M3S12 BIOSTATISTICS - EXERCISE 1
SOLUTIONS

(a) Crude mortality rates

Colorado: $\dfrac{D_C}{N_C} = \dfrac{527}{53808} = 9.79 \times 10^{-3}$ 9.79 per 1000 live births

Louisiana: $\dfrac{D_L}{N_L} = \dfrac{872}{73967} = 11.79 \times 10^{-3}$ 11.79 per 1000 live births

(b) Race-specific mortality rates for the two states

	Colorado			
Race	Number of Infant Deaths	Number of Births	Infant Death Rate	Rate per 1000 live births
Black	52	3166	$\frac{52}{3166} = 16.42 \times 10^{-3}$	16.42
White	469	48805	$\frac{469}{48805} = 9.61 \times 10^{-3}$	9.61
Other	6	1837	$\frac{6}{1837} = 3.27 \times 10^{-3}$	3.27
Total	527	53808	$\frac{527}{53808} = 9.79 \times 10^{-3}$	9.79

	Louisiana			
Race	Number of Infant Deaths	Number of Births	Infant Death Rate	Rate per 1000 live births
Black	526	29670	$\frac{526}{29670} = 17.73 \times 10^{-3}$	17.73
White	343	42749	$\frac{343}{42749} = 8.02 \times 10^{-3}$	8.02
Other	3	1548	$\frac{3}{1548} = 1.94 \times 10^{-3}$	1.94
Total	872	73967	$\frac{872}{73967} = 11.79 \times 10^{-3}$	11.79

Clearly the rates for the three different race categories seem to differ, and the pattern of variability is similar but not identical in the two states.

(c) Direct standardization: need to produce a standardized rate by using the formula

$$R_A^{(S)} = \sum_{k=1}^{3} w_k^{(S)} R_{Ak} = \sum_{k=1}^{3} \left(\frac{N_{Sk}}{N_S}\right) R_{Ak} = \sum_{k=1}^{3} \left(\frac{N_{Sk}}{N_S}\right)\left(\frac{D_{Ak}}{N_{Ak}}\right)$$

- S is the standardizing population, here the entire US population
- N_{Sk} is the number in race stratum k, $k = 1, 2, 3$, and $N_S = N_{S1} + N_{S2} + N_{S3}$
- A is the population of interest, here the populations of Colorado, and then Louisiana
- N_{Ak} is the number in race stratum k, $k = 1, 2, 3$, and $N_A = N_{A1} + N_{A2} + N_{A3}$ for the population in each state
- D_{Ak} is the number of deaths in race stratum k, $k = 1, 2, 3$, where $D_A = D_{A1} + D_{A2} + D_{A3}$

Hence, for Colorado

$$\begin{aligned} R_C^{(S)} &= \sum_{k=1}^{3} \left(\frac{N_{Sk}}{N_S}\right) R_{Ck} \\ &= \left(\frac{641567}{3809394} \times \frac{52}{3166}\right) + \left(\frac{2992488}{3809394} \times \frac{469}{48805}\right) + \left(\frac{175339}{3809394} \times \frac{6}{1837}\right) = 10.47 \times 10^{-3} \end{aligned}$$

or 10.47 per 1000 live births.

For Louisiana

$$\begin{aligned} R_L^{(S)} &= \sum_{k=1}^{3} \left(\frac{N_{Sk}}{N_S}\right) R_{Lk} \\ &= \left(\frac{641567}{3809394} \times \frac{526}{29670}\right) + \left(\frac{2992488}{3809394} \times \frac{343}{42749}\right) + \left(\frac{175339}{3809394} \times \frac{3}{1548}\right) = 9.38 \times 10^{-3} \end{aligned}$$

or 9.38 per 1000 live births.

Hence the **directly standardized rate ratios** for the two states compared to the US population are

$$\begin{aligned} SRR(C;S) &= \frac{R_C^{(S)}}{R_S^{(S)}} = \frac{R_C^{(S)}}{R_S} = \frac{10.47 \times 10^{-3}}{38408/3809394} = 1.04 \\ SRR(L;S) &= \frac{R_L^{(S)}}{R_S^{(S)}} = \frac{R_L^{(S)}}{R_S} = \frac{9.38 \times 10^{-3}}{38408/3809394} = 0.93 \end{aligned}$$

These can be compared with the **crude rate ratios**.

$$\begin{aligned} CRR(C;S) &= \frac{R_C}{R_S} = \frac{527/53808}{38408/3809394} = 0.97 \\ CRR(L;S) &= \frac{R_L}{R_S} = \frac{872/73967}{38408/3809394} = 1.17 \end{aligned}$$

that is, the standardization calculation has reversed the direction of the difference compared with the US population.

For indirect standardization: need to produce a standardized rate by using the formula

$$R_S^{(A)} = \sum_{k=1}^{3} w_k^{(A)} R_{Sk} = \sum_{k=1}^{3} \left(\frac{N_{Ak}}{N_A}\right) R_{Sk} = \frac{1}{N_A}\sum_{k=1}^{3} N_{Ak} R_{Sk} = \frac{E_A}{N_A}$$

that is, the **expected** number of deaths in population A, E_A divided by the number of people in that population, N_A. The observed rate is $R_A = D_A/N_A$, and the race-specific rate in race stratum k is

$$R_{Sk} = \left(\frac{D_{Sk}}{N_{Sk}}\right).$$

Hence, for Colorado

$$\begin{aligned} R_S^{(C)} &= \sum_{k=1}^{3} \left(\frac{N_{Ck}}{N_C}\right) R_{Sk} \\ &= \left(\frac{3166}{53808} \times \frac{11461}{641567}\right) + \left(\frac{48805}{53808} \times \frac{25810}{2992488}\right) + \left(\frac{1837}{53808} \times \frac{1137}{175339}\right) = 9.10 \times 10^{-3} \end{aligned}$$

or 9.10 per 1000 live births.

For Louisiana

$$\begin{aligned} R_S^{(L)} &= \sum_{k=1}^{3} \left(\frac{N_{Lk}}{N_L}\right) R_{Sk} \\ &= \left(\frac{29670}{73967} \times \frac{11461}{641567}\right) + \left(\frac{42749}{73967} \times \frac{25810}{2992488}\right) + \left(\frac{1548}{73967} \times \frac{1137}{175339}\right) = 12.29 \times 10^{-3} \end{aligned}$$

or 12.29 per 1000 live births.

Hence the **indirectly standardized rate ratios**, or **standardized mortality ratios,** for the two states compared to the US population are

$$\begin{aligned} SMR(C;S) &= \frac{R_C^{(C)}}{R_S^{(C)}} = \frac{D_C/N_C}{R_S^{(C)}} = \frac{527/53808}{9.10 \times 10^{-3}} = 1.08 \\ SMR(L;S) &= \frac{R_L^{(L)}}{R_S^{(L)}} = \frac{D_L/N_L}{R_S^{(L)}} = \frac{872/73967}{12.29 \times 10^{-3}} = 0.96 \end{aligned}$$

Again, the standardization calculation has reversed the direction of the difference compared with the US population, in contrast to the crude rate ratios.

This illustrates the need for standardization in the analysis of such data; a comparison of crude rates does not take into account the impact of the confounding factor, race. It is not possible in the above analysis to determine whether there is statistical evidence that there is a **significant** difference between states, or between each state and the US population as a whole; to achieve this, a formal statistical analysis using hypothesis testing must be carried out.