Introductory STATISTICS

9TH EDITION

Neil WEISS

Chapter 13

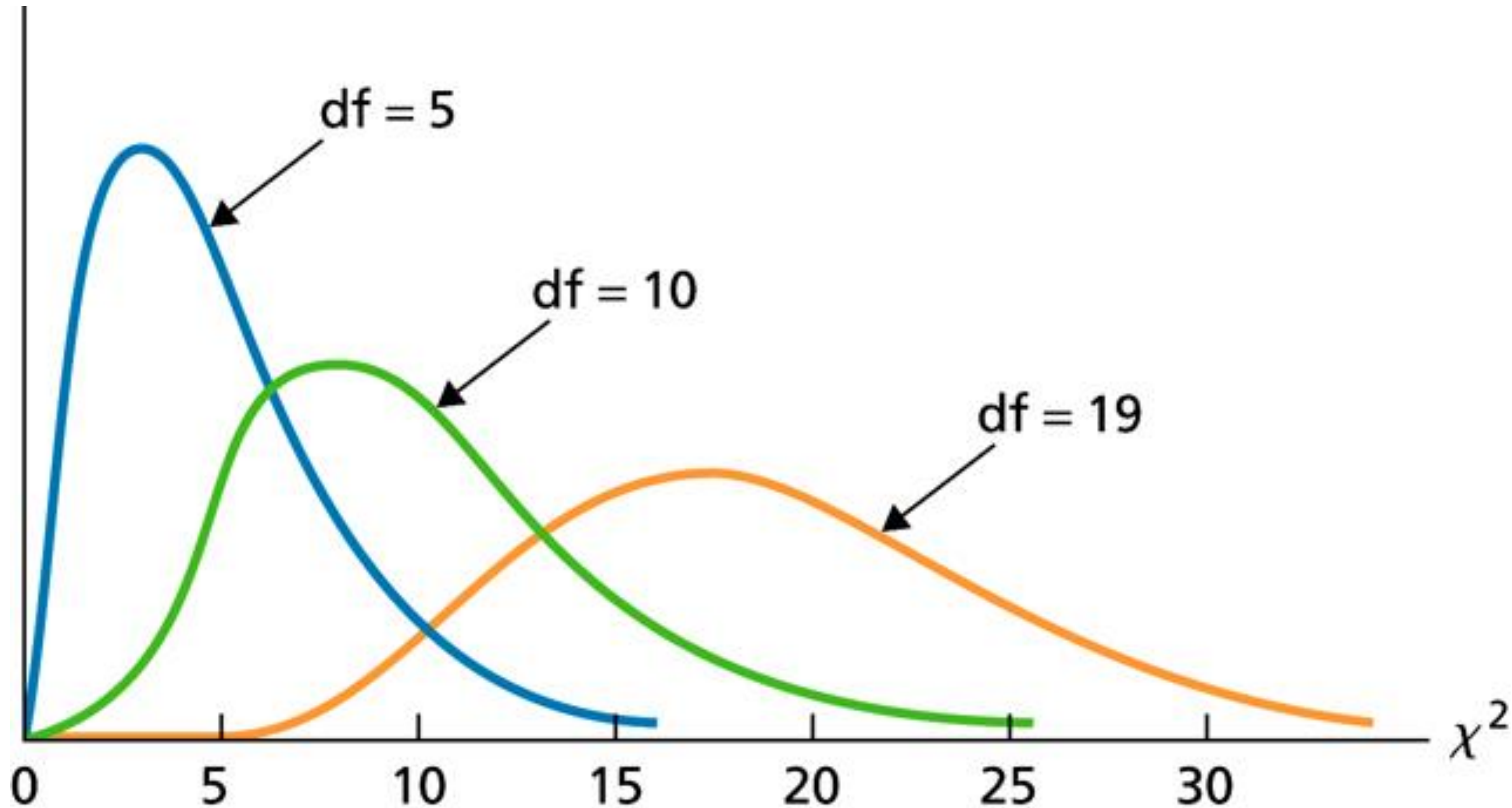
Chi-Square Procedures

Section 13.1

The Chi-Square Distribution

Figure 13.1

χ^2-curves for df = 5, 10, and 19



Key Fact 13.1

Basic Properties of χ^2-Curves

Property 1: The total area under a χ^2-curve equals 1.

Property 2: A χ^2-curve starts at 0 on the horizontal axis and extends indefinitely to the right, approaching, but never touching, the horizontal axis.

Property 3: A χ^2-curve is right skewed.

Property 4: As the number of degrees of freedom becomes larger, χ^2-curves look increasingly like normal curves.

Section 13.2

Chi-Square Goodness-of-Fit Test

Table 13.3

Expected frequencies if last year's violent-crime distribution is the same as the 2000 distribution

Type of violent crime	Expected frequency
Murder	5.5
Forcible rape	31.5
Robbery	143.0
Agg. assault	320.0

Table 13.4

Calculating the goodness of fit

Type of violent crime x	Observed frequency O	Expected frequency E	Difference $O - E$	Square of difference $(O - E)^2$	Chi-square subtotal $(O - E)^2/E$
Murder	3	5.5	−2.5	6.25	1.136
Forcible rape	37	31.5	5.5	30.25	0.960
Robbery	154	143.0	11.0	121.00	0.846
Agg. assault	306	320.0	−14.0	196.00	0.613
	500	500.0	0		3.555

Procedure 13.1

Chi-Square Goodness-of-Fit Test

Purpose To perform a hypothesis test for the distribution of a variable

Assumptions

1. All expected frequencies are 1 or greater
2. At most 20% of the expected frequencies are less than 5
3. Simple random sample

Step 1 **The null and alternative hypotheses are, respectively,**

H_0: The variable has the specified distribution

H_a: The variable does not have the specified distribution.

Step 2 **Decide on the significance level, α.**

Step 3 **Compute the value of the test statistic**

$$\chi^2 = \Sigma(O - E)^2/E,$$

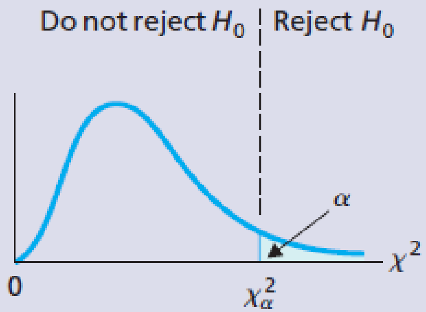
where O and E represent observed and expected frequencies, respectively. Denote the value of the test statistic χ_0^2.

Procedure 13.1 (cont.)

CRITICAL-VALUE APPROACH OR **P-VALUE APPROACH**

Critical-Value Approach

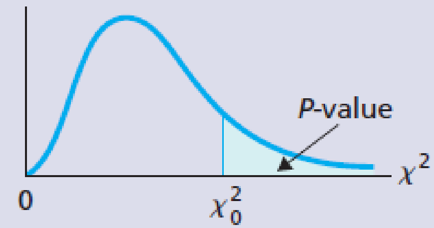
Step 4 The critical value is χ^2_α with $\text{df} = c - 1$, where c is the number of possible values for the variable. Use Table VII to find the critical value.



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

P-Value Approach

Step 4 The χ^2-statistic has $\text{df} = c - 1$, where c is the number of possible values for the variable. Use Table VII to estimate the P-value, or obtain it exactly by using technology.



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

Section 13.3

Contingency Tables; Association

Table 13.7

Political party affiliation and class level for students in introductory statistics

Student	Political party	Class level	Student	Political party	Class level
1	Democratic	Freshman	21	Democratic	Junior
2	Other	Junior	22	Democratic	Senior
3	Democratic	Senior	23	Republican	Freshman
4	Other	Sophomore	24	Democratic	Sophomore
5	Democratic	Sophomore	25	Democratic	Senior
6	Republican	Sophomore	26	Republican	Sophomore
7	Republican	Junior	27	Republican	Junior
8	Other	Freshman	28	Other	Junior
9	Other	Sophomore	29	Other	Junior
10	Republican	Sophomore	30	Democratic	Sophomore
11	Republican	Sophomore	31	Republican	Sophomore
12	Republican	Junior	32	Democratic	Junior
13	Republican	Sophomore	33	Republican	Junior
14	Democratic	Junior	34	Other	Senior
15	Republican	Sophomore	35	Other	Sophomore
16	Republican	Senior	36	Republican	Freshman
17	Democratic	Sophomore	37	Republican	Freshman
18	Democratic	Junior	38	Republican	Freshman
19	Other	Senior	39	Democratic	Junior
20	Republican	Sophomore	40	Republican	Senior

Table 13.8

Preliminary contingency table for political party affiliation and class level

		Class level				
		Freshman	Sophomore	Junior	Senior	**Total**
Party	Democratic	𝍷	𝍷𝍷𝍷𝍷	𝍸	𝍷𝍷𝍷	
	Republican	𝍷𝍷𝍷𝍷	𝍸 𝍷𝍷𝍷	𝍷𝍷𝍷𝍷	𝍷𝍷	
	Other	𝍷	𝍷𝍷𝍷	𝍷𝍷𝍷	𝍷𝍷	
	Total					

Table 13.9

Contingency table for political party affiliation and class level

		Class level				
		Freshman	Sophomore	Junior	Senior	**Total**
Party	Democratic	1	4	5	3	13
	Republican	4	8	4	2	18
	Other	1	3	3	2	9
	Total	6	15	12	7	40

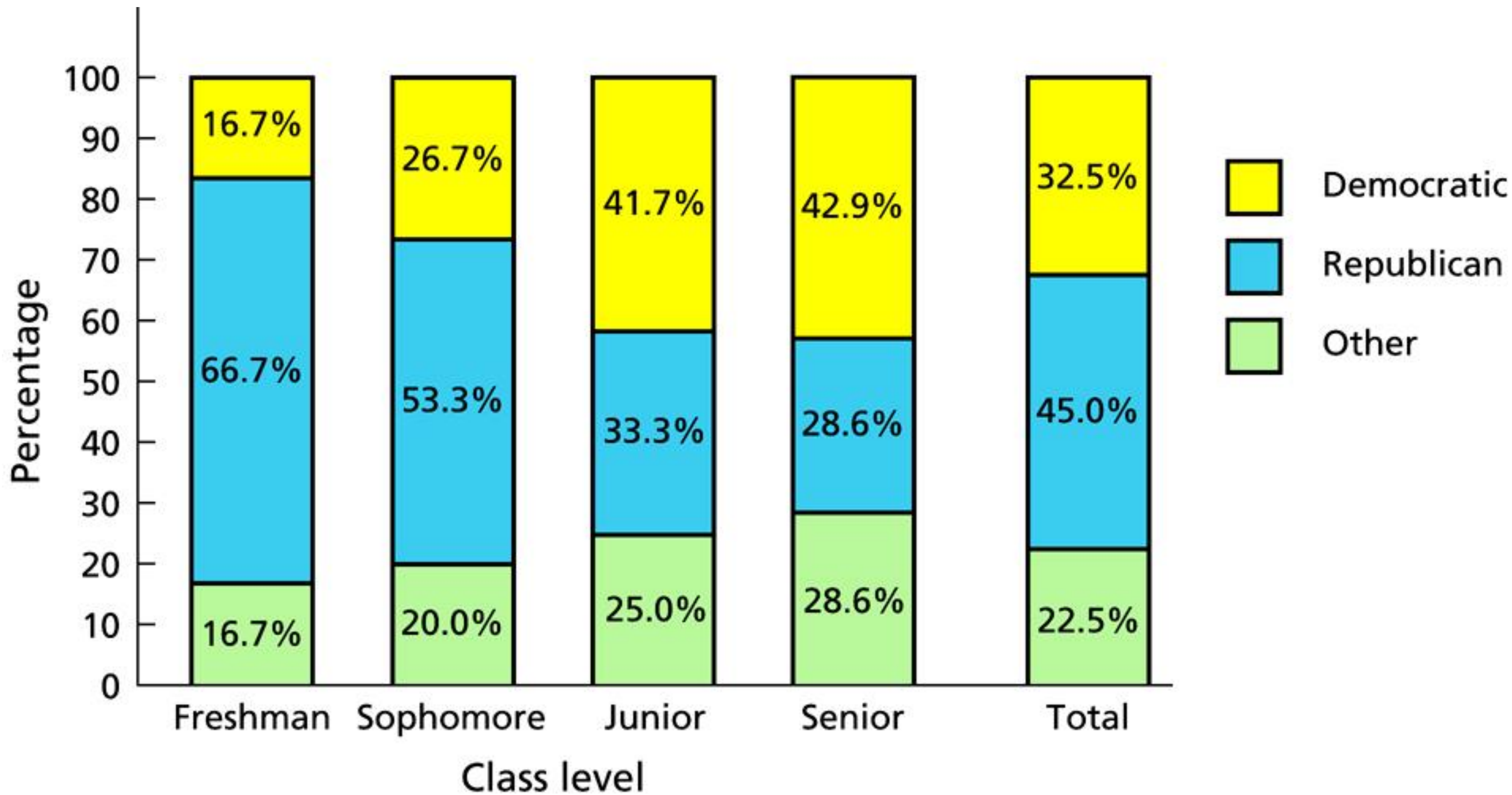
Table 13.10

Conditional distributions of political party affiliation by class level

Party		Class level				
		Freshman	Sophomore	Junior	Senior	**Total**
Party	Democratic	0.167	0.267	0.417	0.429	0.325
	Republican	0.667	0.533	0.333	0.286	0.450
	Other	0.167	0.200	0.250	0.286	0.225
	Total	1.000	1.000	1.000	1.000	1.000

Figure 13.4

Segmented bar graph for the conditional distributions and marginal distribution of political party affiliation



Section 13.4

Chi-Square Independence Test

Table 13.13

Contingency table of marital status and alcohol consumption for 1772 randomly selected U.S. adults

Marital status	Drinks per month: Abstain	1–60	Over 60	**Total**
Single	67	213	74	354
Married	411	633	129	1173
Widowed	85	51	7	143
Divorced	27	60	15	102
Total	590	957	225	1772

Table 13.14

Observed and expected frequencies for marital status and alcohol consumption (expected frequencies printed below observed frequencies)

		Drinks per month			
		Abstain	1–60	Over 60	**Total**
Marital status	Single	67 117.9	213 191.2	74 44.9	354
	Married	411 390.6	633 633.5	129 148.9	1173
	Widowed	85 47.6	51 77.2	7 18.2	143
	Divorced	27 34.0	60 55.1	15 13.0	102
	Total	590	957	225	1772

Procedure 13.2

Chi-Square Independence Test

Purpose To perform a hypothesis test to decide whether two variables are associated

Assumptions

1. All expected frequencies are 1 or greater
2. At most 20% of the expected frequencies are less than 5
3. Simple random sample

Step 1 **The null and alternative hypotheses are, respectively,**

$$H_0\text{: The two variables are not associated.}$$
$$H_a\text{: The two variables are associated.}$$

Step 2 **Decide on the significance level, α.**

Step 3 **Compute the value of the test statistic**

$$\chi^2 = \Sigma(O - E)^2/E,$$

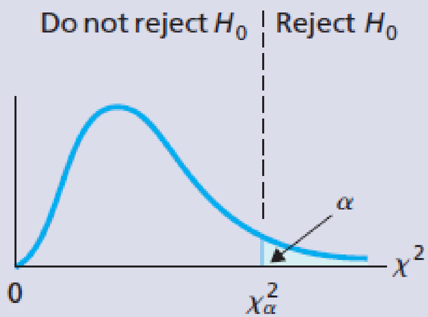
where O and E represent observed and expected frequencies, respectively. Denote the value of the test statistic χ_0^2.

Procedure 13.2 (cont.)

CRITICAL-VALUE APPROACH OR **P-VALUE APPROACH**

CRITICAL-VALUE APPROACH

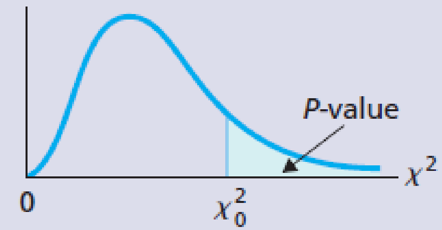
Step 4 The critical value is χ^2_α with df $= (r - 1) \times (c - 1)$, where r and c are the number of possible values for the two variables. Use Table VII to find the critical value.



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

P-VALUE APPROACH

Step 4 The χ^2-statistic has df $= (r - 1)(c - 1)$, where r and c are the number of possible values for the two variables. Use Table VII to estimate the P-value, or obtain it exactly by using technology.



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.