**Chapter 5**

**Discrete Probability Distributions**

**Learning Objectives**

1. Understand the concepts of a random variable and a probability distribution.

2. Be able to distinguish between discrete and continuous random variables.

3. Be able to compute and interpret the expected value, variance, and standard deviation for a discrete random variable.

4. Be able to construct an empirical discrete distribution from available data.

5. Be able to compute the covariance and correlation coefficient for a bivariate empirical discrete distribution.

6. Be able to compute and work with probabilities involving a binomial probability distribution.

7. Be able to compute and work with probabilities involving a Poisson probability distribution.

8. Know when and how to use the hypergeometric probability distribution.

**Solutions:**

1. a. Head, Head (H,H)

Head, Tail (H,T)

Tail, Head (T,H)

Tail, Tail (T,T)

b. *x* = number of heads on two coin tosses

c.

|  |  |
| --- | --- |
| Outcome | Values of *x* |
| (H,H) | 2 |
| (H,T) | 1 |
| (T,H) | 1 |
| (T,T) | 0 |

d. Discrete. It may assume 3 values: 0, 1, and 2.

2. a. Let *x* = time (in minutes) to assemble the product.

b. It may assume any positive value: *x* > 0.

c. Continuous

3. Let Y = position is offered

N = position is not offered

a. *S* = {(Y,Y,Y), (Y,Y,N), (Y,N,Y), (N,Y,Y), (Y,N,N), (N,Y,N), (N,N,Y), (N,N,N)}

b. Let N = number of offers made; N is a discrete random variable.

c.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Experimental Outcome | (Y,Y,Y) | (Y,Y,N) | (Y,N,Y) | (N,Y,Y) | (Y,N,N) | (N,Y,N) | (N,N,Y) | (N,N,N) |
| Value of N | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 0 |

4. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

5. a. *S* = {(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)}

b.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Experimental Outcome | (1,1) | (1,2) | (1,3) | (2,1) | (2,2) | (2,3) |
| Number of Steps Required | 2 | 3 | 4 | 3 | 4 | 5 |

6. a. values: 0,1,2,...,20

discrete

b. values: 0,1,2,...

discrete

c. values: 0,1,2,...,50

discrete

d. values: 0  *x*  8

continuous

e. values: *x* > 0

continuous

7. a. *f* (*x*)  0 for all values of *x*.

 *f* (*x*) = 1 Therefore, it is a proper probability distribution.

b. Probability *x* = 30 is *f* (30) = .25

c. Probability *x*  25 is *f* (20) + *f* (25) = .20 + .15 = .35

d. Probability *x* > 30 is *f* (35) = .40

8. a.

|  |  |
| --- | --- |
| *x* | *f* (*x*) |
| 1 | 3/20 = .15 |
| 2 | 5/20 = .25 |
| 3 | 8/20 = .40 |
| 4 | 4/20 = .20 |
|  | Total 1.00 |

b.



c. *f* (*x*)  0 for *x* = 1,2,3,4.

 *f* (*x*) = 1

9. a. There are a total of 26,975 unemployed persons in the data set. Each probability *f*(*x*) is computed by dividing the number of months of unemployment by 26,975. For example, *f* (1) = 1029/26,975 = .0381. The complete probability distribution is as follows.

|  |  |
| --- | --- |
| *x* | *f* (*x*) |
| 1 | .0381 |
| 2 | .0625 |
| 3 | .0841 |
| 4 | .0992 |
| 5 | .1293 |
| 6 | .1725 |
| 7 | .1537 |
| 8 | .1330 |
| 9 | .0862 |
| 10 | .0415 |

b. 

c. Probability 2 months or less = *f* (1) + *f* (2) = .0381 + .0625 = .1006

Probability more than 2 months = 1 - .1006 = .8994

d. Probability more than 6 months = *f* (7) + *f* (8) + *f* (9) + *f* (10) = .1537 + .1330 + .0862 + .0415 = .4144

10. a.

|  |  |
| --- | --- |
| *x* | *f* (*x*) |
| 1 | 0.05 |
| 2 | 0.09 |
| 3 | 0.03 |
| 4 | 0.42 |
| 5 | 0.41 |
|  | 1.00 |

b.

|  |  |
| --- | --- |
| *x* | *f* (*x*) |
| 1 | 0.04 |
| 2 | 0.10 |
| 3 | 0.12 |
| 4 | 0.46 |
| 5 | 0.28 |
|  | 1.00 |

c. *P*(4 or 5) = *f* (4) + *f* (5) = 0.42 + 0.41 = 0.83

d. Probability of very satisfied: 0.28

e. Senior executives appear to be more satisfied than middle managers. 83% of senior executives have a score of 4 or 5 with 41% reporting a 5. Only 28% of middle managers report being very satisfied.

11. a.

|  |  |
| --- | --- |
| Duration of Call |  |
| *x* | *f* (*x*) |
| 1 | 0.25 |
| 2 | 0.25 |
| 3 | 0.25 |
| 4 | 0.25 |
|  | 1.00 |

b.



c. *f* (*x*)  0 and *f* (1) + *f* (2) + *f* (3) + *f* (4) = 0.25 + 0.25 + 0.25 + 0.25 = 1.00

d. *f* (3) = 0.25

e. *P*(overtime) = *f* (3) + *f* (4) = 0.25 + 0.25 = 0.50

12. a. Yes; *f* (*x*)  0.  *f* (*x*) = 1

b. *f* (500,000) + *f* (600,000) = .10 + .05 = .15

c. *f* (100,000) = .10

13. a. Yes, since *f* (*x*)  0 for *x* = 1,2,3 and  *f* (*x*) = *f* (1) + *f* (2) + *f* (3) = 1/6 + 2/6 + 3/6 = 1

b. *f* (2) = 2/6 = .333

c. *f* (2) + *f* (3) = 2/6 + 3/6 = .833

14. a. *f* (200) = 1 - *f* (-100) - *f* (0) - *f* (50) - *f* (100) - *f* (150)

= 1 - .95 = .05

This is the probability MRA will have a $200,000 profit.

b. *P*(Profit) = *f* (50) + *f* (100) + *f* (150) + *f* (200)

= .30 + .25 + .10 + .05 = .70

c. *P*(at least 100) = *f* (100) + *f* (150) + *f* (200)

= .25 + .10 +.05 = .40

15. a.

|  |  |  |
| --- | --- | --- |
| *x* | *f* (*x*) | *x f* (*x*) |
| 3 | .25 | .75 |
| 6 | .50 | 3.00 |
| 9 | .25 | 2.25 |
|  | 1.00 | 6.00 |

*E*(*x*) = ** = 6

b.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *x* | *x - * | (*x - *)2 | *f* (*x*) | (*x - *)2 *f* (*x*) |
| 3 | -3 | 9 | .25 | 2.25 |
| 6 | 0 | 0 | .50 | 0.00 |
| 9 | 3 | 9 | .25 | 2.25 |
|  |  |  |  | 4.50 |

*Var*(*x*) = **2 = 4.5

c. ** =  = 2.12

16. a.

|  |  |  |
| --- | --- | --- |
| *y* | *f* (*y*) | *y f* (*y*) |
| 2 | .2 | .4 |
| 4 | .3 | 1.2 |
| 7 | .4 | 2.8 |
| 8 | .1 | .8 |
|  | 1.0 | 5.2 |

*E*(*y*) = ** = 5.2

b.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *y* | *y - * | (*y - *)2 | *f* (*y*) | (*y - *)2 *f* (*y*) |
| 2 | -3.20 | 10.24 | .20 | 2.048 |
| 4 | -1.20 | 1.44 | .30 | .432 |
| 7 | 1.80 | 3.24 | .40 | 1.296 |
| 8 | 2.80 | 7.84 | .10 | .784 |
|  |  |  |  | 4.560 |



17. a. Total Student = 1,518,859

*x* = 1 *f*(1) = 721,769/1,518,859 = .4752

*x* = 2 *f*(2) = 601,325/1,518,859 = .3959

*x* = 3 *f*(3) = 166,736/1,518,859 = .1098

*x* = 4 *f*(4) = 22,299/1,518,859 = .0147

*x* = 5 *f*(5) = 6730/1,518,859 = .0044

b. *P*(*x* > 1) = 1 – *f*(1) = 1 - .4752 = .5248

Over 50% of the students take the SAT more than 1 time.

c. *P*(*x* >3) = *f*(3) + *f*(4) + *f*(5) = .1098 + .0147 + .0044 = .1289

d./e.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f* (*x*) | *x f* (*x*) | *x* - ** | (*x* - **)2 | (*x* - **)2 *f* (*x*) |
| 1 | .4752 | .4752 | -.6772 | .4586 | .2179 |
| 2 | .3959 | .7918 | .3228 | .1042 | .0412 |
| 3 | .1098 | .3293 | 1.3228 | 1.7497 | .1921 |
| 4 | .0147 | .0587 | 2.3228 | 5.3953 | .0792 |
| 5 | .0044 | .0222 | 3.3228 | 11.0408 | .0489 |
|  |  | 1.6772 |  |  | .5794 |

*E*(*x*) = Σ *x f*(*x*) = 1.6772

The mean number of times a student takes the SAT is 1.6772, or approximately

1.7 times.





18. a/b.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f* (*x*) | *xf* (*x*) | *x* - ** | (*x* - **)2 | (*x* - **)2 *f* (*x*) |
| 0 | 0.04 | 0.00 | -1.84 | 3.39 | 0.12 |
| 1 | 0.34 | 0.34 | -0.84 | 0.71 | 0.24 |
| 2 | 0.41 | 0.82 | 0.16 | 0.02 | 0.01 |
| 3 | 0.18 | 0.53 | 1.16 | 1.34 | 0.24 |
| 4 | 0.04 | 0.15 | 2.16 | 4.66 | 0.17 |
| Total | 1.00 | 1.84 |  |  | 0.79 |
|  |  |  |  |  |  |
|  |  | *E*(*x*) |  |  | *Var*(*x*) |

c/d.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *y* | *f* (*y*) | *yf* (*y*) | *y* - ** | (*y* - **)2 | (*y* - **)2 *f* (*y*) |
| 0 | 0.00 | 0.00 | -2.93 | 8.58 | 0.01 |
| 1 | 0.03 | 0.03 | -1.93 | 3.72 | 0.12 |
| 2 | 0.23 | 0.45 | -0.93 | 0.86 | 0.20 |
| 3 | 0.52 | 1.55 | 0.07 | 0.01 | 0.00 |
| 4 | 0.22 | 0.90 | 1.07 | 1.15 | 0.26 |
| Total | 1.00 | 2.93 |  |  | 0.59 |
|  |  |  |  |  |  |
|  |  | *E*(*y*) |  |  | *Var*(*y*) |

e. The number of bedrooms in owner-occupied houses is greater than in renter-occupied houses. The expected number of bedrooms is 2.93 - 1.84 = 1.09 greater. And, the variability in the number of bedrooms is less for the owner-occupied houses.

19. a. *E*(*x*) =  *x f* (*x*) = 0 (.56) + 2 (.44) = .88

b. *E*(*x*) =  *x f* (*x*) = 0 (.66) + 3 (.34) = 1.02

c. The expected value of a 3 - point shot is higher. So, if these probabilities hold up, the team will make more points in the long run with the 3 - point shot.

20. a.

|  |  |  |
| --- | --- | --- |
| *x* | *f (x)* | *xf (x)* |
| 0 | .85 | 0 |
| 500 | .04 | 20 |
| 1000 | .04 | 40 |
| 3000 | .03 | 90 |
| 5000 | .02 | 100 |
| 8000 | .01 | 80 |
| 10000 | .01 | 100 |
| Total | 1.00 | 430 |

The expected value of the insurance claim is $430. If the company charges $430 for this type of collision coverage, it would break even.

b. From the point of view of the policyholder, the expected gain is as follows:

Expected Gain = Expected claim payout – Cost of insurance coverage

= $430 - $520 = -$90

The policyholder is concerned that an accident will result in a big repair bill if there is no insurance coverage. So even though the policyholder has an expected annual loss of $90, the insurance is protecting against a large loss.

21. a. *E*(*x*) =  *x f* (*x*) = 0.05(1) + 0.09(2) + 0.03(3) + 0.42(4) + 0.41(5) = 4.05

b. *E*(*x*) =  *x f* (*x*) = 0.04(1) + 0.10(2) + 0.12(3) + 0.46(4) + 0.28(5) = 3.84

c. Executives: **2 =  (*x* - **)2 *f*(*x*) = 1.25

Middle Managers: **2 =  (*x* - **)2 *f*(*x*) = 1.13

d. Executives: ** = 1.12

Middle Managers: ** = 1.07

e. The senior executives have a higher average score: 4.05 vs. 3.84 for the middle managers. The executives also have a slightly higher standard deviation.

22. a. *E*(*x*) =  *x f* (*x*) = 300 (.20) + 400 (.30) + 500 (.35) + 600 (.15) = 445

The monthly order quantity should be 445 units.

b. Cost: 445 @ $50 = $22,250

Revenue: 300 @ $70 = 21,000

$ 1,250 Loss

23. a. Rent Controlled: *E*(*x*) = 1(.61) + 2(.27) + 3(.07) + 4(.04) + 5(.01) = 1.57

Rent Stabilized: *E*(*x*) = 1(.41) + 2(.30) + 3(.14) + 4(.11) + 5(.03) + 6(.01) = 2.08

b. Rent Controlled:

*Var*(*x*) = (-.57)2.61+ (.43)2.27+ (1.43)2.07+ (2.43)2.04+ (3.43)2.01= .75

Rent Stabilized:

*Var*(*x*) = (-1.08)2.41+ (-.08)2.30+ (.92)2.14+ (1.92)2.11+ (2.92)2.03+ (3.92)2.01= 1.41

c. From the expected values in part (a), it is clear that the expected number of persons living in rent stabilized units is greater than the number of persons living in rent controlled units. For example, comparing a building that contained 10 rent controlled units to a building that contained 10 rent stabilized units, the expected number of persons living in the rent controlled building would be 1.57(10) = 15.7 or approximately 16. For the rent stabilized building, the expected number of persons is approximately 21. There is also more variability in the number of persons living in rent stabilized units.

24. a. Medium *E*(*x*) =  *x f* (*x*)

= 50 (.20) + 150 (.50) + 200 (.30) = 145

Large: *E*(*x*) =  *x f* (*x*)

= 0 (.20) + 100 (.50) + 300 (.30) = 140

Medium preferred.

b. Medium

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *x* | *f* (*x*) | *x* - ** | (*x* - **)2 | (*x* - **)2 *f (x*) |
| 50 | .20 | -95 | 9025 | 1805.0 |
| 150 | .50 | 5 | 25 | 12.5 |
| 200 | .30 | 55 | 3025 | 907.5 |

**2 = 2725.0

Large

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *y* | *f* (*y*) |  | *y* - ** | (*y* - **)2 | (*y* - **)2 *f (y*) |
| 0 | .20 |  | -140 | 19600 | 3920 |
| 100 | .50 |  | -40 | 1600 | 800 |
| 300 | .30 |  | 160 | 25600 | 7680 |

**2 = 12,400

Medium preferred due to less variance.

25. a. 







b.

|  |  |
| --- | --- |
| *x + y* | *f (x +y)* |
| 130 | .2 |
| 80 | .5 |
| 100 | .3 |

c.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x + y* | *f(x +y)* | *(x + y)f(x + y)* | *x + y – E(x + y)* |  |  |
| 130 | .2 | 26 | 34 | 1156 | 231.2 |
| 80 | .5 | 40 | -16 | 256 | 128.0 |
| 100 | .3 | 30 | 4 | 16 | 4.8 |
| *E(x + y) =* | | 96 |  | *Var(x + y) =* | 364 |

d. 

*Var(x)* = 61 and *Var(y)* =129 were computed in part (a), so

 



The random variables *x* and *y* are positively related. Both the covariance and correlation coefficient are positive. Indeed, they are very highly correlated; the correlation coefficient is almost equal to 1.

e. 

*Var(x) + Var(y) =* 61 + 129 = 190

The variance of the sum of *x* and *y* is greater than the sum of the variances by two times the covariance: 2(87) = 174. The reason it is positive is that, in this case the variables are positively related. Whenever two random variables are positively related, the variance of the sum of the randomly variables will be greater than the sum of the variances of the individual random variables.

26. a. The standard deviation for these two stocks is the square root of the variance.

 

Investments in Stock 1 would be considered riskier than investments in Stock 2 because the standard deviation is higher. Note that if the return for Stock 1 falls 8.45/5 = 1.69 or more standard deviation below its expected value, an investor in that stock will experience a loss. The return for Stock 2 would have to fall 3.2 standard deviations below its expected value before an investor in that stock would experience a loss.

b. Since *x* represents the percent return for investing in Stock 1, the expected return for investing $100 in Stock 1 is $8.45 and the standard deviation is $5.00. So to get the expected return and standard deviation for a $500 investment we just multiply by 5.

Expected return ($500 investment) = 5($8.45) = $42.25

Standard deviation ($500 investment) = 5($5.00) = $25.00

c. Since *x* represents the percent return for investing in Stock 1 and *y* represents the percent return for investing in Stock 2, we want to compute the expected value and variance for .5*x* + .5*y*.

*E*(.5*x* + .5*y*) = .5*E*(*x*) + .5*E*(*y*) = .5(8.45) + .5(3.2) = 4.225 + 1.6 = 5.825





d. Since *x* represents the percent return for investing in Stock 1 and *y* represents the percent return for investing in Stock 2, we want to compute the expected value and variance for .7*x* + .3*y*.

*E*(.7*x* + .3*y*) = .7*E*(*x*) + .3*E*(*y*) = .7(8.45) + .3(3.2)=5.915 + .96 = 6.875





e. The standard deviations of *x* and *y* were computed in part (a). The correlation coefficient is given by



There is a fairly strong negative relationship between the variables.

27. a. Dividing each of the frequencies in the table by the total number of restaurants provides the joint probability table below. The bivariate probability for each pair of quality and meal price is shown in the body of the table. This is the bivariate probability distribution. For instance, the probability of a rating of 2 on quality and a rating of 3 on meal price is given by *f(*2, 3) = .18. The marginal probability distribution for quality, *x,* is in the rightmost column. The marginal probability for meal price, *y*, is in the bottom row.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Meal Price (*y)* | | |  |
| Quality (*x)* | 1 | 2 | 3 | Total |
| 1 | 0.14 | 0.13 | 0.01 | 0.28 |
| 2 | 0.11 | 0.21 | 0.18 | 0.50 |
| 3 | 0.01 | 0.05 | 0.16 | 0.22 |
| Total | 0.26 | 0.39 | 0.35 | 1 |

b. *E(x)* = 1(.28) + 2(.50) + 3(.22) = 1.94

*Var(x)* = .28(1 – 1.94)2 + .50(2- 1.94)2 + .22(3 – 1.94)2 = .4964

c. *E(y)* = 1(.26) + 2(.39) + 3(.35) = 2.09

*Var(y)* = .26(1 – 2.09)2 + .39(2 – 2.09)2 + .35(3 – 2.09)2 = .6019

d. 

Since, the covariance is positive we can conclude that as the quality rating goes up, the meal price goes up. This is as we would expect.

e. 

With a correlation coefficient of .5221 we would call this a moderately positive relationship. It is not likely to find a low cost restaurant that is also high quality. But, it is possible. There are 3 of them leading to 

28. a. Marginal distribution of Direct Labor Cost

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *y* | *f (y)* | *yf (y)* | *y - E(y)* | *(y-E(y))2* | *(y-E(y))2f (y)* |
| 43 | .3 | 12.9 | -2.3 | 5.29 | 1.587 |
| 45 | .4 | 18 | -.3 | .09 | .036 |
| 48 | .3 | 14.4 | 2.7 | 7.29 | 2.187 |
|  |  | 45.3 |  | *Var(y)=* | 3.81 |
|  |  | *E(y)* = 45.3 |  | *=* | 1.95 |

b. Marginal distribution of Parts Cost

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f (x)* | *xf (x)* | *x - E(x)* | *(x-E(x))2* | *(x-E(x))2f (x)* |
| 85 | .45 | 38.25 | -5.5 | 30.25 | 13.6125 |
| 95 | .55 | 52.25 | 4.5 | 20.25 | 11.1375 |
|  |  | 90.5 |  | *Var(x)=* | 24.75 |
|  |  | *E(x)* = 90.5 |  | *=* | 4.97 |

c. Let *z = x + y* represent total manufacturing cost (direct labor + parts).

|  |  |
| --- | --- |
| *z* | *f (z)* |
| 128 | .05 |
| 130 | .20 |
| 133 | .20 |
| 138 | .25 |
| 140 | .20 |
| 143 | .10 |
|  | 1.00 |

d. The computation of the expected value, variance, and standard deviation of total manufacturing cost is shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *z* | *f (z)* | *zf (z)* | *z - E(z)* | *(z-E(z))2* | *(z-E(z))2f(z)* |
| 128 | .05 | 6.4 | -7.8 | 60.84 | 3.042 |
| 130 | .20 | 26 | -5.8 | 33.64 | 6.728 |
| 133 | .20 | 26.6 | -2.8 | 7.84 | 1.568 |
| 138 | .25 | 34.5 | 2.2 | 4.84 | 1.21 |
| 140 | .20 | 28 | 4.2 | 17.64 | 3.528 |
| 143 | .10 | 14.3 | 7.2 | 51.84 | 5.184 |
|  |  | 135.8 |  | *Var(z)=* | 21.26 |
|  |  | *E(z)* = 135.8 |  | *=* | 4.61 |

e. To determine if *x* = parts cost and *y* = direct labor cost are independent, we need to compute the covariance .



Since the covariance is not equal to zero, we can conclude that direct labor cost is not independent of parts cost. Indeed, they are negatively correlated. When parts cost goes up, direct labor cost goes down. Maybe the parts costing $95 come from a different manufacturer and are higher quality. Working with higher quality parts may reduce labor costs.

f. The expected manufacturing cost for 1500 printers is



The total manufacturing costs of $198,350 are less than we would have expected. Perhaps as more printers were manufactured there was a learning curve and direct labor costs went down.

29. a. Let *x* = percentage return for S&P 500

*y* = percentage return for Core Bond fund

The formula for computing the correlation coefficient is given by



In this case, we know the correlation coefficient and both standard deviations, so we want to rearrange this formula to find the covariance.



b. Letting *r* = portfolio percentage return, we have *r* = .5*x* + .5*y*. The expected return for a portfolio with 50% invested in the S&P 500 and 50% invested in Core Bonds is



We are given and , so and . We can now compute



Then 

So, the expected return for our portfolio is 5.41% and the standard deviation is 9.78%.

c. Letting *r* = portfolio percentage return, we have *r* = .2*x* + .8*y*. The expected return for a portfolio with 20% invested in the S&P 500 and 80% invested in Core Bonds is



We are given and , so and . We can now compute



Then 

So, the expected return for our portfolio is 5.63% and the standard deviation is 3.71%.

d. Letting *r* = portfolio percentage return, we have *r* = .8*x* + .2*y*. The expected return for a portfolio with 80% invested in the S&P 500 and 20% invested in Core Bonds is



We are given and , soand .

We can now compute



Then 

So, the expected return for our portfolio is 5.19% and the standard deviation is 15.43%.

e. The portfolio in part (c), investing 20% in the S&P500 fund and 80% in the Core Bond fund has the largest expected return: 5.63%.

The portfolio in part (c), investing 20% in the S&P 500 fund and 80% in the Core Bond fund also has the smallest standard deviation: 3.71%.

Since the portfolio in part (c) has the highest expected return and the smallest standard deviation (our measure of risk) it is the preferred investment choice.

30. a. Let *x* = percentage return for S&P 500

*y* = percentage return for Core Bond fund

*z* = percentage return for REITs

The formula for computing the correlation coefficient is given by



In this case, we know the correlation coefficients and the 3 standard deviations, so we want to rearrange the correlation coefficient formula to find the covariances.

S&P 500 and REITs: 

Core Bonds and REITS: 

b. Letting *r* = portfolio percentage return, we have *r* = .5*x* + .5*y*. The expected return for a portfolio with 50% invested in the S&P 500 and 50% invested in REITs is



We are given and , soand . We can now compute



Then 

So, the expected return for our portfolio is 9.055% and the standard deviation is 19.89%.

c. Letting *r* = portfolio percentage return, we have *r* = .5*y* + .5*z*. The expected return for a portfolio with 50% invested in Core Bonds and 50% invested in REITs is



We are given and , soand . We can now compute



Then 

So, the expected return for our portfolio is 9.425% and the standard deviation is 11.63%.

d. Letting *r* = portfolio percentage return, we have *r* = .8*y* + .2*z*. The expected return for a portfolio with 80% invested in Core Bonds and 20% invested in REITs is



From part (c) above, we have and . We can now compute



Then 

So, the expected return for our portfolio is 7.238% and the standard deviation is 4.94%.

e. The expected returns and standard deviations for the 3 portfolios are summarized below.

|  |  |  |
| --- | --- | --- |
| **Portfolio** | **Expected Return (%)** | **Standard Deviation** |
| 50% S&P 500 & 50% REITs | 9.055 | 19.89 |
| 50% Core Bonds & 50% REITs | 9.425 | 11.63 |
| 80% Core Bonds & 20% REITs | 7.238 | 4.94 |

The portfolio from part (c) involving 50% Core Bonds and 50% REITS has the highest return. Using the standard deviation as a measure of risk, it also has less risk than the portfolio from part (b) involving 50% invested in an S&P 500 index fund and 50% invested in REITs. So the portfolio from part (b) would not be recommended for either type of investor.

The portfolio from part (d) involving 80% in Core Bonds and 20% in REITs has the lowest standard deviation and thus lesser risk than the portfolio in part (c). We would recommend the portfolio consisting of 50% Core Bonds and 50% REITs for the aggressive investor because of its higher return and moderate amount of risk.

We would recommend the portfolio consisting of 80% Core Bonds and 20% REITS to the conservative investor because of its low risk and moderate return.

31. a.



b. 

c. 

d. 

e. *P*(*x*  1) = *f* (1) + *f* (2) = .48 + .16 = .64

f. *E*(*x*) = *n p*  = 2 (.4) = .8

*Var*(*x*) = *n p*  (1 - *p*) = 2 (.4) (.6) = .48

** =  = .6928

32. a. *f* (0) = .3487

b. *f* (2) = .1937

c. *P*(*x*  2) = *f* (0) + *f* (1) + *f* (2) = .3487 + .3874 + .1937 = .9298

d. *P*(*x*  1) = 1 - *f* (0) = 1 - .3487 = .6513

e. *E*(*x*) = *n p*  = 10 (.1) = 1

f. *Var*(*x*) = *n p*  (1 - *p*) = 10 (.1) (.9) = .9

** = = .95

33. a. *f* (12) = .1144

b. *f* (16) = .1304

c. *P*(*x*  16) = *f* (16) + *f* (17) + *f* (18) + *f* (19) + *f* (20)

= .1304 + .0716 + .0278 + .0068 + .0008 = .2374

d. *P*(*x*  15) = 1 - *P* (*x*  16) = 1 - .2374 = .7626

e. *E*(*x*) = *n p* = 20(.7) = 14

f. *Var*(*x*) = *n p* (1 - *p*) = 20 (.7) (.3) = 4.2

** = = 2.0494

34. a. 

b. *P*(at least 2) = 1 - *f*(0) - *f*(1)

=

= 1 - .2084 - .3735 = .4181

c. 

35. a. 





b. *P*(*x*  3) = 1 - *f* (0) - *f* (1) - *f* (2)







*P*(*x*  3) = 1 - .0282 - .1211 - .2335 = .6172

36. a. Probability of a defective part being produced must be .03 for each part selected; parts must be selected independently.

b. Let: D = defective

G = not defective



c. 2 outcomes result in exactly one defect.

d. *P*(no defects) = (.97) (.97) = .9409

P (1 defect) = 2 (.03) (.97) = .0582

P (2 defects) = (.03) (.03) = .0009

37. a. Yes. Since the employees are selected randomly, *p* is the same from trial to trial and the trials are independent. The two outcomes per trial are loyal and not loyal.

Binomial *n* = 10 and *p* = .25



b. 

c. 

d. Probability (*x* > 2) = 1 - *f* (0) - *f* (1)

From part (b), *f*(0) = .0563



Probability (*x* > 2) = 1 - *f* (0) - *f* (1) = 1 - (.0563 + .1877 ) = .7560

38. a. .90

b. *P*(at least 1) = *f* (1) + *f* (2)



*P*(at least 1) = .18 + .81 = .99

Alternatively

*P*(at least 1) = 1 – *f* (0)



Therefore, *P*(at least 1) = 1 - .01 = .99

c. *P*(at least 1) = 1 - *f* (0)



Therefore, *P*(at least 1) = 1 - .001 = .999

d. Yes; *P*(at least 1) becomes very close to 1 with multiple systems and the inability to detect an attack would be catastrophic.

39. a. Using the 20 golfers in the Hazeltine PGA Championship, the probability that a PGA professional golfer uses a Titleist brand golf ball is *p* = 14/20 = .6

For the sample of 15 PGA Tour players, use a binomial distribution with *n* = 15 and *p* = .6.

*f* (10) ==.1859

Or, using the binomial tables, *f* (10) = .1859

b. *P*(*x* > 10) = *f* (11) + *f* (12) + *f* (13) + *f* (14) + *f* (15)

Using the binomial tables, we have

.1268 + .0634 + .0219 + .0047 + .0005 = .2173

c. *E*(*x*) = *np* = 15(.6) = 9

d. *Var*(*x*) = **2 = *np*(1 - *p*) = 15(.6)(1 - .6) = 3.6

** = = 1.8974

40. a. 





b. *P*(*x*  3) = 1 - *f* (0) - *f* (1) - *f* (2)







*P*(*x*  3) = 1 - .0072 - .0423 - .1150 = .8355

41. a. *f* (0) + *f* (1) + *f* (2) = .0115 + .0576 + .1369 = .2060

b. *f* (4) = .2182

c. 1 - [ *f* (0) + *f* (1) + *f* (2) + *f* (3) ] = 1 - .2060 - .2054 = .5886

d. ** = *n p* = 20 (.20) = 4

42. a. *p* = ¼ = .25







b. *P*(*x*  2) = 1 – *f*(0) – *f*(1)

Using the binomial tables *f*(0) = .0032 and *f*(1) = .0211

*P*(*x*  2) = 1 – .0032 - .0211 = .9757

c. Using the binomial tables *f*(12) = .0008

And, with *f* (13) = .0002, *f* (14) = .0000, and so on, the probability of finding that 12 or more investors have exchange-traded funds in their portfolio is so small that it is highly unlikely that *p* = .25. In such a case, we would doubt the accuracy of the results and conclude that *p* must be greater than .25.

d. ** = *n p* = 20 (.25) = 5

43. *E*(*x*) = *n p* = 35(.23) = 8.05 (8 automobiles)

*Var*(*x*) = *n p* (1 - *p*) = 35(.23)(1-.23) = 6.2

** = = 2.49

44. a. 

b. 

c. 

d. *P*(*x*  2) = 1 - *f* (0) - *f* (1) = 1 - .0498 - .1494 = .8008

45. a. 

b. ** = 6 for 3 time periods

c. 

d. 

e. 

f. 

46. a. ** = 48 (5/60) = 4



b. ** = 48 (15 / 60) = 12



c. ** = 48 (5 / 60) = 4 I expect 4 callers to be waiting after 5 minutes.



The probability none will be waiting after 5 minutes is .0183.

d. ** = 48 (3 / 60) = 2.4



The probability of no interruptions in 3 minutes is .0907.

47. a. 30 per hour

b. ** = 1 (5/2) = 5/2



c. 

48. a. 

b. probability = 1 - [*f*(0) + *f*(1)]



probability = 1 - [.0009 + .0064] = .9927

c. *μ* = 3.5



probability = 1 - *f*(0) = 1 - .0302 = .9698

d. probability = 1 - [*f*(0) + *f*(1) + *f*(2) + *f*(3) + *f*(4)]

= 1 - [.0009 + .0064 + .0223 + .0521 + .0912] = .8271

Note: The Poisson tables were used to compute the Poisson probabilities *f*(0), *f*(1), *f*(2), *f*(3) and *f*(4) in part (d).

49. a. 

b. *f* (0) + *f* (1) + *f* (2) + *f* (3)

*f* (0) = .000045 (part a)



Similarly, *f* (2) = .002267, *f* (3) = .007567

and *f* (0) + *f* (1) + *f* (2) + *f* (3) = .010329

c. 2.5 arrivals / 15 sec. period Use ** = 2.5



d. 1 - *f* (0) = 1 - .0821 = .9179

50. Poisson distribution applies

a. ** = 1.25 per month

b. 

c. 

d. *P*(More than 1) = 1 - *f* (0) - *f* (1) = 1 - 0.2865 - 0.3581 = .3554

51. a. 



b. *P*(*x*  2) = 1 - *f* (0) - *f* (1)



*P*(*x*  2) = 1 - .0498 - .1494 = .8008

c. *µ* = 3 per year

*µ* = 3/2 = 1.5 per 6 months

d. 

52. a. 

b. 

c. 

d. 

e. Note *x* = 4 is *greater than* *r* = 3. It is not possible to have *x*  = 4 successes when there are only 3 successes in the population. Thus, *f*(4) = 0. In this exercise, *n* is greater than *r*. Thus, the number of successes *x* can only take on values up to and including *r* = 3. Thus, *x* = 0, 1, 2, 3.

53. 

54. Hypergeometric Distribution with *N* = 10 and *r* = 7

a. = .5250

b. Compute the probability that 3 prefer football.



*P*(majority prefer football) = *f* (2) + *f* (3) = .5250 + .2917 = .8167

55. Parts a, b & c involve the hypergeometric distribution with *N* = 52 and *n* = 2

a. *r* = 20, *x* = 2



b. *r* = 4, *x* = 2



c. *r* = 16, *x* = 2



d. Part (a) provides the probability of blackjack plus the probability of 2 aces plus the probability of two 10s. To find the probability of blackjack we subtract the probabilities in (b) and (c) from the probability in (a).

*P*(blackjack) = .1433 - .0045 - .0905 = .0483

56. *N* = 60 *n* = 10

a. *r* = 20 *x* = 0



=  .0112

b. *r* = 20 *x* = 1

.0725

c. 1 - *f* (0) - *f* (1) = 1 - .0112 - .0725 = .9163

d. Same as the probability one will be from Hawaii; .0725.

57. a. 

b. 

c. 

d. 

58. Hypergeometric with *N* = 10 and *r* = 3.

a. *n* = 3, *x* = 0



This is the probability there will be no banks with increased lending in the study.

b. *n* = 3, *x* = 3



This is the probability there all three banks with increased lending will be in the study. This has a

very low probability of happening.

c. *n* = 3, *x* = 1



*n* = 3, *x* = 2



|  |  |
| --- | --- |
| *x* | *f*(*x*) |
| 0 | 0.2917 |
| 1 | 0.5250 |
| 2 | 0.1750 |
| 3 | 0.0083 |
| Total | 1.0000 |

*f*(1) = .5250 has the highest probability showing that there is over a .50 chance that there will be exactly one bank that had increased lending in the study.

d. *P*(*x* > 1) = 

There is a reasonably high probability of .7083 that there will be at least one bank that had increased lending in the study.

e. 



59. a/b/c.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f* (*x*) | *xf* (*x*) | *x* - ** | (*x* - **)2 | (*x* - **)2 *f* (*x*) |
| 1 | 0.07 | 0.07 | -2.12 | 4.49 | 0.31 |
| 2 | 0.21 | 0.42 | -1.12 | 1.25 | 0.26 |
| 3 | 0.29 | 0.87 | -0.12 | 0.01 | 0.00 |
| 4 | 0.39 | 1.56 | 0.88 | 0.77 | 0.30 |
| 5 | 0.04 | 0.20 | 1.88 | 3.53 | 0.14 |
| Total | 1.00 | 3.12 |  |  | 1.03 |
|  |  |  |  |  |  |
|  |  | *E*(*x*) |  |  | *Var*(*x*) |

** =  = 1.01

d. The expected level of optimism is 3.12. This is a bit above neutral and indicates that investment managers are somewhat optimistic. Their attitudes are centered between neutral and bullish with the consensus being closer to neutral.

60. a/b.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f* (*x*) | *xf* (*x*) | *x* - ** | (*x* - **)2 | (*x* - **)2 *f* (*x*) |
| 1 | 0.24 | 0.24 | -2.00 | 4.00 | 0.97 |
| 2 | 0.21 | 0.41 | -1.00 | 1.00 | 0.21 |
| 3 | 0.10 | 0.31 | 0.00 | 0.00 | 0.00 |
| 4 | 0.21 | 0.83 | 1.00 | 1.00 | 0.21 |
| 5 | 0.24 | 1.21 | 2.00 | 4.00 | 0.97 |
| Total | 1.00 | 3.00 |  |  | 2.34 |
|  |  |  |  |  |  |
|  |  | *E*(*x*) |  |  | *Var*(*x*) |

c. For the bond fund categories: *E*(*x*) = 1.36 *Var*(*x*) = .23

For the stock fund categories: *E*(*x*) = 4 *Var*(*x*) = 1.00

The total risk of the stock funds is much higher than for the bond funds. It makes sense to analyze these separately. When you do the variances for both groups (stocks and bonds), they are reduced.

61. a.

|  |  |
| --- | --- |
| *x* | *f* (*x*) |
| 9 | .30 |
| 10 | .20 |
| 11 | .25 |
| 12 | .05 |
| 13 | .20 |

b. *E*(*x*) = *x f* (*x*)

= 9(.30) + 10(.20) + 11(.25) + 12(.05) + 13(.20) = 10.65

Expected value of expenses: $10.65 million

c. *Var*(*x*) = (*x* - **)2 *f* (*x*)

= (9 - 10.65)2 (.30) + (10 - 10.65)2 (.20) + (11 - 10.65)2 (.25)

+ (12 - 10.65)2 (.05) + (13 - 10.65)2 (.20) = 2.13

d. Looks Good: *E*(Profit) = 12 - 10.65 = 1.35 million

However, there is a .20 probability that expenses will equal $13 million and the college will run a deficit.

62. a. There are 600 observations involving the two variables. Dividing the entries in the table shown by 600 and summing the rows and columns we obtain the following.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Reading Material (*y*) | | |  |
| Snacks (*x)* | 0 | 1 | 2 | Total |
| 0 | .0 | .1 | .03 | .13 |
| 1 | .4 | .15 | .05 | .6 |
| 2 | .2 | .05 | .02 | .27 |
| Total | .6 | .3 | .1 | 1 |

The entries in the body of the table are the bivariate or joint probabilities for *x* and *y*. The entries in the right most (Total) column are the marginal probabilities for *x* and the entries in the bottom (Total) row are the marginal probabilities for *y*.

The probability of a customer purchasing 1 item of reading materials and 2 snack items is given by *f*( *x* = 1, *y* = 2) =.05.

The probability of a customer purchasing 1 snack item only is given by *f*(*x* = 1, *y* = 0) = .40.

The probability *f*(*x* = 0, *y* = 0) = 0 because the point of sale terminal is only used when someone makes a purchase.

b. The marginal probability distribution of *x* along with the calculation of the expected value and variance is shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | *f (x)* | *xf (x)* | *x - E(x)* | *(x - E(x))2* | *(x-E(x))2f (x)* |
| 0 | 0.13 | 0 | -1.14 | 1.2996 | 0.1689 |
| 1 | 0.60 | 0.6 | -0.14 | 0.0196 | 0.0118 |
| 2 | 0.27 | 0.54 | 0.86 | 0.7396 | 0.1997 |
|  |  | 1.14 |  |  | 0.3804 |
|  |  | *E*(*x*) |  |  | *Var*(*x*) |

We see that *E(x)* = 1.14 snack items and *Var(x)* = .3804.

c. The marginal probability distribution of *y* along with the calculation of the expected value and variance is shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *y* | *f (y)* | *yf (y)* | *y - E(y)* | *(y - E(y))2* | *(y-E(y))2f (y)* |
| 0 | 0.60 | 0 | -0.5 | 0.25 | 0.15 |
| 1 | 0.30 | 0.3 | 0.5 | 0.25 | 0.075 |
| 2 | 0.10 | 0.2 | 1.5 | 2.25 | 0.225 |
|  |  | 0.5 |  |  | 0.45 |
|  |  | *E*(*y*) |  |  | *Var*(*y*) |

We see that *E(y)* = .50 reading materials and *Var(y)* = .45.

d. The probability distribution of *t* = *x* + *y* is shown below along with the calculation of its expected value and variance.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *t* | *f (t)* | *tf (t)* | *t-E(t)* | *(t-E(t))2* | *(t-E(t))2f (t)* |
| 1 | 0.50 | 0.5 | -0.64 | 0.4096 | 0.2048 |
| 2 | 0.38 | 0.76 | 0.36 | 0.1296 | 0.0492 |
| 3 | 0.10 | 0.3 | 1.36 | 1.8496 | 0.1850 |
| 4 | 0.02 | 0.08 | 2.36 | 5.5696 | 0.1114 |
|  |  | 1.64 |  |  | 0.5504 |
|  |  | *E(t)* |  |  | *Var(t)* |

We see that the expected number of items purchased is *E(t)* = 1.64 and the variance in the number of purchases is *Var(t)* = .5504.

e. From part (b), *Var(x)* = .3804. From part (c), *Var(y)* = .45. And from part (d), *Var(x* + *y)* = *Var(t)* = .5504. Therefore,



To compute the correlation coefficient, we must first obtain the standard deviation of *x* and *y*.



So the correlation coefficient is given by



The relationship between the number of reading materials purchased and the number of snacks purchased is negative. This means that the more reading materials purchased the fewer snack items purchased and vice versa.

63. a. The All World stock fund would be considered the more risky because it has a larger standard deviation. Indeed, the stock fund will experience a loss if the return is one standard deviation of 18.9% below the mean or 7.80%.

b. To answer this question we need to compute the expected value, variance, and standard deviation of .75*x* + .25*y*.



To compute the variance and standard deviation of the portfolio, we need to first compute the variance for *x* and *y*.





and 

Expected return ($10,000 investment) = 10,000(.07225) = $722.50

Standard deviation ($10,000 investment) = 10,000(.1406) = $1406

c. To answer this question we need to compute the expected value, variance, and standard deviation of .25*x* + .75*y*.



Expected return ($10,000 investment) = 10,000(.06075) = $607.50

Standard deviation ($10,000 investment) = 10,000(.054386) = $543.86

d. I would recommend the portfolio in part (b) for an aggressive investor because it has a larger return.

I would recommend the portfolio in part (c) for a conservative investor because it has a smaller standard deviation and is, thus, less risky.

64. a. *n* = 20 and *x* = 3



b. *n* = 20 and *x* = 0



c. *E*(*x*) = *n p* = 2000(.05) = 100

The expected number of employees is 100.

d. ** = *np* (1 - *p*) = 2000(.05)(.95) = 95

** =  = 9.75

65. a. We must have *E*(*x*) = *np* ≥ 20

For the 18-34 age group, *p* = .26.

*n*(.26) ≥ 20

*n* ≥ 76.92

Sample at least 77 people to have an expected number of home owners at least 20 for this age group.

b. For the 35-44 age group, *p* = .50.

*n*(.50) ≥ 20

*n* ≥ 40

Sample at least 40 people to have an expected number of home owners at least 20 for this age group.

c. For the 55 and over age group, *p* = .88.

*n*(.88) ≥ 20

*n* ≥ 22.72

Sample at least 23 people to have an expected number of home owners at least 20 for this age group.

d. 

e. 

66. Since the shipment is large we can assume that the probabilities do not change from trial to trial and use the binomial probability distribution.

a. *n* = 5



b. 

c. 1 - *f* (0) = 1 - .9510 = .0490

d. No, the probability of finding one or more items in the sample defective when only 1% of the items in the population are defective is small (only .0490). I would consider it likely that more than 1% of the items are defective.

67. a. *E*(*x*) = *np* = 100(.041) = 4.1

b. *Var*(*x*) = *np*(1 - *p*) = 100(.041)(.959) = 3.93



68. a. *E*(*x*) = 800(.30) = 240

b. 

c. For this one *p* = .70 and (1-*p*) = .30, but the answer is the same as in part (b). For a binomial probability distribution, the variance for the number of successes is the same as the variance for the number of failures. Of course, this also holds true for the standard deviation.

69. ** = 15

*P*(20 or more arrivals) = *f* (20) + *f* (21) + · · ·

= .0418 + .0299 + .0204 + .0133 + .0083 + .0050 + .0029

+ .0016 + .0009 + .0004 + .0002 + .0001 + .0001 = .1249

70. ** = 1.5

*P*(3 or more breakdowns) = 1 - [ *f* (0) + *f* (1) + *f* (2) ].

1 - [ *f* (0) + *f* (1) + *f* (2) ]

= 1 - [ .2231 + .3347 + .2510]

= 1 - .8088 = .1912

71. ** = 10 *f* (4) = .0189

72. a. 

b. *f* (3) + *f* (4) + · · · = 1 - [ *f* (0) + *f* (1) + *f* (2) ]



Similarly, *f* (1) = .1494, *f* (2) = .2240

∴ 1 - [ .0498 + .1494 + .2241 ] = .5767

73. Hypergeometric *N* = 52, *n* = 5 and *r* = 4.

a. 

b. 

c. 

d. 1 - *f* (0) = 1 - .6588 = .3412

74. a. Hypergeometric distribution with *N* = 10, *n* =2, and *r* = 7.



b. 

c. 