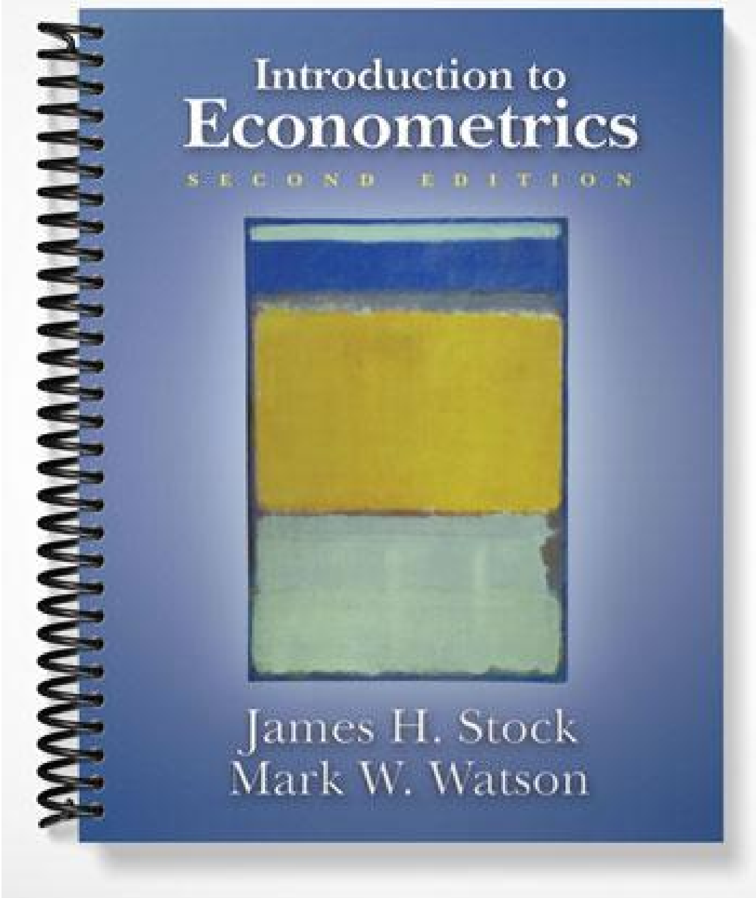
TEST BANK


Introduction to
Econometrics
SECOND EDITION
James H. Stock
Mark W. Watson

Chapter 2
Review of Probability

Multiple Choice

1) The probability of an outcome

a. is the number of times that the outcome occurs in the long run.
b. equals $M \times N$, where M is the number of occurrences and N is the population size.
c. is the proportion of times that the outcome occurs in the long run.
d. equals the sample mean divided by the sample standard deviation.

Answer: c

2) The probability of an event A or B (Pr(A or B)) to occur equals

a. $\Pr(A) \times \Pr(B)$.
b. $\Pr(A) + \Pr(B)$ if A and B are mutually exclusive.
c. $\frac{\Pr(A)}{\Pr(B)}$.
d. $\Pr(A) + \Pr(B)$ even if A and B are not mutually exclusive.

Answer: b

3) The cumulative probability distribution shows the probability

a. that a random variable is less than or equal to a particular value.
b. of two or more events occurring at once.
c. of all possible events occurring.
d. that a random variable takes on a particular value given that another event has happened.

Answer: a

4) The expected value of a discrete random variable

a. is the outcome that is most likely to occur.
b. can be found by determining the 50% value in the c.d.f.
c. equals the population median.
d. is computed as a weighted average of the possible outcome of that random variable, where the weights are the probabilities of that outcome.

Answer: d

5) Let Y be a random variable. Then var(Y) equals

a. $\sqrt{E[(Y-\mu_Y)^2]}$.
b. $E[|(Y-\mu_Y)|]$.
c. $E[(Y-\mu_Y)^2]$.
d. $E[(Y-\mu_Y)]$.

Answer: c

6) The *skewness* of the distribution of a random variable Y is defined as follows:

a. $\dfrac{E[(Y^3-\mu_Y)]}{\sigma_Y^2}$
b. $E[(Y-\mu_Y)^3]$
c. $\dfrac{E[(Y^3-\mu_Y^3)]}{\sigma_Y^3}$
d. $\dfrac{E[(Y-\mu_Y)^3]}{\sigma_Y^3}$

Answer: d

7) The skewness is most likely positive for one of the following distributions:

a. The grade distribution at your college or university.
b. The U.S. income distribution.
c. SAT scores in English.
d. The height of 18-year-old females in the U.S.

Answer: b

8) The *kurtosis* of a distribution is defined as follows:

a. $\dfrac{E[(Y-\mu_Y)^4)}{\sigma_Y^4}$

b. $\dfrac{E[(Y^4-\mu_Y^4))}{\sigma_Y^2}$

c. $\dfrac{skewness}{\text{var}(Y)}$

d. $E[(Y-\mu_Y)^4)$

Answer: a

9) For a normal distribution, the *skewness* and *kurtosis* measures are as follows:

a. 1.96 and 4
b. 0 and 0
c. 0 and 3
d. 1 and 2

Answer: c

10) The conditional distribution of Y given $X = x$, $\Pr(Y = y \mid X = x)$, is

a. $\dfrac{\Pr(Y = y)}{\Pr(X = x)}$.

b. $\sum_{i=1}^{l} \Pr(X = x_i, Y = y)$.

c. $\dfrac{\Pr(X = x, Y = y)}{\Pr(Y = y)}$

d. $\dfrac{\Pr(X = x, Y = y)}{\Pr(X = x)}$.

Answer: d

11) The conditional expectation of Y given X, $E(Y \mid X = x)$, is calculated as follows:

a. $\sum_{i=1}^{k} y_i \Pr(X = x_i \mid Y = y)$
b. $E[E(Y \mid X)]$
c. $\sum_{i=1}^{k} y_i \Pr(Y = y_i \mid X = x)$
d. $\sum_{i=1}^{l} E(Y \mid X = x_i) \Pr(X = x_i)$

Answer: c

12) Two random variables X and Y are independently distributed if all of the following conditions hold, with the exception of

a. $\Pr(Y = y \mid X = x) = \Pr(Y = y)$.
b. knowing the value of one of the variables provides no information about the other.
c. if the conditional distribution of Y given X equals the marginal distribution of Y.
d. $E(Y) = E[E(Y \mid X)]$.

Answer: d

13) The correlation between X and Y

a. cannot be negative since variances are always positive.
b. is the covariance squared.
c. can be calculated by dividing the covariance between X and Y by the product of the two standard deviations.
d. is given by $corr(X,Y) = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)\operatorname{var}(Y)}$.

Answer: c

14) Two variables are uncorrelated in all of the cases below, with the exception of

a. being independent.
b. having a zero covariance.
c. $|\sigma_{XY}| \leq \sqrt{\sigma_X^2 \sigma_Y^2}$.
d. $E(Y \mid X) = 0$.

Answer: c

15) $\text{var}(aX + bY) =$

a. $a^2\sigma_X^2 + b^2\sigma_Y^2$
b. $a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2$
c. $\sigma_{XY} + \mu_X\mu_Y$
d. $a\sigma_X^2 + b\sigma_Y^2$

Answer: b

16) To standardize a variable you

a. subtract its mean and divide by its standard deviation.
b. integrate the area below two points under the normal distribution.
c. add and subtract 1.96 times the standard deviation to the variable.
d. divide it by its standard deviation, as long as its mean is 1.

Answer: a

17) Assume that Y is normally distributed $N(\mu, \sigma^2)$. Moving from the mean (μ) 1.96 standard deviations to the left and 1.96 standard deviations to the right, then the area under the normal p.d.f. is

a. 0.67
b. 0.05
c. 0.95
d. 0.33

Answer: c

18) Assume that Y is normally distributed $N(\mu, \sigma^2)$. To find $\Pr(c_1 \le Y \le c_2)$, where $c_1 < c_2$ and $d_i = \frac{c_i - \mu}{\sigma}$, you need to calculate $\Pr(d_1 \le Z \le d_2) =$

a. $\Phi(d_2) - \Phi(d_1)$
b. $\Phi(1.96) - \Phi(-1.96)$
c. $\Phi(d_2) - (1 - \Phi(d_1))$
d. $1 - (\Phi(d_2) - \Phi(d_1))$

Answer: a

19) If variables with a multivariate normal distribution have covariances that equal zero, then

a. the correlation will most often be zero, but does not have to be.
b. the variables are independent.
c. you should use the χ^2 distribution to calculate probabilities.
d. the marginal distribution of each of the variables is no longer normal.

Answer: b

20) The Student *t* distribution is

a. the distribution of the sum of *m* squared independent standard normal random variables.
b. the distribution of a random variable with a chi-squared distribution with *m* degrees of freedom, divided by *m*.
c. always well approximated by the standard normal distribution.
d. the distribution of the ratio of a standard normal random variable, divided by the square root of an independently distributed chi-squared random variable with *m* degrees of freedom divided by *m*.

Answer: d

21) When there are ∞ degrees of freedom, the t_∞ distribution

a. can no longer be calculated.
b. equals the standard normal distribution.
c. has a bell shape similar to that of the normal distribution, but with “fatter” tails.
d. equals the χ^2_∞ distribution.

Answer: b

22) The sample average is a random variable and

a. is a single number and as a result cannot have a distribution.
b. has a probability distribution called its sampling distribution.
c. has a probability distribution called the standard normal distribution.
d. has a probability distribution that is the same as for the $Y_1,...,Y_n$ i.i.d. variables.

Answer: b

23) To infer the political tendencies of the students at your college/university, you sample 150 of them. Only one of the following is a simple random sample: You

a. make sure that the proportion of minorities are the same in your sample as in the entire student body.
b. call every fiftieth person in the student directory at 9 a.m. If the person does not answer the phone, you pick the next name listed, and so on.
c. go to the main dining hall on campus and interview students randomly there.
d. have your statistical package generate 150 random numbers in the range from 1 to the total number of students in your academic institution, and then choose the corresponding names in the student telephone directory.

Answer: d

24) The variance of $\bar{Y}, \sigma_{\bar{Y}}^2$, is given by the following formula:

a. σ_Y^2.
b. $\frac{\sigma_Y}{\sqrt{n}}$.
c. $\frac{\sigma_Y^2}{n}$.
d. $\frac{\sigma_Y^2}{\sqrt{n}}$.

Answer: c

25) The mean of the sample average $\bar{Y}, E(\bar{Y})$, is

a. $\frac{1}{n}\mu_Y$.
b. μ_Y.
c. $\frac{\mu_Y}{\sqrt{n}}$.
d. $\frac{\sigma_Y}{\mu_Y}$ for $n > 30$.

Answer: b

26) In econometrics, we typically do not rely on exact or finite sample distributions because

a. we have approximately an infinite number of observations (think of re-sampling).
b. variables typically are normally distributed.
c. the covariances of Y_i, Y_j are typically not zero.
d. asymptotic distributions can be counted on to provide good approximations to the exact sampling distribution (given the number of observations available in most cases).

Answer: d

27) Consistency for the sample average $\bar{Y}$ can be defined as follows, with the exception of

a. $\bar{Y}$ converges in probability to μ_Y.
b. $\bar{Y}$ has the smallest variance of all estimators.
c. $\bar{Y} \xrightarrow{p} \mu_Y$.
d. the probability of $\bar{Y}$ being in the range $\mu_Y \pm c$ becomes arbitrarily close to one as n increases for any constant $c > 0$.

Answer: b

28) The central limit theorem states that

a. the sampling distribution of $\frac{\bar{Y} - \mu_Y}{\sigma_{\bar{Y}}}$ is approximately normal.
b. $\bar{Y} \xrightarrow{p} \mu_Y$.
c. the probability that $\bar{Y}$ is in the range $\mu_Y \pm c$ becomes arbitrarily close to one as n increases for any constant $c > 0$.
d. the t distribution converges to the F distribution for approximately $n > 30$.

Answer: a

29) The central limit theorem

a. states conditions under which a variable involving the sum of $Y_1,...,Y_n$ i.i.d. variables becomes the standard normal distribution.
b. postulates that the sample mean $\bar{Y}$ is a consistent estimator of the population mean μ_Y.
c. only holds in the presence of the law of large numbers.
d. states conditions under which a variable involving the sum of $Y_1,...,Y_n$ i.i.d. variables becomes the Student *t* distribution.

Answer: a

30) The covariance inequality states that

a. $0 \le \sigma_{XY}^2 \le 1$.
b. $\sigma_{XY}^2 \le \sigma_X^2 \sigma_Y^2$.
c. $\sigma_{XY}^2 - \sigma_X^2 \le \sigma_Y^2$.
d. $\sigma_{XY}^2 \le \frac{\sigma_X^2}{\sigma_Y^2}$.

Answer: b

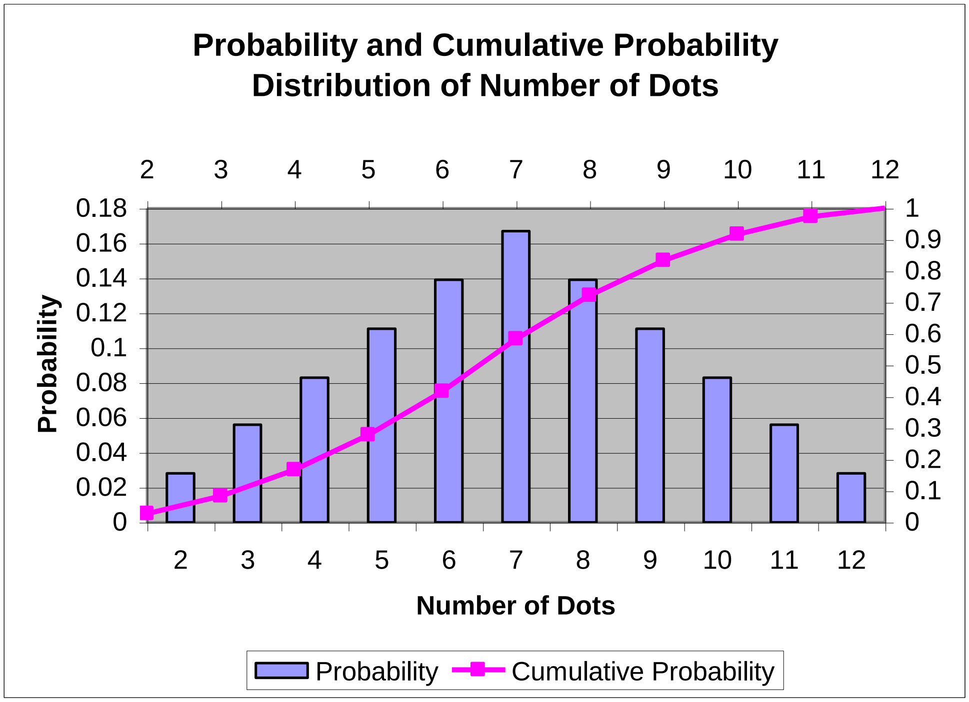
Essays and Longer Questions

1) Think of the situation of rolling two dice and let *M* denote the sum of the number of dots on the two dice. (So *M* is a number between 1 and 12.)

(a) In a table, list all of the possible outcomes for the random variable *M* together with its probability distribution and cumulative probability distribution. Sketch both distributions.

Answer:

Outcome (sum of dots)	2	3	4	5	6	7	8	9	10	11	12
Probability distribution	0.028	0.056	0.083	0.111	0.139	0.167	0.139	0.111	0.083	0.056	0.028
Cumulative probability distribution	0.028	0.083	0.167	0.278	0.417	0.583	0.722	0.833	0.912	0.972	1.000



(b) Calculate the expected value and the standard deviation for M.

Answer: 7.0; 2.42.

(c) Looking at the sketch of the probability distribution, you notice that it resembles a normal distribution. Should you be able to use the standard normal distribution to calculate probabilities of events? Why or why not?

Answer: You cannot use the normal distribution (without continuity correction) to calculate probabilities of events, since the probability of any event equals zero.

2) What is the probability of the following outcomes?

(a) $\Pr(M = 7)$
(b) $\Pr(M = 2 \text{ or } M = 10)$
(c) $\Pr(M = 4 \text{ or } M \neq 4)$
(d) $\Pr(M = 6 \text{ and } M = 9)$
(e) $\Pr(M < 8)$
(f) $\Pr(M = 6 \text{ or } M > 10)$

Answer: (a) 0.167 or $\frac{6}{36} = \frac{1}{6}$; (b) 0.111 or $\frac{4}{39} = \frac{1}{9}$; (c) 1; (d) 0; (e) 0.583; (f) 0.222 or $\frac{8}{36} = \frac{2}{9}$.

3) Probabilities and relative frequencies are related in that the probability of an outcome is the proportion of the time that the outcome occurs in the long run. Hence concepts of joint, marginal, and conditional probability distributions stem from related concepts of frequency distributions.

You are interested in investigating the relationship between the age of heads of households and weekly earnings of households. The accompanying data gives the number of occurrences grouped by age and income. You collect data from 1,744 individuals and think of these individuals as a population that you want to describe, rather than a sample from which you want to infer behavior of a larger population. After sorting the data, you generate the accompanying table:

Joint Absolute Frequencies of Age and Income, 1,744 Households

		Age of head of household			
	X_1	X_2	X_3	X_4	X_5
Household Income	16-under 20	20-under 25	25-under 45	45-under 65	65 and >
Y_1 \$0-under \$200	80	76	130	86	24
Y_2 \$200-under \$400	13	90	346	140	8
Y_3 \$400-under \$600	0	19	251	101	6
Y_4 \$600-under \$800	1	11	110	55	1
Y_5 \$800 and >	1	1	108	84	2

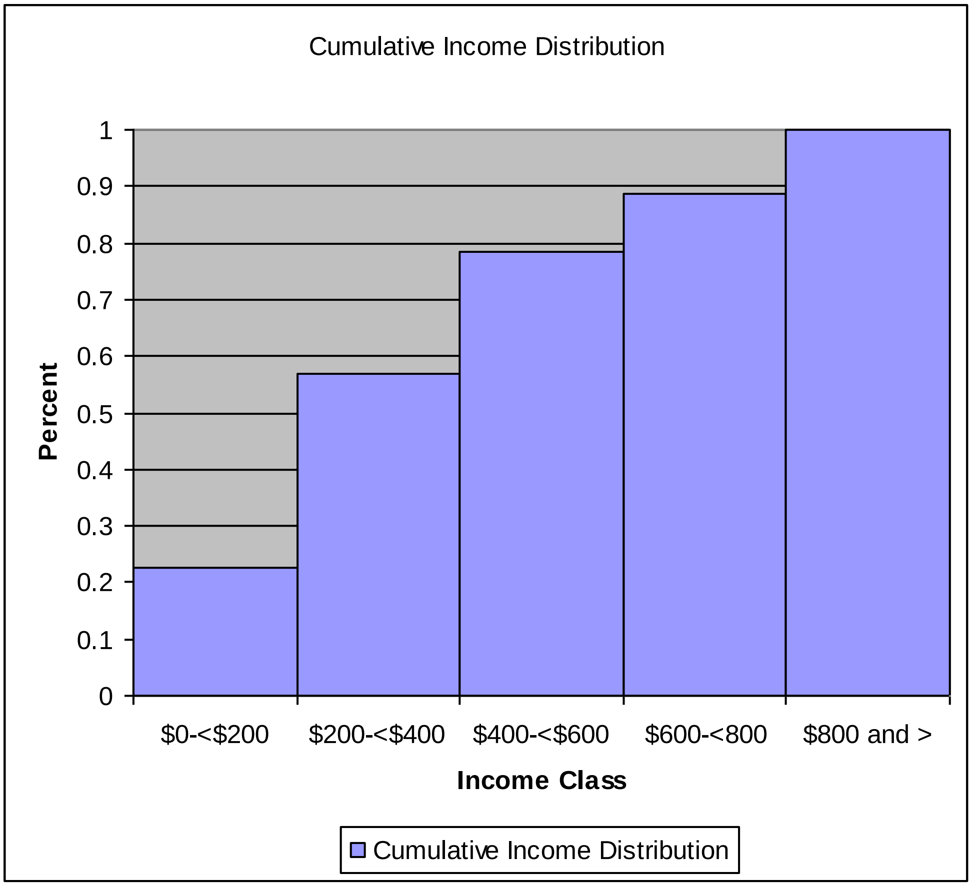
The median of the income group of \$800 and above is \$1,050.

(a) Calculate the joint relative frequencies and the marginal relative frequencies. Interpret one of each of these. Sketch the cumulative income distribution.

Answer: The joint relative frequencies and marginal relative frequencies are given in the accompanying table. 5.2 percent of the individuals are between the age of 20 and 24, and make between \$200 and under \$400. 21.6 percent of the individuals earn between \$400 and under \$600.

Joint Relative and Marginal Frequencies of Age and Income, 1,744 Households

		Age of head of household				
	X_1	X_2	X_3	X_4	X_5	
Household Income	16-under 20	20-under 25	25-under 45	45-under 65	65 and >	**Total**
Y_1 \$0-under \$200	0.046	0.044	0.075	0.049	0.014	0.227
Y_2 \$200-under \$400	0.007	0.052	0.198	0.080	0.005	0.342
Y_3 \$400-under \$600	0.000	0.011	0.144	0.058	0.003	0.216
Y_4 \$600-under \$800	0.001	0.006	0.063	0.032	0.001	0.102
Y_5 \$800 and >	0.001	0.001	0.062	0.048	0.001	0.112



(b) Calculate the conditional relative income frequencies for the two age categories 16-under 20, and 45-under 65. Calculate the mean household income for both age categories.

Answer: The mean household income for the 16-under 20 age category is roughly \$144. It is approximately \$489 for the 45-under 65 age category.

Conditional Relative Frequencies of Income and Age 16-under 20, and 45-under 65, 1,744 Households

		Age of head of household
	X_1	X_4
Household Income	16-under 20	45-under 65
Y_1 \$0-under \$200	0.842	0.185
Y_2 \$200-under \$400	0.137	0.300
Y_3 \$400-under \$600	0.000	0.217
Y_4 \$600-under \$800	0.001	0.118
Y_5 \$800 and >	0.001	0.180

(c) If household income and age of head of household were independently distributed, what would you expect these two conditional relative income distributions to look like? Are they similar here?

Answer: They would have to be identical, which they clearly are not.

(d) Your textbook has given you a primary definition of independence that does not involve conditional relative frequency distributions. What is that definition? Do you think that age and income are independent here, using this definition?

Answer: $\Pr(Y = y, X = x) = \Pr(Y = y)\Pr(X = x)$. We can check this by multiplying two marginal probabilities to see if this results in the joint probability. For example, $\Pr(Y = Y_3) = 0.216$ and $\Pr(X = X_3) = 0.542$, resulting in a product of 0.117, which does not equal the joint probability of 0.144. Given that we are looking at the data as a population, not a sample, we do not have to test how "close" 0.117 is to 0.144.

4) Math and verbal SAT scores are each distributed normally with $N(500,10000)$.

(a) What fraction of students scores above 750? Above 600? Between 420 and 530? Below 480? Above 530?

Answer: Pr(Y>750) = 0.0062; Pr(Y>600) = 0.1587; Pr(420<Y<530) = 0.4061; Pr(Y<480) = 0.4270; Pr(Y>530) = 0.3821.

(b) If the math and verbal scores were independently distributed, which is not the case, then what would be the distribution of the overall SAT score? Find its mean and variance.

Answer: The distribution would be $N(1000, 20000)$, using equations (2.29) and (2.31) in the textbook. Note that the standard deviation is now roughly 141 rather than 200.

(c) Next, assume that the correlation coefficient between the math and verbal scores is 0.75. Find the mean and variance of the resulting distribution.

Answer: Given the correlation coefficient, the distribution is now $N(1000, 35000)$, which has a standard deviation of approximately 187.

(d) Finally, assume that you had chosen 25 students at random who had taken the SAT exam. Derive the distribution for their average math SAT score. What is the probability that this average is above 530? Why is this so much smaller than your answer in (a)?

Answer: The distribution for the average math SAT score is $N(500, 400)$. $\Pr(\bar{Y} > 530)$ = 0.0668. This probability is smaller because the sample mean has a smaller standard deviation (20 rather than 100).

5) The following problem is frequently encountered in the case of a rare disease, say AIDS, when determining the probability of actually having the disease after testing positively for HIV. (This is often known as the accuracy of the test given that you have the disease.) Let us set up the problem as follows: Y = 0 if you tested negative using the ELISA test for HIV, Y = 1 if you tested positive; X = 1 if you have HIV, X = 0 if you do not have HIV. Assume that 0.1 percent of the population has HIV and that the accuracy of the test is 0.95 in both cases of (i) testing positive when you have HIV, and (ii) testing negative when you do not have HIV. (The actual ELISA test is actually 99.7 percent accurate when you have HIV, and 98.5 percent accurate when you do not have HIV.)

(a) Assuming arbitrarily a population of 10,000,000 people, use the accompanying table to first enter the column totals.

	Test Positive (Y=1)	Test Negative (Y=0)	**Total**
HIV (X=1)			
No HIV (X=0)			
Total			10,000,000

Answer:

	Test Positive (Y=1)	Test Negative (Y=0)	**Total**
HIV (X=1)			10,000
No HIV (X=0)			9,990,000
Total			10,000,000

(b) Use the conditional probabilities to fill in the joint absolute frequencies.

Answer:

	Test Positive (Y=1)	Test Negative (Y=0)	**Total**
HIV (X=1)	9,500	500	10,000
No HIV (X=0)	499,500	9,490,500	9,990,000
Total			10,000,000

(c) Fill in the marginal absolute frequencies for testing positive and negative. Determine the conditional probability of having HIV when you have tested positive. Explain this surprising result.

Answer:

	Test Positive (Y=1)	Test Negative (Y=0)	**Total**
HIV (X=1)	9,500	500	10,000
No HIV (X=0)	499,500	9,490,500	9,990,000
Total	509,000	9,491,000	10,000,000

$\Pr(X=1|Y=1) = 0.0187$. Although the test is quite accurate, there are very few people who have HIV (10,000), and many who do not have HIV (9,999,000). A small percentage of that large number (499,500/9,990,000) is large when compared to the higher percentage of the smaller number (9,500/10,000).

(d) The previous problem is an application of Bayes' theorem, which converts $\Pr(Y = y \mid X = x)$ into $\Pr(X = x \mid Y = y)$. Can you think of other examples where $\Pr(Y = y \mid X = x) \neq \Pr(X = x \mid Y = y)$?

Answer: Answers will vary by student. Perhaps a nice illustration is the probability to be a male given that you play on the college/university men's varsity team, versus the probability to play on the college/university men's varsity team given that you are a male student.

6) You have read about the so-called catch-up theory by economic historians, whereby nations that are further behind in per capita income grow faster subsequently. If this is true systematically, then eventually laggards will reach the leader. To put the theory to the test, you collect data on relative (to the United States) per capita income for two years, 1960 and 1990, for 24 OECD countries. You think of these countries as a population you want to describe, rather than a sample from which you want to infer behavior of a larger population. The relevant data for this question is as follows:

Y	X_1	X_2	$Y \times X_1$	Y^2	X_1^2	X_2^2
0.023	0.770	1.030	0.018	0.00053	0.593	1.0609
0.014	1.000	1.000	0.014	0.00020	1.000	1.0000
….	….	….	….	….	….	….
0.041	0.200	0.450	0.008	0.00168	0.040	0.2025
0.033	0.130	0.230	0.004	0.00109	0.017	0.0529
0.625	13.220	17.800	0.294	0.01877	8.529	13.9164

where X_1 and X_2 are per capita income relative to the United States in 1960 and 1990 respectively, and Y is the average annual growth rate in X over the 1960-1990 period. Numbers in the last row represent sums of the columns above.

(a) Calculate the variance and standard deviation of X_1 and X_2. For a catch-up effect to be present, what relationship must the two standard deviations show? Is this the case here?

Answer: The variances of X_1 and X_2 are 0.0520 and 0.0298 respectively, with standard deviations of 0.2279 and 0.1726. For the catch-up effect to be present, the standard deviation would have to shrink over time. This is the case here.

(b) Calculate the correlation between Y and X_1. What sign must the correlation coefficient have for there to be evidence of a catch-up effect? Explain.

Answer: The correlation coefficient is –0.88. It has to be negative for there to be evidence of a catch-up effect. If countries that were relatively ahead in the initial period and in terms of per capita income grow by relatively less over time, then eventually the laggards will catch-up.

7) Following Alfred Nobel's will, there are five Nobel Prizes awarded each year. These are for outstanding achievements in Chemistry, Physics, Physiology or Medicine, Literature, and Peace. In 1968, the Bank of Sweden added a prize in Economic Sciences in memory of Alfred Nobel. You think of the data as describing a population, rather than a sample from which you want to infer behavior of a larger population. The accompanying table lists the joint probability distribution between recipients in economics and the other five prizes, and the citizenship of the recipients, based on the 1969-2001 period.

Joint Distribution of Nobel Prize Winners in Economics and Non-Economics Disciplines, and Citizenship, 1969-2001

	U.S. Citizen ($Y=0$)	Non-U.S. Citizen ($Y=1$)	Total
Economics Nobel Prize ($X=0$)	0.118	0.049	0.167
Physics, Chemistry, Medicine, Literature, and Peace Nobel Prize ($X=1$)	0.345	0.488	0.833
Total	0.463	0.537	1.00

(a) Compute $E(Y)$ and interpret the resulting number.

Answer: $E(Y)=0.537$. 53.7 percent of Nobel Prize winners were non-U.S. citizens.

(b) Calculate and interpret $E(Y \mid X=1)$ and $E(Y \mid X=0)$.

Answer: $E(Y \mid X=1)=0.586$. 58.6 percent of Nobel Prize winners in non-economics disciplines were non-U.S. citizens. $E(Y \mid X=0)=0.293$. 29.3 percent of the Economics Nobel Prize winners were non-U.S. citizens.

(c) A randomly selected Nobel Prize winner reports that he is a non-U.S. citizen. What is the probability that this genius has won the Economics Nobel Prize? A Nobel Prize in the other five disciplines?

Answer: There is a 9.1 percent chance that he has won the Economics Nobel Prize, and a 90.9 percent chance that he has won a Nobel Prize in one of the other five disciplines.

(d) Show what the joint distribution would look like if the two categories were independent.

Answer:

Joint Distribution of Nobel Prize Winners in Economics and Non-Economics Disciplines, and Citizenship, 1969-2001, under assumption of independence

	U.S. Citizen ($Y = 0$)	Non= U.S. Citizen ($Y = 1$)	Total
Economics Nobel Prize ($X = 0$)	0.077	0.090	0.167
Physics, Chemistry, Medicine, Literature, and Peace Nobel Prize ($X = 1$)	0.386	0.447	0.833
Total	0.463	0.537	1.00

8) A few years ago the news magazine *The Economist* listed some of the stranger explanations used in the past to predict presidential election outcomes. These included whether or not the hemlines of women's skirts went up or down, stock market performances, baseball World Series wins by an American League team, etc. Thinking about this problem more seriously, you decide to analyze whether or not the presidential candidate for a certain party did better if his party controlled the house. Accordingly you collect data for the last 34 presidential elections. You think of this data as comprising a population which you want to describe, rather than a sample from which you want to infer behavior of a larger population. You generate the accompanying table:

Joint Distribution of Presidential Party Affiliation and Party Control of House of Representatives, 1860-1996

	Democratic Control of House ($Y = 0$)	Republican Control of House ($Y = 1$)	Total
Democratic President ($X = 0$)	0.412	0.030	0.441
Republican President ($X = 1$)	0.176	0.382	0.559
Total	0.588	0.412	1.00

(a) Interpret one of the joint probabilities and one of the marginal probabilities.

Answer: 38.2 percent of the presidents were Republicans and were in the White House while Republicans controlled the House of Representatives. 44.1 percent of all presidents were Democrats.

(b) Compute $E(X)$. How does this differ from $E(X \mid Y = 0)$? Explain..

Answer: $E(X)$ = 0.559. $E(X \mid Y = 0)$ = 0.701. $E(X)$ gives you the unconditional expected value, while $E(X \mid Y = 0)$ is the conditional expected value.

(c) If you picked one of the Republican presidents at random, what is the probability that during his term the Democrats had control of the House?

Answer: $E(X) = 0.559$. 55.9 percent of the presidents were Republicans. $E(X \mid Y = 0) = 0.299$. 29.9 percent of those presidents who were in office while Democrats had control of the House of Representatives were Republicans. The second conditions on those periods during which Democrats had control of the House of Representatives, and ignores the other periods.

(d) What would the joint distribution look like under independence? Check your results by calculating the two conditional distributions and compare these to the marginal distribution.

Answer:

Joint Distribution of Presidential Party Affiliation and Party Control of House of Representatives, 1860-1996, under the Assumption of Independence

	Democratic Control of House ($Y = 0$)	Republican Control of House ($Y = 1$)	Total
Democratic President ($X = 0$)	0.259	0.182	0.441
Republican President ($X = 1$)	0.329	0.230	0.559
Total	0.588	0.412	1.00

$$\Pr(X = 0 \mid Y = 0) = \frac{0.259}{0.588} = 0.440 \text{ (there is a small rounding error).}$$

$$\Pr(Y = 1 \mid X = 1) = \frac{0.230}{0.559} = 0.411 \text{ (there is a small rounding error).}$$

9) The expectations augmented Phillips curve postulates

$$\Delta p = \pi - f(u - \bar{u}),$$

where Δp is the actual inflation rate, π is the expected inflation rate, and u is the unemployment rate, with "–" indicating equilibrium (the NAIRU – Non-Accelerating Inflation Rate of Unemployment). Under the assumption of static expectations ($\pi = \Delta p_{-1}$), *i.e.*, that you expect this period's inflation rate to hold for the next period ("the sun shines today, it will shine tomorrow"), then the prediction is that inflation will accelerate if the unemployment rate is below its equilibrium level. The accompanying table below displays information on accelerating annual inflation and unemployment rate differences from the equilibrium rate (cyclical unemployment), where the latter is approximated by a five-year moving average. You think of this data as a population which you want to describe, rather than a sample from which you want to infer behavior of a larger population. The data is collected from United States quarterly data for the period 1964:1 to 1995:4.

Joint Distribution of Accelerating Inflation and Cyclical Unemployment, 1964:1-1995:4

	$(u-\bar{u})>0$ $(Y=0)$	$(u-\bar{u})\geq 0$ $(Y=1)$	Total
$\Delta p - \Delta p_{-1} > 0$ $(X=0)$	0.156	0.383	0.539
$\Delta p - \Delta p_{-1} \leq 0$ $(X=1)$	0.297	0.164	0.461
Total	0.453	0.547	1.00

(a) Compute $E(Y)$ and $E(X)$, and interpret both numbers.

Answer: $E(Y) = 0.547$. 54.7 percent of the quarters saw cyclical unemployment.
$E(X) = 0.461$. 46.1 percent of the quarters saw decreasing inflation rates.

(b) Calculate $E(Y \mid X=1)$ and $E(Y \mid X=0)$. If there was independence between cyclical unemployment and acceleration in the inflation rate, what would you expect the relationship between the two expected values to be? Given that the two means are different, is this sufficient to assume that the two variables are independent?

Answer: $E(Y \mid X=1) = 0.356$; $E(Y \mid X=0) = 0.711$. You would expect the two conditional expectations to be the same. In general, independence in means does not imply statistical independence, although the reverse is true.

(c) What is the probability of inflation to increase if there is positive cyclical unemployment? Negative cyclical unemployment?

Answer: There is a 34.4 percent probability of inflation to increase if there is positive cyclical unemployment. There is a 70 percent probability of inflation to increase if there is negative cyclical unemployment.

(d) You randomly select one of the 59 quarters when there was positive cyclical unemployment ($(u - \bar{u}) > 0$). What is the probability there was decelerating inflation during that quarter?

Answer: There is a 65.6 percent probability of inflation to decelerate when there is positive cyclical unemployment.

10) The accompanying table shows the joint distribution between the change of the unemployment rate in an election year and the share of the candidate of the incumbent party since 1928. You think of this data as a population which you want to describe, rather than a sample from which you want to infer behavior of a larger population.

Joint Distribution of Unemployment Rate Change and Incumbent Party's Vote Share in Total Vote Cast for the Two Major-Party Candidates, 1928-2000

	$(Incumbent - 50\%) > 0$ ($Y = 0$)	$(Incumbent - 50\%) \leq 0$ ($Y = 1$)	Total
$\Delta u > 0$ ($X = 0$)	0.053	0.211	0.264
$\Delta u \leq 0$ ($X = 1$)	0.579	0.157	0.736
Total	0.632	0.368	1.00

(a) Compute and interpret $E(Y)$ and $E(X)$.

Answer: $E(Y) = 0.368$; $E(X) = 0.736$. The probability of an incumbent to have less than 50% of the share of votes cast for the two major-party candidates is 0.368. The probability of observing falling unemployment rates during the election year is 73.6 percent.

(b) Calculate $E(Y \mid X = 1)$ and $E(Y \mid X = 0)$. Did you expect these to be very different?

Answer: $E(Y \mid X = 1) = 0.213$; $E(Y \mid X = 0) = 0.799$. A student who believes that incumbents will attempt to manipulate the economy to win elections will answer affirmatively here.

(c) What is the probability that the unemployment rate decreases in an election year?

Answer: $\Pr(X = 1) = 0.736$.

(d) Conditional on the unemployment rate decreasing, what is the probability that an incumbent will lose the election?

Answer: $\Pr(Y = 1 | X = 1) = 0.213$.

(e) What would the joint distribution look like under independence?

Answer:

Joint Distribution of Unemployment Rate Change and Incumbent Party's Vote Share in Total Vote Cast for the Two Major-Party Candidates, 1928-2000 under Assumption of Statistical Independence

	$(Incumbent - 50\%) > 0$ ($Y = 0$)	$(Incumbent - 50\%) \leq 0$ ($Y = 1$)	Total
$\Delta u > 0$ ($X = 0$)	0.167	0.097	0.264
$\Delta u \leq 0$ ($X = 1$)	0.465	0.271	0.736
Total	0.632	0.368	1.00

11) The table accompanying lists the joint distribution of unemployment in the United States in 2001 by demographic characteristics (race and gender).

Joint Distribution of Unemployment by Demographic Characteristics, United States, 2001

	White ($Y = 0$)	Black and Other ($Y = 1$)	Total
Age 16-19 ($X = 0$)	0.13	0.05	0.18
Age 20 and above ($X = 1$)	0.60	0.22	0.82
Total	0.73	0.27	1.00

(a) What is the percentage of unemployed white teenagers?

Answer: $\Pr(Y = 0, X = 0) = 0.13$.

(b) Calculate the conditional distribution for the categories "white" and "black and other."

Answer: **Conditional Distribution of Unemployment by Demographic Characteristics, United States, 2001**

	White ($Y=0$)	Black and Other ($Y=1$)
Age 16-19 ($X=0$)	0.18	0.19
Age 20 and above ($X=1$)	0.82	0.81
Total	1.00	1.00

(c) Given your answer in the previous question, how do you reconcile this fact with the probability to be 60% of finding an unemployed adult white person, and only 22% for the category "black and other."

Answer: The original table showed the joint probability distribution, while the table in (b) presented the conditional probability distribution.

Mathematical and Graphical Problems

1) Think of an example involving five possible quantitative outcomes of a discrete random variable and attach a probability to each one of these outcomes. Display the outcomes, probability distribution, and cumulative probability distribution in a table. Sketch both the probability distribution and the cumulative probability distribution.

Answer: Answers will vary by student. The generated table should be similar to Table 2.1 in the text, and figures should resemble Figures 2.1 and 2.2 in the text.

2) The height of male students at your college/university is normally distributed with a mean of 70 inches and a standard deviation of 3.5 inches. If you had a list of telephone numbers for male students for the purpose of conducting a survey, what would be the probability of randomly calling one of these students whose height is

(a) taller than 6'0"?
(b) between 5'3" and 6'5"?
(c) shorter than 5'7", the mean height of female students?
(d) shorter than 5'0"?
(e) taller than Shaq O'Neal, the center of the Miami Heat, who is 7'1" tall? Compare this to the probability of a woman being pregnant for 10 months (300 days), where days of pregnancy is normally distributed with a mean of 266 days and a standard deviation of 16 days.

Answer: (a) $\Pr(Z > 0.5714) = 0.2839$;
(b) $\Pr(-2 < Z < 2) = 0.9545$ or approximately 0.95;
(c) $\Pr(Z < -0.8571) = 0.1957$;
(d) $\Pr(Z < -2.8571) = 0.0021$;
(e) $\Pr(Z > 4.2857) = 0.000009$ (the text does not show values above 2.99 standard deviations, Pr(Z>2.99 = 0.0014) and $\Pr(Z > 2.1250) = 0.0168$.

3) Calculate the following probabilities using the standard normal distribution. Sketch the probability distribution in each case, shading in the area of the calculated probability.

(a) $\Pr(Z < 0.0)$
(b) $\Pr(Z \leq 1.0)$
(c) $\Pr(Z > 1.96)$
(d) $\Pr(Z < -2.0)$
(e) $\Pr(Z > 1.645)$
(f) $\Pr(Z > -1.645)$
(g) $\Pr(-1.96 < Z < 1.96)$
(h) $\Pr(Z < 2.576$ or $Z > 2.576)$
(i) $\Pr(Z > z) = 0.10$; find z.
(j) $\Pr(Z < -z$ or $Z > z) = 0.05$; find z.

Answer: (a) 0.5000
(b) 0.8413
(c) 0.0250
(d) 0.0228
(e) 0.0500
(f) 0.9500
(g) 0.0500
(h) 0.0100
(i) 1.2816
(j) 1.96

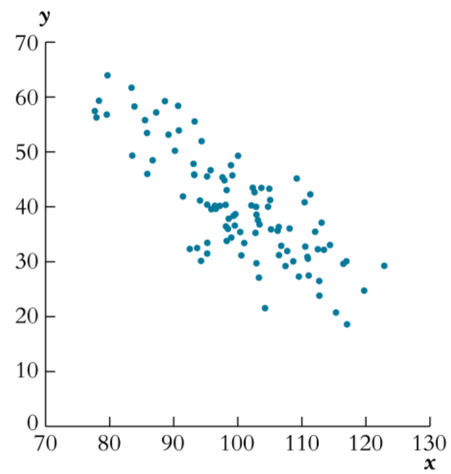
4) Using the fact that the standardized variable Z is a linear transformation of the normally distributed random variable Y, derive the expected value and variance of Z.

Answer: $Z = \frac{Y - \mu_Y}{\sigma_Y} = -\frac{\mu_Y}{\sigma_Y} + \frac{1}{\sigma_Y} Y = a + bY$, with $a = -\frac{\mu_Y}{\sigma_Y}$ and $b = \frac{1}{\sigma_Y}$. Given (2.29) and (2.30) in the text, $E(Z) = -\frac{\mu_Y}{\sigma_Y} + \frac{1}{\sigma_Y}\mu_Y = 0$, and $\sigma_Z = \frac{1}{\sigma_Z^2}\sigma_Z^2 = 1$.

5) Show in a scatterplot what the relationship between two variables X and Y would look like if there was

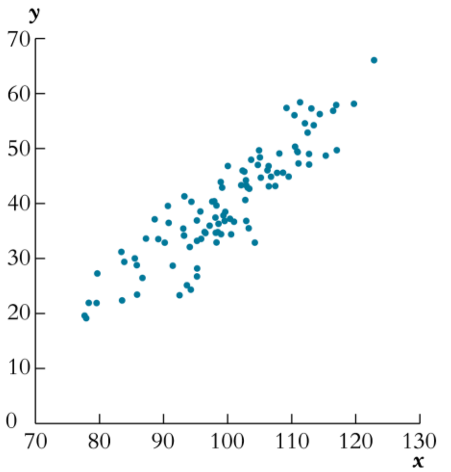
(a) a strong negative correlation

Answer:



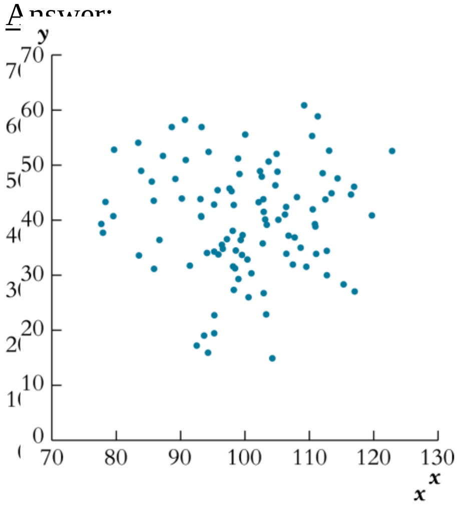
(b) a strong positive correlation

Answer:



(c) no correlation

Answer:



6) What would the correlation coefficient be if all observations for the two variables were on a curve described by $Y = X^2$?

Answer: The correlation coefficient would be zero in this case, since the relationship is non-linear.

7) Find the following probabilities:

(a) Y is distributed χ_4^2. Find $\Pr(Y > 9.49)$.

Answer: 0.05.

(b) Y is distributed t_∞. Find Pr(Y > –0.5).

Answer: 0.6915.

(c) Y is distributed $F_{4,\infty}$. Find Pr(Y < 3.32).

Answer: 0.99.

(d) Y is distributed $N(500,10000)$. Find Pr(Y > 696 or Y < 304).

Answer: 0.05.

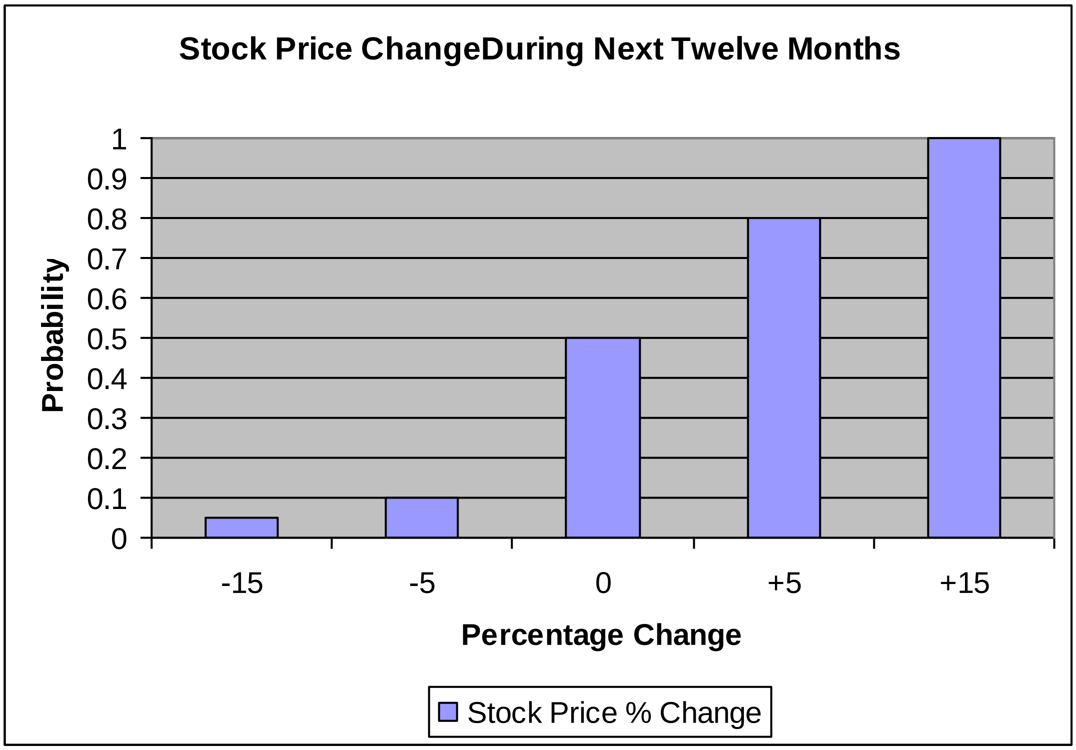
8) In considering the purchase of a certain stock, you attach the following probabilities to possible changes in the stock price over the next year.

Stock Price Change During Next Twelve Months (%)	Probability
+15	0.2
+5	0.3
0	0.4
-5	0.05
-15	0.05

What is the expected value, the variance, and the standard deviation? Which is the most likely outcome? Sketch the cumulative distribution function.

.

Answer: $E(Y) = 3.5$; $\sigma_Y^2 = 8.49$; $\sigma_Y = 2.91$; most likely: 0.



9) You consider visiting Montreal during the break between terms in January. You go to the relevant Web site of the official tourist office to figure out the type of clothes you should take on the trip. The site lists that the average high during January is -7^0 C, with a standard deviation of 4^0 C. Unfortunately you are more familiar with Fahrenheit than with Celsius, but find that the two are related by the following linear function:

$$C = \frac{5}{9}(F - 32).$$

Find the mean and standard deviation for the January temperature in Montreal in Fahrenheit.

Answer: Using equations (2.29) and (2.30) from the textbook, the result is 19.4 and 7.2.

10) Two random variables are independently distributed if their joint distribution is the product of their marginal distributions. It is intuitively easier to understand that two random variables are independently distributed if all conditional distributions of *Y* given *X* are equal. Derive one of the two conditions from the other.

Answer: If all conditional distributions of *Y* given *X* are equal, then

$$\underline{\Pr(Y=y \mid X=1)=\Pr(Y=y \mid X=2)=\ldots=\Pr(Y=y \mid X=l)}.$$

But if all conditional distributions are equal, then they must also equal the marginal distribution, i.e.,

$$\Pr(Y=y \mid X=x)=\Pr(Y=y).$$

Given the definition of the conditional distribution of *Y* given $X = x$, you then get

$$\Pr(Y=y \mid X=x)=\frac{\Pr(Y=y, X=x)}{\Pr(X=x)}=\Pr(Y=y),$$

which gives you the condition

$$\Pr(Y=y, X=x)=\Pr(Y=y)\Pr(X=x).$$

11) There are frequently situations where you have information on the conditional distribution of *Y* given *X*, but are interested in the conditional distribution of *X* given *Y*. Recalling $\Pr(Y=y \mid X=x)=\frac{\Pr(X=x, Y=y)}{\Pr(X=x)}$, derive a relationship between $\Pr(X=x \mid Y=y)$ and $\Pr(Y=y \mid X=x)$. This is called Bayes' theorem.

Answer: Given $\Pr(Y=y \mid X=x)=\frac{\Pr(X=x, Y=y)}{\Pr(X=x)}$,

$\Pr(Y=y \mid X=x)\times\Pr(X=x)=\Pr(X=x, Y=y)$;

similarly $\Pr(X=x \mid Y=y)=\frac{\Pr(X=x, Y=y)}{\Pr(Y=y)}$ and

$\Pr(X=x \mid Y=y)\times\Pr(Y=y)=\Pr(X=x, Y=y)$. Equating the two and solving for $\Pr(X=x \mid Y=y)$ then results in

$$\Pr(X=x \mid Y=y)=\frac{\Pr(Y=y \mid X=x)\times\Pr(X=x)}{\Pr(Y=y)}.$$

12) You are at a college of roughly 1,000 students and obtain data from the entire freshman class (250 students) on height and weight during orientation. You consider this to be a population that you want to describe, rather than a sample from which you want to infer general relationships in a larger population. Weight (*Y*) is measured in pounds and height (*X*) is measured in inches. You calculate the following sums:

$$\sum_{i=1}^{n} y_i^2 = 94{,}228.8,\ \sum_{i=1}^{n} x_i^2 = 1{,}248.9,\ \sum_{i=1}^{n} x_i y_i = 7{,}625.9$$

(small letters refer to deviations from means as in $z_i = Z_i - \bar{Z}$).

(a) Given your general knowledge about human height and weight of a given age, what can you say about the shape of the two distributions?

Answer: Both distributions are bound to be normal.

(b) What is the correlation coefficient between height and weight here?

Answer: 0.703.

13) Use the definition for the conditional distribution of *Y* given $X = x$ and the marginal distribution of *X* to derive the formula for $\Pr(X = x, Y = y)$. This is called the multiplication rule. Use it to derive the probability for drawing two aces randomly from a deck of cards (no joker), where you do not replace the card after the first draw. Next, generalizing the multiplication rule and assuming independence, find the probability of having four girls in a family with four children.

Answer: $\frac{4}{52} \times \frac{3}{51} = 0.0045$; 0.0625 or $\left(\frac{1}{2}\right)^4 = \left(\frac{1}{16}\right)$.

14) The systolic blood pressure of females in their 20s is normally distributed with a mean of 120 with a standard deviation of 9. What is the probability of finding a female with a blood pressure of less than 100? More than 135? Between 105 and 123? You visit the women's soccer team on campus, and find that the average blood pressure of the 25 members is 114. Is it likely that this group of women came from the same population?

Answer: Pr(Y<100) = 0.0131; Pr(Y>135) = 0.0478; Pr(105<Y<123) = 0.6784; $\Pr(\bar{Y} < 114) = \Pr(Z < -3.33) = 0.0004$. (The smallest z-value listed in the table in the textbook is –2.99, which generates a probability value of 0.0014.) This unlikely that this group of women came from the same population.

15) Show that the correlation coefficient between *Y* and *X* is unaffected if you use a linear transformation in both variables. That is, show that $corr(X,Y) = corr(X^*,Y^*)$, where $X^* = a + bX$ and $Y^* = c + dY$, and where *a*, *b*, *c*, and *d* are arbitrary non-zero constants.

Answer: $corr(X^*,Y^*) = \frac{\text{cov}(X^*,Y^*)}{\sqrt{\text{var}(X^*)}\sqrt{\text{var}(Y^*)}} = \frac{bd\,\text{cov}(X,Y)}{\sqrt{b^2\,\text{var}(X)}\sqrt{d^2\,\text{var}(Y)}} corr(X,Y)$.

16) The textbook formula for the variance of the discrete random variable *Y* is given as

$$\sigma_Y^2 = \sum_{i=1}^{k} (y_i - \mu_Y)^2 p_i \,.$$

Another commonly used formulation is

$$\sigma_Y^2 = \sum_{i=1}^{k} y_i^2 p_i - \mu_y^2 \,.$$

Prove that the two formulas are the same.

Answer: $\sigma_Y^2 = \sum_{i=1}^{k} (y_i - \mu_Y)^2 p_i = \sum_{i=1}^{k} (y_i^2 + \mu_Y^2 - 2\mu_Y y_i) p_i = \sum_{i=1}^{k} (y_i^2 p_i + \mu_Y^2 p_i - 2\mu_Y y_i p_i)$.

Moving the summation sign through results in

$\sigma_Y^2 = \sum_{i=1}^{k} y_i^2 p_i + \mu_Y^2 \sum_{i=1}^{k} p_i - 2\mu_Y \sum_{i=1}^{k} y_i p_i$. But $\sum_{i=1}^{k} p_i = 1$ and $\mu_Y = \sum_{i=1}^{k} y_i p_i$, giving you the second expression after simplification.

17) The *Economic Report of the President* gives the following age distribution of the United States population for the year 2000:

United States Population By Age Group, 2000

Outcome (age category)	Under 5	5-15	16-19	20-24	25-44	45-64	65 and over
Percentage	0.06	0.16	0.06	0.07	0.30	0.22	0.13

Imagine that every person was assigned a unique number between 1 and 275,372,000 (the total population in 2000). If you generated a random number, what would be the probability that you had drawn someone older than 65 or under 16? Treating the percentages as probabilities, write down the cumulative probability distribution. What is the probability of drawing someone who is 24 years or younger?

Answer: $\Pr(Y < 16 \text{ or } Y > 65) = 0.35$;

Outcome (age category)	Under 5	5-15	16-19	20-24	25-44	45-64	65 and over
Cumulative probability distribution	0.06	0.22	0.28	0.35	0.65	0.87	1.00

$\Pr(Y \leq 24) = 0.35.$

18) The accompanying table gives the outcomes and probability distribution of the number of times a student checks her e-mail daily:

Probability of Checking E-Mail

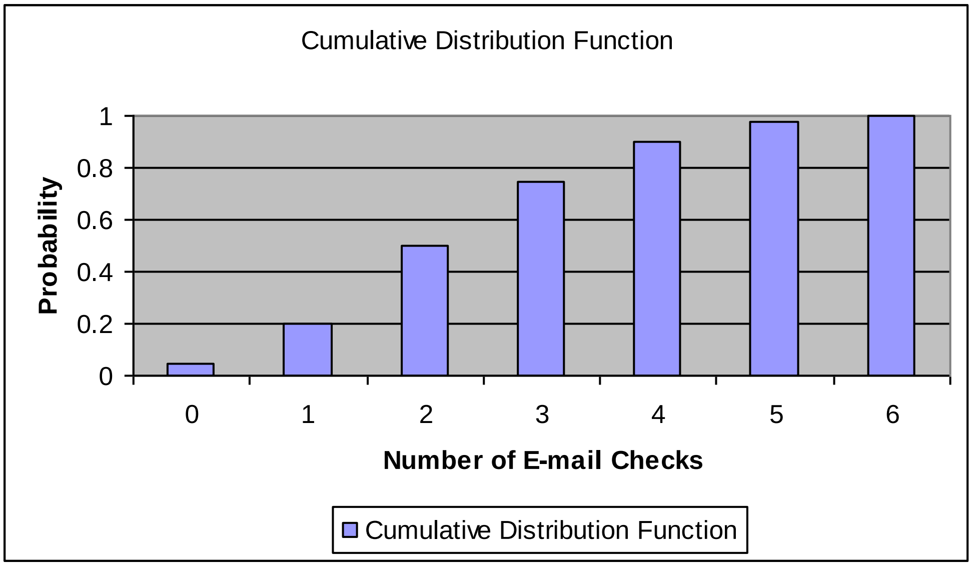
Outcome (number of e-mail checks)	0	1	2	3	4	5	6
Probability distribution	0.05	0.15	0.30	0.25	0.15	0.08	0.02

Sketch the probability distribution. Next, calculate the c.d.f. for the above table. What is the probability of her checking her e-mail between 1 and 3 times a day? Of checking it more than 3 times a day?

Answer:

Outcome (number of e-mail checks)	0	1	2	3	4	5	6
Cumulative probability distribution	0.05	0.20	0.50	0.75	0.90	0.98	1.00

$\Pr(1 \leq Y \leq 3) = 0.70$; $\Pr(Y > 0.25)$.



19) The accompanying table lists the outcomes and the cumulative probability distribution for a student renting videos during the week while on campus.

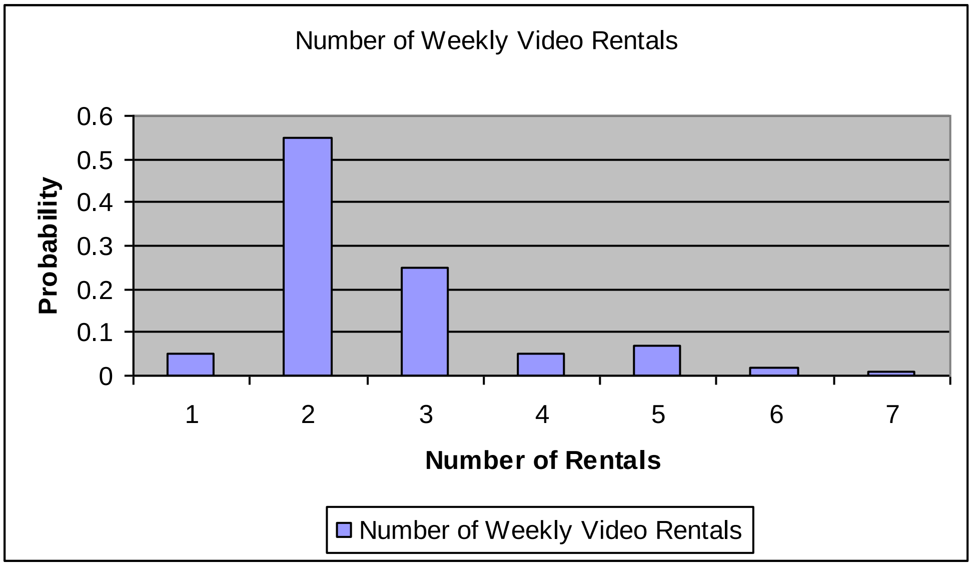
Video Rentals per Week during Semester

Outcome (number of weekly video rentals)	0	1	2	3	4	5	6
Probability distribution	0.05	0.55	0.25	0.05	0.07	0.02	0.01

Sketch the probability distribution. Next, calculate the cumulative probability distribution for the above table. What is the probability of the student renting between 2 and 4 a week? Of less than 3 a week?

Answer: The cumulative probability distribution is given below. The probability of renting between two and four videos a week is 0.37. The probability of renting less than three a week is 0.85.

Outcome (number of weekly video rentals)	0	1	2	3	4	5	6
Cumulative probability distribution	0.05	0.60	0.85	0.90	0.97	0.99	1.00



20) The textbook mentioned that the mean of $Y, E(Y)$ is called the first moment of Y, and that the expected value of the square of $Y, E(Y^2)$ is called the second moment of Y, and so on. These are also referred to as moments about the origin. A related concept is moments about the mean, which are defined as $E[(Y-\mu_Y)^r]$. What do you call the second moment about the mean? What do you think the third moment, referred to as "skewness," measures? Do you believe that it would be positive or negative for an earnings distribution? What measure of the third moment around the mean do you get for a normal distribution?

<u>Answer</u>: The second moment about the mean is the variance. Skewness measures the departure from symmetry. For the typical earnings distribution, it will be positive. For the normal distribution, it will be zero.

21) Explain why the two probabilities are identical for the standard normal distribution: $\Pr(-1.96 \leq X \leq 1.96)$ and $\Pr(-1.96 < X < 1.96)$.

<u>Answer</u>: For a continuous distribution, the probability of a point is zero.