Two-Sample Hypothesis Testing

- Simple comparative experiments:
  study the differences that two different conditions, a/k/a treatments or levels, have on response
- in the context of completely randomized experiment, may lead to cause and effect.

Tests on Difference in μ, Known σ

- process is pretty much identical to single-population tests on mean!

$$H_0: \mu_1 - \mu_2 = \Delta_0$$

(Δ_0: hypothesized value of difference in means; often zero!)

- sometimes we see $H_0: \mu_1 = \mu_2$

$$H_1: \mu_1 - \mu_2 \neq \Delta_0$$

test statistic:

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

.. critical values for fixed-α test:

$$\pm Z_{\alpha/2}$$

.. p-value found same way as single-population test on u

ex: two primer formulations

formulation #1: original

formulation #2: some new ingredient, supposedly shortens drying time

-- known pop. standard deviation of $\sigma = 8$ minutes, (assumed) unaffected by new ingredient

- paint ten test specimens with #1 and ten with #2, in random order

- important!!!

- What might happen if we paint all ten #1's first, then all ten #2's?

  - for example, what if environmental conditions changed?

    - humidity or temp increase?

→ confounding variables

- randomization helps to mitigate them

test data: $\bar{x}_1 = 121$ minutes

$\bar{x}_2 = 112$ minutes

- at first glance, it looks like the new ingredient does shorten drying time!

.. test

$\not\Delta \equiv \mu_1 - \mu_2 = 0$ (Δ_0)

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 > \mu_2$ ②

We expect $\bar{x}_1 > \bar{x}_2$

$\alpha = 0.05$

[fixed-α]

test statistic:

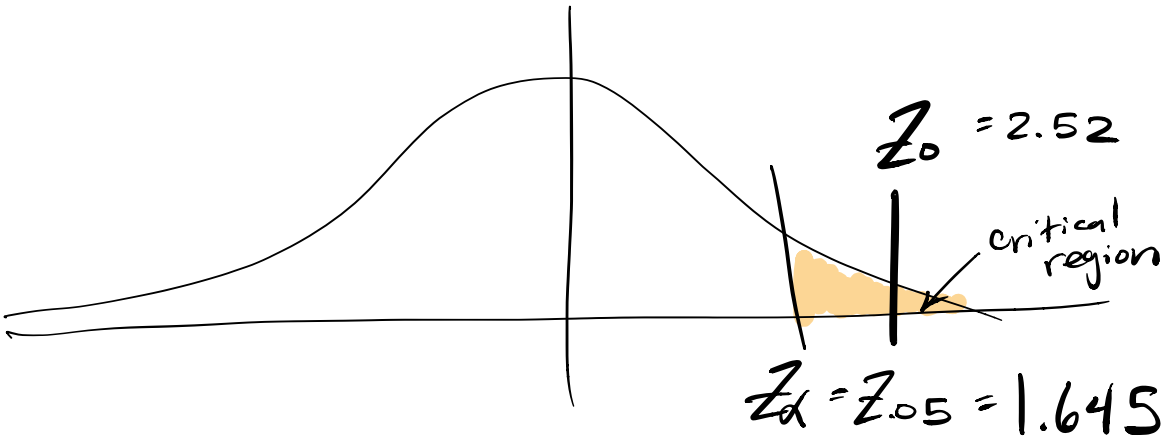
$$Z_0 = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (\Delta_0 \to 0)$$

$$= \frac{121 - 112 - 0}{\sqrt{\frac{8^2}{10} + \frac{8^2}{10}}}$$

$Z_0 = 2.52$ ← pretty out there!

Over 2½ standard deviations

Critical value: $+Z_\alpha$ for upper one-sided H_1

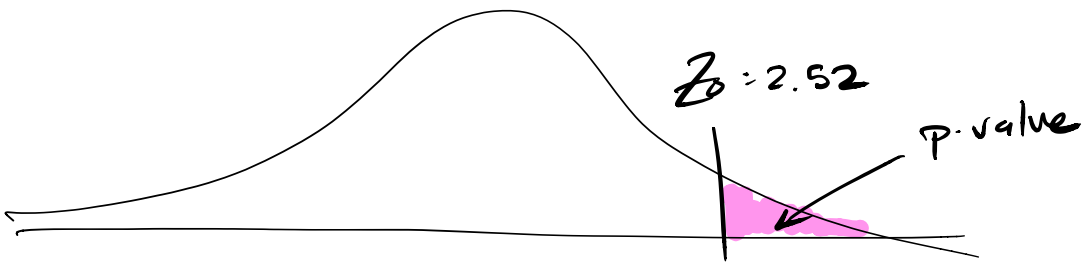


$Z_0 > +Z_\alpha$, reject H_0 @ $\alpha = 0.05$

that means are equal

∴ data suggests $\mu_1 > \mu_2$, new ingredient does shorten drying time

∴ if we had used p-value:



$$\text{p-value} = P(Z > +2.52)$$

$$= 1 - P(Z < +2.52)$$

because z-table is cumulative distribution

$$= 0.005868$$

$0.005868 < 0.05$

$\therefore$ reject H_0 @ $\alpha = 0.05$

C.I. on Difference in μ, Known σ

$$\mu_1 - \mu_2 : \quad \bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ex: tensile strength tests performed on two different grades of aluminum used for aircraft wing spars

∴ assume known pop. standard deviations;

$\sigma_1 = 1 \ kg/mm^2$ $\quad \sigma_2 = 1.5 \ kg/mm^2$

test data: $\bar{X}_1 = 87.6 \ kg/mm^2$ $\quad \bar{X}_2 = 74.5 \ kg/mm^2$

$n_1 = 10$ $\quad n_2 = 12$

∴ write a 90% C.I. on the difference in mean tensile strength for the two grades of Al

∴ need $Z_{\alpha/2}$ for 90% C.I.

$\Downarrow$

$\alpha = 0.10$

$\therefore Z_{\alpha/2} = Z_{.05} = 1.645$

$\mu_1 - \mu_2$: $87.6 - 74.5 \pm 1.645\sqrt{\frac{1^2}{10} + \frac{1.5^2}{12}}$

$$12.22 < \mu_1 - \mu_2 < 13.98 \quad kg/mm^2$$

∴ C.I. is all positive ; ∴

test data suggests $\mu_1 > \mu_2$

∴ if the C.I. had included zero

we would have failed to reject

$H_0 : \mu_1 - \mu_2 = 0$

∴ if C.I. contains Δ_0, fail to reject

∴ we can select sample size to get desired width!

$$n = \left(\frac{Z_{\alpha/2}}{E}\right)^2 \left(\sigma_1^2 + \sigma_2^2\right) \quad \text{(round up)}$$

... where E is the "single-sided error"

ex: What if C.I. on difference in mean tensile strengths of grades of aluminum used in wing spars needs to have width of ± 0.5 kg/mm²?

then $E = 0.5$

$$n = \left(\frac{1.645}{0.5}\right)^2 \left(1^2 + 1.5^2\right)$$

$$= 35.12 \rightarrow \text{round up to } 36$$

for each sample!!

∴ $n_1 = 36$, $n_2 = 36$