Comparing Two Means

A tiny bit of review:

Which statement is NOT true about one-sample t-test?

A. It is used to compare the sample mean of a variable to hypothesized value.

B. The test only assumes that data are random sample from the population.

C. The test statistic is t.

D. If the null hypothesis is true, then t should have a t-distribution with n - 1 df

What does the two-sample t-test method do?

----

a) Compares the <u>means</u> of a numerical variable between <u>two dependent</u> groups

b) Compares the <u>standard error</u> of a numerical variable between <u>two independent</u> groups

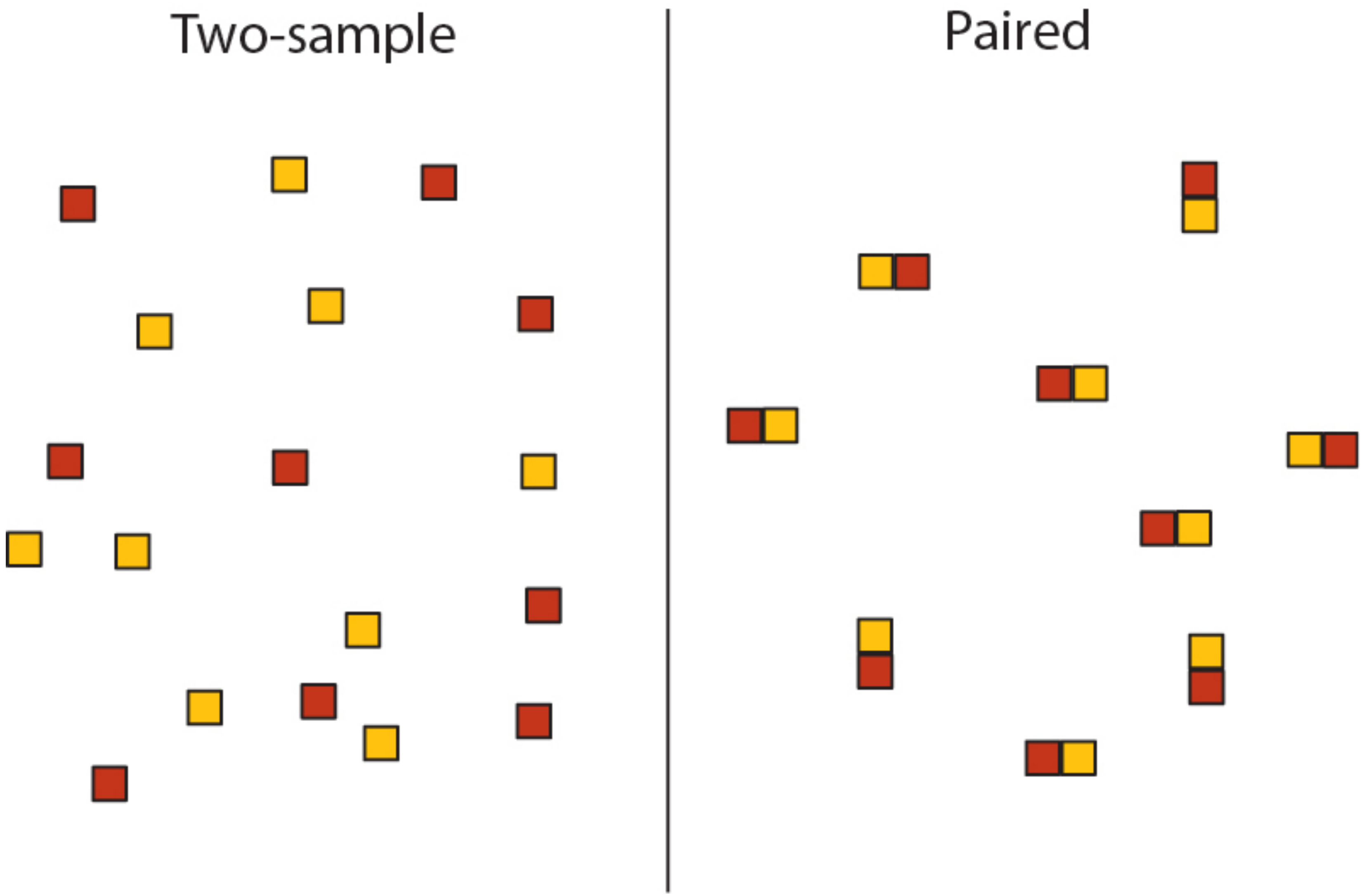
c) Compares the <u>mean</u>s of a numerical variable between <u>two independent</u> groups

d) Compares the <u>means</u> of a numerical variable <u>between two similar populations</u>

- Tests with one categorical and one numerical variable
  - Goal: compare the mean of a numerical variable among different groups

- Examine two major study design comparisons:
  - Paired Design
  - Two-Sample Design

Paired vs. Two Sample Comparisons

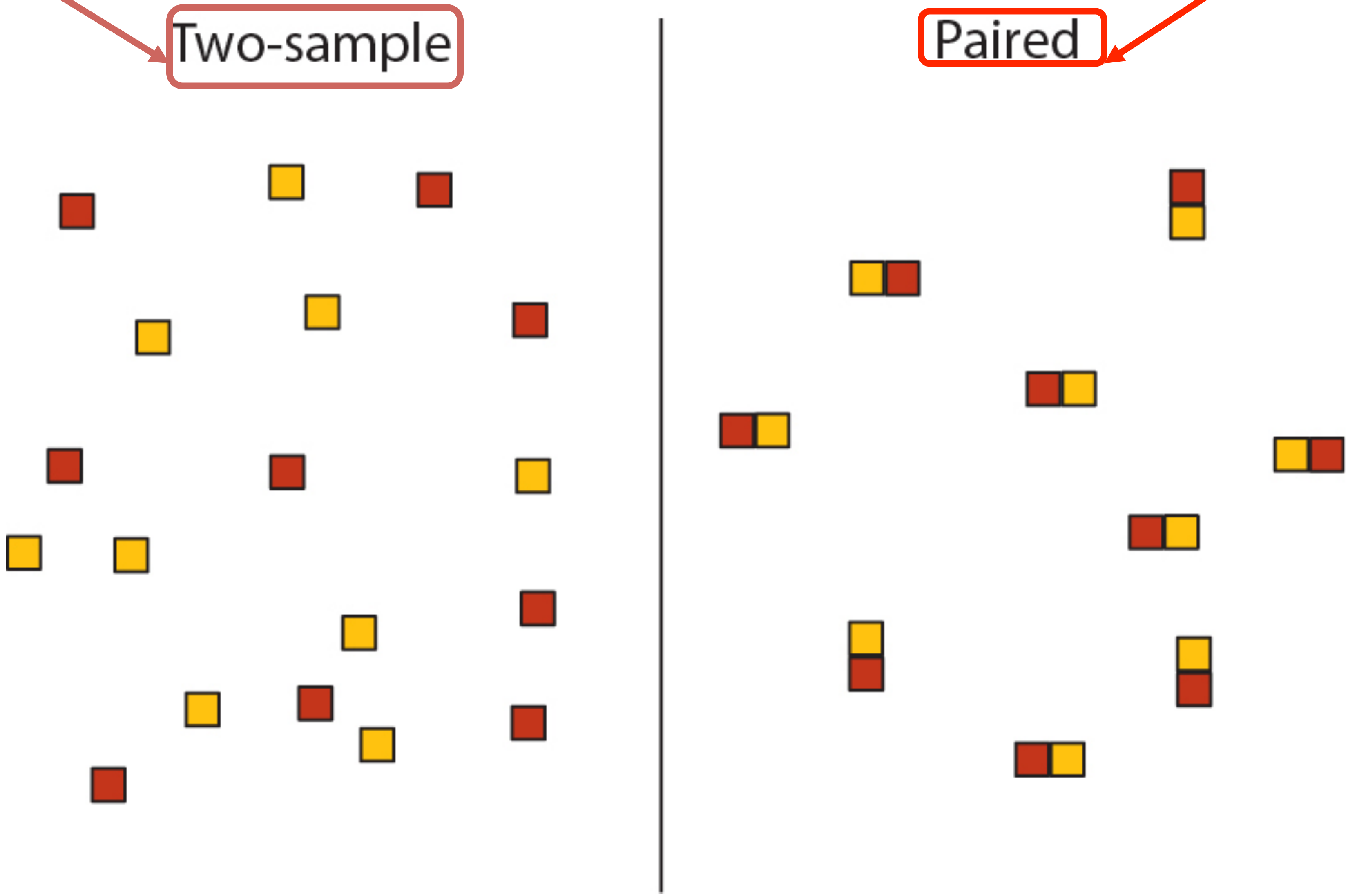


Comparing Two Means

Easier

Better

Paired vs. Two Sample Comparisons



Do scores on a test of science achievement differ for female and male 8th grade students?

a. Test one mean against a hypothesized constant.

b. Test the difference between two means (independent samples).

c. Test the difference in means between two paired or dependent samples.

d. Use a chi-squared test of association.

Do boy/girl twins differ in reading achievement?

a. Test one mean against a hypothesized constant.

b. Test the difference between two means (independent samples).

c. Test the difference in means between two paired or dependent samples.

d. Use a chi-squared test of association.

Paired Design:

- Data from two groups are paired

- Each member of the pair shares everything in common with the other *except* for the tested categorical variable
  - Reduces effects of (hidden) confounding variables

- One-to-one correspondence between the individuals in the two groups

Paired Design:

Examples:

- Before and After treatments
- Identical twins: one with treatment and one without
- One arm given treatment (sunscreen) the other arm is not on the same individual
- Testing effects of treatment (smoking) in a sample of patients, each of which is compared to a non-treatment (nonsmoker) closely matched by age, weight, ethnic background and socioeconomic condition*

Paired Design:

The sampling unit ***is the pair: one member with a treatment and a second with a different treatment***

- two measurements must be reduced to a single number which is the mean of the difference between the two measurements

  - » Ex: If there are 20 individuals grouped into 10 pairs and $\boldsymbol{n = 10}$

Paired Design:

Paired t-test

- Compares the mean of the differences to a value given in the null hypothesis
  - » Often tests the null hypothesis that the mean difference of paired measurements is equal to "0"
- **For each pair, calculate the difference. *The paired t-test is then simply the one-sample t-test on the differences*.**

Paired Design:

- Estimating the mean difference:
  - **Assumptions:**
    - Difference between members of each pair have a normal distribution
      - Distribution of the single measurements on each sampling unit do not need to be normally distributed - just the differences have to be normally distributed
    - Pairs are chosen at random

Example:

- No Smoking Day in Great Britain on the second Wednesday of March. Compared to the previous Wednesday, does (voluntarily) not smoking for a day affect injury rate?

Paired Design:

Year	Injuries Previous Wednesday	Injuries "No Smoking" Wednesday
1987	516	540
1988	610	620
1989	581	599
1990	586	639
1991	554	607
1992	632	603
1993	479	519
1994	583	560
1995	445	515
1996	522	556

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_A: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on “No smoking” Wednesday and a regular Wednesday, $\mu_d \neq 0$

Test statistic:

Assumptions: d_i are normally distributed

- paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

- Calculate $\bar{d}$, $SE_{\bar{d}}$

- Remember: *n* is the number of pairs, which are the independent sampling units

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Test statistic:

– paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

Year	Injuries Previous Wed.	Injuries "No Smoking" Wed.	d_i
1987	516	540	24
1988	610	620	10
1989	581	599	18
1990	586	639	53
1991	554	607	53
1992	632	603	-29
1993	479	519	40
1994	583	560	-23
1995	445	515	70
1996	522	556	34

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Test statistic:

- paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

Year	Injuries Previous Wed.	Injuries "No Smoking" Wed.	d_i
1987	516	540	24
1988	610	620	10
1989	581	599	18
1990	586	639	53
1991	554	607	53
1992	632	603	-29
1993	479	519	40
1994	583	560	-23
1995	445	515	70
1996	522	556	34

$\bar{d} = 25$

$SE_{\bar{d}} = 10.22$

n = 10

$dof = 9$

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Test statistic:

- paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

$$\boxed{\begin{array}{l} \bar{d} = 25 \\ SE_{\bar{d}} = 10.22 \\ dof = 9 \end{array}}$$

$$t = \frac{25 - 0}{10.22} = 2.45$$

Critical Values and Significance levels:

$\alpha = 0.05$

$t_{0.05(2),\,9} > 2.26$

We can reject the H_0: there is a difference in injury rate between Days on which people smoke and those on which people don't

Paired Design: No Smoking Wednesday Injuries

H_0: There is no difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d = 0$

H_0: There is a difference in number of injuries experienced on "No smoking" Wednesday and a regular Wednesday, $\mu_d \neq 0$

Test statistic:

- paired t-test

$$\boxed{t = \frac{\bar{d} - \mu_{d_0}}{SE_{\bar{d}}}}$$

$$\boxed{\begin{array}{l} \bar{d} = 25 \\ SE_{\bar{d}} = 10.22 \\ dof = 9 \end{array}}$$

$$t = \frac{25 - 0}{10.22} = 2.45$$

Critical Values and Significance levels:

$\alpha = 0.05$

$t_{0.05(2),\,9} > 2.26$

We can reject the H_0: there is a difference in injury rate between Days on which people smoke and those on which people don't

$$\boxed{\bar{d} - t_{\alpha(2),df} SE_{\bar{d}} < \mu_d < \bar{d} + t_{\alpha(2),df} SE_{\bar{d}}}$$

$$1.9 < \mu_d < 48.1$$

Real life example:

https://news.brown.edu/articles/2016/12/smoking

Indoor tobacco legislation is associated with fewer emergency department visits for asthma exacerbation in children

These data were presented in abstract form at the Annual Scientific Meeting of the American Academy of Allergy, Asthma, and Immunology; Houston, Texas; February 22, 2015.

Christina E. Ciaccio, MD, MSc, Tami Gurley-Calvez, PhD, Theresa I. Shireman, PhD, RPh

Background

During the past 3 decades, numerous cities and states have adopted laws that ban smoking in public indoor spaces. The rationale for these policies has been to protect nonsmokers from the adverse health effects of secondhand smoke.

Objective

To determine whether the implementation of indoor smoking legislation is associated with a decrease in emergency department **visits for asthma in children.**

Methods

This retrospective analysis used a natural experiment to estimate the impact of clean indoor air legislation on the rate of emergency department admissions for asthma exacerbation in children. Data were obtained from the Pediatric Health Information System. A Poisson regression was used for analyses and controlled for age, sex, race, payer source, seasonality, and secular trends.

Results

Asthma emergency department visits were captured from 20 hospitals in 14 different states plus the District of Columbia from July 2000 to January 2014 (n = 335,588). Indoor smoking legislation, pooled across all cities, was associated with a decreased rate of severe asthma exacerbation (adjusted rate ratio 0.83, 95% confidence interval 0.82–0.85, $P < .0001$).

In more directly relevant-to-this-lecture analysis, they also compared the 3 years prior to the ban and the 3 years after the ban at each hospital. As well, they controlled for any long term decline in smoking by comparing 6 years prior to a randomly generated date (January 1st, 2007) and the 6 years after. No significant decline in emergency room visits due to asthma attacks was found (significance was only found when comparing dates before the ban to dates after the ban).

Conclusion

Indoor tobacco legislation is associated with a decrease in emergency department visits for asthma exacerbation. Such legislation should be considered in localities that remain without this legislation to protect the respiratory health of their children.

Example:

A farmer decides to try out a new fertilizer on a test plot containing 10 stalks of corn. Before applying the fertilizer, he measures the height of each stalk. Two weeks later, he measures the stalks again. The stalks would have grown an average of four inches during that time even without the fertilizer. Did the fertilizer help? Use $\alpha = 0.05$

Stalk	1	2	3	4	5	6	7	8	9	10
Before	35.5	31.7	31.2	36.3	22.8	28.0	24.6	26.1	34.5	27.7
After	45.3	36.0	38.6	44.7	31.4	33.5	28.8	35.8	42.9	35.0

A. Yes, the fertilizer helped!
B. No, the fertilizer did not help the corn grow taller.

Example:

A farmer decides to try out a new fertilizer on a test plot containing 10 stalks of corn. Before applying the fertilizer, he measures the height of each stalk. Two weeks later, he measures the stalks again. The stalks would have grown an average of five inches during that time even without the fertilizer. Did the fertilizer help? Use α =0.05

Stalk	1	2	3	4	5	6	7	8	9	10
Before	35.5	31.7	31.2	36.3	22.8	28.0	24.6	26.1	34.5	27.7
After	45.3	36.0	38.6	44.7	31.4	33.5	28.8	35.8	42.9	35.0
Difference	**9.8**	**4.3**	**7.4**	**8.4**	**8.6**	**5.5**	**4.2**	**9.7**	**8.4**	**7.3**

Step 1:

H_0: $\mu_{After-Before} \leq 4$

H_A: $\mu_{After-Before} > 4$

sample mean difference = 73.6/10 = 7.36

$s^2 = [(9.8-7.36)^2+(4.3-7.36)^2+\ldots+(7.3-7.36)^2]/(10-1) = 4.22$

$s = \sqrt{4.216} = 2.05$

Example:

A farmer decides to try out a new fertilizer on a test plot containing 10 stalks of corn. Before applying the fertilizer, he measures the height of each stalk. Two weeks later, he measures the stalks again. The stalks would have grown an average of four inches during that time even without the fertilizer. Did the fertilizer help? Use $\alpha = 0.05$

Step 2:

$$t = \frac{7.36 - 4}{2.05/\sqrt{10}} = 5.18$$

critical value of t?

* test is one tailed so $t_{0.05,9} = 1.83$

Step 3:

since $t > t_{0.05,9}$, the null hypothesis can be rejected so the fertilizer caused more growth than would normally occur in the two week time frame.

Step 4:

Always include a confidence interval for the difference so that you can confirm your conclusion:

$$7.36-1.83*(2.05/\sqrt{10}) < \mu_d < 7.36+1.83*(2.05/\sqrt{10})$$

$$6.17 < \mu_d < 8.55$$

(note: this 95% CI does not contain '4')