Nov. 24. Hypothesis Testing

Data analysis { statistical hypothesis testing
confidence interval estimate

A statistical hypothesis is a statement about the parameters of one or more populations.

Null hypothesis: H_0 — a statement/claim to be tested that is assumed to be true until evidence to the contrary has been obtained.

Alternative hypothesis: H_1 — a statement/claim to be tested for which we are trying to find evidence

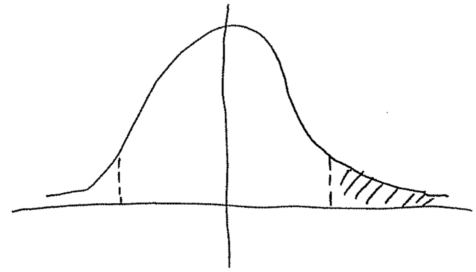
Hypothesis testing — a procedure leading to a decision about a particular hypothesis

<1> Taking a random sample
<2> Computing a test statistic from the sample
<3> making a decision about null hypothesis from test statistic

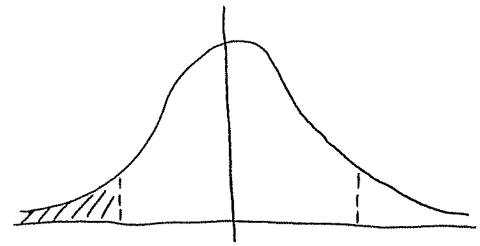
---

Consider yourself as a member of jury:
You must assume that the defendant is innocent until proven guilty.
H_0 : The defendant is innocent
H_1 : The defendant is guilty.
You must have significant evidence in order to reject H_0.

- one-sided hypothesis test

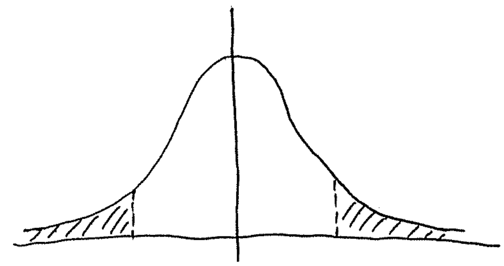


Right-tailed test $\begin{cases} H_0: \text{ parameter} = \text{some value} \\ H_1: \text{ parameter} > \text{some value} \end{cases}$



Left-tailed test $\begin{cases} H_0: \text{ parameter} = \text{some value} \\ H_1: \text{ parameter} < \text{some value} \end{cases}$

- Two-sided hypothesis test



$\begin{cases} H_0: \text{ parameter} = \text{some value} \\ H_1: \text{ parameter} \neq \text{some value} \end{cases}$

Eg1: The recommended daily allowance of sodium is 2400 mg.
A nutritionist claims that one serving of ready-to-eat soup contains more that 800 mg of sodium, on average.

$H_0: \mu = 800\text{ mg}$ null hypothesis: a statement of no difference, no change, no effect.

$H_1: \mu > 800\text{ mg}$. right-tailed test

Eg2: According to fueleconomy.gov
Volkswagen claims that their diesel vehicles have an average gam mileage of 29.5 mpG in city driving.

$H_0: \mu = 29.5\text{ MPG}$

$H_1: \mu \neq 29.5\text{ MPG}$. Two-sided test

Eg3. According to the U.S. Department of Education, of those who graduated from high school from October 2005 to October 2006, 66% were attending college in October 2006. However, Bureau of Labor statistics report that the percentage is lower.

$H_0: p = 0.66$

$H_1: p < 0.66$ left-tailed test

Type I error: rejecting the null hypothesis H_0 when ~~its~~ is true

Type II error: failing to reject the null hypothesis H_0 when it is false

$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

$\beta = P(\text{type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false})$

Decision	H_0 is True	H_0 is False
Fail to reject H_0	no error	Type II error
Reject H_0	Type I error	no error

Eg1: $H_0: \mu = 800$
$H_1: \mu > 800$

Type I error: we believe one serving of ready-to-eat soup has more than 800 mg sodium on average (rejecting the null hypothesis) when the average amount of sodium is actually 800 mg (null hypothesis is true)

Type II error: we believe one serving of ready-to-eat soup has 800 mg sodium on average (fail to reject null hypothesis) but the average amount of sodium is actually more than 800 mg (null hypothesis is false)

Eg2. $\begin{cases} H_0: \mu = 29.5 \text{ MPG} \\ H_1: \mu \neq 29.5 \text{ MPG} \end{cases}$

Type I error? we believe that the average MPG is not 29.5 (rejecting the null hypothesis) when it is actually 29.5 MPG (null hypothesis is true)

Type II error? we believe that the average MPG is 29.5 (fail to reject the null hypothesis) when it is not 29.5 MPG (null hypothesis is false)

Eg3 $\begin{cases} H_0: p = 0.66 \\ H_1: p < 0.66 \end{cases}$

Type I error? we believe that the percentage is less than 66% (rejecting the null hypothesis) when it is actually 66% (null hypothesis is true).

Type II error? we believe that the percentage is 66% (fail to reject the null hypothesis) when it is actually less than 66% (null hypothesis is false)

* Critical region & acceptance region

$$\begin{cases} H_0: \mu = 50 \\ H_1: \mu \neq 50 \end{cases}$$

random sample: $X_1, X_2, \cdots, X_n$

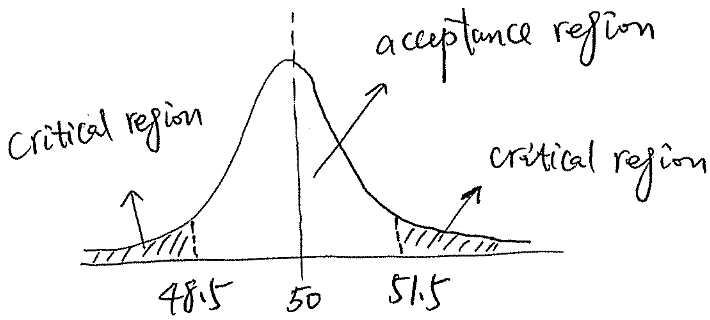
$\bar{X}$ — sampling variations

Critical region: $\bar{x} < 48.5$ and $\bar{x} > 51.5$, reject null hypothesis

Acceptance region: $48.5 \le \bar{x} \le 51.5$, fail to reject null hypothesis

Citical values: boundaries between critical region & acceptance region.

48.5, 51.5



* $\alpha = P(\text{Type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

$\beta = P(\text{Type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false})$.

$1-\beta$ is called power, which is the prob. of rejecting the null hypothesis H_0 when the alternative hypothesis is true (H_1)

Eg. How to calculate α-error, β-error, and power

A manufacturer is interested at the output voltage of a power supply for the laptop.

output voltage is assumed to be normally distributed with standard deviation 0.25 volts.

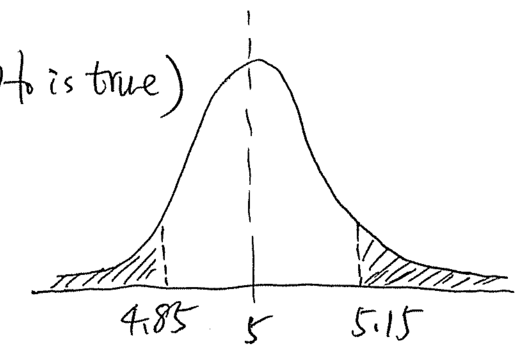
The manufacturer wishes to test the mean of output voltage

$$\begin{cases} H_0: \mu = 5 \text{ volts} \\ H_1: \mu \neq 5 \text{ volts} \end{cases}$$

collected 8 samples.

(a). If the acceptance region is $4.85 \le \bar{X} \le 5.15$, find α error

$\alpha = P(\text{type I error}) = P(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$



$\alpha = P(\bar{X} < 4.85 \text{ when } \mu = 5)$

$\quad + P(\bar{X} > 5.15 \text{ when } \mu = 5)$

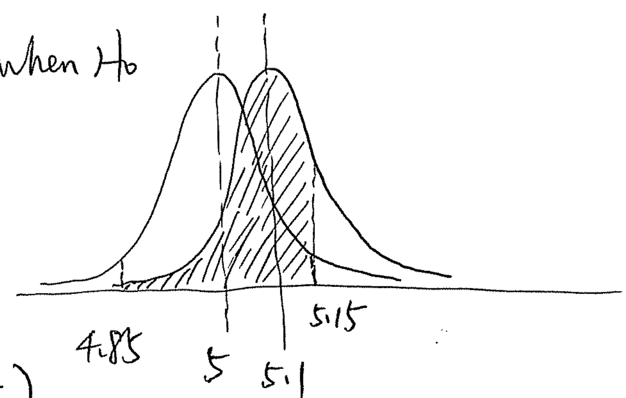
$$\alpha = P\left(\frac{\bar{X}-5}{\sigma/\sqrt{n}} < \frac{4.85-5}{\sigma/\sqrt{n}}\right) + P\left(\frac{\bar{X}-5}{\sigma/\sqrt{n}} > \frac{5.15-5}{\sigma/\sqrt{n}}\right)$$

$\alpha = P(Z < -1.7) + P(Z > 1.7)$

$= 0.04457 \times 2 = 0.08914.$

(b). If the true mean of output voltage is 5.1 volts, find power

$\beta = P(\text{type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false}).$



$\beta = P(4.85 \le \bar{X} \le 5.15 \text{ when } \mu = 5.1)$

$$= P\left(\frac{4.85-5.1}{0.25/\sqrt{8}} \le \frac{\bar{X}-5.1}{0.25/\sqrt{8}} \le \frac{5.15-5.1}{0.25/\sqrt{8}}\right)$$

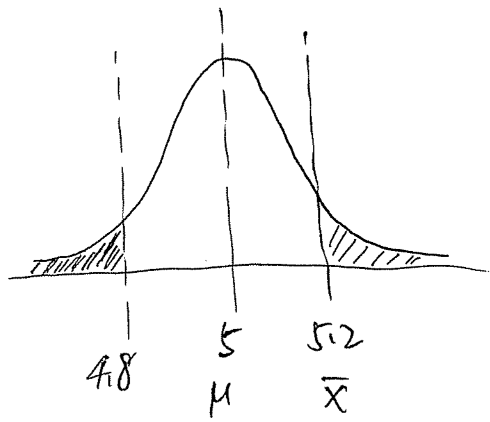
$= P(-2.83 \le Z \le 0.566)$

$= P(Z \le 0.566) - P(Z \le -2.83) = 0.71566 - 0.00233 = 0.71333$

$1-\beta = 0.2867$

(c) If the observed sample mean $\bar{x} = 5.2$, calculate the p-value

p-value: the smallest level of significance that would lead to the rejection of null hypothesis H_0 given the data.



$$\text{p-value} = 1 - P(4.8 < \bar{X} < 5.2)$$

$$= 1 - P\left(\frac{4.8-5}{0.25/\sqrt{8}} < \frac{\bar{X}-5}{0.25/\sqrt{8}} < \frac{5.2-5}{0.25/\sqrt{8}}\right)$$

$$= 1 - P(-2.26 < Z < 2.26)$$

$$= 1 - [P(Z<2.26) - P(Z<-2.26)]$$

$$= 2 \times P(Z<-2.26)$$

$$= 2 \times 0.0119$$

$$= 0.0228$$

Central limit theorem.

A random sample $X_1, X_2, \cdots, X_n$ has been taken from the population.

The population has a mean μ and Variance σ^2

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} \sim \text{Normal}\left(\mu, \frac{\sigma^2}{n}\right)$$

Tests on the mean, Variance known

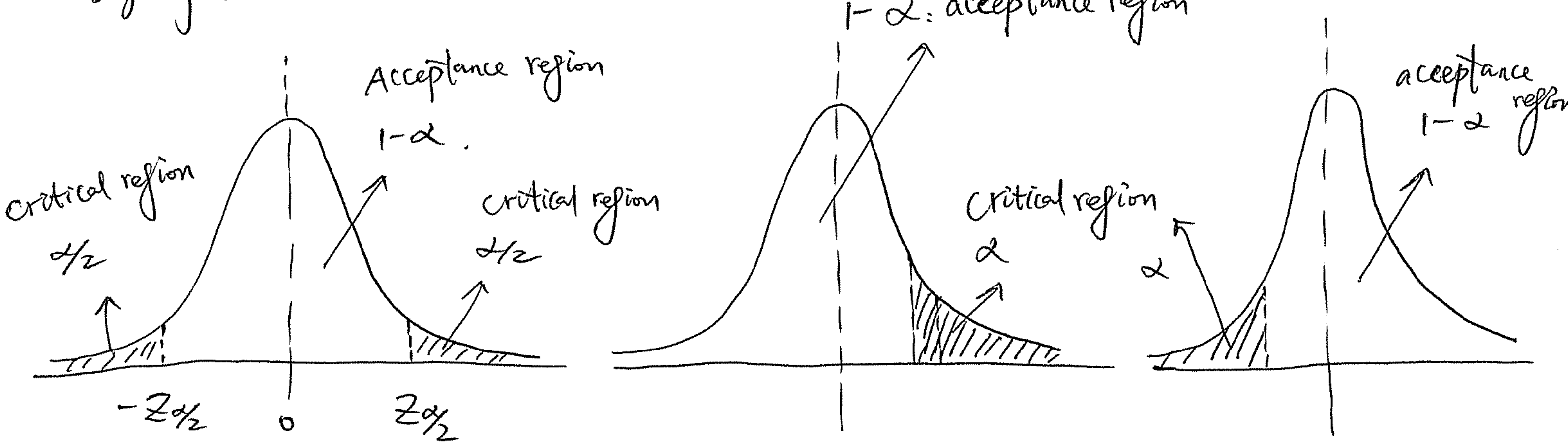
classical approach

Null hypothesis: $H_0 : \mu = \mu_0$

Test statistic: $Z_0 : \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

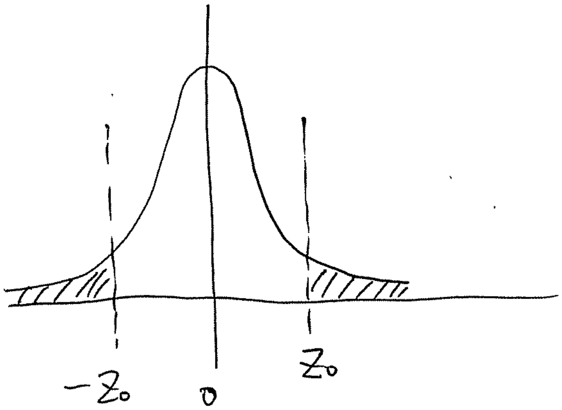
Alternative hypothesis:	rejection criteria
$H_1 : \mu \neq \mu_0$	$Z_0 > Z_{\alpha/2}$ or $Z_0 < -Z_{\alpha/2}$
$H_1 : \mu > \mu_0$	$Z_0 > Z_\alpha$
$H_1 : \mu < \mu_0$	$Z_0 < -Z_\alpha$

significance level α.



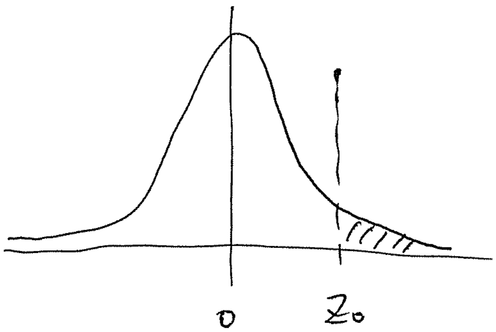
p-value approach

Z_0 is the test statistic $\qquad Z_0 = \dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$



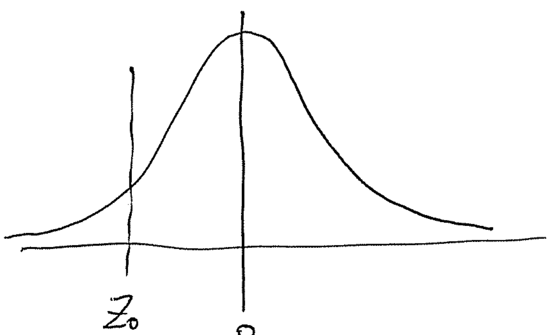
Two-tailed test: $H_0: \mu=\mu_0 \quad H_1: \mu\neq\mu_0$

$$p = 2\cdot\left[1-P(Z<Z_0)\right] = 2(1-\phi(Z_0))$$



right-tailed test $H_0: \mu=\mu_0 \quad H_1: \mu>\mu_0$

$$p = 1-\phi(Z_0).$$



left-tailed test $H_0: \mu=\mu_0 \quad H_1: \mu<\mu_0$

$$p = P(Z<Z_0) = \phi(Z_0)$$

STEPS: <1> Formulate null hypothesis H_0 & alternative hypothesis H_1

<2>. Identify a test statistic

<3> Compute the p-value

The smalled the p-value, the stronger evidence against null hypothesis

<4>. Compare p-value to an acceptance significance level α. If $p\leq\alpha$, the null hypothesis is rejected and the alternative hypothesis is valid.

Eg. Ex 9-38.

The battery life is known to be approximately normal distributed with standard deviation $\sigma = 1.25$ hours. A random sample of 10 batteries has a mean life of $\bar{x} = 40.5$ hours

(a) Is there evidence to support the claim that battery life exceeds 40 hours? use $\alpha = 0.05$

<1> Formulate H_0 & H_1.

$$\begin{cases} H_0: \mu = 40 \\ H_1: \mu > 40 \end{cases} \qquad \alpha = 0.05$$

<2> Calculate the test statistic Z_0

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - 40}{1.25/\sqrt{10}} = 1.26$$

<3> reject H_0 if $Z_0 > Z_\alpha$

$$Z_\alpha = Z_{0.05} = 1.65$$

<4>. making decision

$1.26 < 1.65$ do not reject H_0 and conclude the battery life is not significantly greater than 40 at significance level $\alpha = 0.05$.

(b). $p\text{-value} = 1 - \Phi(Z_0) = 1 - \Phi(1.26) = 1 - 0.8962 = 0.1038$

$p\text{-value} > \alpha$, fail to reject H_0.

what if the battery life does not equal to 40 hours ?

(a). <1> Formulate H_0 & H_1

$H_0: \mu = 40$

$\alpha = 0.05$

$H_1: \mu \neq 40$

<2> calculate the test statistic

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{40.5 - 40}{1.25/\sqrt{10}} = 1.26$$

<3>. reject H_0, if $Z_0 > Z_{\alpha/2}$ or $Z_0 < -Z_{\alpha/2}$

$\alpha = 0.05$

$Z_{0.025} = 1.96$ $\quad -Z_{0.025} = -1.96$

<4>. making decision

$Z_0 = 1.26$ $\quad -Z_{0.025} < Z_0 < Z_{0.025}$

fail to reject H_0,

(b). $p\text{-value} = 2(1 - \phi(Z_0)) = 2 \times 0.1038 = 0.2076$

$p\text{-value} > \alpha$ $\quad \alpha = 0.05$

fail to reject H_0.

Confidence Interval and Hypothesis Testing

1) *Confidence Interval*

A machine is set up such that the average content of juice per bottle equals μ.
A sample of 100 bottles yields an average content of 48oz.
Calculate a 90% and a 95% confidence interval for the average content.
Assume that the population standard deviation σ = 5oz.

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

100(1-α)%	90%	95%	99%
$z_{\alpha/2}$	1.645	1.96	2.576

$$90\%:\ 48 \pm 1.645 \times \frac{5}{\sqrt{100}}$$
$$95\%:\ 48 \pm 1.96 \times \frac{5}{\sqrt{100}}$$

2) *Sample size*

What sample size is required to make sure the margin of error (MOE) is within 0.5oz at the 95% confidence level? (±0.5 oz)
Assume that the population standard deviation σ = 5oz.

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 0.5$$
$$1.96 \times \frac{5}{\sqrt{\mathrm{n}}} = 0.5$$
$$n = (1.96*5/0.5)^2 = 368.64 \sim 369$$

3) *Hypothesis Testing*

A machine is set up such that the average content of juice per bottle equals μ. A sample of 36 bottles yields an average content of 51.5oz. Test the hypothesis that the average content per bottle is 50oz at the 5% significance level.
Assume that the population standard deviation σ = 5oz.

<u>Classical approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: $\mu=50$ H_1: $\mu\neq 50$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{51.5 - 50}{5/\sqrt{36}} = 1.8$$

(c) For the two sided test, reject H_0 if $Z_0 > Z_{\alpha/2}$ or $Z_0 < -Z_{\alpha/2}$

$Z_{\alpha/2} = Z_{0.025} = 1.96$

$-1.96 < 1.8 < 1.96$ $-Z_{\alpha/2} < Z_0 < Z_{\alpha/2}$ within the acceptance region and the null hypothesis cannot be rejected

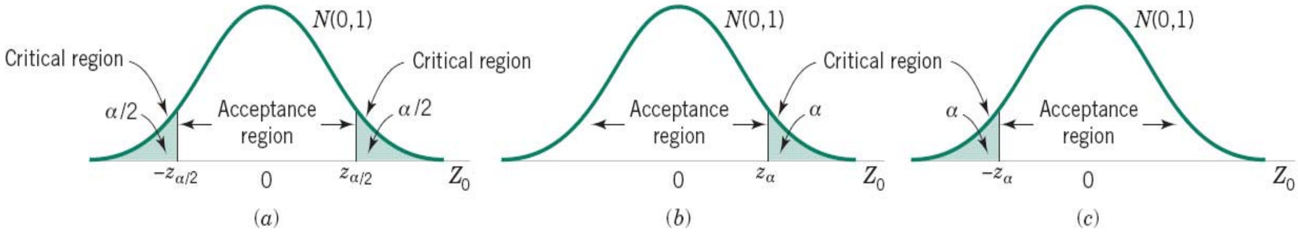


Figure 9-7 The distribution of Z_0 when H_0: $\mu = \mu_0$ is true, with critical region for (a) the two-sided alternative $H_1 : \mu \neq \mu_0$, (b) the one-sided alternative $H_1 : \mu > \mu_0$, and (c) the one-sided alternative $H_1 : \mu < \mu_0$.

Null hypothesis: H_0: $\mu = \mu_0$

Test statistic: $Z_0 = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$

Alternative hypothesis	Rejection criteria
H_1: $\mu \neq \mu_0$	$z_0 > z_{\alpha/2,n-1}$ or $z_0 < -z_{\alpha/2,n-1}$
H_1: $\mu > \mu_0$	$z_0 > z_{\alpha,n-1}$
H_1: $\mu < \mu_0$	$z_0 < -z_{\alpha,n-1}$

<u>P-value approach:</u>

The ***P*-value** is the smallest level of significance that would lead to rejection of the null hypothesis H_0 with the given data.

$$P = \begin{cases} 2[1 - \Phi(|z_0|)] & \text{for a two-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu \neq \mu_0 \\ 1 - \Phi(z_0) & \text{for a upper-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu > \mu_0 \\ \Phi(z_0) & \text{for a lower-tailed test: } H_0\text{: } \mu = \mu_0 \quad H_1\text{: } \mu < \mu_0 \end{cases} \qquad (9\text{-}15)$$

Steps:

(a) Formulate null hypothesis H_0 and alternative hypothesis H_1

(b) Calculate the test statistics

(c) Computer the P-value

The smaller the P-value is, the stronger evidence against the null hypothesis

(d) Compare the P-value to an acceptance significance level α. If p-value< α, the null hypothesis is rejected and the alternative hypothesis is valid.

Steps:

(a) Formulate H_0 and H_1

H_0: $\mu=50$ $\quad$ H_1: $\mu\neq50$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{51.5 - 50}{5/\sqrt{36}} = 1.8$$

(c) For the two sided test, p = 2[1-Φ(Z_0)]=2*(1-0.9641)=2*0.0359=0.0718.

(d) P-value>0.05, the null hypothesis cannot be rejected

4) ***The impact of sample size***

A machine is set up such that the average content of juice per bottle equals μ. A sample of 100 bottles yields an average content of 51.2oz. Test the hypothesis that the average content per bottle is 50oz at the 5% significance level.
Compare the conclusion to that based on the 36 bottles sample.
Assume that the population standard deviation σ = 5oz.

<u>Classical approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: μ=*50* H_1: $\mu \neq$*50*

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{51.2 - 50}{5/\sqrt{100}} = 2.4$$

(c) For the two sided test, reject H_0 if Z_0>$Z_{\alpha/2}$ or Z_0<-$Z_{\alpha/2}$

$Z_{\alpha/2}$ = $Z_{0.025}$=1.96

2.4>1.96 Z_0>$Z_{\alpha/2}$ the null hypothesis is rejected and the alternative hypothesis is valid.

<u>P-value approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: μ=*50* H_1: $\mu \neq$*50*

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{51.2 - 50}{5/\sqrt{100}} = 2.4$$

(c) For the two sided test, p = 2[1-Φ(Z_0)]=2*(1-0.9918)=2*0.0082=0.0164

P-value<0.05, the null hypothesis is rejected and the alternative hypothesis is valid.

5) ***Right-tailed test***

A machine is set up such that the average content of juice per bottle equals μ. A sample of 10 bottles yields an average content of 40.5oz. Is there evidence to support the claim that the average content per bottle exceeds 40oz (5% significance level)?
Assume that the population is approximately normally distributed with standard deviation σ = 1.25oz.

<u>Classical approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: $\mu=40$ H_1: $\mu>40$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{40.5 - 40}{1.25/\sqrt{10}} = 1.2649$$

(c) For the right-tailed test, reject H_0 if $Z_0>Z_\alpha$

$Z_\alpha = Z_{0.05}=1.65$

$1.2649<1.65$ the null hypothesis can not be rejected.

<u>P-value approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: $\mu=40$ H_1: $\mu>40$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{40.5 - 40}{1.25/\sqrt{10}} = 1.2649$$

(c) For the right-tailed test, $p = 1-\Phi(Z_0)=1-0.8962=0.1038$

P-value>0.05, the null hypothesis can not be rejected.

6) ***Left-tailed test***

The manager claims that the average content of juice per bottle is less than 50oz. The machine operator disagrees. A sample of 100 bottles yields an average content of 49oz per bottle. Does this sample allow the manager to claim he is right (5% significance level)?

Assume that the population standard deviation $\sigma = 5$oz.

<u>Classical approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: $\mu=50$ H_1: $\mu<50$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{49 - 50}{5/\sqrt{100}} = -2$$

(c) For the left-tailed test, reject H_0 if $Z_0<-Z_\alpha$

$Z_\alpha = Z_{0.05}=1.65$

$-2<-1.65$ the null hypothesis is rejected and the alternative hypothesis is valid.

<u>P-value approach:</u>

Steps:

(a) Formulate H_o and H_1

H_0: $\mu=50$ H_1: $\mu<50$

(b) Calculate the test statistic Z_0

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{49 - 50}{5/\sqrt{100}} = -2$$

(d) For the left-tailed test, $p = \Phi(Z_0)= \Phi(-2)=0.02275$

P-value<0.05, the null hypothesis is rejected and the alternative hypothesis is valid.