(Lecture five)

❖ **Data accuracy and testing methods**

The goal of the population census is to obtain accurate, comprehensive data about the total population, but due to multiple factors, the data may be inaccurate, and there are different methods, the most important of which are:

1. Internal consistency methods
2. External consistency methods

1. Internal consistency methods

Internal consistency methods include the following tests:

a. Total numbers accuracy test:

This is done by comparing the population in the current census with the population in the previous census, extracting the growth rate, as well as comparing the number in a city with the number in the previous census, the birth rate, deaths and growth rates in this city, in other words, comparing and examining the data with each other.

ويتم ذلك من خلال مقارنة عدد السكان في التعداد الحالي مع عدد السكان في التعداد السابق، واستخراج معدل النمو، وكذلك مقارنة العدد في مدينة ما مع العدد في التعداد السابق، ومعدل المواليد والوفيات ومعدلات النمو في هذه المدينة، بمعنى آخر، مقارنة وفحص البيانات مع بعضها البعض.

b. Test distribution by gender

It is known that the proportions of males and females are relatively equal in society, except in the case of migration. Likewise, the proportions are also similar in the case of deaths, while the proportion of males in the case of births is greater, then the number of deaths decreases, and then the number becomes equal after that. Based on that, it is not possible Statistical results may come in a way that contradicts these patterns, and if necessary, they will be considered.

من المعروف ان نسب الذكور والاناث متساوية بصورة نسبية في المجتمع، الا في حالة الهجرة، كذلك تكون النسبية متماثلة ايضا في حالة الوفيات، في حين ان نسبة الذكور في حالة الولادات هو أكثر ثم يتناقص العدد الوفيات ثم يتساوى العدد بعد ذلك، وبناءً على ذلك لا يمكن للنتائج الاحصائية ان تأتي بصورة مخالفة لهذه الانماط، وإذا استدعى الامر، فسوف يتم النظر فيها.

c. Test distribution by Age:

a. The population of young age groups is larger than the population of older age groups, due to the annual increase in births and deaths in older age groups.

b. The number of males at birth is greater than the number of females, then it decreases until it is equal and then less than that of females in advanced age groups. These percentages may decrease due to wars and migration.

c. The population in the age group between (25 - 29) years, for example, in a certain census will be less than the population in the age group (25 - 35) years in the subsequent census, that is, after ten years, due to deaths that can be determined using the life table. If it is found that the number is the same or higher, this indicates an error in the data, unless there is migration into the country for this age group.

d. Comparing the population of children with their birth records, if they do not match, then we use appropriate life tables to estimate the number of surviving births.

عدد السكان لفئات العمر الصغيرة هو أكبر من عدد السكان لفئات العمر الكبيرة، بسبب الزيادة السنوية في الولادات وحدوث لوفيات في الفئات العمر الكبيرة

ب. عدد الذكور عند الولادة هو أكبر من عدد الاناث، ثم يتناقص الى ان يتساوى ثم يقل عن الاناث في فئات العمر المتقدمة، وقد تقل هذه النسب بسبب الحروب والهجرة.

ج. ان عدد السكان في الفئة العمرية بين (25 – 29) سنة على سبيل المثال، في تعداد معين سيكون اقل من في الفئة من (25 - 35) سنة في التعداد اللاحق، اي بعد عشرة سنوات، وذلك بفعل الوفيات التي يمكن تحديدها باستخدام جدول الحياة، فاذا ما وجد ان العدد نفسه او زاد عنه، فان ذلك يؤشر الى وجود خطأ في البيانات، الا إذا كان هنالك هجرة الى داخل البلد لهذه الفئة العمرية.

د. مقارنة تعداد الاطفال بسجلات ولاداتهم، إذا لم يتم التطابق، عندئذ نستخدم جداول حياة مناسبة لتقدير عدد الباقين على قيد الحياة من المواليد.

2. External consistency methods

This can be done by examining the census with each other or comparing it either with vital and special statistics data for a specific country or region, revealing its trends and trying to find the causes, or comparing it with standard population models, noting that the researcher must be careful, as not every deviation from the model is considered an error, but rather He must note the

conditions of the country or region he is studying. One of the most important methods of external comparison is:

Balanced equation

The balanced equation method is one of the most important methods that can be used in evaluating census data and estimating errors:

$$\boldsymbol{P_1 = P_0 + B + D + I - E}$$

Where :

P_1: The total population in the current census

P_0: Total population in the previous census

B: Number of live births during the period between the two censuses

D: Number of deaths during the period between the two censuses

I: number of inward migrants (arrivals) during the period between the two censuses

E: number of immigrants abroad (displaced persons) during the period between the two censuses