Lesson 9.2.3

9-47. a. Identify: This is a chi-squared independence test because there is one sample with two different categorical variables (grade-level and laugh-status). H_0: Grade-level and laugh-status are independent. H_A: Grade-level and laugh-status are not independent. Sample evidence for H_A: observed frequencies do not match expected frequencies. $\alpha = 0.05$.

Check conditions: Random, to avoid bias. It is stated that samples was random. Independent trials, samples and large counts, so the sampling distribution is ≈ chi-squared. It is assumed that the students are independent of one another and all expected counts are higher than 5 (the smallest is 16.4).

Calculate: $\chi^2 = 9.6154$, p-value = 0.0221.

Conclude: Because the p-value is less than the level of significance, we reject H_0. The evidence seems to indicate that there is an association between grade-level and laugh-status.

b. See table below for contributions. The largest contribution is the "11^{th} grade/does not laugh" category with a contribution of 2.8824.

	9^{th} grade	10^{th} grade	11^{th} grade	12^{th} grade
Laughs	0.6974	0.5814	1.1395	0.306
Does not laugh	1.764	1.4706	2.8824	0.7741

9-48. Identify: This is a chi-squared goodness of fit test. H_0: The distribution of scores follows the binomial model based on students guessing. H_A: The distribution of scores does not follow the binomial model based on students guessing. Sample evidence for H_A: observed frequencies do not match expected frequencies. $\alpha = 0.05$.

Check conditions: Random, to avoid bias. It is stated that the sample is random. Independent trials, samples and large counts, so the sampling distribution is ≈ chi-squared. It is assumed that samples are independent of one another and all expected frequencies > 5.

Calculate: $\chi^2 = 1.634$, p-value = 0.6516.

Conclude: Because the p-value is not less than 5%, students cannot conclude that the students differed significantly from purely guessing.

9-49. Identify: This is a chi-squared test for homogeneity of proportion because students have several samples and the same categories being compared across the various samples. H_0: The proportion of people who would purchase the new product is the same in each of the three cities. H_A: At least one city differs in the proportion of people who would purchase the new product. Sample evidence for H_A: observed frequencies do not match expected frequencies. $\alpha = 0.05$.
Check conditions: Random, to avoid bias. It is stated that the samples are random. Independent trials, samples and large counts, so the sampling distribution is ≈ chi-squared. It is assumed that samples are independent of one another and all expected frequencies > 5.
Calculate: $\chi^2 =$ 5.6059, p-value = 0.0606.
Conclude: Because the p-value is not less than 5%, students cannot conclude that any of the cities differ significantly in the proportion of people who would purchase the new product.

9-50. A chi-squared goodness of fit test is used when there is a single categorical variable (a flat list) with counts that are being compared to some external set of proportions or ideal. A chi-squared independence test is used when a single sample is taken from a population, and two categorical variables are being analyzed for independence. A chi-squared test for homogeneity of proportions is used when independent samples are taken from two or more populations, and the same categorical variable is being compared across the multiple populations.

9-51. Identify: Chi-squared test for independence. H_0: The customers' purchasing choices are independent of the employees' customer prompts. H_A: The customers' purchasing choices are associated with the employees' customer prompts. Sample evidence for H_A: observed and expected frequencies do not match.
Check conditions: Random, to avoid bias. It is stated that the sample gathered was a random sample. Independent trials and large counts, so the sampling distribution is ≈ chi-squared. Assumed that each customer choice was independent. All expected counts are higher than 5.
Calculate: $\chi^2 = 13.162$, df = 2, p-value = 0.0014.
Conclude: Because the p-value is less than the level of significance ($\alpha = 0.05$), students reject H_0. Students conclude that the customers' purchasing choices are associated with the employees prompting the customers at 5% significance.

9-52. a. $b = r\left(\frac{S_y}{S_x}\right) = 1.625$. Using the means of the explanatory and response variables as a point on the LSRL, 13.87 = a + 1.625(26.77). Solving for a = –29.633 dollars/lb, making the LSRL equation: $\hat{y}$ = –29.633 + 1.625x, where x is baking time and $\hat{y}$ is predicted cost to make.

b. For every one minute increase in baking time the predicted value of cost to make will increase by 1.625 dollars/lb.

c. $\hat{y}$ = –29.633 + 1.625(21.83) = 5.84 dollars/lb

d. The predicted value at x = 28.1 minutes is $\hat{y}$ = –29.633 + 1.625(28.1) = 16.03 dollars/lb. The residual is $y_{\text{observed}} - y_{\text{predicted}}$. residual = 19.14 – 16.03 = 3.11 dollars/lb

e. The scatterplot shows that a curved model is more appropriate.

f. No. This would be an extrapolation of the data. In fact the predicted cost is negative!

9-53. a. $P(X \geq 70) = 1 - \text{binomcdf}(100, 0.62, 69) = 0.0595$

b. $P(X < 50) = \text{binomcdf}(100, 0.62, 49) = 0.00551$

c. $P(5 \text{ before miss}) = (0.62^5)(0.38) = 0.0348$

