Welcome to summer of nytd!

Session starts at 12pm EST

- Please turn your video off and mute your line
- This session is being recorded
- See ZOOM Help Center for connection issues: https://support.zoom.us/hc/en-us
  - If issues persist and solutions cannot be found through Zoom contact aa17@cornell.edu

Summer of NYTD, 2018

National Data Archive for Child Abuse and Neglect
Bronfenbrenner Center for Translational Research
Cornell University

Introduction

- Summer Schedule:
  - August 8th — Introduction
  - August 15th — Data Structure
  - August 22nd — Expert Presentation I
  - **August 29th — Expert Presentation II**
  - September 5th — Linking to NCANDS & AFCARS
  - September 12th — Research Presentation I
  - September 19th — Research Presentation II

Question Format

- Generally, if you have q question please jot it down and then send it in at the end of the presentation through the chat box during the Q&A
  - However, if you have a **specific clarifying question** about what is happening in the moment, you can ask the question into the text box.

Weighting in the NYTD Outcomes Survey

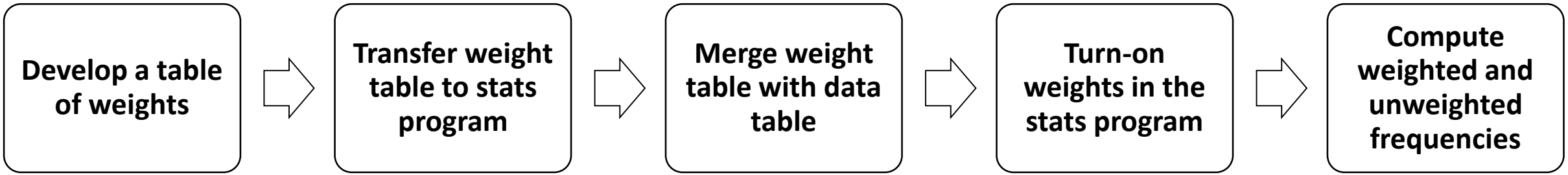
National Data Archive on Child Abuse and Neglect

August 29, 2018

Why Weighting is Done

- Who do we need info about? Kids who age out.
- The respondents to the NYTD Outcomes Survey are not a random sample of kids who are likely to age out of foster care (the population)
- They may differ in important characteristics
- These differences may bias statistical inferences about the population
- Weighting will bring the proportion of subgroups (e.g. Sex, Race) back to the proportions in the population

Today's Webinar Demonstration



Computing Weights

- Compute proportion of pop for each cell in population table
- Compute proportion of pop for each cell in respondent table
- Compute expected count for each cell in respondent table

Example Excel Table to Calculate Weights

Sex	RaceEthn	Pop	%Pop	Expected	Cohort (Observed)	Weight
1	1	6176	=C2/Population	=D2*Cohort	3601	=E2/F2
1	2	4973	=C3/Population	=D3*Cohort	2499	=E3/F3
1	3	175	=C4/Population	=D4*Cohort	104	=E4/F4
1	4	122	=C5/Population	=D5*Cohort	65	=E5/F5
1	5	37	=C6/Population	=D6*Cohort	18	=E6/F6
1	6	563	=C7/Population	=D7*Cohort	342	=E7/F7
1	7	2895	=C8/Population	=D8*Cohort	1369	=E8/F8
1	99	268	=C9/Population	=D9*Cohort	89	=E9/F9
2	1	5676	=C10/Population	=D10*Cohort	3315	=E10/F10
2	2	4282	=C11/Population	=D11*Cohort	2236	=E11/F11
2	3	208	=C12/Population	=D12*Cohort	119	=E12/F12
2	4	129	=C13/Population	=D13*Cohort	61	=E13/F13
2	5	44	=C14/Population	=D14*Cohort	14	=E14/F14
2	6	522	=C15/Population	=D15*Cohort	286	=E15/F15
2	7	2763	=C16/Population	=D16*Cohort	1383	=E16/F16
2	99	271	=C17/Population	=D17*Cohort	96	=E17/F17
		=SUM(C2:C17)	=C18/Population	=SUM(E2:E17)	=SUM(F2:F17)	

SPSS syntax

```
* Encoding: UTF-8.
GET
  FILE='D:\Temp\Outcomes_C11W3v2.sav'.
DATASET NAME NYTD.
frequencies Wave.
* It's only the FY11 Cohort that is going to get weights, so we select for that group and exclude the rest of the file.
select if ( Wave = 1 and FY11Cohort = 1 ).
frequencies Wave.
save outfile = 'D:\Temp\Outcomes_C11_Weighted.sav'.

GET DATA /TYPE=XLSX
  /FILE='Q:\NDACAN\Holdings\Data\NYTD\Summer of NYTD\Session 4 - Weights\Weights_SexRace.xlsx'
  /SHEET=name 'SexRace'
  /CELLRANGE=full
  /READNAMES=on
  /ASSUMEDSTRWIDTH=32767.
EXECUTE.
```

SPSS syntax (cont'd)

```
* Name the Weights Table.
DATASET NAME Weights WINDOW=FRONT.
dataset activate Weights.
formats Weight (F8.3).

* Reminder: Clear the last record.

DATASET ACTIVATE NYTD.
STAR JOIN
  /SELECT t0.Wave, t0.StFIPS, t0.St, t0.Recnumbr, t0.Repdate, t0.DOB, t0.Amiakn, t0.Asian,
    t0.Blkafram, t0.Hawaiipi, t0.White, t0.Raceunkn, t0.Racedcln, t0.Hisorgin, t0.OutcmRpt,
    t0.OutcmDte, t0.OutcmFCS, t0.CurrFTE, t0.CurrPTE, t0.EmplySklls, t0.SocSecrty, t0.EducAid,
    t0.PubFinAs, t0.PubFoodAs, t0.PubHousAs, t0.OthrFinAs, t0.HighEdCert, t0.CurrEnroll, t0.CnctAdult,
    t0.Homeless, t0.SubAbuse, t0.Incarc, t0.Children, t0.Marriage, t0.Medicaid, t0.OthrHlthIn,
    t0.MedicalIn, t0.MentlHlthIn, t0.Prescripin, t0.Baseline, t0.FY11Cohort, t0.Elig19, t0.Elig21,
    t0.SampleState, t0.InSample, t0.Responded, t0.Race, t0.StFCID
   , t1.Weight
  /FROM * AS t0
  /JOIN 'Weights' AS t1
    ON t0.RaceEthn=t1.RaceEthn
    AND t0.Sex=t1.Sex
  /OUTFILE FILE=*.
```

SPSS syntax (cont'd)

```
* Custom Tables.
CTABLES
  /VLABELS VARIABLES=RaceEthn Sex Weight DISPLAY=BOTH
  /TABLE RaceEthn [C] > Sex [C] > Weight [S][MEAN, COUNT F40.0]
  /CATEGORIES VARIABLES=RaceEthn Sex ORDER=A KEY=VALUE EMPTY=INCLUDE.

save outfile = 'D:\Temp\Outcomes_C11_Weighted.sav'.

FREQUENCIES CurrEnroll.
WEIGHT BY Weight.
FREQUENCIES CurrEnroll.
WEIGHT OFF.

FREQUENCIES CnctAdult.
WEIGHT BY Weight.
FREQUENCIES CnctAdult.
WEIGHT OFF.

FREQUENCIES Incarc.
WEIGHT BY Weight.
FREQUENCIES Incarc.
WEIGHT OFF.
```

Formula 1 for Weights

$$W_{ij} = \frac{E_{ij}}{O_{ij}}$$

Where *W* is the weight of cell *ij* (Sex/Race in demo)
E_{ij} is cell *ij*'s expected count (based on pop proportion)
O_{ij} is cell *ij*'s observed count (the Cohort)

Formula 2 for Weights

$$W_{ij} = \frac{\%Pij}{\%Rij}$$

Where *W* is the weight of cell *ij* (Sex/Race in demo)
$\%P_{ij}$ is cell *ij*'s proportion of the Population
$\%R_{ij}$ is cell ij's proportion of the respondents (the Cohort)

Using Info from AFCARS

- NYTD is unusual in the we know a lot of information about the non-respondents because they are in Foster Care and therefor AFCARS
- You can link the two files and pull-in weighting variables from AFCARS
- I'll demonstrate the computation of the weights
- The rest of the process is the same as with the previous demo

Questions?

- Michael Dineen—Research Support Specialist II and Manager of NYTD
  - med39@cornell.edu

Questions received in the video chat window:

- **Is the info from AFCARs what you use for the first weighting example? Where do those populations come from?**
- **When would you use AFCARS linking versus not?**
- **If the weighted frequency table is similar to the unweighted frequency table for a particular variable, which table should you utilize / report?**
- **Do any of the existing cohorts already have weights created?**
- **In your first example, you only used sex and race to create the weights, so the expected and actual numbers used to calculate the weights were fairly large. The second example only had one or two respondents in some of the categories. Does this difference impact the strength of the weights? Is there a limit to how many factors can/should be used to create a weight?**

Next week…

- Date: Wednesday September 5th from 12pm - 1pm
- Presenter: Michael Dineen, NDACAN at Cornell University
- Topic: Linking NYTD with AFCARS and NCANDS