Generalized Beta Regression Models for Random LGD

Xinzheng Huang,
TU Delft

PORTFOLIO CREDIT LOSS

Credit loss of a portfolio consisting of n obligors

$$L = \sum_{i=1}^{n} EAD_i \times LGD_i \times D_i$$

- Default indicator D_i, Probability-of-Default (PD) $p_i = P(D_i = 1)$.
- Exposure-at-Default (EAD): the size of a loan.
- Loss-Given-Default (LGD): proportion of the EAD that will be lost if a default occurs. LGD$\in [0, 1]$.

Portfolio credit loss

$$L = \sum_{i=1}^{n} EAD_i \times LGD_i \times D_i$$

1. the dependence between defaults: latent factor model that introduces systematic risk in default
2. the dependence between LGDs
3. the dependence between PD and LGD

Default Model: Vasicek Gaussian one-factor model

- The standardized asset log-return X_i is normally distributed.
- Default occurs if $X_i < \gamma_i$ where γ_i is a known default threshold.
- X is decomposed into a systematic part Y and an idiosyncratic part Z such that for obligor i,

$$X_i = \sqrt{\rho}Y + \sqrt{1-\rho}Z_i, \tag{1}$$

  where Y and all Z_i are i.i.d standard normal random variables.
- ρ is called the asset correlation.

LGD Models

- Current Industry practice
  - Dependence between LGDs, dependence between PD and LGD ignored.
  - constant LGD: Vasicek model, CreditRisk^{+}.
  - beta distributed LGD: KMV Portfolio Manager, CreditMetrics.
- Empirical evidence: Hu & Perraudin (2002); Altman et al. (2005)
  - LGD is positively correlated to the default rate; LGD is high when the default rate is high.
  - LGD is subject to systematic risk.

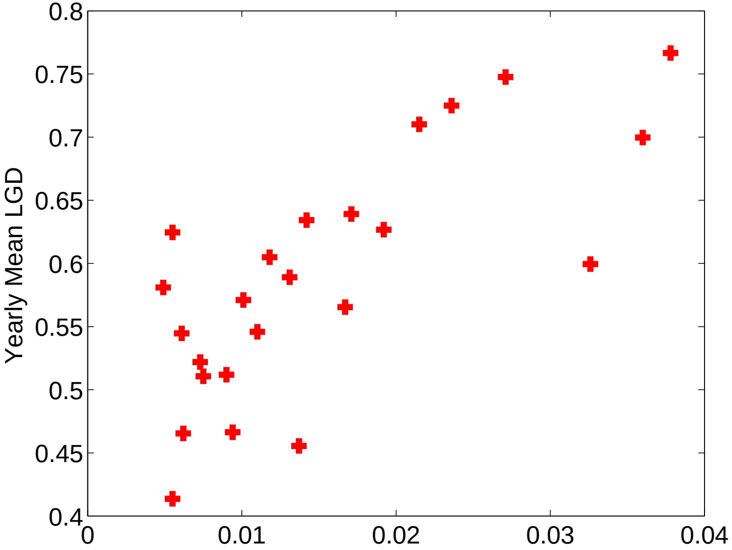


FIGURE: Yearly average default rate and mean LGD.

LGD MODELS

- Frye (2000)

$$LGD = \mu + \sigma\xi, \quad \xi = \sqrt{\tilde{\rho}}Y + \sqrt{1-\tilde{\rho}}\varepsilon,$$

- Pykhtin (2003)

$$LGD = \left(1 - e^{\mu+\sigma\xi}\right)^{+},$$

- Andersen & Sidenius (2004)

$$LGD = \Phi(\mu + \sigma\xi)$$

- Duellmann & Trapp (2004); Roesch & Scheule (2005)

$$LGD = \frac{1}{1 + \exp(\mu + \sigma\xi)}$$

- Giese (2006);Bruche & Gonzalez-Aguado (2008)

$$LGD \sim Beta(\alpha(Y), \beta(Y))$$

LGD Models

A new framework: Generalized Beta Regression Models

- Easy to interpret
- Accommodates better heteroscedastic errors
- Broad enough to fit the historical data and empirical testing of different models can be unified
- Allows efficient numerical procedures (normal approximation, saddlepoint approximation) to calculate the portfolio loss distribution.

PARAMETERIZATION OF A BETA DISTRIBUTION

- probability density function

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}.$$

- mean

$$\mu = E(X) = \frac{\alpha}{\alpha+\beta},$$

- variance

$$\sigma^2 = Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{\mu(1-\mu)}{\alpha+\beta+1} = \frac{\mu(1-\mu)}{\varphi+1}.$$

- Dispersion parameter $\varphi = \alpha+\beta$.

Beta Regression Model: Ferrari & Cribari-Neto (2004)

- Parameterization of a beta distribution: mean μ and dispersion φ.
- Mean Model (GLM)
  - a linear predictor η

    $$\eta = \alpha\zeta$$

    where ζ is a vector of explanatory variables and α is a vector of the corresponding regression coefficients.
  - a monotonic, differentiable link function g

    $$g(\mu) = \eta.$$

- Dispersion parameter φ: a nuisance parameter

BETA REGRESSION MODEL

- A parsimonious one-factor model: $\zeta = [1, Y]'$, where Y is the factor also drives PD.

$$LGD \sim Beta\left(g^{-1}(a_1 + a_2 Y), \varphi\right)$$

- Choices of link function: $g^{-1}(\cdot) : \mathbb{R} \rightarrow [0,1]$.
  logit link
  $$\mu = \frac{e^{\eta}}{1+e^{\eta}},\ \eta = \log\left(\frac{\mu}{1-\mu}\right),$$
  the probit link
  $$\mu = \Phi(\eta),\ \eta = \Phi^{-1}(\mu).$$
  cauchit,complementary log-log, scaled probit...

Generalized Beta Regression Models

1. Jointly modeling mean and dispersion (JGLM)

$$h(\varphi) = b\zeta, \quad \varphi = e^{b\zeta}$$

$$LGD \sim Beta\left(g^{-1}(a_1 + a_2 Y), e^{b_1 + b_2 Y}\right)$$

2. Mixed model: random effect in the linear predictor (GLMM)

$$g(\mu) = \eta = a\zeta + \nu, \quad \nu \sim N(0, \sigma_\nu^2)$$

$$LGD \sim Beta\left(g^{-1}(a_1 + a_2 Y + \nu), \varphi\right)$$

INFERENCE

- Method of moments and least squares
- Maximum likelihood

$$\ell(\mu,\varphi;x) = (\mu\varphi - 1)\log(x) + [(1-\mu)\varphi - 1]\log(1-x) + \\ + \log\Gamma(\varphi) - \log\Gamma(\mu\varphi) - \log\Gamma[(1-\mu)\varphi]$$

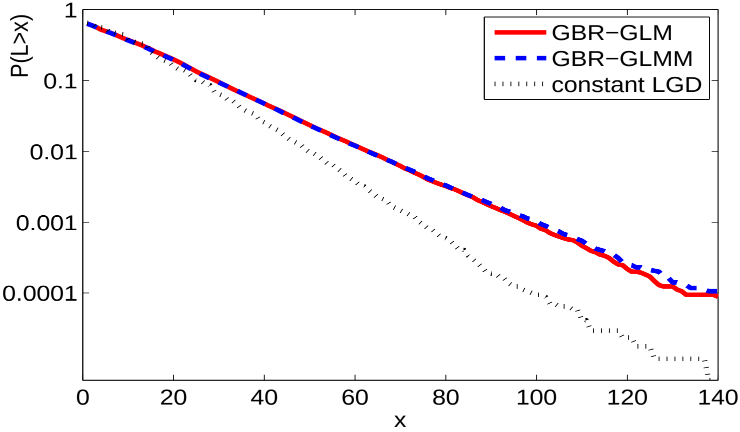
Marginal likelihood in GLMM

$$\ell_m = \sum_{t=1}^{T} \log\left(\int \prod_{k=1}^{K_t} L(a,\varphi,\zeta_t,v_t;x_{t,k}) p_{\sigma_v}(v_t) d v_t \right)$$

- Model selection: Likelihood ratio test, AIC, BIC

	GLM	JGLM	GLMM
a_1	0.3459 (0.0359)	0.3471	0.3319
a_2	-0.3213 (0.0298)	-0.3246	-0.3307
φ	3.0276 (0.1149)	-	3.3240
b_1	-	1.0879	-
b_2	-	-0.0306	-
σ_v	-	-	0.2943
-2ℓ	-402.74	-403.34	-468.78
AIC	-398.74	-395.34	-460.78
BIC	-381.67	-375.25	-440.69

TABLE: Estimates of the models with logit link for mean LGD.



P(L>x)
1
0.1
0.01
0.001
0.0001
0
20
40
60
80
100
120
140
x
GBR−GLM
GBR−GLMM
constant LGD

Saddlepoint approximation to the loss distribution

The conditional CGF and its derivatives up to second order are

$$\kappa(t,Y) = \sum_{i=1}^{n} \log\left[1 - p_i + p_i\, {}_1F_1(\alpha_i, \alpha_i+\beta_i; w_i t)\right].$$

$$\kappa'(t,Y) = \sum_{i=1}^{n} \frac{w_i p_i\, {}_1F_1(\alpha_i+1, \alpha_i+\beta_i+1; w_i t)}{1 - p_i + p_i\, {}_1F_1(\alpha_i, \alpha_i+\beta_i; w_i t)} \frac{\alpha_i}{\alpha_i+\beta_i},$$

$$\kappa''(t,Y) = \sum_{i=1}^{n} \left\{ \frac{w_i^2 p_i \alpha_i(\alpha_i+1)\, {}_1F_1(\alpha_i+2, \alpha_i+\beta_i+2; w_i t)}{(\alpha_i+\beta_i)(\alpha_i+\beta_i+1)[1 - p_i + p_i\, {}_1F_1(\alpha_i, \alpha_i+\beta_i; w_i t)]} - \frac{w_i^2 p_i^2 \alpha_i^2\, {}_1F_1(\alpha_i+1, \alpha_i+\beta_i+1; w_i t)^2}{(\alpha_i+\beta_i)^2[1 - p_i + p_i\, {}_1F_1(\alpha_i, \alpha_i+\beta_i; w_i t)]^2} \right\}.$$

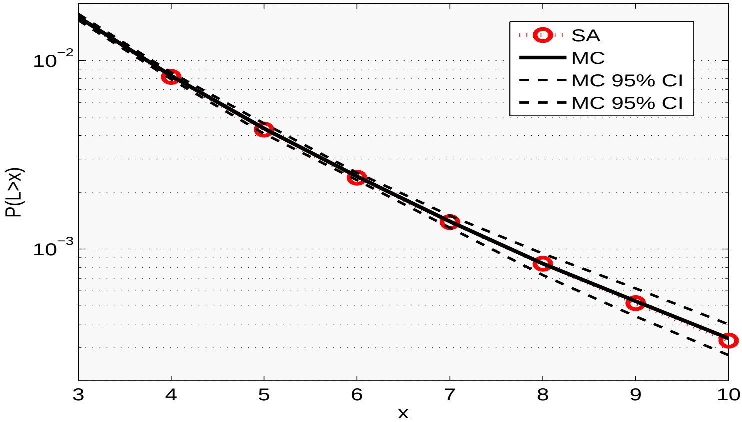


FIGURE: The loss distribution obtained from the saddlepoint approximation (SA) compared to results based on Monte Carlo (MC) simulation of two hundred thousand scenarios. The MC 95% confidence interval (CI) are based on the standard deviation calculated using 10 simulated sub-samples of 20 thousand scenarios each.