Microeconometrics: Alternative Distributions

Ethan Kaplan

1 Poisson, Geometric and Negative Binomial Distributions

- Start with a Bernoulli Distribution: an event happens with probability $(p, 1-p)$
  - Mean: $p * 1 + (1-p) * 0 = p$
  - Variance: $p(1-p)^2 + (1-p)(0-p)^2 = p(1-p)$; SD: $SD = \sqrt{p(1-p)}$
- Generalized Binomial:
  - Random Variable $= X$
  - With n trials: $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
  - Mean: np
  - Variance: $np(1-p)$; SD: $\sqrt{np(1-p)}$

- Poisson: Count variable ("law of rare events")
  - $\lim_{n\to\infty}\binom{n}{k}p^k(1-p)^{n-k}=\frac{e^{-\mu}\mu^k}{k!}$ with fixed $np=\mu$ ($\implies p\to 0$)
  - Mean: $E(N)=\sum_{k=0}^{\infty}kP(N=k)=\sum_{k=1}^{\infty}ke^{-\mu}\frac{\mu^k}{k!}$ but $\sum_{k=1}^{\infty}ke^{-\mu}\frac{\mu^k}{k!}=\mu$
  - Variance: $V(N)=\mu$; SD: $SD(N)=\sqrt{\mu}$
- Geometric Distribution: Waiting Time (First Event)
  - Realization Probability: $P(T=n)=(1-p)^{n-1}p$
  - Mean: $E(T)=\sum_{n=1}^{\infty}nP(T=n)=\frac{1}{p}$
  - Variance: $V(T)=\frac{(1-p)}{p^2}$; SD: $SD(T)=\frac{\sqrt{1-p}}{p}$

- Negative Binomial: (Generalized Geometric Dist.)
  - $P(T_r = t) =$Joint prob. of $r-1$ successes in first $t-1$ trials and trial t success
  - Therefore: $P(T_r = t) = \binom{t-1}{r-1} p^{r-1} (1-p)^{t-r} p$ but $\binom{t-1}{r-1} p^{r-1} (1-p)^{t-r} p = \binom{t-1}{r-1} p^{r} (1-p)^{t-r}$
  - Mean: $E(T_r) = \frac{r}{p}$
  - Variance: $V(T_r) = \frac{r(1-p)}{p^2}$; SD: $SD(T_r) = \frac{\sqrt{r(1-p)}}{p}$

2 Estimation of Poisson and Negative Binomial Distributions

- Why Poisson or Negative Binomial?
  - Only positive numbers (but can use lognormal distribution)
  - Can Improve on SEs
    * In general, binomial distribution has $\lim_{p \to \infty} V(\hat{p}) = 0$
    * Case of $p = 0$ or $p = 1$, can reject based upon one observation.
    * Can improve SEs by using binomial structure for all binomial family derivatives
  - Lowers under-estimation of zeros (allows for right skewness)

- Why not Poisson or Negative Binomial
  - Overdispersion: SEs too low if not bernoulli and with essentially zero probability
    * In Poisson, $V_{Data}(N) > E(N) = V(N)$
  - Still too few zeros?
  - Especially true with Poisson

3 Poisson Regression

- What is being estimated?

  - The determinants of the expected number of counts
  - $\frac{e^{-\mu_i}\mu_i^k}{k!}$ where $\ln \mu_i = X_i'\beta$
  - $\frac{e^{-e^{X_i'\beta}}\left(e^{X_i'\beta}\right)^k}{k!}$
  - $E\left(Y_i|X_i\right) = V\left(Y_i|X_i\right) = \mu_i$

- Estimation Methods

  - NLS: $\min_\beta \sum_{i=1}^{N}\left[y_i - \frac{e^{-e^{X_i'\beta}}\left(e^{X_i'\beta}\right)^{y_i}}{y_i!}\right]^2$
  - Maximum Likelihood:

* $\min_\beta \ln L = \min_\beta \sum_{i=1}^{N} \ln \frac{e^{-e^{X_i'\beta}}\left(e^{X_i'\beta}\right)^{y_i}}{y_i!}$

* $= \min_\beta \sum_{i=1}^{N} \left[y_i X_i'\beta - e^{X_i'\beta} - \ln y_i!\right]$

* Moment Conditions

  · $\sum_{i=1}^{N} \left[k_i - e^{X_i'\beta}\right] X_i' = 0$

* Variance of Estimate: $V\left(\hat{\beta}_P\right) = \left(\sum_{i=1}^{N} \mu_i x_i x_i'\right)^{-1}$

- Test for overdispersion:

  – Model variance: $V\left(y_i | X_i\right) = \mu_i + \alpha g\left(\mu_i\right)$

  – $g\left(\mu_i\right)$ usually $= \mu_i$ or μ_i^2

- Run OLS

$$\frac{(y_i - \hat{\mu}_i)^2 - y_i}{\hat{\mu}_i} = \alpha \frac{g(\hat{\mu}_i)}{\hat{\mu}_i} + u_i$$

- Test $\alpha = 0$

- Robust Standard Errors (overdispersion correction)
  - Define $\hat{\alpha} = (n-k)^{-1} \sum_{i=1}^{N} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$
  - Then $V(y_i|X_i) = \hat{\alpha} E(y_i|X_i) = \hat{\alpha} e^{x_i' \beta}$

- Compute marginal effects: $\frac{\partial E(y_i|X_i)}{\partial X_i} = \beta_i e^{x_i' \beta}$
  - Note that marginal effects depend upon X_i

4 Negative Binomial Regression

- Now suppose that μ_i is random as in specification error
  - $\ln \mu = x_i'\beta + \epsilon_i = \ln \lambda_i + \ln u_i$

- The conditional distribution of f then is:
  - $f(y_i|x_i, \lambda_i) = \frac{e^{-\mu_i\lambda_i}(\lambda_i\mu_i)^{y_i}}{y_i!}$

- Integrating, we get:
  - $f(y_i|x_i) = \int_0^\infty \frac{e^{-\mu_i\lambda_i}(\lambda_i\mu_i)^{y_i}}{y_i!} g(u_i)\, du_i$

- We now assume a distribution for g :
  - $g(\lambda_i) = \frac{\theta^\theta}{\Gamma(\theta)} e^{-\theta\mu_i\mu_i^{\theta-1}}$

- We can now solve for f :
  - $f\left(y_i|x_i\right)=\frac{\Gamma(\theta+y_i)}{\Gamma(y_i+1)\Gamma(\theta)}r_i^{y_i}\left(1-r_i\right)^{\theta}$
  - where $r_i=\frac{\lambda_i}{\lambda_i+\theta}$
- This can be estimated with MLE or NLLS
- The mean is $E\left(y_i|x_i\right)=\lambda_i$
- The variance is $V\left(y_i|x_i\right)=\lambda_i\left(1+\left(\frac{1}{\theta}\right)\lambda_i\right)$
  - Remember that θ is estimated so this decouples the mean and variance
  - The Poisson Model is the restriction of $\frac{1}{\theta}=0$
  - This can be tested using a Wald or LRT Test

- Regression (OLS, Poisson, Negative Binomial) of numbers of donors to presidential, senate and house of representative campaigns during 2000 electoral cycle on population and presence of Fox News in 2000:
  - Spec. A includes just a constant, log population, and a dummy for Fox News
  - Spec. B adds number of donors in 1996:

	Population	*Fox News*
OLS A	124.5 (68.8)	24.8 (1.26)
Poisson A	2.2 (1092.1)	40.2 (180.7)
Neg Bin. A	30.9 (29.9)	16.9 (6.8)
OLS B	−3.7 (8.36)	1.1 (0.27)
Poisson B	1.9 (776.0)	20.9 (100.9)
Neg. Bin. B	8.8 (19.5)	5.7 (4.6)

- Regression of numbers of donors to presidential, senate and house of representative campaigns during 2000 electoral cycle on population and presence of Fox News in 2000 (robust standard errors):
  - Spec. A includes just a constant, log population, and a dummy for Fox News
  - Spec. B adds number of donors in 1996:

	Population R	*Fox News R*
OLS A	124.5 (1.9)	24.8 (2.0)
Poisson A	2.2 (6.8)	40.2 (2.1)
Neg Bin. A	30.9 (13.5)	16.9 (4.5)
OLS B	−3.7 (1.0)	1.1 (0.2)
Poisson B	1.9 (6.4)	20.9 (1.6)
Neg. Bin. B	8.8 (10.1)	5.7 (4.6)