Probability Distributions

Random Variable

- Random variable
 - Outcomes of an experiment expressed numerically
 - e.g.: Throw a die twice; Count the number of times the number 6 appears (0, 1 or 2 times)



Discrete Random Variable

 It is one which can vary only by finite "jumps" and cant manifest every conceivable fractional value.

Ex: No. of children in a family, No.of phone calls, Toss a coin five times;
Count the No. of tails (0, 1, 2, 3, 4, or 5 times)

Teres a to fine

Discrete Random Variable with a Finite Number of Values

A <u>discrete random variable</u> may assume either a finite number of values or an infinite sequence of values.

Let x = number of TVs sold at the store in one day, where x can take on 5 values (0, 1, 2, 3, 4)

We can count the TVs sold, and there is a finite upper limit on the number that might be sold (which is the number of TVs in stock).

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.

Continuous Random Variable

• A <u>continuous random variable</u> may assume any numerical value in an interval or collection of intervals.

Example: Height, Weight

Random Variables

Question	Random Variable x	Туре		
Family size	x = Number of dependents reported on tax return	Discrete		
Distance from home to store	x = Distance in miles from home to the store site	Continuous		
Own dog or cat	<pre>x = 1 if own no pet; = 2 if own dog(s) only; = 3 if own cat(s) only; = 4 if own dog(s) and cat(s)</pre>	Discrete		

The <u>probability distribution</u> for a random variable describes how probabilities are distributed over the values of the random variable.

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

Two types of <u>discrete probability distributions</u> will be introduced.

First type: uses the rules of assigning probabilities to experimental outcomes to determine probabilities for each value of the random variable.

Second type: uses a special mathematical formula to compute the probabilities for each value of the random variable.

The probability distribution is defined by a <u>probability function</u>, denoted by f(x), that provides the probability for each value of the random variable.

The required conditions for a discrete probability function are:

$$f(x) \ge 0$$
$$\Sigma f(x) = 1$$

There are three methods for assign probabilities to random variables: the classical method, the subjective method, and the relative frequency method.

The use of the relative frequency method to develop discrete probability distributions leads to what is called an <u>empirical discrete distribution</u>.

- Example: JSL Appliances
 - Using past data on TV sales, ...
 - a <u>tabular representation</u> of the probability distribution for TV sales was developed.



Example: JSL Appliances



Graphical representation of probability distribution

In addition to tables and graphs, a formula that gives the probability function, f(x), for every value of x is often used to describe the probability distributions.

Several discrete probability distributions specified by formulas are the discrete-uniform, binomial, Poisson, and hypergeometric distributions.

Discrete Uniform Probability Distribution

The <u>discrete uniform probability distribution</u> is the simplest example of a discrete probability distribution given by a formula.

The discrete uniform probability function is

$$f(x)=1/n$$

the values of the random variable are equally likely

where:

n = the number of values the random variable may assume

Expected Value

The <u>expected value</u>, or mean, of a random variable is a measure of its central location.

 $E(x) = \mu = \Sigma x f(x)$

The expected value is a weighted average of the values the random variable may assume. The weights are the probabilities.

The expected value does not have to be a value the random variable can assume.

Variance and Standard Deviation

The <u>variance</u> summarizes the variability in the values of a random variable.

$$Var(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x)$$

The variance is a weighted average of the squared deviations of a random variable from its mean. The weights are the probabilities.

The <u>standard deviation</u>, σ , is defined as the positive square root of the variance.

Expected Value

Example: JSL Appliances

<u>x</u>	<u>f(x)</u>	$\underline{xf(x)}$		
0	.40	.00		
1	.25	.25		
2	.20	.40		
3	.05	.15		
4	.10	<u>.40</u>		
E(x) = 1.20				

expected number of TVs sold in a day

Variance

Example: JSL Appliances

	x	x - µ	$(x - \mu)^2$	f(x)	$(x-\mu)^2 f(x)$	
	Ő	-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	TVs
		Variance	of daily sal	$es = \sigma^2$	$=1\overline{660}$	squared

Standard deviation of daily sales = 1.2884 TVs

Bivariate Distributions

A probability distribution involving two random variables is called a <u>bivariate probability distribution</u>.

Each outcome of a bivariate experiment consists of two values, one for each random variable.

Example: rolling a pair of dice

When dealing with bivariate probability distributions, we are often interested in the relationship between the random variables.

A company asked 200 of its employees how they rated their benefit package and job satisfaction. The crosstabulation below shows the ratings data.

Benefits <u>Jo</u> Package (x)	BenefitsJob Satisfaction (y) Package (x) 123							
1	28	26	4	58				
2	22	42	34	98				
3	2	10	32	44				
Total	52	78	70	200				

The bivariate empirical discrete probabilities for benefits rating and job satisfaction are shown below.

Benefits <u>Jo</u> Package (x)	BenefitsJob Satisfaction (y)Package (x)123							
1	.14	.13	.02	.29				
2	.11	.21	.17	.49				
3	.01	.05	.16	.22				
Total	.26	.39	.35	1.00				

Expected Value and Variance for Benefits Package, x

<u>X</u>	<u>f(x)</u>	<u>xf(x)</u>	<u>x -</u> <u>E(x)</u>	$\frac{(x - E(x))^2}{E(x)^2}$	$(x - E(x))^2 f(x)$
1	0.29	0.29	-0.93	0.8649	0.250821
2	0.49	0.98	0.07	0.0049	0.002401
3	0.22	<u>0.66</u>	1.07	1.1449	<u>0.251878</u>
	E(x) =	1.93		Var(x) =	0.505100

Expected Value and Variance for Job Satisfaction, y

¥	<u>f(y)</u>	<u>yf(y)</u>	<u>y -</u> <u>E(y)</u>	$\frac{(y - y)^2}{E(y)^2}$	<u>(y - E(y))²f(y)</u>
1	0.26	0.26	-1.09	1.1881	0.308906
2	0.39	0.78	-0.09	0.0081	0.003159
3	0.35	<u>1.05</u>	0.91	0.8281	<u>0.289835</u>
	E(y) =	2.09		Var(y) =	0.601900



Binomial Distribution-Assumptions

- "n" Identical & finite trials
 - e.g.: 15 tosses of a coin; 10 light bulbs taken from a warehouse
- Two mutually exclusive outcomes on each trial
 - e.g.: Heads or tails in each toss of a coin; defective or not defective light bulb
- Trials are independent of each other
 - The outcome of one trial does not affect the outcome of the other.
- Constant probability for each trial
 - e.g.: Probability of getting a tail is the same each time a coin is tossed

Binomial Distribution

 A random variable X is said to follow Binomial distribution if it assumes only non negative values and its function is given by

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

- x: No. of "successes" in a sample, x=0,1,2...,n.
- n: Number of trials.
- p: probability of each success.
- q: probability of each failure (q=1-p)

•Notation: $X \sim B(n, p)$.

Binomial Probability Function



Number of experimental outcomes providing exactly *x* successes in *n* trials Probability of a particular sequence of trial outcomes with x successes in *n* trials

Example: Evans Electronics

Evans Electronics is concerned about a low retention rate for its employees. In recent years, management has seen a turnover of 10% of the hourly employees annually.

Thus, for any hourly employee 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 at random, what is the probability that 1 of them will leave the company this year?

Example: Evans Electronics

The probability of the first employee leaving and the second and third employees staying, denoted (S, F, F), is given by

p(1-p)(1-p)

With a .10 probability of an employee leaving on any one trial, the probability of an employee leaving on the first trial and not on the second and third trials is given by

 $(.10)(.90)(.90) = (.10)(.90)^2 = .081$

- Example: Evans Electronics
- Two other experimental outcomes also result in one success and two failures. The probabilities for all three experimental outcomes involving one success follow.

ExperimentalIOutcomeI(S, F, F)I(F, S, F)I(F, S, F)I(F, F, S)I

Probability of <u>Experimental Outcome</u> p(1-p)(1-p) = (.1)(.9)(.9) = .081 (1-p)p(1-p) = (.9)(.1)(.9) = .081 (1-p)(1-p)p = (.9)(.9)(.1) = .081Total = .243

Example: Evans Electronics

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

Using the probability function

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

$$f(1) = \frac{3!}{1!(3-1)!} (0.1)^1 (0.9)^2 = 3(.1)(.81) = (.243)^2$$

Using a tree diagram



Binomial Probabilities and Cumulative Probabilities

Statisticians have developed tables that give probabilities and cumulative probabilities for a binomial random variable.

These tables can be found in some statistics textbooks.

With modern calculators and the capability of statistical software packages, such tables are almost unnecessary.

Using Tables of Binomial Probabilities



Binomial Distribution Characteristics

 $\sigma = \sqrt{np(1-p)} = \sqrt{5(.1)(1-.1)} = .6708$

• Mean

Ex:

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

EX: $\mu = np = 5(.1) = .5$

 Variance and standard deviation-

$$\sigma^{2} = np(1-p)$$
$$\sigma = \sqrt{np(1-p)}$$



Example: Evans Electronics

• Expected Value

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

Variance

~

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

Standard Deviation

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

Example

 The experiment: Randomly draw red ball with replacement from an urn containing 10 red balls and 20 black balls.

- Use S to denote the outcome of drawing a red ball and F to denote the outcome of a black ball.
- Then this is a binomial experiment with p = 1/3.
- <u>Q</u>: Would it still be a binomial experiment if the balls were drawn without replacement?

Practice Questions

- Consider a binomial experiment with n=10 and p=.10
- Compute f(0), f(1), f(2).
- Compute P(x≤2)
- Compute $P(x \ge 1)$
- Compute E(x)
- Compute Var(x) and s.d

Questions for practice

- Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.
- A & B play a game in which their chances of winning are in the ratio 3:2. find A's chance of winning at least 3 games out of the 5 games played.
- Mr. Gupta applies for a personal loan of Rs 150,000 from a nationalized bank to repair his house. The loan offer informed him that over the years bank has received about 2920 loan applications per year and that the prob. Of approval was on average, about 0.85. Mr. Gupta wants to know the average and standard deviation of the number of loans approved per year.

The incidence of a certain disease in such that on the average 20% of workers suffer from it. If 10 workers are selected at random, find the probability that

- 1. Exactly 2 workers suffer from the disease.
- 2. Not more than 2 workers suffer from the disease.
 - Bring out the fallacy, if any
- 1. The mean of a binomial distribution is 15 and its standard deviation is 5.
- Find the binomial distribution whose
 mean is 6 and variance is 4.

- The probability that an evening college student will be graduate is 0.4.Determine the probability that out of 5 students
- None will be graduate.
- One will be graduate.
- At least one will be graduate.
- Multiple choice test consists of 8 questions and three answers to each question (of which only one is correct). A student answers each question by rolling a balanced die and marking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4 and the third answer if he gets 5 or 6. To get a distinction, student must secure at least 75% correct answers. If there is no negative marking, what is the probability that student secures a distinction?

- When a new machine is functioning properly, only 3% of the items produced are defective. Assume that we will randomly select two parts produced on the machine and that we are interested in the number of defective parts found.
- a. Describe the conditions under which this situation would be a binomial experiment.
- b. Draw a tree diagram showing this problem as a two-trial experiment.
- c. How many experimental outcomes result in exactly one defect being found?
- d. Compute the probabilities associated with finding no defects, exactly one defect, and two defects.

- 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

$$P(1 \text{ defect}) = 2(.03)(.97) = .0582$$

- A Harris Interactive survey for InterContinental Hotels & Resorts asked respondents, "When traveling internationally, do you generally venture out on your own to experience culture, or stick with your tour group and itineraries?" The survey found that 23% of the respondents stick with their tour group (USA Today, January 21, 2004).
 - a. In a sample of six international travelers, what is the probability that two will stick with their tour group?
 - b. In a sample of six international travelers, what is the probability that at least two will stick with their tour group?
- c. In a sample of 10 international travelers, what is the probability that none will stick with the tour group?

TABLE C Binomial probabilities

-						p	1.4		57	
12	k	.01	.02	.03	.04	.05	.05	.07	.05	09
2	0	.9801	.9604	9409	.9216	.9025	8836	.8649	.8464	8281
	- 1	.0198	.0392	.0582	.0768	.0950	.1128	1307	1472	103
	.2	.0001	.0004	.0009	.0016	.0025	.0036	.0049	.0064	3008
8	0	.9703	.9412	.9127	3847	.8574	.8306	8044	.7787	1755
	1	.0294	.0576	.0847	1106	.1354	.1590	.1810	2031	- 323
	2		.0012	.0026	.0046	.0071	.0102	0137	.0177	1123
	3				.0001	.0001	,0002	.0003	,0005	000
£	0	.9606	.9224	8853	.8493	.8145	.7807	.7481	.7164	685
	3	.0388	.0753	.1095	.1416	.1715	.1993	,2252	2492	1.273
	2	,0006	.0023	.0051	.0088	.0135	.0191	.0254	.0325	.040
	3			,0001	,000,2	.0005	.0008	.0013	-00Ha	100
								17.05		
500	0	.9510	-9039	8587	.8154	7738	7339	,6957		12.17
	1	.0480	.0922	-1328	.1699	.2036	2342	2018	2000	240
	1.2	.0010	.00.58	0008.2	10142	10214	0279	0030	0043	1000
			10001	30003	'iwwo	1992.8.2	00005	0001	0407	100
	5									
-	0	0415	8858	8330	7828	7351	.6899	.6470	6064	=6
	1	0571	1085	1546	1957	.2321	2642	2922	3104	145
	2	.0014	.0055	0120	.0204	.0305	.0422	.0550	0688	1,080
	. 3.	1.1.2.2.2.2.2	.0002	.0005	.0011	.0021	.0036	.0055	.0090	
	-4					.0061	,0002	.0003	:0005	
	5								1.0.0	
	6									
7	0	.9321	8681	.8080	.7514	.6983	.6485	.6017	.5578	151
	21	.0659	.1240	.1749	.2192	.2373	.2897	3170	.3396	
	2	,0020	.0076	0162	.0274	.0406	,0555	.0716	.0556	-10
	3		10004	.0005	.0019	.00.56	.0059	0050	.0128	101
	-4				1000	.0002	.0004	1000.5	AKET I-	- 100
									-OCHIA	-202
	7									
e.		9727	85/18	7837	7214	6634	6096	5596	5132	
	1	0746	1389	1979	2405	2793	3113	3370	3570	31
	2	0026	.0099	0210	.0351	.0515	.0695	.0888	1087	5.2
	3	.0001	.0004	.0013	.0029	.0054	.0089	.0134 .	.01.59	02
	-4	Contraction of the	a state and a	10001	.0002	:0004	.0007	.0013	0021	- 200
	5							.0001	.0001	135
	D.									
	7									
	38									

Binomial Distribution-Different situations

Random experiments and random variable

- **1.** Flip a coin 10 times. Let X = number of heads obtained.
- 2. A worn machine tool produces 1% defective parts. Let X = number of defective parts in the next 25 parts produced.
- 3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
- 4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X = the number of bits in error in the next five bits transmitted.

Situations Contd..

Random experiments and random variables

- 5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X = the number of questions answered correctly.
- 6. In the next 20 births at a hospital, let X = the number of female births.
- 7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X = the number of patients who experience improvement.

Poisson Distribution

A Poisson distributed random variable is often useful in estimating the number of occurrences over a <u>specified interval of time or space</u>

It is a discrete random variable that may assume an <u>infinite sequence of values</u> (x = 0, 1, 2, ...).

Examples of Poisson distributed random variables:

the number of leaks in 100 miles of pipeline

the number of vehicles arriving at a toll booth in one hour

Bell Labs used the Poisson distribution to model the arrival of phone calls.

Two Properties of a Poisson Experiment

1. The probability of an occurrence is the same for any two intervals of equal length.

2. The occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval.

Poisson Probability DistributionPoisson Probability Function

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

where:

x = the number of occurrences in an interval f(x) = the probability of x occurrences in an interval $\mu =$ mean number of occurrences in an interval e = 2.71828 $x! = x(x - 1)(x - 2) \dots (2)(1)$

Poisson Probability Function

Since there is no stated upper limit for the number of occurrences, the probability function f(x) is applicable for values x = 0, 1, 2, ... without limit.

In practical applications, x will eventually become large enough so that f(x) is approximately zero and the probability of any larger values of xbecomes negligible.

Example: Mercy Hospital

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?

Example: Mercy Hospital

Using the probability function

 $\mu = 6/\text{hour} = 3/\text{half-hour}, \ x = 4$ $f(4) = \frac{3^4 (2.71828)^{-3}}{44} = 0.1680$

Example: Mercy Hospital





A property of the Poisson distribution is that the mean and variance are equal.

 $\mu = \sigma^2$

Variance for Number of Arrivals During 30-Minute Periods

$$\mu = \sigma^2 = 3$$

Poisson Distribution Characteristics

Mean

$$\mu = E(X) = \lambda$$
$$= \sum_{i=1}^{N} X_i P(X_i)$$

Variance & Standard deviation

$$\sigma^2 = \lambda \qquad \sigma = \sqrt{\lambda}$$



Example

Arrivals at a bus-stop follow a Poisson distribution with an average of 4.5 every quarter of an hour.

(assume a maximum of 20 arrivals in a quarter of an hour) and calculate the probability of fewer than 3 arrivals in a quarter of an hour. The probabilities of 0 up to 2 arrivals can be calculated directly from the formula

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

with $\lambda = 4.5$

So p(0) = 0.01111

Similarly p(1)=0.0611 and p(2)=0.2381

So the probability of fewer than 3 arrivals is 0.01111+ 0.0611 + 0.2381 =0.31031

			P	olsson	Distri	bution	Table			
λ-	0.5	1.0	1.5	2.0	25	3.0	3.5	4.0	4.5	5
X-0	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0
1	0.9098	0.7358	0.5578	0.4060	0.2373	0.1991	0.1359	0.0916	0.0611	0.0
2	0.9856	0.9197	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1
3	0.9982	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2
4	0.9998	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4
5	1.0000	0.9994	0.9994	0.9955	0.9834	0.9161	0.8576	0.7851	0.7029	0.6
6	1.0000	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7
7	1.0000	1.0000	0.9998	0.9989	0.9958	6.9381	0.9733	0.9489	0.9134	0.8
8	1.0000	1.0000	1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9
9	1.0000	1.0000	1.0000	1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9
10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9972	0.9933	0.9

Questions for practice

- Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year?
- If 3% of the electric bulbs manufactured by a company are defective, find the probability that in a sample of 100 bulbs

a) 0 b) 5 bulbs c) more than 5

d) between 1 and 3

e) less than or equal to 2 bulbs are defective

 Comment on the following for a Poisson distribution with Mean = 3 and s.D = 2

In a Poisson distribution if p(2)= 2/3 p(1).
 Find
 i) mean ii) standard deviation
 iii)P(3) iv) p(x > 3)

A car hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which
i) Neither car is used.
ii) Some demand is refused. Phone calls arrive at the rate of 48 per hour at the reservation desk for Regional Airways.

- a. Compute the probability of receiving three calls in a 5-minute interval of time.
- b. Compute the probability of receiving exactly 10 calls in 15 minutes.
- c. Suppose no calls are currently on hold. If the agent takes 5 minutes to complete the current call, how many callers do you expect to be waiting by that time? What is the probability that none will be waiting?
- d. If no calls are currently being processed, what is the probability that the agent can take 3 minutes for personal time without being interrupted by a call?

46. a. $\mu = 48(5/60) = 4$

$$f(3) = \frac{4^3 e^{-4}}{3!} = \frac{(64)(.0183)}{6} = .1952$$

b.
$$\mu = 48(15/60) = 12$$

$$f(10) = \frac{12^{10} e^{-12}}{10!} = .1048$$

c. $\mu = 48 (5 / 60) = 4$ I expect 4 callers to be waiting after 5 minutes.

$$f(0) = \frac{4^{\circ} e^{-4}}{0!} = .0183$$

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

d.
$$\mu = 48 (3 / 60) = 2.4$$

$$f(0) = \frac{2.4^{\circ} e^{-2.4}}{0!} = .090$$

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

Airline passengers arrive randomly and independently at the passenger-screening facility at a major international airport. The mean arrival rate is 10 passengers per minute.

- Compute the probability of no arrivals in a one-minute period.
- b. Compute the probability that three or fewer passengers arrive in a one-minute period.
- c. Compute the probability of no arrivals in a 15-second period.
- d. Compute the probability of at least one arrival in a 15-second period.

Poisson type Situations

- Number of deaths from a disease.
- No. of suicides reported in a city.
- No. of printing mistakes at each page of a book.
- Emission of radioactive particles.
- No. of telephone calls per minute at a small business.
- No. of paint spots per new automobile.
- No. of arrivals at a toll booth
- No. of flaws per bolt of cloth