## Υπενθύμιση κάποιων ορισμών και προτάσεων

### Αθροιστική συνάρτηση κατανομής (και ολοκλήρωση ως προς).

(cumulative distribution function, cdf)

*  τυχαία μεταβλητή. Η κατανομή της περιγράφεται πλήρως από την **αθροιστική συνάρτηση κατανομής** 

, αύξουσα, συνεχής από δεξιά

* Αν  **διακριτή**, δηλαδή  τότε

**η είναι κλιμακωτή και πηδάει κατά  στα ** (σχήμα...)

* **Ορίζουμε** για :



### Έννοιες σύγκλισης τυχαίων μεταβλητών

Έστω  ακολουθία τυχαίων μεταβλητών, π.χ., όπου n αριθμός παρατηρήσεων.

#### Σύγκλιση κατά νόμο (convergence in distribution)

**Ορισμός: Γράφουμε****τότε και μόνο τότε όταν**



π.χ. για .

Τότε **Κ.Ο.Θ** : , δηλαδή (εξ ορισμού)



**Εφαρμογή:** Κατασκευή ασυμπτωτικών δ.ε. για  (αν  γνωστό):

Από  έπεται:



Και άρα



έπεται 

έπεται 

έπεται 

#### Σύγκλιση κατά πιθανότητα (convergence in probability)

**Ορισμός: Γράφουμε** **τότε και μόνο τότε όταν**



Π.χ. με . Τότε  (**ασθενής νόμος μεγάλων αριθμών**)

#### Κανόνες

* 
* Αν συνεχής τότε:



* Αν, όπου C σταθερά, έπεται

***Εφαρμογή***

* Ειδικά : αν  έπεται 
* Κατασκευή ασυμπτωτικών δ.ε. για  (αν  άγνωστο):

Από  και  έπεται

* 

Από δώ συνεχίζουμε όπως πριν με  στη θέση του  και παίρνουμε:



### Σύμβολα τάξης μεγέθους ακολουθιών

Έστω ακολουθία πραγματικών αριθμών.

Λέμε

*  αν  φραγμένη, δηλ 
*  αν 

Επίσης, αν  δεύτερη ακολουθία, με λέμε

*  αν ** φραγμένη,**

δηλ. η  τείνει στο 0 τουλάχιστο τόσο γρήγορα όσο η .

Ή **η  έχει τάξη μεγέθους το πολύ όσο η **

*  αν ,

δηλ. η  τείνει στο 0 ταχύτερα από την,

άρα για μεγάλα n είναι **ή  είναι αμελητέα σε σχέση με τη **

Ή **η  έχει τάξη μεγέθους πολύ μικρότερη της **

### Σύνολα εμπιστοσύνης για ποσότητα θ

Όπως ένα διάστημα εμπιστοσύνης για , έτσι μπορούμε να κατασκευάσουμε "σύνολα εμπιστοσύνης" αν το  μας είναι π.χ. συνάρτηση (η cdf για παράδειγμα)

Αν , όπου  το σύνολο των δυνατών τιμών του , και έστω , το οποίο  μπορεί να εξαρτάται από 

#### Ορισμός

*  είναι**σύνολο εμπιστοσύνης *πεπερασμένου μεγέθους δείγματος*** αν



* ***Ομοιόμορφο*** **ασυμπτωτικό σύνολο εμπιστοσύνης** , αν

 , όπου η κατανομή των 

* ***Σημειακό* ασυμπτωτικό σύνολο εμπιστοσύνης** αν



* + **Προτιμάμε τα «πεπερασμένου δείγματος»,** αν μπορούμε να τα έχουμε γιατί έχουν ακριβή πιθανότητα κάλυψης .
  + **Αν όχι προτιμάμε τα «ομοιόμορφα ασυμπτωτικά»,** καθώς για δοσμένο μεγάλο  μέγεθος δείγματος έχουμε εγγυημένη «κάλυψη» για όλα τα ταυτόχρονα.
  + **Στα μη ομοιόμορφα ασυμπτωτικά** μπορεί το  που χρειάζεται για να πετύχουμε συγκεκριμένη κάλυψη να είναι «όλο και μεγαλύτερο» για «όλο και πιο ακραία»

#### Παραδείγματα

για μορφή συνόλων εμπιστοσύνης **όταν  είναι συνάρτηση**

«**Ζώνη εμπιστοσύνης**»=

«**Φάκελος εμπιστοσύνης**» δίνεται από 

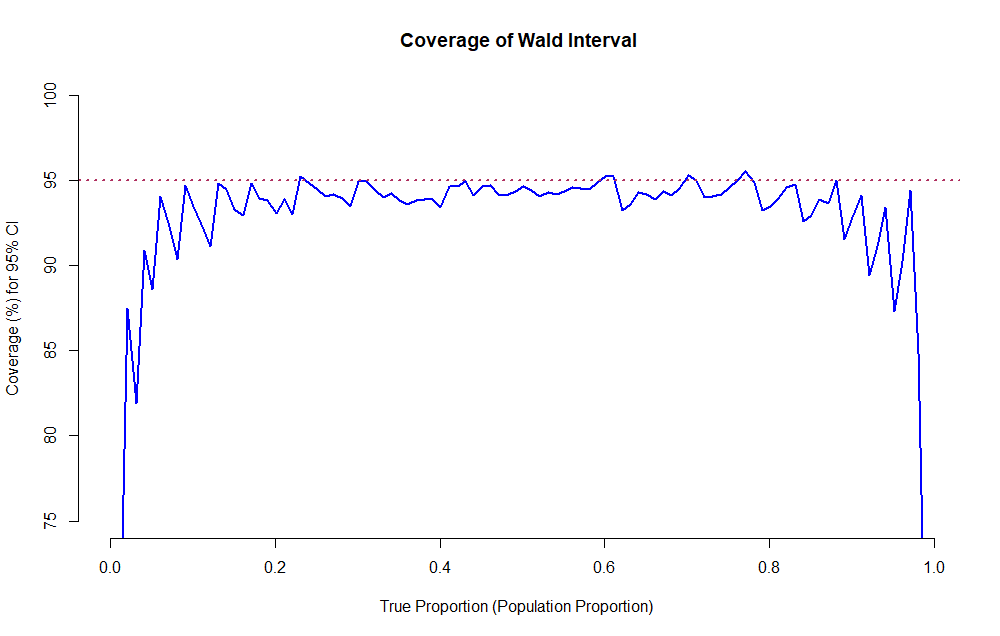


#### Άλλο παράδειγμα: διαστήματα εμπιστοσύνης για διωνυμική πιθανότητα

* Έστω  Bernoulli .
* Tότε  είναι διωνυμική 
* Θέλουμε διαστήματα εμπιστοσύνης για το p.

##### Διαστήματα με κανονική προσέγγιση.

* Αν 
* Τότε επίσης 
* παίρνουμε ένα  δ.ε. από , όπου  το  ποσοστιαίο σημείο της .
* Αυτό το δ.ε. μπορεί να είναι ***μη ομοιόμορφο***: η σύγκλιση εξαρτάται από p. Άρα ***σημειακό ασυμπωτικό δ.ε***.

***n=100, M=10000***

##### Άλλη «απλούστερη» προσέγγιση για δ.ε. πεπερασμένου δείγματος.

Ανισότητα Hoeffding: 

Θέλω   
άρα  δ.ε. πεπερασμένου δείγματος.

### 

##### Delta-method (παρένθεση)

Αν  παραγωγίσιμη στο  και :

Από έπεται 

διότι



Π.χ Αν  τότε 

##### Διαστήματα με (κανονική προσέγγιση) στον arcsin μετασχηματισμό.

Μπορεί να υπολογίσει κανείς ότι 

Επίσης ισχύει ότι

,

μία σύγκλιση που --για p όχι πολύ κοντά στο 0 και στο 1-- είναι ταχύτερη από αυτήν της . Αυτή δίνει - δ.ε. για το :



και αντιστρέφοντας τον μετασχηματισμό παίρνουμε  δ.ε. για το :



##### Διαστήματα κατά Clopper-Pearson (πεπερασμένου δείγματος)

(Στην R: binom.test)

Πρόκειται για δ.ε. που **στηρίζονται στον ακριβή τύπο της διωνυμικής κατανομής** και όχι σε ασυμπτωτική προσέγγιση.

Έστω κάποιο . Τότε η πιθανότητα να πάρω μεγαλύτερο ή ίσο αριθμό επιτυχιών από αυτόν που παρατήρησα είναι .

**Πάρε ως δ.ε. εκείνα τα , τα οποία, αν τα παίρναμε ως μηδενική σε δίπλευρο έλεγχο, οριακά θα τα απορρίπταμε σε επιπεδο . Δηλαδή**

Διαλέγω ως **κάτω όριο του δ.ε.** ένα  τέτοιο ώστε η πιθανότητα υπό  να παρατηρήσω  ή **πιο ακραία μεγάλες τιμές** να ισούται με (ή να είναι μεγαλύτερη του ) :

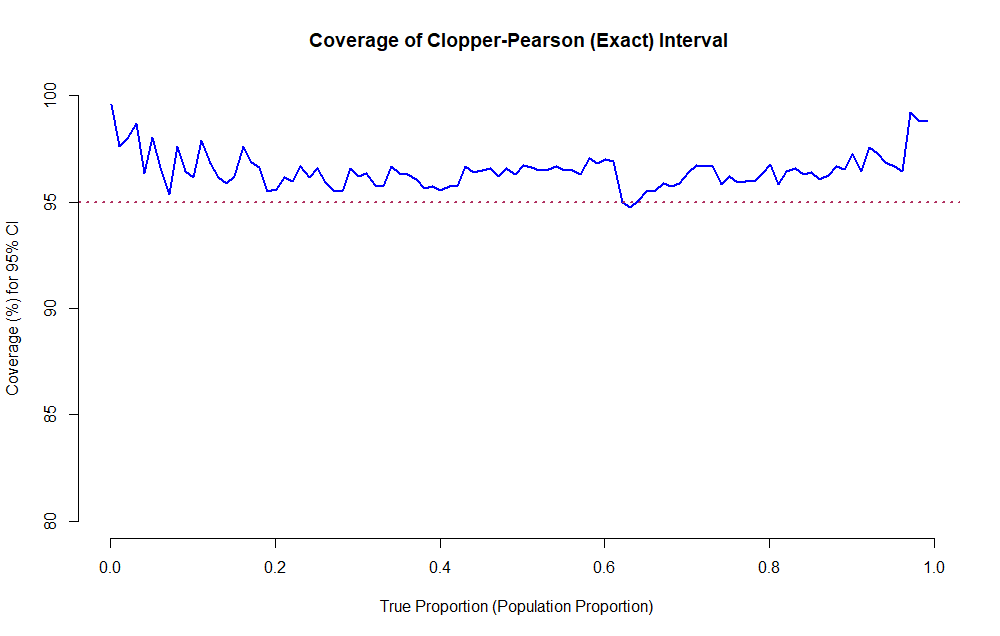
.

Αντιστοιχα:  είναι το μικρότερο  υπό το οποίο η πιθανότητα να πάρω  ή μεγαλύτερες εκτιμήτριες είναι (μεγαλύτερη ή) ίση του .

Παρομοίως από δεξιά: διαλέγω ως άνω όριο του δ.ε. ένα τέτοιο ώστε η πιθανότητα υπό  να παρατηρήσω  ή **πιο ακραία μικρές τιμές** να ισούται (ή να είναι μεγαλύτερη ) του :

.

Τότε το δ.ε. δίνεται από έχει πιθανότητα κάλυψης τουλάχιστο  για οποιοδήποτε .

***n=100, M=10000***