \text{-a.e.} \end{aligned}$$

Theorem 17 (Existence of Predictable Versions of Intensities) *Let N_t be a point process with an $\mathcal{F}_t$-intensity λ_t. Then, one can find an $\mathcal{F}_t$-intensity $\widetilde{\lambda}_t$ that it predictable.*

Now when we talk about the $\mathcal{F}_t$-intensity of the point process N_t (as opposed to an $\mathcal{F}_t$-intensity), we are referring to the (unique) predictable version.

2.2.6 Change of Filtration

A very important result, which forms the foundation of everything that one does when working on enlarged (credit) filtrations –generated by the default events of a credit portfolio– is the change of filtration theorem. As one switches between the enlarged portfolio filtration and the individual single-name filtrations (or other sub-basket filtrations), one needs to pay special attention

to the intensities that are used as they have a fundamental impact on all the conditional expectation calculations that one performs.

We state the change of filtration theorem next.

Theorem 18 (Change of Filtration for Intensities) *Let N_t be a point process with the $\mathcal{F}_t$-intensity λ_t. Let $\mathcal{G}_t$ be a sub-filtration of N_t smaller than $\mathcal{F}_t$, i.e.,*

$$\mathcal{F}_t^N \subset \mathcal{G}_t \subset \mathcal{F}_t, \quad t \geq 0.$$

Then, the process N_t admits a $\mathcal{G}_t$-intensity μ_t defined by

$$\mu_t(\omega) = \left(\frac{\lambda_u d\mathbb{P} du}{d\mathbb{P} du} \right)(t, \omega), \quad on\ \mathcal{P}(\mathcal{G}_t).$$

Loosely speaking, this can be re-stated as: if N_t is a point process with the $\mathcal{F}_t$-intensity λ_t, and if $\mathcal{G}_t$ is a sub-filtration of N_t, which is smaller than $\mathcal{F}_t$, then $\mu_t = \mathbb{E}\left[\lambda_t \,|\mathcal{G}_t\right]$ is the $\mathcal{G}_t$-intensity of N_t.

Now, looking more closely at a multivariate point process, we can describe the (discrete) conditional probability density of the embedded mark process $(Z_n, n \geq 0)$ in terms of the point process intensities. In practice, this is a useful property that usually comes in handy when we wish to implement efficient Monte-Carlo simulation algorithms for multi-name baskets.

Theorem 19 *Let $(T_n, Z_n, n \geq 0)$ be an m-variate point process, and let $N_t(i)$, for $1 \leq i \leq m$, be its associated counting processes. Let $\mathcal{F}_t$ be a filtration of the form*

$$\mathcal{F}_t = \mathcal{F}_0 \vee \left(\bigvee_{i=1}^{m} \mathcal{F}_t^{N(i)} \right),$$

where is $\mathcal{F}_t^{N(i)}$ the filtration of the process $N_t(i)$. Suppose that, for each $1 \leq i \leq m$, $N_t(i)$ admits the $\mathcal{F}_t$-intensity $\lambda_t(i)$. Then, for all $n \geq 1$,

$$\frac{\lambda_{T_n}(i)}{\lambda_{T_n}} = \mathbb{P}\left(Z_n = i \,\middle|\, \mathcal{F}_{T_n^-} \right), \text{ on } \{T_n < \infty\},$$

where $\lambda_t = \sum_{i=1}^m \lambda_t(i)$ is the -intensity of the process $N_t = \sum_{i=1}^m N_t(i)$.

Broadly speaking, the ratio $\frac{\lambda_t(i)}{\sum_{i=1}^m \lambda_t(i)}$ is the probability of having a jump of type i, at time t, conditional on $\mathcal{F}_{t^-}$, and knowing that we have a jump in one

of the m point processes $N_t(j)$ at time t. We could write it heuristically as

$$\begin{aligned}\mathbb{P}(dN_t(i) = 1 \,|\mathcal{F}_{t^-}, dN_t = 1) &= \frac{\mathbb{P}(dN_t(i) = 1, dN_t = 1 \,|\mathcal{F}_{t^-})}{\mathbb{P}(dN_t = 1 \,|\mathcal{F}_{t^-})} \\ &= \frac{\mathbb{P}(dN_t(i) = 1 \,|\mathcal{F}_{t^-})}{\mathbb{P}(dN_t = 1 \,|\mathcal{F}_{t^-})} \\ &= \frac{\lambda_t(i)}{\lambda_t}.\end{aligned}$$

2.2.7 Random Time Change

In the same way that any continuous local martingale can be represented as a (continuous) time-changed Brownian motion, there is a similar property for point processes, which can be re-casted as time-changed Poisson processes.

The basic result is given in the next theorem.

Theorem 20 (Time-Changed Poisson Process) *Let N_t be a point process with the $\mathcal{F}_t$-intensity λ_t and the $\mathcal{G}_t$ -intensity $\widetilde{\lambda}_t$, where $\mathcal{F}_t$ and $\mathcal{G}_t$ are filtrations of N_t such that*

$$\mathcal{F}_t^N \subset \mathcal{G}_t \subset \mathcal{F}_t.$$

Suppose that $N_\infty = \infty$, a.s. Define for each t, the $\mathcal{G}_t$-stopping time $\theta(t)$ as

$$\int_0^{\theta(t)} \widetilde{\lambda}_s ds = t.$$

Then, the point process $\widetilde{N}_t$ defined by the time change $\theta(t)$,

$$\widetilde{N}_t = N_{\theta(t)},$$

is a standard Poisson process (i.e., with $\mathcal{G}_t$-intensity 1)

Having defined, mathematically, the intensity for a general point process, on a given filtration–and appreciated the subtleties around the choice of filtration-, to proceed further, we need to provide an "operational" tool to construct these quantities and relate them to each other in some way. This is achieved by using the concept of copula functions, which we describe next.

2.3 Copulas

Default correlation has been, for a long time, a very ambiguous concept – shrouded in mystery and often misunderstood, or at least misinterpreted in one way or another. This state of fuzziness is probably due to the fact that our minds are trained to think in terms of Gaussian distributions. A multivariate Gaussian distribution is completely determined by its pair-wise correlations (and variances or its covariance matrix). This fact is very specific to normal distributions. When we are talking about "correlating default events" we are trying to specify the multivariate distribution of a set of Bernoulli variables which cannot be achieved by looking solely at the pairwise correlations; a more general tool is needed.

Consider two random variables (T_1, T_2), with a bivariate distribution $F(t_1, t_2)$, and marginals $F_1(t_1)$, $F_2(t_2)$; then, we have the following properties:

$$\begin{aligned} F(t_1, +\infty) &= F_1(t_1), \; F(+\infty, t_2) = F_2(t_2); \\ F(t_1, -\infty) &= F(-\infty, t_2) = F(-\infty, -\infty) = 0; \\ F(+\infty, +\infty) &= 1. \end{aligned}$$

Furthermore, the measure of the rectangle $[x_1, x_2] \times [y_1, y_2]$ is positive and given by:

$$\mathbb{P}(x_1 \leq T_1 \leq x_2, \; y_1 \leq T_2 \leq y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0.$$

When T_1 and T_2 are independent, the bivariate distribution is simply the product of the (univariate) marginals:

$$F(t_1, t_2) = F_1(t_1) \times F_2(t_2).$$

The problem of determining a bivariate distribution from its marginals has an infinite number of solutions. In particular, we have an upper and lower bound that are solutions to this problem. Fréchet (1957) has shown that the following condition holds:

$$\max(F_1(t_1) + F_2(t_2) - 1, 0) \leq F(t_1, t_2) \leq \min(F_1(t_1), F_2(t_2)).$$

This family of solutions can be parametrized elegantly by using the formalism of Copula functions. We present here a short summary of the main results regarding copulas (we refer to Embrechts et al. (2003) for more details).

2.3.1 Sklar's Theorem

Basically, a copula function is a function that links a set of univariate marginal distributions to a complete multivariate distribution.

The formal mathematical definition is given below (see Nelsen 1999).

Definition 21 (*Copula*) An n-dimensional copula is any function $C : [0, 1]^n \rightarrow [0, 1]$ with the following properties

- C is grounded, i.e., $C(u_1, \ldots, u_n) = 0$ for all $(u_1, \ldots, u_n) \in [0, 1]^n$ such that $u_k = 0$ for at least one k;
- C is n-increasing, i.e., the C-volume of all n-boxes whose vertices lie in $[0, 1]^n$ is positive:

$$\sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1+\cdots+i_n} C\left(u_1^{i_1}, \ldots, u_n^{i_n}\right) \geq 0,$$

  for all $\left(u_1^1, \ldots, u_n^1\right)$ and $\left(u_1^2, \ldots, u_n^2\right)$ in $[0, 1]^n$ with $u_k^1 \leq u_k^2, 1 \leq k \leq n$;
- C has margins C_k, which satisfy $C_k(u_k) = C(1, \ldots, 1, u_k, 1 \ldots, 1) = u_k$ for all u_k in $[0, 1]$.

This definition ensures that C is a multivariate uniform distribution.

For our purposes, we shall use the following (equivalent) operational definition.

Definition 22 (*Copula Function*) Let $U_1, U_2, \ldots, U_n$ be a set of n uniform random variables. Then, the joint distribution function

$$C(u_1, u_2, \ldots, u_n) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_n \leq u_n)$$

is called a Copula function.

The link between copulas and the construction of multivariate distributions is given by Sklar's theorem.

For a set of n random variables $X_1, X_2, \ldots, X_n$ with univariate distributions $F_i(x_i) = \mathbb{P}(X_i \leq x_i)$, we can define their multivariate distribution

$F(x_1, x_2, \ldots, x_n)$ by a choice of Copula function as follows:

$$F(x_1, x_2, \ldots, x_n) \triangleq C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)).$$

To see that, it suffices to write:

$$\begin{aligned} C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) &= \mathbb{P}(U_1 \leq F_1(x_1), U_2 \leq F_2(x_2), \ldots, U_n \leq F_n(x_n)) \\ &= \mathbb{P}\left(F_1^{-1}(U_1) \leq x_1, F_2^{-1}(U_2) \leq x_2, \ldots, F_n^{-1}(U_n) \leq x_n\right) \\ &= \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \\ &= F(x_1, x_2, \ldots, x_n). \end{aligned}$$

The converse result is also true. Sklar (1959) has shown that any multivariate distribution can be expressed as a Copula function.

Theorem 23 (Sklar's theorem) *Let F be an n-dimensional distribution function with margins $F_1, \ldots, F_n$. Then, there exists an n-copula C such that for all $(x_1, \ldots, x_n) \in \overline{\mathbb{R}}^n$,*

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

If $F_1, \ldots, F_n$ are all continuous, then C is unique. Otherwise, C is uniquely determined on $Ran\ F_1 \times \cdots \times Ran\ F_n$. Conversely, if C is an n-copula and $F_1, \ldots, F_n$ are distribution functions, then the function F defined above is an n-dimensional distribution function with margins $F_1, \ldots, F_n$.

The mixed kth-order partial derivatives of a copula function C, $\frac{\partial^k C(\mathbf{u})}{\partial u_1 \ldots \partial u_k}$, exist for almost all $\mathbf{u}$ in $[0, 1]^n$; moreover, the partial derivatives are always bounded between 0 and 1,

$$0 \leq \frac{\partial^k C(\mathbf{u})}{\partial u_1 \ldots \partial u_k} \leq 1.$$

Now, every copula function C can be decomposed into its absolutely continuous part and its singular part:

$$C(u_1, \ldots, u_n) = A_C(u_1, \ldots, u_n) + S_C(u_1, \ldots, u_n),$$

where

$$\begin{aligned} A_C(u_1, \ldots, u_n) &= \int_0^{u_1} \ldots \int_0^{u_n} \frac{\partial^n C(s_1, \ldots, s_n)}{\partial s_1 \ldots \partial s_n} ds_1 \ldots ds_n, \\ S_C(u_1, \ldots, u_n) &= C(u_1, \ldots, u_n) - A_C(u_1, \ldots, u_n). \end{aligned}$$

If $C = A_C$ on $[0,1]^n$, then C is said to be absolutely continuous; in this case, it will have a density distribution $\frac{\partial^n C(u_1,\ldots,u_n)}{\partial u_1 \ldots \partial u_n}$.

If $C = S_C$ on $[0,1]^n$, then C is said to be singular, and will have zero density, $\frac{\partial^n C(u_1,\ldots,u_n)}{\partial u_1 \ldots \partial u_n} = 0$, almost everywhere in $[0,1]^n$.

The Marshall-Olkin copula, which we will study later, is an important example of a copula function with a regular continuous part and a singular part on the diagonals.

Frechet-Hoeffding Bounds. Define the functions M^n, Π^n and W^n, on $[0,1]^n$, as

$$\begin{aligned} M^n(\mathbf{u}) &= \min(u_1, \ldots, u_n), \\ \Pi^n(\mathbf{u}) &= u_1 \ldots u_n, \\ W^n(\mathbf{u}) &= \max(u_1 + \cdots + u_n - n + 1, 0). \end{aligned}$$

Note that the functions M^n and Π^n are n-copula functions, for all $n \geq 2$; the function W^n, on the other hand, is not a copula function for any $n \geq 3$.

The upper and lower bounds for a copula function are given by the Frechet-Hoeffding bounds inequality (Fréchet 1957).

Theorem 24 (Frechet-Hoeffding Bounds) *Let C be an n-copula function, then for every $\boldsymbol{u}$ in $[0,1]^n$, we have*

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u}).$$

The whole question now is: what is the best choice of copula function for our purposes.

The traditional copula used in the market explicitly or implicitly is the Gaussian Copula. It has the advantage of being easy to simulate and its correlation parameters happen to have a nice interpretation in the Firm Asset Value approach. That is effectively what is used, for example, in the CreditMetrics model (see Gupton et al. 1997). Other approaches are also possible: we can assume alternatively that individual obligor defaults are driven by the so-called "Shock models", where common market factors trigger the joint defaults of multiple credits simultaneously. The copula function that originates from this model is known as the Marshall-Olkin copula function.

2.3.2 Dependence Concepts

As we have seen previously, linear correlation (also known as Pearson's correlation) is not sufficient to quantify the dependence between two random

variables. The tool that one should use to specify a well-posed dependence structure is the copula function, and as such, any metric, which better captures the random variables distributional dependence features, ought to be based on the copula function itself. Some of the most popular copula-based (dependence) metrics include: Kendall's Tau, Spearman's rho and the (upper) tail dependence coefficient. We review each one in turn.

Linear Correlation. The linear correlation (or Pearson's correlation) coefficient is defined as follows.

Definition 25 (*Linear Correlation*) Let X and Y be two random variables with non-zero finite variances, then the linear correlation coefficient $\rho(X, Y)$ is

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}},$$

where $\text{Cov}(X, Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X]\mathbb{E}[Y]$ is the covariance of X and Y; $\text{Var}(X)$ and $\text{Var}(Y)$ are the variances of X and Y.

Pearson's correlation is a measure of linear dependence. In particular, if we have perfect linear dependence between the random variables X and Y, i.e., when $Y = aX + b$, for some fixed coefficients $a \neq 0$ and b, then the linear correlation is exactly equal to one: $|\rho(X, Y)| = 1$; and the converse result is also true.

Linear correlation is a natural dependence measure for elliptical distributions (such as the multivariate normal or the multivariate t-distribution). For all other distributions, the linear correlation coefficient can be very misleading. Even for elliptical distributions, it only really makes sense for the Gaussian distribution; for t-distributions, where we have heavier tails, the behaviour in the tail is parametrized differently and cannot be captured through the simplistic linear correlation coefficient.

Concordance. Let (X, Y) and $\left(\widetilde{X}, \widetilde{Y}\right)$ be two pairs of random variables with identical marginal distributions.

The probability of concordance between (X, Y) and $\left(\widetilde{X}, \widetilde{Y}\right)$ is given by

$$\mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) > 0\right);$$

similarly, the probability of discordance is

$$\mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) < 0\right).$$

The difference between the probabilities of concordance and discordance can be expressed in terms of the copula functions (see Nelsen 1999).

Theorem 26 (Difference of Concordance and Discordance Probabilities) *Let* (X, Y) *and* $\left(\widetilde{X}, \widetilde{Y}\right)$ *be two pairs of continuous random variables, with common marginal distributions* (F, G)*, and bivariate joint distributions* H *and* $\widetilde{H}$ *respectively. Their bivariate copula functions are* C *and* $\widetilde{C}$ *respectively:* $H(x, y) = C(F(x), G(y))$, $\widetilde{H}(x, y) = \widetilde{C}(F(x), G(y))$. *Let* Q *denote the difference between the probabilities of concordance and discordance*

$$Q = \mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) > 0\right) - \mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) < 0\right),$$

then, we have

$$Q = Q\left(C, \widetilde{C}\right) = 4\int\int_{[0,1]^2} \widetilde{C}(u, v)\, dC(u, v) - 1.$$

This result will be used next to derive the expressions of the Kendall tau and Spearman rho coefficients.

Kendall's Tau and Spearman's Rho. The most important copula-based concordance measures, to be used with distributions other than the normal one, are: Kendall's tau and Spearman's rho.

Definition 27 (*Kendall's Tau*) For two random variables X and Y, Kendall's tau is defined as

$$\tau(X, Y) = \mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) > 0\right) - \mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - \widetilde{Y}\right) < 0\right),$$

where $\left(\widetilde{X}, \widetilde{Y}\right)$ is an independent copy of the pair (X, Y).

This can be written in terms of the copula function.

Theorem 28 *Let X and Y be two continuous random variables with copula C, then their Kendall's tau coefficient is given by*

$$\tau(X, Y) = Q(C, C) = 4\int\int_{[0,1]^2} C(u, v)\, dC(u, v) - 1.$$

Within the same class of concordance measures, we have the Spearman's rho coefficient, which is defined as follows.

Definition 29 (*Spearman's Rho*) For two random variables X and Y, Spearman's rho is defined as

$$\rho_S(X, Y) = 3\mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - Y'\right) > 0\right) - \mathbb{P}\left(\left(X - \widetilde{X}\right)\left(Y - Y'\right) < 0\right),$$

where $\left(\widetilde{X}, \widetilde{Y}\right)$ and $\left(X', Y'\right)$ are independent copies of the pair (X, Y).

Expressed in terms of the copula function, we have the following result.

Theorem 30 *Let X and Y be two continuous random variables with copula C, then their Spearman's rho coefficient is given by*

$$\rho_S(X, Y) = Q(C, \Pi) = 12 \int \int_{[0,1]^2} uvdC(u, v) - 3 = 12 \int \int_{[0,1]^2} C(u, v)\, dudv - 3.$$

If X and Y have marginal distributions F and G, we can use their uniform variates $U = F(X)$ and $V = G(Y)$, and can re-write the Spearman rho measure as

$$\begin{aligned}\rho_S(X, Y) &= 12 \int \int_{[0,1]^2} uvdC(u, v) - 3 = 12\mathbb{E}[UV] - 3 \\ &= \frac{\mathbb{E}[UV] - \frac{1}{4}}{\frac{1}{12}} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} \\ &= \rho\left(F(X), G(Y)\right).\end{aligned}$$

Tail Dependence. Another important concept, especially for heavy-tailed distributions, is the tail dependence coefficient: it quantifies the amount of joint dependence in the tail of the distribution.

Definition 31 (*Tail Dependence*) Let X and Y be two continuous random variables with marginal distributions F and G. The upper tail dependence coefficient is defined as

$$\lim_{u \uparrow 1} \mathbb{P}\left(Y > G^{-1}(u) \,\middle|\, X > F^{-1}(u)\right) = \lambda_U,$$

if the limit $\lambda_U \in [0, 1]$ exists.

If $\lambda_U > 0$, we say that X and Y are asymptotically dependent in the upper tail.

If $\lambda_U = 0$, we say that X and Y are asymptotically independent in the upper tail.

We can re-write the upper tail conditional probability $\mathbb{P}\left(Y > G^{-1}(u) \,\middle|\, X > F^{-1}(u)\right)$ as

$$\frac{1 - \mathbb{P}\left(X \leq F^{-1}(u)\right) - \mathbb{P}\left(Y \leq G^{-1}(u)\right) + \mathbb{P}\left(X \leq F^{-1}(u), Y \leq G^{-1}(u)\right)}{1 - \mathbb{P}\left(X \leq F^{-1}(u)\right)},$$

which gives an alternative definition in terms of the copula function.

Definition 32 (*Copula Tail Dependence*) The upper tail dependence λ_U for a bivariate copula function C is defined as

$$\lim_{u\uparrow 1}\frac{1-2u+C(u,u)}{1-u}=\lambda_U,$$

if the limit λ_U exists.

Similarly, we can also define a lower tail dependence measure in a symmetric way

$$\lim_{u\downarrow 0}\frac{C(u,u)}{u}=\lambda_L.$$

2.3.3 Elliptical Copulas

Elliptical distributions are typically the most commonly used multivariate distribution functions. They enjoy many of the multivariate normal distribution tractability features; and they can also be simulated very easily.

Elliptical copulas are the copula functions generated from elliptical distributions.

Definition 33 (*Elliptical Distributions*) Let $\mathbf{X}$ be an n-dimensional random variable; Fix a real vector $\mu\in\mathbb{R}^n$, a positive definite, symmetric matrix $n\times n$ matrix Σ, and a real function $\phi(.)$. We say that $\mathbf{X}$ has an elliptical distribution, $\mathbf{X}\sim E_n(\mu,\Sigma,\phi)$, with parameters μ, Σ, ϕ if the characteristic function of the vector $\mathbf{X}-\mu$ is a function of the quadratic form $\mathbf{t}^T\Sigma\mathbf{t}$:

$$\varphi_{X-\mu}(\mathbf{t})=\mathbb{E}\left[\exp\left(\mathbf{t}^T(\mathbf{X}-\mu)\right)\right]=\phi\left(\mathbf{t}^T\Sigma\mathbf{t}\right).$$

The most important elliptical copulas are the Gaussian copula and the t-copula.

Gaussian Copula. The copula of the n-variate normal distribution with Gaussian correlation matrix R is given by

$$C_R^G(\mathbf{u})=\Phi_R^n\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_n)\right),$$

where Φ_R^n denotes the n-dimensional multivariate standard normal distribution with correlation matrix R; Φ^{-1} denotes the inverse of the univariate standard normal distribution. In the bivariate case, we can write the Gaussian

copula function as

$$C_R^G(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-R_{12}^2}} \exp\left(-\frac{x^2 - 2R_{12}xy + y^2}{2\left(1-R_{12}^2\right)}\right) dxdy.$$

Student t-copula. If the vector $\mathbf{X}$ can be represented as

$$\mathbf{X} \stackrel{d}{=} \mu + \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z},$$

where $\mu \in \mathbb{R}^n$, and the random variables $S \sim \chi_\nu^2$ and $Z \sim \mathcal{N}_n(\mathbf{0}, \Sigma)$ are independent, then $\mathbf{X}$ has an n-dimensional multivariate t_ν-distribution, with mean μ (for $\nu > 1$) and covariance matrix $\frac{\nu}{\nu-2}\Sigma$ (for $\nu > 2$).

The (Student) t-copula is then defined as

$$C_{\nu,R}^t(\mathbf{u}) = t_{\nu,R}^n\left(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_n)\right),$$

where $R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$, is $t_{\nu,R}^n$ is the n-dimensional multivariate t-distribution with parameters (ν, R), and is the inverse of the univariate t_ν-distribution with ν degrees of freedom. In the bivariate case, the expression of the copula function is

$$C_{\nu,R}^t(u_1, u_2) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-R_{12}^2}} \left(1 + \frac{x^2 - 2R_{12}xy + y^2}{\nu\left(1-R_{12}^2\right)}\right)^{-\frac{\nu+2}{2}} dxdy.$$

2.3.4 Archimedean Copulas

An interesting class of copula functions used in finance insurance applications to model asymmetric losses and gains is the Archimedean copula. Unlike their elliptical counterparts, which are derived from a given family of multivariate distributions, Archimedean copulas are instead constructed, by hand, from a set of generator functions, thereby creating a rich variety of dependence structures. In addition to the asymmetric properties offered by these copulas, they also enjoy easy-to-implement closed-form expressions, which is very useful in numerical applications.

We start with the definition of a pseudo-inverse function needed to construct the Archimedean copula.

Definition 34 (*Pseudo-inverse*) Let $\varphi : [0,1] \to [0,\infty]$ be a continuous, strictly decreasing function, with $\varphi(1) = 0$. The pseudo-inverse of φ is the function $\varphi^{[-1]} : [0,\infty] \to [0,1]$ defined as

$$\varphi^{[-1]}(x) = \begin{cases} \varphi^{-1}(x), \text{ for } 0 \leq x \leq \varphi(0), \\ 0, \text{ if } \varphi(0) < x \leq \infty. \end{cases}$$

We can now give a general definition of bivariate Archimedean copulas (see Nelsen 1999).

Definition 35 (*Bivariate Archimedean Copula*) Let $\varphi : [0,1] \to [0,\infty]$ be a continuous, strictly decreasing function, with $\varphi(1) = 0$, and let $\varphi^{[-1]}$ be its pseudo-inverse. Define the bivariate function $C : [0,1]^2 \to [0,1]$ by

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)).$$

Then, the function C is a copula if and only if is φ convex. Copulas of this form are called Archimedean copulas; and φ is the generator of the copula. If $\varphi(0) = \infty$, we say that φ is a strict generator and C is a strict Archimedean copula.

We give a few popular examples.

Example 36 (Gumbel Copula). Let $\varphi(x) = (-\ln x)^{\theta}$, where $\theta \geq 1$. Its copula function $C_\theta(u,v)$ is called a Gumbel copula:

$$C_\theta(u,v) = \exp\left(-\left((-\ln u)^{\theta} + (-\ln v)^{\theta}\right)^{\frac{1}{\theta}}\right).$$

Example 37 (Clayton Copula). Let $\varphi(x) = \frac{x^{-\theta}-1}{\theta}$, where $\theta \in [-1,\infty)$, $\theta \neq 0$. Its copula function $C_\theta(u,v)$ is called a Clayton copula:

$$C_\theta(u,v) = \max\left(\left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}, 0\right).$$

Example 38 (Frank Copula). Let $\varphi(x) = -\ln \frac{e^{-\theta x}-1}{e^{-\theta}-1}$, where $\theta \in \mathbb{R}$, $\theta \neq 0$. Its copula function $C_\theta(u,v)$ is called a Frank copula:

$$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right).$$

To generalize to an n-dimensional Archimedean copula, we can construct, by extension, the function C^n as

$$C^n(\mathbf{u}) = \varphi^{[-1]}(\varphi(u_1) + \cdots + \varphi(u_n)),$$

but we also need to show that it is indeed a copula function under some conditions.

First, we define what it is meant by a completely monotone function: we say that a function $g(x)$ is completely monotone on the interval I, if it has derivatives of all orders, which alternate in sign, i.e., it satisfies

$$(-1)^k \frac{d^k}{dx^k} g(x) \geq 0,$$

for all $k \geq 0$, and all x in the interior of the interval I.

We can now state the following theorem from Kimberling (1974), which gives necessary and sufficient conditions for the function C^n to be a copula.

Theorem 39 (Kimberling 1974) *Let $\varphi : [0,1] \to [0,\infty]$ be a continuous, strictly decreasing function, with $\varphi(1) = \infty$ and $\varphi(1) = 0$, and let $\varphi^{[-1]}$ be its pseudo-inverse. The n-dimensional function $C^n : [0,1]^n \to [0,1]$ defined by*

$$C^n(\mathbf{u}) = \varphi^{[-1]}(\varphi(u_1) + \cdots + \varphi(u_n)),$$

is a n-copula, for all $n \geq 2$, if and only if φ^{-1} is completely monotone on $[0,\infty)$.

Some n-dimensional examples are given below.

Example 40 (Gumbel Copula). Let $\varphi_i(t) = (-\ln t)^{\theta_i}$, with $\theta_i \geq 1$, for $1 \leq i \leq n$, be the generators of the Gumbel copula. The n-dimensional extension of the Gumbel family of copula functions is an n-copula if $\theta_1 \leq \cdots \leq \theta_n$.

Example 41 The Archimedean copula family defined with generators $\varphi_i(t) = \left(t^{-1} - 1\right)^{\theta_i}$, for $\theta_i \geq 1$, is indeed an n-copula if $\theta_1 \leq \cdots \leq \theta_n$.

2.3.5 Marshall-Olkin Copulas

Next, we discuss the general class of Marshall-Olkin copula functions. We start with the bivariate case, then generalize to the n-dimensional case.

Suppose we have a system with two components subject to some independent shocks, which can trigger the failure of one of the components separately

or both components at the same time. Thus, we have three independent Poisson processes with intensities $(\lambda_1, \lambda_2, \lambda_{12})$ respectively. Their corresponding first jump times are $(\theta_1, \theta_2, \theta_{12})$: they are independent exponentially-distributed variables with parameters $(\lambda_1, \lambda_2, \lambda_{12})$ respectively. We denote by τ_1 and τ_2 the failure time of the two components.

The joint survival probability function of the two components, $\overline{H}(T_1, T_2)$, is given by

$$\begin{aligned}\overline{H}(T_1, T_2) &= \mathbb{P}(\tau_1 > T_1, \tau_2 > T_2) = \mathbb{P}(\theta_1 > T_1)\,\mathbb{P}(\theta_2 > T_2)\,\mathbb{P}(\theta_{12} > \max(T_1, T_2)) \\ &= \exp(-\lambda_1 T_1)\exp(-\lambda_2 T_2)\exp(-\lambda_{12}\max(T_1, T_2)).\end{aligned}$$

Similarly, the univariate survival probabilities, $\overline{F_1}(T_1)$ and $\overline{F_2}(T_2)$, are

$$\begin{aligned}\overline{F_1}(T_1) &= \mathbb{P}(\tau_1 > T_1) = \mathbb{P}(\theta_1 > T_1)\,\mathbb{P}(\theta_{12} > \max(T_1, T_2)) = \exp(-\lambda_1 T_1 - \lambda_{12}\max(T_1, T_2)), \\ \overline{F_2}(T_2) &= \mathbb{P}(\tau_2 > T_2) = \mathbb{P}(\theta_2 > T_2)\,\mathbb{P}(\theta_{12} > \max(T_1, T_2)) = \exp(-\lambda_2 T_2 - \lambda_{12}\max(T_1, T_2)).\end{aligned}$$

Define the ratios $\alpha_1 = \frac{\lambda_{12}}{\lambda_1+\lambda_{12}}$ and $\alpha_2 = \frac{\lambda_{12}}{\lambda_2+\lambda_{12}}$, and substitute the univariate (survival) marginals into the bivariate (survival) distribution function

$$\overline{H}(T_1, T_2) = \overline{F_1}(T_1)\,\overline{F_2}(T_2)\min\left(\left(\overline{F_1}(T_1)\right)^{-\alpha_1}, \left(\overline{F_2}(T_2)\right)^{-\alpha_2}\right).$$

This leads to the family of Marshall-Olkin copula functions

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min\left(u_1^{1-\alpha_1}u_2, u_1 u_2^{1-\alpha_2}\right).$$

This copula function has both an absolutely continuous part

$$A_{C_{\alpha_1,\alpha_2}} = \frac{\partial^2}{\partial u_1 \partial u_2} C_{\alpha_1,\alpha_2}(u_1, u_2) = \begin{cases} u_1^{-\alpha_1}, \text{ if } u_1^{\alpha_1} > u_2^{\alpha_2}, \\ u_2^{-\alpha_2}, \text{ if } u_1^{\alpha_1} > u_2^{\alpha_2}; \end{cases}$$

and a singularity on the region defined by the curve $\left\{u_1^{\alpha_1} = u_2^{\alpha_2}\right\}$, where we have a simultaneous failure of both components at some time θ_{12}.

Kendall's tau, Spearman's rho and upper tail dependence, for this copula function, can be computed easily and are given by

$$\begin{aligned}\tau\left(C_{\alpha_1,\alpha_2}\right) &= \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2}, \\ \rho_S\left(C_{\alpha_1,\alpha_2}\right) &= \frac{3\alpha_1\alpha_2}{2\alpha_1 + 2\alpha_2 - \alpha_1\alpha_2}, \\ \lambda_U &= \min(\alpha_1, \alpha_2).\end{aligned}$$

For the n-dimensional generalization, we have n components, and $2^n - 1$ common shocks, which can trigger the failure of one or more components in the system. We denote by Π_n the set of all non-empty subset of $\{1, \ldots, n\}$. We have $2^n - 1$ independent Poisson shocks N^π, $\pi \in \Pi_n$, with intensities λ_π, which can trigger the failure of the components in the subset π only. Their first jump times are θ_π. The failure time of each individual component is then given by: $\tau_i = \min\{\theta_\pi : i \in \pi\}$.

The (τ_i, τ_j)-bivariate marginal of the Marshall-Olkin copula is also a Marshall-Olkin copula with parameters

$$\alpha_i = \frac{\sum_{\pi: i\in\pi,\, j\in\pi} \lambda_\pi}{\sum_{\pi: i\in\pi} \lambda_\pi} \text{ and } \alpha_j = \frac{\sum_{\pi: i\in\pi,\, j\in\pi} \lambda_\pi}{\sum_{\pi: j\in\pi} \lambda_\pi};$$

the Kendall tau and Spearman rho coefficients are given by

$$\tau\left(C_{\alpha_i,\alpha_j}\right) = \frac{\alpha_i\alpha_j}{\alpha_i + \alpha_j - \alpha_i\alpha_j} \text{ and } \rho_S\left(C_{\alpha_i,\alpha_j}\right) = \frac{3\alpha_i\alpha_j}{2\alpha_i + 2\alpha_j - \alpha_i\alpha_j}.$$

The n-dimensional Marshall-Olkin copula function provides a very rich joint dependence structure, with enough flexibility to capture the granular joint probabilities of every combination of sub-defaults; but with 2^n combinations to deal with, the problem explodes rapidly as the number of names grows. We shall see later in Chap. 7 that we can build a parsimonious parametrization of the model, which gives the desired default clustering properties that we need while maintaining numerical tractability.

References

R. Barlow, F. Proschan, *Statistical Theory of Reliability and Life Testing* (Silver Spring, Maryland, 1981)

P. Brémaud, *Point Processes and Queues: Martingale Dynamics* (Springer, New York, 1980)

P. Embrechts, F. Lindskog, A. McNeil, Modelling dependence with copulas and applications to risk management, in *Handbook of Heavy Tailed Distributions in Finance*, ed. by S.T. Rachev (Amsterdam, Elsevier/North-Holland, 2003)

M. Fréchet, Les tableaux de corrélation dont les marges sont données. Annales de l'Université de Lyon, Sciences Math ématiques et Astronomie **20**, 13–31 (1957)

G.M. Gupton, C.C. Finger, M. Bhatia, *CreditMetrics-Technical Document* (Morgan Guaranty Trust Co, New York, 1997)

H. Joe, *Multivariate Models and Dependence Concepts* (Chapman & Hall, London, 1997)

N. Johnson, S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions* (Wiley, New York, 1972)
C. Kimberling, A Probabilistic Interpretation of Complete Monotonicity. Aequationes Mathematicae **10**, 152–164 (1974)
D. Lando, On cox processes and credit risky securities. Rev. Deriv. Res. **2**(2/3), 99–120 (1998)
F. Lindskog, A. McNeil, Common poisson shock models: applications to insurance and credit risk modelling. ASTIN Bullet. **33**(2), 209–238 (2003)
A. Marshall, I. Olkin, A multivariate exponential distribution. J. Am. Stat. Assoc. **62**, 30–44 (1967)
R. Nelsen, *An Introduction to Copulas* (Springer, New York, 1999)
A. Sklar, Fonctions de répartition à *n* dimensions et leurs marges. Publications de l'Institut de Statistique de l'Universit é de Paris **8**, 229–231 (1959)
S. Watanabe, Additive functionals of Markov processes and Levy systems. Jpn. J. Math. **34**, 53–79 (1964)

Springer

http://www.springer.com/978-3-319-60972-0

Credit Correlation
Theory and Practice
Elouerkhaoui, Y.
2017, XXIV, 456 p. 67 illus., Hardcover
ISBN: 978-3-319-60972-0