AP Statistics – Unit 7 Review: Inference for Means

1. Overview

Goal: Estimate or test a population mean (μ) or a difference in means ($\mu_1 - \mu_2$) using sample data.

Types:

- One-sample z/t-interval or z/t-test for μ
- Two-sample t-interval or t-test for $\mu_1 - \mu_2$
- Paired t-interval or t-test for μ_d (mean of differences)

2. When to Use $n \leq 0.1N$ Rule

The **independence condition** is checked when sampling *without replacement* from a finite population. If $n \leq 0.1N$, treat observations as independent. If data come from random assignment to treatments, independence is built in and the 10% rule is not needed.

3. Conditions

One-sample:

- Random sample / random assignment
- Independence: $n \leq 0.1N$ (if no replacement)
- Normality: Population normal OR $n \geq 30$ (CLT) OR sample roughly symmetric/no outliers

Two-sample:

- Random samples / random assignment for each group
- Independence: within and between groups ($n_1 \leq 0.1N_1$, $n_2 \leq 0.1N_2$ if no replacement)
- Normality: Both populations normal OR both $n \geq 30$ OR both samples roughly symmetric/no outliers

Paired:

- Random sample / random assignment of pairs
- Independence: Pairs independent of each other
- Normality: $n_{\text{pairs}} \geq 30$ OR roughly symmetric/no outliers

4. One-Sample Procedures

If σ known: z-test / z-interval. If σ unknown: t-test / t-interval.

$$\text{CI: } \bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

$$\text{Test: } t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

df $= n - 1$.

5. Two-Sample Procedures

Always use t-distribution (unknown σ's).

$$\text{CI: } (\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{Test: } t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

df from software or conservative $\min(n_1 - 1, n_2 - 1)$.

6. Paired Procedures

Work with differences $d = X_1 - X_2$:

$$\text{CI: } \bar{d} \pm t^* \frac{s_d}{\sqrt{n_{\text{pairs}}}}$$

$$\text{Test: } t = \frac{\bar{d} - \mu_{d,0}}{s_d/\sqrt{n_{\text{pairs}}}}$$

df $= n_{\text{pairs}} - 1$.

7. Steps for Inference

State: Define parameter, state H_0 and H_a (test) or confidence level (interval).
Plan: Choose procedure, check conditions.
Do: Calculate statistic, df, and CI or p-value.
Conclude: Interpret in context.

8. Common Pitfalls

- Forgetting 10% condition when sampling without replacement
- Using s when σ is known or vice versa
- Treating paired data as two-sample
- Not checking normality for small n

Unit 7: Procedures, Formulas, Conditions, TI-84 Steps

Procedure	Formula	Conditions	TI-84 Steps
1-sample z Test	$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	SRS, independent observations, σ known, normal population or large n	`STAT` → `TESTS` → `Z-Test`, choose `Stats` or `Data`
1-sample z Interval	$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$	Same as above	`STAT` → `TESTS` → `Z-Interval`
1-sample t Test	$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}},\ df = n-1$	SRS, independent observations, σ unknown, normal population or large n	`STAT` → `TESTS` → `T-Test`
1-sample t Interval	$\bar{x} \pm t^*_{n-1} \frac{s}{\sqrt{n}}$	Same as above	`STAT` → `TESTS` → `TInterval`
2-sample t Test	$t = \frac{(\bar{x}_1-\bar{x}_2)-\Delta_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$	Two SRSs or random assignment, independent groups, each sample $\leq 10\%$ of population, normal populations or large n	`STAT` → `TESTS` → `2-SampTTest`, choose `Stats` or `Data`, `Pooled: No`
2-sample t Interval	$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	Same as above	`STAT` → `TESTS` → `2-SampTInt`, `Pooled: No`
Paired t Test	$t = \frac{\bar{d}-\mu_{d,0}}{s_d/\sqrt{n}},\ df = n-1$	Paired design (matched pairs or before/after), analyze differences, differences nearly normal or large n	Enter differences in `L1`, then `STAT` → `TESTS` → `T-Test` using `L1`
Paired t Interval	$\bar{d} \pm t^*_{n-1} \frac{s_d}{\sqrt{n}}$	Same as above	Enter differences in `L1`, then `STAT` → `TESTS` → `TInterval` using `L1`