

One-Sample Hypothesis Tests (8.1-8.4)

1. Introduction

- (a) The value of a population parameter can be estimated from sample data as either a single number (a point estimate), or as an interval of plausible values (a confidence interval). Methods to do both of these have been previously discussed.
- (b) Sometimes, however, the objective of an investigation is not to estimate the value of a population parameter, but to decide which of two contradictory claims about the parameter is correct. This type of statistical inference is called **hypothesis testing** and is the subject of current interest.

2. **Tests of Hypotheses:** These are tests that help us decide between two contradictory claims about the value of a population parameter. They provide a method for using sample data to decide whether one hypothesis, the Null Hypothesis, should be rejected.

- (a) **Statistical Hypotheses:** These are the claims or assertions about the value of a population parameter. Initially, one claim, the Null Hypothesis, is taken to be true and then tested against a contradictory claim, the Alternative Hypothesis. The initially favored claim will not be rejected in favor of the alternative claim unless sample evidence strongly contradicts it, providing strong support for the alternative hypothesis.
- (b) **Null Hypothesis, H_0 :** One of the two possible conclusions to the experiment formulated as a statement about the value of a population parameter. H_0 typically represents the status quo or established theory. A simple example is:

$$H_0 : \mu = 7.5$$

- (c) **Alternative Hypothesis, H_A :** The other possible conclusion to the experiment is formulated as a hypothesis that contradicts H_0 . Only one of these hypotheses can be true. The Alternative hypothesis is often referred to as the research hypothesis, which can only be regarded as true when the evidence against established theory, H_0 , is sufficiently strong. Choose one of three possible alternative hypotheses for use with the above null hypothesis:

$$H_A : \mu \neq 7.5 \quad ; \quad H_A : \mu < 7.5 \quad ; \text{or} \quad H_A : \mu > 7.5$$

- (d) **Conclusion:** Only when sample data strongly suggest that μ equals something other than 7.5 should the null hypothesis be rejected. Possible conclusions from hypothesis testing are either "reject H_0 " or "do not reject H_0 ".
- (e) **Test Statistic:** A statistic calculated from the sample, sometimes in standardized form, that is used as a measure of the magnitude of departure from the Null Hypothesis. Generally, the larger the value of the test statistic, the stronger the evidence against H_0 .
- (f) **Rejection Region:** Set of all test statistic values for which H_0 will be rejected. When the value of the test statistic falls within the rejection region, H_0 will be rejected.
- (g) **pValue:** The probability of observing a value of the test statistic at least as contradictory to the Null hypothesis as that computed from the sample data. It is an indicator of the weight of evidence against the Null hypothesis.

3. Example: Suppose Jane weighs herself once a week for 12 weeks and records the following weights in pounds.

145.1 152.3 143.2 147.8 149.4 147.2
 151.7 146.3 149.3 150.2 151.2 152.7

Assume her weight follows a normal distribution with a known standard deviation, $\sigma = 3$. Last year Jane's mean weight was 147.6 lbs. Has she gained weight such that her mean weight is greater this year than last year? Perform a hypothesis test to answer this question.

$$H_0 : \mu = 147.6$$

$$H_A : \mu > 147.6$$

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

$$pVal = P(Z \geq z_0) = 1 - \Phi(z_0)$$

- $z_0 = \frac{\bar{x}-147.6}{3.0/\sqrt{12}} = 1.462621$
 - $pValue = 1 - \Phi(1.463) = 0.0718$
 - As this $pval > .05$ we would not have strong evidence to reject H_0 . So we conclude that the sample evidence is not strong enough to reject the Null hypothesis that her mean weight is 147.6 lbs, at significance level $\alpha = .05$. However, at significance level $\alpha = .10$, H_0 would be rejected.
4. **Errors in Hypothesis Testing:** Two common errors that arise in hypothesis testing are called Type I and Type II errors. Both require some consideration when designing a hypothesis test.

- **Type I Error:** The error of rejecting H_0 when it is actually true.
 - **Significance Level, α :** Tail area to the right of a critical test statistic value that identifies start of the rejection region. This is the probability of committing a type I error.
 - $\alpha = P(\text{Type I Error}) = P(\text{rejecting } H_0 \text{ when it is true})$
 - For test statistic Z : $\alpha = P(Z \leq c)$ when $Z \sim N(0, 1)$
 - From Table A.3: When $\alpha = .05$ then $c = z_{.05} = 1.645$.
 - When the value of α is chosen, c is fixed.
 - Type I error can be managed by selecting the value of α to be used for the test.
- **Type II Error:** The error of not rejecting H_0 when it is false.
 - $\beta = P(\text{Type II Error}) = P(\text{ not rejecting } H_0 \text{ when it is not true})$
 - Type II error can be managed at a chosen value of α , by changing the sample size, n .
 - Increasing n , reduces β . Although mentioned in the text, we will not perform these calculations in this course.

5. Tests of Significance for Population Mean, μ

- (a) **Significance Tests:** These hypothesis tests examine whether an observed difference between a population mean, μ , and a chosen numerical value, μ_0 , is statistically significant. The question is "Could the observed difference have reasonably occurred "just by chance" from selecting a random sample". If yes, then H_0 cannot be refuted. If the observed difference could not have occurred just by chance, this is evidence that H_0 is not true, and that effect observed in the sample is present in the population, so the difference between μ and μ_0 is statistically significant.
- (b) The logic of significance testing goes like this.
- Select a statistic that measures the effect of interest.
 - Regarding the test statistic as a random variable, determine its sampling distribution under the condition that H_0 is true.
 - Compute the statistic value from the observations and compare this value with the sampling distribution, developed under the condition that H_0 is true.
 - If the statistic value, computed from the observations, falls in the main body of the distribution, it could have easily occurred just by chance alone.
 - A value falling in either tail of the distribution, would rarely occur just by chance, and is evidence that something other than chance (something significant) is operating. The *pValue* provides a quantitative measure of this.
 - Use the pValue to evaluate the strength of evidence against H_0 by comparing it with a predetermined value of the significance level, α .
 - Unless otherwise specified in the problem statement, use $\alpha = .05$ as the course default significance level.
 - For pValue $\leq \alpha$, reject H_0 .
 - For pValue $> \alpha$, do not reject H_0 .

6. Hypothesis tests for $H_0 : \mu = \mu_0$, where μ_0 is some specified value.

- (a) The test statistic used is the standardized form of \bar{x} , which is

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- (b) When H_0 is true: $\mu = \mu_0$ so that

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Note: z_0 indicates that z is being evaluated under the condition that H_0 is true.

- (c) The sampling distribution for \bar{X} , estimator for μ , is again the starting point. When the value of σ is not known, its value is estimated using the sample standard deviation, s . Using s instead of σ creates the following variations in the sampling distribution for \bar{X} .

Std. Variable	Distribution
$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$	always $\sim N(0, 1)$
$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$	approx. $\sim N(0, 1)$ for large n , ($n > 40$)
$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$	$\sim t(\nu)$ where $\nu = n - 1$ for small n , ($n < 40$)

(d) Computation of the Test Statistic for three cases of problems

Case	Conditions	Test Stat.	Distribution
1	Normal Pop. known σ	$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$	$\sim N(0, 1)$
2	σ unknown large n , ($n > 40$)	$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$	$\sim N(0, 1)$ (approx.)
3	Normal Pop. σ unknown small n , ($n < 40$)	$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$	$\sim t(\nu)$ $\nu = n - 1$

(e) Calculation of pValue for the three cases of problems

Case	H_A	Test	pValue
1 & 2	$\mu > \mu_0$	Upper tail	pVal= $1 - \Phi(z)$
	$\mu < \mu_0$	Lower tail	pVal= $\Phi(z)$
	$\mu \neq \mu_0$	Two tail	pVal= $2(1 - \Phi(z))$
3	$\mu > \mu_0$	Upper tail	pVal= $1 - F(t_0)$
	$\mu < \mu_0$	Lower tail	pVal= $F(t_0)$
	$\mu \neq \mu_0$	Two tail	pVal= $2(1 - F[t_0])$

(f) The relationship between Strength of Evidence against H_0 and pValue

pValue	Strength of Evidence against H_0
$pVal > .10$	little to none
$.05 < pVal \leq .10$	weak
$.01 < pVal \leq .05$	strong
$.001 < pVal \leq .01$	very strong
$pVal \leq .001$	very very strong

7. One-Sample Z-test(Case 2)

- (a) The manufacturer of Bic Pens claims that the mean lifetime for their pens is 245 hours of writing. To test the manufacturer's claim, a random sample of 50 pens were tested for writing lifetime. The sample mean lifetime was 246.18 hr. and the sample standard deviation was 3.60 hr. Does this sample support the manufacturer's claim at $\alpha = .05$?
- (b) Using the above data, now test:

$$H_0 : \mu = 245$$

$$H_A : \mu \neq 245$$

$$Z_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$pVal = 2P(Z \geq |Z_0|) = 2(1 - \Phi(|Z_0|))$$

- Here, $\bar{x} = 246.18$ $n = 50$ and $s = 3.60$
- $z_0 = \frac{246.18-245}{3.60/\sqrt{50}} = 2.318$
- $pValue = 2(1 - \Phi(2.318)) = 2(1 - 0.9898) = .0204 < .05$
- We have strong evidence against H_0 . So we conclude that the manufacturer's claim is not true. The mean writing lifetime of Bic Pens is not 245 hr.

8. One-Sample t-test(Case 3)

- (a) Suppose we evaluate vitamin C levels (mg/100 gm) in 8 batches of corn soy blend(CSB) from a production run and get:

26 31 23 22 11 22 14 31

- (b) Using the above data, now test:

$$H_0 : \mu = 30$$

$$H_A : \mu \neq 30$$

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$pVal = 2P(t \geq |t_0|) = 2(1 - F(|t_0|))$$

- Here, $\bar{x} = 22.50$ $n = 8$ and $s = 7.19$
- $t_0 = \frac{22.50-30}{7.19/\sqrt{8}} = -2.9504$ with $df = n - 1 = 7$
- $t_{.05,7} = 1.895$, and $|t_0| = 2.9504 > t_{.05,7}$
- $pValue < .05$
- We we have strong evidence against H_0 . So we conclude that the mean Vitamin C level in CSB is not equal to 30.

9. Properties of the t_ν - Distribution

- (a) Let t_ν denote the probability density function for a t statistic with ν degrees of freedom.
- (b) Each t_ν curve is centered on 0 and symmetric about its center.

- (c) It is slightly more spread out than the standard normal (z) curve. As ν increases, the spread of t_ν decreases.
- (d) As $\nu \rightarrow \infty$, $t_\nu \rightarrow N(0, 1)$ in shape. So the Standard Normal distribution is just the t distribution with ∞ degrees of freedom.
- (e) A Critical t value, designated $t_{\alpha, \nu}$, indicates the value of the t statistic with ν degrees of freedom with an area under the curve to the right equal to α .

10. Tests of Significance for Population Proportion, p

- (a) **Population Proportion:** The proportion of individuals in a population with a specific attribute, often labeled as S for success.

Population Proportion	Sample Proportion	Sampling Distribution
$p = X_{pop}/N$	$\hat{p} = X/n$	Small n : Binomial Large n : Normal

- The Null Hypothesis takes the form:

$$H_0 : p = p_0$$

where p_0 is a hypothesized numerical value for population proportion.

- The Alternative Hypothesis is chosen from the three possible alternative hypotheses for use with the above null hypothesis:

$$H_A : p \neq p_0 \quad ; \quad H_A : p < p_0 \quad ; \quad \text{or} \quad H_A : p > p_0$$

- **Conclusion:** Only when sample data strongly suggest that p equals something other than p_0 should the null hypothesis be rejected.
- **Large Sample Test Statistic:** The sampling distribution for \hat{p} , estimator for p , is formulated based on the sampling distribution for \hat{p} in its standardized form.

Std. Variable	Distribution
$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$	approx. $\sim N(0, 1)$

Valid for large n : $(np \geq 10)$ and $n(1 - p) \geq 10$

- When H_0 is true: $p = p_0$

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \quad \text{approx.} \quad \sim N(0, 1)$$

Valid for large n : $(np_0 \geq 10)$ and $n(1 - p_0) \geq 10$

- Calculation of pValue

H_A	Test	pValue
$p > p_0$	Upper tail	pVal= $1 - \Phi(z)$
$p < p_0$	Lower tail	pVal= $\Phi(z)$
$p \neq p_0$	Two tail	pVal= $2(1 - \Phi(z))$

(a) Example: A manufacturer of nickel-hydrogen batteries randomly selects 100 nickel plates for test cells, cycles them a specified number of times, and determines that 14 of the plates have blistered.

i. Does this provide compelling evidence that more than 10% of all plates blister under such circumstances? Use appropriate hypothesis test to reach conclusion for a .05 significance level.

- Let p denote the population proportion of nickel plates that blister under current circumstances.

$$H_0 : p = 0.10$$

$$H_A : p > 0.10$$

- This is an upper tail test.
- Can we use the normal approximation? Check to see if $np_0 \geq 10$ and $n(1 - p_0) \geq 10$, with $n = 100$ and $p_0 = 0.10$.

$$np_0 = 100(0.10) = 10$$

$$n(1 - p_0) = 100(1 - 0.10) = 90$$

so normal approximation ok.

- $\hat{p} \sim N(p, p(1 - p)/n)$, so that $Z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$
- When H_0 is true: $p = 0.10$, so the test statistic is

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

- From the sample: $n = 100$, $X = 14$ and $\hat{p} = 0.14$

$$Z_0 = \frac{0.14 - 0.10}{\sqrt{0.10(1 - 0.10)/100}} = 1.33$$

- pVal = $1 - \Phi(1.33) = 1 - 0.9082 = 0.0918 > .05$
- Do not reject H_0 at $\alpha = .05$.
- Could have committed a Type II error, since we did not reject H_0 .