hw1

Mohammad Mottahedi

January 13, 2016

Problem 1 (16 points). Which scale of measurement is most appropriate for the following variables? (a) Treatment group (treatment or control) in a clinical trial?

answer: Binary variable.

(b) Age range in a clinical trial?

answer: Interval variables.

(c) Geographic location (North America, Southeast Asia, Europe)?

answer: Nominal.

(d) Michelin stars?

answer: Ordinal.

Problem 2 (16 points). Suppose in a certain area, the number of ships per 10 by 10 kilometer region of the ocean has a Poisson distribution with parameter 0.2.

(a) What is the probability of observing 5 ships in one 10 by 10 kilometer region?

$p(y) = \frac{e^{-\mu}\mu^y}{y!} = \frac{e^{-0.2}0.2^5}{5!} = 2.183282 \times 10^{-6}, \mu = 0.2, y = 5$

(b) What is the probability of observing no more than 5 ships in a 100 by 100 kilometer region?

There are 100 (independent) 10x10 km^2 region in a 100x100 km^2 region, number of ships in each of the 100 regions has possion distrubiton $\mu = 0.2$ (X_i). The sum $X_1 + ... + X_{100}$ has the poisson distribuiton with mean $\mu_{total} = \mu_1 + ... + \mu_{100} = 20$.

$p(y) = 7.1906779 \times 10^{-5}$

Problem 3 (18 points). When the 2000 General Social Survey asked subjects whether they would be willing to accept cuts in their standard of living to protect the environment, 344 of 1170 subjects said Yes. Note: Some helpful R and SAS code is provided on ANGEL. See hw1. R and hw1_ help. SAS .

(a) Estimate the population proportion who would say Yes.

estimated population proportion: $\hat{p} = \frac{344}{1170} = 0.2940171$

estimated standard error: $\hat{SE} = \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 0.0133196$

(b) Conduct significance test to determine whether a majority of the population would say Yes. Report and interpret the p-value.

$H_0 : p <= 0.5$

$H_A : p > 0.5$

$Z = \frac{\overline{p}-0.5}{\hat{SE}} = \frac{0.02940171-0.05}{0.0133196} = -15.4646461$

$p - value = 3.005134 \times 10^{-54} < 0.05$

(c) Construct and interpret 99% confidence interval for the population proportion who would say Yes. Here you should construct a classical confi- dence interval, i.e. based on Wald test; later we will see some problems with this approach

$\hat{p} = \frac{344}{1170} = 0.2940171$

$\hat{SE} = \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 0.0133196$

(0.2630311 , 0.3250031)