Chapter 10

Inferences Involving Two Populations

Chapter 10 Overview

- 10-2 Inferences Concerning the Mean Difference Using Two Dependent Samples
- 10-3 Inferences Concerning the Difference between Means Using Two Independent Samples

Independent vs Dependent Samples

- The battle of the sexes can take on many forms. When the topic of which gender is the better, faster, or safer driver comes up on campus, the battles could get quite competitive!
- Once the dust clears one could also ask, "Who drives the longest commute to college?" The length of commute can be measured in distance (miles) or in time (minutes); and there are many factors that play a role for commuting students.
- Do they live at home? Do they work a part-time or full-time job? Do they have family obligations?

Independent vs Dependent Samples

- Male students and female students are two populations. In this chapter we are going to study the procedures for making inferences about two populations.
- When comparing two populations, we need two samples, one from each population.
- Two basic kinds of samples can be used: independent and dependent.

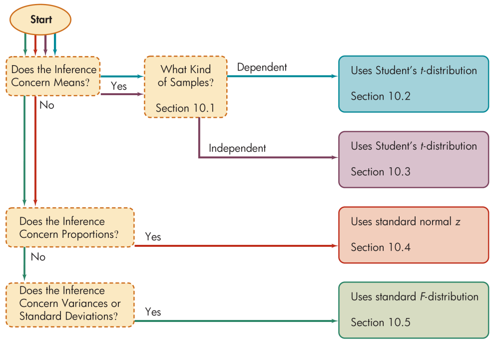
Independent vs Dependent Samples

- The dependence or independence of two samples is determined by the sources of the data.

- A **source** can be a person, an object, or anything else that yields a data value. If the same set of sources or related sets are used to obtain the data representing both populations, we have **dependent samples**.

- If two unrelated sets of sources are used, one set from each population, we have **independent samples**.

Independent vs Dependent Samples



"Road Map" to Two Population Inferences

Figure 10.1

Example 1 – *Dependent Versus Independent Samples*

- A test will be conducted to see whether the participants in a physical-fitness class actually improve in their level of fitness. It is anticipated that approximately 500 people will sign up for this course.

- The instructor decides that she will give 50 of the participants a set of tests before the course begins (a pre-test), and then she will give another set of tests to 50 participants at the end of the course (a post-test).

Example 1 – *Dependent Versus Independent Samples* cont'd

Two sampling procedures are proposed:

- Plan A: Randomly select 50 participants from the list of those enrolled and give them the pre-test. At the end of the course, make a second random selection of size 50 and give them the post-test.

- Plan B: Randomly select 50 participants and give them the pre-test; give the same set of 50 the post-test when they complete the course.

Example 1 – *Dependent Versus Independent Samples* cont'd

- Plan A illustrates independent sampling; the sources (the class participants) used for each sample (pre-test and post-test) were selected separately.

- Plan B illustrates dependent sampling; the sources used for both samples (pre-test and post-test) are the same.

Independent vs Dependent Samples

- Typically, when both a pre-test and a post-test are used, the same subjects participate in the study.

- Thus, pre-test versus post-test (before versus after) studies usually use dependent samples.

Example 2 – *Dependent Versus Independent Samples*

- A test is being designed to compare the wearing quality of two brands of automobile tires.

- The automobiles will be selected and equipped with the new tires and then driven under "normal" conditions for 1 month. Then a measurement will be taken to determine how much wear took place.

Example 2 – *Dependent Versus Independent Samples* cont'd

Two plans are proposed:

- Plan C: A sample of cars will be selected randomly, equipped with brand A tires, and driven for 1 month. Another sample of cars will be selected, equipped with brand B tires, and driven for 1 month.

- Plan D: A sample of cars will be selected randomly, equipped with one tire of brand A and one tire of brand B (the other two tires are not part of the test), and driven for 1 month.

Example 2 – *Dependent Versus Independent Samples*

cont'd

- We suspect that many other factors must be taken into account when testing automobile tires—such as age, weight, and mechanical condition of the car; driving habits of drivers; location of the tire on the car; and where and how much the car is driven.

- However, at this time we are trying only to illustrate dependent and independent samples. Plan C is independent (unrelated sources), and plan D is dependent (common sources).

10.2 Inferences Concerning the Mean Difference Using Two Dependent Samples

The procedures for comparing two population means are based on the relationship between two sets of sample data, one sample from each population.

When dependent samples are involved, the data are thought of as "paired data."

The data may be paired as a result of being obtained from "before" and "after" studies or from matching two subjects with similar traits to form "matched pairs."

Inferences Concerning the Mean Difference Using Two Dependent Samples

The pairs of data values are compared directly to each other by using the difference in their numerical values.

The resulting difference is called a **paired difference.**

Paired Difference

$$d = x_1 - x_2 \tag{10.1}$$

Using paired data this way has a built-in ability to remove the effect of otherwise uncontrolled factors.

Inferences Concerning the Mean Difference Using Two Dependent Samples

The wearing ability of a tire is greatly affected by a multitude of factors: the size, weight, age, and condition of the car; the driving habits of the driver; the number of miles driven; the condition and types of roads driven on; the quality of the material used to make the tire; and so on.

We create paired data by mounting one tire from each brand on the same car.

Since one tire of each brand will be tested under the same conditions, using the same car, same driver, and so on, the extraneous causes of wear are neutralized.

Procedures and Assumptions for Inferences Involving Paired Data

A test was conducted to compare the wearing quality of the tires produced by two tire companies. All the aforementioned factors had an equal effect on both brands of tires, car by car.

One tire of each brand was placed on each of six test cars. The position (left or right side, front or back) was determined with the aid of a random-number table.

Procedures and Assumptions for Inferences Involving Paired Data

Table 10.1 lists the amounts of wear (in thousandths of an inch) that resulted from the test.

Car	1	2	3	4	5	6
Brand A	125	64	94	38	90	106
Brand B	133	65	103	37	102	115

Amount of Tire Wear **[TA10-01]**

Table 10.1

Since the various cars, drivers, and conditions were the same for each tire of a paired set of data, it makes sense to use a third variable, the paired difference ***d***.

Procedures and Assumptions for Inferences Involving Paired Data

Our two dependent samples of data may be combined into one set of *d* values, where *d* = B – A.

Car	1	2	3	4	5	6
$d = B - A$	8	1	9	−1	12	9

The difference between the two population means, when dependent samples are used (often called **dependent means**), is equivalent to the **mean of the paired differences**.

Procedures and Assumptions for Inferences Involving Paired Data

Therefore, when an inference is to be made about the difference of two means and paired differences are used, the inference will in fact be about the mean of the paired differences.

The sample mean of the paired differences will be used as the point estimate for these inferences.

In order to make inferences about the mean of all possible paired differences μ_d, we need to know the *sampling distribution* of $\bar{d}$.

Procedures and Assumptions for Inferences Involving Paired Data

When paired observations are randomly selected from normal populations, the paired difference, d = $x_1 - x_2$ will be approximately normally distributed about a mean μ_d with a standard deviation of σ_d.

This is another situation in which the *t*-test for one mean is applied; namely, we wish to make inferences about an unknown mean (μ_d) where the random variable (d) involved has an approximately normal distribution with an unknown standard deviation (σ_d).

Procedures and Assumptions for Inferences Involving Paired Data

Inferences about the mean of all possible paired differences μ_d are based on samples of n dependent pairs of data and the ***t*-distribution** with $n - 1$ degrees of freedom (df), under the following assumption:

Assumption for inferences about the mean of paired differences μ_d The paired data are randomly selected from normally distributed populations.

Confidence Interval Procedure

The $1 - \alpha$ **confidence interval for estimating the mean difference** μ_d is found using this formula:

Confidence Interval for Mean Difference (Dependent Samples)

$$\bar{d} - t(\text{df}, \alpha/2) \cdot \frac{s_d}{\sqrt{n}} \quad \text{to} \quad \bar{d} + t(\text{df}, \alpha/2) \cdot \frac{s_d}{\sqrt{n}}, \text{ where df} = n - 1 \qquad (10.2)$$

where $\bar{d}$ is the mean of the sample differences:

$$\bar{d} = \frac{\Sigma d}{n} \qquad (10.3)$$

Confidence Interval Procedure

and s_d is the standard deviation of the sample differences:

$$s_d = \sqrt{\frac{\Sigma d^2 - \frac{(\Sigma d)^2}{n}}{n - 1}} \qquad (10.4)$$

Confidence Interval Procedure

Note: This confidence interval is quite wide, in part because of the small sample size. We know from the central limit theorem that as the sample size increases, the standard error (estimated by $s_d/\sqrt{n}$) decreases.

Hypothesis-Testing Procedure

When we test a null **hypothesis about the mean difference,** the test statistic used will be the difference between the sample mean $\bar{d}$ and the hypothesized value of μ_d, divided by the estimated **standard error**.

This statistic is assumed to have a *t*-distribution when the null hypothesis is true and the assumptions for the test are satisfied.

Hypothesis-Testing Procedure

The value of the **test statistic** $t\star$ is calculated as follows:

Test Statistic for Mean Difference (Dependent Samples)

$$t\star = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}, \text{ where df} = n - 1 \qquad (10.5)$$

Note: A hypothesized mean difference, μ_d, can be any specified value. The most common value specified is zero; however, the difference can be nonzero.

Hypothesis-Testing Procedure

Statistical significance does not always have the same meaning when the “practical” application of the results is considered. In the preceding detailed hypothesis test, the results showed a statistical significance with a p-value of 0.002—that is, 2 chances in 1000.

However, a more practical question might be: “Is lowering the pulse rate by this small average amount, estimated to be 1.07 beats per minute, worth the risks of possible side effects of this medication?” Actually, the whole issue is much broader than just this one issue of pulse rate.

Example 3: Hypothesis Testing

A physical education director claims by taking a special vitamin, a weight lifter can increase his strength. Eight athletes are selected and given a test of strength, using the standard bench press. After 2 weeks of regular training, supplemented with the vitamin, they are tested again. Test the effectiveness of the vitamin regimen at $\alpha = 0.05$. Each value in the data represents the maximum number of pounds the athlete can bench-press. Assume that the variable is approximately normally distributed.

Athlete	1	2	3	4	5	6	7	8
Before (X_1)	210	230	182	205	262	253	219	216
After (X_2)	219	236	179	204	270	250	222	216

Example 3: Hypothesis Testing

Athlete	1	2	3	4	5	6	7	8
Before (X_1)	210	230	182	205	262	253	219	216
After (X_2)	219	236	179	204	270	250	222	216

Step 1: State the hypotheses and identify the claim.

H_0: $\mu_D = 0$ and H_1: $\mu_D < 0$ (claim)

Step 2: Find the critical value.

The degrees of freedom are $n - 1 = 8 - 1 = 7$. The critical value for a left-tailed test with $\alpha = 0.05$ is $t = -1.895$.

Example 3: Hypothesis Testing

Step 3: Compute the test value.

Before (X_1)	After (X_2)	$D = X_1 - X_2$	D^2
210	219	-9	81
230	236	-6	36
182	179	3	9
205	204	1	1
262	270	-8	64
253	250	3	9
219	222	-3	9
216	216	0	0
		$\Sigma D = -19$	$\Sigma D^2 = 209$

$$\bar{D} = \frac{\sum D}{n} = \frac{-19}{8} = -2.375$$

$$s_D = \sqrt{\frac{n\sum D^2 - \left(\sum D\right)^2}{n(n-1)}} = \sqrt{\frac{8 \cdot 209 - (-19)^2}{8 \cdot 7}} = 4.84$$

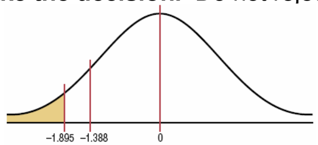
Example 3: Hypothesis Testing

Step 3: Compute the test value.

$$\bar{D} = -2.375, \quad s_D = 4.84$$

$$t = \frac{\bar{D} - \mu_D}{s_D / \sqrt{n}} = \frac{-2.375 - 0}{4.84/\sqrt{8}} = -1.388$$

Step 4: Make the decision. Do not reject the null.



Step 5: Summarize the results.

There is not enough evidence to support the claim that the vitamin increases the strength of weight lifters.

Example 4: Hypothesis Testing

A dietitian wishes to see if a person's cholesterol level will change if the diet is supplemented by a certain mineral. Six subjects were pretested, and then they took the mineral supplement for a 6-week period. The results are shown in the table. (Cholesterol level is measured in milligrams per deciliter.) Can it be concluded that the cholesterol level has been changed at α = 0.10? Assume the variable is approximately normally distributed.

Subject	1	2	3	4	5	6
Before (X_1)	210	235	208	190	172	244
After (X_2)	190	170	210	188	173	228

Example 4: Hypothesis Testing

Subject	1	2	3	4	5	6
Before (X_1)	210	235	208	190	172	244
After (X_2)	190	170	210	188	173	228

Step 1: State the hypotheses and identify the claim.

H_0: μ_D = 0 and H_1: $\mu_D \neq 0$ (claim)

Step 2: Find the critical value.

The degrees of freedom are 5. At α = 0.10, the critical values are ±2.015.

Example 4: Hypothesis Testing

Step 3: Compute the test value.

Before (X_1)	After (X_2)	$D = X_1 - X_2$	D^2
210	190	20	400
235	170	65	4225
208	210	-2	4
190	188	2	4
172	173	-1	1
244	228	16	256
		ΣD = 100	ΣD^2 = 4890

$$\bar{D} = \frac{\sum D}{n} = \frac{100}{6} = 16.7$$

$$s_D = \sqrt{\frac{n\sum D^2 - \left(\sum D\right)^2}{n(n-1)}} = \sqrt{\frac{6 \cdot 4890 - (100)^2}{6 \cdot 5}} = 25.4$$

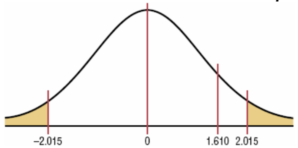
Example 4: Hypothesis Testing

Step 3: Compute the test value.

$$\bar{D} = 16.7, \; s_D = 25.4$$

$$t = \frac{\bar{D} - \mu_D}{s_D / \sqrt{n}} = \frac{16.7 - 0}{25.4/\sqrt{6}} = 1.610$$

Step 4: Make the decision. Do not reject the null.



Step 5: Summarize the results.

There is not enough evidence to support the claim that the mineral changes a person's cholesterol level.

Example 5: Confidence Interval

Find the 90% confidence interval for the difference between the means for the data in Example 4.

$$\bar{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \bar{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

$$16.7 - 2.015\cdot\frac{25.4}{\sqrt{6}} < \mu_D < 16.7 + 2.015\cdot\frac{25.4}{\sqrt{6}}$$

$$16.7 - 20.89 < \mu_D < 16.7 + 20.89$$

$$-4.19 < \mu_D < 37.59$$

Since 0 is contained in the interval, the decision is to not reject the null hypothesis H_0: $\mu_D = 0$.

Class Activity

- Find the 95% confidence interval for μ_d given $n = 26$, $\bar{d} = 6.3$, and $s_d = 5.1$.

Class Activity: Testing the Difference Between 2 Means

Wind Speeds The average wind speed in Casper, Wyoming, has been found to be 12.7 miles per hour, and in Phoenix, Arizona, it is 6.2 miles per hour. To test the relationship between the averages, the average wind speed was calculated for a sample of 31 days for each city. The results are reported below. Is there sufficient evidence at $\alpha = 0.05$ to conclude that the average wind speed is greater in Casper than in Phoenix?

	Casper	Phoenix
Sample size	31	31
Sample mean	12.85 mph	7.9 mph

10.3 Inferences Concerning the Difference between Means Using Two Independent Samples

When comparing the means of two populations, we typically consider the difference between their means, $\mu_1 - \mu_2$ (often called "**independent means**").

The inferences about $\mu_1 - \mu_2$ will be based on the difference between the observed sample means, $\bar{x}_1 - \bar{x}_2$.

This observed difference, $\bar{x}_1 - \bar{x}_2$, belongs to a sampling distribution with the characteristics described in the following statement.

Inferences Concerning the Difference between Means Using Two Independent Samples

If independent samples of sizes n_1 and n_2 are drawn randomly from large populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively, then the sampling distribution of $\bar{x}_1 - \bar{x}_2$, the difference between the sample means, has

1. mean $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$ and
2. standard error $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)}$. **(10.6)**

If both populations have normal distributions, then the sampling distribution of $\bar{x}_1 - \bar{x}_2$ will also be normally distributed.

Inferences Concerning the Difference between Means Using Two Independent Samples

The preceding statement is true for all sample sizes as long as the populations involved are normal and the population variances σ_1^2 and σ_2^2 are known quantities. However, as with inferences about one mean, the variance of a population is generally an unknown quantity.

Therefore, it will be necessary to estimate the standard error by replacing the variances σ_1^2 and σ_2^2, in formula (10.6) with the best estimates available—namely, the sample variances s_1^2 and s_2^2. The *estimated standard error* will be found using the following formula:

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

Inferences Concerning the Difference between Means Using Two Independent Samples

Inferences about the difference between two population means, $\mu_1 - \mu_2$, will be based on the following assumptions.

Assumptions for inferences about the difference between two means, $\mu_1 - \mu_2$ The samples are randomly selected from normally distributed populations, and the samples are selected in an independent manner.

No assumptions are made about the population variances.

Inferences Concerning the Difference between Means Using Two Independent Samples

Since the samples provide the information for determining the standard error, the ***t*-distribution** will be used as the test statistic. The inferences are divided into two cases.

Case 1: The *t*-distribution will be used, and the number of degrees of freedom will be calculated.

Case 2: The *t*-distribution will be used, and the number of degrees of freedom will be approximated.

Inferences Concerning the Difference between Means Using Two Independent Samples

Case 1 will occur when you are completing the inference *using a computer or statistical calculator and the statistical software or program calculates the number of degrees of freedom* for you.

The calculated value for df is a function of both sample sizes and their relative sizes, and both sample variances and their relative sizes.

The value of df will be a number between the smaller of $df_1 = n_1 - 1$ or $df_2 = n_2 - 1$, and the sum of the degrees of freedom, $df_1 + df_2 = [(n_1 - 1) + (n_2 - 1)] = n_1 + n_2 - 2$

Inferences Concerning the Difference between Means Using Two Independent Samples

Case 2 will occur when you are completing the inference *without the aid of a computer or calculator and its statistical software package.*

Use of the *t*-distribution with the smaller of $df_1 = n_1 - 1$ or $df_2 = n_2 - 1$ will give *conservative* results.

Because of this approximation, the true level of confidence for an interval estimate will be slightly higher than the reported level of confidence; or the true *p*-value and the true level of significance for a hypothesis test will be slightly less than reported.

Inferences Concerning the Difference between Means Using Two Independent Samples

The gap between these reported values and the true values will be quite small, unless the sample sizes are quite small and unequal or the sample variances are very different. The gap will decrease as the samples increase in size or as the sample variances become more alike.

Since the only difference between the two cases is the number of degrees of freedom used to identify the *t*-distribution involved, we will study case 2 first.

Inferences Concerning the Difference between Means Using Two Independent Samples

Note

$A > B$ ("*A* is greater than *B*") is equivalent to $B < A$("*B* is less than *A*"). When the difference between *A* and *B* is being discussed, it is customary to express the difference as "larger – smaller" so that the resulting difference is positive: $A - B > 0$.

Expressing the difference as "smaller – larger" results in $B - A < 0$ (the difference is negative) and is usually unnecessarily confusing. Therefore, it is recommended that the difference be expressed as "larger – smaller."

Confidence Interval Procedure

We will use the following formula to calculate the end points of the **1 – α confidence interval**.

Confidence Interval for the Difference between Two Means (Independent Samples)

$$(\bar{x}_1 - \bar{x}_2) - t(\text{df}, \alpha/2) \cdot \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \quad \text{to} \quad (\bar{x}_1 - \bar{x}_2) + t(\text{df}, \alpha/2) \cdot \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \qquad (10.8)$$

where df is either calculated or is the smaller of df_1 or df_2.

Hypothesis-Testing Procedure

When we test a null **hypothesis about the difference between two population means**, the test statistic used will be the difference between the observed difference of the sample means and the hypothesized difference of the population means, divided by the estimated standard error.

The test statistic is assumed to have approximately a *t*-distribution when the null hypothesis is true and the normality assumption has been satisfied.

Hypothesis-Testing Procedure

The calculated value of the **test statistic** is found using this formula:

Test Statistic for the Difference between Two Means (Independent Samples)

$$t\bigstar = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}} \qquad (10.9)$$

where df is either calculated or is the smaller of df_1 or df_2

Hypothesis-Testing Procedure

Note: A hypothesized difference between the two population means, $\mu_1 - \mu_2$, can be any specified value. The most common value specified is zero; however, the difference can be nonzero.

Example 6

The average size of a farm in Indiana County, Pennsylvania, is 191 acres. The average size of a farm in Greene County, Pennsylvania, is 199 acres. Assume the data were obtained from two samples with standard deviations of 38 and 12 acres, respectively, and sample sizes of 8 and 10, respectively. Can it be concluded at $\alpha = 0.05$ that the average size of the farms in the two counties is different? Assume the populations are normally distributed.

Step 1: State the hypotheses and identify the claim.

H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$ (claim)

Example 6

Step 2: Find the critical values.

Since the test is two-tailed, a = 0.05, and the variances are unequal, the degrees of freedom are the smaller of $n_1 - 1$ or $n_2 - 1$. In this case, the degrees of freedom are 8 – 1 = 7. Hence, from Table 6, the critical values are -2.365 and 2.365.

Step 3: Find the test value.

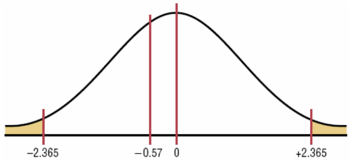
$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(191 - 199) - (0)}{\sqrt{\dfrac{38^2}{8} + \dfrac{12^2}{10}}}$$

$$= -0.57$$

Example 6

Step 4: Make the decision.

Do not reject the null hypothesis.



Step 5: Summarize the results.

There is not enough evidence to support the claim that the average size of the farms is different.

Example 7

Find the 95% confidence interval for the difference between the means for the data in Example 6.

$$\left(\bar{X}_1 - \bar{X}_2\right) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(191 - 199) - 2.365\sqrt{\frac{38^2}{8} + \frac{12^2}{10}} < \mu_1 - \mu_2 < (191 - 199) + 2.365\sqrt{\frac{38^2}{8} + \frac{12^2}{10}}$$

$$-41.0 < \mu_1 - \mu_2 < 25.0$$

Class Activity: Testing the Difference Between 2 Means

Assessed Home Values A real estate agent wishes to determine whether tax assessors and real estate appraisers agree on the values of homes. A random sample of the two groups appraised 10 homes. The data are shown here. Is there a significant difference in the values of the homes for each group? Let $\alpha = 0.05$. Find the 95% confidence interval for the difference of the means.

Real estate appraisers	Tax assessors
$\overline{X}_1 = \$83,256$	$\overline{X}_2 = \$88,354$
$s_1 = \$3256$	$s_2 = \$2341$
$n_1 = 10$	$n_2 = 10$