TABLE E7.8 Computations for Chi-Square Tests of Three Distributions

Interval	Observed	Theoretical Frequencies			$\Sigma(n_i-e_i)^2/e_i$		
Res. Str./Yield Str.	Frequency	Norm	Lognorm	Shifted Gamma	Norm	Lognorm	Shifted Gamma
< 0.111	15	32.6	6.5	23.2	9.48	10.90	2.88
0.111–0.222	72	45.4	73.2	61.0	15.60	0.02	1.97
0.222–0.333	88	67.0	95.9	78.1	6.57	0.66	1.25
0.333–0.444	54	71.6	66.0	66.5	4.33	2.17	2.35
0.444–0.555	38	55.4	37.1	44.2	5.44	0.02	0.87
0.555–0.666	31	31.0	19.5	24.9	0.00	6.75	1.51
0.666–0.777	16	12.5	10.1	12.4	0.96	3.40	1.03
0.777–0.888	3	3.7	5.3	5.7	0.12	0.99	1.27
> 0.888	3	0.9	6.3	4.0	4.81	1.75	0.25
$\sum$	320	320	320	320	47.3	26.7	13.4

From Table E7.8, we see that among the three distributions, the shifted gamma distribution gives the lowest value for $\sum(n_i - e_i)^2/e_i$. Also, at the significance level of 1% and a d.o.f. of $f = 9 - 4 = 5$, we obtain from Table A.4 the critical value of $c_{0.99,5} = 15.09$ for the normal and lognormal distributions, whereas for the shifted gamma distribution $f = 9-5 = 4$ and $c_{0.99,4} = 13.28$. Therefore, according to the chi-square test, only the shifted gamma distribution (among the three distributions) is approximately valid at the 1% significance level for modeling the probability distribution of residual stresses in wide-flange beams. ◄

It may be emphasized that because there is some arbitrariness in the choice of the significance level α, the chi-square goodness-of-fit test (as well as the Kolmogorov–Smirnov and the Anderson–Darling methods, described subsequently in Sects. 7.3.2 and 7.3.3) may not provide absolute information on the validity of a specific distribution. For example, it is conceivable that a distribution acceptable at one significance level may be unacceptable at another significance level; this can be illustrated with the shifted gamma distribution of Example 7.8, in which the distribution is valid at the 1% significance level but will not be valid at the 5% level.

In spite of this arbitrariness in the selection of the significance level, however, such statistical goodness-of-fit tests remain useful, especially for determining the relative goodness-of-fit of two or more theoretical distribution models, as illustrated in Examples 7.7 and 7.8. Moreover, these tests should be used only to help verify the validity of a theoretical model that has been selected on the basis of other prior considerations, such as through the application of appropriate probability papers, or even visual inspection of an appropriate PDF with the available histogram.

7.3.2 The Kolmogorov–Smirnov (K–S) Test for Goodness-of-Fit

Another widely used goodness-of-fit test is the *Kolmogorov–Smirnov* or K–S test. The basic premise of this test is to compare the experimental cumulative frequency with the CDF of an assumed theoretical distribution. If the maximum discrepancy between the experimental and theoretical frequencies is larger than normally expected for a given sample size, the theoretical distribution is not acceptable for modeling the underlying population; conversely, if the discrepancy is less than a critical value, the theoretical distribution is acceptable at the prescribed significance level α.