

5.6: Computing a Cumulative Distribution (Section 4.2.2)

Let's compute a cumulative distribution for the `SleepHrsNight` variable in NHANES. This looks very similar to what we saw in the previous section.

```
# create summary table for relative frequency of different
# values of SleepHrsNight

SleepHrsNight_cumulative <-
  NHANES_unique %>%
  # drop NA values for SleepHrsNight variable
  drop_na(SleepHrsNight) %>%
  # remove other variables
  dplyr::select(SleepHrsNight) %>%
  # group by values
  group_by(SleepHrsNight) %>%
  # create summary table
  summarize(AbsoluteFrequency = n()) %>%
  # create relative and cumulative frequencies
  mutate(
    RelativeFrequency = AbsoluteFrequency / sum(AbsoluteFrequency),
    CumulativeDensity = cumsum(RelativeFrequency)
  )

kable(SleepHrsNight_cumulative)
```

SleepHrsNight	AbsoluteFrequency	RelativeFrequency	CumulativeDensity
2	9	0.00	0.00
3	49	0.01	0.01
4	200	0.04	0.05
5	406	0.08	0.13
6	1172	0.23	0.36
7	1394	0.28	0.64
8	1405	0.28	0.92
9	271	0.05	0.97
10	97	0.02	0.99
11	15	0.00	1.00
12	17	0.00	1.00

This page titled [5.6: Computing a Cumulative Distribution \(Section 4.2.2\)](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Russell A. Poldrack](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.