

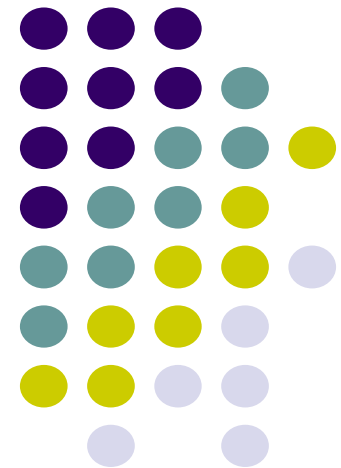


SAS Code for Calculation of the 95% Confidence Interval of Standardized Rates

Fu-Lin Wang, B.Med., MPH, PhD
Epidemiologist
Adjunct Assistant Professor

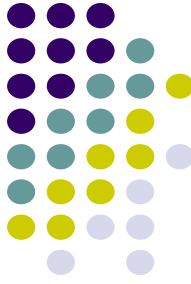
Fu-lin.wang@gov.ab.ca

Tel. (780)422-1825



Public Health Surveillance & Environmental Health Branch, AHW
Community Health Sciences, the University of Calgary

Background



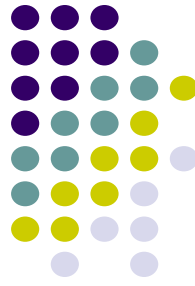
- **Rate: A measure of the frequency of occurrence of an event at a given time period. It is used to measure the magnitude of a problem.**
- **It may be expressed as**

$$\text{Rate} = \frac{\text{number of events}}{\text{population at risk}} \quad \times \quad \text{K (a multiplier)}$$

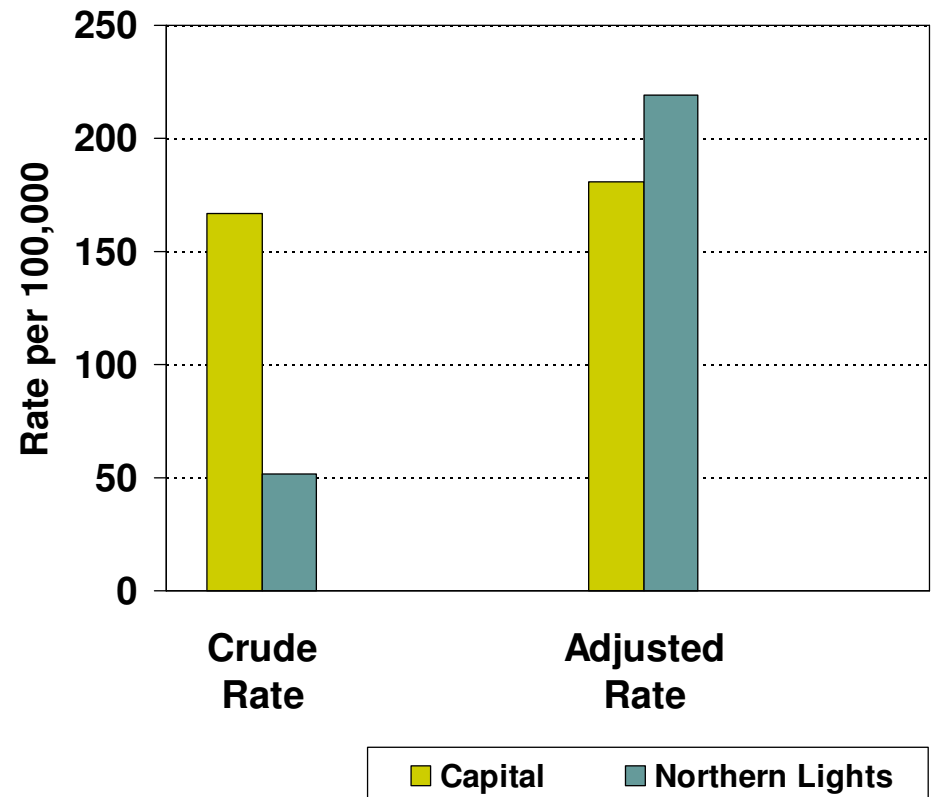
e.g., In 2001, the death rate from neoplasm is 167 per 100 000 population in the Capital Region compared with 52 in Northern Lights

- **Can we draw the conclusion that residents living in the Capital region more likely die from neoplasm?**

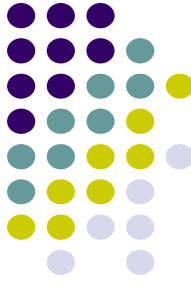
Which region has a higher rate?



- In 2001, the death rate from neoplasm in the Capital Region was 167 per 100000 population, compared to 52 in the Northern Lights Region.
- After controlling the effect of sex and age, the death rate was 181 and 219 for the Capital and Northern Lights Region, respectively.



Background



- **A standardized (adjusted) rate: a rate that is corrected for the differences in age/sex (or other confounding factors) compositions of populations being compared across regions or over time.**
- **A standardized rate is used for various comparisons (e.g., across regions, over time) and is not valid to measure the actual magnitude of a problem.**
- **There are often two ways of standardization:**
 - **The direct method: the specific rates (age/sex specific) in the study population are averaged, using as weights of the distribution of a standard population.**
 - **The indirect method: used when specific rates in the study population are unknown or unstable. The specific rates of a standard population are used as weights the distribution of the study population.**

Background



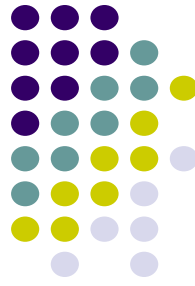
- **A statistical inference: The statistical process of passing from observations to generalizations, with calculated degrees of uncertainty.**
- **The 95% confidence interval: The computed interval with a 95% probability, i.e., the true value of a variable such as a rate, proportion, or mean contained within the interval. Thus, in statistical inference, there is a 5% degree of uncertainty.**
- **The key for estimating the 95% confidence interval is to calculate the standard error (SE) of a standardized rate.**
- **The method of calculating SE depends on the distribution of the standardized rate in populations, such as normal distribution, Poisson distribution, etc.**

Background



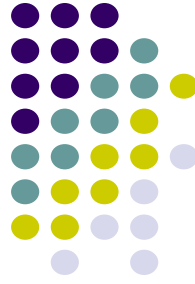
- **Several methods have been proposed for the calculation of SE for standardized rate.**
- **For surveillance purpose, we decide to use Carriere's method (Am J Epidemiol, 1994; 140:472-82).**

Major Tasks



- **Data Preparation: making sure the variables (i.e., age group, sex) to be analyzed are consistent in the type (character or numeric), length, and code values across the**
 - Numerator
 - Denominator
 - Standard population
- **Count number of events and populations by study variables (i.e., age, sex, year, region) save as data tables**
- **Prepare weights (sex/age group) of the standard population**
- **Link tables with counts of events, population and weights by sex and age groups and calculate expected number of events by sex/age group**

Data Preparation



Data Structure of Three Tables: Population, Weight, and Death

Pop_SA_RHA		WT_Can96		Case_SA_RHA	
Sex	Age_Grp	Sex	Age_Grp	Sex	Age_Grp
F	00	F	00	F	00
F	01	F	01	F	01
.....
F	19	F	19	F	19
M	00	M	00	M	00
M	01	M	01	M	01
.....
M	19	M	19	M	19
C	00	C	00	C	00
C	01	C	01	C	01
.....
C	19	C	19	C	19

Count Population and Link WT Table



```
proc sql;
```

```
create table Pop_SA_RHA as
```

```
select distinct a.year, a.sex, a.age_grp, a.rha_res,
```

```
sum(n_pop) as pop, wt_96
```

```
from pop_my as a left join WT_Can96 as b
```

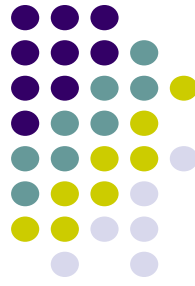
```
on      a.sex      = b.sex and
```

```
       a.age_grp = b.age_grp
```

```
order by rha_res, a.sex, a.age_grp, a.year;
```

```
Quit;
```

Count Number of Death



```
proc sql;
```

```
create table Case_SA_RHA as
```

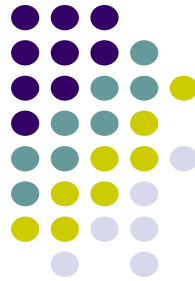
```
select distinct a.year, Grp119HSB, a.sex,  
a.age_grp, a.rha_res,
```

```
Count(*) as N_death
```

```
from Death_8303
```

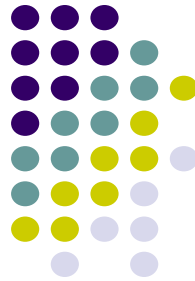
```
Group by rha_res, sex, age_grp, Grp119HSB, year  
order by rha_res, sex, age_grp, Grp119HSB, year;
```

Calculating Sex-Age Specific Rate



```
proc sql;
create table Rate_SA as
select distinct a.year, Grp119HSB, a.sex, a.age_grp, a.rha_res,
      pop, N_death, wt_96,
      N_death/pop*100000 as R_Age format=10.1,
      (wt_96/100)*N_death/pop*100000 as SR_Age format=10.1
from Pop_SAS_RHA as a left join Case_SA_RHA as b
on      a.year      = b.year and
      a.sex        = b.sex and
      a.age_grp    = b.age_grp and
      a.rha_res    = b.rha_res and
      Grp119HSB = '01'
order by a.rha_res, a.sex, a.age_grp, Grp119HSB, a.year;
```

Major Tasks – Calculating 95%CI



- **Calculating the variance of the standardized rate by sex/age group, then sum up**
- **Calculating of SMR**
- **Calculating standard error of SMR**
- **Calculating 95%CI of SMR**

Calculating Standardized Rate (SMR) and Variance - 1



```
data R_A_SE;
  set Rate_SA;
  if N_death >= 1 then ncase = N_death;
  else ncase = 0.1; /* adj zero case for Var_case calculation */

  s_case = (ncase - ncase*ncase/pop)*(pop/(pop- 0.5));
  var_case =
  (((wt_96/100)*(wt_96/100))/(pop*pop))*s_case;
  drop wt_96;

run;
```

Calculating Standardized Rate (SMR) and Variance - 2



```
proc sql;
  create table SR_95Cla as
  select distinct Grp119HSB, a.year, a.sex, a.rha_res, pop_yr as pop,
    sum(N_Death) as N_Death,
    sum(N_Death)/pop_yr*100000 as Rate_cr format = 10.1,
    sum(SR_Age) as SMR format=10.1,
    sum(var_case) as Var_case
  from R_A_SE1 as a left join pop_all as b
    on   a.year = b.year and
        a.sex  = b.sex and
        a.rha_res = b.rha_res
  group by a.sex, a.year, a.rha_res, Grp119HSB
  order by grp119HSB, a.sex, a.rha_res, a.year;
```

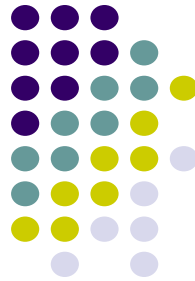
Calculating 95%CI of SMR



```
data SR_95CI (drop= var_case);
  set SR_95CIa;
  if Grp119HSB = ' ' then delete;
  SE_SMR = sqrt(Var_case) *100000;
  if SMR - (1.96*se_smr) < 0 then Low_95CI = 0.00; /* recode invalid value */
  else Low_95CI = SMR - (1.96*SE_SMR);
  Up_95CI = SMR + (1.96*SE_SMR);

format SE_SMR comma7.2
       pop comma10.
       N_Death comma6.
       Low_95CI 10.2
       Up_95CI 10.2;
label  SMR = 'Age/Sex-Standardized Rate (1/100000), Canadian Census 1996'
       SE_SMR='Standard Error of SMR'
       Rate_Cr = 'Crude Death Rate (1/100000) by Cause, Year, Sex, and RHA';
run;
```

Output to HTML format



```
ods html body='\\path\Rate_Std_HSB_chk_8604.html' style=SASweb;  
ods listing close;  
proc print data = SR_95CI noobs;  
where Grp119HSB in ('001','002','003','073') and rha_res = 'AB';  
  var Grp119HSB Year Sex RHA_Res Pop N_Death Rate_Cr SMR SE_SMR  
  Low_95CI Up_95CI;  
format  sex $sexfmt.  
        SE_SMR comma7.2  
        pop comma10.  
        N_Death comma6.  
        RHA_Res $rhafmt.  
        Low_95CI 10.2  
        Up_95CI 10.2;  
title2 'Age/Sex-Adjusted Death Rate by Selected Causes of Death, Year, and  
        Sex, Alberta, 1986-2003';  
run;  
quit;  
ods listing;  
ods html close;
```



Questions?

Pease contact:

Fu-lin.Wang@gov.ab.ca



Thank you!!