Lecture 4

Several Useful Discrete Distributions

Random Variables

- A **random variable** is a function corresponding to the outcome of an experiment, is a chance or random event.
- Random variables can be **discrete** or **continuous.**

- **Examples:**
  - ✓ x = SAT score for a randomly selected student
  - ✓ x = number of people in a room at a randomly selected time of day
  - ✓ x = weight of a fish drawn at random

Probability Distributions for Discrete Random Variables

Probability distribution of a discrete random variable *x* , is a graph, table or formula that gives

- possible values of *x*
- probability $p(x)$ associated with each value *x*.

We must have

$$0 \leq p(x) \leq 1 \text{ and } \sum p(x) = 1$$

Example 1

- Toss a fair coin once,
- Define ***x* = number of heads.**
- Find distribution of ***x***

		$\underline{x}$
H	1/2	1
T	1/2	0

$$P(x = 0) = 1/2$$
$$P(x = 1) = 1/2$$

x	$p(x)$
0	1/2
1	1/2

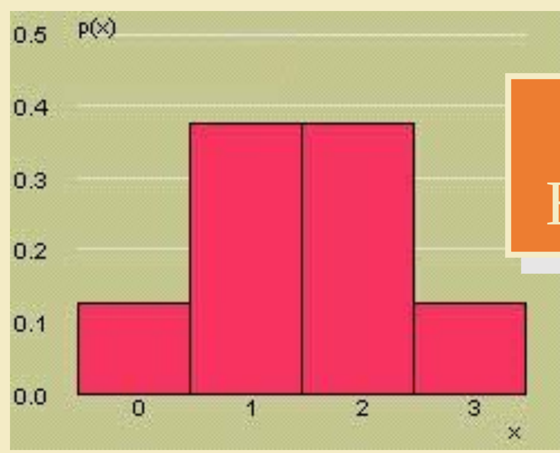
Example 2

- Toss a fair coin three times and define ***x* = number of heads.**

		x
HHH	1/8	3
HHT	1/8	2
HTH	1/8	2
THH	1/8	2
HTT	1/8	1
THT	1/8	1
TTH	1/8	1
TTT	1/8	0

$P(x = 0) = 1/8$
$P(x = 1) = 3/8$
$P(x = 2) = 3/8$
$P(x = 3) = 1/8$

x	$p(x)$
0	1/8
1	3/8
2	3/8
3	1/8



Probability Histogram for x

Expected Value of Random Variable

- Let *x* be a discrete random variable with probability distribution *p*(*x*). Then the **expected value**, denoted by *E*(*x*), is defined by

$$\text{Expected Value (Expectation, Mean } \mu)$$

$$E(x) = \mu = \sum xp(x)$$

Example

- Toss a fair coin 3 times and record *x* the number of heads.

x	$p(x)$	$xp(x)$
0	1/8	0(1/8)=0
1	3/8	1(3/8)=0.375
2	3/8	2(3/8)=0.75
3	1/8	3(1/8)=0.375
		Total 1.5

$$E(x) = \sum xp(x) = 1.5$$

Mean & Standard Deviation

- Let x be a discrete random variable with probability distribution $p(x)$. Then the mean, variance and standard deviation of x are given as

$$\text{Mean}: \mu = \sum x p(x)$$

$$\text{Variance}: \sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\text{Standard deviation}: \sigma = \sqrt{\sigma^2}$$

$$\text{Variance}: \sigma^2 = \sum x^2 p(x) - \mu^2$$

Example

- Toss a fair coin 3 times and record *x* the number of heads.
- Find variance by the definition formula.

x	$p(x)$	$(x-\mu)^2 p(x)$
0	**1/8**	**$(0-1.5)^2(1/8)=.28125$**
1	**3/8**	**$(1-1.5)^2(3/8)=.09375$**
2	**3/8**	**$(2-1.5)^2(3/8)=.09375$**
3	**1/8**	**$(3-1.5)^2(1/8)=.28125$**
		Total .75

$$\mu = E(x) = \sum xp(x) = 1.5$$

$$\sigma^2 = \sum (x-\mu)^2 p(x)$$
$$= 0.75$$

Example

- Toss a fair coin 3 times and record *x* the number of heads.

- Find the variance by the computational formula.

$$\mu = E(x) = \sum xp(x) = 1.5$$

x	$p(x)$	$x^2p(x)$
0	1/8	$0^2(1/8)=0$
1	3/8	$1^2(3/8)=0.375$
2	3/8	$2^2(3/8)=1.5$
3	1/8	$3^2(1/8)=1.125$
		Total 3

$$\sigma^2 = \sum x^2 p(x) - \mu^2$$
$$= 3 - 1.5^2$$
$$= 0.75$$

Introduction

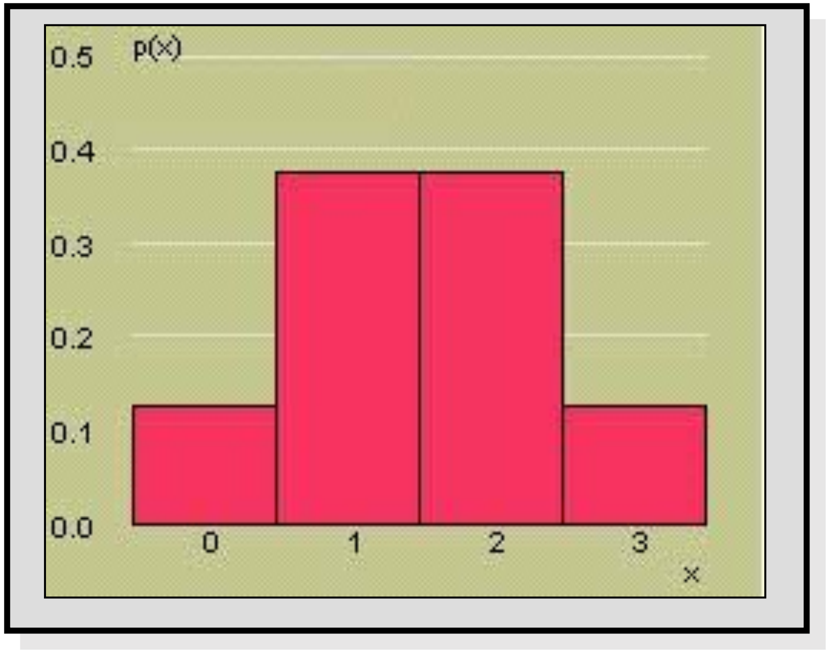
- Discrete random variables take on only a finite or countably infinite number of values.
- Two discrete probability distributions serve as models for a large number of practical applications:

✓The **binomial** random variable

✓The **Poisson** random variable

The Binomial Random Variable

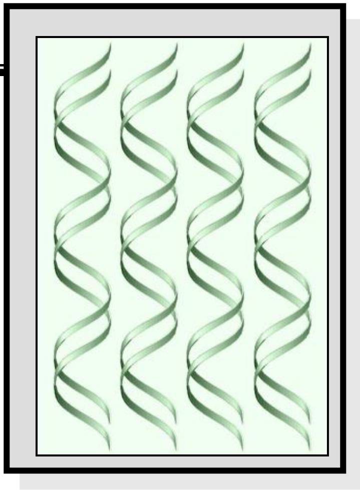
- The **coin-tossing experiment** is a simple example of a **binomial random variable.** Toss a fair coin n = 3 times and record x = number of heads.



x	$p(x)$
0	1/8
1	3/8
2	3/8
3	1/8

The Binomial Random Variable

- Many situations in real life resemble the coin toss, but the coin is not necessarily fair, so that $P(H) \neq 1/2$.



- **Example:** A geneticist samples 10 people and counts the number who have a gene linked to Alzheimer's disease.

- Coin: Person
- Head: Has gene
- Tail: Doesn't have gene
- Number of tosses: $n = 10$
- P(H): P(has gene) = proportion in the population who have the gene.

The Binomial Experiment

1. The experiment consists of ***n* identical trials.**
2. Each trial results in **one of two outcomes**, success (S) or failure (F).
3. The probability of success on a single trial is p and **remains constant** from trial to trial. The probability of failure is $q = 1 - p$.
4. The trials are **independent**.
5. We are interested in ***x*, the number of successes in *n* trials.**

The Binomial Probability Distribution

- For a binomial experiment with *n* trials and probability *p* of success on a given trial, the probability of *k* successes in *n* trials is

$$P(x=k)=C_k^n p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k} \text{ for } k=0,1,2,\ldots n.$$

$$\text{Recall} \quad C_k^n = \frac{n!}{k!(n-k)!}$$

$$\text{with } n! = n(n-1)(n-2)\ldots(2)1 \text{ and } 0! \equiv 1.$$

The Mean and Standard Deviation

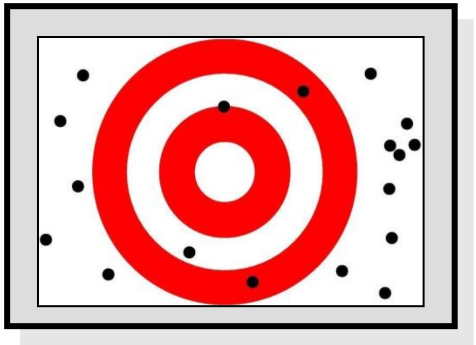
- For a binomial experiment with *n* trials and probability *p* of success on a given trial, the measures of center and spread are:

$$\text{Mean}: \mu = np$$

$$\text{Variance}: \sigma^2 = npq$$

$$\text{Standard deviation}: \sigma = \sqrt{npq}$$

Example



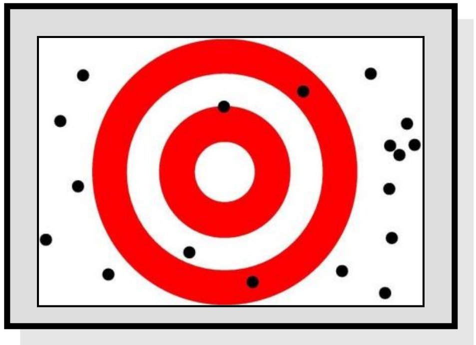
A marksman hits a target 80% of the time. He fires five shots at the target. What is the probability that exactly 3 shots hit the target?

n = 5 *success* = hit *p* = .8 *x* = # of hits

$$P(x=3)=C_3^n p^3 q^{n-3}=\frac{5!}{3!2!}(.8)^3(.2)^{5-3}$$

$$=10(.8)^3(.2)^2=.2048$$

Example



What is the probability that more than 3 shots hit the target?

$$P(x>3)=C_4^5 p^4 q^{5-4}+C_5^5 p^5 q^{5-5}$$

$$=\frac{5!}{4!1!}(.8)^4(.2)^1+\frac{5!}{5!0!}(.8)^5(.2)^0$$

$$=5(.8)^4(.2)+(.8)^5=.7373$$

The Poisson Random Variable

- The Poisson random variable *x* is a model for data that represent the number of occurrences of a specified event in a given unit of time or space.

- Examples:
  - The number of calls received by a switchboard during a given period of time.
  - The number of machine breakdowns in a day
  - The number of traffic accidents at a given intersection during a given time period.

The Poisson Probability Distribution

- ***x*** is the number of events that occur in a period of time or space during which an average of **μ** such events can be expected to occur. The probability of *k* occurrences of this event is

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

For values of *k* = 0, 1, 2, … The mean and standard deviation of the Poisson random variable are

Mean: μ

Standard deviation: $\sigma=\sqrt{\mu}$

Example

The average number of traffic accidents on a certain section of highway is two per week. Assume that the number of accidents follow a Poisson distribution with $\mu = 2$. Find the probability of

(a) exactly one accident during a one-week period.

(b) at most three accidents on this section of highway during a 2-week period

Cumulative Probability Tables

You can use the **cumulative probability tables** to find probabilities for selected Poisson distributions.

✓Find the column for the correct value of μ.

✓The row marked "k" gives the cumulative probability, $P(x \le k) = P(x = 0) + \ldots + P(x = k)$

Example

What is the probability that there is exactly 1 accident?

				μ		
k	2.0	2.5	3.0	3.5	4.0	4.5
0	.135	.082	.055	.033	.018	.011
1	.406	.287	.199	.136	.092	.061
2	.677	.544	.423	.321	.238	.174
3	.857	.758	.647	.537	.433	.342
4	.947	.891	.815	.725	.629	.532
5	.983	.958	.916	.858	.785	.703
6	.995	.986	.966	.935	.889	.831
7	.999	.996	.988	.973	.949	.913
8	1.000	.999	.996	.990	.979	.960
9		1.000	.999	.997	.992	.983
10			1.000	.999	.997	.993
11				1.000	.999	.998
12					1.000	.999
13						1.000

P(x = 1) = P($x \le 1$) – P($x \le 0$)
= .406 - .135
= .271

Check from formula: P(x = 1) = .2707

Example

k	$\mu = 2$
0	.135
1	.406
2	.677
3	.857
4	.947
5	.983
6	.995
7	.999
8	1.000

What is the probability that 8 or more accidents happen?

P($x \geq$ 8) = 1 - P(x < 8)
= 1 – P($x \leq$ 7)
= 1 - .999 = .001

Example

A life insurance salesman sells on the average 3 life insurance policies per week. Use Poisson's law to calculate the probability that in a given week he will sell

a.Some policies

b. 2 or more policies but less than 5 policies.

c. Assuming that there are 5 working days per week, what is the probability that in a given day he will sell one policy?

Here, $\mu = 3$

(a) "Some policies" means "1 or more policies". We can work this out by finding 1 minus the "zero policies" probability:

$$P(X > 0) = 1 - P(x_0)$$

Now $P(X) = \dfrac{e^{-\mu}\mu^x}{x!}$ so $P(x_0) = \dfrac{e^{-3}3^0}{0!} = 4.9787 \times 10^{-2}$

Therefore the probability of 1 or more policies is given by:

$$\begin{aligned} \text{Probability} &= P(X \geq 0) \\ &= 1 - P(x_0) \\ &= 1 - 4.9787 \times 10^{-2} \\ &= 0.95021 \end{aligned}$$

Example

A life insurance salesman sells on the average 3 life insurance policies per week. Use Poisson's law to calculate the probability that in a given week he will sell

a. Some policies

b. 2 or more policies but less than 5 policies.

c. Assuming that there are 5 working days per week, what is the probability that in a given day he will sell one policy?

(b) The probability of selling 2 or more, but less than 5 policies is:

$$
\begin{aligned}
P(2 \leq X < 5) \\
&= P(x_2) + P(x_3) + P(x_4) \\
&= \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} \\
&= 0.61611
\end{aligned}
$$

(c) Average number of policies sold per day: $\frac{3}{5} = 0.6$

So on a given day, $P(X) = \frac{e^{-0.6}(0.6)^1}{1!} = 0.32929$

Example

A company makes electric motors. The probability an electric motor is defective is 0.01. What is the probability that a sample of 300 electric motors will contain exactly 5 defective motors?

Example

A company makes electric motors. The probability an electric motor is defective is 0.01. What is the probability that a sample of 300 electric motors will contain exactly 5 defective motors?

The average number of defectives in 300 motors is $\mu = 0.01 \times 300 = 3$

The probability of getting 5 defectives is:

$$P(X) = \frac{e^{-3}3^5}{5!} = 0.10082$$

Example

If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.

Example

If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.

The average number of failures per week is: $\mu = \frac{3}{20} = 0.15$

"Not more than one failure" means we need to include the probabilities for "0 failures" plus "1 failure".

$$P(x_0) + P(x_1) = \frac{e^{-0.15} 0.15^0}{0!} + \frac{e^{-0.15} 0.15^1}{1!} = 0.98981$$

Example

Vehicles pass through a junction on a busy road at an average rate of 300 per hour.

a.Find the probability that none passes in a given minute.

b.What is the expected number passing in two minutes?

c.Find the probability that this expected number actually pass through in a given two-minute period.

Example

Vehicles pass through a junction on a busy road at an average rate of 300 per hour.
a.Find the probability that none passes in a given minute.
b.What is the expected number passing in two minutes?
c.Find the probability that this expected number actually pass through in a given two-minute period.

The average number of cars per minute is: $\mu = \frac{300}{60} = 5$

(a) $P(x_0) = \frac{e^{-5}5^0}{0!} = 6.7379 \times 10^{-3}$

(b) Expected number each 2 minutes = $E(X) = 5 \times 2 = 10$

(c) Now, with $\mu = 10$, we have: $P(x_{10}) = \frac{e^{-10}10^{10}}{10!} = 0.12511$

Example

Twenty sheets of aluminum alloy were examined for surface flaws. The frequency of the number of sheets with a given number of flaws per sheet was as follows:

Number of flaws	Frequency
0	4
1	3
2	5
3	2
4	4
5	1
6	1

What is the probability of finding a sheet chosen at random which contains 3 or more surface flaws?

Example

The total number of flaws is given by:

$$(0 \times 4) + (1 \times 3) + (2 \times 5) + (3 \times 2) + (4 \times 4) + (5 \times 1) + (6 \times 1) = 46$$

So the average number of flaws for the 20 sheets is given by:

$$\mu = \frac{46}{20} = 2.3$$

The required probability is:

$$\begin{aligned}\text{Probability} &= P(X \geq 3) \\ &= 1 - (P(x_0) + P(x_1) + P(x_2)) \\ &= 1 - \left(\frac{e^{-2.3}2.3^0}{0!} + \frac{e^{-2.3}2.3^1}{1!} + \frac{e^{-2.3}2.3^2}{2!}\right) \\ &= 0.40396\end{aligned}$$

Key Concepts

I. The Binomial Random Variable

1. Five characteristics: *n* identical independent trials, each resulting in either success *S* or failure *F*; probability of success is *p* and remains constant from trial to trial; and *x* is the number of successes in *n* trials.

2. Calculating binomial probabilities

   a. Formula: $\boxed{P(x=k)=C_k^n p^k q^{n-k}}$

   b. Cumulative binomial tables

   c. Individual and cumulative probabilities using Minitab

3. Mean of the binomial random variable: $\mu = np$

4. Variance and standard deviation: $\sigma^2 = npq$ and $\sigma = \sqrt{npq}$

Key Concepts

II. The Poisson Random Variable

1. The number of events that occur in a period of time or space, during which an average of μ such events are expected to occur

2. Calculating Poisson probabilities

   a. Formula: $\boxed{P(x=k)=\dfrac{\mu^k e^{-\mu}}{k!}}$

   b. Cumulative Poisson tables

   c. Individual and cumulative probabilities using Minitab

3. Mean of the Poisson random variable: $E(x) = \mu$

4. Variance and standard deviation: $\sigma^2 = \mu$ and $\boxed{\sigma = \sqrt{\mu}}$

5. Binomial probabilities can be approximated with Poisson probabilities when $np < 7$, using $\mu = np$.

Thank You