Financial Risk: Credit Risk, Lecture 2

Financial Risk, Chalmers University of Technology and University of Gothenburg,
Göteborg
Sweden

Content of lecture

- Short recapitulation of the mixed binomial model
- Discussion of the loss distribution in the mixed binomial model and how to use the LPA theory to find approximation for the loss for large portfolios
- Recapitulation of Value-at-Risk and Expected shortfall and its use in the mixed binomial loss model
- Study of a mixed binomial loss model with a beta distribution
- Study of a mixed binomial loss model with a logit-normal distribution
- Discussion of correlations etc.

Recap of the mixed binomial model

Consider a homogeneous credit portfolio model with m obligors, and where each obligor can default up to fixed time point, say T. Each obligor have identical credit loss at a default, say ℓ. Here ℓ is a constant.

- Let X_i be a random variable such that

$$X_i = \begin{cases} 1 & \text{if obligor } i \text{ defaults before time } T \\ 0 & \text{otherwise, i.e. if obligor } i \text{ survives up to time } T \end{cases} \tag{1}$$

- Let Z be a random variable, discrete or continuous, that represents some common background variable affecting all obligors in the portfolio.
- Since we consider a homogeneous credit portfolio, $X_1, X_2, \ldots X_m$ are identically distributed. Furthermore, we assume the following:

  Conditional on Z, the random variables $X_1, X_2, \ldots X_m$ are independent and each X_i have default probability $p(Z) \in [0, 1]$, that is

$$\mathbb{P}[X_i = 1 \mid Z] = p(Z) \tag{2}$$

  so that $\mathbb{P}[X_i = 1] = \bar{p}$ for each obligor i where $\bar{p}$ is given by

$$\bar{p} = \mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i \mid Z]] = \mathbb{E}[p(Z)] \tag{3}$$

Recap of the mixed binomial model, cont.

- Note that (2) and (3) 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

$$\bar{p} = \mathbb{E}\left[p(Z)\right] = \int_{-\infty}^{\infty} p(z) f_Z(z) dz. \tag{4}$$

- Recall that we want to find the loss distribution in the homogeneous credit portfolio specified on the previous slide.
- 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 $\mathbb{P}\left[L_m = k\ell\right] = \mathbb{P}\left[N_m = k\right]$, it is enough to study N_m.

The mixed binomial model, cont.

- Since $X_1, X_2, \dots X_m$ are conditionally independent given Z, we have

$$\mathbb{P}\left[N_m = k \,|\, Z\right] = \binom{m}{k} p(Z)^k (1 - p(Z))^{m-k}$$

- Hence, we have

$$\mathbb{P}\left[N_m = k\right] = \mathbb{E}\left[\mathbb{P}\left[N_m = k \,|\, Z\right]\right] = \mathbb{E}\left[\binom{m}{k} p(Z)^k (1 - p(Z))^{m-k}\right] \qquad (5)$$

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

$$\mathbb{P}\left[N_m = k\right] = \int_{-\infty}^{\infty} \binom{m}{k} p(z)^k (1 - p(z))^{m-k} f_Z(z) dz. \qquad (6)$$

- We want to find the loss distribution $F_{L_m}(x) = \mathbb{P}\left[L_m \leq x\right]$ for $x \in [0, \infty)$, or in fact for $x \in [0, \ell \cdot m]$ (why ?)

The loss distribution in a mixed binomial model

- Note that for any positive x we have that

$$F_{L_m}(x) = \mathbb{P}[L_m \le x] = \mathbb{P}[\ell N_m \le x] = \mathbb{P}\left[N_m \le \frac{x}{\ell}\right] = \mathbb{P}\left[N_m \le \left\lfloor \frac{x}{\ell} \right\rfloor\right] \quad (7)$$

where $\lfloor y \rfloor$ is the integer part of y rounded downwards, e.g $\lfloor 3.14 \rfloor = 3$.

- For $n = 0, 1 \ldots, m$ then $\mathbb{P}[N_m \le n] = \sum_{k=0}^{n} \mathbb{P}[N_m = k]$ which in (7) yields

$$F_{L_m}(x) = \sum_{k=0}^{\lfloor \frac{x}{\ell} \rfloor} \mathbb{P}[N_m = k] \quad (8)$$

where $\mathbb{P}[N_m = k]$ is computed by (5).

- If Z is a continuous random variable on $\mathbb{R}$ with density $f_Z(z)$ then $\mathbb{P}[N_m = k]$ is computed by (6) and this in (8) renders that

$$F_{L_m}(x) = \sum_{k=0}^{\lfloor \frac{x}{\ell} \rfloor} \int_{-\infty}^{\infty} \binom{m}{k} p(z)^k (1 - p(z))^{m-k} f_Z(z) dz. \quad (9)$$

- Note that $F_{L_m}(x)$ in (8) or (9) will be piece-wise constant (i.e. flat) on each interval $[0, \ell[, [\ell, 2\ell[\ldots [(m-1)\ell, m\ell[, [m\ell, \infty[$ (why ?)

The loss distribution in a mixed binomial model, cont.

- Note the formula for the loss distribution in (8) or (9) is rather tedious and will fail for large values of m (why ?)

- Fortunately, there is a very convenient approximation of the loss distribution $F_{L_m}(x) = \mathbb{P}[L_m \leq x]$ when m is "large"

- Recall that $F(x)$ is the distrib. function of $p(Z)$, i.e $F(x) = \mathbb{P}[p(Z) \leq x]$ and from last lecture we know that for any $x \in [0,1]$ it holds that

$$\mathbb{P}\left[\frac{N_m}{m} \leq x\right] \to F(x) = \mathbb{P}[p(Z) \leq x] \quad \text{as } m \to \infty \tag{10}$$

- We also have that

$$F_{L_m}(x) = \mathbb{P}[L_m \leq x] = \mathbb{P}[\ell N_m \leq x] = \mathbb{P}\left[\frac{N_m}{m} \leq \frac{x}{\ell m}\right]$$

and this in (10) then implies that

$$F_{L_m}(x) \to F\left(\frac{x}{\ell m}\right) \quad \text{as } m \to \infty$$

The loss distribution in a mixed binomial model, cont.

- Hence, if m is "large" we have the following approximation for the loss distribution $F_{L_m}(x) = \mathbb{P}[L_m \leq x]$

$$F_{L_m}(x) \approx F\left(\frac{x}{\ell m}\right) \quad \text{if } m \text{ is "large"}. \tag{11}$$

for any $x \in [0, \ell m]$ and where $F(x) = \mathbb{P}[p(Z) \leq x]$.

- So if m is "large" we can approximate $F_{L_m}(x) = \mathbb{P}[L_m \leq x]$ with $F\left(\frac{x}{\ell m}\right)$ instead of numerically compute the involved expression in the RHS of (9)

- This will be very useful when computing different risk measures for credit portfolios, such as Value-at-Risk and expected shortfall

- Let us define/recap the concept of Value-at-Risk and expected shortfall

Value-at-Risk

- We now define/recap the risk measure Value-at-Risk, abbreviated VaR and the below definition holds for any type of loss L (loss for equity risk, loss for credit risk, loss operational risk etc etc)

Definition of Value-at-Risk

Given a loss L and a confidence level $\alpha \in (0, 1)$, then $\mathrm{VaR}_\alpha(L)$ 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

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

where $F_L(x)$ is the distribution of L.

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

$$\mathrm{VaR}_\alpha(aL + b) = a\mathrm{VaR}_\alpha(L) + b \tag{12}$$

Example of Value-at-Risk when L is continuous r.v.

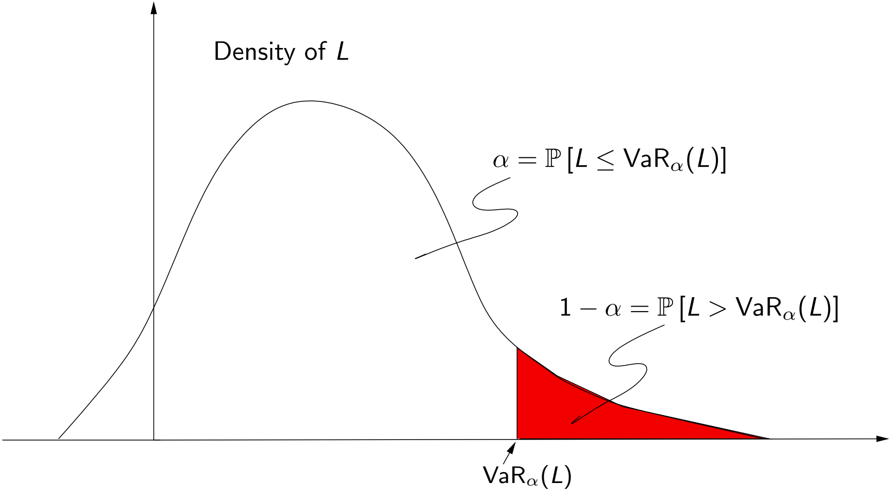


Figure: Visualization of definition of $\mathrm{VaR}_\alpha(L)$ when L is a continuous random variable. The red region has the area $1 - \alpha$

Value-at-Risk, cont.

- Note that Value-at-Risk is defined for a fixed time horizon, so the above definition should also come with a time period, e.g, if the loss L is over one day, then we talk about a one-day $\text{VaR}_\alpha(L)$.
- In market risk, typically the underlying period studied for the loss is 1 day or 10 days.
- In credit risk and in operational risk, one typically consider $\text{VaR}_\alpha(L)$ for the loss over one year.
- Typical values for α are $95\%, 99$ or 99.9%, that is $\alpha = 0.95, \alpha = 0.99$ or $\alpha = 0.999$
- Note that VaR, by definition, does not give any information about *"how bad things can get"*, i.e. the severity of the loss L which may occur with probabilitiy $1 - \alpha$
- We will later shortly discuss the expected shortfall which is a measure that captures the severity of the loss L, given that $L > \text{VaR}_\alpha(L)$.

Value-at-Risk, cont.

- Hence, by definition, $\text{VaR}_{\alpha}(L)$ for a period T have the following interpretation: "We are α % certain that our loss L will not be bigger than $\text{VaR}_{\alpha}(L)$ dollars up to time T"
- However, we should keep in mind that this sentence can be very misleading for several reasons.
- One major reason is that $\text{VaR}_{\alpha}(L)$ is computed under an assumption of how the loss will be distributed, i.e. we use a specific model for L, and this naturally leads to model risk
- One typical example of model risk when computing $\text{VaR}_{\alpha}(L)$ is that $F_L(x) = \mathbb{P}[L \leq x]$ is assumed to have a distribution, which maybe (most likely) not will match the "true" distribution of L, which obviously is difficult to know for sure.

Inverse and generalized inverse functions

- Recall that a function $f(x)$ is strictly monotonic if it is strictly increasing or strictly decreasing
- Recall from your first year calculus course, that a strictly monotonic function $f(x)$ has a unique and well defined inverse $f^{-1}(x)$ such that

$$\begin{array}{lll} 1. & f^{-1}(f(x)) = x\,, & \text{for all } x \text{ in f-s domain} \\ 1. & f(f^{-1}(y)) = y\,, & \text{for all } y \text{ in f-s range} \end{array}$$

- If the function $f(x)$ is monotonic (i.e. not strictly monotonic) then the concept of a inverse function has to be readjusted
- Let us here focus on a nondecreasing function $F(x)$.
- Since $F(x)$ is nondecreasing, it may be "flat" for some regions in its domain (see e.g. example on bottom on slide 6)
- This means that in these "flat" regions we can no longer find a unique inverse function to $F(x)$, so the concept of an inverse function must here be redefined. Let us do this.

Inverse and generalized inverse functions, cont

Definition of generalized inverse for a nondecreasing function

Let $F(x)$ be a nondecreasing function on $\mathbb{R}$, i.e. $F(x) : \mathbb{R} \to \mathbb{R}$. The generalized inverse $F^{\leftarrow}$ to F is then defined as

$$F^{\leftarrow}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\} \tag{13}$$

with the convention that inf of the empty set is ∞, i.e $\inf \emptyset = \infty$.

- Note that if $F(x)$ is a strictly increasing function then $F^{\leftarrow} = F^{-1}$, that is the generalized inverse $F^{\leftarrow}(y)$ will simply be the "usual" inverse $F^{-1}(y)$ defined as on the previous slides

By using the generalized inverse we can now define the α-quantile $q_\alpha(F)$ to a distribution function $F(x)$ as

$$q_\alpha(F) = F^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}, \quad 0 < \alpha < 1. \tag{14}$$

Generalized inverse, α-quantile and VaR

Hence, in view of the definition of a α-quantile (as a generalized inverse) $q_\alpha(F)$ and the definition of Value-at-Risk $\mathrm{VaR}_\alpha(L)$ we conclude that:

- Value-at-Risk $\mathrm{VaR}_\alpha(L)$ is the α-quantile $q_\alpha(F_L)$ of the loss distribution $F_L(x)$ where $F_L(x) = \mathbb{P}[L \leq x]$, that is

$$\mathrm{VaR}_\alpha(L) = F_L^{\leftarrow}(\alpha) = q_\alpha(F_L) \tag{15}$$

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), $F_L(x)$ will not have any "flat" regions, so that $F_L^{\leftarrow}$ will be the usual inverse function F_L^{-1}, and we then have that

$$\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha) = q_\alpha(F_L) \tag{16}$$

Hence, if we can find an analytical expression for the inverse function $F_L^{-1}(y)$, we can then due to (16) also find an analytical expression for the risk-measure Value-at-Risk $\mathrm{VaR}_\alpha(L)$

Value-at-Risk when L is a continuous random variable

- If the loss L is a continuous random variable so that $F_L(x)$ is strictly increasing and continuous, we have that $F_L^{-1}(y)$ is also continuous, and thus well defined and by definition

- Furthermore, from the definition of an inverse function (see previous slides) we have that $F_L(F_L^{-1}(y)) = y$ for all y such that $0 < y < 1$.

- From (16) we have
$$\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha) \tag{17}$$
so we then conclude that
$$F_L(\mathrm{VaR}_\alpha(L)) = F_L(F_L^{-1}(\alpha)) = \alpha \tag{18}$$
that is,
$$F_L(\mathrm{VaR}_\alpha(L)) = \alpha \tag{19}$$
or alternatively,
$$\mathbb{P}\left[L \leq \mathrm{VaR}_\alpha(L)\right] = \alpha \tag{20}$$

Example of Value-at-Risk when L is continuous r.v.

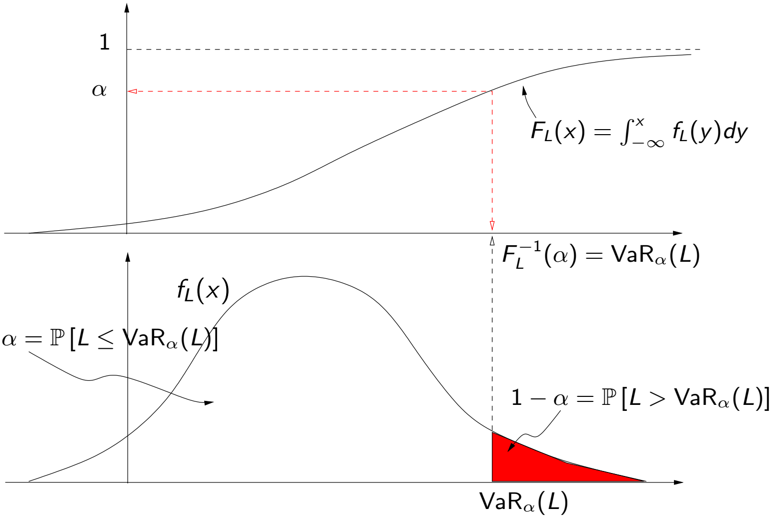


Figure: Visualization of definition of $\mathrm{VaR}_\alpha(L)$ when L is a continuous random variable. The red region has the area $1 - \alpha$

Value-at-Risk for static credit portfolios

- Consider mixed binomial model with m obligors and individual credit loss ℓ.
- By linearity of VaR, see Equation (12), we can w.l.o.g assume that the size of each loan is one monetary unit and that the loss ℓ is in %
- Let $F(x) = \mathbb{P}[p(Z) \leq x]$ where $p(Z)$ is the mixing distribution where Z can be a discrete or continuous random variable
- If we use the exact loss distribution $F_{L_m}(x)$ in (8) or (9) we compute VaR via the generalized inverse of $F_{L_m}(x)$
- However, if m is "large" and Z is a continuous random variable so that $F(x)$ and $F^{-1}(x)$ are continuous, we combine Equation (11) and (16) to get

$$\mathrm{VaR}_\alpha(L) \approx \ell \cdot m \cdot F^{-1}(\alpha) \tag{21}$$

- If m is "large" and Z is a discrete random variable we combine Equation (11) and (15) to get that

$$\mathrm{VaR}_\alpha(L) \approx \ell \cdot m \cdot F^{\leftarrow}(\alpha) \tag{22}$$

where $F^{\leftarrow}(x)$ is the generalized inverse of $F(x) = \mathbb{P}[p(Z) \leq x]$.

Expected shortfall

The expected shortfall $\text{ES}_\alpha(L)$ is defined as

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

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

$$\text{ES}_\alpha(L) = \mathbb{E}\left[\, L \mid L \geq \text{VaR}_\alpha(L)\right]$$

Let $F(x) = \mathbb{P}\left[p(Z) \leq x\right]$ where $p(Z)$ is the mixing distribution and Z is a continuous random variable so that $F(x)$ and $F^{-1}(x)$ are continuous,

Hence, for the same static credit portfolio as on the two previous slides, when m is large we have the following approximation formula for $\text{ES}_\alpha(L)$

$$\text{ES}_\alpha(L) \approx \frac{\ell \cdot m}{1-\alpha}\int_\alpha^1 F^{-1}(u)du$$

Mixed binomial models: the beta distribution

- One example of a mixing binomial model is to let $p(Z) = Z$ where Z is a beta distribution, $Z \sim \text{Beta}(a, b)$, which can generate heavy tails.
- 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

$$f_Z(z) = \frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1} \quad a, b > 0, \quad 0 < z < 1 \tag{23}$$

where

$$\beta(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1} dz = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{24}$$

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

$$\Gamma(y) = \int_0^\infty t^{y-1} e^{-t} dt \tag{25}$$

which satisfies the relation

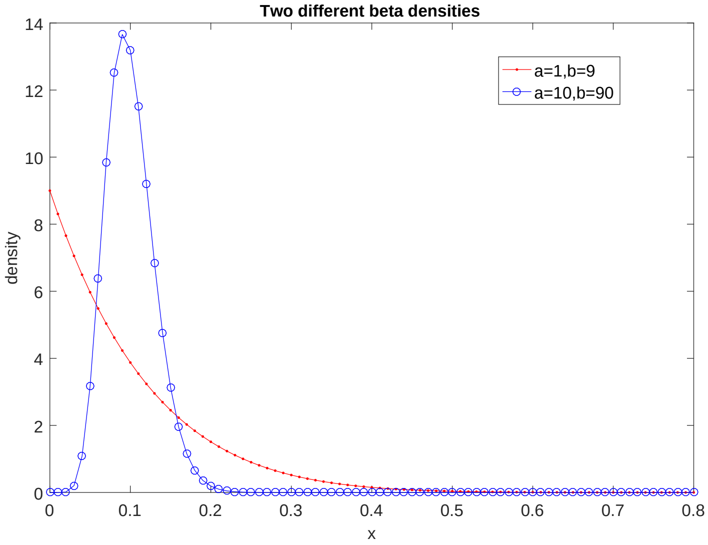
$$\Gamma(y+1) = y\Gamma(y) \tag{26}$$

for any y.

Mixed binomial models: the beta distribution, cont.

- By using Equation (24) and (26) one can show that $\beta(a, b)$ satisfies the recursive relation
$$\beta(a+1, b) = \frac{a}{a+b}\beta(a, b).$$
- Also note that (23) implies that $\mathbb{P}[0 \leq Z \leq 1] = 1$, that is $Z \in [0, 1]$ with probability one.
- If Z has beta distribution with parameters a and b, then by using Equation (24) and (26) one can show that
$$\mathbb{E}[Z] = \frac{a}{a+b} \quad \text{and} \quad \mathbb{E}\left[Z^2\right] = \frac{a(a+1)}{(a+b)(a+b+1)}$$
so the above equations together with definition of $\mathrm{Var}(Z)$ implies that $\mathrm{Var}(Z) = \frac{ab}{(a+b)^2(a+b+1)}$.
- By varying the parameters a and b, the density $f_Z(z)$ can take on quite different shapes (see next slide). Recall that $f_Z(z)$ is given by
$$f_Z(z) = \frac{1}{\beta(a, b)} z^{a-1}(1-z)^{b-1} \quad a, b > 0, \quad 0 < z < 1$$

Mixed binomial models: the beta distribution, cont.



Mixed binomial models: the beta distribution, cont.

- Consider a mixed binomial model where $p(Z) = Z$ has beta distribution with parameters a and b. Then, by using (6) one can show that

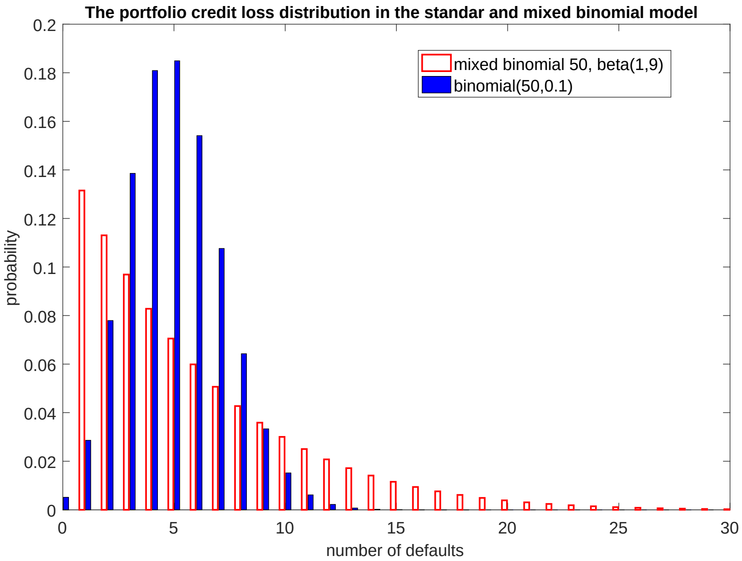
$$\mathbb{P}\left[N_m = k\right] = \binom{m}{k} \frac{\beta(a+k, b+m-k)}{\beta(a,b)}. \tag{27}$$

- It is possible to create **heavy tails** in the distribution $\mathbb{P}\left[N_m = k\right]$ by choosing the parameters a and b properly in (27). This will then imply more realistic probabilities for extreme loss scenarios, compared with the standard binomial loss distribution (see figure on next page).
- Furthermore, since $p(Z) = Z$, the distribution of $\frac{N_m}{m}$ converges to the distribution of the beta distribution, i.e

$$\mathbb{P}\left[\frac{N_m}{m} \leq x\right] \to \frac{1}{\beta(a,b)} \int_0^x z^{a-1}(1-z)^{b-1} dz \quad \text{as } m \to \infty \tag{28}$$

and for large m we use (28) instead of the exact method via (27).

Mixed binomial models: the beta distribution, cont.



Mixed binomial models: logit-normal distribution

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

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

where $\sigma > 0$ and $Z \sim N(0,1)$, that is Z is a standard normal random variable. Note that $p(Z) \in [0,1]$.

- Furthermore, if $x \in (0,1)$ then $p^{-1}(x)$ is well defined and given by

$$p^{-1}(x) = \frac{1}{\sigma}\left(\ln\left(\frac{x}{1-x}\right) - \mu\right). \tag{29}$$

- The mixing distribution $F(x) = \mathbb{P}\left[p(Z) \leq x\right] = \mathbb{P}\left[Z \leq p^{-1}(x)\right]$ for a logit-normal distribution is then given by

$$F(x) = \mathbb{P}\left[Z \leq p^{-1}(x)\right] = \int_{-\infty}^{p^{-1}(x)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = N(p^{-1}(x))$$

where $p^{-1}(x)$ is given as in Equation (29) and $N(x)$ is the distribution function of a standard normal distribution.

Mixed binomial models: logit-normal distribution, cont.

- Furthermore, the distribution of $\frac{N_m}{m}$ converges to $N(p^{-1}(x))$, that is

$$\mathbb{P}\left[\frac{N_m}{m} \leq x\right] \to N(p^{-1}(x)) \quad \text{as } m \to \infty \tag{30}$$

where $x \in (0,1)$ and $p^{-1}(x)$ is given as in Equation (29).

- In a mixed binomial model with logit-normal distribution as above, it is difficult to find closed formulas for quantities such as
  - $\mathbb{P}[X_i = 1] = \mathbb{E}[p(Z)]$,
  - $\text{Var}(X_i) = \mathbb{E}[p(Z)](1 - \mathbb{E}[p(Z)])$
  - $\text{Cov}(X_i, X_j) = \mathbb{E}\left[p(Z)^2\right] - \mathbb{E}[p(Z)]^2 = \text{Var}(p(Z))$ for $i \neq j$

- Hence, in the mixed binomial model with logit-normal distribution, the above quantities have to be determined with a computer

- Next lecture we will study a third mixed binomial model inspired by the Merton model.

Correlations in mixed binomial models

- Recall the definition of the correlation $\text{Corr}\,(X, Y)$ between two random variables X and Y, given by

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

  where $\text{Cov}(X, Y) = \mathbb{E}\,[XY] - \mathbb{E}\,[X]\,\mathbb{E}\,[Y]$ and $\text{Var}\,(X)) = \mathbb{E}\left[X^2\right] - \mathbb{E}\,[X]^2$.

- Furthermore, also recall that $\text{Corr}\,(X, Y)$ may sometimes be seen as a measure of the "dependence" between the two random variables X and Y.

- Now, let us consider a mixed binomial model as presented previously.

- We are interested in finding $\text{Corr}\,(X_i, X_j)$ for two pairs i, j in the portfolio (by the homogeneous-portfolio assumption this quantity is the same for any pair i, j in the portfolio where $i \neq j$).

- Below, we will therefore for notational convenience simply write ρ_X for the correlation $\text{Corr}\,(X_i, X_j)$.

Correlations in mixed binomial models, cont.

- Recall from previous slides that $\mathbb{P}\left[X_i = 1 \,|\, Z\right] = p(Z)$ where $p(Z)$ is the mixing variable.
- Furthermore, we also now that

$$\operatorname{Cov}(X_i, X_j) = \mathbb{E}\left[p(Z)^2\right] - \bar{p}^2 \quad \text{and} \quad \operatorname{Var}(X_i) = \bar{p}(1 - \bar{p}) \tag{31}$$

  where $\bar{p} = \mathbb{E}\left[p(Z)\right]$.
- Thus, the correlation ρ_X in a mixed binomial models is then given by

$$\rho_X = \frac{\mathbb{E}\left[p(Z)^2\right] - \bar{p}^2}{\bar{p}(1 - \bar{p})} \tag{32}$$

  where $\bar{p} = \mathbb{E}\left[p(Z)\right] = \mathbb{P}\left[X_i = 1\right]$ is the default probability for each obligor.
- Hence, the correlation ρ_X in a mixed binomial is completely determined by the fist two moments of the mixing variable $p(Z)$, that is $\mathbb{E}\left[p(Z)\right]$ and $\mathbb{E}\left[p(Z)^2\right]$.