THEORETICAL NOTIONS OF STATISTICAL ESTIMATION

Prof. Constantin ANGHELACHE PhD
Bucharest University of Economic Studies/„Artifex„ University of Bucharest
Assoc. prof. Mădălina-Gabriela ANGHEL PhD
„Artifex„ University of Bucharest
Ihab Jweida S J JWEIDA Ph.D Student
Bucharest University of Economic Studies
Marius Popovici Ph.D Student
Bucharest University of Economic Studies
Emilia Stanciu Ph.D Student
Bucharest University of Economic Studies

Abstract:
This article will address traditional assessment methods, such as maximum likelihood, useful when it is known. Conversely, it is not known, we can use nonparametric methods exploiting specific property that require the involvement of a distribution functions.
Models with discrete variables and partially observed models are usually estimated by maximum likelihood method. We will address some models presented in the first section of this chapter and will use their traditional presentation as a model of parametric indices.
We will address the theory of regression observation asymptotically unbiased estimator using positive and analyzing their effectiveness. We will further analyze the types of errors that are generated from regression and heteroscedastic.

Key words: traditional methods of estimation, distribution functions, parametric indices, logarithmic probabilities, binary selection model

Introduction:
In the following article we apply the idea of dichotomous models. In this regard, we apply binary model selection can take the form of a linear model with indices. Will be used to maximize the likelihood logarithmic regression as non-parametric estimation procedure and the other will be based on minimizing. We analyze the model using "Tobit" where we assume that the remains are normally distributed. We will also address the probability logarithmic forms, standard if we have a mixture of discrete and continuous distributions, but can be maximized using a common method for obtaining recurring MLE.

Methodology and data

In practice econometric can be used different estimation methods. First, the study estimates non parametric models represented models without any assumption indices function indices. Second semi-parametric estimation study involving some form of function indices. Finally, we discuss maximum likelihood estimation.

- **Estimated nonparametric**

  The models presented can be written as models of indices:

$$E^{\theta}(y_i|z_i) = \psi\left(\lambda' z_i\right) = \varphi(z_i),$$

where ψ it is a function of $\Re$ per $\Re$ and $\lambda' z_i$, is a linear combination of elements of z_i. Suppose that is different and z_i, has a continuous density. In addition, to solve the problem of non-identified pair (ψ, λ), bring λ to the normal form, setiing Moreover, we will use here some of its results.

Provided the applicable form from normal and equality

$$E\left[\frac{\partial\varphi(z)}{\partial z_j}\right] = \lambda_j E\left[\frac{\partial\psi(u)}{\partial u}\right],$$

for all $j = 1,\dots q$, it follows that

$$\lambda_j = \frac{E\left[\frac{\partial\varphi(z)}{\partial z_j}\right]}{E\left[\frac{\partial\varphi(z)}{\partial z_1}\right]}.$$

Using this result, the estimate is determined by the following steps:

- Estimate φ by using kernel estimator $\hat{\varphi}_n$.
- Estimate λ_j on

$$\widehat{\lambda}_{jn} = \frac{\sum_{i=1}^{n} \frac{\partial\widehat{\varphi}_n(z_i)}{\partial z_j}}{\sum_{i=1}^{n} \frac{\partial\widehat{\varphi}_n(z_i)}{\partial z_1}}.$$

- Non-parametrical regression y per $\hat{\lambda}_n' z$ where $\hat{\lambda}_n$ is vector which $\hat{\lambda}_{jn}'$, resulting in:

$$\widehat{\psi}_n(u) = \frac{\sum_{i=1}^{n} y_i K\left(\frac{u-\widehat{\lambda}_n' z_i}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{u-\widehat{\lambda}_n' z_i}{h_n}\right)}$$

Another estimation procedure is based on minimizing $E[(\varphi(z)-\psi(\lambda' z))^2]$ paying attenion to ψ to λ. It is described as follows:

- Find an estimate of ψ

$$\widehat{\psi}_{\lambda n}(u) = \frac{\sum_{i=1}^{n} y_i K\left(\frac{u-\lambda' z_i}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{u-\lambda' z_i}{h_n}\right)}.$$

- Estimate λ by

$$\widehat{\lambda}_n = \arg\min \sum_{l=1}^{n} \left(y_l - \frac{\sum_{i=1}^{n} K\left(\frac{\lambda' z-\lambda' z_i}{h_n}\right) y_i}{\sum_{i=1}^{n} K\left(\frac{\lambda' z-\lambda' z_i}{h_n}\right)} \right)$$

- An estimator $\hat{\psi}_n$ is obtained by replacing λ with $\hat{\lambda}_n$ in (11.21).

It can be shown that these estimates of λ are consistent with the rate $\sqrt{n}$ şi and asymptotically normal, and that this test can be implemented. Thus, to test the null hypothesis $H_0 : R\lambda = r_0$, Wald relationship

$$W = \left(R\widehat{\lambda}_n - r_0\right)' \left(R\,\widehat{\Sigma}_{\lambda_n} R'\right)^{-1} \left(R\widehat{\lambda}_n - r_0\right)$$

asymptotically follows X^2, where $\hat{\lambda}_n$ and are, respectively, consistent estimators of λ and of matrix asymptotic covariates Σ_λ.

- **Estimated semi-parametric maximum likelihood**

In some cases, such as models of binary choice model indices

$$E^{\theta}\left(y_i \mid z_i\right) = \psi\left(\lambda' z_i\right)$$

it is such that the function ψ has a property distribution functions. When ψ is known, , it should be used to estimate traditional methods such as maximum likelihood. On the contrary, when ψ is unknown, we return to specific nonparametric methods exploiting the property that ψ is a function of distribution.

We apply this idea to dichotomous models in the following example:

We return to the model of binary selection (11.2) which takes the form of a linear model with indices

- $$E^{\theta}(y_i \mid z_i) = \psi\left(\lambda' z_i\right) = \Pr(y_i = 1 \mid z_i),$$

where ψ has all the properties of distribution functions. If ψ would be known, λ can be estimated through maximizing the logarithmical probability:

$$L(\lambda) = \sum_{i=1}^{n}\left[y_i \ln \psi\left(\lambda' z_i\right) + (1 - y_i)\ln\left(1 - \psi\left(\lambda' z_i\right)\right)\right].$$

Since it is known, we replace ψ with non-parametric estimator $\hat{\psi}_{ni}$

$$\widehat{\psi}_{ni} = \frac{\frac{1}{nh_n}\sum_{j\neq i} y_j J_{nj} K\left(\frac{z-\lambda' z_j}{h_n}\right)}{\frac{1}{nh_n}\sum_{j\neq i} J_{nj} K\left(\frac{z-\lambda' z_j}{h_n}\right)}$$

with

$$J_{nj} = \left|\begin{array}{l} 1 \quad \text{Daca} \quad z_j \in A_{nz} \\ 0 \end{array}\right.$$

where

$$A_{nz} = \left\{ z_i \,\middle|\, \left\| z_i - z_i^* \right\| < 2h_n,\ z_i^* \in A_z \right\},$$

$$A_z = \left\{ z_i \,\middle|\, \psi(\lambda' z_i) \geq \eta,\ \lambda \in B \right\},$$

and B is a compact set of all the categories λ *s*. Therefore, maximize

$$\sum_{i=1}^{n} \left[y_i \ln \widehat{\psi}_{ni}\left(\lambda' z_i\right) + (1 - y_i) \ln\left(1 - \widehat{\psi}_{ni}\left(\lambda' z_i\right)\right) \right]$$

Containing the estimator $\hat{\lambda}_n$ from λ. It can be shown that this estimator is consistent and asymptotically normal

- **Maxim likelihood estimation**

Models with discrete variables and partially observed models are usually estimated by maximum likelihood method. We return to some models presented in the first section of this chapter and use their traditional presentation, ie, not as a model of non-parametric indices.

Consider a traditional depiction of a dichotomous model, where it is assumed that the variable y_i has two values (0 or 1) (y_i = 0 or 1) so

$$\begin{cases} \Pr(y_i = 1) = \psi(\lambda' z_i) \\ \Pr(y_i = 0) = 1 - \psi(\lambda' z_i) \end{cases}$$

Therefore

$$E^{\theta}\left(y_i \,|z_i\right) = 0\left[1 - \psi(\lambda' z_i)\right] + 1\left[\psi(\lambda' z_i)\right] = \psi(\lambda' z_i).$$

Thus Consider the model of "Tobit" defined by (11.6) and (11.7) and assume that the remains are normally distributed. Logarithmic probability is:

$$\begin{aligned} \ln L_n &= \sum_{y_i=0} \ln \mathrm{F_N}\left(-\frac{\lambda' z_i}{\sigma}\right) - \frac{1}{2}\sum_{y_i>0}\left[\ln 2\pi\sigma^2 + \frac{(y_i - \lambda' z_i)^2}{\sigma^2}\right] \\ &= \sum_{y_i=0} \ln \mathrm{F_N}\left(-\frac{\lambda' z_i}{\sigma}\right) - \frac{n_1}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{y_i>0}\left(y_i - \lambda' z_i\right)^2, \end{aligned} \qquad (11.24)$$

wheree n_1 is a positive number y_i. Logarithmic likelihood has a non-standard shape because we have a mixture of discrete and continuous distributions, but can be maximized using a common method for obtaining recurring MLE. Usually a non-parameterized version of (11.24) and is being studied $\gamma = \dfrac{\lambda}{\sigma}$ and $\delta = \dfrac{1}{\sigma}$ i.e.,

$$\ln L_n = \sum_{y_i=0} \ln \mathrm{F_N}\left(-\gamma' z_i\right) - \frac{n_1}{2}\ln 2\pi + n_1 \ln \delta - \frac{1}{2}\sum_{y_i>0}\left(\delta y_i - \gamma' z_i\right)^2.$$

Return to the pattern of imbalance represented by the equation (11.16). We consider for example the case in which the variation of the value is defined by

$$\Delta z_t^* = z_{t+1}^* - z_t^*.$$

Two cases are possible. If $y_t^S > y_t^D$ then $\Delta z_t^* < 0$ from (11.15), and $\Delta z_t^* = \lambda_2^\Delta (y_t - y_t^S)$ (as $y_t = y_t^S$). In this case, assuming normality of rest u_t^D and u_t^S in (11.13), we can write

$$\Delta z_t^* | y_t \sim N\left(\lambda_2^\Delta \left(y_t - z_t^{S\prime}\lambda^S - z_t^* \lambda^{*S}\right), \lambda_2^{\Delta 2}\sigma_S^2\right)$$

and

$$y_t \sim N\left(z_t^{D\prime}\lambda^D + z_t^* \lambda^{*D}, \sigma_D^2\right)$$

Similarly, where we have $y_t^S \leq y_t^D$:

$$\Delta z_t^* | y_t \sim N\left(\lambda_1^\Delta \left(-y_t + z_t^{D\prime}\lambda^D + z_t^* \lambda^{*D}\right), \lambda_1^{\Delta 2}\sigma_D^2\right)$$

and

$$y_t \sim N\left(z_t^{S\prime}\lambda^S + z_t^* \lambda^{*S}, \sigma_S^2\right)$$

Indicated by S_1 assemblage n_1 comments so that $\Delta z_t^* < 0$, and by S_2 assemblage of n_2 other observations ($n_1 + n_2 = n$). Logarithmic probability is:

$$\begin{aligned} \ln L_n = & -n \ln (2\pi\sigma_S\sigma_D) - n_1 \ln \lambda_2^\Delta - n_2 \ln \lambda_1^\Delta \\ & - \frac{1}{2\sigma_D^2}\sum_{S_1}\left(y_t - z_t^{D\prime}\lambda^D - z_t^*\lambda\right)^2 \\ & - \frac{1}{2\sigma_S^2}\sum_{S_2}\left(y_t - z_t^{S\prime}\lambda^S + z_t^*\lambda^{*S}\right)^2 \\ & - \frac{1}{2\lambda_2^{\Delta 2}\sigma_S^2}\sum_{S_1}\left(\Delta z_t^* - \lambda_2^\Delta\left(y_t - z_t^{S\prime}\lambda^S - z_t^*\lambda^{*S}\right)\right)^2 \\ & - \frac{1}{2\lambda_1^{\Delta 2}\sigma_D^2}\sum_{S_2}\left(\Delta z_t^* + \lambda_1^\Delta\left(y_t - z_t^{D\prime}\lambda^D - z_t^*\lambda^{*D}\right)\right)^2 . \end{aligned}$$

The model based on the sampling equations (11.17), (11.18) and (11.19):

$$y_i = \begin{vmatrix} y_i^{(0)*} & Daca \quad y_i^{(1)*} > 0 \\ 0 & \end{vmatrix}$$

with

$$y_i^{(0)*} = \lambda^{(0)\prime} z_i^{(0)} + \varepsilon_i$$

and

$$y_i^{(1)*} = \lambda^{(1)\prime} z_i^{(1)} + \nu_i$$

Moreover, we assume:

$$\begin{pmatrix} \varepsilon_i \\ \nu_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}\right).$$

The likelyhood of the function is given by:

$$L_n = \prod_{y_i=0} F_N\left(-\frac{\lambda^{(1)\prime} z_i^{(1)}}{\sigma_1}\right) \prod_{y_i \neq 0} \frac{1}{\sigma_0} f_N\left(\frac{y_i - \lambda^{(0)\prime} z_i^{(0)}}{\sigma_0}\right)$$

$$\times \prod_{y_i \neq 0} F_N\left(\frac{1}{\sqrt{1-\rho^2}}\left(\frac{\lambda^{(1)\prime}}{\sigma_1} z_i^{(1)} + \frac{\rho}{\sigma_0}\left(y_i - \lambda^{(0)\prime} z_i^{(0)}\right)\right)\right).$$

Since the pair $(\lambda^{(1)}, \sigma_1)$ is not identified, we change parameters:

$\delta_0 = \frac{1}{\sigma_0}, \quad c_0 = \frac{\lambda^{(0)}}{\sigma_0}$ şi $c_1 = \frac{\lambda^{(1)}}{\sigma_1}.$

Thus:

$$L_n = \prod_{y_i=0} F_N\left(-c_1' z_i^{(1)}\right) \prod_{y_i \neq 0} \delta_0 f_N\left(\delta_0 y_i - c_0' z_i^{(0)}\right)$$

$$\times \prod_{y_i \neq 0} F_N\left(\frac{1}{\sqrt{1-\rho^2}}\left(c_1' z_i^{(1)} + \rho\left(\delta_0 y_i - c_0' z_i^{(0)}\right)\right)\right)$$

Probabilitatea poate fi maximizată prin metodele obişnuite. Dar observăm că este de asemenea posibilă folosirea unei proceduri de estimare în două etape:

- We consider the „probit" model associated with the previous model

$$w_i = \begin{vmatrix} 1 \ \ Daca \ \ y_i^{(1)*} > 0 \\ 0 \end{vmatrix}$$

then

$$\Pr(w_i = 1) = \Pr\left(y_i^{(1)*} > 0\right) = F_N\left(\frac{\lambda^{(1)\prime}}{\sigma_1} z_i^{(1)}\right) = F_N\left(c_1' z_i^{(1)}\right).$$

It is, therefore, possible to estimete c_1 by $\hat{c}_{1n}$.

- We consider y_i positive:

$$E^{\theta}(y_i | y_i > 0) = \lambda^{(0)\prime} z_i^{(0)} + \rho\sigma_0 \frac{f_N\left(\frac{\lambda^{(1)\prime}}{\sigma_1} z_i^{(1)}\right)}{F_N\left(\frac{\lambda^{(1)\prime}}{\sigma_1} z_i^{(1)}\right)}$$

Thus:

$$E^{\theta}(y_i | y_i > 0) = \lambda^{(0)\prime} z_i^{(0)} + \rho\sigma_0 \frac{f_N\left(c_1' z_i^{(1)}\right)}{F_N\left(c_1' z_i^{(1)}\right)}.$$

Either:

$$\widehat{\delta}_{in} = \frac{f_N\left(\widehat{c}_{1n}' z_i^{(1)}\right)}{F_N\left(\widehat{c}_{1n}' z_i^{(1)}\right)}.$$

Regression of the positive observation y_i per $z_i^{(0)}$ and $\hat{\delta}_{in}$ leads to asymptotically unbiased estimator $\lambda_i^{(0)}$ and $\rho\sigma_0$, but they are not effective.

Indeed, it can be shown that regression errors are heteroscedastic

Finally, to estimate σ_0, consider the rests of the former estimated regressions

$$\widehat{\eta}_i = y_i - \widehat{\lambda}^{(0)\prime} z_i^{(0)} - \widehat{\rho\sigma_0}\widehat{\delta}_{in}.$$

Whereas:

$$Var^{\theta}(y_i | y_i \neq 0) = Var^{\vartheta}\left(y_i^{(0)} | y_i^{(1)*} > 0\right)$$
$$= \sigma_0^2 + (\rho\sigma_0)^2 \left[-c_1' z_i^{(1)} \frac{f_N\left(c_1' z_i^{(1)}\right)}{F_N\left(c_1' z_i^{(1)}\right)} - \left(\frac{f_N\left(c_1' z_i^{(1)}\right)}{F_N\left(c_1' z_i^{(1)}\right)} \right)^2 \right],$$

We can estimate $\sigma_0{}^2$ by:

$$\widehat{\sigma}_0^2 = \frac{1}{n_1} \sum_{y_i \neq 0} \widehat{\eta}_i^2 + \frac{(\widehat{\rho\sigma_0})^2}{n_1} \sum_{y_i \neq 0} \left[\widehat{\delta}_{in} \widehat{c}_{1n}' z_i^{(1)} + \widehat{\delta}_{in}^2 \right],$$

here n_1 is number y_i different from zero. Thus we get consistent and asymptotically normal estimators.

Conclusion

In this work were analyzed estimation methods traditional as well as maximum likelihood, useful when known but noted that in contrast, it is not known, has been successfully used nonparametric methods specify which exploits the property that is necessary involvement of a distribution functions. It has validated the methods applied dichotomous face and shoulder was verified that the binary model selection can take the form of a linear model with indices. It was used to maximize the likelihood logarithmic regression as non-parametric estimation procedure and the other will be based on minimizing. We have also seen how we can use the model "Tobit" where we presuspus debris are normally distributed. Also forms were addressed logarithmic likelihood, if we have a mixture of standard discrete and continuous distributions, but can be maximized using a common method for obtaining recurring MLE. Models with discrete variables and partially observed models were estimated by maximum likelihood method rule.

References

1. Anghelache, C., Anghel, M.G. (2016). *Bazele statisticii economice. Concepte teoretice şi studii de caz*, Editura Economică, Bucureşti
2. Anghelache, C., Anghel, M.G. (2015). *Statistică. Teorie, concepte, indicatori şi studii ce caz*, Editura Artifex, Bucureşti
3. Anghelache, C., Marinescu, R.T. (2010). *Estimation of the Moment Through the Perturbation Parameters*, Metalurgia International, nr. 4/2010, pp. 131-133
4. Anghelache, C., Marinescu, R.T., Manole M. (2009). *Estimarea şi utilizarea funcţiei de regresie prin nucleu*, Supliment al Revistei Române de Statistică, nr. 12/2009, pp. 103-106
5. Anghelache, C. (2008). *Tratat de statistică teoretică şi economică*, Editura Economică, Bucureşti
6. Arcidiacono, P., Miller, R.A. (2011). *Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity*, Econometrica 79 (November 2011), pp.1823–1867
7. Bosq, D. (2012). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, Springer Science & Business Media
8. Campbell, J.Y., Thompson, S.B. (2008*). Predicting excess stock returns out of sample: Can anything beat the historical average?*, Review of Financial Studies 21 (4), pp. 1509–1531
9. Dodge, Y. (editor) (2006). *The Oxford dictionary of statistical terms*, Oxford University Press
10. Müller, U.K. (2007). *A Theory of Robust Long-Run Variance Estimation*, Journal of Econometrics, 141, pp. 1331—1352.
11. Newbold, P., Karlson, L.W., Thorne, B. (2010). *Statistics for Business and Economics*, 7th ed., Pearson Global Edition, Columbia, U.S
12. Newey, W., Powell, J. (2003). *Instrumental variable estimation of nonparametric models*, Econometrica, pp. 1565–1578