Warm Up

AP MC Practice

EXPERIMENTAL DESIGN 2

An advantage to using surveys as opposed to experiments is that

(A) surveys are generally cheaper to conduct.

(B) it is generally easier to conclude cause and effect from surveys.

(C) surveys are generally not subject to bias.

(D) surveys involve use of randomization.

(E) surveys can make use of stratification.

Answer: (A) Surveys are generally cheaper and quicker to conduct than experiments; however surveys are subject to bias, and it is very difficult to conclude cause and effect from surveys. Experiments also use randomizati in the form of random assignment to treatments. Blocking in experimental design corresponds to stratification in sampling design.

EXPERIMENTAL DESIGN 3

A company wishes to survey what people think about a new product it plans to market. They decide to randomly sample from their customer database as this includes phone numbers and addresses. This procedure is an example of which type of sampling?

(A) Cluster (B) Convenience (C) Simple random

(D) Stratified (E) Systematic

Answer: (B) Convenience samples are based on choosing individuals who are easy to reach. A typical example is sampling based on interviews at a shopping mall. Data obtained from convenience samples tends to be highly unrepresentative of the entire population. In this example, while using the company database is convenient, the resulting data tells nothing about what people outside the customer base think about the new product.

EXPERIMENTAL DESIGN 9

A sales representative wishes to survey her client base of 47 companies. She has 47 business cards, all of the identical size, from her contacts in the companies, and decides to drop them all in a small box, shake them up, and reach in to pick 5 cards for her sample. This procedure is an example of which type of sampling?

(A) Cluster (B) Convenience (C) Simple random

(D) Stratified (E) Systematic

Answer: (C) A simple random sample (SRS) is one in which every possible sample of the desired size has an equal chance of being selected. In this case, every possible sample of five companies has an equal chance of being selected. Note that even though it is also true that each company has an equal chance of being selected, this by itself would not ensure that we have an SRS.

EXPERIMENTAL DESIGN 1

Each of the 30 major league baseball teams carries a 40-person roster. A sample of 60 players (5 percent of all 1,200 players) is to be randomly selected to undergo drug tests. To do this, each team is instructed to put their 40 names in a hat and randomly draw two names. Will this method result in a simple random sample of the 1,200 baseball players?

(A) Yes, because each player has the same chance of being selected.

(B) Yes, because each team is equally represented.

(C) Yes, because this is an example of stratified sampling, which is a special case of simple random sampling.

(D) No, because the teams are not chosen randomly.

(E) No, because not each group of 60 players has the same chance of being selected.

Answer: **(E) In a simple random sample, every possible group of the given size has to be equally likely to be selected, and this is not true here. For example, with this procedure, it will be impossible for all the Cubs to be together in the final sample. This procedure is an example of stratified sampling, but stratified sampling does not result in simple random samples.**

EXPERIMENTAL DESIGN 24

Before taking an exam, students either went to bed at their normal times or were sleep deprived for 4 or 8 hours. Half of each group were given a caffeine pill before taking the exam. Determine the number of factors, levels for each, and number of treatments.

(A) One factor with two levels, five treatments

(B) Two factors, one with one and one with two levels, three treatments

(C) Two factors, one with two and one with three levels, five treatments

(D) Two factors, one with two and one with three levels, six treatments

(E) Three factors, each with two levels, six treatments

Answer: (D) Two factors, sleep deprivation (three levels) and caffeine (two levels), with $3 \times 2 = 6$ treatments.

10.2 Day 2

Today, you will continue to deal with *differences of means* (μ_1-μ_2), but instead of creating confidence intervals you will switch back to conducting a hypotheses test instead.

10.2 Day 2

Incorporating Components of a Hypothesis Test for a Difference of Two means

1. **Which version of the exam is harder?** A college professor has created two different versions of a final exam. She wonders if the difficulty is the same for each version. To find out, she randomly assigns 80 student volunteers to two groups: 40 students take Version A, and 40 students take Version B. Here are the results:

Work up through part d

Version	Mean	SD
A	84.2	8.9
B	79.9	12.3

Do these data give convincing evidence at the $\alpha = 0.05$ significance level of a difference in the true mean score on the final exam for Version A and Version B for students like the ones in the study?

Need your AP Formula sheet

(a) State appropriate hypotheses for performing a significance test. Be sure to define the parameters of interest. Try to incorporate subscripts that go with the context of the situation.

$$H_0: \mu_A - \mu_B = 0$$

$$H_a: \mu_A - \mu_B \neq 0$$

where μ_A = the true mean score on Version A of the final exam for students like the ones in the study and μ_B = the true mean score on Version B of the final exam for students like the ones in the study.

(b) Good news. The Random, 10%, and Normal/Large Sample conditions are exactly the same for significance tests for a difference in means as they were for a confidence interval for a difference in means. WooHoo! Go ahead and check the conditions.

Random: The 80 subjects were randomly assigned to Version A or Version B. ✓

Normal/Large Sample: $n_A = 40 \geq 30$ and $n_B = 40 \geq 30$. ✓

(c) The table of information has been updated below to include the standard deviation and the sample size of each of the randomly assigned groups.

Version	Number of students	Mean	SD
A	40	84.2	8.9
B	40	79.9	12.3

Explain why the sample results give some evidence for the alternative hypothesis.

(c) The table of information has been updated below to include the standard deviation and the sample size of each of the randomly assigned groups.

Version	Number of students	Mean	SD
A	40	84.2	8.9
B	40	79.9	12.3

Explain why the sample results give some evidence for the alternative hypothesis.

The observed difference in the sample means is $\bar{x}_A - \bar{x}_B = 84.2 - 79.9 = 4.3$, which gives some evidence in favor of Ha because $4.3 \neq 0$

(d) Looking only at your AP Formula Sheet, attempt to find/create the specific formula for the standardized test statistic that you will need in this situation. Try to avoid looking at other sources for now. Mr. Cedarlund will verify the formula.

$\rightarrow$

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) • (standard deviation of statistic)

$(\bar{x}_A - \bar{x}_B)$

$\mu_A - \mu_B = 0$

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$ → see below

Confidence interval: statistic $\pm$ (critical value) • (standard deviation of statistic)

Statistic	Standard Deviation of Statistic
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

but use s_A s_B

$$t = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}}$$

(d) Looking only at your AP Formula Sheet, attempt to find/create the specific formula for the standardized test statistic that you will need in this situation. Try to avoid looking at other sources for now. Mr. Cedarlund will verify the formula.

$\mu_A - \mu_B$

$$t = \frac{(\bar{x}_A - \bar{x}_B) - 0}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

or

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(e) First calculate the standardized test statistic and the P-value using Option 2

Option 2: Use the smaller degrees of freedom. Then use Table B or t_{cdf} to calculate the P-value

(f) Now Re-calculate the standardized test statistic and the P-value using Option 1

(Option 1: Use 2-SampTTest on your calculator. Report the t statistic, P-value, and the df.)

(e) First calculate the standardized test statistic and the P-value using Option 2
(Option 2: Use the smaller degrees of freedom. Then use Table B or t_{cdf} to calculate the P-value)

$$t = \frac{(84.2 - 79.9) - 0}{\sqrt{\frac{8.9^2}{40} + \frac{12.3^2}{40}}} = 1.79 \rightarrow df = 39 \rightarrow \text{TABLE B} \rightarrow$$

P-Value between .025 and .05
Two Tailed
2(.025) and 2(.05)
$.05 < \text{P-Value} < 0.10$

or t_{cdf}[lower 1.79, upper 1000, 39 (df)] x2
= .081

(e) First calculate the standardized test statistic and the P-value using Option 2
(Option 2: Use the smaller degrees of freedom. Then use Table B or t_{cdf} to calculate the P-value)

$$t = \frac{(84.2 - 79.9) - 0}{\sqrt{\frac{8.9^2}{40} + \frac{12.3^2}{40}}} = 1.79 \rightarrow df = 39 \rightarrow \text{TABLE B} \rightarrow$$

P-Value between .025 and .05
Two Tailed
2(.025) and 2(.05)
$.05 < \text{P-Value} < 0.10$

or t_{cdf}[lower 1.79, upper 1000, 39] x2
= .081

(f) Now Re-calculate the standardized test statistic and the P-value using Option 1
(Option 1: Use 2-SampTTest on your calculator. Report the t statistic, P-value, and the df.)

2-SampTTest gives $t = 1.79$ and P-value = .078 using df = 71.05

(g) What conclusion would you make?

(g) What conclusion would you make?

Because the P-Value of $0.078 > \alpha = .05$, we fail to reject H_0

There is not convincing evidence of a difference in the true mean score on Version A and Version B of the final exam for students like the ones in the study.

Requirements for **Hypotheses Tests** for a difference of Means

State

- State Hypotheses (using subcripts and be informative and helpful)
- Give significance level
- Define both parameters

Plan

- Identify the procedure
- State and check conditons (In experiments, we are not sampling w/o replacement so don't check 10% cond.)

Do

- Give statistics
- Give standardized test statistics
- Give P-value

Option 1: *example:* $2SampTTest\ gives\ t = 3.27\ and\ P - Value = 0.076\ using\ df = 13.2$

Option 2: *example:* $df = 12$

Table B: $P - Value\ is\ between\ 0.05\ and\ 0.10$

Or with technology: $tcd\ (lower: 3.27, upper: 1000, df: 12)$=0.065

Plan

Two-sentence structure

- Compare P-value to significance level, then reject or fail to reject H_0.
- Evidence for H_a in context.

It's time for:

"Hot Potato INFERENCE"

(1) PAIRS

(2) ONe paper

(3) ONe person starts with it
The other observes/coaches

(4) Until....
Mr. C says.. Rotate!!!

Put it All Together - **Significance Test for a Difference in Means**

How quickly do synthetic fabrics such as polyester decay in landfills? A researcher buried polyester strips in the soil for different lengths of time, then dug up the strips and measured the force required to break them. Breaking strength is easy to measure and is a good indicator of decay. Lower strength means the fabric has decayed more. For one part of the study, the researcher buried 10 strips of polyester fabric in well-drained soil in the summer. The strips were randomly assigned to two groups: 5 of them were buried for 2 weeks and the other 5 were buried for 16 weeks. Here are the breaking strengths in pounds:

Group 1 (2 weeks)	118	126	126	120	129
Group 2 (16 weeks)	124	98	110	140	110

Do the data give convincing evidence that polyester decays more in 16 weeks than in 2 weeks, on average

Group 1 (2 weeks)	118	126	126	120	129
Group 2 (16 weeks)	124	98	110	140	110

$\bar{x}_1 = 123.8$ $s_1 = 4.60$
$\bar{x}_2 = 116.4$ $s_2 = 16.09$

Do the data give convincing evidence that polyester decays more in 16 weeks than in 2 weeks, on average

STATE

$H_0: \mu_1 - \mu_2 = 0$
$H_a: \mu_1 - \mu_2 > 0$

where μ_1 = true mean breaking strength for polyster fabric buried for 2 weeks

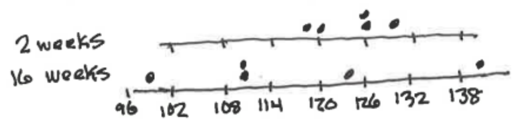
μ_2 = true mean breaking strength after 16 weeks

PLAN

Two-Sample t test for $\mu_1 - \mu_2$

Random: strips randomly assigned to the treatments ✓

Normal/Large Counts: Dot plots show no strong skewness and no outliers ✓



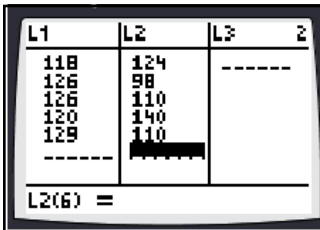
DO

$\bar{x}_1 = 123.8$ $\bar{x}_2 = 116.4$
$s_1 = 4.60$ $s_2 = 16.09$
$n_1 = 5$ $n_2 = 5$

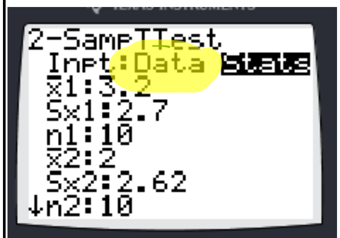
2SampTTest gives $t = 0.99$ and P-Value $= .1857$ using $df = 13.2$

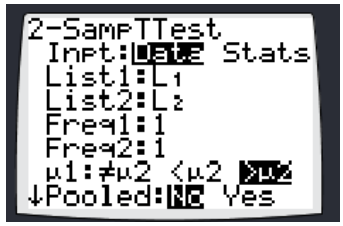
CONCLUDE

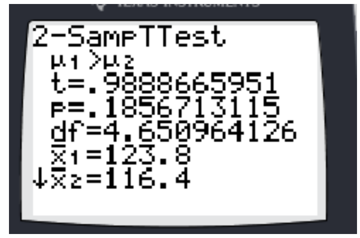
Because the P-Value of $0.1857 > \alpha = .05$, we fail to reject Ho. We do not have convincing evidence that the true mean breaking strength of polyester fabric that is buried for 2 weeks is greater than the same fabric that is buried for 16 weeks



Entering Data
for 2 SampTTest







A Word about "Pooled"

Most software offers a choice of two-sample *t* procedures. One is often labeled "unequal" variances; the other, "equal" variances.

The **unequal variance** procedures use our formula for the two-sample *t* interval and test. *This interval and test are valid whether or not the population variances are equal.*

The **equal variance** procedures assumes the two population distributions have the same variance. This procedure combines (the statistical term is ***pools***) the two sample variances to estimate the common population variance. The resulting statistic is called the *pooled two-sample t statistic.*

A Word about "Pooled"

Most software offers a choice of two-sample *t* procedures. One is often labeled "unequal" variances; the other, "equal" variances.

The **unequal variance** procedures use our formula for the two-sample *t* interval and test. *This interval and test are valid whether or not the population variances are equal.*

The **equal variance** procedures assumes the two population distributions have the same variance. This procedure combines (the statistical term is ***pools***) the two sample variances to estimate the common population variance. The resulting statistic is called the *pooled two-sample t statistic.*

Our advice: *Never use the pooled t procedures if you have technology that will carry out Option 1.*

∴ cross out

(I) Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n}$$

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

$$\cancel{s_p = \sqrt{\frac{\left(n_1-1\right)s_1^2 + \left(n_2-1\right)s_2^2}{\left(n_1-1\right)+\left(n_2-1\right)}}}$$

10.2....51-57(odd), 67, 69-72

and study pp. 673-682 !