Regressão e Modelos para Séries Temporais com Base na EGEDD

Distribuição Dagum Exponencial Generalizada Exponencializada

Vítor Bernardo Silveira Pereira

Motivação

- A indentificação de uma **distribuição** apropropriada é muito importante para a **análise estatística**;
- A distribuição Dagum, idealizada por Camilo Dagum, tem ganhado muita atenção na sua utilidade para modelagem da distribuição de renda e análise de confiabilidade:
  - Distribuição exponenciada **Kumaraswamy-Dagum** (Silva et al., 2015);
  - Distribuição **Dagum Exponencial Generalizada Exponencializada** proposta em (Nasiru et al., 2019);
  - **Gompertz Generalizada**, proposta por El-Gohary et al. (2013), entre outros...
- Para modelagem de média, geralmente se utiliza, a **Regressão Linear Normal** ou **Modelos Lineares Generalizados**;
- Na prática, alguns dados não se adequam a essas distribuições, então nos últimos anos vêm sendo propostos **modelos alternativos**;

Motivação

- Nos novos modelos, tem-se uma reparametrização em termos da média ou mediana, para possibilitara interpretação dos parâmetros em termos de medidas de posição ou precisão;
- Com a reparametrização, uma estrutura de regressão é incluída para **modelar a média ou a médiana**, assim tendo a mesma abordagem dos MLG'S;
- Na **EGEDD**, no entanto, **não há forma fechada para média e variância**;
- Em Séries Temporais, a abordagem mais usual é a Box e Jenkins, no entanto também estão surgindo modelos alternativos;
- Assim a proposta do projeto, é **desenvolver novos modelos para distribuições com suporte nos reais**, baseando-se na EGEDD e na distrubuição Dagum.
- **Aplicações:** principalmente na área da modelagem de dados de sobrevivência.

Propriedades da EGED

Função de distribuição acumulada

$$F_Y(y) = 1 - \left\{1 - \left[1 - \left(1 - \left(1 + \alpha y^{-\delta}\right)^{-\sigma}\right)^{\gamma}\right]^{\eta}\right\}^{\lambda}$$

Suporte

$$y > 0$$

Espaço Paramétrico

$$\alpha, \sigma, \lambda, \delta, \eta, \gamma > 0$$

Os parâmetros $\sigma, \lambda, \delta, \eta, \gamma$ são parâmetros de forma e o parâmetro α é de escala.

Função Quantílica

$$Q_Y(\tau) = \left\{ \frac{1}{\alpha} \left[\left(1 - \left(1 - \left(1 - (1-\tau)^{\frac{1}{\lambda}} \right)^{\frac{1}{\eta}} \right)^{\frac{1}{\gamma}} \right)^{-\frac{1}{\sigma}} - 1 \right] \right\}^{-\frac{1}{\delta}}, \text{ onde } \quad \tau \in (0,1).$$

Função densidade de probabilidade

$$f_Y(y) = \frac{\alpha\sigma\lambda\delta\eta\gamma\left(1+\alpha y^{-\delta}\right)^{-\sigma-1}\left[1-\left(1+\alpha y^{-\delta}\right)^{-\sigma}\right]^{\gamma-1}}{y^{\delta+1}\left\{1-\left[1-\left(1-\left(1+\alpha y^{-\delta}\right)^{-\sigma}\right)^{\gamma}\right]^{\eta}\right\}^{1-\lambda}}$$

$$+\frac{[1-(1-(1+\alpha y^{-\delta})^{-\sigma})^{\gamma}]^{\eta-1}}{y^{\delta+1}\left\{1-\left[1-\left(1-\left(1+\alpha y^{-\delta}\right)^{-\sigma}\right)^{\gamma}\right]^{\eta}\right\}^{1-\lambda}} \quad y>0$$

Não há fórmula fechada para a média e variância

Submodelos da EGEDD

Distribuição	α	λ	σ	δ	η	γ
EGDD	α	1	σ	δ	η	γ
DD	α	1	σ	δ	1	1
EGEBD	1	λ	σ	δ	η	γ
DB	1	1	σ	δ	1	1
EGBD	1	1	σ	δ	η	γ
EGEFD	α	λ	1	δ	η	γ
EGFD	α	1	1	δ	η	γ
FD	α	1	1	δ	1	1

Notamos que existes parâmetros melhores para fazer as reparametrizações, sem perder a generalidade da EGED como o δ.

Reparametrizações

$$\mu = \left\{ \frac{1}{\alpha} \left[\left(1 - \left(1 - \left(1 - (1-\tau)^{\frac{1}{\lambda}} \right)^{\frac{1}{\eta}} \right)^{\frac{1}{\gamma}} \right)^{-\frac{1}{\sigma}} - 1 \right] \right\}^{-\frac{1}{\delta}}$$

Repametrização por α

$$\alpha = \frac{1}{\mu^{-\delta}} \left\{ \frac{1}{\alpha} \left[\left(1 - \left(1 - \left(1 - 0,5^{\frac{1}{\lambda}} \right)^{\frac{1}{c}} \right)^{\frac{1}{d}} \right)^{-\frac{1}{\sigma}} - 1 \right] \right\}.$$

Reparametrização por δ

$$\delta = -\frac{1}{\log(\mu)} \log \left\{ \frac{1}{\alpha} \left[\left(1 - \left(1 - \left(1 - (1-\tau)^{\frac{1}{\lambda}} \right)^{\frac{1}{\eta}} \right)^{\frac{1}{\gamma}} \right)^{-\frac{1}{\sigma}} - 1 \right] \right\}$$

Vítor Bernardo Silveira Pereira - Email: vitorpereira3115@gmail.com

Modelo de Séries Temporais Com Regressão

Incluindo a estrutura de regressão pelo quantil, por:

$$g(\mu_t) = x_t^\top \beta = \sum_{j=1}^{k} x_{tj}\beta = \eta_t, \quad t = 1, \ldots, n,$$

$$com \quad \beta = (\beta_1, \ldots, \beta_k)^\top$$

Função de ligação

A função de ligação é $g(\mu_t) = \log(\mu_t)$, então temos:

$$\mu_t = \exp\left(x_t{}^\top \beta\right) = \exp(\eta_t), \quad \text{para} \quad t = 1, \cdots, n.$$

Assim, ficaremos com o vetor de parâmetros, sendo:

$$\theta = (\alpha, \sigma, \lambda, \eta, \gamma, \beta)$$

Em que μ_t é o quantil da observação y_t no tempo t e

Modelo ARMA(p,q)

Modelo geral para μ:

$$g(\mu_t) = x_t^\top \beta + u_t$$

Sabendo que pela estrutura arma temos uma relação que inclui o componente u_t: $u_t = \frac{\phi(\beta)}{\theta(\beta)}\epsilon_t$

$$g(\mu_t) = x_t^\top \beta + \frac{\phi(\beta)}{\theta(\beta)}\epsilon_t$$

Onde:

$$\phi(\beta) = -\sum_{j=0}^{p} \phi_j z^j, \quad \theta(\beta) = -\sum_{j=0}^{q} \theta_j z^j$$

Desenvolvendo um pouco algébricamente, conseguimos:

$$\phi(\beta)(g(\mu_t) - x_t^\top \beta) = \phi(\beta)\epsilon_t$$

$$\left(1 \ - \ \phi_1 \beta \ - \ \phi_2 \beta^2, \ \ldots, \ \phi_p \beta^p \ \right) \left(g\left(\mu_t\right) \ - x_t^\top\right) \ = \ \phi\left(\beta\right) \epsilon_t$$

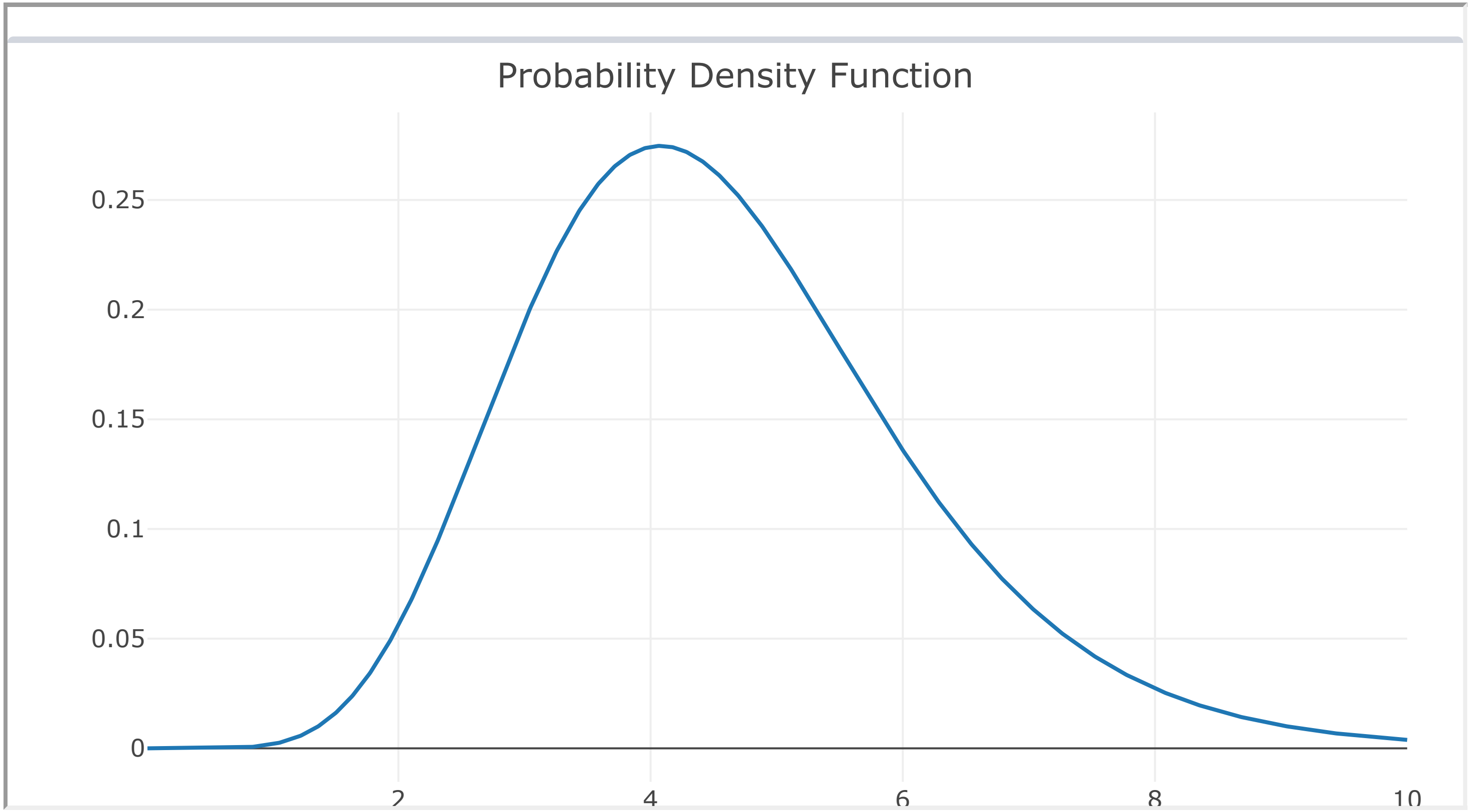
Até chegar no:

Modelo final ARMA:

$$g(\mu_t) - x_t^\top \beta = \sum_{p=1}^{n} \phi_p (g(\mu_t) - x_t^\top \beta) + \phi(\beta)\epsilon_t$$

Outros Procedimentos Estatísticos:

- Construção de testes de hipótese para verificar distribuição dos dados, verificando se é para a EGEDD ou um submodleo;
- Utilização de métodos númericos para execução de simulações (Newton-Rapson, Optim, GosoInp);
- Utilização de Simulações de Monte Carlo para estimação de parâmetros e análise de diagnósticos dos modelos;


Probability Density Function
0.25
0.2
0.15
0.1
0.05
0
2
4
6
8
10