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 Random numbers come from population w/ $\mu = 0.5$ & $\sigma = 0.2887$. A command to generate 100 random #s gives outcomes w/ $\bar{x} = 0.4365$.

Test $H_0: \mu = 0.5$, $H_a: \mu \neq 0.5$

(a) z test statistic?

(b) (use Table C) Is z significant at $\alpha = 0.05$ level?

At $\alpha = 0.01$ level?

(c) Does the test give good evidence against the H_0?

Chp 15 Thinking about Inference

We must be concerned w/ whether or not simple conditions are met. We can overcome not knowing σ. But we have to check if ① population distribution is reasonably normal and ② we have an SRS!

★ "mathematical theorems are true; statistical methods are effective when used with judgement."

② The deliberate use of chance means laws of probability apply to outcomes ⇒ statistical inference can be useful. To decide if it's SRS, consider these:

(a) Problems w/ nonresponse or dropouts from experiments

(b) Need different methods for different experiment designs. (z procedures don't work for more complicated experiments that are not SRS).

(c) There is no cure for flaws like voluntary response surveys or uncontrolled experiments.

① This condition is less essential; because z procedures are based on normality of sample mean $\bar{x}$.

Ex 1 If you observe 50 consecutive shoppers at a grocery store, recording how much each shopper spends, is it reasonable to assume this is SRS? Why or why not?

Chp 15 (cont)

Ex 2

Rate that movie. A professor interested in the opinions of college-age adults about a new hit movie asks the 25 students in her course on documentary filmmaking to rate the entertainment value of the movie on a scale of 0 to 5. Which of the following is the most important reason why a confidence interval for the mean rating by all college-age adults based on these data is of little use? Comment briefly on each reason to explain your answer.

(a) The course is small, so the margin of error will be large.

(b) Many of the students in the course will probably refuse to respond.

(c) The students in the course can't be considered a random sample from the population of all college-age adults.

---

margin of error (for CI)

$$m = \frac{z^* \sigma}{\sqrt{n}}$$

this gets smaller when

1. z^* gets smaller (i.e. lower CI)
2. σ is smaller.
3. n gets larger; large samples allow more precise estimates for μ.

If we want a certain m value, $n = \left(\frac{z^* \sigma}{m}\right)^2$ choose n accordingly

* m in CI covers only random sampling errors. (it doesn't take other difficulties into account)

M1070

Chp 15 (cont)

Ex 3 SRS from 654 women for BMI, with $\bar{x} = 26.8$. We know $\sigma = 7.5$.

(a) Give three CI for μ using 90%, 95%, and 99% CIs.

(b) What are margins of error for those CIs?

(c) Suppose we had SRS of only 100 women. What is m for that 95% CI?

(d) What is m for 95% CI if n = 400? n = 1600?

Chp 15 (cont)

How small does p-value have to be to have convincing evidence against H_0? Consider:

(i) How plausible is H_0? If H_0 is something people believe to be true, then strong evidence (small p-value) is needed to reject H_0.

(ii) Consequences of rejecting H_0 – ex if rejecting H_0 is financially or personally costly, you want strong evidence (small p) to reject H_0.

☆ evidence against H_0 is stronger when H_a is one-sided (because its based on data plus information about direction of deviations from H_0).

Sample size affects statistical significance

- n large ⇒ very small population effects can be highly significant (since $\sigma/\sqrt{n}$ is so small)
- n small ⇒ large population effects can fail to be significant (since $\sigma/\sqrt{n}$ is large so $\bar{x}$ distribution is spread out).
- statistical significance ≠ practical significance

Ex 4 For same BMI study as in Ex 3, $\sigma = 7.5$. How large a sample do we need to estimate BMI mean μ to w/in ± 1 w/ 95% CI?

Chp 15 (cont)

★ higher power gives better chance of detecting effect when it's really there

Power

power of a test against a specific H_a is the probability that we'll reject H_0 at chosen α-value when alternative value of parameter is true.

(having high power means it's highly sensitive to deviations from H_0.)

How large should n be?

- for small α, need larger n.
- for high power, need larger n.
- 2-sided H_a needs bigger n than 1-sided H_a.
- detecting small effect requires bigger n than detecting large effect

Truth (population)

sample conclusion	H_0 true	H_a true	
	Type 1 error	correct	reject H_0
	correct	Type 2 error	fail to reject H_0

Type 1 error: reject H_0 when it's true

Type 2 error: fail to reject H_0 when H_a true

$P(\text{Type 1 error}) = \alpha$

$P(\text{Type 2 error}) = 1 - \text{power}$

Ex 5 Taking a pregnancy test at home produces a result. Give H_0 & H_a. Describe Type 1 & Type 2 errors in terms of "false positive" and "false negative."