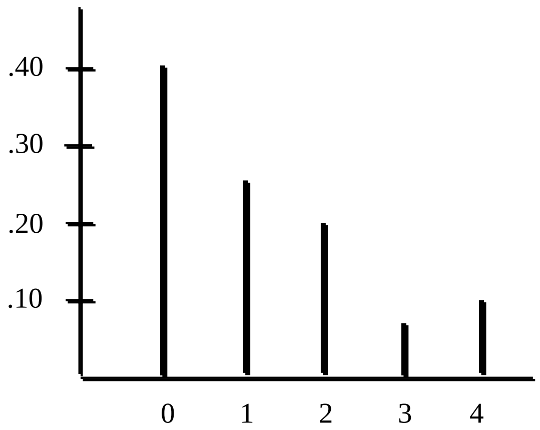
Chapter 5
Discrete Probability Distributions

- Random Variables
- Discrete Probability Distributions
- Expected Value and Variance
- Binomial Probability Distribution
- Poisson Probability Distribution
- Hypergeometric Probability Distribution



Random Variables

- A random variable is a numerical description of the outcome of an experiment.
- A random variable can be classified as being either discrete or continuous depending on the numerical values it assumes.
- A discrete random variable may assume either a finite number of values or an infinite sequence of values.
  - A random variable is discrete if it can assume only a countable number of values.
- A continuous random variable may assume any numerical value in an interval or collection of intervals.
  - A random variable is continuous if it can assume an uncountable number of values.

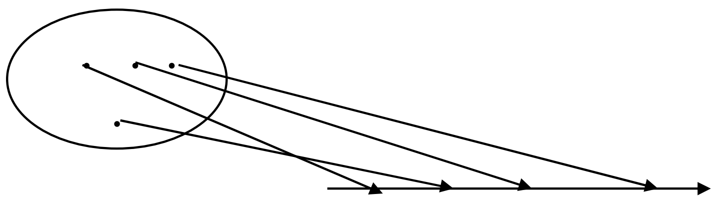
Random Variables

- DEFN:

A random variable is a function X: S→R

(From the sample space S to the real numbers R) which assigns a real number X(e) to each sample point $e \in S$:

X consists of all these arrows



Ex. X(TT)=0

Example: JSL Appliances

- Discrete random variable with a finite number of values

  Let x = number of TV sets sold at the store in one day

  where x can take on 5 values (0, 1, 2, 3, 4)

- Discrete random variable with an infinite sequence of values

  Let x = number of customers arriving in one day

  where x can take on the values 0, 1, 2, . . .

  We can count the customers arriving, but there is no finite upper limit on the number that might arrive.

Discrete Probability Distributions

- The <u>probability distribution</u> for a random variable describes how probabilities are distributed over the values of the random variable.
- A table, formula, or graph that lists all possible values a discrete random variable can assume, together with associated probabilities, is called *a discrete probability distribution.*
- The probability distribution is defined by a <u>probability function</u>, denoted by $f(x)$, which provides the probability for each value of the random variable.
- The required conditions for a discrete probability function are:

$$f(x) \geq 0$$

$$\sum f(x) = 1$$

- We can describe a discrete probability distribution with a table, graph, or equation.

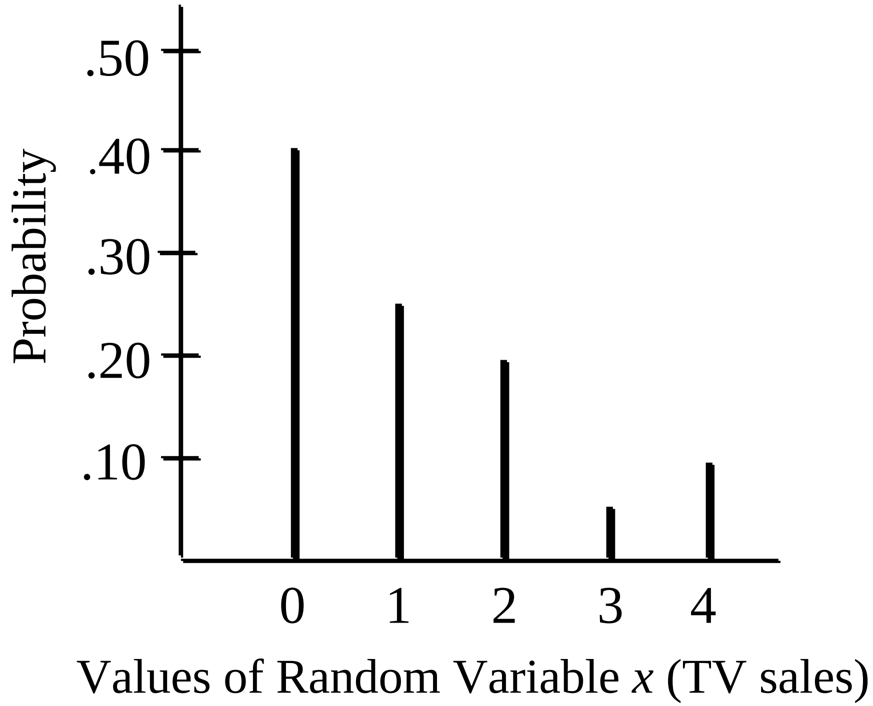
Example: JSL Appliances

- Using past data on TV sales (below left), a tabular representation of the probability distribution for TV sales (below right) was developed.

Units Sold	Number of Days	x	$f(x)$
0	80	0	.40
1	50	1	.25
2	40	2	.20
3	10	3	.05
4	20	4	.10
	200		1.00

Example: JSL Appliances

- Graphical Representation of the Probability Distribution



Slide 7

Discrete Uniform Probability Distribution

- The discrete uniform probability distribution is the simplest example of a discrete probability distribution given by a formula.
- The discrete uniform probability function is

$$\boxed{f(x) = 1/n}$$

where:

n = the number of values the random variable may assume

- Note that the values of the random variable are equally likely.

Slide 8

Expected Value and Variance

- The expected value, or mean, of a random variable is a measure of its central location.
- Expectation is a theoretical version of average
  - Expected value of a discrete random variable:

$$\boxed{E(x) = \mu = xf(x)}$$

- The variance summarizes the variability in the values of a random variable.
  - Variance of a discrete random variable:

$$\boxed{\mathrm{Var}(x) = \sigma^2 = (x - \mu)^2 f(x)}$$

- The standard deviation, σ, is defined as the positive square root of the variance.

Example: JSL Appliances

- Expected Value of a Discrete Random Variable

x	$f(x)$	$xf(x)$
0	.40	.00
1	.25	.25
2	.20	.40
3	.05	.15
4	.10	.40
		$E(x) =$ 1.20

The expected number of TV sets sold in a day is 1.2

Example: JSL Appliances

- Variance and Standard Deviation of a Discrete Random Variable

x	x -1.2	$(x - 1.2)^2$	$f(x)$	$(x - 1.2)^2 f(x)$
0	-1.2	1.44	.40	.576
1	-0.2	0.04	.25	.010
2	0.8	0.64	.20	.128
3	1.8	3.24	.05	.162
4	2.8	7.84	.10	.784
				1.660

The variance of daily sales is 1.66 TV sets squared.

The standard deviation of sales is 1.2884 TV sets.

Bernoulli Probability Distribution

- Properties of a Bernoulli Experiment
  - Two outcomes, success and failure, are possible on the trial.
  - The probability of a success, denoted by *p*.
- Bernoulli random variable

  X = 1, if success;

  = 0, if failure.

  P(X=1)=P(success)=p; P(X=0)=P(failure)=1-p

$$f(x) = p^{x}(1-p)^{(1-x)}$$

$$x = 0,1$$

Slide 12

Binomial Probability Distribution

- Properties of a Binomial Experiment
  - The experiment consists of a sequence of *n* identical trials.
  - Two outcomes, success and failure, are possible on each trial.
  - The probability of a success, denoted by *p*, does not change from trial to trial.
  - The trials are independent.

Example: Evans Electronics

- Binomial Probability Distribution

Evans is concerned about a low retention rate for employees. On the basis of past experience, management has seen a turnover of 10% of the hourly employees annually. Thus, for any hourly employees chosen at random, management estimates a probability of 0.1 that the person will not be with the company next year.

Choosing 3 hourly employees a random, what is the probability that 1 of them will leave the company this year?

Let: p = .10, n = 3, x = 1

Binomial Probability Distribution

- Binomial Probability Function

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

where:

$f(x)$ = the probability of x successes in n trials

n = the number of trials

p = the probability of success on any one trial

Binomial Probability Distribution

$$f(x)=\binom{n}{x}p^x(1-p)^{(n-x)}$$

$$q=1-p$$

$$(p+q)^n=\binom{n}{0}p^0(1-p)^{(n)}+\binom{n}{1}p^1(1-p)^{(n-1)}+\cdots+\binom{n}{n-1}p^{n-1}(1-p)^1+\binom{n}{n}p^n(1-p)^0$$

$X=x$: 0 1 $\cdots$ *n-1* *n*

$$\sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{(n-x)}=(p+(1-p))^n=1$$

Slide 16

Example: Evans Electronics

- Using the Binomial Probability Function

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

$$f(1) = \frac{3!}{1!(3-1)!} (0.1)^1 (0.9)^2$$

$= (3)(0.1)(0.81)$

$= .243$

Slide 17

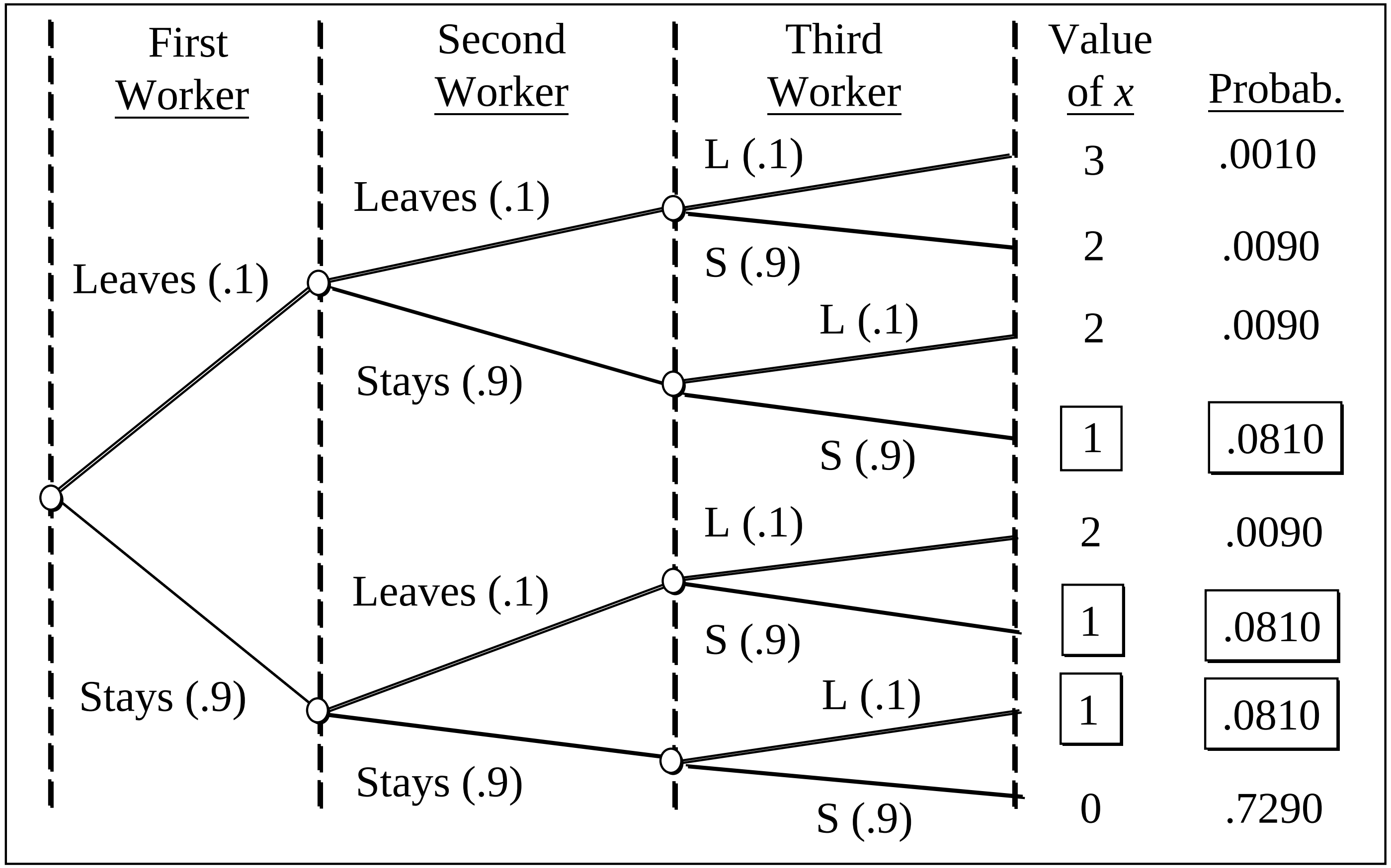
Example: Evans Electronics

- Using the Tables of Binomial Probabilities

		p								
n	***x***	.10	.15	.20	.25	.30	.35	.40	.45	.50
3	0	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250

Example: Evans Electronics

- Using a Tree Diagram



Slide 19

Binomial Probability Distribution

- Expected Value

$$\boxed{E(x) = \mu = np}$$

- Variance

$$\boxed{\mathrm{Var}(x) = \sigma^2 = np(1 - p)}$$

- Standard Deviation

$$\boxed{SD(x) = \sigma = \sqrt{np(1-p)}}$$

Slide 20

Binomial Probability Distribution

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)}$$

$$q = 1 - p$$

$$(p+q)^n = \binom{n}{0} p^0 (1-p)^{(n)} + \binom{n}{1} p^1 (1-p)^{(n-1)} + \cdots + \binom{n}{n-1} p^{n-1} (1-p)^1 + \binom{n}{n} p^n (1-p)^0$$

$$X = x \quad : \quad 0 \qquad\qquad 1 \qquad\qquad \cdots \quad n\text{-}1 \qquad\qquad n$$

$$\sum_{x=0}^{n} x f(x) = \sum x \binom{n}{x} p^x (1-p)^{(n-x)} = \sum x \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

$$= \sum \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{(n-x)} = np \sum \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{(n-x)}$$

$$= np \sum \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-x)} = np$$

Slide 21

Binomial Probability Distribution

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)}$$

$$E(x) = \sum_{x=0}^{n} x f(x) = np$$

$$E(x^2) = E(x(x-1)) + E(x) = (n(n-1))\, p^2 + np$$

$$= np((n-1)\, p + 1) = np(np - p + 1)$$

$$Var(x) = E(x^2) - (E(x))^2 = np(np - p + 1) - (np)^2 = np(1-p)$$

Slide 22

Example: Evans Electronics

- Binomial Probability Distribution
  - Expected Value

    $E(x)$ = np =3(.1) = .3 employees out of 3
  - Variance

    Var(x) = np(1-p) = 3(.1)(.9) = .27
  - Standard Deviation

$$SD(x) = \sigma = \sqrt{3(.1)(.9)} = .52 \text{ employees}$$

- What is the expected number of customers who used a credit card?
  - E(X) = np = 20(.30) = 6
- Find the probability that exactly 14 customers did not use a credit card.
  - Let Y be the number of customers who did not use a credit card.
    P(Y=14) = P(X=6) = P(X<=6) - P(x<=5) = .608 - .416 = .192
- Find the probability that at least 9 customers did not use a credit card.
  - Let Y be the number of customers who did not use a credit card.
    P(Y>=9) = P(X<=11) = .995

Poisson Probability Distribution

- Properties of a Poisson Experiment
  - The probability of an occurrence is the same for any two intervals of equal length.
  - The occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval.
- The Poisson experiment typically fits cases of rare events that occur over a fixed amount of time or within a specified region
- Typical cases
  - The number of errors a typist makes per page
  - The number of customers entering a service station per hour
  - The number of telephone calls received by a switchboard per hour.

Poisson Probability Distribution

- Poisson Probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where:

$f(x)$ = probability of x occurrences in an interval

= mean number of occurrences in an interval

e = 2.71828

Slide 26

- The Poisson Random Variable
  - The Poisson variable indicates the number of successes that occur
    during a given time interval
    or in a specific region in a Poisson experiment
  - Shape of the distribution is skewed right

- Probability Distribution of the Poisson Random Variable.

$$P(X = x) = p(x) = \frac{e^{-\mu}\mu^{x}}{x!} \qquad x = 0,1,2\ldots$$

$$E(X) = V(X) = \mu$$

$$\sum_{L=0}^{\infty} \frac{\mu^{L}}{L!} = e^{\mu} \qquad\qquad E(X^{2}) = \mu^{2} + \mu$$

Slide 27

$$P(X = x) = p(x) = \frac{e^{-\mu}\mu^{x}}{x!} \qquad x = 0,1,2\ldots \qquad \sum_{L=0}^{\infty} \frac{\mu^{L}}{L!} = e^{\mu}$$

$$E(X) = V(X) = \mu$$

$$E(X^2) = \mu^2 + \mu$$

$$E(X) = \sum x \frac{\mu^x e^{-\mu}}{x!} = \sum \frac{\mu^x e^{-\mu}}{(x-1)!} = \mu \sum \frac{\mu^{x-1} e^{-\mu}}{(x-1)!}$$

$$= \mu(e^{-\mu} \sum \frac{\mu^{x-1}}{(x-1)!}) = \mu$$

$$E(X^2) = E(X(X-1)) + E(X) = \mu^2 + \mu$$

$$Var(X) = E(X^2) - (E(X))^2 = \mu^2 + \mu - \mu^2 = \mu$$

Slide 28

Example: Mercy Hospital

- Using the Poisson Probability Function

  Patients arrive at the emergency room of Mercy Hospital at the average rate of 6 per hour on weekend evenings. What is the probability of 4 arrivals in 30 minutes on a weekend evening?

= 6/hour = 3/half-hour, x = 4

$$f(4) = \frac{3^4 (2.71828)^{-3}}{4!} = .1680$$

Example: Mercy Hospital

- Using the Tables of Poisson Probabilities

	μ									
x	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	..0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081

Poisson Approximation of the Binomial

- When n is very large, binomial probability table may not be available.
- If p is very small ($p < .05$), we can approximate the binomial probabilities using the Poisson distribution.
- Use m = np and make the following approximation:

$$P(X_{\text{Binomial}} = x) \cong P(X_{\text{Poisson}} = x)$$

(for n is large, p is small, $np \le 7$)

With parameters n and p

With $\mu = np$

- Example
  - A warehouse has a policy of examining 50 sunglasses from each incoming lot, and accepting the lot only if there are no more than two defective pairs.
  - What is the probability of a lot being accepted if, in fact, 2% of the sunglasses are defective?
  - Solution
    - This is a binomial experiment with $n = 50$, $p = .02$.
    - Tables for $n = 50$ are not available; $p<.05$; thus, a Poisson approximation is appropriate

    $[m = (50)(.02) = 1]$

    - $P(X_{poisson} <= 2) = .920$

    (true binomial probability = .922)

Hypergeometric Probability Distribution

- The hypergeometric distribution is closely related to the binomial distribution.
- With the hypergeometric distribution, the trials are not independent, and the probability of success changes from trial to trial.
- Rule of thumb

  Binomial distribution can be used when the sample size is small compared to the size of population

  ($n \leqq 0.05N$) or $n/N \leqq 0.05$

Slide 33

Hypergeometric Probability Distribution

- Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \quad \text{for } 0 \le x \le r$$

where: $f(x)$ = probability of x successes in n trials

n = number of trials

N = number of elements in the population

r = number of elements in the population labeled success

E(X) = n (r / N)

V(X)= [(N - n) / (N - 1)] n (r / N) (1 – r / N)

Example: Neveready

- Hypergeometric Probability Distribution

    Bob Neveready has removed two dead batteries from a flashlight and inadvertently mingled them with the two good batteries he intended as replacements. The four batteries look identical.

    Bob now randomly selects two of the four batteries. What is the probability he selects the two good batteries?

Example: Nevereadý

■ Hypergeometric Probability Distribution

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{2}{2}\binom{2}{0}}{\binom{4}{2}} = \frac{\left(\frac{2!}{2!0!}\right)\left(\frac{2!}{0!2!}\right)}{\left(\frac{4!}{2!2!}\right)} = \frac{1}{6} = .167$$

where:

x = 2 = number of good batteries selected

n = 2 = number of batteries selected

N = 4 = number of batteries in total

r = 2 = number of good batteries in total

Expected Value and Variance

- The expected value
  - Given a discrete random variable **X** with values x_i, that occur with probabilities $p(x_i)$, the expected value of **X** is

$$E(X) = \sum_{\text{all } x_i} x_i \cdot p(x_i)$$

  - The expected value of a random variable **X** is the weighted average of the possible values it can assume, where the weights are the corresponding probabilities of each x_i.

Laws of Expected Value

- $E(c) = c$
- $E(c\mathbf{X}) = cE(\mathbf{X})$
- $E(\mathbf{X} + \mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y})$
  $E(\mathbf{X} - \mathbf{Y}) = E(\mathbf{X}) - E(\mathbf{Y})$
- $E(\mathbf{XY}) = E(\mathbf{X})E(\mathbf{Y})$ if **X** and **Y** are independent random variables.

Slide 38

Variance

- Let **X** be a discrete random variable with possible values x_i that occur with probabilities $p(x_i)$, and let $E(x_i) = \mu$. The variance of **X** is defined to be

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_{\text{all } x_i} (x_i - \mu)^2 p(x_i)$$

- The variance is the weighted average of the squared deviations of the values of **X** from their mean μ, where the weights are the corresponding probabilities of each x_i.

- Standard deviation
  - The standard deviation of a random variable **X**, denoted s, is the positive square root of the variance of **X**.
- Example 6.5
  - The total number of cars to be sold next week is described by the following probability distribution

x	0	1	2	3	4
p(x)	.05	.15	.35	.25	.20

  - Determine the expected value and standard deviation of **X,** the number of cars sold.

Slide 40

$$E(X) = \mu = \sum_{i=1}^{5} x_i p(x_i)$$

x	0	1	2	3	4
p(x)	.05	.15	.35	.25	.20

$$= 0(0.05) + 1(0.15) + 2(0.35) + 3(0.25) + 4(0.20)$$

$$= 2.40$$

$$V(X) = \sigma^2 = \sum_{i=1}^{5} (x_i - 2.4)^2 p(x_i)$$

$$= (0-2.4)(.05) + (1-2.4)(.15) + (2-2.4)(.35)$$

$$+ (3-2.4)(.25) + (4-2.4)(.20) = 1.24$$

$$\sigma = \sqrt{1.24} = 1.11$$

Slide 41

- Example
  - With the probability distribution of cars sold per week (example 6.5), assume a salesman earns a fixed weekly wages of \$150 plus \$200 commission for each car sold.
  - What is his expected wages and the variance of the wages for the week?
  - Solution
    - The weekly wages is $Y = 200X + 150$
    - $E(Y) = E(200X+150) = 200E(X)+150= 200(2.4)+150=630$ \$.
    - $V(Y) = V(200X+150) = 200^2V(X) = 200^2(1.24) = 49{,}600$ $\$^2$

Bivariate Distributions

- To consider the relationship between two random variables, the **bivariate (or joint) distribution** is needed.

- **Bivariate probability distribution**
  - The probability that X assumes the value x, and Y assumes the value y is denoted

$$p(x,y) = P(X=x,\ Y = y)$$

The joint probabilities function satisfies the following conditions:

1. $0 \leq p(x,y) \leq 1$
2. $\sum_{\text{all } x} \sum_{\text{all } y} p(x,y) = 1$

Slide 43

- Example
  - Xavier and Yvette are two real estate agents. Let X and Y denote the number of houses that Xavier and Yvette will sell next week, respectively.

The bivariate (joint) probability distribution

	X			
Y	0	1	2	p(y)
0	.12	.42	.06	.60
1	.21	.06	.03	.30
2	.07	.02	.01	.10
p(x)	.40	.50	.10	1.00

p(0,0)

p(0,1)

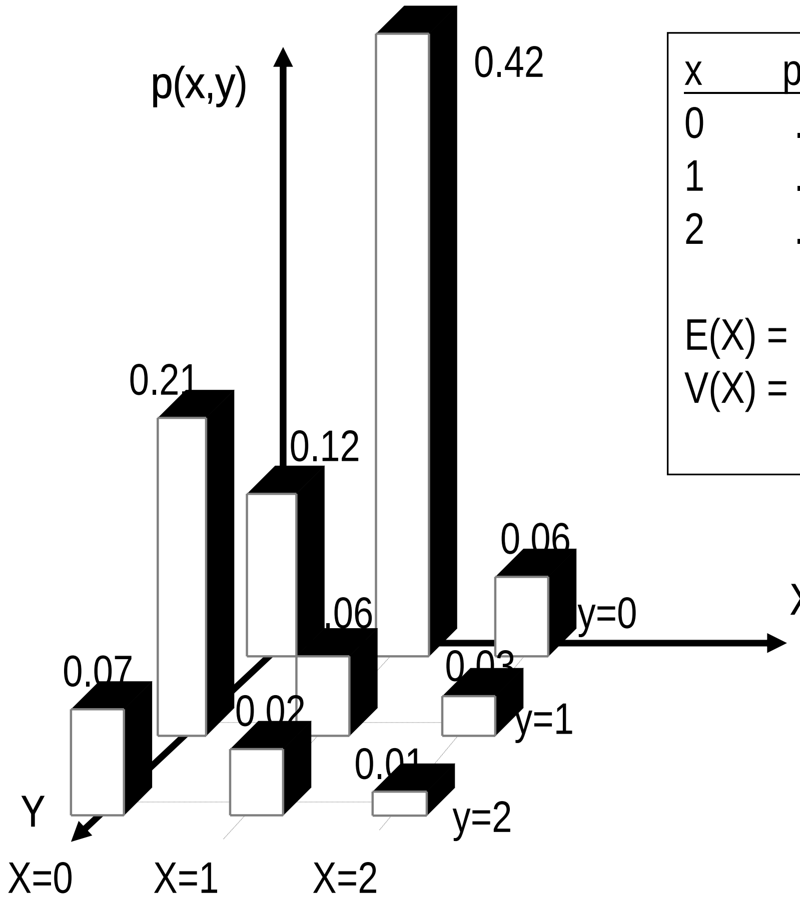
p(0,2)

P(Y=1), the marginal probability.

P(X=0)

The marginal probability

Slide 44



x	p(x)	y	p(y)
0	.4	0	.6
1	.5	1	.3
2	.1	2	.1

E(X) = .7 E(Y) = .5

V(X) = .41 V(Y) = .45

Slide 45

Calculating Conditional Probability

Y	X = 0	X = 1	X = 2	p(y)
0	.12	.42	.06	.60
1	.21	.06	.03	.30
2	.07	.02	.01	.10
p(x)	.40	.50	.10	1.00

$$P(X = x \mid Y = y) = \frac{P(X = x \text{ and } Y = y)}{P(Y = y)}$$

- Example - continued

$$P(X = 0 \mid Y = 1) = \frac{P(X = 0 \text{ and } Y = 1)}{P(Y = 1)} = \frac{.21}{.30} = .7$$

$$P(X = 1 \mid Y = 1) = \frac{P(X = 1 \text{ and } Y = 1)}{P(Y = 1)} = \frac{.06}{.30} = .2$$

$$P(X = 2 \mid Y = 1) = \frac{P(X = 2 \text{ and } Y = 1)}{P(Y = 1)} = \frac{.03}{.30} = .1$$

The sum is equal to 1.0

Slide 46

Conditions for independence

- Two random variables are said to be independent when

$$\boxed{P(X=x|Y=y)=P(X=x) \text{ or } P(Y=y|X=x)=P(Y=y).}$$

- This leads to the following relationship for independent variables

$$\boxed{P(X=x \text{ and } Y=y) = P(X=x)P(Y=y)}$$

- Example
  - Since $P(X=0|Y=1)=.7$ but $P(X=0)=.4$, The variables X and Y are **not independent.**

Slide 47

- Additional example
  - The table below represent the joint probability distribution of the variable X and Y. Are the variables X and Y independent

	X		
Y	1	2	p(y)
1	.28	.42	.7
2	.12	.18	.3
P(y)	.40	.60	

Compare the other two pairs.
Yes, the two variables are independent

Find the marginal probabilities of X and Y.
Then apply the multiplication rule.

P(X=1)P(Y=1) = .40(.70) = .28

P(X=1 and Y=1) = .28

P(X=1)P(Y=2) = .40(.30) = .12

P(X=1 and Y=2) = .12

Slide 48

The sum of two variables

- To calculate the probability distribution for a sum of two variables X and Y observe the example below.
- Example 6.7 - continued
  - Find the probability distribution of the total number of houses sold per week by Xavier and Yvette.
  - Solution
    - X+Y is the total number of houses sold. X+Y can have the values 0, 1, 2, 3, 4.
    - We find the distribution of X+Y as demonstrated next.

P(X+Y=0) = P(X=0 and Y=0) = .12

P(X+Y=1) = P(X=0 and Y=1)+ P(X=1 and Y=0) =.21 + .42 = .63

P(X+Y=2) = P(X=0 and Y=2)+ P(X=1 and Y=1)+ P(X=2 and Y=0)
= .07 + .06 + .06 = .19

The probabilities P(X+Y)=3 and P(X+Y) =4 are calculated the same way. The distribution follows

……..

……..

x +y	0	1	2	3	4
p(x+y)	.12	.63	.19	.05	.01

	X			
Y	0	1	2	p(y)
0	.12	.42	.06	.60
1	.21	.06	.03	.30
2	.07	.02	.01	.10
p(x)	.40	.50	.10	1.00

Slide 50

- Expected value and variance of X+Y
  - When the distribution of X+Y is known (see the previous example) we can calculate E(X+Y) and V(X+Y) directly using their definitions.
  - An alternative is to use the relationships
    - $E(aX+bY) = aE(X) + bE(Y)$;
    - $V(aX+bY) = a^2V(X) + b^2V(Y)$ if X and Y are independent.
    - When X and Y are not independent, (see the previous example) we need to incorporate the covariance in the calculations of the variance V(aX+bY).

Covariance

- The covariance is a measure of the degree to which two random variables tend to move together.

$$\mathrm{COV}(X,Y) = \sum_{\text{Over all } x,y} (x-\mu_x)(y-\mu_y)p(x,y)$$

$$= E[(X - \mu_x)(Y - \mu_y)]$$

$$= E(XY) - \mu_x\mu_y$$

The coefficient of correlation

$$\rho = \frac{\mathrm{COV}(X,Y)}{\sigma_x \sigma_y}$$

- Example
  - Find the covariance of the sales variables X and Y, then calculate the coefficient of correlation.

  - Solution
    - Calculation of the expected values:
      $\mu_x = \sum x_i p(x_i) = 0(.4)+1(.5)+2(.1)=.7$
      $\mu_y = \sum y_i p(y_i) = 0(.6)+1(.3)+2(.1)=.5$

    - Calculation of the covariance:
      $COV(X,Y) = \Sigma(x - \mu_x)(y - \mu_y)p(x,y)$
      $= (0-.7)(0-.5)(.12)+(0-.7)(1-.5)(.21)+$
      $(0-.7)(2-5)(.07) +\ldots+(2-..7)(2-.5)(.01)= -.15$

      There is a negative relationship between the two variables

Slide 53

- To find how strong the relationship between X and Y is we need to calculate the coefficient of correlation.

  - Calculation of the standard deviations of X and Y

    $V(X) = \sum (x_i - \mu_x)^2 p(x_i) = (0-.7)^2(.4)+(1-.7)^2(.5)+(2-.7)^2(.1)=.41$

    $\sigma_x = [V(X)]^{1/2} = .64$

    In a similar manner we have V(Y) = .45

    $\sigma_y = [.45]^{1/2}=.67$

  - Calculation of ρ

$$\rho = \frac{COV(X,Y)}{\sigma_x \sigma_y} = \frac{-.15}{(.64)(.67)} = -.35$$

There is a relatively weak negative relationship between X and Y .

Slide 54

- The variance of the sum of two variables X and Y can now be calculated using

$$V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCOV(X,Y)$$
$$= a^2V(X) + b^2V(Y) + 2ab\ \rho\sigma_x\sigma_y$$

Slide 55

■ Example Investment portfolio diversification

- An investor has decided to invest equal amounts of money in two investments.

	Mean return	Standard dev.
Investment 1	15%	25%
Investment 2	27%	40%

- Find the expected return on the portfolio
- If r = 1, .5, 0 find the standard deviation of the portfolio.

- Solution
  - The return on the portfolio can be represented by

$$R_p = w_1R_1 + w_2R_2 = .5R_1 + .5R_2$$

The relative weights are proportional to the amounts invested.

  - Thus, $E(R_p) = w_1E(R_1) + w_2E(R_2)$

$$=.5(.15) + .5(.27) = .21$$

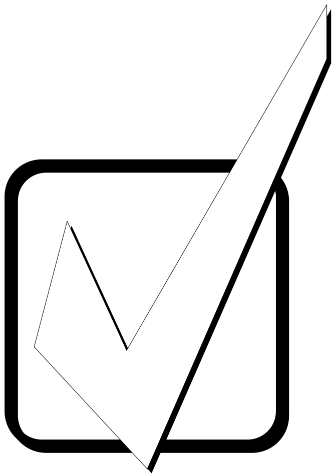
  - The variance of the portfolio return is

$$V(R_p) = w_1^2V(R_1) + w_2^2V(R_1) + 2w_1w_2\rho\sigma_1\sigma_2$$

- Substituting the required coefficient of correlation we have:
  - For r = 1 : $V(R_p) = .1056$ $\sigma_p = .3250$
  - For r = .5: $V(R_p) = .0806$ $\sigma_p = .2839$
  - For r = 0: $V(R_p) = .0556$ $\sigma_p = .2358$

Larger diversification is expressed by smaller correlation.
As the correlation coefficient decreases, the standard deviation decreases too.

End of Chapter 5



Slide 59