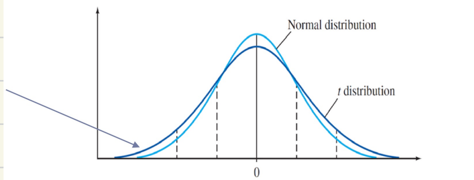
# Student's t Distribution

- random variable $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ (n samples from a normal population)

  has a density function called Student's t distribution

  $\bar{x}$, s are random variables

- t & z are similar
- t's has heavier tails than z as it involves 2 random variables while z only involves 1 ($\bar{x}$)



Standard normal z and the t distribution with 5 degrees of freedom

**Figure 10.1**

- $t \longrightarrow z$ for $\mu = 0$,
  $\sigma^2 = 1$
  n is large $\Longrightarrow s \to \sigma$

- $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$ (n-1) degrees of freedom for sample variance $s^2$

the degrees of freedom determine the shape of t dist

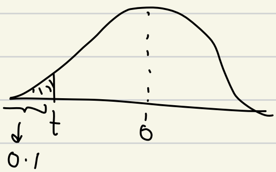
- assumptions :

→ randomly selected sample

→ normally distributed population (t doesn't work if the pop is not normally distributed

Q. Suppose you have a sample of size $n = 10$ from a normal distribution. Find a value of $t$ such that only 1% of all values of $t$ will be smaller.

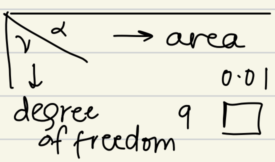
$df = n - 1 = 9$   specify the correct t distribution.




t distribution is symmetric about 0

t table : $t_{\alpha.\nu}$   $Pr(T_\nu > t_{\alpha:\nu}) = \alpha$

$\nu$ = degree's of freedom



# Small sample Inferences : Population Mean

## Small sample Hypothesis Test ($n < 30$)

1. $H_0 : \mu = \mu_0$

2. $H_a : \mu \neq \mu_0$ or $\mu \gtrless \mu_0$

3. Test statistic . $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

4. Rejection region : Reject $H_0$ when

$t > t_\alpha$ one tail $\quad$ $t > t_{\alpha/2}$ two tail $\quad$ p value $< \alpha$

all values of $t$, $t_\alpha$, $t_{\alpha/2}$ are based on $(n-1)$ degrees of freedom

Confidence Interval. : $\bar{x} \pm t_{\alpha/2} \dfrac{s}{\sqrt{n}}$

Q.

A new process for producing synthetic diamonds can be operated at a profitable level only if the average weight of the diamonds is greater than .5 karat. To evaluate the profitability of the process, six diamonds are generated, with recorded weights .46, .61, .52, .48, .57, and .54 karat.

Do the six measurements present sufficient evidence to indicate that the average weight of the diamonds produced by the process is in excess of .5 karat?

$H_0$ : $\mu = 0.5$

$H_a$ : $\mu > 0.5$

t.s. : $\bar{x} = 0.53$

0.07
0.08
0.01
0.04
0.05
0.01

$\sqrt{0.00312}$
$= 0.059$

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

$s = 0.059$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 1.32$$

$t > t_\alpha$

choose confidence level. and then find $t_\alpha$ accordingly.

Q.

Labels on 1-gallon cans of paint usually indicate the drying time and the area that can be covered in one coat. One manufacturer claims that a gallon of its paint will cover 400 square feet of surface area. To test this claim, a random sample of ten 1-gallon cans of white paint were used to paint 10 identical areas using the same kind of equipment. The actual areas (in square feet) covered by these 10 gallons of paint are:

| | | | | |
|---|---|---|---|---|
| 310 | 311 | 412 | 368 | 447 |
| 376 | 303 | 410 | 365 | 350 |

$H_0 : \mu = 400$

$H_a : \mu \neq 400$

$n = 10$

$\bar{x} = 365.2$

$s = 48.417$

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

$t = \frac{\bar{x} - \mu_0}{s / \sqrt{10}} = -2.27$ $\quad t = \pm 2.27$

2 tailed. p value : $p(t \leq -2.27) + p(t \geq 2.27)$

p value $< 0.05$

$\alpha = 5\%$

p value $< 0.05 \rightarrow$ reject. $\therefore$ can reject

confidence interval

$\bar{x} \pm t_{\alpha/2} \left( \frac{s}{\sqrt{n}} \right)$

$365.2 \pm 2.262 \left( \frac{48.417}{\sqrt{10}} \right)$

$365.2 \pm 34.62 \leftarrow$ the values b/w which averge area lies

## Small sample Interferences for Diff b/w Two Population Means: Independent Random Samples

- both pops are normal
- $x_1$, $\mu_1$, $\sigma$, $n_1$
- $x_2$, $\mu_2$, $\sigma$, $n_2$ → same $\sigma$
- $H_0$ : $\mu_1 - \mu_2 = 0$
  $H_A$ : $\mu_1 - \mu_2 \neq 0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- $df = n_1 + n_2 - 2$
- $$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
- assuming samples are random and independent from normally distributed populations with equal variances.
- if variences are unequal, approximate t dist.
  ↑ do this if $\frac{\text{larger } s^2}{\text{smaller } s^2} > 3$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad df \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

# quiz pts to remember

- when trying to estimate $n$ given confidence interval, for proportion's, and if you don't know $\hat{p}$, use 0.5 as it has the highest variability.

| | Runners | | Cyclists | |
|---|---|---|---|---|
| Condition | Mean | Standard Deviation | Mean | Standard Deviation |
| Before Exercise | 255.63 | 115.48 | 173.8 | 60.69 |
| After Exercise | 284.75 | 132.64 | 177.1 | 64.53 |
| Difference | 29.13 | 21.01 | 3.3 | 6.85 |

and Cyclists II Refer to Exercise10.27. In additio... | Chegg.com

$n = 10$

runner : $H_0 : \mu_1 - \mu_2 = 0$     $1 \rightarrow$ after

$H_a : \mu_1 - \mu_2 > 0$     $2 \rightarrow$ before

$s_1 = 115.48$

$s_2 = 132.64$

$$s_d = \sqrt{\frac{(n-1)(s_1^2) + (n-1)(s_2)^2}{2n-2}}$$

$s_d = 124.356$

- $H_0$ , $H_a$
  if one tail : $Z_\alpha$
  if two tail : $Z_{\alpha/2}$

- reject $H_0$ is $p \leq \alpha$ or $Z$ test statistic $> Z_\alpha$

- $Z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ or $\dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ $\dfrac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$

- $Z = \dfrac{\bar{x}_1 - \bar{x}_2 - (\bar{\mu}_1 - \bar{\mu}_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$ or $\dfrac{\bar{x}_1 - \bar{x}_2 - (\bar{\mu}_1 - \bar{\mu}_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

  $\downarrow$

  $\dfrac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}}$

- p value : $\text{Prob}(x = \bar{x} \mid H_0 \text{ is true})$

- type 1 : reject $H_0$ when true
  $P(\text{type } 1) = \alpha$

  type 2 = accept $H_0$ when false
  $P(\text{type } 2) = \beta$

- 0.1 —— 0.05 —— 0.01 ← reject, highly significant

  not rejected
  not statistically significant

  not rejected
  → statistically significant

  reject
  statistically significant

confidence interval

- $\bar{x} \pm (z_\alpha)\left(s/\sqrt{n}\right)$ ← if Ho is in the interval, you can't reject it.

or

$$\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

assumptions:

1. can apply CLT
2. pop normal or $n > 30$
3. $n\hat{p}$, $n\hat{q} > 5$

- $\alpha \downarrow$, $P(\text{type } 1)\downarrow$, $P(\text{type } 2)\uparrow$, power$\downarrow$

# t - dist

- $t = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$
  - $n < 30$ & $\sigma$ unknown
  - pop has to be normal

- $t \to z$ for $\mu = 0, \sigma^2 = 1$, large $n$

- $s = \sqrt{\dfrac{\Sigma (x_i - \bar{x})^2}{n-1}}$

- $df = n - 1$

- reject Ho when : $t > t_\alpha$ , $p < \alpha$

- choose confidence int , if Ho value lies outside it, then reject.

- $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$
  - $df = n_1 + n_2 - 2$
  - $\dfrac{s_1}{s_2} < 3$

$s^2 = \dfrac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}$