P.6 A venture capital company feels that the rate of return (X) on a proposed investment is approximately normally distributed with a mean of 40% and a standard deviation of 10%.

a. Find the probability that the return X will exceed 55%.

b. The banking firm who will fund the venture sees the rate of return differently, claiming that venture capitalists are always too optimistic. They perceive that the distribution of returns is $V = 0.8X - 5\%$, where X is the rate of return expected by the venture capital company. If this is correct, find the probability that the return V will exceed 55%.

P.7 At supermarkets sales of "Chicken of the Sea" canned tuna vary from week to week. Marketing researchers have determined that there is a relationship between sales of canned tuna and the price of canned tuna. Specifically, $SALES = 50000 - 100\ PRICE$. $SALES$ is measured as the number of cans per week and $PRICE$ is measured in cents per can. Suppose $PRICE$ over the year can be considered (approximately) a normal random variable with mean $\mu = 248$ cents and standard deviation $\sigma = 10$ cents.

a. Find the expected value of $SALES$.

b. Find the variance of $SALES$.

c. Find the probability that more than 24,000 cans are sold in a week. Draw a sketch illustrating the calculation.

d. Find the $PRICE$ such that $SALES$ is at its 95th percentile value. That is, let $SALES_{0.95}$ be the 95th percentile of $SALES$. Find the value $PRICE_{0.95}$ such that $P\,(SALES > SALES_{0.95}) = 0.05$.

P.8 The Shoulder and Knee Clinic knows that their expected monthly revenue from patients depends on their level of advertising. They hire an econometric consultant who reports that their expected monthly revenue, measured in \$1000 units, is given by the following equation $E(REVENUE|ADVERT) = 100 + 20\,ADVERT$, where $ADVERT$ is advertising expenditure in \$1000 units. The econometric consultant also claims that $REVENUE$ is normally distributed with variance $\text{var}(REVENUE|ADVERT) = 900$.

a. Draw a sketch of the relationship between expected $REVENUE$ and $ADVERT$ as $ADVERT$ varies from 0 to 5.

b. Compute the probability that $REVENUE$ is greater than 110 if $ADVERT = 2$. Draw a sketch to illustrate your calculation.

c. Compute the probability that $REVENUE$ is greater than 110 if $ADVERT = 3$.

d. Find the 2.5 and 97.5 percentiles of the distribution of $REVENUE$ when $ADVERT = 2$. What is the probability that $REVENUE$ will fall in this range if $ADVERT = 2$?

e. Compute the level of $ADVERT$ required to ensure that the probability of $REVENUE$ being larger than 110 is 0.95.

P.9 Consider the U.S. population of registered voters, who may be Democrats, Republicans or independents. When surveyed about the war with ISIS, they were asked if they strongly supported war efforts, strongly opposed the war, or were neutral. Suppose that the proportion of voters in each category is given in Table P.8:

TABLE P.8 Table for Exercise P.9

		War Attitude		
		Against	Neutral	In Favor
Political Party	Republican	0.05	0.15	0.25
	Independent	0.05	0.05	0.05
	Democrat	0.35	0.05	0

a. Find the marginal probability distributions for war attitudes and political party affiliation.

b. What is the probability that a randomly selected person is a political independent given that they are in favor of the war?

c. Are the attitudes about war with ISIS and political party affiliation stati

d. For the attitudes about the war assign the numerical values $AGAINST = 1$, $NEUTRAL = 2$, and $IN\ FAVOR = 3$. Call this variable WAR. Find the expected value and variance of WAR.

e. The Republican party has determined that monthly fundraising depends on the value of WAR from month to month. In particular the monthly contributions to the party are given by the relation (in millions of dollars) $CONTRIBUTIONS = 10 + 2 \times WAR$. Find the mean and standard deviation of $CONTRIBUTIONS$ using the rules of expectations and variance.

P.10 A firm wants to bid on a contract worth $80,000. If it spends $5000 on the proposal it has a 50–50 chance of getting the contract. If it spends $10,000 on the proposal it has a 60% chance of winning the contract. Let X denote the net revenue from the contract when the $5000 proposal is used and let Y denote the net revenue from the contract when the $10,000 proposal is used.

X	$f(x)$
−5,000	0.5
75,000	0.5

y	$f(y)$
−10,000	0.4
70,000	0.6

a. If the firm bases its choice solely on expected value, how much should it spend on the proposal?

b. Compute the variance of X. [*Hint*: Using scientific notation simplifies calculations.]

c. Compute the variance of Y.

d. How might the variance of the net revenue affect which proposal the firm chooses?

P.11 Prior to presidential elections citizens of voting age are surveyed. In the population, two characteristics of voters are their registered party affiliation (republican, democrat, or independent) and for whom they voted in the previous presidential election (republican or democrat). Let us draw a citizen at random, defining these two variables.

$$PARTY = \begin{cases} -1 & \text{registered republican} \\ 0 & \text{independent or unregistered} \\ 1 & \text{registered democrat} \end{cases}$$

$$VOTE = \begin{cases} -1 & \text{voted republican in previous election} \\ 1 & \text{voted democratic in previous election} \end{cases}$$

a. Suppose that the probability of drawing a person who voted republication in the last election is 0.466, and the probability of drawing a person who is registered republican is 0.32, and the probability that a randomly selected person votes republican given that they are a registered republican is 0.97. Compute the joint probability $\text{Prob}[PARTY = -1, VOTE = -1]$. Show your work.

b. Are these random variables statistically independent? Explain.

P.12 Based on years of experience, an economics professor knows that on the first principles of economics exam of the semester 13% of students will receive an A, 22% will receive a B, 35% will receive a C, 20% will receive a D, and the remainder will earn an F. Assume a 4 point grading scale (A = 4, B = 3, C = 2, D = 1, and F = 0). Define the random variable $GRADE = 4, 3, 2, 1, 0$ to be the grade of a randomly chosen student.

a. What is the probability distribution $f(GRADE)$ for this random variable?

b. What is the expected value of $GRADE$? What is the variance of $GRADE$? Show your work.

c. The professor has 300 students in each class. Suppose that the grade of the ith student is $GRADE_i$ and that the probability distribution of grades $f(GRADE_i)$ is the same for all students. Define $CLASS\_AVG = \sum_{i=1}^{300} GRADE_i/300$. Find the expected value and variance of $CLASS\_AVG$.

d. The professor has estimated that the number of economics majors coming from the class is related to the grade on the first exam. He believes the relationship to be $MAJORS = 50 + 10CLASS\_AVG$. Find the expected value and variance of $MAJORS$. Show your work.

P.13 The LSU Tigers baseball team will play the Alabama baseball team in a weekend series of two games. Let $W = 0, 1$, or 2 equal the number of games LSU wins. Let the weekend's weather be designated Cold or Not Cold. Let $C = 1$ if the weather is cold and $C = 0$ if the weather is not cold. The joint probability function of these two random variables is given in Table P.9, along with space for the marginal distributions.

Using integration, the area under the curve $f(x) = 2x$ and above the x-axis between the limits $x = a$ and $x = b$ is obtained by finding the **definite integral** of $f(x) = 2x$. To use the fundamental theorem of calculus, we need the indefinite integral. Using the power rule, Integral Rule 4, we obtain

$$\int 2xdx = 2\int xdx = 2\left[\frac{1}{2}x^2 + C\right] = x^2 + 2C$$
$$= x^2 + C_1 = F(x) + C_1$$

where $F(x) = x^2$ and the constant of integration is C_1. The area we seek is given by

$$\int_a^b 2xdx = F(b) - F(a) = b^2 - a^2 \qquad \text{(A.14)}$$

This is the same answer we obtained in (A.13) using geometry.

Many times the algebra is abbreviated, because the constant of integration does not affect the definite integral. You will see for definite integrals

$$\int_a^b 2xdx = x^2\Big|_a^b = b^2 - a^2$$

The vertical bar notation means: evaluate the expression first at b and subtract from it the value of the expression at a.

A.5 Exercises

A.1 Each of the following formulas, (1), (2), and (3), represents a supply or demand relation.

(1) $Q = -3 + 2P$ where $P = 10$
(2) $Q = 100 - 20P$ where $P = 4$
(3) $Q = 50P^{-2}$ where $P = 2$

a. Calculate the slope of each function at the given point.
b. Interpret the slope found in (a). Do the slopes change for different values of P and Q? Is it a supply curve (positive relationship) or a demand curve (inverse relationship)?
c. Calculate the elasticity of each function at the given point.
d. Interpret the elasticity found in (c). Do the elasticities change for different values of P and Q?

A.2 The infant mortality rate ($MORTALITY$) for a country is related to the annual per capita income ($INCOME$, U.S. \$1000) in that country. Three relationships that may describe this relationship are

(1) $\ln(MORTALITY) = 7.5 - 0.5\ln(INCOME)$
(2) $MORTALITY = 1400 - 100INCOME + 1.67INCOME^2$
(3) $MORTALITY = 1500 - 50INCOME$

a. Sketch each of these relationships between $MORTALITY$ and $INCOME$ between $INCOME =$ [illegible] and $INCOME = 30$.
b. For each of these relationships, calculate the elasticity of infant mortality with respect to income if (i) $INCOME = 1$, (ii) $INCOME = 3$, and (iii) $INCOME = 25$.

A.3 Suppose the rate of inflation INF, the annual percentage increase in the general price level, is related to the annual unemployment rate $UNEMP$ by the equation $INF = -3 + 7 \times (1/UNEMP)$.

a. Sketch the curve for values of $UNEMP$ between 1 and 10.
b. Where is the impact of a change in the unemployment rate the largest?
c. If the unemployment rate is 5%, what is the marginal effect of an increase in the unemployment rate on the inflation rate?

A.4 Simplify the following expressions:

a. $x^{2/3}x^{2/7}$
b. $x^{2/3} \div x^{2/7}$
c. $\left(x^6y^4\right)^{-1/2}$

A.19 Suppose your wage rate is determined by

$$WAGE = -19.68 + 2.52EDUC + 0.55EXPER - 0.007EXPER^2$$

where *EDUC* is years of schooling and *EXPER* is years of work experience. Using calculus, what value of *EXPER* maximizes *WAGE* for a person with 16 years of education? Show your work.

A.20 Suppose wages are determined by the following equation. *EDUC* = years of education, *EXPER* = years of work experience, and *FEMALE* = 1 if person is female, 0 otherwise.

$$WAGE = -23.06 + 2.85EDUC + 0.80EXPER - 0.008EXPER^2 - 9.21FEMALE$$
$$+ 0.34(FEMALE \times EDUC) - 0.015(EDUC \times EXPER)$$

Find $\partial WAGE/\partial EDUC$ for a female with 16 years of schooling and 10 years of experience. Show your work.

C.19 Does having a household member with an advanced degree increase household income relative to a household that includes a member having only a college degree? Use the sample data file *cex5_small* to explore this question.

a. Construct a histogram of incomes for households that include a member with an advanced degree. Construct another histogram of incomes for households that include a member with a college degree. What do you observe about the shape and location of these two histograms?

b. In the sample that includes a member with an advanced degree, what percentage of households have household incomes greater than $10,000 per month? What is the percentage for households that include a member having a college degree?

c. Test the null hypothesis that the population mean income for households including a member with an advanced degree, μ_{ADV}, is less than, or equal to, $9,000 per month against the alternative that it is greater than $9,000 per month. Use the 5% level of significance.

d. Test the null hypothesis that the population mean income for households including a member with a college degree, μ_{COLL}, is less than, or equal to, $9,000 per month against the alternative that it is greater than $9,000 per month. Use the 5% level of significance.

e. Construct 95% interval estimates for μ_{ADV} and μ_{COLL}.

f. Test the null hypothesis $\mu_{ADV} \le \mu_{COLL}$ against the alternative $\mu_{ADV} > \mu_{COLL}$. Use the 5% level of significance. What is your conclusion?

C.20 How much variation is there in household incomes in households including a member with an advanced degree? Use the sample data file *cex5_small* to explore this question. Let σ^2_{ADV} denote the population variance.

a. Test the null hypothesis $\sigma^2_{ADV} = 2500$ against the alternative $\sigma^2_{ADV} > 2500$. Use the 5% level of significance. Clearly state the test statistic and the rejection region. What is the **p-value** for this test?

b. Test the null hypothesis $\sigma^2_{ADV} = 2500$ against the alternative $\sigma^2_{ADV} < 2500$. Use the 5% level of significance. Clearly state the test statistic and the rejection region. What is the *p*-value for this test?

c. Test the null hypothesis $\sigma^2_{ADV} = 2500$ against the alternative $\sigma^2_{ADV} \neq 2500$. Use the 5% level of significance. Clearly state the test statistic and the rejection region.

C.21 School officials consider performance on a standardized math test acceptable if 40% of the population of students score at least 500 points. Use the sample data file *star5_small* to explore this topic.

a. Compute the sample proportion of students enrolled in regular-sized classes who score 500 points or more. Calculate a 95% interval estimate of the population proportion. Based on this interval can we reject the null hypothesis that the population proportion of students in regular-sized classes who score 500 points or better is $p = 0.4$?

b. Test the null hypothesis that the population proportion p of students in a regular-sized class who score 500 points or more is less than or equal to 0.4 against the alternative that the true proportion is greater than 0.4. Use the 5% level of significance.

c. Test the null hypothesis that the population proportion p of students in a regular-sized class who score 500 points or more is equal to 0.4 against the alternative that the true proportion is less than 0.4. Use the 5% level of significance.

d. Repeat parts (a)–(c) for students in small classes.

C.22 Consider two populations of Chinese chemical firms: those who export their products and those who do not. Let us consider the sales revenue for these two types of firms. Use the data file *chemical_small* for this exercise. It contains data on 1200 firms in 2006.

a. The variable *LSALES* is ln(*SALES*). Construct a histogram for this variable and test whether the data are normally distributed using the Jarque–Bera test with 10% level of significance.

b. Create the variable *SALES* = exp(*LSALES*). Construct a histogram for this variable and test whether the data are normally distributed using the Jarque–Bera test with 10% level of significance.

c. Consider two populations of firms: those who export ($EXPORT = 1$) and those who do not ($EXPORT = 0$). Let μ_1 be the population mean of *LSALES* for firms that export, and let μ_0 be the population mean of *LSALES* for firms that do not export. Estimate the difference in means $\mu_1 - \mu_0$ and interpret this value. [*Hint*: Use the properties of differences in log-variables.]

d. Test the hypothesis that the means of these two populations are equal. Use the test that assumes the population variances are unequal. What do you conclude?

C.23 Does additional education have as large a payoff for females as males? Use the data file *cps5* to explore this question. If your software does not permit using this larger sample use *cps5_small*.

a. Calculate the sample mean wage of females who have 12 years of education. Calculate the sample mean wage of females with 16 years of education. What did you discover?

b. Calculate a 95% interval estimate for the population mean wage of females with 12 years of education. Repeat the calculation for the wages of females with 16 years of education. Do the intervals overlap?

c. Calculate the sample mean wage of males who have 12 years of education. Calculate the sample mean wage of males with 16 years of education. What did you discover? How does the difference in wages for males compare to the difference of wages for females in part (a)?

d. Calculate a 95% interval estimate for the population mean wage of males with 12 years of education. Repeat the calculation for the wages of males with 16 years of education. Does the interval for males with 12 years of education overlap with the comparable interval for females? Does the interval for males with 16 years of education overlap with the comparable interval for females?

e. Denote the population means of interest by μ_{F16}, μ_{F12}, μ_{M16}, μ_{M12} where F and M denote female and male, and 12 and 16 denote years of education. Estimate the parameter $\theta = (\mu_{F16} - \mu_{F12}) - (\mu_{M16} - \mu_{M12})$ by replacing population means by sample means.

f. Calculate a 95% interval estimate of θ. Based on the interval estimate, what can you say about the benefits of the addition of four years of education for males versus females? Use the 97.5 percentile from the standard normal, 1.96, when calculating the interval estimate.

C.24 How much does the variation in wages change when individuals receive more education? Is the variation different for males and females? Use the data file *cps5* to explore this question. If your software does not permit using this larger sample use *cps5_small*.

a. Calculate the sample variance of wages of females who have 12 years of education. Calculate the sample variance of wages of females who have 18 years of education. What did you discover?

b. Carry out a two-tail test, using a 5% level of significance, of the hypothesis that the variance of wage is the same for females with 12 years of education and females with 18 years of education.

c. Calculate the sample variance of wages of males who have 12 years of education. Calculate the sample variance of wages of males who have 18 years of education. What did you discover?

d. Carry out a two-tail test, using a 5% level of significance, of the hypothesis that the variance of wage is the same for males with 12 years of education and males with 18 years of education.

e. Carry out a two-tail test of the null hypothesis that the mean wage for males with 18 years of education is the same as the mean wage of females with 18 years of education. Use the 1% level of significance.

C.25 What happens to the household budget share of necessity items, like food, when total household expenditures increase? Use data file *malwai_small* for this exercise.

a. Obtain the summary statistics, including the median and 90th percentile, of total household expenditures.

b. Construct a 95% interval estimate for the proportion of income spent on food by households with total expenditures less than or equal to the median.

c. Construct a 95% interval estimate for the proportion of income spent on food by households with total expenditures more than or equal to the 90th percentile.

d. Summarize your findings from parts (b) and (c).

e. Test the null hypothesis that the population mean proportion of income spent on food by households is 0.4. Use a two-tail test and the 5% level of significance. Carry out the test separately using the complete sample, and using the samples of households with total expenditures less than or equal to the median, and again for households whose total expenditures are in the top 10%.

C.26 At the famous Fulton Fish Market in New York City sales of Whiting (a type of fish) vary from day to day. Over a period of several months, daily quantities sold (in pounds) were observed. These data are in the data file *fultonfish*.

a. Using the data for Monday sales, test the null hypothesis that the mean quantity sold is greater than or equal to 10,000 pounds a day, against the alternative that the mean quantity sold is less than 10,000 pounds. Use the $\alpha = 0.05$ level of significance. Be sure to (i) state the null and alternative hypotheses, (ii) give the test statistic and its distribution, (iii) indicate the rejection region, including