Microeconometrics
Count Regression Models

Pedro Portugal

NOVA School of Business and Economics

Spring 2025

Counting

Count Variables

- Number of patents
- Entry and exit of firms
- Number of visits to a recreational park
- Number of published papers in scientific journals
- Number of fixed-term contracts
- Number of kids
- Number of kids with a college degree
- "Teenagers" in the workforce

Siméo-Dennis Poisson (1781-1840)

Poisson Distribution

Poisson MLE

- The natural stochastic model for counts is a Poisson point process for the occurrence of the event of interest
- This implies a **Poisson distribution** for the number of occurrences of the event, with probability mass function

$$Pr[Y = y] = \frac{e^{-\mu}\mu^y}{y!}, \quad y = 0, 1, 2, ...$$

where μ is the intensity or rate parameter. The first two moments are

$$E[Y] = \mu$$
$$V[Y] = \mu$$

- The **Poisson regression model** is derived from the Poisson distribution by parameterizing the relation between μ and covariates $\mathbf{x}$

Poisson MLE

- The standard assumption is to use the exponential mean parametrization

$$\mu_i = \exp(\mathbf{x}_i'\beta), \qquad i = 1, \ldots, N$$

- Then the probability mass function can be written as

$$Pr[Y = y] = \frac{\exp(-\exp(\mathbf{x}'\beta))\exp(\mathbf{x}'\beta)^y}{y!}$$

and $E[Y|\mathbf{x}] = \exp(\mathbf{x}'\beta)$ and $Var[Y|\mathbf{x}] = \exp(\mathbf{x}'\beta)$

- The Poisson regression is intrinsically heteroskedastic

Poisson MLE

The log-likelihood function is

$$\ln L_N(\beta) = \sum_{i=1}^{N} \{ y_i \mathbf{x}_i' \beta - \exp(\mathbf{x}_i' \beta) - \ln y_i! \}$$

The **Poisson MLE**, $\widehat{\beta}_P$, is the solution to K nonlinear equations corresponding to the first-order condition for maximum likelihood,

$$\sum_{i=1}^{N} (y_i - \exp(\mathbf{x}_i' \beta)) \mathbf{x}_i = \mathbf{0}$$

Poisson MLE

By asymptotic theory the estimator $\widehat{\beta}_P$ is consistently estimated for β and asymptotically normal with sample covariance matrix

$$V[\widehat{\beta}_P] = \left(\sum_{i=1}^{N} \mu_i \mathbf{x}_i \mathbf{x}_i' \right)^{-1}$$

Therefore,

$$\sqrt{N}(\widehat{\beta} - \beta) \xrightarrow{d} \mathcal{N}\left[0, p\lim \left(\frac{1}{N} \sum \exp(\mathbf{x}_i'\beta)\mathbf{x}_i\mathbf{x}_i' \right)^{-1}\right]$$

Poisson MLE

For linear models, with $E[y|\mathbf{x}] = \mathbf{x}'\beta$, the coefficients β are interpreted as the effect of a one-unit change in regressors on the conditional mean

For any model with exponential conditional mean, differentiation yields

$$\frac{\partial E[y|\mathbf{x}]}{\partial x_j} = \beta_j \exp(\mathbf{x}'\beta)$$

where the scalar x_j denotes the jth regressor

Table 1. ESTIMATED COUNT REGRESSION MODELS FOR PLANT OPENINGS (n=1830)

	POISSON MODEL		NEGATIVE BINOMIAL MODEL		FIXED EFFECTS POISSON MODEL	
	(1)	(2)	(3)	(4)	(5)	(6)
Pwage(t)	-0,715	-0,772	-0,730	-0,794	0,157	-2,804
	(-8.967)	(-9.588)	(-5.607)	(-6.503)	(0.897)	(11.733)
Pminwage(t-1)	-0,561	-0,216	-0,384	-0,092	-1,365	0,125
	(-5.630)	(-2.091)	(-2.890)	(-0.723)	(-10.105)	(0.735)
Nplants(t-1)	1,019	1,033	1,004	1,029	0,602	-0,043
	(77.658)	(78.388)	(40.268)	(44.883)	(10.810)	(-0.657)
Avesize(t-1)	0,219	0,241	0,165	0,186	0,364	0,174
	(13.041)	(14.088)	(5.803)	(6.813)	(6.159)	(2.072)
Conc4(t-1)	0,001	0,002	0,001	0,002	-0,010	-0,003
	(1.283)	(2.569)	(0.468)	(1.593)	(-5.381)	(-1.074)
Age(t-1)	-0,103	-0,093	-0,078	-0,074	-0,040	0,003
	(-29.943)	(-26.151)	(-14.418)	(-14.123)	(-5.298)	(0.357)
GDP growth(t-1)	8,910		8,370		10,105	
	(25.870)		(14.168)		(27.963)	
Constant	4,784	5,432	4,895	5,665		
	(6.323)	(7.116)	(3.998)	(4.917)		
Year Dummies	No	Yes	No	Yes	No	Yes
Sigma			0,086	0,061		
			(10.863)	(9.695)		
Log-likelihood	-4386,683	-4221,440	-4111,95	-4039,015	-3285,727	-3072,623

Table 2. ESTIMATED COUNT REGRESSION MODELS FOR PLANT CLOSINGS (n=1830)

	POISSON MODEL		NEGATIVE BINOMIAL MODEL		FIXED EFFECTS POISSON MODEL	
	(1)	(2)	(3)	(4)	(5)	(6)
Wage(t-1)	0,274	0,245	0,331	0,287	1,244	0,717
	(3.757)	(3.339)	(3.279)	(3.138)	(7.081)	(3.016)
Pminwage(t-1)	0,122	0,118	0,187	0,156	0,769	0,614
	(1.114)	(1.048)	(1.426)	(1.276)	(3.997)	(3.079)
Nplants(t-1)	1,025	1,029	1,022	1,024	1,329	1,450
	(83.405)	(83.318)	(62.245)	(73.113)	(16.491)	(18.571)
Avesize(t-1)	-0,069	-0,069	-0,079	-0,075	-0,297	-0,359
	(-4.302)	(-4.241)	(-3.760)	(-4.016)	(-3.817)	(-4.228)
Conc4(t-1)	0,004	0,004	0,003	0,004	0,007	0,009
	(4.895)	(5.156)	(3.501)	(4.303)	(3.212)	(3.767)
Age(t-1)	-0,010	-0,010	-0,011	-0,011	0,019	0,009
	(-3.495)	(-3.342)	(-2.815)	(3.232)	(1.549)	(0.698)
GDP growth(t)	-1,798		-1,827		-2,084	
	(-6.214)		(-4.745)		(-6.147)	
Constant	-4,817	-4,635	-5,313	-4,996		
	(-6.842)	(-6.573)	(-5.423)	(-5.682)		
Year Dummies	No	Yes	No	Yes	No	Yes
Sigma			0,017	0,009		
			(5.471)	(4.026)		
Log-likelihood	-3937,587	-3912,193	-3903,252	-3888,912	-3001,294	-2974,964

Mixed Poisson Model

Now let $\lambda(\mathbf{x}, \beta) = \exp(\mathbf{x}'\beta + \varepsilon)$, with $E(\exp(\varepsilon)) = 1$ and $V(\exp(\varepsilon)) = \sigma^2$

Then,

$$P(Y = y|\mathbf{x}, \varepsilon) = \frac{\exp[-\exp(\mathbf{x}'\beta + \varepsilon)]\exp(\mathbf{x}'\beta + \varepsilon)^y}{y!}$$

And

$$P(Y = y|\mathbf{x}) = \int \left[\frac{\exp[-\exp(\mathbf{x}'\beta + \varepsilon)]\exp(\mathbf{x}'\beta + \varepsilon)^y}{y!}\right] g(\varepsilon) d\varepsilon$$

Overdispersion

- One fundamental problem with the Poisson regression model is that the distribution is parameterized in terms of μ so that all moments of y are a function of μ

- In many applications a Poisson density predicts the probability of a zero count to be considerably less than is actually observed in the sample – **excess zeros** problem, as there are more zeros in the data than the Poisson predicts

- For count data the variance usually exceeds the mean, a feature called **overdispersion**: $Var(y|\mathbf{x}) > E(y|\mathbf{x})$, and the Poisson model implies equality of the variance and the mean (equidispersion)

Overdispersion Test

Consider again $\lambda(\mathbf{x}, \beta) = \exp(\mathbf{x}'\beta + \varepsilon)$

The null hypothesis of no overdispersion is: $H_0 : \sigma^2 = 0$
We can write,

$$E\left[\left(Y - \exp(\mathbf{x}'\beta)\right)^2 - Y\right] = 0$$

with sample version given by

$$\frac{1}{N}\sum\left[\left(y - \exp(\mathbf{x}'\beta)\right)^2 - y\right] = 0$$

Then the t-statistic under the null hypothesis:

$$\widehat{T} = \sum \frac{\left(y - \exp(\mathbf{x}'\widehat{\beta})\right)^2 - y}{\sqrt{2\sum \exp(2\mathbf{x}'\widehat{\beta})}} \sim \mathcal{N}(0,1)$$

Pseudo Maximum Likelihood Estimation

$$E(Y|\mathbf{x}) = E_{\varepsilon}\left[\exp(\mathbf{x}'\beta + \varepsilon)\right] = \exp(\mathbf{x}'\beta)$$

$$V(Y|\mathbf{x}) = \left[1 + \sigma^2 \exp(\mathbf{x}'\beta)\right] \exp(\mathbf{x}'\beta)$$

The PML estimator $\widehat{\beta}_{PML}$ is the solution to:

$$\sum_{i=1}^{N} \left[y_i - \exp(\mathbf{x}_i'\widehat{\beta}_{PML})\right]\mathbf{x}_i = \mathbf{0}$$

And is asymptotically normally distributed as follows:

$$\sqrt{N}(\widehat{\beta}_{PML} - \beta) \xrightarrow{d} \mathcal{N}\left[0, p\lim A^{-1}BA^{-1}\right]$$

where $A = \frac{1}{N}\sum_i \widehat{\lambda}\mathbf{x}_i\mathbf{x}_i'$ and $B = \frac{1}{N}\sum_i (y_i - \widehat{\lambda}_{PML})^2\mathbf{x}_i\mathbf{x}_i'$

OLS and Poisson Estimates of a Fertility Equation

Dependent Variable: *children*		
Independent Variable	Linear (OLS)	Exponential (Poisson QMLE)
educ	−.0644 (.0063)	−.0217 (.0025)
age	.272 (.017)	.337 (.009)
age^2	−.0019 (.0003)	−.0041 (.0001)
evermarr	.682 (.052)	.315 (.021)
urban	−.228 (.046)	−.086 (.019)
electric	−.262 (.076)	−.121 (.034)
tv	−.250 (.090)	−.145 (.041)
constant	−3.394 (.245)	−5.375 (.141)
Log-likelihood value	—	−6,497.060
R-squared	.590	.598
$\hat{\sigma}$	1.424	.867

Negative Binomial Model

- Suppose the distribution of a random count y is Poisson, conditional on the parameter λ, so that $f(y|\lambda) = \exp(-\lambda)\lambda^y/y!$
- Suppose now that the parameter λ is random. In particular, let $\lambda = \mu\nu$, where μ is a deterministic function of $\mathbf{x}$, for example $\exp(\mathbf{x}'\beta)$, and $\nu > 0$ is iid with density $g(\nu|\alpha)$
- Note that $E[\lambda|\mu] = \mu$ if $E[\nu] = 1$
- If $f(y|\lambda)$ is the Poisson density and $g(\nu)$ is the gamma density with $E[\nu] = 1$ and $V[\nu] = 1/\delta$, $\delta > 0$, the **negative binomial** as a mixture model is given by:

$$h[y|\mu,\delta] = \frac{\Gamma(\alpha^{-1}+y)}{\Gamma(\alpha^{-1})\Gamma(y+1)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu}\right)^{\alpha^{-1}}\left(\frac{\mu}{\mu+\alpha^{-1}}\right)^{y}$$

where $\alpha = 1/\delta$ and $\Gamma(.)$ denotes the gamma integral

Negative Binomial Model

The first two moments of the negative binomial distribution are

$$E[y|\mu, \alpha] = \mu$$
$$V[y|\mu, \alpha] = \mu(1 + \alpha\mu)$$

A standard variant of the negative binomial used in regression applications specifies $\mu_i = \exp(\mathbf{x}_i'\beta)$ and lets α to be a parameter to be estimated, in which case the conditional variance function is quadratic in the mean

Negative Binomial Model

Let $y_i|\mathbf{x}_i, c_i \sim Poisson(c_i\mu(\mathbf{x}_i'\beta))$ and let $Var(c_i) = \eta^2$

Then the log-likelihood function is given by:

$$\ln L(\beta, \eta^2) = \sum_{i=1}^{N} \left[\eta^{-2} \ln\left(\frac{\eta^{-2}}{\eta^{-2} + \exp(\mathbf{x}_i'\beta)}\right) + \right.$$
$$\left. + y_i \ln\left(\frac{\exp(\mathbf{x}_i'\beta)}{\eta^{-2} + \exp(\mathbf{x}_i'\beta)}\right) + \ln\left(\frac{\Gamma(y_i + \eta^{-2})}{\Gamma(\eta^{-2})\Gamma(y_i + 1)}\right)\right]$$

Zero-Inflated Poisson Models

$$P(Y=0|\mathbf{x}) = \pi + (1-\pi)\exp(-\lambda)$$

$$P(Y=y|\mathbf{x}) = (1-\pi)\frac{\exp(-\lambda)\lambda^y}{y!}, \qquad ,y=1,2,...$$

$$E(Y|\mathbf{x}) = (1-\pi)\lambda \qquad\qquad V(Y|\mathbf{x}) = (1+\pi\lambda)E(Y|\mathbf{x})$$

$$\ln L(\lambda) = (1-s)\ln\left[\pi+(1-\pi)\exp(-\lambda)\right] + s\left[\ln(1-\pi)-\lambda+y\ln(\lambda)-\ln(y!)\right]$$

$$s = \begin{cases} 0 & y=0 \\ 1 & y>0 \end{cases}$$

Hurdle Model

$$P(Y = 0) = \exp(-\lambda^*)$$

$$P(Y = y) = \frac{(1 - \exp(-\lambda^*))}{(1 - \exp(-\lambda^\dagger))} \frac{\exp(-\lambda^\dagger)(\lambda^\dagger)^y}{y!}, y > 0$$

$$\ln L(\lambda^*, \lambda^\dagger) = (1-s)(-\lambda^*) + s \ln[1 - \exp(-\lambda^*)] + s \ln \left[\frac{\exp(-\lambda^\dagger)(\lambda^\dagger)^y}{(1 - \exp(-\lambda^\dagger))y!} \right]$$

$$s = \begin{cases} 0 & y = 0 \\ 1 & y > 0 \end{cases}$$

Truncated Models

$$P[Y = y|\mathbf{x}, y > 0] = \frac{P[Y = y|\mathbf{x}]}{1 - P[Y = 0|\mathbf{x}]}, y = 1, 2, ...$$

$$E[Y|\mathbf{x}, y > 0] = \frac{E[Y|\mathbf{x}]}{1 - P[Y = 0|\mathbf{x}]}$$

In-situ sampling

$$P^{is}(Y = y|\mathbf{x}) = \frac{P[Y = y|\mathbf{x}]y}{E[Y|\mathbf{x}]}$$

$$P[Y = y|\mathbf{x}] = \frac{\exp[\lambda(\mathbf{x}, \beta)]\lambda(\mathbf{x}, \beta)^y}{y!}$$

$$P^{is}(Y = y|\mathbf{x}) = \frac{\exp[\lambda(\mathbf{x}, \beta)]\lambda(\mathbf{x}, \beta)^{y-1}}{(y-1)!}$$

Semi-parametric estimation: Latent-class models

$$P[Y = y|\mathbf{x}] = \sum_{q=1}^{Q} \frac{\exp[-\exp(\mathbf{x}'\beta + \alpha_q)]\exp(\mathbf{x}'\beta + \alpha_q)}{y!}\pi(q)$$

Proportions

Let $y_i \sim P(\mu_i)$ where $\mu_i = \exp(\mathbf{x}_i'\beta)N_i$

We can write

$$\ln(\mu_i) = \mathbf{x}_i'\beta + \ln(N_i)$$

Hirings

Variables	no random effects		with random effects	
	Coef	Std	Coef	Std
year 1988 (dummy)	-0.057	0.009	-0.036	0.010
year 1989 (dummy)	-0.065	0.009	-0.043	0.010
firm size	-0.099	0.003	-0.106	0.004
firm hiring rate	-0.474	0.021	-0.455	0.027
market concentration (Herfindahl index)	-0.970	0.045	-0.977	0.062
1=±			0.228	
Log-likelihood	-96177		-92433	
N	99608		99608	

Table 5: SHARE OF TEENAGERS HIRED BY CONTINUING FIRMS, POISSON REGRESSION MODEL. Notes: Controlling for the industry (7 dummies), and public or foreign ownership of the company. Firm size is in logs.

New Firms

Variables	Coef	Std
year 1988 (dummy)	-0.042	0.018
year 1989 (dummy)	-0.041	0.018
firm size	-0.113	0.005
market concentration (Herfindahl index)	-2.880	0.174
N	38138	
Log-likelihood	-29936	

Table 7: Share of teenagers hired by new firms, Poisson regression model. Notes: Controlling for the industry (7 dummies), and public or foreign ownership of the company. Firm size is in logs.

Plant Closures

Variables	Coef	Std
year 1988 (dummy)	0.050	0.023
year 1989 (dummy)	0.025	0.023
firm size	-0.106	0.007
market concentration (Herfindahl index)	-2.941	0.257
N	19203	
Log-likelihood	-15023	

Table 8: SHARE OF TEENAGERS DISMISSED FROM FIRMS CLOSING DOWN, POISSON REGRESSION MODEL. Notes: Controlling for the industry (7 dummies), and public or foreign ownership of the company. Firm size is in logs.

Separations

Variables	no random effects		with random effects	
	Coef	Std	Coef	Std
year 1988 (dummy)	-0.132	0.010	-0.150	0.010
year 1989 (dummy)	-0.111	0.009	-0.140	0.010
firm size	-0.112	0.003	-0.113	0.005
firm hiring rate	-0.123	0.016	-0.121	0.020
market concentration (Herfindahl index)	-0.184	0.054	-1.140	0.075
1=±			0.379	
Log-likelihood	-97770		-93045	
N	125397		125397	

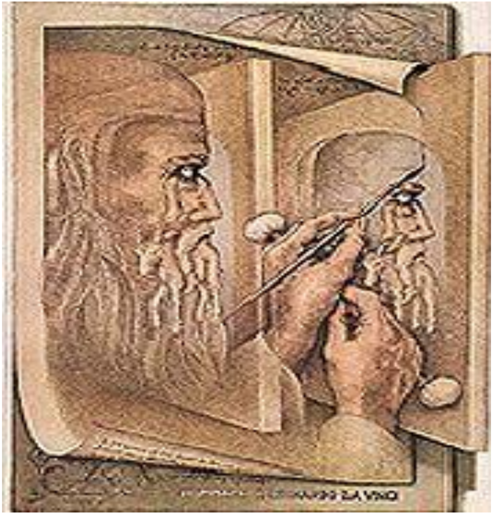
Table 6: SHARE OF TEENAGERS SEPARATED FROM CONTINUING FIRMS, POISSON REGRESSION MODEL. Notes: Controlling for the industry (7 dummies), and public or foreign ownership of the company. Firm size is in logs.

Binomial Regression

$$E(y_i) = N_i p(\mathbf{x}_i'\beta)$$

$$\ln L(\beta) = y_i \ln p(\mathbf{x}_i'\beta) + (N_i - y_i) \ln(1 - p(\mathbf{x}_i'\beta))$$

Models of Count Data



	Beta-Binomial Regression Model		
	(A)	(B)	(C)
Log Wages	0.226*	0.178*	0.189*
	(0.048)	(0.045)	(0.046)
Wage dispersion	-0.004***	-0.002	-0.003
	(0.002)	(0.002)	(0.002)
Fringe benefits	0.065	0.041	0.052
	(0.061)	(0.061)	(0.062)
Training	-0.044	0.399*	-0.426
	(0.182)	(0.143)	(0.369)
Attrition	4.951*	4.440**	4.303**
	(1.877)	(1.845)	(1.850)
Quits	1.187*	0.962*	0.995*
	(0.318)	(0.331)	(0.331)
Tenure 1		-1.340*	-1.399*
		(0.096)	(0.096)
Tenure 2		1.180*	1.210*
		(0.146)	(0.146)
Training * Tenure 1			1.650*
			(0.484)
Firm age			
less than 2 years	-0.048	0.117	0.116
	(0.112)	(0.114)	(0.114)
2 - 5 years	0.058	-0.037	-0.058
	(0.079)	(0.081)	(0.081)
Firm size			
500-999 employees	0.006	-0.003	0.003
	(0.074)	(0.074)	(0.074)
1000 or more employees	-0.051	-0.056	-0.041
	(0.080)	(0.080)	(0.080)
Worker age			
25-44 years	-0.088	-0.446**	-0.452**
	(0.197)	(0.195)	(0.195)
45-64 years	-1.085*	-1.302*	-1.299*
	(0.167)	(0.188)	(0.188)
Constant	-3.744*	-2.937*	-3.065*
	(0.345)	(0.331)	(0.345)
Alpha	1.697	1.618	1.615
	(0.031)	(0.030)	(0.029)
Qualification Level	Yes	Yes	Yes
Log likelihood	-19857.0	-19733.9	-19729.1
Observations	7601	7601	7601