

## EXERCISES

- 12.3.1** The following table shows the distribution of uric acid determinations taken on 250 patients. Test the goodness-of-fit of these data to a normal distribution with  $\mu = 5.74$  and  $\sigma = 2.01$ . Let  $\alpha = .01$ .

Uric Acid Determination	Observed Frequency	Uric Acid Determination	Observed Frequency
< 1	1	6 to 6.99	45
1 to 1.99	5	7 to 7.99	30
2 to 2.99	15	8 to 8.99	22
3 to 3.99	24	9 to 9.99	10
4 to 4.99	43	10 or higher	5
5 to 5.99	50		
Total			250

- 12.3.2** The following data were collected on 300 eight-year-old girls. Test, at the .05 level of significance, the null hypothesis that the data are drawn from a normally distributed population. The sample mean and standard deviation computed from grouped data are 127.02 and 5.08.

Height in Centimeters	Observed Frequency	Height in Centimeters	Observed Frequency
114 to 115.9	5	128 to 129.9	43
116 to 117.9	10	130 to 131.9	42
118 to 119.9	14	132 to 133.9	30
120 to 121.9	21	134 to 135.9	11
122 to 123.9	30	136 to 137.9	5
124 to 125.9	40	138 to 139.9	4
126 to 127.9	45		
Total			300

- 12.3.3** The face sheet of patients' records maintained in a local health department contains 10 entries. A sample of 100 records revealed the following distribution of erroneous entries:

Number of Erroneous Entries Out of 10	Number of Records
0	8
1	25
2	32
3	24
4	10
5 or more	1
Total	100

Test the goodness-of-fit of these data to the binomial distribution with  $p = .20$ . Find the  $p$  value for this test.

- 12.3.4** In a study conducted by Byers et al. (A-2), researchers tested a Poisson model for the distribution of activities of daily living (ADL) scores after a 7-month prehabilitation program designed to prevent functional decline among physically frail, community-living older persons. ADL measured the ability of individuals to perform essential tasks, including walking inside the house, bathing, upper and lower body dressing, transferring from a chair, toileting, feeding, and grooming. The scoring method used in this study assigned a value of 0 for no (personal) help and no difficulty, 1 for difficulty but no help, and 2 for help regardless of difficulty. Scores were summed to produce an overall score ranging from 0 to 16 (for eight tasks). There were 181 subjects who completed the study. Suppose we use the authors' scoring method to assess the status of another group of 181 subjects relative to their activities of daily living. Let us assume that the following results were obtained.

$X$	Observed Frequency $X$	Expected Frequency	$X$	Observed Frequency $X$	Expected Frequency
0	74	11.01	7	4	2.95
1	27	30.82	8	3	1.03
2	14	43.15	9	2	0.32
3	14	40.27	10	3	0.09
4	11	28.19	11	4	0.02
5	7	15.79	12 or more	13	0.01
6	5	7.37			

Source: Hypothetical data based on procedure reported by Amy L. Byers, Heather Allore, Thomas M. Gill, and Peter N. Peduzzi, "Application of Negative Binomial Modeling for Discrete Outcomes: A Case Study in Aging Research," *Journal of Clinical Epidemiology*, 56 (2003), 559–564.

Test the null hypothesis that these data were drawn from a Poisson distribution with  $\lambda = 2.8$ . Let  $\alpha = .01$ .

- 12.3.5** The following are the numbers of a particular organism found in 100 samples of water from a pond:

Number of Organisms per Sample		Number of Organisms per Sample	
	Frequency		Frequency
0	15	4	5
1	30	5	4
2	25	6	1
3	20	7	0
Total			100

Test the null hypothesis that these data were drawn from a Poisson distribution. Determine the  $p$  value for this test.

Score	Number of Applicants	Score	Number of Applicants
30–34	19	60–64	8
35–39	25	65–69	4
Total			175

Do these data provide sufficient evidence to indicate that the population of scores is not normally distributed? Let  $\alpha = .05$ . What is the  $p$  value for this test?

20. A local health department sponsored a venereal disease (VD) information program that was open to high-school juniors and seniors who ranged in age from 16 to 19 years. The program director believed that each age level was equally interested in knowing more about VD. Since each age level was about equally represented in the area served, she felt that equal interest in VD would be reflected by equal age-level attendance at the program. The age breakdown of those attending was as follows:

Age	Number Attending
16	26
17	50
18	44
19	40

Are these data incompatible with the program director's belief that students in the four age levels are equally interested in VD? Let  $\alpha = .05$ . What is the  $p$  value for this test?

21. A survey of children under 15 years of age residing in the inner-city area of a large city were classified according to ethnic group and hemoglobin level. The results were as follows:

Ethnic Group	Hemoglobin Level (g/100 ml)			Total
	10.0 or Greater	9.0–9.9	< 9.0	
A	80	100	20	200
B	99	190	96	385
C	70	30	10	110
Total	249	320	126	695

Do these data provide sufficient evidence to indicate, at the .05 level of significance, that the two variables are related? What is the  $p$  value for this test?

22. A sample of reported cases of mumps in preschool children showed the following distribution by age:

Age (Years)	Number of Cases
Under 1	6
1	20
2	35
3	41
4	48
Total	150