Economics Bulletin

Volume 30, Issue 4

# Hermite regression analysis of multi-modal count data

David E Giles
*University of Victoria, Canada*

## Abstract

We discuss the modeling of count data whose empirical distribution is both multi-modal and over-dispersed, and propose the Hermite distribution with covariates introduced through the conditional mean. The model is readily estimated by maximum likelihood, and nests the Poisson model as a special case. The Hermite regression model is applied to data for the number of banking and currency crises in IMF-member countries, and is found to out-perform the Poisson and negative binomial models.

I am grateful to participants in a seminar at Simon Fraser University, at the MODSIM07 Congress at the University of Canterbury, and a referee for their helpful comments on preliminary versions of this work.

**Citation:** David E Giles, (2010) "Hermite regression analysis of multi-modal count data", *Economics Bulletin*, Vol. 30 no.4 pp. 2936-2945.

**Submitted:** Aug 19 2010. **Published:** November 09, 2010.

## 1. Introduction

Typically, models for count data (*i.e.*, data that take only non-negative integer values) are based on distributions that do not allow for multi-modality. Obvious examples are the Poisson and negative binomial distributions. This seriously limits the usefulness of such models. We discuss a way of broadening the class of discrete distributions that are used in this field by adopting the Hermite distribution proposed by Kemp and Kemp (1965). Apart from Giles (2007), this distribution has not been used in the econometrics literature to date. Further, it appears that when it has been used in other fields, no consideration has been given to introducing covariates into the model, as is done with the conventional Poisson and negative binomial regression models.

Two particularly appealing features of the Hermite distribution are that it is capable of modeling multi-modal count data without any modification, and simultaneously it can account for over-dispersion in the sample. Some background discussion is proved in the next section. Section 3 presents the Hermite regression model, with the covariates introduced in a more satisfactory manner than in Giles (2007). An application involving currency and banking crises is provided in section 4, and section 5 concludes. Our results suggest that the Hermite distribution, parameterized to incorporate covariates, offers considerable potential for modeling discrete economic data.

## 2. Modeling Count Data

Consider the Poisson distribution, with p.m.f.:

$$\Pr[Y = y] = \exp(-\lambda)\lambda^{y} / y! \quad ; \quad y = 0, 1, 2, \dots \tag{1}$$

This distribution is "equi-dispersed" as $\lambda$ ( > 0) is both its mean and variance. In contrast, many data are "over-dispersed", in that their variance exceeds their mean, so this reduces the usefulness of the Poisson distribution. Allowing the variance to be modeled in turn by a gamma distribution, leads to the familiar negative binomial distribution, which can capture over-dispersion in the data.

In linear regression we "explain" the (conditional) mean of the dependent variable as a function of parameters and covariates, so it is natural to introduce covariates into the model by assigning:

$$\lambda = \exp(x'\beta) \quad , \tag{2}$$

so that $\lambda > 0$. Maximum likelihood estimation of the parameters is then straightforward, as the log-likelihood function is strictly concave (as it is also for the negative binomial model). In the ensuing discussion, it is important to recognize that the Poisson model, and standard variants that allow for over-dispersion, cannot describe multi-modal data.[1]

[1] More correctly, if $\lambda$ is *integer*, then the Poisson distribution has equally high modes at $\lambda$ and $(\lambda - 1)$, but never at non-adjacent values. If $\lambda$ is non-integer, the single mode occurs at $[\lambda]$, the integer part of $\lambda$.

The *zero-inflated* Poisson (ZIP) regression model (*e.g.*, Mullahy 1986) is a modified Poisson regression model that allows for an excess of zero counts.[2] This phenomenon is widely encountered in practice and it may (or may not) result in an empirical distribution that is bimodal and/or over-dispersed. The situation where the data exhibit an excess of counts at *several* integer values has received little attention. Santos Silva and Covas (2000) and Hellström (2006) used modified double-hurdle models for this problem. Melkersson and Rooth (1999) considered an extended ZIP model when there was count inflation at the values zero and two; and Giles (2007) discussed a full generalization of the ZIP model to allow for count-inflation at multiple values. We now consider an alternative distribution for modeling count data that allows for both multi-modality and over-dispersion, namely the Hermite distribution (Kemp and Kemp 1965).

## 3. A More General Model

The Hermite distribution is a generalized Poisson distribution, taking its name from the fact that its probabilities and factorial moments can be expressed in terms of the coefficients of (modified) Hermite polynomials. The bivariate Poisson and the Poisson-binomial distributions are special cases of the Hermite distribution. An Hermite variate also arises as the sum of an ordinary Poisson variate and an independent Poisson 'doublet' variate; and the distribution of the sum of a finite number of *correlated* Poisson variates is also Hermite (McKendrick, 1926 and Maritz 1952).

One convenient expression for the p.m.f. for the Hermite distribution is:

$$p_r = \Pr(Y = r) = \exp\{-(\alpha+\gamma)\} \sum_{l=0}^{[r/2]} \frac{\alpha^{r-2l}\gamma^l}{(r-2l)!l!} \quad ; \quad r = 0, 1, 2, \ldots; \ \alpha, \gamma > 0 . \qquad (3)$$

The mean and variance are $(\alpha + 2\gamma)$ and $(\alpha + 4\gamma)$ respectively, and $[x]$ denotes the integer part of $x$. Unless $\gamma = 0$ (implying a Poisson distribution) we have over-dispersion. In all generalized Poisson distributions the probabilities follow some recursion scheme. In the case of the Hermite distribution:

$$\begin{aligned} p_0 &= \exp\{-(\alpha+\gamma)\} \\ p_1 &= (\alpha\, p_0) \\ p_{r+1} &= (\alpha\, p_r + 2\gamma\, p_{r-1})/(r+1) \quad ; \quad r = 1, 2, \ldots \end{aligned} \qquad (4)$$

Apart from Giles (2007) we are not aware of any previous discussion of introducing covariates into models based on the Hermite distribution. Here, we achieve this in a new and natural way that follows the approach adopted by Ferrari and Cribari-Neto (2004) for regression based on the beta distribution. The parameterization in (3) and (4) can be
modified by assigning[3]

[2] The negative binomial regression model may also be extended to allow for zero-inflation of the data in a corresponding and straightforward manner.

[3] Alternative parameterizations were used by Giles (2007) in his empirical illustrations, but these were less intuitive than the one suggested here, and are not recommended.

$$\mu = \alpha + 2\gamma \;;\; \phi = 2\gamma \quad , \tag{5}$$

so

$$\alpha = \mu - \phi \;;\; \gamma = \phi / 2 . \tag{6}$$

Then, (3) and (4) become

$$p_r = \Pr(Y = r) = \exp\{-(\mu - \phi/2)\} \sum_{l=0}^{[r/2]} \frac{(\mu - \phi)^{r-2l} (\phi/2)^l}{(r-2l)!l!} \;;\; r = 0, 1, 2, \ldots \tag{7}$$

and

$$\begin{aligned} p_0 &= \exp\{-(\mu - \phi/2)\} \\ p_1 &= (\mu - \phi) p_0 \\ p_{r+1} &= \{(\mu - \phi) p_r + \phi\, p_{r-1})\} / (r+1) \;; \quad r = 1, 2, \ldots \end{aligned} \tag{8}$$

As $E(Y) = \mu$ and $Var.(Y) = \mu + \phi$, for a given value for the mean of the distribution, $\phi$ is a dispersion parameter – the variance increases with $\phi$. Figures 1 and 2 show the Hermite p.m.f. for particular choices of $\mu$ and $\phi$, and illustrate its ability to exhibit multi-modality.

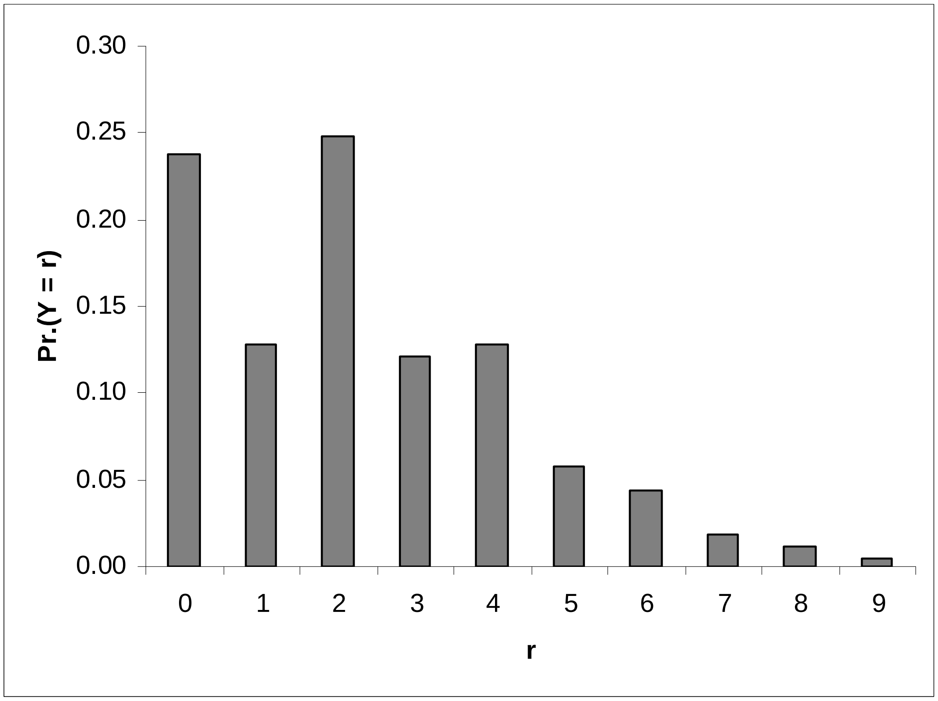


**Figure 1: Probability mass function for Hermite distribution; $\mu$ = 3.1, $\phi$ = 3.0**

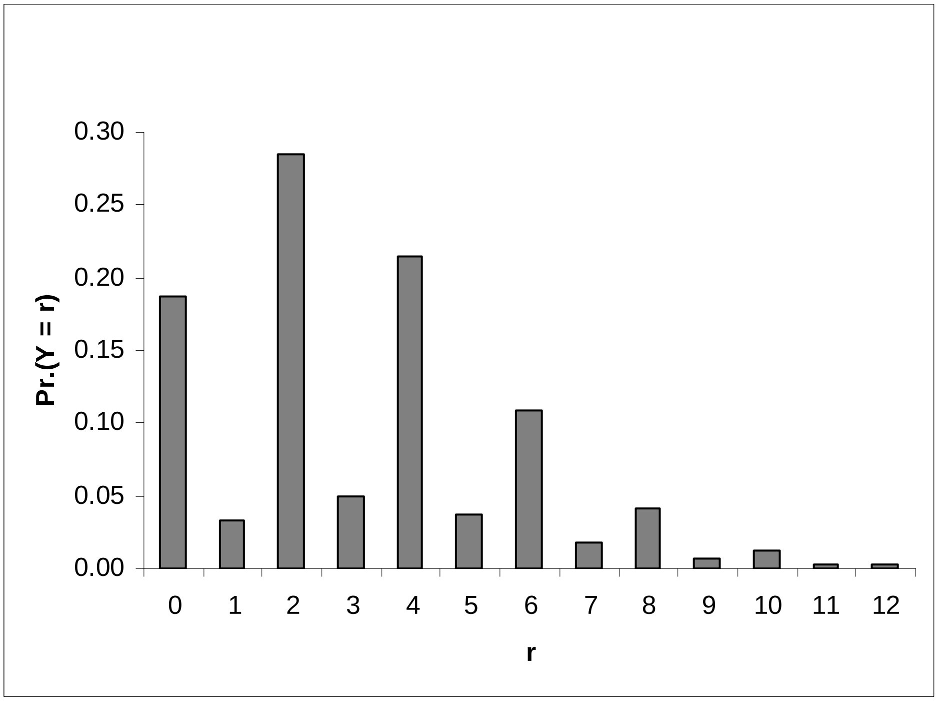


**Figure 2: Probability mass function for Hermite distribution; $\mu$ = 2.2, $\phi$ =1.8**

The introduction of covariates is achieved by mimicking (2) and setting $\mu_i = \exp(x_i'\beta)$ in (7) and (8). The marginal effects of continuous covariates take the same form as in the Poisson regression model, namely for the $k^{\text{th}}$ covariate:

$$\partial E[y_i \mid x_i] / \partial x_{ik} = \exp(x_i'\beta)\beta_k \ . \qquad (9)$$

The coefficients and the corresponding marginal effects have the same signs. As usual, in the case of a "dummy" covariate, the marginal effect is the difference between the conditional mean of the dependent variable, $\mu_i = \exp(x_i'\beta)$, when the dummy variable is unity, and this mean when the dummy variable is zero. The other regressors can be set to their sample mean values.

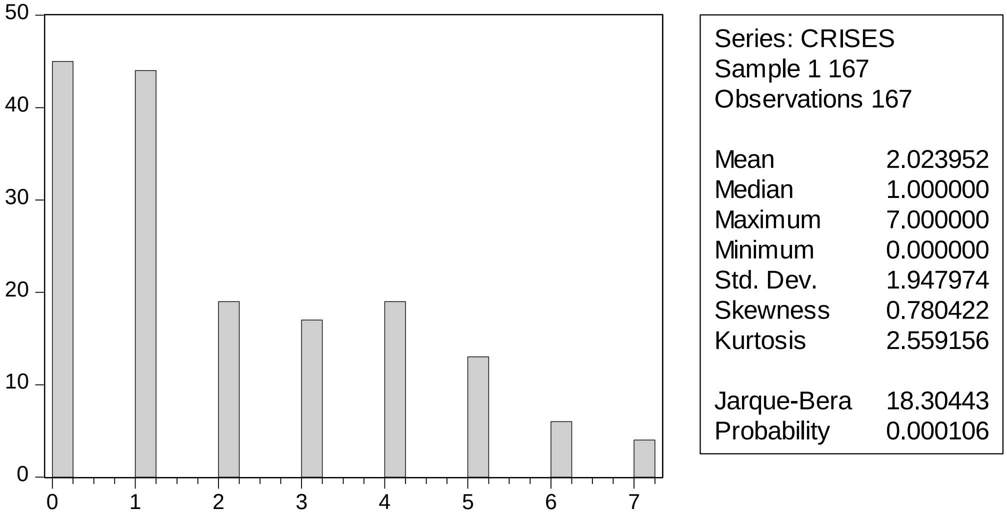
The formulae in (8) facilitate the construction of the log-likelihood function. Assuming $n$ independent observations,

$$\log L = \sum_{i=1}^{n} \Pr.(Y_i = y_i \mid x_i) = \sum_{i=1}^{n}\sum_{r=0}^{r_{\max}} I_r(y_i)\log\{p_r(\beta,\phi \mid x_i)\} \ , \qquad (10)$$

where $I_r(y_i)$ is an indicator function taking the value unity if $y_i = r$, and zero otherwise; $r_{\max}$ is the highest count value in the sample; and $p_r(.)$ comes from (8) with $\mu = \mu_i = \exp(x_i'\beta)$. This log-likelihood is easily programmed in standard econometrics packages.

## 4. Modelling Currency and Banking Crises

We now apply the Hermite regression model to data on the number of currency and banking crises occurring in all IMF member countries between 1970 and 1999. These data are supplied by Ghosh *et al*. (2002). The currency crisis data are based on the Glick and Hutchinson (2001) classification, which measures exchange market pressure by using monthly changes in both the real exchange rate and the level of foreign reserves. The data for banking crises are from Alexander *et al*. (1997) and Glick and Hutchinson (2001). Based on the raw data we have constructed a data-set for the combined number of such crises for each of 167 countries. The frequency distribution for these crises is shown in Figure 3. Its characteristics are typical of those of an Hermite distribution, and the sample data are over-dispersed.



```
Series: CRISES
Sample 1 167
Observations 167

Mean          2.023952
Median        1.000000
Maximum       7.000000
Minimum       0.000000
Std. Dev.     1.947974
Skewness      0.780422
Kurtosis      2.559156

Jarque-Bera   18.30443
Probability   0.000106
```

**Figure 3: Number of currency and banking crises in IMF-member countries**

Source: Ghosh *et al*. (2002), data CD.

Some initial estimation results (ignoring covariates) are given in Table 1. Maximum likelihood estimation of the Hermite model was undertaken by coding a LOGL object in the EViews package (Quantitative Micro Software 2007). The (more restricted) Poisson model is rejected in favour of the negative binomial model using a likelihood ratio test ($p$ = 0), and also using the Wald test of the hypothesis that the exponential of the shape parameter is zero.[4] Recalling that the Hermite distribution collapses to the Poisson distribution when $\gamma = 0$ (and hence $\phi = 0$), we also reject the Poisson model in favour of the Hermite model using a likelihood ratio test ($p$ = 0) and a Wald test ($z$-test). The Hermite model is favoured on the basis of the AIC values. The corresponding actual and fitted counts are shown in Table 2, where the Hermite model provides the best fit to the data for six of the eight counts and the second-best fit for one of the remaining categories. The fit of the Hermite model is especially impressive at the higher counts.

[4] The negative binomial distribution collapses to the Poisson distribution, and the over-dispersion vanishes, as $\eta^2 \rightarrow 0$ . See Quantitative Micro Software (2007, p.248). The Wald test statistic is 0.7731 ($p$ = 0.3793).

**Table 1: Maximum likelihood estimation results (no covariates)**

| | Poisson | NegBin | Hermite |
|---|---|---|---|
| $log(\lambda)$ | 0.7051 | 0.7051 | |
| | (0.0743) | (0.0743) | |
| $log(\eta^2)$ | | -0.5905 | |
| | | (0.2404) | |
| $\mu$ | | | 2.0070 |
| | | | (0.1253) |
| $\phi$ | | | 1.0710 |
| | | | (0.2092) |
| *logL* | -339.5187 | -316.8905 | -305.4095 |
| *AIC* | 4.0781 | 3.8190 | 3.6816 |

Asymptotic standard errors appear in parentheses. These are Huber-White robust standard errors in the case of the Poisson and NegBin models. $log(\eta^2)$ is the shape parameter, as defined in the EViews package (Quantitative Micro Software 2007, p.248).

**Table 2: Actual and predicted counts (currency and banking crises)**

| *r* | Actual | Poisson | NegBin | Hermite |
|---|---|---|---|---|
| 0 | 45 | 83 | 92 | 38 |
| 1 | 44 | 58 | 47 | 36 |
| 2 | 19 | 21 | 18 | 37 |
| 3 | 17 | 45 | 7 | 24 |
| 4 | 19 | 1 | 2 | 16 |
| 5 | 13 | 0 | 1 | 14 |
| 6 | 6 | 0 | 0 | 7 |
| 7 | 4 | 0 | 0 | 5 |

Predicted values are rounded to the nearest integer.

In Table 3 we report some estimated count-data regression models. We focus on covariates associated with the (*de jure*) exchange rate regime in place (pegged, intermediate or floating); and income levels (upper, upper middle, lower middle or lower). We use two dummy variables: DUMIC, which is unity only if the country is an upper middle income country; DUMFLT, which is unity only if a country experienced one or more crises under a (*de jure*) intermediate exchange rate regime; and DUMFLT, which is unity only if a country experienced one or more crises under a (*de jure*) floating exchange rate regime. Although we experimented with various other covariates, these were found to be statistically insignificant. The Poisson model is rejected in favour of the negative binomial model on the basis of a likelihood ratio test ($p$ = 0), but the opposite conclusion is reached using the Wald test of the hypothesis that the exponential of the shape parameter is zero.(In this case the Wald test statistic is 23.3186, with $p$ = 0). The Poisson model is also rejected in favour of the Hermite model (a) on the basis of a likelihood ration test ($p$ = 0) and the $z$-test of the hypothesis that $\phi = 0$. The Hermite regression models again dominate both of the other two models on the basis of the AIC values, with the simpler Hermite model (b) being preferred overall on the basis of its AIC and the insignificance of the interaction term's coefficient ($\beta_4$).

**Table 3: Maximum likelihood estimation results (models with covariates)**

| | **Poisson** | **NegBin** | **Hermite** | |
|---|---|---|---|---|
| | | | **(a)** | **(b)** |
| $\beta_1$ | 0.2742 | 0.2595 | 0.2430 | 0.2396 |
| [const.] | (0.0974) | (0.0963) | (0.0892) | (0.0892) |
| $\beta_2$ | 0.2878 | 0.3128 | 0.2933 | 0.3101 |
| [DINT] | (0.0599) | (0.0667) | (0.0930) | (0.0917) |
| $\beta_3$ | 0.8676 | 0.8776 | 0.8908 | 0.9259 |
| [DFLT] | (0.1175) | (0.1138) | (0.1770) | (0.1643) |
| $\beta_4$ | 0.3149 | 0.3068 | 0.3072 | |
| [DFLT × DUMIC] | (0.1385) | (0.1531) | (0.3667) | |
| $\phi$ | | | 0.4212 | 0.4266 |
| | | | (0.1500) | (0.1488) |
| $log(\eta^2)$ | | -2.5034 | | |
| | | (1.1373) | | |
| $logL$ | -283.0836 | -282.3039 | -272.7315 | -273.5905 |
| $AIC$ | 3.4381 | 3.4408 | 3.3261 | 3.3244 |

Asymptotic standard errors appear in parentheses. These are Huber-White robust standard errors in the case of the Poisson and NegBin models. $log(\eta^2)$ is the shape parameter, as defined in the EViews package (Quantitative Micro Software 2007, p.248).

The results for the Hermite model (b) can be interpreted in terms of the marginal effects associated with the covariates. As these covariates are dummy variables the marginal effects are computed as described in section 3, yielding values of 0.46 and 2.15 respectively for the intermediate and floating exchange rate dummies. So, *ceteris paribus*, moving from a (*de jure*) pegged exchange rate to an intermediate exchange rate would have led a country to experience the same number of crises, or perhaps one more crisis, over the period in question. Similarly, moving from a pegged exchange rate to a (*de jure*) floating exchange rate would have led a country to experience approximately two more crises over this period. This last result accords with the findings of Domaç and Martinez Peria (2003), based on a logit model for *banking* crises.[5]

## 5. Conclusions

The Hermite distribution provides a useful basis for modeling count data whose empirical distribution is multi-modal, and it nests the Poisson distribution. Covariates can be incorporated into the model in a natural way, and the associated maximum likelihood estimates of the parameters are readily obtained. We find that the Hermite regression model out-performs other standard count data models in our illustrative application involving the incidence of financial crises under different exchange rate regimes.

[5] Our own conclusions are not affected when banking and currency crises are modeled separately.

# References

Alexander, W. E., Davies, J. M., Ebrill, L. P. and C-J. Lindgren (1997) *Systematic Bank Restructuring and Economic Policy*, International Monetary Fund: Washington, D.C..

Domaç, I. and M. S. Martinez Peria (2003) “Banking crises and exchange rate regimes: is there a link?” *Journal of International Economics* **61**, 41-72.

Ferrari, S. L. P. and F. Cribari-Neto (2004) “Beta regression for modelling rates and proportions” *Journal of Applied Statistics* **31**, 799-815.

Ghosh, A. R., Gulde, A.-M. and H. C. Wolf (2002) *Exchange Rate regimes - Choices and Consequences*, MIT Press: Cambridge MA.

Giles, D. E. A. (2007) “Modeling inflated count data” in *MODSIM 2007 International Congress on Modelling and Simulation* by L. Oxley and D. Kulasiri, Eds., Modelling and Simulation Society of Australia and New Zealand: Christchurch, N.Z., 919-925.

Glick, R. and M. Hutchinson (2001) “Banking and currency crises” in *Financial Crises in Emerging Markets* by R. Glick, R. Moreno and M. Spiegel, Eds., Cambridge University Press: Cambridge, 35-69.

Hellström, J. (2006) “A bivariate count data model for tourism demand” *Journal of Applied Econometrics* **21**, 213-226.

Kemp, C. D. and A. W. Kemp (1965) “Some properties of the ‘Hermite’ distribution” *Biometrika* **52**, 381-394.

McKendrick, A. G. (1926) “Applications of mathematics to medical problems” *Proceedings of the Edinburgh Mathematical Society* **44**, 98-130.

Maritz, J. S. (1952) “Note on a certain family of discrete distributions” *Biometrika* **39**, 196-198.

Mullahy, J. (1986) “Specification and testing of some modified count data models” *Journal of Econometrics* **33**, 341-365.

Melkersson, M. and D-O. Rooth (2000) “Modeling female fertility using inflated count data models” *Journal of Population Economics* **13**, 189-203.

Quantitative Micro Software (2007) *EViews 6 User’s Guide II*, Quantitative Micro Software: Irvine, CA.

Santos Silva, J. M. C. and F. Covas (2000) “A modified hurdle model for completed fertility” *Journal of Population Economics* **13**, 173-188.