# Lecture 5

## One Sample t-test

Data: $y_1, y_2, \ldots, y_n$

$$t = \frac{\bar{y} - \mu}{s/\sqrt{n}} \qquad s^2 := \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

if $Y \sim N(\mu, \sigma^2 I)$, then

$t \sim$

$t \sim t_{n-1}$

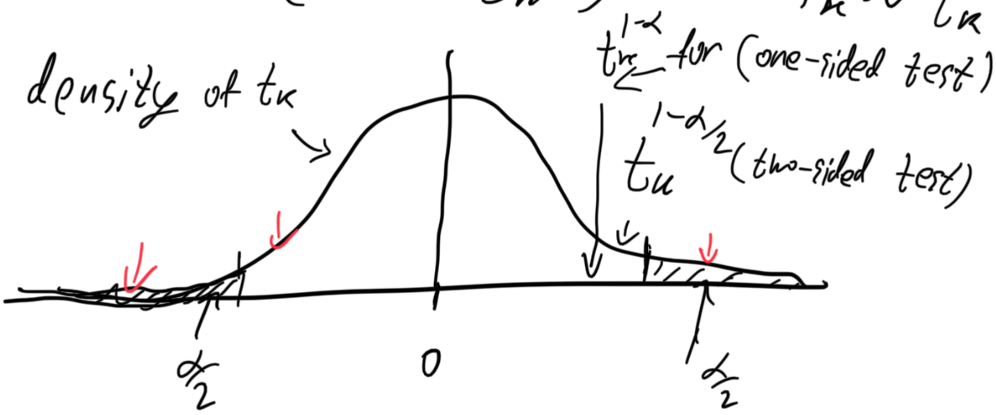
we used this fact to test hypotheses of the form

$$H_0 : \mu = \mu_0 \in \mathbb{R}$$

we reject $H_0$ if $|t| > t_{n-1}^{1-\alpha/2}$

$t_k^{1-\alpha}$ is the $1-\alpha$ quantile of the $t$ dist. over $k$ DoF:

$$\alpha = \Pr(T_k > t_k^{1-\alpha}) \qquad T_k \sim t_k$$




$$P = \Pr(T_{n-1} > |t_{obs}|)$$

---

## Significance

- It is useful to think that

p measures the sample size"
(Richard Olshen)

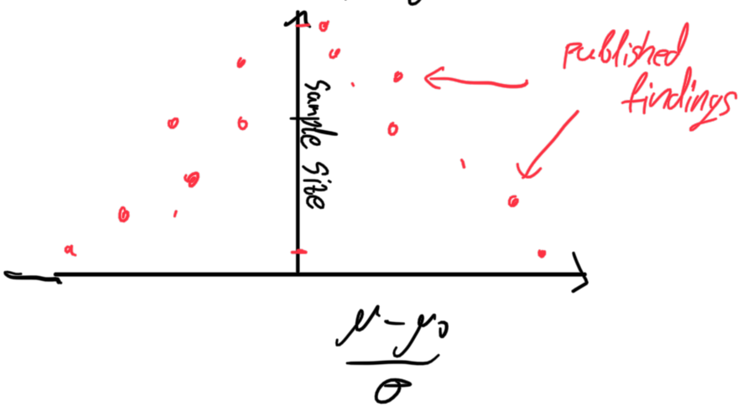
indeed $p \asymp e^{-k \cdot n}$ $k$ only depends on $\mu, \mu_0, \sigma$

"effect size" $\sim \frac{\mu_0 - \mu}{\sigma}$

- As the sample size n goes up, the ability to find smaller statistically significant effects increase

- Non significance $\not\Rightarrow$ Non-effect

- "Funnel Plot"




## Practical vs. statistical significance

- Practical signif. is about the magnitude of the effect $\left|\frac{\mu-\mu_0}{\sigma}\right|$

- A small p-value ( say $p = 10^{-20}$ ) may be impressive, but can also be because you have a huge sample size (but small effect).

- For example CBS data.
  the mean difference seems to have low practical value, but it is statistically significant because the very large sample size.

- Art Owen proposed to summarize things in the following table.

| | Practically Sig. | practically Insig. |
|---|---|---|
| Stat. Sig | learn something useful | n is too large |
| Stat. Insig | n is too small | probably keep $H_0$ |

---

## Confidence Interval

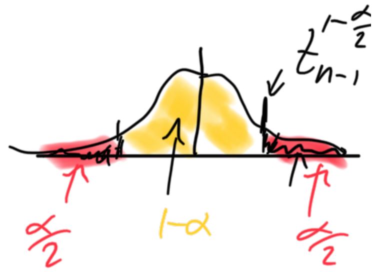
Motivation: a single object that captures both statistical & practical significance.

$$T = \frac{\bar{Y} - \mu}{\phantom{x}} \sim t_{n-1}$$

true mean

$S/\sqrt{n}$

$$\boxed{1-\alpha = \Pr\left(|T| < t_{n-1}^{1-\alpha/2}\right)}$$




$$= \Pr\left(-t_{n-1}^{1-\alpha/2} < T < t_{n-1}^{1-\alpha/2}\right)$$

$$= \Pr\left(-t_{n-1}^{1-\alpha/2} < \frac{\bar{Y}-\mu}{S/\sqrt{n}} < t_{n-1}^{1-\alpha/2}\right)$$

$$= \Pr\left(\underbrace{\bar{Y} - \frac{S}{\sqrt{n}} t_{n-1}^{1-\alpha/2}}_{L} < \mu < \underbrace{\bar{Y} + \frac{S}{\sqrt{n}} t_{n-1}^{1-\alpha/2}}_{U}\right)$$

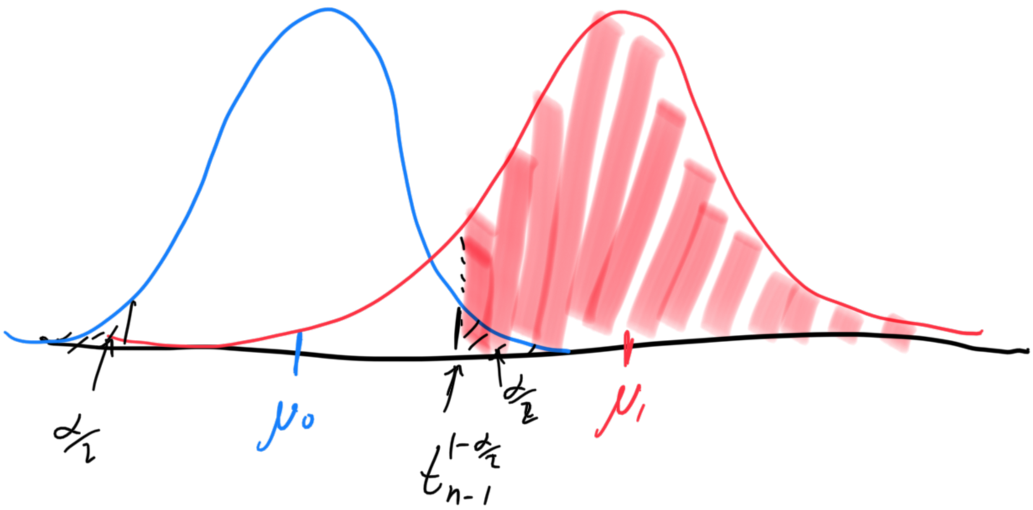
- This is $1-\alpha$ confidence interval for $\mu$
  The range $(L, U)$ s.t. $\Pr(L \le \mu \le U) \ge 1-$

- $(L, U)$ is a <u>random</u> interval that contains the true value of $\mu$ $1-\alpha$ of the "time"

- confidence intervals and p-values have a dual property: We can construct a confidence interval out of all of the p-values of $\mu_0$ that would not have been rejected.

# Statistical Power

Power of a test is the chance of rejecting $H_0$ under $\underline{H_1}$. Usually denoted $1-\beta$. That is, $\beta$ is the prob. of Type II error = Prob. of making a false negative




For the t-test with

$$t = \frac{\bar{Y}-\mu_0}{s/\sqrt{n}} \qquad \begin{array}{l} H_0: \; Y_i \sim N(\mu_0, \sigma^2) \\ H_1: \; Y_i \sim N(\mu, \sigma^2) \end{array}$$

$$1-\beta = \Pr\left(|t| > t_{n-1}^{1-\frac{\alpha}{2}} \mid H_1\right)$$

$$= \Pr\left(t^2 > \left(t_{n-1}^{1-\frac{\alpha}{2}}\right)^2 \mid H_1\right)$$

$$= \Pr\left(n\frac{(\bar{Y}-\mu_0)^2}{s^2} > F_{1,n-1}^{1-\alpha} \mid H_1\right)$$

$$= \Pr\left[\frac{\left(\overbrace{\frac{\sqrt{n}(\bar{Y}-\mu)}{\sigma}}^{Z} + \frac{\sqrt{n}(\mu-\mu_0)}{\sigma}\right)^2}{s^2/\sigma^2} > F_{1,n-1}^{1-\alpha} \mid H_1\right]$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

$$=: \Pr\left[\frac{\left(Z + \sqrt{n}\left(\frac{\mu-\mu_0}{\sigma}\right)\right)^2}{V/(n-1)} (*) > F_{1,n-1}^{1-\alpha} \mid H_1\right]$$

Where $Z \sim N(0,1)$ $\quad Z = \frac{\sqrt{n}(\bar{Y}-\mu)}{\sigma}$

$V \sim \chi^2_{n-1}$ $\quad V = \frac{s^2}{\sigma^2}(n-1)$

$\lambda = n\left(\frac{\mu-\mu_0}{\sigma}\right)$

$\frac{N^2(\sqrt{\lambda},1)}{\frac{1}{k}\chi^2_k}$ is known as the <u>non-central F dist.</u>

with 1 DoF over k DoF

(Denoted $F_{1,k}(\lambda)$) $\quad \lambda = \left(\sqrt{n}\left(\frac{\mu-\mu_0}{\sigma}\right)\right)^2$

- Example:

Suppose you belive that $\mu_0 = 10$ and have some reason to guess that $\sigma = 5$

We are interested in wheather $|\mu-\mu_0| \geq 1$.
We want to test significance at $\alpha = 0.05$, and also want to accept at most $\beta = 0.05$ false negatives. What is the minimal sample size $n$ allowing us to do so?

We have

$$\lambda = \left(\frac{\mu-\mu_0}{\sigma}\sqrt{n}\right)^2 = \frac{n}{25}$$

We want to have

(*)

$$1-\beta = 0.95 \leq Pr\left[F'_{1,n-1}\left(\frac{n}{25}\right) > F^{0.95}_{1,n-1}\right]$$

The smallest $n$ is $n = 327$

other $(\beta, n)$ pairs:

| Power | $\beta$ | $n$ |
|---|---|---|
| 0.95 | 0.05 | 327 |
| 0.80 | 0.2 | 199 |
| 0.185 | 0.815 | 30 |

---

Why t-test works?

$t = \frac{\bar{X}-\mu}{\phantom{x}}$

$$t = \frac{\bar{Y} - \mu}{S/\sqrt{n}}$$

By CLT, $\sqrt{n}(\bar{Y} - \mu) \xrightarrow{d} N(0, \sigma^2)$

By Law of large numbers

$$S^2 \xrightarrow{p} \mathbb{E}[S^2] = \sigma^2 \quad \text{as } n \to \infty$$

so $t \xrightarrow{d} N(0,1)$

this is true even if

$Y_i$ are non-normally dist.

---

$\boxed{\text{Non - Normality}}$

$$Y_i \overset{iid}{\sim} F, \quad F \text{ is not normal}$$

$$t = \sqrt{n}\,\frac{\bar{Y} - \mu}{S}$$

$$\text{corr}(\bar{Y}, S^2) \longrightarrow \frac{\gamma}{\sqrt{\kappa + 2}}$$

- In addition,

$$\mathbb{E}[t] = \frac{-\gamma}{2\sqrt{n}} + O\left(\frac{1}{n}\right)$$

$$Var(t) = 1 + \frac{1}{n}\left(2 + \frac{7}{4}r^2\right) + O\left(\frac{1}{n^2}\right)$$

- The variance converges much faster than the mean

— Confidence interval:

$$Pr\left(|t| < z^{1-\frac{\alpha}{2}}\right) = 1 - \alpha + O\left(\frac{1}{n}\right)$$

- Confidence interval provides the right coverage quite quickly.