## [7] Count Data Models

(1) Some Discrete Probability Density Functions

- Binomial Distribution:
  - Tossing a coin m times.
  - p = probability of having head from a trial.
  - y = # of having heads from n trials (y = 0, 1, ... , m).
  - $f_b(y|n) = \begin{pmatrix} m \\ y \end{pmatrix} p^y(1-p)^{1-y} = \frac{m!}{y!(m-y)!} p^y(1-p)^{m-y}$.
  - E(y) = mp; var(y) = mp(1-p).

- Poisson Distribution:
  - Let p = μ/m for a binomial distribution.
  - $f_p(y) = \lim_{m\to\infty} f_b(y|m) = \frac{\mu^y}{y!} e^{-\mu}$, y = 0, 1, 2, ... .
  - E(y) = var(y) = μ.

**Digression to the Chocolate Chip Cookies problem:**

- Wish to put at least one CC on a cookie with 95%.
- The CC machine locates CC's in a cookie following a Poisson Distribution with $\mu$.

  $\rightarrow$ $Pr(y = 0) < 0.05 \rightarrow e^{-\mu} < 0.05 \rightarrow \mu > -\ln(0.05) = 2.995$.

  $\rightarrow$ You need a machine with $\mu$ = at least 3.

- Geometric Distribution:
  - y = # of trials until having a head.
  - $f_g(y) = (1-p)^{y-1}p$, y = 1, ....
  - $E(y) = 1/p$; $var(y) = (1-p)/p^2$.

- Negative Binomial Distribution:
  - y = # of trials until having r heads.
  - $f_n(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$, y = r, r+1, ...
  - $E(y) = r/p$; $var(y) = r(1-p)/p^2$.

(2) Basic Poisson Regression Model

- Assume $y_i$ i.i.d. with Poisson($\mu_i$), where $\mu_i = \exp(x_i'\beta)$.

  → EX: $y_i$ = # of visiting doctors.

- $f(y_i | x_i) = \frac{e^{-\mu_i}(\mu_i)^{y_i}}{y_i!} = \frac{e^{-e^{x_i'\beta}}(e^{x_i'\beta})^{y_i}}{y_i!}$.

- $\ln f(y_i | x_i) = -e^{x_i'\beta} + y_i(x_i'\beta) - \ln(y_i!)$.

- $l_N(\beta) = \Sigma_{i=1}^N \left\{ y_i(x_i'\beta) - e^{x_i'\beta} - \ln(y_i!) \right\}$.

- $\frac{\partial l_N(\beta)}{\partial \beta} = \Sigma_{i=1}^N x_i'(y_i - e^{x_i'\beta})$.

  → The Poisson ML estimator of β can be viewed as a GMM estimator based on

$$E\left(x_i'(y_i - e^{x_i'\beta})\right) = 0.$$

  → This moment condition is valid as long as $E(y_i | x_i) = e^{x_i'\beta}$.

  → It means that the Poisson ML estimator of β is consistent even if the Poisson assumption is incorrect.

- $B_N(\beta) = \Sigma_{i=1}^N x_i x_i'(y_i - e^{x_i'\beta})^2$; $H_N(\beta) = -\Sigma_{i=1}^N e^{x_i'\beta} x_i x_i'$.

- If the Poisson assumption is truly correct, use

$$[-H_N(\hat{\beta}_{POI-ML})]^{-1} \text{ or } [B_N(\hat{\beta}_{POI-ML})]^{-1}$$

  as an estimate of $Cov(\hat{\beta}_{POI-ML})$.

- If you are not sure, use

$$[-H_N(\hat{\beta}_{POI-ML})]^{-1} B_N(\hat{\beta}_{POI-ML}) [-H_N(\hat{\beta}_{POI-ML})]^{-1}.$$

- For the measures of goodness of fit, see Greene.

- In fact, we can estimate β by NLLS applied to $y_i = e^{x_i'\beta} + \varepsilon_i$ with heteroskedastic error terms.

**Digression to NLLS with Heteroskedastic Errors:**

- $y_i = h(x_i, \beta) + \varepsilon_i$.
- Let $H(x_i, \beta) = \dfrac{\partial h(x_i, \beta)}{\partial \beta}$.
- The NLLS estimator of β ($\hat{\beta}_{NL}$) minimizes:

  $$\Sigma_{i=1}^{N}(y_i - h(x_i, \beta))^2.$$

- $$\begin{aligned} Cov(\hat{\beta}_{NL}) \approx & \left(\Sigma_{i=1}^{N} H(x_i, \hat{\beta}_{NL}) H(x_i, \hat{\beta}_{NL})'\right)^{-1} \\ & \times \Sigma_{t=1}^{T} e_i^2 H(x_i, \hat{\beta}_{NL}) H(x_i, \hat{\beta}_{NL})' \\ & \times \left(\Sigma_{t=1}^{T} H(x_t, \hat{\beta}_{NL}) H(x_t, \hat{\beta}_{NL})'\right)^{-1} \end{aligned}$$

  where $e_i = y_i - h(x_i, \hat{\beta}_{ML})$.

**End of Digression**

(2) Compound Poisson Model (Negative Binomial Model):

- Hausman, Hall, Griliches (HHG, ECON, 1984), and Cameron and Trivedi (CT, JAE, 1986).
- Assume that the $y_i$ follow Poisson($\lambda_i$), where $\lambda_i = e^{x_i'\beta+\alpha_i} = \mu_i e^{\alpha_i}$,

  $\mu_i = e^{x_i'\beta}$, $E(e^{\alpha_i}) = 1$, and the $e^{\alpha_i}$ follow a Gamma distribution:

$$f_{gamma}(\eta) = \frac{\theta^\theta}{\Gamma(\theta)} e^{-\theta\eta}\eta^{\theta-1},$$

  where $0 < \eta < \infty$ and $\Gamma(\theta) = \int_0^\infty t^{\theta-1}e^{-t}dt$.

  → Here, $\alpha_i$ is an unobservable individual effect.

**Digression to Gamma distribution:**

- The most general form of the Gamma density function is given:

$$f_{gen-gamma}(y|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1}e^{-y/\beta},$$

  where y is a continuous positive random variable ( y > 0).
- E(y) = αβ and var(y) = $\alpha\beta^2$.
- $f_{gamma}(\eta)$ is obtained by setting y = η, α = θ and β = 1/θ.

  → Choose α = θ and β = 1/θ to make E(η) = 1 [a normalization.]

**End of Digression**

- Note $f(y_i | x_i, u_i) = \frac{e^{-e^{x_i'\beta+\alpha_i}}(e^{x_i'\beta+\alpha_i})^{y_i}}{y_i!} = \frac{e^{-e^{x_i'\beta}u_i}(e^{x_i'\beta}u_i)^{y_i}}{y_i!}$.

- Then,

$$f(y_i | x_i) = \int_0^\infty f(y_i | x_i, u_i) f_{gamma}(u_i) du_i$$
$$= \frac{\Gamma(\theta + y_i)}{\Gamma(y_i + 1)\Gamma(\theta)} r_i^\theta (1 - r_i)^{y_i},$$

where $r_i = \frac{\theta}{e^{x_i'\beta} + \theta}$, $\Gamma(s) = (s-1)\Gamma(s-1)$, and $\Gamma(s) = (s-1)!$ if s is an integer. Thus, when θ is a positive integer,

$$f(y_i | x_i) = \binom{(\theta + y_i) - 1}{\theta - 1} r_i^\theta (1 - r_i)^{(\theta + y_i) - \theta},$$

→ This is the form of the negative binomial distribution.

→ Compound Poisson = Negative binomial distribution!

- Since $(\theta+y_i)$ follows Neg-Bin,

$E(\theta+y_i) = \theta/(1-r_i) = e^{x_i'\beta} + \theta$ → $E(y_i) = e^{x_i'\beta}$.

$var(y_i) = var(\theta+y_i) = \theta(1-r_i)/r_i^2 = e^{x_i'\beta}\left(1 + \frac{1}{\theta}e^{x_i'\beta}\right)$.

- If we allow θ to vary over i and set $\theta_i = \frac{1}{\alpha}e^{x_i'\beta}$, we have

$E(y_i) = e^{x_i'\beta}$; $var(y_i) = (1+\alpha)e^{x_i'\beta}$.

→ This model is called Neg-Bin 1 model (HHG).

- If we set θ = 1/α,

  $$E(y_i) = e^{x_i'\beta};\ \text{var}(y_i) = e^{x_i'\beta}\left(1+\alpha e^{x_i'\beta}\right) = e^{x_i'\beta} + \alpha\left(e^{x_i'\beta}\right)^2.$$

  → This model is called Neg-Bin 2 model (HHG).

- Comment:

  Poisson, Neg-Bin 1 and Neg-Bin 2 assume that $E(y_i) = \exp(x_i'\beta)$. If this mean specification is correct, the Poisson, Neg-Bin 1, Neg-Bin 2 MLE are all consistent as long as the true distribution belongs to the linear exponential family [see Gourieroux, Monfort and Trognon (ECON, 1984).]

(3) Testing Poisson:

- $H_o$: $E(y_i) = \text{var}(y_i) = e^{x_i'\beta}$ (Poisson);

  $H_a$: $E(y_i) = e^{x_i'\beta}$, but $\text{var}(y_i) = e^{x_i'\beta} + \alpha\left(e^{x_i'\beta}\right)^s$ and $\alpha \neq 0$.

  [If s = 1, $H_a$ = Neg-Bin 1. If s = 2, $H_a$ = Neg-Bin 2.]

- For given s, under $H_o$,

  $$E\left(\frac{1}{\mu_i}\{(y_i-\mu_i)^2 - y_i\}\right) = \frac{1}{\mu_i}\left(E\left\{(y_i-\mu_i)^2 - y_i\right\}\right) = 1-1 = 0;$$

  $$\text{var}\left(\frac{1}{\mu_i}\{(y_i-\mu_i)^2 - y_i\}\right) = 2,$$

where $\mu_i = e^{x_i'\beta}$.

- Then, under $H_o$, CLT implies:

$$T_L = \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\mu_i^{s-1}\frac{1}{\mu_i}\left\{\left(y_i - \mu_i\right)^2 - y_i\right\}}{\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\left(\mu_i\right)^{2(s-1)}}\sqrt{2}} \rightarrow \mathrm{N}(0,1).$$

- Since $\mu_i = e^{x_i'\beta}$ is unobservable, we need to use $\hat{\mu}_i = e^{x_i'\hat{\beta}}$, where $\hat{\beta}$ is the Poisson ML estimator. But, we still can show that under $H_o$,

$$\hat{T}_L = \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\hat{\mu}_i^{s-1}\frac{1}{\hat{\mu}_i}\left\{\left(y_i - \hat{\mu}_i\right)^2 - y_i\right\}}{\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\left(\hat{\mu}_i\right)^{2(s-1)}}\sqrt{2}} \rightarrow \mathrm{N}(0,1).$$

- $H_o$ may hold even if the $y_i$ do not follow Poisson. For such cases, use

$$\hat{T}_L' = \frac{\frac{1}{\sqrt{N}}\Sigma_{i=1}^{N}\hat{\mu}_i^{s-1}\frac{1}{\hat{\mu}_i}\left\{\left(y_i - \hat{\mu}_i\right)^2 - y_i\right\}}{\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\left(\hat{\mu}_i\right)^{2(s-1)}}\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\left(\frac{1}{\hat{\mu}_i}\left\{\left(y_i - \hat{\mu}_i\right)^2 - y_i\right\}\right)^2}}.$$

(4) Poisson Model for Panel Data

- Assume $y_{it}$ i.i.d. with Poisson($\lambda_{it}$), where $\lambda_{it} = \exp(x_{it}'\beta+\alpha_i) = \mu_{it}\exp(\alpha_i)$.
- Fixed Effects Model
  - Treat $\alpha_i$ as parameters.
  - Surprisingly, MLE is consistent!
  - $f(y_{it} | x_{it}) = \dfrac{e^{-\lambda_{it}}(\lambda_{it})^{y_{it}}}{y_{it}!}$.
  - $\ln f(y_{it} | x_{it}) = -\lambda_{it} + y_{it}(x_{it}'\beta + \alpha_i) - \ln(y_{it}!)$.
  - $l_{NT}(\beta, \alpha_{1,} \alpha_2, ..., \alpha_N) = \Sigma_{t=1}^{T}\Sigma_{i=1}^{N}\left\{ y_{it}(x_{it}'\beta + \alpha_i) - e^{x_{it}'\beta+\alpha_i} - \ln(y_{it}!)\right\}$.
  - $\dfrac{\partial l_{NT}(\beta, \alpha_1, ..., \alpha_N)}{\partial \alpha_j} = \Sigma_{t=1}^{T}\left\{ y_{it} - e^{\alpha_i}\mu_{it}\right\} = 0$.

$$\rightarrow \alpha_j = \ln\left(\frac{\Sigma_{t=1}^{T} y_{it}}{\Sigma_{t=1}^{T}\mu_{it}}\right).$$

  - Substitute these solutions into $l_{NT}$:

$$l_{NT}^{c}(\beta) = \Sigma_{i=1}^{N}\Sigma_{t=1}^{T} y_{it} \ln p_{it}(\beta),$$

where $p_{it} = \dfrac{\exp(x_{it}'\beta)}{\Sigma_{t=1}^{T}\exp(x_{it}'\beta)}$.

  - The MLE estimator of β based on $l_{NT}^{c}(\beta)$ is consistent even if the true distribution of $y_{it}$ is not Poisson as long long as $E(y_{it}|x_{it},\alpha_i) = \exp(x_{it}'\beta)\exp(\alpha_i)$ [Wooldridge (JEC, 1999)]. For correct covariance matrix of the ML estimator of β, use the robust form $[(H_{NT})^{-1} B_{NT} (H_{NT})^{-1}]$.

- Random Effects Model
  - Assume that the $e^{\alpha_i}$ follow a Gamma Distribution.
  - $$f(y_{i1},...y_{iT} \mid x_{i1},...,x_{iT},u_i) = \prod_{t=1}^{T} f(y_{it} \mid x_{it}, \alpha_i) = \prod_{t=1}^{T} \frac{e^{-\mu_{it}u_i}(\mu_{it}u_i)^{y_{it}}}{y_{it}!}.$$
  - $$f(y_{i1},...,y_{iT} \mid x_{i1},...x_{iT}) = \int_0^\infty f(y_{i1},...,y_{iT} \mid x_{i1},...,x_{iT},u_i) f_{gamma}(u_i) du_i = \frac{\left(\prod_{t=1}^{T} \mu_{it}^{y_{it}}\right)\Gamma\left(\theta + \Sigma_{t=1}^{T} y_{it}\right)}{\left(\Gamma(\theta)\prod_{t=1}^{T} y_{it}!\right)\left(\left(\Sigma_{t=1}^{T}\mu_{it}\right)^{\Sigma_{t=1}^{T} y_{it}}\right)} Q_i^{\theta}(1-Q_i)^{\Sigma_{t=1}^{T} y_{it}}$$

    where $Q_i = \dfrac{\theta}{\theta + \Sigma_{t=1}^{T}\mu_{it}}$.
  - $l_{NT}(\beta,\theta) = \Sigma_{i=1}^{N} f(y_{i1},...,y_{iT} \mid x_{i1},...,x_{iT})$.
  - Can use the Hausman test to determine whether RE or FE is correct.