Skewed regression and the PC prior for the skewness

King Abdullah University of Science and Technology

جامعة الملك عبدالله
للعلوم والتقنية

April 2022

Outline

Introduction

- Symmetric $\xrightarrow{\text{skewness parameter}}$ Non-symmetric
- A continuous random variable X, follows a skew-normal (SN) distribution with location, scale and skewness(shape) parameters ξ, ω and α, respectively, if the probability density function (pdf) is as follows:

$$g(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left[\alpha\left(\frac{x-\xi}{\omega}\right)\right],$$

with $E[X] = \xi + \omega\delta\sqrt{2/\pi}$ and $V[X] = \omega^2\left(1 - 2\delta^2/\pi\right)$, where $\delta = \alpha/\sqrt{1+\alpha^2}$.

Skew probit regression

$$\begin{aligned} y_i &\sim \text{Binomial}(N_i, p_i) \\ p_i &= G(\eta_i), \quad i = 1, \ldots, n \end{aligned}$$

where $G(\cdot)$ is the CDF of the Skew-Normal that depends on (ξ, ω, α).

What are others (most) doing?

- $\xi = 0, \omega = 1$ like in the probit case
- What is the implication?

$$E[X] = \alpha\sqrt{\frac{2}{\pi(1+\alpha^2)}},$$

and variance

$$V[X] = 1 - \frac{2\alpha^2}{\pi(1+\alpha^2)},$$

Anchored link function

$E[X] = 0, V[X] = 1$

$$F(y|\alpha) = \int_{-\infty}^{y} f(x|\alpha)\, dx$$

where

$$f(x|\alpha) = \frac{2}{\omega(\alpha)} \phi\left(\frac{x - \xi(\alpha)}{\omega(\alpha)}\right) \Phi\left[\alpha\left(\frac{x - \xi(\alpha)}{\omega(\alpha)}\right)\right],$$

$$\xi(\alpha) = -\omega\alpha\sqrt{\frac{2}{\pi(1+\alpha^2)}},$$

and

$$\omega(\alpha) = \sqrt{\left(1 - \frac{2\alpha^2}{\pi(1+\alpha^2)}\right)^{-1}}.$$

Identifiability

Most have remarked that when there is no continuous covariate then the intercept and skewness parameters are unidentifiable.

- Solution? Cannot do skew probit regression without continuous covariates.

What is an intercept?

$$\eta_i = \beta_0 + \boldsymbol{\beta}'\boldsymbol{X}_i + \sum_{k=1}^{K} f^k(\boldsymbol{Z}_i), \tag{1}$$

Consider probit regression with one centered covariate X,

- Probit regression:

$$p = \text{Prob}[Y = 1] = \Phi(\beta_0 + \beta_1 X).$$

Now if $\beta_1 X = 0$, then

$$q = \text{Prob}[Y = 1] = \Phi(\beta_0),$$

What is an intercept?

Consider skew probit regression with one centered covariate X,

- Skew probit regression:

$$p = \text{Prob}[Y = 1] = F(\beta_0 + \beta_1 X | \alpha).$$

So if $\beta_1 X = 0$ then

$$q = \text{Prob}[Y = 1] = F(\beta_0 | \alpha)$$

So redefine

$$\beta_0 = F^{-1}(q | \alpha)$$

Priors for α

$$g(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left[\alpha\left(\frac{x-\xi}{\omega}\right)\right], \quad \alpha \in \Re$$

KLD between SN(α) and N

$$\begin{aligned} \text{KLD}(\alpha) &= \int f(x|\alpha) \log \frac{f(x|\alpha)}{f(x|\alpha = 0)} dx \\ &\approx c_1\alpha^6 + c_2\alpha^8 + c_3\alpha^{10}. \end{aligned} \tag{2}$$

So the estimator of α would be $n^{\frac{1}{6}}$ consistent?

Rather consider $\gamma = \alpha^3$.

What does this mean?

KLD between SN(α) and N

$$\begin{aligned} \text{KLD}(\alpha) &= \int f(x|\alpha) \log \frac{f(x|\alpha)}{f(x|\alpha=0)} dx \\ &\approx c_1\alpha^6 + c_2\alpha^8 + c_3\alpha^{10}. \end{aligned} \tag{2}$$

So the estimator of α would be $n^{\frac{1}{6}}$ consistent?
Rather consider $\gamma = \alpha^3$.
What does this mean?

KLD between SN(α) and N

$$\begin{aligned} \text{KLD}(\alpha) &= \int f(x|\alpha) \log \frac{f(x|\alpha)}{f(x|\alpha=0)} dx \\ &\approx c_1\alpha^6 + c_2\alpha^8 + c_3\alpha^{10}. \end{aligned} \tag{2}$$

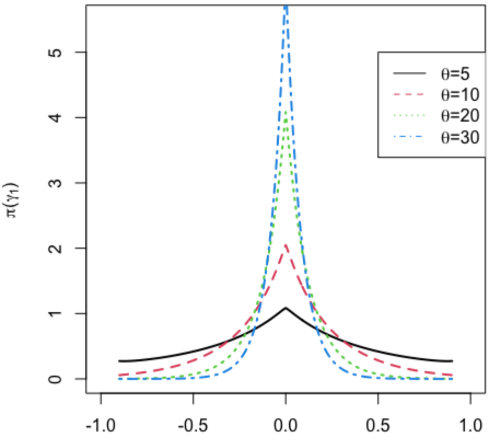
So the estimator of α would be $n^{\frac{1}{6}}$ consistent?
Rather consider $\gamma = \alpha^3$.
What does this mean?

PC prior for $\alpha/\gamma/\gamma_1$

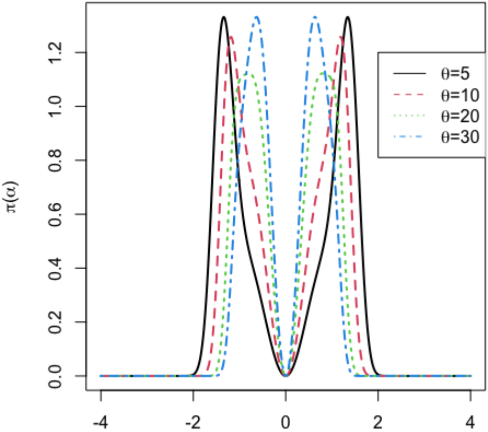
$$\gamma_1 = \frac{(4-\pi)\left(\sqrt{\frac{2\delta^2}{\pi}}\right)^3}{2(1-\frac{2\delta^2}{\pi})^{\frac{3}{2}}}, \tag{3}$$

where $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ and $-0.99527 < \gamma_1 < 0.99527$.

PC prior for γ_1



PC prior for α



Bernoulli regression

Here we focus our attention on samples of size 200 of binary trials, and the scenario's we consider are:

1. $q = \frac{1}{2}, \beta_1 = 1, \gamma_1 = -\frac{2}{3}(\alpha = -10), N_i = 1$
2. $q = \frac{1}{2}, \beta_1 = 1, \gamma_1 = 0(\alpha = 0), N_i = 1$

We consider the PC prior as well as the Gaussian prior for the skewness parameter.

Results for binary trials

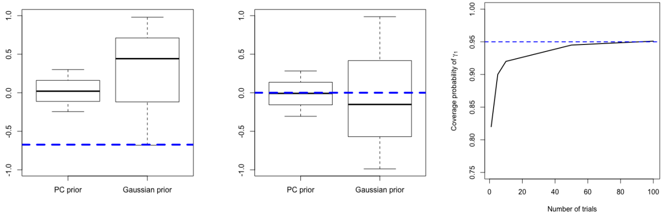
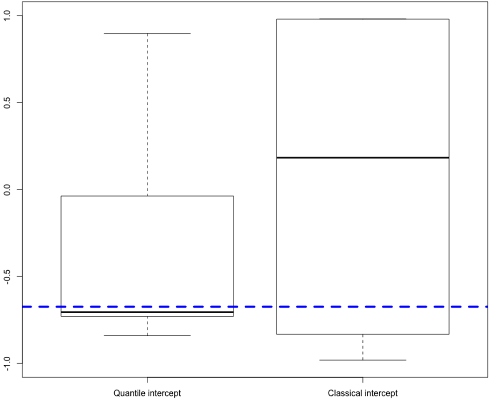


Figure: 95% credible intervals for γ_1 with $n_i = 1$ and $\gamma_1 = -\frac{2}{3}$ (left) or $\gamma_1 = 0$ (middle). Coverage probabilities for γ_1 under scenario 1 as N_i increases (right)

Old vs new intercept

$$q = 0.4, \beta_1 = 0.1, \gamma_1 = -\frac{2}{3} \rightarrow \beta_0 = -0.125$$



Thank you • شكرا

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology