Chp 14 | Introduction to Inference

Statistical Inference uses probability theory to know how trustworthy our conclusions are.

☆ Statistical Inference: drawing conclusions about a population from sample data

Confidence Interval (CI)

level C CI for a parameter
① is calculated from the data
estimate ± margin of error
② level C gives probability that parameter is in that interval in repeated samples

Simple Conditions for Inference about a mean

① We have an SRS from population. There is no nonresponse or other difficulty.
② population distribution is $N(\mu, \sigma)$
③ We don't know μ, but we do know σ.
* these will be assumed true for our problems, even though they're unrealistic

CI for mean of Normal population

CI for μ: $\left[\bar{x} - \frac{z^*\sigma}{\sqrt{n}}, \bar{x} + \frac{z^*\sigma}{\sqrt{n}}\right]$

z^* is critical value z-score corresponding to probability we want
level: 90% $z^* = 1.645$; 95% $z^* = 1.960$ 99% $z^* = 2.576$

Ex1 Here is sample data from population distribution w/ $N(\mu, 0.2)$. 5.32 4.88 5.1 4.73 5.15 4.75
Give 90% CI for μ.

(more space on next page)

Chp 14 (cont)

Ex1 (cont)

Ex2 A student reads that a 95% CI for BMI of young American women is 26.8 ± 0.6. The student claims it means "95% of all young women have BMI between 26.2 + 27.4." Is this correct? Explain.

Chp 14 (cont)

Tests of Significance

☆ an outcome that would rarely happen if a claim were true is good evidence that the claim is false

Vocab: H_0 = null hypothesis = claim tested by a statistical test; usually a statement of "no effect" or "no difference". (ex $H_0: \mu = 0$)

H_a = alternative hypothesis = claim about population; (must exclude H_0, i.e. must be different from H_0)

one-sided H_a: parameter is larger or smaller than H_0 value (ex: $H_a: \mu < 0$)

two-sided H_a: parameter different from H_0 value (ex $H_a: \mu \neq 0$)

null hypothesis like legal (criminal) hypothesis that a person is "innocent until proven guilty"

H_0: innocence

H_a: not innocent

- lawyers must provide convincing evidence against H_0

Ex3

Women's heights. The average height of 18-year-old American women is 64.2 inches. You wonder whether the mean height of this year's female graduates from your local high school is different from the national average. You measure an SRS of 78 female graduates and find that $\bar{x} = 63.1$ inches. What are your null and alternative hypotheses?

Ex 4

Grading a teaching assistant. The examinations in a large accounting class are scaled after grading so that the mean score is 50. The professor thinks that one teaching assistant is a poor teacher and suspects that his students have a lower mean score than the class as a whole. The TA's students this semester can be considered a sample from the population of all students in the course, so the professor compares their mean score with 50. State the hypotheses H_0 and H_a.

Chp 14 (cont)

test statistic: calculated from sample data; measures how far data is different from what we'd expect if H_0 were true; large values of test statistic
$\Rightarrow$ data not consistent w/ H_0.

p-value: probability that test stat would be as extreme or more extreme than actual observed value, assuming H_0 true; small $p \Rightarrow$ stronger evidence against H_0; large $p \Rightarrow$ fail to reject H_0

☆ failing to find evidence against H_0 does not imply H_0 is true

Statistically significant at level α: if p-value is as small or smaller than α, then data is statistically significant at level α.

☆ "significant" means "not likely to happen by chance."

z-test for population mean
Draw SRS, size n, from $N(\mu,\sigma)$ population (assume σ known). To test $H_0: \mu = \mu_0$, calculate

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$ (one sample z-statistic)

① $H_a: \mu > \mu_0$ $\quad p =$ probability that random variable $\geq z$

② $H_a: \mu < \mu_0$ $\quad p =$ " " " $\leq z$

③ $H_a: \mu \neq \mu_0$ $\quad p = 2$ times probability that random variable $\geq |z|$.
(i.e. $rv \geq z$ or $rv \leq -z$)

Chp 14 (cont)

Ex 5 Find the values of the one-sample z statistic.

(a) $H_0: \mu = 0$ random variable $X \sim N(0, 60)$
$H_a: \mu > 0$ get $\bar{x} = 6.1$ from a sample of size 8

$z = ?$
p-value = ?

(b) $H_0: \mu = 48$ random variable $X \sim N(48, 25)$
$H_a: \mu \neq 48$ get $\bar{x} = 58$ from a sample of size 100

$z = ?$
p-value = ?

Chp 14 (cont)

Ex 6 A test of $H_0: \mu = 1$, $H_a: \mu \neq 1$ has test statistic $z = 1.776$. Is this test significant at the 5% level ($\alpha = 0.05$)? Is it significant at 1% level ($\alpha = 0.01$)?

Ex 7

#14.23 pg 383