Lesson 9.1.2

9-10. See answers in bold in table below.

1st digit	1	2	3	4	5	6	7	8	9
count	**42**	**21**	**17**	**11**	**6**	**4**	**6**	**8**	**5**

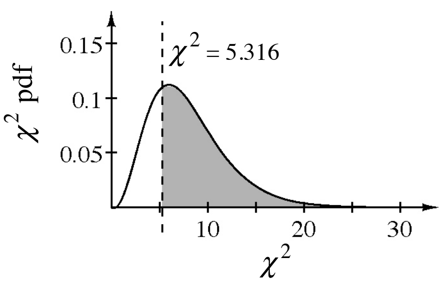
9-11. Expected counts: 36.120, 21.132, 14.998, 11.628, 9.504, 8.028, 6.960, 6.144, 5.496

9-12. Null hypothesis: The first digits of the tree height measurements are distributed according to Benford's law. Alternative hypothesis: The first digits of the tree height measurements are not distributed according to Benford's law. At least one digit is out of proportion.

9-13. Random sampling to avoid bias: The tree heights were stated to be randomly gathered. Independent trials and large counts, so the sampling distribution is ≈ chi-squared: It can be assumed that each observation was independent. The sample is large enough to expect at least 5 "yes" in each category.

9-14. Contributions: 0.957, 0.001, 0.271, 0.034, 1.291, 2.021, 0.132, 0.561, 0.045. $\chi^2 = 5.313$

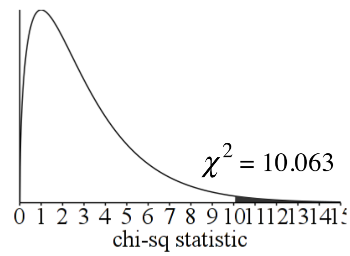
9-15.



p-value = 0.7237

9-16. Because the p-value of 0.7237 is so high (much higher than a standard 5% significance level), there is not sufficient evidence to reject the null hypothesis. We cannot conclude that the first digits of the white oak tree heights differ from Benford's law.

9-17. <u>Identify</u>: Chi-squared goodness of fit test where H_0 = giraffe weight is normally distributed and H_A: giraffe weight is not normally distributed. Sample evidence for H_A: observed and expected frequencies do not match.
<u>Check conditions</u>: Random, to avoid bias—random selection was given in the description. Independent trials and large Counts, so the sampling distribution is ≈ chi-squared: assumed that there are more than $10 \cdot 116$ in the population. Expected frequencies are all greater than 5.
<u>Calculate</u>: df = 4 − 1 = 3,
$\chi^2 = \frac{(29-31.3)^2}{31.3} + \ldots + \frac{(29-16.7)^2}{16.7} \approx 11.58$.
p-value = x^2cdf(11.58, E99, 3) ≈ 0.009. See diagram at right.
<u>Conclude</u>: The p-value is less than 0.01 so reject the null hypothesis and conclude at 1% significance that giraffe weights are not normally distributed.



9-18. a. H_0: The proportions of candy are as claimed by the company. H_A: At least one proportion differs from the proportion reported by the company.

b. The observed and expected counts do not match.

c. It was stated that the candy was randomly gathered. The 165 pieces of candy in the sample are assumed to be "small" compared to the population of all candy. All expected counts are higher than 5.

d. $\chi^2 = 4.461$, df = 4, p-value = 0.3472

e. Because the p-value (0.3427) is not less than the level of significance ($\alpha = 0.05$), students do not have sufficient evidence to reject H_0. Students cannot conclude that the proportions of candy by color differ from those reported by the company.

9-19. a. Assuming the null hypothesis is indeed true, the probability of observing a difference as large as or larger than $\frac{42}{60} - \frac{34}{60}$ or $\frac{8}{60}$ is 0.0648. Therefore since the p-value is larger than 0.05 you do not have enough evidence to reject the null hypothesis that the percent of players with high blood pressure is the same for online gamers and offline gamers.

b. Since the null hypothesis was not rejected, a Type II error could have occurred. This means the null hypothesis should have been rejected and that there really is a difference between the percentage of online gamers and offline gamers in terms of blood pressure level.

9-20. a. $r = 0.9754$ indicating a strong positive linear association between the costs billed to clients and litigation time, which is confirmed by an inspection of the scatterplot.

b. $R^2 = 0.9513$, meaning that 95% of the variation observed in client bills is explained by a linear relationship with litigation time. In other words, knowing the litigation time associated with each client time reduces the error in predicting what they are billed by 95% over using just the mean cost to client as a predictor.

c. When making predictions about client bills by using litigation times, one can expect to typically be off by 0.491 (or $491).

