Negative Binomial Regression

One Problem with Poisson model is that $Var(Y_i|X) = E(y_i|X_i)$

One alternative is the NEGBIN
Basically you introduce and individual, unobservable effect into the regression

log of conditional mean $\log(\mu_i) = X_i'\beta + e_i$

$\log(\mu_i) = \log\lambda_i + \log u_i$

μ MU $\quad$ u v_i

Poisson $\quad Prob(Y_i = y_i) = \dfrac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$

let $\mu_i = \lambda_i u_i$

$\log(\mu_i) = \log\lambda_i + \log u_i$
$= X_i'\beta + e_i$

e_i is either spec. error or in linear regression or unmeasured heterogeneity (cross sect)

conditional

$$f(y_i \mid x_i \mu_i) = \frac{e^{-\lambda_i \mu_i} (\lambda_i \mu_i)^{y_i}}{y_i!}$$

unconditional exp.

$$f(y_i \mid x_i) = \int_0^\infty f(y_i \mid x_i, \mu_i)\, g(\mu_i)\, d\mu_i$$

So, you pick a density, $g(\mu_i)$, and you are in business.

Common choice is the GAMMA dist for $\mu_i = e^{\varepsilon_i}$

you have to omit the constant for purposes of identification After the appropriate Normalization

$$g(\mu_i) = \frac{\theta^\theta}{\Gamma(\theta)} e^{-\theta \mu_i} \mu_i^{\theta - 1}$$

Plug this in to the uncond density

$$\int_0^\infty \frac{e^{-\lambda_i \mu_i}(\lambda_i \mu_i)^{y_i}}{y_i!} \cdot \frac{\theta^\theta}{\Gamma(\theta)} e^{-\theta\mu_i} \mu_i^{\theta-1} \, d\mu_i$$

Pull out the non-μ_i stuff

$$\frac{\lambda_i^{y_i}\theta^\theta}{y_i!\,\Gamma(\theta)} \int_0^\infty e^{-\lambda_i\mu_i}(\cancel{\lambda_i}\mu_i)^{y_i} e^{-\theta\mu_i} \mu_i^{\theta-1} \, d\mu_i$$

$$\frac{\lambda_i^{y_i}\theta^\theta}{y_i!\,\Gamma(\theta)} \int_0^\infty \mu_i^{y_i+\theta-1} e^{-(\lambda_i+\theta)\mu_i} \, d\mu_i$$

NOTE: ~~[illegible]~~

For integers $(y_i!) = \Gamma(y_i+1)$ ✓

$$\frac{\lambda_i^{y_i}\theta^\theta}{\Gamma(y_i+1)\,\Gamma(\theta)} \int_0^\infty \mu_i^{y_i+\theta-1} e^{-(\lambda_i+\theta)\mu_i}$$

note:

$$\int_0^\infty e^{i\mu x} x^{k-1} \, dx = \frac{\Gamma(k)}{(-i\mu)^k} \qquad \lambda > 0$$

Integral

$$= \frac{\Gamma(\theta+y_i)}{(\lambda_i+\theta)^{\theta+y_i}}$$

$$\therefore \quad \frac{\theta^{\theta}\lambda_i^{y_i}\,\Gamma(\theta+y_i)}{\Gamma(y_i+1)\,\Gamma(\theta)\,(\lambda_i+\theta)^{\theta+y_i}}$$

$$\frac{\Gamma(\theta+y_i)}{\Gamma(y_i+1)\,\Gamma(\theta)} \cdot \frac{\theta^{\theta}\lambda_i^{y_i}}{(\lambda_i+\theta)^{\theta+y_i}}$$

$$\frac{\theta^{\theta}\lambda_i^{y_i}}{(\lambda_i+\theta)^{\theta}(\lambda_i+\theta)^{y_i}}$$

$$= \left(\frac{\theta}{\lambda_i+\theta}\right)^{\theta} \cdot \left(\frac{\lambda_i}{\lambda_i+\theta}\right)^{y_i}$$

$$\left(\frac{\theta}{\lambda_i+\theta}\right)^{\theta} \cdot r_i^{y_i}$$

$$= (1-r_i)^{\theta}\, r_i^{y_i}$$

$$1 - r_i = \quad 1 - \frac{\lambda_i}{\lambda_i+\theta}$$

$$(\lambda_i+\theta-\lambda_i)/\lambda_i+\theta_i$$

$$\frac{\Gamma(\theta + y_i)}{\Gamma(y_i + 1)\,\Gamma(\theta)} \; r_i^{y_i} (1 - r_i)^{\theta}$$

where $r_i = \dfrac{\lambda_i}{\lambda_i + \theta}$

This is the neg bin. dist!

$E(y_i|\lambda_i) = \lambda_i$

$Var(y_i | \lambda_i, \theta) = \lambda_i (1 + (\frac{1}{\theta}) \lambda_i)$

Form JPDF and
Use MLE.

Test for Poisson using Likelihood Rat test

$H_0: \theta = 0$

$H_A: \theta \neq 0$

LM Test

$H_0: \theta = 0$ Poisson

$H_A: \theta \neq 0$ Neg BIN

Remember, only the restricted model has to be estimated, $\therefore$ you can estimate the simpler Poisson model and do this test before resorting to the NegBIN.

$$LM = \left[\frac{\sum_{i=1}^{T} \hat{w}_i (y_i - \hat{\lambda}_i)^2 - y_i}{\sqrt{2 \sum_{i=1}^{T} \hat{w}_i \hat{\lambda}_i^2}} \right]^2$$

$\hat{w}_i$ depends on the distribution under H_a: in our case it is Neg BIN and $\hat{w}_i = 1$.

In this case it simplifies to

$$LM = \frac{\hat{e}'\hat{e} - n\bar{y}}{2\hat{\lambda}'\hat{\lambda}}$$

where $\hat{e}_i = y_i - \hat{\lambda}_i \qquad \lambda_i = e^{x_i'\beta}$

under H_0: $LM \sim \chi^2_1$