$H_0: \quad \mu = \mu^*$

$H_a: \quad \mu < \mu^*$

$>$

$\neq$

Uniformly most
most powerful
test

$\mu \le \mu_0$

$\mu > \mu_0$

$\parallel$

$\mu = \mu_0$

$\mu > \mu_0$

$$ts = z = \frac{\bar{X} - \mu^*}{\sigma/\sqrt{n}} \sim N(0,1)$$

Assumptions

1. $X \overset{iid}{\sim} N(\mu, \sigma^2)$

2. $n \geq 30 \qquad \bar{X} \sim N(\mu, \sigma^2/n)$ True

   X false

3. σ must be known.

   If unknown we can estimate σ with s so long as $n \geq 30$

$$S = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

What if σ is unknown and $n < 30$ how do we test the null hypothesis?

$$ts = \frac{\bar{X} - \mu^*}{s/\sqrt{n}} \qquad z = \frac{\bar{X} - \mu^*}{\sigma/\sqrt{n}}$$

$n < 30$
σ unknown

$$ts \sim t(n-1)$$

$$\lim_{n \to \infty} t(n-1) \rightarrow z \sim N(0,1)$$

critical or p-value
and reject H_0

Particulate matter from roads can affect the sewer systems in a city. Particularly the size particulate matter. You conduct a study of the Highways in CA and measure the PM. Previous studies indicate that pm. is around 44 μg. You are interested in determining if the pm in CA is different from previous studies.

82.9 56.8 66.5 49.4 105.4

79.5 82.5 50.7 43.0

$s = 20.49$

$H_0: \mu = 44$
$H_a: \mu \neq 44$

$$ts = \frac{\bar{X} - \mu^*}{s/\sqrt{n}}$$

$$ts = \frac{68.52 - 44}{20.49/\sqrt{9}}$$

$\alpha = 0.05$

$$= 3.59$$

$ts = 3.59 \qquad DF = 8$

$$0.005 > p > 0.001$$

$\alpha > p$ Reject H_0

$\alpha = 0.05$

$$0.05 > p$$

$\alpha > p$ Reject H_0

$$t = \frac{N(0,1)}{\sqrt{\chi^2_\nu / \nu}}$$

$\bar{X} \sim N(\mu, \sigma^2/n)$

$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\frac{\dfrac{\bar{X} - \mu}{\sqrt{\cancel{\sigma^2}/n}}}{S\sqrt{\dfrac{\cancel{(n-1)}S^2}{\cancel{\sigma^2}} \Big/ \cancel{n-1}}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

2 samples X Y

$X = \{X_1, \dots, X_n\} \overset{iid}{\sim} N(\mu_x, \sigma_x^2)$

$Y = \{Y_1, \dots, Y_m\} \overset{iid}{\sim} N(\mu_y, \sigma_y^2)$

$n \overset{?}{=} m$

$H_0: \mu_x = \mu_y$

$H_a: \mu_x \ \substack{> \\ < \\ \neq} \ \mu_y$

$\Delta = 0$

Generalize

$H_0: \mu_x - \mu_y = \Delta$

$H_a: \mu_x - \mu \ \substack{> \\ < \\ \neq} \ \Delta$

$n, m > 30$ σ_x^2 σ_y^2 known

$$ts = \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} \sim N(0,1)$$

$X \perp Y$

$n, m > 30$ σ_x^2 σ_y^2 unknown

$$ts = \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \sim N(0,1)$$

$X \perp Y$

n or $m < 30$

n or $m < 30$ σ_x^2 σ_y^2 unknown

$\sigma_x^2 = \sigma_y^2$

$$ts = \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \sim t_{n+m-2} \qquad S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}$$

pooled variance

$\sigma_x^2 \neq \sigma_y^2$

$$ts = \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} \sim t_v$$

$$v = \frac{\left(\frac{S_x^2}{n} + \frac{S_y^2}{m}\right)^2}{\frac{S_x^4}{n^2(n-1)} + \frac{S_y^4}{m^2(m-1)}} = v$$