L10. Tests of Statistical Hypotheses

Besides C.I., a major area of statistical inference, is to use samples to conduct $\boxed{\text{Hypothesis Tests}}$ for parameter values.

$\boxed{\text{EX}}$ m&m packages states the weight is 49.5g. Is that true? Let's buy 10 bags and weigh them.

$\bar{x} = 49.5, \quad S^2 = 0.04$

How to use the sample to answer our question?

Components of Hypothesis Test:

(1) Hypotheses about the parameter

$\boxed{\text{Equality}}$

parameter ↓ ↙ given constant

$H_0: \mu = \mu_0$ — Null Hypothesis

$H_1: \mu \neq \mu_0$ — Alternative Hypothesis

$\mu > \mu_0$

$\mu < \mu_0$

$\boxed{\text{Inequality}}$

(2) Test statistic from the Sample and its sampling distribution when H_0 is true

the importance of equality in H_0

A random sample from $N(\mu, \sigma^2)$

when H_0 is true: $\bar{X} \sim N(\mu = \mu_0, \frac{\sigma^2}{n})$

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

$$t_{stat} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

(3) Decision Rule — significant level α

Test Result	Facts: H_0 is True	Facts: H_1 is true
Fail to reject H_0 (−)	✓	False Negative (Type II Error)
Reject H_0 (+)	False Positive (Type I Error)	✓

In Hypothesis tests, we care more about false positive rates: when we make a decision that the difference is significant, we want it to be reliable $\longleftrightarrow$ Type I error is controlled. significant level

Now let's talk about decision rule using

$H_0: \mu = \mu_0$

$H_1: \mu > \mu_0$ ← If H_1 is true, μ is greater than 49.3 $\boxed{\text{significantly}}$

using a sample mean $\overline{X}$, How big $\overline{X}$ need to be for us to be confident to say $\mu > 49.3$?

$\boxed{\text{Decision rule}}$: $\{\overline{X} \geq a\}$ we reject H_0 and accept H_1

$- \mathbb{P}(\text{reject } H_0 \mid H_0 \text{ is true})$

$= \mathbb{P}(\overline{X} \geq a)$ given $\overline{X} \sim N(\mu_0, \frac{\sigma^2}{n})$

$= \alpha$

what is a ?

$$\mathbb{P}(\bar{X} \geq a) = \mathbb{P}\Big(\underbrace{\frac{\bar{X}-\mu_0}{s/\sqrt{n}}}_{t} \geq \frac{a-\mu_0}{s/\sqrt{n}}\Big)$$

$$= \alpha$$

$$= \mathbb{P}\Big(t \geq \frac{a-\mu_0}{s/\sqrt{n}}\Big)$$

$$\frac{a-\mu_0}{s/\sqrt{n}} = t_\alpha$$

$$a = \mu_0 + t_\alpha \frac{s}{\sqrt{n}}$$

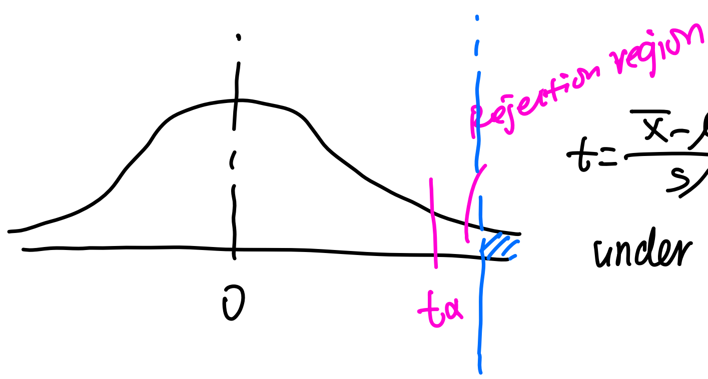
Decision Rule: $\left\{ \bar{X} \geq \mu_0 + t_\alpha \frac{s}{\sqrt{n}} \right\}$

or

for a given α. $\{ t_{stat} \geq t_\alpha \}$

This way, when we reject H_0 and accept H_1 using this decision rule, we only have α $(=0.05, 0.01, 0.1)$ chance of making a false positive decision.

Alternatively,



$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$ under H_0

For any sample with $\bar{x}$ in this rejection region $\iff$ $t_{stat} \geq t_\alpha$

$\iff$ $P(t \geq t_{stat}) \leq \alpha$

Decision Rule*: Reject H_0 and accept H_1 if

$$\text{p-value} = \mathbb{P}(t \geq t_{stat}) \leq \alpha$$

In a summary:

The hypothesis test for one population mean μ,

particularly, t test for μ:

Test for one mean μ

$$t_{stat} = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$$

H_0	H_1	Critical Region, df=n-1	p-value (estimate the bounds)
$\mu = \mu_0$	$\mu > \mu_0$	$t_0 \geq t_\alpha$	$P(t > t_0)$
$\mu = \mu_0$	$\mu < \mu_0$	$t_0 \leq -t_\alpha$	$P(t < t_0)$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert t_0 \rvert \geq t_{\alpha/2}$	$2P(t > \lvert t_0 \rvert)$

$\boxed{\text{Think}}$ For test for population proportion, using the sampling distribution of $\hat{p} = \frac{Y}{n}$:

$$\hat{p} \overset{Appr}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$$

$$Z_{stat} = (\hat{p} - p_0) / \sqrt{p_0(1-p_0)/n}$$

H_0	H_1	Critical Region	p-value
$p = p_0$	$p > p_0$	$Z_0 \geq Z_\alpha$	$P(Z > Z_0)$
$p = p_0$	$p < p_0$	$Z_0 \leq -Z_\alpha$	$P(Z < Z_0)$
$p = p_0$	$p \neq p_0$	$\lvert Z_0 \rvert \geq Z_{\alpha/2}$	$2P(Z > \lvert Z_0 \rvert)$

Test for the equality of two means

$\boxed{\text{Case 1}}$ Two independent random samples from two Normal Distributions

work with the sampling distribution of $\bar{X} - \bar{Y}$:

$$\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$$

then
$$\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{S_{pooled} \cdot \left(\frac{1}{n} + \frac{1}{m}\right)} \sim t(n+m-2)$$

where $S_{pooled} = \sqrt{\dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}$

Define test statistics :

$$t_{stat} = \frac{\bar{X} - \bar{Y}}{S_p(\frac{1}{n} + \frac{1}{m})}$$

H_0	H_1	Critical Region, df=n+m-2	p-value (estimate the bounds)
$\mu_x = \mu_y$	$\mu_x > \mu_y$	$t_0 \ge t_\alpha$	$P(t > t_0)$
$\mu_x = \mu_y$	$\mu_x < \mu_y$	$t_0 \le -t_\alpha$	$P(t < t_0)$
$\mu_x = \mu_y$	$\mu_x \ne \mu_y$	$\lvert t_0 \rvert \ge t_{\alpha/2}$	$2P(t > \lvert t_0 \rvert)$

$\boxed{\text{case 2}}$ X and Y are not independent,

Denote $W = \bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \sigma_W^2)$

unknow

then the test to compare two means

became to test for $\mu_w = 0$ or not.