Introduction to Probability and Statistics Twelve Edition

Chapter 5

Several Useful Discrete Distributions

Introduction

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

✓The **binomial** distribution

✓The **hypergeometric** distribution

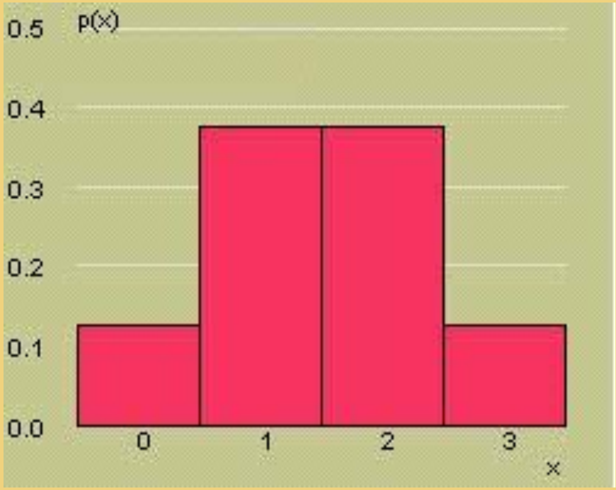
✓The **Poisson** distribution

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.**

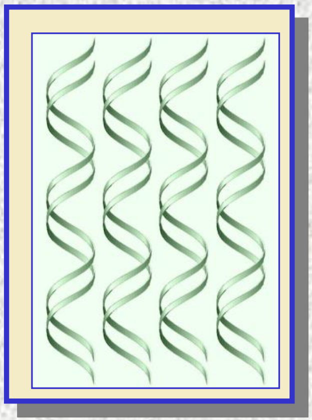
The Binomial Random Variable

- 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

Binomial or Not?



- Very few real life applications satisfy these requirements exactly.

- Select two people from the U.S. population, and suppose that 15% of the population has the Alzheimer's gene.
  - For the first person, p = P(gene) = .15
  - For the second person, $p \approx$ P(gene) = .15, even though one person has been removed from the population.

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

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

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

The Mean and Standard Deviation

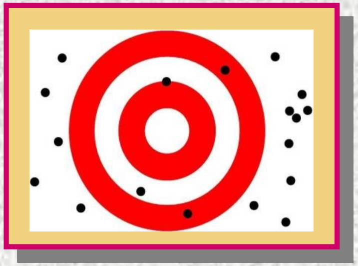
- For a binomial experiment with *n* trials and probability *p* of success on a given trial, the measures of center and spread are:

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

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

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

Example 1.1



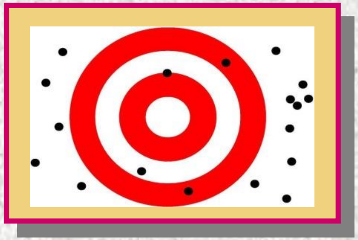
A marksman hits a target 80% of the time. He fires five shots at the target. What is the probability that exactly 3 shots hit the target?

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

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

$$=10(.8)^3(.2)^2=0.2048$$

Example 1.2



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

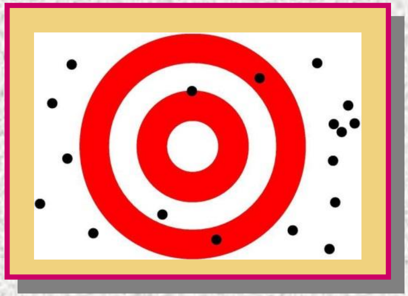
$$P(x>3)=C_4^5 p^4 q^{5-4}+C_5^5 p^5 q^{5-5}$$

$$=\frac{5!}{4!1!}(0.8)^4(0.2)^1+\frac{5!}{5!0!}(0.8)^5(0.2)^0$$

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

$$=0.7373$$

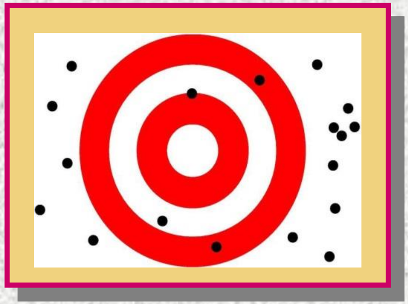
Cumulative Probability Tables



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

- ✓**Find the table for the correct value of *n*.**
- ✓**Find the column for the correct value of *p*.**
- ✓The row marked "*k*" gives the cumulative probability, $P(x \le k) = P(x = 0) + \cdots + P(x = k)$

Example 1.3



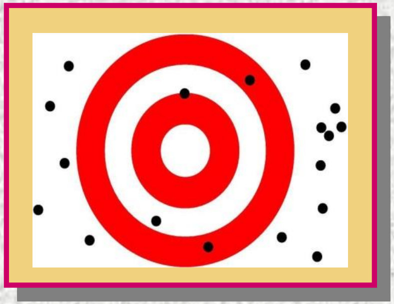
x	$p = .80$
0	0.000
1	0.007
2	0.058
3	0.263
4	0.672
5	1.000

What is the probability that exactly 3 shots hit the target?

$$P(x = 3) = P(x \leq 3) - P(x \leq 2)$$
$$= 0.263 - 0.058$$
$$= 0.205$$

Check from formula:
$P(x = 3) = 0.2048$

Example 1.4



x	$p = .80$
0	0.000
1	0.007
2	0.058
3	0.263
4	0.672
5	1.000

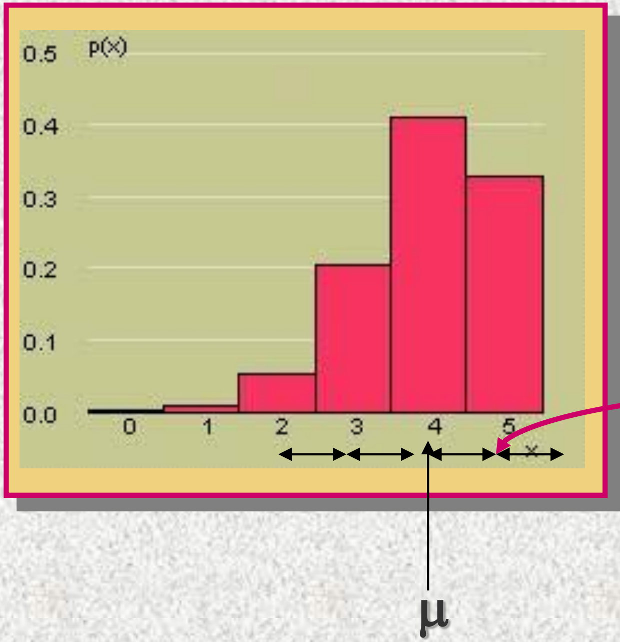
What is the probability that more than 3 shots hit the target?

$$\begin{aligned} P(x > 3) &= 1 - P(x \le 3) \\ &= 1 - 0.263 \\ &= 0.737 \end{aligned}$$

Check from formula:
$P(x > 3) = 0.7373$

Example 1.5

- Here is the probability distribution for x = **number of hits.** What are the mean and standard deviation for x?



$$\text{Mean}: \mu = np = 5(.8) = 4$$

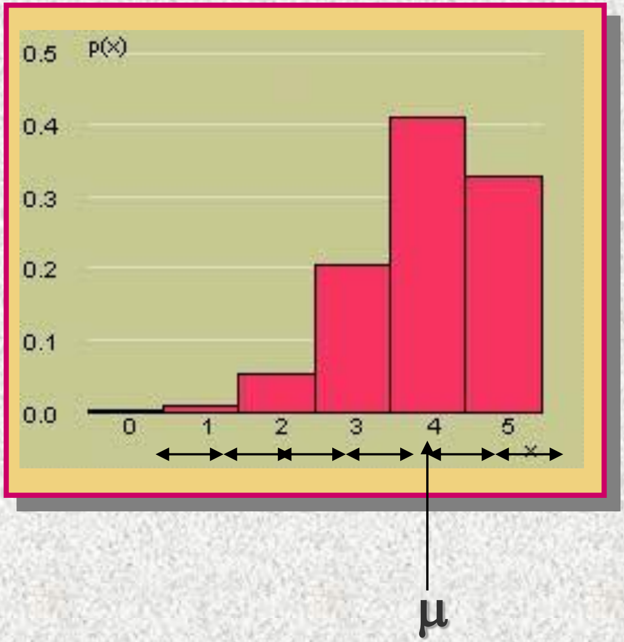
$$\text{Standard deviation}: \sigma = \sqrt{npq}$$

$$= \sqrt{5(.8)(.2)} = .89$$

Example 1.6

- Would it be unusual to find that none of the shots hit the target?

$$\mu = 4; \sigma = .89$$



- The value $x = 0$ lies

$$z = \frac{x-\mu}{\sigma} = \frac{0-4}{0.89} \approx -4.49$$

more than 4 standard deviations below the mean. Very unusual.

Class Activity

1. Consider a binomial random variable with $n = 8$ and $p = 0.7$. Let x be the number of successes in the sample. Use the binomial table to find
   a) $P(x \le 3)$
   b) $P(x = 3)$
   c) $P(x \ge 3)$
   d) $P(x < 3)$
   e) $P(3 \le x \le 5)$
   f) $P(3 < x < 5)$

Class Activity

2. Find the mean and standard deviation for a binomial distribution with these values:
   a) n=1000, p=0.3
   b) n=400, p=0.01

Class Activity

3. On a 12-question multiple-choice test, there are five possible answers, of which one is correct. Suppose that a student guesses on each question. Find
   a) the probability that he gets 3 correct answers,
   b) the probability that he gets at least 3 correct answers.
   c) the mean of correct answers he gets,.
   d) the probability distribution of *x*, where *x* is the number of correct answers.

b) $P(X \geq 3) = P(X=3) + P(X=4) + \dots + P(X=12)$

d) $P(X=k) = C^{12}_{k} (0.2)^{k} (0.8)^{12-k}$, for $k=1,2,\dots,12$

c) $\mu = np = 12(0.2) = 2.4$

Class Activity

4. In a country 30% of the people have a certain blood type. If a group of 15 people are randomly selected, Find $p = 0.3, n = 15$

   a) the probability that at most 4 people will have that blood type. $P(X \le 4) = P(X=0) + P(X=1) + \cdots + P(X=4)$ $= \sum_{x=0}^{4} C_x^{15} (0.3)^x (0.6)^{15-x}$

   b) the probability that none will have that blood type $P(X=0) = C_0^{15} (0.3)^0 (0.7)^{15} = (0.7)^{15}$

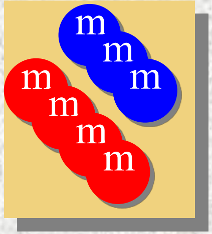
   c) the probability that between 1 and 5 people (inclusively) will have that blood type,

   d) the mean number of people who have this certain blood type. $\mu = np = 15(0.3) = 4.5$

c) $P(1 \le X \le 5) = P(X=1) + P(X=2) + \ldots + P(X=5)$

$= C_1^{15} (0.3)^1 (0.7)^{14} + \ldots + C_5^{15} (0.3)^5 (0.7)^{10}$

The Hypergeometric Probability Distribution



- The "M&M®" problems" from Chapter 4 are modeled by the **hypergeometric distribution**.
- A bowl contains M red candies and N-M blue candies. We randomly select *n* candies all at once from the bowl and record *x* the number of red candies selected. Define a "red M&M®" to be a "success".

The probability of exactly *k* successes in *n* trials is

$$P(x=k)=\frac{C_k^M C_{n-k}^{N-M}}{C_n^N}$$

for $k \le n$, $k \le M$, and $n\text{-}k \le N\text{-}M$.

The Mean and Variance

The mean and variance of the hypergeometric random variable *x* resemble the mean and variance of the binomial random variable:

$$\text{Mean}: \mu = n\left(\frac{M}{N}\right)$$

$$\text{Variance}: \sigma^2 = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Example 2.1

A package of 8 AA batteries contains 2 batteries that are defective. A student randomly selects four batteries and replaces the batteries in his calculator. What is the probability that all four batteries work?

Success = working battery

N = 8

M = 6

***n* = 4**

$$P(x=4) = \frac{C_4^6 C_0^2}{C_4^8}$$

$$= \frac{6(5)/2(1)}{8(7)(6)(5)/4(3)(2)(1)}$$

$$= \frac{15}{70}$$

Example 2.2

What are the mean and variance for the number of batteries that work?

$$\mu = n\left(\frac{M}{N}\right) = 4\left(\frac{6}{8}\right) = 3$$

$$\sigma^2 = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$= 4\left(\frac{6}{8}\right)\left(\frac{2}{8}\right)\left(\frac{4}{7}\right) = 0.4286$$

Class Activity

5. **Defective Computer Chips.** A piece of electronic equipment contains six computer chips, two of which are defective. Three computer chips are randomly chosen for inspection.

   a) What is the probability that one computer chip is defective? $\frac{C_1^2 C_2^4}{C_3^6} = 0.6$

   b) Find the probability distribution of *x*, the number of defective chips found. $\frac{C_k^2 C_{3-k}^4}{C_3^6}$

   c) What is the probability that one computer chip is non-defective? $\frac{C_2^2 C_1^4}{C_3^6}$

   d) What is the mean number of defective chips?

$\mu = 3\left(\frac{2}{6}\right) = 1$ chip

choose 3

defective

Class Activity

6) A random committee of size 4 is selected from 7 doctors and 2 nurses.

a) Write a formula for the probability distribution of the random variable *x*, representing the number of doctors on the committee.

b) On the average how many doctors are on the committee?

a) $\frac{C^7_k C^2_{4-k}}{C^9_4}$ for $c = 2, 3, 4$

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 $\boldsymbol{\mu}$ such events can be expected to occur. The probability of *k* occurrences of this event is

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

The mean and standard deviation of the Poisson random variable are

Mean: μ

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

Example 3.1

The average number of traffic accidents on a certain section of highway is two per week. Find the probability of exactly one accident during a one-week period.

$$P(x=1) = \frac{\mu^k e^{-\mu}}{k!} = \frac{2^1 e^{-2}}{1!} = 2e^{-2} = 0.2707$$

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 3.2

x	$\mu = 2$
0	0.135
1	0.406
2	0.677
3	0.857
4	0.947
5	0.983
6	0.995
7	0.999
…	…

What is the probability that there is exactly 1 accident?

$$\begin{aligned} \mathbf{P(x = 1)} &= P(x \le 1) - P(x \le 0) \\ &= 0.406 - 0.135 \\ &= 0.271 \end{aligned}$$

Check from formula:
$P(x = 1) = 0.2707$

Example 3.3

x	$\mu = 2$
0	0.135
1	0.406
2	0.677
3	0.857
4	0.947
5	0.983
6	0.995
7	0.999
…	…

What is the probability that 8 or more accidents happen?

$$\begin{aligned} P(x \geq 8) &= 1 - P(x < 8) \\ &= 1 - P(x \leq 7) \\ &= 1 - \ 0.999 = 0.001 \end{aligned}$$

The Poisson Probability Distribution

Let ***μ*** be the average number of times an event occurred in a period of time (or space)

and ***y*** be the number of occurrences in ***t*** periods of time, then

$$P(y = k) = \frac{(\mu t)^k e^{-\mu t}}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

The mean of y is ***μt.***

The variance of y is ***μt.***

Example 3.4

The average number of traffic accidents on a certain section of highway is two per week. Find the probability that there are 3 accidents in any month. Assume that there are 4 weeks per month.

Let y be the number of traffic accidents per 4 weeks.

$$P(y=3)=\frac{(\mu t)^3 e^{-(\mu t)}}{3!} = \frac{8^3 e^{-8}}{3!} = 0.0286$$

Example 3.4 (cont.)

What is the probability that there is 1 accident in a day?

$\mu = 2$, $\mu t = \frac{2}{7}$

$$P(w=1) = \frac{(\frac{2}{7})^1 e^{-2/7}}{1!} = 0.215$$

What is the probability of at least 1 accident per two weeks?

$\mu t = 4$ accidents per 2 weeks

$P(u \geq 1) = P(u=1) + P(u=2) + \dots$

or $P(u \geq 1) = 1 - P(u=0)$

$$= 1 - \frac{4^0 e^{-4}}{0!} = 0.9817$$

Table

$P(u \geq 1) = 1 - P(u=0) = 1 - 0.018$

$= 0.982$

Class Activity

7. Consider a Poisson random variable with $\mu=2.5$. Use the Poisson table to calculate the following probabilities:

   a) $P(x < 2)$ → P(0) + P(1) = 0.287

   b) $P(x = 0)$ 0.082

   c) $P(x = 3)$ 0.758 − 0.544 = 0.214

   d) $P(2 \leq x \leq 6)$ P(2) + P(3) + P(4) + P(5) + P(6)

   = P(6) − P(1) − P(0) = 0.986 − 0.287

   = 0.699

Class Activity

8. A secretary makes 0.5 typing error per page on the average. Suppose that the number of typing errors has a Poisson probability distribution.

   a) Find the probability that she makes 3 errors per one page. $\frac{0.5^3 e^{-0.5}}{3!}$

   b) Find the probability that she makes more than 2 errors per one page. $1-\frac{0.5^2 e^{-0.5}}{2!}-\frac{0.5\, e^{-0.5}}{1!}-\frac{0.5^0 e^{-0.5}}{0!}$

   c) Find the probability that she makes 5 errors per 4 pages. $\frac{2^5 e^{-2}}{5!}$