

## Statistical Methods in Epidemiology

### Lab 7

#### Matched Case–Control Studies

##### Στόχος:

Να κατανοηθεί η ανάλυση εξομοιωμένων μελετών τύπου case–control: α) Με χρήση Mantel–Haenszel, β) Με conditional logistic regression, και γ) η διάκριση μεταξύ confounding, effect modification και matching.

#### I. Matched Studies

Σε μελέτες τύπου matched case–control, κάθε ασθενής (case) αντιστοιχίζεται με έναν ή περισσότερους μάρτυρες (controls) με σκοπό τον έλεγχο συγχυτικών παραγόντων μέσω σχεδιασμού. Τα δεδομένα πρέπει να καταγράφονται σε ατομικό επίπεδο, με τις ακόλουθες βασικές μεταβλητές:

- **d:** κατάσταση (1 = ασθενής, 0 = μάρτυρας),
- **set:** αναγνωριστικό εξομοιωμένου συνόλου (matched set).

Σε μελέτη 1:m, κάθε set περιλαμβάνει ακριβώς 1 case και m controls, επομένως θα υπάρχουν 1 + m εγγραφές σε κάθε set.

Η βασική αναλυτική μέθοδος για matched δεδομένα είναι η Mantel–Haenszel για ζεύγη/σύνολα, όπου η αναλογία πιθανοτήτων υπολογίζεται αποκλειστικά από τα *discordant pairs/sets*.

##### i) Matched case–control studies (1:1)

Στη μελέτη των **Breslow και Day (1980)** για τη σχέση χρήσης οιστρογόνων και καρκίνου του ενδομητρίου. Η αντιστοίχιση έγινε ως προς:

- ηλικία ( $\pm 1$  έτος),
- οικογενειακή κατάσταση,
- κοινότητα διαμονής και περίοδο εγκατάστασης.

Οι μεταβλητές **age3** (τριμερής ηλικιακή κατηγοριοποίηση) και **agegrp** (πενταετείς ομάδες) εξασφαλίζουν ότι τα matched pairs έχουν ίδια ηλικία με βάση το matching.

Η αρχική ανάλυση πραγματοποιήθηκε σε υποσύνολο 1:1 (dataset: *bdendo11*). Το πλήρες dataset (*bdendo*) περιλαμβάνει 1:4 αντιστοίχιση.

Το πλήρες dataset (*bdendo*) περιλαμβάνει τις εξής μεταβλητές:

Μεταβλητή Περιγραφή	
<b>set</b>	Αριθμητικό αναγνωριστικό του εξομοιωμένου συνόλου
<b>d</b>	Κατάσταση (1 = ασθενής, 0 = μάρτυρας)
<b>gall</b>	Παθήσεις χοληδόχου κύστης (Όχι/Ναι)
<b>hyp</b>	Υπέρταση (Όχι/Ναι)
<b>ob</b>	Παχυσαρκία (Όχι/Ναι)
<b>est</b>	Χρήση οιστρογόνων (Όχι/Ναι)
<b>dur</b>	Διάρκεια θεραπείας με συζευγμένα οιστρογόνα (ταξινομημένη: 0 < 1 < 2 < 3 < 4)
<b>non</b>	Χρήση μη οιστρογόνων (Όχι/Ναι)
<b>duration</b>	Μήνες θεραπείας με οιστρογόνα (αριθμητική)
<b>age</b>	Ηλικία (αριθμητική)
<b>cest</b>	Δόση συζευγμένων οιστρογόνων (ταξινομημένη: 0 < 1 < 2 < 3)
<b>agegrp</b>	Κατηγορία 5ετίας: 55–59, 60–64, 65–69, 70–74, 75–79, 80–84
<b>age3</b>	Ευρεία κατηγορία ηλικίας: <64, 65–74, 75+

1. Ανοίξτε το αρχείο **bdendo11.dta**. Ποιες μεταβλητές περιλαμβάνει;

```
. use bdendo11, clear
. desc
```

2. Εξετάστε την επίδραση των οιστρογόνων με τη μέθοδο **Mantel–Haenszel**. Βρείτε:

α) τον αριθμό των ζευγών όπου το *case* είναι εκτεθειμένο και ο *control* μη εκτεθειμένος και

β) τον αριθμό των ζευγών όπου συμβαίνει το αντίστροφο.

Επαληθεύστε ότι η εκτίμηση (OR) Mantel–Haenszel ισούται με το πηλίκο αυτών των δύο αριθμών.

```
. mhdods d est set
. sort set d est
. list set d est in 1/10
. by set: gen new1 = est[_n+1] - est
```

3. Εξετάστε αν μπορεί να εκτιμηθεί η επίδραση της μεταβλητής *age3*. Τι παρατηρείτε;

```
. mhdods d age3 set
```

4. Εξετάστε εάν η ηλικία (**age3**) τροποποιεί την επίδραση των οιστρογόνων (effect modification).

```
. mhdods d est set, by(age3)
```

5. Εξετάστε την επίδραση των οιστρογόνων προσαρμόζοντας για υπέρταση (*hyp*) με Mantel–Haenszel. Πόσο αυξάνεται ο αριθμός των strata και πόσα συμβάλλουν στην εκτίμηση;

```
. mhdods d est set hyp
```

## ii) Matched case–control studies (1:4)

1. Εκτιμήστε την επίδραση των οιστρογόνων και αναφέρετε πόσα sets συνεισφέρουν στην εκτίμηση.
2. Προσπαθήστε να εκτιμήσετε την επίδραση της age3.
3. Ελέγξτε αν η age3 τροποποιεί την επίδραση των οιστρογόνων.
4. Εκτιμήστε την επίδραση των οιστρογόνων προσαρμόζοντας για hgr. Πόσα επιπλέον strata δημιουργούνται και πόσα συνεισφέρουν;

## II. Υπό συνθήκη λογιστική παλινδρόμηση (Conditional Logistic Regression)

1. Διαβάστε τα δεδομένα από το bdendo11 και εξετάστε την επίδραση των οιστρογόνων με **conditional logistic regression**.

```
. clogit d i.est, group(set) or nolog
```

2. Προσπαθήστε να εκτιμήσετε την επίδραση της age3.

```
. sort set d
. xi: clogit d i.age3, group(set) or
```

3. Εξετάστε εάν η επίδραση των οιστρογόνων τροποποιείται από την age3 χρησιμοποιώντας interaction terms. Χρησιμοποιήστε τόσο likelihood ratio test όσο και Wald test.

```
. xi: clogit d i.est i.age3, group(set) or
. est store A
. xi: clogit d i.est*i.age3, group(set) or
. lrtest A
```

```
. xi: clogit d i.est*i.age3, group(set) or nolog  
. testparm _IestXage_2_*
```

4. Επαναλάβετε τον έλεγχο αλληλεπίδρασης χρησιμοποιώντας την age ως αριθμητική μεταβλητή.

```
. xi: clogit d i.est*age, group(set) or nolog  
. est store A  
. xi: clogit d i.est age, group(set) or nolog  
. lrtest A
```

5. Εξετάστε την επίδραση των οιστρογόνων προσαρμόζοντας για hyp. Υπάρχει απώλεια πληροφορίας;

```
. xi: clogit d i.est i.hyp, group(set) or nolog
```

6. Εξετάστε αν η υπέρταση τροποποιεί την επίδραση των οιστρογόνων.

```
. xi: clogit d i.est*i.hyp, group(set) or nolog  
. est store A  
. xi: clogit d i.est i.hyp, group(set) or nolog  
. lrtest A
```

7. Αναλύστε τη μελέτη 1:4 με conditional logistic regression. Συγκρίνετε με τα αποτελέσματα Mantel–Haenszel.

### III. Conditional vs Απλής Λογιστικής Παλινδρόμησης

1. Αναλύστε δεδομένα λέπρας: εκτιμήστε την επίδραση του BCG προσαρμόζοντας για ηλικία (κατηγορική) με απλή logistic regression.

```
. xi: logit d i.bcg i.age, or nolog
```

2. Επαναλάβετε την ανάλυση με conditional logistic regression χρησιμοποιώντας τεχνητή μεταβλητή grp=1.

```
. gen grp = 1  
. xi: clogit d i.bcg i.age, group(grp) or
```

3. Συγκρίνετε τα δύο αποτελέσματα.