FORMULA SHEET FOR FINANCIAL RISK
ALLOWED TO BE USED ON THE EXAM

1. Extreme value statistics

Generalized Pareto cumulative distribution function:

$$H(x) = \begin{cases} 1-(1+\frac{\gamma}{\sigma}x)^{-1/\gamma} & \text{for} \quad x \geq 0, & \text{if } \gamma_j > 0 \\ e^{-\frac{x}{\sigma}} & \text{for} \quad x \geq 0, & \text{if } \gamma_j = 0 \\ 1-(1+\frac{\gamma}{\sigma}x)^{-1/\gamma} & \text{for} \quad x \geq 0 \text{ and } x < -\frac{\sigma}{\gamma}, & \text{if } \gamma_j < 0 \end{cases}$$

Generalized Extreme Value cumulative distribution function:

$$G(x) = \begin{cases} \exp\{-(1+\frac{\gamma}{\sigma}(x-\mu))^{-1/\gamma}\} & \text{for} \quad x \geq \mu - \frac{\sigma}{\gamma}, & \text{if } \gamma_j > 0 \\ e^{-e^{-\frac{x-\mu}{\sigma}}} & & \text{if } \gamma_j = 0 \\ \exp\{-(1+\frac{\gamma}{\sigma}(x-\mu))^{-1/\gamma}\} & \text{for} \quad x < \mu - \frac{\sigma}{\gamma}, & \text{if } \gamma_j < 0 \end{cases}$$

Poisson process:

A counting process $N(t)$ is a Poisson process if

- The numbers of events which occur in disjoint time intervals are mutually independent
- $N(t+s) - N(s)$ has a Poisson distribution for any $s, t \geq 0$, i.e.

$$\mathbb{P}\left[N(s+t) - N(s) = k\right] = \frac{\lambda^k}{k!}e^{-\lambda t}, \quad \text{for any} \quad s, t \geq 0 \text{ and } k = 0, 1, 2, \ldots$$

Here λ is the "intensity parameter". One interpretation is that λ is the expected number of events in any interval of length 1.

ML inference:

With $l(\theta)$ denoting the log likelihood function, the expected and observed information matrices are

$$I(\theta) = E_\theta(-\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j}l(\theta)) \quad \text{and} \quad \mathrm{I}(\theta) = (-\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j}l(\theta)),$$

respectively. $I(\theta)$ can be estimated by $\mathrm{I}(\hat{\theta})$ where $\hat{\theta}$ are the ML estimates of the parameters θ. The ML estimate $\hat{\theta} = \hat{\theta_1}, \ldots \hat{\theta_d}$ asymptotically has a mean zero multivariate normal distribution with covariance matrix $I(\theta)^{-1}$.

Partition the parameter vector θ into two parts, $\theta = (\theta_1, \theta_2)$ and write θ_2^* for the value of of θ_2 which maximises $l(\theta) = l(\theta_1, \theta_2)$ over θ_2 for θ_1. A Likelihood Ratio (LR) test then rejects the null hypothesis that θ_1 takes the value θ_1^0 at the significance level α if

$$2(l(\hat{\theta}) - l(\theta_1^0, \hat{\theta}_2)) > \chi^2_\alpha(d-p),$$

where $\chi^2_\alpha(d-p)$ is the $1-\alpha$ quantile of the χ^2-distribution with $d-p$ degrees of freedom, where p and d are the dimensions (=lengths) of the vectors θ and θ_2, respectively.

Dependence and the extremal index:

The extremal index, θ is obtained as $1/\{$asymptotic mean cluster length$\}$. If $L_1, L_2, \ldots$ is a stationary stochastic process with marginal cumulative distribution function $F(x)$ and extremal index θ and $M_n = \max\{L_1, L_2, \ldots, L_n$ then asymptotically

$$\mathbb{P}\left[M_n \leq x\right] = F(x)^{\theta n}.$$

2. Value-at-Risk and expected shortfall

Definition of Value-at-Risk:

Given a loss L and a confidence level $\alpha \in (0,1)$, the $100 \times \alpha\%$ Value-at-Risk, denoted $\mathrm{VaR}_\alpha(L)$ is the α-quantile of the distribution function $F_L(x) = \mathbb{P}[L \leq x]$, that is

$$\mathrm{VaR}_\alpha(L) = F_L^{\leftarrow}(\alpha) \tag{2.1}$$

where $F_L^{\leftarrow}(x)$ is the generalized inverse of $F_L(x)$. Hence, $\mathrm{VaR}_\alpha(L)$ is given by the smallest number y such that the probability that the loss L exceeds y is no larger than $1-\alpha$, that is

$$\begin{aligned}\mathrm{VaR}_\alpha(L) &= \inf\{y \in \mathbb{R} : \mathbb{P}[L > y] \leq 1-\alpha\}\\ &= \inf\{y \in \mathbb{R} : 1 - \mathbb{P}[L \leq y] \leq 1-\alpha\}\\ &= \inf\{y \in \mathbb{R} : F_L(y) \geq \alpha\}\end{aligned}$$

where $F_L(x) = \mathbb{P}[L \leq x]$ is the distribution of L.

In the case when $F_L(x) = \mathbb{P}[L \leq x]$ is continuous, and thus strictly increasing (i.e. the loss L is a continuous random variable), then $F_L^{\leftarrow}(x)$ will be the inverse function $F_L^{-1}(x)$, and we have

$$\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha) \tag{2.2}$$

which means that $\mathrm{VaR}_\alpha(L)$ is the solution x_α to the equation

$$F_L(x_\alpha) = \alpha.$$

Definition of Expected shortfall: Given a loss L and a confidence level $\alpha \in (0,1)$, the $100\times\alpha\%$ expected shortfall, denoted $\mathrm{ES}_\alpha(L)$ is defined as

$$\mathrm{ES}_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(L)du$$

and if L is a continuous random variable one can show that

$$\mathrm{ES}_\alpha(L) = \mathbb{E}[L \mid L \geq \mathrm{VaR}_\alpha(L)] = \frac{1}{1-\alpha}\int_{\mathrm{VaR}_\alpha(L)}^\infty x f_L(x)dx$$

where $f_L(x)$ is the density of the loss L.

Linearity of Value-at-Risk and Expected shortfall: Let L be a loss and let $a > 0$ and $b \in \mathbb{R}$ be constants. Then

$$\mathrm{VaR}_\alpha(aL+b) = a\mathrm{VaR}_\alpha(L) + b \tag{2.3}$$

and

$$\mathrm{ES}_\alpha(aL+b) = a\mathrm{ES}_\alpha(L) + b. \tag{2.4}$$

The relations (2.3) and (2.4) are often useful in practical computations.

3. The mixed binomial model

Let Z be a random variable on $\mathbb{R}$ and let $p(x) : \mathbb{R} \mapsto [0,1]$ be a function. Define the random variable $p(Z) \in [0,1]$ with mean $\bar{p}$, that is

$$\mathbb{E}[p(Z)] = \bar{p}. \tag{3.1}$$

If Z is a continuous random variable with density $f_Z(z)$ then

$$\mathbb{E}[p(Z)] = \int_{-\infty}^\infty p(z) f_Z(z)dz = \bar{p}. \tag{3.2}$$

Let $X_1, X_2, \ldots X_m$ be identically distributed random variables such that $X_i = 1$ if obligor i defaults before time T and $X_i = 0$ otherwise. Furthermore, *conditional on* Z, the random variables

$X_1, X_2, \dots X_m$ are *independent* and each X_i have default probability $p(Z)$ so $\mathbb{P}[X_i = 1 \mid Z] = p(Z)$. We then get that

$$\mathbb{P}[X_i = 1] = \mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i \mid Z]] = \mathbb{E}[p(Z)] = \bar{p}$$

where the last equality is due to (3.1). Next, letting all losses be the same and constant given by, say ℓ, then the total credit loss in the portfolio at time T, called L_m, is

$$L_m = \sum_{i=1}^{m} \ell X_i = \ell \sum_{i=1}^{m} X_i = \ell N_m \quad \text{where } N_m = \sum_{i=1}^{m} X_i$$

thus, N_m is the *number* of defaults in the portfolio up to time T. Since

$$\mathbb{P}[L_m = k\ell] = \mathbb{P}[N_m = k]$$

it is enough to study N_m. Since the random variables $X_1, X_2, \dots X_m$ are conditionally independent, given the outcome Z, we have

$$\mathbb{P}[N_m = k \mid Z] = \binom{m}{k} p(Z)^k (1 - p(Z))^{m-k}.$$

Hence, we have

$$\mathbb{P}[N_m = k] = \mathbb{E}[\mathbb{P}[N_m = k \mid Z]] = \mathbb{E}\left[\binom{m}{k} p(Z)^k (1 - p(Z))^k\right] \tag{3.3}$$

which holds regardless if Z is a discrete or continuous random variable.

If Z is a continuous random variable on $\mathbb{R}$ with density $f_Z(z)$ then

$$\mathbb{P}[N_m = k] = \int_{-\infty}^{\infty} \binom{m}{k} p(z)^k (1 - p(z))^{m-k} f_Z(z) dz. \tag{3.4}$$

3.1. **Some examples of mixing distributions.** Below we list three examples of mixing distributions frequently used in the industry:

Example 1: A mixed binomial model with $p(Z) = Z$ where Z is a beta distribution, $Z \sim \text{Beta}(a, b)$ and by definition of a beta distribution it holds that $\mathbb{P}[0 \le Z \le 1] = 1$ so that $p(Z) \in [0, 1]$. We say that a random variable Z has beta distribution, $Z \sim \text{Beta}(a, b)$, with parameters a and b, if it's density $f_Z(z)$ is given by

$$f_Z(z) = \frac{1}{\beta(a, b)} z^{a-1} (1 - z)^{b-1} \quad a, b > 0, \quad 0 < z < 1 \tag{3.1.1}$$

where

$$\beta(a, b) = \int_0^1 z^{a-1} (1 - z)^{b-1} dz = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}. \tag{3.1.2}$$

Here $\Gamma(y)$ is the Gamma function defined as

$$\Gamma(y) = \int_0^{\infty} t^{y-1} e^{-t} dt \tag{3.1.3}$$

which satisfies the relation

$$\Gamma(y + 1) = y\Gamma(y) \tag{3.1.4}$$

for any y. By using Equation (3.1.2) and (3.1.4) one can show that $\beta(a, b)$ satisfies the recursive relation

$$\beta(a + 1, b) = \frac{a}{a + b} \beta(a, b).$$

Example 2: Another possibility for mixing distribution $p(Z)$ is to let $p(Z)$ be a logit-normal distribution. This means that

$$p(Z) = \frac{1}{1 + \exp(-(\mu + \sigma Z))}$$

where μ and σ are constants with $\sigma > 0$ and Z is a standard normal. Note that $p(Z) \in [0, 1]$.

Example 3: The mixed binomial model inspired by the Merton model with $p(Z)$ given by

$$p(Z) = N\left(\frac{N^{-1}(\bar{p}) - \sqrt{\rho}Z}{\sqrt{1-\rho}}\right) \tag{3.1.5}$$

where Z is a standard normal and $N(x)$ is the distribution function of a standard normal distribution. Furthermore, $\rho \in (0,1)$ and $\bar{p} = \mathbb{P}[X_i = 1]$. Note that $p(Z) \in [0,1]$.

3.2. **Large Portfolio Approximation (LPA) for mixed binomial models.** The following theorem is very useful when considering the loss distribution for a large credit portfolio, i.e. when m is large.

Theorem 3.1. *With notation as above, for any $x \in [0,1]$ it holds that*

$$\mathbb{P}\left[\frac{N_m}{m} \leq x\right] \to \mathbb{P}[p(Z) \leq x] \quad \textit{when} \quad m \to \infty. \tag{3.2.1}$$

The distribution $\mathbb{P}[p(Z) \leq x]$ is called the Large Portfolio Approximation (LPA) to the distribution of $\frac{N_m}{m}$.

Hence, the above result implies that in a mixed binomial model, the distribution of the fractional number of defaults $\frac{N_m}{m}$ in the portfolio converges to the distribution of the random variable $p(Z)$ as $m \to \infty$. Furthermore, if $p(Z)$ has heavy tails, then the random variable $\frac{N_m}{m}$ will also have heavy tails, as $m \to \infty$, which then implies a strong default dependence in the credit portfolio.

Example: In the mixed binomial model inspired by the Merton model with $p(Z)$ given by (3.1.5), we have

$$\mathbb{P}\left[\frac{N_m}{m} \leq x\right] \to N\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}N^{-1}(x) - N^{-1}(\bar{p})\right)\right) \quad \text{as } m \to \infty \tag{3.2.2}$$

where the right hand side in (3.2.2) thus is the LPA distribution in the mixed binomial Merton model.