

Standardisation

One of the purposes of statistical analysis is to place the data in perspective and this is often done by comparing absolute numbers or more often proportions.

Comparisons of the absolute numbers has limitations as it does not take into account the size of the population exposed to the risk of events.

Standardisation is a technique which controls for the compositional variation between the groups being compared. Thus standardised rates gives a 'true' comparison of the phenomena studied.

An example of standardisation in the ABS is age standardisation where the age distribution could lead to misleading results. Age-standardising your data removes the effects of the different age structure of the two groups you are comparing.

Example: Age-standardising is appropriate when comparing health outcomes of Indigenous and non-Indigenous persons. Australia's Indigenous population is considerably younger (on average) than the non-Indigenous population, and there is a close relationship between age and health related issues.

Before undertaking any standardisation it is important to understand the question you are trying to answer. If for example, you are trying to compare different groups and searching for underlying differences other than age, then age standardisation is appropriate.

Standardising your population may not always be appropriate. Care should be taken when deciding to use standardisation. Essentially, standardisation is useful in the following situations:

- ◆ Comparing an aggregate level variable such as a rate or proportion among two or more populations or the same population over time.
- ◆ Comparing a variable which takes on different values between subgroups.
- ◆ The intention is to minimise compositional differences between populations.

Example of age standardisation

The concept of age standardisation may become clearer using a real-life example.

Let us calculate the age standardised rate of those with year 12 schooling for those individuals residing in the top 80 % of Census Districts (CD's) ranked by SEIFA's index of advantage and disadvantage. This example is based on 2001 Population Census data.

First off, we need 3 sets of information:

1. The frequency of the condition (those with year 12 schooling) for the specific category (top 80% of CD's) broken down by age (eg 5 year age categories in this case)
2. The population of the specific category (all individuals residing in top 80% CD's).
3. The standard population.

STEP 1: Frequencies

Create a spreadsheet with the following information, with the category specific frequencies. eg those with a year 12 education.

	Frequency
Age 15-19	301,935
Age 20-24	695,209
Age 25-29	699,226
Age 30-34	612,243
Age 35-39	532,534
Age 40-44	522,726
Age 45-49	452,460
Age 50-54	372,180
Age 55-59	243,371
Age 60-64	160,373
Age 65-69	114,088
Age 70-74	98,131
Age 75-79	71,561
Age 80-84	41,263
Age 85+	29,270
Total	4,946,570

STEP 2: Category specific populations

Next insert the category specific population (eg all individuals whose usual place of residence is in the top 80% of CD's ranked by SEIFA)

	Frequency	Category specific population
Age 15-19	301,935	637,367
Age 20-24	695,209	947,013
Age 25-29	699,226	1,023,159
Age 30-34	612,243	1,095,650
Age 35-39	532,534	1,113,683
Age 40-44	522,726	1,119,776
Age 45-49	452,460	1,028,405
Age 50-54	372,180	968,620
Age 55-59	243,371	731,383
Age 60-64	160,373	564,879
Age 65-69	114,088	448,072
Age 70-74	98,131	404,190
Age 75-79	71,561	316,234
Age 80-84	41,263	193,034
Age 85+	29,270	148,447
Total	4,946,570	10,739,912

STEP 3: Age specific rates

Now calculate the age specific rates:

Age-specific rates = Frequency / Category specific populations

	Frequency	Category specific population	Age specific rates
Age 15-19	301,935	637,367	0.4737224
Age 20-24	695,209	947,013	0.7341071
Age 25-29	699,226	1,023,159	0.6833992
Age 30-34	612,243	1,095,650	0.5587943
Age 35-39	532,534	1,113,683	0.4781738
Age 40-44	522,726	1,119,776	0.466813
Age 45-49	452,460	1,028,405	0.4399629
Age 50-54	372,180	968,620	0.3842374
Age 55-59	243,371	731,383	0.3327545
Age 60-64	160,373	564,879	0.2839068
Age 65-69	114,088	448,072	0.2546198
Age 70-74	98,131	404,190	0.2427843
Age 75-79	71,561	316,234	0.2262913
Age 80-84	41,263	193,034	0.2137603
Age 85+	29,270	148,447	0.1971747
Total	4,946,570	10,739,912	

STEP 4: Insert standard population

You need to use standard population information. In this example our standard population is the Estimated Resident Population (ERP) for Australia for 2001. If you were age standardising 2003 data then you would require ERP data for 2003. You

should be able to construct your own measure of ERP using data contained in ABS Catalogue No. 3201.0 (Population by Age and Sex, Australian States and Territories).

	Frequency	Category specific population	Age specific rates	Standard population
Age 15-19	301,935	637,367	0.4737224	1,352,745
Age 20-24	695,209	947,013	0.7341071	1,302,412
Age 25-29	699,226	1,023,159	0.6833992	1,407,081
Age 30-34	612,243	1,095,650	0.5587943	1,466,615
Age 35-39	532,534	1,113,683	0.4781738	1,492,204
Age 40-44	522,726	1,119,776	0.466813	1,479,257
Age 45-49	452,460	1,028,405	0.4399629	1,358,594
Age 50-54	372,180	968,620	0.3842374	1,300,777
Age 55-59	243,371	731,383	0.3327545	1,008,799
Age 60-64	160,373	564,879	0.2839068	822,024
Age 65-69	114,088	448,072	0.2546198	682,513
Age 70-74	98,131	404,190	0.2427843	638,380
Age 75-79	71,561	316,234	0.2262913	519,356
Age 80-84	41,263	193,034	0.2137603	330,050
Age 85+	29,270	148,447	0.1971747	265,235
Total	4,946,570	10,739,912		15,426,042

STEP 5: Expected frequencies

Using these populations, calculate the expected frequencies.

Expected Frequencies = (**Age-Specific Rates** x **Standard populations**).

	Frequency	Category specific population	Age specific rates	Standard population	Expected frequencies
Age 15-19	301,935	637,367	0.4737224	1,352,745	640,826
Age 20-24	695,209	947,013	0.7341071	1,302,412	956,110
Age 25-29	699,226	1,023,159	0.6833992	1,407,081	961,598
Age 30-34	612,243	1,095,650	0.5587943	1,466,615	819,536
Age 35-39	532,534	1,113,683	0.4781738	1,492,204	713,533
Age 40-44	522,726	1,119,776	0.466813	1,479,257	690,536
Age 45-49	452,460	1,028,405	0.4399629	1,358,594	597,731
Age 50-54	372,180	968,620	0.3842374	1,300,777	499,807
Age 55-59	243,371	731,383	0.3327545	1,008,799	335,682
Age 60-64	160,373	564,879	0.2839068	822,024	233,378
Age 65-69	114,088	448,072	0.2546198	682,513	173,781
Age 70-74	98,131	404,190	0.2427843	638,380	154,989
Age 75-79	71,561	316,234	0.2262913	519,356	117,526
Age 80-84	41,263	193,034	0.2137603	330,050	70,552
Age 85+	29,270	148,447	0.1971747	265,235	52,298
Total	4,946,570	10,739,912		15,426,042	7,017,883

Step 5: Age standardised rates

Finally, calculate the age-standardised rate.

$$\text{Standardised Rate} = \frac{\sum(\text{expected frequencies})}{\sum(\text{standard population})}$$

	Frequency	Category specific population	Age specific rates	Standard population	Expected frequencies
Age 15-19	301,935	637,367	0.4737224	1,352,745	640,826
Age 20-24	695,209	947,013	0.7341071	1,302,412	956,110
Age 25-29	699,226	1,023,159	0.6833992	1,407,081	961,598
Age 30-34	612,243	1,095,650	0.5587943	1,466,615	819,536
Age 35-39	532,534	1,113,683	0.4781738	1,492,204	713,533
Age 40-44	522,726	1,119,776	0.466813	1,479,257	690,536
Age 45-49	452,460	1,028,405	0.4399629	1,358,594	597,731
Age 50-54	372,180	968,620	0.3842374	1,300,777	499,807
Age 55-59	243,371	731,383	0.3327545	1,008,799	335,682
Age 60-64	160,373	564,879	0.2839068	822,024	233,378
Age 65-69	114,088	448,072	0.2546198	682,513	173,781
Age 70-74	98,131	404,190	0.2427843	638,380	154,989
Age 75-79	71,561	316,234	0.2262913	519,356	117,526
Age 80-84	41,263	193,034	0.2137603	330,050	70,552
Age 85+	29,270	148,447	0.1971747	265,235	52,298
Total	4,946,570	10,739,912		15,426,042	7,017,883
				Age-standardised rate	0.455