What technique is appropriate?

Look at Chapter 21, Part III Review, page 540-541.

Which type of parameter -- mean or proportion?

- Look at words
- Look at original data
  - Original data are typical numbers: mean.
  - Original data are "yes/no" "male/female" "succeed/fail" or other categorical variable with two values: proportion.

Mean – which type of problem?

- One sample
  - Population standard deviation (sigma) known: $\bar{X}$ has normal dist'n. Chapters 14-15
  - Pop'n st dev unknown and estimated by sample st dev: $\bar{X}$ has t-dist'n. Ch. 17
- Two samples
  If comparing two sample means, then which design? (See problems 18.1-18.4)
  - Matched pairs (p. 444-447. Problems in Ch. 17. A few in Ch. 18.)
  - Two independent samples (Most problems in Ch. 18)

Proportions – which type of problem?

- One sample
- Two sample

Straightforward types of questions to answer:

- Find the probability that … (General: Ch. 11. Sample mean: Chapter 11. Sample proportion. Chapter 19. p. 502-503)
- Find the sample size needed .. (estimate mean: Chapter 14. p. 355-356. estimate proportion: Chapter 19 p. 510-511)
- "Estimate the (mean or proportion)" – Form a confidence interval and interpret it in the context of the problem.
- "Does the data provide significant evidence that …?" or "Do the data provide good evidence that …?" -- Perform a hypothesis test and interpret it in the context of the problem.

Other types of questions to answer:

- Recognize the different techniques needed for "Measure the effect of the new treatment" and "Does the new treatment have an effect?"
- Use a confidence interval to find the conclusion of a two-sided hypothesis test. (Ch. 14, exercises 57 and 58)
- If a result is statistically significant, must it also be practically significant? Make up an example to illustrate this and discuss it. (Page 402-403)
- If a result is practically significant, must it also be statistically significant? Make up an example to illustrate this and discuss it. (pages 402-403)
- What is the difficulty about doing multiple analyses at once? (pages 403)
- When choosing a significance level for a hypothesis test, what considerations might lead a statistician to use some level besides 0.05? What considerations lead them to use 0.05? (Type I erroe and Type II error. pages 409-410.)
  If the Type I error is worse, use small alpha, such as 0.01.
  If the Type II error is worse, use a relatively large alpha, such as 0.10.
  If both are equally bad, use a medium-sized alpha, such as 0.05.)
- What are some advantages of reporting the result of a hypothesis test just by the p-value instead of "reject the null hypothesis " or "fail to reject the null hypothesis"? What are the disadvantages of reporting it in that way? (Read the conclusion of Example 14.9 on page 379)
- Give an example where these two statements of a conclusion would convey something different to the reader: "Fail to reject the null hypothesis" and "Accept the null hypothesis."

Conditions (Where to look)

It is crucial that YOU be able to check the conditions for any statistical technique you perform. Use this summary to help you prepare your notes on what conditions are needed for the techniques covered in Chapters 14-21.

One mean. If we DO know the pop'n st dev, use normal table (*z*) and the work from chapters 14 and 15

Conditions: Simple conditions on p. 360, EXCEPT that we recall pictures for Central Limit Theorem from Ch 11 and **modify condition 2** as follows:

Pop'n normal – the dist'n of $\bar{X}$ is normal for all sample sizes.

Pop'n strongly skewed – the dist'n of $\bar{X}$ is normal if *n* at least 25

Pop'n not strongly skewed – the dist'n of $\bar{X}$ is normal for smaller *n*, maybe as small as ten or so.

One mean. If we don't know pop'n standard deviations, then $\bar{X}$ has *t* dist'n. Ch. 17. Conditions: Read and learn these on p.458

Two means – Two samples, from a matched pairs design. Ch. 17

Method: Find differences and analyze those differences using one-sample techniques. Use $\bar{X}$ of differences, which has t dist'n.

Conditions: Same as for one-sample mean.

Two means - Two independent samples: Since we don't know pop'n standard deviations, then $\bar{X}_A - \bar{X}_B$ has *t* dist'n. Ch. 18. Conditions: See p 483 (Pay most attention to the second paragraph: the sum of the two sample sizes.)

One proportion: The dist'n of $\hat{p}$ is normal (*z*). Ch. 19

Conditions: Start on p 505. Two conditions:

(1) SRS

(2) sample size – depends on what technique. See pages listed

a. large sample confidence interval p 505

b.plus-four confidence interval p 508

c. hypothesis tests p. 513

Two proportions: The dist'n of $\hat{p}_B - \hat{p}_A$ is normal (*z*). Ch. 20

Conditions:

(1) SRS (Two independent separate SRS's from two pop'ns OR one SRS from one pop'n, which is then divided into two groups. Described on p. 527 (first whole paragraph.))

(2) sample size – depends on what technique. See pages listed.

a. Large sample confidence interval p. 526

b. Plus-four confidence interval p. 528

c. Hypothesis tests p. 532