Model:

One Sample: $X_1, \dots, X_n$ iid ~~$N(\mu, \sigma^2)$~~

R.V.s with $E(X_i) = \mu$

$Var(X_i) = \sigma^2$

ALWAYS UNKNOWN
$\Rightarrow$ PARAMETER OF INTEREST

MAY BE

KNOWN $= \sigma_0^2$

Z-test

$$\frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}$$

: to test $H_0: \mu = \mu_0$.

KNOWN ST. ERROR SINCE

$$\sqrt{Var(\bar{X})} = \frac{\sigma_0}{\sqrt{n}}.$$

$\sim N(0,1)$ if H_0 true.

SO LONG AS $\bar{X}$ is (approx.) normally distributed

X_i's normal

"large n" Central Limit Thm

UNKNOWN.

MUST ESTIMATE ST. ERROR USING

$\frac{S}{\sqrt{n}}$ where $S = \sqrt{\frac{1}{n-1}\sum(X_i - \bar{X})^2}$

t-test

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

if H_0 true

SO LONG AS X_i's NORMAL.

Two-Sample Model
- Paired T-test
- 2 "indep sample" t-test

Paired test: 1-sample t-test on differences.

2 indep samples.

$X_1, \dots, X_m \sim N(\mu_X, \sigma_X^2)$

$Y_1, \dots, Y_n \sim N(\mu_Y, \sigma_Y^2)$

Parameter of interest is $\mu_X - \mu_Y$.

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

"TRUE" St. Error is $\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}$.

2 Sub-cases:

1) $\sigma_X^2 = \sigma_Y^2 \left(= \sigma^2 \text{ say}\right)$.

~~We est~~

Then St. Error $= \sigma \sqrt{\frac{1}{m} + \frac{1}{n}}$

We estimate σ ~~this by~~ using

$$S_p = \sqrt{\frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2}}$$

Then the t-statistic is

$$\frac{(\bar{X} - \bar{Y}) - \underbrace{(\mu_{X0} - \mu_{Y0})}_{\text{often} = 0}}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t_{m+n-2}$$

if

$\mu_X - \mu_Y = \mu_{X0} - \mu_{Y0}$

2) $\sigma_X^2 \neq \sigma_Y^2$: Then we estimate the st. error using.

$$\sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}$$

The t-stat is

$$\frac{\bar{X} - \bar{Y} - (\mu_{X0} - \mu_{Y0})}{\sqrt{\frac{S_X^2}{m} + \frac{S_Y^2}{n}}} \sim ??$$

under $H_0: \mu_X - \mu_Y = \mu_{X0} - \mu_{Y0}$.

approx t-dist'n with "data-dependent" degrees of freedom

"Welch test"