Exam 2 Sample

STA209-04: Applied Statistics

April 5, 2019

Formulas

Statistic	Standard Error
$\hat{p}$	$\sqrt{\frac{p(1-p)}{n}}$
$\bar{x}$	$\frac{\sigma}{\sqrt{n}}$
$\hat{p}_1 - \hat{p}_2$	$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
$\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\bar{x}_d$	$\frac{\sigma_d}{\sqrt{n_d}}$

Other Formula(s)

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i}$$

Confidence Level	80%	90%	95%	99%
z	1.282	1.645	1.960	2.576
$t_{df=5}$	1.476	2.015	2.571	4.030
$t_{df=10}$	1.372	1.812	2.228	2.764
$t_{df=15}$	1.341	1.753	2.131	2.602
$\chi^2_{df=1}$	1.640	2.710	3.840	6.630
$\chi^2_{df=2}$	3.220	4.610	5.990	9.210
$\chi^2_{df=3}$	4.640	6.250	7.810	11.340
$\chi^2_{df=4}$	5.990	7.780	9.490	13.280

1) [30 pts] In the late 1980s, the University of California recruited over 3,000 school-aged children for a study on the effects of ground-level ozone on the prevalence of asthma. Recruited children had no history of asthma, and were recruited from schools in 12 different southern California communities. Researchers followed the recruited children for five years and recorded those children who were medically diagnosed with asthma.

The following tables summarize demographic data for 1571 children from 5 of the 12 communities sampled.

Table 1: Observed counts of white children, male children and children from families with income greater than or equal to $50,000 within each community.

Community	Total	N White	N male	N with family income $\geq$ **$50,000**
Alpine	298	250	148	112
Long Beach	325	123	156	101
Riverside	369	167	174	79
Santa Maria	300	139	144	39
Upland	279	194	138	183

Table 2: Average and standard deviation of daily ozone concentration measurements within each community. The last column provides the number of measurements over which the average or standard deviation was computed.

Community	Average ozone (ppb)	Standard deviation	N days measured
Alpine	48.7	10.4	15
Long Beach	18.3	6.3	16
Riverside	34.0	6.7	16
Santa Maria	18.4	5.6	15
Upland	31.5	8.9	16

Table 3: Asthma diagnoses by sex for each community.

Community	Asthma (male)	Asthma (female)
Alpine	16	9
Long Beach	12	6
Riverside	17	10
Santa Maria	15	14
Upland	18	12

a) The US Environmental Protection Agency (EPA) states that an average ozone exposure of 40 ppb may have a detrimental effect on one's health. According to Table 1, the average ozone exposure in Alpine over a 15 day period was 48.7. Since this estimate is above the EPA threshold, can we conclude that the health of Alpine children is at risk? If so, why? If not, what can we do (statistically) to make this determination?

b) Suppose that we were interested in computing a confidence interval for the average ozone concentration in Riverside. Assuming a fixed confidence level, how would the width of the interval compare between treating the provided standard deviation as a sample estimate versus the true population parameter?

c) Among the five communities, Riverside and Santa Maria appear to be the least affluent in that they have the lowest proportions of children from families with incomes greater than $50,000. Is it reasonable to conclude that the these two communities are equal in their lack of wealth? Explain and justify your response.

d) Regardless of your answer to the previous question, are the ozone levels between Riverside and Santa Maria different from one another? Explain and justify your response.

e) Suppose that we were interested in comparing the average ozone concentrations between each of the <u>twelve</u> communities sampled for this study. One approach to this kind of analysis would be to perform a statistical test of the difference in averages for each possible pair of communities. What concerns, if any, would you have with this approach?

f) Construct a table describing the counts of children with and without asthma within each community at the end of the five-year followup period.

g) Is there a relationship between community and asthma diagnoses? Explain and justify your response.

h) Based on your response to the previous question, does the affluence of a community appear to have a role?