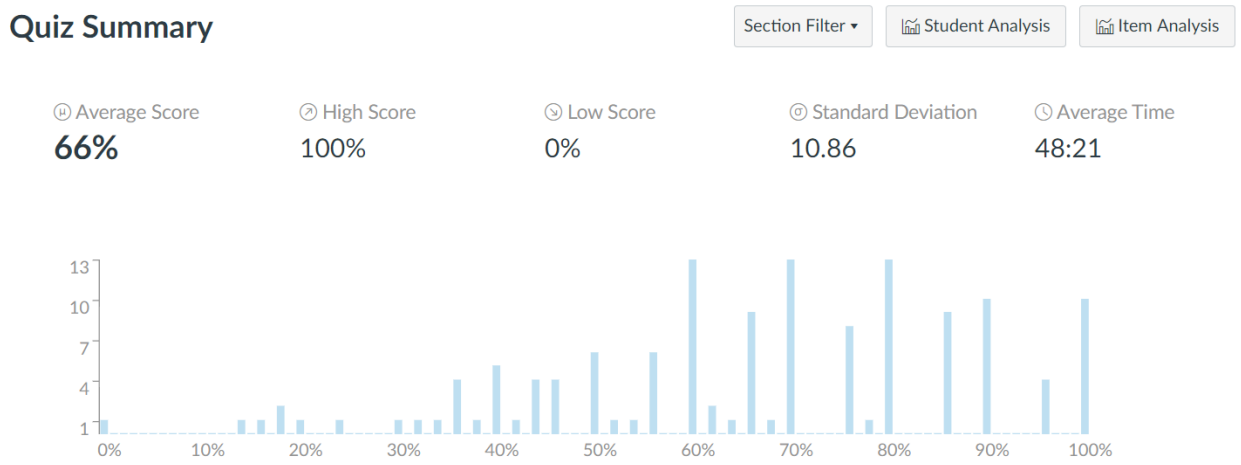
Packet 4: Interval Estimation

Chap 7.4 Sample Size Calculation

In statistical consulting, a common question is "how many samples are needed?"

The answer is often like: how precise you want? As sample size increases, we get more accurate result. However, collecting a large amount of samples costs time and money.

Example 1: Suppose that $X_1, \ldots, X_n$ are are final exam scores i.i.d. from $N(\mu, \sigma^2)$, σ^2 is known, e.g. $\sigma = 11$ from past exams.



We are interested in estimating μ, and want to be fairly confident than $\bar{X} \pm 5$ contains μ, then how many students do we need?

If σ^2 is known, a $100(1-\alpha)\%$ confidence interval for μ is $\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Sample size for μ

The sample size necessary for estimating a population mean μ with $100(1-\alpha)\%$ confidence and error no larger than ϵ is:

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{\epsilon^2}.$$

Since sample size n needs to be an integer, we round it up.

When σ^2 is unknown, we often need to estimate σ^2 by sample variance S^2 from a preliminary data,

$$n = \frac{z_{\alpha/2}^2 S^2}{\epsilon^2}.$$

Example 2: assume the length of trout in Spring Creek follows $N(\mu, \sigma^2)$,, where σ^2 is unknown. From 29 fish, $S^2 = 34.9$. To estimate μ within error $\epsilon = 0.5$ at 95% confidence, what is the minimum sample size n?

Factors that affect the sample size:

Sample sizes in clinical trails:

Phase 1 Screening for safety. Testing within a small group of people (20–80) to evaluate safety, determine safe dosage ranges, and begin to identify side effects. A drug's side effects could be subtle or long term, or may only happen with a few people, so phase 1 trials are not expected to identify all side effects.

Phase 2 Establishing the efficacy of the drug, usually against a placebo. Testing with a larger group of people (100–300) to see if it is effective and to further evaluate its safety. The gradual increase in test group size allows less-common side effects to be progressively sought.

Phase 3 Final confirmation of safety and efficacy. Testing with large groups of people (1,000–3,000) to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow it to be used safely.

Phase 4 Safety studies during sales. Post-marketing studies delineate additional information, including the treatment's risks, benefits, and optimal use. As such, they are ongoing during the drug's lifetime of active medical use. (Particularly relevant after approval under FDA Accelerated Approval Program)

Sample size for p

Suppose $X_1, \ldots, X_n$ are i.i.d. Bernoulli trial with success probability p. $X_i = 0$ (failure) or 1 (success) and $P(X_i = 1) = p$.

We know that the C.I. is $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, and want to find n that leads to small errors.

Example 3: Suppose we know that the unemployment rate is 7.9% in September 2020.

We wish to update this rate for October, and be 99% confident that the new estimated rate is within 0.001 of the true value.

Often, we do not know p and there are two solutions:

1. Guess the value of p, say p^* based on prior knowledge or use a pilot study to find p^*,
$$n = \frac{z^2_{\alpha/2} p^*(1-p^*)}{\epsilon^2}.$$
2. We know that when $p = 0.5$, the value of $p(1-p)$ is maximized,
$$n = \frac{z^2_{\alpha/2} 0.5(1-0.5)}{\epsilon^2} = \frac{z^2_{\alpha/2}}{4\epsilon^2}.$$

Example 4: To estimate the success rate p by flipping the coin n times, and we want to control $\epsilon = 0.05$ with 95% confidence.

Factors that affect the sample size:

More Examples: 7.4-1, 7.4-2, 7.4-3, 7.4-4