

- Κατάρχας αλληλοσυγκρίνει με το λ της συνάρτησης του MLE.
- Υποθέτουμε ότι το στατιστικό υποδείχθηκε είναι αυτως εφελδμευμένο.
(το δε αναπριβα την πραγματική κατάσταση)
- Υποθέτουμε (χωρίς ευνοίας κ' με κάποια απώλεια γενικότητας) ότι εργαζόμαστε στα πλαίσια της ανεξαρτησίας μεταξύ των στοιχείων του δείγματος (η θεωρητική ανεξαρτησία όπως στο υπόδειγμα Probit που έχουμε αναπτύξει.)

— Οπότε η συνάρτηση πιθανοφάνειας υπάρχει την μορφή:

$$l_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ln p(z(i), \theta) \quad \left(\begin{array}{l} \text{στη περίπτωση που έχουμε δεδομένα} \\ \text{σε κάποιες μεταβλητές, έστω } x(i), \\ \text{αυτό βεβαιότερα έχει την μορφή} \\ l_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ln p(z(i), x(i), \theta) \end{array} \right)$$

— Μπορούμε να εκφράσουμε τον MLE, $\hat{\theta}_n$, ισοδύναμα (κι βολικά για την διερεύνηση της συνέπειας) ως $\hat{\theta}_n \in \underset{\theta \in \Theta}{\operatorname{argmax}} l_n(\theta_0) - l_n(\theta)$

- Εξούτως οτι $l_n(b_0) - l_n(b) = \frac{1}{n} \sum_{i=1}^n \ln p(z(i), b_0) - \frac{1}{n} \sum_{i=1}^n \ln p(z(i), b)$

$$= \frac{1}{n} \sum_{i=1}^n \ln \frac{P(z(i), b_0)}{P(z(i), b)}$$

[Για σιμπλωση της διαδικασίας αμελ. έχουμε αυτοτόμως οτι

$$l_n(b_0) - l_n(b) = \frac{1}{n} \sum_{i=1}^n \ln \frac{P(z(i), x(i), b_0)}{P(z(i), x(i), b)}]$$

- $\forall \pi_0$ κάποιες προϋποθέσεις που έχω να κάνω $\psi \in \Pi$ πιθανοθεωρητικές

ιδιότητες των στοιχείων του δείγματος μας (π.χ. υπολογισμός των z στο υπόδειγμα, σταθερότητα για τις μεταβλητές δέσμευσης, αναμενόμενες ιδιότητες του $P(z(i), \theta)$ ως συνάρτηση του θ , κ.ο.κ.) είναι δυνατόν να γίνει εφαρμογή ενός κατάλληλου νόμου μεγάλων αριθμών που να εφαρμόζεται

$$\ell_n(\theta_0) - \ell_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ln \frac{P(z(i), \theta_0)}{P(z(i), \theta)} \xrightarrow{P} \mathbb{E}_{\theta_0} \left(\ln \frac{P(z(i), \theta_0)}{P(z(i), \theta)} \right)$$

$\xrightarrow{\text{από τον νόμο των μεγάλων αριθμών}} \text{ως προς την } P(z(i), \theta_0)$
 Το π.μ.α. υπολογίζεται ως προς θ

- Προβλεπόμενα για την διερεύνηση της γενετικής του h_n , θα
προσπαθήσουμε, χωρίς ιδιαίτερες τεχνικές δεξιότητες να εξημερώσουμε
για παραγωγή την γενετική θεωρία θα τους ευχαριστούμε πολύ M:

κ έχουμε ότι
$$\mathbb{E}_{\mathbf{z}} \left(\ln \frac{p(\mathbf{z}_{(1)}, \mathbf{b}_0)}{p(\mathbf{z}_{(1)}, \mathbf{b})} \right) = KL(p(\mathbf{z}_{(1)}, \mathbf{b}_0), p(\mathbf{z}_{(1)}, \mathbf{b}))$$

(στην περίπτωση που έχουμε τις μεταβλητές δέσμευσης το όριο θα είναι στο πιο περιττού κ θα έχει την μορφή:

$$\mathbb{E}_{\mathbf{x}_{(1)}} \left[KL(p(\mathbf{z}_{(1)}, \mathbf{x}_{(1)}, \mathbf{b}_0), p(\mathbf{z}_{(1)}, \mathbf{x}_{(1)}, \mathbf{b})) \right]$$

αναμενόμενη
 στην κατάσταση του $\mathbf{x}_{(1)}$

- Στην περίπτωση που το όριζο υψήριο είναι το

$$KL(P(z_{(1)}, b_0), P(z_{(1)}, b))$$

Επείδη $KL(P, q) \geq 0$ κ' παίρνουμε την τιμή 0 αν $P = q$

τότε το αλμπτωζιό υψήριο θα ελαχιστοποιείται αν

$P(z_{(1)}, b_0) = P(z_{(1)}, b)$, αν αυτό ισχύει αν $b = b_0$ τότε το b_0 είναι το μοναδικό ελαχιστοποιούν σημείο του αλμπτωζιού μας κριτηρίου.

Υπό οποιες περισσότερες προϋποθέσεις (π.χ. Θ συμπαγές, $k' P(z_{(1)}, \theta)$
 συνεχής συνάρτηση του θ , ή το ασυμπτωτικό κριτήριο υφίσταται ως προς θ
 k' το Θ υφίσταται σύνολο), το θ_0 θα είναι k' ξεχωριστό k' συνεπώς
 βάση της γενικής μας θεωρίας $\theta_n \xrightarrow{P} \argmin_{\theta \in \Theta} KL(P(z_{(1)}, \theta), P(z_{(1)}, \theta)) = \theta_0$.

— Στην περίπτωση που έχουμε την εξέταση k' το ασυμπτωτικό κριτήριο
 έχει την μορφή $\mathbb{E}_{x_{(1)}} (KL(P(z_{(1)}, x_{(1)}, \theta_0), P(z_{(1)}, x_{(1)}, \theta)))$ εφόσον ισχύει ότι

$$P(z_{(1)}, x_{(1)}, b_0) = P(z_{(1)}, x_{(1)}, b) \quad \text{αν } b = b_0 \quad \text{κ}$$

Ισχύουν οι αντίστοιχες αναμενόμενες τιμές για το $\ominus$, την

$P(z_{(1)}, x_{(1)}, b)$ και τότε κ $\pi_{\alpha,1}$:

$$b_n \xrightarrow{P} \argmin_{b \in \Theta} \mathbb{E}_{x_{(1)}} \left[KL(P(z_{(1)}, x_{(1)}, b_0), P(z_{(1)}, x_{(1)}, b)) \right] = b_0.$$

- Ας δούμε πως εξειδικεύονται (εν γένει) τα παρατίω στο

Παράδειγμα Probit που έχουμε αναπτύξει, όταν αυτό είναι καλώς εξειδικευμένο:
(πρωφανώς εδώ έχουμε κι βελτιστοποίηση)

- Υπόμνημα
$$l_n(\theta) = \frac{1}{n} \sum_{i=1}^n \left[y_{(i)} \ln \phi(x_{(i)} \theta) + (1 - y_{(i)}) \ln (1 - \phi(x_{(i)} \theta)) \right]$$

οπότε
$$l_n(\theta_0) - l_n(\theta) = \frac{1}{n} \sum_{i=1}^n \left[y_{(i)} \ln \frac{\phi(x_{(i)} \theta_0)}{\phi(x_{(i)} \theta)} + (1 - y_{(i)}) \ln \frac{1 - \phi(x_{(i)} \theta_0)}{1 - \phi(x_{(i)} \theta)} \right]$$

Υπό τις παραπάνω προϋποθέσεις θα έχουμε ότι:

$$l_n(b_0) - l_n(b) \xrightarrow{P} \underbrace{\mathbb{E} \left(Y_{(1)} \ln \frac{\phi(X_{(1)} b_0)}{\phi(X_{(1)} b)} + (1 - Y_{(1)}) \ln \frac{1 - \phi(X_{(1)} b_0)}{1 - \phi(X_{(1)} b)} \right)}_A \quad (*)$$

(τοπ. συ. ως προς b)

$$(*) = \mathbb{E}(A) \stackrel{\text{I.I.E.}}{=} \mathbb{E} \left(\mathbb{E}(A / X_{(1)}) \right) \quad \text{και έχουμε ότι}$$

$$\mathbb{E}(A / X_{(1)}) = \mathbb{E}(Y_{(1)} / X_{(1)}) \ln \frac{\phi(X_{(1)} b_0)}{\phi(X_{(1)} b)} + \mathbb{E}(1 - Y_{(1)} / X_{(1)}) \ln \frac{1 - \phi(X_{(1)} b_0)}{1 - \phi(X_{(1)} b)} =$$

$$= \phi(x_{(1)} \theta_0) \ln \frac{\phi(x_{(1)} \theta_0)}{\phi(x_{(1)} \theta)} + (1 - \phi(x_{(1)} \theta_0)) \ln \frac{1 - \phi(x_{(1)} \theta_0)}{1 - \phi(x_{(1)} \theta)}$$

$$= KL(\text{Ber}(\theta_0; x_{(1)}), \text{Ber}(\theta; x_{(1)}))$$

— Συνδυάζετε ότι $Y_{(1)} / X_{(1)}$ ακολουθεί κατανομή $\text{Ber}(\theta_0; x_{(1)})$

που ορίζεται από

$$P(Y_{(1)} = 1 / X_{(1)}) = \phi(x_{(1)} \theta_0)$$

— Το υποσύνολο Θ αποτελείται από τις κατανομές

$$P(Y_{(1)} = 0 / X_{(1)}) = 1 - \phi(x_{(1)} \theta_0)$$

$$\text{Ber}(\theta; x_{(1)}) \theta \in \Theta, \text{ με } P(Y_{(1)} = 1 / X_{(1)}) = \phi(x_{(1)} \theta)$$

$$P(Y_{(1)} = 0 / X_{(1)}) = 1 - \phi(x_{(1)} \theta)$$

{Πουθενά υπό τις παραπάνω προϋποθέσεις το αλυσπώζωο
μνημείο θα είναι 70

$$E_{X(1)} \left(KL \left(\text{Ber}(b_0; X(1)), \text{Ber}(b; X(1)) \right) \right)$$

Είναι άρατιον η $P(Y(1)=1/X(1))$ για την $\text{Ber}(b_0; X(1))$ να ισούται
με την $P(Y(1)=1/X(1))$ για την $\text{Ber}(b, X(1))$,

Στοιχείως αλχοζουγιάζε $\psi \in \Gamma_0$ πότε είναι δυνατόν να ισχύει

ψ ισχύει $\phi(x_{(1)} b_0) = \phi(x_{(1)} b) \stackrel{\psi \phi_{\text{v.a.}}}{\implies} \implies \psi \perp \perp$

$$x_{(1)} b_0 = x_{(1)} b \quad (\Leftrightarrow) \quad x_{(1)} (b_0 - b) = 0 \quad (*)$$

η (*) θα ισχύει αν το $b_0 - b = 0_{p \times 1} \Leftrightarrow b = b_0$ αν οι τυχαίες μεταβλητές που απαιτούν το τυχαιο δίδινος $x_{(1)}$ είναι γραμμικά ανεξάρτητες
 Είναι δυνατόν να αποδειχθεί ότι το παραπάνω θα ισχύει αν rank $E(x_{(1)} x_{(1)}) = p$

Σπουδώς, σφύβου εβαφθαγίβουε την σύγκριση του $l_n(b_0) - l_n(b)$

στην προηγούμενη αναφερόμενη τιμή, σφύβου νομικ $E(X_{(1)}' X_{(1)}) = p$ κ

το $\ominus$ αυρτό, τότε $b_n \xrightarrow{P} b_0$.

Σπινζρέφοντας στην δεύτερη περίπτωση ως προτάθηκε να χρησιμοποιήσουμε χωρίς πολλές λεπτομέρειες το τι θα συνέβαινε στην περίπτωση που το Παρατηρητικό Σταθιστικό μας υπόδειγμα ήταν κακώς εξειδικευμένο, επιπλέον π.χ. ως εργαζούμε υποθέτουμε παρατηρείς συνήκες με την παραγωγή μας αύξηση, δηλ. διεκπεύει ανεφάρτητα, ουσιαστικά κ.ο.κ., επιπλέον η παραγωγή αυτών των $\chi_{(1)} / \chi_{(2)}$ ως που υποθέτουμε DGP, δεν βρίσκεται στο υπόδειγμα.

— Ότι τις αντίστοιχες προϋποθέσεις θα ήταν δυνατόν να ισχύει

ότι η τροποποίηση πιθανόφανε θα συνέχισε με τον κατάλληλο τρόπο

$$\text{στην } \mathbb{E} \left(KL(DGP, P(y_{1:n}, x_{1:n}, \beta)) \right) \quad (*)$$

— Αν για την παραπάνω έχουμε ότι υπάρχει $\beta^* \in \Theta$ τέτοιο ώστε

να είναι το μοναδικό ε' ξεχωριστό ελαχιστοποιούν συνάρτηση της (*), τότε

$\beta_n \xrightarrow{P} \beta^*$, το β^* ονομάζεται ψευδο-αληθινή τιμή της παραμέτρου (pseudo-true value)

↓ Το B^* επισημαίνει εκείνη την κατανομή στο υποδείγμα Y_{AS} που βρίσκεται εγγύτερα στην DGP ως προς την ανυπόθεση απόσταση KL.

— Α> Προβλεπόμενα να αποκτηθούμε μια αίσθηση για το B^* στην περίπτωση του καυώς εφεδικοποιημένου Probit που εφέραμε στο αντίστοιχο πείραμα Monte Carlo: $\varepsilon_{ci}/\chi_{ci} \sim \varepsilon$, όπου το ε είναι "άμετα φερόμενο".

Στην συζευξη των περιπτώσεων η DGP θα είναι Bernoulli:

Υπ' την $P(Y_{(1)} = 1 | X_{(1)})$ να δίνεται από την $F_{t_v}(X_{(1)}; \theta_0)$

όπου F_{t_v} η οριστική συνάρτηση της t_v .

Υπό τις προϋποθέσεις που υπέστησαμε η τροποποιημένη πιθανότητα θα συχνησκει

στην $\mathbb{E} \left(KL \left(\text{Ber}(X_{(1)}, \theta_0, F_{t_v}), \text{Ber}(X_{(1)}, \theta, \Phi) \right) \right)$
όπου $\text{Ber}(X_{(1)}, \theta_0, F_{t_v})$ είναι η παρατήρηση, ενώ $\text{Ber}(X_{(1)}, \theta, \Phi)$ είναι καμία για την οποία

Ισχύει ότι $P(Y=1|X) = \Phi(X\beta)$

- Πώς μπορούμε να αποκτήσουμε κάποια πληροφορία για το αν υπάρχει το β^* ή για το ψ τι ισχύει. Ας προτείνουμε να το εξετάσουμε όπως εξής:

Είναι δυνατόν να δείξω ότι όταν το ψ είναι "μικρό" $\rho = \frac{z}{\sigma}$ $d\rho = \frac{dz}{\sigma}$

$$t_v \approx N\left(0, \frac{\hat{v}}{v-2}\right) \Rightarrow \int_{t_v} (X_{(1)}\beta_0) \approx \int_{-\infty}^{X_{(1)}\beta_0} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{p^2}{2\sigma^2}} dp =$$

$$\int_{-\infty}^{\frac{x_{(1)} b_0}{b}} \frac{1}{\sqrt{2\pi}} e^{-p^2/2} dp = \Phi\left(\frac{x_{(1)} b_0}{b}\right) = \Phi\left(\frac{\sqrt{v-2}}{\sqrt{v}} x_{(1)} b_0\right)$$

$$= \Phi\left(x_{(1)} \sqrt{\frac{v-2}{v}} b_0\right)$$

Συμπεραίνει το ασυμπτωτικό μας κριτήριο για $v \in \mathcal{X}_{\text{αξ}}$ $\vee$ είναι $\frac{1}{\pi}$ περίπου 16%,

$$\mathbb{E}\left[KL\left(\text{Ber}\left(x_{(1)}, \sqrt{\frac{v-2}{v}} b_0, \phi\right), \text{Ber}\left(x_{(1)}, b, \phi\right)\right)\right]$$

κ' σφίγγουν ισχύει ότι $\text{rank} \left[\begin{pmatrix} X_{(1)}' & X_{(1)} \end{pmatrix} \right] = p$ κ' $\Theta = \mathbb{R}^p$

Τότε το πρόβλημα θα ελαχιστοποιείται μοναδικά κ' βγαπίζει

$$\text{αν } \beta^* = \sqrt{\frac{v-2}{v}} b$$

↳ γραμμικά υπεραριθμικά
του b .