Chapter 08

Hypothesis Testing: Two-Sample Inference

Introduction

In a two-sample hypothesis-testing problem, the underlying parameters of two different populations, *neither of whose values is assumed known,* are compared.

For example, if we want to study the relationship between oral contraceptive (OC) use and blood pressure in women, two different experimental designs may be used to assess this relationship.

Longitudinal Study Design

1. Identify a group of nonpregnant, premenopausal women of childbearing age (16-49 years) who are not currently OC users, and measure their blood pressure, which will be called the *baseline blood pressure*.
2. Rescreen these women 1 year later to ascertain a subgroup who have remained nonpregnant throughout the year and have become OC users. This is the study population.
3. Measure the blood pressure of the study population at the follow-up visit.
4. Compare the baseline and follow-up blood pressure of the women in the study population to determine the difference between blood pressure levels of women when they were using the pill at follow-up and when they *were not* using the pill at baseline.

Thus, in a **longitudinal or follow-up** study, the same group of people is followed *over time*. Also, this represents a *paired-sample* design because each woman is used as her own control.

Cross-Sectional Study

1. Identify both a group of OC users and a group of non-OC users among nonpregnant, premenopausal women of childbearing age (16-49 years), and measure their blood pressure.
2. Compare the blood pressure level between the OC users and nonusers.

In a **cross-sectional** study, the participants are seen at only one point in time. This study represents an *independent-sample* design because two completely different groups of women are being compared. A cross-sectional study is also less expensive than a follow-up study.

Paired sample: when each data point in the first sample is matched and is related to a unique data point in the second sample.

Paired samples may represent two sets of measurements on the same people or on different people who are chosen on an individual basis using matching criteria, such as age and sex, to be very similar to each other.

Independent samples: when the data points in one sample are unrelated to the data points in the second sample.

For the example under discussion, paired-study design is probably more definitive because most influencing factors present at first screening will also be there at the second screening and will not influence the comparison of BP levels. However, a control group of non-OC users would completely rule out possible causes of BP change.

The second type of study can only be considered suggestive because other confounding factors may influence BP and cause an apparent difference to be found when none is actually present.

The Paired *t* Test

Table 8.1 SBP levels (mm Hg) in 10 women while not using (baseline) and while using (follow-up) OCs

i	SBP level while not using OCs (x_{i1})	SBP level while using OCs (x_{i2})	d_i^*
1	115	128	13
2	112	115	3
3	107	106	–1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	–2
10	115	117	2

$^*d_i = x_{i2} - x_{i1}$

Assume that the systolic blood pressure (SBP) of the *i*th woman is normally distributed at baseline with mean μ_i and variance σ^2 and at follow-up with mean $\mu_i+\Delta$ and variance σ^2.

Δ: mean difference in SBP between follow-up and baseline.

If $\Delta = 0$, difference is 0; if $\Delta > 0$, then OC pills associated with increased mean SBP; if $\Delta < 0$, then OC pills associated with lowered mean SBP.

$H_0: \Delta = 0$ vs. $H_1: \Delta \neq 0$; μ_1 is unknown; difference $d_i = x_{i2} - x_{i1}$

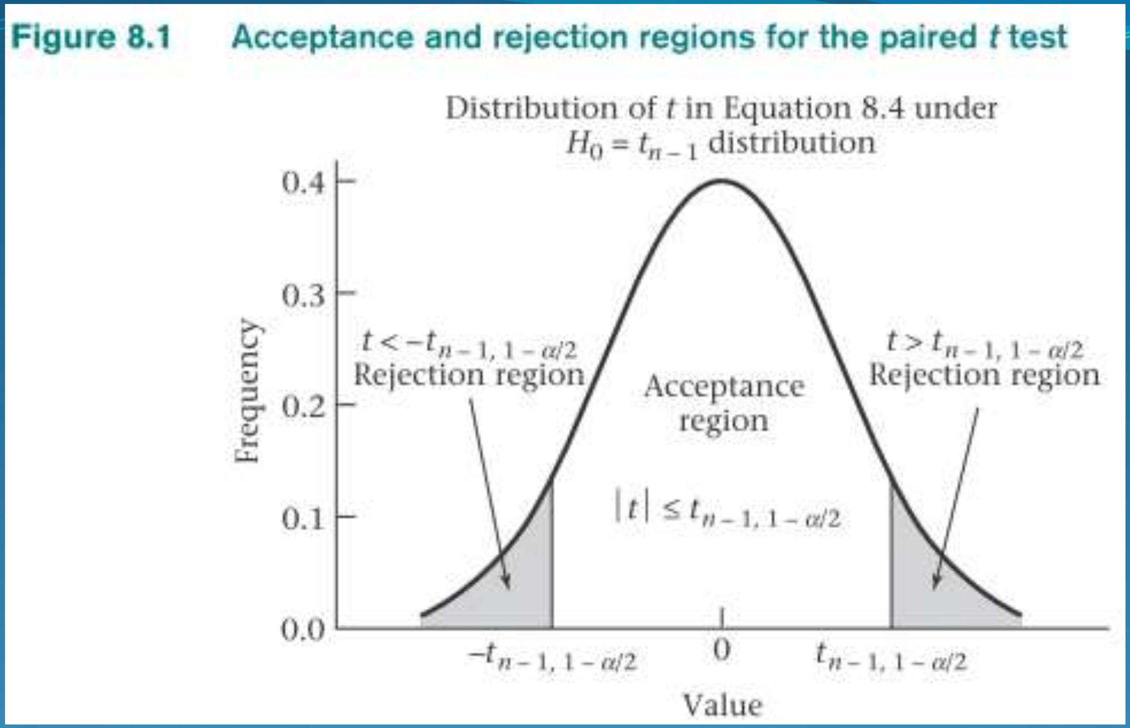
Although BP levels are different for each woman, the difference in BP between baseline and follow-up have the same mean and variance over the entire population of women. Thus, it can be considered a one-sample t test based on the differences (d_i).

Mean difference $\bar{d} = (d_1 + d_2 + \ldots + d_n)/n$

Denote the test statistic $\bar{d}/(s_d/\sqrt{n})$ by t, where s_d is the sample standard deviation of the observed differences:

$$s_d = \sqrt{\left[\sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} d_i\right)^2 \Big/ n\right] \Big/ (n-1)}$$

n = number of matched pairs



If $t > t_{n-1,1-\alpha/2}$ or $t < -t_{n-1,1-\alpha/2}$ then Ho is rejected.

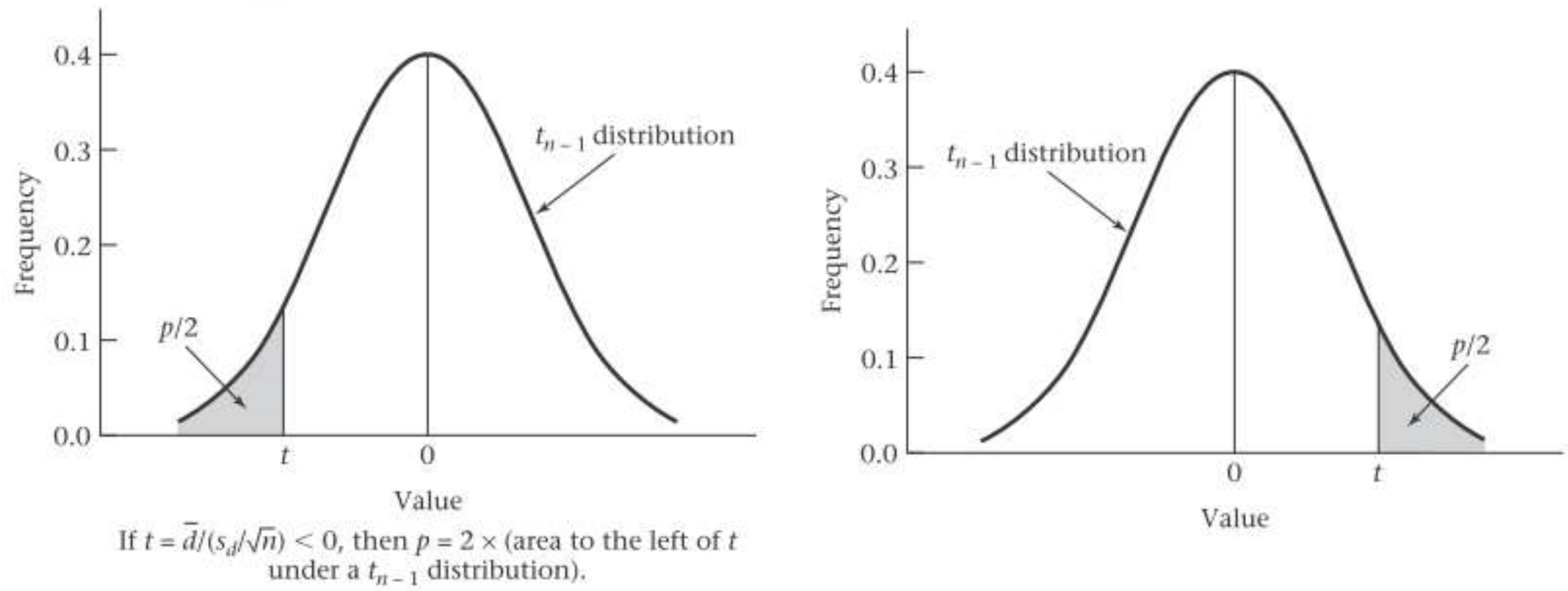
If $-t_{n-1,1-\alpha/2} \le t \le t_{n-1,1-\alpha/2}$ then Ho is accepted.

Computation of the p-Value for the Paired *t* Test

If $t < 0$, $p = 2 \times$ [the area to the left of $t = \overline{d}/(s_d/\sqrt{n})$ under a t_{n-1} distribution]

If $t \geq 0$, $p = 2 \times$ [the area to the right of t under a t_{n-1} distribution]

Figure 8.2 Computation of the *p*-value for the paired *t* test



Interval Estimation for the Comparison of Means from Two Paired Samples

Confidence Interval for the True Difference (Δ) Between the Underlying Means of Two Paired Samples (Two-Sided)

A two-sided 100% × (1-α) CI for the true mean difference (Δ) between two paired samples is given by

$$(\bar{d} - t_{n-1,1-\alpha/2}\, s_d/\sqrt{n},\ \bar{d} + t_{n-1,1-\alpha/2}\, s_d/\sqrt{n})$$

Two-Sample *t* Test for Independent Samples with Equal Variances

If the difference between the two sample means

$\bar{x}_1 - \bar{x}_2$ is far from 0, then H_0 will be rejected; otherwise, it will be accepted.

Because the two samples are independent, $\bar{X}_1 - \bar{X}_2$ is normally distributed with mean $\mu_1 - \mu_2$ and variance $\sigma^2(1/n_1 + 1/n_2)$. In symbols,

$$\bar{X}_1 - \bar{X}_2 \sim N\left[\mu_1 - \mu_2, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]$$

Under H_0, $\mu_1 = \mu_2$

$$\bar{X}_1 - \bar{X}_2 \sim N\left[0, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]$$

If σ^2 were known, then $\overline{X}1 - \overline{X}2$ could be divided by $\sigma\sqrt{(1/n_1 + 1/n_2)}$, to get

$$\frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$$

However, σ^2 in general is unknown and must be estimated from the data using sample variances s_1^2 and s_2^2. The pooled estimate of the variance from two independent samples is given by

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Suppose we wish to test the hypothesis H_0: $\mu_1 = \mu_2$ vs. H_1: $\mu_1 \neq \mu_2$ with a significance level of α for two normally distributed populations, where σ^2 is assumed to be the same for each population. Compute the test statistic

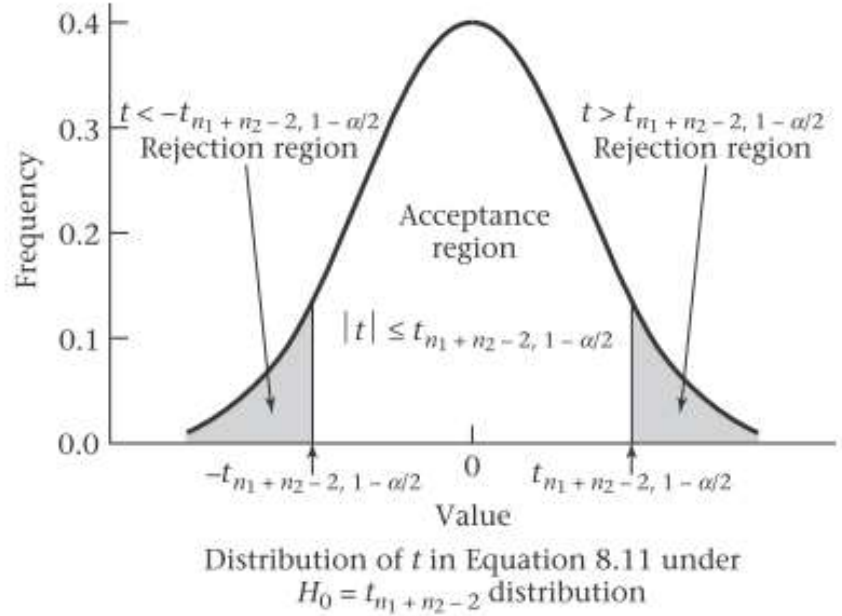
$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right]/(n_1 + n_2 - 2)}$

If $t > t_{n1+n2-2,1-\alpha/2}$ or $t < -t_{n1+n2-2,1-\alpha/2}$ then Ho is rejected.

If $-t_{n1+n2-2,1-\alpha/2} \leq t \leq t_{n1+n2-2,1-\alpha/2}$ then Ho is accepted.

Figure 8.3 Acceptance and rejection regions for the two-sample *t* test for independent samples with equal variances



Computation of the p-value for the Two-Sample t Test for Independent Samples with Equal Variances

Compute the test statistic

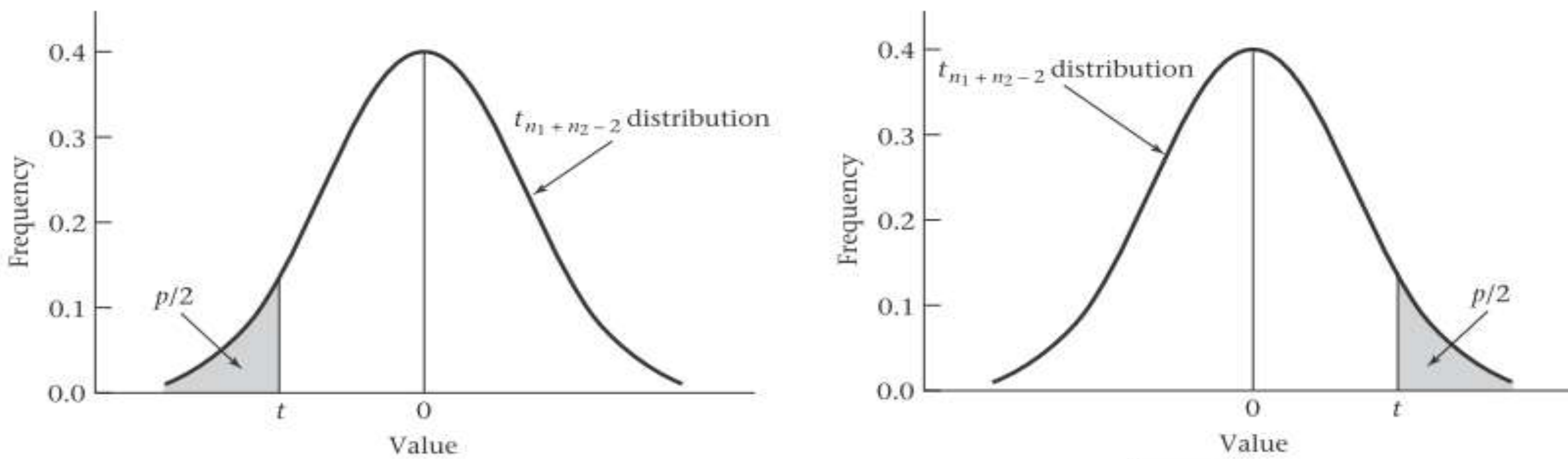
$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right]/(n_1 + n_2 - 2)}$

If $t \le 0$, $p = 2 \times$ (area to the left of t under a $t_{n1+n2-2}$ distribution)

If $t > 0$, $p = 2 \times$ (area to the right of t under a $t_{n1+n2-2}$ distribution)

Computation of the p-value for the two-sample t test for independent samples with equal variances



If $t = (\bar{x}_1 - \bar{x}_2)\Big/\left(s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) \le 0$, then $p = 2 \times$ (area to the left of t under a $t_{n_1+n_2-2}$ distribution).

If $t = (\bar{x}_1 - \bar{x}_2)\Big/\left(s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) > 0$, then $p = 2 \times$ (area to the right of t under a $t_{n_1+n_2-2}$ distribution).

Interval Estimation for the Comparison of Means from Two Independent Samples (Equal Variance Case)

A two-sided 100% × (1-α) CI for the true mean difference μ_1 - μ_2 based on two independent samples is given by

$$\left(\bar{x}_1 - \bar{x}_2 - t_{n_1+n_2-2,1-\alpha/2}s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}_1 - \bar{x}_2 + t_{n_1+n_2-2,1-\alpha/2}s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

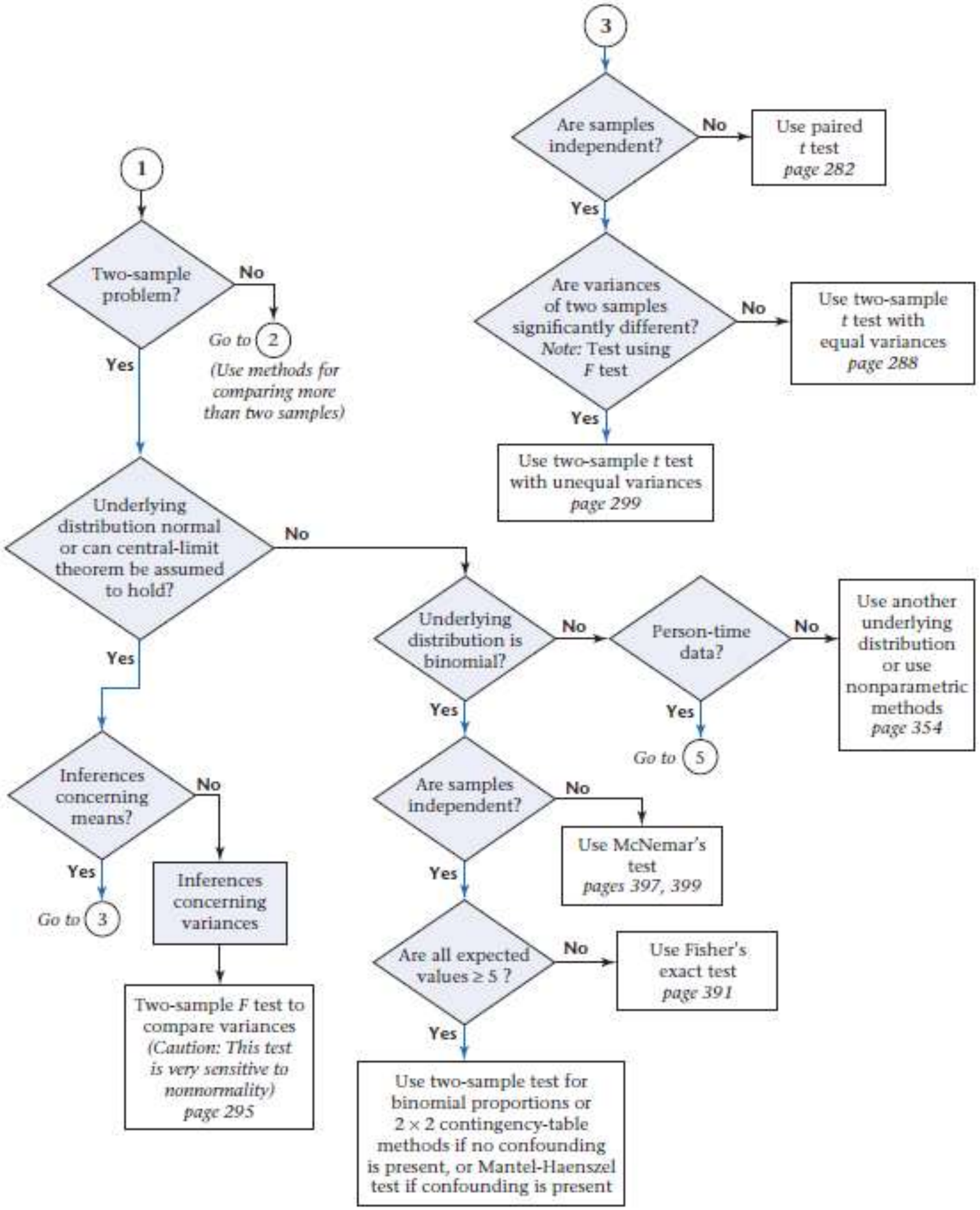
If the interval is wide then it indicates that a much larger sample is needed to accurately assess the true mean difference.

Summary

In this chapter, we discussed

- Methods of hypothesis testing for comparing the means and variances of two samples that are assumed to be normally distributed.
- Paired t test and F test: In a two-sample problem, if the two samples are paired then the paired t test is appropriate. If the samples are independent, then the F test for the equality of two variances is used to decide whether the variances are significantly different.
- If the variances are not significantly different, then the two-sample t test with equal variances is used.

FIGURE 8.13 Flowchart summarizing two-sample statistical inference—normal-theory methods



The End